

REVIEW

The variants question: What is the problem?

Daive Zella¹  | Marta Giovanetti^{2,3}  | Francesca Benedetti¹  |
 Francesco Unali⁴ | Silvia Spoto⁵  | Michele Guarino⁶ | Silvia Angeletti⁷  |
 Massimo Ciccozzi⁸ 

¹Department of Biochemistry and Molecular Biology, Institute of Human Virology and Global Virus Network Center, University of Maryland School of Medicine, Baltimore, Maryland, USA

²Laboratório de Flavivírus, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

³Laboratório de Genética Celular e Molecular, ICB, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

⁴Area Comunicazione e Brand Management, University Campus Bio-Medico of Rome, Rome, Italy

⁵Department of Diagnostic and Therapeutic Medicine, University Campus Bio-Medico of Rome, Rome, Italy

⁶Department of Gastrointestinal Diseases, Campus Bio-Medico University, Rome, Italy

⁷Unit of Clinical Laboratory Science, University Campus Bio-Medico of Rome, Rome, Italy

⁸Medical Statistic and Molecular Epidemiology Unit, University of Biomedical Campus, Rome, Italy

Correspondence

Massimo Ciccozzi, Medical Statistic and Molecular Epidemiology Unit, University of Biomedical Campus, Rome, Italy.
 Email: M.ciccozzi@unicampus.it

Abstract

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) originated in Wuhan, China in early December 2019 has rapidly widespread worldwide. Over the course of the pandemic, due to the advance of whole-genome sequencing technologies, an unprecedented number of genomes have been generated, providing both invaluable insights into the ongoing evolution and epidemiology of the virus and allowing the identification of hundreds of circulating genetic variants during the pandemic. In recent months variants of SARS-CoV-2 that have an increased number of mutations on the Spike protein have brought concern all over the world. These have been called “variants of concerns” (VOCs), and/or “variants of interests” (VOIs) as it has been suggested that their genome mutations might impact transmission, immune control, and virulence. Tracking the spread of emerging SARS-CoV-2 variants is crucial to inform public health efforts and control the ongoing pandemic. In this review, a concise characterization of the SARS-CoV-2 mutational patterns of the main VOCs and VOIs circulating and cocirculating worldwide has been presented to determine the magnitude of the SARS-CoV-2 threat to better understand the virus genetic diversity and its potential impact on vaccination strategy.

KEYWORDS

intervention strategies, pandemic, SARS-CoV-2, vaccine, VOCs, VOIs

1 | INTRODUCTION

“Whether it is nobler to the soul to endure the outrages, stones, and darts of unfair fortune, or take up arms against a sea of trouble and fight to disperse them” quoting a historic sentence from Hamlet in a famous drama by William Shakespeare (Act III, Scene I) written in 1600–1602, where Hamlet's many doubts led him to claim, “the obstacles (in that sleep of death ... from the development of this mortal life...) must induce us to reflect.”

We are living today an important phase of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic where we identify different mutations and variants of the virus circulating all over the world.

Zoonoses are infectious diseases transmitted from animals to humans that can evolve to become efficiently transmissible human-to-human infections. The current pandemic caused by the SARS-CoV-2 virus is believed to have originated from a wildlife food market in China's Wuhan city towards the end of 2019.^{1,2} Current evidence points to its

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Journal of Medical Virology* Published by Wiley Periodicals LLC

origin from a bat-borne virus and the global pandemic represents the first time that the virus has been transmitted into humans.^{3,4}

Generally, the rates of nucleotide substitution of RNA viruses are fast, and this rapid evolution is mainly shaped by natural selection. This high error rate and the consequent rapidly evolving virus populations,⁵ which could lead to the accumulation of amino acid mutations, might affect the transmissibility of the virus, its cell tropism, and pathogenicity. It would unfortunately also present daunting challenges for the design of effective vaccines and diagnostic assays. Fortunately, however, until now the observed diversity among SARS-CoV-2 sequences has been low.

Coronaviruses such as SARS-CoV-2 are relatively stable thanks to a proofreading mechanism that operates during replication. Many genomic studies have nevertheless revealed changes in their genomes, including mutations and deletions. As the terms mutation, variant, and strain are often used interchangeably in describing the epidemiology of SARS-CoV-2, the distinctions appear to be crucial. The genetic material of SARS-CoV-2 is RNA. To replicate, and therefore establish the infection, the virus must hijack the host cell and use the cell's machinery to duplicate itself. Errors often occur during this entire process which is the RNA replication. This results in viruses that are similar but not exact copies of the original one. Those errors are called mutations, and viruses with these mutations are called variants. Variants could differ by a single or many mutations. A variant is referred to as a strain when it starts to present distinct physical properties. Such differences could involve a variant binding to a different cell receptor, or replicating more quickly, or transmitting more efficiently, and enhancement in its virulence. Essentially, all strains are variants, but not all variants are strains.

The first SARS-CoV-2 variant with a D614G substitution in the spike protein emerged early in the pandemic, between January and February 2020.^{6–10} Over a period of several months, the D614G mutation

replaced the initial SARS-CoV-2 strain identified in Wuhan, China, and by June 2020 became the dominant form of the virus circulating globally. Studies highlighted the role of this variant which appears to confer a fitness advantage to the virus possibly associated with the improvement of replication and/or transmission in humans.^{8,11,12}

Currently, four variants (B.1.1.7 also known as 20I/501Y.V1 or VOC 202012/01 or Alpha variant, B.1.351 also known as 20H/501Y.V2 or VOC 202012/02 or Beta variant, P.1 also known as 20J/501Y.V3 or Gamma variant and the B.1.617.2 also known as Delta variant) carrying several mutations in the receptor-binding domain (RBD) of the spike (S) protein, raise concerns about their potential to shift the dynamics and public health impact of the pandemic.^{13–16} They appear potentially associated with (i) increased transmissibility, (ii) propensity for re-infection, (iii) escape from neutralizing antibodies, and (iv) increased affinity for the human angiotensin-converting enzyme 2 (ACE2) receptor.^{16,17} (Figure 1 and Table 1).

The Alpha variant contains 23 nucleotide substitutions and it is not phylogenetically related to the SARS-CoV-2 virus circulating in the UK at the time the variant was detected. Recently epidemiological modeling, phylogenetic and clinical findings suggest that the Alpha variant has increased transmissibility due to a mutation called N501Y, which allows SARS-CoV-2 to bind more readily to the human receptor ACE2, the entry point for SARS-CoV-2 to a wide range of human cells. However, preliminary analyses also indicate that there is no change in disease severity or occurrence of reinfection. Another mutation in this variant of concern (VOC), the deletion at position 69/70del, was found to affect the performance of some diagnostic polymerase chain reaction (PCR) assays with an S gene target. However, most PCR assays in use worldwide use multiple targets, and therefore the impact of the variant on diagnostics is not anticipated to be significant. Laboratory evaluation has also demonstrated no significant impact on the performance of antigen-based lateral flow devices (Figure 1 and Table 1).

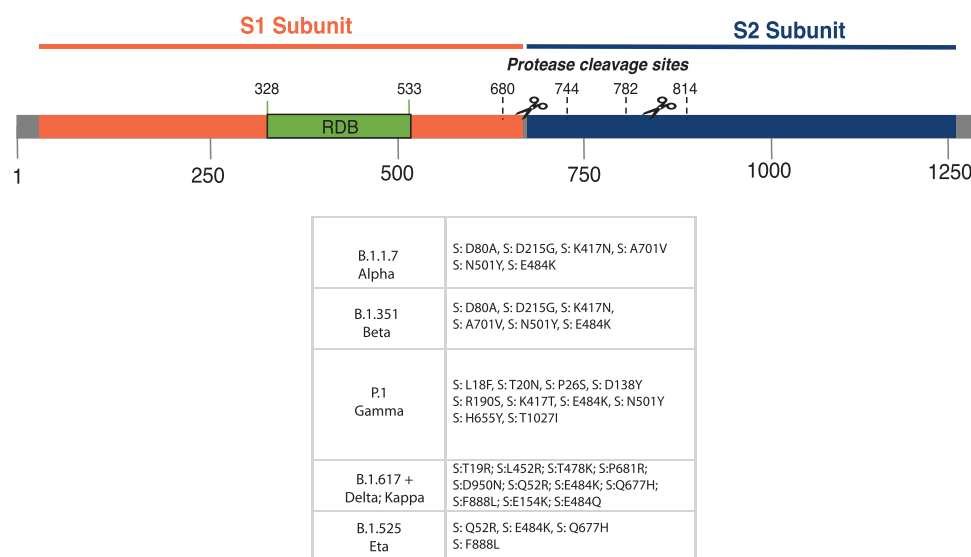


FIGURE 1 Schematic representation of S1 and S2 subunits of SARS-CoV-2 with the main mutations indicated for each VOC and VOI. The “+” symbol indicates the B.1.617.1 and B.1.617.2 variants. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; VOC, variants of concerns; VOI, variants of interests

TABLE 1 List of the main variants of concerns (VOC) and variants of interests (VOI) with the new WHO nomenclature, the old nomenclature based on lineages, the defining SNPs, the earliest documented samples dates, the location of the first detection, and the classification status

Variants	WHO nomenclature	Defining SNPs	Earliest sample date	First detected	Status
B.1.1.7	Alpha	aa:orf1ab:T1001I; aa:orf1ab:A1708D; aa:orf1ab:I2230T del:11288:9; del:21765:6; del:21991:3 aa:S:N501Y; aa:S:A570D; aa:S:P681H; aa:S:T716I; aa:S:S982A; aa:S:D1118H; aa:Orf8:Q27*; aa:Orf8:R52I; aa:Orf8:Y73C aa:N:D3L; aa:N:S235F	Sep 2020	UK	VOC
B.1.351	Beta	aa:E:P71L; aa:N:T205I; aa:orf1a:K1655N; aa:S:D80A aa:S:D215G; aa:S:K417N; aa:S:A701V; aa:S:N501Y; aa:S:E484K	May 2020	South Africa	VOC
P.1	Gamma	aa:orf1ab:S1188L aa:orf1ab:K1795Q; del:11288:9 aa:S:L18F; aa:S:T20N; aa:S:P26S; aa:S:D138Y; aa:S:R190S aa:S:K417T; aa:S:E484K; aa:S:N501Y; aa:S:H655Y; aa:S:T1027I aa:orf3a:G174C; aa:orf8:E92K; aa:N:P80R	May 2020	Brazil	VOC
B.1.617.2	Delta	aa:S:T19R; aa:S:L452R; aa:S:T478K; aa:S:P681R; aa:S:D950N aa:ORF3a:S26L; aa:M:I82T; aa:ORF7a:V82A; aa:ORF7a:T120I aa:N:D63G; aa:N:R203M; aa:N:D377Y	Oct 2020	India	VOC
B.1.525	Eta	aa:orf1ab:L4715F; aa:S:Q52R; aa:S:E484K; aa:S:Q677H aa:S:F888L; aa:E:L21F; aa:E:I82T; del:11288:9; del:21765:6 del:28278:3	Dec 2020	Nigeria	VOI
B.1.617.1	Kappa	aa:orf1ab: T1567I; aa:orf1ab: T3646A; aa:orf1ab: M5753I aa:orf1ab: K6711R; aa:S: E154K; aa:S: L452R; aa:S: E484Q aa:S: P681R; aa:orf3a: S26L; aa:orf7a: V82A; aa:N: R203M	Oct 2020	India	VOI

Abbreviation: SNPs, single nucleotide polymorphisms.

The Beta variant, also called the 501Y.V2 variant (because of the presence of the N501Y mutation), was first notified on December 18, 2020 by national authorities in South Africa. This variant is defined by eight mutations in the spike protein, including three substitutions (K417N, E484K, and N501Y) at residues in its RBD. In a few weeks, this variant largely replaced other SARS-CoV-2 viruses circulating in the Eastern Cape, Western Cape, and KwaZulu-Natal provinces. Recent studies suggest this variant is associated with a higher viral load, which may suggest the potential for increased transmissibility, and or virulence, allowing differential clinical outcomes. Further investigations are needed to understand the impact on transmission, clinical severity of infection, laboratory diagnostics, therapeutics, vaccines, or public health preventive measures (Figure 1 and Table 1).

The Gamma variant, first identified in January 2021 in travelers from the Amazonas state (North of Brazil) who arrived in Japan,¹⁵ harbors a constellation of 17 unique mutations, including three in the RBD of the spike protein (K417T, E484K, and N501Y). It thus

immediately raised concerns to public health authorities over the risk of its unknown potential of faster spreading and/or worsening of coronavirus disease 2019 (COVID-19) clinical outcomes (Figure 1 and Table 1).

The Kappa variant, first identified in India in October 2020, has raised global concern, after being reported in many countries worldwide. This variant harbors a constellation of 12 unique mutations, including three in the RBD of the spike protein (T19R, L452R, T478K, P681R, and D950N) (Figure 1 and Table 1).

Since the identification of these VOCs, the combination between the unprecedented number of cases and more than 1.4 million genomes has allowed the identification of many other circulating genetic variants during the pandemic, containing lineage-specific mutations including single point mutations and/or genetic deletions,^{5,18-21} most of them described as VOIs, including the B.1.525 (Eta variant) and the B.1.617.1 (Kappa variant) (Figure 1 and Table 1).

The Kappa variant, first identified in India in October 2020, has raised global concern, after being reported in many countries

worldwide.²² This variant contains two key mutations to the outer spike portion of the virus, referred to as E484Q and L452R, which appear to have never been reported together. The identification of this double mutation in key areas of the virus's spike protein may allow the virus to escape the immune system (Figure 1 and Table 1).

The Eta variant has been described as a VOI for the first time in mid-December 2020 in Nigeria and later then in many other countries. This variant presents a notable group of missense mutations believed to be of particular importance due to their potential for increased transmissibility, virulence, and reduced effectiveness of vaccines.¹⁶ Also known as 20A/S:484K, the Eta variant harbors some genetic signatures related to diverse putative effects on viral fitness, which are shared by other VOCs, highlighting the potential impact of dissemination in countries where it has not yet been detected: (i) Q677H—described as modulating transmissibility; (ii) Δ 144—associated with immune escape; (iii) Δ 106-108—detected already in Alpha, Beta, and Gamma variants; (iv) E484K—present in Beta, Gamma, and Zeta variants; (v) N439K—found in Y453F, B.1.141, and B.1.258 variants; (vi) two deletions Δ H69/ Δ V70 detected in Alpha variant²³ (Figure 1 and Table 1).

S1 spike is the major SARS-CoV-2 protein allowing virus entry into human cells expressing ACE2 receptor through a defined RBD.²⁴ It has been shown that antibodies generated against the spike protein can block the entry of SARS-CoV-2 into host cells, thus preventing infection.²⁵ For this reason, many efforts to generate safe and efficacious COVID-19 vaccines have been focused on targeting the spike protein. Changes in its sequence and structure are thus particularly relevant because they may hamper the ability of antibodies to bind and block the entry of the virus into the host cells.

The S1 mutation is in a position (AA 614; SD2 region) close to the S1 junction with the S2 component of the Spike protein. The mutation appears to diminish protein stability and change the conformation of the spike glycoprotein, thus enhancing receptor-binding capacity, reducing the shedding, and increasing infectivity.¹²

Among other mutated proteins potentially important we also list Nsp1, which is also known as the leader protein and is central in the inhibition of the antiviral innate immune response. In particular, in cells infected by SARS-CoV-2, it has been shown to play a central role in reducing the expression of IFN- α , and thus it is considered one of the most important determinants of viral pathogenicity.²⁶ By analyzing genomes alignments from all over the world, our group identified a deletion of nine nucleotides in positions 686–694 (AA position 241–243) of the Nsp1 protein.⁵ Even if this deletion was present in only a few genomes from different geographic areas, structural prediction modeling suggested an alteration in the C-terminal tail structure. It is possible that viruses harboring this deletion are likely to be less pathogenic than commonly observed viral strains, though more studies are needed to confirm this theory. To further support this concept, we note that two common endemic human coronaviruses, HCoV-OC43¹⁶ and HCoV-229E,²⁷ have extensive deletions in the C-terminal region of Nsp1.

To provide further information about how quickly the virus could potentially increase its genetic variability, we analyzed the

Open Reading Frame 1ab gene of SARS-CoV-2 to look for the presence of mutations that could have been caused by selective pressure on the virus and that could influence its ability to infect the host.²⁸ We found some potential sites under positive selective pressure and some other stabilizing mutations in the nsp2 protein that probably could explain why this virus is more contagious than SARS-CoV.

Recently our group investigated SARS-CoV-2 transmission dynamics in Italy, one of the countries hit first and hardest by the pandemic, by coupling phylodynamic analysis of viral genetic and epidemiological data. Our study revealed that multiple SARS-CoV-2 lineages, largely linked to outside introductions, cocirculated during early epidemic spread, initially characterized by large transmission clusters concomitant with a high number of infections. Subsequent implementation of three-phase nationwide lockdown measures greatly reduced infection and hospitalization cases during the summer, yet we demonstrate the persistence of viral spread among few, much smaller clusters acting as “hidden reservoirs.” Mathematical modeling shows that increased mobility among residents eventually catalyzed the coalescence of such clusters, thus increasing the number of infections and igniting a new epidemic wave. Our results suggest that timely identification and characterization of the epidemic reservoirs' number, size, and distribution may help to increase the effectiveness of containment measures and provide guidance for future vaccine deployment strategies.²⁹ The probable failure of molecular surveillance systems to monitor the spread of different virus lineages in many countries has given rise to a large amount and spread of different genetic variants of SARS-CoV-2. Fortunately, there has been relatively limited evidence of virus mutations having a significant functional effect on the virus. Of importance is a data sharing and online platform to track in real-time the lineages spread, giving new information to researchers all over the world.³⁰ It is of importance to maintain the mutations landscape of SARS-CoV-2 under control globally to better understand the effect on the infectivity and antigenicity of the variants. Fortunately, most mutations are of little importance from the evolutionary point of view, and no consequence we have seen, but sometimes, the virus can acquire a mutation that gives it an advantage over other strains. In this case, the virus evolution can lead to a consequence, such as greater infectivity and scarce efficacy in antibody response. The Spike mutations can potentially facilitate better affinity or binding and enable easier entry into the host cell, as in the case of the D614G mutation. The RBD in the spike protein is the most variable part of the coronavirus genome.² Mutations can putatively also render the virus resistant to neutralization by host antibodies.

A number of vaccines have been approved by several regulatory agencies worldwide, and at the time of writing the ones with the most scientific information available and most broadly used are BNT162b2 (Pfizer BioNTech),³¹ mRNA-1273 (Moderna),³² Ad26.COV2.S (Janssen),³³ AZD1222 (AstraZeneca),³⁴ and Sputnik-V (Gam-COVID-Vac).³⁵ The efficacy seems to be pretty high for all the vaccines, with 70%–95% protection against mild to severe COVID-19 symptoms and almost total protection against death.

However, several strains of SARS-CoV-2 carrying mutations in the spike protein were recently identified, such as the Alpha,^{36,37} the Gamma,^{14,15} and the Beta variant.¹⁴ Mutations, forming variants that maintain virulence and viral fitness need to be identified and monitored to inform the future of COVID-19 vaccines and therapeutics.

There are a number of preliminary studies aimed at evaluating the efficacy of each vaccine against these variants. Preliminary data seem to indicate that some vaccines are still effective, though they may need to be updated periodically.^{38–41} The hope that three vaccines recently approved by the FDA for emergency use could determine the end of the SARS-CoV-2 pandemic has been supported by the evidence that these vaccines showed an efficacy superior to 85%. This hope has been dampened by the identification of viral variants with mutated spike protein, which in all vaccines is the viral antigen used for active immunization, thus worrying the public opinion with the suspicion that these variants could lower the vaccine's efficacy. So far evidence came from *in vitro* observations showing a 10-fold decrease of neutralization antibody in presence of the viral variants, which in turn raised doubts and questions about vaccine efficacy against mutated SARS-CoV-2. However, more data are needed to validate these observations, and studies are also ongoing by challenging vaccinated individuals with active viruses and monitoring the infection development.⁴²

Vaccine ineffectiveness is so worrying because it represents the real chance to fight the SARS-CoV-2 pandemic. The emergence of viral variants threatening this unique opportunity is the consequence of scientific evidence for mutations inducing enhanced virulence, re-infection, and resistance to monoclonal or polyclonal antibodies therapy. In the case of antibody resistance, the T cell response should be evaluated to understand if the whole immune response could be affected or maintained. Contextually, vaccine production companies are acting by remodeling the first vaccine in the light of the new variants' appearance, even though this could be a continuous evolution phenomenon driven by the selective pressure exerted by the immune response and target viral therapy to which the virus is ever more subjected.⁴³

The real meaning of SARS-CoV-2 variants and their impact on the vaccination campaign is the most important factor to consider. The viral variants affecting the spike protein represent the most important threat to vaccine efficacy, because of the potential decrease in antibody efficacy. This possibility seems improbable until population vaccine coverage will not be sufficient. Moreover, it is important to evaluate the size of neutralization activity change in case of infection sustained by viral strains carrying spike protein mutations. The neutralizing antibody titers routinely determined should be considered. In addition, the occurrence of a viral mutation increasing a determined characteristic of the virus could on the other side affect and decrease another property of the viral strains. On this basis, what could be useful now for the virus should not be in the next future making it less fit and more susceptible to the host immune response. For this reason, it is very important to consider SARS-CoV-2 evolution dynamics and host immune response by long-term analysis.⁴⁴

2 | CONCLUSIONS

The COVID-19 pandemic has stressed our health care systems in unprecedented ways and underlined once more the important role of the studies regarding viral molecular evolution to identify single points mutations and recombination events to assess the following possibilities: that two different SARS-CoV-2 strains may coinfect the same cell; and/or a SARS-CoV-2 strain might have acquired new traits like virulence and drug susceptibility directly from other strains^{21,24}; and/or the adaptability of SARS-CoV-2 to human immune system might be significantly strengthened through genetic recombination. For these reasons, the accuracy of diagnosis based on serologic and molecular biology assays might be compromised by this genetic variability,²⁵ and also the transmission tracking based on the phylogenetic tree could be misleading since the topology of mutation route is a network rather than a tree.^{45–47}

Additionally, it is necessary to understand the phenotypic impact of the SARS-CoV-2 mutations generating variants. Through infectivity assays and neutralization assays, it is possible to understand the effects of emerging mutations on ACE2 binding and NAb binding.

It is unlikely that the early viral mutations observed in Europe and then in the United States once the virus emerged from China radically influenced its fitness. Moreover, their contribution to lethality is difficult to determine, since at the beginning of a pandemic event, the virus is likely to be very aggressive and its sequences would tend to be more homogenous. On the other hand, the early mutations we observed in the viral polymerase gene could have affected its processivity and fidelity, further increasing its mutation rate and the generation of viral clades progressively more heterogeneous. The emergence of subsequent specific patterns of mutations, concomitant with the decline in case fatality rate, likely follows the principle of homoplasmy and suggests a converged evolution due to the accumulation of mutations over time. This would, in turn, lead to the rapidly progressive and convergent adaptation of the virus to the human host. Additional confirmation and the biological significance of such mutations need to be determined. Nonetheless, it is tempting to speculate that they may contribute to the loss of virulence of SARS-CoV-2.

What we have known all along is the ability of the virus to mutate, without warning and without changing its characteristics in the interaction with humans. Mutations are significant only in a few cases and become interesting only when they are part of its strategy of adaptation and survival.

In parallel to the scientific discussion over time in the society of real-time information, faced with a "minute-by-minute" story of the main global event (a planetary emergency in a hyper-connected world), the traditional approach of the media has proved to be quite ineffective in explaining this phenomenon. Unfortunately, very often, the journalistic-chronicle angle has prevailed over the scientific-dissemination point of view.

The discovery of variants was indeed considered as a sudden event and presented with astonishment when it was an

unpredictable (but expected) event that only deserved to be accurately explained.

A first observation: when did we start talking about variants? We started when concerns emerged about the so-called “English variant” (Alpha variant). Until that moment topics of mutations and variants had always remained in the background. Mutations were just a possibility that the virus was not yet expressing one important way, not least because for several months many insignificant variants had been observed. For months since the end of 2019, the virus remained almost unchanged, still in the form that has made it highly contagious and not very lethal: most of the mutations that have occurred (by March 2020 there were more than 160) did not give the virus an evolutionary advantage over humans and were therefore essentially irrelevant.

One year after the outbreak, each country is still struggling to control the pandemic by implementing several different measures. Identifying variants of concern have now become crucial, even though meaningful epidemiological data are still lacking in many cases. A few of the most advanced countries have developed effective epidemiological surveillance systems, while others are still in the process of expanding their current ones. For this reason, data coming from these few countries with an adequate surveillance system tends to influence all the others.

It thus is clear that only when epidemiological surveillance centers will be developed at the national level and then interconnected globally, it will be possible to control the virus in the appropriate ways. This will also allow us to better manage the spreading of variants of concern, and to adopt targeted measures to contain the contagion.

Adaptive mutations in the SARS-CoV-2 genome could alter its pathogenic potential, and at the same time would increase the difficulty of drug and vaccine development. This contribution will not deal in detail with the mass of molecular information now available for SARS-CoV-2. It will rapidly summarize the information on its evolutionary and structural features that could be useful for the development of vaccines.

Finally, we must remember that identifying these mutations is relevant for the design of antiviral drugs and vaccines updates. For this reason, timely viral detection and sequence analysis, more precise and reliable tracking methods, prompt implementation of measures of social distancing, are fundamental to quickly recognize and contain new emerging clusters of infection.

ACKNOWLEDGMENT

Open access funding enabled and organized by Projekt DEAL.

CONFLICTS OF INTEREST

All the authors declare that there are no conflicts of interest.

AUTHOR CONTRIBUTIONS

Conceptualization: Davide Zella and Massimo Ciccozzi. *Writing—original draft preparation:* Davide Zella, Marta Giovanetti, Francesca

Benedetti, and Massimo Ciccozzi. *Writing—review and editing:* Davide Zella, Marta Giovanetti, Francesca Benedetti, Francesco Unali, Silvia Spoto, Michele Guarino, Silvia Angeletti, and Massimo Ciccozzi. *Supervision:* Massimo Ciccozzi.

ORCID

Davide Zella  <http://orcid.org/0000-0001-5576-5770>

Marta Giovanetti  <http://orcid.org/0000-0002-5849-7326>

Francesca Benedetti  <http://orcid.org/0000-0001-8720-090X>

Silvia Spoto  <http://orcid.org/0000-0002-8799-4997>

Silvia Angeletti  <http://orcid.org/0000-0002-7393-8732>

Massimo Ciccozzi  <http://orcid.org/0000-0003-3866-9239>

REFERENCES

1. Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020;579(7798):265-269. Erratum in: *Nature*. 2020 Apr;580(7803):E7. <https://doi.org/10.1038/s41586-020-2008-3>
2. Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020; 579(7798):270-273. <https://doi.org/10.1038/s41586-020-2012-7>
3. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. *Nat Med*. 2020;26(4):450-452. <https://doi.org/10.1038/s41591-020-0820-9>
4. Benvenuto D, Giovanetti M, Ciccozzi A, Spoto S, Angeletti S, Ciccozzi M. The 2019-new coronavirus epidemic: evidence for virus evolution. *J Med Virol*. 2020;92(4):455-459. <https://doi.org/10.1002/jmv.25688>
5. Benedetti F, Snyder GA, Giovanetti M, et al. Emerging of a SARS-CoV-2 viral strain with a deletion in nsp1. *J Transl Med* 2020;18(1): 329.
6. Benvenuto D, Giovanetti M, Salemi M, et al. The global spread of 2019-nCoV: a molecular evolutionary analysis. *Pathog Glob Health* 2020;114(2):64-67.
7. Pachetti M, Marini B, Benedetti F, et al. Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. *J Transl Med*. 2020;18(1):179.
8. Korber B, et al. Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2 [published online ahead of print April 29, 2020]. *bioRxiv*. 2020.
9. Korber B, Fischer WM, Gnanakaran S, et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell*. 2020;182(4):812-827.
10. Benvenuto D, Demir AB, Giovanetti M, et al. Evidence for mutations in SARS-CoV-2 Italian isolates potentially affecting virus transmission. *J Med Virol*. 2020;92(10):2232-2237.
11. Pleskoff O, Sol N, Labrosse B, Alizon M. Human immunodeficiency virus strains differ in their ability to infect CD4+ cells expressing the rat homolog of CXCR-4 (fusin). *J Virol* 1997;71(4):3259-3262.
12. Zhang L, Jackson CB, Mou H, et al. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nat Commun*. 2020;11(1):6013.
13. Rambaut A, Loman N, Pybus O, et al. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. <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>. Accessed December 21, 2020.
14. Tegally H, Wilkinson E, Giovanetti M, et al. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature*. 2021;592(7854): 438-443. <https://doi.org/10.1038/s41586-021-03402-9>

15. Faria NR, Mellan TA, Whittaker C, et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science*. 2021;372:815-821. <https://doi.org/10.1126/science.abh2644>
16. Emerging SARS-CoV-2 Variants. Centers for Disease Control and Prevention, Centers for Disease Control and Prevention. www.cdc.gov/coronavirus/2019-ncov/more/science-and-research/scientific-brief-emerging-variants.html
17. (a) Vasques Nonaka CK, Franco MM, Gräf T, et al. Genomic evidence of SARS-CoV-2 reinfection involving E484K spike mutation. *Brazil Emerg Infect Dis*. 2021;27(5):1522-1524.; (b) Weisblum Y, Schmidt F, Zhang F, et al. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *Elife*. 2020;19:e61312.
18. Holland LA, Kaelin EA, Maqsood R, et al. An 81 nucleotide deletion in SARS-CoV-2 ORF7a identified from sentinel surveillance in Arizona (January to March 2020). *J Virol*. 2020;94:e00711-20.
19. Bal A, Destras G, Gaymard A, et al. Molecular characterization of SARS-CoV-2 in the first COVID-19 cluster in France reveals an amino acid deletion in nsp2 (Asp268del). *Clin Microbiol Infect* 2020; 26(7):960-962.
20. Islam MR, Hoque MN, Rahman MS, et al. Genome-wide analysis of SARS-CoV-2 virus strains circulating worldwide implicates heterogeneity. *Sci Rep* 2020;10(1):14004.
21. Phan T. Genetic diversity and evolution of SARS-CoV-2. *Infect Genet Evol*. 2020;81:104260.
22. Gupta RK. Will SARS-CoV-2 variants of concern affect the promise of vaccines? *Nat Rev Immunol*. 2021;21:1-2. <https://doi.org/10.1038/s41577-021-00556-5>
23. SARS-CoV-2 variants of concern and variants under investigation in England. Technical briefing 7, Public Health England. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/972247/Variants_of_Concern_VOC_Technical_Briefing_7_England.pdf
24. Vinson V. How SARS-CoV-2 binds to human cells. *Science*. 2020; 367(6485):1438.
25. Huo J, Le Bas A, Ruza RR, et al. Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nat Struct Mol Biol* 2020;27(9):846-854.
26. Jauregui AR, Savalia D, Lowry VK, Farrell CM, Wathelet MG. Identification of residues of SARS-CoV nsp1 that differentially affect inhibition of gene expression and antiviral signaling. *PLoS One*. 2013; 8(4):e62416.
27. Farsani SM, Dijkman R, Jebbink MF, et al. The first complete genome sequences of clinical isolates of human coronavirus 229E. *Virus Genes*. 2012;45(3):433-439.
28. Angeletti S, Benvenuto D, Bianchi M, Giovanetti M, Pascarella S, Ciccozzi M. COVID-2019: the role of the nsp2 and nsp3 in its pathogenesis. *J Med Virol* 2020;92(6):584-588.
29. Giovanetti M, Cella E, Benedetti F, et al. SARS-CoV-2 shifting transmission dynamics and hidden reservoirs limited the efficacy of public health interventions in Italy. *Commun Biol*. 2020;4(489):1-9.
30. Hadfield J, Megill C, Bell SM, et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics*. 2018;34(23):4121-4123. <https://doi.org/10.1093/bioinformatics/bty407>
31. Polack FP, Thomas SJ, Kitchin N, et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N Engl J Med* 2020;383(27): 2603-2615.
32. Baden LR, El Sahly HM, Essink B, et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N Engl J Med*. 2021;384(5): 403-416.
33. Sadoff J, Le Gars M, Shukarev G, et al. Interim results of a phase 1-2a trial of Ad26.COV2.S Covid-19 vaccine. *N Engl J Med* 2021; 384:1824-1835.
34. Voysey M, Clemens S, Madhi SA, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet*. 2021;397(10269):99-111.
35. Logunov DY, Dolzhikova IV, Shcheblyakov DV, et al. Safety and efficacy of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine: an interim analysis of a randomised controlled phase 3 trial in Russia. *Lancet*. 2021;397(10275):671-681.
36. Davies NG, Abbott S, Barnard RC, et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science*. 2021; 372(6538)
37. Volz E, Mishra S, Chand M, et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature*. 2021;593:266-269.
38. Wang Z, Schmidt F, Weisblum Y, et al. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature*. 2021;592: 616-622.
39. Garcia-Beltran WF, Lam EC, St. Denis K, et al. Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity. *Cell*. 2021;184:2372-2383.
40. Kuzmina A, Khalaila Y, Voloshin O, et al. SARS-CoV-2 spike variants exhibit differential infectivity and neutralization resistance to convalescent or post-vaccination sera. *Cell Host Microbe* 2021;29: 522-528.
41. Madhi SA, Baillie V, Cutland CL, et al. Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 Variant. *N Engl J Med* 2021;384:1885-1898.
42. Luchsinger LL, Hillyer CD. Vaccine efficacy probable against COVID-19 variants. *Science*. 2021;371(6534):1116. <https://doi.org/10.1126/science.abg9461>
43. Gómez CE, Perdiguero B, Esteban M. Emerging SARS-CoV-2 variants and impact in global vaccination programs against SARS-CoV-2/COVID-19. *Vaccines*. 2021;9(3):243. <https://doi.org/10.3390/vaccines9030243>
44. Lauring AS, Hodcroft EB. Genetic variants of SARS-CoV-2-what do they mean? *JAMA*. 2021;325(6):529-531. <https://doi.org/10.1001/jama.2020.27124>
45. Gu Z, Gao Q, Faust EA, Wainberg MA. Possible involvement of cell fusion and viral recombination in generation of human immunodeficiency virus variants that display dual resistance to AZT and 3TC. *J Gen Virol* 1995;76(10):2601-2605.
46. Moutouh L, Corbeil J, Richman DD. Recombination leads to the rapid emergence of HIV-1 dually resistant mutants under selective drug pressure. *Proceedings of the National Academy of Sciences of the United States of America*. 1996;93(12):6106-6111.
47. Blackard JT, Cohen DE, Mayer KH. Human immunodeficiency virus superinfection and recombination: current state of knowledge and potential clinical consequences. *Clin Infect Dis*. 2002;34(8): 1108-1114. <https://doi.org/10.1086/339547>

How to cite this article: Zella D, Giovanetti M, Benedetti F, et al. The variants question: what is the problem? *J Med Virol*. 2021;1-7. <https://doi.org/10.1002/jmv.27196>