**Effect of using Green Fluorescent Protein double-stranded RNA as non-target negative control in *Nasonia vitripennis*** **RNA interference assays**

Julien Rougeot1$, Yidong Wang1, Eveline C. Verhulst1

1Laboratory of Entomology, Wageningen University, P.O. Box 16, 6700 AA Wageningen, The Netherlands.

$Corresponding author: julien.rougeot@wur.nl.

Keywords: RNA interference, *Nasonia*, RNA sequencing

**Supplementary Files**

**Supplementary Table 1:** Statistics on alignment and mapping for all RNAseq samples.

**Supplementary Table 2:** Differential expression analysis between water-injected samples and uninjected samples as calculated by DESeq2 (see Supplementary Methods, Data analysis). Genes up-regulated in water-injected samples (log2Fold Change≥0 and *padj*<0.05) are highlighted in turquoise whereas down-regulated genes in water-injected samples (log2Fold Change≤0 and *padj*<0.05) are highlighted in red.

**Supplementary Table 3:** Differential expression analysis between *GFP*-i samples and control samples as calculated by DESeq2 (see Supplementary Methods, Data analysis). Genes up-regulated in *GFP*-i samples (log2Fold Change≥0 and *padj*<0.05) are highlighted in turquoise whereas down-regulated genes in *GFP*-i (log2Fold Change≤0 and *padj*<0.05) are highlighted in red. Genes in bold are differentially expressed genes described as potentially involved in microtubule and sperm functions.

**Supplementary Table 4:** Results from blastn (https://blast.ncbi.nlm.nih.gov/) with the *GFP* dsRNA sequence as query and all *Nasonia vitripennis* mRNA as subjects. Settings used were “somewhat similar sequences” and a word size of 7 without filtering for low complexity regions.

**Supplementary Methods:**

*Nasonia vitripennis* rearing

The *N. vitripennis* lab strain AsymCX cured with *Wolbachia* infection (Werren et al., 2010) was used throughout the experiments. Wasps were reared on *Calliphora sp.* hosts and cultured at 25 °C at a L16:D8 cycle.

Larval RNAi and sample collection

Larval RNAi knockdown was induced in AsymCX males in second instar larval stage (approximately 3 days after egg laying) (Werren et al., 2009). *GFP* dsRNA was generated from the vector *pOPINEneo-3C-GFP* (Addgene plasmid # 53534; <https://www.addgene.org/53534/>; RRID: Addgene\_53534, gift from Ray Owens). Amplification by PCR using GoTaq Flexi DNA polymerase (Promega) with *GFP\_RNAi\_F* (*5’-GTGACCACCTTGACCTACG-3’*) and *GFP\_RNAi\_R* (*5’-TCTCGTTGGGGTCTTTGCT-3’*) primers produced a 460 base-pair long amplicon covering 64% of the *Emerald GFP* CDS. The sequence of the amplified fragment is 5' -GTGACCACCTTGACCTACGGCGTGCAGTGCTTCGCCCGCTACCCCGA CCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAAGGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGA- 3'. This PCR product was then amplified again in two separate PCR to add the minimal T7 promoter sequence (5’-TAATACGACTCACTATAGGG-3’) to either ends of the amplicon. The two templates were used in separate reactions to transcribe both sense and anti-sense RNA molecules, which were then mixed and annealed in equimolar amounts, using the MEGAscript RNAi kit (Ambion, Austin, Texas, USA) according to the manufacturer's protocol. Before injection, second instar male larvae were placed on a petri-dish covered with 1% agarose in 1x PBS. Larvae were injected in the posterior part with 4 μg/μl *GFP* dsRNA diluted in milliQ water mixed with red food dye in the ratio 9:1, milliQ water with red food dye, or non-injected as control. Injections were performed using custom made capillary needles and a IM300 Microinjecctor (Narishige) according to the protocol by Werren et al. (2009). Larvae were injected with the dsRNA mixture until a red spot appeared in their posterior end. After injection, the larvae were transferred back to host (6-8 per host) and sealed with the host shell. After 5 to 6 days post-injection, three samples (five individuals per sample) of white pupal stage males per category were collected, snap-frozen in liquid nitrogen, and stored at −80 °C until RNA extraction.

RNA extraction

Total RNA of larvae pools was extracted with Quick-RNA Tissue/Insect Kit coupled with on column DNase treatment (ZymoResearch - R2030) according to manufacturer's protocol. RNA was resuspended in 10 µl DNase RNase free water and RNA concentration was measured on fluorometer (Qubit 2.0 – Life Technologies).

Library preparation

Library preparation was performed at the Novogene (HK) Company Limited. Briefly, 1µg of total RNA was provided to Novogene and absence of RNA degradation was checked on a bioanalyzer (Agilent 2000 - Agilent). Then, mRNA was enriched using oligo(dT) primers and randomly fragmented using a fragmentation buffer. First strand cDNA synthesis was performed using random hexamers followed by second strand synthesis by nick-translation using a customized synthesis buffer (dNTP, RNase H, Esccherichia coli polymerase I). Double-stranded DNA was subsequently purified, end-repaired, A-tailed and ligated to sequencing adapters. Size selection was then performed and library was amplified by PCR. Library concentration was quantified on a fluorometer (Qubit 2.0 – Life Technologies) and insert size was analyzed on a bioanalyzer (Agilent 2100 – Agilent). Sequencing libraries were paired-end sequenced (150 bp read-length) on an Illumina NovoSeq platform.

Data analysis

First, raw data were trimmed with trimmomatic (version 0.39) to remove adapter contamination and low quality data with the following options:

ILLUMINACLIP: TruSeq3-PE-2.fa:2:30:10:2:keepBothReads LEADING:3 TRAILING:3 SLIDINGWINDOW:4:5 MINLEN:136

Sample alignments and read counts per gene were retrieved using GeneCounts quantification method from STAR (Dobin et al., 2013) version 2.6.1b and the Nvit\_psr\_1.1 *N. vitripennis* genome version with RefSeq annotation GCF\_009193385.1 as reference as follows:

STAR –runThreadN 12 –runMode alignReads –genomeDir /STAR\_genome\_NCBI\_RefSeq\_PSR1/ --readFilesIn sampleX\_R1\_paired.fq.gz sampleX\_R2\_paired.fq.gz --readFilesCommand zcat --outSAMtype BAM SortedByCoordinate --outFileNamePrefix \_trimmed\_Star\_NCBI\_RefSeq\_ --quantMode GeneCounts

Differential expression analysis was calculated with DESeq2 (Love et al., 2014) version 1.14.1 in R version 3.5.0. Genes with differential expression were selected when their adjusted *P*-value (Padj; Benjamini-Hochberg correction) was less than 0.05. Pearson correlation were calculated with the R function “cor” (Stats package version 3.5.0) on the rlog transformed count matrix generated by DEseq2 and using the method BLIND. The results were plotted on a heatmap using the R pheatmap package (version 1.0.12) and samples were ordered by hierarchical clustering. PCA plot were generated with the plotPCA function from DESeq2 on the count matrix after variance-stabilizing transformation using the method BLIND. Gene Ontology analyses on selected genes were performed using DAVID bioinformatics resources (Huang da et al., 2009) version 6.8 and only GO terms with a False Discovery Rate (FDR) less than 5 were considered significant.

For comparison between *N. vitripennis* and *A. mellifera* gene expression, homologs were systematically search for between species by tblastx (<https://blast.ncbi.nlm.nih.gov/>) and genes with the highest percentage homology were selected. For comparison between the *GFP* dsRNA sequence used in this study and the sequences of all *N. vitripennis* endogenous mRNA, a blastn (<https://blast.ncbi.nlm.nih.gov/>) was performed with option for “somewhat similar sequences” and a word size of 7 without filtering for low complexity regions. RNAi off-target prediction was also done using the online website WaspAtlas (http://cyverse.warwick.ac.uk:3000/tools/RNAi) and entering the complete *GFP* dsRNA sequence (Davies & Tauber, 2015).

Bibliography

Davies, N. J., & Tauber, E. (2015). WaspAtlas: a Nasonia vitripennis gene database and analysis platform. *Database (Oxford), 2015*. doi:10.1093/database/bav103

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., . . . Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics, 29*(1), 15-21. doi:10.1093/bioinformatics/bts635

Huang da, W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc, 4*(1), 44-57. doi:10.1038/nprot.2008.211

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol, 15*(12), 550. doi:10.1186/s13059-014-0550-8

Werren, J. H., Loehlin, D. W., & Giebel, J. D. (2009). Larval RNAi in Nasonia (parasitoid wasp). *Cold Spring Harb Protoc, 2009*(10), pdb prot5311. doi:10.1101/pdb.prot5311

Werren, J. H., Richards, S., Desjardins, C. A., Niehuis, O., Gadau, J., Colbourne, J. K., . . . Gibbs, R. A. (2010). Functional and evolutionary insights from the genomes of three parasitoid Nasonia species. *Science, 327*(5963), 343-348. doi:10.1126/science.1178028