1 Genome-wide identification of *Aedes albopictus* long noncoding RNAs 2 and their association with dengue and zika virus infection

3
4 **Azali Azlan¹, Sattam M. Obeidat¹, Kunitaa Theva Das², Muhammad Amir Yunus² and Ghows
5 Azzam^{1*}**

6
7
8 ¹*School of Biological Sciences, Universiti Sains Malaysia, 11800 Penang, Malaysia*
9 ²*Infectomics Cluster, Advanced Medical & Dental Institute, Universiti Sains Malaysia, Bertam, 13200
10 Kepala Batas, Pulau Pinang, Malaysia.*

11
12
13
14
15
16
17 *Corresponding author
18 Email: ghows@usm.my (G.A)

19
20
21
22
23
24 **Key words: *Aedes albopictus*, long non-coding RNA, dengue, Zika**

25 Abstract

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

43

44

45

46

47 **Author summary**

48

49 *Ae. albopictus* is an important vector of arboviruses such as dengue and Zika. Studies on virus-
50 host interaction at gene expression and molecular level are crucial especially in devising methods to
51 inhibit virus replication in *Aedes* mosquito. Previous reports showed that, besides protein-coding genes,
52 noncoding RNAs such as lncRNAs are also involved in virus-host interaction. In this study, we report a
53 comprehensive catalog of novel lncRNA transcripts in the genome of *Ae. albopictus*. We also show that
54 the expression of lncRNAs was altered upon infection with dengue and Zika. Additionally, depletion of
55 certain lncRNAs resulted in increased replication of dengue and Zika; hence, suggesting potential
56 association of lncRNAs in virus infection. Results of this study provide a new avenue to the investigation
57 of mosquito-virus interactions that may potentially pave way to the development of novel methods in
58 vector control.

59

60 **Introduction**

61 The Asian tiger mosquito, *Aedes albopictus* (*Ae. albopictus*) is an important vector of arboviruses
62 such as dengue virus (DENV), Chikungunya virus (CHIKV), and Zika virus (ZIKV). Due to its
63 invasiveness and aggressive spread, *Ae. albopictus* has widespread geographic distribution; hence, poses
64 serious health threat globally. Although *Ae. albopictus* is considered as a less competent vector of DENV
65 than *Aedes aegypti* (*Ae. aegypti*), *Ae. albopictus* was responsible for dengue outbreaks in Hawaii, China,
66 and Europe, primarily because of its fast expansion across the globe [1]. Due to its epidemiological
67 importance, *Ae. albopictus* has become a key focus for many scientific investigations.

68 Long noncoding RNAs (lncRNAs) can be defined as RNA species of more than 200 nucleotides
69 (nt) that do not encode protein [2]. Similar to protein-coding mRNAs, lncRNAs are the products of Pol
70 II, and they undergo polyadenylation, capping and alternative splicing [3]. Due to their mRNA-like
71 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 [4], [5]. Besides, research in *Drosophila melanogaster* (*D. melanogaster*) provide evidence on the functional roles of lncRNAs in spermatogenesis [6] and neurogenesis [7].

Meanwhile, studies in *Ae. aegypti* demonstrated that lncRNAs might be involved in early embryonic development and virus infection [8], [9]. 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 [8]. Besides, RNAi-mediated silencing of *Ae. aegypti* lncRNA (lncRNA_1317) resulted in enhanced DENV replication in *Ae. aegypti* cells, suggesting the potential role of lncRNAs in virus-host interaction [9]. 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* [9], the expression level of *Ae. albopictus* lncRNAs was altered during virus infection. Furthermore, CRISPR-Cas9 induced mutation of selected lncRNAs in *Ae. albopictus* cell 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 [10], SRP050258 [11], SRP018112 [12], SRP077936 [13], SRP056407 [14], and SRP164784 [15]. All 148 RNA-seq libraries used in this study were listed in **S1 Data**. FASTQC was used to check the quality of raw reads. Adapters and low quality reads (<20 Phred score) were clipped using Trimmomatic version 0.38 [16].

Mapping and transcriptome assembly

Reads were individually mapped onto *Ae. albopictus* genome (Genome version: AalbF2, assembly: GCA_006496715.1, NCBI) using HISAT2 version 2.1.0 [17]. The resulting alignment files were used as input in Stringtie software (version 1.0.1) for transcriptome assembly [18]. 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”, “i”, and “x” were kept for lncRNA prediction.

Novel lncRNA prediction

TransDecoder [19] 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. The remaining transcripts were then subjected to coding potential assessment analysis. CPC2 [20], PLEK [21], and CPAT [22] were used to evaluate coding potential of the transcripts. For CPAT analysis, a coding probability threshold of < 0.3 was set [8], [9]. 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 TPM in at least 5 RNA-seq libraries were kept. The expression was determined by featureCounts [23], 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 CO₂. 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% CO₂. 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.

146

147 **Tissue Culture Infectious Dose 50 (TCID₅₀) assay**

148 TCID₅₀ assay was conducted to determine infectious virus titre using a method that has been previously
 149 described [24]–[26]. For each determination, a 96-well plate culture of confluent BHK-2 was inoculated
 150 with a 150 µl mixture of 100 µl medium and 50 µl of serial 10-fold dilutions of clarified cell culture
 151 supernatant media. Each row included 6 wells of the same dilution and two negative controls. Plates were
 152 incubated at at 37°C and 5% CO₂ for 5 days. Each well was observed for the presence or absence of
 153 cytopathic effect (CPE). The determination of TCID₅₀/ml was performed using the Reed and Muench
 154 method [27].

155

156 **Virus infection, RNA extraction and sequencing**

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

165

166 **Identification of differentially expressed lncRNAs upon infection**

167 RNA-seq libraries generated in this study were mapped to the *Ae. albopictus* genome (AalbF2, assembly:
 168 GCA_006496715.1, NCBI) using HISAT2 [17]. The expression was quantified using featureCounts [23].
 169 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 [28] in R/Bioconductor environment.

172

173 **Reverse transcription and real-time quantitative PCR (qPCR)**

174 cDNA synthesis was performed with one µg of total RNA in 20 µl reactions using the iScript Reverse
175 Transcription Supermix (Bio-Rad Laboratories, California; cat. no. 1708840). qPCR was performed
176 using iTaq™ Universal SYBR® Green Supermix (Bio-Rad Laboratories, California; cat. no. 1725120),
177 in a total of 20 µl reactions. All experiments were done in three biological and technical replicates. *RPS17*
178 and *RPL32* were used as housekeeping genes, and all qPCR runs were carried out on the BioRad CFX96
179 qPCR platform. The 2-ΔΔCT method [29] was used to estimate the relative expression of each lncRNA,
180 and the results were statistically analyzed using student's t-tests. List of primers used in this study can
181 be found in **S6 Table**.

182

183 **Functional annotation and enrichment**

184 Functional annotation was performed using Pannzer [30] using amino acid sequences of lncRNA-target
185 genes as input. Gene enrichment analysis was done using g:Profiler [31] and g:SCS threshold was used
186 for multiple testing correction.

187

188 **CRISPR-Cas9 induced mutation of selected lncRNAs**

189 sgRNA targeting exon 1 of lncRNA_27639.2 and XR_003899061.1 were designed using CHOPCHOP
190 [32]. The sgRNAs were cloned into the *Drosophila* CRISPR-Cas9 vector pAc-dU6-sgRNA-Cas9 [33],
191 which was a kind gift from Ji-Long Liu (Addgene plasmid 49330). This plasmid has been shown to work
192 well in *Ae. aegypti* (Aag2) cell [34]. Detailed protocol for sgRNA cloning was previously described [33].
193 C6/36 cells (~3x10⁶) 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 ~5x10⁵ 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 for twice, allowing the cells to break 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.

205

206

207 **Results**

208 **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 Data**. 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. lncRNA prediction pipeline used here were modified and adapted from previous works [8], [9], [35]–[41]. Our lncRNA prediction pipeline (**Fig 1**) began with mapping the raw reads against the *Ae. albopictus* genome (AalbF2, assembly: GCA_006496715.1, NCBI) using Hisat2 [17], followed by transcriptome assembly by Stringtie [18]. The resulting transcriptome assemblies were merged into a single assembly, which was then compared to the reference annotation (Aalbo_primary.1_genomic.gtf, NCBI) using Gffcompare (<https://github.com/gpertea/gffcompare>). We selected transcripts with class code “i”, “u”, and “x”

(35,863 transcripts), which indicate intronic, intergenic, and antisense to reference genes respectively [8], [40].

We then filtered the transcripts harboring open-reading frame (ORF) and coding potential. ORF prediction was performed using TransDecoder [19], while coding potential assessment was performed using three different algorithms, namely CPC2 [20], CPAT [22], and PLEK [21]. A total of 22,057 transcripts did not have 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 ($E\text{-value} < 10^3$). 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 **S1 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 [8], [35], [36], [40]. *Ae. albopictus* lncRNAs were found to be significantly shorter in size compared to protein-coding transcripts (Kolmogorov-Smirnov test (KS-test) $p\text{-value} < 2.2e-16$, **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 [8], [35], [42]. 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 [8], [9], [35], [36], [40]. 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 [8], [9], [43], we investigated the validity of this in *Ae. albopictus*. To assess the level of sequence similarity, we performed BLASTN of *Ae. albopictus* lncRNAs against lncRNA transcripts found in other organisms including *Ae. aegypti*, *D. melanogaster*, *M. musculus*, *C. elegans*, and humans. lncRNA sequences of *Ae. aegypti* were taken from previous work [8], while lncRNA sequences of the remaining species were downloaded from NONCODE, a database for ncRNAs, especially lncRNAs [44]. BLAST algorithm bit score was used as an indicator for sequence similarity [8], [9]. In general, protein-coding transcripts were found to be more conserved than lncRNAs (**Fig 3A**).

Ae. albopictus lncRNAs showed high degree of sequence similarity (E-value $< 10^{-5}$) with *Ae. aegypti*, indicating that they are likely genus specific (**Fig 3B**). However, the fraction of *Ae. albopictus* lncRNAs that displayed high level of sequence similarity (E-value $< 10^{-5}$) with *Ae. aegypti* were very low. For instance, only 819 (7.5%) novel lncRNAs and 617 (7.9%) known lncRNAs shared high sequence similarity with *Ae. aegypti*. On the other hand, 30,298 (75.6%) *Ae. albopictus* protein-coding transcripts demonstrated high degree of sequence similarity with *Ae. aegypti* protein-coding sequences.

267

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

269 Studies on *Ae. aegypti* transcriptomes provide evidence that lncRNAs could be involved in
270 DENV and ZIKV-mosquito interaction [9], [45]. To examine the relevance of this to *Ae. albopictus*, we
271 generated paired-end RNA-seq libraries from larval-derived *Ae. albopictus* cell line (C6/36) which had
272 been infected with DENV and ZIKV. A total of 9 samples were sequenced using Illumina HiSeq 2000
273 platform, and they comprised of three replicates of mock, DENV, and ZIKV-infected C6/36 cells. In
274 total, 426 million paired-end reads were generated in this study (**Table 1**). More than 70% of the reads
275 (average of 35.5 million reads per sample) mapped to *Ae. albopictus* reference genome (**Table 1**),
276 providing sufficient coverage for genome-wide transcriptomic analysis, specifically in detecting
277 differentially expressed transcripts. Hierarchical clustering based on the expression values (CPM) of all
278 the expressed genes showed that the replicates were grouped into their three distinct groups (**S1 Figure**).
279 This unsupervised clustering also showed that the data formed two main clusters, i.e., uninfected, and
280 virus-infected (DENV and ZIKV) samples.

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

283 In this study, we considered genes with fold change (FC) of more than or equal to 2 in either
284 direction, and false discovery rate (FDR) of less than 0.05 as differentially expressed. A total of 433 and
285 338 differentially expressed lncRNAs ($FC \geq |2|$, $FDR < 0.05$) were identified respectively in DENV and
286 ZIKV samples. In DENV-infected cells, 298 lncRNAs were found to be upregulated, while the remaining
287 136 transcripts were downregulated. Meanwhile, 216 and 122 lncRNAs were upregulated and
288 downregulated respectively in C6/36 cells infected with ZIKV. Volcano plot analysis was used to depict
289 the distribution of lncRNAs with their corresponding P-value and log FC (**Figure 4A**), and hierarchical
290 clustering analysis was performed to cluster the differentially expressed lncRNAs among different
291 replicates (**Figure 4B**). Similar lncRNAs were found to be differentially expressed - 11 lncRNAs were
292 upregulated and 12 lncRNAs were downregulated in both DENV and ZIKV infected cells (**Table 2**). The

293 similar changes seen in these 23 lncRNAs post-infection suggest that they may be key players, regardless
 294 of the type of viral infection. On the other hand, the number of differentially expressed protein-coding
 295 mRNAs were higher, i.e., 1,718 and 1,098 mRNAs were differentially expressed ($FC \geq |2|$, $FDR < 0.05$)
 296 in DENV and ZIKV-infected cells respectively. List of differentially expressed lncRNAs can be found
 297 in **S2 Table**.

298 We randomly selected ten differentially expressed lncRNAs from RNA-seq analysis for qPCR.
 299 The ten chosen candidates were upregulated or downregulated in both DENV and ZIKV-infected
 300 samples. We validated the differentially expressed lncRNAs using qPCR from the same RNA samples
 301 used for Illumina sequencing. *RPS17* and *RPL32* were used as housekeeping genes [46], [47]. Our
 302 preliminary tests showed that the expression level of the two housekeeping genes were unchanged in
 303 DENV, and ZIKV-infected C6/36 cells; hence, qualifying them as suitable references in this study.
 304 Overall, the qPCR results of all the ten lncRNAs were consistent with RNA-seq data; hence, validating
 305 the bioinformatic approach used in this study (**Fig 5**)

307 **Functional analysis of the differentially expressed lncRNAs**

308 To determine the putative regulatory roles of *Ae. albopictus* lncRNAs upon virus infections, we
 309 searched for protein-coding genes located within 100 kb (upstream and downstream) from differentially
 310 expressed lncRNAs as potential *cis*-regulated genes [48]–[50] and found a total of 198 and 153 coding
 311 genes which fulfilled that criteria in the DENV and ZIKV transcriptome respectively. As lncRNAs also
 312 have the ability to regulate genes in trans or distal regions [48], [51]–[53] we searched for protein-coding
 313 genes that were co-expressed with the differentially expressed lncRNAs based on Pearson correlation
 314 (coefficient > 0.95 or < -0.95 , $P\text{-value} < 0.05$). In DENV transcriptome, 110 coding genes were co-
 315 expressed with the differentially expressed lncRNAs. Meanwhile, we detected 38 coding genes having
 316 high correlation in expression with differentially expressed lncRNAs in ZIKV transcriptome. We
 317 categorized the co-expressed genes as putative trans-regulated genes in *Ae. albopictus* cells upon DENV

318 and ZIKV infection. Results of co-expression analysis between lncRNAs and protein-coding mRNAs
319 can be found in **S3 Table**.

320 We selected the potential *cis* and *trans*-regulated genes for functional annotation and enrichment
321 analyses. Previous reports demonstrated that upon virus infection in *Aedes* mosquitoes, there were
322 changes in the expression level of host immune-related genes [13], [45], [54]–[56]. By conducting
323 functional annotation using Pannzer [30] we discovered 13 lncRNAs that potentially regulated genes
324 involved in immune-related functions (**S4 Table**). Based on gene ontology (GO), these genes were
325 implicated in innate immune response (GO:0045087), defense response to Gram-positive (GO:0050830)
326 and Gram-negative (GO:0050829), defense response to bacterium (GO:0042742), regulation of defense
327 response to virus (GO:0050688), virus maturation (GO:0019075), and innate immune response
328 (GO:0045087).

329 Previous studies reported that genes encoding clip-domain serine protease, and metalloproteinase
330 were differentially expressed in *Aedes* mosquitoes upon virus infection. For example, clip-domain serine
331 protease and metalloproteases were found to be significantly upregulated and downregulated respectively
332 upon DENV infection in *Ae. albopictus* midguts [13]. Meanwhile, another study in *Ae. aegypti* reported
333 that metalloproteinase and serine protease showed an increase in expression upon infection with ZIKV
334 [45]. In this study, we discovered lncRNAs that potentially regulate genes that were also involved in
335 serine-type endopeptidase activity (GO:004252, GO:004253, GO:004254) and metalloendopeptidase
336 activity (GO:004222).

337 We then performed functional enrichment analysis of lncRNA-regulated genes using gProfiler
338 [31]. We discovered that 32 and 22 GO terms were significantly enriched (P-value < 0.05) in ZIKV and
339 DENV-transcriptome respectively (Supplementary information 8, Supplementary Figure 2). A total of 8
340 GO terms were significant and shared by both transcriptomes, suggesting that lncRNAs may target genes
341 that perform similar regulatory functions upon DENV and ZIKV infections (**S5 Table**). Among the
342 significant GO terms were cellular processes, ion/protein binding, metabolic processes, and DNA binding.

343

344 **Mutation of selected lncRNAs and enhancement of DENV and ZIKV replication**

345 To verify the role of *Ae. albopictus* lncRNAs in virus infection, we used CRISPR-Cas9 to induce
 346 mutation of selected lncRNAs in C6/36 cells following DENV and ZIKV infection. We discovered that
 347 CRISPR-Cas9 induced mutations in two lncRNAs loci resulted in the increase of DENV-1 and ZIKV
 348 infectious titer by at least two folds after three dpi (**Fig 6A**). One of the lncRNAs was novel (transcript
 349 ID: lncRNA_27639.2, loci ID: XLOC_029733) and another was known lncRNA (XR_003899061.1,
 350 ncbi gene ID: LOC115270134). Indels for both lncRNAs were confirmed by sequencing the PCR
 351 products spanning the cleavage sites, and the expression level lncRNAs after CRISPR-Cas9 induced
 352 mutation were validated by RT-qPCR (**S2 Figure**). Reduction in the level of expression indicated that
 353 mutations induced by CRISPR-Cas9 resulted in the silencing of the two aforementioned lncRNAs in
 354 C6/36 cells.

355 A previous report showed that the expression of *Ae. aegypti* lincRNA_1317 increased
 356 substantially following the progression of infection [9]. When this lncRNA was silenced by RNAi, it led
 357 to an enhancement of DENV replication in *Ae. aegypti* cell, implying the involvement of certain lncRNAs
 358 in viral infection. We next analyzed our lncRNAs to identify if they also played a key role in anti-viral
 359 response. Based on RNA-seq and RT-qPCR analysis, both lncRNAs (lncRNA_27639.2 and
 360 XR_003899061.1) were upregulated following infection with DENV and ZIKV at 3 dpi (**Fig 5**). We
 361 investigated the level of expression at early (1 dpi) and late (5 dpi) stage of DENV and ZIKV infection.
 362 Interestingly, the expression of both lncRNAs displayed an increasing pattern along the progression of
 363 infection (**Fig 6B**), indicating possible involvement of lncRNA_27639.2 and XR_003899061.1 in
 364 defense against viruses in *Ae. albopictus*.

365 lncRNA_27639.2 might potentially regulate a gene that encodes for solute carrier family 35
 366 member G1 (LOC109431891). This gene was located 1.9 kb downstream of lncRNA_27639.2 on the
 367 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 [57]. 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 [58]. Interestingly, calcium homeostasis, clathrin-mediated endocytosis, and vesicular trafficking are essential for viral replication in host cells [59]–[61]; thereby, further supporting the idea that lncRNA_27639.2 and XR_003899061.1 might be involved in antiviral response.

374

375 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* [9], [45]. 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* [15]. 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 [1]. Therefore, in this study, we used the latest genome version, AalbF2 assembly [62], which had reduced scaffold number, and produced a substantially complete gene set with lower duplications as compared to the previous genomes of *Ae. albopictus* [1], [63], [64].

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 [8], [9], [35], [42],

[43], [65], [66]. 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 [67].

Unlike protein-coding genes, low sequence conservation of lncRNAs makes functional prediction challenging. However, there are research that suggests that lncRNAs could regulate the expression of their neighboring protein-coding genes [49]; 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 [48], [49] in trans or in distal regions. 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 [13], [55], [68]. 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) (**S4 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 [45], [69], [70]. 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 [45], [69], [70].

417 Reports on the association between silencing of host lncRNAs and replication of pathogens are
418 still limited. One study reported that RNAi-mediated silencing of *Ae. aegypti* lincRNA_1317 led to
419 enhancement of DENV replication [9]. A similar observation was made in HIV-1-infected human T cells,
420 where the expression of lncRNA called NEAT1 was altered, and upon silencing of this lncRNA, the
421 replication of virus was observed to be increased [71]. The depletion of two lncRNAs by CRISPR-Cas9
422 resulted in increased DENV and ZIKV infection in our studies, further validated the involvement of
423 lncRNAs in virus infection. Therefore, this study provides strong evidence on the possible regulatory
424 role of lncRNAs in virus infection of many host cells including humans.

425 Depletion of novel lncRNA, lncRNA_27639.2 (loci id: XLOC_029733), and known lncRNA,
426 XR_003899061.1 (ncbi gene id: LOC115270134) in C6/36 cell led to enhancement of DENV-1 and
427 ZIKV by at least two folds. Functional annotation revealed that lncRNA_27639.2 has a potential to
428 regulate solute carrier family 35 member G1 (LOC109431891), whereas known lncRNA,
429 XR_003899061.1, may regulate clathrin interactor 1 protein (LOC115269840). Solute carrier family 35
430 member G1 has been implicated in intracellular calcium sensing and homeostasis [57], a process that has
431 been shown to be vital for virus replication [61]. For instance, infection of West Nile Virus (WNV)
432 triggers an influx of calcium ion, and this process seems to be required for initial stage of infection and
433 viral replication. This WNV-induced influx of calcium ion is probably necessary for rearrangement of
434 the ER membrane to accommodate virus replication [72].

435 Even though there is no report that specifically links solute carrier family 35 member G1 to virus
436 infection, its role in calcium homeostasis, which is important in virus replication, may implicate this gene
437 in virus-host interaction. On the other hand, clathrin interactor 1 protein is essential for trafficking via
438 clathrin-coated vesicles, and is localized at the trans-Golgi network [73]. Clathrin-mediated endocytosis
439 is required for virus to enter cells including *Ae. albopictus* cells [59], [74], [75]. For example, treatment
440 of C6/36 cells with chlorpromazine, an inhibitor of clathrin assembly, led to inhibition of DENV
441 replication and entry [74]. Meanwhile, inhibition of clathrin-mediated endocytosis via RNAi-mediated

442 silencing, or drug treatment (chlorpromazine and pitstop2) severely impaired ZIKV entry into human
443 glioblastoma cell T98G cell [76].

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

452

453

454

455 **Acknowledgements**

456 We would like to thank all our collaborators and colleagues for the discussion and the work conducted
457 in this lab. This study was funded by the Universiti Sains Malaysia Research University Grant
458 (1001/PBIOLOGI/811320 and 1001/PBIOLOGI/8011064) and ScienceFund Grant
459 (305/PBIOLOGI/613238).

460

461

462 **Figure Legend**

463 **Figure 1: Overview of lncRNA identification pipeline**

464 **Figure 2: Length and GC content *Ae. albopictus* lncRNAs**

465 A) Sequence length distribution of lncRNA and protein-coding transcript. B) GC content of lncRNA,
466 protein-coding transcript, 3'UTR, and 5'UTR

467 **Figure 3: Conservation analysis of *Ae. albopictus* lncRNAs**

468 A) Similarity bit score of of lncRNA and protein-coding mRNA. In general, protein-coding mRNA is
469 more conserved than lncRNA. B) The Venn diagram shows the number of *Ae. albopictus* lncRNAs that
470 display high conservation with other species according to BLASTN algorithm (E-value < 10-e5)

471 **Figure 4: lncRNA expression profiles of DENV and ZIKV-infected *Ae. albopictus* cells**

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

477 **Figure 5: RT-qPCR validation of the differentially expressed lncRNAs**

478 Expression levels were normalized to uninfected samples. Bars indicate mean +/- standard error of the
479 mean (SEM) of three independent experiments. Student's *t*-test was used to assess the statistical
480 significance. All have p-value of < 0.05, and those of p-value < 0.01 are denoted with **. RNA-seq fold
481 change was computed by edgeR.

482 **Figure 6: Possible involvement of lncRNAs in DENV and ZIKV replication in C6/36 cell**

483 A) Titer of ZIKV and DENV-1 after 3 dpi as determined by TCID50 assay. In mock sample, C6/36 cells
484 were transfected with transfection reagent only (Lipofectamine 2000). Empty vector refers to the
485 *Drosophila* CRISPR-Cas9 vector (pAc-dU6-sgRNA-Cas9) without sgRNA targeting any gene. B)
486 Expression level of lncRNA_27639.2 and XR_003899061.1 at 1, 3 and 5 dpi. At 1 dpi, the fold change
487 between infected and uninfected samples were statistically insignificant. For A) and B), bars indicate
488 mean +/- SEM of three independent experiments. Student's *t*-test was used to assess the statistical
489 significance. * p<0.05; **p<0.01.

490 **Table 1: Summary of RNA-seq data**

491 **Table 2: Differentially expressed lncRNAs that are common in DENV-1 and ZIKV transcriptome**

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

493 **S1 Data: List of 148 RNA-seq libraries used in this study**

494 **S1 Table: Genomic coordinates of *Ae. albopictus* novel lncRNAs**

495 **S2 Table: List of the differentially expressed lncRNAs**

496 **S3 Table: Results of co-expression analysis between lncRNAs and protein-coding genes**

497 **S4 Table: lncRNAs that potentially regulate immune-related genes**

498 **S5 Table: Significant GO terms found in DENV and ZIKV transcriptome**

499 **S6 Table: Primers used in this study**

500 **S1 Figure: Hierarchical clustering of RNA-seq samples**

501 Hierarchical clustering was done using the Euclidean distance method associated with complete linkage

502 **S2 Figure: Validation of CRISPR-Cas9 induced mutation in lncRNA loci**

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

514

515 **References**

- [1] X.-G. Chen *et al.*, “Genome sequence of the Asian Tiger mosquito, *Aedes albopictus*, reveals insights into its biology, genetics, and evolution,” *Proc. Natl. Acad. Sci.*, vol. 112, no. 44, pp. E5907–E5915, Nov. 2015, doi: 10.1073/pnas.1516410112.
- [2] P. Kapranov *et al.*, “RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription,” *Science*, vol. 316, no. 5830, pp. 1484–1488, Jun. 2007, doi: 10.1126/science.1138341.
- [3] T. Derrien *et al.*, “The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression,” *Genome Res.*, vol. 22, no. 9, pp. 1775–1789, Sep. 2012, doi: 10.1101/gr.132159.111.
- [4] V. Tripathi *et al.*, “Long noncoding RNA MALAT1 controls cell cycle progression by regulating the expression of oncogenic transcription factor B-MYB,” *PLoS Genet.*, vol. 9, no. 3, p. e1003368, Mar. 2013, doi: 10.1371/journal.pgen.1003368.
- [5] V. Tripathi *et al.*, “The Nuclear-Retained Noncoding RNA MALAT1 Regulates Alternative Splicing by Modulating SR Splicing Factor Phosphorylation,” *Mol. Cell*, vol. 39, no. 6, pp. 925–938, Sep. 2010, doi: 10.1016/j.molcel.2010.08.011.
- [6] K. Wen *et al.*, “Critical roles of long noncoding RNAs in *Drosophila* spermatogenesis,” *Genome Res.*, vol. 26, no. 9, pp. 1233–1244, Sep. 2016, doi: 10.1101/gr.199547.115.
- [7] A. L. McCorkindale *et al.*, “A gene expression atlas of embryonic neurogenesis in *Drosophila* reveals complex spatiotemporal regulation of lncRNAs,” *Development*, vol. 146, no. 6, p. dev175265, Mar. 2019, doi: 10.1242/dev.175265.
- [8] A. Azlan, S. M. Obeidat, M. A. Yunus, and G. Azzam, “Systematic identification and characterization of *Aedes aegypti* long noncoding RNAs (lncRNAs),” *Sci. Rep.*, vol. 9, no. 1, Dec. 2019, doi: 10.1038/s41598-019-47506-9.
- [9] K. Etebari, S. Asad, G. Zhang, and S. Asgari, “Identification of *Aedes aegypti* Long Intergenic Non-coding RNAs and Their Association with Wolbachia and Dengue Virus Infection,” *PLoS Negl. Trop. Dis.*, vol. 10, no. 10, p. e0005069, Oct. 2016, doi: 10.1371/journal.pntd.0005069.
- [10] S. Gamez, I. Antoshechkin, S. C. Mendez-Sanchez, and O. S. Akbari, “The Developmental Transcriptome of *Aedes albopictus*, a Major Worldwide Human Disease Vector,” *G3 Genes Genomes Genet.*, vol. 10, no. 3, pp. 1051–1062, Mar. 2020, doi: 10.1534/g3.119.401006.
- [11] X. Huang, M. F. Poelchau, and P. A. Armbruster, “Global Transcriptional Dynamics of Diapause Induction in Non-Blood-Fed and Blood-Fed *Aedes albopictus*,” *PLoS Negl. Trop. Dis.*, vol. 9, no. 4, p. e0003724, Apr. 2015, doi: 10.1371/journal.pntd.0003724.
- [12] M. F. Poelchau, J. A. Reynolds, C. G. Elsik, D. L. Denlinger, and P. A. Armbruster, “RNA-Seq reveals early distinctions and late convergence of gene expression between diapause and quiescence in the Asian tiger mosquito, *Aedes albopictus*,” *J. Exp. Biol.*, vol. 216, no. 21, pp. 4082–4090, Nov. 2013, doi: 10.1242/jeb.089508.
- [13] H. Tsujimoto, K. A. Hanley, A. Sundararajan, N. P. Devitt, F. D. Schilkey, and I. A. Hansen, “Dengue virus serotype 2 infection alters midgut and carcass gene expression in the Asian tiger

mosquito, *Aedes albopictus*,” *PLOS ONE*, vol. 12, no. 2, p. e0171345, Feb. 2017, doi: 10.1371/journal.pone.0171345.

- [14] C. J. Esquivel, B. J. Cassone, and P. M. Piermarini, “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*, vol. 4, p. e1784, Mar. 2016, doi: 10.7717/peerj.1784.
- [15] Y. Xu *et al.*, “Differentiation of Long Non-Coding RNA and mRNA Expression Profiles in Male and Female *Aedes albopictus*,” *Front. Genet.*, vol. 10, 2019, doi: 10.3389/fgene.2019.00975.
- [16] A. M. Bolger, M. Lohse, and B. Usadel, “Trimmomatic: a flexible trimmer for Illumina sequence data,” *Bioinforma. Oxf. Engl.*, vol. 30, no. 15, pp. 2114–2120, Aug. 2014, doi: 10.1093/bioinformatics/btu170.
- [17] D. Kim, B. Langmead, and S. L. Salzberg, “HISAT: a fast spliced aligner with low memory requirements,” *Nat. Methods*, vol. 12, no. 4, pp. 357–360, Apr. 2015, doi: 10.1038/nmeth.3317.
- [18] M. Perte, G. M. Perte, C. M. Antonescu, T.-C. Chang, J. T. Mendell, and S. L. Salzberg, “StringTie enables improved reconstruction of a transcriptome from RNA-seq reads,” *Nat. Biotechnol.*, vol. 33, no. 3, pp. 290–295, Mar. 2015, doi: 10.1038/nbt.3122.
- [19] B. J. Haas *et al.*, “De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity,” *Nat. Protoc.*, vol. 8, no. 8, Aug. 2013, doi: 10.1038/nprot.2013.084.
- [20] Y.-J. Kang *et al.*, “CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features,” *Nucleic Acids Res.*, vol. 45, no. W1, pp. W12–W16, Jul. 2017, doi: 10.1093/nar/gkx428.
- [21] A. Li, J. Zhang, and Z. Zhou, “PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme,” *BMC Bioinformatics*, vol. 15, no. 1, p. 311, Sep. 2014, doi: 10.1186/1471-2105-15-311.
- [22] L. Wang, H. J. Park, S. Dasari, S. Wang, J.-P. Kocher, and W. Li, “CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model,” *Nucleic Acids Res.*, vol. 41, no. 6, p. e74, Apr. 2013, doi: 10.1093/nar/gkt006.
- [23] Y. Liao, G. K. Smyth, and W. Shi, “featureCounts: an efficient general purpose program for assigning sequence reads to genomic features,” *Bioinformatics*, vol. 30, no. 7, pp. 923–930, Apr. 2014, doi: 10.1093/bioinformatics/btt656.
- [24] T. Atieh, C. Baronti, X. de Lamballerie, and A. Nougairède, “Simple reverse genetics systems for Asian and African Zika viruses,” *Sci. Rep.*, vol. 6, Dec. 2016, doi: 10.1038/srep39384.
- [25] J. Li *et al.*, “Enzyme-Linked Immunosorbent Assay-Format Tissue Culture Infectious Dose-50 Test for Titrating Dengue Virus,” *PLOS ONE*, vol. 6, no. 7, p. e22553, Jul. 2011, doi: 10.1371/journal.pone.0022553.

- [26] A. Nougairede, L. D. Fabritus, F. Aubry, E. A. Gould, E. C. Holmes, and X. de Lamballerie, "Random Codon Re-encoding Induces Stable Reduction of Replicative Fitness of Chikungunya Virus in Primate and Mosquito Cells," *PLOS Pathog.*, vol. 9, no. 2, p. e1003172, Feb. 2013, doi: 10.1371/journal.ppat.1003172.
- [27] L. J. Reed and H. Muench, "A SIMPLE METHOD OF ESTIMATING FIFTY PER CENT ENDPOINTS," *Am. J. Epidemiol.*, vol. 27, no. 3, pp. 493–497, May 1938, doi: 10.1093/oxfordjournals.aje.a118408.
- [28] M. D. Robinson, D. J. McCarthy, and G. K. Smyth, "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data," *Bioinformatics*, vol. 26, no. 1, pp. 139–140, Jan. 2010, doi: 10.1093/bioinformatics/btp616.
- [29] K. J. Livak and T. D. Schmittgen, "Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2- $\Delta\Delta$ CT Method," *Methods*, vol. 25, no. 4, pp. 402–408, Dec. 2001, doi: 10.1006/meth.2001.1262.
- [30] P. Koskinen, P. Törönen, J. Nokso-Koivisto, and L. Holm, "PANNZER: high-throughput functional annotation of uncharacterized proteins in an error-prone environment," *Bioinformatics*, vol. 31, no. 10, pp. 1544–1552, May 2015, doi: 10.1093/bioinformatics/btu851.
- [31] J. Reimand, M. Kull, H. Peterson, J. Hansen, and J. Vilo, "g:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments," *Nucleic Acids Res.*, vol. 35, no. Web Server issue, pp. W193–W200, Jul. 2007, doi: 10.1093/nar/gkm226.
- [32] K. Labun, T. G. Montague, M. Krause, Y. N. Torres Cleuren, H. Tjeldnes, and E. Valen, "CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing," *Nucleic Acids Res.*, vol. 47, no. W1, pp. W171–W174, Jul. 2019, doi: 10.1093/nar/gkz365.
- [33] A. R. Bassett, C. Tibbit, C. P. Ponting, and J.-L. Liu, "Mutagenesis and homologous recombination in *Drosophila* cell lines using CRISPR/Cas9," *Biol. Open*, vol. 3, no. 1, pp. 42–49, Jan. 2014, doi: 10.1242/bio.20137120.
- [34] M. Varjak *et al.*, "Aedes aegypti Piwi4 Is a Noncanonical PIWI Protein Involved in Antiviral Responses," *mSphere*, vol. 2, no. 3, Jun. 2017, doi: 10.1128/mSphere.00144-17.
- [35] A. Azlan, M. A. Halim, and G. Azzam, "Genome-wide identification and characterization of long intergenic noncoding RNAs in the regenerative flatworm *Macrostomum lignano*," *Genomics*, Aug. 2019, doi: 10.1016/j.ygeno.2019.07.016.
- [36] B. Chen, Y. Zhang, X. Zhang, S. Jia, S. Chen, and L. Kang, "Genome-wide identification and developmental expression profiling of long noncoding RNAs during *Drosophila* metamorphosis," *Sci. Rep.*, vol. 6, p. 23330, Mar. 2016, doi: 10.1038/srep23330.
- [37] F. Liu, D. Guo, Z. Yuan, C. Chen, and H. Xiao, "Genome-wide identification of long non-coding RNA genes and their association with insecticide resistance and metamorphosis in diamondback moth, *Plutella xylostella*," *Sci. Rep.*, vol. 7, Nov. 2017, doi: 10.1038/s41598-017-16057-2.
- [38] J.-W. Nam and D. P. Bartel, "Long noncoding RNAs in *C. elegans*," *Genome Res.*, vol. 22, no. 12, pp. 2529–2540, Dec. 2012, doi: 10.1101/gr.140475.112.

- [39] M. Wang, W. Zhao, L. Gao, and L. Zhao, “Genome-wide profiling of long non-coding RNAs from tomato and a comparison with mRNAs associated with the regulation of fruit ripening,” *BMC Plant Biol.*, vol. 18, no. 1, Dec. 2018, doi: 10.1186/s12870-018-1300-y.
- [40] Y. Wu *et al.*, “Systematic Identification and Characterization of Long Non-Coding RNAs in the Silkworm, *Bombyx mori*,” *PLOS ONE*, vol. 11, no. 1, p. e0147147, Jan. 2016, doi: 10.1371/journal.pone.0147147.
- [41] J. Zhang *et al.*, “Genome-wide identification, putative functionality and interactions between lncRNAs and miRNAs in Brassica species,” *Sci. Rep.*, vol. 8, no. 1, p. 4960, Mar. 2018, doi: 10.1038/s41598-018-23334-1.
- [42] M. N. Cabili *et al.*, “Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses,” *Genes Dev.*, vol. 25, no. 18, pp. 1915–1927, Sep. 2011, doi: 10.1101/gad.17446611.
- [43] K. Etebari, M. J. Furlong, and S. Asgari, “Genome wide discovery of long intergenic non-coding RNAs in Diamondback moth (*Plutella xylostella*) and their expression in insecticide resistant strains,” *Sci. Rep.*, vol. 5, no. 1, pp. 1–14, Sep. 2015, doi: 10.1038/srep14642.
- [44] Y. Zhao *et al.*, “NONCODE 2016: an informative and valuable data source of long non-coding RNAs,” *Nucleic Acids Res.*, vol. 44, no. D1, pp. D203–D208, Jan. 2016, doi: 10.1093/nar/gkv1252.
- [45] K. Etebari *et al.*, “Global Transcriptome Analysis of *Aedes aegypti* Mosquitoes in Response to Zika Virus Infection,” *mSphere*, vol. 2, no. 6, pp. e00456-17, Nov. 2017, doi: 10.1128/mSphere.00456-17.
- [46] N. Dzaki, K. N. Ramli, A. Azlan, I. H. Ishak, and G. Azzam, “Evaluation of reference genes at different developmental stages for quantitative real-time PCR in *Aedes aegypti*,” *Sci. Rep.*, vol. 7, p. 43618, 16 2017, doi: 10.1038/srep43618.
- [47] N. Dzaki and G. Azzam, “Assessment of *Aedes albopictus* reference genes for quantitative PCR at different stages of development,” *PLOS ONE*, vol. 13, no. 3, p. e0194664, Mar. 2018, doi: 10.1371/journal.pone.0194664.
- [48] M. R. Bakhtiarzadeh and S. A. Salami, “Identification and Expression Analysis of Long Noncoding RNAs in Fat-Tail of Sheep Breeds,” *G3 Genes Genomes Genet.*, vol. 9, no. 4, pp. 1263–1276, Apr. 2019, doi: 10.1534/g3.118.201014.
- [49] F. Kopp and J. T. Mendell, “Functional Classification and Experimental Dissection of Long Noncoding RNAs,” *Cell*, vol. 172, no. 3, pp. 393–407, 25 2018, doi: 10.1016/j.cell.2018.01.011.
- [50] Y. Wan *et al.*, “Systematic identification of intergenic long-noncoding RNAs in mouse retinas using full-length isoform sequencing,” *BMC Genomics*, vol. 20, no. 1, p. 559, Jul. 2019, doi: 10.1186/s12864-019-5903-y.
- [51] A. Jandura and H. M. Krause, “The New RNA World: Growing Evidence for Long Noncoding RNA Functionality,” *Trends Genet.*, vol. 33, no. 10, pp. 665–676, Oct. 2017, doi: 10.1016/j.tig.2017.08.002.

- [52] F. P. Marchese, I. Raimondi, and M. Huarte, “The multidimensional mechanisms of long noncoding RNA function,” *Genome Biol.*, vol. 18, no. 1, p. 206, Oct. 2017, doi: 10.1186/s13059-017-1348-2.
- [53] K. W. Vance and C. P. Ponting, “Transcriptional regulatory functions of nuclear long noncoding RNAs,” *Trends Genet.*, vol. 30, no. 8, pp. 348–355, Aug. 2014, doi: 10.1016/j.tig.2014.06.001.
- [54] Y. I. Angleró-Rodríguez, H. J. MacLeod, S. Kang, J. S. Carlson, N. Jupatanakul, and G. Dimopoulos, “Aedes aegypti Molecular Responses to Zika Virus: Modulation of Infection by the Toll and Jak/Stat Immune Pathways and Virus Host Factors,” *Front. Microbiol.*, vol. 8, 2017, doi: 10.3389/fmicb.2017.02050.
- [55] S. Sim and G. Dimopoulos, “Dengue Virus Inhibits Immune Responses in Aedes aegypti Cells,” *PLOS ONE*, vol. 5, no. 5, p. e10678, May 2010, doi: 10.1371/journal.pone.0010678.
- [56] C. T. Smartt, D. Shin, and B. W. Alto, “Dengue serotype-specific immune response in Aedes aegypti and Aedes albopictus,” *Mem. Inst. Oswaldo Cruz*, vol. 112, no. 12, pp. 829–837, Dec. 2017, doi: 10.1590/0074-02760170182.
- [57] G. Krapivinsky, L. Krapivinsky, S. C. Stotz, Y. Manasian, and D. E. Clapham, “POST, partner of stromal interaction molecule 1 (STIM1), targets STIM1 to multiple transporters,” *Proc. Natl. Acad. Sci.*, vol. 108, no. 48, pp. 19234–19239, Nov. 2011, doi: 10.1073/pnas.1117231108.
- [58] I. G. Mills *et al.*, “EpsinR an AP1/clathrin interacting protein involved in vesicle trafficking,” *J. Cell Biol.*, vol. 160, no. 2, pp. 213–222, Jan. 2003, doi: 10.1083/jcb.200208023.
- [59] A. Agrelli, R. R. de Moura, S. Crovella, and L. A. C. Brandão, “ZIKA virus entry mechanisms in human cells,” *Infect. Genet. Evol.*, vol. 69, pp. 22–29, Apr. 2019, doi: 10.1016/j.meegid.2019.01.018.
- [60] N. Shrivastava, S. Sripada, J. Kaur, P. S. Shah, and D. Cecilia, “Insights into the Internalization and Retrograde Trafficking of Dengue 2 Virus in BHK-21 Cells,” *PLOS ONE*, vol. 6, no. 10, p. e25229, Oct. 2011, doi: 10.1371/journal.pone.0025229.
- [61] Y. Zhou, T. K. Frey, and J. J. Yang, “Viral calciomics: Interplays between Ca²⁺ and virus,” *Cell Calcium*, vol. 46, no. 1, pp. 1–17, Jul. 2009, doi: 10.1016/j.ceca.2009.05.005.
- [62] U. Palatini *et al.*, “Improved reference genome of the arboviral vector Aedes albopictus,” *bioRxiv*, p. 2020.02.28.969527, Feb. 2020, doi: 10.1101/2020.02.28.969527.
- [63] V. Dritsou *et al.*, “A draft genome sequence of an invasive mosquito: an Italian *Aedes albopictus*,” *Pathog. Glob. Health*, vol. 109, no. 5, pp. 207–220, Jul. 2015, doi: 10.1179/2047773215Y.00000000031.
- [64] J. R. Miller *et al.*, “Analysis of the Aedes albopictus C6/36 genome provides insight into cell line utility for viral propagation,” *GigaScience*, vol. 7, no. 3, Mar. 2018, doi: 10.1093/gigascience/gix135.

- [65] H. Hezroni, D. Koppstein, M. G. Schwartz, A. Avrutin, D. P. Bartel, and I. Ulitsky, “Principles of Long Noncoding RNA Evolution Derived from Direct Comparison of Transcriptomes in 17 Species,” *Cell Rep.*, vol. 11, no. 7, pp. 1110–1122, May 2015, doi: 10.1016/j.celrep.2015.04.023.
- [66] Z. Tang *et al.*, “Comprehensive analysis of long non-coding RNAs highlights their spatio-temporal expression patterns and evolutionary conservation in *Sus scrofa*,” *Sci. Rep.*, vol. 7, p. 43166, Feb. 2017, doi: 10.1038/srep43166.
- [67] J. Bohlin, E. Skjerve, and D. W. Ussery, “Investigations of Oligonucleotide Usage Variance Within and Between Prokaryotes,” *PLOS Comput. Biol.*, vol. 4, no. 4, p. e1000057, Apr. 2008, doi: 10.1371/journal.pcbi.1000057.
- [68] M. Bonizzoni, W. A. Dunn, C. L. Campbell, K. E. Olson, O. Marinotti, and A. A. James, “Complex Modulation of the *Aedes aegypti* Transcriptome in Response to Dengue Virus Infection,” *PLOS ONE*, vol. 7, no. 11, p. e50512, Nov. 2012, doi: 10.1371/journal.pone.0050512.
- [69] T. M. Colpitts *et al.*, “Alterations in the *Aedes aegypti* Transcriptome during Infection with West Nile, Dengue and Yellow Fever Viruses,” *PLOS Pathog.*, vol. 7, no. 9, p. e1002189, Sep. 2011, doi: 10.1371/journal.ppat.1002189.
- [70] R. kiran Vedururu *et al.*, “RNASeq Analysis of *Aedes albopictus* Mosquito Midguts after Chikungunya Virus Infection,” *Viruses*, vol. 11, no. 6, Jun. 2019, doi: 10.3390/v11060513.
- [71] H. Zhao, M. Chen, S. B. Lind, and U. Pettersson, “Distinct temporal changes in host cell lncRNA expression during the course of an adenovirus infection,” *Virology*, vol. 492, pp. 242–250, May 2016, doi: 10.1016/j.virol.2016.02.017.
- [72] S. V. Scherbik and M. A. Brinton, “Virus-Induced Ca²⁺ Influx Extends Survival of West Nile Virus-Infected Cells,” *J. Virol.*, vol. 84, no. 17, pp. 8721–8731, Sep. 2010, doi: 10.1128/JVI.00144-10.
- [73] C. Kalthoff, S. Groos, R. Kohl, S. Mahrhold, and E. J. Ungewickell, “Clint: A Novel Clathrin-binding ENTH-Domain Protein at the Golgi,” *Mol. Biol. Cell*, vol. 13, no. 11, pp. 4060–4073, Sep. 2002, doi: 10.1091/mbc.e02-03-0171.
- [74] E. G. Acosta, V. Castilla, and E. B. Damonte, “Functional entry of dengue virus into *Aedes albopictus* mosquito cells is dependent on clathrin-mediated endocytosis,” *J. Gen. Virol.*, vol. 89, no. Pt 2, pp. 474–484, Feb. 2008, doi: 10.1099/vir.0.83357-0.
- [75] L. DeTulleo and T. Kirchhausen, “The clathrin endocytic pathway in viral infection,” *EMBO J.*, vol. 17, no. 16, pp. 4585–4593, Aug. 1998, doi: 10.1093/emboj/17.16.4585.
- [76] M. Li *et al.*, “Characterization of Zika Virus Endocytic Pathways in Human Glioblastoma Cells,” *Front. Microbiol.*, vol. 11, 2020, doi: 10.3389/fmicb.2020.00242.

Figure 1

A

B

Figure 2

A

B

Figure 3

A

B

Figure 4

A

B

Figure 5

Figure 6