IRSTI 578.834.1 https://doi.org/10.26577/EJE.2022.v71.i2.05

B.S. Usserbayev*, Ye.D. Burashev, N.S. Kozhabergenov, A.M. Melisbek, M.Zh. Shirinbekov, Ye.O. Abduraimov, K.D. Zakarya

[1]Ministry of Education and Science of the Republic of Kazakhstan Science Committee
«Research Institute for Biological Safety Problems»,
Zhambyl region, Kordai district, uts.Gvardeiskiy, Kazakhstan
*e-mail: usserbayev.bekbolat@mail.ru

DYNAMICS OF THE SPREAD OF *SARS-COV*-2 VARIANTS AND CLADES

In the XX century, discovered Human Coronavirus, including HCoV-229E, HCoV-NL63 from the genus Alphacoronavirus and HCoV-OC43, HCoV-HKU1 from the genus Betacoronavirus have long been considered insignificant pathogens for humans, causing infections in the upper respiratory tract in adults. However, at the beginning of the XXI century, highly pathogenic coronaviruses of the severe acute respiratory syndrome (SARS-CoV) and coronavirus of the Middle East respiratory syndrome (MERS-CoV) caused global epidemics and eventually led to death. In December 2019, Wuhan, China, revealed another new Human CoV-2019. On March 11, 2020, WHO announced a new global pandemic of Human CoV – 2019 due to its global spread within a short time. As of September 23, 2021, according to WHO data, a total of 229,858,719 confirmed cases of COVID-19 and 4,713,543 deaths were registered. The emergence and global spread of SARS-CoV-2 coronavirus variants remain largely unknown. But the spread of COVID-19 around the world increases interest in the disclosure of genomic mutations of a new variant of the coronavirus SARS-CoV-2. Data on the epidemiology of diseases, including the number of cases and mortality in different parts of the world and the Republic of Kazakhstan, are presented. The schematic nomenclature of GISAID, Nextstrain, and Pango lines is compared. The sequence of the SARS-CoV-2 genome submitted to the GISAID database for the Republic of Kazakhstan was analyzed concerning genomic clades and their geographical, age and gender distribution.

Key words: COVID-19, SARS-CoV-2, clade.

Б.С. Усербаев*, Е.Д. Бурашев, Н.С. Кожабергенов, А.М. Мелисбек,
М.Ж. Ширинбеков, Е.О. Абдураимов, К.Д. Закарья

«Биологиялық қауіпсіздік проблемаларының ғылыми-зерттеу институты» ШЖҚ РМК ҚР БҒМ ҒК,
Жамбыл облысы, Қордай ауданы, Гвардейский құк, Қазақстан
*e-mail: usserbayev.bekbolat@mail.ru

SARS-CoV-2 варианттары мен клайдтарының таралуының динамикасы

Аннотация. XX ғасырда ашылған Human Coronavirus, соның ішінде Alphacoronavirus тұқымдасынан болған HCoV-229E, HCoV-NL63 және Betacoronavirus тұқымдасынан болған HCoV-OC43, HCoV-HKU1 ересек адамдардың жоғары тыныс алу жолында инфекция тудырғанмен, ұзақ уақыт бойы адамдар үшін қауіпсіз патоген саналды. Алайда XXI ғ. басында туындаған жоғары патогенді, жануар резервуарынан шыққан коронавирустар ауыр жедел респираторлық синдром (SARS-CoV) және таяу шығыс респираторлық синдромының коронавирусы (MERS-CoV) әлемдік эпидемия тудырып, соңы өлімге алып келді. 2019 ж. желтоқсан айында Ухан қ. ҚХР, тағы бір жаңа Human CoV – 2019 анықталды. Аз уақыт ішінде ғаламдық таралуына орай, 11 наурыз 2020 ж ДДСҰ жаңа Human CoV – 2019 әлемдік пандемия жариялады. 2021 жылғы 23 қыркүйектегі ДДСҰ мәліметтері бойынша 229,858,719 адам COVID-19 жұқтырғаны расталып, 4,713,543 адамның өлім жағдайы тіркелген**.** SARS-CoV-2 коронавирусының пайда болуы және ғаламдық таралуы әзірше белгісіз. Бірақ COVID-19-ның бүкіл әлемде таралуы SARS-CoV-2 коронавирусының жаңа нұсқасының геномдық мутациясын ашуға деген қызығушылықты арттыруда. Бұл мақалада аурулар эпидемиологиясының деректері, соның ішінде әлем мен ҚР-ның әртүрлі бөліктеріндегі жағдайлар мен өлім-жітім саны ұсынылған. GISAID, Nextstrain және линия Pango бойынша схемалық номенклатурасы салыстырылды. ҚР бойынша GISAID дерекқорына ұсынылған SARS-CoV-2 геномының геномдық клайдтарға және олардың географиялық, жас және гендерлік таралуына қатысты реттілігі талданды.

Түйін сөздер: COVID-19, SARS-CoV-2 , клайд.

Б.С. Усербаев*, Е.Д. Бурашев, Н.С. Кожабергенов, А.М. Мелисбек,
М.Ж. Ширинбеков, Е.О. Абдураимов, К.Д. Закарья
РГП на ПХВ «Научно-исследовательский институт проблем биологической безопасности», КН МОН РК,
Жамбылская область, Кордайский район, пгт. Гвардейский, Казахстан
*e-mail: usserbayev.bekbolat@mail.ru

Динамика распространения вариантов и клайдов *SARS-CoV*-2

В XX веке обнаруженные Human Coronavirus, включая HCoV-229E, HCoV-NL63 из рода Alphacoronavirus и HCoV-OC43, HCoV-HKU1 из рода Betacoronavirus долгое время считался несущественными патогенами для человека, вызывая инфекции в верхних дыхательных путях у взрослых. Однако в начале XXI в. высокопатогенные коронавируса тяжелого острого респираторного синдрома (SARS-CoV) и коронавируса ближневосточного респираторного синдрома (MERS-CoV) вызывали глобальные эпидемии, а в конце концов привели к смерти. В декабре 2019 года. г. Ухань, КНР, выявлен еще один новый Human CoV-2019. 11 марта 2020 года ВОЗ объявила о новой глобальной пандемии Human CoV – 2019 в связи с ее глобальным распространением в течение короткого времени. По состоянию на 23 сентября 2021 г. согласно данным ВОЗ зарегистрировано в общей сложности 229,858,719 подтвержденных случаев COVID-19 и 4,713,543 летальных исходов. Возникновение и глобальное распространение вариантов коронавируса SARS-CoV-2 остаются в значительной степени неизвестными. Но распространение COVID-19 по всему миру увеличивает интерес к раскрытию геномных мутаций нового варианта коронавируса SARS-CoV-2. В этой обзорной статье представлены данные эпидемиологии заболеваний, в том числе количество случаев и летальности в разных частях мира и РК. Сопоставлена схематическая номенклатура GISAID, Nextstrain и линий Pango. Проанализирована последовательность генома SARS-CoV-2, представленного в базу данных GISAID по РК, в отношении геномных клайдов и их географического, возрастного и гендерного распределения.

Ключевые слова: COVID-19, SARS-CoV-2, клайд.

Abbreviations

SARS-CoV-2 – Severe acute respiratory syndrome-related coronavirus 2; *COVID*-19 – Coronavirus disease 2019; *WHO*-World Health Organization.

Introduction

Coronaviruses (CoV; Coronaviridae family) are a group of very diverse, enveloped, single-stranded *RNA* viruses that cause respiratory and intestinal diseases of varying severity in a certain types of animals, including humans *[1].* According to the classification of the *International Committee on Taxonomy of Viruses (ICTV),* currently, according to phylogenetic clustering, the *Orthocoronavirinae (Tab.1)* subfamily includes four genera of coronavirus*: Alphacoronavirus, Betacoronavirus, Deltacoronavirus, Gammacoronavirus* [2,3].

According to Table 1, before the new variant of the *SARS-COV-2* coronavirus in 2019, only six representatives of the coronavirus were known to have infected people and caused respiratory diseases. HCoV-229E and *HCoV-NL63* from the genus *Alphacoronavirus* and *HCoV-OC43*, *HCoV-HKU1* from the genus *Betacoronavirus* cause only mild upper respiratory tract diseases*. Severe Acute Respiratory Syndrome coronavirus* (*SARS-CoV*) and the *Middle East Respiratory Syndrome coronavirus (MERS-CoV)* can infect the lower respiratory tract and cause a severe respiratory syndrome in humans [3-7].

In December 2019, several patients with pneumonia of unknown etiology were identified in Wuhan, Hubei Province, China. With the help of molecular genetic analysis revealed a new etiological agent coronavirus called “severe acute respiratory syndrome coronavirus 2 (*SARS–CoV*-2)”. *SARS-CoV*-2 is a single-stranded positive (+) *RNA* virus, measuring 29903 bp (GenBank number MN908947) in length and encodes 9860 amino acids [8,9]. *SARS-CoV-2* belongs to the subgenus *Sarbecovirus,* genus *Betacoronavirus,* subfamily *Orthocoronavirinae,* belong to the family *Coronaviridae* in the order *Nidovirales* [2].

Table 1 – Taxonomy of the *subfamily Orthocoronavirinae* (https://talk.ictvonline.org/).

Order	***Subor-der***	***Fam-ily***	***Sub-family***	***Genus***	***Subgenus***	***Species***
Nidovirales	*Cornidovirineae*	*Coronaviridae*	*Orthocoronavirinae*	*Alphacoronavirus*	*Duvinacovirus*	***HCoV 229E***
					Setracovirus	***Human coronavirus NL63*** *NL63-related bat coronavirus strain BtKYNL63-9b*
					Colacovirus	*Bat coronavirus CDPHE15*
					Decacovirus	*Bat coronavirus HKU10* *Rhinolophus ferrumequinum alphacoronavirus HuB-2013*
					Luchacovirus	*Lucheng Rn rat coronavirus*
					Minacovirus	*Mink coronavirus 1*
					Minunacovirus	*Miniopterus bat coronavirus 1* *Miniopterus bat coronavirus HKU8*
					Myotacovirus	*Myotis ricketti alphacoronavirus Sax-2011*
					Nyctacovirus	*Nyctalus velutinus alphacoronavirus SC-2013* *Pipistrellus kuhlii coronavirus 3398*
					Pedacovirus	*Porcine epidemic diarrhea virus*
						Scotophilus bat coronavirus 512
					Rhinacovirus	*Rhinolophus bat coronavirus HKU2*
					Soracovirus	*Sorex araneus coronavirus T14*
					Sunacovirus	*Suncus murinus coronavirus X74*
					Tegacovirus	*Alphacoronavirus 1*
				Betacoronavirus	*Sarbecovirus*	*SARS-CoV;* ***SARS-CoV-2***
					Embecovirus	***HCoVHK1; HCoVOC43*** *Betacoronavirus 1* *China Rattus coronavirus HKU24* *Murine coronavirus* *Myodes coronavirus 2JL14*
					Merbecovirus	***MERS-CoV*** *Hedgehog coronavirus 1* *Pipistrellus bat coronavirus HKU5* *Tylonycteris bat coronavirus HKU4*
					Nobecovirus	*Eidolon bat coronavirus C704* *Rousettus bat coronavirus GCCDC1* *Rousettus bat coronavirus HKU9*
					Hibecovirus	*Bat Hp-betacoronavirus Zhejiang2013*
				Delta coronavirus	*Andecovirus*	*Wigeon coronavirus HKU20*
					Buldecovirus	*Bulbul coronavirus HKU11* *Common moorhen coronavirus HKU21* *Coronavirus HKU15* *Munia coronavirus HKU13* *White-eye coronavirus HKU16*
					Herdecovirus	*Night heron coronavirus HKU19*
				Gamma coronavirus	*Brangacovirus*	*Goose coronavirus CB17*
					Cegacovirus	*Beluga whale coronavirus SW1*
					Igacovirus	*Avian coronavirus* *Avian coronavirus 9203* *Duck coronavirus 2714*

Methods

Analytical review of the literature

An analytical review of the literature was carried out with the help of *PubMed Central* and *GISAID*, to identify relevant foreign articles published to date. Search queries included *coronavirus, severe acute respiratory syndrome* coronavirus 2, *2019-nCoV, SARS-CoV-2, SARS-CoV, MERS-CoV* and *COVID-19*.

Statistical data of *COVID-19*

Data on the epidemiology of the disease, including the total number of confirmed cases and the total number of deaths in different parts of the world and the Republic of Kazakhstan, were obtained from the *WHO COVID-19* situation monitoring panel, available at *https://covid19.who.int*.

Schematic comparison of nomenclatures of the *SARS-COV-2*

For a global schematic comparison of the nomenclature of coronavirus infection, the following main databases were used *https://www.gisaid.org/, https://clades.nextstrain.org/ and https://cov-lineages.org/lineage_list.html* (Table 1).

Metadata of the *SARS-CoV-2* genome

Genome analysis included geographic location information, human gender, human age, and viral genome clades. The global distribution of genomes included in this study was as follows: In the current study, the global distribution of SARS cov 2 genomes according to the GASAID database was obtained from six regions of the world, specifically Asia (n = 234,132), Africa (n = 49,929), Europe (n = 2,093,605), North America (n = 1,203,992) South America (n =72,290), Oceania (n = 32,250) (*https://www.epicov.org/*) (Figure 3).

Sequences of all SARS-CoV-2 genomes by the Republic of Kazakhstan were sent to the GISAID database (*https://www.gisaid.org/CoV2020/*), accessed on September 23, 2021. Only the genomes of viruses isolated from humans and those for which genomic clades were indicated (n = 397) were selected for analysis (Figure4).

For age-based comparisons in this article, the indicator of people's ability to work was used *https://geographyofrussia.com. www.gisaid.org* the number of relevant patients of the Republic of Kazakhstan (397) was given on the website and, depending on age, divided into 3 groups: including children (up to 14 years old), adults (15-64 years old), elderly people (65 years and older) and people of unknown age (Table 4).

Results and discussion

Static data of a new type of coronavirus infection in the World

On March 11, 2020, 114 countries registered more than 118000 cases of infection and about 4000 deaths. Taking into account these results, the number and geographical distribution prompted the *World Health Organization* (*WHO*) to declare the *COVID*-19 outbreak a pandemic caused by *SARS-CoV*-2 [10]. As of September 23, 2021, according to *WHO* data, a total of 229,857,955 confirmed cases of *COVID*-19 and 4,713,530 deaths were registered (Figure 1). The estimated global mortality rate (*CFR*) was 2.05% [11].

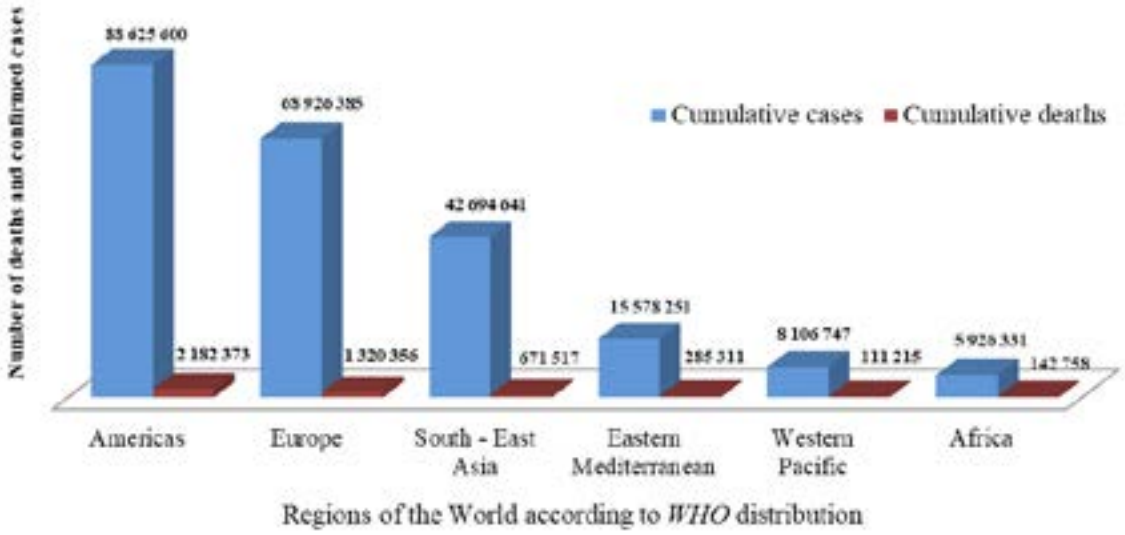

Figure 1 – The number of deaths and confirmed cases of *COVID-19* in the World.

As can be seen in Figure 1, according to *WHO* data, the largest number of *COVID*-19 cases from all regions was registered on the American continents, where 88625600 infected *COVID*-19 were detected. The number of deaths was 2,182,373. This is followed by Europe, where a total of 68,926,385 cases of *COVID*-19 were registered, 1,320,356 deaths, and the estimated mortality rate (*CFR*) was 1.9%. Africa had a high estimated mortality rate (*CFR*) of 2.4% [11].

Schematic comparison of nomenclatures of the *SARS-COV-2*.

Since the onset of coronavirus infection at the end of 2019, simultaneous circulation of various variants of the *SARS-CoV*-2 virus has been established. On May 31 of this year, *WHO* identified the main variants of coronavirus infection, marking them with Greek symbols [10,12]. Currently, there are many different variants of *SARS-CoV*-2 that can be combined into larger groups, including lines or branches [13], but three main nomenclatures for *SARS-CoV*-2 have been proposed, which are *GISAID (https://www.gisaid.org/)*, *Nextstrain (https://clades.nextstrain.org/)* and *Pango lines (https://cov-lineages.org/)*. A schematic comparison of the nomenclatures according to the main database is presented in Table 2.

Table 2 – Schematic comparison of *GISAID*, Nextstrain and *Pango line* nomenclatures

WHO label	Pango lineages	Nextstrain clade	GISAID clade
-	A.1-A.6	19B	S
-	B.3- B.7; B.9 -B.10; B.13-B.16	-	L
-		19A	O
-	B.2	-	V
-	B.1.5-B.1.72: B.1.620; B.1.160; B.1.258; B.1.22.1	20A	G
Eta	B.1.525	21D	
Kappa	B.1.617.1	21B	
Delta	B.1.617.2; AY.1; AY.2	21A	GK
-	B.1.9; B.1.13; B.1.22; B.1.26; B.1.37; B.1.367.	20C	GH
-		20G	
Beta	B.1.351; B.1.351.2; B.1.351.3	20H	
Epsilon	B.1.427; B.1.429	21C	
Iota	B.1.526	21F	
-	B.1.1; B.1.1.519; B.1.1.318; B.1.1.277; B.1.1.302.	20B	GR
-	C.36.3; C.36.3.1	20D	
Mu	B.1.621	21H	
-	B.1.1	20F	
Lambda	C.37	21G	
Gamma	P.1; P.1.1;P.1.2	20J	
Zeta	P.2	21E	
Alpha	B.1.1.7	20I	GRY
-	B.1.177	20E	GV

Note: « – » The official name is not approved by WHO

The data in Table 2 indicate that in September of this year, 10 main warehouses (*S, O, L, V, G, GH, GR, GV, GK* and *GRY*) were presented according to the *GISAID* database. In 2017, Hadfield et al. We have developed a Text string database designed to track the evolution of pathogens in real time. As of September 2021, *Nextstrain* has identified 20 main slides (19A-B, 20A-20J and 21A, 21H). In November 2020, Rambaut et al. proposed *PANGOLIN lineages* that track the transmission and

spread of *SARS-CoV*-2. The development team wrote about this in the article "A dynamic nomenclature proposal for *SARS-CoV*-2 lineage to assist genomic epidemiology" [14].

Over time, all viruses, including *SARS-CoV*-2 change, but most mutations practically do not affect the properties of the virus. Some mutational changes affect the characteristics of the virus, such as the rate of its spread, the severity of the associated disease or the effectiveness of vaccines, therapeutic drugs, increased morbidity, deaths, etc. Variants that relate to one or more of these criteria may be designated as "options of concern" or "designated options of interest". Options of concern include *Alpha (GRY), Beta (GH/501Y.V2), Gamma (GR), Delta (GK). Eta (G/484K.V3), Iata (GH/253 G.V1), Kappa (G/452 R.V3), Lambda (GR/452 Q.V1)* and *Mu (GR)* belong to the designated *SARS-CoV*-2 variants of interest [15,16].

According to the *GISAID* system of nomenclature for the exchange of all influenza data, currently most of the sequenced *SARS-CoV*-2 genomes have been grouped into one of the ten main clades [17]. Such clades include L, to which the reference strain of the *SARS-CoV*-2 virus belongs (Wuhan-Hu-1, *GISIAD* access identifier: EPI_ISL_402124 [3]), *S, O, V, G, GH, GR, GV, GK* and *GRY*. *SARS-CoV*-2 clades differ little from the original reference strain. Clade "*S*" differs from the original strain *S76S* in *NSP4*; *L84S* in *ORF8*. Clade "*V*" will be determined by the corresponding current mutations *L37F* in *NSP6* and *G251V* in *ORF3a*, respectively [18]. Currently, the total mutational variability of *SARS-COV*-2 clades according to *GISAID* data is presented in Table 3.

Table 3 – Current definition of characteristic mutations of SARS-CoV-2 phylogenetic categorization systems (www.gisaid.org)

GISAID clade	*Pango lineages*	Amino acid sequence substitutions
G, GK & GV	A, B & B.2 B.1, B.1.617.2, AY.* & B.1.177	N501S; V445F; K458R; E484Q; Y473H; V445A; G504S; S494L; F456L; G446V; S477N F490L; Y495H; N501T; G476S; T478R; G446A; K417M; N501Y; G446L; V503A; P499R; E484K; E484V; K417R K458N; ; ; Q498R; G504V; E484G K417N; V503F; Y453F; A475S; E484A; T500F; S477G; Q493E; K417E; A475V; G485V; F490S; ; G476D; P499L; F490V; V445I; L455F; G446S; E484D; P499S
GH	B.1.*	Q493R; K417N; N501Y; E484K; N439K; ; G485R; Y453H;
GR	B.1.1.1 & B.1.1.7	K417T; F486I; N501Y; E484K; ; Y449H; ; F490S; Q493K; ; Q498R;

As can be seen from Table 3, the substitution of amino acid sequences *T478K, F490F* and *E484K* was repeated by *G, GK, GV, GH, GR* clades of *SARS-COV*-2. The alternation of amino acid sequence substitutions *S494P* and *Q498R* was repeated in *G, GK, GV, GR* clades, and the alternation of amino acid sequence substitutions *S477I, E484G* and *K417N* was repeated in *G, GK, GV* clades. The alternation of amino acid sequence substitutions *S477I*, *E484G* and *K417N* was repeated in *GH, GR* clades.

Continental distribution of the clades *SARS-CoV-2*.

According to *GISAID*, the most common treasure in the world is *GK* (32,6%), followed by *GRY* (25,3%), *GR* (13,9%), *GH* (12,5%), *G* (8.9%) and *GV* (4,7%). Lower prevalence was noted among *O, L, S* and *V* warehouses, 1,1% 0,1%, 0,18% and 0,18% were identified, respectively. Analysis of the continental distribution of *SARS-CoV*-2 clades (Figure 3) showed that the *GK* clade is often identified among genomes from two continents, namely from Europe (36,02%) and North America (30,02%). The GH warehouse dominated in Africa (32,47%). *GR* clade was most common in Asia (35,1%), South America (70,2%) and Oceania (44,1%) (*https://www.gisaid.org/CoV2020/*).

On March 13, 2020, the first cases of the previously unknown coronavirus *SARS-CoV*-2, which causes atypical pneumonia, were officially recorded in Kazakhstan [19]. On September 23, 2021, according to *WHO* data, a total of 944,733 cases, 15,503 cases of death were registered in Kazakhstan [11]. Clade *GRY* is most common in our country, where its share is 43,5%. Other contributions, including *GR, G, O, S* and *GK*, were relatively 18,6%, 17,6%, 8,06%, 8,06%, 4,7% and 5%, respectively (Figure 4). A lower prevalence was noted among the warehouses of *GH, GV* and *L* and identified 1,5%, 0,25% and 0,5% of the total genomes presented, respectively (*https://www.gisaid.org/CoV2020/*).

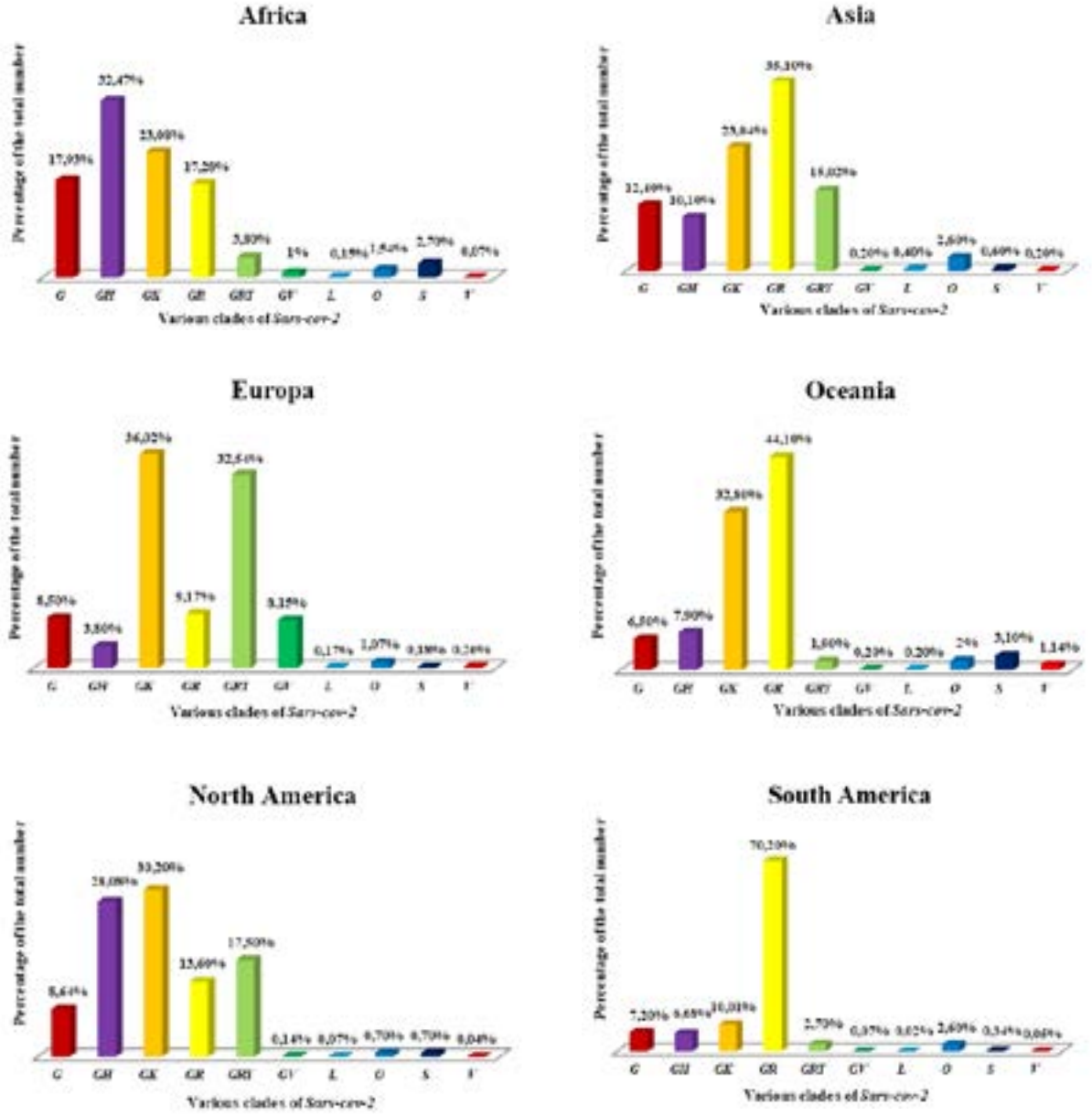

Figure 3 – Continental distribution of various *SARS-CoV*-2 warehouses

As shown in Figure 4, the distribution of *SARS-CoV*-2 clades across the Republic of Kazakhstan showed that the *GRY* clade is most often identified among the genomes represented from seven regions, namely from Almaty region (70,7%), Karaganda region (59%), Kostanay region (53%), Zhambyl region (53%), West Kazakhstan Region (50%), Mangystau region (50%), Pavlodar region (100%). Clade *G* was most common in the region of Aktobe region (60%), Kyzylorda region (45%), Turkestan region (100%). The *GR* clade is the most registered among genomes in Atyrau region (68%) and East Kazakhstan Region (57%). In the Petropavlovsk region, the *GK* clade prevailed (60%). The treasure of *O* and *S* is widespread among the dwarves in the city of Nur-Sultan 38% and 25%, respectively (*https://www.gisaid.org/CoV2020/*).

Distribution of the corresponding genomes according to the *GISAID* database belonging to different clades in different age groups, 397 genomes were analyzed, which were found in the Republic of Kazakhstan (Figure 5).

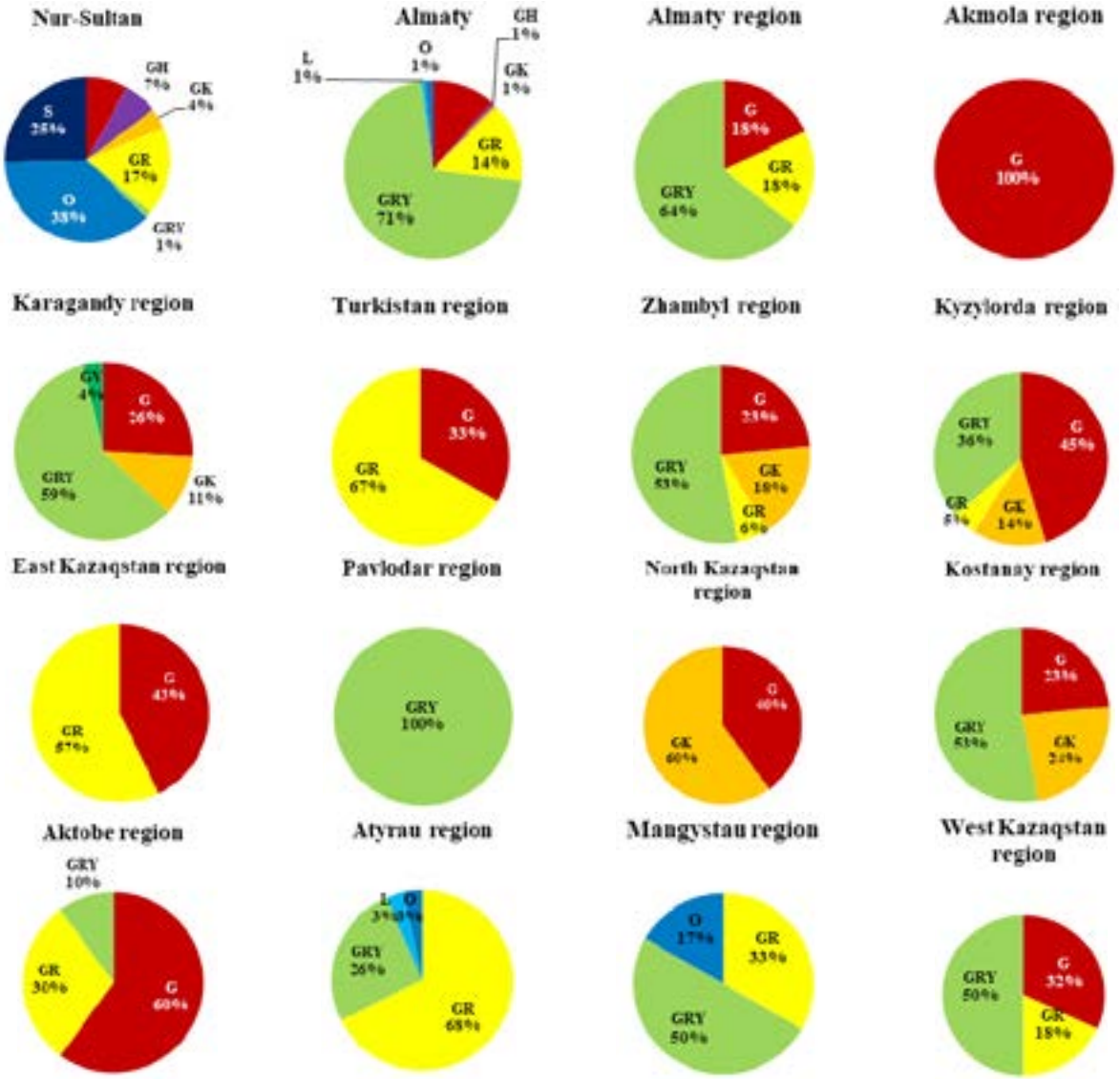

Figure 4 – Analysis of the distribution of various *SARS-CoV*-2 clades across the Republic of Kazakhstan

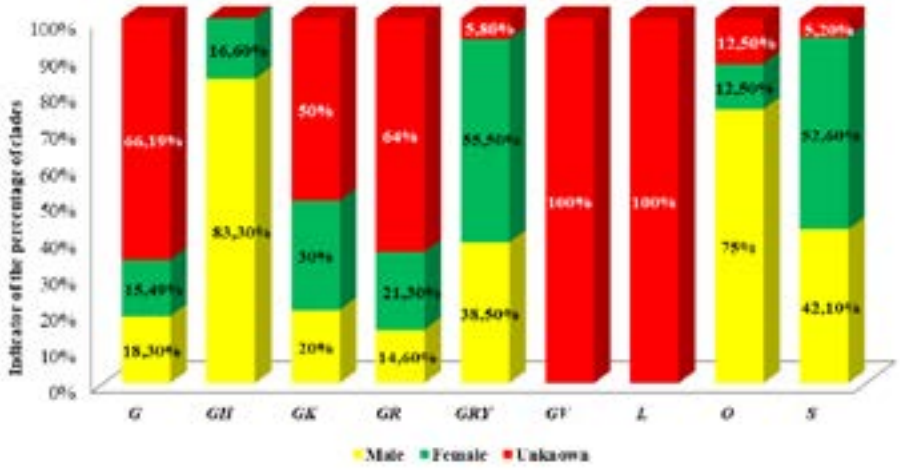

Figure 5 – The spread of the *SARS-CoV*-2 clade to various gender groups

As can be seen from Figure 5, as a result of the analysis of the distribution of *SARS-CoV*-2 warehouses among males of the most common clades *GH* and *O*, their share was 83,3% and 75%, respectively. The *CRY* and S clades are often identified among the female, provided genomes for 55,5% and 52,6 %. The remaining treasures *G* (66,19 %,), *GK* (50%), *GR* (64%), *GV* (100%) and L (100%) correspond to an unknown clade (*https://www.gisaid.org/CoV2020/*).

According to *GISAID* data, up to September 23, 397 *SARS-COV*-2 isolates belonging to different clades in different age groups were registered in the Republic of Kazakhstan, the indicator of which is provided in Table 4 (*https://www.gisaid.org/CoV2020/*).

Table 4 – Age group distribution of the *SARS-CoV*-2 clade

Age groups	*SARS-CoV-2* clades								
	G	*GH*	*GK*	*GR*	*GRY*	*GV*	*L*	*O*	*S*
Children	8,5%	33,3%	10%	18,9%	12,1%	-	-	-	21,05%
Adults	57,1%	66,6%	45%	40,5%	69,9%	-	-	90,6%	68,4%
Elderly	14,2%	-	10%	9,4%	16,7%	100%	-	-	5,2%
Unknown	20%	-	35%	31,08%	1,1%	-	100%	9,3%	5,2%

Note: « – » genomes not found in the *GISAID* database

As can be seen in Table 4, *SARS-CoV*-2 clades registered in *GISAID* were more common in adults. *GV* clade is the most common among the elderly compared to others. As shown in table 4, clade *L* is the most common among unknown age people.

Conclusion

In December 2019, several patients with pneumonia of unknown etiology were identified for the first time in Wuhan, China. With the help of molecular genetic analysis , a new variant of coronavirus was identified from an unknown etiological agent. The rapid spread of a new variant of the *SARS-CoV*-2 coronavirus prompted *WHO* to declare the *COVID*-19 outbreak a pandemic. Currently, according to *WHO* data, a total of 229,858,719 confirmed cases of *COVID*-19 and 4,713,543 deaths have been registered in the world. In the Republic of Kazakhstan, there were a total of 944,733 confirmed cases, 15,503 cases of death [11].

This analysis provides statistical data on the ongoing evolution of *SARS-CoV*-2. According to *GISAID* data, it presents 10 clades, including *G, GH, GK, GR, GR, GV, L, O* and *S*. Currently, the *SARS-CoV*-2 clade is the most common worldwide clade GK (32,6%), followed by *GRY* (25,3%), followed by *GR* (13,9%). The most common genomes in the Republic of Kazakhstan are the *GRY* clade (43,5%), followed by the *GR* clades (18,6%), and *G* (17,6%) (*https://www.gisaid.org/CoV2020/*).

Analysis of the distribution of *SARS-CoV*-2 clades among males in the Republic of Kazakhstan is most common clades *GH* and *O*, and among the female clades *GRY* and *S* were often identified. The spread of various *SARS-CoV*-2 warehouses was more common in adults. Clade *GV* was found only in the elderly. Clade *L* belongs only to people of unknown age.

The work was carried out within the framework of the grant funding project AP09058338 "Study of antiviral activity of drugs in relation to the SARS-COV-2 virus in vitro and molecular epidemiological analysis of circulating COVID-19 strains".

Reference

1 Benjamin W.N., Brian D.A., Craig Y., Joel D.Q., Gretchen O., Peter K., Ronald A.M., Mark Y., Michael J.B. «Supramolecular Architecture of Severe Acute Respiratory Syndrome Coronavirus Revealed by Electron Cryomicroscopy». Journal of Virology no 80 (2006): 7918–7928.

2 Virus Taxonomy: 2021 Release. https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-viruses2011/w/posrna_ viruses /222/

3 Fehr A.R., Perlman S. «Coronaviruses: an overview of their replication and pathogenesis». Methods Mol Biol no 1282 (2015): 1–23.

4 van der Hoek L., Pyrc K., Jebbink M.F., Vermeulen-Oost W., Berkhout R.J.M, Wolthers K.C., Wertheim-van Dillen P.M.E., Kaandorp J., Spaargaren J., Berkhout B. «Identification of a new human coronavirus». Nat Med no 10 (2004): 368–373..

5 Woo P.C.Y., Lau S.K.P., Chu C.-m., Chan K.-h., Tsoi H.-w., Huang Y., Wong B.H.L., Poon R.W.S., Cai J.J., Luk W.-k., Poon L.L.M., Wong S.S.Y., Guan Y., Peiris J.S.M., Yuen K.-Y. «Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia». // J Virol. no 79 (2005): 884–895

6 Paules C.I., Marston H.D., Fauci A.S. «Coronavirus infections—more than just the common cold». JAMA no 323 (2020): 707–708.

7 Shereen M.A., Khan S., Kazmi A., Bashir N., Siddique R. «COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses». Journal of Advanced Research 24 (2020): 91–98.

8 Wu F., Zhao S., Yu B., Chen Y.-M., Wang W., Song Z.-G., Hu Y., Tao Z.-W., Tian J.-H., Pei Y.-Y., Yuan M.-L., Zhang Y.-L., Dai F.-H., Liu Y., Wang Q.-M., Zheng J.-J., Xu L., Holmes E.C., Yong-Zhen. «A new coronavirus associated with human respiratory disease in China». Nature 579 (2020): 265–269.

9 Zhu N., Zhang D., Wang W., Li X., Yang B., Song J., Zhao X., Huang B., Shi W., Lu R., Niu P., Zhan F., Ma X., Wang D., Xu W., Wu G., Gao G.F., Tan W. «A novel coronavirus from patients with pneumonia in China, 2019». N Engl J Med no 382 (2020): 727-733.

10 WHO Director-General's opening remarks at the media briefing on COVID-19 – 11 March 2020 https://www.who.int/

11 WHO Coronavirus Control Panel (COVID-19) https://www.who.int/

12 WHO announces simple, easy-to-say labels for SARS-CoV-2 Variants of Interest and Concern https://www.who.int/news/item/31-05-2021-who-announces-simple-easy-to-say-labels-for-sars-cov-2-variants-of-interest-and-concern

13 Koyama T., Platta D., Paridaa L. «Variant analysis of SARS-CoV-2 genomes». Bull World Health Organ no 98 (2020): 495–504.

14 Rambaut A., Holmes E.C., O'Toole A., Hill V., McCrone J.T., Ruis C., du Plessis L., Pybus O.G. «A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology». Nat Microbiol 5 (2020): 1403–1407.

15 Tracking SARS-CoV-2 variants https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/

16 SARS-CoV-2 Variant Classifications and Definitions https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-classifications.html

17 Hamed S.M., Elkhatib W.F., Khairalla A.S., Noreddin A.M. «Global dynamics of SARS CoV-2 clades and their relation to COVID 19 epidemiology». Scientific Reports no 11 (2021): 1-8.

18 Mercatelli D., Giorgi F.M. «Geographic and genomic distribution of SARS-CoV-2 mutations». Front Microbiol no 11 (2020): 1-13.

19 Batyrshin T. V. «Kazahstane zafiksirovany pervye dva sluchaja zabolevanija koronavirusom [The first two cases of coronavirus have been recorded in Kazakhstan].» Sputnik Kazahstan. https://ru.sputnik.kz/20200313/kazakhstan-pervyy-sluchay-koronavirus-13067637.html. (Data obrashhenija 09.20.2021). – (In Russian)

Reference

1 Benjamin W.N., Brian D.A., Craig Y., Joel D.Q., Gretchen O., Peter K., Ronald A.M., Mark Y., Michael J.B. Supramolecular Architecture of Severe Acute Respiratory Syndrome Coronavirus Revealed by Electron Cryomicroscopy // Journal of Virology. – 2006. – Vol. 80, №16. – P. 7918–7928.

2 Virus Taxonomy: 2021 Release. https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-viruses2011/w/posrna_ viruses /222/

3 Fehr A.R., Perlman S. Coronaviruses: an overview of their replication and pathogenesis // *Methods Mol Biol*. – Vol 1282, №2015. – P. 1–23.

4 van der Hoek L., Pyrc K., Jebbink M.F., Vermeulen-Oost W., Berkhout R.J.M, Wolthers K.C., Wertheim-van Dillen P.M.E., Kaandorp J., Spaargaren J., Berkhout B. Identification of a new human coronavirus. // Nat Med. – 2004. – Vol. 10, № 4. – P. 368–373.

5 Woo P.C.Y., Lau S.K.P., Chu C.-m., Chan K.-h., Tsoi H.-w., Huang Y., Wong B.H.L., Poon R.W.S., Cai J.J., Luk W.-k., Poon L.L.M., Wong S.S.Y., Guan Y., Peiris J.S.M., Yuen K.-Y. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia // J Virol. – 2005. – Vol. 79, № 2. – P. 884–895.

6 Paules C.I., Marston H.D., Fauci A.S. Coronavirus infections—more than just the common cold // JAMA. – 2020. – Vol. 323, №8. – P. 707–708.

7 Shereen M.A., Khan S., Kazmi A., Bashir N., Siddique R. COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses // Journal of Advanced Research. – 2020. – Vol. 24. – P. 91–98.

8 Wu F., Zhao S., Yu B., Chen Y.-M., Wang W., Song Z.-G., Hu Y., Tao Z.-W., Tian J.-H., Pei Y.-Y., Yuan M.-L., Zhang Y.-L., Dai F.-H., Liu Y., Wang Q.-M., Zheng J.-J., Xu L., Holmes E.C., Yong-Zhen. A new coronavirus associated with human respiratory disease in China. // Nature. – 2020. – Vol. 579. – P. 265–269.

9 Zhu N., Zhang D., Wang W., Li X., Yang B., Song J., Zhao X., Huang B., Shi W., Lu R., Niu P., Zhan F., Ma X., Wang D., Xu W., Wu G., Gao G.F., Tan W. A novel coronavirus from patients with pneumonia in China, 2019. // N Engl J Med. – 2020. – Vol. 382, № 8. – P. 727-733.

10 WHO Director-General's opening remarks at the media briefing on COVID-19 – 11 March 2020 https://www.who.int/

11 WHO Coronavirus Control Panel (COVID-19) https://www.who.int/

12 WHO announces simple, easy-to-say labels for SARS-CoV-2 Variants of Interest and Concern https://www.who.int/news/item/31-05-2021-who-announces-simple-easy-to-say-labels-for-sars-cov-2-variants-of-interest-and-concern

13 Koyama T., Platta D., Paridaa L. Variant analysis of SARS-CoV-2 genomes // Bull World Health Organ. – 2020. – Vol. 98, № 7. – P. 495–504.

14 Rambaut A., Holmes E.C., O'Toole A., Hill V., McCrone J.T., Ruis C., du Plessis L., Pybus O.G. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology // Nat Microbiol. – 2020. Vol 5. – P. 1403–1407.

15 Tracking SARS-CoV-2 variants https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/

16 SARS-CoV-2 Variant Classifications and Definitions https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-classifications.html

17 Hamed S.M., Elkhatib W.F., Khairalla A.S., Noreddin A.M. Global dynamics of SARS CoV-2 clades and their relation to COVID 19 epidemiology // Scientific Reports. – 2021. – Vol. 11, №8435. – P. 1-8.

18 Mercatelli D., Giorgi F.M. Geographic and genomic distribution of SARS-CoV-2 mutations // Front Microbiol. – 2020. – Vol. 11, №1800. – P. 1-13.

19 Батыршин Т. В Казахстане зафиксированы первые два случая заболевания коронавирусом. Sputnik Казахстан. https://ru.sputnik.kz/20200313/kazakhstan-pervyy-sluchay-koronavirus-13067637.html. (Дата обращения 09.20.2021)