



# Occurrence of a substitution or deletion of SARS-CoV-2 spike amino acid 677 in various lineages in Marseille, France

Philippe Colson<sup>1,2</sup> · Jeremy Delerce<sup>1</sup> · Emilie Burel<sup>1</sup> · Mamadou Beye<sup>1</sup> · Pierre-Edouard Fournier<sup>1,2</sup> · Anthony Levasseur<sup>1,2</sup> · Jean-Christophe Lagier<sup>1,2</sup> · Didier Raoult<sup>1,2</sup>

Received: 25 June 2021 / Accepted: 21 October 2021 / Published online: 27 November 2021  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

## Abstract

Great concerns have been raised about SARS-CoV-2 variants over the past six months. At the end of 2020, an increasing incidence of spike substitutions Q677H/P was described in the USA, which involved six independent lineages. We searched for changes to this amino acid in the sequence database of SARS-CoV-2 genomes obtained at the IHU Méditerranée Infection (Marseille, France) from 3634 patients sampled between February 2020 and April 2021. In seven genomes (0.2%), we found a deletion of five amino acids at spike positions 675–679 (QTQTN) including Q677, and in 76 genomes (2.3%) we found a Q677H substitution. The 83 genomes were classified in ten different Pangolin lineages. Genomes with a spike Q677 deletion were obtained from respiratory samples collected in six cases between 28 March 2020 and 12 October 2020 and in one case on 1 February 2021. The Q677H substitution was found in genomes all obtained from respiratory samples collected from 19 January 2021 and were classified in seven different lineages. Most of these genomes (41 cases) were of UK variant. Two others were classified in the B.1.160 Pangolin lineage (Marseille-4 variant) which was first detected in July 2020 in our institute but was devoid of this substitution until 19 January 2021. Also, eight genomes were classified in the A.27/Marseille-501 lineage which was first detected in our institute in January 2021 and which either harboured or did not harbour the Q677H substitution. Thus, the spike Q677H substitution should be considered as another example of convergent evolution, as it is the case of spike substitutions L18F, E484K, L452R, and N501Y which also independently appeared in various lineages.

**Keywords** SARS-CoV-2 · COVID-19 · Variant · Q677 · Emergence

## Introduction

One year after SARS-CoV-2 was declared a pandemic by the WHO, some of its variants with various combinations of amino acid substitutions or deletions in the spike protein have taken centre stage [1]. This protein leads to virus entry into human respiratory cells [2, 3]. It is also the major target of neutralising antibodies which are elicited post-infection

or vaccine immunisation, and is the target of most vaccine strategies implemented to date [4]. As a result of our surveillance of genomic epidemiology, our institute has observed and described the emergence of a dozen SARS-CoV-2 variants since summer 2020, after the rate of diagnoses fell to almost zero for several weeks between May and June 2020. These included the Marseille-1 and Marseille-4 variants later assigned to Pangolin lineages B.1.416 and B.1.160, respectively, and the Marseille-501 variant corresponding to lineage A.27 [5–8]. The variants recently considered to be of greatest concern are those carrying amino acid substitutions N501Y and/or E484K within the spike protein [9, 10], as they have increased affinity for the ACE2 cellular receptor, decreased sensitivity to neutralising antibodies, and may escape the immune responses elicited by the vaccines currently used in Western countries [2, 3, 11]. Nevertheless, various other SARS-CoV-2 variants have been reported as emerging worldwide.

Edited by William Dundon.

✉ Didier Raoult  
didier.raoult@gmail.com

<sup>1</sup> IHU Méditerranée Infection, 19-21 boulevard Jean Moulin, 13005 Marseille, France

<sup>2</sup> Microbes Evolution Phylogeny and Infections (MEPHI), Aix-Marseille University, Institut de Recherche pour le Développement (IRD), Assistance Publique - Hôpitaux de Marseille (AP-HM), 27 boulevard Jean Moulin, 13005 Marseille, France

At the end of 2020, the incidence of variants carrying the Q677H, or Q677P substitutions in the spike increased, mainly in the USA, where the first sequences originated from Louisiana [12]. Interestingly, these strains were reported as belonging to six independent sublineages, including one, two, and three in Nextstrain clades 20A, 20B, and 20G, respectively. Amino acid 677 of the SARS-CoV-2 spike protein is only three codons upstream of the polybasic/furin cleavage site of S1/S2 spike domains, which is critical for SARS-CoV-2 pathogenesis, as this cleavage induces a spike conformational change that favours the binding to the Angiotensin-Converting Enzyme 2 (ACE2) cellular SARS-CoV-2 receptor and may enhance viral infection [12–16]. We looked in our SARS-CoV-2 genome sequence database which we have been contributing to since February 2020, in order to analyse the prevalence and genotypic patterns of viruses mutated at codon 677 of the spike.

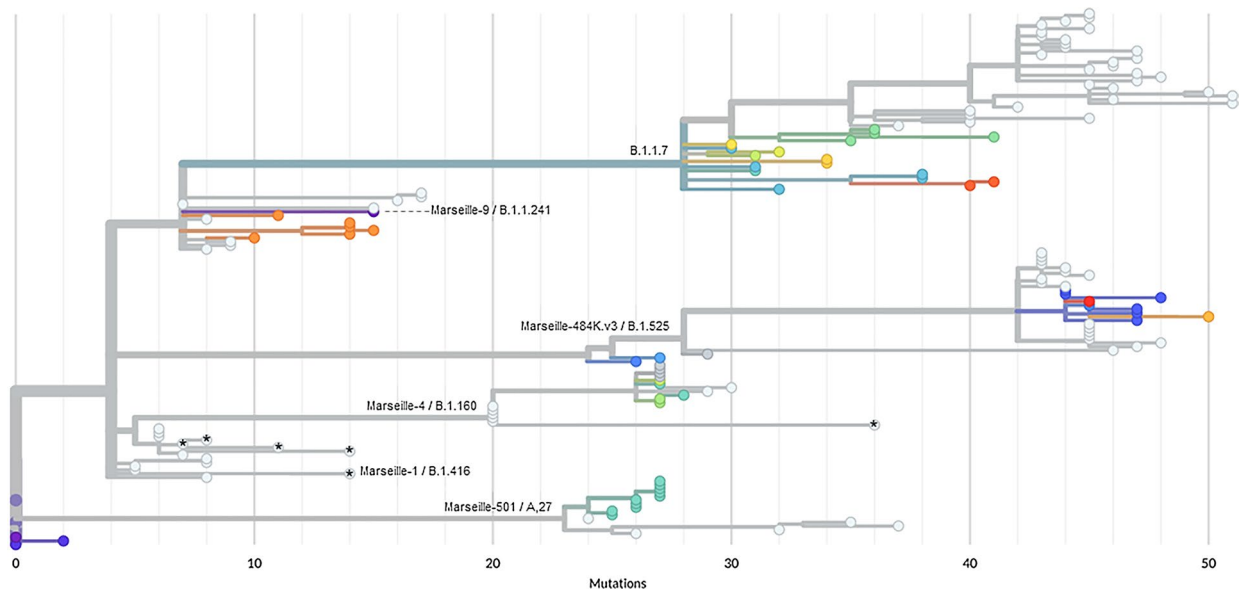
## Results and discussion

SARS-CoV-2 genome sequences analysed here were obtained by next-generation sequencing from viral RNA extracted from 200 µL of nasopharyngeal swab fluid collected from patients for the diagnosis of SARS-CoV-2 infection, using the EZ1 Virus Mini Kit v2.0 on an EZ1 Advanced XL instrument (Qiagen, Courtaboeuf, France) or the KingFisher Flex system (Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer's recommendations [17]. Extracted viral RNA was reverse-transcribed using SuperScript IV (ThermoFisher Scientific) prior to cDNA second strand synthesis with Klenow Fragment DNA polymerase (New England Biolabs, Beverly, MA, USA), with the LunaScript RT SuperMix kit (New England Biolabs), or according to the COVIDSeq protocol (Illumina Inc.). Next-generation sequencing was performed as previously described [17] using either the Illumina Nextera XT paired-end procedure on a MiSeq instrument (Illumina Inc., San Diego, CA, USA), until mid-March 2021; the Oxford Nanopore technology (ONT) on a GridION instrument (Oxford Nanopore Technologies Ltd., Oxford, UK) with cDNA amplification using a multiplex PCR protocol with the ARTIC nCoV-2019 V3 Panel primers (Integrated DNA technologies, Coralville, IA, USA) according to the ARTIC procedure (<https://artic.network/>), between mid-March and mid-April 2021; or the Illumina COVIDSeq protocol on a NovaSeq 6000 instrument (Illumina Inc), since mid-April. Genome consensus sequences were obtained as previously described [17] by mapping on the SARS-CoV-2 genome GenBank accession no. NC\_045512.2 (Wuhan-Hu-1 isolate) with the CLC Genomics workbench v.7 (<https://digitalinsights.qiagen.com/>) or the Minimap2 software [18]. Sequences described in the present study have been

deposited on the GISAID sequence database (<https://www.gisaid.org/>) [19] (Supplementary Table S1). The phylogenetic tree was based on the SARS-CoV-2 genomes obtained in our laboratory and on the 67 genomes the most similar to them retrieved using the GISAID BLAST tool (<https://www.epicov.org/epi3/frontend#4ee9c>) from the GISAID database [18], and on additional reference genomes corresponding to major SARS-CoV-2 variants or to the Wuhan-Hu-1 isolate. The tree was built using the Nextstrain tool (<https://docs.nextstrain.org/projects/ncov/en/latest/index.html>) [20] that performs maximum-likelihood phylogeny using IQ-TREE [21] then visualised with Auspice (<https://docs.nextstrain.org/projects/auspice/en/stable/>). The nature and number of nucleotide changes in the SARS-CoV-2 genomes and of amino acid changes in the SARS-CoV-2 proteins were obtained using the Nextclade tool (<https://clades.nextstrain.org/results>). Classification into Marseille variants was performed using an in house Python script based on sets of mutations. Classification into Nextstrain clades was performed with the Nextclade online tool at URL: <https://clades.nextstrain.org/>) and classification into Pangolin lineages was performed using the Pangolin online tool at URL: <https://cov-lineages.org/pangolin.html> [22].

A codon change or deletion was detected in 83 (2.5%) of 3364 SARS-CoV-2 genomes obtained from respiratory samples collected from different patients between February 2020 and April 2021. These 83 genomes were classified in ten different Pangolin lineages, namely A, A.27 (Marseille-501), B.1, B.1.1, B.1.1.10, B.1.1.241 (Marseille-9), B.1.1.7, B.1.160 (Marseille-4), B.1.416 (Marseille-1), B.1.525 (Marseille-484 K.v3) in three, six, four, three, three, one, 38, three, one, and 21 cases, respectively (Fig. 1; Supplementary Table S1).

In seven of the 3634 patients (0.2%), a deletion of five amino acids was observed at spike protein positions 675–679 (QTQTN) which include amino acid 677 (Supplementary Table S1). These sequences were obtained from respiratory samples collected in six cases between 28 March 2020 and 12 October 2020 and in one case on 1 February 2021. They involved three, one, and three viral strains classified in Nextstrain clades 20A, 20B, and 20G, respectively. Pangolin lineages were B.1 ( $n = 4$  cases), B.1.416 (1), B.1.1.241 (1), and B.1.160 (1). The genome obtained in 2021 was a Marseille-4 variant/B.1.160 [7]. This deletion of spike amino acids 675–679 (QTQTN) was reported to abrogate the enzymatic cleavage of the S protein and the efficiency of both TMPRSS2 and TMPRSS13 for facilitating this cleavage, and to possibly restrict late-phase viral replication in Vero E6 cells [23–25]. In addition, it was reported to be present at a greater frequency from culture supernatant, in 12 (50%) out of 24 isolated viruses, than from respiratory samples, in three (4%) of 68 cases [24]. This suggests that selection pressures on SARS-CoV-2 regarding infection and/



**Fig. 1** Phylogenetic tree of SARS-CoV-2 genomes harbouring a deletion or substitution from Q to H of spike amino acid 677 and obtained from patients diagnosed with SARS-CoV-2 in our institute. The phylogenetic tree was built using the Nextstrain tool (<https://docs.nextstrain.org/projects/ncov/en/latest/index.html>) [20] that performs maximum-likelihood phylogeny using IQ-TREE [21], and was visualised with Auspice (<https://docs.nextstrain.org/projects/auspice/en/stable/>). The tree incorporated 69 of the SARS-CoV-2 genomes described here and obtained in our laboratory; 67 genomes corresponding to the

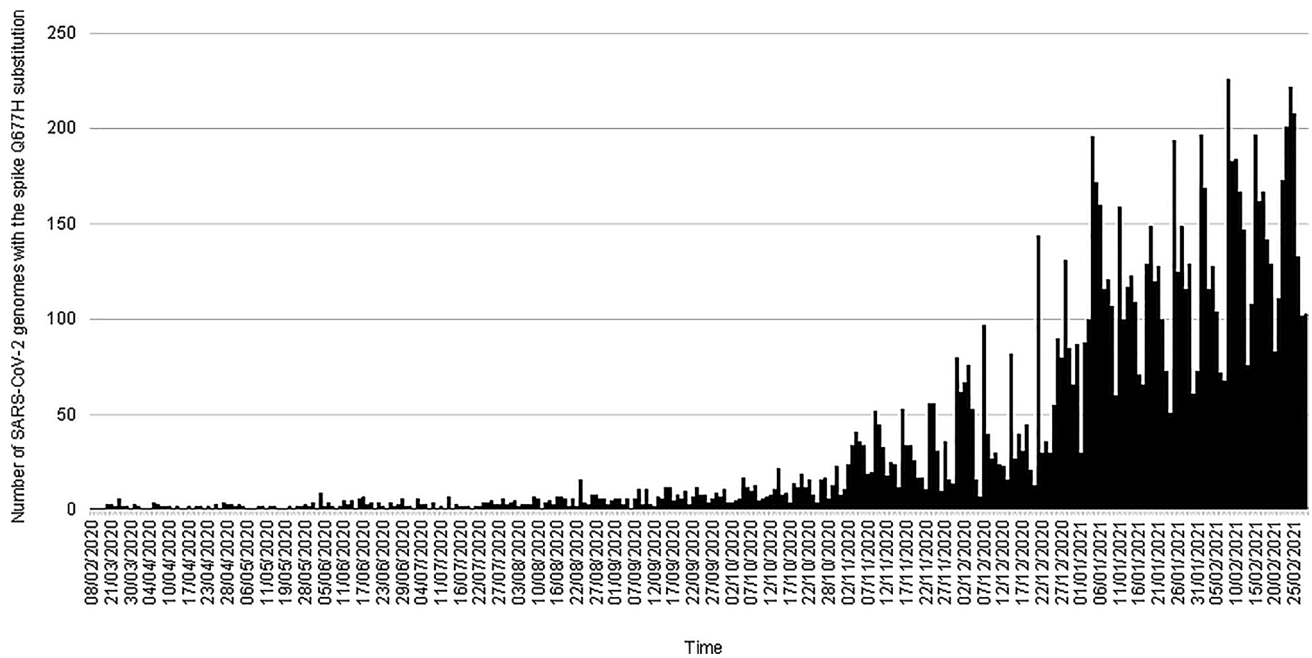
best hits of those obtained in the present study, retrieved using the GISAID BLAST tool (<https://www.epicov.org/epi3/frontend#4ee9c>) from the GISAID database [18]; and additional reference genomes corresponding to major SARS-CoV-2 variants or to the Wuhan-Hu-1 isolate. SARS-CoV-2 genomes harbouring a Q677 deletion are indicated by a black asterisk. This figure is adapted from screenshots of an output of the Nextstrain tool (<https://docs.nextstrain.org/projects/ncov/en/latest/index.html>) [20]. Sequences described in the present study are labelled with the light grey colour

or replication differ in vitro and in vivo. In addition, the rare identification of this deletion in the clinical samples in this latter study and in the present work further suggests that this QTQTN sequence is under strong purifying selection in vivo. It is also worthy to note that the SARS-CoV-2-like bat strain RmYN02 was reported to exhibit a similar QTQT deletion in the spike protein [26].

In 76 patients (2.3%), a spike Q677H substitution was observed in the 3634 SARS-CoV-2 genomes obtained in our institute. These changes involved viral genomes that were all obtained from respiratory samples collected from 19 January 2021, and were classified as belonging to seven different Pangolin lineages (Supplementary Table S1). A majority of these genomes were of Pangolin lineage B.1.1.7, in 41 cases (54% of the 76 genomes and 7.2% of all B.1.1.7 genomes in our database of 3364 SARS-CoV-2 genomes). In addition, two were classified in the Marseille-4/B.1.160 lineage that was first detected in our institute in July 2020 [7] and accounted for 573 (17%) genomes in our database, although its members did not harbour this spike Q677H substitution before 19 January 2021. Also, eight genomes were classified in the Marseille-501/A.27 lineage that was first detected in our institute in January 2021, accounted for 18 genomes in our database and which may or may not harbour Q677H substitution [8]. Codon changes from CAG to

CAT or CAC leading to this Q677H substitution deoptimise the viral codon usage relative to that of the human genome. Indeed, in the human genome, CAG, CAT, and CAC usage frequencies are 34.2, 10.9, and 15.1, respectively (<https://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=9606>), which represent 3.1-fold and 2.3-fold decreases. Worldwide, 13659 SARS-CoV-2 genomes were found to encode this amino acid substitution Q677H according to the CoV-GLUE online tool (<http://cov-glue.cvr.gla.ac.uk/>) [27] (Fig. 2). They were classified in 229 different Pangolin lineages, the majority (77%) being classified in lineages B.1.2 ( $n = 5537$ ), B.1.1.7 (1434), B.1.525 (952), B.1.170 (818), B.1.234 (714), B.1.1.316 (560), and B.1.1.284 (512). These genomes were obtained from clinical samples collected in 84 countries, mostly in the USA (89%;  $n = 7801$ ), England (1866), Japan (610), Denmark (497), Canada (379), Switzerland (314), Germany (309), India (194), and Egypt (122).

Therefore, the spike Q677H substitution should be considered as another example of convergent evolution, in addition to spike amino acid substitutions N501Y [28], L452R [29], and L18F [30] which also independently appeared in various lineages. Q677H is notably part of the substitutions that emerged in strains of the B.1.1.7 [31] and B.1.351 [32] variants of concern. Here, we report the presence of this substitution in SARS-CoV-2 of B.1.146/Marseille-1,



**Fig. 2** Numbers of SARS-CoV-2 genomes harbouring the spike Q677H mutation worldwide according to timeline. Data were collected from the Cov-Glue online tool (<http://cov-glue.cvr.gla.ac.uk/>) [27]

B.1.160/Marseille-4, B.1.1.241/Marseille-9, and B.1.525/Marseille-484 K.v3 lineages, which are different *bona fide* variants [17] although all B.1-derived lineages. In addition, we report the presence of the spike Q677H substitution in the A.27/Marseille-501 variant, which is a A-like lineage [8]. Therefore, we found this substitution in lineages that separated at different stages, including early, in the pandemic. Congruently, Hodcroft et al. reported the emergence late 2020 of variants harbouring substitutions Q677H/P in the spike of viruses that were classified into sublineages of Nextstrain clades 20B and 20G as well as into Nextstrain clade 20A [12], and the Q677H substitution has been identified in the B.1.525 lineage first described in Nigeria [33] and as a fast growing mutation in variants of concern [31].

Such convergent evolution is deemed to be the result of positive selection, and suggests that these amino acid substitutions at spike position 677 confer an evolutionary advantage to the virus [28]. The Q677H substitution in the SARS-CoV-2 spike protein is in the close vicinity of the polybasic RRAR furin-cleavage site of the spike S1/S2 boundary, possibly impacting binding between the spike receptor binding domain and ACE2 [12, 15, 16]. It has been hypothesised that histidine protonation in Q677H could induce a conformational switch that may affect the accessibility to protease of this site, which may enhance the cleavage at the S1/S2 junction and viral entry efficiency [12, 15, 16]. Structural analyses demonstrated that the spike amino acid 677 was located in subdomain SD2 of each of the protomers forming the homotrimeric spike protein, at the beginning of a very

flexible loop (residues 675–690) and in the very vicinity of two experimentally observed O-glycosylation sites at T676 and T678 [34–36]. Zeng et al. used a luciferase-bearing lentiviral pseudotype-based neutralisation assay to assess the effect of this spike Q677H substitution [37]. They did not evidence a significant effect on the cleavage of the spike protein. In contrast, they reported that this amino acid substitution increased viral infectivity and syncytium formation. In addition, they reported that when present in variants of concern of the B.1.1.7 and P1 lineages, it increased viral infectivity by 150 and 26%, respectively, and it decreased susceptibility to neutralisation by serum samples from recipients of the Moderna mRNA-1273 and Pfizer BNT162b2 vaccines of 22 and 29%, respectively. Finally, viral neutralisation in the presence of a monoclonal conformation-dependent antibody targeting the spike receptor binding domain led to a 50% reduction in neutralisation of 677H-harbouring spike relative to 677Q-harbouring spike (wild type), and of Q677H/D614G-harbouring spike relative to D614G only-harbouring spike. This suggested an alteration of the conformation of the spike receptor binding domain. Thus, overall, the functional consequences of the spike Q677H substitution and its epistatic interactions with other amino acid substitutions located inside or outside the receptor binding domain of the spike protein are currently not precisely deciphered. Taking into account the growing prevalence of this substitution in distinct SARS-CoV-2 variants of interest or of concern worldwide and its possible impact on viral infectivity and immune escape, previous data justify carrying out more



investigations and monitoring its evolution through genomic surveillance.

Combined with previous findings, present data highlight the great genetic variability of SARS-CoV-2 and warrant broader genomic surveillance of SARS-CoV-2 in order to gain a better insight of the epidemiology and evolution of this virus at the global, national, and local scales.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11262-021-01877-2>.

**Author contributions** Conceptualization, DR and PC; methodology, PC, and AL; investigation, PC, JD, EB, MB, PEF, AL, and JCL; writing, PC and DR. All authors have read and agreed to the published version of the manuscript.

**Funding** This work was supported by the French Government under the “Investments for the Future” program managed by the National Agency for Research (ANR), Méditerranée-Infection 10-IAHU-03 and was also supported by Région Provence Alpes Côte d’Azur and European funding FEDER PRIMMI (Fonds Européen de Développement Régional-Plateformes de Recherche et d’Innovation Mutualisées Méditerranée Infection), FEDER PA 0000320 PRIMMI.

## Declarations

**Conflict of interest** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**Ethical approval** This study has been approved by the ethics committee of our institution (N°2020-016-03).

## References

- Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S et al (2020) Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* 581:215–220
- Greaney AJ, Loes AN, Crawford KHD, Starr TN, Malone KD, Chu HY et al (2021) Comprehensive mapping of mutations to the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human serum antibodies. *Cell Host Microbe* 29:463–476
- Cele S, Gazy I, Jackson L, Hwa SH, Tegally H, Lustig G, Giandhari J, Pillay S et al (2021) Escape of SARS-CoV-2 501Y.V2 from neutralization by convalescent plasma. *Nature* 593:142–146
- Cohen JI, Burbelo PD (2020) Reinfection with SARS-CoV-2: implications for vaccines. *Clin Infect Dis*. <https://doi.org/10.1093/cid/ciaa1866>
- Fournier PE, Colson P, Levasseur A, Gautret P, Luciani L, Bedotto M et al (2021) Genome sequence analysis enabled deciphering the atypical evolution of COVID-19 in Marseille, France. *IHU Pre-Prints*. <https://doi.org/10.35088/kmct-tj43>
- Colson P, Levasseur A, Gautret P, Fenollar F, Thuan HV, Delerce J et al (2021) Introduction into the Marseille geographical area of a mild SARS-CoV-2 variant originating from sub-Saharan Africa: an investigational study. *Travel Med Infect Dis* 40:101980
- Fournier PE, Colson P, Levasseur A, Devaux CA, Gautret P, Bedotto M et al (2021) Emergence and outcomes of the SARS-CoV-2 “Marseille-4” variant. *Int J Infect Dis* 106:228–236
- Colson P, Levasseur A, Delerce J, Pinault L, Dudouet P, Devaux C et al (2021) Spreading of a new SARS-CoV-2 N501Y spike variant in a new lineage. *Clin Microbiol Infect*. <https://doi.org/10.1016/j.cmi.2021.05.006>
- Faria NR, Mellan TA, Whittaker C, Claro IM, Candido DDS, Mishra S et al (2021) Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* 372:815–821
- Singh J, Rahman SA, Ehtesham NZ, Hira S, Hasnain SE (2021) SARS-CoV-2 variants of concern are emerging in India. *Nat Med*. <https://doi.org/10.1038/s41591-021-01397-4>
- Wang P, Liu L, Iketani S, Luo Y, Guo Y, Wang M et al (2021) Increased resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7 to antibody neutralization. *Cell Host Microbe* 29:747–751
- Hodcroft EB, Domman DB, Oguntuyo K, Snyder DJ, Diest MV, Densmore KH et al (2021) Emergence in late 2020 of multiple lineages of SARS-CoV-2 spike protein variants affecting amino acid position 677. *medRxiv*. <https://doi.org/10.1101/2021.02.12.21251658>
- Johnson BA, Xie X, Kalveram B, Lokugamage KG, Muruato A, Zou J et al (2020) Furin cleavage site is key to SARS-CoV-2 pathogenesis. *bioRxiv*. <https://doi.org/10.1101/2020.08.26.268854>
- Xia S, Lan Q, Su S, Wang X, Xu W, Liu Z et al (2020) The role of furin cleavage site in SARS-CoV-2 spike protein-mediated membrane fusion in the presence or absence of trypsin. *Signal Transduct Target Ther* 5:92
- Wrobel AG, Benton DJ, Xu P, Roustan C, Martin SR, Rosenthal PB et al (2020) SARS-CoV-2 and bat RaTG13 spike glyco-protein structures inform on virus evolution and furin-cleavage effects. *Nat Struct Mol Biol* 27:763–767
- Hoffmann M, Kleine-Weber H, Pöhlmann S (2020) A multibasic cleavage site in the spike protein of SARS-CoV-2 is essential for infection of human lung cells. *Mol Cell* 78:779–784
- Colson P, Fournier PE, Chaudet H, Delerce J, Giraud-Gatineau A, Houhamdi L et al (2021) Analysis of SARS-CoV-2 variants from 24,181 patients exemplifies the role of globalisation and zoonosis in pandemics. *medRxiv*. <https://doi.org/10.1101/2021.09.10.21262922>
- Li H (2018) Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100
- Alm E, Broberg EK, Connor T, Hodcroft EB, Komissarov AB, Maurer-Stroh S et al (2020) Geographical and temporal distribution of SARS-CoV-2 clades in the WHO European Region, January to June 2020. *Euro Surveill* 20(25):2001410
- Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, Sagulenko P, Bedford T, Neher RA (2018) Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 34:4121–4123
- Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274
- Rambaut A, Holmes EC, O’Toole Á, Hill V, McCrone JT, Ruis C et al (2020) A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol* 5:1403–1407
- Cantuti-Castelvetri L, Ojha R, Pedro LD, Djannatian M, Franz J, Kuivanen S et al (2020) Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. *Science* 370:856–860
- Liu Z, Zheng H, Lin H, Li M, Yuan R, Peng J et al (2020) Identification of common deletions in the spike protein of severe acute respiratory syndrome coronavirus 2. *J Virol* 94:e00790–e820
- Laporte M, Stevaert A, Raeymaekers V, Vandeput J, Marchand-Casas I, Thibaut HJ et al (2021) The SARS-CoV-2 and other human coronavirus spike proteins are fine-tuned towards

- temperature and proteases of the human airways. *PLoS Pathog* 17:e1009500
26. Zhou H, Chen X, Hu T, Li J, Song H, Liu Y et al (2020) A novel bat coronavirus reveals natural insertions at the S1/S2 cleavage site of the spike protein and a possible recombinant origin of HCoV-19. *bioRxiv*. <https://doi.org/10.1101/2020.03.02.974139>
  27. Singer JB, Gifford RJ, Cotten M, Robertson DL (2020) CoV-GLUE: a web application for tracking SARS-CoV-2 genomic variation. *Preprints* 2020:2020060225
  28. Martin D, Weaver S, Tegally H, San JE, Shank SD, Wilkinson E et al (2021) The emergence and ongoing convergent evolution of the N501Y lineages coincided with a major global shift in the SARS-Cov-2 selective landscape. *medRxiv*. <https://doi.org/10.1101/2021.02.23.21252268>
  29. Tchesnokova V, Kulakesara H, Larson L, Bowers V, Rechkina E, Kisiela D et al (2021) Acquisition of the L452R mutation in the ACE2-binding interface of spike protein triggers recent massive expansion of SARS-Cov-2 variants. *bioRxiv*. <https://doi.org/10.1101/2021.02.22.432189>
  30. Grabowski F, Kochanczyk M, Lipniacki T (2021) L18F substrain of SARS-CoV-2 VOC-202012/01 is rapidly spreading in England. *medRxiv*. <https://doi.org/10.1101/2021.02.07.21251262>
  31. Grabowski F, Preibisch G, Gizinski S, Kochanczyk M, Lipniacki T (2021) SARS-CoV-2 variant of concern 202012/01 has about twofold replicative advantage and acquires concerning mutations. *Viruses* 13:392
  32. Edara VV, Norwood C, Floyd K, Lai L, Davis-Gardner ME, Hudson WH et al (2021) Infection- and vaccine-induced antibody binding and neutralization of the B.1.351 SARS-CoV-2 variant. *Cell Host Microbe* 29:516–521
  33. Ozer EA, Simons LM, Adewumi OM, Fowotade AA, Omoruyi EC, Adeniji JA et al (2021) Coincident rapid expansion of two SARS-CoV-2 lineages with enhanced infectivity in Nigeria. *medRxiv*. <https://doi.org/10.1101/2021.04.09.21255206>
  34. Bagdonaitė I, Thompson AJ, Wang X, Søggaard M, Fougeroux C, Frank M et al (2021) Site-specific O-glycosylation analysis of SARS-CoV-2 spike protein produced in insect and human cells. *Viruses* 13:551
  35. Sanda M, Morrison L, Goldman R (2021) N- and O-glycosylation of the SARS-CoV-2 spike protein. *Anal Chem* 93:2003–2009
  36. Li X, Zhang L, Chen S, Ji W, Li C, Ren L (2021) Recent progress on the mutations of SARS-CoV-2 spike protein and suggestions for prevention and controlling of the pandemic. *Infect Genet Evol* 93:104971
  37. Zeng C, Evans JP, Faraone JN, Qu P, Zheng YM, Saif L et al (2021) Neutralization of SARS-CoV-2 variants of concern harboring Q677H. *MBio* 12:e0251021

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.