International Journal of Academic Research and Development
www.academicjournal.in
ISSN: 2455-4197
Received: 29-08-2021, Accepted: 13-09-2021, Published: 29-09-2021
Volume 6, Issue 5, 2021, Page No. 33-39

Functional genome analysis and drug screening by bioinformatics methods in SARS-CoV-2

Sudheer Menon
Department of Bioinformatics, Bharathiar University, Coimbatore, Tamil Nadu, India

Abstract
Rapid order of sequence for virus species, analysis of functional genome and screening for effective drugs by bioinformatics methods are considered as significant tasks. Some bioinformatics methods for instance, sequence alignment, motif recognition, ORF identification, secondary and tertiary structure prediction, the entire genome of SARS-CoV-2 was completely analyzed. To discover effective drugs, the boundaries of binding sites were determined by SeeSAR. Moreover, potential miRNAs were anticipated by RNA base-pairing. After anticipating by using WebMGA, NCBI and GeneMark and comparison, an aggregrate of 8 credible ORFs were recognized. Even the entire genome have incredible difference with other CoVs, each ORF has high homology with SARS-CoVs (>90%). Moreover, domain creation in each ORFs was almost similar to SARS. In the DrugBank database, just 7 potential drugs were screened dependent on the sequence search module. Further anticipated restricting boundaries between drug and ORFs uncovered that 2-(N-Morpholino)-ethanesulfonic acid could bind 1# ORF in 4 distinct regions ideally. In the meantime, both benzyl (2-oxopropyl) carbamate and 4-(dimehylamina) benzoic acid have been exhibted to inhibit SARS-CoV disease effectively. Strangely, 2 miRNAs (miR-1307-3p and miR-3613-5p) were anticipated to prevent virus replication by means of focusing 3′-UTR of the genome or as biomarkers. After taking everything into account, the novel COVID-19 might have association with SARS. Drugs which are used to treat SARS may also be proved as a successful against the novel virus. Furthermore, changing miRNA expression might turn into a potential therapeutic schedule.

Keywords: functional genome analysis, drug screening, SARS-CoV-2

Introduction

The recent development of a novel COVID (SARS-CoV-2) caused an outbreak of surprising viral pneumonia in Wuhan, a central city of China and then spread out all around the world. CoVs are common microbes with highly infectious, from humans to animals like mouse, bat, livestock, avian and many other wild animals can be infected in the gastrointestinal, hepatic, respiratory, and central nervous system. Since the outbreaks of the serious acute respiratory syndrome (SARS) in 2002 and the Middle East respiratory syndrome (MERS) in 2012, the chance of COVID transmission from animal to human has been demonstrated. Notably, by means of deep sequencing and etiological examinations, SARS-CoV-2 has been recognized as a novel Covid like SARS-CoV. CoVs are inter-related to the subfamily Coronavirinae in the family of Coronaviridae of the order Nidovirales. The single-stranded genome of CoVs is a positive-sense RNA with 5′-cap structure and 3′-poly-A tail. The genomic RNA is utilized as a template to translate polyprotein 1a/1 ab, the non-structural proteins (nsps) to form a replication-transcription complex in double-membrane vesicles. Subsequently, a set of subgenomic RNAs (sgRNAs) are integrated by RTC in a discontinuous transcription way.

Genomes and subgenomes of Covid contain minimum 6 open reading frames (ORFs). The primary ORF (ORF1a/b), which is about 2/3 of genome length, encodes 16 non-structural proteins (nsp1-16) in general. Such polypeptides will be functioned into 16 nsps by virally encoded protease. Hydrophobic transmembrane spaces are present in nsp3, nsp4, and nsp6 to secure the nascent pp1a/pp1ab polyproteins to membranes once RTC arranges. Other ORFs on the 1/3 genome closer to 3′ terminus encodes minimum 4 main structural proteins: named as spike (S), membrane (M), envelope (E), and nucleocapsid (N) proteins. Other than these 4 main structural proteins, distinctive CoVs encode specific structural and frill proteins, for instance 3a/b protein. Every structural and frill proteins are translated from the sgRNAs RNAs of CoVs. Moreover, a 5′ untranslated region (UTR) and 3′-UTR were also distinguished in the SARS-CoV-2 genome. Subsequently, studies about microRNA might be essential and critical.

Beside these, various cellular proteins have been displayed to interact with CoVs RNA. These are heterogeneous nuclear ribonucleoprotein A1, polypyrimidine tract binding protein, poly (A)-binding protein, and the last one is mitochondrial aconitase. Understanding of the genome-structure-function relationship in SARS-CoV-2 is significant for the identification of all possible anti-viral inhibitors and immunization targets. Recent fast progress in sequencing technologies, bioinformatics methodologies and tools has empowered a more in-depth view of the structure and working of viral networks, supporting the characterization of arising viruses. Bioinformatics analysis of viruses includes the overall tasks related to any sequences analysis, including the distinguishing ORFs, homology searching, gene functional prediction, sequence arrangement, and motif and epitope acknowledgement.

The anticipation for features such as trans membrane domains, protein secondary and tertiary structure are significant for analyzing the structure-working relationship of viral proteins encoding. Biochemical pathway analysis can assist with demonstrating information at the biological frameworks level. Virus-related bioinformatics databases incorporate those inter-connected with taxonomy, homologous protein families, viral sequences, structures, or

committed to explicit viruses such as flu. These computational projects give a resource for genomics and proteomics studies in virology research and are valuable for understanding viral diseases, as well as for the design and advancement of anti-viral agents.

Methods and Materials

- **RNA sequencing an data adjustment**

The order of SARS-CoV-2s was acquired from NCBI, which was given by Dr. Zhang, an educator from Fudan University. Hence, the most common way of sequencing and information calibration should refer to Dr. Zhang's article.

Sequencing reads were first connector- and quality-managed using the Trimmomatic program. The excess reads were collected de novo utilizing both the Megahit and Trinity program with default boundary settings. To recognize possible aetiologic agents present in the sequenced information, the abundance of the assembled contigs was first assessed as the normal counts using the RSEM program executed in Trinity. Non-human reads, produced by filtering host reads utilizing the human genome by Bowtie2, were utilized for the RSEM bounty assessment.

- **Virus Genome Characterization and Investigation**

Understanding the structure-working relationship in viruses is significant for discovering expected antiviral inhibitors and vaccine targets. Databases and bioinformatics methods that contain genomic, proteomic, and functional data have become essential for virology studies. According to our examination, all used databases and techniques were listed in Table given below. All parameter changes were referring to the references. Here, the cycles of some significant analyses would be presented.

Database and bioinformatics tools for virology studies

Table 1

Tool name	Function	URL
GeneMark	ORF identification	http://opal.biology.gatech.edu/GeneMark/genemarks.cgi
ORF Finder	ORF identification	http://www.ncbi.nlm.nih.gov/gorf/gorf.html
BLAST	Homology searching	http://www.ncbi.nlm.nih.gov/BLAST/
SMART	Pattern/motif recognition	http://smart.embl-heidelberg.de/
IEDB	Epitope analysis	http://www.immuneepitope.org
SPLIT 4.0	Protein secondary structure prediction	http://split.pmfst.hr/split/4/
SWISS-MODEL	3-D structure modeling	http://swissmodel.expasy.org/
I-TASSER	3-D structure modeling	https://zhanglab.ccmb.med.umich.edu/
Drugbank	Drug prediction	https://www.drugbank.ca/
SeeSAR	Drug prediction	https://www.biosolveit.de/SeeSAR/

GeneMark, inside an iterative Hidden Markov model based calculation, the accuracy of gene begin prediction can be improved by consolidating models of protein-coding and non-coding regions and models of administrative sites closer to gene start. It can be utilized for a recently sequenced prokaryotic genome prediction using a non-supervised preparing system. After the ORF anticipation, the homologous comparison with the FASTA amino acid sequence was performed. In “Choose Search Set”, the non-excess protein sequences database was picked without organism and prohibited conditions. In “Program Selection”, the blastp algorithm was picked. After that, the distance tree was built with a quick least evolution strategy. The maximum sequence contrast is 0.85. Drug screening was executed for each ORF coding sequence in the Sequence Search module of DrugBank with the help of FASTA format sequence. The BLAST parameters were agreeable. The predicted value is 0.00001 and the reword for match is 1. The number of cost to open or extend a difference is -1, and the value of penalty for mismatch is -3 approximately. Perform gapped arrangement and filter question sequence should be checked properly. Then, at that point the molecular docking reenactments of proteins and drug molecules were performed by SeeSAR technique. The affinities, which are phys-chem properties, torsional ‘heat’ and explorable gap were determined to evaluate the chance of connection between protein and drug.

Result

- **Homology between SARS-CoV-2 and other CoVs**

To quickly understand the genomic trademark and decide the developmental relationships, genome homologous alignments with recently identified CoVs are important to perform. The entire genome sequence of SARS-CoV-2, provided by Dr. Zhang was kept in record by the National Center for Biotechnology Information (NCBI). We predicted phylogenetic trees dependent on the nucleotide groupings of the entire genome sequence.

The alignment result suggested that there is a significant difference between the whole genomes of SARS-CoV-2 and other CoVs. Even the most homologous species just have less than 90% repetitive sequence (Bat SARS-like coronavirus isolate bat-SL-CoVZC45 and CoVZXC21, complete genome). Furthermore, the result displayed that SARS and SARS-CoV-2 are distantly

Related (<82.34%). It may conclude that the new virus is not evolved directly from SARS, but we cannot deny that there is a potential relation between the two viruses.

As well known, when the virus infects the host, the hereditary material leaves the capsid followed by replication and then gathers. Since the presence of sub genome in COVID’s, the genome of new gathered CoVs may be changed. Address this issue; the phylogenetic trees of every primary ORF should be built. By anticipation of ORFs in 3 standard databases (NCBI, GeneMark and WebMGA), an aggregate of 15 ORFs were recognized.

Unexpectedly, at the position from 266 to 21555, each database anticipates a different result. What's more, the most significant non-structural protein ORF 1a/b was encoded dependent on this sequence. For making it sure, the accuracy of the subsequent examination, a total of 8 ORFs were anticipated by minimum two databases would become objects for further examination.

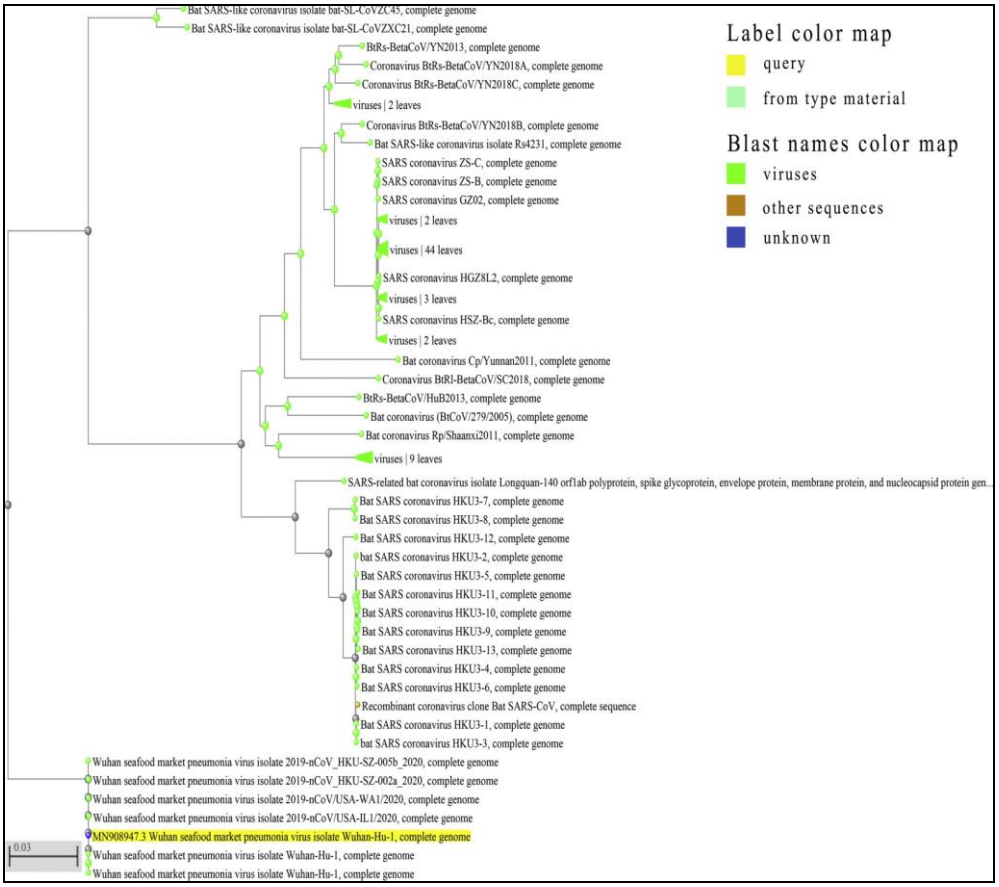

Fig 1: Phylogenetic tree of COVID's whole genome. The tree method is fast minimum evolution and the maximum sequencing difference is 0.75.

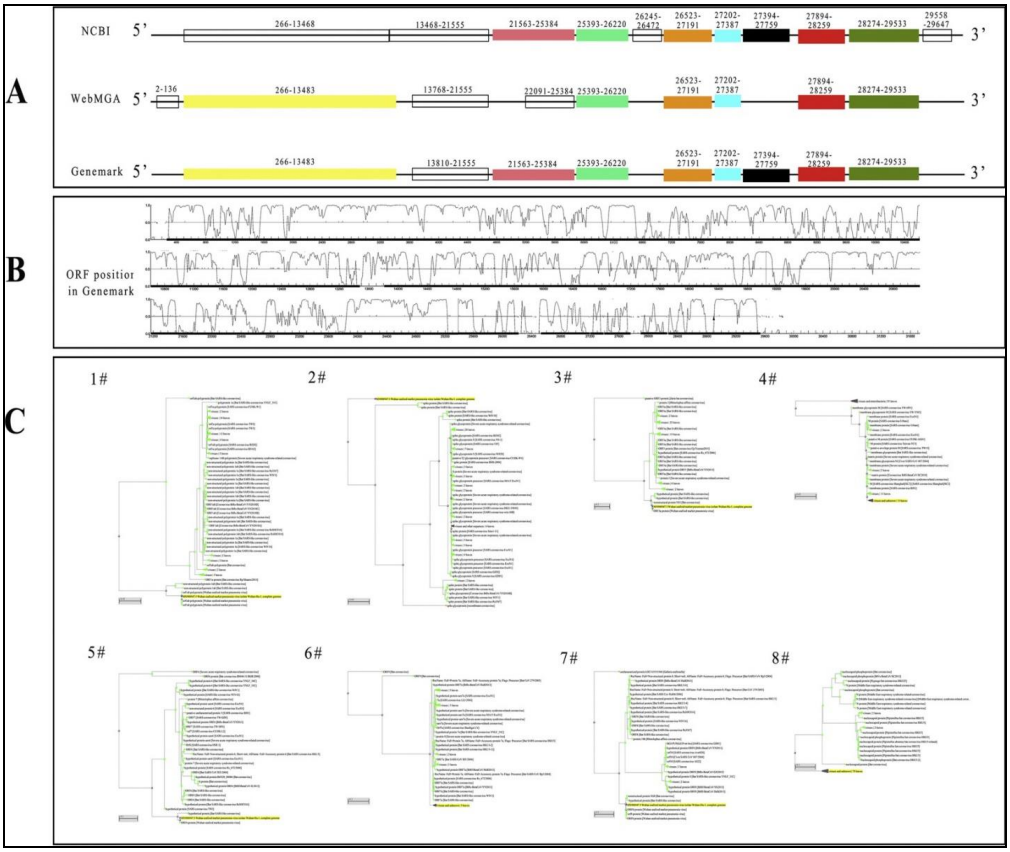

Fig 2: All the anticipated ORFs in SARS-CoV-2. **(A)** The sketched map of all 6 ORFs in the genome. The whole genome was anticipated by three databases which are NCBI, WebMGA and Genemark databases, respectively. The color boxes are representing various ORFs appeared in minimum 2 databases. The hollow boxes are representing the ORFs which just came on screen one time. **(B)** The peak value of the ORF in Genemark. The point where the fluctuation occurs presents ORF. **(C)** The phylogenetic trees of all ORFs. The tree method is quick minimum evolution and the maximum sequencing gap is 0.75. The yellow item is representing each specific ORF in SARS-CoV-2.

- **Structural and Functional ORF**

As indicated by the outcome of phylogenetic trees, it was tracked down that each ORF was similar to the particular genes in other CoVs like SARS, yet the gaps are real. To confirm the structure and working of polypeptide chains encoded by each ORF, the secondary and tertiary structure of polypeptide were anticipated (Fig. 3A, B and S1).

1. In 1#ORF: Total 21 positions of transmembrane helix were located (643–657; 700–716; 1723–1739; 1758–1773; 2219–2235; 2276–2310; 2379–2409; 2414–2438; 2822–2846; 3090–3118; 3127–3152; 3172–3202; 3630–3658; 3662–3676; 3684–3711; 3732–3750; 3758–3772; 3794–3822; 3826–3853; 3916–3932; 3961–3976)
2. Total 3 helixes in 2#ORF (50–74; 748–762; 1275–1299)
3. Same in 3#ORF, total 3 helixes in (85–112; 126–151; 156–180)
4. 4 helixes in 4#ORF (67–89; 97–121; 127–152; 184–200)
5. 1 helix in 5#ORF (54–74)
6. 2 helixes in 6#ORF (52–67; 146–167)
7. 1 helix in 7#ORF (52–67)
8. In 8#ORF, no helix was detected.

The polypeptides with transmembrane helix might play a part in virus disease and the transmembrane spaces can be identified by the immune system. They are acceptable candidates for incorporation in viral immunizations. In this manner, in view of the recorded protein templates in the RCSB PDB database, homology modeling of all polypeptide was performed in SWISS-MODEL.

Especially in the demonstrating of 3# polypeptide, 3 subunits collect into a protein named Spike protein also known as S, in other CoVs. S protein was frequently viewed as the most significant structural element plays an important role in antigen acknowledgement. At the point where homology modeling of these 8 polypeptides, just 1#, 2#, 6# and 8# polypeptide could be coordinated with similar templates (seq identity＞90%) while the others can't coordinate with a template with credible identity (＞30%). The outcome might signify 3#, 4#, 5# and 7# polypeptides may overlay into novel proteins.

The role of polypeptide and mature protein not just relies upon secondary structure and tertiary structure. The domains established by specific amino acid sequences are also pivotal. In this way, the fundamental domains in each ORFs were anticipated via the SMART database (Fig. 3C and Table 2). In the predicted outcome, it was recognized that a few domains play main roles in entry into the host cell like Spike_rec_bind domain in 2# ORF. When CoVs infected host, immune reaction showed up. Hence, the estimation of epitopes, the pieces of antigens collaborating with receptors of the immune system are significant for understanding viral diseases and discovering anti-viral targets. The top3 B cell-linear and irregular epitopes of every polypeptide have been shown in figure S2-9.

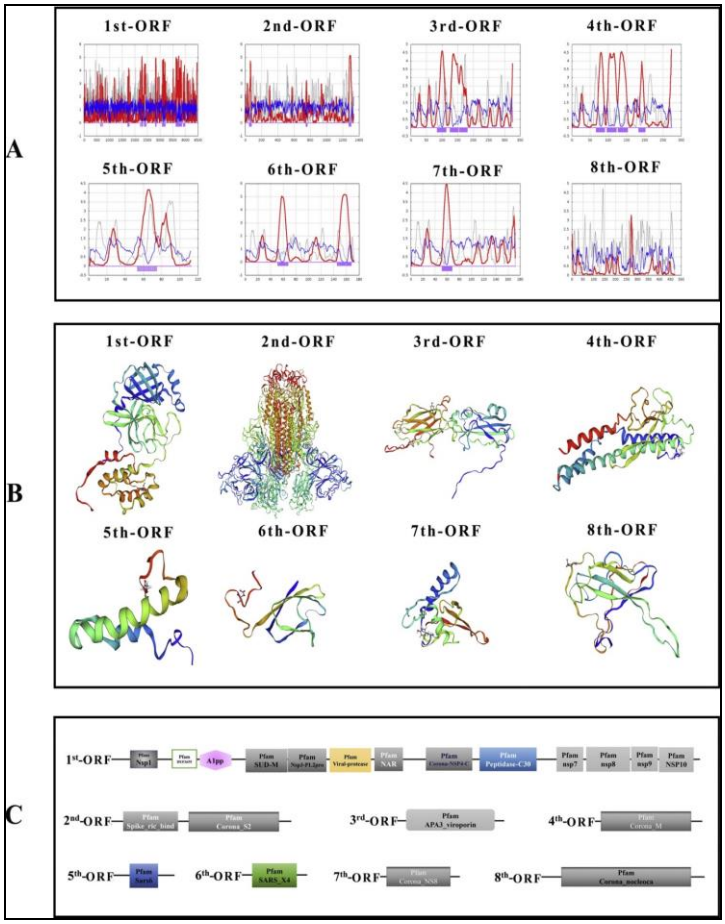

Fig 3: Structural and functional anticipation of all ORFs. **(A)** The secondary structure of all polypeptides. The red line is presenting the transmembrane helix preference; the blue line represents beta preference; the gray line represents the modified hydrophobic moment index (INDA index); the violet boxes represent predicted transmembrane helix position (DIG index). **(B)** The tertiary structure of each polypeptide (3D view). C, Domain of each ORF.

Table to Represent the Function of Each Domain in ORFs.

Table 2

No. ORF	Domain name	Function	Reference or GO item
1#ORF	Nsp1	Mediate RNA replication and processing	
1#ORF	DUF3655	Identifies the N terminus of Nsp3	N/A
1#ORF	A1pp	Bind ADP-ribose	
1#ORF	SUD-M	Identifies Nsp3	
1#ORF	Nsp3_PL2pro	cysteine-type endopeptidase activity	GO:0004197
1#ORF	Viral_protease	proteolytic processing of the replicase polyprotein, transferase activity, cysteine-type endopeptidase activity, omega peptidase activity	GO:0016740, GO:0004197, GO:0008242
1#ORF	NAR	nucleic acid binding	GO:0003676
1#ORF	Corona_NSP4_C	involved in protein-protein interactions	
1#ORF	Peptidase_C30	viral protein processing	GO:0019082
1#ORF	Nsp7	transferase activity, cysteine-type endopeptidase activity, omega peptidase activity	GO:0016740, GO:0004197, GO:0008242
1#ORF	Nsp8	cysteine-type endopeptidase activity, transferase activity, omega peptidase activity	GO:0004197, GO:0016740, GO:0008242
1#ORF	Nsp9	viral genome replication, RNA binding	GO:0019079, GO:0003723
1#ORF	Nsp10	viral genome replication, zinc ion binding, RNA binding	GO:0019079, GO:0008270, GO:0003723
2#ORF	Spike_rec_bind	aids viral entry into the host cell	
2#ORF	Corona_S2	receptor-mediated virion attachment to host cell, membrane fusion an integral component of membrane, viral envelope	GO:0046813, GO:0061025, GO:0016021, GO:0019031
3#ORF	APA3_viroporin	modulate virus release	
4#ORF	Corona_M	implicated in virus assembly, viral life cycle	GO:0019058
5#ORF	Sars6	42 to 63 amino acids, uncharacterised	N/A
6#ORF	SARS_X4	binding activity to integrin I domains	
7#ORF	Corona_NS8	typically between 39 and 121 amino acids, uncharacterised	N/A
8#ORF	Corona_nucleoca	viral nucleocapsid	GO:0019013

Targeted Drug Prediction

Though the genome role was detected and the process of CoVs multiplication has been notable. The screening of medicable drugs had been an issue. To accomplish the reason of treatment in the most limited time, screening existing drugs is more applicable than designing new ones. As indicated by amino acid sequence grouping result of every polypeptide, just 1# polypeptide could be bound with the 7 discovered drugs that given in figure S11.

Through determined affinities, phys-chem properties, torsional 'heat' and explorable space, an exhaustive analysis of the limiting capacity of every drug with a particular binding region was performed as shown in figure 4. It was shown that (2-(N-morpholino)-ethanesulfonic acid could tie with No. 1, 2, 3 and 5 regions with a bunch of desired boundaries. 4-(dimethylamino) benzoic acid and benzyl (2-oxopropyl) carbamate both could tie with all areas even without ideal boundaries. All 7 drugs were attempted to be utilized for relieving SARS, yet the efficacy and mechanism are still unsure.

Figure 4: 3D view of every drug binding site within 1# polypeptide. The blue line is representing the peptide chain. In the center of each little diagram is the drug molecule. The blank space means that the drug cannot bind with the peptide chain. The dotted lines represent intermolecular forces.

In addition to recognized ORFs, in the entire genome, there are 2 UTRs situated at 5′ end and 3′ end respectively. Also known, the multiplication of viruses frequently requires getting the transcription and translation systems within host cells. Accordingly, it was assumed that miRNAs in host cells may bind to the virus which brings about RNA debasement and inhibited translation.

In the miRBase database, through grouping anticipation, miR-1307-3p and miR-3613-5p might bind to 3′-UTR of SARS-CoV-2s. Unexpectedly, recent studies have distinguished in lung cancer patients; both miR-1307-3p and miR-3613-5p were down regulated altogether. Moreover, it was shown tyrosine kinase inhibitors (TKI) could upregulate miR-1307-3p which were utilized to treat non-small cellular lung cancer.Moreover, both 2 miRNAs cannot just become drug target, they can also be viewed as biomarkers in checking of viral pneumonia.

Discussion

Generally, virus emergence frequently prompts to significant outbreaks. In December 2019, an epidemic occurred by SARS-CoV-2 has guaranteed countless lives in China. Confronted with the major situation, a comprehensive examination of the SARS-CoV-2 in a brief time was imperative. In the current review, it was confirmed that main bat-SL-CoVZC45 and CoVZXC21 have more homology with SARS-CoV-2 (<90%). But, every ORF has higher homology with SARS after analyzed. Strangely, the anticipated outcomes of ORF were distinctive like 266bp-21555bp region which encode the main non-structural protein of CoVs as shown in figure 2A. In case there are 3 encode manners, actually, the genome might generate 5 unique species which will incredible difficulties to the treatment.

The entire process of novel drug growth to clinical impact tests will invest a lot of energy. Subsequently, if there some current drugs could be verified therapeutic, it will incredibly improve the prognosis of patients. An aggregate of 7 drugs were anticipated to bind with 1# polypeptide. Outstandingly, the 1# polypeptide could shape the replicate polyprotein 1a

(ORF1a). Generally, ORF1a and ORF1b (transcriptase) regularly form a heterodimer (ORF1ab) which might serve particular roles in destructiveness, virus-cell interactions and adjustments of defense reaction. In ORF1a, the detailed examination revealed evidence of versatile mutations exceeded structural proteins. For instance, the ORF1a polyprotein showed a pace of non- synonymous replacements similar to that in the S gene.

Conclusion

In between the screened drugs, the 2-(N-morpholino)-ethanesulfonic acid has the ideal binding efficiency with ORF1a with the help of the 3D structural fitting. Unfortunately, this hypothesis needs sufficient proof. Beside this, 2-(N-Morpholino)-ethanesulfonic acid is known as a part of

Crystallization buffer. Such buffer partiles including Tris, Hepas, glycerinum, DMSO and water typically showup in crystal structures however make no sense for drug discovery of virus.

In this way, however the anticipated consequences of 2-(N-Morpholino)-ethanesulfonic acid were great, it doesn't mean 2-(N-Morpholino)-ethanesulfonic acid own powerful virus suppression. Conversely, both benzyl (2-oxopropyl) carbamate and 4-(dimehylamina) benzoic acid have bene exhibited to inhibit SARS-CoV infection effectively. Interestingly, the 2-[(2,4-dichloro-5-methylphenyl) sulfonyl] -1,3-dinitro-5-(trifluoromrthyl) benzene couldn't bind with given region in 1# polypeptide which might be caused by 2 reasons

1. It has greater molecular weight and a more convoluted structure, results in a lower affinity with the polypeptide.
2. There are 2 chlorine particles in this drug which may lead to a weakened identified ability between drug and peptide chain.

Eventually, the host defense against viral infection is dependent on the individual defense mechanism system. That is why; the determination of viral antigens is also very significant, which is also a prerequisite for complete vaccine development. In our observation, the B cell epitope of every peptide chain has been anticipated for a better understanding of the infection process of SARS-CoV-2.

References

1. Yang X. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir. Med*,2020:8:475-481.
2. Sbrissa D, Naisan G, Ikonomov OC, Shisheva A. Apilimod, a candidate anticancer therapeutic, arrests not only PtdIns (3,5)P_2 but also PtdIns5P synthesis by PIKfyve and induces bafilomycin A1-reversible aberrant endomembrane dilation. *PLoS ONE* 13, e0204532, 2018.
3. Billich A. Drug evaluation: apilimod, an oral IL-12/IL-23 inhibitor for the treatment of autoimmune diseases and common variable immunodeficiency. *IDrugs*,2007:10:53-59.
4. Cai X. PIKfyve, a class III PI kinase, is the target of the small molecular IL-12/IL-23 inhibitor apilimod and a player in Toll-like receptor signaling. *Chem. Biol*,2013:20:912-921.
5. Sudheer Menon. "Preparation and computational analysis of Bisulphite sequencing in Germfree Mice" *International Journal for Science and Advance Research In Technology*,2020:6(9):557-565.
6. Sudheer Menon, Shanmughavel Piramanayakam and Gopal Agarwal. "Computational identification of promoter regions in prokaryotes and Eukaryotes" EPRA International Journal of Agriculture and Rural Economic Research (ARER),2021:9(7)21-28.
7. Sudheer Menon. "Bioinformatics approaches to understand gene looping in human genome" EPRA International Journal of Research & Development (IJRD),2021:6(7):170-173.
8. Sudheer Menon. "Insilico analysis of terpenoids in Saccharomyces Cerevisiae" International Journal of Engineering Applied Sciences and Technology,2021:6(1). ISSN No. 2455-2143, PP(43-52).
9. Thyaga BPC, Gowda S, Patil S, Srikantiah C, Suresh KP. Future trends of COVID-19 disease outbreak in different states in India: a periodic regression analysis. Highlights Biosci, 2020, 3.
10. Vetter P, Fischer WA, Schibler M, Jacobs M, Bausch DG, Kaiser L. Ebola virus shedding and transmission: review of current evidence. J Infect Dis,2016:214:S177-S184. pmid:27443613
11. Girard MP, Tam JS, Assossou OM, Kieny MP. The 2009 A (H1N1) influenza virus pandemic: A review. Vaccine,2010:28:4895-4902. pmid:20553769
12. Josset L, Menachery VD, Gralinski LE, Agnihothram S, Sova P, Carter VS et al. Cell host response to infection with novel human coronavirus EMC predicts potential antivirals and important differences with SARS coronavirus. MBio,2013:4:e00165-13. pmid:23631916
13. Peiris JS. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet*,2003:361:1319-1325.
14. Sudheer Menon. "Computational analysis of Histone modification and TFBs that mediates gene looping" Bioinformatics, Pharmaceutical, and Chemical Sciences (RJLBPCS),2021:7(3):53-70.
15. Sudheer Menon Shanmughavel Piramanayakam, Gopal Prasad Agarwal. "FPMD-Fungal promoter motif database: A database for the Promoter motifs regions in fungal genomes" EPRA International Journal of Multidisciplinary research,2021:7(7):620-623.
16. Sudheer Menon, Shanmughavel Piramanayakam, Gopal Agarwal. Computational Identification of promoter regions in fungal genomes, International Journal of Advance Research, Ideas and Innovations in Technology,2021:7(4):908-914.
17. Sudheer Menon, Vincent Chi Hang Lui, Paul Kwong Hang Tam. Bioinformatics methods for identifying hirschsprung disease genes, International Journal for Research in Applied Science & Engineering Technology (IJRASET),2021:9(VII):2974-2978.
18. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med*,2012:367:1814-1820.
19. Wang LF, Shi Z, Zhang S, Field H, Daszak P, Eaton BT. Review of bats and SARS Emerg Infect Dis,2006:12(12):1834-1840

20. Ge XY, Li JL, Yang XL *et al.* Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor Nature,2013:503(7477):535-538.
21. Chen Y, Guo D. Molecular mechanisms of coronavirus RNA capping and methylation Virol Sin,2016:31(1):3-11.
22. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses,2019:17(3):181-192.
23. Sudheer Menon. Bioinformatics approaches to understand the role of African genetic diversity in disease, International Journal of Multidisciplinary Research In Science, Engineering and Technology (IJMRSET),2021:4(8):1707-1713.
24. Sudheer Menon. Comparison of High-Throughput Next generation sequencing data processing pipelines, International Research Journal of Modernization in Engineering Technology and Science (IRJMETS),2021:3(8):125-136.
25. Sudheer Menon. Evolutionary analysis of SARS-CoV-2 genome and protein insights the origin of the virus, Wuhan, International Journal of Creative Research Thoughts (IJCRT),2021:9(8):b696-b704.
26. Sudheer Menon, Vincent Chi Hang Lui, Paul Kwong Hang Tam. A step-by-step work flow of Single Cell RNA sequencing data analysis, International Journal for Scientific Research and Development (IJSRD),2021:9(6):1-13.
27. Cauchemez S, Van Kerkhove MD, Riley S, Donnelly CA, Fraser C, Ferguson NM. Transmission scenarios for Middle East Respiratory Syndrome Coronavirus (MERS-CoV) and how to tell them apart Euro Surveill :Bull Eur Sur Les Mal Transmissibles Eur Commun Dis Bull, 2013, 18(24).
28. Chen Y, Liu Q, Guo D. Emerging Coronaviruses: Genome Structure, Replication, and Pathogenesis, 2020.
29. Sudheer Menon. Computational characterization of Transcription End sites in Human Genome, International Journal of All Research Education and Scientific Methods (IJRESM),2021:9(8):1043-1048.
30. Sudheer Sivasankaran Menon, Shanmughavel Piramanayakam. Insilico prediction of gyr A and gyr B in *Escherichia coli* insights the DNA-Protein interaction in prokaryotes, International Journal of Multidisciplinary Research and Growth Evaluation, (IJMRD),2021:2(4):709-714.
31. Sudheer Menon, Vincent Chi Hang Lui, Paul Kwong Hang Tam. Bioinformatics tools and methods to analyze single cell RNA sequencing data, International Journal of Innovative Science and Research Technology, (IJISRT),2021:6(8):282-288.
32. Sudheer Menon. Computational genome analysis for identifying Biliary Atresia genes, International Journal of Biotechnology and Microbiology, (IJBM),2021:3(2):29-33.
33. Goldman JD. Remdesivir for 5 or 10 days in patients with severe Covid-19. *N. Engl. J. Med*,2020:383:1827-1837.
34. Hung IF. Triple combination of interferon beta-1b, lopinavir-ritonavir, and ribavirin in the treatment of patients admitted to hospital with COVID-19: an open-label, randomised, phase 2 trial. *Lancet*,2020:395:1695-1704.
35. Riva L. Discovery of SARS-CoV-2 antiviral drugs through large-scale compound repurposing. *Nature*,2020:586:113-119.
36. Gopal M, Padayatchi N, Metcalfe JZ, O'Donnell MR. Systematic review of clofazimine for the treatment of drug-resistant tuberculosis. *Int. J. Tuberc. Lung Dis*,2013:17:1001-1007.
37. Schaad-Lanyi Z, Dieterle W, Dubois JP, Theobald W, Vischer W. Pharmacokinetics of clofazimine in healthy volunteers. *Int. J. Lepr. Other Mycobact. Dis*,1987:55:9-15.
38. Sudheer Menon. Recent Insilco advancements in genome analysis and characteristics of SARS-Cov2. International Journal of Biology Research, (IJBR),2021:6(3):50-54.
39. Sudheer Menon. Bioinformatics methods for identifying Human disease genes, International Journal of Biology Sciences, (IJBR),2021:3(2):1-5.
40. Sudheer Menon. SARS-CoV-2 Genome structure and protein interaction map, insights to drug discovery, International Journal of Recent Scientific Research, (IJRSR),2021:12(8):42659-42665.
41. Sudheer Menon Insilico Insights to Mutational and Evolutionary aspects of SARS-Cov2, International Journal of Multidisciplinary Research and Development, (IJMRD),2021:8(8):167-172.
42. Clausen TM. SARS-CoV-2 infection depends on cellular heparan sulfate and ACE2. *Cell*,2020:83:1043-1057.
43. Lu L. Structure-based discovery of Middle East respiratory syndrome coronavirus fusion inhibitor. *Nat. Commun*,2014:5:3067.
44. Xia S. A pan-coronavirus fusion inhibitor targeting the HR1 domain of human coronavirus spike. *Sci. Adv*,2019:5:eaav4580.
45. Chen J. Structural basis for helicase-polymerase coupling in the SARS-CoV-2 replication–transcription complex. *Cell*,2020:182:1560-1573.
46. Chan JF-W. Simulation of the clinical and pathological manifestations of coronavirus disease (COVID-19) in golden Syrian hamster model: implications for disease pathogenesis and transmissibility. *Clin. Infect. Dis*,2019:71:2428-2446(2020).