

## RESEARCH ARTICLE

# Molecular epidemiology of SARS-CoV-2 in Cyprus

Jan Richter<sup>1,2\*</sup>, Pavlos Fanis<sup>2,3</sup>, Christina Tryfonos<sup>1,2</sup>, Dana Koptides<sup>1,2</sup>, George Krashias<sup>1,2</sup>, Stavros Bashiardes<sup>1,2</sup>, Andreas Hadjisavvas<sup>2,4</sup>, Maria Loizidou<sup>2,4</sup>, Anastasis Oulas<sup>2,5</sup>, Denise Alexandrou<sup>6</sup>, Olga Kalakouta<sup>6</sup>, Mihalis I. Panayiotidis<sup>2,4</sup>, George M. Spyrou<sup>2,5</sup>, Christina Christodoulou<sup>1,2</sup>

**1** Molecular Virology Department, Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus, **2** The Cyprus School of Molecular Medicine, Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus, **3** Molecular Genetics, Function & Therapy Department, Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus, **4** Cancer Genetics, Therapeutics & Ultrastructural Pathology Department, Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus, **5** Bioinformatics Department, Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus, **6** Medical and Public Health Services, Ministry of Health, Nicosia, Cyprus

\* richter@cing.ac.cy



## OPEN ACCESS

**Citation:** Richter J, Fanis P, Tryfonos C, Koptides D, Krashias G, Bashiardes S, et al. (2021) Molecular epidemiology of SARS-CoV-2 in Cyprus. PLoS ONE 16(7): e0248792. <https://doi.org/10.1371/journal.pone.0248792>

**Editor:** Binod Kumar, Loyola University Health System, UNITED STATES

**Received:** March 26, 2021

**Accepted:** June 22, 2021

**Published:** July 21, 2021

**Copyright:** © 2021 Richter et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All sequences obtained were deposited at the GISAID EpiCov database accession numbers EPI\_ISL\_1164626 – EPI\_ISL\_1164751 ([www.gisaid.org](https://www.gisaid.org)).

**Funding:** The study was supported by seed funding from The Cyprus Institute of Neurology and Genetics.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

Whole genome sequencing of viral specimens following molecular diagnosis is a powerful analytical tool of molecular epidemiology that can critically assist in resolving chains of transmission, identifying of new variants or assessing pathogen evolution and allows a real-time view into the dynamics of a pandemic. In Cyprus, the first two cases of COVID-19 were identified on March 9, 2020 and since then 33,567 confirmed cases and 230 deaths were documented. In this study, viral whole genome sequencing was performed on 133 SARS-CoV-2 positive samples collected between March 2020 and January 2021. Phylogenetic analysis was conducted to evaluate the genomic diversity of circulating SARS-CoV-2 lineages in Cyprus. 15 different lineages were identified that clustered into three groups associated with the spring, summer and autumn/winter wave of SARS-CoV-2 incidence in Cyprus, respectively. The majority of the Cypriot samples belonged to the B.1.258 lineage first detected in September that spread rapidly and largely dominated the autumn/winter wave with a peak prevalence of 86% during the months of November and December. The B.1.1.7 UK variant (VOC-202012/01) was identified for the first time at the end of December and spread rapidly reaching 37% prevalence within one month. Overall, we describe the changing pattern of circulating SARS-CoV-2 lineages in Cyprus since the beginning of the pandemic until the end of January 2021. These findings highlight the role of importation of new variants through travel towards the emergence of successive waves of incidence in Cyprus and demonstrate the importance of genomic surveillance in determining viral genetic diversity and the timely identification of new variants for guiding public health intervention measures.

## Introduction

Corona Virus Disease-2019 (COVID-19) is a pulmonary disease caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) coronavirus [1]. SARS-CoV-2 was first detected in December 2019 in Wuhan city in the Hubei province of China in a patient with acute pneumonia [2, 3]. Following global spread of the virus, on March 11, 2020 the World Health Organization (WHO) characterized COVID-19 as a pandemic, which is still ongoing and as of February 24, 2021 over 110 million confirmed cases and 2.47 million deaths were reported worldwide (<https://covid19.who.int/>).

The complete genome of SARS-CoV-2 isolated from a patient in Wuhan, China was initially published on Jan 5, 2020 [2] and since then analysis of viral sequences worldwide is continuous with more than 614,000 complete genomes currently available in public databases such as the GISAID (<https://www.gisaid.org/>). The global real-time tracing of the viral spread through whole genome sequencing is important in the race to timely identify the emergence of novel SARS-CoV-2 variants that change the transmission, antigenic properties and/or pathogenicity of the virus [4]. The recent identification of three major Variants of Concern (VOC), with increased transmissibility and the potential to reduce vaccine effectiveness has led to increased surveillance efforts worldwide [5, 6].

The SARS-CoV-2 genome consists of a single-stranded positive-sense RNA of around 30 kb in size with a 5' cap and 3'-polyA tail. It contains six major open reading frames (ORFs) that encode 27 different proteins, in which four are structural proteins: Envelope (E), Membrane (M), Nucleocapsid (N) and Spike (S). The E protein function includes virion assembly and morphogenesis and regulates cell stress response and apoptosis by promoting inflammation [7]. The M protein plays a key role in the virion assembly within the host cells [8], while the N protein, can bind to the RNA genome, interact with the viral membrane protein, and play a critical role in enhancing the efficiency of virus replication, transcription and assembly throughout viral infection [9]. The most important factor that mediates virus entry and a primary determinant of cell tropism and pathogenesis of SARS-CoV-2 is the S protein [10]. S protein has two functional subunits, S1 and S2. S1 subunit contains the receptor-binding domain (RBD) which binds directly to the host receptor angiotensin-converting enzyme 2 (ACE2), enabling virus entry into host cells [11]. The S2 subunit is responsible for virus cell fusion [12]. Therefore, the S protein defines the infectivity of the virus and its transmissibility in the host cell [13]. As this protein is the major antigen inducing protective immune responses [14, 15], all vaccines developed or there are under development are directed against it.

The Republic of Cyprus is one of the countries in Europe least affected by the COVID-19 pandemic, presumably due to its rapid and effective response strategy that included high number of COVID-19 tests, effective tracing and isolation of cases and their contacts together with preventive measures (e.g. social distancing, wearing face masks and hand washing) [16]. In addition, being an island guaranteed high effectiveness of airport closure with regard to importing of new cases. The first two cases of COVID-19 in Cyprus were identified on March 9, 2020 and since then 33,567 confirmed cases and 230 deaths were documented (<https://covid19.ucy.ac.cy/>).

The aim of our study was to analyse and evaluate the genomic diversity of circulating SARS-CoV-2 lineages in Cyprus. We report 133 full genome sequences obtained from individuals tested positive by RT-PCR between March 2020 and January 2021 and evaluate their characteristics and relationship with the progression of the pandemic in Cyprus. Public sharing and use of genomic analyses are important tools for disease surveillance systems to track and trace acquired COVID-19 cases for more accurate decision making and appropriate public health action.

## Materials and methods

### Sample collection

The Department of Molecular Virology of Cyprus Institute of Neurology and Genetics was assigned as the reference laboratory for SARS-CoV-2 by the Cyprus Ministry of Health of the Republic of Cyprus. More than 180,000 samples from public as well as private hospitals were received and analysed since March 2020. For this retrospective observational study 133 SARS-CoV-2 positive nasopharyngeal swabs obtained from patients referred for diagnostic purposes between March 2020 and January 2021 were selected for full genome sequencing. A random, unselected approach for sample selection was taken without pre-screening for variants of interest to avoid sampling bias. The viral RNA had been previously detected using a qRT-PCR assay and all samples selected had a cycle threshold value (Ct) lower than 30. The study was approved by the Cyprus National Bioethics committee (EEBK 21.1.01.03). According to the approval, as samples were fully anonymized, the Bioethics committee waived the requirement for informed consent.

### Next generation sequencing

Total RNA was extracted from 200  $\mu$ l of nasopharyngeal swab fluid in a final volume of 50  $\mu$ l, using the MagMAX Viral/Pathogen Nucleic Acid Isolation Kit (Applied Biosystems) on a KingFisher™ Flex Purification System (Thermo Fisher Scientific). Libraries were prepared using the QIAseq SARS-CoV-2 Primer Panel in conjunction with the QIAseq FX DNA Library Kit (Qiagen) according to manufacturer's instructions. In brief, viral RNA was reverse transcribed to synthesize cDNA using random hexamers. cDNAs were amplified in two high-fidelity multiplex PCR reactions using two different primer pools that together cover the entire SARS-CoV-2 genome. The two enriched pools per sample were combined, purified using AMPure XP beads (Beckman Coulter) and quantified using the Qubit dsDNA HS Assay kit (Invitrogen). Fragmentation, end-repair and A-tailing was performed in a combined reaction per sample using 200 ng DNA. Next, Illumina platform-specific adapters were ligated to both ends of the DNA fragments. Library size selection and purification were carried out using AMPure XP beads (Beckman Coulter) in two rounds (0.8x and 1x respectively). The libraries were quality analyzed using the Agilent High Sensitivity D1000 ScreenTape on a 2200 TapeStation system (Agilent) and quantified using the Qubit dsDNA HS Assay kit (Invitrogen). Equimolar quantities of libraries were pooled (24 samples/pool) and sequenced on a Illumina MiSeq sequencer. All sequences obtained were deposited at the GISAID EpiCov database accession numbers EPI\_ISL\_1164626 –EPI\_ISL\_1164751 ([www.gisaid.org](http://www.gisaid.org)).

### Mapping, alignment and lineage assignment

The Burrows-Wheeler Aligner (BWA), version: 0.7.15 was used to map the raw reads to the coronavirus reference genome Wuhan-Hu-1 (NCBI ID:NC\_045512.2) [17]. Duplicate reads, which are likely to be the results of PCR bias, were marked using Picard version: 2.6.0 (<http://broadinstitute.github.io/picard/>). Samtools, version: 0.1.19, was used for additional BAM/SAM file manipulations [18]. SNPs, amino acid replacements and indels were verified using the CoV-Glue web interface [19]. Aligned reads were validated using the Integrative Genomics Viewers v.2.9.2 and consensus sequences were extracted [20]. SARS-CoV-2 lineages of the 133 sequences were assigned using the dynamic nomenclature tool PANGOLIN (<https://pangolin.cog-uk.io>) [21].

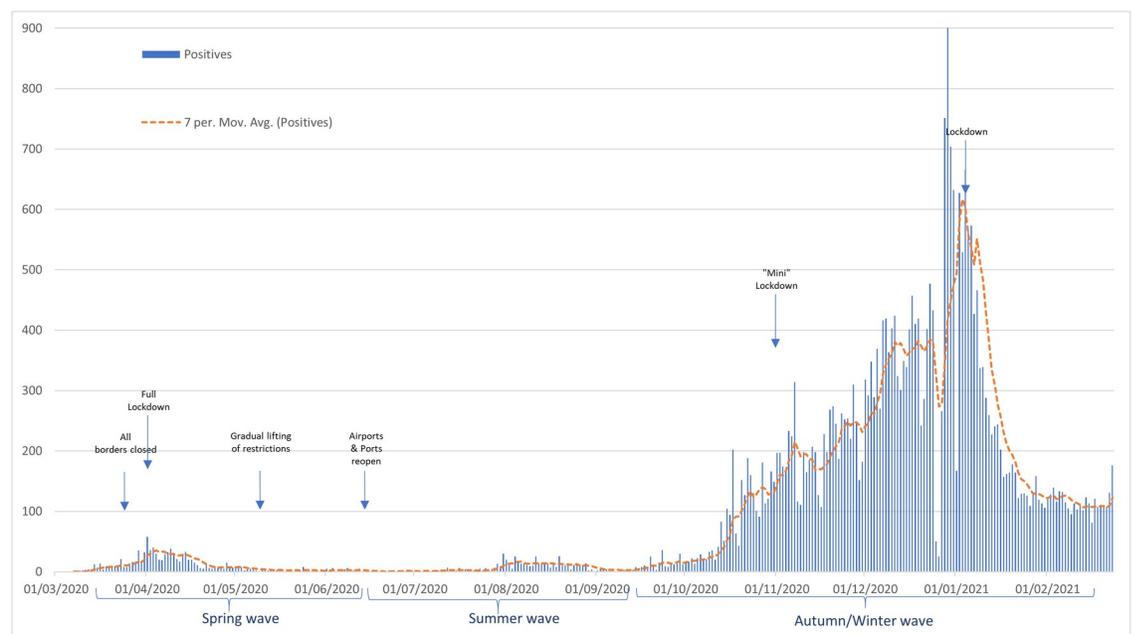
## Phylogenetic analysis

All phylogenetic analyses were conducted in MEGA7 [22]. The alignment of consensus sequences was performed using MUSCLE. Bayesian information criterion (BIC) scores were calculated for different models to determine the best fitting nucleotide substitution model. In addition, jModelTest [23] was used for evaluating the best fitting nucleotide substitution model under BIC yielding the same result (TN93model + gamma distributed rates). This model was subsequently used to construct a Maximum Likelihood phylogenetic tree with 1000 bootstrap replicates. All positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons.

## Results and discussion

Since the beginning of the pandemic and until the 24/2/2021, the Republic of Cyprus reported 33,567 SARS-CoV-2 confirmed cases (3.63% of the population) along with 230 associated casualties (24.2/100,000) (<https://covid19.ucy.ac.cy/>). Based on the number of positives identified (see Fig 1) the course of the pandemic in Cyprus so far can be distinguished into three main phases: a first phase during Spring 2020, starting with the first two cases identified on the 9<sup>th</sup> of March that peaked in the first week of April that was contained with a nationwide lockdown and lasted approximately until June; a second phase during the summer that started after the re-opening of the airports; and a third phase (which is still continuing) during autumn/winter that peaked at the beginning of January and accounted for the majority of all cases registered so far.

Overall, between March 11<sup>th</sup> 2020 and January 29<sup>th</sup> 2021, 15 different lineages were identified among the 133 samples sequenced (Table 1). As shown in Fig 3 the phylogenetic analysis is in support of the PANGOLIN lineages assignment and reveals a picture of independent distinct, importations that are followed by local transmission (Fig 3). In the bubble chart in Fig 2



**Fig 1. Daily number of positive SARS-CoV-2 cases in Cyprus identified between March 2020 and January 2021** (<https://covid19.ucy.ac.cy/>). The start of the implementation as well as the lifting of important governmental pandemic control measures are indicated.

<https://doi.org/10.1371/journal.pone.0248792.g001>

three groups of lineages are clearly distinguishable and associate with the three phases of the SARS-CoV-2 pandemic. Overall, 345 SNPs were identified within the 133 genomes when compared to the Wuhan-Hu-1 reference genome. [S1 Table](#) lists all SNPs identified along with the frequency of each SNP observed in each lineage.

During the spring wave four variants (B.1.1.74, B.1.1.277, B.1.235 and B.1.1.251) were identified reflecting several independent importation events followed by local transmission. However, concomitant with the decline in reported positive cases following the country-wide lockdown, by the end of June these SARS-CoV-2 variants were no longer detected and none of these re-emerged during the second or the third wave.

Following the re-opening of the airports, the summer months of July and August were characterized by multiple distinct introductions of a variety of new lineages (B.1.36, B.1.236, B.1.2, B.1.1.67, B.1.1.288, B.1.1.192), which marked the beginning of a small second wave ([Fig 3](#)). The majority of the samples belonged to lineage B.1.2 and B.1.236 accounting for 86% of the samples in that wave, however, all these lineages again subsequently vanished. The B.1.2 lineage has been otherwise observed only infrequently in Europe, but has been very common in the United States and in Canada [[24](#)].

In September, two new lineages were identified for the first time, which were set to dominate the autumn/winter wave until December, namely the B.1.177 and the B.1.258 variants.

The B.1.177 lineage, which is characterized by the A222V spike mutation, was previously been shown to have emerged in Spain in early summer and subsequently became widespread across Europe as well as Canada, accounting for the majority of sequences by autumn 2020 [[25](#), [26](#)]. However, no evidence of increased transmissibility of this variant was reported [[27](#)]. In Cyprus, this lineage was first identified on September 19<sup>th</sup> and continued to circulate until the end of January at a relatively low but constant frequency fluctuating between 4% and 15%.

In September, the B.1.258 lineage was introduced, which spread rapidly and largely dominated the autumn/winter wave with a peak prevalence of 86% during the months of November

**Table 1. Time course of SARS-CoV-2 lineages identified in Cyprus between March 2020 and January 2021.**

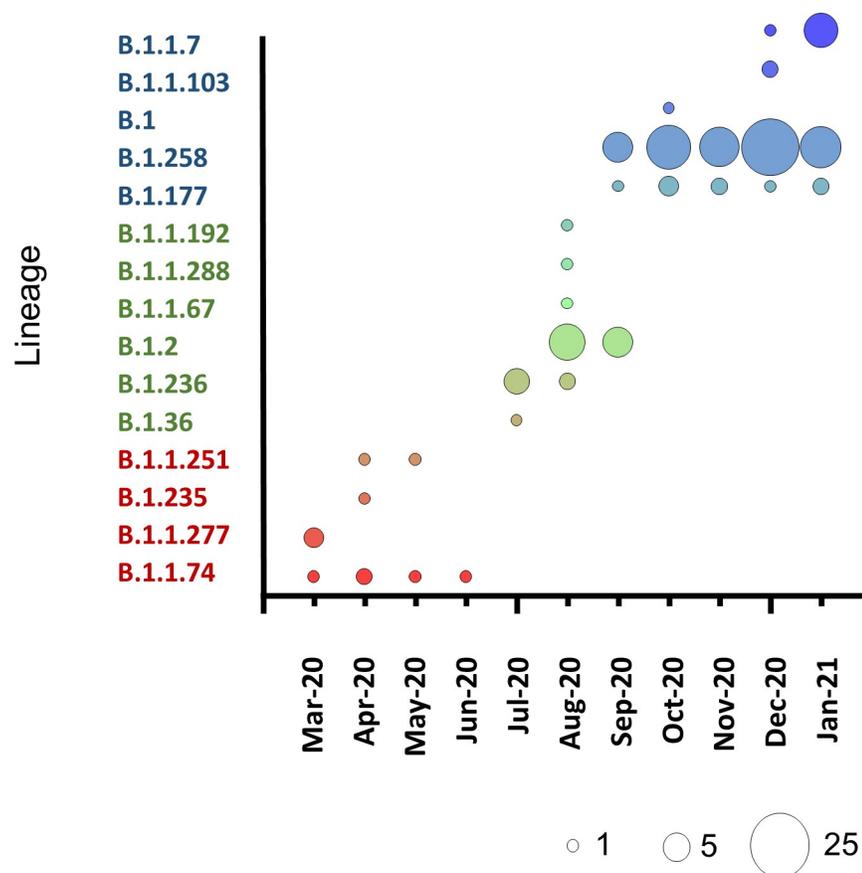
	Lineage	First detected	Last detected	Mar-20	Apr-20	May-20	Jun-20	Jul-20	Aug-20	Sep-20	Oct-20	Nov-20	Dec-20	Jan-21	Total
Spring Wave	B.1.1.74	11/03/20	10/06/20	1	2	1	1								5
	B.1.1.277	22/03/20	25/03/20	3											3
	B.1.235	17/04/20	17/04/20		1										1
	B.1.1.251	27/04/20	05/05/20		1	1									2
Summer wave	B.1.36	11/07/20	11/07/20					1							1
	B.1.236	17/07/20	18/08/20					5	2						7
	B.1.2	05/08/20	23/09/20						10	7					17
	B.1.1.67	13/08/20	13/08/20						1						1
	B.1.1.288	19/08/20	19/08/20						1						1
	B.1.1.192	27/08/20	27/08/20						1						1
Autumn/Winter wave	B.1.177	16/09/20	26/01/21							1	3	2	1	2	9
	B.1.258	23/09/20	27/01/21							7	16	12	25	13	73
	B.1.1.103	28/12/20	29/12/20										2		2
	B.1.1.7	28/12/20	27/01/21										1	9	10
				4	4	2	1	6	15	15	19	14	29	24	133

<https://doi.org/10.1371/journal.pone.0248792.t001>

and December. All samples belonging to the B.1.258 variant identified in Cyprus contain the six-nucleotide  $\Delta$ H69/V70 deletion in the S gene, which was recently proposed to be labelled B.1.258 $\Delta$  [28].

This deletion, which has been shown to enhance viral infectivity [29] has arisen at least six times independently and frequently followed receptor binding AA replacements (i.e. N501Y, N439K, Y453F) [28]. In addition, the B.1.258 variant is characterised by the N439K mutation that confers an increased binding affinity to the hACE2 receptor and leads simultaneously to immune escape from a panel of neutralizing monoclonal antibodies as well as from sera of persons recovered from infection [30].

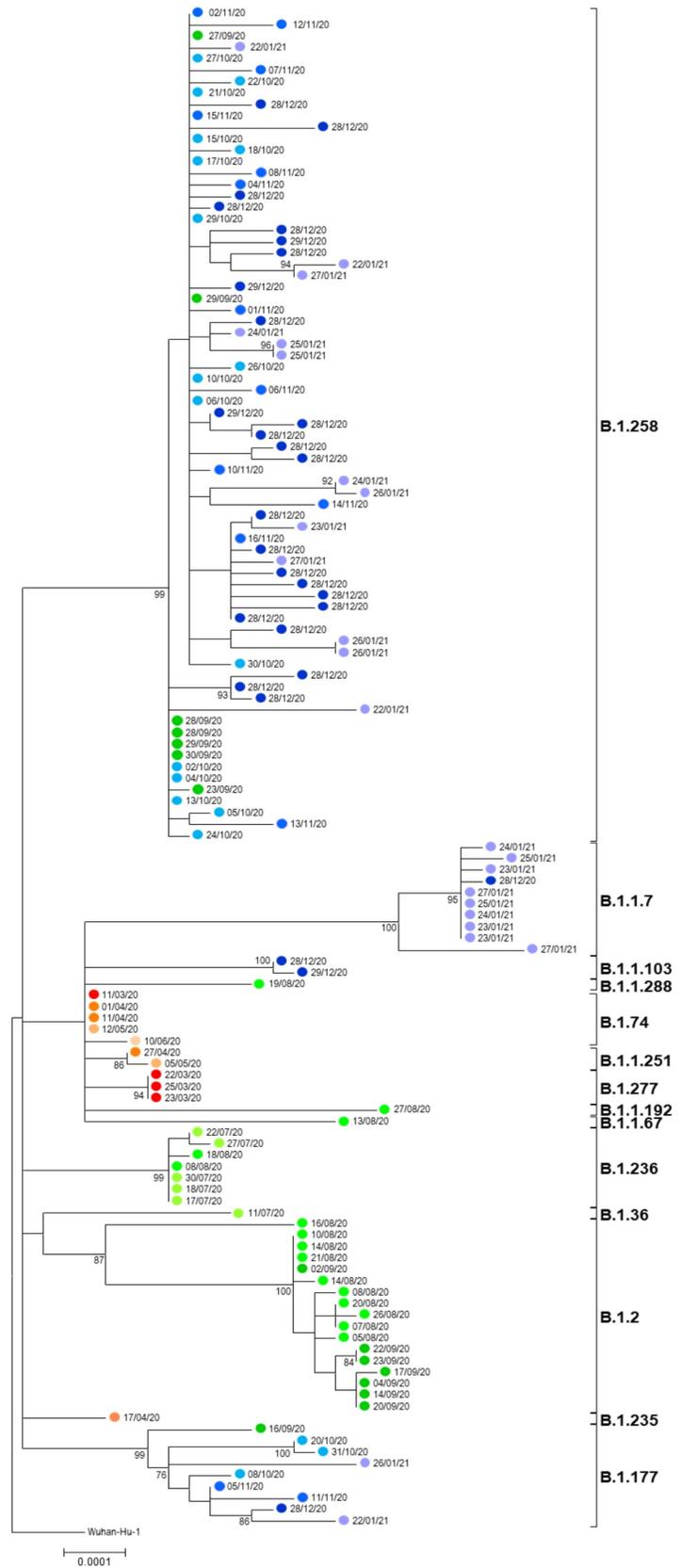
By the end of December, the B.1.1.7 UK variant, also known as 20I/501Y.V1 and Variant of Concern 202012/01 (VOC-202012/01), was identified for the first time in the Cypriot sample set. This lineage was first reported on December 20<sup>th</sup> in the United Kingdom and drew immediate attention due to its rapid spread and increased transmissibility [31]. The earliest sampled genome of the B.1.1.7 variant dates back to September 20<sup>th</sup>. Since then it has grown rapidly in frequency in the UK becoming the most prevalent lineage there with a similar development also observed in a variety of other European countries, including Ireland, Spain and Greece [32]. It is characterised by an unusual large number of mutations including N501Y in the receptor-binding domain that increases ACE2 receptor affinity and P681H that creates a furin cleavage site between S1 and S2, which was previously to promote transmission. It is



**Fig 2. Bubble chart illustrating the pattern of SARS-CoV-2 lineages identified per month.** The size of the circle is proportional to the number of samples.

<https://doi.org/10.1371/journal.pone.0248792.g002>

- Mar 20 ●
- Apr 20 ●
- May 20 ●
- Jun 20 ●
- Jul 20 ●
- Aug 20 ●
- Sep 20 ●
- Oct 20 ●
- Nov 20 ●
- Dec 20 ●
- Jan 21 ●



**Fig 3. Phylogenetic analysis of full genome sequences of 133 Cypriot strains aligned against the reference genome hCoV-19/Wuhan/Hu-1/2019 (NC\_045512.2).** Sample names consist of internal laboratory code followed by the date of sampling. The percentages of replicate trees in which the associated taxa clustered together in the standard bootstrap test (1000 replicates) are shown next to the branches. Only bootstrap values >70% are shown.

<https://doi.org/10.1371/journal.pone.0248792.g003>

speculated that it may have originated in a chronically infected immunocompromised person. In our sample set the frequency of this variant increased in a similar rapid manner from 3.4% in December to 38% by the end of January despite declining case numbers highlighting its superior fitness.

Similar studies have been reported from a plethora of countries around the world including the UK, Ireland, Russia, India, Qatar etc. [26, 33–36]. These studies provide valuable information on the routes of importation of new strains and the emergence of critical mutations that change key parameters such as transmissibility or virulence, as was recently highlighted by the discovery of the B.1.617 variant and its B.1.617.2 sub-lineage in India [37], which led to a surge of SARS-CoV-2 infections in Delhi. It could be shown that B.1.617.2 exhibits a very high transmissibility, as much as 50% greater than B.1.1.7, while no increase in the case-fatality ratio (CFR) was observed.

In Cyprus, the CFR after an initial peak exhibited a falling trend that stabilised over time (see S1 Fig). However, it is difficult to draw conclusions about the virus pathogeny from the CFR alone, as reason for the initially high CFR may lie in the lower testing capacity at the beginning of the pandemic leading to an underestimation of the total number positive cases. Other critical factors that were shown to impact on CFR and likely contributed to the observed reduction are hospital capacity and COVID-19 treatment protocols [38].

## Conclusions

Whole genome sequencing of viral specimens following molecular diagnosis is a powerful analytical tool of molecular epidemiology that can critically assist in resolving chains of transmission, identifying of new variants or assessing pathogen evolution and allows a real-time view into the dynamics of a pandemic [39]. In this study we describe for the first time the changing pattern of circulating SARS-CoV-2 lineages in Cyprus since its appearance between March 2020 and the end of January 2021. Distinct lineages of SARS-CoV-2 contributing to three separate waves of infections reflective of the epidemiological pattern were observed.

The global real-time tracing of the viral spread through whole genome sequencing has led recently to the identification of three major Variants of Concern (VOC). Lineage B.1.1.7 (20I/501Y.V1, VOC-202012/01) better known as the UK variant possesses an unusual high number of mutations is believed to be more transmissible than the wild-type virus. Lineage B.1.351 (501.V2, 20H/501Y.V2, VOC-202012/02) first detected and reported in South Africa in early October 2020, shares several mutations with B.1.1.7 and is feared to reduce to some extent vaccine effectiveness [40]. In Brazil, the P.1 (20J/501Y.V3, VOC-202101/02) lineage emerged in December 2020 [41]. The P.1 lineage contains ten mutations in the spike protein that may affect its ability to be recognized by antibodies [42, 43].

Of these three VOC only the B.1.1.7 variant has so far been identified in Cyprus, which after being encountered for the first time at the end of December was able to reach 37% prevalence within one month. The majority of the Cypriot samples analysed belonged to the B.1.258Δ lineage, which has been first detected at the beginning of August in Switzerland and has since then spread to numerous European countries where it became one of the most prevalent lineages by December including the Czech Republic, Slovakia and Sweden [28, 44]. The N439K receptor binding motif characteristic of this lineage has previously been reported to confer an

increased binding affinity to the hACE2 receptor and simultaneously leading to immune escape from a panel of neutralizing monoclonal antibodies as well as from sera of persons recovered from infection [30]. The association of mutations found in the B.1.258Δ lineage with increased fitness and immune evasion along with the high prevalence in several European countries stipulates further characterisation of this variant. A continuous surveillance of SARS-CoV-2 by whole genome sequencing continues to be critical for timely detection of emerging variants, identifying transmission modes and guiding public health intervention.

## Supporting information

**S1 Fig. Case-fatality rate (CFR) in Cyprus from March 2020 until January 2021.** The CFR was calculated as the ratio between cumulative number of confirmed deaths and cumulative confirmed cases. The daily number of cases are shown for comparison.  
(PDF)

**S1 Table. Frequency of SNPs in lineages observed.** Single nucleotide polymorphisms (SNPs) were identified in the genomes with respect to the reference sequence (NC\_045512). The frequency of SNPs is calculated by the number of sequences with the SNP/total number of sequences per lineage. SNPs are ordered by their position on the genome. Number of sequences of each lineage are indicated in brackets in the header. SNPs with frequency >75% are marked in green, SNPs with frequency between 50% and 75% are marked in yellow.  
(PDF)

## Author Contributions

**Conceptualization:** Jan Richter, Christina Christodoulou.

**Data curation:** Anastasis Oulas, George M. Spyrou.

**Formal analysis:** Jan Richter, Pavlos Fanis, George Krashias, Anastasis Oulas, Christina Christodoulou.

**Investigation:** Pavlos Fanis, Christina Tryfonos, Dana Koptides, Andreas Hadjisavvas, Maria Loizidou, Denise Alexandrou, Olga Kalakouta, Mihalis I. Panayiotidis.

**Methodology:** Jan Richter, Christina Tryfonos, Dana Koptides, Stavros Bashiardes.

**Project administration:** Christina Christodoulou.

**Resources:** Christina Christodoulou.

**Software:** Anastasis Oulas.

**Supervision:** Christina Christodoulou.

**Validation:** George M. Spyrou.

**Writing – original draft:** Jan Richter, Pavlos Fanis, Christina Tryfonos.

**Writing – review & editing:** Dana Koptides, George Krashias, Stavros Bashiardes, Andreas Hadjisavvas, Maria Loizidou, Anastasis Oulas, Denise Alexandrou, Olga Kalakouta, Mihalis I. Panayiotidis, George M. Spyrou, Christina Christodoulou.

## References

1. Gorbalenya AE, Baker SC, Baric RS, de Groot RJ, Drosten C, Gulyaeva AA, et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2.

- Nature Microbiology. Nature Research; 2020. pp. 536–544. <https://doi.org/10.1038/s41564-020-0695-z> PMID: 32123347
2. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020; 579: 265–269. <https://doi.org/10.1038/s41586-020-2008-3> PMID: 32015508
  3. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med*. 2020; 382: 727–733. <https://doi.org/10.1056/NEJMoa2001017> PMID: 31978945
  4. consortiumku.muitrosnocgoc@tcatnoc TC-19 GU (COG-U. An integrated national scale SARS-CoV-2 genomic surveillance network. *The Lancet Microbe*. 2020; 1: e99–e100. [https://doi.org/10.1016/S2666-5247\(20\)30054-9](https://doi.org/10.1016/S2666-5247(20)30054-9) PMID: 32835336
  5. Mitze T, Rode J. Early assessment of epidemiological trends associated with SARS-CoV-2 variants of concern in Germany. medRxiv. 2021; 2021.02.16.21251803. <https://doi.org/10.1101/2021.02.16.21251803>
  6. Challen R, Brooks-Pollock E, Read JM, Dyson L, Tsaneva-Atanasova K, Danon L. Risk of mortality in patients infected with SARS-CoV-2 variant of concern 202012/1: matched cohort study. *BMJ*. 2021; 372: n579. <https://doi.org/10.1136/bmj.n579> PMID: 33687922
  7. Liu DX, Yuan Q, Liao Y. Coronavirus envelope protein: A small membrane protein with multiple functions. *Cellular and Molecular Life Sciences*. *Cell Mol Life Sci*; 2007. pp. 2043–2048. <https://doi.org/10.1007/s00018-007-7103-1> PMID: 17530462
  8. Arndt AL, Larson BJ, Hogue BG. A Conserved Domain in the Coronavirus Membrane Protein Tail Is Important for Virus Assembly. *J Virol*. 2010; 84: 11418–11428. <https://doi.org/10.1128/JVI.01131-10> PMID: 20719948
  9. McBride R, van Zyl M, Fielding BC. The coronavirus nucleocapsid is a multifunctional protein. *Viruses*. MDPI AG; 2014. pp. 2991–3018. <https://doi.org/10.3390/v6082991> PMID: 25105276
  10. Belouzard S, Millet JK, Licitra BN, Whittaker GR. Mechanisms of coronavirus cell entry mediated by the viral spike protein. *Viruses*. *Viruses*; 2012. pp. 1011–1033. <https://doi.org/10.3390/v4061011> PMID: 22816037
  11. Xiao X, Chakraborti S, Dimitrov AS, Gramatikoff K, Dimitrov DS. The SARS-CoV S glycoprotein: Expression and functional characterization. *Biochem Biophys Res Commun*. 2003; 312: 1159–1164. <https://doi.org/10.1016/j.bbrc.2003.11.054> PMID: 14651994
  12. Wan Y, Shang J, Graham R, Baric RS, Li F. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J Virol*. 2020; 94. <https://doi.org/10.1128/JVI.00127-20> PMID: 31996437
  13. Hulswit RJG, de Haan CAM, Bosch BJ. Coronavirus Spike Protein and Tropism Changes. *Advances in Virus Research*. Academic Press Inc.; 2016. pp. 29–57. <https://doi.org/10.1016/bs.avir.2016.08.004> PMID: 27712627
  14. Du L, He Y, Zhou Y, Liu S, Zheng BJ, Jiang S. The spike protein of SARS-CoV—A target for vaccine and therapeutic development. *Nature Reviews Microbiology*. *Nat Rev Microbiol*; 2009. pp. 226–236. <https://doi.org/10.1038/nrmicro2090> PMID: 19198616
  15. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*. 2020; 181: 281–292.e6. <https://doi.org/10.1016/j.cell.2020.02.058> PMID: 32155444
  16. Quattrocchi A, Mamais I, Tsioutis C, Christaki E, Constantinou C, Koliou M, et al. Extensive Testing and Public Health Interventions for the Control of COVID-19 in the Republic of Cyprus between March and May 2020. *J Clin Med*. 2020; 9: 3598. <https://doi.org/10.3390/jcm9113598> PMID: 33171651
  17. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009; 25: 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324> PMID: 19451168
  18. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; 25: 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943
  19. Singer JB, Gifford RJ, Cotten M, Robertson DL. CoV-GLUE: A Web Application for Tracking SARS-CoV-2 Genomic Variation. 2020 [cited 1 Mar 2021]. <https://doi.org/10.20944/preprints202006.0225.v1>
  20. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nature Biotechnology*. *Nat Biotechnol*; 2011. pp. 24–26. <https://doi.org/10.1038/nbt.1754> PMID: 21221095
  21. Rambaut A, Holmes EC, O’Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol*. 2020; 5: 1403–1407. <https://doi.org/10.1038/s41564-020-0770-5> PMID: 32669681

22. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol.* 2016; 33: 1870–1874. <https://doi.org/10.1093/molbev/msw054> PMID: 27004904
23. Darriba D, Taboada GL, Doallo R, Posada D. JModelTest 2: More models, new heuristics and parallel computing. *Nat Methods.* 2012; 9: 772. <https://doi.org/10.1038/nmeth.2109> PMID: 22847109
24. Pater AA, Bosmeny MS, Barkau CL, Ovington KN, Chilamkurthy R, Parasrampur M, et al. Emergence and Evolution of a Prevalent New SARS-CoV-2 Variant in the United States. *bioRxiv.* 2021; 2021.01.11.426287. <https://doi.org/10.1101/2021.01.11.426287>
25. Guthrie JL, Teatero S, Zittermann S, Chen Y, Sullivan A, Rilkoﬀ H, et al. Detection of the novel SARS-CoV-2 European lineage B.1.177 in Ontario, Canada. *medRxiv.* medRxiv; 2020. p. 2020.11.30.20241265. <https://doi.org/10.1101/2020.11.30.20241265>
26. Pwg M, Leon GA, Barra de E, Patrick Mallon P. Whole-genome sequencing of SARS-CoV-2 in the Republic of Ireland during waves 1 and 2 of the pandemic. *medRxiv.* 2021; 2021.02.09.21251402. <https://doi.org/10.1101/2021.02.09.21251402>
27. Hodcroft EB, Zuber M, Nadeau S, Crawford KHD, Bloom JD, Velesler D, et al. Emergence and spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *medRxiv.* medRxiv; 2020. p. 2020.10.25.20219063. <https://doi.org/10.1101/2020.10.25.20219063> PMID: 33269368
28. Brejová B, Hodorová V, Boršová K, Čabanová V, Reizigová L, Paul ED, et al. B.1.258 $\Delta$ 69, a SARS-CoV-2 variant with  $\Delta$ 69 $\Delta$ 70 in the Spike protein circulating in the Czech Republic and Slovakia. 2021 [cited 2 Mar 2021]. Available: <http://arxiv.org/abs/2102.04689>
29. Kemp SA, Datir RP, Collier DA, Ferreira IATM, Carabelli A, Harvey W, et al. Recurrent emergence and transmission of a SARS-CoV-2 Spike deletion  $\Delta$ H69/V70. *bioRxiv.* bioRxiv; 2020. p. 2020.12.14.422555. <https://doi.org/10.1101/2020.12.14.422555>
30. Thomson EC, Rosen LE, Shepherd JG, Spreafico R, da Silva Filipe A, Wojcechowskyj JA, et al. Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. *Cell.* 2021; 184. <https://doi.org/10.1016/j.cell.2021.01.037> PMID: 33621484
31. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations—SARS-CoV-2 coronavirus / nCoV-2019 Genomic Epidemiology—Virological. [cited 2 Mar 2021]. Available: <https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563>
32. RAPID RISK ASSESSMENT: SARS-CoV-2 increased circulation of variants of concern and vaccine rollout in the EU/EEA, 14th update. [cited 2 Mar 2021]. Available: <https://www.ecdc.europa.eu/sites/default/files/documents/RRR-covid-19-14th-update-15-feb-2021.pdf>
33. du Plessis L, McCrone JT, Zarebski AE, Hill V, Ruis C, Gutierrez B, et al. Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. *Science* (80-). 2021; 371: 708–712. <https://doi.org/10.1126/science.abf2946> PMID: 33419936
34. Matsvay A, Klink G V, Safina KR, Nabieva E, Garushyants SK, Biba D, et al. Genomic epidemiology of SARS-CoV-2 in Russia reveals recurring cross-border transmission throughout 2020. *medRxiv.* 2021; 2021.03.31.21254115. <https://doi.org/10.1101/2021.03.31.21254115>
35. Pattabiraman C, Habib F, Harsha PK, Rasheed R, Prasad P, Reddy V, et al. Genomic epidemiology reveals multiple introductions and spread of SARS-CoV-2 in the Indian state of Karnataka. *PLoS One.* 2020; 15. <https://doi.org/10.1371/journal.pone.0243412> PMID: 33332472
36. Benslimane FM, Al Khatib HA, Al-Jamal O, Albatash D, Boughattas S, Ahmed AA, et al. One year of SARS-CoV-2: Genomic characterization of COVID-19 outbreak in Qatar. *medRxiv.* 2021; 2021.05.19.21257433. <https://doi.org/10.1101/2021.05.19.21257433>
37. Dhar MS, Marwal R, Ponnusamy K, Jolly B, Bhojar RC, Fatihi S, et al. Genomic characterization and Epidemiology of an emerging SARS-CoV-2 variant in Delhi, India. *medRxiv.* 2021; 2021.06.02.21258076. <https://doi.org/10.1101/2021.06.02.21258076>
38. Ergönül Ö, Akyol M, Tanrıöver C, Tiemeier H, Petersen E, Petrosillo N, et al. National case fatality rates of the COVID-19 pandemic. *Clin Microbiol Infect.* 2021; 27: 118–124. <https://doi.org/10.1016/j.cmi.2020.09.024> PMID: 32979575
39. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, et al. NextStrain: Real-time tracking of pathogen evolution. *Bioinformatics.* 2018; 34: 4121–4123. <https://doi.org/10.1093/bioinformatics/bty407> PMID: 29790939
40. Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, et al. Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *medRxiv.* medRxiv; 2020. p. 2020.12.21.20248640. <https://doi.org/10.1101/2020.12.21.20248640>

41. Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings—SARS-CoV-2 coronavirus / nCoV-2019 Genomic Epidemiology—Virological. [cited 4 Mar 2021]. Available: <https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586>
42. Jangra S, Ye C, Rathnasinghe R, Stadlbauer D, Krammer F, Simon V, et al. The E484K mutation in the SARS-CoV-2 spike protein reduces but does not abolish neutralizing activity of human convalescent and post-vaccination sera. medRxiv Prepr Serv Heal Sci. 2021 [cited 4 Mar 2021]. <https://doi.org/10.1101/2021.01.26.21250543> PMID: 33532796
43. Liu Z, VanBlargan LA, Rothlauf PW, Bloyet LM, Chen RE, Stumpf S, et al. Landscape analysis of escape variants identifies SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. bioRxiv. bioRxiv; 2020. <https://doi.org/10.1101/2020.11.06.372037> PMID: 33442690
44. Brandt1 C, Spott1 R, Hölzer2 M, Kühnert3 D, Fuchs2 S, Lohde1 M, et al. Molecular epidemiology of SARS-CoV-2—a regional to global perspective. medRxiv. 2021; 2021.01.25.21250447. <https://doi.org/10.1101/2021.01.25.21250447>