RESEARCH ARTICLE

Genome-wide identification of *Aedes albopictus* long noncoding RNAs and their association with dengue and Zika virus infection

Azali Azlan[1], **Sattam M. Obeidat**[1], **Kumitaa Theva Das**[2], **Muhammad Amir Yunus**[2], **Ghows Azzam**[1]*

1 School of Biological Sciences, Universiti Sains Malaysia, Penang, Malaysia, **2** Infectomics Cluster, Advanced Medical & Dental Institute, Universiti Sains Malaysia, Bertam, Kepala Batas, Pulau Pinang, Malaysia

* ghows@usm.my

Abstract

The Asian tiger mosquito, *Aedes albopictus* (*Ae. albopictus*), is an important vector that transmits arboviruses such as dengue (DENV), Zika (ZIKV) and Chikungunya virus (CHIKV). Long noncoding RNAs (lncRNAs) are known to regulate various biological processes. Knowledge on *Ae. albopictus* lncRNAs and their functional role in virus-host interactions are still limited. Here, we identified and characterized the lncRNAs in the genome of an arbovirus vector, *Ae. albopictus*, and evaluated their potential involvement in DENV and ZIKV infection. We used 148 public datasets, and identified a total of 10, 867 novel lncRNA transcripts, of which 5,809, 4,139, and 919 were intergenic, intronic and antisense respectively. The *Ae. albopictus* lncRNAs shared many characteristics with other species such as short length, low GC content, and low sequence conservation. RNA-sequencing of *Ae. albopictus* cells infected with DENV and ZIKV showed that the expression of lncRNAs was altered upon virus infection. Target prediction analysis revealed that *Ae. albopictus* lncRNAs may regulate the expression of genes involved in immunity and other metabolic and cellular processes. To verify the role of lncRNAs in virus infection, we generated mutations in lncRNA loci using CRISPR-Cas9, and discovered that two lncRNA loci mutations, namely XLOC_029733 (novel lncRNA transcript id: lncRNA_27639.2) and LOC115270134 (known lncRNA transcript id: XR_003899061.1) resulted in enhancement of DENV and ZIKV replication. The results presented here provide an important foundation for future studies of lncRNAs and their relationship with virus infection in *Ae. albopictus*.

Author summary

Ae. albopictus is an important vector of arboviruses such as dengue and Zika viruses. Studies on virus-host interaction at gene expression and molecular level are crucial especially in devising methods to inhibit virus replication in *Aedes* mosquitoes. Previous reports

---

have shown that, besides protein-coding genes, noncoding RNAs such as lncRNAs are also involved in virus-host interaction. In this study, we report a comprehensive catalog of novel lncRNA transcripts in the genome of *Ae. albopictus*. We also show that the expression of lncRNAs was altered upon infection with dengue and Zika. Additionally, depletion of certain lncRNAs resulted in increased replication of dengue and Zika; hence, suggesting potential association of lncRNAs in virus infection. Results of this study provide a new avenue to the investigation of mosquito-virus interactions, especially in the aspect of noncoding genes.

Introduction

The Asian tiger mosquito, *Aedes albopictus* (*Ae. albopictus*) is an important vector of arboviruses such as dengue virus (DENV), Chikungunya virus (CHIKV), and Zika virus (ZIKV). Due to its invasiveness and aggressive spread, *Ae. albopictus* has widespread geographic distribution; hence, poses serious health threat globally. Although *Ae. albopictus* is considered as a less competent vector of DENV than *Aedes aegypti* (*Ae. aegypti*), *Ae. albopictus* was responsible for dengue outbreaks in Hawaii, China, Europe, and Central Africa, particularly Gabon [1–7]. Due to its epidemiological importance, *Ae. albopictus* has become a key focus for many scientific investigations.

Long noncoding RNAs (lncRNAs) are typically defined as RNA species of more than 200 nucleotides (nt) that do not encode protein [8]. Similar to protein-coding mRNAs, lncRNAs are the products of Pol II, and they undergo polyadenylation, capping and alternative splicing [9]. Due to their mRNA-like features, lncRNAs are usually represented in RNA-seq datasets. Once considered as junk and noise in RNA-seq data, findings from many studies provide evidence that lncRNAs perform various biological functions. For example, metastasis associated in lung adenocarcinoma transcript 1 (*Malat1*), is important for alternative splicing of mRNA transcripts, and gene regulation [10,11]. In addition, research in *Drosophila melanogaster* (*D. melanogaster*) provides evidence on the functional roles of lncRNAs in spermatogenesis [12] and neurogenesis [13].

Meanwhile, studies in *Ae. aegypti* demonstrated that lncRNAs might be involved in early embryonic development and virus infection [14,15]. A fraction of *Ae. aegypti* lncRNAs was highly expressed in blood-fed ovaries, where the process of oogenesis takes place. The same set of lncRNAs were also found to be enriched in early embryo (0–8 hour), indicating that they were maternally inherited, and performed critical functions during maternal to zygotic transition stage [14]. Besides, RNAi-mediated silencing of *Ae. aegypti* lncRNA (lincRNA_1317) resulted in enhanced DENV replication in *Ae. aegypti* cells, suggesting the potential role of lncRNAs in virus-host interaction [15]. We speculate that, similar to *Ae. aegypti*, lncRNAs in *Ae. albopictus* may participate in various regulatory functions including virus infection.

Here, we report genome-wide identification of lncRNAs in *Ae. albopictus*, and their possible involvement in DENV and ZIKV infection. In this present work, we defined a high-confident set of 10, 867 novel lncRNA transcripts, derived from 9,298 genomic loci. Our newly predicted *Ae. albopictus* lncRNAs shared similar characteristics with other species such as shorter in length, lower GC content, and lower sequence conservation compared to protein-coding transcripts. To investigate the transcriptional response of lncRNAs upon virus infection, we generated RNA-seq libraries of *Ae. albopictus* cells infected with DENV and ZIKV. Similar to previous studies in *Ae. aegypti* [15,16], the expression level of *Ae. albopictus* lncRNAs was altered during virus infection. Furthermore, CRISPR-Cas9 induced mutation of

selected lncRNAs in *Ae. albopictus* cells led to enhancement of DENV and ZIKV replication, suggesting the potential role of *Ae. albopictus* lncRNAs in antiviral response. Collectively, the data presented here provide valuable resources for future studies of lncRNAs in *Ae. albopictus* and facilitate a better understanding of the potential role of lncRNAs in virus-host interactions.

Materials and methods

Preparation of public RNA-seq datasets

A total of 148 public RNA-seq datasets were downloaded from NCBI Sequence Reads Archive (SRA) with accession number SRP219966, [17] SRP050258 [18], SRP018112 [19], SRP077936 [20], SRP056407 [21], and SRP164784 [22]. All 148 RNA-seq libraries used in this study were listed in **S1 Table**. FASTQC was used to check the quality of raw reads. Adapters and low quality reads (<20 Phred score) were removed using Trimmomatic version 0.38 [23].

Mapping and transcriptome assembly

Reads were individually mapped onto the *Ae. albopictus* genome (Genome version: AalbF2, assembly: GCA_006496715.1, NCBI) using Hierarchical indexing for spliced alignment of transcripts (HISAT2) version 2.1.0 [24]. The resulting alignment files were used as input in Stringtie software (version 1.0.1) for transcriptome assembly [25]. Annotation file (GTF format) of *Ae. albopictus* from NCBI (Aalbo_primary.1_genomic.gtf) was used to guide the transcriptome assembly. A minimum of 200 bp was set for the length of the assembled transcripts. Then, the Stringtie output files were merged into a single transcriptome assembly using Stringtie merge. The resulting merged assembly was then compared to the reference annotation using Gffcompare (https://github.com/gpertea/gffcompare). Transcripts with class code "u" (intergenic), "i" (intronic), and "x" (antisense) were kept for lncRNA prediction. These class codes were chosen because lncRNAs were categorized based on their positions relative to protein-coding genes provided in the reference annotation [14,26–29].

Novel lncRNA prediction

TransDecoder [30] was used to predict the presence of long open reading frame (ORF) within transcripts with class code "u", "i", and "x". A minimum of 100 amino acids was set as threshold of ORF length in TransDecoder analysis, and transcripts having long ORF were removed. Long ORF was defined as transcripts having more than 100 amino acids. The remaining transcripts were then subjected to coding potential assessment analysis. Coding Potential Calculator 2 (CPC2) [31], Predictor of long noncoding RNAs and messenger RNAs based on an improved *k*-mer scheme (PLEK) [32], and Coding Potential Assessment Tool (CPAT) [33] were used to evaluate coding potential of the transcripts. For CPAT analysis, a coding probability threshold of < 0.3 was set [14,15]. Transcripts that passed CPAT threshold, and detected as noncoding by CPC2 and PLEK, were kept for subsequent analyses. To avoid false positive prediction, BLASTX algorithm was used to search for sequence similarity of the potential lncRNA transcripts in Swissprot database. Transcripts with high sequence similarity (E-value < 1e-5) were removed. Transcripts without strand information, and those located less than 1 kb from protein-coding genes on the same strand were removed. Finally, to remove lowly expressed transcripts, transcripts that had an expression of more than 1 transcript per million (TPM) in at least 5 RNA-seq libraries were kept. The expression was determined by featureCounts [34], and alignment files (BAM format) generated by HISAT2 were used as input.

Cell culture and virus

C6/36 cells (ATCC: CRL-1660) were cultured in Leibovitz's L-15 medium (Gibco, 41300039) supplemented with 10% Fetal Bovine Serum (FBS, Gibco, 10270) and 10% Tryptose Phosphate Broth (TPB) solution (Sigma, T9157). The C6/36 cells were grown at 25°C without CO_2. BHK-21 cells (ATCC: CCL-10) were used for virus titration. BHK-21 cells were cultured in Dulbecco's modified Eagles Medium (DMEM, Gibco, 11995065) supplemented with 10% FBS (Gibco, 10270) at 37°C and 5% CO_2. In this study, DENV serotype 1 (DENV-1) of Hawaiian strain, and ZIKV of strain H/PF/2013 were used. Both viruses were propagated in C6/36 cells. Virus titer of DENV and ZIKV was determined using tissue culture infectious dose 50 (TCID50) assay. DENV-1 (Hawaiian strain) was a gift from David Perera, University Malaysia Sarawak, and ZIKV (H/PF/2013 strain) was a gift from Shee Mei Lok, Duke-NUS Medical School, Singapore.

Tissue Culture Infectious Dose 50 (TCID50) assay

TCID50 assay was conducted to determine infectious virus titre using a method that has been previously described [35–37]. For each determination, a 96-well plate culture of confluent BHK-2 was inoculated with a 150 μl mixture of 100 μl medium and 50 μl of serial 10-fold dilutions of clarified cell culture supernatant media. Each row included 6 wells of the same dilution and two negative controls. Plates were incubated at at 37°C and 5% CO_2 for 5 days. Each well was observed for the presence or absence of cytopathic effect (CPE). The determination of TCID50/ml was performed using the Reed and Muench method [38].

Virus infection, RNA extraction and sequencing

C6/36 cells were infected with DENV-1 and ZIKV at a multiplicity of infection (MOI) of 0.25. At 3 days post infection (dpi), the culture medium was removed, and total RNA extraction was carried out using miRNeasy Mini Kit 50 (Qiagen, 217004) according to the manufacturer's protocol. RNA concentration was determined using ND-2000 instrument (NanoDrop Technologies), and RNA integrity was determined using 2100 Bioanalyzer (Agilent). Libraries were constructed using Illumina TruSeq RNA Sample Prep kit (Illumina, San Diego) using one microgram (μg) of total RNA as starting material. The prepared libraries were sequenced on the Illumina HiSeq 2000 platform to generate 150 bp paired-end reads. Raw reads were deposited in NCBI with accession SRP221722 (https://www.ncbi.nlm.nih.gov/sra/?term=SRP221722).

Identification of differentially expressed lncRNAs upon infection

RNA-seq libraries generated in this study were mapped to the *Ae. albopictus* genome (AalbF2, assembly: GCA_006496715.1, NCBI) using HISAT2 [24]. The expression was quantified using featureCounts [34]. GTF file annotated with the newly predicted lncRNA was used as reference. Read counts computed by featureCounts were then subjected to differential expression analysis using edgeR [39] in R/Bioconductor environment.

Reverse transcription and real-time quantitative PCR (qPCR)

cDNA synthesis was performed with one μg of total RNA in 20 μl reactions using the iScript Reverse Transcription Supermix (Bio-Rad Laboratories, California; cat. no. 1708840). Quantitative PCR was performed using iTaq™ Universal SYBR Green Supermix (Bio-Rad Laboratories, California; cat. no. 1725120), in a total of 20 μl reactions. All experiments were done in three biological and technical replicates. *RPS17* and *RPL32* were used as reference genes, and all qPCR runs were carried out on the BioRad CFX96 qPCR platform. The $2^{-\Delta\Delta CT}$ [40] method

was used to estimate the relative expression of each lncRNA, and the results were statistically analyzed using student's t-tests. List of primers used in this study can be found in **S8 Table**.

Functional annotation and enrichment

Functional annotation was performed using Pannzer [41] using amino acid sequences of lncRNA-target genes as input. Gene enrichment analysis was done using g;Profiler [42] and g: SCS threshold was used for multiple testing correction.

CRISPR-Cas9 induced mutation of selected lncRNAs

Single guide RNA (sgRNA) targeting exon 1 of lncRNA_27639.2 and XR_003899061.1 were designed using CHOPCHOP [43]. The sgRNAs were cloned into the *Drosophila* CRISPR-Cas9 vector pAc-dU6-sgRNA-Cas9 [44], which was a kind gift from Ji-Long Liu (Addgene plasmid 49330). This plasmid has been shown to work well in *Ae. aegypti* (Aag2) cell [45]. Detailed protocol for sgRNA cloning was previously described [44]. C6/36 cells ($\sim3x10^6$) were transfected with 2.5 μg of gRNA-containing plasmid using Lipofectamine 2000 (Life Technology). Plasmid without gRNA (empty vector) was used as control. After 72 hours post-transfection, Cas9-expressing cells were selected using puromycin (5 μg/mL) for 5 days. Cells were then resuspended and $\sim5x10^5$ cells were added to each well of 12-well plate. Cells were allowed to settle for ~2 hour, before being infected with DENV-1 (Hawaiian strain) and ZIKV (H/PF/2013 strain) at MOI of 0.25. After 3 dpi, cells were frozen and thawed twice, causing the cells to rupture and release viral particles. The samples were then centrifuged, and the cell culture supernatant containing viral particles were used for TCID50 assay. Expression level of lncRNA after CRISPR-Cas9 induced mutation was evaluated by qPCR. RNA isolation and reverse transcription were carried out as previously described. Meanwhile, indel mutation within lncRNA loci was confirmed by sequencing of the PCR products spanning the cleavage sites. The PCR products were cloned into pGEM-T Easy vector (Promega, cat. no: A1360), and ten colonies were selected for sequencing.

Results

Identification of novel lncRNAs in *Ae. albopictus*

To identify novel lncRNAs in *Ae. albopictus*, a total of 148 RNA-seq libraries were used as raw inputs. The list of all 148 RNA-seq datasets used in this study can be found in **S1 Table**. The RNA-seq libraries used were derived from many tissues, organs, developmental stages and cell lines; hence, allowing our lncRNA prediction to be more comprehensive. The lncRNA prediction pipeline used here (**Fig 1**) were modified and adapted from previous works [14,15,26,46–51]. After mapping raw reads to the reference genome using HISAT2, we assembled and merged the assemblies into a single unified transcriptome. Using Gffcompare, we selected transcripts with class code "i", "u", and "x" (35,863 transcripts), which indicate intronic, intergenic, and antisense to reference genes respectively [14,26]. In this study, we were able to recover all *Ae. albopictus* annotated transcripts including known lncRNAs (7,609) using Stringtie; hence, validating our transcriptome assembly method.

We then filtered the transcripts harboring open-reading frame (ORF) and coding potential. The ORF prediction was performed using TransDecoder [30], while coding potential assessment was performed using three different algorithms, namely CPC2 [31], CPAT [33], and PLEK [32]. A total of 22,057 transcripts did not have an ORF as determined from TransDecoder. Of these, 16,630 transcripts passed coding potential assessment by CPC2, CPAT, and PLEK. To avoid false positive prediction, we performed BLASTX on these 16,630 transcripts against Swissprot database. We removed transcripts having significant BLASTX hits (E-
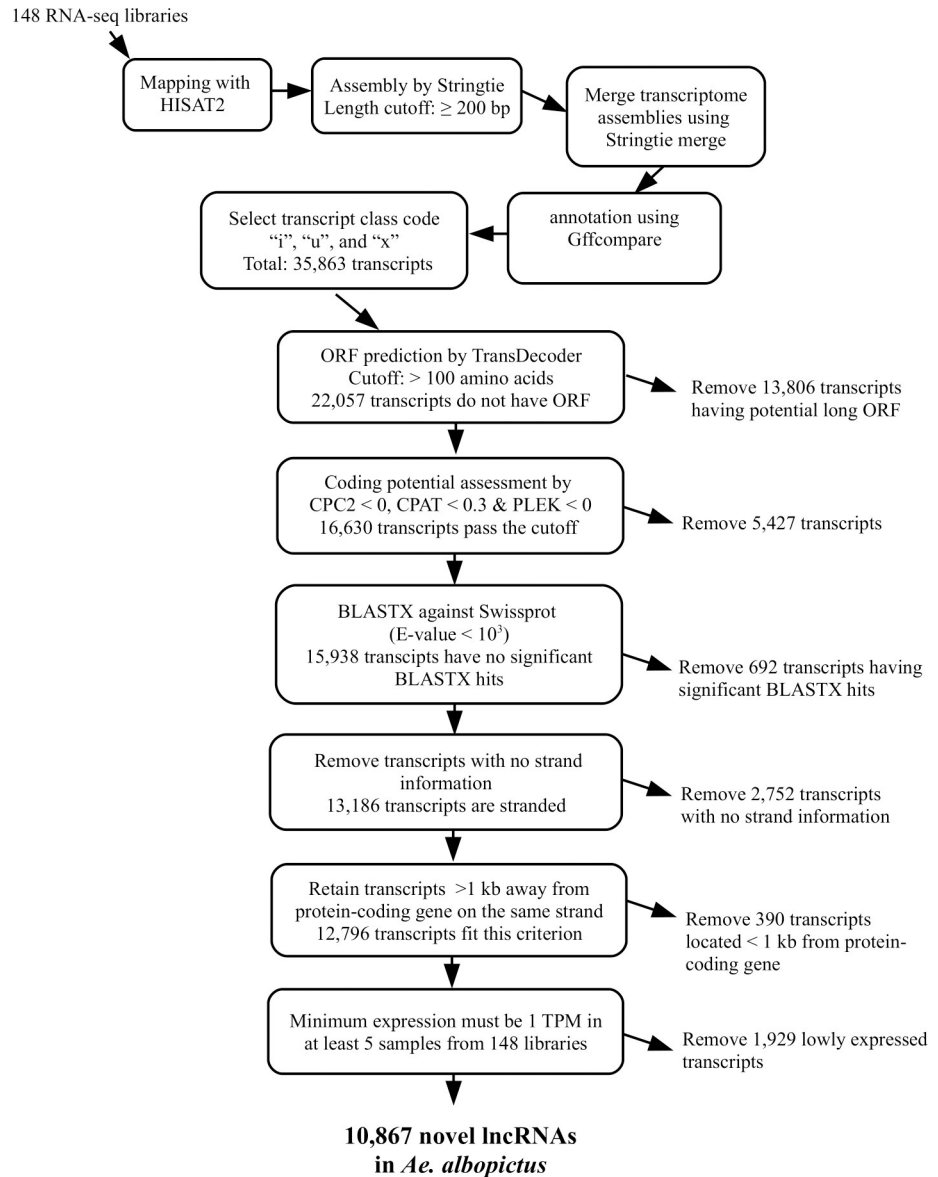
148 RNA-seq libraries

Mapping with HISAT2 → Assembly by Stringtie Length cutoff: ≥ 200 bp → Merge transcriptome assemblies using Stringtie merge

annotation using Gffcompare

Select transcript class code "i", "u", and "x" Total: 35,863 transcripts

ORF prediction by TransDecoder Cutoff: > 100 amino acids 22,057 transcripts do not have ORF → Remove 13,806 transcripts having potential long ORF

Coding potential assessment by CPC2 < 0, CPAT < 0.3 & PLEK < 0 16,630 transcripts pass the cutoff → Remove 5,427 transcripts

BLASTX against Swissprot (E-value < 10³) 15,938 transcripts have no significant BLASTX hits → Remove 692 transcripts having significant BLASTX hits

Remove transcripts with no strand information 13,186 transcripts are stranded → Remove 2,752 transcripts with no strand information

Retain transcripts >1 kb away from protein-coding gene on the same strand 12,796 transcripts fit this criterion → Remove 390 transcripts located < 1 kb from protein-coding gene

Minimum expression must be 1 TPM in at least 5 samples from 148 libraries → Remove 1,929 lowly expressed transcripts

10,867 novel lncRNAs in *Ae. albopictus*

Fig 1. Overview of lncRNA identification pipeline.

https://doi.org/10.1371/journal.pntd.0008351.g001

value $< 10^3$). Next, we removed unstranded transcripts and those located less than 1 kb from protein-coding genes. Finally, we only retained transcripts that had minimum expression of 1 TPM in at least five out of the 148 RNA-seq libraries used in this study. This was done to remove lowly expressed transcripts which might be falsely identified as novel lncRNAs. In total, 10, 867 putative novel lncRNA transcripts, which derived from 9,298 genomic loci, were identified in *Ae. albopictus* genome. Of these 10,867 transcripts, 5,809 were located in the intergenic regions, 4,139 were intronic, while the remaining 919 were antisense to annotated genes. Full list of predicted novel lncRNAs and their genomic coordinates can be found in **S2 Table**. Current *Ae. albopictus* annotation catalogs 7,609 lncRNA transcripts, and we classified these transcripts as known lncRNAs. Therefore, in *Ae. albopictus* genome, the number of both novel and known lncRNAs were 18,476.

Ae. albopictus lncRNAs shared similar features with other species

Next, we determined the similarities of *Ae. albopictus* lncRNAs to other species, such as shorter in length, low GC content, and low sequence conservation [14,26,46,47]. *Ae. albopictus* lncRNAs were found to be significantly shorter in size compared to protein-coding transcripts (Kolmogorov-Smirnov test (KS-test) *p*-value < 2.2e-16, **Fig 2A**). The average length of novel and known lncRNAs in *Ae. albopictus* were 569.67 bp and 763.68 bp respectively, whereas protein-coding transcripts had a mean size of 2650.1 bp. Short in size is a typical feature of lncRNAs; hence, our findings are congruent with lncRNAs identified in other living organisms [14,46,52]. Besides, *Ae. albopictus* known and novel lncRNAs exhibited slightly lower GC content in comparison to protein-coding transcripts (**Fig 2B**). This lower GC or AT enriched characteristic is another notable feature of lncRNA transcripts [14,15,26,46,47]. Mean GC content of novel lncRNAs, known lncRNAs and protein-coding transcripts were 42%, 44%, and 48% respectively. We also analyzed GC content of other noncoding sequences in the genome such as 5'UTR and 3'UTR. We discovered that both of them had lower GC content in comparison to protein-coding transcripts. The average GC content of 5'UTR and 3'UTR were 44% and 36% respectively.

As previous studies showed that lncRNAs demonstrate low sequence conservation even among closely related species [14,15,48], we investigated the validity of this in *Ae. albopictus*. To assess the level of sequence similarity, we performed used BLASTN algorithm to investigate the level of similarity among *Ae. albopictus* lncRNAs sequences with other closely related insect genomes such as *Ae aegypti*, *D. melanogaster*, *Anopheles gambiae* (*An. gambiae*), *Anopheles darlingi* (*An. darlingi*), and *Culex quinquefasciatus* (*Cx. quinquefasciatus*) [14,15]. BLAST algorithm bit score was used as an indicator for sequence similarity [14,15]. We used BLASTP algorithm to assess the level of similarity among *Ae. albopictus* proteins with other insects including *Ae. aegypti*, *D. melanogaster*, *An. gambiae*, *An. darlingi*, and *Cx. quinquefasciatus* In general, protein-coding transcripts were found to be more conserved than lncRNAs (**Fig 3A**). Some *Ae. albopictus* lncRNAs showed high
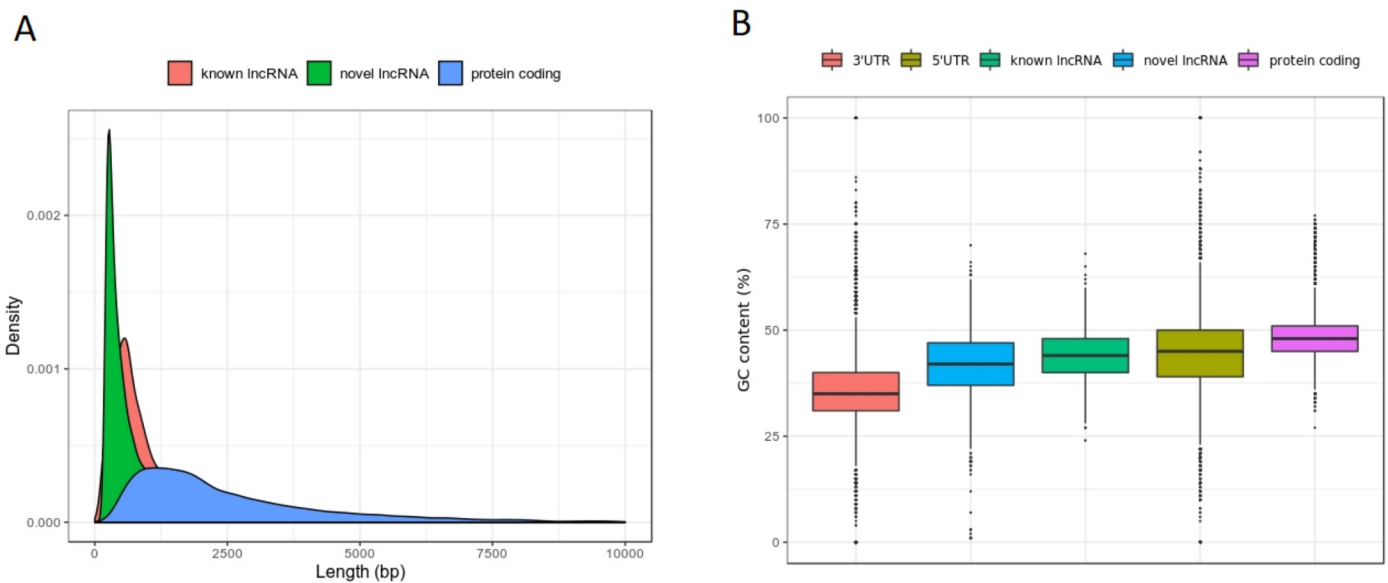

Fig 2. Length and GC content *Ae. albopictus* lncRNAs. A) Sequence length distribution of lncRNA and protein-coding transcript. B) GC content of lncRNA, protein-coding transcript, 3'UTR, and 5'UTR.
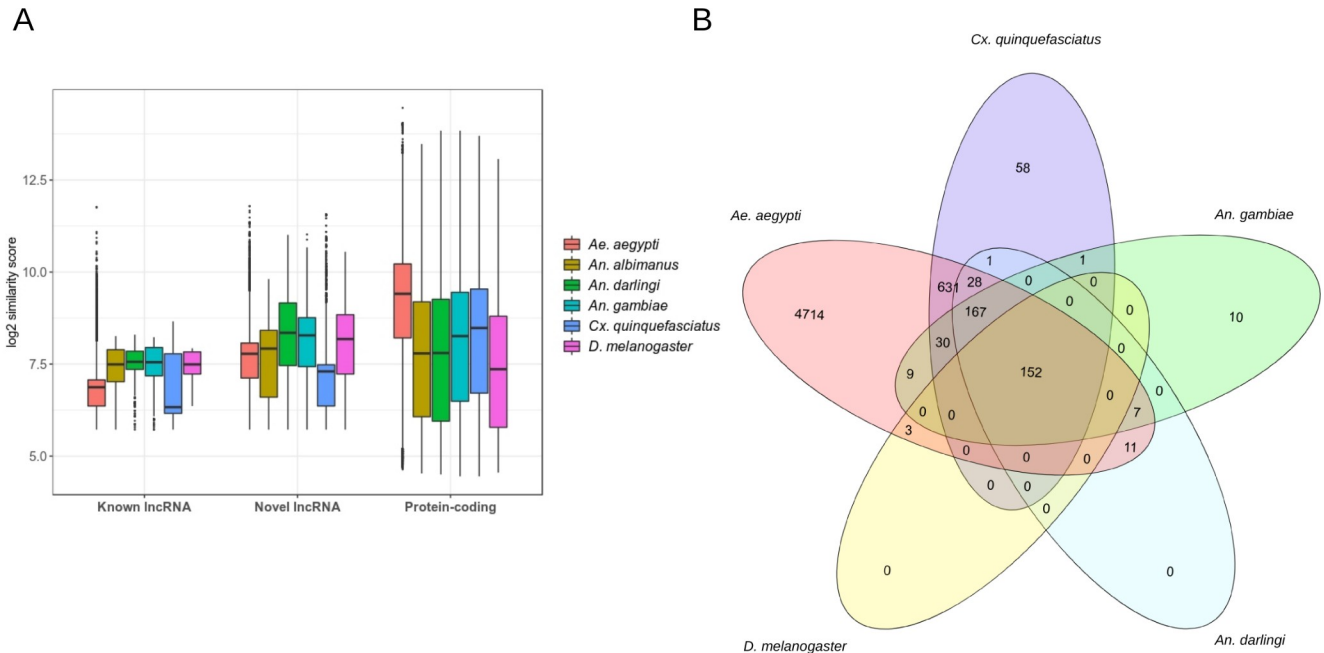
Fig 3. Conservation analysis of *Ae. albopictus* lncRNAs. A) Similarity bit score of of lncRNA and protein sequences. In general, protein-coding mRNA is more conserved than lncRNA. B) The Venn diagram shows the number of *Ae. albopictus* lncRNAs that display high conservation with other species according to BLASTN algorithm (E-value < 10-e5).

degree of sequence similarity (E-value $< 10-^5$) with *Ae. aegypti*, presumably because the two species are very closely related. (**Fig 3B**). However, the fraction of *Ae. albopictus* lncRNAs that displayed high level of sequence similarity (E-value $< 10-^5$) with *Ae. aegypti* were quite low. For instance, only 3,676 (33.8%) novel lncRNAs and 2,053 (29%) known lncRNAs shared high sequence similarity with *Ae. aegypti*. On the other hand, 38,790 (96.7%) *Ae. albopictus* proteins demonstrated high degree of sequence similarity with *Ae. aegypti* protein sequences.

Overview of sequencing data generated from DENV and ZIKV-infected *Ae. albopictus* cells

Studies on *Ae. aegypti* transcriptomes provide evidence that lncRNAs could be involved in DENV and ZIKV-mosquito interaction [15,16]. To examine the relevance of this to *Ae. albopictus*, we generated paired-end RNA-seq libraries from larval-derived *Ae. albopictus* cell line (C6/36) which had been infected with DENV and ZIKV. A total of 9 samples were sequenced using Illumina HiSeq 2000 platform, and they comprised of three replicates of mock, DENV, and ZIKV-infected C6/36 cells. In total, 426 million paired-end reads were generated in this study (**S3 Table**). More than 70% of the reads (average of 35.5 million reads per sample) mapped to *Ae. albopictus* reference genome (**S3 Table**), providing suffi- cient coverage for genome-wide transcriptomic analysis, specifically in detecting differen- tially expressed transcripts. Hierarchical clustering based on the expression values (CPM) of all the expressed genes showed that the replicates were grouped into their three distinct groups (**S1 Fig**). This unsupervised clustering also showed that the data formed two main clusters, i.e., uninfected, and virus-infected (DENV and ZIKV) samples.

Differential expression of *Ae. albopictus* lncRNAs upon DENV and ZIKV infection

In this study, we considered genes with fold change (FC) of more than or equal to 2 in either direction, and false discovery rate (FDR) of less than 0.05 as differentially expressed. A total of 433 and 338 differentially expressed lncRNAs (FC ≥ |2|, FDR < 0.05) were identified respectively in DENV and ZIKV samples. In DENV-infected cells, 298 lncRNAs were found to be upregulated, while the remaining 136 transcripts were downregulated. Meanwhile, 216 and 122 lncRNAs were upregulated and downregulated respectively in C6/36 cells infected with ZIKV. Volcano plot analysis was used to depict the distribution of lncRNAs with their corresponding P-value and log FC (**Fig 4A**), and hierarchical clustering analysis was performed to cluster the differentially expressed lncRNAs among different replicates (**Fig 4B**). Similar lncRNAs were found to be differentially expressed; 11 lncRNAs were upregulated and 12
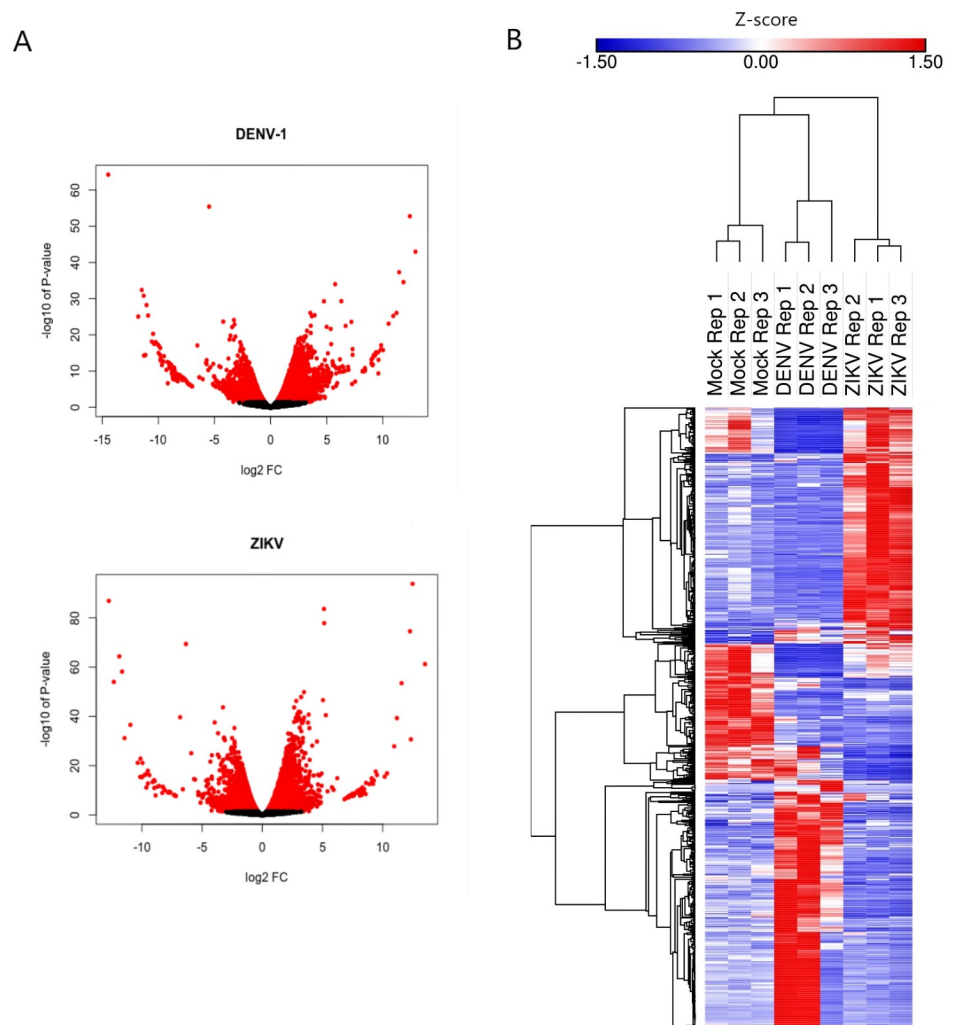

Fig 4. lncRNA expression profiles of DENV and ZIKV-infected *Ae. albopictus* cells. A) Volcano plot of differentially expressed *Ae. albopictus* lncRNAs in DENV-1 and ZIKV-infected C6/36 cells. Dots with red color represent lncRNAs with P-value < 0.05. B) Hierarchical clustering of differentially expressed *Ae. albopictus* lncRNA was done in Morpheus (https://software.broadinstitute.org/morpheus), using one minus Pearson correlation metric with average linkage method.

https://doi.org/10.1371/journal.pntd.0008351.g004

lncRNAs were downregulated in both DENV and ZIKV infected cells (**Table 1**). The similar changes seen in these 23 lncRNAs post-infection suggest that they may be key players, regardless of the type of viral infection. On the other hand, the number of differentially expressed protein-coding mRNAs were higher, i.e., 1,718 and 1,098 mRNAs were differentially expressed (FC ≥ |2|, FDR < 0.05) in DENV and ZIKV-infected cells respectively. List of differentially expressed lncRNAs can be found in **S4 Table**.

We randomly selected ten differentially expressed lncRNAs from RNA-seq analysis for qPCR. The ten chosen candidates were upregulated or downregulated in both DENV and ZIKV-infected samples. We validated the differentially expressed lncRNAs using qPCR from the same RNA samples used for Illumina sequencing. *RPS17* and *RPL32* were used as reference genes [53,54]. Our preliminary tests showed that the expression level of the two reference genes were unchanged in DENV, and ZIKV-infected C6/36 cells; hence, qualifying them as suitable references in this study. We also performed qPCR on subsets of lncRNAs that were differentially expressed only in DENV-infected cells. Similar experiments were done for ZIKV-infected cells. Overall, the qPCR results of all lncRNAs were consistent with RNA-seq data; hence, validating the bioinformatic approach used in this study (**S2 Fig**).

Functional analysis of the differentially expressed lnRNAs

To determine the putative regulatory roles of *Ae. albopictus* lncRNAs upon virus infections, we searched for protein-coding genes located within 100 kb (upstream and downstream) from

Table 1. Differentially expressed lncRNAs that are common in DENV-1 and ZIKV transcriptome.

Transcript ID	DENV (logFC)	ZIKV (logFC)
lncRNA_27639.2	2.876	3.008
lncRNA_54618.1	2.145	3.807
lncRNA_59941.1	2.689	3.133
XR_002128316.2	2.090	3.035
lncRNA_49250.1	2.229	2.067
lncRNA_56589.2	2.092	2.205
XR_003895025.1	2.950	5.359
XR_003896798.1	2.603	3.027
lncRNA_13375.1	3.159	4.348
XR_003899061.1	3.227	2.319
lncRNA_61942.3	2.689	3.173
lncRNA_64141.1	-5.695	-4.666
lncRNA_19975.1	-2.020	-2.326
lncRNA_46959.1	-2.314	-2.980
XR_003895284.1	-2.942	-6.826
lncRNA_50256.1	-5.361	-4.201
XR_003893587.1	-2.649	-2.396
lncRNA_32881.1	-3.292	-2.492
lncRNA_6487.1	-2.456	-3.299
lncRNA_13794.1	-5.479	-3.143
XR_003892459.1	-2.068	-2.094
lncRNA_1008.1	-2.637	-2.058
lncRNA_53241.1	-2.374	-2.513

The value of logFC was determined by edgeR. All have FDR < 0.05

https://doi.org/10.1371/journal.pntd.0008351.t001

differentially expressed lncRNAs as potential *cis*-regulated genes [28,55] and found a total of 198 and 153 coding genes which fulfilled that criteria in the DENV and ZIKV transcriptome respectively. As lncRNAs also have the ability to regulate genes in trans or distal regions [28,55–57], we searched for protein-coding genes that were co-expressed with the differentially expressed lncRNAs based on Pearson correlation (coefficient > 0.95 or < - 0.95, P-value < 0.05). In the transcriptome of DENV-infected cells, 110 coding genes were co-expressed with the differentially expressed lncRNAs. Meanwhile, we detected 38 coding genes having high correlation in expression with differentially expressed lncRNAs in ZIKV transcriptome. We categorized the co-expressed genes as putative trans-regulated genes in *Ae. albopictus* cells upon DENV and ZIKV infection. Results of co-expression analysis between lncRNAs and protein-coding mRNAs can be found in **S5 Table**.

We selected the potential *cis* and *trans*-regulated genes for functional annotation and enrichment analyses. Previous reports demonstrated that upon virus infection in *Aedes* mosquitoes, there were changes in the expression level of host immune-related genes [16,20,58–60]. By conducting functional annotation using Pannzer [41], we discovered 13 lncRNAs that potentially regulated genes involved in immune-related functions (**S5 Table**). Based on gene ontology (GO), these genes were implicated in innate immune response (GO:0045087), defense response to Gram-positive (GO:0050830) and Gram-negative (GO:0050829), defense response to bacterium (GO:0042742), regulation of defense response to virus (GO:0050688), virus maturation (GO:0019075), and innate immune response (GO:0045087).

Previous studies reported that genes encoding clip-domain serine protease, and metalloproteinase were differentially expressed in *Aedes* mosquitoes upon virus infection. For example, clip-domain serine protease and metalloproteases were found to be significantly upregulated and downregulated respectively upon DENV infection in *Ae. albopictus* midguts [20]. Meanwhile, another study in *Ae. aegypti* reported that metalloproteinase and serine protease showed an increase in expression upon infection with ZIKV [16]. In this study, we discovered lncRNAs that potentially regulate genes that were also involved in serine-type endopeptidase activity (GO:004252, GO:004253, GO:004254) and metalloendopeptidase activity (GO:004222).

We then performed functional enrichment analysis of lncRNA-regulated genes using g;Profiler [42]. We discovered that 32 and 22 GO terms were significantly enriched (P-value < 0.05) in ZIKV and DENV-transcriptome respectively. A total of 8 GO terms were significant and shared by both transcriptomes, suggesting that lncRNAs may target genes that perform similar regulatory functions upon DENV and ZIKV infections (**S7 Table**). Among the significant GO terms were cellular processes, ion/protein binding, metabolic processes, and DNA binding.

Mutation of selected lncRNAs and enhancement of DENV and ZIKV replication

To verify the role of *Ae. albopictus* lncRNAs in virus infection, we used CRISPR-Cas9 to induce mutation of selected lncRNAs in C6/36 cells following DENV and ZIKV infection. We discovered that CRISPR-Cas9 induced mutations in two lncRNAs loci resulted in the increase of DENV-1 and ZIKV infectious titer by at least two folds after three dpi (**Fig 5A**). One of the lncRNAs was novel (transcript ID: lncRNA_27639.2, loci ID: XLOC_029733) and another was known lncRNA (XR_003899061.1, ncbi gene ID: LOC115270134). Indels for both lncRNAs were confirmed by sequencing the PCR products spanning the cleavage sites, and the expression level lncRNAs after CRISPR-Cas9 induced mutation were validated by RT-qPCR (**S3 Fig**). Reduction in the level of expression indicated that mutations induced by CRISPR-Cas9 resulted in the silencing of the two aforementioned lncRNAs in C6/36 cells.
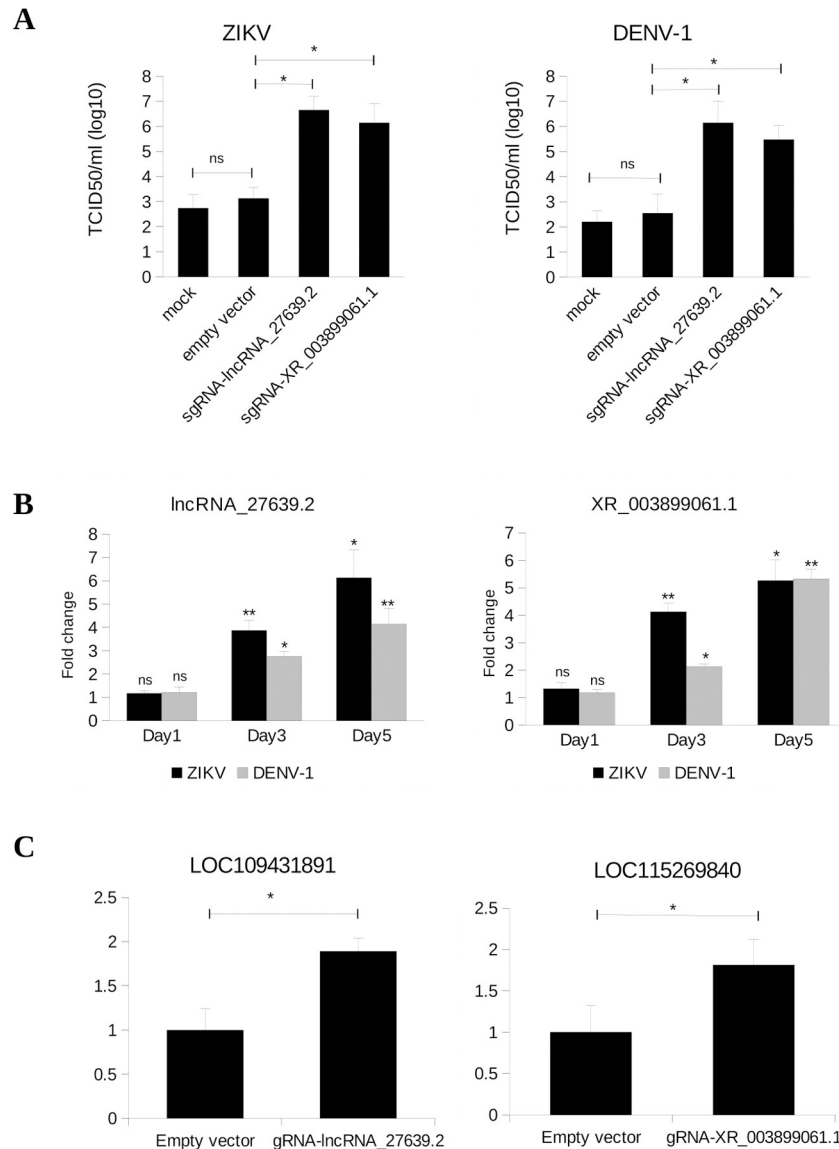
Fig 5. Possible involvement of lncRNAs in DENV and ZIKV replication in C6/36 cell. A) Titer of ZIKV and DENV-1 after 3 dpi as determined by TCID50 assay. In mock sample, C6/36 cells were transfected with transfection reagent only (Lipofectamine 2000). Empty vector refers to the *Drosophila* CRISPR-Cas9 vector (pAc-dU6-sgRNA-Cas9) without sgRNA targeting any gene. B) Expression level of lncRNA_27639.2 and XR_003899061.1 at 1, 3 and 5 dpi. At 1 dpi, the fold change between infected and uninfected samples were statistically insignificant. C) Expression level of LOC109431891 and LOC115269840 in C6/36 cells that had mutations in lncRNA_27639.2 and XR_003899061.1 respectively. Expression levels were normalized to empty vector samples. The expression of both genes (LOC109431891 and LOC115269840) increased in lncRNA-mutated C6/36 cells compared to empty vector samples. For A), B), and C), bars indicate mean +/- SEM of three independent experiments. Student's *t*-test was used to assess the statistical significance. * p<0.05; **p<0.01.

https://doi.org/10.1371/journal.pntd.0008351.g005

A previous report showed that the expression of *Ae. aegypti* lincRNA_1317 increased substantially following the progression of infection [15]. When this lncRNA was silenced by RNAi, it led to an enhancement of DENV replication in *Ae. aegypti* cell, implying the involvement of certain lncRNAs in viral infection. We next analyzed our lncRNAs to idenfify if they also played a key role in anti-viral response. Based on RNA-seq and RT-qPCR analysis, both

lncRNAs (lncRNA_27639.2 and XR_003899061.1) were upregulated following infection with DENV and ZIKV at 3 dpi (**S2 Fig**). We investigated the level of expression at early (1 dpi) and late (5 dpi) stage of DENV and ZIKV infection. Interestingly, the expression of both lncRNAs displayed an increasing pattern along the progression of infection (**Fig 5B**), indicating possible involvement of lncRNA_27639.2 and XR_003899061.1 in defense against viruses in *Ae. albopictus*.

lncRNA_27639.2 might potentially regulate a gene that encodes for solute carrier family 35 member G1 (LOC109431891). This gene was located 1.9 kb downstream of lncRNA_27639.2 on the same strand; hence, making it a potential *cis*-regulated gene. GO annotation of solute carrier family 35 member G1 (LOC109431891) showed that it participates in calcium sensing and homeostasis, and it is a part of endoplasmic reticulum [61]. Meanwhile, XR_003899061.1, was located in *cis* to and 61 kb away from a gene encoding clathrin interactor 1 protein (LOC115269840), which is involved in clathrin assembly and trafficking [62]. Interestingly, calcium homeostasis, clathrin-mediated endocytosis, and vesicular trafficking are essential for viral replication in host cells [63–65]; thereby, further supporting the idea that lncRNA_27639.2 and XR_003899061.1 might be involved in antiviral response.

We then investigated the expression of LOC109431891 and LOC115269840 in C6/36 cells that had CRISPR-Cas9 induced mutations in lncRNA_27639.2 and XR_003899061.1 respectively. Our qPCR results showed that the expression of both LOC109431891 and LOC115269840 increased nearly two-folds in C6/36 cells harboring mutations in the lncRNA loci (**Fig 5C**). Although the fold change of both genes was only approximately two-folds, the changes were statistically significant (P-value < 0.05). Therefore, we speculate that lncRNA_27639.2 and XR_003899061.1 may be involved in antiviral response by limiting the expression of LOC109431891 and LOC115269840 respectively.

Discussion

Even though protein-coding genes have always been the central focus in many host-virus interaction studies, recently, there are few investigations that provided evidence on the possible involvement of lncRNAs in *Aedes* mosquitoes, especially *Ae. aegypti* [15,16]. Next-generation sequencing and bioinformatics analyses give scientists the opportunity to comprehensively identify lncRNAs in various species. Here, we provide a comprehensive annotation of novel lncRNAs in *Ae. albopictus* using the latest genome version (AalbF2). A previous study reported a total of 2,623 lncRNAs in *Ae. albopictus* [22]. Nonetheless, the identification was done using previous version of genome assembly (AaloF1). Due to the utilization of short-read sequencing and high levels of repetitive DNA sequences, AaloF1 assembly becomes highly fragmented and consists of more than 150, 000 scaffolds [66]. Therefore, in this study, we used the latest genome version, AalbF2 assembly [67], which had reduced scaffold number, and produced a substantially complete gene set with lower duplications as compared to the previous genomes of *Ae. albopictus* [66,68,69].

In this study, we provided a comprehensive list of lncRNAs in *Ae. albopictus* genome. We used a total of 148 RNA-seq libraries in our lncRNA prediction pipeline. The libraries were generated from various *Ae. albopictus* tissues, body parts, and developmental stages; hence, making a more comprehensive pipeline. Examination of genomic features showed that our newly predicted *Ae. albopictus* lncRNAs shared similar characteristics with lncRNAs from other species [14,15,46,52,70–72]. The characteristics include shorter in size, lower GC content, and lower sequence conservation as compared to protein-coding genes. These findings reflect the reliability of the lncRNA prediction pipeline used in this study. Shorter in length is a common feature of ncRNAs in many species, presumably due to the fact that, unlike protein-

coding gene, lncRNAs are not required to have ORF, stop codon, 5'UTR and 3'UTR. Novel lncRNAs predicted in this study had lower GC content compared to protein-coding genes, which is in congruent with previous finding that concludes noncoding regions of the genome are usually composed of lower GC [73].

Unlike protein-coding genes, low sequence conservation of lncRNAs makes functional prediction challenging. However, studies of lncRNAs suggests they may regulate the expression of their neighboring protein-coding genes [74]; thereby, enabling the prediction of the function of lncRNAs. Another widely used method that facilitates the prediction of lncRNA's target is co-expression analysis by identifying protein-coding genes that exhibit similar expression patterns in trans or in distal regions [28,56,57]. By identifying both cis and trans-regulated genes, we managed to functionally predict lncRNA functions following DENV and ZIKV infection in C6/36 cells.

Immune-related genes have always been implicated in host-virus interaction in *Aedes* mosquitoes [20,59,75]. Interestingly, our results on *Ae. albopictus* revealed a subset of upregulated lncRNAs that potentially regulated genes with GO terms related to immune response, or defense response to viruses and bacteria (GO:0050688, GO:0050830, GO:0050829, GO:0045087, GO:0042742, GO:0098586) (**S6 Table**). This implies an association between lncRNA overexpression and antiviral response. Functional enrichment analysis done in this study further supports the link between lncRNAs and virus infection, which is corroborated by previous studies in *Aedes* mosquitoes [16,76,77]. For example, GO terms associated with metabolic process, ion/protein binding, membrane-bounded organelle, and DNA binding were found to be significantly enriched in DENV and ZIKV infection, all of which have been identified in other transcriptomic studies in *Aedes* mosquitoes [16,76,77].

Reports on the association between silencing of host lncRNAs and replication of pathogens are still limited. One study reported that RNAi-mediated silencing of *Ae. aegypti* lincRNA_1317 led to enhancement of DENV replication [15]. A similar observation was made in HIV-1-infected human T cells, where the expression of lncRNA called NEAT1 was altered, and upon silencing of this lncRNA, the replication of virus was observed to be increased [78]. The depletion of two lncRNAs by CRISPR-Cas9 resulted in increased DENV and ZIKV infection in our studies, further validated the involvement of lncRNAs in virus infection.

Depletion of novel lncRNA, lncRNA_27639.2 (loci id: XLOC_029733), and known lncRNA, XR_003899061.1 (ncbi gene id: LOC115270134) in C6/36 cell led to enhancement of DENV-1 and ZIKV by at least two folds. Functional annotation revealed that lncRNA_27639.2 has a potential to regulate solute carrier family 35 member G1 (LOC109431891), whereas known lncRNA, XR_003899061.1, may regulate clathrin interactor 1 protein (LOC115269840). Solute carrier family 35 member G1 has been implicated in intracellular calcium sensing and homeostasis [61], a process that has been shown to be vital for virus replication [65]. For instance, infection of West Nile Virus (WNV) triggers an influx of calcium ion, and this process seems to be required for initial stage of infection and viral replication. This WNV-induced influx of calcium ion is probably necessary for rearrangement of the ER membrane to accommodate virus replication [79].

Even though there is no report that specifically links solute carrier family 35 member G1 to virus infection, its role in calcium homeostasis, which is important in virus replication, may implicate this gene in virus-host interaction. On the other hand, clathrin interactor 1 protein is essential for trafficking via clathrin-coated vesicles, and is localized at the trans-Golgi network [80]. Clathrin-mediated endocytosis is required for virus to enter cells including *Ae. albopictus* cells [63,81,82]. For example, treatment of C6/36 cells with chlorpromazine, an inhibitor of clathrin assembly, led to inhibition of DENV replication and entry [81]. Meanwhile, inhibition of clathrin-mediated endocytosis via RNAi-mediated

silencing, or drug treatment (chlorpromazine and pitstop2) severely impaired ZIKV entry into human glioblastoma cell T98G cell [83].

In this study, we discovered that the expression level of LOC109431891 and LOC115269840 increased in C6/36 cells that had mutations in lncRNA_27639.2 and XR_003899061.1 respectively (**Fig 5C**). As previously described, solute carrier family 35 member G1 and clathrin interactor 1 protein are suggested to be involved in virus replication. Therefore, we hypothesized that lncRNA_27639.2 and XR_003899061.1 may be involved in antiviral response by limiting the expression level of LOC109431891 and LOC115269840 respectively in C6/36 cells. The functional studies presented here were solely based on immortalized C6/36 cells. One of the major limitations of using cell line in research is the results do not always accurately represent the actual mechanisms in the whole organism. We believed that the lncRNA responses to arbovirus infection in *Ae. albopictus* mosquitoes *in vivo* are more complex. However, results from this study provide foundations for forthcoming explorations of lncRNA functions and their involvement in virus infection in *Aedes* mosquitoes.

Taken together, we provided a comprehensive genome-wide identification and characterization of *Ae. albopictus* lncRNAs and their involvement in DENV and ZIKV infection. Although our understanding on the functions of lncRNAs in *Aedes* mosquitoes is still limited, findings from this study may help future investigations. Further studies are required to functionally dissect the roles of lncRNA_27639.2 and XR_003899061.1 in DENV and ZIKV infection in not only *Ae. albopictus* cells, but in whole adult mosquitoes as well. We hypothesized that there are more lncRNAs related to antiviral response, so further characterization and functional studies should be done. We hope that this work will be an important resource for genetics and molecular studies of lncRNAs in *Ae. albopictus*.

Supporting information

S1 Table. List of 148 RNA-seq libraries used in this study.
(XLSX)

S2 Table. Genomic coordinates of *Ae. albopictus* novel lncRNAs.
(XLSX)

S3 Table. Summary of RNA-seq data.
(DOCX)

S4 Table. List of the differentially expressed lncRNAs.
(XLSX)

S5 Table. Results of co-expression analysis between lncRNAs and protein-coding genes.
(XLSX)

S6 Table. lncRNAs that potentially regulate immune-related genes.
(XLSX)

S7 Table. Significant GO terms found in DENV and ZIKV transcriptome.
(XLSX)

S8 Table. Primers used in this study.
(DOCX)

S1 Fig. Hierarchical clustering of RNA-seq samples. Hierarchical clustering was done using the Euclidean distance method associated with complete linkage.
(TIFF)

S2 Fig. RT-qPCR validation of the differentially expressed lncRNAs. A) Quantitative PCR validation of the differentially expressed lncRNAs in both DENV-1 and ZIKV-infected C6/36 cells. B) Quantitative PCR validation of the differentially expressed lncRNAs in the transcriptome of DENV-1 and ZIKV-infected cells alone. Expression levels were normalized to uninfected samples. Bars indicate mean +/- standard error of the mean (SEM) of three independent experiments. We used Bonferroni method for pairwise multiple comparison. RNA-seq fold change was computed by edgeR.
(TIFF)

S3 Fig. Validation of CRISPR-Cas9 induced mutation in lncRNA loci. A) As determined by RT-qPCR, CRISPR-Cas9 induced mutation resulted in lower expression level of lncRNA_27639.2 and XR_003899061.1 in C6/36 cells. Expression levels were normalized to empty vector sample. All changes in expression level shown in the figure are statistically significant with P-value < 0.05. Bars indicate mean +/- SEM of three independent experiments. Student's *t*-test was used for all comparisons. B) Sequencing of indel mutations within lncRNA loci after 5 days of puromycin selection. PCR products flanking the cleavage site (black triangle) were cloned and sequenced from C6/36 cells transfected with vector containing lncRNA-specific sgRNAs. Majority of the clones showed deletions at the expected cleavage sites. Empty vector refers to the *Drosophila* CRISPR-Cas9 vector (pAc-dU6-sgRNA-Cas9) without gRNA targeting any gene. In mock sample, the C6/36 cells were transfected with transfection reagent only (Lipofectamine 2000). In both mock and empty vector samples, the clones showed wild type sequence. PAM is highlighted in blue, while target sequence in yellow.
(TIFF)

Author Contributions

Conceptualization: Azali Azlan, Kumitaa Theva Das, Muhammad Amir Yunus, Ghows Azzam.

Data curation: Azali Azlan, Sattam M. Obeidat.

Formal analysis: Azali Azlan.

Funding acquisition: Ghows Azzam.

Investigation: Azali Azlan, Sattam M. Obeidat, Ghows Azzam.

Methodology: Azali Azlan, Kumitaa Theva Das, Muhammad Amir Yunus, Ghows Azzam.

Supervision: Ghows Azzam.

Validation: Azali Azlan.

Visualization: Azali Azlan.

Writing – original draft: Azali Azlan, Ghows Azzam.

Writing – review & editing: Azali Azlan, Kumitaa Theva Das, Muhammad Amir Yunus, Ghows Azzam.

References

1. Leroy EM, Nkoghe D, Ollomo B, Nze-Nkogue C, Becquart P, Grard G, et al. Concurrent Chikungunya and Dengue Virus Infections during Simultaneous Outbreaks, Gabon, 2007. Emerg Infect Dis. 2009; 15: 591–593. https://doi.org/10.3201/eid1504.080664 PMID: 19331740

2. Caron M, Paupy C, Grard G, Becquart P, Mombo I, Nso BBB, et al. Recent Introduction and Rapid Dissemination of Chikungunya Virus and Dengue Virus Serotype 2 Associated With Human and Mosquito

Coinfections in Gabon, Central Africa. Clin Infect Dis. 2012; 55: e45–e53. https://doi.org/10.1093/cid/cis530 PMID: 22670036

3. Paupy C, Delatte H, Bagny L, Corbel V, Fontenille D. Aedes albopictus, an arbovirus vector: From the darkness to the light. Microbes and Infection. 2009; 11: 1177–1185. https://doi.org/10.1016/j.micinf.2009.05.005 PMID: 19450706

4. Cancrini G, Frangipane di Regalbono A, Ricci I, Tessarin C, Gabrielli S, Pietrobelli M. Aedes albopictus is a natural vector of Dirofilaria immitis in Italy. Veterinary Parasitology. 2003; 118: 195–202. https://doi.org/10.1016/j.vetpar.2003.10.011 PMID: 14729167

5. Rai KS. Aedes Albopictus in the Americas. Annual Review of Entomology. 1991; 36: 459–484. https://doi.org/10.1146/annurev.en.36.010191.002331 PMID: 2006869

6. Bonizzoni M, Gasperi G, Chen X, James AA. The invasive mosquito species Aedes albopictus: current knowledge and future perspectives. Trends Parasitol. 2013; 29: 460–468. https://doi.org/10.1016/j.pt.2013.07.003 PMID: 23916878

7. Ab K. Global distribution and continuing spread of Aedes albopictus. Parassitologia. 1995; 37: 91–97. PMID: 8778670

8. Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, et al. RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. Science. 2007; 316: 1484–1488. https://doi.org/10.1126/science.1138341 PMID: 17510325

9. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. Genome Res. 2012; 22: 1775–1789. https://doi.org/10.1101/gr.132159.111 PMID: 22955988

10. Tripathi V, Shen Z, Chakraborty A, Giri S, Freier SM, Wu X, et al. Long noncoding RNA MALAT1 controls cell cycle progression by regulating the expression of oncogenic transcription factor B-MYB. PLoS Genet. 2013; 9: e1003368. https://doi.org/10.1371/journal.pgen.1003368 PMID: 23555285

11. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, et al. The Nuclear-Retained Noncoding RNA MALAT1 Regulates Alternative Splicing by Modulating SR Splicing Factor Phosphorylation. Molecular Cell. 2010; 39: 925–938. https://doi.org/10.1016/j.molcel.2010.08.011 PMID: 20797886

12. Wen K, Yang L, Xiong T, Di C, Ma D, Wu M, et al. Critical roles of long noncoding RNAs in Drosophila spermatogenesis. Genome Res. 2016; 26: 1233–1244. https://doi.org/10.1101/gr.199547.115 PMID: 27516619

13. McCorkindale AL, Wahle P, Werner S, Jungreis I, Menzel P, Shukla CJ, et al. A gene expression atlas of embryonic neurogenesis in Drosophila reveals complex spatiotemporal regulation of lncRNAs. Development. 2019; 146: dev175265. https://doi.org/10.1242/dev.175265 PMID: 30923056

14. Azlan A, Obeidat SM, Yunus MA, Azzam G. Systematic identification and characterization of Aedes aegypti long noncoding RNAs (lncRNAs). Scientific Reports. 2019; 9. https://doi.org/10.1038/s41598-019-47506-9 PMID: 31434910

15. Etebari K, Asad S, Zhang G, Asgari S. Identification of Aedes aegypti Long Intergenic Non-coding RNAs and Their Association with Wolbachia and Dengue Virus Infection. PLOS Neglected Tropical Diseases. 2016; 10: e0005069. https://doi.org/10.1371/journal.pntd.0005069 PMID: 27760142

16. Etebari K, Hegde S, Saldaña MA, Widen SG, Wood TG, Asgari S, et al. Global Transcriptome Analysis of Aedes aegypti Mosquitoes in Response to Zika Virus Infection. Fernandez-Sesma A, editor. mSphere. 2017; 2: e00456–17. https://doi.org/10.1128/mSphere.00456-17 PMID: 29202041

17. Gamez S, Antoshechkin I, Mendez-Sanchez SC, Akbari OS. The Developmental Transcriptome of Aedes albopictus, a Major Worldwide Human Disease Vector. G3: Genes, Genomes, Genetics. 2020; 10: 1051–1062. https://doi.org/10.1534/g3.119.401006 PMID: 31964684

18. Huang X, Poelchau MF, Armbruster PA. Global Transcriptional Dynamics of Diapause Induction in Non-Blood-Fed and Blood-Fed Aedes albopictus. PLOS Neglected Tropical Diseases. 2015; 9: e0003724. https://doi.org/10.1371/journal.pntd.0003724 PMID: 25897664

19. Poelchau MF, Reynolds JA, Elsik CG, Denlinger DL, Armbruster PA. RNA-Seq reveals early distinctions and late convergence of gene expression between diapause and quiescence in the Asian tiger mosquito, Aedes albopictus. Journal of Experimental Biology. 2013; 216: 4082–4090. https://doi.org/10.1242/jeb.089508 PMID: 23913949

20. Tsujimoto H, Hanley KA, Sundararajan A, Devitt NP, Schilkey FD, Hansen IA. Dengue virus serotype 2 infection alters midgut and carcass gene expression in the Asian tiger mosquito, Aedes albopictus. PLOS ONE. 2017; 12: e0171345. https://doi.org/10.1371/journal.pone.0171345 PMID: 28152011

21. Esquivel CJ, Cassone BJ, Piermarini PM. A de novo transcriptome of the Malpighian tubules in non-blood-fed and blood-fed Asian tiger mosquitoes Aedes albopictus: insights into diuresis, detoxification, and blood meal processing. PeerJ. 2016; 4: e1784. https://doi.org/10.7717/peerj.1784 PMID: 26989622

22. Xu Y, Dong Y, Xu Y, Lai Z, Jin B, Hao Y, et al. Differentiation of Long Non-Coding RNA and mRNA Expression Profiles in Male and Female Aedes albopictus. Front Genet. 2019;10. https://doi.org/10.3389/fgene.2019.00010 PMID: 30815010

23. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170 PMID: 24695404

24. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nature Methods. 2015; 12: 357–360. https://doi.org/10.1038/nmeth.3317 PMID: 25751142

25. Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nature Biotechnology. 2015; 33: 290–295. https://doi.org/10.1038/nbt.3122 PMID: 25690850

26. Wu Y, Cheng T, Liu C, Liu D, Zhang Q, Long R, et al. Systematic Identification and Characterization of Long Non-Coding RNAs in the Silkworm, Bombyx mori. PLOS ONE. 2016; 11: e0147147. https://doi.org/10.1371/journal.pone.0147147 PMID: 26771876

27. Rinn JL, Chang HY. Genome Regulation by Long Noncoding RNAs. Annual Review of Biochemistry. 2012; 81: 145–166. https://doi.org/10.1146/annurev-biochem-051410-092902 PMID: 22663078

28. Bakhtiarizadeh MR, Salami SA. Identification and Expression Analysis of Long Noncoding RNAs in Fat-Tail of Sheep Breeds. G3: Genes, Genomes, Genetics. 2019; 9: 1263–1276. https://doi.org/10.1534/g3.118.201014 PMID: 30787031

29. Pauli A, Valen E, Lin MF, Garber M, Vastenhouw NL, Levin JZ, et al. Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. Genome Res. 2012; 22: 577–591. https://doi.org/10.1101/gr.133009.111 PMID: 22110045

30. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. Nat Protoc. 2013; 8. https://doi.org/10.1038/nprot.2013.084 PMID: 23845962

31. Kang Y-J, Yang D-C, Kong L, Hou M, Meng Y-Q, Wei L, et al. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. Nucleic Acids Res. 2017; 45: W12–W16. https://doi.org/10.1093/nar/gkx428 PMID: 28521017

32. Li A, Zhang J, Zhou Z. PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. BMC Bioinformatics. 2014; 15: 311. https://doi.org/10.1186/1471-2105-15-311 PMID: 25239089

33. Wang L, Park HJ, Dasari S, Wang S, Kocher J-P, Li W. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res. 2013; 41: e74. https://doi.org/10.1093/nar/gkt006 PMID: 23335781

34. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014; 30: 923–930. https://doi.org/10.1093/bioinformatics/btt656 PMID: 24227677

35. Atieh T, Baronti C, de Lamballerie X, Nougairède A. Simple reverse genetics systems for Asian and African Zika viruses. Scientific Reports. 2016; 6: 1–7. https://doi.org/10.1038/s41598-016-0001-8 PMID: 28442746

36. Nougairede A, Fabritus LD, Aubry F, Gould EA, Holmes EC, Lamballerie X de. Random Codon Re-encoding Induces Stable Reduction of Replicative Fitness of Chikungunya Virus in Primate and Mosquito Cells. PLOS Pathogens. 2013; 9: e1003172. https://doi.org/10.1371/journal.ppat.1003172 PMID: 23436995

37. Li J, Hu D, Ding X, Chen Y, Pan Y, Qiu L, et al. Enzyme-Linked Immunosorbent Assay-Format Tissue Culture Infectious Dose-50 Test for Titrating Dengue Virus. PLOS ONE. 2011; 6: e22553. https://doi.org/10.1371/journal.pone.0022553 PMID: 21799894

38. Reed LJ, Muench H. A simple method of estimating fifty percent endpoints. Am J Epidemiol. 1938; 27: 493–497. https://doi.org/10.1093/oxfordjournals.aje.a118408

39. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010; 26: 139–140. https://doi.org/10.1093/bioinformatics/btp616 PMID: 19910308

40. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2$-\Delta\Delta$CT Method. Methods. 2001; 25: 402–408. https://doi.org/10.1006/meth.2001.1262 PMID: 11846609

41. Koskinen P, Törönen P, Nokso-Koivisto J, Holm L. PANNZER: high-throughput functional annotation of uncharacterized proteins in an error-prone environment. Bioinformatics. 2015; 31: 1544–1552. https://doi.org/10.1093/bioinformatics/btu851 PMID: 25653249

42. Reimand J, Kull M, Peterson H, Hansen J, Vilo J. g:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. Nucleic Acids Res. 2007; 35: W193–W200. https://doi.org/10.1093/nar/gkm226 PMID: 17478515

43. Labun K, Montague TG, Krause M, Torres Cleuren YN, Tjeldnes H, Valen E. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. Nucleic Acids Res. 2019; 47: W171–W174. https://doi.org/10.1093/nar/gkz365 PMID: 31106371

44. Bassett AR, Tibbit C, Ponting CP, Liu J-L. Mutagenesis and homologous recombination in Drosophila cell lines using CRISPR/Cas9. Biology Open. 2014; 3: 42–49. https://doi.org/10.1242/bio.20137120 PMID: 24326186

45. Varjak M, Maringer K, Watson M, Sreenu VB, Fredericks AC, Pondeville E, et al. Aedes aegypti Piwi4 Is a Noncanonical PIWI Protein Involved in Antiviral Responses. mSphere. 2017; 2: e00144–17. https://doi.org/10.1128/mSphere.00144-17 PMID: 28497119

46. Azlan A, Halim MA, Azzam G. Genome-wide identification and characterization of long intergenic noncoding RNAs in the regenerative flatworm Macrostomum lignano. Genomics. 2019 [cited 16 Jan 2020]. https://doi.org/10.1016/j.ygeno.2019.07.016 PMID: 31381967

47. Chen B, Zhang Y, Zhang X, Jia S, Chen S, Kang L. Genome-wide identification and developmental expression profiling of long noncoding RNAs during Drosophila metamorphosis. Scientific Reports. 2016; 6: 23330. https://doi.org/10.1038/srep23330 PMID: 26996731

48. Liu F, Guo D, Yuan Z, Chen C, Xiao H. Genome-wide identification of long non-coding RNA genes and their association with insecticide resistance and metamorphosis in diamondback moth, Plutella xylostella. Sci Rep. 2017;7. https://doi.org/10.1038/s41598-017-00035-9 PMID: 28127057

49. Nam J-W, Bartel DP. Long noncoding RNAs in C. elegans. Genome Res. 2012; 22: 2529–2540. https://doi.org/10.1101/gr.140475.112 PMID: 22707570

50. Wang M, Zhao W, Gao L, Zhao L. Genome-wide profiling of long non-coding RNAs from tomato and a comparison with mRNAs associated with the regulation of fruit ripening. BMC Plant Biology. 2018;18. https://doi.org/10.1186/s12870-018-1232-6 PMID: 29352810

51. Zhang J, Wei L, Jiang J, Mason AS, Li H, Cui C, et al. Genome-wide identification, putative functionality and interactions between lncRNAs and miRNAs in Brassica species. Scientific Reports. 2018; 8: 4960. https://doi.org/10.1038/s41598-018-23334-1 PMID: 29563515

52. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. 2011; 25: 1915–1927. https://doi.org/10.1101/gad.17446611 PMID: 21890647

53. Dzaki N, Ramli KN, Azlan A, Ishak IH, Azzam G. Evaluation of reference genes at different developmental stages for quantitative real-time PCR in Aedes aegypti. Sci Rep. 2017; 7: 43618. https://doi.org/10.1038/srep43618 PMID: 28300076

54. Dzaki N, Azzam G. Assessment of Aedes albopictus reference genes for quantitative PCR at different stages of development. PLOS ONE. 2018; 13: e0194664. https://doi.org/10.1371/journal.pone.0194664 PMID: 29554153

55. Wan Y, Liu X, Zheng D, Wang Y, Chen H, Zhao X, et al. Systematic identification of intergenic long-noncoding RNAs in mouse retinas using full-length isoform sequencing. BMC Genomics. 2019; 20: 559. https://doi.org/10.1186/s12864-019-5903-y PMID: 31286854

56. Jandura A, Krause HM. The New RNA World: Growing Evidence for Long Noncoding RNA Functionality. Trends in Genetics. 2017; 33: 665–676. https://doi.org/10.1016/j.tig.2017.08.002 PMID: 28870653

57. Vance KW, Ponting CP. Transcriptional regulatory functions of nuclear long noncoding RNAs. Trends Genet. 2014; 30: 348–355. https://doi.org/10.1016/j.tig.2014.06.001 PMID: 24974018

58. Angleró-Rodríguez YI, MacLeod HJ, Kang S, Carlson JS, Jupatanakul N, Dimopoulos G. Aedes aegypti Molecular Responses to Zika Virus: Modulation of Infection by the Toll and Jak/Stat Immune Pathways and Virus Host Factors. Front Microbiol. 2017;8. https://doi.org/10.3389/fmicb.2017.00008 PMID: 28144237

59. Sim S, Dimopoulos G. Dengue Virus Inhibits Immune Responses in Aedes aegypti Cells. PLOS ONE. 2010; 5: e10678. https://doi.org/10.1371/journal.pone.0010678 PMID: 20502529

60. Smartt CT, Shin D, Alto BW. Dengue serotype-specific immune response in Aedes aegypti and Aedes albopictus. Memórias do Instituto Oswaldo Cruz. 2017; 112: 829–837. https://doi.org/10.1590/0074-02760170182 PMID: 29211244

61. Krapivinsky G, Krapivinsky L, Stotz SC, Manasian Y, Clapham DE. POST, partner of stromal interaction molecule 1 (STIM1), targets STIM1 to multiple transporters. PNAS. 2011; 108: 19234–19239. https://doi.org/10.1073/pnas.1117231108 PMID: 22084111

62. Mills IG, Praefcke GJK, Vallis Y, Peter BJ, Olesen LE, Gallop JL, et al. EpsinR an AP1/clathrin interacting protein involved in vesicle trafficking. J Cell Biol. 2003; 160: 213–222. https://doi.org/10.1083/jcb.200208023 PMID: 12538641

63. Agrelli A, de Moura RR, Crovella S, Brandão LAC. Zika virus entry mechanisms in human cells. Infection, Genetics and Evolution. 2019; 69: 22–29. https://doi.org/10.1016/j.meegid.2019.01.018 PMID: 30658214

64. Shrivastava N, Sripada S, Kaur J, Shah PS, Cecilia D. Insights into the Internalization and Retrograde Trafficking of Dengue 2 Virus in BHK-21 Cells. PLOS ONE. 2011; 6: e25229. https://doi.org/10.1371/journal.pone.0025229 PMID: 21991304

65. Zhou Y, Frey TK, Yang JJ. Viral calciomics: Interplays between Ca2+ and virus. Cell Calcium. 2009; 46: 1–17. https://doi.org/10.1016/j.ceca.2009.05.005 PMID: 19535138

66. Chen X-G, Jiang X, Gu J, Xu M, Wu Y, Deng Y, et al. Genome sequence of the Asian Tiger mosquito, Aedes albopictus, reveals insights into its biology, genetics, and evolution. PNAS. 2015; 112: E5907–E5915. https://doi.org/10.1073/pnas.1516410112 PMID: 26483478

67. Palatini U, Masri RA, Cosme LV, Koren S, Thibaud-Nissen F, Biedler JK, et al. Improved reference genome of the arboviral vector Aedes albopictus. Genome Biology. 2020; 21: 215. https://doi.org/10.1186/s13059-020-02141-w PMID: 32847630

68. Miller JR, Koren S, Dilley KA, Puri V, Brown DM, Harkins DM, et al. Analysis of the Aedes albopictus C6/36 genome provides insight into cell line utility for viral propagation. Gigascience. 2018;7. https://doi.org/10.1093/gigascience/gix135 PMID: 29329394

69. Dritsou V, Topalis P, Windbichler N, Simoni A, Hall A, Lawson D, et al. A draft genome sequence of an invasive mosquito: an Italian Aedes albopictus. Pathogens and Global Health. 2015; 109: 207–220. https://doi.org/10.1179/2047773215Y.0000000031 PMID: 26369436

70. Etebari K, Furlong MJ, Asgari S. Genome wide discovery of long intergenic non-coding RNAs in Diamondback moth (Plutella xylostella) and their expression in insecticide resistant strains. Scientific Reports. 2015; 5: 1–14. https://doi.org/10.1038/srep14642 PMID: 26411386

71. Hezroni H, Koppstein D, Schwartz MG, Avrutin A, Bartel DP, Ulitsky I. Principles of Long Noncoding RNA Evolution Derived from Direct Comparison of Transcriptomes in 17 Species. Cell Reports. 2015; 11: 1110–1122. https://doi.org/10.1016/j.celrep.2015.04.023 PMID: 25959816

72. Tang Z, Wu Y, Yang Y, Yang Y-CT, Wang Z, Yuan J, et al. Comprehensive analysis of long non-coding RNAs highlights their spatio-temporal expression patterns and evolutionary conservation in Sus scrofa. Scientific Reports. 2017; 7: 43166. https://doi.org/10.1038/srep43166 PMID: 28233874

73. Bohlin J, Skjerve E, Ussery DW. Investigations of Oligonucleotide Usage Variance Within and Between Prokaryotes. PLOS Computational Biology. 2008; 4: e1000057. https://doi.org/10.1371/journal.pcbi.1000057 PMID: 18421372

74. Kopp F, Mendell JT. Functional Classification and Experimental Dissection of Long Noncoding RNAs. Cell. 2018; 172: 393–407. https://doi.org/10.1016/j.cell.2018.01.011 PMID: 29373828

75. Bonizzoni M, Dunn WA, Campbell CL, Olson KE, Marinotti O, James AA. Complex Modulation of the Aedes aegypti Transcriptome in Response to Dengue Virus Infection. PLOS ONE. 2012; 7: e50512. https://doi.org/10.1371/journal.pone.0050512 PMID: 23209765

76. Colpitts TM, Cox J, Vanlandingham DL, Feitosa FM, Cheng G, Kurscheid S, et al. Alterations in the Aedes aegypti Transcriptome during Infection with West Nile, Dengue and Yellow Fever Viruses. PLOS Pathogens. 2011; 7: e1002189. https://doi.org/10.1371/journal.ppat.1002189 PMID: 21909258

77. Vedururu R kiran, Neave MJ, Tachedjian M, Klein MJ, Gorry PR, Duchemin J-B, et al. RNASeq Analysis of Aedes albopictus Mosquito Midguts after Chikungunya Virus Infection. Viruses. 2019; 11. https://doi.org/10.3390/v11060513 PMID: 31167461

78. Zhao H, Chen M, Lind SB, Pettersson U. Distinct temporal changes in host cell lncRNA expression during the course of an adenovirus infection. Virology. 2016; 492: 242–250. https://doi.org/10.1016/j.virol.2016.02.017 PMID: 27003248

79. Scherbik SV, Brinton MA. Virus-Induced Ca2+ Influx Extends Survival of West Nile Virus-Infected Cells. Journal of Virology. 2010; 84: 8721–8731. https://doi.org/10.1128/JVI.00144-10 PMID: 20538858

80. Kalthoff C, Groos S, Kohl R, Mahrhold S, Ungewickell EJ. Clint: A Novel Clathrin-binding ENTH-Domain Protein at the Golgi. MBoC. 2002; 13: 4060–4073. https://doi.org/10.1091/mbc.e02-03-0171 PMID: 12429846

81. Acosta EG, Castilla V, Damonte EB. Functional entry of dengue virus into Aedes albopictus mosquito cells is dependent on clathrin-mediated endocytosis. J Gen Virol. 2008; 89: 474–484. https://doi.org/10.1099/vir.0.83357-0 PMID: 18198378

82. DeTulleo L, Kirchhausen T. The clathrin endocytic pathway in viral infection. EMBO J. 1998; 17: 4585–4593. https://doi.org/10.1093/emboj/17.16.4585 PMID: 9707418

83. Li M, Zhang D, Li C, Zheng Z, Fu M, Ni F, et al. Characterization of Zika Virus Endocytic Pathways in Human Glioblastoma Cells. Front Microbiol. 2020; 11. https://doi.org/10.3389/fmicb.2020.00242 PMID: 32210929