



الجمهورية الجزائرية الديمقراطية الشعبية

**République Algérienne Démocratique et Populaire**

**N série:**

وزارة التعليم العالي والبحث العلمي

**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique**

جامعة الشهيد حممة لخضر الوادي

**Université Echahid Hamma Lakhdar - El OUED**

كلية علوم الطبيعة والحياة

**Faculté des Sciences de la Nature et de la Vie**

قسم البيولوجيا الخلوية والجزيئية

**Département de biologie Cellulaire et Moléculaire**

## MEMOIRE DE FIN D'ETUDE

En vue de l'obtention du diplôme de Master Académique en Sciences  
biologiques

Spécialité : Toxicologie

### THEME

***Revue de la littérature sur le Covid-19 à l'aide  
d'outils bioinformatiques***

Présenté Par :

M<sup>elle</sup> Lechelah Rania

M<sup>elle</sup> Ghenabzia Abir

Devant le jury composé de :

**Présidente :** M<sup>me</sup> Nadji Nassima

M.A.A, Université d'El Oued.

**Examinatrice :** M<sup>me</sup> Allouch Djennat

M.A.A, Université d'El Oued.

**Promoteur :** Dr.BENALI Abdelhai

M.C.B, Université d'El Oued.

Année universitaire 2020/2021



## **Remerciement**

*Premièrement et dernièrement, tout le remercie à Dieu qui nous a donné la patience,  
le courage et la force pour réaliser ce travail.*

*Nous remercions **Dr.BENALI Abdelhai** notre encadreur de mémoire qui nous a  
permis de réaliser ce travail dans les meilleures conditions ...*

*Merci aussi pour nous avoir fait partager votre expérience et votre culture  
scientifique et pour votre confiance.*

*Travailler sous votre direction a été un plaisir et un honneur.*


*Nos sincères remerciements s'adressent également aux membres de jury :*

***Mme NADJI Nassima** ; Maitre de Conférence "A" à l'Université d'El Oued  
d'avoir accepté de présider la soutenance.*

***Mme ALIOUCHE Djanat** , Maitre Assistant "A" à l'Université  
d'El Oued d'avoir accepté d'examiner notre travail.*

*Nos vifs remerciements et notre gratitude à tous les enseignants du  
Département des Sciences de la Nature et de Vie de l'Université EChahid HAMMA  
LAKHDAR, El-Oued.*

*Merci à tout les personnes qui nous ont aidés à arrive à ce stade.*





## *Dédicaces*

*Au nom du dieu le clément et le miséricordieux louange à ALLAH le tout puissant.  
Je dédie ce modeste travail à l'âme de très chère mère « **Zaara** » A toi ma regrettée maman  
à celle qui m'a donné la vie, le symbole de tendresse, qui s'est sacrifiée pour mon bonheur  
et ma réussite,*

*malgré ton affection incomparable, le monde a été Jaloux de toi ma mère,  
tu as semé ce que tu ne devait pas récolter alors que tu m'encourager que j'aie  
plus loin que possible dans mes études mais de qui le sort m'a séparé  
prématurément avant même mon premier couronnement universitaire  
et en fin tu as disparue de ce monde ingrat d'amour.*

*Quoique je puisse dire et écrire, je ne pourrais exprimer ma grande,  
affection et ma profonde reconnaissance. J'espère ne jamais te décevoir.*

*« **L'amour d'une maman, c'est le baume miracle qui guérit toutes les blessures,  
maman est un porte-bonheur pour la vie** »*

*Paix à son âme qu'ALLAH lui fasse Miséricorde et lui ouvre les portes du paradis, Amine  
A mon très cher père « **Abdelsalam** » qui a été mon ombre durant toutes les années des études,  
et qui a veillé tout au long de ma vie à m'encourager, à me donner l'aide et à me protéger,  
Que dieu le garde et le protège.*

*A mes adorables sœurs :**Hanane, Aicha, Ouidad, Safa, Maroua***

*A mes chers frères :**Mohammed Ali, Ahmed Faouzi***

*A mes amies et A tous ceux qui me sont chères*

*A tous ceux qui m'aiment. A tous ceux que j'aime*

*Louange à Dieu tout puissant, qui m'a permis de voir ce jour tant attendu.*



***Lechelah Rania***



## ***Dédicaces***

*ALLAH , les moments ne peuvent être beaux que par ton mentionner et tes  
Remerciements, et au-delà ne peut être qu' avec ton pardon, et le paradis est pas  
Merveilleux sauf pour te voir.*

*Je dédie ce travail:*

*A qui Dieu a donné le prestige et le dignité, à qui porte fièrement son nom, vos paroles  
resteront les étoiles guidées aujourd'hui, demain et pour toujours, au grand cœur,*

*Mon cher père **GHENABZIA MILOUD***

*A mon ange dans la vie, au sens de amour et de la tendresse, au sourire de la vie et au  
secret de l'existence, à qui sa prière était le secret de mon succès,*

*à ma bien-aimée mère **GUENOUA ZINEB***

*A mon soutien, ma force et mon refuge après Dieu, aux les cœurs purs, mes frères*

***TAHA et MED LARBI et MED BACHIR***

*Et mes sœurs **SOUMIA et HANAN et SIHAM et MOUNA***

*Mon fiancé **OUSSAMA** est cher a mon cœur*

*A ma copine **ABRAR***

***Ghenabzia Abir***



## **Abstract**

SARS-CoV-2 is a novel coronavirus identified as the cause of the 2019 coronavirus disease (COVID-19) which started in Wuhan, China at the end of 2019 and has spread around the world. Coronaviruses encode four major structural proteins, namely spike (S), membrane (M), envelope (E) and nucleocapsid (N).

Our current global situation requires an urgent need for rapid vaccine development while adhering to strict SARS-CoV-2 vaccine safety guidelines. While there is strong optimism for late stage vaccine candidates and imminent regulatory approval, the field should also be prepared for a potential mutation of SARS-CoV-2 as its seasonal recurrences. In addition, we need to develop countermeasures against other emerging pathogens. In particular, the structural fragility and suboptimal immunogenicity of many vaccine candidates should be addressed.

The NCBI houses a series of biotechnology and biomedicine databases and is an important resource for bioinformatics tools and services. A lot of important information is provided by the GenBank format such as sequence length, GenBank division, accession number. In this work, we used the BLASTN program to identify and characterize the sequence of SARS-CoV-2 compared to the sequences of SARS-CoV and Bat-CoV; RaTG13.

From the alignment by BLASTN, we can note that a sequence identity at the nucleotide level equal to 82.3% for SARS-CoV and 92.96% for Bat-CoV; RaTG13. Also, the sequence of Bat-CoV; RaTG13 matches our SARS-CoV-2 sequence better than the SARS-CoV sequence. Our work and previous studies of the literature confirm that the proteins SARS-CoV-2 and Bat-CoV; RaTG13 share high identity and similarity to SARS-CoV-2 and SARS-CoV proteins.

**Keywords:**SARS-CoV-2,SARS-CoV,Bat-CoV,RaTG13,NCBI.

## Résumé

Le SARS-CoV-2 est un nouveau coronavirus identifié comme la cause de la maladie à coronavirus de 2019 (COVID-19) qui a débuté à Wuhan, en Chine à la fin de 2019 et s'est propagé dans le monde entier. Les coronavirus codent pour quatre protéines structurales majeures, à savoir la pointe (S), la membrane (M), l'enveloppe (E) et la nucléocapside (N). Notre situation mondiale actuelle nécessite un besoin urgent de développement rapide de vaccins tout en respectant des directives strictes la sécurité des vaccins contre le SRAS-CoV-2.

Bien qu'il y ait un fort optimisme pour candidats vaccins de stade avancé et approbation réglementaire imminente, le sur le terrain doit également être préparé pour une mutation potentielle du SRAS-CoV-2 comme ses récurrences saisonnières. De plus, nous devons développer des contre-mesures contre d'autres agents pathogènes émergents. En particulier, la fragilité structurelle et immunogénicité sous-optimale de nombreux vaccins Les candidats doivent être adressés.

Le NCBI abrite une série de bases de données relatives à la biotechnologie et à la biomédecine et constitue une ressource importante pour les outils et services bioinformatiques.

De nombreuses informations importantes sont fourni par le format GenBank telles que la longueur de la séquence, la division GenBank, le numéro d'accession. Dans ce travail, nous avons utilisé le programme BLASTN pour identifier et caractériser la séquence du SARS-CoV-2 par rapport aux séquences du SARS-CoV et du Bat-CoV; RaTG13.

D'après l'alignement par BLASTN, nous pouvons noter qu'une identité de séquence au niveau nucléotidique égal à 82,3% pour le SRAS- CoV et 92,96 % pour le Bat-CoV ; RaTG13. Aussi, la séquence de Bat-CoV; RaTG13 correspond mieux à notre séquence du SRAS-CoV-2 qu'à la séquence du SRAS-CoV. Notre travail avec les études antérieures de la littérature confirment que les protéines SARS-CoV-2 et Bat-CoV; RaTG13 partagent une identité et une similitude élevées par rapport aux protéines SARS-CoV-2 et SARS-CoV.

**Mots-clés :** SARS-CoV-2,SARS-CoV,Bat-CoV,RaTG13,NCBI.

## الملخص

SARS-CoV-2 هو فيروس كورونا جديد تم تحديده على أنه سبب مرض فيروس كورونا 2019 (COVID-19) الذي بدأ في مدينة ووهان بالصين في نهاية عام 2019، وانتشر في جميع أنحاء العالم. يقوم فيروس كورونا بتشفير أربعة بروتينات هيكلية رئيسية ، وهي الطرف (S) والغشاء (M) المغلف (E) والنيوكليوكابسيد (N).

يتطلب وضعنا العالمي الحالي حاجة ملحة لتطوير لقاح سريع مع الالتزام بإرشادات سلامة لقاح SARS-CoV-2 الصارمة. في حين أن هناك تفاؤلاً قوياً بالمرحلة المتأخرة من اللقاحات والموافقة التنظيمية الوشيكة ، يجب أيضاً أن نكون مستعدين لطفرة محتملة في فيروس SARS-CoV-2 كتكرار موسمي. بالإضافة إلى ذلك ، نحن بحاجة إلى تطوير تدابير مضادة ضد مسببات الأمراض الناشئة الأخرى. على وجه الخصوص ، ينبغي معالجة الهشاشة الهيكلية والمناعة دون المستوى الأمثل للعديد من اللقاحات المرشحة.

يضم NCBI سلسلة من قواعد بيانات التكنولوجيا الحيوية والطب الحيوي وهو مورد مهم لأدوات وخدمات المعلوماتية الحيوية. يتم توفير الكثير من المعلومات المهمة من خلال تنسيق GenBank مثل طول التسلسل وقسمه ورقم الانضمام. في هذا العمل ، استخدمنا برنامج BLASTN لتحديد وتوصيف تسلسل SARS-CoV-2 مقارنة بتسلسلات SARS-CoV و Bat-CoV و RaTG13.

من خلال المحاذاة بواسطة BLASTN ، يمكننا ملاحظة أن نسبة التطابق (هوية التسلسل) على مستوى النوكليوتيدات تساوي 82.3% لـ SARS-CoV و 92.96% لـ Bat-CoV- RaTG13. أيضاً ، تسلسل Bat-CoV- RaTG13 ؛ يطابق تسلسل SARS-CoV-2 بشكل أفضل من تسلسل SARS-CoV. يؤكد عملنا بالإضافة للدراسات السابقة للأدبيات أن بروتينات SARS-CoV-2 و Bat-CoV- RaTG13 ؛ تشترك في هوية عالية وتشابه مع بروتينات مقارنة مع SARS-CoV.

الكلمات المفتاحية : SARS-CoV-2,SARS-CoV,Bat-CoV,RaTG13,NCBI .

## *Liste des abréviations*

- OMS** : L'*Organisation mondiale de la santé*
- SARS** : *Severe acute respiratory syndrome*
- CoV** : coronavirus
- ICTV** : The *International Committee on Taxonomy of Viruses*
- Kb** : kilobyte
- ARN** : acide ribonucléique
- Nm** : nanomètre
- SDRA** : *Syndrome de détresse respiratoire aiguë*
- PEDv** : *Porcine epidemic diarrhea virus*
- FIPV** : Fractional doses of the inactivated poliovirus vaccine
- MHV** : Mouse Hepatitis Virus
- °C** : Température
- PCR** : *Polymerase Chain Reaction*
- TDR** : *Test de diagnostic rapide*
- TMD** : Emporomandibular disorders
- Mm** : *Micromètre*
- INSPQ** : *Institut national de santé publique du Québec*
- AVC** : *Accident vasculaire cérébral*
- ARNm** : Acide ribonucléique messenger
- UTRs** : Untranslated regions
- GNRA** : La tétraboucle GNRA a une paire de bases guanine-adénine
- VIH** : Virus de l'immunodéficience humaine
- Å** : Angstrom
- ARNr** : Acide ribonucléique ribosomique
- IBV** : *Infectious Bronchitis Virus*
- FCoV** : Feline Coronavirus
- ACE2** : *Angiotensin-converting enzyme 2*
- ORF** : open reading frame
- IFN** : nterférons
- IDR** : Une intradermoréaction
- NSP** : neuroendocrine-specific protein
- RBD** : Receptor Binding Domain
- RaTG13** : Bat coronavirus
- CRP** : La protéine C-réactive
- IL** : Les *interleukines*

**TNF** : Les facteurs de nécrose tumorale

**CDC** : *Centers for Disease Control and Prevention*

**LAMP** : Loop-mediated Isothermal Amplification

**ADN** : Acide désoxyribonucléique

**IgM** : Immunoglobulines M

**IgG** : Immunoglobulines G

**IgA** : Immunoglobulines A

**ELISA** : Enzyme-linked immunosorbent assay

**NGS** : Séquençage nouvelle génération

**FDA** : *Food and Drug Administration*

**HCQ** : Hydroxychloroquine

**IMP** : *Institute of Molecular Pathology*

**FFP** : *Fresh frozen plasma*

**PRRs** : *Porcine reproductive and respiratory syndrome*

**TLRs** : Les récepteurs Toll-like

**RIG-I** : Retinoic acid-inducible gene I

**MDA5** : Melanoma differentiation-associated protein 5

**IRFs** : Interferon regulatory factors

**MCP1** : *Monocyte chemoattractant protein-1*

**NK** : Les cellules *natural killer*

**CD** : Cluster de différenciation

**NKG2A** : Natural Killer Cell Receptor

**TCR** : Le récepteur des cellules T

**HLA** : *The human leukocyte antigen*

**AICD** : Automatic implantable cardioverter-defibrillator

**MIS-C** : Multisystem inflammatory syndrome in children

**DABK** : Analysis of responses to kallidin

**ROS** : *Reactive oxygen species*

**MBL** : *The Marine Biological Laboratory*

**MASP** : Mannose-associated serine protease 1

**TFPI** : *Le Tissue factor pathway inhibitor*

**Th** : Les lymphocytes T auxiliaires

**PEG** : Polyethylene glycol

**DOTMA** : 1,2-di-O-octadecenyl-3-trimethylammonium propane

**DOTAP** : 1,2-dioleoyl-3-trimethylammonium-propane

**PBAE** : Poly-beta-amino-esters

**DSPC** : Distearoylphosphatidylcholine

**PSA** : L'antigène spécifique de prostate

**CMC** : Carboxyméthylcellulose

**VLP** : *Virus-like particles*

**CMH** : *Le complexe majeur d'histocompatibilité*

**SC** : Sous-cutanée

**IN** : Intramusculaire

**IV** : Intraveineuse

**ID** : Intradermique

**IM** : Intramusculaire

**IT** : Intratrachéale

**EGFP** : *The enhanced green fluorescent protein*

**NCBI** : *The National Center for Biotechnology Information*

**BLAST** : *The Basic Local Alignment Search Tool*

**blastn** : Nucleotide BLAST

**WIV** : *Wuhan Institute of Virology*

**BALF** : Bronchoalveolar lavage fluid

**hACE2** : Transgenic mouse model for coronavirus

**HEK** : Human embryonic kidney 293 cells

**Ph** : *Le potentiel hydrogène*

# Liste des figures

## Liste des figures

Figure	Titre	Page
I.1	Modélisation du Corona Virus	7
I.2	Les symptômes de la Covid-19	8
I.3	Le scanner ( Imagerie) thoracique du covid 19	13
I.4	La transmission du cycle de SRAS-COV-2	14
I.5	Transmission et cycle de vie du SARS-CoV-2.	16
I.6	L'importance de isolement et la quarantaine	17
I.7	Courbes des Tendances quotidiennes et nouvelles contamination	18
I.8	Comparaison statistiques entre l'Algérie avec le reste du monde	19
I.9	La période d'incubation de la maladie à coronavirus 2019 (COVID-19) à partir de cas confirmés déclarés publiquement : estimation et application	21
II.1	Les structures primaire, secondaire et tertiaire de l'ARN s2m du SRAS	26
II.2	Représentations stéréo de la structure de l'ARN s2m du SRAS	29
II.3	Interactions structurelles tertiaires dans l'ARN s2m du SRAS	30
II.4	Structure de SARS-CoV-2	34
II.5	Organisation structurelle et génomique complète du coronavirus 2 du syndrome respiratoire aigu sévère (SRAS-CoV-2).	36
II.6	Présentation schématique de la structure du génome du SRAS-CoV-2.	41
II.7	Diagrammes de dessins animés illustrant trois orientations du monomère de pointe coloré	44
II.8	Structures globales des glycoprotéines de pointe RaTG13 et RBD	45
II.9	Affinités de liaison et interactions de l'ACE2 humain avec différents RBD	46
III.1	Pathogenèse et résultats de COVID-19	51
III.1	Courbe de valeur S polyS neutrophile et des lymphocyte au cours des infection sars-cov-2	58
III.2	Effet de l'infection par le SRAS-CoV-2 sur la fonction des cellules endothéliales, la coagulation systémique et la thrombose	64
III.3	Les effets du COVID-19 sur le rein	70
IV.1	Diagnostic protocol recommended for COVID-19	83

## Liste des figures

<b>IV.2</b>	Le mécanisme d'action de différentes thérapeutiques contre le COVID-19.	83
<b>V.1</b>	Various vaccine platforms for vaccination against SARS-CoV-2	102
<b>V.2</b>	Réponse immunitaire humorale et cellulaire des ARNm-1273 et BTN-162b1.	109
<b>V.3</b>	mRNA vaccine delivery systems applicable for COVID-19 vaccine development.	110
<b>V.4</b>	Le vaccin à ADN INO-4800 d'Inovio Pharmaceutical.	113
<b>V.5</b>	Réponses immunitaires humorales et cellulaires observées	116
<b>V.6</b>	Courbe du les personnes vaccinées en Algérie	124
<b>VI.1</b>	COMPROTEIN, le premier logiciel de bioinformatique	127
<b>VI.2</b>	Représentation de l'algorithme d'alignement global Needleman-Wunsch	128
<b>VI.3</b>	Une station de travail de bureau HP-9000 exécutant le système basé sur Unix HP-UX	129
<b>VI.4</b>	Réseau de séquenceurs capillaires ABI 373 au NHGRI (1993)	130
<b>VI.5</b>	Quantité totale de séquences sur les bases de données NCBI GenBank et WGS au fil du temps.	131
<b>VI.6</b>	Recherche sur Covid19 à partir de la base de données Nucleotide (14 juin 2021).	134
<b>VI.7</b>	Section Header au format GenBank du SARS-CoV-2 (14 juin 2021).	135
<b>IV.8</b>	Sections Caractéristiques et Origine au format GenBank du SARS-CoV-2 (14 juin 2021).	138
<b>IV.9</b>	Format FASTA du SARS-CoV-2 (14 juin 2021).	141
<b>IV.10</b>	Format graphique du SARS-CoV-2 (14 juin 2021).	142
<b>VII.1</b>	Les différents programmes BLAST disponibles via la page d'accueil du serveur Web NCBI.	145
<b>VII.2</b>	Mise en place de notre recherche BLASTN de la séquence SARS-CoV-2	146
<b>VII.3</b>	Les descriptions statistiques font partie de BLASTN entre le SARS-CoV-2	148
<b>VII.4</b>	Partie récapitulative graphique de BLASTN entre le SARS-CoV-2	149
<b>VII.5</b>	Une partie de l'alignement BLASTN entre le SARS-CoV-2	150
<b>VII.6</b>	Dot Plot partie de BLASTN entre le SARS-CoV-2	154

## Liste des figures

---

<b>VII.7</b>	Caractérisation du génome du 2019-nCoV	156
<b>VII.8</b>	Bat SL-CoV RaTG13 utilise hACE2 et RaACE2 pour l'entrée du virus	160
<b>VII.9</b>	Alignment of partial amino acid sequences of RaTG13 and SARS-CoV-2 S proteins	162

## Liste des tableaux

### *Liste des Tableaux*

<b>Tableau</b>	<b>Titre</b>	<b>Page</b>
<b>I.1</b>	Classification de Coronavirus	11
<b>I.2</b>	Statistique du nombre de cas, de décès et de guérisons, liés au coronavirus (COVID-19) en Algérie	17
<b>II.1</b>	Homologie de séquence entre le SARS-CoV-2 et d'autres souches de coronavirus	37
<b>II.2</b>	Pourcentage d'identité entre les protéines du SRAS-CoV-2 et de la famille des Coronaviridae	39
<b>II.3</b>	Fonctions des protéines non structurales du SARS-CoV-2	48
<b>V.1</b>	Vaccins non viraux COVID19 actuellement en cours de développement	103
<b>VII.1</b>	Les protéines étudiées des trois virus	164
<b>VII.2</b>	Les pourcentages d'identité et de similarité des protéines du SARS-CoV-2 par rapport à ceux du SARSCoV et du RaTG13 (coronavirus de chauve-souris)	165

## *SOMMAIRE*

<b>Dédicaces</b>	
<b>Remerciements</b>	
<b>Résumé</b>	
<b>Liste des figures</b>	
<b>Liste des tableaux</b>	
<b>Introduction</b>	1
<b>Chapitre I : Généralité sur le Covid 19</b>	
<b>I.1. Définition de covid 19</b>	6
<b>I.2. Les symptômes de la Covid-19</b>	7
I.2.1. Le temps met-on à ressentir les symptômes	8
I.2.2. Durée de la maladie	9
<b>I.3. Propriétés générale de Covid19</b>	9
I.3.1. Classification	10
I.3.2. Caractères physico-chimiques	11
I.3.3. Les Teste sur Covid-19	11
<b>I.4. Mode de transmission</b>	13
I.4.1. Transmission par aerosols	13
I.4.2. Transmission indirecte	14
I.4.3. La transmission par les selles et autres liquides biologiques	15
I.4.4. La transmission de la mère à l'enfant	15
I.4.5. La transmission zoonotique	15
I.4.5. L'isolement et la quarantaine	16
<b>I.5. Données statistique</b>	17

## sommaire

I.5.1. Comparaison de l'Algérie avec le reste du monde	18
I.5.2. Taux de reproduction et facteur de dispersion	19
<b>I.6. Période d'incubation</b>	20
<b>I.7. Contagiosité</b>	21
<b>I.8. Présentations cliniques</b>	22
<b>I.9. Complications associées à la COVID-19</b>	22
I.9.1. Complications cardiovasculaires	22
I.9.2. Complications pulmonaires	22
I.9.3. Complications neurologiques	23
I.9.4. Autres complications possibles	23
<b>Chapitre II: Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV</b>	
<b>II.1. SARS-CoV (syndrome respiratoire aigu sévère)</b>	25
II.2.1. Les structures de l'ARN s2m du SRAS	25
II.2.1.1. Les structures primaire, secondaire et tertiaire	26
II.2.2. Analyse de séquence de l'élément s2m conservé	28
II.2.3. La structure cristalline de l'élément ARN s2m du SRAS	28
<b>II.2. SARS-CoV-2</b>	30
II.2.1. Structure de SARS CoV-2	32
II.2.1.1. Glycoprotéine S	31
II.2.1.2. Protéine M	32
II.2.1.3. Protéine E	33
II.2.1.4. Protéine N	33
II.2.1.5. nsps et protéines accessoires	33
II.2.2. Aperçu de l'organisation génomique	34
II.2.3. Séquençage du génome	35
II-2.3.1 Protéines Conservées	36
II.2.3.2 Domaine de liaison au récepteur (RBD)	38
II.2.3.3. Organisation génomique du SRAS - CoV2	39

## sommaire

II.2.4 Mutations récentes du SRAS-CoV-2	41
<b>II.3. bat coronavirus RATG13 (bat COV RATG13)</b>	43
II.3.1-structure BAT COV RATG13	43
<b>Chapitre III: La réponse immunitaire covid-1920</b>	
<b>III.1. La réponse immunitaire SARS-CoV-2</b>	50
III.1.1. Réponse immunitaire innée	50
III.1.1.1 Mécanisme d'action	52
III.1.2.Réponse immunitaire adaptative	54
<b>III.2.Effets du SARS-CoV-2 sur l'endothélium</b>	61
<b>III.3.Activation excessive des suppléments</b>	66
<b>III.4.Microangiopathie et thrombose vasculaire</b>	68
<b>III.5.Implication des reins dans COVID-19</b>	69
<b>Chapitre IV:Diagnostic et thérapeutique sur SARS COV-2</b>	
<b>IV.1Diagnostic COVID – 19</b>	75
IV.1.1.Présentation clinique	75
IV.1.2.-Tests de dépistage non spécifiques du COVID - 19 chez les patients exposés	76
IV.1.3.Résultats radiologiques	77
IV.1.4.Diagnostic moléculaire	78
IV.1.4.1.Détection d'acide nucléique	79
IV.1.4.2.Technique basée sur les puces à ADN	81
IV.1.5.Diagnostic immunologique	82
IV.1.6.Nouvelles techniques	84
IV.1.6.1.CRISPR technique	84
IV.1.6.2.Technique basée sur la LAMPE	85
IV.1.6.3.Technologie Penn RAMP	87
IV.1.6.4. PCR numérique de gouttelettes	87
IV.1.6.5. Technique basée sur le séquençage de nouvelle génération (NGS)	88
<b>IV.2. Thérapeutique COVID-19</b>	79

## sommaire

IV.2.1.Chloroquine et Hydroxychloroquine	90
IV.2.2.Remdesivir	92
IV.2.3.Dexaméthasone	93
IV.2.4.Inhibiteurs Ivermectin, 3CLpro et PLpro	93
IV.2.5.Famotidine	94
IV.2.6.Inhibiteurs TMPRSS2 et furine	94
IV.2.7.Plasma de convalescence et traitement par anticorps ciblés	95
IV.2.8. Thérapies immunitaires	97
<b>Chapitre V: Vaccins contre le covid-19</b>	
<b>V.1 Vaccins ANm</b>	104
V.1.1 Vaccins ARNm au stade clinique	104
V.1.2 Considérations de conception pour les vaccins à ARNm	106
<b>V.2 Vaccins ADN</b>	110
V.2.1 Vaccins ADN de stade clinique	110
1.2.2 Stratégies alternatives pour l'administration de vaccins à ADN :	114
<b>V.3 Vaccins protéiques</b>	115
1.4 Vaccins contre les particules de type virus (VLP)	118
1.5 Vaccins à base de peptides	119
1.6 Perspectives futures des systèmes de distribution de vaccins :	119
1.6.1Systèmes d'administration adjuvants :	119
1.6.2 Voies de vaccination	120
1.6.3 Stabilité thermique	122
1.6.4 Administration ciblée des vaccins sur les ganglions lymphatiques	123
1.7. Statistiques et recherche Vaccins contre le coronavirus (COVID-19)	123
<b>Chapitre VI:Étude sur COVID-19 en utilisant la bioinformatique</b>	
<b>VI.1. Une brève histoire de la bioinformatique</b>	126
VI .1.1. Origines de la bioinformatique (1950 – 1970)	127

## sommaire

---

VI.1.2. Changement de paradigme de l'analyse des protéines à l'analyse de l'ADN (1970-1980)	128
VI.1.3. Avancées parallèles en biologie et en informatique (1980-1990)	129
VI.1.4. Génomique et autoroute de l'information (1990-2000)	129
VI.1.5. Bioinformatique à haut débit (2000-2010)	130
VI.1.6. Perspectives présentes et futures (2010-Aujourd'hui)	131
<b>VI.2. Centre national d'information sur la biotechnologie (NCBI)</b>	132
VI.2.1. introduction	132
VI.2.2. Ressources populaires de NCBI	132
<b>VI.3. Base de données nucléotidiques</b>	133
VI.3.1. Format GenBank	134
VI.3.1.1. Section HEADER	136
VI.3.1.2. Section FEATURES	139
VI.3.1.3. Partie ORIGINE	140
VI.3.2. Format FASTA	140
VI.3.3. Format Graphic (Graphical Sequence Panel)	141
<b>Chapter VII: Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN</b>	
<b>VII.1. Introduction</b>	144
<b>VII.2. L'interface Web NCBI BLAST</b>	145
<b>VII.3. Compararison entre SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN</b>	147
VII.3.1. Descriptions	149
<b>VII.4. Graphic Summary</b>	150
<b>VII.5. Alignements</b>	150

## **sommaire**

---

<b>VII.6. Méthode Dot Plot (Méthode graphique)</b>	153
<b>VII.7. Études antérieures de la littérature sur le SARS-CoV-2 (en 2020 et 2021)</b>	155
<b>Conclusion générale</b>	
<b>Références bibliographiques</b>	
<b>Résumé et mots-clés</b>	

# **Introduction générale**

## Introduction generale

---

Les coronavirus (CoV) font partie d'une grande famille de virus à ARN à un seul brin, enveloppés, touchant à la fois l'homme et différentes espèces animales.

Un nouveau coronavirus responsable du syndrome respiratoire aigu sévère (SARS-CoV-2). Le SARS-CoV-2 appartient à la famille des coronavirus (CoV), dénomination liée à la « couronne » que forment certaines protéines à la surface de ces virus. Il a été identifié pour la première fois à Wuhan en Chine, en décembre 2019. Plusieurs coronavirus sont déjà connus pour être capables d'infecter les humains : trois coronavirus saisonniers responsables de symptômes hivernaux sans gravité (rhumes), le SARS-CoV responsable du syndrome respiratoire aigu sévère (SRAS) et le MERS-CoV responsable d'une atteinte respiratoire potentiellement sévère (Middle East Respiratory Syndrome).

Dans cette mémoire, nous avons fait une lecture charnière sur le virus du SRAS, en étudiant ses caractéristiques et en identifiant ses composants. Nous avons également appliqué les bases de données biotechnologiques et biomédicales, qui constituent une ressource importante pour les outils et services bioinformatiques.

Le premier chapitre traitait de la définition du SRAS CoV-2, de ses symptômes, méthodes de prévention, de sa classification, des modes de transmission et des méthodes de dépistage. (Test rapide .PCR. Le scanner thoracique) Isolement et quarantaine pour nous protéger **Complications associées à la COVID-19** cardiovasculaires pulmonaires neurologiques.

Nous attendons avec impatience le chapitre 2 sur, **Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV, Pour cette étude, nous avons pris 3 gènes très similaires.**

En 2<sup>ème</sup> chapitre montrer les coronavirus codent pour quatre protéines structurelles majeures, à savoir la pointe (**S**), la membrane (**M**), l'enveloppe (**E**) et la nucléocapside (**N**), , Ce sont des virus à ARN de sens positif avec une taille de génome d'environ de 30 kpb

L'ARN du virus contient 6 cadres de lecture ouverts (ORF1ab, ORF3a, ORF6, ORF7ab, ORF8, et ORF10).

En 3<sup>ème</sup> chapitre nous présentons la réponse immunitaire au SARS-CoV-2 la réponse immunitaire innée et adaptative joue un rôle essentiel dans la reconnaissance et l'élimination des agents pathogènes étrangers. L'accumulation de preuves suggère qu'une réponse immunitaire sous-optimale ou non maîtrisée pendant l'infection par le

## Introduction generale

---

SRAS-CoV-2 détermine les schémas cliniques, la gravité de la maladie et la progression de COVID-19.

L'identification de signatures immunitaires spécifiques peut fournir de nouvelles informations sur les trajectoires divergentes de la maladie observées chez des patients individuels et aider à la conception d'approches thérapeutiques plus ciblées.

En 4<sup>ème</sup> chapitre nous présentons l'importance de le diagnostic précoce et l'isolement des patients suspects jouent un rôle essentiel dans le contrôle de cette épidémie Covid -19.

Le 5<sup>ème</sup> chapitre concerne le processus de développement d'un vaccin comprend deux étapes clés:

-identifier un antigène et -développer une approche de délivrance pour le dit antigène pour obtenir une immunité cellulaire et humorale robuste.

BioNTech /Pfizer et Moderna ont obtenu une autorisation d'urgence pour leurs vaccins COVID-19 par la Food and Drug Administration (FDA).

il existe un vaccin à base d'adénovirus chimpanzé appelé ChAdOx1 développé par l'Université d'Oxford et AstraZeneca ainsi que Adenovirus 26 vaccins à base de vecteurs de Johnson & Johnson qui provoquent de puissants immuns réponses.

Il y a quatre type de vaccins : Vaccins ARNm, Vaccins ADN ,Vaccins protéiques,

Vaccins contre les particules de type virus Vaccins à base de peptides.

Le chapitre 6 commence par montrer un bref historique de la bioinformatique et donne des informations sur le National Center for Biotechnology Information (NCBI) et leurs ressources populaires. Ensuite, nous avons effectué une recherche dans la base de données Nucleotide sur le SRAS-CoV-2 et donné un détail sur chaque format (GenBank, FASTA et Graphics), nous avons également effectué une comparaison entre le virus SARS-CoV-2 contre les virus SRAS et Bat-CoV; RaTG13 à l'aide de l'outil BLAST.

Dans Le 7<sup>ème</sup> chapitre, nous présentons un examen rapide de l'interface Web NCBI BLAST, puis présentons l'alignement de séquences effectué entre l'ensemble du génome du SARS-CoV-2 par rapport au SARS-CoV et au Bat-CoV ; RaTG13 utilisant le nucléotide-nucléotide BLAST (blastn). En outre, nous incluons certaines

## **Introduction generale**

---

des études importantes trouvées dans la littérature publiée en 2020 et 2021, qui concernent le SARS-CoV-2, le SARS-CoV et le Bat-CoV ; virus RaTG13.

# **Chapitre I**

## **Généralité sur le Covid 19**

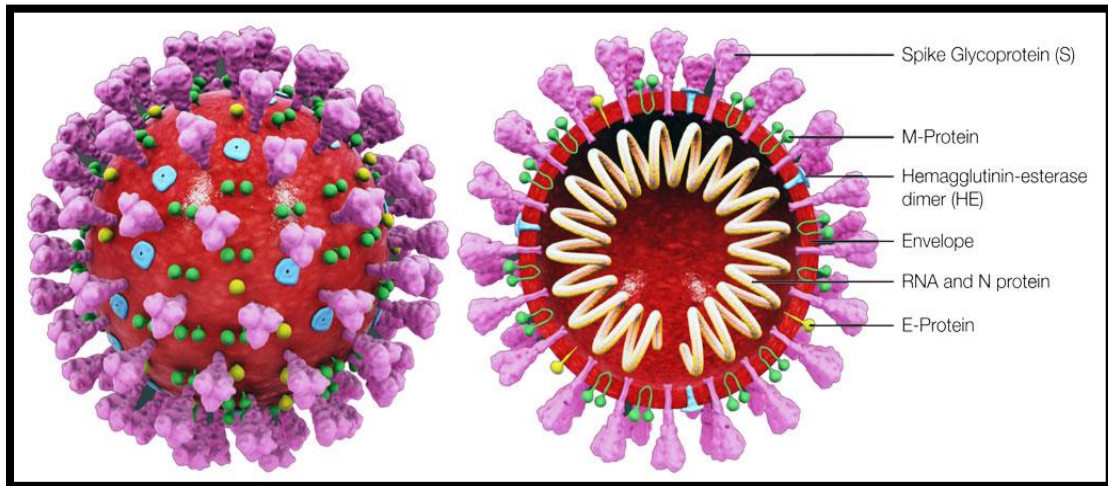
## I.1. Définition de Covid 19

La COVID-19 est la maladie causée par un nouveau coronavirus, le SARS-CoV-2. L'OMS a appris l'existence de ce nouveau virus le 31 décembre 2019 lorsqu'un foyer épidémique de cas de « pneumonie virale » a été notifié à Wuhan, en République populaire de Chine.

Ces virus peuvent provoquer des maladies cliniques chez l'homme qui peuvent s'étendre du simple rhume à des maladies respiratoires plus graves comme le SRAS et le MERS[OMS,2020].

Le SRAS-CoV-2 récemment émergent a fait des ravages en Chine et a provoqué une situation pandémique dans la population mondiale, Ce virus a été proposé pour être désigné / nommé coronavirus 2 du syndrome respiratoire aigu sévère (SRAS-CoV-2) par le Comité International de Taxonomie des Virus (ICTV), qui a déterminé que le virus appartient à la catégorie des coronavirus liés au syndrome respiratoire aigu sévère et a découvert que ce virus est lié au SRAS-CoV[OMS,2020].

Les coronavirus possèdent un génome à ARN non segmenté, simple brin, à sens positif d'environ 30 kb, entouré d'une queue 5'-cap et 3'-poly (A). Le génome du SARS-CoV-2 est long de 29 891 pb, avec une teneur en G + C de 38 %, Ces virus sont entourés d'une enveloppe contenant la nucléocapside virale. Les nucléocapsides des CoV sont disposées en symétrie hélicoïdale, ce qui reflète un attribut atypique dans les virus à ARN de sens positif. Les micrographies électroniques du SRAS-CoV-2 ont révélé un contour sphérique divergent avec un certain degré de pléomorphisme, des diamètres de virions variant de 60 à 140 nm et des pointes distinctes de 9 à 12 nm, donnant au virus l'apparence d'une couronne solaire[OMS,2020].



**Figure I.1 : Modélisation du Corona Virus[OMS,2020].**

## I.2. Les symptômes de la Covid-19

Les symptômes généraux de la COVID-19 sont assez similaires à ceux d'autres infections respiratoires aiguës telles que la grippe saisonnière. Les quatre principaux symptômes sont la fatigue, la fièvre, la toux et les difficultés à respirer. D'autres symptômes sont parfois décrits :

- + courbatures ou mal de dos ;
- + mal de gorge, écoulement nasal ;
- + éventuellement nausées, vomissements ou diarrhée.

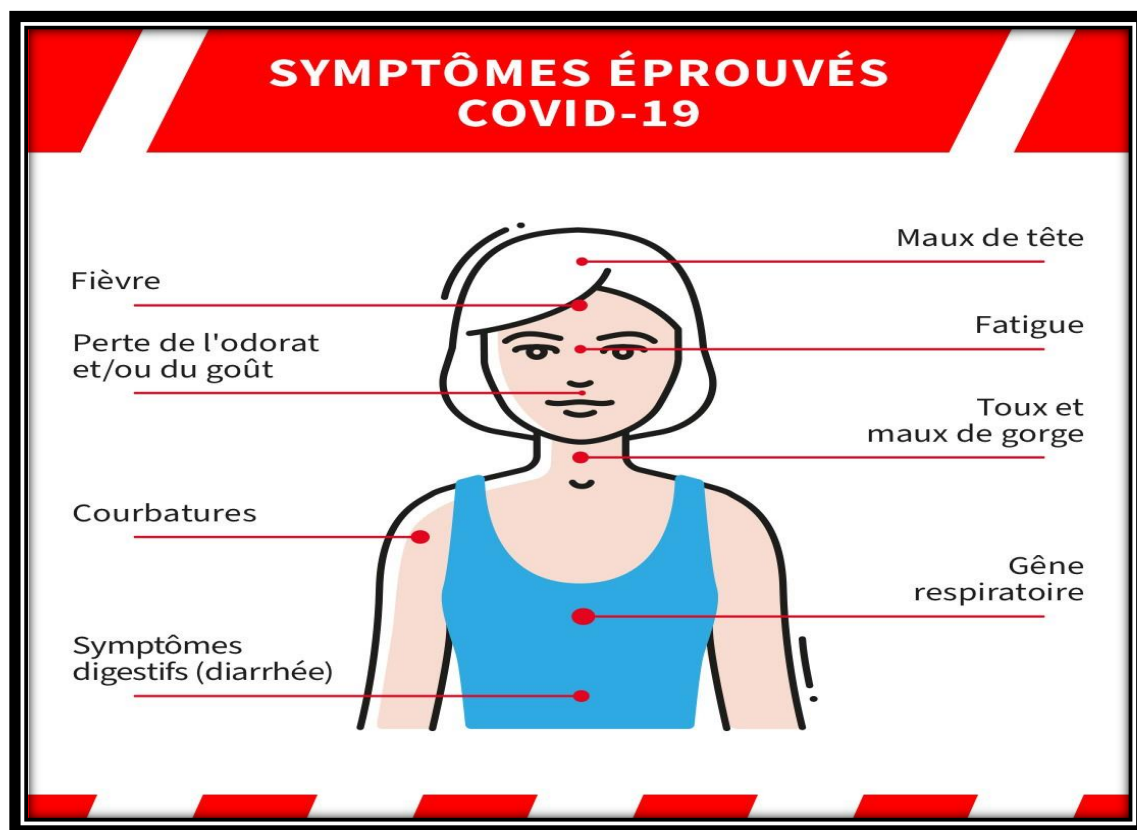
Certains symptômes sont plus caractéristiques de la COVID-19, en particulier la perte soudaine et temporaire de l'odorat (anosmie) et du goût (agueusie). Les lésions du poumon (observés par scanner) sont particulières et légèrement différentes de celles de la grippe saisonnière[INCESS2020].

Lors de formes particulièrement sévères, les patients souffrent de difficultés à oxygéner leur sang et doivent être hospitalisés et placés sous oxygène. Dans les cas les plus graves, ils sont mis en état de coma artificiel et intubés pour amener l'oxygène au plus près du sang.

Toute personne, quel que soit son âge, qui présente de la fièvre et/ou une toux associée à des difficultés à respirer ou à un essoufflement, des douleurs ou une sensation d'oppression dans la poitrine, ou une perte d'élocution ou de mouvement, doit immédiatement consulter un médecin. Si possible, appelez d'abord votre prestataire de soins de santé, votre ligne d'assistance téléphonique ou votre établissement de santé, afin d'être dirigé vers la structure adéquate[OMS,2020].

Parmi les malades qui développent des symptômes, la plupart (environ 80 %) guérissent sans qu'il soit nécessaire de les hospitaliser. Environ 15 % des patients tombent gravement malades et nécessitent une oxygénothérapie, et 5 % des infections sont critiques et exigent des soins intensifs. Les complications entraînant la mort sont notamment l'insuffisance respiratoire, le syndrome de détresse respiratoire aiguë (SDRA), la septicémie et le choc septique, la thromboembolie et/ou une défaillance multiviscérale, provoquant des lésions au cœur, au foie ou aux reins [Han MS, 2020].

Dans de rares cas, les enfants peuvent développer un syndrome inflammatoire grave quelques semaines après l'infection.



**Figure I.2 :** Les symptômes de la Covid-19 [OMS, 2020].

### I.2.1. Le temps met-on à ressentir les symptômes

Le délai entre l'exposition à la COVID-19 et le moment où les symptômes commencent à se manifester est, en moyenne, de 5-6 jours et peut aller de 1 à 14 jours. C'est pourquoi il est conseillé aux personnes qui ont été exposées au virus de rester chez elles, à l'écart des autres, pendant 14 jours, afin d'éviter la propagation du virus, en particulier lorsqu'il est difficile de se faire tester [OMS, 2020].

## I.2.2. Durée de la maladie

La durée de la maladie consiste en la période où l'altération de l'état de santé se manifeste par un ensemble de symptômes. Les manifestations cliniques des formes légères et modérées de la COVID-19 durent généralement moins de deux semaines, tandis que les formes plus sévères peuvent durer plus longtemps.

Certains symptômes peuvent persister au-delà de deux à trois semaines, il s'agit habituellement de toux (post- infectieuse), de fatigue, de myalgies, de rhinorrhée, de dyspnée et d'anosmie avec ou sans agueusie[Tenforde MW,2020].

La durée des symptômes de la Covid-19 peut varier d'une personne à l'autre mais en moyenne ils sont présents pendant 7 à 10 jours. "C'est ce qu'on appelle la phase aiguë de la maladie. Elle correspond à la durée moyenne d'un épisode de grippe par exemple", précise le Dr Le Flohic. Dans certains cas, les symptômes peuvent durer dans le temps et donner lieu à une forme de Covid long.

On parle de Covid long quand le patient présente un ou plusieurs symptômes initiaux 4 semaines après le début de la maladie, les plus fréquents étant les troubles respiratoires persistants, une grande fatigue, la perte de l'odorat et la perte de goût[OMS,2020].

## I.3. Propriétés générale de Covid19

### I.3.1. Classification

Les coronavirus (CoV) sont des virus à ARN monocaténaire de sens positif (groupe IV de la classification Baltimore) correspondant à la sous-famille *Orthocoronavirinae* de la taxonomie de l'ICTV, dans la famille *Coronaviridae*, et de l'ordre *Nidovirales*. Selon les caractéristiques de leurs séquences protéiques, les CoV sont classés en 4 genres (alpha-CoV, beta-CoV, gamma-CoV et delta-CoV), qui tous contiennent des virus pathogènes pour les mammifères :

- ✚ *Alphacoronavirus*, qui inclut le virus de la diarrhée épidémique porcine (PEDv), le virus de la gastro-entérite transmissible porcine (TGEV), le coronavirus du syndrome de la diarrhée aiguë porcine (SADS-CoV),

le coronavirus canin, le coronavirus entérique félin, le virus de la péritonite infectieuse féline (FIPV) ;

- ✚ *Betacoronavirus*, dont le virus respiratoire-respiratoire du SRAS (SARS-CoV), le SARS-CoV-2, le coronavirus du syndrome respiratoire du Moyen-Orient (MERS-CoV), virus de l'hépatite murine (MHV), coronavirus bovins, virus de la sialodacryoadénite du rat, virus de la sialodacryoadénite porcine, hémagglutinoase porcine, virus de l'hémagglutinoase porcine coronavirus équin. Dans ce genre *Betacoronavirus*, le SARS-CoV et le SARS-CoV-2 appartiennent tous les deux au sous-genre *Sarbecovirus* au sein duquel trois clades distincts ont été identifiés :
  - Clade1: souches "chauve-souris" de Bulgarie et Kenya ;
  - Clade2: SARS-CoV-2 et souches "chauve-souris" de Chine orientale ;
  - Clade3: SARS-CoV et souches "chauves-souris" de Chine du sud-ouest :
- ✚ *Gammacoronavirus*: surtout trouvé chez des oiseaux migrateurs, causant notamment des bronchites ; un *Gammacoronavirus* a été isolé d'un béluga en captivité,
- ✚ *Deltacoronavirus*: connus depuis peu, qui semblent surtout infecter les oiseaux, mais aussi trouvé chez les porcs[OMS,2020].

### voir le tableau suivante (Tableau I.1)

Les CoVs sont divisés en quatre genres, Alphacoronavirus, Betacoronavirus, Gammacoronavirus et Deltacoronavirus.

- ✚ Les genres Alphacoronavirus et Betacoronavirus proviennent de chauves-souris.
- ✚ Les genres Gammacoronavirus et Deltacoronavirus ont évolué à partir de pools génétiques d'oiseaux et de porcs.

Tableau I.1 :Classification de Coronavirus[Hoffmann2020]

Type	Virus
Domaine	Riboviria
Régne	Orthornavirae
Classe	Pisoniviricetes
Ordre	Nidovirales
Sous-ordre	<i>Cornidovirineae</i>
Famille	<i>Coronaviridae</i>
Sous-famille	<i>Orthocoronavirinae</i>

### I.3.2. Caractères physico-chimiques

Comme les autres coronavirus, SRAS CoV-2 est sensible aux rayons ultraviolets et à la chaleur (56 °C) ; peut être efficacement inactivés par les solvants lipidiques, notamment l'éther (75 %), l'éthanol, les désinfectants chlorés (eau de javel), l'acide peroxyacétique et le chloroforme, à l'exception de la Chlorhexidine [Hoffmann2020].

### I.3.3. Les Teste sur Covid-19

Généralement, il y a trois tests connue pour le diagnostique de Covid19, sont le test rapid, le test PCR (A polymerase chain reaction) et le scanner thoracique. Les tests rapides est tests antigéniques de diagnostic rapide (également appelés tests de diagnostic rapide ou TDR) détectent les protéines virales (connues sous le nom d'antigènes).

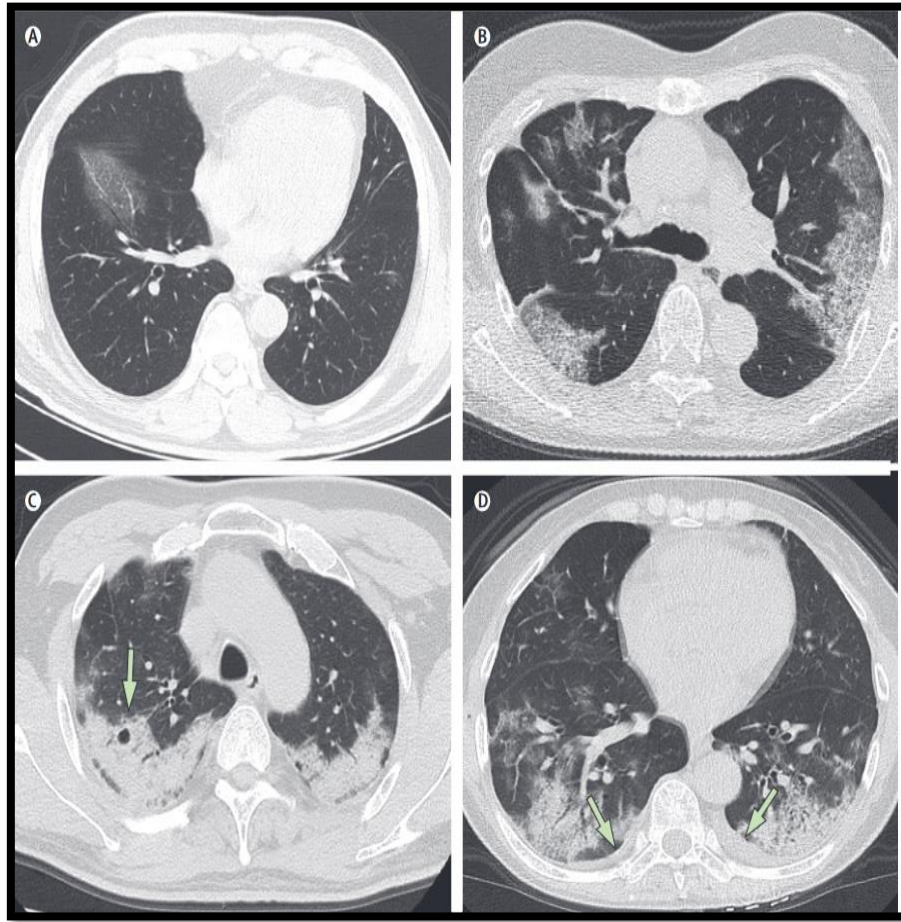
Les tests PCR sont plus efficaces lorsque le virus circule beaucoup dans une communauté et lorsque l'échantillon est prélevé au moment où la personne est la plus contagieuse. Il ya deux cas :

- ✚ Négatifs ne dépassé pas le scanner.
- ✚ Positifs obligé dépassé le scanner.

Le Scanner thoracique (TMD) appareil de radiodiagnostic composé d'un système de tomographie et d'un ordinateur qui en fournit les résultats sous forme d'image[**Pan Y,2020**].

Facile d'accès, rapide à réaliser, avec une forte sensibilité, le scanner thoracique a permis de pal- lier le manque d'outils de dépistage biologique durant les premières semaines de la pandémie en France, mais aussi dans le monde. Des images TDM typiques associant opacités en verre dépoli et *crazy paving* permettent de faire un diagnostic rapide. Cette stratégie a favorisé une gestion des flux bien plus efficiente, en limitant les temps de présence des patients dans les services d'urgence, réduisant ainsi le risque de contaminer ou d'être contaminé, pour les patients comme pour les soignants [**Hoffmann2020**].

- ✚ Le scanner thoracique, sensible, facile d'accès et rapide, a permis de trier et d'orienter les patients des urgences suspects de la
- ✚ Covid-19 et présentant des critères de gravité.
- ✚ Pour réaliser ce triage, le scanner thoracique était réalisé sans injection avec un protocole basse dose (low dose).
- ✚ Le scanner thoracique peut être normal
- ✚ Les trois premiers jours. L'aspect scanographique précoce est celui de plages d'opacités de verre dépoli non systématisées, périphériques,
- ✚ sous-pleurales, multifocales, bilatérales et asymétriques.
- ✚ L'évolution des lésions se fait avec la confluence des plages de verre dépoli, l'organisation de condensations et l'apparition d'opacités linéaires.



**Figure I.3 :** le scanner ( Imagerie) thoracique du covid 19[INSPQ,2020].

#### **I.4. Mode de transmission**

Le virus peut se propager lorsque de petites particules liquides sont expulsées par la bouche ou par le nez quand une personne infectée tousse, éternue, parle, chante ou respire profondément[OMS2020].

##### **I.4.1. Transmission par aérosols**

Le virus SRAS-CoV-2 est excrété par les voies respiratoires et peut se retrouver dans des aérosols de proximité expulsés par la personne infectée. Selon leur taille, ces particules pourront se déposer dans le nez, la gorge, les bronches ou se rendre jusqu'aux alvéoles pulmonaires. Elles peuvent également se déposer sur les conjonctives.

Ces aérosols de proximité sont émis en quantité variable quand la personne respire, parle, chante, tousse, éternue, etc. La distance sur laquelle pourront se déplacer les aérosols de différentes tailles (moins de 100  $\mu\text{m}$ ) varie en fonction de leur diamètre aérodynamique.

Les données expérimentales et épidémiologiques disponibles soutiennent une transmission par aérosols à proximité, c'est-à-dire à moins de deux mètres, mais il n'est pas exclu que, dans certains contextes, les aérosols puissent se déplacer à distance. La distance maximale demeure imprécise, mais il est peu probable que ce soit au-delà de quelques mètres[INSPQ,2020].

#### I.4.2. Transmission indirecte

Les gouttes et aérosols de plus gros diamètre projetés par une personne infectée peuvent se déposer sur les surfaces environnantes [Santé publique Ontario,2020].

La viabilité du SRAS-CoV-2 sur les différentes surfaces varie selon les facteurs environnementaux (température, humidité relative et rayonnement) et le type de surface. Il est possible que le SRAS-CoV-2 puisse survivre jusqu'à quelques jours sur les surfaces [INSPQ,2020].

Bien que la preuve appuyant la transmission par les fomites ne soit pas très robuste, elle reste plausible à proximité des personnes qui sont contagieuses, le risque pouvant être influencé par les pratiques de désinfection, l'hygiène, le respect des mesures sanitaires et la présentation clinique de la personne infectée.

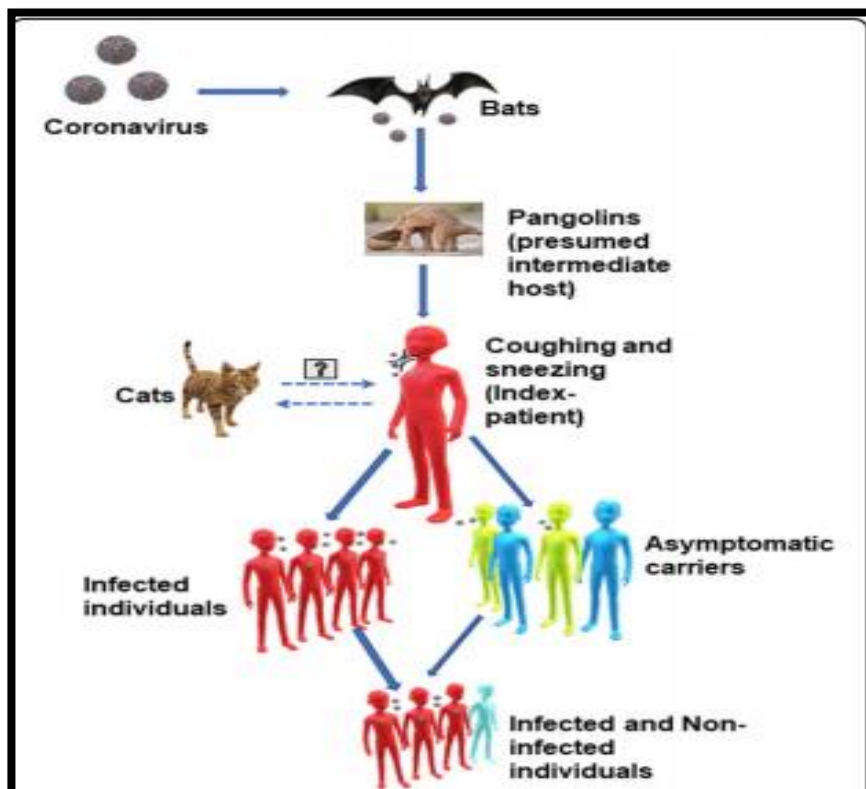


Figure I.4 :La transmission du cycle de SRAS-COV-2[INSPQ,2020].

### I.4.3. La transmission par les selles et autres liquides biologiques

La transmission par voie fécale-orale n'est pas établie, bien que de l'Acide ribonucléique (ARN) du SRAS-CoV-2 ainsi que des virus vivants aient été isolés dans les selles. Également, la transmission par l'urine, le plasma ou le sérum n'a pas été documentée, même si l'ARN du virus a été identifié dans ces liquides biologiques[INSPQ,2020].

### I.4.4. La transmission de la mère à l'enfant

Plusieurs études soulèvent la possibilité d'une transmission verticale du SRAS-CoV-2.

En effet, l'ARN viral a été détecté dans des prélèvements nasopharyngés chez 7 % de nouveau-nés dont la mère était positive au SRAS-CoV-2 (19). Le virus a été détecté en de rares occasions dans le placenta de mères atteintes par la COVID-19 (20). Toutefois, aucun cas n'a été confirmé par la détection du SRAS-CoV-2 dans le tissu ou le sang du cordon ombilical.

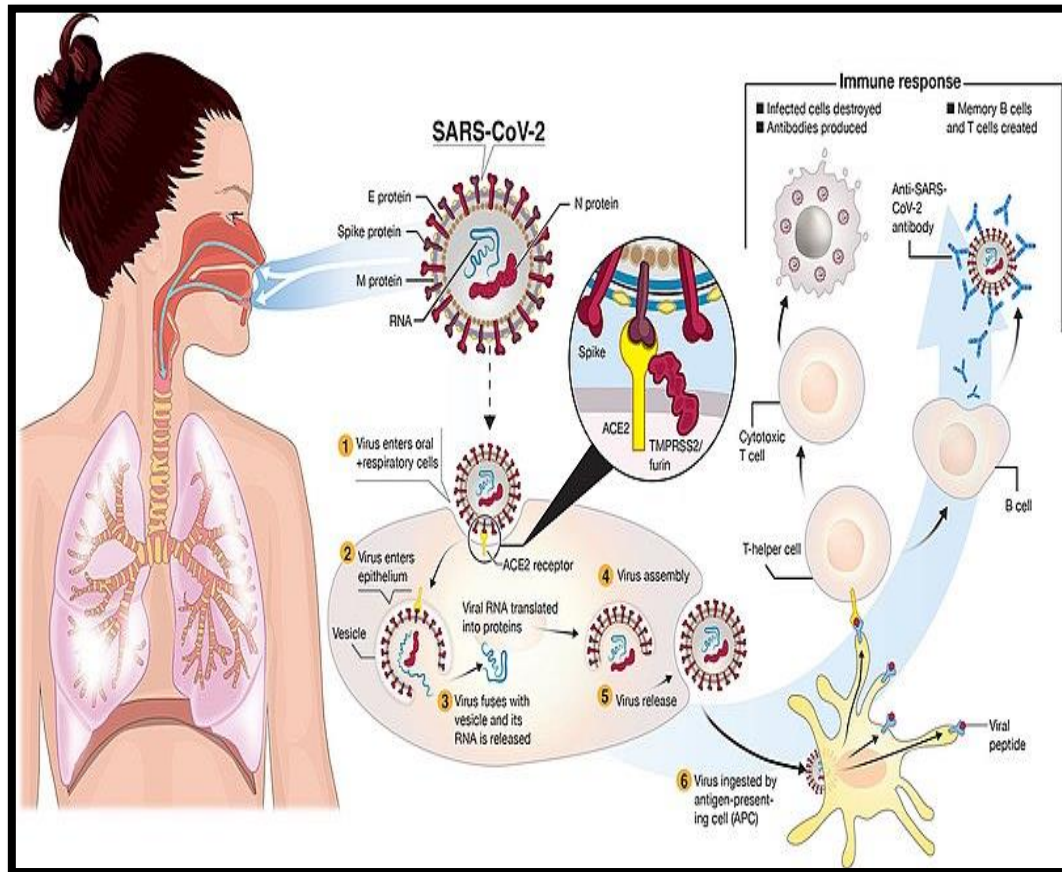
À ce jour, aucune évidence scientifique n'indique une transmission de la COVID-19 par l'allaitement.

Toutefois, deux études de cas ont rapporté la détection par réaction de polymérisation en chaîne avec transcriptase inverse (RT-PCR) de fragments du virus non viables dans le lait maternel de deux femmes[INSPQ,2020].

### I.4.5. La transmission zoonotique

Le réservoir principal de ce virus semble être la chauve-souris fer à cheval, mais plusieurs questions se posent toujours sur la façon dont il a pu être transmis à l'humain [Sallard E,2020].

La transmission du vison vers l'humain a été confirmée dans le contexte des fermes d'élevage de visons infectés par un variant spécifique. Les implications des changements identifiés dans ce variant ne sont pas encore bien comprises (24). Des cas de transmission de l'humain à l'animal ont été rapportés pour différentes espèces incluant le chat, le chien, le furet, le vison, le hamster et le singe (26).



**Figure I.5 :** Transmission et cycle de vie du SARS-CoV-2[INSPQ,2020].

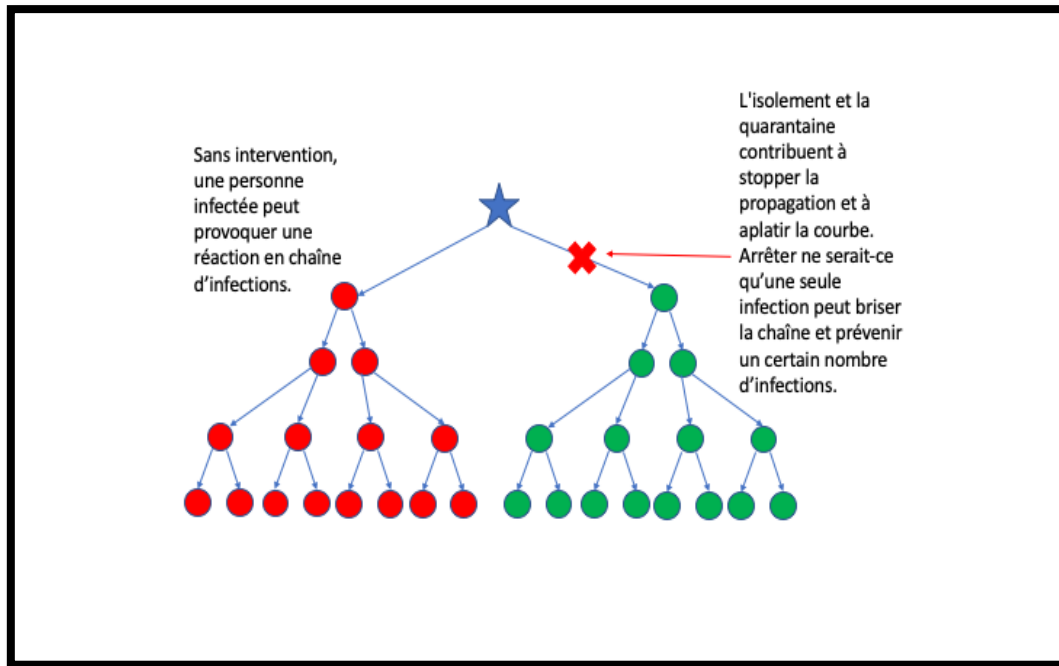
#### I.4.5. L'isolement et la quarantaine

L'isolement et la quarantaine sont deux méthodes permettant de freiner la propagation de la COVID-19.

La quarantaine est imposée aux individus ayant été en contact avec une personne infectée par le SARS-CoV-2, virus responsable de la COVID-19, que celle-ci présente des symptômes ou non. Pendant la quarantaine, vous ne rentrez en contact avec personne car vous avez été exposé au virus et que vous avez pu être infecté. La quarantaine peut s'effectuer dans un établissement désigné ou à domicile. Pour la COVID-19, vous devez rester en quarantaine dans l'établissement désigné ou à domicile pendant 14 jours.

L'isolement est imposé aux personnes qui présentent des symptômes de la COVID-19 ou qui ont été testées positives. Lorsque vous êtes en isolement, vous ne rentrez en contact avec personne et, dans l'idéal, vous vous rendez dans un établissement médical dans lequel vous pouvez recevoir des soins cliniques[Weblex.fr].

Si vous ne pouvez pas vous isoler dans un établissement médical et que vous n'appartenez pas à un groupe risquant de développer une forme grave de la maladie, l'isolement peut avoir lieu à domicile. Si vous présentez des symptômes, vous devez rester isolé pendant au moins 10 jours, plus 3 jours supplémentaires après la disparition des symptômes. Si vous êtes infecté et que vous ne développez pas de symptômes, vous devez rester isolé pendant 10 jours à partir du test positif [Weblex.fr].



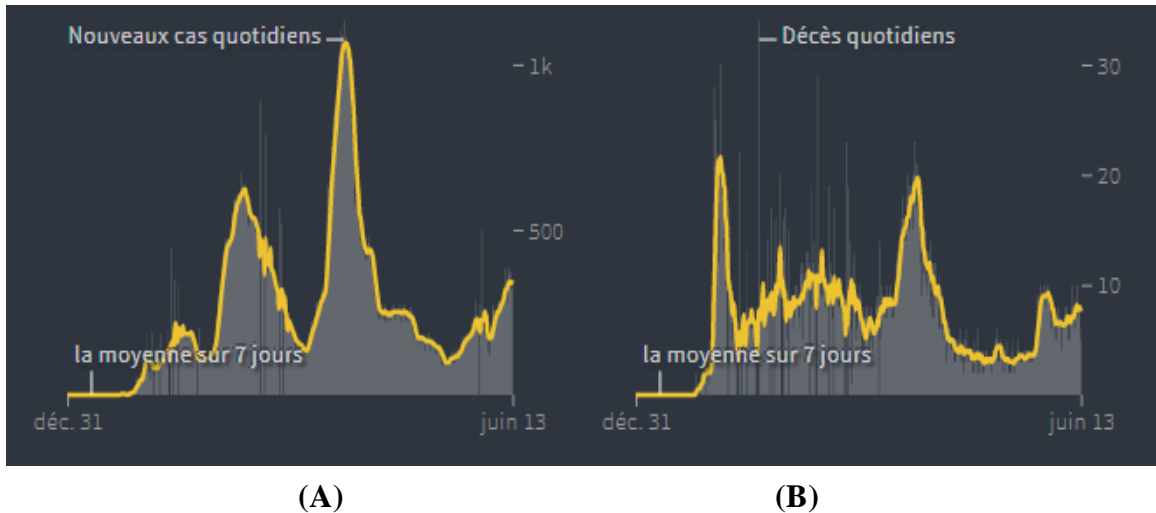
FigureI.6 : l'importance de isolement et la quarantaine[Weblex.fr]..

I.5. Données statistique

En Algérie, 33% du pic de contaminations et en hausse , 5 cas recensés pour 100.000 personnes au cours des 7 derniers jours[graphics.reuters].

Algérie : Les cas de COVID-19 sont en hausse avec 347 nouvelles contaminations recensées en moyenne chaque jour. Cela représente 33% du pic des infections — le nombre moyen le plus élevé de cas de contamination quotidiens a été reporté le 26 novembre.

Il y a eu 133 388 cas de contamination et 3 571 décès liés au coronavirus recensés dans le pays depuis le début de l'épidémie.

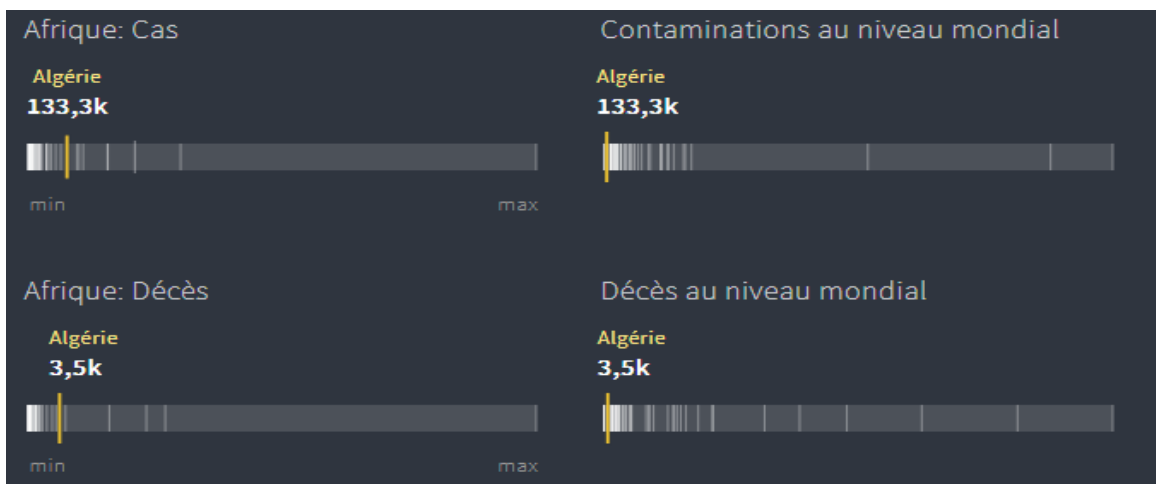


**Figure I.7:** Courbes des Tendances quotidiennes et nouvelles contamination[graphics.reuters].

**I.5.1. Comparaison de l’Algérie avec le reste du monde**

Il n’existe pas de statistiques optimales pour comparer le développement de l’épidémie dans chaque pays. Prendre en compte plusieurs indicateurs permet d’obtenir une vision plus globale de l’impact local du virus.

Ces graphiques présentent plusieurs statistiques - chacune d’entre elles possédant leurs propres forces et faiblesses - qui indiquent les différentes manières dont l’épidémie s’est développée dans chaque pays en comparaison avec sa région et le reste du monde[graphics.reuters].



**Figure I.8 :** Comparaison statistiques entre l’Algérie avec le reste du monde[graphics.reuters].

**TABLEAU I.2 :** Statistique du nombre de cas,de décès et de guérisons,liés au coronavirus (COVID-19) en Algérie[Salzberger B,2020].

<b>Date</b>	<b>Total cas cumulés 1) (variation J- )</b>	<b>Total décès cumulés</b>	<b>Nbre décès du jour  (en % variation J-1)</b>	<b>Guérisons cumulées</b>
<b>13/06</b>	<b>133 388 (+00%)</b>	<b>3 571 (+00%)</b>	<b>6</b>	<b>92 852</b>
<b>12/06</b>	<b>133 070 (+00%)</b>	<b>3 565 (+00%)</b>	<b>5</b>	<b>92 620</b>
<b>11/06</b>	<b>132 727 (+00%)</b>	<b>3 560 (+00%)</b>	<b>8</b>	<b>92 384</b>
<b>10/06</b>	<b>132 355 (+00%)</b>	<b>3 552 (+00%)</b>	<b>8</b>	<b>92 132</b>
<b>09/06</b>	<b>132 034 (+00%)</b>	<b>3 544 (+00%)</b>	<b>7</b>	<b>91 894</b>
<b>08/06</b>	<b>131 647 (+00%)</b>	<b>3 537 (+00%)</b>	<b>10</b>	<b>91 651</b>
<b>07/06</b>	<b>131 283 (+00%)</b>	<b>3 527 (+00%)</b>	<b>9</b>	<b>91 413</b>
<b>06/06</b>	<b>130 958 (+00%)</b>	<b>3 518 (+00%)</b>	<b>8</b>	<b>91 198</b>
<b>05/06</b>	<b>130 681 (+00%)</b>	<b>3 510 (+00%)</b>	<b>6</b>	<b>90 995</b>
<b>04/06</b>	<b>130 361 (+00%)</b>	<b>3 504 (+00%)</b>	<b>7</b>	<b>90 767</b>

**I.5.2. Taux de reproduction et facteur de dispersion**

Le taux de reproduction est l’un des indicateurs permettant d’estimer le niveau de transmission du virus au sein d’une population. Le taux de reproduction de base (R0) correspond à la moyenne de personnes infectées (cas secondaires) par un cas au sein d’une population naïve (sans immunité).

L’estimation consensuelle du R 0 pour la COVID-19 est entre 2 et 3 [Salzberger B,2020].

Le taux de reproduction effectif ( $R_t$ ) correspond au nombre moyen de cas secondaires produits par un cas infecté au sein d'une population dans laquelle circule déjà la maladie et où des mesures sont mises en place.

L'interprétation des taux de reproduction doit toujours être effectuée en fonction du contexte épidémiologique et en rapport avec les autres indicateurs disponibles incluant les taux de vaccination.

L'étude de grandes éclosions a également permis de mettre en lumière la dynamique hétérogène de transmission du virus et le concept d'évènements de super-propagation. En effet, certains cas ont été identifiés comme responsables d'un grand nombre de cas secondaires (agrégats).

Des auteurs ont rapporté que 15 à 24 % des cas de COVID-19 sont à l'origine de 80 % de l'ensemble des cas secondaires [Adam DC, 2020].

Cette hétérogénéité de la propagation du virus est illustrée par un autre indicateur, le facteur de dispersion ( $k$ ).

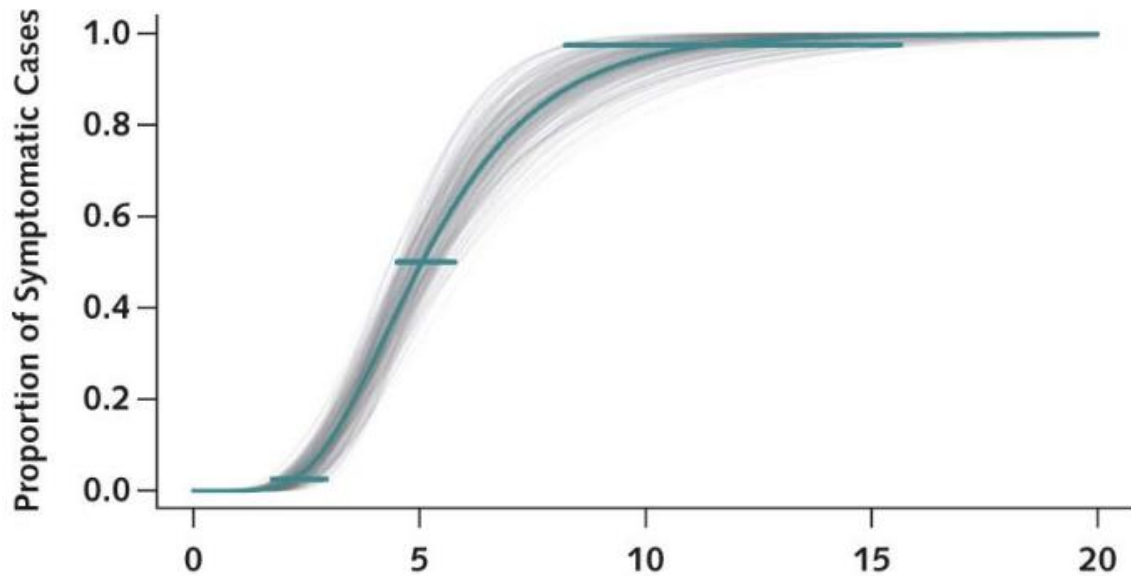
### **I.6. Période d'incubation**

La période d'incubation est le temps qui s'écoule entre l'exposition à l'agent infectieux et l'apparition des symptômes. Le temps médian pour l'incubation du SRAS-CoV-2 est de cinq jours (intervalle de 2-14 jours).

Des périodes d'incubation allant jusqu'à 17 jours ont été rapportées chez les personnes âgées et les enfants.

Les estimations combinées de quatre revues systématiques et méta-analyses pour les percentiles de la période d'incubation médiane sont :

- 50 e percentile : **5,1** jours (IC à 95 %, 4,5 à 5,8) à **5,4** jours (IC à 95 %, 5,0 à 5,7);
- 75 e percentile : **6,7** jours (IC à 95 %, 5,7 à 7,9) à **8,5** jours (IC à 95 %, 7,9 à 9,1);
- 90 e percentile : **9,7** jours (IC à 95 %, 8,1 à 11,6);
- 95 e percentile : **11,2** jours (IC à 95 %, 10,7 à 11,8) à **11,7** jours (IC à 95 %, 9,7 14,2);
- 97,5 e percentile : **11,5** jours (IC à 95 %, 8,2 à 15,6) à **16,5** jours (IC à 95 %, 14,8 à 18,3).



**Figure I.9:**La période d'incubation de la maladie à coronavirus 2019 (COVID-19) à partir de cas confirmés déclarés publiquement : estimation et application[INSPQ2020].

### I.7.Contagiosité

La période de contagiosité est la durée de temps pendant laquelle la personne infectée peut transmettre la maladie à une autre personne. En se basant sur la détection de l'ARN du virus dans les échantillons respiratoires et la capacité à faire pousser le virus en culture, la période estimée de contagiosité maximale du SRAS-CoV-2 se situe autour de la journée d'apparition des symptômes; elle débute quelques jours avant et décline de manière graduelle pendant les jours suivants[OMS,2020].

Une recension des études visant à établir la durée de la période infectieuse a montré que le virus viable a été cultivé à partir d'échantillons respiratoires de 1 à 6 jours avant le début des symptômes. Chez les cas ayant une maladie légère, l'excrétion virale a été documentée jusqu'à 8-10 jours après le début des symptômes. Le pic de la charge virale se situe dans les jours précédant le début de la maladie jusqu'à la première semaine après. Chez les personnes présentant une maladie grave ou chez les personnes immunosupprimées, l'excrétion virale peut être prolongée, du virus viable ayant été retrouvé jusqu'à 18-32 jours après le début de la maladie[FontanaL,2020].

Une étude épidémiologique de grande envergure réalisée à Taiwan a montré que les cas secondaires avaient été exposés pendant la période précédant l'apparition des

symptômes chez les cas index et jusqu'à cinq jours après, aucune transmission n'ayant été documentée lorsque l'exposition survenait six jours ou plus après le début de la maladie chez le cas index [INSPQ2020].

Une revue rapide de la littérature de l'INSPQ indique que les personnes asymptomatiques transmettent peu le SRAS-CoV-2 en comparaison avec les personnes symptomatiques. Une méta-analyse de 54 études évalue le taux d'attaque secondaire des contacts domiciliaires à 18,0 % (IC à 95 %, 14,2 % à 22,1 %) lorsque le cas index est symptomatique tandis qu'il est de 0,7 % (IC à 95 %, 0 % à 4,9 %) lorsque le cas index est asymptomatique[Mizrahi B,2020].

### **I.8. Présentations cliniques**

La COVID-19 est principalement une maladie respiratoire. Elle peut toutefois se présenter comme une maladie inflammatoire, occasionner des thromboses et provoquer des atteintes neurologiques, digestives, cardiaques, hépatiques, oculaires, et cutanées[Bonny V,2020]. L'étendue des manifestations cliniques de la COVID-19 va de l'absence de symptômes à des symptômes légers, modérés ou graves et au décès.

Une revue rapide de l'INSPQ mentionne que 15 % à 30 % des personnes infectées restent complètement asymptomatiques[Sun J,2020]. Selon les études, une grande variabilité de proportion de cas asymptomatiques est toutefois rapportée.

Chez l'adulte, l'anosmie d'apparition brutale (35 %) et la dysgueusie ou agueusie (44 %) sont les symptômes les plus discriminants chez l'adulte. Une étude québécoise évalue que les personnes présentant de l'anosmie et/ou dysgueusie ont 62,9 fois plus de risque d'être positif au test de dépistage du SRAS-CoV-2 comparativement à ceux ne présentant pas ces symptômes caractéristiques [OMS,2020].

### **I.9.Complications associées à la COVID-19**

Il a été établi que l'infection au SRAS-CoV-2 peut causer diverses complications graves persistantes ou survenant au-delà de six semaines.

#### **I.9.1. Complications cardiovasculaires:**

Insuffisance cardiaque; Évènements thrombotiques;Syndrome inflammatoire multisystémique de l'enfant (MIS-C)[ INSPQ,2020].

#### **I.9.2. Complications pulmonaires :**

Syndrome respiratoire aigu sévère;Syndrome restrictif résiduel; Fibrose pulmonaire interstitielle; Persistance de la difficulté respiratoire due à des dommages alvéolaires.

### **I.9.3. Complications neurologiques:**

Atteinte cognitive due à une hypercytokinémie;Accident vasculaire cérébral (AVC) [**Santé publique Ontario,2020**]; Dysfonctionnement olfactif (anosmie) et gustatif (agueusie ou dysgeusie)[**CDC,2020**].

### **I.9.4. Autres complications possibles :**

Septicémie avec ou choc septique;Défaillance multisystémique;Atteintes rénales (ex. : insuffisance rénale aiguë); Atteintes hépatiques; Diminution de la concentration et de la mobilité des spermatozoïdes; Autres séquelles en lien avec les soins intensifs administrés.

# **Chapitre II**

## **Revue génétique sur les SRAS-CoV et SARS-CoV- 2 Bat-CoV**

## **Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV**

---

Le SRAS-CoV-2 est un virus à ARN qui est devenu un problème majeur de santé publique après l'épidémie du syndrome respiratoire du Moyen-Orient-CoV (MERS-CoV) et du syndrome respiratoire aigu sévère-CoV (SARS-CoV) en 2002 et 2012, respectivement. Dans ce chapitre nous verrons une revue de la littérature sur les propriétés génétiques du covid19, tel que la structure.

### **II.1. SARS-CoV (syndrome respiratoire aigu sévère)**

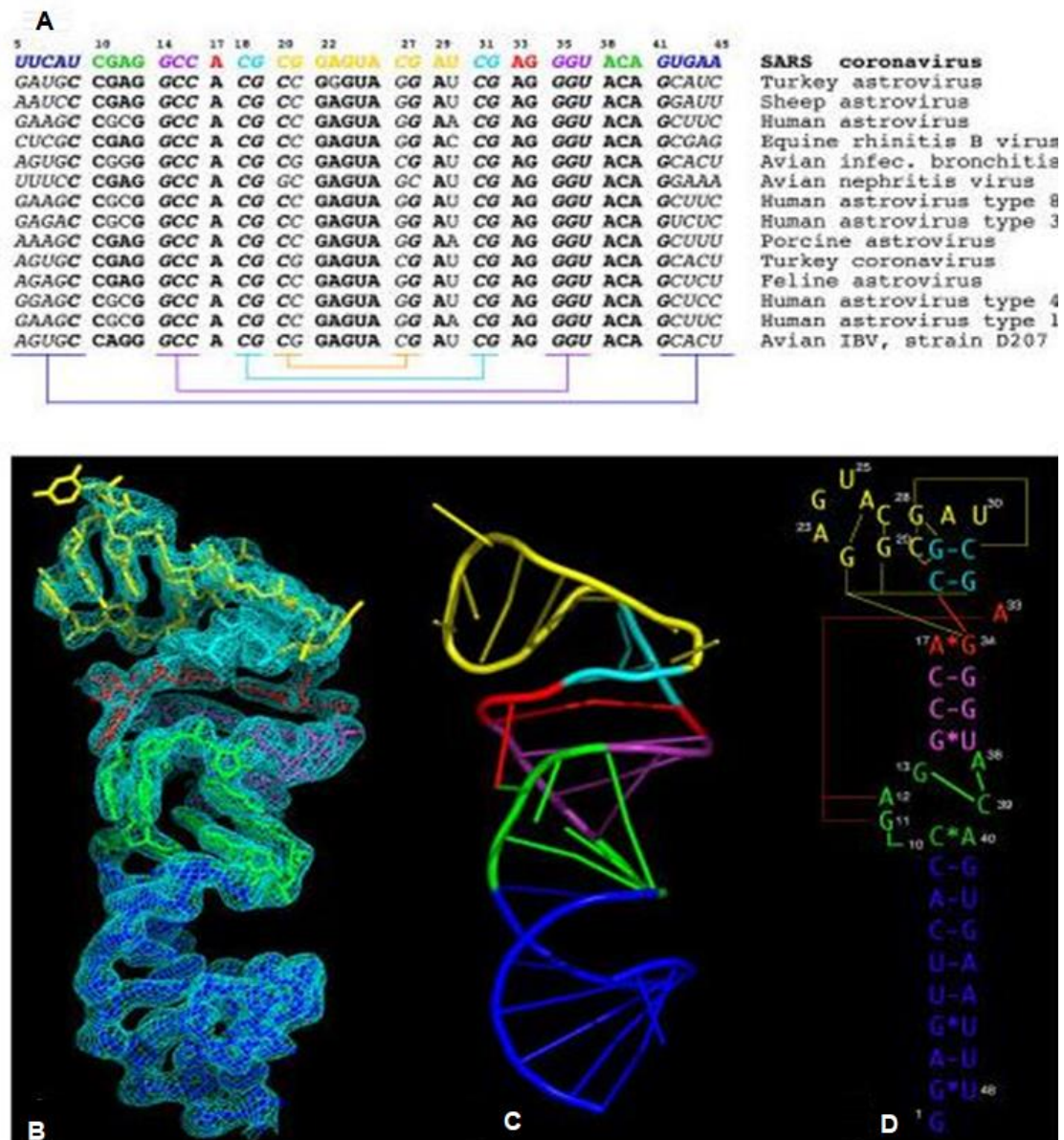
#### **II.2.1. Les structures de l'ARN s2m du SRAS**

Le virus qui *cause* le SRAS, comme d'autres coronavirus et astrovirus pathogènes, possède un génome d'ARN à brin plus-sens linéaire qui a une coiffe méthylée de 5 et une queue poly-A de 3.

##### **II.2.1.1. Les structures primaire, secondaire et tertiaire**

La réplicase virale est traduite directement à partir de l'ARN génomique à brin sens, puis crée une copie complémentaire complète (brin sens moins) de l'ARN génomique, ainsi qu'un ensemble imbriqué d'ARNm subgénomiques plus courts ayant 3' communs UTRs (régions non traduites). Ces 3' UTRs partagent tous avec l'ARN génomique du SRAS un élément de 32 nucléotides, immédiatement en amont de la queue 3 A poly-A (résidus 29590-29621)[**Iqra Mehmood,2021**] , à l'origine appelé motif tige-boucle II (s2m) astrovirus chez l'homme. L'élément s2m est l'élément d'ARN le plus conservé parmi les coronavirus et les astrovirus qui le contiennent (**Figure II.1**).

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV



**Figure II.1** : Les structures primaire, secondaire et tertiaire de l'ARN s2m du SRAS.

La partie A, de la figure 2.2 montre une comparaison phylogénétique de séquences d'ARN s2m de diverses espèces de coronavirus et d'astrovirus. La séquence d'ARN du SRAS est codée par couleur pour correspondre au schéma de couleurs utilisé tout au long. Les séquences conservées sont mises en évidence en caractères gras et les séquences co-variantes impliquées dans l'appariement de bases hélicoïdal d'ARN conventionnel sont indiquées en italique. Les compléments de séquence sont indiqués à l'aide de parenthèses codées par couleur.

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

---

La partie **B**, représente la carte de densité électronique expérimentale SIRAS en phase platine et aplatie au solvant de 2,7 Å a un contour de 1,25 écart quadratique moyen. La carte a permis un traçage sans ambiguïté de la molécule d'ARN car la densité était sans ambiguïté pour tous les atomes du squelette et toutes les bases nucléotidiques à l'exception de U(25), U(30) et U(48).

La partie **C**, pour un diagramme de ruban correspondant mettant en évidence le pli inhabituel[**Michael P. Robertson,2005**].

La partie **D**, c'est un représentation schématique de la structure secondaire de l'ARN s2m, avec des interactions structurelles tertiaires indiquées comme des contacts à longue distance. Le diagramme schématique est conçu pour se rapprocher de la représentation du pli. La structure pentaloop de type GNRA est représentée en jaune, les hélices d'ARN de forme A sont représentées en bleu et violet, le renflement asymétrique à trois purines est en rouge et la bulle de sept nucléotides est en vert. Les contacts tertiaires à longue portée sont indiqués par de fines lignes rouges et jaunes.

Les analyses de génomique structurale standard se concentrent sur l'obtention des structures tridimensionnelles des protéines codées dans un génome et sur l'identification d'une fonction protéique inconnue basée sur une homologie tridimensionnelle avec des structures protéiques de fonction connue .

Cependant, il est également impératif d'identifier et d'élucider les structures tridimensionnelles des produits géniques non protéiques, y compris les divers ARN requis pour le traitement de l'ARNm, la synthèse des proté à ARN, y compris des agents pathogènes tels que le VIH et le SRAS, il devient ines et d'autres fonctions cellulaires. Dans le cas des virus qui possèdent un génome essentiel d'élargir encore plus la portée des analyses de génomique structurale pour inclure les interactions tertiaires d'ARN biologiquement pertinentes qui se produisent dans le génome d'ARN lui-même.

Les éléments d'ARN génomique ayant le plus haut degré de conservation sont les plus susceptibles d'être cruciaux pour l'évolution, la croissance et la réplication de ces virus, et demandent donc le plus d'attention de la part de ceux qui cherchent à comprendre la pathogenèse virale de l'ARN et à concevoir des antiviraux appropriés. Drogues.

En utilisant la cristallographie aux rayons X, ils ont résolu la structure tridimensionnelle de l'ARN s2m du virus du SRAS à une résolution de 2,7-Å. La

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

---

structure révèle un virage à 90 ° spectaculaire et plusieurs nouvelles interactions tertiaires supplémentaires. Bien que la séquence et la structure tridimensionnelle de l'ARN s2m soient toutes deux uniques, la comparaison du repliement global de l'ARN s2m du SRAS aux structures tertiaires de l'ARN connues révèle que le repliement du squelette de l'ARN s2m imite celui de la boucle 530 de l'ARNr 16S, nous permettant d'émettre l'hypothèse que la fonction biologique de s2m dans le SRAS et les virus apparentés est basée sur le mimétisme macromoléculaire de cette région de l'ARN ribosomique. La boucle d'ARN ribosomique 530 et les protéines qui s'y lient sont impliquées dans l'initiation de la traduction, suggérant que le rôle du s2m dans le SRAS pourrait également impliquer l'initiation de la traduction. Plus précisément, nous proposons, sur la base d'arguments d'homologie structurelle, que l'ARN s2m du SRAS pourrait se lier au facteur d'initiation de la traduction eucaryote 1A (eIF-1A) de l'hôte pour détourner la machinerie traductionnelle de l'hôte à utiliser par le virus, ou pour se lier à d'autres protéines de régulation de la traduction ayant des plis similaires à des fins similaires[Angeletti,2020].

### II.2.2. Analyse de séquence de l'élément s2m conservé

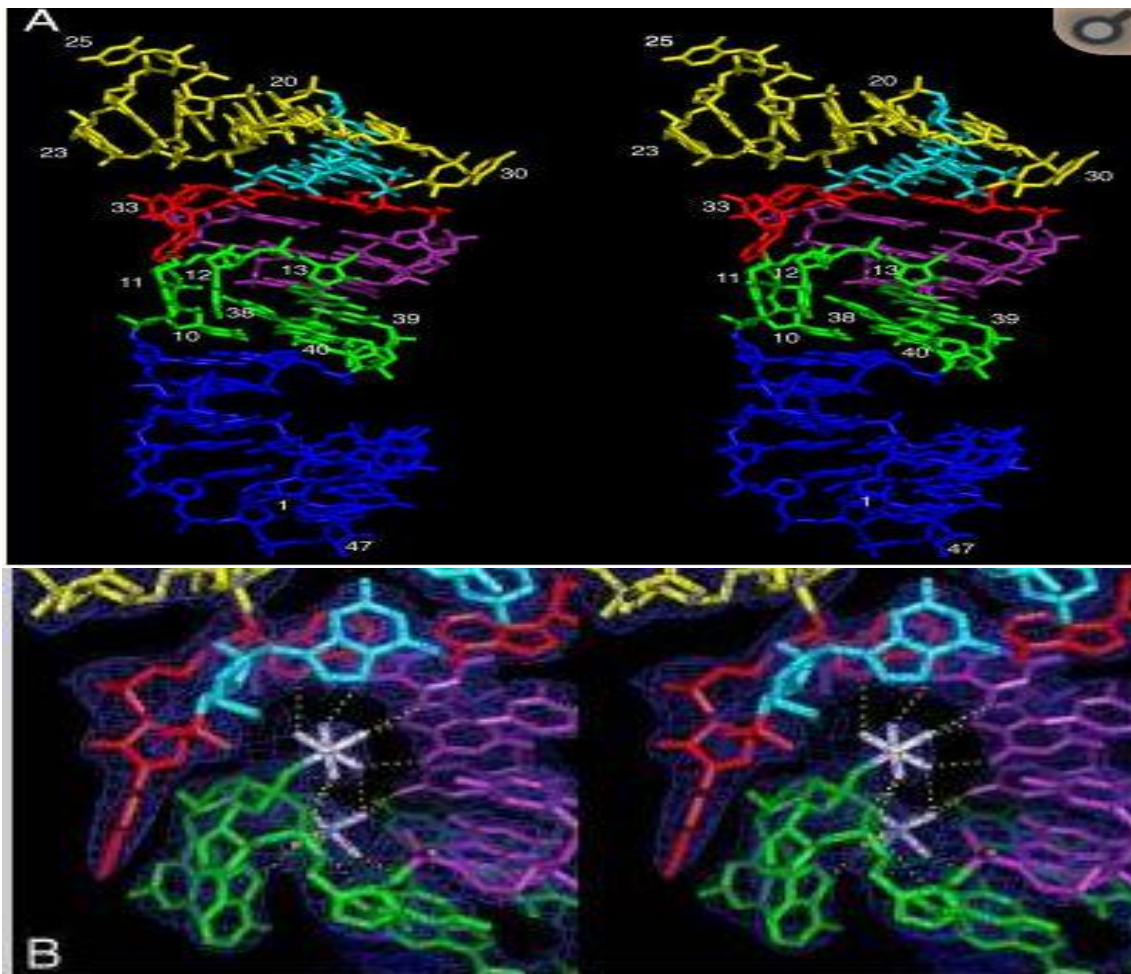
Robertson et al. (2005) ont aligné les séquences génomiques disponibles les plus récentes de coronavirus et d'astrovirus et analysé les modèles de conservation au sein de l'élément s2m (**Figure II.2**). Remarquablement, environ 75 % de cette séquence est absolument invariante entre les espèces virales (nucléotides indiqués en gras sur la **figure II.1a**) et une grande partie de la variation qui se produit préserve les éléments structurels secondaires (nucléotides indiqués en italique sur la **figure II.1a**). De plus, ils ont analysé 38 variantes séquencées du SRAS et constaté que le motif est absolument conservé dans chacune d'elles. Aucune insertion ou délétion ne semble tolérée, ce qui indique que cette région forme une structure tertiaire d'ARN hautement conservée qui est universellement requise pour la fonction virale.

### II.2.3. La structure cristalline de l'élément ARN s2m du SRAS

En utilisant la transcription in vitro, ils ont préparé et cristallisé une construction de 48 nucléotides contenant l'élément s2m de 45 nucléotides. Ils ont résolu la structure cristalline à une résolution de 2,7-Å en utilisant un seul dérivé isomorphe/anormal du platine et obtenu une carte de densité électronique aplatie au solvant facilement interprétable (**figures II.1.B – D.2A**) La qualité de la densité électronique nous a permis d'adapter sans ambiguïté la séquence d'ARN s2m à la carte

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

et de construire un modèle de la structure tertiaire inhabituelle. La carte initiale était pratiquement impossible à distinguer de la carte finale  $3F_o - 2F_c$  calculée à l'aide de phases de la structure d'ARN raffinée, indiquant que les phases expérimentales de remplacement isomorphe unique avec diffusion anormale (SIRAS) initialement obtenues étaient assez précises). Deux complexes  $Mg^{2+}$  + hydratés bien ordonnés liés au squelette phosphate de l'ARN sont également facilement observables dans la carte de densité électronique initiale (**figure II.2B**).



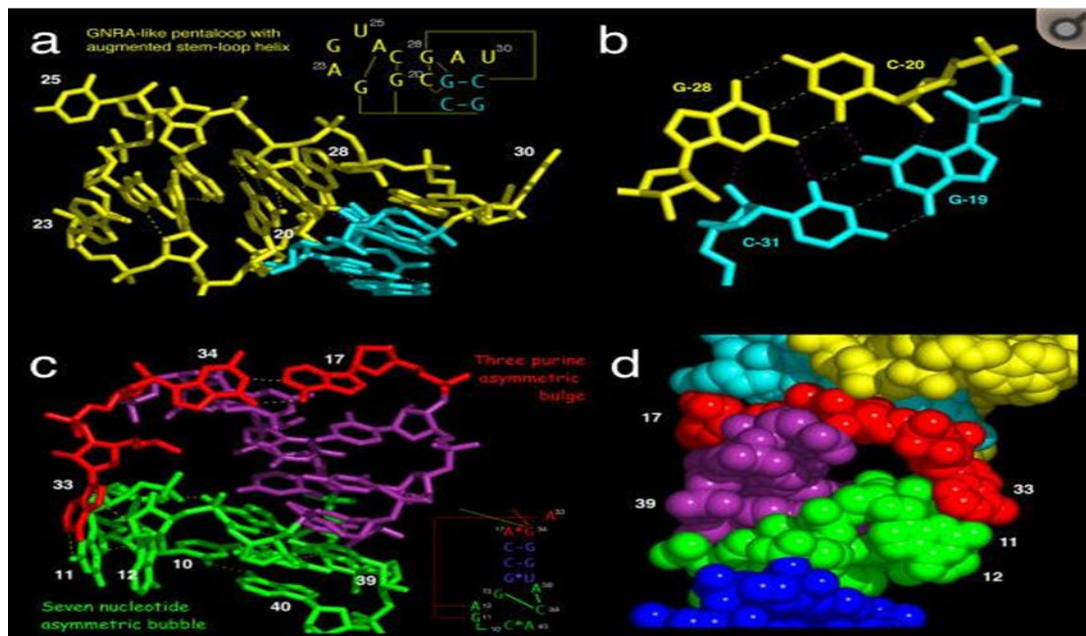
**Figure II.2** : Représentations stéréo de la structure de l'ARN s2m du SRAS

(A) La structure tridimensionnelle globale de l'ARN s2m du SRAS et (B) une vue détaillée des contacts tertiaires et les sites de liaison  $[Mg(H_2O)_5]^{2+}$  dans le contexte de la carte de densité électronique en phase expérimentale (bleu foncé). Les ions complexes  $[Mg(H_2O)_5]^{2+}$ , représentés par des octaèdres blancs, se lient aux atomes d'oxygène phosphate pro-R et pro-S de A (12). Un vaste réseau de liaisons

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

hydrogène potentielles entre les molécules d'eau coordonnées par le métal et l'ARN est représenté par des lignes pointillées jaunes.

La structure cristalline du domaine s2m de l'ARN du SRAS révèle plusieurs nouveaux éléments structuraux tertiaires (**figure II.3**). Trois régions de l'ARN canonique de forme A sont indiquées dans diverses nuances de bleu, et trois régions de structure inhabituelle, y compris les interactions tertiaires, sont représentées en vert, rouge et jaune. Le pli tridimensionnel réel de l'ARN est illustré sur la figure 1C, la figure 1D étant conçue pour représenter schématiquement ce pli ainsi que les contacts structuraux secondaires et tertiaires qui le stabilisent. La figure 2A montre un diagramme stéréo correspondant dans lequel tous les atomes non hydrogène sont présents [Pachetti, 2020].



**Figure II.3:** Interactions structurelles tertiaires dans l'ARN s2m du SRAS

La partie A, présente un gros plan de la structure pentaloop avec l'hélice croissante, représentée en jaune, et la jonction perpendiculaire formée avec la tige en forme de A, représentée en cyan. Les liaisons hydrogène roses indiquent une liaison hydrogène en quartet de base, comme indiqué en (B). Le coude à 90° ainsi formé est facilité par une courbure très prononcée de la colonne vertébrale impliquant des résidus non appariés 29 et 30.

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

---

La partie **B** est concernant la formation de la jonction de deux hélices perpendiculaires est facilitée par un quatuor de base composé de deux paires G – C.

La partie **C**, représente l'appariement inhabituel entre A (17) et G (34) facilite la formation d'un contact tertiaire à longue portée entre A (33) du renflement asymétrique à trois purines et G (11) et A (12) des sept -bulle asymétrique de nucléotide. A (38) forme un triple de base avec C (39) et G (13), forçant G (11) et A (12) hors de l'hélice principale.

La partie **D**, c'est une représentation de remplissage d'espace de la région illustrée en (C), mais tournée d'environ 180 °. Un tunnel est créé par les contacts tertiaires entre A (33) du renflement asymétrique purine (rouge), G (11) et A (12) de la bulle de sept nucléotides (vert) et la région hélicoïdale entre eux (violet). L'oxygène phosphate non pontant de G (11) et A (12) tapisse la surface de la cavité, créant une région chargée négativement dans laquelle on observe la liaison des ions Mg<sup>2+</sup>.

### II.2. SARS-CoV-2

#### II.2.1. Structure de SARS CoV-2

Les coronavirus codent pour quatre protéines structurelles majeures, à savoir la pointe (**S**), la membrane (**M**), l'enveloppe (**E**) et la nucléocapside (**N**), qui sont décrites en détail ci-dessous.

##### II.2.1.1. Glycoprotéine S

La protéine du coronavirus S est une grande protéine transmembranaire virale multifonctionnelle de classe I. La taille de cette abondante protéine S varie de 1160 acides aminés (IBV, virus de la bronchite infectieuse, chez les volailles) à 1400 acides aminés (FCoV, coronavirus félin) [Tsai,2020]. Il se situe dans un trimère à la surface du virion, donnant au virion une apparence de couronne ou de couronne. Fonctionnellement, il est nécessaire à l'entrée des particules de virion infectieuse dans la cellule par interaction avec divers récepteurs cellulaires de l'hôte .

De plus, il agit comme un facteur critique pour le tropisme tissulaire et la détermination de la gamme d'hôtes. Notamment, la protéine S est l'une des protéines immunodominantes vitales des CoVs capables d'induire des réponses immunitaires de l'hôte. Les ectodomaines de toutes les protéines CoVs sont des organisations de domaine similaires, divisées en deux sous-unités, s1 et s2. Le premier, S1, aide à la liaison au récepteur hôte, tandis que le second, s2, représente la fusion. Le premier

## **Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV**

---

(s1) est en outre divisé en deux sous-domaines, à savoir le domaine n-terminal (ntd) et le domaine c-terminal (ctd). Ces deux sous-domaines agissent comme des domaines de liaison au récepteur, interagissant efficacement avec divers récepteurs de l'hôte. Le s1 ctd contient le motif de liaison au récepteur (RBM).

Dans chaque protéine de pointe de coronavirus, le trimérique S1 se situe au-dessus de la tige trimérique S2[Lu,2020]. Récemment, des analyses structurales des protéines S de COVID-19 ont révélé 27 substitutions d'acides aminés dans un tronçon de 1273 acides aminés. Six substitutions sont situées dans le domaine de liaison au récepteur-RBD (acides aminés 357 à 528), tandis que quatre substitutions sont dans le RBM au niveau du CTD du domaine S1. Il est à noter qu'aucun changement d'acide aminé n'est observé dans le RBM, qui se lie directement au récepteur de l'enzyme de conversion de l'angiotensine-2 (ACE2) dans le SRAS-CoV. À l'heure actuelle, l'accent est mis sur la connaissance du nombre de différences nécessaires pour modifier le tropisme de l'hôte. La comparaison des séquences a révélé 17 changements non synonymes entre la séquence précoce du SRAS-CoV-2 et les isolats ultérieurs du SRAS-CoV. Les changements ont été trouvés dispersés sur le génome du virus, avec neuf substitutions dans ORF1ab, ORF8 (4 substitutions), le gène de pointe (3 substitutions) et ORF7a (substitution unique).

Notamment, les mêmes changements non synonymes ont été trouvés dans un groupe familial, indiquant que l'évolution virale s'est produite au cours de la transmission de personne à personne. De tels événements d'évolution adaptative sont fréquents et constituent un processus constamment en cours une fois que le virus se propage parmi de nouveaux hôtes. Même si aucun changement fonctionnel ne se produit dans le virus associé à cette évolution adaptative, une surveillance étroite des mutations virales qui se produisent lors de la transmission interhumaine ultérieure est justifiée.

### **II.2.1.2. Protéine M**

La protéine M est la protéine virale la plus abondante présente dans la particule de virion, donnant une forme définie à l'enveloppe virale. Il se lie à la nucléocapside et agit comme un organisateur central de l'assemblage du coronavirus. Les protéines M du coronavirus ont une teneur en acides aminés très diverse, mais conservent une similitude structurelle globale entre différents genres. La protéine M possède trois domaines transmembranaires, flanqués d'une courte extrémité aminée à l'extérieur du

## **Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV**

---

virion et d'une longue extrémité carboxy à l'intérieur du virion. Dans l'ensemble, l'échafaudage viral est maintenu par l'interaction M-M. Il est à noter que la protéine M du SARS-CoV-2 n'a aucune substitution d'acide aminé par rapport à celle du SARS-CoV [Iqra Mehmood,2021].

### **II.2.1.3. Protéine E**

La protéine E du coronavirus est la plus énigmatique et la plus petite des principales protéines structurales. Il joue un rôle multifonctionnel dans la pathogenèse, l'assemblage et la libération du virus. C'est un petit polypeptide membranaire intégral qui agit comme une viroporine (canal ionique). L'inactivation ou l'absence de cette protéine est liée à la virulence altérée des coronavirus en raison de changements de morphologie et de tropisme. La protéine E se compose de trois domaines, à savoir un court amino-terminal hydrophile, un grand domaine transmembranaire hydrophobe et un domaine C-terminal efficace. La protéine SARS-CoV-2 E révèle une constitution d'acides aminés similaire sans aucune substitution [Pachetti,2021].

### **II.2.1.4. Protéine N**

La protéine N du coronavirus est polyvalente. Parmi plusieurs fonctions, il joue un rôle dans la formation de complexes avec le génome viral, facilite l'interaction de la protéine M nécessaire lors de l'assemblage du virion et améliore l'efficacité de la transcription du virus.

Il contient trois domaines distincts et hautement conservés, à savoir, un NTD, un domaine de liaison à l'ARN ou une région de liaison (LKR) et un CTD. La NTD se lie à l'extrémité 3' du génome viral, peut-être via des interactions électrostatiques, et présente de fortes divergences tant en longueur qu'en séquence. Le LKR chargé est riche en sérine et en arginine et est également connu sous le nom de domaine SR (sérine et arginine). Le LKR est capable d'interagir directement avec l'interaction ARN in vitro et est responsable de la signalisation cellulaire. Il module également la réponse antivirale de l'hôte en agissant comme un antagoniste de l'interféron (IFN) et de l'interférence ARN. Par rapport à celle du SRAS-CoV, la protéine N du SRAS-CoV-2 possède cinq mutations d'acides aminés, dont deux dans la région intrinsèquement dispersée (IDR ; positions 25 et 26), une chacune dans la NTD (position 103), LKR (position 217) et CTD (position 334).

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

### II.2.1.5. nsp et protéines accessoires

En plus d'importantes protéines structurales, le génome du SRAS-CoV-2 contient 15 nsp, nsp1 à nsp10 et nsp12 à nsp16, et 8 protéines accessoires (3a, 3b, p6, 7a, 7b, 8b, 9b et ORF14). Toutes ces protéines jouent un rôle spécifique dans la réplication virale. Contrairement aux protéines accessoires du SRAS-CoV, le SARS-CoV-2 ne contient pas de protéine 8a et possède une protéine 8b plus longue et une protéine 3b plus courte. Les protéines accessoires nsp7, nsp13, enveloppe, matrice et p6 et 8b n'ont pas été détectées sans substitution d'acides aminés par rapport aux séquences d'autres coronavirus.

La structure virale du SRAS-CoV-2 est illustrée à la figure suivante (Figure II.4).

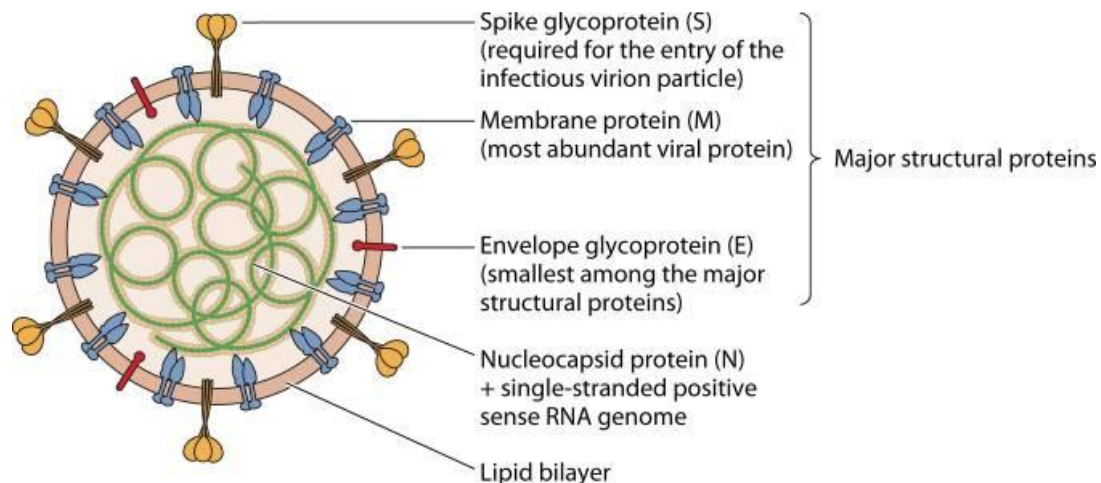


Figure II.4 : Structure de SARS-CoV-2

### II.2.2. Aperçu de l'organisation génomique

Les coronavirus, qui appartiennent à la famille des Coronaviridae, sont des virus enveloppés et pléomorphes, Ce sont des virus à ARN de sens positif avec une taille de génome d'environ de 30 kpb qui semble être la plus grande taille pour un virus à ARN, contenant 50 capuchon et 30 poly A-queue. Les coronavirus ont une nucléocapside hélicoïdale et flexible. La membrane de ces virus contient une glycoprotéine membranaire, une protéine enveloppée et une protéine de pointe tandis que l'ARN est entouré de nucléocapside [Tsai.2020].

L'ARN du virus contient 6 cadres de lecture ouverts (ORF1ab, ORF3a, ORF6, ORF7ab, ORF8, et ORF10).

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

---

Les deux tiers du génome du virus comprennent l'ORF 1a/1b et le reste un tiers du code génomique est utilisé pour M (membrane), S (pointe), N (nucléocapside) et E protéines structurales virales (enveloppées) [Tsai.2020].

La transcription a été réalisée par la synthèse de sgRNA (sub-genomic RNA) et complexe de réplication-transcription (RTC), enveloppé dans des vésicules à double membrane. La terminaison de la transcription s'est produite par les séquences régulatrices de la transcription qui sont présentes entre les cadres de lecture ouverts (ORF). Il y a 6 ORF dans le génome du SRAS-CoV-2, comme discuté ci-dessus [Tsai.2020].

Une mutation de décalage du cadre de lecture dans ORF1a et ORF1b produit des polypeptides (pp1a et pp1ab), qui sont ensuite traités par des protéases virales telles que la principale protéases (Mpro), protéases de type chymotrypsine (3CLpro), ou par des protéases de type papaïne pour la production de protéines non structurales (nsps)

Outre 1a et 1b lecture ouverte cadres (ORF), tous les autres ORF sont responsables de la production de protéines structurelles (protéines de membrane, de nucléocapside, enveloppées et de pointe), comme le montre la **figure II.4**.

Grâce à l'analyse des séquences du SARS-CoV-2 et du SARS-CoV, les scientifiques ont proposé un mutation dans la protéine de pointe responsable du saut du virus des animaux aux humains [Angeletti2020].

De même, certaines mutations ont également été trouvées dans des séquences protéiques qui conduire à la formation de protéines avec une modification des résidus d'acides aminés. Par exemple, à position 723, à la place de la glycine il y a une sérine, alors qu'en position 1010 il y a de la proline au lieu de l'isoleucine (Angeletti.2020). La récurrence potentielle de la maladie dépend de l'évolution du virus en raison de l'accumulation de mutations dans le génome viral au fil du temps.

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

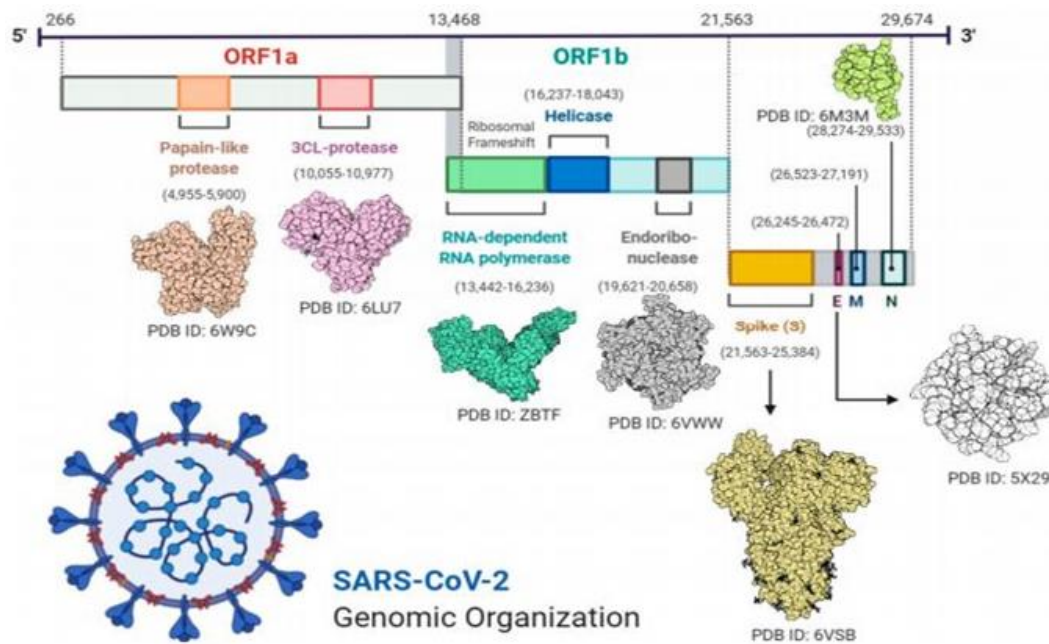


Figure II.5. Organisation structurale et génomique complète du coronavirus 2 du syndrome respiratoire aigu sévère (SRAS-CoV-2).

### II.2.3. Séquençage du génome

Grâce à l'analyse de la séquence génomique, il a été confirmé que bien que le SARS-CoV-2 a de nombreuses similitudes avec le SRAS-CoV et d'autres coronavirus apparentés, c'est un nouveau virus (**Tableau II.1**). Le virus a transmis l'organisme hôte des animaux aux humains avec certaines modifications / mutations uniques. L'analyse du séquençage du génome suggère que la plupart des contigs / reads viraux avaient une similitude avec le génome du bêta-coronavirus. Le SARSCoV-2 a des niveaux de similitude de 96,20 % et 88,00 % avec le SARSr-CoV précédemment publié (RaTG13) et bat-SL-CoVZC45, respectivement. Le séquençage du génome du SARSCoV-2 d'une autre étude a indiqué une similitude de séquence de 69,58 % et 82,45 % avec Les génomes du MERS-CoV et du SARS-CoV, respectivement (**Lu,2020**)

Dix séquences du génome viral obtenus à partir de 9 patients présentaient une identité de séquence de 99,98 %.

Dans une autre étude, les séquences des huit échantillons de patients présentaient une identité de séquence de 99,98 % entre eux dans l'ensemble génome [**Lu,2020**].

Recherche BLASTn des séquences SARS-CoV-2 identifiées des correspondances du les virus les plus proches connus auparavant : le bêta-coronavirus de type SRAS

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

d'origine chauve-souris, bat-SL-CoVZC45 (88 % d'identité de séquence ; 99 % de couverture des requêtes) et bat-SL-CoVZXC21 (88 % d'identité de séquence ; 98 % de couverture des requêtes). Dans 5 régions de gènes (7, M, N, 14 et E), l'identité de séquence était supérieure à 90 % avec 98,7 % comme niveau le plus élevé pour l'enveloppe (E) inconfortable. Le gène Spike (S) a démontré la plus faible identité de séquence à 75 %. cependant, l'identité de séquence dans les régions des gènes 1a et 1b était respectivement de 90 % et 87 %. Thé la majorité des protéines codées par le SARS-CoV-2 étaient très similaires aux protéines codées par coronavirus liés aux chauves-souris avec quelques insertions et suppressions [Lu,2020].

**Tableau II.1.** Homologie de séquence entre le SARS-CoV-2 et d'autres souches de coronavirus

Coronaviruses Strains	Sequence Similarity
SARSr-CoV; RaTG13	96.20%
bat-SL-CoVZC45	88.00%
bat-SL-CoVZXC21	88.00%
SARS-CoV	82.45%
SARS-HCoV Tor2	82.00%
SARS-HCoV BJ01	82.00%
MERS-CoV	69.58%
HCoV-OC43	68.93%
HCoV-HKU1	67.59%
HCoV-229E	65.04%
HCoV-NL63	65.11%

Cependant, la protéine 13 et la protéine S a révélé une identité de 73,2 % et 80 % avec d'autres protéines virales dérivées de chauve-souris, respectivement. Le SRAS-CoV-2 a codé une grande protéine de pointe, ce qui est une distinction majeure est l'un des coronavirus SARS-CoV-2, SARS-CoV, MERS-CoV et autres dérivés de chauve-souris. Le SARS-CoV-2 a la même organisation génomique que bat-SL-CoVZXC21, SARS-CoV, et bat-SL-CoVZC45, comme révélé par comparaison des régions codantes prédites.

Dix régions codantes ont été identifiées, dont E, M, N, S, 10ab, 9, 8, 7, 3 et 1ab

### Analyse phylogénétique

L'analyse phylogénétique des génomes du SRAS-CoV-2 obtenus à partir des premiers échantillons de patients a suggéré une similitude dans l'organisation des séquences avec les bêta-coronavirus tels que 5' UTR (région non traduite), le complexe de réplicase (orflab), 4 gènes (M, N, S et E ), 3' UTR (régions non

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

---

traduites<sup>1</sup>) et quelques ORF non structuraux non identifiés (cadres de lecture ouverts). Au lieu d'avoir une similitude de séquence avec les bêta-coronavirus découverts chez les chauves-souris, le SARSCoV-2 est distinct du SARS-CoV, ainsi que du MERS-CoV. Un autre élément de preuve indiquant sa nouveauté est que l'identité de séquence dans les domaines de réplicase conservés (ORF 1ab) est inférieure à 90% entre le SARS-CoV-2 et d'autres membres des bêta-coronavirus et du sous-genre sarbeco-virus de la famille des Coronaviridae.

### II-2.3.1 Protéines Conservées:

La protéine S est responsable de la fusion membranaire et de la liaison au récepteur. C'est aussi critique dans le contrôle de la capacité de transmission du virus et du tropisme de l'hôte. La protéine S du SARSCoV-2 possède deux domaines, à savoir les domaines S1 et S2. Le domaine S1 est responsable de la liaison au récepteur, tandis que le domaine S2 pour la fusion membranaire. Il a été rapporté que une protéase cellulaire (furine) est responsable du clivage des sites S1/S2 et ce clivage est nécessaire à l'entrée du virion dans les cellules pulmonaires humaines et la fusion cellulaire facilitée par la protéine S. Les domaines S1 et S2 du SRAS-CoV-2 ont une similarité de séquence de 93 % et 68 % avec batSL-CoVZXC21 et bat-SL-CoVZC45, respectivement [Lu,2020].

Parmi les sarbécocoronavirus, des variations d'acides aminés dans la protéine S ont été identifiées. Bien que le SARS-CoV et le SARS-CoV-2 appartiennent à différents clades dans l'arbre phylogénétique, ils ont 50 acides aminés conservés dans le domaine S1 de la protéine S. Cependant, le MERS-CoV présente des différences mutationnelles dans S protéines. La plupart de ces événements mutationnels se produisent dans le domaine C-terminal. Plusieurs autres les protéases sont également impliquées dans différents processus, tels que l'entrée du virion, la maturation de polyprotéine et assemblage de différentes particules de virion. Outre la protéine S, une variété de SARS-CoV-2 d'autres protéines présentent une similitude avec les protéines d'autres Coronaviridae membres de la famille, comme indiqué dans le **tableau 2**.

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

**Tableau II.2.** Pourcentage d'identité entre les protéines du SRAS-CoV-2 et de la famille des Coronaviridae.

Gene	SARS NC_004718.3	Bat MG772934.1	Bat DQ022305.2
ORF1ab	86.12%	95.15%	85.78%
ORF3a	72.36%	92.00%	72.99%
ORF6	68.85%	93.44%	67.21%
ORF7a	85.25%	88.43%	88.52%
ORF7b	81.40%	93.02%	79.07%
ORF8	30.16%	94.21%	57.02%
ORF10	72.45%	73.20%	74.23%
S (Spike)	75.96%	80.32%	76.04%
E (Envelope)	94.74%	100%	94.74%
M (Membrane)	90.54%	98.65%	90.99%
N (Nucleo-capsid)	90.52%	94.27%	89.55%

### II.2.3.2 Domaine de liaison au récepteur (RBD)

Le RBD du SARS-CoV-2 se trouve dans le domaine C-terminal de la protéine de pointe comme dans SARS-CoV, Bat CoV HKU4 et MERS-CoV . Il a également été signalé que le SRAS-CoV-2 utilise l'ACE2 (enzyme de conversion de l'angiotensine) comme récepteur cellulaire pour l'entrée dans l'humain cellules. À partir de l'analyse phylogénétique, il a été constaté qu'au niveau du génome, le SARS-CoV-2 est étroitement lié à bat-SL-CoVZXC21 et bat-SL-CoVZC45, bien que le RBD du SARS-CoV-2 est très similaire au SRAS-CoV. Cependant, les résidus clés du domaine de liaison au récepteur responsables de la liaison du récepteur étaient différents dans le SRAS-CoV-2 par rapport au SRAS-CoV. D'après les études ci-dessus, il est à nouveau établi que bien que le SARSCoV-2 présente une grande similitude avec le MERS-CoV, le SARS-CoV et d'autres dérivés de chauves-souris coronavirus, il s'agit d'une nouvelle version du coronavirus et est responsable d'une infection qui est se répandre dans le monde.

### II.2.3.3. Organisation génomique du SRAS - CoV2

Le SARS-CoV-2 est un virus à ARN positif simple brin de ~ 29,9 Ko en taille. Le SARS-CoV-2 possède 14 cadres de lecture ouverts (ORF), qui codent pour 27 protéines différentes. Il a 5 ' région non traduite (UTR), la réplication complexe (ORF1a et ORF1b), gène Spike (S), Enveloppe gène (E), gène de la membrane (M), gène de la nucléocapside (N),

3 ' UTR, plusieurs ORF non structuraux non identifiés et une queue poly (A) . Le gène ORF1a est situé à le 5'UTR, code pour la polyprotéine pp1a, qui contient 10

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

nsp. Le gène ORF1b, situé à côté d'ORF1a, code pour la polyprotéine pp1ab qui contient 16 nsp. Les protéines pp1ab et pp1a subissent un clivage autoprotéolytique pour former le complexe de réplication virale. Le 3'UTR contient les quatre gènes de structure et gènes accessoires. Les gènes accessoires sont répartis entre les gènes de structure et leur fonction est pour la plupart inconnus. Le SARS-CoV-2 est un virus enveloppé non segmenté d'un diamètre de 50 à 200 nm. Structurellement, il a une enveloppe lipidique à double couche, y compris la glycoprotéine de Spike (S), la protéine d'enveloppe (E), Glycoprotéine membranaire (M) et protéine nucléocapside (N) [Pachetti,2020].

Le génome viral ayant un RBD pour l'interaction avec les récepteurs des cellules hôtes est couvert par le Spike glycoprotéine. La glycoprotéine Te M est responsable de l'assemblage de particules virales a trois domaines, le domaine cytoplasmique, le domaine transmembranaire et le domaine hydrophile N. Te Enveloppe protéine est signalé à jouer un rôle dans la pathogenèse car il interagit avec la protéine liée aux jonctions serrées PALS1. Toi La protéine de la nucléocapside emballe le génome viral dans un complexe ribonucléoprotéique.

La nucléocapside, une phosphoprotéine joue un rôle dans la réplication du génome viral et la voie de signalisation cellulaire (Figure II.6).

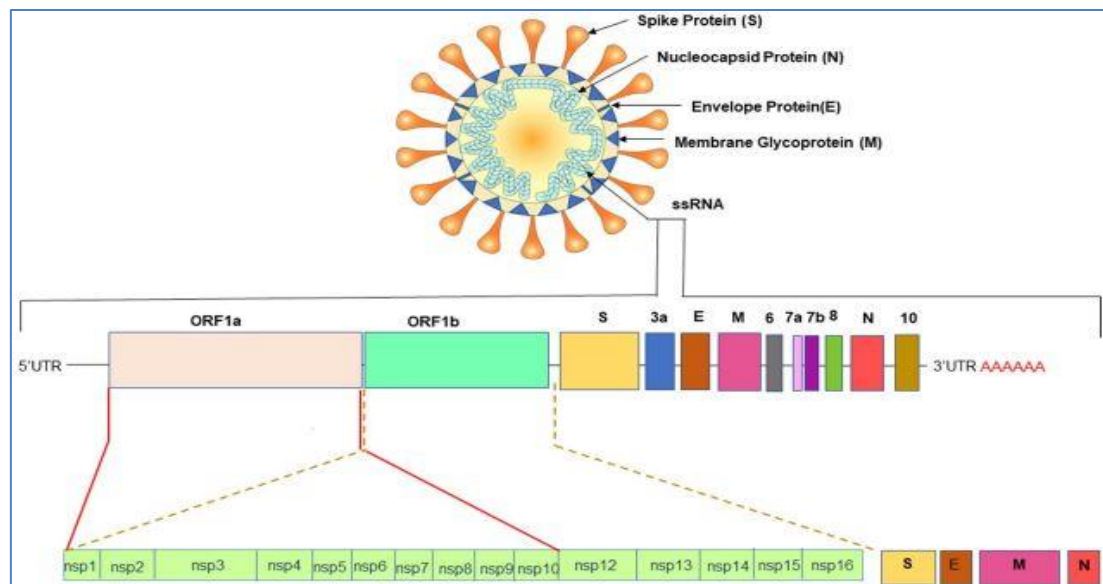


Figure II 6. Présentation schématique de la structure du génome du SRAS-CoV-2.

Le SARS-CoV-2 a une structure sphérique. Le virus a une enveloppe lipidique externe, recouvert de glycoprotéine de pointe. Le SARS-CoV-2 représente une

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

---

organisation typique du génome du Betacoronavirus. Le génome à ARN complet comprend environ 29 903 nucléotides et possède un complexe de réplicase (composé d'ORF1a et d'ORF1b) au 5'UTR.

L'ORF1a encode pour nsp1–nsp10, tandis que ORF1b code pour nsp1–nsp16. Quatre gènes qui codent pour les protéines structurales : gène Spike, gène Enveloppe, Membrane gène, gène nucléocapside et une queue poly (A) à la 3'UTR. Les gènes accessoires sont répartis entre les gènes de structure.

### II.2.4 Mutations récentes du SRAS-CoV-2

Les virus à ARN ont un taux de mutation élevé mais pour survivre, ils atteignent un équilibre de adaptations via des mutations et un génome capable de se répliquer. Par exemple, le SARCOV de type sauvage a signalé moins d'une mutation dans un génome séquencé (fréquence d'erreur de  $2 \times 10^{-5}$ ), alors que le SARS-CoV ExoN (-) a révélé 10 mutations dans un génome séquencé (fréquence d'erreur de  $3 \times 10^{-4}$ ). Comme d'autres virus à ARN, le SRAS-Cov-2 a une forte le taux de mutation et ses multiples variantes ont été rapportés dans le monde (**Pachetti,2020**). La recherche a révélé que 93 mutations ont été observées sur l'ensemble du génome du SRAS-CoV-2. Dans toutes les grandes protéines d'enveloppe, à l'exception de la protéine d'enveloppe, 42 mutations faux-sens ont été identifiées. Là étaient 29 mutations faux-sens dans la polyprotéine codée par ORF1ab, 8 dans la glycoprotéine de la surface de la pointe, 1 dans la protéine matricielle et quatre dans la protéine de la nucléo-capside. La glycoprotéine de surface du pic a un rôle important dans la liaison avec les cellules hôtes et régule la réponse de l'hôte. C'est également la cible clé de la neutralisation des anticorps. Mutation dans la glycoprotéine présente à la surface de la pointe peut induire des changements conformationnels importants qui peuvent modifier l'antigénicité. Dans une autre étude, 5775 génomes différents des mutations ont été signalées, dont 2 insertions dans le cadre, 11 suppressions de décalage du cadre, 36 variantes stop-gainées, 66 insertions dans la région non codante, 100 suppressions en trame, 142 délétions dans les régions non codantes, 484 mutations dans les régions non codantes, 1965 mutations synonymes et 2969 mutations faux-sens .

Le Royaume-Uni a fait face à une augmentation spectaculaire des cas de COVID-19 qui a entraîné une augmentation du nombre de enquêtes épidémiologiques et virologiques. La séquence du génome viral a révélé que un nombre important de

## **Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV**

---

cas appartiennent à un groupe phylogénétique nouvellement signalé. Le nouveau variante du SRAS-CoV-2 (B.1.1.7) a de multiples mutations de la protéine de pointe, (suppression 69-70, P681H, T716I, S982A, suppression 144, N501Y, D614G, A570D, D1118H) en plus de divers mutations dans différentes régions génomiques. Bien qu'il soit clairement compris et prédit que les virus continueront d'évoluer en raison de mutations conduisant au développement de nouvelles variantes, des recherches préliminaires au Royaume-Uni indiquent que cette variante est nettement plus infectieux que ses variantes précédentes, ayant le potentiel d'augmenter la reproduction nombre de 0,4 ou plus avec une transmissibilité accrue jusqu'à 70 pour cent. Ce nouveau La variante est apparue à l'époque de l'année où la mixité sociale s'est traditionnellement accrue. À ce stade, il n'y a aucune preuve d'une sévérité accrue de l'infection liée à ce nouveau variante du coronavirus. Cette variante se répand également dans différents pays à travers le monde. Une autre variante connue sous le nom de B.1.351 est apparue indépendamment en Afrique du Sud. Il présente de nombreuses similitudes avec la variante B.1.1.7 et a été détecté début octobre 2020. Cette variante s'est également propagée aux États-Unis et des cas causés par cette variante ont été signalés là-bas début janvier 2021. Une variante connue sous le nom de P.1 est apparue au Brésil début janvier 2021.

Il a été découvert pour la première fois chez des voyageurs lors de leur dépistage régulier au Japon. Cette variante a des mutations supplémentaires qui peuvent affecter la capacité d'un anticorps à identifier SRAS-CoV-2. Ces souches nouvellement identifiées semblent se propager plus rapidement. C'est prévu pour éventuellement augmenter les cas de coronavirus dans le monde, ce qui pourrait entraîner un plus grand fardeau sur les hôpitaux et plus de décès (Centers for Disease Control and Prevention).

### **II.3. bat coronavirus RATG13 (bat COV RATG13)**

Le coronavirus de chauve-souris RaTG13 est un bêtacoronavirus semblable au SRAS qui infecte la chauve-souris en fer à cheval *Rhinolophus affinis*.

RaTG13 est un virus à ARN à brin positif avec une membrane externe. Son génome est d'environ 29 800 nucléotides. Le génome code pour une réplicase (ORF1a/1b) et quatre protéines structurelles ; comprenant une protéine de pointe (S), une protéine de membrane (M), une protéine de membrane externe (E) et une protéine

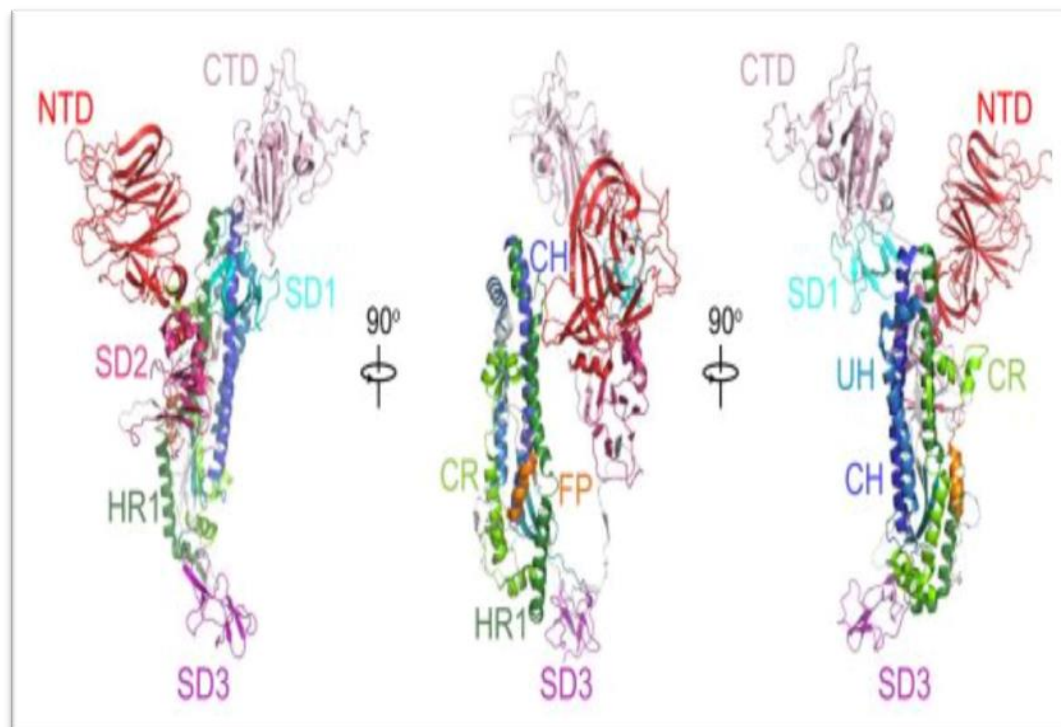
## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

de capside (N); et cinq protéines auxiliaires non structurales, dont NS3, NS6, NS7a, NS7b et NS8, qui sont courantes dans les coronavirus [Shuyuan Zhang, 2021].

### II.3.1-structure BAT COV RATG13

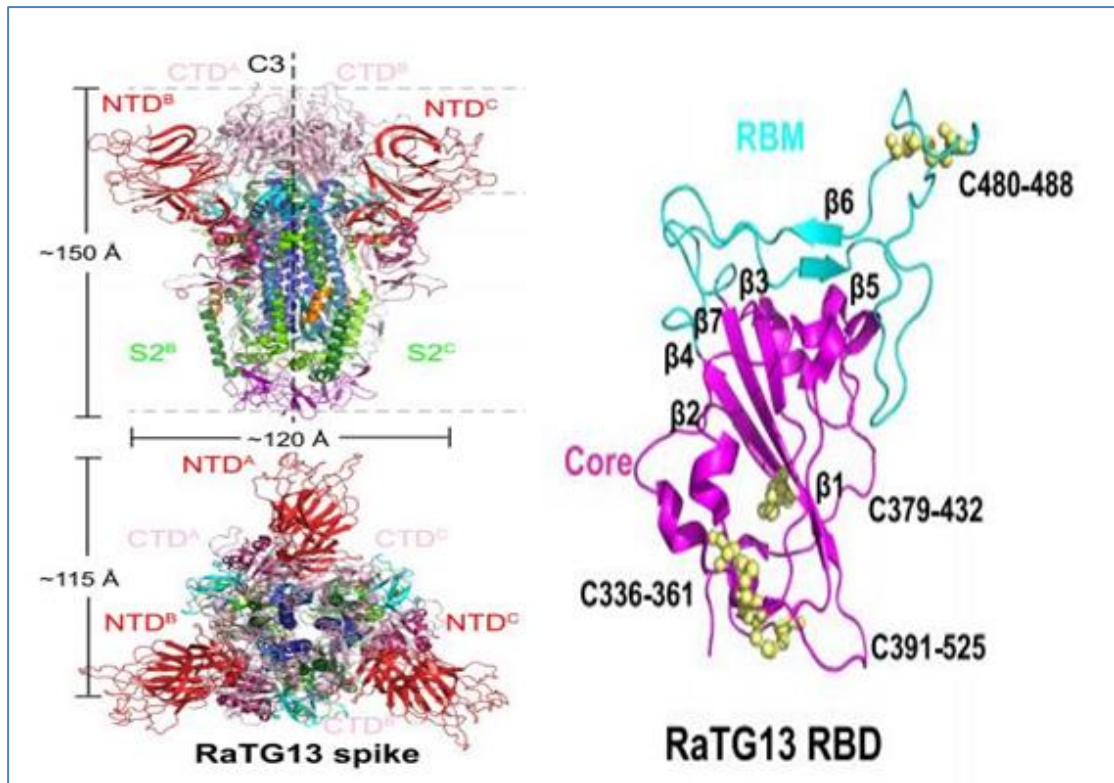
Structures globales des pointes de homotrimérique et du pic ressemblent à la structure de pré-fusion précédemment rapportée du pic de coronavirus (Figure II.7). pointe avoir une forme de champignon (~150Å de hauteur et ~120Å de largeur), composé d'un capuchon principalement formé par un brin une tige principalement formée par des hélices (Fig. 1a). Comme les autres coronavirus, le RaTG13 les monomères de pointe sont composés des sous-unités S1 et S2 avec un site de clivage de la protéase entre eux.

La structure les composants du pic comprennent le domaine N-terminal (NTD), Domaine C-terminal (CTD), sous-domaine 1 (SD1) et sous-domaine 2 (SD2) dans la sous-unité S1 ; et l'hélice amont (UH), fusion peptide (FP), région de connexion (CR), répétition heptade 1 (HR1), hélice centrale (CH), épingle à cheveux  $\beta$  (BH), sous-domaine 3 (SD3) et heptade répéter 2 (HR2) dans la sous-unité S2 (Shuyuan Zhang, 2021)



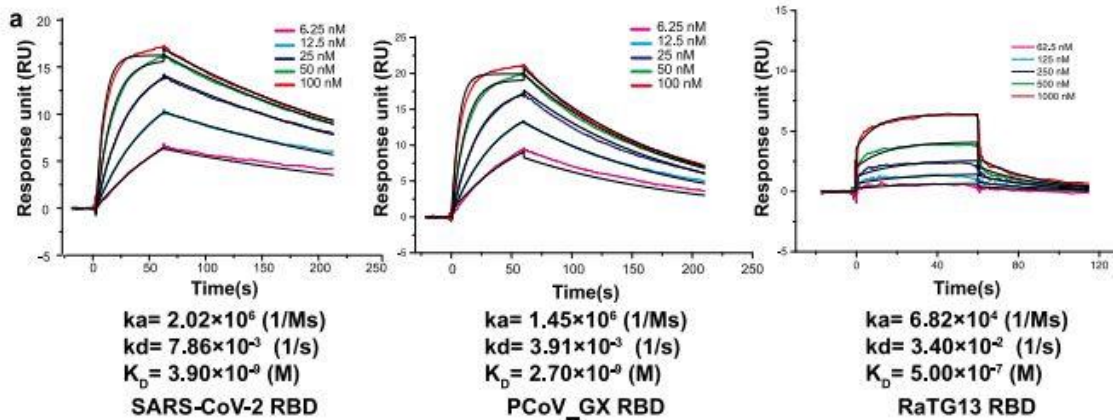
**Figure II.7.** Diagrammes de dessins animés illustrant trois orientations du monomère de pointe coloré

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

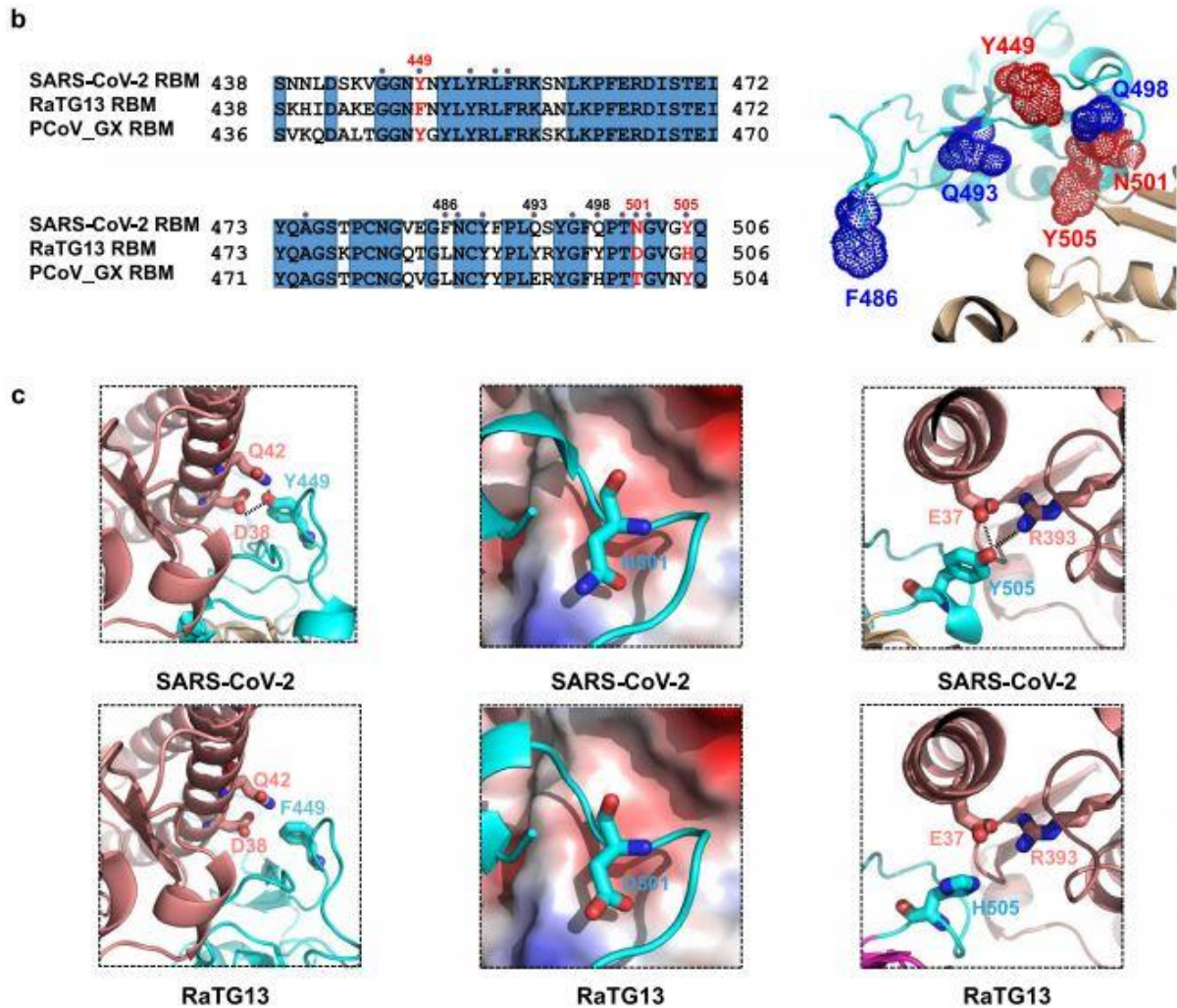


**Figure II.8** Structures globales des glycoprotéines de pointe RaTG13 et RBD

Structures RBD des protéines de pointe RaTG13. Le RBD RaTG13 sont montrés en vue latérale. Le cœur RaTG13 RBD est coloré en magenta.



## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV



## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

---

(c) Sites importants représentant différentes forces de liaison de hACE2 avec les RBD RaTG13 et SARS-CoV-2 (interface SARS-CoV-2 RBD-hACE2 : ID PDB 6M0J). Les RBD sont montrés sous forme de bande dessinée cyan et le hACE2 est montré dans une bande dessinée de bleu ou une surface avec un potentiel électrostatique. Liaisons hydrogène entre SARS-CoV-2 Y449 et hACE2 D38 et Q42 seraient abolies après la mutation Y à F dans le RBM RaTG13 (deux panneaux les plus à gauche). SRAS-CoV-2 La mutation N501 en D501 dans RaTG13, le contact avec la surface locale chargée négativement de hACE2 défavoriserait la liaison de hACE2 avec le RaTG13 RBD (deux panneaux du milieu). Les liaisons hydrogène entre SARS-CoV-2 Y505 et hACE2 E37 et R393 seraient abolies après mutation Y à H dans le RBM RaTG13 (deux panneaux les plus à droite)[Shuyuan Zhang,2021].

## Chapitre II Revue génétique sur les SRAS-CoV et SARS-CoV-2 Bat-CoV

**Tableau II.3:** Fonctions des protéines non structurales du SARS-CoV-2 [Meghana Rastogi, 2020]

Protéine	Fonction
Nsp1	Interfère avec la liaison de l'ARNm et supprime les fonctions immunitaires de l'hôte Ancre le complexe de réplication aux membranes cellulaires Dégrade l'ARNm de l'hôte en interagissant avec la sous-unité ribosomique 40S humaine
Nsp2	Héberge des mutations qui la rendent plus contagieuse Pourrait jouer un rôle dans la modulation de la survie des cellules hôtes ; également connu sous le nom d'homologue p65
Nsp3	Papain-like protease 2 (PL2pro) impliquée dans le clivage protéolytique
Nsp4	Responsable de la formation de la double vésicule membranaire lors de la réplication Ancre le complexe réplication-transcription virale aux membranes du réticulum endoplasmique
Nsp5	Protéases (3CLpro, Mpro) impliquées dans le clivage des polypeptides
Nsp6	Empêche l'expansion de l'autophagosome, aide à la formation de vésicules à double membrane; supprime la signalisation IFN-I
Nsp7	Forme un hexadécamère avec nsp8 et agit comme une primase dans la réplication virale
Nsp8	Agit comme une primase avec nsp7
Nsp9	Agit comme protéine de liaison à l'ARNss
Nsp10	Joue un rôle dans la méthylation de la coiffe d'ARNm viral. Stimule les activités nsp14 3'-5' exoribonucléase et 2'-O-méthyltransférase (NSP16)
Nsp12	Sous-unité catalytique de l'ARN polymérase dépendante de l'ARN ; Catalyse la synthèse d'ARN viral en utilisant nsp7 et nsp8 comme cofacteurs
Nsp13	Activité hélicase et NTPase : hydrolyser les NTP et dérouler l'ARN et l'ADN duplex avec une queue simple brin 5' dans une direction 5' à 3' Un puissant antagoniste de l'interféron
Nsp14	Guanine-N7 méthyltransférase, un complexe multienzymatique Agit des deux côtés ARNsb et ARNdb dans une direction 3' -> 5' Un puissant antagoniste de l'interféron Il joue un rôle dans la réplication du génome, la synthèse et la recombinaison d'ARN sous-génomique
Nsp15	Il s'agit d'une endoribonucléase spécifique à l'uridylylate d'ARN nidoviral (NendoU); joue un rôle dans la réplication et la transcription virales Un puissant antagoniste de l'interféron
Nsp16	Agit en tant que 2'-O-méthyltransférase qui médie la méthylation du 2'-O-ribose de l'ARNm cap à la structure 5'-cap des ARNm viraux

**CHAPITRE III**  
**réponse immunitaire au**  
**SARS-CoV-2**

### III.1. La réponse immunitaire SARS-CoV-2

La réponse immunitaire innée et adaptative joue un rôle essentiel dans la reconnaissance et l'élimination des agents pathogènes étrangers [Luca. P,2020].

L'accumulation de preuves suggère qu'une réponse immunitaire sous-optimale ou non maîtrisée pendant l'infection par le SRAS-CoV-2 détermine les schémas cliniques, la gravité de la maladie et la progression de COVID-19.

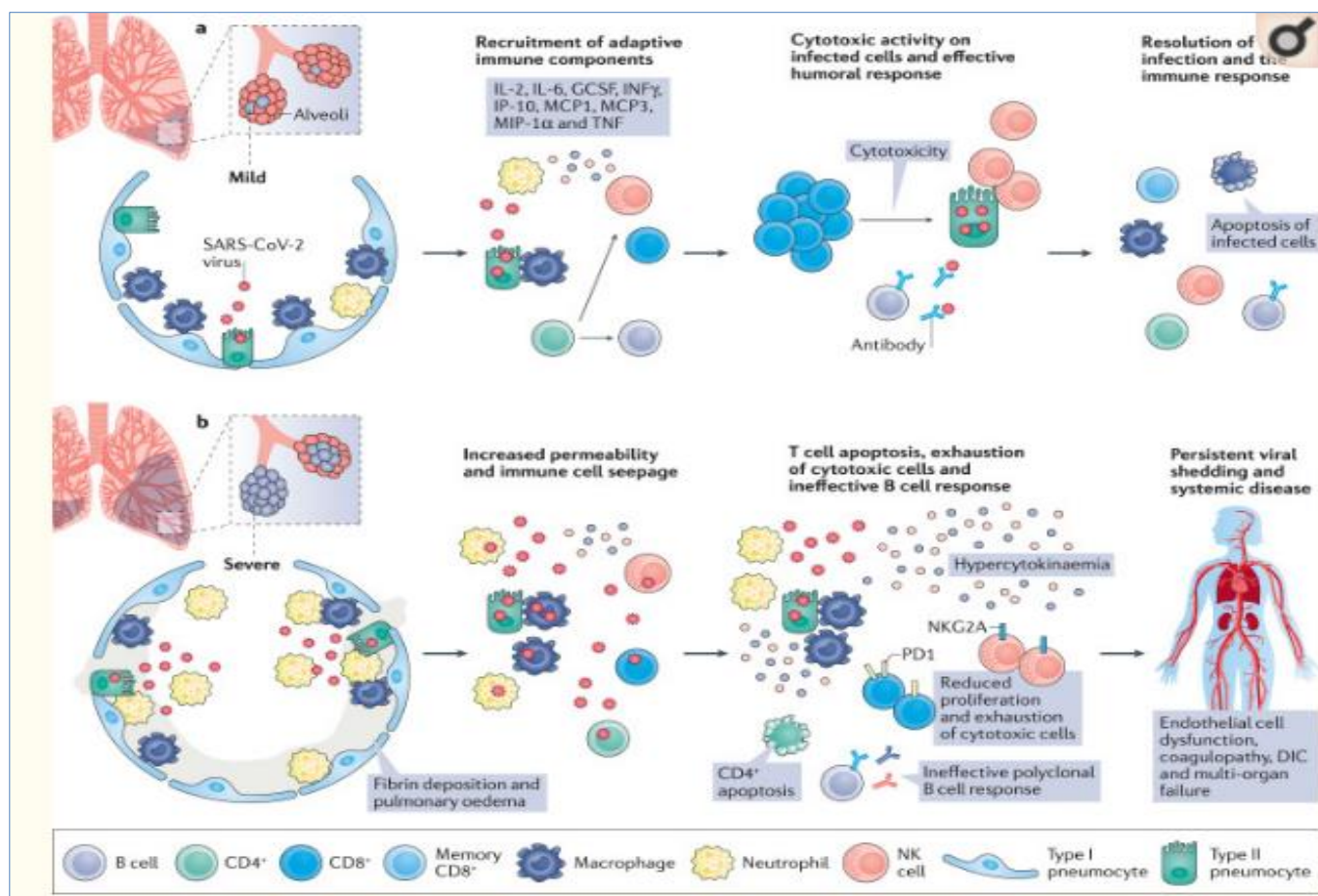
L'identification de signatures immunitaires spécifiques peut fournir de nouvelles informations sur les trajectoires divergentes de la maladie observées chez des patients individuels et aider à la conception d'approches thérapeutiques plus ciblées [Luca. P,2020].

#### III.1.1. Réponse immunitaire innée

Le système immunitaire inné est la première ligne de défense contre les infections virales. Lorsque les virus colonisent leurs cellules cibles, l'ARN viral simple brin, comme celui du SRAS-CoV-2, active les cellules immunitaires innées en engageant l'un des nombreux récepteurs de reconnaissance de formes intracellulaires (*PRRs*), tels que les récepteurs Toll-like (*TLRs*), Les récepteurs similaires au gène I inducible par l'acide rétinoïque (*RIG-I*) et le gène 5 associé à la différenciation du mélanome (*MDA5*), qui détectent les structures d'ARN aberrantes qui se forment souvent lors de la réplication virale

Après l'activation du PRR, les cascades de signalisation moléculaire aboutissent à l'activation de facteurs de transcription en aval, tels que les facteurs régulateurs d'interféron (*IRFs*) et le facteur nucléaire- $\kappa$ B (*NF- $\kappa$ B*) [Luca.P,2020].

Ces facteurs de transcription déclenchent les défenses antivirales cellulaires initiales en induisant l'activation transcription des interférons de type I et III (*IFN-I* et *IFN-III*) et des gènes stimulés par les interférons (*ISG*), ainsi que des cytokines et des chimiokines (*Figure III.1a*) [Luca. P,2020].



**Figure III.1.** Pathogénèse et résultats de CoVID-19 [Luca, P,2020].

La partie "a" (**Figure III.1 a**) présente la suite à une infection des poumons (SRAS-CoV-2) induit la mort des cellules épithéliales, en particulier des pneumocytes de type II, dans le cadre du cycle de réplication virale. L'augmentation des cytokines pro-inflammatoires dans les poumons conduit au recrutement de leucocytes, propageant ainsi la réponse inflammatoire locale. Parmi ces cytokines, l'interleukine-2 (**IL-2**), **IL -6**, facteur de stimulation des colonies de granulocytes (**GCSF**) interféron- $\gamma$  (**IFN $\gamma$** ), protéine inductible **IFN $\gamma$  10 (IP-10)**, protéine chimiotactique monocyte 1 et 3 (**MCP1** et **3**), protéine inflammatoire macrophage **1 $\alpha$  (MIP-1 $\alpha$ )** ) et le facteur de nécrose tumorale (**TNF**) stimulent la réponse immunitaire adaptative. A ce stade, une infiltration de lymphocytes (cellules **T CD4 +** et **CD8 +**) et de cellules tueuses naturelles (**NK**) est nécessaire pour assurer une réponse de défense optimale contre le SRAS-CoV-2. Médiat des lymphocytes **T CD4 +** La production d'anticorps par les cellules B et améliore également les réponses des cellules **T CD8 +** effectrices et **NK** lors d'infections virales[Luca, P,2020].

Concernant la partie "b" (**Figure III.1 b**), pour des raisons inconnues, mais peut-être une prédisposition individuelle ou des différences de charges virales au cours d'une infection primaire par le SRAS-CoV-2, certains patients présentent une maladie plus grave. Une réponse immunitaire inadaptée, caractérisée par une lymphopénie et une suppression des lymphocytes **T CD4 +**, est probablement responsable du mauvais pronostic de ces patients. En l'absence d'activation robuste des lymphocytes **T CD4 +**, les lymphocytes **B** génèrent une réponse anticorps polyclonal qui peut être inefficace pour neutraliser le SARS-CoV-2.

Un nombre accru de cellules **T** épuisées qui expriment des niveaux élevés de protéine de mort cellulaire programmée 1 (**PD1**), suggère une diminution de la prolifération et de l'activité des cellules **T CD8 +**. De même, les cellules **NK** présentent des niveaux accrus du membre A inhibiteur **CD94 – NK** du groupe 2 (**NKG2A**). Une activité cytotoxique altérée entraîne une excrétion virale persistante qui amplifie l'activation des macrophages et des neutrophiles, conduisant à la production massive de cytokines (un processus appelé hypercytokinémie). Chez ces patients, le COVID-19 se manifeste comme une maladie grave, consistant en une pneumonie avancée et un syndrome de détresse respiratoire aiguë. La génération d'un excès de cytokines et une infection virale persistante entraînent des lésions vasculaires systémiques, une coagulation intravasculaire disséminée (CID) et une défaillance des organes vitaux, notamment les reins et le cœur [**Luca. P,2020**].

### III.1.1.2 Mécanisme d'action

Cette vaste réponse antivirale a exercé une pression sélective sur les virus pour qu'ils développent d'innombrables contre-mesures pour éviter les réponses immunitaires de l'hôte, dont la plupart sont mal comprises. Mécaniquement, **NSP1** du SRAS-CoV-2 supprime les réponses immunitaires innées dépendantes de **RIG-I** en se liant à la sous-unité ribosomique 40 S, entraînant l'arrêt de la traduction de l'ARNm de l'hôte [**Wang. D,2020**].

Le SRAS-CoV-2 peut également échapper aux mécanismes de défenses immunologiques grâce à la liaison indirecte d'**ORF9b** à la protéine de signalisation antivirale mitochondriale (**MAVS**) via la translocase de la membrane mitochondriale externe 70 (**TOM70**) 48. La localisation de **MAVS** sur la membrane mitochondriale externe est essentielle pour l'activation de **l'IFN-I** après l'infection de virus à ARN, et il est possible que la liaison du SARS-CoV-2 au **MAVS** empêche sa translocation et l'activation de **l'IFN-I**. À l'appui de cette hypothèse, l'infection par le SRAS-CoV-2

est associée à des niveaux d'**IFN-I** et d'**IFN-III** plus faibles et à une réponse **ISG** modérée par rapport à d'autres virus respiratoires, bien qu'elle soit caractérisée par une production vigoureuse de cytokines inflammatoires [Wang.D,2020].

Conformément aux résultats décrits ci-dessus, les anomalies de laboratoire les plus fréquentes chez les patients présentant une progression défavorable du COVID-19 incluent un profil de cytokine caractérisé par des niveaux accrus d'**IL-2**, **IL-6**, facteur de stimulation des colonies de granulocytes (**G-CSF**), Protéine inductible par l'**IFN $\gamma$**  10 (**IP-10**, également appelée **CXCL10**), protéine chimiotactique des monocytes 1 et 3 (**MCP1** et **MCP3**, également appelées **CCL2** et **CCL7**, respectivement), protéine inflammatoire des macrophages **1 $\alpha$**  (**MIP-1 $\alpha$** , également appelée **CCL3**) et le facteur de nécrose tumorale (**TNF**). Parmi ces cytokines et chimiokines dérégulées, l'évaluation combinée d'**IP-10** et de **MCP3** avait la valeur prédictive la plus élevée de mauvais résultats chez les patients atteints de COVID-19[Wang.D,2020].

Ces résultats rappellent le syndrome hyper-inflammatoire, la lymphohistiocytose hémophagocytaire secondaire et, par conséquent, la réponse immunitaire à l'infection par le SRAS-CoV-2 était initialement considérée comme impliquant une forme de tempête de cytokines, similaire à celle trouvée chez les patients atteints d'autres infections virales ou la septicémie, ainsi que chez les patients atteints de maladies auto-immunes telles que l'arthrite. Cependant, ce concept a depuis été revisité et la contribution des cytokines au COVID-19 remise en question. Dans COVID-19, l'activation des macrophages semble faire partie intégrante de l'initiation et de la propagation de cette réaction hyper-inflammatoire (**Figure III. 1b**).

Une étude préliminaire a détecté l'**ACE2** à la surface des macrophages et a montré que les protéines de pointe du SRAS-CoV-2 interagissent avec les macrophages **CD68+**, suggérant une infection virale directe de ces cellules.

Le séquençage de l'ARN unicellulaire des cellules immunitaires du liquide de lavage broncho-alvéolaire de patients atteints de COVID-19 a suggéré l'expansion des monocytes pro-inflammatoires et des macrophages dérivés des monocytes positifs à la Ficoline, parallèlement à une diminution des macrophages alvéolaires réparateurs résidant dans les tissus chez les patients atteints de graves maladie par rapport à celle observée dans les cas modérés. Une autre étude élégante de séquençage d'ARN monocellulaire a confirmé un enrichissement de l'ARN du SRAS-CoV-2 dans la population de macrophages pulmonaires de patients infectés, et une analyse

différentielle de l'expression des gènes a révélé que les macrophages infectés avaient une expression de chimiokine plus élevée que les macrophages non infectés [Wang, D,2020].

De plus, l'utilisation de la métabolomique non ciblée a montré que l'infection par le SRAS-CoV-2 déclenche de profondes altérations du phénotype et de la fonction des macrophages en régulant à la baisse l'expression des apolipoprotéines qui sont importantes pour une polarisation correcte des macrophages vers le phénotype anti-inflammatoire. Mécaniquement, l'interactome virus-hôte et l'analyse protéomique des cellules sanguines périphériques de patients présentant divers symptômes de COVID-19 suggèrent que **NSP9** et **NSP10** du SRAS-CoV-2 interagissent avec le répresseur NF-kB NKRF, qui induit un système immunitaire inné trop exubérant. réponse à l'infection par le SRAS-CoV-2 .

Une analyse du transcriptome des cellules du liquide de lavage broncho alvéolaire de huit patients atteints de COVID-19 a suggéré que l'infection par le SRAS-CoV-2 provoque une infiltration de neutrophiles plus prononcée que d'autres formes de pneumonie<sup>65</sup>. Conformément à ces résultats, la neutrophilie est prédictive de mauvais résultats chez les patients atteints de COVID-19 de plus, des études publiées sous forme de préimpression ont identifié un rapport neutrophiles/lymphocytes élevé comme facteur de risque indépendant de maladie grave<sup>66</sup>. Une nouvelle étude a également indiqué que l'activation des neutrophiles et la formation de pièges extracellulaires de neutrophiles (NET) peuvent jouer un rôle de premier plan dans la propagation de la libération sévère de cytokines observée dans les cas graves de COVID-19, représentant une réponse de défense innée auparavant non reconnue, mais puissante qui peut aggraver les lésions pulmonaires chez les patients atteints, Ces résultats suggèrent qu'une immunité innée déséquilibrée peut exacerber la réponse hyper-inflammatoire dans les cas graves de la maladie (**Figure III. 1b**) [Wang, D,2020].

### III.1.2. Réponse immunitaire adaptative

La lymphopénie est une caractéristique commune du COVID-19 sévère, caractérisée par des nombres absolus considérablement réduits de cellules **CD4** + et en particulier de cellules **CD8** + **T27** cet épuisement est en corrélation avec la gravité du COVID-19 et la mortalité associée, Une analyse protéomique des cellules sanguines périphériques de patients atteints de COVID-19 a signalé une diminution du nombre de cellules T, associée à des indices réduits d'activation et de fonction des

cellules T, y compris une expression plus faible des sous-unités du récepteur des cellules T (TCR) (**CD3e**, **CD3g**, **CD247**, **TRAC** et **TRBC1**) chez les patients atteints d'une maladie grave que chez les patients atteints d'une maladie bénigne. [Qin, C,2020]

Niveaux des molécules de surface des cellules T (**CD4**, **CD8a**, **CD8b** et **CD2**), des stimulateurs de la migration des cellules T (**DDP4**), des kinases de signalisation du TCR (**ZAP70**, **LCK** et **FYN**) et des molécules du CMH de classe II (**HLA-DRA**, **HLA-DRB1**, **HLA-DRB4** et **HLA-DRB5**) étaient également significativement plus faibles chez les patients atteints d'une maladie grave, indiquant une altération globale de la réponse immunitaire adaptative, Une diminution du nombre de cellules T périphériques peut résulter d'une migration accrue vers les poumons ou d'autres tissus infectés. [Qin, C,2020]

Cependant, bien que les cellules **CD4 + T** puissent être détectées dans les tissus pulmonaires de patients décédés atteints de COVID-19, les cellules **CD8 + T** sont rarement observées, Les cellules **CD8 + T** périphériques de patients atteints de COVID-19 expriment des niveaux élevés de marqueurs d'épuisement, y compris la protéine de mort cellulaire programmée 1 (**PD1**) et la mucine-3 d'immunoglobuline des cellules T (**TIM3**) à noter, ce modèle d'expression était plus prononcé chez les patients nécessitant des soins intensifs que chez les patients atteints d'une maladie bénigne, En plus de confirmer le phénotype épuisé des cellules **CD8 + T** chez les patients atteints de COVID-19, une autre étude a démontré que les cellules tueuses naturelles (NK) étaient également épuisées avec une expression accrue de l'hétérodimère de récepteur inhibiteur **CD94 - NK** groupe 2 membre A (**NKG2A**).

Fait intéressant, le nombre de cellules **CD8 + T** et de cellules NK a été restauré chez les patients convalescents, avec une expression réduite de **CD94 – NKG2A73**, Ces résultats suggèrent que l'augmentation des niveaux de cytokines inflammatoires induite par l'infection par le SRAS-CoV-2 pourrait entraîner l'épuisement des cellules T et NK, qui par la suite ne parviennent pas à détruire les cellules infectées. [Qin, C, 2020].

Cependant, la lymphopénie des lymphocytes T peut également résulter de la mort cellulaire induite par le SRAS-CoV-2. Les résultats d'un examen post-mortem de la rate et des ganglions lymphatiques de patients atteints de COVID-19, publiés sous forme de préimpression, ont identifié des niveaux élevés d'apoptose des lymphocytes T et une expression accrue du récepteur de mort FAS, suggérant peut-être que la mort

cellulaire induite par l'activation (AICD) est probablement responsable de la déplétion des lymphocytes T chez les patients atteints de COVID-19 sévère.

L'analyse par cytométrie en flux a également révélé une expression de FAS significativement plus élevée dans les cellules T circulantes de patients atteints de COVID-19 que chez les témoins [Qin.C,2020].

Il convient de noter qu'un rapport antérieur indiquait que l'ACE2 n'est pas exprimé par les cellules T, excluant la possibilité que le SRAS-CoV-2 ait un effet toxique direct sur ces cellules et soutenant davantage l'idée que l'épuisement des cellules T dans COVID-19 résulte de l'AICD, quelle que soit la cause sous-jacente, cependant, l'absence d'une réponse antivirale efficace des cellules T et des cellules NK pourrait jouer un rôle essentiel dans le maintien d'une charge virale élevée et l'hyperactivation de la réponse innée (Figure . III.1b). Une infection grave causée par le SRAS-CoV-2 a été comparée à une pneumonie lymphopénique communautaire un phénotype immunologique spécifique associé à un risque élevé de mortalité. Ainsi, les défauts des lymphocytes T dans la réponse des lymphocytes T pourraient contribuer à l'excrétion virale soutenue et à la pathogenèse du dysfonctionnement de plusieurs organes chez les patients atteints de COVID-19 avec une maladie cliniquement sévère.

Malgré la réponse altérée des lymphocytes T associée à un COVID-19 sévère, une étude américaine a documenté la présence de lymphocytes T CD4 + et CD8 + circulants spécifiques au SRAS-CoV-2 chez 100 % et 70 %, respectivement, des patients qui avaient guéri de la maladie. La réponse fonctionnelle des cellules CD4 + T contre le SRAS-CoV-2 a été identifiée comme une réponse T helper (TH) classique de type 1, qui était robuste contre la protéine de pointe, la protéine membranaire et la protéine de la nucléocapside. Des réponses supplémentaires des cellules CD4 + T ciblaient NSP3, NSP4, ORF3a et ORF8 chez les patients convalescents atteints de COVID-19.

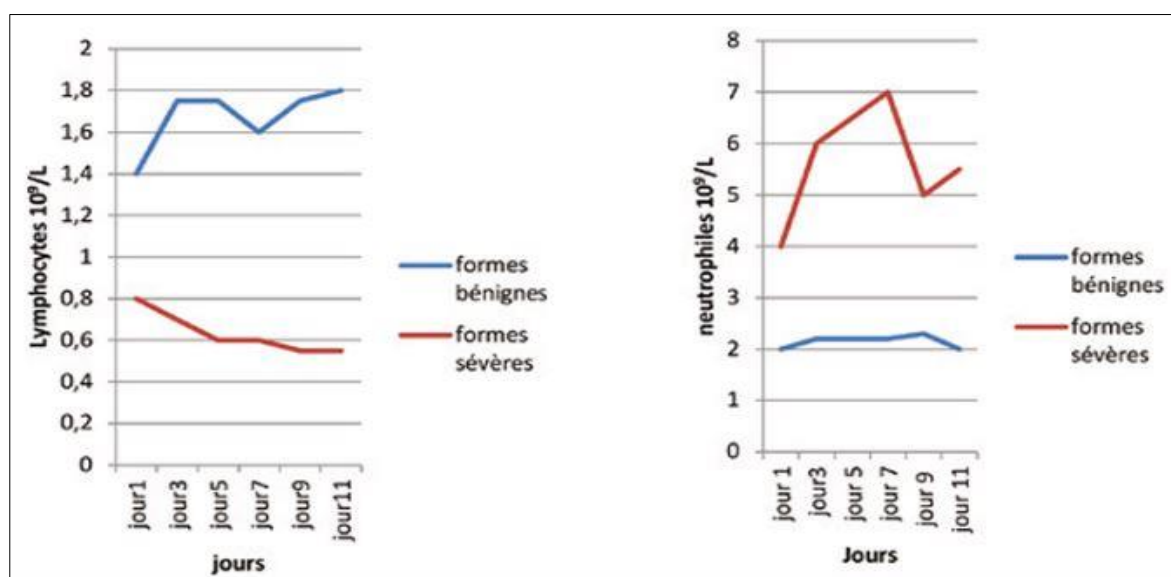
Les cellules CD8 + spécifiques au SRAS-CoV-2 étaient IFN $\gamma$  + et reconnaissaient les protéines de pointe et de membrane, et au moins huit ORF différents du SRAS-CoV-2

La stimulation ex vivo avec des peptides viraux a induit des cellules CD4 + et CD8 + T spécifiques au SRAS-CoV-2 dans le sang des patients atteints de COVID-19 sévère, ces lymphocytes spécifiques du SRAS-CoV-2 étaient déjà détectables à

l'admission en soins intensifs. Notamment, ces études ont systématiquement identifié la présence de cellules T spécifiques au SRAS-CoV-2 dans un large sous-ensemble de personnes convalescentes séronégatives, suggérant que les cellules T pourraient représenter un indicateur plus sensible de l'exposition au SRAS-CoV-2 que les anticorps.

Une analyse d'échantillons collectés de 2015 à 2018 avant l'émergence de la pandémie de COVID-19 a détecté des cellules T spécifiques au SRAS-CoV-2, y compris des cellules T qui reconnaissent NSP7 et NSP13, chez des individus non infectés, suggérant une éventuelle réactivité croisée avec le rhume Coronavirus. Typiquement, une réponse robuste des cellules T est dirigée vers seulement un petit sous-ensemble de peptides potentiellement immunogènes dans une protéine étrangère, et la contrainte majeure qui détermine cette immunodominance est dictée par le traitement antigénique [Qin, C, 2020].

Dans le SRAS, la protéine de pointe du SRAS-CoV représentait près des deux tiers de la réactivité des cellules CD4 + T alors qu'aucune réactivité n'a été trouvée contre les protéines de la nucléocapside et de la membrane. En revanche, les protéines de pointe, de nucléocapside et de membrane du SRAS-CoV-2 semblent être des protéines immunogènes co-dominantes, et des réponses marquées des cellules CD4 + T dirigées contre les protéines NSP et ORF ont également été rapportées.



**Courbe.III.1** valeurS polyS necluiareS neutrophile et des lymphocyte au cours des infection sars-cov-2

L'efficacité et la durée de la protection assurée par l'immunité cellulaire ne sont encore pas définies mais il semble qu'au vu de l'apparition de cellules souches mémoires, l'immunité cellulaire pourrait persister plus longtemps que l'immunité humorale.

Les lymphocytes CD4, CD8 et NK présentent des marqueurs d'activation et d'exhaustion ainsi qu'une perte de leur multifonctionnalité, notamment en termes de cytotoxicité.

L'absence d'une réponse claire des cellules **CD4 + T** immunodominantes au SRAS-CoV-2 pourrait indiquer un traitement antigénique aberrant et un dérèglement conséquent de la réponse immunitaire des cellules T. À l'appui de cette hypothèse, la fréquence des lymphocytes T mémoire spécifiques du virus et le nombre d'épitopes de virus immunodominants auxquels ils ont répondu étaient significativement plus élevés chez les patients atteints d'une maladie grave que chez les patients atteints d'une maladie bénigne<sup>82,84</sup>, ce qui soulève la question de savoir pourquoi ces virus- les cellules T spécifiques ne sont pas capables de contrôler la maladie.

Contrairement aux cellules T, très peu d'études ont démontré un effet du SRAS-CoV-2 sur la biologie des cellules B. À ce jour, les études concernant la fonction des cellules B dans COVID-19 se sont principalement concentrées sur la génération d'anticorps spécifiques contre le SRAS-CoV-2. Dans une étude portant sur 285 patients atteints de COVID-19, la séroconversion pour **IgG** et **IgM** contre la nucléocapside et la protéine de pointe s'est produite simultanément ou séquentiellement et leurs titres ont atteint un plateau dans les 6 jours suivant la séroconversion[**Qin.C,2020**].

Dans les 19 jours suivant l'apparition des symptômes, 100 % des patients étaient **IgG**-positifs. Sur la base de la réponse humorale au SRAS et au MERS, le développement d'anticorps neutralisants contre la protéine de pointe SARS-CoV-2 est probablement l'approche la plus prometteuse pour assurer l'immunité protectrice contre le COVID-19.

Les primates non humains réinfectés par le SRAS-CoV-2 moins d'un mois après leur première infection ont affiché une charge virale réduite et ont empêché la récurrence de la maladie par rapport aux réponses des animaux primo-infectés ces effets bénéfiques étaient associés à de puissantes réponses d'anticorps neutralisants. De plus, le traitement avec deux puissants anticorps monoclonaux humains qui

reconnaissent la protéine de pointe a protégé les macaques rhésus contre un défi avec le SRAS-CoV-2, tel que rapporté par une étude de préimpression<sup>94</sup> - une découverte qui suggère fortement qu'une réponse humorale protectrice peut être induite par la vaccination[**Qin.C,2020**].

Cependant, la mesure dans laquelle les anticorps induits naturellement contre la rechute de la maladie peuvent fournir une protection de longue durée est encore inconnue. La longévité de la réponse humorale contre le SRAS-CoV-2 chez l'homme peut être transitoire, comme le démontre la découverte que 40 % des patients asymptomatiques et 13 % des patients symptomatiques sont devenus négatifs pour les **IgG** anti-Spike au début de la phase de convalescence. Comme le montre une étude de préimpression, l'évaluation longitudinale des patients infectés par le SRAS-CoV-2 a signalé une baisse rapide des niveaux d'anticorps neutralisants au cours d'une période de suivi de 4 mois, avec des titres approchant le seuil inférieur de positivité chez plusieurs individus. La décroissance des anticorps anti-SRAS-CoV-2 a été particulièrement rapide chez les patients atteints de COVID-19 avec une maladie bénigne. Cependant, ces résultats ont été remis en cause par une étude dans la population islandaise montrant que les anticorps contre le SRAS-CoV-2 sont restés stables pendant plus de 4 mois. Plus important encore, les anticorps neutralisants des patients atteints de COVID-19 sont peu mutés, ce qui suggère qu'ils se lient au SARS-CoV-2 avec une faible affinité en raison d'une maturation d'affinité limitée en effet, ces anticorps ont une activité neutralisante modeste, ciblant des épitopes en dehors du domaine de liaison au récepteur, Cette réponse anticorps disparate peut indiquer l'échec à développer une réponse humorale protectrice robuste et durable contre le SRAS-CoV-2,À l'appui d'une réponse dérégulée des cellules B à l'infection par le SRAS-CoV-2, les analyses des cellules B circulantes ont montré une expansion des plasmablastes oligoclonaux et une réduction des fréquences des cellules B mémoire chez les patients atteints de COVID-19 sévère par rapport aux réponses chez les patients atteints d'une maladie bénigne ou chez des individus sains[**Luca.P,2020**].

Les séquences d'anticorps des plus grands clones de cellules B détectées chez les patients atteints de COVID-19 sévère étaient variables en termes de niveaux d'hypermutation somatique et d'utilisation du gène VH, suggérant une réponse polyclonale des cellules B102, pouvant survenir via une voie extra-folliculaire, comme indiqué dans une étude de préimpression. Les niveaux de plasmablaste chez les patients atteints de COVID-19 n'étaient pas en corrélation avec les niveaux

d'anticorps contre la protéine de pointe, suggérant que ces cellules ont généré des anticorps contre des antigènes autres que la protéine de pointe ou ont généré des anticorps avec une faible spécificité et affinité.

Étonnamment, les anticorps anti-SARS-CoV-2 sont élevés chez les patients gravement malades atteints de COVID-19 alors qu'ils seraient faibles chez la plupart des patients guéris<sup>82</sup>. Dans une étude de préimpression sur des personnes atteintes de COVID-19 léger, les titres d'anticorps neutralisants étaient corrélés positivement avec l'âge et les taux sanguins de protéine C réactive, mais corrélés négativement avec le nombre de lymphocytes des patients au moment de l'admission. Malheureusement, cette étude n'a pas évalué la corrélation potentielle entre la charge virale et les symptômes de la maladie avec des titres d'anticorps neutralisants chez des patients individuels pour évaluer le rôle pathogène réel de ces anticorps. Dans ce contexte, une étude qui a suggéré que la réponse humorale contre la protéine de pointe du SRAS-CoV-2 peut augmenter la gravité des lésions pulmonaires en biaisant les réponses des macrophages, en abrogeant les réponses de cicatrisation et en favorisant le recrutement et l'accumulation des macrophages dans le parenchyme pulmonaire, est d'un intérêt particulier[Luca.P,2020].

Il a également été démontré que les niveaux relatifs d'*IgA* et d'*IgG* sont nettement plus élevés chez les patients gravement malades atteints de COVID-19 que chez les patients atteints d'une maladie plus bénigne, ce qui suggère que les *IgA* peuvent avoir des effets synergiques avec les *IgG* en favorisant la cytotoxicité cellulaire dépendante des anticorps et l'exacerbation du COVID-19 .

Ensemble, les preuves disponibles indiquent qu'une variété de processus dérégulés se produisent dans l'immunité innée et adaptative, y compris la réponse cellulaire et humorale à l'infection par le SRAS-CoV-2 (**Figure III.1**). Ces résultats indiquent la nécessité d'études supplémentaires pour mieux comprendre l'immunopathologie COVID-19 et pour aider à la conception du vaccin optimal et des interventions thérapeutiques efficaces[Luca. P,2020].

### III.2.Effets du SARS-CoV-2 sur l'endothélium

Les cellules endothéliales contribuent à la préservation d'une hémostasie normale en maintenant l'intégrité de la paroi vasculaire et en équilibrant la fibrinolyse par l'expression d'inhibiteurs de coagulation et d'enzymes lytiques de caillot sanguin, et par le maintien de la glycocalyx. De plus en plus de preuves suggèrent que la perte de l'intégrité de la barrière vasculaire et le développement d'un endothélium pro-

coagulant contribuent à l'initiation et à la propagation du SDRA dans COVID-19, en induisant une endothéliite et en médiant l'infiltration cellulaire inflammatoire dans les poumons [Grifoni. A, 2020].

Une proportion importante de patients atteints de COVID-19 sévère, y compris les jeunes et les enfants, développent des complications veineuses et artérielles. À New York, sur une période de 2 semaines du 23 mars au 7 avril 2020, les chercheurs ont observé une multiplication par sept des accidents vasculaires cérébraux des gros vaisseaux chez les patients de moins de 50 ans par rapport à une période de 2 semaines au cours des 12 mois précédentes. Bien que les premières preuves suggèrent que les enfants pourraient être moins exposés au risque d'infection par le SRAS-CoV-2 que les adultes en raison de la plus faible expression de l'ACE2 dans l'épithélium nasal, un groupe sans précédent de syndrome de choc hyper-inflammatoire a été observé dans le sud-est de l'Angleterre, au Royaume-Uni. une période de 10 jours à la mi-avril 2020 chez huit enfants présentant une exposition familiale probable au COVID-19[Grifoni.A, 2020].

Ce syndrome a été décrit comme un état pathologique avec des caractéristiques similaires à la maladie de Kawasaki - une maladie rare qui affecte les vaisseaux sanguins, y compris ceux du cœur. Tous les enfants ont été testés négatifs pour le SRAS-CoV-2 lors d'un lavage broncho-alvéolaire ou d'aspirations nasopharyngées, mais les tests de laboratoire ont révélé une infection et une inflammation, notamment des niveaux élevés de protéine C-réactive, de ferritine et de D-dimères. Des preuves plus récentes suggèrent que ce syndrome, maintenant appelé syndrome inflammatoire multisystémique chez l'enfant (MIS-C), présente des caractéristiques hématologiques et immunologiques uniques qui sont distinctes de celles du COVID-19 typique chez l'adulte et de la maladie de Kawasaki. Des cas similaires de MIS-C ont été observés chez 17 patients de moins de 18 ans à New York<sup>120</sup> et à Bergame<sup>121</sup>, l'une des régions les plus durement touchées du nord de l'Italie. Une étude rétrospective de 35 enfants atteints de MIS-C en France a décrit une insuffisance cardiaque aiguë résultant de l'état inflammatoire sévère causé par l'infection par le SRAS-CoV-2 aucun de ces enfants n'avait de maladie cardiovasculaire sous-jacente[Grifoni. A,2020].

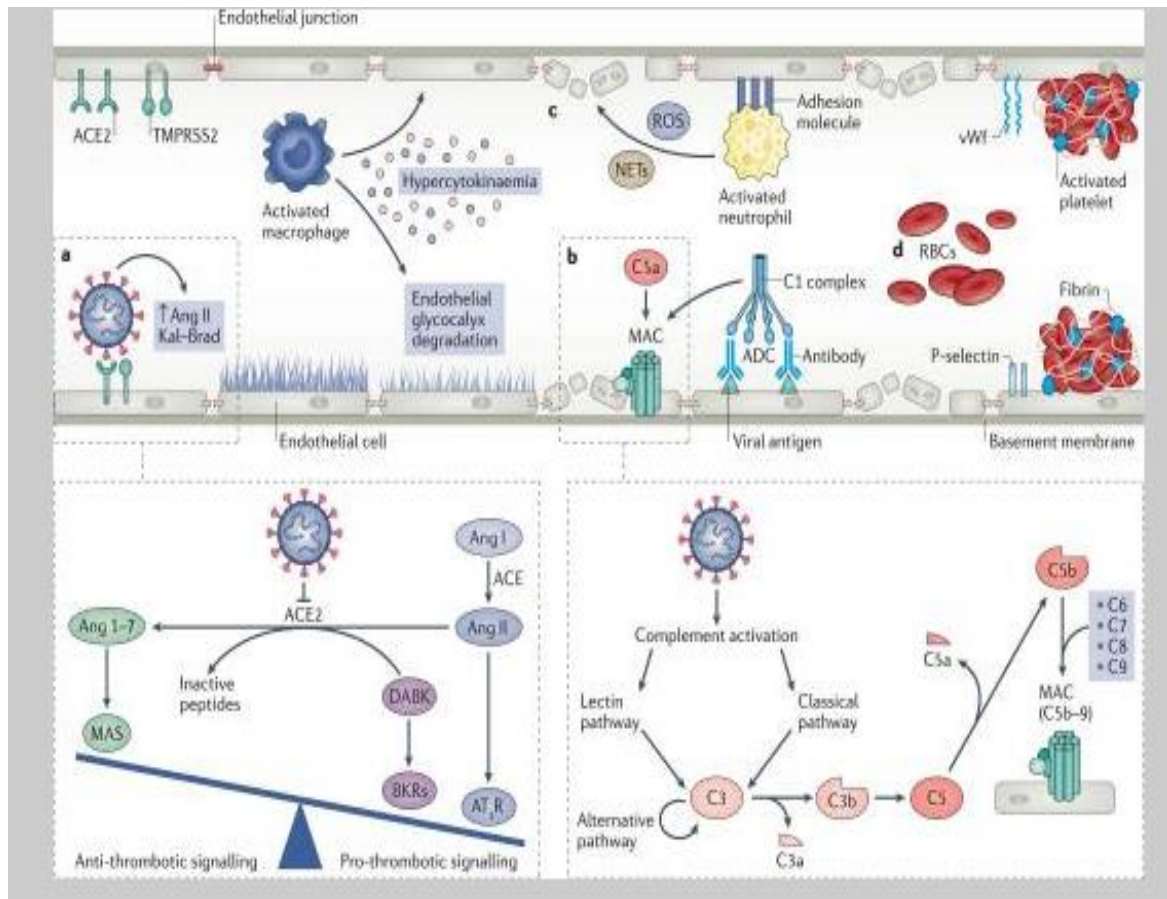
Le SRAS-CoV-2 peut altérer l'homéostasie vasculaire en infectant directement les cellules endothéliales via l'ACE2 (**Figure. III.2a**). À l'appui de cette proposition, le SARS-CoV-2 peut infecter avec succès des organoïdes de vaisseaux sanguins

humains modifiés, confirmant le tropisme du SARS-CoV-2 pour les cellules endothéliales. Cette découverte est encore corroborée par des analyses de microscopie électronique des tissus post-mortem, qui ont montré que le SRAS-CoV-2 peut infecter les cellules endothéliales pulmonaires et induire une endothéliite, au moins dans un sous-ensemble de patients gravement malades. Après liaison par le SARS-CoV-2, l'ACE2 est internalisée et régulée négativement sur les cellules endothéliales, ce qui favorise la progression des processus inflammatoires et pro-fibrotiques dans le poumon déclenchés par l'hyperactivité locale de l'angiotensine II (Ang II). Étant donné que l'ACE2 convertit l'Ang II en Ang (1-7), qui agit sur le récepteur MAS, il est concevable qu'une réduction de l'expression de l'ACE2 par les cellules endothéliales en réponse à l'infection par le SRAS-CoV-2 entraîne une réduction de la génération d'Ang 1-7 et par conséquent, une activation réduite du MAS, favorisant un phénotype local de cellules endothéliales pro-thrombotiques (**Figure. III.2a**). De plus, une expression réduite de l'ACE2 pourrait à son tour activer indirectement le système kallikréine-kinine (KKS), ce qui conduit finalement à une augmentation de la perméabilité vasculaire. En effet, plusieurs études ont documenté qu'un équilibre étroit entre les systèmes KKS et rénine-angiotensine est nécessaire pour réguler la thromborésistance des cellules endothéliales. La réduction de l'ACE2 pulmonaire altère l'inactivation de la bradykinine des-Arg9 (DABK), la laissant libre de se lier à ses récepteurs apparentés (**Figure. III.2a**). De plus, des niveaux élevés de bradykinine chez les patients atteints de COVID-19 pourraient expliquer la majorité des symptômes graves, allant des lésions des vaisseaux sanguins aux complications neurologiques, suggérant que l'inhibition du KKS, par exemple, avec l'antagoniste du récepteur de la bradykinine B2 icatibant, pourrait être un stratégie thérapeutique possible<sup>133,134</sup>. Mécaniquement, la perturbation de la barrière endothéliale par la protéine de pointe SARS-CoV-2 est médiée par des changements dans l'expression de surface de la molécule d'adhésion intracellulaire-1 (*ICAM1*), de la protéine d'adhésion cellulaire vasculaire-1 (*VCAM1*) et de la protéine d'échafaudage de jonction serrée zonula occludens -1 (*ZO-1*), tel que rapporté par une étude de préimpression [**Grifoni. A, 2020**].

Cependant, la découverte que l'ARN viral est rarement détectable dans le sang<sup>19</sup> indique que des facteurs supplémentaires dépendants de l'hôte pourraient contribuer à un dysfonctionnement endothélial systémique et à une vasculopathie dans COVID-19, plutôt que des effets dépendants viraux directs sur les cellules endothéliales. Comme

examiné ailleurs, l'hypercytokinémie et la réponse pro-inflammatoire massive de l'hôte peuvent contribuer au dysfonctionnement endothélial dans le COVID-19, en particulier via les actions de *IL-6* et du *TNF* - dont les niveaux sont fortement augmentés chez les patients gravement malades<sup>113</sup>. Les cytokines exercent leurs effets pathogènes par divers mécanismes. Par exemple, l'*IL-6* peut augmenter la perméabilité vasculaire et favoriser la sécrétion de cytokines pro-inflammatoires par les cellules endothéliales elles-mêmes, amplifiant ainsi la libération de cytokine. Le *TNF* peut activer les glucuronidases qui pourraient dégrader le glycocalyx endothélial mais pourraient également réguler positivement l'acide hyaluronique synthase 2, entraînant un dépôt accru d'acide hyaluronique dans la matrice extracellulaire et favorisant la rétention d'eau [Noris, M,2020].

Des mécanismes supplémentaires, tels qu'une réduction de l'activité de l'oxyde nitrique synthase endothéliale et des niveaux d'oxyde nitrique, ainsi que la libération du facteur de croissance endothéliale vasculaire (*VEGF*) en conséquence de l'hypoxie systémique induite par le SDRA, ont également été proposés comme processus pathogènes clés. Dysfonctionnement endothélial sous-jacent à la suite d'une infection par le SRAS-CoV-2. Ci-dessous, nous nous concentrons sur le rôle possible du système du complément dans l'induction d'un dysfonctionnement endothélial systémique dans COVID-19 [Noris, M.2020].



**Figure III.2.** Effet de l'infection par le SRAS-CoV-2 sur la fonction des cellules endothéliales, la coagulation systémique et la thrombose [Luca, P,2020].

**A)** Le coronavirus 2 du syndrome respiratoire aigu sévère (SRAS-CoV-2) infecte directement les cellules endothéliales en raison de leurs niveaux d'expression élevés de l'enzyme de conversion de l'angiotensine 2 (**ACE2**) et de la protéase sérine transmembranaire 2 (**TMPRSS2**). Après liaison par le SARS-CoV-2, l'ACE2 est internalisé, et l'absence d'ACE2 sur les cellules endothéliales favorise la progression des processus inflammatoires et thrombotiques déclenchés par l'hyperactivité locale de l'angiotensine II (Ang II). L'inhibition de l'ACE2 par la liaison du SARS-CoV-2 réduit la conversion médiée par l'ACE2 de l'Ang II en Ang 1-7, le ligand vasoactif du récepteur MAS. La réduction de l'activation des récepteurs MAS induit un phénotype pro-inflammatoire via une activation accrue des récepteurs de l'angiotensine de type 1 (AT1R). De plus, la réduction des niveaux d'ACE2 limite la dégradation de la bradykinine des-Arg9 (DABK) en peptides inactifs, conduisant finalement à une augmentation de la signalisation pro-thrombotique via l'activation des récepteurs de la bradykinine (BKR) [Noris, M,2020].

**B)** Le SARS-CoV-2 active également le système du complément - une composante intégrale de la réponse immunitaire innée. La cascade du complément peut être activée par trois voies différentes, la classique, la lectine et la voie alternative, qui se résolvent autour de la formation des convertases C3 qui clivent C3, générant le peptide pro-inflammatoire C3a et une grande quantité de C3b qui opsonise les agents pathogènes. C3b forme également la C5 convertase, qui conduit à la libération de la puissante anaphylatoxine C5a, ainsi que du fragment C5b, responsable de la formation du complexe d'attaque membranaire (MAC) C5b-9 sur les cellules cibles, qui est considéré comme le terminal cas d'activation du complément. De plus, des complexes d'anticorps spécifiques au SRAS-CoV-2 et d'antigènes viraux pourraient induire une lésion des cellules endothéliales par l'activation du complexe C1 de la voie classique et l'induction d'une cytotoxicité dépendante des anticorps (ADC [Noris. M, 2020]).

**C)** Les cytokines et chimiokines pro-inflammatoires libérées par les macrophages activés amplifient le cercle vicieux de la perturbation de l'intégrité vasculaire, de la coagulation des vaisseaux et de la thrombose en dégradant le glycocalyx endothélial, en activant le système de coagulation et en atténuant les mécanismes anticoagulants. Le phénotype adhésif des cellules endothéliales induit par les cytokines et les chimiokines inflammatoires favorise l'infiltration des neutrophiles, qui produisent de grandes quantités de médiateurs histotoxiques, y compris les espèces réactives de l'oxygène (ROS) et les pièges extracellulaires des neutrophiles (NET), conduisant finalement à des lésions des cellules endothéliales. [Noris. M,2020]

**D)** Les cellules endothéliales activées initient la coagulation en exprimant la sélectine P, le facteur von Willebrand (vWf) et le fibrinogène, entraînant une liaison plaquettaire massive, la formation de fibrine et la coagulation des globules rouges (GR), entraînant finalement une thrombose systémique et une coagulation intravasculaire disséminée[Noris. M,2020].

### **III.3.Activation excessive des suppléments**

Le système du complément fait partie intégrante de la réponse immunitaire innée et comprend plus de 30 protéines différentes. Ce système peut être activé par trois voies différentes : la classique, la lectine et la voie alternative. Toutes ces voies convergent vers la formation de convertases C3 qui clivent C3 pour générer le peptide pro-inflammatoire C3a et une grande quantité de C3b, qui opsonise les agents pathogènes. C3b forme également la C5 convertase, qui induit la libération de la

puissante anaphylatoxine C5a, ainsi que du fragment C5b qui est responsable de la formation du complexe d'attaque membranaire C5b-9 sur les cellules cibles - l'événement terminal de l'activation du complément. Bien que le complément, comme d'autres composants du système immunitaire inné, aide à contrôler les infections bactériennes et virales, son activation sans restriction - par exemple, dans le contexte d'une infection prolongée par le SRAS-CoV-2 - peut être préjudiciable en causant des dommages directs aux tissus de l'hôte (**Figure. III.2b**).

La plupart des données disponibles sur l'activation du complément dans le contexte d'une infection à coronavirus sont dérivées des épidémies de SRAS et de MERS. Une étude de 2018 sur le SRAS-CoV a démontré que les souris déficientes en C3 présentaient une forme de maladie moins grave que les souris de type sauvage, associée à une infiltration pulmonaire réduite de cellules inflammatoires et à une hypercytokinémie plus faible dans les poumons et les sérums. Le SRAS-CoV active la voie de la lectine en interagissant avec la lectine de liaison au mannose (MBL) du composant du complément - une lectine de type C qui active la sérine protéase 2 associée à la protéine de liaison au mannose (**MASP-2**) - améliorant le dépôt de fragments d'activation C4 sur des cellules infectées par le virus [**Noris. M,2020**].

Chez les souris infectées par le MERS-CoV, les niveaux de C5a circulant ont augmenté après l'infection et l'expression du récepteur C5a (**C5aR**) a été régulée à la hausse dans les pneumocytes et les leucocytes inflammatoires, associée à un fort dépôt de C5b-9 sur les cellules épithéliales bronchiolaires, les pneumocytes et les leucocytes infiltrants. Conformément à ces résultats, les analyses d'immunohistochimie dans le tissu pulmonaire post-mortem de patients infectés par le SRAS-CoV-2 ont révélé une expression élevée des composants du complément MBL, C4, C3 et C5b-9, dans les cellules épithéliales alvéolaires et les pneumocytes, ainsi que dans les cellules inflammatoires, comme le rapporte une étude de prépublication. Comme mentionné ci-dessus, la protéine de nucléocapside du SRAS-CoV-2 peut interagir indirectement avec MASP-2 et initier la voie des lectines du complément, entraînant une activation aberrante du complément et une aggravation des lésions pulmonaires inflammatoires. Les niveaux circulants de C5a étaient plus élevés chez les patients atteints de COVID-19 sévère que chez les patients atteints d'une maladie bénigne qui avaient des niveaux de C5a similaires à ceux des témoins sains [**Noris. M,2020**].

L'utilisation d'une approche métabolomique non ciblée a également démontré une régulation à la hausse des protéines d'activation du complément, y compris des protéines complexes d'attaque membranaire telles que C5, C6 et C8, dans les sérums de 46 patients atteints d'une maladie grave.

En plus de la voie des lectines, la voie classique peut également contribuer aux lésions tissulaires induites par l'infection par le coronavirus. En effet, les complexes immuns formés par les antigènes viraux et les anticorps peuvent activer C1 et initier l'activation du complément. Chez les patients infectés par le SRAS-CoV, le développement d'auto-anticorps contre les cellules épithéliales pulmonaires humaines et les cellules endothéliales a induit une cytotoxicité dépendante du complément. Cependant, on ne sait pas si ce processus est également activé après une infection par le SRAS-CoV-2. L'utilisation d'une approche protéomique clinique à ultra-haut débit a identifié une activation cohérente de la voie classique du complément (*en particulier, les composants du complément C1R, C1S et C8A*) dans les sérums de patients atteints de COVID-19 cliniquement sévère, ainsi qu'une activation accrue de la voie alternative composant, le facteur du complément B (**CFB**) et ses modulateurs CFI et CFH144. Les résultats d'une autre étude suggèrent que des niveaux accrus d'IgA pendant l'infection par le SRAS-CoV-2 pourraient agir en synergie avec les **IgG** pour favoriser la cytotoxicité dépendante des anticorps<sup>108</sup>. Ces chercheurs ont également proposé que COVID-19 pourrait être au moins en partie une maladie médiée par les IgA<sup>108</sup>, induisant éventuellement un phénotype de vascularite via une cytotoxicité médiée par la voie classique [Noris, M,2020].

Bien que les connaissances cliniques sur l'activation du complément à la suite d'une infection par le SRAS-CoV-2 soient limitées, les résultats ci-dessus indiquent que l'activation non limitée du complément - principalement aux phases terminales en aval des cascades du complément - peut contribuer à une lésion des cellules endothéliales, une thrombose et une coagulation intravasculaire, conduisant à la défaillance d'organes multisystémiques chez les patients atteints de COVID-19. [Noris, M.2020]

#### III.4. Microangiopathie et thrombose vasculaire

L'activation du système du complément entraîne une lésion et la mort des cellules endothéliales avec une dénudation vasculaire subséquente et une exposition de la membrane basale thrombogène, ce qui entraîne l'activation des cascades de coagulation. Ces événements entraînent une inflammation, une thrombose

microvasculaire, un œdème vasculaire et des séquelles hémorragiques - qui sont tous des caractéristiques importantes de la pathologie pulmonaire des patients atteints de pneumonie associée au COVID-19. Dans une étude d'autopsie de dix patients atteints de COVID-19, la formation de petits thrombus dans la périphérie pulmonaire a été associée à des foyers d'hémorragie alvéolaire.

Dans une plus grande série d'analyses pulmonaires post-mortem dans le nord de l'Italie, une constatation prédominante était la présence de thrombus plaquettaires-fibrine dans les petits vaisseaux artériels, suggérant en outre que la coagulopathie a un rôle important dans les résultats après l'infection par le SRAS-CoV-2. En effet, les patients infectés par le SRAS-CoV-2 présentent un risque accru de coagulation généralisée des petits et grands vaisseaux, et les complications thromboemboliques veineuses - en particulier les embolies pulmonaires aiguës sont reconnues comme des complications potentiellement mortelles chez les patients atteints de COVID-19[Noris. M,2020].

Les cellules pro-inflammatoires, les cytokines et les chimiokines amplifient le cercle vicieux de la coagulation des vaisseaux et de la thrombose, en agissant simultanément pour activer le système de coagulation et atténuer les mécanismes anticoagulants. Les neutrophiles activés et infiltrés (*Figure. III.2c*) produisent de grandes quantités de médiateurs histotoxiques, y compris des espèces réactives de l'oxygène, ainsi que des *TNE*, qui ont été trouvés dans des microthrombus artériels dans les échantillons pulmonaires post-mortem de quatre patients décédés atteints de COVID-19 (réf .155). L'expression du facteur tissulaire (**TF**) par les monocytes activés stimule la cascade de coagulation et la génération subséquente de thrombine ces processus sont contrebalancés de manière inefficace par l'inhibiteur de la voie du facteur tissulaire (**TFPI**), qui est présent dans le plasma et les plaquettes, et produit principalement par les cellules endothéliales. Parallèlement, les voies anticoagulantes associées aux cellules endothéliales, en particulier le système de la protéine **C**, sont altérées par les cytokines pro-inflammatoires. En réponse à l'activation par l'**IL-1 $\beta$**  et le TNF, les cellules endothéliales favorisent la coagulation en exprimant la P-sélectine, le facteur von Willebrand et le fibrinogène, conduisant à une liaison plaquettaire massive (**Figure. III.2d**).

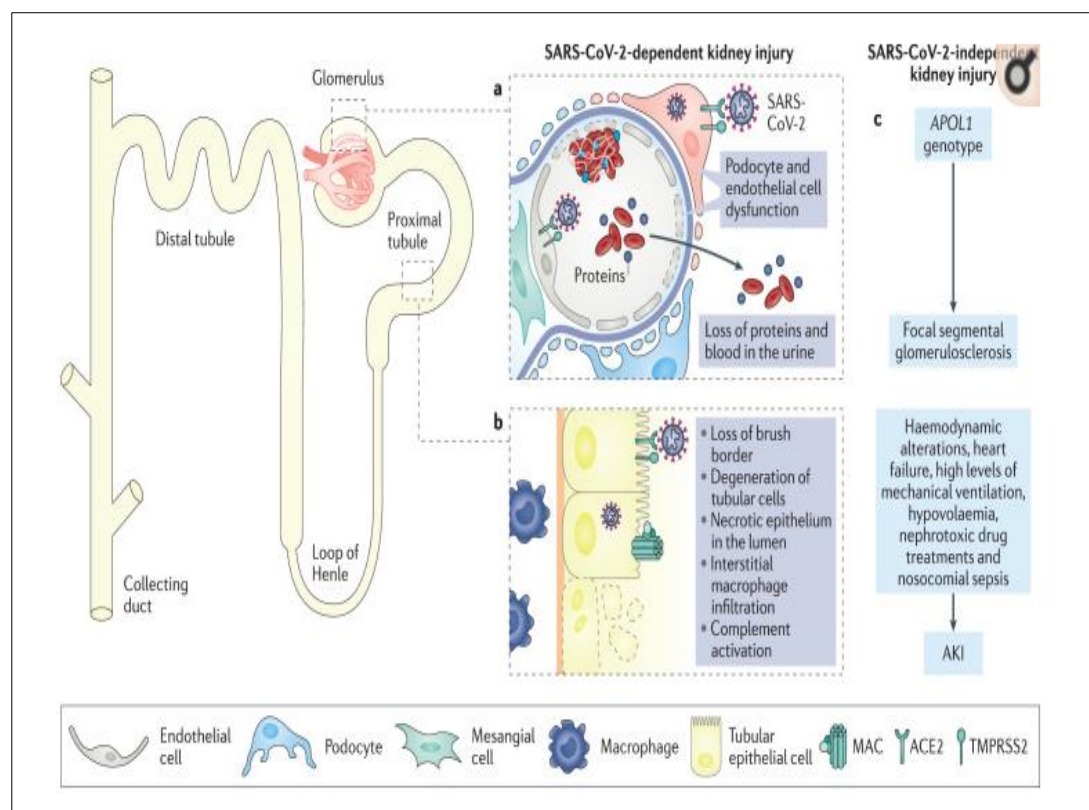
Ce processus est en outre soutenu par la libération de cytokines trophiques par les cellules endothéliales, qui augmentent la production de plaquettes. Les plaquettes recrutées libèrent du VEGF, qui amplifie l'expression du **TF** par les cellules

endothéliales, stimulant davantage la cascade de coagulation. Ensemble, ces événements contribuent à l'augmentation de la génération de fibrine, induisant le dépôt de caillots sanguins dans la microvascularisation (**Figure. III.2**).

#### **III.4.Implication des reins dans COVID-19**

Les poumons sont le site principal de l'infection par le SRAS-CoV-2 et le premier organe affecté par COVID-19. Cependant, l'excrétion virale soutenue et l'hyperactivation de la réponse immunitaire dans les cas graves augmentent considérablement le risque de défaillance systémique multiviscérale. Parmi les organes touchés par COVID-19, le rein est l'un des principaux sites de complications du COVID-19 (**Figure. III.3**), et une fonction rénale anormale est un facteur de risque important de décès chez les patients gravement malades atteints de COVID-19 dans les soins intensifs. Réglage. Dans une étude de cohorte prospective de 701 patients atteints de COVID-19 à Wuhan, en Chine, près de 17 % (113) sont décédés à l'hôpital. A l'admission, 43,9% des patients présentaient une protéinurie et 26,7% d'hématurie, alors que 5,1% des patients ont présenté une insuffisance rénale aiguë (IRA) au cours de l'étude. Des rapports publiés plus tard dans la pandémie décrivaient une incidence plus élevée d'IRA (8 à 17 %), avec les fourchettes les plus élevées chez les personnes gravement malades (14 à 35 %). Chez la plupart des patients atteints de COVID-19, une fonction rénale anormale en présence d'une protéinurie de bas niveau reflète peut-être une lésion tubulaire. [Su. H, 2020]

Dans certains cas, cependant, la protéinurie est abondante ou contient de l'albumine, suggérant une atteinte glomérulaire. Les patients atteints d'insuffisance rénale courent un risque de décès significativement plus élevé que ceux sans insuffisance rénale, et des taux élevés de créatinine sérique de base ou d'azote uréique du sang, le stade AKI, la protéinurie ou l'hématurie étaient des facteurs de risque indépendants de mortalité à l'hôpital. Une revue systématique et une méta-analyse ont suggéré que, chez les patients atteints de COVID-19, l'IRA sévère est un prédicteur clinique de mauvais augure et est associée à une mortalité élevée. En conséquence, la dialyse est un outil important pour la gestion des patients gravement malades atteints d'IRA [Su. H, 2020].



**Figure III.3.** Les effets du COVID-19 sur le rein. [Luca, P,2020]

- Le rein exprime des niveaux élevés d'enzyme de conversion de l'angiotensine 2 (**ACE2**) et de protéase sérine transmembranaire 2 (**TMPRSS2**) et a été identifié comme un organe cible pour l'infection par le coronavirus 2 du syndrome respiratoire aigu sévère (SRAS-CoV-2) par plusieurs études. [Su, H, 2020]

**A)** Dans le glomérule, les podocytes et les cellules endothéliales ont été identifiés comme des sites spécifiques d'infection virale. Le dysfonctionnement des podocytes induit par le SRAS-CoV-2 est susceptible d'induire une altération de la filtration glomérulaire, entraînant une protéinurie et une hématurie, qui sont souvent observées chez les patients atteints de la maladie à coronavirus 2019 (COVID-19). L'infection des cellules endothéliales altère l'hémostase capillaire glomérulaire et induit la formation de thrombus de fibrine.

**B)** Le SRAS-CoV-2 a également été trouvé dans les cellules tubulaires proximales, qui présentent des niveaux élevés d'**ACE2** dans la bordure en brosse apicale. L'infection par le SRAS-CoV-2 entraîne la perte de la bordure en brosse et la dégénérescence vacuolaire dans les cellules épithéliales tubulaires, avec des débris luminaux composés d'épithélium nécrotique avec des preuves d'activation du

complément et de dépôt de complexe d'attaque membranaire (MAC) (C5b-9) sur les cellules tubulaires et infiltration massive de macrophages dans l'interstitium du rein.

C) En outre, un mécanisme non dépendant du virus pourrait également être responsable d'un dysfonctionnement rénal dans le contexte de COVID-19, y compris une contribution possible du génotype de « risque » APOL1 dans l'induction d'une glomérulosclérose segmentaire focale après une infection par le SRAS-CoV-2, comme ainsi que le rôle des facteurs hémodynamiques, du dysfonctionnement cardiaque, des niveaux élevés de ventilation mécanique, de l'hypovolémie, des traitements médicamenteux néphrotoxiques et de la septicémie nosocomiale entraînant une insuffisance rénale aiguë (**IRA**) associée au COVID-19. [Su. H, 2020]

Cependant, de nouvelles preuves suggèrent que le SRAS-CoV-2 pourrait avoir un effet direct sur les cellules rénales. Les cellules épithéliales tubulaires et les podocytes présentent des niveaux élevés d'ACE2 et de TMPRSS2, suggérant qu'ils sont des cibles potentielles du SRAS-CoV-2. L'internalisation de l'ACE2 à la suite d'une infection par le SRAS-CoV-2 pourrait induire un déséquilibre dans le système rénine – angiotensine – aldostérone, avec une augmentation de la signalisation Ang II conduisant à des processus pro-inflammatoires et profibrotiques dans le rein<sup>14,175</sup>. Le rôle critique de l'ACE2 dans l'infection rénale par le SRAS-CoV-2 est souligné par les résultats d'études *in vitro* qui montrent que le SRAS-CoV-2 peut infecter les organoïdes du rein humain et que cette infection peut être limitée par l'expression de l'ACE2 soluble recombinant humain - un récepteur leurre en cours d'évaluation dans le cadre d'essais cliniques<sup>123</sup>. De plus, à l'appui de l'hypothèse selon laquelle le SRAS-CoV-2 exerce un tropisme dans le rein, l'examen au microscope électronique d'échantillons d'autopsie de 26 patients décédés avec COVID-19 a démontré des amas de particules virales dans l'épithélium tubulaire et les podocytes. [Larsen. C, 2020]

Une autre étude qui impliquait une microdissection de tissu autoptique a démontré une charge virale SARS-CoV-2 détectable chez trois des six patients décédés ; dans les trois échantillons positifs pour le SRAS-CoV-2, le virus a été détecté dans tous les compartiments rénaux examinés, avec un ciblage préférentiel des cellules glomérulaires. Cependant, la plupart des analyses du tissu rénal ont utilisé la microscopie électronique, la microscopie optique et/ou l'immunohistochimie, qui ne peuvent pas déterminer de manière concluante si les particules identifiées sont réellement le SARS-CoV-2 ou simplement des structures de type viral. En particulier, l'identification de particules virales dans le tissu rénal par microscopie électronique a

été remise en question étant donné la ressemblance de ces particules avec d'autres structures cellulaires, telles que les vésicules recouvertes de clathrine, et le fait que plusieurs études n'ont pas pu confirmer la présence de virus dans le rein. De plus, les études du SARS-CoV-2 dans les fluides biologiques ont rarement montré une excrétion virale dans l'urine, ce qui soulève la possibilité qu'un réservoir de SARS-CoV-2 n'existe pas dans le rein. Malgré ces divergences et inconnues, une étude élégante a documenté l'enrichissement de l'ARN du SRAS-CoV-2 dans les cellules épithéliales et endothéliales glomérulaires par hybridation in situ<sup>176</sup>. Conformément à cette découverte, une répllication virale active dans différents tissus, y compris le tissu rénal, a été trouvée dans un sous-ensemble de patients atteints de COVID-19, bien qu'il ne soit toujours pas clair si cette répllication active rénale contribue à la charge virale. dans le corps [Larsen. C, 2020].

Comme décrit précédemment, les réponses immunitaires innées et adaptatives contribuent à la pathogenèse de COVID-19. Certaines preuves suggèrent que la modulation de ces processus chez les patients atteints d'insuffisance rénale peut avoir des conséquences sur les résultats de la maladie. Malgré quelques premières petites études non contrôlées montrant le contraire<sup>183,184</sup>, un rapport du Japon a montré une mortalité nettement plus élevée chez les patients sous hémodialyse d'entretien avec COVID-19 que chez les patients non dialysés, bien que cette différence de mortalité puisse être due à l'âge plus avancé de patients sous hémodialyse. Deux autres études menées au Royaume-Uni et en Italie ont également signalé une mortalité significativement plus élevée chez les patients sous hémodialyse que parmi la population non dialysée. Dans l'étude italienne, le SDRA était probablement la cause de décès chez les patients hospitalisés en hémodialyse [Luca. P,2020].

# **Chapitre IV**

**Diagnostic et thérapeutique**

**De SARS COV-2**

## **IV.1 Diagnostic COVID – 19**

Le diagnostic précoce et l'isolement des patients suspects jouent un rôle essentiel dans le contrôle de cette épidémie [Shang J,2020]. La spécificité et la sensibilité des différentes techniques de diagnostic diffèrent selon les populations et les types d'équipements utilisés [Leefflang MM,2013]. Plusieurs procédures ont été recommandées pour le diagnostic de COVID - 19 :

### **IV1.1.Présentation clinique :**

Les symptômes du COVID - 19 sont observés environ 5 jours après l'incubation [Shereen MA,2020]. Le temps médian d'apparition des symptômes de l'incubation de COVID - 19 est de 5,1 jours, et les personnes infectées présentent des symptômes pendant 11,5 jours. Il a été démontré que cette durée a un lien étroit avec le système immunitaire et l'âge du patient.

Les symptômes gastro-intestinaux comprennent la diarrhée, les vomissements et l'anorexie, enregistrés chez près de 40 % des patients. 48, 49 Jusqu'à 10 % des patients présentant des symptômes gastro-intestinaux ne présentent aucun signe de fièvre ou d'infection des voies respiratoires [Wang D,2020].

COVID - 19 a également été lié à une maladie hypercoagulable, augmentant le risque de thrombose veineuse. Il existe également des enregistrements de symptômes neurologiques (tels que fatigue, vertiges et troubles de la conscience), d'accidents vasculaires cérébraux ischémiques et hémorragiques et de lésions musculaires [Mao L,2020].

De nombreux symptômes extrapulmonaires, y compris des manifestations cutanées et oculaires. Des chercheurs italiens ont identifié des manifestations cutanées chez 20% des patients.

Les perspectives cliniques pour les enfants peuvent progressivement s'aggraver en raison d'une insuffisance respiratoire, qui n'a pas pu être corrigée en 1 à 3 jours par l'oxygène traditionnel (c'est-à-dire la sonde nasale) dans les cas graves ; les caractéristiques sont le choc septique, la septicémie, les saignements extrêmes et continus résultant d'anomalies de la coagulation et l'acidose métabolique

Le choc septique pourrait causer des dommages graves et étranges à plusieurs organes, en plus d'une infection pulmonaire grave.

Lorsqu'un dysfonctionnement du système extrapulmonaire survient, y compris les systèmes circulatoire et digestif, un choc septique est probable et le taux de mortalité augmente considérablement[Chen Z-M,2020].

L'accouchement prématuré et l'hypoxie intra-utérine surviennent lorsque le fœtus est privé d'un environnement adéquat d'oxygène. Les symptômes insidieux nécessitent des soins spécifiques chez certains nouveau-nés et prématurés.

Les rapports ont indiqué un bon pronostic pour les enfants dans les 1 ou 2 semaines. Les enfants sont sujets à une réponse hyperinflammatoire au COVID - 19 similaire à la maladie de Kawasaki, qui répond bien à la prise en charge, pour laquelle un nouveau terme est en train d'être inventé[Loomba RS,2020].

En outre, de nombreuses recherches ont révélé la répartition par âge des patients adolescents âgés de 25 à 89 ans. De nombreux patients âgés avaient entre 35 et 55 ans, et moins de cas parmi les nouveau-nés et les nourrissons ont été trouvés. Une analyse de la dynamique initiale de transmission du virus a montré que l'âge médian des patients était de 59 ans, variant de 15 à 89 ans ; la plupart (59 %) étaient des hommes

#### **IV.1.2.-Tests de dépistage non spécifiques du COVID - 19 chez les patients exposés :**

Les résultats de la plupart des tests sanguins sont généralement non spécifiques, mais pourraient aider à déterminer les causes de la maladie. Une formule sanguine complète montre généralement un nombre normal ou faible de globules blancs et de lymphopénie[Malihe Mohamadian,2020].

La protéine C-réactive (CRP) et le taux de sédimentation des érythrocytes étaient généralement augmentés, ce qui de manière optimale serait réévalué aux jours 3, 5 et 7 après l'admission. 1, 57, 58 Les niveaux de créatine kinase plus myoglobine, aspartate aminotransférase et alanine aminotransférase, lactate déshydrogénase, D-dimère et créatine phosphokinase pourraient être augmentés dans les formes graves de la maladie COVID-19. Lors de co-infections virales-bactériennes, les taux de procalcitonine peuvent être élevés. 59, 60 Dans une revue systématique et une étude de méta-analyse[Pormohammad et al,2020]. Ont étudié les résultats de laboratoire accessibles obtenus chez 2361 patients atteints du SRAS - CoV2, les résultats démontrant 26% de leucopénie, 13,3% de leucocytose et 62,5% de lymphopénie. De plus, parmi 2 200 patients, 91 % et 81 % ont révélé une élévation des plaquettes

(thrombocytose) et de la CRP, respectivement. De plus, un examen des études de cas a identifié un diagnostic clinique et une modification des paramètres cliniques chez un homme de 47 ans diagnostiqué avec la maladie de Wuweian.

Pour étudier l'effet du coronavirus pendant la phase aiguë de la maladie, les cytokines/chimiokines plasmatiques facteur de nécrose tumorale (TNF)  $-\alpha$  et interleukine (IL)  $-1\beta$ , IL1RA, IL2, IL4, IL5, IL - 6, IL - 10 , IL13, IL15 et IL17A ont été mesurés[Hossein-Khannazer N,2020]. Une étude a montré que les macrophages et les cellules dendritiques jouent un rôle crucial dans un système immunitaire adaptatif.

Ces cellules contiennent des cytokines et des chimiokines inflammatoires, telles que l'IL-6, l'IL-8, l'IL-12, le TNF- $\alpha$ , la protéine chimiotactique des monocytes-1, le facteur de stimulation des colonies de granulocytes-macrophages et le facteur de stimulation des colonies de granulocytes. Ces réactions inflammatoires pourraient provoquer une inflammation systémique.

Par conséquent, des tests fécaux et urinaires ont été recommandés pour les patients et le personnel de santé afin de détecter une éventuelle transmission alternative. Par conséquent, l'avancement d'outils pour déterminer les différents modes de transmission, y compris les échantillons de selles et d'urine, est urgent pour développer des stratégies d'inhibition et de minimisation de la transmission, ainsi que pour développer des thérapies pour contrôler la maladie[Malihe Mohamadian,2020].

#### **IV.1.3.Résultats radiologiques:**

L'examen aux rayons X du thorax peut afficher diverses caractéristiques ou modèles d'imagerie chez les patients COVID-19 avec une gravité et une durée différentes de la maladie.

Les résultats d'imagerie diffèrent en fonction de l'âge du patient, du stade de la maladie lors du dépistage, de la compétence immunitaire et des protocoles de traitement médicamenteux[Fang Y,2020].

D'autre part, l'imagerie par tomodensitométrie (TDM) est essentielle pour surveiller la progression de la maladie et évaluer l'efficacité thérapeutique. Il peut être réexaminé 1 à 2 jours après l'admission, sur la base du Règlement sur les protocoles de diagnostic et de traitement (RTPR)[Vyakaranam AR,2019].

La marque cardinale de COVID - 19 était des opacités multiples, bilatérales, postérieures et périphériques en verre dépoli avec ou sans consolidation pulmonaire

et, dans les cas graves, des ombres infiltrantes. 68 L'analyse d'autopsie d'un patient COVID - 19 a montré une accumulation de liquide et une formation de membrane hyaline dans les parois alvéolaires, qui peuvent être le principal moteur pathologique de l'opacité du verre dépoli[Pan F,2020].

Cependant, d'autres études ont indiqué que de petites ombres inégales, des modifications pleurales, une ligne curviligne sous-pleurale et des signes de halo inversé sont généralement observés chez les patients COVID - 19. Les lignes intralobulaires et les cloisons interlobulaires épaissies ont été montrées dans un motif de pavage fou sur le sol - fond d'opacité du verre.

En outre, plusieurs lésions lobaires peuvent être trouvées dans le système respiratoire chez les enfants atteints d'une infection grave. Les preuves ont montré que les manifestations de la tomodensitométrie thoracique (œdème pulmonaire) signalées pour COVID - 19 sont généralement proches du SRAS et du MERS

Des données probantes ont indiqué qu'un scanner thoracique initial a un taux de détection plus élevé (environ 98 %) par rapport à la réaction en chaîne par transcriptase inverse-polymérase (RT-PCR) (environ 70 %) chez les patients infectés [Xie et al,2020]. Ont démontré qu'environ 3 % des patients n'avaient pas de RT - PCR primaire positive mais avaient un scanner thoracique positif ; par conséquent, les deux tests sont recommandés pour les patients COVID - 19.

TDM du thorax comprenant une méthode urgente et simple de détection de l'infection initiale au COVID - 19 avec une sensibilité élevée pour un diagnostic rapide et une surveillance de la progression de la maladie chez les patients. Une attention particulière devrait être accordée au rôle des radiologues dans la découverte de nouvelles maladies infectieuses.

#### **IV.1.4.Diagnostic moléculaire :**

Le diagnostic clinique de COVID - 19 se concentre principalement sur les données épidémiologiques, les symptômes cliniques et certaines technologies adjuvantes, telles que la détection des acides nucléiques et les tests immunologiques. De plus, l'isolement du SRAS - CoV - 2 nécessite des équipements à haut débit (niveau de biosécurité - 3) pour assurer la sécurité du personnel.

De plus, les tests sérologiques n'ont pas encore été validés. Dans le domaine du diagnostic moléculaire, il y a trois problèmes principaux : (i) diminuer le nombre de faux négatifs en détectant des quantités minimales d'ARN viral ; (ii) éviter le nombre

de faux positifs grâce à la différenciation correcte des signaux positifs entre les différents agents pathogènes ; et (iii) une grande capacité de test rapide et précis d'un grand nombre d'échantillons en peu de temps[Malihe Mohamadian,2020].

#### **IV.1.4.1-Détection d'acide nucléique :**

Deux technologies largement utilisées pour la détection des acides nucléiques du SRAS - CoV - 2 sont la RT - PCR en temps réel (rRT - PCR) et le séquençage à haut débit. Néanmoins, en raison de la dépendance à l'égard de l'équipement et des coûts élevés, le séquençage à haut débit dans le diagnostic clinique est limité.

L'accès à l'ensemble de la structure du génome du SRAS - CoV - 2 a aidé à la conception d'amorces spécifiques et a introduit les meilleurs protocoles de diagnostic. Dans les premiers rapports publiés sur l'application de la rRT - PCR dans le diagnostic de COVID - 19, le ciblage de la région du gène de pointe (S) du SRAS - COV - 2 a montré une spécificité remarquable et une sensibilité limitée[Corman VM,2020].

Plus tard, la sensibilité de cette technique a été grandement améliorée par l'utilisation de sondes spécifiques pour les autres gènes spécifiques du virus, y compris l'ARN polymérase dépendante de l'ARN (RdRp) dans la région ORF1ab, la nucléocapside (N) et l'enveloppe (E).

Pour une réaction croisée avec d'autres coronavirus humains et prévenir la dérive génétique potentielle du SRAS - CoV - 2, deux cibles moléculaires doivent être impliquées dans ce test : une cible non spécifique pour détecter d'autres CoV, et une cible spécifique pour le SRAS - CoV - 2 . La comparaison des résultats obtenus en ciblant tous les gènes étudiés a montré que le gène RdRp est la cible la plus appropriée avec la sensibilité la plus élevée. Les tests RdRp ont été validés dans environ 30 laboratoires européens utilisant la technologie des acides nucléiques synthétiques

Actuellement,[Chan et al,2020]. Ont proposé un nouveau test RT - PCR ciblant une séquence de RdRp / Hel qui pourrait détecter une faible charge de SRAS - CoV - 2 dans les échantillons des voies respiratoires supérieures, du plasma et de la salive sans aucune réactivité croisée avec d'autres virus respiratoires courants. Bien que les tests recommandés par le CDC aux États-Unis reposent sur deux protéines de nucléocapside N1 et N2, l'OMS recommande le test du gène E comme dépistage de première intention, suivi du test du gène RdRp comme test de confirmation. Sur la

base des preuves les plus récentes, le panel QIAstat - Dx SARS - CoV - 2, un système de PCR multiplex RT - en temps réel ciblant les gènes RdRp et E, reste très sensible malgré les variations nucléotidiques affectant l'hybridation du test PCR.

Généralement, la RT - PCR quantitative (RT - PCR) a une spécificité élevée en tant que test de référence pour le diagnostic final de COVID - 19. Cependant, sa sensibilité pourrait être variable en fonction de la charge virale, de la technique d'extraction d'ARN, de la source d'échantillonnage et du stade de la maladie au cours de le moment de l'échantillonnage. En effet, les résultats faussement positifs de la RT - PCR sont liés à la contamination croisée des échantillons et aux erreurs de manipulation[Malihe Mohamadian,2020].

En revanche, des inexactitudes à n'importe quelle étape de la collecte, du stockage et du traitement des échantillons peuvent conduire à des résultats faussement négatifs. Certaines études ont révélé que les échantillons des voies respiratoires supérieures (bas des narines et de l'oropharynx) sont plus souhaitables pour le test RT - PCR en raison de nombreuses copies virales. En outre, d'autres lacunes des tests RT-qPCR incluent les risques de sécurité biologique résultant de la maintenance et du travail sur des échantillons de patients, ainsi que le processus de détection d'acide nucléique long et lourd.

Pour améliorer les techniques de diagnostic moléculaire du COVID - 19, des méthodes basées sur l'amplification isotherme sont actuellement en cours de développement. L'amplification isotherme à médiation par boucle (LAMP) utilise l'ADN polymérase et 4 à 6 amorces différentes se liant aux séquences distinctes du génome cible [Bernheim A,2020]. Dans les réactions LAMP, l'ADN amplifié est indiqué par la turbidité résultant d'un sous-produit de la réaction, une couleur détectable générée par un colorant sensible au pH, ou la fluorescence produite par un colorant fluorescent. L'approche se produit à une seule température, en moins d'une heure, et avec des signaux de fond minimales. Le test de diagnostic LAMP pour COVID - 19 est plus spécifique et sensible par rapport aux tests conventionnels RT - PCR et ne dépend pas d'équipements de laboratoire spécialisés tels qu'un thermocycleur. Cependant, en raison de la multiplicité des amorces utilisées dans cette méthode, l'optimisation des conditions de réaction présente un défi majeur

#### **IV.1.4.2. Technique basée sur les puces à ADN :**

La puce à ADN est une méthode rapide et à haut débit pour le test COVID - 19. En guise de bref résumé du protocole, l'ARN du coronavirus produira d'abord de l'ADNc marqué avec des sondes spécifiques via une transcription inverse.

L'ADN complémentaire est produit par des matrices d'ARN de coronavirus, puis par un marquage de transcription inverse avec des sondes particulières. Les cibles marquées sont hybridées à la puce à sonde. Les ADN libres sont éliminés en lavant la solution.

Enfin, des sondes particulières identifient l'ARN COVID - 19[Shi et al,2020]. Ont réussi à détecter le SRAS - CoV dans des échantillons de patients. Dans leur étude, [Xu et al,2020] Ont étudié un large éventail de polymorphismes de gènes de pointe avec une grande précision. En outre, d'autres études ont conçu des méthodes de fluorescence et de non-fluorescence pour détecter l'ensemble du genre coronavirus avec une efficacité prometteuse[Jiang et al,2020]. Ont construit un microréseau de protéome du SRAS - CoV - 2 composé de 18 des 28 protéines attendues et l'ont administré à 29 cas de convalescence pour caractériser les réactions d'immunoglobuline (Ig) G et IgM dans les sérums. Il a été révélé que tous ces patients avaient des anticorps IgM et IgG, qui reconnaissent et lient les protéines du SRAS - CoV - 2, en particulier les protéines S1 et N. En plus de ces protéines, des réponses anticorps importantes contre NSP5 et ORF9b sont également reconnues. Le signal IgG spécifique de S1 - est fortement lié à l'âge et aux taux de lactate déshydrogénase et négativement au rapport lymphocytaire[Shen et al,2020]. Fait progresser l'expérience RT - LAMP pour montrer des signaux à l'aide d'une sonde d'extinction avec la même efficacité que le test RT - PCR standard en ce qui concerne l'identification MERS - CoV.

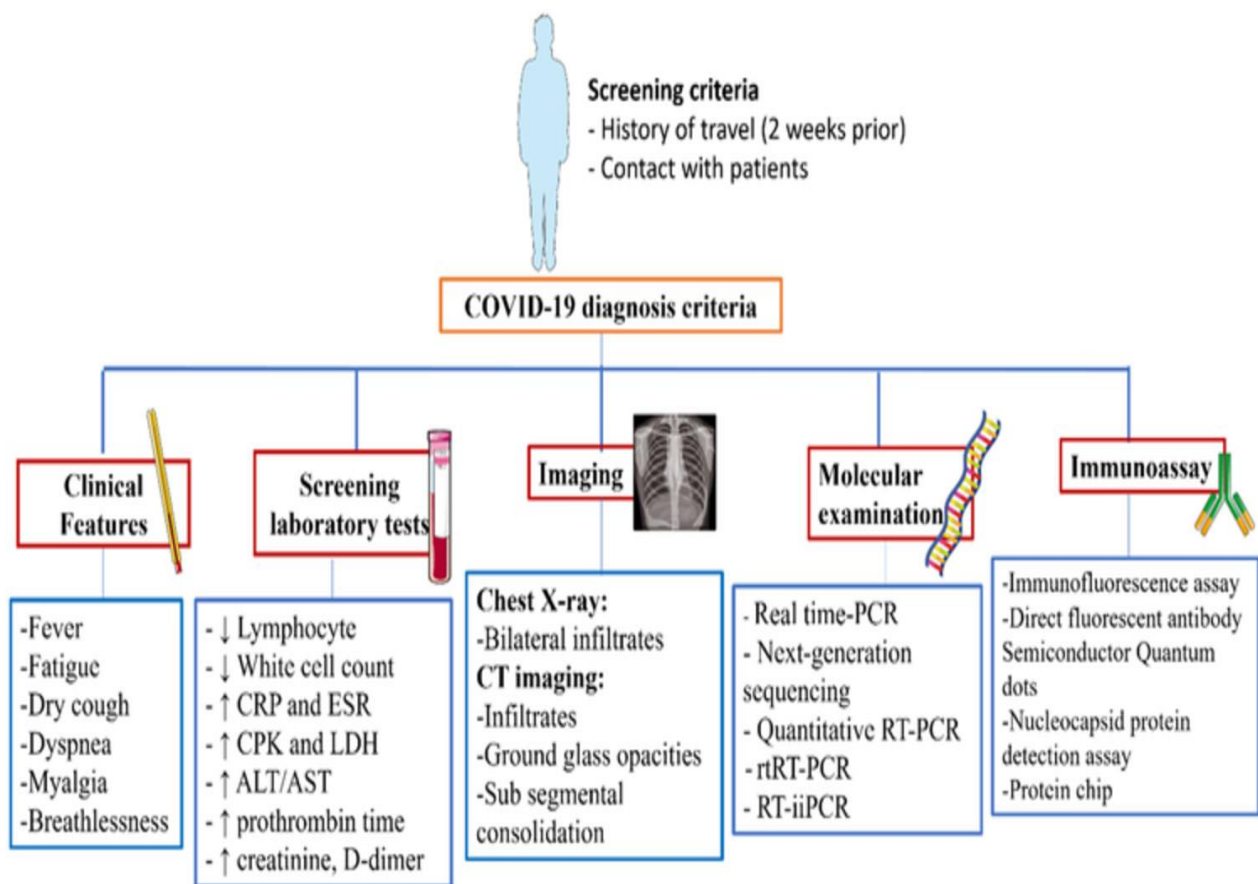
#### **IV.1.5. Diagnostic immunologique :**

La détection des antigènes et les techniques immunologiques peuvent être utilisées pour un diagnostic rapide et rentable tout en offrant une alternative aux méthodes moléculaires.

Les techniques immunologiques comprenant le test d'immunofluorescence, le test d'anticorps à fluorescence directe, le test de détection de protéine de nucléocapside, la puce à protéine, les points quantiques semi-conducteurs et le test de

microneutralisation définissent une liaison entre un antigène viral et un anticorps spécifique[Chan K-H,2013].

Ces méthodes immunologiques sont simples à utiliser mais ont une faible spécificité/sensibilité. Dans le cas du COVID - 19, la morphologie du virus peut être observée par microscopie électronique selon les postulats traditionnels de Koch. Les tests sérologiques peuvent améliorer la détection des coronavirus, de sorte que les antigènes et les anticorps monoclonaux associés peuvent représenter une nouvelle approche diagnostique pour le développement futur (Figure 2).



**Figure01:** Diagnostic protocol recommended for COVID-19

les tests érologiques peuvent être spécifiques à un type d'immunoglobuline, ils peuvent mesurer simultanément les anticorps IgM et IgG, ou ils peuvent être des examens absolus des anticorps, qui mesurent souvent les anticorps IgA. 96 Sur la base de la procédure et du dispositif spécifiques, ces expériences seront généralement effectuées dans les 1 à 2 heures suivant l'arrivée d'un échantillon au laboratoire et son chargement sur la plate-forme appropriée[Guo et al,2020]. Ont indiqué que les

anticorps IgA et IgM ont des taux positifs de 93,0% et 85,5% après 3 à 6 jours, respectivement.

De plus, 78,0 % des anticorps IgG positifs ont été détectés pendant 10 à 18 jours. L'efficacité de détection par un dosage immuno-enzymatique IgM (ELISA) est supérieure à celle de la qPCR après 5,5 jours d'apparition des symptômes.

Après 5 jours, la détection des IgM ELISA est plus efficace qu'une qPCR.

De plus, la combinaison de la PCR et de l'ELISA IgM a augmenté le taux de détection de 98,5% [Xiang et al,2020] Ont testé 63 patients infectés par le SRAS - CoV - 2 admis à l'hôpital Jinyintan de Wuhan, Hubei, Chine. Les échantillons de sérum des patients ont été évalués à l'aide d'une capture ELISA et ELISA indirecte d'IgG. Les résultats de l'étude indiquent que l'IgM était positif avec une précision de 64,3 %, une sensibilité de 44,4 % et une spécificité de 100 % dans 28 des 63 échantillons. L'identification de l'échantillon de 52 cas a également montré un test IgG positif avec une sensibilité de 82,54 %, une spécificité de 100 % et une précision de 88,8 %. De plus, une sensibilité de 87,3 % a été obtenue en utilisant l'analyse combinée IgM et IgG.

Ont évalué les anti-IgM et anti-IgG produits contre la protéine de pointe recombinante et la protéine de nucléocapside du SRAS-CoV-2 chez 397 patients COVID-19 confirmés par PCR et 128 cas négatifs sur huit sites cliniques distincts.

La sensibilité et la spécificité moyennes de l'examen étaient respectivement de 88,5% et 90,5%. Les résultats ont montré une cohérence de détection considérable parmi les différents types d'échantillons de sang veineux et prélevés au doigt.

Par rapport à un seul test IgM ou IgG, l'analyse combinée IgM - IgG a une efficacité et une sensibilité plus élevées. Par conséquent, il est important et urgent d'améliorer plusieurs approches supplémentaires sensibles et spécifiques pour le diagnostic de COVID-19.

#### **IV.1.6.Nouvelles techniques**

##### **IV.1.6.1.CRISPR technique**

La détection des acides nucléiques avec CRISPR - Cas13a / C2c2 est une plateforme de détection moléculaire très rapide, sensible et spécifique, qui peut aider à l'épidémiologie, au diagnostic et au contrôle de la maladie.

De plus, Cas13a/C2c2 peut détecter l'expression de transcrits dans des cellules vivantes et différentes maladies[Zhang et al,2020]. A présenté un protocole pour la

détection de COVID - 19 en utilisant la technique SHERLOCK (Specific High Sensitivity Enzymatic Reporter UnLOCKing) basée sur les diagnostics CRISPR.

Des fragments d'ARN du virus SRAS - CoV - 2 permettent de détecter des séquences cibles d'environ 100 copies. L'expérience est réalisée par amplification isotherme de l'acide nucléique extrait d'échantillons de patients puis amplification de la séquence d'ARN viral via Cas13 et est enfin lue par une bandelette réactive papier en moins d'1 heure [Huang et al,2020]. Ont établi un test basé sur CRISPR par un complexe CRISPR Cas12a/ARNg personnalisé.

Ils ont utilisé une sonde fluorescente pour identifier les amplicons cibles produits par RT - PCR standard ou amplification isotherme par recombinaise polymérase. Cette méthode a montré une détection spécifique à des endroits non équipés des systèmes PCR nécessaires aux tests de diagnostic qPCR en temps réel. L'analyse permet l'identification de SARS - CoV - 2 échantillons positifs avec un temps de réponse test - à - d'environ 50 minutes et une limite de détection de deux copies de chaque échantillon à détecter.

Les résultats du test CRISPR sur des échantillons nasaux prélevés sur des personnes atteintes de COVID - 19 étaient comparables aux données appariées obtenues à partir du test RT approuvé par le CDC - Qpcr[Broughton et autres,2019].

Description Développement d'un test de flux latéral rapide (<40 min), simple et précis basé sur CRISPR-Cas12 pour le diagnostic de l'extrait d'ARN du SRAS-CoV 2 à partir d'un écouvillon nasal. À l'aide d'échantillons de référence synthétiques et d'échantillons cliniques de patients, comprenant des patients diagnostiqués avec COVID-19 et 42 patients atteints d'autres maladies respiratoires, ils ont confirmé le processus.

Cette approche basée sur CRISPR offre une alternative visuelle plus rapide à la méthode de RT-PCR en temps réel du SRAS CoV 2 utilisée par les Centers for Disease Control and Prevention des États-Unis, avec un accord prédictif négatif de près de 100 % et un accord prédictif positif de 95 %.

#### **IV.1.6.2. Technique basée sur la LAMPE**

Loop-mediated isothermal amplification L'amplification isotherme à médiation par boucle (LAMP) est une nouvelle méthode d'amplification isotherme des acides nucléiques avec une grande efficacité. Ceci est utilisé pour amplifier des ARN et des ADN avec une spécificité et une sensibilité élevées en raison de sa fonction

d'amplification exponentielle et de six séquences cibles particulières diagnostiquées par quatre amorces distinctes [Enosawa M, 2003]. Le test LAMP est rapide et ne nécessite pas de réactifs ou d'équipements coûteux. En outre, la méthode d'électrophorèse sur gel est largement utilisée pour l'étude des éléments amplifiés afin de détecter les points finaux. Par conséquent, le test LAMP contribuera à réduire le coût de la détection des coronavirus. Plusieurs stratégies de détection du coronavirus basées sur LAMP sont définies ici, telles que développées et réalisées dans le diagnostic clinique.

Ont rapporté un test LAMP simple dans l'étude sur le SRAS et ont démontré la faisabilité de cette méthode pour la détection du SRAS - CoV. Le site SARS - CoV ORF1b a été sélectionné pour la détection du SRAS et amplifié en présence de six amorces via la réaction LAMP, puis les produits amplifiés ont été évalués par électrophorèse sur gel. Les niveaux de sensibilité et de détection du test LAMP pour le SRAS sont proches de ceux des techniques traditionnelles basées sur la PCR [Pyrc et al, 2020]. A appliqué efficacement LAMP à la détection du HCoV - NL63 avec une sensibilité et une spécificité souhaitables dans les cultures cellulaires mobiles et les échantillons cliniques.

En particulier, une réplique de la matrice d'ARN s'est avérée responsable de la restriction de détection. L'amplification est observée sous forme de colorant fluorescent ou de précipitation de pyrophosphate de magnésium. Ces techniques peuvent être exécutées en temps réel en surveillant la turbidité du pyrophosphate ou de la fluorescence, qui surmontent correctement la restriction de la détection du point final.

A développé un test RT - LAMP bénéfique pour le diagnostic et la surveillance épidémiologique du MERSCoV humain. Cette méthode était hautement spécifique, sans aucune réaction croisée avec d'autres virus respiratoires spécifiques, et détectait aussi peu que copies d'ARN. Par la suite, ils ont développé le test RT - LAMP en révélant un signe à l'aide d'une sonde d'extinction (QProbe), qui a la même efficacité que le test RT - PCR en temps réel habituel en ce qui concerne la détection du MERSCoV.

Sur la base d'autres preuves, une méthode de visualisation des acides nucléiques a été développée qui combine RT - LAMP et une bande de visualisation à flux vertical pour la détection du MERS

#### **IV.1.6.3. Technologie Penn RAMP**

Sur la base de l'efficacité rapportée par [ **Zhang et al**]. en utilisant la LAMP comparativement moins sensible, la sensibilité améliorée de la technique Penn-RAMP obtenue qui est attribuable à un protocole LAMP mis à jour en deux étapes, peut s'avérer considérablement efficace en tant que diagnostic.

Pour amplifier des cibles spécifiques par amplification par recombinaison polymérase, dans laquelle toutes les cibles sont amplifiées simultanément, le Penn-RAMP nécessite une réaction préliminaire avec des amorces LAMP externes. Une prochaine réaction LAMP très précise est alors déclenchée. En particulier, la première étape utilise des amorces LAMP externes F3 et B3, tandis que les quatre autres amorces RAMP sont encore mélangées dans l'étape 2.

Par rapport à la LAMP normale, ce concept « niché » a considérablement amélioré la sensibilité de LAMP d'environ 10 à 100 fois, en particulier lorsque vous travaillez avec des échantillons distillés et bruts. De plus, lorsqu'elle a été étendue à des essais simulés, la méthodologie Penn - RAMP a reçu une note d'approbation de 100 % à 7-10 copies d'ARN viral par réaction, par rapport à une note d'approbation de 100 % pour les 700 copies d'ARN viral nécessaires pour l'analyse PCR.

#### **IV.1.6.4. PCR numérique de gouttelettes :**

Pour l'identification directe et la quantification des cibles d'ADN et d'ARN, la PCR numérique en gouttelettes (ddPCR) comprenait une technique extrêmement sensible. Il a été largement utilisé pour les maladies infectieuses, notamment en raison de sa capacité à identifier quelques copies de génomes viraux avec précision et efficacité. Si l'identification de l'existence virale de bas niveau et/ou résiduelle est appropriée, les données quantitatives ddPCR sont beaucoup plus pertinentes que celles fournies par les tests RT-PCR réguliers. Compte tenu de la nécessité de restreindre (dans la mesure du possible) les résultats faussement négatifs dans le diagnostic de COVID - 19, l'utilisation de la ddPCR peut apporter un soutien vital. Malgré cela, le test ddPCR est encore très rarement étudié en milieu clinique et il n'y a actuellement aucune preuve disponible pour les cas européens.

#### **IV.1.6.5. Technique basée sur le séquençage de nouvelle génération (NGS) :**

Les virus à ARN se présentent sous une grande variété de variétés et sont les spécialistes étiologiques de nombreuses maladies infectieuses humaines et animales importantes.

Les virus à ARN comprenaient la principale variété et sont les agents étiologiques de maladies très infectieuses chez l'homme et l'animal telles que le SRAS, l'hépatite, la grippe et l'IB (bronchite infectieuse aviaire). La technologie NGS à haut débit joue un rôle essentiel dans le diagnostic primaire et précis[Orilio AF,2013].

De plus, la méthode NGS peut détecter si différents types de virus incluaient ou non un agent pathogène. La nouvelle technique rapide des virus par NGS, y compris le séquençage de l'ADN et de l'ARN, a permis de développer l'identification de la diversité virale. L'identification d'une vaste gamme d'agents pathogènes à l'aide des technologies NGS est également importante pour contrôler l'infection virale causée par un nouvel agent pathogène. Ces dernières années, l'avancement de la méthode NGS via le séquençage de l'ARN nous a permis de faire de grands progrès dans la reconnaissance rapide de nouveaux virus à ARN.

Le séquençage de l'ARN détecte simultanément des millions de fragments d'ADN transcrits de manière inverse à partir d'échantillons d'ARN complexes en utilisant des amorces aléatoires[Chen et al,2020]. Ont signalé un nouveau coronavirus de canard utilisant la méthode de séquençage de l'ARN, qui différait de celle du poulet IBV (virus de la bronchite infectieuse). Le nouveau CoV spécifique au canard était une nouvelle espèce possible au sein du genre Gamma - coronavirus, comme le montrent les séquences du gène viral 1b de trois régions.

En conclusion, l'épidémie d'un nouveau virus est apparue fin décembre 2019. Le COVID - 19 s'est immédiatement propagé et a remis en cause la médecine, l'économie et la santé publique dans le monde entier. De nombreuses preuves ont proposé que les récepteurs ACE2 comprenaient des protéines structurales cruciales pour le bourgeonnement du virus et l'entrée dans les cellules hôtes.

La transmission d'hôtes intermédiaires non identifiés à des espèces croisées et la transmission d'humain à humain ont été reconnues. Par conséquent, la détection précoce et l'isolement des patients suspects peuvent jouer un rôle essentiel dans le contrôle de cette épidémie. Actuellement, les méthodes de diagnostic du COVID - 19 sont nombreuses ; il est donc impératif de choisir un protocole de détection adapté. Chacune des techniques décrites a ses inconvénients et avantages spécifiques. L'imagerie CT du thorax et les tests RT - PCR sont recommandés pour les patients COVID - 19. Cependant, l'utilisation de la PCR nécessite divers équipements et un laboratoire bien rodé. LAMP peut être détecté avec de faibles nombres de matrices

d'ADN ou d'ARN en 1 heure. La puce à ADN est une méthode coûteuse pour le diagnostic de COVID - 19, et d'autres méthodes nouvellement développées nécessitent également des investigations supplémentaires pour parvenir à un développement et une détection rapides à l'avenir. Étant donné que le nombre de cas infectés augmente rapidement, de futures études devraient révéler les secrets des voies moléculaires du virus en ce qui concerne le développement de vaccins ciblés et de traitements antiviraux.

## **THERAPEUTIQUE**

### **IV.2. Thérapeutique COVID-19**

En l'absence de médicaments efficaces ayant une efficacité thérapeutique établie contre le SRAS-CoV-2, une attention considérable a été portée à la réutilisation des médicaments actuellement disponibles pour le traitement du COVID-19 [**Maisonnasse P,2020**]. Les essais de médicaments ont inclus des agents antiviraux qui étaient principalement homologués ou développés pour le traitement de la grippe (oseltamivir, favipavir, umifénovir), du VIH (lopinavir, ritonavir, azvudine), de l'hépatite (ribavirine, sofosbuvir), d'Ebola (remdesivir), des infections herpétiques (penciclovir), antibiotiques aux propriétés antivirales ou anti-inflammatoires (azithromycine), antiparasitaires y compris antipaludiques (chloroquine, hydroxychloroquine, ivermectine) et antagoniste de l'histamine-2 à activité antivirale (famotidine). Pour contourner ou augmenter les réponses immunitaires, plusieurs médicaments immunomodulateurs ont également été étudiés dont les stéroïdes, les interférons, la colchicine, les inhibiteurs de l'interleukine (tocilizumab, anakinra) ainsi que le plasma de convalescence.

Les thérapies spécifiquement dirigées contre le SRAS-CoV-2 comprennent l'utilisation de protéines ACE2 humaines recombinantes pour agir en tant que récepteurs leurres pour le virus SARS-CoV-2. Une indication de l'effort de recherche mondial est reflétée par un rapport indiquant qu'il y a plus de 2000 essais cliniques évaluant plus de 700 thérapies potentielles contre le COVID-19 [**Maisonnasse P,2020**]. Cependant, une revue systématique «vivante» dans le British Medical Journal a conclu que: «l'efficacité de la plupart des interventions est incertaine car la plupart des études contrôlées randomisées jusqu'à présent ont été de petite taille et ont d'importantes limites d'étude». - souligné par [**Tikkinen et al,2020**] déclarant : «plusieurs fausses déclarations d'efficacité ont émergé de comparaisons non randomisées.» En revanche, les rapports du groupe de collaboration RECOVERY (Randomized Evaluation of Covid-19 Therapy) au Royaume-Uni ont établi

les avantages d'un traitement de 28 jours avec le stéroïde dexaméthasone pour réduire la mortalité et la progression vers l'assistance ventilatoire chez les patients atteints de COVID-19 ; en revanche, aucun bénéfice n'a été associé à l'association du lopinavir et du ritonavir [Freedberg DE,2020].

#### **IV.2.1.Chloroquine et Hydroxychloroquine**

L'intérêt initial pour la chloroquine et l'hydroxychloroquine était compréhensible car suite à l'épidémie de SRAS en 2002, les données d'études in vitro suggéraient que les médicaments possédaient à la fois des propriétés antivirales et anti-inflammatoires liées à leur capacité à élever le pH endosomal dans la cellule hôte et à inhiber glycosylation de l'hôte ACE2 (**figure 3A**) ; une conclusion étayée par des données de Chine [Gao J,2020], mais contestée par la suite par des données d'études in vivo sur des singes macaques (199). Des études in vitro suggèrent également que la chloroquine et l'hydroxychloroquine pourraient bloquer la liaison du virus à l'ACE2;

cependant, les concentrations requises seraient très probablement toxiques si elles étaient utilisées en clinique et augmenteraient le risque de dysrythmies cardiaques .

De plus, l'hydroxychloroquine et la chloroquine ont été largement utilisées pour traiter le paludisme, le lupus érythémateux et la polyarthrite rhumatoïde et, étant donné que leurs profils pharmacocinétiques et de toxicité étaient bien connus, leur utilisation était considérée comme sûre. Malheureusement, des insuffisances dans la conception de certains des essais cliniques portant sur ces médicaments dans le COVID-19 ont abouti à des résultats contradictoires. En mars 2020, le soutien clinique à l'efficacité d'une association d'hydroxychloroquine avec l'antibiotique macrolide, l'azithromycine, dans la réduction de la charge virale de SAR-CoV-2 chez les patients a été salué et largement diffusé à la suite d'une étude ouverte sans ECR de Marseille [Wang Y, 2020].

Cependant, le petit nombre de patients avec un processus de recrutement douteux et des résultats non congruents ont soulevé des questions quant à l'interprétation des données. Néanmoins, en l'absence d'options alternatives, de nombreuses autorités de santé mondiales ont approuvé l'utilisation de l'hydroxychloroquine, y compris la FDA, cette dernière approuvant l'autorisation d'utilisation précoce de l'hydroxychloroquine le 28 mars 2020.

Par la suite, des essais contrôlés randomisés au Canada et aux États-Unis n'ont pas réussi à démontrer un avantage prophylactique post-exposition. De même, une étude contrôlée

randomisée multicentrique ouverte au Brésil utilisant l'hydroxychloroquine et l'azithromycine n'a pas seulement réussi à démontrer des résultats bénéfiques, mais a également signalé des événements indésirables, notamment des toxicités cardiaque et hépatique.

Pour compliquer le débat scientifique, des études rétrospectives ont produit des résultats contradictoires, une ne rapportant aucun bénéfice de l'hydroxychloroquine ou de l'azithromycine seule, ou en association, pour réduire la mortalité hospitalière, et une autre rapportant une réduction de la mortalité avec l'hydroxychloroquine seule ou en association avec azithromycine par rapport à la norme de soins (SOC) [Bestle D,2020].

Dans cette dernière étude, un nombre beaucoup plus grand de patients traités par hydroxychloroquine et / ou azithromycine a également reçu des stéroïdes, une faiblesse confondante du talon d'Achille apparente dans les études rétrospectives. Fin avril 2020, la FDA a annulé son approbation et déconseillé avec prudence l'utilisation de l'hydroxychloroquine ou de la chloroquine en dehors des essais cliniques ou des hôpitaux.

Enfin, les données de l'essai RECOVERY, qui est de loin le plus grand ECR à ce jour, ont annulé l'efficacité de l'hydroxychloroquine concluant: «Chez les patients hospitalisés pour COVID-19, l'hydroxychloroquine n'était pas associée à une réduction de la mortalité à 28 jours, mais était associée à une augmentation durée du séjour à l'hôpital et risque accru d'évoluer vers une ventilation mécanique invasive ou de mourir ».

Malgré les preuves contre l'utilisation de la chloroquine ou de l'hydroxychloroquine, la controverse se poursuit. Une méta-analyse publiée le 30 septembre 2020 sur le serveur de pré-impression medRxiv a conclu que l'utilisation d'hydroxychloroquine en ambulatoire réduisait le séjour à l'hôpital et la mortalité bien qu'aucun des cinq ECR inclus dans l'analyse ne démontre individuellement un bénéfice [Baum A,2020].

#### **IV.2.2.Remdesivir**

Le médicament antiviral, le remdesivir, a été initialement développé pour l'hépatite C puis testé pour Ebola et le profil pharmacocinétique et de toxicité de ce médicament est également établi bien que limité aux données de petits ECR.

Le remdesivir est un promédicament qui pénètre facilement dans les cellules où il subit une hydrolyse en une forme active par des enzymes estérases en un analogue d'adénosine triphosphorylé qui inhibe de manière compétitive la polymérase dépendante de l'ARN, entraînant la terminaison de la chaîne de répliation (**figure 3A**).

Les résultats *in vitro* indiquent que l'activité antivirale du remdesivir est efficace à des concentrations micromolaires inférieures ou faibles contre d'autres coronavirus, y compris le MERS-CoV, et ensuite étayée par une étude *in vivo* chez des singes rhésus .

Un ECR multicentrique de Chine a comparé les résultats cliniques de 158 patients atteints de COVID-19 sévère traités pendant 28 jours par remdesivir contre 78 patients recevant du SOC.

Les auteurs ont conclu: «Dans cette étude portant sur des patients adultes hospitalisés pour un COVID-19 sévère, le remdesivir n'était pas associé à des bénéfices cliniques statistiquement significatifs.» L'étude n'a pas non plus démontré que le remdesivir réduisait la charge virale. Les résultats d'une étude internationale multicentrique (ACTT-1) de patients hospitalisés COVID-19 ont démontré que le remdesivir par rapport au placebo raccourcissait le temps de récupération de 5 jours, mais ne réduisait pas la mortalité, chez les adultes hospitalisés avec COVID-19 et preuve d'infection des voies respiratoires inférieures. À l'inverse, une étude comparant un traitement de 5 à 10 jours par remdesivir, mais sans bras SOC de comparaison correspondant, n'a trouvé aucune différence dans les résultats. Les résultats d'un troisième ECR ont démontré un bénéfice modeste pour un cours de gestion de 5 jours, mais curieusement pas pour un cours de 10 jours.

Enfin, les résultats intermédiaires de l'ECR international multicentrique OMS SOLIDARITY n'indiquent aucun bénéfice du remdesivir, de l'hydroxychloroquine, du lopinavir ou de l'interféron sur la mortalité globale ou l'instauration de la ventilation et la durée du séjour à l'hôpital. De manière critique, la charge virale en ARN n'a pas été mesurée pour refléter l'efficacité antivirale. En conclusion, les résultats des essais cliniques avec le remdesivir ne montrent pas qu'il réduit la charge virale ou les taux de mortalité, bien que certaines études, mais pas toutes, indiquent que les patients se sont rétablis plus rapidement[Russo V, et al,2020].

#### **IV.2.3.Dexaméthasone**

Bien que les glucocorticoïdes, tels que la dexaméthasone, suppriment la réponse immunitaire de l'hôte et augmentent le risque d'améliorer la réplication virale et d'activer potentiellement une infection latente telle que la tuberculose ou l'hépatite, ils peuvent être bénéfiques pour contrer une tempête de cytokines hyperinflammatoire comme on le voit dans le COVID-19 gravement malade patients (**figure 3B**). Les résultats d'un ECR du groupe d'étude RECOVERY basé au Royaume-Uni ont démontré une efficacité significative pour l'utilisation de la dexaméthasone.

#### **IV.2.4. Inhibiteurs Ivermectin, 3CLpro et PLpro**

Une autre cible médicamenteuse prometteuse est la cystéine protéase, Mpro (ou 3CLpro), une polyprotéine précurseur, qui est essentielle pour le clivage viral. Les avantages antiviraux potentiels de l'ivermectine, un médicament antiparasitaire à large spectre largement utilisé, suscitent un intérêt considérable en tant qu'inhibiteur putatif du complexe protéique importine  $\alpha / \beta 1$  empêchant l'importation nucléaire et inhibant également la protéase 3CLpro empêchant ainsi la réplication virale [Moulton H, et al, 2020].

Il existe une littérature abondante qui soutient l'innocuité de l'ivermectine pour traiter une variété d'infections parasitaires ainsi que des données in vitro et in vivo reflétant les effets anti-SRAS-CoV-2. Bien que n'ayant pas reçu l'approbation de la FDA pour la maladie COVID-19, l'ivermectine a été utilisée pour le traitement de la maladie dans un certain nombre de pays et notamment au Pérou où son utilisation a été liée à une diminution de la mortalité et de la létalité bien qu'en l'absence de données d'après les ECR appropriés, d'autres facteurs non liés au médicament peuvent avoir contribué aux avantages apparents du traitement [Baum A, 2020].

Un essai clinique exploratoire de phase 1 a été approuvé au Royaume-Uni par l'Agence de réglementation des médicaments et des produits de santé pour déterminer l'efficacité thérapeutique de l'ivermectine en tant que prophylactique à faible coût pour prévenir la propagation du COVID-19. Plusieurs autres inhibiteurs de 3CLpro sont également à l'étude, notamment GC376 et son analogue GC373, qui sont actuellement utilisés pour traiter la péritonite infectieuse féline mortelle causée par un coronavirus. Le GC376 et le GC373 se sont avérés être des inhibiteurs très puissants du SARS-CoV-2 dans un protocole de culture cellulaire utilisant des cellules pulmonaires de singe et humaines infectées par le SARS-CoV-2. Les médicaments étaient efficaces avec des valeurs de CI50 dans la gamme nano-molaire sans toxicité apparente [Freedberg DE, 2020].

Une autre cible est la protéase de type papaïne, PLpro, qui joue un rôle important dans l'atténuation de l'inflammation de l'hôte et des voies de signalisation antivirales, y compris l'inhibition de la production d'interféron de type I.

#### **IV.2.5. Famotidine**

La famotidine est un antagoniste des récepteurs de l'histamine 2 (H2) largement utilisé comme antiacide et des études in vitro indiquent que le médicament a des effets inhibiteurs sur plusieurs protéases virales, notamment 3CLpro, PLpro et TMPRSS2. Les données cliniques rétrospectives soutiennent les avantages de la famotidine dans la réduction de la

mortalité chez les patients atteints de COVID-19, indépendamment de son action d'abaissement de l'acide gastrique, comme l'indique l'absence d'avantages associés à l'utilisation d'inhibiteurs de la pompe à protons. Les données d'un certain nombre d'essais cliniques de phase III avec la famotidine sont attendues en 2021.

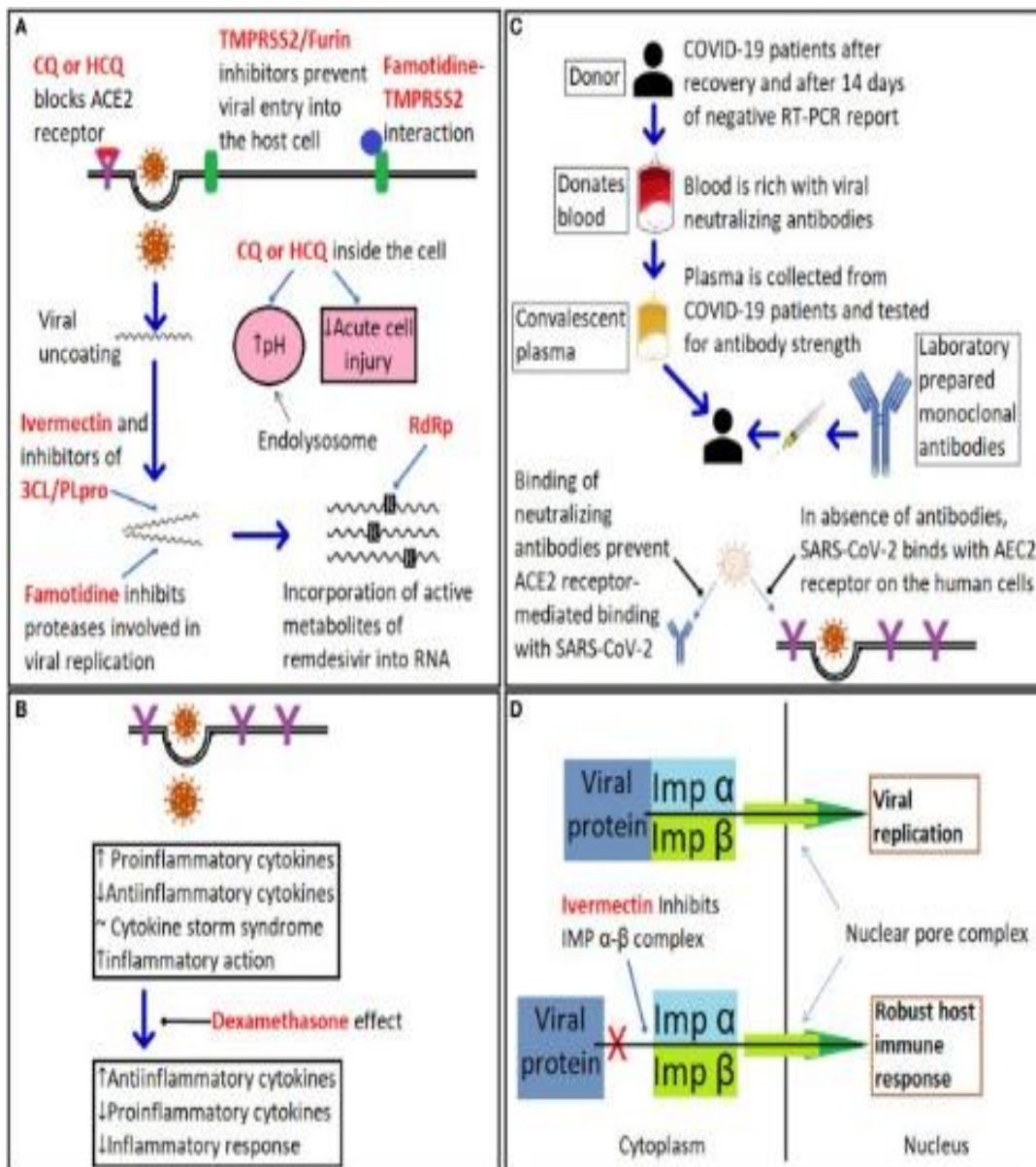
#### **IV.2.6. Inhibiteurs TMPRSS2 et furine**

En outre, il existe plusieurs médicaments expérimentaux, qui semblent prometteurs en tant qu'inhibiteurs potentiels de la sérine protéase qui ciblent les deux enzymes hôtes, TMPRSS2 et furine qui sont détournés par le virus SARS-CoV-2, pour faciliter la liaison à ACE2 et entrer dans la cellule hôte [Bestle D, 2020]. Le mésilate de camostat, un médicament approuvé au Japon pour une utilisation dans l'inflammation pancréatique, et le nafamostat associé, ont également été montrés dans des études in vitro pour inhiber le TMPRSS2, avec des CI50 de 142 et 55 nm, respectivement. Camostat est en cours d'essais cliniques pour le traitement du COVID-19 avec un avec l'identifiant NCT04608266 maintenant en phase III «Un essai randomisé multicentrique pour évaluer l'efficacité et l'innocuité du mésylate de camostat pour le traitement de l'infection par le SRAS-CoV-2 — COVID-19 chez les patients adultes ambulatoires (CAMOVID). » Il convient de noter que, étant donné que le TMPRSS2 et la furine sont importants pour une variété de processus cellulaires chez l'hôte, le risque d'effets secondaires peut être accru avec l'utilisation chronique d'inhibiteurs de ces enzymes.

#### **IV.2.7. Plasma de convalescence et traitement par anticorps ciblés**

Comme mentionné précédemment, les données cliniques soutenant l'utilisation de la thérapie plasmatique de convalescence pour le traitement des patients atteints de COVID-19 n'ont pas démontré de bénéfices positifs; cependant, des progrès ont été réalisés dans l'utilisation d'anticorps synthétiques pour traiter les patients atteints de COVID-19. Eli-Lilly a reçu l'EUA de la FDA pour son anticorps unique, le bamlanivimab (LY-CoV555) le 9 novembre 2020 et le 21 novembre, la société de biotechnologie Regeneron a également reçu l'EUA pour son produit, REGN-CoV2. Le LY-CoV555 est un anticorps unique très puissant isolé de patients convalescents COVID-19 qui se lie à la RBD de la protéine de pointe virale. Le cocktail d'anticorps de Regeneron se compose de deux anticorps, le casirivimab (REGN10933) et l'imdevimab (REGN10987) qui ciblent également et se lient de manière irréversible au site RBD sur la protéine de pointe du SRAS-CoV-2 et reposaient également sur le criblage d'anticorps neutralisants humains à partir du COVID-19 récupéré patients [Baum A, 2020].

Pour le LY-CoV555 et le REGN-CoV2, des données positives ont été publiées indiquant que les cocktails d'anticorps sont particulièrement efficaces chez les patients ayant une charge virale élevée ou dont le système immunitaire n'avait pas répondu au virus, cependant, ils ne sont pas efficaces chez les patients gravement malades hospitalisés. AstraZeneca développe également une thérapie ciblée par anticorps, AZD7442, pour le traitement prophylactique des adultes exposés au COVID-19 avec les données d'un essai clinique de phase III, NCT04507256, attendu en 2021.



**Figure 2:** Le mécanisme d'action de différentes thérapeutiques contre le COVID-19.

(A) illustre le mode d'action de médicaments ciblant le COVID-19, notamment la chloroquine (CQ) et l'hydroxychloroquine (HCQ), qui ont de multiples sites d'action putatifs : (i). récepteur ACE2 pour le SRAS-CoV-2 ; (ii). augmenter le pH de l'endolysosome; et (iii). suppression de la réponse immunitaire. Les sites d'action des inhibiteurs de TMPRSS2 tels que le camostat, la famotidine et les inhibiteurs de la furine sont indiqués ; la famotidine est également un inhibiteur putatif des protéases 3CL/PLpro ; l'ivermectine est un inhibiteur putatif de TMPRSS2 qui inhibe également le complexe importine (IMP)  $\alpha$ - $\beta$  et la réplication virale ; tandis que le remdesivir inhibe l'ARN polymérase virale.

(B) La dexaméthasone supprime l'expression des cytokines pro-inflammatoires.

(C) Résumé du rôle du plasma convalescent et de la thérapie par anticorps monoclonaux.

(D) L'ivermectine inhibe le complexe  $\beta$  d'importine hétérodimérique (IMP) en se liant directement à  $IMP\alpha$ , empêchant ainsi l'importation nucléaire de protéines virales clés (pour plus de détails et de références, voir la section Thérapeutique COVID-19).

#### **IV.2.8. Thérapies immunitaires**

##### **❖ Transfusion plasma**

La transfusion de plasma est utilisée pour les infections aiguës et est une forme d'immunothérapie pour les maladies infectieuses émergentes [Wang Y,2020]. Cette thérapie est basée sur l'administration passive de plasma avec des titres élevés de nAbs, générés de manière endogène par convalescence particuliers. Une telle thérapie est toujours utilisée et a été mise en œuvre avec succès dans les maladies sans traitement ou vaccin spécifique. Lors des dernières épidémies due à Ebola (EBOV), à la grippe A-H1N1 et coronavirus, ce type de thérapie a été mis en œuvre en tant que alternative au traitement [Baum A,2020].

Principalement, transfert de plasma de patients convalescents après Les infections au SRAS-CoV ou au MERS ont montré une réduction de la charge virale et l'amélioration des paramètres cliniques et de laboratoire, avec une diminution de la température, augmentation de la saturation en oxyhémoglobine, amélioration des taux de lymphocytes et diminution de la protéine C-réactive les niveaux. De plus, le plasma frais congelé (FFP) améliore les taux de survie chez les patients.

De même, récupération de COVID-19 la maladie a été observée chez des patients atteints d'une maladie grave après traitement par plasma de convalescence à nAbs élevé obtenus de donneurs survivants. Surtout, lors d'un Traitement FFP, autres produits biologiques impliqués dans la récupération des patients, tels que les cytokines anti-inflammatoires et défensines, sont également transférées. La principale limitation de FFP consiste à déterminer le titre nAb optimal requis pour une neutralisation efficace du virus. Ainsi, l'hémovigilance est essentielle car certains patients pourraient avoir besoin de plus d'une dose de FFP en fonction de la gravité de la maladie, comme décrit pour d'autres maladies virales où la transfusion de plasma a été utilisée. De plus, l'utilisation de FFP présente des risques équivalents pour les effets que toute autre transfusion de composants sanguins, tels que réaction allergique / anaphylactique et infections dues à d'autres microbes[Maisonasse P,2020].

# **Chapitre V**

## **Vaccins contre le covid-19**

En novembre 2020, le SRAS-CoV-2 a causé plus de 54 millions de cas et 1,3 million de décès dans le monde. Les États-Unis seuls ont signalé plus de 11 millions de cas et subi plus de 250 000 vies perdues depuis le premier cas signalé en mars 2020. Il y a des efforts mondiaux pour développer des thérapies et des vaccins contre le SRAS-CoV-2. La thérapie vise à raccourcir la période d'hospitalisation et augmenter la survie des patients infectés, tout en les vaccins prophylactiques visent à générer une immunité protectrice contre SRAS-CoV-2. SRAS-CoV-2 partage une similarité de séquence élevée avec le SRAS-CoV et prend une voie d'entrée pour infecter les cellules humaines [Y. Wan, 2020]. Autrement dit, le domaine de liaison au récepteur (RBD) de la sous-unité S1 de la protéine S interagit avec l'enzyme de conversion de l'angiotensine 2 (ACE2), suivie d'une fusion membranaire médiée par la sous-unité S2 [T. Tang, 2020].

Ainsi, la protéine S est un composant essentiel du SRAS-CoV-2 pour l'infection cellulaire. En outre, les anticorps neutralisants fonctionnels (NAb) générés chez les patients COVID-19 se sont avérés principalement épitopes cibles dans la protéine S, suggérant que la protéine S est une cible prometteuse pour la vaccination contre le SRAS-CoV-2.

Le processus de développement d'un vaccin comprend deux étapes clés:

- identifier un antigène et
- développer une approche de délivrance pour le dit antigène pour obtenir une immunité cellulaire et humorale robuste.

Dans le cas du SRAS-CoV-2, des expériences antérieures avec le SRAS-CoV et le MERSCoV ont permis le développement rapide de vaccins candidats. En novembre 2020, plus de 170 vaccins contre le SRAS-CoV-2 sont sous développement tel que suivi par l'Organisation mondiale de la santé, et de ceux-ci, 13 sont en cours d'essais cliniques humains de phase III. Ces vaccins les candidats peuvent être classés en quatre plates-formes vaccinales: virale Vaccins à base de vecteurs, d'ARN, d'ADN et de protéines [D.van Riel, 2020].

Notamment, au moment d'écrire ces lignes, BioNTech / Pfizer et Moderna ont récemment rapporté que leurs vaccins à ARNm présentaient une efficacité de protection > 90% basée sur leurs premières analyses intermédiaires des essais de phase III en cours [Kyung Soo Park, 2020].

BioNTech /Pfizer et Moderna ont obtenu une autorisation d'urgence pour leurs vaccins COVID-19 par la Food and Drug Administration (FDA).

Le virus entier est un type de vaccin conventionnel historiquement utilisé pour maladies majeures telles que la variole, la tuberculose et la fièvre jaune. Ce type de vaccin peut être divisé en deux types principaux: vivant atténué et virus entier inactivé vivant. Les virus inactivés sont plus fréquemment utilisés en raison de leur incapacité à induire une réversion virale.

Dirigé par des chercheurs en Chine, Wuhan Institute of Biological Products / Sinopharm et Sinovac Biotech ont procédé avec succès à des essais cliniques de phase III utilisant le SARS-CoV-2 entier inactivé [A. Koirala,2020].

De plus, il existe un vaccin à base d'adénovirus chimpanzé appelé ChAdOx1 développé par l'Université d'Oxford et AstraZeneca ainsi que Adenovirus 26 vaccins à base de vecteurs de Johnson & Johnson qui provoquent de puissantes immunes réponses[T.T. Le,2020].

En plus de ces vaccins à base de vecteurs viraux, un vaccin non viral plateformes, rendues possibles par les progrès de la nanomédecine et de la distribution de vaccins technologies, en sont aux derniers stades des essais cliniques (tableau 1). Ils comprennent l'ARNm et l'ADN codant pour les antigènes protéiques du SRAS-CoV-2 comme ainsi que des vaccins à base d'antigènes protéiques.

Les glycoprotéines membranaires des virus corona, appelées protéines de pointe, sont responsables de l'entrée virale dans les cellules hôtes, présentant ainsi une cible idéale pour les vaccins contre le SRAS-CoV-2[L. Du, 2009].

Vaccins basés sur l'ARNm et l'ADN sur les moyens de fournir du matériel génétique codant pour des candidats antigéniques dans les cellules hôtes, mais en raison de l'absorption cellulaire limitée et de l'instabilité de l'ARNm et l'ADN nus, ils nécessitent l'utilisation de véhicules de délivrance ou d'appareils d'électroporation.

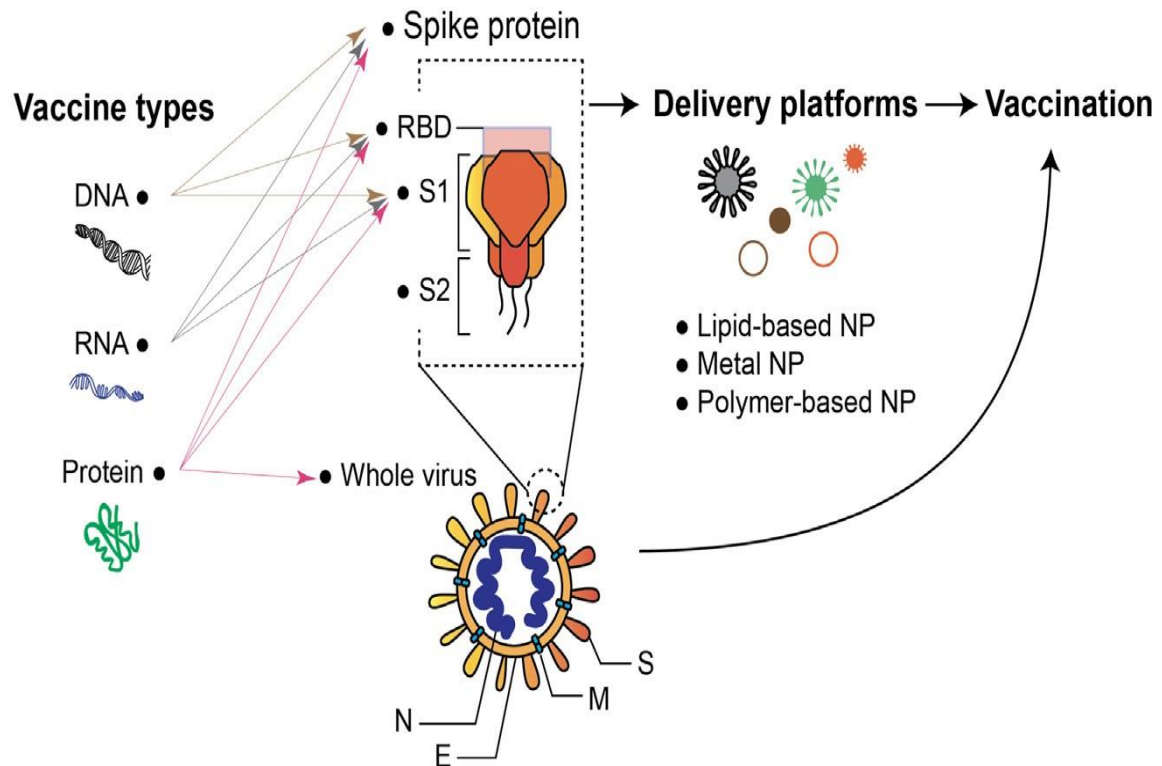
En revanche, les vaccins à base de protéines nécessitent la synthèse et la caractérisation des antigènes protéiques dans le cadre de la production de vaccins. Étant donné que les antigènes protéiques sont synthétisés par les cellules et sécrétés dans le milieu de culture sous des formes solubles, il est souvent difficile d'assurer un repliement adéquat des protéines et maintenir leur antigénicité.

Par conséquent, des techniques de recombinaison sont souvent impliquées lors de la conception de vecteurs génétiques pour conférer une stabilité aux protéines antigéniques avant les vecteurs. sont transfectées dans les cellules cibles.

Par rapport à l'ensemble des virus ou des vaccins à base de vecteurs viraux, des antigènes de sous-unités utilisés dans des vaccins à base d'ARNm, d'ADN et de protéines provoquent des réponses immunitaires plus faibles, nécessitant ainsi une co-administration d'adjuvants. Les adjuvants sont des agents immunostimulants ajoutés aux vaccins pour renforcer le système immunitaire réponse, et ils sont généralement sous la forme de ligands qui interagissent avec des récepteurs de reconnaissance de formes (PRR) sur des cellules présentatrices d'antigène (APC).

Une fois engagée, la signalisation en aval dans les APC déclenche diverses voies immunitaires qui conduisent à une activation immunitaire plus forte.

Dans cette revue, nous nous concentrons sur les principaux candidats vaccins non viraux actuellement en cours de développement contre le SRAS-CoV-2. Nous discutons du rôle de leurs plates-formes d'administration et présentent des facteurs cruciaux pour le développement de vaccins non viraux efficaces contre le SRAS-CoV-2, y compris immunogénicité, adjuvants, administration in vivo, sécurité, ainsi que les obstacles à surmonter pour un déploiement rapide et une vaccination de masse.



**FigureV.1:** Various vaccine platforms for vaccination against SARS-CoV-2[KyungSoo Park,2010]

**Tableau V.1 :** Vaccins non viraux COVID19 actuellement en cours de développement [Kyung Soo Park, 2010]

Vaccine type	Developer	Vaccine	Antigen	Formulation/Delivery system	Delivery route	adjuvant	Status	Clinical trial
Virus-like particles (VLP)	Medicago	CoVLP	Self-assembled VLP from S protein	Recombinant Coronavirus Virus-Like Particle (CoVLP)	IM	AS03 or CpG 1018	Phase 1	NCT04450004
Recombinant protein	Novavax	NVX-CoV2373	Prefusion state full-length S protein	Spontaneous nanoparticle formation	IM	Matrix-M (saponin-based)	Phase 2b	NCT04368988
	Clover Biopharmaceuticals	SCB-2019	S protein trimer	Patented Trimer-Tag technology used for trimer formation	IM	AS03 or CpG 1018	Phase 1	NCT04405908
	University of Queensland	UQ-CSL V451	Prefusion state S protein trimer	"Molecular clamp" used for trimer formation	IM	MF59 (Squalene-based)	Phase 1	ISRCTN51232965
DNA	University of Pittsburgh	PittCoVacc	S1 subunit protein	Microneedle patch	cutaneous	none	Preclinical	NA
	Inovio Pharmaceuticals/International Vaccine Institute	INO-4800	S protein	Electroporation	ID	None	Phase 1/2	NCT04336410
	AnGes, Inc.	AC0301-COVID19	S protein	Plasmid	IM	None	Phase 1/2	NCT04447781
	Symvivo	bacTRL-Spike	S protein	Bifidobacterium longum	Oral	None	Phase 1	NCT04334980
	Genexine Consortium	CX-19	S protein	Electroporation	IM	None	Phase 1/2	NCT04445389
	Osaka University/AnGes/ Takara Bio	AC0301-COVID19	S protein	Plasmid	IM	unknown	Phase 1/2	NCT04463472
mRNA	Moderna/NIAD	mRNA-1273	A full-length, prefusion stabilized spike (S) protein of COVID-19	Lipid nanoparticle	IM	None	Phase 3	NCT04470427
	BioNTech/Fosun Pharma/Pfizer	BNT-162b2	Optimized SARS-CoV-2 full-length spike protein or optimized SARS-CoV-2 receptor-binding domain (RBD)	Lipid nanoparticle	IM	None	Phase 3	NCT04368728
		BNT-162a1					Phase 1/2	
		BNT-162b1						
	Arcturus/Duke-NUS	ARCT-021	Self-replicating mRNA encoding the prefusion spike protein of 2019-nCoV	LUNAR® lipid nanoparticle	IM	None	Phase 1/2	NCT04480957
	People's Liberation Army (PLA) Academy of Military Sciences/ Walvax Biotech.	ARCov	SARS-CoV-2 receptor-binding domain (RBD)	Lipid nanoparticle	IM	None	Phase 1	ChiCTR2000034112
	Imperial College London	SARS-CoV-2 sARNA LNP	SARS-CoV-2 spike protein	Lipid nanoparticle	IM	None	Phase 1	ISRCTN17072692
	Curevac AG	CvxCov	full-length spike protein of SARS-CoV-2	Lipid nanoparticle	IM	None	Phase 1	NCT04449276

## V.1 Vaccins ARNm :

### V.1.1 Vaccins ARNm au stade clinique :

Les vaccins à ARNm sont les vaccins COVID-19 les plus avancés en phase Essais cliniques III . Bien que les vaccins à ARNm soient un nouveau technologie sans produit approuvé sur le marché, les vaccins à ARNm présentent un certain nombre d'avantages par rapport à d'autres plates-formes[K.A. Hajj, 2017] :

✚ les vaccins à ARNm sont sûrs. Par rapport aux vaccins à base d'ADN, L'ARNm ne s'intègre pas dans le génome de l'hôte, minimisant ainsi le problème de génotoxicité. Par rapport aux virus inactivés ou aux vaccins vivants à base de vecteurs, la production d'un vaccin à ARNm synthétique est contrôlée de la qualité avec un risque réduit de contamination biologique pendant production . De plus, l'ARNm lui-même ainsi que l'ARNm lipidique les supports sont biodégradables.

✚ les vaccins à ARNm ont démontré leur efficacité pour induire une forte réponse immunitaire humorale et cellulaire dans des études précliniques [B. Petsch,2012] et cliniques [M. Alberer, 2017].

✚ Le vaccin à ARNm peut être conçu et fabriqué dans un court laps de temps pour répondre aux besoins d'épidémie de pandémie. La chimie bien définie fait de l'ARNm vaccin adapté au processus de fabrication de pointe utilisé dans l'industrie pharmaceutique.

La capacité de production à grande échelle jusqu'à un milliard de doses font des vaccins à ARNm les pionniers de la Développement du vaccin COVID-19 [A.M. Reichmuth,2016].

Premièrement, l'ARNm sans formulation appropriée est instable et peut être rapidement dégradé par les RNases omniprésentes in vitro et in vivo.

Deuxièmement, l'ARNm est imperméable aux membranes cellulaires et ne peut pas être efficacement internalisé dans le cytosol pour la traduction.

Troisièmement, l'ARNm pourrait activer le système immunitaire inné et potentiellement induisent une inflammation et une toxicité. Quatrièmement, les vaccins à ARNm actuels nécessitent une chaîne du froid stricte pour le stockage et la distribution, pose un défi logistique pour le déploiement dans les pays en développement.

Pour résoudre ces problèmes, il est crucial de concevoir des technologies de livraison d'ARNm appropriées qui peuvent protéger l'ARNm de la dégradation, de la

navette ARNm pour cibler le cytosol cellulaire et augmenter la stabilité thermique, les profils d'innocuité et l'efficacité du vaccin ARNm.

La vitesse de développement des vaccins contre le SRAS-CoV-2 a été sans précédent[ **Lisa A.Jackson,2020**],et cela se reflète bien dans les progrès de l'ARNm développement de vaccins contre le SRAS-CoV-2. Dès le 13 janvier, Moderna et le NIH a finalisé la séquence d'ARNm-1273 codant pour la protéine de pointe stabilisée par préfusion de SARS-CoV-2 et s'est orienté vers la clinique fabrication, seulement 4 jours après que l'OMS a annoncé un cas de pneumonie inconnu trouvé à Wuhan et 2 jours après qu'une équipe chinoise a signalé la séquence génétique du COVID-19. Un court 88 jours plus tard, ARNm-1273 doses ont été administrées aux premiers participants aux essais cliniques sur l'homme.

De plus, il y a beaucoup plus d'ARNm vaccins actuellement en essais cliniques de phase I / II (**tableau V.1**).

Basé sur précliniques et cliniques et études rapportées à ce jour, les vaccins à ARNm sont bien toléré et induisent des réponses immunitaires efficaces[**M.J. Mulligan,2020**].

Pour mRNA-1273[**Lisa A. Jackson,2020**],l'essai clinique de phase I de 47 participants a révélé de solides réponses NAb après deux vaccinations, atteignant des niveaux similaires de activité neutralisante observée dans les sérums de convalescence de patients atteints de COVID-19.

De plus, les réponses des lymphocytes T CD4 asymétriques Th1 et un faible niveau La réponse des lymphocytes T CD8 à S-2P a été observée (**FigV.2a-d**). Aucun effet indésirable grave n'a été observé, mais un effet indésirable systémique et local léger ont été notés. Conformément aux résultats humains, la vaccination à l'ARNm-1273 de primates non humains a induit une NAb robuste et une asymétrie de Th1 Réponses des lymphocytes T CD4, protégeant les animaux de la provocation par le SRAS-CoV-2.

Pour le BTN-162b1 de BioNTech , les vaccinations prime-boost en Phase I / II essai clinique a induit de fortes réponses NAb, qui étaient supérieures à les taux observés dans les sérums humains convalescents de COVID-19. En particulier, La vaccination BTN-162b1 a généré de forts CD4 + spécifiques à la RBD et Réponses des lymphocytes T CD8 +, qui peuvent cibler et lyser les cellules infectées avec mémoire à long terme (Fig. 2e-h). Semblable à l'ARNm-1273 de Moderna, il n'y a eu

aucun effet indésirable grave, mais des effets secondaires légers ont été notés pour le BTN-162b1 de BioNTech.

Dans l'ensemble, les résultats cliniques à un stade précoce indiquent que les vaccins à ARNm sont bien tolérés et induisent une forte et les réponses immunitaires cellulaires contre COVID19.

### V.1.2 Considérations de conception pour les vaccins à ARNm :

Comme le montre le tableau 1, il est intéressant de noter que tous les vaccins à ARNm COVID-19 en cours de développement clinique sont délivrés par des nanoparticules lipidiques (LNP).

Les LNP, encapsulant l'ARNm dans une structure lipidique solide, sont composés de quatre composants[N. Pardi, 2018]: lipides cationiques ou ionisables pour Complexation d'ARNm, cholestérol pour stabiliser la nanoparticule, aide des phospholipides pour aider à la formation et à la libération intracellulaire, et Lipides PEGylés pour réduire les interactions non spécifiques. Les LNP présentent les avantages suivants en tant que vecteurs d'ARNm non viraux:

- ✚ la LNP encapsule et condense efficacement l'ARNm;
- ✚ La LNP favorise la délivrance intracellulaire deARNm au cytosol en augmentant l'absorption cellulaire et en déclenchant l'endosomal s'échapper;
- ✚ LNP augmente la stabilité de l'ARNm en les protégeant de la dégradation dans les espaces extracellulaires;
- ✚ LNP sont composés de biocompatiblesmatériaux adaptés à l'usage humain;
- ✚ La LNP de qualité GMP est synthétisée à grande échelle.

Développer un vecteur LNP efficace pour la délivrance de l'ARNm de COVID-19 vaccin, deux facteurs critiques doivent être pris en considération. La première est le choix d'un lipide cationique ou ionisable approprié [K.A. Hajj,2017] .

Compte tenu de l'état de la pandémie d'urgence, en réutilisant ces lipides ionisables performants pour une La formulation du vaccin à ARNm COVID-19 serait une bonne option. Le deuxième facteur optimise le cholestérol, les phospholipides auxiliaires et la composition lipid PEG ainsi que leur rapport relatif en LNP.

Cependant, le processus d'optimisation des formulations implique généralement, beaucoup plus de variables et nécessite des ressources importantes. Pour rendre le processus d'optimisation plus efficace, des méthodologies de conception d'expériences (DOE) ont été appliquées, y compris les deux conceptions factorielles fractionnelles et dépistage définitif[Kyung SooPark,2020].

Sur cette base, la LNP contenant C12-200 a été optimisée pour la livraison d'ARNm d'érythropoïétine et une amélioration de puissance de 7 fois a été observée.

Outre les LNP, les liposomes cationiques pourraient également être utilisés pour Applications vaccinales à base d'ARNm, telles que largement utilisées agent de transfection, lipofectamine composée de (N- (1- (2,3-dioléyloxy) propyl) -N- (2- (sperminecarboxamido) éthyl) -N, trifluoroacétate de Ndiméthylammonium (DOSPA) et (DOPE)] ainsi que des liposomes contenant du 1,2-di-O-octadécényl3 triméthylammonium-propane(DOTMA)et du 1,2-dioléoyl-3

triméthylammoniumpropane (DOTAP). En particulier, un Une étude récente a rapporté un vaccin à ARNm contre le mélanome prêt à l'emploi qui nécessite un simple mélange du composant ARN avec un liposome cationique suspension avant injection[**Kyung SooPark,2020**].

Le système de délivrance d'ARNm optimisé a activé les APC et codé divers antigènes du cancer pour induire fortes réponses des lymphocytes T spécifiques de l'antigène, conduisant à une régression tumorale.

Une autre catégorie importante de système de délivrance d'ARNm est basée sur particules hybrides polymère et polymère / lipide. Par exemple, un polyplex basé sur l'Île-du-Prince-Édouard système a été utilisé pour délivrer un réplicon ARN auto-amplifiant (RepRNA) vaccins codant pour l'hémagglutinine et la nucléocapside du virus de la grippe.

Une nanoparticule de dendrimère-ARNm avec revêtement lipide-PEG a été utilisée comme vaccin à dose unique contre les infections mortelles multiples, y compris Ebola virus, grippe H1N1 et Toxoplasma gondii. Poly (lactique-co-glycolique acide) (PLGA) a également été incorporé dans un système de délivrance d'ARNm lipidique, appelé TT3-LLN, pour la délivrance d'ARNm dans des lignées cellulaires humaines[**Y. Dong,2016**].

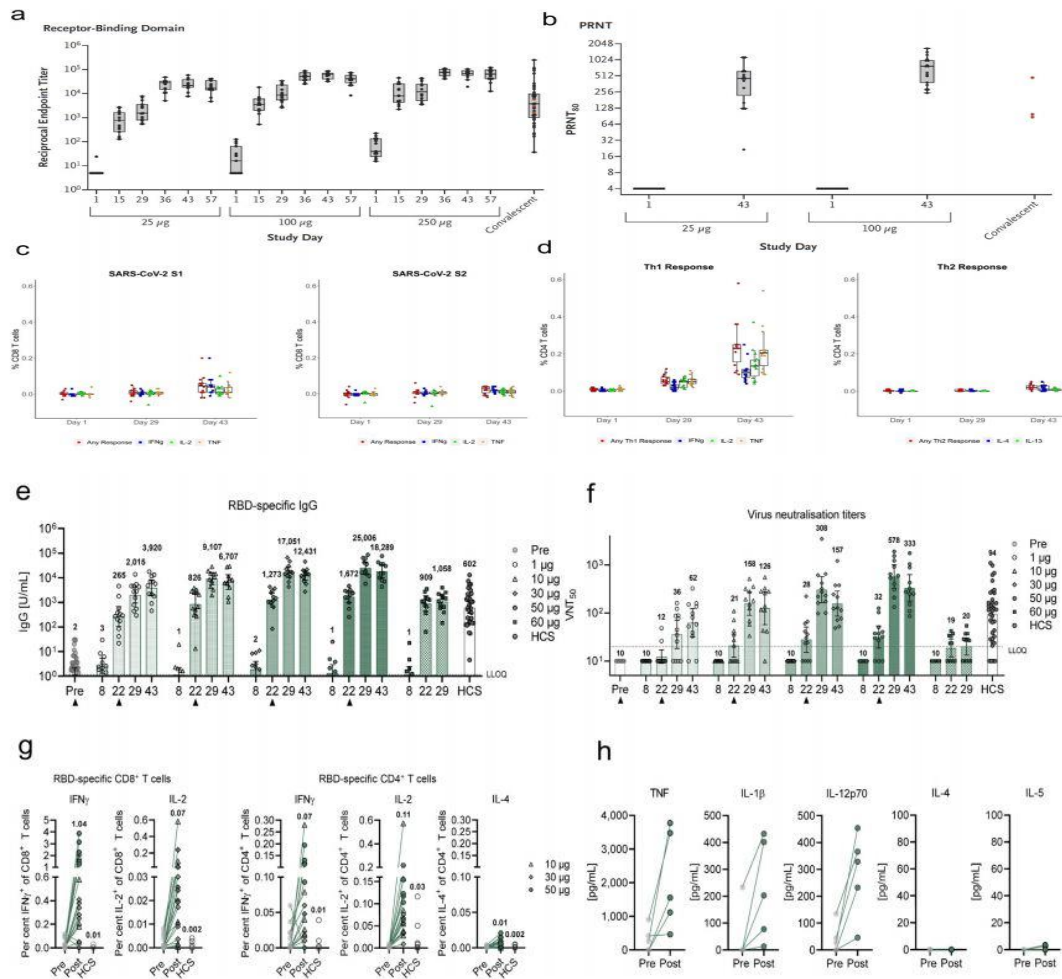
CARTs est une classe de polymères cationiques oligo (carbonate-β-α-amino ester), qui se complexent avec l'ARNm et facilitent la délivrance intracellulaire. Après Le complexe ARNm-CART est internalisé, les CART subissent une auto-immolation dégradation pour libérer l'ARNm pour la traduction des protéines. Cette conception a favorisé l'échappement endosomique et la traduction des protéines à la fois in vitro et in vivo. Par exemple, une nanoparticule de pinceau polymère formée en mélangeant des poly (β-amino esters) (PBAE) avec du cholestérol, DSPC (1,2-distéaroyl-sn-glycéro-3-phosphocholine), mPEG2000-DMG (1,2-dimyristoyl-sn-glycéro-3-

phosphoéthanolamine-N- [méthoxy (polyéthylène glycol) -2000]), et l'ARNm a été rapporté pour l'ARNm demandes de livraison. La co-formulation de PBAE et de lipide-PEG a effectivement délivré de l'ARNm aux poumons, et une structure noyau-enveloppe formé par encapsulation du complexe PBAE / ARNm dans un liposome à double couche a été rapporté pour des applications vaccinales à base d'ARNm [S. Persano,2017 ].

D'autres systèmes de délivrance, tels que les micelles et les émulsions, ont été utilisés pour les applications de distribution d'ARNm. Par exemple, une nano micelle PEG-Pasp (TEP) -Chol a été utilisé pour l'administration systémique du vaccin à ARNm comme traitement du cancer du pancréas, et un système micellaire basé sur les conjugués ramifiés PEI-acide stéarique (PSA) a été utilisé pour Administration du vaccin ARNm du VIH. Chez les souris, les rats, les lapins et les non-humains primates, CNE formulé avec un vaccin à ARNm auto-amplifiant induit réponses humorales et des lymphocytes T efficaces.

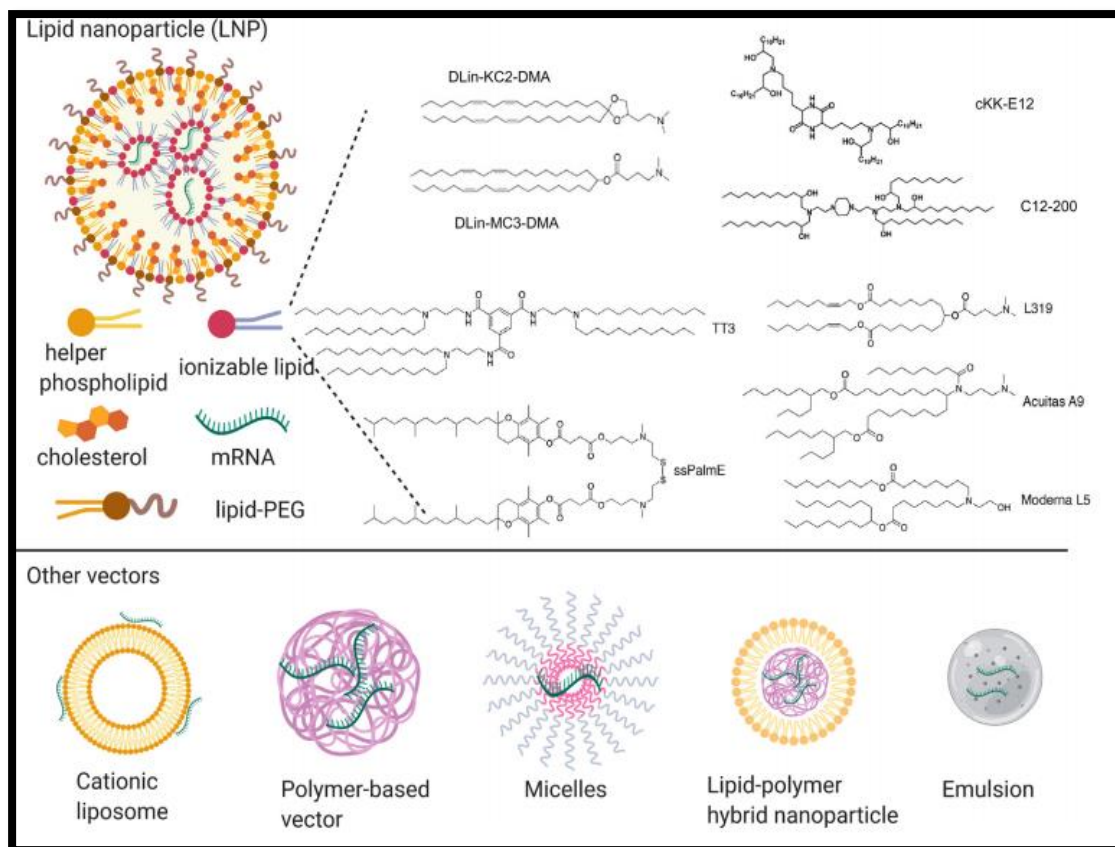
Dans l'ensemble, le vaccin à ARNm est une technologie puissante pour lutter contre un épidémie pandémique, comme le COVID-19. Malgré l'optimisme général sur les vaccins à ARNm contre COVID-19, d'autres études sont nécessaires pour contre-mesures efficaces contre d'autres agents pathogènes émergents et pour cas potentiels de mutations ou de réapparitions saisonnières du SRAS-CoV-2. Les systèmes d'administration de vaccins à ARNm standardisés prêts à l'emploi doivent être optimisé pour :

- ✚ emballer et protéger efficacement le codage de l'antigène ARNm;
- ✚ améliorer la stabilité thermique du vaccin, permettant le stockage et l'expédition à température ambiante;
- ✚ ont des propriétés adjuvantes appropriées, induisant efficacement une réponse immunitaire cellulaire et humorale contre un large éventail d'agents pathogènes; et
- ✚ être en mesure de réaliser une production de masse contrôlée de qualité pour les besoins mondiaux à un coût abordable.



**Figure V.2 :** Réponse immunitaire humorale et cellulaire des ARNm-1273 et BTN-162b1.

a-d, réponse immunitaire humorale (a-b) et réponse immunitaire cellulaire (c-d) de l'ARNm-1273 contre le SRAS-CoV2. e-h, réponse immunitaire humorale (e-f) et réponse immunitaire cellulaire (g-h) de l'ARNm-1273 contre le SRAS-CoV-2 . [M.J. Mulligan,2020].



**Figure V.3 :** mRNA vaccine delivery systems applicable for COVID-19 vaccine development.

## V.2 .Vaccins ADN

### V.2.1 .Vaccins ADN de stade clinique

Les vaccins à ADN sont basés sur des plasmides bactériens qui codent pour le vaccin antigènes dirigés par des promoteurs eucaryotes. Contrairement aux antigènes protéiques, le plasmide (vaccin à ADN) doit pénétrer dans le noyau de cellules, y compris les APC [Kyung Soo Park, 2020].

Une fois à l'intérieur du noyau, l'expression de ces gènes codés par plasmide permet la synthèse d'antigènes étrangers, qui font l'objet d'une surveillance immunitaire et déclenchent ensuite des réponses immunitaires humorales et cellulaires. Les avantages des vaccins à ADN sont leur facilité de fabrication, de stockage et de sécurité.

La fabrication d'ADN plasmidique est rapide, avec un lot terminé en 2-4 semaines. En termes de stockage, l'ADN plasmidique, contrairement aux chromosomiques ADN, renature facilement et rapidement dans de nombreuses conditions sans perte d'activité biologique, la rendant plus similaire à une petite molécule en termes de stockage et d'analyse qu'un produit biologique. Vaccins ADN

n'ont montré aucun effet indésirable significatif dans les essais cliniques [Kyung Soo Park, 2020].

Chacun des ces avantages prouvent que des vaccins à ADN peuvent être développés et déployés rapidement. Cependant, les vaccins à ADN présentent également des inconvénients potentiels.

De plus, une inflammation chronique peut survenir en raison du vaccin en continu stimulant une réponse immunitaire humorale ou l'induction de tolérance à l'antigène protéique produit.

De plus, un autre défi majeur pour les vaccins à ADN est que les vaccins à ADN en général ont induit une immunogénicité limitée chez les primates non humains et humains, par rapport aux résultats observés chez la souris [R. Garg, 2017]. Co-livraison de plasmides d'ADN et les plasmides codant pour des cytokines, des chimiokines ou des molécules costimulatrices ont eu un effet positif sur la réponse immunitaire [R. Garg, 2017].

Les nanotransporteurs sont capables de conditionner efficacement les plasmides d'ADN, le protéger de l'agrégation des protéines sériques et empêcher l'endocytose / lysosomal dégradation. Applications des injecteurs à jet et administration bactérienne sont généralement sûrs et faciles à administrer.

Certains des premiers exemples de vaccins à ADN étaient ceux contre grippe et le premier essai clinique de phase I avec un vaccin à ADN était pour un vaccin anti-VIH-1. Depuis lors, il y a eu de nombreux essais cliniques utilisant des vaccins à ADN contre le cancer, les maladies et affections cardiovasculaires, infectieuses, neurologiques et oculaires, soulignant leur polyvalence. En réponse à la pandémie

Plusieurs entreprises et groupes de développement de la recherche, y compris Inovio Pharmaceuticals, Symvivo, AnGes, Genexine et Zydus Cadila, ont franchi cette prochaine étape en développant et en testant un vaccin à base d'ADN candidats aux essais cliniques. Inovio Pharmaceuticals a développé un vaccin et dispositif d'administration ciblant la protéine SARS-CoV-2 S (INO-4800).

Inovio a tiré parti des expériences antérieures dans le développement d'un vaccin MERSCoV (INO-4700), qui partage une structure globale de repli protéique avec SARS-CoV-2. Sujets traités par INO-4700 en Phase I / II aux essais cliniques ont montré des NAb durables, des réponses immunitaires des lymphocytes T et un taux de séroconversion de 96% .

Les deux candidats vaccins utilisent les dispositifs CELLECTRA exclusifs de la société pour la livraison directe d'ADN dans la peau. Contrairement aux autres approches ADN et acide nucléique, CELLECTRA surmonte une limitation clé en utilisant une brève impulsion électrique pour ouvrir de manière réversible de petits pores dans les cellules qui permettent aux plasmides d'entrer et produire l'antigène ciblé. De plus, le CELLECTRA 3PSP est un petit appareil portable et portatif qui peut fabriquer et stocké en grandes quantités avec un minimum d'entretien, caractéristiques qui sont souhaitables dans un contexte de pandémie.

Des études précliniques chez des souris et des cobayes immunisés par INO4800 ont montré que l'anti-SRAS-CoV-2 la liaison Abs a bloqué la liaison de l'ACE2, le principal récepteur de l'entrée cellulaire SARSCoV 2.

Les animaux ont provoqué des abdos dans les poumons après une seule immunisation intradermique, empêchant la réplication virale dans les poumons en cas de contestation avec le SRAS-CoV-2 . Dans l'essai clinique de phase I avec INO-4800, 40 adultes en bonne santé âgés de 18 à 50 ans ont reçu deux doses de 1 mg ou Doses de 2 mg à quatre semaines d'intervalle. Le vaccin INO-4800 a induit des réponses humorales et cellulaires équilibrées, 94% des participants présentant des réponses immunologiques basées sur l'humorale (liaison et neutralisation) et T réponses immunitaires cellulaires.

D'autres vaccins candidats à base d'ADN sont à l'étude, certains des qui ont des stratégies de livraison intéressantes. L'administration orale du vaccin est intéressante car elle se traduit par une amélioration de la sécurité, de l'observance du patient et coûts réduits , et le tractus gastro-intestinal avec plus de 300 m<sup>2</sup> de la surface muqueuse pourrait servir d'organe inductif immunitaire . Par exemple, Symvivo a développé un candidat vaccin bacTRL-Spike, qui est une capsule de gel lyophilisée de bactéries probiotiques génétiquement modifiées (bifidobactéries)qui peuvent coloniser l'intestin, se lier directement à l'intestin cellules épithéliales et répliquent, sécrètent et délivrent le plasmide de manière constitutive Molécules d'ADN codant pour la protéine SARS-CoV-2 S .

Parce que le vaccin est un médicament vivant, les taux de délivrance de gènes sont soutenus tout au long de la vie de la colonie bactérienne, résultant en un gène étendu expression dans toute la muqueuse épithéliale . Le vaccin oral est hautement avantageux pour la vaccination de masse, et cette plate-forme est l'un des deux

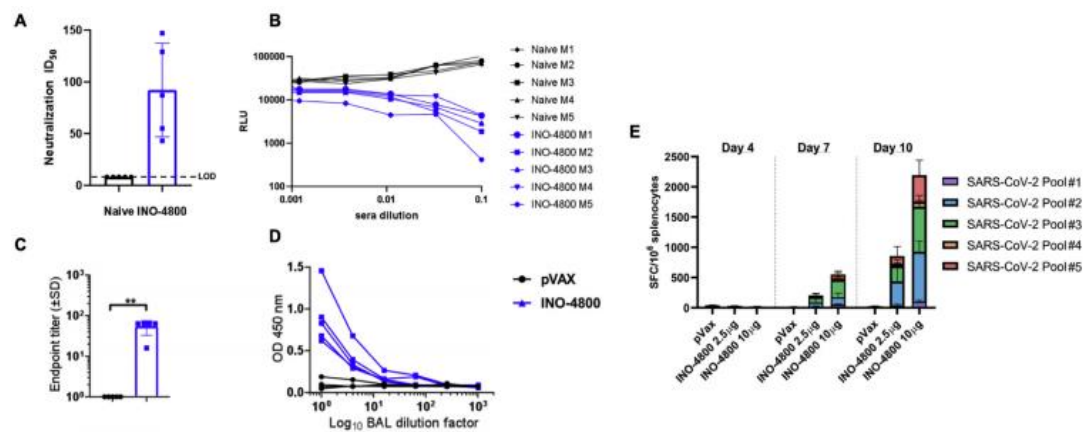
candidats vaccins COVID-19 administrés par voie orale actuellement en essais cliniques humains.

Un autre candidat vaccin à ADN avec une livraison stratégique sans aiguille est le GX-19 de Genexine, qui devrait débiter une phase Essai clinique II a bientôt .Le GX-19 est fourni par un injecteur à jet de PharmaJet, qui utilise un courant de fluide à haute pression et étroit pour pénétrer dans le la peau au lieu d'une aiguille.

La technologie Pharmjet s'est avérée sûr et efficace avec les vaccins commerciaux, y compris la grippe, le ROR, Vaccins contre le VPH et la polio [Kyung Soo Park,2020].

En outre, la livraison de porcs à base d'ADN vaccin antigrippal utilisant l'injecteur à jet sans aiguille PharmaJet Tropis a donné lieu à des réponses immunitaires comparables à celles du vaccin administré via aiguille et seringue, suivie d'une électroporation.

Le vaccin AnGes est administré par voie intramusculaire et fait actuellement l'objet d'essais cliniques de phase I / II. Le vaccin de Zydus est administré par voie intradermique et actuellement dans un essai clinique de phase II et entrant probablement Essais cliniques de phase III d'ici fin 2020 ou début 2021 [Kyung Soo Park,2020].



**FigureV.4 :** Le vaccin à ADN INO-4800 d'Inovio Pharmaceutical déclenche une réponse immunitaire chez les souris BALB/c.

(A) Neutralisation ID<sub>50</sub> chez les souris naïves et immunisées INO-4800 et (B) unités de luminescence relative (RLU) pour les sérums de souris naïves et vaccinées. (C) Liquide de lavage broncho-alvéolaire dosé pour les anticorps IgG spécifiques à la protéine SARS-CoV-2 Spike par ELISA. (D) Courbes de dilution BAL avec des

valeurs brutes de DO 450. (E) Réponses des lymphocytes T mesurées par IFN- $\gamma$  ELISpot dans les splénocytes avec des pools de peptides chevauchants couvrant la protéine Spike SARS-CoV-2.

### **V.2.2. Stratégies alternatives pour l'administration de vaccins à ADN**

Il existe d'autres stratégies d'administration d'ADN aux stades précliniques applicables au COVID-19. Une méthode pour améliorer l'efficacité in vivo se fait par l'addition d'adjuvants ou de plasmides codés par adjuvant. Lorsqu'il est co-délivré avec des plasmides codés par l'antigène, des adjuvants sont sécrétés dans la région environnante, où ils sont capables de stimuler à la fois les APC locales et les cellules des ganglions lymphatiques drainants (LN), ce qui entraîne réponses immunitaires plus fortes et plus durables . Une autre méthode consiste à appliquer des micro-aiguilles (MN) par voie transcutanée, ce qui permet le vaccin à ADN à déposer sur l'épiderme riche en cellules immunitaires et derme .

Comparé au vaccin à ADN soluble, le vaccin MN a provoqué des fréquences 3 fois plus élevées d'anticorps sériques IgG1 spécifiques de l'antigène, un excès de cellules T CD8 cytotoxiques 3 fois plus élevé et un mélanome inhibé métastases pulmonaires . Alternativement, l'administration orale d'un vaccin à ADN a été examinée dans des études précliniques. Contrairement au bacTRL-Spike de Symvivo vaccin, plasmides d'ADN M1 codant pour la protéine matricielle du virus H1N1 ont été encapsulés dans des liposomes cationiques et administrés par voie orale, entraînant une augmentation significative des titres d'IgG, de l'activité des lymphocytes T et de la mémoire immunologique .

En cas de contestation intranasale avec homologue virus grippal A, les taux viraux étaient inférieurs au niveau détectable dans le poumons de souris immunisées avec les liposomes d'ADNp M1 par voie orale, indiquant protection contre les défis respiratoires . Compte tenu de l'infection respiratoire par SARSCoV-2, les candidats-vaccins qui offrent une protection respiratoire améliorée sont d'un intérêt particulier.

Pris ensemble, les vaccins à ADN sont une stratégie vaccinale prometteuse contre le SRAS-CoV-2. Ils ont un design simple, facile et rapide processus de fabrication, et ils sont rentables et stables avec un longue durée de vie. Le principal inconvénient des vaccins à ADN est leur immunogénicité relativement plus faible par rapport à d'autres plates-formes vaccinales, mais d'autres études sont en cours pour améliorer cet aspect.

grâce à diverses stratégies de livraison. Bien qu'il n'y ait pas d'approbation Vaccin à base d'ADN sur le marché, les vaccins à ADN peuvent s'avérer être une plate-forme efficace pour la vaccination contre le SRAS-CoV-2.

### V.3 .Vaccins protéiques

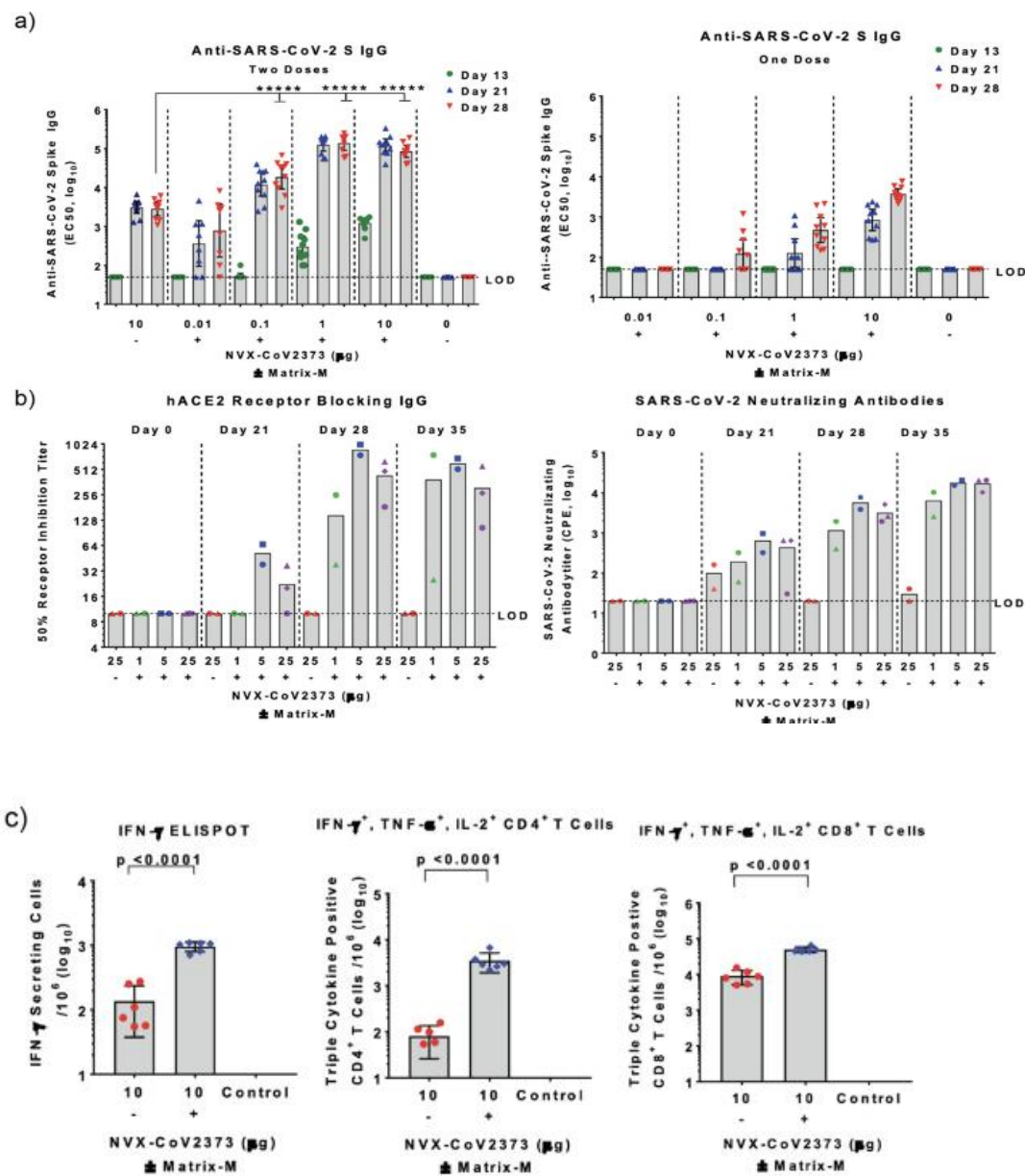
Par rapport aux vaccins à base d'ADN et d'ARN, qui ne nécessitent que la séquence génétique du virus, les vaccins à base de protéines nécessitent autres étapes techniques et de purification. De plus, les vaccins protéiques nécessitent des schémas posologiques multiples avec des adjuvants pour obtenir de fortes réponses immunitaires. Malgré ces obstacles, il existe de nombreux vaccins avant la préparation de l'antigène, comme le système d'expression de protéines cellulaires in vitro contre le SRAS-CoV-2 actuellement sous Essais cliniques de phase I, II et III.

Novavax est l'une des principales sociétés de vaccins dans une course contre SRAS-CoV-2. Le vaccin de Novavax utilise une protéine S pleine longueur (NVXCoV2373) génétiquement modifiée pour une plus grande stabilité structurale par rapport avec la protéine S de type sauvage (WT). Le variant de protéine S utilisé dans le L'étude avait une taille hydrodynamique plus petite de 27,2 nm ainsi qu'une distribution de taille, de masse et de forme plus uniforme (indice de polydispersité de 0,25- 0,29) par rapport à la taille WT de 69,53 nm et PDI de 0,46, malgré les deux ayant le même poids moléculaire . De plus, les mutations générées dans la séquence génétique sont considérées comme cruciales pour la production fiable d'une structure de protéine S pré-fusion stable et la capacité à résister au stress du pH et de la température après l'administration in vivo, tous contribuant à une réponse immunitaire plus efficace [Kyung Soo Park,2020].

Etudes précliniques réalisées avec NVX-CoV2373 chez la souris et le babouin ainsi que chez un humain de phase I / II un essai clinique a montré des titres élevés d'IgG anti-S protéine (IgG anti-S) avec activité neutralisante puissante (Fig V.5a) . L'essai clinique de phase I / II ont montré que le vaccin était bien toléré chez les adultes en bonne santé âgés de 18 et 59 ans lorsqu'il est utilisé dans un schéma à deux doses de 5 µg et 25 µg de la protéine S et de l'adjuvant Matrix M, respectivement. À noter, la co-injection avec l'adjuvant a considérablement amélioré l'efficacité du vaccin, permettant une économie de dose potentielle, ce qui est un avantage crucial pour vaccination de masse à l'avenir. Les anticorps sériques ont bloqué la liaison entre les protéines hACE2 et S et ont également pu neutraliser le virus(FigV.5b).

En plus de la réponse humorale, de fortes réponses cellulaires ont été observés, comme le montrent les tests ELISPOT et de coloration intracellulaire réalisée chez la souris (**FigureV.5c**).

En plus de Novavax, Clover Biopharmaceuticals, basé à Chine, se concentre sur la structure trimérique des protéines S natives sur le SRAS CoV-2. En utilisant cette protéine S trimérique en combinaison avec AS03 ou CpG 1018 plu Alun comme adjuvants, la société a récemment lancé un essai clinique de phase I (NCT04405908).



**FigureV.5** : Réponses immunitaires humorales et cellulaires observées chez les souris et les babouins après vaccination avec NVX-CoV2373 plus Matrix-M.

(a) Le régime à deux doses de NVX-CoV2373 (à gauche) a provoqué une réponse IgG sérique plus forte chez la souris, par rapport au régime à dose unique (à droite). L'ajout de Matrix-M a encore amplifié les réponses.

(b) Une étude de vaccination chez des babouins a indiqué la génération d'IgG sériques neutralisantes qui bloquaient l'interaction entre l'ACE2 et la protéine S (à gauche) et l'infection virale des cellules Vero E6 in vitro (à droite).

(c) Réponses immunitaires cellulaires examinées par dosage ELISPOT (à gauche) et dosage de coloration intracellulaire (à droite).

En plus des vaccins mentionnés ci-dessus dans les essais cliniques, il existe d'autres études actuellement en phase d'essais précliniques. Une étude qui a utilisé microneedle array (MNA) pour l'administration de MERS-CoV et SARS-CoV-2 Les protéines sous-unitaires en font partie . Le MNA a été synthétisé avec la carboxyméthylcellulose (CMC), qui est conçue pour se dissoudre par voie intracutanée après application sur la peau, libérant ainsi la teneur en protéines.

Lorsqu'ils sont testés sur des souris, les vaccins MERS-CoV et SARS-CoV-2 MNA ont pu induire des niveaux significatifs d'IgG sériques. Anticorps induits par le vaccin MERS-CoV ont pu neutraliser le virus in vitro, ce qui indique l'efficacité du MNA pour l'administration du vaccin à base de protéine S.

Lors de l'administration de protéines solubles en association avec un adjuvant (s) a démontré son efficacité in vivo, d'autres améliorations peuvent être apportées grâce à l'utilisation de la nanotechnologie. La nanotechnologie peut augmenter la stabilité des protéines solubles, tout en réalisant une livraison ciblée aux LN pour immunogénicité améliorée. Dans un contexte de pandémie, améliorer les l'administration peut se traduire par une réduction de la quantité d'antigène dans une dose, ce qui est avantageux pour tenter de maximiser la distribution des vaccins.

Semblable aux protéines solubles, la co-administration avec des adjuvants améliore encore efficacité des vaccins à base de nanoparticules. Par exemple, un système de nanoparticules lipidiques multicouches a été utilisé pour co-administrer l'antigène protéique et un adjuvant monophosphoryl lipide A (MPLA), un agoniste du récepteur-4 de type Toll utilisé dans d'autres vaccins approuvés par la FDA . Comparé à la formulation soluble, le vaccin nanoparticulaire a suscité beaucoup plus de réponses immunitaires humorales et cellulaires puissantes lors d'une injection sous-cutanée dans des études précliniques. Dans le même esprit, le phosphate de calcium Des nanoparticules à base de (CaP) ont été utilisées pour délivrer du CpG et un

antigène viral dérivé de l'hémagglutinine du virus de la grippe A, qui a empêché infection virale chez la souris.

Lorsque le vaccin était administré par voie intrapéritonéale ou intranasale, les nanoparticules étaient efficacement absorbées par les CD, conduisant par la suite à des réponses immunitaires médiées par les lymphocytes T. Une autre étude a utilisé les nanoparticules de CaP comme antigène tumoral / porteur de CpG pour traiter le cancer colorectal dans un modèle murin .

Le vaccin a induit une réponse immunitaire médiée par l'interféron de type I qui a augmenté la fréquence des cellules T CD8 + spécifiques de l'antigène tumoral modèle et a exercé un meilleur contrôle de la tumeur, par rapport à la formulation soluble. Ces exemples indiquent la polyvalence des systèmes de nanoparticules comme plates-formes de distribution de vaccins et montrent leur potentiel pour la vaccination contre d'autres agents pathogènes émergents.

#### **V.4. Vaccins contre les particules de type virus (VLP)**

Comme discuté, les virus vivants atténués ou inactivés provoquent des l'activation immunitaire, mais en raison de la réversion virale potentielle et de la variation de lots à lots, une technologie alternative est recherchée pour imiter structures virales.

Les particules de type virus (VLP) peuvent être conçues pour exprimer les protéines de surface ou les séquences d'acides nucléiques du virus natif sans risque de réplication ou d'infection. Bien que les VLP puissent être catégorisées en tant que vaccin protéique recombinant, ils maintiennent généralement la conformation des protéines virales, ce qui est avantageux par rapport aux autres vaccins à protéines sous-unitaires en termes d'antigénicité et d'immunogénicité [Kyung Soo Park,2020].

De plus, selon la façon dont la particule est conçue, elle peut potentiellement transporter d'autres types d'immunopotentiateurs pour améliorer la réponse immunitaire.

Medicago, une société pharmaceutique canadienne, a généré VLP en utilisant une méthode à base de plantes où un gène synthétique contenant une partie des gènes du SRAS-CoV-2 ont été transfectés dans une espèce de tabac en utilisant un vecteur bactérien. Ces plantes expriment alors des VLP qui peuvent être purifiées à travers plusieurs étapes de traitement. L'entreprise a entamé une phase I essai clinique (NCT04450004), test de l'injection intramusculaire de Corona particules de type virus (CoVLP) en combinaison avec CpG 1018 ou Adjuvants AS03.

L'entreprise utilisait auparavant une technologie similaire pour synthétiser les VLP pour les vaccins antigrippaux. L'un des essais cliniques de phase III a testé l'efficacité de leur vaccin contre la grippe chez des participants de plus de 65 ans (NCT03739112) et a démontré l'efficacité du VLP système de vaccination dans un groupe d'âge plus vulnérable au COVID-19 .

### **I.5 Vaccins à base de peptides**

Une autre stratégie consiste à utiliser des peptides comme immunogène, qui ont une structure relativement simple et stable par rapport aux protéines. Cependant, les peptides souffrent souvent d'une activation immunitaire sous-optimale en raison de leur courte longueur de séquences d'acides aminés fournissant des épitopes au système immunitaire cellulaire.

Cette limitation peut être améliorée en identifiant la région immunodominante d'un antigène, par ex. la sous-unité ou le récepteur domaine de liaison de la protéine S, suivi de la génération de peptides à partir de ces régions. Généralement, la séquence d'acides aminés de l'antigène est segmentée en plusieurs séquences courtes, et chaque segment est étudié pour sa immunogénicité [Kyung Soo Park, 2020].

Une étude similaire a été subie récemment pour examiner l'immunogénicité régionale de la protéine SARSCoV-2 S . En outre, il y a eu des tentatives de conception de vaccins multi-épitopes à base de peptides via des approches *in silico*.

Les épitopes candidats des protéines de la sous-unité SARS-CoV-2 sont criblés par calcul qui permet de prédire leur stabilité, leurs interactions avec les molécules du complexe majeur d'histocompatibilité (CMH) et l'immunogénicité .

Bien que ces études nécessitent la synthèse et les tests réels des vaccins pour valider leur efficacité, elles suggèrent une alternative possible aux vaccins protéiques de plus grande taille, qui peuvent fournir une meilleure sécurité et spécificité.

## **V.6. Perspectives futures des systèmes de distribution de vaccins**

### **V.6.1. Systèmes d'administration adjuvants**

Les adjuvants déclenchent les PRR sur les cellules immunitaires adaptatives et en fonction du type de PRR un adjuvant vise à activer, différentes voies de réponses immunitaires sont déclenchées. Par conséquent, l'utilisation d'un adjuvant puissant et bien adapté peut grandement améliorer l'efficacité du vaccin. Actuellement, de nombreux développeurs de vaccins COVID-19 utilisent des adjuvants dans leurs vaccins, y compris AS03 ( $\alpha$ -tocophérol et squalène de GSK dans une huile dans l'eau émulsion), CpG 1018 (agoniste du TLR-9 basé sur l'ADN de Dynavax), et MF59

(squalène de Novartis dans une émulsion huile-dans-eau), qui ont Démontré pour améliorer considérablement l'efficacité du vaccin en préclinique et en clinique études [Kyung Soo Park,2020].

Pour les vaccins à ARNm, l'ARNm lui-même pourrait se lier à certains récepteurs de reconnaissance de formes (PRR), tels que TLR 3, 7, 8, RIG-I, PKR, OAS et MDA5, qui à son tour induisent une activation immunitaire innée, un interféron de type I et production de cytokines pro-inflammatoires, fournissant des effets adjuvants. Cependant, l'activation de ces récepteurs pourrait déclencher un mécanisme antiviral naturel pour inhiber la traduction de l'ARNm exogène via la phosphorylation de eIF2 $\alpha$  et surexpression de la RNase L. Il existe donc un équilibre entre pour l'activation immunitaire innée induite par l'ARNm et la traduction de l'ARNm.

Les vaccins à ADN ont également le potentiel d'induire une activation immunitaire innée puisque de nombreux plasmides d'ADN produits dans des bactéries peuvent contenir motifs CpG non méthylés . En fait, des motifs CpG avaient été intentionnellement ajoutés aux squelettes plasmidiques d'ADN en tant qu'adjuvants pour améliorer le efficacité du vaccin.

Il a été démontré que l'inclusion de CpG améliore considérablement la réponse immunitaire médiée par les lymphocytes T spécifiques de l'antigène, qui souris protégées de la provocation avec des cellules de mélanome de souris .

De plus, de nombreux adjuvants sont en cours de développement avec un grand potentiel pour améliorer les vaccins à base de protéines. Un TLR-7/8 synthétique agoniste, 3M-052, a été utilisé comme adjuvant avec le VIH-1 protéine d'enveloppe gp140 (Env) dérivée du clade C 1086.C pour la vaccination contre le VIH-1.

### **1.6.2 Voies de vaccination :**

Bien que les vaccins COVID-19 en cours de développement utilisent principalement injection intramusculaire (IM) (tableau 1), les autres voies de vaccination doivent être envisagée, y compris l'injection intradermique (ID), sous-cutanée (SC), intranasale (IN) et intraveineuse (IV).

Seuls quelques rapports ont comparé différents itinéraires pour Vaccination à l'ARNm. En termes d'expression des protéines, la délivrance médiée par la LNP de l'ARNm de luciférase de luciole a été examinée chez la souris après injection via 6 itinéraires différents . Fait intéressant, à une dose de 5  $\mu$ g d'ARNm / injection, tandis que l'injection IV a produit la plus grande quantité totale de protéines, ID et l'injection

IM a prolongé la durée d'expression des protéines jusqu'à 10 jours, et l'injection ID fonctionnait mieux à une faible dose d'ARNm (0,1 µg /dose). Cette étude a également rapporté que l'administration intratrachéale (IT) induisait un niveau élevé d'expression des protéines dans les poumons, qui sont l'organe cible de l'infection par le SRAS CoV-2. Un vaccin informatique et de même un le vaccin intranasal offrirait probablement une meilleure protection immunitaire les poumons.

De nombreuses études ont étudié les effets des voies d'administration sur l'efficacité des vaccins à ADN. Lorsque plusieurs invasifs ou des voies d'administration non invasives ont été testées en délivrant un plasmide d'ADN codant pour l'antigène de surface de l'hépatite B à des souris et à des non-humains. Primates, il a été constaté que les voies d'administration avaient une grande influence sur la puissance et la nature des réponses immunitaires .

Pour Vaccins ADN qui sont en cours d'essais cliniques, injections IM ou ID, suivis par électroporation sont les méthodes de livraison les plus couramment utilisées. .

Dans une étude murine, l'injection ID d'ADN codant pour EGFP a suscité une mais seulement des réponses transitoires d'anticorps et de lymphocytes T CD8 cytotoxiques à EGFP, alors que l'injection IM induit plus faible mais plus durable réponses immunitaires jusqu'à 5 semaines après la dernière vaccination .

Aussi, lorsque les technologies d'administration de nanoparticules sont appliquées au vaccin à ADN, Diverses voies de vaccination peuvent être utilisées, y compris IM, ID, IV et SC .

Cependant, les protéines contournent la nécessité d'un tel prise en compte car ils sont facilement traités par les APC sans doivent être transcrits et traduits. Par conséquent, pour les vaccins à base de protéines, diverses voies d'administration autres que le SC traditionnel Les itinéraires de messagerie instantanée pourraient offrir des avantages supplémentaires. Par exemple, lorsqu'un vaccin antigrippal a été administré par voie IM ou IN, les deux ont obtenu une augmentation des cellules sécrétant des anticorps dans les organes lymphoïdes secondaires, tandis que la production d'IgA dans les tissus muqueux et la localisation dans le les poumons n'étaient prédominants que pour la vaccination IN .

Étant donné que les tissus muqueux sont l'endroit où se produisent la plupart des infections et des transmissions, ces études mettent en évidence l'avantage des protéines à base de protéines.

vaccins pour la voie muqueuse de vaccination et l'induction de la muqueuse immunité.

### **V.6.3 Stabilité thermique**

Un autre domaine actif de la formulation d'ARNm est d'augmenter la température stabilité des vaccins à ARNm, afin de permettre le stockage et l'expédition en moins conditions restreintes. Pour le stockage et le transport à long terme, les vaccins ARNm actuels de BioTech / Pfizer et Moderna dans les essais de phase III exigent  $-80^{\circ}\text{C}$  et  $-20^{\circ}\text{C}$ , respectivement, et cette exigence de la chaîne du froid limite grandement leur distribution et augmente le coût. Cependant, avant la recherche a montré que les vaccins à ARNm peuvent être conservés dans un forme [Kyung Soo Park,2020].Après lyophilisation, un vaccin à ARNm nu s'est révélé stable jusqu'à 10 mois dans un stockage à  $4^{\circ}\text{C}$

En utilisant un système de livraison, cela peut être encore amélioré. La plateforme d'ARN lyophilisé de CureVac peut être conservée 3 ans à  $5-25^{\circ}\text{C}$  et 6 mois à  $40^{\circ}\text{C}$ . Récemment, un vaccin à ARNm COVID-19 thermiquement stable, Il a été rapporté que l'ARCoV induisait des réponses immunitaires humorales et cellulaires prometteuses contre le SRAS-CoV-2 chez les souris et les primates non humains.

Cette formulation de vaccin à ARNm LNP pourrait être conservée pendant une semaine à température ambiante. Même si les vaccins à ARNm non congelés ne sont pas Prêt pour l'utilisation clinique en ce moment pour COVID-19, futures études sur Protection de l'ARNm par l'utilisation de systèmes de délivrance ou la co-administration avec Les inhibiteurs de RNase peuvent rendre cela possible.

Les vaccins à ADN sont considérés comme étant relativement plus stables thermiquement que l'ARNm. L'ADN lyophilisé est généralement stable à température ambiante, et comme mentionné ci-dessus, ils conservent leur structure d'origine après renaturation avant utilisation .

Un vaccin à ADN codant pour la glycoprotéine Ebola Testé dans un essai clinique s'est avéré être une température stable, démontrant la facilité relative de distribution de vaccins à base d'ADN à la parties du monde qui ne disposent pas de l'équipement nécessaire pour entretenir basses températures requises .

En revanche, les vaccins à base de protéines sont vulnérables à la déformation structurelle ou à la dégradation due à la fois aux fluctuations de température et au stockage,aux températures ambiantes.

La plupart des vaccins à base de protéines approuvés par la FDA, en solution ou sous forme lyophilisée, sont recommandés pour le stockage à 2-8 ° C ou à une température inférieure à zéro .

Dans une tentative d'empêcher la dénaturation thermique des protéines antigéniques, un antigène de la tuberculose (Ag85b) a été encapsulé dans une cage de silice, ce qui a entraîné une augmentation significative de l'antigène, stabilité en dehors des conditions réfrigérées ou lyophilisées .

#### **V.6.4 Administration ciblée des vaccins sur les ganglions lymphatiques**

Les ganglions lymphatiques sont des tissus cibles essentiels pour l'administration du vaccin. Pour la réponse immunitaire cellulaire, les cellules T sont amorcées par les CD et activées dans les LN. Pour Réponse immunitaire humorale, les centres germinaux dans les LN sont cruciaux pour maturation d'affinité des anticorps et changement d'isotype. Par conséquent, le ciblage Les LN peuvent grandement améliorer la réponse immunitaire.À cette fin, les vaccins ARNm et ADN pourraient être directement injectés dans LN ou délivrés par des nanoparticules de taille, de propriétés de surface et de charge appropriées pour un drainage efficace des LN[Kyung SooPark,2020].

L'injection intra LN directe nécessite un guide d'imagerie par ultrasons et a donné des résultats prometteurs dans les essais cliniques .

De même, lorsque les vaccins à ADN étaient directement injectés dans les LN périphériques, les réponses immunitaires étaient significativement plus grande, offrant une meilleure protection des souris contre la provocation tumorale, par rapport aux vaccins délivrés par des voies conventionnelles telles que la MI et ID .

#### **V.7. Statistiques et recherche Vaccins contre le coronavirus (COVID-19)**

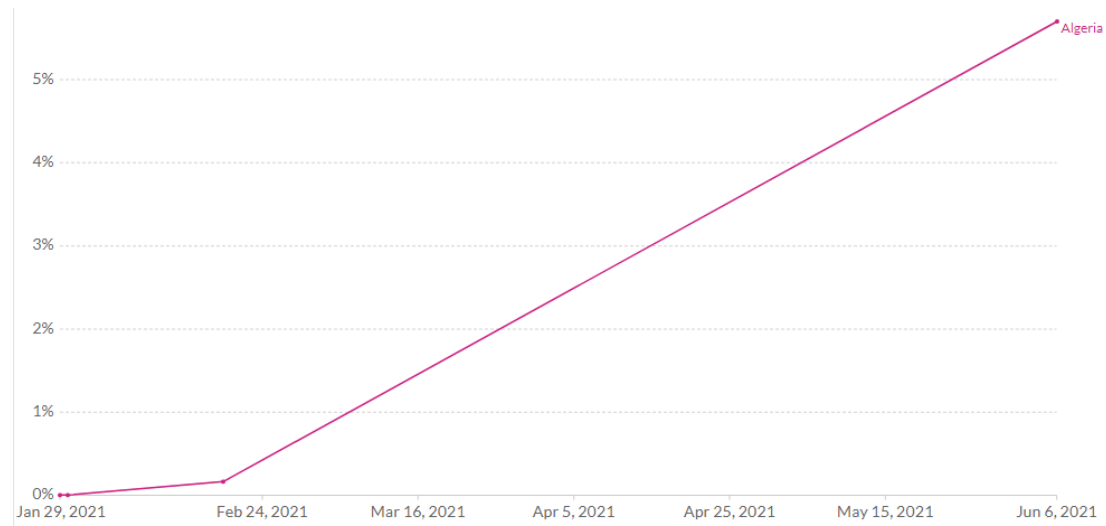
Pour mettre fin à cette pandémie, une grande partie du monde doit être immunisée contre le virus. Le moyen le plus sûr d'y parvenir est d'utiliser un vaccin. Les vaccins sont une technologie sur laquelle l'humanité s'est souvent appuyée dans le passé pour réduire le nombre de décès dus aux maladies infectieuses.

Moins de 12 mois après le début de la pandémie de COVID-19, plusieurs équipes de recherche ont relevé le défi et développé des vaccins qui protègent du SRAS-CoV-2, le virus qui cause le COVID-19 [Ourworldindata].

Maintenant, le défi est de rendre ces vaccins disponibles pour les gens du monde entier. Il sera essentiel que les habitants de tous les pays, et pas seulement des pays riches, bénéficient de la protection requise.

21,1% de la population mondiale a reçu au moins une dose d'un vaccin COVID-19. 2,5 milliards de doses ont été administrées dans le monde et 33,9 millions sont désormais administrées chaque jour.

Seulement 0,8 % des habitants des pays à faible revenu ont reçu au moins une dose.



**Figure V.1 :** Courbe de les personnes vaccinées en Algérie[Ourworldindata].

**Chapitre VI**

**Étude sur COVID-19 en**

**utilisant la**

**bioinformatique**

La bioinformatique est un domaine interdisciplinaire qui développe des méthodes et des outils logiciels pour comprendre les données biologiques, en particulier lorsque les ensembles de données sont volumineux et complexes. En tant que domaine scientifique interdisciplinaire, la bioinformatique combine la biologie, l'informatique, l'ingénierie de l'information, les mathématiques et les statistiques pour analyser et interpréter les données biologiques. La bioinformatique a été utilisée pour des analyses *in silico* de requêtes biologiques en utilisant des techniques mathématiques et statistiques.

La bioinformatique comprend des études biologiques qui utilisent la programmation informatique dans le cadre de leur méthodologie, ainsi qu'une analyse spécifique « pipelines » qui sont utilisées à plusieurs reprises, en particulier dans le domaine de la génomique. Les utilisations courantes de la bioinformatique comprennent l'identification de gènes candidats et de polymorphismes nucléotidiques simples (SNP). Souvent, une telle identification est faite dans le but de mieux comprendre la base génétique de la maladie, les adaptations uniques, les propriétés souhaitables (en particulier chez les espèces agricoles) ou les différences entre les populations. De manière moins formelle, la bioinformatique essaie également de comprendre les principes d'organisation au sein des séquences d'acides nucléiques et de protéines, appelées protéomique. [Lesk2013].

Ce chapitre commence par montrer un bref historique de la bioinformatique et donne des informations sur le National Center for Biotechnology Information (NCBI) et leurs ressources populaires. Ensuite, nous avons effectué une recherche dans la base de données Nucleotide sur le SRAS-CoV-2 et donné un détail sur chaque format (GenBank, FASTA et Graphics), nous avons également effectué une comparaison entre le virus SARS-CoV-2 contre les virus SRAS et Bat-CoV; RaTG13 à l'aide de l'outil BLAST.

### **VI.1. Une brève histoire de la bioinformatique**

La bioinformatique est souvent décrite comme étant à ses balbutiements, mais les ordinateurs sont devenus des outils importants en biologie moléculaire au début des années 1960. Une décennie avant que le séquençage de l'ADN ne devienne réalisable, les biologistes informaticiens se sont concentrés sur les données rapidement accumulées de la biochimie des protéines. Sans les avantages des superordinateurs ou

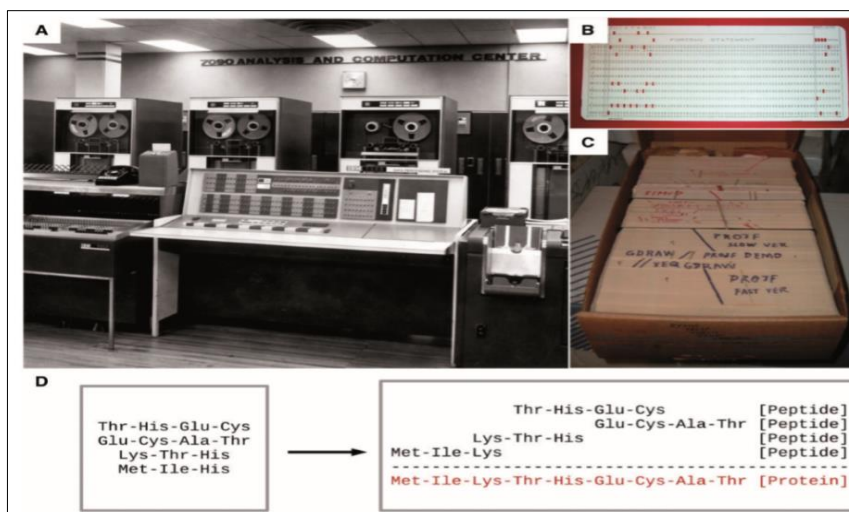
## Chapitre VI Étude sur COVID-19 en utilisant la bioinformatique

des réseaux informatiques, ces scientifiques ont jeté d'importantes bases conceptuelles et techniques pour la bioinformatique d'aujourd'hui [Joel.b , 2000].

### VI.1.1. Origines de la bioinformatique (1950 – 1970)

Margaret Dayhoff (1925-1983) était une physico-chimiste américaine qui a été la pionnière de l'application de méthodes informatiques au domaine de la biochimie. Dayhoff avait largement utilisé des méthodes de calcul pour sa thèse de doctorat en électrochimie et avait vu le potentiel des ordinateurs dans les domaines de la biologie et de la médecine. En 1960, elle devient directrice associée de la National Biomedical Resource Foundation. Là, elle a commencé à travailler avec Robert S. Ledley, un physicien qui a également cherché à apporter des ressources informatiques aux problèmes biomédicaux [Dayhoff1962].

De 1958 à 1962, tous deux ont combiné leur expertise et développé COMPROTEIN, «un programme informatique complet pour l'IBM 7090» conçu pour déterminer la structure primaire des protéines à l'aide des données de séquençage du peptide d'Edman. Ce logiciel, entièrement codé en FORTRAN sur cartes perforées, est la première occurrence de ce que l'on appellerait aujourd'hui un assembleur de séquences de novo (*Figure VI.1*).



**Figure VI.1** : COMPROTEIN, le premier logiciel de bioinformatique [Jeff G. 2019]

COMPROTEIN ce logiciel, entièrement codé en FORTRAN sur cartes perforées, est la première occurrence de ce que l'on appellerait aujourd'hui un assembleur de séquences de novo.

(A) Un ordinateur central IBM 7090, pour lequel COMPROTEIN a été conçu pour fonctionner. (B) Une carte perforée contenant une ligne de code FORTRAN (le langage COMPROTEIN a été écrit avec). (C) Le code source d'un programme entier dans des cartes perforées. (D) Un aperçu simplifié de l'entrée de COMPROTEIN (c'est-à-dire des séquences peptidiques d'Edman) et de la sortie (une séquence de protéine consensus).

### VI.1.2. Changement de paradigme de l'analyse des protéines à l'analyse de l'ADN (1970-1980)

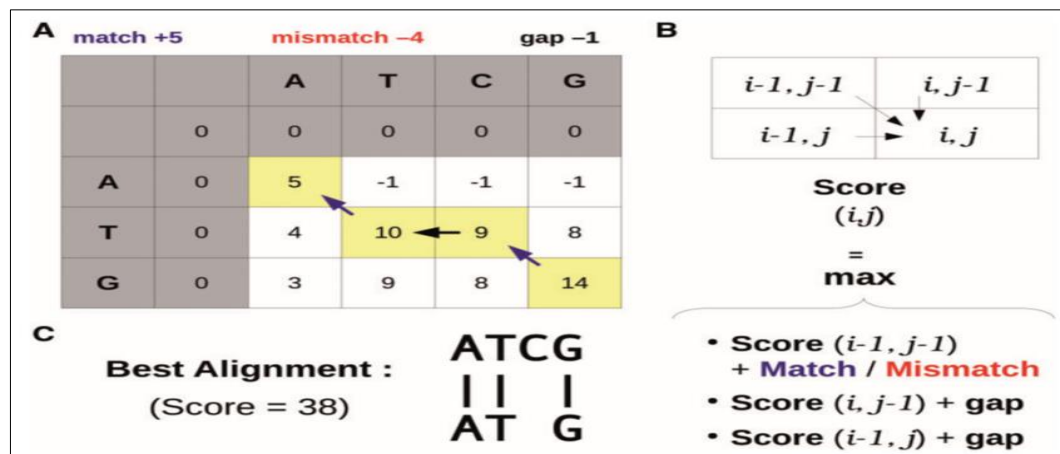
Cette période est caractérisé par :

- Le premier algorithme d'alignement de séquences multiples (MSA) publié était une généralisation de l'algorithme Needleman-Wunsch, qui impliquait l'utilisation d'une matrice de notation (1970). [Proc. N,1985]

- La première méthode de séquençage d'ADN à être largement adoptée : fut la méthode de séquençage Maxam-Gilbert en 1976. [Proc. N,1977]

- Le premier logiciel dédié à l'analyse de Sanger séquence a été publié par Roger Staden en 1979. [Nucleic Acids Res,1979]

- Un modèle de changement évolutif des protéines par Dayhoff et al.,. Dans : Atlas de la séquence et de la structure des protéines. [Fondation nationale de la recherche biomédicale, 1978].



**Figure. VI.2 :** Représentation de l'algorithme d'alignement global Needleman-Wunsch

(A) Un alignement optimal entre deux séquences est trouvé en trouvant le chemin optimal sur une matrice de notation calculée avec des points de correspondance et de non-concordance (ici +5 et -4), et une pénalité d'écart (ici -1). (B) Chaque cellule (i, j) de la matrice de notation est calculée avec une fonction

maximale, basée sur le score des cellules voisines. (C) Le meilleur alignement entre les séquences ATCG et ATG, en utilisant les paramètres mentionnés en (A).  
Remarque : aucune pénalité d'écart initial et final n'a été définie dans cet exemple.

[JeffG, 2019]

### VI.1.3. Avancées parallèles en biologie et en informatique (1980-1990)

Cette période spécifiée en manipulant l'ADN avec la réaction en chaîne par polymérase (PCR), qui permet d'amplifier l'ADN sans procédures de clonage [Mulis and Faloon, 1987]. En 1984, Genetics Computer Group (Wisconsin Univ.) a publié la suite logicielle éponyme « GCG ». Ce fut la première collection de logiciels développée pour l'analyse de séquences. [Nucleic Acids Res, 1984].

Une autre suite de manipulation de séquences a été développée après GCG la même année, le logiciel "DNASTAR", qui pouvait être exécuté sur PC. Dans les années 1984-1985, d'autres suites de manipulation de séquences ont été développées, pour fonctionner sur les ordinateurs Apple II et Macintosh. [Nucleic Acids Res 1984].



**Figure VI.3** : Une station de travail de bureau HP-9000 exécutant le système basé sur Unix HP-UX

Surtout, c'est durant cette période que les bases de données de séquences EMBL, GenBank (NCBI) et DDBJ (DNA Data Bank of Japan) se sont unies (1986-1987) afin de standardiser le formatage des données. Aujourd'hui, cette union existe toujours et est maintenant représentée par l'International Nucleotide Sequence Database Collaboration (INSDC) [<http://www.insdc.org/>]

### VI.1.4. Génomique et autoroute de l'information (1990-2000)

A cette époque, Tim Berners-Lee (CERN) a lancé le World Wide Web (www), au début des années 90. En outre, le Laboratoire européen de biologie moléculaire (EMBL) a été mis en ligne sur certaines bases de données telles que SWISS-PROT et

REBASE en 1993. En 1992, la base de données GenBank est devenue la responsabilité du Centre national d'information sur la biotechnologie (NCBI), avant qu'elle ne soit sous LANL [JeffG, 2019]

Le NCBI en 1994 a été rendu disponible en ligne (y compris **BLAST**), Genomes (1995), PubMed (1997) et Human Genome (1999).

En 1995, le premier séquençage complet du génome d'un organisme libre a été séquencé. Cependant, le tournant qui a commencé l'ère génomique, a été la publication du génome humain au début du 21<sup>ème</sup> siècle.

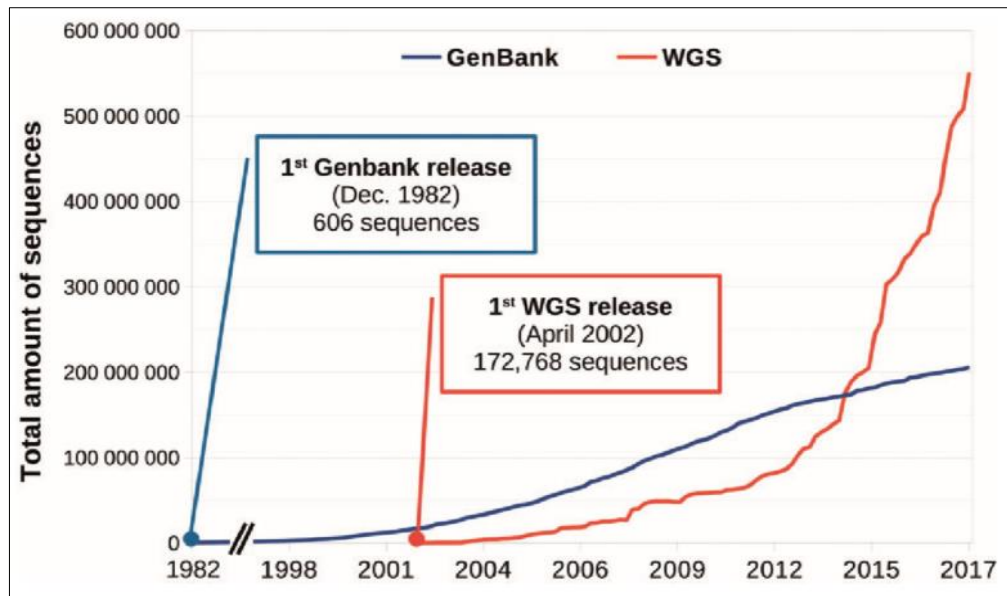


**Figure VI.4. :** Réseau de séquenceurs capillaires ABI 373 au NHGRI (1993)

### VI.1.5. Bioinformatique à haut débit (2000-2010)

Séquençage de deuxième génération, également appelé séquençage de nouvelle génération (NGS), qui a commencé avec la technologie de pyroséquençage « 454 ». Cette technologie a permis de séquencer des milliers à des millions de molécules d'ADN en une seule machine.

Depuis 2008, l'arrivée des technologies de séquençage massivement parallèle. Cela a entraîné une augmentation exponentielle des séquences dans les bases de données publiques et une nouvelle infrastructure de référentiel est apparue pour les organismes modèles[Mardis. E,2011].



**Figure VI.5.** : Quantité totale de séquences sur les bases de données NCBI GenBank et WGS au fil du temps.

Le nombre de génomes provisoires/incomplets a dépassé le nombre de séquences de génomes complètes en juin 2014, et le nombre de génomes continue de croître de façon exponentielle.

### VI.1.6. Perspectives présentes et futures (2010-Aujourd'hui)

A cette période, l'émergence de chercheurs spécialisés dans ce domaine de la Bioinformatique. De plus, il y a une augmentation significative de :

- (i) des outils conviviaux, souvent disponibles via des serveurs Web intégratifs comme Galaxy..., et
- (ii) aider les communautés telles que SEQanswers et BioStar

Il existe également un besoin explosif de bioinformaticiens sur le marché du travail dans les secteurs académique, privé et gouvernemental.

L'International Society for Computational Biology (ISCB) a publié des lignes directrices et des recommandations sur les compétences de base qu'un bioinformaticien devrait avoir dans son programme (utilisation, application et analyse) [ *PLoSComputBiol2014*].

### VI.2. Centre national d'information sur la biotechnologie (NCBI)

#### VI.2.1. introduction

Le National Center for Biotechnology Information (NCBI) fait partie de la National Library of Medicine (NLM) des États-Unis, une branche des National Institutes of Health (NIH). Il est approuvé et financé par le gouvernement des États-Unis. Le NCBI est situé à Bethesda, Maryland et a été fondé en 1988 par une législation parrainée par le sénateur Claude Pepper.

Le NCBI abrite une série de bases de données relatives à la biotechnologie et à la biomédecine et constitue une ressource importante pour les outils et services bioinformatiques. Les principales bases de données incluent GenBank pour les séquences d'ADN et PubMed, une base de données bibliographique pour la littérature biomédical.

D'autres bases de données incluent la base de données NCBI Epigenomics. Toutes ces bases de données sont disponibles en ligne via le moteur de recherche *Entrez*. [<https://www.ncbi.nlm.nih.gov/>]

Le NCBI était responsable de la mise à disposition de la base de données de séquences d'ADN *GenBank* depuis 1992. GenBank se coordonne avec des laboratoires individuels et d'autres bases de données de séquences telles que celles du Laboratoire européen de biologie moléculaire (EMBL) et de la Banque de données ADN du Japon (DDBJ).

#### VI.2.2. Ressources populaires de NCBI

##### PubMed

PubMed® comprend plus de 32 millions de citations pour la littérature biomédicale de MEDLINE, des revues de sciences de la vie et des livres en ligne. Les citations peuvent inclure des liens vers le contenu en texte intégral de PubMed Central et des sites Web des éditeurs.

##### Nucléotide

La base de données Nucleotide est une collection de séquences provenant de plusieurs sources, dont GenBank, RefSeq, TPA et PDB. Les données sur le génome, les gènes et les séquences de transcription constituent la base de la recherche et de la découverte biomédicales

### **Basic Local Alignment Search Tool (BLAST)**

BLAST trouve des régions de similarité locale entre les séquences. Le programme compare les séquences de nucléotides ou de protéines aux bases de données de séquences et calcule la signification statistique des correspondances. BLAST peut être utilisé pour déduire des relations fonctionnelles et évolutives entre les séquences ainsi que pour aider à identifier les membres des familles de gènes.

### **Protein**

La base de données de protéines est une collection de séquences provenant de plusieurs sources, y compris des traductions de régions codantes annotées dans GenBank, RefSeq et TPA, ainsi que des enregistrements de SwissProt, PIR, PRF et PDB. Les séquences protéiques sont les déterminants fondamentaux de la structure et de la fonction biologiques.

### **Genome**

Cette ressource organise les informations sur les génomes, y compris les séquences, les cartes, les chromosomes, les assemblages et les annotations.

### **PubMed Central® (PMC)**

Le PMC est une archive gratuite en texte intégral de la littérature de revues biomédicales et des sciences de la vie à la National Library of Medicine (NIH/NLM) des National Institutes of Health des États-Unis.

### **PubChem**

PubChem est la plus grande collection au monde d'informations chimiques librement accessibles. Recherchez des produits chimiques par nom, formule moléculaire, structure et autres identifiants. Trouvez des propriétés chimiques et physiques, des activités biologiques, des informations sur la sécurité et la toxicité, des brevets, des citations de la littérature et plus encore.

### **Gène**

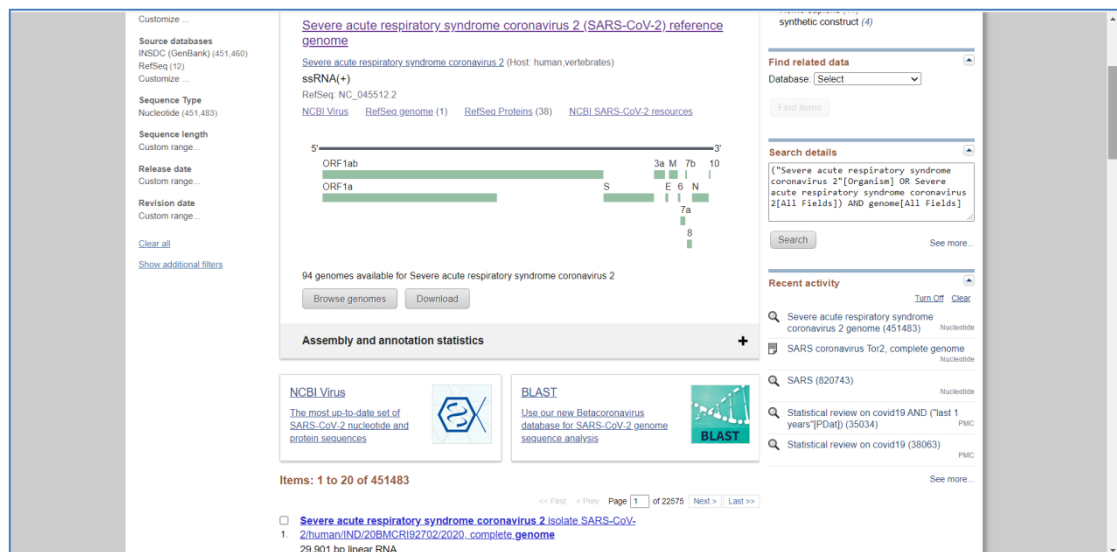
Gene intègre des informations provenant d'un large éventail d'espèces. Un enregistrement peut inclure une nomenclature, des séquences de référence (RefSeqs), des cartes, des voies, des variations, des phénotypes et des liens vers des ressources spécifiques au génome, au phénotype et au locus dans le monde entier.

### **VI.3. Base de données nucléotidiques**

Nous avons utilisé la base de données Nucleotide pour obtenir la séquence complète du génome de COVID-19 en utilisant des mots clés similaires tels que

## Chapitre VI Étude sur COVID-19 en utilisant la bioinformatique

covid19, SARS-CoV-2 ou génome du coronavirus 2 du syndrome respiratoire aigu sévère. La figure ci-dessous (**Figure VI.6**) montre le résultat obtenu concernant le Covid19 en utilisant la recherche dans cette base de données (14 juin 2021).



The screenshot displays the NCBI Nucleotide search results for 'Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) reference genome'. The interface includes a left sidebar with filters, a main search results area showing 94 genomes available, and a detailed view of the first result: '2|human|ND:20BMCRI82702|2020, complete genome'. The detailed view includes a protein domain diagram with ORF1a and ORF1ab, and a list of related search results.

**Figure VI.6:** Recherche sur Covid19 à partir de la base de données Nucleotide (14 juin 2021).

Toujours, la recherche dans la base de données Nucleotide fournit une séquence de requêtes dans trois formats différents : (1) GenBank, (2) FASTA et (3) Formats graphiques.

### VI.3.1. Format GenBank

GenBank est une base de données publique complète de séquences nucléotidiques et d'annotations bibliographiques et biologiques. GenBank est construit et distribué par le National Center for Biotechnology Information (NCBI), une division de la National Library of Medicine (NLM), située sur le campus des US National Institutes of Health (NIH) à Bethesda, MD, États-Unis [Benson.D,2010].

NCBI construit GenBank principalement à partir de la soumission de données de séquences par les auteurs et de la soumission en masse d'étiquettes de séquences exprimées (EST), de séquences d'enquête sur le génome (GSS) et d'autres données à haut débit provenant de centres de séquençage. L'Office des brevets et des marques des États-Unis fournit également des séquences de brevets délivrés. GenBank participe avec la base de données de séquences de nucléotides du Laboratoire européen de biologie moléculaire (EMBL) et la banque de données d'ADN du Japon (DDBJ) en tant que partenaire de l'International Nucleotide Sequence Database Collaboration (INSDC), qui échange des données quotidiennement pour garantir

## Chapitre VI Étude sur COVID-19 en utilisant la bioinformatique

qu'une collection uniforme et complète de les informations sur les séquences sont disponibles dans le monde entier. NCBI rend les données GenBank disponibles gratuitement sur Internet, via FTP et une large gamme de services de récupération et d'analyse basés sur le Web. [Sayers ,E 2010]. Le format GenBank de la séquence du « *génom de référence du coronavirus du syndrome respiratoire aigu sévère 2 (SARS-CoV-2)* » est (Figure VI.7) :

GenBank Send to: ▾

### Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome

NCBI Reference Sequence: NC\_045512.2  
[FASTA](#) [Graphics](#)

---

Go to: ☺

LOCUS NC\_045512 29903 bp ss-RNA linear VRL 18-JUL-2020

DEFINITION Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome.

ACCESSION NC\_045512

VERSION NC\_045512.2

DBLINK BioProject: [PRJNA485481](#)

KEYWORDS RefSeq.

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

ORGANISM [Severe acute respiratory syndrome coronavirus 2](#)  
Viruses; Riboviria; Orthornavirae; Pisuviricota; Pisoniviricetes;  
Nidovirales; Coronaviridae; Orthocoronavirinae;  
Betacoronavirus; Sarbecovirus.

REFERENCE 1 (bases 1 to 29903)

AUTHORS Wu, F., Zhao, S., Yu, B., Chen, Y. M., Wang, W., Song, Z. G., Hu, Y.,  
Tao, Z. W., Tian, J. H., Pei, Y. Y., Yuan, M. L., Zhang, Y. L., Dai, F. H.,  
Liu, Y., Wang, Q. M., Zheng, J. J., Xu, L., Holmes, E. C. and Zhang, Y. Z.

TITLE A new coronavirus associated with human respiratory disease in  
China

JOURNAL Nature 579 (7798), 265-269 (2020)

PUBMED [32015508](#)

REMARK Erratum:[Nature. 2020 Apr;580(7803):E7. PMID: 32296181]

REFERENCE 2 (bases 13476 to 13503)

AUTHORS Baranov, P. V., Henderson, C. M., Anderson, C. B., Gesteland, R. F.,  
Atkins, J. F. and Howard, M. T.

TITLE Programmed ribosomal frameshifting in decoding the SARS-CoV genome

JOURNAL Virology 332 (2), 498-510 (2005)

PUBMED [15680415](#)

REFERENCE 3 (bases 29728 to 29768)

AUTHORS Robertson, M. P., Igel, H., Baertsch, R., Haussler, D., Ares, M. Jr. and  
Scott, W. G.

TITLE The structure of a rigorously conserved RNA element within the SARS  
virus genome

JOURNAL PLoS Biol. 3 (1), e5 (2005)

PUBMED [15630477](#)

REFERENCE 4 (bases 29609 to 29657)

AUTHORS Williams, G. D., Chang, R. Y. and Brian, D. A.

TITLE A phylogenetically conserved hairpin-type 3' untranslated region  
pseudoknot functions in coronavirus RNA replication

JOURNAL J. Virol. 73 (10), 8349-8355 (1999)

PUBMED [10482585](#)

REFERENCE 5 (bases 1 to 29903)

CONSRM NCBI Genome Project

TITLE Direct Submission

JOURNAL Submitted (17-JAN-2020) National Center for Biotechnology  
Information, NIH, Bethesda, MD 20894, USA

REFERENCE 6 (bases 1 to 29903)

AUTHORS Wu, F., Zhao, S., Yu, B., Chen, Y. M., Wang, W., Hu, Y., Song, Z. G.,

**Figure VI.7** : Section Header au format GenBank du SARS-CoV-2 (14 juin 2021).

L'enregistrement GenBank comprend trois sections : (1) En-tête (Locus, Définition, Accession, Organisme (Taxonomie), Citations...etc.), (2) Tableau des caractéristiques (Codon de départ et d'arrêt, SNP, Séquence AA...etc. ) et (3) Origine (Séquence)[[ncbi.genbank](#)].

### VI.3.1.1. Section HEADER

Le champ LOCUS contient un certain nombre d'éléments de données différents, notamment le nom du locus, la longueur de la séquence, le type de molécule, la division GenBank et la date de modification.


Le nom du locus est NC\_045512, qui est à l'origine conçu pour aider à regrouper les entrées avec des séquences similaires.

La longueur de séquence est le nombre de paires de bases de nucléotides (ou de résidus d'acides aminés) dans l'enregistrement de séquence. Dans cet exemple, la longueur de la séquence est de 29903 pb.


Le type de molécule qui a été séquencée. Dans cet exemple, le type de molécule est ss-RNA. Chaque enregistrement GenBank doit contenir des données de séquences contiguës d'un seul type de molécule. Les divers types de molécules peuvent comprendre l'ADN génomique, l'ARN génomique, l'ARN précurseur, l'ARNm (ADNc), l'ARN ribosomique, l'ARN de transfert, le petit ARN nucléaire et le petit ARN cytoplasmique.

La division GenBank à laquelle appartient un enregistrement est indiquée par une abréviation à trois lettres. Dans cet exemple, la division GenBank est VRL, ce qui signifie séquences virales.

Date de modification : la date dans le champ LOCUS est la date de la dernière modification. L'exemple d'enregistrement présenté ici a été modifié pour la dernière fois le 18-JUL-2020.

 **DEFINITION** est une brève description de la séquence ; comprend des informations telles que l'organisme source, le nom du gène/nom de la protéine ou une description de la fonction de la séquence (si la séquence n'est pas codante). Si la séquence a une région codante (CDS), la description peut être suivie d'un qualificatif de complétude, tel que "cds complets".

Dans cet exemple la DEFINITION est « isolat de coronavirus 2 du syndrome respiratoire aigu sévère Wuhan-Hu-1, génome complet ».

 **ACCESSION** est l'identifiant unique d'un enregistrement de séquence. Un numéro d'accès s'applique à la notice complète et est généralement une combinaison d'une ou de plusieurs lettres et de chiffres, par exemple une seule lettre suivie de cinq chiffres ou deux lettres suivies de six chiffres. Certaines accessions peuvent être plus longues, selon le type d'enregistrement de séquence.

Le numéro d'accèsion dans cet exemple est NC\_045512. C'est-à-dire ne pas changer, même si les informations de la notice sont modifiées à la demande de l'auteur. Parfois, cependant, un numéro d'accèsion original peut devenir secondaire par rapport à un numéro d'accèsion plus récent, si les auteurs font une nouvelle soumission qui combine des séquences précédentes, ou si pour une raison quelconque une nouvelle soumission remplace un enregistrement antérieur.

- ✚ **RÉFÉRENCE** : Publications des auteurs de la séquence qui traitent des données rapportées dans l'enregistrement. Les références sont automatiquement triées dans la notice en fonction de la date de publication, en affichant en premier les références les plus anciennes.
- ✚ **PUBMED** : l'identifiant PubMed (PMID) est une référence qui inclut les identifiants PubMed contenant des liens entre l'enregistrement de séquence et l'enregistrement PubMed correspondant. Inversement, les notices PubMed qui contiennent le(s) numéro(s) d'accèsion dans le champ SI (identifiant de source secondaire) contiennent des liens vers la ou les notices de séquence.

```

FEATURES             Location/Qualifiers
     source            1..29903
                        /organism="Severe acute respiratory syndrome coronavirus
                        2"
                        /mol_type="genomic RNA"
                        /isolate="Wuhan-Hu-1"
                        /host="Homo sapiens"
                        /db_xref="taxon:2697049"
                        /country="China"
                        /collection_date="Dec-2019"
     5'UTR            1..265
     ..
     gene              266..21555
                        /gene="ORF1ab"
                        /locus_tag="GU280_gp01"
                        /db_xref="GeneID:43740578"
     CDS               join(266..13468,13468..21555)
                        /gene="ORF1ab"
                        /locus_tag="GU280_gp01"
                        /ribosomal_slippage
                        /note="pp1ab; translated by -1 ribosomal frameshift"
                        /codon_start=1
                        /product="ORF1ab polyprotein"
                        /protein_id="YP_009724389.1"
                        /db_xref="GeneID:43740578"
                        /translation="MESLVPGFNEKTHVQLSLPVLQVRDLVLRGFGDSVEEVLSEARQ
                        HLKDGTCGLVEVEKGVLPQLEQPYVFIKRS DARTAPHGHVMV LVAELEGIQYGRSGE
                        TLGLVLPVHVEIPVAYRKVLLRKNNGKAGGHSYGADLKSFDLGDDELGTPDYEDFQEN
                        WNTKHS SGTRELMRELNNGGAYTRYVDNFCGPDGYPLECIKDLLARAGKASCTLSEQ
                        LDFIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDTFNGECPNFWFP
                        LNSIIKTIQPRVEKKKLDGFMGRIRSVYPVAPSNENQMC LSTLMKCDHCGETSWQTF
                        ..
     mat_peptide       13442..13480
                        /gene="ORF1ab"
                        /locus_tag="GU280_gp01"
                        /product="nsp11"
                        /note="produced by pp1a only"
                        /protein_id="YP_009725312.1"
     stem_loop         13476..13503
                        /gene="ORF1ab"
                        /locus_tag="GU280_gp01"
                        /inference="COORDINATES:
                        profile:Rfam-release-14.1:RF00507,Infernal:1.1.2"
                        /function="Coronavirus frameshifting stimulation element
                        stem-loop 1"
     stem_loop         13488..13542
                        /gene="ORF1ab"
                        /locus_tag="GU280_gp01"
                        /inference="COORDINATES:
                        profile:Rfam-release-14.1:RF00507,Infernal:1.1.2"
                        /function="Coronavirus frameshifting stimulation element
                        stem-loop 2"
     gene              21563..25384
                        /gene="S"
                        /locus_tag="GU280_gp02"
                        /gene_synonym="spike glycoprotein"
                        /db_xref="GeneID:43740568"
     CDS               21563..25384
                        /gene="S"
                        /locus_tag="GU280_gp02"
                        /gene_synonym="spike glycoprotein"
                        /note="structural protein; spike protein"
                        /codon_start=1
                        /product="surface glycoprotein"
                        /protein_id="YP_009724390.1"
                        /db_xref="GeneID:43740568"
                        /translation="MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVVYDPDKVFR
                        SSVLHSTQDLFLPFFSNVTNFHAIHVS GTNGTKRFDNPVLPFNDGVYFAST EKSNIIR
                        GHIFGTTLD SKTQSL L I VNNATN VV I K V C E F Q C N D P F L G V Y Y H K N K S W H E S F R V Y
                        SSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFNIDGYFKIYSKHTPINLVRDL PQ
                        GFSALEPLVDLPIGINITRFQTL LALHRSYLP GDS SSGMTAGAAA YVGYLQPRFTL
                        LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITN
                        LCPFGEVFNATRFASVYAWNRKRISNVCADYSVLYNSASFSTFKCYGVSPTKLNDLCLF
                        TNVYADSFVIRGDEVRIAPGQTGKIADYNYKLPDDFTGCVIAWNSNLD SKVGGVNYN
                        YLYRFLFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCFYPLQSYGFQPTNGVGYQPY
                        RVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFG
                        RDIADTTDAVRDPQTL EILDITPCSFGGVSVITPGTNTSNQAVLVYQDVNCTEVPVAI
                        HADQLTPTRVYVSTGSNVFQTRAGCLIGA EHVNSY ECDIPIGAGICASYQTQNSPR
                        RARSVASQSI IAYTMSLGAENSVAYSNNSIAIPTNFTISVTT EILPVSMTKTSVDCM
                        YICGDS TECSNLL LQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFG
                        GFNFSQLPDPSPKSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDI AARDLICAKFN
                        GLTVLPPLLTD EMI AQYTSAL LAGTIITSGWTFGAGAA LQIPFAMQ MAYRFNGIGVTVN
                        VLYENQKLIANQFNSAIGKIQDLSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGA
                        ISSV LNDILSR LDKVEAEVQIDRLITGRLOS LQTYVTQQLIRAAEIRASANLAATKMS
                        ECVLGQSKRVDFCGKGYHLHSFPQSAPHGVVFLHVTYVPAQEKNTTAPAICHGDKAH
                        FPREGVFVSNGTWVFTQRNFYEPQIITTDNTFVSGNCDVVIGIVMNTVYDPLQPELD
                        SFKEELDKYFNHTSPDVLGDISGINASV VNIQKEIDRLNEVAKNLNESLIDLQELG
                        KYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCHTSCCSCLKGCCSCGSCCKFPEDDSE
                        PVLKGVK LHYT"
                        ..
ORIGIN
1 attaaagggt tataccttcc caggtaacaa accaaccaac tttcgatctc ttgtagatct
61 gttctctaaa cgaactttaa aatctgtgtg gctgtcactc ggctgcacgc ttagtgcaact
121 cagcgagtat aattaataac taattactgt cgttgacagg acacgagtaa ctcgctctac
181 tcttcgaggc tgcctacggt ttcgtccgtg ttgcagccga tcactcagcac atctaggttt
241 cgtccggggt tgaccgaaag gtaaatgga gagcctgtgc cctggtttca acgagaaaaa
301 acacgtccaa ctacagttgc ctgttttaca gggttcgcgac gtgctcgtac gggccttgg
361 agactccgtg gaggaggctc tatcagaggc acgtcaacat cttaaagatg gcacttgtgt

```

**Figure IV.8 :** Sections Caractéristiques et Origine au format GenBank du SARS-CoV-2 (14 juin 2021).

### VI.3.1.2. Section FEATURES

Informations sur les gènes et les produits géniques, ainsi que les régions d'importance biologique signalées dans la séquence. Ceux-ci peuvent inclure des régions de la séquence qui codent pour des protéines et des molécules d'ARN, ainsi qu'un certain nombre d'autres caractéristiques [**ncbi. genbank**].

**source** : élément obligatoire dans chaque enregistrement qui résume la longueur de la séquence, le nom scientifique de l'organisme source et le numéro d'identification du taxon. Peut également inclure d'autres informations telles que l'emplacement sur la carte, la souche, le clone, le type de tissu, etc., si fourni par le demandeur.

**CDS** : séquence codante ; région de nucléotides qui correspond à la séquence d'acides aminés dans une protéine (l'emplacement comprend les codons de départ et d'arrêt). La fonction CDS comprend une traduction d'acides aminés. Les auteurs peuvent spécifier la nature du CDS en utilisant le qualificatif `"/evidence=experimental"` ou `"/evidence=not_experimental"`.

Les auteurs sont également encouragés à annoter la caractéristique d'ARNm, qui comprend la région non traduite 5' (5'UTR), les séquences codantes (CDS, exon) et la région non traduite 3' (3'UTR).

**Translation**: la traduction des acides aminés correspondant à la séquence codante des nucléotides (CDS). Dans de nombreux cas, les traductions sont conceptuelles. Notez que les auteurs peuvent indiquer si la SDC est basée sur des preuves expérimentales ou non expérimentales.

**gène** : une région d'intérêt biologique identifiée comme un gène et pour laquelle un nom a été attribué. L'étendue de base de la caractéristique du gène dépend des caractéristiques 5' et 3' les plus éloignées. AF165912 et AF090832 sont des exemples supplémentaires d'enregistrements qui montrent la relation entre les caractéristiques des gènes et d'autres caractéristiques telles que l'ARNm et le CDS.

**Complement** : Indique que la caractéristique est située sur le brin complémentaire.

**Other Features**: exemples d'autres enregistrements qui montrent une variété de caractéristiques biologiques ; un format graphique est également disponible pour chaque enregistrement de séquence et représente visuellement les caractéristiques annotées :

AF165912 (gène, promoteur, signal TATA, ARNm, 5'UTR, CDS, 3'UTR)

AF090832 (liaison aux protéines, gène, 5'UTR, ARNm, CDS, 3'UTR)

L00727 (ARNm épissés alternativement)

### VI.3.1.3. Partie ORIGINE

L'ORIGINE peut être laissée en blanc, peut apparaître comme « Non signalé » ou peut donner un pointeur local vers le début de la séquence, impliquant généralement un site de clivage de restriction déterminé expérimentalement ou le locus génétique (si disponible). Cette information n'est présente que dans les enregistrements plus anciens. Les données de séquence commencent sur la ligne immédiatement en dessous d'ORIGIN [**ncbi. genbank**].

### VI.3.2. Format FASTA

En bioinformatique et en biochimie, le format FASTA est un format textuel permettant de représenter des séquences de nucléotides ou des séquences d'acides aminés, dans lesquelles les nucléotides ou les acides aminés sont représentés à l'aide de codes à une lettre. Le format permet également aux noms de séquences et aux commentaires de précéder les séquences.

Une séquence au format FASTA commence par une description sur une seule ligne, suivie de lignes de données de séquence. La ligne de description (defline) se distingue des données de séquence par un symbole supérieur à (">") au début. Il est recommandé que toutes les lignes de texte comportent moins de 80 caractères [**FASTA**].

La séquence du « *génomome de référence du coronavirus du syndrome respiratoire aigu sévère 2 (SARS-CoV-2)* » au format FASTA est (**Figure IV.9**) :

```
FASTA Send to: ▾  
  
Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome  
NCBI Reference Sequence: NC_045512.2  
GenBank Graphics  
>NC_045512.2 Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome  
ATTTAAAGGTTTATACCTTCCAGGTAACAACAACCAACCTTTCGATCTCTTGATAGCTGTTCTCTAAA  
CGAAGCTTTAAATCTGTGTGGCTGTCACCTCGGCTGCATGCTTAGTGCACCTACGCAAGTATAAATTAATAAC|  
TAATTACTGTCTGTGACAGGACACAGTAACCTCGTCTATCTTCTGCAAGGCTGCTTACGGTTTCGTCGGTG  
TTGCAGCCGATCATCAGCACATCTAGGTTTCGTCGGGTGTGACCGAAAAGTAAAGATGGAGAGCCTTTGTC  
CCTGGTTTCAACGAGAAAACACACGTCCTCAACTCAGTTTGCCTGTTTACAGGTTTCGCGACGCTCGTAC  
GTGGCTTTGGAGACTCCGTGGAGGAGTCTTATCAGAGGCACGTCAACATCTTAAAGATGGCACTTGTGG  
CTTAGTAGAAGTTGAAAAGGCGTTTGGCTCAACTTGAACAGCCCTATGTGTTTCAACAACTTCCGAT  
GCTCGAACTGCACCTATGGTCAATGTTATGGTTGAGCTGGTAGCAGAAGCTCGAAGGCATTCAGTACGGTC  
GTAGTGTGTAGACACTTGGTGTCTTGTCCCTCATGTGGGCGAAATACCAAGTGGCTTACCGCAAGGTTCT  
TCTTCGTAAGAAGCGTAATAAAGGAGCTGGTGGCATAAGTTACGGCGCCGATCTAAAGTCATTTGACTTA  
GGCAGCAGCTTGGCAGTGCCTTATGAAGATTTTCAAGAAAACGGAACACATAAAGATAGCAGTGGTG  
TTACCCGTGAACTCATCGCTGAGCTTAAACGGAGGGGCATACACTCGCTATGTGATAACAACCTTGTGG  
CCCTGATGGCTACCTCTTGGATGCAATTAAGACCTTACGACGCTGCTGGTAAAGCTTTCATGCACCTTTG  
TCCGAAACACTGGACTTATTGACACTAAGAGGGGTGATACTGTGCGTGAACATGAGCATGAAATTTG  
CTTGGTACCGGAACGTTCTGAAAAGAGCTATGAATTCAGACACCTTTTGAATTAATTTGGCAAAAGAA  
ATTTGACACCTTCAATGGGGAATGTCCTAAATTTTGATTTCCCTTAAATCCATAATCAAGACTATTCAA  
CCAAGGTTTGAAGAAAAGCTTGTAGGCTTTATGGGTAGAAATTCGATCTGTCTATCCAGTTGCGTCAAC  
CAAATGAATGCAACCAATGTCCTTCAACTCTCATGAAGTGTGATCATTGTGGTGAAGCTTCAATGGCA  
GACGGGCGATTTTGTAAAGCCACTTGCGAATTTGTGGCAGTGAAGATTTGACTAAAGAAAGTGCCTACT  
ACTTGTGGTTACTTACCCAAAATGCTGTTGTTAAATTTTATGTCCAGCATGTCACAATTCAGAAGTAG  
GACCTGAGCATAGCTTGGCGAATACCATAAATGAATCTGGCTTGAACCACTTCTCGTAAGGGTGGTCCG  
CACTATTGCTTGGAGGCTGTGTCTCTTATGTTGGTTCGCAATAACAAGTGGCTTATGGGTTCCA  
CGTGTAGGCTTAACATAGGTTGTAACCATACAGGTGTTGTTGGAGAAGGTTCCGAAGGCTTAATGACA  
ACCTTCTTGAATACTCAAAAAGAGAAAAGTCAACATCAATATTGTTGGTGACTTTAAACTTAAATGAAGA  
GATCGCCATTATTTGGCATCTTTTCTGCTTCCACAAGTCTTTTGTGGAAGTGTGAAAGGTTTGGAT  
TATAAAGCATTCAAAACAAATGTTGAATCCTGTGGTAAATTTAAAGTTACAAAAGGAAAAGCTAAAAAAG  
GTGCTGGAATATTGGTGAACAGAAAATCAATACTGAGTCTCTTATGCAATTTGCATCAGAGGCTGCTCG  
GTTGTACGATCAATTTTCTCCGCACCTTGAACCTGCTCAAAATCTGTGGCTGTTTACAGAAGGCC  
GCTATAACAATACTAGATGGAATTTACAGTATTACTGAGACTCATTGATGCTATGATGTTTACATCTG  
ATTTGGCTACTAACAATCTAGTTGTAATGGCTTACATTACAGGTTGGTGTGTTGAGTGCATTCGCAGTG  
GCTAACTAACATCTTTGGCACTGTTTATGAAAACCTCAACCCGCTTGAATGGCTTGAAGAGAAAGTTT  
AAGGAAGGTTAGAGTTTCTTAGAGACGGTTGGGAAATTTGTAATTTATCTCAACCTGTGCTTGTGAAA  
TTGTCGGTGGCAAAATTTGTCACCTGTGCAAAAGAAAATTAAGGAGAGTGTTCAGACATCTTTAAGCTTGT
```

**Figure IV.9** : Format FASTA du SARS-CoV-2 (14 juin 2021).

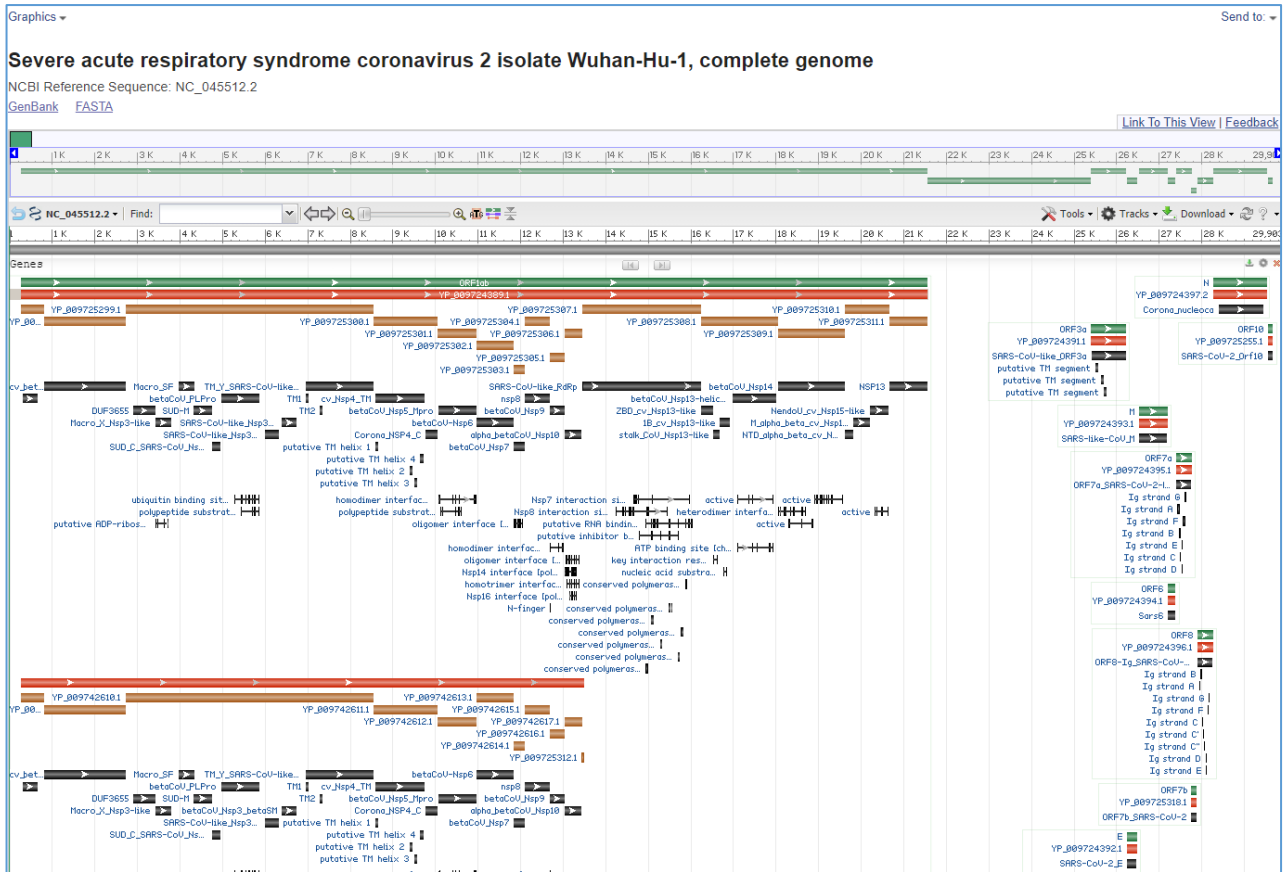
Les lignes vides ne sont pas autorisées au milieu de l'entrée FASTA. Les séquences doivent être représentées dans les codes d'acides aminés et d'acides nucléiques IUB/IUPAC standard, avec ces exceptions : les lettres minuscules sont acceptées et sont mappées en majuscules ; un seul tiret ou tiret peut être utilisé pour représenter un espace de longueur indéterminée ; et dans les séquences d'acides aminés, U et (sont des lettres acceptables (voir ci-dessous). Avant de soumettre une demande, tous les chiffres numériques dans la séquence de requête doivent être soit supprimés, soit remplacés par des codes de lettres appropriés (par exemple, N pour un résidu d'acide nucléique inconnu ou X pour un résidu d'acide aminé inconnu) [FASTA.]

### VI.3.3. Format Graphic (Graphical Sequence Panel)

Le panneau de séquence graphique est le composant central du visualiseur de séquence graphique. Il est composé de pistes contenant des données de séquence biologique annotées ou alignées sur une séquence d'acide nucléique ou de protéine. La règle en haut du panneau indique les coordonnées sur la séquence affichée.

Les pistes de données peuvent être supprimées en cliquant sur le bouton x rouge à l'extrémité droite de chaque piste. L'ordre des pistes peut être réorganisé par glisser-déposer.

# Chapitre VI Étude sur COVID-19 en utilisant la bioinformatique



**Figure IV.10 :** Format graphique du SARS-CoV-2 (14 juin 2021).

L'icône d'engrenage sur le côté droit de chaque piste vous permet d'accéder aux paramètres d'affichage de piste pour cette piste. De nouvelles pistes peuvent être ajoutées dans le panneau Configuration des pistes, accessible via l'icône d'engrenage sur les barres d'outils supérieure ou inférieure.

L'icône de téléchargement sur le côté droit de la piste vous permet d'accéder à l'option de téléchargement de piste pour cette piste. Cette icône n'est visible que pour les pistes téléchargeables depuis le visualiseur graphique[ncbi.nlm].

# **Chapter VII**

**Une étude comparative du  
SARS-CoV-2 versus SARS-  
CoV et Bat-CoV ; RaTG13  
avec l'outil BLASTN**

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

---

Dans ce chapitre, nous présentons un examen rapide de l'interface Web NCBI BLAST, puis présentons l'alignement de séquences effectué entre l'ensemble du génome du SARS-CoV-2 par rapport au SARS-CoV et au Bat-CoV ; RaTG13 utilisant le nucléotide-nucléotide BLAST (blastn). En outre, nous incluons certaines des études importantes trouvées dans la littérature publiée en 2020 et 2021, qui concernent le SARS-CoV-2, le SARS-CoV et le Bat-CoV ; virus RaTG13.

### VII.1. Introduction

En bioinformatique, BLAST (outil de recherche d'alignement local de base) est un algorithme et un programme de comparaison d'informations de séquences biologiques primaires, telles que les séquences d'acides aminés de protéines ou les nucléotides de séquences d'ADN et/ou d'ARN. Une recherche BLAST permet à un chercheur de comparer une protéine ou une séquence de nucléotides (appelée requête) avec une bibliothèque ou une base de données de séquences, et d'identifier les séquences de la base de données qui ressemblent à la séquence de requête au-dessus d'un certain seuil. Par exemple, suite à la découverte d'un gène auparavant inconnu chez la souris, un scientifique effectuera généralement une recherche BLAST du génome humain pour voir si les humains portent un gène similaire ; BLAST identifiera les séquences du génome humain qui ressemblent au gène de la souris sur la base de la similitude de séquence [WikiBLAST].

BLAST est en fait une famille de programmes (tous inclus dans l'exécutable blastall). Ceux-ci inclus:

✚ **Nucléotide-nucléotide BLAST (blastn)** : Ce programme, étant donné une requête ADN, renvoie les séquences ADN les plus similaires de la base de données ADN que l'utilisateur spécifie.

✚ **Protéine-protéine BLAST (blastp)** : Ce programme, étant donné une requête de protéine, renvoie les séquences de protéines les plus similaires de la base de données de protéines que l'utilisateur spécifie.

✚ **Protéine de traduction nucléotidique à 6 cadres (blastx)**: Ce programme compare les produits de traduction conceptuelle à six cadres d'une séquence d'interrogation de nucléotides (les deux brins) à une base de données de séquences de protéines.

✚ **Traduction de nucléotide à 6 cadres-traduction de nucléotide à 6 cadres (tblastx)**: Ce programme est le plus lent de la famille BLAST. Il traduit la séquence nucléotidique de la requête dans les six cadres possibles et la compare aux traductions à

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

six cadres d'une base de données de séquences nucléotidiques. Le but de tblastx est de trouver des relations très éloignées entre les séquences nucléotidiques.

✚ **Traduction protéine-nucléotide à 6 cadres (tblastn)** : ce programme compare une requête de protéine aux six cadres de lecture d'une base de données de séquences nucléotidiques [WikiBiotech].

BLAST est populaire car il permet d'identifier rapidement les régions de similarité locale entre deux séquences. Plus important encore, BLAST utilise un cadre statistique robuste qui peut déterminer si l'alignement entre deux séquences est statistiquement significatif. Dans ce didacticiel, nous utiliserons l'interface Web BLAST du National Center for Biotechnology Information (NCBI) pour nous aider à aligner la séquence SARS-CoV-2 par rapport à SARS-CoV et Bat-CoV ; RaTG13.

### VII.2. L'interface Web NCBI BLAST

Avant de commencer à analyser une séquence inconnue, nous devons d'abord nous familiariser avec l'interface Web NCBI BLAST. Ouvrez une nouvelle fenêtre de navigateur Web et accédez à la page principale de BLAST à l'adresse <http://blast.ncbi.nlm.nih.gov/Blast.cgi>. Dans ce tutoriel, nous n'utiliserons que quelques-uns des outils disponibles. Si vous souhaitez en savoir plus sur les options avancées disponibles (telles que Mes comptes NCBI) sur l'interface BLAST, cliquez à tout moment sur le bouton « Aide » en haut de la page (**Figure VII.1**).

COVID-19 is an emerging, rapidly evolving situation.  
Get the latest public health information from CDC: <https://www.cdc.gov/covid>  
Get the latest research from NIH: <https://www.nih.gov/coronavirus>

**Basic Local Alignment Search Tool**  
BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

**Search Betacoronavirus Database**  
We have created a new BLAST database focused on the SARS-CoV-2 (Severe acute respiratory syndrome coronavirus 2) Sequences. For further detail please visit [NCBI GenBank](#).  
Mon, 03 Feb 2020 10:00:00 EST [More BLAST news...](#)

**Web BLAST**

**Nucleotide BLAST**  
nucleotide ► nucleotide

**blastx**  
translated nucleotide ► protein

**tblastn**  
protein ► translated nucleotide

**Protein BLAST**  
protein ► protein

**BLAST Genomes**  
Enter organism common name, scientific name, or tax id

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

**Figure VII.1** : Les différents programmes BLAST disponibles via la page d'accueil du serveur Web NCBI.

Quatre des cinq programmes BLAST courants sont disponibles dans la section « Web BLAST » de la page d'accueil NCBI BLAST (**Figure VII.2**). Le programme tblastx, qui traduit la requête de nucléotides et la base de données de nucléotides lorsqu'il effectue les comparaisons de séquences, n'est pas répertorié dans la section « Web BLAST ».

The screenshot shows the NCBI BLASTN web interface. At the top, there is a navigation bar with 'BLAST' and 'blastn suite' links, and a search bar. Below this is a red banner with COVID-19 information. The main content area is titled 'Align Sequences Nucleotide BLAST'. It contains two main sections: 'Enter Query Sequence' and 'Enter Subject Sequence'. The 'Enter Query Sequence' section has a text input field containing 'U\_845112.2', a 'Clear' button, and a 'Query subrange' section with 'From' and 'To' fields. Below this is an 'Or, upload file' section with a 'Choose File' button and a 'Job Title' field containing 'BLAST\_COVID19\_vs\_SARS'. The 'Enter Subject Sequence' section has a text input field containing 'U\_894718.3', a 'Clear' button, and a 'Subject subrange' section with 'From' and 'To' fields. Below this is a 'Program Selection' section with radio buttons for 'Highly similar sequences (megablast)', 'More dissimilar sequences (discontiguous megablast)', and 'Somewhat similar sequences (blastn)'. At the bottom, there is a 'BLAST' button and a 'Search nucleotide sequence using Megablast (Optimize for highly similar sequences)' option. A 'Show results in a new window' checkbox is also present.

**Figure VII.2** : Mise en place de notre recherche BLASTN de la séquence SARS-CoV-2 contre le SARS-CoV et Bat-CoV-RaTG13

Cependant, nous pouvons accéder à ce programme en cliquant sur l'un des programmes BLAST dans la section « Web BLAST », puis sur l'onglet « tblastx » dans le formulaire de recherche NCBI BLAST. W peut également aligner deux (ou plus) séquences à l'aide du service 'blast 2 sequences' (bl2seq) dans la section 'Specialized BLAST' de la page principale NCBI BLAST( **Figure VII.2**).

Remarque : la recherche BLAST (BLASTN, BLASTP..etc.) peut prendre quelques minutes lorsque le serveur Web NCBI est occupé. Une fois la recherche terminée, une nouvelle page Web apparaîtra avec le rapport BLAST.

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

### VII.3. Compararison entre SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

Dans cette partie nous avons exécutée l'alignement entre le sequence du genome complete pour le virus SARS-CoV-2 contre les sequences des genomes complètes des virus SARS-CoV et Bat-CoV ; RaTG13. Cet alignement a été effectuée avec l'utilisation du Nucleotide-nucleotide BLAST (blastn). Où BLASTN étant donné une requête ADN, renvoie les séquences ADN les plus similaires de la base de données ADN que l'utilisateur spécifie. Le resultat de l'alignement a été obtenue sous la forme de quatre parties essentiel qui sont 'Description' Graphic Summary', Alignements' et 'Dot Plot' (voire la figure VII.3).

#### VII.3.1. Descriptions

Cet onglet affiche la liste des séquences de la base de données qui ont une homologie de séquence significative avec notre séquence (**Figure VII.3**). Par défaut, les résultats sont triés par leur valeur E dans l'ordre croissant, où les valeurs E inférieures indiquent des résultats plus significatifs. Nous pouvons cliquer sur les entêtes des colonnes pour trier les résultats par les autres colonnes. Nous pouvons également utiliser le menu déroulant « Select columns» sur la barre d'outils principale pour afficher ou masquer chaque colonne.

The screenshot shows the BLASTN results interface. At the top, there are navigation links: '< Edit Search', 'Save Search', and 'Search Summary'. On the right, there are links for 'How to read this report?', 'BLAST Help Videos', and 'Back to Traditional Results Page'. The main content area is divided into two sections: 'Job Title' and 'Filter Results'. The 'Job Title' section shows 'COVID19 vs SARS' and various search parameters like RID, Program, Query ID, Query Descr, Query Length, Subject ID, Subject Descr, Subject Length, and Other reports. The 'Filter Results' section has input fields for 'Percent Identity', 'E value', and 'Query Coverage', along with 'Filter' and 'Reset' buttons. Below these sections, there are tabs for 'Descriptions', 'Graphic Summary', 'Alignments', and 'Dot Plot'. The 'Descriptions' tab is active, showing a table of sequences producing significant alignments. The table has columns for 'Description', 'Scientific Name', 'Max Score', 'Total Score', 'Query Cover', 'E value', 'Per. Ident', 'Acc. Len', and 'Accession'. One sequence is listed: 'SARS coronavirus Tor2, complete genome' with a scientific name of 'SARS coronavirus Tor2' and an accession number of 'NC\_004718.3'. The table also shows a 'Download' button, a 'Select columns' dropdown menu, and a 'Show' dropdown menu set to '100'.

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> SARS coronavirus Tor2, complete genome	SARS coronavirus Tor2	15175	22568	88%	0.0	82.30%	29751	NC_004718.3

(a)

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

Job Title: COVID19 vs Bat-CoV; RaTG13

RID: CH42RM8811N Search expires on 06-16 19:34 pm Download All

Program: Blast 2 sequences Citation

Query ID: NC\_045512.2 (nucleic acid)

Query Descr: Severe acute respiratory syndrome coronavirus 2 isolate V...

Query Length: 29903

Subject ID: MN996532.2 (nucleic acid)

Subject Descr: Bat coronavirus RaTG13, complete genome

Subject Length: 29855

Other reports: MSA viewer

Filter Results: Percent Identity, E value, Query Coverage

Sequences producing significant alignments

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Bat coronavirus RaTG13, complete genome	Bat coronavirus RaTG13	48758	48758	99%	0.0	96.14%	29855	MN996532.2

(b)

**Figure VII.3 :** Les descriptions statistiques font partie de BLASTN entre le SARS-CoV-2 en tant que séquence de requête et la séquence de sujet pour (a) SARS-CoV et (b) Bat-CoV ; RaTG13.

D'après la liste des résultats BLASTN, nous pouvons noter que tous les alignements sont cohérents avec notre séquence de requête (SARS-CoV-2): tous les blocs d'alignement présentent un degré élevé d'homologie de séquence, une identité de séquence au niveau nucléotidique égal à **82,3%** pour le SRAS- CoV et **96,14 %** pour le Bat-CoV ; RaTG13.

En cliquant sur le numéro d'accèsion dans le tableau, une nouvelle page apparaîtra avec l'enregistrement GenBank de la séquence. Cliquer sur la description du hit nous amènera à l'alignement correspondant dans la sortie BLAST. Alternativement, vous pouvez cliquer sur l'onglet « Alignements » pour passer au premier alignement.

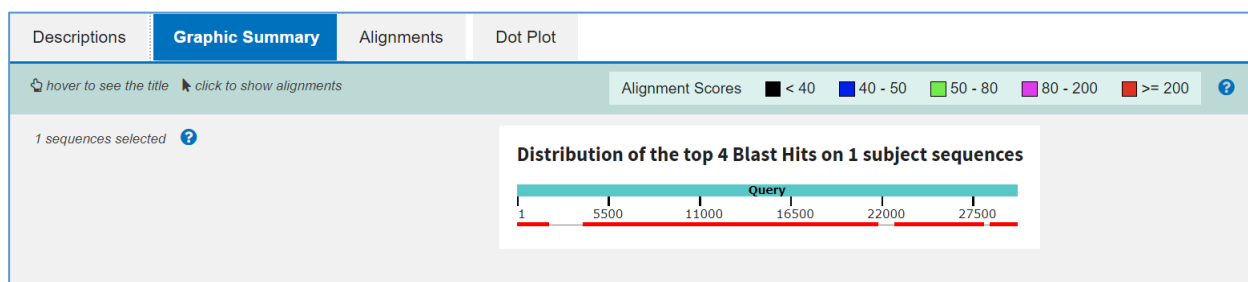
En plus d'examiner les enregistrements pour des séquences individuelles, vous pouvez également examiner plusieurs enregistrements de séquence en cochant la case à côté de chaque correspondance. Le contenu des autres onglets sera mis à jour automatiquement en fonction de votre sélection. Vous pouvez utiliser le menu déroulant "Download" sur la barre d'outils principale pour télécharger les résultats sélectionnés dans plusieurs formats (par exemple, FASTA, GenBank, Hit Table).

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

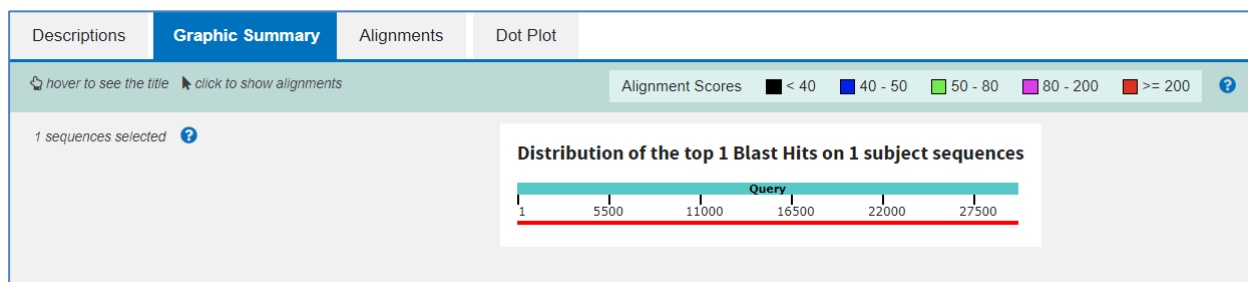
### VII.4. Graphic Summary

Le résumé graphique montre les alignements (sous forme de cases colorées) des correspondances de la base de données avec notre séquence de requête (barre rouge pleine sous la clé de couleur). Comme son nom l'indique, BLAST est conçu pour identifier des régions locales de similarité de séquence.

Cela signifie que BLAST peut signaler plusieurs régions discrètes de similarité de séquence entre une séquence de requête et une séquence de sujet dans une base de données. Par exemple, si une séquence d'ARNm épissé (mature) est alignée sur la séquence génomique inconnue, nous nous attendrions à voir plusieurs blocs d'alignement (dont beaucoup correspondent probablement à des exons transcrits) dans notre sortie BLAST. Les régions de la séquence génomique sans alignement significatif qui se situent entre ces exons sont susceptibles d'être des introns.



(a)



(b)

**Figure VII.4 :** Partie récapitulative graphique de BLASTN entre le SARS-CoV-2 en tant que séquence de requête et la séquence de sujet pour (a) SRAS-CoV et (b) Bat-CoV ; RaTG13.

La couleur des cases correspond au score (S) de l'alignement, le rouge représentant les scores d'alignement les plus élevés. Généralement, plus le score d'alignement est élevé, plus le coup est important. Lorsque vous déplacez votre souris sur un coup BLAST, la définition et le score du coup seront affichés dans la zone de

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

---

texte au-dessus de l'aperçu graphique. Lorsque vous cliquez sur une case, vous passez directement à l'alignement d'ADN réel associé à ce coup BLAST [communityedu].

Les résultats d'alignement des séquences locales de BLASTN montrent que le SARS-CoV correspond au SARS-CoV-2 dans quatre parties (blocs) le long des deux séquences (de SARS-CoV-2 et SARS-CoV) totalisant **88%** de la séquence de requête, avec la deuxième parcelle (bloc) étant le plus long des quatre. En revanche, les résultats montrent que Bat-CoV ; RaTG13 correspond au SARS-CoV-2 dans un segment unitaire qui couvre la plupart des deux chaînes, représentant **99%** de la séquence de requête. Ainsi, on peut dire que Bat-CoV ; RaTG13 est plus similaire au SARS-CoV-2 en termes de série de bases azotées (basées sur BLASTN).

Dans ce cas, basé sur le programme BLASTN (base de données), nous remarquons que le Bat-CoV ; La séquence RaTG13 correspond mieux à notre séquence de SARS-CoV-2 qu'à la séquence SARS-CoV. Nous voyons également que Bat-CoV ; Les correspondances de séquences RaTG13 couvrent presque toute la longueur de notre séquence de requête.

### VII.5. Alignements

Cet onglet contient les alignements entre les hits BLAST sélectionnés dans l'onglet Descriptions et la séquence de requête. Les alignements de séquences nous montrent à quel point notre séquence de requête correspond à la séquence de sujet dans la base de données.

Les alignements sur différentes séquences de sujets dans la base de données sont séparés par une barre d'outils bleue qui contient des options pour manipuler les résultats d'alignement et pour récupérer des informations supplémentaires pour ce coup BLAST spécifique (**Figure VII.5 a et b**). Par exemple, nous pouvons utiliser le menu déroulant "Télécharger" de cette barre d'outils pour obtenir la séquence FASTA ou l'enregistrement GenBank pour un hit spécifique. Nous pouvons utiliser les liens de navigation sur le côté droit de la barre d'outils pour naviguer rapidement vers le hit BLAST suivant ou précédent.

De plus, nous pouvons cliquer sur le lien « Graphics » pour examiner l'emplacement de chaque bloc d'alignement par rapport à la séquence du sujet.

Chaque bloc d'alignement commence par un résumé qui inclut le score Max et la valeur attendue (la signification statistique de l'alignement), l'identité de la séquence (nombre de bases identiques entre la requête et la séquence sujet), le nombre

## **Chapter VII                    Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN**

---

de lacunes dans l'alignement et le orientation de la séquence de requête par rapport à la séquence sujet. L'alignement se compose de trois lignes : la séquence de requête, la séquence de correspondance et la séquence de sujet (**Figure VII.5 a et b**).

La séquence de correspondance consiste en une combinaison de caractères « | » et d'espaces vides. Le caractère « | » désigne des bases identiques entre la requête et les séquences sujet. Un espace vide dans la séquence d'appariement indique une base non appariée. Le caractère « - » dans la requête ou la séquence d'objets indique un écart dans l'alignement (**Figure VII.5 a et b**).

# Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

Descriptions Graphic Summary **Alignments** Dot Plot

Alignment view Pairwise  CDS feature  Restore defaults

1 sequences selected

Download GenBank Graphics sort by: E value

**SARS coronavirus Tor2, complete genome**  
 Sequence ID: [NC\\_004718.3](#) Length: 29751 Number of Matches: 4  
[See 1 more title\(s\)](#) [See all Identical Proteins\(IPG\)](#)

Range 1: 3883 to 21505 [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
15175 bits(8217)	0.0	14581/17716(82%)	187/17716(1%)	Plus/Plus

Query 3956 AAAATCAAAGCTTGTGTGAAGAGTTACAACAACCTCTGGAAGAAAC TAAGTTCTCTACA 4015  
 Sbjct 3883 AAAATTAAGGCTTGCAATGATGAGGTTACCAACAACCTGGAAGAAAC TAAGTTCTTACC 3942

Query 4016 GAAAACCTGTACT-TTATAT--TGACATTAATGGCAATCTTCATCCA-GATTCTGCCAC 4071  
 Sbjct 3943 AATAA--GTTACTCTGTTTCTGATATCAATGGTAAGCTTTA-CCATGATCT-CAGA 3997

Query 4072 TCTTG-TT--A-GTGACATTGACATCACCTTCTAAAGAAAGATGCTCCATATATAGTGG 4127  
 Sbjct 3998 ACATGCTTAGAGGTGA-A-GATATGCTTCTTCTGAGAAGGATGCACCTTACATGGTAG 4054

.....

Query 21502 ATTAGAGAAAACAAGAGTTGTTATTTCTAGTGATGTTCTGTTAAACAACATAACGAAC 21561  
 Sbjct 21432 ATTAGAGAAAACAAGAGTTGTTGTTCAAGTGATATCTTGTAAACAACATAACGAAC 21491

Query 21562 AATGTTGTTTTTCTT 21577  
 Sbjct 21492 A-TGTTATTTT-CTT 21505

.....

Range 2: 22429 to 27798 [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#) [First Match](#)

Score	Expect	Identities	Gaps	Strand
3930 bits(2128)	0.0	4336/5417(80%)	92/5417(1%)	Plus/Plus

Query 22539 TTGTTAGATTTCCTAATAATTACAACAACCTTGGCCCTTTTGGTGAAGTTTTTAAAGCCACCA 22598  
 Sbjct 22429 TTGTGAGATTCCCTAATAATTACAACAACCTTGTGCTCTTTGGAGAGTTTTTAAATGCTACTA 22488

Query 22599 GATTGTCATCTGTTTATGCTTGGAAACAGGAAGAAATCAGCAACTGTGTTGCTGATTATT 22658  
 Sbjct 22489 AATTCCCTCTGCTATGATGATGGAGAGAAAAAATTTCAATGTTGTTGCTGATTACT 22548

.....

Query 27841 GGTTCCTCACTTGAACCTGCAAGATC-ATAATGAAACTTGT-C-AC-GCCTAAACGAACATG 27896  
 Sbjct 27723 GGTTTTCACTCGAAATCCAGGATCTAGAA-GAACCTTGATACCAAGCTTAAACGAACATG 27781

Query 27897 AAA-TT-Tc-TTGTTTT 27910  
 Sbjct 27782 AAACCTTCTATTGTTTT 27798

.....

Range 3: 28103 to 29751 [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#) [First Match](#)

Score	Expect	Identities	Gaps	Strand
2137 bits(1157)	0.0	1491/1653(90%)	19/1653(1%)	Plus/Plus

Query 28257 TAAACGAACAACATAAAATGCTGTAATGGACCCCA--AAATCAGCGAAATGACCCCC 28313  
 Sbjct 28103 TAAACGAACAATAAAATGCTGTAATGGACCCCAATCAACCAACGATAGTCCCCCC 28162

Query 28314 GCATTACGTTTGGTGGACCTCAGATTCACTGGCAGTAAACAGAAATGGAGAACGCAGTG 28373  
 Sbjct 28163 GCATTACATTTGGTGGACCCACAGATTCACTGACAAATAACAGAAATGGAGAACGCAGTG 28222

.....

Query 29862 GAGAATGACAAAAAAAAAAAAAAAAAAAAAAAA 29894  
 Sbjct 29719 GAGAATGACAAAAAAAAAAAAAAAAAAAAAAAA 29751

.....

Range 4: 6 to 1922 [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#) [First Match](#)

Score	Expect	Identities	Gaps	Strand
1325 bits(717)	0.0	1538/1938(79%)	41/1938(2%)	Plus/Plus

Query 6 AGGTTTATACCTTCCAGGTAACAACAACAACCAACTTTCATCTCTGTAGATCTGTTCT 65  
 Sbjct 6 AGGTTTATACCTACCCAGG-AA-AAGCAACAAC-CTCGATCTCTGTAGATCTGTTCT 62

Query 66 CTAAACGAACCTTAAAACTTGTGGCTGTCACTCGGCTGCATGCTTATAGTCACTCACCC 125  
 Sbjct 63 CTAAACGAACCTTAAAACTTGTGTAGCTGTGCTCGGCTGCATGCTTATAGTCACTCACCC 122

.....

(a)

Descriptions Graphic Summary **Alignments** Dot Plot

Alignment view Pairwise  CDS feature  Restore defaults

1 sequences selected

Download GenBank Graphics

**Bat coronavirus RaTG13, complete genome**  
 Sequence ID: [MN99532.2](#) Length: 29855 Number of Matches: 1

Range 1: 1 to 29855 [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
48758 bits(26403)	0.0	28723/29877(96%)	24/29877(0%)	Plus/Plus

Query 1 ATTAAGGTTTATACCTTCCAGGTAACAACAACAACAACCAACTTTCATCTCTGTAGATCT 60  
 Sbjct 1 ATTAAGGTTTATACCTTCCAGGTAACAACAACAACAACAACCAACTTTCATCTCTGTAGATCT 60

Query 61 GTTCTCTAAACGAACCTTAAAACTGTGTGGCTGTCACTGGCTGCATGCTTGTAGTGCAT 120  
 Sbjct 61 GTTCTCTAAACGAACCTTAAAACTGTGTGGCTGTCACTGGCTGCATGCTTGTAGTGCAT 120

Query 121 CACGAGTATAATTAATAACTAATTAAGTCTGTTGACAGGACACGAGTAACCTGCTATC 180  
 Sbjct 121 CACGAGTATAATTAATAACTAATTAAGTCTGTTGACAGGACACGAGTAACCTGCTATC 180

Query 181 TTCTGCAGGCTGCTACGGTTTCTGCTGTTGACAGGACATCATCAGCAGCATTAAGTGT 240  
 Sbjct 181 TTCTGCAGGTTGCTACGGTTTCTGCTGTTGACAGGACATCATCAGCAGCATTAAGTGT 240

Query 241 CGTCCGGGTGTGACCAAAAGTAAAGTGAAGAGCTTGTCCCTGGTTTCAACGAGAAAAAC 300  
 Sbjct 241 CGTCCGGGTGTGACCAAAAGTAAAGTGAAGAGCTTGTCCCTGGTTTCAACGAGAAAAAC 300

Query 301 ACAGGTCGAACCTGCTGCTGTTTACAGGTCGCGAGCTGCTGCTAGGCTGGTGG 360  
 Sbjct 301 ACAGGTCGAACCTGCTGCTGTTTACAGGTCGCGAGCTGCTGCTAGGCTGGTGG 360

Query 361 AGACTCCGTTGAGGAGCTTATCAGAGGACGCTCAACATCTTAAAGATGACCATTTGG 420  
 Sbjct 361 AGACTCCGTTGAGGAGCTTATCAGAGGACGCTCAACATCTTAAAGATGACCATTTGG 420

Query 421 CTTAGTGAAGTTGAAAAAGGCGTTTGTCTCAACTGAACAGCCCTATGTTTTCATCAA 480  
 Sbjct 421 CTTAGTGAAGTTGAAAAAGGCGTTTGTCTCAACTGAACAGCCCTATGTTTTCATCAA 480

Query 481 ACGTTCGAGTCTGAACTGCACCTCATGCTCATGTTATGTTGAGCTGGTAGCAGAACT 540  
 Sbjct 481 ACGTTCGAGTCTGAACTGCACCTCATGCTCATGTTATGTTGAGCTGGTAGCAGAACT 540

Query 541 CGAAGCATTGAGTACGGTGTAGTGGTGAAGACACTTGGTCTCTTCCCTCATGTGGG 600  
 Sbjct 541 TAATGGCATTGAGTACGGTGTAGTGGTGAAGACACTTGGTCTCTTCCCTCATGTGGG 600

Query 601 CGAAATACCAAGTGGCTTACCGCAAGGTTCTTCTGTAAGAACGGTAAAGAGAGCTGG 660  
 Sbjct 601 CGAAATACCAAGTGGTTACCGCAAGGTTCTTCTGTAAGAACGGTAAAGAGAGCTGG 660

Query 661 TGCCATAGTACCGGCGGCTTAAAGTCACTTGAAGTGGGAGCAGCTTGGCCTGTA 720  
 Sbjct 661 TGCCATAGTACCGGCGGCTTAAAGTCACTTGAAGTGGGAGCAGCTTGGCCTGTA 720

Query 721 TCCTTATGAAGATTTTCAAGAAAACGGAACATAACATAGCAGTGGTGTACCCTGTA 780  
 Sbjct 721 TCCTTATGAAGATTTTCAAGAAAACGGAACATAACATAGCAGTGGTGTACCCTGTA 780

.....

Query 28739 GCAATCTGCTACAACCTTCTCAAGGAACAACATTCGCAAAAGGCTTCTACGCGAAGGG 28798  
 Sbjct 28720 GCAATCTGCTACAACCTTCTCAAGGAACAACATTCGCAAAAGGCTTCTACGCGAAGGG 28779

Query 28799 AGCAGAGGCGGCACTCAAGCTTCTTCTGCTTCTCATCACTAGTGCAGCAAGTTCAGA 28858  
 Sbjct 28780 AGCAGAGGCGGCACTCAAGCTTCTTCTGCTTCTCATCACTAGTGCAGCAAGTTCAGA 28839

Query 28859 AATCAACTCCAGGCAAGTGAAGGAACTTCTCTGTAGAAAGGCTTGGCAATGGCGT 28918  
 Sbjct 28840 AATCAACTCCAGGCAAGTGAAGGAACTTCTCTGTAGAAAGGCTTGGCAATGGCGT 28899

Query 28919 GATGCTGCTTGTCTTGTCTGCTGTGACAGATTAAGCAAGCTTGAAGCAAAAGTCT 28978  
 Sbjct 28900 GATGCTGCTTGTCTTGTCTGCTGTGACAGATTAAGCAAGCTTGAAGCAAAAGTCT 28959

Query 28979 GTTAAAGGCAACAACAAGGCAAGCTCAACTGCTCAAGAAATCTGCTGCAAGGCTTCT 29038  
 Sbjct 28960 GTTAAAGGCAACAACAAGGCAAGCTCAACTGCTCAAGAAATCTGCTGCAAGGCTTCT 29019

Query 29039 AAGAAGCCTCGGCAAAAAGTACTGCCACTAAGCAATCAAGTGAACACAGCTTCTCGG 29098  
 Sbjct 29020 AAGAAGCCTCGGCAAAAAGTACTGCCACTAAGCAATCAAGTGAACACAGCTTCTCGG 29079

Query 29099 AGAGCTGCTCAGAACAAACCAAGAAACTTTGGGATCAAGAACTAATCAGGCAAGGA 29158  
 Sbjct 29080 AGAGCTGCTCAGAACAAACCAAGAAACTTTGGGATCAAGAACTAATCAGGCAAGGA 29139

Query 29159 ACTGATTACAACATTTGGCGCAATTTGCAAAATTTGGCCAGGCTTCAAGCTTCTCT 29218  
 Sbjct 29140 ACTGATTACAACATTTGGCGCAATTTGCAAAATTTGGCCAGGCTTCAAGCTTCTCT 29199

Query 29219 GGAATGTCGCGCATTTGGCATGGAAGTCAACCTTGGGAAAGTGGTGGACTACAGAGT 29278  
 Sbjct 29200 GGAATGTCGCGCATTTGGCATGGAAGTCAACCTTGGGAAAGTGGTGGACTACAGAGT 29259

Query 29279 GCCATCAAAATGGATGACAAAGTCAAAATTTCAAGATCAAGTATTTTGTGTAAGAAG 29338  
 Sbjct 29260 GCCATCAAAATGGATGACAAAGTCAAAATTTCAAGATCAAGTATTTTGTGTAAGAAG 29319

Query 29339 CATATTGACGCATACAACAACTTCCCAACAGAGGCTTAAAGAGCAAAAAGAGAAAG 29398  
 Sbjct 29320 CATATTGACGCATACAACAACTTCCCAACAGAGGCTTAAAGAGCAAAAAGAGAAAG 29379

Query 29399 GCTGATGAACTCAAGCTTACCGAGAGACAGAAAGCAACCAACTGTGACTTCTCTT 29458  
 Sbjct 29380 GCTGATGAACTCAAGCTTACCGAGAGACAGAAAGCAACCAACTGTGACTTCTCTT 29439

Query 29459 CCTGCTGAGATTTGGATGATTTCTTCAACAATTTGCAAACTTCAAGAGAGCTGCTGAC 29518  
 Sbjct 29440 CCTGCTGAGATTTGGATGATTTCTTCAACAATTTGCAAACTTCAAGAGAGCTGCTGAC 29499

Query 29519 TCAACTCAGGCTAACTCATGACAGCACCAAGGAGATGGCTATATAAAGGTTTTT 29578  
 Sbjct 29500 TCAACTCAGGCTAACTCATGACAGCACCAAGGAGATGGCTATATAAAGGTTTTT 29559

Query 29579 GCTTTTCCGTTTACGATATAGTCTACTTCTGTGAGAAATGAATCTGTAACTACATA 29638  
 Sbjct 29560 GCTTTTCCGTTTACGATATAGTCTACTTCTGTGAGAAATGAATCTGTAACTACATA 29619

Query 29639 GCACAGTAGATGATGATTAACCTTAACTCACTAGCAATCTTAACTAGTGTGTAACAT 29698  
 Sbjct 29620 GCACAGTAGATGATGATTAACCTTAACTCACTAGCAATCTTAACTAGTGTGTAACAT 29679

Query 29699 TAGGGAGACTTGAAGAGCCACACATTTTCCAGGAGCCAGCGGAGTACGATGAGT 29758  
 Sbjct 29680 TAGGGAGACTTGAAGAGCCACACATTTTCCAGGAGCCAGCGGAGTACGATGAGT 29739

Query 29759 GTACAGTGAACAATGCTAGGGAGAGCTGCTATATGAGAGAGCCCTAATGTGTAATAA 29818  
 Sbjct 29740 GTACAGTGAACAATGCTAGGGAGAGCTGCTATATGAGAGAGCCCTAATGTGTAATAA 29799

Query 29819 ATTTTATAGTGCATCCCATGTTATTAATAGCTTCTAGGAGAAATGAC 29875  
 Sbjct 29800 ATTTTATAGTGCATCCCATGTTATTAATAGCTTCTAGGAGAAATGAC 29855

(b)

**Figure VII.5 :** Une partie de l'alignement BLASTN entre le SARS-CoV-2 en tant que séquence de requête et la séquence de sujet pour (a) SARS-CoV et (b) Bat-CoV ; RaTG13.

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

---

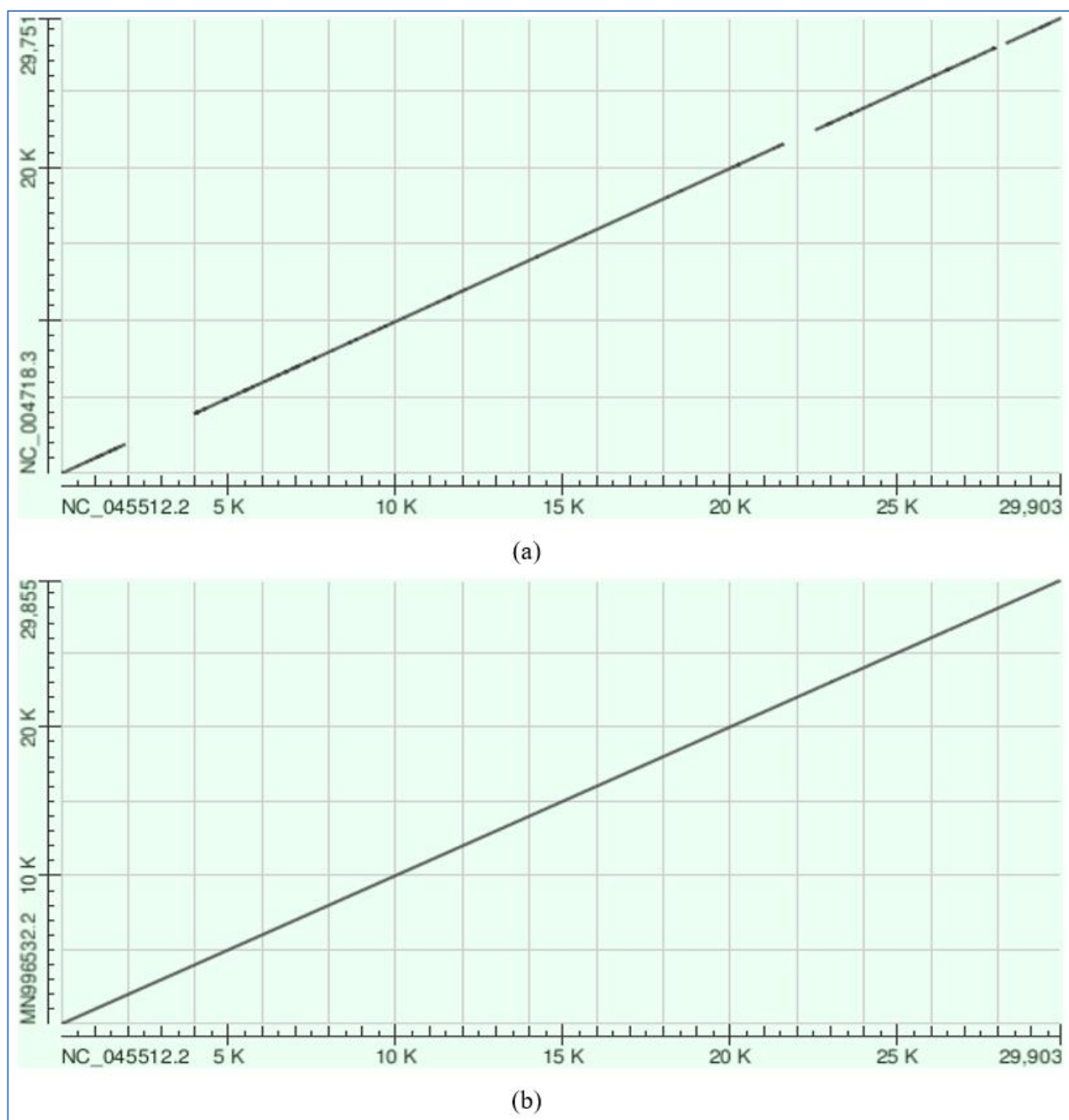
Par défaut, NCBI BLAST masque automatiquement les séquences de faible complexité (séquences avec beaucoup des mêmes bases) dans la séquence de requête, par ex. BLAST n'essaiera pas de faire correspondre ces régions aux séquences de la base de données. Ces bases masquées peuvent apparaître sous forme de lettres minuscules grises ou de X selon les paramètres de recherche BLAST.

Maintenant que nous avons une meilleure compréhension de l'organisation d'un rapport BLAST, nous sommes prêts à interpréter nos résultats de blastn. À partir de l'aperçu graphique et de la liste de descriptions (**figures VII.5 a et b**), nous remarquons que les deux résultats montrent des alignements significatifs avec un degré élevé de matches (avec des valeurs E de 0,0). Ces hits ont également des régions d'homologie sur l'ensemble de notre séquence de requête (**Figure VII.5 a et b**).

Les coordonnées des blocs d'alignement, triées par valeur E, entre SARS-CoV-2 et SARS-CoV par rapport à notre séquence d'ARNm sujet (coordonnées de la séquence de requête entre parenthèses) produisent la liste suivante de blocs d'alignement : *plage1* de 3883 à 21505 ( 3956 à 21577), *plage2* de 22429 à 27798 (22539 - 27910), *plage3* de 28103 à 29751 (28257 - 29894) et *plage4* de 6 à 1922 (6 à 1923). D'autre part, l'alignement (trié par valeur E) entre SARS-CoV-2 et Bat-CoV ; RaTG13 n'a produit qu'une seule liste de bloc d'alignement (*plage*) de 1 à 29855 (1 à 29875).

### VII.6. Méthode Dot Plot (Méthode graphique)

En bioinformatique, Dot Plot est une méthode graphique permettant de comparer deux séquences biologiques et d'identifier les régions de similitude étroite après alignement des séquences. C'est un type de tracé de récurrence.



**Figure VII.6 :** Dot Plot partie de BLASTN entre le SARS-CoV-2 en tant que séquence de requête et la séquence de sujet pour (a) SRAS-CoV et (b) Bat-CoV ; RaTG13.

La méthode de Dot Plot compare deux séquences en organisant une séquence sur l'axe des X et une autre sur l'axe des Y d'un tracé. Lorsque les résidus des deux séquences correspondent au même endroit sur le tracé, un point est tracé à la position correspondante. Notez que les séquences peuvent être écrites en arrière ou en avant, cependant les séquences sur les deux axes doivent être écrites dans le même sens. A noter également, que la direction des séquences sur les axes déterminera la direction

## **Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN**

---

de La ligne sur le Dot Plot. Une fois les points tracés, ils se combinent pour former des lignes.

La proximité des séquences en similitude déterminera à quel point la ligne diagonale est proche de ce qu'est un graphique montrant une courbe démontrant une relation directe. Cette relation est affectée par certaines caractéristiques de séquence telles que les décalages de trame, les répétitions directes et les répétitions inversées. Les décalages de cadre incluent les insertions, les suppressions et les mutations. La présence de l'une de ces caractéristiques, ou la présence de plusieurs caractéristiques, entraînera le tracé de plusieurs lignes dans différentes possibilités de configurations, en fonction des caractéristiques présentes dans les séquences.

Une caractéristique qui entraînera un résultat très différent sur le tracé de points est la présence de régions de faible complexité. Les régions de faible complexité sont des régions de la séquence avec seulement quelques acides aminés, ce qui à son tour provoque une redondance dans cette région petite ou limitée. Ces régions se trouvent généralement autour de la diagonale et peuvent ou non avoir un carré au milieu du Dot Plot [wikiorg].

### **VII.7. Études antérieures de la littérature sur le SARS-CoV-2 (en 2020 et 2021)**

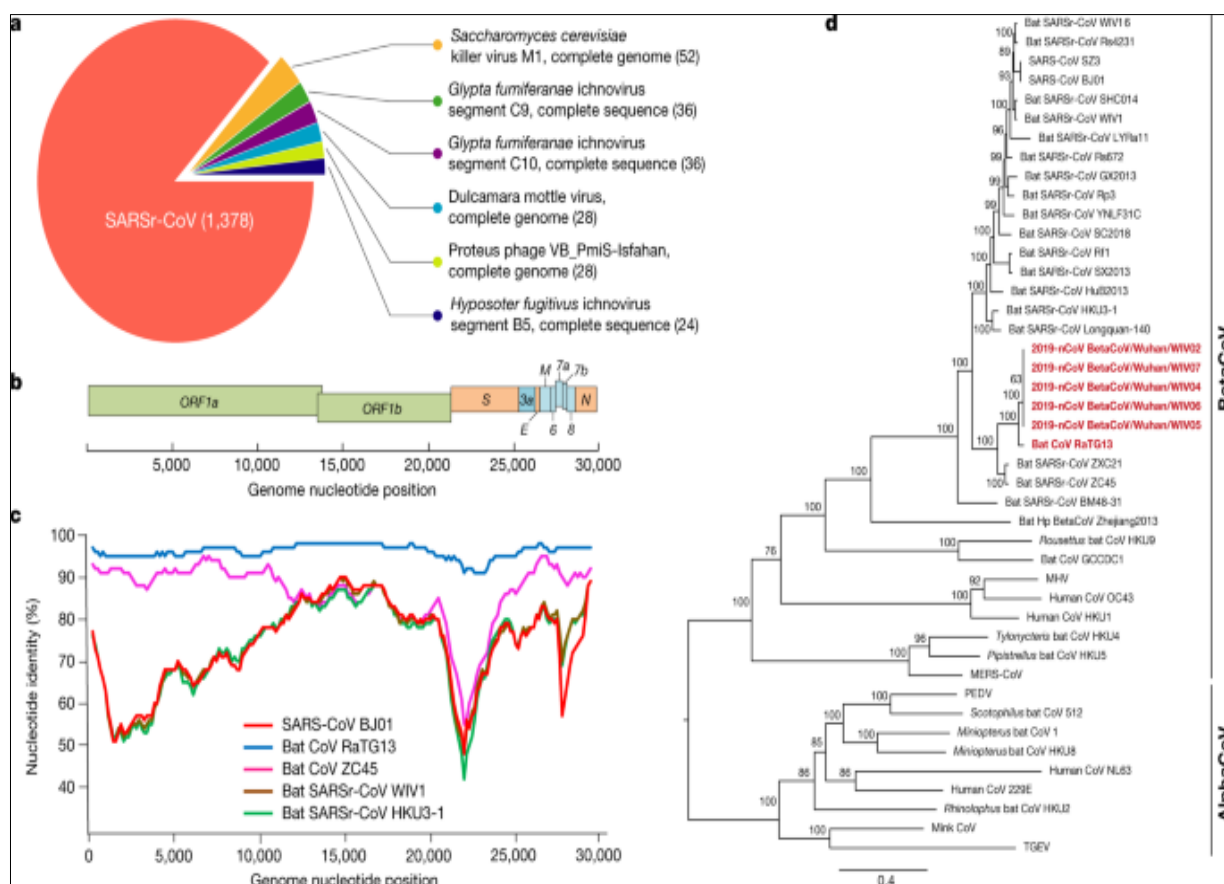
Dans une étude précédente par P. Zhou et al. [Zhou et al. , 2020], sur l'épidémie de pneumonie associée à un nouveau coronavirus d'origine probable de chauve-souris, échantillons de sept patients atteints de pneumonie sévère (dont six vendeurs ou livreurs du marché des fruits de mer), qui ont été admis à l'unité de soins intensifs de l'hôpital Wuhan Jin Yin-Tan au début de l'épidémie, ont été envoyés au laboratoire de l'Institut de virologie de Wuhan (WIV) pour le diagnostic de l'agent pathogène en cause (**Tableau de données étendu 1**). En tant que laboratoire étudiant le CoV, ils ont d'abord utilisé des amorces PCR pan-CoV pour tester ces échantillons<sup>13</sup>, étant donné que l'épidémie s'est produite en hiver et sur un marché - le même environnement que les infections par le SRAS. Ils ont trouvé cinq échantillons positifs par PCR pour les CoV. Un échantillon (WIV04), prélevé dans le liquide de lavage bronchoalvéolaire (BALF), a été analysé par analyse métagénomique en utilisant un séquençage de nouvelle génération pour identifier les agents étiologiques potentiels. Sur les 10 038 758 lectures totales, dont 1 582 lectures totales ont été

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

retenu après filtrage des lectures du génome humain — 1 378 séquences (87,1 %) correspondaient à la séquence du SARSr-CoV (**figure VII.6.a**).

Par assemblage de novo et PCR ciblée, ils ont obtenu un génome CoV de 29 891 paires de bases qui partageait 79,6 % d'identité de séquence avec le SRAS-CoV BJ01 (numéro d'accès GenBank AY278488.2). Une couverture élevée du génome a été obtenue en remappant le total des lectures sur ce génome (**Extended Data Figure VII.6**).

Cette séquence a été soumise à GISAID (<https://www.gisaid.org/>) (numéro d'accès EPI\_ISL\_402124). Après le nom donné par l'Organisation mondiale de la santé (OMS), ils l'appellent provisoirement le nouveau coronavirus 2019 (2019-nCoV). Quatre autres séquences génomiques complètes du 2019-nCoV (WIV02, WIV05, WIV06 et WIV07) (numéros d'accès GISAID EPI\_ISL\_402127-402130) qui étaient identiques à plus de 99,9 % les unes aux autres ont ensuite été obtenues à partir de quatre patients supplémentaires utilisant le séquençage de nouvelle génération et PCR.



**Figure VII.7:** Caractérisation du génome du 2019-nCoV [Zhou et al.,2020].

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

---

Le génome du virus se compose de six grands cadres de lecture ouverts (ORF) communs aux coronavirus et à un certain nombre d'autres gènes accessoires (**Figure VII.7 a**). Une analyse plus approfondie indique que certains des gènes 2019-nCoV partageaient moins de 80% d'identité de séquence de nucléotides avec le SRAS-CoV. Cependant, les séquences d'acides aminés des sept domaines de réplicase conservés dans ORF1ab qui ont été utilisés pour la classification des espèces de CoV étaient identiques à 94,4% entre 2019-nCoV et SARS-CoV, suggérant que les deux virus appartiennent à la même espèce, SARSr-CoV.

P. Zhou et al., ont rapporté que [**Zhou et al., 2020**], le génome du virus se compose de six cadres de lecture ouverts (ORF) majeurs qui sont communs aux coronavirus et à un certain nombre d'autres gènes accessoires (**Figure VII.7 b**). Une analyse plus approfondie indique que certains des gènes 2019-nCoV partageaient moins de 80% d'identité de séquence de nucléotides avec le SRAS-CoV. Cependant, les séquences d'acides aminés des sept domaines de réplicase conservés dans ORF1ab qui ont été utilisés pour la classification des espèces de CoV étaient identiques à 94,4% entre 2019-nCoV et SARS-CoV, suggérant que les deux virus appartiennent à la même espèce, SARSr-CoV.P.

Zhou et al., [**Zhou et al., 2020**] ont ensuite découvert qu'une courte région d'ARN polymérase dépendante de l'ARN (RdRp) d'un coronavirus de chauve-souris (BatCoV RaTG13) - qui avait déjà été détectée chez *Rhinolophus affinis* de la province du Yunnan - a montré une identité de séquence élevée avec 2019-nCoV. Nous avons effectué un séquençage complet sur cet échantillon d'ARN (numéro d'accès GISAID EPI\_ISL\_402131). L'analyse Simplot a montré que le 2019-nCoV était très similaire dans l'ensemble du génome à RaTG13 (**Figure VII.7 c**), avec une identité globale de la séquence du génome de 96,2 %. En utilisant les séquences génomiques alignées de 2019-nCoV, RaTG13, SARS-CoV et des SARSr-CoV de chauve-souris précédemment signalés, aucune preuve d'événements de recombinaison n'a été détectée dans le génome de 2019-nCoV. L'analyse phylogénétique du génome complet et des séquences de gènes de RdRp et de pointe (S) a montré que, pour toutes les séquences, le RaTG13 est le parent le plus proche du 2019-nCoV et forment une lignée distincte des autres SARSr-CoV (**Figure VII.7 d**)

Des séquences complètes du génome ont été obtenues à partir de cinq patients à un stade précoce de l'épidémie. Les séquences sont presque identiques et partagent

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

---

79,6% d'identité de séquence avec le SRAS-CoV. De plus, nous montrons que le 2019-nCoV est identique à 96 % au niveau du génome entier à un coronavirus de chauve-souris [Zhou et al. , 2020].

L'analyse par paires de séquences protéiques de sept domaines protéiques non structuraux conservés montre que ce virus appartient à l'espèce SARSr-CoV. De plus, le virus 2019-nCoV isolé du liquide de lavage broncho-alvéolaire d'un patient gravement malade pourrait être neutralisé par les sérums de plusieurs patients. Notamment, P. Zhou et al. [Zhou et al., 2020] ont confirmé que le 2019-nCoV utilise le même récepteur d'entrée cellulaire - l'enzyme de conversion de l'angiotensine II (ACE2) - que le SRAS-CoV [Zhou et al 2020].

L. Lv et al. [Lv et al. 2020] ont effectué des analyses génomiques comparatives révèlent un modèle de mutation spécifique entre le coronavirus humain SARS-CoV-2 et le Bat-CoV RaTG13. Ils ont d'abord observé que la proportion de substitutions synonymes était similaire à celle des substitutions non synonymes au sein des souches SARS-CoV-2 actuellement disponibles. Selon l'hypothèse de la dérive aléatoire [Castellano et al., 2018], ces différences de nucléotides entre les différentes souches du SRAS-CoV-2 peuvent résulter principalement d'une évolution neutre. Bref, aucun facteur puissant n'existe pour forcer le SARS-CoV-2 à évoluer dans une certaine direction.

Cependant, des précautions strictes doivent être prises contre les facteurs puissants qui peuvent provoquer une variation directionnelle du SRAS-CoV-2 à la fois dans l'environnement naturel et pendant le traitement de l'infection.

Deuxièmement, nos résultats ont montré que les mutations synonymes sont considérablement élevées entre le SARS-CoV-2 et le RaTG13. La proportion relative de substitutions synonymes entre le SRAS-CoV-2 humain et son origine animale possible (RaTG13) est beaucoup plus élevée que celle entre les autres coronavirus humains et leurs sources animales potentielles. Ces résultats ont indiqué que les souches du SRAS-CoV-2 pourraient subir une sélection purificatrice plus forte après avoir divergé de leur ancêtre commun. Fait intéressant, les mutations nucléotidiques ont été enrichies dans la transition T:C. Le modèle de mutation spécifique peut être causé par l'inactivation de l'ARN 30 à 50 ExoN [Smith et al., 2013] . Cette mutation T:C accrue implique que l'ExoN du SARS-CoV-2 peut être désactivé par rapport à celui du RaTG13. De plus, le traitement par ARN mutagène 5-fluorouracile (5-FU)

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

---

peut également augmenter les transitions U:C et A:G [Smith et al.,2013]. Par conséquent, les mécanismes sous-jacents de ces mutations potentielles entre le SRAS-CoV-2 et le RaTG13 nécessitent une enquête plus approfondie à l'avenir.

Un résumé de l'étude de L. Lv et al.[Lv et al. 2020], grâce à une analyse comparative complète entre le SRAS-CoV-2 et d'autres coronavirus, ils ont découvert que les mutations synonymes étaient considérablement élevées entre le SRAS-CoV-2 et le RaTG13 par rapport à d'autres souches de coronavirus, et que les mutations nucléotidiques étaient enrichies dans la transition T:C . Étant donné que le SARS-CoV-2 est censé provenir du Bat-CoV RaTG13, la substitution synonyme accrue entre le SARS-CoV-2 et la souche RaTG13 suggère que le SARS-CoV-2.

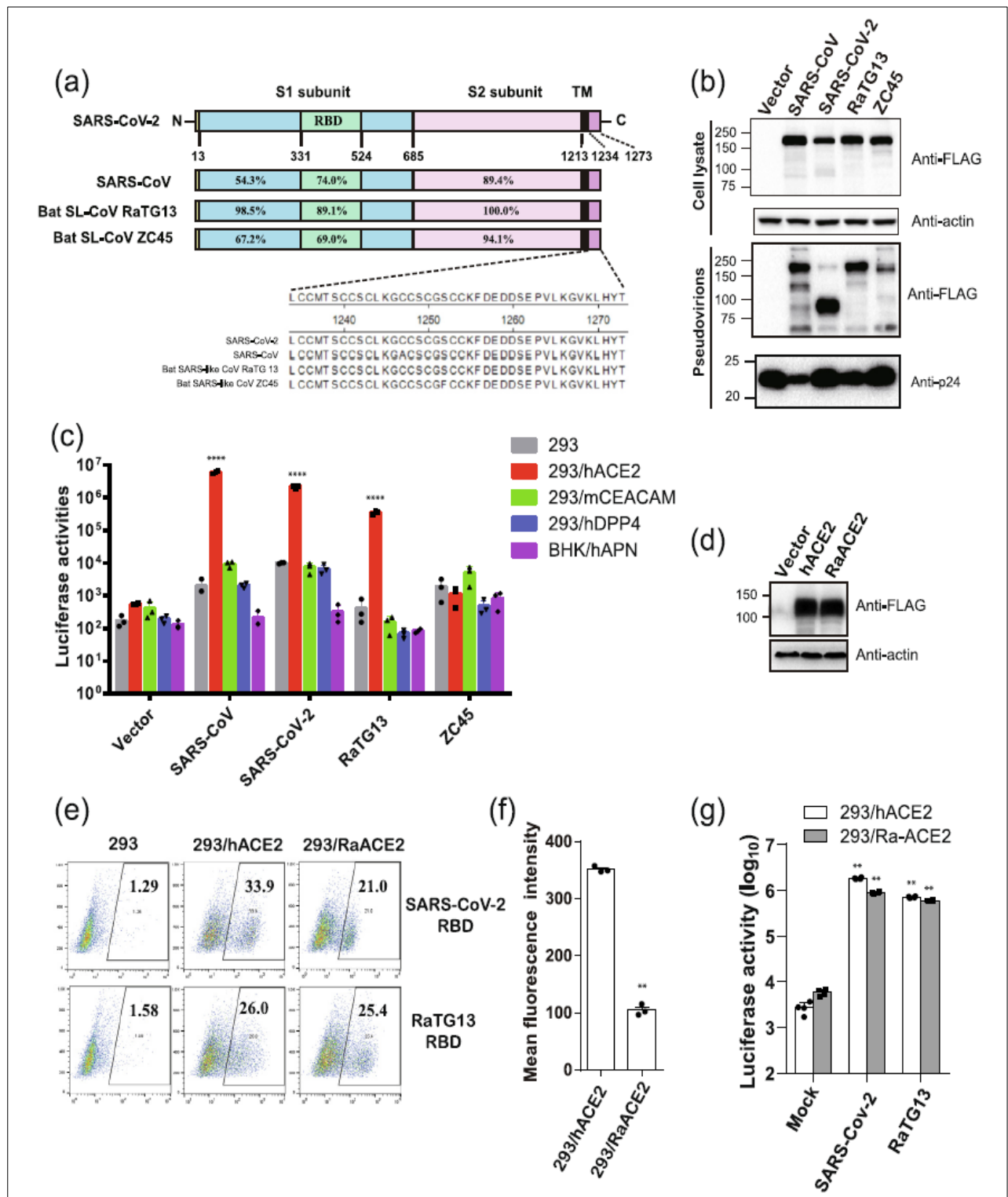
Le génome doit être soumis à une sélection négative (purifiante) rigoureuse. De plus, le mécanisme qui sous-tend l'augmentation des mutations T:C nécessite une enquête plus approfondie.[Lv et al. 2020].

Dans une étude intitulée "Le Rhinolophus affinis bat ACE2 et plusieurs orthologues animaux sont des récepteurs fonctionnels pour le coronavirus de chauve-souris RaTG13 et SARS-CoV-2", P. Li et al. [Li et al., 2021] ont rapporté que le coronavirus de chauve-souris (CoV) RaTG13 partage l'identité de séquence génomique la plus élevée avec le coronavirus 2 du syndrome respiratoire aigu sévère (SRAS-CoV-2) parmi tous les coronavirus connus, et utilise également l'enzyme de conversion de l'angiotensine humaine 2 ( hACE2) pour l'entrée du virus. Ainsi, on pense que le SARS-CoV-2 est originaire de la chauve-souris. Cependant, si le SRAS-CoV-2 a émergé des chauves-souris directement ou via un hôte intermédiaire reste insaisissable.

Ici, P. Li et al. [Li et al., 2021] ont découvert que Rhinolophus affinis bat ACE2 (RaACE2) est un récepteur d'entrée pour le SARSCoV-2 et le RaTG13, bien que la liaison de RaACE2 au domaine de liaison au récepteur (RBD) du SARSCoV-2 soit nettement plus faible que celui de hACE2. Ils ont en outre évalué les activités des récepteurs des ACE2 de 16 espèces animales diverses supplémentaires pour RaTG13, SARS-CoV et SARS-CoV-2 en termes de liaison à la protéine S, de fusion membranaire et d'entrée de pseudovirus. Ils ont découvert que la protéine de pointe (S) RaTG13 est significativement moins fusogène que le SRAS-CoV et le SARS-CoV-2, et sept des seize ACE2 différents fonctionnent comme des récepteurs d'entrée

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

pour les trois virus, indiquant que les trois virus pourraient avoir un hôte large rage (figure VII.7).



**Figure VII.8 :** Bat SL-CoV RaTG13 utilise hACE2 et RaACE2 pour l'entrée du virus [Li et al. 2021].

## **Chapter VII                      Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN**

---

(a) Un diagramme schématique de la longueur totale des différentes protéines CoV S et les identités de séquence d'acides aminés de chaque région sont indiqués aux endroits correspondants. S1, sous-unité de liaison au récepteur ; S2, sous-unité de fusion membranaire ; TM, domaine transmembranaire.

(b) Détection des protéines S du SARS-CoV, SARS-CoV-2, bat SL-CoV RaTG13 et ZC45 dans les lysats cellulaires et les pseudovirions par Western blot (WB). Les cellules HEK293T transfectées avec un vecteur vide ou des plasmides codant pour les protéines CoV S indiquées ont été lysées 40 h après la transfection. Les protéines S dans les lysats cellulaires et les pseudovirions ont été soumises à une analyse WB par transfert avec un anticorps monoclonal de souris anti-FLAG M2. L'actine et le gag-p24 ont servi de témoins de chargement (lysats cellulaires, panneau supérieur, pseudovirions, panneau inférieur). La protéine S complète est d'environ 180 kD, tandis que la protéine S clivée est d'environ 90 kD. Des expériences ont été faites trois fois et le représentant a été montré.

(c) Entrée des pseudovirions RaTG13 S sur différents récepteurs CoV. Les cellules ont été inoculées par centrifugation avec les pseudovirions indiqués. 48 h après l'inoculation, l'efficacité de la transduction a été déterminée par la mesure des activités luciférase. cellules HEK293 (gris), HEK293/hACE2 (rouge), cellules HEK293 exprimant de manière stable hACE2; 293/mCEACAM (vert), cellules HEK293 exprimant de manière stable mCEACAM, le récepteur MHV ; 293/hDPP4 (bleu), cellules HEK293 exprimant de manière stable hDPP4, le récepteur MERS-CoV. BHK/hAPN (violet), cellules BHK exprimant de manière stable hAPN, le récepteur hCoV-229E. Les expériences ont été réalisées en triple et répétées au moins trois fois. Un représentant est affiché avec des barres d'erreur indiquant SEM.

(d) Expression de la protéine ACE2 de *Rhinolophus affinis* dans des cellules HEK 293. Les cellules HEK 293 transfectées de manière transitoire avec les plasmides codant pour les protéines hACE2 ou *Rhinolophus affinis* ACE2 (RaACE2) marquées par FLAG ont été lysées 40 h après la transfection. L'expression des protéines ACE2 a été détectée par un anticorps monoclonal de souris anti-FLAG M2.

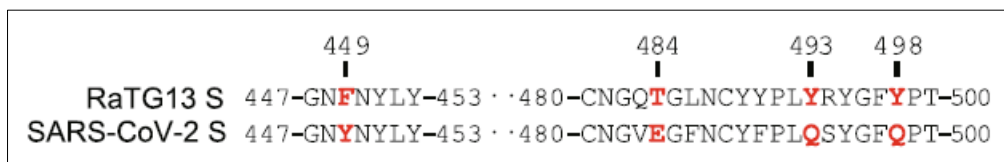
(e) Liaison de hACE2 et RaACE2 par les RBD SARSCoV-2 et RaTG13. Les cellules HEK 293 exprimant de manière transitoire les protéines hACE2 ou RaACE2 ont été incubées avec SARS-CoV-2 RBD ou RaTG13 RBD sur glace, suivies d'anticorps de lapin anti-His tag et d'IgG de chèvre anti-lapin conjugué Alexa-488, et analysées par

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

cytométrie en flux. Les expériences ont été faites trois fois, et un représentant est montré. (f) Intensités de fluorescence moyennes des cellules fermées positives pour le SARS-CoV-2 RBD se liant aux cellules 293/hACE2 et 293/RaACE2 en (e).

g) Entrée des pseudovirions des protéines SARS-CoV, SARS-CoV-2 et RaTG13 S sur des cellules 293/RaACE2. Des expériences ont été faites trois fois, et un représentant est montré avec des barres d'erreur indiquant SEM. \*P < 0,05 ; \*\*P < 0,001 (comparé au contrôle par ANOVA suivi du test t de comparaisons multiples de Dunnett).

Il est à noter que les pseudovirions RaTG13 S peuvent utiliser la souris, mais pas la pangoline ACE2, pour l'entrée du virus, tandis que les pseudovirions SARS-CoV-2 S peuvent utiliser la pangoline, mais pas la souris, l'ACE2 pénètre efficacement dans les cellules. L'analyse de mutagenèse a révélé que les résidus 484 et 498 dans les protéines RaTG13 et SARS-CoV-2 S jouent un rôle essentiel dans la reconnaissance des ACE2 murins et humains. Dans la figure suivante (**Figure VII.7**), les résidus 449, 484, 493 et 498 sont marqués en rouge. Détection de protéines S mutantes dans des lysats cellulaires et des pseudovirions par Western blot à l'aide d'un anticorps monoclonal de souris anti-FLAG M2.



**Figure VII.9** : Alignment of partial amino acid sequences of RaTG13 and SARS-CoV-2 S proteins [Li et al2021].

Enfin, P. Li et al. ont rapporté que deux ACE2 polymorphes de chauve-souris *Rhinolophus sinicus* ont montré des sensibilités différentes à l'entrée du virus par les pseudovirions RaTG13 et SARS-CoV-2 S, suggérant une coévolution possible. En outre, leurs résultats offrent une meilleure compréhension du mécanisme d'entrée du coronavirus, de la gamme d'hôtes et de la coévolution d'hôte de virus [Li et al2021].

Mohammed Elimam [Mohammed2021] a réalisé une étude importante intitulée « Les pourcentages de similarité et d'identité de la protéine SARS-CoV-2 avec les protéines SARS-CoV et BatCoV RaTG13 peuvent être utilisés comme indicateurs de l'origine du virus ». Cette étude a examiné la similarité des séquences d'acides aminés et les pourcentages d'identité de 10 protéines dans le SARS-CoV-2, le SARS-CoV et le

## **Chapter VII                    Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN**

---

coronavirus de chauve-souris *Rhinolophus afnis* (BatCoV RaTG13). Les protéines étudiées étaient la polyprotéine 1ab, la protéine de pointe, orf3a, la protéine d'enveloppe, la protéine membranaire, orf6, orf7a, orf7b, orf8 et la protéine de la nucléocapside. Le service d'alignement de séquences en ligne de The European Molecular Biology Open Software Suite (EMBOSS) a été utilisé pour déterminer les pourcentages de similarité et d'identité des protéines dans les trois virus.

La polyprotéine 1ab du SARS-CoV-2, du SARS-CoV et du BatCoV RaTG13 a été étudiée. De plus, les protéines structurales et accessoires trouvées dans le SRAS-CoV-2 et le BatCoV RaTG13 ont été étudiées, y compris la protéine de pointe (S), orf3, la protéine d'enveloppe (E), la protéine membranaire (M), orf6, orf7a, orf7b, orf8, et la protéine de nucléocapside (N) (**tableau 2**). Les séquences d'acides aminés ont été obtenues sur le site du National Center for Biotechnology Information (NCBI) [NCBIGOV] (**tableau 1**).

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

**Tableau VII.1** :Les protéines étudiées des trois virus [Mohammed2021].

	Protein		SARS-CoV-2	SARS-CoV	RaTG13
1	1ab polyprotein	NCBI Code	YP_009724389.1	NP_828849.7	QHR63299.1
		Gene location	266..21555	265..21485	251..21537
		Amino acid number	7096	7073	7095
2	S protein	NCBI Code	YP_009724390.1	YP_009825051.1	QHR63300.2
		Gene location	21492..25259	21492..25259	21545..25354
		Amino acid number	1273	1255	1269
3	Orf3	NCBI Code	YP_009724391.1	YP_009825052.1	QHR63301.1
		Gene location	25393..26220	25268..26092	25363..26190
		Amino acid number	275	274	275
4	E protein	NCBI Code	YP_009724392.1	YP_009825054.1	QHR63302.1
		Gene location	26245..26472	26117..26347	26215..26442
		Amino acid number	75	76	75
5	M protein	NCBI Code	YP_009724393.1	YP_009825055.1	QHR63303.1
		Gene location	26523..27191	26398..27063	26493..27158
		Amino acid number	222	221	221
6	Orf6	NCBI Code	YP_009724394.1	YP_009825056.1	QHR63304.1
		Gene location	27202..27387	26913..27265	27169..27354
		Amino acid number	61	63	61
7	Orf7a	NCBI Code	YP_009724395.1	YP_009825057.1	QHR63305.1
		Gene location	27394..27759	27273..27641	27360..27725
		Amino acid number	121	122	121
8	Orf7b	NCBI Code	YP_009725318.1	YP_009825058.1	QHR63306.1
		Gene location	27756..27887	27638..27772	27722..27853
		Amino acid number	43	44	43
9	Orf8	NCBI Code	YP_009724396.1	YP_009825059.1 YP_009825060.1	QHR63307.1
		Gene location	27894..28259	27779..27898, 27864..28118	27860..28225
		Amino acid number	121	39, 84	121
10	N protein	NCBI Code	YP_009724397.2	YP_009825061.1	QHR63308.1
		Gene location	28274..29533	28120..29388	28240..29499
		Amino acid number	419	422	419

The number of amino acids and their sequences for the proteins were obtained from the National Center for Biotechnology Information (NCBI) site accessible at <https://www.ncbi.nlm.nih.gov/protein>

La polyprotéine 1ab du SARS-CoV-2, du SARS-CoV et du BatCoV RaTG13 est composée respectivement de 7096, 7073 et 7095 acides aminés (**tableau 1**). L'identité et la similitude de la séquence d'acides aminés de la polyprotéine 1ab du SARSCoV-2 et du BatCoV RaTG13 étaient respectivement de 98,5% et 99,1%. Les pourcentages d'identité et de similarité de la polyprotéine 1ab du SARS-CoV-2 et du SARS-CoV étaient respectivement de 86,2 % et 92,9 %. Les résultats montrent que le SRAS-CoV-2 provient très probablement de la chauve-souris *Rhinolophus afnis*, et non d'une variante du SRAS-CoV modifiée en laboratoire (**tableau VII.2**).

## Chapter VII Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN

**Tableau VII.2:** Les pourcentages d'identité et de similarité des protéines du SARS-CoV-2 par rapport à ceux du SARSCoV et du RaTG13 (coronavirus de chauve-souris)

Protein		Identity %	Similarity %	
1	1 ab polyprotein	SARS-CoV-2 and SARS-CoV	86.2	92.9
		SARS-CoV-2 and RaTG13	98.5	99.1
2	Spike protein	SARS-CoV-2 and SARS-CoV	76	86
		SARS-CoV-2 and RaTG13	97.4	98.4
3	Orf3a (NS3)	SARS-CoV-2 and SARS-CoV	72.4	85.1
		SARS-CoV-2 and RaTG13	97.8	98.9
4	E protein	SARS-CoV-2 and SARS-CoV	94.7	96.1
		SARS-CoV-2 and RaTG13	100	100
5	M protein	SARS-CoV-2 and SARS-CoV	90.5	96.4
		SARS-CoV-2 and RaTG13	99.5	99.5
6	Orf6 (NS6)	SARS-CoV-2 and SARS-CoV	68.9	88.5
		SARS-CoV-2 and RaTG13	100	100
7	Orf7a (NS7a)	SARS-CoV-2 and SARS-CoV	85.2	90.2
		SARS-CoV-2 and RaTG13	97.5	99.2
8	Orf7b (NS7b)	SARS-CoV-2 and SARS-CoV	85.4	90.2
		SARS-CoV-2 and RaTG13	97.7	97.7
9	Orf8 (NS8)	SARS-CoV-2 and SARS-CoV	38.9	77.8
			44.4	66.7
10	N protein	SARS-CoV-2 and RaTG13	95	95.9
		SARS-CoV-2 and SARS-CoV	90.5	94.3

La protéine de pointe du SARS-CoV-2 contient 1273 acides aminés, tandis que la protéine de pointe du SARS-CoV contient 1255 acides aminés et celle de BatCoV RaTG13 contient 1269 acides aminés (**tableau 1**). La protéine de pointe du SARSCoV-2 et celle du SARS-CoV a un pourcentage d'identité de 76% et un pourcentage de similarité de 86% (**tableau 2**). Les pourcentages d'identité et de similarité de la protéine de pointe du SRAS-CoV-2 et de la protéine de pointe de RaTG13 sont respectivement de 97,4 % et 98,4 % (**tableau 2**). Les pourcentages d'identité et de similarité de la protéine de pointe du SARS-CoV-2 et du RaTG13 sont plus élevés que ceux de la protéine de pointe du SARS-CoV-2 et du SARS-CoV.

Les résultats de Mohammed Elimam [**Mohammed2021**] ont montré que les pourcentages de similarité et d'identité des protéines SARS-CoV-2 et BatCoV RaTG13 étaient tous deux supérieurs à 95 %, tandis que les pourcentages d'identité et de similarité du SARS-CoV-2 et du SARS-CoV étaient tous deux supérieurs à 38 %. Les protéines SARS-CoV-2 et le BatCoV RaTG13 partagent une identité et une similitude élevées par rapport aux protéines SARS-CoV-2 et SARS-CoV. Les résultats de cette étude ont prouvé l'utilité de déterminer les pourcentages d'identité et de similarité des protéines pour déterminer l'origine des virus.

## **Chapter VII                    Une étude comparative du SARS-CoV-2 versus SARS-CoV et Bat-CoV ; RaTG13 avec l'outil BLASTN**

---

[Mohammed2021] Mohammed Elimam Ahamed Mohammed. Les pourcentages de similarité et d'identité des protéines SARS-CoV-2 avec les protéines SARS-CoV et BatCoV RaTG13 peuvent être utilisés comme indicateurs de l'origine du virus. Journal des protéines et de la protéomique.

# **Conclusion générale**

## Conclusion générale

---

Le SARS-CoV-2 est un nouveau coronavirus identifié comme la cause de la maladie à coronavirus de 2019 (COVID-19) qui a débuté à Wuhan, en Chine à la fin de 2019 et s'est propagé dans le monde entier. Il s'agit de virus à ARN, regroupés en quatre sous-familles *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus* et *Deltacoronavirus*. Les symptômes généraux de la COVID-19 sont assez similaires à ceux d'autres infections respiratoires aiguës telles que la grippe saisonnière. Il y a trois tests connus pour le diagnostic de Covid19, sont le test rapide, le test PCR

Les coronavirus codent pour quatre protéines structurelles majeures, à savoir la pointe (S), la membrane (M), l'enveloppe (E) et la nucléocapside (N). La protéine du coronavirus S est une grande protéine transmembranaire virale multifonctionnelle de classe I. La taille de cette abondante protéine S varie de 1160 acides aminés (IBV, virus de la bronchite infectieuse, chez les volailles) à 1400 acides aminés (FCoV, coronavirus félin).

L'ARN du virus contient 6 cadres de lecture ouverts (ORF1ab, ORF3a, ORF6, ORF7ab, ORF8, et ORF10). Les deux tiers du génome du virus comprennent l'ORF 1a/1b et le reste un tiers du code génomique est utilisé pour M (membrane), S (pointe), N (nucléocapside) et E protéines structurales virales (enveloppées). Le RBD du SARS-CoV-2 se trouve dans le domaine C-terminal de la protéine de pointe comme dans SARS-CoV, Bat CoV HKU4 et MERS-CoV. Il a également été signalé que le SARS-CoV-2 utilise l'ACE2 (enzyme de conversion de l'angiotensine) comme récepteur cellulaire pour l'entrée dans l'humain cellules.

Le rôle protecteur de la réponse immunitaire adaptative de l'hôte au cours de l'infection par le SARS-CoV-2 est devenu une question critique en l'absence d'un traitement spécifique, d'un vaccin préventif ou d'une immunothérapie. Au cours de l'infection par le SARS-CoV-2, la réponse immunitaire contribuerait à la défense de l'hôte dans la majorité des cas, mais serait responsable de sa pathogénèse chez certains malades. Notamment, au cours des formes sévères, un déséquilibre entre les réponses immunitaires innée et adaptative pourrait être fatal. Au cours de la COVID-19, de nombreuses questions se posent sur la génération de l'immunité spécifique contre les diverses protéines du virus, la cinétique, la fonction des anticorps, ainsi que la qualité des réponses des lymphocytes effecteurs CD4<sup>+</sup> et CD8<sup>+</sup> pour la protection de l'hôte. L'étude bio-informatique des épitopes T et B des coronavirus a soulevé la question de l'immunité croisée entre le SARS-COV-2 et d'autres coronavirus sources

## Conclusion générale

---

d'infection bénigne ou responsables de pneumopathies graves telles que le MERS-CoV et le SARS-CoV.

Le diagnostic précoce et l'isolement des patients suspects jouent un rôle essentiel dans le contrôle de cette épidémie. La spécificité et la sensibilité des différentes techniques de diagnostic diffèrent selon les populations et les types d'équipements utilisés. Plusieurs procédures ont été recommandées pour le diagnostic de COVID - 19 tel que : (1) caractéristique cliniques, (2) test de dépistage en laboratoire, (3) imagerie, (4) examen moléculaire dosage et (5) immunologique.

L'expérience antérieure avec des maladies similaires causées par le SRAS-CoV et le MERS-CoV a jeté les bases d'un vaccin accéléré contre le SRAS-CoV-2 développement. Notre situation mondiale actuelle nécessite un besoin urgent de développement rapide de vaccins tout en respectant des directives strictes la sécurité des vaccins contre le SRAS-CoV-2. Bien qu'il y ait un fort optimisme pour candidats vaccins de stade avancé et approbation réglementaire imminente, le sur le terrain doit également être préparé pour une mutation potentielle du SRAS-CoV-2 comme ses récurrences saisonnières. De plus, nous devons développer des contre-mesures contre d'autres agents pathogènes émergents. En particulier, la fragilité structurelle et immunogénicité sous-optimale de nombreux vaccins Les candidats doivent être adressés.

En tant que domaine scientifique interdisciplinaire, la bioinformatique combine la biologie, l'informatique, l'ingénierie de l'information, les mathématiques et les statistiques pour analyser et interpréter les données biologiques. La bioinformatique a été utilisée pour des analyses *in silico* de requêtes biologiques en utilisant des techniques mathématiques et statistiques. La bioinformatique comprend des études biologiques qui utilisent la programmation informatique dans le cadre de leur méthodologie, ainsi qu'une analyse spécifique « pipelines » qui sont utilisées à plusieurs reprises, en particulier dans le domaine de la génomique. Les utilisations courantes de la bioinformatique comprennent l'identification de gènes candidats et de polymorphismes nucléotidiques simples (SNP).

Le National Center for Biotechnology Information (NCBI) fait partie de la National Library of Medicine (NLM) des États-Unis, une branche des National Institutes of Health (NIH). Le NCBI abrite une série de bases de données relatives à la biotechnologie et à la biomédecine et constitue une ressource importante pour les

## Conclusion générale

---

outils et services bioinformatiques. Les principales bases de données incluent GenBank pour les séquences d'ADN, et PubMed une base de données bibliographique pour la littérature biomédicale.

La base de données Nucleotide est une collection de séquences provenant de plusieurs sources, dont GenBank, RefSeq, TPA et PDB. Nous avons utilisé la base de données Nucleotide pour obtenir le génome complet de COVID-19 en utilisant des mots clés similaires tels que covid19, SARS-CoV-2 ou le génome du coronavirus 2 du syndrome respiratoire aigu sévère via une recherche dans la base de données de nucléotides. La recherche dans la base de données Nucleotide fournit une séquence de requêtes dans trois formats différents : (1) *GenBank*, (2) *FASTA* et (3) *Graphic*. Le format GenBank est une base de données publique complète de séquences nucléotidiques et prenant en charge les annotations bibliographiques et biologiques. Le format FASTA est un format textuel pour représenter des séquences de nucléotides ou des séquences d'acides aminés, dans lequel les nucléotides ou les acides aminés sont représentés à l'aide de codes à une lettre. Le format Graphic est le composant principal pour visualiser de séquence graphiquement.

Le format de la base de données GenBank Nucleotide nous a fourni de nombreuses informations importantes telles que la longueur de la séquence (nombre de paires de bases de nucléotides) du génome complet du SRAS-CoV-2 est de 29903 pb et la division GenBank est VRL. De plus, le numéro d'accession est NC\_045512, tel qu'il a été modifié pour la dernière fois le 18-JUL-2020 ...etc. Dans notre travail, nous avons utilisé le programme BLASTN pour identifier et caractériser la séquence complète du génome du SARS-CoV-2 par rapport aux séquences complètes du génome du SARS-CoV et du Bat-CoV ; RaTG13.

D'après la liste des résultats BLASTN, nous pouvons noter que tous les alignements sont cohérents avec notre séquence de requête (SARS-CoV-2): tous les blocs d'alignement présentent un degré élevé d'homologie de séquence, une identité de séquence au niveau nucléotidique égal à 82,3% pour le SRAS- CoV et 92,96 % pour le Bat-CoV ; RaTG13. Aussi, la séquence de Bat-CoV; RaTG13 correspond mieux à notre séquence du SRAS-CoV-2 qu'à la séquence du SRAS. Nous voyons également que, le correspondance de séquence Bat-CoV; RaTG13 couvrent presque toute la longueur de notre séquence de requête.

Des études antérieures de la littérature confirment que l'analyse par paires de séquences protéiques de sept domaines de protéines non structurales conservées

## Conclusion générale

---

montre que ce virus appartient à l'espèce SARSr-CoV. De plus, le virus 2019-nCoV isolé du liquide de lavage broncho-alvéolaire d'un patient gravement malade pourrait être neutralisé par les sérums de plusieurs patients. Notamment, il est confirmé que le 2019-nCoV utilise le même récepteur d'entrée cellulaire - l'enzyme de conversion de l'angiotensine II (ACE2) - que le SRAS-CoV. Les résultats de la littérature ont montré une proportion accrue de substitutions synonymes et la transition T:C entre le SARS-CoV-2 et le RaTG13.

Références  
bibliographiques

## Références bibliographiques

---

1. Institut national d'excellence en santé et services sociaux (INESSS). COVID-19 – Indications et critères d'accès au test de détection moléculaire du SARS-CoV-2 (21 juillet). Réponse rapide. 2020. Version du 21 juillet 2020. Repéré sur : <https://www.inesss.qc.ca/covid-19/investigation-procedures-diagnostiques/indications-et-criteres-dacces-autes-test-de-detection-moleculaire-du-sars-cov-2.html>
2. Fiche épidémiologique et clinique de la COVID-19, INSPQ, Mise à jour 4 mars 2021 – version 3.1
3. Han MS, Choi EH, Chang SH, Jin B-L, Lee EJ, Kim BN, et al. Clinical characteristics and viral RNA detection in children with coronavirus disease 2019 in the Republic of Korea. *JAMA pediatrics*. 2020. Repéré sur : <http://jamanetwork.com/article.aspx?doi=10.1001/jamapediatrics.2020.3988>
4. Organisation mondiale de la santé (OMS). SRAS-CoV-2 variants 2020. Version du 31 décembre 2020. Repéré sur : <https://www.who.int/csr/don/31-december-2020-sars-cov2-variants/en/>
5. Agence ontarienne de protection et de promotion de la santé (Santé publique Ontario). COVID-19 UK Variant VOC202012/01 – What We Know So Far. 2020. Version du 29 décembre 2020. Repéré sur : <https://www.publichealthontario.ca/-/media/documents/ncov/covid-wwksf/2020/12/what-we-know-uk-variant.pdf>
6. Organisation mondiale de la santé (OMS). Transmission of SARS-CoV-2: implications for infection prevention precautions. Scientific brief. 2020. Version du 9 juillet 2020. Repéré sur : <https://www.who.int/newsroom/commentaries/detail/transmission-of-sars-cov-2-implications-for-infection-prevention-precautions>
7. Institut national de santé publique du Québec (INSPQ). COVID-19 : Modes de transmission et mesures de prévention et de protection contre les risques, incluant le rôle de la ventilation. Questions-réponses. 2021. Version du 13 janvier 2021. Repéré sur : <https://www.inspq.qc.ca/covid-19/environnement/modes-transmission>
8. (<https://www.weblex.fr/>).
9. (<https://graphics.reuters.com/>).
10. Salzberger B, Buder F, Lampl B, Ehrenstein B, Hitzenbichler F, Holzmann T, et al. Epidemiology of SARS-CoV-2. *Infection*. 2020:1-7. Repéré sur : <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7543961/>

## Références bibliographiques

---

11. Adam DC, Wu P, Wong JY, Lau EHY, Tsang TK, Cauchemez S, et al. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. *Nature Medicine*. 2020;26(11):1714-9. Repéré sur : <https://doi.org/10.1038/s41591-020-1092-0>.
12. Fontana L, Villamagna AH, Sikka MK, McGregor JC. Understanding viral shedding of SARS-CoV-2: review of current literature. *Infection Control & Hospital Epidemiology*. 2020:1-35. Repéré sur : <https://doi.org/10.1017/ice.2020.1273>
13. Mizrahi B, Shilo S, Rossman H, Kalkstein N, Marcus K, Barer Y, et al. Longitudinal symptom dynamics of COVID-19 infection in primary care. medRxiv. 2020. Repéré sur : <https://www.medrxiv.org/content/10.1101/2020.07.13.20151795v2>
14. Centers for Disease Control and Prevention (CDC). Interim Guidelines for COVID-19 Antibody Testing 2020. Version du 1er août 2020. Repéré sur : [https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antibody-testsguidelines.html#anchor\\_1590264247573](https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antibody-testsguidelines.html#anchor_1590264247573)
15. Tsai, P.H.; Wang, M.L.; Yang, D.M.; Liang, K.H.; Chou, S.J.; Chiou, S.H.; Lin, T.H.; Wang, C.T.; Chang, T.J. Genomic variance of Open Reading Frames (ORFs) and Spike protein in severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). *J. Chin. Med. Assoc.* 2020, 83, 725–732
16. Angeletti, S.; Benvenuto, D.; Bianchi, M.; Giovanetti, M.; Pascarella, S.; Ciccozzi, M. COVID-2019: The role of the nsp2 and nsp3 in its pathogenesis. *J. Med. Virol.* 2020, 92, 584–588.
17. Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding. *Lancet* 2020, 395, 565–574
18. Pachetti, M.; Marini, B.; Benedetti, F.; Giudici, F.; Mauro, E.; Storici, P.; Masciovecchio, C.; Angeletti, S.; Ciccozzi, M.; Gallo, R.C.; et al. Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. *J. Transl. Med.* 2020, 18, 1–9.
19. Luca, P. et al. Immunity, endothelial injury and complement-induced coagulopathy in COVID-19, (2020)
20. Wang, D. et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* 323, 1061–1069 (2020)

## Références bibliographiques

---

21. Vabret, N. et al. Immunology of COVID-19: current state of the science. *Immunity* 52, 910–941 (2020)
22. Qin, C. et al. Dysregulation of immune response in patients with COVID-19 in Wuhan, China. *Clin. Infect. Dis.* 71, 762–768 (2020)
23. Grifoni, A. et al. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* 181, 1489–1501.e15 (2020).
24. Noris, M., Benigni, A. & Remuzzi, G. The case of complement activation in COVID-19 multiorgan impact. *Kidney Int.* 98, 314–322 (2020).
25. Su, H. et al. Renal histopathological analysis of 26 postmortem findings of patients with COVID-19 in China. *Kidney Int.* 98, 219–227 (2020).
26. Larsen, C. P., Bourne, T. D., Wilson, J. D., Saqqa, O. & Sharshir, M. A. collapsing glomerulopathy in a patient with COVID-19. *Kidney Int. Rep.* 5, 935–939 (2020).
27. 88. Gao J, Tian Z, Yang X. Breakthrough: chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. *Biosci Trends.* (2020) 14:72–3. doi: 10.5582/bst.2020.01047
28. 199. Maisonnasse P, Guedj J, Contreras V, Behillil S, Solas C, Marlin R, et al. Hydroxychloroquine use against SARS-CoV-2 infection in non-human primates. *Nature.* (2020) 585:584–7. doi: 10.1038/s41586-020-2558-4
29. 211. Wang Y, Zhang D, Du G, Du R, Zhao J, Jin Y, et al. Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. *Lancet.* (2020) 395:1569–78. doi: 10.1016/S0140-6736(20)31022-9
30. 227. Freedberg DE, Conigliaro J, Wang TC, Tracey KJ, Callahan MV, Abrams JA, et al. Famotidine use is associated with improved clinical outcomes in hospitalized COVID-19 patients: a propensity score matched retrospective cohort study. *Gastroenterology.* (2020) 159:1129–31.e1123. doi: 10.1053/j.gastro.2020.05.053
31. 228. Bestle D, Heindl MR, Limburg H, Van Lam van T, Pilgram O, Moulton H, et al. TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Sci Alliance.* (2020) 3:e202000786. doi: 10.26508/lsa.202000786

## Références bibliographiques

---

- 32.** 233. Baum A, Fulton BO, Wloga E, Copin R, Pascal KE, Russo V, et al. Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science*. (2020) 369:1014–8. doi: 10.1126/science.abd0831
- 33.** Y. Wan, J. Shang, R. Graham, R.S. Baric, F. Li, Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus, *J. Virol.* 94 (2020).
- 34.** T. Tang, M. Bidon, J.A. Jaimes, G.R. Whittaker, S. Daniel, Coronavirus membrane fusion mechanism offers a potential target for antiviral development, *Antivir. Res.* 178 (2020) 104792.
- 35.** D. van Riel, E. de Wit, Next-generation vaccine platforms for COVID-19, *Nat. Mater.* 19 (2020) 810–812 Kyung Soo Park, Xiaoqi Sun, Marisa E. Aikins, James J. Moon, Non-viral COVID-19 vaccine delivery systems, 2020 Elsevier
- 36.** A. Koirala, Y.J. Joo, A. Khatami, C. Chiu, P.N. Britton, Vaccines for COVID-19: The current state of play, *Paediatr. Respir. Rev.* 35 (2020) 43–49
- 37.** T.T. Le, Z. Andreadakis, A. Kumar, R.G. Roman, S. Tollefsen, M. Saville, S. Mayhew, The COVID-19 vaccine development landscape, *Nat. Rev. Drug Discov.* 19 (2020) 305–306.
- 38.** L. Du, Y. He, Y. Zhou, S. Liu, B.J. Zheng, S. Jiang, The spike protein of SARS-CoV--a target for vaccine and therapeutic development, *Nat. Rev. Microbiol.* 7 (2009) 226–236.
- 39.** K.A. Hajj, K.A. Whitehead, Tools for translation: non-viral materials for therapeutic mRNA delivery, *Nat. Rev. Mater.* 2 (2017) 1–17.
- 40.** M. Alberer, U. Gnad-Vogt, H.S. Hong, K.T. Mehr, L. Backert, G. Finak, R. Gottardo, M.A. Bica, A. Garofano, Sven Dominik Koch, Mariola Fotin-Mleczek, Ingmar Hoerr, Ralf Clemens, Frank von Sonnenburg, Safety and immunogenicity of a mRNA rabies vaccine in healthy adults: an open-label, non-randomised, prospective, first-in-human phase 1 clinical trial, *Lancet* 390 (2017) 1511–1520
- 41.** A.M. Reichmuth, M.A. Oberli, A. Jaklenec, R. Langer, D. Blankschtein, mRNA vaccine delivery using lipid nanoparticles, *Ther. Deliv.* 7 (2016) 319–334.
- 42.** Lisa A. Jackson, Evan J. Anderson, Nadine G. Rouphael, Paul C. Roberts, Mamodikoe Makhene, Rhea N. Coler, Michele P. McCullough, James D. Chappell, Mark R. Denison, Laura J. Stevens, Andrea J. Pruijssers, Adrian McDermott, Britta Flach, Nicole A. Doria-Rose, Kizzmekia S. Corbett, Kaitlyn M. Morabito, Sijy O'Dell, Stephen D. Schmidt, Phillip A. Swanson, Marcelino Padilla, John R.

## Références bibliographiques

---

Mascola, Kathleen M. Neuzil, Hamilton Bennett, Wellington Sun, Etza Peters, Mat Makowski, Jim Albert, Kaitlyn Cross, Wendy Buchanan, Rhonda Pikaart-Tautges, Julie E. Ledgerwood, Barney S. Graham, John H. Beigel, An mRNA vaccine against SARSCoV-2 - preliminary report, *N. Engl. J. Med.* 383 (2020) 1920–1931.

**43.** M.J. Mulligan, K.E. Lyke, N. Kitchin, J. Absalon, A. Gurtman, S.P. Lockhart, K. Neuzil, V. Raabe, R. Bailey, K.A.J.M. Swanson, Phase 1/2 study to describe the safety and immunogenicity of a COVID-19 RNA vaccine candidate (BNT162b1) in adults 18 to 55 years of age: interim report, *Nature* 586 (2020) 589–593.

**44.** S. Persano, M.L. Guevara, Z. Li, J. Mai, M. Ferrari, P.P. Pompa, H. Shen, Lipopolyplex potentiates anti-tumor immunity of mRNA-based vaccination, *Biomaterials* 125 (2017) 81–89.

**45.** <https://ourworldindata.org/>

**46.** [[https://en.wikipedia.org/wiki/Basic\\_Local\\_Alignment\\_Search\\_Tool](https://en.wikipedia.org/wiki/Basic_Local_Alignment_Search_Tool)]

**47.** [[https://en.wikipedia.org/wiki/BLAST\\_\(biotechnology\)](https://en.wikipedia.org/wiki/BLAST_(biotechnology))]

**48.** [[https://community.gep.wustl.edu/repository/course\\_materials\\_WU/annotation/Introduction\\_NCBI\\_BLAST.pdf](https://community.gep.wustl.edu/repository/course_materials_WU/annotation/Introduction_NCBI_BLAST.pdf)]

[[https://community.gep.wustl.edu/repository/course\\_materials\\_WU/annotation/Introduction\\_NCBI\\_BLAST.pdf](https://community.gep.wustl.edu/repository/course_materials_WU/annotation/Introduction_NCBI_BLAST.pdf)]

**50.** [[https://en.wikipedia.org/wiki/Dot\\_plot\\_\(bioinformatics\)](https://en.wikipedia.org/wiki/Dot_plot_(bioinformatics))]

**51.** Peng Zhou, Xing-Lou Yang, Xian-Guang Wang, Ben Hu, Lei Zhang, Wei Zhang et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* | Vol 579 | 12 March 2020. <https://doi.org/10.1038/s41586-020-2012-7>.

**52.** Peng Zhou, Xing-Lou Yang, Xian-Guang Wang, Ben Hu, Lei Zhang, Wei Zhang et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 270-273/17/11.2020

**53.** Longxian Lv, Gaolei Li, Jinhui Chen, Xinle Liang and Yudong Li. Comparative Genomic Analyses Reveal a Specific Mutation Pattern Between Human Coronavirus SARS-CoV-2 and Bat-CoV RaTG13. *Front. Microbiol.* 11:584717. doi: 10.3389/fmicb.2020.584717.

**54.** Pei Li, Ruixuan Guo, Yan Liu, Yingtao Zhang, Jiabin Hua et al. Le Rhinolophus affinis bat ACE2 et plusieurs orthologues animaux sont des récepteurs fonctionnels pour le coronavirus de chauve-souris RaTG13 et SARS-CoV-2. *Bulletin scientifique* 66 (2021) 1215-1227.

**55.** <https://doi.org/10.1016/j.scib.2021.01.011>.

## Références bibliographiques

---

56. (<https://www.ncbi.nlm.nih.gov/protein>)
57. <https://doi.org/10.1007/s42485-021-00060-3>.
58. Lesk, A. M. (26 juillet 2013). "Bioinformatique". Encyclopédie Britannica.
59. joel b. hagen the origins of bioinformatics, nature reviews genetic 01 December 2000
60. [Dayhoff and Ledley. Comprotein. New York, NY: ACM,1962,262–74].
61. Jeff Gauthier et al. "A brief history of bioinformatics". Brief Bioinform. 2019 Nov 27;20(6):1981-1996. doi: 10.1093/bib/bby063. PMID: 30084940
62. Jeff Gauthier et al. "A brief history of bioinformatics". Brief Bioinform. 2019 Nov 27;20(6):1981-1996. doi: 10.1093/bib/bby063. PMID: 30084940
63. Jeff Gauthier et al. "A brief history of bioinformatics". Brief Bioinform. 2019 Nov 27;20(6):1981-1996. doi: 10.1093/bib/bby063. PMID: 30084940
64. Mardis E. A decade's perspective on dna sequencing technology. nature, 470: 198-203. 2011
65. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res. 2010;38:D46–D51.
66. <https://www.ncbi.nlm.nih.gov/genbank/samplerrecord/>
67. <http://genetics.bwh.harvard.edu/pph/FASTA.html>
68. <https://www.ncbi.nlm.nih.gov/tools/sviewer/graphicalpanel/>
69. Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, Dicuccio M, Edgar R, Federhen S, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2010;38:D5–D16]
70. K B Mulis, F A Faloon. Specific synthesis of DNA in vitro via polymerase-catalyzed chain reaction. Methods Enzymol 1987; 155:335-50.
71. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. [Nucleic Acids Res, 2010.
72. [https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE\\_TYPE=BlastDocs&DOC\\_TYPE=BlastHelp](https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=BlastHelp).