A species-specific lncRNA modulates the reproductive ability of the asian tiger mosquito

Alexandros Belavilas-Trovas[1], Maria-Eleni Gregoriou[1†‡],
Spyros Tastsoglou[2,3†], Olga Soukia[1], Antonis Giakountis[1] and
Kostas Mathiopoulos[1]*

[1]Laboratory of Molecular Biology and Genomics, Department of Biochemistry & Biotechnology,
University of Thessaly, Larissa, Greece, [2]DIANA-Lab, Department of Computer Science and Biomedical
Informatics, University of Thessaly, Lamia, Greece, [3]Hellenic Pasteur Institute, Athens, Greece

Long non-coding RNA (lncRNA) research has emerged as an independent
scientific field in recent years. Despite their association with critical cellular
and metabolic processes in plenty of organisms, lncRNAs are still a largely
unexplored area in mosquito research. We propose that they could serve as
exceptional tools for pest management due to unique features they possess.
These include low inter-species sequence conservation and high tissue
specificity. In the present study, we investigated the role of ovary-specific
lncRNAs in the reproductive ability of the Asian tiger mosquito, *Aedes
albopictus*. Through the analysis of transcriptomic data, we identified several
lncRNAs that were differentially expressed upon blood feeding; we called these
genes Norma (NOn-coding RNA in Mosquito ovAries). We observed that
silencing some of these Normas resulted in significant impact on mosquito
fecundity and fertility. We further focused on Norma3 whose silencing resulted
in 43% oviposition reduction, in smaller ovaries and 53% hatching reduction of
the laid eggs, compared to anti-GFP controls. Moreover, a significant
downregulation of 2 mucins withing a neighboring (~100 Kb) mucin cluster
was observed in smaller anti-Norma3 ovaries, indicating a potential mechanism
of in-*cis* regulation between Norma3 and the mucins. Our work constitutes the
first experimental proof-of-evidence connecting lncRNAs with mosquito
reproduction and opens a novel path for pest management.

KEYWORDS

Aedes albopictus, tiger mosquito, RNAi pest control, lncRNAs (long non-coding RNAs),
species-specific control

1 Introduction

The remarkable progress of next-generation sequencing and genomics technologies
that took place during the past 20 years revealed an unexpected world of transcribed, non-
coding (nc) genomic elements that by far exceed in numbers the protein-coding
transcripts (Claverie, 2005). Long non-coding RNAs (lncRNAs) represent one class of
functional ncRNA transcripts, characterized by species specificity and tissue-specific

expression patterns. LncRNA transcripts are longer than 200 nucleotides, they are mainly transcribed by unique genes and most of them are subject to post-transcriptional modifications (splicing, poly-A tail, and C-cap), although they have limited or no protein-coding potential.

Various research studies in eukaryotic organisms have highlighted the role of lncRNAs in essential biological processes and different modes have been proposed regarding their action. These modes include, but are not limited to, 1) guiding or decoying transcription factors, 2) acting as scaffolds for chromatin modifying complexes, 3) functioning as sponges for miRNAs, 4) regulating post-transcriptional mRNA modifications [reviewed in (Marchese et al., 2017)].

Due to the absence of coding capacity, lncRNAs demonstrate a notable lack of nucleotide sequence conservation even among closely related species, which results in a high number of unique, species or genus specific lncRNAs across eukaryotic organisms (Pang et al., 2006; Bhartiya and Scaria, 2016). This low sequence conservation along with the aforementioned multifunctionality provide extraordinary difficulties in the development of computational tools that would predict lncRNA targets or potential modes of action. Nonetheless, despite this sequence variation, lncRNAs from different organisms may exhibit structural or functional conservation, highlighting their conserved role in essential biological pathways (Ponjavic et al., 2007; Diederichs, 2014; Tavares et al., 2019). Indeed, unlike other non-coding RNAs (e.g., miRNAs) that hybridize to their targets through sequence complementarity patterns, the functionality of lncRNAs mainly results from their secondary structure motifs (Kino et al., 2010; Ding et al., 2014; Chillón and Pyle, 2016; Smola et al., 2016).

The high sequence divergence of lncRNAs renders them as ideal candidates for the development of species-specific population control approaches against organisms of interest. Achieving species-level specificity is a point of major importance for the development of novel pest management approaches especially against insect pests, such as mosquitoes. Current insect control approaches are mainly based on chemical insecticides that pose serious threats to public health and biodiversity due to their neurotoxic action against other species, either mammals (including human) (Costa et al., 2008) or off-target insects (Desneux et al., 2007). Beneficial insects, contributing vitally to the stability of ecosystems and to agriculture, are severely harmed by the main classes of pesticides, even by those that are considered safe for humans. This issue arises due to the lack of species-specific mechanisms underlying insecticide action. Both organophosphates and carbamates target acetylcholinesterase (AChE) which is conserved among insects (Kwong, 2002), while pyrethroids target conserved voltage-sensitive sodium channels (Soderlund, 2010). Moreover, neonicotinoid pesticides which are considered non-toxic for vertebrates, act as agonists of nicotinic acetylcholine receptors (nAChRs). However, all insects are vulnerable to these pesticides due to the conserved nAChRs sequences and their critical role in signal transduction (Simon-Delso et al., 2015). In addition to the lack of species-specificity in the currently used chemical insecticides, there is also the huge and long-standing problem of insecticide resistance. The way out of this unfavorable situation is the development of novel pesticides that target alternative gene classes that could lead to more effective pest management (Sparks et al., 2021).

The Asian tiger mosquito, *Aedes albopictus*, is also the target of various control approaches, as it is a cosmopolitan vector of several lethal arboviruses, including dengue, Zika and chikungunya (Martinet et al., 2019). *Ae. albopictus* emerged from the tropical and sub-tropical regions of south-east Asia and rapidly expanded throughout the world due to its exceptional ability to adapt in different environmental conditions (Benedict et al., 2007). Its vectorial status arose especially after its connection with the major outbreaks of Chikungunya virus in 2005-07 in La Reunion (Pialoux et al., 2007) and in 2007 in Italy (Rezza et al., 2007; Angelini et al., 2008), while it was also associated with the first autochthonous cases of dengue fever in France in 2010, 2013, and 2015 (La Ruche et al., 2010; Marchand et al., 2013; Succo et al., 2016). It is certain that in the coming years its expansion will continue, and estimates indicate that by 2050 half of the world population will be exposed to disease-spreading mosquitoes, such as *Ae. albopictus*, due to climate change and global warming (Kraemer et al., 2019).

In the present study we explore the potential of targeting lncRNAs to control insect populations. LncRNAs could be used as species-specific molecular targets for the development of next-generation pesticides (e.g., RNAi pesticides (Fletcher et al., 2020)) or be part of the rapidly growing synthetic biology systems, such as SIT and gene drives (Caragata et al., 2020). The present study focuses on the investigation of lncRNAs that are related to *Ae. albopictus* reproduction, due to the significance of reproduction in population suppression approaches. The reproductive process in females is triggered by the consumption of a blood meal (BM) which activates a cascade of metabolic signaling pathways that lead to the development and production of eggs. We sought to identify lncRNAs that influence the reproductive process, aiming at the disruption of oogenesis and the reduction of mosquito fecundity and fertility.

2 Materials and methods

2.1 Mosquito rearing

An *Aedes albopictus* laboratory line was established from wild mosquitoes which were collected from the region of Thessaly, as described elsewhere (Ioannou et al., 2021), and were reared in the insectary facility of the Department of Biochemistry &

Biotechnology at the University of Thessaly. Adult mosquitoes were reared at 26 ± 1°C with a relative humidity of 60−70%, under a 14 h:10 h light/dark photoperiod. They were fed with 10% sucrose solution, while female mosquitoes were blood-fed (BF) from a human arm to initiate their gonotrophic cycles.

2.2 RNA extraction, reverse transcription and real-time PCR

Ovaries and other tissues were dissected under the microscope and their total RNA was extracted using Extrazol (BLIRT S.A., Gdańsk, PL), according to the manufacturer's instructions. The integrity of the RNA was assessed by agarose gel electrophoresis. Total RNA was treated with DNase I (Thermo Fisher Scientific, Waltham, MA, United States) and 1 μg of RNA was reverse transcribed to cDNA by using oligo-dT primers and MMLV-RT (Invitrogen, Waltham, MA, United States), according to the manufacturer's instructions. Each biological replicate corresponds to tissues collected from individual mosquitoes. We preferred to study tissues from individual mosquitoes, rather than pooling them together, in order to be able to assess within population variability. All qPCR assays were performed in CFX96 Real-Time Thermal Cycler (Bio-Rad, Hercules, CA, United States). All amplifications were performed with two technical replicates and the relative gene expression was analyzed by using the $2^{-\Delta\Delta Ct}$ method (Livak) through the CFX Manager™ software. Specific primers to amplify the genes identified by the transcriptomic analysis were designed using PrimerQuest™Tool. Their target specificity was verified through Primer-BLAST (Ye et al., 2012) against the RefSeq mRNA database of *Aedes albopictus*. Primers that lacked any sequence homology to other transcripts were selected. Two endogenous ribosomal housekeeping genes (RpL32, RpS17) were used for the normalization of the results (Dzaki and Azzam, 2018). The average expression stability value (M-value) of the reference genes for each biological sample was determined and samples that exhibited M-value <0.5 and Coefficient of Variance <0.25 were accepted for further analysis. Primers for qPCR are shown in Supplementary Table S2.

2.3 *In Vitro* double-stranded (ds) RNA synthesis-dsRNA treatment

Target templates for *in vitro* transcription were generated using gene specific primers with the respective recognition site for T7 RNA polymerase designed by the eRNAi web platform (Horn and Boutros, 2010). Thermodynamic parameters of the primers were tested through the web platform OligoAnalyzer (https://eu.idtdna.com/calc/analyzer). We selected primer sets that exhibited a minimum ΔG value of −9 kcal/mol for homodimer and heterodimer formation, while their hairpin structures did not exceed the primer Tm-10°C. The target specificity of the primers was assessed by Primer-BLAST (Ye et al., 2012) against the RefSeq mRNA database of *Aedes albopictus*. The off-target effect of dsRNAs (T7 amplicons) that were introduced for silencing was evaluated by examining for potential homologies of their siRNAs with other transcripts of *Aedes albopictus*. We obtained all possible 21-mers (i.e., sliding window with width = 21 and step = 1) and subjected them to remote blastn analysis (argument -task "blastn-short" for short input queries) against RefSeq mRNA database of *Aedes albopictus*. We proceeded with dsRNAs/siRNAs that did not contain 21-mers matching any other transcript of *Aedes albopictus,* other than their lncRNA target. dsRNA was synthesized by using the MEGAscript RNAi kit (Ambion, Austin, TX) from a dsDNA template by incubating at 37°C for 16 h. dsDNA was produced by PCR with cDNA template and gene-specific primers with the T7 RNA polymerase promoter (TAATACGACTCACTATAGGG) attached to their 5' end. dsRNA was purified with standard phenol/chloroform protocol, its integrity was measured by agarose gel electrophoresis and its concentration was quantified by the Q3000 Spectrophotometer (Quawell, San Jose, CA). Green fluorescent protein (GFP) dsRNA was used as a control. For silencing with each one of the Norma genes, inseminated *Ae. albopictus* adult females were used. The mosquitoes were collected 5 days after eclosion and were anesthetized on CO_2. Then, 64.4 nl of dsRNA solution (5,000 ng/μl) was injected into their thorax using the Nanojet II microinjector (Drummond, Birmingham, AL), under a Leica™ stereoscope. At least 25 individual females (i.e., biological replicates) were used for silencing with each Norma gene. Two controls were used for comparison: one was non-injected females (untreated) and the other was GFP-dsRNA injected females (anti-GFP). Both control populations were reared under the same conditions with the Norma-dsRNA injected ones. Primers for dsRNA production are shown in Supplementary Table S2.

2.4 Phenotypic assays

Five-day-old mosquitoes injected with either GFP-dsRNA or Norma3-dsRNA were transferred to Bugdorm™ 17.5 × 17.5 × 17.5 cm cages (MegaView Science Co., Talchung, Taiwan) immediately after injection, where they were fed with 10% sucrose solution for 24 h to recover. An additional sample of same age, non-injected mosquitoes (untreated) was also maintained. The mosquitoes were left to starve for 12 h and subsequently they were blood-fed (36 h post-injection). All blood meals took place in the same timeframe, between morning and noon, to avoid the influence of circadian variability among the replicates. Fully engorged, blood-fed mosquitoes were reared in the presence of 10% sucrose.

2.4.1 Ovarian measurement

Ovaries were dissected 60 h post-blood meal by detaching the soft cuticle between the fifth and sixth abdominal segment and pulling the terminal segments with fine forceps onto a drop of phosphate buffered saline (PBS). Pictures of the ovaries were captured with a Leica™ stereoscope (Leica Microsystems, Wetzlar, Germany) and the length of the long axis of the ovarian follicle was measured by FIJI software (Schindelin et al., 2012).

2.4.2 Oviposition

Four days after blood feeding individual mosquitoes were placed in polystyrene fly vials that contained a filter paper attached to a wet cotton ball and were left into the vial for 2 days in the insectary to lay their eggs. Additional moisture was added regularly to the vials to keep the filter paper wet. Two days later, mosquitoes were removed from the vials and the total number of eggs deposited by each individual mosquito was counted under a stereoscope. Mosquitoes that deceased during the process were excluded from the analysis.

2.4.3 Hatching assay

Laid eggs were dried and then stored into sealed petri dishes that contained a wet cotton ball as a source of humidity. Each petri dish hosted the eggs laid by one individual mosquito. Three days (72 h) post egg laying, 30 ml of hatching broth, that was prepared as described elsewhere (Maïga, 2017), were added into each petri dish. Eggs that were stored inside the petri dishes were incubated with the hatching broth for 14 days. Hatching broth was freshly replaced every 5 days. The long 14-day period was preferred because shorter periods led to reduced hatching of the eggs, probably due to diapause effect. Emerged larvae from each egg batch were counted daily and were removed from petri dishes. Finally, the hatch rate was estimated as the total number of emerged larvae divided by the total number of laid eggs:

$$hatch\ rate = \frac{emerged\ larvae}{laid\ eggs} \times 100\%$$

2.4.4 Bleaching assay

The filter papers that contained the eggs that were laid by individual dsRNA-injected females were collected, placed separately into petri dishes and left under moisture for 72 h to complete their embryogenesis. Then they were dechorionated in order to visualize their internal segments, based on a clarification methodology described elsewhere (Trpiš, 1970). A fresh Trpiš solution (3gr $NaClO_2$, 2 ml glacial acetic acid, in 1 L distilled H_2O) was prepared for each experiment and was added into the petri dishes that hosted eggs that were collected from mosquitoes injected with dsRNA against Norma3 or GFP. The eggs were incubated with Trpiš solution for 40 min at room temperature and then were immediately washed with PBS and placed gently with a soft brush into a drop of PBS under a Leica™ stereoscope (Leica Microsystems, Wetzlar, Germany) for visualization.

2.5 RNA-seq data pre-processing and primary analysis

Raw FASTQ files from a detailed, publicly available, developmental transcriptomic dataset by Gamez and colleagues were retrieved from the Sequence Read Archive (SRA) study entry SRP219966 (Gamez et al., 2020), using SRA toolkit. Quality assessment, identification of adapters and overrepresented contaminants was performed by employing FASTQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/) and the EMBL-EBI's Minion application (http://www.dev.ebi.ac.uk/enright-dev/kraken/reaper/src/reaper-latest/doc/minion.html). Quality trimming and adapter trimming of reads was performed with cutadapt. Pre-processed FASTQ files were mapped to the GCF_001876365.2_canu_80X_arrow2.2 *Ae. albopictus* genome assembly utilizing STAR splice-aware aligner (v2.7.9a) (Dobin et al., 2013) using the ENCODE standard options provided in the STAR manual. The respective transcript annotation file, derived from the NCBI Eukaryotic Genome Annotation Pipeline (Gnomon gene predictions), was also provided (--sjdbGTFfile argument) to effectively account for known splice junctions, while the genome index was accordingly parameterized to the read length (--sjdbOverhang 49). Transcript-level expression was calculated with RSEM v1.3.1. (Li and Dewey, 2011).

2.6 LncRNA annotation

Accession numbers with the "XR_" prefix, annotated as non-coding RNA (ncRNA) were retained from the GCF_001876365.2 gene models. In order to obtain estimates of the transcripts' coding potential, CPC2 (Kang et al., 2017) and FEELnc (Wucher et al., 2017) tools were employed. Both tools were executed at default settings (in the absence of any known lncRNAs, FEELnc "--mode = shuffle" method was chosen and the "XM_" (mRNA) transcripts were also provided as input for training). Transcripts were queried locally with BLASTn v2.13.0 (Altschul et al., 1990), at default settings, against Reference RNA Sequences (RefSeq_rna) from all Hexapoda species (Taxonomy ID: 6960). Each transcript was annotated with respect to the number of species it presented hits to, besides *Ae. albopictus*. FEELnc classifier was also used to characterize all "XR_" transcripts as genic and intergenic, according to their genomic localization and strand, respective to other transcripts. Lastly, RSEM-derived counts from all analyzed datasets in Section 2.6 were TMM-normalized (Robinson and Oshlack, 2010), log_2-

transformed and utilized to obtain *tau* (τ) tissue specificity indices (Yanai et al., 2005) (https://github.com/roonysgalbi/tispec). For a transcript *x* with expression x_i across *n* tissues/contexts, *tau* is obtained using its expression normalized by its maximal expression as follows:

$$tau = \frac{\sum_1^n (1 - \hat{x}_i)}{n - 1}; \ \hat{x}_i = \frac{x_i}{\max(x_i)_{1 \le i \le n}}$$

After deriving the global *tau* for a transcript, its per-context specificity fraction was calculated by multiplying *tau* with the maximal-normalized expression metric $\hat{x}_i$. "XR_" transcripts were annotated with their context (i.e., developmental stage, tissue and time-point) *tau* specificity index, as well as per-context fraction. Heatmaps were created using R package pheatmap, using Euclidean distance as a metric and complete-linkage hierarchical clustering where applicable. Annotation results are provided unfiltered in Supplementary Data S1.

2.7 Norma3 expression correlation and clustering analysis

For correlation analysis, samples where Norma3 presented non-zero expression (Transcripts-Per-Million, TPM units) were selected and transcripts presenting zero TPM across all samples were filtered out. Pearson correlation testing and False Discovery Rate adjustment of *p*-values were conducted in R. Clustering analysis was performed by providing TPM values to DPGP (McDowell et al., 2018), which applies Dirichlet Process to nonparametrically determine the optimal number of expression trajectory clusters and Gaussian Process to model the trajectories of expression through time. Following DPGP recommendations, a limited list of transcripts was subjected to clustering analysis instead of the whole transcriptome. For that purpose, read counts from post-blood meal (PBM) ovary samples (12, 24, 36, 48, 60, and 72 h) were imported in R. In the absence of replicates, 12–24 h, 36–48 h and 60–72 h time points were grouped into "early", "middle" and "late" groups respectively. EBSeq (v1.30.0) (Leng et al., 2013) was utilized to assign transcripts to patterns of differential expression, at a relaxed posterior probability ≥75%. DPGP was executed at default settings.

2.8 Statistical analysis

Data were presented as mean ± SEM of independent biological replicates, unless otherwise noted. Distribution patterns of the samples were evaluated through Shapiro-Wilk test (Shapiro and Wilk, 1965) and those populations that followed normal distribution were analyzed by two-tailed unpaired Student's t-test. Samples that did not pass normality test were analyzed by nonparametric two-tailed Mann-Whitney *U* test (Mann and Whitney, 1947). *p*-values of ≤0.05 were considered as significantly different. All analyses were performed through GraphPad Prism 8 software and all values are displayed in Supplementary Data S3.

3 Results

3.1 Annotation of predicted non-coding RNA in *Ae. albopictus*

In order to query for *Ae. albopictus* lncRNAs that are potentially implicated in mosquito reproduction, we initiated our computational analysis on the respective NCBI gene models, which contain 8571 transcripts that are predicted to belong to the ncRNA class ("gbkey = ncRNA"). Bioinformatics approaches were adopted to annotate further these RNAs regarding their coding potential, species specificity, genomic localization, as well as their context-expression specificity within *Ae. albopictus*.

Coding potential estimates of the predicted ncRNAs were calculated with CPC2 and FEELnc. CPC2 (Kang et al., 2017) is a well-known species-agnostic Support Vector Machine model that makes use of Fickett score, open reading frame (ORF) length and integrity and isoelectric point to predict coding potential. In contrast, FEELnc (Wucher et al., 2017) accepts as input known coding and non-coding transcripts from a species of interest and trains a species-aware coding potential model (Random Forests) using as predictors the ORF coverage of transcripts, k-mer composition (1-, 2-, 3-, 6-, 9-, and 12-mers) and transcript length. When known non-coding RNA sequences are not available, under shuffle mode, FEELnc shuffles the provided protein coding sequences but preserves 7-mer frequencies, to create a faux set of non-coding transcripts to use in model training-testing, a method shown to fare adequately in benchmarks against real lncRNAs (Wucher et al., 2017). The FEELnc model output is a score between 0 (i.e., no coding potential) and 1 (coding potential), while its optimal cutoff point that maximizes sensitivity and specificity is calculated by means of 10-fold cross validation. CPC2 annotated 8289 "ncRNA" transcripts as non-coding and 282 as potentially coding. FEELnc classification yielded 8096 and 475 transcripts labeled as non-coding and coding respectively, at an optimal cutoff point of 0.43 (Figure 1A). CPC2 and FEELnc predictions were juxtaposed with each other, highlighting 7,900 and 86 transcripts (92 and 1% of total) concordantly classified as non-coding and coding RNA respectively, and identifying 589 predictions (7% of total) that were discrepant between the two approaches (Figure 1B, Supplementary Data S1).
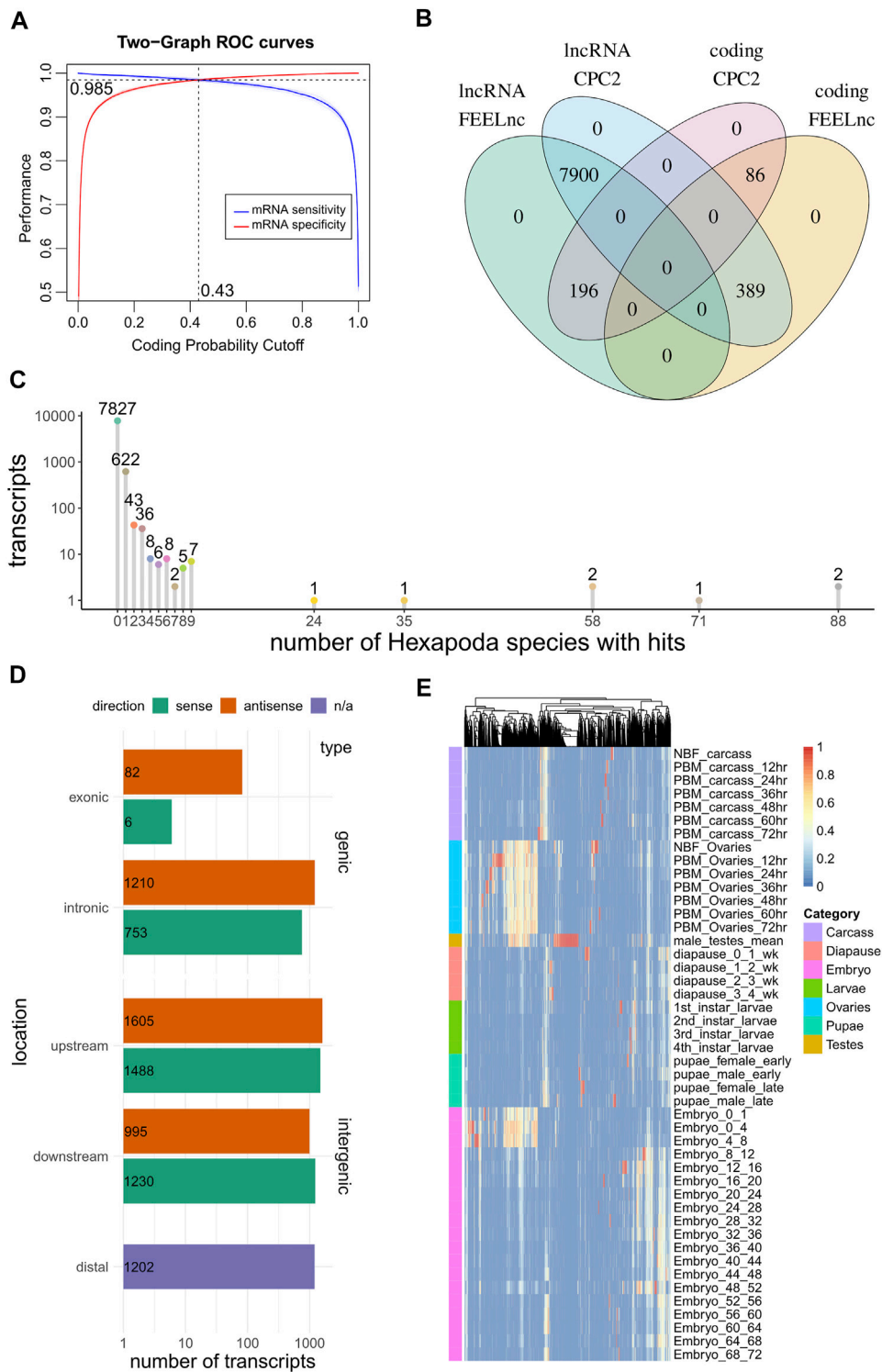
FIGURE 1
Computational analysis of RNAs predicted to be non-coding via automated NCBI analysis. **(A)** A coding potential prediction model (FEELnc Random Forests) was trained specifically on *Ae. albopictus* sequences. The optimal cutoff to discriminate coding from non-coding sequences was set at the point where sensitivity and specificity was maximized (10-fold cross validation). **(B)** Overlap of coding potential predictions between FEELnc and CPC2 models. **(C)** Depiction of number of transcripts presenting hits to other Hexapoda species. Transcripts are tallied by the number of species in which they presented BLASTn hits (*x*-axis). **(D)** Genomic localization of lncRNAs, per type (genic/intergenic), subcategory (exonic/intronic/upstream/downstream/distal (i.e., >100 kb afar from other transcripts) and strand relative to close/overlapping elements (sense/antisense). **(E)** Heatmap of per-context fractions for all lncRNAs presenting *tau* > 0.5.

FIGURE 2
Expression of Norma transcripts across mosquito developmental contexts and oviposition rate change after dsRNA-treatment against them. **(A)** Heatmap of Norma expression levels in carcass, ovary, testis, diapause, larva, pupa and embryo samples from a publicly available transcriptomic dataset by Gamez *et al.* Expression is depicted as $\log_2$-transformed TMM-normalized counts. **(B)** Impact of dsRNA-treatment, against each Norma separately, in oviposition rate. Each dot corresponds to the number of eggs that were laid individually 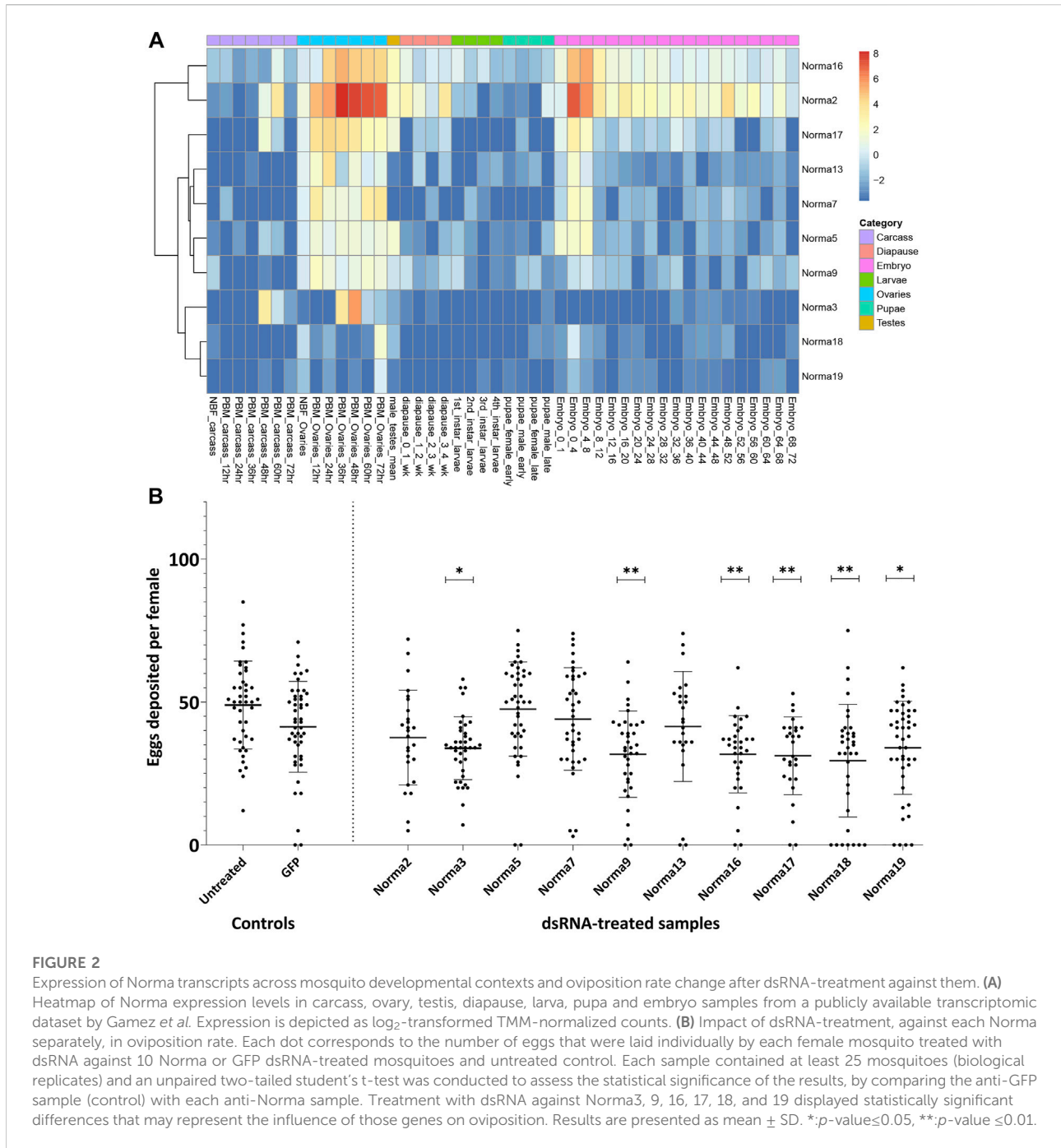by each female mosquito treated with dsRNA against 10 Norma or GFP dsRNA-treated mosquitoes and untreated control. Each sample contained at least 25 mosquitoes (biological replicates) and an unpaired two-tailed student's t-test was conducted to assess the statistical significance of the results, by comparing the anti-GFP sample (control) with each anti-Norma sample. Treatment with dsRNA against Norma3, 9, 16, 17, 18, and 19 displayed statistically significant differences that may represent the influence of those genes on oviposition. Results are presented as mean $\pm$ SD. *:*p*-value≤0.05, **:*p*-value ≤0.01.

In order to identify species-specific lncRNAs, transcript homology queries were performed (local BLASTn) against all Hexapoda Reference RNA sequences. Each *Ae. albopictus* "ncRNA" transcript was annotated with respect to the number of other species it presented hits against. Out of the total transcripts, 7,827 (91%) were found to exhibit no similarity with RNAs of any other Hexapoda species (Figure 1C).

FEELnc was also utilized to annotate transcripts' localization with regard to transcripts annotated as protein coding (Figure 1D). In total, 2051 instances (24%) were found to overlap protein coding transcripts in sense (9%) or antisense (15%) orientation. The remaining 76% was divided among intergenic transcripts that presented non-overlapping neighboring genes within 100 kb of their loci (5,318 transcripts, 62%) and those that were found to exist in

genomic regions not harboring other genes (annotated as "distal" intergenic, 1,202 instances, 14%).

Finally, in order to obtain expression metrics with regard to these ncRNAs in discrete developmental stages and tissues, a publicly available developmental transcriptomic dataset of *Ae. albopictus* (Gamez et al., 2020) was analyzed from scratch (SRA accession number SRP219966). Briefly, the dataset captured expression estimates from adult ovaries of non-blood-fed (NBF, fed with a 10%-sucrose solution) and post-blood meal (PBM, at 12-24-36-48-60-72 time-points) insects, carcasses (i.e., female bodies without the ovaries, also NBF and PBM at the same time-points), adult testes, diapause, larvae, pupae and embryo samples at numerous time-points. In the absence of replicates (with the exception of testes which were in duplicate), no attempt to assess differential expression was performed. Instead, log_2-transformed TMM-normalized counts were used to obtain *tau* indices (Yanai et al., 2005), which have been recently shown to perform consistently as a tissue specificity metric (Kryuchkova-Mostacci and Robinson-Rechavi, 2017). Briefly, *tau* constitutes a per-transcript tissue specificity unit ranging from 0 (no specificity) to 1 (highest specificity), while the fractions of each tissue that contributed to calculation of *tau* can also be obtained. Within the current dataset that is composed of a number of distinct tissues, developmental stages and time-points, we denoted that *tau* be regarded as a context-specificity index. The fractions of 3,119 transcripts exhibiting *tau* > 0.5 are presented as a heatmap (Figure 1E) for all available contexts, outlining the existence of numerous instances of intermediate-to-high specificity in ovary contexts (i.e., yellow-to-red transcripts within ovary samples).

We focused our attention on 984 transcripts that were 1) annotated as having no/limited coding potential by both CPC2 and FEELnc, 2) presented no sequence similarity to known transcripts of other Hexapoda and 3) exhibited *tau* > 0.5 and *tau fraction* > 0.5 in at least one PBM ovary time-point. Aiming to experimentally scrutinize a limited number of lncRNAs, out of this subset we selected 10 ovary-specific lncRNAs that were upregulated upon blood-feeding and presented limited or no expression in other developmental samples (Supplementary Table S1; Figure 2A). We termed these lncRNAs Norma (NOn-coding RNAs in Mosquito ovAries).

3.2 Phenotypic impact of norma genes silencing

In order to clarify the potential role of the 10 Norma genes, we evaluated their phenotypic impact in reproduction through RNAi silencing. Given their expression pattern, we reasoned that knocking down Norma genes would mostly impact mosquito oviposition. To assess this, we generated *in vitro* dsRNA against

each of the ten Norma lncRNAs which we injected into 5-day old inseminated adult females. We then provided a blood meal to the injected mosquitoes and monitored them for oviposition. We counted the number of eggs that were laid by each individual mosquito treated with any of the ten Norma-dsRNA (anti-Norma) and compared oviposition with a GFP-dsRNA sample (anti-GFP). A non-injected, blood-fed sample (untreated) was also present in the assay to monitor the effect of the environmental conditions. We observed that females injected with six different Norma-dsRNAs laid fewer eggs compared to GFP-dsRNA injected control, indicating the potential influence of silencing the corresponding lncRNAs to reproduction (Figure 2B). Specifically, dsRNA against Norma3, Norma9, Norma16, Norma17, Norma18 & Norma19 exhibited statistically significant reduction of oviposition rates compared to the GFP control (*p*-value ≤0.05). We focused our downstream analyses on Norma3 because it also presented some further striking features: an ovary-specific expression profile along with a sharp upregulation pattern in the post-vitellogenic time-points. Other lncRNAs that influenced oviposition (Norma9, Norma16, Norma17, Norma18, and Norma19) will be the subject of a future investigation.

3.3 Norma3 expression analysis and RNAi-knockdown

Initially, we determined the detailed expression profile of Norma3 by examining its stage- and tissue-related patterns in samples collected from our laboratory mosquito strain. We collected ovaries at the same time-points as the ones that were described in the RNA seq dataset that we analyzed. Specifically, we collected ovaries from NBF and PBM (12, 24, 36, 48, 60, and 72 h) time-points. In addition, we dissected and stored other tissues from the same time-points (carcass, midgut, Malpighian tubes, and head). Norma3 exhibited basal expression in NBF and early PBM samples, while its expression abruptly increased >2000-fold in 60 h-PBM ovaries, compared to the expression in NBF ovaries, and then significantly dropped to 200-fold in 72 h-PBM ovaries, compared to NBF (Figure 3A). At the same time-point (60 h PBM), Norma3 displayed basal expression in all other examined tissues (carcass, midgut, Malpighian tubes, head) (Figure 3A). The expression pattern was in accordance with publicly available RNA seq data (Gamez et al., 2020). Subsequently, we estimated the efficiency of Norma3 silencing in the ovaries of female mosquitoes injected with Norma3-dsRNA (anti-Norma3), compared with mosquitoes injected with GFP-dsRNA (anti-GFP). We collected ovaries from individual mosquitoes injected with either Norma3-dsRNA or GFP-dsRNA 60 h PBM and measured the effect of RNAi on the expression levels of Norma3. Results showed an average expression drop of >50% in anti-Norma3 replicates compared to anti-GFP (Figure 3B), which presented adequate statistical
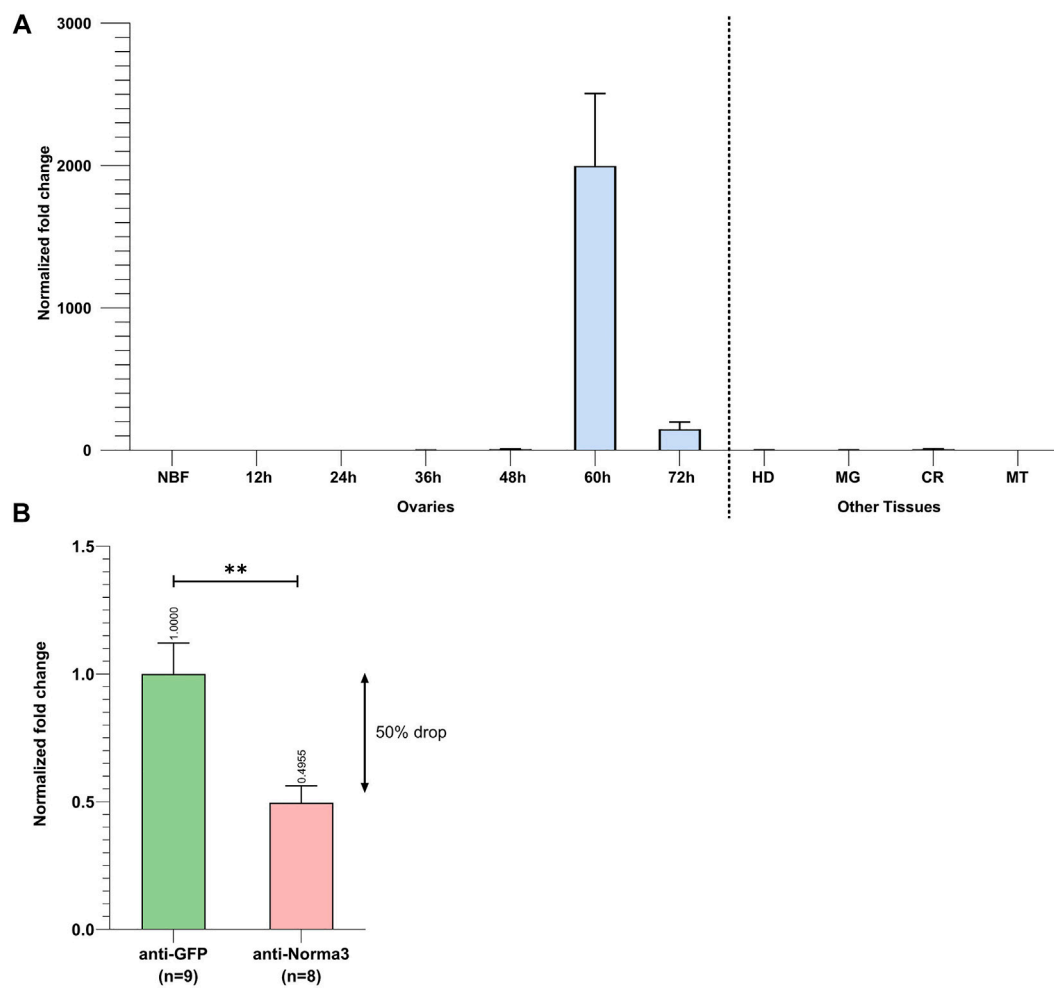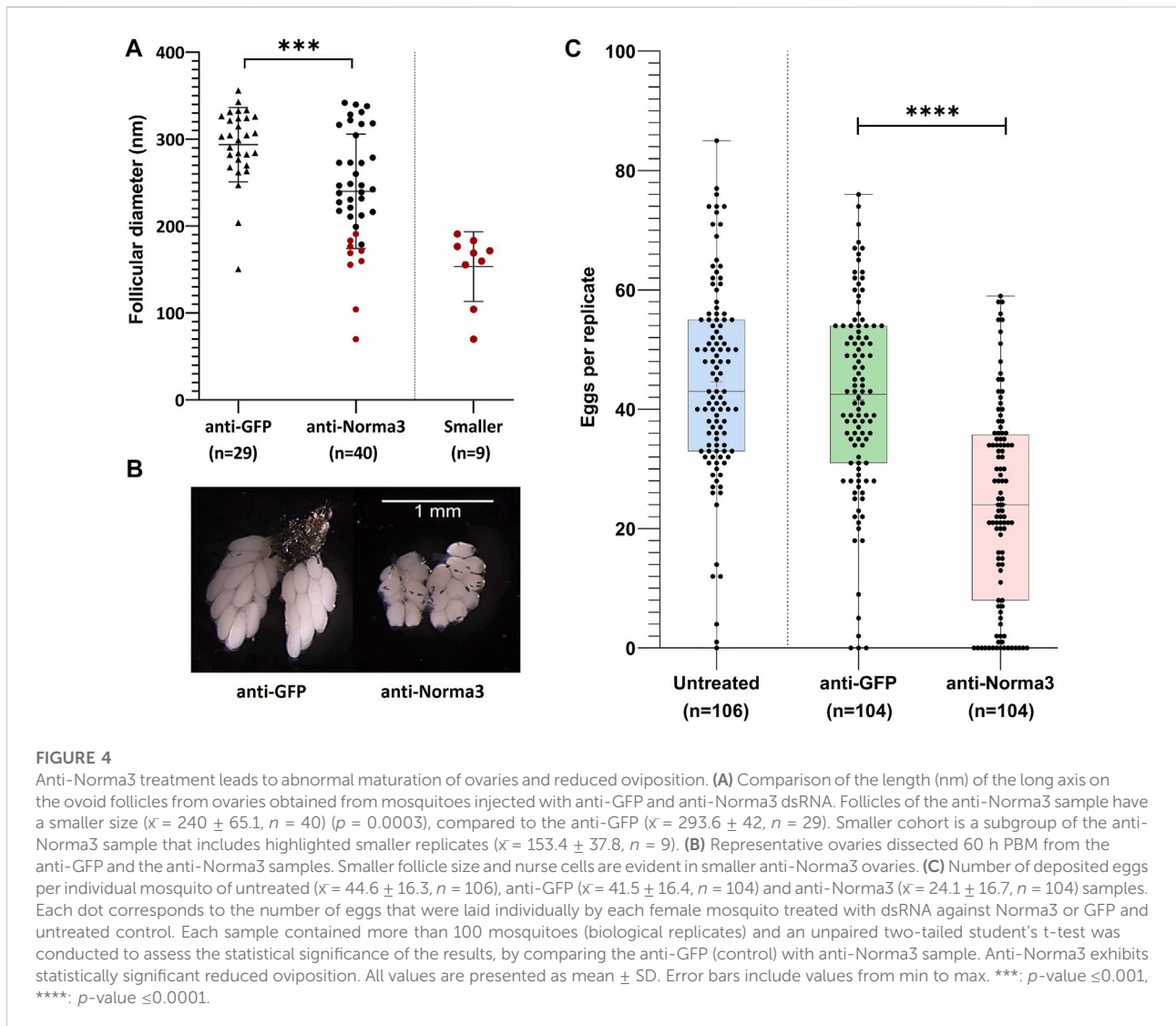
FIGURE 3
Spatiotemporal expression of Norma3 and its silencing efficacy. **(A)** Expression profile of Norma3 among non-blood fed (NBF) ovaries, post-blood meal (PBM) ovaries and other tissues collected 60 h PBM. Ovaries of blood-fed mosquitoes were collected every 12 h upon a blood meal, ranging from 12 to 72 h PBM when egg development is completed. Ovary samples (NBF & PBM) contained ovaries collected from 5-6 individual mosquitoes (biological replicates). Norma3 exhibits a basal expression in NBF, 12, 24, 36, and 48 h PBM ovaries, which abruptly increases at 60 h PBM leading to a fold change of >2,000 at 60 h PBM ($\bar{x} = 2,387 \pm 455.8$, $n = 5$) which drops to 200-fold in 72 h-PBM ($\bar{x} = 199.6 \pm 46.6$, $n = 6$), compared to NBF ovaries ($\bar{x} = 1 \pm 0.27$, $n = 6$). The other tissues that are presented were collected 60 h PBM. OV = Ovaries, HD = Head, MG = Midgut, CR = carcass, MT = Malpighian tubes. Each sample contained tissues collected from three individual mosquitoes (biological replicates). Fold change is presented relatively to NBF ovaries. All values were normalized with ribosomal genes RpL32 & RpS17 and are presented as mean $\pm$ SEM. **(B)** Relative quantification of Norma3 expression in replicates of the anti-Norma3 dsRNA vs. anti-GFP dsRNA samples. Ovaries were collected 60 h PBM, the time point when Norma3 peaks its expression. Each sample contains 8-9 biological replicates. Average expression of Norma3 in anti-GFP replicates was set as 1 ($\bar{x} = 100 \pm 11.4$, n = 9) and the overall expression drop of anti-Norma3 replicates, compared to anti-GFP, was measured to 50% ($\bar{x} = 49.5 \pm 6.2$, n = 8, $p = 0.0031$). All values were normalized with ribosomal genes RpL32 & RpS17 and are presented as mean $\pm$ SEM, **:p-value ≤ 0.01.

significance (Student's t-test, $p = 0.0031$) to support our experimental pipeline.

3.4 Impact of Norma3 silencing on reproduction

To more deeply characterize the impact of Norma3 silencing on *Ae. albopictus* reproductive ability, we examined various phenotypic traits that are connected to reproduction. First, we looked at the ovary morphology at 60 h PBM by measuring the long axis of the ovoid follicle of ovaries dissected by individuals of the anti-Norma3 dsRNA and anti-GFP dsRNA samples. We repeated the dissection process three times and detected that ovaries of anti-GFP sample presented an average length of 293.6 nm, while ovaries of anti-Norma3 presented a significantly lower length of 240 nm (Mann-Whitney test, $p = 0.003$) (Figure 4A). Interestingly, anti-Norma3 sample included

FIGURE 4

Anti-Norma3 treatment leads to abnormal maturation of ovaries and reduced oviposition. **(A)** Comparison of the length (nm) of the long axis on the ovoid follicles from ovaries obtained from mosquitoes injected with anti-GFP and anti-Norma3 dsRNA. Follicles of the anti-Norma3 sample have a smaller size ($\bar{x}$ = 240 $\pm$ 65.1, n = 40) (p = 0.0003), compared to the anti-GFP ($\bar{x}$ = 293.6 $\pm$ 42, n = 29). Smaller cohort is a subgroup of the anti-Norma3 sample that includes highlighted smaller replicates ($\bar{x}$ = 153.4 $\pm$ 37.8, n = 9). **(B)** Representative ovaries dissected 60 h PBM from the anti-GFP and the anti-Norma3 samples. Smaller follicle size and nurse cells are evident in smaller anti-Norma3 ovaries. **(C)** Number of deposited eggs per individual mosquito of untreated ($\bar{x}$ = 44.6 $\pm$ 16.3, n = 106), anti-GFP ($\bar{x}$ = 41.5 $\pm$ 16.4, n = 104) and anti-Norma3 ($\bar{x}$ = 24.1 $\pm$ 16.7, n = 104) samples. Each dot corresponds to the number of eggs that were laid individually by each female mosquito treated with dsRNA against Norma3 or GFP and untreated control. Each sample contained more than 100 mosquitoes (biological replicates) and an unpaired two-tailed student's t-test was conducted to assess the statistical significance of the results, by comparing the anti-GFP (control) with anti-Norma3 sample. Anti-Norma3 exhibits statistically significant reduced oviposition. All values are presented as mean $\pm$ SD. Error bars include values from min to max. ***: p-value $\leq$0.001, ****: p-value $\leq$0.0001.
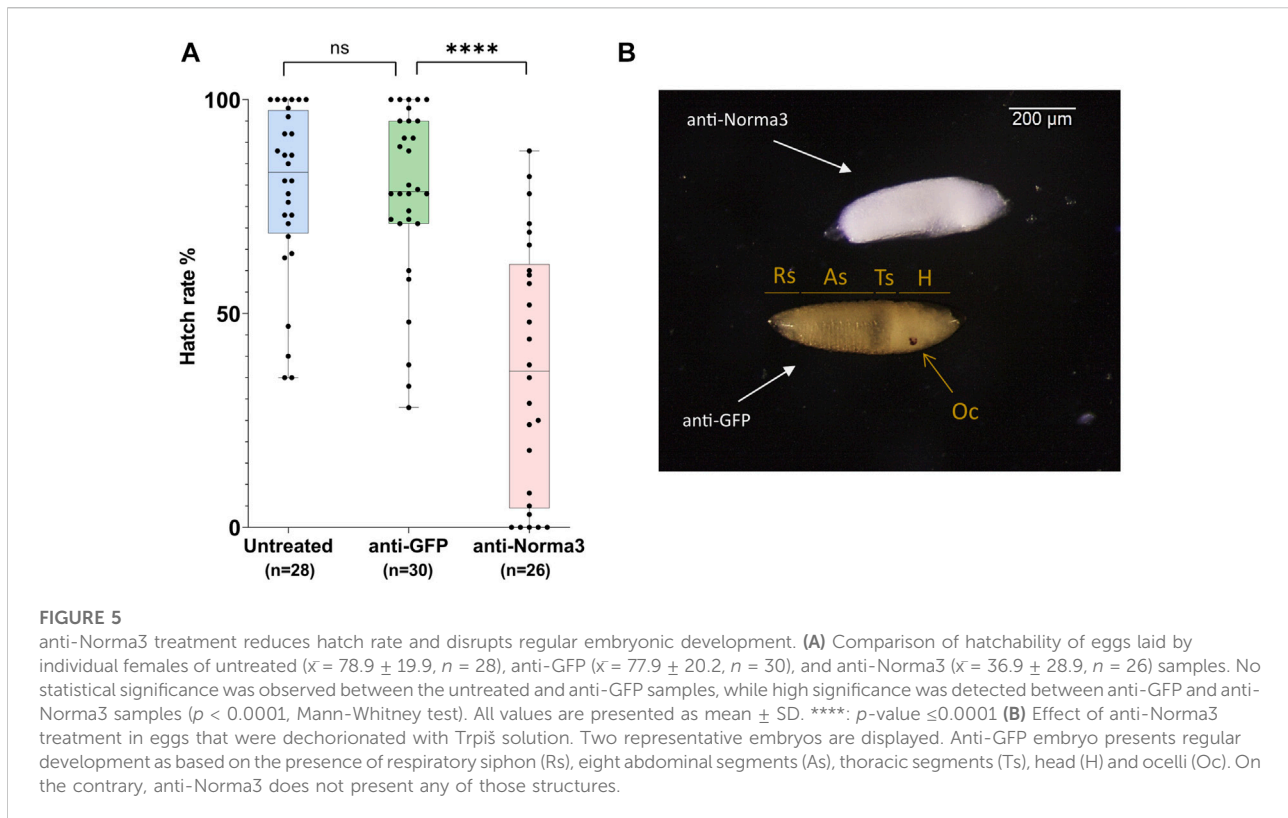
several much smaller ovaries that had a shorter follicular diameter (Supplementary Data S3), while there was evidence of the presence of nurse cells along with their oocyte, indicating their delayed development (Figure 4A,B, Supplementary Figure S1B). Nine anti-Norma3 smaller ovaries that displayed a mean length of 153.4 nm were collected and stored for further analysis.

Afterwards, we estimated the effect of the Norma3-dsRNA on oviposition by counting the number of eggs that were laid individually by each female of anti-Norma3, anti-GFP and untreated samples. Mosquitoes of the untreated control laid an average of 44.6 eggs, while mosquitoes of the anti-GFP control laid an average of 41.6 eggs. On the other hand, mosquitoes of the anti-Norma3 treatment laid an average of 24.1 eggs. Interestingly, 14% (n = 15) of the anti-Norma3 replicates laid zero eggs, while 26% (n = 27) laid

fewer than 10 eggs. In contrast, 3% (n = 3) of the anti-GFP and 1% of the untreated samples (n = 1) laid 0 eggs, while 6% (n = 6) of the anti-GFP and 3% of the untreated samples (n = 3) laid fewer than 10 eggs (Supplementary Data S3). The reduction of the average oviposition rate between anti-Norma3 and anti-GFP was 43% and exhibited high statistical significance (Student's t-test, p < 0.0001) (Figure 4C).

Subsequently, we addressed the hatch rate of untreated, anti-GFP and anti-Norma3 samples. We counted the number of larvae that emerged from the eggs that were laid by each individual mosquito of the studied samples and we divided by the total amount of eggs that were laid by each mosquito. Untreated mosquitoes displayed an average hatch rate of 79%, anti-GFP mosquitoes presented a similar hatch rate 78%, while anti-Norma3 mosquitoes had a much lower rate of

FIGURE 5
anti-Norma3 treatment reduces hatch rate and disrupts regular embryonic development. **(A)** Comparison of hatchability of eggs laid by individual females of untreated ($\bar{x}$ = 78.9 $\pm$ 19.9, n = 28), anti-GFP ($\bar{x}$ = 77.9 $\pm$ 20.2, n = 30), and anti-Norma3 ($\bar{x}$ = 36.9 $\pm$ 28.9, n = 26) samples. No statistical significance was observed between the untreated and anti-GFP samples, while high significance was detected between anti-GFP and anti-Norma3 samples ($p < 0.0001$, Mann-Whitney test). All values are presented as mean $\pm$ SD. ****: p-value ≤ 0.0001 **(B)** Effect of anti-Norma3 treatment in eggs that were dechorionated with Trpiš solution. Two representative embryos are displayed. Anti-GFP embryo presents regular development as based on the presence of respiratory siphon (Rs), eight abdominal segments (As), thoracic segments (Ts), head (H) and ocelli (Oc). On the contrary, anti-Norma3 does not present any of those structures.
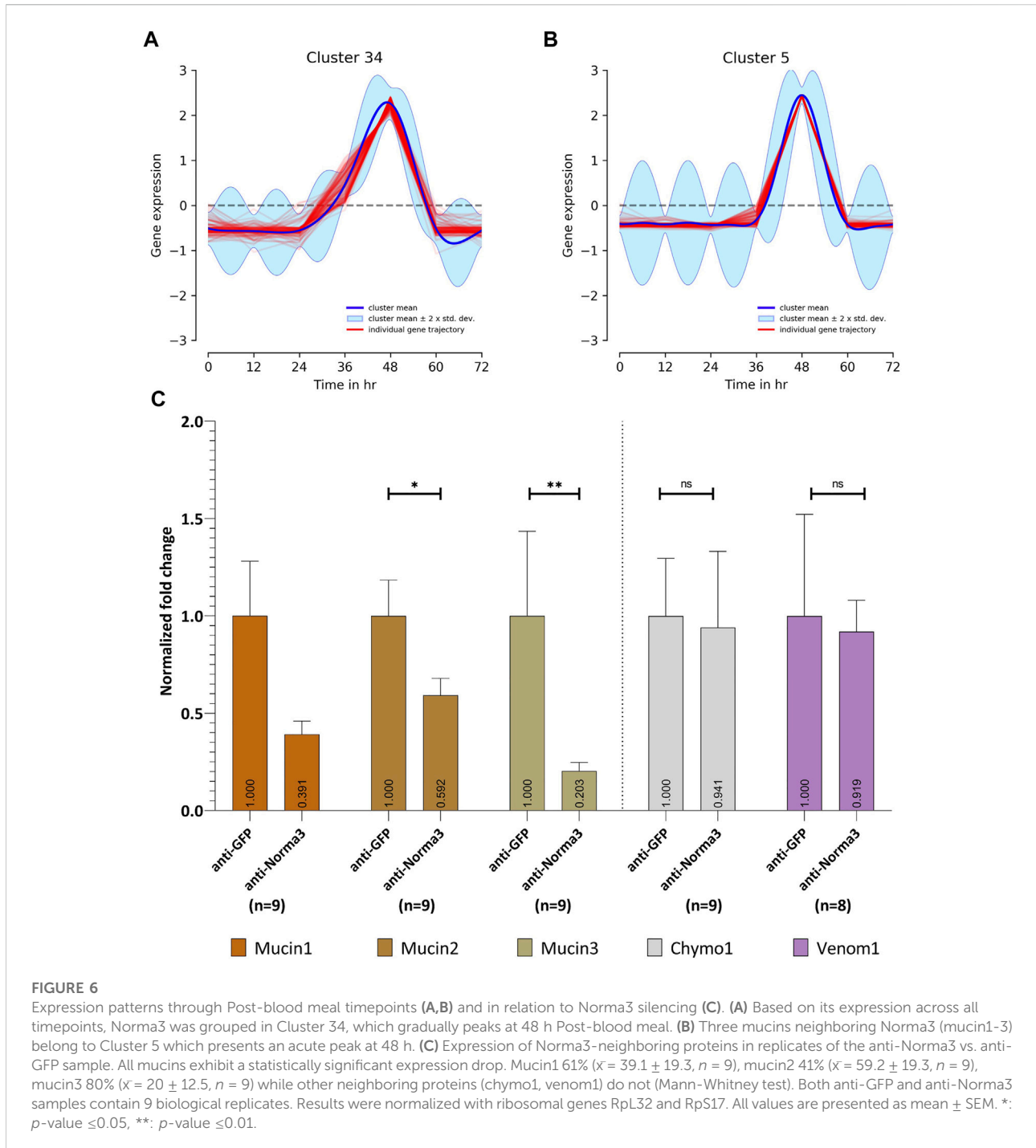
37%. (Figure 5A). The difference of the hatch rate between anti-GFP and anti-Norma3 was about 53% and presented high statistical significance (Mann-Whitney test, $p < 0.0001$).

Finally, we attempted to visualize possible larval defects in the anti-Norma3 treated sample. For this, we rendered 72 h eggs transparent by bleaching and observed the larvae under the microscope. We investigated the entire batches of eggs/embryos laid by each mosquito treated with dsRNA either against Norma3 or GFP and detected significant changes. There was a constant percentage of 75-85% that presented the deficient phenotype in anti-Norma3 sample, while the respective percentage of anti-GFP embryos was 10-15%. We dechorionated the embryos laid by 10 anti-Norma3 and 10 anti-GFP mosquitoes to validate the result. Representative images of dechorionated embryos that were laid by different mosquitoes of each sample and displayed the characteristic phenotype are presented in Figure 5B and Supplementary Figure S2. In anti-GFP embryos the eight abdominal segments, the thoracic segments, the respiratory siphon, the head and the ocelli were clearly visible indicating regular development of the embryo. On the other hand, anti-Norma3 embryos did not present any of the anticipated normal structures; instead, they exhibited a defective appearance of an undifferentiated mass (Figure 5B, Supplementary Figure S2B).

3.5 Regulatory interplay of Norma3 with neighboring mucins

Since lncRNAs often regulate coding genes in their genomic neighborhood (Rinn and Chang, 2012; Engreitz et al., 2016; Joung et al., 2017; Khyzha et al., 2019; Xing et al., 2021), we investigated the possible association of Norma3 with protein-coding genes in its vicinity. We returned to the available transcriptomic study containing NBF and PBM ovary datasets (Gamez et al., 2020) and assessed the correlation of Norma3 expression against the rest of the transcriptome. Within the region from 100 kb upstream to 100 kb downstream of Norma3, we identified 8 annotated transcripts (i.e., 4 mucins, 1 venom protein, and 3 chymotrypsin inhibitors) which all exhibited robust positive correlation with Norma3 (Pearson correlation coefficient >0.96, maximum FDR = 2.17e-9) (Supplementary Data S2). Subsequently, we reduced the transcriptome to transcripts which were more likely to present any change in ovaries among early (i.e., 12–24 h), middle (36–48 h) and late (60–72 h) time points (posterior-probability EBSeq >75%). This transcript set ($n = 10,314$) was subjected to clustering analysis based on their expression over the entire time course (i.e., NBF, 12, 24, 36, 48, 60, and 72 h PBM). Norma3 was grouped together with 114 other genes (111 protein-coding and 3 long non-coding) in a cluster of genes that was not expressed in NBF and early PBM

FIGURE 6
Expression patterns through Post-blood meal timepoints **(A,B)** and in relation to Norma3 silencing **(C)**. **(A)** Based on its expression across all timepoints, Norma3 was grouped in Cluster 34, which gradually peaks at 48 h Post-blood meal. **(B)** Three mucins neighboring Norma3 (mucin1-3) belong to Cluster 5 which presents an acute peak at 48 h. **(C)** Expression of Norma3-neighboring proteins in replicates of the anti-Norma3 vs. anti-GFP sample. All mucins exhibit a statistically significant expression drop. Mucin1 61% ($\bar{x}$ = 39.1 $\pm$ 19.3, n = 9), mucin2 41% ($\bar{x}$ = 59.2 $\pm$ 19.3, n = 9), mucin3 80% ($\bar{x}$ = 20 $\pm$ 12.5, n = 9) while other neighboring proteins (chymo1, venom1) do not (Mann-Whitney test). Both anti-GFP and anti-Norma3 samples contain 9 biological replicates. Results were normalized with ribosomal genes RpL32 and RpS17. All values are presented as mean $\pm$ SEM. *: p-value $\leq$0.05, **: p-value $\leq$0.01.

samples but initiated low transcription at 36 h and peaked at the end of vitellogenesis (around 48 h PBM) (Figure 6A). Among the 8 positively correlated neighbor transcripts, three mucins belonged to a cluster that presented an intense expression peak at 48 h (Figure 6B).

We then generated their detailed expression profile in NBF and PBM ovaries. We confirmed the expression pattern of five of the neighboring transcripts via qPCR. Three were annotated as mucin-2 and mucin-2-like (mucin1, 2, 3), one was annotated as cysteine-rich venom protein 6-like (venom1) and one as chymotrypsin inhibitor (chymo1). All five genes presented basal expression in NBF, 12, 24, 36, and 48 h ovaries. The expression of the three mucins presented a sharp increase between 12,000 (mucin2) and ~40,000-fold (mucin3) in 60 h

PBM ovaries, while chymo1 increased ~10,000-fold in 72 h PBM and venom1 3,000-fold in 60 h PBM ovaries (Supplementary Figure S3). To further assess the potential impact of Norma3 on the mucins we assayed the expression of the three mucins in a small sample of smaller ovaries (Figure 4A) of the anti-Norma3 60 h-PBM sample (that we collected as described in Section 3.4). Statistically significant downregulation between 40 to 80% of two out of the three mucin genes was detected in the smaller anti-Norma3 ovaries compared to anti-GFP control (Figure 6C). Downregulation of mucin3 presented the most significant effect, while no significant effect was detected on other protein-coding transcripts located in the same genomic region (mucin1, chymo1, venom1). Nevertheless, since we only tested nine smaller anti-Norma3 ovaries, we acknowledge that this result serves as a preliminary indication of the potential impact of Norma3 on mucin expression that should be validated with larger datasets.

4 Discussion

Long non-coding RNAs have arisen during the last decades as a fascinating field of research due to their unique features and intriguing mechanisms of action in the absence of a protein product. One entirely unexplored field of potential lncRNA applications is pest management and mosquito control. Due to the very low sequence conservation among lncRNAs (Diederichs, 2014; Tavares et al., 2019), such genes could be ideal as species-specific targets in new generation genetic control approaches. However, up to date there has not been sufficient research on specific lncRNAs to serve as efficient molecular targets with potential utility in pest management. In the tiger mosquito *Ae. albopictus*, a competent vector of multiple arboviruses, only a few studies regarding ncRNAs have been published so far (Gu et al., 2019; Azlan et al., 2021; Betting et al., 2021) and none presented significant data related to the impact of particular lncRNAs in physiological systems. Our aim was to provide the first proof-of-evidence regarding the role of specific lncRNAs in a vital biological process of the mosquito with potential utility for pest control. We focused on the reproductive system because of its significance for mosquito's viability and the broad applications that it may offer in mosquito control.

After computationally annotating predicted lncRNAs regarding their coding potential, genomic localization, species and developmental context specificity (Figure 1), we identified 10 species-specific lncRNAs which were overexpressed in *Ae. albopictus* ovaries upon blood-feeding (Figure 2A) and set out to further explore their potential role in reproduction. We deployed a loss-of-function RNAi-mediated pipeline and investigated the changes that occurred in the fecundity of female mosquitoes upon silencing of each lncRNA (Figure 2B). We focused on an antisense intergenic lncRNA, that we named Norma3, because its

targeting provoked a robust phenotypic effect. Moreover, it exhibited an ideal expression profile: basal expression in most of the NBF/early PBM time points followed by a sharp increase by the end of vitellogenesis (Figure 3A). Its expression was also highly ovary-specific (Figure 3A), while its nucleotide sequence did not display any similarities to annotated genes of other species. It is worth mentioning that we observed a slight discordance between expression peaks of RNA seq data and qRT-PCR, probably due to different sampling methodologies or adaptation of the local *Ae. albopictus* strain to our laboratory conditions.

Our successful gene silencing approach (Figure 3B) resulted in significant reduction in oviposition and egg hatching (Figure 4C; Figure 5A) of the anti-Norma3 dsRNA-treated mosquitoes. We further associated this fertility reduction with defective ovaries (Figure 4B), smaller ovary follicle size (Figure 4A) and obvious malformations of 72-h embryos of the anti-Norma3 treated mosquitoes (Figure 5B). In addition, we attempted to obtain insights on the potential mode of action of Norma3. Long non-coding RNAs operate in a variety of different modes and uncovering their functional role is not a trivial task. Unlike protein-coding genes that share conserved domains due to sequence homology, the features of lncRNAs mainly arise from their secondary structures which perform complex interplay with DNA regions, RNAs or proteins (Pang et al., 2006; Ding et al., 2014). One of most frequent roles of lncRNAs is to regulate the expression of coding genes and, oftentimes, such genes are localized in the vicinity of a lncRNA (Rinn and Chang, 2012; Engreitz et al., 2016; Joung et al., 2017; Khyzha et al., 2019; Xing et al., 2021). To explore this possibility, we performed *in silico* analysis to highlight coding genes that possessed similar expression profiles as Norma3 and were localized in the same region ±100 kb up-or-down stream of Norma3 (Supplementary Data S2). Our search resulted in eight coding genes which were annotated as mucins, chymotrypsin inhibitors and cysteine-rich venom-like proteins, while all of them contained a trypsin inhibitor-like (TIL) cysteine rich domain. According to our qPCR results, the expression of two of these mucins was severely diminished upon injection of anti-Norma3 dsRNA (Figure 6C). While this evidence is circumstantial, it may indicate a possible interplay between Norma3 and the two neighboring mucins. Mucins are proteins that are characterized by domains that contain repetitive sequences of proline, threonine and serine (PTS domains) and heavy O-glycosylation (Tran and Ten Hagen, 2013) which are present in most metazoan (Lang et al., 2007). While mucins are the principal components of mucus and mucous membranes, they carry diverse roles from lubrication to cell signaling to forming chemical barriers.

To our knowledge, there are no direct studies associating mucins with oviposition in mosquitoes. Ovary-specific mucins have been highlighted in *Ae. albopictus* by a recent study that reported the upregulation of multiple mucins upon blood-feeding in females and speculated a potential role of these

genes in oviposition (Deng et al., 2020). Nonetheless, various ovary-specific mucins have been identified in other insects and studies have revealed their role in ovary development and reproduction. In particular, mucin-like proteins have been associated with the eggshell, a multilayered structure which is formed during oogenesis that protects and nurtures the developing embryo prior to its arrest (Osterfield et al., 2017). In *Drosophila melanogaster* three putative eggshell genes code for proteins with mucin-like domains (Muc4B, Muc11D, and Muc12Ea). Muc4B has been suggested to be a component of the wax layer of the embryo, while Muc11D and Muc12Ea potentially act as mediators of chorion hardening and coating for passage of the embryo through the oviduct (Tootle et al., 2011). Another ovary-specific mucin (NlESMuc) that was identified in plant grasshopper *Nilaparvata lugens* was also related to the eggshell and was proven essential for its fecundity. Specifically, the RNAi-mediated targeting of NlESMuc caused reduced oviposition, lower egg production and less egg hatching (Lou et al., 2019). Finally, a study in the lepidopteran *Spodoptera exigua* presented an ovary-specific mucin-like protein called Se-Mucin1 that was associated with choriogenesis. In the absence of Se-Mucin1, females exhibited reduced fecundity and the hatch rate of the eggs was also significantly impaired, while SEM analysis of the eggshell structures revealed that they were remarkably malformed (Ahmed et al., 2021).

Through our analysis we collected pieces of circumstantial evidence that indicates a possible interplay between Norma3 and two neighboring mucins (mucin 2, 3). Firstly, both RNA seq analysis (Figures 6A,B) and qPCR data (Figure 3A, Supplementary Figure S3) demonstrate tightly linked sharp expression increases of Norma3 and the three mucins at 60 h post-blood meal. Moreover, we detected a potential influence of anti-Norma3 treatment in the expression drop of mucins 2 & 3 that could be associated to the developmental delay of ovaries (Figure 6C). Anti-Norma3 treatment led to a downregulation of their expression, especially in mucin3 which exhibited the most intense and statistically significant effect (*p*-value: 0.0019). However, the current sample size of anti-Norma3 and anti-GFP ovaries should be enlarged in order to validate the finding. Given these preliminary results, we speculate that Norma3 may act as a positive regulator of the mucin cluster and its silencing could lead to inhibition of their ovary specific-expression and possibly to disruption of the reproductive ability of *Ae. albopictus*. *Cis*-acting lncRNAs are one of the most dominant lncRNA classes and the majority of them overlap with enhancers elements (Gil and Ulitsky, 2020). We also presume that Norma3 and the mucin proteins are involved in the formation of the eggshell, a hypothesis which is based on the well-studied role of mucins in other insects (Tootle et al., 2011; Osterfield et al., 2017; Lou et al., 2019; Ahmed et al., 2021), but also relies on the relevant phenotypic outcomes that were provoked by targeting an eggshell-related protein (EOF-1) in the relative species *Ae. aegypti* (Isoe et al., 2019). Further research is necessary to verify the hypotheses on both

the impact of mucins on mosquito reproduction and their regulatory association with Norma3.

Our study presents a conceptual pipeline and a proof-of-principle towards novel approaches of insect pest control. It begins with the discovery of lncRNAs involved in the regulation of a physiological system that is fundamental for species survival and propagation, the reproductive system, and it showcases that down-regulating a particular lncRNA results in the damage of that system, reducing insect fecundity and fertility. Our results suggest that species-specificity of lncRNAs renders them preferred targets of RNAi-based pesticides. In fact, RNAi technology has offered a new and hopeful prospect of ecologically friendly approaches to insect control since it can minimize off-target effects. Low sequence conservation and high species-specificity of lncRNAs provide extra added value towards that end. Nonetheless, delivery methods of RNAi-based insecticides still pose major challenges (Yu et al., 2013; Niu et al., 2018) Till now, transgenic plants producing dsRNAs against vital insect genes has been the most straightforward and efficient application of RNAi technology for insect control (Nowara et al., 2010). However, this approach entails public acceptance problems due to the transgenic nature of the plants (Herman et al., 2021). Alternatively, spraying of stabilized dsRNA is recently being considered and actively researched [reviewed in (Rank and Koch, 2021)]. Examples come for the use of this technology against plant pests [e.g., sprays of dsRNA-producing *E. coli* against 28-spotted ladybird (Wu et al., 2021) or dsRNA sprays against Colorado potato beetle (Mehlhorn et al., 2020)], but one can easier envision house sprayings against biting mosquitoes due to greater dsRNA stability inside a house environment. Furthermore, reproductive system related lncRNAs, such as Norma3, could also be exploited in cutting-edge gene drive approaches aiming to suppress disease vectors by reducing female fertility (Hammond et al., 2016; Kyrou et al., 2018; North et al., 2020; Simoni et al., 2020). Given the available sequencing data in several insect species of public health or agricultural importance (or the affordability of obtaining such data from any organism of choice), this pipeline can be adopted to any given species and yield novel species-specific targets for pest control, thus addressing one of the most difficult challenges of the pesticide industry for species-specificity and environmental safety.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and

institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

AB-T, M-EG, AG and KM conceived the study; AB-T and M-EG organized and designed the bioassays; AB-T. and OS performed the phenotypic and in vitro assays; AB-T conducted the statistical analysis; ST and AG analyzed RNA sequencing data; ST performed lncRNA annotation and clustering analysis; and AB-T and KM wrote the paper. All authors edited and reviewed the manuscript.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Acknowledgments

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2022.885767/full#supplementary-material

References

Ahmed, S., Seo, K., and Kim, Y. (2021). An ovary specific mucin is associated with choriogenesis mediated by prostaglandin signaling in Spodoptera exigua. Arch. Insect Biochem. Physiol. 106 (1), e21748. doi:10.1002/arch.21748

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. J. Mol. Biol. 215 (3), 403–410. doi:10.1016/s0022-2836(05)80360-2

Angelini, P., Macini, P., Finarelli, A., Po, C., Venturelli, C., Bellini, R., et al. (2008). Chikungunya epidemic outbreak in Emilia-Romagna (Italy) during summer 2007. Parassitologia 50 (1/2), 97–98.

Azlan, A., Obeidat, S. M., Theva Das, K., Yunus, M. A., and Azzam, G. (2021). Genome-wide identification of Aedes albopictus long noncoding RNAs and their association with dengue and zika virus infection. PLoS Negl. Trop. Dis. 15 (1), e0008351. doi:10.1371/journal.pntd.0008351

Benedict, M. Q., Levine, R. S., Hawley, W. A., and Lounibos, L. P. (2007). Spread of the tiger: global risk of invasion by the mosquito Aedes albopictus. Vector-borne zoonotic Dis. 7 (1), 76–85. doi:10.1089/vbz.2006.0562

Betting, V., Joosten, J., Halbach, R., Thaler, M., Miesen, P., Van Rij, R. P., et al. (2021). A piRNA-lncRNA regulatory network initiates responder and trailer piRNA formation during mosquito embryonic development. RNA 27, 1155. doi:10.1261/rna.078876.121.078121-1172

Bhartiya, D., and Scaria, V. (2016). Genomic variations in non-coding RNAs: structure, function and regulation. Genomics 107 (2-3), 59–68. doi:10.1016/j.ygeno.2016.01.005

Caragata, E., Dong, S., Dong, Y., Simoes, M., Tikhe, C., Dimopoulos, G., et al. (2020). Prospects and pitfalls: next-generation tools to control mosquito-transmitted disease. Annu. Rev. Microbiol. 74, 455–475. doi:10.1146/annurev-micro-011320-025557

Chillón, I., and Pyle, A. M. (2016). Inverted repeat Alu elements in the human lincRNA-p21 adopt a conserved secondary structure that regulates RNA function. Nucleic Acids Res. 44 (19), 9462–9471. doi:10.1093/nar/gkw599

Claverie, J.-M. (2005). Fewer genes, more noncoding RNA. Science 309 (5740), 1529–1530. doi:10.1126/science.1116800

Costa, L. G., Giordano, G., Guizzetti, M., and Vitalone, A. (2008). Neurotoxicity of pesticides: a brief review. Front. Biosci. 13 (4), 1240. doi:10.2741/2758

Deng, F., Wu, S., Wu, Y., Liu, X., Wu, P., Zhai, Z., et al. (2020). Identification of mucins and their expression in the vector mosquito Aedes albopictus. J. Vector Ecol. 45 (2), 297–305. doi:10.1111/jvec.12400

Desneux, N., Decourtye, A., and Delpuech, J.-M. (2007). The sublethal effects of pesticides on beneficial arthropods. Annu. Rev. Entomol. 52, 81–106. doi:10.1146/annurev.ento.52.110405.091440

Diederichs, S. (2014). The four dimensions of noncoding RNA conservation. Trends Genet. 30 (4), 121–123. doi:10.1016/j.tig.2014.01.004

Ding, Y., Tang, Y., Kwok, C. K., Zhang, Y., Bevilacqua, P. C., Assmann, S. M., et al. (2014). In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. Nature 505 (7485), 696–700. doi:10.1038/nature12756

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

Dzaki, N., and Azzam, G. (2018). Assessment of *Aedes albopictus* reference genes for quantitative PCR at different stages of development. *PloS One* 13 (3), e0194664. doi:10.1371/journal.pone.0194664

Engreitz, J. M., Haines, J. E., Perez, E. M., Munson, G., Chen, J., Kane, M., et al. (2016). Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* 539 (7629), 452–455. doi:10.1038/nature20149

Fletcher, S. J., Reeves, P. T., Hoang, B. T., and Mitter, N. (2020). A perspective on RNAi-based biopesticides. *Front. Plant Sci.* 11, 51. doi:10.3389/fpls.2020.00051

Gamez, S., Antoshechkin, I., Mendez-Sanchez, S. C., and Akbari, O. S. (2020). The developmental transcriptome of *Aedes albopictus*, a major worldwide human disease vector. *G3 Genes Genomes Genetics* 10 (3), 1051–1062. doi:10.1534/g3. 119.401006

Gil, N., and Ulitsky, I. (2020). Regulation of gene expression by cis-acting long non-coding RNAs. *Nat. Rev. Genet.* 21 (2), 102–117. doi:10.1038/s41576-019-0184-5

Gu, J., Xu, Y., Dong, Y., Lai, Z., Jin, B., Hao, Y., et al. (2019). Differentiation of long non-coding RNA and mRNA expression profiles in male and female *Aedes albopictus*. *Front. Genet.* 10, 975. doi:10.3389/fgene.2019.00975

Hammond, A., Galizi, R., Kyrou, K., Simoni, A., Siniscalchi, C., Katsanos, D., et al. (2016). A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*. *Nat. Biotechnol.* 34 (1), 78–83. doi:10. 1038/nbt.3439

Herman, R. A., Storer, N. P., Anderson, J. A., Amijee, F., Cnudde, F., Raybould, A., et al. (2021). Transparency in risk-disproportionate regulation of modern crop-breeding techniques. *GM Crops Food* 12 (1), 376–381. doi:10.1080/21645698.2021. 1934353

Horn, T., and Boutros, M. (2010). E-RNAi: a web application for the multi-species design of RNAi reagents-2010 update. *Nucleic acids Res.* 38 (2), W332–W339. doi:10.1093/nar/gkq317

Ioannou, C. S., Hadjichristodoulou, C., Mouchtouri, V. A., and Papadopoulos, N. T. (2021). Effects of selection to diflubenzuron and Bacillus thuringiensis var. Israelensis on the overwintering successes of *Aedes albopictus* (Diptera: Culicidae). *Insects* 12 (9), 822. doi:10.3390/insects12090822

Isoe, J., Koch, L. E., Isoe, Y. E., Rascón, A. A., Brown, H. E., Massani, B. B., et al. (2019). Identification and characterization of a mosquito-specific eggshell organizing factor in *Aedes aegypti* mosquitoes. *PLoS Biol.* 17 (1), e3000068. doi:10.1371/journal.pbio.3000068

Joung, J., Engreitz, J. M., Konermann, S., Abudayyeh, O. O., Verdine, V. K., Aguet, F., et al. (2017). Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. *Nature* 548 (7667), 343–346. doi:10.1038/nature23451

Kang, Y.-J., Yang, D.-C., Kong, L., Hou, M., Meng, Y.-Q., Wei, L., et al. (2017). CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic acids Res.* 45 (W1), W12–W16. doi:10.1093/nar/gkx428

Khyzha, N., Khor, M., DiStefano, P. V., Wang, L., Matic, L., Hedin, U., et al. (2019). Regulation of CCL2 expression in human vascular endothelial cells by a neighboring divergently transcribed long noncoding RNA. *Proc. Natl. Acad. Sci. U. S. A.* 116 (33), 16410–16419. doi:10.1073/pnas.1904108116

Kino, T., Hurt, D. E., Ichijo, T., Nader, N., and Chrousos, G. P. (2010). Noncoding RNA gas5 is a growth arrest–and starvation-associated repressor of the glucocorticoid receptor. *Sci. Signal.* 3 (107), ra8. doi:10.1126/scisignal.2000568

Kraemer, M. U., Reiner, R. C., Brady, O. J., Messina, J. P., Gilbert, M., Pigott, D. M., et al. (2019). Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. *Nat. Microbiol.* 4 (5), 854–863. doi:10.1038/s41564-019-0376-y

Kryuchkova-Mostacci, N., and Robinson-Rechavi, M. (2017). A benchmark of gene expression tissue-specificity metrics. *Brief. Bioinform.* 18 (2), 205–214. doi:10. 1093/bib/bbw008

Kwong, T. C. (2002). Organophosphate pesticides: biochemistry and clinical toxicology. *Ther. drug Monit.* 24 (1), 144–149. doi:10.1097/00007691-200202000-00022

Kyrou, K., Hammond, A. M., Galizi, R., Kranjc, N., Burt, A., Beaghton, A. K., et al. (2018). A CRISPR–Cas9 gene drive targeting doublesex causes complete population suppression in caged *Anopheles gambiae* mosquitoes. *Nat. Biotechnol.* 36 (11), 1062–1066. doi:10.1038/nbt.4245

La Ruche, G., Souarès, Y., Armengaud, A., Peloux-Petiot, F., Delaunay, P., Desprès, P., et al. (2010). First two autochthonous dengue virus infections in metropolitan France, September 2010. *Eurosurveillance* 15 (39), 19676. doi:10.2807/ese.15.39.19676-en

Lang, T., Hansson, G. C., and Samuelsson, T. (2007). Gel-forming mucins appeared early in metazoan evolution. *Proc. Natl. Acad. Sci. U. S. A.* 104 (41), 16209–16214. doi:10.1073/pnas.0705984104

Leng, N., Dawson, J. A., Thomson, J. A., Ruotti, V., Rissman, A. I., Smits, B. M., et al. (2013). EBSeq: an empirical bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* 29 (8), 1035–1043. doi:10.1093/bioinformatics/btt087

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinforma.* 12 (1), 323. doi:10.1186/1471-2105-12-323

Lou, Y.-H., Shen, Y., Li, D.-T., Huang, H.-J., Lu, J.-B., Zhang, C.-X., et al. (2019). A mucin-like protein is essential for oviposition in Nilaparvata lugens. *Front. Physiol.* 10, 551. doi:10.3389/fphys.2019.00551

Maïga, H. (2017). Guidelines for routine colony maintenance of Aedes mosquito species. FAO-IAEA.

Mann, H. B., and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* 18, 50–60. doi:10.1214/aoms/1177730491

Marchand, E., Prat, C., Jeannin, C., Lafont, E., Bergmann, T., Flusin, O., et al. (2013). Autochthonous case of dengue in France, October 2013. *Eurosurveillance* 18 (50), 20661. doi:10.2807/1560-7917.es2013.18.50.20661

Marchese, F. P., Raimondi, I., and Huarte, M. (2017). The multidimensional mechanisms of long noncoding RNA function. *Genome Biol.* 18 (1), 206. doi:10.1186/s13059-017-1348-2

Martinet, J.-P., Ferté, H., Failloux, A.-B., Schaffner, F., and Depaquit, J. (2019). Mosquitoes of north-western europe as potential vectors of arboviruses: a review. *Viruses* 11 (11), 1059. doi:10.3390/v11111059

McDowell, I. C., Manandhar, D., Vockley, C. M., Schmid, A. K., Reddy, T. E., Engelhardt, B. E., et al. (2018). Clustering gene expression time series data using an infinite Gaussian process mixture model. *PLoS Comput. Biol.* 14 (1), e1005896. doi:10.1371/journal.pcbi.1005896

Mehlhorn, S. G., Geibel, S., Bucher, G., and Nauen, R. (2020). Profiling of RNAi sensitivity after foliar dsRNA exposure in different European populations of Colorado potato beetle reveals a robust response with minor variability. *Pesticide Biochem. Physiol.* 166, 104569. doi:10.1016/j.pestbp.2020. 104569

Niu, J., Taning, C. N. T., Christiaens, O., Smagghe, G., and Wang, J.-J. (2018). "Rethink RNAi in insect pest control: challenges and perspectives," in *Advances in insect physiology* (Elsevier), 1–17. doi:10.1016/bs.aiip.2018.07.003

North, A. R., Burt, A., and Godfray, H. C. J. (2020). Modelling the suppression of a malaria vector using a CRISPR-Cas9 gene drive to reduce female fertility. *BMC Biol.* 18 (1), 98. doi:10.1186/s12915-020-00834-z

Nowara, D., Gay, A., Lacomme, C., Shaw, J., Ridout, C., Douchkov, D., et al. (2010). HIGS: host-induced gene silencing in the obligate biotrophic fungal pathogen blumeria graminis. *Plant Cell* 22 (9), 3130–3141. doi:10.1105/tpc. 110.077040

Osterfield, M., Berg, C. A., and Shvartsman, S. Y. (2017). Epithelial patterning, morphogenesis, and evolution: drosophila eggshell as a model. *Dev. cell* 41 (4), 337–348. doi:10.1016/j.devcel.2017.02.018

Pang, K. C., Frith, M. C., and Mattick, J. S. (2006). Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends Genet.* 22 (1), 1–5. doi:10.1016/j.tig.2005.10.003

Pialoux, G., Gaüzère, B.-A., Jauréguiberry, S., and Strobel, M. (2007). Chikungunya, an epidemic arbovirosis. *Lancet Infect. Dis.* 7 (5), 319–327. doi:10.1016/s1473-3099(07)70107-x

Ponjavic, J., Ponting, C. P., and Lunter, G. (2007). Functionality or transcriptional noise? evidence for selection within long noncoding RNAs. *Genome Res.* 17 (5), 556–565. doi:10.1101/gr.6036807

Rank, A. P., and Koch, A. (2021). Lab-to-Field transition of RNA spray applications-how far are we? *Front. Plant Sci.* 12, 755203. doi:10.3389/fpls. 2021.755203

Rezza, G., Nicoletti, L., Angelini, R., Romi, R., Finarelli, A., Panning, M., et al. (2007). Infection with chikungunya virus in Italy: an outbreak in a temperate region. *Lancet* 370 (9602), 1840–1846. doi:10.1016/s0140-6736(07)61779-6

Rinn, J. L., and Chang, H. Y. (2012). Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* 81, 145–166. doi:10.1146/annurev-biochem-051410-092902

Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11 (3), R25. doi:10. 1186/gb-2010-11-3-r25

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9 (7), 676–682. doi:10.1038/nmeth.2019

Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* 52 (3/4), 591. doi:10.2307/2333709

Simon-Delso, N., Amaral-Rogers, V., Belzunces, L. P., Bonmatin, J.-M., Chagnon, M., Downs, C., et al. (2015). Systemic insecticides (neonicotinoids and fipronil): trends, uses, mode of action and metabolites. *Environ. Sci. Pollut. Res.* 22 (1), 5–34. doi:10.1007/s11356-014-3470-y

Simoni, A., Hammond, A. M., Beaghton, A. K., Galizi, R., Taxiarchi, C., Kyrou, K., et al. (2020). A male-biased sex-distorter gene drive for the human malaria vector *Anopheles gambiae*. *Nat. Biotechnol.* 38 (9), 1054–1060. doi:10.1038/s41587-020-0508-1

Smola, M. J., Christy, T. W., Inoue, K., Nicholson, C. O., Friedersdorf, M., Keene, J. D., et al. (2016). SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the Xist lncRNA in living cells. *Proc. Natl. Acad. Sci. U. S. A.* 113 (37), 10322–10327. doi:10.1073/pnas.1600008113

Soderlund, D. M. (2010). "Toxicology and mode of action of pyrethroid insecticides," in *Hayes' handbook of pesticide toxicology* (Elsevier), 1665–1686. doi:10.1016/B978-0-12-374367-1.00077-X

Sparks, T. C., Storer, N., Porter, A., Slater, R., and Nauen, R. (2021). Insecticide resistance management and industry: the origins and evolution of the Insecticide Resistance Action Committee (IRAC) and the mode of action classification scheme. *Pest Manag. Sci.* 77 (6), 2609–2619. doi:10.1002/ps.6254

Succo, T., Leparc-Goffart, I., Ferré, J.-B., Roiz, D., Broche, B., Maquart, M., et al. (2016). Autochthonous dengue outbreak in Nîmes, south of France, July to September 2015. *Eurosurveillance* 21 (21), 30240. doi:10.2807/1560-7917.es.2016.21.21.30240

Tavares, R. C., Pyle, A. M., and Somarowthu, S. (2019). Phylogenetic analysis with improved parameters reveals conservation in lncRNA structures. *J. Mol. Biol.* 431 (8), 1592–1603. doi:10.1016/j.jmb.2019.03.012

Tootle, T. L., Williams, D., Hubb, A., Frederick, R., and Spradling, A. (2011). Drosophila eggshell production: identification of new genes and coordination by pxt. *PloS One* 6 (5), e19943. doi:10.1371/journal.pone.0019943

Tran, D. T., and Ten Hagen, K. G. (2013). Mucin-type O-glycosylation during development. *J. Biol. Chem.* 288 (10), 6921–6929. doi:10.1074/jbc.r112.418558

Trpiš, M. (1970). A new bleaching and decalcifying method for general use in zoology. *Can. J. Zool.* 48 (4), 892–893. doi:10.1139/z70-158

Wu, J. J., Mu, L. L., Kang, W. N., Ze, L. J., Shen, C. H., Jin, L., et al. (2021). RNA interference targeting ecdysone receptor blocks the larval–pupal transition in Henosepilachna vigintioctopunctata. *Insect Sci.* 28 (2), 419–429. doi:10.1111/1744-7917.12777

Wucher, V., Legeai, F., Hedan, B., Rizk, G., Lagoutte, L., Leeb, T., et al. (2017). FEELnc: a tool for long non-coding RNA annotation and its application to the dog transcriptome. *Nucleic Acids Res.* 45 (8), e57. doi:10.1093/nar/gkw1306

Xing, L., Xi, Y., Qiao, X., Huang, C., Wu, Q., Yang, N., et al. (2021). The landscape of lncRNAs in Cydia pomonella provides insights into their signatures and potential roles in transcriptional regulation. *BMC genomics* 22 (1), 4. doi:10.1186/s12864-020-07313-3

Yanai, I., Benjamin, H., Shmoish, M., Chalifa-Caspi, V., Shklar, M., Ophir, R., et al. (2005). Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* 21 (5), 650–659. doi:10.1093/bioinformatics/bti042

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., Madden, T. L., et al. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinforma.* 13 (1), 134. doi:10.1186/1471-2105-13-134

Yu, N., Christiaens, O., Liu, J., Niu, J., Cappelle, K., Caccia, S., et al. (2013). Delivery of dsRNA for RNAi in insects: an overview and future directions. *Insect Sci.* 20 (1), 4–14. doi:10.1111/j.1744-7917.2012.01534.x