In Search of Novel Coronavirus 19 Therapeutic Targets

Jyothsna Gundlapally, Amit Kumar, Amita Kashyap, Ajit Kumar Saxena, Aryasekhar Sanyal

BioAxis DNA Research Centre, Hyderabad, India
All India Institute of Medical Sciences, Patna, India
Loreal Research and Innovation, Bangalore, India

Email: jyothsna@dnares.in, amit.kumar@dnares.in, amita@dnares.in, aksaxena1@rediffmail.com, aryasekhar.sanyal@rd.loreal.com

Received: 25th March 2020, Accepted: 27th March 2020, Published: 28th March 2020

Abstract
Novel Coronavirus 2019 pandemic has become a nightmare of the year 2019-20. It is affecting both health and wealth across the world. It has become a great challenge for the entire human race to protect itself from the viral outbreak. This is time for the entire Scientific community to come together and undertake studies and contribute in conducting research on CoVID 19 and possible solutions to defeat this killer, million times smaller than humans, as even minute information can also play a very important role in fighting against the Virus. The current work is aimed to analyze the genome of CoVID 19 and compare its evolutionary relation with the other species of viruses that are known to cause respiratory disorders.

Viral membrane proteins and proteins involved in replication of viral genetic material play an integral part in virus–host interactions. These classes of protein are often the best candidates for antiviral drug and vaccine targets. Disrupting these proteins may be an effective means to inhibit the growth and disintegrate the virus. Taking advantage of the recent release of some of the gene sequences and the genome of Novel Coronavirus 2019 by NCBI GenBank and the agility provided by Insilico Bioinformatics tools, the current work aimed to study the evolutionally conserved regions of the genome of the CoVid 19. The comparison of the complete genome and specifically the coding gene sequence for membrane proteins and proteins involved in viral replication of MN908947.3 Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome- 2019 (better known as Covid 19), Isolated from China, was conducted. 25 viruses including commonly known respiratory tract pathogen were selected for the current study. Sequence similarity analysis and comparative study results revealed that the viral membrane protein, M protein Shares similarity only with the other corona group of viruses and (MERS) and not with HCOV. Moreover, the complete genome comparison revealed the presence of a specific conserved gene region shared by MERS and CoVid 19, which was further analyzed. Identifying the commonly shared gene regions can immensely aid in identifying druggable target and help in development of appropriate therapy protocol or medication for 2019 novel Coronavirus.

Keywords
COVID 19, Coronavirus, M Protein, COVID Drug Target, Polyprotein

Introduction
COVID19 pandemic started on 31st December 2019 at Wuhan, China [1]. Till date it has affected 192 countries and territories around the world [2]. A large number of Mortalities were recorded in China, Iran, Italy and Republic of Korea [3]. According to the reports released by WHO as on 23rd March 2020 a total of 294,110 individuals were tested CoVid19 positive. The number of deaths stood at 129,44 at the time of current study. In view of the outbreak of pandemic several Governments, WHO, CDC etc have launched digital platforms for COVID related information and updates. According to the government website https://www.mygov.in/covid-19 people of all age groups can be the targets of corona infection. However, children, senior citizens and individuals with prior health issues like asthma, Diabetes etc are at a higher risk due to their weaker immune system [4]. The virus is completely resistant to cold weather conditions. Even the high temperature produced by hand dryers is not effective in killing the viral cells. The major symptoms to identify/associated with the infection are dry cough, headache, high body temperature. The viral incubation period is of 1-14 days [5]. The global death rate for COVID 19 is 4%. It is a highly contagious virus which can easily spread through saliva, air droplets generated during cough and other body fluids.

Coronaviruses are a group of single stranded RNA viruses and contain club shaped spikes projecting from the viral surface [6]. These are a varied group of viruses belonging to the order Nidovirales. They have 4 different groups [7]. They are known to infect ether animals or humans. Several species are known for their pathogenic effect on cats,
Dogs etc. [8]. In order to increase the awareness of the safety measures and cut down the spread of COVID 19 several websites are exploring the Do’s and Not’s for the Pandemic [14]

Materials and Methods

1. Gene and Genome Sequence Retrieval of Novel CoV19:
Novel CoVid 19 being new to the world all the gene and protein sequences of the isolates are not readily available in the publically accessible databases. However, GenBank [9] of NCBI has released the genomes of some novel CoVid 19 isolates. the Complete Genome of this virus was analyzed with special focus on the M protein coding regions.

2. BLAST Analysis:
The M protein obtained is first compared with the 25 different organisms selected for the study. These organisms included both Coronavirus group and another non-corona group. In addition to complete genome of this virus, the M protein coding gene sequences was also subjected for the comparison with the above 25 organisms individually. Nucleotide BLAST [10] (BLASTN) was used for the comparison for all the gene sequences. BLASTP was also used in the study, for the Identification of Translated protein product of MERS gene region.

3. TRANSLATE:
TRANSLATE is a tool from EXPASY server which can produce translated protein products for the input DNA sequences using different frames for translation. MERS CoV gene region showing similarity with the Novel CoV19 was analyzed using this tool to identify the possible protein product. The protein product obtained was further subjected for BLASTP analysis to identify the protein coded by the sequence.

4. SMART Analysis:
It is a tool from EMBL database. SMART results are completely based on Published research and recorded database. This tool is used for the Protein sequence annotation providing in depth analysis of the input protein sequence. The tool enables the identification of Functional Domains along with their location and function, Signal Peptides conserved or regions of importance.

5. CLUSTAL OMEGA:
CLUSTAL OMEGA tool from EMBL was used to evaluate the phylogenetic relation among the selected organisms based on MSA and Neighbor joining method.

Results and Discussion

1. The table 1 shows the comparative analysis between of M Protein and complete genome of MN908947.3 Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome- 2019, Isolated from China with 25 selected organisms.

| S. No | Protein Tested | Test Organisms Selected for the Study | Identity | Remarks |
|-------|----------------|--------------------------------------|----------|---------|
| 1     | Complete Genome M protein | SARS-CoV 2 isolate 2019 nCoV/Italy | Whole Genome not available | Shares identity to the M protein |
|       |                 |                                      | 595/595  100% Identical |         |
| 2     | Complete Genome M protein | SARS-CoV-2/Hu/DP/Kng/19-027 RNA, complete genome | Greater Identity | Genetically close |
|       |                 |                                      | 595/595  100% Identical |         |
| 3     | Complete Genome M protein | SARS-CoV-2/IQTC03/human/2020/CHN, complete genome | Greater Identity | Genetically close |
|       |                 |                                      | 595/595  100% Identical |         |
| 4     | Complete Genome M protein | SARS-CoV2 isolate 2019-nCoV/USA-CA9/2020, complete genome | Greater Identity | Genetically close |
|       |                 |                                      | 595/595  100% Identical |         |
| 5     | Complete Genome M protein | SARS-CoV-2/WH-09/human/2020/CHN, complete genome | Greater Identity | Genetically close |
|       |                 |                                      | 595/595 |         |
|   | Complete Genome | M protein | Genetic Identity | Genetically Close |
|---|----------------|-----------|------------------|-------------------|
| 6 | SARS-CoV-2/Yunnan-01/human/2020/CHN, complete genome | Greater Identity | Genetically close |
| 7 | Wuhan seafood market pneumonia virus genome assembly, chromosome: whole genome | Greater Identity | Genetically close |
| 8 | SARS-CoV2019 nCoV/Italy-INMI1 membrane glycoprotein (M) gene, partial CDS | Whole genome is not available | Shares identity to the M protein |
| 9 | Bat coronavirus RaTG13, complete genome | Greater Identity | Genetically close |
| 10 | Bat SARS-like coronavirus isolate bat-SL-CoVZXC21, complete genome | Greater Identity | Genetically close |
| 11 | Bat SARS-like coronavirus isolate bat-SL-CoVZC45, complete genome | Greater Identity | Genetically close |
| 12 | Pangolin coronavirus isolate MP789 genomic sequence | Greater Identity | Genetically close |
| 13 | Bat coronavirus Cp/Yunnan2011, complete genome | Greater Identity | Genetically close |
| 14 | Bat SARS coronavirus Rp3, complete genome | Greater Identity | Genetically close |
| 15 | Human Influenza A Virus (taxid:11320) | No similarity | Genetically Diverse |
| 16 | human hepatitis C virus HCV (taxid:11103) | No similarity | Genetically Diverse |
| 17 | Mycoplasma peripneumoniae (taxid:2102) | No similarity | Genetically Diverse |
| 18 | Chlamydia pneumoniae | No similarity | Genetically Diverse |
| 19 | Bordetella pertussis | No similarity | Genetically Diverse |
|   | M protein | (taxid:520) |   |   |   |
|---|-----------|-------------|---|---|---|
| 20 | Complete Genome | Rhinoviruses (taxid:12059) | No similarity | Genetically Diverse |
|   | M protein |   |   |   | No similarity |
| 21 | Complete Genome | MERS coronavirus (taxid:1335626) | 767/767 | 70.58% Identity | Lower query coverage and 70.58% identity |
|   | M protein |   | No similarity to M protein |
| 22 | Complete Genome | HCoV-SARS (taxid:694009) | 100% identical 55208/55208 | Shares highest identity both genomic and M protein level |
|   | M protein |   | 100% Similarity (595/595) |
| 23 | Complete Genome | Human respiratory syncytial virus (taxid:11250) | No similarity | Genetically Diverse |
|   | M protein |   | No similarity |
| 24 | Complete Genome | Human immunodeficiency virus 1 HIV-1 (taxid:11676) | No similarity | Genetically Diverse |
|   | M protein |   | No similarity |
| 25 | Complete Genome | HCoV-229E (taxid:11137) | No Similarity | Genetically Diverse |
|   | M protein |   | No similarity to M protein |

**Table 1: Above Table Shows the Summary of Comparison between the Genome of Novel Coronavirus 2019 and 25 Selected Organisms**

The columns show that HCoV shares highest shares highest similarity with the query genome with query coverage of 99% and identity being 100%. Thus it is the most closely related virus to the Novel CoV19. Further the Middle East Respiratory Syndrome MERS genome shares only a localized gene region identity with query coverage of 14% an identity 70.58% identity. This conserved gene region is further annotated in the study.

**2. Phylogenetic Tree Using CLUSTAL OMEGA of the Above Sequences that Share Genome Similarity with Novel Coronavirus 2019**

The method used is neighbor joining method for the construction of tree.

![Fig 1: Showing the Phylogenetic Tree of Covid19 with the Other Group Under Comparison.](image)

Labels 1,2,3.. in the given Phylogenetic Tree indicates organism number as shown in the table 1. The prefix 0 indicates Query genome. The phylogenetic tree was constructed based on the Neighbor joining method and is shown as above figure.
3. BLAST Result Showing the Similarity of MERS Genome with Novel Coronavirus 2019

In the above Fig 2a and 2b the results of BLAST analysis are provided showing the genomic region of MERS that shares 14% query coverage and 70.58% identity with the Novel CoV19 genome.

4. EXPASY TRANSLATE [11] Result Showing the Translated Product of the Above Gene Region: 5prime to 3 prime frame 2

5. BLASTP Result of Above Protein
The following conclusion can be made based on the above results:
In contrast to the previous results related to the M protein when the whole genome of novel CoV19 (query) is compared with the whole genomes of the 25 organisms, both HCoV SARS and MERS Coronavirus show a greater degree of similarity.

6. Domain Analysis of ORF1ab:
In order to study the significance and function of ORF1ab SMART analysis was performed.

**Fig 4a:** Shows all the Domains present in the ORF1ab Polyprotein

It was found that the function of ORF 1 polyprotein 1ab is as follows:
The PP1ab - Orf1ab polyprotein is completely specific for Coronavirus group [13]. Thus, it can be a better drug target compared to M protein. It is a Multifunctional protein involved in the transcription and replication of viral RNAs. It contains the proteinases responsible for the cleavages of the polyprotein.

**Fig 4b:** Highlights Other Regions that are important in the Protein Sequence
(The above figure shows the other non-domain regions of importance)

**Fig 4c:** Coronavirus Proteases which are involved in Processing of Polyproteins
(RCSB PDB has declared Coronavirus proteases as molecule of the month)
Suggestible Therapy and Medication Based on Genome Similarity:
Based on the homology existing between the genomes of the above group and especially the ORF1ab gene having a major role in viral replication, it can be concluded that the medication and therapy used for MERS and other Corona group infections can be tried on the patients with COVID 19. Most important is to use certain compounds that can drug the ORF1ab gene of Coronavirus to inhibit its development.

Some of the Medications Suggestible for Treating COVID 19 According to our Hypothesis are:
Methylprednisolone, Kaletra (400 mg ritonavir and 100 mg lopinavir) a protease inhibitor [21], SKP2 inhibitors can also be used to disintegrate the viral genome.

According to the reports released on 24th March 2020, the ICMR recommends the use of Hydrochloroquine for the prophylaxis against COVID19 and also for the high risk health care workers [15].

Why Hydroxycholoquine?
The current study tried to evaluate the reason for the use of Hydroxychloquine against COVID19.

It was found that the same chemical was used for the treatment of Dengue Virus [16] as shown by Li-Fong Wang et.al. in March 2015.

Julie Dyall et.al in Dec 2018 [17] also detailed the use of this chemical in the treatment of MERS. The article states the use of both Hydroxychloroquine sulfate and Mefloquine for the treatment of both SARS and MERS infections.

It was also reported that the antimalarial agents, chloroquine (CQ), amodiaquine, and mefloquine have activity against SARS-CoV and MERS-CoV in vitro [18, 19, 20].

Further in the research of Li hualan 2020 [22], it was stated that Chloroquine would effectively bind to the ORF1ab there by preventing its binding with the beta chain of Hemoglobin. This would finally prevent the virus from causing respiratory illness. This would add on to the importance of targeting the ORF1ab Protein in the virus.

Conclusion
Current findings involve the comparison of gene encoding M protein of Novel Coronavirus 19 with the selected list of 25 closest organisms as per available literature. Our results reveal a high degree of M Protein similarity to all the Species of SARS and other Corona group of viruses. HCoV shares highest degree of similarity with the COVID genome with sequence query coverage of 99% and identity of 100%. However, it was noted that MERS do not share any similarity to the specified gene coding for M Protein. Further the complete genome of Novel CoV 19 was considered for Sequence analysis. The data revealed a specific region of MERS sharing identity greater than 70%. The study was extended to identify the protein coded by this region. EXPASY TRANSLATE was used to translate this gene region to its protein followed by BLAST analysis. The gene region was found to code for ORF1ab polyprotein of MERS. The ORF1ab poly protein is a Coronavirus specific protein having a major role in the replication of RNA and multiplication of viruses. What we can infer at this stage is that irrespective of the similarity for the whole genome MERS and Novel CoV share a considerable degree of identity for ORF1ab. Thus all the medication and therapy capable of fighting against MERS can be tried for Novel CoV19 victims. In support of the conclusion it was also reported that Chloroquine can be used to bind to the ORF1ab to prevent its efficacy in binding to the beta chain of hemoglobin. A brief data mining and analysis was performed related to the use of Hydroxychloroquine for fighting against COVID19. As shown in the study homology of the ORF1ab between MERS and COVID 19 suggests the use of same drugs for both the treatments. Thus hydroxycholquine can be used for COVID 19 as it was previously used for MERS and SARS. However the use of these medications for Novel CoV 19 is hypothesized only on the basis of previously available data and analysis, they have to be validated by clinical trials to confirm their efficacy. Any such medication shall only be considered for use only after consultation with qualified medical practitioner, their improper doses and combination has proved to be lethal in several patients.
Our work is extending towards Rational lead molecule identification, Vaccine screening and Drug Development studies using the current finding of ORF1ab domain.

References
1) “Facts about COVID-19”, European Centre for Disease Prevention and Control, March, 2020
2) “COVID-19 CORONAVIRUS PANDEMIC” Worldometer, March 2020, https://www.worldometers.info/coronavirus/
3) VINCE MCLEOD, “COVID-19: A History of Coronavirus, March 16: 2020; Website: Lab Manager, Lab Health and Safety
4) https://www.mygov.in/covid-19
5) Sanjaya Senanayake, “Coronavirus: how long does it take to get sick? How infectious is it? Will you always have a fever? COVID-19 basics explained”, The Conversation.com, March 20, 2020 http://theconversation.com/coronavirus-how-long-does-it-take-to-get-sick-how-infectious-is-it-will-you-always-have-a-fever-covid-19-basics-explained-132963
6) Anthony R. Fehr “Coronaviruses: An Overview of Their Replication and Pathogenesis”, Methods Mol Biol. 2015; 1282: 1–23. PMID: 25720466
7) He F1, Deng Y1, Li W2, “Coronavirus Disease 2019 (COVID-19): What we know?”, Journal of Medical Virology, 2020 Mar 14, PMID: 32170865
8) Amy Carotenuto, “Flagler Humane Society: Can the coronavirus affect animals?” Journalonline.com, March 3: 2020
9) https://www.ncbi.nlm.nih.gov/genbank/sars-cov-2-seqs/
10) Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410. PubMed
11) Gasteiger E., Gattiker A., Hoogland C., Ivanyi I., Appel R.D., Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis Nucleic Acids Res. 31:3784-3788(2003).
12) Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z, Woolsey J. Drugbank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res. 2006 Jan 1;34 (Database issue):D668-72. 16381955
13) Rachel L. Graham, Jennifer S. Sparks, Lance D. Eckerle, Amy C. Sims, Mark R. Denison, Virus Res. Author manuscript; available in PMC 2009 Feb 9. Published in final edited form as: Virus Res. 2008 Apr; 133(1): 88–100. Published online 2007 Mar 29. doi: 10.1016/j.virusres.2007.02.017
14) https://ieeecovidmove.in/
15) “Advisory on the use of Hydroxychloroquin as Prophylaxis for SARS CoV2 Infection” https://www.mohfw.gov.in/pdf/AdvisoryontheuseofHydroxychloroquinasprophylaxisforSARSCoV2infection.pdf
16) Li-Fong Wang et.al, “Hydroxychloroquine-Inhibited Dengue Virus Is Associated with Host Defense Machinery”, Journal of Interferon & Cytokine Research, 2015 Mar 1; 35(3): 143–156, PMID: 25321315
17) Julie Dyall et.al, “Middle East Respiratory Syndrome and Severe Acute Respiratory Syndrome: Current Therapeutic Options and Potential Targets for Novel Therapies”, Drugs. 2017 Dec; 77(18): 1935–1966., PMID: 29143192
18) de Wilde AH, Jochmans D, Posthuma CC, Zevenhoven-Dobbe JC, van Nieuwkoop S, Bestebroer TM, et al. Screening of an FDA-approved compound library identifies four small-molecule inhibitors of Middle East respiratory syndrome coronavirus replication in cell culture. Antimicrob Agents Chemother. 2014 Aug;58(8):4875–84
19) Dyall J, Coleman CM, Hart BJ, Venkataraman T, Holbrook MR, Kindrachuk J, et al. Repurposing of clinically developed drugs for treatment of Middle East respiratory syndrome coronavirus infection. Antimicrob Agents Chemother. 2014 Aug;58(8):4885–93.
20) Keyaerts E, Vlijgen L, Maes P, Neyts J, Van Ranst M. In vitro inhibition of severe acute respiratory syndrome coronavirus by chloroquine. Biochem Biophys Res Commun. 2004 Oct 8;323(1):264–8.
21) Tai DY, “Pharmacologic treatment of SARS: current knowledge and recommendations”, Annals of Academy of Medicine, Singapore, 2007 Jun;36 (6):438-43.
22) Wenzhong, Liu; Hualan, Li (2020): COVID-19: Attacks the 1-Beta Chain of Hemoglobin and Captures the Porphyrin to Inhibit Human Heme Metabolism. ChemRxiv. Preprint. https://doi.org/10.26434/chemrxiv.11938173.v4