International Journal on Emerging Technologies 11(3): 1157-1163(2020)

ISSN No. (Print): 0975-8364
ISSN No. (Online): 2249-3255

In Silico based Whole Genome Phylogenetic Analysis of Novel Coronavirus (SARS-CoV-2)

Raghunath Satpathy
School of Biotechnology, Gangadhar Meher University,
Amruta Vihar, Sambalpur (Odisha), India.

(Corresponding author: Raghunath Satpathy)
(Received 17 April 2020, Revised 22 June 2020, Accepted 23 June 2020)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Novel coronavirus (SARS-CoV-2) began in the Wuhan city in China that spread worldwide in the pandemic manner and create a global concern in recent times. It is also associated with significant global economic losses. However, the source of origin of this novel virus is still a debate. In this work, a genomic perspective has been taken to study the phylogeny of the whole virus genome by retrieving the recent available genomic data from the NCBI database. In the first part of the work, the published genome sequence of Wuhan seafood market pneumonia virus isolates Wuhan-Hu-1 complete genome was retrieved and corresponding homologous sequences were obtained by BLAST N tool. Further Clustal Omega was used to multiple sequence alignment (MSA) study followed by phylogenetic analysis by MEGAX software. The whole genome-based phylogenetic study reveals about the origin of the Wuhan-Hu-1 virus. The Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1 genome is showing an evolutionary relationship with Bat Corona Virus RaTG13 genome sequence that shares 96.12 % of sequence identity. Further 31 numbers of the country-wise whole genome of SARS-CoV-2 were retrieved from the NCBI server (available genome sequence up to 15th May 2020) and analysed to establish the phylogenetic relationship as well as to predict the probable spreading pattern among the strains.

Keywords: Novel coronavirus (SARS-CoV-2), phylogenetic analysis, whole genome analysis, sequence alignment, comparative genomics

I. INTRODUCTION

Human coronaviruses belong to the order Nidovirales, family Coronaviridae, and genus Alpha coronavirus or Beta coronavirus [1]. The emerging SARS-CoV-2, a beta coronavirus, can cause COVID-19, officially named by the World Health Organization (WHO) on February 11, 2020. SARS-CoV-2 is highly infectious; the entire population is generally susceptible, and respiratory droplets and contact are the main routes of transmission. It is generally believed that the incubation period is between 3 to 7 days on average, with 1 day as the shortest and 14 days longest. SARS-CoV-2 is highly contagious and the current outbreak of COVID-19 remains severe in China and has been designated as a Public Health Emergency of International Concern by the WHO [2-4]. In December 2019, Wuhan city of, China, one was observed as the centre of an outbreak of novel pneumonia of Unknown cause. And it was further confirmed by the Chinese health authorities that the cause of the disease is due to the novel coronavirus (COVID-19) [5, 6]. Recently many cases of COVID-19 infection have been reported in many countries. And it is expected that the out break is due to the *cross-species* phenomena of the animal coronavirus. Unfortunately, till today no specific effective drugs (and/or) vaccines have been reported. However, due to advancement of sequencing technologies, complete genomic sequences of these viruses have been made available to the databases such as NCBI nucleotide database also the subsequent research update details can be obtained from https://www.nih.gov/health-information/coronavirus [7-8].

The availability of complete genome sequence data for these deadliest viruses provides the opportunity for the use of such data for phylogenetic reconstruction. As the complete genome sequences contain phylogenetic information at several levels hence the relationship among other viruses and their outbreak strategies can be predicted. Therefore, the research on the origin of this novel virus will be an asset to the scientific community for further analysis. The new genotypes of the virus have been evolved by the rapid evolution in the gene sequences. So, the sequencing of whole genomes, genes or gene fragments is more and more commonly used for understanding the outbreak of the disease. By constructing the phylogenetic profile by considering only gene sequences may not predict accurate result because the gene might have undergone different mutation process such as deletion, duplication, horizontal gene transfer and so on. Therefore, one of the most effective methods of using is the whole genome sequence information. Another advantage is this provides the opportunities for the construction of whole genome phylogenetic trees, lead to the characterization of the novel virus strains. In addition to this, it leads to elucidate the virus-host interactions and cross-species transmission mechanism followed by the understanding of the evolutionary rates and ultimately predict a clear and concrete estimation of epidemiological parameters associated with the disease [9-11].

In this paper, we analyzed the phylogeny of the novel corona 2019 virus by considering the whole genome and efforts have been made to establish the relationship among the novel virus and their spreading pattern.

II. MATERIALS AND METHODS

A. Retrieval of whole genome sequence data

Wuhan seafood market pneumonia virus isolates Wuhan-Hu-1, complete genome was taken from NCBI having GENBANK ID MN908947.3 and having 29903 bpss-RNA linear DNA and sequences were available from 18th March 2020 [12].

B. Phylogenetic analysis

Homologous genomic sequence was retrieved by BLASTN algorithm available at (https://blast.ncbi.nlm.nih.gov/) by considering the default parameter [13]. 100 homologous sequences were retrieved based on Expectation (E) value, query coverage and identity percentage value. Further, the multiple sequence alignment was conducted by using the Clustal omega program available at EBI server (European Bioinformatics Institute) server (https://www.ebi.ac.uk/Tools/msa/clustalo). Clustal Omega is a new version of multiple sequence alignment (MSA) program that uses Hidden Markov Model (HMM) profile-profile comparison techniques for the alignments between three or more sequences and generates a guide tree [14]. The phylogenetic tree was constructed by using the Unweighted Pair Group Method with Arithmetic mean (UPGMA) method with 100 bootstraps implemented in MEGA X(Molecular Evolutionary Genetics Analysis), across computing platforms [15]. Version 10 of the MEGA software enables cross-platform use, running natively on Windows and Linux systems (https://www.megasoftware.net/).Further Wg-VISTA server (http://genome.lbl.gov/vista/wgvista/about.shtml) was used for comparative annotation about ORF positioning and conserved non-coding of the genome sequences related to Wuhan-Hu-1, complete genome.

Tracing of phylogenetic relationship among the Sars Co-V2 genomes obtained from different countries

The latest genome sequences from the NCBI (https://www.ncbi.nlm.nih.gov/genbank/sars-cov-2-seqs/) in the form of *one sequence to one country*, because almost of the sequence of a country shares more than 90% similarity with each other [9, 16]. Therefore, randomly sequences were selected from the countries where whole genome sequences are available in the database. All total of 31 numbers of whole genome sequences were obtained from the NCBI till available up to 15-05-2020. MSA was performed by Clustal Omega tool available online and phylogenetic tree was built by UPGMA method implemented in the MEGA X tool as stated in the above section with 1000 bootstrap analysis.

III. RESULTS AND DISCUSSIONS

All the severe acute respiratory syndrome coronavirus 2 isolates in 2019 and 2020 are showing a common origin and show a close relationship with the Bat Corona Virus RaTG13 genome (Fig. 1). Also, the base composition of the genomes were computed and it was observed that they are distributed in almost similar manner, shown in the Table 1.

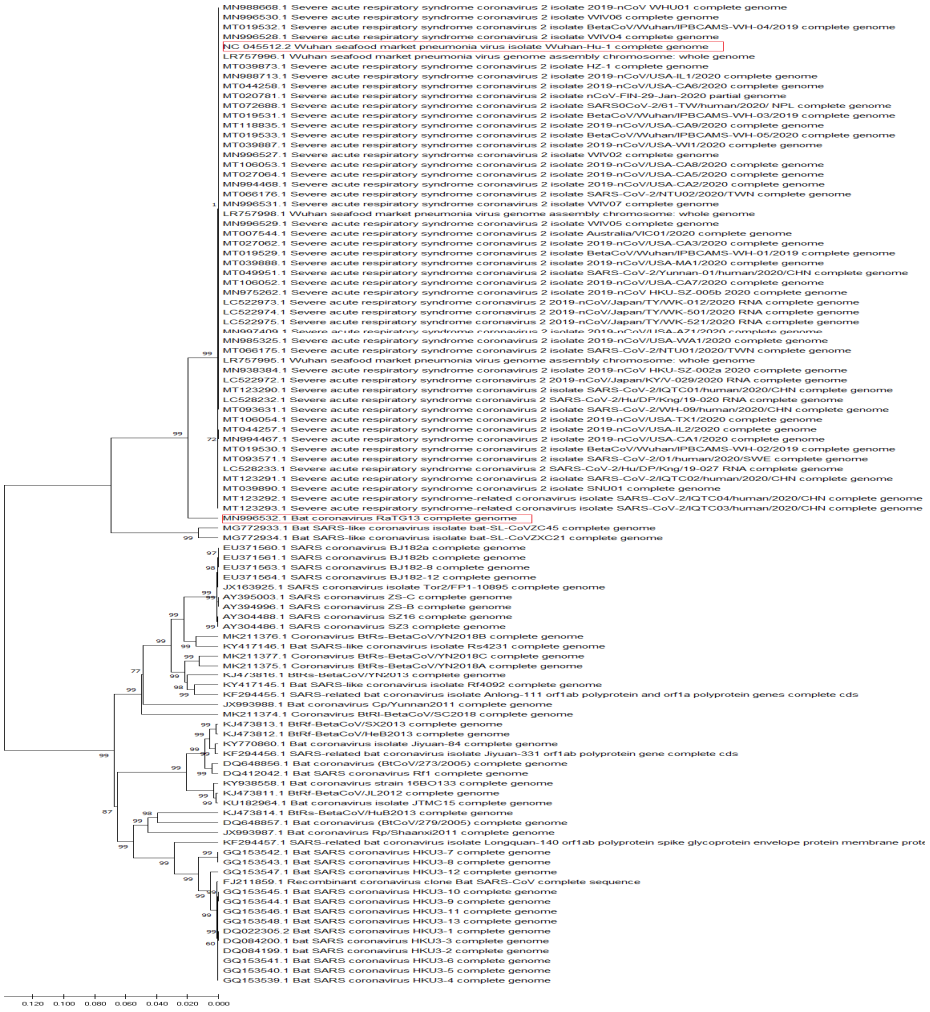

Fig. 1. Showing phylogenetic tree constructed by UPGMA method with highlighted genomes of Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1 complete genome and Bat coronavirus RaTG13.

Table 1: Showing the comparison of genome composition of two genome.

S. No.	Sequence details	% T(U)	% C	% A	% G	Total base pairs
1.	Bat coronavirus RaTG13 complete genome	32.0	18.4	29.9	19.5	29855
2.	Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1 complete genome	32.0	18.3	29.9	19.6	29903
3.	Average of all the 100 sequences considered for the study	31.4	19.1	29.2	20.2	29594

The Bat coronavirus RaTG13 complete genome shares 99% query coverage and 96.12 % of sequence identity with the query sequence obtained from the BLAST N output. Similarly, Bat SARS-like Corona Virus isolates show a common cluster and having and having 94-95% of query coverage and having 88-89% sequence identity was observed. All the Bat SARS coronavirus HKU types shows a common origin from their genome analysis and shares 88-89% of query coverage and 80-82% of sequence identity with the novel corona virus Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1 (Fig. 1). The evolutionary origin of the novel Corona Virus 2019 and Bat coronavirus RaTG13 genome has been obtained in the present phylogenetic analysis and highlighted red in Fig. 1. Further, comparative genome analysis of the above two genome was performed by using the https://genomevolution.org/coge/GEvo.pl server and the program uses the BLAST Z algorithm to generate the alignment results with default parameters shown in Fig. 2 [17]. The Fig. 2 indicates about the difference in the available gene structures and orders about the two genomes. A recent study establishes the fact that the Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1 contains, a variable number of open reading frames (ORFs). About two-thirds of this genomic RNA, is located in the first ORF (ORF1a/b) translates two polyproteins, pp1a and pp1ab, and encodes 16 non-structural proteins (NSP), while the remaining ORFs encode accessory and structural proteins. The remaining part of virus genome encodes four important structural proteins, viz. *small envelope (E) protein, spike (S),* nucleocapsid (N) protein, *glycoprotein, matrix (M) protein*, including other accessory proteins that interfere with the host innate immune response. However, the Bat Corona virus RatG13 contains proteins like orf1abpolyprotein, gene S (code for spike glycoprotein), non-structural proteins like NS3, NS6, NS7a, NS 7b, NS 8, envelope protein, membrane protein, and nucleo capsid protein. [18, 20]. Despite of their variability in the genomic composition of the two selected genomes, the gene orders were obtained as similar in nature. Further annotation of the two genomes were carried out by VISTA server (http://genome.lbl.gov/vista/wgvista/about.shtml) whole genome alignment pipeline. To compute the conserved non-coding sequences (CNS) region among two selected genomes (Table 2).

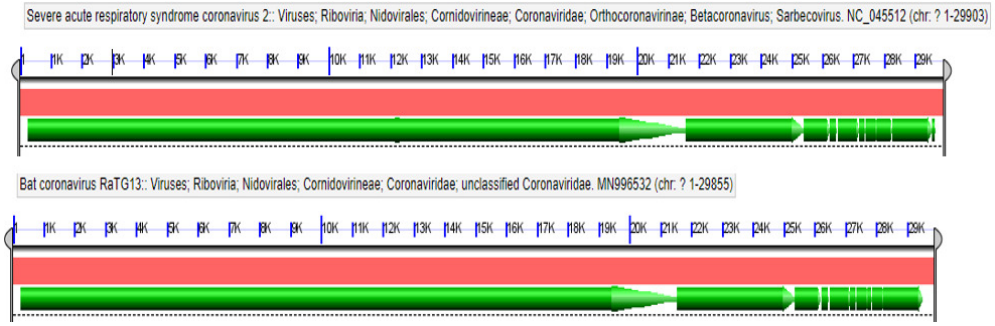

Fig. 2. Showing alignment of the two genome: ORFs has been represented in arrow.

Table 2: Showing the Conserved Non-coding sequences between two genome obtained from Wg-Vista server.

S. No.	Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1 complete genome (Accession: MN_908947.3)		Bat coronavirus RaTG13 complete genome (Accession : NC_045512.2)		No. of base (bp)
	From	To	From	To	
1.	1	250	16	265	250
2.	251	21537	266	21555	21290
3.	21545	25354	21563	25384	3822
4.	25363	26190	25393	26220	828
5.	26215	26442	26245	26472	228
6.	26493	27158	26523	27191	669
7.	27169	27354	27202	27387	186
8.	27360	27853	27394	27887	494
9.	27860	28225	27894	28259	366
10.	28240	29499	28240	29533	1260
11.	29524	29640	29558	29674	117
12.	29641	29855	29675	29890	216

Table 2 shows the Conserved non-coding sequences (CNS) available in two genomes. The CNS is basically the non-coding (functionless) DNA sequence distributed in the genome and show high similarity in the base sequence between the related species. In this case, the comparative genomic analysis indicated 12 no. of position wise distribution of this CNS, indicated about the molecular closeness of the two genomes.

Similar to the case for other SARS-CoV, the bat might be the probable origin for SARS-CoV-2 because of SARS-CoV-2 shares 96% whole-genome with a bat coronavirus (CoV), as supported by the recent research in addition to our phylogeny based study and sequence identity [20, 21]. It also supports the fact that, the novel coronavirus is not a laboratory constructed virus but originated from Bat coronavirus by some evolutionary mechanism to be established. Since any kind of wildlife like a bat can act as reservoirs/host of virus can give rise to the source of new emerging diseases by crossing the interspecies barrier [22-24]. The recent emergence of the novel coronavirus (SARS-CoV2) can be considered as an important phenomena to study the mechanism of origin and surveillance of the virus in the wild animals, also the relevant molecular mechanism of crossing the interspecies barrier.

Table 3: Showing the country wise retrieved sequences up to 15th May 2020.

S. No.	Accession No.	Name Of The Complete Genome	Length (bp)	Date of Sequence Release	Country
1.	MN908947	SARS-CoV-2 Wuhan-Hu-1, complete genome.	29903	30-MAR-2020	China
2.	MT419860	SARS-CoV 2/human/USA/CA-UCSF-UC36/2020	29841	01-MAY-2020	USA
3.	MT447175	SARS-CoV-2/human/THA/SI206587-NT/2020, complete genome.	29828	08-MAY-2020	Thailand
4.	MT439597	SARS-CoV-2/human/IND/nimh-3970/2020	29812	07-MAY-2020	India
5.	MT428553	SARS-CoV-2/human/KAZ/NCB-3/2020, complete genome.	29903	05-MAY-2020	Kazakhstan
6.	MT396266	SARS-CoV-2/mink/NLD/1/2020, complete genome.	29880	28-APR-2020	Nederland
7.	MT374116	SARS-CoV-2/human/TWN/CGMH-CGU-20/2020, complete genome.	29901	24-APR-2020	Taiwan
8.	MT372483	SARS-CoV-2/human/MYS/186197/2020	29481	23-APR-2020	Malaysia
9.	MT371574	SARS-CoV-2/human/CZE/CzechiaMotol_1967/2020	29756	23-APR-2020	Czech Republic
10.	MT371050	SARS-CoV-2/human/LKA/COV486/2020	29903	23-APR-2020	Sri Lanka
11.	MT365032	SARS-CoV-2/human/HKG/HKU-904a/2020	29887	05-MAY-2020	Hong Kong
12.	MT320891	SARS-CoV-2/human/IRN/HGRC-1.1-IPI-8206/2020.	29822	16-APR-2020	Iran
13.	MT320538	SARS-CoV-2/human/FRA/KRAROB/2020.	29882	24-APR-2020	France
14.	MT304476	SARS-CoV-2/human/KOR/BAACH_2719/2020	29882	07-APR-2020	South Korea
15.	MT292582	SARS-CoV-2/human/ESP/Valencia26/2020	29782	06-APR-2020	Spain
16.	MT256924	SARS-CoV-2/human/COL/79256_Antioquia/2020	29782	13-APR-2020	Colombia
17.	MT192773	SARS-CoV-2/human/VNM/nCoV-19-02S/2020	29890	06-APR-2020	Viet Nam
18.	MT126808	SARS-CoV-2/human/BRA/SP02/2020,	29876	06-APR-2020	Brazil
19.	MT093571	SARS-CoV-2/human/SWE/01/2020,	29886	06-APR-2020	Sweden
20.	MT077125	SARS-CoV-2/human/ITA/INMI1/2020,	29785	11-APR-2020	Italy
21.	MT072688	SARS-CoV-2/human/NPL/61-TW/2020,	29811	06-APR-2020	Nepal
22.	MT020781	SARS-CoV-2 nCoV-FIN-29-Jan-2020	29806	17-MAR-2020	Finland
23.	LC542976	SARS-CoV-2 nCoV-TKYE6968_2020	29903	22-APR-2020	Japan
24.	MT240479	SARS-CoV-2/human/PAK/Gilgit1/2020	29836	06-APR-2020	Pakistan
25.	MT324062	SARS-CoV-2/human/ZAF/R03006/2020	29903	13-APR-2020	South Africa
26.	MT276597	SARS-CoV-2/human/ISR/ISR_JP0320/2020	29851	06-APR-2020	Israel
27.	MT263074	SARS-CoV-2/human/PER/Peru-10/2020	29856	06-APR-2020	Peru
28.	MT394864	SARS-CoV-2/human/DEU/NRW-53/2020	29782	28-APR-2020	Germany
29.	MT327745	SARS-CoV-2/human/TUR/ERAGEM-001/2020	29832	13-APR-2020	Turkey
30.	MT476385	SARS-CoV-2/human/BGD/CHRF_0001/2020	29902	15-MAY-2020	Bangladesh
31.	MT459979	SARS-CoV-2/human/SRB/Novi Pazar-363/2020	29782	13-MAY-2020	Serbia

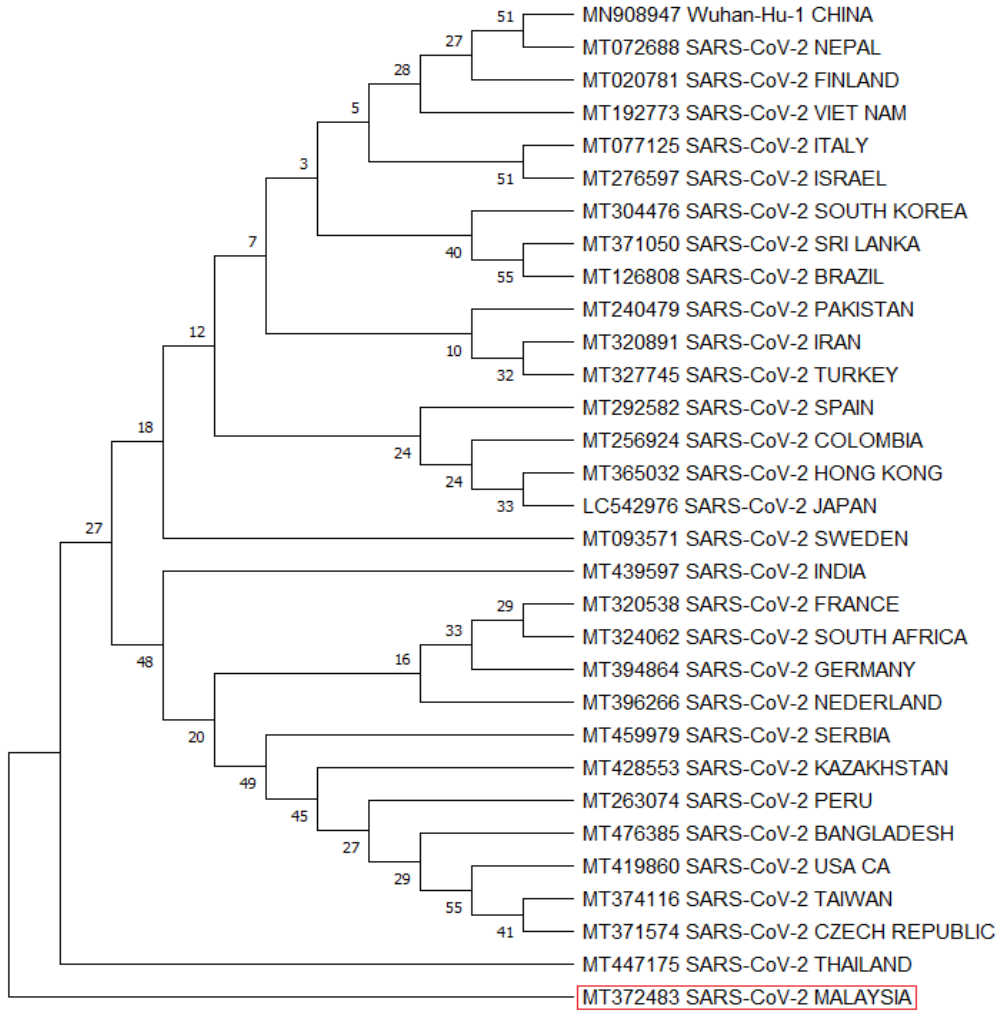

Fig. 3. Showing phylogenetic tree of the selected country-wise viral strains and out group is highlighted.

Phylogenetic Tree analysis of the country –wise selected strains) indicates about the world wide relationship between the strains of the virus as shown in the Fig. 3. From the top of the phylogenetic tree, there is a close relationship with the strains between Iran and Turkey was observed that is related to the viral strain of Pakistan. Also, Italy and Israel strains are closely related. Next, there exists a close relationship with the strains of China and Nepal, to which the Finland and Viet Nam strains are related. The strains of Japan and Hong Kong shows a close evolutionary relationship, to which the strains of Columbia and Spain are related. All to the above the Sweden strain is showing its relatedness. In the next cluster, the France and South-African strains show a good relationship to which Germany and Netherland are related. Similarly, the strains of USA-CA and Czech Republic is showing the close relationship in a cluster, further, these are related to Taiwan, Kazakhstan, Bangladesh, Peru, Serbia are related and to which the Indian strain is related. The viral strain of Thailand is showing common ancestor to all the above sequence and the strain of Malaysia is projected as the out-group that indicated that this particular strain is quite different in the origin respect to the other selected strains.

Due to lack of availability of sufficient whole-genome sequence data of the virus in the database for all the infected countries, although reported for ~ 213 countries till date, only selected, 31 strains (one strain per one country) were analysed in this work. The severe pandemic situations that are associated with the spread of novel coronavirus is evidenced by the whole genome phylogram (Fig. 3). It was also evidenced that the spreading of the virus occurring very rapidly across the globe as most of the genome sequences accounts for its evolutionary relationship with geographically closely placed and distant countries also [25-27]. In addition to this, it was also noticeable that, the spread of the virus within the countries were evident as the selected sequences of the viral strains country sequences matched with the sequences of the other countries (Fig. 4). Such an intercontinental transmission might be possible because of the increased migration event of people across the world [28]. Similarly, the probable origin of the virus from China (Wuhan) was clustered across mostly in case of Asian and some of the European Countries and therefore supports the fact that China might be a hub for international SARS-CoV-2 virus entry as well as the exchange of the same [29, 30]. The novel coronaviruses tend to change their genome due to mutation, probably at a fairly high rate, therefore allows to cluster the virus into more closely related viruses and distantly out-group. All the selected sequences in this work are closely related, but obtaining the out-group for the Malaysian strain may be due to independent mutation to be analysed further. The phylogenetic tree based analysis of the genomic sequences might be used to reconstruct the

transmission tree of the virus. Since there are only 8 mutations has been reported recently in a ~30, 000 base pairs of SARS-CoV-2 virus sequence, the information predict that the probable origin of the virus has been originated between mid-November and early December 2019 [31-33]. Since the flow of SARS –CoV-2 sequence information from the country-wise genome sequencing experiment across the globe is going updated in the database so more interesting fact can be deduced in future by analysing these data.

IV. CONCLUSION AND FUTURE PROSPECTS

In the first part of the study, we presented the computer-based phylogenetic analysis of a novel coronavirus, Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1 complete genome. UPGMA based phylogenetic tree was constructed by considering 100 whole-genome homologous sequences after the multiple sequence alignment. The study establishes a clear-cut genomic origin of these viruses from the Bat coronavirus RaTG13, also the fact was supported by the recent research. Further analysis of the whole genome comparison shows the arrangement of gene order and ORF position discuss their relationships at a deep level. It can be hypothesised that the novel coronavirus (SARS-CoV-2) and their hosts are subject to evolutionary pressure. So further study is warranted for the individual protein level study to analyse to compute evolutionary selection pressure for better understanding of the evolutionary mechanism that takes part in the evolution of the novel virus. In addition to this (protein) structural level analysis is to be done to analyse the effect of the mutation in the protein structure, for a better understanding of the drug molecule binding efficacy. In the second part of the study, the phylogeny of the selected country based viral strain available in the database up to 15th May 2020 indicated about the evolution of the selected viral strains are in an agreement with diversity and homology. As the pandemic nature of the virus will more likely be a severe threat to mankind, therefore, the phylogenetic tree based analysis of sequences would help us to understand the flow of severity and transmission of the virus across the globe. It has been evidenced that the SARS Co V2 virus entered and distributed more prominently from China followed by Asian countries and subsequently followed by other European foreign countries. However, the unique origin of the Malaysian Strain is to be studied thoroughly to understand its origin. Not only the genomic analysis, further study about the protein sequence and structure can also be implemented to identify the conserved amino acids of host receptor that are involved for the recognition by SARS-CoV virus spike protein, that regulate the cross-species transmissions of SARS-CoV. Therefore, future study about the comparative host protein receptor sequences might be fruitful to uncover the key amino acid residues that interact in the receptor in many species. So that, it will lead to providing more possibility of finding alternative /intermediate hosts for the origin of this novel virus.

Conflict of Interest. No.

ACKNOWLEDGEMENT

I am thankful to administration of Gangadhar Meher University, Sambalpur for the encouragement and support.

REFERENCES

[1]. Hui, D. S., Azhar, E. I., Madani, T. A., Ntoumi, F., Kock, R., Dar, O., & Zumla, A. (2020). The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health. The latest 2019 novel coronavirus outbreak in Wuhan, China. *International Journal of Infectious Diseases*, *91*, 264-266.

[2]. Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases*. https://doi.org/10.1016/S1473-3099 (20)30120-1.

[3]. Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., & Xing, X. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England Journal of Medicine*.. https://www.nejm.org/doi/10.1056/NEJMoa2001316.

[4]. Vashisht G. & Prakash R (2020). Predicting the Rate of Growth of the Novel Corona Virus 2020, *International Journal on Emerging Technologies,11*(3): 19-25.

[5]. Munster, V. J., Koopmans, M., van Doremalen, N., van Riel, D., & de Wit, E. (2020). A novel coronavirus emerging in China—key questions for impact assessment. *New England Journal of Medicine*, *382*(8), 692-694.

[6]. Zhu, H., Wang, L., Fang, C., Peng, S., Zhang, L., Chang, G., & Zhou, W. (2020). Clinical analysis of 10 neonates born to mothers with 2019-nCoV pneumonia. *Translational pediatrics*, *9*(1), 51. https://doi.org/10.21037/tp.2020.02.06.

[7]. Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J. ... & Niu, P. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine*. DOI: 10.1056/NEJMoa2001017

[8]. Hong, H., Wang, Y., Chung, H. T., & Chen, C. J. (2020). Clinical characteristics of novel coronavirus disease 2019 (COVID-19) in newborns, infants and children. *Pediatrics & Neonatology*. doi: 10.1016/j.pedneo.2020.03.001.

[9]. Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H. ... & Bi, Y. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, *395*(10224), 565-574.

[10]. Ren, L. L., Wang, Y. M., Wu, Z. Q., Xiang, Z. C., Guo, L., Xu, T., Jiang, Y. Z., Xiong, Y., Li, Y. J., Li, X. W., Li, H., Fan, G. H., Gu, X. Y., Xiao, Y., Gao, H., Xu, J. Y., Yang, F., Wang, X. M., Wu, C., Chen, L., … Wang, J. W. (2020). Identification of a novel coronavirus causing severe pneumonia in human: a descriptive study. *Chinese medical journal*, *133*(9), 1015–1024. https://doi.org/10.1097/CM9.0000000000000722

[11]. Herniou, E. A., Luque, T., Chen, X., Vlak, J. M., Winstanley, D., Cory, J. S., & O'Reilly, D. R. (2001). Use of whole genome sequence data to infer baculovirus phylogeny. *Journal of virology*, *75*(17), 8117-8126.

[12]. Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2011). GenBank. *Nucleic acids research*, *39*(Database issue), D32.

[13]. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, *215*(3), 403-410.

[14]. Sievers, F., & Higgins, D. G. (2018). Clustal Omega for making accurate alignments of many protein sequences. *Protein Science*, *27*(1), 135-145.

[15]. Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution*, *35*(6), 1547-1549.

[16]. Ceraolo, C., & Giorgi, F. M. (2020). Genomic variance of the 2019-nCoV coronavirus. *Journal of medical virology*, *92*(5), 522-528.
[17]. Schwartz, S., Kent, W. J., Smit, A., Zhang, Z., Baertsch, R., Hardison, R. C., & Miller, W. (2003). Human–mouse alignments with BLASTZ. *Genome research*, *13*(1), 103-107.
[18]. Song, Z., Xu, Y., Bao, L., Zhang, L., Yu, P., Qu, Y., & Qin, C. (2019). From SARS to MERS, thrusting coronaviruses into the spotlight. *Viruses*, *11*(1), 59. Doi: https://doi.org/10.3390/v11010059.
[19]. Cui, J., Li, F., & Shi, Z. L. (2019). Origin and evolution of pathogenic coronaviruses. *Nature reviews Microbiology*, *17*(3), 181-192.
[20]. Guo, Y. R., Cao, Q. D., Hong, Z. S., Tan, Y. Y., Chen, S. D., Jin, H. J., ... & Yan, Y. (2020). The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak–an update on the status. *Military Medical Research*, *7*(1), 1-10.
[21]. Paraskevis, D., Kostaki, E. G., Magiorkinis, G., Panayiotakopoulos, G., Sourvinos, G., & Tsiodras, S. (2020). Full-genome evolutionary analysis of the novel corona virus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event. *Infection, Genetics and Evolution*, *79*, 104212.
[22]. https://www.scienceboard.net/index.aspx?sec=sup&sub=cell&pag=dis&ItemID=465
[23]. Zhang, T., Wu, Q., & Zhang, Z. (2019). Probable Pangolin Origin of 2019-nCoV Associated with Outbreak of COVID-19. *CURRENT-BIOLOGY-D-20-00299*
[24]. Poon, L. L. M., Chu, D. K. W., Chan, K. H., Wong, O. K., Ellis, T. M., Leung, Y. H. C., ... & Guan, Y. (2005). Identification of a novel coronavirus in bats. *Journal of virology*, *79*(4), 2001-2009.
[25]. Shen, Z., Xiao, Y., Kang, L., Ma, W., Shi, L., Zhang, L., Zhou, Z., Yang, J., Zhong, J., Yang, D., Guo, L., Zhang, G., Li, H., Xu, Y., Chen, M., Gao, Z., Wang, J., Ren, L., & Li, M. (2020). Genomic diversity of SARS-CoV-2 in Coronavirus Disease 2019 patients. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*, ciaa203. Advance online publication. https://doi.org/10.1093/cid/ciaa203
[26]. Malik, Y. S., Sircar, S., Bhat, S., Sharun, K., Dhama, K., Dadar, M., & Chaicumpa, W. (2020). Emerging novel coronavirus (2019-nCoV)—current scenario, evolutionary perspective based on genome analysis and recent developments. *Veterinary quarterly*, *40*(1), 68-76.
[27]. Forster, P., Forster, L., Renfrew, C., & Forster, M. (2020). Phylogenetic network analysis of SARS-CoV-2 genomes. *Proceedings of the National Academy of Sciences*, *117*(17), 9241-9243
[28]. Lin, Q., Zhao, S., Gao, D., Lou, Y., Yang, S., Musa, S. S., & He, D. (2020). A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action. *International journal of infectious diseases*, *93*, 211-216.
[29]. Phelan, A. L., Katz, R., & Gostin, L. O. (2020). The novel coronavirus originating in Wuhan, China: challenges for global health governance. *Jama*, *323*(8), 709-710.
[30]. Kruse, R. L. (2020). Therapeutic strategies in an outbreak scenario to treat the novel coronavirus originating in Wuhan, China. *F1000Research*, *9*, 72. https://doi.org/10.12688/f1000research.22211.2
[31]. Thompson R. N. (2020). Novel Coronavirus Outbreak in Wuhan, China, 2020: Intense Surveillance Is Vital for Preventing Sustained Transmission in New Locations. *Journal of clinical medicine*, *9*(2), 498. https://doi.org/10.3390/jcm9020498
[32]. Cheng, Z. J., & Shan, J. (2020). 2019 Novel coronavirus: where we are and what we know. *Infection*, *48*(2), 155–163. https://doi.org/10.1007/s15010-020-01401-y.
[33]. Xu, Y. (2020). Unveiling the Origin and Transmission of 2019-nCoV. *Trends in microbiology*, *28*(4), 239–240. https://doi.org/10.1016/j.tim.2020.02.001.

How to cite this article: Satpathy, R. (2020). *In Silico* based Whole Genome Phylogenetic Analysis of Novel Coronavirus (SARS-CoV-2). *International Journal on Emerging Technologies*, *11*(3): 1157–1163.