Supplementary Data

Supplementary Material and Methods

(a) Animal sampling, rearing, and experimental design

Adult *Coleomegilla maculata* (more than 4000) were collected in 2009 at Saint-Mathieu-de-Beloeil, Québec, Canada. Ladybeetles were reared in plastic mesh boxes (946 ml, Ziploc) and fed *ad libitum* with pollen and aphids as previously described [1]. Emerging *Dinocampus coccinellae* parasitoids were used to start a laboratory colony from Québec. Parasitoids from Poland, Japan, and The Netherlands were kindly provided by P. Ceryngier (Polish Academy of Science, Lomianki, Poland), S. Koyama (Tokyo University of Agriculture and Technology, Japan), and J. Stam (WU Plant Sciences, Wageningen, The Netherlands), respectively. Adult *C. maculata* were exposed daily in a mesh cage (35.5 cm × 20 cm × 19 cm in height) to female *D. coccinellae* at a 2:1 ratio for 24 h. Parasitized *C. maculata* were transferred into plastic mesh boxes and fed as described above until egression. All insects were reared in a growth chamber (Conviron E15) at 24±1°C, 50% relative humidity, and 16 L : 8 D photoperiod.

From the ladybeetle colony, we collected heads and abdomens (including thorax) from 12 healthy adult males and 12 healthy adult females *C. maculata* (samples healthy, He; Figure 1). The sex of *C. maculata* was determined by observation of genitalia. Hundreds of *C. maculata* were exposed to female *D. coccinellae*. For parasitized hosts, *D. coccinellae* larvae and heads and abdomens of *C maculata* were collected from 12 individuals of each sex at three time points before egression (i.e., 5, 13, or 20 days at 25°C; samples before egression, Be). For 12 more individuals of each sex, we let the parasitoid prepupa egress from the host, and we collected prepupa and ladybeetle heads and abdomens before the prepupa had spun a cocoon (samples after egression, Ae). Finally, we collected 12 male and 12 female ladybeetles that had been left undisturbed until *D. coccinellae* egression and that survived the bodyguard manipulation (samples recovering, R). Again, heads and abdomens were collected. Some ladybeetles were resistant to parasitism, and an encapsulated parasitoid was found within the ladybeetle abdomen at dissection 25 days post oviposition (two

samples, Res). The abdomens and heads of these resistant ladybeetles were collected separately. Four adult *D. coccinellae* from Québec were sampled and four prepupa (Ae) (two pools of two individuals) of *D. coccinellae* colonies from Japan, Poland, and the Netherlands were also sampled.

All samples were quick-frozen in liquid nitrogen and stored in liquid nitrogen until RNA extraction.

(b) Ethic Statement

No special permit was necessary for collection of wild ladybeetles and parasitoids

(c) Total RNA extraction

Total RNA was extracted from individual tissues (abdomens, Ab) or from pools of four individuals (larvae, L and heads, H). Only abdomens were processed further for Res samples. First, cell disruption was carried out by grinding tissues in liquid nitrogen with mortar and pestle. The resulting powder was mixed with 500 µl of Trizol reagent. Abdomen tissue powder resuspended in Trizol was transferred to tubes containing ceramic beads and subjected to mechanical homogenization using a MagNA Lyzer instrument (Roche) for 30s at a speed of 6,000. Abdomen samples were kept on ice for 1 min and homogenization was repeated twice before transferring samples into new tubes. Abdomen, head, and larva samples were then stored at -80°C.

Total RNA was extracted from samples according to the manufacturer's instructions. Briefly, 100 µl of chloroform was added to the 500 µl of Trizol extract and homogenized for 15 min at room temperature. Samples were centrifuged for 15 min at 12,000×g and the aqueous phase was extracted and transferred into a new tube. RNA was precipitated in 500 µl of isopropanol and centrifuged 10 min at 12,000×g. RNA pellets were finally washed twice with 500 µl of cold ethanol (70%) and resuspended in water. RNA concentrations were determined using an ND-1000 spectrophotometer (Nanodrop Technologies) at 260 nm using the conversion factor 1 OD = 40 µg/ml RNA. Six

biological replicates (three males and three females; pools of four individuals) were used for Ab, H and L from conditions He, Be D20, Ae, and R. For condition Be D5 and D13, four biological replicates of single individuals were used for Ab and two biological replicates of pools of four individuals were used for T and L. Four adult wasps were used as biological replicates.

(d) cDNA library construction and Illumina sequencing

For RNA sequencing, equimolar amounts of pools of males and females (total of 24 individuals with 12 males and 12 females) were pooled for each category leading to a total of 12 samples: Abdomen He, Abdomen Be D20, Abdomen Ae, Abdomen R, Head He, Head Be D20, Head Ae, Head R, two samples of Larva Be D20, and two samples of Larva Ae. Contaminating DNA was removed from pools of total RNA with an Ambion DNA-free kit following the standard procedure. The absence of trace genomic DNA was confirmed by gel electrophoresis. RNA concentrations were determined using an ND-1000 spectrophotometer (Nanodrop Technologies) at 260 nm using the conversion factor 1 OD = 40 μg/ml RNA. RNA integrity was controlled on the Agilent bioAnalyzer using RNA 6000 Nano kits (Agilent Technologies) according to manufacturer's instructions without consideration for the RNA Integrity Number (RIN) [2].

The twelve cDNA libraries were generated using the TruSeq RNA Sample Prep Kit V2 for transcriptome sequencing on the Illumina HiSeq2000 platform. Paired-end 100 bp cDNA libraries were generated for each sample category. Four samples were multiplexed for each lane. Library construction and sequencing were performed by commercial service provider (MGX, Montpellier Genomix, c/o Institut de Génomique fonctionnelle, Montpellier, France). Each library was validated using a DNA1000 Chip on a 2100 BioAnalyser (Agilent Technologies) and quantified by qPCR using a Kapa SYBR Fast qPCR Kit (Applied Biosystems 7500 Technologies). Only reads that passed the default high quality filtering performed by the Illumina pipeline were conserved for further analyses. Raw fastq files were submitted to NCBI's sequence read archive

(http://trace.ncbi.nlm.nih.gov/Traces/sra/) under the reference PRJNA227418 for *C. maculata* samples and PRJNA227420 for *D. coccinellae* samples.

(e) *De novo* assembly of transcriptomes

All assemblies were performed on a server with 16 cores and 96 GB of random access memory. Before assembly, reads were cleaned using a workflow with the FastX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) that includes the removal of sequencing artefacts, sequence trimming, and adapter clipping. This procedure removed about 5% of the reads from each library. We used the python script provided by Oases (version 0.2.05; http://www.ebi.ac.uk/~zerbino/oases/) that runs velvet (version 1.2.02; http://www.ebi.ac.uk/~zerbino/velvet/) for individual single-k assemblies of short reads using the de Bruijn graph algorithm and exploits paired-end information using Oases (version 0.2.05; http://www.ebi.ac.uk/~zerbino/oases/) to construct transcript isoforms and compute a merged assembly. For each sample, we performed the *de novo* assembly of transcriptome using velvetg with k-mer lengths of 37, 39, 41, 43, 45, 47, 49, 53, 57 and 61. The parameters considered for optimal assembly included number of contigs produced, maximum contig length, total length of all contigs and N50. We observed an inverse correlation between k-mer length and the number of contigs and the total length of contigs. However, the maximum length of contigs and the N50 value increased up to k-mer 49 and started decreasing from k-mer 57. The optimal assembly was obtained with k-mer length of 53, The *de novo* assembly of the transcriptome was performed using k-mers 45, 49, 53, 57, and 61. The merged assemblies had between 95,082 and 157,214 transcripts. All eight assemblies of healthy and parasitized *C. maculata* were then merged using CD-HIT-EST and produced a final transcriptome of 161,750 transcripts. The four assemblies of *D. coccinellae* larvae were merged to produce a final transcriptome of 115,321 transcripts. The *de novo* assembled transcriptomes are publicly available at http://2ei.univ-perp.fr/telechargement/transcriptomes/ALL_Cocc_95.zip and http://2ei.univ-

perp.fr/telechargement/transcriptomes/ALL_Larve_95.zip for parasitized *C. maculata* and *D. coccinellae* transcriptomes, respectively.

(f) Differential gene expression analysis from RNAseq data

Each sample was mapped against the three reference transcriptomes using BWA [3] and for each transcript we obtained raw counts that corresponded to the number of reads mapped to the transcript. Samtools sort and view were used to generate sorted bam files and a table of read counts for each transcript [4]. For differential gene expression analysis with count data using a binomial negative distribution without replication, we used the DESeq package in R with an adjusted p value < 0.05 [5]. We performed all pair comparisons between samples He, Be, Ae and R for both ladybeetle abdomens and heads. A total of 9461 transcripts appeared differentially expressed between at least two experimental conditions.

(g) Functional annotation

The functional annotation of *de novo* assembled transcripts was performed using Blast 2 GO as follows: i) an initial annotation with BLASTX (against the nonredundant NCBI database; e-value at 1.10^{-6}); ii) assignment of Gene Ontology terms (GO; http://www.geneontology.org/); iii) protein domain searches using InterProscan; and iv) enzyme annotation using the Kyoto Encyclopedia of Genes and Genomes (KEGG).

(h) Reverse transcription

cDNA were obtained using the revertAid premium First strand cDNA synthesis kit (Thermo Scientific). RNA was quantified for each sample just prior Reverse transcription. Viral cDNA of both polarities were transcribed with primers containing a 5' tag sequence (Table S4) [6]. Forward primers were used to transcribe cDNA from the negative-strand viral genome while reverse primers were

used to transcribe cDNA from the positive-strand viral genome. For gene expression analysis, *C. maculata* transcripts were reverse transcribed with random primers. Briefly, primers (at a final concentration of 500 nM for gene expression and 50 nM for strand specific RT-qPCR) were incubated with RNA (500 ng for gene expression and 500 μg for strand specific RT-qPCR) for 5 min at 65°C and then placed on ice for 2 min. cDNA was synthesized with a reverse aid premium enzyme at 25°C for 10 min and at 50°C for 20 min, and then heat inactivated at 85°C for 5 min. cDNAs used in the generation of standard curves were serially diluted and stored at -20°C until use.

(i) *Dinocampus coccinellae* Paralysis Virus (DcPV) genome sequencing

A draft genome of *Dinocampus coccinellae* Paralysis Virus from Québec was generated by aligning the *de novo* generated transcripts with other picorna-like viruses. This allowed us to generate nine primer pairs in order to cover the maximum length of the draft genome (Table S2; primer pairs 1 to 9). Reverse transcription was performed with random primers as described above. Next, polymerase chain reaction (PCR) was performed using the GoTaq Hot Start Polymerase (Promega) following the standard recommended procedure. PCR was run for 35 cycles with an annealing temperature of 55°C for 30 sec and elongation at 72°C for 80 sec, followed by a final elongation at 72°C for 5 min. PCR products were sequenced by automated sequencing (GATC Biotech, Konstanz, Germany).

To complete the 5' and 3' ends of the genome sequence, we performed a Rapid Amplification for cDNA End (RACE) using the SMART RACE cDNA amplification kit (Clontech Laboratories). Briefly, the extracted RNA was reverse-transcribed with PrimeScript reverse transcriptase (Takara Bio Inc.) using different modified oligo dT primers for the 5' and 3' RACE according to the manufacturer's instructions. PCR of cDNA ends was performed with the Advantage 2 PCR system (Invitrogen) using gene-specific primers (Table S2). Each fragment was cloned into the pCR4-TOPO vector (Invitrogen), and 10 independent clones from each fragment were sent for sequencing (GATC

Biotech). The Sequencher 4.5 program (Gen codes) was used to assemble and align DNA fragments issued from the sequencing of PCR products and RACE-PCR products.

The DcPV genome sequence reported here has been deposited in the EMBL Genbank database under accession number KF843822.

Partial genome sequences of DcPV from Japan, The Netherlands, and Poland were obtained using degenerated primers (Table S2; primer pair 10). The PCR was run for 35 cycles with an annealing temperature ranging from 42 to 58°C for 30 sec and elongation at 72°C for 80 sec, followed by a final elongation at 72°C for 5 min. PCR products were sequenced by automated sequencing (GATC Biotech) and showed no difference with sequences from Québec.

DcPV abundance analysis from RNAseq data

Each sample was mapped against the reference genome of DcPV using BWA[3]. We obtained raw counts that corresponded to the number of reads mapped to the genome sequence. For visualization of differences in abundance, data were expressed as reads per kilobase per million of mapped reads (rpkm) as previously described [7].

(j) In vitro synthesis of negative- and positive-strand viral cDNA standard

To generate a strand-specific standard curve for single-stranded RT-qPCR, positive- and negative-strand cDNA were ligated within a pCR 4-TOPO vector using the TOPO TA Cloning Kit for Sequencing from Invitrogen. Plasmids were produced by cloning into Transform One Shot TOP10 competent cells grown overnight in Lysogeny Broth containing 50 ug/ml ampicillin. Plasmids were isolated using the PureLink Quick Plasmid Miniprep Kit following the manufacturer's instructions. The presence of positive- and negative-strand inserts in the plasmid was tested by Sanger sequencing of PCR products. DNA quantity was determined using an Epoch spectrophotometer (Biotek). The cloned fragments had a total molecular weight of 272.8 g/mol for a length of 4,134 bp. Thus 50 ng of plasmid DNA equals approximately 1.1×10^{10} molecules. Ten-fold

serial dilutions (10^9, 10^8, 10^7, 10^6, 10^5, 10^4, 10^3, 10^2, 10^1, 10 copies / µl) were used to generate standard curves. The standard curves generated for absolute quantification of cDNA copies of positive- and negative-strand virus had slopes of -1.55 and -1.52, respectively, and coefficient of determination (R^2) of 0.991 and 0.998 (Figure S6).

(k) Reference genes

For RT-qPCR analysis of differentially expressed genes of *C. maculata*, we had to identify reference genes with stable expressions across all biological samples. The reference genes were chosen among the non-differentially expressed transcripts in *C. maculata*. Transcripts with a mean rpkm value that exceeded five (the mean of all expression values) were considered as potential reference genes for RT-qPCR. For each of these transcripts, the rpkm values were logarithmically transformed. The coefficient of variation (CV; the ratio of the standard deviation to the mean) was used to compare the degree of variation between transcripts within different means of expression [8], and the maximum fold change (MFC, ratio of the maximum and minimum value) was used to reflect the minor variations of those candidates [9]. Genes were considered as good candidate references for RT-qPCR when they had a CV $\leq$ 0.05, a MFC $\leq$ 1.20, and a high mean rpkm value (preferably, among the top 10%). Finally, RT-qPCR was performed to identify the best candidate control genes among the 10 tested genes as previously described [2] (Figure S8). Among *C. maculata* transcripts, the elongation factor 2 and carbonic anhydrase had the most stable log (mean Ct/sample Ct). Primer pairs used are provided in Table S5.

(l) Quantitative PCR

Amplification reactions were performed in 1X Absolute Sybr green Master Mix (Applied Biosystems), forward and reverse primers (50 pmol each), and cDNA in a final volume of 10 µl. The standard cycling conditions were 95°C for 10 min followed by 45 cycles of 95°C for 10 sec, 60°C for

10 sec, and 72°C for 15 sec. Specific quantification of positive and negative strands was obtained with primers containing a 5'tag sequence as previously described [7, 11]. Quantitative PCRs were performed with the appropriate forward, reverse, or tag-specific primer pairs (Tables S4 and S5). The specificity of the negative- and positive-strand primer pairs was confirmed by melting curve analysis at the end of each qPCR run, and amplicons were verified by Sanger sequencing of the PCR product. For positive- and negative-strand virus quantification, Ct values were compared to the standard curve generated previously in order to evaluate the number of cDNA copies. The comparative threshold cycle (Ct) method was used to quantify copy numbers of the target gene in tissues. PCR amplification efficiency (E; $E=10^{(-1/slope)}$) for each primer pair was determined by linear regression analysis of a dilution series [10]. The ΔCt was calculated using the elongation factor 2 and carbonic anhydrase as reference genes. Relative gene expression was obtained by calculating the $2^{-\Delta Ct}$. The values were square root transformed to obtain a normal distribution before statistical analyses.

(m) Statistical analyses

The number of cDNA copies calculated for positive and negative strand RNA virus quantification were natural log transformed. Then, the normal distribution was confirmed with the Shapiro Wilk normality test for each dataset (see supplementary note 3). Two-sided Student's t-tests with Welch correction were used to compare virus quantification values and relative gene expression. The variation was considered significant for a q-value below 0.05 after Bonferroni correction (Supplementary note 3).

(n) Phylogenetic analyses

Phylogenetic analysis of the RdRp domains of DcPV, other Iflaviruses, Dicistroviruses and Picornaviruses was conducted from the alignement of 30 RdRp sequences by using the Maximum Likelyhood Method based in MEGA5 [11]. Initial tree(s) for the heuristic search were obtained

automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances [12]. The percentage of trees in which the taxa clustered together is shown next to the branches (based on 1000 bootstraps) for branches supported by more than 50% [13]. Branch lengths are proportional to the number of changes. DcPV is located in the Iflavirus cluster (Total list of abbreviations used and accession number are provided in Table S3)

(o) Transmission electron microscopy

Samples were fixed overnight at 4°C in 2% glutaraldehyde in 0.1 M sodium cacodylate buffer, pH 7.4. They were rinsed (3 × 15 min each) in 0.1 M sodium cacodylate buffer and fixed for 1h in 2% osmium tetroxide buffer. Following two washes in water, the samples were dehydrated in a graded EtOH series (50%, 75%, 95%, 100%, 100%; 3 × 10 min each) and infiltrated overnight in fresh 100% resin. Samples were placed into fresh resin in moulds, and polymerised in an oven at 60°C for 48h. Ultrathin 0.1 µm sections made with an LKB Ultrotome on copper grids were stained with uranyl acetate and lead citrate and examined using a Hitachi H7100 electron microscope at 80 KV.

Supplementary Note 1: The Bodyguard behavior: a neurological disorder

Previous studies have demonstrated that the ladybeetle *Coleomegilla maculata* parasitized by D*inocampus. coccinellae* displays an unusual behavior throughout the wasp pupation. It has been described as follows: "…the coccinellid, which is partially paralysed, displays a grasping behavior on top of the cocoon and twitches at irregular intervals, especially when disturbed" [1]. The ladybeetle never walks away from the stressor. In addition, paralysed ladybeetles were never seen flying or feeding, even under laboratory conditions when food is available in close proximity. Some ladybeetles survived the behavior manipulation and the associated starvation, and gradually recovered a normal behavior. As described by Maure et al. [14], the recovering ladybeetle walks slowly and its legs shake at irregular intervals. Overall, these observations suggest neurological disorders. We used behavioral assays to evaluate the level of paralysis of the parasitized ladybeetle and the extent to which the locomotor behavior is affected. The assays include a modified forced swimming test and a righting test, which involve motor structures similar to walking. These tests had previously been employed to study the host behavioral manipulation by the jewel wasp [15].

The forced swimming test was used to evaluate motor activity. Ladybeetles were placed in a closed pool of water, and swimming behavior was recorded with a camera. A healthy ladybeetle presents a 'breaststroke' movement during active swimming (Figure S1, Video S1). It uses first its hind legs (steps 1 to 2), and then its middle legs (steps 2 to 3) and front legs (steps 2 to 3)(Figure S1, Video S1). This allows healthy ladybeetles to swim very quickly. The behavior of the paralysed ladybeetle is markedly different: 9 out of 11 individuals remained static and their legs twitched at irregular intervals (Video S2). A single individual showed some ability to swim and reached the side of the pool more slowly than a healthy individual. It was able to climb out, but it remained motionless afterward. Finally, all recovering ladybeetles showed some ability to swim, although leg movements were slower than for healthy individuals and leg coordination was not perfect, resulting in a slower stroke (Video S3).

The righting test consisted of placing the ladybeetle ventral side up and measuring the length of time it took to turn over and stand upright. Healthy ladybeetles successfully righted themselves almost instantly (Video S4) whereas no parasitized ladybeetle successfully turned over within the first 20 seconds (Video S5). Finally, recovering ladybeetles were able to right themselves but took longer than healthy ladybeetles (Video S6).

Parasitized ladybeetles present numerous symptoms characteristic of neurological disorders: tremors, slow limited movement, and gait disturbance. These symptoms remain, although less intense, in recovering ladybeetles.

Supplementary Note 2: Molecular characterisation of DcPV

A combination of PCR, Race-PCR, cloning, and sequencing analyses allowed characterisation of the *Dinocampus coccinelae* paralysis Virus (DcPV) genome from *D. coccinellae* larvae from Québec, Canada. This genome is composed of 10,168 nt, not including the poly(A) tail. The sequence primers used are provided in Table S2. The DcPV genome has a low G + C content (32.9% A, 15.6% C, 19.2% G, and 31.8% U) and contains one large Open Reading Frame (ORF), 820–9840 nt, encoding a predicted polyprotein of 3007 aa (Figure 2, Figure S2). The structural proteins are encoded by the 5' half of the sequence, and the non-structural helicase, protease, and RNA-dependent RNA polymerase (RdRp) are encoded by the 3' half of the sequence. The 5' and the 3' non-translated regions (NTR) have lengths of 819 and 325 nt, respectively (without the poly(A) tail) (Figure 2, Figure S2).

Based on the predicted polyprotein, the predicted structural and non-structural DcPV proteins contain functional motifs and domains shared by picorna-like viruses and picornaviruses of the order Picornavirales (Figure 2) [16]. DcPV RdRp (aa 2920-3220) contains the eight domains originally identified by Koonin et al. [17] and found in all RdRp from these families of viruses. The three domains of Helicase described by Koonin et al. [17] (aa 1834-1986) were also detected, but the most interesting feature is the presence of the nucleoside triphosphate binding residue GxxGxGKS of domain A that is highly conserved. In addition, both the cysteine protease motif GxCG and the substrate-binding residues GxHxxG described by Gorbalenya and Koonin [18] of the 3C protease are conserved (aa 2555-2730) (Figure S3). In addition, VP2, VP3, and VP1 were found in positions 596–778, 872–1076, and 1296–1494 of the polyprotein, respectively (Figure 2, Figure S2), and the key characteristic motifs of picornaviruses were also identified: NxNxFQxG for VP2, $WxGxLx_3FxFx_7Gx_5YxP$ for VP3, and FxRG for VP1 (Figure S4) [19].

Protease cleavage sites have been identified in DWV [20] and PnPV [21]. Sequence alignment and manual screening for classic patterns of viral 3C proteases allowed the prediction of protease

cleavage sites of DcPV (Figure S2). All predicted sites are suitably located for the cleavage of functional viral proteins, including the predicted VP2, VP3, VP1, Hel, Pro, and RdRp. The majority of putative cleavage sites share some similarities (three times PQ/MD, three times PQ/ME, one time PQ/ID, one time PQ/YK, one time PQ/LH), providing insight into the involvement of 3C protease in DcPV polyprotein processing (Figure S2). The VP4/VP1 processing site does not share the same consensus sequence (GLNR/DNPP), suggesting that it could involve another protease family (Figure S3). This last motif (NX/DXP motif), located between VP4 and VP1, was also found in other Iflaviruses [20, 22-24]. In addition, a leader protein (L) has been reported upstream of the capsid region in different picorna-like viruses [19, 20, 22, 24]. The predicted cleavage site between L and the VP2 site is unclear in DcPV (PQ/MD or QM/DV), but it is highly similar to the one predicted in DWV: PE/MD or EM/DN [20]. Finally, a VpG similar to DWV VpG was found in the DcPV polyprotein upstream of the protease C3 [20]. The 27 amino acid DcPV VpG contains the Y residue necessary for RNA interaction (Figure S2) [25].

The secondary RNA structure prediction was conducted for the DcPV 5' NTR of 819 nt using the RNAfold algorithm [26]. The DcPV 5' NTR displays 10 putative dominant structural features. Most of them were hairpins with a branched, cloverleaf-like structures (structure VI) (Figure S5). Such cloverleaf-like structures have also been found in other Iflaviruses, such as VDW and VDV, and in entero-/rhinoviruses such as Poliovirus [24, 27]. The first and third hairpin structures had the 5'-AUUU-3' loop similar to hairpin structures found in other Iflaviruses (Figure S5) [24, 28]. These conserved structural features suggest that the DcPV IRES is located in this region.

Supplementary Note 3: Statistical analyses

Regarding the timepoint Be D13, we collected either eggs or larvae depending on the sample. In eggs, none, or very limited amount of virus was quantified whereas in larvae, the virus was abundant. This resulted in a high variance in positive and negative strand DcPV genomes at this time point and the sample did not fit the normal distribution. This sample was not included in the statistical analyses of viral genomes.

Timepoint Be D5 did not fit the normal distribution for the gene Ago 2 and this samples was not included in the statistical analysis.

For the gene r2d2, the distribution was nor normal when abdomen and head samples were analyzed together. However, a normal distribution was obtained for abdomen and head samples separately. It may be that the regulation of r2d2 gene expression in these two tissues is based on different mechanisms. Since no comparison of expression between head and abdomen sample was needed, the analysis was not impacted.

For all other datasets, the normal distribution of data was confirmed and all samples were included for statistical analyses using the two-sided student T test with Welch correction (See Results below; significant differences are indicated in **bold**).

Statistics for Positive Strand viral genomes

Data normalization: Natural logarithm

Shapiro-Wilk normality test
caluclated after removal of samples E/L D13
W = 0.9674. p-value = 0.08831

Two-sample T test with Welch correction and Bonferroni correction

in *D. coccinella*e		statistic	p-value	**q-value**
E D5	L Be D20	67,09	0	**0**
L Be D20	L Ae	1,44	0.183741	0.551223
L Ae	Adult	4,43	0.002532	**0.007596**
in *C. maculata*				
Abdomens				
He	Res	3,07	0.74688	0.13015
He	Be D5	2,86	0.75151	0.115
Be D5	Be D20	25,45	0.00002	**0**
Be D20	Ae	3,71	0.30437	0.05655

Ae	R	-0,61	0.08453	2.78545
Heads				
He	Be D5	3,33	0.00238	0.52896
Be D5	Be D20	20,09	0.0001	**0.02392**
Be D20	Ae	2,65	0.07603	0.17536
Ae	R	-5,22	0.00726	**0.02724**
Abdomens versus Heads				
Ab Be D20	H Be D20	-5,02	0.05131	0.10262

Statistics for Negative Strand replication intermediates

Data normalization: Natural logarithm

Shapiro-Wilk normality test
caluclated after removal of samples E/L D13
W = 0.9764. p-value = 0.1874

Two-sample T test with Welch correction and Bonferroni correction

in *D. coccinella*e		statistic	p-value	**q-value**
E D5	L Be D20	44,54	0	**0**
L Be D20	L Ae	-0,05	0.964487	2.893461
L Ae	Adult	5,17	0.007307	**0.021921**

in *C. maculata*				
Abdomens				
He	Res	-0,34	0.02603	3.7344
He	Be D5	0,33	0.0234	3.75755
Be D5	Be D20	-10,70	0	**0.0001**
Be D20	Ae	-1,08	0.01131	1.52185
Ae	R	1,95	0.55709	0.42265
Heads				
He	Be D5	9,62	0.13224	**0.00952**
Be D5	Be D20	28,03	0.00598	**0.0004**
Be D20	Ae	2,14	0.04384	0.30412
Ae	R	-4,42	0.00681	**0.02904**
Abdomens versus Heads				
Ab Be D20	H Be D20	2,30	0.001	**0.002**

Statistics for Ago2 gene expression

Data normalization: Square root

Shapiro-Wilk normality test

calculated after removal of sample Ab Be D5
W = 0.9816. p-value = 0.5342

Two-sample T test with Welch correction and Bonferroni correction

in *C. maculata*

Abdomens		statistic	p-value	**q-value**
He	Res	-4,92	0.0031	**0.0155**
He	Be D13	1,66	0.13624	0.6812
He	Be D20	1,46	0.17541	0.87705
He	Ae	1,70	0.12889	0.64445
He	R	1,42	0.19806	0.9903
Heads				
He	Be D5	7,78	0.00037	**0.00185**
He	Be D13	7,88	0.00134	**0.0067**
He	Be D20	1,13	0.01935	0.09675
He	Ae	-1,30	0.36087	1.80435
He	R	0,00	0.05823	0.29115

Statistics for C3PO gene expression
Data normalization: Square root

Shapiro-Wilk normality test
W = 0.9864. p-value = 0.7351

Two-sample T test with Welch correction and Bonferroni correction

in *C. maculata*

Abdomens		statistic	p-value	**q-value**
He	Res	-2,27	0.21978	1.31868
He	Be D5	0,05	0.95943	5.75658
He	Be D13	-0,80	0.46249	2.77494
He	Be D20	-2,90	0.01656	0.09936
He	Ae	-1,11	0.29479	1.76874
He	R	-5,70	0.00035	**0.0021**
Heads				
He	Be D5	0,01	0.99438	4.9719
He	Be D13	-0,03	0.97552	4.8776
He	Be D20	0,66	0.53762	2.6881
He	Ae	-0,21	0.83845	4.19225
He	R	-0,38	0.71275	3.56375

Statistics for Dicer2 gene expression

Data normalization: Square root

Shapiro-Wilk normality test
W = 0.9802. p-value = 0.4261

Two-sample T test with Welch correction and Bonferroni correction

in *C. maculata*

Abdomens		statistic	p-value	**q-value**
He	Res	-3,85	0.01083	0.06498
He	Be D5	1,12	0.30212	1.81272
He	Be D13	2,41	0.04274	0.25644
He	Be D20	2,08	0.0772	0.4632
He	Ae	4,02	0.00655	**0.0393**
He	R	3,78	0.00847	0.05082
Heads				
He	Be D5	9,53	0.00023	**0.00115**
He	Be D13	7,22	0.00108	**0.0054**
He	Be D20	2,04	0.07949	0.39745
He	Ae	0,99	0.35778	1.7889
He	R	2,31	0.0789	0.3945

Statistics for PI3K gene expression

Data normalization: Square root

Shapiro-Wilk normality test
W = 0.9795. p-value = 0.396

Two-sample T test with Welch correction and Bonferroni correction

in *C. maculata*

Abdomens		statistic	p-value	**q-value**
He	Res	-11,25	0.0001	**0.0006**
He	Be D5	-1,62	0.15709	0.94254
He	Be D13	-2,16	0.06525	0.3915
He	Be D20	2,02	0.07699	0.46194
He	Ae	1,53	0.16661	0.99966
He	R	-0,15	0.88333	5.29998
Heads				
He	Be D5	12,78	0.00026	**0.0013**

He	Be D13	10,34	0.00015	**0.00075**
He	Be D20	1,27	0.25005	1.25025
He	Ae	-0,62	0.55896	2.7948
He	R	-0,19	0.85107	4.25535

Statistics for R2D2 gene expression
Data normalization: Square root

Shapiro-Wilk normality test
Abdomens Only
W = 0.9592. p-value = 0.2308
Shapiro-Wilk normality test
Heads Only
W = 0.9325. p-value = 0.07939

Two-sample T test with Welch correction and Bonferroni correction

in *C. maculata*

Abdomens		statistic	p-value	**q-value**
He	Res	-6,60	0.00059	**0.00354**
He	Be D5	0,84	0.44645	2.6787
He	Be D13	0,85	0.427	2.562
He	Be D20	1,87	0.09269	0.55614
He	Ae	2,62	0.03022	0.18132
He	R	2,47	0.04907	0.29442
Heads				
He	Be D5	1,77	0.20002	1.0001
He	Be D13	3,56	0.03611	0.18055
He	Be D20	6,77	0.00042	**0.0021**
He	Ae	7,38	0.00051	**0.00255**
He	R	6,76	0.00015	**0.00075**

Statistics for Toll7 gene expression
Data normalization: Square root

Shapiro-Wilk normality test
W = 0.9716. p-value = 0.1664

Two-sample T test with Welch correction and Bonferroni correction

in *C. maculata*

Abdomens		statistic	p-value	**q-value**
He	Res	-14,40	0.00007	**0.00042**
He	Be D5	-2,11	0.08695	0.5217
He	Be D13	-2,26	0.05661	0.33966
He	Be D20	2,51	0.03171	0.19026
He	Ae	3,12	0.0138	0.0828
He	R	-0,03	0.97638	5.85828
Heads				
He	Be D5	12,94	0.00005	**0.00025**
He	Be D13	10,12	0.00025	**0.00125**
He	Be D20	1,57	0.16823	0.84115
He	Ae	-0,10	0.92624	4.6312
He	R	0,21	0.84262	4.2131

Figure S1. The 'breaststroke' movements of ladybeetles during active swimming. We identified three steps involving, in order, the hind, middle, and front legs.

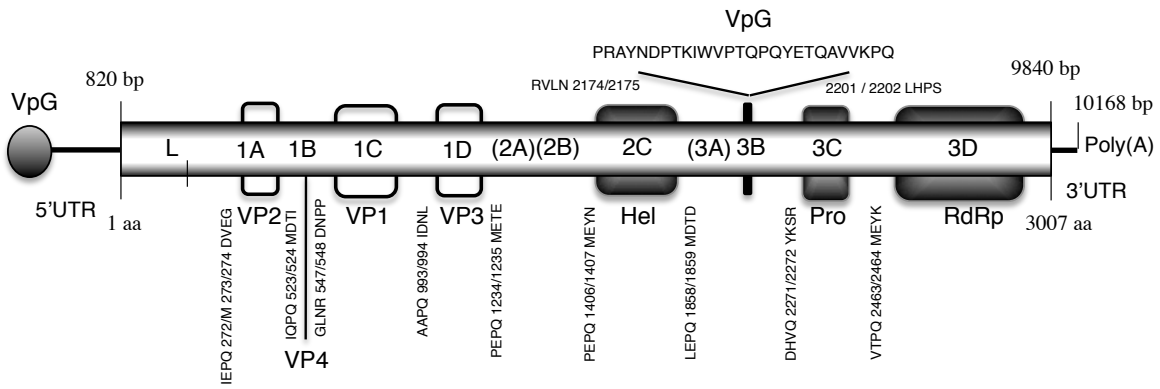

Figure S2 : Schematic diagram of the predicted DcPV genome structure. Numbers on the top indicate nucleotide positions, numbers on the bottom indicate amino acid positions, and the long shaded box represents the single Open Reading Frame (ORF). Predicted protease cleavage sites are indicated together with their position and surrounding amino acid sequences. A leader sequence was found upstream of the viral capsid proteins (VP). Predicted proteins are indicated using the L434 nomenclature system. Boxes indicate the position of recognizable protein domains of structural proteins (VP 1 to 4; open boxes) and non-structural proteins (dark boxes). A putative VpG sequence was found downstream of the Helicase (Hel) and upstream of the protease (Pro) and RNA-dependent RNA polymerase (RdRp).

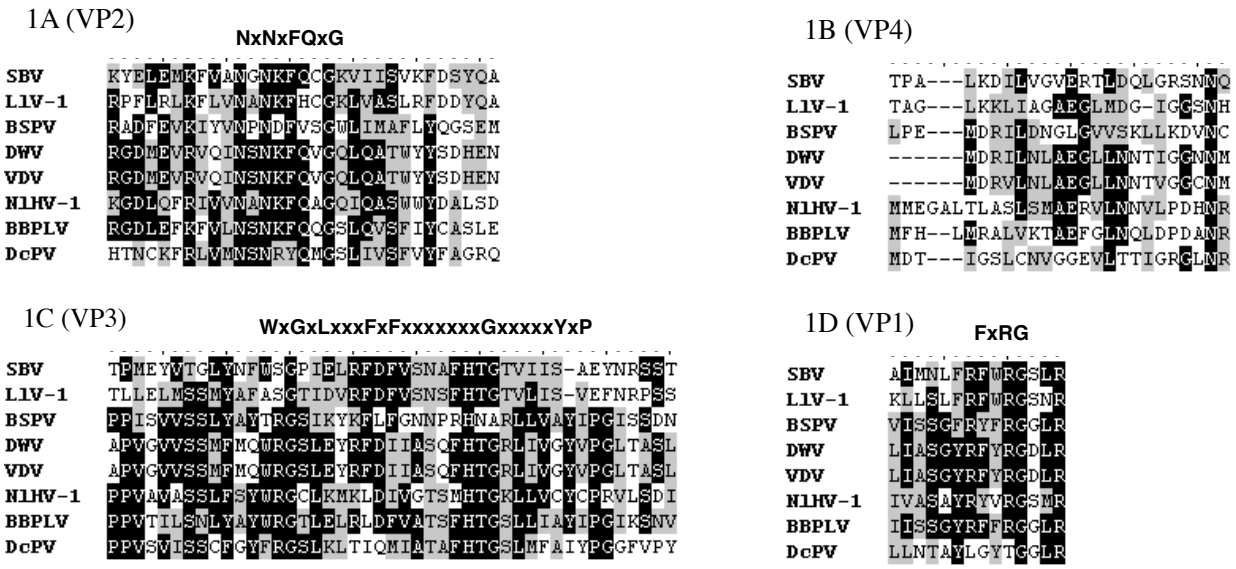

Figure S3. Multiple sequence alignment of the structural proteins of DcPV and other Iflaviruses. The key motifs present in capsid proteins 1A, 1C, and 1D of picornaviruses are indicated [19]. The predicted full-length sequence of capsid protein 1B is provided.

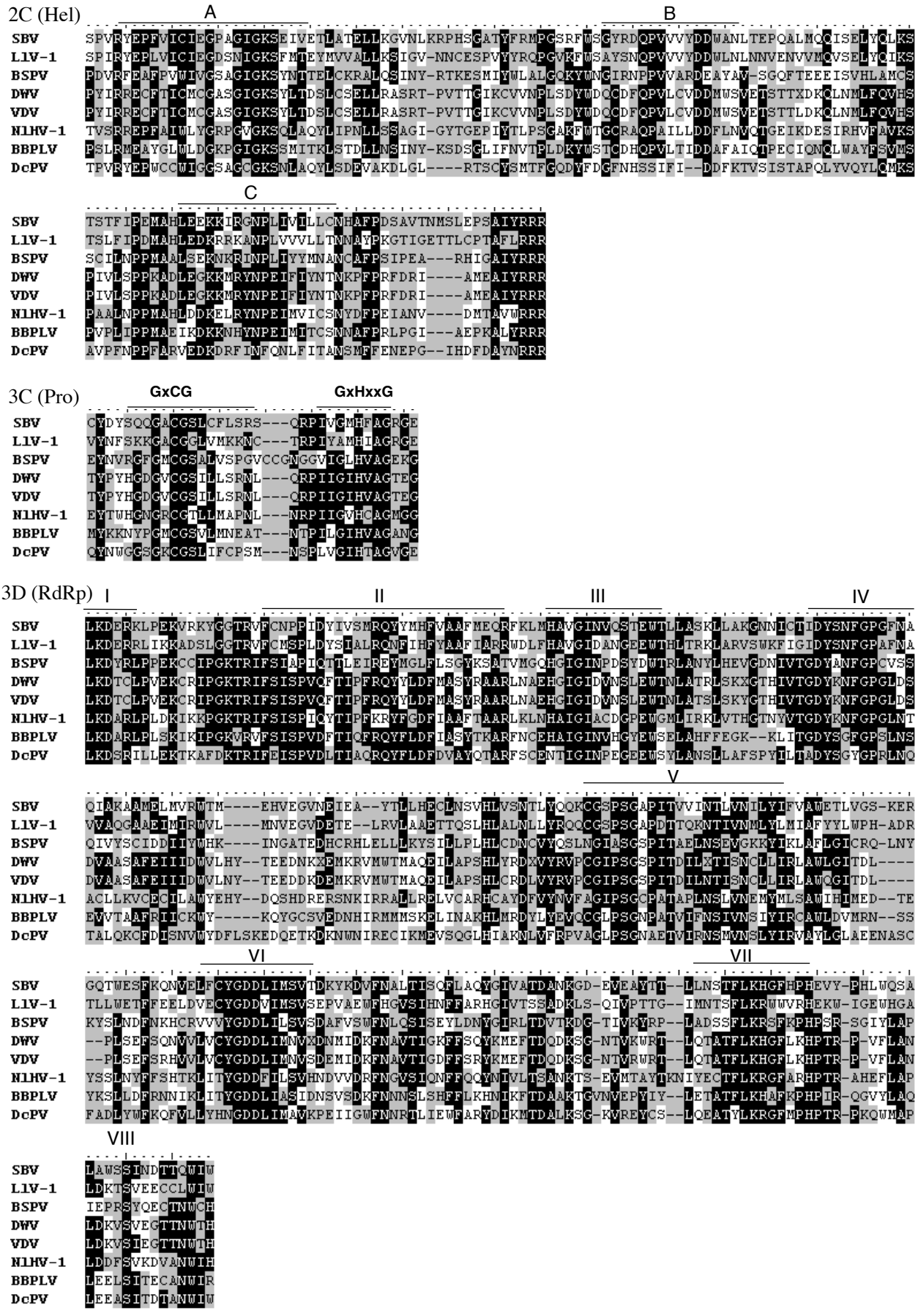
2C (Hel)
A
B
C
3C (Pro)
GxCG
GxHxxG
3D (RdRp)
I
II
III
IV
V
VI
VII
VIII
SBV
LlV-1
BSPV
DWV
VDV
NlHV-1
BBPLV
DcPV

Figure S4. Multiple sequence alignment of the non-structural proteins of DcPV and other Iflaviruses. The putative Helicase contained the three domains originally identified by Koonin and Dolja [17], but the second and third domains had weak homologies. The putative protease had one core domain of picornaviruses and the substrate-binding residue [18]. RdRp had all eight motifs of the core region of described by Koonin and Dolja [17].

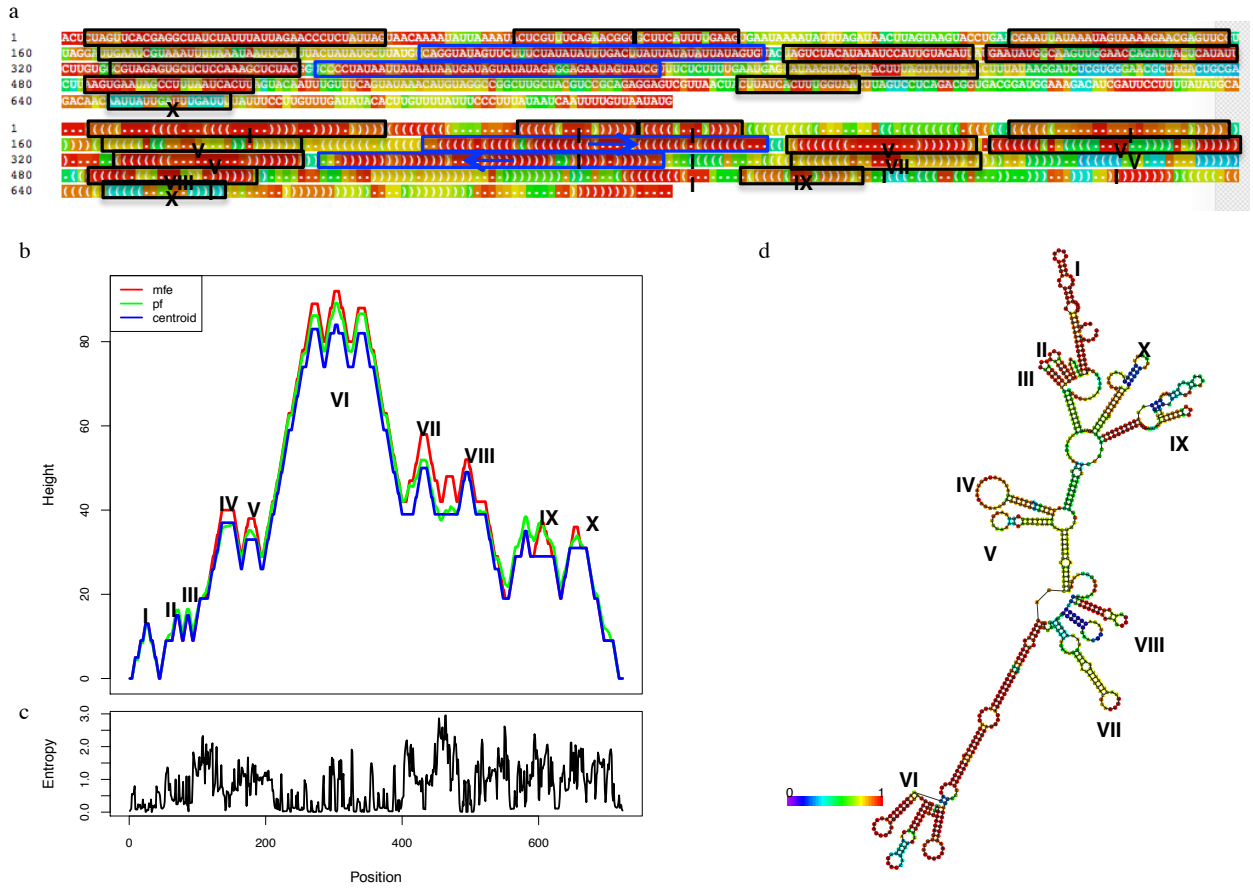

Figure S5. The optimal secondary structure of the 5'UTR as predicted using the RNAfold software. (a) Maximum free energy structure in Vienna format. Rectangles indicate the 10 predicted hairpins. (b) Mountain plot representation of the maximum free energy (MFE, red) structure, the thermodynamic ensemble of RNA structures (pf, green), and the centroid structure (centroid, blue). (c) Entropy for each position. (d) Maximum free energy structure encoding base pair probabilities.

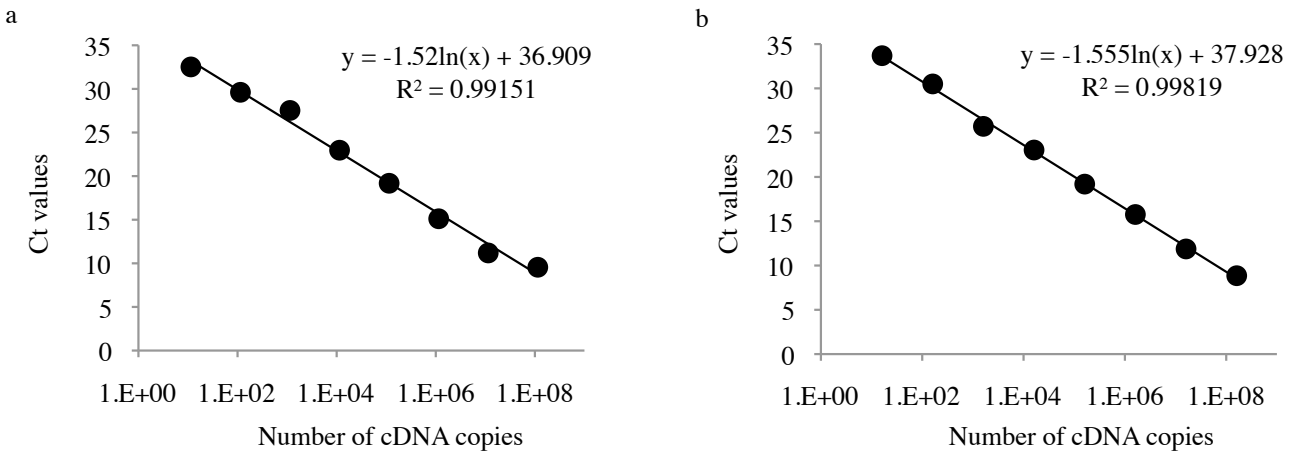

Figure S6. Standard curves for negative- (**a**) and positive- (**b**) strand RT-qPCR.

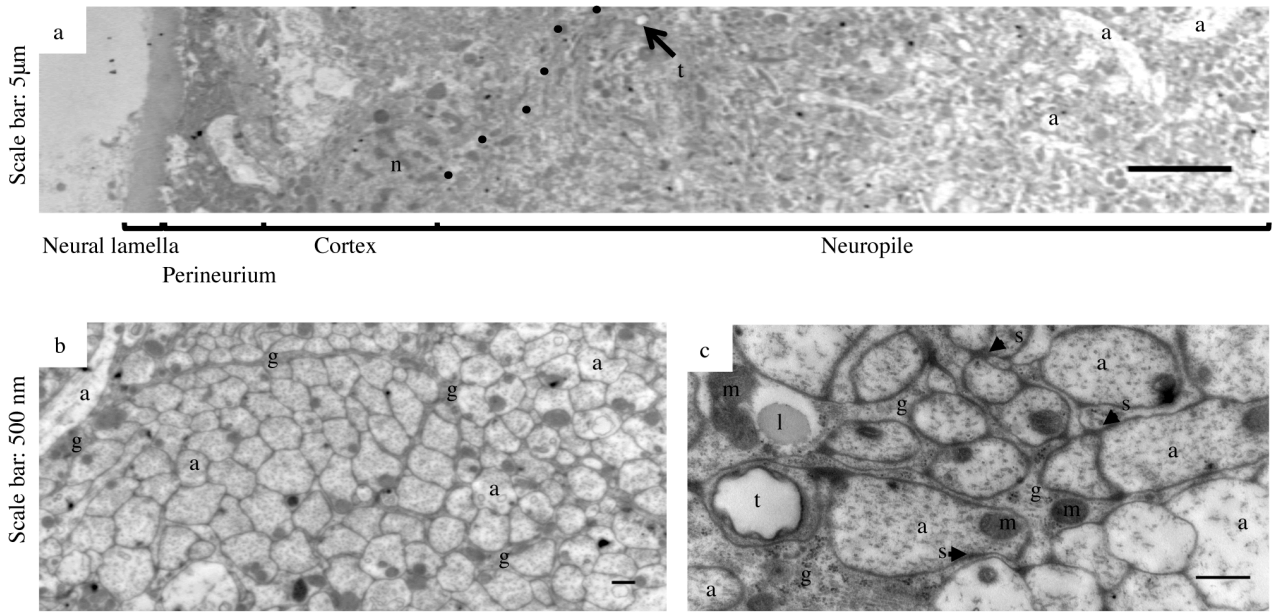

Figure S7 : General structure of the cerebral ganglia of a healthy ladybeetle. (A) Architecture of the cerebral ganglia. The dotted line separates the cortex from the neuropile. (B, C) Glia wrapped around axon bundles and individual axons. a, axon; m, mitochondria; g, glia; t, trachea; n, nucleus; l, lipid droplet, s: synapse. Scale bars represent 5 µm (A) or 500 nm (B and C).

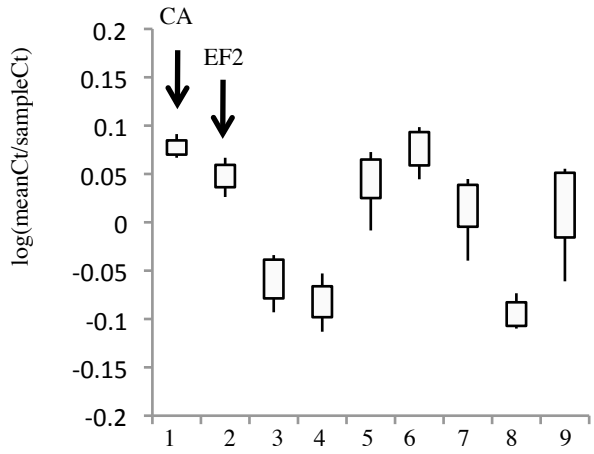

Figure S8. **Selection of reference genes for RT-qPCR analyses of *C. maculata* genes**. The Figure shows the distribution of log(meanCt/sampleCt) of the nine pre-selected candidate reference genes. The two reference genes chosen had the most stable meanCt/sampleCt. CA: Carbonic anhydrase; EF2: Elongation Factor 2

Table S1: Number of reads in each library*. The table provides the number of sequenced paired end (PE) reads (raw data), the number of PE reads that passed the default high quality filtering performed by the Illumina pipeline (Quality filter) and the number of PE reads used for assembly and mapping after read trimming (Read trimming).

	Raw data	Quality filter	Read trimming
Ab He	45,333,036	42,185,603	39,703,653
Ab Be D20	40,580,985	37,734,776	35,667,579
Ab Ae	41,533,795	37,584,672	35,463,692
Ab R	43,643,475	40,533,588	38,232,208
H He	46,969,093	43,783,704	41,245,004
H Be D20	46,952,616	43,658,295	41,087,390
H Ae	51,725,733	48,225,632	45,478,449
H R	45,686,729	42,522,845	40,043,104
L Be D20 1	38,537,049	35,614,855	33,775,620
L Be D20 2	48,259,946	43,761,195	41,636,805
L Ae 1	52,455,633	47,663,804	45,304,485
L Ae 2	65,841,503	59,715,331	56,875,865

***** For RNA sequencing, equimolar amounts of pools of males and females were pooled for each category leading to a total of 12 samples: AbHe; Abdomens of healthy ladybeetles, Ab Be D20; Abdomens of parasitized ladybeetle collected before egression, 20 days after oviposition, Ab Ae; Abdomens of parasitized ladybeetle collected immediately after egression, Ab R Abdomens of parasitized ladybeetle recovering from the manipulation, H He; Heads of healthy ladybeetles, H Be D20; Heads of parasitized ladybeetle collected before egression, 20 days after oviposition, H Ae; Heads of parasitized ladybeetle collected immediately after egression, H R Heads of parasitized ladybeetle recovering from the manipulation, L Be D20 1 and 2; replicates of Larva dissected out of ladybeetles before egression, 20 days after oviposition and L Ae 1 and 2; replicates of Larva collected after egression from the ladybeetle.

Table S2. Nucleotide sequences of primers used for PCR sequencing and RACE-PCR of the DcPV.

F1	TAACTTAGTAAGTACCTGAACG
R1	CGACCTTTAATGCTAGATCATA
F2	GAGCAATTTCCTCCACAATGC
R2	GCAACCGATGTCTTTAGTTCC
F3	ATGTACCAGCAAAGGAGAGC
R3	GCTATTACCTATGGAGTATATG
F4	GTGTCTGATGGATGGTCTGGT
R4	GGATACATTGTCCAGTAGATG
F5	TTAGAGTGGAAGTGCCTTGG
R5	GCAAGCTCTGAGCGTGAAC
F6	ATTGAAGAGAGTATTTCATCGC
R6	CCGTCGATAGTAGCGTAATTT
F7	TGCTGATCGGTGTCAAAAGC
R7	CCAGATTCGTGTCATCAAGC
F8	TGTGGAGTGGGTTATGTACC
R8	GGTGGATCAGGAAAATGTGG
F9	GAAGGATTTCAAGATGCTGC
R9	GCACAAAGGCAGTATTTCTTAG
F10	YMTTAAAATGACKGAYGC
R10	GTTCCYTCCCADACATGA
5’ RACE	AAGTCGCAGTCTAGCGTTCCCACG
3’ RACE	GCTCCCAATGTGTTGTACCAAGCTCG

Table S3. Picornaviruses, Dicistroviruses, and Iflaviruses included in the phylogenetic analysis and their Genbank accession numbers.

Abbreviation	Virus species name	Genbank Number
ALPV	Aphid lethal paralysis virus	NC_004365
RPV	*Rhopalosiphum pad*i virus	NC_001874
HCV	*Homalodisca coagulat*a virus-1	NC_008029
TV	*Triatoma* virus	NC_003783
TSV	Taura syndrome virus	JX094350
MCD	Mud crab dicistrovirus	NC_014793
BQCV	Black queen cell virus	NC_003784
HPV	*Himetobi* P virus	NC_003782
DCV	Drosophila C virus	NC_001834
PSIV	*Plautia stali* intestine virus	NC_003779
SIV	*Solenopsis invicta* virus	NC_006559
ABPV	Acute bee paralysis virus	NC_002548
NvitV-1	*Nasonnia vitripennis* virus isolate 1	FJ790486
VcPLV	*Venturia canescens* picorna-like virus	AY534885
BbPLV	*Brevicoryne brassicae* picorna-like virus	NC_009530
DWV	Deformed wing virus	NC_004830
KV	Kakugo virus	NC_005876
VDV	*Varroa destructor* virus	NC_006494
SBV	Sacbrood virus	NC_002066
EoV	*Extropis obliqua* virus	NC_005092
PnV	*Perina nuda* virus	AF323747

SeIV-1	*Spodoptera exigua* Iflavirus-1	JN_091707
SeIV-2	*Spodoptera exigua* Iflavirus-2	JN_870848
IFV	Infectious Flatcherie Virus	AB_000906
BSPV	Bee Slow Paralysis Virus	EU_035616
NlHV-1	*Nilaparvata lugens* Honeydew Virus	AB_766259
LlV-1	*Lygus lineolaris* Virus-1	JF_720348
EMCV	Encephalomyocarditis virus	NC_001479
Poliovirus	Poliovirus	NC_002058

Table S4. Nucleotide sequences of primers used for single-strand reverse transcription (RT) and qPCR. The non-DcPV sequence (5'Tag) is shown in boldface.

Negative strand detection	
Forward for RT	**GGCCGTCATGGTGGCGAATAA**GCTATACTCAGGCAGCAC
Forward for qPCR = Tag S	**GGCCGTCATGGTGGCGAATAA**
Reverse for qPCR	AGTACCCGAAATCGAACAG
Positive strand detection	
Reverse for RT	**GGCCGTCATGGTGGCGAATAA**AGTACCCGAAATCGAACAG
Forward for qPCR	GCTATACTCAGGCAGCAC
Reverse for qPCR = Tag S	(see sequence above)

Table S5. Nucleotide sequences of primers used for RT-qPCR analysis.

F_CA	TGACCAGTGGGAAATGGAT
R_CA	CGTCCTACATACACCAACTTGAA
F_Ef2	AACGACTCATTTACTGGCAA
R_Ef2	TTATGGAGCCCGTATATCTCTG
F_Dicer2	TAGTGAAATGGATGAACAATTAGAGT
R_Dicer2	TCTTTACCTATCAAATCTCCACCC
F_Ago2	ATGACTATTGAGGAGTATTTCCAG
R_Ago2	GGTTGACCGCTTGATGT
F_R2D2	CTACACTTGGATCATCAGGTTC
R_R2D2	TCGGCTGTTGGGATAGG
F_C3PO	TGAACGTCACCAACCAAG
R_C3PO	CATTAAGGCACTCTACCGA
F_Toll7	ACATGGAATGGCTACAACTAA
R_Toll7	TTCGTAAGTACATAGAAAGTCGC
F_PI3K	ATGTTGGAGAAATTACCGCC
R_PI3K	GGTTTCCAAGAGCGAAGCATA

References

1. Maure F., Brodeur J., Ponlet N., Doyon J., Firlej A., Elguero E., Thomas F. 2011 The cost of a bodyguard. *Biol Lett* **7**(6), 843-846.
2. Dheilly N.M., Lelong C., Huvet A., Favrel P. 2011 Development of a Pacific oyster (*Crassostrea gigas*) 31,918-feature microarray: identification of reference genes and tissue-enriched expression patterns. *BMC Genomics* **12**(1), 468.
3. Li H., Durbin R. 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**(14), 1754-1760.
4. Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R., Subgroup G.P.D.P. 2009 The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**(16), 2078-2079. (doi:10.1093/bioinformatics/btp352).
5. Anders S., Huber W. 2010 Differential expression analysis for sequence count data. *Genome Biology* **11**(10), R106.
6. .
7. Mortazavi A., Williams B.A., McCue K., Schaeffer L., Wold B. 2008 Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**(7), 621-628.
8. Novak J.P., Sladek R., Hudson T.J. 2002 Characterization of variability in large-scale gene expression data: implications for study design. *Genomics* **79**(1), 104-113.
9. de Jonge H.J.M., Fehrmann R.S.N., de Bont E.S.J.M., Hofstra R.M.W., Gerbens F., Kamps W.A., de Vries E.G.E., van der Zee A.G.J., te Meerman G.J., ter Elst A. 2007 Evidence Based Selection of Housekeeping Genes. *PLoS ONE* **2**(9), e898.
10. Muller P.Y., Janovjak H., Miserez A.R., Dobbie Z. 2002 Processing of gene expression data generated by quantitative real-time RT-PCR. *Biotechniques* **32**(6), 1372-1374, 1376, 1378-1379.
11. Tamura K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S. 2011 MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* **28**(10), 2731-2739. (doi:10.1093/molbev/msr121).
12. Jones D.T., Taylor W.R., Thornton J.M. 1992 The rapid generation of mutation data matrices from protein sequences. *Computer applications in the biosciences : CABIOS* **8**(3), 275-282. (doi:10.1093/bioinformatics/8.3.275).
13. Felsenstein J. 1985 Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**, 783-791.
14. Maure F., Doyon J., Thomas F., Brodeur J. 2014 Host behavior manipulation as an evolutionary route toward attenuation of parasitoid virulence *J Evol Biol* **27**(12), 2871-2875.
15. Libersat F., Gal R. 2013 What can parasitoid wasps teach us about decision-making in insects? *J Exp Biol* **216**(1), 47-55.
16. Le Gall O., Christian P., Fauquet C., King A.Q., Knowles N., Nakashima N., Stanway G., Gorbalenya A. 2008 Picornavirales, a proposed order of positive-sense single-stranded RNA viruses with a pseudo-T=3 virion architecture. *Arch Virol* **153**(4), 715-727.
17. Koonin E.V., Dolija V.V., Morris T.J. 1993 Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Crit Rev Biochem Mol Biol* **28**(5), 375-430.
18. Gorbalenya A.E., Koonin E.V. 1989 Viral proteins containing the purine NTP-binding sequence pattern. *Nucleic Acids Res* **17**(21), 8413-8438.
19. Liljas L., Tate J., Lin T., Christian P., Johnson J.E. 2002 Evolutionary and taxonomic implications of conserved structural motifs between picornaviruses and insect picorna-like viruses. *Arch virol* **147**(1), 59-84.
20. Lanzi G., de Miranda J.R., Boniotti M.B., Cameron C.E., Lavazza A., Capucci L., Camazine S.M., Rossi C. 2006 Molecular and Biological Characterization of Deformed Wing Virus of Honeybees (*Apis mellifera* L.). *J Virol* **80**(10), 4998-5009.

21. Wu C.-Y., Lo C.-F., Huang C.-J., Yu H.-T., Wang C.-H. 2002 The complete genome sequence of *Perina nuda* Picorna-like Virus, an insect-infecting RNA Virus with a genome organization similar to that of the mammalian Picornaviruses. *Virology* **294**(2), 312-323.
22. Choi H., Sasaki T., Tomita T., Kobayashi M., Kawase S. 1992 Processing of structural polypeptides of Infectious flatcherie virus of silkworm, Bombyx mori: VP1 and VP4 are derived from VP0. *J Invertebr Pathol* **60**(2), 113-116.
23. Isawa H., Asano S., Sahara K., Lizuka T., Bando H. 1998 Analysis of genetic information of an insect picorna-like virus, infectious flacherie virus of silkworm: evidence for evolutionary relationships among insect, mammalian and plant picorna(-like) viruses. *Arch Virol* **143**(1), 127-143.
24. Ongus J.R., Roode E.C., Pleij C.W.A., Vlak J.M., van Oers M.M. 2006 The 5' non-translated region of Varroa destructor virus 1 (genus Iflavirus): structure prediction and IRES activity in *Lymantria dispar* cells. *J Gen Virol* **87**(11), 3397-3407.
25. Weitz M., Baroudy B.M., Maloy W.L., Ticehurst J.R., Purcell R.H. 1986 Detection of a genome-linked protein (VPg) of hepatitis A virus and its comparison with other picornaviral VPgs. *J Virol* **60**(1), 124-130.
26. Gruber A.R., Lorenz R., Bernhart S.H., Neuböck R., Hofacker I.L. 2008 The Vienna RNA Websuite. *Nucleic Acids Res* **36**(suppl 2), W70-W74.
27. Witwer C., Rauscher S., Hofacker I.L., Stadler P.F. 2001 Conserved RNA secondary structures in Picornaviridae genomes. *Nuc Acids Res* **29**(24), 5079-5089. (doi:10.1093/nar/29.24.5079).
28. Ryabov E.V. 2007 A novel virus isolated from the aphid *Brevicoryne brassicae* with similarity to Hymenoptera picorna-like viruses. *J Gen Virol* **88**(9), 2590-2595.