



COMPARATIVE STUDY ON GENOMIC SEQUENCE OF SARS-COV2, SARS- COV1, H1N1 AND H3N2 VIRUSES

¹Dhanalakshmi.G,²Uzma farheen, ³Anadi

¹Assistant Professor,

Department of Biochemistry,

Padmashree institute of management and sciences, Bangalore, Karnataka.

Abstract: Viruses cause various diseases and syndrome in humans. Some have even evolved with time to show infection in humans. Human Beta coronaviruses of SARS-CoV-2, SARS-CoV, and influenza virus are very common pandemic caused viruses. Even the symptoms, transmission and the severity of the disease also very close to each other. But, current pandemic caused SARS-CoV-2 infection and mortality rate was too high. At present, there are no specific, effective drugs or vaccines for SARS-CoV-2. But, there are some drugs, vaccines, treatment and prevention of influenza virus are available. Based on this information, this study highlighted to compare the genomic sequence and GC content of SARS-CoV-2, SARS-CoV-1 and influenza viruses. From this study we conclude SARS-CoV-2 sequence has high genomic variety of due to this reason it's get consecutive transmutation.

Keywords: SARS-CoV-2, COVID 19, SARS-CoV-1, influenza viruses, H1N1, H3N2

I. INTRODUCTION:

The novel coronavirus 2019 (SARS-CoV-2), Severe acute respiratory syndrome coronavirus (SARS-CoV), and influenza A viruses are major pathogens that primarily target the human respiratory system. Diseases associated with their infections vary from mild respiratory illness to acute pneumonia, respiratory failure and even death.

The current pandemic was caused by new betacoronavirus (SARS-CoV-2). It was identified in china at December 2019, WHO was announced a pandemic in March 2020 (WHO, 2020). These cases were characterized by acute pneumonia-associated symptoms, such as fever, dry cough, chills, shortness of breath, and muscle pain (Lu et al., 2020). It spreads mainly through droplets of salivation or release from the nose henceforth requested to cover nose and mouth while coughing or sneezing and to maintain social distance in common places (Dharmendra Kumar et al., 2020) for this reason this outbreak is rapidly spread worldwide. As per WHO's report on 25th December, 7,81,94,947 individuals are infected and death rate is more than 17,36,752 occur in world. In India total number of infected cases are 1,01,46,845 and death rate was 1,47,097. (<https://worldhealthorg.shinyapps.io/covid/>).

Severe acute respiratory syndrome (SARS) is a viral respiratory disease caused by a SARS-associated coronavirus. This outbreak was identified in china at 2003 and spread to 4 other countries, most of the cases had animal exposure before developing the disease. Some investigations disclosed that SARS-CoV strains were transmitted to palm civets from other animals (Guan et al., 2003; Wang et al., 2004). Hu et al., suggest

that bats may be the natural reservoirs for the virus and that palm civets are only intermediate hosts (Hu et al., 2017).

Influenza virus also called “flu”virus.H1N1 is a subtype of Influenza A virus well known outbreaks of H1N1 strains (*pdm09* virus, also known as the novel influenza A virus) in humans include the 2009 swine flu pandemic, as well as 1918 (Spanish influenza)flu pandemic. H3N2 pandemic was caused by Hong Kong flu virus in 1968; due to these 1 million deaths occur in worldwide (Viboud et al., 2005). It contains glycoprotein like haemagglutinin and neuraminidase, for this reason, they are described as H1N1,H1N2 etc. depending on the type of H or N antigens that express with metabolic synergy (Zeinab Abdelrahman et al., 2020). The flu genome is made of seven unique pieces called segments, when two diverse virus co-taint a similar cell, they can deliver an offspring containing the genetic material of both parental strains (<https://www.ias.edu>). This cycle was the core of pandemics, when novel virus produced by the reassortment can communicate from human to human. The H1N1 virus arrows a complicated assortment process whose story is still not fully elucidated (Raul Rabadan 2006).

The SARS and influenza viruses have almost similar symptoms and the way of transmission. Based on this reason and for better understanding of the current COVID-19 pandemic, we planned to compare the sequence of SARS COV -1, SARS-CoV-2, H1N1 and H3N2 viruses.

II. METHODOLOGY:

A genome sequence was downloaded from the viral genome sequences database of the National Center for Biotechnology Information (NCBI), and different comparative analyses were performed. For this analysis, reference genome sequence of SARS-CoV-2 (NC_045512) and SARS-CoV-1 (DQ640652) was isolated from China. An influenza virus of H1N1 (NC_026438.1) and H3N2 (NC_007370.1) sequence was isolated from California and New York was downloaded.

Sequence alignment: Samples were aligned to the reference genome in the nucleotide BLAST using the Basic Local Alignment Search Tool (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) to extract the differences between the genome variants. Quantitative study of the sequence alignment was done by studying the query cover, E value and Identify percentage.

Comparison of the GC content: The Genomic GC content was calculated among all the sequences. It was calculated by the GC content calculator of Biologics International Corp (BIC) (<https://www.biologicscorp.com/tools/GC Content>). FASTA format sequence was entered in the box provided (in FASTA format). The window size was selected for 30 and the sequence was submitted to obtain the results. GC content is also called G+C ratio or GC-ratio and was calculated as a percentage. GC-content percentage was calculated as $\text{Count (G + C) / Count (A + T + G + C) * 100\%}$.

III. RESULT AND DISCUSSION:

Sequence comparison: The sequence comparison of SARS CoV2, CoV1, H1N1 and H3N2 Length, query coverage, E value and Percentage identity was showed in **table 1**. The quantitative measurement to determine the similarities between the sequences species which are closely related to each other show higher percentage identity and thus reflect relatedness.

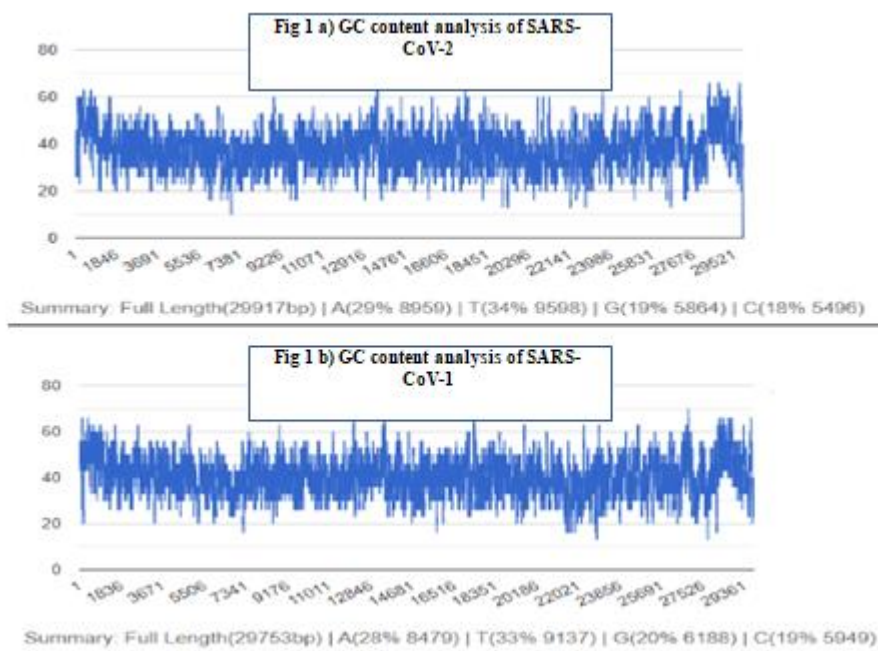
Table1: Sequence comparison of SARS COV 2, SARS COV 1, H1N1 and H3N2

Name of the Strains	GeneBank Acession Number	Length (bp)	Query Cover (%)	E value	Identidy with reference sequence (%)
SARS-CoV-2	NC_045512.2	29903	100	0	100.00
SARS-CoV-1	DQ640652	29746	98	0	87.34
H1N1	NC_026438.1	2187	13.64	0.6	84.00
H3N2	NC_007373.1	2015	14.84	1.7	82.62

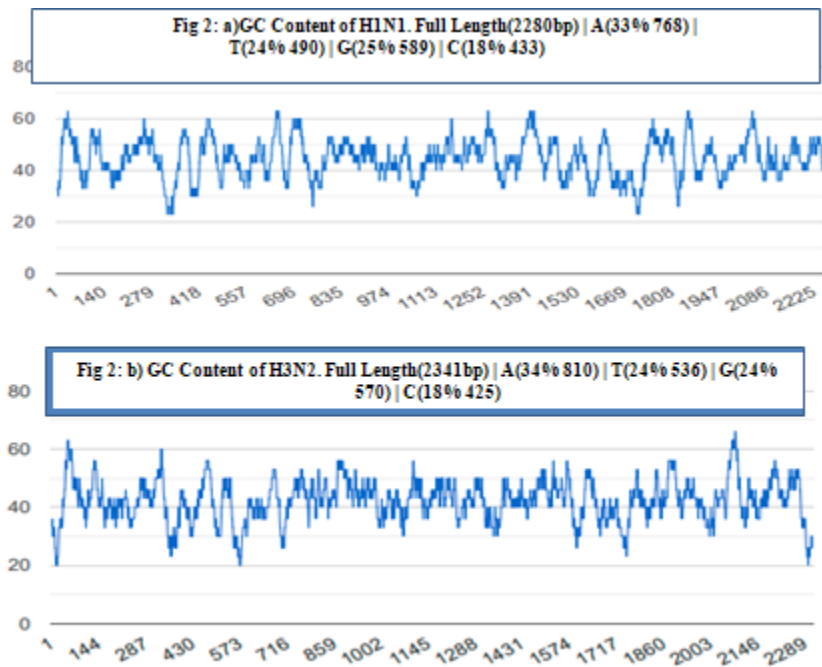
From the outcome, SARS-CoV-1 has 98% of the query coverage and only 87.34% of identity with the SARS-CoV-2 sequence. However H1N1 and H3N2 sequence length (2187, 2015) and also have less query cover percentage compare with other sequences. Eskild Petersen reported that H3N2 did not match with SARS COV 2 and showed less similarity (Eskild Petersen et al., 2020).

This sequence homology explains that both SARS-CoV-2 and SARS-CoV-1 have bats serving as reservoir hosts for its progenitor (Kristian G. Andersen et al., 2020). Jiabao Xu et al speculate the purpose behind the distinction in SARS COV 2 and influenza virus was that it has a more limited middle hatching period and a more limited serial interval than SARS COV 2 virus. With this understanding that influenza can spread quicker than SARS COV 2 (Jiabao Xu et al., 2020).

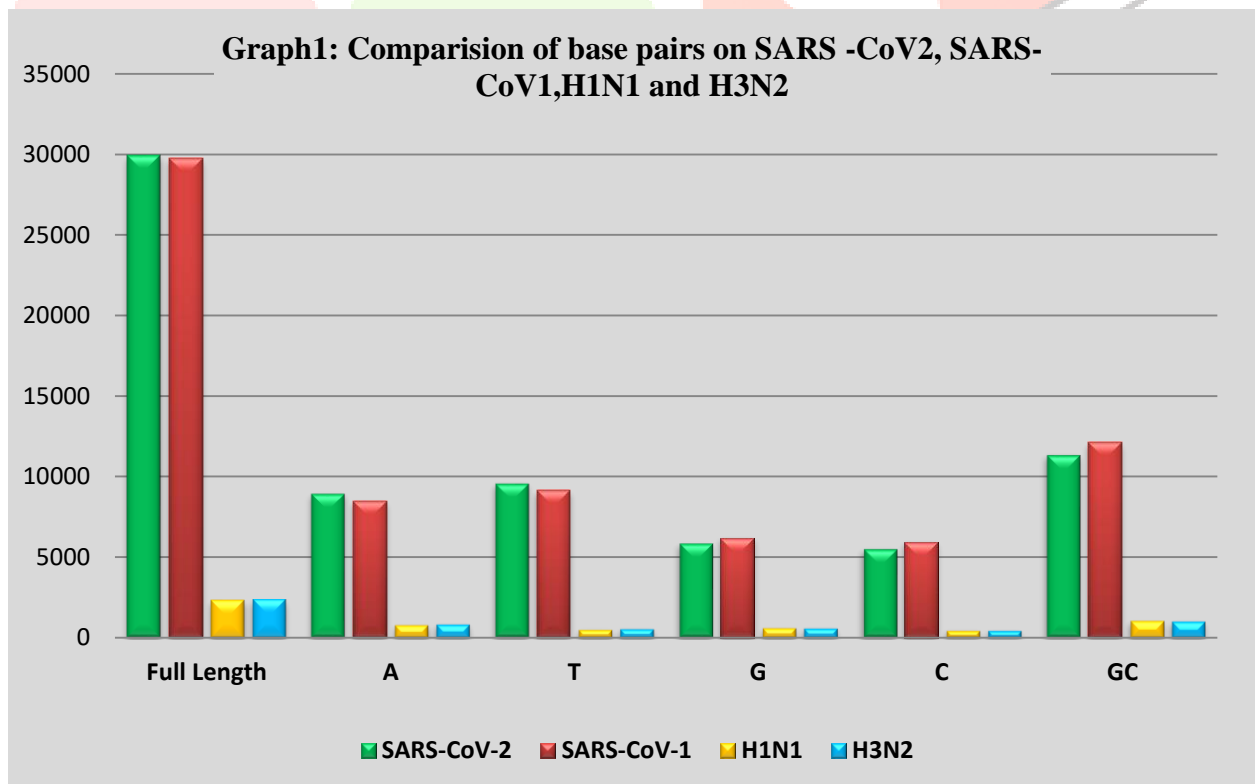
Comparison of GC content: The genomic GC content was calculated for all the sequences: SARS-CoV-2 reference genome (NC_045512), SARS-CoV-1 (DQ640652), H1N1 (NC_026438.1) and H3N2 (NC_007373.1).



The GC content is important as it directly affects the viral codon adaptation index. This study can help in predicting the efficiency of viral gene expression. The GC content of SARS-CoV-1 (39%) was higher than the SARS-CoV-2 (37%) (Fig 1. a and b).



However the A(29%) and T (34%) content was higher in SARS-CoV-2. H1N1 and H3N2 GC content was shown in **Fig 2. a and b**. In overall comparison, GC content of H1N1 (43%) and H3N2 (42%) was higher than the other sequences (**Graph 1**).



GC content fundamentally influences genome working. High GC content gives solidness and a mind boggling quality guideline framework to the creature (Petr smarda, 2014).

GC content interacts with genome 37 polarities and codon usage bias in RNA viruses. Genome composition and codon usage bias are particularly interesting in RNA viruses because the same RNA may be used as mRNA, genome, or anti-genome. The replication of RNA genome is also very different from DNA replication of the host using different polymerase enzymes and taking place in different environments, which may contribute to the mutational bias that drives the genome composition (Auewarakul, 2005).

IV. CONCLUSION:

The SARS-CoV-2 pandemic has already caused extreme horribleness and mortality in elderly peoples higher than in the pandemic influenza. The signs and side effects of SARS-CoV2, SARS-CoV1 and influenza virus are similar. The majority of them contaminate the respiratory tract through direct contact and aviation route droplets and cause comparable clinical manifestations like fever, cough, and sore throat. Extreme contamination prompts bilateral pneumonia, intense respiratory misery disorder (ARDS), respiratory failure, and even death.

Based on this study, it speculates that the SARS CoV 2(NC_0445512) reference sequence length (29903) was highly differing from the H1N1 (NC_026438.1) and H3N2 (NC_007373.1) of 2187 and 2015 respectively. Even in the identity percentage and GC content also have a significant difference. From this outcome we expect that, high genomic variety of SARS CoV 2 has opportunities to get successive transformation. May be it's a one of the major reason to not able to get a success with the pharmacological research such as medicine and vaccine preparation. We should go on with to be attentive for the emergence of new variants or sub strains of SARS CoV 2 should be gathered at one place for better understanding. It's highly essential for further research and develop categorize more effective methods of treatment and prevention to better control the COVID- 19 pandemic.

REFERENCES:

1. Auewarakul, Composition bias and genome polarity of RNA viruses. *Virus Res*, 2005,109(1);33-37.
2. Dharmendra Kumar, Rishabha Malviya, Pramod Kumar Sharma ., Corona Virus: A Review of COVID-19. *Eurasian Journal of Medicine and Oncology*. 2020. 4(1):8–25.DOI: 10.14744/ejmo.2020.51418.
3. Eskild Petersen , Marion Koopmans , Unyeong Go, Davidson H Hamer ,Nicola Petrosillo , Francesco Castelli , etal., Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics, *Lancet Infect Dis*, 2020,20(9):e238-e244. doi: 10.1016/S1473-3099(20)30484-9.
4. Guan Y, Zheng BJ, He YQ, Liu XL, Zhuang ZX, Cheung CL, et al. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science*. (2003) 302:276–8. doi: 10.1126/science.1087139.
5. Hu B, Zeng LP, Yang XL, Ge XY, Zhang W, Li B, et al. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog*. (2017) 13:e1006698. doi: 10.1371/journal.ppat.1006698.
6. Jiabao Xu, Shizhe Zhao, Tieshan Teng, Abualgasim Elgaili Abdalla, Wan Zhu, Longxiang Xie et al., Systematic Comparison of Two Animal-to-Human Transmitted Human Coronaviruses: SARS-CoV-2 and SARS-CoV, *Viruses*. 2020 Feb; 12(2): 244.
7. Jon Bohlin, Vegard Eldholm, John H. O. Pettersson, Ola Brynildsrud and Lars Snipen, The nucleotide composition of microbial genomes indicates differential patterns of selection on core and accessory genomes. *BMC Genomics*. 2017, 18, 63-67.
8. Kristian G. Andersen, Andrew Rambaut, W. Ian Lipkin, Edward C. Holmes and Robert F. Garry, The proximal origin of SARS-CoV-2, *Nature Medicine* , 2020, (26), 450–45.

9. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*. (2020) 395:565–74. doi: 10.1016/S0140-6736(20)30251-8.
10. Petr Smarda, Ecological and evolutionary significance of genomic GC content diversity in monocots. *Proceeding of the National Academy of Sciences of the United States of America*, 2014, 201, 68-72.
11. Raul Rabadan, Reassortment events in the evolution of influenza H1N1 virus. *Journal of General Virology*, 2009, 90-105.
12. Viboud C, Grais RF, Lafont BA, Miller MA, Simonsen L. Multinational impact of the 1968 Hong Kong influenza pandemic: evidence for a smoldering pandemic. *J Infect Dis*. (2005) 192:233–48. doi: 10.1086/431150.
13. Wang M, Xu HF, Zhang ZB, Zou XZ, Gao Y, Liu XN, et al. Analysis on the risk factors of severe acute respiratory syndromes coronavirus infection in workers from animal markets]. *Zhonghua Liu Xing Bing Xue Za Zhi*. (2004) 25:503–5.
14. WHO, *Coronavirus disease (COVID-19) pandemic*. [Online] Available at: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>, 2020, [Accessed 23 May 2020].
15. Zeinab Abdelrahman, Mengyuan Li, and Xiaosheng Wang, Comparative Review of SARS-CoV-2, SARS-CoV, MERS-CoV, and Influenza A Respiratory Viruses, 2020, (11) 552909.

