Parasitology

cambridge.org/par

Deciphering the behaviour manipulation imposed by a virus on its parasitoid host: insights from a dual transcriptomic approach

J. Varaldi and D. Lepetit

Laboratoire de Biométrie et Biologie Evolutive (UMR CNRS 5558), University Lyon 1 – University of Lyon, 43 boulevard du 11 novembre 1918, 69622 Villeurbanne cedex, France

Research Article

Cite this article: Varaldi J, Lepetit D (2018). Deciphering the behaviour manipulation imposed by a virus on its parasitoid host: insights from a dual transcriptomic approach. *Parasitology* **145**, 1979–1989. https://doi.org/10.1017/S0031182018000835

Received: 5 October 2017
Revised: 28 March 2018
Accepted: 23 April 2018
First published online: 28 May 2018

Key words:
Behaviour manipulation; GABA; odorant receptor; potassium channel; RNAseq; superparasitism; symbiosis

Author for correspondence:
J. Varaldi, E-mail: julien.varaldi@univ-lyon1.fr

Abstract

Behaviour manipulation imposed by parasites is a fascinating phenomenon but our understanding is still very limited. We studied the interaction between a virus and the parasitic wasp *Leptopilina boulardi* that attacks *Drosophila* larvae. Wasps usually refrain to lay eggs into already parasitized hosts (superparasitism avoidance). On the contrary, females infected by the Leptopilina boulardi Filamentous Virus (LbFV) are much more incline to superparasitize. Interestingly, the host-sharing induced by this behaviour modification leads to the horizontal transmission of the virus, thus increasing its fitness at the expense of that of the wasp. To better understand the mechanisms underlying this behaviour manipulation, we studied by RNA sequencing the meta-transcriptome of LbFV and the parasitic wasp both in the abdomen and in the head. We found that the abundance of viral transcripts was independent of the wasp strain but strongly differed between tissues. Based on the tissue pattern of expression, we identified a set of 20 viral genes putatively involved in the manipulation process. In addition, we identified a set of wasp genes deregulated in the presence of the virus either in the abdomen or in the head, including genes with annotations suggesting involvement in behaviour (i.e. Potassium-channel protein). This dataset gives new insights into the behaviour manipulation and on the genetic basis of superparasitism in parasitoids.

Introduction

In innumerable situations, organisms have to take a decision that may have consequences on their reproductive success. At principle, we may believe that these decisions are taken by the focal individual, and may have been optimized by natural selection (Charnov, 1976). However, it is now understood that the phenotype of individuals may also be affected by their symbiotic partners (we refer here to symbiosis as any type of interaction ranging from parasitism to mutualism, following its initial definition by De Bary, 1879). In particular, when the symbiont has the opportunity to be transmitted independently of the host genes through horizontal transmission, natural selection may favour the evolution of manipulation strategies: in such cases, the symbiont may force the host to take inappropriate decisions that favours its own transmission. For instance, the parasite *Toxoplasma gondii* is able to transform the innate aversion of its rodent host for cat odor into an attraction for cat odors, thus favouring its trophic (horizontal) transmission to its definitive host (Berdoy *et al.*, 2000). A recent paper also shows how the bacterial gut microbiota modifies the food choice decisions taken by *Drosophila* towards more carbohydrates and less amino acids (Leitão-Gonçalves *et al.*, 2017), raising the fascinating possibility that microbiota may manipulate host alimentary choices.

In spite of the long list of host–parasite systems where parasitic behaviour manipulation has been observed, the molecular mechanisms allowing parasites to manipulate the behaviour of their hosts are poorly known. To our knowledge, parasites genes directly or indirectly involved in the behavioural manipulation have been unequivocally identified only in a few systems involving baculoviruses responsible for the tree-top disease of their caterpillar hosts: when the caterpillar is infected, it climbs to the top of the plant where it eventually dies and liquefies, thus releasing its viral particles. This behaviour modification is expected to enhance virus dissemination either by directly increasing virus spread to new larvae or through the mediation of birds that can transport viruses over long distances (reviewed in Han *et al.*, 2015). It has been shown that a viral ecdysteroid glycosyl transferase (egt) is involved in tree-top disease (Hoover *et al.*, 2011; Ros *et al.*, 2015), which could be achieved directly or indirectly *via* prolonging the time to death (Han *et al.*, 2015). Apart from this example, the molecular determinants involved in such fascinating phenomenon are far from being identified, and this deserves further investigation in other systems.

Due to their peculiar lifestyle, parasitoids have frequently been used to study the optimality of foraging decisions, because of the tight link between foraging efficiency and fitness in these organisms (Godfray, 1994). In particular, parasitoids may encounter already parasitized hosts in their environment and have to decide whether or not to accept such hosts of lower value.

In most conditions, parasitoid females refuse to lay additional eggs into an already parasitized hosts. This decision makes sense since the second parasitoid egg is usually at a

© Cambridge University Press 2018

CAMBRIDGE UNIVERSITY PRESS

competitive disadvantage in 'superparasitized' hosts. However, females do superparasitize on some occasions. This paradoxical decision was first interpreted as being the consequence of 'errors' on the part of the egg-laying females (Van Lenteren and Bakker, 1975) but was later re-interpreted as an adaptive decision when females have no better solution around, i.e. when the overall environment quality is low. This last observation led researchers to formulate an adaptive interpretation of this behaviour: laying a second egg in a host may still pay if this egg has a non-nul probability to win the competition for the possession of the host, which is the case. Several experiments and theoretical models, in particular, done using *Drosophila* parasitoids as a model system, globally validated this interpretation (Van Alphen and Visser, 1990; Gandon *et al.*, 2006).

Surprisingly, however, we found that superparasitism decision is under the control of an inherited DNA virus in the parasitoid of *Drosophila Leptopilina boulardi*. The virus, called LbFV for Leptopilina boulardi Filamentous Virus, forces the females to accept already parasitized hosts, contrary to their uninfected counterparts. Interestingly, the host sharing of infected and uninfected parasitoids leads to the horizontal transmission of the virus (Varaldi *et al.*, 2003). Thus, the virus directly benefits from the behaviour alteration it induces, since superparasitism situations allow it to colonize new uninfected parasitoid lineages or to replace a pre-existing viral strain (Martinez *et al.*, 2015). In addition to horizontal transmission under superparasitism, the virus is vertically transmitted (from mother to offspring) with very high efficiency (although <100%) and reaches high prevalence (~95%) in natural populations (Patot *et al.*, 2010). This is consistent with the observation that LbFV is relatively abundant in the ovaries of the wasp, as attested by microscopy investigations (Varaldi *et al.*, 2006). The effect of the virus is mostly restricted to superparasitism decision. For instance, infected females are still able to discriminate between good hosts species and bad hosts species (Varaldi *et al.*, 2009) in spite of the fact that this decision probably involves similar processes: it is believed that the wasp senses its host using chemoreceptors that are present on its ovipositor and other organs (Van Lenteren *et al.*, 2007; Ruschioni *et al.*, 2015), integrate it into the central nervous system and 'decide' whether or not to accept such host. Finally, LbFV also brings a slight protection for the wasp egg against the immune reaction of some *Drosophila* strains (Martinez *et al.*, 2012*b*), underlying the multidimensionality of the relationship between the virus and the wasp. The virus has a large circular genome (~110 kb) containing 108 predicted ORFs, most of them having no homologs in public databases. However, LbFV is related, although very distantly, to Hytrosaviridae, and represents a possible new virus family (Lepetit *et al.*, 2016).

In order to decipher the molecular basis of the behaviour manipulation, we studied both the transcriptome of the virus and that of the wasp in response to infection. Our final aim is to identify the effector genes in the virus genome and the targeted genes in the wasp genome. Because we could not rely on the annotation to identify candidate effector genes involved in the behaviour manipulation, we decided to define 2 non-exclusive transcript abundance patterns that would suggest that a viral gene is involved in the behaviour manipulation. We hypothesized that the virus may manipulate the wasp behaviour either by interfering at the level of the abdomen (disturbing information intake possibly at the level of the ovipositor) and/or at the level of the central nervous system (disturbing the final oviposition decision after proper information intake). Our rational was that a viral transcript involved in the behaviour manipulation (effector) should be (i) present in adult females because this is where and when manipulation takes place and (ii) either have an abdomen-skewed abundance (if it acts on information uptake) or (iii) have a head-skewed abundance (if it acts on information processing in the central nervous system). In other words, viral transcripts that would not be present in adult female, or viral transcripts that would be equally abundant in both tissues (like we may expect for instance for apoptotic inhibitors) are excluded from the candidate list. Based on the same rationale, wasp genes that are targeted by the effectors of the virus should be specifically deregulated (either up or downregulated) by the virus in the tissue in which manipulation takes place. In other words, wasp genes that would be differentially expressed in response to infection in the same direction in both tissues are excluded from our target list since they may correspond to immunity genes rather than manipulation genes.

To define these effectors and targets lists, we studied the pattern of expression of both viral and wasp genes in the abdomen and the head of young adult females through RNA sequencing. This work was performed on 4 wasp genetic backgrounds, either infected by the same viral strain or uninfected since we knew from previous work that the manipulation intensity may vary according to wasp genotype (Martinez *et al.*, 2012*a*).

Materials and methods

Wasp strains and transfer of LbFV

Leptopilina boulardi strains originating from France (StFoy near Lyon, Lb84, near Avignon), Italy (Sienna) and Ivory Coast (Lamto, G495) were used in the experiments. The virus LbFV was isolated from wasps originating from Gotheron (near Valence, France) and was transferred into the 4 wasp genetic backgrounds by means of superparasitism taking benefit of the fact that superparasitism conditions allow for the horizontal transmission of the virus. As *L. boulardi* is a haplo-diploid species, fertilized eggs develop into females whereas unfertilized ones develop into males. Hence, virus horizontal transmission can easily be monitored by obtaining *Drosophila* hosts successively parasitized by a mated mother (recipient line) and a virgin infected mother (donor line). As females emerging from these hosts are necessarily offspring of the mated mother, the occurrence of horizontal transmission is simply tested by polymerase chain reaction (PCR) detection of the virus in those females. In practice, one mated parasitoid female (1- to 2-day-old) of the recipient line of interest was placed with 100 *Drosophila* eggs for 24 h in a rearing vial. Then, the female was removed and replaced by 4 virgin females (2- to 3- day-old) of the donor line of interest for an additional 24 h. The success of horizontal transmission was then evaluated using PCR test on emerging females (see Patot *et al.*, 2009 for details). For each vial, mothers of the recipient and donor lines were PCR checked for viral infection. This setup led to the production of 4 wasp genetic backgrounds either infected or not by the same virus isolate (4 pairs of strains). Rearing and experiments were performed under a 12L:12D photoperiod at 25 ± 1 °C, 60% humidity, using as hosts a *D. melanogaster* strain originating from Sainte-Foy-lès-Lyon (near Lyon, France) fed with a standard diet.

Phenotypic tests

The effect of the virus LbFV on superparasitism behaviour was measured according to a standard procedure: 1 or 2 days-old females were placed in a Petri dish containing 10 first instar *D. melanogaster* larvae on an agar layer with live yeast. 48 h later, 4 hosts from each Petri dish were dissected to count for the number of parasite eggs/larvae. The superparasitism phenotype of each female was measured as the mean number of eggs laid per parasitized larva.

https://doi.org/10.1017/S0031182018000835 Published online by Cambridge University Press

Library preparation for RNAseq experiment

Total RNA was extracted using the RNeasy kit (Qiagen). The 2–3 days old adult females were dissected under the microscope and pools of 20–30 heads or 20–30 abdomens were homogenized in the lyzing buffer following the RNeasy kit instructions (Qiagen) for total RNA extraction. Library preparation was performed by Eurofins using regular Illumina protocol for paired-end reads (size selection 150–400 bp). The cDNAs were sequenced on an Illumina HiSeq 2000/2500 using the high-output run mode. We obtained between 7 and 24 M of paired-end reads (2 × 100 bp) per biological sample (mean = 15 M). Two biological replicates were performed for each condition (4 wasp genotypes × 2 infection status × 2 tissues × 2 replicates = 32 samples). The raw RNAseq reads have been deposited on ncbi (SRA accession SRP128635).

Bioinformatic pipeline for RNAseq data

The reads were mapped against a draft of the genome of *L. boulardi,* strain Sienna complemented by the genomic sequence of LbFV (Lepetit *et al.*, 2016) This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession PQAT00000000. The version described in this paper is version PQAT01000000. The wasp draft has been obtained after assembling genomic DNA reads (Illumina paired-end reads 2 × 100 bp, 46× coverage) with IDBA_ud (Peng *et al.*, 2012) followed by scaffolding using assembled RNAseq data by running the software L_RNA_scaffolder (Xue *et al.*, 2013). The assembly obtained was judged sufficiently good to analyse gene expression, since we were able to find 1044 complete genes out of 1066 ubiquitous eukaryotic genes (more precisely genes that are shared by at least 90% of eukaryotes) that we looked for using the pipeline BUSCO (Simão *et al.*, 2015). Four genes were detected as duplicated and 8 as fragmented. Only 10 were considered as missing (<1%) which is in the order of magnitude of apparent missing genes for excellent assemblies such as the human genome assembly. The present RNAseq reads were mapped using Hisat2 v2.2.1 (Kim *et al.*, 2015) with default parameters; then cufflinks v2.2.1 (Trapnell *et al.*, 2010) was used to generate a gtf file for each biological sample. The proportion of reads successfully mapped on the genome was between 95 and 97% for all samples, except for the 4 head samples of Lb84 (~85%) raising the possibility that this strain could be too divergent to the other 3 strains to allow a good estimation of its gene expression. However, the fact that the 4 abdomen samples of the same strain (Lb84) did behave similarly compared with other samples (~96% mapping success) ruled out this possibility. In addition, we calculated a synonymous distance between all 4 strains using the BUSCO gene set. Lb84 was neither particularly distant from Sienna (used for the genome sequencing) nor to other strains (Fig. S1) confirming our conclusion. All cufflinks gtf files obtained for the 32 samples were merged together using cuffmerge v2.2.1 to identify wasp genes in a unified gtf file. Ht-seq v0.6.0 (Anders *et al.*, 2015) was then used to count the number of reads corresponding to each gene using this unified gtf file complemented by the virus gtf file deduced from the *de novo* prediction presented in Lepetit *et al.* (2016). Default parameters were used. In total 16 679 wasp genes were identified, among which 16 023 were mapped by more than 2 reads in the head and 14 518 in the abdomen. The virus genome contains 108 putative genes. From this analysis, we generated a global count table containing both the viral and wasp loci for each sample (16 787 × 32 table). The wasp assembled transcripts have been submitted to ncbi (tsa submission number: GGGI01000000).

Statistical analysis for differential expression

The count table was analysed using the DESeq2 R package (Love *et al.*, 2014) which is based on the negative binomial distribution of count data. In order to take into consideration the differences in library sizes between the samples, we estimated the 'size factor' on the whole table, before splitting the dataset in 2 sub-tables corresponding to viral genes (108 × 32) and wasp genes (16 679 × 32).

For the virus transcripts abundance table, we discarded uninfected samples and included both the wasp genetic background and the tissue in a linear model (design = formula (~genotype + tissue)). We successively tested for the genotype and tissue effect.

For the wasp dataset, we were interested in identifying wasp genes that would be specifically deregulated by the virus presence either in the head or the abdomen. For this purpose, we did 2 distinct analysis. For the first analysis, we split the dataset into head and abdomen datasets. For each separate analysis (head, abdomen), we included both the genotype of the wasp and its infection status in a linear model (design = formula (~genotype + infection)). We then considered genes as valid candidates if there was a significant 'infection' effect either in the head or in the abdomen (but not when the 'infection' effect was detected in both head and abdomen). In the second analysis, we analysed the whole dataset in a unique linear model (design = formula (~genotype + tissue + infection + tissue:genotype + genotype:infection + tissue:infection)). We then tested for a significant tissue:infection interaction that would indicate that the infection effect is different across tissues (after accounting for baseline differences, accounting for each tissue having a genotype-specific baseline and for infection having a genotype-specific baseline). After visual inspection of their expression profile candidates identified by this second approach were either classified as 'abdomen-altered' or 'head-altered' or discarded. We filtered-out the genes showing clear up-regulation in one tissue and downregulation in the other. We thus selected genes that are considered as specifically altered in one tissue. The final candidates obtained from the second analysis were merged with the candidates identified in the first analysis for further analysis and discussion.

To test multi-level or interaction factors (i.e. genotype or the interaction tissue:infection), we compared the likelihood of the complete and reduced model using LRT tests (nbinomLRT function). Factors with only 2 levels (i.e. infection) were tested using classical wald tests. Virus genes were considered as differentially expressed between factor levels when the corrected *P* value (using the Benjamini and Hochberg method) were below 0.05. In the first analysis, wasp genes were considered as conclusively differentially expressed in response to virus presence when the corrected *P* value (using the Benjamini and Hochberg method) associated with the virus presence was below 0.05 and when the differential of expression was consistent among all 4 wasp genetic backgrounds (measured by the sign of the log2foldChange). For the second analysis, genes with a significant interaction term tissue:infection were considered in the next step. For all plots, we calculated the statistic 'Transcripts per Million' (TPM) which has the property of summing up to 1 million for each library. It estimates the number of transcripts for a particular gene out of 1 million transcripts in the tissue.

Annotation of the wasp transcripts

For each locus identified by cufflinks, we took the first transcript reported by cuffmerge and blasted (blastx) it against a local refseq protein database downloaded the 26/02/2018. The output of the blast (evalue threshold 10e-5) was used to feed the blast2GO software using default parameters. This gave us a crude list of GO terms associated with each locus. In addition, we performed an *in silico* translation of each mRNA using transdecoder. The predicted proteins were scanned for conserved domains by searching the pfam database using hmmscan v3.1b2 (Finn *et al.*, 2016) and

https://doi.org/10.1017/S0031182018000835 Published online by Cambridge University Press

transmembrane domains were searched for using the tmhmm webserver v2.0 (http://www.cbs.dtu.dk/services/TMHMM/).

Results

Phenotype

In order to study properly the effect of the virus on wasp gene expression, we first infected 4 parasitoid genetic backgrounds (Italy, France, Ivory Coast) with the same LbFV isolate (isolated from wasps captured in Valence, France). This gave us 4 pairs of uninfected and infected strains with similar genetic backgrounds. Two generations after the horizontal transfer of the virus, we tested the phenotype of all strains, either infected or not. As expected, the virus induced the typical superparasitism behaviour in the reference wasp strain (Sienna origin, Italy) used in most of our previously published experiments. Similarly, the virus induced strong superparasitism tendency in the other 3 genetic backgrounds (Fig. 1). However, some wasp genetic backgrounds were more prone to the behaviour manipulation than others (StFoy *vs* G495, pairwise *t*-test with Bonferroni correction: $P = 0.026$), which was expected from previous results (Martinez *et al.*, 2012*a*). Importantly, wasps lay only one egg per oviposition meaning that the strains differ in their decision to accept already parasitized hosts for oviposition rather than by the clutch size per oviposition (Varaldi *et al.*, 2005).

A two-sided transcriptomic analysis

We studied the transcripts abundance of both viral and wasp genes in all 4 genetic backgrounds in the presence or absence of the virus. This was achieved by mapping the RNAseq reads of each library on a draft-assembly of *L. boulardi* genome complemented with the genome of LbFV (see methods).

Abundance of LbFV transcripts in wasp tissues: searching for effectors of the manipulation

We hypothesized that the virus manipulates the behaviour of the wasp either by acting at the level of the abdomen, possibly interfering with the acquisition of information from its host, or at the level of the central nervous system, possibly interfering with the decision after correct information uptake. Based on this rationale, our aim was to identify virus genes with particular abundance pattern: either particularly abundant in the abdomen, or in the head. We thus compared the viral transcripts abundance in the head and in the abdomen of the wasps.

The total abundance of LbFV transcripts (proxied by the cumulative TPM, transcripts per million) was much higher in the abdomen compared with the head, whatever the wasp genetic background ($F_{(1,8)} = 192.5$, $P < 10^{-7}$). The same result is observed for raw number of reads mapped to the LbFV genome (mean number ± standard error: Abdomen = 3880 ± 662; head = 527 ± 42). This indicates that the RNA corresponding to LbFV genes are more abundant in the abdomen compared with the head, possibly due to higher viral density and/or higher overall expression of LbFV genes in the abdomen. In the abdomen, the whole set of viral transcripts represented on average 0.14% of all transcripts in cells of *L. boulardi* (min = 0.08%, max = 0.19%, var = 0.0016). This percentage is much lower in the head (average = 0.03%, min = 0.02%, max = 0.04%, var = 5.7×10^{-5}). The virus transcripts thus represent a very tiny proportion of the transcripts in the cell in the range of the fraction represented by baculovirus transcripts at the very beginning of infection (~1 h post-infection, ~0.1%), but is of orders of magnitude lower if we compare with the proportion of viral transcripts found in a baculovirus-infected cell at late infection (~80%, Chen *et al.*, 2013). Out of the 108 putative LbFV ORFs, we did not detect any transcripts in any condition for 13 of them. Note that since the viral genome density may differ between tissues, and possibly between wasp strains, we do not know whether the observed differences in viral transcripts abundance arise from differences in viral genome load and/or from differences in transcriptional activity.

Wasp strain effect

In spite of the phenotypic differences observed among wasp genetic background regarding the intensity of the behaviour manipulation, we did not detect any 'wasp strain' effect on the transcripts abundance for any LbFV genes. Accordingly, the clustering performed on the columns (samples) grouped together abdomen samples on one side and head samples on the other side irrespective of the wasp strain (Fig. 2).

Tissue effect

The overall viral transcripts abundance pattern was clearly different in the head and the abdomen: 17 transcripts displayed significant abdomen-skewed abundance profiles whereas 3 transcripts showed significant head-skewed profiles (all P values <0.05 after correction for multiple tests, Fig. 2, Table S1).

ORF89 was the most abundant LbFV transcript in the abdomen (on average TPM = 1146) but was much less abundant in the head (TPM = 6.75). It represented on average 78% of all LbFV transcripts in the abdomen and only 2% in the head. Consistently with this result, this transcript was the major constituent of a subtracted library of cDNA that we did previously to compare infected and uninfected strains (Patot *et al.*, 2009). The strong abundance of ORF89 transcripts in the abdomen explained most of the difference in LbFV transcripts total abundance between the head and the abdomen: the total abundance of all LbFV transcripts was approximately 4 times higher in the abdomen compared with the head ($F_{(1,8)} = 192.5$, $P < 10^{-7}$) and this effect cancelled out when we removed ORF89 from the analysis ($F_{(1,8)} = 0.035$, $P = 0.85$). ORF89 is predicted to contain a signal peptide, suggesting that it is excreted. In total, 17 viral transcripts were over-represented in the abdomen compared

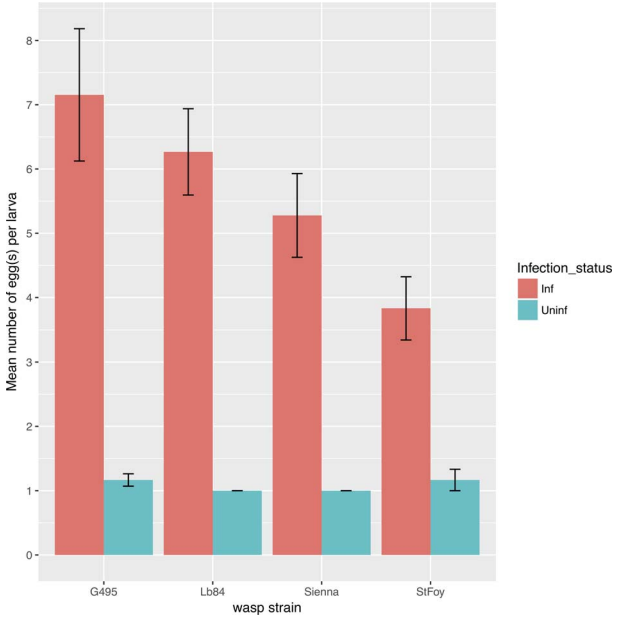


Fig. 1. Phenotype of the wasp strains used in the experiments. Error bars indicate standard errors.

https://doi.org/10.1017/S0031182018000835 Published online by Cambridge University Press

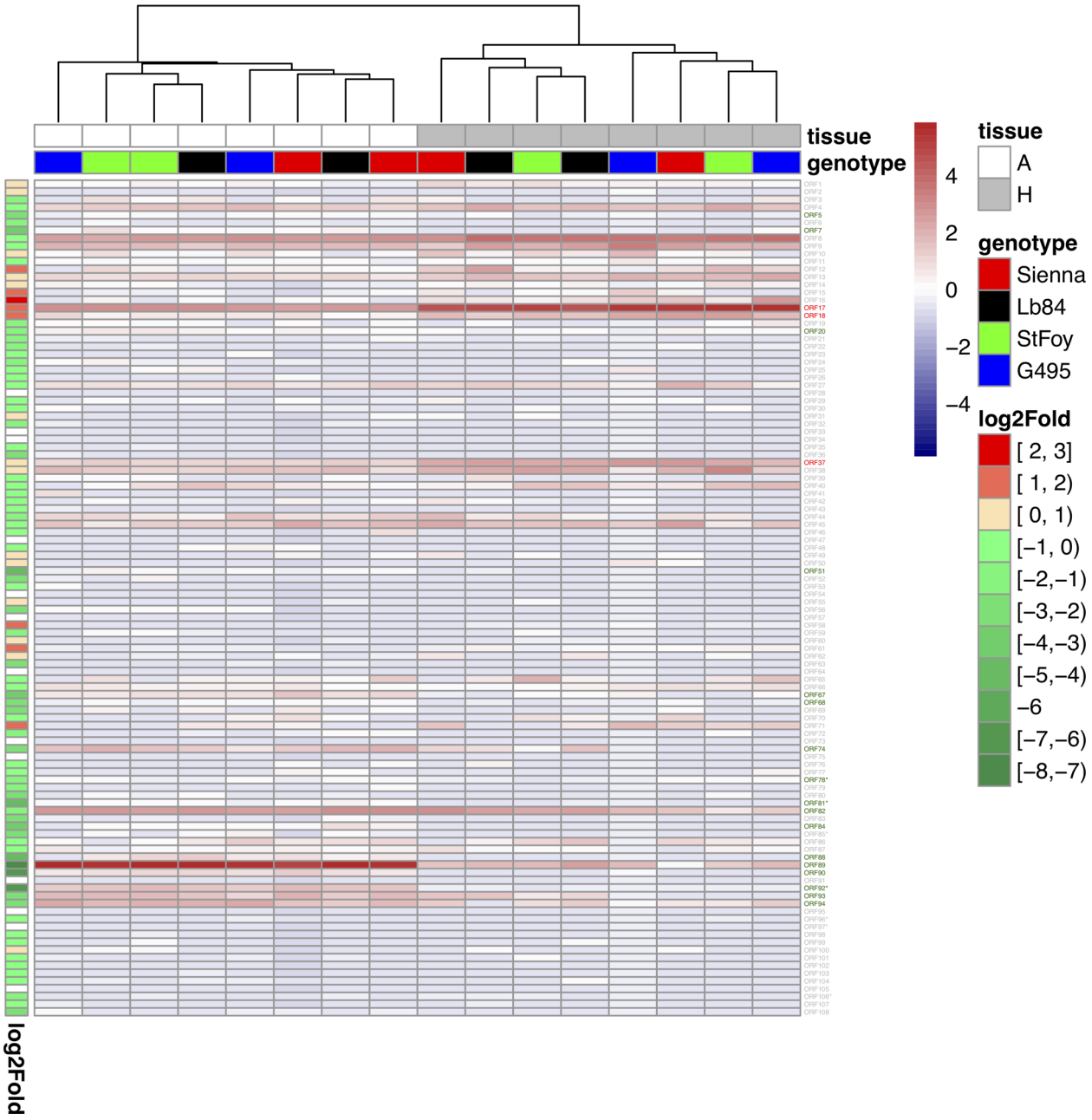


Fig. 2. Heatmap of virus transcripts abundance measured as transcripts per million (TPM) in head and abdomen of *L. boulardi*. ORF are ordered according to their position in the virus genome. The samples have been clustered in an unsupervised manner using the default algorithm of the pheatmap R function (complete linkage method on Euclidian distance). ORFs written in green have a significant abdomen-skewed expression profile whereas ORFs written in red have a significant head-skewed expression profile (corrected *P* values<0.05). TPM values were centered and reduced according to samples in order to increase readability. This remove the high variance among samples on the overall expression (especially between head and abdomens), which better illustrate the relative expression profile of the virus genes in these tissues. Log2Fold changes are represented on the left scale with levels of red indicating over-expression in the head compared with the abdomen and levels of green indicating the opposite.

with the head (Fig. 3A) but they have no homologs in public databases. The only exception is ORF68 which resembles a nuclease/helicase (Psi-blast, data not shown).

Three LbFV transcripts were significantly over-represented in the head compared with the abdomen (Fig. 3B): ORF17, ORF18 and ORF37 (*P* values wald test: 7×10^{-17}, 3×10^{-4}, 0.01, respectively). ORF37 contains a hydrolase domain (NUDIX). Nudix hydrolases are found in all classes of organism and hydrolyse a wide range of organic pyrophosphates, including nucleoside di- and triphosphates, dinucleoside and diphosphoinositol polyphosphates, nucleotide sugars and RNA caps, with varying degrees of substrate specificity. Finally, the 2 ORFs that show similarities with inhibitors of apoptosis have similar transcript abundance in both tissues (Fig. 3C).

Expression of wasp genes: searching for targets of the manipulation

A total of 16 679 wasp loci were identified on the whole RNAseq dataset. Within each tissue, most of the variance in wasp gene expression was explained by the geographic origin of the wasp with the 2 French strains (Lb84, StFoy) having similar overall expression pattern but distinct from the other strains (Fig. S2). On the contrary, the virus did not have a strong effect on the overall wasp expression pattern, if we rely on the principal component analysis (PCA) (Fig. S2A–C). Nevertheless, by analysing the dataset by 2 approaches (separate analysis in the head and abdomen, and test of the interaction tissue:infection, see methods) we identified a set of wasp genes whose expression is deregulated by the

https://doi.org/10.1017/S0031182018000835 Published online by Cambridge University Press

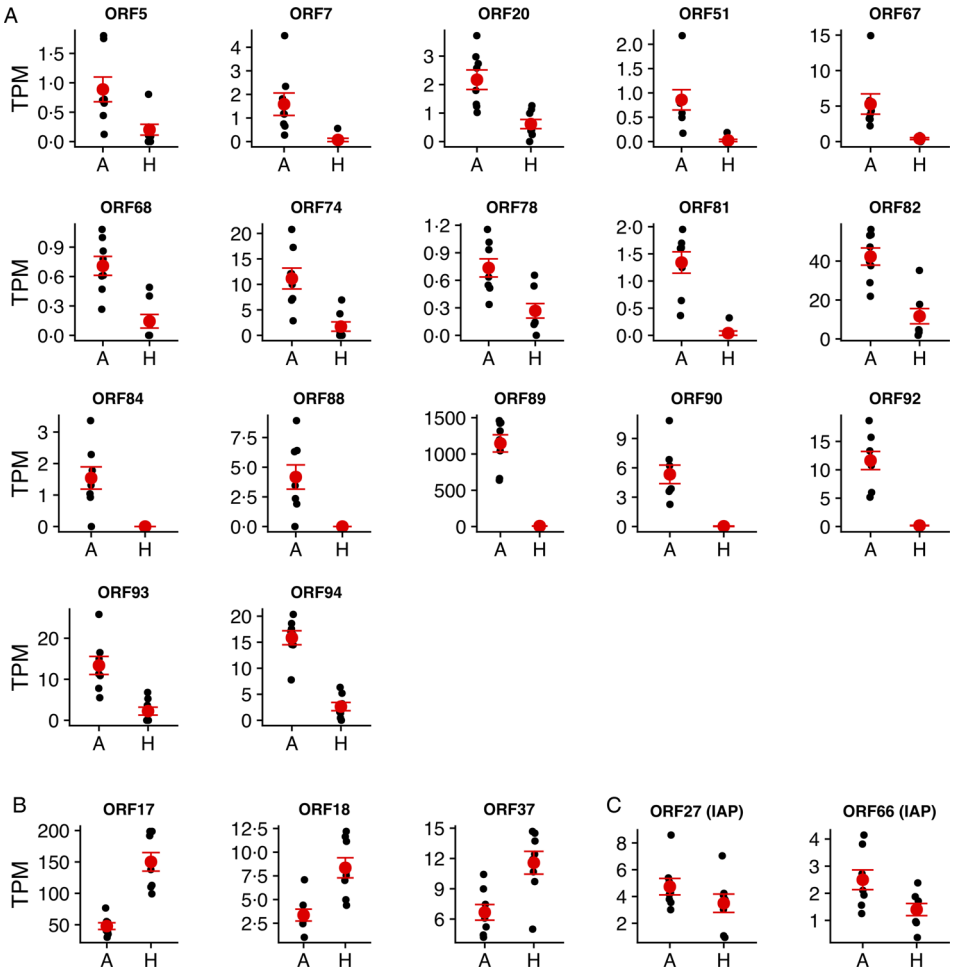


Fig. 3. Candidate viral effector genes displaying (A) abdomen-skewed abundance profiles or (B) head-skewed expression profiles. (C) ORFs 27 and 66, which are putative apoptosis inhibitors, are presented as non-differentially expressed controls. TPM stands for transcripts per million. Error bars indicate standard errors.

virus either in the head or in the abdomen (Fig. S2D and E). In total, we identified 153 wasp genes whose expression is either 'abdomen-altered' or 'head-altered' by the virus (among which 144 were identified by the first analysis, 9 by the second and 2 by both).

In the abdomen, we identified 76 genes up-regulated and 59 genes downregulated in presence of the virus, corresponding to 0.9% of loci identified in the abdomen (Fig. 4A and Table S2). In the head, 13 genes were upregulated compared with 5 downregulated corresponding to 0.1% of all expressed genes in the head (Fig. 4B and Table S3). Overall, the fold-changes associated with viral infection were relatively low (most of them were inferior to 2) although significant at $P < 0.05$ after correction for multiple testing. Approximately 2/3 (13/18) of the head-specifically DE genes were annotated (based on blast analysis) which is similar to the global annotation rate (67%). Annotation was slightly better for the abdomen-specifically DE genes (115/135, 85%). This annotation allows for a cautious functional interpretation of the results. There was no clear gene ontology (GO) enrichment in the candidate gene sets compared with representative control gene sets (non differentially expressed genes with a similar distribution of expression level) neither in the head nor in the abdomen (not shown). However, we identified genes deregulated that had annotation suggesting a possible link with the wasp behaviour (Table 1).

In the abdomen

All proteins discussed in this section have been obtained using the first analysis (separate analysis, see methods). We identified 5 transcripts putatively encoding proteins of particular interest (Fig. 5A): a protein involved in the morphogenesis of receptors (XLOC_012210), a small protein interacting with pheromones (XLOC_007994), 2 proteins involved in the excitability of neurons (XLOC_009885, XLOC_016502), and a protein involved in the recycling of neurotransmitter (XLOC_009407).

XLOC_012210 is highly similar to the transmembrane protein crumbs, both in sequence and structure (presence of a long extracellular domain and a short intracellular one, 3 laminin domains and several EGF motifs, Fig. S3A and Table S2). Crumb proteins are highly conserved proteins found in insects as well as in nematodes or vertebrates. They are involved in several developmental processes including photoreceptor organization. Indeed, mutations in crumbs are associated with retina degeneration in human and with progressive light-induced degeneration of photoreceptors in *Drosophila* (Bazellieres *et al.*, 2009). The orthologous of crumbs in *L. boulardi* is specifically downregulated in infected abdomens.

XLOC_007994 is predicted to encode a protein belonging to the Insect pheromone-binding family A10/OS-D. It is specifically upregulated in the abdomen (Fig. 5A). Pheromone-binding proteins are involved in the solubilization and transport of odorants, which are generally hydrophobic. They play a major role in presenting chemical signals (pheromone, odors) to transmembrane receptor proteins in neurons that subsequently translate the chemical signal into an action potential (Vieira and Rozas, 2011). The virus is associated with an abdomen-specific increase in the abundance of this transcript.

Two transcripts encoding proteins involved in the transport of cations are specifically downregulated in the abdomen of infected wasps (XLOC_009885, XLOC_016502, Fig. 5A, Table S2). XLOC_009885 is predicted to be a Calcium-activated BK potassium channel alpha subunit (Pfam domain PF03493). The protein

https://doi.org/10.1017/S0031182018000835 Published online by Cambridge University Press

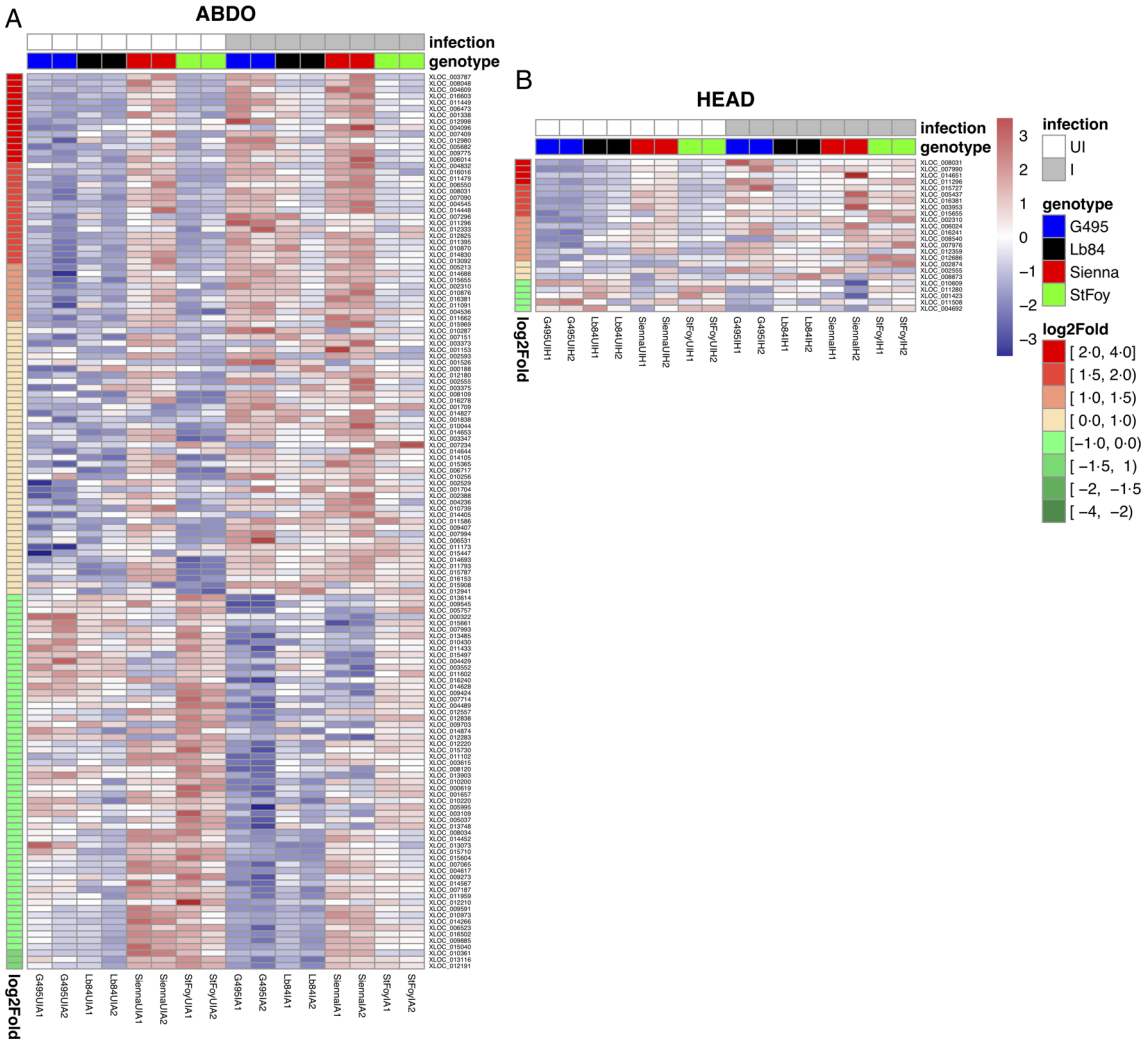


Fig. 4. Heatmap of wasp gene expression showing specific de-regulation in relation to virus presence in (A) the abdomen (B) the head. The level of transcription was measured as transcripts per million (TPM). The data were centered and reduced according to the rows (loci), which reduced the strong variance between genes, increasing readability. In the strain names (in column), the first letter indicate the wasp genetic background (Lb84, G495, StFoy, Sienna) the next one indicate their LbFV infection status (I: infected, UI: uninfected), the next one indicate the tissue (A: abdomen) and the number indicates the biological replicate number (1 or 2).

presumably encoded by XLOC_016502 belongs to the Cation transporting ATP-ase P-type family. These proteins are simple single catalytic membrane-bound enzymes that use ATP hydrolysis to drive the transport of cations across membranes (Palmgren and Nissen, 2011). Notably, the abundance of both transcripts is highly correlated among all 32 samples (Fig. S4, $P = 1.6 \times 10^{-16}$), suggesting a common transcriptional regulation and a possible common function.

After proper induction of an action potential in a neuron, neurotransmitters will be released in a synaptic cleft. For most neurotransmitters (all but acetycholine), this release is accompanied by a fast re-uptake of the neurotransmitter by the presynaptic neuron, which terminates the signalling. This phenomenon is essential for the precision of the nervous influx. Proteins responsible for this recycling are trans-membrane transporters that depend on the co-transport of Na+ or Cl− for obtaining the energy required for neurotransmitter transport against its concentration gradient (Lodish *et al.*, 2000). We found that XLOC_009407, which encodes such a protein, has 8 predicted transmembrane domains (Fig. S3A), and is upregulated in presence of the virus, specifically in the abdomen (Fig. 5A).

In the head

We considered 2 genes as candidates here, out of 18 (Table S3). Both were obtained using the second analysis, relying on the test of the interaction tissue:infection (see methods). Based on the observation of the expression data in both the head and the abdomen (Fig. 5B), each of these genes were classified as specifically downregulated in presence of the virus in the head, although they also showed the opposite trend in the abdomen (thus explaining why the interaction term was significant). However, considering the very low level of expression in the abdomen (TPM < 5) for both genes, we considered that the expression trend observed in the abdomen was not very reliable, contrary to what is observed in the head. We thus considered them as 'specifically head-altered'.

First, we identified a putative gamma-aminobutyric acid receptor (XLOC_011280). This protein is predicted to contain 4 transmembrane domains (Fig. S3B), which is the typical structure for gamma-aminobutyric acid receptors. Gamma-aminobutyric acid, also known as GABA, is the major fast inhibitory neurotransmitter in the nervous system of vertebrates and insects.

Second, we found that a protein belonging to the yellow/Major Royal Jelly Protein family is downregulated in presence of the virus in the head (XLOC_010609). This protein family is well known because honeybee queens are fed with a diet exclusively composed of such proteins. They are secreted by hypopharyngal glands of young nurse bees, and play a nutritional role. However, MRJP are particularly expressed in the honeybee brain (Peixoto *et al.*, 2009) and may be involved in honeybee behaviour and possibly in the development of learning abilities (Hojo *et al.*, 2010).

https://doi.org/10.1017/S0031182018000835 Published online by Cambridge University Press

https://doi.org/10.1017/S0031182018000835 Published online by Cambridge University Press

Table 1. List of candidate genes identified

Locus	First blast hit description	Length	Blast evalue	Mean_sim	pfam_target_name	pfam_accession	pfam_evalue	pfam_description	log2Fold Change.abdo	padj.abdo	log2Fold Change.head	padj.head	padj_interaction_analysis
XLOC_012210	protein crumbs isoform X3 [XP_015599267.1]	8479	0	81.15	EGF/Laminin_G_1/hEGF/Laminin_G_2/EGF_CA/EGF_2/EGF/hEGF/EGF_2/EGF_CA/FXa_inhibition/Laminin_G_2/EGF/hEGF/FXa_inhibition/EGF_2/EGF_3	PF00008.26/PF00054.22/PF12661.6/PF02210.23/PF07645.14/PF07974.12/PF00008.26/PF12661.6/PF07974.12/PF07645.14/PF14670.5/PF02210.23/PF00008.26/PF12661.6/PF14670.5/PF07974.12/PF12947.6	1.2×10^{-60}/2.6×10^{-44}/3.7×10^{-36}/3.4×10^{-33}/9.7×10^{-11}/5.1×10^{-9}/9.5×10^{-52}/1.9×10^{-30}/3.4×10^{-12}/3.8×10^{-09}/1.3×10^{-6}/3.5×10^{-5}/4.7×10^{-23}/7.6×10^{-12}/1.8×10^{-6}/2.6×10^{-5}/8×10^{-5}	EGF-like domain/Laminin G domain/Human growth factor-like EGF /Laminin G domain/Calcium-binding EGF domain/EGF-like domain/EGF-like domain/Human growth factor-like EGF /EGF-like domain/Calcium-binding EGF domain/Coagulation Factor Xa inhibitory site/Laminin G domain/EGF-like domain/Human growth factor-like EGF /Coagulation Factor Xa inhibitory site/EGF-like domain/EGF domain	−0.611	**0.045**	0.148	1.000	1.000
XLOC_007994	DC-STAMP domain-containing protein 1 [XP_011153575.1]	2016	1.30×10^{-46}	86.55	OS-D	PF03392.12	2.50×10^{-37}	Insect pheromone-binding family, A10/OS-D	0.400	**1.7×10^{-4}**	0.195	1.000	1.000
XLOC_009885	potassium channel subfamily T member 2 isoform X6 [XP_012232623.1]	12 034	0	89.3	BK_channel_a/Ion_trans_2/PP2C/CAP/CAP	PF03493.17/PF07885.15/PF00481.20/PF00188.25/PF00188.25	2.2×10^{-29}/5.3×10^{-9}/3.2×10^{-64}/9.4×10^{-13}/4.8×10^{-8}	Calcium-activated BK potassium channel alpha/Ion channel/Protein phosphatase 2C/Cysteine-rich secretory protein family /Cysteine-rich secretory protein family	−0.855	**0.001**	−0.071	1.000	1.000
XLOC_016502	probable cation-transporting ATPase 13A3 isoform X2 [XP_012139702.1]	7262	0	85.3	Hydrolase/Cation_ATPase/P5-ATPase/E1-E2_ATPase/Cation_ATPase_N	PF00702.25/PF13246.5/PF12409.7/PF00122.19/PF00690.25	9.2×10^{-22}/2.4×10^{-6}/4.2×10^{-24}/7.4×10^{-19}/0.00096	haloacid dehalogenase-like hydrolase/Cation transport ATPase (P-type) /P5-type ATPase cation transporter /E1-E2 ATPase/Cation transporter/ATPase, N-terminus	−0.844	**3.7×10^{-4}**	0.021	1.000	1.000
XLOC_009407	excitatory amino acid transporter 1 [XP_012276012.1]	3177	5.25×10^{-154}	77.75	SDF	PF00375.17	4.50×10^{-123}	Sodium:dicarboxylate symporter family	0.403	**0.012**	−0.110	1.000	0.425
XLOC_011280	gamma-aminobutyric acid receptor alpha-like [XP_006610609.1]	2547	2.58×10^{-108}	82.1	Neur_chan_memb/Neur_chan_LBD	PF02932.15 PF02931.22	1.5×10^{-28}/1.4×10^{-13}	Neurotransmitter-gated ion-channel transmembrane region /Neurotransmitter-gated ion-channel ligand binding domain	0.169	0.919	−0.535	0.176	**0.023**
XLOC_010609	yellow-b precursor [NP_001154989.1]	2239	1.67×10^{-164}	81.45	MRJP	PF03022.15	3.10×10^{-88}	Major royal jelly protein	0.744	0.665	−0.230	1.000	**0.026**

The adjusted *P* values of the separate analysis are given (padj.abdo and padj.), together with the adjusted *P* value associated with the interaction tissue × infection obtained from the second analysis (p_adj_interaction_analysis). The Blast2GO analysis did not lead to GO term mapping for any of these mRNA. Bold indicates *p*-values < 0.05.

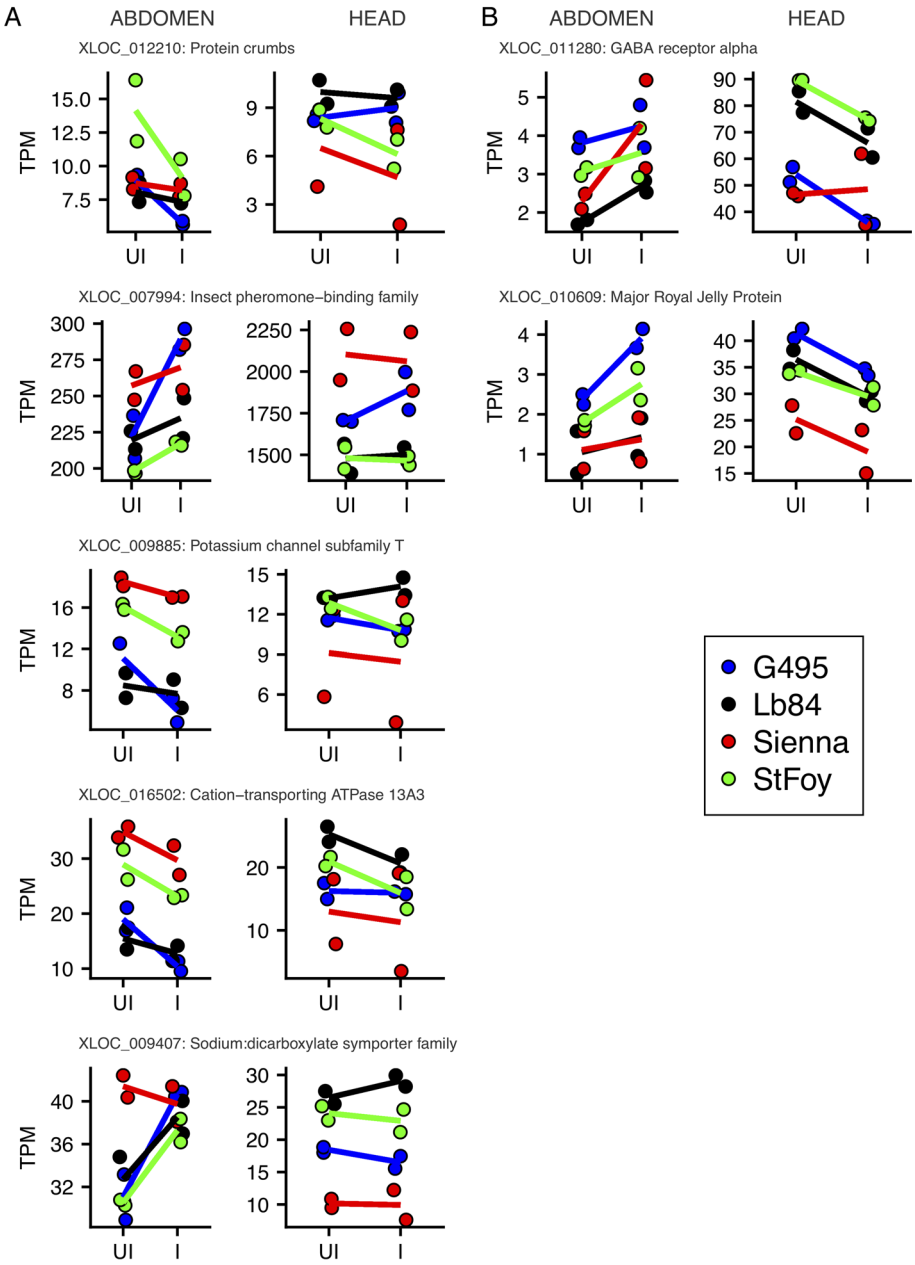


Fig. 5. Wasp gene expression for candidates showing deregulation in presence of the virus in particular (A) in the abdomen, or (B) in the head. UI: uninfected, I: infected. The annotation derived from the first blast hit is indicated. Colors indicate the identity of the wasp strain. All genes presented here where obtained by the first analysis except XLOC_011280 and XLOC_010609 that were obtained by the second one (see method). All genes obtained by the first analysis were filtered to keep only locus showing consistent trend for all 4 wasp strains. The apparent discrepancy observed for XLOC_009407 (Sienna/Abdomen) is explained by the fact that the statistic presented here (TPM) is different from the one used to measure the trend in our analysis (measured by the sign of the log2-fold in DESEq2 contrast, see methods for details).

Correlation with the phenotype

We tested the existence of a correlation between each candidate wasp gene expression level and the observed phenotype of the strain (Fig. 1). From the phenotypic data, we would expect larger virus-induced transcriptomic effects in strain G495 than in strain StFoy for instance since G495 is more manipulated than StFoy. None of the candidate genes were significantly correlated with the phenotypes (after correction for multiple testing). However, this test has very little statistical power since only 4 phenotypic values are available (since we do not have individual phenotypes of the females used in the RNAseq experiment but rather on their sisters). The coefficient correlation and its associated raw *P* value (not corrected) can be found in the last 4 columns of the Supplementary Tables S2 and S3.

Discussion

In order to decipher the molecular basis of the behaviour manipulation induced by the virus LbFV on its parasitoid host, we studied the transcriptome of both the virus and the wasp in a dual RNAseq experiment. We identified 17 viral genes showing an abdomen-skewed expression profile and 3 viral genes having a head-skewed expression profile. Unfortunately, we have only very few cues on the possible functions associated with these proteins because LbFV is very distantly related to other known viruses (Lepetit *et al.*, 2016). Nevertheless, we noticed that one transcript has particular high abundance in wasp abdomen (with no known homologs in public databases, ORF89) and may be a key effector for the manipulation. We previously identified 2 LbFV genes bearing a Jumonji domain and having a peculiar evolutionary history. Those genes (ORF11 and 13) probably derive from an ancestral wasp gene that have been captured by an ancestor of the virus, with a subsequent gene duplication (Lepetit *et al.*, 2016). Based on their particular evolutionary history and on their putative involvement in lysine demethylation of histones, we previously speculated that this gene could have been co-opted by the virus to manipulate the behaviour of the wasp, as it as been observed in baculovirus–Lepidoptera interactions (Katsuma *et al.*, 2012). Neither ORF11, nor ORF13 are part of the candidate list in the present analysis, although ORF13 shows a tendency for head-skewed transcript abundance (Fig. 2).

On the wasp side, we utilized the annotations derived from the blast outputs combined with pfam domain searches to identify a set of candidate genes in the abdomen ($n = 5$) and in the

https://doi.org/10.1017/S0031182018000835 Published online by Cambridge University Press

head ($n = 2$) possibly targeted by the virus. Although at that stage we cannot present a definitive scenario explaining the extended phenotype, we propose in the following section a hypothetical one.

The central nervous system of insects is a point of convergence for multiple excitatory and inhibitory channels of information, informing the insect over its environment (Barron *et al.*, 2015). Such integration may be mediated by the convergence of several excitatory and inhibitory synapses on the same neuronal population that may elicit or not the corresponding behaviour depending on the balance of both types of stimulations. In our case, we can hypothesize that a set of neurons located in the central nervous system are involved in the elicitation of the stereotypical oviposition program, as observed in *Drosophila* (Yang *et al.*, 2008). Those neurons would receive sensory input both from excitatory and inhibitory synapses. Excitatory synapses would inform the central nervous system that for instance a suitable larva has been encountered. On the contrary, inhibitory synapses would preclude the elicitation of oviposition behaviour since they would inform the wasp that the encountered larva is not suitable, for instance, because it is already parasitized. When the excitation in the efferent neuron reaches a threshold, the oviposition program would be elicited. In the following, we discuss our results under this general hypothesis.

A hypothetical inhibitory pathway for oviposition

We will focus here on a supposedly inhibitory pathway, linked to the perception of the presence of a parasitoid inside the encountered *Drosophila* larva. This pathway is expected to inhibit the initiation of the typical oviposition program observed in this species (Varaldi *et al.*, 2005). We found 2 results that may lead to a reduction in the strength of such inhibitory pathway. First, the observation of an abdomen-specific upregulation of a Sodium: dicarboxylate symporter (XLOC_009407), which is involved in recycling of neurotransmitters released in the synaptic cleft, may translate into a disruption of the transmission from presynaptic to postsynaptic neurons. Indeed, neurons over-producing such protein may be less able at eliciting an action potential in the postsynaptic neurons because of an excessive or an excessively fast recycling of the neurotransmitter in the synaptic cleft. If the presynaptic neuron is involved in the detection of parasitoid larvae/eggs inside the *Drosophila* larva, we then predict that this molecular manipulation will limit the transmission of the signal to the postsynaptic neuron, ultimately reducing superparasitism avoidance. Second, we found that the virus is associated with a downregulation of a GABA receptor (XLOC_011280) in the head of the wasp. GABA is the major fast inhibitory neurotransmitter in the nervous system of vertebrates and insects. Indeed, its release in the synaptic space leads to the opening of the pentameric transmembrane chloride channel within GABA receptors, thereby releasing negative charged chloride ions into the neuron. This entrance of negatively charged ions leads to a hyperpolarization, which reduces its excitability (Palmer and Harvey, 2014). If we assume that this phenomenon occurs in the inhibitory channel involved in superparasitism avoidance, then a reduction in the transcription of this putative GABA receptor subunit should impede the transmission of the 'inhibitory signal'. Interestingly, changing the ion-selectivity of an inhibitory anion channel into an excitatory cation channel led to a switch behaviour in the worm *Caenorhabditis elegans* (Pirri *et al.*, 2015), demonstrating how the balance of excitatory *vs* inhibitory synapses may profoundly affect animal decisions. Likewise, knockdown of a GABA receptor, in *Drosophila* increased the consumption of non-appetitive and appetitive substances, beyond satiation (Cheung and Scott, 2017).

A hypothetical excitatory pathway for oviposition

We found that a pheromone-binding protein (XLOC_007994) is specifically upregulated in presence of the virus in the abdomen. We expect that this increase should translate into a superior capacity of infected females to detect the corresponding odor. If the detected signal is linked to the overall quality of the encountered host, then it can be predicted that this transcriptomic change over-stimulates the excitatory pathway, ultimately increasing the tendency of females to lay eggs in the encountered host. In the same line, we found that 2 transcripts encoding proteins involved in the transport of cations are specifically downregulated in the abdomen of infected wasps (XLOC_009885, XLOC_016502).

XLOC_009885 is predicted to be a BK channel. BK channels are characterized by their large conductance for potassium ions through cell membranes and are activated by high concentration of Ca2+ or by changes in membrane action potential. The opening of BK channels leads to a passive diffusion of K+ ions along the electrochemical gradient. In normal physiological situations, this opening leads to a massive expulsion of K+ ions outside the cells, leading to hyperpolarization in neurons, thus reducing their excitability (Miller, 2000). Because both XLOC_016502 and XLOC_009885 are downregulated, we may expect an increase in the excitability of neurons in the infected female, through a modification of the intracellular cation concentration. If this increase in excitability concerns the 'pathway for oviposition', then it may participate in increasing the tendency of females to accept the host.

Under the proposed scenario the virus would be able to turn parasitized hosts into acceptable hosts by manipulating the balance between inhibitory and excitatory pathways in favour of the latter. This should lead to increased superparasitism.

Our results identified a protein belonging to the yellow/Major Royal Jelly Protein family (XLOC_010609) as a candidate gene. Its possible involvement in a behaviour phenotype is plausible (Peixoto *et al.*, 2009). Interestingly, the genome of the parasitoid wasp *Nasonia vitripennis* encodes the greatest number of copies belonging to this protein family (Buttstedt *et al.*, 2014), raising the possibility that yellow/Major Royal Jelly protein family fulfills a particular function in parasitoids. In line with the *Nasonia* genome, using a pfam search, we found 12 locus containing a protein/MRJP domain in the genome of *L. boulardi* (not shown). We can hypothesize that this protein plays a role in superparasitism avoidance.

It is clear that at that stage we can neither rule out the possibility that other genes not discussed here are important players in the extended phenotype, nor that we identified genes involved in other aspects of the interaction. Also, we cannot completely exclude that our candidate list contains some false positives, although we tried to be stringent in our criterions. In order to further test the involvement of the candidate genes in the behaviour manipulation, future work will aim at manipulating their expression by RNAi or CrisprCas9 techniques. These experiments should allow one to identify the effector genes responsible for the manipulation in the virus genome and the genes targeted in the wasp genome/transcriptome. This work could eventually help in understanding not only the mechanisms of the behaviour manipulation but also the genetic basis of superparasitism in parasitoids.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0031182018000835

Acknowledgements. This work was supported by a grant from the Agence Nationale de la Recherche (ANR) to JV (11-JSV7-0011 Viromics). The bioinformatic work was performed using the computing facilities of the CC LBBE/PRABI. We thank Prof. Valérie Raymond and anonymous reviewers

https://doi.org/10.1017/S0031182018000835 Published online by Cambridge University Press

for helpful comments on a previous draft of the manuscript and Michael Love for statistical advice.

Conflict of interest. None.

Ethical standards. Not applicable.

References

Anders S, Pyl PT and Huber W (2015) HTSeq – a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England)* **31**, 166–169.

Barron AB, Gurney KN, Meah LFS, Vasilaki E and Marshall JAR (2015) Decision-making and action selection in insects: inspiration from vertebrate-based theories. *Frontiers in Behavioral Neuroscience* **9**, 363.

Bazellieres E, Assemat E, Arsanto J-P, Le Bivic A and Massey-Harroche D (2009) Crumbs proteins in epithelial morphogenesis. *Frontiers in Bioscience* **14**, 2149–2169.

Berdoy M, Webster JP and Macdonald DW (2000) Fatal attraction in rats infected with *Toxoplasma gondii*. *Proceedings Biological Sciences/the Royal Society* **267**, 1591–1594.

Buttstedt A, Moritz RFA and Erler S (2014) Origin and function of the major royal jelly proteins of the honeybee (*Apis mellifera*) as members of the yellow gene family. *Biological Reviews* **89**, 255–269.

Charnov EL (1976) Optimal foraging, the marginal value theorem. *Theoretical Population Biology* **9**, 129–136.

Chen YR, Zhong S, Fei Z, Hashimoto Y, Xiang JZ, Zhang S and Blissard GW (2013) The transcriptome of the baculovirus autographa californica multiple nucleopolyhedrovirus in trichoplusia ni cells. *Journal of Virology* **87**, 6391–6405.

Cheung SK and Scott K (2017) GABAA receptor-expressing neurons promote consumption in *Drosophila melanogaster*. *PLoS ONE* **12**, e0175177.

De Bary HA (1879) De la symbiose. *Revue internationale des Sciences* **3**, 301–309.

Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate JG and Bateman A (2016) The Pfam protein families database – towards a more sustainable future. *Nucleic Acids Research* **44**, D279–D285.

Gandon S, Rivero A and Varaldi J (2006) Superparasitism evolution: adaptation or manipulation? *The American Naturalist* **167**, E1–E22.

Godfray H (1994) *Parasitoids: Behavioral and Evolutionary Ecology*. Princeton, NJ: Princeton University Press.

Han Y, van Oers MM, van Houte S and Ros VID (2015). Virus-induced behavioural changes in insects. In Mehlhorn H (ed.), *Host Manipulations by Parasites and Viruses*. Cham: Springer International Publishing, pp. 149–174 doi: 10.1007/978-3-319-22936-2_10.

Hojo M, Kagami T, Sasaki T, Nakamura J and Sasaki M (2010) Reduced expression of major royal jelly protein 1 gene in the mushroom bodies of worker honeybees with reduced learning ability. *Apidologie* **41**, 194–202.

Hoover K, Grove M, Gardner M, Hughes DP, McNeil J and Slavicek J (2011) A gene for an extended phenotype. *Science (New York, NY)* **333**, 1401.

Katsuma S, Koyano Y, Kang W, Kokusho R, Kamita SG and Shimada T (2012) The baculovirus uses a captured host phosphatase to induce enhanced locomotory activity in host caterpillars. *PLoS Pathogens* **8**, e1002644.

Kim D, Langmead B and Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* **12**, 357–360.

Leitão-Gonçalves R, Carvalho-Santos Z, Francisco AP, Fioreze GT, Anjos M, Baltazar C, Elias AP, Itskov PM, Piper MDW and Ribeiro C *et al.* (2017) Commensal bacteria and essential amino acids control food choice behavior and reproduction. *PLoS Biology* **15**, e2000862.

Lepetit D, Gillet B, Hughes S, Kraaijeveld K and Varaldi J (2016) Genome sequencing of the behavior manipulating virus LbFV reveals a possible new virus family. *Genome Biology and Evolution* **8**, 3718–3739.

Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D and Darnell D (2000) *Molecular Cell Biology*, 4th Edn. New York: W. H. Freeman.

Love MI, Huber W and Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 31.

Martinez J, Fleury F and Varaldi J (2012a) Heritable variation in an extended phenotype: the case of a parasitoid manipulated by a virus. *Journal of Evolutionary Biology* **25**, 54–65.

Martinez J, Fleury F and Varaldi J (2015) Competitive outcome of multiple infections in a behavior-manipulating virus/wasp interaction. *Ecology and Evolution* **5**(24), 5934–5945.

Martinez J, Duplouy A, Woolfit M, Vavre F, O'Neill SL and Varaldi J (2012b) Influence of the virus LbFV and of *Wolbachia* in a host-parasitoid interaction. *PLoS ONE* **7**, e35081.

Miller C (2000) An overview of the potassium channel family. *Genome Biology* **1**(4), reviews0004.1–0004.5.

Palmer MJ and Harvey J (2014) Honeybee Kenyon cells are regulated by a tonic GABA receptor conductance. *Journal of Neurophysiology* **112**, 2026–2035.

Palmgren MG and Nissen P (2011) P-Type ATPases. *Annual Review of Biophysics* **40**, 243–266.

Patot S, Lepetit D, Charif D, Varaldi J and Fleury F (2009) Molecular detection, penetrance, and transmission of an inherited virus responsible for behavioral manipulation of an insect parasitoid. *Applied and Environmental Microbiology* **75**, 703–710.

Patot S, Martinez J, Allem R, Gandon S, Varaldi J and Fleury F (2010) Prevalence of a virus inducing behavioural manipulation near species range border. *Molecular Ecology* **19**, 2995–3007.

Peixoto LG, Calábria LK, Garcia L, Capparelli FE, Goulart LR, de Sousa MV and et Espindola FS (2009) Identification of major royal jelly proteins in the brain of the honeybee Apis mellifera. *Journal of Insect Physiology* **55**, 671–677.

Peng Y, Leung HCM, Yiu SM and Chin FYL (2012) IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics (Oxford, England)* **28**, 1420–1428.

Pirri JK, Rayes D and Alkema MJ (2015) A change in the ion selectivity of ligand-gated ion channels provides a mechanism to switch behavior. *PLoS Biology* **13**, e1002238.

Ros VID, van Houte S, Hemerik L and van Oers MM (2015) Baculovirus-induced tree-top disease: how extended is the role of egt as a gene for the extended phenotype? *Molecular Ecology* **24**, 249–258.

Ruschioni S, van Loon JJA, Smid HM and van Lenteren JC (2015) Insects can count: sensory basis of host discrimination in parasitoid wasps revealed. *PLoS ONE* **10**, e0138045.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV and Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics (Oxford, England)* **31**, 3210–3212.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ and Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology* **28**, 511–515.

Van Alphen JJ and Visser ME (1990) Superparasitism as an adaptive strategy for insect parasitoids. *Annual Review of Entomology* **35**, 59–79.

Van Lenteren JC and Bakker K (1975) Discrimination between parasitised and unparasitised hosts in the parasitic wasp *Pseudeucoila bochei*: a matter of learning. *Nature* **254**, 417–419.

Van Lenteren JC, Ruschioni S, Romani R, Van Loon JJA, Qiu YT, Smid HM, Isidoro N and Bin F (2007) Structure and electrophysiological responses of gustatory organs on the ovipositor of the parasitoid *Leptopilina heterotoma*. *Arthropod Structure & Development* **36**, 271–276.

Varaldi J, Fouillet P, Ravallec M, Lopez-Ferber M, Boulétreau M and Fleury F (2003) Infectious behavior in a parasitoid. *Science* **302**, 1930–1930.

Varaldi J, Fouillet P, Boulétreau M and Fleury F (2005) Superparasitism acceptance and patch-leaving mechanisms in parasitoids: a comparison between two sympatric wasps. *Animal Behaviour* **69**, 1227–1234.

Varaldi J, Ravallec M, Labrosse C, Lopez-Ferber M, Boulétreau M and Fleury F (2006) Artifical transfer and morphological description of virus particles associated with superparasitism behaviour in a parasitoid wasp. *Journal of Insect Physiology* **52**, 1202–1212.

Varaldi J, Patot S, Nardin M and Gandon S (2009) A virus-shaping reproductive strategy in a *Drosophila* parasitoid. *Advances in Parasitology* **70**, 333–363.

Vieira FG and Rozas J (2011) Comparative genomics of the odorant-binding and chemosensory protein gene families across the arthropoda: origin and evolutionary history of the chemosensory system. *Genome Biology and Evolution* **3**, 476–490.

Xue W, Li J-T, Zhu Y-P, Hou G-Y, Kong X-F, Kuang Y-Y and Sun X-W (2013) L_RNA_scaffolder: scaffolding genomes with transcripts. *BMC Genomics* **14**, 1–1.

Yang C-H, Belawat P, Hafen E, Jan LY, Jan Y-N (2008) Drosophila egg-laying site selection as a system to study simple decision-making processes. *Science* **319**, 1679–1683.

https://doi.org/10.1017/S0031182018000835 Published online by Cambridge University Press