White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Leeds.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/161723/

Version: Accepted Version

Article:
Sironi, M, Hasnain, SE, Rosenthal, B et al. (2020) SARS-CoV-2 and COVID-19: A genetic, epidemiological, and evolutionary perspective. Infection, Genetics and Evolution, 84. 104384. ISSN: 1567-1348

https://doi.org/10.1016/j.meegid.2020.104384

© 2020, Elsevier. All rights reserved. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/.

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

SARS-CoV-2 and COVID-19: A Genetic, Epidemiological, and Evolutionary Perspective

Manuela Sironi, Seyed E. Hasnain, Benjamin Rosenthal, Tung Phan, Fabio Luciani, Marie-Anne Shaw, M. Anice Sallum, Marzieh Ezzaty Mirhashemi, Serge Morand, Fernando González-Candelas, on behalf of the Editors of Infection, Genetics and Evolution.

Manuela Sironi: Bioinformatics Unit, Scientific Institute IRCCS E. MEDEA, Bosisio Parini (LC), Italy. manu.sir@gmail.com

Seyed E. Hasnain: JH Institute of Molecular Medicine, Jamia Hamdard, Tughlakabad, New Delhi, India. seyedhasnain@gmail.com

Benjamin M. Rosenthal: Animal Parasitic Disease Laboratory, Agricultural Research Service, United States Department of Agriculture, Beltsville, MD, USA. benjamin.rosenthal@ars.usda.gov

Tung Phan: Division of Clinical Microbiology, University of Pittsburgh and University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA. phantg@upmc.edu

Fabio Luciani: University of New South Wales, Sydney, 2052, New South Wales, Australia. luciani@unsw.edu.anu

Marie-Anne Shaw: Leeds Institute of Medical Research at St James's, School of Medicine, University of Leeds, Leeds, United Kingdom. genmas@leeds.ac.uk

M. Anice Sallum: Departamento de Epidemiologia, Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo, Brazil. masallum@usp.br

Marzieh Ezzaty Mirhashemi: University of Massachusetts Medical School, Worcester, Massachusetts, 01655-0112, USA. Marzieh.Mirhashemi@umassmed.edu

Serge Morand: Institute of Evolution Science of Montpellier, Case Courier 064, F-34095, Montpellier, France. serge.morand@univ-montp2.fr

Fernando González-Candelas: Joint Research Unit Infection and Public Health FISABIO-University of Valencia, Institute for Integrative Systems Biology (I2SysBio) and CIBER in Epidemiology and Public Health. Valencia, Spain. fernando.gonzalez@uv.es

ABSTRACT

In less than five months, COVID-19 has spread from a small focus in Wuhan, China to more than 4.5 million people in almost every country in the world, dominating the concern of most governments and public health systems. The social and political distresses caused by this epidemic will certainly impact our world for a long time to come. Here, we synthesize lessons from a range of scientific perspectives rooted in epidemiology, virology, genetics, ecology and evolutionary biology so as to provide perspective on how this pandemic started, how it is developing, and how best we can stop it.

KEYWORDS

coronavirus; coevolution; host susceptibility; immune system; pandemics; phylodynamics

Introduction

Pathogen X, the hypothetical unknown potentially devastating microorganism capable of causing a major pandemic (Friedrich 2018), now has a name: Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Despite warnings and preparedness efforts by WHO and other health agencies, the rapid spread of COVID-19, which in less than 4 months has moved from affecting a few persons in Wuhan (Hubei province, China) to more than 5 million people in almost every country in the world (Coronavirus Research Center, https://coronavirus.jhu.edu/map.html visited on May 21th, 2020) has caught by surprise most governments and public health systems. The result cannot be evaluated yet but the over 300,000 officially recognized deaths and the economic, social and political distresses caused by this epidemic will certainly impact our world in the coming months, probably years. In light of the wave of disinformation and mistrust by wide sectors in the public towards experts and scientists' views on how to stop this pandemic, we feel the need to bring a scientific perspective on the source, causes, uses, and possible ways to halt the spread of this new coronavirus from a basic science perspective, that marking the core nature of this journal, blending Genetics, Virology, Epidemiology, and Evolutionary Biology ("Population Genetics and Population Biology: What Did They Bring to the Epidemiology of Transmissible Diseases? An E-Debate" 2001).

The origin of SARS-CoV-2 and COVID-19

The novel human coronavirus (SARS-CoV-2), responsible for the current COVID-19 pandemic, was first identified in December 2019, in the Hubei province of China (Zhu et al. 2020). After SARS-CoV (severe acute respiratory syndrome coronavirus) and MERS-CoV (Middle East respiratory syndrome Coronavirus), SARS-CoV-2 is the third highly pathogenic coronavirus to emerge and spread in human populations. Phylogenetic analyses showed that, as SARS-CoV, SARS-CoV-2 is a member of the *Sarbecovirus* subgenus (genus *Betacoronavirus*) (Zhou et al. 2020). Recently, the International Committee on Taxonomy of Viruses (ICTV) indicated that SARS-CoV-2 is to be classified within the species *Severe acute*

respiratory syndrome-related coronavirus (Coronaviridae Study Group of the Inte...). MERS-CoV (subgenus *Merbecovirus*) is more distantly related to SARS-CoV-2 than SARS-CoV.

In the aftermath of the SARS and MERS epidemics, intense efforts were devoted to identify the animal reservoirs of these viruses and to reconstruct the chain of events that led to the human spillovers. It is now known that both viruses originated in bats and were transmitted to humans by intermediate hosts (Normile and Enserink 2003; Killerby et al. 2020). In particular, the progenitor of SARS-CoV emerged through recombination among bat viruses and subsequently infected palm civets and other small carnivores, eventually spilling over to humans (Cui et al. 2019). MERS-CoV was most likely transmitted from bats to dromedary camels decades earlier than the first human cases were registered (Müller et al. 2014). MERS-CoV displays limited human-to-human spread, and most cases resulted from independent zoonotic transmission from camels (Cui et al. 2019). It is thus unsurprising that the closest known relatives of SARS-CoV-2 are two bat-coronavirus (BatCoV RmYN02 and BatCoV RaTG13) identified in horseshoe bats (*Rhinolophus malyanus and R. affinis, respectively*) ("A Novel Bat Coronavirus Closely Related to SARS-CoV-2 Contains Natural Insertions at the S1/S2 Cleavage Site of the Spike Protein" 2020)(Zhou et al. 2020). Both BatCoVs display an average nucleotide identity of ~95% with SARS-CoV-2, although with variations through the genome, and they were detected in bats sampled in Yunnan province, China, in 2019 and 2013, before the first detection of SARS-CoV-2 in humans (Zhou et al. 2020). Both the place and timing of BatCoV RmYN02 and RaTG13 detection, as well as their levels of identity with SARS-CoV-2, indicate that these viruses are not direct progenitors of the SARS-CoV-2 epidemic strain, but clearly support the view that the latter had an ultimate bat origin (Zhou et al. 2020) (Figure 1A).

In analogy to SARS-CoV and MERS-CoV, several lines of evidence suggest that an intermediate host was responsible for the cross-species transmission of SARS-CoV-2 to humans. First, most although not all, early COVID-19 detected cases were associated with the Huanan seafood and wildlife market in Wuhan city, where several mammalian species were traded (Huang et al. 2020). This is reminiscent of the circumstances associated with the initial phases of SARS-CoV spread, as palm civets were sold in wet markets and their meat consumed (Cui, Li, and Shi 2019). Second, *in vitro* experiments have shown that, in

Figure 1 (see legend after References).

addition to bats, SARS-CoV-2 can infect cells from small carnivores and pigs (Zhou et al. 2020). Experimental *in vivo* infection and transmission in ferrets and cats was also reported ("Infection and Rapid Transmission of SARS-CoV-2 in Ferrets" 2020)(J. Shi et al. 2020). Third, viruses very closely related (85.5% to 92.4% sequence similarity) to SARS-CoV-2 were very recently detected in Malayan or Sunda pangolins (*Manis javanica*) illegally imported in Southern China (Lam et al. 2020). The analysis of these viral genomes indicated that pangolins host at least two sub-lineages of sarbecoviruses, which are referred as the Guangdong and Guangxi lineages after the locations where the animals were sampled (Lam et al. 2020) (Figure 1A and B). Viruses in the Guangdong lineage share high similarity with SARS-CoV-2 in the receptor-binding motif of the spike protein. This same region is instead the most divergent between BatCoV RaTG13 and SARS-CoV-2 (Lam et al. 2020; (Zhou et al. 2020). This observation is very relevant, as the binding affinity between the spike protein and the cognate cellular receptor (angiotensin-converting enzyme 2, ACE2, in the case of SARS-CoV and SARS-CoV-2) is a major determinant of coronavirus host range (Haijema et al. 2003;Kuo et al. 2000; McCray et al. 2007; Moore et al. 2004; Schickli et al. 2004). Indeed, coronaviruses use a surface spike glycoprotein to attach to host receptors and gain entry into cells (Walls et al. 2020). It is a homo-trimeric protein formed by two subunits, S1 and S2. The S1 subunit contains an N-terminal domain connected by a linker of variable length to the receptor-binding domain (RBD). In various coronaviruses, the N-terminal domain (NTD) and RBD contribute to define host range (Haijema et al. 2003;Kuo et al. 2000; McCray et al. 2007; Moore et al. 2004; Schickli et al. 2004; (Lu, Wang, and Gao 2015). The S2 domain participates in membrane fusion (Duquerroy et al., 2005). Comparison of the complete spike protein, as well as of the NTD, indicated higher similarity of SARS-CoV-2 with RaTG13 than with pangolin coronaviruses (Figure 1A). Domain based sequence analysis indicates that sequence variations are majorly confined to the S1-domain (~16% variable sites) than to the S2-domain (Figure 1C), and mutations were observed in both NTD and RBD. Interestingly, an insertion (YLTPGD) is present only in the NTD of SARS-CoV-2, RaTG13, and Guangxi pangolin coronaviruses, but absent in related bat coronaviruses (in Guangdong pangolin viruses the region is not fully covered by sequencing) (Figure 1D). This motif forms a conformational

cluster at the exposed NTD regions of the spike trimer and may contribute to determine host range. However, as mentioned above, within the RBD, SARS-CoV-2 shows higher identity with pangolin viruses belonging to the Guangdong lineage than to RaTG13 (Figure 1B and 1D). Structural analysis of the binding interface between SARS-CoV-2 RBD and human ACE2 shows a strong network of polar contacts (PDB id: 6m0j) (Lan et al. 2020). SARS-CoV-2 RBD mediates these polar interactions through Lys417, Gly446, Tyr449, Asn487, Gln493, Gln498, Thr500, Asn501, Gly502 and Tyr505 with ACE2 (Figure 1E). These residues, which participate in polar interactions with the host protein, are almost completely conserved with pangolin coronaviruses of the Guangdong lineage (Figure 1D).

One possible explanation for this observation is that the RBDs of SARS-CoV-2 and Guangdong pangolin viruses have been progressively optimized through natural selection (convergent evolution) to bind ACE2 molecules from humans and pangolins (and possibly other non-bat mammalian species) (T. T.-Y. Lam et al. 2020). An alternative possibility is that recombination events among coronaviruses hosted by bats, pangolins, and possibly other mammals originated the progenitor of SARS-CoV-2 (T. T.-Y. Lam et al. 2020; Cagliani et al. 2020). It is presently impossible to disentangle these two alternative scenarios, and only the sequencing of additional related sarbecoviruses might eventually clarify the evolutionary history of SARS-CoV-2 RBD. It is also worth mentioning here that, as previously noted (Andersen et al. 2020), the similarity of the SARS-CoV-2 RBD with that of viruses only recently sequenced from pangolins can be regarded as a major evidence against the circulating theory that SARS-CoV-2 is the result of deliberate human manipulation.

In any case, these data do not necessarily imply that pangolins had a role in the emergence of SARS-CoV-2 and in its spread to humans, as these animals might have in turn contracted infection from a bat or other reservoir. Moreover, the SARS-CoV-2 spike protein displays a unique feature that is not shared with either BatCoV RaTG13 or the pangolin viruses, namely the presence of a furin cleavage site insertion (PRRA) at the S1-S2 junction (Walls et al. 2020) (Figures 1D and 1F). This feature, also absent in SARS-CoV, was suggested to increase viral infectivity and/or pathogenicity (Walls et al. 2020)(Andersen et al. 2020). A three amino acid insertion (PAA) is however present at the S1/S2 junction of the RmYN02 spike protein (Figure 1F) (A Novel Bat Coronavirus Closely Relat...). Because the insertions in the two viruses differ in sequence, they were most likely acquired independently. It is presently unknown how and

when SARS-CoV-2 acquired the furin cleavage site, but it is equally unexplored whether it affects any viral phenotype or if it contributed to adaptation to humans or other hosts. Importantly, though, the presence of a similar insertion in a virus isolated from wild bats is another strong indication in favor of a natural animal origin of SARS-CoV-2 (A Novel Bat Coronavirus Closely Relat...).

Where did adaptation to humans occur?

Although, for the reasons mentioned above, an as-yet unidentified intermediate host is likely to have played a role in the zoonotic transmission of SARS-CoV-2 in the Wuhan market or elsewhere, the possibility that SARS-CoV-2 was directly transmitted from bats to humans cannot be discarded. Indeed, serological surveys on people living in proximity to bat colonies in Yunnan province indicated that direct bat-to-human transmission of SARS-CoV related coronaviruses might occur (Wang et al. 2018). In general, wild animal trade might reduce the ecological barriers separating humans from coronavirus hosts, including these bats or other mammals. However, the observation that some early COVID-19 cases had no apparent epidemiological link to the Huanan seafood and wildlife market (Huang et al. 2020) opens the possibility that the virus originated elsewhere and that the crowded market only contributed to the spreading of the epidemic. Clearly, wide sampling of coronavirus diversity in different mammals in China and neighboring countries will be necessary to track down the progenitor(s) of SARS-CoV-2.

The scenarios described above implicitly imply that, whatever the source of the zoonotic transmission, SARS-CoV-2 emerged and adapted in another host, eventually spilling over to humans (Andersen et al. 2020). It cannot however be excluded that, before being recognized in December 2019, SARS-CoV-2 had been circulating for a while in humans, maybe causing mild symptoms. The acquisition of some key viral mutations during unrecognized human transmission might have then fostered the current pandemic strain, characterized by sustained human-to-human transmissibility and virulence. In this case, the identification of such mutations and the assessment of the role of the above-mentioned furin cleavage site insertion would be of paramount importance. It should nonetheless be kept in mind that extreme caution should be exerted when ascribing phenotypic effect to mutations that arise during viral spreads (Grubaugh et al. 2020). Here, again, MERS-CoV and SARS-CoV have

lessons to teach. Whereas in the case of SARS-CoV, changes in the spike protein RBD contributed to the adaptation of the virus to human cells (W. Li et al. 2005); (Wu et al. 2012); (Qu et al. 2005), MERS-CoV adaptation to our species occurred with limited changes in the RBD (Cotten et al. 2014; Forni et al. 2015). However, during an outbreak of MERS-CoV in South Korea, viral strains carrying point mutations in the spike RBD emerged and spread (Y. Kim et al. 2016). Notably, whereas these RBD mutations were found to decrease rather than increase binding to the cellular receptor, they facilitated viral escape from neutralizing antibodies (Y. Kim et al. 2016; Y.-S. Kim et al. 2019) (Rockx et al. 2010; Kleine-Weber et al. 2019). This suggests that changes in the spike protein do not necessarily arise as an adaptation to optimize receptor-binding affinity.

Even more emblematic is the case of SARS-CoV ORF8, encoding an accessory viral protein. In the early stages of the epidemic, SARS-CoV strains acquired a 29-nucleotide deletion in ORF8 (Chinese SARS Molecular Epidemiology Consortium 2004). Together with the observation that the encoded protein is fast evolving in SARS-CoV strains, this finding was taken to imply that deletions in ORF8 were driven by natural selection and favored human infection (Lau et al. 2015). The evidence for adaptation was subsequently not confirmed, and recent data indicate that the 29-nucleotide deletion most likely represents a founder effect causing fitness loss in bat and human cells (Forni et al. 2017; Muth et al. 2018). This observation clearly indicates that a mutation sweeping at high frequency in the viral population does not necessarily represent a selectively advantageous change.

Clearly, huge gaps remain in our understanding of SARS-CoV emergence and adaptation to our species. Most likely, acquisition of inserts in the spike protein and remodeling of the RBD did not occur in a single animal-to-human jump event. The majority of relevant changes might have been present in a reservoir/intermediate species and only minor changes may then have been required to gain full transmissibility in humans.

Alternatively, the progressive optimization from multiple zoonotic events may have followed short events of human-to-human transmission over an extended period of time. This has been earlier observed during MERS transmission, entailing repeated jumps of MERS-CoV from camels (Cui, Li, and Shi 2019). Accumulation of mutations in NTD and RBD during global and local transmissions further strengthens the likelihood that the progenitor virus may have

been transmitted among humans over time, adapting to become more efficiently transmitted in people.

In a nutshell, the missing knowledge about the reservoir/intermediate host of SARS-CoV-2, as well as about the extent of its distribution in wild and domestic animals, raises concerns about zoonotic transmissions. If pre-adaptation routinely occurs as these viruses incubate in animal reservoirs, there is a high probability for new pandemics to recur. Thorough surveillance of animal populations and early diagnosis in people will be necessary to prevent future COVID-19 like pandemics.

Ecological factors

Zoonotic spillover of infectious pathogens is threatening socioeconomic development and public health worldwide (Jones et al. 2008). The emergence and rapid worldwide propagation of the SARS-CoV-2 coronavirus has led to the COVID-19 pandemic. The chain of events that facilitated the zoonotic spillover likely required the alignment of ecological, epidemiological and behavioral determinants that allowed a precursor lineage of a bat virus to trespass a series of barriers, possibly after establishing infections in intermediate hosts, before evolving to become capable of efficient human-to-human transmission. The mechanisms involved in the spillover and evolution of this coronavirus as a human pathogen remain unclear, including contrasting information regarding the immediate reservoir or other hosts, including horseshoe bats (*Rhinolophus sinicus*) or the ant and termite-feeders, critically endangered pangolin species (*Manis javanica*). The latter is the most trafficked animal worldwide and illegally imported from Malaysia into China for trade in markets (J. F.-W. Chan et al. 2020; Huang et al. 2020; Xu 2020; Zhou et al. 2020). Two lineages related to that of SARS-CoV-2 were identified in Malayan pangolins trafficked into China, showing that these animals can potentially be involved in the zoonotic transfer to humans (T. T.-Y. Lam et al. 2020).

Two scenarios have been proposed to explain the cross-species transfer and evolution of the new betacoronavirus to a human-to-human transmission. The precursor of the SARS-CoV-2, cycling in an animal host before the zoonotic transfer, might have evolved the bind to an

ACE2 receptor similar to that of humans (Andersen et al. 2020). For this to have occurred, a high host population density would have been needed to achieve efficient natural selection for this important phenotype. The second scenario proposed by Andersen and colleagues hypothesizes that natural selection for this trait occurred in humans after the virus established itself in human beings, continuously adapting before transmission caught the attention of the medical and public health communities. In either scenario, the new coronavirus emergence required a combination of successful mutations that enabled the virus to cause infection in humans, and ecological and behavior determinants that enabled the virus to breach host species barriers and then spread in large and dense human populations before being recognized as the agent of a potential pandemic. Such an alignment of genetic, ecological, epidemiological, and behavioral factors was proposed by Geoghegan and Holmes (2017) and Plowright et al. (2017) to explain virus disease emergence at the human-animal interface.

Zoonotic transfer of pathogens to humans has been attributed to changes in natural environments for expanded global food production (Rohr et al. 2019; Wolfe, Dunavan, and Diamond 2007), climate change (Zinsstag et al. 2018), habitat degradation (Allen et al. 2017), biodiversity loss (Felicia Keesing et al. 2010; F. Keesing, Holt, and Ostfeld 2006), wildlife trade (Smith et al. 2009), and changes in the distribution and prevalence of host reservoirs and their parasites (Vanwambeke, Linard, and Gilbert 2019). In fragmented forest landscapes, generalist reservoir species are capable of adapting to new ecological conditions, becoming more abundant, expanding their distribution range, and accumulating parasites and pathogens. In this scenario, the host species richness and high population density increase the human-animal reservoir contact rate and the exposure to parasites (Borremans et al. 2019; Johnson et al. 2020). In addition, the close proximity of people to forest ecosystems, increased human population density, poor living conditions, low socioeconomic condition and human mobility facilitate human exposure to wild animals and the probability of spillover events (Geoghegan and Holmes 2017; Wilkinson et al. 2018).

The probability of a spillover event establishing new human infection is influenced by (1) the pathogen pressure (that is the amount of pathogen available at a given time and space), (2) human and reservoir host behavior (mediating the risk of human exposure to the pathogen),

and (3) factors linked to human susceptibility to infection, together with parasite dose and route of exposure (Plowright et al. 2017).

Decline in wildlife populations caused by predatory hunting activities, decreased habitat quality and habitat loss driven by extensive deforestation has been linked to increased probability of zoonotic virus spillover to humans at the human-animal interface (Geoghegan and Holmes 2017). In these human-dominated landscapes, primates and bats are reservoir hosts of more viruses than other mammal species, increasing the potential for new infections in humans to become established (Johnson et al. 2020). The potential risk of SARS-CoV spillover from horseshoe bat population to humans was demonstrated by Menachery et al. (2015) using a reverse genetics system entailing a chimeric virus expressing the bat SHCO14 in a mouse-adapted backbone. This study illustrated scenarios for the emergence of bat SARS-CoV in humans: infection of an intermediate nonhuman host might be followed by human infection. Direct bat-human transmission could be followed by selection in the human population (distinct from closely-related viruses circulating in the source host). In a third scenario, the circulation of quasi-species pools in the animal reservoir might maintain multiple virus strains, some of which capable of causing infection in humans without the need for additional mutations. Alphacoronaviruses and betacoronavirues were identified in free-ranging bats from Myanmar, showing the potential for zoonotic virus emergence in humans in close contact with sylvatic animals in forest areas disturbed by ongoing process of changes in land use (Valitutto et al. 2020).

Climate variability anomalies

Climate variability is known to affect the outbreaks of many infectious diseases (Morand et al. 2013). Vector-borne diseases such as Murray Valley encephalitis, Rift Valley fever, Ross River virus disease and dengue, among many others, have all been linked with anomalies in climate El Niño Southern Oscillation (ENSO) (Anyamba et al. 2019; Nicholls 1993). Two emerging viral diseases transmitted by bats have been found associated with El Niño events: Hendra virus in Australia (McFarlane, Becker, and Field 2011; Giles et al. 2018) and Nipah virus in Malaysia (Daszak et al. 2013). The recent emergence of SARS-CoV-2 has also followed an important El Niño event (NOAA 2019, https://www.climate.gov/news-features/blogs/enso/august-2019-el-ni%C3%B1o-update-stick-fork-it), which has particularly affected China.

In Asia and the Pacific, eight new viruses transmitted by bats have emerged in humans and livestock since the 1990s (Table 1). Eight of these nine outbreaks of newly emerging bat-borne diseases appear associated with El Niño - La Niña events using values of 'NINO 3.4' index (Figure 2) retrieved from the National Oceanic and Atmospheric Administration (NOAA, https://www.noaa.gov). Four bat-borne viruses have emerged during an El Niño phase and four during a La Niña phase according to the monthly classification of the ENSO provided by NOAA. Only Kampar virus has emerged during a neutral phase, although following closely a La Niña event.

**** Place Table 1 about here ****

This observation, although needing more in depth analysis, suggests that viral diseases transmitted by bats seem likely driven by ENSO climatic anomalies. Abnormal rainfall, temperature, and vegetation development, whether above or below normal condition during El Niño - La Niña events, are known to create appropriate ecological conditions for pathogens, their reservoirs and vectors that may enhance transmission, risk of spill-over, emergence, and propagation of disease clusters (Anyamba et al. 2019). Stresses induced by climate variability may have a profound effect on disease dynamics in wild animal populations, mostly in relation to immune or behavioral changes (Subudhi, Rapin, and Misra 2019).

Figure 2

Then, climate anomalies by their effects on food shortage, behavioral mobility, and modulation of the immune system of bats are likely to increase the risks of disease emergence by putting them in contact with other animals, wild or livestock, and favoring viral spillover. A condition that has prevailed in 2019 for SARS-CoV2.

Lessons from Molecular Epidemiology

We know by now that SARS-CoV2 transmission has occurred by droplets through human to human contact similar to SARS and MERS (Han et al. 2020). Spread by aerosols is also

possible but it is still under investigation. Some studies have reported the onset of gastrointestinal symptoms in patients upon admission and some reported detection of virus nucleic acid in fecal samples of patients. SARS-CoV-2 RNA was also detected in oesophagus, stomach, duodenum and rectum specimens for both two severe patients. In contrast, only duodenum was positive in one of the four non-severe patients. These findings suggest a possible oral-fecal transmission route of infection (L. Lin et al. 2020; Hindson 2020).

Tracking changes in the virus has shed bright light on its global spread. Indeed, viral sequencing proved key in establishing transmission where only imported cases from travelers had been suspected. How can we be sure? We briefly overview the methods that sustain these claims and why we can be confident that they, and existing evidence, more than suffice to teach us important lessons about this virus' global spread.

Evolutionary trees (phylogenies) illustrate relationships among biological lineages. The tools that build such trees "work backwards" from existing sequence data, inferring the genealogical or phylogenetic relationships among them, and reconstructing which and when changes arose.

There a few basic methodologies with a plethora of different implementations which are beyond the scope of this review to detail (Boussau and Daubin 2010; Holder and Lewis 2003). Nevertheless, the currently most popular and reliable methods are based either on maximum likelihood or Bayesian inference. Some of these methods can incorporate temporal information, at the tips and/or at internal nodes which, along with some restrictions on constance of the evolutionary rates, global or local, can be used to date particular nodes, including the root or Most Recent Common Ancestor (MRCA) of the sequences in the tree. Differently from most other phylogenetic analyses, the time of sampling of most sequences derived in the molecular epidemiology analyses of SARS-CoV-2 is known. This allows the use of time-stamped phylogenetic trees (Neher and Bedford 2018). In these, the purely phylogenetic information is enriched with sampling time data which improves the use of genetic information from pathogens to track their spread in time and space. When the analyses include hundreds or even thousands of sequences we face a problem in analyzing the data and visualizing the results. Some convenient, easy-to-use solutions have been provided by the NextStrain Platform (https://nextstrain.org). A

complete description of the tree building method used in NextStrain is available elsewhere (Hadfield et al. 2018), but most readers need only understand that millions of alternative trees are considered and only the most compatible with the data is represented. Nothing guarantees this tree to be correct in every detail because other, similar trees enjoy nearly equivalent support. .

To fully appreciate how successfully the recent evolution of SARS-CoV-2 has been captured in the NextStrain platform, it is worth first reflecting on the strengths and weaknesses in the time-stamped tree. Figure 3 describes the changes that SARS-CoV-2 viruses underwent since COVID-19 became widely recognized as a viral pathogen in December, 2019, through mid-April, 2020. It depicts relationships among hundreds of viral sequences, sub-sampled from the thousands of sequences available at the time of writing, as rendered for real-time evaluation by the Nextstrain consortium (https://nextstrain.org/ncov) using data from GISAID.org.

Figure 3

One feature establishes, without question, the usefulness of this tree: contemporaneously sampled viruses (each depicted by its own color) are perfectly chronologically organized. In this rendering, each sample (circle) was colored to indicate when it was sampled. The "rainbow" pattern means that each virus has undergone a nearly equal amount of change. Note, too, that the viruses do not "bunch up" at the tips of the tree, but are situated throughout. We need not guess when and where ancestors lived, because we can see them. Lastly, no branches protrude to any great extent, eliminating the fear of being misled by tree artifacts introduced by recombination. Although recombination contributed to the origins of this viral group (Ji et al. 2020; Boni et al., n.d.), and future recombinations cannot be discounted, this does not appear to have been a major force over this five month period. By so faithfully reproducing these samples' known chronology, this tree should earn our trust.

The total number of points at any column (encompassing different time intervals) informs about the sampling success at those times. There are only a few sequences from the early weeks (before January 15) followed by a clear increase until the beginning of February

(around February 5). Then, Chinese researchers (see below) almost completely ceased to deposit sequences in GISAID and only a few sequenced were being obtained from other countries. This period corresponds to the silent spread of COVID-19 throughout most countries. Asymptomatic infections were not detected, hence no virus samples were available for sequencing and we lack detailed information on how the virus was spreading. This period also corresponds to a bout of diversification that can be observed in the upper part of the tree, with many new sublineages stemming from pre-existing lineages. The times of these diversification events are inferred from the application of a constant rate of evolution of 8 × 10^-4 subs per site per year. The constancy of this evolutionary rate is difficult to be evaluated at this stage in the evolutionary history of the virus, given the very short time elapsed since its appearance. By the end of February, the situation changed drastically, as many more sequences became available for analysis when the pandemic exploded and cases accumulated rapidly in many countries.

See below the same tree (Figure 4), now colored to show each sample's geographical origins (as shown on the accompanying map). Note that isolates from China dominate the oldest branches of the tree. The first isolates sequenced came from Wuhan, China (where the epidemic was first recognized); therefore, it is logical that they should appear first.

Figure 4 about here ****

What other epidemiological inferences can we draw from tracing the path of this virus' evolution?

Human infections may have commenced months before the recognized outbreak in Wuhan. Most human pathogens originate as pathogens of other animals; they occupy a spectrum, emerging from those exclusively transmitted among animals to those primarily transmitted among other animals, to those that have lost their dependence on non-human reservoirs and, ultimately, those no longer capable of being transmitted excepting from person to person (Wolfe, Dunavan, and Diamond 2007). The human drama now unfolding began, in earnest, in the latter months of 2019; the tree above records the virus's

subsequent proliferation and spread. But, what can be said about preceding events?

Serological evidence suggests that people living the vicinity of bats harboring SARS-like viruses are frequently exposed to such infections (Wang et al. 2018). More surveillance will be needed to truly understand how often, and under what circumstances, such exposures take place. Nor can we yet appreciate what proportion of such exposures eventually lead to a virus becoming established as a human pathogen.

Acknowledging this broader context, what do the phylogenetic data say about the proximate origins of SARS-CoV-2? More surveillance among the viruses of suspected animal reservoirs (bats, most especially) will be needed to confidently estimate when and where this virus first took a firm foothold in people. Early reports (T. T.-Y. Lam et al. 2020; Zhou et al. 2020; "A Novel Bat Coronavirus Closely Related to SARS-CoV-2 Contains Natural Insertions at the S1/S2 Cleavage Site of the Spike Protein" 2020) identified BatCoV RaTG13 and RmYN02 as the closest isolates to SARS-CoV-2 and it has been proposed that these lineages diverged between 40 and 70 years ago (Boni et al., n.d.). Assuming constancy in the evolutionary rate, which is also problematic as discussed above, a few months may have elapsed between human acquisition of the virus and its recognition by the medical and public health communities in December 2019. Molecular clock analyses suggest a time for the most recent common ancestor of all current SARS-CoV-2 lineages in October-December 2019 (van Dorp et al. 2020) If further sampling identifies more genetically proximate enzootic viruses, then the estimated times for separation and diversification may likewise change. Lacking such additional sampling, it seems fair to suspect from the available evidence that an exclusively human chain of transmission may have gone undetected for some months before galvanizing public attention.

Some lineages **took root outside of China.** Note that clusters of viruses define foci of transmission. Some of these represent expansions outside of China. For example, a group almost exclusively composed of red dots (near the bottom of the figure) indicate a viral lineage that took root in the United States. These mostly came from the state of Washington (though Wyoming, Virginia, and other states are included in the group). Within this clusters, a few dots different to red indicate likely exports from the USA to other countries, such a Australia, Taiwan, Iceland and India. Nevertheless, inferences about epidemiological

processes drawn exclusively from genome information must be taken very cautiously (Villabona-Arenas, Hanage, and Tully 2020)

Long-distance travel spread closely related viral isolates. Although most of the isolates in the group discussed above were sampled from patients in the American Pacific Northwest, a few blue dots in that group denote Australian isolates, demonstrating air travel as a powerful disseminating force. Specific introductions [i.e., from China to Germany, from Germany to Italy, from Italy to Iceland (Gudbjartsson et al. 2020)] that accord with known travel histories are recapitulated in phylogenetic networks reconstructed from viral RNA sequence. For example, the phylogeny accords with other evidence that the virus was introduced to Europe by early January (Olsen et al. 2020; Spiteri et al. 2020; Stoecklin et al. 2020).

Community transmission occurred weeks before local outbreaks were recognized. Early efforts in the United States to manage the virus focused on limiting travel from suspected endemic regions, and early testing was limited to travelers returned from such places. Meanwhile, the virus had already established footholds, as discovered when nasal swabs collected for monitoring community transmission of influenza were discovered to harbor the RNA of SARS-CoV-2. Nearly identical viruses were found in residents of Seattle who had no direct contact with one another. This sounded the alarm that the virus was actively circulating in the community, a fact that would soon be tragically borne out by a critical surge in illness and death (Bedford et al., n.d.; Fauver et al. 2020).

Multiple, independent introductions of the virus. A parallel process was soon to play out in New York. Based on the viral phylogeny, it appears the virus was established in New York at least seven times, primarily from Europe, while travel restrictions focused on perceived threats from China, South Korea, Italy and Iran (Gonzalez-Reiche et al., n.d.). The occurrence of red dots in nearly every clade of the tree (even clades composed mostly of European isolates) underscores this point. This is equally true for Europe, because nearly every clade of the tree includes green dots (denoting samples of European origin). This tree includes relatively few isolates from Australia (blue dots), but the same phenomenon is evident: Australia received virtually every basic type of SARS-CoV-2.

The Global South has not been sampled nearly enough. A complete description of the virus

evolution would necessitate vast and representative sampling. With diagnosed cases (as of mid-April) numbering in the millions, and in the face of untold numbers of undiagnosed cases, how much confidence should we place in trees reconstructed from only hundreds or thousands of isolates? Biologists have means to assess internal consistency, enabling many facets of the recorded history to be viewed by now as impervious to egregious sampling error. Nonetheless, it is clear that the available evidence neglects to inform us much about the epidemic. This is for lack of attention, not for lack of virus. Even in the most affluent of societies, the virus is exploiting preexisting social inequities (Yancy 2020). As the virus takes root in Africa, South America, and other under-resourced regions, evolutionary epidemiologists should assume responsibility for understanding and communicating all that can be learned about how, where, when, and why the pandemic is unfolding there. The toll of this disease will be multiplied where patients cannot access therapies and where clinicians cannot protect themselves. There, the behavior of the virus may or may not parallel patterns so far described. We will not know if we do not look.

Diagnostic Testing for SARS–CoV-2/COVID-19

SARS-CoV-2 was first detected in human bronchoalveolar lavage (BAL) specimens by unbiased Illumina and nanopore sequencing technologies, and a real-time RT-PCR assay using pan-betacoronavirus degenerate primers targeting a highly conserved RdRp region (N. Zhu et al. 2020). The SARS-CoV-2 isolation was performed in different cell lines, including human airway epithelial cells, Vero E6, and Huh-7. Cytopathic effects were observed in human airway epithelial cells 4 days after inoculation, but not in Vero E6 and Huh-7 (N. Zhu et al. 2020). For safety, CDC recommends that clinical virology laboratories should not attempt viral isolation from clinical specimens collected from COVID-19 patients under investigation (PUIs).

Because SARS-CoV-2 is a newly discovered virus and its genome is very divergent from those of HCoV-229E, -NL63, -OC43, and -HKU1, several commercially available multiplex NAAT tests (BioFire FilmArray Respiratory Panel, ePlex Respiratory Pathogen Panel, NxTAG Respiratory Pathogen Panel, RespiFinderSmart22kit, and others.) for the detection of respiratory organisms in clinical virology laboratories were predicted no cross-reactivity with

SARS-CoV-2 (Phan 2020). At that time, several different protocols for laboratory-developed tests (LDTs) were developed, and they were available on the WHO website (https://www.who.int/who-documents-detail/molecular-assays-to-diagnose-covid-19-summary-table-of-available-protocols). While the CDC assay was designed for specific detection of three different regions of the N gene, the assay developed by Corman et al. (2020) aimed to amplify three different regions of the RdRp, N and E genes.

Since SARS-CoV-2 continues to spread around the world, and the number of potential cases increases rapidly, faster and more-accessible testing is extremely needed. Thus, many companies are racing against the clock to develop commercial test kits to detect SARS-CoV-2 more quickly and accurately. More than fifty RT-PCR-based diagnostic tests are currently available in the US that have been granted the FDA's Emergency Use Authorization (https://www.fda.gov/medical-devices/emergency-situations-medical-devices/emergency-use-authorizations#covid19ivd). These tests can qualitatively identify SARS-CoV-2's RNA in the lower respiratory tract (bronchoalveolar lavage, sputum, and tracheal aspirate), and upper respiratory tract (nasopharyngeal and oropharyngeal swabs). Recently, the FDA approved emergency use for a portable, fast, swab test for SARS-CoV-2 which can provide results in less than 15 minutes. The IDNOW COVID-19 (Abbott, Illinois) can be used at a point-of-care in doctor's offices, urgent care and hospitals. Lateral flow immunoassays are another rapid, point-of-care diagnostic test, which has been widely used. The Sofia 2 SARS Antigen FIA is the first COVID-19 antigen test to be granted the FDA's Emergency Use Authorization. This lateral flow immunofluorescent sandwich assay is used with the Sofia 2 instrument intended for the qualitative detection of the nucleocapsid protein antigen from SARS-CoV-2 in nasopharyngeal and nasal swabs.

Another diagnostic approach would be to devise blood tests for antibodies against SARS-CoV-2. Many companies around the world have raced to develop antibody tests. The qSARS-CoV-2 IgG/IgM Rapid Test (Cellex Inc.) was the first antibody test being approved by the FDA. This test is intended to qualitatively detect IgG and IgM antibodies against the SARS-CoV-2 in human serum, plasma, and whole blood. Together with increased availability of commercial RT-PCR-based diagnostic tests, important questions about immunity to the novel coronavirus will be answered soon. The use of serological tests is complementary to those based on direct detection of viral RNA because they indicate that the individual has

developed specific immunological response to the virus, which usually takes a few days after infection; hence, a negative result does not necessarily mean that a person is not infected - this might have occurred too recently to have developed the immune response, and a positive results does not necessarily mean that the person has an active infection. Consequently, these tests are essential to obtain a more precise estimate of the total number of infections in a population and in which proportion they are expected to be immune to a future infection by the virus. Its main use is in epidemiological surveillance rather than in clinical diagnostics. Early application of community surveys for seroprevalence have suggested far more widespread community exposure than has been supposed based merely on the reporting of diagnosed cases, suggesting for example that by mid-April as many as one in five residents of New York City may have been exposed, with diminishing exposure rates less proximate to the outbreak's epicenter.

Recently, clustered regularly interspaced short palindromic repeats (CRISPR) based nucleic acid detection technology has emerged as a powerful tool with the advantages of rapidity, simplicity and a low cost (X. Wang et al. 2020). A recent study reported that the CRISPR-based assay was used to detect SARS-CoV-2 in the patient's pharyngeal swab in China (Ai et al. 2020). The turn-around time (TAT) of CRISPR was 2 hours, much faster than RT–PCR (3 hours) and mNGS (24 hours) (Ai et al. 2020). The FDA authorized a COVID-19 test that uses the gene-editing technology CRISPR. Sherlock CRISPR SARS-CoV-2 Kit has been developed by Sherlock Biosciences Inc.

Human susceptibility to SARS-CoV-2

Knowledge of genetic variation, at both individual and population levels, may further our understanding of disease transmission and pathogenesis, enabling identification of individuals at high risk of infection and sequelae. More broadly, this may inform drug design and vaccine development. However, previous coronavirus epidemics of recent years, namely SARS-CoV in 2003 and MERS-CoV in 2012, did not lend themselves to such evaluation owing to their rapid progression, acute onset, and relatively small reach. The current SARS-CoV-2 pandemic appears unfettered by any population-level protective immunity to cross-reactive epitopes. The much greater number of individuals affected, together with the availability of

high throughput technologies, will stimulate rapid research on variation in host susceptibility. Indeed, the number of preprints appearing online, prior to peer review, is unprecedented. Other respiratory viruses where host susceptibility has received attention include Respiratory Syncytial Virus (RSV) (Tahamtan et al. 2019) and Influenza A (Nogales and DeDiego 2019; Clohisey and Baillie 2019). However, diseases caused by these viruses are sufficiently different that inferences concerning the role of genetics with respect to coronavirus infection cannot be directly applied.

Genetic control of susceptibility and approaches to study

Although host genetics is likely to be important for the outcome of infection, because of the nature of the disease, we only have preliminary information on the heritability of susceptibility to coronavirus infections, i.e. estimates of the proportion of phenotypic variability due to genetic factors. Heritability, a time and place-specific measurement for a population, provides an indication of likely success in the hunt for susceptibility loci, and low values for heritability of <20% would indicate that very large sample sizes might be needed to detect genetic effects. Williams et al. (2020) have reported high heritability, in the range 40-50%, of COVID-19 symptoms in a classical twin study (Williams et al. 2020). In addition, the apparent differences in susceptibility seen between ethnic groups could be due, in part, to human genetic variability. Niedzwiedz et al. (2020) report a higher risk of confirmed SARS-CoV-2 infection, not accounted for by variables such as socioeconomic differences, in some minority ethnic groups as included in the UK Biobank (Niedzwiedz et al. 2020). The genetic contribution to the outcome of infection will be complex, since some of the underlying health conditions associated with death from coronavirus infection are known to have a significant genetic contribution.

The difficulties of family-based studies, whilst controlling for exposure, leaves us with case control association methodology. Ultimately, the high numbers of affected people will facilitate genome wide association studies (GWAS), or even studies of whole genome or exome sequence, particularly where genotypic information is already available, as is the case for UK Biobank participants. In the short term, we are likely to see a number of candidate

gene studies. For susceptibility to coronavirus infection *per se*, loci coding for viral receptors provide obvious candidate genes. Nevertheless, the major question is why some individuals develop life-threatening immune-mediated pathologies, where knowledge of genetic variation may contribute to our understanding of key mediators of the immune response.

Most relevant, albeit limited, information on the allelic diversity contributing to susceptibility/resistance to date has used sampling from the SARS CoV 2003 epidemic, and a little information has been gleaned from the MERS-CoV 2012 epidemic. Sample sizes are small, and results sometimes conflicting, hence no firm conclusions can be drawn. There are several studies identifying genes of interest from mouse models (Kane and Golovkina 2019) and some information from interspecies comparisons (Hou et al. 2010; Sironi et al. 2015). To date, candidate genes for published human studies have included pathogen receptor genes and loci controlling innate and adaptive immunity.

Receptors

S1 spike proteins of SARS-CoV and SARS-CoV-2 bind to angiotensin converting enzyme 2 (ACE2) on cell membranes, catalysing the cleavage of the vasoconstrictor angiotensin II and countering the activity of ACE. MERS-CoV binds to the dipeptidyl peptidase 4 (DPP4, CD26) an immunoregulatory serine exopeptidase. Particularly considering indications of relevance from cross-species comparisons (Sironi et al. 2015), there has been a surprising lack of interest in these 2 candidate loci for susceptibility to infection *per se.* A small, low-powered, case control study, with information on anti-SARS-CoV antibody status, did not show any associations between SARS phenotypes and *ACE2* polymorphisms in a Vietnamese population (Itoyama et al. 2005). Genes coding for functionally associated molecules such as transmembrane serine protease 2 (*TMPRSS2*), which cleaves and activates viral spike glycoproteins, are also worthy of study (Iwata-Yoshikawa et al. 2019). Lopera et al. (2020) have just reported a PheWAS of 178 quantitative phenotypes, including cytokine and cardio-metabolic markers but not specific SARS-CoV-2 infection markers, in relation to *ACE2* and *TMPRSS2* variation (Lopera et al. 2020).

MHC

Amongst immune response related loci, MHC class I and class II allelic associations are to be expected, particularly through MHC class I restriction of CD8+ T cells (M. Lin et al. 2003; M. H. L. Ng et al. 2004; S.-F. Wang et al. 2011; Keicho et al. 2009). MHC associations are relevant for susceptibility to disease *per se*, disease pathogenesis and response to vaccination. There are a few follow up studies e.g. of *HLA A**0201 restricted SARS-CoV epitopes for CD8+ T cells (Zhao, Wang, and Wu 2011), but investigation of such highly polymorphic loci needs to be more comprehensive.

Other loci

Other loci, with some information on allelic associations, include those coding for ICAM3 (DC-SIGN, CD209) (Kelvin Y. K. Chan et al. 2007; K. Y. K. Chan et al. 2010) and DC-SIGNR (CD209L) (H. Li et al. 2008) and widely studied loci such as *MBL* (Zhang et al. 2005; Ip et al. 2005) and *CD14*. In particular, the inflammatory response associated with SARS suggests a number of candidate loci e.g. *AHSG* (X. Zhu et al. 2011), *IFNG* (Chong et al. 2006), *CD14* (Yuan et al, 2007) and *CCL5* (Rantes) (M. W. Ng et al. 2007). Nevertheless, some relatively small studies have resulted in some conflicting findings being noted e.g. for MBL (Yuan et al, 2005) and DC-SIGNR (H. Li et al. 2008).

And from mice

More recently, loci of interest have been identified using mouse models, after infection with SARS-CoV, where pathology can be well studied. These include *Trim55* and *Ticam2* (Kane and Golovkina 2019). *Trim55* codes for an E3 ubiquitin ligase present in smooth muscle around blood vessels, affecting lung pathology by controlling airways and immune cell infiltration. Deficiency was relevant to lung injury although susceptibility alleles were not reported (Gralinski et al. 2015). *Ticam2* knockout mice were highly susceptible to disease with some evidence of allelic heterogeneity. Ticam2 is an adaptor for MyD88-independent TLR4 signaling contributing to innate immunity (Gralinski et al. 2017). These genes require

complementary studies in human populations.

Choice of phenotypes and genotypes

To date, phenotypes employed for human genetics have been limited: susceptibility to infection *per se,* some measures of morbidity and mortality. For both candidate gene studies and GWAS, finding totally 'resistant' individuals, for an unaffected control group, is hard since for COVID19, despite the occurrence of asymptomatic individuals, many show very mild symptoms of disease. There is scope for development of phenotypes, particularly relating to the most pertinent issues of pathogenesis and control of immune responsiveness. Perhaps immediate research could focus on critical and potentially useful phenotypes e.g. severity of disease sufficient to require a ventilator, or antibody responses.

Identification of a susceptibility gene with no a priori hypothesis is difficult for an active acute viral infection, but candidate genes, potentially controlling these phenotypes, would include those from both innate and adaptive immunity. Published GWAS, which might be of interest for suggesting susceptibility loci, tend to have used phenotypes from chronic rather than acute conditions e.g. COPD. Whether conducting genotype/phenotype studies or simply measuring immune responses, age and sex are clearly major factors in the outcome of infection with SARS-CoV-2, and will require careful consideration in all analyses.

Immunity and ageing

Immunity and pathology will be influenced by many factors including genetics, age, sex, viral dose and underlying health problems. The interactions between these factors will also be important (Bupp et al. 2018). COVID19 particularly affects the elderly, with age a key consideration in any study design, as are sex differences, which are known to be important for respiratory infections in general (Kadel and Kovats 2018). Although both innate and adaptive immunity are required to mount an effective immune response, and despite the tools for investigation of the relevant loci being available, the lack of candidate gene studies for coronavirus infection even prior to COVID19, including information on MHC associations,

has been noted. Not only will variability in amino acid sequence be important, such as that influencing MHC restriction, but also the levels of immune mediators, possibly driven by promoter variation. Timing of responses is another area for exploration. Rockx et al. (2009) discuss molecular mechanisms of age related susceptibility in coronavirus infected mice, and development of pathology via the host response. Aged mice showed a greater number of differentially expressed genes when infected than young mice, including elevated interferon and cytokine genes, indicating a different host response kinetics in old versus young mice.

A number of immune response molecules/genes are of interest, including Toll-like receptors (TLRs), proinflammatory cytokines and the appropriate signalling pathways. Both adaptive and innate immunity are recognized to decline with age (Goronzy and Weyand 2013; Nikolich-Žugich 2018), and although, in the past, adaptive immunity has been the focus of attention, there is now cumulative evidence for a decline in innate immunity with age. Changes occur in neutrophil, macrophage, dendritic cell and NK populations, and there are changes in TLR expression, with an overall distinct but not uniformly lower expression in later life, particularly in macrophages, and consequent major changes in TLR mediated responses (Solana et al. 2012). The innate immune system of the frail elderly is often in a state of heightened inflammation, but this higher basal production of proinflammatory cytokines does not necessarily translate to a generally elevated and/or effective response (Fulop et al. 2018). In particular, the incidence of and mortality from lung infections increase sharply with age, with such infections often leading to worse outcomes, prolonged hospital stays and life-threatening complications, such as sepsis or acute respiratory distress syndrome. The review by Boe et al (2017) covers research on bacterial pneumonias and pulmonary viral infections and discusses age-related changes in innate immunity that contribute to the higher rate of these infections in older populations.

Regulation of adaptive immune function appears to be diminished in old versus young adults, and decreased TLR responses are associated with the inability to mount protective antibody responses to flu vaccine (Kollmann et al. 2012). In addition, the pre-existing B cell repertoire restricts the quantitative response in the elderly (Goronzy and Weyand 2013). B

cell maintenance and function in aging has been well reviewed and is of particular relevance when considering responsiveness to vaccines evoking primary versus memory humoral responses (Kogut et al. 2012). The age-related rise in proinflammatory cytokines is also associated with reduced response to vaccination (Oh, Lee, and Shin 2019). It is known that vaccination with TLR5 targeting adjuvants in elderly patients enhances flu vaccine responsiveness without increasing inflammation (Goronzy and Weyand 2013).

We know multiple host factors, including age and sex, are important for disease outcome of coronavirus infection and a systems approach has been suggested for the study of disease pathogenesis (Schäfer, Baric, and Ferris 2014). However, to date, we do not know the relative importance of factors such as host and pathogen genetic variability. Whilst we can draw some comparisons with other similar acute respiratory infections such as influenza, SARS-CoV-2 presents unique and urgent challenges. There is an unprecedented opportunity to obtain large sample sizes and the need to understand the role of human genetics in the outcome of infection with SARS-CoV-2 is well recognized (Kaiser 2020). Information on infection with SARS-CoV-2 is being added to health data e.g. inclusion in the UK Biobank with its 500,000 volunteers (www.ukbiobank.ac.uk). The scale of the current pandemic is stimulating free access to human genetic studies as they progress and data sharing for combined analyses, with multiple, diverse partners from academia and industry e.g. The COVID19 Host Genetics Initiative (www.covid19hg.org). Hopefully, useful information will be produced at a rate not previously seen for a transmissible disease.

Host response and considerations for vaccine development

There is limited knowledge on the mechanisms driving immune response against COVID-19. A recent work (Thevarajan et al. 2020) showed that a strong adaptive response across B and T cells was mounted during symptomatic phase in a 47-year-old woman from Wuhan, Hubei province, China, that successfully cleared the virus within two weeks. Increased antibody-

secreting cells (ASCs), follicular helper T cells (TFH cells), activated CD4+ T cells and CD8+ T cells and immunoglobulin M (IgM) and IgG antibodies that bound the COVID-19-causing coronavirus SARS-CoV-2 were detected in blood before symptomatic recovery and with a peak around day 7-9 since onset of symptoms. These immunological changes persisted after full resolution of symptoms at day 20. These findings, based on a single subject, support the hypothesis that a rapid multi-factorial immune response against COVID-19 can be mounted within a week and with evidence of recruitment of T and B cells as well as macrophages before resolution of symptoms, and that this rapid response may correlate with positive clinical outcome.

Severe cases, on the other hand, seem to have disrupted immune responses. Several early case reports have demonstrated the presence of cytokines release storms, which demonstrate the presence of severe inflammatory responses in these patients characterised by increased interleukin (IL)-2, IL-7, granulocyte-colony stimulating factor, interferon-γ inducible protein 10, monocyte chemoattractant protein 1, macrophage inflammatory protein 1-α, and tumor necrosis factor-α.6 (Mehta et al. 2020).

Immunogenomics analysis of Bronchoalveolar Lavage Fluid (BALF) and also from blood samples across two different studies from China (Wen et al., n.d.; Liao et al., n.d.) also revealed clear immunological features characterizing severe cases. The overall message based on these early studies point towards the polarization of a macrophage response in the lungs associated with severe disease. Conversely, patients that recovered had increased cytotoxic T cell response compared to severe cases. Specifically, analysis of bronchoalveolar lavage fluid (BALF) from 3 severe and 3 mild COVID-19 patients using single-cell RNA sequence (scRNA-seq) combined with T cell receptor sequences revealed a distinct population of monocyte-derived FCN1+ macrophages in BALF samples from patients with severe disease, while patients with mild disease presented increased levels of FABP4+ alveolar macrophages (Liao et al., n.d.). These cells were found highly inflammatory and strongly associated with cytokine storms. Notably, immunogenomics analysis on blood samples from 5 patients that had early recovery versus 5 patients that recovered late (Wen et al., n.d.) confirmed the findings in BALF, namely that mild disease was associated to clonally expanded cytotoxic CD8+ T cells, which suggest that a specific T cell response may be at play in controlling COVID-19 infection in the lungs.

Several vaccine trials are underway, carrying great promise for sustainable solutions to future epidemics (https://www.who.int/blueprint/priority-diseases/key-action/list-of-candidate-vaccines-developed-against-sars.pdf) with so far 33 vaccines listed (accessed on April 25th, 2020).

From a vaccine point of view, it remains to be seen whether a strong humoral response mediated by neutralizing antibodies can be successfully mounted without side effects, and whether T cell responses could be also generated by a vaccine. Based on current knowledge it appears that severe cases are associated with a sustained, virus-specific immune response, with evidence of a hyper-inflammatory response causing severe lung damage. The fact that all arms of the immune responses are activated in patients that clear the virus, while in severe disease cytokine storms occur along with lymphopenia, suggests that a broad T and B cell-based vaccine should be considered (Y. Shi et al. 2020).

Data on SARS-CoV-2 so far revealed that several epitopes can be targeted by neutralizing antibodies response. Previous work on a SARS-CoV/macaque models investigating host response in the productively infected lungs revealed unexpected interactions between the presence of anti–spike IgG (S-IgG) prior to viral clearance and the Alveolar macrophages (Liu et al. 2019). The latter population was found to undergo functional polarization in acutely infected macaques, demonstrating simultaneously both proinflammatory and wound-healing characteristics. However, the presence of S-IgG prior to viral clearance , abrogated wound-healing responses and promoted MCP1 and IL-8 production and proinflammatory monocyte/macrophage recruitment and accumulation. The authors commented that patients who died of SARS displayed similarly accumulated pulmonary proinflammatory, absence of wound-healing macrophages, and faster neutralizing antibody responses, whereas blockade of FcγR reduced such effects. Therefore, vaccine development should consider possible host responses that exacerbate disease, and alternative therapies should be also considered in combination with vaccines to avoid virus-mediated lung injury.

Perspectives and conclusions

The current COVID-19 pandemic has impacted on many aspects of social, economic,

personal, and behavioral matters of our daily life. We still do not know when we will completely recover from these effects and when life will return to be “normal”. It may never be the same as it used to be or, with luck, an effective vaccine becomes widely available and the return takes only a few months. But the pandemic has also exposed scientists and the processes of scientific advancement, by focusing on what the experts had to say about the new virus, how to control its spread, how to treat those infected, or what to expect if different social measures should be taken. It has shown that science advances through uncertainties and that different opinions and interpretations are frequent, especially when moving through unexplored fields. And a new pathogen is almost inevitably one such field and when it has such a dramatic impact on the health of so many people, the need for urgent answers puts an extraordinary, additional pressure on the daily work of scientists.

In little more than four months, we have learned a lot about SARS-CoV-2 and COVID-19. We know that it is a new coronavirus, closely related to those usually found in bats and, quite oddly, in an endangered species, Malayan pangolins. But we do not know how the jump from a yet unknown intermediate species to humans occurred and how the virus was capable of being so easily and efficiently transmitted among individuals of our species. From what we have learned from other zoonotic spillovers, it seems clear that several factors have concurred in this jump, including ecological, cultural, and possible behavioral. We know that recombination is frequent in coronaviruses, including sarbecoviruses, but we do not know whether this process played a significant role in the emergence of SARS-CoV-2 as a new pathogen. All the evidence gathered so far undermines the possibility that the virus was created in, or escaped from, a laboratory proximate to the city where the initial infections were detected.

SARS-CoV-2 is highly transmissible, even by asymptomatic and presymptomatic infected persons, which makes its spread very difficult to control. This can be observed in its fast expansion in just a few weeks to almost every country in the world and the poor efficiency of border closures, as implemented by most governments, to stop it. When these have been put in action the virus was already circulating within borders. Genomic epidemiology of the virus is allowing a “live” monitoring of the viral spread and has revealed differentiation in several lineages, some clearly resulting from ongoing local circulation, which also show the frequent exchange across borders even from distant countries and continents. Globalization

shows its dark side in this fast spreading pandemic.

There are substantial differences in the natural history of infection. These were observed from the very start of the epidemic, with a much more severe presentation in aged persons and those with additional pathologies. Those differences have persisted but, as the number of infected persons has grown continuously, the cases of serious, even fatal, infections in younger persons have become more frequent. Nevertheless, there are still many unknowns around the differences in susceptibility, progression, and outcome of the infection by SARS-CoV-2 that, as in other infectious diseases, probably depend on a complex interaction of host, pathogen, and environmental factors. Several large-scale analyses encompassing all three kinds of factors are underway and hopefully will shed light on this critical point.

Currently, there are many additional unknowns highly relevant for the possibility of recovering a more or less “normal” lifestyle. There is no specific and highly efficient treatment yet, there are several vaccine candidates in the making but we are still at least months away from their availability for the general population. We do not even know whether those infected by SARS-CoV-2 will develop a strong immune response or a weak one and how long it will last nor whether the virus will have a seasonal behavior or will circulate all year round. In the meantime, the most important measures to prevent a generalized spread of the virus and a collapse of health systems, the main reason behind the social and economic disruption caused by COVD-19, are still social distancing, contact tracing and testing.

Acknowledgements

We thank Michel Tibayrenc (editor-in-chief Infection, Genetics and Evolution) for launching the project of this article, and for inspiring its conception. Tung Phan acknowledges support from the Division of Clinical Microbiology, University of Pittsburgh Medical Center, USA. Manuela Sironi was supported by the Italian Ministry of Health ("Ricerca Corrente 2019-2020"). Fernando González-Candelas was supported by project BFU2017-89594R from MICIN (Spanish Government).

REFERENCES

Ai, Jing-Wen, Yi Zhang, Hao-Cheng Zhang, Teng Xu, and Wen-Hong Zhang. 2020. "Era of Molecular Diagnosis for Pathogen Identification of Unexplained Pneumonia, Lessons to Be Learned." *Emerging Microbes & Infections* 9 (1): 597–600.

Allen, Toph, Kris A. Murray, Carlos Zambrana-Torrelio, Stephen S. Morse, Carlo Rondinini, Moreno Di Marco, Nathan Breit, Kevin J. Olival, and Peter Daszak. 2017. "Global Hotspots and Correlates of Emerging Zoonotic Diseases." *Nature Communications* 8 (1): 1124.

Andersen, Kristian G., Andrew Rambaut, W. Ian Lipkin, Edward C. Holmes, and Robert F. Garry. 2020. "The Proximal Origin of SARS-CoV-2." *Nature Medicine*. https://doi.org/10.1038/s41591-020-0820-9.

"A Novel Bat Coronavirus Closely Related to SARS-CoV-2 Contains Natural Insertions at the S1/S2 Cleavage Site of the Spike Protein." 2020. *Current Biology: CB*, May. https://doi.org/10.1016/j.cub.2020.05.023.

Anyamba, Assaf, Jean-Paul Chretien, Seth C. Britch, Radina P. Soebiyanto, Jennifer L. Small, Rikke Jepsen, Brett M. Forshey, et al. 2019. "Global Disease Outbreaks Associated with the 2015–2016 El Niño Event." *Scientific Reports*. https://doi.org/10.1038/s41598-018-38034-z.

Bedford, Trevor, Alexander L. Greninger, Pavitra Roychoudhury, Lea M. Starita, Michael Famulare, Meei-Li Huang, Arun Nalla, et al. n.d. "Cryptic Transmission of SARS-CoV-2 in Washington State." https://doi.org/10.1101/2020.04.02.20051417.

Boe, D. M., L. A. Boule, and E. J. Kovacs. 2017. "Innate Immune Responses in the Ageing Lung." *Clinical & Experimental Immunology*. https://doi.org/10.1111/cei.12881.

Boni, Maciej F., Philippe Lemey, Xiaowei Jiang, Tommy Tsan-Yuk Lam, Blair Perry, Todd Castoe, Andrew Rambaut, and David L. Robertson. n.d. "Evolutionary Origins of the SARS-CoV-2 Sarbecovirus Lineage Responsible for the COVID-19 Pandemic." https://doi.org/10.1101/2020.03.30.015008.

Borremans, Benny, Christina Faust, Kezia R. Manlove, Susanne H. Sokolow, and James O. Lloyd-Smith. 2019. "Cross-Species Pathogen Spillover across Ecosystem Boundaries: Mechanisms and Theory." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 374 (1782): 20180344.

Boussau, Bastien, and Vincent Daubin. 2010. "Genomes as Documents of Evolutionary

History." *Trends in Ecology & Evolution* 25 (4): 224–32.
Bupp, Melanie R. Gubbels, Melanie R. Gubbels Bupp, Tanvi Potluri, Ashley L. Fink, and Sabra L. Klein. 2018. "The Confluence of Sex Hormones and Aging on Immunity." *Frontiers in Immunology*. https://doi.org/10.3389/fimmu.2018.01269.
Cagliani, Rachele, Diego Forni, Mario Clerici, and Manuela Sironi. 2020. "Computational Inference of Selection Underlying the Evolution of the Novel Coronavirus, SARS-CoV-2." *Journal of Virology*, April. https://doi.org/10.1128/JVI.00411-20.
Chadha, Mandeep S., James A. Comer, Luis Lowe, Paul A. Rota, Pierre E. Rollin, William J. Bellini, Thomas G. Ksiazek, and Akhilesh Mishra. 2006. "Nipah Virus-Associated Encephalitis Outbreak, Siliguri, India." *Emerging Infectious Diseases* 12 (2): 235–40.
Chan, Jasper Fuk-Woo, Shuofeng Yuan, Kin-Hang Kok, Kelvin Kai-Wang To, Hin Chu, Jin Yang, Fanfan Xing, et al. 2020. "A Familial Cluster of Pneumonia Associated with the 2019 Novel Coronavirus Indicating Person-to-Person Transmission: A Study of a Family Cluster." *The Lancet* 395 (10223): 514–23.
Chan, Kelvin Y. K., Johannes C. Y. Ching, M. S. Xu, Annie N. Y. Cheung, Shea-ping Yip, Loretta Y. C. Yam, Sik-to Lai, et al. 2007. "Association ofICAM3Genetic Variant with Severe Acute Respiratory Syndrome." *The Journal of Infectious Diseases*. https://doi.org/10.1086/518892.
Chan, K. Y. K., M. S. Xu, J. C. Y. Ching, V. S. Chan, Y. C. Ip, L. Yam, C. M. Chu, et al. 2010. "Association of a Single Nucleotide Polymorphism in the CD209 (DC-SIGN) Promoter with SARS Severity." *Hong Kong Medical Journal = Xianggang Yi Xue Za Zhi / Hong Kong Academy of Medicine* 16 (5 Suppl 4): 37–42.
Chinese SARS Molecular Epidemiology Consortium. 2004. "Molecular Evolution of the SARS Coronavirus during the Course of the SARS Epidemic in China." *Science* 303 (5664): 1666–69.
Chong, Wai Po, W. K. Eddie Ip, Gloria Hoi Wan Tso, Man Wai Ng, Wilfred Hing Sang Wong, Helen Ka Wai Law, Raymond W. H. Yung, et al. 2006. "The Interferon Gamma Gene Polymorphism 874 A/T Is Associated with Severe Acute Respiratory Syndrome." *BMC Infectious Diseases*. https://doi.org/10.1186/1471-2334-6-82.
Chua, Kaw Bing, Kenny Voon, Gary Crameri, Hui Siu Tan, Juliana Rosli, Jennifer A. McEachern, Sivagami Suluraju, Meng Yu, and Lin-Fa Wang. 2008. "Identification and Characterization of a New Orthoreovirus from Patients with Acute Respiratory Infections." *PLoS ONE*. https://doi.org/10.1371/journal.pone.0003803.
Chua, K. B., G. Crameri, A. Hyatt, M. Yu, M. R. Tompang, J. Rosli, J. McEachern, et al. 2007. "A Previously Unknown Reovirus of Bat Origin Is Associated with an Acute Respiratory Disease in Humans." *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.0701372104.
Clohisey, Sara, and John Kenneth Baillie. 2019. "Host Susceptibility to Severe Influenza A Virus Infection." *Critical Care / the Society of Critical Care Medicine* 23 (1): 303.
Cotten, Matthew, Simon J. Watson, Alimuddin I. Zumla, Hatem Q. Makhdoom, Anne L. Palser, Swee Hoe Ong, Abdullah A. Al Rabeeah, et al. 2014. "Spread, Circulation, and Evolution of the Middle East Respiratory Syndrome Coronavirus." *mBio* 5 (1). https://doi.org/10.1128/mBio.01062-13.
Cui, Jie, Fang Li, and Zheng-Li Shi. 2019. "Origin and Evolution of Pathogenic Coronaviruses." *Nature Reviews. Microbiology* 17 (3): 181–92.
Daszak, P., C. Zambrana-Torrelio, T. L. Bogich, M. Fernandez, J. H. Epstein, K. A. Murray, and H. Hamilton. 2013. "Interdisciplinary Approaches to Understanding Disease

Emergence: The Past, Present, and Future Drivers of Nipah Virus Emergence." *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1201243109.

Dorp, Lucy van, Mislav Acman, Damien Richard, Liam P. Shaw, Charlotte E. Ford, Louise Ormond, Christopher J. Owen, et al. 2020. "Emergence of Genomic Diversity and Recurrent Mutations in SARS-CoV-2." *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, May, 104351.

Fauver, Joseph R., Mary E. Petrone, Emma B. Hodcroft, Kayoko Shioda, Hanna Y. Ehrlich, Alexander G. Watts, Chantal B. F. Vogels, et al. 2020. "Coast-to-Coast Spread of SARS-CoV-2 in the United States Revealed by Genomic Epidemiology." *Public and Global Health*. medRxiv.

Forni, Diego, Rachele Cagliani, Mario Clerici, and Manuela Sironi. 2017. "Molecular Evolution of Human Coronavirus Genomes." *Trends in Microbiology* 25 (1): 35–48.

Forni, Diego, Giulia Filippi, Rachele Cagliani, Luca De Gioia, Uberto Pozzoli, Nasser Al-Daghri, Mario Clerici, and Manuela Sironi. 2015. "The Heptad Repeat Region Is a Major Selection Target in MERS-CoV and Related Coronaviruses." *Scientific Reports* 5 (September): 14480.

Friedrich, M. J. 2018. "WHO's Blueprint List of Priority Diseases." *JAMA*. https://doi.org/10.1001/jama.2018.5712.

Fulop, Tamas, Jacek M. Witkowski, Fabiola Olivieri, and Anis Larbi. 2018. "The Integration of Inflammaging in Age-Related Diseases." *Seminars in Immunology*. https://doi.org/10.1016/j.smim.2018.09.003.

Geoghegan, Jemma L., and Edward C. Holmes. 2017. "Predicting Virus Emergence amid Evolutionary Noise." *Open Biology* 7 (10). https://doi.org/10.1098/rsob.170189.

Giles, John R., Peggy Eby, Hazel Parry, Alison J. Peel, Raina K. Plowright, David A. Westcott, and Hamish McCallum. 2018. "Environmental Drivers of Spatiotemporal Foraging Intensity in Fruit Bats and Implications for Hendra Virus Ecology." *Scientific Reports* 8 (1): 9555.

Gong, Lang, Jie Li, Qingfeng Zhou, Zhichao Xu, Li Chen, Yun Zhang, Chunyi Xue, Zhifen Wen, and Yongchang Cao. 2017. "A New Bat-HKU2–like Coronavirus in Swine, China, 2017." *Emerging Infectious Diseases*. https://doi.org/10.3201/eid2309.170915.

Gonzalez-Reiche, Ana S., Matthew M. Hernandez, Mitchell Sullivan, Brianne Ciferri, Hala Alshammary, Ajay Obla, Shelcie Fabre, et al. n.d. "Introductions and Early Spread of SARS-CoV-2 in the New York City Area." https://doi.org/10.1101/2020.04.08.20056929.

Goronzy, Jörg J., and Cornelia M. Weyand. 2013. "Understanding Immunosenescence to Improve Responses to Vaccines." *Nature Immunology*. https://doi.org/10.1038/ni.2588.

Gralinski, Lisa E., Martin T. Ferris, David L. Aylor, Alan C. Whitmore, Richard Green, Matthew B. Frieman, Damon Deming, et al. 2015. "Genome Wide Identification of SARS-CoV Susceptibility Loci Using the Collaborative Cross." *PLoS Genetics* 11 (10): e1005504.

Gralinski, Lisa E., Vineet D. Menachery, Andrew P. Morgan, Allison L. Totura, Anne Beall, Jacob Kocher, Jessica Plante, et al. 2017. "Allelic Variation in the Toll-Like Receptor Adaptor Protein Ticam2 Contributes to SARS-Coronavirus Pathogenesis in Mice." *G3: Genes|Genomes|Genetics*. https://doi.org/10.1534/g3.117.041434.

Gudbjartsson, Daniel F., Agnar Helgason, Hakon Jonsson, Olafur T. Magnusson, Pall Melsted, Gudmundur L. Norddahl, Jona Saemundsdottir, et al. 2020. “Early Spread of SARS-Cov-2 in the Icelandic Population.” *Epidemiology*. medRxiv.

Guindon, Stéphane, and Olivier Gascuel. 2003. “A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood.” *Systematic Biology* 52 (5): 696–704.

Hadfield, James, Colin Megill, Sidney M. Bell, John Huddleston, Barney Potter, Charlton Callender, Pavel Sagulenko, Trevor Bedford, and Richard A. Neher. 2018. “Nextstrain: Real-Time Tracking of Pathogen Evolution.” *Bioinformatics*. https://doi.org/10.1093/bioinformatics/bty407.

Han, Qingmei, Qingqing Lin, Zuowei Ni, and Liangshun You. 2020. “Uncertainties about the Transmission Routes of 2019 Novel Coronavirus.” *Influenza and Other Respiratory Viruses*, March. https://doi.org/10.1111/irv.12735.

Hindson, Jordan. 2020. “COVID-19: Faecal–oral Transmission?” *Nature Reviews Gastroenterology & Hepatology*. https://doi.org/10.1038/s41575-020-0295-7.

Holder, Mark, and Paul O. Lewis. 2003. “Phylogeny Estimation: Traditional and Bayesian Approaches.” *Nature Reviews Genetics*. https://doi.org/10.1038/nrg1044.

Hou, Yuxuan, Cheng Peng, Meng Yu, Yan Li, Zhenggang Han, Fang Li, Lin-Fa Wang, and Zhengli Shi. 2010. “Angiotensin-Converting Enzyme 2 (ACE2) Proteins of Different Bat Species Confer Variable Susceptibility to SARS-CoV Entry.” *Archives of Virology*. https://doi.org/10.1007/s00705-010-0729-6.

Huang, Chaolin, Yeming Wang, Xingwang Li, Lili Ren, Jianping Zhao, Yi Hu, Li Zhang, et al. 2020. “Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China.” *The Lancet* 395 (10223): 497–506.

“Infection and Rapid Transmission of SARS-CoV-2 in Ferrets.” 2020. *Cell Host & Microbe*, April. https://doi.org/10.1016/j.chom.2020.03.023.

Ip, W. K. Eddie, W. K. Eddie Ip, Kwok Hung Chan, Helen K. W. Law, Gloria H. W. Tso, Eric K. P. Kong, Wilfred H. S. Wong, et al. 2005. “Mannose-Binding Lectin in Severe Acute Respiratory Syndrome Coronavirus Infection.” *The Journal of Infectious Diseases*. https://doi.org/10.1086/429631.

Itoyama, Satoru, Naoto Keicho, Minako Hijikata, Tran Quy, Nguyen Chi Phi, Hoang Thuy Long, Ha Le Dang, et al. 2005. “Identification of an Alternative 5’-Untranslated Exon and New Polymorphisms of Angiotensin-Converting Enzyme 2 Gene: Lack of Association with SARS in the Vietnamese Population.” *American Journal of Medical Genetics Part A*. https://doi.org/10.1002/ajmg.a.30779.

Iwata-Yoshikawa, Naoko, Tadashi Okamura, Yukiko Shimizu, Hideki Hasegawa, Makoto Takeda, and Noriyo Nagata. 2019. “TMPRSS2 Contributes to Virus Spread and Immunopathology in the Airways of Murine Models after Coronavirus Infection.” *Journal of Virology* 93 (6). https://doi.org/10.1128/JVI.01815-18.

Ji, Wei, Wei Wang, Xiaofang Zhao, Junjie Zai, and Xingguang Li. 2020. “Cross-species Transmission of the Newly Identified Coronavirus 2019-nCoV.” *Journal of Medical Virology*. https://doi.org/10.1002/jmv.25682.

Johnson, Christine K., Peta L. Hitchens, Pranav S. Pandit, Julie Rushmore, Tierra Smiley Evans, Cristin C. W. Young, and Megan M. Doyle. 2020. “Global Shifts in Mammalian Population Trends Reveal Key Predictors of Virus Spillover Risk.” *Proceedings. Biological Sciences / The Royal Society* 287 (1924): 20192736.

Jones, Kate E., Nikkita G. Patel, Marc A. Levy, Adam Storeygard, Deborah Balk, John L.

Gittleman, and Peter Daszak. 2008. “Global Trends in Emerging Infectious Diseases.” *Nature*. https://doi.org/10.1038/nature06536.

Kadel, Sapana, and Susan Kovats. 2018. “Sex Hormones Regulate Innate Immune Cells and Promote Sex Differences in Respiratory Virus Infection.” *Frontiers in Immunology* 9 (July): 1653.

Kaiser, Jocelyn. 2020. “How Sick Will the Coronavirus Make You? The Answer May Be in Your Genes.” *Science*. https://doi.org/10.1126/science.abb9192.

Kane, Melissa, and Tatyana V. Golovkina. 2019. “Mapping Viral Susceptibility Loci in Mice.” *Annual Review of Virology* 6 (1): 525–46.

Keesing, Felicia, Lisa K. Belden, Peter Daszak, Andrew Dobson, C. Drew Harvell, Robert D. Holt, Peter Hudson, et al. 2010. “Impacts of Biodiversity on the Emergence and Transmission of Infectious Diseases.” *Nature* 468 (7324): 647–52.

Keesing, F., R. D. Holt, and R. S. Ostfeld. 2006. “Effects of Species Diversity on Disease Risk.” *Ecology Letters*. https://doi.org/10.1111/j.1461-0248.2006.00885.x.

Keicho, Naoto, Satoru Itoyama, Koichi Kashiwase, Nguyen Chi Phi, Hoang Thuy Long, Ha Le Dang, Vo Van Ban, et al. 2009. “Association of Human Leukocyte Antigen Class II Alleles with Severe Acute Respiratory Syndrome in the Vietnamese Population.” *Human Immunology*. https://doi.org/10.1016/j.humimm.2009.05.006.

Killerby, Marie E., Holly M. Biggs, Claire M. Midgley, Susan I. Gerber, and John T. Watson. 2020. “Middle East Respiratory Syndrome Coronavirus Transmission.” *Emerging Infectious Diseases* 26 (2): 191–98.

Kim, Yeon-Sook, Abdimadiyeva Aigerim, Uni Park, Yuri Kim, Ji-Young Rhee, Jae-Phil Choi, Wan Beom Park, et al. 2019. “Sequential Emergence and Wide Spread of Neutralization Escape Middle East Respiratory Syndrome Coronavirus Mutants, South Korea, 2015.” *Emerging Infectious Diseases* 25 (6): 1161–68.

Kim, Yuri, Shinhye Cheon, Chan-Ki Min, Kyung Mok Sohn, Ying Jin Kang, Young-Je Cha, Ju-Il Kang, et al. 2016. “Spread of Mutant Middle East Respiratory Syndrome Coronavirus with Reduced Affinity to Human CD26 during the South Korean Outbreak.” *mBio* 7 (2): e00019.

Kleine-Weber, Hannah, Mahmoud Tarek Elzayat, Lingshu Wang, Barney S. Graham, Marcel A. Müller, Christian Drosten, Stefan Pöhlmann, and Markus Hoffmann. 2019. “Mutations in the Spike Protein of Middle East Respiratory Syndrome Coronavirus Transmitted in Korea Increase Resistance to Antibody-Mediated Neutralization.” *Journal of Virology* 93 (2). https://doi.org/10.1128/JVI.01381-18.

Kogut, Igor, Jean L. Scholz, Michael P. Cancro, and John C. Cambier. 2012. “B Cell Maintenance and Function in Aging.” *Seminars in Immunology*. https://doi.org/10.1016/j.smim.2012.04.004.

Kollmann, Tobias R., Ofer Levy, Ruth R. Montgomery, and Stanislas Goriely. 2012. “Innate Immune Function by Toll-like Receptors: Distinct Responses in Newborns and the Elderly.” *Immunity*. https://doi.org/10.1016/j.immuni.2012.10.014.

Lam, Sai Kit, and Kaw Bing Chua. 2002. “Nipah Virus Encephalitis Outbreak in Malaysia.” *Clinical Infectious Diseases*. https://doi.org/10.1086/338818.

Lam, Tommy Tsan-Yuk, Marcus Ho-Hin Shum, Hua-Chen Zhu, Yi-Gang Tong, Xue-Bing Ni, Yun-Shi Liao, Wei Wei, et al. 2020. “Identifying SARS-CoV-2 Related Coronaviruses in Malayan Pangolins.” *Nature*, March, 1–6.

Lan, Jun, Jiwan Ge, Jinfang Yu, Sisi Shan, Huan Zhou, Shilong Fan, Qi Zhang, et al. 2020. “Structure of the SARS-CoV-2 Spike Receptor-Binding Domain Bound to the ACE2

Receptor." *Nature*, March. https://doi.org/10.1038/s41586-020-2180-5.
Lau, Susanna K. P., Yun Feng, Honglin Chen, Hayes K. H. Luk, Wei-Hong Yang, Kenneth S. M. Li, Yu-Zhen Zhang, et al. 2015. "Severe Acute Respiratory Syndrome (SARS) Coronavirus ORF8 Protein Is Acquired from SARS-Related Coronavirus from Greater Horseshoe Bats through Recombination." *Journal of Virology* 89 (20): 10532–47.
Liang, Wan-Nian, Yong Huang, Wan-Xin Zhou, Lei Qiao, Jian-Hui Huang, and Zheng-Lai Wu. 2003. "Epidemiological Characteristics of an Outbreak of Severe Acute Respiratory Syndrome in Dongcheng District of Beijing from March to May 2003." *Biomedical and Environmental Sciences: BES* 16 (4): 305–13.
Liao, Minfeng, Yang Liu, Jin Yuan, Yanling Wen, Gang Xu, Juanjuan Zhao, Lin Chen, et al. n.d. "The Landscape of Lung Bronchoalveolar Immune Cells in COVID-19 Revealed by Single-Cell RNA Sequencing." https://doi.org/10.1101/2020.02.23.20026690.
Li, H., N. L-S Tang, P. K-S Chan, C-Y Wang, D. S-C Hui, C. Luk, R. Kwok, et al. 2008. "Polymorphisms in the C-Type Lectin Genes Cluster in Chromosome 19 and Predisposition to Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) Infection." *Journal of Medical Genetics*. https://doi.org/10.1136/jmg.2008.058966.
Lin, Lu, Xiayang Jiang, Zhenling Zhang, Siwen Huang, Zhenyi Zhang, Zhaoxiong Fang, Zhiqiang Gu, et al. 2020. "Gastrointestinal Symptoms of 95 Cases with SARS-CoV-2 Infection." *Gut*, April. https://doi.org/10.1136/gutjnl-2020-321013.
Lin, Marie, Hsiang-Kuang Tseng, Jean A. Trejaut, Hui-Lin Lee, Jun-Hun Loo, Chen-Chung Chu, Pei-Jan Chen, et al. 2003. "Association of HLA Class I with Severe Acute Respiratory Syndrome Coronavirus Infection." *BMC Medical Genetics*. https://doi.org/10.1186/1471-2350-4-9.
Liu, Li, Qiang Wei, Qingqing Lin, Jun Fang, Haibo Wang, Hauyee Kwok, Hangying Tang, et al. 2019. "Anti-Spike IgG Causes Severe Acute Lung Injury by Skewing Macrophage Responses during Acute SARS-CoV Infection." *JCI Insight* 4 (4). https://doi.org/10.1172/jci.insight.123158.
Li, Wenhui, Chengsheng Zhang, Jianhua Sui, Jens H. Kuhn, Michael J. Moore, Shiwen Luo, Swee-Kee Wong, et al. 2005. "Receptor and Viral Determinants of SARS-Coronavirus Adaptation to Human ACE2." *The EMBO Journal* 24 (8): 1634–43.
Lopera, Esteban, Adriaan van der Graaf, Pauline Lanting, Marije van der Geest, Lifelines Cohort Study, Jingyuan Fu, Morris Swertz, et al. 2020. "Lack of Association between Genetic Variants at ACE2 and TMPRSS2 Genes Involved in SARS-CoV-2 Infection and Human Quantitative Phenotypes." *medRxiv*, April, 2020.04.22.20074963.
Lu, Guangwen, Qihui Wang, and George F. Gao. 2015. "Bat-to-Human: Spike Features Determining 'Host Jump' of Coronaviruses SARS-CoV, MERS-CoV, and beyond." *Trends in Microbiology* 23 (8): 468–78.
McFarlane, Rosemary, Niels Becker, and Hume Field. 2011. "Investigation of the Climatic and Environmental Context of Hendra Virus Spillover Events 1994-2010." *PloS One* 6 (12): e28374.
Mehta, Puja, Daniel F. McAuley, Michael Brown, Emilie Sanchez, Rachel S. Tattersall, and Jessica J. Manson. 2020. "COVID-19: Consider Cytokine Storm Syndromes and Immunosuppression." *The Lancet* 395 (10229): 1033–34.
Menachery, Vineet D., Boyd L. Yount Jr, Kari Debbink, Sudhakar Agnihothram, Lisa E. Gralinski, Jessica A. Plante, Rachel L. Graham, et al. 2015. "A SARS-like Cluster of Circulating Bat Coronaviruses Shows Potential for Human Emergence." *Nature Medicine* 21 (12): 1508–13.

Morand, Serge, Katharine A. Owers, Agnes Waret-Szkuta, K. Marie McIntyre, and Matthew Baylis. 2013. “Climate Variability and Outbreaks of Infectious Diseases in Europe.” *Scientific Reports*. https://doi.org/10.1038/srep01774.

Muth, Doreen, Victor Max Corman, Hanna Roth, Tabea Binger, Ronald Dijkman, Lina Theresa Gottula, Florian Gloza-Rausch, et al. 2018. “Attenuation of Replication by a 29 Nucleotide Deletion in SARS-Coronavirus Acquired during the Early Stages of Human-to-Human Transmission.” *Scientific Reports* 8 (1): 15177.

Neher, Richard A., and Trevor Bedford. 2018. “Real-Time Analysis and Visualization of Pathogen Sequence Data.” *Journal of Clinical Microbiology* 56 (11). https://doi.org/10.1128/JCM.00480-18.

Ng, Man Wai, Gangqiao Zhou, Wai Po Chong, Loretta Wing Yan Lee, Helen Ka Wai Law, Hongxing Zhang, Wilfred Hing Sang Wong, et al. 2007. “The Association of RANTES Polymorphism with Severe Acute Respiratory Syndrome in Hong Kong and Beijing Chinese.” *BMC Infectious Diseases*. https://doi.org/10.1186/1471-2334-7-50.

Ng, Margaret H. L., Kin-mang Lau, Libby Li, Suk-hang Cheng, Wing Y. Chan, Pak K. Hui, Benny Zee, Chi-bon Leung, and Joseph J. Y. Sung. 2004. “Association of Human-Leukocyte-Antigen Class I (B*0703) and Class II (DRB1*0301) Genotypes with Susceptibility and Resistance to the Development of Severe Acute Respiratory Syndrome.” *The Journal of Infectious Diseases*. https://doi.org/10.1086/421523.

Nicholls, N. 1993. “El Niño-Southern Oscillation and Vector-Borne Disease.” *The Lancet*. https://doi.org/10.1016/0140-6736(93)92368-4.

Niedzwiedz, Claire L., Catherine A. O’Donnell, Bhautesh D. Jani, Evangelia Demou, Frederick K. Ho, Carlos Celis-Morales, Barbara I. Nicholl, et al. 2020. “Ethnic and Socioeconomic Differences in SARS-CoV2 Infection in the UK Biobank Cohort Study.” *medRxiv*, April, 2020.04.22.20075663.

Nikolich-Žugich, Janko. 2018. “The Twilight of Immunity: Emerging Concepts in Aging of the Immune System.” *Nature Immunology*. https://doi.org/10.1038/s41590-017-0006-x.

Nogales, Aitor, and Marta L. DeDiego. 2019. “Host Single Nucleotide Polymorphisms Modulating Influenza A Virus Disease in Humans.” *Pathogens*. https://doi.org/10.3390/pathogens8040168.

Normile, Dennis, and Martin Enserink. 2003. “SARS in China. Tracking the Roots of a Killer.” *Science* 301 (5631): 297–99.

Oh, Soo-Jin, Jae Kyung Lee, and Ok Sarah Shin. 2019. “Aging and the Immune System: The Impact of Immunosenescence on Viral Infection, Immunity and Vaccine Immunogenicity.” *Immune Network*. https://doi.org/10.4110/in.2019.19.e37.

Olsen, Sonja J., Meng-Yu Chen, Yu-Lun Liu, Mark Witschi, Alexis Ardoin, Clémentine Calba, Pauline Mathieu, et al. 2020. “Early Introduction of Severe Acute Respiratory Syndrome Coronavirus 2 into Europe.” *Emerging Infectious Diseases* 26 (7). https://doi.org/10.3201/eid2607.200359.

Phan, Tung. 2020. “Novel Coronavirus: From Discovery to Clinical Diagnostics.” *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 79 (April): 104211.

Plowright, Raina K., Colin R. Parrish, Hamish McCallum, Peter J. Hudson, Albert I. Ko, Andrea L. Graham, and James O. Lloyd-Smith. 2017. “Pathways to Zoonotic Spillover.” *Nature Reviews. Microbiology* 15 (8): 502–10.

“Population Genetics and Population Biology: What Did They Bring to the Epidemiology of Transmissible Diseases? An E-Debate.” 2001. *Infection, Genetics and Evolution:*

Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases 1 (2): 161–66.

Qu, Xiu-Xia, Pei Hao, Xi-Jun Song, Si-Ming Jiang, Yan-Xia Liu, Pei-Gang Wang, Xi Rao, et al. 2005. "Identification of Two Critical Amino Acid Residues of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein for Its Variation in Zoonotic Tropism Transition via a Double Substitution Strategy." *The Journal of Biological Chemistry* 280 (33): 29588–95.

Rockx, Barry, Tracey Baas, Gregory A. Zornetzer, Bart Haagmans, Timothy Sheahan, Matthew Frieman, Matthew D. Dyer, et al. 2009. "Early Upregulation of Acute Respiratory Distress Syndrome-Associated Cytokines Promotes Lethal Disease in an Aged-Mouse Model of Severe Acute Respiratory Syndrome Coronavirus Infection." *Journal of Virology*. https://doi.org/10.1128/jvi.00127-09.

Rockx, Barry, Eric Donaldson, Matthew Frieman, Timothy Sheahan, Davide Corti, Antonio Lanzavecchia, and Ralph S. Baric. 2010. "Escape from Human Monoclonal Antibody Neutralization Affects in Vitro and in Vivo Fitness of Severe Acute Respiratory Syndrome Coronavirus." *The Journal of Infectious Diseases* 201 (6): 946–55.

Rohr, Jason R., Christopher B. Barrett, David J. Civitello, Meggan E. Craft, Bryan Delius, Giulio A. DeLeo, Peter J. Hudson, et al. 2019. "Emerging Human Infectious Diseases and the Links to Global Food Production." *Nature Sustainability* 2 (6): 445–56.

Schäfer, Alexandra, Ralph S. Baric, and Martin T. Ferris. 2014. "Systems Approaches to Coronavirus Pathogenesis." *Current Opinion in Virology*. https://doi.org/10.1016/j.coviro.2014.04.007.

Selvey, Linda A., Rachel M. Wells, Joseph G. McCormack, Anthony J. Ansford, Keith Murray, Russell J. Rogers, Peter S. Lavercombe, Paul Selleck, and John W. Sheridan. 1995. "Infection of Humans and Horses by a Newly Described Morbillivirus." *Medical Journal of Australia*. https://doi.org/10.5694/j.1326-5377.1995.tb126050.x.

Shi, Jianzhong, Zhiyuan Wen, Gongxun Zhong, Huanliang Yang, Chong Wang, Baoying Huang, Renqiang Liu, et al. 2020. "Susceptibility of Ferrets, Cats, Dogs, and Other Domesticated Animals to SARS-Coronavirus 2." *Science*, April. https://doi.org/10.1126/science.abb7015.

Shi, Yufang, Ying Wang, Changshun Shao, Jianan Huang, Jianhe Gan, Xiaoping Huang, Enrico Bucci, Mauro Piacentini, Giuseppe Ippolito, and Gerry Melino. 2020. "COVID-19 Infection: The Perspectives on Immune Responses." *Cell Death and Differentiation* 27 (5): 1451–54.

Sironi, Manuela, Rachele Cagliani, Diego Forni, and Mario Clerici. 2015. "Evolutionary Insights into Host–pathogen Interactions from Mammalian Sequence Data." *Nature Reviews Genetics*. https://doi.org/10.1038/nrg3905.

Smith, Katherine F., Michael Behrens, Lisa M. Schloegel, Nina Marano, Stas Burgiel, and Peter Daszak. 2009. "Ecology. Reducing the Risks of the Wildlife Trade." *Science* 324 (5927): 594–95.

Solana, Rafael, Raquel Tarazona, Inmaculada Gayoso, Olivier Lesur, Gilles Dupuis, and Tamas Fulop. 2012. "Innate Immunosenescence: Effect of Aging on Cells and Receptors of the Innate Immune System in Humans." *Seminars in Immunology*. https://doi.org/10.1016/j.smim.2012.04.008.

Spiteri, Gianfranco, James Fielding, Michaela Diercke, Christine Campese, Vincent Enouf, Alexandre Gaymard, Antonino Bella, et al. 2020. "First Cases of Coronavirus Disease 2019 (COVID-19) in the WHO European Region, 24 January to 21 February 2020." *Euro*

Surveillance: Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin 25 (9). https://doi.org/10.2807/1560-7917.ES.2020.25.9.2000178.

Stoecklin, Sibylle Bernard, Patrick Rolland, Yassoungo Silue, Alexandra Mailles, Christine Campese, Anne Simondon, Matthieu Mechain, et al. 2020. "First Cases of Coronavirus Disease 2019 (COVID-19) in France: Surveillance, Investigations and Control Measures, January 2020." *Eurosurveillance*. https://doi.org/10.2807/1560-7917.es.2020.25.6.2000094.

Subudhi, Sonu, Noreen Rapin, and Vikram Misra. 2019. "Immune System Modulation and Viral Persistence in Bats: Understanding Viral Spillover." *Viruses*. https://doi.org/10.3390/v11020192.

Tahamtan, Alireza, Fatemeh Sana Askari, Louis Bont, and Vahid Salimi. 2019. "Disease Severity in Respiratory Syncytial Virus Infection: Role of Host Genetic Variation." *Reviews in Medical Virology* 29 (2): e2026.

Thevarajan, Irani, Thi H. O. Nguyen, Marios Koutsakos, Julian Druce, Leon Caly, Carolien E. van de Sandt, Xiaoxiao Jia, et al. 2020. "Breadth of Concomitant Immune Responses prior to Patient Recovery: A Case Report of Non-Severe COVID-19." *Nature Medicine* 26 (4): 453–55.

Valitutto, Marc T., Ohnmar Aung, Kyaw Yan Naing Tun, Megan E. Vodzak, Dawn Zimmerman, Jennifer H. Yu, Ye Tun Win, et al. 2020. "Detection of Novel Coronaviruses in Bats in Myanmar." *PloS One* 15 (4): e0230802.

Vanwambeke, Sophie O., Catherine Linard, and Marius Gilbert. 2019. "Emerging Challenges of Infectious Diseases as a Feature of Land Systems." *Current Opinion in Environmental Sustainability* 38 (June): 31–36.

Villabona-Arenas, Ch Julián, William P. Hanage, and Damien C. Tully. 2020. "Phylogenetic Interpretation during Outbreaks Requires Caution." *Nature Microbiology*, May, 1–2.

Walls, Alexandra C., Young-Jun Park, M. Alejandra Tortorici, Abigail Wall, Andrew T. McGuire, and David Veesler. 2020. "Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein." *Cell*, March. https://doi.org/10.1016/j.cell.2020.02.058.

Wang, Sheng-Fan, Kuan-Hsuan Chen, Marcelo Chen, Wen-Yi Li, Yen-Ju Chen, Ching-Han Tsao, Muh-Yong Yen, Jason C. Huang, and Yi-Ming Arthur Chen. 2011. "Human-Leukocyte Antigen Class I Cw 1502 and Class II DR 0301 Genotypes Are Associated with Resistance to Severe Acute Respiratory Syndrome (SARS) Infection." *Viral Immunology*. https://doi.org/10.1089/vim.2011.0024.

Wang, Xinjie, Pinpin Ji, Huiying Fan, Lu Dang, Wenwei Wan, Siyuan Liu, Yanhua Li, et al. 2020. "CRISPR/Cas12a Technology Combined with Immunochromatographic Strips for Portable Detection of African Swine Fever Virus." *Communications Biology* 3 (1): 62.

Wen, Wen Wen, Wenru Su, Hao Tang, Wenqing Le, Xiaopeng Zhang, Yingfeng Zheng, et al. n.d. "Immune Cell Profiling of COVID-19 Patients in the Recovery Stage by Single-Cell Sequencing." https://doi.org/10.1101/2020.03.23.20039362.

Wilkinson, David A., Jonathan C. Marshall, Nigel P. French, and David T. S. Hayman. 2018. "Habitat Fragmentation, Biodiversity Loss and the Risk of Novel Infectious Disease Emergence." *Journal of the Royal Society, Interface / the Royal Society* 15 (149). https://doi.org/10.1098/rsif.2018.0403.

Williams, Frances M. K., Maxim Freydin, Massimo Mangino, Simon Couvreur, Alessia Visconti, Ruth C. E. Bowyer, Caroline I. Le Roy, et al. 2020. "Self-Reported Symptoms of Covid-19 Including Symptoms Most Predictive of SARS-CoV-2 Infection, Are

Heritable." *medRxiv*, April, 2020.04.22.20072124.
Wolfe, Nathan D., Claire Panosian Dunavan, and Jared Diamond. 2007. "Origins of Major Human Infectious Diseases." *Nature*. https://doi.org/10.1038/nature05775.
Wu, Kailang, Guiqing Peng, Matthew Wilken, Robert J. Geraghty, and Fang Li. 2012. "Mechanisms of Host Receptor Adaptation by Severe Acute Respiratory Syndrome Coronavirus." *The Journal of Biological Chemistry* 287 (12): 8904–11.
Xu, Yifei. 2020. "Unveiling the Origin and Transmission of 2019-nCoV." *Trends in Microbiology* 28 (4): 239–40.
Yancy, Clyde W. 2020. "COVID-19 and African Americans." *JAMA: The Journal of the American Medical Association*, April. https://doi.org/10.1001/jama.2020.6548.
Zaki, Ali M., Sander van Boheemen, Theo M. Bestebroer, Albert D M, and Ron A. M. Fouchier. 2012. "Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia." *New England Journal of Medicine*. https://doi.org/10.1056/nejmoa1211721.
Zhang, Hongxing, Gangqiao Zhou, Lianteng Zhi, Hao Yang, Yun Zhai, Xiaojia Dong, Xiumei Zhang, Xue Gao, Yunping Zhu, and Fuchu He. 2005. "Association between Mannose-Binding Lectin Gene Polymorphisms and Susceptibility to Severe Acute Respiratory Syndrome Coronavirus Infection." *The Journal of Infectious Diseases*. https://doi.org/10.1086/491479.
Zhao, Kai, Hui Wang, and Changyou Wu. 2011. "The Immune Responses of HLA-A*0201 Restricted SARS-CoV S Peptide-Specific CD8 T Cells Are Augmented in Varying Degrees by CpG ODN, PolyI:C and R848." *Vaccine*. https://doi.org/10.1016/j.vaccine.2011.06.100.
Zhou, Peng, Xing-Lou Yang, Xian-Guang Wang, Ben Hu, Lei Zhang, Wei Zhang, Hao-Rui Si, et al. 2020. "A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin." *Nature* 579 (7798): 270–73.
Zhu, Na, Dingyu Zhang, Wenling Wang, Xingwang Li, Bo Yang, Jingdong Song, Xiang Zhao, et al. 2020. "A Novel Coronavirus from Patients with Pneumonia in China, 2019." *The New England Journal of Medicine* 382 (8): 727–33.
Zhu, Xiaohui, Yan Wang, Hongxing Zhang, Xuan Liu, Ting Chen, Ruifu Yang, Yuling Shi, et al. 2011. "Genetic Variation of the Human α-2-Heremans-Schmid Glycoprotein (AHSG) Gene Associated with the Risk of SARS-CoV Infection." *PloS One* 6 (8): e23730.
Zinsstag, Jakob, Lisa Crump, Esther Schelling, Jan Hattendorf, Yahya Osman Maidane, Kadra Osman Ali, Abdifatah Muhummed, et al. 2018. "Climate Change and One Health." *FEMS Microbiology Letters* 365 (11). https://doi.org/10.1093/femsle/fny085.

Grubaugh ND, Petrone ME, Holmes EC. We shouldn't worry when a virus mutates during disease outbreaks. Nat Microbiol 2020;5;529-30. https://doi.org/10.1038/s41564-020-0690-4

Table 1. Emergence of bat-borne viruses in Asia (Middle East, China, South Asia, Southeast Asia) and Australia in relation to El Niño Southern Oscillation (ENSO) driven climate anomalies. Note that the existence of lag time between the index ENSO 3.4 and its effects on a country or region may vary from less than one month (Australia) to two months (Southeast Asia), three months (South Asia, Middle East) and up to four-six months (China).

Emergence / outbreaks	Intermediate host	Date, location	ENSO	Reference
Hendra	Horse	Sep 1994, Australia	El Niño	(Selvey et al. 1995)
Nipah	Swine	Sep 1998, Malaysia	La Niña	(S. K. Lam and Chua 2002)
Nipah	Unknown	Jan 2001, India	La Niña	(Chadha et al. 2006)
SARS	Civet cat	Nov 2002, China	El Niño	(Liang et al. 2003)
Melaka	Unknown	Mar 2006, Malaysia	La Niña	(K. B. Chua et al. 2007)
Kampar	Unknown	Aug 2006, Malaysia	Neutral Phase	(Kaw Bing Chua et al. 2008)
MERS-CoV	Camel	Apr 2012, Middle East	La Niña	(Zaki et al. 2012)
HKU2 2017	Swine	Oct 2016, China	La Niña	(Gong et al. 2017)
COVID-12	Unknown	Dec 2019, China	El Niño	(N. Zhu et al. 2020)

LEGENDS TO FIGURES

Figure 1. Comparative analysis of SARS-CoV-2 with other coronaviruses. Maximum-likelihood phylogenetic trees of the NTD (A) and RBD (B) regions of SARS-CoV-2 (red), RaTG13 (green), RmYN02 (light blue), Pangolin coronaviruses (Guangdong lineage, grey; Guangxi lineage, orange), and other Asian (dark blue) and non-Asian (brown) bat coronaviruses belonging to the *Sarbecovirus* subgenus. Trees are based on amino acid sequences and were built using PhyML (Guindon and Gascuel 2003). Trees are mid-point rooted. (C) Combined variability in S1 (grey) and S2 (red) domains of SARS-CoV-2 when compared to RaTG13 and pangolin coronaviruses spike sequences. (D) Sequence alignments showing absence of the YLTPGD insert in bat sarbecoviruses, and the sequence of the RBD region involved in the interaction with ACE2. (E) The position of YLTPGD inserts forming conformational clusters (red spheres) at the NTD of SARS-CoV-2 spike protein is shown (left). The ribbon structure of the spike protein-ACE2 interaction surface is represented to show polar interactions (right). Polar interactions were analyzed using PyMol using PDB id: 6m0j (Lan et al. 2020). (F) Alignment of the region carrying the polybasic amino acid insertion (red) at the S1/S2 cleavage site. GenBank/GISAID accessions for the sequences included in trees are: NC_045512.2 (SARS-CoV-2), MN996532.1(RaTG13), EPI_ISL_412977 (RmYN02), MT084071.1 (MP789 or Guangdong 1), EPI_ISL_410544 (Guangdong P2S), MT040334.1 (GX-P1E),MT072865.1 (GX-P3B), MT040335.1 (GX-P5L), KY417148 (Rs4247), DQ071615.1 (Rp3), GQ153547.1 (HKU3-12), GQ153542 (HKU3-7), MK211378.1 (BtRs-BetaCoV/YN2018D), DQ648856.1 (BtCoV/273/2005), JX993987.1 (Rp/Shaanxi2011), KJ473816 (BtRs-BetaCoV/YN2013), MG772933 (CoVZC45), MG772934 (CoVZXC21), KY417151.1 (Rs7327), KF569996 (LYRa11), , NC_014470.1 (BM48-31/BGR/2008), KY352407.1 (BtKY72).

Figure 2. Onset of bat-borne viral diseases in Asia and the Pacific since 1990 (see Table 2) in relation to El Niño Southern Oscillation (ENSO) driven climate anomalies using 'NINO 3.4' index retrieved from the National Oceanic and Atmospheric Administration (NOAA, https://www.noaa.gov).

Figure 3. Molecular phylogeny of *Mycobacterium tuberculosis* showing typical some degree of concordance with known chronology of sampling. Source: https://nextstrain.org/tb/global?c=num_date&l=rect&m=div

Figure 4. Maximum Likelihood phylogenetic reconstruction of SARS-CoV-2 isolates deposited to GISAID.org and rendered by Nextstrain.org. Note the absolute concordance of predicted and known chronological sampling.

Figure 5. The same reconstruction of SARS-CoV-2 phylogeny, now denoted by geography. Isolates originated and initially diversified in China (purple), followed by multiple and independent introductions to Oceania (blue), Europe (green and yellow), and North America (red). Less information is known about Africa, India, South America, and other populations of major concern in the Global South.

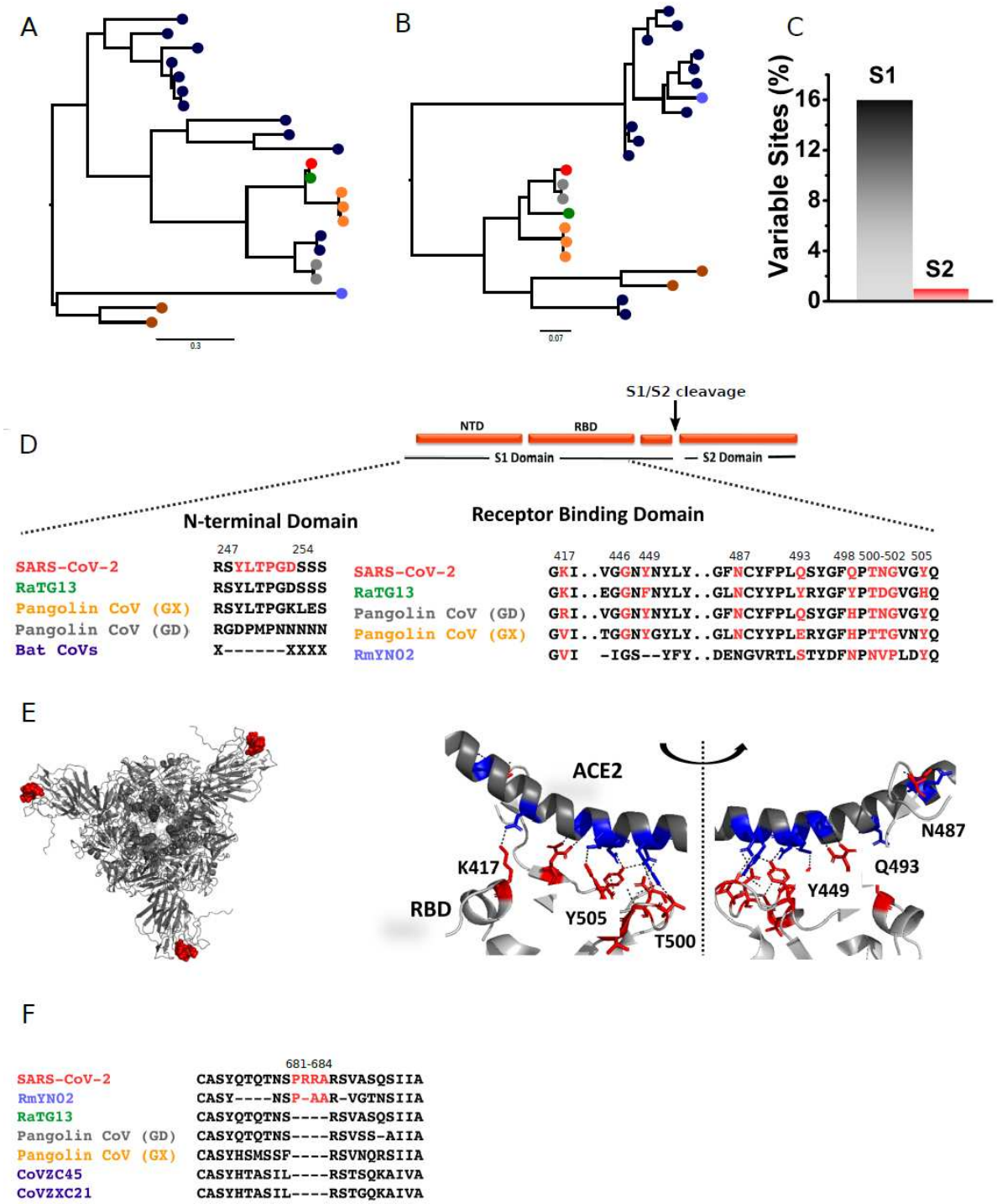
A
B
C
Variable Sites (%)
S1
S2
D
S1/S2 cleavage
NTD
RBD
S1 Domain
S2 Domain
N-terminal Domain
Receptor Binding Domain
SARS-CoV-2 RSYLTPGDSSS
RaTG13 RSYLTPGDSSS
Pangolin CoV (GX) RSYLTPGKLES
Pangolin CoV (GD) RGDPMPNNNNN
Bat CoVs X------XXXX
SARS-CoV-2 GKI..VGGNYNYLY..GFNCYFPLQSYGFQPTNGVGYQ
RaTG13 GKI..EGGNFNYLY..GLNCYYPLYRYGFYPTDGVGHQ
Pangolin CoV (GD) GRI..VGGNYNYLY..GFNCYFPLQSYGFHPTNGVGYQ
Pangolin CoV (GX) GVI..TGGNYGYLY..GLNCYYPLERYGFHPTTGVNYQ
RmYN02 GVI -IGS--YFY..DENGVRTLSTYDFNPNVPLDYQ
E
ACE2
K417
RBD
Y505
T500
N487
Q493
Y449
F
SARS-CoV-2 CASYQTQTNSPRRARSVASQSIIA
RmYN02 CASY----NSP-AAR-VGTNSIIA
RaTG13 CASYQTQTNS----RSVASQSIIA
Pangolin CoV (GD) CASYQTQTNS----RSVSS-AIIA
Pangolin CoV (GX) CASYHSMSSF----RSVNQRSIIA
CoVZC45 CASYHTASIL----RSTSQKAIVA
CoVZXC21 CASYHTASIL----RSTGQKAIVA

Figure 1.

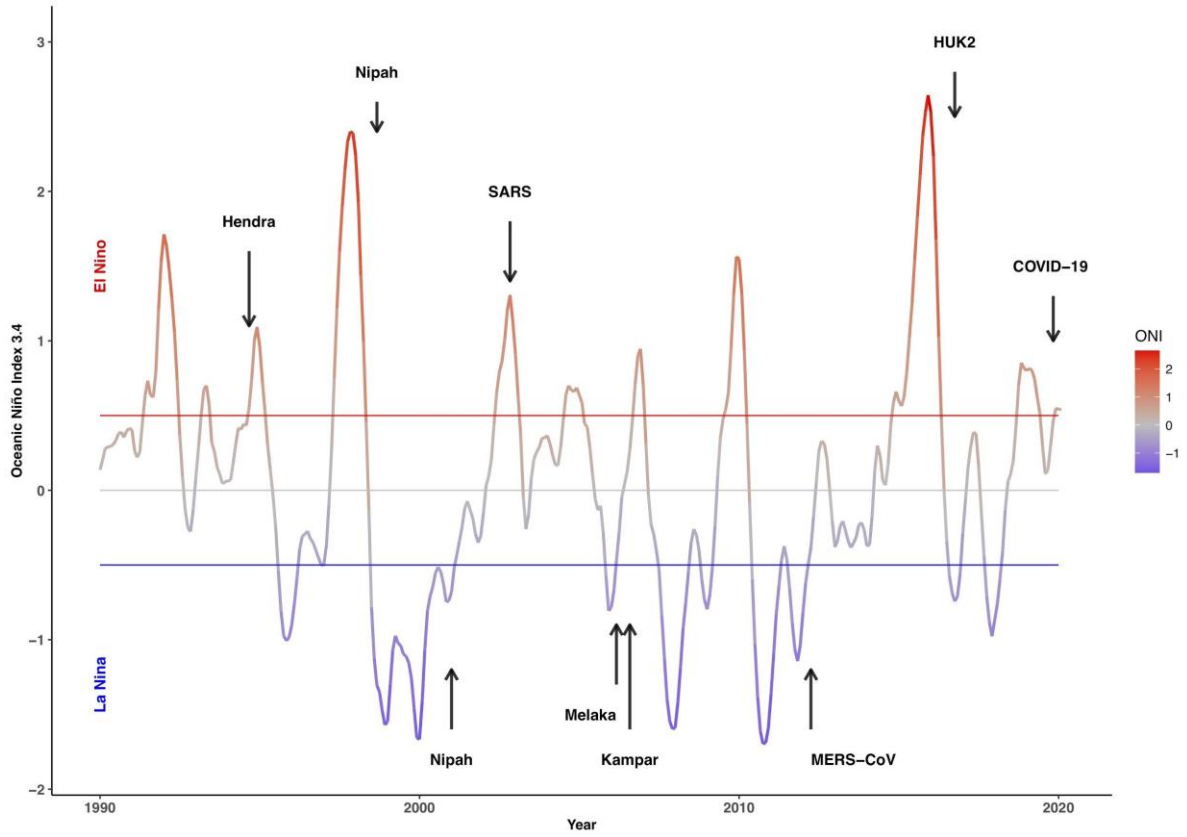
Oceanic Niño Index 3.4
3
2
1
0
−1
−2
El Nino
La Nina
Hendra
Nipah
SARS
HUK2
COVID−19
Nipah
Melaka
Kampar
MERS−CoV
1990
2000
2010
2020
Year
ONI
2
1
0
−1

Figure 2.

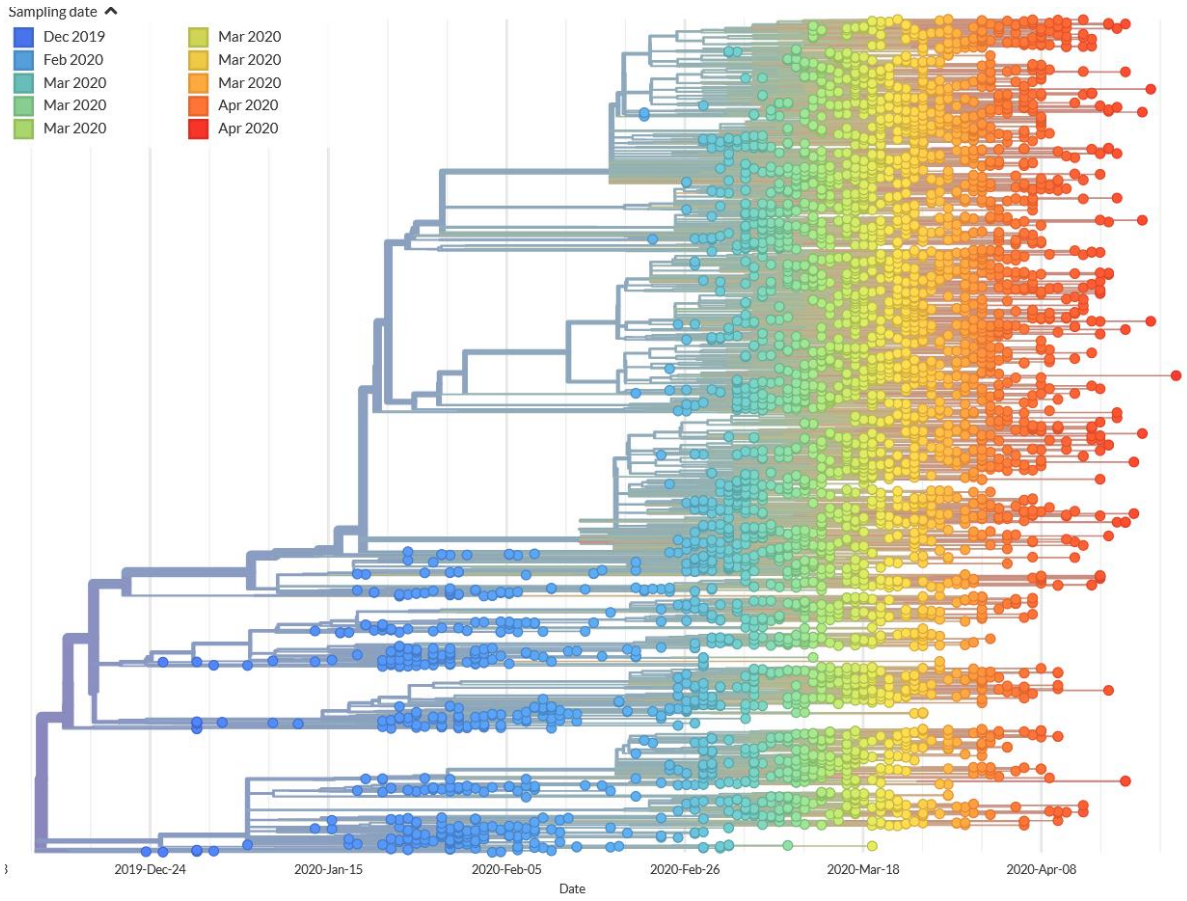
Sampling date
Dec 2019
Feb 2020
Mar 2020
Mar 2020
Mar 2020
Mar 2020
Mar 2020
Mar 2020
Apr 2020
Apr 2020
2019-Dec-24
2020-Jan-15
2020-Feb-05
2020-Feb-26
2020-Mar-18
2020-Apr-08
Date

Figure 3.

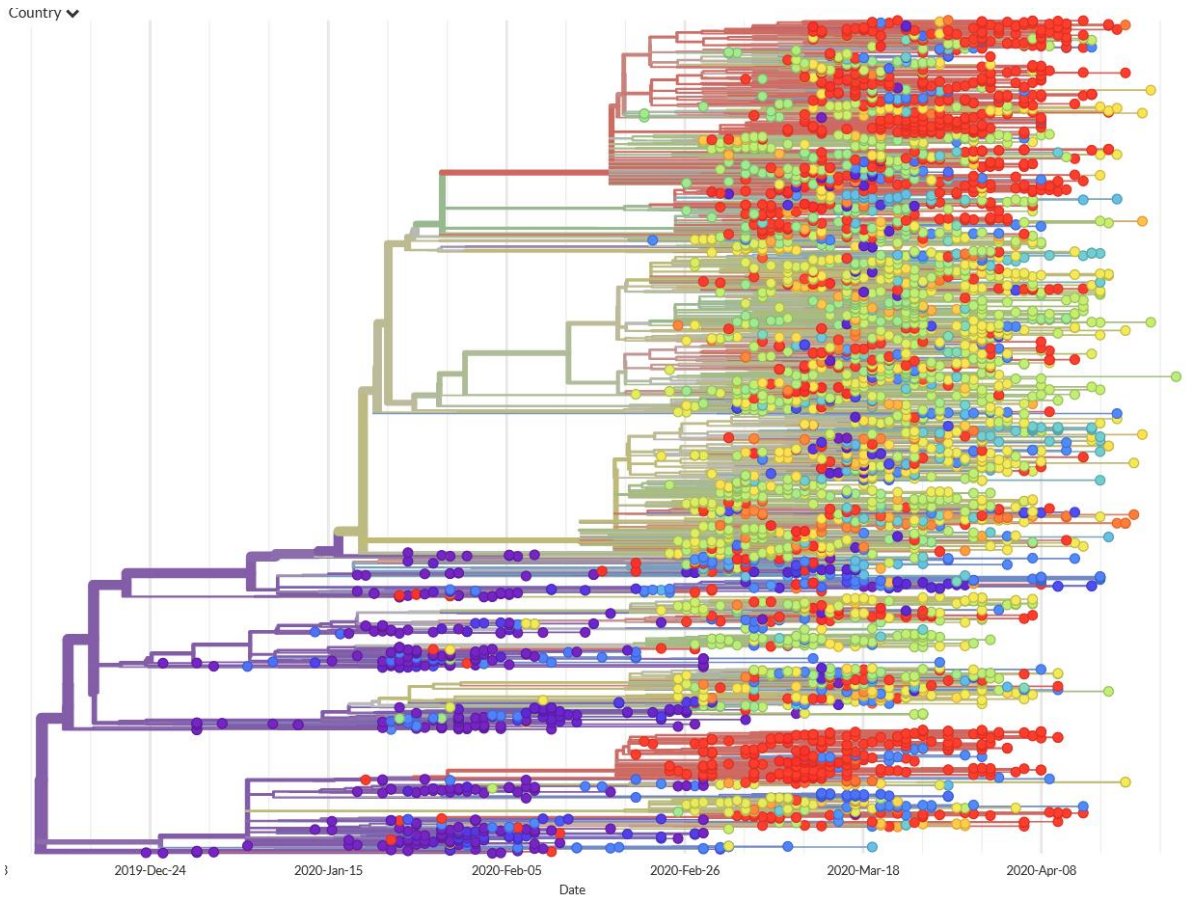
Country
2019-Dec-24
2020-Jan-15
2020-Feb-05
2020-Feb-26
2020-Mar-18
2020-Apr-08
Date

Figure 4.