Review

Origin and evolution of SARS-CoV-2

Isabel Pagani[1], **Silvia Ghezzi**[1], **Simone Alberti**[1], **Guido Poli**[2,3], **Elisa Vicenzi**[1,a]

[1] Viral Pathogenesis and Biosafety Unit, Division of Immunology, Transplantation, and Infectious Diseases, San Raffaele Scientific Institute, Via Olgettina, 58, Milan, Italy
[2] Human Immuno-Virology (H.I.V.) Unit, Division of Immunology, Transplantation, and Infectious Diseases, San Raffaele Scientific Institute, Via Olgettina, 58, Milan, Italy
[3] Vita-Salute San Raffaele University School of Medicine, Via Olgettina, 58, Milan, Italy

© The Author(s) 2023

Abstract SARS-CoV-2 is a novel coronavirus that emerged in China at the end of 2019 causing the severe disease known as coronavirus disease 2019 (COVID-19). SARS-CoV-2, as to the previously highly pathogenic human coronaviruses named SARS-CoV, the etiological agent of severe acute respiratory syndrome (SARS), has a zoonotic origin, although SARS-CoV-2 precise chain of animal-to-human transmission remains undefined. Unlike the 2002–2003 pandemic caused by SARS-CoV whose extinction from the human population was achieved in eight months, SARS-CoV-2 has been spreading globally in an immunologically naïve population in an unprecedented manner. The efficient infection and replication of SARS-CoV-2 has resulted in the emergence of viral variants that have become predominant posing concerns about their containment as they are more infectious with variable pathogenicity in respect to the original virus. Although vaccine availability is limiting severe disease and death caused by SARS-CoV-2 infection, its extinction is far to be close and predictable. In this regard, the emersion of the Omicron viral variant in November 2021 was characterized by humoral immune escape and it has reinforced the importance of the global monitoring of SARS-CoV-2 evolution. Given the importance of the SARS-CoV-2 zoonotic origin, it will also be crucial to monitor the animal-human interface to be better prepared to cope with future infections of pandemic potential.

1 Introduction

Since the beginning of 2020, our lives have been profoundly changed by the emergence of the worldwide SARS-CoV-2 spreading [1, 2] that can cause a disease known as Coronavirus disease 2019 (COVID-19) involving pneumonia and multiorgan failure [3, 4] with a significant fatality rate that varies among countries and estimated to be approximately 1–2% of all infected individuals who have not been vaccinated [5]. SARS-CoV-2 belongs to the large family of coronaviruses that are enveloped viruses, characterized by a single-stranded positive RNA genome of approximately 30,000 nucleotides in size, the largest known genome for RNA viruses [6]. Despite coronaviruses are widely present in the animal world, in humans, seven pathogenic coronaviruses have been identified in the last 20 years [7]. Out of these, three have raised special concern as they are highly pathogenic causing severe pneumonia and death. The first human coronavirus of public health concern was SARS-CoV that emerged in the Guangdong Province of China in 2002 [8, 9]. Approximately 8,000 cases and 700 deaths were reported during the virus's eight-month circulation in 29 countries and five continents, but, by the end of July 2003, the virus disappeared from the human population, thanks to the containment measures [10, 11].

In 2012, another coronavirus, named Middle East respiratory syndrome (MERS)-CoV, emerged in Saudi Arabia [12, 13]. Unlike SARS-CoV, this virus is still present in the human population, albeit infections are sporadic and limited to the Arabic peninsula with only 12 cases reported in 2021 (https://www.ecdc.europa.eu/en/publications-data/geographical-distribution-confirmed-mers-cov-cases-1-january-2021-2-august-2021).

At the end of 2019, a novel SARS-CoV-2 emerged in China. In contrast to the previous human coronaviruses, it has been spreading very efficiently by infecting more than 670 million individuals worldwide causing more than 6.8 million deaths at the time of this report (https://coronavirus.jhu.edu/map.html). Although both SARS-CoV and SARS-CoV-2 are airborne viruses [14], SARS-CoV is transmitted during the symptomatic phase of the infection whereas SARS-CoV-2 can be transmitted prior to the onset of symptoms and from asymptomatic individuals [15], who represent approximately 40–45% of all infected individuals [16], as

[a] e-mail: vicenzi.elisa@hsr.it (corresponding author)

Springer

Fig. 1 Scenario of SARS-CoV-2
natural origin
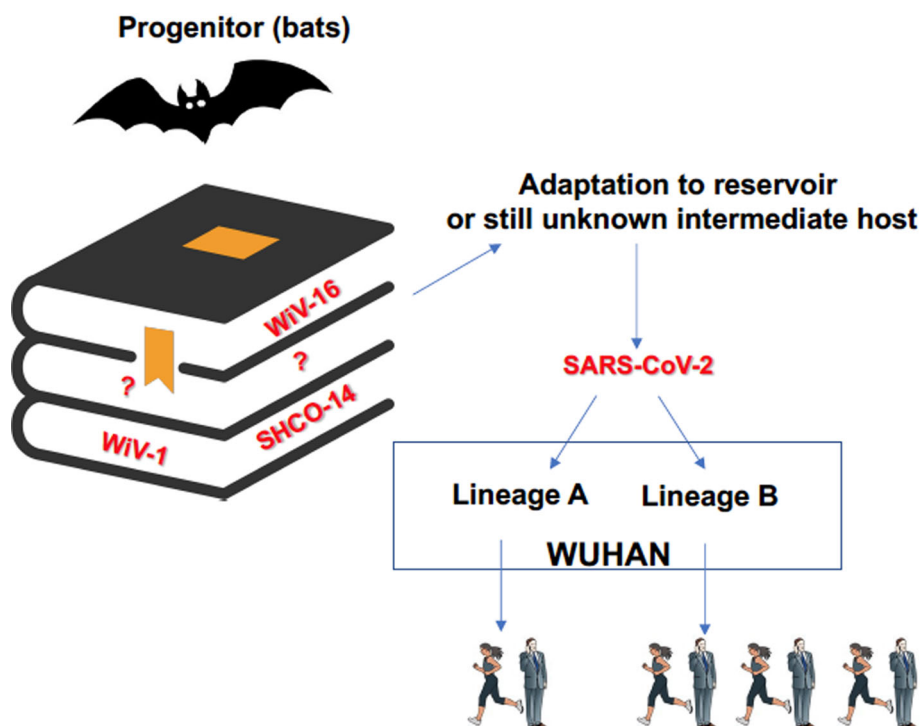
efficiently as symptomatic individuals. This SARS-CoV-2 peculiarity poses a challenge for the public health systems to contain the spreading of the infection. Importantly, however, effective vaccines have been developed in a record time (11 months) to provide global containment of this pandemic [17, 18].

2 Origin of SARS-CoV-2

All previous highly pathogenic human coronaviruses have had a zoonotic origin. SARS-CoV, which shares *ca*. 79% homology at the nucleotide sequence level with SARS-CoV-2, was linked to live animals sold at the markets in Foshan, Guangdong Province, China [19, 20]. Viral isolates sharing 99.8% homology with SARS-CoV were obtained from a few Himalayan palm civets and one raccoon dog and animal traders without a SARS diagnosis were found to have high levels of antibodies (Abs) against SARS-CoV in late 2002 [21]. However, several years of research surveying wild animals revealed that bats are infected with the ancestral viruses of SARS including SARS-CoV-2 and, thus, bats serve as a main reservoir of many SARS-like viruses [22]. The first evidence came from the isolation of the closest known relatives to SARS-CoV but also SARS-CoV-2 from horseshoe (*Rhinolophus*) bats that colonize caves in the Yunnan Province of mainland China [1, 23]. However, the notable geographic gap between the Yunnan caves and the location of the first emergence of SARS-CoV and SARS-CoV-2 highlights the difficulty of precisely trace the virus's movements and the importance to sampling other regions and countries. In this regard, viruses more closely related to SARS-CoV-2 than those from the Yunnan region [24] have been documented in bats but also in pangolins from multiple locations in South-East Asia, including China, Thailand, Cambodia and Japan [25–27]. As a significant evolution gap exists between SARS-CoV-2 and the closest-related animal viruses, it is unlikely that SARS-CoV-2 spilled over directly from bats to humans. Likely, the spillover event(s) occurred in the context of locations where high-density population is in contact with live animals susceptible to SARS-CoV-2 infection and sold alive for food in the Wuhan markets that were the early and major epicenters of SARS-CoV-2. However, these highly related viruses were not identified in the animals present in the Wuhan markets during the initial phase of COVID-19 pandemic (Fig. 1). This gap has fostered the speculation of a virus escape from the Wuhan Institute of Virology (WIV) where the research on SARS-related viruses were intensified after the 2002–2003 SARS-CoV global spread. In this regard, the analysis of the early Wuhan sequences identified two independent lineages of SARS-CoV-2 that were simultaneously distributed at different Wuhan wildlife markets and likely transmitted by a still unidentified wild-caught or farmed animals sold in the markets [28]. To fully demonstrate the natural origin of SARS-CoV-2, it will be fundamental to extensively survey the presence of coronaviruses in wild animals and studying their evolution and ecology.

Another important aspect of SARS-CoV-2 ecology is the cross-species transmission [29]. Indeed, SARS-CoV-2 can be transmitted from human-to-animal as documented in farmed minks [30], dogs and cats [31] as well as in lions and tigers in zoos [32]. Upon animal infection, the human virus may evolve and adapt to the new host. Viral evolution occurred in minks in the Netherlands and
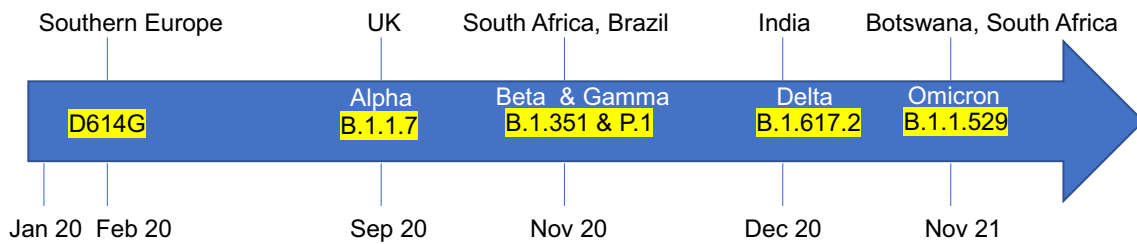
Fig. 2 Significant SARS-CoV-2 variants and countries and dates of first detection

Denmark [33] where the human-to-animal transfer resulted in the introduction of an adaptive mutation in the receptor binding domain (RBD) of the viral Spike protein. A tyrosine was replaced by phenylalanine at position 453 (Y453F) causing an increase in the Spike protein affinity for the viral entry receptor angiotensin converting enzyme-2 (ACE2) [33]. Furthermore, infection of animals with human SARS-CoV-2 may pose a serious problem as recombination events may occur among animal and human coronaviruses [34] to generate novel hybrid viruses with pandemic potential if they spillover in humans with no or partial immunity.

3 SARS-CoV-2 evolution

SARS-CoV-2, similarly to all RNA viruses, is prone to accumulating mutations in its genome during each cycle of replication. However, coronaviruses encode a 3′-to-5′-exonuclease that allows high fidelity replication by the viral RNA-dependent-RNA polymerase [35, 36]. This proofreading activity is thought to be necessary for maintaining the integrity of a long viral RNA genome and limiting the error rate imposed by viral RNA polymerases [37]. However, SARS-CoV-2 replicates very efficiently in infected hosts as demonstrated by the elevated viral loads measured in the nasopharyngeal swabs [38]. From December 2019 to November 2020, the evolution rate of SARS-CoV-2 suggests that its genome acquires approximately 2 mutations per month [39, 40]. However, over the course of the pandemic, several mutations have emerged, particularly in the Spike protein, implying that SARS-CoV-2 evolution rate is higher than predicted. By late February 2020 a single non-synonymous substitution of an aspartic acid into glycine was detected in the Spike protein at position 614 (D614G) [41] (Fig. 2).

Firstly, a minority of the sequences (5%) available in the data base, particularly in southern Europe, had the 614G, but with the spreading of the virus, this variant become the most prevalent genotype worldwide [41]. Importantly, in vitro studies have demonstrated that the glycine residue in position 614 confers a higher efficiency in viral entry and replication than that of Spikes with the aspartic acid residue [41, 42]. Furthermore, competition studies with original and mutant viruses generated by reverse genetics clearly demonstrated the higher fitness of the virus with the 614G than 614D [42]. These results were confirmed in vivo in a model of SARS-CoV-2 infection in Syrian hamsters [42]. Despite the increased infectivity and viral load in infected people, increased COVID-19 disease severity was not observed in association with D614G mutation, although infection rates were extended to younger patients [43]. Importantly, the D614G variant remained sensitive to the neutralizing activity of sera from individuals who have received vaccine preparation based on the original Wuhan Spike sequence and human monoclonal Abs targeting the Spike protein [44].

With the unfolding pandemic, an unprecedented amount of full genome sequencing data was made available via the Global Initiative on Sharing All Influenza Data (GISAID—https://www.gisaid.org/). In September 2020, the analysis of the continuous increase in genomic data highlighted the emergence of sequences characterized by a higher number of mutations relative to the previous circulating SARS-CoV-2. A new lineage defined as B.1.1.7 was identified in the United Kingdom (UK). This new variant was characterized by 14 non-synonymous and 6 synonymous substitutions and 3 deletions along the entire genome that mapped to a single monophyletic branch of the phylogenetic tree [43, 45]. Eight of these mutations, including substitutions and deletions, were positioned in the Spike protein that interacts with the cellular receptor to mediate viral entry into target cells [46–49]. Of note is the mutation in position 501 of the Spike protein where an asparagine residue is changed into a tyrosine residue (N501Y). This position is the principal contact residue in the RBD and it was shown to increase the Spike affinity for ACE2 by sixfold [50, 51].

The emergence of the B.1.1.7 variant has paved the way for more attention to the SARS-CoV-2 evolution as, by November 2020, the virus had started to mutate more drastically. Variants of concern (VOCs) with mutations of note in the Spike protein started to emerge, being characterized by increased transmissibility and changes in antigenicity, thus, raising the threat level to the public health. The three most notable VOCs came from South Africa (B.1.351) [52], Brazil (P.1) [53] and India (B.1.617.2) [54]. However, an abundance of variants of interest (VOIs) including B.1.525 (Nigeria), B.1.526 (New York), B.1.427/B.1.429 (California), B.1.258 (Scotland), and A.23.1 (Liverpool) have emerged which collectively now account for more than 90% of the sequenced viruses [55]. Given the complexity of the multiple variants, recently, the WHO has changed the nomenclature of VOCs and VOIs with the letters of the Greek alphabet re-naming B.1.1.7 as Delta, B.1.351 as Beta, P.1 as Gamma and B.1.617.2 as Delta (Fig. 2) whereas VOI were renamed for B.1.525 as Eta, B.1.526 as Iota, B.1.617.1 as Kappa, C.37 as Lambda and B.1.621 as Mu.

Worth noting, however, is the Delta variant that emerged in India at the end of 2020 [56] and since then it has been spreading in and outside India to become the predominant variant in Europe and worldwide [57, 58]. Epidemiological data indicate that the Delta variant is 40 to 60% more transmissible than Delta and, importantly, Delta variant replicates more efficiently than the previous variants as demonstrated by the viral loads in the nasopharyngeal swab reported to be *ca.* 1000-fold higher than that of the infection with the previous variants [59]. The higher efficiency of Delta to infect and replicate is explained by a set of three mutations in the Spike RBD. Firstly, a lysine to asparagine mutation at position 417 is present in some, but not all sequences. This mutation has been associated with conformational changes in the Spike protein that may result in immune escape [60]. The second mutation, a leucine to arginine change at position 452 increases the affinity for ACE2 receptor and the third, a threonine to lysine substitution at position 478, is common to the B.1.1.519 lineage, and it has been also predicted to increase RBD/ACE2 binding affinity and enable immune escape [54]. In addition, a proline to arginine substitution is present at position 681 adjacent to the furin cleavage site of the Spike protein that increases the efficiency of its cleavage to mediate viral fusion and initiate infection.

Unexpectedly, however, a new VOC, named B.1.1.529, was detected for the first time in Botswana and South Africa in November 2021 and later renamed Omicron [61]. This VOC has immediately raised concerns due to its high number of mutations, particularly in the Spike protein. Multiple changes have been detected within the two immunogenic regions in S1 (N-terminal domain (NTD and RBD) including a 3-amino acid insertion and an accumulation of mutations in proximity of the furin cleavage site including the combination of N679K and P681H mutations. Overall, these mutations render this virus more infectious than previous variants [62] with a short doubling time [63]. However, in vivo experimental infection of human ACE2 transgenic mice and Syrian hamsters has shown that Omicron causes attenuated disease as compared with D614G and Beta variants due to an impaired capacity to spread in the lower respiratory tract [64]. Importantly, unlike previous VOC, the Omicron mutations have lowered the neutralization activity of the Abs raised by vaccination by *ca.* 22-fold [65] rendering the human population more susceptible to infection. The rapid spread of Omicron has replaced the previous Delta variant in a few weeks worldwide and signs of viral evolution have been detected as determined by the emergence of sub-variants [66]. After the initial BA.1 emersion, the BA.2 sub-lineage has increased in many areas globally in a few weeks whereas the BA.3 has not taken off yet globally. While BA.2 and BA.3 share many of the changes present in BA.1, they are also characterized by unique mutations that define distinct groups in the phylogenetic analysis of multiple sequences [61]. Thus, the rapid spread, evolution and immune evasion of Omicron with its sub-variants continue to represent a concern for its containment.

4 Concluding remarks

SARS-CoV-2 is a novel coronavirus with a likely zoonotic origin albeit the precise spill over event(s) has not been elucidated. Animal-to-human transmission was documented for other highly pathogenic coronaviruses whereas human-to-animal transmission had been documented for SARS-CoV-2. This interspecies transmission has raised concern as coronaviruses can easily adapt to their host to become more infectious and can recombine with animal coronavirus generating novel coronaviruses with pandemic potential.

Regarding SARS-CoV-2 adaptation in humans, several variants named as VOCs and VOIs have emerged in a relatively short period of time. The mechanism under which VOCs and VOIs have emerged is unclear [67]. There is convincing evidence that the variants likely develop in chronic infection of immunocompromised hosts receiving convalescent sera [68–70]. In these conditions, a suboptimal immune response to the virus is a perfect milieu for the selection of variants characterized by higher efficiency of viral replication and immune evasion. The availability of an unprecedented sequence data set and their analysis have supported the presence of signatures of positive selection dependent on the immune response to the virus. VOCs and VOIs are characterized by several defining and converging mutations that have emerged in a relatively short period suggesting that the evolution rate of SARS-CoV-2 is several folds higher than at the beginning of the pandemic when the evolution rate was estimated to be two mutations per month. This continuous evolution that reminds the antigenic drift observed in Influenza A virus [29] is a topic of concern for the effectiveness of the current vaccines. A rapid and broad immunization campaign, particularly focusing on the poor regions of the world will be pivotal to contain the current SARS-CoV-2 pandemic.

Declarations

Conflict of interest The authors have no conflicts of interest to declare.

References

1. P. Zhou, X.L. Yang, X.G. Wang, B. Hu, L. Zhang, W. Zhang, H.R. Si, Y. Zhu, B. Li, C.L. Huang, H.D. Chen, J. Chen, Y. Luo, H. Guo, R.D. Jiang, M.Q. Liu, Y. Chen, X.R. Shen, X. Wang, X.S. Zheng, K. Zhao, Q.J. Chen, F. Deng, L.L. Liu, B. Yan, F.X. Zhan, Y.Y. Wang, G.F. Xiao, Z.L. Shi, Nature **579**(7798), 270 (2020)
2. Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K.S.M. Leung, E.H.Y. Lau, J.Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T.T.Y. Lam, J.T. Wu, G.F. Gao, B.J. Cowling, B. Yang, G.M. Leung, Z. Feng, N Engl J Med **382**(13), 1199 (2020)
3. N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao, W. Tan, I. China Novel Coronavirus, and T. Research, N Engl J Med **382**(8), 727 (2020)
4. E. Terpos, I. Ntanasis-Stathopoulos, I. Elalamy, E. Kastritis, T.N. Sergentanis, M. Politou, T. Psaltopoulou, G. Gerotziafas, M.A. Dimopoulos, Am J Hematol **95**(7), 834 (2020)
5. E. Petersen, M. Koopmans, U. Go, D.H. Hamer, N. Petrosillo, F. Castelli, M. Storgaard, S. Al Khalili, L. Simonsen, Lancet Infect Dis **20**(9), 238 (2020)
6. A. Saberi, A.A. Gulyaeva, J.L. Brubacher, P.A. Newmark, A.E. Gorbalenya, PLoS Pathog **14**(11), e1007314 (2018)
7. S.R. Weiss, J Exp Med **217**(5), e20200537 (2020)
8. T.G. Ksiazek, D. Erdman, C.S. Goldsmith, S.R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J.A. Comer, W. Lim, P.E. Rollin, S.F. Dowell, A.E. Ling, C.D. Humphrey, W.J. Shieh, J. Guarner, C.D. Paddock, P. Rota, B. Fields, J. DeRisi, J.Y. Yang, N. Cox, J.M. Hughes, J.W. LeDuc, W.J. Bellini, L.J. Anderson, S.W. Group, N Engl J Med **348**(20), 1953 (2003)
9. C. Drosten, S. Gunther, W. Preiser, S. van der Werf, H.R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R.A. Fouchier, A. Berger, A.M. Burguiere, J. Cinatl, M. Eickmann, N. Escriou, K. Grywna, S. Kramme, J.C. Manuguerra, S. Muller, V. Rickerts, M. Sturmer, S. Vieth, H.D. Klenk, A.D. Osterhaus, H. Schmitz, H.W. Doerr, N Engl J Med **348**(20), 1967 (2003)
10. E. Vicenzi, F. Canducci, D. Pinna, N. Mancini, S. Carletti, A. Lazzarin, C. Bordignon, G. Poli, M. Clementi, Emerg Infect Dis **10**(3), 413 (2004)
11. L.F. Wang, B.T. Eaton, Curr Top Microbiol Immunol **315**, 325 (2007)
12. A.M. Zaki, S. van Boheemen, T.M. Bestebroer, A.D. Osterhaus, R.A. Fouchier, N Engl J Med **367**(19), 1814 (2012)
13. A. Zumla, D.S. Hui, S. Perlman, The Lancet **386**(9997), 995 (2015)
14. M. Klompas, M.A. Baker, C. Rhee, JAMA **324**(5),441 (2020)
15. R. Wölfel, V.M. Corman, W. Guggemos, M. Seilmaier, S. Zange, M.A. Müller, D. Niemeyer, T.C. Jones, P. Vollmar, C. Rothe, M. Hoelscher, T. Bleicker, S. Brünink, J. Schneider, R. Ehmann, K. Zwirglmaier, C. Drosten, C. Wendtner, Nature **581**(7809), 465(2020)
16. D.P. Oran, E.J. Topol, Ann Intern Med **173**(5), 362 (2020)
17. F.P. Polack, S.J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J.L. Perez, G. Perez Marc, E.D. Moreira, C. Zerbini, R. Bailey, K.A. Swanson, S. Roychoudhury, K. Koury, P. Li, W.V. Kalina, D. Cooper, R.W. Frenck, L.L. Hammitt, O. Tureci, H. Nell, A. Schaefer, S. Unal, D.B. Tresnan, S. Mather, P.R. Dormitzer, U. Sahin, K. U. Jansen, W. C. Gruber, C. C. T. Group, N Engl J Med **383**(27), 2603 (2020)
18. L.R. Baden, H.M. El Sahly, B. Essink, K. Kotloff, S. Frey, R. Novak, D. Diemert, S.A. Spector, N. Rouphael, C.B. Creech, J. McGettigan, S. Khetan, N. Segall, J. Solis, A. Brosz, C. Fierro, H. Schwartz, K. Neuzil, L. Corey, P. Gilbert, H. Janes, D. Follmann, M. Marovich, J. Mascola, L. Polakowski, J. Ledgerwood, B.S. Graham, H. Bennett, R. Pajon, C. Knightly, B. Leav, W. Deng, H. Zhou, S. Han, M. Ivarsson, J. Miller, T. Zaks, and C.S. Group, N Engl J Med **384**(5), 403 (2021)
19. Y. Guan, B.J. Zheng, Y.Q. He, X.L. Liu, Z.X. Zhuang, C.L. Cheung, S.W. Luo, P.H. Li, L.J. Zhang, Y.J. Guan, K.M. Butt, K.L. Wong, K.W. Chan, W. Lim, K.F. Shortridge, K.Y. Yuen, J.S. Peiris, L.L. Poon, Science **302**(5643), 276 (2003)
20. H.D. Song, C.C. Tu, G.W. Zhang, S.Y. Wang, K. Zheng, L.C. Lei, Q.X. Chen, Y.W. Gao, H.Q. Zhou, H. Xiang, H.J. Zheng, S.W. Chern, F. Cheng, C.M. Pan, H. Xuan, S.J. Chen, H.M. Luo, D.H. Zhou, Y.F. Liu, J.F. He, P.Z. Qin, L.H. Li, Y.Q. Ren, W.J. Liang, Y.D. Yu, L. Anderson, M. Wang, R.H. Xu, X.W. Wu, H.Y. Zheng, J.D. Chen, G. Liang, Y. Gao, M. Liao, L. Fang, L.Y. Jiang, H. Li, F. Chen, B. Di, L.J. He, J.Y. Lin, S. Tong, X. Kong, L. Du, P. Hao, H. Tang, A. Bernini, X.J. Yu, O. Spiga, Z.M. Guo, H.Y. Pan, W.Z. He, J.C. Manuguerra, A. Fontanet, A. Danchin, N. Niccolai, Y.X. Li, C.I. Wu, G.P. Zhao, Proc Natl Acad Sci USA **102**(7), 2430 (2005)
21. C. Centers for Disease and Prevention, MMWR Morb Mortal Wkly Rep **52**(41), 986 (2003)
22. W. Li, Z. Shi, M. Yu, W. Ren, C. Smith, J.H. Epstein, H. Wang, G. Crameri, Z. Hu, H. Zhang, J. Zhang, J. McEachern, H. Field, P. Daszak, B.T. Eaton, S. Zhang, L.F. Wang, Science **310**(5748), 676 (2005)
23. B. Hu, L.P. Zeng, X.L. Yang, X.Y. Ge, W. Zhang, B. Li, J.Z. Xie, X.R. Shen, Y.Z. Zhang, N. Wang, D.S. Luo, X.S. Zheng, M.N. Wang, P. Daszak, L.F. Wang, J. Cui, Z.L. Shi, PLoS Pathog **13**(11), e1006698 (2017)
24. H. Zhou, J. Ji, X. Chen, Y. Bi, J. Li, Q. Wang, T. Hu, H. Song, R. Zhao, Y. Chen, M. Cui, Y. Zhang, A.C. Hughes, E.C. Holmes, W. Shi, Cell **184**(17), 4380 (2021)
25. S. Wacharapluesadee, C.W. Tan, P. Maneeorn, P. Duengkae, F. Zhu, Y. Joyjinda, T. Kaewpom, W.N. Chia, W. Ampoot, B.L. Lim, K. Worachotsueptrakun, V.C. Chen, N. Sirichan, C. Ruchisrisarod, A. Rodpan, K. Noradechanon, T. Phaichana, N. Jantarat, B. Thongnumchaima, C. Tu, G. Crameri, M.M. Stokes, T. Hemachudha, L.F. Wang, Nat Commun **12**(1), 972 (2021)
26. V. Hul, D. Delaune, E.A. Karlsson, A. Hassanin, P.O. Tey, A. Baidaliuk, F. Gámbaro, V.T. Tu, L. Keatts, J. Mazet, C. Johnson, P. Buchy, P. Dussart, T. Goldstein, E. Simon-Lorière, V. Duong, bioRxiv, 2021.01.26.428212 (2021)
27. S. Mallapaty, Nature **588**(7836), 15 (2020)
28. R.F. Garry. https://virological.org/t/early-appearance-of-two-distinct-genomic-lineages-of-sars-cov-2-in-different-wuhan-wildlife-markets-suggests-sars-cov-2-has-a-natural-origin/691 (2021)
29. A. Telenti, A. Arvin, L. Corey, D. Corti, M.S. Diamond, A. Garcia-Sastre, R.F. Garry, E.C. Holmes, P.S. Pang, H.W. Virgin, Nature **596**(7873), 495 (2021)
30. N. Oreshkova, R.J. Molenaar, S. Vreman, F. Harders, B.B. Oude Munnink, R.W. Hakze-van der Honing, N. Gerhards, P. Tolsma, R. Bouwstra, R.S. Sikkema, M.G. Tacken, M.M. de Rooij, E. Weesendorp, M.Y. Engelsma, C.J. Bruschke, L.A. Smit, M. Koopmans, W.H. van der Poel, A. Stegeman, Euro Surveill **25**(23), 2001005 (2020)
31. E.I. Patterson, G. Elia, A. Grassi, A. Giordano, C. Desario, M. Medardo, S.L. Smith, E.R. Anderson, T. Prince, G.T. Patterson, E. Lorusso, M.S. Lucente, G. Lanave, S. Lauzi, U. Bonfanti, A. Stranieri, V. Martella, F. Solari Basano, V.R. Barrs, A.D. Radford, U. Agrimi, G.L. Hughes, S. Paltrinieri, N. Decaro, Nat Commun **11**(1), 6231 (2020)
32. D. McAloose, M. Laverack, L. Wang, M.L. Killian, L.C. Caserta, F. Yuan, P.K. Mitchell, K. Queen, M.R. Mauldin, B.D. Cronk, S.L. Bartlett, J.M. Sykes, S. Zec, T. Stokol, K. Ingerman, M.A. Delaney, R. Fredrickson, M. Ivancic, M. Jenkins-Moore, K. Mozingo, K. Franzen, N.H. Bergeson, L. Goodman, H. Wang, Y. Fang, C. Olmstead, C. McCann, P. Thomas, E. Goodrich, F. Elvinger, D.C. Smith, S. Tong, S. Slavinski, P.P. Calle, K. Terio, M.K. Torchetti, D.G. Diel, mBio **11**(5) (2020)

33. B.B. Oude Munnink, R.S. Sikkema, D.F. Nieuwenhuijse, R.J. Molenaar, E. Munger, R. Molenkamp, A. van der Spek, P. Tolsma, A. Rietveld, M. Brouwer, N. Bouwmeester-Vincken, F. Harders, R. Hakze-van der Honing, M.C.A. Wegdam-Blans, R.J. Bouwstra, C. GeurtsvanKessel, A.A. van der Eijk, F.C. Velkers, L.A.M. Smit, A. Stegeman, W.H.M. van der Poel, M.P.G. Koopmans, Science **371**(6525), 172 (2021)

34. R.L. Graham, R.S. Baric, J Virol **84**(7), 3134 (2010)

35. M.R. Denison, R.L. Graham, E.F. Donaldson, L.D. Eckerle, R.S. Baric, RNA Biol **8**(2), 270 (2011)

36. E.C. Smith, N.R. Sexton, M.R. Denison, Annu Rev Virol **1**(1), 111 (2014)

37. E.C. Holmes, Trends Microbiol **11**(12), 543 (2003)

38. R. Wolfel, V.M. Corman, W. Guggemos, M. Seilmaier, S. Zange, M.A. Muller, D. Niemeyer, T.C. Jones, P. Vollmar, C. Rothe, M. Hoelscher, T. Bleicker, S. Brunink, J. Schneider, R. Ehmann, K. Zwirglmaier, C. Drosten, C. Wendtner, Nature **581**(7809), 465 (2020)

39. S. Duchene, L. Featherstone, M. Haritopoulou-Sinanidou, A. Rambaut, P. Lemey, G. Baele, Virus Evol **6**(2), veaa061 (2020)

40. M. Worobey, J. Pekar, B.B. Larsen, M.I. Nelson, V. Hill, J.B. Joy, A. Rambaut, M.A. Suchard, J.O. Wertheim, P. Lemey, Science **370**(6516), 564 (2020)

41. B. Korber, W.M. Fischer, S. Gnanakaran, H. Yoon, J. Theiler, W. Abfalterer, N. Hengartner, E.E. Giorgi, T. Bhattacharya, B. Foley, K.M. Hastie, M.D. Parker, D.G. Partridge, C.M. Evans, T.M. Freeman, T.I. de Silva, C.-G.G. Sheffield, C. McDanal, L.G. Perez, H. Tang, A. Moon-Walker, S.P. Whelan, C.C. LaBranche, E.O. Saphire, D.C. Montefiori, Cell **182**(4), 812 (2020)

42. J.A. Plante, Y. Liu, J. Liu, H. Xia, B.A. Johnson, K.G. Lokugamage, X. Zhang, A.E. Muruato, J. Zou, C.R. Fontes-Garfias, D. Mirchandani, D. Scharton, J.P. Bilello, Z. Ku, Z. An, B. Kalveram, A.N. Freiberg, V.D. Menachery, X. Xie, K.S. Plante, S.C. Weaver, P.Y. Shi, Nature **592**(7852), 116 (2021)

43. E. Volz, S. Mishra, M. Chand, J.C. Barrett, R. Johnson, L. Geidelberg, W.R. Hinsley, D.J. Laydon, G. Dabrera, Á. O'Toole, R. Amato, M. Ragonnet-Cronin, I. Harrison, B. Jackson, C.V. Ariani, O. Boyd, N.J. Loman, J.T. McCrone, S. Gonçalves, D. Jorgensen, R. Myers, V. Hill, D.K. Jackson, K. Gaythorpe, N. Groves, J. Sillitoe, D.P. Kwiatkowski, COVID-19 Genomics UK (COG-UK) consortium, S. Flaxman, O. Ratmann, S. Bhatt, S. Hopkins, A. Gandy, A. Rambaut, N.M. Ferguson. Nature **593**(7858), 266 (2021)

44. Y. Liu, J. Liu, H. Xia, X. Zhang, C.R. Fontes-Garfias, K.A. Swanson, H. Cai, R. Sarkar, W. Chen, M. Cutler, D. Cooper, S.C. Weaver, A. Muik, U. Sahin, K.U. Jansen, X. Xie, P.R. Dormitzer, P.Y. Shi, N Engl J Med **384**(15), 1466 (2021)

45. R. Rambaut, N.J. Loman, O.G. Pybus, W. Barclay, J. Barrett, A. Carabelli, T. Connor, T. Peacock, D.L. Robertson, E. Volz. https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563. (2020)

46. M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Kruger, T. Herrler, S. Erichsen, T.S. Schiergens, G. Herrler, N.H. Wu, A. Nitsche, M.A. Muller, C. Drosten, S. Pohlmann, Cell **181**(2), 271 (2020)

47. D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.L. Hsieh, O. Abiona, B.S. Graham, J.S. McLellan, Science **367**(6483), 1260 (2020)

48. F.A. Lempp, L. Soriaga, M. Montiel-Ruiz, F. Benigni, J. Noack, Y.J. Park, S. Bianchi, A.C. Walls, J.E. Bowen, J. Zhou, H. Kaiser, A. Joshi, M. Agostini, M. Meury, E. Dellota, Jr., S. Jaconi, E. Cameroni, J. Martinez-Picado, J. Vergara-Alert, N. Izquierdo-Useros, H.W. Virgin, A. Lanzavecchia, D. Veesler, L. Purcell, A. Telenti, D. Corti, Nature **598**(7880), 342 (2021)

49. A.C. Walls, Y.J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, D. Veesler, Cell **181**(2), 281 (2020)

50. D.A. Collier, A. De Marco, I. Ferreira, B. Meng, R.P. Datir, A.C. Walls, S.A. Kemp, J. Bassi, D. Pinto, C. Silacci-Fregni, S. Bianchi, M.A. Tortorici, J. Bowen, K. Culap, S. Jaconi, E. Cameroni, G. Snell, M.S. Pizzuto, A.F. Pellanda, C. Garzoni, A. Riva, C.-N. B. C.-. Collaboration, A. Elmer, N. Kingston, B. Graves, L.E. McCoy, K.G.C. Smith, J.R. Bradley, N. Temperton, L. Ceron-Gutierrez, G. Barcenas-Morales, C.-G.U. Consortium, W. Harvey, H.W. Virgin, A. Lanzavecchia, L. Piccoli, R. Doffinger, M. Wills, D. Veesler, D. Corti, R.K. Gupta, Nature **593**(7857), 136 (2021)

51. H. Gu, Q. Chen, G. Yang, L. He, H. Fan, Y.Q. Deng, Y. Wang, Y. Teng, Z. Zhao, Y. Cui, Y. Li, X.F. Li, J. Li, N.N. Zhang, X. Yang, S. Chen, Y. Guo, G. Zhao, X. Wang, D.Y. Luo, H. Wang, X. Yang, Y. Li, G. Han, Y. He, X. Zhou, S. Geng, X. Sheng, S. Jiang, S. Sun, C.F. Qin, Y. Zhou, Science **369**(6511), 1603 (2020)

52. H. Tegally, E. Wilkinson, R.R. Lessells, J. Giandhari, S. Pillay, N. Msomi, K. Mlisana, J. Bhiman, M. Allam, A. Ismail, S. Engelbrecht, G. Van Zyl, W. Preiser, C. Williamson, F. Pettruccione, A. Sigal, I. Gazy, D. Hardie, M. Hsiao, D. Martin, D. York, D. Goedhals, E.J. San, M. Giovanetti, J. Lourenco, L.C.J. Alcantara, T. de Oliveira, Nat Med **27**(3), 440 (2021)

53. D.S. Candido, I.M. Claro, J.G. de Jesus, W.M. Souza, F.R.R. Moreira, S. Dellicour, T.A. Mellan, L. du Plessis, R.H.M. Pereira, F.C.S. Sales, E.R. Manuli, J. Theze, L. Almeida, M.T. Menezes, C.M. Voloch, M.J. Fumagalli, T.M. Coletti, C.A.M. da Silva, M.S. Ramundo, M.R. Amorim, H.H. Hoeltgebaum, S. Mishra, M.S. Gill, L.M. Carvalho, L.F. Buss, C.A. Prete, Jr., J. Ashworth, H.I. Nakaya, P.S. Peixoto, O.J. Brady, S.M. Nicholls, A. Tanuri, A.D. Rossi, C.K.V. Braga, A.L. Gerber, C.G.A. P. de, N. Gaburo, Jr., C.S. Alencar, A.C.S. Ferreira, C.X. Lima, J.E. Levi, C. Granato, G.M. Ferreira, R.S. Francisco, Jr., F. Granja, M.T. Garcia, M.L. Moretti, M.W. Perroud, Jr., T. Castineiras, C.S. Lazari, S.C. Hill, A.A. de Souza Santos, C.L. Simeoni, J. Forato, A.C. Sposito, A.Z. Schreiber, M.N.N. Santos, C.Z. de Sa, R.P. Souza, L.C. Resende-Moreira, M.M. Teixeira, J. Hubner, P.A.F. Leme, R.G. Moreira, M.L. Nogueira, D.G. Brazil-Uk Centre for Arbovirus Discovery, N. Epidemiology Genomic, N.M. Ferguson, S.F. Costa, J.L. Proenca-Modena, A.T.R. Vasconcelos, S. Bhatt, P. Lemey, C.H. Wu, A. Rambaut, N.J. Loman, R.S. Aguiar, O.G. Pybus, E.C. Sabino, N.R. Faria, Science **369**(6508), 1255 (2020)

54. P. Mlcochova, S. Kemp, M.S. Dhar, G. Papa, B. Meng, I. Ferreira, R. Datir, D. A. Collier, A. Albecka, S. Singh, R. Pandey, J. Brown, J. Zhou, N. Goonawardane, S. Mishra, C. Whittaker, T. Mellan, R. Marwal, M. Datta, S. Sengupta, K. Ponnusamy, V. S. Radhakrishnan, A. Abdullahi, O. Charles, P. Chattopadhyay, P. Devi, D. Caputo, T. Peacock, D. C. Wattal, N. Goel, A. Satwik, R. Vaishya, M. Agarwal, S.-C.-G. C. Indian, C. Genotype to Phenotype Japan, C.-N. B. C.-. Collaboration, A. Mavousian, J. H. Lee, J. Bassi, C. Silacci-Fegni, C. Saliba, D. Pinto, T. Irie, I. Yoshida, W.L. Hamilton, K. Sato, S. Bhatt, S. Flaxman, L.C. James, D. Corti, L. Piccoli, W.S. Barclay, P. Rakshit, A. Agrawal, R.K. Gupta, Nature **599**(7883), 114 (2021)

55. A. Rambaut, E.C. Holmes, A. O'Toole, V. Hill, J.T. McCrone, C. Ruis, L. du Plessis, O.G. Pybus, Nat Microbiol **5**(11), 1403 (2020)

56. S. Cherian, V. Potdar, S. Jadhav, P. Yadav, N. Gupta, M. Das, P. Rakshit, S. Singh, P. Abraham, S. Panda, and Nic team, Microorganisms **9**(7), 1542 (2021)

57. I. Torjesen, BMJ **373**, n1445 (2021)

58. C. Chakraborty, A. R. Sharma, M. Bhattacharya, G. Agoramoorthy, and S. S. Lee, mBio **12**(4), e0114021 (2021)

59. B. Li, A. Deng, K. Li, Y. Hu, Z. Li, Q. Xiong, Z. Liu, Q. Guo, L. Zou, H. Zhang, M. Zhang, F. Ouyang, J. Su, W. Su, J. Xu, H. Lin, J. Sun, J. Peng, H. Jiang, P. Zhou, T. Hu, M. Luo, Y. Zhang, H. Zheng, J. Xiao, T. Liu, R. Che, H. Zeng, Z. Zheng, Y. Huang, J. Yu, L. Yi, J. Wu, J. Chen, H. Zhong, X. Deng, M. Kang, O. G. Pybus, M. Hall, K. A. Lythgoe, Y. Li, J. Yuan, J. He, and J. Lu, Nat Commun **13**(1), 460 (2022)

60. L. Yurkovetskiy, X. Wang, K.E. Pascal, C. Tomkins-Tinch, T.P. Nyalile, Y. Wang, A. Baum, W.E. Diehl, A. Dauphin, C. Carbone, K. Veinotte, S.B. Egri, S.F. Schaffner, J.E. Lemieux, J.B. Munro, A. Rafique, A. Barve, P.C. Sabeti, C.A. Kyratsous, N.V. Dudkina, K. Shen, J. Luban, Cell **183**(3), 739 (2020)

61. R. Viana, S. Moyo, D.G. Amoako, H. Tegally, C. Scheepers, C.L. Althaus, U.J. Anyaneji, P.A. Bester, M.F. Boni, M. Chand, W.T. Choga, R. Colquhoun, M. Davids, K. Deforche, D. Doolabh, L. du Plessis, S. Engelbrecht, J. Everatt, J. Giandhari, M. Giovanetti, D. Hardie, V. Hill, N.Y. Hsiao, A. Iranzadeh, A. Ismail, C. Joseph, R. Joseph, L. Koopile, S.L. Kosakovsky Pond, M.U.G. Kraemer, L. Kuate-Lere, O. Laguda-Akingba, O. Lesetedi-Mafoko, R.J. Lessells, S. Lockman, A.G. Lucaci, A. Maharaj, B. Mahlangu, T. Maponga, K. Mahlakwane, Z. Makatini, G. Marais, D. Maruapula, K. Masupu, M. Matshaba, S. Mayaphi, N. Mbhele, M.B. Mbulawa, A. Mendes, K. Mlisana, A. Mnguni, T. Mohale, M. Moir, K. Moruisi, M. Mosepele, G. Motsatsi, M.S. Motswaledi, T. Mphoyakgosi, N. Msomi, P.N. Mwangi, Y. Naidoo, N. Ntuli, M. Nyaga, L. Olubayo, S. Pillay, B. Radibe, Y. Ramphal, U. Ramphal,

J.E. San, L. Scott, R. Shapiro, L. Singh, P. Smith-Lawrence, W. Stevens, A. Strydom, K. Subramoney, N. Tebeila, D. Tshiabuila, J. Tsui, S. van Wyk, S. Weaver, C.K. Wibmer, E. Wilkinson, N. Wolter, A.E. Zarebski, B. Zuze, D. Goedhals, W. Preiser, F. Treurnicht, M. Venter, C. Williamson, O.G. Pybus, J. Bhiman, A. Glass, D.P. Martin, A. Rambaut, S. Gaseitsiwe, A. von Gottberg, T. de Oliveira, Nature **603**(7902), 679 (2022)

62. E. Callaway, H. Ledford, Nature **600**(7888), 197 (2021)
63. S.S.A. Karim, Q.A. Karim, Lancet **398**(10317), 2126 (2021)
64. P.J. Halfmann, S. Iida, K. Iwatsuki-Horimoto, T. Maemura, M. Kiso, S.M. Scheaffer, T.L. Darling, A. Joshi, S. Loeber, G. Singh, S.L. Foster, B. Ying, J.B. Case, Z. Chong, B. Whitener, J. Moliva, K. Floyd, M. Ujie, N. Nakajima, M. Ito, R. Wright, R. Uraki, P. Warang, M. Gagne, R. Li, Y. Sakai-Tagawa, Y. Liu, D. Larson, J.E. Osorio, J.P. Hernandez-Ortiz, A.R. Henry, K. Ciouderis, K.R. Florek, M. Patel, A. Odle, L. R. Wong, A.C. Bateman, Z. Wang, V.V. Edara, Z. Chong, J. Franks, T. Jeevan, T. Fabrizio, J. DeBeauchamp, L. Kercher, P. Seiler, A. S. Gonzalez-Reiche, E.M. Sordillo, L.A. Chang, H. van Bakel, V. Simon, g. Consortium Mount Sinai Pathogen Surveillance study, D.C. Douek, N.J. Sullivan, L.B. Thackray, H. Ueki, S. Yamayoshi, M. Imai, S. Perlman, R.J. Webby, R.A. Seder, M.S. Suthar, A. Garcia-Sastre, M. Schotsaert, T. Suzuki, A.C.M. Boon, M.S. Diamond, Y. Kawaoka, Nature **603**(7902), 687 (2022)
65. S. Cele, L. Jackson, D.S. Khoury, K. Khan, T. Moyo-Gwete, H. Tegally, J.E. San, D. Cromer, C. Scheepers, D.G. Amoako, F. Karim, M. Bernstein, G. Lustig, D. Archary, M. Smith, Y. Ganga, Z. Jule, K. Reedoy, S.H. Hwa, J. Giandhari, J.M. Blackburn, B.I. Gosnell, S.S. Abdool Karim, W. Hanekom, S.A. Ngs, C.-K. Team, A. von Gottberg, J.N. Bhiman, R.J. Lessells, M.S. Moosa, M.P. Davenport, T. de Oliveira, P.L. Moore, A. Sigal, Nature **602**(7898), 654 (2022)
66. S. Iketani, L. Liu, Y. Guo, L. Liu, Y. Huang, M. Wang, Y. Luo, J. Yu, M.T. Yin, M.E. Sobieszczyk, Y. Huang, H.H. Wang, Z. Sheng, D.D. Ho, Nature **604**(7906), 553 (2022)
67. W.T. Harvey, A.M. Carabelli, B. Jackson, R.K. Gupta, E.C. Thomson, E.M. Harrison, C. Ludden, R. Reeve, A. Rambaut, C.-G.U. Consortium, S.J. Peacock, D.L. Robertson, Nat Rev Microbiol **19**(7), 409 (2021)
68. S.A. Kemp, D.A. Collier, R.P. Datir, I. Ferreira, S. Gayed, A. Jahun, M. Hosmillo, C. Rees-Spear, P. Mlcochova, I.U. Lumb, D.J. Roberts, A. Chandra, N. Temperton, C.-N. B. C.-. Collaboration, C.-G. U. Consortium, K. Sharrocks, E. Blane, Y. Modis, K.E. Leigh, J.A.G. Briggs, M.J. van Gils, K.G.C. Smith, J.R. Bradley, C. Smith, R. Doffinger, L. Ceron-Gutierrez, G. Barcenas-Morales, D.D. Pollock, R.A. Goldstein, A. Smielewska, J.P. Skittrall, T. Gouliouris, I.G. Goodfellow, E. Gkrania-Klotsas, C.J.R. Illingworth, L.E. McCoy, R.K. Gupta, Nature **592**(7853), 277 (2021)
69. B. Choi, M.C. Choudhary, J. Regan, J.A. Sparks, R.F. Padera, X. Qiu, I.H. Solomon, H.H. Kuo, J. Boucau, K. Bowman, U.D. Adhikari, M.L. Winkler, A.A. Mueller, T.Y. Hsu, M. Desjardins, L.R. Baden, B.T. Chan, B.D. Walker, M. Lichterfeld, M. Brigl, D.S. Kwon, S. Kanjilal, E.T. Richardson, A.H. Jonsson, G. Alter, A.K. Barczak, W.P. Hanage, X.G. Yu, G.D. Gaiha, M.S. Seaman, M. Cernadas, J.Z. Li, N Engl J Med **383**(23), 2291 (2020)
70. V.A. Avanzato, M.J. Matson, S.N. Seifert, R. Pryce, B.N. Williamson, S.L. Anzick, K. Barbian, S.D. Judson, E.R. Fischer, C. Martens, T.A. Bowden, E. de Wit, F.X. Riedo, V.J. Munster, Cell **183**(7), 1901 (2020)