Charurut Somboonwit · Paul Shapshak · Pandjassarame Kangueane · S. Balaji · John T. Sinnott · Lynette J. Menezes · Asa Oxner *Editors*

Global Virology IV: Viral Disease Diagnosis and Treatment Delivery in the 21st Century

Springer

Global Virology IV: Viral Disease Diagnosis and Treatment Delivery in the 21st Century

Charurut Somboonwit • Paul Shapshak
Pandjassarame Kangueane • S. Balaji
John T. Sinnott • Lynette J. Menezes
Asa Oxner
Editors

Global Virology IV: Viral Disease Diagnosis and Treatment Delivery in the 21st Century

Springer

Editors
Charurut Somboonwit
Department of Internal Medicine
Division of Infectious Disease and International Medicine
University of South Florida
Morsani College of Medicine
Tampa, FL, USA

Pandjassarame Kangueane
Irulan Sandy Annex
Biomedical Informatics
Pondicherry, Pondicherry, India

John T. Sinnott
Department of Internal Medicine
University of South Florida
Morsani College of Medicine
Tampa, FL, USA

Asa Oxner
University of South Florida
Division of General Internal Medicine
Morsani Center for Advanced Healthcare
Tampa, FL, USA

Paul Shapshak
Department of Internal Medicine
University of South Florida
Morsani College of Medicine
Tampa, FL, USA

S. Balaji
Department of Biotechnology
Manipal Institute of Technology
Manipal Academy of Higher Education
Manipal, Karnataka, India

Lynette J. Menezes
University of South Florida
Morsani College of Medicine
Tampa, FL, USA

ISBN 978-3-031-57368-2 ISBN 978-3-031-57369-9 (eBook)
https://doi.org/10.1007/978-3-031-57369-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The twenty-first century brings notable changes to clinical virology, mirrored by rapid advancements across science and technology. *Global Virology IV: Viral Disease Diagnosis and Treatment Delivery in the 21st Century* reflects this evolving landscape, offering insights into innovative methods enhancing our approach to viral diseases.

This book is intended for a broad audience, from medical students to experienced professionals. It provides an update on the latest advances in virology and sheds light on trends influencing clinical practice.

Recent global events, including pandemics and viral outbreaks, have highlighted the importance of innovation in our field, especially for the care of vulnerable populations. While we have yet to understand viruses like COVID-19 fully, this volume focuses on the underlying principles and advancements that are likely to have a significant impact.

This book presents a series of new concepts and methods. Respectful of its predecessors, this book builds upon the foundation laid by GVI through GVIII. We prioritize exploring new ideas and practices, ensuring fresh and relevant content. The chapters bridge the gap between current practice and future applications. This book is not a mere repetition of past volumes but a step to guide readers to the frontier of clinical virology.

With this volume, we invite you to explore the ongoing evolution of clinical virology, which remains as dynamic and challenging as ever. We hope to provide a helpful resource that sparks curiosity, informs practice, and contributes to improving global health.

Welcome to the state of virology today and the directions it may head in the future.

Tampa, FL, USA — Charurut Somboonwit
Tampa, FL, USA — Paul Shapshak
Pondicherry, Pondicherry, India — Pandjassarame Kangueane
Manipal, Karnataka, India — S. Balaji
Tampa, FL, USA — John T. Sinnott
Tampa, FL, USA — Lynette J. Menezes
Tampa, FL, USA — Asa Oxner

Introduction

Global Virology IV: Viral Disease Diagnosis and Treatment Delivery in the 21st Century explores our relentless battle against viral diseases. This volume will explore the historical context and then contemporary scientific advancements. Hopefully, a coherent narrative of our species' response to viral threats will be seen, and notable progress in treatment and prevention strategies will be presented.

We begin with a detailed analysis of emerging viruses, zeroing in on the genetic complexities of SARS-CoV-2 variants and their broader public health ramifications. The astute reader will note the focus shift to mankind's enduring struggle with perhaps our threat—the orthomyxoviruses. Always a dynamic challenge this pathogen continues to evolve.

Our microscope then explores Ebola's intricate virology, examines strategies for outbreak management, and discusses preventive measures. In the context of the Zika virus, the book highlights the recent leaps in genetic understanding and the diagnostic and therapeutic innovations triggered by this emerging virus.

Subsequent chapters delve into the omnipresent oncogenic human papillomavirus and its impact on genital and extragenital disease. Then, we focus on the problem of hepatitis C in specific populations and the most recent advances in treatment.

Our authors explore the intersection of emerging viral infections with health and social issues by further examining NeuroAIDS, the mental health implications of antiretroviral therapy, and even the effects of marijuana on people living with HIV.

The discussion addresses the challenges immunocompromised individuals face when dealing with viral infections, focusing on those undergoing hematopoietic cell transplants. It explores the repercussions of COVID-19 for cancer patients and stem cell transplant recipients and impacts of the Epstein-Barr Virus's infection pathways and prevention approaches, and considers the effects of Human Herpesviruses 6A, 6B, 7, and 8 on stem cell transplantation. The text discusses the management of infections such as Cytomegalovirus, Herpes Simplex, Varicella Zoster Virus, and Hemophagocytic Lymphohistiocytosis, and notes the progress in developing antiviral therapies.

The final chapters provide an update on the latest antiviral medications, from cidofovir to brincidofovir, and introduce the innovative domain of radioimmunotherapy, blending radiology with immunotherapy to create new treatments.

Finally, *Global Virology IV: Viral Disease Diagnosis and Treatment Delivery in the 21st Century* is an informative resource on virology, charting the scientific progress and ongoing efforts to tackle viral diseases.

Contents

The Genome Sequence Analysis of SARS-CoV-2 Variants of Concern

Adithi Somayaji and S. Balaji

Abstract The coronavirus pandemic has been at the forefront of the news, with many variants emerging with varied infectivity and transmissibility. These variants are the result of mutations in the virus. The evolutionary relationships of various SARS-CoV-2 variants were analyzed with a major focus on the Omicron variant. There were 61 sequence representatives obtained from GISAID and GenBank databases and were aligned and the phylogenetic trees were represented using MEGAX software. Evolutionary substitution models were utilized in conjunction with clustering approaches such as UPGMA, Neighbor-Joining (NJ), Minimum Evolution, and Maximum Composite Likelihood. The sequences of the omicron variant showed a close phylogenetic relationship with the Alpha variant UK (MZ376737.1), whereas the others showed significant distance from the Omicron variant.

Keywords SARS-CoV-2 · Genome analysis · Phylogenetics · Variants · Mutation · Omicron

1 Introduction

In the 1960s, scientists discovered that viruses found in some unhealthy animals had a similar bristly shape and spiky protein protrusions as a virus that causes the common cold in humans. These viruses resembled the solar corona under the electron microscope, prompting researchers to create the moniker coronaviruses in 1968. Coronaviruses were thought to only cause minor clinical manifestations in humans

A. Somayaji · S. Balaji (✉)
Department of Biotechnology, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India
e-mail: adithi.somayaji@learner.manipal.edu; s.balaji@manipal.edu

C. Somboonwit et al. (eds.), *Global Virology IV: Viral Disease Diagnosis and Treatment Delivery in the 21st Century*,
https://doi.org/10.1007/978-3-031-57369-9_1

until the 2003 outbreak of severe acute respiratory syndrome (SARS), which revealed how easily these adaptable viruses might kill people [1]. Now, SARS-CoV-2, a novel RNA virus from the *Coronaviridae* family, has resulted in a global outbreak, with the World Health Organization (WHO) declaring it a pandemic on March 11, 2020 [2].

The coronavirus genomes are the largest RNA viruses, with 30,000 genetic bases. SARS-CoV-2, unlike its close relatives, may infect human cells at various target sites, with the throat and lungs being the most common targets. Coronaviruses are also one of the few RNA viruses with a genetic proofreading system, which prevents the virus from accumulating mutations that would weaken it [1]. The receptor-binding domain (RBDs) of coronavirus spike proteins is critical for the virus's ability to penetrate human cells, with the SARS-CoV-2 binding domain being especially effective [3]. The spike proteins are the most distinct and important features of SARS-CoV-2 (and coronaviruses in general). These spike proteins were discovered to aid virus entry via the angiotensin-converting enzyme II (ACE2) receptor [4]. Additionally, proprotein convertase furin preactivates SARS-CoV-2 entry into the cells, lowering its reliance on target cell proteases. The SARS-CoV-2 can maintain efficient cell penetration while avoiding immunological detection due to the RBD's increased binding affinity to hACE2 and the furin preactivation of the spike protein [5]. The SARS-CoV-2 is reported to infect people within less than a fortnight of exposure, with an average incubation time of 4–5 days. Symptoms such as cough, cold, fever, weariness, and weakness are common. However, severe cases lead to lung damage, pneumonia, and mortality.

2 COVID-19 Variants and Implications

The coronavirus genome changes and evolves leading to variants. These variants are due to the accumulation of such changes or mutations in the viral genome. Many environmental mutagens modify the genomic composition of SARS-CoV-2 including metal ions, UV radiation, and endogenous elements of organisms [6]. The mutations in the RBD significantly improve the affinity of binding in the RBD-hACE2 complex and are connected to quick dissemination in human populations, and the term variant of concern is utilized [7]. At present, there are five variants of concern—Alpha, Beta, Gamma, Delta, and Omicron.

The alpha strain (B.1.1.7), first identified in the UK in September 2020 and drove the UK's second and third waves, is considered a variant of concern by the WHO [8]. The B.1.1.7 has gradually developed a foothold in the USA because of its high rate of transmission and lethality, which is 30–70% more than the original strain identified in Wuhan, China [9]. The efficacies of vaccines were reported in literature against this strain: 93.7% with the Pfizer-BioNTech vaccine, 100% with Moderna, 85.6% with Novavax vaccine, 71–91% Sinovac, and 74.5% Oxford-AstraZeneca vaccine [10].

The beta strain (B.1.351) was intially discovered in May 2020 in South Africa with higher infectivity in the younger age groups than earlier versions. Because of the E484K mutation, this strain is fatal and is the primary cause of the third wave in South Africa. Since then, the strain has been detected in 80 countries. This strain has been observed to evade several existing vaccines [8]. The beta variant showed complete resistance to therapeutically important monoclonal antibodies and neutralizing antibodies in the plasma of SARS-CoV-2 recovering patients in another in vitro research [9]. Early tests suggest that the Pfizer vaccine is slightly less effective against beta than the wild-type SARS-CoV-2 vaccine (about 75%), although both Moderna and Pfizer claim that their vaccines are still 95% effective against severe sickness and mortality [10].

The delta strain (B.1.617.2) originated in India (June 2021), had accounted for more than 80–90% of all cases in the UK and the USA. The devastating second wave of infections in India has been driven by the delta variant [8]. The infection rate of children and teens was five times more than older adults, and most of the infections were among unvaccinated individuals. It is also possible that some vaccines are less effective at neutralizing the delta version. According to a study, the delta variant is 60% more transmissible and it is more infective in the airways due to the increased amount of virus in the infected person (1260 times more than people infected with wild-type SARS-CoV-2), i.e., they can expel an increased amount of virus into the air [10]. The effectiveness of vaccines varies from 67% with Oxford-AstraZeneca, 88% with Pfizer-BioNTech, as well as 90% according to Sputnik V manufacturer [10].

The gamma strain (P.1) emerged in Manaus, Brazil (November 2020), was responsible for two Coronavirus outbreaks. Based on the patients' data, it was suggested that this strain is twice as transmissible as earlier strains. Only 54–79% protection was reported in this variation, resulting in a lethal second wave [9]. On November 26, 2021, WHO identified the variant Omicron (B.1.1.529) as a variant of concern (VOC) [11]. There were more than 30 mutations in the Omicron variant that have resulted in amino acid alterations in the spike protein sequences, 15 of which are situated in the RBD, which is critical for viral-cell contact mediated by the ACE-2 receptor. The Omicron spike gene sequence has been used to make inferences about the transmission rate. These findings revealed a group of alterations near the S1-S2 furin cleavage region that could improve viral infectivity. Furthermore, docking experiments revealed that a combination of mutations in the RBD would result in this variant's high binding affinity for human ACE2. This needs to be confirmed via in vitro studies [12].

There are variants categorized as "variants of interest" and "variants under monitoring." For instance, Lambda and Mu are designated as variants of interest as of February 2, 2022, while GR (B.1.1.318), GR (C.1.2), and GH/490R are classified as variants under monitoring. A genetic change linked to modifications in receptor binding, lower neutralization by antibodies established against a previous infection or vaccination, a projected increase in transmissibility or disease severity, reduced treatment efficacy, or potential diagnostic impact are all factors that contribute to make a variant, a variant of interest [13] (Table 1).

Table 1 Variant-specific data [14–16]

Variant	Pango Lineage	Notable mutations	Amino acid number	Amino acid change	Mutation location	Comments	Announced/ estimated appearance on	Country of origin
Alpha	B.1.1.7	S:N501Y	501	Asparagine to Tyrosine	Spike protein	May increase ACE2 binding Common to Beta B.1.351, Gamma P.1, Mu B.1.621, Omicron	14 Dec 2020	UK
		S:H69-	69	Histidine to deleted	Spike protein	Might affect antibody recognition, perhaps affecting some antibody-therapy treatments or immunity Common to Eta B.1.525, S:126A B.1.620, Omicron		
		N:D3L	3	Aspartic acid to leucine	Nucleocapsid	The N:D3L and N:S235F mutations in the nucleocapsid appear in the unstructured portions of the N-terminal domain and linker sections, respectively, outside the dimer and RNA interaction interfaces However, mutation research predicted that N:S235F confers a stabilizing impact on the nucleocapsid		
		N:S235F	235	Serine to Phenylalanine				

Variant	Pango Lineage	Notable mutations	Amino acid number	Amino acid change	Mutation location	Comments	Announced/ estimated appearance on	Country of origin
Beta	B.1.351	S:E484K	484	Glutamic acid to Lysine	Spike protein	S:E484 mutations may diminish convalescent serum neutralization substantially Common to Kappa B.1.617.1, Gamma, V3 P.1, Eta B.1.525, Iota B.1.526, Mu B.1.621, S:126A B.1.620, Omicron	Dec 2020	South Africa
		N:T205I	205	Threonine to Isoleucine	Nucleocapsid			
Gamma	P.1	N:P80R	80	Proline to Arginine	Nucleocapsid		Jan 2021	Brazil
Delta	B.1.617	S:L452R	452	Leucine to Arginine	Spike protein	14 of 35 RBD monoclonal antibodies had their neutralizing activity lowered or eliminated by S:L452R Common in Kappa B.1.617.1, Epsilon B.1.427/9	Late 2020	India
		S:P681H	681	Proline to Arginine	Spike protein	Might diminish the recognition by class 3 antibodies Located near the furin cleavage site, which could be crucial for immunological recognition Common in Kappa B.1.617.1, Omicron		
		S:G142D	142	Glycine to Aspartic acid	Spike protein	Found in the spike protein's N-terminal domain, is an antibody escape mutant that has been found in viruses cultured in the presence of a monoclonal antibody		

(continued)

Table 1 (continued)

Variant	Pango Lineage	Notable mutations	Amino acid number	Amino acid change	Mutation location	Comments	Announced/ estimated appearance on	Country of origin
Omicron	B.1.1.529	S:H655Y S:N679K S:P681H	655 679 681	Histidine to tyrosine Asparagine to Lysine Proline to histidine	Spike protein	A cluster of mutations associated with an increased rate of transmission at the S1-S2 Furin cleavage site	Nov 2021	Multiple countries
		S:Q498R S:N501Y	498 501	Glutamine to arginine Asparagine to tyrosine	Spike protein	Associated with increased binding affinity to ACE-2 receptor		
Lambda	C.37	S:L452Q	452	Leucine to Glutamine	Spike protein	Similar to S:L452R	Late 2020	South America
		S:R246-toS:G252-	246 252	Arginine and Glycine respectively to deleted	Spike	Is a characteristic 7 amino acid deletion		

Variant	Pango Lineage	Notable mutations	Amino acid number	Amino acid change	Mutation location	Comments	Announced/ estimated appearance on	Country of origin
Mu	B.1.621	S:Y144T and S:Y145S	144 145	Tyrosine to Threonine and Serine respectively	Spike protein	Spike protein has a 3 nucleotide/1 amino acid insertion with an insertion “N” at position 146	Jan 2021	Colombia
		S:D950N	950	Glutamic acid to Asparagine	Spike protein			
Iota	B.1.526	S:A701V	701	Alanine to Valine	Spike protein	S:A701V is close to the furin cleavage site, while S:D253G is in an N-terminal domain region that is a target for neutralizing antibodies S:A701V is common to Beta B.1.351 S:D253G is common to Lambda C.37	Late 2020	USA
		S:D253G	253	Aspartic acid to Glycine	Spike protein	S:A701V is close to the furin cleavage site, while S:D253G is in an N-terminal domain region that is a target for neutralizing antibodies. S:A701V is common to Beta B.1.351 S:D253G is common to Lambda C.37		
Kappa	B.1.617.1	S:L452R, S:E484Q, and S:P681R	452 484 681	Leucine to Arginine Glutamic acid to Glutamine Proline to Arginine	Spike protein	Moderately lower sensitivity to vaccine-elicited antibodies and may have decreased cell-entry efficiency	Oct 2020	India
Eta	B.1.525	S:Y144-	144	Tyrosine to deleted	Spike protein	Associated with certain antibody escape Common to Alpha B.1.1.7, Mu B.1.621, S:126A B.1.620, Omicron	Dec 2020	Multiple countries
		S:H69-	69	Histidine to deleted	Spike protein	Might affect antibody recognition, perhaps affecting some antibody-therapy treatments or immunity Common to Alpha B.1.1.7, S:126A B.1.620, Omicron		
		S:V70-	70	Valine to deleted	Spike protein	Common to Alpha B.1.1.7, S:126A B.1.620, Omicron		

3 Materials and Methods

3.1 Sequence Retrieval

The genome sequences of COVID-19 variants were retrieved from the GenBank and GSIAD database. One full sequence of each variant from at least one country was collected from the database as of February 2, 2022 (Table 2), with NC_045512 being the Wuhan reference sequence. The sequences were kept to a minimum of 61 genome sequences to avoid handling errors during analysis.

3.2 Sequence Alignment

The 60 representative genomes, which included the Wuhan reference sequence, were stored in FASTA format. The sequences were aligned in MEGA X software using ClustalX 2.1 option (default parameters) [17] and as well as Clustal Omega (online).

3.3 Codon Usage, Frequency Analysis, and Heatmaps

The data on codon usage (standard), nucleotide, and amino acid frequencies were obtained using MEGAX. The tab-separated value data sets (.csv files) of nucleotide and amino acid frequencies were used as input for ClustVis (https://biit.cs.ut.ee/clustvis/), an online program for creating heatmaps [18]. The unit variance method of row scaling and the singular value decomposition method of imputation was set as default options.

3.4 Phylogenetic Tree Construction

The MEGAX standalone package was used to obtain the phylogenetic tree using the Neighbor-Joining, UPGMA, minimum evolution, and maximum likelihood using composite maximum likelihood model and p-distances. The sequence file of the 60 representative genomes was uploaded to MEGAX to obtain the overall phylogenetic trees. The bootstrapping was enabled, and resampling was set to 100. The pairwise distance matrix was obtained using MEGAX. The .csv file of the matrix was used as input for ClustVis with default options.

Table 2 The analyzed dataset is displayed, along with accession numbers, sampled geographic location, variant type, year of collection, sequence length, and GC%

Sl. No	Accession No.	Location	Variant	Year (collection)	Length	GC%
1	NC_045512	Wuhan	Reference sequence	2019	29,903	37.9728
2	EPI_ISL_6756515	Japan	Alpha variant	2021	29,763	37.9800
3	MZ413979.1	Bangladesh	Alpha variant	2021	29,766	37.9567
4	MW509795.1	Egypt	Alpha variant	2021	29,793	37.9500
5	MW822594.1	Germany	Alpha variant	2020	29,851	37.9567
6	MT991088.1	Hong Kong	Alpha variant	2020	29,793	37.9669
7	MZ268634	India	Alpha variant	2020	29,790	37.9837
8	MZ376737.1	UK	Alpha variant	2020	29,840	37.9500
9	OK319445.1	USA	Alpha variant	2021	29,817	37.9500
10	EPI_ISL_5416540	Japan	Beta variant	2021	29,764	37.9721
11	MZ562483.1	Bangladesh	Beta variant (B.1.351.3)	2021	29,835	37.9433
12	MW580244.1	France	Beta variant (B.1.351.2)	2021	29,812	37.9770
13	MZ314998.1	Germany	Beta variant	2021	29,827	37.9601
14	MZ317890.1	India	Beta variant	2021	29,784	37.9635
15	MZ413998.1	Pakistan	Beta variant	2021	29,885	37.9365
16	OU471655.1	Slovakia	Beta variant	2021	29,885	38.0411
17	MZ202314.1	South Africa	Beta variant	2020	29,783	37.9584
18	MZ277392.1	Taiwan	Beta variant (B.1.351.3)	2021	29,880	37.9433
19	OU212934.1	UK	Beta variant	2021	29,885	37.9500
20	MW617734.1	USA	Beta variant	2021	29,812	37.9466
21	MZ212516.1	USA	Beta variant (B.1.351.3)	2021	29,881	37.9365
22	MZ213758.1	USA	Beta variant (B.1.351.2)	2021	29,880	37.9466
23	EPI_ISL_6832166	Japan	Delta variant	2021	29,769	37.9556
24	OK104648.1	Egypt	Delta variant (B.1.617.2)	2021	29,903	37.9331
25	MZ340535.1	India	Delta variant (B.1.617.2)	2021	29,802	37.9601
26	LC643049.1	Japan	Delta variant (B.1.617.2)	2021	29,882	37.9196
27	OU659423.1	UK	Delta variant (B.1.617.2)	2021	29,890	37.9669
28	MZ436591.1	USA	Delta variant (AY.2)	2021	29,844	37.9365
29	MZ491704.1	USA	Delta variant (AY.3)	2021	29,779	37.9331

(continued)

Table 2 (continued)

Sl. No	Accession No.	Location	Variant	Year (collection)	Length	GC%
30	MZ892623.1	Uzbekistan	Delta variant (B.1.617.2)	2021	29,825	37.9534
31	OD990459	UK	Epsilon variant (B.1.427)	2021	29,903	37.9095
32	MW795884	USA	Epsilon variant (B.1.427)	2021	29,904	37.9196
33	MZ970734	Djibouti	Eta variant	2021	29,709	37.9635
34	MW598416	Ghana	Eta variant	2021	29,888	37.9669
35	OD976033	UK	Eta variant	2021	29,830	37.9635
36	MW974897	USA	Eta variant	2021	29,709	37.9669
37	EPI_ISL_6228367	Japan	Gamma variant	2021	29,768	39.3140
38	MZ020420.1	Bangladesh	Gamma variant	2021	29,853	37.9348
39	MZ477753.1	Brazil	Gamma variant	2021	29,735	37.9601
40	MW642248	Italy	Gamma variant	2021	29,853	37.9298
41	OU385891.1	UK	Gamma variant	2021	29,898	38.0985
42	MW963199.1	USA	Gamma	2021	29,777	37.9534
43	MW967498.1	USA	Gamma (P.1.2)	2021	29,897	37.9196
44	MZ611957.1	Venezuela	Gamma	2021	29,810	37.9365
45	EPI_ISL_6887009	France	GH490R	2021	29,724	37.9559
46	EPI_ISL_6910522	USA	hCoV	2021	29,807	37.9609
47	MW984875	USA	Iota variant (B.1.526)	2021	29,704	37.9837
48	MZ315140	Germany	Kappa variant	2021	29,781	37.9736
49	MZ314697	India	Kappa variant	2021	29,801	37.9871
50	MZ571142	Morocco	Kappa variant	2021	29,818	37.9635
51	OU391429	UK	Kappa variant	2021	29,902	38.3144
52	MZ257515	USA	Kappa variant	2021	29,876	37.9736
53	MZ275289.1	Peru	Lambda variant (C.37)	2021	29,781	37.9365
54	EPI_ISL_4470504	Japan	Mu GH	2021	29,781	37.9537
55	OU476023	Slovakia	Mu variant(B.1.621)	2021	29,899	37.9196
56	MZ348460	USA	Mu variant(B.1.621)	2021	29,839	37.9888
57	MZ710933	USA	Mu variant(B.1.621)	2021	29,881	37.9365
58	OK124281	USA	Mu variant (B.1.621)	2021	29,693	37.9466
59	OL677199.1	Canada	Omicron variant	2021	29,684	37.9497
60	EPI_ISL_6640916	Botswana	Omicron variant	2021	29,684	37.9813

4 Results and Discussion

4.1 Codon Usage, Frequency Analysis, and Heatmaps

For 60 genomes, the average frequency (in percentage) of A, C, G, and T was 32.182, 18.3362, 29.865, and 19.6152, respectively. The average computations did not include unresolved bases ("N"), which can occur owing to weak signal quality, base call errors, and other factors which still need to be eliminated while trimming [19].

The GC% of the entire dataset was calculated to be 37.98%, while the GT% was computed to be 51.79% due to the higher frequency of Thymine bases. The amino acid frequency and nucleotide frequency of the 60 sequences are depicted as heatmaps in Fig. 1a, b respectively. The nucleotide heat map (Fig. 1a) likewise shows that T/U is the most common, followed by A. The most common codon was UUU (354) for phenylalanine, while the least common was CCG (25.7) for alanine, according to the codon usage frequency with UGU (314.9) also observed under the high-frequency bracket. Several codons come under a low-frequency bracket, including CGA, CGC, etc.

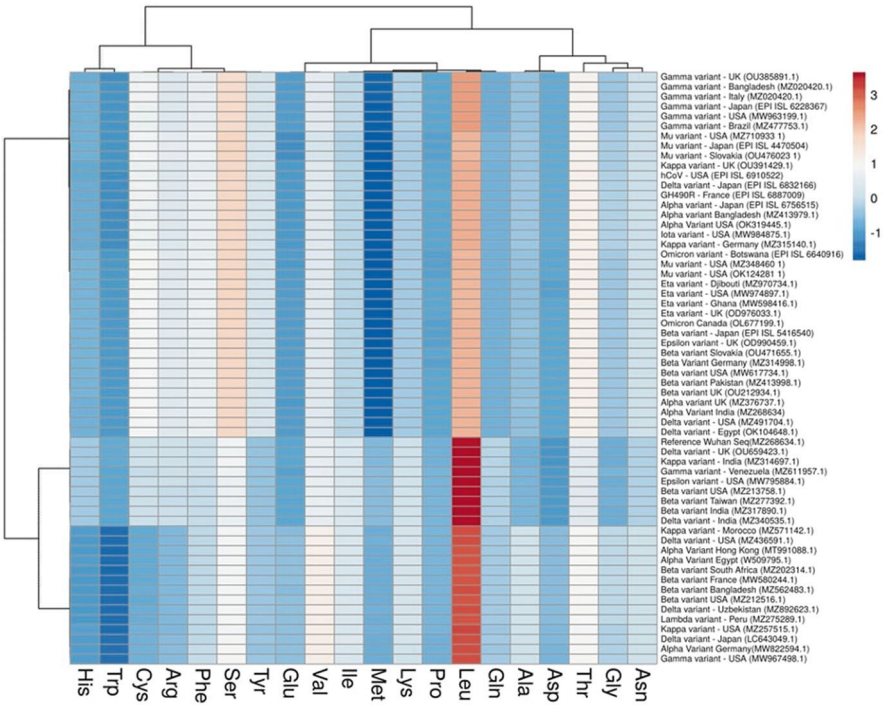

Fig. 1 (**a**) Amino acid frequency. (**b**) Nucleotide frequency

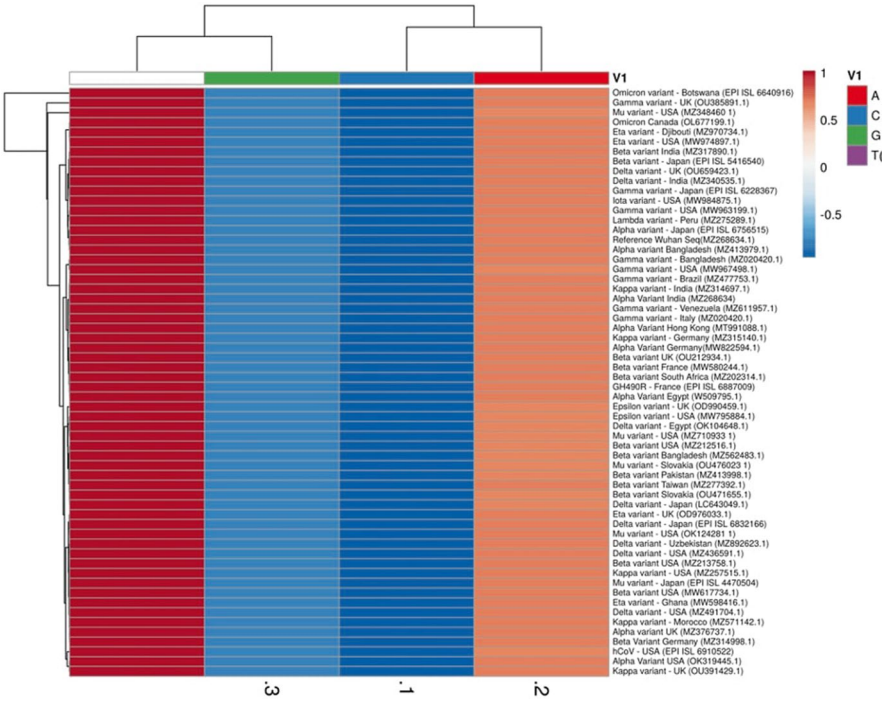

Fig. 1.1 (continued)

The amino acid heatmap (Fig. 1b) indicates a relatively leucine-rich pattern, followed by serine and threonine. The low-frequency bracket color of dark blue confirms the lack of ambiguous bases in the sequences. The least abundant amino acid was tryptophan (2.2963%), and the most abundant amino acid was computed to be leucine (11.673%), as shown in the figure. In both the heatmaps, three different clusters were observed with varied locations and variants. In the first cluster, all the variants of concern including omicron and at least one of the other variants were observed. Most of the sequences isolated from the USA were present in this cluster. The second cluster comprises the delta, gamma, beta, epsilon, and kappa from varied geographic locations. The reference sequence was also observed in the second cluster. In the third cluster, all the variants of concern, except omicron and Kappa and lambda variants, were observed. The geographical patterns were only readily observed in some of the clusters.

4.2 Phylogenetic Tree Construction

Two omicron variants were used to conduct the phylogenetic analysis: one from Botswana (the first sequenced omicron genome) and the other from Canada. They were aligned with alpha, beta, gamma, delta, mu, lambda, GH490R, kappa, eta, and

epsilon variant genome sequences sequenced at various locations as well as SARS-CoV-2 USA isolates using online clustal omega website and MEGAX software. The phylogenetic analysis of the omicron variants was done using UPGMA/Maximum composite likelihood (MCL), UPGMA/p-distance, Neighbor-joining (NJ)/MCL, Maximum likelihood (ML)/MCL, and Minimum evolution/p-distance using MEGAX software as shown in Figs. 2, 3, 4, 5, and 6. The analysis conducted in this paper is like the study by Kandeel et al. 2021 [20].

All three phylogenetic trees showed a close relationship with the Alpha variant UK (MZ376737.1) and the Omicron variants. When developing the tree, the UPGMA model assumes all lineages advance at the same rate, and the mutation rate is ignored. The pairwise distance is used to build the tree. During tree construction, the NJ, on the other hand, examines the pace of evolution. The aggregate of linked log-likelihoods is referred to as a composite likelihood. Because the evolutionary links among the sequences cause all pairwise distances in a distance matrix to have correlations, the sum of their log-likelihoods is a composite likelihood. This distance is just the proportion (p) of nucleotide sites that differ between the two sequences being compared. As the study [20] suggests, a single evolution model cannot be used due to complex evolutionary methods of viruses and single-gene variations.

The distance matrix also shows the minimum distance between omicron variants and the alpha UK variant as shown in the heatmap (Fig. 7). Using NCBI nucleotide Blast, the identity percentage between Alpha variant and Omicron sequences was calculated, and Omicron Botswana-Alpha UK showed 99.69% identity and Omicron_Canada-Alpha_UK showed 99.97% identity. The reference sequence and the omicron variants showed significant distance (although not the highest), implying that the omicron variant is highly mutated. The Omicron variants are phylogenetically distant from the other variants, as shown by all four methods. However, in the UPGMA/MCL and UPGMA/p-distance methods, the two omicron sequences were more closely related than in the NJ/MCL and ML/MCL methods, where the omicron_Canada and alpha_UK were more closely related. This could be due to UPGMA models assuming the same evolutionary speed of all lineages and ignoring the mutation rates.

The delta and kappa variants were almost always clustered together, which could be due to their Indian origins. The Mu variant USA (MZ710993) and the GH490R_ France variant showed a significant phylogenetic distance from other sequences. The beta variant India sequence showed the closest relationship with the reference sequence in all the phylogenetic methods.

Contrary to the study [20] from where we used the same sequences (GISAID obtained sequences), we observed that the Alpha variant Japan was not closely related to the omicron variants and instead had a closer relationship with some of the beta variant sequences (Slovakia, Japan, Pakistan, UK). This could be due to the higher number of sequences used, the different methods of phylogenetic tree construction, or changes in the models used to calculate distances.

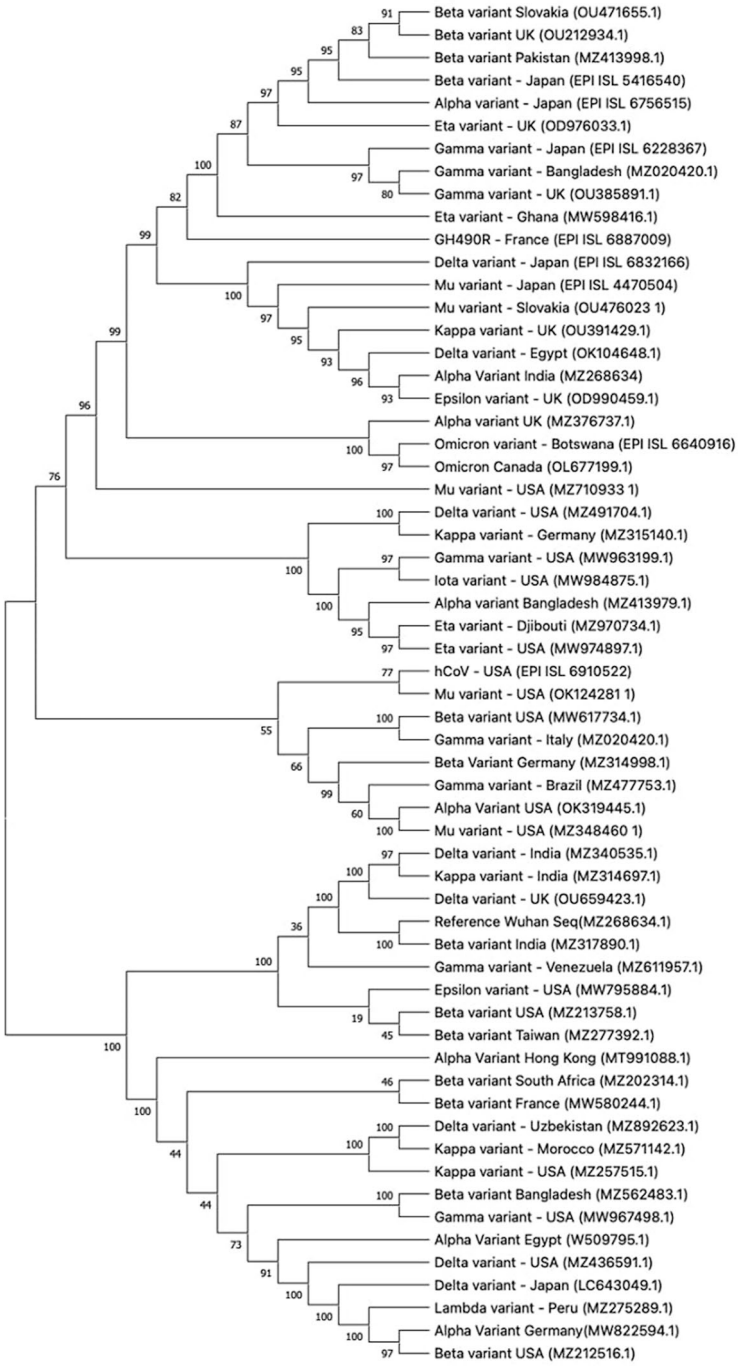

Fig. 2 Phylogenetic tree developed using UPGMA/Maximum composite likelihood model. Software used: MEGAX

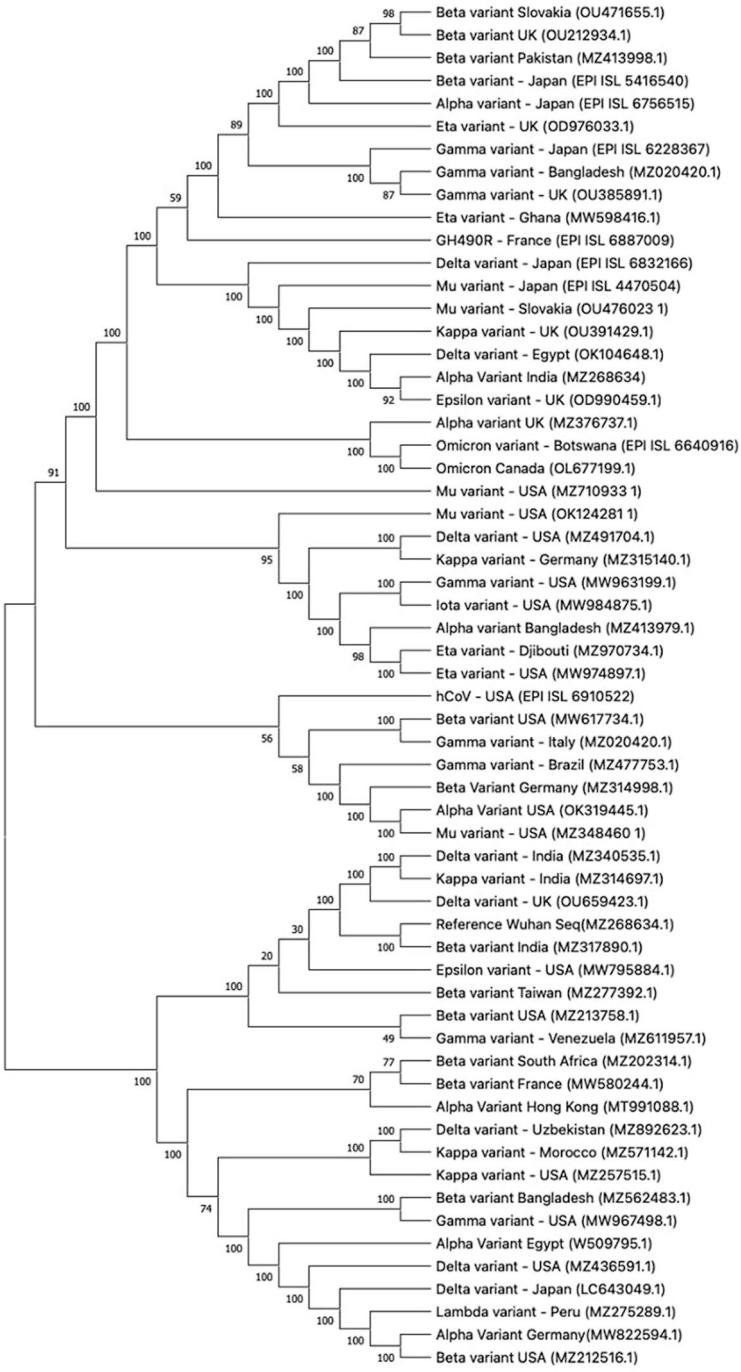

Fig. 3 Phylogenetic tree developed using UPGMA/p-distance model. Software used: MEGAX

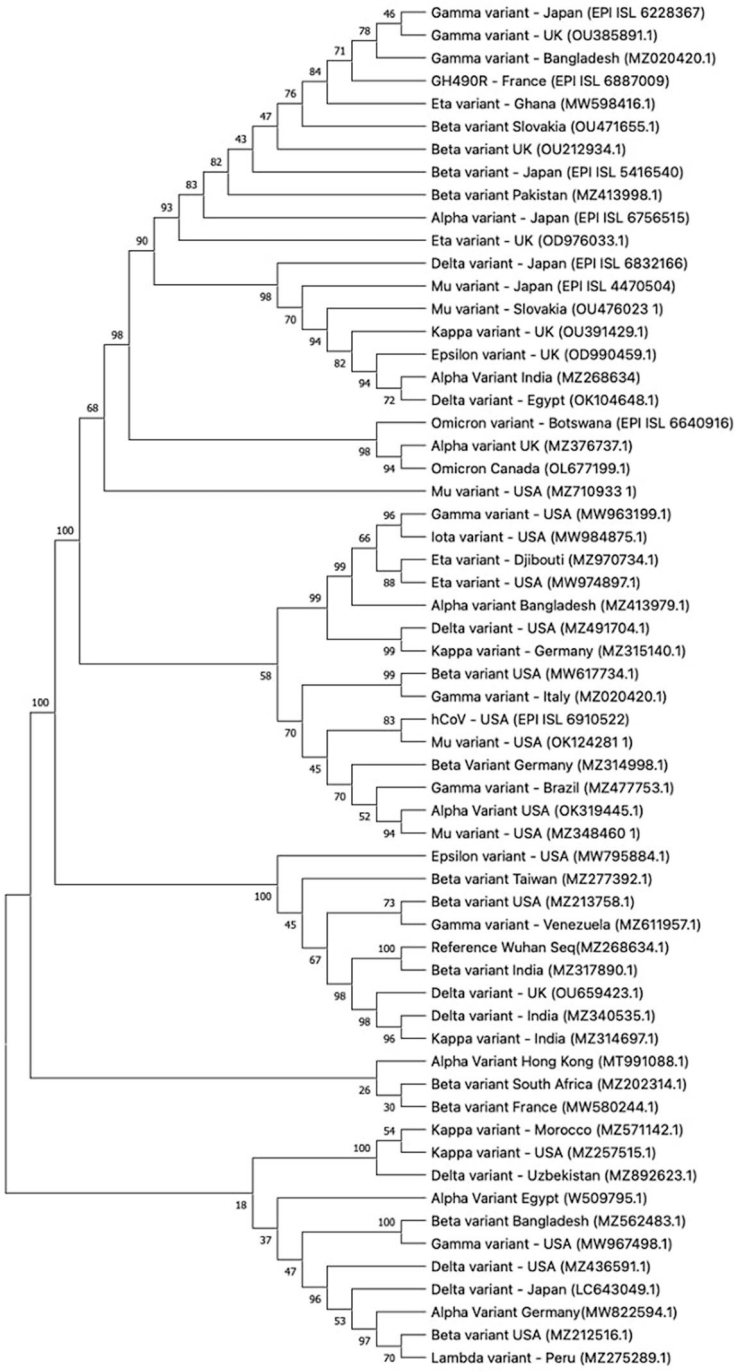

Fig. 4 Phylogenetic tree developed using Neighbor-Joining/Maximum composite likelihood model. Software used: MEGAX

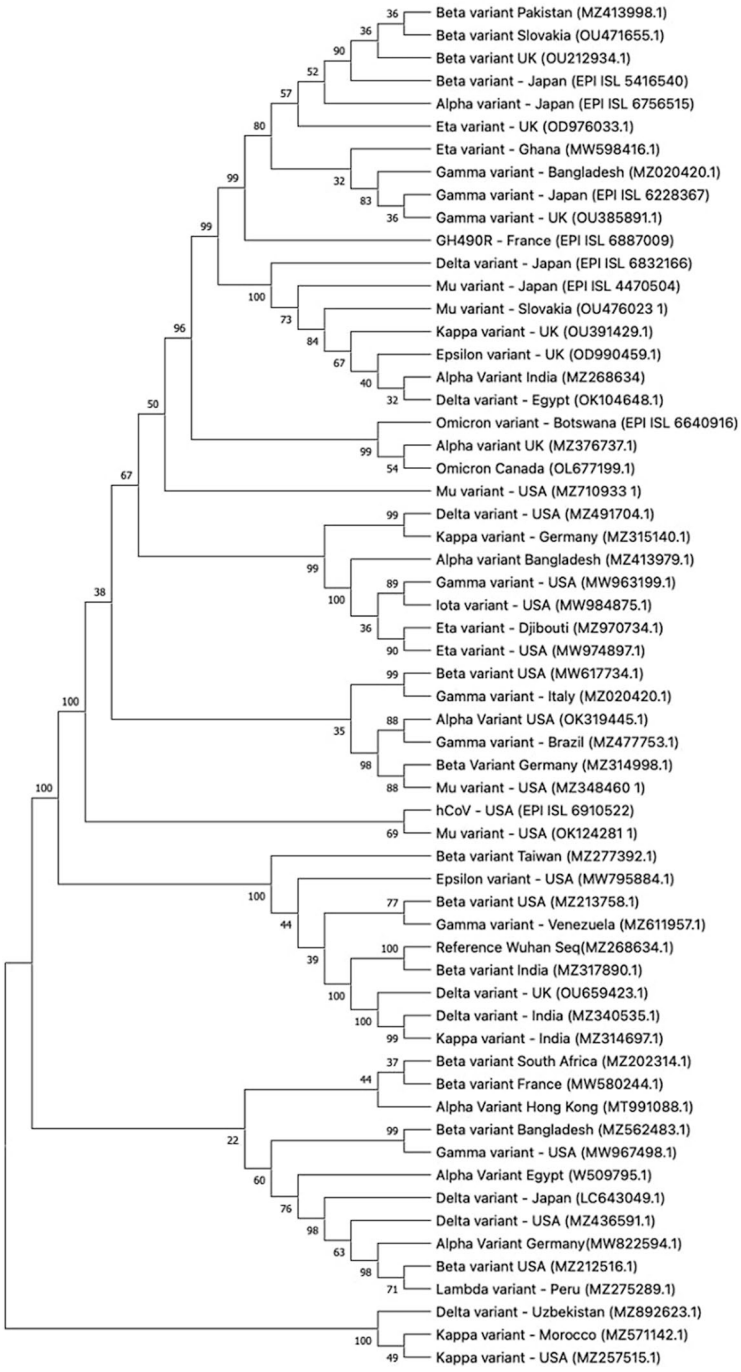

Fig. 5 Phylogenetic tree developed using Maximum Likelihood/Maximum composite likelihood model. Software used: MEGAX

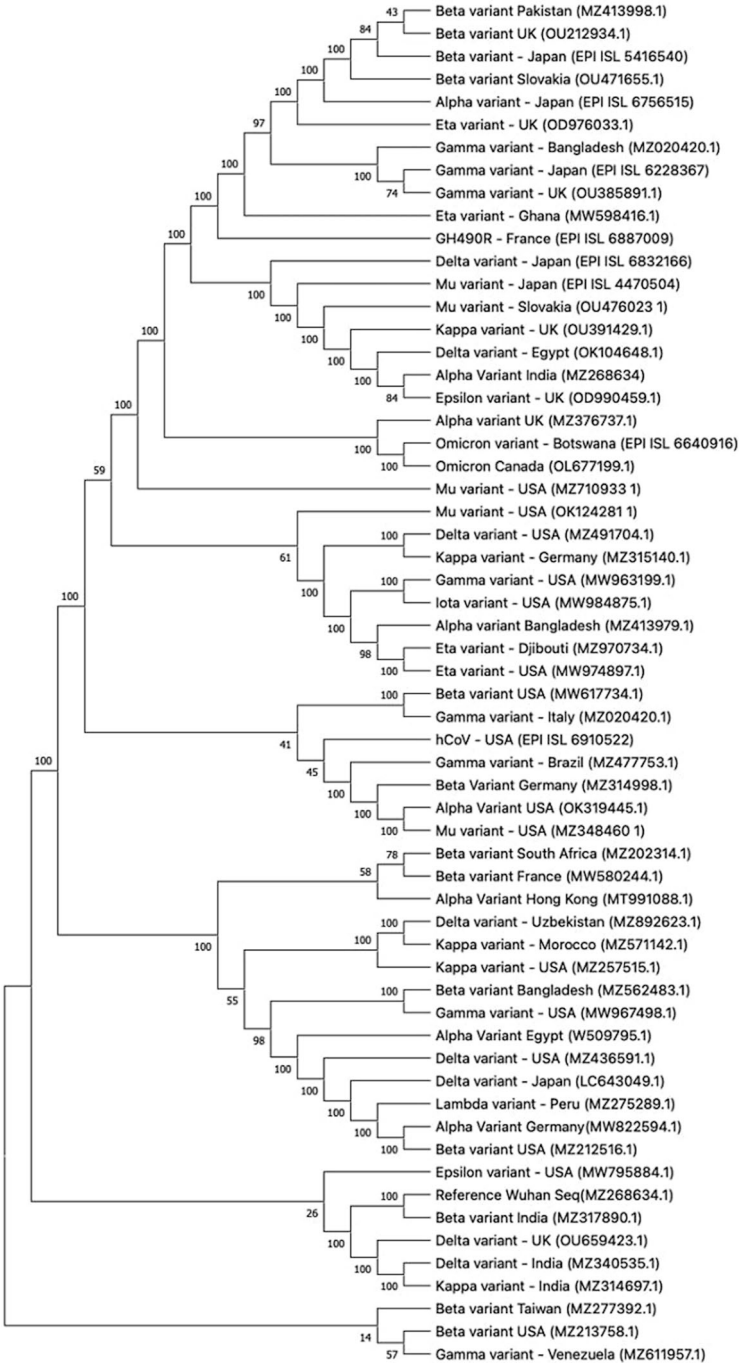

Fig. 6 Phylogenetic tree developed using Minimum evolution/p-distance model. Software used: MEGAX

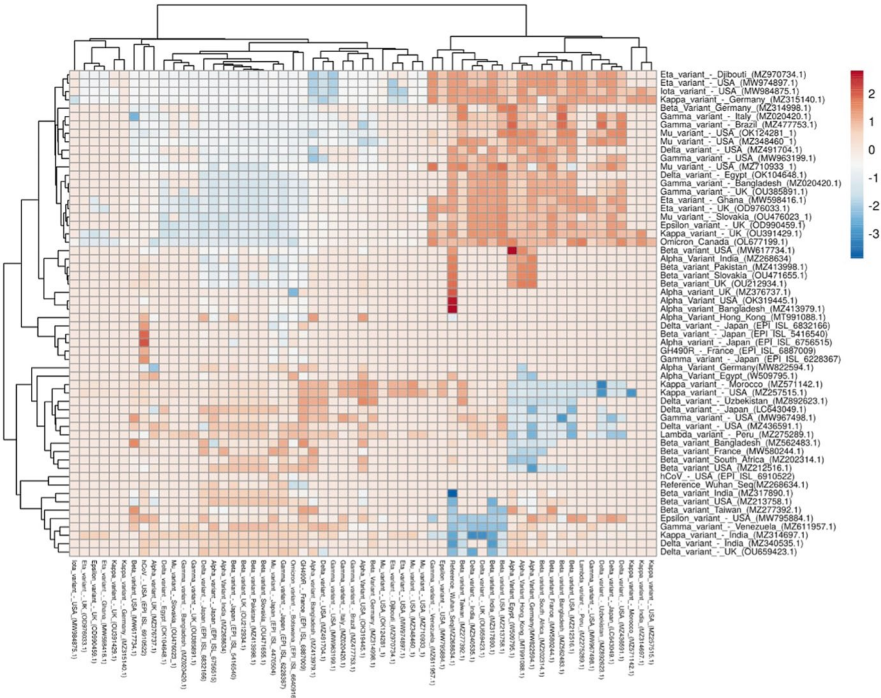

Fig. 7 Pairwise distance matrix heatmap. Software used: ClustVis (Online)

Moreover, similar to the phylogenetic results of the study by Yamin et al., the phylogenetic trees did not have any intermediate branches of evolution from the omicron variant (exception of Alpha variant UK), implying that the variant evolved differently from all the variants [21]. In addition, similar to their conclusion regarding the relationship between the delta variant and omicron, this study revealed a weak evolutionary relationship. In contrast to previous findings regarding gamma and omicron variants, this study did not find any relationship between gamma and omicron variants, which may be due to the differences in the analysis methods. Moreover, the covariant data demonstrates that the omicron variant has a very distinct mutation profile compared to other variants [14].

Acknowledgments The authors would like to acknowledge the support received from Manipal Institute of Technology and Manipal Academy of Higher Education.

Declaration of Competing Interests The authors declare no conflict of interest.

References

1. Cyranoski D (2020) Profile of a killer: the complex biology powering the coronavirus pandemic. In: Nature. https://www.nature.com/articles/d41586-020-01315-7. Accessed 20 Jun 2021
2. Behera P, Mazumder A, Arora M, et al (2020) SARS-CoV-2 epidemic in India: Epidemiological features and in silico analysis of the effect of interventions. F1000Res. https://doi.org/10.12688/f1000research.23496.1
3. Zhang T, Wu Q, Zhang Z (2020) Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak. Current Biology 30:1346–1351.e2
4. Zhou H, Chen X, Hu T, et al (2020) A Novel Bat Coronavirus Closely Related to SARS-CoV-2 Contains Natural Insertions at the S1/S2 Cleavage Site of the Spike Protein. Current Biology 30:2196–2203.e3
5. Wang MY, Zhao R, Gao LJ, Gao XF, Wang DP, Cao JM (2020) SARS-CoV-2: Structure, Biology, and Structure-Based Therapeutics Development. Frontiers in Cellular and Infection Microbiology. https://doi.org/10.3389/fcimb.2020.587269
6. Sanjuán R, Domingo-Calap P (2016) Mechanisms of viral mutation. Cellular and Molecular Life Sciences 73:4433–4448
7. (2022) Variant of concern. In: Wikipedia. https://en.wikipedia.org/wiki/Variant_of_concern. Accessed 20 Jun 2021
8. Roy B, Dhillon J, Habib N, Pugazhandhi B (2021) Global variants of COVID-19: Current understanding. Journal of Biomedical Sciences 8:8–11
9. Robishaw JD, Alter SM, Solano JJ, Shih RD, DeMets DL, Maki DG, Hennekens CH (2021) Genomic surveillance to combat COVID-19: challenges and opportunities. The Lancet Microbe 2:e481–e484
10. Mahase E (2021) Covid-19: How many variants are there, and what do we know about them?: Video 1. BMJ n1971
11. Classification of omicron (B.1.1.529): SARS-CoV-2 variant of concern. https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern. Accessed 10 Jan 2022
12. Ferré VM, Peiffer-Smadja N, Visseaux B, Descamps D, Ghosn J, Charpentier C (2022) Omicron SARS-CoV-2 variant: What we know and what we don't. Anaesthesia Critical Care and Pain Medicine. https://doi.org/10.1016/j.accpm.2021.100998
13. (2021) SARS-CoV-2 Variant Classifications and Definitions. In: Centres for disease control and prevention. https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-classifications.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fvariants%2Fvariant-info.html. Accessed 1 Jan 2022
14. Overview of Variants/Mutations. In: GISAID. https://covariants.org/variants. Accessed 1 Jan 2022
15. Outbreak.info. https://outbreak.info. Accessed 7 Sep 2021
16. Detect and analyze variants of SARS-CoV-2. https://cov-spectrum.org/explore/Switzerland/AllSamples/AllTimes. Accessed 2 Sep 2021
17. Stecher G, Tamura K, Kumar S (2020) Molecular evolutionary genetics analysis (MEGA) for macOS. Molecular Biology and Evolution 37:1237–1239
18. Metsalu T, Vilo J (2015) ClustVis: A web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. Nucleic Acids Research 43:W566–W570
19. Seetharaman B, Ramachandran A, Nandy K, Paul S (2017) Sequence accuracy in primary databases: A case study on HIV-1B. In: Global Virology II - HIV and NeuroAIDS. Springer New York, pp 779–822
20. Kandeel M, Mohamed MEM, Abd El-Lateef HM, Venugopala KN, El-Beltagi HS (2021) Omicron variant genome evolution and phylogenetics. Journal of Medical Virology. https://doi.org/10.1002/jmv.27515
21. Sun Y, Lin W, Dong W, Xu J (2022) Origin and evolutionary analysis of the SARS-CoV-2 Omicron variant. Journal of Biosafety and Biosecurity 4:33–37

Influenza: Clinical Challenges in the Twenty-First Century

Mindy M. Sampson, Christopher M. Polk, Anupama Neelakanta, and Catherine L. Passaretti

Abstract Influenza has had a significant impact on society and human health and may have had a similar impact in earlier historical times, if post-hoc diagnoses are correct. Significant progress has been made within diagnostics, vaccines, treatment, and infection control practices for influenza, which has resulted in improved control of this virus at both the seasonal epidemic and pandemic levels. Despite advances in knowledge of influenza, it is important that we continue to improve access to healthcare and address disparities between many populations. Poor outcomes occur when people infected with influenza have less access to healthcare or refuse it. With improved understanding of viral evolution, it is also critical for public health agencies to plan for continual future influenza epidemics and pandemics.

Keywords Influenza · Pandemic · Epidemics · H1N1 · Vaccine · Age · Race · Sex · Diagnostic testing · Baloxivir · Oseltamavir · Infection control · Healthcare · Disparities · Minorities

M. M. Sampson (✉)
Department of Internal Medicine, Stanford University, Palo Alto, CA, USA
e-mail: mindysam@stanford.edu

C. M. Polk
Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA
e-mail: Christopher.Polk@atriumhealth.org

A. Neelakanta · C. L. Passaretti
Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA

Department of Internal Medicine, Atrium Health, Charlotte, NC, USA
e-mail: Anupama.Neelakanta@atriumhealth.org; Catherine.Passaretti@atriumhealth.org

C. Somboonwit et al. (eds.), *Global Virology IV: Viral Disease Diagnosis and Treatment Delivery in the 21st Century*,
https://doi.org/10.1007/978-3-031-57369-9_2