BIOL 425 Final Report: The dengue virus and its ability to adapt to different host species.

Tiana Joseph

5-21-2023

Table of Contents

Introduction

Viruses are continually developing and adapting to their host, they evade their defense mechanism and diminish their immune system. They then take over the host's own cell processes for their own advantage. They make modifications to their genetic makeup, and

their ability to adapt to new hosts. Though viruses can switch between species they face extraordinary stressors since they have to retain their ability to infect different types of organisms.

Background

The Dengue virus is a viral infection that spreads from mosquitoes to people. The virus infects about 390 million and kills about 10,000 people annually. The dengue virus is a single strand of RNA, and its genome is mutating and evolving rapidly. Sequencing allows for the detection and quantification of RNA for a genetic diagnosis. Today high accuracy sequencing methods like PrimerID or Circular Sequencing (CirSeq) allow for the control of sequencing errors through barcoding, template circulation, and amplification. CirSeq specifically control errors in sequencing by outlining individual allele in viruses, demonstrating its ability to adapt to different species. It also detects the mutations and deletion of alleles and helps identify their role in the virus's adaptive ability. This experiment explores how the Dengue virus switches and adapts to different host species.

Biological hypothesis

What is the significance between the dengue virus's genetic makeup and its ability to adapt to different host species?

Null Hypothesis: (H0): No effect equal μ, there is no difference between the virus's genetic makeup and its ability to adapt to a different host.

Significance

This is significant because studying the genetic makeup of the dengue virus helps understand the ways it adapts to new host species. Also, by identifying and targeting distinct mutations, scientists can explore ways to block or inhibit those mutations. This could potentially restrict the dengue virus's capability to adapt in humans. Essentially leading to a decrease in its impact on public health, through the creation of strategies and policies to control the spread.

Materials and Methods

Focus-forming assay is used to measure the ability of the virus as it infects, replicates in host cells, and quantifies viral infection based on foci formation.

Quantitative real-time PCR (qRT-PCR) was used to measure the number of RNA molecules in a sample and quantify gene expression levels.

Samples

The genome of the Human hepatoma-derived cell lines from Huh7 and Aedes albopictus derived from C6/36 Asian tiger mosquito cells. Sample sizes were about 5x10^6 or 5000000 cells. For the experiment, the positive controls were the human hepatoma cell and the Aedes albopictus cell line. The Negative controls were the Huh7.5.1 cells, human hepatoma-derived HepG2 cells, and African Green Monkey epithelial-derived Vero cell line.

Experimental procedure

Illumina's CirSeq package was used to sequence the circular, viral RNA genome and identify rare and low-frequency genetic variants of the RNA virus.

CirSeq v2 package was used to determine variant base calls and allele frequencies.

R/RStudio/R packages were used to do the calculation and create the graph.

Schrödinger's pyMol2 molecular was used to visualize the system for visual of the fitness values for the non-synonymous mutations on available dengue pdb structures.

TMpred program was used to identify transmembrane regions and their orientation.

Statistical methods

The statistical analysis was performed using the R/RStudio platform.

A t-test was used to determine the significant difference in the RNA content of the tested host and the adapted host.

The scatterplot was used to determine any trends between the RNA content and the passages of the Human cell line and Mosquito cell line.

Exploratory data analysis content

```
knitr::opts_chunk$set(echo = TRUE)

# loads the library
library(tidyverse)

library(reshape2)

library(rstatix)

library(ggpubr)
library(dplyr)

#reads the files
fig1 <- read_tsv("group_5_dengue_genomics/elife-61921-fig1-data1-v3.txt",
show_col_types = FALSE)

#filters set A then save in variable
Set_A = fig1 %>% filter(set == "A" )

#creating graph from the data in Set A
# plot including meaning and confidence intervals

# x-axis is the passage and  y-axis id the  mean1.
#The mean was colored by the adaptedhost.
#Dots,a line,polygon and jitters  was added to show the mean,SEM and the CI.
#the plot was scaled to show the 9 passage and then separated based on the
tested host
fig1_Set_A = ggplot(Set_A, aes(x = passage, y = mean1, color = AdaptedHost))
+ geom_line() + geom_point() + geom_polygon(stat = "summary", fun.data =
mean_cl_boot, aes(fill = AdaptedHost), alpha=0.1, color = NA, show.legend =
FALSE)  + geom_jitter() + facet_wrap(~TestedHost) + scale_x_continuous(breaks
= 1:9)

# Titles were added to the graph, axis, and the font and  size was adjusted
fig1_Set_A + scale_color_discrete(name="Mean and SEM") + theme_bw() +
```

```
labs(title ="Assayed Cell Line", y="Cellular RNA content (AUs)", x="Passage")
+ theme(text = element_text(face="bold", size=12), axis.text =
element_text(face="bold", size=12), axis.title = element_text(face="bold",
size=12), legend.text=element_text(face="bold", size=12),
legend.title=element_text(face="bold", size=12), legend.position="top")

#Test for the tested host
t.test(data=Set_A, mean1 ~ TestedHost)

#t-test for the adapted host
t.test(data=Set_A, mean1 ~ AdaptedHost)
```

Results

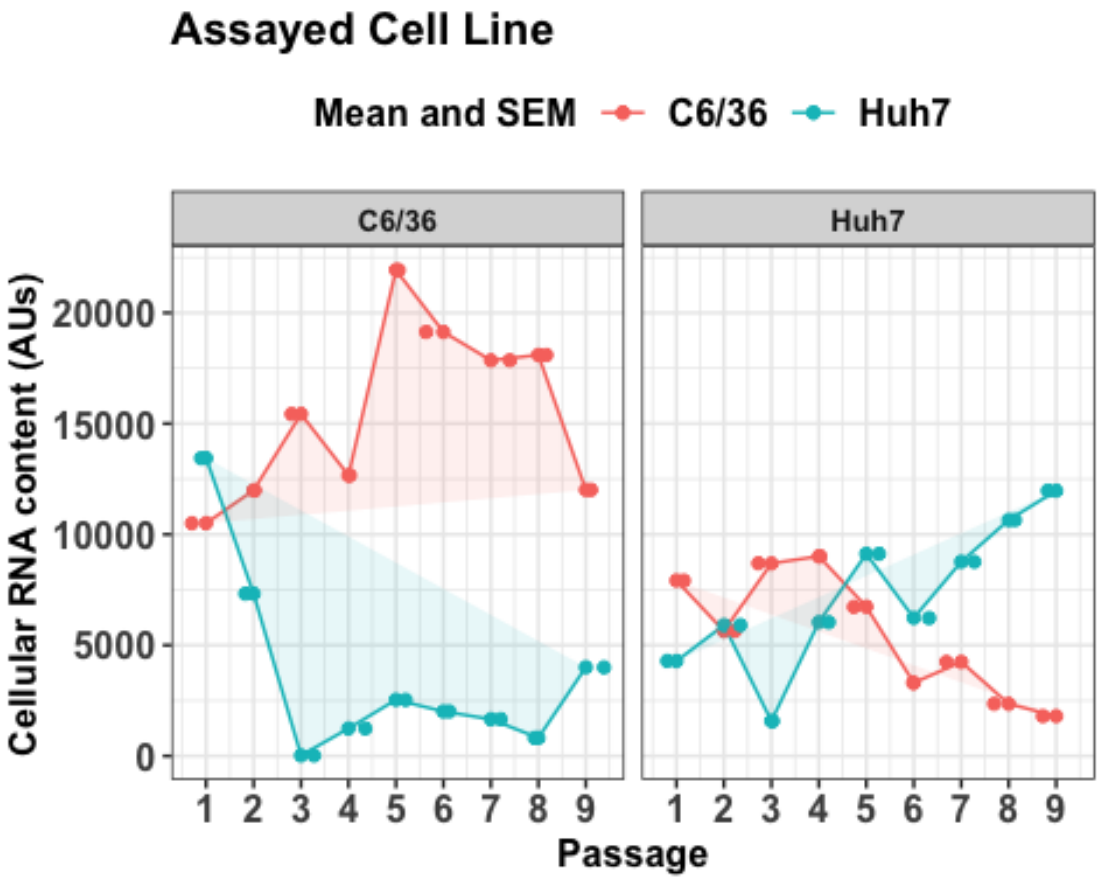

Figure 1: This graph is an analysis of the dengue virus RNA content and the passages of the Human and Mosquito cell line.

The x-axis is the passages of the different cell lines, and the y-axis is the mean of the cellular RNA content. The lines and dots represent the mean and the shaded area represents the SEM replication of all RNA content obtained in the Adapted host for each passage then separated based on the Tested host. The graph shows the fitness gains and, as the fitness measures increase over time, it shows a significant adaptive ability, however when cultivated by the alternative host there is a loss of adaptive ability.

```
        Welch Two Sample t-test

data:  mean1 by TestedHost
t = 1.7472, df = 22.76, p-value = 0.09409
alternative hypothesis: true difference in means between group C6/36 and group
Huh7 is not equal to 0
95 percent confidence interval:
 -599.7048 7093.6717
sample estimates:
mean in group C6/36  mean in group Huh7
           9594.835            6347.852
```

The t-test shows that the p-value of the Tested Host is equal to 0.09409, which is more than 0.05. Therefore, there is no significant difference between the means of the C6/36 cell line and the Huh7 cell line, The test also shows the averages of the mean, and the C6/36 cell line has the highest average of 6347.852.

```
        Welch Two Sample t-test

data:  mean1 by AdaptedHost
t = 2.9422, df = 29.703, p-value = 0.006264
alternative hypothesis: true difference in means between group C6/36 and group
Huh7 is not equal to 0
95 percent confidence interval:
 1557.253 8634.873
sample estimates:
mean in group C6/36  mean in group Huh7
          10519.375            5423.312
```

The t-test shows that the p-value of the Adapted Host is equal to 0.006264, which is less than 0.05. Therefore, there is a significant difference between the means of the C6/36 cell line and the Huh7 cell line. The t-test also shows the averages of the mean and the C6/36 cell line has the highest average of 10519.375

Conclusions

The scatter plot showed the RNA content of the virus in the mosquito cell line increased, indicating a mutation and possible increased adaptability. However, when the virus was cultivated in a human host cell, the RNA content decreased, suggesting another mutation change and a potential decrease in adaptability. Comparing the mean of the Tested host showed no significant difference between the mosquito cell line and the human cell. Therefore, we fail to reject our null hypothesis and there isn't enough evidence to support a change in the genetic makeup for the Tested Host. Nevertheless, when comparing the mean of the Adapted Host, there is a significant difference. This implies the genetic makeup of the dengue virus does change. This suggests it influences its ability to adapt to different host species. Additionally, we can see the virus has its strongest adaptability in the C6/36 cell. This show C6/36 cell is the best fitted host for the virus.

Biological conclusions

The genetic makeup of the dengue virus does influence its ability to adapt to different host species. By identifying the specific gene we can remove or block the gene. The concept of adaptability is linked to the frequency mechanics of individual alleles in the host species. Therefore lethal and deleterious alleles are held to low frequencies by negative selection, while beneficial mutations increase in frequency due to positive selection. Hence the frequency of a given allele over time and the type of mutation determines its adaptiveness.

Future work

Only four cell lines were cultivated for this experiment. What other cell can we use to test the dengue virus adaptability? What specific condition was the gene in that triggered a lethal or deleterious allele? What other cell line shows a decline in the RNA content when passaged?

References

Patrick T Dolan, Shuhei Taguwa, Mauricio Aguilar Rangel, Ashley Acevedo, Tzachi Hagai, Raul Andino, Judith Frydman (2021) Principles of dengue virus evolvability derived from genotype-fitness maps in human and mosquito cells eLife 10:e61921; https://doi.org/10.7554/eLife.61921

Data file

URL: https://purl.stanford.edu/gv159td5450

Code repository

GitHub: https://github.com/ptdolan/Dolan_Taguwa_Dengue_2020 [copy archived at https://archive.softwareheritage.org/swh:1:rev:adbf0dd213c5c9b422e55a9d97aeae9e7e64279f/]. Sequencing Data has been deposited as BioProject: PRJNA669406.