Genetics & Biodiversity Journal

Journal homepage: https://journals.univ tlemcen.dz/GABJ/index.php/GABJ

ISSN: 2588-185X

OPEN ACCESS

Original Research Paper

Studying mutations of SARS-Cov-2 different variants (Alpha, Beta, Delta, Gamma, Omicron)

DEROUICHE Louiza[1*], LATRACH Rabia[2], SEDDAOUI Rania[2]

[1]Higher School of Food Sciences & Agri-food Industries Algiers, Algeria;
[2]Saad Dahleb Blida 1 University, Faculty of Natural and Life Sciences, Algeria.

***Corresponding Author**: Derouiche Louiza,Higher School of Food Sciences & Agri-food Industries Algiers, Algeria; **E-mail:**derouiche_fatma@yahoo.fr

Article history: Received: September 23rd 2022; Revised: November 26th 2022; Accepted: December 15th 2022

Abstract

Being pushed by natural selection, random genetic drift, gene editions, and receptor immunity response, viruses develop constantly through mutations affecting different genes and leading to genetic diversity and producing new variants. In order to know well how a mutation could have an impact on the possibility of being infected, on transmission, and on aggressivity of SARS-CoV-2 it would be important to study these mutations. To be able to carry out a comparative study between variants and undergone mutations over many countries in the world, we've dealt with many genomic sequences that have been rapidly accumulated in the GenBank since January 2020, and published by many laboratories over the world. These sequences allowed us to establish phylogenetical trees using a strong bioinformatic tool, just enhanced to study Covid which is MEGA version 11. Distribution of shifted sequences of different variants over the world within phylogenetical trees shows that the overwhelming majority of detected mutations are accumulated in the 5 known variants Alpha (B.1.1.7), Beta (B.1.351), Delta (B.1.617.2), Gamma (P.1) et Omicron (B.1.1.529), especially within their most variable genes, structural genes of which are N (Nucleocapsid protein) and S (Spike glycoprotein) added to functional ones ORF (Open Reading Frame : ORF1ab, ORF3a) ; hence, variants holding these mutations are the most dominant and the most infectious this time in the world.

Keywords: SARS-CoV-2; genes; mutations; variants; phylogenetic.

الملخص

يتطور الفيروس باستمرار من خلال الطفرات التي تؤثر على الجينات المختلفة ويؤدي إلى التنوع الجيني وإنتاج متغيرات جديدة، وذلك بسبب الانتقاء الطبيعي والانحراف الجيني العشوائي وإصدار الجينات واستجابة المناعة للمستقبلات. من أجل معرفة كيف يمكن أن يكون للطفرة تأثير على إمكانية الإصابة بالعدوى، وعلى انتقال عدوانية فيروس سارس كوف 2، سيكون من المهم دراسة هذه الطفرات. من أجل أن نكون قادرين على إجراء دراسة مقارنة بين المتغيرات والطفرات في العديد من البلدان في العالم، تعاملنا مع العديد من التسلسلات الجينية التي تراكمت بسرعة في بنك الجينات منذ يناير 2020، ونشرتها العديد من المختبرات حول العالم. سمحت لنا هذه التسلسلات بتأسيس أشجار النشوء والتطور باستخدام أداة قوية للمعلومات الحيوية، تم تحسينها للتو لدراسة الكوفيد وهو الإصدار 11 من برنامج ميجا. يوضح توزيع التسلسلات المتغيرة من المتغيرات المختلفة حول العالم داخل الأشجار النشوء والتطور أن الغالبية العظمى من الطفرات المكتشفة تتراكم في 5 متغيرات معروفة ألفا (ب.1.1.7) وبيتا (ب.1.351) ودلتا (ب.1.617.2) وجاما (پ.1) وأوميكرون (ب.1.1.529)، خاصةً في جيناتها الأكثر تغيرًا، الجينات الهيكلية منها تضاف إلى الجينات الوظيفية.وبالتالي، فإن المتغيرات التي تحمل هذه الطفرات هي الأكثر انتشارًا والأكثر عدوى في العالم

الكلمات المفتاحية: سارس كوف 2، الجينات، الطفرات، المتغيرات، النشوء والتطور

Introduction

Covid-19 is a disease caused by SRAS-CoV-2 and might be a major threat to public health all over the world because of its great deal of infections and death. As the rate of morbidity and mortality continues increasing because the main drugs that may prevent the expansion of Covid-19 are not available, despite the availability of some vaccines during the first half of 2020, particularly from March to May, the number of people infected with Covid-19 increased fiercely, while states and countries afford great efforts to deal with the new pandemic virus (Wu and *al.*, 2020).

The first case of SARS-CoV-2 has been discovered on December 30th, 2019 in the Wuhan district of China and was declared later as a pandemic (Hang and *al.*, 2020). The main symptoms are coughing, high fever, temperature, and breathing difficulties, these symptoms may go forward until suffocation and organic failure. Accordingly, the disease may cause a systemic inflammation transmitted by infection and sneezing droplets, but other transmission modes such as airborne and fecal-oral transmission are also risen (Wang and *al.*, 2020).

SARS-CoV-2 is a virus covered by a genome with RNA simple positive polarity of about 29,8 to 29,9 kb with 10 open reading frames (ORF) coding about 30 proteins. The ORF1ab code of 16 non-structural proteins among which RdRp, the other ORF codes merely 4 structural proteins; glycoprotein S, which is the epitope of neutralizing antibodies and ACE receptor attachment site, the protein N (nucleoprotein) covering protein E and the matrix M. Coding genes for RdRp and structural proteins are the wanted targets of PCR test. Being a virus with RNA mutations is natural and may produce variants of interest such as Alpha, Bêta, Gamma, Delta, and Omicron (Xu and *al.*, 2020; Choi and *al.*, 2021).

The appearance of SRAS-CoV-2 new variants with quick infection transmission may lead to longer a pandemic as it may have damaging consequences on public health and the economy (Chen and *al.*, 2021). The target of this research work is to study variants and mutations detected within their genomic sequences, for this reason, we've referred to bioinformatics in order to be able to get phylogenetical trees.

Material and methods

The goal of this research relies merely on the use of the bioinformatic software MEGA version 11 for analyzing DNA sequences extracted from molecular biology's data, which is GenBank, for the sake of making phylogenetical trees.

Getting data from GenBank

All DNA sequences studied in our research have been downloaded from GenBank. GenBank is a database made and distributed by National Center for Biotechnology Information (NCBI). It is a collection of all nucleotides and protein sequences and bibliographical and biological annotations available for people (Sayers and *al*., 2020).

These registered results are born from direct communications of the original authors' database with DNA sequence, those who give their files in order to make them available for everybody or to submit them to the publishing process (Clark and *al.*, 2016).

DNA sequences processing from GenBank by MEGA11 software

For the sake of DNA sequence processing, there are many software able to generate phylogenetical trees among which MEGA11 software. Molecular Evolutionary Genetics Analysis (MEGA) is a software that allows the exploration and analysis of nuclides' and proteins' sequences (Tamura, 2011).

We've used the MEGA 11 version to get a phylogenetical tree after many processing and analyzing steps that could be summed up in Introducing DNA sequences under FASTA form; these sequences' alignment by Clustal W; selection of the method adopted to get phylogenetical tree and parameters order of the selected method, in our case we've used UPGMA method (Unweighted Pair Group Method with Arithmetic mean). This simple method allows distance matrix transformation (between different organisms, populations or nucleotides' sequences) within a deep-rooted tree based on similarities between pairs of sequences with a Bootstrap of 1000 (Kumar and *al.*, 2018; Thompson and *al.*, 1994).

Results and Discussion

Data extracted from GenBank

5353266SARS-COV-2DNAsequences published on GenBank have been processed and classified into four groups representing the most variable four genes of Coronavirus (Derouiche and *al*., 2021), studied over different countries around the world, which are either structural genes N (Nucleocapsid protein) and S (Spike glycoprotein), or rather functional ones ORF (Open Reading Frame: ORF1ab (ORF1a and ORF1b), ORF3a).

Table 01. Determination of Coronavirus variants and genes' mutations starting from DNA sequences extracted from GenBank.

Genes	Mutations	Country	Variants
S	A570D, T716I	Kenitra, Morocco	Alpha B.1.1.7
	P681R	Dhaka, Bangladesh	
	D1118H, L730F, T1001	Lima, Peru	
	Délé H 69 /V70	Philippines	
	S982A	Manila, philippines	
	E484K, D614G	Mexico	
	R408I	Delhi, India	
	D178H, L452R	USA	
	N501Y, Délé H 69 /V70	Berlin, Germany	
	P681H	New York	
	D614G	Germany	
	E484K, D614G	Switzerland	
	P681H, K417T	Geneva, Switzerland	
	K417N, E484K, N501Y	South Africa	Beta B.1.351
	E484K	Germany	
	D614G	Malaysia	
	P681R	Mohakhali, Bangladesh	Delta B.1.617.2
	E156G, Délé F157	Chittagong, Bangladesh	
	T19R, G142D, E156GT478K, D614G, D950N	India	
	L452R	Delhi, India	
	D950N, T95I	Egypt	
	E484K	Denmark	
	T19R, G142D, E156G	Kolkata, Bangladesh	
	T19R, G142D, R158G, L452R, T478K, K417T, D614G, P681R, D950N	Morocco	
	Délé H 69 /V70	Kerala, India	
	D614G	Venezuela	Gamma P.1
	E484K	Uruguay	
	N501Y	Germany	
	E661D	Brazil	
	G339D, N440K, S477N, T478K	South Africa	Omicron B.1.1.529
	Délé H 69 /V70, N501Y	Bangladeshi	
	Q493R, G496S, Q498R, N501Y	Australia	
	S371P, S373P, S477A	Rabat, Morocco	
	G142D, N211I, G339D, S371L, S373P, S375F, G446S	China	
	N764K, D496Y, N856K, H655Y	Belgium	
ORF1ab	P472L	Malaysia	Alpha B.1.1.7
	P4804P, L5784A	Kenitra, Morocco	
	T1001I	Dhaka, Bangladesh	
	T1001I, A4489V, P4619L	Hyderabad, India	
	I2230T, T1001I, P314L	New York	
	N764K	Lebanon	
	S1433P, P472L, D6909G	Peru	
	P4619 L, P4715L	Bangladesh	
	T1001I, I2230T, P314L	Mexico	
	T58I, L37F, P314L	Israel	
	A4489V	Bangladesh 2	
	L2780F, R1383K	Thailand	
	I2230T, M429I	India	
	M902I	USA	
	A1708D, P2908I, M902I, T1001I	Delhi, India	
	L2780F	Lima, Peru	
	T1001I, I2230T, A1708D	Colombia	Beta B.1.351
	P309L, P314L, P4619L	China	
	T1001I, P2046L, P314L, I2230T, P4619L	Egypt	Delta B.1.617.2
	I2230T	Casablanca, Morocco	

	S370L, K977 Q, L71F, P323L	Uruguay	Gamma P.1
	T1001I	Venezuela	
	P314L, R226K	Brazil	
	T3255I, T1001I, P314L, P4619L	Belgium	Omicron B.1.1.529
	Y655T, L856A, P924P, I2230T	Rabat, Morocco	
N	S235F	Kenitra, Morocco	Alpha B.1.1.7
	S235P	Lima, Peru	
	P314L	Lebanon	
	S235F, R203K, G204R	New York	
	G204R	Mexico	
	Y73C	Dhaka, Bangladesh	
	G204R	Chittagong, Bangladesh	
	L139F	Thailand	
	P344S	Hyderabad, India	
	P13L, D3L	USA	
	L37F, N1410T, A116V	United Kingdom	
	P344S	Peru	
	D63G, P344S	Egypt	Delta B.1.617.2
	A211V, T141I, G204R	Mohakhali, Bangladesh	
	R203k	Bangladesh	
	D377Y	Morocco	
	P80R	Singapore	
	G215C	Mexico	
	D63G	Thailand	
	R203K, K204R	Venezuela	Gamma P.1
	P80R, P344S	Brazil	
	R203K, G204R	Morelos, Mexico	
	D63G, R203K, G215C, D377Y	Australia	Omicron B.1.1.529
	R203K, G204R	Rebat, Morocco	
	P240S	Belgium	
ORF3a	W128L, L140V, G251V	Delhi, India	Alpha B.1.1.7
	S26L	Maharashtra, India	
	S171L, G251V, S26L	Philippines	
	L275F	Mexico	
	I82T	New York	
	G251V	United Kingdom	
	Q57H, W131C, W193L, T223I	Israel	
	Q57H	India	
	F15L	Thailand	
	Q57H	India	Beta B.1.351
	S171L, Q57H, S26L, G251V	USA	
	Q57H	USA	Delta B.1.617.2
	Q57H, S26L	Egypt	
	Q57H, G251V	Denmark	
	S26L	Mexico	
	S26L	Morocco	
	S253P, Q57H	Uruguay	Gamma P.1
	S253P, S26L, G251V	Brazil	
	T64T, Q57H, G251V	Rabat, Morocco	Omicron B.1.1.529
	S92L, D155Y, S26L	France	

These genes have passed by many mutations, which lead to the appearance of 5 variants: the English variant (Alpha B.1.1.7) appeared for the first time in the United Kingdom (September 2020) ; the South African one (Beta B.1.351) appeared for the first time in South Africa (May 2020) ; the Indian one (Delta B.1.617.2) appeared for the first time in India (October 2020) ; the Brazilian one (Gamma P.1) appeared for the first time in Brazil (November 2020) ; and the Omicron variant (B.1.1.529) appeared for the first time on November 2021 and affected many countries around the world (WHO, 2022).

Obtained results from DNA sequences' processing in different countries representing genes, mutations, and variants are given classified in table 01.

Phylogenetical analysis

MEGA 11 software allowed us to find out 4 phylogenetical trees for 4 different genes: N, S, ORF1ab and ORF3a. Every gene is represented by many nucleotide sequences according to the selected country.

-Phylogenetical tree of the gene S

The phylogenetical tree (Fig. 1) has been realized based on the gene S of different variants of Coronavirus that affected many countries around the world. Because of the quick appearance of many variants, WHO has established a system of classification to distinguish newborn variants of SARS-CoV-2, many variants of SARS-CoV-2, having the same mutations, appeared in many countries such as Alpha, Béta, Gamma, Delta, and Omicron ; researches showed that variant is to expand rapidly because their transmissibility is too high and they are more infectious, but they do not seem to cause dangerous illness or a high rate of mortality or even symptoms different from those already observed since the beginning of the pandemic (WHO, 2022).

Different groups come up within the gene S tree phylum are:

Alpha (B.1.1.7) : The Alpha variant group contains 13 sequences that have almost the same size and contains different mutations in several countries, including the sequences of Kenitra (Morocco) with a size of 21529 up to 25341 bp (A570D, T716I), Dhaka (Bangladesh) from 21500 to 25325 bp (P681R), Lima (Peru) from 21384 to 25196 bp (D1118H, L730F, T1001), Philippines from 21320 to 25486 (Délé H69/V70), Manila (Philippines) from 21524 to 25336 bp (S982A), Mexico from 21516 to 25328 bp (E484K, D614G), Delhi (India) from 21550 to 25368bp (R408I), USA from 21563 to 25384bp (D178H, L452R), Berlin (Germany) from 21510 to 25330bp (N501Y, Dele H69/V70), New York from 21320 to 25360bp (P681H), Germany from 21559 to 25392bp (D614G), Switzerland from 21557 to 25378bp (E484K, D614G), Geneva (Switzerland) from 21503 to 25315bp (P681H, K417T).

Beta (B.1.351): The Beta variant group contains 3 sequences with different sizes and mutations in the following countries: South Africa ranges between 21504 and 25316bp (K417N, E484K, N501Y), Germany from 21524 to 25336bp (E484K), Malaysia from 21509 to 25330bp (D614G).

Delta (B.1.617.2): The sequences of the Delta variant represent the mutations and the sizes according to the following countries: Mohakhali (Bangladesh) varies between 21561 and 25376 bp (P681R), Chittagong (Bangladesh) from 21509 to 25324bp (E156G, Dele F157), India from 21515 to 25324 bp (T19R, G142D, E156G,T478K, D614G, D950N), Delhi (India) from 21223 to 25250 (L452R), Denmark from 21538 to 25353bp (E484K), Kolkata (Bangladesh) from 21525 to 25355bp (T19R, G142D, E156G), Morocco from 21443 to 25258 bp (T19R, G142D, R158G, L452R, T478K, K417T, D614G, P681R, D950N), Kerala (India) from 21529 to 25344 bp (Délé H 69/V70), Egypt from 21534 to 25355 bp (D950N, T95I).

Gamma (P.1): Gamma variant sequences include sizes and mutations in the following countries: Venezuela ranges from 21516 to 25337bp (D614G), Uruguay from 21561 to 25382bp (E484K), Germany from 21552 to 25364bp (N501Y), Brazil from 21543 to 25322bp (E661D).

Omicron (B.1.1.529): The Omicron variant group includes sequences from different countries with distinct sizes and mutations which are: South African varies between 21446 and 25258bp (G339D, N440K, S477N, T478K), Bangladesh from 21500 to 25312bp (Délé H 69/V70, N501Y), Australia from 21383 to 25816bp (Q493R, G496S, Q498R, N501Y), Morocco from 21526 to 25338bp (S371P, S373P, S477A), China from 21497 to 25309bp (G142D, N211I, G339D, S371L, S373P, S375F, G446S), Belgium from 21497 to 25309bp (N764K, D496Y, N856K, H655Y).

According to the gene S mutations' number shown in table 1, we can observe that the S gene is the most variable because the mutation number within this gene is too high compared with the mutations' number of N, ORF1ab et ORF3a genes. These mutations are distributed on its whole genetical code, among these mutations are those that are situated within the RBD field such as K417N, L452R, E484K and N501Y; they have more impact on the strength of the virus, on its transmissibility and even on its ability to escape from the immune system, thus, variants englobing these mutations are the most dominant and the most infectious in the world (Nelson and *al*., 2021).

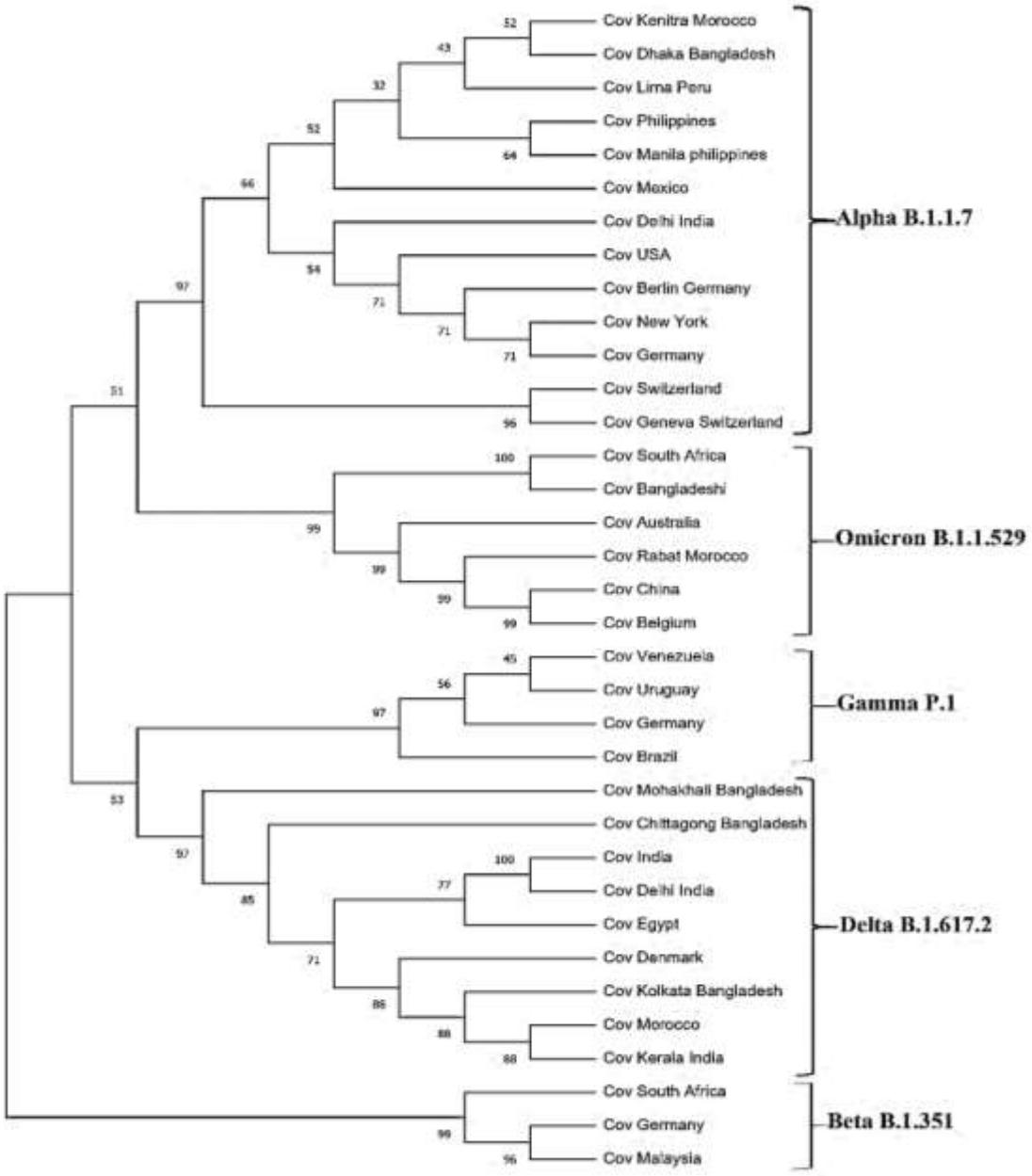

Figure 1. Phylogenetical tree of the gene S representing different variants of Coronavirus.

-Phylogenetical tree of the ORF1ab gene

The phylogenetical tree (Fig. 2) has been realized based on ORF1ab gene sequences of different Coronavirus variants that have affected many countries in the world. Exactly like the tree of the gene S, the ORF1ab gene tree is constituted of 5 variants Alpha (B.1.1.7), Béta (B.1.351), Omicron (B.1.1.529), Delta (B.1.617.2) and Gamma (P.1).

Sequences of different variants come up within the tree of ORF1ab gene and their mutations are:

Alpha (B.1.1.7): Malaysia varies between 225 and 13530bp (P472L), Kenitra (Morocco) from 241 to 13449bp (P4804P, L5784A), Dhaka (Bangladesh) from 256 to 13464bp (T1001I), Hyderabad (India) from 264 to 13481bp (T1001I, A4489V, P4619L), New York from 215 to 13423bp (I2230T, T1001I,

P314L), Lebanon from 220 to 13256bp (N764K), Peru from 212 to 13501bp (S1433P, P472L, D6909G), Bangladesh from 215 to 13423bp (P4619L, P4715L), Mexico from 228 to 21508 bp (T1001I, I2230T, P314L), Israel from 264 to 13481 bp (T58I, L37F, P314L), Bangladesh 2 from 250 to 13265 bp (A4489V), Thailand from 234 to 13523 bp (L2780F, R1383K), India from 244 to 13421bp (I2230T, M429I), USA from 230 to 13423bp (M902I), Delhi (India) from 253 to 21542bp (A1708D, P2908I, M902I, T1001I), Lima (Peru) from 225 to 13420bp (L2780F).

Beta (B.1.351): The Colombian sequence varies between 251 and 21540bp (T1001I, A1708D, I2230T), China from 212 to 13417bp (P309L, P314L, P4619L).

Delta (B.1.617.2): The Egyptian sequence varies between 234 and 13454bp (T1001I, P2046L, P314L, I2230T, P4619L), Casablanca (Morocco) from 241 to 21530bp (I2230T).

Gamma (P.1): The sequence of Uruguay varies between 264 and 21553bp (S370L, K977Q, L71F, P323L), Venezuela from 228 to 21508bp (T1001I), Brazil from 272 to 21546bp (P314L, R226K).

Omicron (B.1.1.529): The sequence of Belgium varies between 212 and 13417bp (T3255I, T1001I, P314L, P4619L), Rabat (Morocco) from 241 to 13446bp (Y655T, L856A, P924P, I2230T).

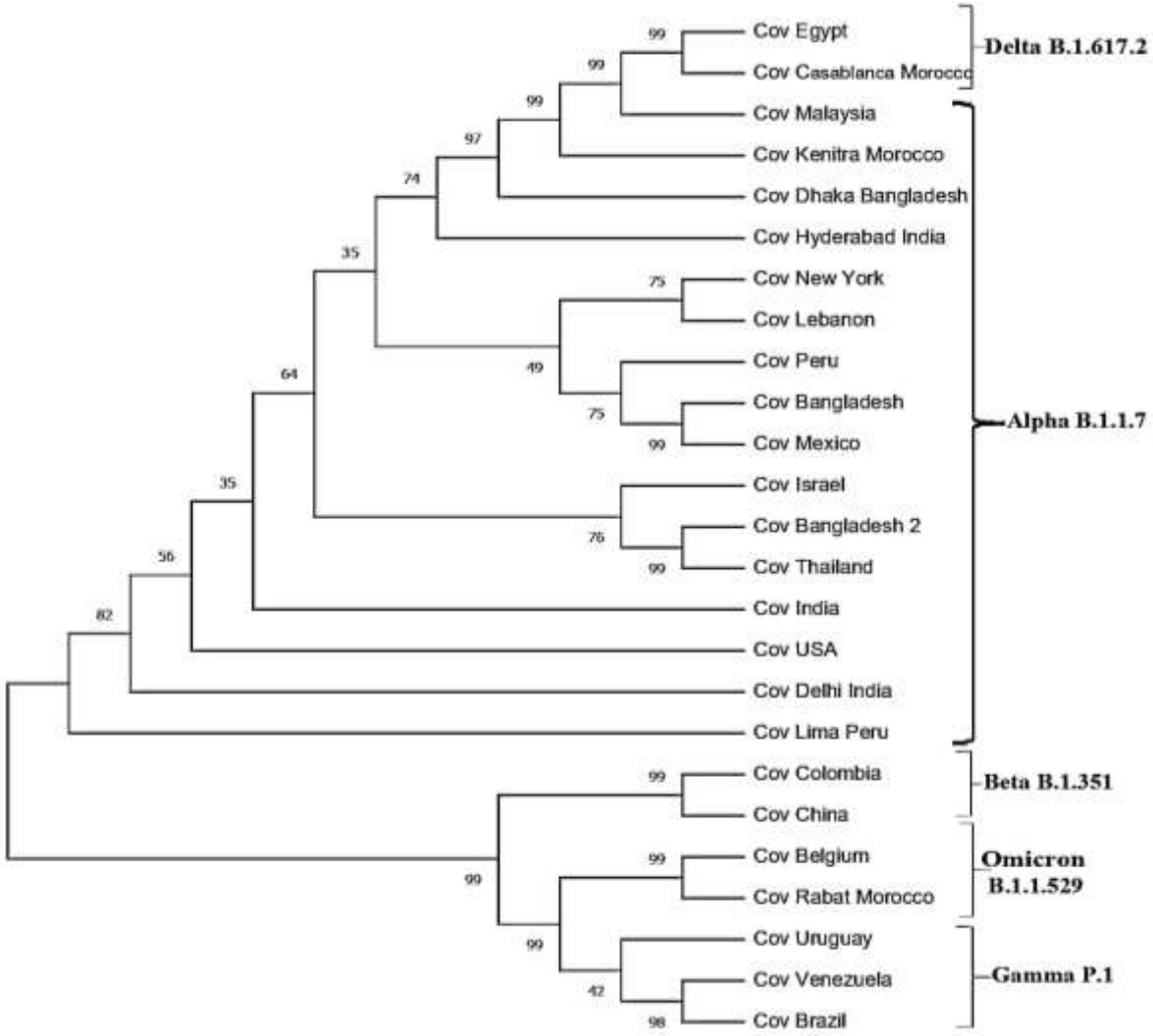

Figure 2. ORF1ab phylogenetical tree representing sequences of different Coronavirus variants.

Many mutations have been detected for ORF1ab gene among them those situated within ORF1ab field: I2230T, P314L, T1001I and P4619L; according to Zárate and *al.* (2022) these mutations are the most harmful, they have an impact on strength of the virus (transmissibility, escape, resistance to vaccine and drugs).

Phylogenetical tree of the N gene

The phylogenetical tree (Fig. 3) has been realized based on nucleocapsid gene N of different variants. Because of the lack of data on GenBank concerning the Beta variant, the gene N tree constitutess only four variants Alpha, Delta, Gamma, and Omicron.

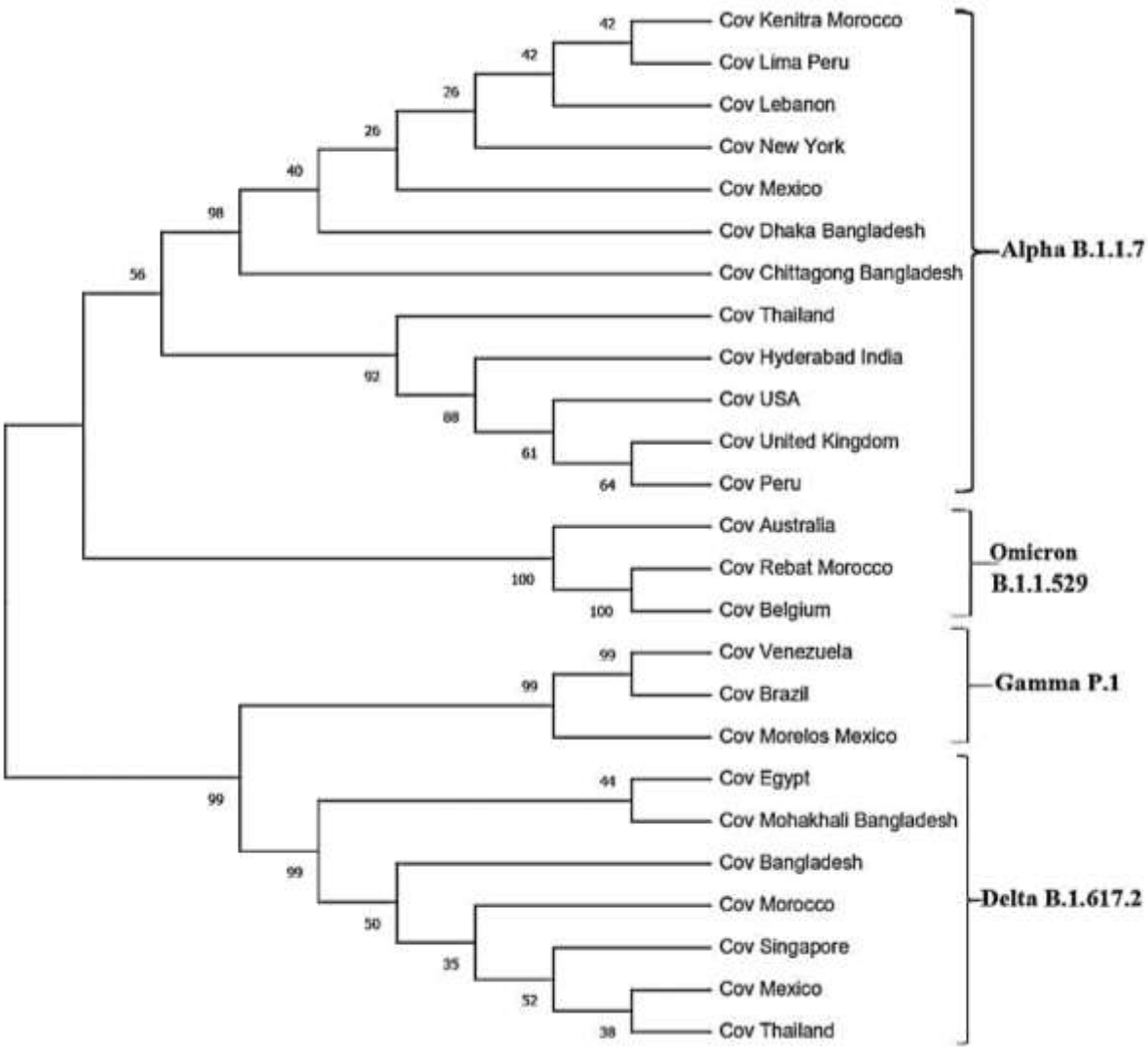

Figure 3. Phylogenetical tree of nucleocapsid N representing sequences of different Coronavirus variants.

Sequences of different variants come up within the tree of N nucleocapsid gene with their measurements and their mutations are:

Alpha (B.1.1.7) : The sequence of Kenitra (Morocco) varies between 28230 and 29456bp (S235F), Lima (Peru) from 28220 to 29479bp (S235P), Lebanon from 28440 to 29203bp (P314L), New York from 28205 to 29464bp (S235F, R203K, G204R), Mexico from 28217 to 29476bp (G204R), Dhaka (Bangladesh) from 28245 to 29504bp (Y73C), Chittagong (Bangladesh) from 28204 to 29463bp (G204R), Thailand from 28235 to 29494bp (L139F), Hyderabad (India) from 28258 to 29517bp (P344S), USA from 28220 to 29479bp (P13L, D3L), United Kingdom from 28274 to 29533bp (L37F, N1410T, A116V), Peru from 28085 to 29344bp (P344S).

Delta (B.1.617.2): The Egyptian sequence varies between 28245 and 29504bp (D63G, P344S), Mohakhali (Bangladesh) from 28268 to 29527bp (A211V, T141I, G204R), Bangladesh from 28207 to 29466bp (R203K), Morocco from 28249 to 29508bp (D377Y), Singapore from 28220 to 29321bp (P80R), Mexico from 28223 to 29482bp (G215C), Thailand from 28245 to 29504bp (D63G).

Gamma (P.1): The Venezuela sequence varies between 228 and 21508bp (R203K, K204R), Brazil from 272 to 21546bp (P80R, P344S), Morelos (Mexico) from 264 to 21553bp (R203K, G204R).

Omicron (B.1.1.529): The Australian sequence varies between 230 and 21556bp (D63G, R203K, G215C, D377Y), Rabat (Morocco) from 241 to 21446bp (R203K, G204R), Belgium from 212 to 13417bp (P240S).

This gene is with a lot of mutations distributed on its whole genetical code we've as G204R, R203K, and P344S. According to Cavanaugh and *al.* (2021) mutations of the protein, N is a major shifter of interactions between virus and host, increasing viral charge and confirming N protein potential; within this protein, these mutations could enhance its function by setting off genes' overexpression which may rise the effect ofSARS-COV-2 pathogenesis.

-Phylogenetical tree of the ORF3a gene

Exactly like the gene S tree and gene ORF1ab phylogenetical tree of the gene ORF3a (Fig. 4) is constituted of five variants (Alpha, Beta, Delta, Gamma, Omicron), which represent mutations that appeared in many countries.

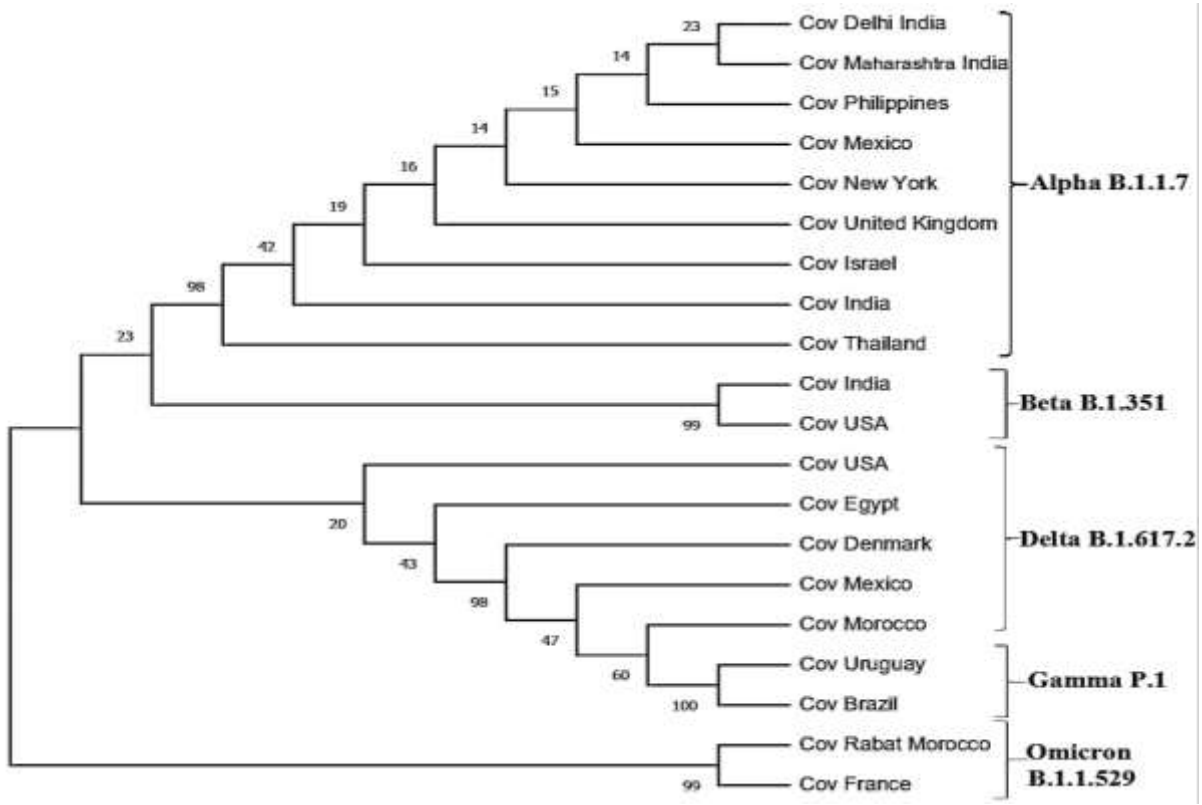

Figure 4. Phylogenetical tree of ORF3a gene representing sequences of different Coronavirus variants.

Sequences of different variants that come up within the ORF3a gene tree are:

Alpha (B.1.1.7): Sequence sizes are as follows: Sequence size Delhi (India) from 25220 to 26456bp (W128L, L140V, G251V), Maharashtra (India) from 25336 to 26122bp (S26L), Philippines from 25355 to 26182bp (S171L, G251V, S26L), Mexico from 25432 to 26222bp (L275F), New York from 25324 to 26151bp (I82T), United Kingdom from 25393 to 26220bp (G251V), Israel from 25391 to 26218bp (Q57H, W131C, W193L, T223I), India from 25364 to 26191bp (Q57H), Thailand from 25361 to 26188bp (F15L).

Beta (B.1.351): The Indian sequence varies between 25393 and 26220 bp (Q57H), USA from 25141 to 25968 bp (S171L, Q57H, S26L, G251V).

Delta (B.1.617.2): The sequence of the USA varies between 25339 and 26166 bp (Q57H), Egypt from 25364 to 26191 bp (Q57H, S26L), Denmark from 25362 to 26189 bp (Q57H, G251V), Mexico from 25346 to 26173 bp (S26L), Morocco from 25368 to 26195bp (S26L).

Gamma (P.1): The Uruguay sequence varies between 25391 and 26218bp (S253P, Q57H), Brazil from 25384 to 26211bp (S253P, S26L, G251V).

Omicron (B.1.1.529): The Rabat sequence (Morocco) varies between 25347 and 26174bp (T64T, Q57H, G251V), the French sequence from 25293 to 26120bp (S92L, D155Y, S26L).

ORF3a protein is with many mutations among which those situated within the ORF3 field (Q57H, S26L, and G251V). These mutations have an impact on the seriousness of the disease or vaccine efficiency. It has been speculated however that they lead to higher concentrations in superior air paths increasing the way virus transmission rate by about 50% (Shrivastava and *al*., 2022).

Covid-19 in Algeria

As well as other parts of the world Algeria faced virus expansion in its various variants embodied in presence of different mutations (Zeghbib and *al*., 2021). In the scope of SARS-CoV-2 sequencing action held by the Pasteur Institute in Algeria and research laboratories for the detection of existing variants by PCR dosage, the presence of five detected variants has been reported Alpha, Beta, Delta, Gamma and Omicron. These variants infected many countries in the same period of time called 'wave' which we've mentioned in the four obtained phylogenetical trees for the gene S (Fig. 1), gene ORF1ab (Fig. 2), gene N (Fig. 3), and the gene ORF3a (Fig. 4).

First wave: Alpha variant

Detected for the first time by WHO in the United Kingdom; appeared in Algeria in the first wave on February 25th, 2020, the first case of the Alpha variant had been introduced by an Italian worker for the farm Eni, who came from Lombardie, one of the most infected areas in Italy, arrived to the camp of Menzel Ledjmet East in Wargla district on February 18th. The peak of this wave indicated 453 dead and 4154 cases confirmed by PCR. This variant showed 15 mutations of the gene S (Tab. I) among which two major mutations which are N501Y and D614G, identified in our phylogenetical tree of the gene S in Berlin in Germany (N501Y), Mexico, Germany and Switzerland (D614G)**.** The N501Y mutation implies replacing Adenine by Thymine in position 23063, which results in replacing Asparagine amino acid with Tyrosine amino acid in position 501. It modifies lightly S protein extremity, which increases its affinity for receptors ACE2 enhancing the way viral attachment and making the virus enter the host cells thus, increasing transmission (Larsen and *al*., 2020; Ketfi and *al*., 2020).

D614G mutation corresponds to the substitution of a Glycine by an Aspartic in position 614 due to a transition of Adenine by Guanine in position 23403. It is situated out of the RDB site and thus it doesn't modify the affinity of link with the receptor ACE2 or neutralize the virus by antibodies, however, it may increase the thickness of the S protein by preserving its structural integrity and avoid loss of the subunit S1, leading consequently to increasing the deal of virus transmissibility that explains being the most overspread mutation in the world, detected in this time in quite all circulating variants (Zhang and *al*., 2020) ; next to that works of Laamrti and *al.* (2020) confirmed that all SARS-Cov-2 variants of North Africa bear the same mutation D614G located on the gene S.

Wave two: Beta variant

The second wave is linked to the Beta variant discovered for the first time in South Africa and appeared in Algeria in June 2020; the first infection known of Beta was the result of a sample taken on June 9th. According to Pasteur Institute, from the beginning to the end of this wave the whole number of cases was 11300 with 2985 deads. This new variant is called Beta or 501Y.V2 because of the mutation it bears N501Y and which is detected in many countries around the world; in our gene S tree, it has been detected in South Africa. N501Y mutation is located within spike protein in the proximity of the split area, which has a major role in merging viral membranes and cellular ones (Greaney and *al*., 2020).

Third wave: Delta variant

In the third wave appeared the Indian variant Delta. Pasteur Institute of Algeria declared that undertaken sequencing activities carried out to control SARS-CoV-2 variants allowed to confirm 7 first cases of Delta variant in Algiers since May 3rd, 2021. With this variant Algeria passed by an increasing number of found cases as well as the rate of mortality, the whole number of infection cases confirmed by PCR has overcome 14036 and 4504 deads.

This variant is the most infectious, the harshest, and the more transmissible than the English one, and it shows less sensibility towards a vaccine, it is known by two specific mutations L452R and E484K that affect the gene S (Larsen and *al*., 2020). L452R mutation has been detected in our S gene tree within Indian DNA sequences (Delhi) and within that of Morocco; it corresponds to replacing Thymine with Guanine in position 1355 seen as a substitution of Threonine by Arginine in position 19. This mutation is implicated in increasing transmissibility and immune escape (Kannan et *al*., 2021). E484K mutation shown in the S gene tree identified in the Danish sequence; leads to replacing Guanine with Adenine in position 1450 which implies replacing Glutamic amino-acid with Lysine in position 484. This mutation allows an eventual escape from antibodies especially polyclonal antibodies derived from infection or earlier vaccination (Larsen and *al*., 2020).

Fourth wave: Gamma variant

The fourth wave is due to the P1 variant (Gamma) identified for the first time in Brazil, showing 10% increased transmissibility in the world compared with B.1.1.7 (Alpha) and B.1.351 (Beta) (Campbell and *al.*, 2021), introduced to Algeria on July 6^{th,} 2021, provoking about 13036 new infections and 6276 deads. This variant holds merely three mutations within the protein Spike (E484K, N501Y and D614G) which makes it highly transmissible and more seriously unable to be stopped by the natural immune brought by an earlier infection by the very first originSARS-CoV-2 (Larsen and *al*., 2020). E484K mutation already found in gene S tree variants Alpha Beta and Delta is found as well within Gamma through a sequence of Uruguay. E484K mutation increases somehow the amount of antibodies necessary to prevent cells' infection (Wise and *al*., 2021). Nevertheless, D614G mutation increases the thickness of S protein on the surface of viral particles and links with the receptorACE2 (Zhang and *al*., 2020, Benvenuto and *al*., 2020); this last mutation is marked in gene S tree by a sample from Venezuela, next the N501Y mutation is marked by a German nucleotide sequence.

Fifth wave: Omicron variant

The fifth wave is due to the Omicron variant, which appeared for the first time in WHO in South Africa; it has been mentioned in Algeria on December 14th, 2021, the first case of the variant was a foreign person who came back to Algeria and has been submitted to an antigenic test in the airport and the result was positive. The wave of the Omicron variant is more infectious but less dangerous than Delta, with a record number of infections in the world as well as in Algeria (17119 new infections and 2503 deads pointed out till February 2022). This variant has a strangely high mutations number among which thirty or so in spike protein, the virus key to enter the organism (Nelson and *al*., 2021), among these mutations T478K represented in our research work by South African sequence and N501Y mutation represented by two sequences that of Bangladesh and that of Australia. T478K mutation disturbs monoclonal antibodies' fixation on the protein S, which is due to the difference in lateral chains' length between Lysine (the longer) and threonine (amino-acid number 478 of the protein S, a Tyrosine has replaced a Lysine within the South African sequence)(Nonaka et *al*., 2021).

Conclusion

All virus kinds including SARS-CoV-2 continue shifting and the majority of mutations have only a little effect or even no effect on the characteristics of the virus. However, some mutations are so worrying and constitute a challenge against prevention and SARS-COV-2 cure. A big part of vaccine types against SARS-COV-2 already approved or those on clinical trial period are conceived starting from the basis of the most variable genes S, ORF1ab, N, and ORF3a. However, the mutation frequency of these genes is very high because it is in the first range against the immune system, a thing which makes virus adapt and develop through mutations able to give it capacities of neutralizing and immune escape, and hence, the task of preparing vaccine becomes more difficult. Through this research work, we could observe that SARS-COV-2 develops starting from its genome's variants and mutations; a big part of detected mutations accumulate in variants Alpha B.1.1.7, Béta B.1.351, Delta B.1.617.2, Gamma P.1 and Omicron B.1.1.529, especially within the most variable genes ; these genes are with mutations distributed on the whole genetic code, among these mutations K417N, L452R, E484K, N501Y, I2230T, P314L, T1001I, P4619L, G204R, R203K, P344S, Q57H, S26L, G251V, T478K, and D614G; having more impact on the fierceness of virus and its transmissibility as well as on its ability to escape from the immune system. Thus, variants holding these mutations are now the most dominant and the most infectious in the world and Algeria as well.

Author's Contributions

Author 1: Data analysis, literature search, drafting the article.
Author 2: Data analysis.
Author 3: Data analysis.

References

Benevento D., Banu Demir A., Giovanetti M., Bianchi M., Angeletti S., Pascarella S., Cauda R., Ciccozzi M., Antonio C., 2020. Evidence for mutations in SARS-CoV-2 Italian isolates potentially affecting virus transmission. J Med Virol ;92(10); P: 2232-2237.

Campbell F., Archer B., Laurenson-Schafer H., Jinnai Y., Konings F., Batra N., Pavlin B., Vandemaele K., Van Kerkhove M. D., Jombart T., Morgan O., le Polain de Waroux O., 2021. Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. Eurosurveillance; Volume 26, Issue 24.

Chen Y., Lieu Q., Guo D., 2021. Emerging coronaviruses: Genome structure, replication, and pathogenesis. J. Med. Virol. 92(4), P: 418-423.

Choi J. Y., Smith D. M., 2021. SARS-CoV-2 Variants of Concern. Yonsei Med. J.; 62(11) P: 961-968.

Clark K., Karsch-Mizrachi I., Lipman D. J., Ostell J., Sayers E. W., 2016. GenBank. Nucleic Acids Research, Volume 44, Issue D1, P: D67–D72.

Derouiche L., Benzayed Y., Belmihoub M., Derouiche F., 2021. A study of genetic variants of SARS-CoV-2 using bioinformatics tools. Genetics and Biodiversity Journal, 6 (1); P: 137-148.

Greaney A.J., Starr T.N., Gilchuk P., Zost S.J., Binshtein E., Loes A.N., Hilton S.K., Huddleston J., Eguia R., Crawford K.H.D., Dingens A.S., Nargi R.S., Sutton R.E., Suryadevara N., Rothlauf P.W., Liu Z., Whelan S.P.J., Carnahan R.H., Crowe Jr J.E., Bloom J.D., 2021. Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. Cell Host Microbe, Volume 29, Issue 1; P: 44-57.

Huang C., Wang Y., Li X., Ren L., Zhao J., Hu Y., Zhang L., Fan G., Xu J., Gu X., Cheng Z., Yu T., Xia J., Wei Y., Wu W., Xie X., Yin W., Li H., Liu M., Xiao Y., Gao H., Guo L., Xie J., Wang G., Jiang R., Gao Z., Jin Q., Wang J., Cao B., 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet, Volume 395, Issue 10223, P: 497-506.

Kannan S., Spratt A., Cohen A., Naqvi S., Chand H., Quinn T., 2021. Evolutionary analysis of the Delta and Delta Plus variants of the SARS-CoV-2 viruses. J Autoimmun. 124:102715.

Ketfi A., Chabati O., Chemali S., Mahjoub M., Gharnaout M., Touahri R., Djenouhat K., Selatni F., Ben Saad H., 2020. Profil Clinique, biologique et radiologique des patients Algériens hospitalisés pour COVID 19 : données préliminaires. Pan Afr Med J; 35 (Suppl 2); P: 77.

Kumar S., Stecher G., Li M., Knyaz C., Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular biology and evolution, 35 (6), P:1547-1549.

Laamarti M., Chemao-Elfihri M.W., Kartti S., Laamarti R., Allam L., Ouadghiri M., Smyej I., Rahoui J., Benrahma H., Diawara I., Alouane T., Essabbar A., Siah S., Karra M., El Hafidi N., El Jaoudi R., Sbabou L., Nejjari C., Amzazi S., Mentag R., Belyamani L., Ibrahimi A., 2020. Genome Sequences of Six SARS-CoV-2 Strains Isolated in Morocco, Obtained Using Oxford Nanopore MinION Technology. Microbiol Resour Announc, 9(32), e00767-20, P: 8.

Larsen C. S., Paludan S.R., 2020. Corona's new coat: SARS-CoV-2 in Danish minks andimplications for travel medicine. Travel Med Infect Dis 38.

Nelson G., Buzko O., Spilman P., Niazi K., Rabizadeh S., Soon-Shiong P., 2021. Molecular dynamic simulation reveals E484K mutation enhances spike RBD-ACE2 affinity and the combination of E484K, K417N and N501Y mutations (501Y.V2 variant) induces conformational change greater than N501Y mutant alone. Potentially resulting in an escape mutant, BioRxiv The preprint server for biology. P: 13.

Nonaka C.K.V., Franco M.M., Gräf T., de Lorenzo Barcia C.A., de Ávila Mendonça R.N., Felix de Sousa K.A., Neiva L.M.C., Fosenca V., Mendes A.V.A., de Aguiar R.S., Giovanetti M., 2021. Genomic Evidence of a Sars-Cov-2 Reinfection Case with E484K Spike Mutation in Brazil. Emerg Infect Dis; 27(5); P: 1522–1524.

Sayers E. W., Cavanaugh M., Clark K., Pruitt K. D., Schoch C. L., Sherry S. T., Karsch-Mizrachi I. K., 2020. GenBank. Nucleic Acids Res. 49(D1): D92-D96.

Shrivastava S., Mhaske S.T., Modak M.S., Virkar R.G., Pisal S.S., Mishra A.C., Arankalle V.A., 2022. Emergence of two distinct variants of SARS-CoV-2 and an explosive second wave of COVID-19: the experience of a tertiary care hospital in Pune, India. Arch Virol. 167(2); P: 393-403.

Tamura K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S., 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Likelihood, Distance, and Parsimony methods. Molecular Biology and Evolution. Vol. 28, Iss. 10, P: 2731-2739.

Cavanaugh A.M., Fortier S., Lewis P., Arora V., Johnson M., George K., Tobias J., Lunn S., Miller T., Thoroughman D., Spicer K.B., 2021. COVID-19 Outbreak Associated with a SARS-CoV-2 R.1 Lineage Variant in a Skilled Nursing Facility After Vaccination Program - Kentucky, March 2021. MMWR Morb Mortal Wkly Rep, 70(17). P: 639-643.

Thompson J. D., Higgins D. G., Gibson T. J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22(22), P: 4673-80.

Wang Y., Zhang D., Du G., Du R., Zhao J., Jin Y., Fu S., Gao L., Cheng Z., Lu Q, Hu Y., Luo G., Wang K., Lu Y., Li H., Wang S., Ruan S., Yang C., Mei C., Wang Y., Ding D., Wu F., Tang X., Ye X., Ye Y., Liu B., Yang J., Yin W., Wang A., Fan G., Zhou F., Liu Z., Gu X, Xu J., Shang L., Zhang Y., Cao L., Guo T., Wan Y., Qin H., Jiang Y., Jaki T., Hayden F.G., Horby P.W., Cao B., Wang C., 2020. Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial, 16; 395(10236). P:1569-1578.

Wise J., 2021. Covid-19: The E484K mutation and the risks it poses. BMJ; 372: n359.

WHO, 2022. SARS-CoV-2 variant tracking. World Health Organization.

Wu D., Lu J., Liu Y., Zhang Z., Luo L., 2020. Positive effects of COVID-19 control measures on influenza. International Journal of Infectious Diseases. Volume 95, P: 345-346.

Xu J., Zhao S., Teng T., Abdalla A. E., Zhu W., Xie L., Wang Y., Guo X. 2020. Systematic comparison of Two animal-to-human Transmitted Human Coronaviruses: SARS-CoV-2 and SARS-CoV. Viruses, 22;12(2). P: 244.

Zárate S., Taboada B., Muñoz-Medina J.E., Iša P., Sanchez-Flores A., Boukadida C., Herrera-Estrella A., Mojica N.S., Rosales-Rivera M., Gómez-Gil B., Salas-Lais A.G., Santacruz-Tinoco C.E., Montoya-Fuentes H., Alvarado-Yaah J.E., Molina-Salinas C.M., Espinoza-Ayala G.E., Enciso-Moreno J.A., Gutiérrez-Ríos R.M., Loza A., Moreno-Contreras J., García-López R., Rivera-Gutierrez X., Comas-García A., Wong-Chew R.M., Jiménez-Corona M.E, Del Angel R.M., Vazquez-Perez J.A, Matías-Florentino M., Pérez-García M., Ávila-Ríos S., Castelán-Sánchez H.G., Delaye L., Martínez-Castilla L.P., Escalera-Zamudio M., López S., Arias C.F., 2022. The Alpha Variant (B.1.1.7) of SARS-CoV-2 Failed to Become Dominant in Mexico. Microbiology Spectrum, 10(2): e0224021.

Zhang H., Penninger J. M., Li Y., Zhong N., Slutsky A. S., 2020. Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target. Intensive Care Med; 46(4), P: 586-590.

Zeghbib S., Somogyi B.A., Zana B., Kemenesi G., Herczeg R., Derrar F., Jakab F., 2021. The Algerian Chapter of SARS-CoV-2 Pandemic: An Evolutionary, Genetic, and Epidemiological Prospect. Viruses ;13(8). P:1525.