SARS-COV-2: INSIGHT INTO THE EMERGING GENETIC VARIANTS

Neli Korsun[1], Ivelina Trifonova[1], Veselin Dobrinov[1], Iliyana Grigorova[1], Svetla Angelova[1], Ivaylo Alexiev[2]

[1]National Laboratory "Influenza and ARD", [2]NCL "HIV", Department of Virology, National Center of Infectious and Parasitic Diseases, Sofia

ABSTRACT:
SARS-CoV-2 is a highly contagious virus, which appeared in China in late 2019, spread rapidly and caused the largest pandemic in the last 100 years. Despite the intensive research, there is no specific antiviral drug currently. Effective vaccines have been developed in a short time and they are already widely used. As a RNA virus, SARS-CoV-2 mutates constantly, and several thousand genetic variants have emerged in the course of the pandemic, some of which are associated with increased infectivity, reinfection risk, reduced activity of therapeutic antibodies and reduced effectiveness of vaccines. This review highlights the features in SARS-CoV-2 structure and replication cycle that would help to understand the significance of individual mutations contained in the emerging genetic variants and to predict the impact of mutations on viral transmissibility, disease severity, diagnostics, therapeutics or immune escape. The main characteristics of the variants of concern are presented.
Keywords: *SARS-CoV-2, mutation, genetic variant*
Category of the manuscript: mini review

INTRODUCTION
At the end of December 2019, the World health Organization (WHO) reported cases of severe pneumonia caused by an unknown agent in Wuhan, Hubei Province, China (1). The causative agent was identified as a new, hitherto unknown coronavirus, which received the name SARS-CoV-2 from the International Committee on Taxonomy of Viruses, and the disease it caused was named COVID-19. In a short time, SARS-CoV-2 spread globally, causing an unprecedented health and economic crisis. On January 30th, 2020, the WHO declared the SARS-CoV-2 epidemic a public health emergency, and a global pandemic was declared on March 11th. Since December 31 of 2019 and as of week 22/2021, 174 032 728 confirmed cases of COVID-19 have been reported worldwide, including 3 738 030 deaths (https://who.int). The long-term circulation of SARS-CoV-2, affecting millions of individuals, has been accompanied by appearance of several thousand mutations in the viral genome, most of which do not affect the biological properties and behavior of virus. Mutations that are associated with increased virus infectivity; increased risk of severe disease, hospitalization, and death; increased frequency of reinfection; failure of some diagnostic tests or reduced efficacy of therapeutic agents and vaccines are of concern. Detailed understanding of the structure, key genomic elements, and replication cycle of the virus is needed to assess the significance of the particular mutations identified in circulating viruses and to predict changes in the behavior of viruses as a result of these mutations.

Classification and origin of SARS-CoV-2
SARS-CoV-2 belongs to the subfamily *Coronavirinae*, family *Coronaviridae* of the order Nidovirales. This subfamily includes four genera: α, β, γ and δ. Among them, α and β coronaviruses (CoVs) only infect mammals, while γ and δ mainly infect birds (2). To this day, seven types of CoVs able to infect humans are identified: four endemic, low pathogenic CoVs (229E, NL63, OC43, and HKU1) causing 15-30% of common colds, and three highly pathogenic CoVs: Middle East Respiratory

ADDRESS FOR CORRESPONDENCE:
Neli Korsun,
44A Stoletov Blvd, 1233 Sofia;
e-mail: neli_korsun@abv.bg;
mobile: 0885413620

Syndrome Coronavirus (MERS-CoV), Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV), and Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) capable to cause severe respiratory illnesses such as pneumonia and acute respiratory distress syndrome (ARDS), leading to death (3). The CoVs 229E and NL63 belong to genus α coronavirus and the remaining are members of genus β coronavirus. SARS-CoV and SARS-CoV-2 belong to the subgenus: Sarbecovirus. The three highly pathogenic CoVs originate from bat CoVs. Palm civets and racoon dogs have been recognized as intermediate hosts for SARS-CoV and dromedary camels for MERS-CoV. Pangolin was found to be a potential intermediate host for SARS-CoV-2. During the 2002-2003 epidemic, SARS-CoV caused 8 096 disease cases in 26 countries including 774 deaths (mortality rate of 9.6%) then disappeared from circulation. MERS-CoV emerged in the Arabian Peninsula in September 2012, causing 2566 confirmed cases of infection in 27 countries and 882 deaths (mortality rate of 34%). SARS-CoV-2 is responsible for the current pandemic that caused enormous human casualties, serious consequences in many areas of human society and changed the world, as we know it.

Structure of SARS-CoV-2 and key viral factors

SARS-CoV-2 is an enveloped virus with nonsegmented, single-stranded positive-sense RNA genome, whose structure resembles that of other CoVs. Among known RNA viruses, CoVs have one of the largest genomes, ranging between 26-32 kilobases (kb) in size. The coronaviral virions are spherical or moderately pleomorphic particles with a diameter of approximately 60 to 140 nm with crown-shaped 9-12 nm-long spikes on their surface formed by the S glycoprotein. The envelope is a lipid bilayer derived from the host cell membrane in which virus proteins S, M and E are embedded. The membrane (M) protein, the most abundant structural protein, is responsible for shaping the virions. It plays an important role in viral assembly along with

E and N proteins. The envelope (E) protein, the smallest structural protein, is present in small quantities in the virion. The fourth structural protein N binds the viral RNA forming a helical nucleocapsid, located in the core of the viral particle. It is required for packaging viral RNA into the viral particle during viral assembly and acts as an interferon (IFN) inhibitor.

The genome of SARS-CoV-2 is approximately 29.9 kb of size and comprises of a 5'-cap structure and a 3'-poly-A tail. It includes 5' and 3'-untranslating regions (UTR) consisting of 265 and 229 nt, respectively, and 15 open reading frames (ORFs) encoding at least 29 proteins. The first ORFs (ORF1a/b), located at the 5'end of the genome, cover about two-thirds of the entire genome length and encode 16 nonstructural proteins (nsp1-16), involved in viral replication and transcription. Remaining ORFs encode the four major structural proteins: S, M, E, and N and nine accessory proteins (3a, 3b, 6, 7a, 7b, 8b, 9a, 9b, and ORF14) participating in the assembly of viral particles (2,4,5).

The SARS-CoV-2 S protein plays a crucial role in the initial steps of viral infection - it mediates receptor recognition, cell attachment, and entry into host cells. It is the major viral antigen used as a key target for vaccines and therapeutic antibodies, therefore its structure and functions have been intensively studied (6). It exists as a homotrimer, each monomer of which is about 180 kDa and contains 1273 amino acids. The spike protein is comprised of three segments: large ectodomain, transmembrane anchor and a short intracellular tail. It is a heavily glycosylated protein with 22 host derived N-linked glycans, which serve as a shield reducing access of antibodies to specific epitopes. In the native state, the S protein exists as an inactive precursor consisting of two subunits: a distal S1 subunit (14–685 amino acid residues), which is responsible for receptor recognition and binding, and a proximal S2 subunit (a.a. 686–1273), which mediates membrane fusion and entry of virus

into host cell (7). The S protein contains several functional regions (domains): a signal peptide (SP) (a.a. 1–13) located at the N-terminus, an N-terminal domain (NTD) (a.a. 14–305) and a receptor-binding domain (RBD) (a.a. 319–541) in the S1 subunit; a fusion peptide (FP) (a.a. 788–806), heptapeptide repeat sequence 1 (HR1) (a.a. 912–984), HR2 (a.a. 1163–1213), transmembrane domain (TM) (a.a. 1213–1237), and cytoplasm tail (CT) (a.a. 1237–1273) in the S2 subunit (Figure 1) (2,8). The S protein binds to the receptors of sensitive cells via RBD, which is a key target for the most potent neutralizing antibodies (nAbs) of host. Due to its surface location, NTD is the least conserved region and is another potential target for nAbs. FP consists of 15-20 conserved amino acids. HR1 and HR2 are composed of a repetitive heptapeptide HPPHCPC. They form a six-helical bundle (6-HB), which plays an assential role in viral fusion and entry [7,9]. Unlike SARS-CoV and MERS-CoV, SARS-CoV-2 has a unique polybasic "RRAR" (Arg-Arg-Ala-Arg) cleavage site at the junction of S1 and S2 subunits (a.a. 682–685), which enables effective cleavage by furin and furin-like proteases (10). The presence of this furin recognition region enhances cellular tropism and transmissibility of SARS-CoV-2 due to the broad cellular expression of furin proteases (6). Such furin cleavage site is also present in highly pathogenic avian influenza viruses and is associated with their pathogenicity.
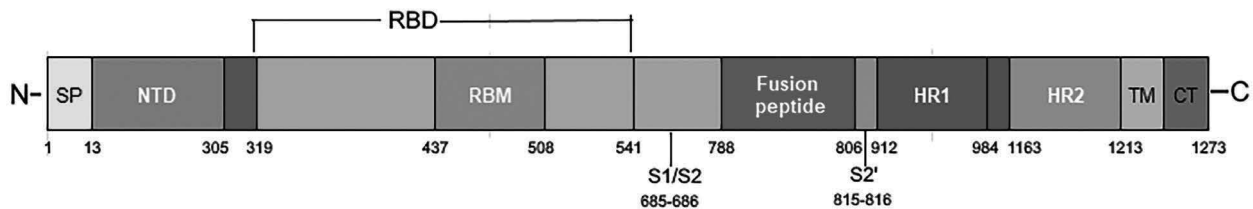
Figure 1. Schematic of SARS-CoV-2 S protein

SARS-CoV-2 replication

The first step in SARS-CoV-2 infection is binding of the S protein to the cell surface receptor, mediated by the RBD. SARS-CoV and SARS-CoV-2 use the angiotensin-converting enzyme 2 (ACE-2) as a cellular receptor, which is distributed in the lung, intestine, heart, and kidney. MERS-CoV recognizes the dipeptidyl peptidase 4 receptor. The RBDs of S1 subunits undergo hinge-like movements between two states ("up" or "down"), in which the residues that bind ACE2 receptor are transiently exposed or hidden. Based on the hinge-like movements of RBDs, S protein monomers are displayed in "open" or "closed" orientations: "open", receptor-accessible conformation (RBD up) for receptor binding and "closed", receptor-inaccessible conformation, (RBD down) for immune evasion (11). Within the RBD, there is a receptor binding motif (RBM) (a.a. 437-508), which makes direct contacts with the peptidase domain of ACE2. 17 amino acid residues in RBM are in contact with 20 amino acids in ACE2, six of these 17 residues - L455, F486, Q493, S494, N501 and Y505, are crucial for efficient binding to the ACE-2 receptor (4). RBD of SARS-CoV-2 recognizes and binds to the ACE2 with more than 10-fold higher affinity than the RBD of SARS-CoV (6). This may explain the higher contagiousness and transmissibility of SARS-CoV-2 as compared to SARS-CoV. Binding of RBD with the ACE-2 receptor is followed by a proteolytic cleavage of the S protein in two consecutive steps: in the S1/S2 junction (a.a. 685-686) ("priming" cleavage) and in the so-called S2' site located immediately after the FP (a.a. 815-816) ("activation" cleavage) (5,6). The first cleavage results in the separation of RBD from FP; the other cleavage leads to the exposure of FP. SARS-CoV-2 S protein is cleaved

by human cell proteases such as furin, trypsin, cathepsins and transmembrane protease serine 2 (TMPRSS2) (7). After the cleavages, the S1 subunit is dissociated and the S2 subunit undergoes dramatic conformational changes, in which the FP is exposed and triggers the fusion of the viral membrane with the host cell membrane. SARS-CoV-2 enters susceptible cells through endocytosis, in which fusion of the viral and endosome membranes leads to the release of the viral nucleocapsid into the cell cytoplasm. After entering the cellular cytosol, 5'-proximal ORF1a and ORF1b of the genomic RNA are translated to produce two large polyproteins (pp1a and pp1ab), where pp1ab is produced via a ribosomal frameshift mechanism. These polyproteins are processed by virally encoded proteases into 16 non-structural proteins, which form a replication-transcription complex in double-membrane vesicles (4,5). During the replication, full-length (-) RNA copies of the genome are produced, which are used as templates for full-length (+) RNA genomes. A characteristic feature of the CoV family is the synthesis of multiple negative-sense subgenomic RNA intermediates, which serve as templates for the production of subgenomic mRNAs. They are subsequently translated to produce virus-specific structural and accessory proteins (12). The subgenomic mRNAs share the same leader sequence of 70–90 nucleotides at their 5' ends and the same 3' ends. Subgenomic N mRNA is the most abundantly generated transcript in SARS-CoV-2 infected cells, which makes the N gene one of the most suitable targets for detection of SARS-CoV-2. Following translation, structural M, E and S proteins migrate along the secretory pathway into the endoplasmic reticulum-Golgi intermediate compartment (ERGIC). N proteins bind to the daughter's genomic RNAs and form helical nucleocapsids that interact with the other structural proteins to form mature viral particles. Following assembly and budding into the lumen of the ERGIC, virions are realesed from the infected cells through exocytosis.

Genetic variants of SARS-CoV-2

As a RNA virus, SARS-CoV-2 undergoes frequent mutations, in spite of some proof-reading capacity of its RNA polymerase complex (nsp14 protein acts as 3'-5'exoribonuclease) (3). Mutations in the circulating viruses are evaluated compared to the reference strain Wuhan-Hu-1 (GenBank accession MN908947), which comprises a 29,903-bp–long RNA and is the first virus with sequenced complete genome. Due to the important role of S protein in the early phase of coronaviral infection and the fact that it is the major target of most vaccines and therapeutic agents, mutations in this protein are of particular concern. *A group of SARS-CoV-2 that share the same inherited set of distinctive mutations compared to the reference strain but do not show significant changes in their phenotypic characteristics (such as antigenic properties, virulence, features in the epidemiology of infections) is defined as a genetic variant.* Over the course of the pandemic, several thousands of SARS-CoV-2 variants have arisen. The variants are grouped into larger groups called clades. Several nomenclatures of the clades have been proposed. Rambaut et al. identify five major lineages: A, B, B.1, B.1.1, and B.1.177 (https://cov-lineages.org). The Global Initiative on Sharing All Influenza Data (GISAID) distinguishes 7 clades: O, S, L, V, G, GH, and GR (www.gisaid.org). Nextstrain identifies five clades - 19A, 19B, 20A, 20B, and 20C (https://nextstrain.org/ncov). Genetic variants, which are associated with one or more of the following changes: increase in transmissibility or detrimental change in COVID epidemiology; and/or increase in virulence or change in clinical disease presentation; and/or decrease in effectiveness of available diagnostics, vaccines and therapeutics, have been accepted by the WHO as *variants of concern* (VOCs). VOCs include multiple mutations in S protein and at least one mutation in the RBD. The following

variants have been designated as VOCs: Alpha (B.1.1.7, GRY or 20I/S:501Y.V1); Beta (B.1.351, GH/501Y.V2 or 20H/S:501Y.V2); Gamma (P1, GR/501Y.V3 or 20J/S:501Y.V3) and Delta (B.1.617.2, G/452R.V3 or 21A/S:478K) (https://who.int).

B.1.1.7 variant was earliest documented in September 2020 in the UK. After detection of the first cases, it surged at an exponential rate and quickly spread in other countries. This variant harbors 23 mutations leading to 17 amino acid changes among which 8 are located in the S protein: deletion 69-70, deletion 144, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H. The role of the following mutations has been established:

— The N501Y is one of the six key amino acid residues in the RBM determining the tight interaction of SARS-CoV-2 with the ACE2 receptor. This mutation is associated with increased infectivity and is also presented in the B.1.351 and P.1 variants.

— The P681H is located immediately adjacent to the furin cleavage site (a.a. 682–685). It has been speculated that this mutation increases transmissibility by facilitating cleavage of the S precursor.

— The H69–V70 and Y144/145 deletions are located in the NTD, which contains important antigenic sites, targets of many neutralizing antibodies. Mutations in this region increase the ability of the virus to escape immune recognition. The H69–V70 deletion is responsible for a negative result in some RT-PCR assays targeting the S-gene (S-gene drop-out) (13).

— The D614G mutation emerged in eastern China early in the pandemic and then quickly spread around the world, displacing other CoVs that did not have this mutation. The D614G mutation makes the SARS-CoV-2 more infectious, but it does not associate with increased disease severity or escape from host immunity. It increases the ability of the virus to bind to the ACE2 receptor and stabilizes the interaction between the S1 and S2 subunits of the spike, leading to increased transmissibility (14).

The B.1.1.7 variant is ~50% more infectious than other variants in circulation (15). Based on studies in the UK, it is associated with increased risk of hospitalizations and increased case fatality rate (16). There is evidence of minimal impact on the neutralization activity of convalescent and post-vaccination sera (17).

B.1.351 variant was initially detected in in South Africa in September 2020, from where it spread to 94 countries. It contains 23 mutations leading to 17 amino acid changes, including 9 changes in the S protein: L18F, D80A, D215G, 242–244 del, R246I, K417N, E484K, N501Y, D614G, and A701V. Mutations near the tip of the S protein include:

— N501Y, which helps the virus to bind more tightly to the ACE2 receptor.

— K417N, which is located in the RBD on the tip of the spike and helps the virus interact more tightly with human cells.

— E484K, which is located in the RBM and is potentially associated with antigenic change and immune escape. This mutation creates resistance to neutralizing antibodies contained in convalescent plasma and reduces the activity of some neutralizing monoclonal antibodies. It also increases the affinity of the S protein for the ACE2 receptor.

— Combination of E484K, K417N, and N501Y mutations leads to the most significant changes in the structure of the RBD allowing the virus escape antibody neutralization (18).

— L18F, D80A, D215G, 242-244 del, and R246I are located in the NTD, which is a preferential target of antibodies (14).

P1 variant was first detected in Japan in four travelers who arrived in Tokyo from Amazonas, Brazil. This lineage emerged in November 2020 in Manaus, the largest city in Brazil's Amazon region, where it has become the dominant circulating virus. As of June 4, 2021 it has been detected in 53 countries. P.1 variant contains approximately 35 mutations leading to 17 amino

acid changes including 11 changes in S protein: L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, H655Y, T1027I, and V1176F. Some of these mutations: L18F, K417T, E484K, N501Y, and D614G, are located in the same positions of S protein as in the B.1.351 variant.

Both variants, B.1.351 and P1 are associated with increased infectivity, significantly reduced susceptibility to some monoclonal antibodies and reduced neutralization ability of antibodies generated by a previous natural infection or vaccination (18,19). With regard to P1, there is an evidence of an increased risk of reinfections. At the end of 2020, this variant has widespread in Manaus, Brazil, where more than 75% of residents have been infected with SARS-CoV-2 earlier (20).

B.1.617.2 variant first appeared in India in October 2020 and is one of the variants responsible for the high morbidity of COVID-19 in this country. It carries the following mutations in S protein: T19R, 156del, 157del, R158G, L452R, T478K, D614G, P681R, and D950N (21).

- The L452R is also found in the B.1.427/429 variants that are widespread in California. This change in the RBD increases the affinity of the spike protein for the ACE-2 receptor and decreases recognition by antibodies that present in covalescent plasma as well as by some therapeutic monoclonal antibodies.
- The P681R is similar to P681H and is located immediately adjacent to furin-cleavage site.
- T19R, 156del, 157del and R158G are located in the NTD that is extensively mutated and is a target of many neutralizing antibodies.

Variants of interest (VOIs) represent groups of genetically changed viruses in which genetic changes are associated with established or suspected phenotypic implications; and they have been identified to cause multiple SARS-CoV-2 cases or they have been detected in multiple countries (https://who.int). Such VOIs are Eta/Iota (B.1.525/B.1.526) - spread in Nigeria and New York; Epsilon (B.1.427/429) - in California; Kappa (B.1.617.1/B.1.617.3) - in India, and others.

DIAGNOSTIC TESTING

For diagnosing current or recent SARS-CoV-2 infection, viral tests, including nucleic acid amplification tests (NAATs) and antigen tests, are used. NAATs detect one or more viral genes and are characterized by high sensitivity and specificity. Antigen tests have high specificity but are less sensitive than NAATs. To confirm infection with a specific genetic variant, whole genome sequencing (WGS) or sequencing of selected parts of the viral genome should be performed. WGS allows the identification of mutations in various viral genes and the detection of VOCs. Full or partial S-gene sequencing is a cheaper and faster method than WGS. To distinguish the circulating variants, sequenced region must cover at least the entire N-terminal and RBD (amino acid 1-541, 1623 bp), at best the entire S gene. For early detection of known VOCs, diagnostic screening PCR-based assays have been developed. A negative or significantly weaker positive S-gene result with positive results for the other gene targets can be used as an indicator of potential circulation of B.1.1.7 (Alpha) variant. The S-gene target failure is not exclusive to B.1.1.7 - it can be observed in other non-VOC-variants but does not occur for Beta, Gamma and Delta variants. Specific RT-PCR assays identifying VOC specific amino acid substitutions (e.g. spike N501Y, K417N, E484K, L452R) have been developed. Appropriate positive controls should be used.

CONCLUSIONS

Novel genetic variants of SARS-CoV-2 will emerge in the future as a result of the adaptation of the virus to the human population with selection of mutations that improve viral replication and transmissibility or permit the virus to escape from adaptive immune responses. As the proportion of the vaccinated and previously infected individuals worldwide increases, the evolution

of SARS-CoV-2 will be increasingly driven by the immune pressure of the human population. The course of the pandemic will depend on the effectiveness of the vaccines against emerging genetic variants, on the strength and duration of the immunity they create.

ACKNOWLEDGMENTS

REFERENCES

1. Zhou P. *A pneumonia outbreak associated with a new coronavirus of probable bat origin*. Nature, 2020, 579: 270-273.
2. Satarker S, Nampoothiri M. *Structural Proteins in Severe Acute Respiratory Syndrome Coronavirus-2*. Arch Med Res, 2020.
3. Al-Qaaneh AM, Alshammari T, Aldahhan R, Aldossary H, Alkhalifah ZA, Francis Borgio J. *Genome composition and genetic characterization of SARS-CoV-2*. Saudi J Biol Sci. 2021;28(3):1978-1989.
4. Shamsi A, et al. *Potential drug targets of SARS-CoV-2: From genomics to therapeutics*. Int J Biol Macromol. 2021;177:1-9.
5. Kadam SB, Sukhramani GS, Bishnoi P, Pable AA, Barvkar VT. *SARS-CoV-2, the pandemic coronavirus: Molecular and structural insights*. J Basic Microbiol. 2021;61(3):180-202.
6. Pilley TS. *Gene of the month: the 2019-nCoV/SARS-CoV-2 novel coronavirus spike protein*. J Clin Pathol. 2020;73(7):366-369.
7. Duan L, Zheng Q, Zhang H, Niu Y, Lou Y, Wang H. *The SARS-CoV-2 Spike Glycoprotein Biosynthesis, Structure, Function, and Antigenicity: Implications for the Design of Spike-Based Vaccine Immunogens*. Front Immunol, 2020;11:576622.
8. Mittal A, Manjunath K, Ranjan RK, Kaushik S, Kumar S, Verma V. *COVID-19 pandemic: Insights into structure, function, and hACE2 receptor recognition by SARS-CoV-2*. PloS Pathog. 2020:16(8): e1008762.
9. Rehman H, Ahmad I. *COVID-19: a wreak havoc across the globe*. Arch Physiol Biochem. 2020;1-13.
10. Coutard B, Valle C, de Lamballerie X, Canard B, Seidah NG, Decroly E. *The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade*. Antiviral Res. 2020;176:104742.
11. Wang L, Xiang Y. *Spike Glycoprotein-Mediated Entry of SARS Coronaviruses*. Viruses. 2020,12(11):1289.
12. Hatmal MM, Alshaer W, Al-Hatamleh MAI, Hatmal M, Smadi O, Mutasem O. Taha MO et al. *Comprehensive Structural and Molecular Comparison of Spike Proteins of SARS-CoV-2, SARS-CoV and MERS-CoV, and Their Interactions with ACE2*. Cells. 2020; 9(12):2638.
13. Gómez CE, Perdiguero B, Esteban M. *Emerging SARS-CoV-2 Variants and Impact in Global Vaccination Programs against SARS-CoV-2/COVID-19*. Vaccines (Basel). 2021;9(3):243.
14. Guo S, Liu K, Zheng J. *The Genetic Variant of SARS-CoV-2: would It Matter for Controlling the Devastating Pandemic*? Int J Biol Sci. 2021;17(6):1476-1485.
15. Davies NG, Abbott S, Barnard RC, et al. *Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England*. MedRXiv 2021.
16. Horby P, Huntley C, Davies N, Edmunds J, Ferguson N, Medley G, et al. *NERVTAG: Note on B.1.1.7 Severity, 20 January 2021. 2021. Scientific Advisory Group for Emergencies, GOV, UK*.
17. Boehm E, Kronig I, Neher RA, Eckerle I, Vetter P, Kaiser L, et al. *Novel SARS-CoV-2 variants: the pandemics within the pandemic.* Clin Microbiol Infect. 2021;S1198-743X(21)00262-7
18. Jia Z, Gong W. *Will Mutations in the Spike Protein of SARS-CoV-2 Lead to the Failure of COVID-19 Vaccines?* J Korean Med Sci. 2021;36(18):e124
19. Wang Z., Schmidt F., Weisblum Y., Muecksch F., Barnes C.O., Finkin S., Schaefer-Babajew D., Cipolla M., Gaebler C., Lieberman J.A. *mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants*. Nature. 2021;592(7855):616-622.
20. Li R, Liu J, Zhang H. *The challenge of emerging SARS-CoV-2 mutants to vaccine development*. J Genet Genomics. 2021 Apr 20.
21. ECDC. *Emergence of SARS-CoV-2 B.1.617 variants in India and situation in the EU/EEA*. 11 May, 2021.