Transcriptional and Translational Dynamics of Zika and Dengue Virus Infection

Md Shazid Hasan, Dylan Harris, Jahanara Sonia, Weixia Deng

Background and Introduction:

Translation, the key event in all living things in which the genetic code is decoded as proteins are synthesized, has been underappreciated as a regulatory stage of gene expression. Historically, gene expression monitoring has focused on the levels of mRNA (the transcriptome) determined by methods that continue to be more informative, such as the widely used high throughput RNA-sequencing (RNA-seq) that alone is not sufficient to determine protein expression level. Proteomics study can address this questions regarding protein expression, but measuring protein levels combined with RNA-seq provides only an indirect measure of translation efficiency and does not account for differences in protein stability. To study and understand translation regulation, ribosome profiling is one of the powerful tools. Ribosome Profiling or RiboSeq is a RNAseq-based powerful deep sequencing technique that sequences the RNA occupied by translating ribosome instead of sequencing the total mRNA. This can provide a panaroma of whole translation process.

In this study, the authors have used ribosome profiling to understand translation regulation in virus infection. Zika virus (ZIKV) and dengue virus (DENV) belongs to the family Flaviviridae and contain single-stranded plus-sense RNA genomes. Both viruses are transmitted to humans by tropical Aedes mosquitos that are increasingly reaching first world regions. ZIKV and DENV share over 90% of their genome sequences. Also, previous studies on the transcriptional and immunological effects of DENV and ZIKV have revealed that both viruses induce a classical Type I interferon anti-viral response. In this study, they have done high-resolution ribosome profiling and RNA deep sequencing (RNA-seq) to define the gene expression and mRNA translation dynamics of the viral and host genomes during ZIKV and DENV infection of human neuronal progenitor cells (hNPCs). Their data highlights the cellular stress response and the activation of RNA translation and polyamine metabolism during DENV and ZIKV infection.

In this project, we tried to reproduce and understand the differentially expressed gene analysis from this study, both in transcriptional and translational level.

Overall design

They performed ribosome footprinting and RNA sequencing on human neuronal progenitor cells (hNPCs) cells either uninfected or infected with ZIKV or DENV in two biological replicates that were collected 72 hours post infection.

Source of the raw data

Data Link

In the source, unifected samples are indicated as A1 and B1, ZIKV infected samples are indicated as A2, B2 and DENV infected samples are indicated as A3 and B3. Then the raw sequences were processed.

Technical Details:

We were able to replicate most figures that did not require a reference genome using `ggplot2` 3.4.2 in `R` version 4.3.0.

Difficulties

- Rerunning analysis leads to differing adjusted p values. Different statistical software was used by the Singh et al. meaning a different algorithm is used to correct false discovery rate.
- Some rounding was used by the paper resulting in slight differences in figures even when their p adjusted was used
- A large portion of the figures in this paper are generated through Ensembl which compares the RNA sequences to the human genome giving them labels
- No documentation makes this difficult

Figure 1 B and C

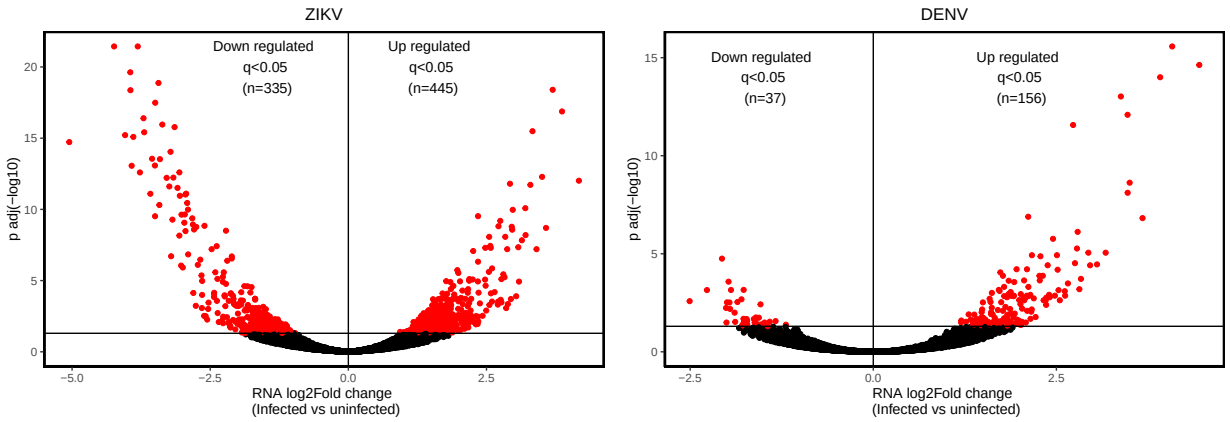

Figure 1B & C: To reproduce Figure 1B and C we used log2fold change in RNA expression (uninfected vs infected) in the X axis and p(adjusted value) in Y axis. We found upregulation of 445 mRNAs and downregulation of 335 mRNAs in case of ZIKV infection. Also, we found upregulation of 156 mRNAs and downregulation of 37mRNAs in DENV infection, which is the same as the study.

The x-axis is `log2FoldChange` and y-axis is `padj` from `GSE207347_ZIKV_DESeq2_result_name.txt.gz` and `GSE207347_DENV_DESeq2_result_name.txt.gz` files.

Figure 2 A and B

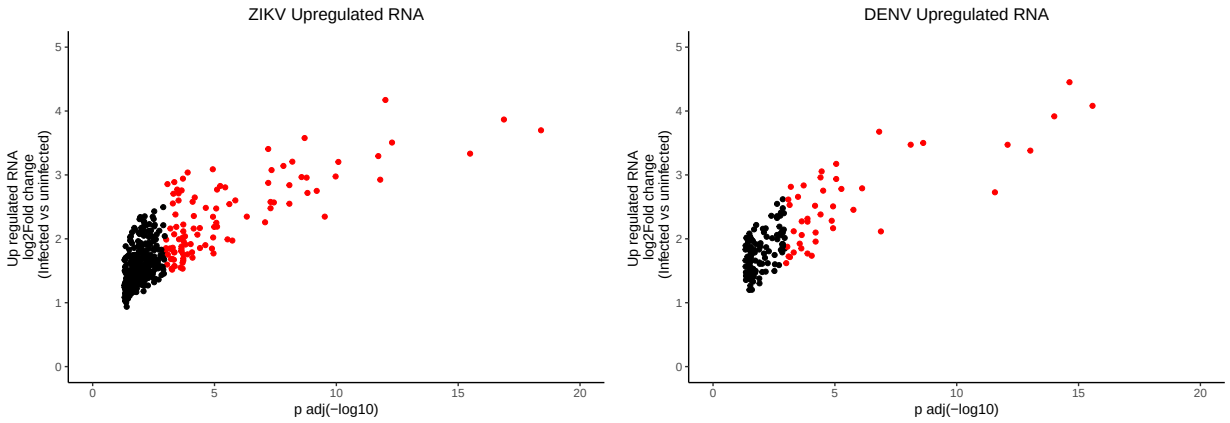

Figure 2A & 2B: Here, we tried to analyze the upregulated RNA in ZIKV(fig. 2A) and DENV (fig. 2B) infected samples compared to the uninfected samples, using DSeq2 processed data. Although, in their plots they highlighted the genes that are upregulated both in ZIKV and DENV infection, our plot resembles their plot showing 104 (ZIKV infection) and 44 (DENV infection) upregulated RNAs.

The x-axis is `padj` and y-axis is `log2FoldChange` > 0 from `GSE207347_ZIKV_DESeq2_result_name.txt.gz` and `GSE207347_DENV_DESeq2_result_name.txt.gz` files.

Figure 3 A and B

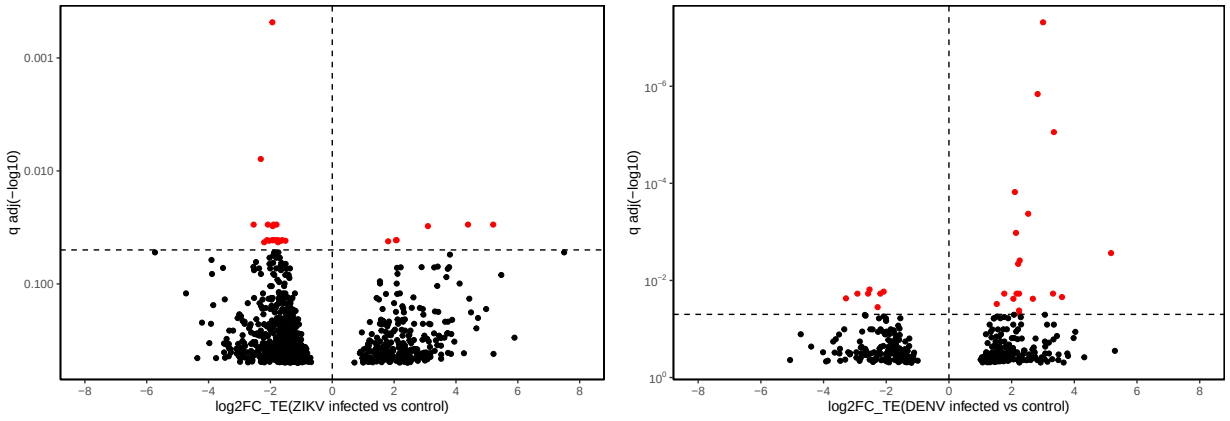

Figure 3A and 3B: These figures indicate translational changes induced by ZIKV and DENV infection. Here, to reproduce the data we have used the ribodiff dataset where we have used log2 fold changes in Translational Efficiency in X axis. We also had to cut off all the p values greater than 0.05. For Y axis, q adusted value is used where we had to use reverse log10 scale to model the graph the same as the study's.

The x-axis is `log2FC_TE_DrugTreated_vs_Control` and y-axis is `padj` < 0.5.

from `GSE207347_A1B1_vs_A2B2_ZIKV_ribodiff_name.txt.gz`

and `GSE207347_A1B1_vs_A2B2_DENV_ribodiff_name.txt.gz` files.

Figure 2 E

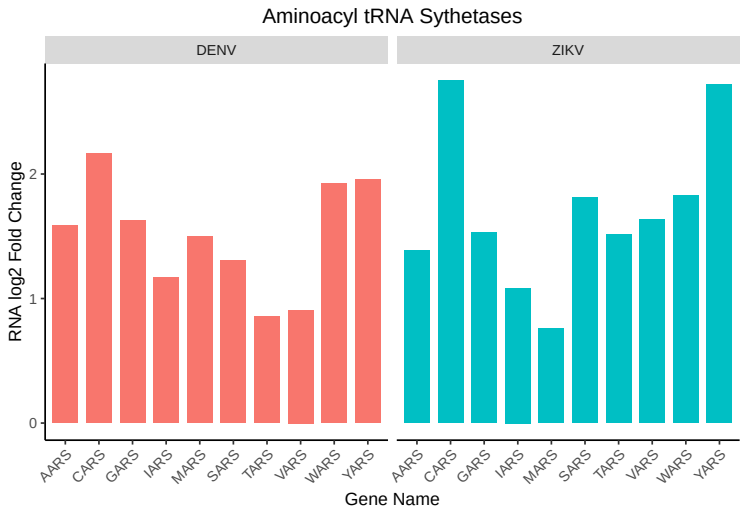

Figure 2E: Here, we analyzed RNA fold change of ZIKV and DENV induced aminoacyl tRNA synthetases (ARS) genes. Interestingly, here, our plot showed difference compared to the actual study. For example, TARS genes showed non-significant decrease in RNA abundance in their study, while ours still shows an increase which is also indicated by their data tables.

The x-axis is all the aminoacyl tRNA synthetases and the y-axis is the `log2FoldChange` in RNA from `GSE207347_ZIKV_DESeq2_result_name.txt.gz` and `GSE207347_DENV_DESeq2_result_name.txt.gz` files.

Test (Figure 1 F)

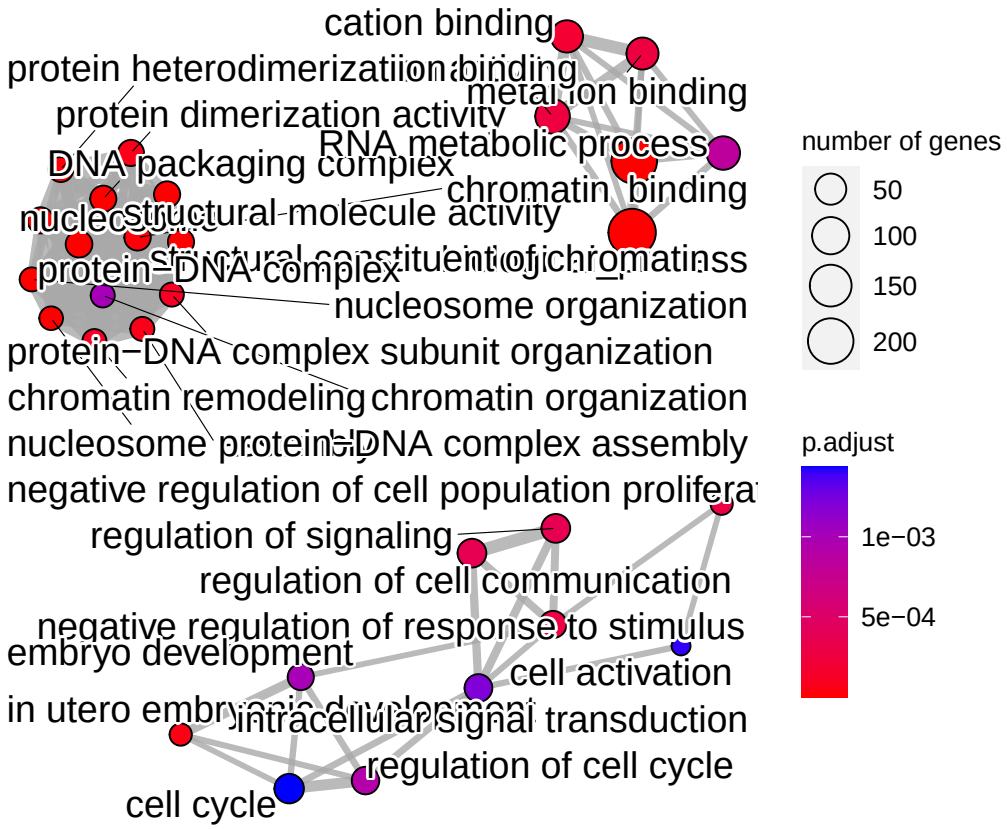

Figure 1F: This figure is the most interesting part of our analysis. Here we tried to reproduce their SRING analysis of gene clusters to understand interaction between downregulated genes. Although we ended up producing a very different figure, the clusters we got were similar to the study's. Their study showed clusters of genes associated with cell cycle and histone regulation, where the clusters we found also showed clusters associated with cell cycle and chromatin organization.

By following an online tutorial **Link**, we use `clusterProfiler` ver. 4.8.0 package, the genome wide annotation for human `org.Hs.eg.db` and `enrichplot` ver. 1.20.0 available on Bioconductor ver. 3.17 to perform the gene set enrichment analysis and plot the enrichment map.