

République algérienne démocratique et populaire
Ministère de l'enseignement supérieur et de la recherche scientifique
Université SAAD DAHLEB- Blida 1



Faculté des Sciences de la Nature et de la Vie

Département De Biologie

Mémoire de fin d'étude

En vue de l'obtention du diplôme de Master dans le domaine SNV

Spécialité : Génétique

Thème

**Etude par outils bio-informatiques de
quelques mutations de la protéine Spike
du corona virus variant omicron**

Présenté par :

Guelai Radja

Devant le jury :

Nom

Grade/lieu

Qualité

Mohamed Said .R

MCA/Blida

président

Arkam.F

MCB/Blida

examinatrice

Abdul Hussain.A

MCA/Blida

promotrice

Afrit S

assistante /USDB1

Co-promotrice

Promotion 2021/202



Remerciements

" الْحَمْدُ لِلَّهِ الَّذِي بِنِعْمَتِهِ تَتِمُّ الصَّالِحَات "

Je remercie, en premier lieu, Allah, le tout puissant, de m'avoir permis et accorder la volonté, la patience et le courage pour réaliser ce travail. Je présente mon immense gratitude et mes remerciements, les plus sincères, à madame Abdul Hussein de m'avoir fait confiance et d'avoir acceptées d'encadrer ce travail, et à Mlle Afrit de sa disponibilité, ses orientations pertinentes et avisées et surtout sa patience et gentillesse. Je remercie ces deux femmes surtout pour leurs qualités humaines qui ont constitué un apport considérable, sans lequel, ce travail n'aurait pas vu le jour. Je tiens à exprimer mes sincères remerciements à tous les professeurs et employés de l'université, et qui par leurs compétences et sérieux m'ont permis de poursuivre mes études

Un spécial merci pour monsieur mohamed said Je vous suis très reconnaissante pour tout votre travail. Je vous remercie pour tout ce que vous faites pour les étudiants. Il est rare de rencontrer des personnes aussi dévouées et dignes de confiance. Vos efforts ne passent pas inaperçus. Je remercie aussi madame arkam d'avoir accepté d'examiner cet humble travail.

Mes remerciements s'adressent aussi à mes parents, ma famille et toutes les personnes qui m'ont Soutenu, de près ou de loin, pendant mes études et durant la réalisation de ce mémoire de master.





Dédicaces

Je dédie ce modeste travail A mes chers parents en témoignage de ma profonde gratitude et de mon incontestable reconnaissance, pour tous les sacrifices qu'ils me contentent, toute la confiance qu'ils m'accordent et tout l'amour dont ils m'entourent.

A ma chère sœur Fella et mes deux frères Abdelkader et Abdelghani que je ne trouverais jamais assez de mots pour leurs exprimer mon amour, En leurs espérant le plein succès.

A mes cousins, cousines, amis et à tous les collègues de Master 2
génétique

Et tous ceux qui m'aiment

Et à tous ceux qui me sont chers.



Résumé

D'après l'OMS la pandémie de COVID-19 causée par le virus SARS-CoV-2 a entraîné plus de 600 millions d'infections et 6,5 millions de décès dans le monde. La variante Omicron récemment apparue a soulevé de sérieuses inquiétudes quant à la réduction de l'efficacité des vaccins et des anticorps neutralisants en raison de son grand nombre de mutations notamment au niveau de la protéine de surface connue par (SPIKE PROTEINE).

Dans cette étude, on a sélectionné 15 mutations les plus connues et les plus présentes chez le variant omicron : (A67V, T95I, S371L, S373P, S375F, N404K, G446S, S477N, E484A, Q493R, Q498R, N501Y, Y505H, N856K, N969K) toute ces mutations causent des substitutions d'acide aminé, on a utilisé une combinaison d'approches in silico conventionnelles et avancées basées sur des réseaux de neurones pour prédire comment ces mutations affecteraient la protéine de pointe.

On a étudié l'impact de ces mutations chez le variant omicron de différent coté (phylogénétique, physicochimique, fonctionnel, structural, interactionnel) et d'utiliser plusieurs et différents outils bio-informatique. On a aussi analysé l'interaction (RBD-ACE2) en effectuant un amarrage moléculaire.

Les résultats ont démontré une augmentation de l'alcalinité de la protéine d'un point isoélectrique de 1.5 à 7, une modification de l'hydrophobicité, variations des résidus fonctionnels, le taux de distance entre les 5 variant VOC connues et que le variant Alpha étant le plus proche au variant omicron on a déduit aussi que le variant omicron contenant les 15 mutations est distant de 6.03577 paire de base par rapport à la (WIV04) séquence de GISAID utilisé comme séquence référence

Les résultats de l'amarrage indique que le RBD du variant Omicron avait une affinité plus élevée que la référence. De plus, des simulations de dynamique moléculaire de tous les atomes ont conclu que le RBD de la variante Omicron présente un réseau d'interaction plus dispersé puisque les mutations ont entraîné une augmentation du nombre d'interactions hydrophobes et de liaisons hydrogène avec hACE2

Mots clé : bio-informatique, proteinespike, sarscovide 2, variant omicron, mutations

Abstract

The COVID-19 pandemic caused by the SARS-CoV-2 virus has led to more than 600 million infections and 6.5 million of deaths worldwide. The recently occurred Omicron variant raised serious concerns about reducing the efficacy of vaccines and neutralization antibodies due to its large number of mutations especially in the surface protein known as (SPIKE PROTEINE).

In this study, we have selected the 15 most known and most present mutations in the omicron variant: (A67V, T95I, S371L, S373P, S375F, N404K, G446S, S477N, E484A, Q493R, Q498R, N501Y, Y505H, N856K, N969K) all these mutations cause amino acid substitutions, we used a combination of conventional and advanced in silico approaches based on neural networks to predict how these mutations would affect the spike protein.

We studied the impact of these mutations in the omicron variant from different sides (phylogenetic, physicochemical, functional, structural, interactional) and we used several and different bioinformatics tools. We also analyzed the interaction (RBD-ACE2) by performing a molecular docking

The results demonstrated an increase in the alkalinity of the protein an isoelectric point 1.5 to 7 , a modification of the hydrophobicity, the variations of the functional residues, the rate of distance between the 5 known VOC variants and that the Alpha variant being the closest to the omicron variant. We also deduced that the omicron variant containing the 15 mutations is 6.03577 base pairs away from (WIV04) the GISAID sequence used as reference sequence.

The results revealed that the RBD of the Omicron variant had a higher affinity than the reference. Additionally, all-atom molecular dynamics simulations concluded that the Omicron variant RBD exhibits a more dispersed interaction network since the mutations resulted in an increase in the number of hydrophobic interactions and hydrogen bonds with hACE2

Key words: bio-informatics, spikeproteine, sarscovid2, variant omicron, mutations

ملخص

تسببت جائحة كوفيد 19 الناجم عن فيروس (سارس-كوفيد 2) في أكثر من 600 مليون إصابة و 6.5 مليون حالة وفاة في جميع أنحاء. ظهرت العديد من المتغيرات الرئيسية لـ(سارس كوفيد 2) وطُرحت تحديات في السيطرة على الوباء أثار متغير "اوميكرون" مخاوف جدية بشأن انخفاض فعالية اللقاحات وتحبيد الأجسام المضادة نظرًا لوجود عدد كبير من الطفرات خاصة في البروتين السطحي المعروف باسم (سبايك بروتين)

في هذه الدراسة ، اخترنا 15 طفرة من أشهر الطفرات وأكثرها انتشارًا في المتغير (اوميكرون)

(A67V, T95I, S371L, S373P, S375F, N404K, G446S, S477N, E484A, Q493R, Q498R, N501Y, Y505H, N856K, N969K)

كل هذه الطفرات تسبب بدائل في الحمض الاميني . الأحماض الأمينية

استخدمنا مزيجًا من الأساليب التقليدية والمتقدمة في السيليكو استنادًا إلى الشبكات العصبية للتنبؤ بكيفية تأثير هذه الطفرات على البروتين قمنا أيضًا بتحليل التفاعل (مجال ربط المستقبلات- الأنجيوتنسين2) عن طريق إجراء الالتحام الجزيئي. لقد قمنا بدراسة تأثير هذه الطفرات في المتغير (اوميكرون) من جوانب مختلفة (النشوء والتطور، والفيزياء الكيميائية، والوظيفية، والهيكلية، والتفاعلية) واستخدام العديد من أدوات المعلوماتية الحيوية المختلفة

أظهرت النتائج زيادة في قلوية البروتين ارتفاع نقطة تساوي الكهرباء من 1.5 إلى 7 ، وتعديل الكراهية للماء ، وتغيرات في المخلفات الوظيفية ، ومعدل المسافة بين المتغيرات الخمسة المعروفة للكوفيد وأن المتغير ألفا هو الأقرب إلى متغير اوميكرون تم أيضًا استنتاج أن متغير اوميكرون الذي يحتوي على 15 طفرة يبعد ب 6.03577 زوج قاعدي عن المرجع

و درسنا أيضًا التغيير الهيكلي لالبروتين سبايك عند المتغير اوميكرون. والتفاعل بين (مجال ربط المستقبلات-

كشفت النتائج أن مجال ربط المستقبلات للمتغير اوميكرون كان له تقارب الأنجيوتنسين2)

بالإضافة إلى ذلك ، خلصت عمليات محاكاة الديناميكيات الجزيئية الشاملة إلى أن (مجال ربط أعلى من المرجع.

المستقبلات) للمتغير اوميكرون يعرض شبكة تفاعل أكثر تشتتًا بحيث أن أدت الطفرات إلى زيادة عدد التفاعلات الكارهة

للماء والروابط الهيدروجينية مع - الأنجيوتنسين2

كلمات مفتاحية : المعلوماتية الحيوية , بروتين سبايك , سارس-كوفيد 2 , المتغير اوميكرون , الطفرات

Liste des abréviations

ACE-2 : Angiotensine Conversion Enzyme -2

ssRNA : Single Stranded RNA

RBD : Recepteur Binding Domaine

HR1/2 : Heptade Repeted

TM : TransMembrane

TMPRSS2 : Transmembranaire Protéase à sérine 2

NTD : N Terminal Domaine

CTD : C Terminal Domaine

FP : Fusion Peptide

OMS : Organistaion Mondial de la Santé

VOC : Variant Of Conserne

VOI : Variant Of Interest

VUM : Virus Under Monitoring

GISAID : Global Initiative Sharing Avian Influenza Data

ORF : Open Reading Frame

PCR : Polymerase Chain Reaction

SGTF : S Gene Target Failure

NCBI : National Center for Biotechnology Information

PDB :Protein Data Bank

Liste des tableaux.

Tableau 1. Statistiques covid-19 dans le monde entier en Algérie et à Blida mai 2022	14
Tableau 2. Différents symptômes du covid-19, grippe et le froid (FORTUNE STAFF, s. d.)	14
Tableau 3. Résultats physicochimiques.	50
Tableau 4. Résultats de l'analyse fonctionnelle et de la stabilité.	51
Tableau 5. Charge mol des : (1R42), (6M0J) et (7WK6) (Résultat personnel 2022)	53
Tableau 6. Résultat docking RBD- hACE2. (Résultats personnel 2022)	54

Liste des figures.

Figure 1 Taxonomie Coronaviridae. (Fehr et Perlman 2015).....	3
Figure 2 Structure sars-cov2, génome et domaine fonctionnel de la protéine S.....	4
Figure 3 Explication de la nomenclature omicron et les autres variants.(OMS2022)	9
Figure 4 Prévalence de quelques mutations à travers les lignées.	12
Figure 5 les mutations de pointe de 5 variants (VOC) (He et al. 2021).....	14
Figure 6 Mutations de pointe chez les variants Delta et Omicron (He et al. 2021)	15
Figure 7 Répartition variants du sars-cov2 dans le monde du 12-2019 au -05-2022 (oms 2022).....	16
Figure 8 les variants du sars-cov2 dans le monde pour le mois de mai 2022 (GISAID 2022)	17
Figure 9 variants sars-cov2 en Afrique, A : du 12-2019 au 05-2022, B le 5.2022 (GISAID 2022)	18
Figure 10 Stratégie du diagnostic covid-19 (Drobysh et al. 2022)	20
Figure 11 Exemple du résultat de la recherche sur GISAID (gisaid.org).....	29
Figure 12 Exemple de la transcription par Expasy (expasy.org)	30
Figure 13 Page de face GISAID (gisaid.org).....	31
Figure 14 Séquence nucléotidique de la Spike protéine (gisaid.org).....	31
Figure 15 Manipulation des séquences par sms (bioinformatics.org/sms2/).....	32
Figure 16 l'outil de Numération du génome par sms (bioinformatics.org/sms2)	32
Figure 17 Résultat numération du génome par sms (bioinformatics.org/sms2)	32
Figure 18 Etapes d'introduire les séquences sur Mega 11 (Megasoftware).....	33
Figure 19 Résultat alignement par mega 11 (Meg software).....	34
Figure 20 Etape de la construction de l'arbre Phylogénétique par mega11 (megasoftware)	34
Figure 21 A et B explication de la lecture des Colonne (megasoftware)	35
Figure 22 Exemple figure de face de l'utilisation pepstats sur emboss (ebi.ac.uk)	36
Figure 23 Le Lien utilisé de DynaMUT.....	37
Figure 24 Le choix d'analyse proposer par DynaMut.....	37
Figure 25 Exemple d'introduction des mutations sur DynaMut	38
Figure 26 Exemple du résultat donné par DynaMut	39
Figure 27 Exemple résultat 3D proposé par DynaMut	39
Figure 28 Lien deepDDG utilisé.....	40
Figure 29 Figure de face du serveur SNAP2.....	41
Figure 30 la prediction des mutation neutral et qui affect par SNAP2.....	42
Figure 31 recherche de l'effet de la mutation de l'AA 67	42
Figure 32 exemple de la mutation Y505H	43
Figure 33 la sequence 7N1U par pymol.....	44

Sommaire

Introduction

Partie theorique

I. Généralités :	2
1. Sars covid 2 :	2
1.1 Définition :	2
1.2 Taxonomie :	3
1.3 Structure :	3
1.4 .cycle de vie :	5
2. Fonctionnement de la protéine spike :	5
3. Le Variant omicron	8
3.1. Historique	8
3.2 Nomenclature :	9
3.3. Caractéristiques :	9
3.4. Génome et mutation :	10
3.5. Lignées et sous-lignées omicron :	15
4. Epidémiologie :	16
5. symptômes covid	19
6. diagnostique sars-cov-2 :	20
7. prévention et conseil covid-19 :	23

partie pratique

Introduction	25
Définition des outils bio-informatique utilisés	25
Matériels et méthodes	27
Analyses et résultats	59
Conclusion	72
References	73

Introduction

La maladie à coronavirus (COVID-19), causée par le virus du syndrome respiratoire aigu sévère coronavirus 2 (SRAS-CoV-2), s'est rapidement propagée dans le monde entier et a été déclarée pandémie mondiale en tant qu'urgence de santé publique de portée internationale, qui a causé de graves effets négatifs sur la santé et l'économie dans le monde. (Y.-C. Wu, Chen, et Chan 2020) (Ozili et Arun 2020)

Plusieurs vaccins ont été développés et appliqués pour empêcher la propagation des virus SARS-CoV-2 (Mathieu et al. 2021). Cependant, ces efforts sont remis en question par des variantes virales émergentes dues à des mutations (Tregoning et al. 2021). Parmi les variants majeurs, plusieurs souches ont été qualifiées de "variantes préoccupantes (VOC)" par l'Organisation mondiale de la santé (OMS).

Le 26 novembre 2021, l'OMS a nommé une nouvelle variante (B.1.1.529) pour être Omicron, désignée comme COV. Un total de 30 mutations de la glyco-protéine Spike du variant Omicron ont été identifiées dont 15 localisées au niveau de son domaine de liaison au récepteur et certaines de ces mutations sont aussi présentes dans d'autres VOCs.

Une infection par le SRAS-coV-2 commence par la liaison de la protéine de pointe virale à l'enzyme de conversion de l'angiotensine 2 (ACE2), suivie du traitement protéolytique de la protéine de pointe trimérique en sous-unités S1 et S2 par la sérine protéase furine (Li et al. 2021). À son tour, S2 aide à la fusion de la membrane cellulaire virale avec l'hôte, conduisant à l'entrée virale dans le cytoplasme cellulaire via l'endocytose. La protéine de pointe n'est pas seulement le partenaire de liaison du récepteur ACE2 (enzyme de conversion de l'angiotensine 2) (Shang et al. 2020). Mais aussi site majeur d'antigénicité, donc cible de nombreux anticorps ou médicaments, il est crucial d'étudier les impacts des mutations sur cette protéine.

Plusieurs études ont rapporté une affinité plus élevée pour le récepteur ACE2 dans la protéine de pointe mutante d'Omicron que dans le SRAS-CoV-2 de type sauvage, ainsi qu'une plus grande capacité à échapper au système immunitaire, ce qui pourrait entraîner une transmissibilité virale plus élevée (Lupala et al. 2022)

Au meilleur de notre connaissance, aucune des études n'a examiné les effets détaillés sur les propriétés physicochimiques, structurales ou fonctionnelles de la protéine de pointe en raison du grand nombre de mutations trouvées dans la protéine de pointe. Par conséquent, le but de notre étude était d'utiliser une approche *in silico* pour analyser de manière exhaustive 15 mutations de la protéine de pointe du variant Omicron par rapport au variant de référence afin de distinguer les différences d'un point de vue physicochimique, structurel, phylogénétique et fonctionnel. De plus, nous avons effectué un amarrage protéine-protéine avec le récepteur ACE2 afin d'étudier l'effet des mutations sur les interactions protéiques.



Partie 1

Théorique



I. Généralités :

1. Sars covid 2 :

1.1 Définition :

Coronavirus 2 du syndrome respiratoire aigu sévère (SARS-CoV-2) est un virus hautement transmissible et pathogène de la famille des *Coronaviridae* qui est apparu en fin 2019 et a provoqué une pandémie de maladie respiratoire aiguë, appelée « maladie à coronavirus 2019 » : COVID-19 (Hui et al. 2020) (J. T. Wu, Leung, et Leung 2020) , Les coronavirus sont un groupe diversifié de virus infectant de nombreux animaux différents, et ils peuvent provoquer des infections respiratoires légères à graves chez l'homme (Cui, Li, et Shi 2019). En 2002 et 2012, respectivement, deux coronavirus hautement pathogènes d'origine zoonotique, le coronavirus du syndrome respiratoire aigu sévère (SRAS-CoV) et le coronavirus du syndrome respiratoire du Moyen-Orient (MERS-CoV), sont apparus chez l'homme et ont provoqué une maladie respiratoire mortelle, faisant des coronavirus émergents une nouvelle préoccupation de santé publique au XXI^e siècle (Cui, Li, et Shi 2019).

Les coronavirus sont des virus à ARN enveloppés qui sont largement distribués parmi les humains, les autres mammifères et les oiseaux et qui provoquent des maladies respiratoires, entériques, hépatiques et neurologiques (Weiss et Leibowitz 2011) (Cui, Li, et Shi 2019). Six espèces de coronavirus sont connues pour causer des maladies humaines. Quatre virus - 229E, OC43, NL63 et HKU1 - sont répandus et provoquent généralement des symptômes de rhume chez les personnes immunocompétentes. Les deux autres souches : le coronavirus du syndrome respiratoire aigu sévère (SARS-CoV) et le coronavirus du syndrome respiratoire du Moyen-Orient (MERS-CoV)

- **Protéine S** 1: La protéine S du SRAS-CoV-2 est une glycoprotéine structurelle transmembranaire de fusion de classe I qui est composée de sous-unités S1 et S2 (Huang et al. 2020). C'est un homotrimère d'une taille de 180 à 200 kDa (Yesudhas, Srivastava, et Gromiha 2021) et d'une longueur totale comprise entre 1273 et 1300 acides aminés (Duan et al. 2020). Les résidus d'acides aminés (1-13) constituent un peptide signal. Il s'agit d'une structure en forme de clou de girofle (Huang et al. 2020) . faisant saillie à la surface du virus lui permettant de transporter la liaison virale à la cellule hôte et la fusion (Kumar et Al Khodor 2020). C'est l'unité de base du virus qui

reconnait et se lie au récepteur de la cellule hôte ACE-2 faisant ainsi de la protéine S une cible principale pour neutraliser les anticorps

1.2 Taxonomie :

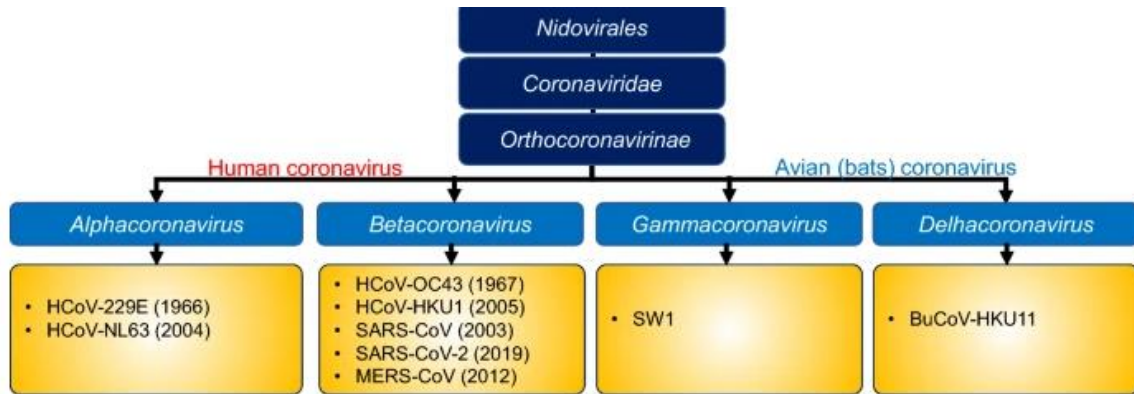


Figure 1 Taxonomie Coronaviridae. (Fehr et Perlman 2015)

1.3 Structure :

Les COVS2 sont des virus à ARN à brin positif avec une nucléocapside et une enveloppe (Chen et al. 2020). Un virion SARS-CoV-2 mesure environ 50 à 200 nm de diamètre (Licastro et al. 2020) et possède un génome +ssRNA3

d'environ 29,9 kb de longueur le plus grand virus à ARN connu avec une structure a 5'-cap et 3'-poly- A-queue et possèdent 14 cadres de lecture ouverts putatifs (ORF) codant pour 27 protéines A. Wu et al. 2020; C. Wu et al. 2020). Le virion a quatre protéines structurales, connues sous le nom de protéines S (pointe), E (enveloppe), M (membrane) et N (nucléocapside) ; la protéine N contient le génome de l'ARN et les protéines S, E et M créent ensemble l'enveloppe virale (Hoffmann et al. 2020). La glycoprotéine-S de pointe facilite la fixation du virus au récepteur de l'enzyme de conversion de l'angiotensine 2 (ACE2) et fusionne avec la membrane de la cellule hôte Hoffmann et al. 2020). Le SARS-CoV-2 utilise ensuite les sérines protéases TMPRSS2 (transmembrane protéase serine 2) pour l'amorçage de la protéine S, infectant les cellules cibles (Jaimes, Millet, et Whittaker 2020). Les protéines de pointe de SARS-CoV-2 contiennent deux sous-unités ; Sous-unité de liaison au récepteur S1 et sous-unité de fusion S2, pour médier la liaison du virion à la protéine réceptrice et

Initier la fusion membranaire. Les sous-unités S1 et S2 sont divisées par le site de clivage S . Pour faciliter la fixation du virion au récepteur et fusionner avec la membrane cellulaire, la protéine de pointe doit être clivée par des protéases cellulaires du site de clivage S1/S2 . Fait intéressant, l'analyse moléculaire des protéines S a identifié une insertion au site S1/S2, qui est absente dans d'autres SARS-CoV , bien que l'importance de cette insertion soit encore inconnue, il semble que cette insertion unique apporte un gain- avantage de la fonction pour une infection cellulaire facile et une propagation efficace dans tout l'hôte humain (Chen et al. 2020).

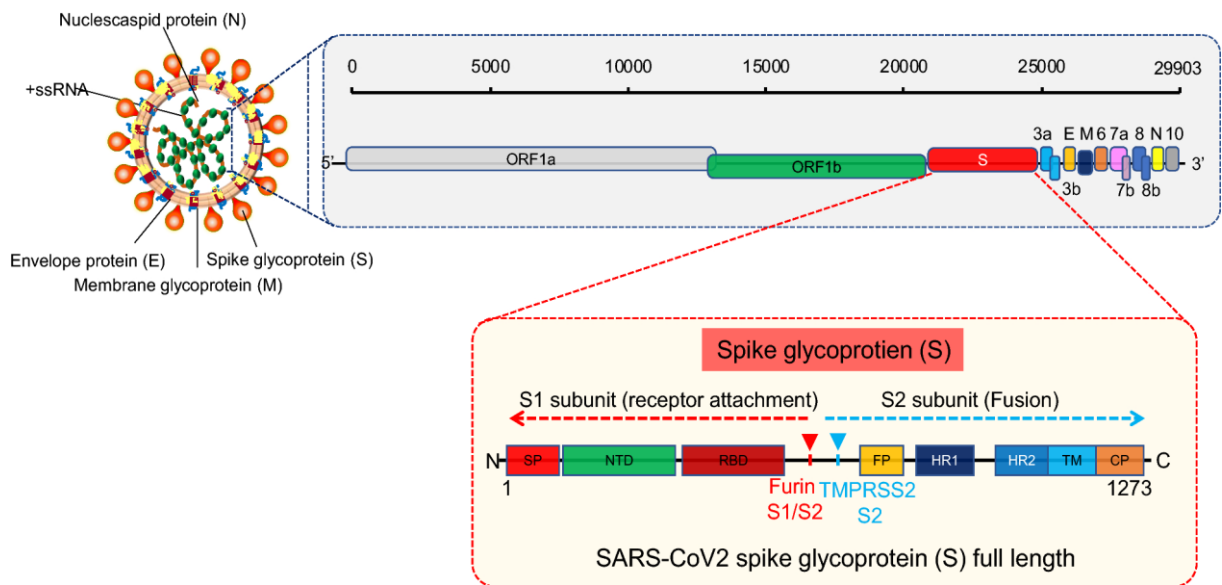


Figure 2 Structure sars-cov2, génome et domaine fonctionnel de la protéine S.

1.4 .cycle de vie :

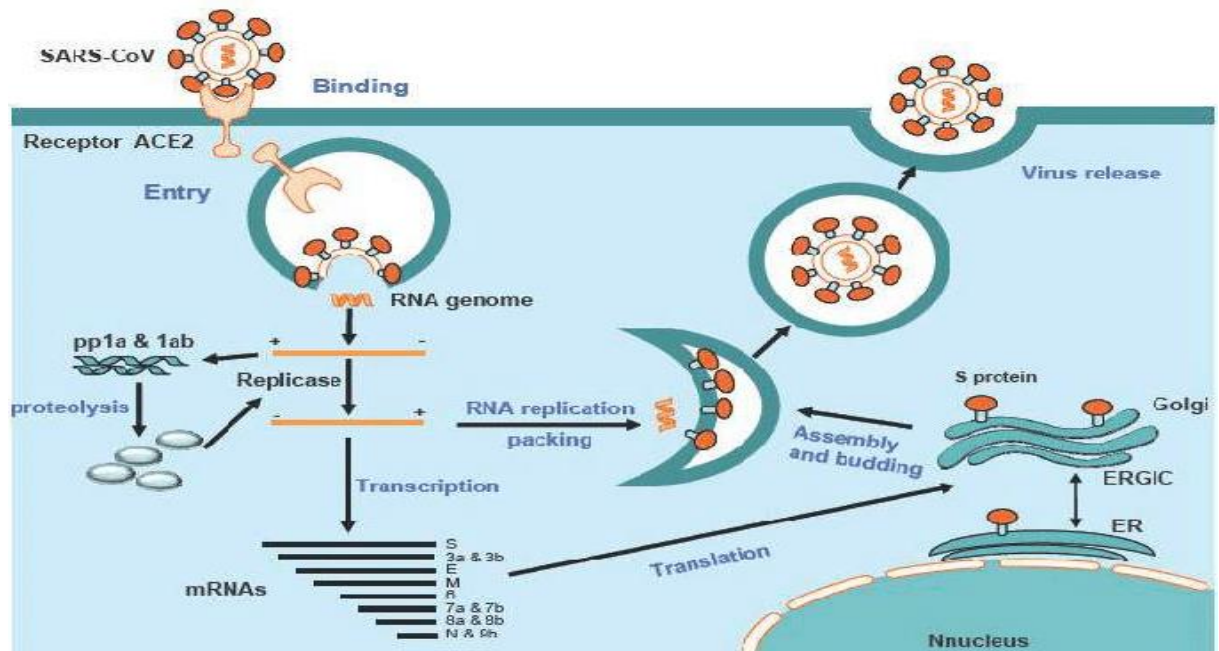


Figure 3 Cycle de vie du SARS-CoV-2 dans la cellule de l'hôte (Alouane et al. 2020).

- 1 : attachement du virus à sa cellule cible grâce à la protéine S du virus
- 2 : le virus est englobé dans une vésicule, on dit qu'il y a endocytose du virus. L'enveloppe du virus fusionne alors avec la vésicule, ce qui libère le génome viral dans la cellule. Cette fusion est déclenchée par la protéine S
- 3 : expression du génome viral et synthèse de protéines virales. En parallèle, le génome viral se réplique
- 4 : début de l'assemblage des nouveaux virus (association des protéines virales, du génome ARN et des lipides de l'enveloppe virale)
- 5 : le virus se trouve dans une vésicule qui migre jusqu'à la membrane cellulaire
- 6 : la vésicule fusionne avec la membrane cellulaire ce qui libère les nouveaux virus.

Un virus infectant une cellule permet la production de plusieurs centaines de nouveaux virus..

2. Fonctionnement de la protéine spike :

La spike protéine est une protéine transmembranaire de type I qui est présente au sein de l'enveloppe sous forme d'un trimère homogène. Chaque monomère se compose de deux sous-unités fonctionnelles, S1 et S2. La sous-unité S1 contient un domaine RBD (Receptor

Binding Domain) qui se fixe au récepteur ACE2 sur la membrane de la cellule pulmonaire hôte. La sous-unité S2 est activée par la liaison de S1 avec ACE2 et contribue à la fusion membranaire. La longueur totale moyenne de la protéine S est de 1273 résidus et se compose d'un peptide signal de 13 résidus situés à l'extrémité N-terminale de la sous-unité S1 (14–685 résidus) et de la sous-unité S2 (686– 1273 résidus). Les deux dernières régions sont respectivement responsables de la liaison au récepteur et de la fusion membranaire

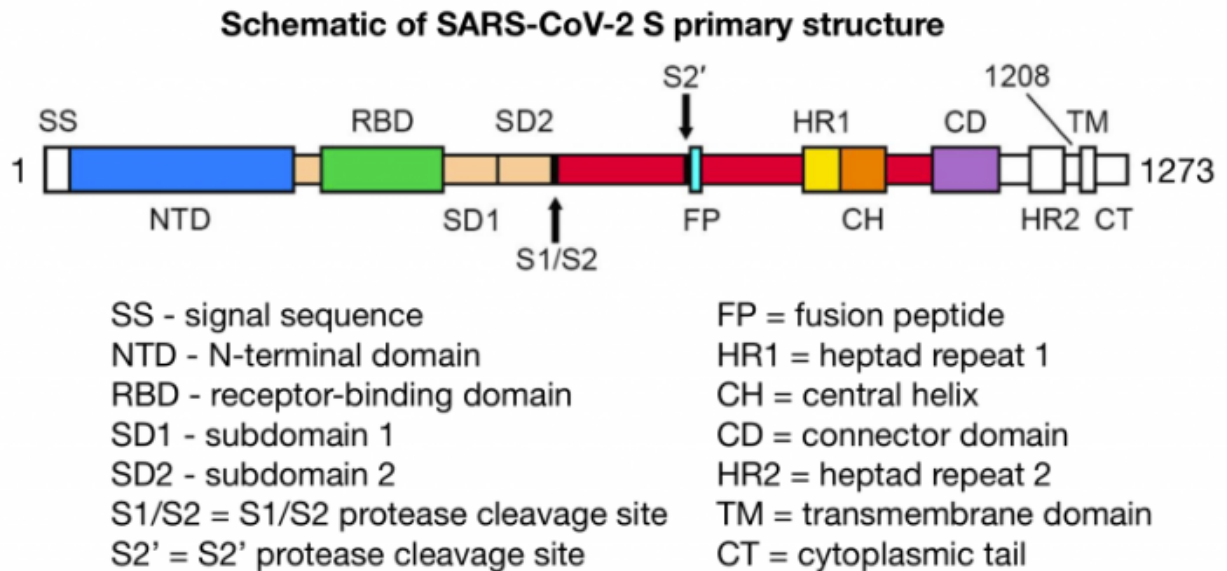


Figure 4 Schéma de la protéine Spike du sars-cov-2

Dans la sous-unité S1, il y a un domaine N-terminal (14 à 305 résidus) et un domaine de liaison au récepteur RBD (319 à 541 résidus). La sous-unité S2 comporte le peptide de fusion (788–806 résidus), la séquence de répétition heptapeptidique 1 (HR1) (912–984 résidus), HR2 (1163–1213 résidus), le domaine TM (1213–1237 résidus) et le domaine du cytoplasme (1237– 1273 résidus)

A l'état natif, la protéine S existe en tant que précurseur inactif. Lors d'une infection virale, les protéases des cellules cibles activent la protéine S en la coupant en sous-unités S1 et S2, ce qui est nécessaire pour activer le domaine de fusion membranaire après l'entrée virale dans les cellules cibles. Semblable à d'autres coronavirus, la protéine S est clivée en sous-unités S1 et S2 par des protéases cellulaires, et la sérine protéase TMPRSS24.

Est utilisée comme amorce protéique, Le RBD situé dans la sous-unité S1 se lie au récepteur cellulaire ACE2 dans la région de l'aminopeptidase N. La région S1 contient le NTD 5 et le CTD, et les détails atomiques à l'interface de liaison démontrent des substitutions de résidus clés dans SARS-CoV-2-CTD. En outre, l'interface de liaison la protéine S CTD 6 a plus de résidus qui interagissent directement avec le récepteur ACE2 que SARS-RBD (21 contre 17), et une plus grande surface est enterrée avec la protéine S CTD dans complexe avec ACE2 qu'avec la protéine S RBD.

La sous-unité S2, composée successivement des domaines FP, HR1, HR2, TM et d'une fusion de domaine cytoplasmique (CT), est responsable de la fusion et de l'entrée du matériel génétique (ARN) du virus à l'intérieur de la cellule hôte

FP est un court segment de 15 à 20 résidus conservés de la famille virale. Il est principalement composé de résidus hydrophobes, tels que la glycine (G) ou l'alanine (A), qui s'ancrent à la membrane cible lorsque la protéine S adopte la conformation de la préépine. Des recherches antérieures ont montré que la FP joue un rôle essentiel dans la médiation de la fusion membranaire en perturbant et en reliant les bicouches lipidiques de la membrane de la cellule hôte. La HR1 et HR2 sont composés d'un heptapeptide répétitif: HPPHCPC, où H est un résidu hydrophobe ou traditionnellement volumineux, P est un résidu polaire ou hydrophile et C est un autre résidu chargé. HR1 et HR2 forment le faisceau à six hélices (6-HB), qui est essentiel pour la fusion virale et la fonction d'entrée de la sous-unité S2. HR1 est situé à l'extrémité C-terminale d'un FP hydrophobe, et HR2 est situé à l'extrémité N-terminale du domaine TM. Le domaine TM en aval ancre la protéine S à la membrane virale et la sous-unité S2 se termine par une queue CT. Le RBD se lie à ACE2 et S2 change de conformation en insérant FP dans la membrane de la cellule cible, en exposant la bobine enroulée de la préhairpin du domaine HR1 et en déclenchant l'interaction entre le domaine HR2 et le trimère HR1 pour former 6-HB, apportant ainsi l'enveloppe virale et membrane cellulaire à proximité pour la fusion et l'entrée virales. HR1 forme un assemblage homotrimérique dans lequel trois rainures hydrophobes hautement conservées sur la

surface qui se lie à HR2 sont exposées. Le domaine HR2 forme à la fois une hélice rigide et une boucle flexible pour interagir avec le domaine HR1. Dans la conformation en épingle à cheveux post fusion des CoV, il existe de nombreuses interactions fortes entre les domaines HR1 et HR2 à l'intérieur de la région hélicoïdale, qui est appelée «région centrale de fusion» (régions HR1core etHR2core, respectivement).

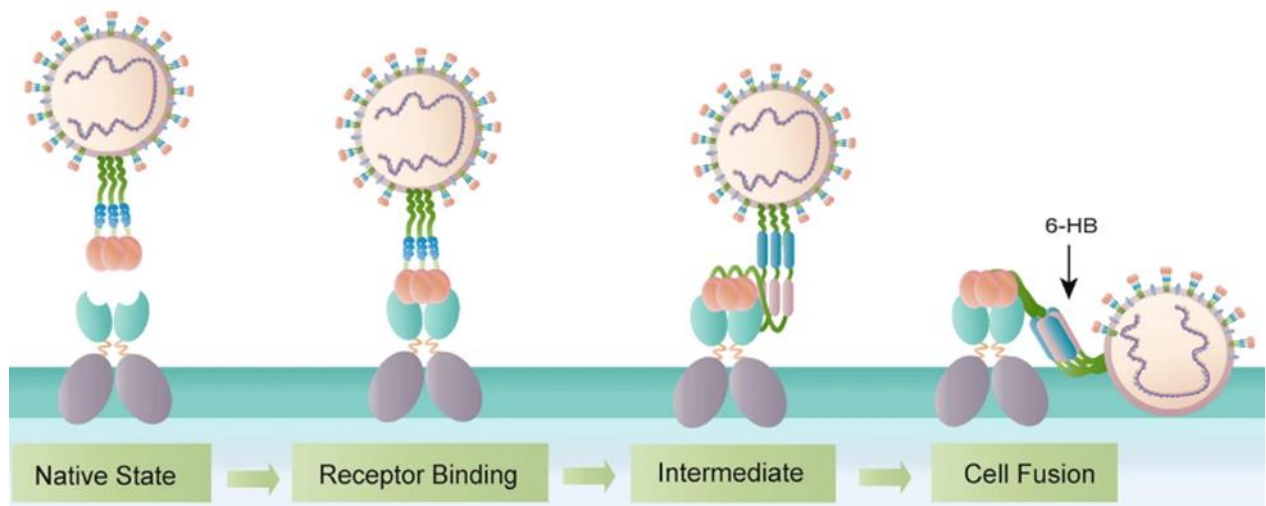


Figure 5 Le processus de liaison et de fusion virus-cellule médié par la protéine S (Huang et al. 2020).

3. Le Variant omicron

3.1. Historique

Le 11 mars 2020, l'OMS a annoncé publiquement l'épidémie de SRAS-CoV-2 comme une pandémie mondiale. La pandémie de maladie à coronavirus 2019 (COVID-19) est en plein essor depuis près de trois ans. Plus de 500 millions de cas confirmés ont été signalés selon les statistiques de l'Organisation mondiale de la santé (OMS), dont plus de six millions de décès. Le virus original du coronavirus 2 du syndrome respiratoire aigu sévère (SRAS-CoV-2) qui a été identifié fin 2019 avait évolué et diverses variantes sont apparues. Afin de prioriser la surveillance et la recherche de ces variantes, l'OMS les a classés en trois catégories : les variants préoccupants (VOC), les variants d'intérêt (VOI) et les variants sous surveillance (VUM). Les quatre VOC précédents incluent Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1) et Delta (B.1.617.2) (22). Ils ont tous entraîné une nouvelle vague de pandémie et des milliers de morts dans plus d'un pays et d'une région, et même dans le monde entier. Le 26 novembre 2021, une nouvelle variante

nommée Omicron (B.1.1.529) a été désignée comme cinquième COV par l'OMS, ce qui a immédiatement soulevé des inquiétudes mondiales

3.2 Nomenclature :

. L'Organisation mondiale de la santé réserve une désignation en lettres grecques pour les « variante préoccupants »(23)

. Le 26 novembre, le groupe consultatif technique de l'OMS sur l'évolution du virus du SRAS-CoV-2 a déclaré que la lignée PANGO B.1.1.529 était un variant préoccupant et l'a désigné par la lettre grecque omicron. L'OMS a sauté les lettres précédentes nu et xi dans l'alphabet grec pour éviter toute confusion avec les similitudes du mot anglais « nouveau » et du nom de famille chinois X (Marcus Parekh et al., s. d.), Le projet GISAID lui a attribué l'identifiant de clade GR/484A et le projet Nextstrain lui a attribué l'identifiant de clade 21K.

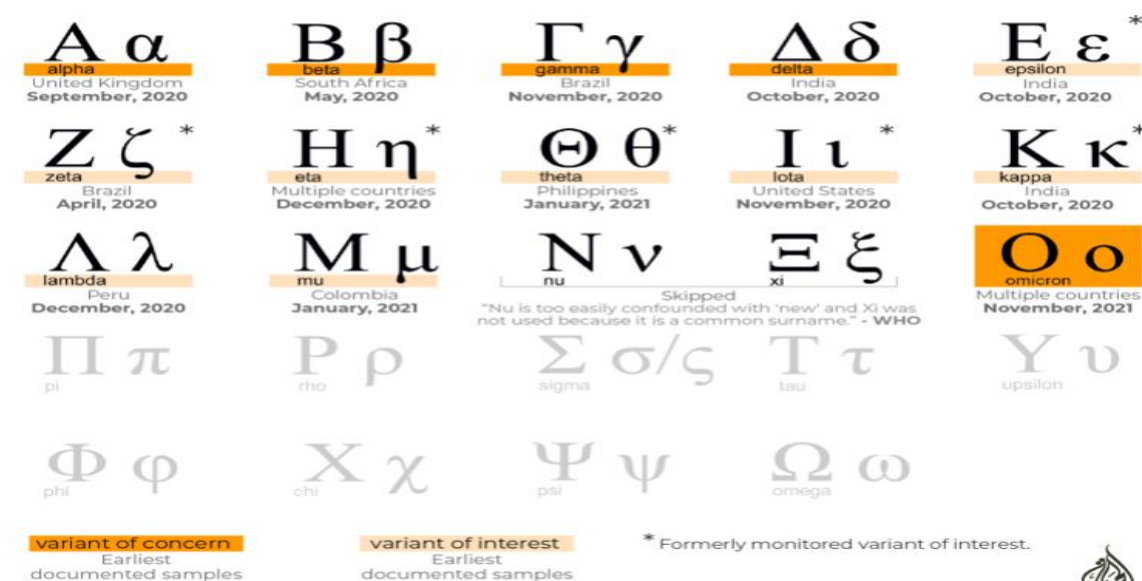


Figure 3 Explication de la nomenclature omicron et les autres variants.(OMS2022)

3.3. Caractéristiques :

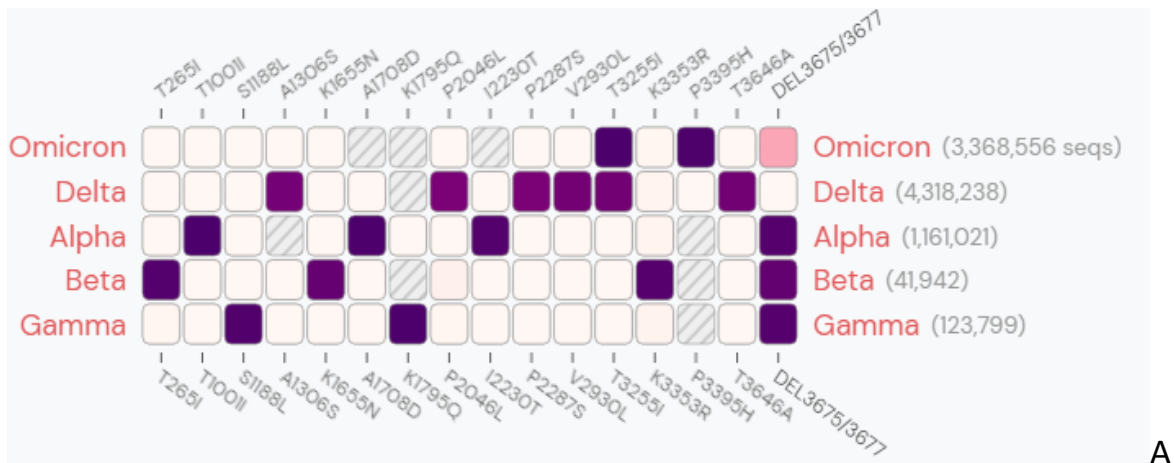
Selon l'OMS, le risque de contracter une forme grave est moins important qu'avec les autres formes du Covid-19, mais sa propagation est plus rapide (Garcia-Beltran et al. 2022). Une étude de décembre 2021 sur des pseudo virus estime qu'il est quatre fois plus infectieux (et non pas contagieux) que le type naturel et deux fois plus que le variant Delta.

Les études initiales en Afrique du Sud publiées en décembre 2021, sur un nombre réduit de patients, mettent en évidence que l'échappement immunitaire est « bien plus important » que celui observé pour les précédents variantes. L'action des anticorps produits par les personnes vaccinées (après deux doses du vaccin de Pfizer).est quarante fois moins bonne contre le variant omicron que contre la souche historique, Selon les premières études européennes, le même schéma vaccinal entraînerait une perte d'efficacité totale ou seulement sept fois inférieure. Toutefois, une troisième dose de vaccin rétablirait une protection partielle (Planas et al. 2022).

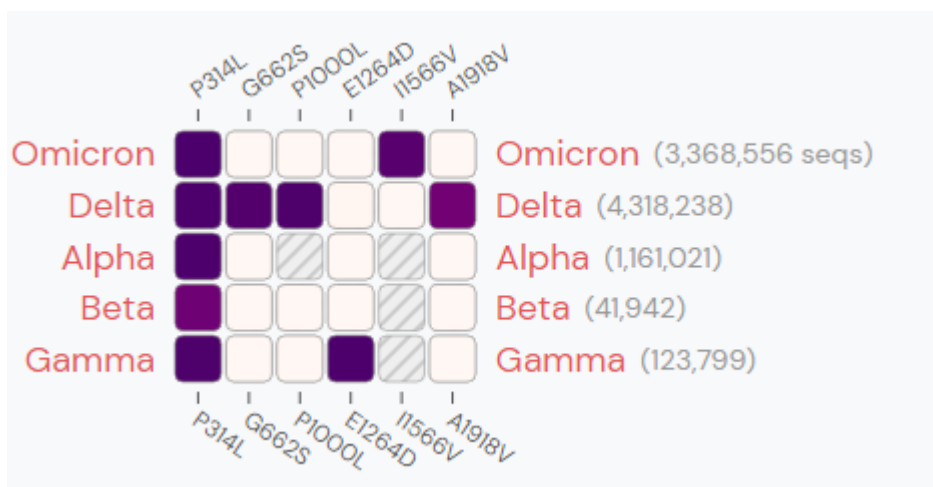
Une étude prépubliée en décembre 2021 menée sur neuf anticorps monoclonaux utilisés en clinique ou en phase de développement préclinique montre que six d'entre eux sont totalement inefficaces face au variant Omicron, tandis que les trois autres voient leur efficacité diminuée de 3 à 80 fois par rapport à leur action sur le variant Delta. Les tests menés sur des sujets un mois après une troisième dose du vaccin Pfizer montrent son efficacité contre le variant Omicron, mais également que ce variant nécessite une quantité d'anticorps de 5 à 31 fois supérieure à celle nécessaire pour le variant Delta pour être neutralisé. La durée de cette efficacité n'est toujours pas connue. La même étude montre par contre que des patients ayant été symptomatiques il y a plus de 12 mois ou ayant reçu les deux doses du vaccin Pfizer il y a plus de cinq mois ne sont quasiment pas protégés contre le variant Omicron (He et al. 2021; Planas et al. 2022).

3.4. Génome et mutation :

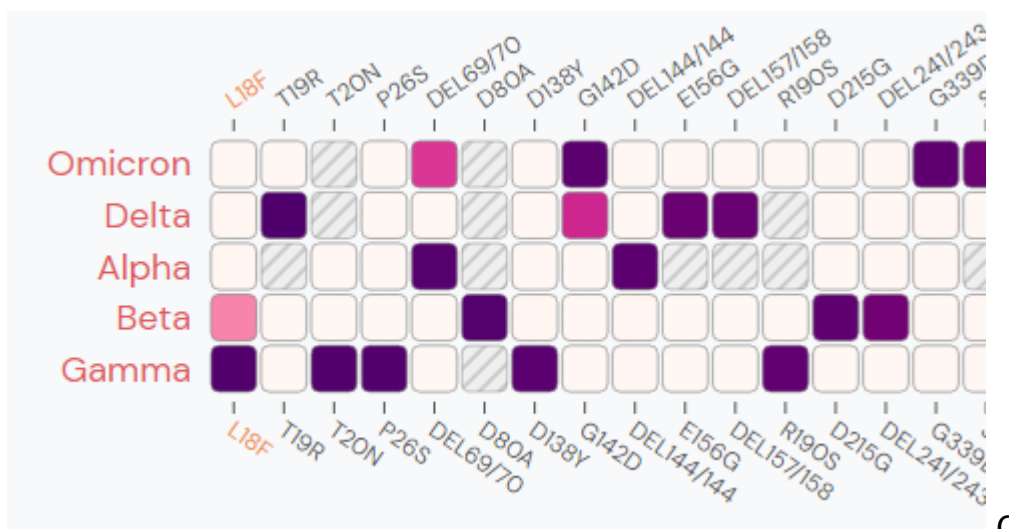
L'analyse des séquences génomiques du variant Omicron a révélé un nombre élevé de mutations non synonymes, dont plusieurs dans la pointe qui se sont avérées impliquées dans la transmissibilité, l'incision de la maladie et l'évasion immunitaire. Au total, plus de 60 substitutions/délétions/insertions ont été identifiées dans la variante Omicron, faisant d'Omicron une variante possédant le plus grand nombre de sites de mutation de toutes les variants du SARS-CoV-2 caractérisées à ce jour (Korber et al. 2020).



A



B



C

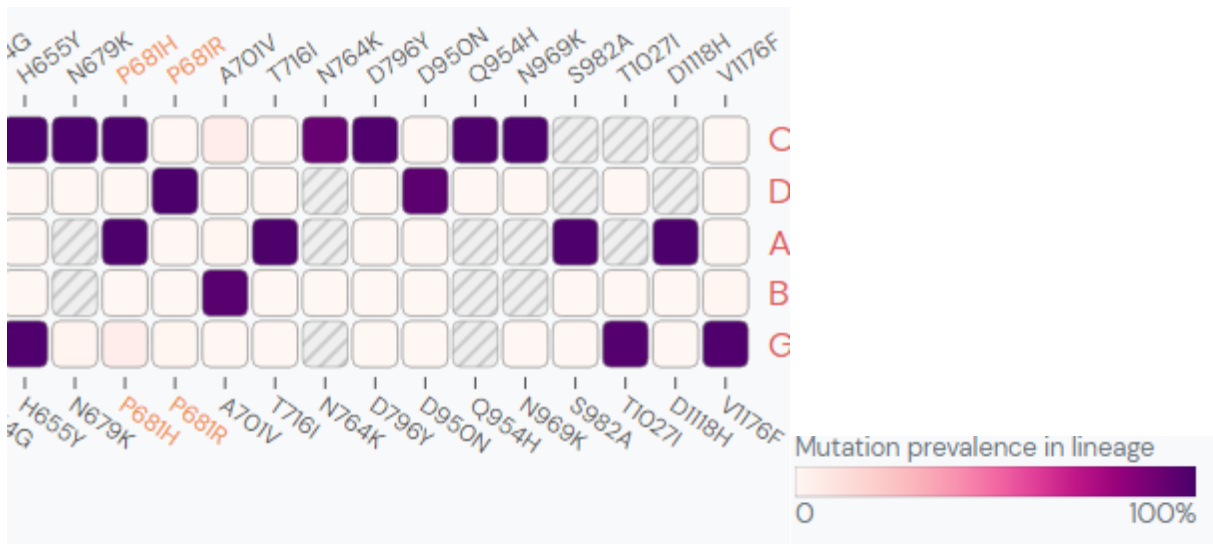
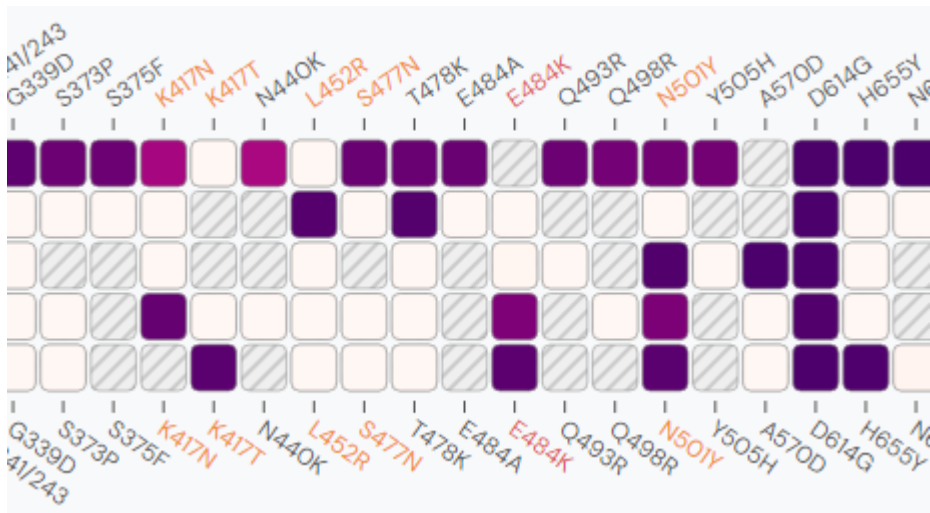


Figure 4 Prévalence de quelques mutations à travers les lignées, A ORF1a, B ORF1b, C Spike protéine, on rapporte actuellement que les changements de substitution et de suppression, en raison des défis informatiques liés à l'identification des insertions dont plus de 5 millions de séquences chaque jour.

La plupart des séquences Omicron contiennent également une insertion de 3 acides aminés (EPE) en position 214 dans la protéine Spike. Plus de la moitié des mutations totales d'Omicron identifiées sont accumulées dans la protéine Spike. Ceux-ci incluent 30 substitutions de A67V, T95I, Y145D, L212I, G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501H, T65K, YG679, YG679, YG679, YG679, YG679, D796Y, N856K, Q954H, N969K et L981F, trois délétions de H69/V70, G142/V143/Y144 et N211, et une insertion de trois acides aminés (EPE) en position 214

(dans certains rapports, les changements sont décrits comme la délétion V143/Y144/Y145 en combinaison avec G142D et la délétion L212 en combinaison avec N211I).

Il est à noter que les cinq variants contiennent le changement d'acide aminé D614G en pointe et la P314L en ORF1b.

. Le variant Omicron partage également N501Y avec les variantes Alpha, Beta et Gamma. On pense que cette mutation améliore la liaison entre la protéine de pointe et l'enzyme de conversion de l'angiotensine 2 (ACE2) et induit une transmissibilité plus élevée. Lorsqu'elle est combinée à la délétion H69/V70, la transmissibilité pourrait être encore augmentée

. L'Omicron possède également des mutations N679K et P681H près du site de clivage de la furine. L'incorporation d'acides aminés basiques autour du site de clivage de la furine pourrait faciliter le clivage de la pointe en S1 et S2, améliorant ainsi la fusion et l'infection virale. En fait, la mutation P681H a également été identifiée dans le variant Alpha . Cette mutation a été suggérée pour améliorer l'infectiosité du SRAS-CoV-2 (Yan et al. 2020).

Le domaine de liaison au récepteur de pointe (RBD) est l'entité virale authentique qui reconnaît le récepteur ACE2 pour méditer l'entrée du virus (Yuan et al. 2021). Alors que la variante Delta ne possède que les mutations L452R et T478K dans le RBD, 15 mutations ont été accumulées dans le RBD de la variante Omicron. Parmi ces substitutions, on observe qu'un groupe de résidus se localise à proximité du récepteur ACE2 lié. La manière dont ces mutations affecteraient la liaison au récepteur reste cependant à étudier à l'avenir.

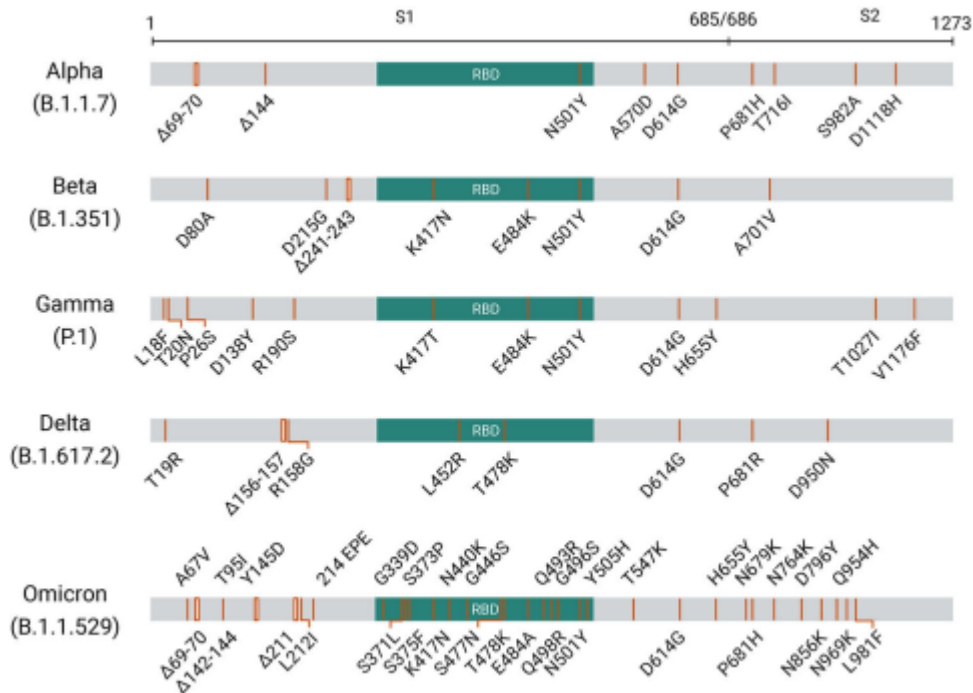


Figure 5 les mutations de pointe de 5 variants (VOC) (He et al. 2021)

. Il est à noter que le pic RBD est la cible principale des anticorps neutralisants et que l’Omicron a accumulé 15 substitutions dans la région RBD. Nous avons noté que plusieurs sites antigéniques ont été caractérisés dans RBD, avec les sites RBS-A, RBS-B, RBS-C, CR302 et S309 .Toutes les 15 mutations identifiées dans Omicron Spike RBD peuvent être localisées sur un ou plusieurs de ces sites antigéniques, indiquant une résistance potentielle d'Omicron à un ou plusieurs des anticorps monoclonaux ciblant ces sites. Quant au traitement par anticorps en usage clinique, le cocktail constitué de LY-CoV555 (également appelé Bamlanivimab) et de LY-CoV016 (également appelé Etesevimab) a été autorisé en urgence. Des études antérieures ont révélé que les mutations aux positions 484 et 417 du pic sont associées à l'évasion immunitaire (Starr et al. 2021), et que les variants bêta et gamma pourraient échapper à la neutralisation de LY-CoV555 (due à E484K) et LY-CoV016 (due à K417N/T)(Starr et al. 2021).Étant donné que la variante Omicron contient également des mutations E484A et K417N, il est probable qu'Omicron résisterait également à ces deux anticorps.

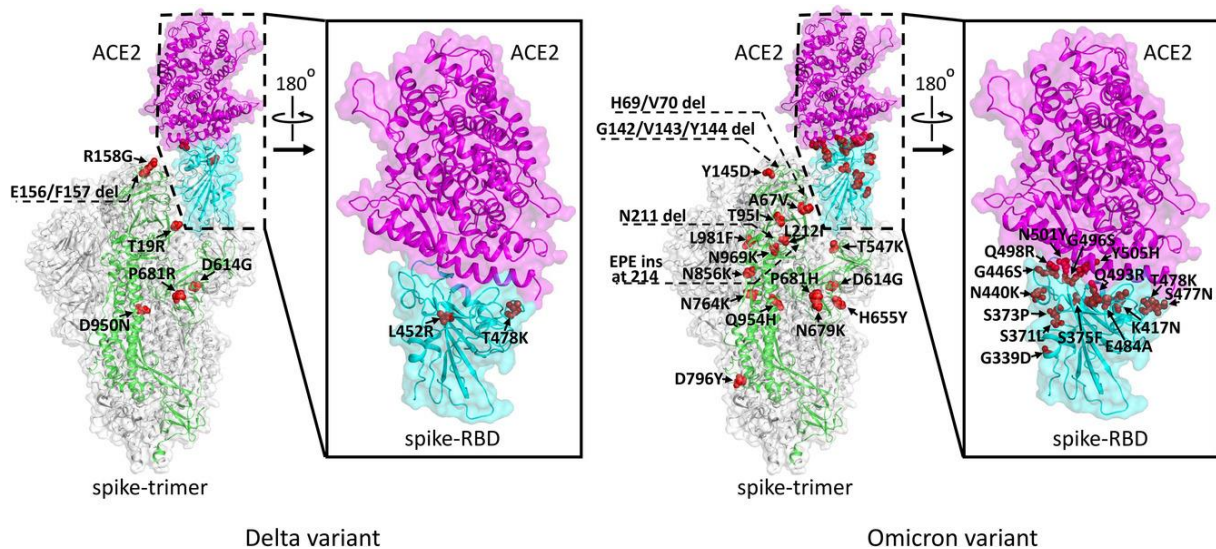


Figure 6 Mutations de pointe chez les variants Delta et Omicron (He et al. 2021)

Les structures sont représentées sur la base de la structure trimère de pointe de cryo-microscopie électronique du code 6VYB de la banque de données protéiques (PDB) et de la structure complexe cristalline RBD/ACE2 du code PDB 6LZG. Un protomère du trimère de pointe est mis en évidence en vert et son domaine de liaison au récepteur (RBD) en cyan. Le récepteur de l'enzyme de conversion de l'angiotensine 2 (ACE2) lié est de couleur magenta. Les mutations définies sur la base de covariants sont étiquetées.

3.5. Lignées et sous-lignées omicron :

Depuis son apparition Plusieurs lignées et sous-lignées omicron sont apparue partout dans le monde

-les lignées omicron : B.1.1.529 (sud Afrique et Botswana)

BA.1 Alias of B.1.1.529.1, / pango-designation issue #361

BA.2 Alias of B.1.1.529.2, / pango-designation issue #361

BA.3 Alias of B.1.1.529.3, / pango-designation issue #367

BA.4 Alias of B.1.1.529.4, / pango-designation issue #517

BA.5 Alias of B.1.1.529.5, /pango-designation issue #517

XE (recombinaison BA.1 et BA.2) pango-designation issue #454.

4. Epidémiologie :

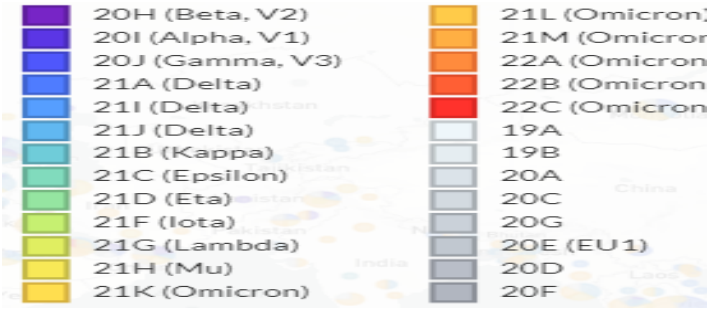
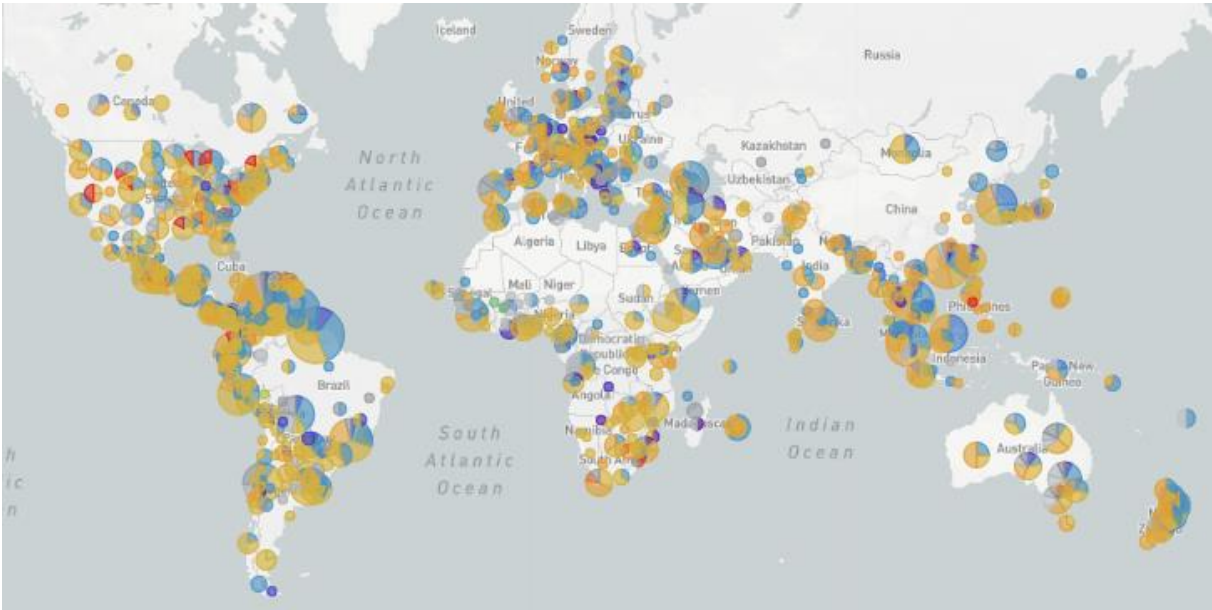


Figure 7 Répartition variants du sars-cov2 dans le monde du 12-2019 au -05-2022 (oms 2022).

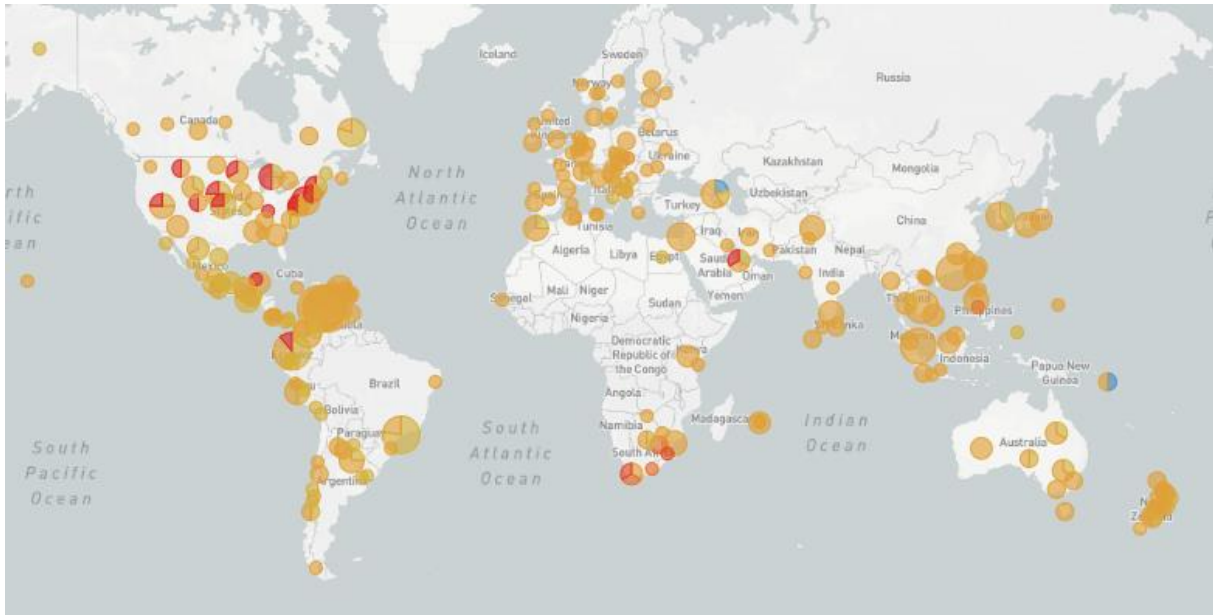
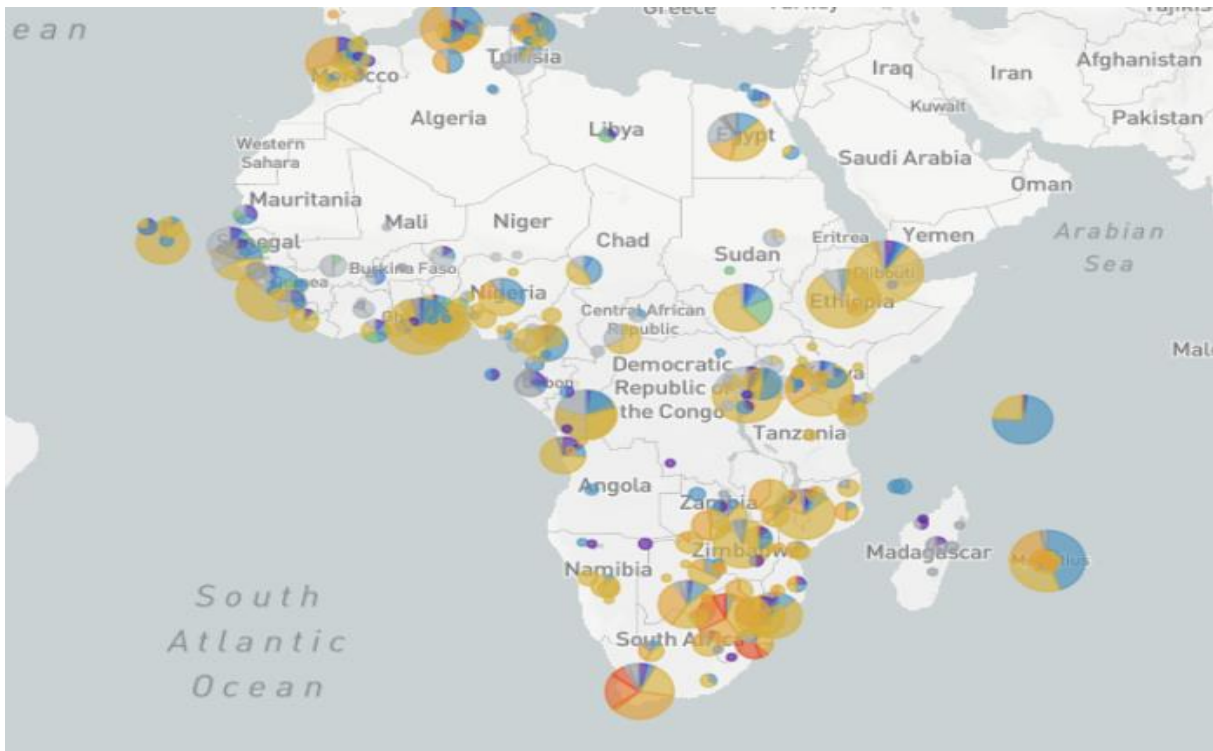


Figure 8 les variants du sars-cov2 dans le monde pour le mois de mai 2022 (GISAID 2022)



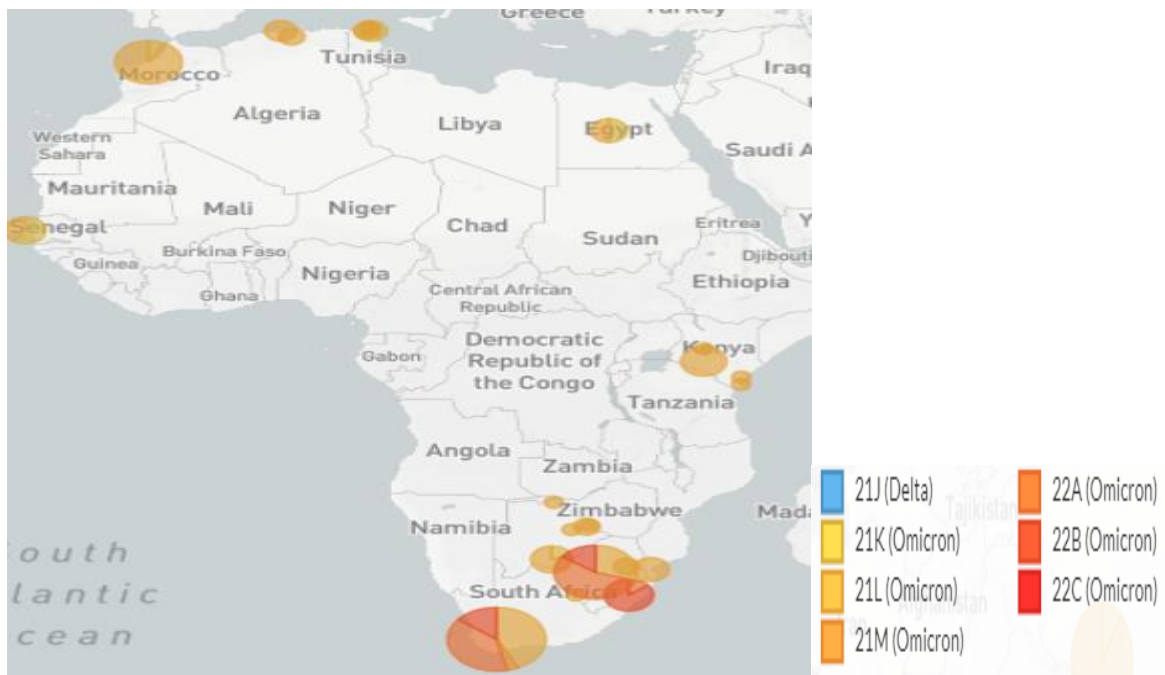


Figure 9 variants sars-cov2 en Afrique, A : du 12-2019 au 05-2022, B le 5.2022 (GISAID 2022)

Tableau 1 statistiques covid-19 dans le monde entier en Algérie et à Blida (OMS mai 2022)

	Nombre total	Nouveau cas /jours	cas / 1million de personne	Décès
Monde	523 450 042	700 168	67 318	6 271 630
Algérie	265 834	6	6 182	6 875
Blida	4 435	/	/	131

D'après la repartions géographique et le tableau des statistiques on remarque :

- une grande diminution des nombre de cas part apport aux années précédentes
- la variante omicron se transmet d'une façon très rapide
- large dominance de la variante omicron dans le monde entier

5. symptômes covid

une différence exacte des symptômes du covid-19 a ceux d'une grippe est difficile à faire , le tableau suivant issue de l'article «Omicron, Delta, cold, or flu? The most common symptoms » publier par le journal « fortune » le 12 janvier 2012 indiquant d'une façon global ces différence et le taux de leur apparitions.

Tableau 1 Différents symptômes du covid-19 , grippe et le froid (FORTUNE STAFF, s. d.)

symptômes	Delta	Omicron	Grippe	Attrapé froid
La toux	Fréquente (sèche)	Peu Fréquente	Fréquente (sèche)	Fréquente (grasse)
Nez qui coule	Fréquente	Fréquente	peu Fréquente	Fréquente
éternuement	No	Fréquente	No	Fréquente
mal de gorge	Fréquente	Fréquente	Peu Fréquente	Fréquente
essoufflement	Fréquente	No	No	No
fièvre	Fréquente	Peu Fréquente	Fréquente	Courte
sueurs nocturnes	No	peu De fois	No	No
des frissons	Fréquente	Peu Fréquente	Fréquente	No
Mal de tête	Fréquente	Fréquente	Fréquente	rare
perte d'odorat	Très fréquente	Peu Fréquente	NO	NO
Fatigue	Fréquente	Fréquente	Fréquente	Peu de fois
Durée d'apparition (jours)	4-5	2-3	1-4	1-3

- Il est a noté que les symptômes se varie d'une personne a une autres selon plusieurs critères (l'Age, le sexe, comorbidité, la population, le vaccin ...).
- On remarque que la variante omicron à moins de symptômes grave que la variante Delta
- Le tableau représente les symptômes les plus connues, mais il existe bien d'autre facteur comme la diarrhée, courbature ...

6. diagnostique sars-cov-2 :

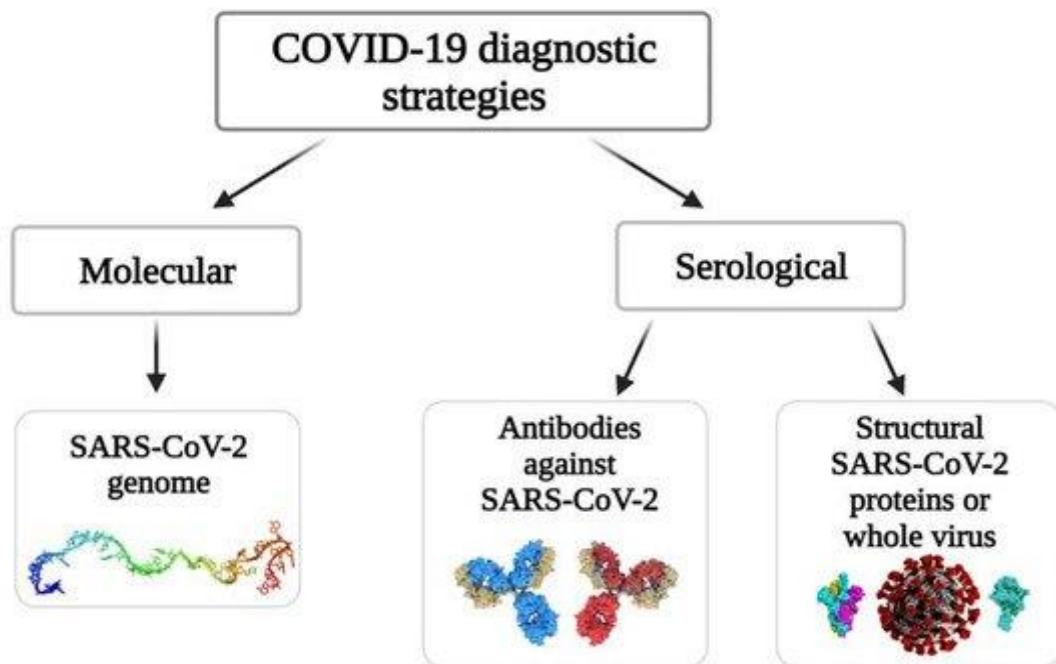


Figure 10 Stratégie du diagnostic covid-19 (Drobysh et al. 2022)

Il existe trois stratégies générales actuellement utilisées pour détecter le SRAS-CoV-2 et diagnostiquer le COVID-19, à savoir les tests moléculaires basés sur la détermination de l'ARN viral, les tests antigéniques basés sur la détermination des protéines virales et les tests d'anticorps basés sur la détermination d'anticorps spécifiques contre les protéines virales.

➤ **Comment identifier le variant « omicron » ?**

Comme lors de l'identification du variant Alpha au Royaume-Uni, l'équipe du laboratoire en Afrique du Sud a détecté un nombre croissant d'échantillons avec un échec d'amplification de la cible du gène S en utilisant le Thermo Fisher TaqPath COVID-19 Combo Kit comme test de diagnostic. Le séquençage de ces échantillons a permis d'identifier le nouveau variant B.1.1.529 du SRAS-CoV-2, désormais appelé Omicron, qui présente plus de 30 mutations sur le gène S. La prévalence relative de chaque mutation dans les séquences d'Omicron n'est pas encore connue car les données de séquençage sont actuellement limitées.

Cette note technique développée par l'OMS a pour but de fournir des informations aux États membres sur l'utilisation des tests PCR 7 comme outil de criblage du variant Omicron. Le criblage par PCR peut guider la sélection d'échantillons à séquencer pour permettre l'identification définitive d'Omicron. Des tests de criblage par PCR spécifiques à ce variant sont rapidement produits par les fabricants

Globalement, le VOC Delta a été identifié comme circulant largement dans les pays africains.

- Dans un premier temps, il est suggéré d'examiner les VOC identifiés au cours des 4 à 8 dernières semaines dans votre pays, si cette information est disponible, afin d'identifier le VOC prévalent en circulation.

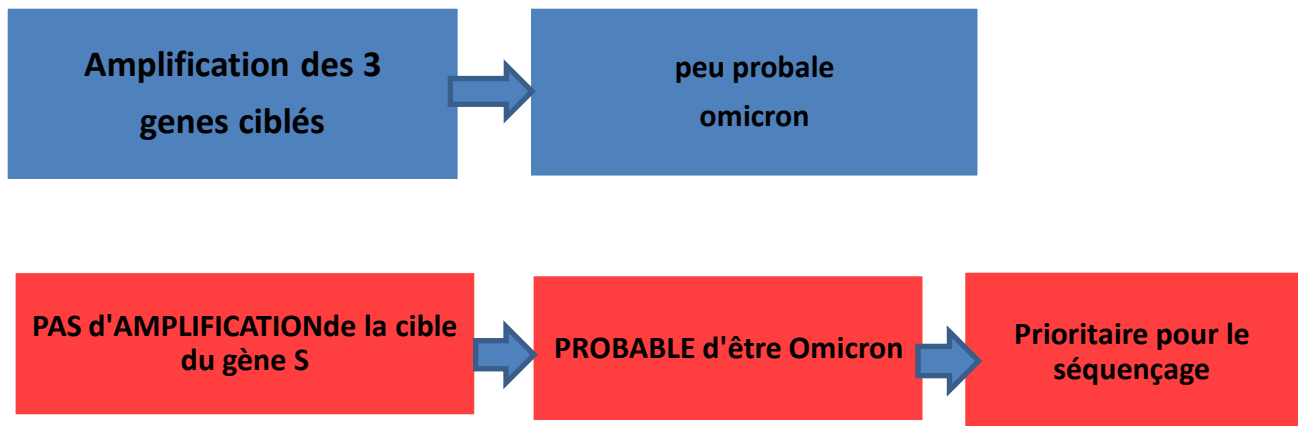
- Dans un deuxième temps, identifiez les tests de diagnostic utilisés pour détecter le SRAS-CoV-2 et les réactifs de PCR de criblage des variantes que vous possédez

Le criblage de la présence probable du VOC Omicron peut être effectué en identifiant l'échec d'amplification de la cible du gène S (S Gene Target Failure - SGTF) ou en utilisant des tests qui ciblent des mutations spécifiques.

➤ **CRIBLAGE À PARTIR DE L'ÉCHEC D'AMPLIFICATION DE LA CIBLE DU GÈNE S (SGTF) :**

Si vous utilisez le Thermo Fisher TaqPath COVID-19 Combo Kit (test PCR avec 3 cibles) comme réactif de diagnostic, interprétez

Les résultats ci-dessous comme une indication qu'un échantillon pourrait être un variant Omicron.



Pour ABI7500, Quant Studio 5 ou autre. De plus, ces machines doivent être calibrées pour les deux fluorophores qui sont la propriété de Thermo Fisher

. Tous les virus Omicron ne présentent pas la mutation DEL69/70. Au 02/12/2021, 5% des séquences Omicron dans GISAID ne présentaient pas la mutation DEL69/70. Il est donc important de surveiller l'évolution épidémiologique de la délétion si elle est utilisée comme méthode de criblage.

Pouvoir utiliser ce test, vous devez utiliser un thermo cycleur ABI, par exemple

CRIBLAGE PAR PCR POUR EXCLURE DELTA

Si Delta est/était le variant prédominant circulant dans votre pays, utilisez un test de criblage PCR (par exemple EscapePlex ou un test de mutation unique L452R) capable d'identifier le variant Delta. Analyser les résultats des tests en considérant les interprétations ci-dessous :

- Un résultat POSITIF identifiera le variant Delta
- Un résultat NEGATIF exclura Delta mais peut indiquer la présence d'Omicron ou d'un autre variant. Ce spécimen doit être prioritaire pour le séquençage

Le séquençage ciblé du gène S par NGS ou séquençage Sanger, ou le séquençage du génome entier par NGS fournira une Identification précise d'Omicron (oms 2022)

7. prévention et conseil covid-19 :

Recommandations de mesures à prendre par les pays :

Omicron ayant été désigné variant préoccupant, l'OMS recommande aux pays de prendre plusieurs mesures, notamment :

- *le renforcement de la surveillance et du séquençage des cas (OMS.2022).
- *la mise en commun des séquences du génome sur des bases de données accessibles au public (GISAID).
- * le signalement à l'OMS des premiers cas ou foyers épidémiques.
- *la réalisation d'enquêtes sur le terrain et d'évaluations en laboratoire pour mieux comprendre si Omicron présente des caractéristiques de transmission ou de virulence différentes ou s'il a une incidence sur l'efficacité des vaccins, des traitements, des produits de diagnostic ou des mesures sociales et de santé publique. (oms.2022)

En outre, il est primordial que les inégalités en matière d'accès aux vaccins contre la COVID-19 soient corrigées de toute urgence afin de garantir que les groupes vulnérables, y compris les agents de santé et les personnes âgées, reçoivent partout dans le monde leurs deux premières doses, et qu'ils bénéficient d'un accès équitable au traitement et au diagnostic. (oms.2022)

Recommandations de mesures à prendre par les personnes :

- *ouvrir les fenêtres pour améliorer la ventilation. (oms.2022)
- *éviter les espaces mal ventilés ou bondés. (oms.2022)
- *se laver les mains.(oms.2022)
- * tousser ou à éternuer dans son coude ou dans un mouchoir. (oms.2022)
- * se faire vacciner et ne pas dépasser la durée entre les doses. (oms.2022)
- *une troisième dose de vaccin surtout pour lutter contre le variant omicron(oms.2022).
- *consulté les informations se disponibles sur les plateformes de médias numériques et les réseaux sociaux de l'OMS. (oms.2022)
- *isolement en cas de symptômes (oms.2022)



B

Partie

Expérimentale.



1. Introduction

La bio-informatique est un domaine interdisciplinaire qui développe des méthodes et des outils « logiciels » pour comprendre les données biologiques, en particulier lorsque les données sont volumineuses et complexes. En tant que domaine scientifique interdisciplinaire, la bio-informatique combine la biologie, la chimie, la physique, l'informatique, l'ingénierie de l'information, les mathématiques et les statistiques pour analyser et interpréter les données biologiques. Ce large et multiple choix d'analyse nous a permis pour notre étude de passer de l'analyse physico-chimique, fonctionnelle à l'analyse structural tridimensionnelle et même à l'analyse phylogénétique. Le but de ces analyses est de comprendre le changement et l'impact mutationnel de la protéine Spike causé par : A67V, T95I, S371L, S373P, S375F, N404K, G446S, S477N, E484A, Q493R, Q498R, N501Y, Y505H, N856K, N969K d'un point de vue (physico-chimique, structural, fonctionnel et phylogénétique), en utilisant plusieurs et différents logiciels

Définition des outils bio-informatique utilisés:

- ❖ **Gene Bank** : est une la base données de séquences génétiques gérée par le National Institutes of Health. C'est une collection annotée de l'ensemble des séquences ADN publiquement accessibles. Elle fait partie de l'International Nucleotide Sequence Database Collection aux côtés de la DNA Databank of Japan et de l'European Molecular Biology Laboratory, et échange quotidiennement des données avec ces deux organismes.
- ❖ **NCBI** : Centre National de l'information Biotechnologique, fait progresser la science et la santé en donnant accès à des informations biomédicales, génomiques et des outils bio-informatique
- ❖ **MEGA** : logiciel de l'évolution Moléculaire et l'analyse Génétique ,c'est une application de bureau conçue pour l'analyse comparative des séquences des gènes homologues provenant de familles multigéniques ou de différentes espèces, avec un accent particulier sur la déduction des relations évolutives et

des modèles d'évolution de l'ADN et des protéines. En plus des outils d'analyse statistique des données, MEGA fournit de nombreuses fonctionnalités pratiques pour l'assemblage d'ensembles des données des séquences à partir de fichiers ou de référentiels Web, et il comprend des outils pour la présentation visuelle des résultats obtenus sous la forme d'arbres phylogénétiques interactifs et matrices de distance évolutives

- ❖ **Clustal Omega** : Clustal Omega est un nouveau programme d'alignement de séquences multiples qui utilise des arbres guides ensemenés et des techniques de profil-profil HMM pour générer des alignements entre trois séquences ou plus. Pour l'alignement de deux séquences, veuillez plutôt utiliser nos outils d'alignement de séquences par paires.
- ❖ **EXPASY** : SIB Windows serveur, qui se concentre sur les protéines et la protéomique, et donne accès à une variété de bases de données et d'outils d'analyse. L'un des atouts majeurs d'ExpASy est le haut degré d'intégration et d'inter connectivité qu'il établit entre toutes les bases de données et services disponibles.
- ❖ **SMS** : est une collection de programmes JavaScript permettant de générer, de formater et d'analyser de courtes séquences d'ADN et de protéines. Il est couramment utilisé par les biologistes moléculaires, pour l'enseignement et pour les tests de programmes et d'algorithmes.
- ❖ **DeepDDG** : un serveur web prédit le changement de stabilité de mutations ponctuelles de protéines à l'aide de réseaux neuronaux
- ❖ **Dynamut** : un serveur web Met en œuvre deux approches de mode normal distinctes et bien établies, qui peuvent être utilisées pour analyser et visualiser la dynamique des protéines en échantillonnant les conformations et évaluer l'impact des mutations sur la dynamique et la stabilité des protéines résultant du changement d'entropie vibrationnelle.
- ❖ **SNAP 2** : un serveur web permet la Prédiction des effets fonctionnels des variants des séquences
- ❖ **GISAID** : signifie (Global Initiative on Sharing Avian Influenza Data) , offre un accès libre aux données génomiques des virus de la grippe et du coronavirus responsable de la pandémie de COVID-19.

- ❖ **PDB9** : signifie (protein data base) est une collection mondiale de données sur la structure tridimensionnelle de macromolécules biologiques : protéines, essentiellement, et acides nucléiques.
- ❖ **Pymol** : est un logiciel libre de visualisation de structures chimiques en 3D créé par Warren DeLano.
- ❖ **HDOCK** : un serveur web pour Amarrage protéine-protéine et protéine-ADN/ARN

Matériels et méthodes :

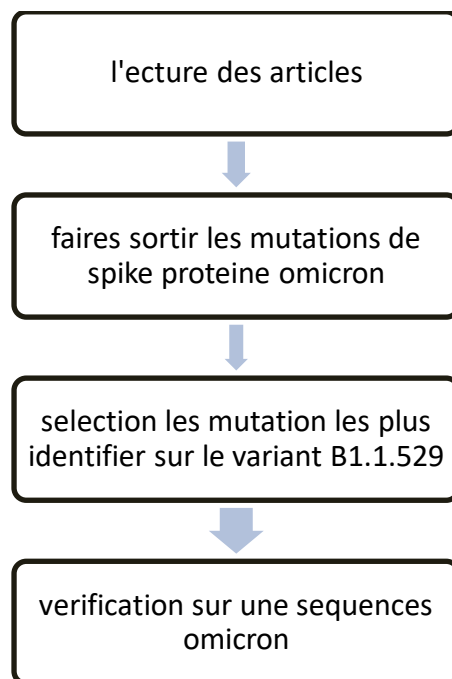
Cette étude est réalisée par un pc Samsung, Windows 7 professionnel, 4.0G de Ram avec système d'exploitation 32 bits

Pour utiliser le logiciel pymol et HDOCK on a travaillé avec un pc plus performant : pc Asus, Windows 10, 8.0G de Ram avec un système d'exploitation 64 bits.

A. Choix des mutations :

Pour cette étude le choix des mutations est effectué à partir des analyses d'articles selon des critères précis, les plus importantes :

- Les mutations les plus présentes chez le variant omicron.
- Les mutations-causant un changement de fonction ou de structure



Les mutations pour mon étude : A67V, T95I, S371L, S373P, S375F, N404K, G446S, S477N, E484A, Q493R, Q498R, N501Y, Y505H, N856K, N969K

B. Données et Etapes :

- La séquence de la protéine de pointe de référence du SRAS-CoV-2 a été obtenue de GISAID « WIV04 » et le premier génome complet d'Omicron de GISAID avec le numéro d'accession EPI_ISL_6640916.

- le génome omicron a été annoté et transcrit par Expasy, après un alignement par paires des séquences de protéines de pointe a ensuite été effectué avec Clustal Omega pour détecter la séquence protéique (Spike protéine) sur le génome EPI_ISL_6640916.

➤ On consulte page GISAID, on clique sur search après entre le numéro d'accèsion

Endemic coronavirus causing COVID-19

previously unknown human coronavirus (hCoV-19) was first detected in late 2019 in patients in the city of Wuhan, who suffered from respiratory illnesses including atypical pneumonia, an illness that has since become known as coronavirus disease (COVID-19). The coronavirus originated from an animal host and is closely related to the virus responsible for the Severe Acute Respiratory Syndrome coronavirus (SARS).

On 10. January 2020, the first virus genomes and associated data were publicly shared via GISAID. The World Health Organization announced on 11. March 2020 the first coronavirus pandemic. As the

Virus name:	hCoV-19/Botswana/R40B59_BHP_3321001248/2021
Accession ID:	EPI_ISL_6640916
Type:	betacoronavirus
Clade:	GRA
Pango Lineage:	BA.1.17 (Pango v.4.0.6 PANGO-v1.9), Omicron (BA.1-like) (Scorpio)
AA Substitutions:	Spike A67V, Spike D614G, Spike D796Y, Spike E484A, Spike G142D, Spike G339D, Spike G446S, Spike G496S, Spike H69del, Spike H655Y, Spike ins214EPE, Spike K417N, Spike L212I, Spike N211del, Spike N440K, Spike N501Y, Spike N679K, Spike N764K, Spike N856K, Spike N969K, Spike P681H, Spike Q493R, Spike Q498R, Spike Q954H, Spike S371L, Spike S373P, Spike S375F, Spike S477N, Spike T95I, Spike T478K, Spike T547K, Spike V70del, Spike V143del, Spike Y144del, Spike Y145del, Spike Y505H, E T9I, M A63T, M D3G, M Q19E, N E31del, N G204R, N P13L, N R32del, N R203K, N S33del, NSP3 A1892T, NSP3 K38R, NSP3 L1266I, NSP3 S1265del, NSP3 V1069I, NSP4 T492I, NSP5 P132H, NSP6 G107del, NSP6 I189V, NSP6 L105del, NSP6 S106del, NSP12 P323L, NSP14 I42V
Variant:	VOC Omicron GRA (B.1.1.529+BA.*) first detected in Botswana/Hong Kong/South Africa
Passage details/history:	Original
Sample information	
Collection date:	2021-11-11
Location:	Africa / Botswana / South East / Greater Gaborone / Gaborone
Host:	Human
Additional location information:	
Gender:	Female
Patient age:	57
Patient status:	Live
Specimen source:	Nasopharyngeal and oropharyngeal swab
Additional host information:	
Sampling strategy:	
Outbreak:	
Last vaccinated:	
Treatment:	
Sequencing technology:	Nanopore MinION
Assembly method:	
Coverage:	
Comment:	⚠ Stretches of NNNs (1.31% of overall sequence). Insertion of 6 nucleotides when compared to the reference WIV04

Figure 11 Exemple du résultat de la recherche sur GISAID (gisaid.org)

On utilise le logiciel Expasy pour obtenir la séquence protéique parce que GISAID nous donne la séquence en séquence nucléotidique

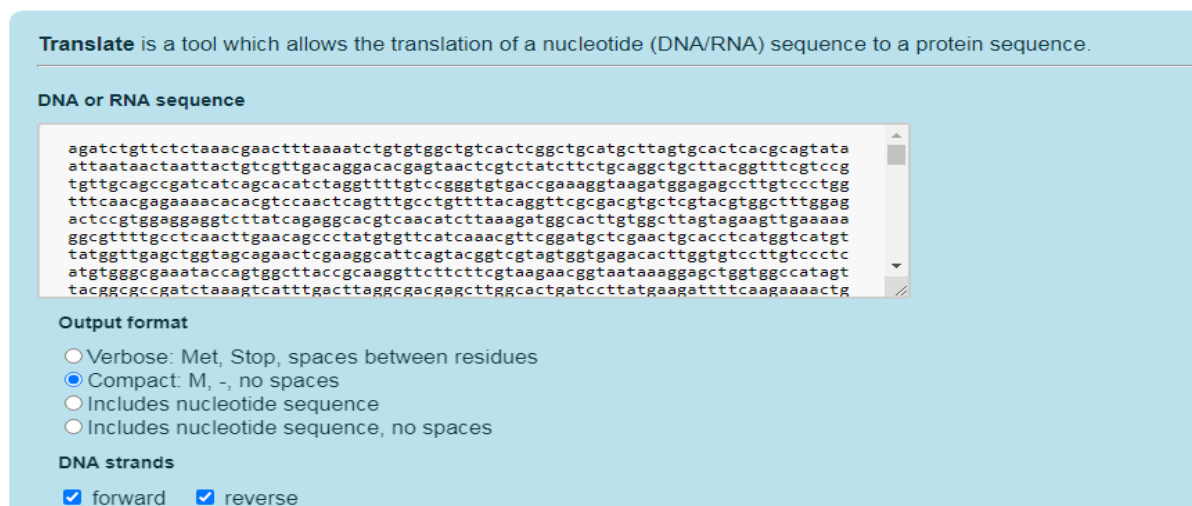


Figure 12 Exemple de la transcription par Expasy (expasy.org)

- On enregistre la séquence protéique du génome entier

Donnée arbre phylogénétique :

Les séquences de l'arbre phylogénétique sont issues de GISAID :

Variant	N° GISAID
Sequence Reference	WIV04 was isolated by The Wuhan Institute
Alpha (3.1.1.7-like)	EPI_ISL_675651511
Beta (13.1.351-like)	EPI_ISL_541654011
Delta (13.1.617.2-like)	EPI_ISL_133676820
Omicron (BA.1-like)	EPI_ISL_664091611
Gamma (P.1-like)	EPI_151_6228367t1

Un alignement du génome de chaque variant avec la séquence référence (Spike protéine) est obligatoire pour identifier la séquence précise de la protéine Spike pour chaque variant.

Registered Users EpiFlu™ EpiCoV™ EpiPox™ My profile

EpiCoV™ | Search | Downloads | Upload

Pandemic coronavirus causing COVID-19

A previously unknown human coronavirus (hCoV-19) was first detected in late 2019 in patients in the city of Wuhan, who suffered from respiratory illnesses including atypical pneumonia, an illness that has become known as coronavirus disease (COVID-19). The coronavirus originated from an animal host and is closely related to the virus responsible for the Severe Acute Respiratory Syndrome coronavirus (SARS).

On 10 January 2020, the first virus genomes and associated data were publicly shared via GISAID. The World Health Organization announced on 11 March 2020 the first coronavirus pandemic. As the pandemic progresses, scientists from around the globe are tracking the virus and its genome sequences to ensure optimal virus diagnostic tests, to track and trace the ongoing outbreak and to identify potential intervention options. Several analyses to assist with these efforts are offered here, including sequence alignments, diagnostic primer and probe coordinates, 3D protein models, drug targets, phylogenetic trees and many more.

Figure 13 Page de face GISAID (gisaid.org)

- On clique sur le rectangle jaune qui représente la Spike protéine du génome référence

WIV04 is representative of and identical to the early outbreak sequences. WIV04 was isolated by the Wuhan Institute of Virology from a bronchoalveolar lavage fluid (BALF) collected at the Wuhan Jinyintan Hospital in Hubei Province on 30th December 2019 from a symptomatic patient, a retailer working at the Huanan Seafood Wholesale Market.

BALF was used for RNA extraction and metagenomic next-generation sequencing (NGS). The consensus sequence was obtained by de-novo assembly (Zheng-Li Shi et al 2020).

Download FASTA

Evidence in the Literature

1. Denissov, A. D., Williamson, M. K., Lewis, S., et al. Characterization of the transcriptome and proteome of SARS-CoV-2 reveal passage-induced in-frame deletion of the furin-like cleavage site in the spike glycoprotein. *Genome Med* 12, 68 (2020) DOI: 10.1186/s13073-020-00763-0

Nucleotide Sequence

Click sequence to download FASTA

```

ATGTTTGTGTTTCTTGTTTATGGCCACTAGTCTCTAGTCAAGTGTAAATCTTCAACACAGAACTCAATACCCTGTCATACACATAATCTTTTACACAGTGGTGTATACCCCTGACAAAC
CTCAGTTTTACATTCAACTCAGGACTTGTCTTACCTTCTTTTCCAATGTTACTTGGTCCATGCTATACATGCTCTGGGACCAATGGTACTAAGAGGTTTGATAACCCCTGCTTACCATTT
TTTATTTTGGCTTCCACTGAGAGTCTAACATAATAAGAGGCTGGATTTTGGTACTACTTTAGATTCGAAGACCCAGTCCCTACTTATTGTAAATACCGCTACTAATGTTGTTTAAAGTCTC
TTTGTAAATGATCCGATTTTGGGTGTTTATACACACAAACACAAAGTTGGATGGAAAGTGAAGTTCAGAGTTTATCTAGTGGGAAATATTCACATTTTGGACATGATGCTCTCAGCTTTT
TGAAGGAAACAGGGTAATTTCAAATACTTAGGGAATTTGTGTTAAGAATATTGATGGTATTTTAAATAATATTTAAGCACACGCCCTATTAAATTTAGTGGCTGATCCCTCAGGGTTTT
AACCATGGTAGATTTGCCAATAGGATTTAACATCACTAGGTTTCAAACTTTACTTGTCTTACATAGAGATTTATGGACTCTGGTGAATCTCTTCAAGTGGACACCTGGTCTCAGCTTT
TATCTTCAACCTAGGACTTTTCTATTAAAATAAATGAAATGAAACATTACAGATGCTGTAGACTGTGCACCTTGAACCTCTCAGAAACAAAGTGTACGTTGAAATCTTCACTGTAGAAAT

```

Figure 14 Séquence nucléotidique de la Spike protéine (gisaid.org)

On télécharge la séquence référence, ainsi que le génome de tous nos variants utilisés, on lance un alignement par « clustal.omega » des génomes avec la séquence omicron référence pour définir la séquence Spike dans chaque variant

Dans une nouvelle page web on lance le serveur « sms » et on clique sur « group dna »



Sequence Manipulation Suite: Version 2

- Formal Conversion
 - Combine FASTA
 - EMBL to FASTA
 - EMBL Feature Extractor
 - EMBL Trans Extractor
 - Filter DNA
 - Filter Protein
 - GenBank to FASTA
 - GenBank Feature Extractor
 - GenBank Trans Extractor
 - One to Three
 - Range Extractor DNA
 - Range Extractor Protein
 - Reverse Complement
 - Split Codons
 - Split FASTA
 - Three to One
 - Window Extractor DNA
 - Window Extractor Protein
- Sequence Analysis
 - Codon Plot
 - Codon Usage
 - CpG Islands
 - DNA Molecular Weight
 - RNA Pattern Find

- The Sequence Manipulation Suite is a collection of JavaScript programs for generating molecular biologists, for teaching, and for program and algorithm testing.
- See the about the Sequence Manipulation Suite page for more information about ind
- You can easily mirror the Sequence Manipulation Suite on your own web site, or you
- This version of the Sequence Manipulation Suite represents a complete re-write of th
- previous version of the Sequence Manipulation Suite can still be accessed.
- Send questions and comments to stothard@ualberta.ca.

- Protein Molecular Weight
- Protein Pattern Find
- Protein Stats
- Restriction Digest
- Restriction Summary
- Reverse Translate
- Translate

Sequence Figures

- Color Align Conservation
- Color Align Properties
- Group DNA
- Group Protein
- Primer Map

Figure 15 Manipulation des séquences par sms (bioinformatics.org/sms2/)

- On Place à chaque fois le génome de notre variant, et on clique sur « submit »
- La numérotation de tout le génome, nous facilitera l'identification de la séquence (Spike protéine)



Sequence Manipulation Suite:

Group DNA

Group DNA adjusts the spacing of DNA sequences and adds numbering. You can specify the group size (the number of bases per group), as well as the r output of this program can serve as a convenient reference, since the numbering and spacing allows you to quickly locate specific bases.

Paste the raw sequence or one or more FASTA sequences into the text area below. Input limit is 100,000,000 characters.

Submit Clear Reset

- Use a group size of and display bases per line.
- Don't show the reverse strand.
- Show numbering to the left of the sequence.
- Start the numbering at

*This page requires JavaScript. See browser compatibility.
*You can mirror this page or use it off-line.

Figure 16 l'outil de Numérotation du génome par sms (bioinformatics.org/sms2)

Group DNA results

Results for 29763 residue sequence "Untitled" starting "AGATCTGTTC"

```

1  AGATCTGTTC  TCTAAACGAA  CTTTAAATC  TGTGTGGCTG  TCACTCGGCT  GCATGCTTAG  TGCACACAG  CAGI
81  AATAACTAAT  TACTGTCTGT  GACAGGACAC  GAGTAACTCG  TCTATCTTCT  GCAGGCTGCT  TACGSTTTCG  TCCG
161  AGCCGATCAT  CAGCACATCT  AGGTTTGTGC  CGGGTGTGAC  CGAAAGGTAA  GATGGAGAGC  CTTGTCCCTG  GTTI
241  GAAAACACAC  GTCCAACTCA  GTTTCCTGTG  TTTACAGGTT  CCGACCTGCG  TCGTACCTGG  CTTTGGAGAC  TCCG
321  AGGTCTTATC  AGAGGCACGT  CAACATCTTA  AAGATGGCAC  TTGTGGCTTA  GTAGAAAGTTG  AAAAAGGCGT  TTTC
401  CTGGAACAGC  CCTATGTGTT  CATCAAACGT  TCGGRTGTCT  GRACTGCACC  TCAATGGTCA  GTTATGGTTG  AGCI
481  AGAARTCGAA  GGCATTCAAT  ACGGTCGTAG  TGGTGAGACA  CTTGGTGTCC  TTGTCCCTCA  TGTGGGGGAA  ATAC
561  CTTACCGCAA  GGTTCCTCTT  CGTAAGAACG  GTAATAAAGG  AGCTGGTGGC  CATAGTTACG  GCGCCGATCT  AAAG
641  GACTTAGGCG  ACGAGCTTGG  CACTGATCCT  TATGAAGATT  TTCAAGAAAA  CTGGAACACT  AAACATAGCA  GTGG
721  CCGTGAACTC  ATGCGTGAGC  TTAACGGGAG  GGCATACACT  CGCTATGTCT  ATAACAACCT  CTGTGGCCCT  GATG
801  CTCTTAGAGT  CATTAAGAGC  CTCTAGACAC  GTGCTGGTRA  AGCTTCATGC  ACTTGTCTGT  AACAACCTGA  CTTT
881  ACTAAGAGGG  GTGTATACTG  CTGCCGTGAA  CATGAGCATG  AAATTGCTTG  GTACACGGAA  CGTTCGTAAA  AGAG
961  ATTGCAGACA  CCTTTTGAAA  TTAATTTGGC  AAAGAAATTT  GACACCTTCA  ATGGGGAATG  TCCAAATTTT  GTAI
1041  TAATTTCCAT  AATCAAGACT  ATTCAACCAA  GGGTTGAAA  GAAAAGGCTT  GATGGCTTTA  TGGGTAGAA  TCGA
1121  TATCCAGTTC  CGTCACCAA  TGAATGCRAC  CAAATGTGCC  TTTCAACTCT  CATGAAGTGT  GATCATTGTG  GTGA
1201  ATGGCAGAGC  GGCATTTTGG  TTAAGCCAC  TTGCGAATTT  TGTGGCAGTG  AGAATTTGAC  TAAAGAAGGT  GCCA
1281  GTGGTTACTT  ACCCCAAAAT  GCTGTTGTTA  AAATTTATTG  TCCAGCATGT  CACAATTCAG  AAGTAGGACC  TGAG
1361  CTTGCCGAAT  ACCATAATGA  ATCTGGCTTG  AAAAOCATTG  TTCGTAAGGG  TGGTGGCACT  ATTGCCCTTG  GAGG
1441  GTTCTCTTAT  GTTGGTGGCC  ATAACAAGTG  TGCCATTGGG  GTTCCACGTG  CTAGCGCTAA  CATAGGTTGT  AACG
1521  GTTCTCTTAT  GTTGGTGGCC  ATAACAAGTG  TGCCATTGGG  GTTCCACGTG  CTAGCGCTAA  CATAGGTTGT  AACG

```

Figure 17 Résultat numérotation du génome par sms (bioinformatics.org/sms2)

- On copiera pour chaque variant la séquence de la protéine Spike et on l'insert dans un fichier bloc note de notre ordinateur
- On clique sur enregistré
- **Réalisation de l'arbre phylogénétique par Mega.11 :**

L'alignement des séquences est effectuer par (MEGA 11) est un outil intégré pour l'alignement automatique et manuel des séquences, on d'abord télécharger nos séquences en format « fasta » après on a sélectionné « protéine séquence alignement » et on introduit nos séquence

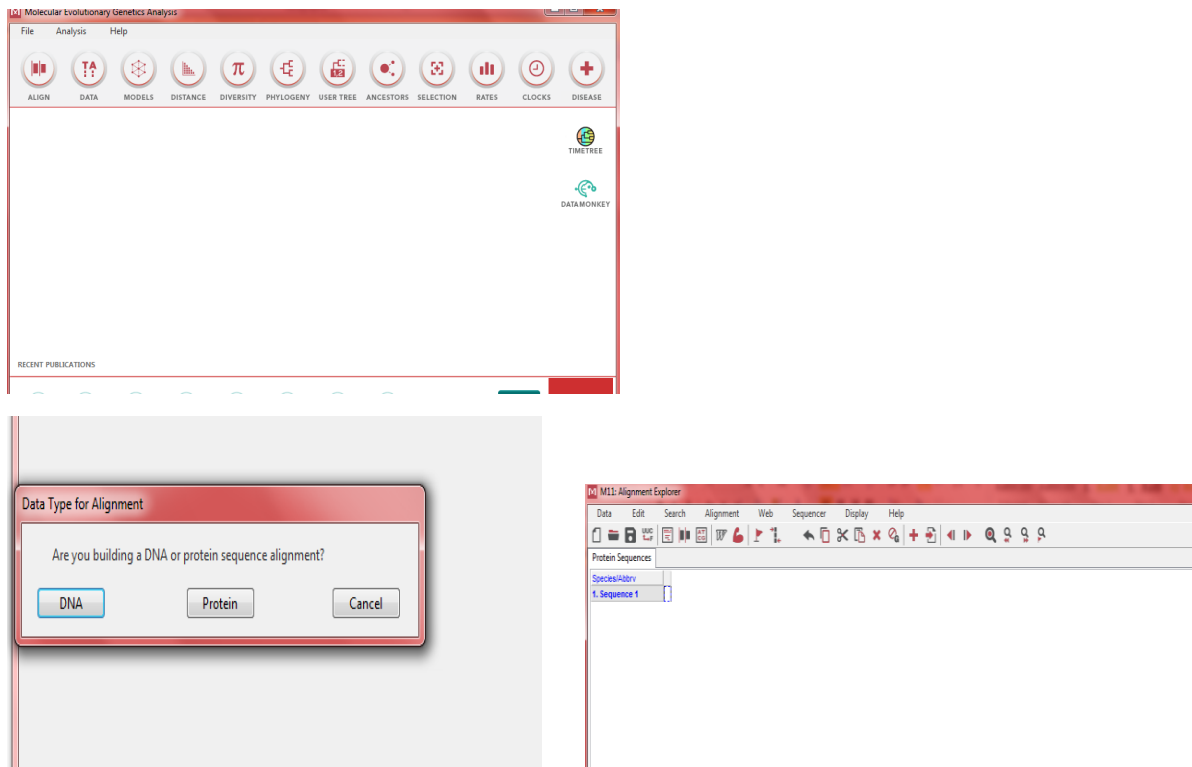


Figure 18 Etapes d'introduire les séquences sur Mega 11 (Megasoftware)

- Placez toute vos séquences.

L'alignement des séquences nous a permis de bien visionné et comparé chaque emplacement d'AA avec les autres variant et aussi de savoir qu'elle type de mutations est apparue

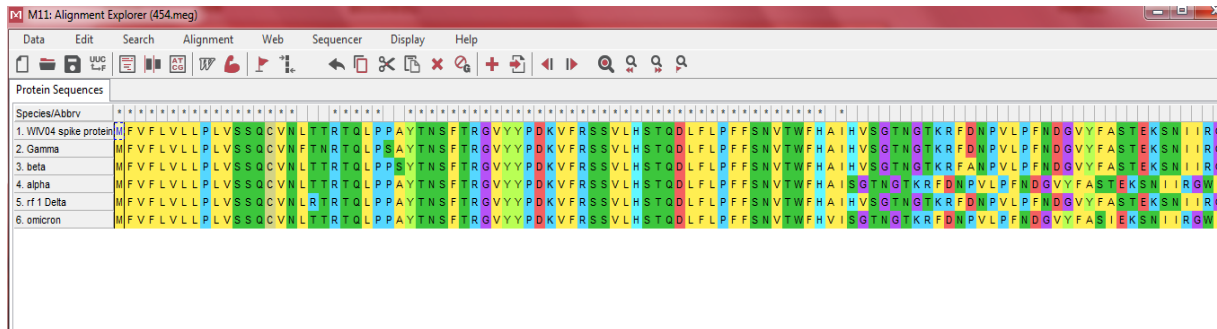


Figure 19 Résultat alignement par mega 11 (Meg software)

- On Clique sur « data » puis sur (export alignement) et on choisit (Mega format) notre alignement sera enregistré en format mega.
- On revient sur la page d'accueil MEGA et on clique sur « phylogeny »

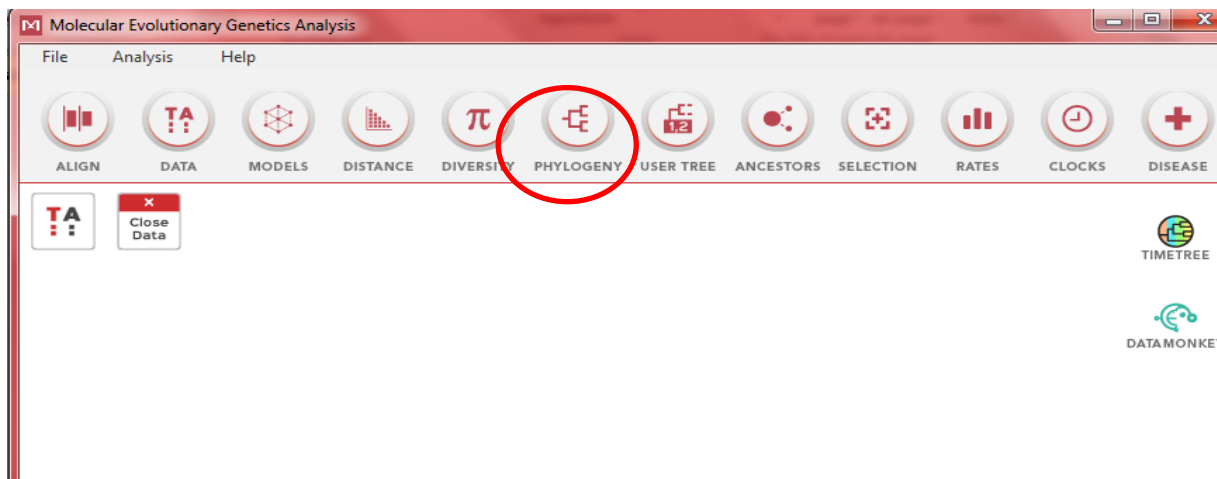


Figure 20 Etape de la construction de l'arbre Phylogénétique par mega11 (megasoftware)

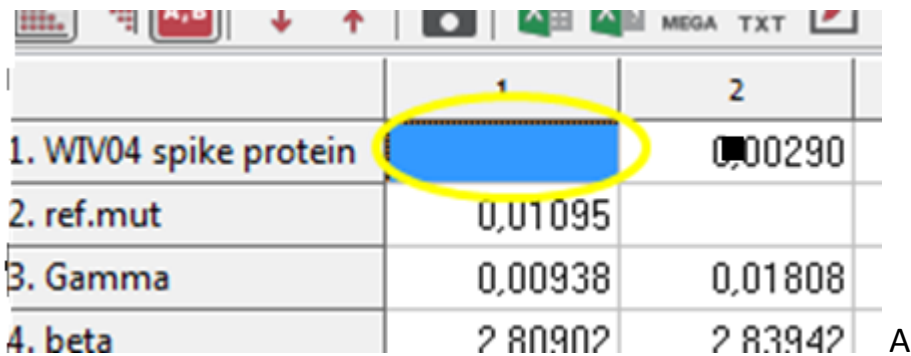
- On Clique sur (construction/test UPGMA tree.)
- On Choisit les reconstructions adaptées à notre recherche
- On Clique sur « OK »

Pour obtenir le tableau de « distance par paires » :

- On ouvre le logiciel « mega 11 »
- On choisit notre alignement déjà fait
- Et on clique sure distance

- On clique sur « compute pairwise distance »
- Et choisit les critères soulignés
- On clique sur « ok »
- Une fois les résultats obtenus (tableau dans la partie « résultats et discussion ») on clique sur le carré rouge en haut de la page pour faciliter la lecture du tableau

Exemple pour la première Colonne en bleu.



	1	2
1. WIV04 spike protein	0,00290	
2. ref.mut	0,01095	
3. Gamma	0,00938	0,01808
4. beta	2 80902	2 83942

On obtient on bas de la page :

[1,1] (WIV04 spike protein-WIV04 spike protein) / Amino: JTT matrix-based B

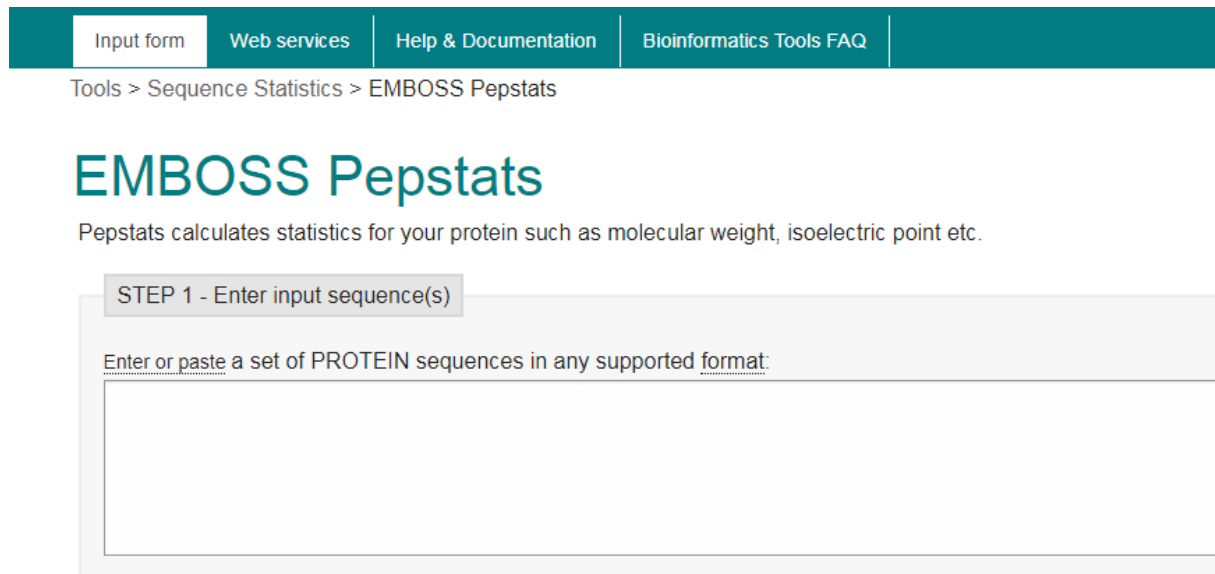
Figure 21 A et B explication de la lecture des Colonne (megasoftware)

➤ **Analyse physicochimique :**

La protéine de pointe (Spike protéine) des variantes de référence (avec et sans les mutations étudiées) et omicron, a été soumise à des analyses de séquence préliminaires pour distinguer leurs différences physicochimiques. La composition en acides aminés, le poids moléculaire, la distribution des résidus chargés, l'hydrophobie, l'indice aliphatique, l'indice d'instabilité et quelques autres paramètres sont calculés principalement à l'aide du serveur en ligne EMBOSS Pepstats, on a travaillé avec la séquence référence de GISAID « WIV04 » (avec et sans nos 15 mutations), et la séquence référence omicron de GISAID (EPI_ISL_6640916) (avec et sans nos 15 mutations)

On lance le serveur emboss-pepstats partir de Google

On met nos séquence dans le carré réservé et on clique sur submit, ici c'est pour obtenir des calculs statistique, le poids moléculaire ainsi que le point isoélectrique



Input form | Web services | Help & Documentation | Bioinformatics Tools FAQ

Tools > Sequence Statistics > EMBOSSE Pepstats

EMBOSS Pepstats

Pepstats calculates statistics for your protein such as molecular weight, isoelectric point etc.

STEP 1 - Enter input sequence(s)

Enter or paste a set of PROTEIN sequences in any supported format:

Figure 22 Exemple figure de face de l'utilisation pepstats sur emboss (ebi.ac.uk)

On enregistre nos résultats et en reviens en arrière pour cliquer sur (sequence statistics).

On obtient un large choix d'analyse à effectuer

Pour notre étude on choisit :

*Pepwindow

*Pepinfo

*Pepstats

➤ **Analyse des propriétés fonctionnelles :**

Afin de déterminer l'effet des mutations sur la stabilité des protéines, Dyanmut et DeepDDG ont été utilisés, où Dyanmut a utilisé une analyse en mode normal et des représentations graphiques des structures protéiques, et DeepDDG a utilisé des réseaux de neurones pour prédire l'effet des mutations sur la stabilité des protéines.

On a utilisé la protéine Spike avec le pdb : 7N1U

On a utilisé le SNAP2 pour évaluer l'impact des mutations sur la fonction.

Analyse par Dynamut :

-on tape « Dyanmut» et on clique sur :

<http://biosig.unimelb.edu.au/dyna...> · Traduire cette page

DynaMut - Biosig Lab - University of Melbourne

Figure 23 Le Lien utilisé de DynaMUT

On clique sur (mutation effect prediction)

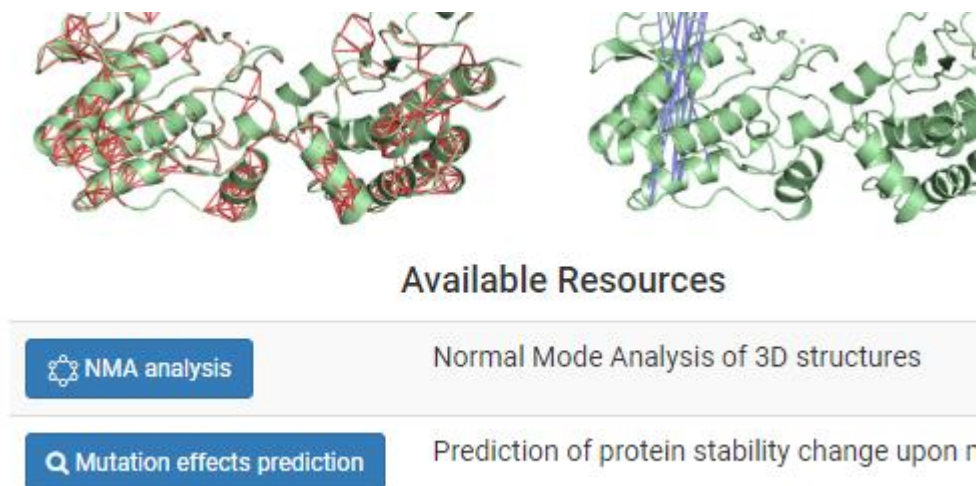


Figure 24 Le choix d'analyse proposer par DynaMut

-on choisit « single mutations » pour traiter chaque mutation appart

Single Mutation

Provide a wild-type structure*
Submit a molecule in [PDB format](#).

Wild-type (Ex.: 1U46) OR **PDB Accession**

Choisir un fichier
Aucun fichier choisi

7N1U

Mutation details

Mutation*

E484A

Chain*

A

Email ⓘ (optional)

your@email.com

▶ Run prediction

Figure 25 Exemple d'introduction des mutations sur DynaMut

-on a travaillé avec le code PDB : 7N1U de la séquence Spike protéine

-on met à chaque fois une mutation et on clique sur submit

(Il faut vérifier que la séquence utilisé ne contient pas déjà cette mutation)

Pour les mutations on écrit la lettre de l'AA non muté +numéro dans la séquence +AA muté

-on clique sur « run prediction »

Le résultat de l'exemple utilisé (mutation E484A Chain A) et ainsi :

Info! Your results will be available for **7 days** after the job is processed.

Run another prediction

Submission details

Wild-type: **GLU**

Position: **484**

Mutant: **ALA**

Chain: **A**

$\Delta\Delta G$ Predictions

Interatomic Interactions

Deformation and Fluctuation Analysis

Prediction Outcome

$\Delta\Delta G$: **-0.117 kcal/mol (Destabilizing)**

NMA Based Predictions

$\Delta\Delta G$ ENCoM: **-0.061 kcal/mol (Destabilizing)**

Other Structure-Based Predictions

$\Delta\Delta G$ mCSM: **-0.274 kcal/mol (Destabilizing)**

$\Delta\Delta G$ SDM: **0.150 kcal/mol (Stabilizing)**

$\Delta\Delta G$ DUET: **-0.022 kcal/mol (Destabilizing)**

Figure 26 Exemple du résultat donné par DynaMut

Suivi par une représentation graphique très utile pour la compréhension, les résidus colorés en bleu indiquant une rigidification de la structure

Δ Vibrational Entropy Energy Between Wild-Type and Mutant

$\Delta\Delta S_{\text{vib}}$ ENCoM: **0.076 kcal.mol⁻¹.K⁻¹ (Increase of molecule flexibility)**

Δ Vibrational Entropy Energy | Visual representation



Amino acids colored according to the vibrational entropy change upon mutation. **BLUE** represents a rigidification of the structure and **RED** a gain in flexibility

Figure 27 Exemple résultat 3D proposé par DynaMut

Suivi par toutes les informations qui nous permettent de bien comprendre et surtout de visualiser le changement causé par la mutation étudiée

- mutation pour la version DynaMut2 on clique sur Wild-type et on choisit « mutant », pour visualiser avec et sans la

Analyse par DeeDDG :

-on tape « deepddg » sur Google et on clique sur :

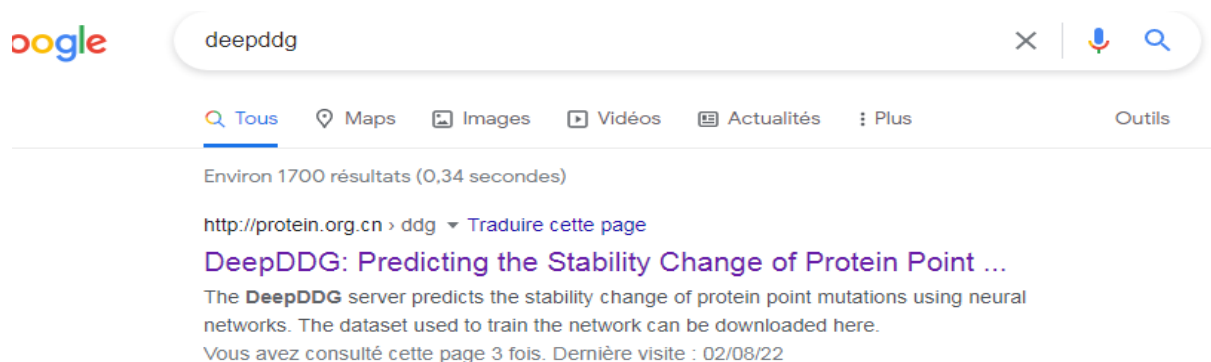


Figure 28 Lien deepDDG utilisé

-on met le code pdb, on a utilisé « 7N1U »

-on choisit « DeepDDG (non integrated model) » et (all possible mutations)

-et on clique sur submit

Pour les deux types d'analyse (Dynamut et DeepDDG) on a mis notre adresse mail pour recevoir les résultats, parce que l'analyse prend beaucoup de temps

Le résultat obtenu est ainsi :

```
#chain WT ResID Mut ddG (kcal/mol, >0 is stable, <0 is unstable)
A Q 14 A -0.608
A Q 14 R -0.398
A Q 14 N -0.527
A Q 14 D -0.483
A Q 14 C -0.520
A Q 14 E -0.432
A Q 14 G -0.750
A Q 14 H -0.485
A Q 14 I -0.349
A Q 14 L -0.464
A Q 14 K -0.448
A Q 14 M -0.354
A Q 14 F -0.481
A Q 14 P -0.631
A Q 14 S -0.564
A Q 14 T -0.675
A Q 14 W -0.376
A Q 14 Y -0.485
A Q 14 V -0.611
A C 15 A -1.291
A C 15 R -1.251
A C 15 N -2.033
A C 15 D -1.674
A C 15 Q -1.226
A C 15 E -1.940
A C 15 G -1.597
A C 15 H -1.485
A C 15 I -0.942
A C 15 L -1.428
A C 15 K -1.696
A C 15 M -1.141
A C 15 F -1.404
A C 15 P -1.617
A C 15 S -0.700
A C 15 T -1.674
A C 15 W -0.832
A C 15 Y -0.927
A C 15 V -1.650
A V 16 A -1.787
A V 16 R -1.876
```

Et on choisit uniquement les mutations qui nous intéressent.

Analyse par SNAP2 :

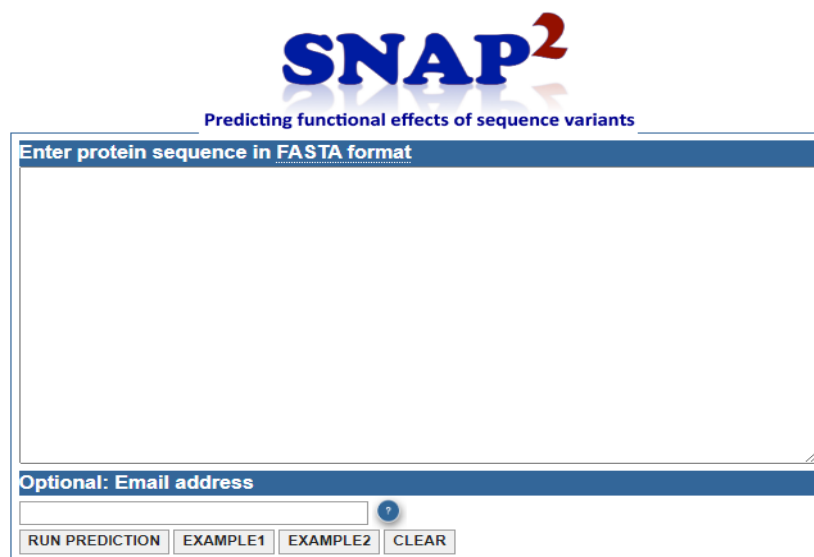
-on tape SNAP2 sur google , et on clique sur :

<https://roslab.org> > [snap2web](#) ▾ [Traduire cette page](#)

SNAP2 - Predicting functional effects of sequence variants

SNAP2: Prediction of functional effects of sequence variants.

-resultat :



The image shows the SNAP2 web interface. At the top, the logo "SNAP²" is displayed in blue and red, with the tagline "Predicting functional effects of sequence variants" below it. The main form area has a blue header that says "Enter protein sequence in FASTA format". Below this is a large, empty text input field. Underneath the input field is another blue header that says "Optional: Email address", followed by a small text input field. At the bottom of the form are four buttons: "RUN PREDICTION", "EXAMPLE1", "EXAMPLE2", and "CLEAR".

Figure 29 Figure de face du serveur SNAP2

on place la sequence spike proteine refrence (WIV04) en format « fasta »

-on clique sur « run prediction » :

Le resultat et ainsi :

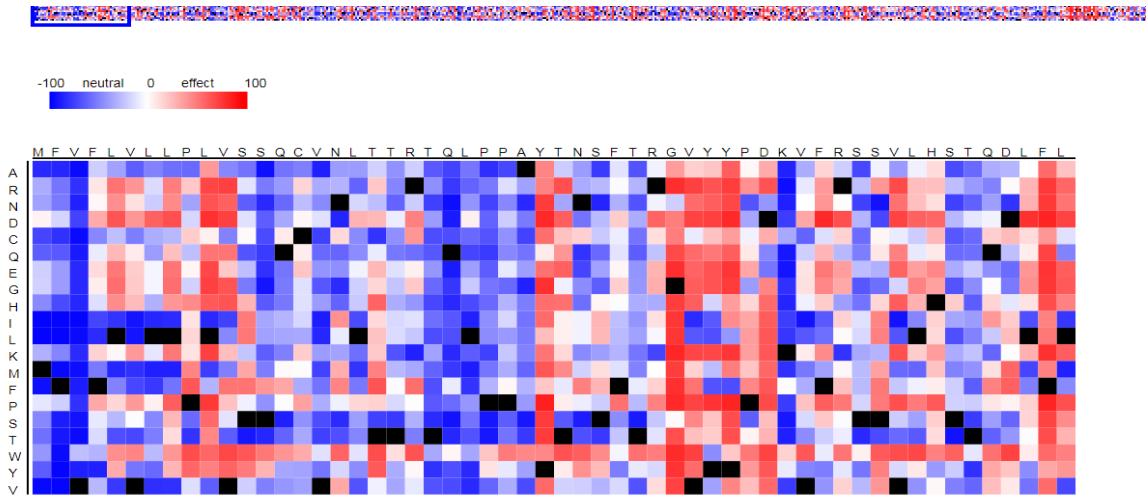


Figure 30 la prediction des mutation neutral et qui affect par SNAP2

La lecture de cette figure se fait à l'aide de la clé en haut montrant la dégradation des couleurs et si la mutation affecte ou elle est neutre, la lecture de AA non muté et sur la ligne orientale en haut, le AA muté et sur la ligne verticale à gauche

Jump to residue position: Variants per page:

Wildtype Amino Acid◆	Position◆	Variant Amino Acid◆	Predicted Effect◆	Score◆	Expected Accuracy◆
A	67	Q	effect	16	59%
A	67	M	effect	17	59%
A	67	K	neutral	-21	61%
A	67	P	neutral	-24	61%
A	67	C	neutral	-32	66%
A	67	H	effect	36	66%
A	67	L	neutral	-37	66%
A	67	Y	effect	31	66%
A	67	R	effect	45	71%
A	67	E	effect	42	71%
A	67	F	effect	46	71%
A	67	N	neutral	-45	72%
A	67	I	neutral	-50	72%
A	67	D	effect	50	75%
A	67	W	effect	63	80%
A	67	G	neutral	-66	82%
A	67	T	neutral	-80	87%
A	67	S	neutral	-82	93%
A	67	V	neutral	-85	93%
A	67	A	neutral	-99	97%

Figure 31 recherche de l'effet de la mutation de l'AA 67

on met à chaque fois dans « jump to residue position » le numéro indiquant la position de notre mutation précise Et on cherche les mutations qui nous intéressent

Exemple d'une mutation qui affecte la fonction.

Jump to residue position: 505 Go! Variants per page: 20

Wildtype Amino Acid	Position	Variant Amino Acid	Predicted Effect	Score	Expected Accuracy
Y	505	C	effect	6	53%
Y	505	H	effect	5	53%
Y	505	I	neutral	-18	57%
Y	505	L	neutral	-15	57%
Y	505	W	neutral	-13	57%
Y	505	M	effect	15	59%
Y	505	T	effect	13	59%
Y	505	Q	effect	29	63%
Y	505	A	effect	31	66%
Y	505	S	effect	39	66%

Figure 32 exemple de la mutation Y505H

Analyse structural et 3D :

Pour visualiser la forme 3D de notre protéine avec et sans mutations, et étudier chaque mutation et son changement structural , ainsi que l'effet des mutations étudiées sur l'interaction avec le récepteur (hACE 2) , on utilise le logiciel Pymol

Le programme permet une multitude d'activités, y compris: la fusion et la séparation des molécules, sélections arbitraires des molécules, un rendu d'images de haute qualité, etc.

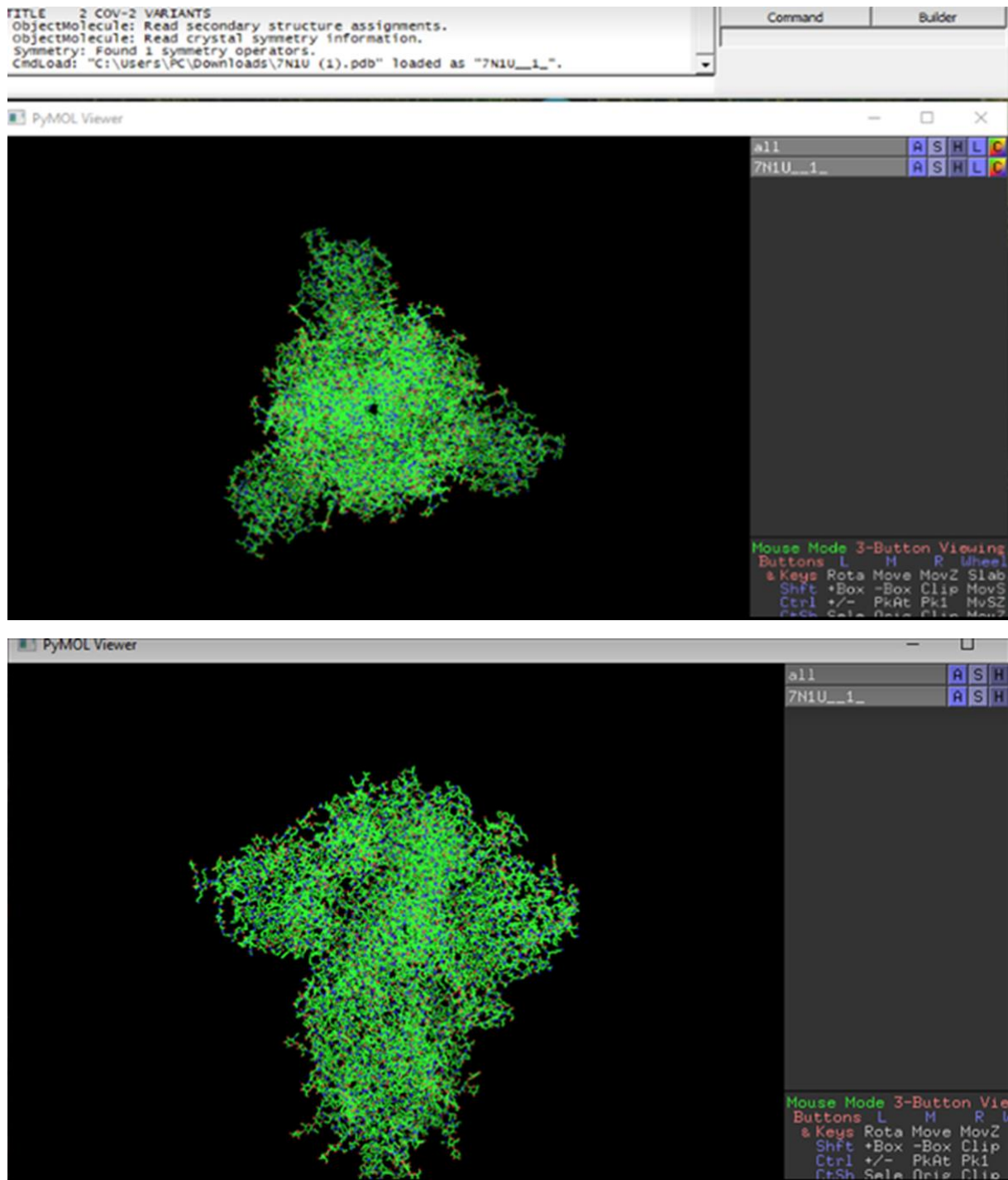


Figure 33 la sequence 7N1U par pymol

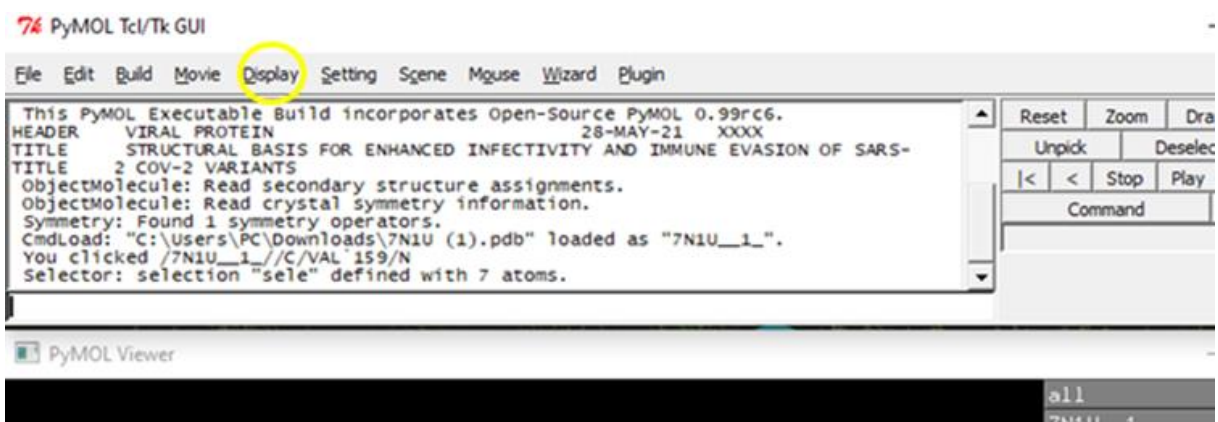


Les lettres qui apparaissent ont la signification suivante :

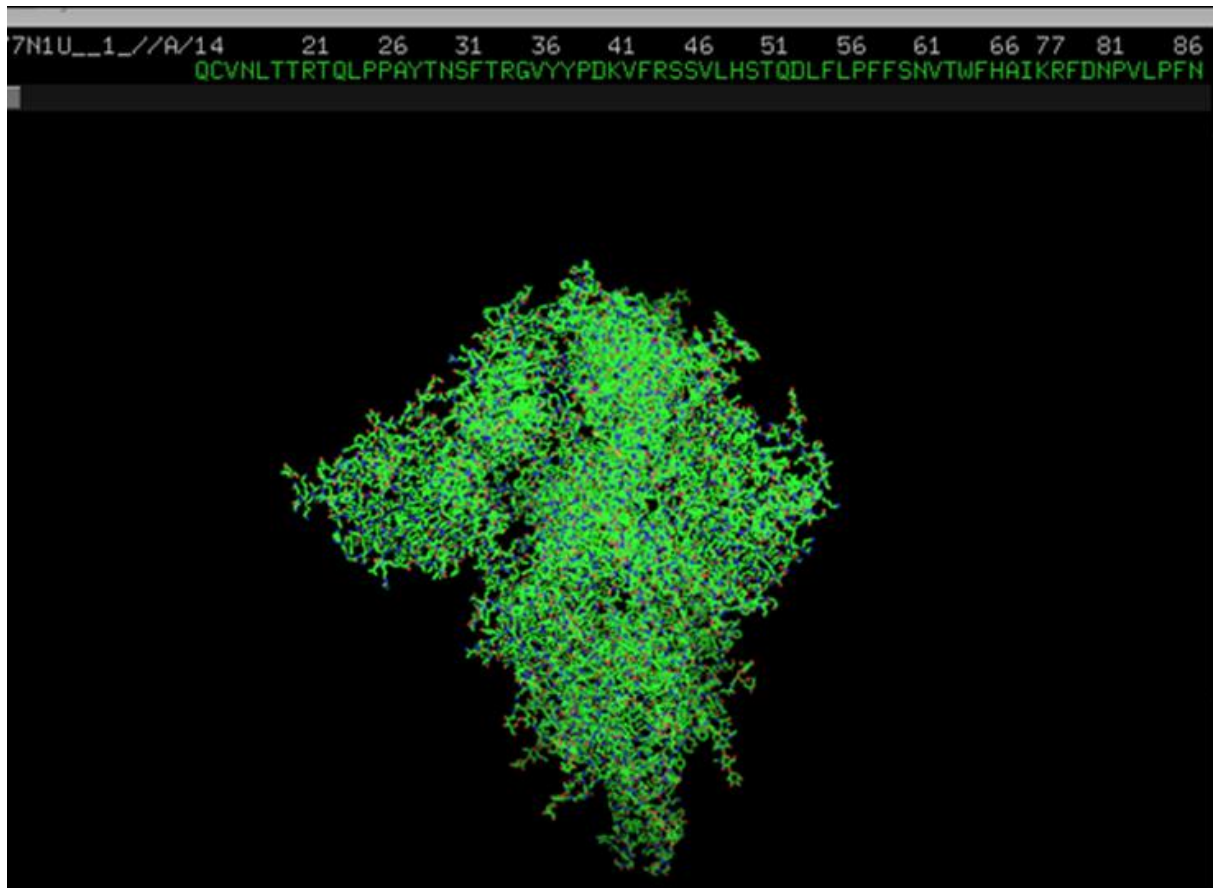
A: action / S: show / H: hide / L: label / C: color

-la première des choses qui nous intéresse c'est l'emplacement de chaque mutation en version 3D pour cela

-on clique sur « Display »

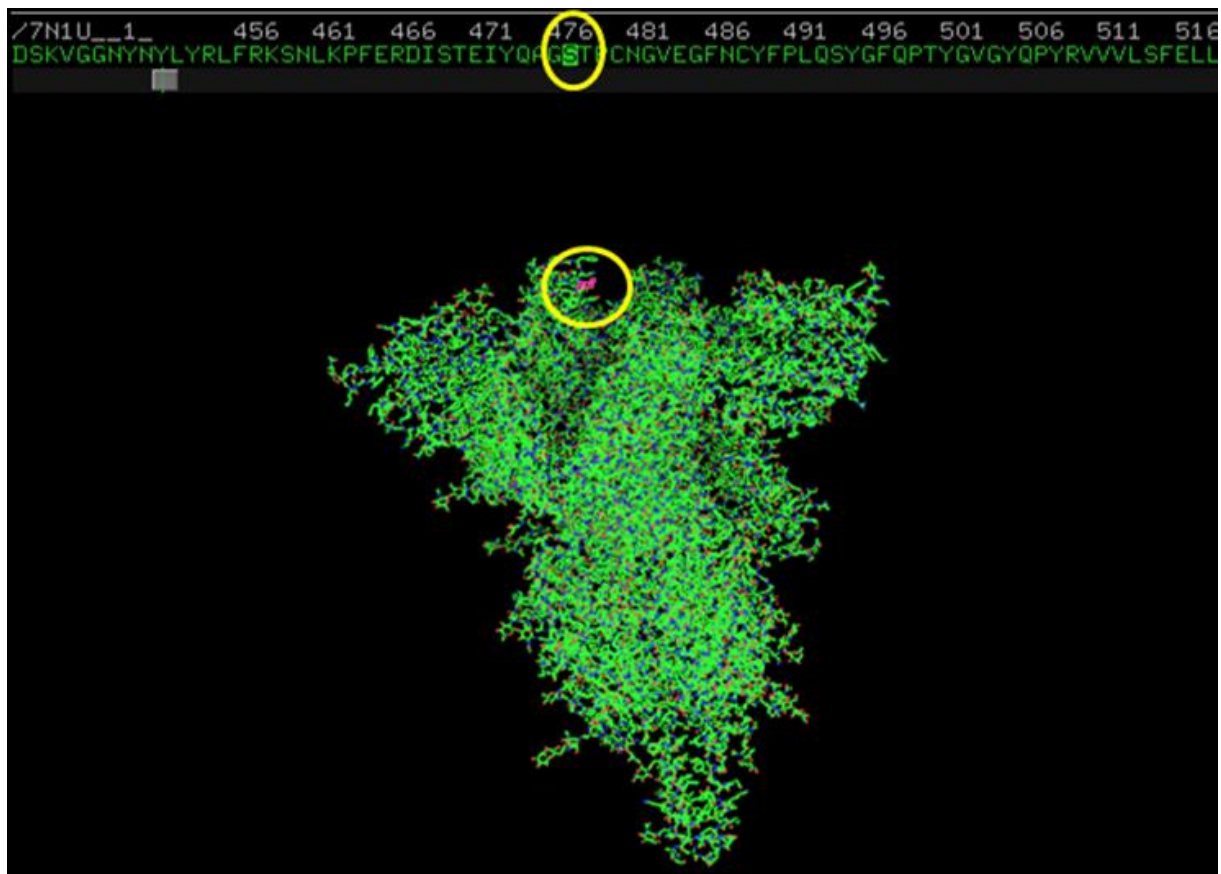


-on clique sur séquence, pour obtenir la séquence d'AA de notre protéine avec leurs positions.

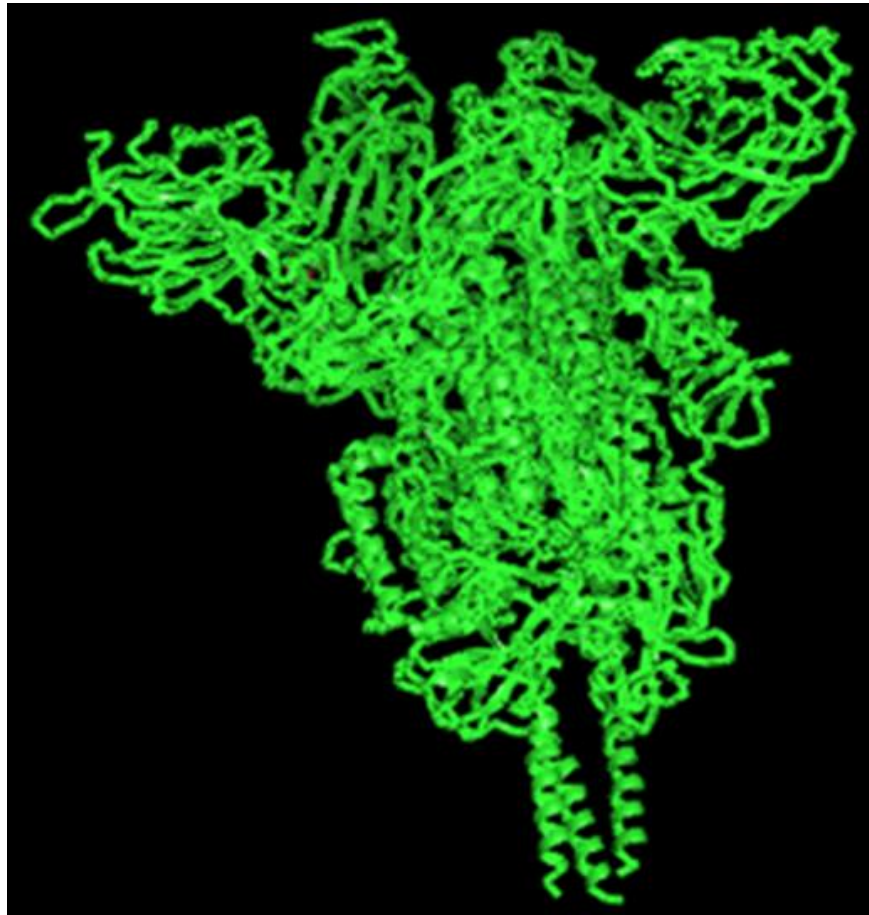


A chaque fois on clique sur un résidu de notre choix pour visualiser son positionnement dans la structure 3d mentionné par une couleur différente

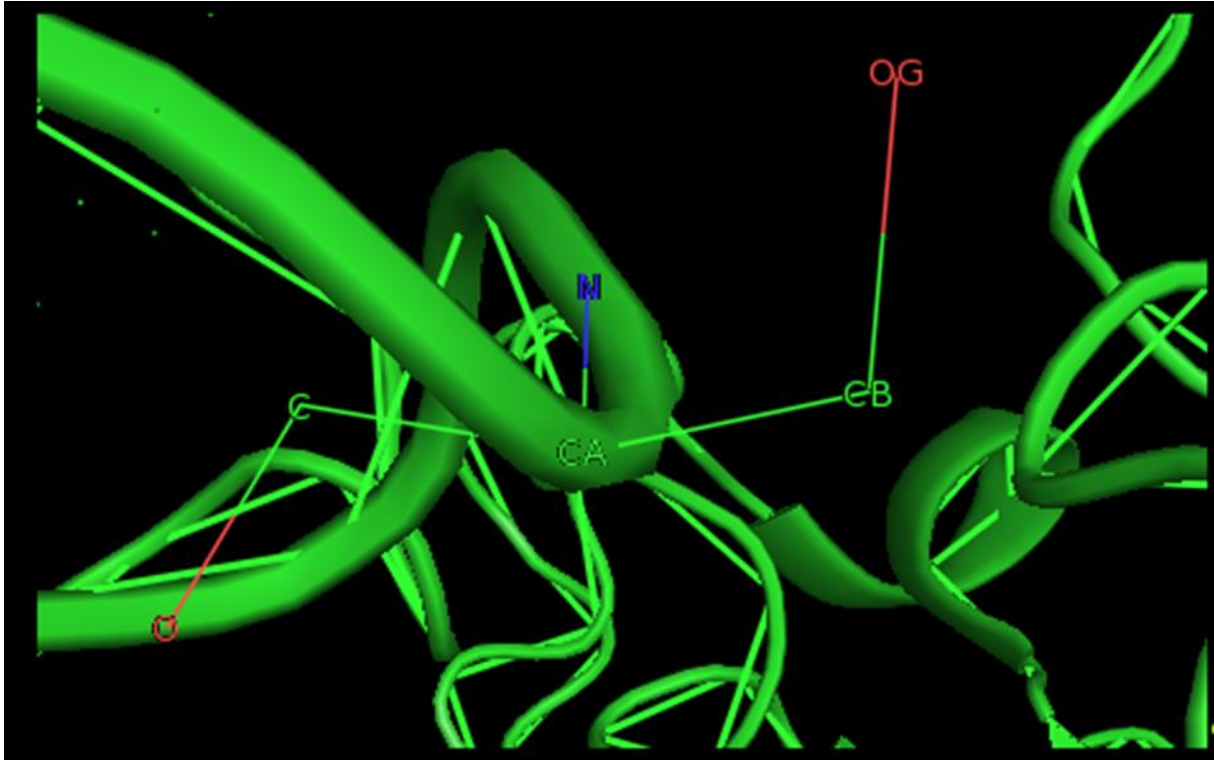
Exemple : résidu 477 de la séquence non muté.



-la couleur violet indique son positionnement.



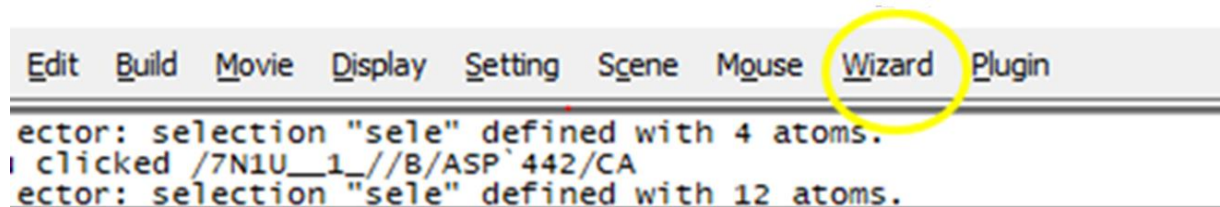
- pour bien visualiser nos mutations on clique sur zoom, Exemple : pour la position 477.
- on clique sur « L » et on clique sur atome Name.



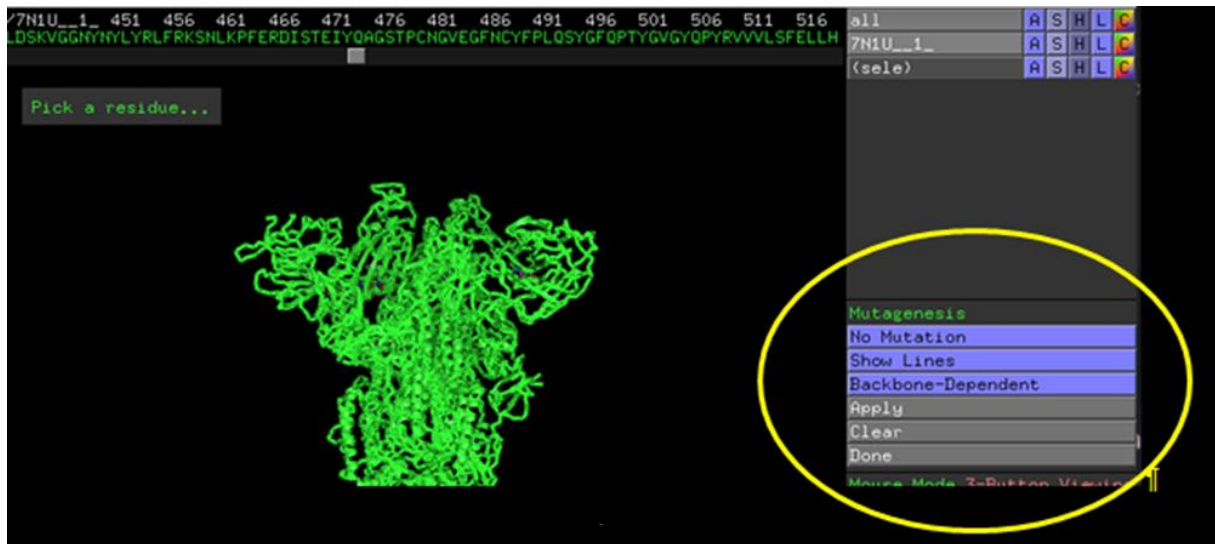
Pour visualiser les mutations une par une et toute à la fois

-on clique sur « display » après on clique sur « sequence »

-on clique sur « Wizard » après sur « Mutagenesis »



-on obtient un tableau à gauche :



- on clique sur le résidu voulue de la séquence en haut exemple : S477.
- on clique sur « No mutation » .
- on choisit AA muté exemple de notre étude : N, donc la mutation est nommé S477N.

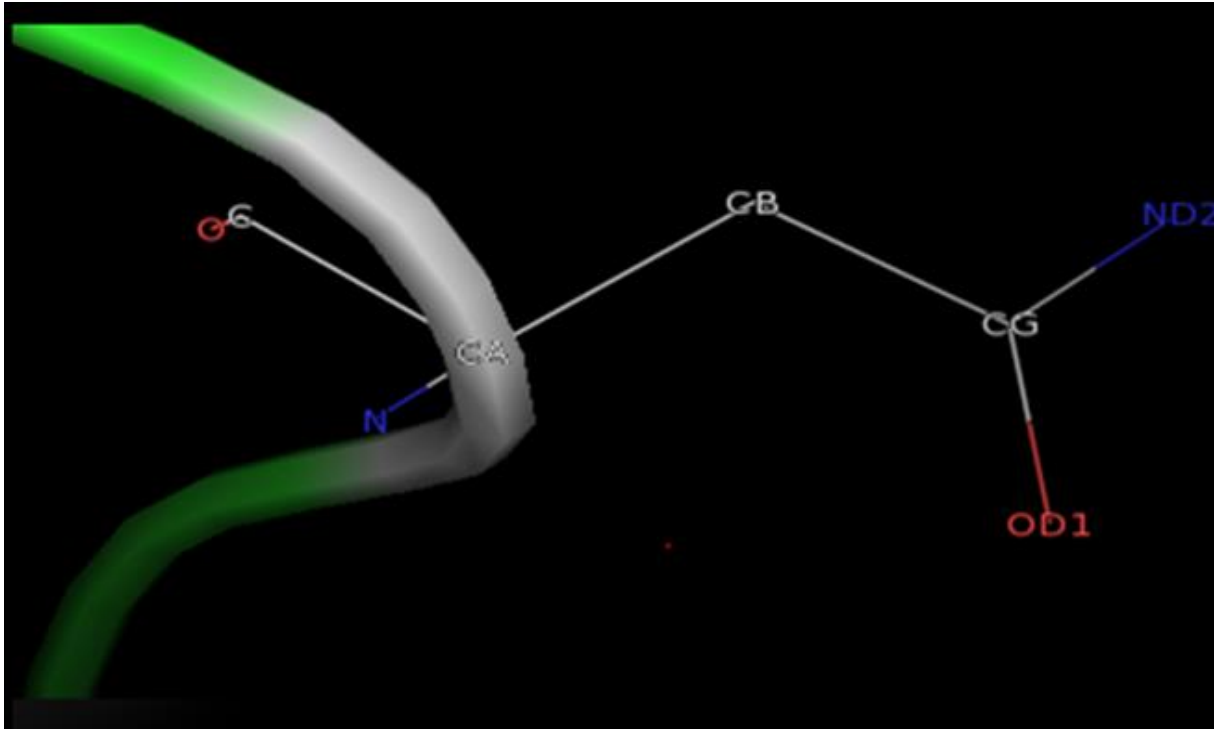


-on clique sur « Apply » , après sur « done »

La mutation est indiquée en blanc sur la séquence et même sur la structure.

```
'7N1U__1_ 456 461 466 471 476 481 486 491 496 501 506 511 516  
ISKVGGNYLYRLFRKSNLKPFERDISTEIQAGNTPCNGVEGFNCYFPLQSYGFQPTYGVGYQPYRVVLSFELLHA
```

-on zoom sur le residu muté et en clique sur « l » après sur « atom Name ».



-on a fait pareille pour chaque mutation, et avec toute les mutations à la fois

Pour l'étude de l'interaction moléculaire avec l'hACE2, on a travaillé avec la séquence « 6MOJ ».

Biological Assembly 1

6M0J

Crystal structure of SARS-CoV-2 spike receptor-binding domain bound with ACE2

PDB DOI: 10.2210/pdb6M0J/pdb

Classification: **VIRAL PROTEIN/HYDROLASE**

Organism(s): Homo sapiens, Severe acute respiratory syndrome coronavirus 2

Expression System: Trichoplusia ni

Mutation(s): No

Deposited: 2020-02-21 Released: 2020-03-18

Deposition Author(s): Wang, X., Lan, J., Ge, J., Yu, J., Shan, S.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION

Resolution: 2.45 Å

R-Value Free: 0.227

R-Value Work: 0.192

R-Value Observed: 0.194

wwPDB Validation

Metric Percentile Ranks

Rfree	Percentile Ranks
Clashscore	Percentile Ranks
Ramachandran outliers	Percentile Ranks
Cis-peptide outliers	Percentile Ranks

View: Structure | 1D-3D View | in Density | Validation Report

-on télécharge la séquence et on l'ouvre avec « pymol ».

PyMOL Viewer

all A S H L

6M0J A S H L

(sele) A S H L

Mouse Mode 3-Button Viewing

Buttons L M R Wheel

& Keys Rota Move MovZ Sla

Shft +Box -Box Clip Mov

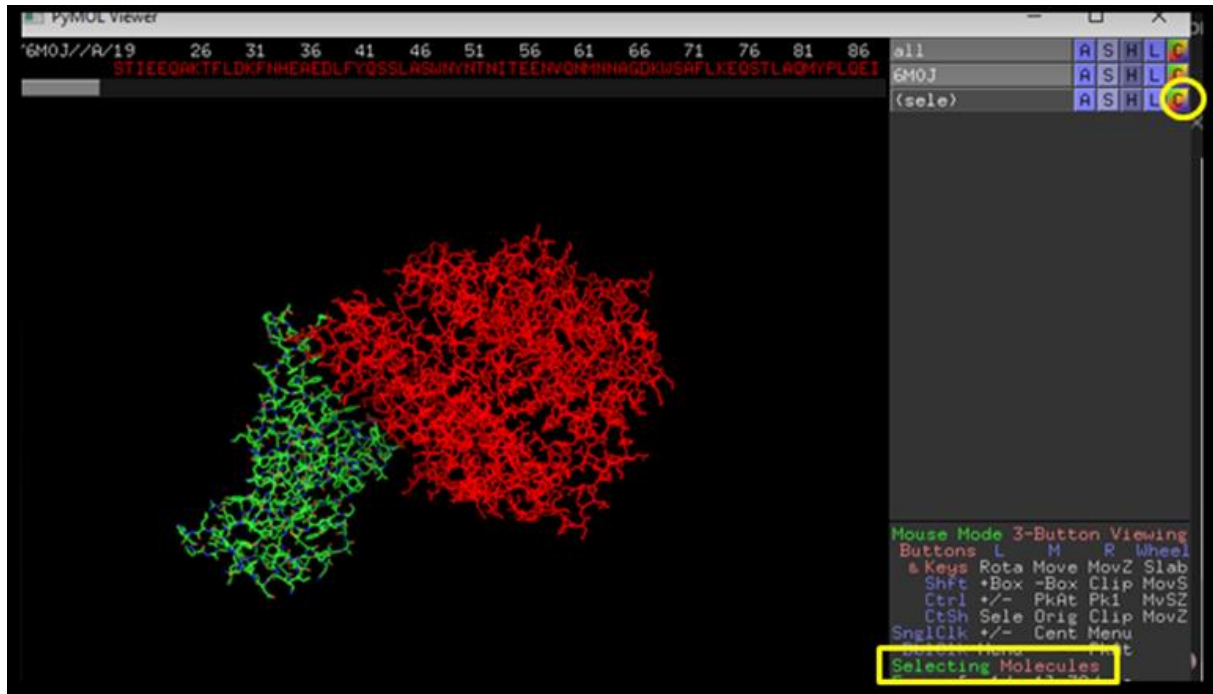
Ctrl +/- PkAt Pk1 MvS

-On clique sur « Display » après sur séquence pour avoir la séquence d'AA

-on différencie la couleur d'RBD de l'hACE2

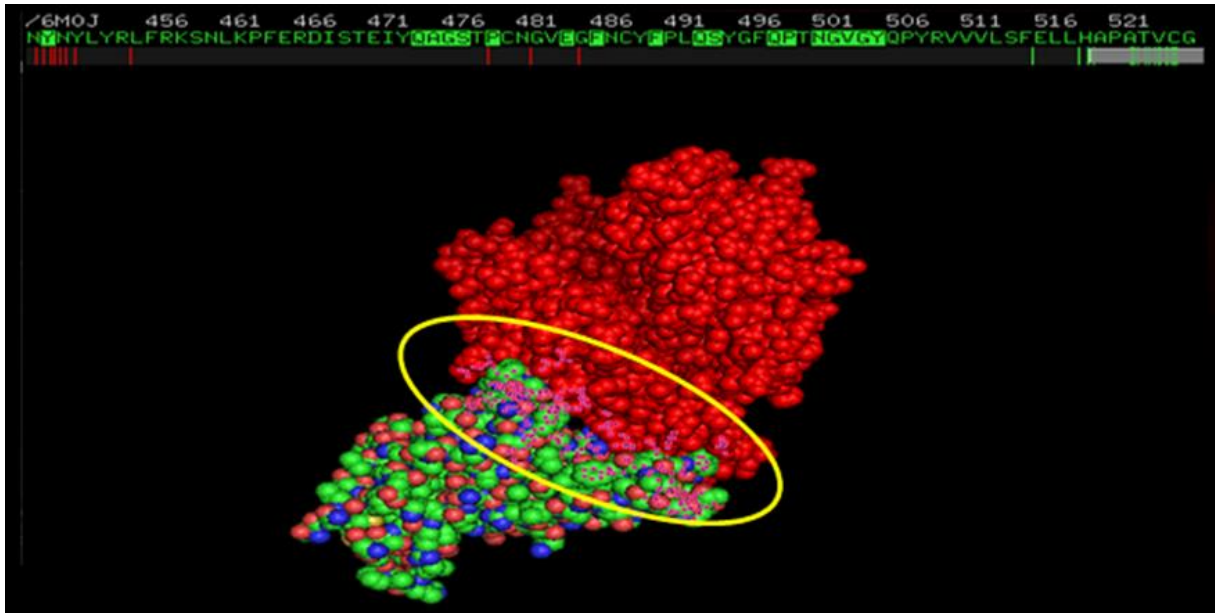
-on clique sur (selecting) et on choisit molécules

- on clique sur la structure et une des molécules sera sectionnée
- on clique sur (c) de (sel) et on choisit la couleur, dans l'exemple (rouge).

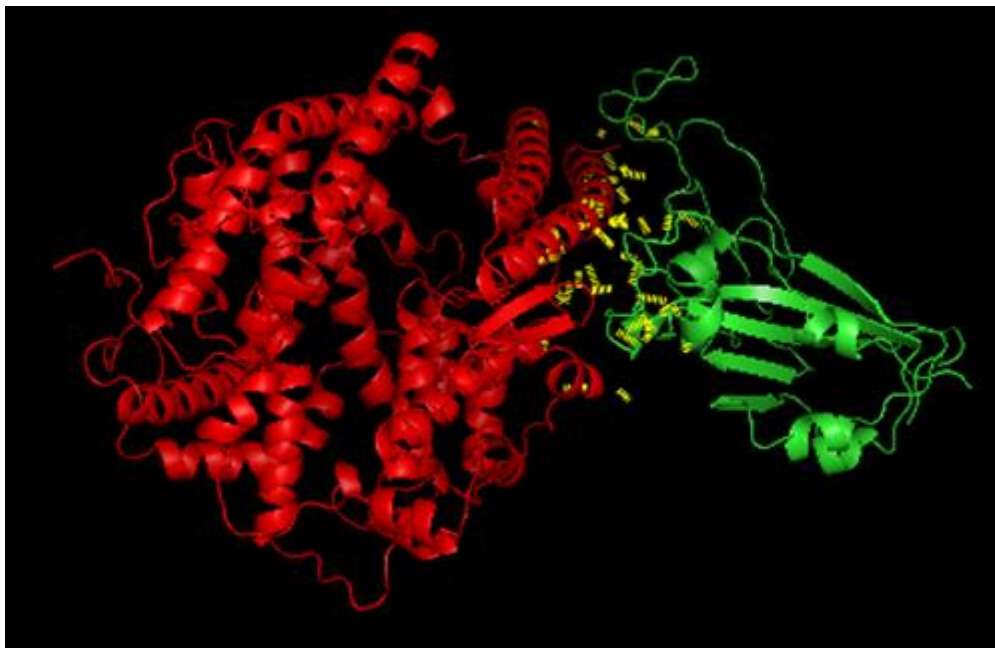


Pour étudier et visualiser la liaison entre les deux molécules avant et après mutation on :

- on affiche toujours les séquences en haut
- on clique sur (H) et on choisit (everything), après on clique sur (S) de (all) et on clique sur (surface)
- on clique sur (sélection) et on choisit résidu
- on choisit a vue d'œil les résidus d'interaction des deux molécules.

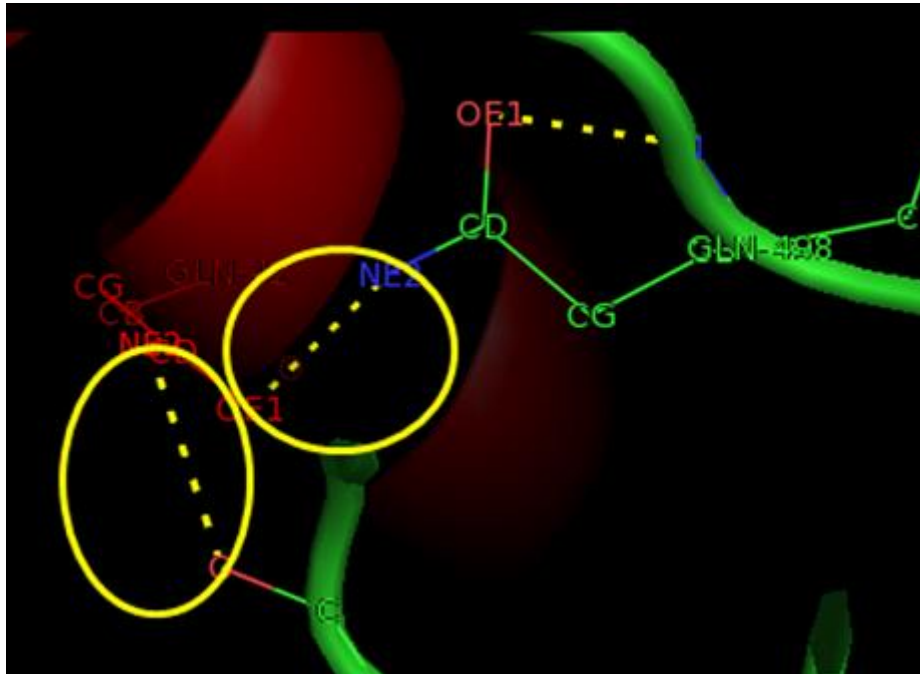


- on clique sur (rename selecting) pour enregistré et on pourras après étudier l'interaction de chaque résidus appart
- on clique sur (s) et on choisit (ribbon ou cartoon)
- on clique sur (A) de (sele qu'on a nommé interaction) , et on clique sur (find) après sur (polar contact) après sur (to authors atoms in object).



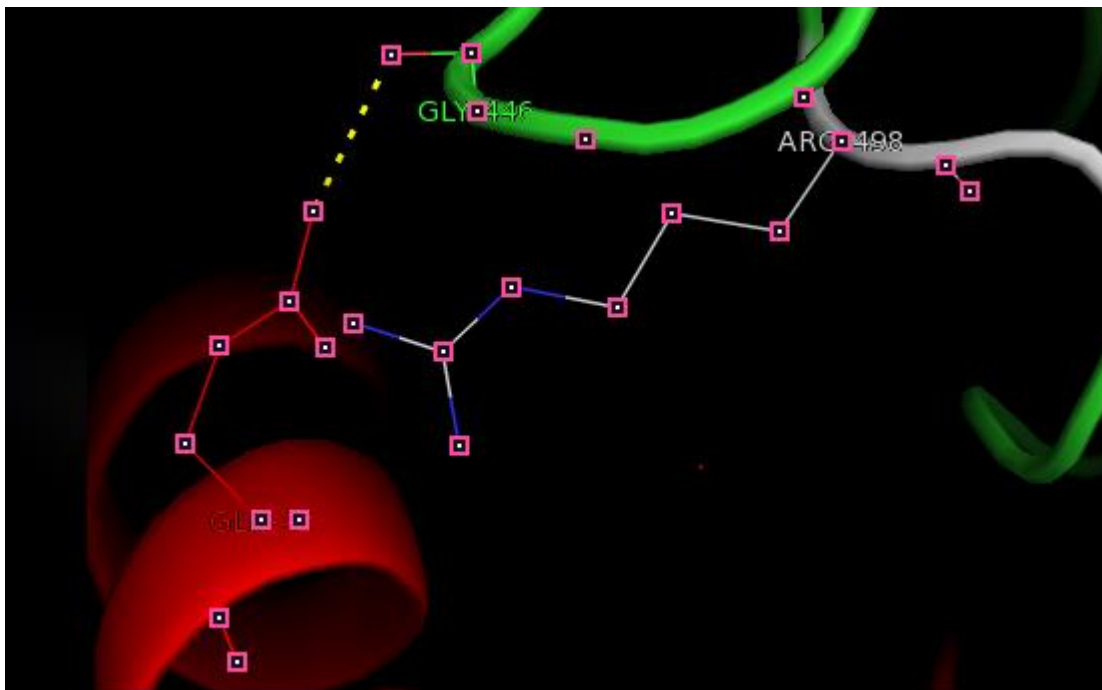
-on clique sur zoom pour visualiser entre quelles résidus se trouve exactement :

Exp : interaction hACE L45 avec RBD (G446 et le Q498) sans mutation on fait la même chose avec les autres résidus.



-on introduit les mutations et on refait la même chose :

Exp : mutation (Q498R) : perte de liaison avec le hACE



Toujours pour bien comprendre l'interaction moléculaire on a calculé la charge de la molécule ACE2 issue par le code PDB 1R42 ainsi que la charge RBD et ACE2 du complexe 6M0J et aussi du complexe 7WK6 (liaison RBDomicon –ACE2)

On a également fait un alignement structural pour la protéine complète avec et sans mutation et aussi pour le site RBD avec et sans mutation pour la protéine (omicron on a travaillé avec le code PDB : 7QO7)

Pour effectuer notre alignement il suffit d'introduire les deux structure a aligné au même temps sur pymol et on clique sur « Align »

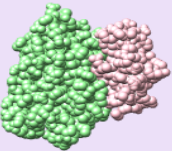
A la fin nous avons effectué un docking moléculaire avant et après les mutations, on a travaillé avec la séquence du code pdb (6M0J)

-on a d'abord séparé le ligand et le RBD avec pymol

-l'introduction des mutations aussi était effectué par pymol

-on à utiliser le HDOCK pour l'amarrage moléculaire

-on introduit nos séquence (ACE2 et RBD)déjà couper par pymol et en enregistré en format PDB.



HDock SERVER

Protein-protein and protein-DNA/RNA docking based on a hybrid algorithm template-based modeling and *ab initio* free docking.

[\[Huang Lab\]](#) [\[HDock\]](#) [\[Help\]](#) [\[Output example\]](#)

Input Receptor Molecule using **ONE** of the following four options: [\[help\]](#)

- Upload your **pdb** file in **PDB format**: Aucun fichier choisi [\[example\]](#)
- OR** provide your **pdb** file in PDB ID:ChainID: (Example: [1CGI:E](#))
- OR** copy and paste your **sequence** below in **FASTA format** (Sample input: [1CGI:E](#), [1HCJ:A](#))

OR upload your **sequence** file in **FASTA format**: Aucun fichier choisi [\[example\]](#)

Input Ligand Molecule using **ONE** of the following four options: [\[help\]](#)

- Upload your **pdb** file in **PDB format**: Aucun fichier choisi [\[example\]](#)
- OR** provide your **pdb** file in PDB ID:ChainID: (Example: [1CGI:I](#))
- OR** copy and paste your **sequence** below in **FASTA format** (Sample input: [1CGI:I](#)) [\[help\]](#)

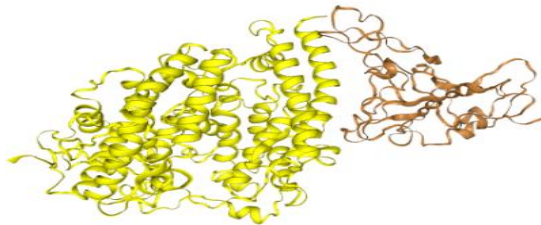
- On clique sur submit
- Exemple du résultat :

Your HDOCK results for job 63133faabee5

Download Files

[Receptor_PDB_file](#) [Ligand_PDB_file](#)
[\[0\]](#) [\[1\]](#) [\[2\]](#) [\[3\]](#) [\[4\]](#) [\[5\]](#) [\[6\]](#) [\[7\]](#) [\[8\]](#) [\[9\]](#) [\[10\]](#) [\[11\]](#) [\[12\]](#) [\[13\]](#) [\[14\]](#) [\[15\]](#) [\[16\]](#) [\[17\]](#) [\[18\]](#) [\[19\]](#) [\[20\]](#)
[Top 10 Predictions](#) [Top 100 Predictions](#) [All the results in a package](#)

* NOTE: The model #0 is predicted by template-based homology modeling.



Receptor style
Cartoon

Receptor color
 Pure
 Rainbow

MODEL No.
Model 0
Model 1
Model 2
Model 3
Model 4
Model 5
Model 6
Model 7
Model 8
Model 9
Model 10

Ligand style
Cartoon

Molecule	PDB ID	Chain ID	Align_length	Coverage	Seq_ID (%)
Receptor	6M0J	E	194	1.000	100.0
Ligand	6M0J	A	597	1.000	100.0

Note: The built model of "Model 0" based on the above PDB complex template has a **high** confidence.

Summary of the Top 10 Models

Rank	1	2	3	4	5	6	7	8	9	10
Docking Score	-310.19	-237.32	-237.06	-231.88	-225.96	-224.20	-223.40	-222.77	-218.58	-217.36
Confidence Score	0.9610	0.8515	0.8508	0.8372	0.8204	0.8152	0.8128	0.8108	0.7976	0.7937
Ligand rmsd (Å)	3.27	88.65	79.42	70.45	42.59	36.83	3.66	53.61	73.28	80.54
Interface residues	model 1	model 2	model 3	model 4	model 5	model 6	model 7	model 8	model 9	model 10

Note: The models are ranked according to the docking scores. Please click [help](#) for the explanations of evaluation metrics

Quality of Docking Structures/Input Data ([Click to Show](#))

on clique sur modèle 1 (interface résidus) pour obtenir les résidus d'interface et leur interaction ,exemple des résultats :

=====
Model 1
=====

Receptor interface residue(s):

LYS	417E	2.599
GLY	446E	2.463
GLY	447E	3.912
TYR	449E	2.375
TYR	453E	3.740
LEU	455E	3.282
PHE	456E	3.133
TYR	473E	3.524
ALA	475E	3.440
GLY	476E	3.381
SER	477E	4.209
GLU	484E	2.558
GLY	485E	4.788
PHE	486E	2.768
ASN	487E	2.946
TYR	489E	2.886
PHE	490E	3.783
GLN	493E	2.063
TYR	495E	4.893
GLY	496E	2.579
PHE	497E	4.671
GLN	498E	2.917
THR	500E	2.730
ASN	501E	3.425
GLY	502E	3.169
TYR	505E	3.275

Ligand interface residue(s):

SER	19A	4.579
GLN	24A	2.946
THR	27A	3.133
PHE	28A	2.886
ASP	30A	2.599
LYS	31A	2.348
HIS	34A	3.282
GLU	35A	2.063
GLU	37A	4.330
ASP	38A	2.462
TYR	41A	2.730
GLN	42A	2.375
LEU	45A	3.532
LEU	79A	3.647
MET	82A	2.909
TYR	83A	2.768
ASN	330A	3.137
LYS	353A	2.579
GLY	354A	3.840
ASP	355A	3.417
ARG	357A	4.223
ARG	393A	4.635

Résultats et discussion

Le variant SARS-CoV-2 Omicron a été désignée comme la variant préoccupante en raison de son émergence rapide dans le monde, qui comprend 30 mutations dans la protéine Spike, et près de la moitié d'entre elles se trouvent dans le domaine de liaison au récepteur. En raison de la perte de séquence, la variante Omicron a 1270 acides aminés au lieu des 1273 acides aminés du pic de référence.

On ressort d'une analyse primaire de la séquence protéique que cette variante contient plus d'arginine, d'histidine, de lysine et d'acide glutamique que la référence, ce qui indique que la protéine de pointe est plus chargée. De plus, ces résidus sont beaucoup plus exposés et contribuent à la liaison avec les récepteurs car leurs pKa sont suffisamment élevés avec les chaînes latérales polaires, qui peuvent former des liaisons hydrogène. D'autre part, l'isoleucine et la phénylalanine sont également présentes en plus grand nombre dans le noyau de la protéine, ce qui rend la protéine de pointe plus hydrophobe que la variante de référence. Ces mutations modifieront ses propriétés physicochimiques et structurelles, ce qui affectera le taux de transmission et la pathogénicité au sein des populations humaines en réduisant la protection médiée par les anticorps

➤ Physicochimiques :

Nos résultats physicochimiques sont représenté dans le tableau suivant indiquant les différences physicochimiques entre les trois séquences (séquence référence, la référence avec mutation et la séquence omicron utilisé)

Tableau 3. Résultats physicochimiques (résultats personnel 2022).

Séquences	références	Référence avec mutation	Omicron
Résidus	1273	1273	1270
Pois moléculaire	141178.47	141421.10	141299.81
Point isoélectrique	6.6146	7	7.2679
Charge	1.5	7.0	9.0
Gravy	-0.079	-0.068	-0.083

instabilité index	33.01	33.53	34.67
Aliphatique index	84.67	85.51	84.95
Résidus Acidique	(B+D+E+Z) : 8.641	(B+D+E+Z) :8.562	(B+D+E+Z) : 8.740
Résidus basique	(H+K+R) :427	(H+K+R) :9.819	(H+K+R) : 10.157
Résidus polaires	45.277	(D+E+H+K+N+Q+R+S+T+Z):44.933	(D+E+H+K+N+Q+R+S+T+Z): 45.118
Résidus non polaires	54.753	(A+C+F+G+I+L+M+P+V+W+Y):55.067	(A+C+F+G+I+L+M+P+V+W+Y) 54.803

+ Pour la séquence omicron sans les 15 mutations de notre étude on a trouvé une charge de 1.5

-pas de différence entre le nombre des résidus de la séquence référence et la séquence avec les mutations étudiées qui est très évident parce que aucune des mutations étudié ne cause une délétion

-le nombre des résidus et le même on remarque une différence dans le poids moléculaire, cette différence de quelques résidus causé par les mutations et par la différence molaire de chaque résidu et son nombre de répétition.

-GRAVY (Grand average of hydropathicity index) a été calculé en utilisant les valeurs d'hydropathie de Kyte et Doolittle , la différence n'est pas très remarquable dans les trois cas la protéine est non-polar

La valeur de la Gravy de l'omicron spike protein indique que la protéine est devenue intrinsèquement légèrement plus hydrophile, ce qui indique les effets des mutations sur l'accessibilité de surface de la protéine en raison de l'altération des propriétés structurales secondaires et tertiaires

-présence d'un changement dans les résidus acide et base , on remarque une augmentation des résidus basiques qui donne le résultat d'une charge « 7 » pour la séquence référence mutée, ce qui facilite encore plus l'attachement avec le substrat hACE2

-les valeurs de l'index de l'instabilité : 33.01 , 33.53 , 34.67 respectivement pour les séquences Spike protéines : référence , référence avec mutation et omicron , sont toutes inférieures à 40 ce qui indique que les trois séquences sont stables . Avec la séquence référence étant plus stable.

-la séquence référence et la séquence référence avec mutation ainsi que la séquence omicron ont un index aliphatique élevé ce qui indique que ces séquences sont thermostables sur une large plage de températures.

➤ **Les propriétés fonctionnelles et de la stabilité :**

DynaMut 2 un serveur amélioré qui combine l'analyse en mode normal avec une représentation graphique de la structure des protéines, pour prédire avec précision et rapidité les effets des mutations ponctuelles uniques et multiples sur la stabilité et la dynamique des protéines

Tableau 4 Résultats de l'analyse fonctionnelles et de la stabilité (résultat personnel 2022)

AA Référence	position	AA muté	Effets.	
			Stabilité	fonctionnalité
A	67	V	Diminue la stabilité	Neutre
T	95	I	Diminue la stabilité	Neutre
S	371	L	Diminue la stabilité	Neutre
S	373	P	Diminue la stabilité	Neutre
S	375	F	Diminue la stabilité	Neutre
N	440	K	Diminue la stabilité	Affecte
G	446	S	Diminue la stabilité	Neutre
S	477	N	Diminue la stabilité	Neutre
E	484	A	Diminue la stabilité	Neutre
Q	493	R	Diminue la stabilité	Neutre
Q	498	R	Diminue la stabilité	Neutre
N	501	Y	Diminue la stabilité	Neutre
Y	505	H	Diminue la stabilité	Affecte
N	856	K	Diminue la stabilité	Neutre

N	969	K	Diminue la stabilité	Affecte
---	-----	---	----------------------	---------

En utilisant une combinaison d'algorithmes de réseau neuronal d'apprentissage en profondeur et de prédictions basées sur la structure, l'effet des mutations sur la stabilité de la protéine de pointe a été prédit. On a observé une diminution de la stabilité pour tous les résidus

L'analyse fonctionnel montre que Les mutations N440K,Y505H et la N969K affecte le fonctionnement de la protéine Spike , le reste des mutation n'affecte pas le fonctionnement -on remarque que parmi les mutations qui affectent la fonction de la protéine la N440K et la Y505H sont localisé au site récepteurs RBD.

🔍 Analyse pymol :

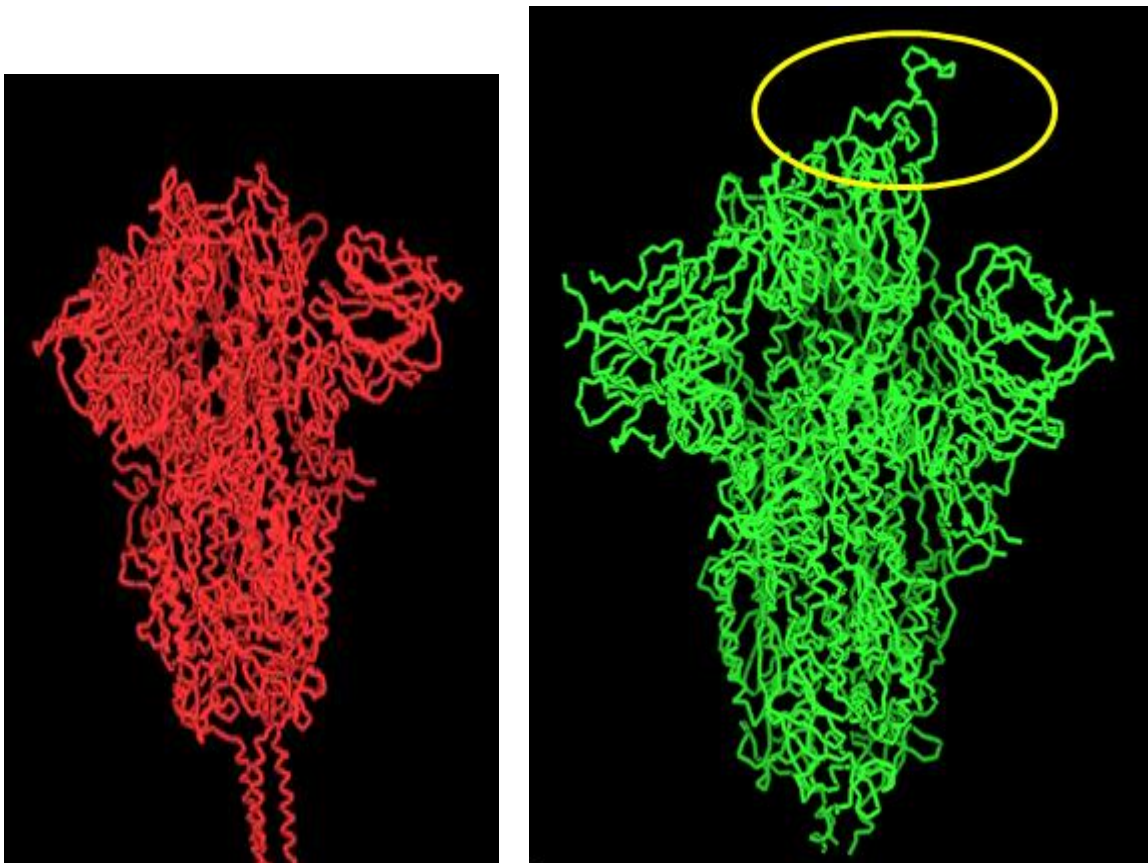


Figure 14 structure 3D de la spike proteine 7N1U (en rouge) et la (7Q07) en vert (image personnel 2022).

Le résultat structural nous a permis de voir d'une façon très claire les trimères, qui apparaissent au microscope électronique comme des protubérances

On remarque une différence structurale entre la séquence 7N1U (en rouge) et la séquence 7QO7 (en vert) dans la partie du site RBD au niveau des résidus (493 et 498), indiqué par le cercle rouge.

D'après la séquence 6M0J le site RBD est situé au environ (332-525)

D'après le positionnement de tous les résidus de notre étude et d'après la séquence 6M0J on remarque que les résidus (67, 95, 856 et 969) sans hors site RBD.

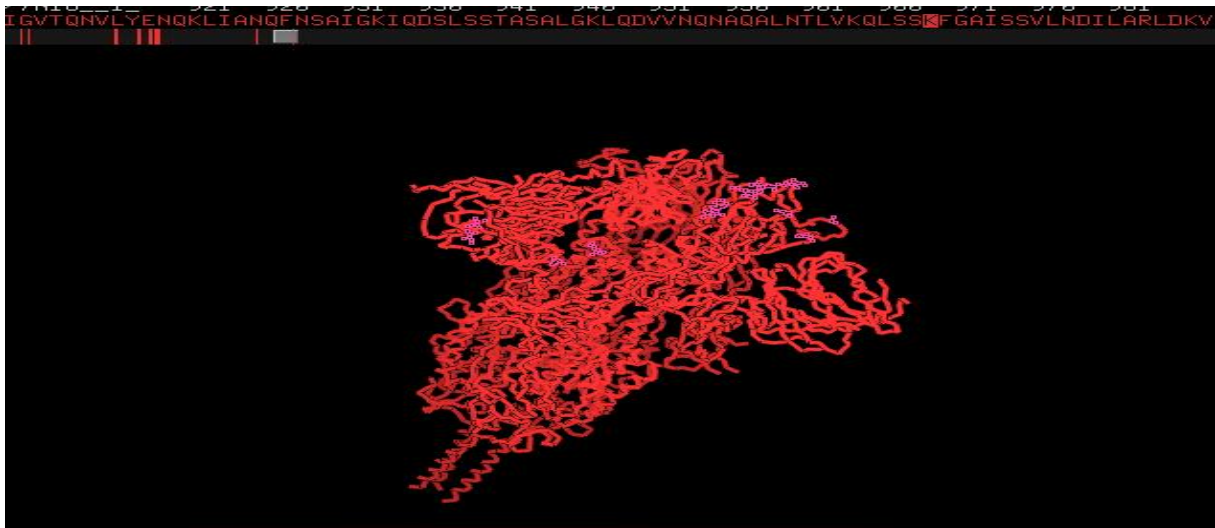


Figure 15. Spike protéine référence utilisé (7N1U) avec les 15 mutations apparaissent en violet (image personnel 2022).

Le calcul des charges par pymol a données :

Tableau 5 Charge mol des : (1R42), (6M0J) et (7WK6) (Résultat personnel 2022).

	ACE2	RBD
1R42	-20	/
6M0J	-28	2
7WK6	-28	5

Les résultats nous permis de déduire L'ACE 2 est considéré comme une molécule négative

La différence de charge des RBD montrant que le RBD omicron est plus basique, ce qui peut améliorer sa propension à se lier aux zones chargées négativement des autres molécules telles que l'ACE2

Résultat de l'alignement :

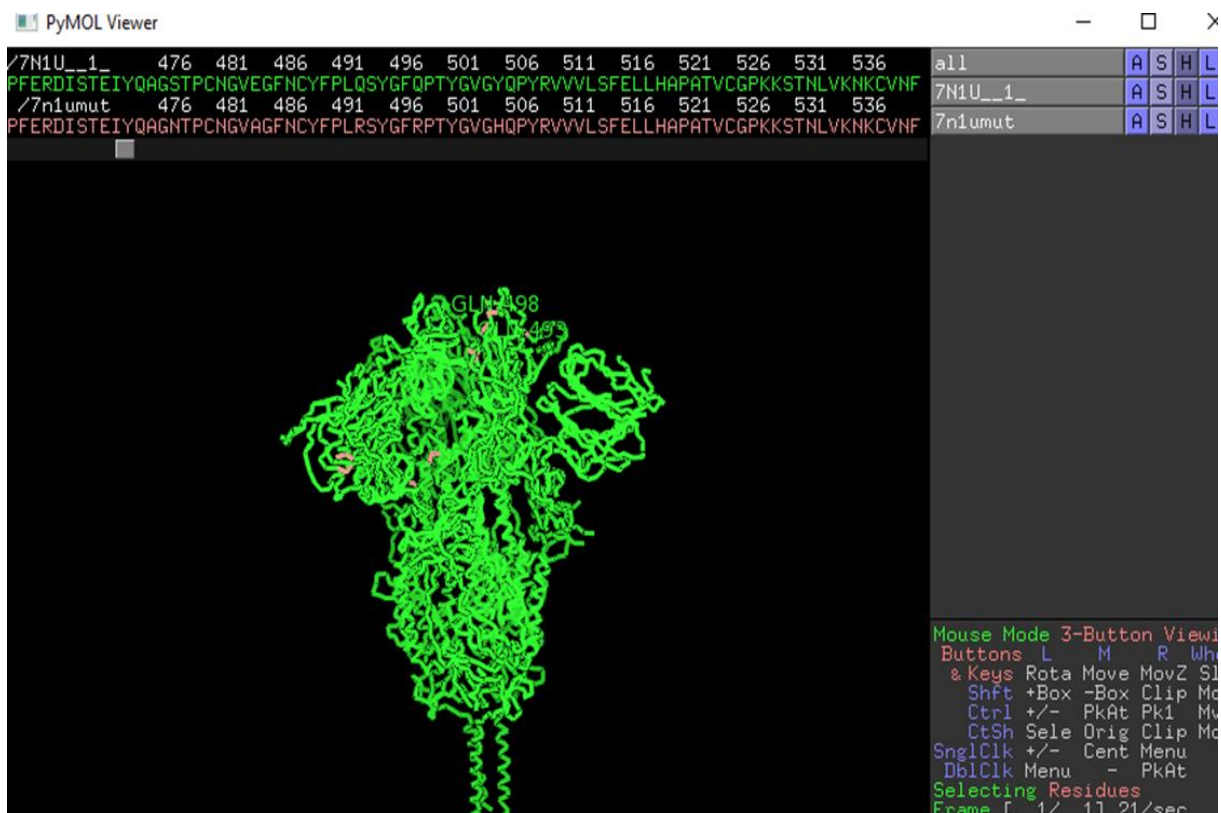


Figure 16 Alignement Spike protéine référence utilisée (7N1U) avec et sans mutations (Résultat personnel 2022)

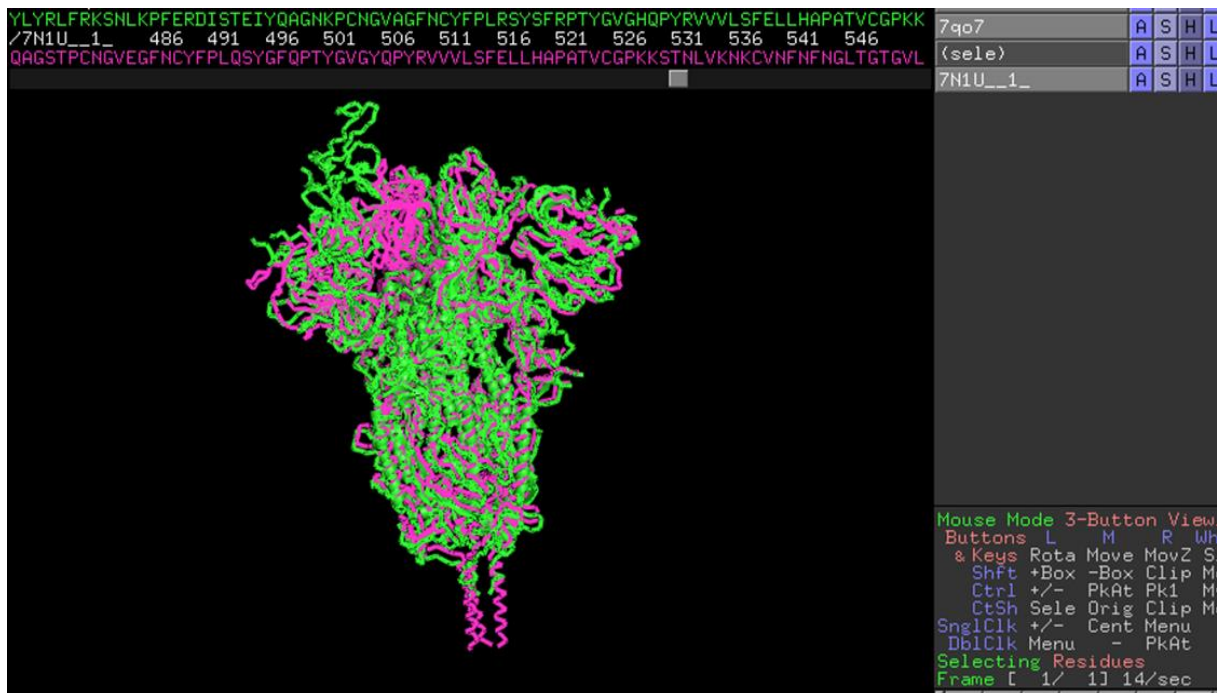


Figure 17 Alignement Spike protéine omicron (7QO7) et la Spike référence utilisé (7N1U). (Résultat personnel 2022)

On a constaté que la forme structural de la protéine Spike (7N1U) sans et avec mutations est presque la même

L'alignement 3D de la protéine Spike (7N1U) avec et sans mutations indique un RMSD=0,000Å

L'alignement 3D de la forme RBD pris de la séquence (6M0J), avec et sans mutations indique un RMSD= 0,000Å

L'alignement 3D de la protéine Spike omicron et la séquence référence utilisé (7QO7-7N1U) indique un

RMSD=2.549 Å

RMSD (root-mean-square-deviation) nous sert à savoir à quel point l'alignement était « bon », Plus la RMSD est faible, meilleur est l'alignement de la structure, un RMSD=0 on considère que c'est la même molécule , un RMSD de 2-2.5 Å signifie une similitude d'environ 82%

D'après l'analyse par HDock on a constaté que l'interaction (RBD –hACE2) possède plusieurs modèles (plus ou moins fort), pour des résultats correctes on a pris en considération seulement les résultats avec un RMSD très faible

Tableau 6. Résultat docking RBD- hACE2. (Sans et avec mutations) (résultat personnel 2022)

Docking RBD-récepteur hACE2 :

Docking RBD-récepteur hACE2 : Docking score	-310.19
Confidence score	0.9610
RMSD	3.27

Docking RBD avec mutation –hACE2 :

Docking RBD-récepteur hACE2 : Docking score	-335.0
Confidence score	0.9759
RMSD	0.67

Les interactions:

RBD sans mutations.

417E 30A 2.599	456E - 30A 3.984	486E - 79A 3.647	493E - 35A 2.063
417E - 34A 4.738	456E - 31A 3.559	486E - 82A 2.909	495E - 38A 4.893
446E - 42A 2.463	473E - 27A 3.524	486E - 83A 2.768	496E - 38A 2.989
446E - 45A 4.389	475E - 19A 4.579	487E - 24A 2.946	496 ^E - 353A 2.579
447E - 42A 3.912	475E - 24A 3.460	487E - 83A 3.281	497E- 353A 4.671
449E - 38A 2.462	475E - 27A 3.440	489E - 27A 4.276	498E - 38A 4.719
449E - 42A 2.375	476E - 24A 3.381	489E - 28A 2.886	498E - 41A 3.036
453E - 34A 3.740	477E - 24A 4.209	489E - 31A 3.709	498E - 42A 3.022
455E - 30A 3.438	484E - 31A 2.558	489E - 83A 2.983	498E - 45A 3.532
455E - 31A 4.371	484E - 35A 4.773	490E - 31A 3.783	498E- 353A 2.917
455E - 34A 3.282	485E - 79A 4.788	493E - 31A 2.348	500E - 41A 2.730
456E - 27A 3.133	486E - 24A 4.576	493E - 34A 4.757	500E - 45A 4.670
500E- 330A 3.137	502E- 354A 3.840		
500 ^E - 355A 3.417	502E- 355A 4.095		
500E- 357A 4.223	505E - 37A 4.330		

501E - 41A 3.532	505E- 353A 3.275
501E- 353A 3.425	505E- 354A 4.230
501E- 355A 4.638	505 ^E - 393A 4.63
502E- 353A 3.169	

RBD avec les 15 mutations de Notre etude :

417E - 30A 2.470	486E - 79A 3.636	498E - 42A 3.827
446E - 42A 4.214	486E - 82A 3.147	498E - 45A 4.381
449E - 38A 2.670	486E - 83A 2.234	500E - 41A 2.581
453E - 34A 2.839	487E - 24A 2.585	500E - 330A 3.044
455E - 30A 3.666	487E - 83A 3.799	500E - 355A 2.793
455E - 31A 3.758	489E - 24A 4.673	500E - 357A 3.814
455E - 34A 3.589	489E - 27A 3.511	501E - 38A 3.048
456E - 26A 4.647	489E - 28A 3.039	501E - 41A 3.147
456E - 27A 2.865	489E - 31A 3.811	501E - 353A 3.421
456E - 28A 4.857	489E - 76A 4.858	501E - 354A 4.704
456E - 30A 3.277	489E - 79A 4.621	501E - 355A 4.230
456E - 31A 3.604	489E - 83A 2.782	502E - 324A 4.763
473E - 27A 3.747	493E - 31A 4.781	502E - 353A 2.930
475E - 19A 2.332	493E - 34A 3.094	502E - 354A 3.012
475E - 23A 3.937	493E - 35A 2.380	502E - 324A 4.763
475E - 24A 2.801	494E - 38A 4.685	502E - 353A 2.930
475E - 27A 3.443	495E - 38A 4.818	502E - 354A 3.012
476E - 19A 3.703	496E - 38A 3.140	502E - 355A 4.017
476E - 24A 2.551	496E - 353A 4.999	503E - 324A 4.172
477E - 19A 4.738	498E - 38A 3.093	503E - 325A 4.216
477E - 24A 4.335	498E - 41A 3.302	503E - 354A 4.670
505E - 37A 2.719	505E - 353A 3.046	505E - 353A 3.046

Les scores d'amarrage (Docking score) sont calculés par la fonction de notation itérative basée sur la connaissance ITScorePP ou ITScorePR. Un score d'amarrage plus négatif signifie

un modèle de liaison plus possible, le résultat -310.19 sans mutations et -335.0 avec mutations montre clairement que l'amarrage avec mutation est plus négative donc un modèle de liaison plus possible

Confidence score indiquent la probabilité de liaison de deux molécules, En gros, lorsque le score de confiance est supérieur à 0,7, les deux molécules seraient très susceptibles de se lier ; lorsque le score de confiance est compris entre 0,5 et 0,7, les deux molécules pourraient se lier ; lorsque le score de confiance est inférieur à 0,5, il est peu probable que les deux molécules se lient. Nos résultats sans et avec mutation sont supérieures à 0.7 ce qui indique une grande susceptibilité de liaison encore plus avec les mutations

La valeur RMSD est utilisée pour valider le protocole d'amarrage, il est utilisé pour comparer la conformation ancrée avec la conformation de référence ou avec une autre conformation ancrée , un RMSD < 2.0 Å correspond à de bonnes solutions d'amarrage d'après nos résultats on considère que l'amarrage de RBD muté représente une bonne reproduction de la pose correcte ,on peut estimer que les mutations augmenté la stabilité du complexe RBD-ACE2 chez le variant omicron

D'après l'interaction des résidus on remarque que le RBD muté forme 5 interactions en plus que RBD référence, on constate aussi que les résidus en positions (446/477/493/498/501/505) sont des résidus essentiels pour l'interaction.

➤ L'ARBRE PHYLOGENETIQUE :

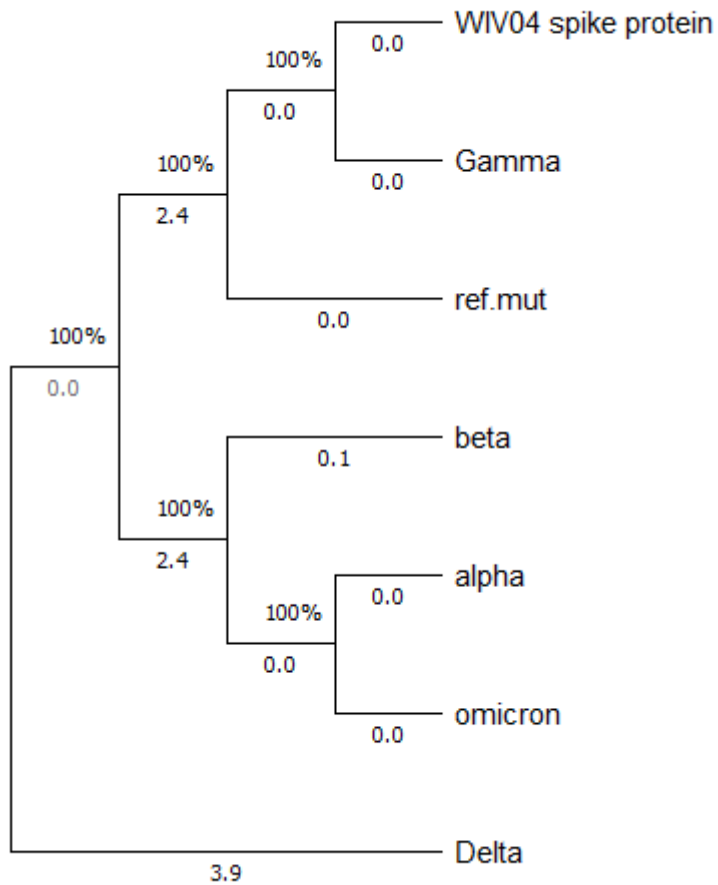


Figure 18 Arbre phylogénétique de la protéine spike des 5 variants (alpha, beta, gamma, delta et omicron) et la protéine référence avec et sans les mutations étudiées (résultat personnel 2022)

	1	2	3	4	5	6	7
1. WIV04 spike protein		0.00279	0.00265	0.19099	0.58112	0.60588	0.64644
2. ref.mut	0.01095		0.00369	0.19343	0.61243	0.58945	0.68474
3. Gamma	0.00938	0.01808		0.19443	0.57158	0.61203	0.64702
4. beta	2.80902	2.83942	2.84503		0.01412	0.49261	0.01577
5. alpha	5.76014	5.84682	5.78710	0.16530		10.77613	0.00875
6. Delta	5.05371	5.02368	5.13014	4.76027	9.45207		9.20589
7. omicron	6.03577	6.11386	6.09866	0.18820	0.08399	8.54708	

Figure 19 distance par paires de la Spike protéine des 5 variantes (alpha , beta ,gamma ,delta et omicron Et la protéine référence avec et sans les mutations (résultat personnel 2022)

Le premier génome séquencé de la variante Omicron a été utilisé pour tracer ses relations phylogénétiques avec d'autres variantes du SRAS-CoV-2, L'échantillon a été prélevé le 11 novembre 2021 au Botswana, la séquence du génome a été soumise le 23 novembre 2021 avec le numéro d'accès. EPI_ISL_6640916. Nous avons vérifié la présence de toutes les mutations de notre étude sur cette séquence

On ajouter à cette étude phylogénétique la séquence référence muté (avec les mutations étudiées), pour estimer la distance causé uniquement par les mutations de notre étude

L'approche UPGMA suppose que toutes les lignées évoluent au même rythme, et le taux de mutation n'est pas pris en compte lors de la construction de l'arbre

La distance évolutive entre une paire de séquences est généralement mesurée par le nombre de substitutions de nucléotides ou d'acides aminés entre elles. Les distances évolutives sont fondamentales pour l'étude de l'évolution moléculaire et sont utiles pour la reconstruction phylogénétique et l'estimation des temps de divergence

le tableau de distance par paires a révélé que le plus grand nombre de mutations a été enregistré avec la variante Omicron. Par rapport à d'autres variantes, le nombre de changements d'A.A par rapport a la séquence référence était dans l'ordre suivant :



WIV04 - Gamma – référence avec mutations -BETA - DELTA - ALPHA -OMICRON

On déduit à partir du tableau de distance par paires que les mutations de notre étude causent une distance de 0.01095 par rapport à la séquence référence de Wuhan

On déduit à partir du tableau de distance par paires que l'omicron a la plus grande distance par rapport à la séquence référence , une distance de 6.03577

L'arbre phylogénétique et Le tableau de distance par paires révèle que alpha est le plus identique au variant omicron.

conclusion

À la lumière du grand nombre de mutations observées dans la protéine de pointe du SARS-CoV-2, Omicron, une nouvelle variante du virus pose de sérieuses inquiétudes. Par conséquent, dans les travaux suivants, nous avons utilisé des méthodes de calcul in silico pour étudier les changements dynamiques de la protéine de pointe du SRAS-CoV-2 causés par 15 mutations précises, Les résultats de notre analyse ont révélé :

- L'importance de la bio-informatique dans le domaine clinique et la recherche scientifique
- la bio-informatique permet un très large domaine d'analyse (physicochimique, structurel, phylogénétique..)
- des changements importants dans la région RBD de la variante Omicron pourraient contribuer à une spécificité de liaison élevée avec hACE2, ce qui peut entraîner un taux de transmission plus élevé et un impact considérable sur la pathogenèse
- la charge moléculaire entre la séquence référence et la séquence du variant omicron est la différence la plus remarquable
- Notre étude ne répond pas à toutes les questions on relations avec ces 15 mutations chez le variant omicron, mais Cette étude sera une excellente opportunité pour la communauté scientifique pour toute enquête plus approfondie

Références

- Alouane, Tarek, Meriem Laamarti, Abdelmunim Essabbar, Mohammed Hakmi, El Mehdi Bouricha, M. W. Chemaou-Elfihri, Souad Kartti, et al. 2020. « Genomic Diversity and Hotspot Mutations in 30,983 SARS-CoV-2 Genomes: Moving Toward a Universal Vaccine for the “Confined Virus”? » *Pathogens (Basel, Switzerland)* 9 (10): E829. <https://doi.org/10.3390/pathogens9100829>.
- Chen, Nanshan, Min Zhou, Xuan Dong, Jieming Qu, Fengyun Gong, Yang Han, Yang Qiu, et al. 2020. « Epidemiological and Clinical Characteristics of 99 Cases of 2019 Novel Coronavirus Pneumonia in Wuhan, China: A Descriptive Study ». *Lancet (London, England)* 395 (10223): 507-13. [https://doi.org/10.1016/S0140-6736\(20\)30211-7](https://doi.org/10.1016/S0140-6736(20)30211-7).
- « Classification of Omicron (B.1.1.529): SARS-CoV-2 Variant of Concern ». s. d. [https://www.who.int/news/item/26-11-2021-classification-of-omicron-\(b.1.1.529\)-sars-cov-2-variant-of-concern](https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern).
- Cui, Jie, Fang Li, et Zheng-Li Shi. 2019. « Origin and Evolution of Pathogenic Coronaviruses ». *Nature Reviews. Microbiology* 17 (3): 181-92. <https://doi.org/10.1038/s41579-018-0118-9>.
- Dess, N. K., J. Patterson, et J. B. Overmier. 1986. « Relative Effectiveness of Concurrent Forward/Backward versus Simple Forward and Simple Backward Pavlovian Conditioning Procedures ». *The American Journal of Psychology* 99 (1): 31-44.
- Drobysch, Maryia, Almira Ramanaviciene, Roman Viter, Chien-Fu Chen, Urte Samukaite-Bubniene, Vilma Ratautaite, et Arunas Ramanavicius. 2022. « Biosensors for the Determination of SARS-CoV-2 Virus and Diagnosis of COVID-19 Infection ». *International Journal of Molecular Sciences* 23 (2): 666. <https://doi.org/10.3390/ijms23020666>.
- Duan, Liangwei, Qianqian Zheng, Hongxia Zhang, Yuna Niu, Yunwei Lou, et Hui Wang. 2020. « The SARS-CoV-2 Spike Glycoprotein Biosynthesis, Structure, Function, and Antigenicity: Implications for the Design of Spike-Based Vaccine Immunogens ». *Frontiers in Immunology* 11: 576622. <https://doi.org/10.3389/fimmu.2020.576622>.
- Fehr, Anthony R., et Stanley Perlman. 2015. « Coronaviruses: An Overview of Their Replication and Pathogenesis ». *Methods in Molecular Biology (Clifton, N.J.)* 1282: 1-23. https://doi.org/10.1007/978-1-4939-2438-7_1.
- Garcia-Beltran, Wilfredo F., Kerri J. St Denis, Angelique Hoelzemer, Evan C. Lam, Adam D. Nitido, Maegan L. Sheehan, Cristhian Berrios, et al. 2022. « mRNA-Based COVID-19 Vaccine Boosters Induce Neutralizing Immunity against SARS-CoV-2 Omicron Variant ». *Cell* 185 (3): 457-466.e4. <https://doi.org/10.1016/j.cell.2021.12.033>.
- He, X. et al. (2021) ‘SARS-CoV-2 Omicron variant: Characteristics and prevention’, *MedComm*, 2(4), pp. 838–845. Available at: <https://doi.org/10.1002/mco2.110>.
- He, Xuemei, Weiqi Hong, Xiangyu Pan, Guangwen Lu, et Xiawei Wei. 2021. « SARS-CoV-2 Omicron Variant: Characteristics and Prevention ». *MedComm* 2 (4): 838-45. <https://doi.org/10.1002/mco2.110>.
- Hoffmann, Markus, Hannah Kleine-Weber, Simon Schroeder, Nadine Krüger, Tanja Herrler, Sandra Erichsen, Tobias S. Schiergens, et al. 2020. « SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor ». *Cell* 181 (2): 271-280.e8. <https://doi.org/10.1016/j.cell.2020.02.052>.
- Huang, Yuan, Chan Yang, Xin-feng Xu, Wei Xu, et Shu-wen Liu. 2020. « Structural and Functional Properties of SARS-CoV-2 Spike Protein: Potential Antivirus Drug

- Development for COVID-19 ». *Acta Pharmacologica Sinica* 41 (9): 1141-49.
<https://doi.org/10.1038/s41401-020-0485-4>.
- Hui, David S., Esam I Azhar, Tariq A. Madani, Francine Ntoumi, Richard Kock, Osman Dar, Giuseppe Ippolito, et al. 2020. « The Continuing 2019-NCoV Epidemic Threat of Novel Coronaviruses to Global Health - The Latest 2019 Novel Coronavirus Outbreak in Wuhan, China ». *International Journal of Infectious Diseases: IJID: Official Publication of the International Society for Infectious Diseases* 91 (février): 264-66.
<https://doi.org/10.1016/j.ijid.2020.01.009>.
- Jaimes, Javier A., Jean K. Millet, et Gary R. Whittaker. 2020. « Proteolytic Cleavage of the SARS-CoV-2 Spike Protein and the Role of the Novel S1/S2 Site ». *IScience* 23 (6): 101212. <https://doi.org/10.1016/j.isci.2020.101212>.
- January 12, 2022. s. d. « Omicron, Delta, cold, or flu? The most common symptoms ». <https://fortune.com/2022/01/12/omicron-delta-cold-and-flu-symptoms-chart/>.
- Kolb, M. J., et W. M. Bourne. 1986. « Supravital Fluorescent Staining of the Corneal Endothelium with Acridine Orange and Ethidium Bromide ». *Current Eye Research* 5 (7): 485-94. <https://doi.org/10.3109/02713688608996370>.
- Korber, Bette, Will M. Fischer, Sandrasegaram Gnanakaran, Hyejin Yoon, James Theiler, Werner Abfalterer, Nick Hengartner, et al. 2020. « Tracking Changes in SARS-CoV-2 Spike: Evidence That D614G Increases Infectivity of the COVID-19 Virus ». *Cell* 182 (4): 812-827.e19. <https://doi.org/10.1016/j.cell.2020.06.043>.
- Kumar, Manoj, et Souhaila Al Khodor. 2020. « Pathophysiology and Treatment Strategies for COVID-19 ». *Journal of Translational Medicine* 18 (1): 353.
<https://doi.org/10.1186/s12967-020-02520-8>.
- Li, Shiqin, Sijia Li, Cyrollah Disoma, Rong Zheng, Mei Zhou, Aroona Razzaq, Pinjia Liu, et al. 2021. « SARS-CoV-2: Mechanism of Infection and Emerging Technologies for Future Prospects ». *Reviews in Medical Virology* 31 (2): e2168.
<https://doi.org/10.1002/rmv.2168>.
- Licastro, Danilo, Sreejith Rajasekharan, Simeone Dal Monego, Ludovica Segat, Pierlanfranco D'Agaro, et Alessandro Marcello. 2020. « Isolation and Full-Length Genome Characterization of SARS-CoV-2 from COVID-19 Cases in Northern Italy ». *Journal of Virology* 94 (11): e00543-20. <https://doi.org/10.1128/JVI.00543-20>.
- Lupala, Cecylia S., Yongjin Ye, Hong Chen, Xiao-Dong Su, et Haiguang Liu. 2022. « Mutations on RBD of SARS-CoV-2 Omicron Variant Result in Stronger Binding to Human ACE2 Receptor ». *Biochemical and Biophysical Research Communications* 590 (janvier): 34-41. <https://doi.org/10.1016/j.bbrc.2021.12.079>.
- Marcus Parekh, ;Poppie Platt, ;, Global Health Security Team, and, et Joe Barnes. s. d. « Coronavirus latest news: EU suspends all flights to southern Africa over omicron Covid variant fears ». *26 November 2021*.
- Mathieu, Edouard, Hannah Ritchie, Esteban Ortiz-Ospina, Max Roser, Joe Hasell, Cameron Appel, Charlie Giattino, et Lucas Rodés-Guirao. 2021. « A Global Database of COVID-19 Vaccinations ». *Nature Human Behaviour* 5 (7): 947-53.
<https://doi.org/10.1038/s41562-021-01122-8>.
- Mittal, Anshumali, Kavyashree Manjunath, Rajesh Kumar Ranjan, Sandeep Kaushik, Sujeet Kumar, et Vikash Verma. 2020. « COVID-19 Pandemic: Insights into Structure, Function, and HACE2 Receptor Recognition by SARS-CoV-2 ». *PLoS Pathogens* 16 (8): e1008762. <https://doi.org/10.1371/journal.ppat.1008762>.
- Ozili, Peterson K, et Thankom Arun. 2020. « Spillover of COVID-19: Impact on the Global Economy ». *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3562570>.

- Planas, Delphine, Nell Saunders, Piet Maes, Florence Guivel-Benhassine, Cyril Planchais, Julian Buchrieser, William-Henry Bolland, et al. 2022. « Considerable Escape of SARS-CoV-2 Omicron to Antibody Neutralization ». *Nature* 602 (7898): 671-75. <https://doi.org/10.1038/s41586-021-04389-z>.
- Shang, Jian, Gang Ye, Ke Shi, Yushun Wan, Chuming Luo, Hideki Aihara, Qibin Geng, Ashley Auerbach, et Fang Li. 2020. « Structural Basis of Receptor Recognition by SARS-CoV-2 ». *Nature* 581 (7807): 221-24. <https://doi.org/10.1038/s41586-020-2179-y>.
- Starr, Tyler N., Allison J. Greaney, Adam S. Dingens, et Jesse D. Bloom. 2021. « Complete Map of SARS-CoV-2 RBD Mutations That Escape the Monoclonal Antibody LY-CoV555 and Its Cocktail with LY-CoV016 ». *BioRxiv: The Preprint Server for Biology*, février, 2021.02.17.431683. <https://doi.org/10.1101/2021.02.17.431683>.
- Tregoning, John S., Katie E. Flight, Sophie L. Higham, Ziyin Wang, et Benjamin F. Pierce. 2021. « Progress of the COVID-19 Vaccine Effort: Viruses, Vaccines and Variants versus Efficacy, Effectiveness and Escape ». *Nature Reviews. Immunology* 21 (10): 626-36. <https://doi.org/10.1038/s41577-021-00592-1>.
- UniProt Consortium. 2021. « UniProt: The Universal Protein Knowledgebase in 2021 ». *Nucleic Acids Research* 49 (D1): D480-89. <https://doi.org/10.1093/nar/gkaa1100>.
- « Weekly epidemiological update on COVID-19 - ». s. d. 25 May 2022.
- Weiss, Susan R., et Julian L. Leibowitz. 2011. « Coronavirus Pathogenesis ». *Advances in Virus Research* 81: 85-164. <https://doi.org/10.1016/B978-0-12-385885-6.00009-2>.
- Wu, Aiping, Yousong Peng, Baoying Huang, Xiao Ding, Xianyue Wang, Peihua Niu, Jing Meng, et al. 2020. « Genome Composition and Divergence of the Novel Coronavirus (2019-NCoV) Originating in China ». *Cell Host & Microbe* 27 (3): 325-28. <https://doi.org/10.1016/j.chom.2020.02.001>.
- Wu, Canrong, Yang Liu, Yueying Yang, Peng Zhang, Wu Zhong, Yali Wang, Qiqi Wang, et al. 2020. « Analysis of Therapeutic Targets for SARS-CoV-2 and Discovery of Potential Drugs by Computational Methods ». *Acta Pharmaceutica Sinica. B* 10 (5): 766-88. <https://doi.org/10.1016/j.apsb.2020.02.008>.
- Wu, Joseph T, Kathy Leung, et Gabriel M Leung. 2020. « Nowcasting and Forecasting the Potential Domestic and International Spread of the 2019-NCoV Outbreak Originating in Wuhan, China: A Modelling Study ». *The Lancet* 395 (10225): 689-97. [https://doi.org/10.1016/S0140-6736\(20\)30260-9](https://doi.org/10.1016/S0140-6736(20)30260-9).
- Wu, Yi-Chi, Ching-Sung Chen, et Yu-Jiun Chan. 2020. « The Outbreak of COVID-19: An Overview ». *Journal of the Chinese Medical Association: JCMA* 83 (3): 217-20. <https://doi.org/10.1097/JCMA.0000000000000270>.
- Yan, Renhong, Yuanyuan Zhang, Yaning Li, Lu Xia, Yingying Guo, et Qiang Zhou. 2020. « Structural Basis for the Recognition of SARS-CoV-2 by Full-Length Human ACE2 ». *Science (New York, N.Y.)* 367 (6485): 1444-48. <https://doi.org/10.1126/science.abb2762>.
- Yang, Tzu-Jing, Pei-Yu Yu, Yuan-Chih Chang, Kang-Hao Liang, Hsian-Cheng Tso, Meng-Ru Ho, Wan-Yu Chen, Hsiu-Ting Lin, Han-Chung Wu, et Shang-Te Danny Hsu. 2021. « Impacts on the Structure-Function Relationship of SARS-CoV-2 Spike by B.1.1.7 Mutations ». Preprint. *Biochemistry*. <https://doi.org/10.1101/2021.05.11.443686>.

- Yesudhas, Dhanusha, Ambuj Srivastava, et M. Michael Gromiha. 2021. « COVID-19 Outbreak: History, Mechanism, Transmission, Structural Studies and Therapeutics ». *Infection* 49 (2): 199-213. <https://doi.org/10.1007/s15010-020-01516-2>.
- Yuan, Meng, Hejun Liu, Nicholas C. Wu, et Ian A. Wilson. 2021. « Recognition of the SARS-CoV-2 Receptor Binding Domain by Neutralizing Antibodies ». *Biochemical and Biophysical Research Communications* 538 (janvier): 192-203. <https://doi.org/10.1016/j.bbrc.2020.10.012>.
- Zhu, Xiaojie, Qi Liu, Lanying Du, Lu Lu, et Shibo Jiang. 2013. « Receptor-Binding Domain as a Target for Developing SARS Vaccines ». *Journal of Thoracic Disease* 5 Suppl 2 (août): S142-148. <https://doi.org/10.3978/j.issn.2072-1439.2013.06.06>.
- Zimmerman, Ricardo Ariel, Patrícia Aline Gröhs Ferrareze, Flavio Adsuara Cadegiani, Carlos Gustavo Wambier, Daniel do Nascimento Fonseca, Andrea Roberto de Souza, Andy Goren, Liane Nanci Rotta, Zhihua Ren, et Claudia Elizabeth Thompson. 2022. « Comparative Genomics and Characterization of SARS-CoV-2 P.1 (Gamma) Variant of Concern From Amazonas, Brazil ». *Frontiers in Medicine* 9 (février): 806611. <https://doi.org/10.3389/fmed.2022.806611>.