

## Importance of Bioinformatics in COVID-19 analysis

Bulbul Ahmed

Dept. of Bioinformatics

Indian Agricultural Research Institute, Library Avenue, PUSA, New Delhi-110012

ARTICLE ID: 012

Bioinformatics approaches helps in identifying genes, proteins and pathways of COVID-19 virus (Dolan et al., 2020). It is also necessary to find out the information associated with other genes, proteins for finding the efficient solution of COVID-19. Various approaches are as follows:

**BLAST:** On performing BLAST of SARS-CoV-2 (Severe acute respiratory syndrome coronavirus 2, accession genome with GCF\_009858895.2 using BLOSUM62 matrix with default parameters in NCBI blast (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) (Caldera *et al.*, 2017).

**Phylogenetic analysis:** On mapping severe acute respiratory syndrome coronavirus 2 reference genome, the top most hit is associated with severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1 and two hits are associated with Wuhan sea food market for top 100 hits. These hits downloaded in multiple-fasta format and translated the reads into amino acid for alignment using MUSCLE (codon) v 10.1.8 (Waterhouse *et al.*, 2009). The next is to construct and visualize the phylogenetic tree using MEGAX v10.1.8 (Kumar *et al.*, 2018) using maximum likelihood with default parameters.

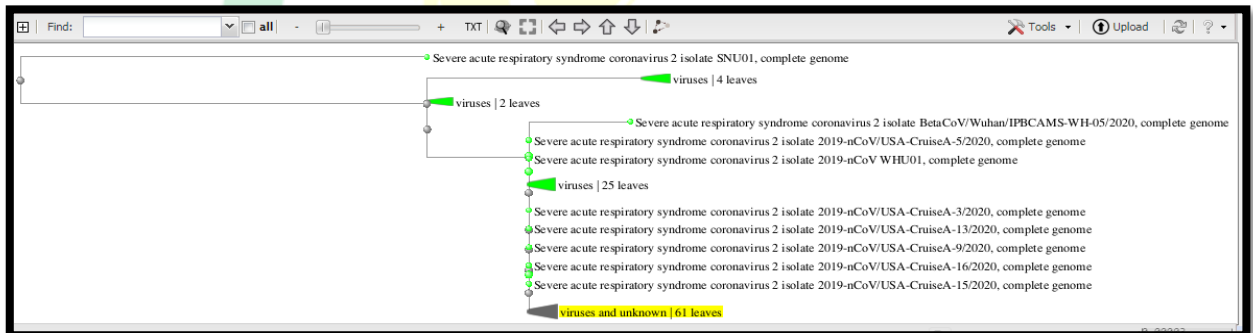
Table 1. Details of SARS COVID-19.

Results Statistics	
Length adjustment	40

Effective length of query	29863
Effective length of database	270197372709
Effective search space	8068904141208867
Effective search space used	8068904141208867

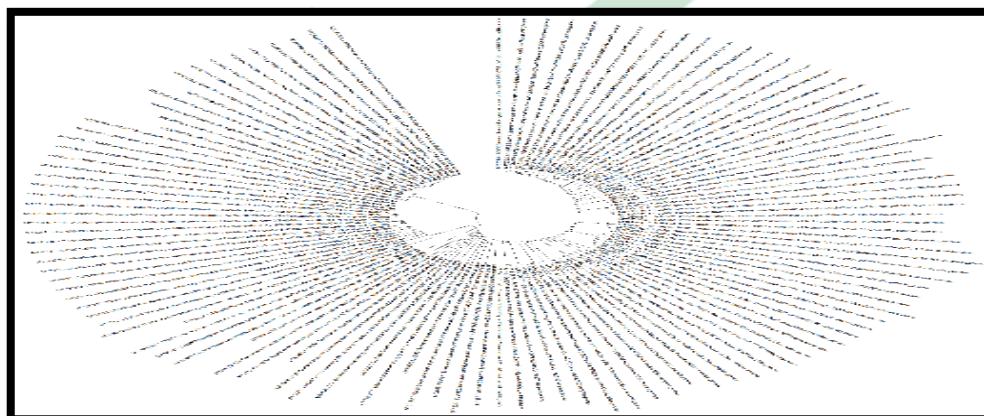
In blast, it is observed that 29863 nucleotides are there in the searched genome and the database found with 270197372709 nucleotides shown in table 1.

In phylogenetic analysis it was observed that nine are related to respiratory whereas 61 are unknown virus that have been generated in the phylogenetic tree. Out of all, 17 virus are seen in one major group i.e. MT192773.1:1-29890, MT163716.1:2-29903, MT263396.1:5-29903, LR757995.1:2-29872, MN988668.1:1-29881 whereas another group has three groups where one group have three viruses with associated with pneumonia virus originated from Wuhan seafood market and SARS-CoV-2 (human severe acute respiratory syndrome coronavirus 2) genome i.e. LR757996.1 :1-29868, MT258379.1:1-29892 S, MT258378.1:1-29892 shown figure 1.



(Figure 1. Major SARS Virus)

All other leaves also associated with severe acute respiratory syndrome coronavirus shown in figure 2).



(Figure 2. Circular tree construction of SARS virus)