International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 12 | December - 2022 Impact Factor: 7.185 ISSN: 2582-3930

A COMPLETE ASSESSMENT AND USE OF BIOINFORMATICS IN THE COVID-19

Dr CK GOMATHY, Assistant Professor, Department of CSE,
SCSVMV Deemed to be University, India
Mr. V.V.S.V.RONIT , Mr.N.S.VIGNESHWARA REDDY,
Mr. VENKATESHWARULU, Mr. S.SAI PRANAV
UG Scholars, Department of CSE, SCSVMV Deemed to be University, India

Abstract

The severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2) that causes the 2019 novel coronavirus illness (COVID-19) is a contagious virus that can spread from person to person. Over 228 million cases had been identified as COVID-19 infections as of September 21, 2021, over more than 200 nations and regions. As it has gradually spread over the world, there have been around 4.69 million fatalities, and the mortality rate has risen to about 2.05%.To combat the COVID-19 pandemic, it is crucial to investigate the properties of the novel coronavirus' genome and proteins, clinical diagnostics, pathogenic mechanisms, and the creation of antiviral medications and vaccines. The COVID-19-related investigations needed to comprehend the pandemic's occurrence are constrained by the limitations of conventional biology methods. Bioinformatics is the application of computational methods and analytical tools in the field of biological research, which has clear benefits in predicting the structure, outcome, function, and evolution of unidentified genes and proteins, as well as in drug and vaccine screening from a large amount of sequence data. Based on the most recent reports of bioinformatics technologies, we have here provided a thorough summary of some of the most significant COVID-19 approaches and applications with an eye on future research for containing the virus pandemic. The SARS-CoV-2 genome can be swiftly retrieved using third-generation sequencing (TGS) and next-generation sequencing (NGS) technology in addition to virus detection. The development of data on the SARS-CoV-2 genome sequences, variations, and haplotypes aids in our understanding of the structure of the genome and proteins, variant calling, mutation, and other biological traits. According to phylogenetic research and sequence alignment, the bat may be the novel coronavirus's native host. The mechanism of the immunological response generated by COVID-19 can be discovered via single-cell RNA sequencing, which is a rich source of information. Angiotensin-converting enzyme 2 (ACE2) can be explored as a possible therapeutic target to treat COVID-19 since it functions as an entrance receptor. The discovery of medications for SARS-CoV-2 can be sped up using molecular dynamics modelling, molecular docking, and artificial

intelligence (AI) technologies of bioinformatics methodologies based on drug databases. In the meanwhile, computational methods such as reverse vaccination, immune informatics, and structural vaccination are helpful in identifying the best vaccinations to prevent COVID-19 infection.

Keywords: COVID-19, SARS-CoV-2, Application, Bioinformatics technology

I. INTRODUCTION

Wuhan, Hubei Province, China, has confirmed many cases of acute respiratory tract infection since December 2019. (C. Huang et al., 2020; Zhu et al., 2020). A serious threat to public health is posed by the widespread pandemic of confirmed cases, which spreads quickly and varies in its leaps in case count (Mahase, 2020; J.T. Wu et al., 2020). On March 11, 2020, the World Health Organization (WHO) proclaimed the epidemic a "International Public Health Emergency." The International Committee's Coronavirus Study Group (CSG) designated the recently discovered -coronavirus (-CoVs) as SARS-CoV-2, or severe acute respiratory syndrome coronavirus type 2. SARS-CoV-2 disease was recognised by WHO as coronavirus disease 2019 (COVID-19) and was thought to be an infectious disease (Zhu et al., 2020).SARS-CoV-2 is currently known to be extremely contagious and quickly spread among people by respiratory droplets. It may also live in the air for up to two hours (Y. Liu et al., 2020; Munster et al., 2020; wei Lu et al., 2020; H. Zhang et al., 2020). Following infection, the incubation period typically lasts 4 to 8 days (Chen et al., 2020a; C. Huang et al., 2020; D. Wang et al., 2020). SARS-CoV-2 can infect people of any age, but older patients and smokers are more likely to develop a severe illness with comorbidities (Chen et al., 2020b; J. Zhang et al., 2020). Importantly, situations, such as human-to-human transfer during the asymptomatic infection period, appear to enhance the likelihood of local transmission (Rothe et al., 2020, Yuan et al., 2006).

II. BIOINFORMATICS TECHNOLOGY

The two highly pathogenic β-CoVs, which are known to cause the fatal pneumonia diseases Middle East respiratory syndrome (MERS) and severe acute respiratory syndrome (SARS), are zoonotic viruses that pose a substantial threat to human health since they primarily infect the lower respiratory tract (Channappanavar and Perlman, 2017, Luk et al., 2019, Ramadan and Shaib, 2019; F. Wu et al., 2020). SARS and MERS had overall fatality rates of 9.6% and 34.4%, respectively, according to the WHO (Munster et al., 2020, World Health Organization, 2020). In contrast, the global mortality rate for SARS-CoV-2 is 2.05%, while the fatality rates for elderly people with chronic illnesses and ICU patients have reached 17–38%. (Chen

et al., 2020a; D. Wang et al., 2020).Global Real Time COVID-19 Outbreak Map from WHO (https://covid19.who.int/) displaying the confirmed cases, fatalities, and vaccination doses given in various nations and areas globally. Human coronavirus (CoVs) infection has no proven cure as of yet, and no medications are advised. However, a lot of research has gone into characterising CoVs, and the medicines and vaccines that target the virus are being thoroughly examined. Traditional biological tools for viral study have a number of drawbacks, including the inability to gather fast and extensive data. Bioinformatics technology has greatly aided in the quick and thorough study of SARS and MERS, and it will also usher in a new era in the investigation of SARS-CoV-2 (Woo et al., 2010).

In the diagnosis of the novel coronavirus, sequencing diagnosis remains the primary technique, and it is crucial to the study of how viruses function. The examination of virus sequencing data using bioinformatics techniques serves as a key foundation for the identification of SARS-CoV-2 in this scientific and technological conflict. SARS-CoV-2 databases offer interactive visualisation, variant calling, quality assessment of the genome and protein structure, and other biological properties. Sequence alignment between different species enables us to comprehend the evolutionary history and features of viruses. It is critically necessary to develop point-of-care testing (POCT) methods that are quick, practical, precise, and sensitive for detecting SARS-CoV-2 in various situations.The characteristics of epithelial and immunological cell types were discovered through bioinformatics analysis of single-cell RNA sequencing of SARS-CoV-2 RNA. Additionally, it can offer high-throughput computational support for predictive and screening to hasten the development of vaccines and specific medications for pandemic treatment and prevention. Through the use of high-performance computing techniques like molecular dynamics simulation, molecular docking, and artificial intelligence, small molecules can be competitively bound to the functional sites of viral proteins, preventing the viral protein from binding to its true substrate and inhibiting viral activity, or the receptors can cause the host to produce immune responses and neutralising antibodies, inhibiting the activity of the virus.

SARS-CoV-2 is a highly pathogenic human pathogen and may be a zoonotic agent, despite the fact that we know relatively little about it. The evidence of the virus's RNA genome sequence obtained through bioinformatics methods is crucial for understanding the characteristics, evolution, detection, pathogenesis, and function of SARS-CoV-2 as well as for the development of antiviral medications and vaccines, which will provide information for making public health decisions. Challenges will continue to be faced in a number of important areas.

III.THE BIOINFORMATICS TECHNOLOGIES RELATING TO COVID-19

Based on reports that are currently available, the following sections offer some of the most crucial techniques and uses of bioinformatics technologies in battling the COVID-19. A description of the workflow is also provided.

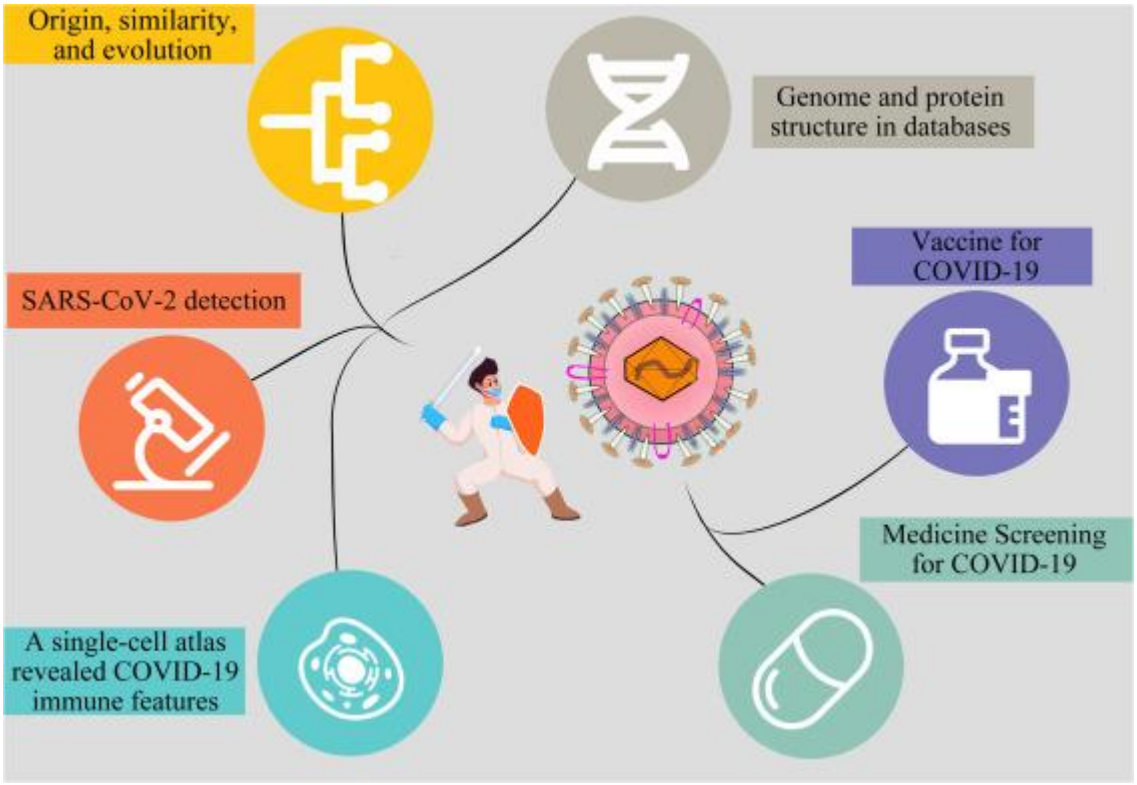

Fig 1: Technologies Relating To COVID-19

IV.GENOME AND PROTEIN STRUCTURE IN DATABASES

The SARS-CoV-2 genome and protein structure quality evaluation, variant calling, and interactive visualisation are provided by significantly updated and publicly accessible databases using the next-generation sequencing (NGS) and third-generation sequencing (TGS) technologies. GISAID (Global Initiative on Sharing Avian Influenza Data, https://www.gisaid.org), National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov), National Microbiology Data Center (NMDC), China National GeneBank (CNGB), https://www.cngb.org), China National Center for Bioinformation (CNCB, https://bigd.big.ac.cn/ncov), and numerous other sources have (Song et al., 2020; W.M. Zhao et al., 2020).The novel coronavirus SARS-CoV-2, which belongs to lineage B -CoVs and is a positive single-stranded RNA virus with a genome length of about 29.9 kb, non-structural protein-coding regions at both ends, and structural

protein-coding regions and non-structural protein-coding regions in the middle, was found to be the cause of COVID-19 after nucleic acid sequence was obtained (Chan et al., 2020; Roujian Lu et al., 2020; Zhu et al., 2020). Positive-sense RNA genome sizes for SARS-CoV and MERS-CoV are 27.9 kb and 30.1 kb, respectively (De Wit et al., 2016).

There are a variable number of open reading frames and 29,870 nucleotides in the SARS-CoV-2 RNA genome (ORFs)(Chen et al., 2020a; Y. Liu et al., 2020; Munster et al., 2020; D. Wang et al., 2020; wei Lu et al., 2020; H. Zhang et al., 2020). The ORFs encode 9860 amino acids with five conventional ORFs on a shared coding strand and in the order of 5′ UTR, excluding the poly (A) tail (GenBank no. MN908947). polypeptide -replicase (orf1a/b, 7096-aa) Glycoprotein spike (S, 1273-aa) - Membrane protein - Envelope protein (E,75-aa) (M, 222-aa) - Protein nucleocapsid (N, 419-aa) the four crucial protein-coding regions S, E, M, and N, which encode structural proteins, are located in the 3′ UTR (Fig. 2A, B) (Chan et al., 2020). Additionally, some accessory genes, including ORF3b and ORF8, have an impact on the host's innate immune response. A wholly new, brief protein without a known function is encoded by ORF3b.The newly discovered ORF8 likely codes for a secreted protein made up of an alpha-helix followed by six-stranded beta sheets that lack any known functional domains or motifs (Chan et al., 2020). The ORF1a and ORF1b genes, which make up about two-thirds of viral RNA, are primarily found in the coding region of non-structural proteins. These genes translate the two replicase polyproteins (pp1a and pp1ab) into 16 non-structural proteins (NSP), referred to as NSP1–16, which are essential for viral transcription and replication (Chan et al., 2020; F. Wu et al., 2020; Zhou et al., 2020).

V.CONCLUSION

The genomes that are sequenced and compared are those of the virus, not of humans. During the month of October 2022, 15.9% of all confirmed molecular COVID-19 cases were sequenced Medical staff and laypeople, such as friends, family members, and coworkers, may also provide information on COVID-19 to the general public. Health practitioners can better educate the public by studying the elements that influence such information sources.COVID-19 might trigger a new age of sustained increase in development assistance for health similar to HIV between 2001 and 2010, when development assistance for health increased by 11·2% per year and annual development assistance for health tripled between 2001 and 2010.Genomic sequencing goes beyond testing for SARS-CoV-2 and allows scientists to classify a virus as a particular variant and determine its lineage. Genomic surveillance has been a key component of public health efforts throughout the COVID-19 pandemic. Read more about the genomic sequencing process.

VI. REFERENCES

1. Dr.C K Gomathy and et al, Machine Learning-Based Clinical Decision Support System, International Journal of Scientific Research in Engineering and Management (IJSREM) Volume: 06 Issue: 10 | October - 2022 Impact Factor: 7.185 ISSN: 2582-3930

2.Dr.C K Gomathy et al,Web Service Composition In A Digitalized Health Care Environment For Effective Communications, Published by International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 5 Issue 4, April 2016, ISSN: 2278 – 1323.

3. Vishnupriya C K and et al,Dimensional and Morphologic Variations of palatal Rugae-a hospital based study among Chennai populations, International Journal Of Science Research, ISSN No: 2277-8179 Volume 7, Issue 7,P.No-19-20,July ’2018

4. Dr.C K Gomathy et al, Machine Learning-Based Clinical Decision Support System, International Journal of Scientific Research in Engineering and Management (IJSREM) Volume: 06 Issue: 10 | October - 2022 Impact Factor: 7.185 ISSN: 2582-3930

5. Dr.C K Gomathy et al,A Review On IOT Based Covid-19 Patient Health Monitor In Quarantine, International Research Journal of Engineering and Technology (IRJET),e-ISSN: 2395-0056 Volume: 08 Issue: 09 | Sep 2021 www.irjet.net p-ISSN: 2395-0072`

6. Dr.C K Gomathy, et al, A Medical Information Security Using Cryptosystem For Wireless Sensor Networks, International Journal Of Contemporary Research In Computer Science And Technology (Ijcrcst) E-Issn: 2395-5325 Volume3, Issue 4, P.No-1-5,April ’2017

7. Dr.C K Gomathy and et al, The Parkinson’s Disease Detection Using Machine Learning Techniques, International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 10 | Oct 2021, e-ISSN: 2395-0056, p-ISSN: 2395-0072.

8.Dr.C K Gomathy, V Geetha , T.Jayanthi, M.Bhargavi, P.Sai Haritha: A Medical Information Security Using Cryptosystem For Wireless Sensor Networks, International Journal Of Contemporary Research In Computer Science And Technology (Ijcrcst) *E*-Issn: 2395-5325 Volume3, Issue 4, P.No-1-5,April ’2017

9. Dr.C K Gomathy and et.al, The Smart Stick Assistant For Visually Challenged People Using Ai Image Recognition, International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 9 | Sep 2021, e-ISSN: 2395-0056, p-ISSN: 2395-0072.

10. Dr. C.K. Gomathy , Dr. V.Geetha ,G.S.V.P.Praneetha , M.Sahithi sucharitha. (2022). Medicine Identification Using Open Cv. Journal of Pharmaceutical Negative Results, 3718–3723. https://doi.org/10.47750/pnr.2022.13.S09.457

11. Dr. V.Geetha ,Dr. C.K. Gomathy , Kommuru Keerthi , Nallamsetty Pavithra. (2022). Diagnostic Approach To Anemia In Adults Using Machine Learning. Journal of Pharmaceutical Negative Results, 3713–3717. https://doi.org/10.47750/pnr.2022.13.S09.456

Author's Profile

- V.V.S.V.RONIT Student, Reg.no.11199A254 B.E. Computer Science, and Engineering, Sri Chandrasekharendra Saraswathi Viswa Maha Vidyalaya Enathur, Kanchipuram, India
- N.S.VIGNESHWARA REDDY Student, Reg.no.11199A259 B.E. Computer Science, and Engineering, Sri Chandrasekharendra Saraswathi Viswa Maha Vidyalaya Enathur, Kanchipuram, India
- VENKATESHWARULU Student, Reg.no.11199A250, B.E. Computer Science, and Engineering, Sri Chandrasekharendra Saraswathi Viswa Maha Vidyalaya Enathur, Kanchipuram, India
- S.SAI PRANAV Student, Reg.no.11199A224 B.E. Computer Science, and Engineering, Sri Chandrasekharendra Saraswathi Viswa Maha Vidyalaya Enathur, Kanchipuram, India
- Dr. C.K. Gomathy is Assistant Professor in Computer Science and Engineering at Sri Chandrasekharendra Saraswathi Viswa MahaVidyalaya, Enathur, Kanchipuram, India. Her area of interest in Software Engineering, Web Services, Knowledge, Management and IOT.