

The Palestinian Covid Conflict

Waleejah Rana

10 December, 2022

Background and Overview

SARS-CoV-2 was the virus that caused the illness COVID-19. It was a member of the coronavirus family of viruses, which also included the viruses that caused the common cold and severe acute respiratory syndrome (SARS). SARS-CoV-2 was highly infectious and spread easily through respiratory droplets, such as when an infected person coughed or sneezed. Symptoms of COVID-19 could range from mild (fever, cough, and difficulty breathing) to severe (pneumonia, organ failure, and death). The virus was first identified in Wuhan, China in December 2019 and quickly spread to countries around the world, leading to the global pandemic. Scientists and health organizations worked to develop effective treatments and vaccines for the illness, and many countries implemented measures to slow the spread of the virus, such as quarantines and travel restrictions. However, the virus continued to spread and had a significant impact on global health and economies. In Palestine, the government implemented measures to prevent the spread of SARS-CoV-2, such as closing schools and public gatherings, and implementing quarantines and travel restrictions. However, the virus still had a significant impact on the country, with many people becoming infected and some losing their lives to the illness. Palestinians were in amidst a cloudy political climate with the people of Israel. The Israeli-Palestinian conflict not only became the reason for the increased cases per day, this conflict expanded SARS-CoV-2 geographically as well as exponentially across the nations. The objective of this analysis was to assess the differences in SARS-CoV-2 data from settlements of Israel in contrast to the settlements of Palestine regarding the Israeli-Palestinian conflict. This was done by utilizing Rmarkdown, git/GitHub, and the Linux command line to analyze the data taken from The National Center for Biotechnology Information (NCBI). I found that the conflict played an immense role in the rise of the virus in Palestinian territories rather than Israeli territories. The Palestinian-Israeli conflict is a long-standing dispute between the Palestinian and Israeli people over land and national identity. The conflict has its roots in the aftermath of World War II, when the British withdrew from Palestine and the United Nations partitioned the land into two states, one for the Jewish people and one for the Arab people. Since then, the conflict has taken many forms, including wars, terrorism, and political negotiations. The main issues at the heart of the conflict are the status of Jerusalem, the right of Palestinian refugees to return to their homes, and the Israeli settlements in the West Bank and Gaza Strip. The conflict has resulted in violence and suffering for both.

Methods

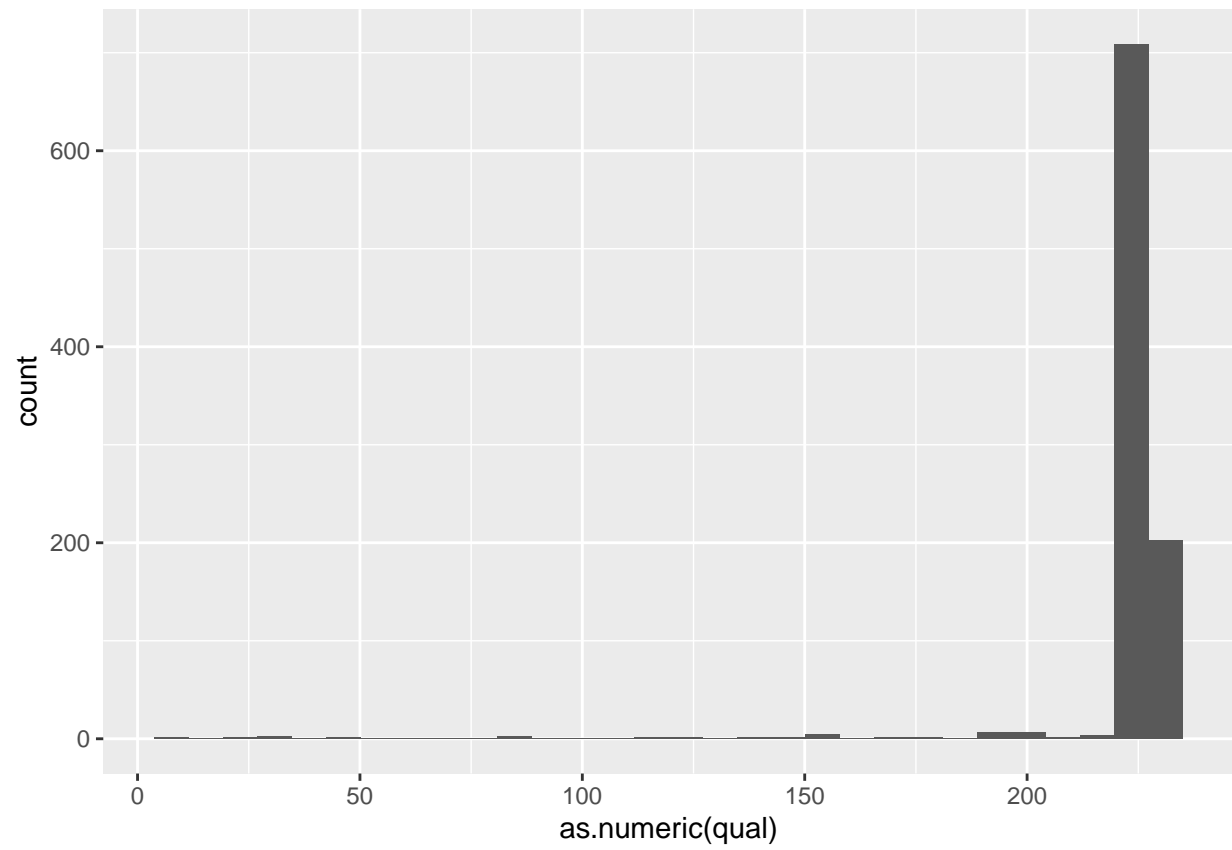
To begin my project I began to search for an adequate dataset by utilizing the National Library of Medicine (). The dataset needed to be a fair amount of samples, anywhere from 50+, and the dataset must have been illumina sequencing data. After finding the perfect sample sized illumina dataset; my Palestinian dataset, I then moved on to running that dataset through the SRA Run selector(). This indicated the amount of runs in the dataset and also allowed me to download the metadata of my dataset. Obtaining a dataset that had high quality data was crucial through this process, to access the quality of the data, the dataset had to be run through a 14-step bash pipeline created by Dr. Zimmerman. Dr. Zimmerman provided each student a unique Tule port container and ID to log into the USF server in order to successfully gain access to the 14 step Bash

pipeline with our own unique dataset. By entering our SRA dataset ID into the pipeline, the output produced a number of trimmed fastq VCF files. The number of files produced depended on the amount of samples, in my case, 69 VCF files were produced. Trimming the dataset was vital in order for the poor quality data to be disposed of properly. I used a libre server known as Cyberduck (), to analyze my data efficiently. After reading some samples, it was evident that some data needed to be trimmed to arrange an easy final output of my data; therefore, by entering into the 5th step of the bash pipeline, "05_trim.sh" I altered the "Minlen" to 50 to have the pipeline adjusted to my sequence data read length and was able to achieve a cleaner output. To start the process of gathering all my tabular data, I needed to invest a great deal in research about my country and the conflicts during the time of my data and how it correlated to SARS-CoV-2. There were plenty of news articles about the how the Israeli-Palestinian conflict left the Palestinian people far less equipped for the pandemic in regards to resources, infrastructure, and even vaccines than the Israeli people (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8583470/>) . There were a number of articles discussing the difference in the amount of vaccines distributed to the Israeli territories in comparison to the Palestinian territories (<https://www.reuters.com/article/factcheck-covid-israelpalestine-idUSL1N2PN1X4>). As a result of these articles, the data to code about this issue fit together like a puzzle. By navigating back on my R (Tule) server, I began to invest my time in the Report.Rmd creating my figures and tables using ggplot2. ## Subsections are ok too

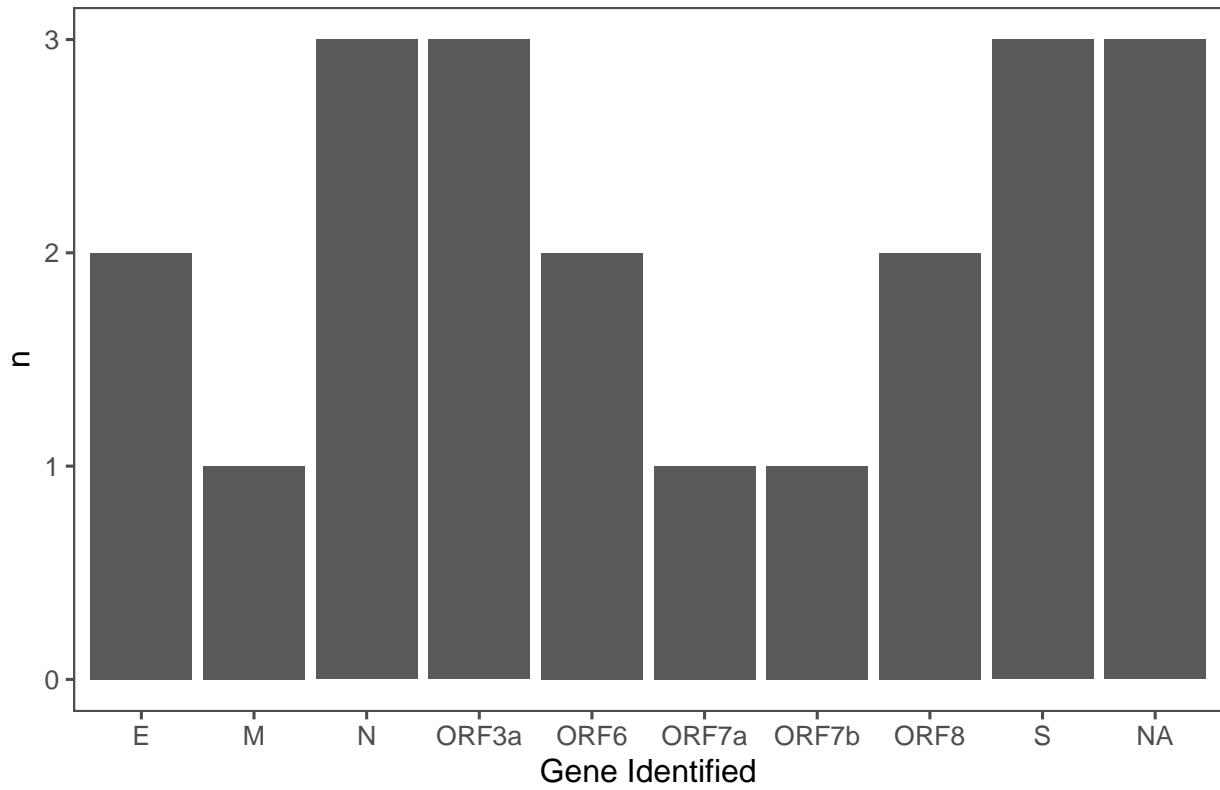
Results and Discussion

After conducting an abundance of research on the effects of the Palestinian-Israeli conflict in regards to Covid-19 cases in Palestine, I have concluded that the conflict had much to do with the improper distribution of vaccinations and the rise on Covid-19 cases in general.. This conclusion is supported by the data collected from the study Google conducted with Palestine, (<https://health.google.com/covid-19/open-data/explorer?loc=PS>), and Israel (explorer) which showed that the number of cases was significantly higher in Palestinian territories than in Israeli territories. The "Israel" country with Jerusalem as the district is also arguably Palestinian territory as well which supported my argument even more. With all of my VCF files and compressed data regarding my SNPs, I automated all of my tables and figures using R code in my Tule server. The first figure I created was a histogram that demonstrated the quality of the SNPs in my dataset. The figure displayed the count of good quality SNPs numerically which showed the validity of the quality of my dataset. My second and third figures compare the amount of Covid-19 vaccinations between the two nations. The significant increase in vaccinations in Israel correlates to the lack of resources distributed in Palestine. The difference in vaccinations was more than 300,000 due to the incorrectly labeled territory of Israel and Palestine. The Jerusalem district in the Israeli territory consists more of the Palestinian people.

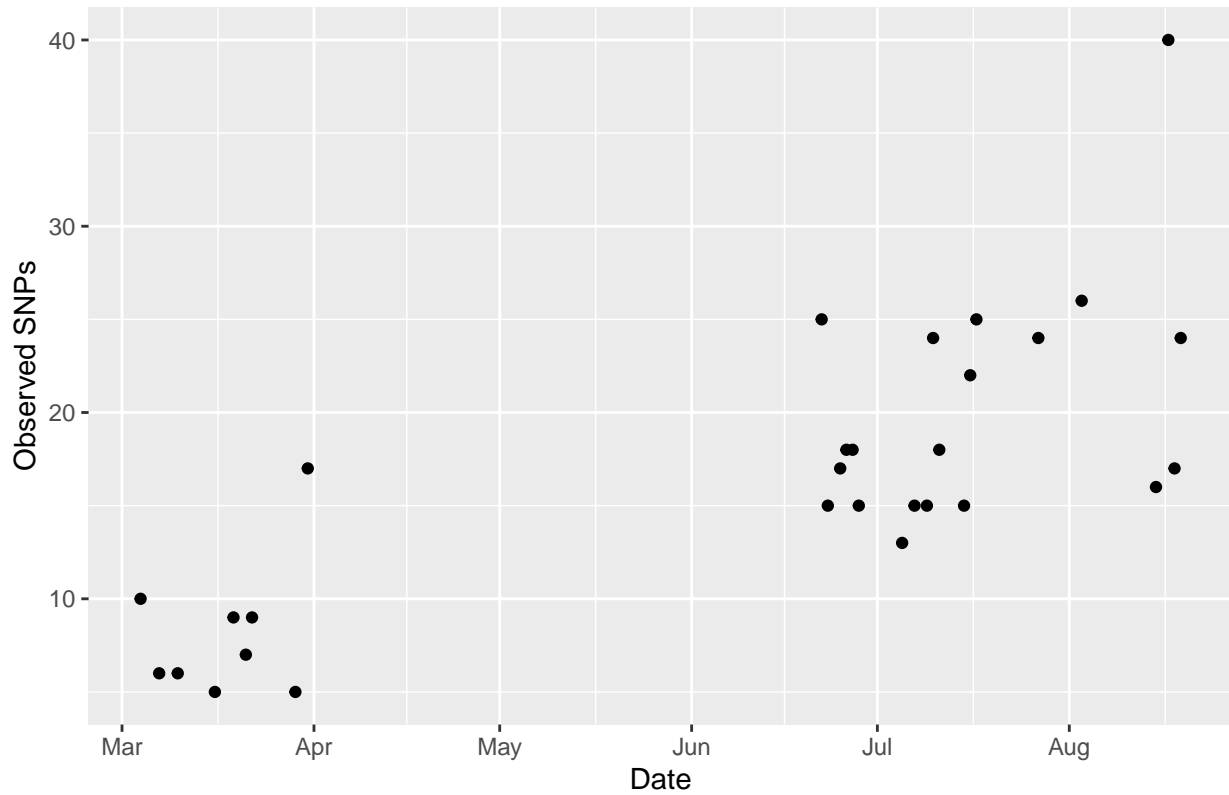
Figures



Average Spot length Within Each Gene



SARS-Cov-2 SNP Changes in Palestine



| Gene | Position | Reference | Variation | city1 | city2 | n |
|------|----------|-----------|-----------|-----------|-----------|----|
| N | 28881 | A | G | Jerusalem | Jerusalem | 15 |
| N | 28881 | A | G | Ramallah | Ramallah | 5 |
| N | 28881 | A | G | Jericho | Jericho | 4 |
| N | 28881 | A | G | Nablus | Der Hatab | 4 |
| N | 28882 | A | G | Jerusalem | Jerusalem | 15 |
| N | 28882 | A | G | Ramallah | Ramallah | 5 |
| N | 28883 | C | G | Jerusalem | Jerusalem | 15 |
| N | 28883 | C | G | Ramallah | Ramallah | 5 |
| N | 29179 | T | G | Jerusalem | Jerusalem | 5 |
| N | 29260 | T | G | Jerusalem | Jerusalem | 11 |

Figure 1: N and S genes have more unique SNPs in the set of samples analyzed.

Tables

| Collection Date | Release Date | Gene | SNP | Position |
|-----------------|--------------|--------------|-------|----------|
| 2020-03-04 | 2020-10-26 | T | NA | 241 |
| 2020-03-04 | 2020-10-26 | T | NA | 3037 |
| 2020-03-04 | 2020-10-26 | T | NA | 14408 |
| 2020-03-04 | 2020-10-26 | T | NA | 15324 |
| 2020-03-04 | 2020-10-26 | G | S | 23403 |
| 2020-03-04 | 2020-10-26 | T | NA | 1059 |
| 2020-03-04 | 2020-10-26 | T | NA | 14805 |
| 2020-03-04 | 2020-10-26 | C | NA | 20755 |
| 2020-03-04 | 2020-10-26 | T | ORF3a | 25563 |
| 2020-03-04 | 2020-10-26 | ACAAAACGTAAG | ORF6 | 27301 |
| 2020-03-07 | 2020-10-26 | T | NA | 241 |
| 2020-03-07 | 2020-10-26 | T | NA | 3037 |
| 2020-03-07 | 2020-10-26 | T | NA | 14408 |
| 2020-03-07 | 2020-10-26 | T | NA | 15324 |
| 2020-03-07 | 2020-10-26 | G | S | 23403 |
| 2020-03-07 | 2020-10-26 | T | ORF7a | 27661 |
| 2020-03-10 | 2020-10-26 | T | NA | 241 |
| 2020-03-10 | 2020-10-26 | A | NA | 1685 |
| 2020-03-10 | 2020-10-26 | T | NA | 3037 |
| 2020-03-10 | 2020-10-26 | T | NA | 14408 |
| 2020-03-10 | 2020-10-26 | T | NA | 15324 |
| 2020-03-10 | 2020-10-26 | G | S | 23403 |
| 2020-03-16 | 2020-10-26 | T | NA | 241 |
| 2020-03-16 | 2020-10-26 | T | NA | 3037 |
| 2020-03-16 | 2020-10-26 | T | NA | 14408 |
| 2020-03-16 | 2020-10-26 | T | NA | 15324 |
| 2020-03-16 | 2020-10-26 | G | S | 23403 |
| 2020-03-19 | 2020-10-26 | T | NA | 241 |
| 2020-03-19 | 2020-10-26 | T | NA | 3037 |
| 2020-03-19 | 2020-10-26 | T | NA | 14408 |

| Gene | Position | Reference | Variation | Occurrences |
|-------|----------|-----------|-----------|-------------|
| ORF3a | 25455 | T | G | 1 |
| ORF3a | 25526 | T | G | 1 |
| ORF3a | 25563 | T | G | 9 |
| ORF3a | 25572 | T | C | 1 |
| ORF3a | 25617 | T | G | 3 |
| ORF3a | 25665 | T | C | 2 |
| ORF3a | 25785 | T | G | 50 |
| ORF3a | 25844 | T | C | 3 |
| ORF3a | 26144 | T | G | 2 |

Table 1: Gene names, locations, and lengths in the SARS-CoV-2 genome. Higher SNP counts in the S and N genes may be related to the larger size of these genes.

Sources Cited