

ARTICOL DE SINTEZĂ

## Biologia virusului SARS-CoV-2: sinteză narativă

Igor Cemortan<sup>1\*</sup>, Valentina Vorobjit<sup>1</sup>, Svetlana Capcelea<sup>1</sup>, Mariana Ulinici<sup>1</sup>, Eugen Ursu<sup>2</sup>, Dan Croitoru<sup>1</sup>, Tatiana Grigoriev<sup>1</sup>

<sup>1</sup>Catedra de biologie moleculară, Universitatea de Stat de Medicină și Farmacie „Nicolae Testemițanu”, Chișinău, Republica Moldova;

<sup>2</sup>Computational Biology of Aging Group, Institutul de Biochimie al Academiei Române, București, România.

Data primirii manuscrisului: 20.05.2020

Data acceptării spre publicare: 16.06.2020

### Autor corespondent:

Igor Cemortan, dr. șt. med., conf. univ.

Catedra de biologie moleculară

Universitatea de Stat de Medicină și Farmacie „Nicolae Testemițanu”

bd. Ștefan cel Mare și Sfânt, 165, Chișinău, Republica Moldova, MD-2004

e-mail: igor.cemortan@usmf.md

### Ce nu este cunoscut, deocamdată, la subiectul abordat

Există numeroase lacune în înțelegerea particularităților ciclului vital al SARS-CoV-2.

### Ipoteza de cercetare

Publicațiile care abordează biologia SARS-CoV-2 ar fi binevenite pentru înțelegerea și managementul COVID-19.

### Noutatea adusă literaturii științifice din domeniu

Articolul însumează o sinteză a articolelor recent publicate referitoare la biologia SARS-CoV-2.

### Rezumat

**Introducere.** Un nou sindrom respirator acut sever (SARS) determinat de SARS-CoV-2 a apărut recent și se răspândește rapid la om, provocând COVID-19.

**Material și metode.** Au fost analizate 38 de articole relevante actuale la tema cercetată.

**Rezultate.** În articolul dat, a fost analizată poziția taxonomică a virusului, particularitățile de organizare a genomului, ciclul vital al virusului. Genomul SARS-CoV-2 prezintă un grad înalt de similaritate cu alte virusuri care cauzează patologii similare SARS-COVID, MERS-COVID. Un rol aparte este acordat proteinelor virale, în special proteinei Spike, care se unește la receptorul enzimei de conversie a angiotensinei 2, determinând pătrunderea virusului în celula gazdă la oameni. Proteina S poate avea un potențial înalt pentru designul preparatelor antivirale.

REVIEW ARTICLE

## Biology of SARS-CoV-2 virus: narrative review

Igor Cemortan<sup>1\*</sup>, Valentina Vorobjit<sup>1</sup>, Svetlana Capcelea<sup>1</sup>, Mariana Ulinici<sup>1</sup>, Eugen Ursu<sup>2</sup>, Dan Croitoru<sup>1</sup>, Tatiana Grigoriev<sup>1</sup>

<sup>1</sup>Chair of molecular biology, Nicolae Testemitanu State University of Medicine and Pharmacy, Chisinau, Republic of Moldova;

<sup>2</sup>Computational Biology of Aging Group, Institute of Biochemistry of Romanian Academy, Bucharest, Romania.

Manuscript received on: 20.05.2020

Accepted for publication on: 16.06.2020

### Corresponding author:

Igor Cemortan, PhD, assoc. prof.

Chair of molecular biology

Nicolae Testemitanu State University of Medicine and Pharmacy

165, Stefan cel Mare si Sfânt ave., Chisinau, Republic of Moldova, MD-2004

e-mail: igor.cemortan@usmf.md

### What is not known yet, about the topic

There are many gaps in understanding the particularities of the life cycle of SARS-CoV-2.

### Research hypothesis

Publications addressing the biology of SARS-CoV-2 would be helpful for COVID-19 understanding and management.

### Article's added novelty on this scientific topic

The article represents a summary of recently published articles on the biology of SARS-CoV-2.

### Abstract

**Introduction.** A new severe acute respiratory syndrome (SARS) caused by SARS-CoV-2 has emerged recently and is spreading rapidly in humans, causing COVID-19.

**Material and methods.** Thirty-eight current relevant articles on the researched topic were analyzed.

**Results.** In this article, we analyzed the taxonomic position of the virus, the particularities of genome organization, and the life cycle of the virus. The SARS-CoV-2 genome has a high degree of similarity to other viruses that cause similar pathologies: SARS-CoV, MERS-CoV. A special role is given to viral proteins, especially Spike protein, which binds to the angiotensin 2 conversion enzyme receptor, causing the virus to enter the human host cell. The Spike protein can have a high potential for the design of antiviral drugs.

**Concluzii.** S-au înregistrat progrese semnificative în studierea multilaterală a genomului SARS-CoV-2, fapt ce permite identificarea exactă a virusului și determină dezvoltarea medicamentelor și a vaccinurilor pentru tratamentul COVID-19. Cooperarea diferitor instituții, centre științifice, guverne și companii farmaceutice este necesară pentru a preveni răspândirea în continuare a COVID-19 și identificarea soluțiilor de tratament.

**Cuvinte cheie:** COVID-19, SARS-CoV-2, genom, ARN, proteina Spike.

## Introducere

Infecția cu noul coronavirus COVID-19, care este încă în curs de desfășurare, a provocat un număr mare de cazuri de boală și decese, însă înțelegerea noastră despre agentul patogen SARS-CoV-2 rămâne, în mare măsură, neclară [1]. În cazul unei infecții virale este important să se stabilească rapid dacă focarul este cauzat de un virus nou sau unul cunoscut anterior. Aceasta ar ajuta în luarea deciziilor referitoare la abordările și acțiunile necesare pentru detectarea agentului cauzal, controlul transmiterii acestuia și limitarea consecințelor potențiale ale epidemiei. Evaluarea noutății unui virus are, de asemenea, implicare în nomenclatura virusului [2]. Pandemia curentă a contribuit la acumularea cunoștințelor noi despre aspectele genetice ale coronavirusurilor. Secvențierea completă a genomului SARS-CoV-2 a demonstrat statutul său taxonomic în calitate de membru al genului *Betacoronavirus*, cu o divergență clară de la SARS-CoV-1 și MERS-CoV, agenți cauzali ai epidemiilor anterioare [3].

## Material și metode

A fost efectuată o căutare sistematică a literaturii pe internet, atât pentru cercetări originale, cât și pentru articole de recenzie în bazele de date biomedicale, prin intermediul *Google Scholar*, *PubMed* și *Medline*, articole ce sunt publicate în limba engleză. Bibliografiile articolelor au fost utilizate pentru a găsi alte subiecte înrudite. După introducerea cuvintelor cheie „COVID-19 genomic epidemiology”, „SARS-CoV-2 biology”, „SARS-CoV-2 mutations”, au fost obținute 729 de rezultate, dintre care au fost selectate 38 de articole relevante, unde 26 au fost publicate în perioada 2019-2020.

## Rezultate

Încă de la începutul anului 2020 întreaga lume s-a axat pe noua pandemie cu COVID-19, și pe analiza agenților care cauzează patologii similare SARS-COVID, MERS-COVID, pentru a stabili particularitățile la nivel de genom și transcriptom [3]. Explorarea și definirea caracterelor biologice ale virusurilor în contextul întregii diversități naturale a speciilor nu a fost niciodată o prioritate. Apariția SARS-CoV-2 ca agent patogen uman în decembrie 2019 poate fi, astfel, percepută ca fiind complet independentă de focarul SARS-CoV din 2002-2003. Deși SARS-CoV-2 nu este un descendent al SARS-CoV și introducerea fiecăruia dintre aceste virusuri în populația umană a fost, probabil, facilitată de factori externi necunoscuți inde-

**Conclusions.** Important progress was achieved in the multilateral study of the SARS-CoV-2 genome, which allows the exact identification of the virus and leads to the development of drugs and vaccines for the treatment of COVID-19. Cooperation between various institutions, scientific centers, governments, and pharmaceutical companies is needed to prevent the further spread of COVID-19 and to identify treatment solutions.

**Key words:** COVID-19, SARS-CoV-2, genome, RNA, Spike protein.

## Introduction

Infection with the new COVID-19 coronavirus, which is still ongoing, has caused a large number of cases of illness and deaths, but our understanding of the SARS-CoV-2 pathogen remains largely unclear [1]. In the case of a viral infection, it is essential to quickly determine if a new or previously known virus causes the outbreak. This would help to make decisions on the approaches and actions needed to detect the causative agent, control its transmission, and limit the potential consequences of the epidemic. Assessing the novelty of a virus is also important for the identification of the virus nomenclature [2]. The current pandemic has contributed to the accumulation of new knowledge about the genetic aspects of coronaviruses. Full genome sequencing of the SARS-CoV-2 has demonstrated its taxonomic status as a member of the genus *Betacoronavirus*, with a clear divergence from SARS-CoV-1 and MERS-CoV, causative agents of previous epidemics [3].

## Material and methods

We have conducted a systematic search of the literature on the Internet, both for original research and review articles published in English in biomedical databases, through *Google Scholar*, *PubMed*, and *Medline*. Bibliographies of the articles were used to find other related topics. After entering the keywords “COVID-19 genomic epidemiology”, “SARS-CoV-2 biology”, “SARS-CoV-2 mutations”, we obtained 729 results, from which we selected 38 relevant articles, of which 26 were published during 2019-2020.

## Results

Since the beginning of 2020, the whole world has focused on the new COVID-19 pandemic, as well as the analysis of agents that cause similar pathologies SARS-CoV, MERS-CoV to establish the particularities at the level of genome and transcriptome [3]. Exploring and defining the biological characteristics of viruses in the context of the entire natural diversity of species has never been a priority. The emergence of SARS-CoV-2 as a human pathogen in December 2019 can, thus, be perceived as completely independent of the SARS-CoV outbreak in 2002-2003. Although SARS-CoV-2 is not really a descendant of SARS-CoV and the introduction of each of these viruses into the human population was probably facilitated by independent unknown external factors, the two viruses are so

pendenți, cele două virusuri sunt atât de apropiate genetic unul de celălalt, încât istoriile și caracteristicile lor evolutive sunt reciproc informative [2].

### Clasificarea virusului, nomenclatura bolii cauzate

În clasificarea virusurilor bazate pe ARN, trebuie să se ia în considerație variabilitatea lor genetică inerentă, ceea ce duce, adesea, la apariția a două sau a mai multor virusuri cu secvențe de genom neidentice, dar similare, considerate variante ale aceluiași virus [2].

Clasificarea curentă a coronavirusurilor include 39 de specii din 27 de subgenuri, cinci genuri și două subfamilii, care aparțin familiei *Coronaviridae*, subordinul *Cornidovirineae*, ordinul *Nidovirales* și domeniul *Riboviria* [4]. Genurile *Alpha*- și *Beta*-coronavirus afectează oamenii și animalele. Virusurile SARS-CoV (betacoronavirus), 229E (alfacoronavirus), HKU1 (betacoronavirus), NL63 (alfacoronavirus), OC43 (betacoronavirus) și MERS-CoV (betacoronavirus) pot cauza infecții la oameni [5]. Cu toate acestea, betacoronavirusurile sunt cel mai important grup, deoarece acestea cuprind cele mai multe virusuri patogene care pot infecta oamenii, inclusiv, SARS-CoV-2, MERS-CoV și SARS-CoV.

Având în vedere cele menționate mai sus, *Grupul de Studiu Coronaviridae* (GSC) al *Comitetului Internațional de Taxonomie a Virusurilor* (ICTV) a decis să urmeze o linie de raționament bazată pe filogenetică pentru a numi acest virus, a cărui ontogeneză este redată în Figura 1 [2]. Numele virusului este „severe acute respiratory syndrome coronavirus 2” – SARS-CoV-2. Specia căreia îi aparține virusul SARS-CoV-2 este asemănătoare coronavirusurilor ce cauzează sindromul respirator acut sever [6].

Denumirea bolii a fost desemnată de Organizația Mondială a Sănătății (OMS) ca și COVID-19, abrevierea de la „coronavirus disease 2019”. Numele virusului a fost anunțat de Organizația Mondială a Sănătății pe 11 februarie 2020 [7].

Diferite coronavirusuri prezintă o gamă diversă de gazde și

geneticamente apropiate unele de altele astfel încât istoriile și caracteristicile lor evolutive sunt reciproc informative [2].

### Clasificarea virusului, nomenclatura bolii cauzate

În clasificarea de RNA-bazate virusuri, variabilitatea lor genetică inerentă trebuie luată în considerare, ceea ce duce adesea la apariția a două sau a mai multor virusuri cu secvențe de genom neidentice, dar similare, considerate variante ale aceluiași virus [2].

Clasificarea curentă a coronavirusurilor include 39 de specii din 27 de subgenuri, cinci genuri și două subfamilii, care aparțin familiei *Coronaviridae*, subordinul *Cornidovirineae*, ordinul *Nidovirales* și domeniul *Riboviria* [4]. Genurile *Alpha*- și *Beta*-coronavirus afectează oamenii și animalele. SARS-CoV (betacoronavirus), 229E (alfacoronavirus), HKU1 (betacoronavirus), NL63 (alfacoronavirus), OC43 (betacoronavirus) și MERS-CoV (betacoronavirus) pot cauza infecții la oameni [5]. Cu toate acestea, betacoronavirusurile sunt cel mai important grup, deoarece acestea cuprind cele mai multe virusuri patogene care pot infecta oamenii, inclusiv, SARS-CoV-2, MERS-CoV și SARS-CoV.

În vederea celor de mai sus, *Grupul de Studiu Coronaviridae* (CSG) al *Comitetului Internațional de Taxonomie a Virusurilor* (ICTV) a decis să urmeze o linie de raționament bazată pe filogenetică pentru a numi acest virus, a cărui ontogeneză este redată în Figura 1 [2]. Numele virusului este „severe acute respiratory syndrome coronavirus 2” – SARS-CoV-2. Specia la care aparține virusul SARS-CoV-2 este asemănătoare coronavirusurilor care cauzează sindromul respirator acut sever [6].

Numele bolii a fost desemnată de Organizația Mondială a Sănătății (OMS) ca și COVID-19, abrevierea de la „coronavirus disease 2019”. Numele virusului a fost anunțat de Organizația Mondială a Sănătății pe 11 februarie 2020 [7].

Diferite coronavirusuri prezintă o gamă diversă de gazde și

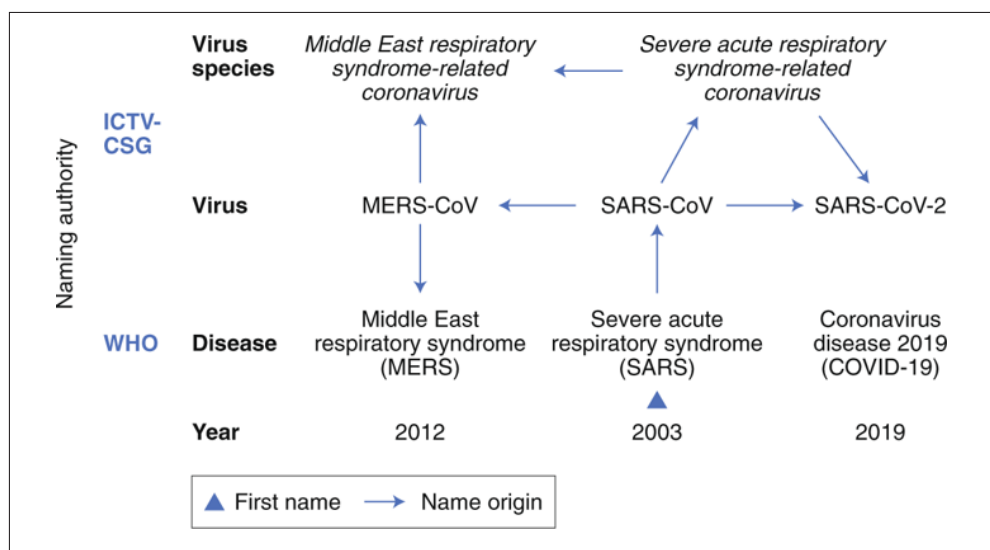


Fig. 1 Filogenia SARS-CoV-2 [2].  
Fig. 1 The phylogeny of SARS-CoV-2 [2].

tropism tisular. Uzual, alfacoronavirusurile și betacoronavirusurile infectează mamiferele, pe când gammacoronavirusurile și deltacoronavirusurile infectează păsările și peștii, dar unele dintre ele pot infecta și mamiferele. Înainte de 2019, existau doar șase coronavirusuri despre care se știa că infectează omul și cauzează boli respiratorii: HCoV-229E, HCoV-OC43, HCoV-NL63 și HKU1 [8].

### **Originea și gazdele intermediare**

Se știe că atât SARS-CoV, cât și MERS-CoV au provenit de la lilieci și au fost transmise oamenilor prin intermediul pisicilor civet și, respectiv, al cămășilor. Printr-o comparație filogenetică a SARS-CoV-2 cu alte coronavirusuri, lilieci au fost considerați gazda nativă a SARS-CoV-2, deoarece noul virus este identic în proporție de 96% cu două coronavirusuri de tip SARS ce infectează lilieci, numite bat-SL-CoVZ45 și bat-SL-CoVZ21 [9-12]. Cu toate acestea, gazda intermediară care a ajutat virusul să treacă de bariera dintre specii pentru a infecta omul rămâne necunoscută și calea de transmitere nu este, deocamdată, elucidată. Ji W. *et al.*, au propus șerpilor în calitate de vectori ai virusului de la lilieci la oameni, care au implicat recombinarea omoloagă în proteina S [13]. Potrivit unui studiu, cercetătorii din Guangzhou, China, au sugerat că pangolinele, folosite adesea în medicina tradițională chineză, sunt gazda intermediară potențială a SARS-CoV-2, bazându-se pe o omologie genetică de 99% dintr-un coronavirus descoperit la pangoline și SARS-CoV-2 [14]. Cu toate acestea, diferența de 1% dintre cele două genomuri este una semnificativă; astfel, sunt așteptate rezultate concludente pentru dovezi concrete [15].

### **Structura și reproducerea virusului**

Coronavirusurile sunt particule cu diametrul de 80-220 nm, acoperite de capsidă, genomul fiind reprezentat de o catenă de ARN. În particula de coronavirus, o nucleoproteină (N) înconjoară genomul ARN pentru a forma o structură spiralată tubulară. Anvelopa virală (A) înconjoară nucleocapsida helicală. Două sau trei proteine structurale sunt asociate cu anvelopa virală. Proteina matricială (M) este încorporată în anvelopa virală, pe când proteina structural Spike (S) este ancorată în anvelopa virală, și este ținta anticorpului de neutralizare. Numeroase betacoronavirusuri, de asemenea, conțin hemaglutinin-esteraza (Figura 2) [16, 17, 18].

Virusul utilizează proteina Spike glicozilată (S) pentru a pătrunde în celula gazdă și la oameni, unde se unește la receptorul pentru enzima de conversie a angiotensinei 2 (ACE2), similar cu SARS-CoV [19]. Genomul viral compus din ARN este eliberat în citoplasmă. După replicare, ARN-ul genomic viral este asociat cu anvelopa glicoproteică și proteinele nucleocapsidale, formează vezicule ce conțin virioni, care pot fuziona cu membrana plasmatică pentru a elibera virusul [20, 21].

Genomul CoV reprezintă o catena liniară unică pozitivă de ARN (+ssARN) (~30 kb) cu un cap 5' și o coadă 3'-polyA. ARN-ul genomic este utilizat în calitate de matriță pentru translarea directă a poliproteinei *1a/1ab* (*pp1a/pp1ab*), alcătuit din proteine non-structurale (*nsp*s) care formează un complex pentru replicare-transcripție [22]. Secvențierea inițială a genomului

tissue tropism. Usually, alfa-coronaviruses and beta-coronaviruses infect mammals, while gamma-coronaviruses and delta-coronaviruses infect birds and fish, but some of them can also infect mammals. Before 2019, there were only six coronaviruses known to infect humans and cause respiratory diseases: HCoV-229E, HCoV-OC43, HCoV-NL63 and HKU1 [8].

### **Origin and intermediate hosts**

It is known that both SARS-CoV and MERS-CoV came from bats and were transmitted to humans via civet cats and camel, respectively. By a phylogenetic comparison of SARS-CoV-2 with other coronaviruses, bats were considered the native host of SARS-CoV-2, because the new virus is 96% identical to two SARS-type coronaviruses that infect bats called bat-SL-CoVZ45 and bat-SL-CoVZ21 [9-12]. However, the intermediate host that helped the virus cross the interspecific barrier to infect humans remains unknown and the route of transmission is not elucidated yet. Ji W. *et al.*, proposed snakes as vectors of the virus from bats to humans involving homologous recombination in S protein [13]. According to a study, researchers in Guangzhou, China, have suggested that pangolins, often used in traditional Chinese medicine, are the potential intermediate host of SARS-CoV-2, based on a 99% genetic homology of a coronavirus found in pangolins, and SARS-CoV-2 [14]. However, the 1% difference between the two genomes is significant; thus, conclusive results are expected for concrete evidence [15].

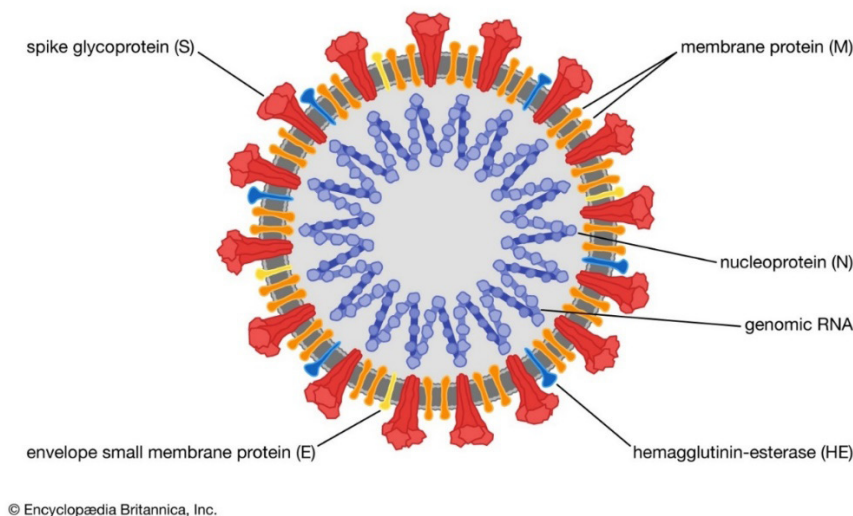
### **The structure and reproduction of the virus**

Coronaviruses are particles with a diameter of 80-220 nm, covered by a capsid, the genome is represented by an RNA strand. In the coronavirus particle, a nucleoprotein (N) surrounds the RNA genome to produce a tubular spiral structure. The viral envelope (E) covers the helical nucleocapsid. Two or three structural proteins are associated with the viral envelope. The matrix protein (M) is embedded in the viral envelope, while the Spike structural protein (S) is anchored in the viral envelope and is the target of the neutralizing antibody. Many beta-coronaviruses also contain hemagglutinin esterase (Figure 2) [16, 17, 18].

The virus uses glycosylated Spike protein (S) to enter the host cell and in humans binds to the angiotensin converting enzyme 2 (ACE2) receptor, similar to SARS-CoV [19]. The viral genome, which consists of RNA, is released into the cytoplasm. After replication, viral genomic RNA is associated with the glycoprotein envelope and nucleocapsid proteins, forming virion-containing vesicles, which can fuse with the plasma membrane to release the virus [20, 21].

The CoV is a positive-sense single-stranded RNA virus, with linear genome (+ssRNA) (~30 kb) with a 5' head and a 3'-polyA tail. Genomic RNA is used as a template for direct translation of polyprotein *1a/1ab* (*pp1a/pp1ab*), which consists of non-structural proteins (*nsp*s) that form a replication-transcription complex [22]. The initial sequencing of the SARS-CoV genome demonstrated taxonomic status as a member of the *Betacoronaviridae* group, with clear divergences from

### Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)



**Fig. 2** Structura coronavirusului SARS-CoV-2, agentul pandemiei COVID-19 [18].  
**Fig. 2** The structure of coronavirus SARS-CoV-2, the cause of the COVID-19 pandemic [18].

SARS-CoV a demonstrat statutul taxonomic în calitate de membru al grupului *Betacoronaviridae*, cu divergențe clare față de SARS-CoV și MERS-CoV. De asemenea, au fost identificate ținte potențiale pentru preparatele farmacologice, ținând cont de gradul de similaritate între hCoV-19 și SARS. Atât proteaza, cât și polimeraza, care sunt ținte pentru preparate farmacologice, sunt înalt conservate între hCoV-19 și SARS (similaritate de 96% și 97%, respectiv). Inhibitorii creați împotriva SARS-CoV, de asemenea, au arătat un potențial de a se uni la proteaza sau polimeraza CoV-19 [3]. Moleculele, precum caveolinele, clatrinele, dinamina, au un rol fundamental în internalizarea virusurilor în celulele gazdă [23].

Glicoproteinele S ale virionului se unesc la receptorul celular pentru ACE2 și intră în celula țintă prin intermediul căii endozomale. În celula-gazdă ARN viral este localizat în citoplasmă. *ORF1a* și *ORF1ab* sunt translate pentru a produce poliproteinele *pp1a* și *pp1ab*, care sunt clivate de către proteazele complexului de replicare / transcripție - RTC. În timpul replicării, RTC induce sinteza unei copii integrale de ARN(-) a genomului viral care reprezintă o matrice pentru producerea de ARN(+). În timpul transcripției un set de ARN sub-genomic (sgARN) este produs prin transcripție discontinuă (transcripție fragmentată). După ce se sintetizează proteinele structurale proprii SARS-CoV-2, nucleocapsidele sunt asamblate în citoplasmă și urmează încorporarea în lumenul reticulului endoplasmatic și compartimentul intermediar al Aparatului Golgi. Virionii sunt eliberați din celulele infectate prin exocitoză [24].

#### Organizarea genomului SARS-CoV-2

Genomul ARN(+) al SARS-CoV-2 diferă de genomurile SARS-CoV și MERS-CoV, având lungimea aproximativă de 29,9 kb, față de 27,9 kb și respectiv 30,1 kb. Genomul SARS-CoV-2 conține 14 faze de lectură deschise (ORF) și codifică 27 proteine. Primul ORF, care constituie 2/3 din genom, determină sin-

SARS-CoV and MERS-CoV. Potential drug targets have also been identified, taking into account the degree of similarity between hCoV-19 and SARS. Both protease and polymerase, which are targets for drugs, are highly conserved between hCoV-19 and SARS (similarity of 96% and 97%, respectively). Inhibitors created against SARS-CoV have also shown a potential to bind to CoV-19 protease or polymerase [3]. Molecules such as caveolin, clathrin, dynamin have a fundamental role in the internalization of viruses into host cells [23].

The S glycoproteins of the virion bind to the cellular receptor for ACE2 and enter the target cell via the endosomal pathway. In the host cell, viral RNA is located in the cytoplasm. *ORF1a* and *ORF1ab* are translated to produce the *pp1a* and *pp1ab* polyproteins, which are cleaved by the proteases of the replication / transcription complex - RTC. During replication, RTC induces the synthesis of an integral copy of RNA(-) of the viral genome, which represents a template for the production of RNA(+). During transcription, a set of sub-genomic RNA (*sgRNA*) is produced by discontinuous transcription (fragmented transcription). After synthesizing the own SARS-CoV-2 structural proteins, the nucleocapsids are assembled in the cytoplasm and followed by budding into the lumen of the endoplasmic reticulum and the intermediate compartment of the Golgi Apparatus. The virions are released from infected cells by exocytosis [24]

#### The genome organization of the COVID-19

The SARS-CoV-2 RNA (+) genome differs from the SARS-CoV and MERS-CoV genomes, having an approximate length of 29.9 kb, compared to 27.9 kb and 30.1 kb, respectively. The SARS-CoV-2 genome contains 14 open reading frames (ORFs) and encodes 27 proteins. The first ORF, which constitutes 2/3 of the genome, determines the synthesis of two polyproteins

teza a două poliproteine, *pp1a* și *pp1ab*, care produc 16 proteine non-structurale: *NSP2* (cistein-proteaza virală), *NSP3* (proteaza asemănătoare cu papaina), *NSP5* (proteaza principală), *NSP12* (ARN-polimeraza-ARN-dependență), *NSP13* (helicaza) și alte NSP-uri care sunt, probabil, implicate în transcripția și replicarea ARN-ului viral. ORF-urile rămase codifică proteine structurale, inclusiv glicoproteina Spike (proteina S), proteina mica anvelopară (E), proteina membranară (M) și proteina nucleocapsidală (N). SARS-CoV-2 posedă proteine accesorii care interferează cu răspunsul imun al gazdei [23]. Ca și alte virusuri bazate pe ARN, coronavirusurile prezintă plasticitate genomică determinată de numeroși factori. ARN polimerazele ARN dependente (*RdRp*) prezintă o rată de mutații mare, ajungând de la 1/1000 până la 1/10.000 de nucleotide în timpul replicării. De asemenea, sunt cunoscute mecanisme de recombinare între genomurile virale. Dat fiind faptul că genomul este relativ mare, rata mutațiilor este, de asemenea, înaltă. Toate acestea contribuie la plasticitatea genomului și diversitatea formelor de coronavirus [25].

#### **Interacțiunea dintre SARS-CoV-2 și receptorii celulari**

Proteinele de fuziune mediază pătrunderea în celula-gazdă care poate fi activată de unirea la receptor, pH scăzut sau o combinație dintre acești 2 factori [26]. Aceste proteine sunt distribuite în 3 clase structurale: I, II și III. Proteina Spike a coronavirusului este o proteină de fuziune virală de clasa I [27] și a fost identificată la toate speciile de Betacoronavirus. Proteinele de fuziune virală de clasa I suportă o succesiune de modificări conformaționale specifice care asigură fuziunea membranară. Studiarea caracteristicilor structurale ale proteinei Spike și interacțiunea acesteia cu receptorul ACE2 la oameni sau alte specii, poate permite obținerea informației cu privire la originea, evoluția, virulența, diagnosticul, terapia și producerea vaccinurilor contra SARS-CoV-2.

SARS-CoV-2 utilizează receptorul ACE2, același receptor, ca și SARS în tractul respirator inferior la om pentru pătrunderea în celulele pulmonare. Serin-proteaza de la suprafața celulară (TMPRSS2), de asemenea, este utilizată de către SARS-CoV-2 pentru atașarea proteinei Spike, ce facilitează fuziunea membranară în celulele-gazdă [28].

#### **Structura proteinei Spike**

Glicoproteina Spike, numită și proteina S, este un homotrimer, dar structural este compusă din două subunități, S1 și S2, care se disting funcțional: S1 este localizată mai distal de membrana virală și este responsabilă pentru interacțiunea cu receptorul, în timp ce S2 este o componentă necesară pentru fuziunea membranelor. S1 poate fi divizată în domeniul N-terminal (NTD), localizat mai distal, și domeniul C-terminal (CTD), ce este situat mai proximal de axul median al proteinei. Ambele pot funcționa în calitate de liganzi pentru receptori [29]. S1 CTD este adesea descris în literatură precum *domeniul de legare a receptorului* (RBD), deoarece SARS-CoV și MERS-CoV le folosesc pentru a se uni de receptori [30]. Proteina S a SARS-CoV-2 folosește CTD pentru a interacționa cu hACE2 [29].

Bineînțeles, structura subunității S2 are o deosebită în-

- *pp1a* and *pp1ab*, which produce 16 non-structural proteins: (NSP) 2 viral cysteine proteases, NSP3 (papain-like protease), NSP5 (main protease), NSP12 (RNA-dependent-RNA-polymerase), NSP13 (helicase) and other NSPs that are likely to be involved in the transcription and replication of viral RNA. The remaining ORFs encode structural proteins, including Spike glycoprotein (S protein), small envelope protein (E), membrane protein (M), and nucleocapsid protein (N). SARS-CoV-2 possesses accessory proteins that interfere with the host's immune response [23]. Like other RNA-based viruses, coronaviruses have genomic plasticity caused by many factors. RNA-dependent RNA polymerases (RdRp) have high mutation rates, reaching from 1/1000 to 1/10,000 nucleotides during replication. Also known mechanisms of recombination between viral genomes. Taking into consideration that the genome is relatively large, the rate of mutations is high too. All these contribute to the plasticity of the genome and the diversity of coronavirus forms [25].

#### **Interaction of SARS-CoV-2 with cellular receptors**

Viral fusion proteins mediate penetration into the host cell that can be activated by receptor binding, low pH, or a combination of these two factors [26]. These proteins are divided into 3 structural classes: I, II and III. Coronavirus Spike protein is a class I virus fusion protein [27] and is found in all identified species of Betacoronavirus. The class I viral fusion proteins undergo a specific sequence of conformational changes that achieve membrane fusion. Exploration of the Spike protein's structural features and its interaction with the ACE2 receptor of humans, or other species, provides insight into SARS-CoV-2 origin, evolution, infectivity, diagnostics, therapeutics, and vaccine development.

SARS-CoV-2 exploits the ACE-2 receptor, the same receptor as SARS, in the human lower respiratory tract for entry into lung cells. Cellular surface serine protease TMPRSS2 is also used by SARS-CoV-2 for priming of spike protein S to facilitate membrane fusion with host cells [28].

#### **Spike protein structure**

The Spike glycoprotein, further referred to as the S protein, is a homotrimer, but structurally it is composed of two subunits, S1 and S2, which are functionally distinct: S1 is located farther from the viral membrane and is responsible for receptor interaction, while S2 is the machinery necessary for membrane fusion. S1 can be further divided into an N-terminal domain (NTD), located farther from the long axis of the molecule, and a C-terminal domain (CTD), which is closer to the midline. Both can function as receptor binders [29]. S1 CTD is commonly referred to in the literature as the *receptor binding domain* (RBD), because SARS-CoV and MERS-CoV use it to bind to receptors [30]. SARS-CoV-2 S protein uses CTD to interact with hACE2 [29]. The S protein is a trimer and thus has three S1 CTD domains. RBD harbors the *receptor binding motif* (RBM).

Of course, the structure of the S2 subunit is also important to investigate, since it is the fusion hardware and thus can set the stage for rational design of coronavirus fusion inhibitors

semnătate pentru a fi investigată, deoarece este o subunitate importantă în cadrul fuziunii și poate reprezenta cheia pentru designul unui inhibitor de fuziune membranară a coronavirusurilor [31]. Proteina S fiind un homotrimer, orice monomer are o subunitate S1 și una S2. Subunitatea S2 a fiecărui monomer posedă două domenii: heptad repeat 1 (HR1) și heptad repeat 2 (HR2) (Figura 3) [32]. HR1 și HR2 din cei 3 monomeri interacționează pentru a alcătui o formațiune numită *six-helix bundle structure* – 6-HB, sau nucleu de fuziune [27]. Structura 6-HB, formată de regiunile HR1 și HR2 în subunitatea S2 a coronavirusului uman joacă un rol cheie în timpul procesului de fuziune membranară, ce o face una din cele mai importante ținte pentru designul preparatelor [31].

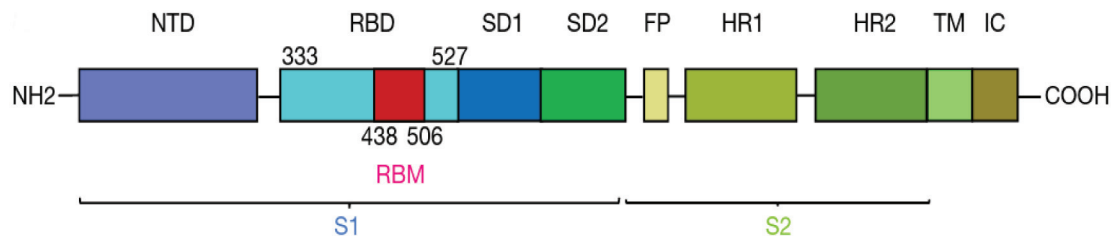
Alt moment cheie este faptul că proteina Spike este o glicoproteină. Trimerii S conțin glicani uniți la capătul N-terminal care sunt importanți atât pentru conformarea corectă, modularea accesibilității proteazelor gazdei cât și neutralizarea anticorpilor (Figura 3) [33]. Proteina S poate prezenta multiple conformații. Per general, se poate de zis că statutul proteinelor de fuziune virală poate fi de prefuziune și postfuziune. Starea predominantă de prefuziune a trimerului S are unul din cei trei RBD pliați într-o conformație accesibilă pentru interacțiunea cu receptorul [34].

[31]. The S protein being a homotrimer, each monomer has a S1 and S2 subunit. The S2 subunit of a monomer further has two domains: heptad repeat 1 (HR1) and heptad repeat 2 (HR2) (Figure 3) [32]. HR1 and HR2 from the 3 monomers interact to form a *six-helix bundle structure*, abbreviated 6-HB, also referred to as the fusion core [27]. The 6-HB structure formed by HR1 and HR2 regions in the S2 subunit of HCoV-2019 plays a key role during the viral membrane fusion process, which makes it one of the most important targets for drug design [31].

Another point concerns the fact that the Spike protein is a glycoprotein. S trimers are extensively decorated with N-linked glycans that are important for proper folding and for modulating accessibility to host proteases and neutralizing antibodies [33]. S protein can be found in multiple conformations. Generally, viral fusion proteins states can be classified as prefusion and postfusion. The predominant prefusion state of the S trimer has one of the three RBDs rotated up in a receptor-accessible conformation [34].

#### **Spike protein preprocessing**

A distinctive feature of the SARS-CoV-2 S glycoprotein is a novel furin cleavage site at the boundary between S1 and



**Fig. 3** Topologia monomerului Spike al SARS-CoV-2. Adaptat după [32].

**Fig. 3** The topology of SARS-CoV-2 spike monomer. Adapted from [32].

#### **Pre-procesarea proteinelor Spike**

O caracteristică distinctivă a glicoproteinei S SARS-CoV-2 este prezența unui situs nou de clivare a furinei la nivelul joncțiunii dintre S1 și S2. Situsul de clivare este procesat de către furină și, probabil, are loc în Aparatul Golgi în timpul biosintezei vitale [33]. Furina este o protează înalt prevalentă, implicată în activarea multiplelor preproteine, precum pro-PTH, pro-albumină și pro-NGF. Situsul nou de clivare poate determina o creștere în tropismul viral.

De notat că toate CoV au domeniul S1 clivat de la S2, deoarece acest pas este necesar pentru demararea infecției, însă există deosebiri în ce privește enzimele și momentul când are loc clivarea [33]. Unele studii raportează că clivarea proteinei S de către cathepsina L are loc în endosomi, ceea ce indică asupra faptului că fuziunea membranară are loc după endocitoza mediată de receptori [35]. Deoarece mai multe studii demonstrează importanța acestor proteaze, ele pot deveni ținte importante pentru elaborarea medicamentelor.

Experimentele prin care se elimină situsul pentru furină în SARS-CoV-2 demonstrează că acesta nu este necesar pentru fuziunea membranară S-mediată, ceea ce indică asupra faptului că clivarea în situsul SARS-CoV-2 S1/S2 poate avea loc cu

S2. This cleavage site is processed during viral biosynthesis, probably in the Golgi Apparatus by furin [33]. Furin is a highly prevalent protease involved in the activation of multiple pre-proteins, such as pro-PTH, pro-albumin and pro-NGF. The new cleavage site may determine an increase in the viral tropism.

It is important to note that all CoVs have the S1 cleaved from S2, since it is a necessary step for infection, but what distinguishes them is the enzymes and moment in time when the cleavage happens [33]. Some studies report that cleavage of the S protein by cathepsin L takes place in endosomes, supporting the view that the membrane fusion happens after receptor-mediated endocytosis [35]. As many studies show the importance of these proteases, they can become emerging targets for drug design.

Experiments which remove the furin cleavage site in SARS-CoV-2 show that it is not necessary for the S-mediated membrane fusion, implying that SARS-CoV-2 S1/S2 cleavage can happen via the same proteases as other SARS-related CoVs, namely TMPRSS2 or cathepsin L [33, 36].

In conclusion, coronavirus infection requires both receptor binding and proteolytic processing [33]. Altered proteolytic

ajutorul acelorași proteaze, ca și în cazul altor CoV, adică cu ajutorul TMPRSS2 sau catepsinei L [33, 36].

Ca și concluzie, infecția cu coronavirus necesită atât atașarea la receptor, cât și procesarea proteolitică [33]. Alterarea procesării proteolitice poate explica modificările în tropism, patogenitate și urmează a fi luate în considerație în cazul transmiterii de la o specie la alta, deoarece unele specii nu dispun de același mecanism de procesare.

### **Interacțiunea dintre proteina Spike și receptor**

SARS-CoV-2 se unește cu o afinitate înaltă de receptorul ACE2 uman și îl folosește pentru intrare în celulele țintă [33]. Receptorii ACE2 pot fi identificați pe suprafața multor tipuri de celule, inclusiv pneumocitele de tip II. SARS-CoV-2 și SARS-CoV recunosc receptorul ACE2 cu afinitate comparabilă [33], însă sunt studii care sugerează o afinitate crescută pentru proteina S proprie SARS-CoV-2. Analiza interfeței dintre RBD caracteristic pentru SARS-CoV și ACE2 a relevat un total de 16 aminoacizi pentru RBD-SARS-CoV ce se află în contact cu 20 de resturi de aminoacizi ai ACE2 [32]. Rezultă că virusurile SARS-CoV au evoluat pentru a recunoaște receptorul ACE2. Totodată, studiile lui Brielle E. *et al.* [37] sugerează că proteina S-SARS-CoV-2 utilizează o strategie proprie de atașare la ACE2.

Interacțiunea S-ACE2 este etapa crucială în infecție, determinând un rol important. Unii anticorpi neutralizanți din plasma pacienților care s-au însănătoșit într-adevăr se atașează la RBD [38], însă nu toți anticorpii neutralizanți sunt capabili de aceasta. În pofida similarității înalte între diferite SARS-CoV, anticorpii care acționează asupra altor specii nu sunt eficienți contra SARS-CoV-2, fapt determinat de particularitățile proteinei S [29].

### **Concluzii**

În pofida eforturilor globale de a ține sub control infecția cu SARS-CoV-2, răspândirea virusului a atins un nivel de pandemie. Elaborarea medicamentelor pentru tratamentul COVID-19 și a vaccinurilor este, deocamdată, în dezvoltare, dar s-au înregistrat progrese semnificative în studierea multilaterală a genomului SARS-CoV-2, fapt ce permite identificarea exactă a acestuia. Cooperarea diferitor instituții, centre științifice, guverne și companii farmaceutice este necesară pentru a preveni răspândirea în continuare a COVID-19 și identificarea soluțiilor de tratament.

### **Contribuția autorilor**

Toți autorii au contribuit egal la elaborarea și scrierea articolului. Versiunea finală a articolului a fost citită și aprobată de toți autorii.

### **Declarația conflictului de interese**

Nimic de declarat.

processing can explain modifications in tropism, pathogenicity, but also is an important consideration when investigating the cross-species viral transmission, since some species may not harbor the necessary machinery for adequate processing.

### **Spike protein receptor interaction**

SARS-CoV-2 binds with high affinity to human ACE2 and uses it as an entry receptor to invade target cells [33]. ACE2 can be found on the surface of multiple cell types, including type II pneumocytes. SARS-CoV-2 and SARS-CoV recognize ACE2 with comparable affinity [33], but there are studies that suggest an increased affinity of the SARS-CoV-2 S protein. Analysis of the interface between the SARS-CoV RBD and ACE2 put in evidence a total of 16 residues of the SARS-CoV RBD in contact with 20 residues of ACE2 [32]. It turns out that SARS-CoV viruses have evolved to recognize the ACE2 receptor. However, a study by Brielle E. *et al.* [37] suggests that the SARS-CoV-2 S protein leverages a different strategy for binding ACE2.

The S-ACE2 interaction is a crucial step during infection, it has significant practical implications. Some neutralizing antibodies from convalescent patients' plasma do indeed bind the RBD [38], but it must be noted that it's not necessary that all neutralizing antibodies are able to do that. Despite the high similarity between different SARS CoV, antibodies that act on other species are not effective against SARS-CoV-2, due to the peculiarities of protein S [29].

### **Conclusions**

Despite global efforts to control the SARS-CoV-2 infection, the spread of the virus has reached a pandemic level. The development of drugs for the treatment of COVID-19 and vaccines is still in development, but significant progress has been made in the multilateral study of the SARS-CoV-2 genome, which allows its exact identification. Cooperation between various institutions, scientific centers, governments and pharmaceutical companies is needed to prevent the further spread of COVID-19 and to identify treatment solutions.

### **Authors' contributions**

The authors contributed equally to the work done. All authors have read and approved the final version of the article.

### **Declaration of conflicting interests**

Nothing to declare.



## Referințe/references

- Li Y, Zhang Z, Yang L, Lian X, Xie Y, Li S. *et al.* The MERS-CoV receptor DPP4 as a candidate binding target of the SARS-CoV-2 spike. *iScience*, 2020; 101160.
- Coronaviridae Study Group of the International Committee on Taxonomy. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.*, 2020; 5 (4): 536-544.
- Rodriguez-Morales A., Balbin-Ramon G., Rabaan A., Sah R., Dhama K. Genomic epidemiology and its importance in the study of the COVID-19 pandemic. *Infez. Med.*, 2020; 28 (2): 139-142.
- Siddell S., Walker P., Lefkowitz E. *et al.* Additional changes to taxonomy ratified in a special vote by the International Committee on Taxonomy of Viruses (October 2018). *Arch. Virol.*, 2019; 164 (3): 943-946.
- Cui J., Li F., Shi Z. Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.*, 2019; 17 (3): 181-192.
- [https://talk.ictvonline.org/International Committee on Taxonomy of Viruses-ICTV](https://talk.ictvonline.org/International%20Committee%20on%20Taxonomy%20of%20Viruses-ICTV) (Accessed at: 10.06.2020).
- [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200211-sitrep-22-ncov.pdf?sfvrsn=fb6d49b1\\_2](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200211-sitrep-22-ncov.pdf?sfvrsn=fb6d49b1_2), Novel Coronavirus (2019-nCoV). Situation Report – 22.
- Chen, Y., Liu Q., Guo D. Emerging coronaviruses: genome structure, replication, and pathogenesis. *J. Med. Virol.*, 2020; 92 (4): 418-423.
- Dong N., Yang X., Ye L., Chen K., Chan E. *et al.* Genomic and protein structure modelling analysis depicts the origin and infectivity of 2019-nCoV, a new coronavirus which caused a pneumonia outbreak in Wuhan, China. *bioRxiv*, 2020; 2020.01.20.913368.
- Xu X., Chen P., Wang J., Feng J., Zhou H., Li X. *et al.* Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Sci. China Life Sci.*, 2020; 63 (3): 457-460.
- 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): 565-574.
- Zhou P., Yang X., Wang X., Hu B., Zhang L., Zhang W. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 2020; 579 (7798): 270-273.
- Ji W., Wang W., Zhao X., Zai J., Li X. Cross-species transmission of the newly identified coronavirus 2019-nCoV. *J. Med. Virol.*, 2020; 92 (4): 433-440.
- Xiao K., Zhai J., Feng Y., Zhou N., Zhang X., Zou J. *et al.* Isolation and characterization of 2019-nCoV-like Coronavirus from Malayan pangolins. *bioRxiv*, 2020; 2020.02.17.951335.
- Yi Y., Lagniton P., Ye S., Li E., Xu R. COVID-19: what has been learned and to be learned about the novel coronavirus disease. *Int. J. Biol. Sci.*, 2020; 16 (10): 1753-1766.
- Brian D., Baric R. Coronavirus genome structure and replication. *Curr. Top Microbiol. Immunol.*, 2005; 287: 1-30.
- Park S. Epidemiology, virology, and clinical features of Severe Acute Respiratory Syndrome coronavirus-2 (SARS-CoV-2; Coronavirus Disease-19). *Clin. Exp. Pediatr.*, 2020; 63 (4): 119-124.
- <https://www.britannica.com/science/coronavirus-virus-group>. Encyclopaedia Britannica.
- Raj V., Mou H., Smits S., Dekkers D., Muller M., Dijkman R. *et al.* Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature*, 2013; 495 (7440): 251-4.
- Del Rio C., Malani P. COVID-19: new insights on a rapidly changing epidemic. *JAMA*, 2020.
- Sevajol M., Subissi L., Decroly E., Canard B., Imbert I. Insights into RNA synthesis, capping and proofreading mechanisms of SARS-coronavirus. *Virus Res.*, 2014; 194: 90-9.
- Abduljalil J., Abduljalil B. Epidemiology, genome, and clinical features of the pandemic SARS-CoV-2: a recent view. *New Microbes New Infect.*, 2020; 35: 100672.
- Dawood A. Mutated COVID-19 may foretells mankind in a great risk in the future. *New Microbes New Infect.*, 2020; 100673.
- Alanagreh L., Alzoughool F., Atoum M. The human coronavirus disease COVID-19: its origin, characteristics, and insights into potential drugs and its mechanisms. *Pathogens*, 2020; 9 (5).
- Randhawa G., Soltysiak M., El Roz H., de Souza C., Hill K., Kari L. Machine learning using intrinsic genomic signatures for rapid classification of novel pathogens: COVID-19 case study. *PLoS One*, 2020; 15 (4): e0232391.
- White J., Delos S., Brecher M., Schornberg K. Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. *Crit. Rev. Biochem. Mol. Biol.*, 2008; 43 (3): 189-219.
- Bosch B., van der Zee R., de Haan C., Rottier P. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J. Virol.* 2003; 77 (16): 8801-11.
- 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.
- 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): 894-904e9.
- Taguchi F., Hirai-Yuki A. Mouse hepatitis virus receptor as a determinant of the mouse susceptibility to MHV infection. *Front. Microbiol.*, 2012; 3: 68.
- Xia S., Liu M., Wang C., Xu W., Lan Q., Feng S. *et al.* Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. *Cell Res.*, 2020; 30 (4): 343-355.
- 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): 215-220.
- Walls A., Park Y., Tortorici M., Wall A., McGuire A., Veesler D. Structure, function and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*, 2020; 181 (2): 281-292 e6.
- Wrapp D., Wang N., Corbett K., Goldsmith J., Hsieh C., Abiona O. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, 2020; 367 (6483): 1260-1263.
- Yan R., Zhang Y., Li Y., Xia L., Guo Y., Zhou Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science*, 2020; 367 (6485): 1444-1448.
- Millet J., Whittaker G. Host cell proteases: critical determinants of coronavirus tropism and pathogenesis. *Virus Res.*, 2015; 202: 120-34.
- Brielle E., Schneidman-Duhovny D., Linial M. The SARS-CoV-2 exerts a distinctive strategy for interacting with the ACE2 human receptor. *Viruses*, 2020; 12 (5).
- Wu Y., Wang F., Shen C., Peng W., Li D., Zhao C. *et al.* A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2. *Science*, 2020.