Bima Journal of Science and Technology, Vol. 8(4B) Jan, 2025 ISSN: 2536-6041

DOI: 10.56892/bima.v8i4B.1171

An Alignment-Free Method for Analyzing COVID-19 Sequence Similarity and Evolutionary Divergence Across Global Epicenters

[1]Adamu Yusha'u*, [2]Ahmed Aliyu, [3]Usman Waziri

[1,2.3]Department of Mathematical Science, Faculty of science, Bauchi State University, Gadau, Nigeria.

Corresponding Author: yadamu@basug.edu.ng

ABSTRACT

In this paper, we developed an alignment-free method for examining the similarity and evolutionary divergence of COVID-19 virus sequences collected from five different countries that are epicenters of the disease, facilitating faster sequence analysis. The concept of similarity presented here is essentially a generalization of the notion of equivalence. Our method employed MATLAB programming to calculate the frequency of occurrence of all nucleotide bases from each DNA sequence, resulting in a dissimilarity matrix. In this matrix, smaller values indicate greater similarity among the COVID-19 virus sequences. The dissimilarities were then converted into a similarity matrix, and using the mathematical concept of fuzzy transitive relations, we ultimately derived a phylogenetic tree for the COVID-19 viruses. this help to trace back a common ancestor or identify evolutionary bottlenecks and divergence events and could help also to understand how different strains have evolved and spread, which has implications for tracking disease outbreaks and developing vaccines or treatments which is the ultimate aim of the research work.

Keywords: Alignment-free method, COVID-19 virus sequences, Dissimilarity matrix, Similarity matrix, Epicenters

INTRODUCTION

Biological systems are incredibly complex due to their durability, stability, flexibility, and adaptability. These systems generate vast amounts of biological data through unknown processes. Interpreting such data involves uncovering interconnected linkages hidden within it. However, due to these challenges, the accuracy of predictions or the insights gained from databases is often limited. This highlights a significant need for further investigation, necessitating the development of new conceptual models and analytical methods. ("36th European Congress of Pathology – Abstracts," 2024) Numerous researchers have attempted to offer scientific solutions for the control of infectious illnesses on a worldwide scale, but we continue to struggle with these diseases in our local communities, which result in a high proportion of infection-related morbidity and mortality in society. China, South Korea, Iran, Italy, France, Germany, Spain, the United States, and Switzerland were listed as the initial COVID-19 epicenters (Bloom & Cadarette, 2019). The dynamics of SARS-CoV-2 dissemination altered to reflect spatial proximities as the epicenter moved from China to other nations, such as Italy (Adegboye et al., 2021).

As of March 2, 2020, an early analysis estimated that the true extent of the outbreak in Wuhan, the epicenter in China, was approximately 56,945 cases, which closely matched the number publicly reported at the time. (Lin et al., 2020). This highlights the challenges of accurately estimating infection levels due to testing capacity limitations,

DOI: 10.56892/bima.v8i4B.1171

especially in the early stages. By June 30, 2020, there were approximately 10 million COVID-19 infections worldwide, with at least one case reported in 215 countries and territories.(Chae et al., 2020). This suggests that, although the epicenters shifted, sporadic outbreaks continued to occur in many countries worldwide. There's no doubt that COVID-19 is a viral disease caused by the SARS-CoV-2 virus that was first identified in December 2019 in Wuhan, China. Since then, it has spread rapidly across the world, leading to a global pandemic cases were being reported in other countries, including South Korea, Japan, and Thailand. The World Health Organization (WHO) declared a Public Health Emergency of International Concern on January 30, 2020(Cucinotta & Vanelli, 2020). The virus continued to spread rapidly, and by March 11, 2020, the WHO declared COVID-19 a global pandemic. Governments around the world began implementing measures such as travel restrictions, social distancing, and lockdowns in an attempt to slow the spread of the virus.(Krishnaratne et al., 2022)

As the pandemic progressed, scientists and healthcare workers worked tirelessly to develop effective treatments and vaccines for COVID-19. The first vaccines were authorized for emergency use in December 2020, and by early 2021, vaccination programs were underway in many countries(Thanh Le et al., 2020).

However, the pandemic continued to have a significant impact on the world. As of April 2023, over 460 million confirmed cases and 6.3 million deaths have been reported worldwide, according to data from the WHO. The ongoing emergence of new variants of the virus, such as the Omicron variant, continues to pose challenges in controlling the pandemic(Chakraborty & Maity, 2020) The phylogenetic tree was constructed using the UPGMA method, a numerical coding technique for DNA sequences, employing dynamic programming. However, these methods require multiple alignments of the sequences and assume a specific evolutionary model(Drummond & Rodrigo, 2000; Jakó et al., 2009),(Yoshida et al., 2022),(Raman Kumar & Vaegae, 2020),(Rust, 2019). In addition to the computational complexity and inherent uncertainty associated with alignment cost criteria, these approaches often fail to provide a comprehensive understanding of phylogeny. Numerous alignment-free methods for phylogenetic analysis have been reported in the literature(Saw et al., 2019),(Chan et al., 2014),(Adamu & mathur, 2023),(Huang et al., 2011),(Bonham-Carter et al., 2014),(Ren et al., 2018) while recent studies propose a novel way to represent genetic sequences as fuzzy sets in the I12 space. This expands upon the approach initially introduced by Torres & Nieto and Sadegh-Zadeh, offering a straightforward method to analyze the similarity of biological sequences. Therefore, we intend to conduct this research in order to come up with a suitable method to help us understand the evolutionary divergence of the covid-19 pandemic taken into account the sequences taken from country that are epicenter of the disease in relation to a sequence taken from Nigeria in order to understand the trend of the disease and where those it evolves after construction the Phylogenetic tree.

MATERIALS AND METHODS

The methodology of the proposed model is presented in the following listed stages

Step I: twenty-three arbitrary sequences of covid-19 were considered from four different countries that are epicenter of covid-19 virus namely China, Brazil, India, United State America and Nigeria. Therefore, complete genome of the aforementioned sequence was considered in this study and tabulated in 'Frequency table' below

DOI: 10.56892/bima.v8i4B.1171

Step 2: From the generated DNA sequences, a unique feature known as frequency of nucleotide base were identified and computed.

Step 3: Four different types of frequencies termed as frequency of Adenine, Cytosine, Guanine, and Thymine designated as "A", C, G, T respectively were identified and used in forming 'Frequency table'.

Step 4: Formation of frequency table can be access in the result and discussion section of this work.

Step 5: From 'Frequency table' in **step 3** above we computed the Euclidean distances among twenty-three (23) Covid-19 virus sequences.

Step 6: Composition table was formed using fuzzy compatibility relationship *R* which is define on *X* in terms of the correct distance function of the Minkowski class using the relation

$R(x_j, x_k) = 1 - \partial(\sum_{i=1}^{p} |x_{ji} - x_{ki}|^q)^{\frac{1}{q}}$ where $R(x_j, x_k) \in X$, $q \in \bar{R}$ and ∂ is a constant that constrains $R(x_j, x_k) \in [0, 1]$, clearly, ∂ is the reverse value of the largest distance in *X*. In general, R is a fuzzy compatibility relation, but not necessarily a fuzzy equivalence relation(Adamu & mathur, 2023). Hence, there is need to determine the transitive closure of R in the next step.

Step 7: Fuzzy Transitivity closure testing on $R(x_j, x_k)$ is done on the finite set X to ascertain equivalence relation.

Step 8: Phylogenetic tree was formed based on clustering method obtained from fuzzy equivalence relation.

Constructing Phylogenetic the Tree

The fuzzy compatibility relationship R obtained can be defined on X in terms of the correct distance function of the Minkowski class by the formula [16].

$$\Re(\overline{A},\overline{B}) = 1 - \partial(d_w(\overline{A},\overline{B})).....(1)$$

For all pairs, $\Re(\overline{A},\overline{B}) \in X$ where $w \in \Re$, ∂ is a constant that ensures $\Re(\overline{A},\overline{B}) \in [0, 1]$, clearly, ∂ is the reciprocal of the largest distance in X . In general, $\Re$ defined by equation (1) is a fuzzy compatibility relation, but not necessarily a fuzzy equivalence relation. Hence, there is need to determine the transitive closure of $\Re$.

Given a relation $\Re(X, X)$, its transitive closure $\overline{\Re}(X, X)$ can be determined by simple algorithm that consists of the following three steps:

$\Re' = \Re \bigcup (\Re o \Re)$

If $\Re' \neq \Re$, *make* $\Re = \Re'$ *and go to step one*

stop $\Re' = \overline{\Re}$

This algorithm applies to both smooth and fuzzy relationships, the form of composition and set union in step I must be however compatible with the definition of transitivity employed. When max-min composition and the max operator for set union are used, we called $\bar{R}$ the transitive max-min closure (Lee, 2001a).

DOI: 10.56892/bima.v8i4B.1171

The max–min product of the relation matrices for M and N. A similarity relation P in X is a fuzzy relation in X which is reflexive, symmetric, and transitive (Lee, 2001b). We can see that the concept of a similarity relation is essentially a generalization of the concept of an equivalence relation.

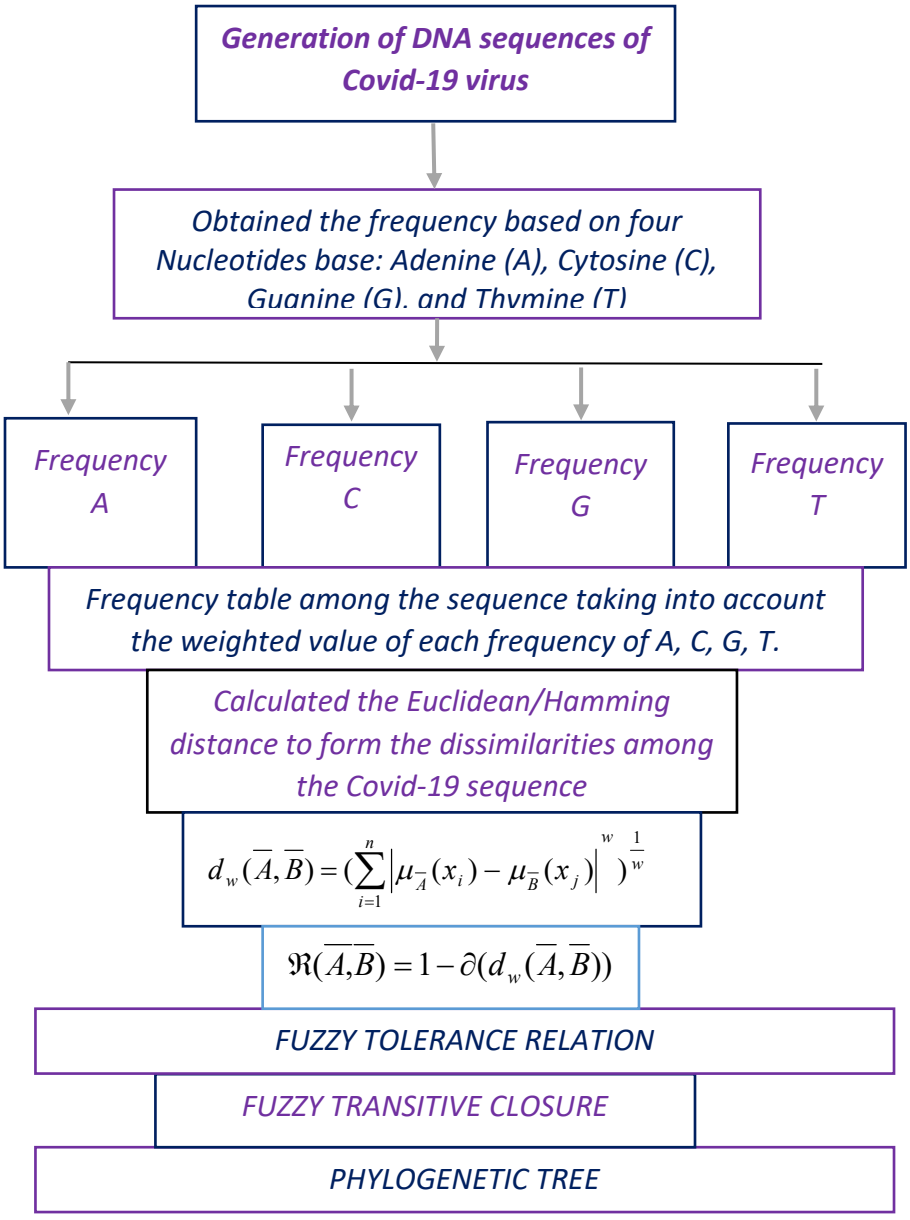

Figure 1: Flowchart of the proposed method.

More specifically, if we define $\overline{\Re} = (\bigcup_{i=1}^{n} \Re^i)$ transitive closure of any fuzzy relation $\overline{\Re}$, then $\overline{\Re}$ is transitive. If $\Re$ is a fuzzy relation characterized by a relation matrix of order n. Now, we can construct a transitive relation from any relation characterized by a relation matrix of order n. In fact, for any such fuzzy relation $\Re$ we compute $\Re^k$ successively.

DOI: 10.56892/bima.v8i4B.1171

according to our procedure above we achieve $\overline{\Re} = \Re^k$. By applying above procedure, analyze the data for $w = 1,2$ In equation (1), firstly, for $w = 2$, which is corresponds to the Euclidean distance [18] represented by:

$$D_{i,j} = \sqrt{\left(Na_i - Na_j\right)^2 + \left(Nc_i - Nc_j\right)^2 + \left(Ng_i - Ng_j\right)^2 + \left(Nt_i - Nt_j\right)^2}$$

In every fuzzy relation R can be uniquely represented in terms of it's $\alpha - cuts$ by the formula $\Re = \bigcup \alpha^{\alpha}.\Re_{\alpha} \in [1,0]$ for any value $\alpha \in [1,0]$, create a crisp equivalence relation that represents the presence the similarity between the elements to the degree α. Each of these equivalence relation forms a partition of X. Let $\pi({}^{\alpha}.\Re)$ denote the partition corresponding to the equivalence relation $({}^{\alpha}.\Re)$ we may say two elements x and y belongs to the same block of this partition if and only if $\pi({}^{\alpha}.\Re)$

R is associated with a sequence of twenty-three (23) nested partition $\pi({}^{\alpha}R)$ for $\alpha \in \wedge_R$ and $\alpha > 0$

The hierarchical clustering method has been used on Fuzzy equivalence relation to obtain the desired phylogenetic tree, we determine a fuzzy comparability relation in term of Euclidean distance and also hamming distance then we apply the algorithms for obtaining nested sequence for $\alpha - cut$ (Bei et al., 2021)

Data Materials

Materials

Firstly, we consider twenty sequences of covid-19 from five different countries that are epicenter of covid-19 namely Brazil, China, Italy, India and united states of America. Due to the fact that covid-19 was originate their which later was spread to the other part of the globe.

Bima Journal of Science and Technology, Vol. 8(4B) Jan, 2025 ISSN: 2536-6041

DOI: 10.56892/bima.v8i4B.1171

S/NO	SYMBOL	ACCESSION NUMBER	ORIGIN	LENTH	FREQUENCY OF 'A'	FREQUENCY OF 'C'	FREQUENCY OF 'G'	FREQUENCY OF 'T'
1	A	NM_001405923	BRAZIL	4385	1429	775	795	1386
2	B	NM_000746	BRAZIL	6149	1329	725	736	1327
3	C	NM_001276379	BRAZIL	4117	1552	1349	1330	1918
4	D	NM_001385055	BRAZIL	2365	659	528	564	614
5	E	NM_001289128	BRAZIL	3943	1048	888	993	1014
6	F	OM065388	CHINA	29903	8954	5492	5863	9594
7	G	OM065387	CHINA	29903	8954	5490	5862	9597
8	H	OM065386	CHINA	29903	8954	5491	5862	9596
9	I	OM065385	CHINA	29903	8954	5492	5863	9594
10	J	OM065376	CHINA	29903	8954	5492	5862	9595
11	K	OP787253	INDIA	29836	8896	5460	5849	9596
12	L	OP781960	INDIA	29726	8881	5445	5835	9562
13	M	OP787551	INDIA	29836	8893	5453	5845	9580
14	N	OP787561	INDIA	29838	8892	5453	5845	9581
15	O	MT358637	INDIA	29903	8952	5488	5865	9598
16	P	OT532241	USA	29254	8756	5350	5739	9409
17	Q	MT020881	USA	29882	8933	5491	5863	9595
18	R	OT947617	USA	9231	2139	2536	2380	2175
19	S	OT947619	USA	7467	1731	2016	1904	1816
20	T	OO781920	USA	13932	6300	5745	5549	5322
21	V	ON564650.1	NIGERIA	29801	10187	6753	7427	5434
22	W	ON564675.1	NIGERIA	29789	6919	6695	7330	8845
23	X	NM_006744.4	NIGERIA	1070	232	274	286	278

Figure 2: Frequency Table

Result and Discussion

The proposed method has been illustrated and validated by constructing phylogenetic tree for twenty-three (23) sequence in frequency table taken into consideration. All the sequences are length between 1-30,000 length presented in **Figure 2** that gives the frequency of each sequence of covid-19 virus based on it four nucleotide bases. i.e. frequency of adenine**(A)**, cytosine**(C)**, guanine**(G)** and thymine**(T)**

For all the twenty- three sequence that been considered.

DOI: 10.56892/bima.v8i4B.1171

Algorithm I: generating frequency of nucleotides base of each sequence of covid-19 under consideration

1. Initialize matrix N of size 20x4 with zeros
2. For from 1 to 20 do:
3. Create file path 'fil'
4. Set delimiter in to ''
5. Read content of file 'fil' into variable 'mydata'
6. Set N [i, 1] to count occurrences of 'a' in 'mydata'
7. Set N [i, 2] to count occurrences of 'c' in 'mydata'
8. Set N [i, 3] to count occurrences of 'g' in 'mydata'
9. Set N [i, 4] to count occurrences of 't' in 'mydata
10. End For

S/NO	SYMBOL	SYMBOL 2	ACCESSION NUMBER	ORIGIN	LENTH	FREQUENCY OF 'A'	FREQUENCY OF 'C'	FREQUENCY OF 'G'	FREQUENCY OF 'T'
1	N[1,1]	A	NM_001405923	BRAZIL	4385	1429	775	795	1386
2	N[1,2]	B	NM_000746	BRAZIL	6149	1329	725	736	1327
3	N[1,3]	C	NM_001276379	BRAZIL	4117	1552	1349	1330	1918
4	N[1,4]	D	NM_001385055	BRAZIL	2365	659	528	564	614
5	N[1,5]	E	NM_001289128	BRAZIL	3943	1048	888	993	1014
6	N[1,6]	F	OM065388	CHINA	29903	8954	5492	5863	9594
7	N[1,7]	G	OM065387	CHINA	29903	8954	5490	5862	9597
8	N[1,8]	H	OM065386	CHINA	29903	8954	5491	5862	9596
9	N[1,9]	I	OM065385	CHINA	29903	8954	5492	5863	9594
10	N[1,10]	J	OM065376	CHINA	29903	8954	5492	5862	9595
11	N[1,11]	K	OP787253	INDIA	29836	8896	5460	5849	9596
12	N[1,12]	L	OP781960	INDIA	29726	8881	5445	5835	9562
13	N[1,13]	M	OP787551	INDIA	29836	8893	5453	5845	9580
14	N[1,14]	N	OP787561	INDIA	29838	8892	5453	5845	9581
15	N[1,15]	O	MT358637	INDIA	29903	8952	5488	5865	9598
16	N[1,16]	P	OT532241	USA	29254	8756	5350	5739	9409
17	N[1,17]	Q	MT020881	USA	29882	8933	5491	5863	9595
18	N[1,18]	R	OT947617	USA	9231	2139	2536	2380	2175
19	N[1,19]	S	OT947619	USA	7467	1731	2016	1904	1816
20	N[1,20]	T	OO781920	USA	13932	6300	5745	5549	5322
21	N[1,21]	V	ON564650.1	NIGERIA	29801	10187	6753	7427	5434
22	N[1,22]	W	ON564675.1	NIGERIA	29789	6919	6694	7330	8845
23	N[1,23]	X	NM_006744.4	NIGERIA	1070	232	274	286	278

Figure 3: frequencies of nucleotide base of twenty-three of covid-19 sequences.

Algorithm II: General minkwoski formula (considered special case of Euclidean distance)

1. Initialize matrix D of size 20x20 with zeros
2. For i from 1 to 20 do:
3. For j from 1 to 20 do:
4. Set D [i, j] to sqrt ((N[i, 1] – N[j, 1])^2 + (N[i, 2] – N[j, 2])^2 + (N[i, 3] – N[j, 3])^2 + (N[i, 4] – N[j, 4])^2)
5. end for
6. end for

DOI: 10.56892/bima.v8i4B.1171

R(i,j)	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	V	W	X
A	0	0.014	0.0956	0.1142	0.0579	1.3112	1.3113	1.3113	1.3112	1.3113	1.3063	1.3023	1.3048	1.3048	1.3113	1.2784	1.31	0.2596	0.1745	0.9302	1.6984	1.6665	0.2009
B	0.014	0	0.1068	0.1013	0.0519	1.3247	1.3248	1.3248	1.3247	1.3248	1.3198	1.3158	1.3183	1.3183	1.3248	1.2919	1.3236	0.2713	0.1852	0.9436	1.7119	1.68	0.1874
C	0.0956	0.1068	0	0.1939	0.1182	1.2305	1.2306	1.2306	1.2305	1.2306	1.2256	1.2215	1.224	1.224	1.2306	1.1977	1.2293	0.1709	0.0904	0.8441	1.6177	1.5855	0.2822
D	0.1142	0.1013	0.1939	0	0.0791	1.4219	1.4219	1.4219	1.4219	1.4219	1.417	1.4129	1.4154	1.4154	1.4219	1.389	1.4207	0.3458	0.257	1.0298	1.8091	1.777	0.0893
E	0.0579	0.0519	0.1182	0.0791	0	1.3455	1.3456	1.3455	1.3455	1.3455	1.3406	1.3365	1.339	1.339	1.3456	1.3127	1.3443	0.2679	0.1792	0.9513	1.7326	1.7005	0.1677
F	1.3112	1.3247	1.2305	1.4219	1.3455	0	0.0004	0.0002	0	0.0001	0.0068	0.0097	0.0076	0.0077	0.0006	0.033	0.0021	1.1061	1.185	0.5045	0.3874	0.3555	1.5107
G	1.3113	1.3248	1.2306	1.4219	1.3456	0.0004	0	0.0001	0.0004	0.0003	0.0067	0.0096	0.0075	0.0076	0.0004	0.0331	0.0021	1.1063	1.1851	0.5048	0.3873	0.3554	1.5108
H	1.3113	1.3248	1.2306	1.4219	1.3455	0.0002	0.0001	0	0.0002	0.0001	0.0067	0.0097	0.0076	0.0076	0.0005	0.033	0.0021	1.1062	1.1851	0.5047	0.3874	0.3555	1.5108
I	1.3112	1.3247	1.2305	1.4219	1.3455	0	0.0004	0.0002	0	0.0001	0.0068	0.0097	0.0076	0.0077	0.0006	0.033	0.0021	1.1061	1.185	0.5045	0.3874	0.3555	1.5107
J	1.3113	1.3248	1.2306	1.4219	1.3455	0.0001	0.0003	0.0001	0.0001	0	0.0068	0.0097	0.0076	0.0076	0.0006	0.033	0.0021	1.1062	1.185	0.5046	0.3874	0.3555	1.5107
K	1.3063	1.3198	1.2256	1.417	1.3406	0.0068	0.0067	0.0067	0.0068	0.0068	0	0.0042	0.0018	0.0017	0.0065	0.0281	0.005	1.1014	1.1802	0.5018	0.3923	0.3603	1.5058
L	1.3023	1.3158	1.2215	1.4129	1.3365	0.0097	0.0096	0.0097	0.0097	0.0097	0.0042	0	0.0025	0.0025	0.0095	0.0239	0.0082	1.0974	1.1762	0.4981	0.3963	0.3644	1.5017
M	1.3048	1.3183	1.224	1.4154	1.339	0.0076	0.0075	0.0076	0.0076	0.0076	0.0018	0.0025	0	0.0001	0.0074	0.0264	0.006	1.0998	1.1786	0.5003	0.3938	0.3619	1.5042
N	1.3048	1.3183	1.224	1.4154	1.339	0.0077	0.0076	0.0076	0.0077	0.0076	0.0017	0.0025	0.0001	0	0.0074	0.0264	0.006	1.0998	1.1786	0.5003	0.3938	0.3619	1.5042
O	1.3113	1.3248	1.2306	1.4219	1.3456	0.0006	0.0004	0.0005	0.0006	0.0006	0.0065	0.0095	0.0074	0.0074	0	0.033	0.002	1.1062	1.1851	0.5048	0.3873	0.3554	1.5108
P	1.2784	1.2919	1.1977	1.389	1.3127	0.033	0.0331	0.033	0.033	0.033	0.0281	0.0239	0.0264	0.0264	0.033	0	0.0318	1.0739	1.1525	0.4788	0.4202	0.3883	1.4779
Q	1.31	1.3236	1.2293	1.4207	1.3443	0.0021	0.0021	0.0021	0.0021	0.0021	0.005	0.0082	0.006	0.006	0.002	0.0318	0	1.1049	1.1838	0.5035	0.3886	0.3567	1.5095
R	0.2596	0.2713	0.1709	0.3458	0.2679	1.1061	1.1063	1.1062	1.1061	1.1062	1.1014	1.0974	1.0998	1.0998	1.1062	1.0739	1.1049	0	0.089	0.6896	1.4915	1.4591	0.4321
S	0.1745	0.1852	0.0904	0.257	0.1792	1.185	1.1851	1.1851	1.185	1.185	1.1802	1.1762	1.1786	1.1786	1.1851	1.1525	1.1838	0.089	0	0.7769	1.5713	1.539	0.3435
T	0.9302	0.9436	0.8441	1.0298	0.9513	0.5045	0.5048	0.5047	0.5045	0.5046	0.5018	0.4981	0.5003	0.5003	0.5048	0.4788	0.5035	0.6896	0.7769	0	0.8577	0.8259	1.1179
V	1.6984	1.7119	1.6177	1.8091	1.7326	0.3874	0.3873	0.3874	0.3874	0.3874	0.3923	0.3963	0.3938	0.3938	0.3873	0.4202	0.3886	1.4915	1.5713	0.8577	0	0.033	1.8979
W	1.6665	1.68	1.5855	1.777	1.7005	0.3555	0.3554	0.3555	0.3555	0.3555	0.3603	0.3644	0.3619	0.3619	0.3554	0.3883	0.3567	1.4591	1.539	0.8259	0.033	0	1.8659
X	0.2009	0.1874	0.2822	0.0893	0.1677	1.5107	1.5108	1.5108	1.5107	1.5107	1.5058	1.5017	1.5042	1.5042	1.5108	1.4779	1.5095	0.4321	0.3435	1.1179	1.8979	1.8659	0

Figure 4: Dissimilarities obtained by using Euclidean distance among covid-19 virus sequences

Algorithm III: For calculating fuzzy compatibility relation among sequences

1. Display matrix D
2. Set rho to 1 / (maximum value in matrix D)
3. Set R to (1 – rho * D)
4. Display "Table R"
5. Display matrix R
6. Initialize matrix composition of size 20x20 with zeros
7. For i from 1 to 20 do:
8. For j from 1 to 20 do:
9. Set maxi to min (R [i, 1], R [1, j]
10. For k from 2 to 20 do:
11. Set tempmin to min (R[i, k], R[k, j])
12. If tempmin > maxi, then:
13. Set maxi to tempmin
14. End If
15. End For

Now, our aims now is to get Fuzzy transitive relation, which can be obtained by performing max-min composition on R i.e. $R^1 = R0R$ using this inequality "$2^k \geq n - 1$"where k is a positive integer value indicate the number of times to performed composition as stated in the Step-1 of the algorithms to get transitive closure, where n is the number of columns. In our study, we have n=23, therefore at most three compositions can be performed to reach to Algorithms III in the stated algorithms of transitive closure.

Now, first composition will be performed for k=1 and the resultant relationship will be obtained as $R^1 = R0R$.

DOI: 10.56892/bima.v8i4B.1171

R(i,j)	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	V	W	X
A	1	0.9926	0.9496	0.9398	0.9695	0.3091	0.3091	0.3091	0.3091	0.3091	0.3117	0.3138	0.3125	0.3125	0.3091	0.3264	0.3097	0.8632	0.908	0.5099	0.1051	0.1219	0.8942
B	0.9926	1	0.9437	0.9466	0.9726	0.302	0.302	0.302	0.302	0.302	0.3046	0.3067	0.3054	0.3054	0.302	0.3193	0.3026	0.8571	0.9024	0.5028	0.098	0.1148	0.9012
C	0.9496	0.9437	1	0.8978	0.9377	0.3516	0.3516	0.3516	0.3516	0.3516	0.3542	0.3564	0.3551	0.3551	0.3516	0.3689	0.3523	0.9099	0.9524	0.5552	0.1477	0.1646	0.8513
D	0.9398	0.9466	0.8978	1	0.9583	0.2508	0.2508	0.2508	0.2508	0.2508	0.2534	0.2555	0.2542	0.2542	0.2508	0.2681	0.2515	0.8178	0.8646	0.4574	0.0468	0.0637	0.953
E	0.9695	0.9726	0.9377	0.9583	1	0.2911	0.291	0.291	0.2911	0.2911	0.2936	0.2958	0.2945	0.2945	0.291	0.3083	0.2917	0.8588	0.9056	0.4988	0.0871	0.104	0.9116
F	0.3091	0.302	0.3516	0.2508	0.2911	1	0.9998	0.9999	1	0.9999	0.9964	0.9949	0.996	0.996	0.9997	0.9826	0.9989	0.4172	0.3756	0.7342	0.7959	0.8127	0.204
G	0.3091	0.302	0.3516	0.2508	0.291	0.9998	1	0.9999	0.9998	0.9999	0.9965	0.9949	0.996	0.996	0.9998	0.9826	0.9989	0.4171	0.3756	0.734	0.7959	0.8127	0.204
H	0.3091	0.302	0.3516	0.2508	0.291	0.9999	0.9999	1	0.9999	0.9999	0.9965	0.9949	0.996	0.996	0.9997	0.9826	0.9989	0.4171	0.3756	0.7341	0.7959	0.8127	0.204
I	0.3091	0.302	0.3516	0.2508	0.2911	1	0.9998	0.9999	1	0.9999	0.9964	0.9949	0.996	0.996	0.9997	0.9826	0.9989	0.4172	0.3756	0.7342	0.7959	0.8127	0.204
J	0.3091	0.302	0.3516	0.2508	0.2911	0.9999	0.9999	0.9999	0.9999	1	0.9964	0.9949	0.996	0.996	0.9997	0.9826	0.9989	0.4172	0.3756	0.7341	0.7959	0.8127	0.204
K	0.3117	0.3046	0.3542	0.2534	0.2936	0.9964	0.9965	0.9965	0.9964	0.9964	1	0.9978	0.999	0.9991	0.9966	0.9852	0.9974	0.4197	0.3782	0.7356	0.7933	0.8101	0.2066
L	0.3138	0.3067	0.3564	0.2555	0.2958	0.9949	0.9949	0.9949	0.9949	0.9949	0.9978	1	0.9987	0.9987	0.995	0.9874	0.9957	0.4218	0.3803	0.7376	0.7912	0.808	0.2087
M	0.3125	0.3054	0.3551	0.2542	0.2945	0.996	0.996	0.996	0.996	0.996	0.999	0.9987	1	0.9999	0.9961	0.9861	0.9968	0.4205	0.379	0.7364	0.7925	0.8093	0.2074
N	0.3125	0.3054	0.3551	0.2542	0.2945	0.996	0.996	0.996	0.996	0.996	0.9991	0.9987	0.9999	1	0.9961	0.9861	0.9968	0.4205	0.379	0.7364	0.7925	0.8093	0.2074
O	0.3091	0.302	0.3516	0.2508	0.291	0.9997	0.9998	0.9997	0.9997	0.9997	0.9966	0.995	0.9961	0.9961	1	0.9826	0.999	0.4171	0.3756	0.734	0.7959	0.8127	0.204
P	0.3264	0.3193	0.3689	0.2681	0.3083	0.9826	0.9826	0.9826	0.9826	0.9826	0.9852	0.9874	0.9861	0.9861	0.9826	1	0.9832	0.4342	0.3927	0.7477	0.7786	0.7954	0.2213
Q	0.3097	0.3026	0.3523	0.2515	0.2917	0.9989	0.9989	0.9989	0.9989	0.9989	0.9974	0.9957	0.9968	0.9968	0.999	0.9832	1	0.4178	0.3763	0.7347	0.7952	0.8121	0.2046
R	0.8632	0.8571	0.9099	0.8178	0.8588	0.4172	0.4171	0.4171	0.4172	0.4172	0.4197	0.4218	0.4205	0.4205	0.4171	0.4342	0.4178	1	0.9531	0.6366	0.2141	0.2312	0.7723
S	0.908	0.9024	0.9524	0.8646	0.9056	0.3756	0.3756	0.3756	0.3756	0.3756	0.3782	0.3803	0.379	0.379	0.3756	0.3927	0.3763	0.9531	1	0.5906	0.1721	0.1891	0.819
T	0.5099	0.5028	0.5552	0.4574	0.4988	0.7342	0.734	0.7341	0.7342	0.7341	0.7356	0.7376	0.7364	0.7364	0.734	0.7477	0.7347	0.6366	0.5906	1	0.5481	0.5648	0.411
V	0.1051	0.098	0.1477	0.0468	0.0871	0.7959	0.7959	0.7959	0.7959	0.7959	0.7933	0.7912	0.7925	0.7925	0.7959	0.7786	0.7952	0.2141	0.1721	0.5481	1	0.9826	0
W	0.1219	0.1148	0.1646	0.0637	0.104	0.8127	0.8127	0.8127	0.8127	0.8127	0.8101	0.808	0.8093	0.8093	0.8127	0.7954	0.8121	0.2312	0.1891	0.5648	0.9826	1	0.0169
X	0.8942	0.9012	0.8513	0.953	0.9116	0.204	0.204	0.204	0.204	0.204	0.2066	0.2087	0.2074	0.2074	0.204	0.2213	0.2046	0.7723	0.819	0.411	0	0.0169	1

Figure 5: Fuzzy compatibility relationships among Covid-19 virus sequences.

Algorithm IV: Calculation of fuzzy transitive relationship by first composition for K=1

1. Initialize matrix composition of size 23x23 with zeros
2. For i from 1 to 20 do:
3. For j from 1 to 20 do:
4. Set maxi to min (R [i, 1], R [1, j]
5. For k from 2 to 20 do:
6. Set tempmin to min (R [i, k], R [k, j])
7. If tempmin > maxi, then:
8. Set maxi to tempmin
9. End If
10. End For
11. Set composition [i, j] to maxi
12. End For
13. End For
14. Display “Composition 1”
15. Display matrix composition
16. For counter from 2 to 19 do:
17. Set R to composition
18. Initialize matrix composition of size 20x20 with zeros
19. For i from 1 to 20 do:
20. For j from 1 to 20 do:
21. Set maxi to min (R [i, 1], R [1, j])
22. For k from 2 to 20 do:
23. Set tempmin to min (R [i, k], R [k, j])
24. If tempmin > maxi, then:
25. Set maxi to tempmin
26. End If
27. End For
28. Set composition [i, j] to maxi
29. End For
30. End For

DOI: 10.56892/bima.v8i4B.1171

R(i,j)	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	V	W	X
A	1	0.9926	0.9496	0.9583	0.9726	0.5099	0.5099	0.5099	0.5099	0.5099	0.5099	0.5099	0.5099	0.5099	0.5099	0.5099	0.5099	0.9099	0.9496	0.6366	0.5099	0.5099	0.9398
B	0.9926	1	0.9496	0.9583	0.9726	0.5028	0.5028	0.5028	0.5028	0.5028	0.5028	0.5028	0.5028	0.5028	0.5028	0.5028	0.5028	0.9099	0.9437	0.6366	0.5028	0.5028	0.9466
C	0.9496	0.9496	1	0.9437	0.9496	0.5552	0.5552	0.5552	0.5552	0.5552	0.5552	0.5552	0.5552	0.5552	0.5552	0.5552	0.5552	0.9524	0.9524	0.6366	0.5481	0.5552	0.9116
D	0.9583	0.9583	0.9437	1	0.9583	0.4574	0.4574	0.4574	0.4574	0.4574	0.4574	0.4574	0.4574	0.4574	0.4574	0.4574	0.4574	0.8978	0.908	0.6366	0.4574	0.4574	0.953
E	0.9726	0.9726	0.9496	0.9583	1	0.4988	0.4988	0.4988	0.4988	0.4988	0.4988	0.4988	0.4988	0.4988	0.4988	0.4988	0.4988	0.9099	0.9377	0.6366	0.4988	0.4988	0.953
F	0.5099	0.5028	0.5552	0.4574	0.4988	1	0.9999	0.9999	1	0.9999	0.9974	0.9964	0.9968	0.9968	0.9998	0.9874	0.999	0.6366	0.5906	0.7477	0.8127	0.8127	0.4172
G	0.5099	0.5028	0.5552	0.4574	0.4988	0.9999	1	0.9999	0.9999	0.9999	0.9974	0.9965	0.9968	0.9968	0.9998	0.9874	0.999	0.6366	0.5906	0.7477	0.8127	0.8127	0.4171
H	0.5099	0.5028	0.5552	0.4574	0.4988	0.9999	0.9999	1	0.9999	0.9999	0.9974	0.9965	0.9968	0.9968	0.9998	0.9874	0.999	0.6366	0.5906	0.7477	0.8127	0.8127	0.4171
I	0.5099	0.5028	0.5552	0.4574	0.4988	1	0.9999	0.9999	1	0.9999	0.9974	0.9964	0.9968	0.9968	0.9998	0.9874	0.999	0.6366	0.5906	0.7477	0.8127	0.8127	0.4172
J	0.5099	0.5028	0.5552	0.4574	0.4988	0.9999	0.9999	0.9999	0.9999	1	0.9974	0.9964	0.9968	0.9968	0.9998	0.9874	0.999	0.6366	0.5906	0.7477	0.8127	0.8127	0.4172
K	0.5099	0.5028	0.5552	0.4574	0.4988	0.9974	0.9974	0.9974	0.9974	0.9974	1	0.9987	0.9991	0.9991	0.9974	0.9874	0.9974	0.6366	0.5906	0.7477	0.8101	0.8127	0.4197
L	0.5099	0.5028	0.5552	0.4574	0.4988	0.9964	0.9965	0.9965	0.9964	0.9964	0.9987	1	0.9987	0.9987	0.9966	0.9874	0.9974	0.6366	0.5906	0.7477	0.808	0.8127	0.4218
M	0.5099	0.5028	0.5552	0.4574	0.4988	0.9968	0.9968	0.9968	0.9968	0.9968	0.9991	0.9987	1	0.9999	0.9968	0.9874	0.9974	0.6366	0.5906	0.7477	0.8093	0.8127	0.4205
N	0.5099	0.5028	0.5552	0.4574	0.4988	0.9968	0.9968	0.9968	0.9968	0.9968	0.9991	0.9987	0.9999	1	0.9968	0.9874	0.9974	0.6366	0.5906	0.7477	0.8093	0.8127	0.4205
O	0.5099	0.5028	0.5552	0.4574	0.4988	0.9998	0.9998	0.9998	0.9998	0.9998	0.9974	0.9966	0.9968	0.9968	1	0.9874	0.999	0.6366	0.5906	0.7477	0.8127	0.8127	0.4171
P	0.5099	0.5028	0.5552	0.4574	0.4988	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	1	0.9874	0.6366	0.5906	0.7477	0.7959	0.8127	0.4342
Q	0.5099	0.5028	0.5552	0.4574	0.4988	0.999	0.999	0.999	0.999	0.999	0.9974	0.9974	0.9974	0.9974	0.999	0.9874	1	0.6366	0.5906	0.7477	0.8121	0.8127	0.4178
R	0.9099	0.9099	0.9524	0.8978	0.9099	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	1	0.9531	0.6366	0.5481	0.5648	0.8632
S	0.9496	0.9437	0.9524	0.908	0.9377	0.5906	0.5906	0.5906	0.5906	0.5906	0.5906	0.5906	0.5906	0.5906	0.5906	0.5906	0.5906	0.9531	1	0.6366	0.5481	0.5648	0.9056
T	0.6366	0.6366	0.6366	0.6366	0.6366	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.6366	0.6366	1	0.7477	0.7477	0.6366
V	0.5099	0.5028	0.5481	0.4574	0.4988	0.8127	0.8127	0.8127	0.8127	0.8127	0.8101	0.808	0.8093	0.8093	0.8127	0.7959	0.8121	0.5481	0.5481	0.7477	1	0.9826	0.411
W	0.5099	0.5028	0.5552	0.4574	0.4988	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.5648	0.5648	0.7477	0.9826	1	0.411
X	0.9398	0.9466	0.9116	0.953	0.953	0.4172	0.4171	0.4171	0.4172	0.4172	0.4197	0.4218	0.4205	0.4205	0.4171	0.4342	0.4178	0.8632	0.9056	0.6366	0.411	0.411	1

Figure 6: Fuzzy transitive relationship among covid-19 virus sequences for K=1.

Algorithm V: Fuzzy transitive relationship among covid-19 virus sequences for K=2

1. repeat
2. Initialize matrix composition of size 23x23 with zeros
3. For i from 1 to 20 do:
4. For j from 1 to 20 do:
5. Set maxi to min (R [i, 1], R [1, j]
6. For k from 2 to 20 do:
7. Set tempmin to min (R [i, k], R [k, j])
8. If tempmin > maxi, then:
9. Set maxi to tempmin
10. End If
11. End For
12. Set composition [i, j] to maxi
13. End For
14. End For
15. Display “Composition 1”
16. Display matrix composition
17. For counter from 2 to 19 do:
18. Set R to composition
19. Initialize matrix composition of size 20x20 with zeros
20. For i from 1 to 20 do:
21. For j from 1 to 20 do:
22. Set maxi to min (R [i, 1], R [1, j])
23. For k from 2 to 20 do:
24. Set tempmin to min (R [i, k], R [k, j])
25. If tempmin > maxi, then:
26. Set maxi to tempmin
27. End If
28. End For
29. Set composition [i, j] to maxi
30. End For
31. End For

DOI: 10.56892/bima.v8i4B.1171

R(i,j)	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	V	W	X
A	1	0.9926	0.9496	0.9583	0.9726	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.9496	0.9496	0.6366	0.6366	0.6366	0.953
B	0.9926	1	0.9496	0.9583	0.9726	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.9496	0.9496	0.6366	0.6366	0.6366	0.953
C	0.9496	0.9496	1	0.9496	0.9496	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.9524	0.9524	0.6366	0.6366	0.6366	0.9496
D	0.9583	0.9583	0.9496	1	0.9583	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.9437	0.9496	0.6366	0.6366	0.6366	0.953
E	0.9726	0.9726	0.9496	0.9583	1	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.9496	0.9496	0.6366	0.6366	0.6366	0.953
F	0.6366	0.6366	0.6366	0.6366	0.6366	1	0.9999	0.9999	1	0.9999	0.9974	0.9974	0.9974	0.9974	0.9998	0.9874	0.999	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
G	0.6366	0.6366	0.6366	0.6366	0.6366	0.9999	1	0.9999	0.9999	0.9999	0.9974	0.9974	0.9974	0.9974	0.9998	0.9874	0.999	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
H	0.6366	0.6366	0.6366	0.6366	0.6366	0.9999	0.9999	1	0.9999	0.9999	0.9974	0.9974	0.9974	0.9974	0.9998	0.9874	0.999	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
I	0.6366	0.6366	0.6366	0.6366	0.6366	1	0.9999	0.9999	1	0.9999	0.9974	0.9974	0.9974	0.9974	0.9998	0.9874	0.999	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
J	0.6366	0.6366	0.6366	0.6366	0.6366	0.9999	0.9999	0.9999	0.9999	1	0.9974	0.9974	0.9974	0.9974	0.9998	0.9874	0.999	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
K	0.6366	0.6366	0.6366	0.6366	0.6366	0.9974	0.9974	0.9974	0.9974	0.9974	1	0.9987	0.9991	0.9991	0.9974	0.9874	0.9974	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
L	0.6366	0.6366	0.6366	0.6366	0.6366	0.9974	0.9974	0.9974	0.9974	0.9974	0.9987	1	0.9987	0.9987	0.9974	0.9874	0.9974	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
M	0.6366	0.6366	0.6366	0.6366	0.6366	0.9974	0.9974	0.9974	0.9974	0.9974	0.9991	0.9987	1	0.9999	0.9974	0.9874	0.9974	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
N	0.6366	0.6366	0.6366	0.6366	0.6366	0.9974	0.9974	0.9974	0.9974	0.9974	0.9991	0.9987	0.9999	1	0.9974	0.9874	0.9974	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
O	0.6366	0.6366	0.6366	0.6366	0.6366	0.9998	0.9998	0.9998	0.9998	0.9998	0.9974	0.9974	0.9974	0.9974	1	0.9874	0.999	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
P	0.6366	0.6366	0.6366	0.6366	0.6366	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	1	0.9874	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
Q	0.6366	0.6366	0.6366	0.6366	0.6366	0.999	0.999	0.999	0.999	0.999	0.9974	0.9974	0.9974	0.9974	0.999	0.9874	1	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
R	0.9496	0.9496	0.9524	0.9437	0.9496	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	1	0.9531	0.6366	0.6366	0.6366	0.9116
S	0.9496	0.9496	0.9524	0.9496	0.9496	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.9531	1	0.6366	0.6366	0.6366	0.9437
T	0.6366	0.6366	0.6366	0.6366	0.6366	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.6366	0.6366	1	0.7477	0.7477	0.6366
V	0.6366	0.6366	0.6366	0.6366	0.6366	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.6366	0.6366	0.7477	1	0.9826	0.6366
W	0.6366	0.6366	0.6366	0.6366	0.6366	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.6366	0.6366	0.7477	0.9826	1	0.6366
X	0.953	0.953	0.9496	0.953	0.953	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.9116	0.9437	0.6366	0.6366	0.6366	1

Figure 7: Fuzzy transitive relationship among covid-19 virus sequences for K=2.

Algorithm VI: Fuzzy transitive relationship among covid-19 virus sequences for K=3

1. Initialize matrix composition of size 20x20 with zeros
2. For i from 1 to 20 do:
3. For j from 1 to 20 do:
4. Set maxi to min (R [i, 1], R [1, j]
5. For k from 2 to 20 do:
6. Set tempmin to min (R [i, k], R [k, j])
7. If tempmin > maxi, then:
8. Set maxi to tempmin
9. End If
10. End For
11. Set composition [i, j] to maxi
12. End For
13. End For
14. Display “Composition 1”
15. Display matrix composition
16. For counter from 2 to 19 do:
17. Set R to composition
18. Initialize matrix composition of size 20x20 with zeros
19. For i from 1 to 20 do:
20. For j from 1 to 20 do:
21. Set maxi to min (R [i, 1], R [1, j])
22. For k from 2 to 20 do:
23. Set tempmin to min (R [i, k], R [k, j])
24. If tempmin > maxi, then:
25. Set maxi to tempmin
26. End If
27. End For
28. Set composition [i, j] to maxi
29. End For
30. End For
31. Display “Composition “+ counter
32. Display matrix composition
33. If composition <= R, then:
34. Break
35. End If
36. End For

For k=3, performance of max-min composition on R^1 (i.e. $R^2 = R^1 o R^1$) has been made which realized in the algorithms V for transitive closure (i.e. .$R^2 = \bar{R}$) and result obtained are shown in figure 8

DOI: 10.56892/bima.v8i4B.1171

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	V	W	X
A	1	0.9926	0.9496	0.9583	0.9726	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.9496	0.9496	0.6366	0.6366	0.6366	0.953
B	0.9926	1	0.9496	0.9583	0.9726	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.9496	0.9496	0.6366	0.6366	0.6366	0.953
C	0.9496	0.9496	1	0.9496	0.9496	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.9524	0.9524	0.6366	0.6366	0.6366	0.9496
D	0.9583	0.9583	0.9496	1	0.9583	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.9496	0.9496	0.6366	0.6366	0.6366	0.953
E	0.9726	0.9726	0.9496	0.9583	1	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.9496	0.9496	0.6366	0.6366	0.6366	0.953
F	0.6366	0.6366	0.6366	0.6366	0.6366	1	0.9999	0.9999	1	0.9999	0.9974	0.9974	0.9974	0.9974	0.9998	0.9874	0.999	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
G	0.6366	0.6366	0.6366	0.6366	0.6366	0.9999	1	0.9999	0.9999	0.9999	0.9974	0.9974	0.9974	0.9974	0.9998	0.9874	0.999	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
H	0.6366	0.6366	0.6366	0.6366	0.6366	0.9999	0.9999	1	0.9999	0.9999	0.9974	0.9974	0.9974	0.9974	0.9998	0.9874	0.999	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
I	0.6366	0.6366	0.6366	0.6366	0.6366	1	0.9999	0.9999	1	0.9999	0.9974	0.9974	0.9974	0.9974	0.9998	0.9874	0.999	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
J	0.6366	0.6366	0.6366	0.6366	0.6366	0.9999	0.9999	0.9999	0.9999	1	0.9974	0.9974	0.9974	0.9974	0.9998	0.9874	0.999	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
K	0.6366	0.6366	0.6366	0.6366	0.6366	0.9974	0.9974	0.9974	0.9974	0.9974	1	0.9987	0.9991	0.9991	0.9974	0.9874	0.9974	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
L	0.6366	0.6366	0.6366	0.6366	0.6366	0.9974	0.9974	0.9974	0.9974	0.9974	0.9987	1	0.9987	0.9987	0.9974	0.9874	0.9974	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
M	0.6366	0.6366	0.6366	0.6366	0.6366	0.9974	0.9974	0.9974	0.9974	0.9974	0.9991	0.9987	1	0.9999	0.9974	0.9874	0.9974	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
N	0.6366	0.6366	0.6366	0.6366	0.6366	0.9974	0.9974	0.9974	0.9974	0.9974	0.9991	0.9987	0.9999	1	0.9974	0.9874	0.9974	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
O	0.6366	0.6366	0.6366	0.6366	0.6366	0.9998	0.9998	0.9998	0.9998	0.9998	0.9974	0.9974	0.9974	0.9974	1	0.9874	0.999	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
P	0.6366	0.6366	0.6366	0.6366	0.6366	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	0.9874	1	0.9874	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
Q	0.6366	0.6366	0.6366	0.6366	0.6366	0.999	0.999	0.999	0.999	0.999	0.9974	0.9974	0.9974	0.9974	0.999	0.9874	1	0.6366	0.6366	0.7477	0.8127	0.8127	0.6366
R	0.9496	0.9496	0.9524	0.9496	0.9496	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	1	0.9531	0.6366	0.6366	0.6366	0.9496
S	0.9496	0.9496	0.9524	0.9496	0.9496	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.9531	1	0.6366	0.6366	0.6366	0.9496
T	0.6366	0.6366	0.6366	0.6366	0.6366	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.7477	0.6366	0.6366	1	0.7477	0.7477	0.6366
V	0.6366	0.6366	0.6366	0.6366	0.6366	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.6366	0.6366	0.7477	1	0.9826	0.6366
W	0.6366	0.6366	0.6366	0.6366	0.6366	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.8127	0.6366	0.6366	0.7477	0.9826	1	0.6366
X	0.953	0.953	0.9496	0.953	0.953	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366	0.9496	0.9496	0.6366	0.6366	0.6366	1

Figure 8: Fuzzy transitive relationship among covid-19 virus sequences for K=3.

Algorithsm VII: for constructing phylogenetic three for propose model

1. ***Input a predefined similarity matrix.***
2. *similarity_matrix = [(matrix of similarity values between samples)]*
3. ***Convert the similarity matrix to a distance matrix.***
4. *- Formula: `distance_matrix = 1 - similarity_matrix`*
5. *for each element in similarity_matrix: distance_matrix[i][j] = 1 - similarity_matrix[i][j]*
6. ***Perform hierarchical clustering using the distance matrix.***
7. *perform hierarchical clustering on distance_matrix using 'average' linkage*
8. ***Plot the resulting hierarchical clusters as a dendrogram to visualize the relationships between samples.***
9. *create dendrogram from the hierarchical clustering results*
10. *label x-axis as 'Samples'*
11. *label y-axis as 'Distance'*
12. *set plot title to 'Phylogenetic Tree'*

Bima Journal of Science and Technology, Vol. 8(4B) Jan, 2025 ISSN: 2536-6041

DOI: 10.56892/bima.v8i4B.1171

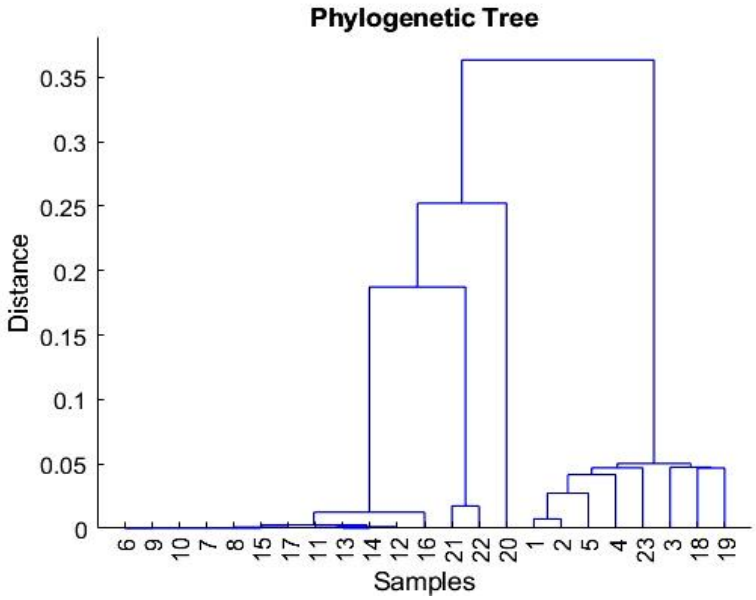

Figure 9: Phylogenetic Tree.

KEYS: 1 correspond to A while 2 correspond to B Subsequently 23 correspond to X in the phylogentic tree above

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	V	W	X

DISCUSSION ON PHYLOGENETIC TREE ON THE PROPOSE MODEL

Cluster Similarity and Evolutionary Relationships

Four first sequence from Brazil and last sequence from Nigeria are closely related, with short branch lengths. This suggests they share a significant portion of their genetic sequences or have evolved recently from a common ancestor. In evolutionary terms, these samples may represent different strains or species that diverged more recently.

F and J Sequence representation from China and Q sequence representation from USA, and O sequence representation from India form a separate cluster with very short branch lengths, implying a high degree of genetic similarity. This could mean they belong to a group that either diverged very recently or have undergone similar selective pressures leading to convergent evolution.

Divergence Between Major Groups

The large distance between the clusters on the left (e.g. F and J Sequence representation from China and Q sequence representation from USA, and O sequence representation from India) and the right (e.g. Four first sequence from Brazil and last sequence from Nigeria) reflects a substantial genetic divergence. These groups likely represent two different evolutionary paths, possibly originating from a common ancestor but diverging significantly due to geographic

DOI: 10.56892/bima.v8i4B.1171

separation, mutations, or environmental pressures.

The joining of these two groups at a higher distance on the dendrogram (around 0.35 on the y-axis) indicates that the two groups share fewer genetic characteristics and have diverged for a longer period.

Intermediate Clusters

P sequence presentation from USA and V&W sequence representation from Nigeria suggests that they form an intermediate evolutionary branch. Their genetic differences are smaller compared to the larger clusters on the left and right, which may suggest that they split from a common ancestor more recently than the major groups but earlier than the tightly clustered samples.

These sequence from USA and that up from Nigeria might be ancestral or transitional forms between the two larger clusters, reflecting intermediate evolutionary stages.

Application of the proposed Tree:

Evolutionary Studies

This phylogenetic tree could be used to study evolutionary relationships, helping researchers trace back to a common ancestor or identify evolutionary bottlenecks and divergence events which is the ultimate aim of the research work.

Medical Applications

The tree could help understand how different strains have evolved and spread, which has implications for tracking disease outbreaks and developing vaccines or treatments.

Acknowledgement

This research project received funding from **TETFUND** through the Institutional Based Research (**IBR**) with reference number: TETFUND/**DR&D/CE/UNI/GADAU/IBR/2023/VOL.1**

REFERENCES

36th European Congress of Pathology – Abstracts. (2024). *Virchows Archiv*, *485*(S1), 1–546. https://doi.org/10.1007/s00428-024-03880-y

Adamu, Y., & mathur, rinku. (2023). An alignment free mathematical model based on fuzzy dissimilarity metric to infer phylogenetic tree among Ebola viruses. *Journal of Emerging Technologies and Innovative Research*, *10*(3), 134–145.

Adegboye, O. A., Adekunle, A. I., Pak, A., Gayawan, E., Leung, D. Hy., Rojas, D. P., Elfaki, F., McBryde, E. S., & Eisen, D. P. (2021). Change in outbreak epicentre and its impact on the importation risks of COVID-19 progression: A modelling study. *Travel Medicine and Infectious Disease*, *40*, 101988. https://doi.org/10.1016/j.tmaid.2021.101988

Bei, H., Mao, Y., Wang, W., & Zhang, X. (2021). Fuzzy Clustering Method Based on Improved Weighted Distance. *Mathematical Problems in Engineering*, *2021*, 1–11. https://doi.org/10.1155/2021/6687202

Bloom, D. E., & Cadarette, D. (2019). Infectious Disease Threats in the Twenty-First Century: Strengthening the Global Response. *Frontiers in Immunology*, *10*, 549. https://doi.org/10.3389/fimmu.2019.00549

Bonham-Carter, O., Steele, J., & Bastola, D. (2014). Alignment-free genetic sequence comparisons: A review of recent approaches by word analysis. *Briefings in Bioinformatics*, *15*(6), 890–905. https://doi.org/10.1093/bib/bbt052

DOI: 10.56892/bima.v8i4B.1171

Chae, S. Y., Lee, K., Lee, H. M., Jung, N., Le, Q. A., Mafwele, B. J., Lee, T. H., Kim, D. H., & Lee, J. W. (2020). Estimation of Infection Rate and Predictions of Disease Spreading Based on Initial Individuals Infected With COVID-19. *Frontiers in Physics*, *8*, 311. https://doi.org/10.3389/fphy.2020.00311

Chakraborty, I., & Maity, P. (2020). COVID-19 outbreak: Migration, effects on society, global environment and prevention. *Science of The Total Environment*, *728*, 138882. https://doi.org/10.1016/j.scitotenv.2020.138882

Chan, C. X., Bernard, G., Poirion, O., Hogan, J. M., & Ragan, M. A. (2014). Inferring phylogenies of evolving sequences without multiple sequence alignment. *Scientific Reports*, *4*(1), 6504. https://doi.org/10.1038/srep06504

Cucinotta, D., & Vanelli, M. (2020). WHO Declares COVID-19 a Pandemic. *Acta Bio Medica Atenei Parmensis*, *91*(1), 157–160. https://doi.org/10.23750/abm.v91i1.9397

Drummond, A., & Rodrigo, A. G. (2000). Reconstructing Genealogies of Serial Samples Under the Assumption of a Molecular Clock Using Serial-Sample UPGMA. *Molecular Biology and Evolution*, *17*(12), 1807–1815. https://doi.org/10.1093/oxfordjournals.molbev.a026281

Huang, G., Zhou, H., Li, Y., & Xu, L. (2011). Alignment-free comparison of genome sequences by a new numerical characterization. *Journal of Theoretical Biology*, *281*(1), 107–112. https://doi.org/10.1016/j.jtbi.2011.04.003

Jakó, É., Ari, E., Ittzés, P., Horváth, A., & Podani, J. (2009). BOOL-AN: A method for comparative sequence analysis and phylogenetic reconstruction. *Molecular Phylogenetics and Evolution*, *52*(3), 887–897. https://doi.org/10.1016/j.ympev.2009.04.019

Krishnaratne, S., Littlecott, H., Sell, K., Burns, J., Rabe, J. E., Stratil, J. M., Litwin, T., Kreutz, C., Coenen, M., Geffert, K., Boger, A. H., Movsisyan, A., Kratzer, S., Klinger, C., Wabnitz, K., Strahwald, B., Verboom, B., Rehfuess, E., Biallas, R. L., … Pfadenhauer, L. M. (2022). Measures implemented in the school setting to contain the COVID-19 pandemic. *Cochrane Database of Systematic Reviews*, *2022*(2). https://doi.org/10.1002/14651858.CD015029

Lee, H.-S. (2001a). An optimal algorithm for computing the max–min transitive closure of a fuzzy similarity matrix. *Fuzzy Sets and Systems*, *123*(1), 129–136. https://doi.org/10.1016/S0165-0114(00)00062-2

Lee, H.-S. (2001b). An optimal algorithm for computing the max–min transitive closure of a fuzzy similarity matrix. *Fuzzy Sets and Systems*, *123*(1), 129–136. https://doi.org/10.1016/S0165-0114(00)00062-2

Lin, Y.-F., Duan, Q., Zhou, Y., Yuan, T., Li, P., Fitzpatrick, T., Fu, L., Feng, A., Luo, G., Zhan, Y., Liang, B., Fan, S., Lu, Y., Wang, B., Wang, Z., Zhao, H., Gao, Y., Li, M., Chen, D., … Zou, H. (2020). Spread and Impact of COVID-19 in China: A Systematic Review and Synthesis of Predictions From Transmission-Dynamic Models. *Frontiers in Medicine*, *7*, 321.

DOI: 10.56892/bima.v8i4B.1171

https://doi.org/10.3389/fmed.2020.00321

Raman Kumar, M., & Vaegae, N. K. (2020). A new numerical approach for DNA representation using modified Gabor wavelet transform for the identification of protein coding regions. *Biocybernetics and Biomedical Engineering*, *40*(2), 836–848. https://doi.org/10.1016/j.bbe.2020.03.007

Ren, J., Bai, X., Lu, Y. Y., Tang, K., Wang, Y., Reinert, G., & Sun, F. (2018). Alignment-Free Sequence Analysis and Applications. *Annual Review of Biomedical Data Science*, *1*(1), 93–114. https://doi.org/10.1146/annurev-biodatasci-080917-013431

Rust, J. (2019). Has Dynamic Programming Improved Decision Making? *Annual Review of Economics*, *11*(1), 833–858. https://doi.org/10.1146/annurev-economics-080218-025721

Saw, A. K., Raj, G., Das, M., Talukdar, N. C., Tripathy, B. C., & Nandi, S. (2019). Alignment-free method for DNA sequence clustering using Fuzzy integral similarity. *Scientific Reports*, *9*(1), 3753. https://doi.org/10.1038/s41598-019-40452-6

Thanh Le, T., Andreadakis, Z., Kumar, A., Gómez Román, R., Tollefsen, S., Saville, M., & Mayhew, S. (2020). The COVID-19 vaccine development landscape. *Nature Reviews Drug Discovery*, *19*(5), 305–306. https://doi.org/10.1038/d41573-020-00073-5

Yoshida, R., Paul, L., & Nesbitt, P. (2022). Stochastic Safety Radius on UPGMA. *Algorithms*, *15*(12), 483. https://doi.org/10.3390/a15120483.