Assignment 1

Abstract:

As a result of studying the genetic variants and mutations that occurred in 1200 genomes of SARS virus and presenting an analytical study, it was found that some mutations in the SARS genome do not happen randomly. And also found the most important good candidates for drug development and treatment of Covid-19 disease through research, where it was found that the coding region at the nucleotide level of NSP13 protein is relatively conserved compared to other protein regions in the ORF1ab gene.

Introduction:

In December 2019, in the Wuhan region of China, the spread of the COVID-19 virus began as a result of the SARS-CoV-2 virus, as this virus and other pathogens spread rapidly and ferociously as a result of the genetic changes, including the sequence difference, adding unknown variables to the immune system.

Whereas, the variation in the genomic structure of the SARS COVID-19 genome occurs according to environmental conditions (ultraviolet rays ... minerals).

It has been scientifically proven that the highest rate of mutations among all living organisms is in viruses, especially single-stranded viruses .. The biggest difference that affects all mutation rates is DNA and RNA viruses. Despite the dangerousness of SARS-Covid-2, we discovered new chemicals and treatments for Covid-19 disease by studying the differences in the genetic sequence at the level of its nucleotides.

Assignment 2

Related work:

(Li et al., 2020): provide an excellent overview of SARS-evolutionary CoV-2's history and the future intermediate transferring species.

(Yoshimoto,2020): described the full collection of SARS-CoV-2 genes and proteins.

(Saha et al. 2020): The virus will develop into a better version of itself by mutations to suit best in the host environment, according to Saha et al. The virus uses mutation as a method to acclimate with its environment.

(Tai et al. 2020): investigate and present all aspects of the virus and disease in the sense of its transmission to humans through the Spike protein, as well as the sequence of molecular and biological functions involved in the process.

(Petropoulos and Makridakis,2020): provide realistic predictions for confirmed coronavirus disease, as well as a study and timeline of the disease's possible consequences for preparation and decision-making.

(Andersen et al.2020): present an analysis of the virus's prominent characteristics and origin hypotheses.

(Khailany et al.2020): presented an analysis of genetic differences and mutation comparisons of identified SARS-CoV-2 genetic data across different time frames and places, Analysing 95 SARS-CoV-2 complete genomes submitted to various databases through April 2020.

(Emameh et al .2020): To classify the proteins in the ORF1ab region of the genome, a data mining and computational analysis of SARS-CoV-2 isolates from oronasopharynx of Iranian patients was presented. SARS-CoV-2 polyproteins are cleaved by virus-encoded cysteine proteinases, which encode 16 nonstructural proteins (NSPs), with ORF1a encoding NSP1 to NSP11 and ORF1b encoding NSP12 to NSP16.

Assignment 3

Methodology:

During the first seven months of 2020, 12 complete genome sequences were collected for SARS-Covid-2, each set contains 100 complete genome sequences. The total number of available sequences in NCBI is shown based on the month of collection for each month, March and April were the peak months in data collection for the genome this virus. Use the results of Clustal Omega and Jalview to extract

the genomic variants in each Si data set with the aim of finding important mutations. It is expected that the virus mutates and changes to itself constantly. Most of the analogous work on SARS-Covid-2 variation is focused on extracting, identifying, and studying specific mutations or mutations within a specific coding region. Other studies focus on the difference in mutations based on geographic regions and chronological progression. Koyama et al (2020) studied the SARS-CoV genome and were able to identify more than 5,700 distinct genetic mutations.

Results:

Thirteen common genetic variants were collected in nearly all ten groups.
The mutation in the ORF1a region are changes in the third nucleotides in each triplet of the codon, which retained the amino acid of the reference genome in all mutations except for one mutation, which is due to the hypothesis of oscillation. NSP13 also contained another mutation that changes this codon from encoding a hydrophilic polar amino acid to a hydrophobic non-polar amino acid, causing this amino acid to encounter inward rather than outward. Autophagy plays an important role in the virulence of this virus, as the autopharynx determines which parts of the cell are for intracellular degradation by the cell particle.
In the second gene, ORF2, which codes for the spike protein, where the Spike protein acts as a binding to the cell receptors to start the virus attack at the cellular level, and a continuous mutation in this region may indicate that the virus has become more effective in transmitting the virus from person to person. Moreover, this protein is important for the body to develop antibodies as an immune response to this virus, as spike protein is the antigen used to make the antibody. The ORF3a third gene contained a missense mutation that changed a previously, highly polar, hydrophilic basic amino acid to a largely non-polar and hydrophobic amino acid, which could lead to the disruption of proteins with different functions and this mutation is rapidly transmitted over time in the dataset from S5 to S7
In the ORF8 gene, the mutation may not affect the protein structure or function, it may be a diffuse mutation because it helps to hide itself in the human body or it may affect the transmission of infection from human to human and this is the reverse direction of the mutation that was studied in ORF3a.

The development of NPS13 mutations is particularly interesting because it affects the NSP13 protein encoded by ORF1ab. Indeed, according to Mirza and Fruin (2020), the genetic coding for NSP13 is highly conserved and should be used to produce inhibitors and treatments for this virus. The mutations present in the NSP13 region are very strange because this region is very protected and another function has been discovered for it as it acts as an antagonist. For interferons.

These three mutations (28881g> a, 28882g> a, 28883g> c) belong to the nucleocapsid protein (N protein) and this protein can lead to immune responses that may lead to progress toward developing a vaccine using this protein.