7995 Final Report

Howard Hu

August 13, 2022

1 Exclusive Summary

This report aims to help my undergraduate school colleague to finish his thesis. We investigate the relationship between the apoptosis rate and COVID virus transfection regent. To achieve this purpose, we first build a linear model and use an ANOVA test to determine the primary contributor to the apoptosis rate. Since the p-value of the experiment group is <2e-16, which is statistically significant, we conclude the experiment group can influence the apoptosis rate of each observation. Because the p-value of the interaction effect between time and group is 0.4397, which is not statistically significant, we conclude that there is no interaction effect between time and group. According to the result of our paired samples t-test, the p-value of the t-test is 0.752, which indicates the mean apoptosis rate of BHK and Vero is different.

2 Introduction

2.1 General Background

In 2019, the Chinese government reported a new member of the Coronaviridae virus family, Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which can damage people's respiratory system. It has caused a worldwide pandemic since it can cause asymptomatic infections. In previous studies, SARS-CoV-2 infection caused apoptosis in cell cultures, but the mechanism has yet to be elucidated. It has been shown that some relative accessory proteins of SARS coronavirus induce apoptosis, and corresponding proteins in SARS-CoV-2 may do the same. In this report, using the plasmid encoding the ORF9b protein of SARS-CoV-2, we investigated the apoptosis process in Vero E6 and BHK cells. The apoptosis process is a term in biology to describe the death process of cells. Transfection experiments and immunostaining showed that the expression of the ORF9b protein leads to apoptosis in both cell lines, changes in morphology, and the formation of apoptotic bodies.

2.2 Objectives

The first objective of Dr.∼ Zhongliu Liu is to figure out the primary contributor to the apoptosis rate. The second objective of Dr.∼ Zhongliu Liu is to find out any statistical significance between two different cells type.

2.3 Exploratory Data Analysis

The data is given by my undergraduate colleague Dr. Zhongliu Liu, who is currently a biology Ph.D. candidate at the University of Notre Dame. His study aims to determine whether coronavirus proteins induce apoptosis (cell death) in cells. We characterized the apoptosis phenomenon in Vero E6 and BHK cell lines with transfection of a plasmid encoding ORF9b protein of SARS-CoV-2. There are 37 groups of Vero E6 cells and 37 groups of BHK cells in our data set. We separate those 74 cells into three different groups: experiment, mock, and control group. For the experiment group, we add plasmid and transfection regent. We add transfection regent for the mock group (negative control group). For the control group, blank control group, we only add water. Then we record the number of apoptosis for Vero E6 cells and BHK cells in 24 and 48 hours.

Firstly, we would like to use a box plot and dot plot to see the distribution of apoptosis rates in two different cells.

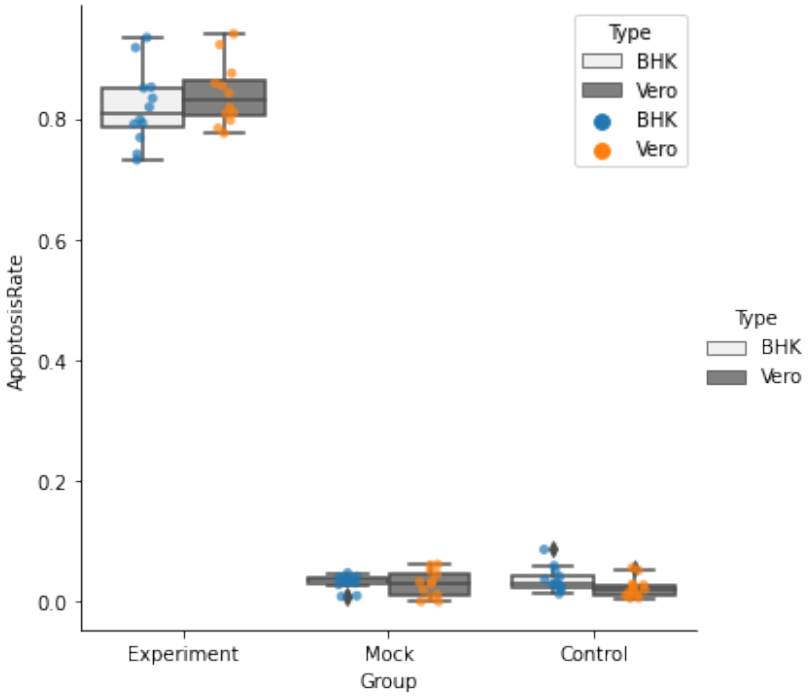

Figure 1: The box plot of three experiment group of BHK and Vero

Figure 1 shows that the median apoptosis rate of the experiment group is much higher than the median apoptosis rate of the mock and control groups. It indicates that adding the transfection regent will increase the apoptosis rate of two cells. Then we would like to two box plots to describe the distribution of mock group vs control group and experiment group.

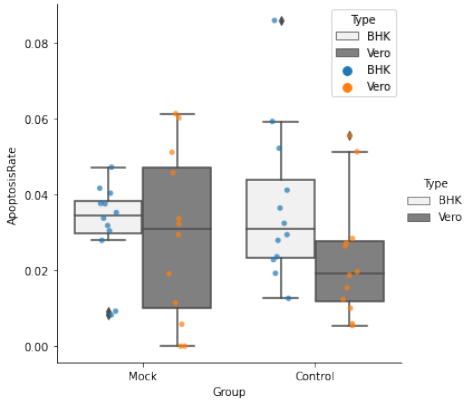

Figure 2: The box plot of mock group vs control group of BHK and Vero

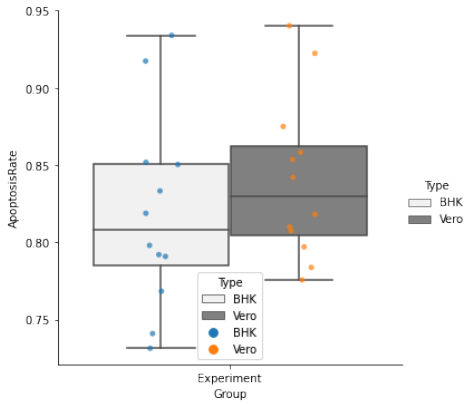

Figure 3: The box plot of experiment group of BHK and Vero

Figure 2 shows that the apoptosis rate difference between the mock and control groups in the two cells is close. Figure 3 shows the median apoptosis rate of Vero is higher than the median apoptosis rate of BHK.

Secondly, we would like to use a box plot to describe the distribution of the apoptosis rate of two cells in two different periods.

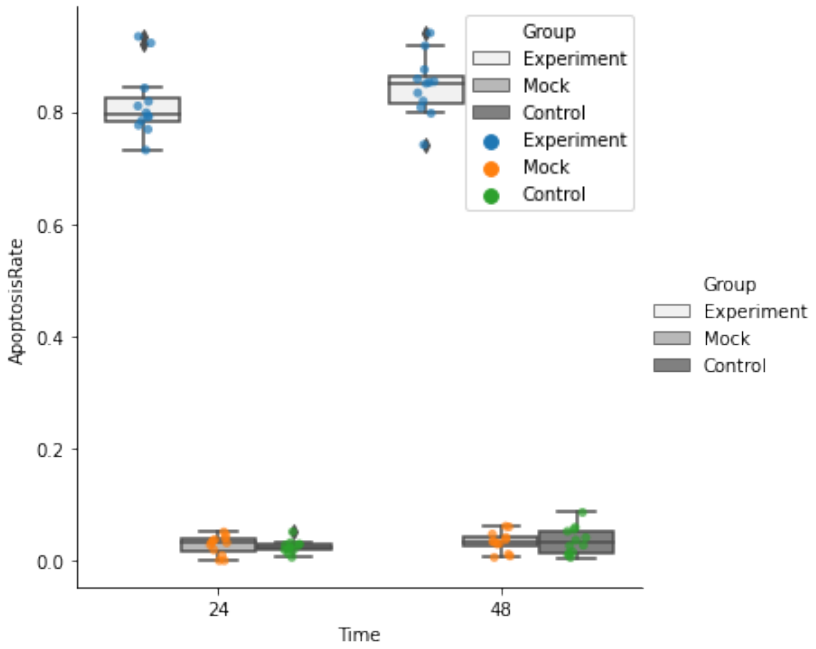

Figure 4: The box plot of three experiment group of BHK and Vero in two time periods

Figure 4 shows the median apoptosis rate of the experiment group is much higher than other groups at the same time. Therefore, we separate the original box plot into two more specified plots.

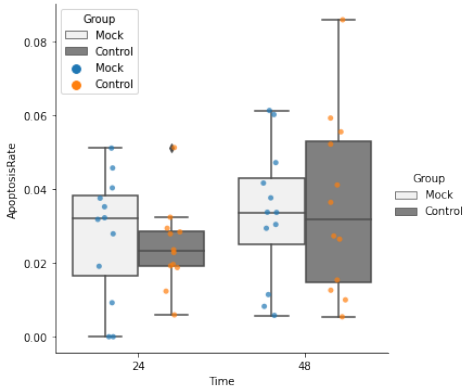

Figure 5: The box plot of mock group vs control group of BHK and Vero in two time periods

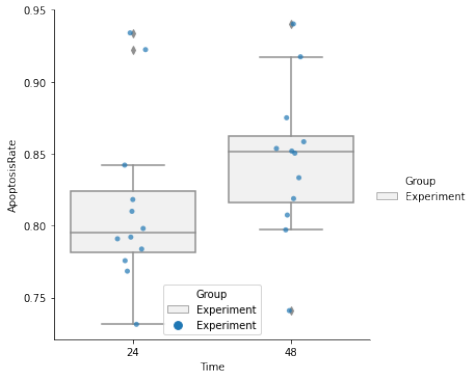

Figure 6: The box plot of experiment group of BHK and Vero in two time periods

Figure 6 shows that if we increase the experiment group's time, the median apoptosis rate will increase from 0.8 to 0.85.
Thirdly, we would like to use the line plot to investigate the relationship between apoptosis rate and time in three different groups.

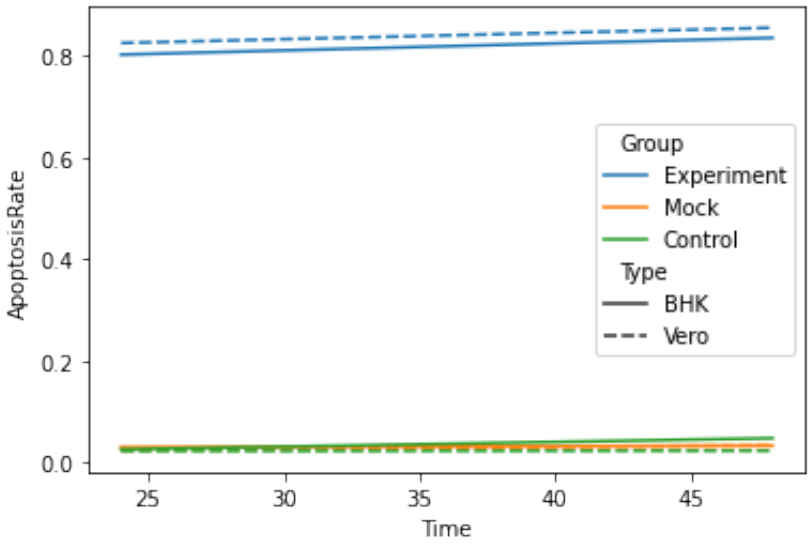

Figure 7: The line plot of the relationship between apoptosis rate vs time in three different groups

Figure 7 indicates that the difference between the two cells is very close. Fourthly, we would like to use a stem and leaf plot to describe the distribution of apoptosis rates in three different groups.

```
The decimal point is 1 digit(s) to the

7 | 34
7 | 78899
8 | 00112234
8 | 55568
9 | 2234
```

Figure 8: Experiment group

```
The decimal point is 2 digit(s) to the

0 | 0068919
2 | 8902244588
4 | 02671
6 | 01
```

Figure 9: Mock group

```
The decimal point is 2 digit(s) to the

0 | 56023599
2 | 0346788926
4 | 11269
6 |
8 | 6
```

Figure 10: Control group

Figure 8 shows the apoptosis rate in the experiment group is continuous and figures 9 and 10 show the apoptosis rate is discrete in mock and control groups.

3 Approach

To achieve our client's primary objective, we start by using the ANOVA test to find out which type of factor is statistically significant.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
as.factor(Time)	1	0.00	0.00	3.62	0.0615
as.factor(Group)	2	10.23	5.11	3908.41	<2e-16
as.factor(Type)	1	0.00	0.00	0.04	0.8491
as.factor(Time):as.factor(Group)	2	0.00	0.00	0.83	0.4397
Residuals	65	0.09	0.00		

From the previous table, we know the p-value of the group effect is <2e-16, which is statistically significant. Then we would like to build a linear regression model as the following formulation.

$$\text{ApoptosisRate} = time + group + type + time * group + e$$

type : The cell type of each observation

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.0235	0.0113	2.08	0.0412
as.factor(Time)48	0.0113	0.0148	0.77	0.4453
as.factor(Group)Experiment	0.7896	0.0148	53.48	<2e-16
as.factor(Group)Mock	0.0032	0.0148	0.22	0.8279
as.factor(Type)Vero	0.0016	0.0085	0.19	0.8491
as.factor(Time)48:as.factor(Group)Experiment	0.0201	0.0209	0.96	0.3390
as.factor(Time)48:as.factor(Group)Mock	-0.0055	0.0209	-0.26	0.7945

From the previous table, we find out the p-value of the experiment group is <2e-16, which is much smaller than our significant level of 0.05. It means the effect of the experiment group is one of the influential factors in our subject's apoptosis rate. Moreover, the p-value of the interaction effect of time×group is not significant. Therefore, we conclude there is no interaction effect of time×group in our experiment.

In addition, to solve the secondary objective of our client, we would like to use a paired samples T-test to find out is there any statistical significance between two different cells type.

Test	Results
1 Paired t-test:	t(35) = -0.32, p = .752, d = -0.08

From the previous table, we know the p-value of the paired t-test is 0.752, which is not statistically significant. Thus, we conclude the mean apoptosis rate of BHK and Vero is different.

4 Conclusion

We build a linear regression model and use an ANOVA test to determine the main contributor to the apoptosis rate. Because the p-value of the experiment group is <2e-16, which is statistically significant, we conclude the experiment group can influence the apoptosis rate of each observation. Since the p-value of the interaction effect between time and group is 0.4397, which is not statistically significant, we conclude that there is no interaction effect between time and group. Based on the result of our paired samples t-test, the p-value of the t-test is 0.752, which indicates the mean apoptosis rate of BHK and Vero is different.