PAPER RESEARCH

Name: Nada Samy Mahmoud.

Level: 200 Bioinformatics.

Paper Name: *COVID SEQUENCE, a New Tool for SARS-CoV-2 Genome Analysis and Visualization: Development and Usability Study.*

Introduction:

What is a coronavirus?

Coronavirus are a large family of viruses that are known to cause illness ranging from the common cold to more severe diseases such as Middle East Respiratory syndrome and severe acute respiratory syndrome “SARS”, coronavirus can transmitted from person to person and to understand its evolution and genetics scientists have sequenced SARS COVID-19 from patients

and make statistics from group of people with different ages and genders so now we have a huge data to take a genomic sequences to keep up with the latest updates developments scientists need to frequently download and clean a new data sets , **COVID -19** SEQ consists of different components : a data analysis that takes **FASTA** sequences and generates variant call sets in **VCF** *"variant call format* "also known as **processed files** : includes a genomic data and **ORF** *"open reading frames"* is the part of reading frame that has the ability to be translated , the pipeline automatically filters low –quality sequences and remove duplicate sequences , performs sequence alignment and identifies ,annotates genetic variant , we use a web server to enable the rapid analysis of sequences without using programming web interface includes an interactive genome visualizer and tabulated displays of genetic variants and ORF predictions further we use command-line interface to facilitate data sharing .

Abstract:

The genomes of SARS **COVID-19** rapidly sequenced and to keep up with updates and evolution scientists want to refresh and re-clean data sets but scientists have limited with Bioinformatics tools and programming to analyse the sequences so to handle these problems they developed **COVID -19** SEQ by using "web server "which facilitate analysing the sequence implemented in python and JavaScript using web server the results when we have a new sequence **COVID -19** SEQ predicts Gene Boundaries ,the locations of genes ,identifies genetic variations ,identifying elements on the genome ,a process called gene prediction and attaching biological information to these elements by using **A command- line interface** is available for high throughput processing so the conclusions scientists have developed the SARS-COV-2 sequences and they handled web service for fast and easy analysis so the web server provides an interactive module for analysis and annotations the genome they thought that understanding the genome will help to know a whole sequences an predict any variations in the sequences .