life

MDPI

Systematic Review

A Global Mutational Profile of SARS-CoV-2: A Systematic Review and Meta-Analysis of 368,316 COVID-19 Patients

Wardah Yusof [1], Ahmad Adebayo Irekeola [1,2], Yusuf Wada [1,3], Engku Nur Syafirah Engku Abd Rahman [1], Naveed Ahmed [1], Nurfadhlina Musa [4], Muhammad Fazli Khalid [5], Zaidah Abdul Rahman [1,6], Rosline Hassan [7], Nik Yusnoraini Yusof [5] and Chan Yean Yean [1,5,*]

1 Department of Medical Microbiology and Parasitology, School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian 16150, Kelantan, Malaysia; wardahyusof@usm.my (W.Y.); irekeola@student.usm.my (A.A.I.); wadayusuf@student.usm.my (Y.W.); engkunursyafirah@student.usm.my (E.N.S.E.A.R.); naveed.malik@student.usm.my (N.A.); drzaidah@usm.my (Z.A.R.)
2 Microbiology Unit, Department of Biological Sciences, College of Natural and Applied Sciences, Summit University, Offa 250101, Kwara State, Nigeria
3 Department of Zoology, Faculty of Life Sciences, Ahmadu Bello University, Zaria 810107, Kaduna State, Nigeria
4 Human Genome Centre, School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian 16150, Kelantan, Malaysia; fadhlina@usm.my
5 Institute for Research in Molecular Medicine (INFORMM), Universiti Sains Malaysia, Kubang Kerian 16150, Kelantan, Malaysia; fazlikhalid@usm.my (M.F.K.); nikyus@usm.my (N.Y.Y.)
6 Hospital Universiti Sains Malaysia, Universiti Sains Malaysia, Kubang Kerian 16150, Kelantan, Malaysia
7 Department of Hematology, School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian 16150, Kelantan, Malaysia; roslin@usm.my
* Correspondence: yychan@usm.my; Tel.: +60-129-011066

Abstract: Since its first detection in December 2019, more than 232 million cases of COVID-19, including 4.7 million deaths, have been reported by the WHO. The SARS-CoV-2 viral genomes have evolved rapidly worldwide, causing the emergence of new variants. This systematic review and meta-analysis was conducted to provide a global mutational profile of SARS-CoV-2 from December 2019 to October 2020. The review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA), and a study protocol was lodged with PROSPERO. Data from 62 eligible studies involving 368,316 SARS-CoV-2 genomes were analyzed. The mutational data analyzed showed most studies detected mutations in the Spike protein ($n = 50$), Nucleocapsid phosphoprotein ($n = 34$), ORF1ab gene ($n = 29$), 5′-UTR ($n = 28$) and ORF3a ($n = 25$). Under the random-effects model, pooled prevalence of SARS-CoV-2 variants was estimated at 95.1% (95% CI; 93.3–96.4%; $I^2 = 98.952\%$; $p = 0.000$) while subgroup meta-analysis by country showed majority of the studies were conducted 'Worldwide' ($n = 10$), followed by 'Multiple countries' ($n = 6$) and the USA ($n = 5$). The estimated prevalence indicated a need to continuously monitor the prevalence of new mutations due to their potential influence on disease severity, transmissibility and vaccine effectiveness.

Keywords: COVID-19; SARS-CoV-2; mutation; mutational profile

check for updates

Citation: Yusof, W.; Irekeola, A.A.; Wada, Y.; Engku Abd Rahman, E.N.S.; Ahmed, N.; Musa, N.; Khalid, M.F.; Rahman, Z.A.; Hassan, R.; Yusof, N.Y.; et al. A Global Mutational Profile of SARS-CoV-2: A Systematic Review and Meta-Analysis of 368,316 COVID-19 Patients. *Life* **2021**, *11*, 1224. https://doi.org/10.3390/life11111224

Academic Editor: Patrick Mercié

Received: 12 October 2021
Accepted: 8 November 2021
Published: 11 November 2021

1. Introduction

Coronavirus Disease-19 (COVID-19) is caused by Severe Acquired Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) [1,2]. Since the SARS-CoV-2 epidemic first reported in Wuhan, China, the clinical features of COVID-19 have evolved, moving from clinically apparent pulmonary or flu-like symptoms to subclinical or even silent infections. The COVID-19 infection could frequently involve an asymptomatic or paucisymptomatic framework, leading to a spread in the general population [3]. COVID-19 produces respiratory distress with mild to severe symptoms, and it is fatal in individuals with a chronic disease or a compromised immune system [4]. Various clinical outcomes in COVID-19

patients have also been documented throughout several other regions across the world. As of 29 September 2021, according to the World Health Organization (WHO), the SARS-CoV-2 pandemic has infected over 232,075,351 individuals across the world, resulting in 4,752,988 fatalities and significant disruptions to regular activities and national economies.

The international scientific community continuously characterized the pathophysiological features of COVID-19, developed diagnostic tools, evaluated immune responses, and identified risk factors for severe illness courses. SARS-CoV-2 clustered outbreaks and super spreading episodes provide a unique challenge to pandemic control [5]. However, the basic characteristics of SARS-CoV-2 genome evolution and transmission dynamics within the human population are still unknown [6]. COVID-19 infection demonstrated related inflammatory state of the upper airway mucosa and olfactory neurotoxic damage. However, to date, a reliable method in the evaluation of the nasal health of post-infection patients is not clear [7].

SARS-CoV-2 genomic sequencing from several geographical regions has recently revealed that the virus quickly changes by accumulating mutations in its genome. It has been proposed that new SARS-CoV-2 variants may adapt better to new geographical locations, making them more potent than the virus that discovered in Wuhan, China.

All viruses' genomes gain mutations over time. However, various variables, including the mutation rate and the effects of mutation on viral dynamics within and between individual hosts, influence the rate of mutation accumulation and its repercussions for transmission and illness in the host population [8]. The combination of these variables determines the development and transmission of viral variations and the evolution of pandemics. Detection of mutations spread worldwide is essential for a better understanding of the viral evolution, bio-pathology and transmission since RNA virus genomes are highly susceptible to mutation [9].

2. Materials and Methods

2.1. Study Design and Protocol

The Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocol (PRISMA-P 2015) guidelines [10] were used as this study's checklist. The study population included individuals with SARS-CoV-2 infection with the main out-come being mutations in the SARS-CoV-2. Reference was made to the Wuhan strain as a comparator A Prospero protocol (No CRD42021229620) was lodged for this study.

2.2. Literature Review

The PROSPERO database and Database of Abstracts of Reviews of Effects (DARE) (http://www.library.UCSF.edu; accessed on 10 January 2021) were searched to ensure no other meta-analysis on the impact of the mutational profile of SARS-CoV-2 on transmissibility and disease severity exists or is ongoing. The literature search was performed using international databases PubMed, Scopus, Science Direct and Google Scholar using the search terms listed in Table S2. Two authors carried out the database search to minimize bias.

2.3. Inclusion and Exclusion Criteria for Studies

Inclusion criteria: (1) Studies reporting on human COVID-19, (2) Studies reporting on SARS-CoV-2 mutations, (3) Studies reporting on SARS-CoV-2 mutations and their association with superspreading events, transmissibility and severity of illness in COVID-19 patients. Exclusion criteria include reviews papers, animal studies, protein characterization studies, studies on environmental sampling, and media reports.

2.4. Quality Assessment

The methodological quality of the included studies was assessed independently by two authors using the Joanna Briggs Institute (JBI) critical appraisal checklist for prevalence data [11]. A score of '1' for "yes" and '0' for other parameters was assigned to attain a total

quality score ranging from '0' to '9'. Studies with an overall score of '7'–'9' were considered sufficient quality (Table S3).

2.5. Data Extraction

Two independent authors performed the data extraction by using standardized forms, which included manuscript title, authors, journal, publication year, countries of study, period of study, number of participants, number of mutated cases, regions of mutations, types of mutations, mutations, viral load, symptoms, severity (mild, moderate, severe, fatal), sample types (nasopharyngeal swab, bronchoalveolar lavage), viral shedding, co-morbidity, mutation detection method, the database used (data downloaded), database accessed and transmissibility.

Studies that analyzed genetic mutations from more than one country were categorized as "multiple countries" rather than the individual countries included. When mutational data from different countries and regions were analyzed as a whole, instead of by specific countries, they were characterized as 'worldwide', and the data were extracted and analyzed in that form to avoid confusion. For regions of mutations labelled as ORF1a, ORF1b, nsp1-14, 3C-like proteinase, RNA-dependent RNA polymerase (RdRp), helicase, 3′-to-5′ exonuclease, endoRNAse, 2′-O-ribose methyltransferase, or leader protein, they were characterized as 'ORF1ab' to simplify analysis. Where more than one article reported mutational data from the same group of sample, record, or patient cohort, only one was counted and selected.

2.6. Data Synthesis and Analysis

Data analysis was conducted using Comprehensive Meta-analysis Software (CMA) (Version 2.0) (https://www.meta-analysis.com/; accessed on 25 July 2021). The pooled prevalence of SARS-CoV-2 variants was calculated and subgroup analysis was done according to country. A random-effect model using the DerSimonian-Laird method of the meta-analysis was employed to determine the pooled estimates of the reported SARS-CoV-2 variants and subtype proportions. A forest plot was subsequently generated to visually summarize details of the individual studies alongside the estimated common effect and degree of heterogeneity. Publication bias was examined using funnel plots (visual aid for detecting bias) and Egger's regression test. Cochran's Q test evaluated the heterogeneities (i.e., variation in study outcomes between studies) of study-level estimates and quantified using I^2 statistics. I^2 values of 25%, 50%, and 75% were considered low, moderate, and high heterogeneity, respectively [12].

Subgroup meta-analysis was used to analyze sources of heterogeneity. A sensitivity test was conducted using the leave-one-out analysis. p-value of <0.001 was considered to be statistically significant for all tests.

3. Results

3.1. Search Result and Eligible Studies

The complete literature search process is displayed in Figure 1. The search strategy initially found 352 articles, after which 325 were left after duplicates removal. Two hundred and fifty-three articles were excluded based on the exclusion criteria. The full-text of 72 articles were assessed for eligibility, and ten were excluded for lack of mutations data or mutations data were not countries-specified. A total of 62 articles were included in the final qualitative synthesis, and finally, 51 articles published between December 2019 and October 2020 were included in the final quantitative synthesis (meta-analysis).

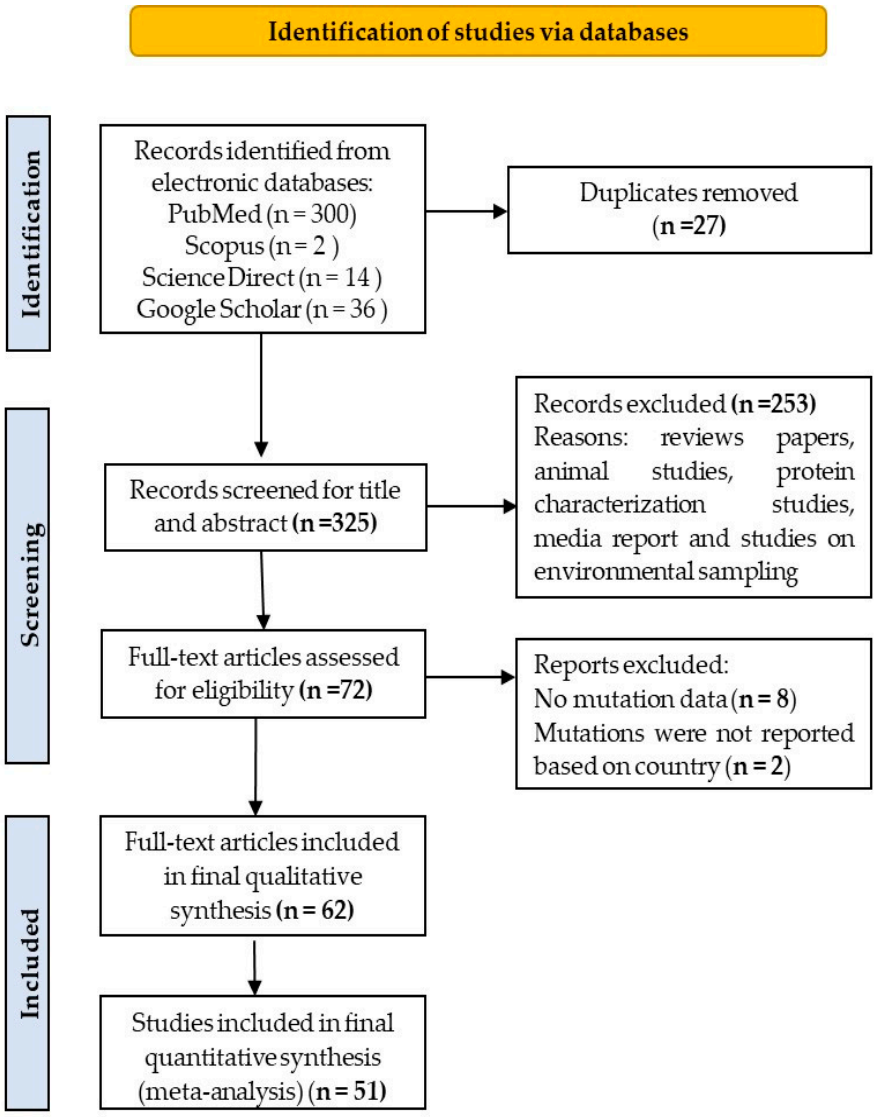

Figure 1. PRISMA flow diagram for the selection og eligible articles included in the study.

3.2. Characteristics of the Eligible Studies

All the eligible studies included in the meta-analyses were of high methodological quality. From 62 studies included from December 2019 to October 2020 (Table 1) [2,13–73], the highest numbers were from Worldwide (n = 10), multiple countries (n = 6) and the USA (n = 5). The 368,316 samples and genomic data analyzed in the studies were detected by quantitative Reverse Transcriptase (qRT-PCR) or DNA sequencing (Sanger, Next-generation, Whole-genome or Nanopore sequencing).

Genomic data from the studies included covered all regions of SAR-CoV-2 (Figure 2). From the mutational data analysed, the studies detected mutations in the Spike (S) protein (n = 50), Nucleocapsid (N) phosphoprotein (n = 34), Open Reading Frame (ORF) 1ab gene (n = 29), 5′-Untranslated region (UTR) (n = 28), ORF3a (n = 25), Membrane (M) glycoprotein (n = 19), ORF7 (n = 10), ORF6 (n = 8), ORF8 (n = 8), ORF10 (n = 8), Envelope (E) protein (n = 7), 3′ UTR (n = 5) and ORF14 (n = 1). The synonymous and missense mutations detected, mostly based on countries and region of mutations, are listed in Table S1.

Table 1. Major characteristics of the included studies.

No	Study ID (Ref)	Country of Study	Period of Study	No. of Participant	No. of Mutated Cases	Mutation Detection Method	Regions of Mutation
1	Akter et al., 2020 [13]	Bangladesh	May–June 2020	3	3	Whole-genome sequencing	ORF1ab, N and S gene
2	Andrés et al., 2020 [14]	Spain	March 2020	18	18	Deep sequencing of S gene	S gene
3	Badua et al., 2020 [15]	Multiple countries	January–May 2020	151	151	NGS	ORF1ab, ORF8, ORF3a, 5′UTR, 3′UTR, ORF6, ORF7a, ORF10, S, E, M and N gene.
4	Barret et al., 2020 [16]	USA	December 2019–May 2020	119	119	NGS	5′UTR, ORF1ab, S gene
5	Bartolini et al., 2020 [17]	Italy	February–March 2020	9	9	NGS (SARS-CoV-2 panel)	ORF1ab, UTR, S, N and M gene,
6	Becerra-Flores 2020 [18]	Worldwide	March–April 2020	NR	NR	NGS	S gene
7	Benvenuto et al., 2020 [19]	Italy	January–April 2020	79	79	NGS	S and N gene
8	Chang et al., 2020 [20]	Multiple countries	NR	10	10	NGS	ORF1ab, ORF8, S and E gene
9	Chen et al., 2020 [21]	China	January–February 2020	10	10	qRT-PCR on ORF1ab and N gene; RNA sequencing	ORF1ab, ORF3a, ORF8, ORF10, S and N gene
10	Cusi et al., 2020 [22]	Italy	March 2020	1	1	Direct RNA and amplicon sequencing	S gene
11	Demİr et al., 2020 [23]	Turkey	March–May 2020	63	63	NGS	ORF1ab, ORF3a, 3′UTR, 5′UTR, S, N and M gene
12	Devendran et al., 2021 [24]	India	as of April 2020	10	10	NGS/WGS	ORF1ab, ORF8, S and N gene
13	Du et al., 2020 [25]	China	January–April 2020	102	102	qRT-PCR, meta-transcriptomic sequencing	5′UTR, ORF1ab, S, ORF3a, ORF8, N gene
14	Elizondo et al., 2020 [26]	Uruguay	March–May 2020	44	44	qRT-PCR, NGS	ORF8, ORF3a, ORF1ab
15	Eskier et al., 2020 [27]	USA and UK	January–March 2020	11,701	11,701	NGS	Whole genome
16	Gómez-Carballa et al., 2020 [28]	Spain	as of June 2020	922	922	NGS	Whole genome
17	Gong et al., 2020 [29]	Taiwan	January–March 2020	20	19	RT-PCR & WGS	ORF1ab, ORF8, ORF3a, S gene, N gene
18	Gupta 2020 [30]	Worldwide	January–April 2020	87	87	NGS	ORF1ab, ORF3a, ORF7a, ORF8, N, S and M gene
19	Hartley et al., 2021 [31]	USA	March–June 2020	200	173	NGS	ORF1ab, S gene
20	Hassan et al., 2020 [32]	India	as of May 2020	128	128	NGS	ORF1, ORF3a, ORF8, ORF7a, S, M and N gene
21	Yang et al., 2020 [33]	Worldwide	December 2019–June 2020	46,414	46,414	NGS	Whole genome
22	Ip et al., 2020 [34]	Hong Kong	January–March 2020	12	1	Sanger sequencing, Nanopore and Illumina sequencing	S gene
23	Islam et al., 2020 [35]	Multiple countries	as of May 2020	444	404	NGS	ORF1ab, N, E, M, S
24	Jacob et al., 2020 [36]	India	until June 2020	>600	NR	NGS/WGS	S gene
25	Jary et al., 2021 [37]	France	January–February 2020	1	1	NGS	ORF3a, ORF7a, ORF6, ORF7b, ORF8, ORF10, N, M and E gene
26	Jenjaroenpun et al., 2021 [38]	USA	July 2020	2	2	Oxford Nanopore Technologies (ONT) MinION sequencing technology	ORF1ab, ORF3a, ORF14 and S gene
27	Khailany et al., 2020 [39]	Worldwide	December 2019–April 2020	95	71	NGS/WGS	ORF1ab, ORF8, ORF3a, ORF10, S, N and M gene
28	Kim et al., 2020 [40]	Worldwide	NR	178	178	NGS	ORF1ab, ORF3, ORF6, ORF7a, ORF7b, ORF8, ORF10 S, M, E and N gene
29	Kim et al., 2020 [41]	Korea	NR	4	4	qRT-PCR and Sanger sequencing	S gene
30	Koyama et al., 2020 [42]	Worldwide	February–May 2020	15,755	10,022	NGS	Whole genome
31	Kozlovskaya et al., 2020 [43]	Russia	March–April 2020	220	220	NGS	ORF1ab, S and N gene
32	Kumar et al., 2020 [44]	Multiple countries	December 2019–March 2020	95	95	NGS/WGS	Whole genome

Table 1. *Cont.*

No	Study ID (Ref)	Country of Study	Period of Study	No. of Participant	No. of Mutated Cases	Mutation Detection Method	Regions of Mutation
33	Laamarti et al., 2020 [45]	Morocco	NR	6	6	Oxford NanoporeTechnologies [ONT]	ORF1ab, S gene, 5′UTR
34	Leung et al., 2021 [46]	Hong Kong	as of February 2020	50	50	Nanopore and NGS	ORF3a, ORF1ab, S gene
35	Ling et al., 2020 [47]	Sweden	February–May 2020	348	348	NGS	5′-UTR, ORF1ab, S, ORF3a, M and N gene
36	McNamara et al., 2020 [48]	USA	March–May 2020	175	175	NGS	S and 3′UTR
37	Micheli et al., 2020 [49]	Italy	February–April 2020	20	20	NGS	M and N gene
38	Nagy et al. 2021 [50]	Worldwide	December 2019–September 2020	149,061	149,061	NGS	ORF1ab, ORF3a, ORF8, ORF6, N and S gene
39	Pachetti et al., 2020 [51]	Worldwide	December 2019–March 2020	220	215	NGS/WGS	Whole genome
40	Parvez et al., 2021 [52]	Bangladesh	as of August 2020	311	311	NGS/WGS	ORF1a, S and N gene
41	Raghav et al., 2020 [53]	India	March–June 2020	202	202	NGS	ORF1ab, 5′-UTR, ORF3a, ORF6, ORF7b, ORF84, ORF10, M, N and S
42	Rito et al., 2020 [54]	Worldwide	May–20	26,869	20,163	NGS/WGS	Whole genome
43	Saha et al., 2020 [55]	India	NR	566	566	NGS	5′UTR, ORF1ab, ORF3a, S, M and N
44	Saha et al., 2020 [56]	Bangladesh	April–July 2020	41	41	NGS	ORF1ab, ORF3a, ORF6, ORF7a, ORF8, Matrix (M gene), S and N gene
45	San et al., 2021 [57]	South Africa	March–June 2020	109	109	NGS	ORF1ab, ORF3a, ORF6, ORF7a, ORF7b, ORF8, ORF10, S, E, M and N gene
46	Skums et al., 2020 [58]	Worldwide	NR	319	274	NGS/WGS	Whole genome
47	Soliman et al., 2021 [59]	Egypt	June 2020	1	1	NGS	ORF1ab and S gene
48	Soratto et al., 2020 [2]	Sweden	April 2020	4	4	NGS	ORF1ab, ORF3a, ORF7a, S and N gene
49	Sun et al., 2020 [60]	China	NR	1	1	RT-PCR	E gene
50	Surleac et al., 2020 [61]	Romania	January–February 2020	25	25	NGS	ORF1ab, S and N gene
51	Taboada et al., 2020 [62]	Mexico	February–March 2020	17	17	NGS	ORF1ab, ORF8 and S gene
52	Toyoshima et al., 2020 [63]	Multiple countries	As of May 2020	12,343	12,343	NGS	ORF1ab, ORF3a, ORF8, S, N and M gene
53	Velasco et al., 2020 [64]	The Phillipines	April–July 2020	23	23	NGS	ORF1ab, ORF6, ORF7a, OORF7b, ORF8, ORF10, S, N and M gene
54	Volz et al., 2021 [65]	UK	January–June 2020	26,986	21,231	NGS	S gene
55	Wang et al., 2020 [66]	USA	July 2020	24,715	24,715	NGS	ORF1ab, ORF3a, ORF8 and S gene
56	Wang et al., 2020 [67]	Multiple countries	as of October 2020	75,775	75,775	NGS	ORF1ab
57	Wang et al., 2020 [68]	Worldwide	as of June 2020	15,140	15,140	NGS	Whole genome
58	Yap et al., 2020 [69]	Multiple countries	January–April 2020	142	112	NGS	ORF1ab, ORF8, S and N gene
59	Yuan et al., 2020 [70]	Worldwide	January–May 2020	11,183	11,183	NGS	Whole genome
60	Zhang et al., 2020 [71]	China	June–July 2020	6	6	NGS	ORF1ab gene, S and N gene
61	Ziegler et al., 2020 [72]	Germany	July 2020	1	1	qRT-PCR, PCR & Sanger sequencing	N gene
62	Zuckerman et al., 2020 [73]	Isreal	March 2020	8	8	qRT-PCR, NGS	5-UTR, ORF1ab, S, ORF3a and N gene

NGS: Next Generation Sequencing; WGS: Whole Genome Sequencing; RT-PCR: Reverse transcriptase PCR; qRT-PCR: Quantitative Reverse Transcription PCR; ORF: Open Reading Frame; S gene: Spike gene; N gene: Nucleocapsid phosphoprotein gene; M gene: Membrane glycoprotein gene; E gene: Envelope gene. (Additional information regarding the reported mutations and their types is provided in Supplementary Table S1). Sequences downloaded from databases such as NCBI and GSAID are presumed to be detected by NGS/WGS unless authors specify. ORF1ab includes Nsp1-14, RdRp, ORF1a, ORF1b, helicase, 3′ to 5′ exonuclease, endoRNAse, 2′-O-ribose methyltransferase.

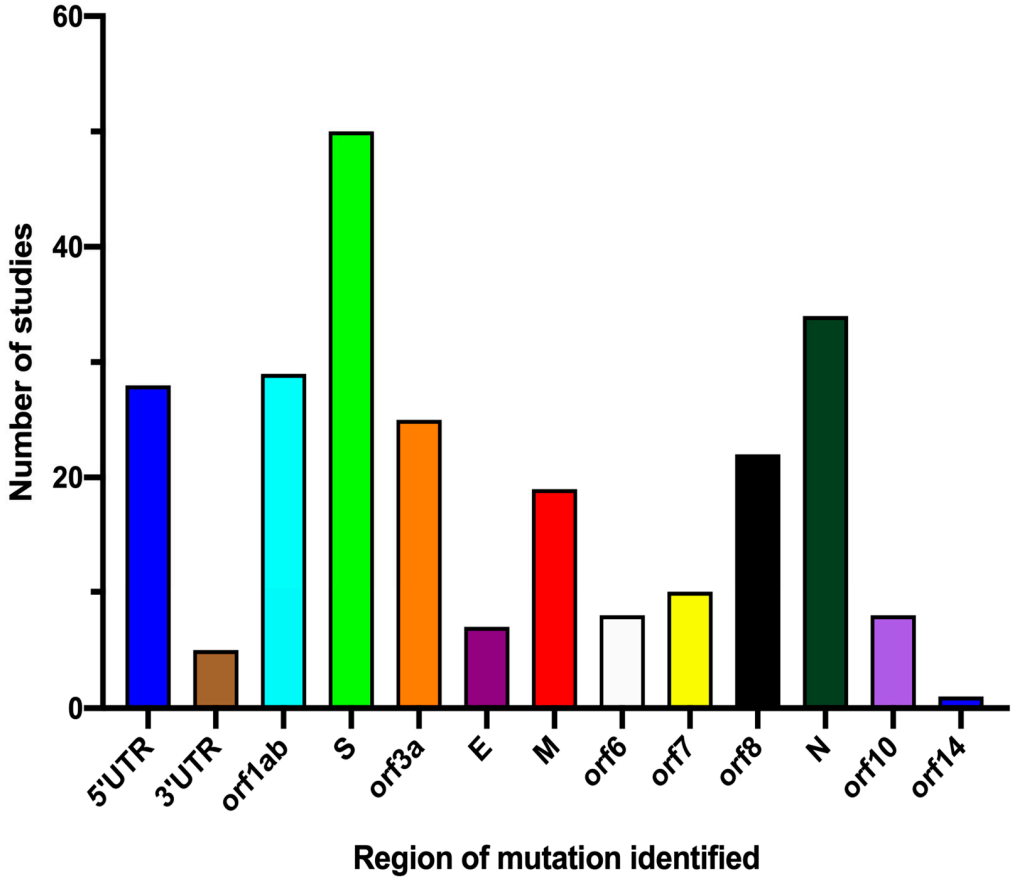

Figure 2. Reported regions of SARS-CoV-2 mutation. Data presented is based on the identification of mutation in any of the highlighted genomic regions from the included studies that presented data on region of mutation (n = 62). Some studies reported more than one region.

3.3. The Pooled Prevalence of SARS-CoV-2 Variants

The pooled prevalence of SARS-CoV-2 variants was estimated at 95.1% (95% CI; 93.3–96.4%; I^2 = 98.952%; p = 0.000) (Figure 3). Random-effects meta-analyses were carried out. Between-study variability was high (t^2 = 0.515; heterogeneity I^2 = 98.952% with heterogeneity chi-square (Q) = 4772.621, degrees of freedom (df) = 50, and p = 0.000). Moreover, publication bias was observed, as shown in the asymmetrical funnel plot (Figure 4). Using the Trim and Fill method and because the random-effects model was utilized, 22 missing studies were imputed to the left side of the mean effect (Figure 5), resulting in a point estimate of 82.5% (95% CI; 77.6–86.4). In addition to the funnel plots, Egger's test was used to confirm the extent of bias (t-value = 1.447; p = 0.07717).

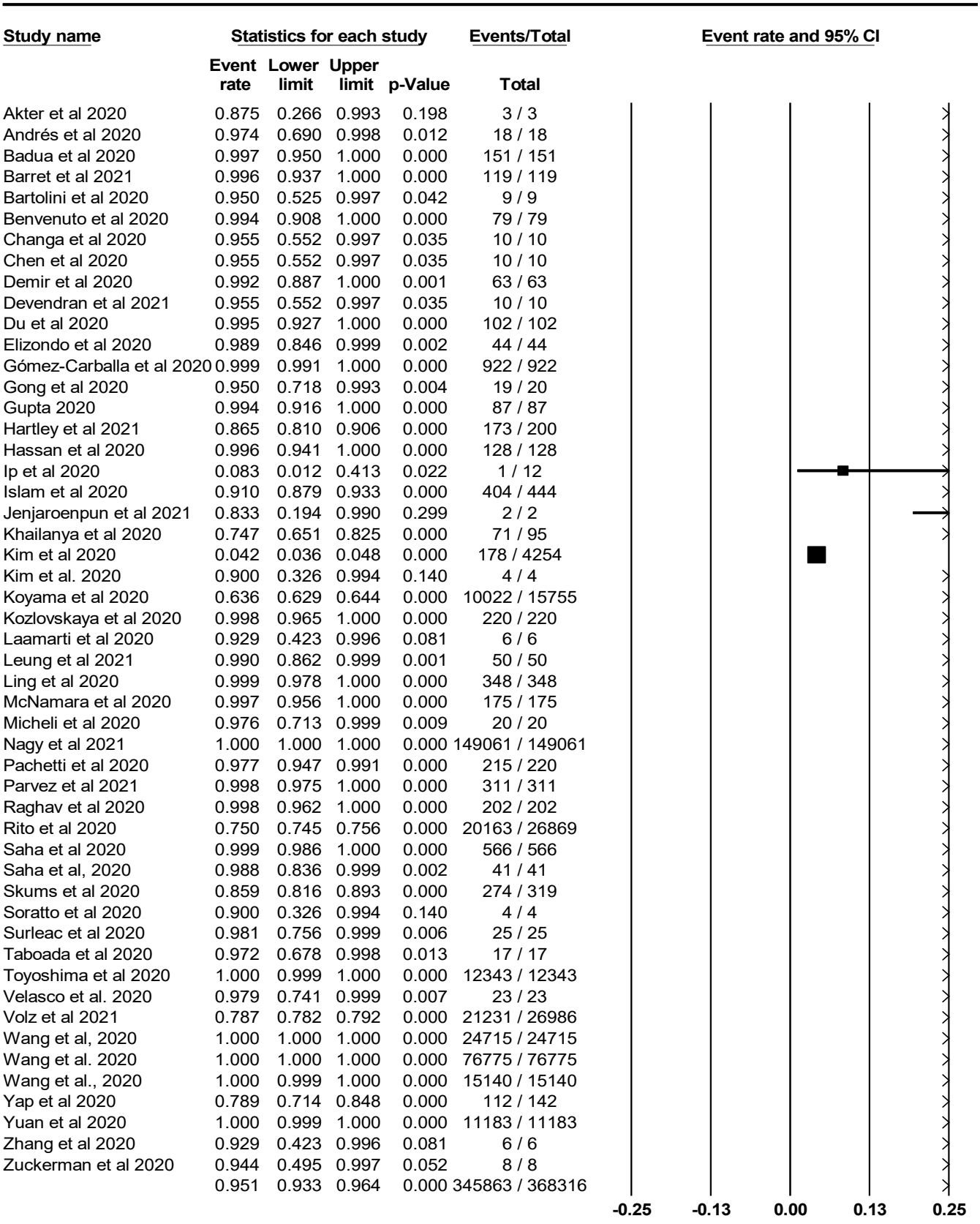

Figure 3. Forest plot showing the pooled prevalence of SARS-CoV-2 variants.

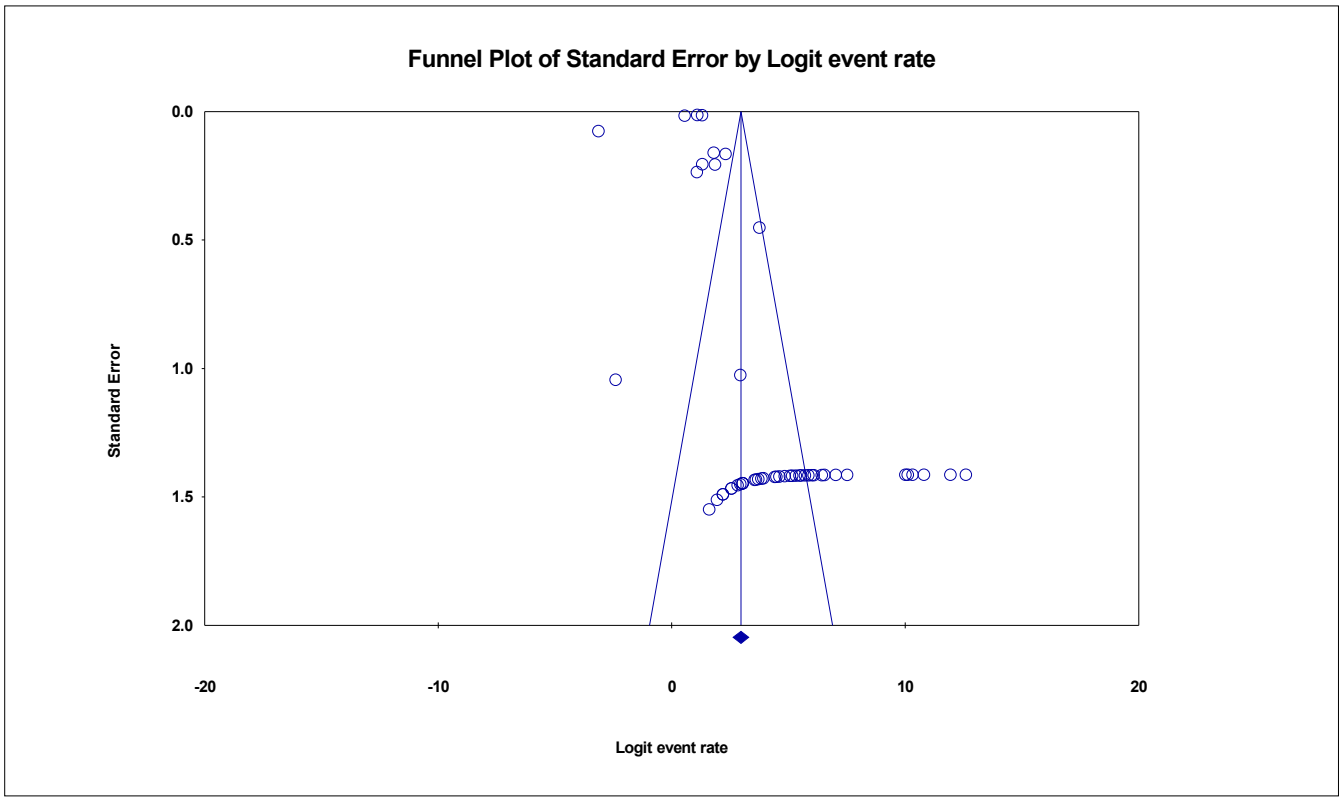

Figure 4. Funnel plot showing publication bias in studies reporting the prevalence of SARS-CoV-2 variants.

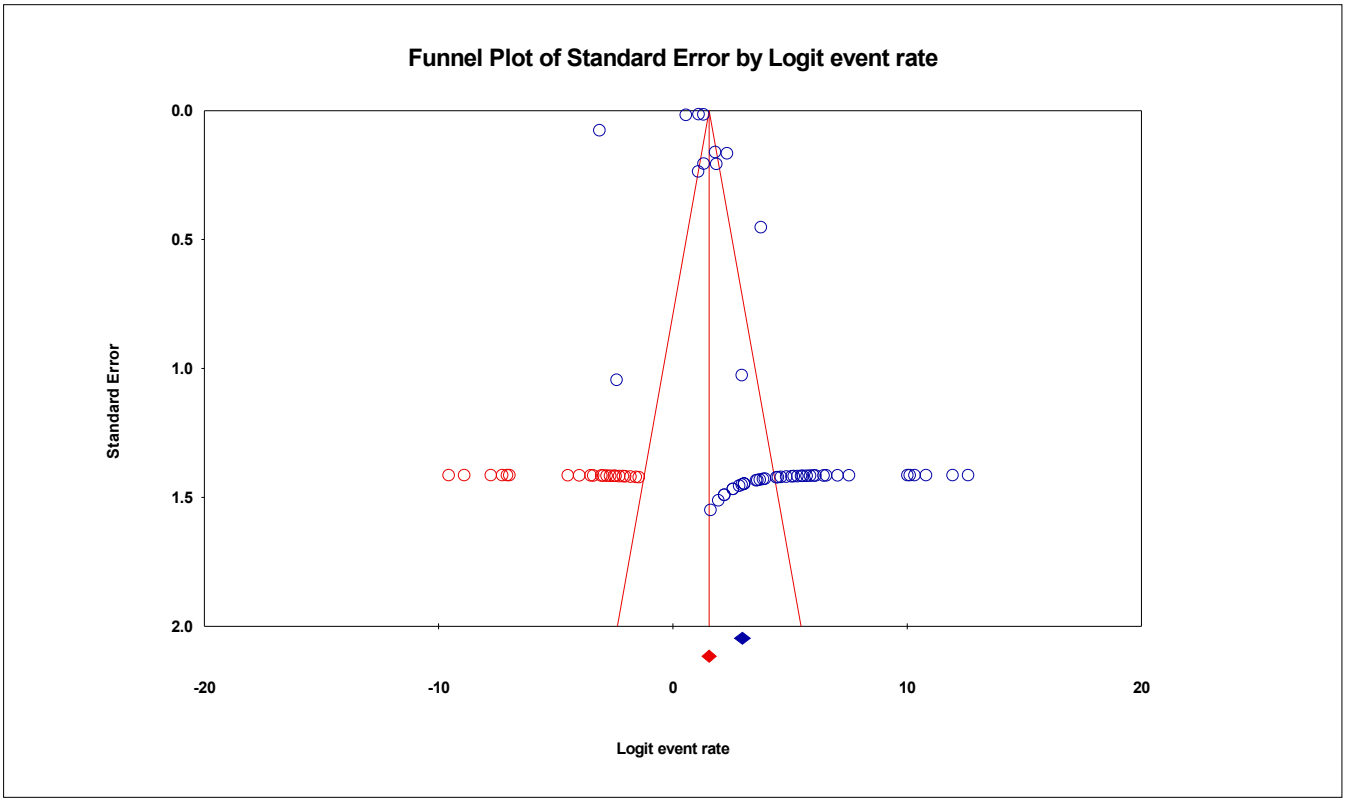

Figure 5. Funnel plot of the prevalence of SARS-CoV-2 variants showing 22 added studies (in Red) in the Trim-and- Fill method.

3.4. Subgroup Meta-Analysis

The result of subgroup meta-analysis by country showed that the majority of the studies were conducted Worldwide (n = 10), followed by studies carried out in Multiple countries (n = 6) and the USA (n = 5). Interestingly, China with three studies had heterogeneity (I^2) of 5.356 and prevalence of 97.5% (CI = 85.1–99.6%), while Italy, with the same number of studies, had heterogeneity of 0.000 and prevalence of 98.1% (CI = 88.2–99.7%). Heterogeneity was highest among studies conducted Worldwide (I^2= 99.747%), which was also trailed by six studies conducted in Multiple countries (I^2= 95.168%) (Table 2). The forest plot is shown in Figure 6.

Table 2. Subgroup analysis for comparison of SARS-CoV-2 variants across the country.

Country of Study	Number of Studies	Prevalence (%)	95% CI	I^2 (%)	Q	Heterogeneity Test DF	p
Bangladesh	3	98.7	91.9–99.8	57.445	4.700	2	0.095
China	3	97.5	85.1–99.6	5.356	2.113	2	0.348
Hong Kong	2	60.3	15.9–92.5	93.675	15.811	1	0.000
India	4	99.6	97.8–99.9	28.063	4.170	3	0.244
Israel	1	94.4	37.4–99.8	-	-	-	1.000
Italy	3	98.1	88.2–99.7	0.000	1.129	2	0.569
Korea	1	90.0	22.9–99.6	-	-	-	1.000
Mexico	1	97.2	56.0–99.9	-	-	-	1.000
Morocco	1	92.9	30.9–99.7	-	-	-	1.000
Multiple countries	6	98.2	95.3–99.3	95.168	103.483	5	0.000
Romania	1	98.1	65.2–99.9	-	-	-	1.000
Russia	1	99.8	94.3–100.0	-	-	-	1.000
Spain	2	99.6	96.3–100.0	73.466	3.769	1	0.052
Sweden	2	98.8	89.0–99.9	77.667	4.478	1	0.034
Taiwan	1	95.0	56.8–99.6	-	-	-	1.000
The Philippines	1	97.9	63.3–99.9	-	-	-	1.000
Turkey	1	99.2	82.5–100.0	-	-	-	1.000
UK	1	78.7	78.2–79.2	-	-	-	1.000
Uruguay	1	98.9	76.7–100.0	-	-	-	1.000
USA	5	98.4	94.8–99.5	92.301	51.954	4	0.000
Worldwide	10	90.3	82.6–94.8	99.747	3553.894	9	0.000
Total	51	97.4	94.4–98.8	98.952	4772.621	50	0.000

3.5. Meta-Regression

Meta-regression was done for the single variable country. Method of moments was used as the computational option, and a scattered plot (Figure 7) was plotted. p-value of '0.000' was obtained for 'Country', indicating the heterogeneity observed in this study, aside from chance, could also be contributed by country.

Group by Subgroup within study	Study name	Country	Event rate	Lower limit	Upper limit	p-Value
Bangladesh	Akter et al 2020	Bangladesh	0.875	0.266	0.993	0.198
Bangladesh	Parvez et al 2021	Bangladesh	0.998	0.975	1.000	0.000
Bangladesh	Saha et al, 2020	Bangladesh	0.988	0.836	0.999	0.002
Bangladesh			0.987	0.918	0.998	0.000
China	Chen et al 2020	China	0.955	0.552	0.997	0.035
China	Du et al 2020	China	0.995	0.927	1.000	0.000
China	Zhang et al 2020	China	0.929	0.423	0.996	0.081
China			0.975	0.851	0.996	0.000
Hong Kong	Ip et al 2020	Hong Kong	0.083	0.012	0.413	0.022
Hong Kong	Leung et al 2021	Hong Kong	0.990	0.862	0.999	0.001
Hong Kong			0.603	0.159	0.925	0.694
India	Devendran et al 2021	India	0.955	0.552	0.997	0.035
India	Hassan et al 2020	India	0.996	0.941	1.000	0.000
India	Raghav et al 2020	India	0.998	0.962	1.000	0.000
India	Saha et al 2020	India	0.999	0.986	1.000	0.000
India			0.996	0.978	0.999	0.000
Isreal	Zuckerman et al 2020	Isreal	0.944	0.495	0.997	0.052
Isreal			0.944	0.373	0.998	0.097
Italy	Bartolini et al 2020	Italy	0.950	0.525	0.997	0.042
Italy	Benvenuto et al 2020	Italy	0.994	0.908	1.000	0.000
Italy	Micheli et al 2020	Italy	0.976	0.713	0.999	0.009
Italy			0.981	0.881	0.997	0.000
Korea	Kim et al. 2020	Korea	0.900	0.326	0.994	0.140
Korea			0.900	0.229	0.996	0.207
Mexico	Taboada et al 2020	Mexico	0.972	0.678	0.998	0.013
Mexico			0.972	0.560	0.999	0.036
Morocco	Laamarti et al 2020	Morocco	0.929	0.423	0.996	0.081
Morocco			0.929	0.309	0.997	0.136
Multiple countries	Badua et al 2020	Multiple countries	0.997	0.950	1.000	0.000
Multiple countries	Changa et al 2020	Multiple countries	0.955	0.552	0.997	0.035
Multiple countries	Islam et al 2020	Multiple countries	0.910	0.879	0.933	0.000
Multiple countries	Toyoshima et al 2020	Multiple countries	1.000	0.999	1.000	0.000
Multiple countries	Wang et al. 2020	Multiple countries	1.000	1.000	1.000	0.000
Multiple countries	Yap et al 2020	Multiple countries	0.789	0.714	0.848	0.000
Multiple countries			0.982	0.953	0.994	0.000
Romania	Surleac et al 2020	Romania	0.981	0.756	0.999	0.006
Romania			0.981	0.652	0.999	0.020
Russia	Kozlovskaya et al 2020	Russia	0.998	0.965	1.000	0.000
Russia			0.998	0.943	1.000	0.000
Spain	Andrés et al 2020	Spain	0.974	0.690	0.998	0.012
Spain	Gómez-Carballa et al 2020	Spain	0.999	0.991	1.000	0.000
Spain			0.996	0.963	1.000	0.000
Sweden	Ling et al 2020	Sweden	0.999	0.978	1.000	0.000
Sweden	Soratto et al 2020	Sweden	0.900	0.326	0.994	0.140
Sweden			0.988	0.890	0.999	0.000
Taiwan	Gong et al 2020	Taiwan	0.950	0.718	0.993	0.004
Taiwan			0.950	0.568	0.996	0.031
The Phillipines	Velasco et al. 2020	The Phillipines	0.979	0.741	0.999	0.007
The Phillipines			0.979	0.633	0.999	0.023
Turkey	Demir et al 2020	Turkey	0.992	0.887	1.000	0.001
Turkey			0.992	0.825	1.000	0.004
UK	Volz et al 2021	UK	0.787	0.782	0.792	0.000
UK			0.787	0.389	0.955	0.146
Uruguay	Elizondo et al 2020	Uruguay	0.989	0.846	0.999	0.002
Uruguay			0.989	0.767	1.000	0.008
USA	Barret et al 2021	USA	0.996	0.937	1.000	0.000
USA	Hartley et al 2021	USA	0.865	0.810	0.906	0.000
USA	Jenjaroenpun et al 2021	USA	0.833	0.194	0.990	0.299
USA	McNamara et al 2020	USA	0.997	0.956	1.000	0.000
USA	Wang et al, 2020	USA	1.000	1.000	1.000	0.000
USA			0.984	0.948	0.995	0.000
Worldwide	Gupta 2020	Worldwide	0.994	0.916	1.000	0.000
Worldwide	Khailanya et al 2020	Worldwide	0.747	0.651	0.825	0.000
Worldwide	Kim et al 2020	Worldwide	0.042	0.036	0.048	0.000
Worldwide	Koyama et al 2020	Worldwide	0.636	0.629	0.644	0.000
Worldwide	Nagy et al 2021	Worldwide	1.000	1.000	1.000	0.000
Worldwide	Pachetti et al 2020	Worldwide	0.977	0.947	0.991	0.000
Worldwide	Rito et al 2020	Worldwide	0.750	0.745	0.756	0.000
Worldwide	Skums et al 2020	Worldwide	0.859	0.816	0.893	0.000
Worldwide	Wang et al., 2020	Worldwide	1.000	0.999	1.000	0.000
Worldwide	Yuan et al 2020	Worldwide	1.000	0.999	1.000	0.000
Worldwide			0.903	0.826	0.948	0.000
Overall			0.974	0.944	0.988	0.000

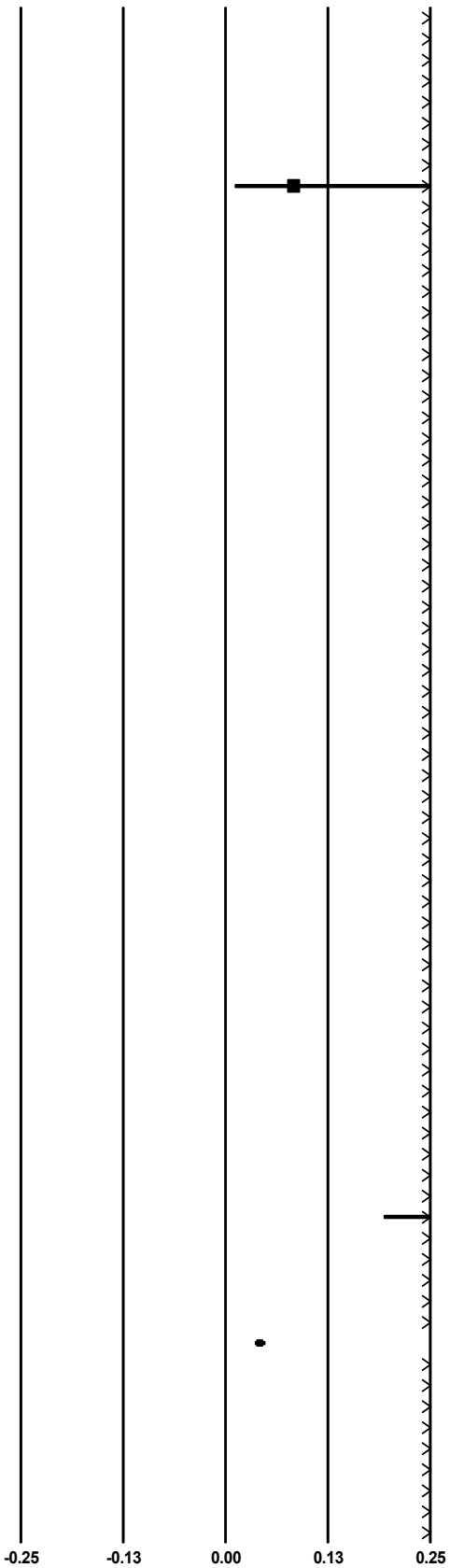

Figure 6. Forest plot showing the subgroup meta-analysis by country.

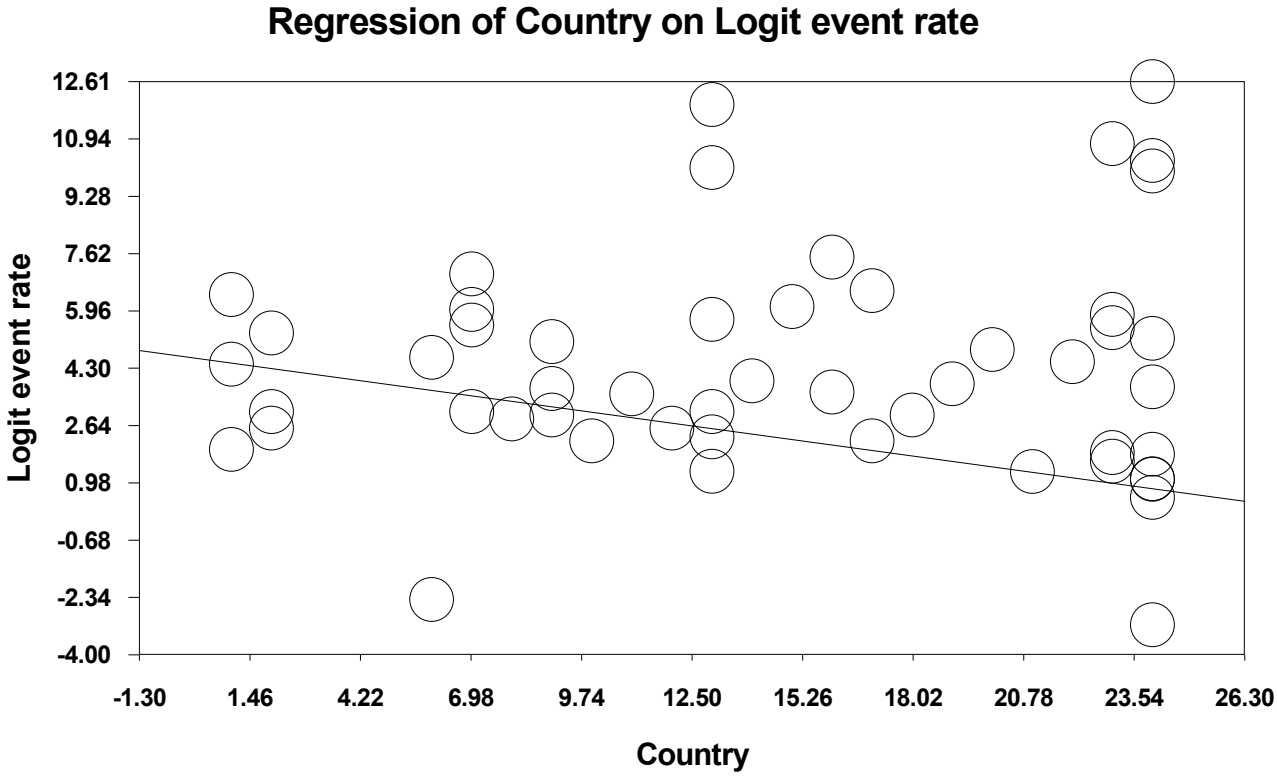

Figure 7. A scattered plot of Country Meta-regression.

4. Discussion

With the high infection numbers worldwide, the SARS-CoV-2 virus has evolved, developed mutations and given rise to new genetic variations with increased infectivity and transmissibility. Efforts are currently being undertaken to characterize the virus and its genomic variability molecularly. Viral mutations and variants around the globe are routinely monitored through sequence-based surveillance, epidemiological analysis and laboratory studies.

This study has examined the mutational profile of SARS-CoV-2 between December 2019 to October 2020 from 62 studies of different continents. The pooled prevalence of SARS-CoV-2 variants in COVID-19 patients' samples estimated by the random-effect model was 95.1%. Upon using the Trim and Fill method to adjust for potential bias, the estimate for the prevalence of the variants was still very high at 82.5%.

The analysis showed that between-study variability was high (I^2 = 98.95%). The subgroup meta-analysis showed that the high heterogeneity was contributed by countries such 'Worldwide' (I^2 = 99.7%), 'Multiple Countries' (I^2 = 95.2%), Hong Kong (I^2 = 93.7%) and USA (I^2 = 92.3%). Only two countries, Italy (I^2 = 0%) and China (I^2 = 5.4%) showed a low heterogeneity score. The different methods used to detect the mutations may contribute to the high heterogeneity, especially in the 'Worldwide' and 'Multiple Countries'. The high heterogeneity could also be attributed to the different regions of the SARS-CoV-2 gene analyzed (Spike protein, ORF1ab, Nucleocapsid polyprotein, ect.) and the type of samples used in the studies. Most of the studies, especially those referred to as 'Worldwide' and 'Multiple Countries', analyzed patients' genomic data downloaded from GISAID's database.

Most of the reported mutations were located at the Spike gene region, followed by the Nucleocapsid gene and ORF1ab gene. The high number of studies reporting on the Spike gene region might be due to its importance in the pathogenicity and transmissibility of the SARS-CoV-2 virus. The Spike (S) gene has two domains: S1 and S2. The S1 domain mediates receptor binding while S2 mediates downstream membrane fusion [74]. The S1

receptor-binding domain (RBD) shows a high affinity for the human ACE2 receptor in the lungs' alveolar type 2 (AT2) cells. Once the virus is attached to the host cell receptor, cleavage occurs between subunits S1 and S2. The subunit S2 will drive the viral and cellular membranes to fuse. The S1 recognizes and binds to the ACE2 receptor, whereas S2 directly facilitates entry into the host cell, making S1 and S2 crucial for infection [14].

Data extracted from publications included in this study showed that a 23403A>G mutation in the S gene, which produced a missense mutation of D614G in the Spike protein, was recorded in 43 out of 62 studies. The D614G substitution is usually linked to three other mutations: a 241C>to-T mutation in the 5′-UTR region, a synonymous 3037C>T mutation, and a non-synonymous 14408C>T mutation at the RNA-dependent RNA polymerase (RdRP) known as P323L or P4715L at ORF1ab gene.

Our data showed that D614G was detected in the European region from middle to late February 2020 [51]. By early March, it had spread rapidly to the United States (US) [51] and the South American region [62]. In east Asia, the D614G variant was found in Thailand from a sample diagnosed with COVID-19 in early March 2020 [69]. While in China, the variant was detected in samples collected from January to April 2020 [25]. By June 2020, D614G was found in every sample sequenced worldwide [75].

The mutation appeared to arise independently to simultaneously sweep across multiple geographic regions, suggestive of natural selection and an adaptive benefit of D614G. However, subsequent sequencing efforts identified the D614G mutation in viruses in several Chinese provinces in late January (first D614G in China: hCoV-19/Zhejiang/HZ103/2020; 24 January 2020), raising the possibility that global spreading of this mutation may result from chance founder events. Viruses carrying 614G mutation could initiate most early transmission events in multiple locations, demonstrating that D614G mutation was not adaptive, despite in vitro data showing its effects on receptor binding [76].

A study of more than 25,000 sequences of the UK population found that viruses bearing 614G mutation are associated with higher viral load and younger age of patients. It appeared to spread faster and seed larger phylogenetic clusters than viruses with 614D; however, no association was found between the presence of the Spike 614G with clinical severity and COVID-19 mortality [65].

In this study, few limitations were identified, including the inability to assess the impact of the identified mutations on patients' viral loads, severity of the disease, and its transmissibility, due to the lack of reported data from the included studies. An understanding of the impact of the mutations on these variables would be invaluable. Furthermore, most of the studies downloaded only viral genomic data extracted from COVID-19 patients from NCBI and GSAID websites, thus limiting our access to the patients' demographic information such as sex and age; and clinical data such as viral loads symptoms, co-morbidities and disease severity. The scarcity of the required data also limited the subgroup meta-analyses that could be conducted.

5. Conclusions

In this study, a systematic review and meta-analysis of studies were conducted to report the global prevalence of SARS-CoV-2 variants, estimated at 95.1%. Although a high heterogeneity was observed, we believe the estimate provides a good indication of the prevalence of SARS-CoV-2 variants worldwide from December 2019 to October 2020. With the fast evolution of the SARS-CoV-2 virus, there is a need to continuously monitor the prevalence of new mutations due to their potential influence on disease severity, transmissibility, resistance to antiviral drugs and vaccine effectiveness.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/life11111224/s1, Table S1: Major characteristics of the included studies; Table S2: Search strategy in four electronic databases; Table S3: Quality of included studies by JBI critical appraisal checklist for studies reporting prevalence data.

Author Contributions: Conceptualization and methodology, C.Y.Y., A.A.I., Y.W., E.N.S.E.A.R., W.Y.; Data extraction, synthesis and interpretation, W.Y., N.A., A.A.I. and Y.W.; Formal analysis, A.A.I. and Y.W.; Writing (original draft preparation), W.Y.; Writing (review and editing), W.Y., A.A.I., Y.W., E.N.S.E.A.R., N.A., N.M., M.F.K., Z.A.R., R.H., N.Y.Y. and C.Y.Y.; W.Y., A.A.I., Y.W., E.N.S.E.A.R., N.A., N.M., M.F.K., Z.A.R., R.H., N.Y.Y. and C.Y.Y. have read and agreed to the published version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Universiti Sains Malaysia (USM) Short Term Grant (grant number 304/PPSP/6315459).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used and/or analyzed during the current study are included in the manuscript.

Acknowledgments: A.A.I., Y.W., E.N.S.E.A.R. are supported by the USM Fellowship Scheme and N.A. is supported by USM's Graduate Research Assistant Scheme.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sahlan, M.; Irdiani, R.; Flamandita, D.; Aditama, R.; Alfarraj, S.; Ansari, M.J.; Khayrani, A.C.; Pratami, D.K.; Lischer, L. Molecular Interaction Analysis of Sulawesi Propolis Compounds with SARS-CoV-2 Main Protease as Preliminary Study for COVID-19 Drug Discovery. *J. King Saud. Univ. Sci.* **2021**, *33*, 101234. [CrossRef]
2. Soratto, T.A.T.; Darban, H.; Bjerkner, A.; Coorens, M.; Albert, J.; Allander, T.; Andersson, B. Four SARS-CoV-2 Genome Sequences from Late April in Stockholm, Sweden, Reveal a Rare Mutation in the Spike Protein. *Microbiol. Resour. Announc.* **2020**, *9*, e00934-20. [CrossRef]
3. Dinnes, J.; Deeks, J.J.; Berhane, S.; Taylor, M.; Adriano, A.; Davenport, C.; Dittrich, S.; Emperador, D.; Takwoingi, Y.; Cunningham, J.; et al. Rapid, Point-of-Care Antigen and Molecular-Based Tests for Diagnosis of SARS-CoV-2 Infection. *Cochrane Database Syst. Rev.* **2021**, *3*, CD013705. [PubMed]
4. Oughtred, R.; Rust, J.; Chang, C.; Breitkreutz, B.J.; Stark, C.; Willems, A.; Boucher, L.; Leung, G.; Kolas, N.; Zhang, F.; et al. The Biogrid Database: A Comprehensive Biomedical Resource of Curated Protein, Genetic, and Chemical Interactions. *Protein Sci.* **2021**, *30*, 187–200. [CrossRef] [PubMed]
5. Shakoor, H.; Feehan, J.; al Dhaheri, A.S.; Ali, H.I.; Platat, C.; Ismail, L.C.; Apostolopoulos, V.; Stojanovska, L. Immune-Boosting Role of Vitamins D, C, E, Zinc, Selenium and Omega-3 Fatty Acids: Could They Help against COVID-19? *Maturitas* **2021**, *143*, 1–9. [CrossRef] [PubMed]
6. Ali, Z.; Jatoi, M.A.; Al-Wraikat, M.; Ahmed, N.; Li, J. Time to Enhance Immunity Via Functional Foods and Supplements: Hope for SARS-CoV-2 Outbreak. *Altern. Ther. Health Med.* **2021**, *27* (Suppl. S1), 30–44. [PubMed]
7. Gelardi, M.; Notargiacomo, M.; Trecca, E.M.C.; Cassano, M.; Ciprandi, G. COVID-19 and Nasal Cytobrush Cytology. *Acta Cytol.* **2020**, *64*, 397–398. [CrossRef]
8. Samudrala, P.K.; Kumar, P.; Choudhary, K.; Thakur, N.; Wadekar, G.S.; Dayaramani, R.; Agrawal, M.; Alexander, A. Virology, Pathogenesis, Diagnosis and in-Line Treatment of COVID-19. *Eur. J. Pharmacol.* **2020**, *883*, 173375. [CrossRef]
9. Qiang, X.L.; Xu, P.; Fang, G.; Liu, W.B.; Kou, Z. Using the Spike Protein Feature to Predict Infection Risk and Monitor the Evolutionary Dynamic of Coronavirus. *Infect. Dis. Poverty* **2020**, *9*, 33. [CrossRef]
10. Shamseer, L.; Moher, D.; Clarke, M.; Ghersi, D.; Liberati, A.; Petticrew, M.; Shekelle, P.; Stewart, L.A. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (Prisma-P) 2015: Elaboration and Explanation. *BMJ Br. Med. J.* **2015**, *349*, g7647. [CrossRef]
11. Munn, Z.; Moola, S.; Lisy, K.; Riitano, D.; Tufanaru, C. Methodological Guidance for Systematic Reviews of Observational Epidemiological Studies Reporting Prevalence and Cumulative Incidence Data. *JBI Evid. Implement.* **2015**, *13*, 147–153. [CrossRef] [PubMed]
12. Higgins, J.P.; Thompson, S.G. Quantifying Heterogeneity in a Meta-Analysis. *Stat. Med.* **2002**, *21*, 1539–1558. [CrossRef] [PubMed]
13. Akter, S.; Banu, T.A.; Goswami, B.; Osman, E.; Uzzaman, M.S.; Habib, M.A.; Jahan, I.; Mahmud, A.S.M.; Sarker, M.M.H.; Hossain, M.S.; et al. Coding-Complete Genome Sequences of Three SARS-CoV-2 Strains from Bangladesh. *Microbiol. Resour. Announc.* **2020**, *9*, e00764-20. [CrossRef] [PubMed]
14. Andrés, C.; Garcia-Cehic, D.; Gregori, J.; Piñana, M.; Rodriguez-Frias, F.; Guerrero-Murillo, M.; Esperalba, J.; Rando, A.; Goterris, L.; Codina, M.G.; et al. Naturally Occurring SARS-CoV-2 Gene Deletions Close to the Spike S1/S2 Cleavage Site in the Viral Quasispecies of Covid19 Patients. *Emerg. Microbes Infect.* **2020**, *9*, 1900–1911. [CrossRef]
15. Badua, C.L.D.; Baldo, K.A.T.; Medina, P.M.B. Genomic and Proteomic Mutation Landscapes of SARS-CoV-2. *J. Med. Virol.* **2020**, *93*, 1702–1721. [CrossRef] [PubMed]

16. Barrett, C.; Bura, A.C.; He, Q.; Huang, F.W.; Li, T.J.X.; Waterman, M.S.; Reidys, C.M. Multiscale Feedback Loops in SARS-CoV-2 Viral Evolution. *J. Comput. Biol.* **2020**, *28*, 248–256. [CrossRef]
17. Bartolini, B.; Rueca, M.; Gruber, C.E.M.; Messina, F.; Carletti, F.; Giombini, E.; Lalle, E.; Bordi, L.; Matusali, G.; Colavita, F.; et al. SARS-CoV-2 Phylogenetic Analysis, Lazio Region, Italy, February–March 2020. *Emerg. Infect. Dis.* **2020**, *26*, 1842–1845. [CrossRef]
18. Becerra-Flores, M.; Cardozo, T. SARS-CoV-2 Viral Spike G614 Mutation Exhibits Higher Case Fatality Rate. *Int. J. Clin. Pract.* **2020**, *74*, e13525. [CrossRef]
19. Benvenuto, D.; Demir, A.B.; Giovanetti, M.; Bianchi, M.; Angeletti, S.; Pascarella, S.; Cauda, R.; Ciccozzi, M.; Cassone, A. Evidence for Mutations in SARS-CoV-2 Italian Isolates Potentially Affecting Virus Transmission. *J. Med. Virol.* **2020**, *92*, 2232–2237. [CrossRef]
20. Chang, T.J.; Yang, D.M.; Wang, M.L.; Liang, K.H.; Tsai, P.H.; Chiou, S.H.; Lin, T.H.; Wang, C.T. Genomic Analysis and Comparative Multiple Sequences of SARS-CoV-2. *J. Chin. Med. Assoc.* **2020**, *83*, 537–543. [CrossRef]
21. Chen, J.; Hilt, E.E.; Li, F.; Wu, H.; Jiang, Z.; Zhang, Q.; Wang, J.; Wang, Y.; Li, Z.; Tang, J.; et al. Epidemiological and Genomic Analysis of SARS-CoV-2 in 10 Patients from a Mid-Sized City Outside of Hubei, China in the Early Phase of the COVID-19 Outbreak. *Front. Public Health* **2020**, *8*, 567621. [CrossRef]
22. Cusi, M.G.; Pinzauti, D.; Gandolfo, C.; Anichini, G.; Pozzi, G.; Santoro, F. Whole-Genome Sequence of SARS-CoV-2 Isolate Siena-1/2020. *Microbiol. Resour. Announc.* **2020**, *9*, e00944-20. [CrossRef] [PubMed]
23. Demir, A.B.; Benvenuto, D.; Abacioğlu, H.; Angeletti, S.; Ciccozzi, M. Identification of the Nucleotide Substitutions in 62 SARS-CoV-2 Sequences from Turkey. *Turk. J. Biol.* **2020**, *44*, 178–184. [CrossRef] [PubMed]
24. Devendran, R.; Kumar, M.; Chakraborty, S. Genome Analysis of SARS-CoV-2 Isolates Occurring in India: Present Scenario. *Indian J. Public Health* **2020**, *64*, S147–S55. [PubMed]
25. Du, P.; Ding, N.; Li, J.; Zhang, F.; Wang, Q.; Chen, Z.; Song, C.; Han, K.; Xie, W.; Liu, J.; et al. Genomic Surveillance of COVID-19 Cases in Beijing. *Nat. Commun.* **2020**, *11*, 5503. [CrossRef] [PubMed]
26. Elizondo, V.; Harkins, G.W.; Mabvakure, B.; Smidt, S.; Zappile, P.; Marier, C.; Maurano, M.; Perez, V.; Mazza, N.; Beloso, C.; et al. SARS-CoV-2 Genomic Characterization and Clinical Manifestation of the COVID-19 Outbreak in Uruguay. *Emerg. Microbes Infect.* **2020**, *10*, 51–65. [CrossRef] [PubMed]
27. Eskier, D.; Suner, A.; Karakülah, G.; Oktay, Y. Mutation Density Changes in SARS-CoV-2 Are Related to the Pandemic Stage but to a Lesser Extent in the Dominant Strain with Mutations in Spike and Rdrp. *PeerJ* **2020**, *8*, e9703. [CrossRef] [PubMed]
28. Gómez-Carballa, A.; Bello, X.; Pardo-Seco, J.; del Molino, M.L.P.; Martinón-Torres, F.; Salas, A. Phylogeography of SARS-CoV-2 Pandemic in Spain: A Story of Multiple Introductions, Micro-Geographic Stratification, Founder Effects, and Super-Spreaders. *Zool. Res.* **2020**, *41*, 605–620. [CrossRef]
29. Gong, Y.N.; Tsao, K.C.; Hsiao, M.J.; Huang, C.G.; Huang, P.N.; Huang, P.W.; Lee, K.M.; Liu, Y.C.; Yang, S.L.; Kuo, R.L.; et al. SARS-CoV-2 Genomic Surveillance in Taiwan Revealed Novel Orf8-Deletion Mutant and Clade Possibly Associated with Infections in Middle East. *Emerg. Microbes Infect.* **2020**, *9*, 1457–1466. [CrossRef]
30. Gupta, A.M.; Chakrabarti, J.; Mandal, S. Non-Synonymous Mutations of SARS-CoV-2 Leads Epitope Loss and Segregates Its Variants. *Microbes Infect.* **2020**, *22*, 598–607. [CrossRef] [PubMed]
31. Hartley, P.; Tillett, R.L.; Xu, Y.; AuCoin, D.P.; Sevinsky, J.R.; Gorzalski, A.; Pandori, M.C.; Buttery, E.; Hansen, H.; Picker, M.; et al. Genomic Surveillance Revealed Prevalence of Unique SARS-CoV-2 Variants Bearing Mutation in the Rdrp Gene among Nevada Patients. *medRxiv* **2020**. [CrossRef]
32. Hassan, S.S.; Choudhury, P.P.; Roy, B.; Jana, S.S. Missense Mutations in SARS-CoV2 Genomes from Indian Patients. *Genomics* **2020**, *112*, 4622–4627. [CrossRef] [PubMed]
33. Yang, H.C.; Chen, C.H.; Wang, J.H.; Liao, H.C.; Yang, C.T.; Chen, C.W.; Lin, Y.C.; Kao, C.H.; Lu, M.J.; Liao, J.C. Analysis of Genomic Distributions of SARS-CoV-2 Reveals a Dominant Strain Type with Strong Allelic Associations. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 30679–30686. [CrossRef]
34. Ip, J.D.; Kok, K.H.; Chan, W.M.; Chu, A.W.; Wu, W.L.; Yip, C.C.; To, W.K.; Tsang, O.T.; Leung, W.S.; Chik, T.S.; et al. Intra-Host Non-Synonymous Diversity at a Neutralizing Antibody Epitope of SARS-CoV-2 Spike Protein N-Terminal Domain. *Clin. Microbiol. Infect.* **2020**, *27*, 1350-e1. [CrossRef] [PubMed]
35. Islam, M.R.; Hoque, M.N.; Rahman, M.S.; Alam, A.R.U.; Akther, M.; Puspo, J.A.; Akter, S.; Sultana, M.; Crandall, K.A.; Hossain, M.A. Genome-Wide Analysis of SARS-CoV-2 Virus Strains Circulating Worldwide Implicates heterogeneity. *Sci. Rep.* **2020**, *10*, 14004. [CrossRef] [PubMed]
36. Jacob, J.J.; Vasudevan, K.; Veeraraghavan, B.; Iyadurai, R.; Gunasekaran, K. Genomic Evolution of Severe Acute Respiratory Syndrome Coronavirus 2 in India and Vaccine Impact. *Indian J. Med. Microbiol.* **2020**, *38*, 210–212. [CrossRef] [PubMed]
37. Jary, A.; Leducq, V.; Malet, I.; Marot, S.; Klement-Frutos, E.; Teyssou, E.; Soulié, C.; Abdi, B.; Wirden, M.; Pourcher, V.; et al. Evolution of Viral Quasispecies During SARS-CoV-2 Infection. *Clin. Microbiol. Infect.* **2020**, *26*, 1560-e1. [CrossRef] [PubMed]
38. Jenjaroenpun, P.; Wanchai, V.; Ono-Moore, K.D.; Laudadio, J.; James, L.P.; Adams, S.H.; Prior, F.; Nookaew, I.; Ussery, D.W.; Wongsurawat, T. Two SARS-CoV-2 Genome Sequences of Isolates from Rural U.S. Patients Harboring the D614g Mutation, Obtained Using Nanopore Sequencing. *Microbiol. Resour. Announc.* **2020**, *10*, e01109-20. [CrossRef]
39. Khailany, R.A.; Safdar, M.; Ozaslan, M. Genomic Characterization of a Novel SARS-CoV-2. *Gene Rep.* **2020**, *19*, 100682. [CrossRef]
40. Kim, J.S.; Jang, J.H.; Kim, J.M.; Chung, Y.S.; Yoo, C.K.; Han, M.G. Genome-Wide Identification and Characterization of Point Mutations in the SARS-CoV-2 Genome. *Osong Public Health Res. Perspect.* **2020**, *11*, 101–111. [CrossRef]

41. Kim, S.; Lee, J.H.; Lee, S.; Shim, S.; Nguyen, T.T.; Hwang, J.; Kim, H.; Choi, Y.O.; Hong, J.; Bae, S.; et al. The Progression of Sars Coronavirus 2 (SARS-CoV2): Mutation in the Receptor Binding Domain of Spike Gene. *Immune Netw.* **2020**, *20*, e41. [CrossRef]
42. Koyama, T.; Platt, D.; Parida, L. Variant Analysis of SARS-CoV-2 Genomes. *Bull. World Health Organ.* **2020**, *98*, 495–504. [CrossRef]
43. Kozlovskaya, L.; Piniaeva, A.; Ignatyev, G.; Selivanov, A.; Shishova, A.; Kovpak, A.; Gordeychuk, I.; Ivin, Y.; Berestovskaya, A.; Prokhortchouk, E.; et al. Isolation and Phylogenetic Analysis of SARS-CoV-2 Variants Collected in Russia During the COVID-19 Outbreak. *Int. J. Infect. Dis.* **2020**, *99*, 40–46. [CrossRef] [PubMed]
44. Kumar, R.; Verma, H.; Singhvi, N.; Sood, U.; Gupta, V.; Singh, M.; Kumari, R.; Hira, P.; Nagar, S.; Talwar, C.; et al. Comparative Genomic Analysis of Rapidly Evolving SARS-CoV-2 Reveals Mosaic Pattern of Phylogeographical Distribution. *mSystems* **2020**, *5*, e00505-20. [CrossRef] [PubMed]
45. Laamarti, M.; Chemao-Elfihri, M.W.; Kartti, S.; Laamarti, R.; Allam, L.; Ouadghiri, M.; Smyej, I.; Rahoui, J.; Benrahma, H.; Diawara, I.; et al. Genome Sequences of Six SARS-CoV-2 Strains Isolated in Morocco, Obtained Using Oxford Nanopore Minion Technology. *Microbiol. Resour. Announc.* **2020**, *9*, e00767-20. [CrossRef] [PubMed]
46. Leung, K.S.; Ng, T.T.; Wu, A.K.; Yau, M.C.; Lao, H.Y.; Choi, M.P.; Tam, K.K.; Lee, L.K.; Wong, B.K.; Ho, A.Y.M.; et al. Territorywide Study of Early Coronavirus Disease Outbreak, Hong Kong, China. *Emerg. Infect. Dis.* **2021**, *27*, 196–204. [CrossRef]
47. Ling, J.; Hickman, R.A.; Li, J.; Lu, X.; Lindahl, J.F.; Lundkvist, Å.; Järhult, J.D. Spatio-Temporal Mutational Profile Appearances of Swedish SARS-CoV-2 During the Early Pandemic. *Viruses* **2020**, *12*, 1026. [CrossRef] [PubMed]
48. McNamara, R.P.; Caro-Vegas, C.; Landis, J.T.; Moorad, R.; Pluta, L.J.; Eason, A.B.; Thompson, C.; Bailey, A.; Villamor, F.C.S.; Lange, P.T.; et al. High-Density Amplicon Sequencing Identifies Community Spread and Ongoing Evolution of SARS-CoV-2 in the Southern United States. *Cell Rep.* **2020**, *33*, 108352. [CrossRef] [PubMed]
49. Micheli, V.; Rimoldi, S.G.; Romeri, F.; Comandatore, F.; Mancon, A.; Gigantiello, A.; Perini, M.; Mileto, D.; Pagani, C.; Lombardi, A.; et al. Geographical Reconstruction of the SARS-CoV-2 Outbreak in Lombardy (Italy) During the Early Phase. *J. Med. Virol.* **2020**, *93*, 1752–1757. [CrossRef]
50. Nagy, Á.; Pongor, S.; Győrffy, B. Different Mutations in SARS-CoV-2 Associate with Severe and Mild Outcome. *Int. J. Antimicrob. Agents* **2020**, *57*, 106272. [CrossRef]
51. Pachetti, M.; Marini, B.; Benedetti, F.; Giudici, F.; Mauro, E.; Storici, P.; Masciovecchio, C.; Angeletti, S.; Ciccozzi, M.; Gallo, R.C.; et al. Emerging SARS-CoV-2 Mutation Hot Spots Include a Novel Rna-Dependent-Rna Polymerase Variant. *J. Transl. Med.* **2020**, *18*, 179. [CrossRef]
52. Parvez, M.S.A.; Rahman, M.M.; Morshed, M.N.; Rahman, D.; Anwar, S.; Hosen, M.J. Genetic Analysis of SARS-CoV-2 Isolates Collected from Bangladesh: Insights into the Origin, Mutational Spectrum and Possible Pathomechanism. *Comput. Biol. Chem.* **2020**, *90*, 107413. [CrossRef] [PubMed]
53. Raghav, S.; Ghosh, A.; Turuk, J.; Kumar, S.; Jha, A.; Madhulika, S.; Priyadarshini, M.; Biswas, V.K.; Shyamli, P.S.; Singh, B.; et al. Analysis of Indian SARS-CoV-2 Genomes Reveals Prevalence of D614g Mutation in Spike Protein Predicting an Increase in Interaction with Tmprss2 and Virus Infectivity. *Front. Microbiol.* **2020**, *11*, 594928. [CrossRef]
54. Rito, T.; Richards, M.B.; Pala, M.; Correia-Neves, M.; Soares, P.A. Phylogeography of 27,000 SARS-CoV-2 Genomes: Europe as the Major Source of the COVID-19 Pandemic. *Microorganisms* **2020**, *8*, 1678. [CrossRef] [PubMed]
55. Saha, I.; Ghosh, N.; Maity, D.; Sharma, N.; Sarkar, J.P.; Mitra, K. Genome-Wide Analysis of Indian SARS-CoV-2 Genomes for the Identification of Genetic Mutation and Snp. *Infect. Genet. Evol.* **2020**, *85*, 104457. [CrossRef]
56. Saha, O.; Hossain, M.S.; Rahaman, M.M. Genomic Exploration Light on Multiple Origin with Potential Parsimony-Informative Sites of the Severe Acute Respiratory Syndrome Coronavirus 2 in Bangladesh. *Gene Rep.* **2020**, *21*, 100951. [CrossRef] [PubMed]
57. San, J.E.; Ngcapu, S.; Kanzi, A.M.; Tegally, H.; Fonseca, V.; Giandhari, J.; Wilkinson, E.; Nelson, C.W.; Smidt, W.; Kiran, A.M.; et al. Transmission Dynamics of SARS-CoV-2 within-Host Diversity in Two Major Hospital Outbreaks in South Africa. *Virus Evol.* **2021**, *7*, veab041. [CrossRef] [PubMed]
58. Skums, P.; Kirpich, A.; Baykal, P.I.; Zelikovsky, A.; Chowell, G. Global Transmission Network of SARS-CoV-2: From Outbreak to Pandemic. *medRxiv* **2020**. [CrossRef]
59. Soliman, M.S.; AbdelFattah, M.; Aman, S.M.N.; Ibrahim, L.M.; Aziz, R.K. A Gapless, Unambiguous Rna Metagenome-Assembled Genome Sequence of a Unique SARS-CoV-2 Variant Encoding Spike S813i and Orf1a A859v Substitutions. *Omics* **2020**, *25*, 123–128. [CrossRef] [PubMed]
60. Sun, Y.S.; Xu, F.; An, Q.; Chen, C.; Yang, Z.N.; Lu, H.J.; Chen, J.C.; Yao, P.P.; Jiang, J.M.; Zhu, H.P. A SARS-CoV-2 Variant with the 12-Bp Deletion at E Gene. *Emerg. Microbes Infect.* **2020**, *9*, 2361–2367. [CrossRef] [PubMed]
61. Surleac, M.; Banica, L.; Casangiu, C.; Cotic, M.; Florea, D.; Sandulescu, O.; Milu, P.; Streinu-Cercel, A.; Vlaicu, O.; Paraskevis, D.; et al. Molecular Epidemiology Analysis of SARS-CoV-2 Strains Circulating in Romania During the First Months of the Pandemic. *Life* **2020**, *10*, 152. [CrossRef]
62. Taboada, B.; Vazquez-Perez, J.A.; Muñoz-Medina, J.E.; Ramos-Cervantes, P.; Escalera-Zamudio, M.; Boukadida, C.; Sanchez-Flores, A.; Isa, P.; Mendieta-Condado, E.; Martínez-Orozco, J.A.; et al. Genomic Analysis of Early SARS-CoV-2 Variants Introduced in Mexico. *J. Virol.* **2020**, *94*, e01056-20. [CrossRef] [PubMed]
63. Toyoshima, Y.; Nemoto, K.; Matsumoto, S.; Nakamura, Y.; Kiyotani, K. SARS-CoV-2 Genomic Variations Associated with Mortality Rate of COVID-19. *J. Hum. Genet.* **2020**, *65*, 1075–1082. [CrossRef]

64. Velasco, J.M.; Chinnawirotpisan, P.; Joonlasak, K.; Manasatienkij, W.; Huang, A.; Valderama, M.T.; Diones, P.C.; Leonardia, S.; Timbol, M.L.; Navarro, F.C.; et al. Coding-Complete Genome Sequences of 23 SARS-CoV-2 Samples from the Philippines. *Microbiol. Resour. Announc.* **2020**, *9*, e01031-20. [CrossRef] [PubMed]
65. Volz, E.; Hill, V.; McCrone, J.T.; Price, A.; Jorgensen, D.; O'Toole, Á.; Southgate, J.; Johnson, R.; Jackson, B.; Nascimento, F.F.; et al. Evaluating the Effects of SARS-CoV-2 Spike Mutation D614g on Transmissibility and Pathogenicity. *Cell* **2021**, *184*, 64–75. [CrossRef]
66. Wang, R.; Chen, J.; Gao, K.; Hozumi, Y.; Yin, C.; Wei, G. Characterizing SARS-CoV-2 Mutations in the United States. *Res. Sq.* **2020**. [CrossRef]
67. Wang, R.; Chen, J.; Hozumi, Y.; Yin, C.; Wei, G.W. Decoding Asymptomatic COVID-19 Infection and Transmission. *J. Phys. Chem. Lett.* **2020**, *11*, 10007–10015. [CrossRef]
68. Wang, R.; Hozumi, Y.; Yin, C.; Wei, G.W. Decoding SARS-CoV-2 Transmission and Evolution and Ramifications for COVID-19 Diagnosis, Vaccine, and Medicine. *J. Chem. Inf. Model* **2020**, *60*, 5853–5865. [CrossRef]
69. Yap, P.S.X.; Tan, T.S.; Chan, Y.F.; Tee, K.K.; Kamarulzaman, A.; Teh, C.S.J. An Overview of the Genetic Variations of the SARS-CoV-2 Genomes Isolated in Southeast Asian Countries. *J. Microbiol. Biotechnol.* **2020**, *30*, 962–966. [CrossRef]
70. Yuan, F.; Wang, L.; Fang, Y.; Wang, L. Global Snp Analysis of 11,183 SARS-CoV-2 Strains Reveals High Genetic Diversity. *Transbound. Emerg. Dis.* **2020**. [CrossRef]
71. Zhang, Y.; Pan, Y.; Zhao, X.; Shi, W.; Chen, Z.; Zhang, S.; Liu, P.; Xiao, J.; Tan, W.; Wang, D.; et al. Genomic Characterization of SARS-CoV-2 Identified in a Reemerging COVID-19 Outbreak in Beijing's Xinfadi Market in 2020. *Biosaf. Health* **2020**, *2*, 202–205. [CrossRef] [PubMed]
72. Ziegler, K.; Steininger, P.; Ziegler, R.; Steinmann, J.; Korn, K.; Ensser, A. SARS-CoV-2 Samples May Escape Detection Because of a Single Point Mutation in the N Gene. *Eurosurveillance* **2020**, *25*, 2001650. [CrossRef] [PubMed]
73. Zuckerman, N.S.; Pando, R.; Bucris, E.; Drori, Y.; Lustig, Y.; Erster, O.; Mor, O.; Mendelson, E.; Mandelboim, M. Comprehensive Analyses of SARS-CoV-2 Transmission in a Public Health Virology Laboratory. *Viruses* **2020**, *12*, 854. [CrossRef] [PubMed]
74. Zhang, L.; Jackson, C.B.; Mou, H.; Ojha, A.; Peng, H.; Quinlan, B.D.; Rangarajan, E.S.; Pan, A.; Vanderheiden, A.; Suthar, M.S.; et al. SARS-CoV-2 Spike-Protein D614g Mutation Increases Virion Spike Density and Infectivity. *Nat. Commun.* **2020**, *11*, 6013. [CrossRef]
75. Yurkovetskiy, L.; Wang, X.; Pascal, K.E.; Tomkins-Tinch, C.; Nyalile, T.P.; Wang, Y.; Baum, A.; Diehl, W.E.; Dauphin, A.; Carbone, C.; et al. Structural and Functional Analysis of the D614g SARS-CoV-2 Spike Protein Variant. *Cell* **2020**, *183*, 739–751. [CrossRef]
76. Lauring, A.S.; Hodcroft, E.B. Genetic Variants of SARS-CoV-2—What Do They Mean? *JAMA* **2021**, *325*, 529–531. [CrossRef] [PubMed]