

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université de Blida 1



Faculté des sciences de la nature et de la vie
Département biologie

Mémoire de fin d'étude

En vue de l'obtention du diplôme de Master
en Sciences de la nature et de la vie

Option : Génétique

Thème :

**Étude comparative entre des séquences d'ADN de
Coronaviridae chez l'homme et chez d'autres espèces
d'animaux.**

Présenter par :

M^{elle} ASSEME Imane

M^{elle} SAIDI Manel

Devant le jury :

Mme AMOKRANE A.

M.C.B USDB 1

Présidente

Mme EDDAKRA A.

M.C.B USDB 1

Examinatrice

Mme DEROUICHE L.

M.C.A ESSAIA

Promotrice

Mme ZATRA Y.

M.C.B USDB 1

Co-promoteur

Promotion : 2021/2022

Remerciements

*Nous remercions **ALLAH** le tout puissant de nous avoir donnée la santé, le courage et la patience d'entamer ce mémoire.*

*Nous adressons nos remerciements à toutes les personnes qui nous ont aidées à réaliser ce travail. En premier lieu notre promotrice **Mme DERIOUCHE** maitre de conférences à l'école supérieure des sciences de l'aliment et de l'industrie agroalimentaires, Alger. Pour sa patience, sa qualité d'encadrement et sa disponibilité durant notre préparation de ce mémoire. Ainsi que notre co-promotrice **Mme ZATRA** maitre de conférences au département de biologie à Blida 1.*

*Notre remerciement s'adresse aux membres de jury : la présidente **Mme AMOKRANE** maitre de conférences au département de biologie à Blida 1, pour avoir accepté de présider ce jury. A l'examinatrice **Mme EDDAIKRA** maitre de conférences au département de biologie à Blida 1, pour avoir accepté d'évaluer ce travail et pour toutes leurs remarques et critiques.*

A ne jamais oublier nos enseignants pour leur générosité et la grande patience dont ils ont fait preuves malgré leurs charges académique et professionnel.

Merci à tous.

Dédicaces

Je dédie ce mémoire à tous ceux qui me sont chers,

***A mes chers parents
ma mère Hamida et mon père
Mohamed***

*Aucune dédicace ne saurait exprimer mon respect, mon amour
et ma considération pour les sacrifices que vous avez consenti
pour mon instruction et mon bien être.*

*Je vous remercie pour tout le soutien et l'amour que vous me
portez depuis mon enfance et j'espère que votre bénédiction
m'accompagnera toujours. Que ce modeste travail soit
l'exaucement de vos vœux tant formulés, le fruit de vos
innombrables sacrifices. Puisse Dieu, le très haut, vous accorde
santé, bonheur et longue vie.*

***A mes chers
Frères Zinou, Abdou et ma sœur Malak
Mon époux Hichem
Ma belle-sœur Nisrine et le neveu
Amirou***

*De ma profonde estime pour l'aide que vous m'apporterez. Vous
me soutenez, réconfortez et encouragez. Puissent nos liens
fraternels se consolider et se renforcer encore plus.*

***À Ma Chère Famille
SELLAMI et ASSEME***

*Je vous remercie énormément personne par
personne de votre soutien et votre
encouragement durant mon long parcours,
que dieux vous protège.*

À Mes chers Amis

*Ma chère binôme Manel, Assia et Asma. A travers mon travail,
je vous transmets mes meilleurs sentiments d'amour.*

Imane.

Dédicaces

Je dédie ce mémoire à tous ceux qui mon chère,

Je remercie ALLAH

Le tout puissant qui nous guide dans notre vie, qui nous aide dans notre chemin je lui dois ce que je suis devenue, ce travail est de lui et grâce à lui.

A Mes Chères parents

Ma mère Dalila et mon père Mohamed

Pour leur patience, leur amour, leur soutien et leurs encouragements.

Aucun hommage ne pourrait être à la hauteur de l'amour dont ils ne cessent de me combler. Que Dieu, vous leur procure la santé, le bonheur et une longue vie.

Et mon défunt grand-père

Qui nous a quittés tôt, que Dieu bénisse son âme avec miséricorde

A ma chère

Sœur Imane

Qui m'a aidé et soutenu profondément précisément dans cette période de réalisation de ce travail. Et mon cher frères Abderrahmane

À Mes chers Amis

Ma chère binôme Imane je vous remercie je te souhaite le meilleur, à mes amis Assia et Asma. Et mon ami de temps d'année Imane A travers mon travail, je vous transmets mes meilleurs sentiments d'amour.

Celui qui ne remercie pas les gens ne remercie pas Dieu

Manel.

Sommaire

INTRODUCTION

CHAPITRE I : RECHERCHE BIBLIOGRAPHIQUES

I.1. Présentation du Coronavirus.....	1
I.1.1. Présentation générale du Coronavirus.....	1
I.1.2. Historique	1
I.1.3. Structure et génome du Coronavirus.....	2
I.1.4. Taxonomie et classification des coronavirus.....	4
I.1.4.1. Structure de l'ordre des Nidovirales.....	5
I.1.5. Coronavirus chez l'hôte.....	6
I.1.5.1 Le Coronavirus dans la cellule.....	7
I.1.5.2. Réplication du génome de Coronavirus.....	7
I.1.5.3. Transcription du Coronavirus.....	8
I.1.5.4. Incubation du Coronavirus	8
I.1.6. Transmission du Coronavirus.....	8
I.1.6.1. Transmission entre les espèces.....	8
I.2. Présentation de la phylogénie.....	10
I.2.1. Arbre phylogénétique.....	11
I.2.2. Les méthodes de reconstruction phylogénétique et les logiciels utilisés.....	12
I.2.2.1. Les méthodes de reconstruction phylogénétique	13
I.2.2.1.1. Méthode de parcimonie.....	13
I.2.2.1.2. Méthode Bayésien inférence	13
I.2.2.1.3. Méthode de Maximum de vraisemblance.....	14

CHAPITRE II : MATERIELS ET METHODES

II.1. Matériel.....	15
II.1.1. Matériel Biologique	15
II.1.2. Matériel non Biologique.....	20
II.2. Méthode.....	21

II.2.1. Outils Bioinformatique	21
II.2.1.1. Outil GenBank.....	21
II.2.1.2. Outil de recherche d'alignement local de base (BLAST).....	28
II.3. Traitement des séquences	32
II.3.1. Outils de MEGA 11.....	32
II.3.2. Outils de GeniGen 2.....	36
 CHAPITRE III : RESULTATS	
III.1. Résultats de l'arbre du gène la glycoprotéine S	43
III.2. Résultats de l'arbre du gène de la protéine N	51
III.2. Résultats de l'arbre du gène RdRp	59
 CHAPITRE IV : DISCUSSION	
IV.1. Comparaison des souches de l'arbre du gène de pointe S (gène structurale)	68
IV.1.1. Comparaison des souches de l'arbre du groupe Alphacoronavirus du gène de pointe S	68
IV.1.2. Comparaison des souches de l'arbre du groupe Deltacoronavirus du gène de pointe S.....	70
IV.1.3. Comparaison des souches de l'arbre du groupe Gammacoronavirus du gène de pointe S.....	71
IV.1.4. Comparaison des souches de l'arbre du groupe Betacoronavirus du gène de pointe S.....	71
IV.2. Comparaison des souches de l'arbre du gène de pointe N (gène structurale)	73
IV.2.1. Comparaison des souches de l'arbre du groupe Betacoronavirus du gène N (gène de nucléocapside).....	73
IV.2.2. Comparaison des souches de l'arbre du groupe Alphacoronavirus du gène N (gène de nucléocapside).....	75
IV.2.3. Comparaison des souches de l'arbre du groupe Gammacoronavirus du gène N (gène de nucléocapside)	75
IV.2.4. Comparaison des souches de l'arbre du groupe Deltacoronavirus du gène N (gène de nucléocapside)	75
IV.3. Comparaison des souches de l'arbre du gène RdRp (gène non structurale)	76
IV.3.1. Comparaison des souches de l'arbre du groupe Betacoronavirus du gène RdRp	76

IV.3.2. Comparaison des souches de l'arbre du groupe Gammacoronavirus du gène RdRp...	78
IV.3.3. Comparaison des souches de l'arbre du groupe Alphacoronavirus du gène RdRp.....	78
IV.2.1. Implication des mutations dans l'évolution.....	79
IV.2.2. Conséquences des mutations.....	80

CONCLUSION

RÉFÉRENCES BIBLIOGRAPHIQUES

ANNEXE

Abréviations

- ACE2:** Angiotensin Converting Enzyme 2.
- BCoV:** Bovine Coronavirus.
- CCoV:** Canine Coronavirus.
- CSG:** Coronavirus Study Group
- DAMBE:** Data Analysis in Molecular Biology and Evolution.
- DCoV:** Duck Coronavirus.
- DDBJ:** Japanese Data Banque.
- DNASP :** DNA Sequence Polymorphism.
- EBI:** European Bioinformatics Institute.
- ECoV:** Equine Coronavirus.
- Fabox:** FASTA Sequence Tool BOX.
- FIPV:** Feline Infectious Peritonitis Virus.
- FCoV :** Feline Coronavirus.
- GiCoV:** Giraffe Coronavirus.
- HCoV-229E:** Human Coronavirus 229E.
- HCoV-HKU1:** Human Coronavirus HKU1.
- HCoV-NL63:** Human Coronavirus NL63.
- HCoV-OC43:** Human Coronavirus OC43.
- Hi-BatCoV HKU10:** Hipposideros Bat Coronavirus HKU10.
- IBV:** Infectious Bronchitis Virus.
- ICTV:** International Committee on Taxonomy of Viruses.
- LANL:** Los Alamos National Laboratory.
- MCMC:** Bayesian Markov Chain Monte Carlo Method.
- MEGA:** Molecular Evolutionary Genetics Analysis.
- MERS-CoV:** Middle East respiratory syndrome coronavirus.
- MHV:** Murine Hepatitis Virus.

ML: Maximum Likelihood.

MRCA: Most Recent Common Ancestor of mammalian CoVs.

NCBI: National Center for Biotechnology Information.

NJ: Arithmetic Averages.

ORF: Open Reading Frame.

PCoV: Pigeon Coronavirus.

PDB: Protein Data Bank.

PEDV: Porcine Epidemic Diarrhea Virus.

PHEV: Porcine Hemagglutinating Encephalomyelitis Virus.

PHYLIP : Phylogeny Inference Package.

RBD: Receptor Binding Domain.

RCoV: Rat Coronavirus.

RdRp: RNA-dependent RNA polymerase.

RDRP: RNA-Dependent RNA Polymerase.

RE: Endoplasmic Reticulum.

SARS: Severe Acute Respiratory Syndrome.

SGD: Saccharomyces Genome Database

TCoV: Turkey Coronavirus.

TEGV: Transmissible Gastroenteritis Virus.

UPGMA: Unweighted Pair Group Method with Arithmetic mean.

WHO: World Health Organization.

WiCoV HKU20: American Duck Coronavirus HKU20.

Liste des figures

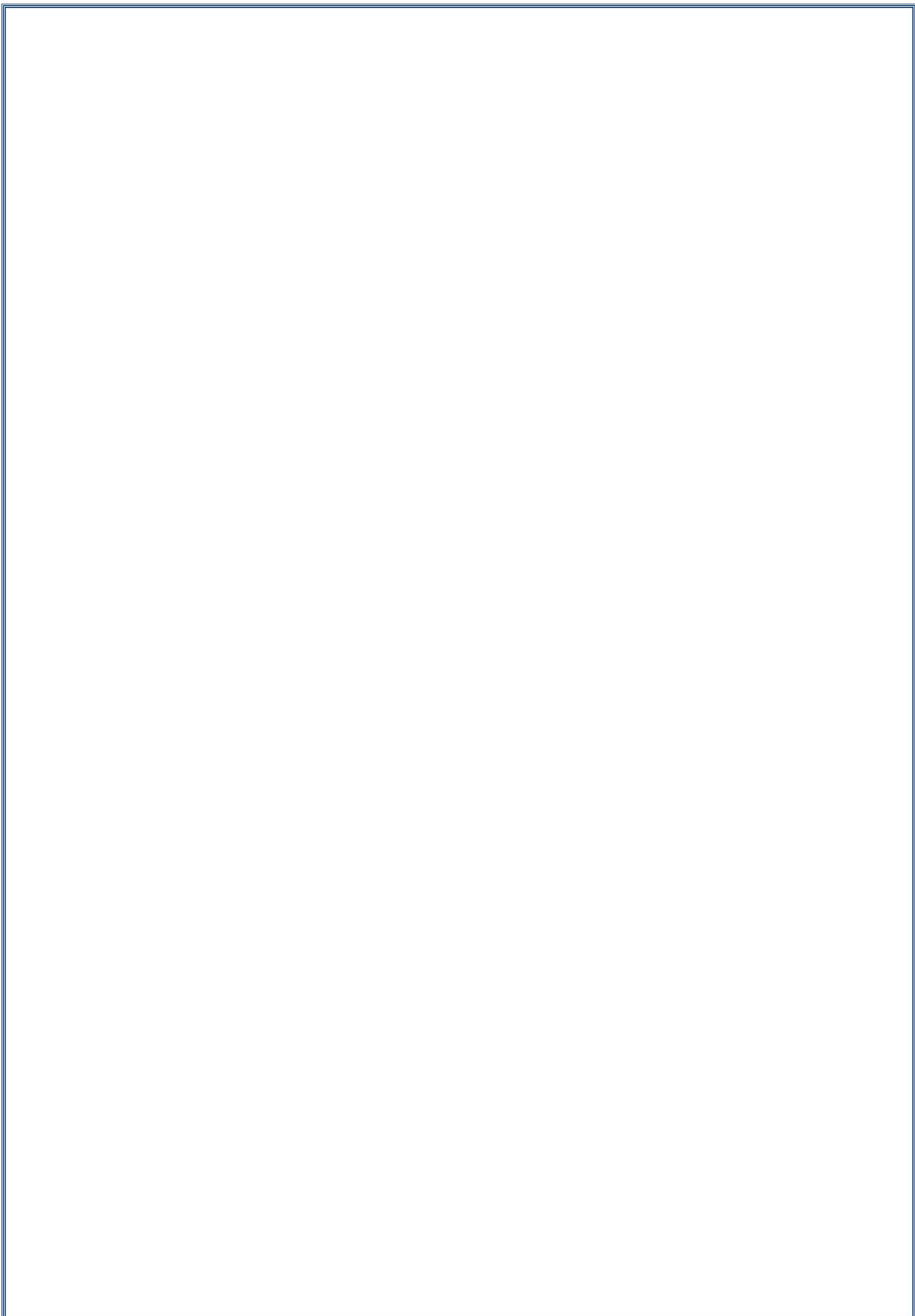
N° de figure	Titre	Page
1	Structure de Coronavirus	3
2	Annotation du génome	3
3	Schéma structurel de l'ordre des Nidovirales	5
4	Classification des Nidovirus et membres prototypes	6
5	Cycle viral du Coronavirus	7
6	Les chauves-souris et les oiseaux comme sources de gènes probables pour l'évolution des coronavirus, sur la base de preuves épidémiologiques, virologiques et phylogénétiques	10
7	Arbre phylogénétique non enraciné	12
8	Arbre phylogénétique enraciné	12
9	Page d'accueil de GenBank.	22
10	Résultat de la recherche de la séquence dans la GenBank.	23
11	les informations du gène dans la GenBank.	23
12	Forme FASTA de la séquence affichée par la GenBank.	24
13	Enregistrement de la forme FASTA.	24
14	Classement des données de la séquence dans la GenBank.	25
15	Résultats de la recherche en texte intégral PMC.	25
16	Résultat de recherche PopSet de la séquence dans la GenBank.	26
17	Résultat de la recherche du groupe de protéines identiques à la séquence.	26
18	Informations de la séquence dans la GenBank.	27
19	Les deux formes d'affichage de la séquence.	27
20	Exemple du graphique d'un gène S de la séquence.	28
21	Page d'accueil de BLAST.	29
22	Recherche des séquences à haute similarités.	29
23	Classement des séquences similaires dans la description.	30
24	Classement des séquences similaire par le résumé graphique.	30
25	Alignement des séquences par paire avec ou sans fonctions CDS.	31

26	Taxonomie des trois cas de figure d'affichage proposé par BLAST.	31
27	Page d'accueil et la barre d'outils de MEGA 11.	32
28	Fenêtre « Alignement Explorer » de MEGA 11.	33
29	Séquences d'ADN format FASTA	33
30	Alignement des séquences par MEGA 11.	34
31	L'ajustement des paramètres de construction de l'arbre par MEGA 11.	36
32	Page d'accueil de GeniGen 2.	37
33	Séquences FASTA dans la GenieGen 2.	37
34	Suppression des séquences non ciblées.	38
35	Renommer les séquences ciblées.	38
36	Choisir un nouveau nom.	39
37	Alignement des séquences sélectionnées.	39
38	Alignement des séquences par GeniGen 2.	40
39	Action d'affichage du tableau de comparaison.	40
40	Tableau de comparaison de GeniGen 2.	41
42	L'arbre phylogénétique du gène S.	43
43	Fragment de l'arbre du gène S comportant les séquences du sous-groupe αa .	44
44	Fragment de l'arbre du gène S comportant les séquences du sous-groupe αb .	45
45	Fragment de l'arbre du gène S comportant les séquences du groupe δ .	48
46	Fragment de l'arbre du gène S comportant les séquences du groupe γ .	48
47	Fragment de l'arbre du gène S comportant les séquences du sous-groupe βb .	47
48	Fragment de l'arbre du gène S comportant les séquences du sous-groupe βa .	47
49	L'arbre phylogénétique du gène N.	53
50	Fragment de l'arbre du gène N comportant les séquences du sous-groupe βa	54
51	Fragment de l'arbre du gène N comportant les séquences du sous-	54

	groupe β b.	
52	Fragment de l'arbre du gène N comportant les séquences du sous-groupe α a.	55
53	Fragment de l'arbre du gène N comportant les séquences du sous-groupe α b.	55
54	Fragment de l'arbre du gène N comportant les séquences du groupe γ .	56
55	Fragment de l'arbre du gène N comportant les séquences du groupe δ .	56
56	L'arbre phylogénétique du gène RdRp.	61
57	Fragment de l'arbre du gène RdRp comportant les séquences du groupe β .	62
58	Fragment de l'arbre du gène RdRp comportant les séquences du groupe γ .	63
59	Fragment de l'arbre du gène RdRp comportant les séquences du sous-groupe α a.	63
60	Fragment de l'arbre du gène RdRp comportant les séquences du sous-groupe α b.	64

Liste des tableaux

N° du tableau	Titre	Page
I	Protéines structurales du coronavirus et leurs fonctions.	4
II	Les séquences d'ADN de Coronavirus utilisé pour la présentation phylogénétique.	15
III	les logiciels utilisés dans l'étude phylogénétique.	20
IV	Les séquences d'ADN de Coronavirus utilisé pour la présentation phylogénétique du gène S.	40
V	Pourcentages d'identités des différentes séquences de souches du groupe Alphacoronavirus (α CoV) qui présentent le gène S.	49
VI	Pourcentages d'identités de différentes séquences de souches du groupe Deltacoronavirus (δ CoV) qui présentent le gène S.	50
VII	Pourcentages d'identités de différentes séquences de 4 souches du groupe Gammacoronavirus (γ CoV) qui présentent le gène S.	50
VIII	Pourcentages d'identités de différentes séquences de 15 souches du groupe Gammacoronavirus (γ CoV) qui présentent le gène S.	51
IX	Les séquences d'ADN de Coronavirus utilisé pour la présentation phylogénétique du gène N.	52
X	Pourcentages d'identités de différentes séquences de souches du groupe Betacoronavirus (Beta CoV) qui présentent le gène N.	57
XI	Pourcentages d'identités de différentes séquences de souches du groupe Alphacoronavirus (Alpha CoV) qui présentent le gène N.	58
XII	Pourcentages d'identités de différentes séquences de souches du groupe Deltacoronavirus (δ CoV) qui présentent le gène N.	59
XIII	Les séquences d'ADN de Coronavirus utilisées pour la présentation phylogénétique du gène RdRp.	59
XIV	Pourcentages d'identités de différentes séquences de souches du groupe Betacoronavirus (Beta CoV) qui présentent le gène RdRp.	66
XV	Pourcentages d'identités de différentes séquences de souches du groupe Gammacoronavirus (γ CoV) qui présentent le gène RdRp.	66
XVI	Pourcentages d'identités de différentes séquences de souches du groupe Alphacoronavirus (Alpha CoV) qui présentent le gène RdRp.	67



Résumé

Les Coronavirus sont des virus à ARN simple brin sens positif, leur nom est dû à l'apparence d'une couronne solaire. Dans cette étude nous avons réalisé une comparaison par une analyse phylogénétique basée sur des séquences nucléotidiques de différentes souches virales extraites de la GenBank qui ont touchés les animaux ainsi que l'homme. Pour réaliser ce travail, des séquences extraites des trois gènes, gène de l'ARN polymérase ARN-dépendante RdRp, gène de nucléocapside N et gène de pointe Sont été analysées à l'aide du logiciel MEGA 11. Nous avons classé les souches de virus pour ces trois gènes selon les espèces animales qu'ils ont infectées et les régions géographiques dans lesquelles ils ont été retrouvés.

Commençant par le gène S dont nous avons travaillé sur 49 séquences nucléotidiques de différentes souches à travers plusieurs pays comme les Pays Bas, le Japon et les Etats-Unis qui touche plusieurs espèces telles que, HCoV 229E, HCoV NL63 infectant l'homme et les chauves-souris ; PRCV et PDCoV infectant le porc. Nous avons trouvé qu'il y a une accumulation des mutations au sein du gène S entre les différentes souches qui peuvent provoquer une transmission du virus de l'animal à l'homme.

L'analyse du gène RdRp a été basé sur 35 séquences nucléotidiques de différentes souches virales ; les résultats obtenus prouvent la présence des souches BatCoV infectant les chauves-souris dans Mayotte et Mozambique, RtCoV infectant les rongeurs en Chine, Slovénie, PHEV qui touche le porc en Italie, BCoV touchant les bovins en Slovénie et FRSCV des furets au Pérou.

Nous avons traité 36 séquences nucléotidiques du gène N de différentes souches virales telles que SARS-CoV-2 qui touche les chauves-souris, HCoV OC43 infectant l'homme, PEDV et PDCoV qui infecte le porc, ces derniers se partagent entre plusieurs pays comme la Chine, l'Allemagne, la Belgique, la France et la Corée du sud.

Sachant que le Coronavirus est considéré comme l'un des plus anciens virus, où il n'est pas surprenant que les mutations de ces gènes conduisent à l'apparition des différentes souches de virus affectant différentes espèces dans des zones géographiques variables. Dans les résultats de notre étude ; nous a démontré que la diversité génétique entre les souches de ce virus est très élevée que ce soit pour les souches qui infectent les animaux ou celles qui infectent l'homme dans différents pays du monde.

Mots clé : Coronavirus, gène, phylogénie.

Abstract

Coronaviruses are positive-sense single-stranded RNA viruses; their name is due to the appearance of a solar corona. In this study, we performed a comparison by phylogenetic analysis based on nucleotide sequences of different viral strains extracted from the GenBank that affected animals as well as humans. To carry out this work, sequences extracted from the three genes, RNA-dependent RNA polymerase gene RdRp, nucleocapsid gene N and spike gene S were analyzed using MEGA 11 software. Here, we have classified the virus strains for these three genes according to the animal species they infected. In addition, the geographic regions in which they found them.

Starting with the S gene where we worked on 49 nucleotide sequences of different strains across several countries such as the Netherlands, Japan and the United States, which affects several species such as, HCoV 229E, HCoV NL63 infecting humans and bats; PRCV and PDCoV infecting pigs. We have found that the transmission of the virus from animals to humans there due an accumulation of mutations within the S gene between different strains.

The analysis of the RdRp gene was based on 35 nucleotide sequences from different viral strains; the results obtained prove the presence of BatCoV strains infecting bats in Mayotte and Mozambique, RtCoV infecting rodents in China, Slovenia, PHEV in pigs in Italy, BCoV affecting cattle in Slovenia and FRSCV in ferrets in Peru.

We have processed 36 nucleotide sequences of the N gene of different viral strains such as SARS-CoV-2 which affects bats, HCoV OC43 which infects humans, PEDV and PDCoV which infects pigs, the latter being shared between several countries as China, Germany, Belgium, France, and South Korea.

Knowing that the Coronavirus is considered one of the oldest viruses, where it is not surprising that the mutations of these genes lead to the appearance of the different strains of viruses affecting different species in variable geographical areas, the results of which of these three genes; have shown us that the genetic diversity between the strains of this virus is very high, whether for the strains that infect animals or those that infect humans in different countries of the world.

Keywords: Coronavirus, gene, phylogeny

ملخص

فيروسات كورونا هي فيروسات RNA أحادية الشريطة إيجابية الاتجاه ، ويرجع اسمها إلى ظهور الهالة الشمسية الخاصة بشكلها. أجرينا في هذه الدراسة مقارنة عن طريق تحليل التطور العرقي بناءً على تسلسلات نيكلوتيدية لسلاسل فيروسية مختلفة أثرت على الحيوانات وكذلك البشر. لتنفيذ هذا العمل ، استخدمنا برنامج MEGA 11 لتحليل التسلسلات المستخرجة من GenBank لثلاثة جينات RdRp و N و S ؛ حيث صنفنا السلالات الفيروسية لهذه الجينات الثلاثة حسب الأنواع الحيوانية المصابة والمنطقة الجغرافية التي تم العثور عليها فيها.

بدأنا من الجين S حيث عملنا على 49 تسلسلاً نيكلوتيدي من سلالات مختلفة مثل HCoV 229E و HCoV NL63 التي تصيب البشر والخفافيش اما PRCV و PDCoV فيصيان الخنازير ، في العديد من البلدان مثل هولندا واليابان والولايات المتحدة. حيث وجدنا أن هناك تراكماً للطفرات داخل الجين S بين السلالات المختلفة التي يمكن أن تتسبب في انتقال الفيروس من الحيوانات إلى البشر.

استند تحليل جين RdRp على 35 تسلسل نيكلوتيدي من سلالات فيروسية مختلفة. الذي أثبتت نتائجه وجود سلالات BatCoV التي تصيب الخفافيش في مايوت وموزمبيق ، RtCoV التي تصيب القوارض في الصين وسلوفينيا ، PHEV بالنسبة للخنازير في إيطاليا و BCoV التي تصيب الماشية في سلوفينيا و FRSCV للقوارض في بيرو. كما قمنا بمعالجة 36 تسلسلاً نيكلوتيدياً من الجين N لسلاسل فيروسية مختلفة مثل SARS-CoV-2 الذي يصيب الخفافيش و HCoV OC43 الذي يصيب البشر و PEDV و وأخيراً PDCoV الذي يصيب الخنازير. فيما تبين ان هذه الفيروسات مشتركة بين عدة دول مثل الصين وألمانيا ، بلجيكا وفرنسا وكوريا الجنوبية.

مع العلم أن فيروس كورونا يعتبر من أقدم الفيروسات ، حيث لا غرابة في أن تؤدي طفرات هذه الجينات إلى ظهور سلالات مختلفة من الفيروسات التي تصيب بدورها أنواعاً مختلفة من الحيوانات في مناطق جغرافية متغيرة ، كما ان نتائج هذه الجينات الثلاثة ؛ أظهرت لنا أن التنوع الجيني بين سلالات هذا الفيروس مرتفع للغاية ، سواء بالنسبة للسلاسل التي تصيب الحيوانات أو تلك التي تصيب البشر في مختلف دول العالم.

الكلمات المفتاحية: فيروس كورونا ، ، الجينات ، علم تطور السلالات.

Introduction

La famille des *Coronaviridae* est une famille qui comporte un groupe monogénique de 11 virus infectant les vertébrés (**Siddell et Coll, 1983**). Selon le comité international de taxonomie des virus, les Coronavirus sont classés sous l'ordre des *Nidovirales*, famille des *Coronaviridae*, sous-famille des *Coronavirinae*. Sur la base de preuves sérologiques précoces et génomiques ultérieures, le *Coronavirinae* est divisé en quatre groupes : Alphacoronavirus, Betacoronavirus, Gammacoronavirus et Deltacoronavirus (**Jasper et al., 2013; Fung, 2019**) et chaque genre se divise en plusieurs souches virales telles que BCoV chez les bovins, CCoV chez les chiens, IBV chez les oiseaux (**Alluwaimi et al; 2020**). Ce virus est parmi les virus émergents les plus importants chez une grande variété d'espèces de mammifères ces dernières années, provoquant divers syndromes de maladies respiratoires, notamment le syndrome respiratoire aigu sévère (SRAS) et le syndrome respiratoire du Moyen-Orient (MERS) (**Marschang et al., 2019**).

Les coronavirus sont des virus enveloppés qui contiennent le plus grand génome linéaire d'ARN viral simple brin de sens positif, allant de 26 à 32 kb de longueur (**Kasmi et al., 2020 ; Naqvi et al., 2020**). L'ARN génomique du coronavirus est de composition constante entre leur virus ou il comprend deux principaux cadres de lecture ouverts (ORF) ; ORF1a et ORF1b, englobant les deux premiers tiers du génome (**Zhang et al., 2021**). Ces deux premiers tiers du génome composait de 2 grands cadres de lecture ouverts qui se chevauchent et qui codent pour 16 protéines non structurales (**Bergmann et al., 2020**). Le tiers restant du génome code pour des ORF qui se chevauchent, ainsi que quatre protéines structurales majeures y compris S (glycoprotéine de pointe), N (protéine de nucléocapside), M (protéine de membrane) et E (protéine d'enveloppe) (**Zhang et al., 2021**).

Avant 2003, il n'y avait que 10 Coronavirus avec des génomes complets disponibles. Après l'épidémie de SARS, jusqu'en décembre 2008, il y a eu un ajout de 16 Coronavirus avec des génomes complets séquencés. Ceux-ci comprennent deux Coronavirus humains (Coronavirus humain NL63 et Coronavirus humain HKU1), 10 autres Coronavirus de mammifères (Coronavirus du SRAS de chauve-souris ; Coronavirus de chauve-souris : Bat-CoV HKU2, Bat-CoV HKU4, Bat-CoV HKU5, Bat-CoV HKU8, Bat-CoV HKU9, Bat-CoV 512/2005, Bat-CoV 1A ; Coronavirus équin et Coronavirus de béluga) et quatre Coronavirus aviaires

(Coronavirus de dinde, Coronavirus bulbul HKU11, Coronavirus de muguet HKU12 et Coronavirus de Munia HKU13) (**Woo et al., 2012**).

Afin de mieux connaître les Coronavirus nous avons réalisé ce travail fondé sur une étude phylogénétique (**Kuhl et al., 2019**) de ce virus, classé comme le virus le plus célèbre et le plus diffus dans le monde entier actuellement. Afin de réaliser ce travail nous avons appliqué les éléments de base de la phylogénie qui reposent sur la construction des arbres phylogénétiques des différentes souches de Coronavirus (**Will et al., 2017**) par un traitement des séquences d'ADN et la comparaison entre ces souches qui ont touché plusieurs espèces.

Notre travail se divise en :

Chapitre I, recherche bibliographique, qui se divise en deux parties : présentation du Coronavirus ; présentation générale, histoire du virus, structure, taxonomie et classification, transmission du virus, propagation dans le monde ; et en parallèle nous avons étudié le Coronavirus en Algérie. Définitions des études phylogénétiques qui aident à la reconstruction phylogénétique des gènes. Exposition des différentes banques de données, les logiciels utilisés et les méthodes de reconstruction phylogénétique.

Le chapitre II, fournit le matériel et les méthodes que nous avons utilisées pour réaliser notre mémoire.

Le chapitre III, rapporte les résultats obtenus ainsi qu'une discussion à la lumière des données bibliographiques.

Enfin, on terminera ce mémoire par une conclusion.

RECHERCHE BIBLIOGRAPHIQUE

I.1. Présentation du Coronavirus

I.1.1. Présentation générale du Coronavirus

Les coronavirus sont parmi les virus émergents les plus importants chez une grande variété d'espèces de mammifères ces dernières années, provoquant divers syndromes de maladies respiratoires, notamment le syndrome respiratoire aigu sévère (SRAS) et le syndrome respiratoire du Moyen-Orient (MERS) (**Marschang et al., 2019**).

Fin décembre 2019, des informations ont fait état d'une mystérieuse pneumonie à Wuhan, en Chine, une ville de 11 millions d'habitants dans la province du Hubei, dans le sud-est du pays. La cause, ont rapidement déterminé les scientifiques chinois, était un nouveau coronavirus éloigné du virus du SRAS (qui avait émergé en Chine en 2003, avant de se propager dans le monde et de tuer près de 800 personnes). Six mois plus tard, plus de dix millions de cas confirmés, la pandémie de COVID-19 est devenue la pire crise de santé publique depuis un siècle (**Callaway et al., 2020**).

Depuis lors, les chercheurs étudiant les coronavirus confrontés à plusieurs reprises à la question de la nouveauté du coronavirus, grave, voire mortelle, a été introduite chez l'homme. D'un réservoir zoonotique : c'est le cas du syndrome respiratoire aigu sévère (SRAS) et, quelques années plus tard, du syndrome respiratoire du Moyen-Orient (MERS) (**Gorbalenya et al., 2020**).

I.1.2. Historique

Dans la recherche menée par (**Gozlan 2020**) l'historique de coronavirus se présente comme suite :

1930 : Le premier coronavirus découvert (Infectious Bronchitis Virus, IBV) est un virus animal, responsable chez les volailles d'une maladie respiratoire (bronchite infectieuse aviaire) qui a été observée pour la première fois dans le Dakota du Nord (Etats-Unis).

1931 : Arthur Frederick Schalk et MC Hawn identifié une nouvelle maladie respiratoire chez des poussins âgés de deux jours à trois semaines.

1946 : Le virus de la gastroentérite du porc (Transmissible Gastroenteritis Virus, TGEV) est identifié.

1949 et 1951: Des chercheurs du Rockefeller Institute (New York), puis une équipe londonienne ont découvert le virus de l'hépatite murine (Murine Hepatitis Virus, MHV) chez une souris paralysée.

1965 : La découverte du premier coronavirus humain (HCoV).

RECHERCHE BIBLIOGRAPHIQUE

1966 : Hamre et Procknow, chercheurs à l'université de Chicago, rapportent avoir cultivé un virus sur des cultures de cellules rénales embryonnaires humaines inoculées avec des prélèvements respiratoires d'étudiants en médecine enrhumés.

1967 : Une autre souche, baptisée HCoV OC43, est découverte aux Etats-Unis par Kenneth McIntosh et ses collègues des instituts nationaux de la santé.

1968 : Le terme « coronavirus » fait officiellement son apparition dans la revue Nature. Ce nouveau groupe de virus est défini à partir de critères essentiellement morphologiques.

1975 : Le comité international de taxonomie des virus (ICTV) approuve la création d'une nouvelle famille de virus, les Coronavirinae.

2012 et 2014 : Aux Etats-Unis, ce coronavirus a ainsi entraîné une diminution de plus de 3 % de l'élevage porcin.

2000 et 2004 : Identifications des séquences génomiques complètes des coronavirus humains HCoV-229E et HCoV-OC43 respectivement.

2003 : L'identification d'un coronavirus (SARS-CoV) comme agent infectieux responsable du Syndrome Respiratoire Aigu Sévère.

2012 : Un sixième coronavirus humain, le MERS-CoV, émerge au Moyen-Orient, soit dix ans après le SARS-CoV.

2020 : On comptait 2 519 cas confirmés de MERS, dont 866 morts, ce qui représente un taux de létalité d'environ 35 %, la première utilisation du terme « pandémie » sur le coronavirus.

I.1.3. Structure et génome du Coronavirus

Les Coronavirus sont des virus enveloppés qui contiennent le plus grand génome linéaire d'ARN viral simple brin de sens positif et pathogènes et le plus grand génome viral à ARN, allant de 26 à 32 kb de longueur (**Kasmi et al., 2020 ; Naqvi et al., 2020**).

Ces virus contiennent des projections de glycoprotéines en forme de pointes à leur surface, qui apparaissent comme une couronne au microscope électronique ; comme il se présente dans la **figure 1** ; par conséquent, ils sont appelés Coronavirus (**Mittal et al., 2020**).

RECHERCHE BIBLIOGRAPHIQUE

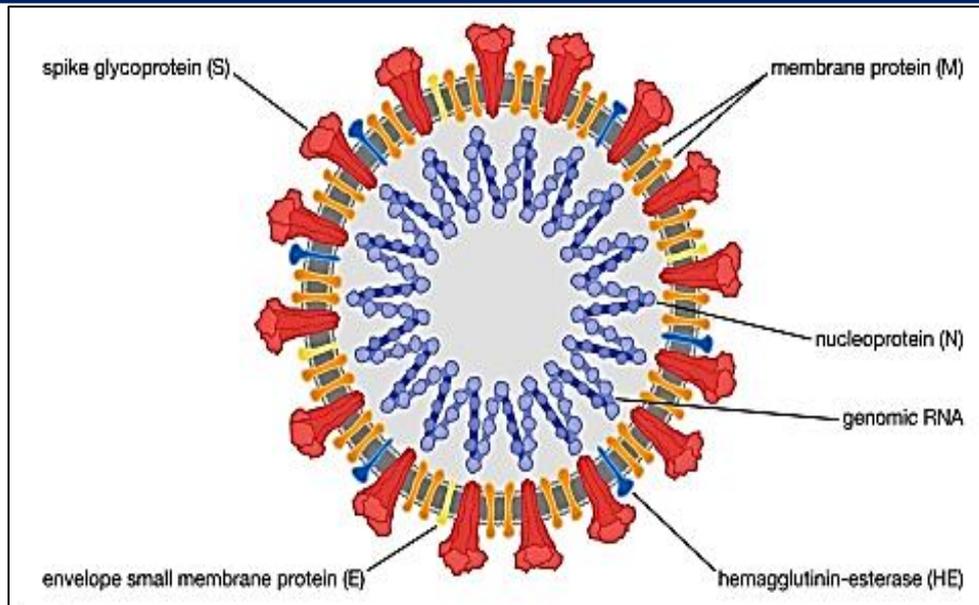


Figure 1 : Structure de Coronavirus (Shaikh *et al.*, 2020).

L'ARN génomique du Coronavirus est de composition constante entre leur virus, prenant le génome de SRAS-CoV-2 comme exemple de coronavirus, qui comprend deux principaux cadres de lecture ouverts (ORF) ; ORF1a et ORF1b, englobant les deux tiers du génome comme le montre la **figure 2** (Zhang *et al.*, 2021).

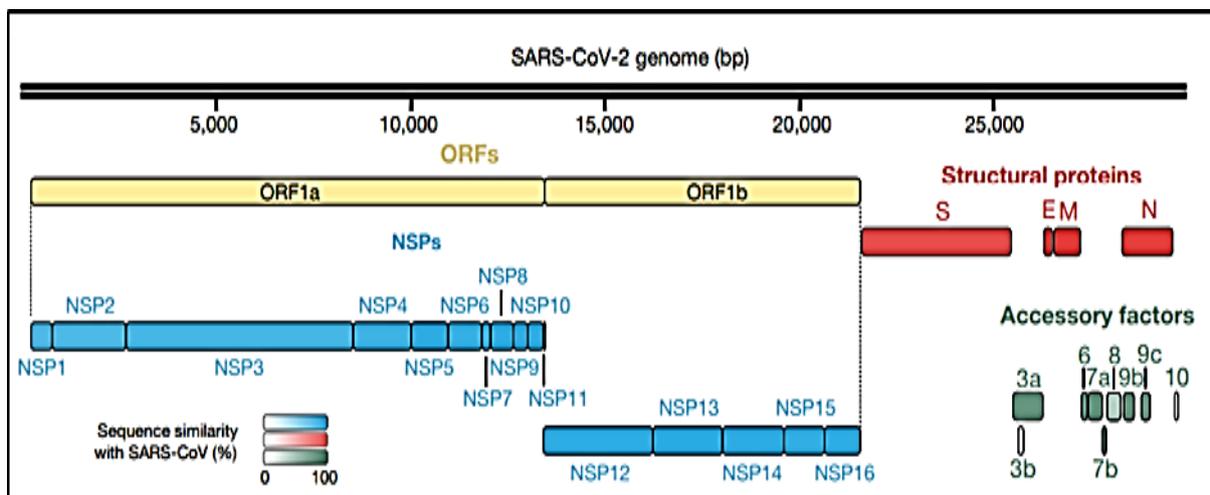


Figure 2 : Annotation du génome du Coronavirus (Gordon *et al.*, 2020).

Les deux premiers tiers du génome se composent de 2 grands cadres de lecture ouverts qui se chevauchent qui codent pour 16 protéines non structurales, y compris les protéases, l'ARN polymérase dépendante de l'ARN (RdRp), l'hélicase à ARN, la primase et d'autres, qui forment le complexe de réplique virale, une plate-forme pour propager les ARNm viraux. Où ces protéines non structurales sont toutes des cibles potentielles pour des thérapies, qui fonctionneraient en théorie contre tous les coronavirus (Bergmann *et al.*, 2020).

RECHERCHE BIBLIOGRAPHIQUE

Le tiers restant du génome qui code à des ORF qui se chevauchent, codant quatre protéines structurales majeures y compris S (glycoprotéine de pointe), N (protéine de nucléocapside), M (protéine de membrane) et E (protéine d'enveloppe), qui sont présentées dans le **tableau I** ainsi que leurs rôles ; et certaines protéines accessoires. La protéine S comprend le peptide signal (SP), le domaine de liaison au récepteur (RBD), le sous-domaine 1 (SD1) et le sous-domaine 2 (SD2) dans la sous-unité S1 et le peptide de fusion (FP), l'heptade répétée 1 (HR1), l'heptade répétée 2 (HR2) et transmembranaire (TM) dans la sous-unité de fusion membranaire (S2). La protéine E, avec M et N, est connue pour faciliter la formation de particules de type viral. 20 Le SRAS-CoV-2 code également pour des protéines accessoires, notamment ORF3, ORF6, ORF7a, ORF7b, ORF8 et ORF9b, qui sont toutes réparties parmi les gènes de structure, présentées dans la **figure 2 (Zhang et al., 2021)**.

Tableau I : Protéines structurales du coronavirus et leurs fonctions (**Cheikh, 2020**).

Protéines structurales	Fonctions des protéines
Protéine de pointe (S)	Fusion virus et cellule hôte par liaison
Protéine membranaire (M)	Transport des nutriments, détermine la forme et la formation de l'enveloppe
Protéine d'enveloppe (E)	Interfère avec la réponse immunitaire de l'hôte
Protéine de nucléocapside (N)	Se lie au génome d'ARN et constitue la nucléocapside
Hémagglutinine-estérase (HE)	Lie les acides sialiques sur la glycoprotéine de surface

I.1.4. Taxonomie et classification des coronavirus

L'ordre des Nidovirales regroupe les familles des *Coronaviridae* *Roniviridae*. Les variations dans la variété, le type et la taille des Nidovirales des protéines structurales provoquent des changements considérables dans la structure et la morphologie des nucléocapsides et des virions, comme il est représenté dans la **figure 3 (Mallikarjunaiah et al., 2021)**.

La recombinaison se produit fréquemment entre le génome des différentes lignées virales qui code pour la protéine de liaison au récepteur de l'hôte, contribuant à l'émergence de nouveaux virus, capables de transmission interspécifique et d'adaptation à de nouveaux hôtes animaux (**Mallikarjunaiah et al., 2021**).

RECHERCHE BIBLIOGRAPHIQUE

I.1.4.1 Structure de l'ordre des Nidovirales

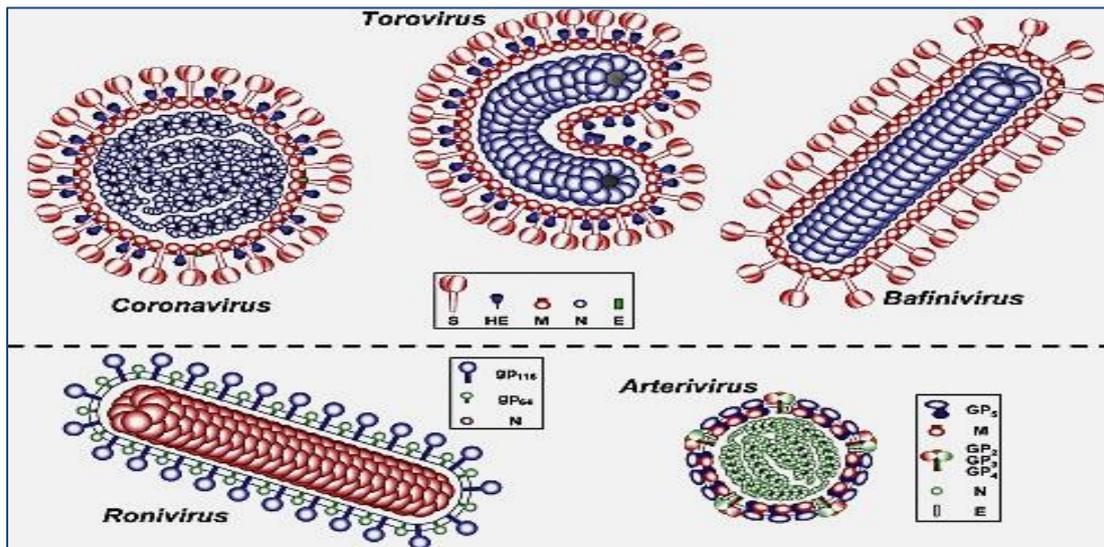


Figure 3: Schéma structurel de l'ordre des Nidovirales (Kasmi *et al.*, 2020).

La taxonomie de la famille de Nidoviral est présentée dans la **figure 4**. Les Coronavirus font partie de l'ordre *Nidovirales* *Arteriviridae* *Coronaviridae* *Roniviridae*. Les virus génériquement connus sous le nom de Coronavirus font partie de l'ordre Alphacoronavirus Betacoronavirus Gammacoronavirus Deltacoronavirus (Britton, 2019). La famille des Coronaviridae a été divisée en Coronavirinae et Torovirinae, Les Torovirus appartiennent à la sous-famille des Torovirinae qui provoque des maladies entériques chez les bovins et éventuellement chez l'homme (qui comprennent l'alpha coronavirus (α -CoV), le bêta coronavirus (β -CoV), le gamma coronavirus). La famille des Arteriviridae comprend les virus porcins et équins, et la famille des Roniviridae comprend les virus des invertébrés (γ -CoV) et delta coronavirus (δ -CoV) (Mallikarjunaiah *et al.*, 2021).

RECHERCHE BIBLIOGRAPHIQUE

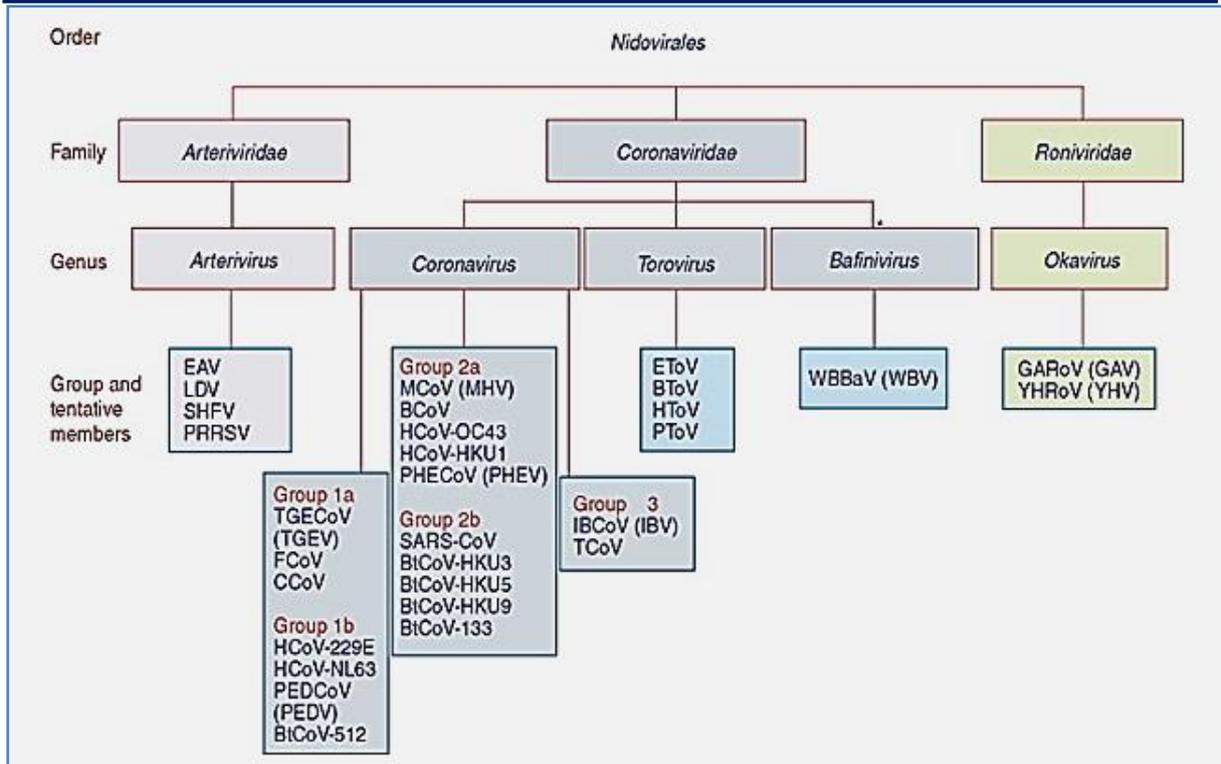


Figure 4 : Classification des Nidovirus et membres prototypes (Snijders, 2008).

L' α -CoV contient le virus de la diarrhée épidémique porcine, le virus de la gastro-entérite transmissible porcine, le coronavirus canin, le virus de la péritonite infectieuse féline, les coronavirus humains 229E (HCoV-229E) et NL63 (HCoV-NL63).

Le β -CoV est divisé en lignées A, B, C et D :

- La lignée A : Comprend les CoV OC43, HKU1, MHV, PHEV, équin, lapin, chameau, bovin et antilope qui code pour une protéine plus petite appelée hémagglutinine estérase, qui est fonctionnellement similaire à la protéine de pointe (S).
- La lignée B : Comprend le SRAS-CoV, les virus de type SRAS des chauves-souris et des civettes palmistes.
- La lignée C : Comprend le MERS-CoV et certains virus dérivés des chauves-souris.
- La lignée D : Ne contient que des Coronavirus dérivés de chauves-souris.

Le γ -CoV contient l'IBV aviaire et le coronavirus de la dinde.

Le δ -CoV comprend plusieurs Coronavirus nouvellement évolués issus des pools génétiques des oiseaux sauvages.

I.1.5. Coronavirus chez l'hôte

Des gouttelettes se déposent sur les mains ou des objets fixes ou mobiles et le virus va être ainsi transmis à une autre personne lorsque les mains entrent en contact avec le nez, les yeux ou la bouche. Le virus reste détectable de quelques heures à quelques jours sur des surfaces

RECHERCHE BIBLIOGRAPHIQUE

inertes selon le type de surface. Ce temps est plus court à des températures supérieures à 30 °C, mais la transmission indirecte par des objets inertes est faible (Mal, 2020).

I.1.5.1. Le Coronavirus dans la cellule

Tout le cheminement du Coronavirus dans la cellule de l'hôte est schématisé dans la **figure 5**. L'entrée du virus implique la liaison du récepteur de l'enzyme de conversion de l'angiotensine 2 (ACE2) et le clivage par la sérine protéase TMPRSS2 pour permettre la fusion avec la membrane de l'hôte. D'autres protéases cellulaires, par exemple la furine, facilitent l'entrée dépendante du pH par la voie endocytaire. Où les voies d'entrées prédominantes sont spécifiques au type de cellule et dépendent de la disponibilité de certaines protéases (Bergmann et Silverman, 2020).

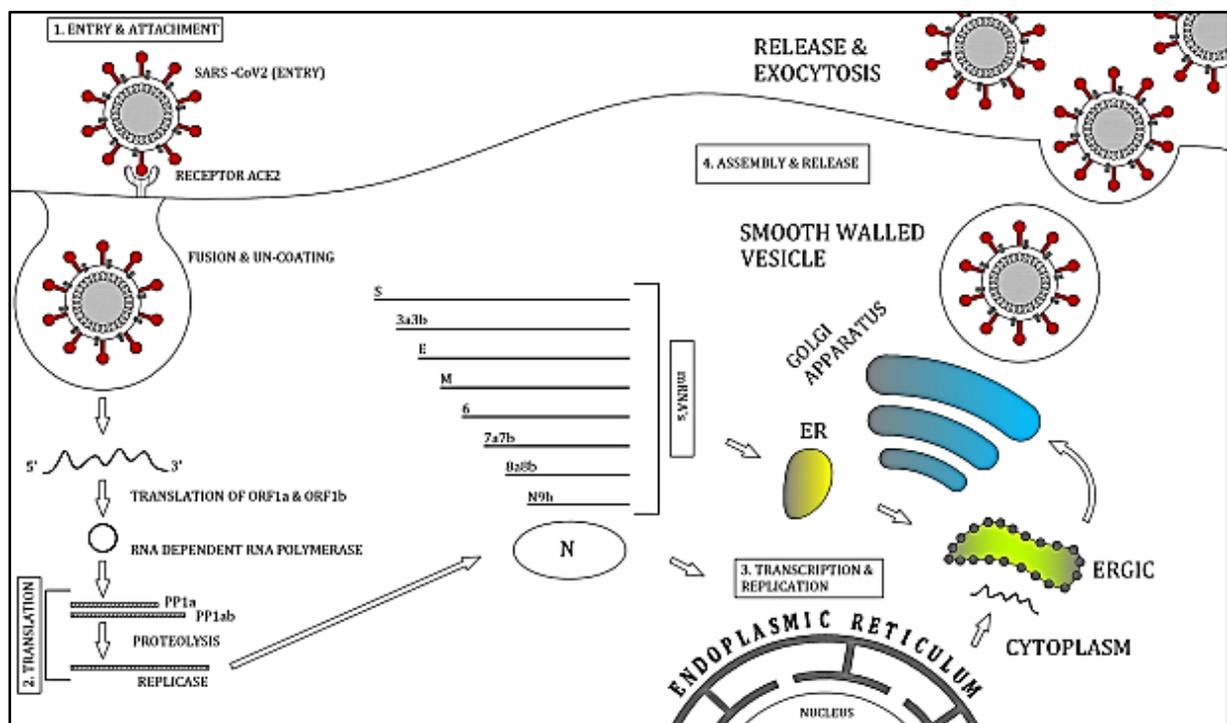


Figure 5: Cycle viral du Coronavirus (Shaikh et al., 2020).

I.1.5.2. Réplication du génome de Coronavirus

La réplication du génome du Coronavirus est médiée par la synthèse d'un ARN à brin négatif, qui est la matrice pour la synthèse des génomes de virus descendants. Des études de cartographie avec des ARN interférents défectueux (DI) du MHV ont indiqué que 470 nt de l'extrémité 5' et 436 nt de l'extrémité 3' sont nécessaires pour la réplication de l'ARN DI (60,73). Les deux extrémités du génome sont nécessaires à la synthèse du brin positif, alors que seuls les 55 derniers nt de l'extrémité 3' et la queue poly (A) sont nécessaires à la synthèse du brin négatif (Enjuanes et al., 2006).

RECHERCHE BIBLIOGRAPHIQUE

I.1.5.3. Transcription du Coronavirus

La transcription du Coronavirus, est en général la transcription dans l'ordre des Nidovirales, est une synthèse d'ARN dépendante de l'ARN qui comprend trois étapes a été proposé : (a) formation de complexes 5'-3' dans l'ARN génomique, (b) balayage d'appariement de bases de l'ARN brin négatif naissant par le leader TRS, et (c) commutation de matrice pendant la synthèse du brin négatif pour compléter le sgRNA moins-brin (Mal, 2020).

I.1.5.4. Incubation du Coronavirus

La période d'incubation est d'environ 5 jours, le plus souvent comprend entre 3 à 7 jours, mais peut aller jusqu'à 14 jours. C'est la raison pour laquelle la durée de 14 jours a été considérée comme la durée de la quarantaine. Faute de traitement antiviral efficace, l'évolution de la maladie est naturelle. Elle se fait vers la guérison en moins de dix jours chez plus de 80 % des malades. Mais chez les 20 % restant, on observe environ 7 à 10 jours après l'apparition des premiers symptômes, une phase d'orage immunitaire (orage cytokinique). Ce phénomène se traduit, d'une part, par une aggravation brutale des symptômes respiratoires qui peuvent conduire à une insuffisance respiratoire nécessitant une oxygénothérapie, voire une ventilation assistée en réanimation. À ce stade, le virus est encore détectable mais pas nécessairement viable (Mal, 2020).

I.1.6. Transmission du Coronavirus

I.1.6.1. Transmission entre les espèces

Les Coronavirus mammifères et aviaires ont des gammes d'hôtes diverses. La datation phylogénétique de la divergence de séquence de l'ARN polymérase dépendante de l'ARN (RdRp) a suggéré que l'ancêtre commun le plus récent (MRCA) des Coronavirus de mammifères est apparu il y a environ 7 000 à 8 000 ans, alors que le MRCA des Coronavirus aviaires remonte à 10 000 ans. Ces résultats sont probablement des sous-estimations car ils ne pouvaient pas tenir compte de la diversité supplémentaire des séquences de virus non découverts. Néanmoins, les estimations actuelles coïncident à peu près avec la dispersion de la population humaine dans le monde il y a environ 50 000 à 100 000 ans et ont fortement augmenté au cours des 10 000 dernières années lors de la première transition historique. Au cours de cette transition, les humains ont commencé diverses activités agricoles, telles que le défrichage des forêts pour l'agriculture et l'élevage, entraînant un changement significatif de l'écologie et de la dynamique des populations de virus en raison de l'intrusion d'habitats fauniques et du mélange intensif de différents hôtes animaux (Jasper *et al.*, 2013).

RECHERCHE BIBLIOGRAPHIQUE

Au cours des deux dernières décennies, trois Coronavirus zoonotiques (SARS-CoV, MERS-CoV et SARS-CoV-2) ont constitué une menace sérieuse pour la santé publique et la sécurité mondiale. De plus, HCoV-HKU1, HCoV-229E, HCoV-NL63, HCoV-OC43 et HECV-4408 sont également d'origine zoonotique (**Kong et al., 2022**).

Un animal sert d'hôte évolutif du Coronavirus s'il abrite un ancêtre étroitement apparenté partageant une homologie au niveau de la séquence nucléotidique dans ce cas le virus ancestral est généralement bien adapté et non pathogène pour cet hôte, ou si l'hôte réservoir héberge le Coronavirus d'une manière continue et à long terme. Dans les deux cas, les hôtes sont naturellement infectés et sont des hôtes naturels du Coronavirus ou de son virus parental, en revanche, si le Coronavirus est nouvellement introduit dans un hôte intermédiaire juste avant ou autour de son introduction à l'homme, il n'est pas bien adapté au nouvel hôte et est souvent pathogène. Cet hôte intermédiaire peut servir de source zoonotique de l'infection et jouer le rôle d'hôte amplificateur en permettant au virus de se répliquer de manière transitoire, puis le transmettre à l'homme pour amplifier l'échelle de l'infection humaine (**Ye et al., 2020**).

Le tropisme de l'hôte des Coronavirus est largement déterminé par la protéine de pointe en raison de ses fonctions de liaison aux récepteurs et de fusion membranaire. Une étape critique dans la transmission inter-espèces des Coronavirus est la capacité à se lier à de nouveaux récepteurs chez un nouvel hôte, qui se produit par la protéine de pointe. En raison de leurs grands génomes d'ARN, les Coronavirus sont génétiquement plus tolérants pour acquérir des mutations non létales, des insertions et des délétions, et une recombinaison dans le gène S, augmentant la capacité des Coronavirus à traverser les barrières des espèces hôtes, comme il est représenté dans la **figure 6 (Kong et al., 2022)**.

RECHERCHE BIBLIOGRAPHIQUE

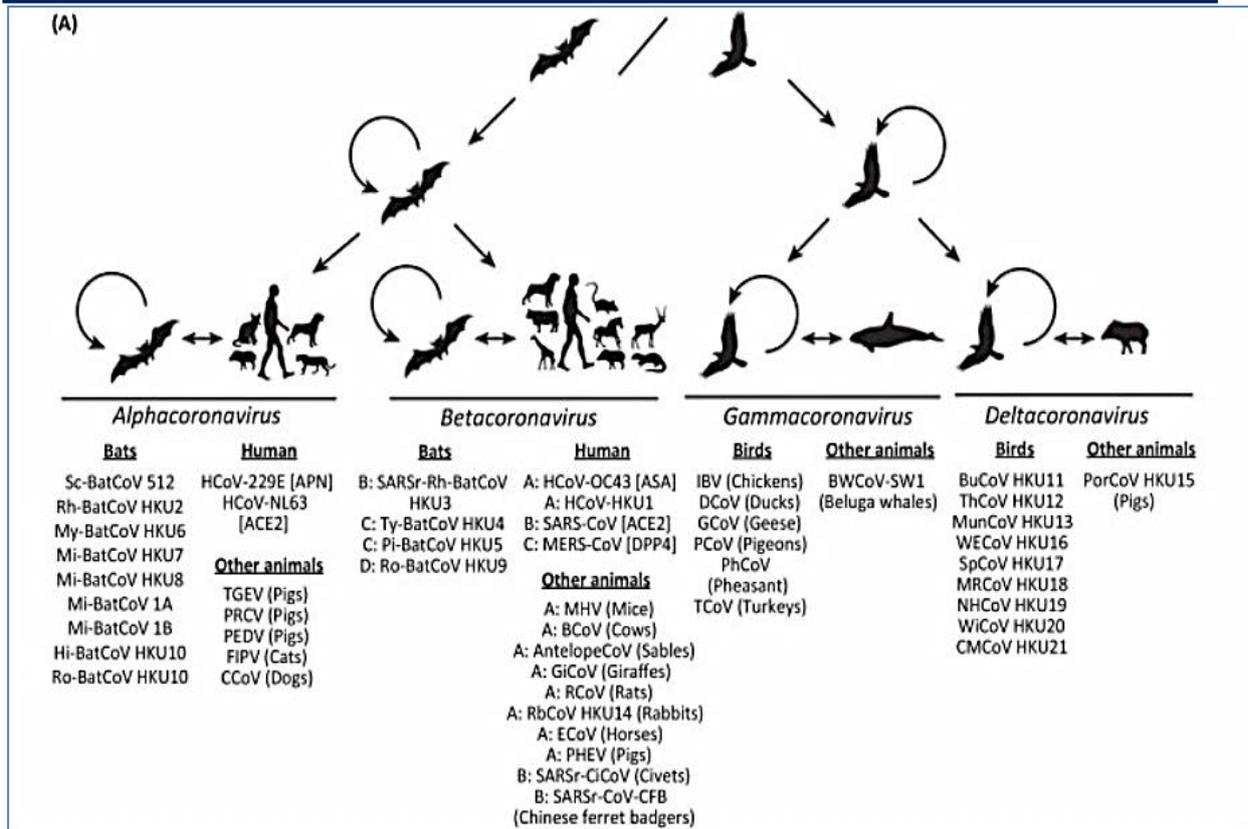


Figure 6 : Les chauves-souris et les oiseaux comme sources de gènes probables pour l'évolution des coronavirus, sur la base de preuves épidémiologiques, virologiques et phylogénétiques (Jasper *et al.*, 2013).

Dans la figure 6 Jasper *et al.*, 2013 :

I.2. Présentation de la phylogénie

La reconstruction phylogénétique est la méthode de choix pour déterminer les relations homologues entre les séquences. Les difficultés à produire des alignements de qualité, qui sont la base de bons arbres, et à automatiser l'analyse des arbres (Frickey et Lupas., 2006).

En raison du grand nombre de séquences impliquées, les analyses phylogénétiques des protomes excluent les étapes manuelles et nécessitent donc un degré élevé d'automatisation dans la sélection des séquences, l'alignement, l'inférence phylogénétique et l'analyse de l'ensemble d'arbres résultant (Frickey, 2004).

La phylogénie moléculaire consiste à reconstruire les relations de parenté entre les séquences de nucléotides ou d'acides aminés. Elle permet d'étudier l'évolution des séquences et la relation de parenté entre espèces. La nature quantitative des relations, ainsi établie, a permis le développement de méthodes et de règles rigoureuses pour la construction des arbres phylogénétiques (Cheikh, 2006).

RECHERCHE BIBLIOGRAPHIQUE

Pour reconstituer des phylogénies, on peut utiliser différents gènes, ou les protéines correspondantes, appelés marqueurs phylogénétiques (**Alauzet., 2009**).

Un marqueur génétique est basé sur la variation dans le génome, cette variation est aussi nommée polymorphisme. Selon le niveau taxonomique étudié, il est essentiel de choisir un type de marqueur qui lui est adapté. Par exemple, si le but visé est de distinguer des groupes d'individus au niveau de l'espèce, les marqueurs utilisés devront être plus conservés que dans le cas des études de populations (**Beaulieu, 2007**). Les polymorphismes ADN peuvent être détectés dans l'ADN nucléaire et l'ADN d'organites, trouvé dans les mitochondries et les chloroplastes. Les marqueurs moléculaires concernent la molécule d'ADN elle-même, comme tels, sont considérés comme des mesures objectives de la variation (**Vicente et al., 2003**).

Les polymorphismes d'ADN mitochondrial (ADNmt) ont été largement utilisés lors des analyses de la diversité phylogénétique et génétique. L'ADNmt haploïde transporté par les mitochondries du cytoplasme cellulaire, possède un mode maternel d'hérédité (les animaux héritent l'ADNmt de leurs mères et non de leurs pères) et un taux de mutation élevé ; il ne se recombine pas. Ces caractéristiques consentent aux biologistes de reconstruire les relations évolutives intra et interraciales par l'évaluation des modèles de mutation de l'ADNmt. Les marqueurs d'ADNmt peuvent également fournir un moyen rapide de détecter l'hybridation entre les espèces et les sous-espèces d'animaux d'élevage (**Nijman et al., 2003**).

I.2.1. Arbre phylogénétique

L'arbre phylogénétique est une construction-clé dans l'histoire de la biologie et de la géologie. Son succès opérationnel ne s'est jamais démenti. Imaginons que nous retracions l'histoire de trois espèces choisies arbitrairement (actuelles ou fossiles). En remontant le temps, nous pouvons espérer relier les deux espèces qui dérivent de la même espèce ancestrale ; en remontant plus loin encore, nous rencontrons l'espèce ancestrale des trois espèces. Le dessin en deux dimensions de ces relations de parenté est un arbre composé d'une succession de branchements. Au reste, on n'a jamais trouvé meilleure façon d'illustrer, en deux dimensions, les relations de parenté entre espèces ou groupes d'espèces, en fonction du temps (verticalement) et de la diversité taxinomique (horizontalement). Dans le cas de la diversité, de façon conventionnelle et le plus souvent symbolique, la dimension horizontale permet aussi une représentation du degré de divergence morphologique de deux branches à partir du point de branchement, ce que Darwin nomme la « somme des modifications » (**Daru et al., 2008**).

RECHERCHE BIBLIOGRAPHIQUE

L'arbre phylogénétique est une représentation graphique de la phylogenèse d'un groupe de taxons (séquences) qui peut se présenter selon deux formes (Benhamdi, 2020) :

- Arbre non enraciné : graphe connexe non cyclique, où il n'y a qu'un seul et unique chemin pour passer d'un sommet à l'autre (Figure 7).

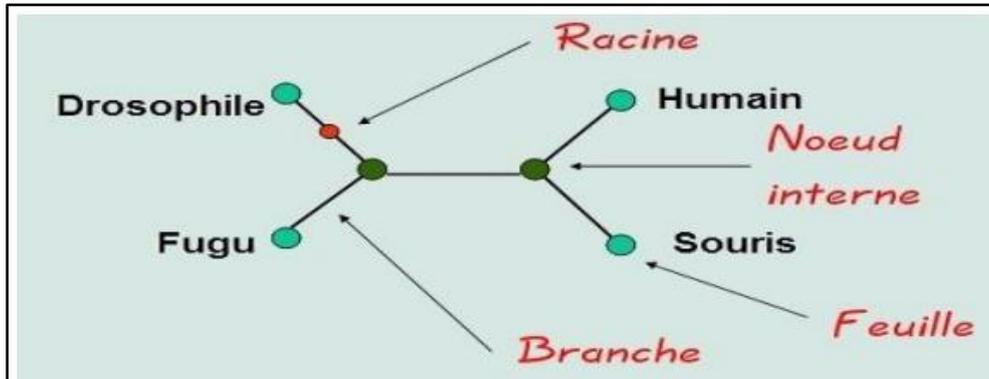


Figure 7 : Arbre phylogénétique non enraciné (Benhamdi, 2020).

- Arbre enraciné : possède une contrainte supplémentaire par rapport au précédent. Présence de liens orientés depuis une origine ou un ancêtre (Figure 8).

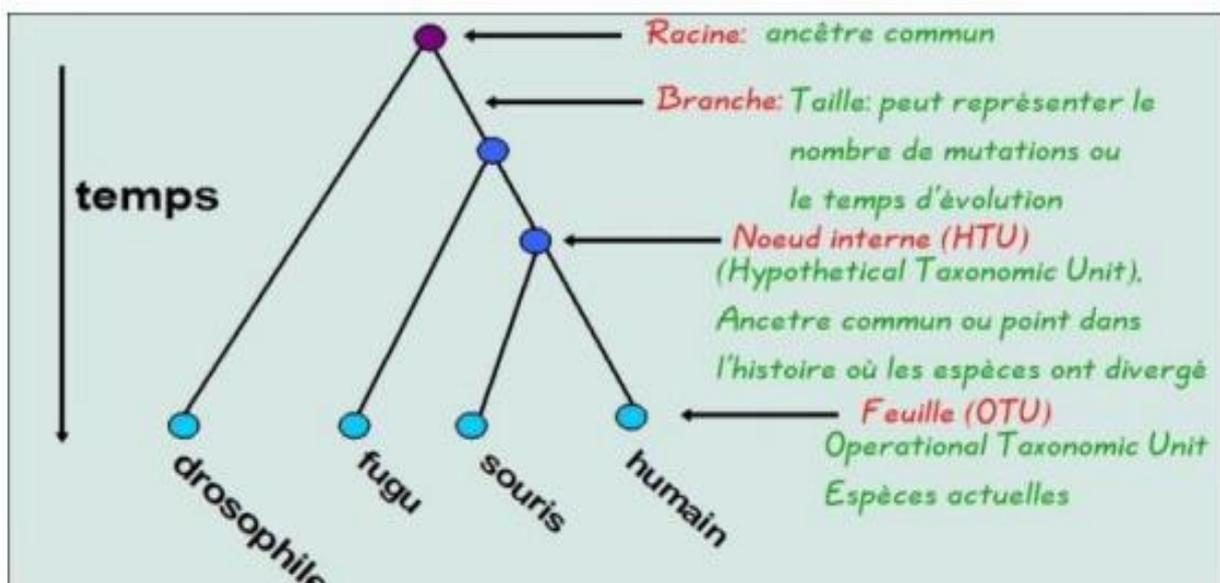


Figure 8 : Arbre phylogénétique enraciné (Benhamdi, 2020).

I.2.2. Les méthodes de reconstruction phylogénétique et les logiciels utilisés

Aujourd'hui, il existe une myriade de logiciels, implantés sur différentes machines depuis le microordinateur jusqu'à des ordinateurs les plus puissants, écrits dans tel ou tel langage, traitant tel ou tel problème.

RECHERCHE BIBLIOGRAPHIQUE

Dans le domaine de la reconstruction phylogénétique, il y a plusieurs programmes traitant les différents aspects de la phylogénie, allant des problèmes de comparaison des séquences, calcul des distances évolutives, reconstruction et visualisation d'arbres phylogénétiques et le Nettoyage de l'alignement. Pour cela nous allons décrire quelques grandes catégories de logiciels utilisés pour les analyses phylogénétiques selon leurs fonctions.

I.2.2.1. Les méthodes de reconstruction phylogénétique

I.2.2.1.1. Méthode de parcimonie

Les méthodes de parcimonie reposent sur l'idée que l'histoire évolutive la plus plausible est celle qui nécessite un minimum de mutations. Ces méthodes recherchent donc parmi tous les arbres possibles, et toutes les séquences possibles aux nœuds internes, la combinaison qui minimise le nombre total de mutations requis pour expliquer les données observées.

Les différentes mutations peuvent avoir des coûts différents, et il est possible d'attribuer des poids différents à chacun des sites.

Dans cette approche, les sites peuvent être traités de manière indépendante. Le nombre de mutations correspondant à un jeu de séquences particulier est égal à la somme des mutations requises par chacun des sites. Les coûts attribués aux mutations sont généralement symétriques (**Renwez et al., 2003**).

Ces parcimonies sont réalisées par des logiciels spécialisés tels que : PHYLIP (PhylogenyInference Package), PAUP, PAML, EDIBLE, PTP, PHYML, PARDOOT, ALIFRITZ, CRUX, PRAP.

I.2.2.1.2. Méthode Bayésien inférence

La méthode de bayes, plus exactement appelée méthode Bayesian Markov Chain Monte Carlo (MCMC), a récemment apporté un nouveau moyen de construire l'arbre phylogénétique le plus probable. Cette méthode semble donner des résultats prometteurs et son emploi s'est dernièrement généralisé.

Le principe de cette méthode est relativement simple : une série d'arbre est construite sur le principe que chaque arbre, très voisin du précédent, doit apporter une amélioration à la construction générale. Ainsi, un premier arbre phylogénétique est construit au hasard. Puis un élément de cet arbre est changé, également au hasard. Ce nouvel élément peut être la longueur d'une branche, ou le changement de position d'un taxon puis les arbres successifs sont conservés selon un algorithme particulier (dit de metropolis –Hastings) qui accepte les arbres au *prorata* de leur apport : si le nouvel arbre apporte une amélioration, il est toujours accepté ;

RECHERCHE BIBLIOGRAPHIQUE

s'il est moins bon de 10% de l'arbre précédent, il n'est accepté que 90% du temps (**Pagel, 2002**).

Cette méthode est réalisée par des logiciels spécialisés tels que : PAML, BAMBE, PAL, Vanilla, MrBayes, Mesquite, PHASE, BEAST, MrBayestree scanners, P4, SIMMAP, IMa2, BALI-Phy.

I.2.2.1.3. Méthode de Maximum de vraisemblance

Les méthodes de maximum de vraisemblance utilisent un modèle mathématique du processus d'évolution des séquences pour définir la probabilité qu'une phylogénie puisse produire les séquences observées. Elles cherchent ensuite la phylogénie pour laquelle cette probabilité est maximale. Pour utiliser une méthode de maximum de vraisemblance, il faut donc être capable de choisir un modèle d'évolution, d'estimer ses paramètres, et de calculer la vraisemblance d'un arbre pour ce modèle (**Ranwez, 2013**).

Cette méthode est réalisée par des logiciels spécialisés tels que : PHYLIP, DAMBE, EMBOSS, PRAP, PAUP*, ARB, Astrée, Creux, Vanilla, McRate, DART.

MATERIEL ET METHODES

La réalisation de ce travail a été faite au niveau de l'université Saad Dahleb à Blida I.

Les analyses phylogéniques par des outils de bio-informatique de la famille de *Coronaviridae* pour la construction d'arbres phylogéniques montre les liens de transmission zoonotique entre les différentes espèces touchées par ce virus.

II. Matériel et Méthodes

II.1. Matériel

II.1.1. Matériel Biologique

Dans cette partie on a utilisé des séquences d'ADN qu'on a obtenu grâce au numéro d'accèsion trouvé ci-dessous :

Tableau II : Les séquences d'ADN de Coronavirus utilisé pour la présentation phylogénétique.

Animal	Espèce	Région	Gène	N° D'accèsion
Alpaga	<i>Vicugna pacos</i>	Etats-Unis	S, ORF1ab , ORF4	JQ410000
Aviaire	Aviaire	Chine	S	AY319651
	<i>Gallus gallus</i>	Corée du sud	S	JQ977698
	<i>Gallus gallus</i>	Chine	S	KM213963
Bovin	Bovin	Japon	N	AB354579
	Bovin	Canada	N	AF220295
	Bovin	Etats-Unis	S	BCU00735
	Bovin	Etats-Unis	S	DQ811784
	Bovin	Corée du sud	N	EU401981
	Vache	Brésil	N	GU808341
	Bétail	Suède	S	KF169908
	Veau	Suède	S	KF169940
	Bovin	France	N	KT318090
	porc	France	RDRP	KT318109
	Bétail	Slovénie	RDRP	KX059608
	Bétail	Slovénie	RDRP	KX059610
	Bétail	Slovénie	RDRP	KX059630

MATERIEL ET METHODES

	<i>Bos taurus</i>	Chine	N	KX266233
	<i>Bos taurus</i>	Brésil	N	MF510415
	<i>Bos taurus</i>	Brésil	N	MF510416
	<i>Bovin</i>	Italie	N	MW074864
Canard	<i>Anas sp</i>	Pologne	RDRP	KJ690951
	<i>Anas sp</i>	Pologne	RDRP	KJ690952
	<i>Anas platyrhynchos</i>	Pologne	RDRP	KJ690953
	<i>Anser fabalis</i>	Pologne	RDRP	KJ690954
	<i>Anser fabalis</i>	Pologne	RDRP	KJ690955
	<i>Mareca penelope</i>	Chine	N	NC_016995
Chien	<i>Canis lupus familiaris</i>	Allemagne	N	KC175339
	Chien	Chine	N	KT852998
Chat	<i>Prionailurus bengalensis</i>	asia	ORF1ab	EF584908
Chauves-souris	<i>Chaerephon pusillus</i>	Mayotte	RDRP	MN183173
	<i>Triaenops after</i>	Kenya	S	KY073746
	<i>Rhinolophus sinicus</i>	Chine	S	KJ473814
	<i>Hipposideros abae</i>	Ghana	S, ORF1ab	KT253259
	<i>Triaenops</i>	Kenya	ORF1ab -S- ORF3- E-M-N- ORFx	KY073746
	<i>Hipposideros</i>	Kenya	ORF1ab -S ORF4- E-M-N- ORF8	KY073747
	<i>Hipposideros vittatus</i>	Kenya	ORF1ab -S	KY073748

MATERIEL ET METHODES

			ORF4- E-M-N- ORF8	
	<i>Rhinolophus pusillus</i>	Chine	S, N	MG772934
	<i>Rhinolophus sp</i>	Mozambique	RDRP	MN183152
	<i>Rhinolophus rhodesiae</i>	Mozambique	RDRP	MN183153
	<i>Rhinolophus lobatus</i>	Mozambique	RDRP	MN183154
	<i>Rhinolophus rhodesiae</i>	Mozambique	RDRP	MN183159
	<i>Miniopterus mossambicus</i>	Mozambique	RDRP	MN183161
	<i>Triaenops Aafter</i>	Mozambique	RDRP	MN183168
	<i>Hipposideros caffer</i>	Mozambique	RDRP	MN183172
	<i>rhinolophus affinis</i>	Chine	N	MN996532
	<i>Triaenops after</i>	Kenya	ORF1ab -S- ORF3- E-M-N- ORFx	NC_032107
	<i>Triaenops after</i>	Kenya	ORF1ab -S- ORF3- E-M-N- ORFx	NC_048216
Chien	Chien	Taiwan	S	GQ477367
	Chien	Corée du sud	N	JX860640
	<i>Canis lupus familiaris</i>	Vietnam	S	LC190906
	<i>Nyctereutes procyonoides</i>	chine	S	EF192155
Equin	<i>Equin</i>	Etats-Unis	S	EF446615
	<i>Equus caballus</i>	Japon	S	LC061272

MATÉRIEL ET MÉTHODES

	<i>Equus caballus</i>	Japon	S	LC061273
	<i>Equus caballus</i>	Japon	S	LC061274
Félin	Félin	Etats-Unis	N	AY994055
	Félin	Thaïlande	N	MW558579
	Félin	Thaïlande	N	MW558584
	Félin	Thaïlande	N	MW558586
	<i>Felis catus</i>	Japon	S	AB535528
	Chat	Pays Bas	S	FJ938058
	Chat	Etats-Unis	S	FJ938061
	Félin	Pays Bas	S	HQ012369
Furet	<i>Melogale moschata</i>	chine	S	EF192156
	furet domestique de Lima	Pérou	RDRP	KR758767
	furet domestique de Lima	Pérou	RDRP	KR758768
Girafe	Girafe	Etats-Unis	S	EF424622
	Girafe	Etats-Unis	S	EF424623
	Girafe	Etats-Unis	S	EF424624
Humain	<i>Homo sapiens</i>	Belgique	N	AY391777
	<i>Homo sapiens</i>	Amsterdam	S	AY567487
	<i>Homo sapiens</i>	Belgique	N	AY903460
	<i>Homo sapiens</i>	Chine	N	JN129834
	<i>Homo sapiens</i>	Chine	N	JN129835
	<i>Homo sapiens</i>	Slovénie	RDRP	KX059640
	<i>Homo sapiens</i>	Slovénie	RDRP	KX059650
	<i>Homo sapiens</i>	Slovénie	RDRP	KX059680
	<i>Homo sapiens</i>	Slovénie	RDRP	KX059690
	<i>Homo sapiens</i>	Slovénie	RDRP	KX059697
	<i>Homo sapiens</i>	Chine	S, N	MT049951
	<i>Homo sapiens</i>	Chine	N	MZ420153
Oiseau sauvage	<i>Spatula discors</i>	Etats-Unis	N	MN379902
	<i>Spatula discors</i>	Etats-Unis	N	MN379904

MATÉRIEL ET MÉTHODES

Pangolin	Pangolin	Chine	S	MT072864
Porc	Cochon	Chine	S	MK040451
	<i>Cochon</i>	CHINE	S	MK040453
	Cochon	Chine	S	MK040454
	Cochon	Chine	S	MK040455
	Porc	Italie	RDRP	MW089337
	Porc	Italie	RDRP	MW089338
	Porc	Italie	RDRP	MW089339
	Porc	Italie	RDRP	MW089340
	Porc	Italie	RDRP	MW089341
	Porc	Italie	RDRP	MW089343
	<i>Cochon</i>	Etats-Unis	S	DQ811785
	<i>Pocin</i>	Etats-Unis	S	DQ811786
	Cochon	Etats-Unis	S	DQ811787
	Cochon	Etats-Unis	S	DQ811788
	Pocin	Etats-Unis	S	DQ811789
	"porcelet"	Chine	N	KC196276
	Porcin	Chine	N	KC210145
	Porcin	Etats-Unis	N	KF468753
	Porcin	Etats-Unis	N	KJ481931
	Porcin	Etats-Unis	N	KJ567050
	Porcin	Corée du sud	N	KM820765
	Porcin	Chine	N	KP757890
	Cochon	Chine	S	KT266822
	<i>Sus scrofa</i>	Chine	S	KT336560
	Cochon	Thaïlande	N	KU051649
	Porc	Chine	S	KU981059
	<i>Cochon</i>	Thaïlande	S	KU984334
	Cochon	Laos	N	KX118627
	Porcin	Japon	N	KY619801
	Porcin	Japon	N	KY619802
Porc	Chine	S	MF041982	

MATERIEL ET METHODES

	Porc	Chine	N	MF431743
Poulet	Poulet	suède	N	JQ088078
	<i>Gallus gallus</i>	Pologne	Génome complet	KT886454
	<i>Gallus gallus</i>	Pologne	S	MK581204
	<i>Gallus gallus</i>	Pologne	Génome complet	MK581205
	<i>Gallus gallus</i>	Pologne	Génome complet	MK581206
	<i>Gallus gallus</i>	Pologne	Génome complet	MK581207
	<i>Gallus gallus</i>	Pologne	Génome complet	MK581208
Rongeur	<i>Rattus tanezumi</i>	Chine	RDRP	KY370055
	<i>Rattus flavipectus</i>	Chine	RDRP	KY370056
	<i>Mus caroli</i>	Chine	RDRP	KY370057
	<i>Mus caroli</i>	Chine	RDRP	KY370058
	<i>Niviventer niviventer</i>	Chine	RDRP	KY370060
	<i>Apodemus agrarius</i>	Chine	RDRP	KY370061
Sanglier	<i>Sus scrofa</i>	Philippine	N	LC384327
	<i>Sus scrofa</i>	Philippine	N	LC384328

II.1.2. Matériel non biologique

Dans la partie matérielle non biologique on a utilisé un nombre de logiciel trouvé dans le tableau III ci-dessous :

Tableau III : les logiciels utilisés dans l'étude phylogénétique.

logiciel	Description	Lien de web
GenBank	une base de données qui contient une collection de toutes les séquences des nucléotides et de protéines et annotations	https://www.ncbi.nlm.nih.gov/genbank/

MATERIEL ET METHODES

	bibliographiques et biologiques disponibles au public.	
BLAST	un algorithme utilisé pour comparer les séquences d'acides aminés de différentes protéines ou les séquences de nucléotides d'acide nucléique.	https://blast.ncbi.nlm.nih.gov/Blast.cgi
MEGA 11	un programme qui permet d'explorer et d'analyser des séquences nucléidiques et protéiques.	https://www.megasoftware.net/
GenieGen 2	un logiciel qui permet de travailler sur des séquences d'ADN, d'ARN ou de protéines.	https://www.pedagogie.ac-nice.fr/svt/productions/geniegen2/

II.2. Méthode

Comme une première étape, la recherche des articles qui cible les séquences d'espèces et souches différentes, était réalisé par PubMed et Google Scholar qui sont l'un des moteurs de recherche utilisés qui offrent l'accès aux articles, thèses, livres à tous les cliniciens, chercheurs et stagiaires intéressés, ainsi qu'au grand public. Ils se concentrent principalement sur la médecine et les sciences biomédicales où ils autorisent le plus grand nombre de mots-clés par recherche (Falagas *et al.*, 2007).

La deuxième étape consiste à extraire plus d'informations à partir des séquences traitées dans ces articles marqués par un numéro d'accession, grâce à la GenBank.

II.2.1. Outils Bioinformatique

II.2.1.1. Outil GenBank

Principe : La GenBank est construite principalement à partir de la soumission de données de séquence par les auteurs et de la soumission en masse de l'étiquette de séquence exprimée

MATERIEL ET METHODES

(EST), de la séquence d'enquête génomique (GSS) et d'autres données à haut débit provenant des centres de séquençage (Benson et al., 2005). Elle est une collection de toutes les séquences des nucléotides et de protéines et annotations bibliographiques et biologiques disponibles au public (Sayers et al., 2021).

Toutes les séquences étudiées de la famille du *Coronaviridae* dans notre travail ont été téléchargées à partir de la GenBank. Nous passons par plusieurs étapes distinctes :

- Construire l'URL. La page d'accueil de la GenBank s'affiche comme elle se présente dans la **figure 9**.

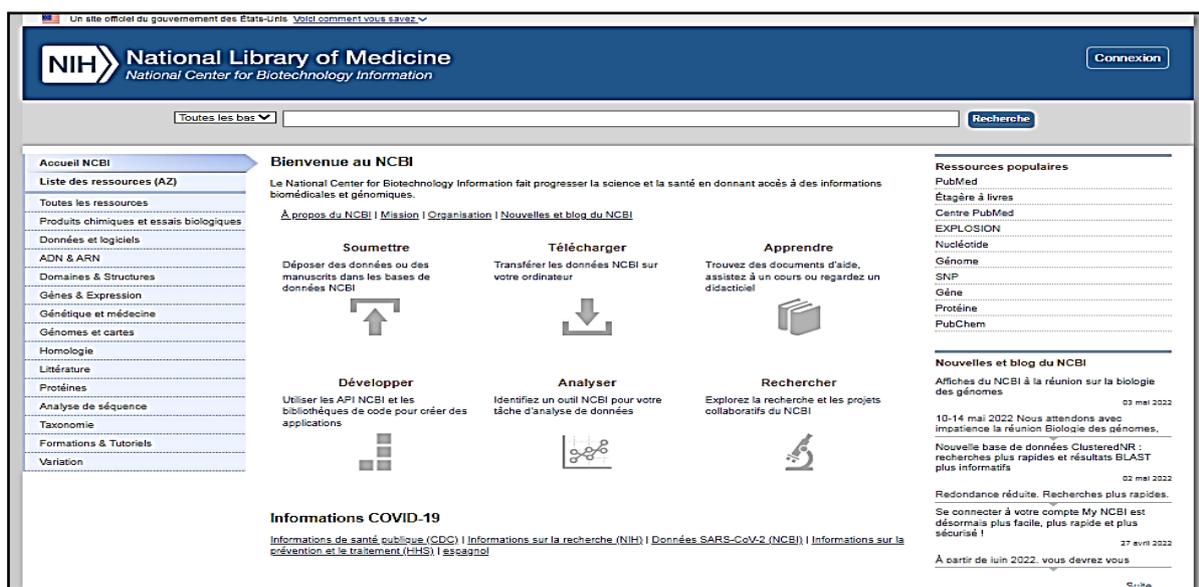


Figure 9 : Page d'accueil de la GenBank.

- Coller le numéro d'accèsion cité dans l'article (nous avons choisis à titre d'exemple le gène de la protéine de pointe d'une séquence sous le numéro d'accèsion MK040451) dans le cadre de recherche. La page affiche d'abord les informations de base de la séquence (**Figure 10**).



Figure 10 : Résultat de la recherche de la séquence dans la GenBank.

MATERIEL ET METHODES

- Dans la **figure 11** s'affiche la page de GenBank qui montre plus d'information sur la séquence recherché, pour obtenir la séquence FASTA on clique sur FASTA encadré en rouge.

GenBank

Send to: Change region shown

Customize view

Porcine deltacoronavirus strain CH-WH-2017 spike protein gene, complete cds

GenBank: MK040451.1

FASTA Graphics PopSet

Go to: ☺

LOCUS MK040451 3480 bp RNA linear VRL 08-MAY-2019

DEFINITION Porcine deltacoronavirus strain CH-WH-2017 spike protein gene, complete cds.

ACCESSION MK040451

VERSION MK040451.1

KEYWORDS .

SOURCE Porcine deltacoronavirus

ORGANISM [Porcine deltacoronavirus](#)
Viruses; Riboviria; Orthornavirae; Pisuviricota; Pisoniviricetes; Nidovirales; Coronaviridae; Orthocoronavirinae; Deltacoronavirus; Buldecovirus.

REFERENCE 1 (bases 1 to 3480)

AUTHORS Gu, J., Mao, Y. and Zhou, J.

TITLE The complete S gene of porcine deltacoronavirus

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 3480)

AUTHORS Gu, J., Mao, Y. and Zhou, J.

TITLE Direct Submission

ANALYZE THIS SEQUENCE

Run BLAST

Pick Primers

Highlight Sequence Features

Find in this Sequence

RELATED INFORMATION

Protein

Taxonomy

PopSet

LINK OUT TO EXTERNAL RESOURCES

Virus Pathogen Resource

Figure 11 : les informations du gène dans la GenBank.

- La forme FASTA de cette séquence est présentée dans la **figure 12**.

Nucléotide Nucléotide Recherche

Avancé Aider

FASTA Envoyer à:

Souche de deltacoronavirus porcin CH-WH-2017 gène de la protéine de pointe, CD complet

GenBank: MK040451.1

Popset [graphics](#) [GenBank](#)

>MK040451.1 Souche de deltacoronavirus porcin CH-WH-2017 gène de la protéine de pointe, CD complet

```
ATGCAGAGAGCTATTGATTATGACCTTACTTTGTCTCGTTGAGCAAAGTTTGCTGATGACTACTCG
ATTGTGCTACCTTCCGGGTGCACATCGCTTCTACATAAACTCAGAGTAATCCAGCAGTTCTCCACTC
GCGGTAATAACTTTGATGTTGGGGTTCTCTGCTGACGCCACTAAGAAGCTTAACCTCTCTCACCA
CTTACTAACTACTTTGCCATTAAATGGCTTCTCAGGATTAACAACACTCATGCTGAATTTTATA
CTAAATAACTAACCACTCTCAGATGTAATCTCTACTACTAGTGGGTACAACCTATAGCTCCCGCGG
TGCCATGGTAAATACCAACATGATGCAATGATGATGATGATGATGATGATGATGATGATGATGATG
CTGTTGAAGTTTGTGGCAACGTGGTGAATAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG
CTGCATACCAACTCTACTTCTCACTTAATCTGAGCTTATGCTTTACTAATAACACTACTT
AGGTATCTTCCCACTGATTAACTGACTTACGGTCTATGCTACTGGGCAAGTTTATGCTAATGGTTAC
CTTTTAGGTACTTTACCTATTACGGTAACTATGTAAGGTTGATCGGGCTCATTTGTCGGCAATAGTG
CCCACCTTGGCCCTTGAACCTAACCGATACACTCATAAACAATCACTAACAATCACTATATGCAAAATCA
TTATTGTTGATAAGTCAGTAGTGTGATCAATAGCATGCCAGCGCTTCTCAGAAAGTGGAGGATGGGTTT
TACTCTGACCTAAATCTCGCTTAGAGCTAGGCAAGCTATTGTTACACTAACCTAACCTCCCTGAGC
TTGAGTAGTGCAGTTAAATTTCTGCAACATGGATTTTGGCAAGCTAGCTTTGAGAGCTTACCAT
CAATGGTAACACCTCTATTGTTGCTACTAAGCTTACTTCAGSCTTGAACCTAACTTATGTTGACAGGT
TGCACTATGATCTGGCACTGATCACTGATGTTTGGACCTCGCAGCAACAATGGCATCTGATCTTCT
CTCAATTTTGTCTAAGCACTGAATCTGGTCTTGTGAGATGAAAATATTGTTTACCTAGCTATTGGAAATTA
CTTGCTAGGCAAGCTTGTATGTTACTGCTAGAGGGCCAGCTCAGCTGGAACCTCTCAGTACAT
GCAACAGCACTTCTAGTGAATCACTGATGTTCTGCACTGACTACACTATCTAGGAGTCTCTGGTACTG
GCATTTCAAGCCTCAGATCTCTATTGCAACAATGGCATAGCATCCTCTCCACAGGTGAGCTCTTA
TGCAATTAATAATAACCACTGGCAAAACCTTCAAGTCTTACGGTGTGAAACCCCTTCTCAACTGAT
```

Modifier la région affichée

Séquence entière

Région sélectionnée parmi :

begin pour end

Mettre à jour la vue

Personnaliser la vue

Analysez cette séquence

Exécutez BLAST

Choisissez des amorces

Mettre en surbrillance les fonctionnalités de la séquence

Informations connexes

Protéine

Taxonomie

PopSet

LinkOut vers des ressources externes

Figure 12: Forme FASTA de la séquence affichée par la GenBank.

La séquence de forme FASTA (alignement rapide) qui se définit par un programme de choix dans les recherches initiales dans les bases de données. Cependant, en raison de sa sensibilité relativement plus élevée dans certains cas, il doit être inclus dans une recherche (**Donkor et al., 2014**). Cette dernière sera enregistrée comme il est expliqué par la **figure 13**.

MATERIEL ET METHODES

The screenshot shows the NIH PMC search results for the query 'MKD40451'. The page displays two search results:

- Porcine Deltacoronaviruses: Origin, Evolution, Cross-Species Transmission and Zoonotic Potential**
Fanzhi Kong, Qihong Wang, Scott P. Kenney, Kwoni Jung, Anastasia N. Vlasova, Linda J. Saif
Pathogens. 2022 Jan; 11(1): 79. Published online 2022 Jan 9. doi: 10.3390/pathogens11010079
PMCID: PMC8778258
[Article](#) [PubReader](#) [PDF-1.3M](#) [Cite](#)
- Detection and spike gene characterization in porcine deltacoronavirus in China during 2016-2018**
Yu Zhang, Yao Cheng, Gang Xing, Jing Yu, Ao Liao, Liuyang Du, Jing Lei, Xue Lian, Jiyong Zhou, Jinyan Gu
Infect Genet Evol. 2019 Sep; 73: 151-158. Published online 2019 Apr 23. doi: 10.1016/j.meegid.2019.04.023
PMCID: PMC7106057
[Article](#) [PubReader](#) [PDF-1.7M](#) [Cite](#)

Figure 15 : Résultats de la recherche en texte intégral PMC.

2. Par la suite dans le cadre de "Gène", la recherche de PopSet nous offre plus de détails sur l'article qui a traité cette séquence et les séquences dans le même ensemble de données (la même souche, la même espèce pour le même gène) (Figure 16).

The screenshot shows the NIH PopSet search results for the query 'MKD40451'. The page displays search results for the gene 'Gène de la protéine de pointe du deltacoronavirus porcin, CD complet'. The results list several sequences from the GenBank database:

Accession Number	Description
MKD40455.1	Souche de deltacoronavirus porcin CH-HA3-2017 gène de la protéine de pointe, CD complet
MKD40454.1	Souche de deltacoronavirus porcin CH-HA2-2017 gène de la protéine de pointe, CD complet
MKD40453.1	Souche de deltacoronavirus porcin CH-HA1-2017 gène de la protéine de pointe, CD complet
MKD40452.1	Souche de deltacoronavirus porcin CH-XS-2018 gène de la protéine de pointe, CD complet
MKD40451.1	Souche de deltacoronavirus porcin CH-WH-2017 gène de la protéine de pointe, CD complet
MKD40450.1	Souche de deltacoronavirus porcin CH-DH2-2017 gène de la protéine de pointe, CD complet
MKD40449.1	Souche de deltacoronavirus porcin CH-DH1-2017 gène de la protéine de pointe, CD complet
MKD40448.1	Souche de deltacoronavirus porcin CH-HX-2019 gène de la protéine de pointe, CD complet
MKD40447.1	Souche de deltacoronavirus porcin CH-FY-2017 gène de la protéine de pointe, CD complet
MKD40446.1	Souche de deltacoronavirus porcin CH-CZ2-2017 gène de la protéine de pointe, CD complet
MKD40445.1	Souche de deltacoronavirus porcin CH-CZ1-2017 gène de la protéine de pointe, CD complet

Figure 16 : Résultat de recherche PopSet de la séquence dans la GenBank.

3. Dans le cadre de "Protéine", la recherche du groupe de protéines identiques à la séquence ciblée s'affiche comme il est présenté dans la figure 17.

MATERIEL ET METHODES

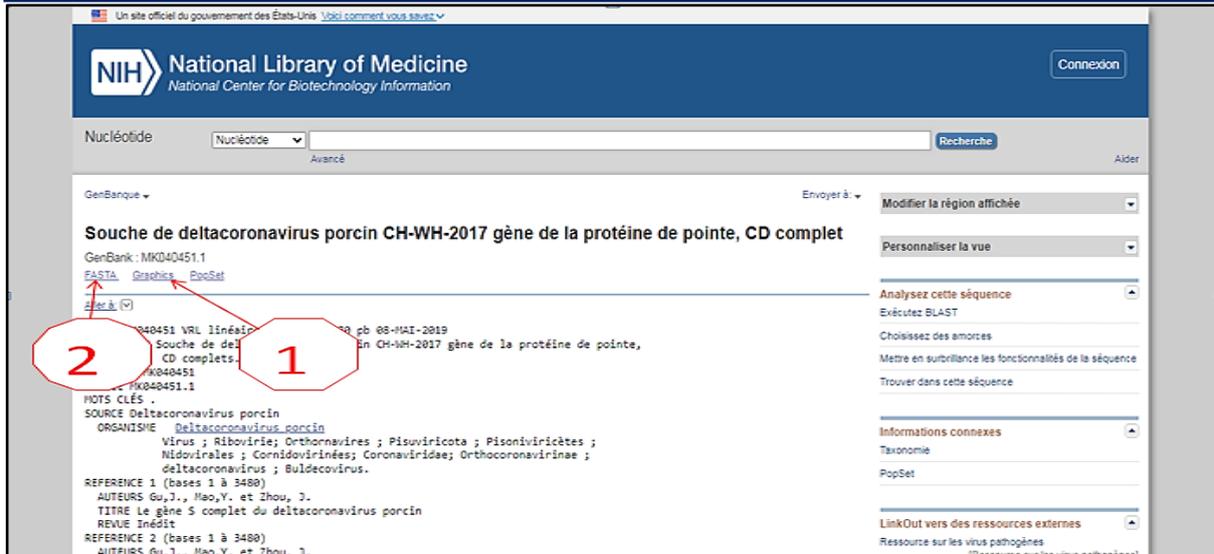


Figure 19 : Les deux formes d'affichage de la séquence.

- La forme graphique du gène S est présentée dans la **figure 20**. Sa position, sa taille en nucléotide (ARN) et en acide aminée (protéine) (région S1 et S2 qui les composent).

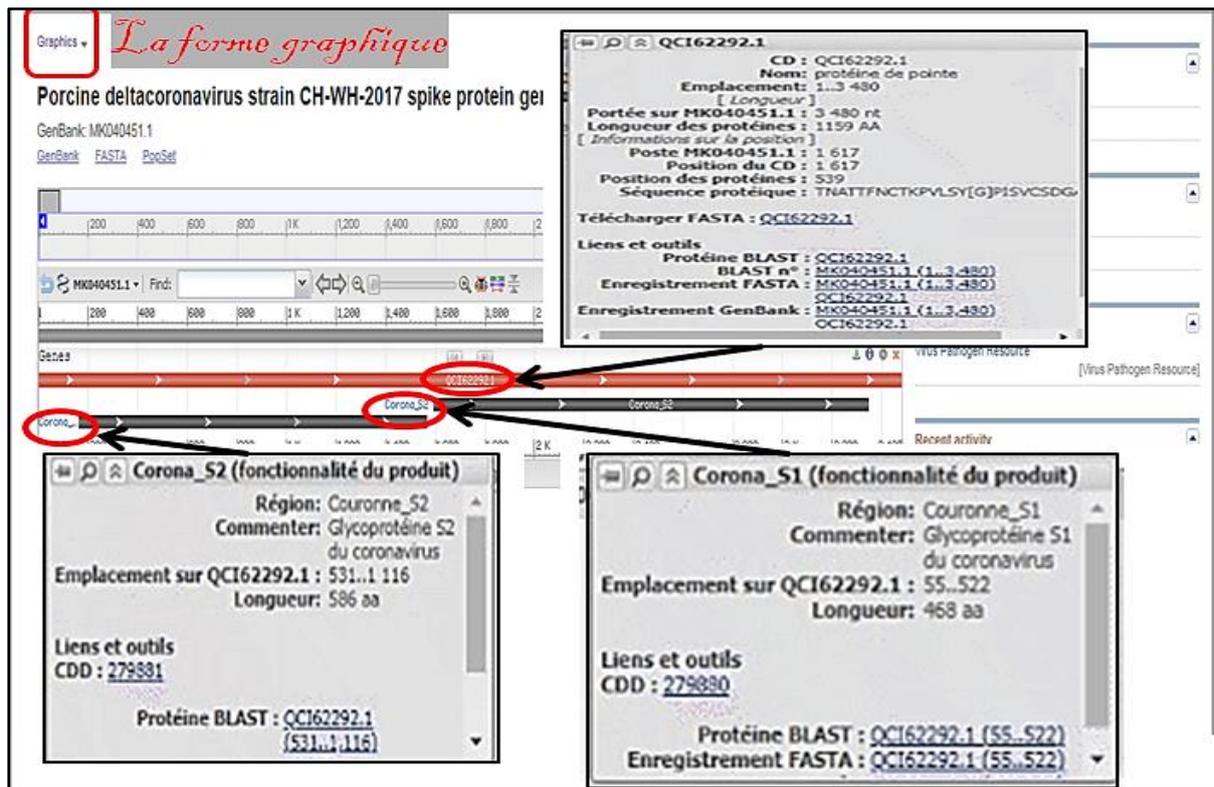


Figure 20: Exemple du graphique d'un gène S de la séquence.

*Dans une étape supplémentaire d'identification de séquence à forte similarité par rapport à celles précédemment étudiées dans le cadre de BLAST.

II.2.1.2. Outil de recherche d'alignement local de base (BLAST)

MATERIEL ET METHODES

Principe : BLAST est un algorithme utilisé pour comparer les séquences d'acides aminés de différentes protéines ou les séquences de nucléotides d'acide nucléique. BLAST a été inventé en 1990 et est depuis devenu le standard de facto dans les outils de recherche et d'alignement (Altschul et al., 1990).

Grâce à une recherche BLAST, on peut comparer une séquence de requête avec une base de données de séquences, et ainsi identifier les séquences de bibliothèque qui partagent une ressemblance avec la séquence de requête au-dessus d'un certain seuil. Sur la base d'une telle comparaison, BLAST peut être utilisé pour atteindre plusieurs objectifs, notamment l'identification des espèces, la localisation des domaines, la cartographie de l'ADN et l'annotation (Altschul et al., 1990).

La page d'accueil de BLAST est représentée dans la **figure 21**.

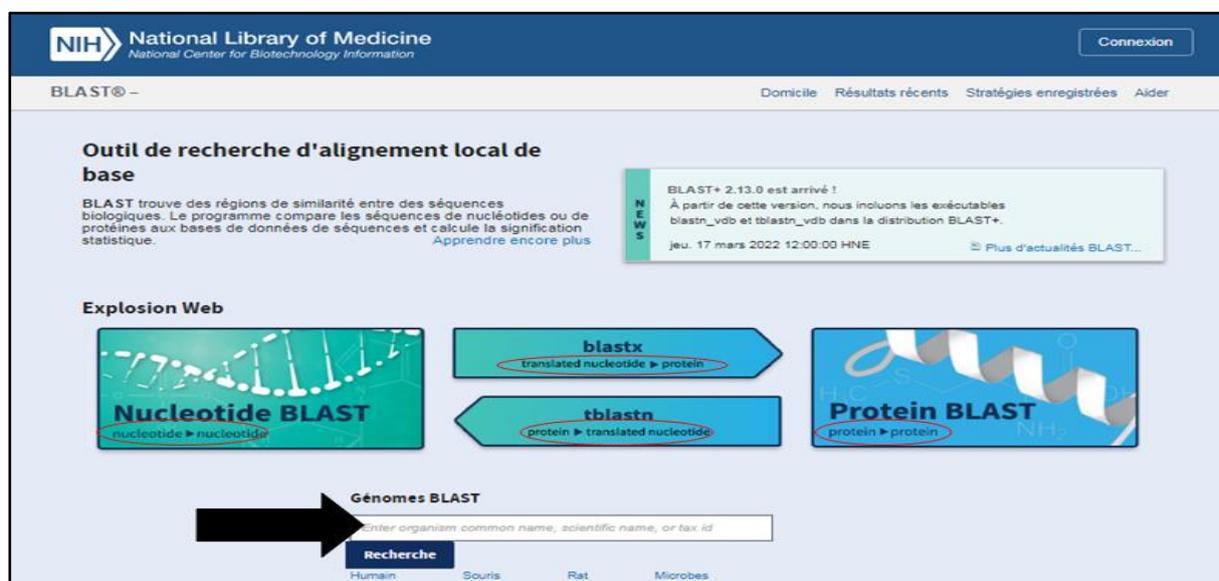


Figure 21 : Page d'accueil de BLAST.

- Nous avons travaillé sur des séquences nucléotidiques, donc on choisit "**nucléotide BLAST**", et nous procéderons au reste des étapes comme indiqué dans la **figure 22**.

MATERIEL ET METHODES

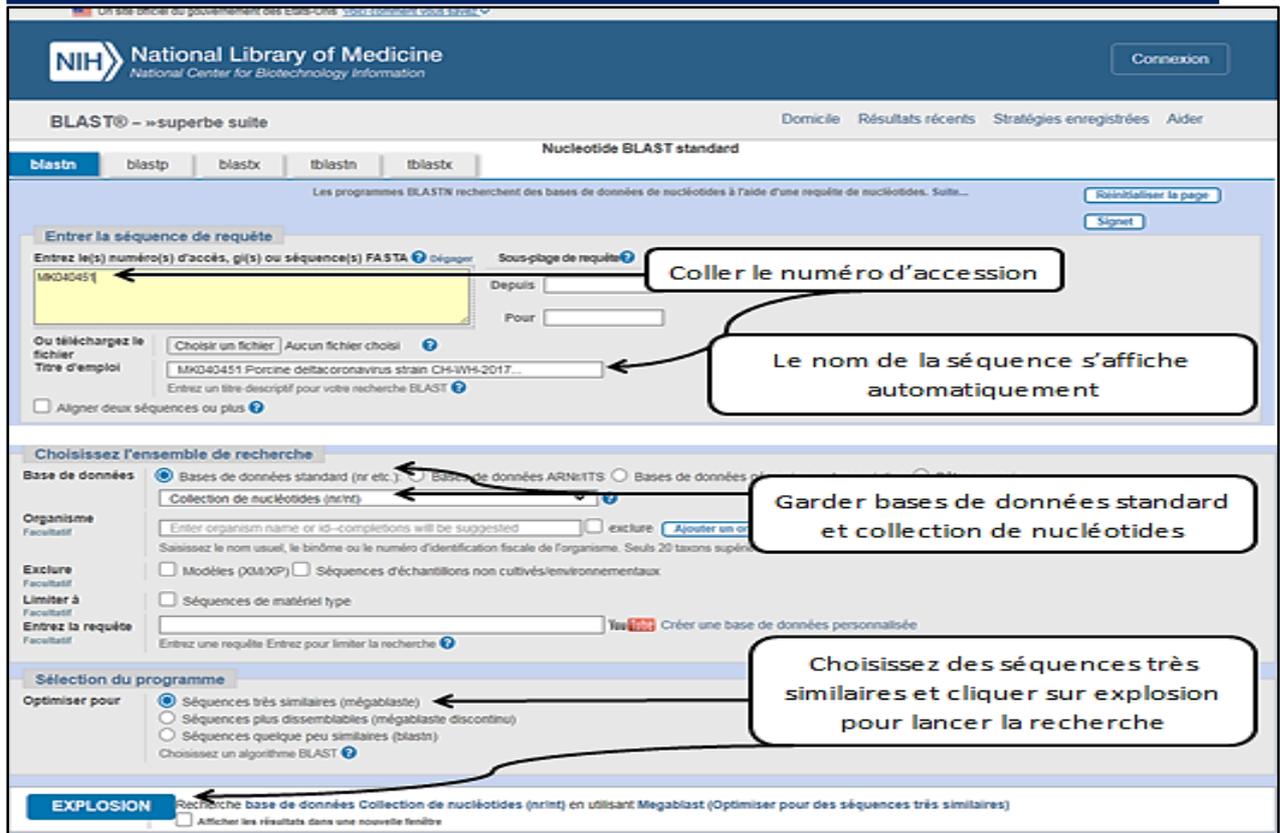


Figure 22 : Recherche des séquences à haute similarités.

- BLAST affiche les informations appropriées aux séquences recherchées en quatre :
 1. "Description" : Les séquences sont classées dans un tableau de colonnes nous mentionnons les principaux : la description (nom complet de la séquence), le nom scientifique (c'est le nom de la souche), par identité (pourcentage d'identité de cette souche avec la souche ciblée) et le numéro d'accèsion (Figure 23).

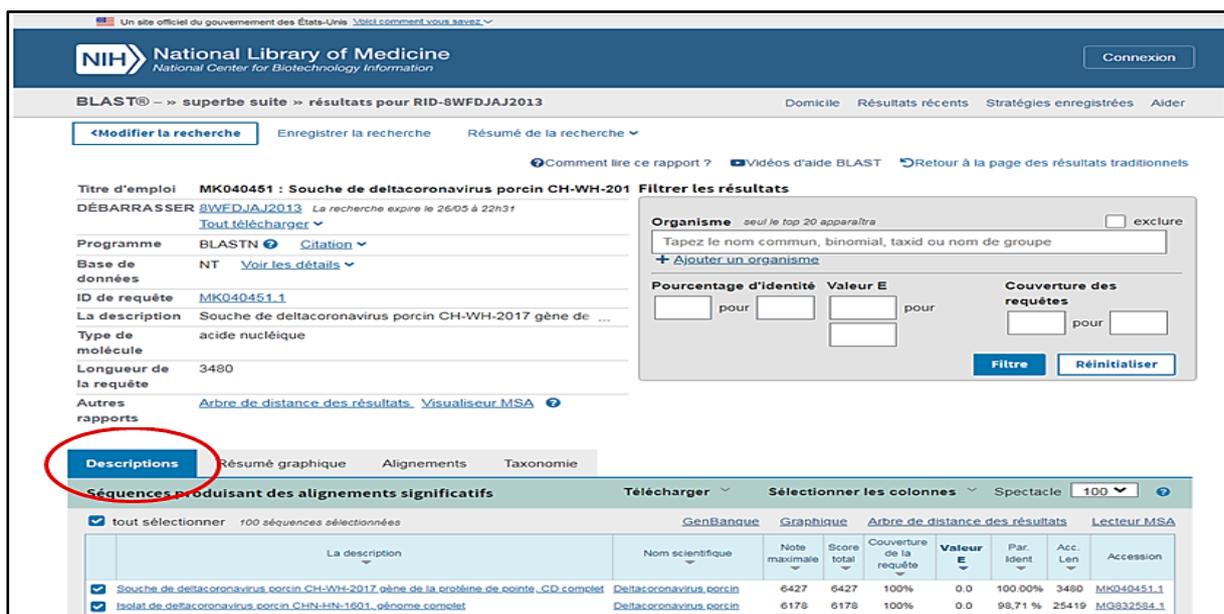


Figure 23 : Classement des séquences similaires dans la description.

MATERIEL ET METHODES

2. "Résumé graphique" : C'est la répartition des séquences par rapport aux scores d'alignements du plus bas au plus fort (dans notre cas tous les scores d'alignements des séquences sont supérieurs à 200 où ils prennent la couleur rouge) (Figure 24).

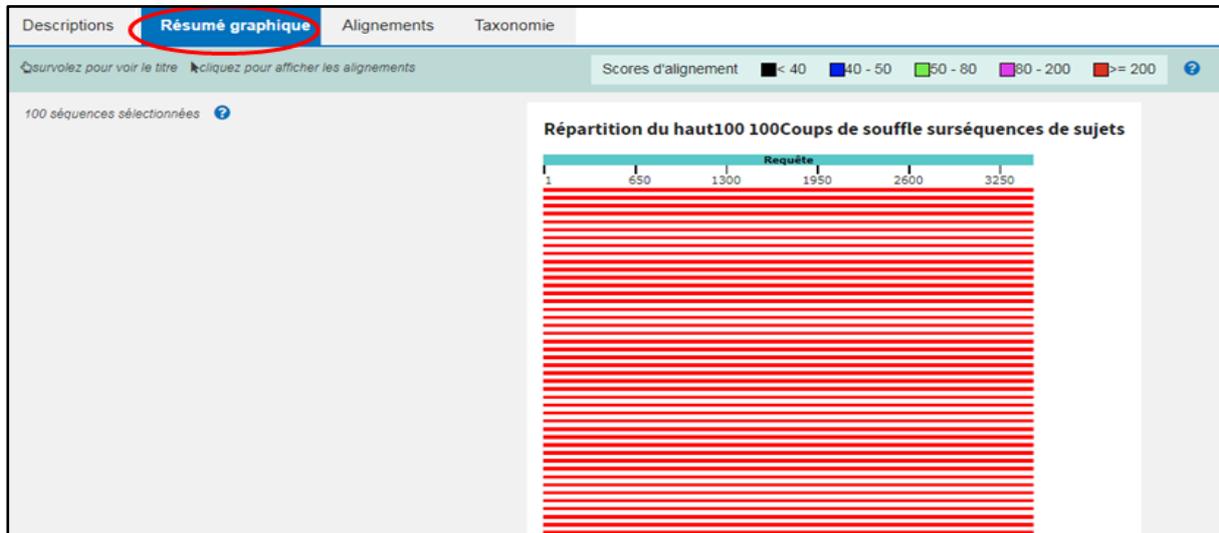


Figure 24 : Classement des séquences similaires par le résumé graphique.

3. "Alignement" des séquences avec ou sans la fonction CDS (Figure 25).

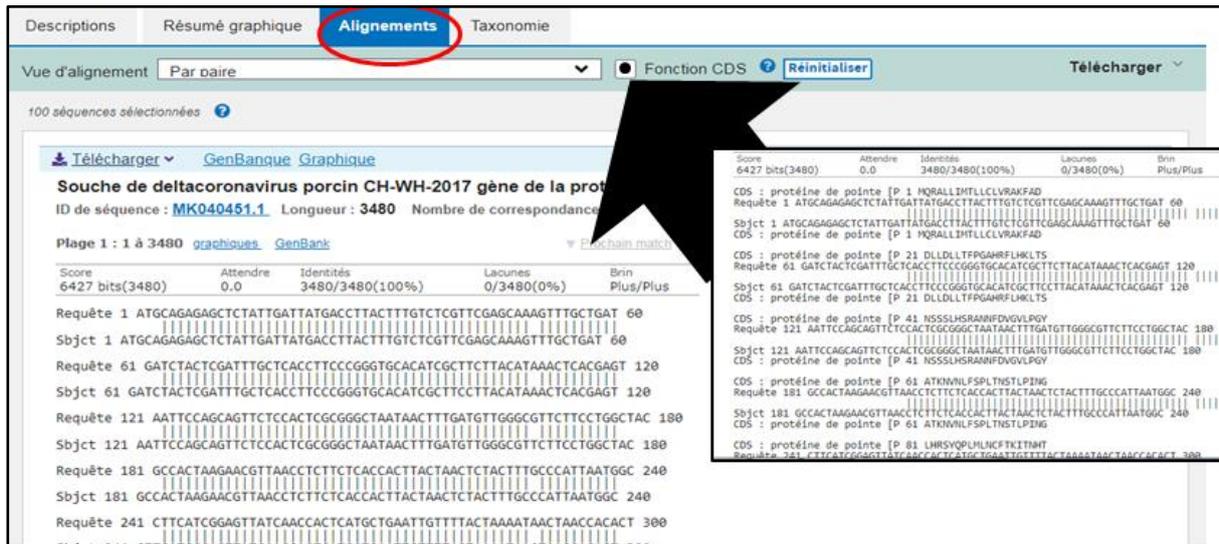


Figure 25 : Alignement des séquences par paire avec ou sans fonctions CDS.

4. "Taxonomie" : Qui s'affiche par rapport à la lignée, organisme ou taxonomie en propre terme. La figure 26 représente les trois cas de figure de taxonomie.

MATERIEL ET METHODES

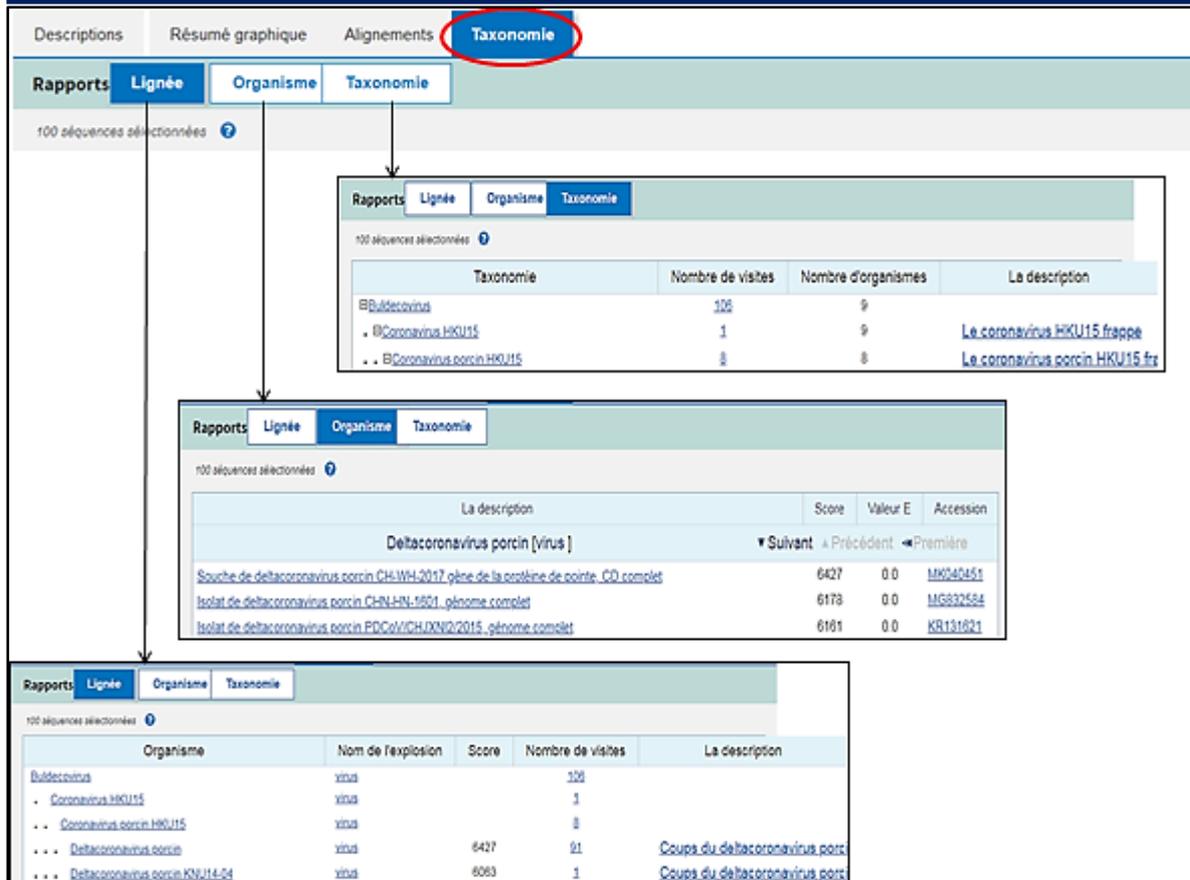


Figure 26 : Taxonomie des trois cas de figure d'affichage proposé par BLAST. Les informations de la GenBank de toutes les séquences seront classées dans l'annexe.

II.3. Traitement des séquences

II.3.1. Outil de MEGA 11

Principe : Moléculaire Génétique évolutive Analyse (MEGA) est un programme qui permet d'explorer et d'analyser des séquences nucléidiques et protéiques. En effet, dans l'étude d'évolution moléculaire il est souvent nécessaire de connaître certaines propriétés statistiques de base telles que les fréquences des différents nucléotides, les fréquences des codons et les rapports de transitions et transversions. Il permet également le calcul des distances évolutives entre les séquences sous différents modèles de substitution (Jukes et Cantor, Kimura à 2 paramètres et Tamura-Nei) et la reconstruction des arbres phylogénétiques selon trois méthodes, UPGMA Neighbor-Joining et Parcimonie (**Tumar et al., 2021**).

La page d'accueil de MEGA 11 présenté par la **figure 27**.

MATERIEL ET METHODES

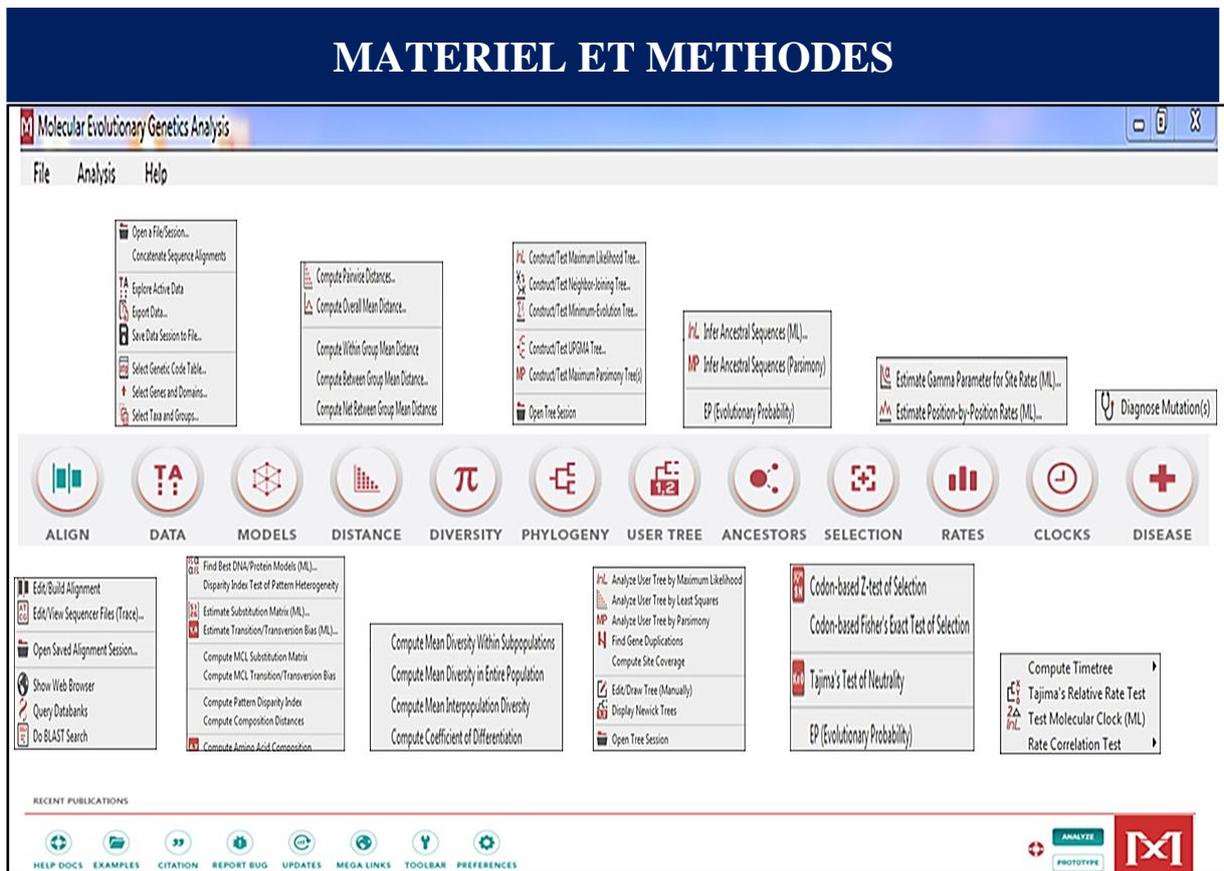


Figure 27 : Page d'accueil et barre d'outils de MEGA 11.

Afin de réaliser nos arbres phylogénétiques nous suivons les étapes :

- Choisir "**Align**".
- Choisir "**Edit/Build / Alignment**".
- Une boîte de dialogue s'affiche, choisir d'abord « **Create a new alignment** " pour un nouvel alignement.
- Ensuite ADN pour que une autre fenêtre " **Alignement Explorer** " comme elle est présentée dans la **figure 28**.

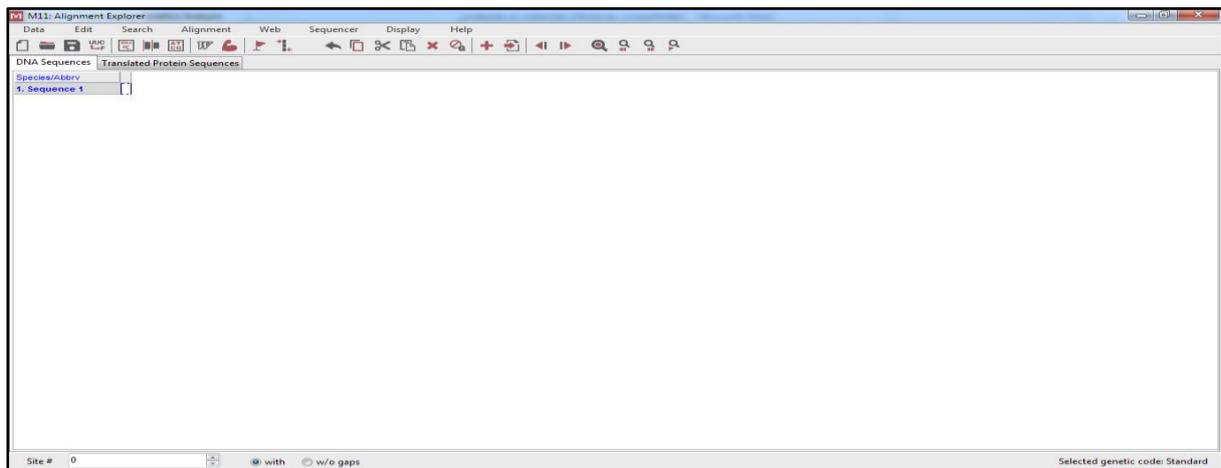


Figure 28 : Fenêtre « Alignement Explorer » de MEGA 11.

Une autre succession d'étapes se poursuit :

MATERIEL ET METHODES

- Aller à "Edit".
- Choisir "Insert Sequence From File".
- Dans la boîte de dialogue qui s'affiche sélectionner les séquences FASTA précédemment télécharger et appuyer sur ouvrir et les séquences seront installées dans la page de MEGA 11. comme se présente dans la **figure 29**.

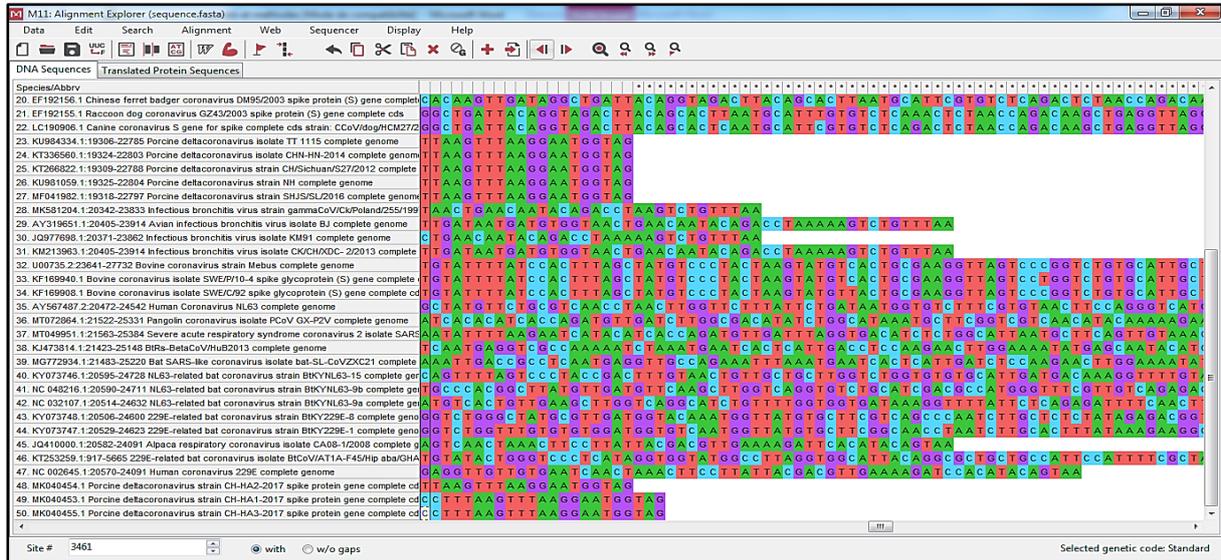


Figure 29 : Séquences d'ADN format FASTA.

- Ajuster la nomination des séquences (elles sont nommées par le numéro d'accèsion suivie du nom de la séquence dans la GenBank, nous avons choisi de les nommer par l'espèce puis le pays et ensuite le virus et la souche virale entre parenthèse).
- Par la suite aller à " **Alignement** ".
- Choisir l'une des méthodes d'alignements (nous avons choisi " **Alignement by clustal W** " car il incorpore un certain nombre d'améliorations à l'algorithme d'alignement, y compris la pondération de séquence, les pénalités d'écart spécifiques à la position et le choix automatique d'une matrice de comparaison de résidus appropriée à chaque étape de l'alignement multiple (**Julie et al., 1994**)).
- Ajuster la taille des séquences (**Figure 30**).

MATERIEL ET METHODES

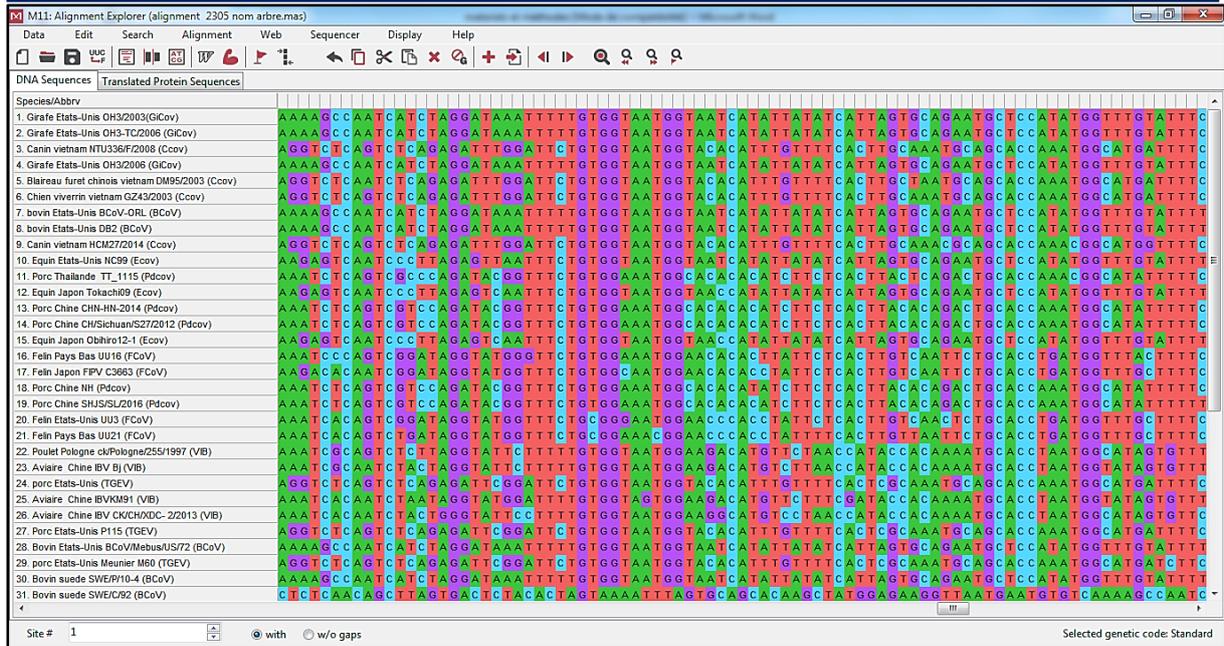


Figure 30 : Alignement des séquences par MEGA 11.

Après avoir un alignement de séquences à des tailles proches, sauvegarder l'alignement dans un fichier. Puis on commence la construction de l'arbre phylogénétique appropriée.

Pour cette phase nous avons suivi les étapes suivantes :

- Ouvrir à nouveau le programme MEGA 11.
- Aller à " **PHYLOGENIE** ".
- Choisir l'une des méthodes classées dans la liste, dans notre cas nous avons choisi les paramètres suivants (**Figure 31**) :
 - **"UPGMA"** comme type de méthode statistique car l'arbre phylogénétique généré par UPGMA n'est pas un arbre aléatoire mais plutôt un type spécial d'arbre qui dépend d'un algorithme simple qui remplit la propriété d'une horloge moléculaire de sorte que le temps total parcouru sur le chemin vers les feuilles de n'importe quel nœud est le même, quel que soit le chemin qu'on choisit et ainsi de suite, si l'arbre phylogénétique d'observation n'a pas cette propriété, il ne s'intégrera pas bien dans l'arbre UPGMA. En plus de la mesure extrêmement précise qui est une condition nécessaire à la production d'un meilleur arbre phylogénétique (**Chuang, 2007**).
 - Test de phylogénie de type : **"Bootstrap"** (**Efron et al., 1993**) qui teste la fiabilité des branches internes des arbres phylogénétiques. Il consiste à effectuer un tirage au hasard avec remise des sites. Dans chaque réplication de "Bootstrap", certains sites peuvent être présents plusieurs fois, tandis que d'autres peuvent être absents, où chaque réplication produit un nouveau alignement "artificiel" qui est utilisé pour construire un arbre

MATERIEL ET METHODES

"artificiel". Pour chaque branche interne on calcule le pourcentage des arbres "artificiels" contenant cette branche. On considère généralement que les branches définies par une valeur de "Bootstrap" supérieure à 95% sont fiables.

- Il faut "**1000 réplifications**" de "Bootstrap" pour que cette méthode soit statistiquement valable (**Cheikh, 2006**).
- Pour le type de substitution nous avons choisis "**nucléotide**" car il s'agit d'un modèle très simple, ne comportant qu'un seul paramètre: le taux de substitution d'une base d'ADN vers une autre. Ce modèle suppose la même probabilité de substitution quelle que soit la nature de la mutation. La fréquence de substitution utilisée par ce modèle est de 25% pour les quatre bases de l'ADN (**Cheikh, 2006**).
- "**Maximum composite likelihood**" choisi pour son aptitude à différencier les augmentations de diversification dans le temps d'un modèle constant (birth-death) où il y a présence constante d'extinction (Rabosky, 2006b). De plus, cette méthode permet de détecter des changements temporels dans les taux de diversification et ainsi d'inférer des événements de radiation rapide (**Crête-Lafrenière, 2009**).
- Pour le modèle nous avons choisis "**transition /transversions**" : Elles constituent près de 70 % des mutations : on distingue les transitions (remplacement d'une base pyrimidique (C ou T) ou purique (A ou G) par une autre base de même nature) et les transversions (remplacement d'une base purique par une base pyrimidique, ou inversement) (**Hanna et al., 2005**)
- "**Uniform rates among sites**" qui signifie le taux uniforme sur les sites et dans "**Same pattern among lineages**" nous avons choisis "**uniforme (homogène)**" car la variation des taux de substitution entre les sites (**Yang, 1996**) et les lignées (**Kolaczkowski et Thornton, 2004**), peut engendrer un signal non phylogénétique qui compétitionne dans la même analyse avec le signal phylogénétique authentique) (**Crête-Lafrenière, 2009**).
- "**Pairwise deletion**" (la délétion par paire) : L'approche de suppression par paires tente d'atténuer la perte de données en utilisant tous les cas disponibles dans la procédure d'analyse des données, produisant un sous-ensemble différent d'observations utilisées pour calculer chaque élément de la matrice de corrélation polychorique (**Shi et al., 2019**).

MATERIEL ET METHODES

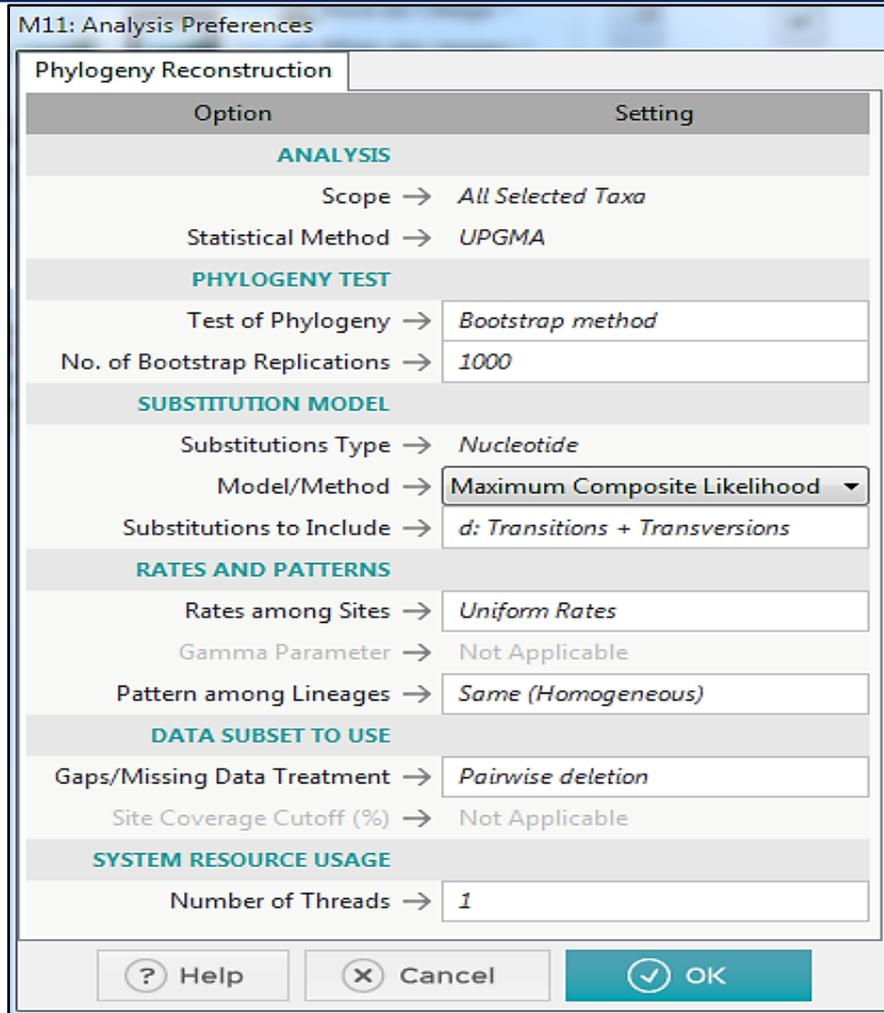


Figure 31 : L'ajustement des paramètres de construction de l'arbre par MEGA 11.

*Pour obtenir les tableaux de comparaison entre les séquences d'ADN nous avons utilisé un logiciel nommé GénieGen 2 (Génie Génétique).

II.3.2. Outil de GenieGen 2

GenieGen2 est un logiciel qui permet de travailler sur des séquences d'ADN, d'ARN ou de protéines. La page d'accueil de la GenieGen 2 s'affiche comme se présente dans la **figure 32**.

MATERIEL ET METHODES

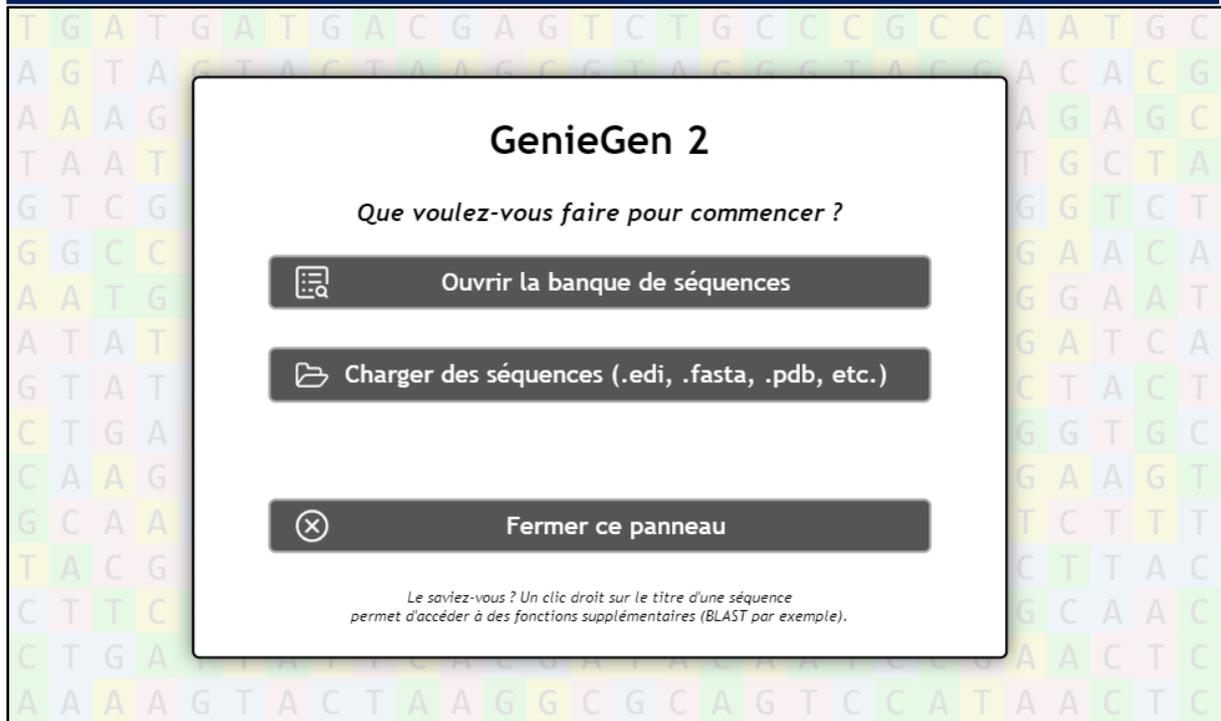


Figure 32 : Page d'accueil de GeniGen 2.

Afin de réaliser nos tableaux de comparaison nous passerons par ces étapes distinctes :

- Construire l'URL.
- Choisir "**charger des séquences (.edi)**" sur le panneau d'accueil pour charger les séquences FASTA précédemment téléchargées.
- Sélectionner les séquences ; une autre fenêtre s'ouvre celle présente dans la **figure 33**.

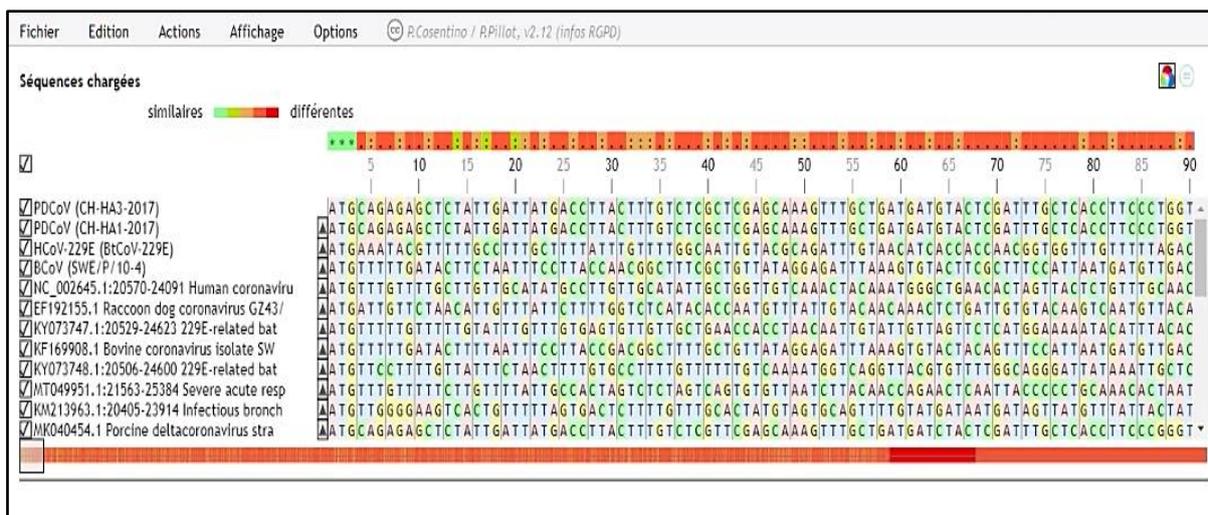


Figure 33 : Séquences FASTA dans la GeniGen 2.

- Ensuite on va trier les séquences, après on choisit "**Edition**" dans la barre d'outils qu'on trouve en haut, puis on sélectionne "**Supprimer les séquences sélectionnées**".

MATERIEL ET METHODES

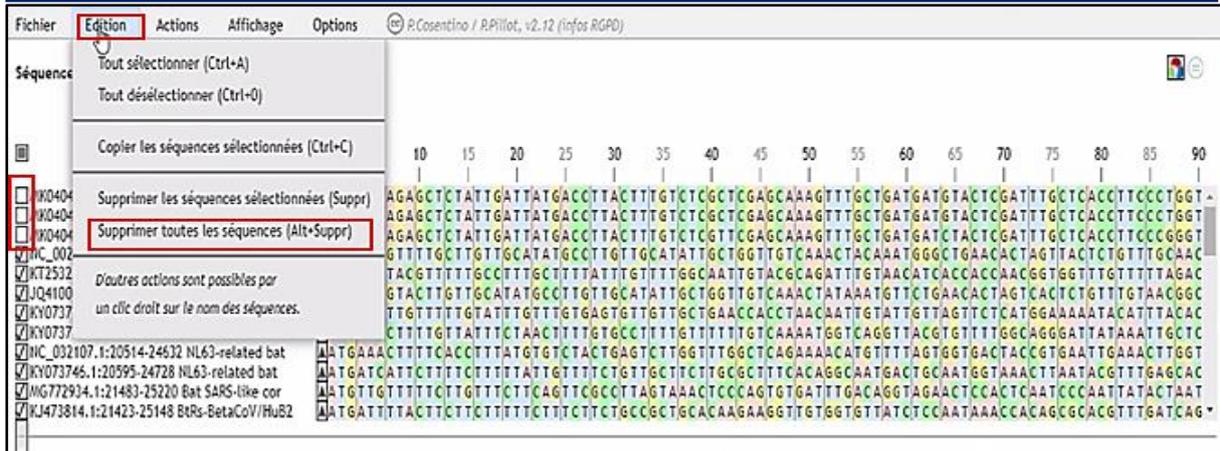


Figure 34 : Suppression des séquences non ciblées.

- On renomme nos séquences, en choisissant "**Renommer**" dans la boîte de dialogue (qui apparaît en cliquant deux fois à gauche), puisque le site n'accepte pas une longue nomination on a opté à donner juste le nom du virus puis celui de la souche. La succession de ces étapes se présente dans les deux **figures 35 et 36**.



Figure 35 : Renommer les séquences ciblées.

MATERIEL ET METHODES

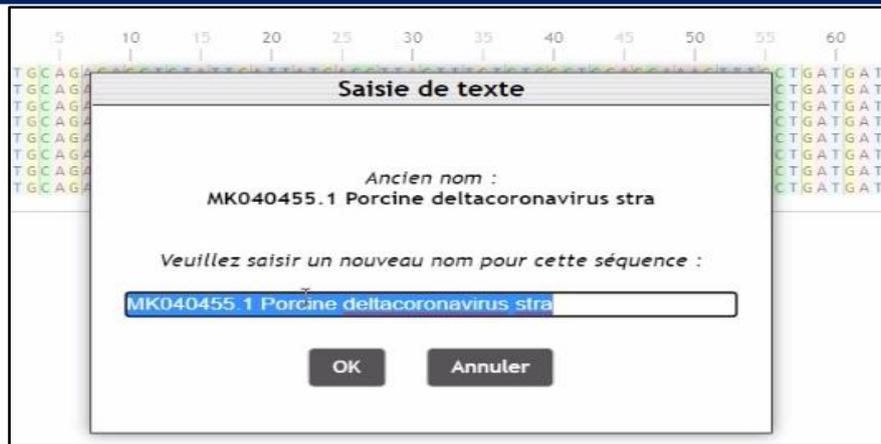


Figure 36 : Choisir un nouveau nom.

- Afin d'aligner les séquences et obtenir les tableaux de comparaison une autre succession d'étapes se poursuit :

- Aller à "**Actions**" puis sélectionné "**Aligner les séquences sélectionnées**".

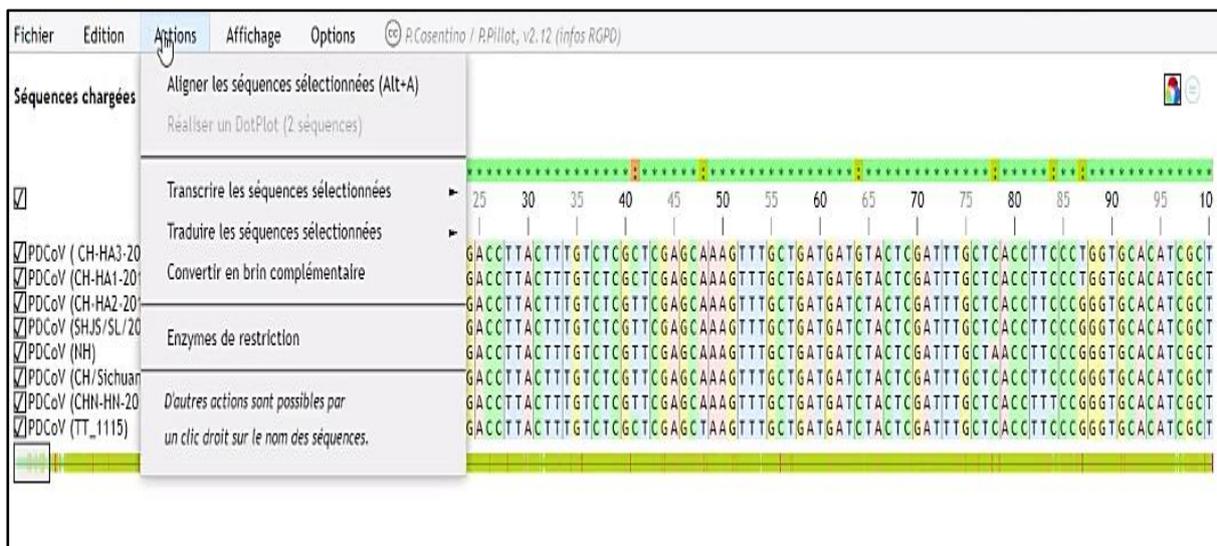


Figure 37 : Alignement des séquences sélectionnées.

Dans la **figure 38** se trouve une représentation d'une ligne de comparaison colorée qui met en évidence les positions des nucléotides (ou les acides aminés). Ceux qui sont identiques : vert ; les différents : Orange ou rouge (selon le degré de différence).

MATERIEL ET METHODES



Figure 38 : Aligement des séquences par GenieGen 2.



Figure 39 : Action d'affichage du tableau de comparaison.

- On choisit
- + "Affichage" puis on sélectionne "Tableau de comparaison" figure 40.

MATERIEL ET METHODES

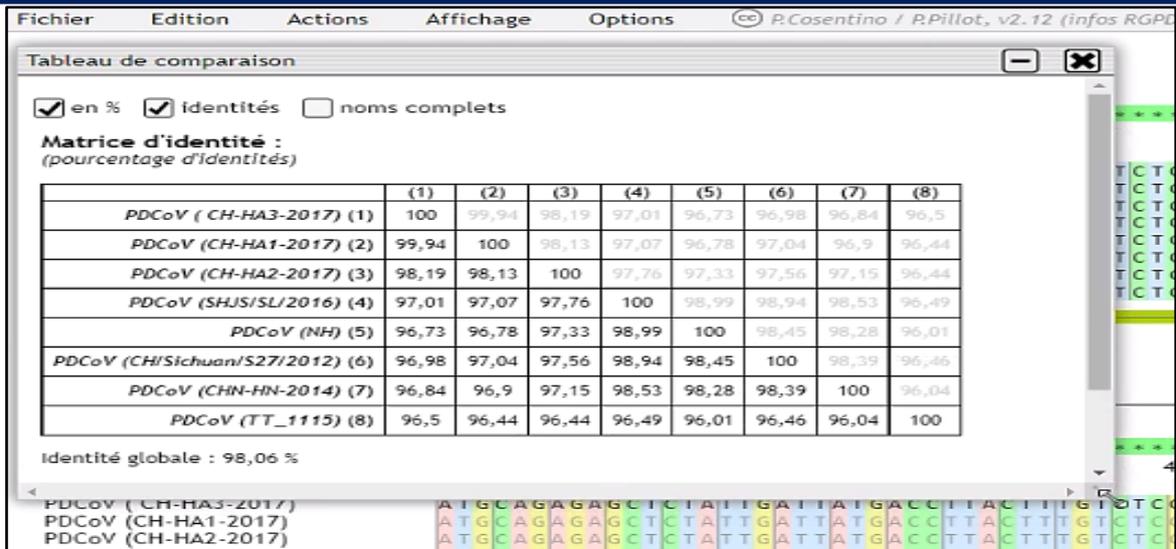


Figure 40 : Tableau de comparaison de GeniGen 2.

RESULTATS

III. Résultats

Dans cette partie, nous avons examiné les résultats obtenus qui reposent sur la comparaison entre les séquences d'ADN de différentes souches virales de Coronavirus. Le traitement des séquences par des logiciels de bio-informatiques nous a permis de construire les arbres phylogénétiques ainsi que de calculer les pourcentages de similarités entre ces séquences, afin de désigner les liens entre les différentes séquences de souches dans le même gène, nous avons calculé les pourcentages de similarités de ces séquences entre elles en moyenne générale au sein du même gène. Cette dernière a été réalisée à l'aide d'un programme spécialisé.

Dans cette étude, nous avons analysé les séquences nucléotidiques de différentes souches de Coronavirus publiées dans la GenBank, mais nous n'avons choisi que 250 séquences nucléotidiques qui présentent les trois gènes : gène de pointe S (gène structurale), gène N (gène de nucléocapside) et gène RdRp (gène non structurale) qui sont en commun entre plusieurs espèces animales infectés.

Dans cette étape qui suit nous présenterons les résultats phylogéniques présentés dans chacun des trois arbres : arbre du gène S, arbre du gène N, arbre du gène RdRp.

III.1. Résultats de l'arbre de la glycoprotéine S

Dans cette partie, nous avons étudié 70 séquences nucléotidiques de différentes souches virales du gène S, puis nous avons ciblé 49 séquences représentatifs de Coronavirus, qui sont mentionnés dans le **tableau II** selon les espèces animales touchées dans plusieurs pays du monde.

Tableau II : Les séquences d'ADN de Coronavirus utilisé pour la présentation phylogénétique du gène S.

Animal	Espèce	Pays	Virus
Porc	<i>Porc</i>	Etats-Unis, chine	TGEV, PDCoV PRCV
	<i>Sus scrofa</i>	Chine	
Chien	Chien	Taiwan	CCoV
	<i>Canis lupus familiaris</i>	Vietnam	
Blaireau furet chinois	<i>Melogale moschata</i>	Chine	CCoV

RESULTATS

Chien viverrin	Nyctereutes procyonoides	Chine	CCoV
Félin	<i>Felis catus</i>	Japon	FCoV
	Chat	Pays Bas, Etats-Unis	
	Félin	Pays Bas	
Chauve-souris	<i>Triaenops after</i>	Kenya	HCoV NL63, HCoV-229E, SARS 2-CoV
	<i>Hipposideros abae</i>	Ghana	
	<i>Hipposideros vittatus</i>	Kenya	
	<i>Hipposideros sp</i>	Kenya	
	<i>Rhinolophus sinicus</i>	Chine	
	<i>Rhinolophus pusillus</i>	Chine	
Humain	<i>Homo sapiens</i>	Amsterdam, Etats- Unis, Chine	HCoV-NL63, HCoV-229E, SARS 2-CoV
Alpaga	<i>Vicugna pacos</i>	Etats-Unis	HCoV-229E
Aviaire	Aviaire	Chine, Corée du sud, Pologne	IBV
Pangolin	Pangolin	Chine	PCoV
Equin	<i>Equus caballus</i>	Japon	ECoV
	Equin	Etats-Unis	
Bovin	Veau	Suède	BCoV
	Bovin	Etats-Unis	
	Bétail	Suède	

RESULTATS

Girafe	Girafe	Etats-Unis	GiCoV
--------	--------	------------	-------

L'arbre phylogénétique (**Figure 35**) tracé par le logiciel MEGA 11 est réalisé sur la base du gène S de 49 différentes souches virales du Coronavirus qui ont touchés plusieurs espèces animales et l'espèce humaine avec une taille de 3915 Pb.

RESULTATS

En raison du grand nombre de séquences, ainsi qu'afin de bien visualiser ce grand arbre nous proposons les figures suivantes qui regroupent les séquences de chaque groupe viral.

Commençant par le groupe α qui se divise dans notre arbre en deux sous-groupes : le sous-groupe αa et le sous-groupe αb ; les deux se représentent dans les figures : **Figure 36** et la **figure 37** successivement.

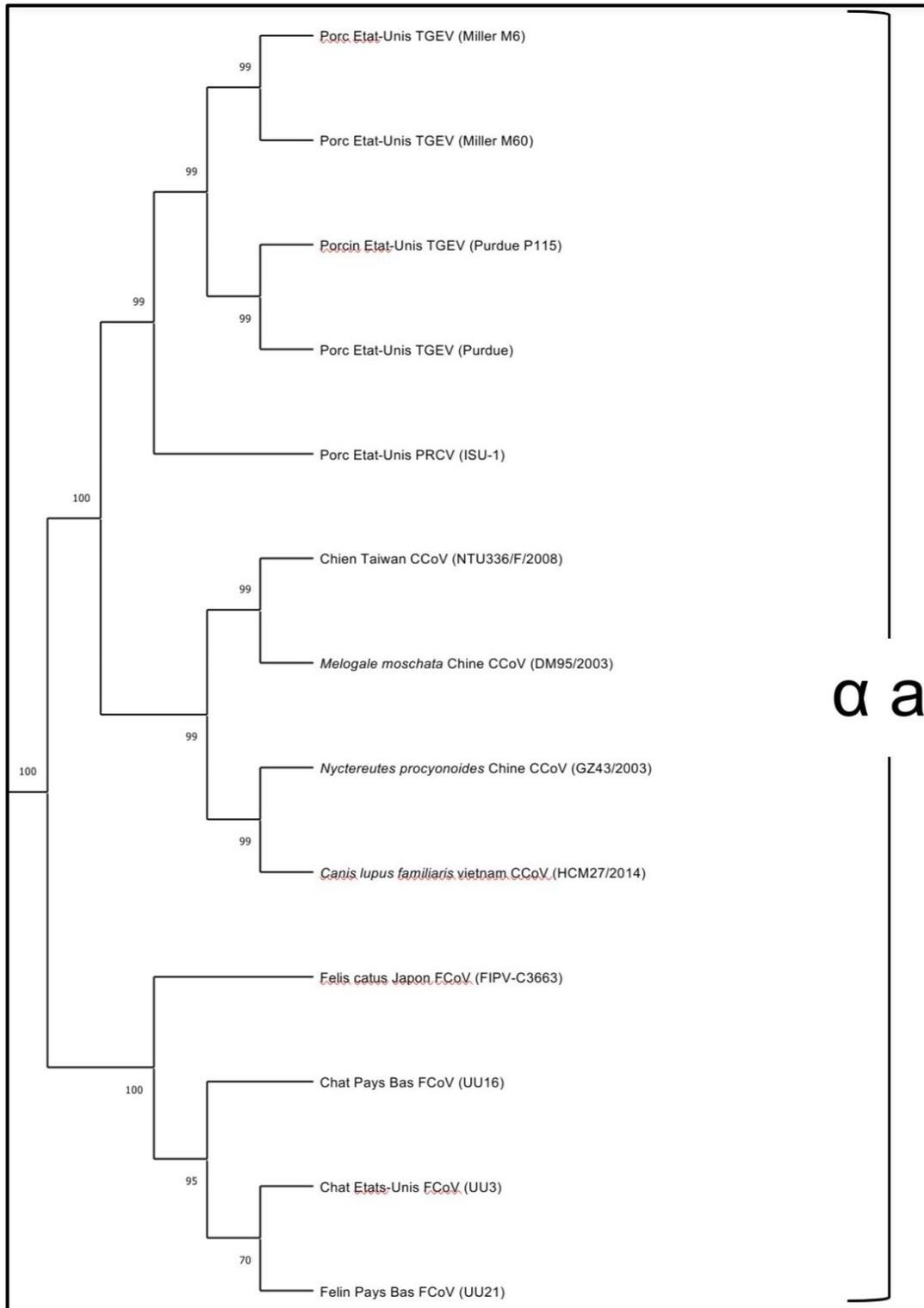


Figure 36 : Fragment de l'arbre du gène S comportant les séquences du sous-groupe αa .

RESULTATS

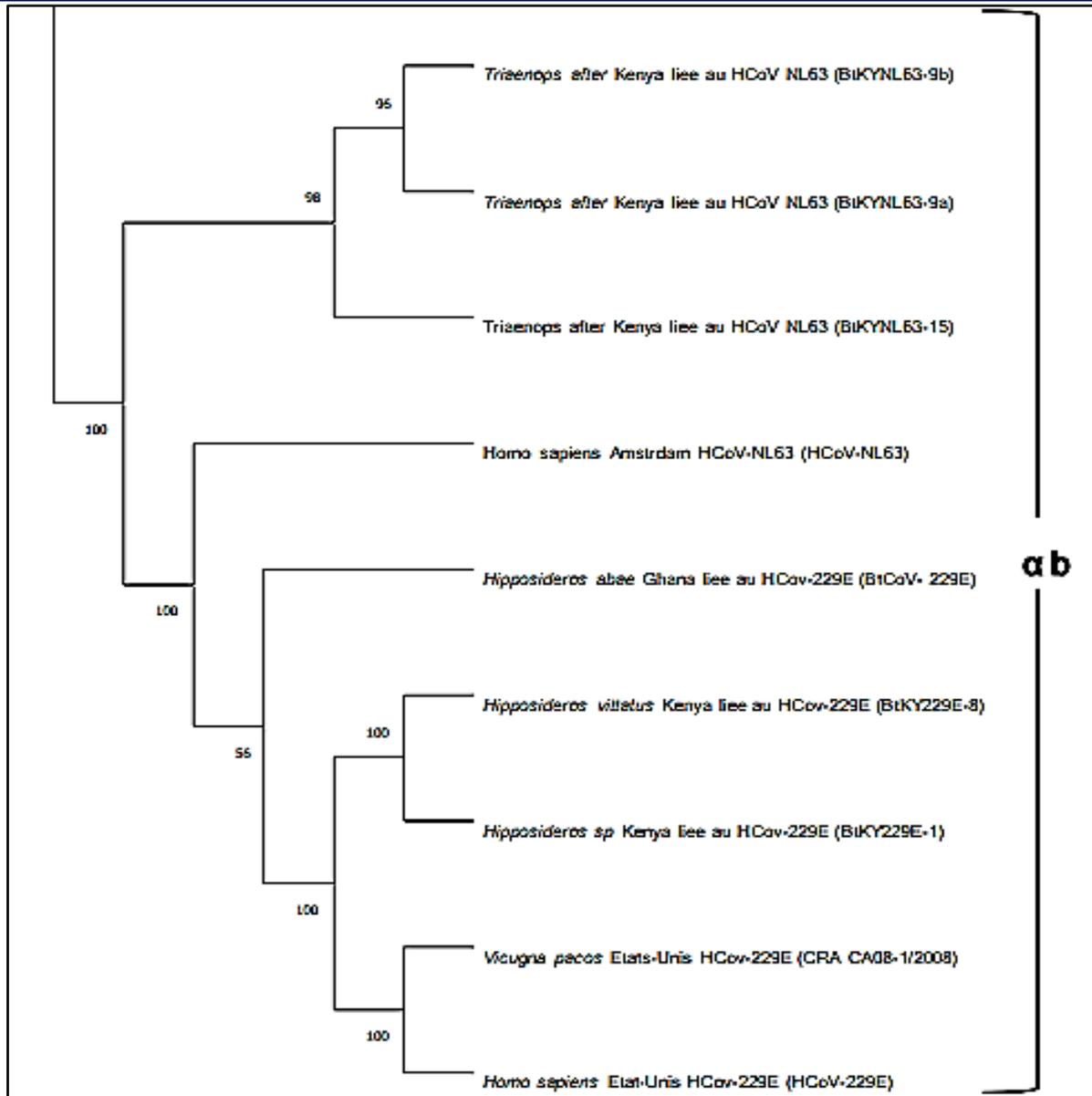


Figure 37 : Fragment de l'arbre du gène S incluant les séquences du sous-groupe αb . Les séquences qui comportent le gène S et qui se regroupent dans le groupe δ ; sont présentées dans la **figure 38**.

RESULTATS

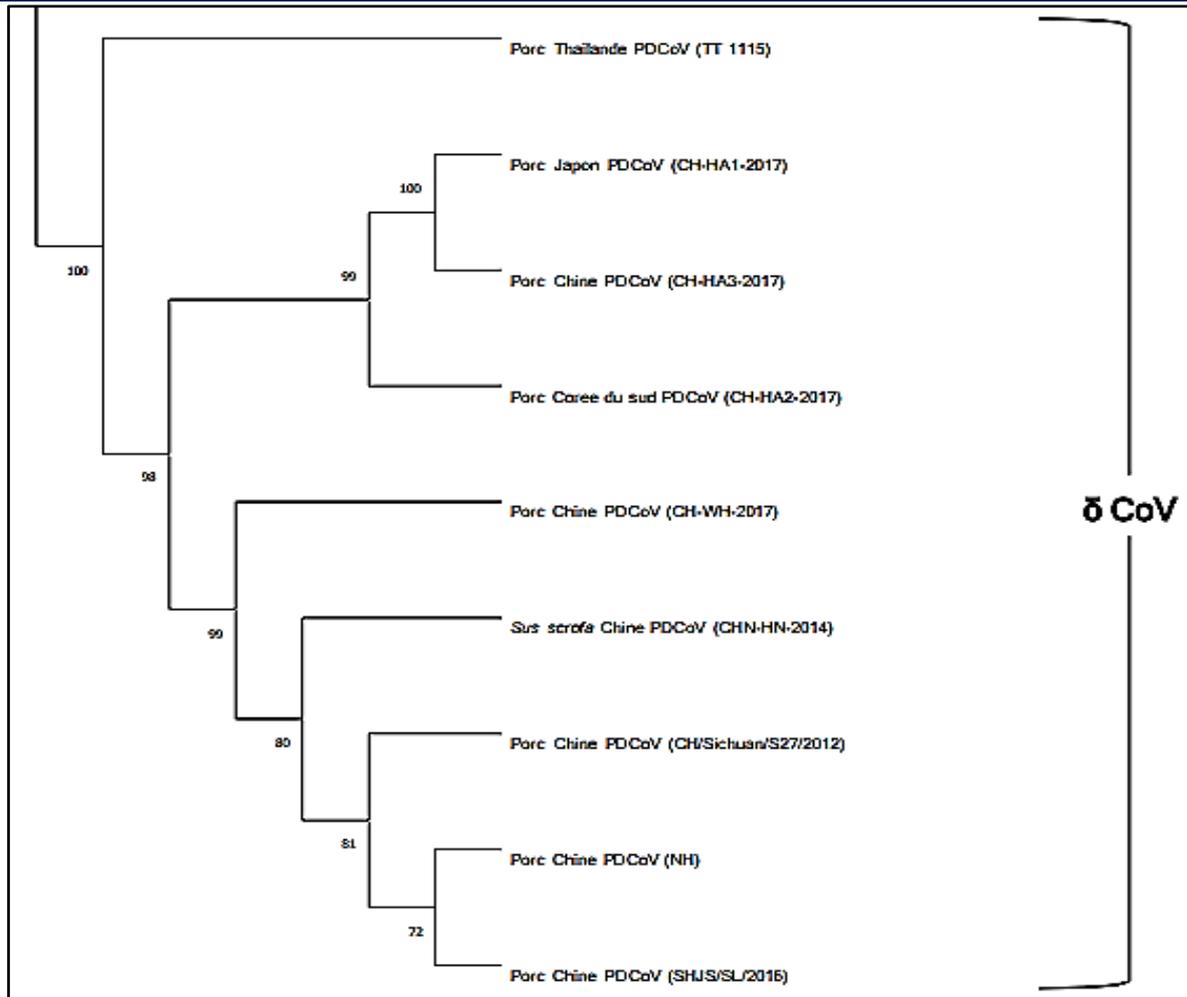


Figure 38 : Fragment de l'arbre du gène S incluant les séquences du groupe δ .

La figure suivante (**Figure 39**) présente le fragment de l'arbre du gène S qui comporte les séquences du groupe γ .

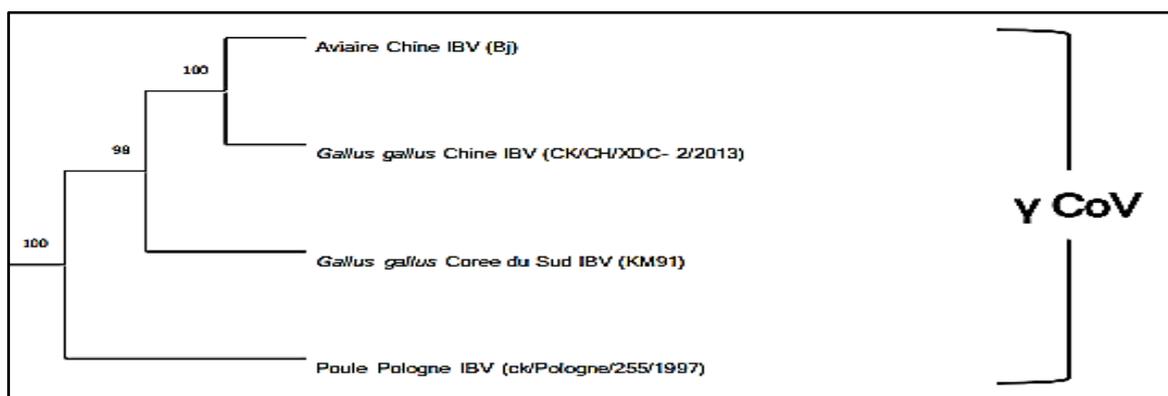


Figure 39 : Fragment de l'arbre du gène S incluant les séquences du groupe γ .

Les séquences du groupe β qui se divise dans notre arbre en deux sous-groupes : le sous-groupe β_b et le sous-groupe β_a ; les deux sont représentés dans les figures : **Figure 40** et la **figure 41** successivement.

RESULTATS

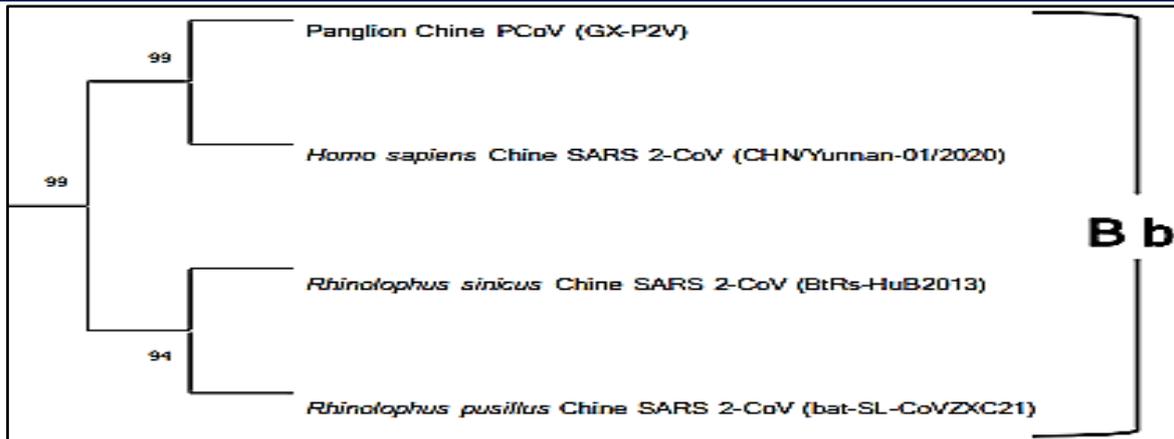


Figure 40 : Fragment de l'arbre du gène S comportant les séquences du sous-groupe β b.

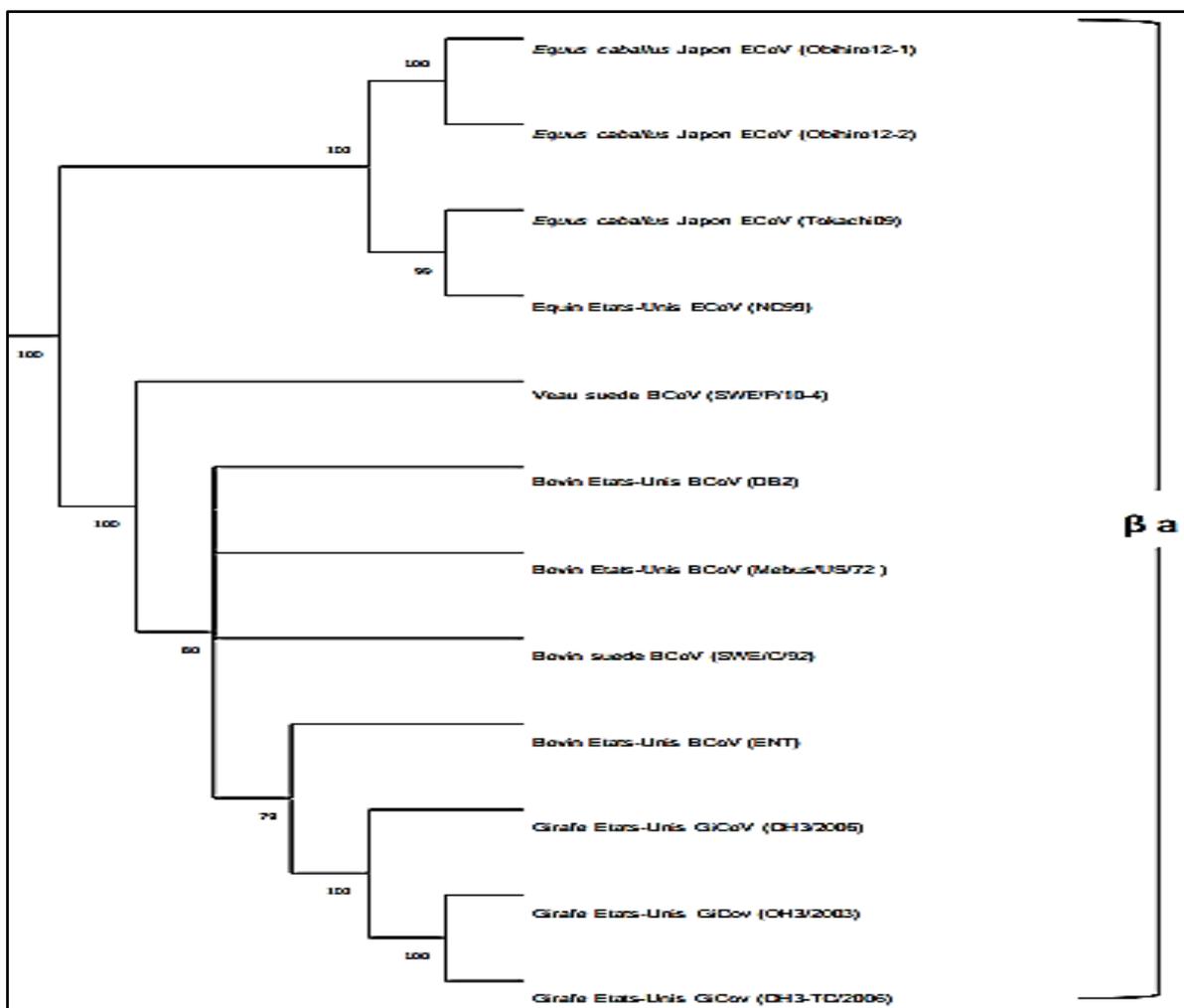


Figure 41 : Fragment de l'arbre du gène S comportant les séquences du sous-groupe β a.

Selon la classification mondiale des souches virales du Coronavirus, ainsi que les séquences d'ADN du gène S. Cet arbre phylogénétique se divise en quatre groupes, Alphacoronavirus, Betacoronavirus, Gamma coronavirus, Deltacoronavirus :

Alphacoronavirus (α CoV) contenant 21 séquences qui se divisent à leurs tours en 2 sous-groupes :

RESULTATS

- **Sous-groupe A** : 12 séquences nucléotidiques classées en :
 - 3 séquences du Gastro-entérite transmissible (**TGEV**) des Etats-Unis.
 - 1 séquence de Coronavirus respiratoire porc (**PRCV**) des Etats-Unis.
 - 4 séquences de Coronavirus Canine (**CCoV**) du Vietnam et de Taïwan qui infecte le chien *Canis lupus familiaris* et celui de la Chine qui infecte le chien viverrin *Nyctereutes procyonoides* et le blaireau furet chinois *Melogale moschata*.
 - 4 séquences de Coronavirus Félin (**FCoV**) une du Japon qui infecte *Felis catus*, 2 des Pays bas et les Etats-Unis.
- **Sous-groupe B** : Comprend 9 séquences nucléotidiques de :
 - 5 séquences de Coronavirus humain 229E (**HCoV-229E**) une qui infecte la chauve-souris *Hipposideros sp* de Kenya, *Hipposideros abae* et *Hipposideros vittatus de* Ghana, et une infecte aussi l'humain et l'alpaga *Vicugna pacos* des Etats-Unis.
 - 4 séquences de Coronavirus humain NL63 (**HCoV-NL63**) une d'Amsterdam infecte l'homme, et 3 du Kenya qui touche la chauve-souris *Triaenops sp*.

Les pourcentages de similarités entre les différentes séquences de 19 souches du groupe Alphacoronavirus (α CoV) sont représentés dans le **tableau III** avec une identité moyenne de 86,8%.

RESULTATS

Tableau III : Pourcentages d'identité des différentes séquences de souches du groupe Alphacoronavirus (α CoV) qui présentent le gène S.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
<i>HCov-229E (HCov-229E) (1)</i>	100	60,23	56,11	69,58	55,1	55,37	51,1	55,04	72,9	73,24	53,54	55,62	56,08	94,26	60,64	55,07	67,02	54,34	55,18
<i>HCov NL63 (BtKYNL63-15) (2)</i>	60,23	100	46,66	55,91	48,04	47,04	44,42	46,66	56,35	55,26	47,9	47,59	48,43	60,33	54,32	48,22	56,19	47,8	45,92
<i>CCoV (HCM27/2014) (3)</i>	56,11	46,66	100	48,96	72,18	68,46	45,15	70,6	51,17	50,02	68,68	76,58	76,31	55,04	45,97	72,54	50,07	48,36	47,01
<i>HCov-229E (BtCoV-229E) (4)</i>	69,58	55,91	48,96	100	50,27	50,05	45,25	48,76	63,72	64,45	48,48	50,09	50,56	68,76	54,46	50,12	61,84	47,94	46,71
<i>TGEV (Miller M60) (5)</i>	55,1	48,04	72,18	50,27	100	65,75	43,96	85,21	50,51	51,02	84,03	80,71	78,65	54,75	47,39	95,29	51,32	48,78	45,6
<i>CCoV (GZ43/2003) (6)</i>	55,37	47,04	68,46	50,05	65,75	100	42,57	64,33	49,83	50,77	63,85	72,04	72,04	54,63	47,93	65,75	52,01	47,25	47,66
<i>FCoV (UU21) (7)</i>	51,1	44,42	45,15	45,25	43,96	42,57	100	42,81	44,18	44,27	42,8	44,08	44,33	50,05	41,23	43,95	45,34	59,43	51,99
<i>TGEV (Miller M6) (8)</i>	55,04	46,66	70,6	48,76	85,21	64,33	42,81	100	50,4	47,83	76,99	75,47	75,38	54,76	47,19	83,07	50,24	48,3	46,57
<i>HCov-229E (BtKY229E-1) (9)</i>	72,9	56,35	51,17	63,72	50,51	49,83	44,18	50,4	100	70,51	49,12	51,05	51,22	72,95	55,08	50,41	65,22	50,52	48,5
<i>HCov-229E (BtKY229E-8) (10)</i>	73,24	55,26	50,02	64,45	51,02	50,77	44,27	47,83	70,51	100	49,98	51,16	51,61	73,2	54,12	49,97	67,55	48,96	48,39
<i>PRCV (ISU-1) (11)</i>	53,54	47,9	68,68	48,48	84,03	63,85	42,8	76,99	49,12	49,98	100	74,08	74,58	52,82	46,2	83,55	50,55	46,95	44,91
<i>CCoV (DM95/2003) (12)</i>	55,62	47,59	76,58	50,09	80,71	72,04	44,08	75,47	51,05	51,16	74,08	100	90,8	54,89	47,53	81,39	52,11	48,5	47,26
<i>CCoV (NTU336/FI/2008) (13)</i>	56,08	48,43	76,31	50,56	78,65	72,04	44,33	75,38	51,22	51,61	74,58	90,8	100	55,5	46,59	79,48	52,84	48,75	47,05
<i>HCov-229E (CA08-1/2008) (14)</i>	94,26	60,33	55,04	68,76	54,75	54,63	50,05	54,76	72,95	73,2	52,82	54,89	55,5	100	59,19	54,87	66,92	53,79	54,24
<i>HCov NL63 (BtKYNL63-9a) (15)</i>	60,64	54,32	45,97	54,46	47,39	47,93	41,23	47,19	55,08	54,12	46,2	47,53	46,59	59,19	100	47,44	56,17	48,09	47,26
<i>TGEV (Purdue P115) (16)</i>	55,07	48,22	72,54	50,12	95,29	65,75	43,95	83,07	50,41	49,97	83,55	81,39	79,48	54,87	47,44	100	51,6	48,37	45,43
<i>HCov NL63 (HCov-NL63) (17)</i>	67,02	56,19	50,07	61,84	51,32	52,01	45,34	50,24	65,22	67,55	50,55	52,11	52,84	66,92	56,17	51,6	100	49,53	47,91
<i>FCoV (FIPV-C3663) (18)</i>	54,34	47,8	48,36	47,94	48,78	47,25	59,43	48,3	50,52	48,96	46,95	48,5	48,75	53,79	48,09	48,37	49,53	100	64,83
<i>FCoV (UU3) (19)</i>	55,18	45,92	47,01	46,71	45,6	47,66	51,99	46,57	48,5	48,39	44,91	47,26	47,05	54,24	47,26	45,43	47,91	64,83	100

Identité globale : 86,8 %

Deltacoronavirus (δ CoV) : Ce groupe représente 9 séquences de Deltacoronavirus du porc (PDCoV) une de la Thaïlande, 7 de la Chine, et une qui infecte le cochon *Sus scrofa*.

Pourcentages de similarités entre les différentes séquences de 9 souches du groupe Deltacoronavirus (δ CoV) avec une identité moyenne de 98,18 % sont représentés dans le **tableau IV**.

RESULTATS

Tableau IV: Pourcentages d'identités de différentes séquences de souches du groupe Deltacoronavirus (δ CoV) qui présentent le gène S.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>PDCoV (CH-HA2-2017) (1)</i>	100	97,64	97,56	97,21	97,39	97,76	98,19	98,13	96,49
<i>PDCoV (CH-WH-2017) (2)</i>	97,64	100	98,16	97,82	98,16	98,45	96,75	96,81	96,15
<i>PDCoV (CH/Sichuan/S27/2012) (3)</i>	97,56	98,16	100	98,45	98,51	98,94	96,98	97,04	96,52
<i>PDCoV (CHN-HN-2014) (4)</i>	97,21	97,82	98,45	100	98,33	98,59	96,84	96,9	96,06
<i>PDCoV (NH) (5)</i>	97,39	98,16	98,51	98,33	100	99,05	96,73	96,78	96,01
<i>PDCoV (SHJS/SL/2016) (6)</i>	97,76	98,45	98,94	98,59	99,05	100	97,01	97,07	96,55
<i>PDCoV (CH-HA3-2017) (7)</i>	98,19	96,75	96,98	96,84	96,73	97,01	100	99,94	96,5
<i>PDCoV (CH-HA1-2017) (8)</i>	98,13	96,81	97,04	96,9	96,78	97,07	99,94	100	96,44
<i>PDCoV (TT 1115) (9)</i>	96,49	96,15	96,52	96,06	96,01	96,55	96,5	96,44	100

Identité globale : 98,18 %

Gammacoronavirus (γ CoV) : Comprend 4 séquences du virus de la bronchite infectieuse aviaire (**IBV**) qui infecte l'aviare *Gallus gallus*. Ces séquences se rendent compte à des souches virales qui se partagent en paies du monde qui sont : une souche présente dans la Corée du sud, une de Pologne et 2 autres souches représentées dans la Chine.

Pourcentages de similarités entre les différentes séquences de souches du groupe Gammacoronavirus (γ CoV) avec une identité globale de 85,99% sont représentés dans le **tableau V**.

Tableau V : Pourcentages d'identité de différentes séquences de 4 souches du groupe Gammacoronavirus (γ CoV) qui présentent le gène S.

	(1)	(2)	(3)	(4)
<i>IBV (CK/CH/XDC- 2/2013) (1)</i>	100	83,79	87,84	80,82
<i>IBV (KM91) (2)</i>	83,79	100	83,04	81,72
<i>IBV (BJ) (3)</i>	87,84	83,04	100	81,28
<i>IBV (ck/Pologne/255/1997) (4)</i>	80,82	81,72	81,28	100

Identité globale : 85,99 %

Betacoronavirus (β CoV) constante 15 séquences divisées en 2 sous-groupes :

➤ **Sous-groupe A :** Il comprend 12 séquences nucléotidiques de :

- 5 séquences de Coronavirus de bovin (**BCoV**) 2 de la Suède et 3 des Etats-Unis.
- 3 séquences de Coronavirus de la girafe (**GiCoV**) des Etats-Unis qui infectent la girafe *Giraffa camelopardalis*.
- 4 séquences de Coronavirus équin (**ECoV**) une des Etats-Unis et 3 du Japon qui infectent *Eqqus caballus*.

➤ **Sous-groupe B :** Il comprend 4 séquences nucléotidiques de :

RESULTATS

- 3 séquences de Coronavirus du syndrome respiratoire aigu sévère (**SRAS-CoV-2**) une de la chine qui touche l'humain, une infectant les chauves-souris *Rhinolophus sinicus* et une touche les chauves-souris *Rhinolophus pusillus*.
- Une séquence de Coronavirus du pangolin (**PCoV**) de chine.

Pourcentages de similarités entre les différentes séquences de souches du groupe Gammacoronavirus (γ CoV) avec une identité moyenne de 85,8% sont représentés dans le **tableau VI**.

Tableau VI : Pourcentages d'identité de différentes séquences de 15 souches du groupe Gammacoronavirus (γ CoV) qui présentent le gène S.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
<i>BCoV (SWE/PI10-4)</i> (1)	100	97,85	97,4	79,81	97,4	54,67	79,67	97,36	79,59	97,9	55,17	56,44	97,51	55,62	97,4
<i>BCoV (SWE/CI92)</i> (2)	97,85	100	97,87	79,89	97,87	53,97	79,77	98,09	79,59	98,51	54,58	54,94	98,14	54,57	97,87
<i>GiCoV (OH3/2006)</i> (3)	97,4	97,87	100	79,87	100	53,71	79,73	97,82	79,68	98,55	54,72	55,41	99,04	54,68	100
<i>ECoV (Tokachi09)</i> (4)	79,81	79,89	79,87	100	79,87	50,26	99,02	79,59	98,53	80,01	49,46	46,67	79,94	48,23	79,87
<i>GiCoV (OH3/2003)</i> (5)	97,4	97,87	100	79,87	100	53,71	79,73	97,82	79,68	98,55	54,72	55,41	99,04	54,68	100
<i>SARS 2-CoV (CHN/Yunnan-01/2020)</i> (6)	54,67	53,97	53,71	50,26	53,71	100	50,18	53,9	49,85	53,79	61,16	44,03	53,71	43,06	53,71
<i>ECoV (NC99)</i> (7)	79,67	79,77	79,73	99,02	79,73	50,18	100	79,4	98,51	79,79	49,62	46,31	79,79	48,09	79,73
<i>BCoV (Mebus/USI72)</i> (8)	97,36	98,09	97,82	79,59	97,82	53,9	79,4	100	79,4	98,63	54,36	55,09	98,07	54,6	97,82
<i>ECoV (Obihiro12-1)</i> (9)	79,59	79,59	79,68	98,53	79,68	49,85	98,51	79,4	100	79,77	49,46	46,47	79,77	47,81	79,68
<i>BCoV (DB2)</i> (10)	97,9	98,51	98,55	80,01	98,55	53,79	79,79	98,63	79,77	100	54,58	55,3	98,92	54,78	98,55
<i>SARS 2-CoV (BtRs-HuB2013)</i> (11)	55,17	54,58	54,72	49,46	54,72	61,16	49,62	54,36	49,46	54,58	100	47,68	54,68	40,98	54,72
<i>SARS 2-CoV (bat-SL-CoVZXC21)</i> (12)	56,44	54,94	55,41	46,67	55,41	44,03	46,31	55,09	46,47	55,3	47,68	100	55,51	35,94	55,41
<i>BCoV (ENT)</i> (13)	97,51	98,14	99,04	79,94	99,04	53,71	79,79	98,07	79,77	98,92	54,68	55,51	100	54,74	99,04
<i>PCoV (GX-P2V)</i> (14)	55,62	54,57	54,68	48,23	54,68	43,06	48,09	54,6	47,81	54,78	40,98	35,94	54,74	100	54,68
<i>GiCoV (OH3-TC/2006)</i> (15)	97,4	97,87	100	79,87	100	53,71	79,73	97,82	79,68	98,55	54,72	55,41	99,04	54,68	100

Identité globale : 85,3 %

III.2. Résultats de l'arbre du gène de la protéine N

Les séquences d'ADN de différentes souches virales du gène N de Coronavirus extraites sont mentionnées dans le **tableau VII** en fonction des espèces animales affectées à travers plusieurs pays dans le monde.

RESULTATS

Tableau VII : Les séquences d'ADN de Coronavirus utilisé pour la présentation phylogénétique du gène N.

Animal	Espèce	Région	Gène	Virus
Porc	Porc	Japon, Chine Etats-Unis, Laos, Thaïlande, Corée de sud	N	PEDV, PDCoV
	<i>Sus scrofa</i>	Philippines		
Félin	Félin	Thaïlande, Etats- Unis	N	FCoV
Bovin	Bovin	Canada, Italie France,	N	BCoV
Canin	<i>Canis lupus familiaris</i>	Allemagne	N	CCoV
	Chien	Chine, Corée du sud	N	CRCoV
Chauve-souris	<i>Rhinolophus sinicus</i>	Chine	N	SARS-CoV 2
	<i>Rhinolophus affinis</i>			
	<i>Rhinolophus pusillus</i>			
Canard siffleur	<i>Mareca penelope</i>	Chine	N	WiCoV
Oiseau sauvage	<i>Spatula discors</i>	Etats-Unis	N	DCoV
Humain	<i>Homo sapiens</i>	Chine, Belgique	N	HCoV-OC43

L'arbre phylogénétique (**Figure 42**) étudié par le logiciel MEGA 11 est réalisé sur la base du gène N de 36 différentes souches virales du Coronavirus qui ont touchés plusieurs espèces animales et l'espèce humaine avec une taille qui varie entre 1112 Pb et 991 Pb.

RESULTATS

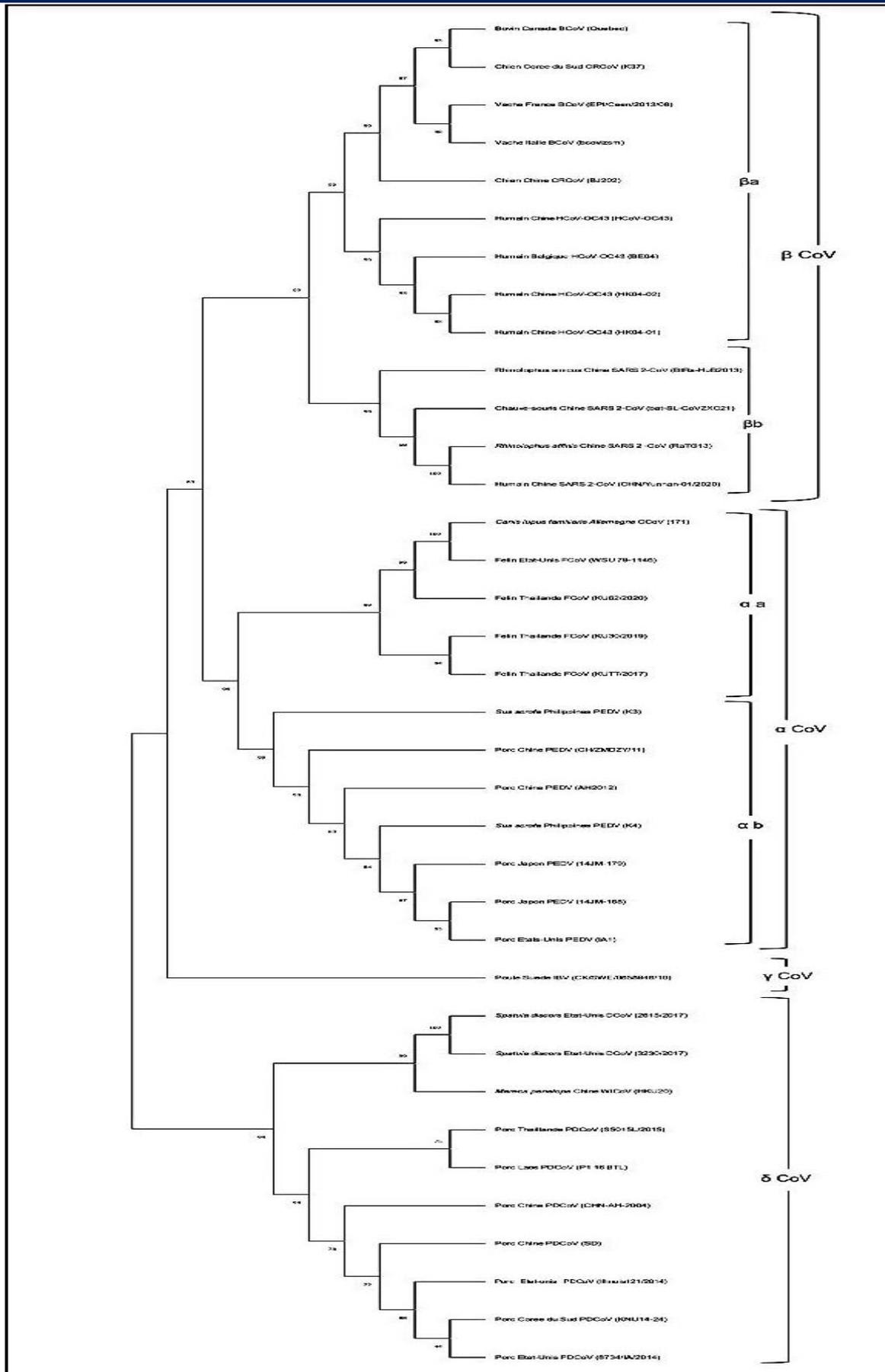


Figure 42 : L'arbre phylogénétique du gène N.

RESULTATS

En raison du grand nombre de séquence, ainsi qu'afin de bien visualiser cet arbre nous proposons les figures suivantes qui présentes les séquences de chaque groupe viral.

Commençant par le groupe β qui se divise dans cet arbre en deux sous-groupes : le sous-groupe βa et le sous-groupe βb les deux sont représentés dans les figures **figure 43** et la **figure 44** successivement.

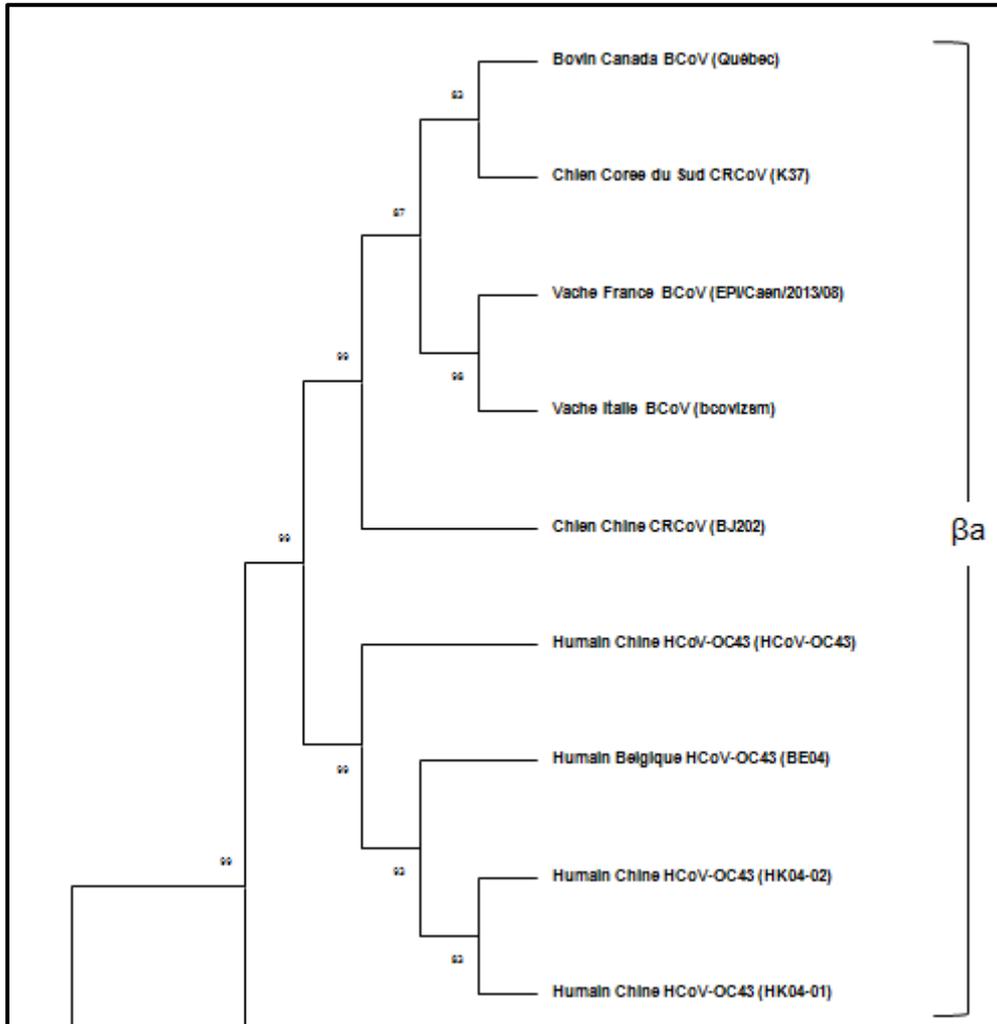


Figure 43 : Fragment de l'arbre du gène N comportant les séquences du sous-groupe βa .

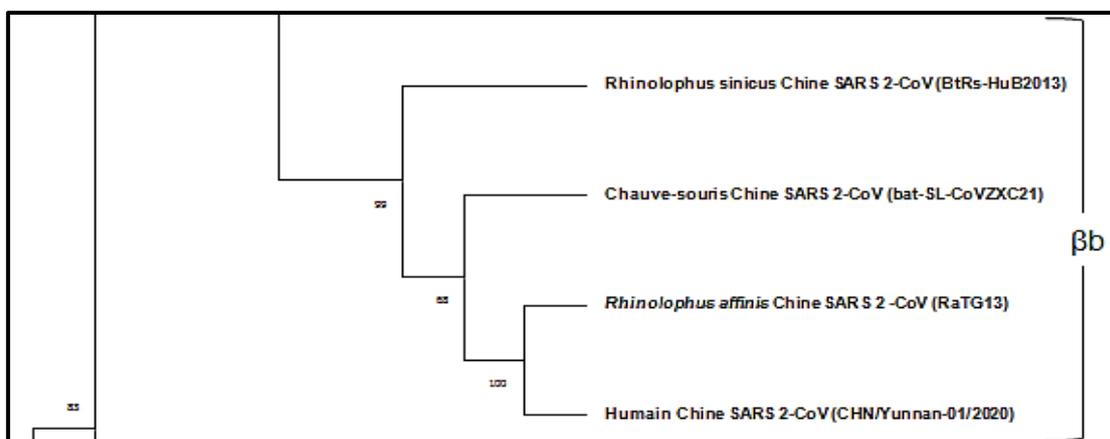


Figure 44 : Fragment de l'arbre du gène N comportant les séquences du sous-groupe βb .

RESULTATS

En deuxième lieu le groupe α qui se divise à son tour dans notre arbre en deux sous-groupes : le sous-groupe αa et le sous-groupe αb les deux sont représentés dans les figures **figure 45** et la **figure 46** successivement.

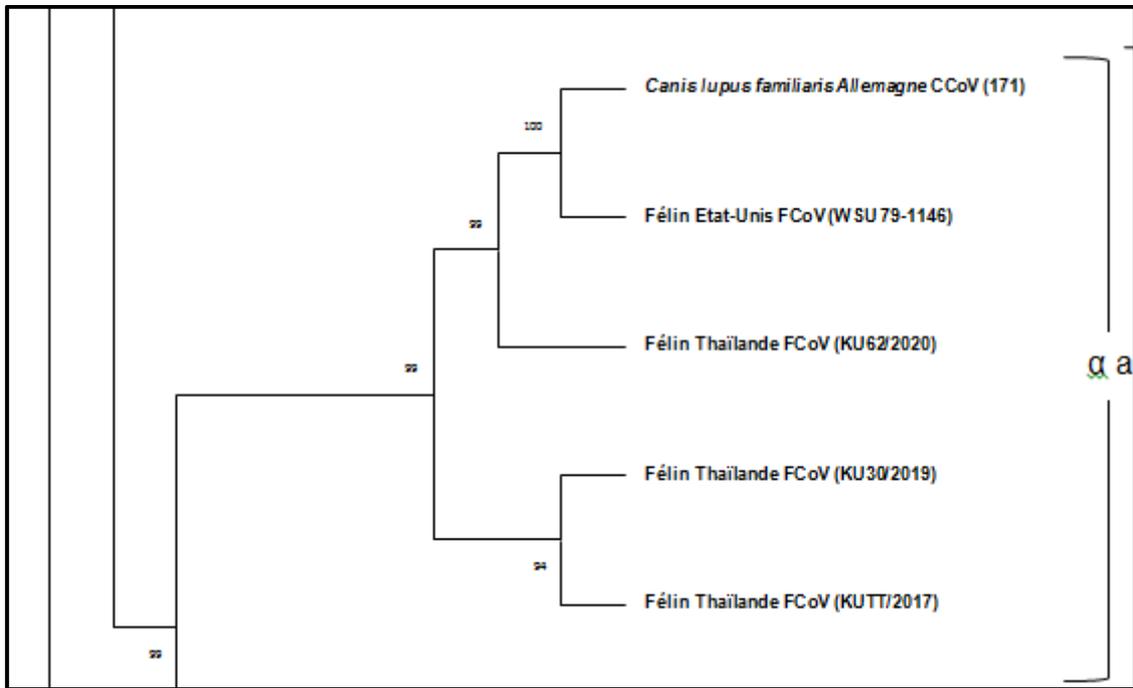


Figure 45 : Fragment de l'arbre du gène N comportant les séquences du sous-groupe αa .

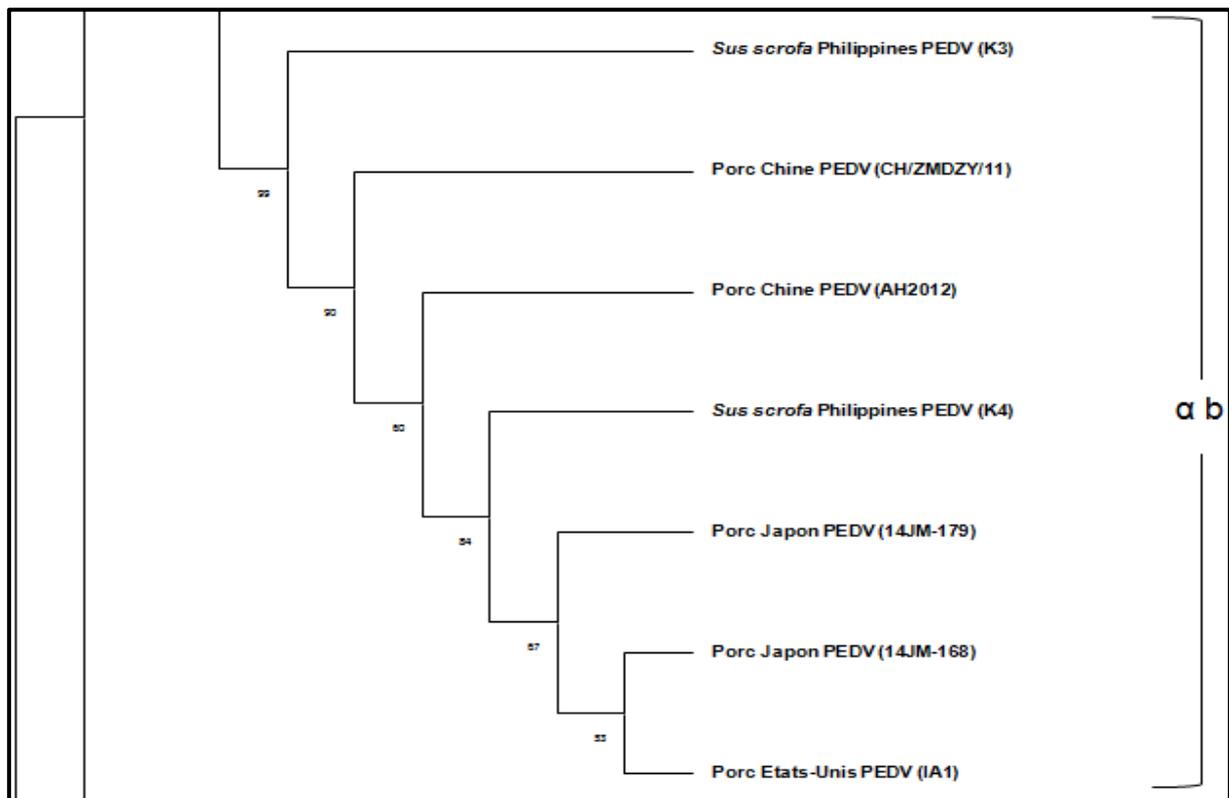


Figure 46 : Fragment de l'arbre du gène N comportant les séquences du sous-groupe αb .

RESULTATS

Par la suite nous avons la **figure 47** qui représente l'emplacement de la séquence présentatrice du groupe γ .

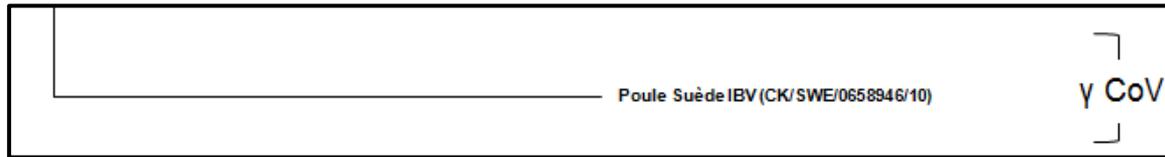


Figure 47 : Fragment de l'arbre du gène N comportant les séquences du groupe γ .

La **figure 48** qui présente le fragment de l'arbre du gène N qui comporte les séquences du groupe δ .

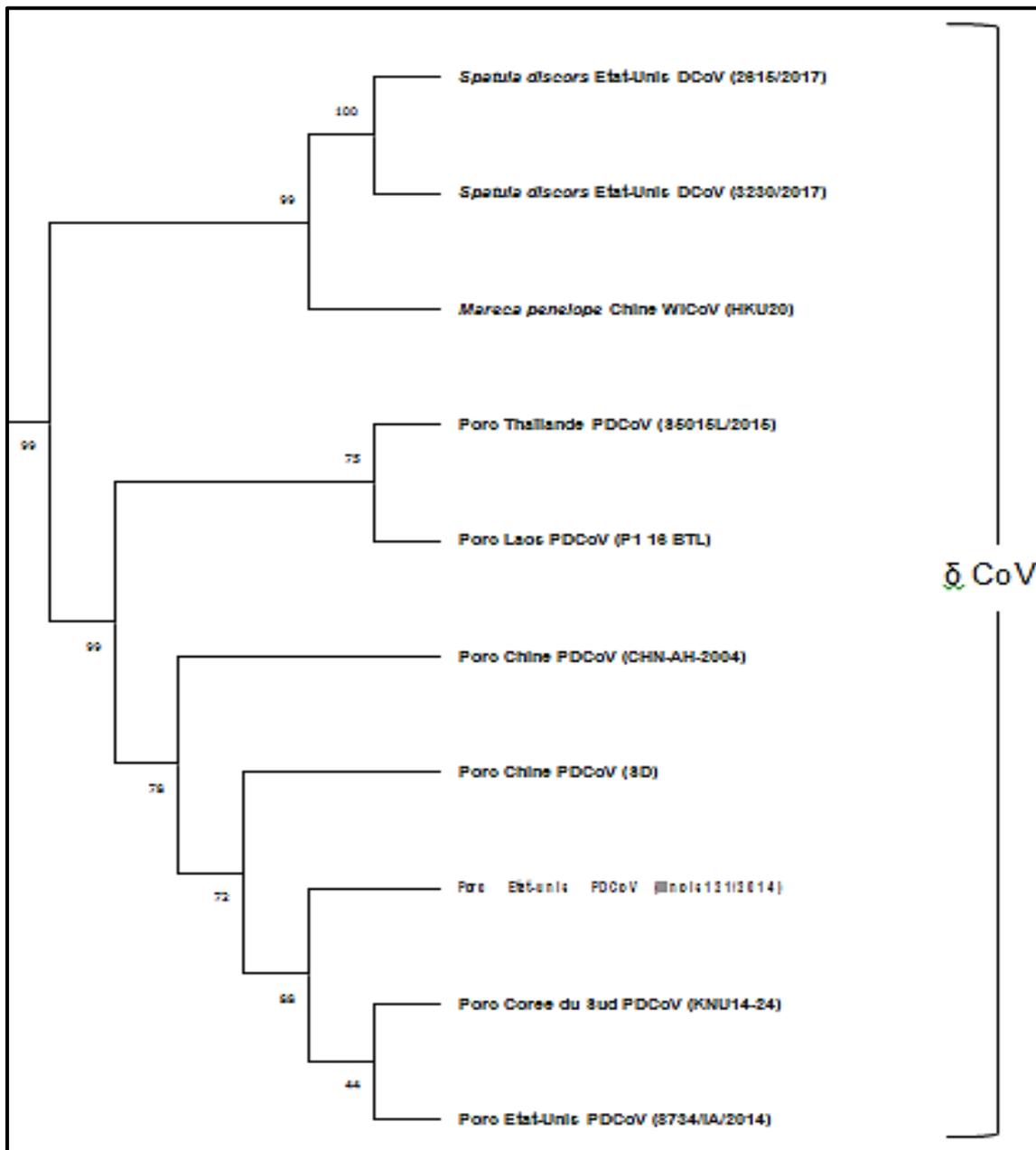


Figure 48 : Fragment de l'arbre du gène N comportant les séquences du groupe δ .

A leur tour les séquences du gène N se divisent en quatre groupes qui sont :

RESULTATS

Betacoronavirus (β CoV) 13 différentes souches de Coronavirus classées en :

- Sous-groupe A : Il comprend 9 séquences nucléotidiques de :
 - 3 séquences de Coronavirus de bovin (BCoV) distribuées en, une de Canada, une de l'Italie et une de France qui infecte les vaches.
 - 2 séquences de Canine respiratoire Coronavirus (CRCoV) venant de, Chine et Corée du Sud.
 - 4 séquences de Coronavirus humain OC43 (HCoV-OC43) distribuées en égalité entre la Belgique et la Chine.
- Sous-groupe B : Il comprend 4 séquences nucléotidiques du Coronavirus du syndrome respiratoire aigu sévère (SRAS-CoV-2) de la Chine qui touche l'humain aussi bien que les chauves-souris *Rhinolophus affinis*, *Rhinolophus pusillus* et *Rhinolophus sinicus*.

Pourcentages de similarités entre les différentes séquences de 13 souches du groupe Betacoronavirus (Beta CoV) avec une moyenne d'identité de 74,85% sont représentés dans le **tableau VIII**.

Tableau VIII : Pourcentages d'identité de différentes séquences de souches du groupe Betacoronavirus (Beta CoV) qui présentent le gène N.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
SARS 2-CoV (<i>bat-SL-CoVZXC21</i>) (1)	100	90,39	91,19	91,19	47,49	47,48	47,09	47,09	47,09	47,4	47,4	47,01	47,09
SARS 2-CoV (<i>BtRs-HuB2013</i>) (2)	90,39	100	88,8	88,17	48,74	48,23	47,92	47,92	47,92	48,47	48,47	48,08	48,08
SARS 2-CoV (<i>CHN/Yunnan-01/2020</i>) (3)	91,19	88,8	100	96,9	47,49	46,93	46,22	46,22	46,22	46,77	46,77	46,54	46,46
SARS 2-CoV (<i>RaTG13</i>) (4)	91,19	88,17	96,9	100	47,91	47,4	46,85	46,85	46,85	47,24	47,24	47,17	46,93
BCoV (<i>bcovizsm</i>) (5)	47,49	48,74	47,49	47,91	100	99,33	96,22	96,22	95,56	98,67	98,67	97,78	96,67
BCoV (<i>EPI/Caen/2013/08</i>) (6)	47,48	48,23	46,93	47,4	99,33	100	96,14	96,07	95,84	98,29	98,66	98,07	96,59
HCoV-OC43 (<i>HK04-01</i>) (7)	47,09	47,92	46,22	46,85	96,22	96,14	100	99,93	99,7	96,21	96,59	96,29	99,11
HCoV-OC43 (<i>HK04-02</i>) (8)	47,09	47,92	46,22	46,85	96,22	96,07	99,93	100	99,78	96,14	96,51	96,21	99,03
HCoV-OC43 (<i>BE04</i>) (9)	47,09	47,92	46,22	46,85	95,56	95,84	99,7	99,78	100	95,92	96,29	95,99	98,81
CRCoV (<i>K37</i>) (10)	47,4	48,47	46,77	47,24	98,67	98,29	96,21	96,14	95,92	100	99,33	98	96,81
BCoV (<i>Quebec</i>) (11)	47,4	48,47	46,77	47,24	98,67	98,66	96,59	96,51	96,29	99,33	100	98,37	97,18
CRCoV (<i>BJ202</i>) (12)	47,01	48,08	46,54	47,17	97,78	98,07	96,29	96,21	95,99	98	98,37	100	96,73
HCoV-OC43 (<i>HCoV-OC43</i>) (13)	47,09	48,08	46,46	46,93	96,67	96,59	99,11	99,03	98,81	96,81	97,18	96,73	100

Identité globale : 74,85 %

Alphacoronavirus (α CoV) : 12 différentes souches de Coronavirus classées en :

- **Sous-groupe A** : Il comprend les séquences nucléotidiques de :
 - Une séquence de Coronavirus Canine (CCoV) de l'Allemagne qui infecte le chien *Canis lupus familiaris*.

RESULTATS

- 4 séquences de Coronavirus Félin (**FCoV**) : 3 de Thaïlande et une des Etats-Unis.

➤ **Sous-groupe B** : Il comprend 7 séquences nucléotidiques du virus de la diarrhée épidémique porc (**PEDV**) 2 de Chine, 2 du Japon, une des Etats-Unis et 2 des Philippines qui infecte *Sus scrofa*.

Pourcentages de similarités entre les différentes séquences de 12 souches du groupe Alphacoronavirus (Alpha CoV) avec une identité globale de 67,67% sont représentées dans le **tableau IX**.

Tableau IX : Pourcentages d'identité de différentes séquences de souches du groupe Alphacoronavirus (Alpha CoV) qui présentent le gène N.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PEDV (AH2012) (1)	100	98,49	99,47	97,02	98,61	43,77	44,28	43,88	45,58	45,67	99,4	99,4
PEDV (CH/ZMDZY/11) (2)	98,49	100	98,87	97,02	98,61	44,08	44,49	44,5	45,94	46,03	98,79	98,79
PEDV (IA1) (3)	99,47	98,87	100	97,66	99,25	43,98	44,49	44,09	45,67	45,76	99,92	99,92
PEDV (K3) (4)	97,02	97,02	97,66	100	97,59	44,35	45,07	44,37	44,83	44,94	97,55	97,55
PEDV (K4) (5)	98,61	98,61	99,25	97,59	100	44,7	45,3	44,6	44,95	45,06	99,15	99,15
FCoV (KUTTI/2017) (6)	43,77	44,08	43,98	44,35	44,7	100	92,5	91,89	90,36	90,25	44,08	43,87
FCoV (KU30/2019) (7)	44,28	44,49	44,49	45,07	45,3	92,5	100	91,03	89,93	89,93	44,39	44,39
FCoV (KU62/2020) (8)	43,88	44,5	44,09	44,37	44,6	91,89	91,03	100	94,35	94,24	44,19	44,19
FCoV (WSU 79-1146) (9)	45,58	45,94	45,67	44,83	44,95	90,36	89,93	94,35	100	99,91	45,76	45,76
CCoV (171) (10)	45,67	46,03	45,76	44,94	45,06	90,25	89,93	94,24	99,91	100	45,85	45,85
PEDV (14JM-168) (11)	99,4	98,79	99,92	97,55	99,15	44,08	44,39	44,19	45,76	45,85	100	99,85
PEDV (14JM-179) (12)	99,4	98,79	99,92	97,55	99,15	43,87	44,39	44,19	45,76	45,85	99,85	100

Identité globale : 67,67 %

Gammacoronavirus (γ CoV) : Comprend une séquence du virus de la bronchite infectieuse aviaire (**IBV**) qui infecte le poulet de la Suède.

Deltacoronavirus (δ CoV) : Représente 10 séquence de :

- 7 séquences du Deltacoronavirus du porc (**PDCoV**) : une de Thaïlande, une du Laos, 2 de Chine, 2 des Etats-Unis et 1 de Corée du Sud
- 2 séquences du Coronavirus du canard (**DCoV**) des Etats-Unis qui infecte *Spatula discors*.
- Une séquence du Coronavirus du canard d'Amérique (**HKU-20**) qui infecte le canard de Chine *Mareca penelope*.

Pourcentages de similarités entre les différentes séquences de 10 souches du groupe Deltacoronavirus (δ CoV) avec une identité moyenne de 75,71% est représenté dans le **tableau X**.

RESULTATS

Tableau X : Pourcentages d'identité de différentes séquences de souches du groupe Deltacoronavirus (δ CoV) présentant le gène N.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>PDCoV(Illinois121/2014)</i> (1)	100	97,86	96,99	98,15	98,83	99,42	99,51	50,41	49,86	55,24
<i>PDCoV (P1_16_BTL)</i> (2)	97,86	100	98,54	98,15	98,25	98,06	98,15	50,14	49,59	54,95
<i>PDCoV (S5015L/2015)</i> (3)	96,99	98,54	100	97,67	97,38	97,18	97,08	50,41	49,86	54,85
<i>PDCoV (CHN-AH-2004)</i> (4)	98,15	98,15	97,67	100	98,74	98,35	98,45	49,32	48,77	54,95
<i>PDCoV (SD)</i> (5)	98,83	98,25	97,38	98,74	100	99,03	98,93	50,14	49,59	54,95
<i>PDCoV(8734/IA/2014)</i> (6)	99,42	98,06	97,18	98,35	99,03	100	99,51	50,68	50,14	55,24
<i>PDCoV (KNU14-24)</i> (7)	99,51	98,15	97,08	98,45	98,93	99,51	100	50,41	49,86	55,24
<i>DCoV (3230/2017)</i> (8)	50,41	50,14	50,41	49,32	50,14	50,68	50,41	100	99,49	84,62
<i>DCoV (2615/2017)</i> (9)	49,86	49,59	49,86	48,77	49,59	50,14	49,86	99,49	100	84,1
<i>WiCoV (HKU20-9243)</i> (10)	55,24	54,95	54,85	54,95	54,95	55,24	55,24	84,62	84,1	100

Identité globale : 75,71 %

III.3. Résultats de l'arbre du gène RdRp

Les séquences d'ADN de différentes souches virales du gène RdRp de Coronavirus extraites sont mentionnées dans le **tableau XI** en fonction des espèces animales affectées à travers plusieurs pays dans le monde.

Tableau XI : Les séquences d'ADN de Coronavirus utilisées pour la présentation phylogénétique du gène RdRp.

Animal	Espèce	Pays	Gène	virus
Chauve-souris	<i>Chaerephon pusillus</i>	Mayotte	RDRP	Bat CoV
	<i>Hipposideros caffer</i>	Mozambique	RDRP	
	<i>Triaenops Aafter</i>	Mozambique	RDRP	
	<i>Miniopterus mossambicus</i>	Mozambique	RDRP	
	<i>Rhinolophus rhodesiae</i>	Mozambique	RDRP	
	<i>Rhinolophus lobatus</i>	Mozambique	RDRP	
	<i>Rhinolophus sp</i>	Mozambique	RDRP	
Canard	<i>Anas sp</i>	Pologne	RDRP	ACoV
	<i>Anas platyrhynchos</i>	Pologne	RDRP	
	<i>Anser fabalis</i>	Pologne	RDRP	
Furet	Furet domestique de Lima	Pérou	RDRP	FRSCV
Bovin	Bétail	Slovénie	RDRP	BCoV
Humain	<i>Homo sapiens</i>	Slovénie	RDRP	HCoV-OC43,

RESULTATS

				HKU1, FRSCV
Rongeur	<i>Mus caroli</i>	Chine	RDRP	RtMc-CoV-1, RtMc-CoV-2, RtRf-CoV-2, RtNn-CoV, RtAa-CoV
	<i>Rattus tanezumi</i>	Chine	RDRP	
	<i>Rattus flavipectus</i>	Chine	RDRP	
	<i>Niviventer niviventer</i>	Chine	RDRP	
	<i>Apodemus agrarius</i>	Chine	RDRP	
Porc	Porc	Italie	RDRP	PHEV

L'arbre phylogénétique (**Figure 49**) étudié par MEGA 11 a été réalisé sur la base du gène RdRp, de 35 séquences de différentes souches virales du Coronavirus qui ont touchés plusieurs espèces animales et l'espèce humaine avec une taille qui varie entre 425 Pb et 440 Pb.

RESULTATS

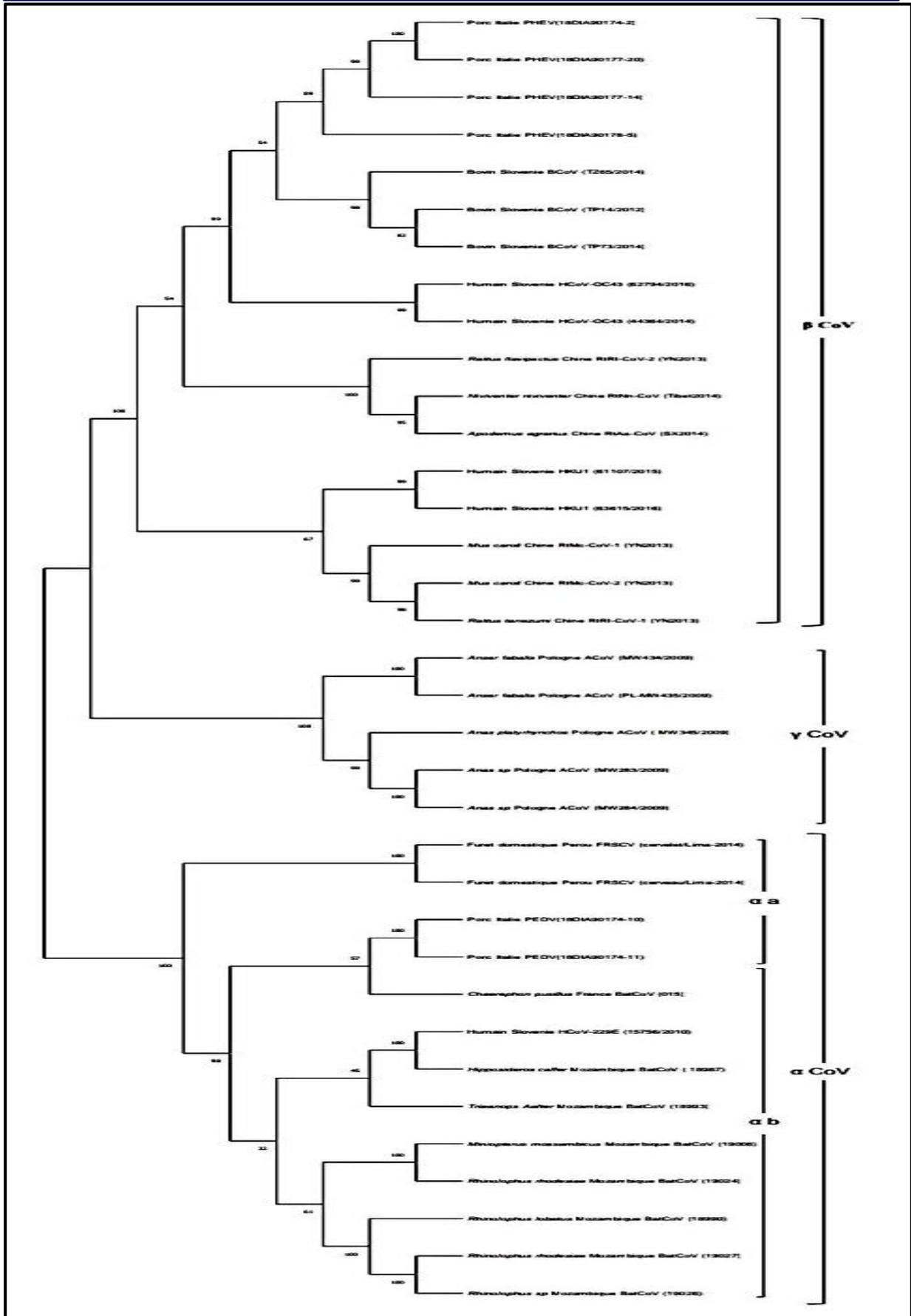


Figure 49 : L'arbre phylogénétique du gène RdRp.

RESULTATS

En raison du grand nombre de séquences, ainsi qu'afin de bien visualiser ce grand arbre nous proposons les figures suivantes qui regroupent les séquences de chaque groupe viral.

Dans la première position le groupe β qui comprend seulement les souches regroupées dans le sous-groupe βb qui se représentent dans la **figure 50**.

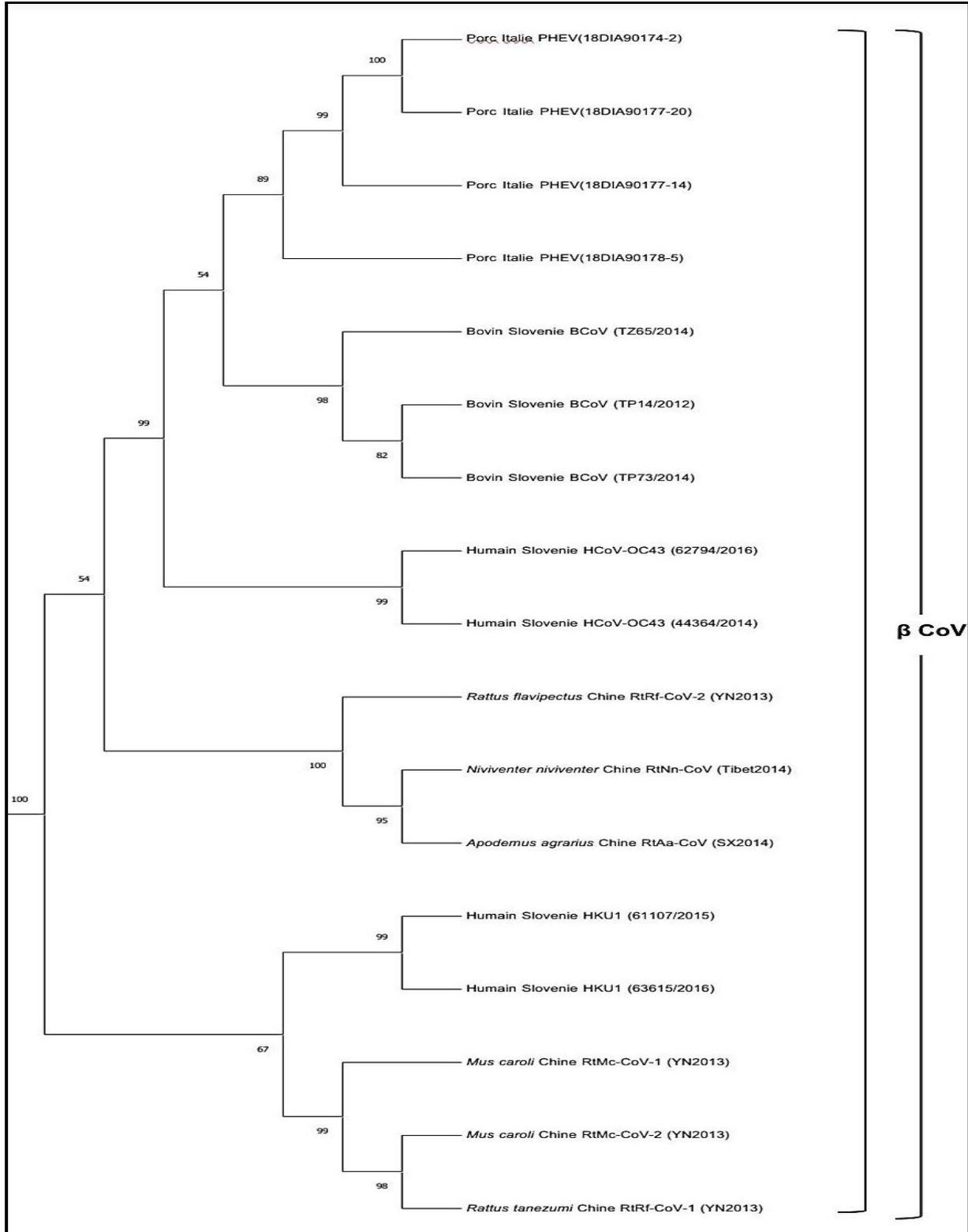


Figure 50 : Fragment de l'arbre du gène RdRp comportant les séquences du groupe β .

RESULTATS

Par la suite la **figure 51** présente le fragment de l'arbre du gène N qui comporte les séquences du groupe γ .

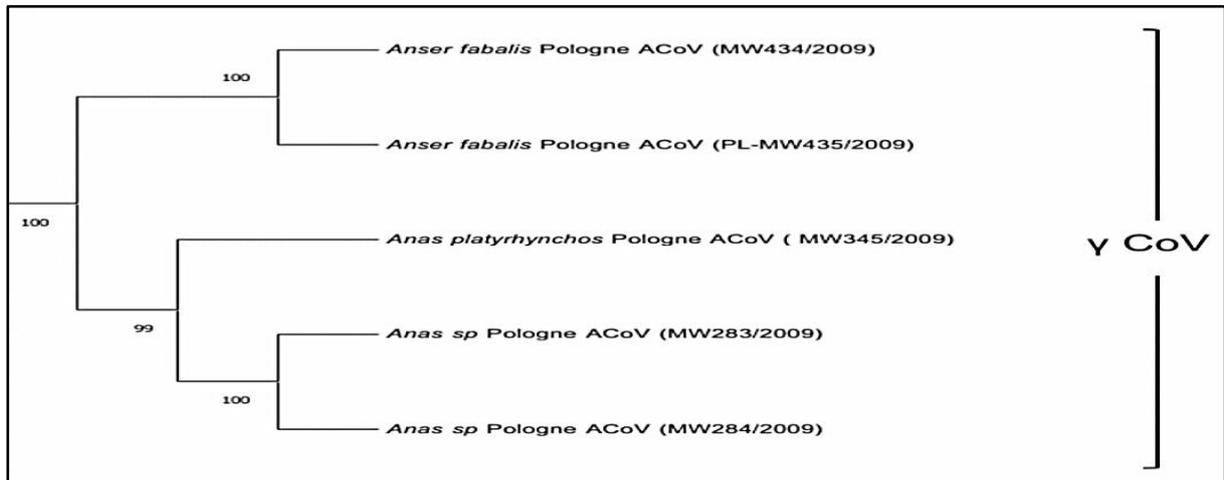


Figure 51 : Fragment de l'arbre du gène RdRp incluant les séquences du groupe γ .

Par la suite le groupe α qui se divise dans notre arbre en deux sous-groupes : le sous-groupe αa et le sous-groupe αb les deux se représentent successivement dans la **figure 52** et la **figure 53**.



Figure 52 : Fragment de l'arbre du gène RdRp incluant les séquences du sous-groupe αa .

RESULTATS

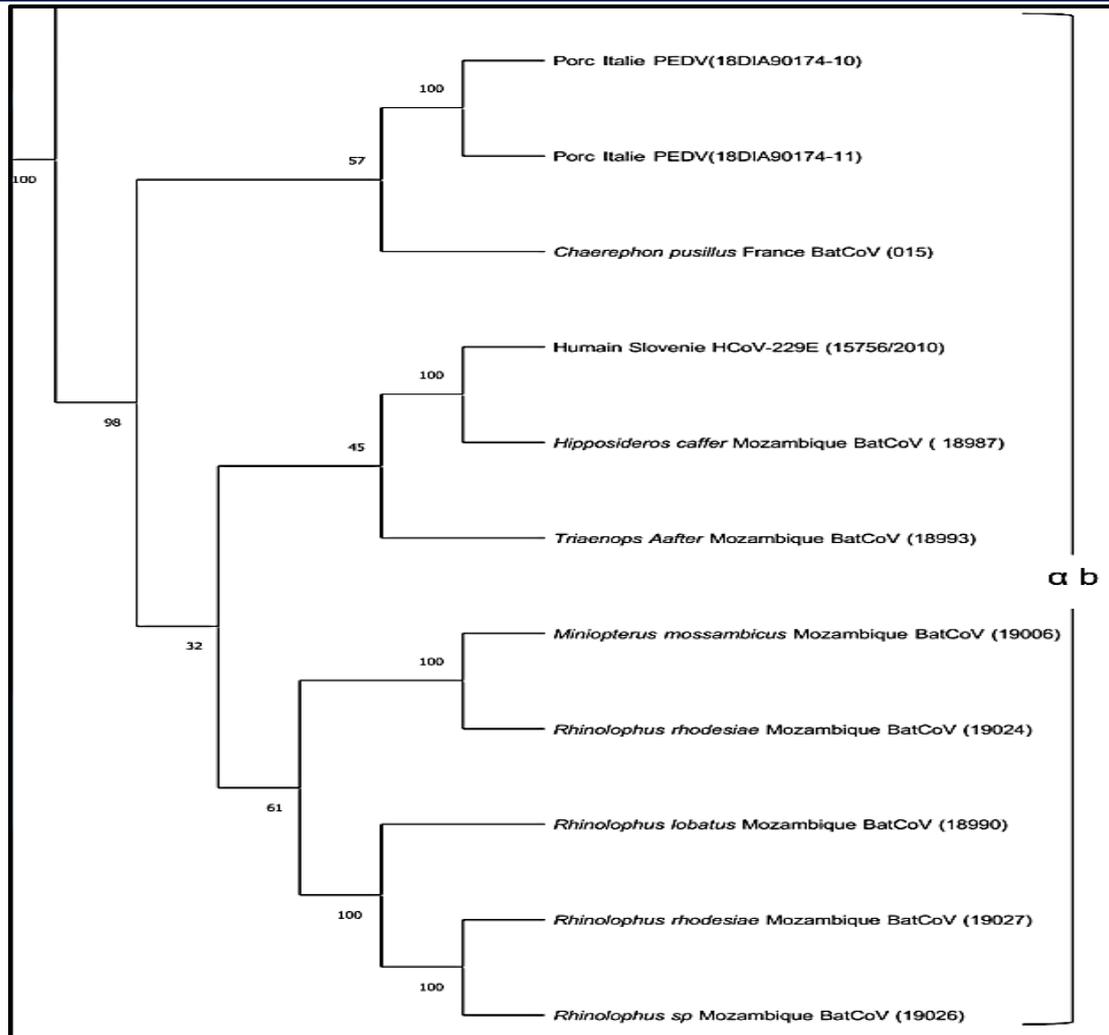


Figure 53 : Fragment de l'arbre du gène RdRp incluant les séquences du sous-groupe $\alpha\beta$.

L'arbre de gène RdRp se divise en plusieurs groupe suivant :

Betacoronavirus (β CoV) : 17 différentes souches de Coronavirus classées dans le sous-groupe B :

- 4 séquences nucléotidiques de virus de l'encéphalomyélite hémagglutinante porc (PHEV) d'Italie.
- 3 séquences de Coronavirus du bovin (BCoV) qui touchent le bétail de Slovénie.
- 2 séquences de Coronavirus humain OC43 (HCoV-OC43) de Slovénie.
- 2 séquences de Coronavirus humain HKU1 (HCoV-HKU1) de Slovénie qui touchent l'humain.

***Des souches non classées dans la GenBank, qui prennent l'emplacement dans l'arbre avec les souches de ce groupe.**

- Une séquence de Coronavirus de rongeur (RtRf-CoV-2) de Chine qui infecte le rongeur *Rattus flavipectus*.

RESULTATS

- Une séquence de Coronavirus de rongeur (RtNn-CoV) de Chine qui infecte le rongeur *Niviventer niviventer*.
- Une séquence de Coronavirus de rongeur (RtAa-CoV) de Chine qui infecte le rongeur *Apodemus agrarius*.
 - Ces souches se regroupent pour une origine commune avec les premières souches déjà classées dans ce groupe viral.
- Une séquence de Coronavirus de rongeur (RtMc-CoV-1) de Chine qui infecte le rongeur *Mus caroli*.
- Une séquence de Coronavirus de rongeur (RtMc-CoV-2) de Chine qui infecte le rongeur *Mus caroli*.
- Une séquence de Coronavirus de rongeur (RtRf-CoV-1) de Chine qui infecte le rongeur *Rattus tanezumi*.
 - Ces trois souches se positionnent en même clade avec les souches du virus HKU1 (qui est déjà classée dans ce groupe viral).

Pourcentages de similarités entre les différentes séquences de 17 souches du groupe Betacoronavirus (Beta CoV) avec une identité globale de 88,85% est représenté dans le **tableau XII**.

RESULTATS

Tableau XII : Pourcentages d'identité de différentes séquences de souches du groupe Betacoronavirus (Beta CoV) qui présentent le gène RdRp.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
BCoV (TP14/2012) (1)	100	96,67	96,67	98,97	99,74	82,05	82,05	83,72	83,16	83,93	84,99	84,22	83,93	96,68	96,43	96,68	96,68
HCoV-OC43 (62794/2016) (2)	96,67	100	100	96,67	96,92	82,05	82,05	83,72	83,16	84,44	85,24	84,73	84,69	95,41	95,66	95,41	95,41
HCoV-OC43 (44364/2014) (3)	96,67	100	100	96,67	96,92	82,05	82,05	83,72	83,16	84,44	85,24	84,73	84,69	95,41	95,66	95,41	95,41
BCoV (TZ65/2014) (4)	98,97	96,67	96,67	100	99,23	82,05	82,05	83,21	83,16	83,93	84,73	84,22	83,93	96,17	96,43	96,17	96,17
BCoV (TP73/2014) (5)	99,74	96,92	96,92	99,23	100	82,05	82,05	83,97	83,42	84,18	84,99	84,48	84,18	96,43	96,68	96,43	96,43
HKU1 (61107/2015) (6)	82,05	82,05	82,05	82,05	82,05	100	100	84,73	84,69	80,87	84,48	79,39	79,08	80,36	81,38	80,36	80,36
HKU1 (63615/2016) (7)	82,05	82,05	82,05	82,05	82,05	100	100	84,73	84,69	80,87	84,48	79,39	79,08	80,36	81,38	80,36	80,36
RtMc-CoV-2 (YN2013) (8)	83,72	83,72	83,72	83,21	83,97	84,73	84,73	100	95,69	83,9	93,41	83,02	82,72	83,21	83,21	83,21	83,21
RtRf-CoV-1 (YN2013) (9)	83,16	83,16	83,16	83,16	83,42	84,69	84,69	95,69	100	83,64	93,42	82,74	82,45	83,17	83,17	83,17	83,17
RtRf-CoV-2 (YN2013) (10)	83,93	84,44	84,44	83,93	84,18	80,87	80,87	83,9	83,64	100	85,26	95,04	94,92	82,91	83,67	82,91	82,91
RtMc-CoV-1 (YN2013) (11)	84,99	85,24	85,24	84,73	84,99	84,48	84,48	93,41	93,42	85,26	100	84,43	84,56	84,96	85,46	84,96	84,96
RtNn-CoV (Tibet2014) (12)	84,22	84,73	84,73	84,22	84,48	79,39	79,39	83,02	82,74	95,04	84,43	100	97,16	83,71	84,46	83,71	83,71
RtAa-CoV (SX2014) (13)	83,93	84,69	84,69	83,93	84,18	79,08	79,08	82,72	82,45	94,92	84,56	97,16	100	83,67	84,42	83,67	83,67
PHEV(18DIA90177-14) (14)	96,68	95,41	95,41	96,17	96,43	80,36	80,36	83,21	83,17	82,91	84,96	83,71	83,67	100	98,23	100	100
PHEV(18DIA90178-5) (15)	96,43	95,66	95,66	96,43	96,68	81,38	81,38	83,21	83,17	83,67	85,46	84,46	84,42	98,23	100	98,23	98,23
PHEV(18DIA90174-2) (16)	96,68	95,41	95,41	96,17	96,43	80,36	80,36	83,21	83,17	82,91	84,96	83,71	83,67	100	98,23	100	100
PHEV(18DIA90177-20) (17)	96,68	95,41	95,41	96,17	96,43	80,36	80,36	83,21	83,17	82,91	84,96	83,71	83,67	100	98,23	100	100

Identité globale : 88,35 %

Gammacoronavirus (γ CoV) : Comprend 5 séquences de Coronavirus de canard (ACoV) de Pologne qui touche le canard *Anser fabalis*, *Anas platyrhynchos* et *Anas sp.*

Pourcentages de similarités entre les différentes séquences de 5 souches du groupe Gammacoronavirus (γ CoV) avec une identité moyenne de 92,04% est représenté dans le **tableau XIII.**

Tableau XIII : Pourcentages d'identité de différentes séquences de souches du groupe Gammacoronavirus (γ CoV) qui présentent le gène RdRp.

	(1)	(2)	(3)	(4)	(5)
ACoV (PL-MW283/2009) (1)	100	99,56	89,05	89,38	92,69
ACoV (PL-MW284/2009) (2)	99,56	100	89,38	89,16	92,71
ACoV (PL-MW434/2009) (3)	89,05	89,38	100	99,71	88,78
ACoV (PL-MW435/2009) (4)	89,38	89,16	99,71	100	88,78
ACoV (MW345/2009) (5)	92,69	92,71	88,78	88,78	100

Identité globale : 92,04 %

RESULTATS

Alphacoronavirus (Alpha CoV) : contenant 13 séquences de souches différentes classées en 2 sous-groupes :

➤ **Sous-groupe A :** Il comprend 4 séquences nucléotidiques de :

- 2 séquences de Coronavirus systémique du furet (**FRSCV**) de Pérou qui infecte le furet *domestique*.

➤ **Sous-groupe B :** Il comprend 9 séquences nucléotidiques de :

- 2 du virus de l'encéphalomyélite hémagglutinante porc (**PEDV**) d'Italie.
- Une séquence de Coronavirus humain 229E (**HCoV-229E**) qui infecte l'humain.
- 8 séquences de Coronavirus de chauve-souris (**Bat CoV**) de France qui touche la chauve-souris *Chaerephon pusillus* et de Mozambique qui infecte aussi les chauves-souris *Hipposideros caffer*, *Triaenops After*, *Miniopterus mossambicus*, *Rhinolophus rhodesiae*, *Rhinolophus lobatus*, *Rhinolophus rhodesiae* et *Rhinolophus sp*

Pourcentages de similarités entre les différentes séquences de 12 souches du groupe Alphacoronavirus (Alpha CoV) avec une identité moyenne de 77,17% sont représentés dans le **tableau XIV**.

Tableau XIV : Pourcentages d'identité de différentes séquences de souches du groupe Alphacoronavirus (Alpha CoV) qui présentent le gène RdRp.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<i>BatCoV (18987) (1)</i>	100	76,7	76,47	73,3	73,3	73,3	67,59	67,34	92,09	92,09	69,1	73,76
<i>BatCoV (19006) (2)</i>	76,7	100	97,29	77,15	77,15	77,15	67,34	67,09	73,72	73,72	69,83	74,21
<i>BatCoV (19024) (3)</i>	76,47	97,29	100	77,83	77,83	77,83	67,34	67,09	73,98	73,98	70,07	75,11
<i>BatCoV (18990) (4)</i>	73,3	77,15	77,83	100	100	100	65,32	65,06	71,43	71,43	71,78	75,79
<i>BatCoV (19027) (5)</i>	73,3	77,15	77,83	100	100	100	65,32	65,06	71,43	71,43	71,78	75,79
<i>BatCoV (19026) (6)</i>	73,3	77,15	77,83	100	100	100	65,32	65,06	71,43	71,43	71,78	75,79
<i>FRSCV (cervelet/Lima-2014) (7)</i>	67,59	67,34	67,34	65,32	65,32	65,32	100	99,75	69,57	69,57	66,08	65,82
<i>FRSCV (cerveau/Lima-2014) (8)</i>	67,34	67,09	67,09	65,06	65,06	65,06	99,75	100	69,31	69,31	65,82	65,57
<i>HCoV-229E (15756/2010)(2) (9)</i>	92,09	73,72	73,98	71,43	71,43	71,43	69,57	69,31	100	100	68,62	69,9
<i>HCoV-229E (15756/2010) (10)</i>	92,09	73,72	73,98	71,43	71,43	71,43	69,57	69,31	100	100	68,62	69,9
<i>PEDV(18DIA90174-10) (11)</i>	69,1	69,83	70,07	71,78	71,78	71,78	66,08	65,82	68,62	68,62	100	74,94
<i>BatCoV (015) (12)</i>	73,76	74,21	75,11	75,79	75,79	75,79	65,82	65,57	69,9	69,9	74,94	100

Identité globale : 77,17 %

DISCUSSION

IV. Discussion

Dans cette étude qui traite par des outils bioinformatique, dont l'étude phylogénique les arbres phylogénétiques basée sur les trois gènes : gène de pointe S (gène structurale), gène N (gène de nucléocapside) et gène RdRp (gène non structurale).

nous avons comparé les différentes séquences de différentes souches virales de Coronavirus (CoV) qui sont présent chez une grande variété d'animaux, chez lesquels ils peuvent provoquer des maladies respiratoires, entériques, hépatiques et neurologiques de gravité variable. Leur tendance à la recombinaison et les taux de mutation intrinsèquement élevés du virus à ARN peuvent leur permettre de s'adapter à de nouveaux hôtes et niches écologiques, cette adaptation est responsable de l'apparition de 4 genres selon le "Study Groupe" de la "International Committee for Taxonomy» ces genres sont Alphacoronavirus, Betacoronavirus, Gammacoronavirus, Deltacoronavirus (**Woo et al., 2012**).

IV.1. Comparaison des souches de l'arbre du gène de pointe S (gène structurale)

Cet arbre phylogénétique se divise en quatre groupes : Alphacoronavirus, Betacoronavirus, Gammacoronavirus, Deltacoronavirus ce qui se confirme par la classification de **Snijders (2008)**.

IV.1.1. Comparaison des souches de l'arbre du groupe Alphacoronavirus du gène de pointe S

Selon **Cui et al. (2019)** Alphacoronavirus n'infecte que les mammifères ; dans notre travail nous avons trouvé des séquences de Coronavirus qui infectent le porc, canin, félin, furet et humain.

Dans les étapes suivantes nous allons présenter ces deux sous-groupes ainsi que les virus et les souches virales qui lui sont propres.

a. Comparaison des souches du sous-groupe $\alpha\alpha$

Ce groupe comporte les séquences nucléotidiques de : Virus de la gastro-entérite transmissible porcine (TGEV), Coronavirus canin (CCoV), le virus de la péritonite infectieuse féline (FCoV), Coronavirus respiratoire porcine (PRCV). Ces virus sont classés dans le sous-groupe $\alpha\alpha$, ceci est confirmé par les recherches de **Snijders (2008)**.

Un ancêtre commun entre ces virus se déduit par le lien entre ces derniers ; les recherches de **Kathy et al. (2002)** montrent que le virus TGEV est antigéniquement lié au virus de la péritonite infectieuse féline (FCoV) et au Coronavirus canin (CCoV) où ils ont déduit que les

DISCUSSION

porcs peuvent être infectés expérimentalement par des Coronavirus de chat et de chien, et de même, les chiens et les chats peuvent être infectés expérimentalement par le TGEV.

- **Virus de la gastro-entérite transmissible porcine (TGEV)**

Ce virus infectant les porcs des États-Unis où il se présente dans l'arbre par les souches : "Miller M60", "Perdue P115", "Purdue virulente". Ils se regroupent en même branche suite à des mutations nucléotidiques produites, ces résultats sont confirmés par les travaux de **Zhang et al. (2007)** qui ont déduit que "Miller M6" et "Purdue virulente" partagent jusqu'à 98,9% l'identité de séquence génomique, sachant que "Miller M60" présente une suppression de 3 nt dans 2385 pb à 2387 pb, résultant en une protéine de pointe de 1448 aa de longueur ainsi que "Perdue P115" qui présente une suppression de 6 nt conduisant à une protéine de pointe à 2 aa plus courte que la souche "Purdue virulente" (1447 aa contre 1449 aa).

- **Coronavirus respiratoire porcin (PRCV)**

Ce virus est présent dans l'arbre par la souche "ISU-1" qui touche les porcs américains, cette dernière se regroupe en même branche avec les souches de TGEV précédemment citées où ils présentent un ancêtre commun, résultats s'expliqués par les recherches de **Zhang et al. (2007)** qui ont conclu que le PRCV et le TGEV partagent une identité nucléotidique élevée allant de 96,2 % à 99,9 % qui se traduit par une délétion dans l'extrémité 5' du gène S du PRCV, ce dernier conduit à un manque ou à une protéine tronquée mais exprimée.

- **Coronavirus canin (CCoV)**

Dans l'arbre ce virus présente les souches : "NTU336/F/2008" de Taïwan, "HCM27/2014" qui infecte le chien *Canis lupus familiaris* et "GZ43/2003" de la Chine qui infecte le chien viverrin *Nyctereutes procyonoides* et le blaireau furet chinois *Melogale moschata* "DM95/2003", ces souches sont classées dans la même branche, ce qui est justifié par les recherches de **Nguyen et al. (2016)** qui montrent que : La souche "HCM27/2014" est similaire à 95,9 % aux souches "GZ43/2003" et "DM95/2003" et à 95,2 % similaire à la souche "NTU336/F/2008".

- **Virus de la péritonite infectieuse féline (FCoV)**

Les souches de FCoV qui se présentent dans l'arbre sont : "FIPV-C3663", "UU3", "UU21" et "UU16" leur groupement est dans la même branche qui donne la probabilité d'une origine commune ce qui est justifié par leurs positions dans l'arbre du gène S ainsi que par les travaux de **Chang et al. (2012)** et **Terada et al (2012)** qui ont conclu l'existence de grandes délétions au sein de la sous-unité S1 correspondant à la région N-terminale qui ont été trouvées dans des variantes du "FIPV" par rapport aux souches "UU16", "UU21" et "C3663".

DISCUSSION

b. Comparaison des souches du sous-groupe *ab*

Ce groupe comporte les séquences nucléotidiques des deux Coronavirus humain 229E (HCoV-229E) et NL63 (HCoV-NL63). Ainsi que dans notre arbre, NL63 (HCoV-NL63) présente : trois séquences de Chauve-souris *Triaenops after* de Kenya : "BtkNL63-9b", "BtkNL63-9a" et "BtkNL63-15" ; cependant les Coronavirus humain 229E (HCoV-229E) se présentent par les souches BtCoV- 229E, BtKY229E-8, BtKY229E-1, CRA CA08-1/2008 et HCoV-229E qui se regroupent dans la même branche ou il est justifié par les recherches de **Tao et al. (2017)** qui montrent une identité génétique entre les séquences de protéines S du premier qui étaient de 53 %. En outre pour les souches "NL63" qui infectent l'humain, les travaux de **Hoek et al. (2004)** ont déduit que les Coronavirus humains NL63 et les Coronavirus humains 229E ont une origine recombinante zoonotique à partir de leur ancêtre commun le plus récent (estimé à environ 1000 ans), avec la présence d'un grand insert de 537 nucléotides dans la partie 5' du gène S de HCoV-NL63, par rapport à celui de HCoV-229 E. Ce qui justifie leur position dans la même branche dans l'arbre.

IV.1.2. Comparaison des souches de l'arbre du groupe Deltacoronavirus du gène de pointe S

Ce groupe classe le virus Deltacoronavirus du porc (PDCoV) comme il est indiqué par les recherches de Snijders (2008) et **Hepojoki et al. (2017)**. Dans notre arbre phylogénique de ce gène, ce virus présente les souches : "TT1115", "CH-HA1-2017", "CH-HA2-2017", "CH-HA3-2017", "CH-WH-2017", "CHN-HN-2017", "CH/Sichuan/S27/2012", "NH" et "SHJS/SL/2016" qui représentent tous la même origine justifiée par les résultats de **Adthakorm et al. (2016)**, la souche "TT1115" était fortement homologue aux souches de PDCoV de Chine (les souches : "CH-WH-2017", "CHN-HN-2017", "CH/Sichuan/S27/2012", "NH" et "SHJS/SL/2016"), avec des similitudes de nucléotides et d'acides aminés de 97,2% à 97,8 % et de 93,0% à 94,0 %, respectivement. De plus, l'analyse génétique basée sur le gène S a démontré qu'ils sont étroitement liés avec des similitudes de 95,6% à 96,7 % et de 95,9 % à 98,1 % au niveau des nucléotides et des acides aminés, respectivement ; qui se résultent de vingt-quatre substitutions au niveau des acides aminés. De plus, "TT1115" possède une suppression d'un acide aminé en position 51, similaire aux souches Chinoises. Concluant que les résultats de cette étude suggèrent que "TT1115" est une nouvelle souche, étroitement liée aux souches de PDCoV de Chine.

DISCUSSION

IV.1.3. Comparaison des souches de l'arbre du groupe Gammacoronavirus du gène de pointe S

Dans notre arbre, ce groupe comprend les deux virus de la bronchite infectieuse (IBV) qui infecte l'aviaire principalement l'espèce de *Gallus gallus* ou il présente les souches : "Bj", "CK/CH/XDC-2/2013", "KM91", "ck/Pologne/255/1997" qui présente une origine commune comme les ont montrés les résultats de **Leghari et al. (2016)** ou le gène S de "CK/CH/XDC-2/2013" provenait de la recombinaison des souches parentales majeures "BJ" et parentales mineures "KM91", Cependant **Domansk-Blicharz et al. (2020)** ont démontré que la souche "255/1997" a été recombinaisonnée avec une souche vaccinale 4/91 (la souche "KM91").

IV.1.4. Comparaison des souches de l'arbre du groupe Betacoronavirus du gène de pointe S

Dans notre arbre phylogénique, ce groupe comprend les séquences nucléotidiques des virus : Les Coronavirus du syndrome respiratoire aigu sévère (SRAS-CoV), le Coronavirus équin (ECoV), Coronavirus bovins (BCoV) et le Coronavirus de girafe (GiCoV). Don tous ces virus sont classés d'après les recherches de **Mavrodiev et al. (2020)** en deux sous-groupes : "sous-groupe a" et "sous-groupe b".

a. Comparaison des souches du sous-groupe β b

Le sous-groupe b dans cette arbres, comprend le Coronavirus du syndrome respiratoire aigu sévère (SRAS-CoV-2) et le Coronavirus du pangolin (PCoV) ce qui est justifié par le classement des groupes adoptés par **Jasper et al. (2013)**.

Coronavirus du pangolin (PCoV) se présente par la souche "GX-P2V" ; cependant le virus de Coronavirus du syndrome respiratoire aigu sévère SARS-CoV représente à leurs tours les souches : "CHN/Yunnan-01/2020", "BtRs-HuB2013" et "bat-SL-CoVZXC21". Dans ces deux virus sont regroupées en même branche indiquant la présence d'un lien entre eux ce qui est interprété par les travaux de **Lohrasbi-Nejad (2021)** qui a déduit qu'un événement de recombinaison a été observé lors de la comparaison de la séquence du génome des souches virales de SRAS-CoV-2 humain, la souche "CHN/Yunnan-01/2020" à ceux de la chauve-souris "BtRs-HuB2013" et " bat-SL-CoVZXC21" ainsi que le souche "GX-P2V" de Coronavirus du pangolin où une recombinaison a eu lieu aux segments des gènes suivant : le gène de la point S, le gène de la membrane nucléotidique M, et le début du gène de la nucléocapside N. D'autres travaux de **Wassenaar et Zou (2020)** indiquent une relation étroite entre la souche de SARS-CoV-2 "CHN/Yunnan-01/2020" et "bat-SL-CoVZXC21". Alors la

DISCUSSION

souche "bat-SL-CoVZXC21 " de chauve-souris pourrait être la source du SARS-CoV-2 "CHN/Yunnan-01/2020" humain.

b. Comparaison des souches du sous-groupe β a

Ce sous-groupe comprend les virus de : le Coronavirus équin (ECoV), Coronavirus bovins (BCoV) et le Coronavirus de girafe (GiCoV); dont les résultats confirmés par les travaux de **Maier et al. (2015)** et **Jasper et al. (2013)**.

• Coronavirus des équins (ECoV)

Ce virus n'infecte que les équins, il se présente seulement dans l'arbre du gène S où il présente les souches suivantes : " Obihiro12-1", "Obihiro12-2", "Tokachi09" et "NC99" qui sont toutes regroupés dans le même groupe et pour le même point d'origine ce qui s'explique par les résultats de **Nemoto et al. (2015)** qui montrent que les trois souches virales d'ECoV "Tokachi09", "Obihiro12-1" et "Obihiro12-2" sont génétiquement similaires à "NC99", ainsi que selon **Oue et al. (2011)** les séquences de "Tokachi09" montraient une grande similarité avec ceux de "NC99" qui s'explique par une délétion de 185 nucléotides à partir de quatre bases après l'extrémité 3'-terminale du gène S de "Tokachi09".

• Coronavirus du bovin (BCoV)

Les souches présentent dans cette arbre sont : BCoV "SWE/P/10-4", "SWE/C/92", "DB2", "Mebus/US/72 " et "ENT" de Suède et des Etats-Unis, toutes ces souches se positionnent en une même branche ce qui indique la présence d'un certain lien entre ces souches-là ceci est expliqué dans les travaux de **Amoroso et al. (2021)** où les deux souches virales "SWE/P/10-4", "SWE/C/92" isolées de suède en les comparant avec la souche "Mebus/US/72" isolée des États-Unis partagent une identité nucléotidique de 99%, ainsi que selon les travaux de **Bidokhti et al. (2013)** une substitution de 78 à 113 nucléotides a été indiqué au niveau du gène S de ces souches, et d'après les résultats issus des travaux de **Hasoksuz et al. (2007)** différentes mutations ont été détectées au niveau du gène S des souches suivantes BCoV "Mebus/US/72, " DB2" et " ENT".

• Coronavirus de girafes GiCoV

L'arbre phylogénétique du gène S présent les souches virales qui infecte les girafes : "OH3/2006", "OH3/2003" et "OH3-TC/2006", ils ont été regroupés près des souches virales de BCoV en indiquant la présence d'un lien de parenté entre ces souches-là, ce qui est confirmée par les travaux de **Hasoksuz et al. (2007)** qui indiquent que ces souches ont été isolées chez trois girafes souffrant de diarrhée légère à sévère aux États-Unis, l'analyse comparative des séquences entières du génome ont révélé la parenté la plus proche entre ces

DISCUSSION

souches et la souche "ENT " de BCoV et une parenté légèrement plus éloignée entre ces souches de GiCoV et la souche "DB2" de BCoV, alors qu'avec la souche "Mebus" de BCoV ces souches de GiCoV étaient les plus éloignées. Ainsi que **Hasoksuz et al. (2007)** ont rapporté une délétion dans la présumé région hypervariable du gène de pointe de "OH3" du virus de GiCoV par rapport à celui des souches de "Mebus" "ENT" "DB2" du virus de BCoV, d'où en peut conclure qu'ils peuvent avoir évolué simultanément à partir d'un ancêtre commun.

IV.2. Comparaison des souches de l'arbre du gène de pointe N (gène structurale)

Cet arbre phylogénétique se divise en quatre groupes : Alphacoronavirus, Betacoronavirus, Gammacoronavirus, Deltacoronavirus ce qui se confirme par la classification de **Snijders (2008)**.

IV.2.1. Comparaison des souches de l'arbre du groupe Betacoronavirus du gène N (gène de nucléocapside)

Dans cet arbre phylogénétique, ce groupe comprend les séquences nucléotidiques des virus : Les Coronavirus bovins (BCoV), Coronavirus respiratoire de canin (CRCoV), le virus corona humain OC43 (HCoV-OC43) et Coronavirus du syndrome respiratoire aigu sévère (SRAS-CoV). Tous ces virus sont classés d'après les recherches de **Mavrodiev et al. (2020)** en deux sous-groupes : "sous-groupe a" et "sous-groupe b".

a. Comparaison des souches du sous-groupe β a

Ce sous-groupe comprend les virus de : Coronavirus bovins (BCoV), le virus corona humain OC43 (HCoV-OC43) et le Coronavirus respiratoire de canin (CRCoV) ; résultats confirmés par les travaux de **Maier et al. (2015)** et **Jasper et al. (2013)**.

Les souches présentes dans l'arbre du gène N sont : "Québec", "EPI/Caen/2013/08", "bcovizsm" ; la souche virale " Québec" se trouve en même branche avec CRCoV "K37" ce qui est interprété dans les travaux de **Lu et al. (2017)** par le fait que la souche virale BCoV "Québec" isolée de Canada est liée plus étroitement au CRCoV "K37" isolé de Corée du Sud. Concernant la souche BCoV "bcovizsm" isolée d'Italie elle est regroupée avec la souche "EPI/Caen/2013/08" en même branche ce qui nous indique la présence d'un lien ; d'après **Amoroso et al. (2021)** le "bcovizsm" présente la plus forte homologie (99%) avec une souche entérique isolée d'un bovin en France "EPI/Caen/2013/08" avec une seule substitution de base A au lieu de C.

DISCUSSION

Cependant cet arbre présente les souches isolées en Chine qui sont : " HCoV-OC43", " HK04-02", "HK04-01" et la souche de Belgique "BE04", où ces souches représentent la même origine, ce point est confirmé par les travaux de **Lau et al. (2011)** qui ont justifié que ces souches sont issues d'une recombinaison, de plus l'analyse de l'horloge moléculaire à l'aide de séquences de ce gène a suggéré que l'ancêtre commun le plus récent de tous les génotypes HCoV-OC43 est apparu dans les années 1950, résultat confirmé aussi par **Forni et al. (2021)** qui montre que la source de ce virus est les rongeurs dont il est largement admis que les bovins étaient les hôtes intermédiaires assurant la transmission du HCoV-OC43 aux humains, ainsi que l'équipe de **Vijgen et al. (2004)** et **Virant et al. (2021)** suggérant que vers 1890, le BCoV a pu sauter la barrière des espèces et il est devenu capable d'infecter les humains ; entraînant l'émergence d'un nouveau type de Coronavirus humain HCoV-OC43.

b. Comparaison des souches du sous-groupe β b

Le sous-groupe b dans ce cas, comprend le Coronavirus du syndrome respiratoire aigu sévère (SRAS-CoV-2) ce qui est justifié par le classement des groupes adoptés par **Jasper et al. (2013)**.

Les séquences nucléotidiques du virus de SARS-CoV représentent les souches : "CHN/Yunnan-01/2020", "RaTG13", "BtRs-HuB2013" et "bat-SL-CoVZXC21". Ces souches sont isolées de la Chine et sont toutes regroupées en même branches indiquant la présence d'un lien entre eux ce qui est interprété par les travaux de **Lohrasbi-Nejad (2021)** qui a déduit qu'un événement de recombinaison a été observé lors de la comparaison de la séquence du génome des souches virales de SRAS-CoV-2 humain, la souche "CHN/Yunnan-01/2020" à ceux de la chauve-souris "BtRs-HuB2013" et " bat-SL-CoVZXC21" ainsi que le souche "GX-P2V" de Coronavirus du pangolin où une recombinaison a eu lieu aux segments des gènes suivant le gène de la point S, le gène de la membrane nucléotidique M, et le début du gène de la nucléocapside N. De plus que leurs résultats ont montré la présence d'événements de recombinaison entre la SRAS-CoV-2 "CHN/Yunnan-01/2020" et "BatCoV RaTG13" avec une similarité de 96% entre leurs génomes. D'autres travaux de **Wassenaar et Zou (2020)** indiquent une relation étroite entre la souche de SARS-CoV-2 "CHN/Yunnan-01/2020" et "bat-SL-CoVZXC21". Alors la souche "bat-SL-CoVZXC21" de chauve-souris pourrait être la source du SARS-CoV-2 "CHN/Yunnan-01/2020" humain.

DISCUSSION

IV.2.2. Comparaison des souches de l'arbre du groupe Alphacoronavirus du gène N (gène de nucléocapside)

Ce groupe viral se divise en deux sous-groupes : sous-groupe α et le sous-groupe β . chacun de ces deux derniers englobe un groupe de virus qui lui propre.

a. Comparaison des souches du sous-groupe α

Ce groupe comporte les séquences nucléotidiques de : Coronavirus canin (CCoV), le virus de la péritonite infectieuse féline (FCoV). Ces virus sont classés dans le sous-groupe α , ceci est confirmé par les recherches de **Snijders (2008)**.

Une séquence nucléotidique de la souche des États-Unis "171" qui infecte le chien *Canis lupus familiaris* ce regroupe dans le même clade avec une souche du virus de la péritonite infectieuse féline (FCoV) nommée par "WSU 79-1146" située dans l'arbre du gène N ce qui est précédemment expliqué par l'origine commune entre CCoV et FCoV prouvé par les recherches de **Kathy et al. (2002)**. Cependant les souches de FCoV nommées par : "KUTT/2017", "KU30/2019", "KU62/2020" et "WSU 79-1146", montrent une origine commune confirmée par les travaux de **Tuanthap et al. (2021)**.

b. Comparaison des souches du sous-groupe β

Ce groupe comporte les séquences nucléotidiques de : virus de la diarrhée épidémique porcine (PEDV). Dans notre arbre du gène N, ce virus n'infecte que les porcs isolés des Philippines, la Chine, le Japon et les Etats-Unis représentées par les séquences des souches : "IA1", "CH/ZMDZY/11", "AH2012", "14JM-168", "14 JM-179" et "K3" et "K4" caractérisé chez les porcs de l'espèce de *Sus scrofa* qui présentent la même origine mais qui se résultent à des mutations positionnées justifié par les travaux de **Léopardi et al. (2020)**.

IV.2.3. Comparaison des souches de l'arbre du groupe Gammacoronavirus du gène N (gène de nucléocapside)

L'arbre du gène N présente seulement la souche "CK/SWE/0658946/10" qu'Abro **et al. (2012)** détectent les points chauds de recombinaison qui n'ont été détecté que dans le gène N (commencent par la position du point d'arrêt au nucléotide 26 874 et se terminent par le point d'arrêt 26 967).

IV.2.4. Comparaison des souches de l'arbre du groupe Deltacoronavirus du gène N (gène de nucléocapside)

Ce groupe se construit à la fois par les trois virus : Deltacoronavirus du porc (PDCoV), Coronavirus du canard (DCoV), Coronavirus du canard d'Amérique (HKU-20) ; comme

DISCUSSION

indiqué par **Snijders (2008)** et **Hepojoki et al. (2017)** ce groupe comprend plusieurs CoV nouvellement évolués issus des pools génétiques des oiseaux sauvages et domestiques.

- **Coronavirus du canard d'Amérique (HKU-20) et Coronavirus du canard (DCoV)**

Ces deux sont présents dans cette arbre où la séquence de HKU-20 regroupé en un ancêtre commun avec les deux souches de DCoV ce qui est justifié par les résultats phylogéniques de **Paim et al. (2019)** qui montrent que les trois δ -CoV étaient les plus étroitement liés au Coronavirus du canard siffleur HKU20 en partageant 83,38 % à 84,39 % d'identité nucléotidique dans le gène N.

- **Deltacoronavirus du porc (PDCoV)**

Ce virus présente les souches : "S5015L/2015", "P1 16 BTL", "CHN-AH-2004", "SD", "Illinois121/2014", "KNU14-04" et "8734/IA/2014" qui se regroupent dans la même branche ; ce qui est justifié par les travaux de **Hu et al. (2015)** qui montrent que ces souches de PDCoV sont étroitement liées. Notamment, la souche PDCoV "KNU14-04", était la plus étroitement liée aux souches émergentes de PDCoV aux États-Unis, suggérant qu'elles peuvent provenir d'une souche ancestrale similaire. Le renforcement des résultats se base sur le pourcentage de similarités calculées dans les travaux d'**Athip et al. (2016)** qui montre que la souche "P1 16 BTL" partage une identité nucléotidique avec la souche d'Amérique "8734/IA/2014" qui est de 97,2% et 97,8% entre le génome complet et la protéine N successivement. Cependant l'identité entre "P1 16 BTL" et les souches chinoise "CHN-AH-2004", "SD" "Illinois121/2014" et "KNU14-04" est près de 97% pour le génome complet et se varie entre 98,1% et 98,7% en comparant leurs gènes N.

IV.3. Comparaison des souches de l'arbre du gène RdRp (gène non structurale)

Cet arbre phylogénétique se divise en trois groupes : Alphacoronavirus, Betacoronavirus, Gammacoronavirus, ce qui se confirme par la classification de **Snijders (2008)**.

IV.3.1. Comparaison des souches de l'arbre du groupe Betacoronavirus du gène RdRp

Dans notre arbre phylogénique, ce groupe comprend les séquences nucléotidiques des virus : Les Coronavirus bovins (BCoV), le virus corona humain OC43 (HCoV-OC43), le Coronavirus humain (HKU-1). Tous ces virus sont classés d'après les recherches de **Mavrodiev et al. (2020)**.

DISCUSSION

- **Virus de porcine encéphalomyélite hémagglutinante PHEV**

Ce virus n'infecte que les porcs qui sont la seule espèce sensible à l'infection naturelle par le PHEV, d'après **Mora-Díaz et al. (2019)** le PHEV circule chez le porc depuis des décennies. Ce virus se trouve dans l'arbre du gène d'RdRp et sous les noms de souches suivantes : "18DIA90174-2", "18DIA90177-20", "18DIA90177-14" et "18DIA90178-5", toutes ces souches de PHEV sont regroupées avec les souches de BCoV ce qui est expliqué par les résultats de **González et al. (2003)** ou ils ont démontré que PHEV est étroitement lié aux Coronavirus canin, bovin, murin, humain et équin.

- **Coronavirus du bovin (BCoV)**

Les souches Slovénie et Italie se présentes dans l'arbre de gène RdRp sont : "TZ65/2014", "TP14/2012" et "TP73/2014". Où ces souches sont regroupées au niveau de mêmes branches dans l'arbre de gène RdRp indiquent la présence d'un lien entre eux ce qui est justifiée par les études menées par **Virant et al. (2021)** qui indiquent que ces souches sont étroitement liées les uns aux autres avec 99,2 à 100% d'identité nucléotidique.

- **Virus du corona humain OC43 (HCoV-OC43)**

Les séquences extraites de l'humain en Slovénie se présentent dans l'arbre du gène RdRp sont : "62794/2016" et "44364/2014". Où la souche "62794/2016" été regroupe en même branche avec la souche "44364/2014" ceci est expliqué par les résultats de **Virant et al. (2021)** que ces deux partagent de 98,7 à 100 % d'identité nucléotidique au niveau du gène RdRp.

- **Coronavirus humain HKU-1**

Dans notre étude ce virus infectant l'humain Slovénie, Dans l'arbre phylogénétique du gène RdRp par les souches présent sont : "61107/2015" et "63615/2016" où ils indiquent qu'il y'a une similarité entre eux et qu'ils partagent un ancêtre commun qui se confirme dans les travaux de **Woo et al. (2005)** par l'analyse phylogénétique qu'elle a montré que ces derniers ont formé des branches distinctes dans l'arbre phylogénétique.

Quant aux séquences des virus "RtMc-CoV-1" et "RtMc-CoV-2" ils infectent le rongeur *Mus caroli* ; et les souches "RtRf-CoV-1", "RtRf-CoV-2", "RtNn-CoV" et "RtAa-CoV" infectent respectivement les espèces de rongeur : *Rattus tanezumi*, *Rattus flavipectus*, *Niviventer niviventer* et *Apodemus agrarius*. Ces séquences qui sont regroupés avec les virus non classés dans la taxonomie de GenBank, dans notre arbre du RdRp ont été regroupés avec les virus de Betacoronavirus de la "lignée a", ceci est confirmée par les travaux de **Wu et al. (2018)** où les

DISCUSSION

analyses phylogénétiques ont révélé que ces virus formaient divers clades évolutifs avec des identités de séquence du gène RdRp comprise entre 88,1% et 98,9%.

IV.3.2. Comparaison des souches de l'arbre du groupe **Gammacoronavirus** du gène **RdRp**

Ce groupe ne comporte que les séquences des souches virales représentant le Coronavirus aviaire (ACoV). Où ce dernier se présente par les souches : "MW434/2009", "MW435/2009", "MW345/2009", "MW283/2009" et "MW284/2009" qui se regroupent en deux groupes, "MW434/2009" et "PL-MW435/2009" dans un groupe, le reste dans un autre groupe, ce qui confirme exactement les résultats phylogénétiques de Domansk-Blicharz *et al.* (2014), qui ont montré aussi que ces cinq oiseaux sauvages positifs à l'ACoV étaient homologues entre 88,1 et 100 %, les trois souches "MW345/2009", "MW283/2009" et "MW284/2009" formaient un cluster commun et leur homologie nucléotidique était comprise entre 92,7 et 100 % ; les deux souches "MW434/2009" et "MW435/2009" formaient un groupe séparé situé à distance des souches IBV/TCoV.

IV.3.3. Comparaison des souches de l'arbre du groupe **Alphacoronavirus** du gène **RdRp**

Ce groupe se présente par ces deux sous-groupes : sous-groupe a et le sous-groupe b

a. Comparaison entre les souches du sous-groupe **aa**

Ce groupe comporte les séquences nucléotidiques du Coronavirus respiratoire porcine (PRCV), dont ceci est confirmé par les recherches de **Snijders (2008)**.

C'est le virus qui touche le furet Pérou, dont il est représenté dans l'arbre du gène RdRp par les souches "furet /cerveau/Lima-2014" et "furet/cervelet/Lima-2014" qui se regroupent dans la même branche c'est ce qui est justifié par les travaux de **Lescano et al. (2015)** qui montrent par l'analyse BLAST que ces deux souches avaient un SNP (polymorphisme nucléotidique unique) dans la séquence amplifiée.

b. Comparaison entre les souches du sous-groupe **ab**

Ce groupe comporte les séquences nucléotidiques de : virus de la diarrhée épidémique porcine (PEDV), Coronavirus de chauve-souris (Bat CoV) et le Coronavirus humain 229E (HCoV-229E).

- **Virus de la diarrhée épidémique porcine (PEDV)**

Dans notre arbre du gène RdRp : la présence des deux souches "18DIA90174-10" et "18DIA90174-11" caractérisée chez les porcs d'Italie qui se positionnent dans la même

DISCUSSION

branche ; ce qui est justifié par le même positionnement phylogénique dans les travaux de **Léopardi et al. (2020)**.

- **Coronavirus de chauve-souris (Bat CoV)**

Ce virus se présente dans l'arbre par les séquences nucléotidiques des souches de Mozambique : "18987", "18993", "19006", "19024", "18990", "19026", "19027" et la souche "015" de France qui ciblent successivement les espèces de Chauve-souris : *Hipposideros caffer*, *Triaenops afer*, *Miniopterus mossambicus*, *Rhinolophus rhodesiae*, *Rhinolophus lobatus*, *Rhinolophus sp* et *Rhinolophus rhodesiae*. Tous s'assemblent pour une origine commune dans l'arbre du gène RdRp qui se rend compte aux travaux de **Joffrin et al. (2020)** avec un taux élevé de similarité de séquences allant de 60,2 % à 99,8 %.

- **Coronavirus humains HCoV-229E**

Une séquence nucléotidique du HCoV-229E qui infecte l'humain isolé de Slovénie dans l'arbre du gène RdRp elle se positionne dans la même branche avec les BatCoV ; ce qui est identique avec les résultats de **Léa et al. (2020)** qui trouvent un pourcentage de 93 % de similitude entre ces deux virus.

IV.2.1. Implication des mutations dans l'évolution

En étudiant les liens entre les différentes souches virales de Coronavirus de notre travail nous avons remarqué qu'un grand nombre de changements nucléotidiques ou protéiques dans la structure des séquences nucléotidiques et des acides aminés de différents gènes influencent sur l'organisation moléculaire de ces derniers, points signalés par **Sci (2005)**, qui définissent les mutations comme agent perturbant la composition génomique (**Figure 54**) qui se traduit finalement par des protéines non fonctionnelles, ou protéine non traduites, ou changée par une autre, suite à un décalage dans le cadre de lecture. Le **tableau XV** nous montre ces mutations et leurs conséquences ; sachant que le génome de Coronavirus comprend deux principaux cadres de lecture ouverts (ORF), ORF1a et ORF1b, englobant les deux tiers du génome, où les deux premiers codent pour 16 protéines non structurales, le tiers restant du génome code pour des ORF qui se chevauchent, codant les quatre protéines structurales majeures (**Bergmann et al., 2020 ; Zhang et al., 2021**).

Afin de déduire l'effet de ces mutations qui touchent les différentes souches ainsi que les différents types de mutations qui touchent les gènes traités ainsi que leurs effets sur la ou les souche(s) qui la concernent. On peut résumer les points suivants :

DISCUSSION

IV.2.2. Conséquences des mutations

- Toutes les mutations citées des conséquences visibles dans l'histoire de l'évolution, car une mutation dans la séquence nucléotidique qui se traduit à la base par un changement dans le cadre de lecture, est donc une protéine changée ou une protéine non fonctionnelle, le changement de la protéine se traduit à son tour par un phénotype différent au premier (**Ports, 1925**).
- Une petite mutation qui touche les gènes se reflète par un changement dans leurs morphologies externes qui lui permettent d'infecter un nouvel hôte et donc on aura une nouvelle souche virale, c'est l'exemple de la mutation de la souche "K3" de PEDV qui présente les substitutions (**Garcia et al. (2018)** et **Van et al. (2018)**) suivantes : C en G à position 188 ; G en A à la position 192 ; C en A au poste 455 ; T en A à la position 457 ; A en T aux positions 487, 494 et 625 ; A en G aux positions 588, 591 et 616 ; et A en C en position 610, une délétion a également été notée aux positions 36, 619, 627 et 708.
- Cependant les grandes mutations (en termes de nucléotides), ou accumulation des mutations, ou bien les recombinaisons qui se passent entre les différentes souches ou les différents virus ; produits de nouveaux virus qui se regroupent dans de nouveaux groupes viraux ; **Ports (1925)** suggère que les variations décrites comme infiniment petites, peuvent s'accumuler dans n'importe quelle direction pour produire de nouvelles espèces, c'est l'exemple du virus de HCoV-NL63 et HCoV-229 E ou l'équipe de **Hoek et al. (2004)** ont détecté la présence d'un grand insert de 537 nucléotides dans la partie 5' du gène S. Ainsi que les événements de recombinaison inter-genre prouvée par les travaux de **Tao et al. (2017)** et **Corman et al. (2015)** qui l'ont déduit entre différents Coronavirus de chauve-souris *Hipposideridae* (BtCoV-229E, BtCoV-229E-8 et BtCoV-229E-1).

Conclusion

L'intérêt de notre étude est fondé sur les liens entre les souches de Coronavirus dans différentes zones géographiques ; nous avons utilisé la phylogénie qui est une méthode utile dans la génétique des populations et l'épidémiologie, qui est une composante essentielle dans la biologie pour l'étude des liens et l'histoire évolutive des organismes et sa diversification sur le plan moléculaire et morphologique. Pour la réalisation de ce travail, nous avons utilisé un outil de bio-informatique permettant l'analyse phylogénétique et la construction de l'arbre phylogénétique, le logiciel MEGA 11. Les analyses ont été appliquées sur plus de 250 séquences dans lesquelles 35 séquences nucléotidiques de gène RdRp, 49 séquences nucléotidiques de séquences de gène S et 36 séquences de gène N ont été sélectionnées pour classer les souches virales de Coronavirus de chaque genre selon les espèces touchées et les zones géographiques où elles ont été trouvées.

Les résultats de MEGA 11 ont démontré que, les séquences nucléotidiques du gène ARN-polymérase-ARN-dépendant (RdRp) étudiées se regroupent en 3 groupes représentés par trois clades différents, alors que les gènes de la protéine de pointe (S) et la protéine (N) se divisent en quatre groupes représentés par quatre clades différents ; les différences en nombre de clades et sous-clades sont justifiées par l'absence des analyses moléculaires de plusieurs souches virales qui peuvent toucher différentes espèces animales. La large répartition de chauve-souris dans le monde et la représentation de 43% de rongeurs des mammifères surtout en Asie justifie la présence de leurs souches virales dans les deux clades Alphacoronavirus et Betacoronavirus. Ainsi que la présence de séquences de coronavirus aviaires (infectant les oiseaux) dans le clade de Gamma Coronavirus, et la répartition des autres souches (infecte les mammifères) dans différents clades.

Nous avons démontré aussi qu'il y a une grande diversité génétique entre les souches de Coronavirus qui est due à l'infidélité des gènes qui se traduit par leurs taux de mutations très élevés qui ont des conséquences visibles dans l'histoire de l'évolution. Car une mutation dans la séquence nucléotidique qui se traduit à la base par un changement dans le cadre de lecture, est donc une protéine changée ou une protéine non fonctionnelle, le changement de la protéine se traduit à son tour par un nouveau phénotype. Où une petite mutation touche les gènes et se reflète par un changement dans leurs morphologies externes qui lui permet de changer de tropisme ou d'infecter un nouvel hôte et donc on aura une nouvelle souche virale. Cependant

nous avons remarqué que le gène le plus altéré à ces changements est le gène qui code pour la protéine de pointe S.

Ce travail nous a permis d'établir des liens et des connexions entre des souches spécifiques, ainsi nous avons pu mettre en évidence la position de l'humain dans la transmission du coronavirus. Alternativement, les humains sont exposés à un risque constant d'infection grâce à la capacité de ce virus à franchir la barrière zoonotique.

Annexe

Animal	<i>Espèce</i>	Région	Gène	N° D'accession
Humain	<i>Homo sapiens</i>	china	ORF8	AY390556
	<i>Homo sapiens</i>	china	ORF8	AY394981
	<i>Homo sapiens</i>	china	ORF8	AY394984
	<i>Homo sapiens</i>	china	ORF8	AY394995
	<i>Homo sapiens</i>	china	ORF8	AY394996
	<i>Homo sapiens</i>	china	ORF8	AY394997

Références

- **Abro S. H., Renström L.H., Ullman K., Belák S., Baule C. 2012.** Characterization and analysis of the full-length genome of a strain of the European QX-like genotype of infectious bronchitis virus. *Archives of virology*, 157(6), 1211–1215, P : 5.
- **Alekseev K. P., Vlasova A. N., Jung K., Hasoksuz M., Zhang X., Halpin R., Wang S., Ghedin E., Spiro D., Saif L. J. 2008.** Bovine-like coronaviruses isolated from four species of captive wild ruminants are homologous to bovine coronaviruses, based on complete genomic sequences. *Journal of virology*, 82(24), 12422–12431, P : 10.
- **Alluwaimi A. M., Alshubaith I. H., Al-Ali A. M., Abohelaika S. 2020.** The Coronaviruses of Animals and Birds: Their Zoonosis, Vaccines, and Models for SARS-CoV and SARS-CoV2. *Frontiers in veterinary science*, 7, 582287, P : 12.
- **Amoroso M. G., Lucifora G., Degli Uberti B., Serra F., De Luca G., Borriello G., De Domenico A., Brandi S., Cuomo M. C., Bove F., Riccardi M. G., Galiero G., Fusco, G. 2020.** Fatal Interstitial Pneumonia Associated with Bovine Coronavirus in Cows from Southern Italy. *Viruses*, 12(11), P : 12.
- **Benson D.A., Karsch-Mizrachi I., Lipman D.J., Ostell J., Wheeler D.L., 2005.** GenBank, Nucleic Acids Research , volume 33, numéro suppl_1, , pages D34 à D38,
- **Bergmann C. C., Silverman, R. H. 2020.** COVID-19: Coronavirus replication, pathogenesis, and therapeutic strategies. *Cleveland Clinic journal of medicine*, 87(6), 321–327, P : 7.
- **Bidokhti M., Trávén M., Krishna N. K., Munir M., Belák S., Alenius S., Cortey M. 2013.** Evolutionary dynamics of bovine coronaviruses: natural selection pattern of the spike gene implies adaptive evolution of the strains. *The Journal of general virology*, 94(Pt 9), 2036–2049, P : 14.
- **Bottard S. 1996.** Application de la méthode du Bootstrap pour l'estimation des valeurs extrêmes dans les distributions de l'intensité des séismes. *Revue de statistique appliquée*, 44(4), 5-17, P : 14.
- **Callaway E., Ledford H., Mallapaty, S. 2020.** Six months of coronavirus : the mysteries scientists are still racing to solve. *Nature*, 583(7815), 178–179, P : 2.
- **Chan J. F., To K. K., Tse H., Jin D. Y., Yuen K. Y. 2013.** Interspecies transmission and emergence of novel viruses: lessons from bats and birds. *Trends in microbiology*, 21(10), 544–555, P : 12.

- **Chung J. Y., Kim H. R., Bae Y. C., Lee O. S., Oem J. K. 2011.** Detection and characterization of bovine-like coronaviruses from four species of zoo ruminants. *Veterinary microbiology*, 148(2-4), 396–401, P : 6.
- **Cook G. W., Benton M. G., Akerley W., Mayhew G. F., Moehlenkamp C., Raterman D., Burgess D. L., Rowell W. J., Lambert C., Eng K., Gu J., Baybayan P., Fussell J. T., Herbold H. D., O'Shea J. M., Varghese T. K., Emerson L. L. 2020.** Structural variation and its potential impact on genome instability: Novel discoveries in the EGFR landscape by long-read sequencing. *PloS one*, 15(1), e0226340, P : 21.
- **Corman V. M., Baldwin H. J., Tateno A. F., Zerbinati R. M., Annan A., Owusu M., Nkrumah E. E., Maganga G. D., Oppong S., Adu-Sarkodie Y., Vallo P., da Silva Filho L. V., Leroy E. M., Thiel V., van der Hoek L., Poon L. L., Tschapka M., Drosten C., Drexler J. F. 2015.** Evidence for an Ancestral Association of Human Coronavirus 229E with Bats. *Journal of virology*, 89(23), 11858–11870, P : 13.
- **Cui J. T., Qiao H., Hou C. Y., Zheng H. H., Li X. S., Zheng L. L., Chen H. Y. 2020.** Characteristics of the spike and ORF3 genes of porcine epidemic diarrhea virus in Henan and Shanxi provinces of China. *Archives of virology*, 165(10), 2323–2333, P : 11.
- **Cui J., Li F., Shi Z. L. 2019.** Origin and evolution of pathogenic coronaviruses. *Nature reviews. Microbiology*, 17(3), 181–192, P : 12.
- **Domanska-Blicharz K., Jacukowicz A., Lisowska A., Wyrostek K., Minta Z. 2014.** Detection and molecular characterization of infectious bronchitis-like viruses in wild bird populations. *Avian pathology : journal of the W.V.P.A*, 43(5), 406–413, P : 11.
- **Du L., He Y., Zhou Y., Liu S., Zheng B. J., Jiang S. 2009.** The spike protein of SARS-CoV--a target for vaccine and therapeutic development. *Nature reviews. Microbiology*, 7(3), 226–236, P : 11.
- **Enjuanes L., Gorbalenya A. E., Groot R. J. de, Cowley J. A., Ziebuhr J., Snijder E. J. 2008.** Nidovirales. In B. W. J. Mahy & M. H. V. V. Regenmortel (Eds.), *Encyclopedia of Virology (Third Edition)* (Third Edition, pp. 419–430). Academic Press, P : 12.
- **Eric S. D., Nicholas T. D., Theophilus K. A. 2014.** Bioinformatics with basic local alignment search tool (BLAST) and fast alignment (FASTA). *Journal of Bioinformatics and sequence analysis*, 6(1), P : 6.
- **Erles K., Toomey C., Brooks H. W., Brownlie J. 2003.** Detection of a group 2 coronavirus in dogs with canine infectious respiratory disease. *Virology*, 310(2), 216–223, P : 8.

- **Fehr A. R., Perlman S. 2015.** Coronaviruses: an overview of their replication and pathogenesis. *Methods in molecular biology (Clifton, N.J.)*, 1282, P : 23.
- **Forni D., Cagliani R., Arrighoni F., Benvenuti M., Mozzi A., Pozzoli U., Clerici M., De Gioia L., Sironi M. 2021.** Adaptation of the endemic coronaviruses HCoV-OC43 and HCoV-229E to the human host. *Virus evolution*, 7(2), veab061, P : 11.
- **Frickey T., Lupas A. N. 2004.** PhyloGenie: automated phylome generation and analysis. *Nucleic acids research*, 32(17), 5231–5238, P : 8.
- **Fung TS., Liu DX. 2019.** Coronavirus humain : interaction hôte-pathogène. Annual review of microbiology, 73 :52*-557
- **Garcia G. G., Aquino M., Balbin M. M., Belotindos L. P., Supnet J. G., Mingala C. N. 2018.** Characterisation of porcine epidemic diarrhea virus isolates during the 2014-2015 outbreak in the Philippines. *Virusdisease*, 29(3), P : 7.
- **González J. M., Gomez-Puertas P., Cavanagh D., Gorbalenya A. E., Enjuanes L. 2003.** A comparative sequence analysis to revise the current taxonomy of the family Coronaviridae. *Archives of virology*, 148(11), P : 29.
- **Gordon D. E., Jang G. M., Bouhaddou M., Xu J., Obernier K., White K. M., O'Meara M. J., Rezelj V. V., Guo J. Z., Swaney D. L., Tummino T. A., Hüttenhain R., Kaake R. M., Richards A. L., Tutuncuoglu B., Foussard, H., Batra J., Haas K., Modak M., Kim M., ... Krogan N. J. 2020.** A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*, 583(7816), P : 48.
- **Hanna N., Parfait B., Vidaud D., Vidaud M. 2005.** Mécanismes et conséquences des mutations [Mutation mechanisms and their consequences]. *Medecine sciences : M/S*, 21(11), P : 11.
- **Herrewegh A. A., Smeenk I., Horzinek M. C., Rottier P. J., de Groot R. J. Feline coronavirus type II strains 79-1683 and 79-1146 originate from a double recombination between feline coronavirus type I and canine coronavirus 1998.** *Journal of virology*, 72(5), 4508–4514, P : 7.
- **Hu D., Zhu C., Ai L., He T., Wang Y., Ye F., Yang L., Ding C., Zhu X., Lv R., Zhu J., Hassan B., Feng Y., Tan W., Wang, C. 2018.** Genomic characterization and infectivity of a novel SARS-like coronavirus in Chinese bats. *Emerging microbes & infections*, 7(1), 154, P : 10.

- **Hu H., Jung K., Vlasova A. N., Chepngeno J., Lu Z., Wang Q., Saif L. J. 2015.** Isolation and characterization of porcine deltacoronavirus from pigs with diarrhea in the United States. *Journal of clinical microbiology*, 53(5), 1537–1548, P : 12.
- **Huang Y., Yang C., Xu X. F., Xu W., Liu S. W. 2020.** Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. *Acta pharmacologica Sinica*, 41(9), 1141–1149, P : 9.
- **Hulswit R. J., de Haan C. A., Bosch B. J. 2016.** Coronavirus Spike Protein and Tropism Changes. *Advances in virus research*, 96, P : 29.
- **Jevšnik Virant M., Černe D., Petrovec M., Paller T., Toplak I. Genetic Characterisation and Comparison of Three Human Coronaviruses (HKU1, OC43, 229E) from Patients and Bovine Coronavirus (BCoV) from Cattle with Respiratory Disease in Slovenia. 2021.** *Viruses*, 13(4), 676, P : 10.
- **Joffrin L., Goodman S. M., Wilkinson D. A., Ramasindrazana B., Lagadec E., Gomard Y., Le Minter G., Dos Santos A., Corrie Schoeman M., Sookhareea R., Tortosa P., Julienne S., Gudo E. S., Mavingui P., Lebarbenchon C. 2020.** Bat coronavirus phylogeography in the Western Indian Ocean. *Scientific reports*, 10(1), 6873, P : 11.
- **Johnson M., Zaretskaya I., Raytselis Y., Merezuk Y., McGinnis S., Madden T. L. 2008.** NCBI BLAST: a better web interface. *Nucleic acids research*, 36(Web Server issue), W5–W9, P : 5.
- **Kasmi Y., Khataby K., Souiri A., Ennaji M. M. 2020.** Coronaviridae: 100,000 years of emergence and reemergence. In *Emerging and reemerging viral pathogens* (pp. 127-149). Academic Press, P : 11.
- **Kenney S. P., Wang Q., Vlasova A., Jung K., Saif L. 2021.** Naturally Occurring Animal Coronaviruses as Models for Studying Highly Pathogenic Human Coronaviral Disease. *Veterinary pathology*, 58(3), 438–452, P :14.
- **Ketfi A., Chabati O., Chemali S., Mahjoub M., Gharnaout M., Touahri R., Djenouhat K., Selatni F., Saad H. B. 2020.** Profil clinique, biologique et radiologique des patients Algériens hospitalisés pour COVID-19: données préliminaires. *The Pan African medical journal*, 35(Suppl 2), 77, P : 10.
- **Kin N., Mischczak F., Lin W., Gouilh M. A., Vabret A., EPICOREM Consortium 2015.** Genomic Analysis of 15 Human Coronaviruses OC43 (HCoV-OC43s) Circulating in

France from 2001 to 2013 Reveals a High Intra-Specific Diversity with New Recombinant Genotypes. *Viruses*, 7(5), 2358–2377, P : 20.

- **Kong F., Wang Q., Kenney S. P., Jung K., Vlasova A. N., Saif L. J. 2022.** Porcine Deltacoronaviruses: Origin, Evolution, Cross-Species Transmission and Zoonotic Potential. *Pathogens (Basel, Switzerland)*, 11(1), 79, P : 17.
- **Laber K. E., Whary M. T., Bingel S. A., Goodrich J. A., Smith A. C., Swindle M. M. 2002.** Chapter 15—Biology and Diseases of Swine. In J. G. Fox, L. C. Anderson, F. M. Loew, & F. W. Quimby (Eds.), *Laboratory Animal Medicine (Second Edition)* (Second Edition, pp. 615–673). Academic Press, P : 59.
- **Lau S. K., Woo P. C., Yip C. C., Tse H., Tsoi H. W., Cheng V. C., Lee P., Tang B. S., Cheung C. H., Lee R. A., So L. Y., Lau Y. L., Chan K. H., Yuen K. Y. 2006.** Coronavirus HKU1 and other coronavirus infections in Hong Kong. *Journal of clinical microbiology*, 44(6), 2063–2071, P : 9.
- **Lau S. K., Lee P., Tsang A. K., Yip C. C., Tse H., Lee R. A., So L. Y., Lau Y. L., Chan K. H., Woo P. C., Yuen, K. Y. 2011.** Molecular epidemiology of human coronavirus OC43 reveals evolution of different genotypes over time and recent emergence of a novel genotype due to natural recombination. *Journal of virology*, 85(21), 11325–11337, P : 13.
- **Leghari R. A., Fan B., Wang H., Bai J., Zhang L., Abro S. H., Jiang, P. 2016.** Full-length genome sequencing analysis of avian infectious bronchitis virus isolate associated with nephropathogenic infection. *Poultry science*, 95(12), 2921–2929, P : 10.
- **Leopardi S., Priori P., Zecchin B., Zamperin G., Milani A., Tonon F., Giorgiutti M., Beato M. S., De Benedictis P. 2020.** Interface between Bats and Pigs in Heavy Pig Production. *Viruses*, 13(1), 4. P : 18.
- **Lescano J., Quevedo M., Gonzales-Viera O., Luna L., Kee M. K., Gregori, F. 2015.** First Case of Systemic Coronavirus Infection in a Domestic Ferret (*Mustela putorius furo*) in Peru. *Transboundary and emerging diseases*, 62(6), 581–585, P : 5.
- **Licitra B. N., Duhamel G. E., Whittaker G. R. 2014.** Canine enteric coronaviruses: emerging viral pathogens with distinct recombinant spike proteins. *Viruses*, 6(8), 3363–3376, P : 14.
- **Lohrasbi-Nejad A. 2022.** Detection of homologous recombination events in SARS-CoV-2. *Biotechnology letters*, 44(3), P : 16.

- **Lorusso A., Desario C., Mari V., Campolo M., Lorusso E., Elia, G., Martella V., Buonavoglia C., Decaro N. 2009.** Molecular characterization of a canine respiratory coronavirus strain detected in Italy. *Virus research*, 141(1), 96–100, P : 5.
- **Lu S., Ye Q., Singh D., Cao Y., Diedrich J. K., Yates J. R., 3rd Villa E., Cleveland D. W., Corbett K. D. 2021.** The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. *Nature communications*, 12(1), 502, P : 15.
- **Lu S., Wang Y., Chen Y., Wu B., Qin K., Zhao J., Lou Y., Tan W. 2017.** Discovery of a novel canine respiratory coronavirus support genetic recombination among betacoronavirus1. *Virus research*, 237, 7–13, P : 7.
- **Lu S., Chen Y., Qin K., Zhou J., Lou Y., Tan W. 2016.** Genetic and antigenic characterization of recombinant nucleocapsid proteins derived from canine coronavirus and canine respiratory coronavirus in China. *Science China. Life sciences*, 59(6), 615–621, P : 7.
- **Madapong A., Saeng-Chuto K., Lorsirigool A., Temeeyasen G., Srijangwad A., Tripipat T., Wegner M., Nilubol D. 2016.** Complete Genome Sequence of Porcine Deltacoronavirus Isolated in Thailand in 2015. *Genome announcements*, 4(3), e00408-16, P : 2.
- **Mavrodiev E. V., Tursky M. L., Mavrodiev N. E., Ebach M. C., Williams D. M. 2020.** On Classification and Taxonomy of Coronaviruses (Riboviria, Nidovirales, Coronaviridae) with special focus on severe acute respiratory syndrome-related coronavirus 2 (SARS-Cov-2). *bioRxiv*, P : 38.
- **Menachery V. D., Yount B. L., Jr Debbink K., Agnihothram S., Gralinski L. E., Plante J. A., Graham R. L., Scobey T., Ge X. Y., Donaldson E. F., Randell S. H., Lanzavecchia A., Marasco W. A., Shi Z. L., Baric R. S. 2015.** A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nature medicine*, 21(12), 1508–1513, P : 8.
- **Mira Fernandes A., Brandão P. E., Dos Santos Lima M., de Souza Nunes Martins M., da Silva T. G., da Silva Cardoso Pinto V., de Paula L. T., Vicente M., Okuda L. H., Pituco, E. M. 2018.** Genetic diversity of BCoV in Brazilian cattle herds. *Veterinary medicine and science*, 4(3), 183–189. Advance online publication, P : 7.

- **Mora-Díaz J. C., Piñeyro P. E., Houston E., Zimmerman J., Giménez-Lirola L. G. 2019.** Porcine Hemagglutinating Encephalomyelitis Virus: A Review. *Frontiers in veterinary science*, 6, 53, P: 12.
- **Naqvi, A.T., Fatima. K., Mohammad. T., Fatima. U., Singh. I. K., Singh. A., Hassan. M. I. 2020.** Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, 1866(10), 165878.
- **Nemoto M., Oue Y., Murakami S., Kanno T., Bannai H., Tsujimura K., Yamanaka T., Kondo T. 2015.** Complete genome analysis of equine coronavirus isolated in Japan. *Archives of virology*, 160(11), 2903–2906, P : 4.
- **Oong X. Y., Ng K. T., Takebe Y., Ng L. J., Chan K. G., Chook J. B., Kamarulzaman, A., Tee K. K. 2017.** Identification and evolutionary dynamics of two novel human coronavirus OC43 genotypes associated with acute respiratory infections: phylogenetic, spatiotemporal and transmission network analyses. *Emerging microbes & infections*, 6(1), e3, P : 13.
- **Oue Y., Ishihara R., Edamatsu H., Morita Y., Yoshida M., Yoshima M., Hatama S., Murakami K., & Kanno, T. 2011.** Isolation of an equine coronavirus from adult horses with pyrogenic and enteric disease and its antigenic and genomic characterization in comparison with the NC99 strain. *Veterinary microbiology*, 150(1-2), P : 7.
- **Paim F. C., Bowman A. S., Miller L., Feehan B. J., Marthaler, D., Saif L. J., & Vlasova A. N. 2019.** Epidemiology of Deltacoronaviruses (δ -CoV) and Gammacoronaviruses (γ -CoV) in Wild Birds in the United States. *Viruses*, 11(10), 897, P : 11.
- **PAPP. T., MARSCHANG., Rachel. E. 2019.** Detection and characterization of invertebrate iridoviruses found in reptiles and prey insects in Europe over the past two decades. *Viruses*, vol. 11, no 7, p. 600.
- **Peng C. 2005.** Distance based methods in phylogenetic tree construction. *NEURAL PARALLEL AND SCIENTIFIC COMPUTATIONS* 15.4 : 547, P : 9.
- **Pérez-Rivera C., Ramírez-Mendoza H., Mendoza-Elvira S., Segura-Velázquez R., Sánchez-Betancourt J. I. 2019.** First report and phylogenetic analysis of porcine deltacoronavirus in Mexico. *Transboundary and emerging diseases*, 66(4), 1436–1441, P : 6.

- **Pillay T. S. 2020.** Gene of the month: the 2019-nCoV/SARS-CoV-2 novel coronavirus spike protein. *Journal of clinical pathology*, 73(7), 366–369, P : 4.
- **Sallard E., Halloy J., Casane D., van Helden J., Decroly É. 2020.** Retrouver les origines du SARS-CoV-2 dans les phylogénies de coronavirus [Tracing the origins of SARS-COV-2 in coronavirus phylogenies]. *Medecine sciences : M/S*, 36(8-9), 783–796, P :5.
- **Sánchez C. M., Gebauer F., Suñé C., Mendez A., Dopazo J., Enjuanes, L. 1992.** Genetic evolution and tropism of transmissible gastroenteritis coronaviruses. *Virology*, 190(1), 92–105, P : 14.
- **Shaikh S. S., Jose A. P., Nerkar D. A., Vijaykumar Kv M., Shaikh S. K. 2020.** COVID-19 pandemic crisis-a complete outline of SARS-CoV-2. *Future journal of pharmaceutical sciences*, 6(1), 116, P: 19.
- **Siddell. S., Coll M. 1983.** Pathogeneses des onfections virales. *Clinique viterinaire xe lamerique de nord(6)* : 104- 107.
- **So R., Chu D., Miguel E., Perera R., Oladipo J. O., Fassi-Fihri O., Aylet G., Ko R., Zhou Z., Cheng M. S., Kuranga S. A., Roger F. L., Chevalier V., Webby R. J., Woo P., Poon L., Peiris M. 2019.** Diversity of Dromedary Camel Coronavirus HKU23 in African Camels Revealed Multiple Recombination Events among Closely Related Betacoronaviruses of the Subgenus Embecovirus. *Journal of virology*, 93(23), e01236-19, P : 18.
- **Song H. D., Tu C. C., Zhang G. W., Wang S. Y., Zheng K., Lei L. C., Chen Q. X., Gao Y. W., Zhou H. Q., Xiang H., Zheng H. J., Chern S. W., Cheng F., Pan C. M., Xuan H., Chen S. J., Luo H. M., Zhou D. H., Liu Y. F., He J. F., Zhao G. P. 2005.** Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. *Proceedings of the National Academy of Sciences of the United States of America*, 102(7), 2430–2435, P : 6.
- **Stecher G., Tamura K., Kumar S. 2020.** Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Molecular biology and evolution*, 37(4), 1237–1239, P : 3.
- **Stott C. J., Temeeyasen G., Tripipat T., Kaewprommal P., Tantituvanont A., Piriyaopngsa J., Nilubol D. 2017.** Evolutionary and epidemiological analyses based on spike genes of porcine epidemic diarrhea virus circulating in Thailand in 2008-2015. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases*, 50, P :

- **Sun W., Wang L., Huang H., Wang W., Cao L., Zhang J., Zheng M., Lu, H. 2020.** Genetic characterization and phylogenetic analysis of porcine deltacoronavirus (PDCoV) in Shandong Province, China. *Virus research*, 278, 197869, P :9.
- **Tamura K., Stecher G., Kumar S. 2021.** MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular biology and evolution*, 38(7), 3022–3027, P : 6.
- **Tao Y., Shi M., Chommanard C., Queen K., Zhang J., Markotter W., Kuzmin I. V., Holmes E. C., Tong S. 2017.** Surveillance of Bat Coronaviruses in Kenya Identifies Relatives of Human Coronaviruses NL63 and 229E and Their Recombination History. *Journal of virology*, 91(5), e01953-16, P : 16.
- **Tao Q., Tamura K., U Battistuzzi F., Kumar S. 2019.** A Machine Learning Method for Detecting Autocorrelation of Evolutionary Rates in Large Phylogenies. *Molecular biology and evolution*, 36(4), P : 14.
- **Terada Y., Shiozaki Y., Shimoda H., Mahmoud H., Noguchi K., Nagao Y., Shimojima M., Iwata H., Mizuno T., Okuda M., Morimoto M., Hayashi T., Tanaka Y., Mochizuki M., Maeda K. 2012.** Feline infectious peritonitis virus with a large deletion in the 5'-terminal region of the spike gene retains its virulence for cats. *The Journal of general virology*, 93(Pt 9), 1930–1934, P : 5.
- **Tian P. F., Jin Y. L., Xing G., Qv L. L., Huang Y. W., Zhou J. Y. 2014.** Evidence of recombinant strains of porcine epidemic diarrhea virus, United States, 2013. *Emerging infectious diseases*, 20(10), 1735–1738, P : 4.
- **Tratner I. 2003.** SRAS: 1.Le virus. *M/S : médecine sciences*, 19(8-9), 885–891, P : 8.
- **Tuanthap S., Chiteafea N., Rattanasrisomporn J., Choowongkomon K. 2021.** Comparative sequence analysis of the accessory and nucleocapsid genes of feline coronavirus strains isolated from cats diagnosed with effusive feline infectious peritonitis. *Archives of virology*, 166(10), 2779–2787, P : 9.
- **Turlewicz-Podbielska H., Pomorska-Mól M. 2021.** Porcine Coronaviruses: Overview of the State of the Art. *Virologica Sinica*, 36(5), 833–851, P : 19.
- **Van der Hoek L., Pyrc K., Jebbink M. F., Vermeulen-Oost W., Berkhout R. J., Wolthers K. C., Wertheim-van Dillen P. M., Kaandorp J., Spaargaren J., Berkhout B. 2004.** Identification of a new human coronavirus. *Nature medicine*, 10(4), 368–373, P : 6.
- **Van der Hoek L. 2007.** Human coronaviruses: what do they cause?. *Antiviral therapy*, 12(4_part_2), 651-658, P : 8.

- **Van Diep N., Sueyoshi M., Norimine J., Hirai T., Myint O., Teh A., Izzati U. Z., Fuke N., Yamaguchi R. 2018.** Molecular characterization of US-like and Asian non-S INDEL strains of porcine epidemic diarrhea virus (PEDV) that circulated in Japan during 2013-2016 and PEDVs collected from recurrent outbreaks. *BMC veterinary research*, 14(1), 96, P : 12.
- **Van Nguyen D., Terada Y., Minami S., Yonemitsu K., Nagata N., LE T. D., Kuwata R., Shimoda H., Maeda K. 2017.** Characterization of canine coronavirus spread among domestic dogs in Vietnam. *The Journal of veterinary medical science*, 79(2), 343–349, P : 7.
- **Velavan T. P., Meyer C. G. 2020.** The COVID-19 epidemic. *Tropical medicine & international health*, 25(3), 278, P : 3.
- **Wassenaar T. M., Zou Y. 2020.** 2019_nCoV/SARS-CoV-2: rapid classification of betacoronaviruses and identification of Traditional Chinese Medicine as potential origin of zoonotic coronaviruses. *Letters in applied microbiology*, 70(5), 342–348, P : 7.
- **Whittaker G. R., André N. M., Millet J. K. 2018.** Improving Virus Taxonomy by Recontextualizing Sequence-Based Classification with Biologically Relevant Data: the Case of the *Alphacoronavirus 1* Species. *mSphere*, 3(1), e00463-17, P : 8.
- **Woo P. C., Lau S. K., Chu C. M., Chan K. H., Tsoi H. W., Huang Y., Wong B. H., Poon R. W., Cai J. J., Luk W. K., Poon L. L., Wong S. S., Guan Y., Peiri J. S., Yuen K. Y. 2005.** Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *Journal of virology*, 79(2), 884–895, P : 12.
- **Woo P. C., Lau S. K., Lam C. S., Lau C. C., Tsang A. K., Lau J. H., Bai R., Teng J. L., Tsang C. C., Wang M., Zheng B. J., Chan K. H., Yuen, K. Y. 2012.** Discovery of seven novel Mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of alphacoronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus. *Journal of virology*, 86(7), 3995–4008, P : 14.
- **Woo P. C., Lau S. K., Wernery U., Wong E. Y., Tsang A. K., Johnson B., Yip C. C., Lau C. C., Sivakumar S., Cai J. P., Fan R. Y., Chan K. H., Mareena R., Yuen K. Y. 2014.** Novel betacoronavirus in dromedaries of the Middle East, 2013. *Emerging infectious diseases*, 20(4), 560–572, P : 13.

- **Wu Z., Lu L., Du J., Yang L., Ren X., Liu B., Jiang J., Yang J., Dong J., Sun L., Zhu Y., Li Y., Zheng D., Zhang C., Su H., Zheng Y., Zhou H., Zhu G., Li H., Chmura A., ... Jin, Q. 2018.** Comparative analysis of rodent and small mammal viromes to better understand the wildlife origin of emerging infectious diseases. *Microbiome*, 6(1), 178, P : 14.
- **Yang D., Ju H., Wang J., Bai Y., Ge F., Liu J., ... Zhu, J. 2017.** Genome sequencing and analysis of a porcine delta coronavirus from eastern China. *Eur. J. Exp. Biol*, 7, 25. P : 7.
- **Yoo D., Deregts D. 2001.** A single amino acid change within antigenic domain II of the spike protein of bovine coronavirus confers resistance to virus neutralization. *Clinical and diagnostic laboratory immunology*, 8(2), 297–302, P : 6.
- **Zhai S. L., Wei W. K., Li X. P., Wen X. H., Zhou X., Zhang H., Lv D. H., Li F., Wang D. 2016.** Occurrence and sequence analysis of porcine deltacoronaviruses in southern China. *Virology journal*, 13, 136, P : 11.
- **Zhang Q., Xiang R., Huo, S., Zhou Y., Jiang, S., Wang Q., Yu, F. 2021.** Molecular mechanism of interaction between SARS-CoV-2 and host cells and interventional therapy. *Signal transduction and targeted therapy*, 6(1), 233, P : 19.
- **Zhang X., Hasoksuz M., Spiro D., Halpin R., Wang S., Stollar S., Janies D., Hadya N., Tang Y., Ghedin E., Saif L. 2007.** Complete genomic sequences, a key residue in the spike protein and deletions in nonstructural protein 3b of US strains of the virulent and attenuated coronaviruses, transmissible gastroenteritis virus and porcine respiratory coronavirus. *Virology*, 358(2), 424–435, P : 12.