



**HAL**  
open science

## Conserved Viral Transcription Plays a Key Role in Virus-Like Particle Production of the Parasitoid Wasp *Venturia canescens*

Alexandra Cerqueira de Araujo, Matthieu Leobold, Annie Bézier, Karine Musset, Rustem Uzbekov, Anne-Nathalie Volkoff, Jean-Michel Drezen, Elisabeth Huguet, Thibaut Josse

### ► To cite this version:

Alexandra Cerqueira de Araujo, Matthieu Leobold, Annie Bézier, Karine Musset, Rustem Uzbekov, et al.. Conserved Viral Transcription Plays a Key Role in Virus-Like Particle Production of the Parasitoid Wasp *Venturia canescens*. *Journal of Virology*, 2022, 96 (13), 10.1128/jvi.00524-22 . hal-03838164

**HAL Id: hal-03838164**

**<https://hal.inrae.fr/hal-03838164>**


Submitted on 21 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Conserved Viral Transcription Plays a Key Role in Virus-Like Particle Production of the Parasitoid Wasp *Venturia canescens*

Alexandra Cerqueira de Araujo,<sup>a</sup> Matthieu Leobold,<sup>a</sup> Annie Bézier,<sup>a</sup> Karine Musset,<sup>a</sup> Rustem Uzbekov,<sup>b,d</sup> Anne-Nathalie Volkoff,<sup>c</sup> Jean-Michel Drezen,<sup>a</sup>  Elisabeth Huguet,<sup>a</sup> Thibaut Josse<sup>a</sup>

<sup>a</sup>Institut de Recherche sur la Biologie de l'Insecte (IRBI), UMR 7261, CNRS - Université de Tours, Tours, France

<sup>b</sup>Université de Tours, Département des Microscopies, Tours, France

<sup>c</sup>Diversité, Génomes & Interactions Microorganismes - Insectes (DGIMI), UMR 1333, Université de Montpellier - INRAE, Montpellier, France

<sup>d</sup>Faculty of Bioengineering and Bioinformatics, Moscow State University, Moscow, Russia

Elisabeth Huguet and Thibaut Josse are co-senior authors.

**ABSTRACT** Nudiviruses are large double-stranded DNA viruses related to baculoviruses known to be endogenized in the genomes of certain parasitic wasp species. These wasp-virus associations allow the production of viral particles or virus-like particles that ensure wasp parasitism success within lepidopteran hosts. *Venturia canescens* is an ichneumonid wasp belonging to the Campopleginae subfamily that has endogenized nudivirus genes belonging to the *Alphanudivirus* genus to produce “virus-like particles” (*Venturia canescens* virus-like particles [VcVLPs]), which package proteic virulence factors. The main aim of this study was to determine whether alphanudivirus gene functions have been conserved following endogenization. The expression dynamics of alphanudivirus genes was monitored by a high throughput transcriptional approach, and the functional role of *lef-4* and *lef-8* genes predicted to encode viral RNA polymerase components was investigated by RNA interference. As described for baculovirus infections and for endogenized nudivirus genes in braconid wasp species producing bracoviruses, a transcriptional cascade involving early and late expressed alphanudivirus genes could be observed. The expression of *lef-4* and *lef-8* was also shown to be required for the expression of alphanudivirus late genes allowing correct particle formation. Together with previous literature, the results show that endogenization of nudiviruses in parasitoid wasps has repeatedly led to the conservation of the viral RNA polymerase function, allowing the production of viruses or virus-like particles that differ in composition but enable wasp parasitic success.

**IMPORTANCE** This study shows that endogenization of a nudivirus genome in a Campopleginae parasitoid wasp has led to the conservation, as for endogenized nudiviruses in braconid parasitoid wasps, of the viral RNA polymerase function, required for the transcription of genes encoding viral particles involved in wasp parasitism success. We also showed for the first time that RNA interference (RNAi) can be successfully used to downregulate gene expression in this species, a model in behavioral ecology. This opens the opportunity to investigate the function of genes involved in other traits important for parasitism success, such as reproductive strategies and host choice. Fundamental data acquired on gene function in *Venturia canescens* are likely to be transferable to other parasitoid wasp species used in biological control programs. This study also renders possible the investigation of other nudivirus gene functions, for which little data are available.

**KEYWORDS** RNA polymerase, *Venturia canescens*, bracovirus, nudivirus, parasitoid wasp, viral domestication, virus-like particles

The endogenization and fixation of viral sequences into eukaryotic DNA is a widespread process (1–3). Most of these embedded viral sequences, called endogenous

**Editor** Colin R. Parrish, Cornell University

**Copyright** © 2022 American Society for Microbiology. All Rights Reserved.

Address correspondence to Elisabeth Huguet, elisabeth.huguet@univ-tours.fr, or Thibaut Josse, thibaut.josse@univ-tours.fr.

The authors declare no conflict of interest.

**Received** 1 April 2022

**Accepted** 4 May 2022

**Published** 9 June 2022

viral elements (EVEs) are now fragmented and nonfunctional, having evolved under neutral selection (4, 5). However, in some cases, endogenized viral genes can confer selective advantages to the host organism by providing new functions that are adaptive (6–10). For example, at least 5 independent domestication events involving almost complete viral genomes have been described in five lineages of Ichneumonoidea parasitoid wasps (11, 12). These integrated viruses have retained dozens of genes allowing the wasps to produce virus-derived particles to ensure the parasitism success of their offspring injected as eggs into an insect host, commonly belonging to Lepidoptera (13–21).

Among the domestication events, ichnoviruses (IV) identified in the genomes of some Ichneumonidae Campopleginae and Banchinae wasps are thought to originate from closely related virus ancestors corresponding to an uncharacterized or extinct viral taxon (19, 22, 23). Bracoviruses (BV) are integrated in the genomes of Braconidae parasitoid wasps belonging to a monophyletic taxonomic group called the microgastroid complex, estimated to comprise over 46,000 species (24). This association originated about 100 million years ago (MYA) following the integration of a nudivirus from the *Betanudivirus* genus into the genome of the microgastroid wasp ancestor (14, 25–27). Infecting insects and crustaceans, exogenous nudiviruses have circular double-stranded DNA (dsDNA) genomes and are a sister group of the well-known baculoviruses (26, 28–30). Integrated nudivirus sequences have been described in several arthropods (31–37); however, the functional consequences of such viral genomic elements are only evident in parasitoid wasps (11, 13, 20, 27, 38). Despite the fact that IV and BV are of completely different viral origin, these viruses have historically been grouped in the same *Polydnaviridae* (PDV) family, and they illustrate a remarkable case of convergent evolution. Both IV and BV genomes lead to the production of enveloped viral particles that contain multiple dsDNA molecules that harbor virulence genes (27, 39–41). The expression of virulence genes in the lepidopteran host leads to host immune suppression, enabling wasp parasitism success (42–44).

Interestingly, more recent viral integration events in Ichneumonoidea parasitoid wasps have been described, which involve integrations of nudiviruses from the *Alphanudivirus* genus (20, 45, 46). The braconid wasp *Fopius arisanus*, which is distantly related to species from the microgastroid complex, was recently shown to have integrated an alphanudivirus, allowing the production of virus-like particles (VLPs), which consist of enveloped capsids that lack DNA (46). *Venturia canescens*, an ichneumonid campoplegine wasp, was also known to produce VLPs (47) supposedly from the endogenous IV present in related species, which could have lost the ability to produce nucleocapsids (48). However, surprisingly, it was instead shown to have relatively recently integrated an alphanudivirus (*Venturia canescens* endogenous nudivirus [VcENV]), which allows the production of immune suppressive *V. canescens* virus-like particles (VcVLPs) consisting of an envelope but lacking a capsid (20). Similar to bracovirus and ichnovirus particles, VcVLPs are produced in the calyx of ovaries and released in the ovary lumen. VcVLPs, however, do not contain nucleic acids but instead package virulence proteins of wasp origin that protect eggs from the lepidopteran host immune system after oviposition (20, 48–52).

The *Nudiviridae* consists of a family of large circular dsDNA arthropod viruses, which diverged over 220 MYA. They share 32 core genes between them (26, 29, 32, 34, 53) and 21 core genes with baculoviruses (11, 21, 26, 29, 34, 54). They include genes involved in essential viral functions, such as DNA replication, viral gene transcription, capsid assembly, and host cell entry. None of the characterized parasitoid wasp-endogenized nudiviruses have retained a functional viral DNA polymerase gene, this gene being either pseudogenized in VcENV or totally missing in BV genomes (13, 20, 45, 55, 56), suggesting that BV DNA replication must rely on polymerases of wasp origin (13, 17). However, parasitoid wasp-endogenized nudiviruses have retained most (BV) or many (VcENV) of the other baculovirus or nudivirus core genes (11, 13, 14, 20, 45). Indeed, the endogenized virus in *V. canescens*, for which a total of 51 nudivirus genes could be identified, has conserved 21 nudivirus core genes, 16 of which are core genes shared with baculoviruses (11, 20, 45).

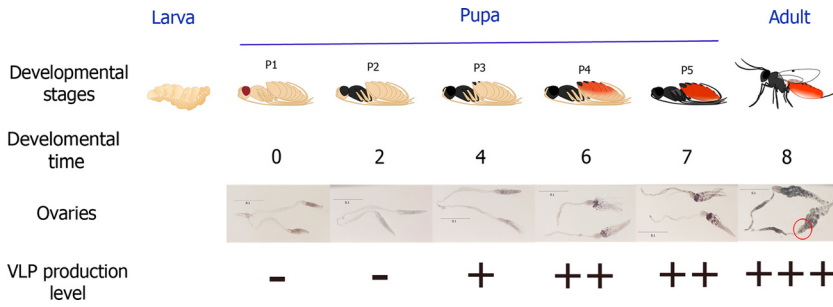
*V. canescens* is, to date, the only wasp in its lineage to have endogenized a nudivirus; this domestication event is thought to be more recent (probably a few MYA) than the betanuvirus integration within the microgastroid wasp ancestor (100 MYA) (20, 25). In accordance, nudiviral genes have been shown to be less dispersed in the *V. canescens* genome than in wasps associated with BVs (20, 38, 56). Interestingly, the integrated alphanudivirus (VcENV) did not take the same evolutionary trajectory as its betanuvirus counterpart in braconid wasps: the genes encoding the structural components of the nucleocapsid have been specifically lost by pseudogenization (20, 45), resulting in the production of VcVLPs that consist of liposome-like particles delivering virulence proteins within the lepidopteran host (20), whereas BVs deliver virulence genes.

In baculoviruses, genes are classified as “early,” “late,” and “very late” genes according to the timing of their expression during cell infection. The cellular RNA polymerase transcribes the early genes that comprise *p47* and *lef-5* genes and the viral RNA polymerase subunit encoding genes (*lef-4*, *lef-8*, and *lef-9*). Then, the viral RNA polymerase ensures the transcription of the late or the very late genes that are involved in capsid assembly (such as *vp91*, also known as *pif-8*, and *p33*) and production of envelope components of the virion, which can play a role in host cell entry (*pif* genes for *per os* infectivity factor) (57–59). Remarkably, several lines of evidence show that nudivirus genes endogenized by braconid wasps have also retained core viral functions after domestication. Indeed, nudivirus gene expression analyses during wasp pupal development revealed early expression of nudivirus genes whose products are required for transcription of other viral late genes involved in particle formation (46, 56, 60). Furthermore, knockdown of the bracovirus *lef-4* and *lef-9* in *Microplitis demolitor* by RNA interference resulted in lowered transcription of some BV genes of nudiviral origin, as expected if BV *lef* genes have retained RNA polymerase function (17). RNA interference (RNAi) experiments also supported a conserved role of VP39, VLF-1, P74 (alias PIF-0), and PIF-1 as structural components of *Microplitis demolitor* BV (MdBV) virions (17).

In this study, we investigated for the first time in the Campopleginae wasp *V. canescens* whether the endogenized alphanudivirus also retained viral core functions. Indeed, proteins potentially involved in viral gene transcription and several PIF proteins in VcENV showed a higher-than-average sequence conservation compared with that of other exogenous alphanudiviruses, suggesting that functional constraints are operating in the context of domestication to maintain viral functions (45). More specifically, to determine whether core viral functions have been retained in *V. canescens*, we first investigated the expression dynamics of alphanudivirus genes in ovaries during wasp pupal development with a high throughput transcriptional approach and compared promoter regions with known baculovirus promoters. Second, we investigated whether the nudivirus genes required to produce the RNA polymerase are essential for the expression of late genes and VLP morphogenesis by developing, for the first time in *V. canescens*, an RNAi approach. The results allowed us to demonstrate that *lef-4* and *lef-8* are involved in nudivirus gene transcription of late genes, showing that the essential RNA polymerase viral function has been retained. Furthermore, electron microscopy analyses of knockdown *lef-4* and *lef-8* wasps revealed strong impacts on virion production. These results, together with previous literature, show experimentally that the nudivirus RNA polymerase is essential in both BV and VLP morphogenesis (17, 20, 45).

## RESULTS

**Dynamics of endogenized alphanudivirus (VcENV) gene expression throughout wasp development.** To determine whether alphanudivirus genes in *V. canescens* are regulated in a temporally coordinated manner as has been observed for nudivirus genes in Braconidae wasps (56, 60), we first investigated the development of ovaries and calyx during wasp pupal development in order to relate the developmental stages of these organs with VcVLP production. A previous study divided the *V. canescens* pupal period into five pupal stages (called P1 to P5) based on pigmentation of pupal cuticle (Fig. 1) (61). Here, we show that the calyx gains in length, measuring

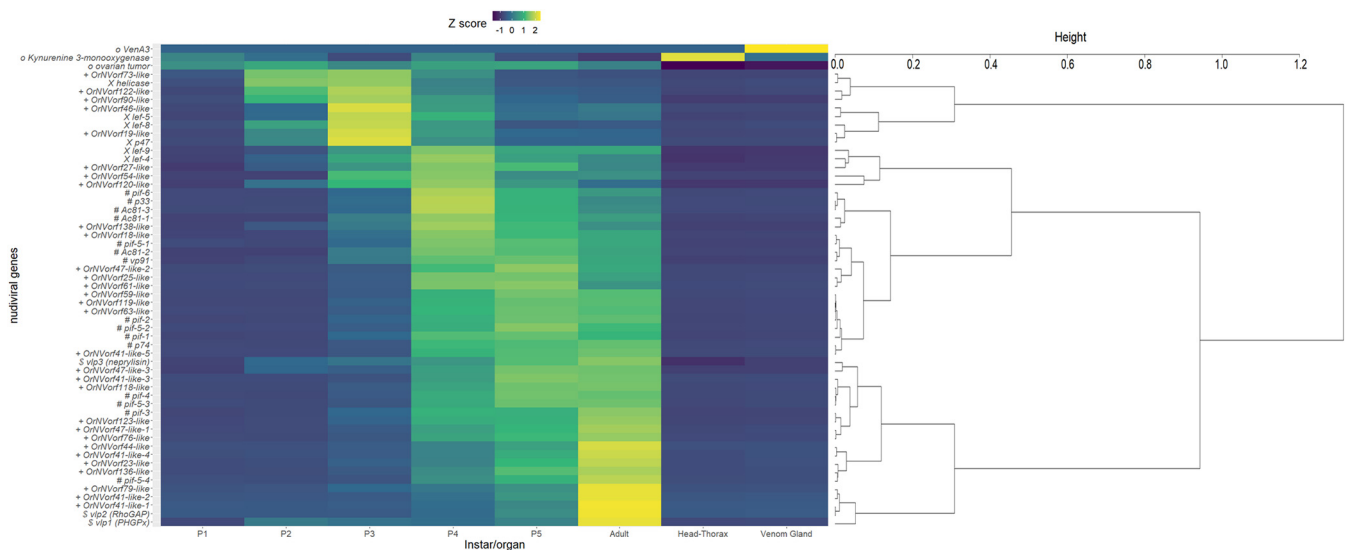


**FIG 1** Schematic representation of each *V. canescens* pupal developmental stage associated with pictures of ovaries and an indication of VcVLP production intensity. Pictures of wasp ovaries at different pupal developmental stages were obtained using a stereo microscope; the scale bar represents 0.1 cm. Calyx position is specified with a circle in one of the adult ovaries. VLP production was reported from TEM observations from Pichon and collaborators (20), in which P4 and P5 stages described in Reineke et al. (61) were merged (-, no production; +, very low production; ++, intermediate production; +++, strong production). Development times postpupation (in days) are indicated here for wasps reared in our conditions at 25°C.

approximately 100 μm in P1 and 200 μm in adult wasps (Fig. 1), as the production of VcVLPs determined by transmission electron microscopy (TEM) increases (20).

Second, in order to study VcENV gene transcription kinetics throughout ovarian development, transcriptome sequencing (RNA-seq) analyses were performed on wasp ovaries of P1 to P5 pupal stages and the adult stage (Fig. 1 and 2). As a negative control, VcENV gene expression was also monitored in adult venom glands and anterior body parts (head and thorax).

The expression of three reference control genes was monitored to assess the quality of the samples. *VenA3* encodes a venom protein conserved among Campopleginae wasps (62) and is expected to be specifically expressed in venom glands of the *V. canescens* wasp. The *kynurenine 3-monoxygenase (kmo)* gene is involved in eye pigmentation in insects (63–65) and is expected to be specifically expressed in the head-thorax (Fig. 2). Finally, the *ovarian tumor (otu)* gene is an ovarian gene required for the development of



**FIG 2** VcENV and VcVLP gene expression in *V. canescens* tissues. A heatmap (left) and the associated hierarchical clustering analysis (right) of endogenized alphanodivirus genes (VcENV) and VcVLP virulence gene expression in *V. canescens* ovaries (P1 to P5 and adults) or adult head-thorax and venom glands are presented. The heatmap displays Z scores calculated from normalized TPM (transcript per million reads). *VenA3*, *kynurenine 3-monoxygenase*, and *ovarian tumor* are cellular wasp genes specifically expressed, respectively, in venom gland, head, and ovaries. Data were obtained from two pools of 15 individuals in each condition, except for head-thorax for which only five individuals were used per pool. Genes were annotated according to their functional class as follows: o, reference genes (three cellular wasp genes); X, transcription and DNA amplification (six VcENV genes); #, envelope components and morphogenesis (15 VcENV genes); +, unknown (30 VcENV genes); \$, VcVLP virulence proteins (3 cellular wasp genes). The dendrogram (right) represents the result of a hierarchical clustering analysis performed on Z score values of VcENV and VcVLP virulence gene expression (the tree was cut with  $k = 4$ ; the four groups are represented with bold vertical bars). Z scores are summarized in Table S2 in the supplemental material.

the female germ line in *Drosophila* and is expected to be specifically expressed in ovaries (66). As expected, *VenA3*, *kmo*, and *otu* were specifically expressed in the venom gland, head-thorax, and ovaries of *V. canescens* wasps (Fig. 2), respectively.

Concerning the 51 VcENV genes, their expression is specific to ovarian tissue (Fig. 2), which is consistent with the fact that VcVLPs are only produced in the calyx (50). Four gene clusters could be obtained by a hierarchical clustering analysis. Genes within the first two clusters were classified as “early expressed,” whereas those in the last two clusters were classified as “late expressed” (Fig. 2). The genes supposedly involved in DNA amplification and in virus transcription, such as *helicase*, *lef-8*, *p47*, and *lef-5*, were clustered among early expressed genes (transcripts per million [TPM] are summarized in Table S1 in the supplemental material). *lef-9* and *lef-4* tended to have an early expression, as their expression started at P2 pupae but reached a maximum later in P4 pupae before decreasing (Fig. 2; see also Table S1). For most of the genes potentially involved in VcVLP morphogenesis and coding for envelope products, their expression followed a late pattern. For these late genes, two subpatterns in expression could be detected: the first one contains genes expressed earlier than other late genes and contains all *pif* genes (except one copy of *pif-5*), *Ac81*, *p33*, and *vp91* (Fig. 2; see also Table S1). The second one contains genes such as *OrNVorf136-like*, *OrNVorf23-like*, and *OrNVorf44-like*, which have a very late expression pattern (Fig. 2). For VcENV genes of unknown function, both early and late expression patterns could be observed. Regarding the wasp genes encoding the virulence proteins contained within the VcVLPs (*vlp1*, *vlp2*, and *vlp3*), they exhibit a late expression pattern, which is consistent with VcVLP proteins being packaged at the late stages of VcVLP production (Fig. 2; see also Table S1). These genes encode, respectively, a glutathione peroxidase (PHGPx), a Rho GTPase-activating protein (RhoGAP), and a metalloprotease (nepylisin), which could in theory be transcribed by the cellular RNA polymerase II, as they are cellular wasp genes and do not have a nudiviral origin.

Taken together, the expression dynamics of these VcENV genes throughout wasp ovarian development are consistent with their proposed functions and consistent with expression dynamics observed for baculovirus genes during infection and for endogenized nudiviruses in Braconidae wasps during BV production (46, 56–58, 60). Indeed, VcENV *lef-4*, *lef-8*, *lef-9*, and *p47* are expressed earlier than genes involved in VcVLP morphogenesis and envelope component production, as expected for genes that would encode the viral RNA polymerase regulating such downstream processes.

**Analysis of VcENV gene promoter regions.** To identify potential promoter motifs involved in early or late expression of endogenized nudivirus genes in *V. canescens*, we first searched for promoter motifs already described in exogenous nudiviruses or in baculoviruses (67–70). In nudiviruses, only one late promoter sequence TTATAGTAT was described in HzNV-1 (*Heliothis zea* nudivirus 1) (67). In VcENV, this motif could only be found in the upstream region of *OrNVorf44-like*. In baculoviruses, early genes that are transcribed by the cellular RNA polymerase II possess a TATA box promoter, followed by CA[G/T]T or CGTGC motifs placed among 40 nucleotides downstream of the TATA box. In contrast, late genes that are transcribed by the viral RNA polymerase harbor a [A/T/G]TAAG motif. The results of promoter motif search in VcENV genes are summarized in Table 1. VcENV gene upstream sequences are shown in Fig. S1 in the supplemental material.

According to the RNA-seq analyses (Fig. 2 and previous paragraph), 14 VcENV genes were classified as “early expressed” and 37 as “late expressed” (see third column of Table 1). Baculovirus motifs were found in 37 out of the 51 VcENV genes. For some VcENV genes (such as *lef-4*, *lef-8*, *p47*, and *pif-4* for example), the motif found was coherent with the expression kinetic pattern described above (Fig. 2; last column of Table 1). In some cases, such as *OrNVorf90-like*, the motif search gave ambiguous results as both motif types (early and late) were found. In contrast, for other VcENV genes, such as *lef-5*, *helicase*, or *pif-6*, for example, the baculovirus motif type found did not match the transcriptomic results. Surprisingly, no motif was detected for a few VcENV genes (*p47* and *OrNVorf59-like*, for example). Taken together, baculovirus promoter motifs were found in 73% of all of the

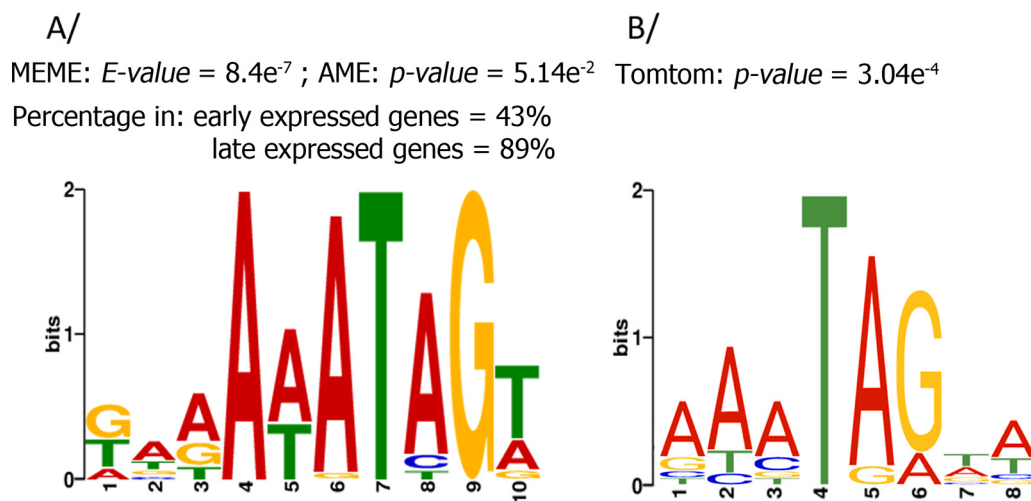
**TABLE 1** Promoter motif detection in VcENV gene upstream sequences

Function predicted from baculoviruses	VcENV gene	Expression kinetic pattern of VcENV gene	Type of baculovirus promoter(s) detected <sup>a</sup>	“Late” motif found by MEME <sup>b</sup>	
DNA replication	<i>helicase</i>	Expressed early	Late		
Transcription	<i>p47</i>	Expressed early	X		
	<i>lef-5</i>	Expressed early	Late		
	<i>lef-4</i>	Expressed early	Early	Late	
	<i>lef-8</i>	Expressed early	Early		
	<i>lef-9</i>	Expressed early	Late	Late	
Unknown	<i>OrNVorf73-like</i>	Expressed early	Late		
	<i>OrNVorf120-like</i>	Expressed early	X		
	<i>OrNVorf90-like</i>	Expressed early	Early/late	Late	
	<i>OrNVorf46-like</i>	Expressed early	Early	Late	
	<i>OrNVorf19-like</i>	Expressed early	X		
	<i>OrNVorf27-like</i>	Expressed early	Late		
	<i>OrNVorf54-like</i>	Expressed early	Late	Late	
Morphogenesis and envelope components	<i>p33</i>	Expressed late	Late	Late	
	<i>pif-6</i>	Expressed late	Early	Late	
	<i>ac81-1</i>	Expressed late	X	Late	
	<i>ac81-2</i>	Expressed late	Early	Late	
	<i>ac81-3</i>	Expressed late	X	Late	
	<i>p74 (pif-0)</i>	Expressed late	Late		
	<i>pif-1</i>	Expressed late	Early	Late	
	<i>pif-2</i>	Expressed late	X	Late	
	<i>pif-3</i>	Expressed late	Late	Late	
	<i>pif-4</i>	Expressed late	Late		
	<i>pif-5-1</i>	Expressed late	Late	Late	
	<i>pif-5-2</i>	Expressed late	Early/late	Late	
	<i>pif-5-3</i>	Expressed late	Late	Late	
	<i>pif-5-4</i>	Expressed late	Late	Late	
	<i>vp91 (pif-8)</i>	Expressed late	Late	Late	
	Unknown	<i>OrNVorf41-like-1</i>	Expressed late	Early	Late
		<i>OrNVorf41-like-2</i>	Expressed late	Late	Late
<i>OrNVorf41-like-3</i>		Expressed late	Early	Late	
<i>OrNVorf41-like-4</i>		Expressed late	X	Late	
<i>OrNVorf41-like-5</i>		Expressed late	Late	Late	
<i>OrNVorf47-like-1</i>		Expressed late	Early/late	Late	
<i>OrNVorf47-like-2</i>		Expressed late	Early/late	Late	
<i>OrNVorf47-like-3</i>		Expressed late	X	Late	
<i>OrNVorf138-like</i>		Expressed late	Late		
<i>OrNVorf18-like</i>		Expressed late	Late	Late	
<i>OrNVorf25-like</i>		Expressed late	X	Late	
<i>OrNVorf61-like</i>		Expressed late	X	Late	
<i>OrNVorf59-like</i>		Expressed late	X		
<i>OrNVorf119-like</i>		Expressed late	X	Late	
<i>OrNVorf63-like</i>		Expressed late	Late	Late	
<i>OrNVorf118-like</i>		Expressed late	X	Late	
<i>OrNVorf123-like</i>		Expressed late	Late	Late	
<i>OrNVorf76-like</i>		Expressed late	Early	Late	
<i>OrNVorf23-like</i>		Expressed late	Late	Late	
<i>OrNVorf136-like</i>		Expressed late	X	Late	
<i>OrNVorf79-like</i>		Expressed late	Late	Late	
<i>OrNVorf44-like</i>	Expressed late	Late/HzNV late	Late		

<sup>a</sup>X, absence of baculovirus promoter sequences; HzNV late, gene for which the HzNV-1 late promoter motif was found.

<sup>b</sup>Empty frames indicate that no motif was found for the respective gene.

VcENV genes. Globally, even if most of the motif search results were congruent with the expression kinetic analysis, it seems hard to predict the expression kinetic pattern of VcENV genes by scanning for baculovirus promoters, as some results seem to be ambiguous and are not congruent with the RNA-seq analyses (Table 1). Similar conclusions had



**FIG 3** Sequence logo identified upstream of VcENV genes using MEME software and *br* gene sequence logo. (A) Significant motif identified from an enriched alignment of late expressed VcENV gene upstream sequences using MEME software. The  $E$  value was calculated by MEME software, and the  $P$  value was calculated using AME software to determine if the motif was significantly more enriched in upstream sequences of late expressed genes. The percentages above indicate the proportion of early and late sequences displaying this motif. (B) Binding site motif of the Broad-Complex protein encoded by the *br* gene (isoform 2) described in *Drosophila melanogaster* (NCBI txid7227) (72).

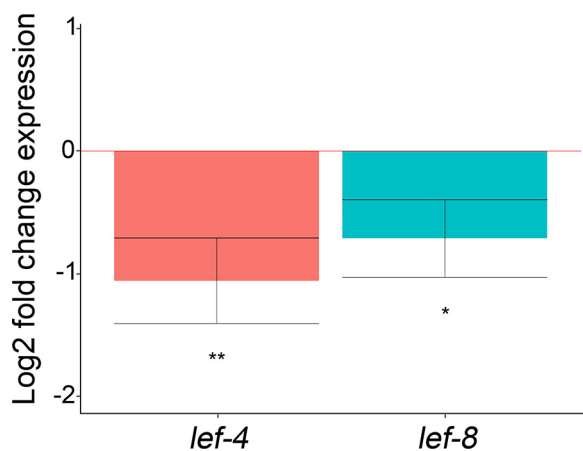
been drawn following the analysis of upstream sequences of MdBV nudivirus genes combined with expression data (38). We therefore investigated for the presence of other motifs in order to find potential regulatory sequences that could better explain the expression kinetic patterns of endogenized alphavirus genes in *V. canescens*.

To this end, a search for new motifs without *a priori*, was performed using MEME software (see Table S3 in the supplemental material). Only one motif was found to have an acceptable  $E$  value inferior to 1 (Fig. 3A,  $E$  value =  $8.4e^{-7}$ , consensus sequence = [G/T][A/T][A/G]A[A/T]ATAG[T/A]). This motif was identified by using the upstream sequences of late expressed genes in which it was shown to be significantly enriched by AME software (adjusted  $P$  value =  $5.14e^{-2}$ ): 33 out of the 37 late expressed genes have this motif in their upstream sequences, whereas only 6 out of the 14 early expressed genes have it (fifth column of Table 1; see also Fig S1). Using Tomtom software and the Jaspar 2018 core database (71), this new motif (Fig. 3A) was shown to align significantly ( $P$  value =  $3.04e^{-4}$ ) to a binding site motif for a Broad-Complex protein encoded by the *Drosophila melanogaster br* gene (isoform 2) (72) (Fig. 3B).

**RNA interference of *lef-4* and *lef-8* early genes has an impact on late VcENV gene expression levels.** In order to determine whether *lef-4* and *lef-8* could encode viral RNA polymerase subunits, thereby regulating the expression of late genes, RNAi technology was applied against them. By interfering with *lef-4* and *lef-8* genes, we expected to reduce the viral RNA polymerase activity and to observe a decrease in late gene expression together with the reduction of VcVLP production. *lef-4* and *lef-8* double-stranded RNA (dsRNA) were synthesized and then injected into hyaline pupae (P1 stage). Controls consisted of injection of *gfp* dsRNA (*dsgfp*) into P1 pupae. To check interference effectiveness, the amount of remaining *lef-4* and *lef-8* RNA and the consequences on late gene expression were measured by reverse transcriptase quantitative PCR (RT-qPCR) on P4 pupae.

First, RNA interference of *lef-4* and *lef-8* was shown to be effective (Fig. 4). Results displayed a reduction of *lef-4* and *lef-8* RNA by one-half and one-third, respectively, compared to the RNA amount measured in control wasps (results are summarized in Table S4 in the supplemental material). In addition, interference of *lef-4* or *lef-8* caused an alteration in the expression of late genes in P4 pupae, with an observable downregulation of *p74*, *pif-4*, and *OrNVorf59-like* expression compared to the *dsgfp* control treatment (Fig. 5A and B; see also Table S4). The expression of *pif-4* was, in particular, drastically altered with, respectively, a 30- to 50-fold reduction (fold change [FC] of





**FIG 4** *lef-4* and *lef-8* gene log<sub>2</sub> fold changes in expression in response to ds*lef-4* (*lef-4* interference) and ds*lef-8* (*lef-8* interference). (Left) *lef-4* interference, log<sub>2</sub> fold change in expression of *lef-4* in P4 pupae after *lef-4* dsRNA injection (*n* ds*gfp* = 14; *n* ds*lef-4* = 14). (Right) *lef-8* interference, log<sub>2</sub> fold change in expression of *lef-8* in P4 pupae after *lef-8* dsRNA injection (*n* ds*gfp* = 29; *n* ds*lef-8* = 26). FC were obtained using  $\Delta\Delta C_T$  from qPCR analysis as  $FC = 2^{(-\Delta\Delta C_T)}$ . Statistically significant results are displayed with asterisks (\*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ). Error bars represent standard error from the mean.

0.03 and 0.02) in *lef-4*- and *lef-8*-interfered wasps. Concerning *vp91*, neither *lef-4* nor *lef-8* seemed to influence its expression.

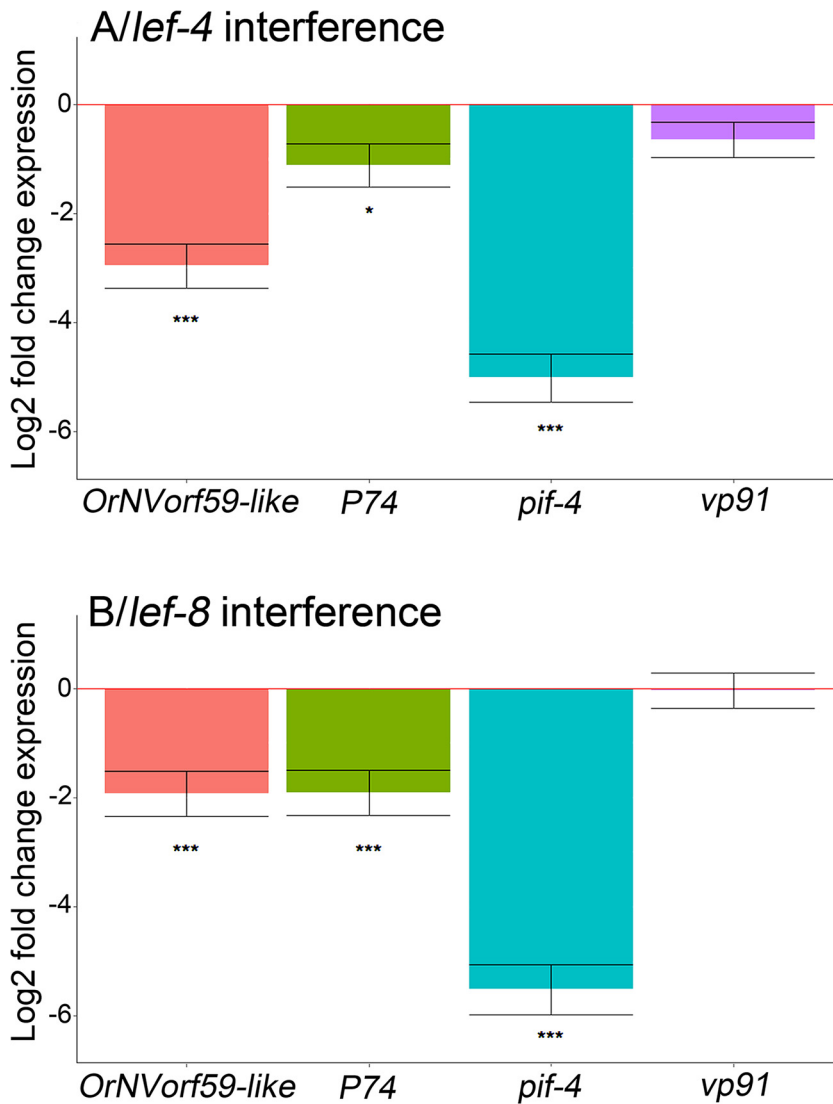
These results show, on the one hand, the effectiveness of the interference on the targeted genes and, on the other hand, the presence of a transcriptional cascade in which late genes (here *p74*, *pif-4*, and *OrNVorf59-like*) clearly require the expression of early expressed genes (*lef-4* and *lef-8*) in order to be transcribed.

#### RNA interference of *lef-4* and *lef-8* early genes affects PIF-4 protein production.

*lef-4* and *lef-8* silencing efficiency to alter the downstream processes involving late expressed genes was then assessed at the protein level by Western blotting analysis. Based on previous quantitative PCR (qPCR) results, an alteration or inhibition of P74, PIF-4, and ORF59 protein synthesis could be expected in wasps injected with *lef* dsRNA. As *pif-4* expression was the most impacted following *lef-4* and *lef-8* RNA polymerase gene interference, the PIF-4 protein was chosen as a candidate to investigate RNAi impact at the protein level by Western blotting analysis in control and silenced wasps.

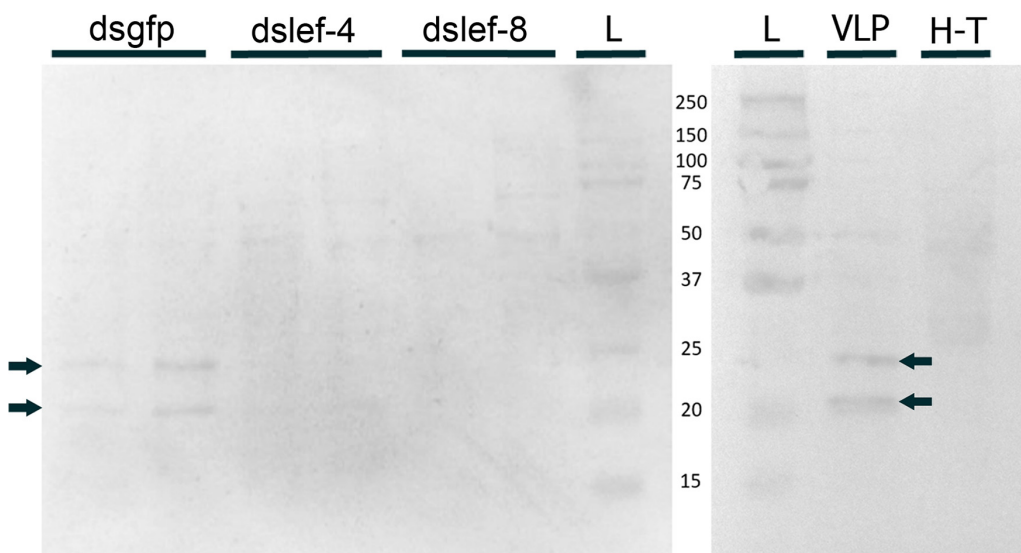
For this purpose, P1 pupae were injected with either *gfp* (control), *lef-4*, or *lef-8* dsRNA. Once wasps reached the adult stage (corresponding to the stage with the highest VcVLP production), the ovaries were dissected and pooled by four individuals. Negative and positive controls corresponding respectively to protein extract from adult head-thorax (with no VLP proteins) and protein extract from purified VLPs of adults were also prepared. Figure 6 shows the results obtained for PIF-4 after the RNAi treatment. In *V. canescens*, *pif-4* is supposed to produce a 236-amino-acid protein with a molecular weight of approximately 27 kDa. In Fig. 6, two bands are visible in the control treatment (ds*gfp*), one displaying a molecular weight of approximately 25 kDa as expected and the other one of approximately 20 kDa, which we interpret as a PIF-4-cleaved product (although bioinformatic analysis did not detect a specific cleavage site in the protein sequence). The two bands observed in the control treatment (ds*gfp*) appear to be specific to PIF-4 protein, since we could observe the same two bands in the positive control (Fig. 6, "VLP" on the right), while no band was observable in the negative control (Fig. 6, "H-T" on the right). We could clearly show that the PIF-4 protein is no longer or barely detectable in pools of *lef-4* and *lef-8* dsRNA-injected wasps. These results show that interference of *lef-4* and *lef-8* not only resulted in downregulation of *pif-4* expression but also in a clear decrease of PIF-4 protein production.

**The RNA interference of *lef-4* and *lef-8* genes has an impact on VcVLP morphogenesis.** As RNAi against *lef-4* and *lef-8* has an impact on PIF-4 production, and as PIF-4 is hypothesized to be a VcVLP component, we investigated by electron microscopy the influence of *lef-4* and *lef-8* RNAi on VcVLP production. The observations



**FIG 5** VcENV gene log<sub>2</sub> fold changes in expression after dsRNA injections of *lef-4* and *lef-8*. (A) *lef-4* interference, log<sub>2</sub> fold changes in gene expression of *OrNVorf59-like*, *p74*, *pif-4*, and *vp91* genes in response to *lef-4* interference in P4 pupae. (B) *lef-8* interference, log<sub>2</sub> fold changes in gene expression of *OrNVorf59-like*, *p74*, *pif-4*, and *vp91* in response to *lef-8* interference in P4 pupae. Numbers of dsRNA injected P4 wasps are *n* ds $\Delta$ gfp = 14 (control), *n* ds $\Delta$ lef-4 = 14, and *n* ds $\Delta$ lef-8 = 13. FC was obtained using  $\Delta\Delta C_T$  from qPCR analysis as FC = 2<sup>(- $\Delta\Delta C_T$ )</sup>. Captions are displayed as indicated in Fig. 4; \*\*\*, *P* < 0.001.

were carried out on the ovaries of adult individuals 9 days post dsRNA (*gfp*, *lef-4*, and *lef-8*) injection. In all conditions, eggs were visible in ovarioles (data not shown). Control *gfp*-interfered wasps displayed the same calyx phenotype as noninterfered wasps, i.e., the injection did not alter VcVLP production (see Fig S2 in the supplemental material). In the ds $\Delta$ gfp control treatment (Fig. 7), VcVLPs were produced in calyx cell nuclei (white arrow in Fig. 7, Ba and Ca) by the virogenic stroma (VS) (a “viral factory,” containing VcVLP components and recruited envelopes, where VcVLP assembly takes place) and transited through the nuclear and plasma membranes by budding to end in the calyx lumen where they accumulated (Fig. 7, Ca). In the nucleus, VcVLP empty envelopes required for VcVLP formation can be observed (white circle in Fig. 7, Ba). Strikingly, in *lef-4*-interfered wasps, there was no production (or a scarce production) of VcVLPs (Fig. 8) despite the presence of VS. Surprisingly, we also observed reproducibly that the calyx lumen was closed compared to that in *gfp*-interfered wasps at the same



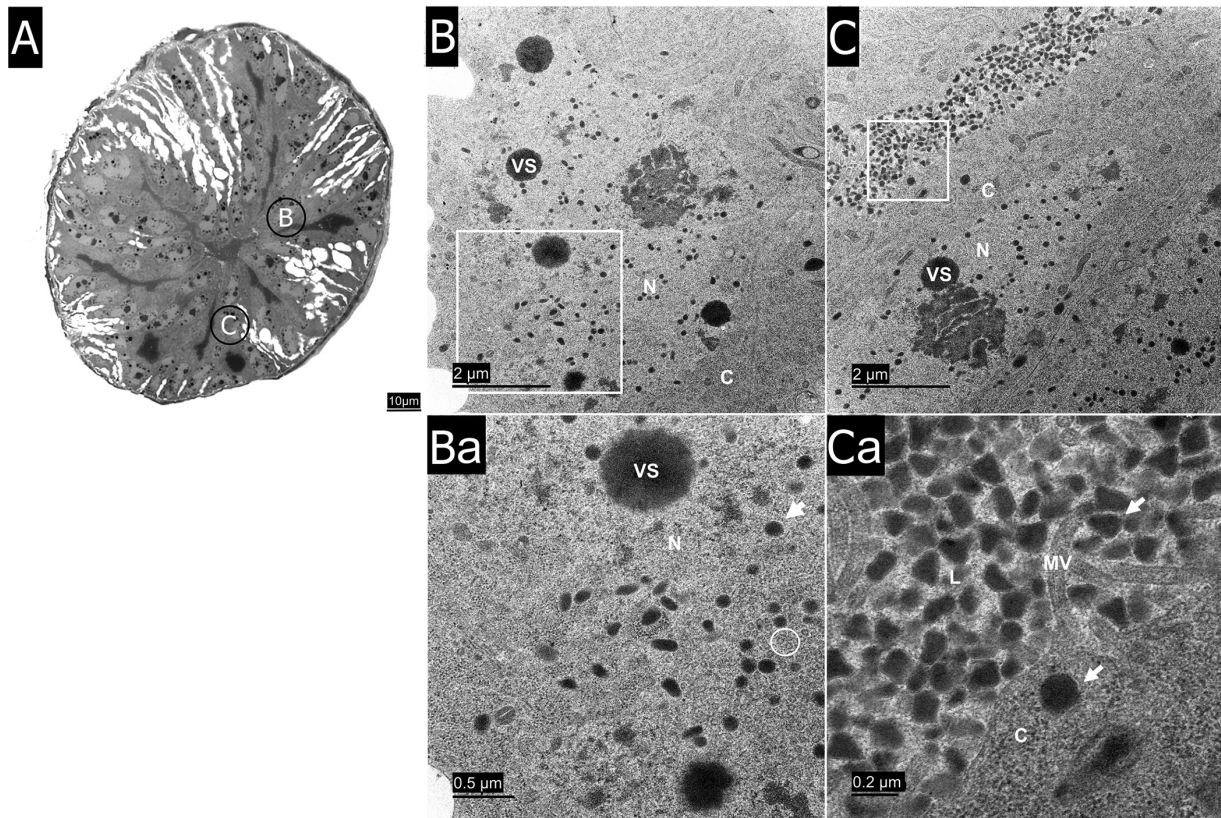
**FIG 6** *lef-4* and *lef-8* silencing impact on PIF-4 production. Western blotting analysis of PIF-4 production (shown with arrows) in ovaries after dsRNA injections (left). Western blotting analysis of PIF-4 production in positive control (VLP) and negative control (H-T) (right). L refers to the protein molecular size ladder (in kDa). dsgfp, dslef-4, and dslef-8 refer to *gfp* dsRNA, *lef-4* dsRNA, and *lef-8* dsRNA injection treatments, respectively. VLP and H-T refer to the purified VLPs and the protein extract from head-thorax, respectively.

developmental stage (compare Fig. 7 with Fig. 8 and Fig. 7, C with Fig. 8, Ca). Furthermore, we were not able to distinguish envelopes around the few VcVLPs found in calyx cells (white arrowhead in Fig. 8, Ba), and no empty envelopes were visible in the cell nucleus, suggesting that envelope components are no longer produced in *lef-4*-interfered wasps. In the dslef-8 treatment (Fig. 9), fewer VLPs were produced by the VS than in the control treatment; indeed, fewer VLPs were visible both in the cells and in the calyx lumen. Both empty envelopes (white circles, Fig. 9, Ba) and empty VcVLPs could be observed in the cell nuclei (black arrows, Fig. 9, Ba and C) of *lef-8*-interfered wasps. Empty VcVLPs are similar in size to normal VcVLPs, but they seem to be devoid of VcVLP content. All together, these results demonstrate the primordial role of *lef-4* and *lef-8* expression in the downstream process involving VcVLP component production and consequently VcVLP morphogenesis and synthesis.

## DISCUSSION

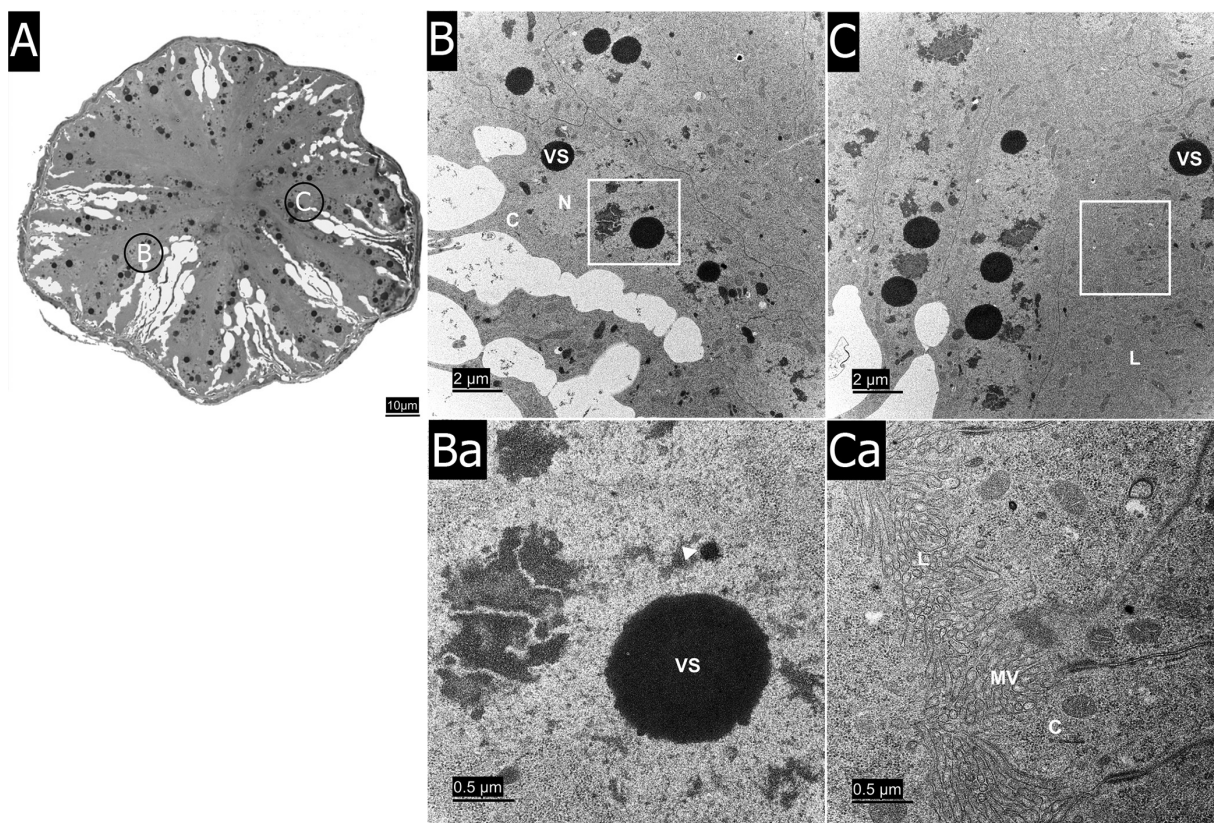
Independent endogenization events of large dsDNA viruses have been observed in several lineages of Ichneumonoidea parasitoid wasps. Key questions regarding these events concern the impact of the wasp genomic environment on viral functions. Our results suggest that the process of viral domestication in the ichneumonid Campopleginae wasp *V. canescens* has allowed the endogenized alphanudivirus to retain essential viral functions. We have shown that VcENV genes have conserved an expression kinetic resembling that of baculovirus homologues during infection and of endogenized nudivirus homologues during particle production in Braconid parasitoid wasps (46, 56, 60). Furthermore, genes involved in viral RNA polymerase synthesis were shown to be required for the expression of *pif* genes involved in VcVLP production and for the correct production of the particles. In accordance with results obtained for MdBV (17), we have shown that VcVLP production is impacted when interfering with *lef-4* and *lef-8* expression. Our results also highlight, as for endogenized betanudivirus genes in *M. demolitor*, that baculovirus promoter motifs have been conserved to some extent, but other regulatory sequences of nudiviral or of wasp origin could be involved. Our results, together with previous literature, show experimentally that the nudivirus RNA polymerase is essential in both BV and VLP morphogenesis, but resulting particles differ because some nudivirus genes retained in BV-producing wasps have been pseudogenized or lost in *V. canescens* (17, 20, 45).

Expression dynamics of VcENV genes throughout wasp ovarian development were



**FIG 7** *V. canescens* adult wasp calyx observed by electron microscopy after green fluorescent protein (GFP) dsRNA injection. (A) Semithin section picture. (B and C) Electron transmission microscopy pictures of calyx portions specified on panel A. (Ba and Ca) Magnified sections (white frame) of panels B and C, respectively. VcVLPs (white arrows) are produced by the virogenic stroma (VS) recruiting VcVLP empty envelopes (white circles) localized in the calyx cell nucleus (N). VLPs pass through the nuclear membrane, the cytoplasm (C), and cytoplasmic membrane microvilli (MV) to end in the calyx lumen (L).

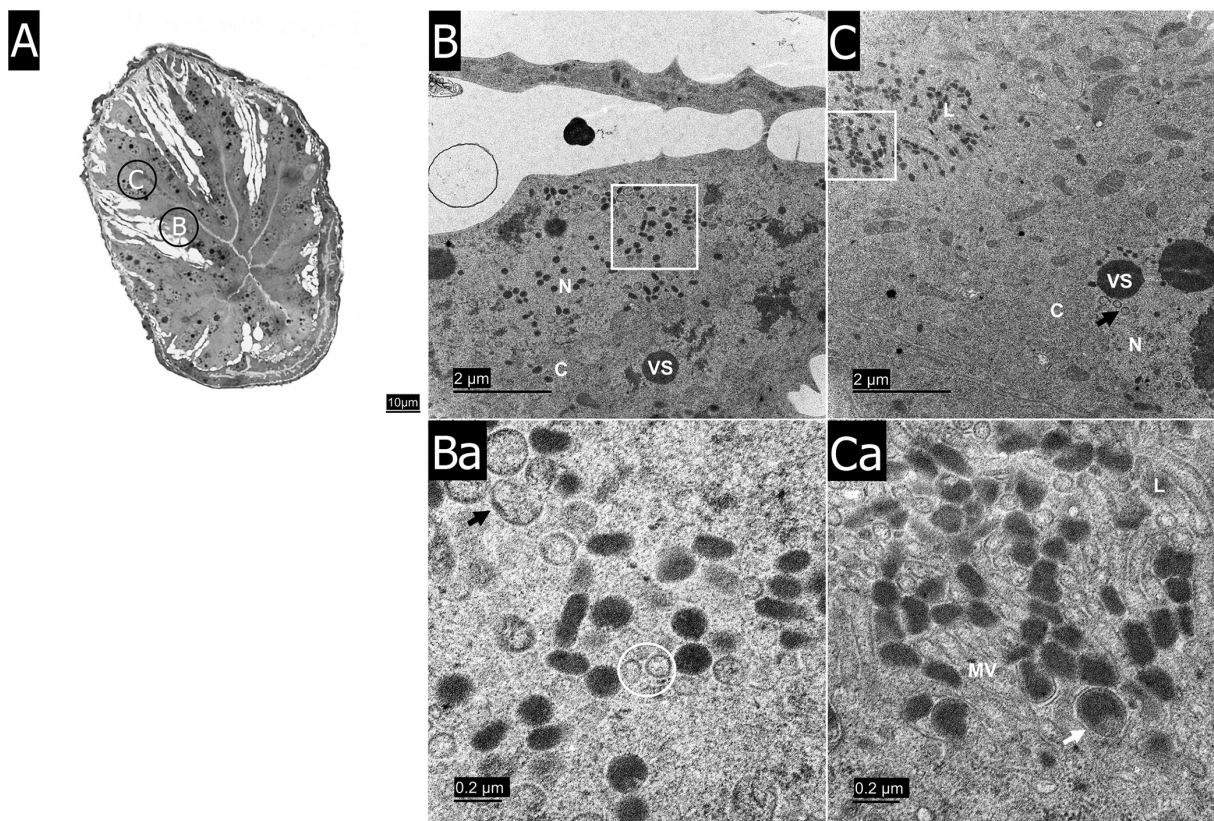
consistent with expression dynamics observed for baculovirus genes during infection and for endogenized nudiviruses in Braconidae wasps (46, 56–58, 60). Indeed, VcENV *lef-4*, *lef-8*, *lef-9*, and *p47* are expressed earlier than genes involved in VcVLP morphogenesis and envelope component production (*pif*), as expected for genes that would encode the viral RNA polymerase regulating such downstream processes. Analysis at all five pupal stages combined with a hierarchical clustering analysis allowed us to refine and observe subtle differences in expression dynamics in *V. canescens*. Indeed, among early expressed genes, *helicase*, *lef-8*, *lef-5*, and *p47* were shown here to be expressed earlier than *lef-9* and *lef-4*. The *helicase* is a key DNA replication protein in baculoviruses (57) and may contribute to VcENV gene cluster amplification as suggested by Leobold and collaborators (45). Cluster amplification may occur early in ovaries to allow the expression of late genes and massive VcVLP production. In the same vein, in baculoviruses, *lef-5* encodes an initiation factor (31, 68, 73) and *lef-8* the catalytic site of the RNA polymerase (74, 75), which could explain their very early expression in *V. canescens*. These differences in expression concerning “early” genes were not observed for the endogenized betanucleoviruses in *M. demolitor* and *Cotesia congregata*, either because less stages were tested and/or because expression dynamics may slightly differ between species (56, 60). Our results show that the *V. canescens pif* genes are, as for baculoviruses and BVs, expressed in the late pupal and adult stages, suggesting that they may similarly encode VcVLP envelope proteins (13, 38, 57, 58, 60). The late expression of *V. canescens* VP91, which in baculoviruses is involved in occlusion body morphogenesis (76), is suggestive that a similar function could be extrapolated for *V. canescens* VP91. Finally, VcENV genes appear to be specifically expressed in



**FIG 8** *V. canescens* adult wasp calyx observed by electron microscopy after *lef-4* dsRNA injection. (A) Semithin section picture; (B and C) Electron transmission microscopy pictures of calyx parts specified in panel A; (Ba and Ca) Magnified sections (white frame) of panels B and C, respectively. Practically no VcVLPs were observable in the nucleus (N) and the cytoplasm (c). A rare VcVLP lacking a visible envelope is indicated by a white arrowhead (Ba), near a virogenic stroma (SV). Only cytoplasmic membrane microvilli (MV) were visible in the calyx lumen (L) location (C and Ca).

ovarian tissues, indicating that host regulation mechanisms that have yet to be uncovered are acting to allow control of viral gene expression in calyx cells.

Baculovirus promoter motifs were found in the upstream sequences of some VcENV genes, but their involvement in gene regulation remains unclear since the presence of these motifs was not always congruent with VcENV gene expression profiles (Table 1 and Fig. 2). Indeed, 27 VcENV genes out of 51 do not possess a baculovirus promoter type that matches with their expression kinetics. This lack of correlation is not totally surprising since similar observations were reported for some baculovirus genes (69) and also for endogenized nudivirus genes in *M. demolitor* (38). The production of polycistronic mRNAs, as recently described in baculoviruses (77), could explain some expression patterns. For example, no baculovirus motif was found in the upstream sequence of *OrNVorf59-like*, but it displays a late expression pattern. The *pif-4* gene, which is located upstream of *OrNVorf59-like*, possesses a late baculovirus promoter; the late expression pattern of these genes could, therefore, be explained by the production of a polycistronic mRNA including these two genes. It is important to note that baculovirus promoters and noncanonical TATA boxes are difficult to predict, especially in the context of an insect genome, as these sequences are small (10-bp maximum). It is also important to bear in mind that the exogenous baculovirus and endogenized alphanudivirus regulatory systems are difficult to compare, as (i) nudivirus regulatory sequences may have diverged from regulatory sequences described in baculoviruses, and (ii) these sequences are not evolving within the same genetic background (the endogenized virus evolving within the wasp genome). Some of the viral regulatory sequences could, thus, be altered and not conserved in the eukaryotic genome. In



**FIG 9** *V. canescens* adult wasp calyx observed by electron microscopy after *lef-8* dsRNA injection. (A) Semithin section picture; (B and C) Electron transmission microscopy pictures of calyx parts specified in panel A; (Ba and Ca) Magnified sections (white frame) of panels B and C, respectively. VcVLPs and empty envelopes are indicated by white arrows (Ca) and a white circle (Ba), respectively. VcVLPs were visible in the nucleus (N), the cytoplasm (c) and the lumen (L) surrounded by cytoplasmic membrane microvilli (MV). Empty VcVLPs were visible in the nucleus (Ba and C) and are indicated by black arrows.

consequence, given the discrepancy between observed baculovirus promoter types and expression kinetics of alpha and betanucleovirus genes (38), we searched for new motifs in VcENV promoter sequences. We discovered a motif enriched in upstream sequences of late expressed genes, which aligned with a binding site of a Broad-Complex protein encoded by the *br* gene of *Drosophila melanogaster*. This gene encodes a C2H2 zinc finger transcriptional regulator protein involved in transcription of late genes required for pupal development (78, 79) for microRNA (miRNA) regulation (80) and for metamorphosis (72, 78, 79). Some VcENV genes may, therefore, have acquired binding sites of a cellular transcription factor, which might enhance viral RNA polymerase activity in order to match the developmental processes of the host parasitoid wasp. However, a functional validation should be performed in order to validate this hypothesis. Taken together, our results and previous work by Burke et al. (38) suggest that some baculovirus promoter sequences were retained during the evolution and the domestication of alphanucleoviruses. Furthermore, the expression of some endogenized nucleovirus genes probably rely also on wasp regulatory sequences. Since other regulatory mechanisms have yet to be described in both baculoviruses and nucleoviruses, further studies would be needed to better describe regulatory sequences in VcENV. Globally, it will be interesting to characterize regulatory motifs of pathogenic nucleoviruses, which probably reflect those of the ancestral virus captured by a *V. canescens* ancestor. Another approach would be to investigate transcription factor binding sites involved in VcENV regulation by using chromatin immunoprecipitation sequencing (ChIP-Seq) analysis targeting LEF-5 protein (a transcription initiation factor in baculovirus [81]) and the homologue of the Broad-Complex protein in *V. canescens*.

Since the expression kinetics of alphanucleovirus genes supposedly involved in the

production of the viral RNA polymerase are in accordance with the expression kinetics described in baculoviruses and certain bracoviruses, we tested the function of two *lef* genes by RNAi, expecting that the knockdown of the viral RNA polymerase will both alter expression of late VcENV genes and VcVLP production. Indeed, by interfering *lef-4* and *lef-8*, we decreased the available amount of *lef-4* and *lef-8* transcripts that are, respectively, supposed to code for the viral RNA capping enzyme subunit (Fig. 4A) (82–84) and the catalytic site subunit of the viral RNA polymerase (Fig. 4B) (74, 75). Subsequently, we investigated the effect of the knockdown of these early expressed genes on expression of late VcENV genes: *lef-4* and *lef-8* RNAi lead also to the downregulation of *p74*, *pif-4*, and *OrNVorf59-like* expression in P4 pupae, where their expression was lowered approximately 15 times, 40 times, and 5 times, respectively (Fig. 5). We could also confirm that *lef-4* and *lef-8* interference impacted the production of the PIF-4 protein. These results are concordant with RNA interference experiments conducted in the *Microplitis demolitor* braconid wasp, in which knockdown of *lef-4* and *lef-9* also led to the downregulation of late genes encoding BV structural proteins (17). In contrast, *vp91* expression was not affected by *lef-4* or *lef-8* interference (Fig. 5; see also Table S3 in the supplemental material). This could suggest that the expression of this gene is either not entirely dependent on the viral RNA polymerase or that the RNAi efficiency is not sufficient to see a significant decrease in *vp91* expression in P4 pupae. Indeed, *vp91* is expressed relatively early, from the P3 pupal stage (Fig. 2), whereas *p74*, *pif-4*, and *OrNVorf59-like* are expressed later from the P4 pupal stage. Regarding *OrNVorf59-like*, the promoter of this gene did not harbor any baculovirus motif, but its expression was nonetheless affected by *lef* gene interference. Therefore, either the expression of this gene relies on another type of promoter not described to date in baculoviruses, or simply not present in baculoviruses, that may correspond to a specific promoter sequence of alphanodiviruses or a wasp promoter as discussed above.

The importance of *lef-4* and *lef-8* expression for VcVLP production was confirmed by electron microscopy. In *lef-4*-interfered individuals, VcVLP production was abolished, the rare particles that were produced seemed to lack an envelope, and no empty envelopes were visible in the nucleus. Previously, Pichon and collaborators (20) described that VS recruit empty envelopes and fill them with their content to form VcVLPs. Despite the fact that *lef-4*-interfered individuals lacked VcVLPs, VS were still observable. This could be explained either by the fact that VS are composed essentially of proteins encoded by genes transcribed by the cellular RNA polymerase II (such as the three virulence proteins [20]) that retain their ability to aggregate within the nucleus or that the effect of RNA interference is partial or decreasing through time, allowing VS formation. In contrast, in *lef-8*-interfered wasps, VcVLP production and release in the calyx lumen were observable but in much smaller amounts compared to control individuals. Also, many empty VcVLPs were observable in the nucleus of *lef-8*-interfered individuals. The observed differences in VcVLP production and phenotype in wasps interfered with *lef-4* or *lef-8* dsRNA could be explained by either their impact on late VcENV gene expression or by the interference efficiency on each target. As no empty envelope was seen in wasps subjected to the *lef-4* dsRNA treatment, it could also mean that the expression of *lef-4* (RNA capping enzyme supposedly involved in RNA stability) is important for the expression of envelope proteins or for envelope recruitment. As we have shown that RNA interference is a suitable approach to silence genes and that *lefs* have conserved their viral RNA polymerase function in *V. canescens*, we could investigate the impact of viral RNA polymerase subunit genes on all nudivirus genes using RNA-seq analyses. Combined with electron microscopy observations, this approach used on other viral RNA polymerase subunit genes might also give better insights on how the viral machinery is functioning in the parasitoid wasp.

Regarding calyx morphology, it was possible to observe in the *lef-4* dsRNA treatment that, in the absence of VcVLPs in the calyx lumen, cell microvilli seem to join to each other and close the calyx lumen. In *lef-8* dsRNA treatment, the calyx lumen was still observable but not as wide as the calyx lumen seen in the control treatment. All of

these observations suggest that the calyx morphology in this species is impacted by VcVLP production, which could by consequence affect egg transit through the ovarian track.

These results point out the practicality of RNA interference experiments for gene function prediction testing in *V. canescens*: for the first time, RNAi was performed in *V. canescens* and allowed to investigate the function of *lef-4* and *lef-8*. In the same way as for *Microplitis demolitor* BV and *Hyposoter didymator* IV, for which RNAi allowed researchers to uncover endogenized viral gene functions (17, 85, 86), RNAi in *V. canescens* will allow us to test alphanudivirus gene predicted functions based on what is known for baculoviruses and bracoviruses and also to investigate the role of alphanudivirus genes with unknown functions. This research also opens the avenue for the investigation of cellular gene function in *V. canescens*. For example, by using RNAi, the function of *vlp1*, *vlp2*, and *vlp3* genes coding for proteins that are contained in VLPs and that have been proposed to play essential roles in host immunity alteration and in VcVLP formation (48, 61, 87–89) could be accurately tested. Our results regarding the evolution of an endogenized alphanudivirus, together with previous data obtained on endogenized betanudiviruses in Braconidae (13, 17, 56, 60), reveal that endogenization of nudiviruses in parasitoid wasps has repeatedly led to conservation of the viral RNA polymerase function, allowing the production of viruses or virus-like particles that differ in composition but enable wasp parasitic success. We extend to virus-like particles the finding that, like bracovirus-associated wasps, viral transcription still plays a key role in particle production, suggesting that conservation of this function could be crucial in maintaining a viral entity after nudivirus endogenization. This is suggestive that strict control and expression of the nudivirus genes together with the loss of the DNA polymerase gene are important steps leading to mutualist association between nudiviruses and host wasps.

Finally, discovering other endogenized nudiviruses in Campopleginae parasitoid wasps phylogenetically related to *V. canescens* will allow us to better understand the evolution of nudiviruses under viral domestication and to investigate the function of alphanudivirus genes together with viral domestication mechanisms.

## MATERIALS AND METHODS

**Parasitoid model. (i) Parasitoid rearing.** The thelytokous strain (i.e., composed only of females) of *V. canescens* (Ichneumonidae, Campopleginae) used for this study originated from a natural population of wasps collected in Valence (France) (20). This strain was bred on *Ephestia kuehniella* (Lepidoptera, Pyralidae), a natural host of *V. canescens*, on organic whole wheat flour in a climatic chamber at 25°C under a 16-h light/8-h dark photoperiod. Wasp pupae were collected using pre-cut cardboard that traps host larvae. For all experiments, first pupal stage wasps (P1) were kept and maintained in petri dishes (25°C, 16h light/8h dark photoperiod) with a piece of moist tissue until they reached the stage of interest.

**(ii) Development stage determination.** Wasp pupae were sorted based on their observed melanization pattern using a binocular magnifier according to Reineke and collaborators (2002) (61) (Fig. 1). Pictures of dissected wasp ovaries were taken with a camera (Leica IC80 HD) using a binocular magnifier for each developmental stage.

**RNA-seq analysis. (i) RNA extraction and sequencing.** Wasps were dissected in order to obtain ovaries (15 pairs per replicate) of all developmental stages (the 5 pupal stages and adult; the age post-pupation for each stage is described in Fig. 1). Venom glands (15 pairs per replicate) and head-thorax (5 per replicate) of adults were also conserved in order to check for absence of ectopic expression of VcENV genes. Two replicates per kind of dissected tissues were analyzed using RNA-seq approaches. Total RNA was first extracted from all samples (16 in total) using TRIzol reagent (Invitrogen) according to the manufacturer's instructions. Only the isopropanol step was replaced by cold absolute ethanol (3 volumes per sample instead of 1), and total RNA were finally resuspended in 30  $\mu$ L RNase-free water. Total RNA was quantified using the Qubit RNA HS assay kit (Invitrogen) designed for the Qubit 2.0 fluorometer. Samples of extracted RNA were then stored at  $-80^{\circ}\text{C}$  until their use.

The sequencing of library preparations and the generation of paired-end reads were performed on an Illumina platform by the Novogene company (UK). Raw data are available in NCBI (BioProject accession number PRJNA739064).

**(ii) Alignment and statistical analyses.** Paired-end reads of each pool were mapped to the reference genome of *V. canescens* (genome assembly, *Venturia canescens* genome v1.0, downloaded from BIPAA database, available at [https://bipaa.genouest.org/sp/venturia\\_canescens/download/genome/v1.0/](https://bipaa.genouest.org/sp/venturia_canescens/download/genome/v1.0/)) using HISAT2 alignment software (90) with default parameters. The featureCounts program (91) with default



parameters from the Rsubread R package (v2.2.6) (92) was then used to access the number of reads mapped to each gene, named as “counts,” of the annotated reference genome.

The generated counts were used to perform gene expression analysis on VcENV genes according to the developmental stage and the studied tissue. First of all, these generated counts were converted into transcripts per million (TPM) by dividing the reads per million (RPK) (i.e., the number of counts/length of the gene) for each gene by a per million scaling factor (total number of RPK  $\times$  1/10<sup>6</sup>).

Genes with a raw count lower than 15 in all libraries and genes for which the total raw count (the sum of count of all libraries) was less than 120 were discarded from the analysis (see Fig. S3 in the supplemental material). TPMs were then normalized by a normalization factor using the edgeR TMM (trimmed means of M values [93]) method (see Fig. S4 in the supplemental material). The reproducibility of the two replicates for each condition was then assessed by a Spearman correlation using the filtered and normalized TPMs (see Fig. S5 in the supplemental material).

To examine the differential gene expression between each condition (developmental stages and organs) and then assess the VcENV gene kinetics, a quasi-likelihood negative binomial generalized log-linear model was fitted to the data after estimation of the common dispersion using edgeR (94, 95). Empirical Bayes quasi-likelihood *F*-tests were performed to identify differentially expressed genes where each condition was compared to each other and the *F*-test *P*-values were then adjusted using false discovery rate (FDR) method (95).

For data visualization, Z scores were calculated for each gene. The Z score corresponds to the number of standard deviations separating the result value from the mean value (96). Z scores for each gene were calculated by dividing the difference between the TPM value obtained for a condition and the TPM mean of all conditions by the standard deviation. A hierarchical clustering analysis was then performed on these Z score values using the stats R package (v4.0.3).

**(iii) Promoter analysis.** In Baculovirus, early genes that are transcribed by the cellular RNA polymerase II have a TATA box promoter followed by a CA[G/T]T or a CGTGC motif placed among 40 nucleotides downstream of the TATA box. In contrast, late genes that are transcribed by the viral RNA polymerase harbor a [A/T/G]TAAG motif (68–70). Early and late promoter motifs were searched for in the 300-bp upstream sequences of VcENV genes; for early promoters, we looked for TATA boxes (TATA[A/T][A/T], allowing two errors in the last 3 nucleotides) with one of the two early motifs (CA[G/T]T and CGTGC) placed among 40 nucleotides downstream of the TATA box. For late promoters, we looked for the [A/T/G]TAAG baculovirus motif and also for the TTATAGTAT motif supposed to regulate late genes in *Heliothis zea* nudivirus-1 (67). These results were compared with VcENV gene expression obtained by RNA-seq.

In parallel, a *de novo* motif search was performed within the 300-bp sequence located upstream of nuclear genes using MEME (Multiple EM for Motif Elicitation [97]). AME (Analysis of Motif Enrichment [98]) and Tomtom (motif comparison tool [99]) software are available in MEME Suite 5.4.1 (<https://meme-suite.org/>). To search for motifs, two databases were analyzed independently. One of the databases included the 300 bp upstream of 14 early expressed genes, and the other one included the 300 bp upstream of 37 genes detected as late expressed (Fig. 2 and Table 1; see also Fig. S1 in the supplemental material). The first step consisted of searching within the two databases for motifs spanning between 4 and 20 nucleotides that were present in at least 70% of the analyzed sequences. To this effect, the MEME software (97) was used in command line with the following parameters: *-dna* (DNA alphabet), *-zoops* (zero or one occurrence per sequence), *-minw* 4, *-maxw* 20 (size of the motifs between 4 and 20 bp), *-nmotifs* 20 (search for 20 different motifs), and either *-minsites* 10 (present in at least 10 sequences) for the “early database” or *-minsites* 26 (present in at least 26 sequences) for the “late database” in accordance with the previously established thresholds. The second step of the analysis consisted in checking if the motifs were well enriched in their respective database. To that purpose, we checked if the early motifs found are enriched in the early database compared to the late database and, conversely, in the case of motifs found in the late expressed sequences. This analysis was performed with the AME software (98) with the default parameters (one-tailed Fisher's exact test). Finally, motifs found enriched according to AME were compared using Tomtom software (99) against the nonredundant JASPAR 2018 core DNA database (71) to find a potential homology with known transcription factor binding site profiles.

**RNA interference approach. (i) Double-stranded RNA synthesis.** In order to silence the VcENV *lef-4* and *lef-8*, RNAi was performed using corresponding dsRNA synthesized from wasp total RNA (100). Individuals were first individually disrupted using a pellet pestle. Total RNA was extracted from entire wasps (4th pupal stage, P4, at 6 days postpupation) using the NucleoSpin RNA kit (Macherey-Nagel) and quantified with a Qubit 2.0 (Invitrogen) using the Qubit RNA HS assay kit (Invitrogen). Entire cDNAs were obtained by retrotranscription of the total RNA (1  $\mu$ g of RNA per sample) with oligo(dT) primer using the QuantiTect reverse transcription kit (Qiagen) according to the manufacturers' instructions.

*lef-4* and *lef-8* cDNAs were specifically amplified by two consecutive PCR, the first one designed to amplify *lef-4* and *lef-8* cDNA (starting at 95°C for 5 min followed by 25 cycles [95°C for 1 min, 58°C for 30 s, and 72°C for 30 s] and ending at 72°C for 10 min) and the second one to add T7 sequences (TAATACGACTCACTATAGGGAGA) on each generated fragment for the transcription (starting at 95°C for 10 min followed by 30 cycles [95°C for 30 s, 58°C for 45 s, and 72°C for 1 min] and ending at 72°C for 10 min). Fifty nanograms of template in the Eppendorf Mastercycler Nexus GX2 thermal cycler was used for each PCR. Transcription was performed using the T7 RNA polymerase enzyme included in the Invitrogen MEGAscript T7 transcription kit (Invitrogen) with the manufacturer's protocol in order to obtain dsRNA of *lef-4* and *lef-8*. For each gene, 150 ng of T7 template (each PCR product added in an equal amount) was used in total to perform the transcription. The transcription was performed in the Eppendorf ThermoMixer C for 16 h at 37°C, and 1  $\mu$ L of Turbo DNase (Invitrogen) was added for a further

15 min at 37°C to degrade the DNA template. One-twentieth volume of NaOAc (3M) and two volumes of ethanol (100%) were then added to precipitate the dsRNA, which was then stored at -80°C for at least 2 h. Thawed dsRNA solutions were centrifuged 15 min at 11,000 rpm. RNA pellets were washed with 70% ethanol and air dried after centrifugation to remove all ethanol. The dsRNA pellets were finally resuspended in 15  $\mu$ L RNase-free water. In parallel, a third dsRNA dedicated to the experimental control was obtained by producing *gfp* dsRNA with the same methods from the pmaxGFP plasmid (LONZA). All primer sequences used for this experiment are summarized in Table S6 in the supplemental material.

Produced dsRNAs were quantified using the Qubit RNA HS assay kit (Invitrogen) and the Qubit 2.0 fluorometer. The dsRNA fragment size was checked by electrophoresis before each injection experiment.

**(ii) dsRNA injection procedures.** Newly synthesized dsRNAs were individually injected into wasps at the 1st pupal stage (hyaline pupae, P1) using the Nanoject III (Drummond) microinjector in order to silence targeted genes and to test their functions in downstream processes such as VcENV gene expression and VcVLP production. Each wasp received 50 nL of dsRNA solution containing 400 ng/ $\mu$ L of either *lef-4*, *lef-8*, or *gfp* (experimental control) dsRNAs. Wasps in the 4th pupal stage (P4 pupae) were then collected for RNA extraction in order to perform qPCR analysis.

For Western blotting and electron microscopy experiments, wasps were collected at emergence (adult stage), as it corresponds to the stage where the VLP production is at its maximum.

**(iii) RT-qPCR analyses.** The impact of *lef-4* and *lef-8* gene silencing was evaluated postinjection at the transcription level by RT-qPCR. For this purpose, wasps at the 4th pupal stage were collected and treated individually. All wasps were injected at the first pupal stage (P1 pupae). Total RNA extraction was performed as previously described for the dsRNA production to the exception that the DNA digestion step was performed twice and total RNA was eluted in 60  $\mu$ L RNase-free water. All of the samples were quantified using the Qubit RNA HS assay kit with the Qubit 2.0 fluorometer. Fifty picograms of kanamycin RNA (Promega) were added to each sample as an external retrotranscriptional control. Reverse transcription was then performed from 1  $\mu$ g RNA per sample using the QuantiTech reverse transcription kit (Qiagen) according to the manufacturers' instructions. Newly synthesized total cDNAs were finally stored at -20°C until their use.

Gene expression study was performed using the QuantStudio 6 real-time PCR system (Thermo Fisher Scientific). The qPCR experiments were run in triplicate (technical replicates) using the standard mode, and reaction mixtures included 3  $\mu$ L cDNA (1  $\mu$ g), 6.5  $\mu$ L MESA blue qPCR mastermix plus (2 $\times$ ) for Sybr assay reagent for SYBR green (Eurogentec), 2.45  $\mu$ L mixed F/R primers (20  $\mu$ M each), and 1.05  $\mu$ L molecular grade water. The expressions of the following different nudiviral genes were evaluated: *lef-4*, *lef-8*, *vp91*, *OrNVorf59-like*, *p74*, and *pif-4*. The following internal control genes were also quantified to normalize gene expression and to check transcription efficiency: *actine*, *gapdh*, and *ef1* as well as kanamycin as an external control. Used primers are presented in Table S7 in the supplemental material.

The following analyses were performed based on the obtained cycle threshold ( $C_T$ ) values using R (version 3.5.3). Data from 14 P4 pupae injected with *gfp* dsRNA (*dsgfp*), 14 with *lef-4* dsRNA (*dslef-4*), and 13 with *lef-8* dsRNA (*dslef-8*) were analyzed to access the expression of *lef-4* and late expressed genes. To verify *lef-8* inactivation, data were generated from 29 (*dsgfp*) and 26 (*dslef-8*) P4 pupae. Targeted gene  $C_T$  values were first normalized with the kanamycin RNA  $C_T$  value (the external control) and then with the reference gene  $C_T$  values (*actine*, *gapdh*, and *ef1*). To this end, an "artificial" reference gene was built from the geometrical mean of the three reference genes (*actine*, *gapdh*, and *ef1*). Normalized  $C_T$  values, named as  $\Delta C_T (=C_{T\text{reference gene}} - C_{T\text{studied gene}})$ , were then analyzed by a linear regression model (R Stats package, v4.0.3) with the interference conditions as factor. For *lef-4* and *lef-8* expression, comparison between the tested treatments (*dslef-4* and *dslef-8*) and the control treatment (*dsgfp*) was performed using a Welch's heteroscedastic *F*-test. For late expressed gene expression, each interference condition (*dsgfp*, *dslef-4*, and *dslef-8*) was compared to the others with a multiple comparison test based on least-squares means of the linear regression model, and *P* values were then adjusted using the Bonferroni method (lsmmeans R package, v2.30.0). The equation  $\Delta\Delta C_T (= \Delta C_{T\text{dsgfp(control)}} - \Delta C_{T\text{dslef-4 or dslef-8}})$  was calculated and used to finally assess the fold change expression ( $FC = 2^{(-\Delta\Delta C_T)}$ ).

**(iv) Protein isolation, SDS-PAGE, and Western blotting.** Proteins were extracted from the dissected ovaries of a pool of four emerging injected wasps using TRIzol reagent (Invitrogen) according to the manufacturer's instructions. This adult stage was chosen for Western blotting but also for electron microscopy experiments, as it corresponds to the stage where VLP production is maximal (20). Proteins forming the head and thorax of noninterfered wasps were also extracted from 4 adults following the same protocol. VLPs were purified from 100 dissected adult calices. Calices were first disrupted using a 1-mL syringe and a 25G ( $\times 5/8''$ ) needle in 20  $\mu$ L Dulbecco's phosphate-buffered saline (DPBS) 1 $\times$  containing 1/100 of protease inhibitor (Sigma, reference P2714). Disrupted calices were centrifuged for 10 min at 770 g at 4°C. The resulting supernatant containing VLPs and calix fluid was submitted to another centrifugation for 10 min, at 15,400  $\times$  g at 4°C allowing us to pellet the VLPs. Proteins were resuspended in 10  $\mu$ L of 0.1% SDS in water and quantified using the Qubit protein assay kit (Invitrogen) designed for the Qubit 2.0 fluorometer.

Before SDS-PAGE electrophoresis, 5  $\mu$ g of purified proteins were mixed with Laemmli buffer (225 mM Tris, 50% glycerol, 12% SDS, 0.5% bromophenol blue, 5%  $\beta$ -mercaptoethanol) and then heated 10 min for VLPs and 5 min for all other samples at 95°C. Denatured proteins were submitted to SDS-PAGE (stacking gel at 5% of acrylamide/bisacrylamide, running gel at 12.5% of acrylamide/bisacrylamide) in running buffer (192 mM glycine, 25 mM Tris, 20% SDS). Proteins were transferred onto a

Immun-Blot polyvinylidene difluoride (PVDF) membrane (Bio-Rad) previously equilibrated for 10 min with absolute ethanol and Towbin buffer (25 mM Tris, 192 mM glycine, 20% ethanol [vol/vol], pH 8.3) using the Trans-Blot Turbo blotting system (Bio-Rad) according to the Standard SD protocol (25 V, up to 1 A, for 30 min) with extra thick blot filter paper  $7 \times 8.4$  cm (Bio-Rad) equilibrated with Towbin buffer for 10 min as ion reservoir stacks. After the transfer, the PVDF membrane was incubated with blocking solution (TNT blocking buffer, 48 mM Tris, 140 mM NaCl, 0.05% Tween 20, pH 7.4; with the addition of 5% dehydrated fat-free milk) over night at 4°C. The PVDF membrane was then washed with TNT blocking buffer for 10 min. After the blocking step, the membrane was incubated with primary antibody against PIF-4 (purified polyclonal antibody from guinea pig, used at 1:2,000 dilution) mixed with blocking solution (containing 1% dehydrated milk) for 90 min at room temperature. PIF-4 polyclonal antibody was custom manufactured by Eurogentec using the peptide sequence Nterm-CPAGKRIPPEELIKIT-Cterm. The membrane was then washed three times with TNT blocking buffer for 10 min and incubated with secondary alkaline phosphatase (AP) conjugated rabbit anti-guinea pig antibodies (dilution = 1:10,000) mixed with blocking solution and 1% of dehydrated milk for 90 min at room temperature. After three new washing steps with TNT blocking buffer, immunodetection was performed using nitroblue tetrazolium/5-bromo-4-chloro-3-indolylphosphate (NBT/BCIP) substrate (Bio-Rad) mixed with alkaline phosphatase buffer (100 mM Tris-HCl, pH 9.0, 150 mM NaCl, 1 mM  $MgCl_2$ ) according to the manufacturer's protocol. The membranes were finally washed in  $H_2O$  to stop the immunodetection reaction and observed using the Gel Doc XR+ Molecular Imager system (Bio-Rad) with the Image Lab software.

**(v) Electron microscopy.** Injected wasps were dissected 1 day after adult emergence to collect their ovaries. A pair of ovaries from a noninjected wasp was also dissected at the same time to ensure observed phenotypes on VLP production and calyx formation were not due to the injection procedure. Four wasps of the same age were dissected per treatment in order to check the reproducibility of the observations.

Samples were fixed in mixture of 2% paraformaldehyde (Merck), 2% glutaraldehyde (Agar Scientific), and 0.1% sucrose in 0.1 M cacodylate buffer (pH 7.4) for 24 h, washed 3 times for 30 min in 0.1 M cacodylate buffer, and postfixed for 90 min with 2% osmium tetroxide (Electron Microscopy Science) in 0.1 M cacodylate buffer. After washing in 0.1 M cacodylate buffer for 20 min and 2 times for 20 min in distilled  $H_2O$ , samples were dehydrated in a graded series of ethanol solutions (50% ethanol 2 times for 15 min; 70% ethanol 2 times for 15 min and a third portion of 70% ethanol for 14 h; 90% ethanol 2 times for 20 min; and 100% ethanol 3 times for 20 min). Final dehydration was performed by 100% propylene oxide (PrOx) (VWR Int) 3 times for 20 min. Then, samples were incubated in PrOx/Epon epoxy resin (Fluka) mixture in a 2:1 ratio for 2 h with closed caps, in PrOx/Epon epoxy resin mixture in a 1:2 ratio for 2 h with closed caps and 90 min with open caps, and in 100% Epon for 16 h at room temperature. Samples were replaced in new 100% Epon and incubated at 37°C for 24 h and at 60°C for 48 h for polymerization.

Semithin cross sections (thickness, 0.8  $\mu m$ ) of ovaries were cut with a Leica Ultracut UCT ultramicrotome (Leica Microsystems GmbH), placed on glass, stained with toluidine blue (Electron Microscopy Science) and embedded in Epon resin (Fluka), which was allowed to polymerize for 48 h at 60°C. The sections were then observed with a Nikon Eclipse 80i microscope (Nikon) connected to DS-Vi1 camera driven by Nis-Element D 4.4 imaging software (Nikon).

Ultrathin sections (thickness, 70 nm) were cut with a Leica Ultracut UCT ultramicrotome (Leica Microsystems GmbH), placed on transmission electron microscope (TEM) one-slot grids (Agar Scientific) coated with Formvar film, and stained 20 min with 5% uranyl acetate (Electron Microscopy Science) and 5 min with Reynolds lead citrate. The sections were then observed at 100 kV with a JEM-1011 TEM (JEOL) connected to a CMOS Rio 9 digital camera driven by Digital Micrograph software (Gatan).

**Data availability.** Raw RNA-seq data are available in NCBI (BioProject accession number [PRJNA739064](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA739064)). All other data required to evaluate the findings are present in the paper and in the supplemental data.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**SUPPLEMENTAL FILE 1**, PDF file, 2.1 MB.

## ACKNOWLEDGMENTS

We thank Manon Lautrec, Francesco Iannielli, Ludovica De Vivo, and Giustina Filosi for their help with RNA interference experiments.

qPCR and TEM experiments have been, respectively, carried out with the technical support of the Genomic and IBiSA Electron Microscopy Facility of University of Tours and University Hospital of Tours. This work was supported by a grant from the Région Centre-Val de Loire to EH (Campovigne 2018-00124135) and by a grant from the Fédération de Recherche en Infectiologie de la Région Centre-Val de Loire to ML. ACA was supported by a Ph.D grant from the French Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation.

Authors thank reviewer comments which greatly helped to improve the manuscript.

## REFERENCES

- Katzourakis A, Gifford RJ. 2010. Endogenous viral elements in animal genomes. *PLoS Genet* 6:e1001191. <https://doi.org/10.1371/journal.pgen.1001191>.
- Holmes EC. 2011. The evolution of endogenous viral elements. *Cell Host Microbe* 10:368–377. <https://doi.org/10.1016/j.chom.2011.09.002>.
- Feschotte C, Gilbert C. 2012. Endogenous viruses: insights into viral evolution and impact on host biology. *Nat Rev Genet* 13:283–296. <https://doi.org/10.1038/nrg3199>.
- Tarlinton RE, Meers J, Young PR. 2006. Retroviral invasion of the koala genome. *Nature* 442:79–81. <https://doi.org/10.1038/nature04841>.
- Arbuckle JH, Medveczky MM, Luka J, Hadley SH, Luegmayr A, Ablashi D, Lund TC, Tolar J, De Meirleir K, Montoya JG, Komaroff AL, Ambros PF, Medveczky PG. 2010. The latent human herpesvirus-6A genome specifically integrates in telomeres of human chromosomes in vivo and in vitro. *Proc Natl Acad Sci U S A* 107:5563–5568. <https://doi.org/10.1073/pnas.0913586107>.
- Mangeney M, Renard M, Schlecht-Louf G, Bouallaga I, Heidmann O, Letzelter C, Richaud A, Ducos B, Heidmann T. 2007. Placental syncytins: genetic disjunction between the fusogenic and immunosuppressive activity of retroviral envelope proteins. *Proc Natl Acad Sci U S A* 104:20534–20539. <https://doi.org/10.1073/pnas.0707873105>.
- Armezzani A, Varela M, Spencer T, Palmarini M, Arnaud F. 2014. “Ménage à trois”: the evolutionary interplay between JSRV, enJSRVs and domestic sheep. *Viruses* 6:4926–4945. <https://doi.org/10.3390/v6124926>.
- Chuong EB, Elde NC, Feschotte C. 2016. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* 351:1083–1087. <https://doi.org/10.1126/science.125497>.
- Frank JA, Feschotte C. 2017. Co-option of endogenous viral sequences for host cell function. *Curr Opin Virol* 25:81–89. <https://doi.org/10.1016/j.coviro.2017.07.021>.
- Horie M, Tomonaga K. 2019. Paleovirology of bornaviruses: what can be learned from molecular fossils of bornaviruses. *Virus Res* 262:2–9. <https://doi.org/10.1016/j.virusres.2018.04.006>.
- Burke GR. 2019. Common themes in three independently derived endogenous nudivirus elements in parasitoid wasps. *Curr Opin Insect Sci* 32:28–35. <https://doi.org/10.1016/j.cois.2018.10.005>.
- Burke GR, Hines HM, Sharanowski BJ. 2021. The presence of ancient core genes reveals endogenization from diverse viral ancestors in parasitoid wasps. *Genome Biol Evol* 13:evab105. <https://doi.org/10.1093/gbe/evab105>.
- Bézier A, Herbinière J, Lanzrein B, Drezen J-M. 2009. Polydnavirus hidden face: the genes producing virus particles of parasitic wasps. *J Invertebr Pathol* 101:194–203. <https://doi.org/10.1016/j.jip.2009.04.006>.
- Bézier A, Annaheim M, Herbinière J, Wetterwald C, Gyapay G, Bernard-Samain S, Wincker P, Roditi I, Heller M, Belghazi M, Pfister-Wilhelm R, Periquet G, Dupuy C, Hugué E, Volkoff A-N, Lanzrein B, Drezen J-M. 2009. Polydnaviruses of braconid wasps derive from an ancestral nudivirus. *Science* 323:926–930. <https://doi.org/10.1126/science.1166788>.
- Volkoff A-N, Jouan V, Urbach S, Samain S, Bergoin M, Wincker P, Demetree E, Cousserans F, Provost B, Coulibaly F, Legeai F, Béliveau C, Cusson M, Gyapay G, Drezen J-M. 2010. Analysis of virion structural components reveals vestiges of the ancestral ichnovirus genome. *PLoS Pathog* 6:e1000923. <https://doi.org/10.1371/journal.ppat.1000923>.
- Strand MR, Burke GR. 2012. Polydnaviruses as symbionts and gene delivery systems. *PLoS Pathog* 8:e1002757. <https://doi.org/10.1371/journal.ppat.1002757>.
- Burke GR, Thomas SA, Eum JH, Strand MR. 2013. Mutualistic polydnaviruses share essential replication gene functions with pathogenic ancestors. *PLoS Pathog* 9:e1003348. <https://doi.org/10.1371/journal.ppat.1003348>.
- Strand MR, Burke GR. 2014. Polydnaviruses: nature’s genetic engineers. *Annu Rev Virol* 1:333–354. <https://doi.org/10.1146/annurev-virology-031413-085451>.
- Béliveau C, Cohen A, Stewart D, Periquet G, Djoumad A, Kuhn L, Stoltz D, Boyle B, Volkoff A-N, Herniou EA, Drezen J-M, Cusson M. 2015. Genomic and proteomic analyses indicate that banchine and campoplegine polydnaviruses have similar, if not identical, viral ancestors. *J Virol* 89:8909–8921. <https://doi.org/10.1128/JVI.01001-15>.
- Pichon A, Bézier A, Urbach S, Aury J-M, Jouan V, Ravallec M, Guy J, Cousserans F, Thézé J, Gauthier J, Demetree E, Schmieler S, Wurmsler F, Sibut V, Poirié M, Colinet D, da Silva C, Couloux A, Barbe V, Drezen J-M, Volkoff A-N. 2015. Recurrent DNA virus domestication leading to different parasite virulence strategies. *Sci Adv* 1:e1501150. <https://doi.org/10.1126/sciadv.1501150>.
- Drezen J-M, Leibold M, Bézier A, Hugué E, Volkoff A-N, Herniou EA. 2017. Endogenous viruses of parasitic wasps: variations on a common theme. *Curr Opin Virol* 25:41–48. <https://doi.org/10.1016/j.coviro.2017.07.002>.
- Volkoff A-N, Cusson M. 2020. The unconventional viruses of ichneumonid parasitoid wasps. *Viruses* 12:1170. <https://doi.org/10.3390/v12101170>.
- Sharanowski BJ, Ridenbaugh RD, Piekarski PK, Broad GR, Burke GR, Deans AR, Lemmon AR, Moriarty Lemmon EC, Diehl GJ, Whitfield JB, Hines HM. 2021. Phylogenomics of Ichneumonoidea (Hymenoptera) and implications for evolution of mode of parasitism and viral endogenization. *Mol Phylogenet Evol* 156:107023. <https://doi.org/10.1016/j.ympev.2020.107023>.
- Rodriguez JJ, Fernández-Triana JL, Smith MA, Janzen DH, Hallwachs W, Erwin TL, Whitfield JB. 2013. Extrapolations from field studies and known faunas converge on dramatically increased estimates of global microgastriine parasitoid wasp species richness (Hymenoptera: Braconidae). *Insect Conserv Divers* 6:530–536. <https://doi.org/10.1111/icad.12003>.
- Murphy N, Banks JC, Whitfield JB, Austin AD. 2008. Phylogeny of the parasitic microgastriid subfamilies (Hymenoptera: Braconidae) based on sequence data from seven genes, with an improved time estimate of the origin of the lineage. *Mol Phylogenet Evol* 47:378–395. <https://doi.org/10.1016/j.ympev.2008.01.022>.
- Thézé J, Bézier A, Periquet G, Drezen J-M, Herniou EA. 2011. Paleozoic origin of insect large dsDNA viruses. *Proc Natl Acad Sci U S A* 108:15931–15935. <https://doi.org/10.1073/pnas.1105580108>.
- Herniou EA, Hugué E, Thézé J, Bézier A, Periquet G, Drezen J-M. 2013. When parasitic wasps hijacked viruses: genomic and functional evolution of polydnaviruses. *Philos Trans R Soc Lond B Biol Sci* 368:20130051. <https://doi.org/10.1098/rstb.2013.0051>.
- Abd-Alla AMM, Cousserans F, Parker AG, Jehle JA, Parker NJ, Vlcek JM, Robinson AS, Bergoin M. 2008. Genome analysis of a *Glossina pallidipes* salivary gland hypertrophy virus reveals a novel, large, double stranded circular DNA virus. *J Virol* 82:4595–4611. <https://doi.org/10.1128/JVI.02588-07>.
- Wang Y, Bininda-Emonds ORP, Jehle JA. 2012. Nudivirus genomics and phylogeny, p 33–52. In Garcia ML, Romanowski V (ed), *Viral genomes - molecular structure, diversity, gene expression mechanisms and host-virus interactions*. InTech, London, UK.
- Jehle JA, Abd-Alla AMM, Wang Y. 2013. Phylogeny and evolution of *Hytrosaviridae*. *J Invertebr Pathol* 112:S62–S67. <https://doi.org/10.1016/j.jip.2012.07.015>.
- Wang Y, Kleespies RG, Huger AM, Jehle JA. 2007. The genome of *Gryllus bimaculatus* nudivirus indicates an ancient diversification of baculovirus-related nonoccluded nudiviruses of insects. *J Virol* 81:5395–5406. <https://doi.org/10.1128/JVI.02781-06>.
- Burand JP, Kim W, Afonso CL, Tulman ER, Kutish GF, Lu Z, Rock DL. 2012. Analysis of the genome of the sexually transmitted insect virus *Helicoverpa zea* nudivirus 2. *Viruses* 4:28–61. <https://doi.org/10.3390/v4010028>.
- Cheng R-L, Xi Y, Lou Y-H, Wang Z, Xu J-Y, Xu H-J, Zhang C-X. 2014. Brown planthopper nudivirus DNA integrated in its host genome. *J Virol* 88:5310–5318. <https://doi.org/10.1128/JVI.03166-13>.
- Bézier A, Thézé J, Gavory F, Gaillard J, Poulain J, Drezen J-M, Herniou EA. 2015. The genome of the nucleopolyhedrosis-causing virus from *Tipula oleracea* sheds new light on the Nudiviridae family. *J Virol* 89:3008–3025. <https://doi.org/10.1128/JVI.02884-14>.
- Cheng R-L, Li X-F, Zhang C-X. 2020. Nudivirus remnants in the genomes of arthropods. *Genome Biol Evol* 12:578–588. <https://doi.org/10.1093/gbe/evaa074>.
- Liu S, Coates BS, Bonning BC. 2020. Endogenous viral elements integrated into the genome of the soybean aphid, *Aphis glycines*. *Insect Biochem Mol Biol* 123:103405. <https://doi.org/10.1016/j.ibmb.2020.103405>.
- Liu S, Sappington TW, Coates BS, Bonning BC. 2021. Nudivirus sequences identified from the southern and western corn rootworms (Coleoptera: Chrysomelidae). *Viruses* 13:269. <https://doi.org/10.3390/v13020269>.
- Burke GR, Walden KKO, Whitfield JB, Robertson HM, Strand MR. 2014. Widespread genome reorganization of an obligate virus mutualist. *PLoS Genet* 10:e1004660. <https://doi.org/10.1371/journal.pgen.1004660>.
- Wyder S, Blank F, Lanzrein B. 2003. Fate of polydnavirus DNA of the egg-larval parasitoid *Chelonus inanitus* in the host *Spodoptera littoralis*. *J Insect Physiol* 49:491–500. [https://doi.org/10.1016/s0022-1910\(03\)00056-8](https://doi.org/10.1016/s0022-1910(03)00056-8).

40. Kroemer JA, Webb BA. 2004. Polydnavirus genes and genomes: emerging gene families and new insights into polydnavirus replication. *Annu Rev Entomol* 49:431–456. <https://doi.org/10.1146/annurev.ento.49.072103.120132>.
41. Espagne E, Dupuy C, Huguët E, Cattolico L, Provost B, Martins N, Poirié M, Periquet G, Drezen JM. 2004. Genome sequence of a polydnavirus: insights into symbiotic virus evolution. *Science* 306:286–289. <https://doi.org/10.1126/science.1103066>.
42. Thoctakittikul H, Beck MH, Strand MR. 2005. Inhibitor  $\kappa$ B-like proteins from a polydnavirus inhibit NF- $\kappa$ B activation and suppress the insect immune response. *Proc Natl Acad Sci U S A* 102:11426–11431. <https://doi.org/10.1073/pnas.0505240102>.
43. Bitra K, Suderman RJ, Strand MR. 2012. Polydnavirus Ank proteins bind NF- $\kappa$ B homodimers and inhibit processing of relish. *PLoS Pathog* 8:e1002722. <https://doi.org/10.1371/journal.ppat.1002722>.
44. Bitra K, Burke GR, Strand MR. 2016. Permissiveness of lepidopteran hosts is linked to differential expression of bracovirus genes. *Virology* 492:259–272. <https://doi.org/10.1016/j.virol.2016.02.023>.
45. Leobold M, Bézier A, Pichon A, Herniou EA, Volkoff A-N, Drezen J-M. 2018. The domestication of a large DNA virus by the wasp *Venturia canescens* involves targeted genome reduction through pseudogenization. *Genome Biol Evol* 10:1745–1764. <https://doi.org/10.1093/gbe/evy127>.
46. Burke GR, Simmonds TJ, Sharanowski BJ, Geib SM. 2018. Rapid viral symbiogenesis via changes in parasitoid wasp genome architecture. *Mol Biol Evol* 35:2463–2474. <https://doi.org/10.1093/molbev/msy148>.
47. Rotheram S. 1967. Immune surface of eggs of a parasitic insect. *Nature* 214:700. <https://doi.org/10.1038/214700a0>.
48. Reineke A, Asgari S, Schmidt O. 2006. Evolutionary origin of *Venturia canescens* virus-like particles. *Arch Insect Biochem Physiol* 61:123–133. <https://doi.org/10.1002/arch.20113>.
49. Salt G. 1965. Experimental studies in insect parasitism XIII. The haemolytic reaction of a caterpillar to eggs of its habitual parasite. *Proc R Soc Lond B* 162:303–318.
50. Rotheram S. 1973. The surface of the egg of a parasitic insect. I. The surface of the egg and first instar larva of *Nemeritis*. *Proc R Soc Lond B* 183:179–194.
51. Bedwin O. 1979. An insect glycoprotein: a study of the particles responsible for the resistance of a parasitoid's egg to the defence reactions of its insect host. *Proc R Soc Lond B* 205:271–286. <https://doi.org/10.1098/rspb.1979.0065>.
52. Feddersen I, Sander K, Schmidt O. 1986. Virus-like particles with host protein-like antigenic determinants protect an insect parasitoid from encapsulation. *Experientia* 42:1278–1281. <https://doi.org/10.1007/BF01946422>.
53. Bézier A, Harichaux G, Musset K, Labas V, Herniou EA. 2017. Qualitative proteomic analysis of *Tipula oleracea* nudivirus occlusion bodies. *J Gen Virol* 98:284–295. <https://doi.org/10.1099/jgv.0.000661>.
54. Herniou EA, Olszewski JA, Cory JS, O'Reilly DR. 2003. The genome sequence and evolution of baculoviruses. *Annu Rev Entomol* 48:211–234. <https://doi.org/10.1146/annurev.ento.48.091801.112756>.
55. Burke GR, Walden KKO, Whitfield JB, Robertson HM, Strand MR. 2018. Whole genome sequence of the parasitoid wasp *Microplitis demolitor* that harbors an endogenous virus mutualist. *G3 (Bethesda)* 8:2875–2880. <https://doi.org/10.1534/g3.118.200308>.
56. Gauthier J, Boulain H, van Vugt JFFA, Baudry L, Persyn E, Aury J-M, Noel B, Bretaudeau A, Legeai F, Warris S, Chebbi MA, Dubreuil G, Duvic B, Kremer N, Gayral P, Musset K, Josse T, Bigot D, Bressac C, Moreau S, Periquet G, Harry M, Montagné N, Boulogne I, Sabeti-Azad M, Maïbèche M, Chertemps T, Hilliou F, Siauxsat D, Amselem J, Luyten J, Capdevielle-Dulac C, Labadie K, Merlin BL, Barbe V, de Boer JG, Marbouty M, Cónsoli FL, Dupas S, Hua-Van A, Le Goff G, Bézier A, Jacquin-Joly E, Whitfield JB, Vet LEM, Smid HM, Kaiser L, Koszul R, Huguët E, Herniou EA, et al. 2021. Chromosomal scale assembly of parasitic wasp genome reveals symbiotic virus colonization. *Commun Biol* 4:104. <https://doi.org/10.1038/s42003-020-01623-8>.
57. Blissard GW, Rohrmann GF. 1990. Baculovirus diversity and molecular biology. *Annu Rev Entomol* 35:127–155. <https://doi.org/10.1146/annurev.ento.35.010190.001015>.
58. Ikeda M, Hamajima R, Kobayashi M. 2015. Baculoviruses: diversity, evolution and manipulation of insects. *Entomol Sci* 18:1–20. <https://doi.org/10.1111/ens.12105>.
59. Rohrmann GF. 2019. Baculovirus molecular biology, 4th ed. National Center for Biotechnology Information, Bethesda, MD.
60. Burke GR, Strand MR. 2012. Deep sequencing identifies viral and wasp genes with potential roles in replication of *Microplitis demolitor* bracovirus. *J Virol* 86:3293–3306. <https://doi.org/10.1128/JVI.06434-11>.
61. Reineke A, Asgari S, Ma G, Beck M, Schmidt O. 2002. Sequence analysis and expression of a virus-like particle protein, VLP2, from the parasitic wasp *Venturia canescens*. *Insect Mol Biol* 11:233–239. <https://doi.org/10.1046/j.1365-2583.2002.00330.x>.
62. Dorémus T, Urbach S, Jouan V, Cousserans F, Ravallec M, Demettré E, Wajnberg E, Poulain J, Azéma-Dossat C, Darboux I, Escoubas J-M, Colinet D, Gatti J-L, Poirié M, Volkoff A-N. 2013. Venom gland extract is not required for successful parasitism in the polydnavirus-associated endoparasitoid *Hyposoter didymator* (Hym. Ichneumonidae) despite the presence of numerous novel and conserved venom proteins. *Insect Biochem Mol Biol* 43:292–307. <https://doi.org/10.1016/j.ibmb.2012.12.010>.
63. Summers KM, Howells AJ, Pyliotis NA. 1982. Biology of eye pigmentation in insects, p 119–166. *In* Berridge M, Treherne JE, Wigglesworth V (ed), *Advances in insect physiology*. Elsevier Academic Press, San Diego, CA.
64. Werren JH, Loehlin DW, Giebel JD. 2009. Larval RNAi in *Nasonia* (parasitoid wasp). *Cold Spring Harb Protoc* 2009:pdb.prot5311. <https://doi.org/10.1101/pdb.prot5311>.
65. Colinet D, Anselme C, Deleury E, Mancini D, Poulain J, Azéma-Dossat C, Belghazi M, Tares S, Pennacchio F, Poirié M, Gatti J-L. 2014. Identification of the main venom protein components of *Aphidius ervi*, a parasitoid wasp of the aphid model *Acyrtosiphon pisum*. *BMC Genomics* 15:342. <https://doi.org/10.1186/1471-2164-15-342>.
66. King RC, Storto PD. 1988. The role of the *otu* gene in *Drosophila* oogenesis. *Bioessays* 8:18–24. <https://doi.org/10.1002/bies.950080106>.
67. Cheng C-H, Liu S-M, Chow T-Y, Hsiao Y-Y, Wang D-P, Huang J-J, Chen H-H. 2002. Analysis of the complete genome sequence of the Hz-1 virus suggests that it is related to members of the *Baculoviridae*. *J Virol* 76:9024–9034. <https://doi.org/10.1128/jvi.76.18.9024-9034.2002>.
68. Passarelli A, Guarino L. 2007. Baculovirus late and very late gene regulation. *Curr Drug Targets* 8:1103–1115. <https://doi.org/10.2174/138945007782151324>.
69. Chen Y-R, 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>.
70. Rohrmann GF. 2014. Baculovirus nucleocapsid aggregation (MNPV vs SNPV): an evolutionary strategy, or a product of replication conditions? *Virus Genes* 49:351–357. <https://doi.org/10.1007/s11262-014-1113-5>.
71. Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, van der Lee R, Bessy A, Chêneby J, Kulkarni SR, Tan G, Baranasic D, Arenillas DJ, Sandelin A, Vandepoel K, Lenhard B, Ballester B, Wasserman WW, Parcy F, Mathelier A. 2018. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res* 46:D260–D266. <https://doi.org/10.1093/nar/gkx1126>.
72. von Kalm L, Crossgrove K, Von Seggern D, Guild GM, Beckendorf SK. 1994. The broad-complex directly controls a tissue-specific response to the steroid hormone ecdysone at the onset of *Drosophila* metamorphosis. *EMBO J* 13:3505–3516. <https://doi.org/10.1002/j.1460-2075.1994.tb06657.x>.
73. Harwood SH, Li L, Ho PS, Preston ÅK, Rohrmann GF. 1998. AcMNPV late expression factor-5 interacts with itself and contains a zinc ribbon domain that is required for maximal late transcription activity and is homologous to elongation factor TFIIS. *Virology* 250:118–134. <https://doi.org/10.1006/viro.1998.9334>.
74. Passarelli AL, Todd JW, Miller LK. 1994. A baculovirus gene involved in late gene expression predicts a large polypeptide with a conserved motif of RNA polymerases. *J Virol* 68:4673–4678. <https://doi.org/10.1128/JVI.68.7.4673-4678.1994>.
75. Titterton JS, Nun TK, Passarelli AL. 2003. Functional dissection of the baculovirus late expression *factor-8* gene: sequence requirements for late gene promoter activation. *J Gen Virol* 84:1817–1826. <https://doi.org/10.1099/vir.0.19083-0>.
76. Zhou F, Kuang W, Wang X, Hou D, Huang H, Sun X, Deng F, Wang H, van Oers MM, Wang M, Hu Z. 2019. The cysteine-rich region of a baculovirus VP91 protein contributes to the morphogenesis of occlusion bodies. *Virology* 535:144–153. <https://doi.org/10.1016/j.virol.2019.06.016>.
77. Moldován N, Tombácz D, Szűcs A, Csabai Z, Balázs Z, Kis E, Molnár J, Boldogkői Z. 2018. Third-generation sequencing reveals extensive polyclonism and transcriptional overlapping in a baculovirus. *Sci Rep* 8:8604. <https://doi.org/10.1038/s41598-018-26955-8>.
78. Guay PS, Guild GM. 1991. The ecdysone-induced puffing cascade in *Drosophila* salivary glands: a *Broad-Complex* early gene regulates intermolt and late gene transcription. *Genetics* 129:169–175. <https://doi.org/10.1093/genetics/129.1.169>.
79. Karim FD, Guild GM, Thummel CS. 1993. The *Drosophila Broad-Complex* plays a key role in controlling ecdysone-regulated gene expression at

- the onset of metamorphosis. *Development* 118:977–988. <https://doi.org/10.1242/dev.1183.977>.
80. Sempere LF, Sokol NS, Dubrovsky EB, Berger EM, Ambros V. 2003. Temporal regulation of microRNA expression in *Drosophila melanogaster* mediated by hormonal signals and *Broad-Complex* gene activity. *Dev Biol* 259:9–18. [https://doi.org/10.1016/s0012-1606\(03\)00208-2](https://doi.org/10.1016/s0012-1606(03)00208-2).
  81. Guarino LA, Dong W, Jin J. 2002. In vitro activity of the baculovirus late expression factor LEF-5. *J Virol* 76:12663–12675. <https://doi.org/10.1128/jvi.76.24.12663-12675.2002>.
  82. Gross CH, Shuman S. 1998. Characterization of a baculovirus-encoded RNA 5'-triphosphatase. *J Virol* 72:7057–7063. <https://doi.org/10.1128/JVI.72.9.7057-7063.1998>.
  83. Li Y, Guarino LA. 2008. Roles of LEF-4 and PTP/BVP RNA triphosphatases in processing of baculovirus late mRNAs. *J Virol* 82:5573–5583. <https://doi.org/10.1128/JVI.00058-08>.
  84. Ogino T, Banerjee AK. 2008. Formation of guanosine(5')tetraphospho(5') adenosine cap structure by an unconventional mRNA capping enzyme of vesicular stomatitis virus. *J Virol* 82:7729–7734. <https://doi.org/10.1128/JVI.00326-08>.
  85. Lorenzi A, Ravallec M, Eychenne M, Jouan V, Robin S, Darboux I, Legeai F, Gosselin-Grenet A-S, Sicard M, Stoltz D, Volkoff A-N. 2019. RNA interference identifies domesticated viral genes involved in assembly and trafficking of virus-derived particles in ichneumonid wasps. *PLoS Pathog* 15: e1008210. <https://doi.org/10.1371/journal.ppat.1008210>.
  86. Lorenzi A, Strand MR, Burke GR, Volkoff A-N. 2022. Identifying bracovirus and ichnovirus genes involved in virion morphogenesis. *Curr Opin Insect Sci* 49:63–70. <https://doi.org/10.1016/j.cois.2021.11.006>.
  87. Hellers M, Beck M, Theopold U, Kamei M, Schmidt O. 1996. Multiple alleles encoding a virus-like particle protein in the ichneumonid endoparasitoid *Venturia canescens*. *Insect Mol Biol* 5:239–249. <https://doi.org/10.1111/j.1365-2583.1996.tb00098.x>.
  88. Asgari S, Reineke A, Beck M, Schmidt O. 2002. Isolation and characterization of a neprilysin-like protein from *Venturia canescens* virus-like particles. *Insect Mol Biol* 11:477–485. <https://doi.org/10.1046/j.1365-2583.2002.00356.x>.
  89. Schmidt O, Li D, Beck M, Kinuthia W, Bellati J, Roberts HLS. 2005. Phenoloxidase-like activities and the function of virus-like particles in ovaries of the parthenogenetic parasitoid *Venturia canescens*. *J Insect Physiol* 51: 117–125. <https://doi.org/10.1016/j.jinsphys.2004.05.006>.
  90. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 37:907–915. <https://doi.org/10.1038/s41587-019-0201-4>.
  91. Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30:923–930. <https://doi.org/10.1093/bioinformatics/btt656>.
  92. Liao Y, Smyth GK, Shi W. 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res* 47:e47. <https://doi.org/10.1093/nar/gkz114>.
  93. Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 11:R25. <https://doi.org/10.1186/gb-2010-11-3-r25>.
  94. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140. <https://doi.org/10.1093/bioinformatics/btp616>.
  95. Chen Y, Lun ATL, Smyth GK. 2016. From reads to genes to pathways: differential expression analysis of RNA-seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Res* 5:1438. <https://doi.org/10.12688/f1000research.8987.1>.
  96. Abdi H. 2007. Z-scores, p 1055–1058. In Salkind N (ed), *Encyclopedia of measurement and statistics*. SAGE Publications, Newbury Park, CA.
  97. Bailey TL, Elkan C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* 2:28–36.
  98. McLeay RC, Bailey TL. 2010. Motif enrichment analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics* 11:165. <https://doi.org/10.1186/1471-2105-11-165>.
  99. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble W. 2007. Quantifying similarity between motifs. *Genome Biol* 8:R24. <https://doi.org/10.1186/gb-2007-8-2-r24>.
  100. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391:806–811. <https://doi.org/10.1038/35888>.