International Journal of Scientific Engineering and Applied Science (IJSEAS) – Volume-10, Issue-12, December 2024
ISSN: 2395-3470
www.ijseas.com

Insilico Prediction of the Function and Pathogenicity of Non-Coding Variants in COVID-19 Patients

Sudheer Menon[1], Shanmughavel Piramanayagam[2], Gopal Prasad Agarwal[3], Amanda Ramer Tait[4], Michelle Ramsay[5], Scott Hazelhurst[6], Keiko Ozato[7] Vincent Chi Hang Lui [8], Paul Kwong Hang Tam [9]

1. Department of Public Health, College of Applied Medical Sciences in Al-Namas, the Bisha University, Saudi Arabia.
2. Department of Bioinformatics Bharathiar University, Coimbatore, India
3. Department of Biochemical Engineering and Biotechnology, Indian Institute of Technology, Hauz Khaz, New Delhi, India.
4. Food Technology department, University of Nebraska, Lincoln, United States of America.
5,6. Wits Bioinformatics, Witwatersrand University, Johannesburg, South Africa.
7. Keiko Ozato, NICHD, National Institutes of Health (NIH), United States of America.
8. Department of Surgery, Li Ka Shing Faculty of Medicine, the University of Hong Kong, Hong Kong.
9. Dr. Li Dak-Sum Research Centre, The University of Hong Kong – Karolinska Institutet Collaboration in Regenerative Medicine, the University of Hong Kong, Hong Kong.

Corresponding Author: Sudheer Menon

Abstract
According to the recent studies on the sequence analysis of severe acute respiratory syndrome coronavirus- 2 (SARS-CoV-2) isolates and genomes of the host from the human population, the host genomes and virus interaction will be discussed. The genetic variants that were found in the genomes of host and coronavirus are ACE2 and TMPRSS2 genes. The host non-coding RNAs and virus interaction rely upon relation to their regulatory roles in target genes and signaling pathways that are crucially related to SARS-CoV-2 infection. There is a clear contribution to the knowledge of COVID-19 pathogenesis and SARS-CoV-2 infection mechanism.
The majority of diseases associated with SNP (Single Nucleotide Polymorphism) are found in non-coding regions of the genome. Despite the fact of several vaccines have been developed against COVID-19, there is a still need for developing non-invasive biomarkers for the determination of disease severity. So, circulating RNAs have capable biomarkers for several diseases. There are three motifs in the genome of SARS-CoV-2 in the 50 UTR leader sequence. The change in the expression of the coding gene and long non-coding RNA was found in SARS-CoV-2 infected cells and tissues in the COVID-19 patients. For determining C0VID-19 severity in patients. In-silico data and clinical validation led to the determination of potential RNA or protein networks as a novel prediction biomarkers.
Keywords: SARS-CoV-2, COVID-19, host genome, viral genome variants, lncRNA,

Introduction

COVID-19 which is commonly known as Coronavirus 2019 is a pandemic infectious disease and has high mortality rates. COVID-19 is caused by SARS-CoV-2 which is called Severe Acute Respiratory Syndrome Coronavirus-2. SomeCOVID-19 diseases' variations in the progression and clinical symptoms have been found in the infected patients of SARS-CoV-2 (Guan et.al, 2020). here research studies about the genomic interaction of SARS-CoV-2 and host cells will be discussed. Especially, this research focus on determining and understanding the role of Single Nucleotide Polymorphisms (SNPs) and other genomic variants in the host genome and virus and also the non-coding RNAs.

Many in-silico methods are frequently used to identify the potential pathogenic non-coding variants in COVID-19 patients. Topical in-silico approaches use machine learning techniques that are directed on the previous data to anticipate the activity without any experimental input (Wells A et.al, 2019, Kircher M et.al, 2014, Abramov S et.al, 2021, Lee D et.al, 2015). In another way, experimental methods are much more advanced than in-silico approaches. One of the most favorable methods that use in-vivo ChIPseq (Chromatin Immunoprecipitation with Sequencing) to compute the effect of non-coding variants in allelic - imbalance of TF binding (Kasowski M et.al, 2013, McVicker G et.al, 2013, Cheng Z et.al, 2021, Baca S.C et.al, 2021, Gusev A et.al, 2018).

Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) proteins or nucleic acid interrelate in a method called host-pathogen interaction with the biomolecules. The crucial role of these highly mutated RNA viruses is the encroachment into the non-coding and coding RNAs of the host. It plays a vital role in the pathogenicity and severity in the host and also has an impact on the host's response during different stages of viral infection. (Demiry A et.al, 2021, Maranon D.G et.al, 2020).

Materials and Methods

Study Population and Clinical Trials

According to a study, there were 131 COVID-19 patients, and more than half of them were admitted to Tor Vergata University Hospital (68%) and others (32%)admitted to Bambino Gesu Children's Hospital of Rome. 114 out of 131 patients showed the clinical symptoms of COVID-19 after positive results of nasooropharyngeal swabs. Other subjects that were not diagnosed with COVID-19 were sent home after keeping a few days under observation. Most of the admitted patients approximately 81 were male age range between 6 to 92 years old, 50 were female age range between 2 to 93 years old, and 13 were children (Verdoni L et.al, 2020). All patients were categorized into four groups.

- Asymptomatic: In this category, those patients were kept in which clinical symptoms were absent.
- Mild: In this mild group, the presence of few symptoms in the patients was observed. But these patients were not requiring ventilation except in some cases.
- Moderate: These patients showed respiratory impairment symptoms that require non-invasive ventilation and Continous Positive Airway Pressure (CPAP) or Bilevel Positive Airway Pressure (BiPAP) cycles.
- Severe: In this severe group, patients showed critical symptoms of the disease such as respiratory failure. These patients require invasive ventilation and intensive care unit (ICU) admission.

Severe Acute Respiratory Syndrome Coronavirus-2 (SARS CoV-2) Genome Variants and Sequence

Single-stranded RNA viruses, such as Severe Acute Respiratory Syndrome Coronavirus-2 have a high mutation rate that causes their genome diversity and the appearance of the variant that helps the viruses to adapt to different environments (Duffy 2018). Coronaviruses other than RNA viruses have a proofreading function which is done by NSP14 exoribonuclease, which is much conserved and crucial to maintaining the viral genome replication (Robson et.al, 2020). To determine the viral genome and other RNA viruses, extensive sequence-based analysis has been done in Severe Acute Respiratory Syndrome Coronavirus-2 isolates (Coppee et.al, 2020, Isabel et.al, 2020, Parlikar et.al, 2020)

There are found similarities between the genome sequences of SARS-CoV-1 and SARS-CoV-2. mostly the variations that were observed in the SARS-CoV-2 are SNPs (Single Nucleotide Polymorphisms) and SNVs (Single Nucleotide Variants). These variations in the coding and non-coding regions of the genome are the main cause of its genetic diversity, virulence, transmissibility, and evolution (Khailany et.al, 2020). Recent studies showed that Severe Acute Respiratory Syndrome Coronavirus-2 genome variants affect the immune escape, clinical characteristics of COVID-19, and its infection rates. So, these variants could be used as epidemiological tools for tracking viruses and also control infections.

Genome coverage	Methodological approach	Main results	Reference
Whole genome	– Genome sequencing and alignment analysis of 7710 GISAID[a] sequences	– Average pairwise difference of 9.6 SNPs between any two genomes – Mutation rate of the global diversity of SARS-Cov2 of ~6 × 10-4 nucleotides/genome/year – 290 aminoacid alterations in the genomes: 232 synonymous and 58 non-synonymous mutations	Khailany et al., 2020
Whole genome	Analysis of SARS-CoV-2 sequences using CoV_GLUE (http://cov-glue.cvr.gla.ac.uk) of 9028 available sequences, including 4973 European sequences	– Divergence of the two main mutations (S-D614G and nsp12-P323L) from the NCBI (NC_045512) retrieved in all continents with only three cases in Asia – Mutations at ORF8-L84S and ORF3a-Q57H (as the third and fourth most frequent mutation, respectively) – Co-evolving of the L84S amino acid substitution with three other mutations: nsp4-F308Y, ORF3a-G196V and N-S197L	Coppée et al., 2020
Whole genome	Genome sequencing and alignment analysis of 94 Genbank genomes	– 156 variants and 116 unique variants across the genome (46 missense, 52 synonymous, 2 insertion, 1 deletion and 14 non-coding alleles) – C > T and or T > C as the most common variants in the ORF1ab (NSP1-NSP16), ORF8 and, N genes	van Dorp et al., 2020
Whole genome	Genome sequencing and alignment analysis of ~660 genomes- NCBI[b] virus database	– Mutations in the S protein (D614G, V483A, L5F, Q675H, H655Y, and S939F) – Substitutions at R203K and G204R in the N protein – Substitutions at L84S, V62L, and S24L in the ORF8 – Non-synonymous mutations in ORF3a (Q57H and G251V – Non-synonymous mutations in ORF1ab (T265I, P4715L, P5828L, and Y5865C)	Laha et al., 2020
Whole genome	RNA sequencing analysis of NCBI RNA-seq data	– A-to-G (59.1%) RNA modifications (caused by RNA deamination) – Non-A-to-G variations, G to A (22,4%) and others (18,5%) (caused by replication errors) – A-to-G alterations in the N (>de 40%), ORF1AB (~35%), S, M, E, ORF3A, ORF8, ORF7A, and ORFA6 genes	Li et al., 2020b
Whole genome	Genome sequencing and alignment analysis of 12,909 genomes/estimation of common ancestor (TMRCA[c]) and mutation rates	– Indication that COVID-19 might have originated earlier than and outside of Wuhan Seafood Market – The genetic polymorphism patterns, including the enrichment of specific haplotypes and the temporal allele frequency trajectories generated from infection clusters, are similar to those caused by evolutionary forces such as natural selection	Liu et al., 2020a
Whole genome	Genome sequencing and alignment analysis of 106 NCBI genomes	– Higher number of mutations in the S protein, Nsp1, RdRp and the ORF8 regions – 47 key point mutations/SNPs located along the entire genome sequence in isolates from 12 different countries – NSP1 and ORF8 as the two hot spots with mutations and deletions	Vankadari, 2020
Whole genome	Genome sequencing and alignment analysis of 167 sequences from 15 distinct geographical locations	– 290 sites with variations (S, M and N genes; orf1ab, orf3a, in the envelope protein-coding gene, orf6, orf8, orf7b and orf10) – 244/290 variants were of a nucleotide substitution (158 transitions and 86 transversions) – High similarity (>99.9%) amongst all locations	Parlikar et al., 2020
Whole genome	Genome sequencing and alignment analysis of 566 genomes from India compared to NCBI	– 933 substitutions, 2449 deletions and 2 insertions, in total 3384 unique point mutations: distributed in 100 clusters of mutations (mostly deletions); 1609 substitution, deletion and insertion point mutations, 64 SNPs in coding regions and 7 in 5′-UTR and 3′-UTR – Largest number of SNPs in coding regions of ORF1ab and Spike protein	Saha et al., 2020
Whole genome	Genome sequencing and alignment analysis of 86 GISAID genomes from 12 countries	– 3 deletions (2 ORF1ab polyprotein and one in the 3′ end of the genome) in the genomes from Japan, USA, and Australia – 42 missense mutations (non-structural and structural proteins): 29 in the ORF1ab polyprotein, 8 in the S glycoprotein, 1 in the matrix protein, and 4 in the nucleocapsid protein	Phan, 2020
Whole genome	Genome sequencing and alignment analysis of 30,366 genomes/software developed by the researcher's group (ODOTool[d])	– 11 variations, with the incidence of over 10% in the 30,366 isolates – 8 of these variations (C1059T, G11083T, C14408T, A23403G, G25563T, G28881A, G28882A, and G28883C) caused amino acid substitutions	Ugurel et al., 2020
Whole genome, D614G mutation (gene spike protein)	Statistical analysis of the D614G mutation of 2795 GISAID genomes from 55 countries	– Amino acid change from an aspartate to a glycine residue at position 614 (D614G)	Isabel et al., 2020

Genome coverage	Methodological approach	Main results	Reference
		– High frequency of the D614G mutation (87%) among Italians isolates – D614G clade report of 954 of 1449 (66%) European isolates and 1237 of 2795 (44%) worldwide isolates	
Whole genome, *ACE2* binding domain	Mutation analysis of 34 human and animal isolates	– 60% of nucleotide variations between human SARS-CoV-2 and bat RaTG13, can be attributed to C > U and U > C substitutions – An accumulation of C > U mutations was observed in SARS-CoV2 variants in the human population, suggesting a significant role in the evolution of the SARS-CoV-2 coronavirus	Matyášek and Kovařík, 2020
Whole genome, Spike protein	Genome sequencing and alignment analysis of 1,325 genomes and 1604 CDS[c] of spike proteins from NCBI database	– 1197 SNPs, classified in 782 clusters – 1604 CDS at the S protein – Two major phylogenyclades A and B with many subclades in the S protein of SARS-CoV-2 circulating worldwide – 23402A > G SNP in 48.2% (the most common)	Singh et al., 2020
Spike gene	Development of a bioinformatics pipeline for Spike amino acid variants-GISAID data	– A spike protein amino acid change at D614G – Association of the D614G variant with high levels of infectivity and viral loads	Korber et al., 2020
ORF8	Evolutionary analysis of ORF8: genetic diversity and genomic rearrangements	– The ORF8 is poorly conserved among coronaviruses with a small number of highly frequent lineages – Nonsense mutations and three main deletions in the ORF8 gene that either remove or significantly change the ORF8 protein, which suggests that SARS-CoV-2 can persist without a functional ORF8 protein	Pereira, 2020
Orf1a, Orf1b, ORF3a ORF6, ORF7a, ORF8, ORF10, S, E, M, N, Sum	Metatranscriptome sequencinganalysis of eight fluid bronchoalveolar lavage from 25 community-acquired pneumonia patients and 20 healthy controls (Wuhan, China)	– No specific polymorphism was described – The median number of intra-host variants (iSNVs) was 1–4 in SARS-CoV-2 infected patients – SARS-CoV-2 evolves *in vivo* after infection, which may affect its virulence, infectivity, and transmissibility	Shen et al., 2020
RdRp, S, and Nsp-2	Sanger sequencing of the NSP-2, NSP-12, and S genes for phylogenetic analysis of 7 cases from Iran	– NSP-2 sequences - highest similarity between Iranian and Wuhan (China) – RdRp and S gene sequences-highest similarity between Iranian and China and USA – No identified differences between Iranian isolates	Tabibzadeh et al., 2020
S, RdRP, RNA primase, nucleoprotein	Genotyping of 558 isolates worldwide	– Mutations in genes encoding the S proteins and RNA polymerase, RNA primase, and nucleoprotein – Classification of the SNPs into four major groups: single mutation in nsp6 (11083G > T) (115%), single mutation in ORF3a (26144G > T) (49%), single mutation in RNA polymerase (nsp8) (8782C > T, 28144 T > C) (140%), and double mutations in S protein and RNA polymerase: (241C > T, 3037C > T, 14408C > T, 23403A > G) (178%; 182%; 182%; 183%) – Predominance of co-mutations (241C > T, 3037C > T, 23403A > G) in isolates from Europe – Estimated transmission of SARS-CoV-2 of 14 generations since its first infection to humans in Dec 2019	Yin, 2020

Table 1: Genome Variants in SARS-CoV-2

Long Non-coding RNA and their Expressions in SARS-CoV-2 Infected Cells

These Non-coding RNAs are heterogeneous transcripts and consist of more than 200 nucleotides without having coding proteins potentials, that's why called Long Non-Coding RNA (lncRNA). These RNAs also interact with DNA, mRNA, Proteins, and chromatin and regulate the levels of the protein-coding gene at the transcription, post-transcription, and post-translation levels.lncRNA helps in stabilizing or destabilizing the mRNA when they interact with mRNA or proteins. When lncRNA interacts with proteins they help in splicing and also indirectly affect the level of mRNA. LncRNA has been widely used in the biological processes, transcription, mRNA splicing, immunity, functions, and inflammation by changing the levels of proteins (Statello L. et.al, 2021, Walther K. et.al, 2021, Yao R-W. et.al, 2019).
Hundreds of long non-coding RNAs' expressions have been found including uncharacterized new transcripts in SARS-CoV-2 infected cells (Turjya R.R. et.al, 2020, Devadoss D. et.al, 2021, Laha S. et.al, 2021, Mukherjee S. et.al, 2021) such as BALF and PBMC from the Coronavirus patients (Moazzam-Jazi, M. et.al, 2021, Cheng J. et.al, 2021, Taheri M. et.al, 2021).

Severe Acute Respiratory Syndrome Coronavirus-2 Susceptibility and Host Genome

Several factors cause the severity and susceptibility of COVID-19, such as smoking, obesity, and alcohol (Yamamoto et.al, 2021). The host genome also plays a vital role in the progression of COVID-19. For the replication of its RNA, SARS-CoV-2 attaches itself to host cells with the

help of its S-proteins (Spike Proteins) and enters the cells to use host machinery (Hoffmann et.al, 2020). ACE2 and TMPRSS2 genes and their variants have been identified as the main molecular markers that help in the genetic susceptibility and resistance to COVID-19.
ACE2 is a transmembrane glycoprotein that consists of about 805 amino acids and there are 256 missense Single Nucleotide Polymorphisms (SNPs). ACE2 gene expression depends on the biological sex and age of the subjects (Ovsyannikova et.al, 2020). TMPRSS2 gene is involved in the proteolytic activity of ACE2 and SARS-CoV-2 S-protein, which leads to viral penetration into the host cells (Torre-Fuentes et.al, 2020).

Discussions

As we know COVID-19 is a serious disease that causes severe acute respiratory syndrome (SARS). There is no treatment for this COVID-19 to date (V'Kovski P . et.al, 2020). COVID-19 is not just a health problem but also, socio-economic dimensions have been into, which entire globe has been affected. There is a crucial need for diagnostic tools or assays for the evaluation of the infection severity, as there are constant mutations taking place in the whole genome and spike proteins of the virus (Harvey, W, T. et.al, 2021).
The severity of Coronavirus infection is very critical, it also affects the mortality and spread rate (Munster V. et.al, 2020). By developing the new and non-invasive biomarkers, early detection and severity of infection could be detected at the very early stage leading to control of the virus spreading. So, there should also be markers for COVID-19 severity detection along with therapy and vaccine development (Zeng F. et.al, 2020).

In-silico data suggest that ACE2 gene variants in the structural region of the proteins have an effect on the binding of pathogens or on the increase of quantitative expression of ACE2 genes (Letko M. et.al, 2020, Yan R. et.al, 2020, Wang Q. et.al, 2020, Hussain M. et.al, 2020). ACE2 receptors have more capability of binding with the spike protein of SARS-CoV-2, approximately 10 times more than SARS-CoV-1 (Cai Y. et.al, 2020).

Conclusions

Clinical efforts and research have been made on the complicated mechanism of SARS-CoV-2 to better understand it. There are also many mechanisms and procedures to be known that are related to immune responses and the pathogenicity of viruses. To define patients' risk groups, both host and viral genome contributed along with the role of non-coding RNAs. With the help of simple and specific tools, we predicted the early COVID-19 disease severity in asymptomatic and symptomatic patients. Studies suggested that there is no such evidence of constant relation of ACE2 gene variants with SARS-CoV-2 susceptibility.

References

1. Abplanalp, W.T.; Fischer, A.; John, D.; Zeiher, A.M.; Gosgnach, W.; Darville, H.; Montgomery, R.; Pestano, L.; Allée, G.; Paty, I.; et al. Effificiency and Target Derepression of An-ti-miR-92a: Results of a First in Human Study. Nucleic Acid Ther. **2020**, 30, 335–345. [CrossRef] [PubMed]
2. Abramov,S., Boytsov,A., Bykova,D., Penzar,D.D., Yevshin,I., Kolmykov,S.K., Fridman,M.V.,

3. Aiewsakun P, et al (2021). SARS-CoV-2 genetic variations associated with COVID-19 pathogenicity. Microb Genom.7(12), PP-1-13
4. Arnold,C.D., Gerlach,D., Stelzer,C., Boryń,Ł.M., Rath,M. and Stark,A. (2013) Genome-wide quantitative enhancer activity maps identified by STARR-seq. Science, **339**, 1074–1077.
5. Bajan, S.; Hutvagner, G. RNA-Based Therapeutics: From Antisense Oligonucleotides to miRNAs. Cells **2020**, 9, 137. [CrossRef]
6. Sudheer Menon (2020) "Preparation and computational analysis of Bisulphite sequencing in Germfree Mice" International Journal for Science and Advance Research In Technology, 6(9) PP (557-565).
7. Sudheer Menon, Shanmughavel Piramanayakam and Gopal Agarwal (2021) "Computational identification of promoter regions in prokaryotes and Eukaryotes" EPRA International Journal of Agriculture and Rural Economic Research (ARER), Vol (9) Issue (7) July 2021, PP (21-28).
8. Sudheer Menon (2021) "Bioinformatics approaches to understand gene looping in human genome" EPRA International Journal of Research & Development (IJRD), Vol (6) Issue (7) July 2021, PP (170-173).
9. Sudheer Menon (2021) "Insilico analysis of terpenoids in Saccharomyces Cerevisiae"international Journal of Engineering Applied Sciences and Technology, 2021 Vol. 6, Issue1, ISSN No. 2455-2143, PP(43-52).
10. Sudheer Menon (2021) "Computational analysis of Histone modification and TFBs that mediates gene looping" Bioinformatics, Pharmaceutical, and Chemical Sciences (RJLBPCS), June 2021, 7(3) PP (53-70).
11. Sudheer Menon Shanmughavel piramanayakam, Gopal Prasad Agarwal (2021) "FPMD-Fungal promoter motif database: A database for the Promoter motifs regions in fungal genomes" EPRA International Journal of Multidisciplinary research,7(7) PP (620-623).
12. Sudheer Menon, Shanmughavel Piramanayakam and Gopal Agarwal (2021) Computational Identification of promoter regions in fungal genomes, International Journal of Advance Research, Ideas and Innovations in Technology, 7(4) PP (908-914).
13. Boyle,A.P., Zhang,Q.C., Zakharia,F., Spacek,D.V., et al. (2013) Extensive variation in chromatin states across humans. Science, **342**, 750–752.
14. Callan,C.G.,Jr, Kinney,J.B., et al. (2012) Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. Nat. Biotechnol., **30**, 271–277.
15. Cantalupo S, et al (2021) Regulatory Noncoding and Predicted Pathogenic Coding Variants of CCR5 Predispose to Severe COVID-19. Int J Mol Sci. 20;22(10):5372.
16. Chattopadhyay, P.; Srinivasa Vasudevan, J.; Pandey, R. Noncoding RNAs: Modulators and modulatable players during infectioninduced stress response. Brief. Funct. Genomics **2021**, 20, 28–41. [CrossRef]
17. Cheng, J., Zhou, X., Feng, W., Jia, M., Zhang, X., An, T., Luan, M., Pan, Y.i., Zhang, S., Zhou, Z., Wen, L., Sun, Y., Zhou, C., 2021. Risk stratification by long non-coding RNAs profiling in COVID-19 patients. J. Cell Mol. Med. 25, 4753–4764. https://doi.org/10.1111/jcmm.16444.
18. Cheng, J.; Zhou, X.; Feng, W.; Jia, M.; Zhang, X.; An, T.; Luan, M.; Pan, Y.; Zhang, S.; Zhou, Z.; et al. Risk stratification by long non-coding RNAs profifiling in COVID-19 patients. J. Cell. Mol. Med. **2021**, 25, 4753–4764. [CrossRef]

19. Cheng, L.; Wang, P.; Tian, R.; Wang, S.; Guo, Q.; Luo, M.; Zhou, W.; Liu, G.; Jiang, H.; Jiang, Q. LncRNA2Target v2.0: A comprehensive database for target genes of lncRNAs in human and mouse. Nucleic Acids Res. **2019**, 47, D140–D144. [CrossRef]
20. Cheng,Z., Vermeulen,M., Rollins-Green,M., DeVeale,B. and Babak,T. (2021) Cis-regulatory mutations with driver hallmarks in major cancers. iScience, **24**. Baca,S.C., Singler,C., Zacharia,S., Seo,J.-H., Morova,T., Hach,F., Ding,Y., Schwarz,T.,
21. Sudheer Menon, Vincent Chi Hang Lui and Paul Kwong Hang Tam (2021) Bioinformatics methods for identifying hirschsprung disease genes, International Journal for Research in Applied Science & Engineering Technology (IJRASET), Volume 9 Issue VII July, PP (2974-2978).
22. Sudheer Menon, (2021), Bioinformatics approaches to understand the role of African genetic diversity in disease, International Journal Of Multidisciplinary Research In Science, Engineering and Technology (IJMRSET), 4(8), PP 1707-1713.
23. Sudheer Menon (2021) Comparison of High-Throughput Next generation sequencing data processing pipelines, International Research Journal of Modernization in Engineering Technology and Science (IRJMETS), 3(8), PP 125-136.
24. Sudheer Menon (2021) Evolutionary analysis of SARS-CoV-2 genome and protein insights the origin of the virus, Wuhan, International Journal of Creative Research Thoughts (IJCRT), 9 (8), PP b696-b704.
25. Sudheer Menon, Vincent Chi Hang Lui and Paul Kwong Hang Tam (2021) A step-by-step work flow of Single Cell RNA sequencing data analysis, International Journal for Scientific Research and Development (IJSRD), 9(6) PP 1-13.
26. Sudheer Menon (2021) Computational characterization of Transcription End sites in Human Genome, International Journal of All Research Education and Scientific Methods (IJRESM), 9(8), PP 1043-1048.
27. Sudheer Sivasankaran Menon and Shanmughavel Piramanayakam (2021) 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), 2(4), PP 709-714.
28. Civenni, G. Targeting Promoter-Associated Noncoding RNA In Vivo. Methods Mol. Biol. **2017**, 1543, 259–270. [CrossRef] [PubMed]
29. Cohen, M.S. Hydroxychloroquine for the Prevention of Covid-19—Searching for Evidence. N. Engl. J. Med. **2020**, 383, 585–586. [CrossRef]
30. Cui, J.; Li, F.; Shi, Z.-L. Origin and evolution of pathogenic coronaviruses. Nat. Rev. Genet. **2019**, 17, 181–192. [CrossRef] [PubMed]
31. Drubay,D., Gautheret,D. and Michiels,S. (2018) A benchmark study of scoring methods for non-coding mutations. Bioinformatics, **34**, 1635–1641.
32. Duffy, S., 2018. Why are RNA virus mutation rates so damn high? PLoS Biol. 16, e3000003. https://doi.org/10.1371/journal.pbio.3000003.
33. Favorov,A.V., Vorontsov,I.E., Baulin,E., et al. (2021) Landscape of allele-specific transcription factor binding in the human genome. Nat. Commun., **12**, 2751.
34. Gao,P., Xia,J.-H., Sipeky,C., Dong,X.-M., Zhang,Q., Yang,Y., Zhang,P., Cruz,S.P., Zhang,K.,Zhu,J., et al. (2018) Biology and Clinical Implications of the 19q13 Aggressive Prostate Cancer Susceptibility Locus. Cell, **174**, 576–589.e18.
35. Sudheer Menon, Vincent Chi Hang Lui and Paul Kwong Hang Tam (2021) Bioinformatics tools and methods to analyze single cell RNA sequencing data,

International Journal of Innovative Science and Research Technology, (IJISRT), 6(8), PP 282-288.

36. Sudheer Menon (2021) Computational genome analysis for identifying Biliary Atresia genes, International Journal of Biotechnology and Microbiology, (IJBM), 3(2), PP 29-33.
37. Sudheer Menon (2021) Recent Insilco advancements in genome analysis and characteristics of SARS-Cov2. International Journal of Biology Research, (IJBR), 6(3), PP 50-54.
38. Sudheer Menon (2021) Bioinformatics methods for identifying Human disease genes, International Journal of Biology Sciences, (IJBR), 3(2), PP 1-5.
39. Sudheer Menon (2021) SARS-CoV-2 Genome structure and protein interaction map, insights to drug discovery, International Journal of Recent Scientific Research, (IJRSR), 12(8), PP 42659-42665.
40. Sudheer Menon (2021) Insilico Insights to Mutational and Evolutionary aspects of SARS-Cov2, International Journal of Multidisciplinary Research and Development, (IJMRD) 8(8), 167-172.
41. Sudheer Menon (2021) Computational biology, machine learning and reverse vaccinology detects the role of conserved Nsp3 Protein and its importance in Covid-19 vaccine development, European Journal of Biotechnology and Bioscience 9(3), PP 95-99.
42. Sudheer Menon (2021) Protein Evolutionary Analysis of SARS-CoV-2 Delta Plus and C.1.2 Insights Virulence and Host Immunity, International Journal of Multidisciplinary Research and Publications (IJMRAP), 4(2), PP 6-13.
43. Garg, A.; Ghoshal, U.; Patel, S.S.; Singh, D.V.; Arya, A.K.; Vasanth, S.; Pandey, A.; Srivastava, N. Evaluation of seven commercial RT-PCR kits for COVID-19 testing in pooled clinical specimens. J. Med. Virol. **2021**, 93, 2281–2286. [CrossRef]
44. Geleris, J.; Sun, Y.; Platt, J.; Zucker, J.; Baldwin, M.; Hripcsak, G.; Labella, A.; Manson, D.K.; Kubin, C.; Barr, R.G.; et al. Observational Study of Hydroxychloroquine in Hospitalized Pa-tients with Covid-19. N. Engl. J. Med. **2020**, 382, 2411–2418. [CrossRef] [PubMed]
45. Guo,Y.A., Chang,M.M. and Skanderup,A.J. (2020) MutSpot: detection of non-coding mutation hotspots in cancer genomes. NPJ Genom Med, **5**, 26.
46. Hai-Bing, Y.; Sivasankaran, M.S.; Ottakandathil, B.R.; Zhong-Luan, W.; Man-Ting, S.; Ho-Yu, C.; Kak-Yuen,W.; Kwong-Hang, T.; Chi-Hang, L. Environmental Toxin Biliatresone-Induced Biliary Atresia-like Abnormal Cilia and Bile Duct Cell Development of Human Liver Organoids. Toxins **2024**, 16, 144.
47. Harvey, W.T.; Carabelli, A.M.; Jackson, B.; Gupta, R.K.; Thomson, E.C.; Harrison, E.M.; Ludden, C.; Reeve, R.; Rambaut, A.; Peacock, S.J.; et al. SARS-CoV-2 variants, spike mutations and immune escape. Nat. Rev. Genet. **2021**, 19, 409–424. [CrossRef] [PubMed]
48. Sudheer Menon (2021) Computing for Genomics and Proteomics in Microbial biotechnology in the identification of SARS COV2 Variants, International Journal of Scientific Engineering and Applied Science (IJSEAS), 7(9), PP 147-175.
49. Sudheer Menon (2021) Functional Genome analysis and drug screening by Bioinformatics methods in SARS-CoV-2. International Journal of Academic Research and Development (IJARD), 6(5), PP 33-39.

50. Sudheer Menon (2021) Advances in Systems biology and immunological data analysis for drug and vaccine development, International Journal of Educational Research and Studies (IJERS), 3 (3), PP 53-64.
51. Sudheer Menon (2021) Whole genome sequencing and Insilico approaches to understand the genetic basis of Rare Diseases, International Journal of Advanced Scientific Research (IJASR), 6(5), PP 7-13.
52. Sudheer Menon (2021) Bioinformatics approaches to identify neurodegenerative diseases by Next generation sequencing data, International Journal of Educational Research and Development (IJERD), 3(3), PP 70-74.
53. Sudheer Menon (2021) Bioinformatics approaches for cross-species liver cancer analysis based single cell RNA sequencing data, International Journal of Molecular Biology and Biochemistry (IJMBB), 3(1), PP 05-13.
54. Sudheer Menon (2021) Genomic sequence of worldwide strains of SARS-CoV-2: insights the Role of Variants in Disease Epidemiology, International Journal of Advanced Research and Development (IJARD), 6(4), PP 14-21.
55. Sudheer Menon (2021) Differential Gene expression analysis by Single Cell RNA sequencing reveals insights to genetic diseases, International Journal of Multidisciplinary Education and Research (IJMER), 6(4), PP 1-7.
56. Sudheer Menon (2021) Protein–protein interactions by exploiting evolutionary information insights the genes and conserved regions in the corresponding human and mouse genome, International Journal of Advanced Multidisciplinary Research (IJARM), 8(9), PP 36-55.
57. Sudheer Menon and Binu Thomas (2021) An insights to Bioinformatics methods in Natural Product Drug Discovery, International Journal of Clinical Biology and Biochemistry (IJCBB), 3(1), PP 26-37.
58. Huang,C.-C.F., Kalita,C., et al. (2021) Genetic determinants of chromatin reveal prostate cancer risk mediated by context-dependent gene regulation. bioRxiv, 1101/2021.05.10.443466.
International Journal of Pharmacy and Biological Sciences (IJPBS), 11(3), PP-104-115.
59. Ivashchenko, A.; Rakhmetullina, A.; Aisina, D. How miRNAs can Protect Humans from Coronaviruses COVID-19, SARS-CoV, and MERS-CoV. Research Square. 2020. Available online: https://assets.researchsquare.com/fifiles/rs-16264/v1/manuscript.pdf (accessed on 18 February 2021).Non-coding RNA **2021**, 7, 14 14 of 16
60. Kasowski,M., Kyriazopoulou-Panagiotopoulou,S., Grubert,F., Zaugg,J.B., Kundaje,A., Liu,Y.,
61. Kircher,M., Witten,D.M., Jain,P., O'Roak,B.J., Cooper,G.M. and Shendure,J. (2014) A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet., **46**, 310–315.
62. Kircher,M., Xiong,C., Martin,B., Schubach,M., Inoue,F., Bell,R.J.A., Costello,J.F., Shendure,J. and Ahituv,N. (2019) Saturation mutagenesis of twenty disease-associated regulatory elements at single base-pair resolution. Nat. Commun., **10**, 3583.
63. Kozomara, A.; Griffifiths-Jones, S. miRBase: Integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. **2010**, 39, D152–D157. [CrossRef]
64. Sudheer Menon and Binu Thomas (2021) Use of multi omics data in precision medicine and cancer research with applications in tumor subtyping, prognosis, and diagnosis, International Journal of Advanced Education and Research (IJAER), 6(5), PP 19-29.

65. Sudheer Menon and Binu Thomas (2021) Gene Expression Analysis, Functional Enrichment, and Network Inference in disease prediction, International Journal of Advanced Educational Research (IJAER), 6(5), PP 6-18.
66. Sudheer Menon and Binu Thomas (2021) Vaccine informatics, systems biology studies on neuro-degenerative diseases, International Journal of Multidisplinary Educational Research IJMER, 10(9-4), PP 85-101.
67. Sudheer Menon (2021) Computational approaches to understand the role of epigenetics and microRNA in Cancer among different populations, International Journal of Multidisciplinary Research Review (IJMDRR) 7(9), PP 69-81.
68. Sudheer Menon and Binu Thomas (2021) Recent advances, methods and opportunities of Deep learning in computational and systems biology International Journal of Applied and Advanced Scientific Research (IJAASR) 6(2), PP 12-25.
69. Sudheer Menon (2021) Computational genome based studies of Complex Genetic Diseases, International Journal of Multidisciplinary Research and Modern Education (IJMRME) 7(2) PP 28-39.
70. Sudheer Menon (2021) Transcriptomic and metabolic pathway analysis reveals the mutated genes p53 and TP53 and their role in different types of cancer, International Journal of Engineering Research and Modern Education (IJERME), 6(2) PP 10-33.
71. Sudheer Menon, Amanda Ramer Tait, Shanmughavel Piramanagakam, Vincent Chi Hang Lui and Paul Kwong Hang Tam (2021) Bisulphite seq data reveals host-microbiota interaction in germfree and conventional mice, International Journal of Bioinformatics and Biological Sciences, 9(1), PP 1-8.
72. Lai, C.-C.; Shih, T.-P.; Ko, W.-C.; Tang, H.-J.; Hsueh, P.-R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and corona-virus disease-2019 (COVID-19): The epidemic and the challenges. Int. J. Antimicrob. Agents. **2020**, 55, 105924. [CrossRef]
73. Lee,D., Gorkin,D.U., Baker,M., Strober,B.J., Asoni,A.L., McCallion,A.S. and Beer,M.A.bioRxiv preprint doi: https://doi.org/10.1101/2022.03.11.483985; this version posted March 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission. (2015) A method to predict the impact of regulatory variants from DNA sequence. Nat. Genet., **47**, 955–961.
74. Liu,L., Sanderford,M.D., Patel,R., Chandrashekar,P., Gibson,G. and Kumar,S. (2019) Biological relevance of computationally predicted pathogenicity of noncoding variants. Nat. Commun., **10**, 330.
75. Liu,S., Liu,Y., Zhang,Q., Wu,J., Liang,J., Yu,S., Wei,G.-H., White,K.P. and Wang,X. (2017) Systematic identification of regulatory variants associated with cancer risk. Genome Biol., Sudheer Menon (2021) Insilico identification of Transcription End Sites in Eukaryotes,

International Journal of Pharmacy and Biological Sciences (IJPBS), 11(3), PP-104-115
76. Sudheer Menon (2021) Computational identification of upregulated genes in Breast, Prostate, and Colorectal Cancer and the Cellular Signaling process, Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences, (RJLBPCS), 7(5), PP 68-83.
77. Sudheer Menon (2021) Structural and Functional Characteristics of miRNAs in colon cancer and the identification of targets by Insilco methods, Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences, (RJLBPCS), 7(6), PP 86-96.
78. Sudheer Menon, Vincent Chi Hang Lui and Paul Kwong Hang Tam (2021, Statistical approach for Predicting Genome Architecture and the major challenges, World Journal of Advanced Research and Reviews (WJARR), 2022, 13(1), PP 433-445.
79. Binu Thomas, Mahesh Mohanan, and Sudheer Menon (2021), Role of edible Plants to balance dietary requirements in human diet, Devagiri Journal of Science, 7(1), PP-27-40
80. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal (2022), Bioinformatics Analysis of Actin molecules and the actin gene by next-generation sequencing data insights cell mobility, International Journal of Advanced Trends in Engineering and Technology, IJATET, 7(1), PP 1-20.
81. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal (2022), CDKN2A gene expression/mutation prognosis by computational methods in the identification of nodular melanoma, International Journal of Scientific Research and Modern Education (IJSRME), 7(1), PP 11-22.
82. Sudheer Menon , Vincent Chi Hang Lui1, Paul Kwong Hang Tam (2022), comparison of computational tools for differential gene expression analysis of rna sequencing and single-cell rna sequencing data, European Journal of Biomedical AND Pharmaceutical sciences, 9(4), PP 1-9.
83. Lundstrom, K. RNA-based drugs and vaccines. Expert Rev. Vaccines **2015**, 14, 253–263. [CrossRef]
84. McVicker,G., van de Geijn,B., Degner,J.F., Cain,C.E., Banovich,N.E., Raj,A., Lewellen,N.,
85. Melnikov,A., Murugan,A., Zhang,X., Tesileanu,T., Wang,L., Rogov,P., Feizi,S., Gnirke,A.,
86. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal (2022), computational characterization of genomewide dna methylation and hydroxymethylation in the mouse hypothalamus by next-generation sequencing data, International Journal of Applied and Advanced Scientific Research (IJAASR), 7(1), PP 9-29.
87. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal (2022), Differential expression analysis of stressinducible candidate genes in response to cold and draught in Coffea Arabica, International Journal of Advanced Trends in Engineering and Technology (IJATET), 7(1), PP 21-38.
88. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal (2022), Identification of immune gene–related IncRNA signature to detect brain tumors, International Journal of Current Research and Modern Education (IJCRME), 7(1), PP 6-20.
89. Sudheer Menon, Shanmughavel piramanayakam and Gopal Prasad Agarwal (2023), “Identification of Molecular Markers Associated With liver steatosis and Treatment

Based on Integrated Transcriptome Analysis by PNPLA3 gene", Journal of Emerging Technologies and Innovative Research (JETIR), 10 (12), PP- g398-g416.

90. Sudheer Menon (2023), Dynamical Modeling of the Core Gene Network and Mutation in Transitional Cell Carcinoma, Journal of Emerging Technologies and Innovative Research (JETIR), 10 (12), PP-h326-h338.
91. Morenikeji, O.B.; Bernard, K.; Strutton, E.; Wallace, M.; Thomas, B.N. Evolutionarily Conserved Long Non-coding RNA Regulates Gene Expression in Cytokine Storm During COVID-19. Front. Bioeng. Biotechnol. **2021**, 8, 582953. [CrossRef]
92. Moyon L et al (2022) Classification of non-coding variants with high pathogenic impact. PLoS Genet 18(4): e1010191
93. Munster, V.; Koopmans, M.; van Doremalen, N.; van Riel, D.; de Wit, E. A Novel Coronavirus Emerging in China—Key Questions for Impact Assessment. N. Engl. J. Med. **2020**, 382, 692–694. [CrossRef]
94. Hai-Bing, Y.; Sivasankaran, M.S.; Ottakandathil, B.R.; Zhong-Luan, W.; Man-Ting, S.; Ho-Yu, C.; Kak-Yuen,W.; Kwong-Hang, T.; Chi-Hang, L. Environmental Toxin Biliatresone-Induced Biliary Atresia-like Abnormal Cilia and Bile Duct Cell Development of Human Liver Organoids. Toxins **2024**, 16, 144.
95. Sudheer Menon, "Neural network prediction of genome wide transcriptome signatures in Adenocarcinoma", International Journal of Creative Research Thoughts (IJCRT), ISSN:2320-2882, Volume.12, Issue 12, pp.b716-b726, December 2024.
96. Sudheer Menon, "Computational Prediction of SARS-CoV-2 Genomic, Proteomic Mutation, and Variants by (NGS) Next-Generation Sequencing Data", International Journal of Science and Research (IJSR), Volume 13 Issue 12, December 2024, pp. 565-573.
97. Myrthil,M., Gilad,Y. and Pritchard,J.K. (2013) Identification of genetic variants that affect histone modifications in human cells. Science, **342**, 747–749.
98. Natarelli, L.; Weber, C. Next-Generation Therapeutic Concepts for Atherosclerosis: Focus on Cell Specifificity and Noncoding RNAs. Thromb. Haemost. **2019**, 119, 1199–1201. [CrossRef]
99. Paniri, A.; Akhavan-Niaki, H. Emerging role of IL-6 and NLRP3 inflflammasome as potential therapeutic targets to combat COVID-19: Role of lncRNAs in cytokine storm modulation. Life Sci. **2020**, 257, 118114. [CrossRef]
100. Poller, W.; Dimmeler, S.; Heymans, S.; Zeller, T.; Haas, J.; Karakas, M.; Leistner, D.-M.; Jakob, P.; Nakagawa, S.; Blankenberg, S.; et al. Non-coding RNAs in cardiovascular diseases: Diagnostic and therapeutic perspectives. Eur. Hear. J. **2018**, 39, 2704–2716. [CrossRef] [PubMed]
101. Saha, I., Ghosh, N., Maity, D., Sharma, N., Sarkar, J.P., Mitra, K., 2020. Genome-wide analysis of Indian SARS-CoV-2 genomes for the identification of genetic mutation and SNP. Infect. Genet. Evol. 85, 104457. https://doi.org/10.1016/j.meegid.2020.104457.
102. Shaath, H.; Alajez, N.M. Identifification of PBMC-based molecular signature associational with COVID-19 disease severity. Heliyon **2021**, 7, e06866. [CrossRef]
103. Shen, L.; Cui, S.; Zhang, D.; Lin, C.; Chen, L.; Wang, Q. Comparison of four commercial RT-PCR diagnostic kits for COVID-19 in China. J. Clin. Lab. Anal. **2021**, 35, e23605. [CrossRef]

104. Shereen, M.A.; Khan, S.; Kazmi, A.; Bashir, N.; Siddique, R. COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses. J. Adv. Res. **2020**, 24, 91–98. [CrossRef]
105. Spisák,S., Lawrenson,K., Fu,Y., Csabai,I., Cottman,R.T., Seo,J.-H., Haiman,C., Han,Y., Lenci,R., Li,Q., et al. (2015) CAUSEL: an epigenome- and genome-editing pipeline for establishing function of noncoding GWAS variants. Nat. Med., **21**, 1357–1363.
106. Taheri, M.; Rad, L.M.; Hussen, B.M.; Nicknafs, F.; Sayad, A.; Ghafouri-Fard, S. Evaluation of expression of VDR-associated lncRNAs in COVID-19 patients. BMC Infect. Dis. **2021**, 21, 588. [CrossRef]
107. Vetter, P.; Kaiser, L.; Calmy, A.; Agoritsas, T.; Huttner, A. Dexamethasone and remdesivir: Finding method in the COVID-19 madness. Lancet Microbe **2020**, 1, e309–e310. [CrossRef]
108. Wang,X., He,L., Goggin,S.M., Saadat,A., Wang,L., Sinnott-Armstrong,N., Claussnitzer,M. and Kellis,M. (2018) High-resolution genome-wide functional dissection of transcriptional regulatory regions and nucleotides in human. Nat. Commun., **9**, 5380.
109. Zacher,B., Michel,M., Schwalb,B., Cramer,P., Tresch,A. and Gagneur,J. (2017) Accurate Promoter and Enhancer Identification in 127 ENCODE and Roadmap Epigenomics Cell Types and Tissues by GenoSTAN. PLoS One, **12**, e0169249
110. Zeng, F.; Dai, C.; Cai, P.; Wang, J.; Xu, L.; Li, J.; Hu, G.; Wang, Z.; Zheng, F.; Wang, L. A comparison study of SARS-CoV-2 IgG antibody between male and female COVID-19 patients: A possible reason underlying different outcome between sex. J. Med. Virol. **2020**, 92, 2050–2054. [CrossRef]
111. Zhang,P., Xia,J.-H., Zhu,J., Gao,P., Tian,Y.-J., Du,M., Guo,Y.-C., Suleman,S., Zhang,Q., Kohli,M., et al. (2018) High-throughput screening of prostate cancer risk loci by single nucleotide polymorphisms sequencing. Nat. Commun., **9**, 2022.
112. Zheng Wang, et al (2022) Performance comparison of computational methods for the prediction of the function and pathogenicity of non-coding variants, Genomics, Proteomics & Bioinformatics, PP 1-35.