# Genome-wide identification of lncRNAs associated with viral infection in Spodoptera frugiperda

Stéphanie Robin, Fabrice Legeai, Véronique Jouan, Mylène Ogliastro, Isabelle Darboux

JOURNAL OF GENERAL VIROLOGY

RESEARCH ARTICLE

Robin *et al.*, *Journal of General Virology* 2023;104:001827

DOI 10.1099/jgv.0.001827

MICROBIOLOGY SOCIETY

OPEN ACCESS

# Genome-wide identification of lncRNAs associated with viral infection in *Spodoptera frugiperda*

Stéphanie Robin[1,2], Fabrice Legeai[1,2], Véronique Jouan[3], Mylène Ogliastro[3] and Isabelle Darboux[3,*]

**Abstract**

The role of lncRNAs in immune defence has been demonstrated in many multicellular and unicellular organisms. However, investigation of the identification and characterization of long non-coding RNAs (lncRNAs) involved in the insect immune response is still limited. In this study, we used RNA sequencing (RNA-seq) to investigate the expression profiles of lncRNAs and mRNAs in the fall armyworm *Spodoptera frugiperda* in response to virus infection. To assess the tissue- and virus-specificity of lncRNAs, we analysed and compared their expression profiles in haemocytes and fat body of larvae infected with two entomopathogenic viruses with different lifestyles, i.e. the polydnavirus HdIV (*Hyposoter didymator* IchnoVirus) and the densovirus JcDV (*Junonia coenia* densovirus). We identified 1883 candidate lncRNAs, of which 529 showed differential expression following viral infection. Expression profiles differed considerably between samples, indicating that many differentially expressed (DE) lncRNAs showed virus- and tissue-specific expression patterns. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment and target prediction analyses indicated that DE-LncRNAs were mainly enriched in metabolic process, DNA replication and repair, immune response, metabolism of insect hormone and cell adhesion. In addition, we identified three DE-lncRNAs potentially acting as microRNA host genes, suggesting that they participate in gene regulation by producing miRNAs in response to virus infection. This study provides a catalogue of lncRNAs expressed in two important immune tissues and potential insight into their roles in the antiviral defence in *S. frugiperda*. The results may help future in-depth functional studies to better understand the biological function of lncRNAs in interaction between viruses and the fall armyworm.

## DATA SUMMARY

The supplementary tables are deposited on Figshare at https://doi.org/10.6084/m9.figshare.21739289.v1 [1].

## INTRODUCTION

In recent years, lncRNAs have attracted considerable interest for their essential role in the modulation and control of virus–host interactions. Although the biological function of lncRNAs remains largely unknown, some studies have elucidated the molecular mechanisms that enable lncRNAs to control viral infections in vertebrates. For example, the lncRNA NEAT1 was demonstrated to exert antiviral effects on the Hantaan virus (HTNV) by enhancing HTNV-induced interferon responses in humans [2]. In some cases, however, viruses can hijack the activity of the host lncRNAs to neutralize antiviral responses and promote viral replication and pathogenesis. This is the case for influenza A virus (IAV), which specifically exploits the human lncRNA IPAN to promote its replication [3]. IAV infection stimulates the expression of IPAN, which then associates with the viral RNA polymerase PB1. The IPAN : PB1 complex protects the viral protein from degradation by host immunity, thereby promoting efficient viral RNA synthesis. In certain cases, unrelated viruses exploit the same lncRNA for optimizing their life cycle. For example, the lncRNA ACOD1 is induced by vesicular stomatitis virus (VSV), DNA virus herpes simplex virus type 1 (HSV-1) and vaccinia virus

Received 14 October 2022; Accepted 29 December 2022; Published 09 February 2023

**Author affiliations:** [1]BIPAA, IGEPP, INRAE, Institut Agro, University of Rennes, Rennes, France; [2]University of Rennes, INRIA, CNRS, IRISA, Rennes, France; [3]INRAE, University of Montpellier, UMR Diversité, Génomes & Interactions Microorganismes-Insectes (DGIMI), Montpellier, France.

***Correspondence:** Isabelle Darboux, isabelle.darboux@inrae.fr

**Keywords:** densovirus; host-virus interactions; long non-coding RNA; polydnavirus; RNA-Seq; *Spodoptera frugiperda*.

The NCBI accession numbers for the sequencing data are SAMN29883271, SAMN29883272, SAMN29883273, SAMN29883274, SAMN29883275, SAMN29883276, SAMN29883277, SAMN29883278, SAMN29883279, SAMN29883280, SAMN29883281, SAMN29883282, SAMN29883283, SAMN29883284, SAMN29883285, SAMN29883286, SAMN29883287 and SAMN29883288.

Ten supplementary tables are available with the online version of this article.

(VACV). Binding of lncRNA-ACOD1 to GOT2, an essential enzyme involved in metabolism, stimulates GOT2 activity, which promotes viral replication. Depletion of lncRNA-ACOD1 led to attenuated replication of all the three viruses in macrophages [4]. For their part, some viruses are also capable of generating lncRNAs to promote viral infection [5, 6]. There are only a limited number of examples, but functional analyses show that these viral lncRNAs act at different steps of the virus life cycle. For example, the noncoding subgenomic flavivirus RNA (sfRNA), produced by West Nile virus and dengue virus, inhibits the antiviral response by neutralizing interferon (IFN) signalling [7]. The human cytomegalovirus (HCMV) produces the nuclear lncRNA RNA4.9, which is important for the initiation of viral DNA replication [8].

From a structural point of view, lncRNAs have more characteristics in common with protein-coding genes than with other non-coding RNAs. Indeed, the majority of lncRNAs are transcribed by RNA polymerase II, although there are a few exceptions that are transcribed by RNA polymerase III. LncRNAs are also spliced, polyadenylated and 5′-end capped. However, they differ from mRNAs in various aspects through their size, specificity, organization and subcellular location. LncRNAs are classically defined as cellular RNAs >200 nucleotides in length. Their size can be very heterogeneous, but their length is on average shorter than that of mRNAs. They are considered to have no translational capacity, with the exception of a few recently described examples that encode for micropeptides [9]. Transcripts of lncRNAs generally contain fewer exons than mRNAs, and fewer isoforms per gene locus. They are generally less expressed than mRNAs, but more specific to the cell type, tissue or developmental stage. For exemple, Cabili *et al.* showed that a high proportion (78%) of the 8195 human intergenic lncRNAs identified from 24 tissues and cell types are tissue-specific, compared with only 19% of protein-coding genes [10]. Another notable characteristic of lncRNAs is that they usually display poor evolutionary conservation among phylogenetically related species. It is therefore challenging to predict the function of lncRNAs from their sequence.

LncRNAs have been identified in the genome of a large number of pluricellular and unicellular species. In eukaryotes, thousands of genes coding for lncRNAs have already been identified and referenced in the NONCODE database (v6.0, http://www.noncode.org) [11, 12]. High-throughput sequencing analyses have revealed the expression of thousands of lncRNAs genes in diverse insect taxa. For example, in *Drosophila melanogaster*, the NONCODE database has listed 15 543 lncRNA genes producing 42 848 lncRNA transcripts. Genome-wide identifications have shown that insect lncRNAs are involved in various important biological processes, including metamorphosis, embryogenesis, neurogenesis and acquisition of blood meal in Diptera [13–16]. LncRNAs may also be involved in more species-specific biological processes, such as the transition from nurses to foragers for the hymenoptera *Apis mellifera*, fecundity and virulence adaptation to resistant rice varieties in the Hemiptera *Nilaparvate lugens*, plant–aphid interactions, or response to thermal stress in the cold-hardy *Diamesa tonsa* [17–20]. In Lepidoptera, lncRNAs may be involved in insect outbreaks, metamorphosis, insecticide resistance and regulation of hydroxyecdysone-induced autophagy [21–25]. The possible role of lncRNAs in immune response has been addressed by analysing the regulation of lncRNAs expression in Coleoptera, Homoptera and Hymenoptera insects following microsporidial infection [26] and viral infection [21, 27]. LncRNAs have also been proposed to be involved in the regulation of innate immune memory in *Tribolium castaneum* [28]. The precise role of insect lncRNAs during viral infection remains largely unknown. In vectors of arboviruses, functional studies have shown that silencing the expression of some lncRNAs could promote viral replication [29, 30]. Recently, Zhang *et al.* identified VINR, a *D. melanogaster* nuclear lncRNA that interacts with the viral suppressor of *Drosophila* C virus RNAi, thereby antagonizing its activity [31]. In Lepidoptera, a study showed that stimulation of the antiviral response by injecting the viral pathogen-associated molecular pattern double-stranded RNA into larvae altered the expression of lncRNAs, suggesting that the latter play a role in the immune response [32].

The objective of our study was to broaden the current knowledge of lncRNAs in the lepidopteran antiviral response by assessing the tissue- and virus-specificity of lncRNAs expression using RNA sequencing (RNA-seq) analysis. To this end, we explored interactions between the fall armyworm *Spodoptera frugiperda* and two viruses, namely the polydnavirus HdIV (*Hyposoter didymator* IchnoVirus) and the densovirus JcDV (*Junonia coenia* densovirus), which differ considerably in their life cycle. HdIV is an enveloped virus. It is an obligate symbiont of the endoparasitoid hymenopteran *H. didymator*. The virus replication occurs exclusively in the female reproductive tract. The packaged genome, 306.9 kb in size, consists of 54 segments of circular double-stranded DNA molecules and has 152 ORFs encoding virulence genes expressed during parasitism [33, 34]. The wasp female co-injects the virus particles with the parasitoid egg into the haemolymph of lepidopteran larvae at the time of parasitism. As the packaged genome does not include any replication-associated genes, HdIV is unable to replicate in lepidopteran host cells. However, it infects a large number of cells and tissues very rapidly, including haemocytes and fat body. The viral proteins neutralize the host immune response, which promotes parasitoid survival. In particular, we previously demonstrated that HdIV infection triggers cell death by apoptosis in the fat body and haemocytes of *S. frugiperda* larvae, allowing elimination of important host immune tissues [35]. The growth and development rate of infected larvae is also disrupted until development of the parasite is complete [36–38]. Of note, intrahaemocoelic injection of HdIV into *S. frugiperda* larvae induces the same symptoms as those observed in parasitized larvae, in terms of immunosuppression and development arrest. JcDV is a small, non-enveloped DNA virus characterized by a single-stranded linear genome of 6 kb, consisting of four ORFs coding for seven proteins. Lepidopteran larvae are infected by ingesting contaminated food. After crossing the intestinal barriers, JcDV infects several other cells and tissues, including haemocytes and fat body. Densonucleosis, due to amplification

of the viral genome, was clearly observed in haemocytes 72 h after infection [39]. In contrast, the fat body shows autophagic and apoptotic-like bodies, suggesting that the mode of action of JcDV is different in these two immune cell compartments. Death of the JcDV-infected *S. frugiperda* larvae, which occurs approximately 5–10 days after initial infection, depending on the initial dose of virus ingested, appears to be due to moult blockage, oxidative stress and anoxia from tracheal obstruction [39]. Direct intrahaemocoelic injection of JcDV is also lethal to caterpillars [39]. The host insect chosen for our study is *S. frugiperda*, a serious pest for many agricultural crops worldwide, with a fully sequenced genome and transcriptome [40, 41]. To identify virus-responsive lncRNAs, we compared the expression profiles of lncRNAs and mRNAs in the major insect immune cells, haemocytes and fat body, infected with HdIV or JcDV. Gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment and co-expression analyses were then performed to predict the function of lncRNAs in response to HdIV or JcDV infection.

Identification of lncRNAs differentially expressed in response to viral infection will provide a molecular biological basis for a better understanding of the roles of lncRNAs in *Spodoptera*–virus interaction. Further, the results will yield new insights into the mechanisms of interactions between insects and the viruses HdIV and JcDV.

## METHODS

### Insect rearing and virus preparation

*S. frugiperda* larvae (corn strain) were maintained in the DGIMI laboratory for many years and reared on the Poitout artificial diet [42] at 25 °C under a photoperiod of 16 h light and 8 h dark. *H. didymator* parasitoids, which carry the polydnavirus symbiont HdIV, were reared on *S. frugiperda* at 26 °C with a 16 h light and 8 h dark photoperiod. HdIV was isolated from *H. didymator* ovaries as described previously [38]. Briefly, after dissection, ovaries were homogenized through a 23 gauge needle. The cell suspension, which contains the viral particles, was then filtered through a 0.45 mm pore size filter. HdIV concentration was classically expressed as wasp-equivalent (weq). One weq is defined as the amount of HdIV collected from the ovary of a single adult female. HdIV preparation was used within 24 h. The JcDV stock was prepared as reported in [39]. In brief, *S. frugiperda* larvae were infected with JcDV by oral ingestion. After 7–10 days, JcDV was extracted from dead larvae, purified on salt density gradient and then dialysed. JcDV concentration was estimated by quantitative PCR and expressed as viral genome (vg) $ml^{-1}$. The virus stock was maintained at −20 °C until use.

### Virus injection, sample collection and RNA isolation

HdIV and JcDV were injected intrahaemocoelically into fifth-instar *S. frugiperda* larvae (1 day old). Nine microlitres of a viral suspension containing 7.2e10 vg of JcDV and 0.5 weq of HdIV were injected per larva. Mock-infected larvae were injected with an equivalent volume of phosphate-buffered saline (PBS). With the dose of JcDV used in our experimental conditions, mortality was observed in some, but not all, JcDV-infected larvae (20–36 % in the three replicates). To avoid RNA degradation, only JcDV-infected larvae showing low signs of pathogenesis were used for analysis. None of the uninfected (PBS-injected) or HdIV-infected larvae died, as expected. The developmental arrest and growth reduction observed in HdIV-infected larvae indicated that HdIV infection was established [37]. After collection of haemolymph, haemocytes were harvested by centrifugation at 800 ***g*** for 5 min and washed in PBS. The cell pellet was then resuspended in the TRK lysis buffer of the EZNA Total RNA kit I (Omega). Fat body was dissected out under a stereomicroscope, washed in PBS and transferred directly into the TRK lysis buffer. Samples were stored at −80 °C until use. Total RNA was extracted using the EZNA Total RNA kit I according to the manufacturer's instructions (Omega). RNA samples were treated with a TURBO DNA-free kit (Ambion) to eliminate potential genomic DNA contamination. RNA concentration and integrity were assessed using a NanoDrop ND-1000 spectrophotometer and an Agilent 2100 Bioanalyzer, respectively. Total RNA with RNA integrity (RIN) scores ≥9 were retained for RNA sequencing. For each replicate, total RNA was extracted from 25–28 control larvae, 34–79 HdIV-infected larvae and 50–58 JcDV-infected larvae.

### Illumina sequencing

Library preparation and RNA sequencing were conducted by the Montpellier GenomiX (MGX) sequencing facility (CNRS, Montpellier, France). Briefly, polyA mRNA was isolated with oligo(dT) beads. After fragmentation, the first cDNA strand was synthesized using random primers and Superscript Reverse Transcriptase II (Invitrogen). The cDNA was then converted into double-stranded cDNA. Eighteen cDNA libraries were generated using the Illumina TruSeq Stranded mRNA sample and sequenced on an Illumina HiSeq 2500 platform, which generated 125 bp paired-end reads. Libraries were validated using the Fragment Analyzer system with the Standard Sensitivity NGS kit. The sequencing data has been deposited in the National Center for Biotechnology Information (NCBI) database under accession nos SAMN29883271, SAMN29883272, SAMN29883273, SAMN29883274, SAMN29883275, SAMN29883276, SAMN29883277, SAMN29883278, SAMN29883279, SAMN29883280, SAMN29883281, SAMN29883282, SAMN29883283, SAMN29883284, SAMN29883285, SAMN29883286, SAMN29883287 and SAMN29883288.

### Mapping, transcriptome assembly and lncRNA identification

The sequencing data were analysed using the bioinformatics pipeline nf-core/rnaseq v3.0 (https://nf-co.re/rnaseq/3.0) [43]. Briefly, after removing low-quality reads and adapter sequences with Trim Galore! version 0.6.6, paired-end clean reads were obtained and aligned to the reference genome sfC.ver6.fa (https://doi.org/10.1038/s42003-020-01382-6) concatenated with HdIV (doi: 10.1186 /s12915-020-00822-3) and JcDV (KC883978.1) sequences, using STAR software version 2.6.1d. The alignment results were assembled using StringTie version 2.1.4 with stringtie_ignore_gtf parameter. Identification of candidate lncRNAs from the assembled transcriptome was based on applied filters: exon number ≥2, transcript length >200 bp, using the FEELnc pipeline version 0.2 [44]. Prediction of the protein-coding potential was also evaluated using FEELnc. A new annotation based on sfC. ver6.fa with HdIV, JcDV, lncRNA and mRNA genes identified by FEELnc was obtained and will be used for the next analysis.

### Conservation analysis of lncRNAs

The sequence conservation of *S. frugiperda* lncRNAs was examined using BLASTN against the genomes of 12 lepidopteran species with e-value cut-off 1E-10. The best hits, covering at least 70 % of the lncRNA transcript, were retained. Genome datasets of the Lepidopteran species, including *Amyelois transitella* (*Amyelois_transitella_v1_-_scaffolds.fa / GCA_001186105.1*), *Bombyx mori* (*ASM15162v1*), *Chilo suppressalis* (*GCA_000636095.1*), *Danaus plexippus* (*GCF_009731565.1*), *Manduca sexta* (*GCA_000262585.1*), *Helicoverpa armigera* (*GCA_002156985.1*), *Plutella xylostella* (*GCF_000330985.1*), *Heliconius melpomene* (*ASM31383v2*), *Operophtera brumata* (GCA_001266575.1), *Plodia interpunctella* (*ASM2298509v1*), *S. exigua* (GCA_011316535.1/ GCA_015679615.1) and *S. litura* (*GCA_002706865.3*), were retrieved from the lepbase database (http://download.lepbase.org/v4/sequence/) and for the most part also the present NCBI genome database.

### Identification of differentially expressed lncRNAs

FeatureCounts version 2.0.1 [45] was used with default parameters except -C -p -M -s 2 --fraction to count mapped reads for genes for each sample. The AskoR library (https://github.com/askomics/askoR), which includes the EdgeR package [46], was then used for the differential expression analysis with the following steps. First, raw counts were converted to CPM values, and genes were filtered on expression level; only genes with CPM >0.5 in at least three libraries were conserved. Then, normalization was done using a TMM (trimmed mean of M-values) method. Differentially expressed transcripts with a $|\log_2$ fold change $\geq|$ and adjusted *P*-value (false discovery rate FDR) <0.05 were assigned at differentially expressed and virus-responsive (HdIV or JcDV) compared to control (PBS).

### Gene functional annotation and enrichment analysis

GO term enrichment analysis was performed with AskoR library using topGO package [47] with default parameters defined in askoR. EggNog-mapper (version: 2.1.6) [48] was used with eggnog-mapper database (version 5.0.2.) for the KEGG pathway annotation. KEGG pathway enrichment analysis was performed using the Fisher test. For each KEGG pathway, we compared the genes total number in KEGG pathway/the genes total number not in KEGG pathway with the DE genes number in KEGG pathway/the DE-genes number not in KEGG pathway.

### Gene clustering for co-expression analysis

Gene clustering for co-expression analysis was performed using the ClustAndGO function based on coseq package [49] included in AskoR library with default parameters defined in askoR excepted coseq_HeatmapOrderSample=True.

### LncRNA target prediction

The biological function targeted by DE-lncRNAs was predicted by identifying neighbouring protein-encoding genes located 1 Mb upstream or downstream of each lncRNA using the last step of the FEELnc pipeline, FEELnc_classifier [50]. Interactions between mRNA and lncRNA were identified by a Spearman correlation analysis, described by Le Béguec *et al.* [51]. For exonic antisense and intergenic antisense lncRNAs, a Spearman correlation coefficient was calculated for each lncRNA–mRNA pair.

### Identification of lncRNAs potentially acting as miRNA host gene

A BLASTN version 2.9.0 of miRNA precursors sequences [52] was done against *S. frugiperda* genome v6.0 with evalue=1e-10 and max_target_seqs=20. Intersections between BLASTN hits of *S. frugiperda* sequences and lncRNA sequences were analysed using bedtools version 2.27.1 [53] and corresponded to lncRNA potential precursor miRNA.

### Real time RT-PCR analysis

A set of lncRNAs were randomly selected for validation using RT-qPCR. DNase-treated RNAs were the same as those used for library preparation. Reverse transcription was performed with oligo(dT) primer and GoScriptTM Reverse Transcriptase following the manufacturer's instructions (Promega). Gene expression analysis was carried out using a LightCycler 480 Instrument II (Roche) and GoTaq qPCR Master Mix (Promega) in a final volume of 3 µl. The programme consisted of an initial hot-start

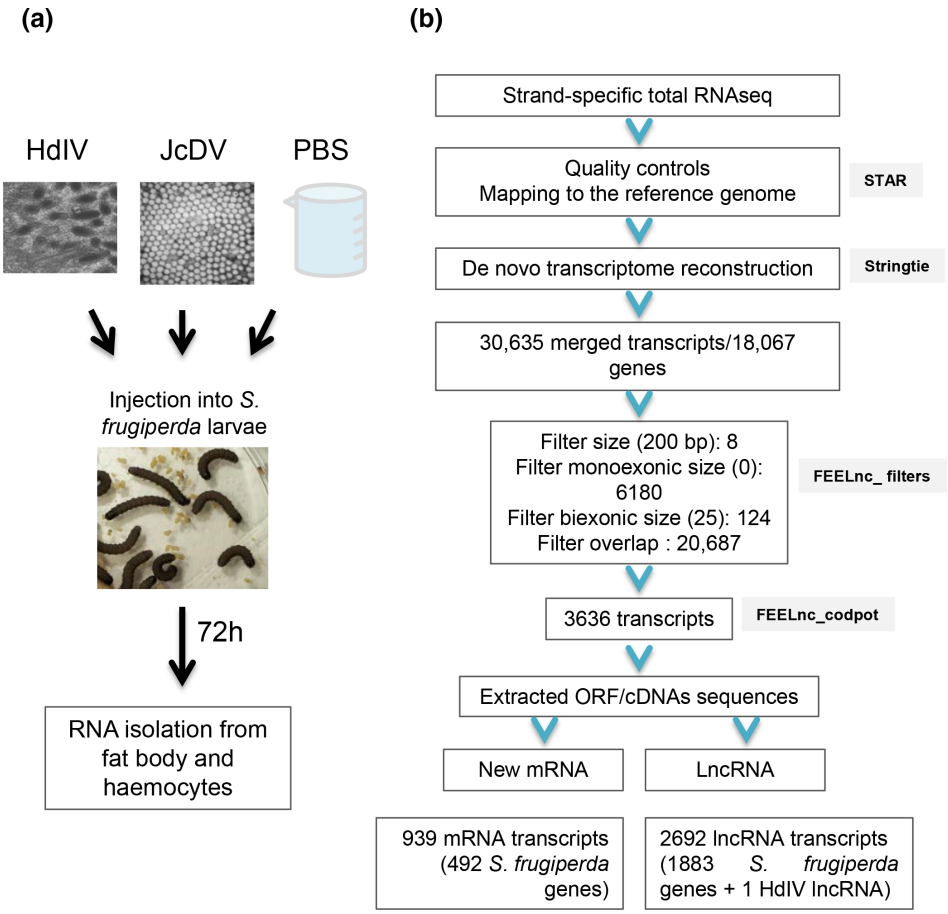


**Fig. 1.** Detailed workflow for lncRNA identification in haemocytes and fat body of *S. frugiperda*. (a) Schematic overview of the experimental procedure. HdIV (0.5 weq/larva) and JcDV (7.2e10 vg/larva) were injected into fifth-instar larvae (1 day old). Larvae injected with PBS were used as uninfected control. Total RNA from haemocytes and fat body was recovered 72h after infection. (b) Overview of the analysis pipeline used for lncRNAs identification in *S. frugiperda* haemocytes and fat body. The number of transcripts filtered by FEELnc_filters are indicated in the rectangle corresponding to this step.

activation step of 95 °C for 2 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 60 s. The primers specific for each lncRNA gene were designed with Primer3 software and are listed in Table S10. Levels of mRNA were normalized relative to the expression of the housekeeping gene *S. frugiperda rpl32* and compared to the PBS-treated larvae using the $2^{-\Delta\Delta Ct}$ method [54]. The fold change in expression was expressed relative to the mean for the group of controls, which was arbitrarily assigned a value of 1. Data were transformed to $\log_2$ scale.

### Statistical analysis

Data were expressed as the mean+/−SD of three independent experiments with three biological replicates. The relative expression levels of genes were analysed using one-way analysis of variance (ANOVA) followed by Tukey's post-hoc test with GraphPad Prism 8.0. A *P*-value <0.05 (two-tailed) was considered statistically significant.

## RESULTS

### Identification of lncRNAs expressed in haemocytes and fat body of *S. frugiperda* larvae

To study the expression profiles of lncRNAs in response to viral infection and to analyse their virus- and tissue-specificity, we injected HdIV and JcDV directly into the haemolymph of *S. frugiperda* larvae and total RNA was extracted from haemocytes and fat body 72 h after infection (Fig. 1a). We generated a total of 18 RNA-seq paired-end libraries, comprising 3 independent biological replicates for haemocytes and 3 others for fat bodies from *S. frugiperda* larvae injected with HdIV, JcDV and PBS (uninfected control). Strand-specific high-throughput parallel Illumina sequencing was then performed. The sequencing data were analysed using the nf-core rnaseq v3.0 pipeline (https://nf-co.re/rnaseq) [43], as schematized in Fig. 1b. The resulting statistical analysis of mapping is summarized in Table S1 (available in the online version of this article). An average of 23.6 million clean

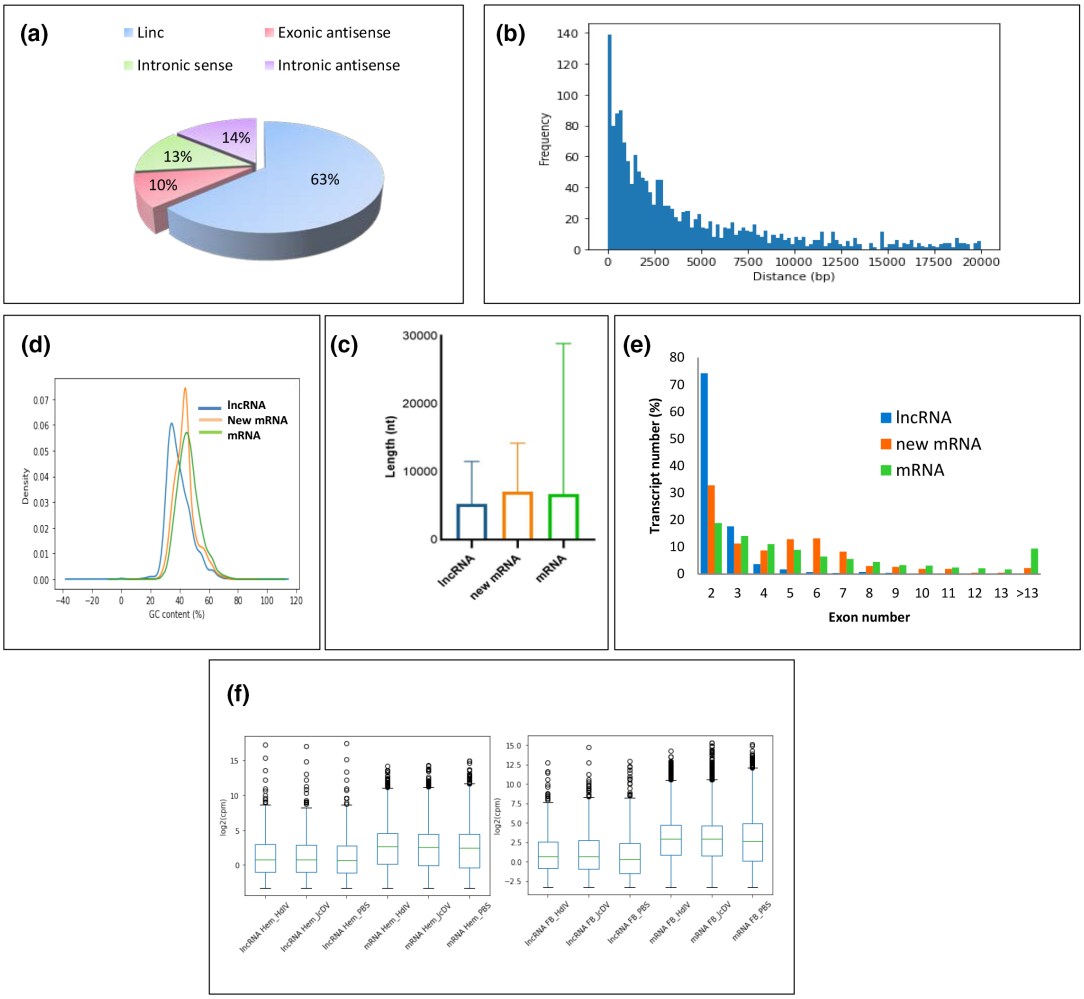


**Fig. 2.** General characteristics of lncRNAs in haemocytes and the fat body of *S. frugiperda* (a) The lncRNAs were classified into four categories based on their genomic locations and their direction of transcription relative to that of protein-coding gene. The pie chart indicates the percentage of lncRNA found in each class. (b) Frequency of lncRNA-mRNA distances (in bp). (c) Comparison of transcript lengths (in nt) between lncRNAs, novel mRNA and mRNAs. (d) Comparison of GC content (in %) between lncRNAs, novel mRNA and mRNAs. (e) Comparison of the number of exons between lncRNAs, novel mRNA and mRNAs. (f) Comparison of expression levels [$\log_2$(c.p.m.)] of lncRNA and mRNAs in fat body (FB, right panel) and haemocytes (Hem, left panel).

reads were obtained from each sample and approximately 18.5 million reads (78.6%) were successfully mapped onto the reference genome OGS6.1 of *S. frugiperda*, and the HdIV and JcDV genomes [33, 55, 56]. A heatmap of the correlation between the samples and multidimensional scaling (MDS) plots, produced from the matrix of CPM, revealed that the samples of the same tissue (fat body or haemocytes) are clustered, and within these two clusters the three replicates of the same condition (HdIV infection, JcDV infection or PBS) are clustered (data not shown). We identified a total of 30 635 transcripts, corresponding to 18 067 genes. To identify potential lncRNAs, FEELnc was applied to the transcriptome assembly, using several filtering steps, including transcript length ≥200 nucleotides (nt) and exon number ≥2 [50]. LncRNA candidates that overlapped with coding genes in the same sense were discarded from the analysis.

## Sequence feature analysis of lncRNAs

We identified a total of 2692 lncRNA transcripts (1884 genes), including 1 lncRNA derived from HdIV, and 939 novel mRNAs (492 genes), from the transcriptome data, including control and virus-infected samples. The most abundant class (63%) of the lncRNAs is located in intergenic regions. The remaining lncRNAs were derived from exonic antisense (10%), intronic sense (13%) and antisense (14%) at protein-coding loci (Fig. 2a). The vast majority of intergenic lncRNAs are located within 10 000 nt of a coding gene (Fig. 2b). The length of lncRNA transcripts ranged from 241 to 55 298 nt, with a mean of 5226 nt and a median of 2709 nt, which is shorter than the length of mRNAs and novel mRNAs (Fig. 2c). We also observed that the GC content in

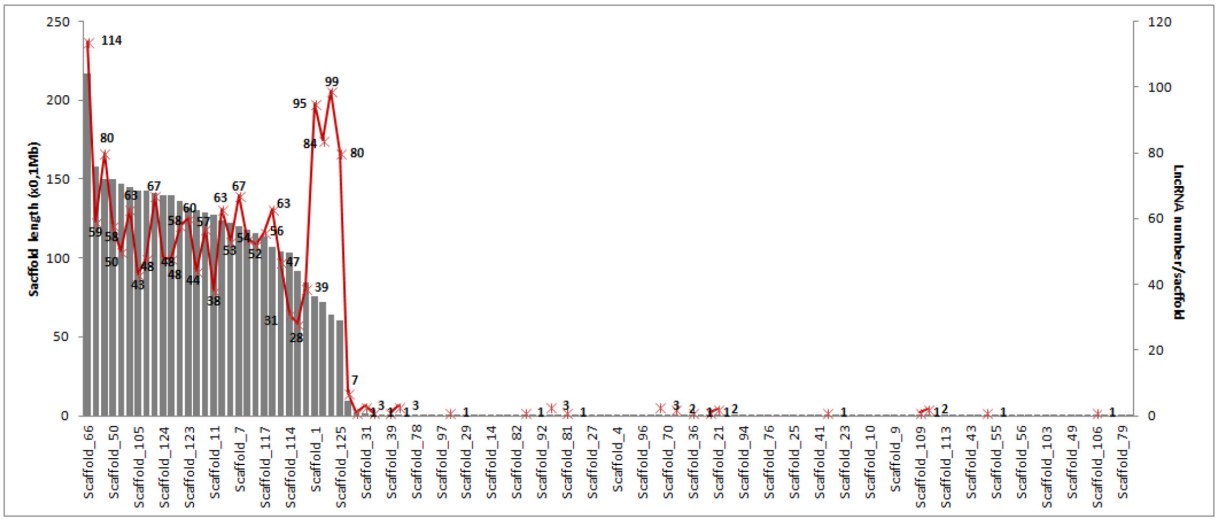


**Fig. 3.** Distribution of *S. frugiperda* lncRNA on the chromosome-sized scaffolds The number of lncRNAs localized on each scaffold versus the length of each scaffold (×0.1Mb) is indicated by a red line and by the grey bars, respectively.

lncRNAs was significantly lower than in mRNAs, i.e. with means of 32 and 44% of GC, respectively (Fig. 2d). Genetic structure analysis revealed that most lncRNAs had two or three exons, fewer than mRNAs, which have a wider distribution of exon numbers (Fig. 2e). Comparison of expression levels also showed that lncRNAs had a lower overall expression level than mRNAs (Fig. 2f).

The new high-quality genome of *S. frugiperda* assembly by Gimenez *et al.* generated 125 chromosome-sized scaffolds, with the 27 longest scaffolds accounting for 90% of the genome assembly [56]. Analysis of the genomic localization of *S. frugiperda* lncRNAs showed that they were distributed among the 27 superscaffolds (Fig. 3). A large number of lncRNA (114 lncRNAs or 6.1% of the 1883 lncRNAs) are located on the longest scaffold, 66, which is 21 694 391 bp in length. However, some smaller scaffolds also contain a significant number of lncRNAs. For example, scaffold 122, with a length of 6 389 370 nt, contains 99 lncRNAs, i.e. 5.3% of lncRNAs. Analysis of the position of lncRNAs on each of the scaffolds did not reveal any obvious clustering (data not shown). This suggests that the lncRNAs were distributed unevenly across the 31 chromosomes of *S. frugiperda*.

## Conservation analysis of lncRNAs in Lepidoptera

We performed a BLASTN analysis using the 2691 *S. frugiperda* lncRNA transcripts to search for the presence of homologues in the published and high-quality genome of 12 other lepidopteran species (*Amyelois transitella*, *Bombyx mori*, *Chilo suppressalis*, *Danaus plexippus*, *Manduca sexta*, *Helicoverpa armigera*, *Plutella xylostella*, *Heliconius melpomene*, *Operophtera brumata* and *Plodia interpunctella*, *Spodoptera exigua* and *Spodoptera litura*). Sequences having a hit with an e-value cut-off <1e-10 and covering at least 70% of the *S. frugiperda* lncRNA sequence were considered homologous. We identified 1104 *S. frugiperda* lncRNAs sharing one hit in at least one of the other lepidopteran species (Table S2). A maximum number of hits was observed for the two closely related species, *S. litura* (4160 hits) and *S. exigua* (3621 hits), while the smallest number of hits were found with the most distantly related species, *C. suppressalis* and *P. xylostella*. We have not found any *S. frugiperda* lncRNAs with a homologous sequence present in all 12 other lepidopteran species. This result suggests that some *S. frugiperda* lncRNAs have orthologues in closely related species (*S. litura* and *S. exigua*), but exhibit low evolutionary sequence conservation in Lepidoptera.

## Identification of lncRNAs regulated by virus infection

We then searched for lncRNAs whose expression is significantly altered in haemocytes and fat body in response to HdIV and JcDV infection. Genes with an overall expression >0.5 counts per million reads mapped (c.p.m.) in at least three libraries were considered significantly transcribed. At this step, the HdIV-derived candidate lncRNA identified at the end of the bioinformatic pipeline (see above, paragraph 'Sequence feature analysis of lncRNAs') was not retained, because its expression level did not meet this criterion. A total of 1403 lncRNAs was finally obtained. Hierarchical clustering analysis showed that they were classified into separated groups, according to the cell type (haemocytes or fat body), and infected vs uninfected samples (Fig. 4a). Using a false discovery rate (FDR) <0.05 and fold change≥2.0 relative to the uninfected control, we found a total of 529 (38%) lncRNAs showing significant changes in expression in response to viral infection. Fig. 4b indicates the number of differentially expressed lncRNAs (DE-lncRNAs) in each individual dataset compared to the corresponding uninfected control. Overall, only a few lncRNA genes were downregulated compared with upregulated ones in response to HdIV and JcDV infection, and more differentially expressed lncRNAs are found in fat body than haemocytes, regardless of the virus (HdIV or JcDV). Venn diagrams depict the

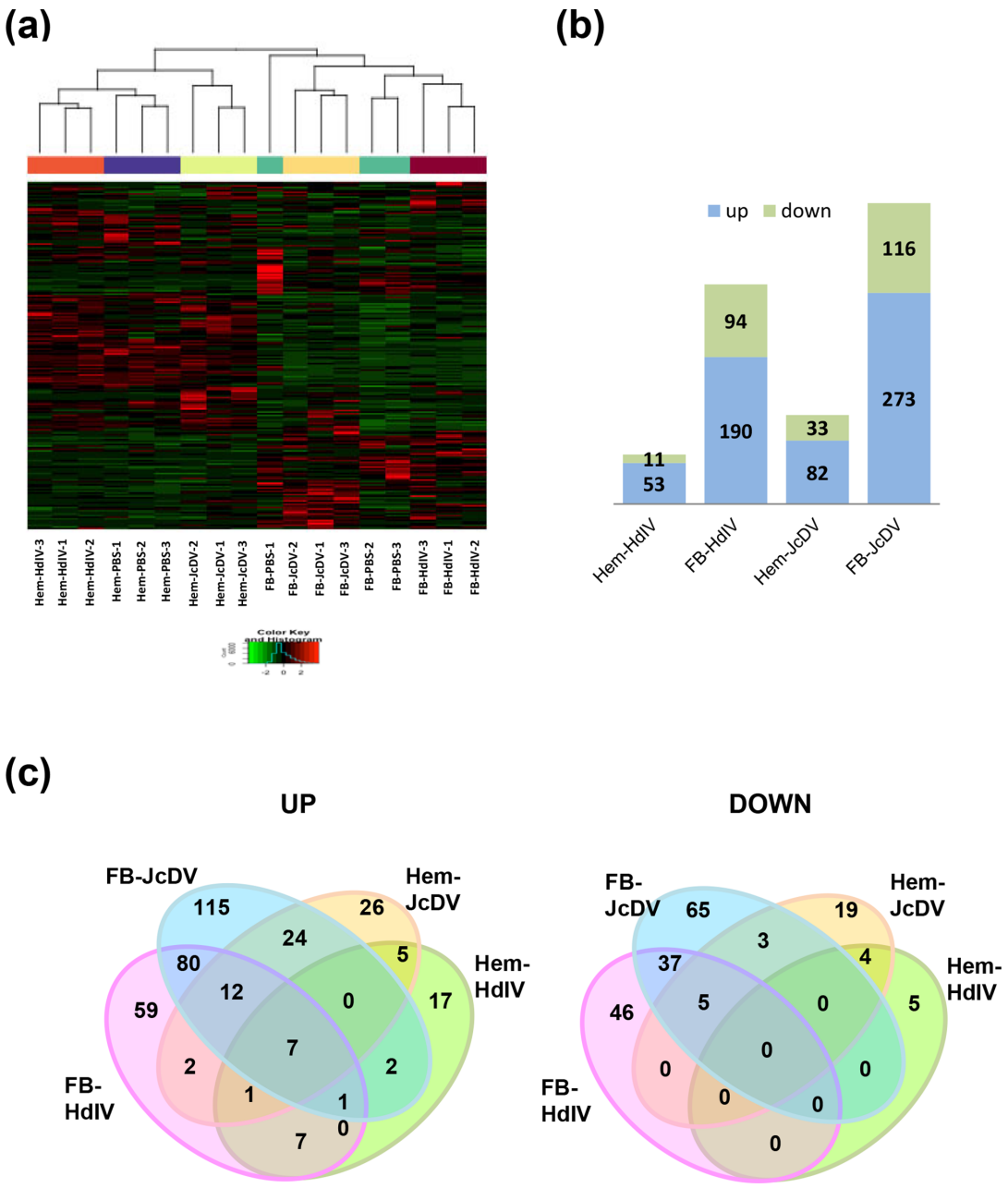


**Fig. 4.** Expression analysis of lncRNAs in response to virus infection. (a) Hierarchical clustering of the 1403 lncRNA expressed in haemocytes and fat body, after infection with HdIV and JcDV or uninfected (PBS) (three replicates per experimental condition). Green indicates low relative expression; red indicates high relative expression. Hem, haemocytes; FB, fat body. Each row represents an individual RNA and each column represents a replicate. (b) The number of DE-lncRNAs (up in blue, down in green) in infected samples relative to uninfected control samples. (c) Venn diagrams showing the distribution of significantly upregulated (left) and down-regulated (right) lncRNAs compared to uninfected samples.

distribution of DE-lncRNAs in the infected samples compared to the corresponding uninfected samples (Fig. 4c). The majority of lncRNAs (88.5%) showed tissue- and/or virus-specific expression patterns. For example, 217 upregulated and 135 downregulated lncRNAs are specifically regulated in only 1 of the 4 experimental conditions. On the other hand, several lncRNAs were regulated by both viruses either in haemocytes or fat body, suggesting that they are involved in the cell type-specific antiviral response. A small group of lncRNAs is also regulated by only one of the two viruses simultaneously in haemocytes and fat body, probably reflecting the specific host–virus interaction. Seven overexpressed lncRNAs were found to be shared between all four infected samples, indicating that they might contribute to a general host response to pathogen infection. A list of the DE-lncRNAs for each individual dataset is provided in Table S3. A set of randomly chosen DE-lncRNAs were selected and analysed for their expression

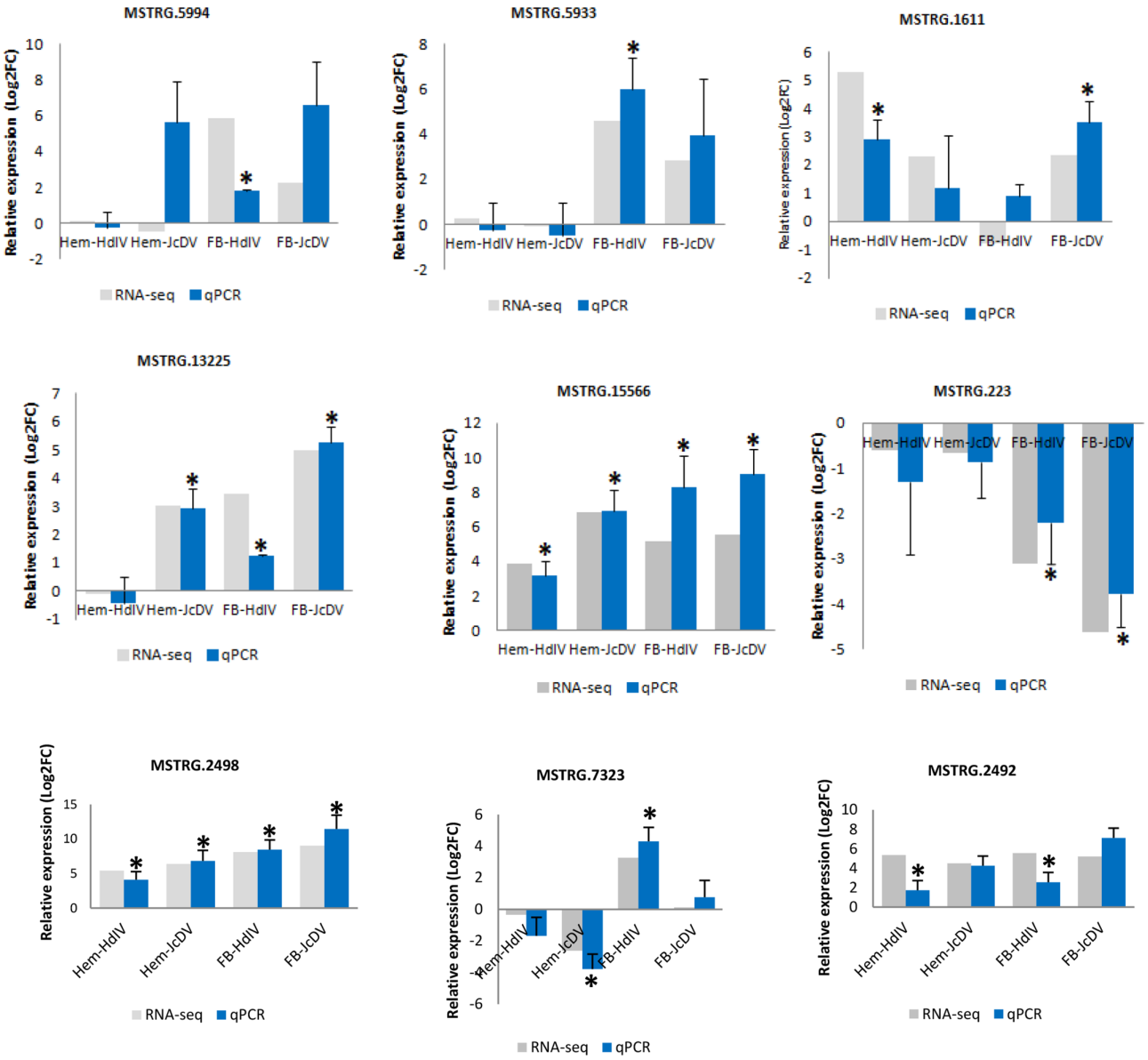


**Fig. 5.** Validation of randomly selected virus-reponsive lncRNAs by RT-qPCR. Graphs show the expression levels ($log_2$ fold change) of randomly selected host lncRNAs in *S. frugiperda* larvae infected with HdIV and JcDV relative to the uninfected control samples. LncRNA expressions were normalized with the housekeeping gene *rpl32*. For each gene, the relative fold change was calculated vs that of PBS-treated control (calibrator = 0) and presented as a bar graph. All the RT-qPCR data are represented as the mean+/−SD from three biological replicates, each with three technical replicates. Asterisks indicate statistically significant differences ($P<0.05$) as compared to PBS-treated samples (control).

levels across all the conditions by RT-qPCR (Fig. 5). The results showed similar trends to the data obtained by RNA-seq, thereby confirming the expression levels observed in the RNA-seq analysis.

## Expression profiling and functional enrichment of DE-mRNAs

To explore the potential function of DE-lncRNAs, we first examined mRNA expression profiles in infected haemocytes and fat body. The same selection filters used above to identify DE-lncRNAs were applied to identify DE protein-coding genes. The analysis revealed that both HdIV and JcDV infection had significant and profound effects on gene expression compared with uninfected samples. Hierarchical clustering analysis showed that the DE-mRNAs were classified into separated groups, according to the cell type and infected vs uninfected samples (Fig. 6a). Similarly to lncRNAs, the expression profiles of host-responsive protein-coding genes are more altered in virus-infected fat body than haemocytes, whatever the virus considered (Fig. 6b). We also noted a greater transcriptome alteration by JcDV than by HdIV. Venn diagrams showed that 80 and 9 mRNAs were overexpressed and

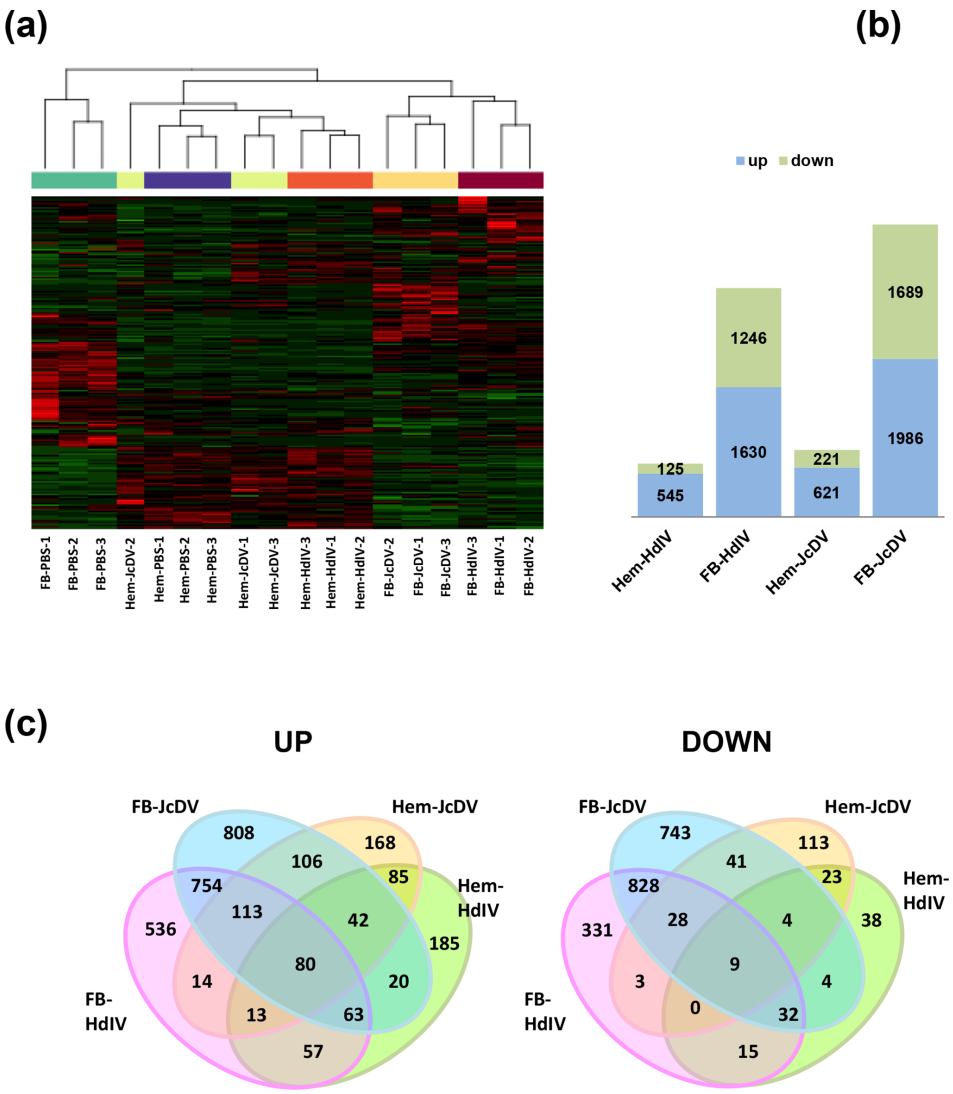


**Fig. 6.** Differentially expressed mRNAs in response to virus infection. (a) Hierarchical clustering of the 10 746 DE-mRNA expressed in haemocytes and fat body challenged with HdIV and JcDV or uninfected (three replicates per experimental condition). Green indicates low relative expression; red indicates high relative expression. Each row represents an individual RNA and each column represents a replicate. (b) The number of DE-mRNAs in infected samples relative to uninfected control samples. (c) Venn diagrams showing the number of significantly upregulated (left) and downregulated (right) mRNA compared to uninfected samples. Hem, haemocytes; FB, fat body.

downregulated, respectively, across all infected samples compared to uninfected samples (Fig. 6c), suggesting that these genes may participate in a general defence response against pathogens. Further, the distribution of other mRNAs across the four samples may reflect specific molecular mechanisms triggered by each virus in haemocytes and fat body.

We then performed GO enrichment analysis on the DE protein-coding genes, according to the three GO categories, namely Biological Processes (BP), Cell Component (CC) and Molecular Function (MF) (Fig. 7). Some terms were shared by samples infected by HdIV and JcDV. For example, fat body infected with HdIV and JcDV was enriched in upregulated genes associated with the two MF terms 'Calcium ion binding' and 'Glycosyltransferase activity'», while HdIV- and JcDV-infected haemocytes were enriched in 'Catalytic activity' and 'Carbohydrate binding'. The shared GO terms associated with downregulated genes were mainly enriched in several metabolic and biosynthetic processes in both HdIV- and JcDV-infected fat body. Besides, some GO terms were virus-specific. JcDV-infected fat body was enriched in two BF terms, 'Cellular response to stimulus' and 'Signal transduction', while HdIV-infected fat body was enriched in 'Regulation of cellular process' and 'Signaling'.

To better understand the biological functions of DE genes, we then performed KEGG pathway enrichment analyses (Table S4). Overall, the results showed enrichment in various metabolic processes (carbohydrate, amino acid, energy, cofactors and vitamins, xenobiotic) in immune tissues infected with both HdIV and JcDV. Specific enrichment in certain pathways was also observed

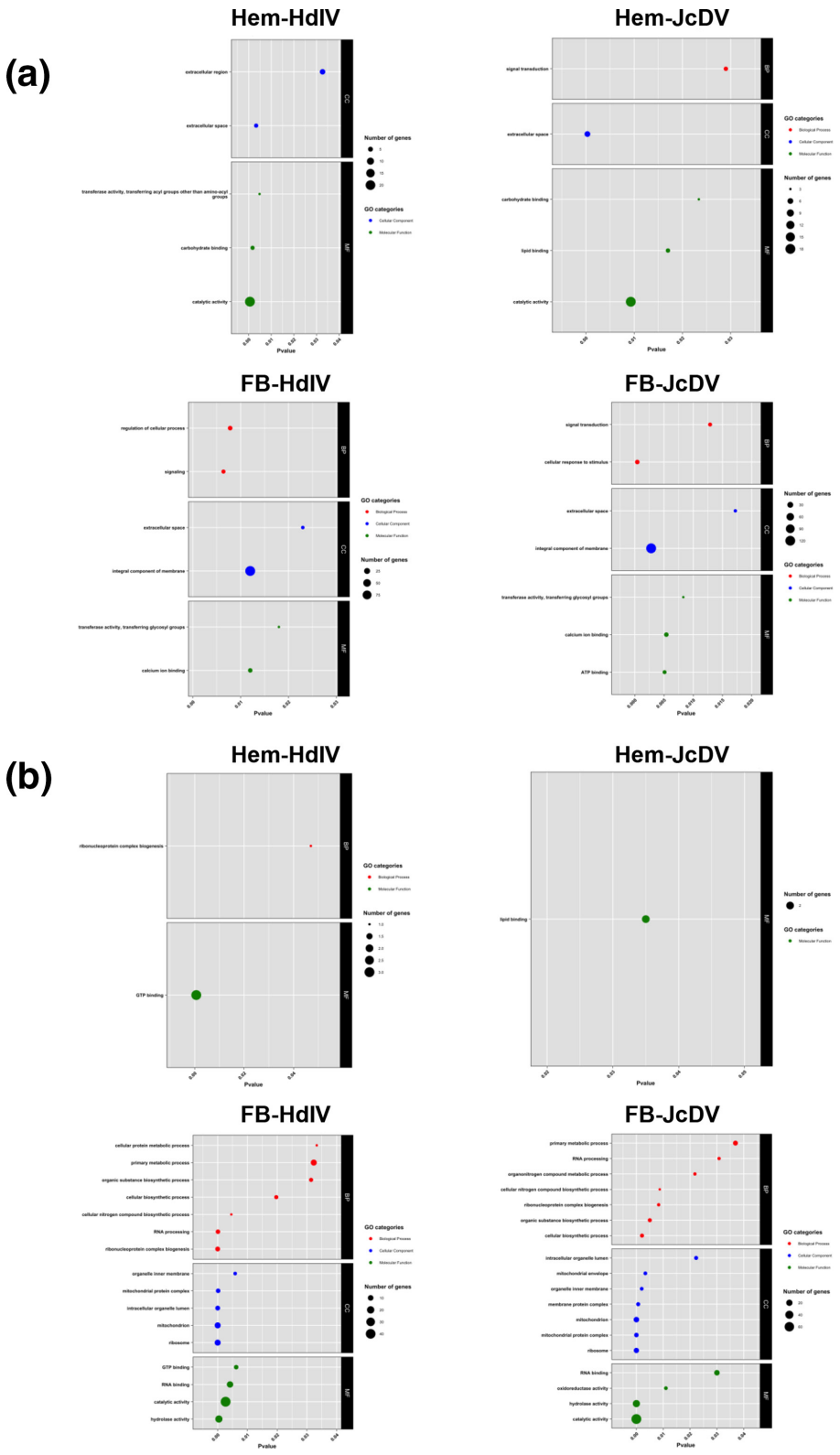


**Fig. 7.** Gene Ontology analysis of differentially expressed mRNAs in response to HdIV and JcDV infection GO analysis was applied within either upregulated (a) or downregulated (b) genes. Red, blue and green represent 'Biological process', 'Cell component' and 'Molecular function' categories, respectively. The number of DE genes in each GO pathway is represented by the size of the dots. The ratio significant/expected is the ratio between the significant values ($P$<0.01) and values expected in the pathway.

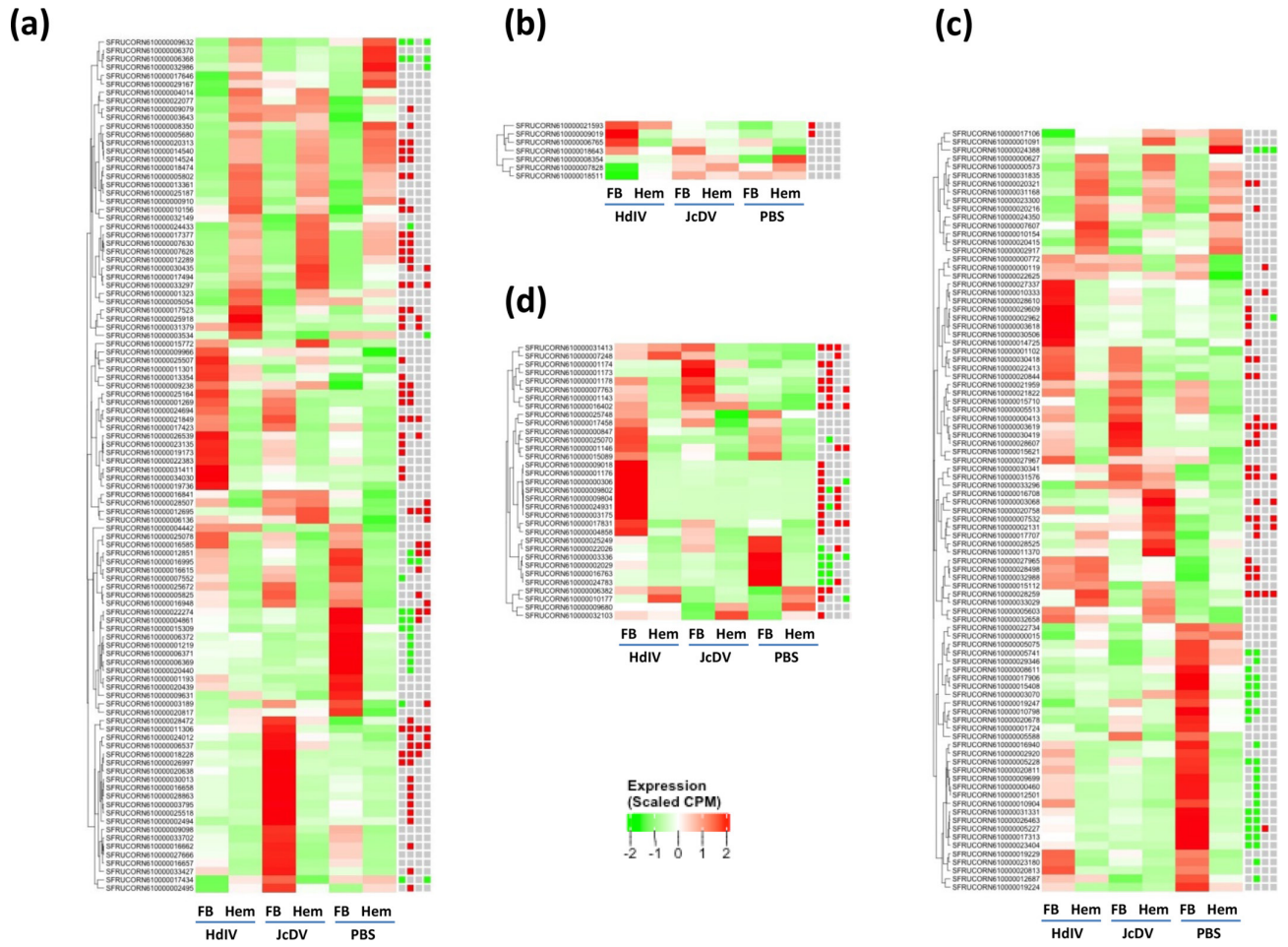


**Fig. 8.** Heatmaps with clustering analysis of the differentially expressed mRNAs involved in immune response (a), RNAi (b), cell death and oxydative stress (c) and molting (d). The squares on the right side of the heatmap indicate the regulatory status of the genes: green and red indicate that the gene is significantly up- and downregulated, respectively, compared to the uninfected control sample; from left to right: FB-HdIV, FB-JcDV, Hem-HdIV and Hem-JcDV.

according to the tissue and the virus. In HdIV-infected fat body, pathways related to cell matrix adhesion and infectious diseases were selectively enriched for upregulated genes, while the downregulated pathways were particularly associated with translation. In HdIV-infected haemocytes, the upregulated genes were significantly enriched within pathways related to hormone synthesis and the immune system. The KEGG pathways enriched in downregulated genes were related to DNA replication and repair and translation. In JcDV-infected fat body, the upregulated genes were enriched in MAPK signalling pathways, immune system, DNA replication and autophagy, while the downregulated genes were mainly enriched in pathways related to translation. JcDV-infected haemocytes were enriched in RNA processing and immune system pathways for upregulated genes, and pathways related in metabolic processes for donwregulated genes.

We also focused on biological functions expected to be altered by HdIV and JcDV pathogenesis, including immune defence, RNA silencing pathways, programmed cell death and the larval moulting process. We found that several immune key genes, including *lysozyme*, *hemocytin*, *spaetzle 5*, *cactus*, *Toll-6* and *Toll-7*, and several genes involved in activation of the phenoloxidase cascade (*PPAF2*, *PPO1*, *PPO2*, *PPAE*), were differentially regulated compared to uninfected samples, according to the tissue and the virus (Fig. 8a). We also found that key factors of the RNA silencing pathways were not differentially expressed during HdIV and JcDV infection, with the exception of *dicer1* and *Piwi-like protein Siwi*, two members of the PIWI-interacting RNA (piRNA) pathway, which were specifically overexpressed in the HdIV-infected fat body (Fig. 8b). Upregulation of these two factors raised the question of their role in the interaction between the insect host and the polydnavirus. Interestingly, some studies have previously revealed the contribution of the piRNA pathway in the antiviral immune response in mosquitoes and lepidopteran cell lines [57–59], suggesting that HdIV triggers a piRNA-based immune defence in the fat body of *S. frugiperda*.

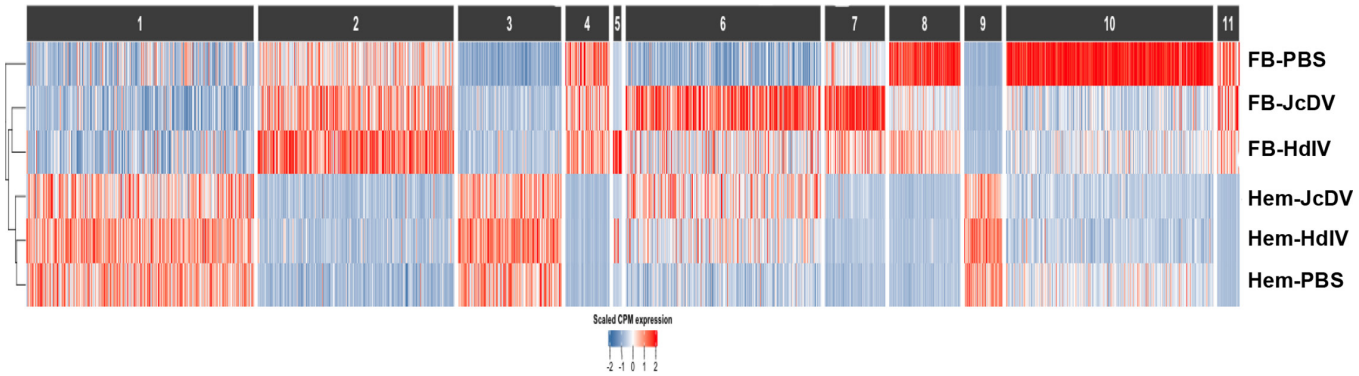


**Fig. 9.** Hierarchical clustering of the 11 co-expressed gene clusters. Blue indicates low relative expression and red indicates high relative expression. Each column represents an individual RNA and each row represents a sample.

Previously, we demonstrated that HdIV and JcDV infection viruses modulate two important programmed cell death pathways, autophagy and apoptosis, respectively. In particular, injection of HdIV into *S. frugiperda* larvae induced apoptosis in haemocytes and fat body 24 h post-infection [35], while infection of the IPLB-Ld652Y cells with JcDV activated autophagy by inhibiting the TOR pathway 4 days after infection [60]. In the present work, we found that many genes involved in positive regulation of apoptosis, such as *dronc*, *caspase 1*, *apaf1*, *wengen*, p*rotein lifeguard 1*, *innexin 2* and *innexin 3*, were upregulated in the two immune tissues 72 h after infection with HdIV and JcDV (Fig. 8c). On the other side, two autophagy-related genes, *atg3* and *atg13,* were downregulated in JcDV-infected fat body, suggesting that JcDV infection does not trigger autophagy in fat body. Further experimental studies will be required to clarify which is the right hypothesis.

Infection of *S. frugiperda* larvae with HdIV and JcDV leads to the inhibition of larval moulting. In both cases, the underlying molecular mechanisms are not well understood. Nevertheless, several studies have shown that polydnaviruses can disrupt hormonal events responsible for the initiation of metamorphosis [39, 61, 62]. By examining the expression profiles of some genes related to this process, we found that several enzymes involved in the juvenile hormone (JH) degradation pathway, such as JH esterase and JH epoxide hydrolase, were downregulated in HdIV- and JcDV-infected fat body, suggesting that maintenance of JH levels in infected larvae prevents normal moulting (Fig. 8d).

Altogether, these results showed that 72 h after infection HdIV and JcDV infections modulated several important metabolic and physiological functions, including the immune response, programmed cell death and the moulting process in haemocytes and fat body of *S. frugiperda*.

## Functional prediction of the DE-lncRNAs: co-expression clustering analysis

We used the coseq (co-expression RNA-seq data) package [49] to identify clusters of genes with similar expression patterns across all samples. This approach is based on the principle of guilt by association, according to which genes that share the same function or are involved in the same regulatory pathway will tend to have similar expression profiles [63]. Thus, the function of co-expressed unknown genes can be predicted by identifying genes with known functions in each cluster. Our coseq analysis revealed 11 coexpression clusters. The distribution of DE-lncRNAs and DE-mRNAs in each cluster is depicted in Fig. 9 and Table S5. All 11 clusters were enriched in lncRNAs, but HdIV transcripts were only observed in clusters 1, 2 and 5, and JcDV transcripts were only detected in cluster 7. Cluster 1 comprised 473 transcripts, including 103 lncRNAs, 361 mRNAs and 9 HdIV transcripts, differentially expressed in at least 1 of the 4 experimental conditions (Table S5). Overall, all transcripts in cluster 1 had higher expression levels in haemocytes than in fat body, suggesting that this cluster groups haemocyte-specific mRNAs and lncRNAs. GO enrichment analysis revealed a significant enrichment of genes involved in 'Catalytic complex' in terms of cellular component, and 'DNA binding' in MF category (Table S6). KEGG analysis showed that cluster 1 is significantly enriched in several metabolic pathways, cellular processes and diseases pathways (Table S6). Cluster 2 grouped 369 transcripts, including 41 DE-lncRNAs, 314 DE-mRNAs and 14 HdIV transcripts. In contrast to cluster 1, all these transcripts have a higher level of expression in fat body than haemocytes, suggesting that they act primarily in fat body. Only the 'Zinc ion binding' GO term is significantly enriched in this cluster (Table S6). According to the KEGG analysis, the pathways associated with 'Adipocytokine signaling pathway', 'Spliceosome', 'Oxidative phosphorylation', ligand–receptor interaction and human diseases, were significantly enriched in this cluster 2 (Table S6). Interestingly, several immune genes, including *lysozyme*, *spaetzle 4 and 5* and *hemolin*, all of which are upregulated in the HdIV-infected fat body, were also clustered. Some lncRNAs in this cluster may therefore be involved in the regulation of immune defence against HdIV. On the other hand, cluster 5 grouped 58 transcripts, including 3 DE-lncRNAs, 16 DE-mRNAs and 39 HdIV transcripts. GO and KEGG enrichment analyses showed that genes from cluster 5 were enriched

in 'Transferase activity, transferring sulphur-containing groups', and associated with an inflammatory response pathway (Table S6). On the other hand, many of the HdIV transcripts were clustered with genes involved in hormone metabolism, suggesting a functional relationship between these two groups of genes (Table S5). Only three lncRNAs are present in cluster 5. Further experimental study is needed to test whether these lncRNAs are regulators of the hormone pathway in *S. frugiperda*. Cluster 7 grouped 226 transcripts, including 3 JcDV transcripts, as well as 30 DE-lncRNAs and 193 DE-mRNAs. GO enrichment analysis showed that all of them were associated with the GO term 'Structural component of cuticle' (Table S6). KEGG analysis revealed only one pathway, which is related to the immune system (Table S6). Indeed, cluster 7 comprised different genes involved in the melanization immune response, such as *phenoloxidase-activating factor 2* and *serine protease inhibitor 27A*, which are upregulated in JcDV-infected haemocytes and fat body.

Overall, the analysis showed that DE-lncRNAs clustered with HdIV and JcDV transcripts were mainly associated with genes enriched in functions related to metabolic process, signal transduction, metabolism of insect hormone and immune system.

### Functional prediction of the DE-lncRNAs: co-expression of lncRNAs with neighbouring protein-coding genes

A great proportion of lncRNAs have been shown to act as *cis* regulators of neighbouring coding genes. To examine whether *S. frugiperda* lncRNAs act in cis to regulate nearby gene expression, we used the same approach developed by Le Béguec *et al.* [51]. For that, we searched for the target protein-coding genes located within a 1 Mb sliding window around each DE-lncRNA, transcribed in divergent and convergent orientations relative to the neighbouring lncRNA, and whose expression was significantly correlated with the DE-lncRNA [51]. Table S7 describes the distribution of the cis-acting interactions for each experimental condition, using an adjusted *P*-value <0.05 and a Spearman correlation coefficient between −0.5 and 0.5 as the cut-off values. The distance between one lncRNA and its neighbouring paired coding gene was found to be quite large, 431 kb on average. In the HdIV-infected fat body, the neighbouring genes were involved in various metabolic processes, regulation of gene expression, RNA processing, protein modification and cell adhesion. We also found some lncRNAs associated with immune genes, such as *MSTRG.5019*, which is co-expressed with *phenoloxidase subunit 2* (*PO2*), a key enzyme involved in the melanization cascade activated by an immune challenge. Both genes were upregulated following HdIV infection, suggesting that *MSTRG.5019* acts as a positive regulator of *PO2*. In the fat body of JcDV-infected larvae, the coding genes are involved in a wider variety of functions, including metabolic processes, processes related to protein modification, localization and folding, vesicle-mediated transport, RNA processing, carbohydrate transport and regulation of gene expression. Some lncRNAs were also co-expressed with immune genes. This is the case for the intergenic lncRNA *MSTRG.11283*, which is associated with *lysozyme* mRNA in both HdIV- and JcDV-infected fat body. Interestingly, one of the co-expressed gene pairs involves the intergenic lncRNA *MSTRG.14514* and *atg3*, a key gene involved in autophagy. MSTRG.14514 was overexpressed in JcDV-infected fat body, while *atg3* was downregulated. This result suggests that the lncRNA negatively regulates the expression of *atg3*, thereby inhibiting autophagy 72 h after larval infection. The analysis also revealed a significant co-expression between the intergenic lncRNA *MSTRG.10979* and *spn28D*, which encodes the serine protease inhibitor 28Dc (*spn28D*), which regulates the phenoloxidase cascade [64]. *MSTRG.10979* was upregulated in JcDV-infected fat body, whereas *spn28D* was downregulated, suggesting that stimulation of *MSTRG.10979* expression by JcDV results in inactivation of the *spn28D* gene. Far fewer co-expressed lncRNA-mRNA pairs were found in haemocytes. In HdIV-infected haemocytes, the protein-coding genes are mainly involved in metamorphosis and moulting, in the regulation of gene expression and in cell adhesion. In JcDV-infected haemocytes, the coding genes are rather involved in metabolism, cell adhesion and oxidoreductase activity.

### Identification of *S. frugiperda* lncRNAs as a potential source of miRNA

LncRNAs can serve as precursors for small non-coding RNAs, including microRNAs (miRNAs) [65]. To determine whether any of the lncRNAs identified in our study produce annotated miRNAs, we performed a BLASTN analysis using the virus-responsive lncRNAs and the published miRNA repertoire of *S. frugiperda* [52]. We identified five DE-lncRNAs whose genes harbour miRNAs. Three of them, MSTRG.4700, MSTRG.8383 and MSTRG.98, were downregulated in at least one of the infected samples compared with uninfected tissues, with the exception of MSTRG.8383, which is upregulated in the JcDV-infected fat body (Table S8). These three lncRNAs can act as miRNA precursors of miR-8–5 p, miR-970–5 p and let-7–5 p precursors, respectively. Two miRs, miR-8–5 p and let-7–5 p, were located within an exon, while miR-970–5 p was embedded within an intron. The candidate mRNA targets of miR-8–5 p, let-7–5 p and miR-970–5 p were predicted using TargetScan and miRanda software (Table S9). Because binding of an individual miRNA to the 3′ untranslated regions of its mRNA targets is known to be an important mechanism for downregulation of target gene expression, we can expect the target genes to be upregulated due to the downregulation of the precursor lncRNAs and conversely downregulated in the case where the precursor lncRNA was upregulated.

## DISCUSSION

The role of lncRNAs in immune defence is the subject of extensive research. Numerous transcriptomic studies have demonstrated changes in the expression of lncRNAs following viral infection. While most of the evidence comes from vertebrates, studies revealing the role of lncRNAs in insect host–pathogen interactions remain scarce [27, 29, 66]. In our study, high-throughput RNA

sequencing was used to analyse the expression profiles of lncRNAs and mRNAs in the two major immune tissues, haemocytes and fat body, of *S. frugiperda* larvae infected with two viruses that have different life cycles.

A total of 1403 candidate lncRNAs, displaying structural features commonly observed in lncRNAs from other organisms, were significantly transcribed in the immune tissues. Among them, 529 lncRNAs showed significant changes in expression following viral infection, compared with uninfected control tissues. Tissue- and/or virus-specific expression patterns were observed for 468 DE-lncRNAs, whose 322 were specifically regulated in only 1 of the 4 experimental conditions. Fat body-specific lncRNAs were more abundant (69.2% upregulated and 79.7% downregulated genes) than haemocyte-specific lncRNAs (13.1 and 16.2% up- and downregulated genes, respectively). In agreement with this, abundant expression of lncRNAs in the fat body compared with other tissues was also observed in honey bees [27]. Our results also indicated that many DE-lncRNAs are specifically regulated by only one of the two viruses. JcDV-specific lncRNAs were more abundant than HdIV-specific ones. A similar trend was also observed for differentially expressed protein-coding genes. This difference could be related to the size of their respective genomes. Indeed, the HdIV genome is 306.9 kbp in size and encodes 152 genes expressed throughout the parasite cycle, which allows the virus to manipulate the host machinery to its advantage. In contrast, JcDV is a virus with a much smaller genome size (~5.8 kbp) and only encodes seven genes. Therefore, it is possible that to establish and complete its infectious cycle JcDV has a greater need to hijack host factors, including mRNA and lncRNAs, which could act as modulators of the latter, for its benefit. The difference in replication capacity between HdIV and JcDV may also explain the observed difference in the number of DE genes and DE-lncRNAs. HdIV transcripts expressed in the parasitized host only encode virulence factors necessary for successful parasitism, i.e. suppression of the host insect immune system to benefit parasite development. In contrast, JcDV replication results in significant production of viral progeny in infected cells. This production may require the involvement of many more host genes.

The ability of lncRNAs to interact with a variety of molecules, including DNA, RNA and proteins, allows them to be involved in a wide range of biological processes, through various mechanisms. Although many insect lncRNAs have been identified, the biological functions of most of them remain poorly understood, especially in non-model insects. In addition, the low conservation of lncRNAs, which is generally observed between species, makes it more difficult to elucidate their biological function [10]. To predict the potential functions of *S. frugiperda* lncRNAs, we first examined the physiological processes altered upon viral infection by analysing enrichment in GO terms and KEGG pathways on DE-mRNA in each expression dataset. The analyses showed that the preferentially enriched pathways were assigned to various metabolic processes. This result was not unexpected, as it is now clear that viruses significantly alter host cell metabolism upon infection. Because viral replication and the life cycle are resource-intensive, viruses have developed various strategies to hijack the host cell machinery to obtain various nutrients and metabolites, including amino acids, fatty acids and nucleotides, and to generate energy [67]. Some GO terms and KEGG pathways were enriched in a virus-specific manner in infected samples, highlighting that the metabolic reprogramming imposed on target cells is pathogen-specific. Indeed, HdIV and JcDV differ in their life cycle, for example in terms of viral product synthesis and replication capacity. The JcDV-infected fat body is enriched in upregulated genes involved in DNA replication, compared with HdIV, a replication-deficient virus. The replicative life cycle of many viruses is known to affect the DNA damage response and repair pathways in infected cells [68]. It is therefore possible that JcDV must activate DNA repair pathways to promote replication of its own genome. In contrast, downregulation of DNA replication and repair pathways was observed in HdIV-infected haemocytes. Interestingly, several studies have demonstrated integration of polydnavirus sequences within the chromosomes of infected lepidopteran host cells and tissues, allowing persistent expression of virulence genes throughout parasitism [69–72]. This mechanism generally creates lesions in the host DNA and requires DNA breakage and joining reactions, possibly followed by the final step of gap repair. While one might assume that, in the case of polydnavirus, activation of this mechanism promotes integration of viral DNA into host chromosomes, this is not what is observed. Further studies are therefore needed to clarify the link between DNA repair mechanisms and the polydnavirus life cycle.

In vertebrates, lncRNAs can actively regulate multiple and various aspects of cellular metabolism [73]. Studies have also shown that viruses use host lncRNAs to manipulate cellular metabolism to regulate, positively or negatively, their replication. For example, Wang *et al.* demonstrated that virus-induced lncRNA-ACOD1 directly binds to the key metabolic enzyme, namely the glutamic oxaloacetic transaminase GOT2, thereby stimulating its catalytic activity and production of its metabolites, which benefits viral replication [4]. Some lncRNAs are also involved in the DNA damage response, especially in cancer cells. Feldstein *et al.* showed that lncRNA ERIC limits apoptosis, a cellular response to DNA damage, induced by the chemotherapeutic drug etoposide [74].

To date, lncRNAs involved in metabolic regulation, viral replication and DNA repair have not been characterized in insects. IBIN, previously annotated as a lincRNA linking immunity and metabolism in *Drosophila*, was subsequently found to encode a putative short peptide [75]. However, considering the regulatory functions of vertebrate lncRNAs, it can be assumed that some DE-lncRNAs may serve the same type of functions to regulate different metabolic processes in virus-infected haemocytes and fat body *S. frugiperda*. The identification of lncRNAs that regulate metabolism during viral infection could benefit from co-expression analyses. Among many other examples, the expression of lncRNA *MSTRG.12547* is negatively correlated with that of the *hexokinase-2* gene, which encodes a critical enzyme of the glycolytic pathway in JcDV-infected fat body. Similarly, both lncRNAs, *MSTRG.17339* and *MSTRG.17340*, are co-expressed with the *adenylate kinase* gene, which is involved in cellular energy

homeostasis and adenine nucleotide metabolism. It may therefore be of interest to characterize the function of these co-expressed lncRNAs in the regulation of HdIV- and JcDV-infected fat body metabolism.

KEGG analysis revealed a significant enrichment of only two immune pathways, namely 'Cytokine–cytokine receptor interaction' and 'Antigen processing and presentation', which were upregulated in all infected tissues, except for the HdIV-infected fat body. This may not seem like much for the transcriptomes of virus-infected tissues. However, samples were prepared 72 h after infection. JcDV replication begins as early as 48 h post-infection [39]. In HdIV-infected caterpillars, expression of viral transcripts is already detected 2 h post-infection [38]. Therefore, at 72 h post-infection of host cells, the activity of both viruses is well underway, and the viruses may have already taken control of the host antiviral response. On the other hand, our transcriptomic analysis revealed the upregulation of the expression of key immune genes regulated in infected host cells, including several components of the Toll pathway, such as *cactus*, *pelle*, *Toll-7*, *Toll-6*, *PGRP-2* and *PGRP-LE* (peptidoglycan-recognition protein LE), as well as genes involved in the melanization process, such as *PPAE* (phenoloxidase-activating enzyme) and *phenoloxidase subunits 1* and *2*, suggesting that the host immune defence is not totally neutralized at this step of infection.

Recent functional analyses have shown that lncRNAs can be involved in different stages of the insect immune response. For example, *LncRNA-CR33942,* upregulated upon infection with *Micrococcus luteus*, physically interacts with the transcription factors, Relish, Dif and Dorsal, thereby inducing transcription of antimicrobial peptides regulated by the Toll and Imd pathways in *Drosophila* [76, 77].

In our study, we found a significant co-expression between *MSTRG.10979* and its neighbouring protein-coding gene *spn28D*. In *Drosophila*, inactivation of *spn28D* causes widespread melanization, a prominent innate immune response, of air-contact tissues, including the tracheae [64]. JcDV pathogenesis also affects tracheae in *S. frugiperda* larvae, but melanization is only observed on the cuticle, some time before death [39]. It will therefore be interesting to evaluate the hypothesis that the cuticle melanization observed in JcDV-infected larvae is under the control of the lncRNA *MSTRG.10979* acting as a negative regulator of *spn28D*.

We have previously shown that HdIV and JcDV manipulate the programmed cell death in infected host cells. JcDV infection induces autophagy in the gypsy moth cells IPLB-Ld652Y [60]. Our analysis indicated that *atg3* may be a cis-target of the *MSTRG.14514* lncRNA. While *atg3* is downregulated, *MSTRG.14514* is upregulated in the JcDV-infected fat body, suggesting negative regulation. Interestingly, a recent study showed that *atg3* could be trans-regulated by the lncRNA MSTRG.20486.1 in the midgut of *Bombyx mori* larvae infected with the cypovirus BmCPV [66]. This result suggests that either negative or positive regulation of *atg3* by lncRNAs following viral infection is a conserved mechanism in Lepidoptera. HdIV induces apoptosis in haemocytes and the fat body 24 h after infection of *S. frugiperda* larvae, allowing the elimination of the immune cells [35]. We found several pro-apoptotic genes upregulated in the HdIV-infected fat body, including *caspase-1* and *dronc*, two key cysteine proteases acting as initiator and executor of apoptosis, respectively. *Caspase-1* and *dronc* are co-expressed with several lncRNAs, in clusters 1 and 6, respectively. These lncRNAs are therefore good candidates for the regulation of these apoptotic genes.

Among the 11 clusters identified by the co-expression clustering analysis, cluster 5 grouped the largest number of HdIV transcripts, with three DE-lncRNAs and several genes associated with larval moulting. Polydnaviruses not only have immunosuppressive effects, preventing the parasitoid from being eliminated by the immune response, but they also affect the endocrine system, causing developmental arrest in the host [78]. Functional analyses have identified some viral products involved in this process. For example, two members of the *vankyrin* gene family produced by the polydnavirus TnBV, carried by the parasitoid *Toxoneuron nigriceps,* are involved in the disruption of the endocrine system of the parasitized host [79, 80]. Interestingly, of the 8 *vankyrin* genes identified in the HdIV genome, 6 were present among the 39 viral transcripts in cluster 5. It will therefore be interesting to evaluate whether the *vankyrins* in cluster 5 manipulate the endocrine system of *S. frugiperda* larvae, under the regulation of one or more of the three lncRNAs grouped in this cluster.

LncRNAs can function as primary miRNA precursors. Among the DE-lncRNAs identified in our study, we found three DE-lncRNAs harbouring miRNAs locus embedded in an intron or an exon, and potentially acting as precursors of miR-let-7–5 p, miR-970–5 p and miR-8–5 p. While the functions of the three miRs are not yet known in *S. frugiperda*, their roles have been examined in various insects. The role of miR-8–5 p in chitin biosynthesis during moulting has been well characterized in the hemipteran insect *N. lugens* [81]. The authors showed that, in response to 20-hydroxyecdysone (20E) signalling, miR-8–5 p expression decreases, while expression of its target gene, *trehalase-2*, the first enzyme involved in the chitin biosynthetic pathway, increases. Since JcDV- and HdIV-infected *S. frugiperda* larvae exhibit moulting defects, we can speculate that downregulation of *MSTRG.4700*, the precursor lncRNA of miR-8–5 p, is responsible of this alteration. A potential role in host–virus interaction has also been proposed for miR-970. In mosquitoes, miR-970 was upregulated in response to flavivirus infection, although its exact function in host–virus interaction is not known [82]. *MSTRG.8383*, the candidate precursor of miR-970, is upregulated in the fat body of JcDV-infected *S. frugiperda*. Thus, it would be interesting to develop future research to elucidate its function during JcDV infection and examine whether this function is conserved in insects. Let-7 was also found to be involved in moulting and metamorphosis in insects. In *B. mori*, BmLet-7 affects moulting by downregulating key regulatory factors in the ecdysone signalling cascade [83]. Several target genes of BmLet-7 have been predicted, including *Abrupt,* the negative regulator of ecdysone signalling pathway. In BmN4 cell lines treated with 20E, the expression of *Abrupt* decreased significantly following induction of the

primary transcript lncR17454 of let-7, suggesting that the latter is involved in the metamorphosis of *B. mori* [84]. In *S. frugiperda*, *MSTRG.98*, the potential precursor of let-7, is significantly downregulated in JcDV- and HdIV-infected fat body, while *Abrupt* was overexpressed in the same tissues. These results suggest that infection of larvae with HdIV and JcDV induces downregulation of the *MSTRG.98* leading to the upregulation of *Abrupt*, thereby altering the moulting process *in S. frugiperda* larvae.

## Conclusion

In summary, using high-throughput RNA sequencing, we compared the expression profiles of lncRNAs between two key immune tissues of *S. frugiperda* larvae, following infection with two viruses displaying a specific life cycle. A total of 1403 lncRNAs were significantly expressed in haemocytes and fat body, of which 529 were differentially expressed between virus-infected and uninfected control tissues. Our analysis revealed that some DE-lncRNAs have virus- and tissue-specific expression patterns. In order to infer the possible biological function of DE-lncRNAs, we combined different *in silico* analyses, including prediction of cis- and trans-targets of DE-lncRNAs and co-expressed gene cluster analysis. The results suggested that the DE-lncRNAs are associated with metabolic process, DNA replication, immune response, the moulting process and other pathways through their regulated target genes. Furthermore, we identified three virus-responsive lncRNAs hosting miRNAs known to be involved in the moulting process in other insects and virus replication. This suggests that the DE-lncRNAs take part in virus-induced pathogenesis through various pathways in *S. frugiperda*.

Overall, our results show that lncRNAs are important players in the interactions between host and viruses, and provide clues on their regulatory functions in viral pathogenesis in *S. frugiperda* larvae. Obviously, the functional importance of lncRNAs following virus infection would require further biological validation.

Funding information
This work was supported by the division Plant Health and Environment of National Research Institute for Agriculture, Food and Environment (INRAE).

Acknowledgements
We thank Gaetan Clabots and Raphaël Bousquet for rearing the *S. frugiperda* larvae and *H. didymator*. We would like to thank Cécile Clouet for providing the densovirus JcDV. We thank the quarantine insect platform (PIQ), a member of the Vectopole Sud network, for providing the infrastructure needed for pest insect experimentations. We also thank Philippe Clair from the qPHD platform (Montpellier GenomiX). We thank Emmanuelle d'Alençon for fruitful discussions about the identification of lncRNAs as a potential precursor of miRNAs.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References

1. Robin S, Legai F, Jouan V, Ogliastro M, Darboux I. Genome-wide identification of lncRNAs associated to viral infection in Spodoptera frugiperda. *Figshare* 2022.
2. Ma H, Han P, Ye W, Chen H, Zheng X, *et al*. The long noncoding RNA NEAT1 exerts antihantaviral effects by acting as positive feedback for RIG-I signaling. *J Virol* 2017;91:e02250-16.
3. Wang J, Zhang Y, Li Q, Zhao J, Yi D, *et al*. Influenza virus exploits an interferon-independent lncRNA to preserve viral RNA synthesis through stabilizing viral RNA polymerase PB1. *Cell Rep* 2019;27:3295–3304.
4. Wang P, Xu J, Wang Y, Cao X. An interferon-independent lncRNA promotes viral replication by modulating cellular metabolism. *Science* 2017;358:1051–1055.
5. Wang Z, Zhao Y, Zhang Y. Viral lncRNA: a regulatory molecule for controlling virus life cycle. *Noncoding RNA Res* 2017;2:38–44.
6. Tycowski KT, Guo YE, Lee N, Moss WN, Vallery TK, *et al*. Viral noncoding RNAs: more surprises. *Genes Dev* 2015;29:567–584.
7. Schuessler A, Funk A, Lazear HM, Cooper DA, Torres S, *et al*. West Nile virus noncoding subgenomic RNA contributes to viral evasion of the type I interferon-mediated antiviral response. *J Virol* 2012;86:5708–5718.
8. Tai-Schmiedel J, Karniely S, Lau B, Ezra A, Eliyahu E, *et al*. Human cytomegalovirus long noncoding RNA4.9 regulates viral DNA replication. *PLoS Pathog* 2020;16:e1008390.
9. Ruiz-Orera J, Messeguer X, Subirana JA, Alba MM. Long noncoding RNAs as a source of new peptides. *Elife* 2014;3:e03523.
10. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, *et al*. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* 2011;25:1915–1927.
11. Liu C, Bai B, Skogerbø G, Cai L, Deng W, *et al*. NONCODE: an integrated knowledge database of non-coding RNAs. *Nucleic Acids Res* 2005;33:D112–5.
12. Bu D, Yu K, Sun S, Xie C, Skogerbø G, *et al*. NONCODE v3.0: integrative annotation of long noncoding RNAs. *Nucleic Acids Res* 2012;40:D210–5.
13. Chen B, Zhang Y, Zhang X, Jia S, Chen S, *et al*. Genome-wide identification and developmental expression profiling of long noncoding RNAs during *Drosophila metamorphosis*. *Sci Rep* 2016;6:23330.
14. McCorkindale AL, Wahle P, Werner S, Jungreis I, Menzel P, *et al*. A gene expression atlas of embryonic neurogenesis in *Drosophila* reveals complex spatiotemporal regulation of lncRNAs. *Development* 2019;146:dev175265.
15. Schor IE, Bussotti G, Maleš M, Forneris M, Viales RR, *et al*. Noncoding RNA expression, function, and variation during *Drosophila Embryogenesis*. *Curr Biol* 2018;28:3547–3561.
16. Azlan A, Halim MA, Mohamad F, Azzam G. Identification and characterization of long noncoding RNAs and their association with acquisition of blood meal in *Culex quinquefasciatus*. *Insect Sci* 2021;28:917–928.
17. Liu F, Shi T, Qi L, Su X, Wang D, *et al*. lncRNA profile of *Apis mellifera* and its possible role in behavioural transition from nurses to foragers. *BMC Genomics* 2019;20:393.
18. Xiao H, Yuan Z, Guo D, Hou B, Yin C, *et al*. Genome-wide identification of long noncoding RNA genes and their potential association with fecundity and virulence in rice brown planthopper, *Nilaparvata lugens*. *BMC Genomics* 2015;16:749.
19. Bernabò P, Viero G, Lencioni V. A long noncoding RNA acts as a post-transcriptional regulator of heat shock protein (HSP70)

synthesis in the cold hardy *Diamesa tonsa* under heat shock. *PLoS One* 2020;15:e0227172.

20. Chen Y, Singh A, Kaithakottil GG, Mathers TC, Gravino M, *et al*. An aphid RNA transcript migrates systemically within plants and is a virulence factor. *Proc Natl Acad Sci* 2020;117:12763–12771.
21. Zhang S, Shen S, Yang Z, Kong X, Liu F, *et al*. Coding and non-coding RNAs: molecular basis of forest-insect outbreaks. *Front Cell Dev Biol* 2020;8:369.
22. Liu F, Guo D, Yuan Z, Chen C, Xiao H. Genome-wide identification of long non-coding RNA genes and their association with insecticide resistance and metamorphosis in diamondback moth, *Plutella xylostella*. *Sci Rep* 2017;7:15870.
23. Zhu B, Xu M, Shi H, Gao X, Liang P. Genome-wide identification of lncRNAs associated with chlorantraniliprole resistance in diamondback moth *Plutella xylostella* (L.). *BMC Genomics* 2017;18:380.
24. Etebari K, Furlong MJ, Asgari S. Genome wide discovery of long intergenic non-coding RNAs in Diamondback moth (*Plutella xylostella*) and their expression in insecticide resistant strains. *Sci Rep* 2015;5:14642.
25. Qiao H, Wang J, Wang Y, Yang J, Wei B, *et al*. Transcriptome analysis reveals potential function of long non-coding RNAs in 20-hydroxyecdysone regulated autophagy in *Bombyx mori*. *BMC Genomics* 2021;22:374.
26. Lopez-Ezquerra A, Mitschke A, Bornberg-Bauer E, Joop G. *Tribolium castaneum* gene expression changes after *Paranosema whitei* infection. *J Invertebr Pathol* 2018;153:92–98.
27. Jayakodi M, Jung JW, Park D, Ahn Y-J, Lee S-C, *et al*. Genome-wide characterization of long intergenic non-coding RNAs (lincRNAs) provides new insight into viral diseases in honey bees *Apis cerana* and *Apis mellifera*. *BMC Genomics* 2015;16:680.
28. Ali A, Abd El Halim HM. Re-thinking adaptive immunity in the beetles: evolutionary and functional trajectories of lncRNAs. *Genomics* 2020;112:1425–1436.
29. Azlan A, Obeidat SM, Theva Das K, Yunus MA, Azzam G. Genome-wide identification of *Aedes albopictus* long noncoding RNAs and their association with dengue and Zika virus infection. *PLoS Negl Trop Dis* 2021;15:e0008351.
30. Etebari K, Asad S, Zhang G, Asgari S. Identification of *Aedes aegypti* long intergenic non-coding RNAs and their association with Wolbachia and Dengue virus infection. *PLoS Negl Trop Dis* 2016;10:e0005069.
31. Zhang L, Xu W, Gao X, Li W, Qi S, *et al*. lncRNA sensing of a viral suppressor of RNAi activates non-canonical innate immune signaling in Drosophila. *Cell Host Microbe* 2020;27:115–128.
32. Guan R, Li H, Zhang H, An S. Comparative analysis of dsRNA-induced lncRNAs in three kinds of insect species. *Arch Insect Biochem Physiol* 2020;103:e21640.
33. Legeai F, Santos BF, Robin S, Bretaudeau A, Dikow RB, *et al*. Genomic architecture of endogenous ichnoviruses reveals distinct evolutionary pathways leading to virus domestication in parasitic wasps. *BMC Biol* 2020;18:89.
34. Dorémus T, Cousserans F, Gyapay G, Jouan V, Milano P, *et al*. Extensive transcription analysis of the Hyposoter didymator Ichnovirus genome in permissive and non-permissive lepidopteran host species. *PLoS One* 2014;9:e104072.
35. Visconti V, Eychenne M, Darboux I. Modulation of antiviral immunity by the ichnovirus HdIV in *Spodoptera frugiperda*. *Mol Immunol* 2019;108:89–101.
36. Barat-Houari M, Hilliou F, Jousset F-X, Sofer L, Deleury E, *et al*. Gene expression profiling of *Spodoptera frugiperda* hemocytes and fat body using cDNA microarray reveals polydnavirus-associated variations in lepidopteran host genes transcript levels. *BMC Genomics* 2006;7:160.
37. Provost B, Jouan V, Hilliou F, Delobel P, Bernardo P, *et al*. Lepidopteran transcriptome analysis following infection by phylogenetically unrelated polydnaviruses highlights differential and common responses. *Insect Biochem Mol Biol* 2011;41:582–591.
38. Clavijo G, Dorémus T, Ravallec M, Mannucci M-A, Jouan V, *et al*. Multigenic families in Ichnovirus: a tissue and host specificity study through expression analysis of vankyrins from *Hyposoter didymator* Ichnovirus. *PLoS One* 2011;6:e27522.
39. Mutuel D, Ravallec M, Chabi B, Multeau C, Salmon J-M, *et al*. Pathogenesis of *Junonia coenia* densovirus in *Spodoptera frugiperda*: a route of infection that leads to hypoxia. *Virology* 2010;403:137–144.
40. Legeai F, Gimenez S, Duvic B, Escoubas J-M, Gosselin Grenet A-S, *et al*. Establishment and analysis of a reference transcriptome for *Spodoptera frugiperda*. *BMC Genomics* 2014;15:704.
41. Gouin A, Bretaudeau A, Nam K, Gimenez S, Aury J-M, *et al*. Two genomes of highly polyphagous lepidopteran pests (*Spodoptera frugiperda*, *Noctuidae*) with different host-plant ranges. *Sci Rep* 2017;7:11816.
42. Poitout S, Bues R. Linolenic acid requirements of lepidoptera Noctuidae Quadrifinae Plusiinae: *Chrysodeixis chalcites* Esp., *Autographa gamma* L.' *Macdunnoughia confusa* Stph., *Trichoplusia ni* Hbn. reared on artificial diets. *Ann Nutr Aliment* 1974;28:173–187.
43. Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, *et al*. The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol* 2020;38:276–278.
44. Legeai F, Derrien T. Identification of long non-coding RNAs in insects genomes. *Curr Opin Insect Sci* 2015;7:37–44.
45. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 2014;30:923–930.
46. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;26:139–140.
47. Alexa A JR. (2021) TopGO: enrichment analysis for gene ontology. version 2.48.0. ed2022. p. R package.
48. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol* 2021;38:5825–5829.
49. Rau A, Maugis-Rabusseau C. Transformation and model choice for RNA-seq co-expression analysis. *Brief Bioinform* 2018;19:425–436.
50. Wucher V, Legeai F, Hédan B, Rizk G, Lagoutte L, *et al*. FEELnc: a tool for long non-coding RNA annotation and its application to the dog transcriptome. *Nucleic Acids Res* 2017;45:e57.
51. Le Béguec C, Wucher V, Lagoutte L, Cadieu E, Botherel N, *et al*. Characterisation and functional predictions of canine long non-coding RNAs. *Sci Rep* 2018;8:13444.
52. Moné Y, Nhim S, Gimenez S, Legeai F, Seninet I, *et al*. Characterization and expression profiling of microRNAs in response to plant feeding in two host-plant strains of the lepidopteran pest *Spodoptera frugiperda*. *BMC Genomics* 2018;19:804.
53. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;26:841–842.
54. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001;25:402–408.
55. Dumas B, Jourdan M, Pascaud AM, Bergoin M. Complete nucleotide sequence of the cloned infectious genome of *Junonia coenia* densovirus reveals an organization unique among parvoviruses. *Virology* 1992;191:202–222.
56. Gimenez S, Abdelgaffar H, Goff GL, Hilliou F, Blanco CA, *et al*. Adaptation by copy number variation increases insecticide resistance in the fall armyworm. *Commun Biol* 2020;3:664.
57. Hess AM, Prasad AN, Ptitsyn A, Ebel GD, Olson KE, *et al*. Small RNA profiling of Dengue virus-mosquito interactions implicates the PIWI RNA pathway in anti-viral defense. *BMC Microbiol* 2011;11:45.
58. Léger P, Lara E, Jagla B, Sismeiro O, Mansuroglu Z, *et al*. Dicer-2- and Piwi-mediated RNA interference in Rift Valley fever virus-infected mosquito cells. *J Virol* 2013;87:1631–1648.

59. Santos D, Verdonckt T-W, Mingels L, Van den Brande S, Geens B, *et al*. PIWI proteins play an antiviral role in lepidopteran cell lines. *Viruses* 2022;14:1442.
60. Salasc F, Mutuel D, Debaisieux S, Perrin A, Dupressoir T, *et al*. Role of the phosphatidylinositol-3-kinase/Akt/target of rapamycin pathway during ambidensovirus infection of insect cells. *J Gen Virol* 2016;97:233–245.
61. Pruijssers AJ, Falabella P, Eum JH, Pennacchio F, Brown MR, *et al*. Infection by a symbiotic polydnavirus induces wasting and inhibits metamorphosis of the moth *Pseudoplusia* includens. *J Exp Biol* 2009;212:2998–3006.
62. Cusson M, Laforge M, Miller D, Cloutier C, Stoltz D. Functional significance of parasitism-induced suppression of juvenile hormone esterase activity in developmentally delayed *Choristoneura fumiferana* larvae. *Gen Comp Endocrinol* 2000;117:343–354.
63. Wolfe CJ, Kohane IS, Butte AJ. Systematic survey reveals general applicability of "guilt-by-association" within gene coexpression networks. *BMC Bioinformatics* 2005;6:227.
64. Scherfer C, Tang H, Kambris Z, Lhocine N, Hashimoto C, *et al*. Drosophila Serpin-28D regulates hemolymph phenoloxidase activity and adult pigmentation. *Dev Biol* 2008;323:189–196.
65. Dykes IM, Emanueli C. Transcriptional and post-transcriptional gene regulation by long non-coding RNA. *Genomics, Proteomics & Bioinformatics* 2017;15:177–186.
66. Zhang Z, Zhao Z, Lin S, Wu W, Tang W, *et al*. Identification of long noncoding RNAs in silkworm larvae infected with *Bombyx mori* cypovirus. *Arch Insect Biochem Physiol* 2021;106:1–12.
67. Mayer KA, Stöckl J, Zlabinger GJ, Gualdoni GA. Hijacking the supplies: metabolism as a novel facet of virus-host interaction. *Front Immunol* 2019;10:1533.
68. Weitzman MD, Fradet-Turcotte A. Virus DNA replication and the host DNA damage response. *Annu Rev Virol* 2018;5:141–164.
69. Chevignon G, Periquet G, Gyapay G, Vega-Czarny N, Musset K, *et al*. Cotesia congregata *Bracovirus* circles encoding *PTP* and Ankyrin genes integrate into the DNA of parasitized *Manduca sexta* hemocytes. *J Virol* 2018;92:e00438-18.
70. Muller H, Chebbi MA, Bouzar C, Périquet G, Fortuna T, *et al*. Genome-wide patterns of *Bracovirus* chromosomal integration into multiple host tissues during parasitism. *J Virol* 2021;95:e0068421.
71. Wang Z, Ye X, Zhou Y, Wu X, Hu R, *et al*. Bracoviruses recruit host integrases for their integration into caterpillar's genome. *PLoS Genet* 2021;17:e1009751.
72. Wang Z-H, Zhou Y-N, Yang J, Ye X-Q, Shi M, *et al*. Genome-wide profiling of *Diadegma semiclausum* ichnovirus integration in parasitized *Plutella xylostella* hemocytes identifies host integration motifs and insertion sites. *Front Microbiol* 2020;11:608346.
73. Kerr AG, Wang Z, Wang N, Kwok KHM, Jalkanen J, *et al*. The long noncoding RNA ADIPINT regulates human adipocyte metabolism via pyruvate carboxylase. *Nat Commun* 2022;13:2958.
74. Feldstein O, Nizri T, Doniger T, Jacob J, Rechavi G, *et al*. The long non-coding RNA ERIC is regulated by E2F and modulates the cellular response to DNA damage. *Mol Cancer* 2013;12:131.
75. Valanne S, Salminen TS, Järvelä-Stölting M, Vesala L, Rämet M. Immune-inducible non-coding RNA molecule lincRNA-IBIN connects immunity and metabolism in *Drosophila melanogaster*. *PLoS Pathog* 2019;15:e1007504.
76. Zhou H, Li S, Pan W, Wu S, Ma F, *et al*. Interaction of *lncRNA-CR33942* with Dif/Dorsal facilitates antimicrobial peptide transcriptions and enhances *Drosophila* toll immune responses. *J Immunol* 2022;208:1978–1988.
77. Zhou H, Wu S, Liu L, Li R, Jin P, *et al*. *Drosophila* relish activating *lncRNA-CR33942* transcription facilitates antimicrobial peptide expression in imd innate immune response. *Front Immunol* 2022;13:905899.
78. Lanzrein B, Treiblmayr K, Meyer V, Pfister-Wilhelm R, Grossniklaus-Bürgin C. Physiological and endocrine changes associated with polydnavirus/venom in the parasitoid-host system Chelonus inanitus-*Spodoptera littoralis*. *J Insect Physiol* 1998;44:305–321.
79. Valzania L, Romani P, Tian L, Li S, Cavaliere V, *et al*. A polydnavirus ANK protein acts as virulence factor by disrupting the function of prothoracic gland steroidogenic cells. *PLoS One* 2014;9:e95104.
80. Ignesti M, Ferrara R, Romani P, Valzania L, Serafini G, *et al*. A polydnavirus-encoded ANK protein has a negative impact on steroidogenesis and development. *Insect Biochem Mol Biol* 2018;95:26–32.
81. Chen J, Liang Z, Liang Y, Pang R, Zhang W. Conserved microRNAs miR-8-5p and miR-2a-3p modulate chitin biosynthesis in response to 20-hydroxyecdysone signaling in the brown planthopper, *Nilaparvata lugens*. *Insect Biochem Mol Biol* 2013;43:839–848.
82. Skalsky RL, Vanlandingham DL, Scholle F, Higgs S, Cullen BR. Identification of microRNAs expressed in two mosquito vectors, *Aedes albopictus* and *Culex quinquefasciatus*. *BMC Genomics* 2010;11:119.
83. Ling L, Ge X, Li Z, Zeng B, Xu J, *et al*. MicroRNA Let-7 regulates molting and metamorphosis in the silkworm, *Bombyx mori*. *Insect Biochem Mol Biol* 2014;53:13–21.
84. Fu Y, Wang Y, Huang Q, Zhao C, Li X, *et al*. Long noncoding RNA lncR17454 regulates metamorphosis of silkworm through let-7 miRNA cluster. *J Insect Sci* 2022;22:12.