A Genome based Detection and Classification of Coronavirus Infection

[1]Ankita Sasmal, [2]Chhanda Ray*

[1]Techno India University Kolkata, West Bengal, INDIA
[2]State Council of Educational Research and Training Kolkata, West Bengal, INDIA

Abstract:
The Coronavirus (COVID-19) infection has become a global threat in recent time. Many researchers have been dedicated to control COVID-19 pandemic. In this paper, an effective method is presented for detection and classification of COVID-19 infection based on genome sequences. First, the COVID-19 infection is detected based on the induction of changes in the DNA microarray gene expression pattern of the host during and after infection and comparing it with DNA sequences of Coronavirus (SARS-CoV-2). In order to analyse DNA microarray gene expression data, a bi-directional string matching algorithm is used and the analytical result is represented in terms of eight-directional chain code sequence. At the end of the work, an approach for categorization of Coronavirus infection is provided based on the distribution probabilities of eight-directional chain code sequences correspond to DNA microarray gene expression data of different Corona viruses by taking random samples from the GenBank. The categorization of Coronavirus infection will be helpful for forecasting rate of mortality, rate of infection, severity of the infection and other issues related to COVID-19.

Keywords: Coronavirus, COVID-19, Distribution Probability, Genome Sequence.

1. Introduction
In the recent time, the Coronavirus disease outbreak, a typical person-to-person transmissible pneumonia caused by the severe acute respiratory syndrome Coronavirus 2 (SARS-COV-2), has become a global threat. There have been more than 45,69,076 confirmed cases of the Coronavirus disease with the death toll of 3,04,794 throughout 210 countries, territories and conveyances, as of May 15, 2020. The first case of Coronavirus disease had been suspected on the 1st of December 2019 in Wuhan, Hubei Province, China, where a total of 40,235 confirmed cases and 909 deaths had been reported till 10th February 2020 [20]. The novel Coronavirus (SARS-COV-2) disease has been named COVID-19 by the World Health Organization (WHO). The COVID-19 outbreak is believed to have caused a public health emergency of international concern by the WHO Director-General on 30th January 2020 [20].

The origin of COVID-19 has not yet been determined. According to the Centres for Disease Control and Prevention, the novel Coronavirus is transmitted from person to person principally by respiratory droplets causing different symptoms such as fever, cough and shortness of breath after a period ranging from 2 to 14 days following infection. However, the precise virologic and epidemiologic characteristics including transmissibility and mortality of novel Coronavirus in human are still unknown. The outbreak of COVID-19 has reached a pandemic state and threatening to be an epidemic. In this context, many researchers have been dedicated to control COVID-19 pandemic [1, 18, 22, 23]. A tentative three-week forecast of the evolution of the COVID-19 outbreak at the epicentre China using some estimated parameters is provided in [1]. A note of caution for the unfolding COVID-19 outbreak of China has also been forecasted in [1]. In [18], a systematic review about the epidemiology, clinical

characteristics, diagnosis, treatment and prevention of COVID-19 has been depicted based on current published evidence. The epidemiology, clinical manifestations, diagnosis, treatments and preventions of the novel Coronavirus is also aggregated and consolidated in [22]. A mathematical model for the transmission of MERS-CoV (Middle East Respiratory Syndrome Coronavirus) in two areas by separating the human population into two groups, namely, susceptible group and infection group is discussed in [23]. In this paper [23], the dynamics of this disease is studied by a compartmental model involving ordinary differential equations and the basic reproductive number is discussed for controlling the outbreak of the disease. A sensitivity analysis of the model [23] is also performed to determine the relative importance of the model parameters to the MERS-CoV transmission.

In order to control the spread of Corona Virus Disease (COVID-19), screening of large numbers of suspected cases for appropriate quarantine and treatment is necessary. It has been stated in [4] that Mesenchymal Stem Cell (MSC) transplantation could be a possible alternative for treating COVID-19 since H7N9 and the Coronavirus disease share similar complications such as Acute Respiratory Distress Syndrome (ARDS), lung failure and corresponding multi-organ dysfunction. A model for the classification of Macrophage Activation Syndrome (MAS) to stratify the MAS-like presentation in COVID-19 pneumonia and to explore the complexities of discerning ARDS from MAS is introduced in [9]. The potential impact of timing of anti-cytokine therapy on viral clearance, intra-pulmonary macrophage activation and emergent pulmonary vascular disease has also been discussed in this paper [9]. Paper [10] analysed the 3CLpro sequence, constructed its 3D homology model and screened it against a medicinal plant library containing 32,297 potential anti-viral phytochemicals/ traditional Chinese medicinal compounds. It has been stated that the above top nine medicinal compounds might serve as potential anti-SARS-CoV-2 lead molecules for further optimisation and drug development process to combat COVID-19. Another work [21] systematically analyzed all the proteins encoded by SARS-CoV-2 genes, compared them with proteins from other Coronaviruses, predicted their structure and built 19 structures by homology modelling. Finally, this paper[21] attempted to discover potential drugs by computational methods to treat SARS-CoV-2 infections.

Fast and accurate diagnostic methods are urgently needed for combating COVID-19 pandemic. In this context, detection of Coronavirus disease is the focus of attention of many researchers [6, 14, 15, 16, 19, 24]. A brief introduction of the general features of SARS-CoV-2 and current knowledge of molecular immune pathogenesis, diagnosis and treatment of COVID-19 on the base of the understanding of SARS-CoV and MERS-CoV infections is depicted in [6]. In [14], a deep learning based methodology is suggested for detection of Coronavirus infected patient using X-ray images. For classification of the Corona affected X-ray images from others using deep features, the support vector machine is used. The current issues and challenges for the laboratory diagnosis of COVID-19 infections are discussed in [15]. In [16], a review on diagnostic and surveillance technologies for SARS-CoV-2 and their performance characteristics is focused. Paper[19] has introduced a deep learning method that could extract COVID-19's graphical features based on COVID-19 radiographical changes in CT images for timely and accurate clinical diagnosis of COVID-19. In [19], the experiment was done on 1065 CT images of pathogen-confirmed COVID-19 cases (325 images) along with previously diagnosed viral pneumonia (740 images) cases and a total accuracy of 89.5% with specificity of 0.88 and sensitivity of 0.87 has been achieved. A protocol for detection of COVID-19 using CRISPR diagnostics has been proposed in [24]. Paper [11] introduced a methodology for cancer identification and gene classification based on DNA microarray gene expression patterns.

In general, nucleic acid testing and CT scans are used for diagnosing and screening COVID-19. Nucleic acid testing is the gold standard for detection of Coronavirus infected patient, but it is time consuming with significant false negative results. Molecular techniques are more accurate than syndromic testing and CT scans since they can target and identify specific pathogens. The

development of molecular techniques is dependent upon understanding the proteomic and genomic composition of the pathogen or the induction of changes in the expression of proteins/genes in the host during and after infection. In this paper, an effective method is introduced to detect COVID-19 infection based on the induction of changes in the DNA microarray gene expression pattern of the host and comparing it with DNA sequences (complementary DNA strands) of SARS-CoV-2. This methodology will also help to classify the COVID-19 infection by matching with DNA sequences of different Corona viruses and thus determines the severity of the infection. Moreover, this technique is also suitable for identification of asymptomatic carriers of SARS-CoV-2.

The contribution of the paper is as follows. Initially, a bi-direction string matching algorithm is used to analyse DNA microarray gene expression data and the analytical result is represented graphically. The shape of the graph corresponds to a DNA microarray gene expression is represented in terms of eight-directional chain code sequence. The infection of COVID-19 is detected based on the changes in DNA microarray gene expression of the patient during and after infection and comparing it with DNA sequences of SARS-CoV-2. Finally, the categorisation of Coronavirus infection is done based on the distribution probabilities of eight-directional chain code sequences correspond to DNA sequences of different Corona viruses by taking random samples from the GenBank.

The paper is organised in the following way. In Section 2, the analysis of DNA microarray gene expression patterns is illustrated and the analytical result is represented graphically. The shape of the graph corresponds to a DNA microarray gene expression pattern is represented in terms of eight-directional chain code sequence in Section 3. The methodology for detection of COVID-19 infection has been depicted in Section 4. Section 5 provides the method for classification of Coronavirus infection and we concluded in Section 6.

2. Analysis of DNA Microarray Gene Expression Pattern

A gene is primarily made up of sequence of triplets of the nucleotides (exons). Deoxyribonucleic acid (DNA) is a nucleic acid that contains the genetic instructions used in the development, reproducing and functioning of all living organisms. DNA strand comprises of a large sequence of nucleotides or bases. One nucleotide contains one nitrogen base, one sugar molecule (deoxyribose) and one phosphate. In general, four nitrogen bases are denoted by the first letter of the corresponding nitrogen base such as A for Adenine, C for Cytosine, G for Guanine and T for Thymine. A linear chain of DNA is paired to a complementary strand. The complementary property stems from the ability of the nucleotides to establish specific pairs A-T and G-C and the pair of complementary strands then forms the double helix. Therefore, each strand holds the entire information and the biochemical machinery guarantees that the information can be copied over and over again even when the original molecule has long since vanished.

DNA microarray represents the gene expression pattern and based on the genetic code the linear string of DNA is translated into amino acids, that is, a protein via mRNA (messenger RNA). For example, the DNA sequence GAACTACACACGTGTAAC codes for the amino acid sequence ELHTCN.

In this paper, an efficient pattern matching algorithm [12] is used initially to detect and count the number of occurrences of all possible patterns comprises of the four characters of the nitrogen bases A, C, G, and T taking three at a time including the repetitions in a given DNA microarray gene expression data. Hence, the total number of possible patterns is 64. For the sake of simplicity, three characters of nitrogen bases is considered at a time instead of taking four characters A, C, G, and T of nitrogen bases. In this algorithm, the text is scanned into chunks of n (n=3) characters at a time which is encoded to a unique number and the corresponding counter of the specific pattern is subsequently incremented. Thus, the number generated by each pattern acts as the key and the encoding is done according to a number system of base 4 since there are only 4 characters in DNA microarray gene expression data. This pattern matching algorithm is ideally suitable for handling huge amount of DNA microarray gene expression data since the complexity of the algorithm is significantly low.

The analytical result of DNA microarray gene expression data is represented graphically for better visibility and understanding. To represent graphically, each quadrant of a circle is partitioned into 16 fragments and a radius is drawn for each fragment since 64 patterns are possible by taking three at a time from four characters of nitrogen bases A, C, G, and T of gene expression data. The radius of the circle is fitted to a high scale value considering the maximum frequency value of patterns that may occur. The frequency value of each pattern of DNA microarray gene expression data is plotted on the corresponding radius and the graph is drawn by connecting all the corresponding points. For a given DNA microarray gene expression data, the analytical result is graphically depicted in the following figure.
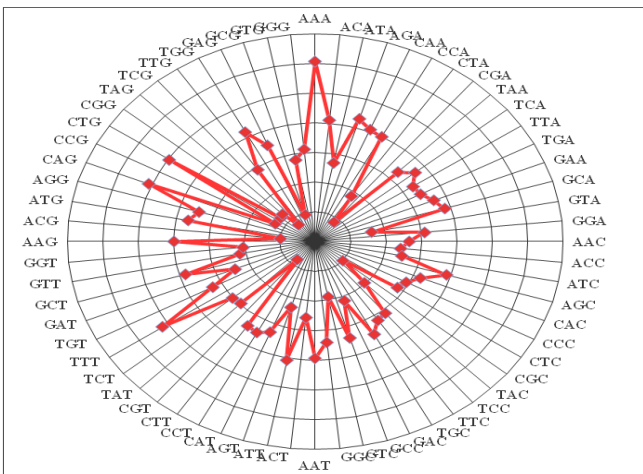

Figure 1: Graphical Representation of DNA Microarray Gene Expression Data

3. Eight –Directional Chain Code Sequence
In this Section, an eight-directional chain code sequence is introduced to determine the shape of the graph corresponds to a DNA microarray gene expression data which is invariant to translation, scaling and rotation. The edges of the graph (shape) are partitioned into line segments by the frequency value of each pattern of the DNA microarray gene expression data. The chain code of a line segment depends on the previous line segment and it is determined by counting the number of direction changes in a counterclockwise direction that separate two adjacent elements of the chain code.
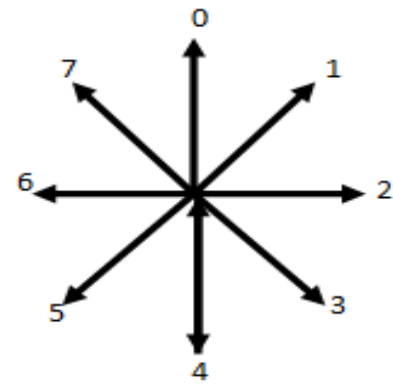

Figure 2: Eight-directional Chain Code

The chain code sequence of a DNA microarray gene expression data also depends on the starting point that can be normalized by treating it as a circular sequence of direction numbers and redefining the starting point. For example, the chain code for the line segment PQ in Fig. 3 is 2 since YP is the previous line segment and the direction has been changed twice from YP to PQ.
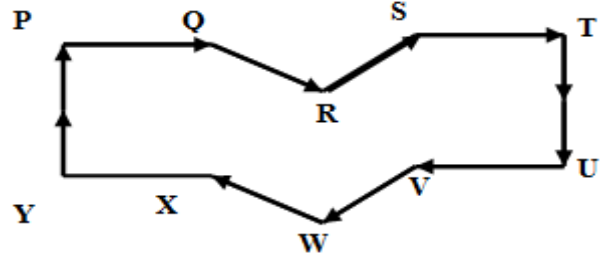

Figure 3: An Example of Chain Code

The chain code sequence for the DNA microarray gene expression data represented in Fig. 1 is shown below assuming AAA as the starting point.

2 0 6 2 0 2 0 4 0 4 6 0 0 2 6 2 0 6 0 2 0 0 2 4 0 2 0 2 4 4 6 0 2 6 2 6 0 2 2 4 2 6 4 0 4 4 0 6 2 6 2 6 4 4 4 6 2 4 0 4 0 4 2 0

For the sake of simplicity, if all 0's are removed from the chain code sequence then the sequence remains same since the graphical representation is invariant to scaling, rotation and scaling with rotation.

4. Detection of Coronavirus (COVID-19) Infection
SARS-CoV-2 has a single-stranded positive sense RNA genome that encodes 27 proteins including an RNA-dependent RNA polymerase (RdRP) and four

structural proteins. For maintaining genome fidelity, RdRP acts in conjunction with non-structural proteins. The four structural proteins of SARS-CoV-2 are the spike surface glycoprotein (S), small envelope protein (E), matrix protein (M), and nucleocapsid protein (N). In Coronavirus, the S gene codes for the receptor-binding spike protein which enables the virus to infect cells. The spike protein mediates receptor binding and membrane fusion which determines host tropism and transmission capabilities. The other three structural proteins are necessary for general functions of Coronavirus and involved in encasing the RNA and/or in protein assembly, budding, envelope formation and pathogenesis.

The line of attack of Coronavirus is roughly classified into six phases, namely, fusion/endocytosis, translation, proteolysis, translation and RNA replication, packing and Virion release. SARS-CoV-2 enters into cell by interacting with the angiotensin converting enzyme 2 (ACE2) in fusion/endocytosis phase. In translation phase, different proteins are prepared from viral RNA. These proteins are broken down into multiple small parts in proteolysis phase and multiple SARS-CoV-2 RNA are prepared from these small proteins in translation and RNA replication phase. In packing phase, cell membrane, spike and other necessary proteins are packed with SARS-CoV-2 RNA and in this way multiple Coronaviruses are produced. Finally, in virion release phase, Coronavirus attached with the human cell with the help of spikes, break through the cell membrane and released.

Transcription is the process through which ribonucleic acid (RNA) is formed from DNA (gene) and through translation amino acids are formed from RNA. There are sequences of nucleotides within the DNA that are spliced out progressively in the process of transcription and translation. To detect SARS-CoV-2 genetically, complementary DNA (cDNA) strands are generated from SARS-CoV-2 RNA by reverse transcription process. The first genome sequence of SARS-CoV-2 was conducted with metagenomic RNA sequencing and the sequence was added to the GenBank sequence repository on 10th January 2020. Since then, more than 1000 sequences have been made available on the GenBank by the researchers across the globe.

The biomedical steps that led to the specific cellular malfunction displayed in cells are to be identified for detection of COVID-19 infection. This process involve the transcription and translation regulatory system, cellular signaling pathways and metabolic pathways that control cell cycle, differentiation, apoptosis and motility. The expression of a gene provides a measure of 'how active' a specific gene under certain biochemical conditions and specific patterns of gene expression occur during different biological states. In this paper, COVID-19 infection is detected based on the induction of changes in the DNA microarray gene expression pattern of the host and comparing it with DNA sequences of Coronavirus (SARS-CoV-2). The graphical representation of DNA microarray gene expression data will facilitate to simultaneously monitor the expression of thousand of genes of the host which provides a molecular biomarker to differentiate normal cells from COVID-19 infected ones.
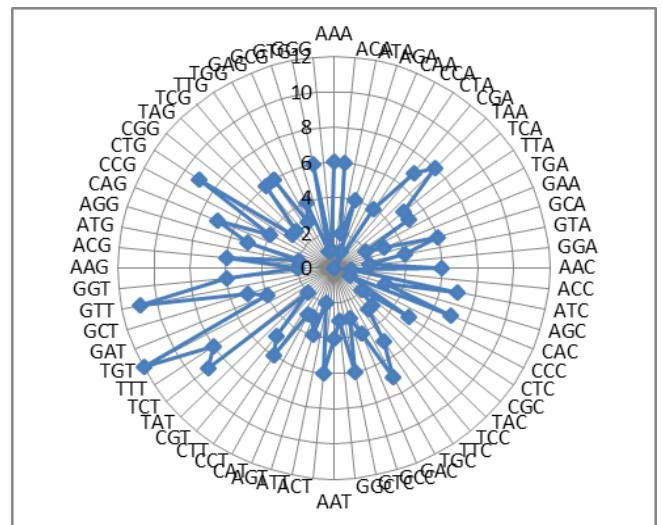

Figure 4: DNA Microarray Gene Expression of SARS-CoV-2

An example of DNA microarray gene expression of SARS-CoV-2 (MT594112.1) infected in India taken from the GenBank is shown below.

The corresponding chain code sequence is 6 1 2 6 2 0 6 1 3 6 2 4 0 4 0 5 4 5 4 4 4 4 2 6 0 4 0 5 3 6 0 2 6 2 0 6 2 2 4 2 6 4 5 0 3 0 6 2 6 0 4 4 4 6 6 2 2 5 3 4 4 6 1.

During COVID-19 infection, the changes in DNA microarray gene expression of the host are easily identified based on the distribution probability of the chain code sequence. The similarity with DNA sequences of SARS-COV-2 is measured by comparing the vector distance between the distribution probabilities of chain code sequences of the host and SARS-CoV-2. In this paper, the vector distance between chain code sequences is computed by using Jeffrey Divergence techniques. Jeffrey Divergence is a modification of Kullback-Leibler Divergence that compares two empirical distributions and the formula for Jeffery Divergence is as follows:

$$D(I, J) = \sum f_i(I) \log(f_i(I)/f_i + f_i(J) \log f_i(J)/f_i,$$

where $f_i = [f_i(I) + f_i(J)]/2$ and I and J are two different DNA microarray gene expression data.

5. Classification of Coronavirus Infection

The classification of Coronavirus infection plays an important role to determine the severity of the infection and thus helps in appropriate care and treatment. Moreover, it will be helpful for forecasting mortality rate, rate of infection and other issues related to COVID-19. In this work, the classification of Coronavirus infection is done based on similarities between the distribution probabilities of different chain code sequences representing DNA microarray gene expression patterns of different Corona viruses. The similarity is measured by comparing the vector distances between distribution probabilities of different chain code sequences taking two at a time. In this paper, one random sample of DNA microarray gene expression of SARS-CoV-2 infected in each of the Countries China, USA, Italy and India has been taken from the GenBank and their corresponding chain code sequences are listed below.

India: 6 1 2 6 2 0 6 1 3 6 2 4 0 4 0 5 4 5 4 4 4 4 4 2 6 0 4 0 5 3 6 0 2 6 2 0 6 2 2 4 2 6 4 5 0 3 0 6 2 6 0 4 4 4 6 6 2 2 5 3 4 4 6 1

China: 5 3 0 1 5 2 6 2 6 4 4 4 4 4 0 5 3 6 4 4 4 4 4 4 4 2 2 0 4 4 6 0 2 0 0 2 4 4 0 4 4 4 4 5 0 3 1 4 4 5 0 4 4 4 4 0 4 0 0 0 4 4 5

Italy: 5 3 1 5 2 2 4 0 0 4 4 4 4 0 4 0 6 2 6 4 4 4 4 4 4 0 0 2 2 0 4 4 6 0 3 5 1 3 4 4 0 4 4 4 4 6 0 2 0 6 3 5 0 3 5 4 6 6 2 2 5 3 4 4

USA: 5 2 1 6 4 0 4 0 0 4 6 2 4 0 4 0 5 3 6 4 4 4 4 4 4 0 0 1 3 6 6 4 6 0 3 5 2 0 6 4 0 4 3 6 3 5 0 3 0 6 3 5 0 4 4 4 6 6 2 2 5 3 4 4

The divergence between distribution probabilities of different countries India, China, Italy and USA taking two at a time are listed in the following.

Table 1: Distribution Probabilities of different categories

	0	1	2	3	4	5	6	7
India	0.15625	0.046875	0.1875	0.0625	0.25	0.078125	0.21875	0
China	0.203125	0.03125	0.09375	0.046875	0.46875	0.09375	0.0625	0
Italy	0.1875	0.03125	0.125	0.09375	0.359375	0.09375	0.109375	0
USA	0.203125	0.03125	0.078125	0.125	0.296875	0.09375	0.171875	0

The divergence between distribution probabilities of different countries India, China, Italy and USA taking two at a time are listed in the following.

Countries	Divergence	Value
India, China	D(India, China)	0.245
India, Italy	D(India, Italy)	0.246
India, USA	D(India, USA)	0.532
China, Italy	D(China, Italy)	0.029
China, USA	D(China, USA)	0.069
Italy, USA	D(Italy, USA)	0.021

From the above result, it is evident that Coronavirus infection that spread out in China, Italy and USA are of very similar category. It is also clear that the characteristics of Coronavirus infection that spread out in India are dissimilar from other countries China, Italy and USA.

6. Conclusion

In this paper, an effective method is introduced for detection and classification of COVID-19 infection based on genome sequences. The COVID-19 infection is detected based on the induction of changes in the DNA microarray gene expression pattern of the host during and after infection and comparing it with DNA sequences of SARS-CoV-2. An approach for classification of Coronavirus infection is also provided at the end of the work which will be helpful for forecasting rate of mortality, rate of infection, severity of the infection and other issues related to COVID-19. However, a limited number of random samples from a limited number of countries, i.e., India, China, Italy and USA have only been taken for categorization of COVID-19 infection. Moreover, symptoms of the patients, other environmental and related factors are not considered in this work for detection of COVID-19 infection.

In future scope, other factors in gene expression data may be considered to achieve better result in detection of COVID-19 infection. Moreover, an appropriate sample set may be taken in categorization of COVID-19 infection for better outcome. Further, a model may be developed to establish a relationship among the category of the Coronavirus infection, the rate of mortality, the rate of infection, severity of the infection and other relevant issues related to COVID-19.

References

[1.] C. Anastassopoulou, L. Russo, A. Tsakris, C. Siettos, "Data-Based Analysis, Modelling and Forecasting of the COVID-19 Outbreak", https://doi.org/10.1101/2020.02.11.2002218 6, 2020.

[2.] F. Azuaje, "Gene expression Patterns and Cancer Classification: A Self-Adaptive and Incremental Neural Approach", Journal of IEEE Engineering in Medicine and Biology, pp 308 -313, 2000.

[3.] D. Barash, D. Comaniciu, "Meanshift Clustering for DNA Microarray Analysis", Proceedings of the IEEE Conference on Computational Systems, pp 578-579, 2004.

[4.] J. Chen et. Al., "Clinical Study of Mesenchymal Stem Cell Treatment for Acute Respiratory Distress Syndrome Induced by Epidemic Influenza A (H7N9) Infection: A Hint for COVID-19 Treatment", https://doi.org/10.1016/j.eng.2020.02.006 , April 2020.

[5.] R. Jana, C. Ray, "Image Registration using Object Shape's Chain Code", Proceedings of 2nd International Congress Image and Signal Processing, pp 1-5, ISBN: 978-1-42444129-7, 2009.

[6.] X. Li, M. Geng, Y. Peng, L. Meng, S. Lu, "Molecular Immune Pathogenesis and Diagnosis of COVID-19", Elsevier Journal of Pharmaceutical Analysis, 2020.

[7.] F.J. Lopez, M. Cuadros, A. Blanco, A. Concha, "Unveiling Fuzzy Associations between Breast Cancer Prognostic Factors and Gene Expression Data", Proceedings of

20th International Workshop on Database and Expert Systems Application, pp 338-342, 2009.

[8.] M.A. Marra, S. J. M. Jones et al, "The Genome Sequence of the SARS-Associated Corona Virus", www.sciencemag.org, SCIENCE Vol. 300, pp 1399-1404, April 2020.

[9.] D. McGonagle, K. Sharif, A. Oregan, C. Bridgewood, "The Role of Cytokines including Interleukin-6 in COVID-19 Induced Pneumonia and Macrophage Activation Syndrome-Like Disease", Elsevier Journal of Autoimmunity Reviews, March-April 2020.

[10.] M. T. U. Qamar, S. M. Alqahtani, M.A. Alamri, L.L. Chen, "Structural Basis of SARS-CoV-2 3CLpro and Anti-COVID-19 Drug Discovery from Medicinal Plants", Journal of Pharmaceutical Analysis, March 2020.

[11.] C. Ray, "Cancer Identification and Gene Classification using DNA Microarray Gene Expression Patterns", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 2, pp 155-160, March 2011.

[12.] C. Ray, S.Tripathi, A. Chatterjee, A. Das, "An efficient Bi-directional String Matching Algorithm for Statistical Estimation", Proceedings of International Symposium on Data, Information, and Knowledge Spectrum, INDIA, pp 73-79, December 2007.

[13.] P.J. Russel, "Fundamentals of Genetics", Second Edition, San Francisco, Addison Wesley Longman Inc., 2000.

[14.] P.K. Sethy, S. K. Behera, "Detection of Coronavirus Disease (COVID-19) based on Deep Features", www.preprints.org, March 2020.

[15.] Y.W. Tang, J.E. Schmitz, D. H. Persing, C.W. Stratton, "The Laboratory Diagnosis of COVID-19 Infection: Current Issues and Challenges", http://jcm.asm.org, April 2020.

[16.] B.Udugama et. Al., "Diagnosing COVID-19: The Disease and Tools for Detection", www.acsnano.org, 2020.

[17.] S. Volinia, N. Mascellani, J. Marchesini, A. Veronese, E. Ormondroyd, H. Alder, J. Palatini, M. Negrini, C. M. Croce, "Genome wide Identification of Recessive Cancer Genes by Combinatorial Mutation Analysis", www.plosone.org,Vol. 3, Issue 10, pp 1-13, 2008.

[18.] L. Wang, Y. Wang, D. Ye, Q. Liu, "Review of the 2019 novel Coronavirus (SARS-Cov-2) based on Current Evidence", Elsevier International Journal of Antimicrobial Agents, April 2020.

[19.] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng, B. Xu, "A Deep Learning Algorithm using CT Images to Screen for Corona Virus Disease (COVID–19)", https://doi.org/10.1101/2020.02.14.2002302 , 2020.

[20.] Organization WH. WHO Statement Regarding Cluster of Pneumonia Cases in Wuhan, China; Available from: https://www.who.int/china/news/detail/09-01-2020-who-statement-regarding-cluster-of-pneumonia-cases-in-wuhan-china, 2020.

[21.] Canrong Wu et. Al., "Analysis of Therapeutic Targets for SARS-CoV-2 and Discovery of Potential Drugs by Computational Methods", http://creativecommons.org/licenses/by-nc-nd/4.0, February 2020.

[22.] D. Wu, T. Wu, Q. Liu, Z. Yang, "The SARS-CoV-2 Outbreak: What We Know", Elsevier International Journal of Infectious Diseases, pp 44-48, 2020.

[23.] B. Yong, L. Owen, "Dynamical Transmission Model of MERS-CoV in Two Areas", Proceedings of AIP conference on Application of Mathematics in Industry and Life, pp 1-7, February 2016.

[24.] F. Zhang, O. O. Abudayyeh, J.S. Gootenberg, "A Protocol for Detection of COVID-19 using CRISPR Diagnostics", v.20200321.