



CURRENT TOPIC / AKTUELNA TEMA

Emerging variants of novel coronavirus – myth and reality

Tanja Jovanović, Marko Janković, Aleksandra Knežević

University of Belgrade, Faculty of Medicine, Institute of Microbiology and Immunology, Belgrade, Serbia

SUMMARY

The new coronavirus has crossed the species barrier leading to the pandemic of COVID-19. The lengthy circulation of the virus within the human population has enabled the development of many new viral variants, some of which are conducive to further pathogen spread. Notable variants are those that contain mutations within the S gene, particularly within the region that codes for the receptor-binding domain (RBD) that links to the hACE-2 receptor. These mutations are responsible for increased viral transmission and influence disease severity, reliability of clinical tests, as well as vaccine efficacy. At present, the variant first identified in the United Kingdom poses the greatest threat in Europe.

Keywords: coronavirus; COVID-19; variant; SARS-CoV-2; emerging virus

INTRODUCTION

For all the queries that remain about the novel coronavirus (CoV) and the disease it causes, scientists have generated a vast amount of knowledge in a very short period of time. Some of them are answered, but some of them are still a mystery. Today, a year after the virus was discovered, the question still stands: what do we really know about SARS-CoV-2?

GENERAL CHARACTERISTICS OF SARS-COV-2

Coronaviridae are enveloped, positive-sense single-stranded RNA viruses in the order of *Nidovirales*. The subfamily *Orthocoronavirinae* is further classified into four CoV genera: *Alfa*-, *Beta*-, *Gamma*- and *Deltacoronavirus* (Figure 1). They infect a variety of animals including birds and mammals. A broad range of coronaviruses are found in bats. They were identified as human pathogens since the 1960s and have been associated with 15–30% of annual respiratory tract infections [1]. Widespread human coronaviruses HCoV-OC43, HCoV-229E and HCoV-HKU1 cause common colds but also lower respiratory tract infections in the youngest and oldest age groups. The HCoV-NL63 is considered to be an important cause of pseudocroup and bronchiolitis in children.

In the last 20 years, two zoonotic coronaviruses have emerged: SARS-CoV and MERS-CoV, discovered in 2002 and 2012, respectively. Both have caused human outbreaks, the Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS). In late 2019, a third novel coronavirus, initially related to a cluster of pneumonia cases

in Wuhan (China), was identified and named SARS-CoV-2 [2].

The origin of SARS-CoV-2 is still a matter of debate. Bioinformatic studies revealed that it has a 96.2% identity with the coronavirus genome isolated from the feces of *Rhinolophus affinis* bats [3]. There are three possible mechanisms by which the bat virus has crossed the species barrier: by natural selection in an animal host before jumping into humans or natural selection in humans after the virus transferred into a human host. The third mechanism, which is the least likely one, is artificial manipulation with the bat virus in a laboratory [4].

The novel virus is closely related to SARS-CoV. However, SARS-CoV-2 genome sequence is more distant from SARS-CoV, with 79% of similarity throughout the whole genome, and more than 90% of the sequence identity for essential enzymes and structural proteins [5, 6]. The main difference lies in the receptor binding domain within the S gene [5, 6]. But characteristics of SARS-CoV-2 in the surface proteins lead to a higher viral load kinetics, enable virus to enhanced rate of transmission during the human infection.

GENETIC VARIABILITY OF SARS-COV-2

The length of the SARS-CoV-2 genome is about 30,000 nucleotides – extremely long for an RNA virus. The replication error correction system with proofreading mechanism limits the mutation rate of the virus.

The first two thirds of the genome correspond to a single *ORF1ab* gene, coding for the polyprotein which is cleaved into 16 non-structural viral proteins. The last third of the

Received • Примљено:
February 6, 2021

Accepted • Прихваћено:
March 29, 2021

Online first: April 6, 2021

Correspondence to:

Tanja JOVANOVIĆ
1 Dr Subotića Starijeg Street
Belgrade 11000, Serbia
tanja.jovanovic@med.bg.ac.rs

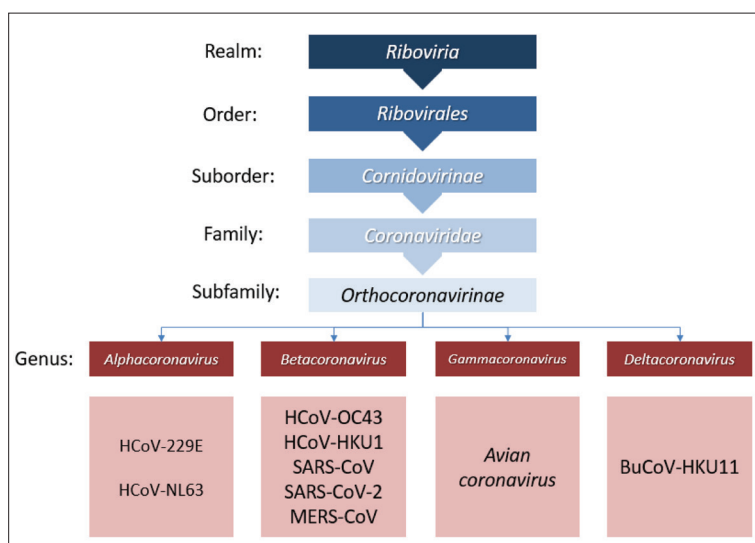


Figure 1. The novel SARS-CoV-2 position within the viral taxonomy

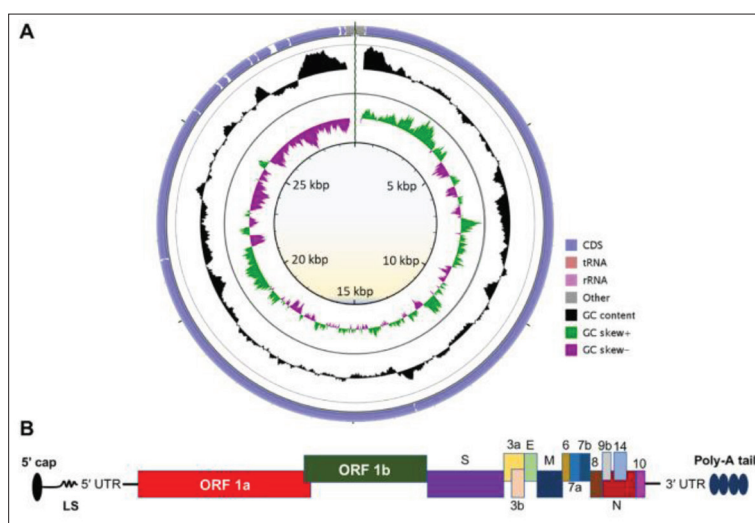


Figure 2. A: The organization of the SARS-CoV-2 reference genome; B: representation of the pathogen's single-stranded positive-sense RNA; source: Naqvi et al. [25]

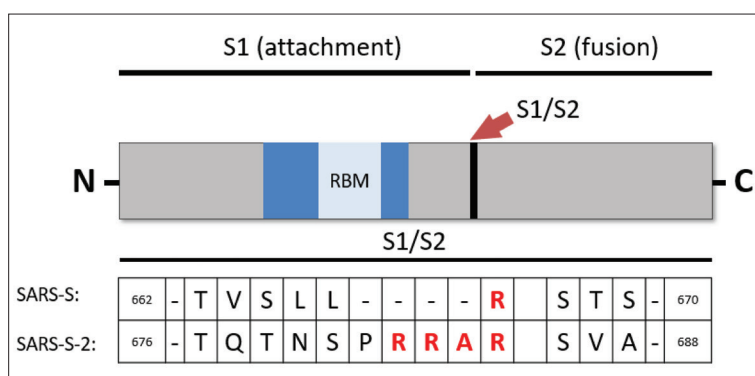


Figure 3. The insertion sequence of SARS-CoV-2 at the cleavage site

genome contains nine genes coding for structural proteins, with the most important being surface (spike, S), membrane (M), envelope (E), and nucleocapsid (N) proteins. Structural proteins aid in the assembly and release of new copies of the virus. The M and E proteins are involved in the formation of the viral envelope, while the N protein

forms a helical ribonucleocapsid complex with the RNA and interacts with viral membrane proteins during pathogen assembly. Accessory genes (3a, 3b, 6, 7a, 7b, 8, 9b, 9c, and 10) present short bits of genome [7, 8]. They are clustered within the structural genes and help the virus to evade the immune system (Figure 2).

The life cycle of SARS-CoV-2 is a very dynamic process. After specific adsorption and membrane fusion, viral genome RNA is released into the cytoplasm, while uncoated RNA translates two polyproteins, pp1a and pp1b, which encode nonstructural proteins and form a replication-transcription complex (RTC) in the vesicle. The RTC continuously replicates and synthesizes a set of subgenomic mRNAs that encodes accessory and structural proteins. After the components of RNA and protein assemble, new viral particles are produced and then released into the extracellular space via exocytosis.

The polymerase, in addition to its canonical RNA-dependent RNA polymerase activity, is able to jump between the different RNA strands. This is a property that plays a key role in the recombination capacity of CoVs and promotes their evolution and host change in the case of dual or mix infection.

THE MAIN ANTIGEN OF SARS-COV-2 – SPIKE PROTEIN

The S protein of the virus is a key factor involved in infection. Like other coronaviruses, the S protein of SARS-CoV-2 mediates receptor recognition, cell adsorption and fusion of viral envelope with plasma membrane – all events responsible for viral entry into the cell [9, 10, 11]. This means that S protein is a major player in the pathogenesis of viral infection, but is also involved in evolution of CoV and crossing of the species barrier [12]. The insertion of short sequences within the S gene is unique for SARS-CoV-2. These insertions add four amino-acids at the precise cleavage site of the protein, immediately upstream of the arginine, which creates a sequence RRAR, corresponding to the consensus recognition motif of the furin protease (Figure 3) [13].

The S protein is a trimer, composed of three identical units. It forms a characteristic crown-like halo surrounding the viral particle. The spikes are coated with polysaccharide molecules to camouflage them, evading surveillance of the host immune system during entry [14].

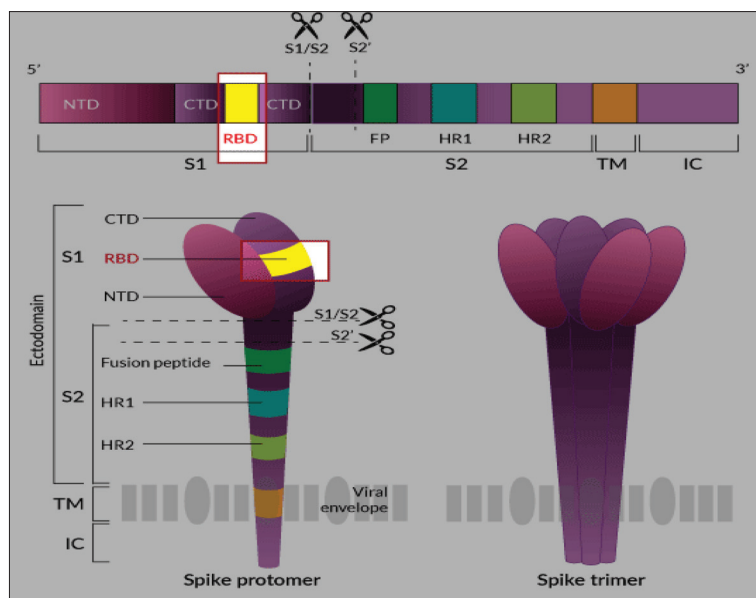


Figure 4. The S protein and its structure

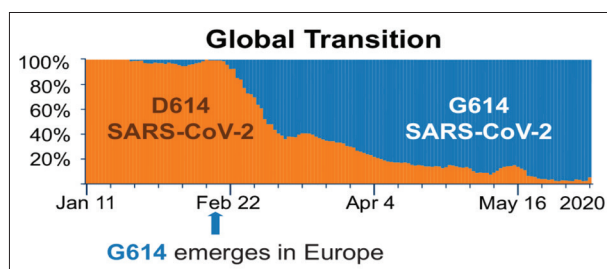


Figure 5. The D614G mutation and its frequency in regard to time

In its native state, S protein exists as an inactive precursor. During viral infection, target cell proteases activate the S protein and cleaved into two subunits (Figure 4). The first cleavage called “priming” generates the S1 and S2 subunits [15]. The second cleavage occurs within the S2 unit and releases the end of the fusion peptide located at the beginning of the S2 subunit [16]. These two proteolytic cleavages are catalyzed by furin and serine transmembrane proteases 2 – TMPRSS2 [15, 16, 17].

The S1 domain contains the receptor-binding domain, or RBD (319–541 residues), which is mainly responsible for binding of the virus to the receptor. The S2 domain mainly contains the heptapeptide repeat sequence, or the HR domain with HR1 (912–984 residues) and HR2 (1163–1213 residues) units, which is closely related to the virus fusion protein (FP – 788–806 residues) [12, 18]. Once the virus interacts with the host cell, extensive structural rearrangement of the S protein occurs, allowing the virus envelope to fuse with the host cell membrane and start replication in the cell.

NEW EMERGING VARIANTS OF SARS-COV-2

Viruses constantly change through mutation. Scientists monitor changes in the virus by sequencing, trying to find the new variants of concern. New variants can change

virulence and infectivity of viruses in the sense of how widely the new variants spread and how the disease differs from the disease caused by other variants. The most important question is how the new variants affect therapy, molecular diagnostic testing for COVID-19, and effectiveness of vaccination.

The first variant of SARS-CoV-2 occurred at the beginning of the pandemic, in late January or early February 2020. It was a D614G substitution (substitution of aspartic acid to glycine at position 614) in the gene encoding the spike protein, but outside of the RBD region [12, 19]. After several months, this specific D614G mutant replaced the initial SARS-CoV-2 strain identified in China and by June 2020 became the dominant form of the virus circulating globally (Figure 5) [19]. This strain has increased infectivity and transmission, but without the effects on clinical illness or effectiveness of commercial laboratory diagnostics tests.

The next new variant of SARS-CoV-2 identified in North Jutland, Denmark, in August and September 2020 has been linked to infection among farmed mink. This specific variant has a combination of mutations not previously observed. The impact of this mutant was that the infected mink transferred the infection to 12 humans, but it did not spread widely [20].

In the United Kingdom, a novel important variant has been identified as SARS-CoV-2 VOC 202012/01 (Variant of Concern, year 2020, month 12, variant 01) or 501Y.V1. In November 2020, a rapid increase in COVID-19 cases overall was associated with the emergence of this new variant in south east, east regions of the UK and London. Retrospectively, the first instance of VOC 202012/01 was identified in a case from September 20, 2020 in the UK [21]. This variant contains 23 nucleotide substitutions and is not phylogenetically related to the SARS-CoV-2 virus circulating in the United Kingdom at the time the variant was detected [21].

The three mutations that have the largest potential biological effect of the UK variants are N501Y, spike deletion 69/70 del and P687H [21, 22]. The N501Y mutation leads to an amino acid change from asparagine to tyrosine at position 501. It is one of six key contact residues within the RBD and has been identified with having increasing binding affinity to human ACE-2 [23].

The 69/70 del mutation is a deletion of six bases in the RNA, leading to the removal of two amino acids at position 69 and 70 of the spike protein. This spike deletion has been described in the context of evasion of the human immune response but has also occurred a number of times in association with other RBD changes [22, 23].

The P681H mutation led to a change from proline to histidine at position 681 [23]. This mutation is immediately adjacent to the furin cleavage site, a known location of biological significance.

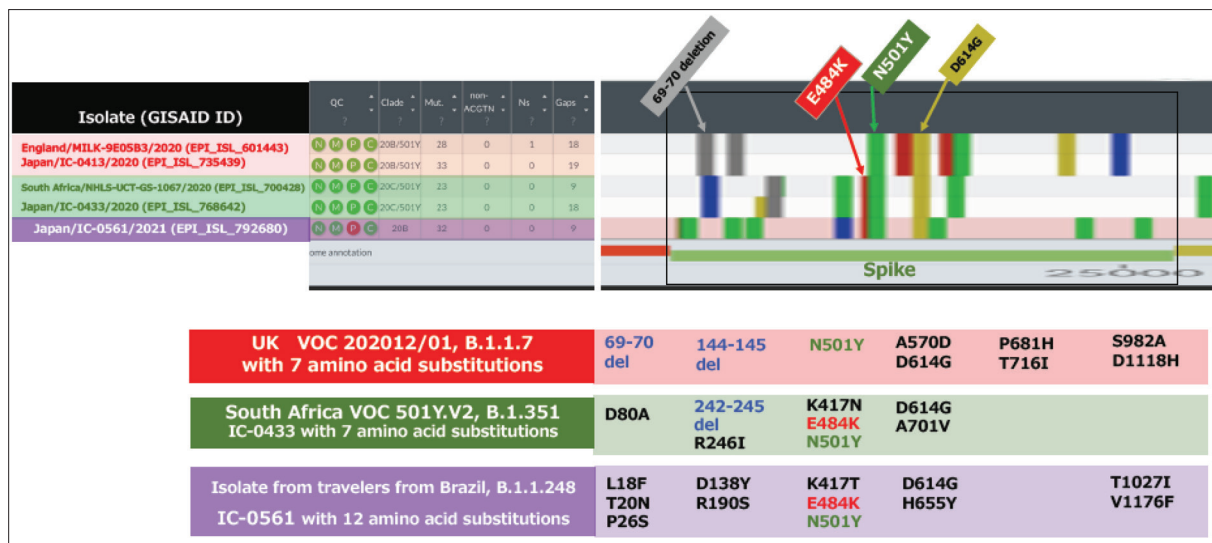


Figure 6. A vis-à-vis presentation of mutations in the UK, South African, and Brazilian variants; source: National Institute of Infectious Diseases, Japan [26]

Molecular virology, epidemiology, and phylogenetic analyses suggest that SARS-CoV-2 VOC 202012/01 has increased transmissibility. There is no evidence that this variant leads to changes in disease severity measured by length of hospitalization and 28-day case fatality, but the analyses have not been completed as yet [23]. This deletion was found to affect the performance of PCR assays with the S gene target. The deletion has no significant impact on the performance of antigen-based tests. Many countries all over the world have been reporting the presence of the British variant [22, 23].

A new variant rapidly spreading in South Africa named 501Y.V2 or the South Africa Variant was identified in December 2020. In addition to N501Y mutation, the new variant has another two mutations, K417N and E484K [23]. The combination of these three mutations results in the highest degree of conformation alteration of S RBD bound to hACE-2 compared to either E484E or N501Z alone. The 501Y.V2 has largely replaced other SARS-CoV-2 viruses circulating in the Eastern Cape, Western Cape, and KwaZulu-Natal provinces.

Retrospective analysis of whole-genome sequences from South Africa indicates that this variant emerged in early August 2020 [24]. Preliminary studies suggest that this variant is associated with a higher viral load, which may suggest a potential for increased transmissibility but no clear evidence of more severe outcomes of the disease [23, 24]. There is some evidence that one of the spike protein mutations, E484K, may affect neutralization by some polyclonal and monoclonal antibodies [23, 24].

Finally, the National Institute of Infectious Diseases in Japan, in January 2021, found a new variant with 17 unique amino acid changes and three deletions in four travelers from Brazil [23, 24]. This variant contains 17 unique amino acid changes with 10 in its S protein, including the three

mutations in the spike protein receptor-binding domain: K417T, E484K, and N501Y. There is some evidence that the mutations in this variant may affect its transmissibility and antigenic profile, which may influence the ability of antibodies generated through a previous natural infection or through vaccination to recognize and neutralize the virus. A representation of mutations in current variants of concern in parallel to one another is presented in Figure 6.

CONCLUSION

The SARS-CoV-2 genome has been fully sequenced. The function of the viral genes and their role in the virus–host interaction is elucidated. Scientists around the world follow the accumulation of mutations occurring in the pathogen's genome, with particular focus on the S gene coding for a protein important for the initial step in host infection. The change in the RBD region of the mentioned gene enabling infection of ACE-2-positive cells is especially significant. The longer the virus circulates within the general population, as well as immunocompromised persons, the number of mutations will increase and novel pathogen variants will arise. Mutations responsible for the faster transmission of SARS-CoV-2 have been identified in the variants of concern, namely the British, South African, and Brazilian variants. This highlights the importance of further sequencing efforts from patient samples, which aids in keeping track of viral characteristics linked to disease course and outcome, as well as in the application of epidemiological measures centered on the prevention of spread of variants of concern.

Conflict of interest: None declared.

REFERENCES

1. Audi A, Allbrahim M, Kaddoura M, Hijazi G, Yassine HM, Zaraket H. Seasonality of Respiratory Viral Infections: Will COVID-19 Follow Suit? *Front Public Health*. 2020;8:567184.
2. Fernandes JD, Hinrichs AS, Clawson H, Gonzalez JN, Lee BT, Nassar LR, et al. The UCSC SARS-CoV-2 Genome Browser. *Nat Genet*. 2020;52(10):991–8.
3. Kaur N, Singh R, Dar Z, Bijarnia RK, Dhingra N, Kaur T. Genetic comparison among various coronavirus strains for the identification of potential vaccine targets of SARS-CoV2. *Infect Genet Evol*. 2020;89:104490.
4. Perer JC, Montagnier L. HIV man-manipulated coronavirus evolution trends. *Int J Res*. 2020;8(07):217–63.
5. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*. 2020;395(10224):565–74.
6. Tai W, He L, Zhang X, Pu J, Voronin D, Jiang S, et al. Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cell Mol Immunol*. 2020;17(6):613–20.
7. Michel CJ, Mayer C, Poch O, Thompson JD. Characterization of accessory genes in coronavirus genomes. *Virology*. 2020;17(1):131.
8. Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, et al. Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host Microbe*. 2020;27(3):325–8.
9. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*. 2020;581(7807):215–20.
10. Kawase M, Shirato K, van der Hoek L, Taguchi F, Matsuyama S. Simultaneous treatment of human bronchial epithelial cells with serine and cysteine protease inhibitors prevents severe acute respiratory syndrome coronavirus entry. *J Virol*. 2012;86(12):6537–45.
11. Shang J, Wan Y, Luo C, Ye G, Geng Q, Auerbach A, et al. Cell entry mechanisms of SARS-CoV-2. *Proc Natl Acad Sci USA*. 2020;117(21):11727–34.
12. Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol*. 2020;5(4):562–9.
13. Dhama K, Patel SK, Sharun K, Pathak M, Tiwari R, Yatoo MI, et al. SARS-CoV-2 jumping the species barrier: Zoonotic lessons from SARS, MERS and recent advances to combat this pandemic virus. *Travel Med Infect Dis*. 2020;37:101830.
14. Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, et al. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell*. 2020;181(4):894–904.e9.
15. Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science*. 2020;369(6501):330–3.
16. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*. 2020;181(2):281–92.e6.
17. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*. 2020;581(7807):215–20.
18. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*. 2020;181(2):271–80.e8.
19. Hou YJ, Chiba S, Halfmann P, Ehre C, Kuroda M, Dinno KH 3rd, et al. SARS-CoV-2 D614G variant exhibits efficient replication *ex vivo* and transmission *in vivo*. *Science*. 2020;370(6523):1464–8.
20. Koopmans M. SARS-CoV-2 and the human-animal interface: outbreaks on mink farms. *Lancet Infect Dis*. 2021;21(1):18–9.
21. World Health Organization [Internet]. SARS-CoV-2 Variant – United Kingdom of Great Britain and Northern Ireland [cited 2021 Feb 1]. Available from: <https://www.who.int/csr/don/21-december-2020-sars-cov2-variant-united-kingdom/en/>
22. World Health Organization [Internet]. SARS-CoV-2 Variants [cited 2021 Feb 1]. Available from: <https://www.who.int/csr/don/31-december-2020-sars-cov2-variants/en/>
23. World Health Organization [Internet]. COVID-19 Weekly Epidemiological Update [cited 2021 Feb 1]. Available from: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20210105-weekly-epi-update_21.pdf?
24. European Centre for Disease Prevention and Control [Internet]. Risk related to spread of new SARS-CoV-2 variants of concern in the EU/EEA – 29 December 2020 [cited 2021 Feb 1]. Available from: <https://www.ecdc.europa.eu/en/publications-data/covid-19-risk-assessment-spread-new-sars-cov-2-variants-eueea>
25. Naqvi AAT, Fatima K, Mohammad T, Fatima U, Singh IK, Singh A, et al. Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. *Biochim Biophys Acta Mol Basis Dis*. 2020;1866(10):165878.
26. National Institute of Infectious Diseases [Internet]. Brief report: New Variant Strain of SARS-CoV-2 Identified in Travelers from Brazil [cited 2021 Apr 2]. Available from: <https://www.niid.go.jp/niid/en/2019-ncov-e/10108-covid19-33-en.html>.

Нове варијанте новог коронавируса – мит и реалност

Тања Јовановић, Марко Јанковић, Александра Кнежевић

Универзитет у Београду, Медицински Факултет, Институт за микробиологију и имунологију, Београд, Србија

САЖЕТАК

Нови коронавирус прескочио је баријеру врсте и довео до пандемије инфекције COVID-19. Дуготрајна циркулација вируса у популацији омогућила је настанак бројних нових варијанти вируса, од којих су неке претеће за даље ширење епидемије. Од посебног су значаја оне варијанте код којих је дошло до мутација у оквиру гена S, посебно у региону за

везивање вируса за рецептор *hACE-2* (RBD). Мутације у овом гену одговорне су за повећану трансмисију вируса, тежину клиничке слике, поузданост дијагностичких тестова као и ефикасност вакцинације. За Европу највећу претњу тренутно представља британски сој вируса SARS-CoV-2.

Кључне речи: коронавирус, COVID-19; варијанта; SARS-CoV-2; нови вирус