

République Algérienne Démocratique et Populaire.
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique.



UNIVERSITE de TLEMCEN
Faculté des Sciences de la Nature et de la Vie et Sciences de la Terre et de l'Univers
Laboratoire de physiologie, physiopathologie et biochimie de la nutrition (PpBioNut)
Département : Biologie

MEMOIRE

Présenté par

BENMADANI Nora

ET

HOUARI Mohammed Charaf Eddine.

En vue de l'obtention du

Diplôme de MASTER Académique.

Spécialité: Génétique.

Thème :

Etude phylogénétique du SARS-COV 2 et comparaison des séquences et identification des positions des mutations.

Soutenu le 10/10/2020, devant le jury composé de :

Qualité	Nom	Grade	Université
Président :	GAOUAR S.B.S	Professeur	Abou-BekrBelkaid. Tlemcen
Encadreur :	TRIQUI C.	MCB	Abou-BekrBelkaid. Tlemcen
Examineur :	BRAHAMI N.	MCA	Abou-BekrBelkaid. Tlemcen

Année universitaire : 2019-2020.

Remerciements

Tout d'abord, Il est primordial de remercier « ALLAH » le Tout-Puissant pour nous avoir donnés la force, la volonté et la patience pour achever ce modeste travail.

*On exprime d'abord nous profonds remerciements, nos vives reconnaissances et nos sincère gratitude à **Mme TRIQUI Chahinez** pour nous avoir permis de travailler sur un sujet passionnant, sa patience, sa disponibilité et surtout ses judicieux conseils, qui ont contribué à alimenter notre réflexion.*

*Nous adressons nos sincères remerciements à **Mr GOUAR Semir Suhel** pour nous avoir honoré en acceptant de présider le jury de ce mémoire.*

*Nous remercions également **Mme BRAHAMI Nabila** de nous fait l'honneur d'accepter d'examiner ce travail.*

Nous exprimons nos reconnaissances et nos remerciements à nos profs qui nous ont suivi pendant notre parcours universitaire.

*Nous remercions également nos amis **Mr HADDAM Youssouf Hadi**, **Mr BENBACHIR Belkacem**, **Mme HAMEL Wissem** et **Mme ABDELBARI Hanaa** pour leur soutien.*

Dédicaces

A nos chers parents, pour tous leurs sacrifices, leur amour, leurs tendresses, leurs soutiens et leurs prières tout au long de mes études,

Nos aimables sœurs et nos frères soutien moral,

A toute nos familles pour leurs soutiens tout au long de mon parcours universitaire

A Tous nos collègues de promotion de biochimie appliquée, pour leurs aides et disponibilités.

A Nos amis qui nous ont permis d'oublier les moments de stress et de découragement.

Tous ceux qui nous sont chers.

Que ce travail soit l'accomplissement de vos vœux tant allégués, et le fruit de votre soutien infailible,

Merci d'être toujours là pour nous.

TABLE DES MATIERS

Remercîments, Dédicaces.

Liste des tableaux, figures.

Introduction générale

Première partie : Revue bibliographique.

I. Chapitre I : Généralités sur le COVID-19.....	05
I.1 Introduction.....	05
I.2 L’histoire des coronavirus :.....	05
I.3 L’origine de SARS-COV-2 et COVID-19	07
I.4 Structure et génome.....	07
I.5 Epidémiologie moléculaire.....	08
I.6 Transmission.....	08
I.7 Contamination, infection cellulaire et cycle de réplication.....	09
I.8 Réponse de l’hôte et considérations pour le développement de vaccins..	10
I.9 Les clades du SARS-COV 2 et mutations.....	11
II. Chapitre II : BIOINFORMATIQUE ET PHYLOGÉNIE.....	14
II.1 Les banques de données biologiques.....	15
II.2 Le séquençage d’ADN	16
II.3 Alignement multiple de séquences.....	16
II.4 Phylogénie	16
II.5 Définition et terminologie.....	17
II.6 Construction d’arbre phylogénétique.....	17

Deuxième partie : Partie expérimentale.

I. Chapitre I : Matériels et méthodes.....	20
I.1 Introduction.....	20
I.2 Bases de données et programmes.....	20
I.2.A GisaiD.....	20
I.2.B Mega 7.....	20
I.2.C Gblocks.....	20
I.2.D GenieGen.....	21
I.3 Les étapes d'analyse	21
I.3.A Téléchargement et organisation des données.....	21
I.3.B Alignement et Filtration des données.....	22
I.3.C Construction de l'arbre phylogénétique.....	23
I.3.D Comparaison des séquences.....	25
II. Chapitre II : Résultats et interprétations.....	27
II.1 Phylogénie du COVID-19.....	27
II.1.A Résultats de la première étape	27
II.1.B Résultats de la deuxième étape.....	33
II.1.C Résultat de la dernière étape.....	35
II.2 Identifications des mutations.....	37
II.3 Comparaison des séquences et identifications des mutations.....	56
II.3.A Comparaison du Groupe I des pays :.....	56
II.3.B Comparaison du Groupe II des pays :.....	58
Conclusion générale.....	60
Références bibliographiques.....	61

Liste des tableaux

Tableau 01: les différentes mutations des séquences étudiés a la construction d'arbre phylogénétique..... 37.

Tableau 02 : les différentes mutations résultantes de la comparaison des séquences du Groupe I des pays..... 57.

Tableau 03: les différentes mutations résultantes de la comparaison des séquences du Groupe II des pays..... 58.

Liste des figures

Figure 1 : Classification des coronavirus.....	06
Figure 2 : Phylogénie, structure et réplication du SARS-COV-2.....	10
Figure 3 : les différents clades du covid-19.....	12
Figure 4 : la distribution des clades du covid-19 au monde.....	13
Figure 5 : les objets fondamentaux d'un arbre phylogénétique.....	17
Figure 6 : téléchargement des données.....	22
FIGURE 7 : Alignement des séquences.....	22
FIGURE 8 : Modification et filtrage des séquences alignés.....	23
FIGURE 9 : le tableau représentant du meilleur caractère comparatif.....	23
FIGURE 10 : Construction de l'arbre phylogénétique.....	24
FIGURE 11 : Alignement des séquences à comparer.....	25
FIGURE 12 : Comparaison des séquences.....	26
FIGURE 13 : L'arbre phylogénétique du premier groupe (USA).....	28
FIGURE 14 : L'arbre phylogénétique du deuxième groupe (ANG).....	29
FIGURE 15 : L'arbre phylogénétique du troisième groupe (Au).....	30
FIGURE 16 : L'arbre phylogénétique du quatrième groupe (FE).....	31
FIGURE 17 : L'arbre phylogénétique du cinquième groupe (ALL).....	32
FIGURE 18 : L'arbre phylogénétique des trois groupes 2(ANG),4(FE) et 5(ALL).....	33
FIGURE 19 : L'arbre phylogénétique du groupe 1 (USA) avec le groupe 3 (AU).....	34
FIGURE 20 : L'arbre phylogénétique final.....	36

INTRODUCTION GENERALE

INTRODUCTION GENERALE

INTRODUCTION GENERALE

Introduction générale

Les épidémies sont des scénarios en évolution par nature, et La pandémie de coronavirus 2019 (COVID-19) n'est pas une exception à la règle, elle est la troisième pandémie causée par le coronavirus les 20 dernières années ; dépassant toute autre épidémie causée par les bêtacoronavirus dans l'histoire humaine (**Qi Liuet al., 2020**).

Le coronavirus 2 du syndrome respiratoire aigu sévère (SARSCoV-2), un nouveau virus de l'ARN de la famille des Coronaviridae, a provoqué une épidémie de coronavirus en 2019 (COVID-19) fin décembre 2019, et s'est rapidement propagée à plus de 214 pays et régions depuis lors.

Le génome du SRAS-CoV-2 a fait l'objet de plusieurs études approfondies pour mieux le comprendre. À l'heure actuelle, des milliers de séquences de virus isolés dans plusieurs pays sont partagées dans la base de données du GISAID. L'examen de ces séquences permet une compréhension détaillée de son contenu génétique, de son évolution et la diversité à l'intérieur et entre les communautés locales. Les informations obtenues sont également essentielles dans la conception du vaccin efficace.

L'objectif de la présente étude était d'étudier et de suivre les différents clades de SRAS-CoV-2 dans 13 pays en utilisant des méthodes moléculaires et phylogénétiques ou plusieurs logiciels bio-informatiques populaires ont été utilisés pour l'alignement et l'analyse des séquences et la construction du l'arbre phylogénique.

Nous avons examiné les séquences de sars-cov-2 dans le but de cartographier les variations structurelles de ce nouveau génome du coronavirus et les modèles de sélection, le cas échéant.

Des séquences de génome de SRAS-CoV-2 humains de 13 pays ont été téléchargées à partir de la base de données GISAID à partir du mois de janvier jusqu'au mois d'aout afin de les étudier et les analyser en construisant l'arbre phylogénique et en comparant les séquences du clade G avec celles du clade O.

Dans la première partie nous allons présenter deux chapitres. Dans le chapitre I on va essayer de donner le maximum de connaissances sur COVID-19, dont on va discuter de ses origines et influences sur l'organisme et sa propagation. Et le chapitre II est divisé en deux parties ; la première partie consiste à décrire les outils bio-informatiques et les banques de séquences

INTRODUCTION GENERALE

biologiques existantes et la seconde partie est réservée pour décrire la synthèse phylogénétique.

La deuxième partie est divisé en deux chapitres, le premier représente les programmes et analyses réalisés, et les différentes interfaces implémentées, et le deuxième détaille les résultats et les interprétations des résultats et discussion.

Enfin on terminera par une conclusion générale.

PREMIERE PARTIE :
CHAPITRE I

CHAPITRE I : GENERALITE SUR COVID-19

I. Chapitre I : Généralités sur le COVID-19 :

I.1 Introduction

Le SARS-COV-2 est un virus de la famille des coronavirus semblable au SARS-COV et au MERS- COV, d'origine zoonotique identifié pour la première fois en décembre 2019 à Wuhan, capitale de la province Chinoise du Hubei. Le virus s'est depuis propagé à l'échelle mondiale, ce qui a entraîné la pandémie actuellement en cours du COVID-19 (**Dorpa LV et al., 2020**). Physiologiquement l'infection par le SARS-CoV-2, peut conduire à une réaction immunitaire inadaptée et à une coagulopathie responsable d'un véritable sepsis viral (**Bonnya V et al., 2020**). Ce virus a une épidémiologie différente, et il se comporte comme les autres virus coronavirus qui provoquent des symptômes ressemblant à un rhume. Ils se répliquent principalement au niveau des voies respiratoires supérieures sans provoquer de brusques symptômes sévères. Les signes cliniques associés à l'infection par le SARS-CoV-2 comprennent, la fièvre, la toux sèche, les maux de tête et dyspnée, et la pneumonie qui peut évoluer vers l'insuffisance respiratoire et la mort (**Jaimes A et al., 2020**). La période d'incubation moyenne du SRAS-COV-2 est de 6,4 jours dans une fourchette de 2,1 à 11,1 jours et le mode principal de l'infection est une transmission interhumaine par contact étroit (**Elayadeth-Meethal M et al., 2020**). Il est noté que ces symptômes peuvent se développer en pneumonie aiguë sévère avec syndrome de détresse respiratoire aiguë (SDRA) (**Elayadeth-Meethal M et al., 2020**), d'où la dangerosité de cette affection.

Le taux de mortalité du SARS-COV-2 est plutôt faible (2,8%) par rapport au SARS-COV (9,14%) et au MERS-COV (34,4%), mais il est remarqué que le SARS-COV-2 cause plus de perturbations économiques et sanitaires au niveau mondial que le SRAS-COV et le MERS-COV en raison de ses chances de propagation rapide (**Elayadeth-Meethal M et al., 2020**).

I.2 L'histoire des coronavirus :

En 1960, le coronavirus a été caractérisé et isolé pour la première fois chez un enfant atteint d'infections des voies respiratoires (**Golzar Hossain Md et al., 2020**). Après, des souches de coronavirus comme le HCOV-OC43 et HCOV-229E ont été identifiées par les chercheurs Hamre et Procknow en 1966 et McIntosh et collaborateurs en 1967 chez des personnes présentant des symptômes d'un coup de froids. Ce nouveau groupe de virus a été nommé coronavirus à la fin des années 1960, son leur nom est tiré de la faite que sous

CHAPITRE I : GENERALITE SUR COVID-19

microscope électronique leur surface donné l'aspect d'une couronne (**Golzar Hossain Md et al., 2020**).

Les coronavirus sont des virus à ARN de la famille des Coronaviridae, la sous-famille orthocoronaviridae, qui est divisée en quatre genres : Alphacoronavirus, Betacoronavirus, Gammacoronavirus et Deltacoronavirus. Le SARS-COV, MERS-COV et SARS-CoV-2 sont tous des bêtacoronavirus, un genre qui comprend de nombreux virus qui infectent les humains, les chauves-souris et d'autres animaux domestiques et sauvages (**Jaimes A et al.,2020**). Les bêtacoronavirus présentent de nombreuses similitudes au sein de l'ORF1ab polyprotéine et la plupart des protéines structurales ; Cependant la protéine S et les protéines accessoires montrent une diversité significative (**Jaimes A et al.,2020**). MERS-COV a maintenu un génome stable depuis son émergence en 2012, contrairement à d'autres coronavirus qui évoluent facilement et peuvent subir des événements de recombinaison notables. Alphacoronavirus et Bêtacoronavirus, y compris le SARS-COV, MERS-COV, et le SARS-COV2, et d'autres coronavirus humains comme HCOV-NL63 seraient originaires de chauves-souris (**Jaimes A et al.,2020**). Alors que, les gammacoronavirus et les deltacoronavirus sont d'origine aviaire, mais sont connus pour infecter à la fois les mammifères et les espèces aviaires (**Jaimes A et al.,2020**).

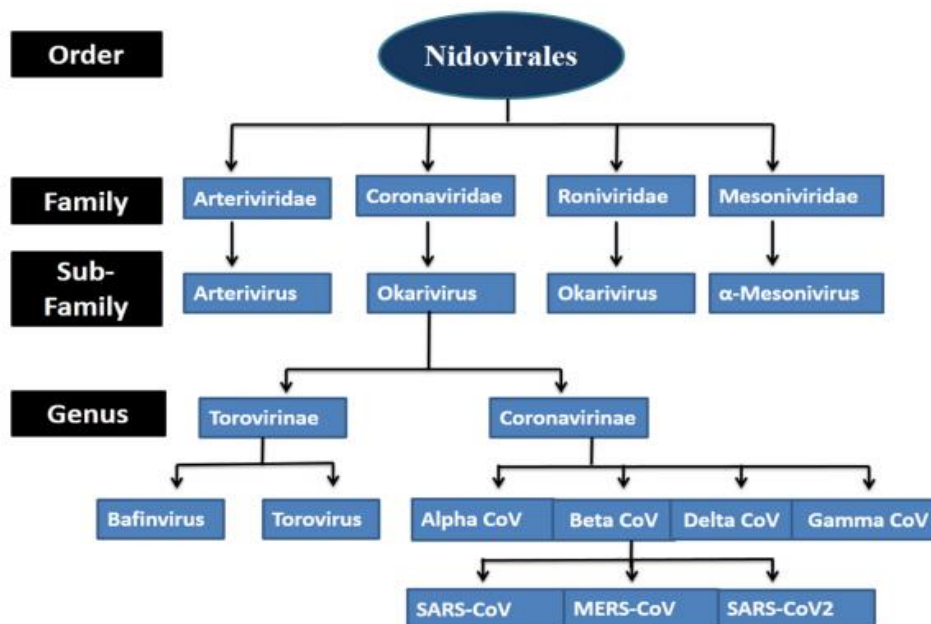


Figure 1 : Classification des coronavirus (Junejoa Y et al., 2020).

CHAPITRE I : GENERALITE SUR COVID-19

I.3 L'origine de SARS-COV-2 et COVID-19

Les virus Corona sont des habitants habituels des animaux sauvages et domestiques. Les virus deviennent infectieux et provoquent des maladies sous deux conditions : lorsque la virulence (capacité du virus à produire la maladie) des virus augmente sous des conditions appropriées telles que le changement du climat ; et / ou lorsque l'immunité de l'hôte (animaux / humains) diminue dans des conditions de stress (**Elayadeth-Meethal M et al., 2020**).

Analyse phylogénétique a révélé que le SARS-CoV-2 est un nouveau bêtacoronavirus distinct du SARS-COV et du MERS-COV. À ce jour, le virus apparenté au SARS-CoV-2 est raTG13, identifié à partir d'une chauve-souris *Rhinolophus affinis* échantillonnée dans la province du Yunnan en 2013 (**Zhou H et al., 2020**). Ce virus partage 96,1 % de l'identité nucléotidique et 92,9 % le gène S, suggérant encore une fois que les chauves-souris jouent un rôle clé en tant que réservoirs de coronavirus. Notamment, plusieurs nouveaux bêtacoronavirus liés au SARS-CoV-2 ont également été identifiés dans les pangolins (*Manis javanica*), qui ont été importés illégalement dans les provinces du Guangxi et du Guangdong, dans le sud de la Chine. Bien que ces Coronavirus de pangolin soient plus éloignés du SARS-CoV-2 que du RaTG13 dans l'ensemble du génome du virus, ils sont très semblables au SARS-CoV-2 dans le domaine de la liaison des récepteurs de la protéine S, y compris aux résidus d'acides aminés responsable de la liaison à ACE2. Il est donc possible que les pangolins jouent un rôle important dans l'écologie et l'évolution de Coronavirus. En effet, la découverte de virus dans les pangolins suggère qu'il y a encore une grande diversité de Coronavirus à échantillonner dans la faune, dont certains peuvent être directement impliqués dans l'émergence de SARS-CoV-2 (**Zhou H et al., 2020**)

I.4 Structure et génome

Les virus possèdent de l'ARN ou de l'ADN comme matériel génétique. Contrairement aux autres cellules, les virus ne contiennent ni noyau ni organites cellulaires comme les mitochondries ou les ribosomes. Le virion du SARS-CoV-2 est 50 à 200 nm de diamètre (**Suzuki Y et al., 2020**), de forme sphérique et enveloppé (**Golzar Md et al., 2020**), il a un génome de type ARN à brin positif simple brin, d'une taille de 26 à 32 kb. (**Elayadeth-Meethal M et al., 2020**).

Le génome comprend deux régions non codantes en 5' et en 3', la partie codante est divisée en plusieurs parties, les deux premiers tiers du génome sont constitués de deux

CHAPITRE I : GENERALITE SUR COVID-19

grandes régions chevauchantes, ORF1a (Open Reading Frame) et ORF1b, codant le complexe de réplication-transcription, dont le gène RNA-dépendant RNA Polymerase (RdRp) qui code l'ARN polymérase ARN dépendante. Le dernier tiers du génome code les protéines de structure qui sont : la glycoprotéine (S), protéine d'enveloppe (E), membrane (M), les protéines nucléocapsides (N) (**Golzar Hossain Md et al., 2020**), Et les protéines accessoires telles que ORF3a, ORF6, ORF7a, ORF7b, ORF8 et ORF10. Et plusieurs protéines non structurales (NSP) sont codées par le cadre de lecture (ORF) (**Sahaa I et al., 2020**).

I.5 Epidémiologie moléculaire

Au 3 juin 2020, un total de 6467229 cas de COVID-19 a été confirmé dans le monde, dont 382 766 décès (**OMS., 2020**). Le COVID-19 en tant que maladie infectieuse respiratoire aiguë, se propage principalement par voie respiratoire, via des gouttelettes, des sécrétions respiratoires émis par une personne infectée, ou par contact direct pour une faible dose. Un niveau significativement élevé de charge virale a également été observée dans la cavité nasale par rapport à la gorge, où la charge virale était la même entre personnes symptomatique et asymptomatiques (**Junejoa Y et al., 2020**). Les patients peuvent être porteurs d'infection même lors de la guérison clinique et peu de personnes peuvent agir comme candidat pour propager l'infection. La période d'incubation de cette maladie varie de 2 à 14 jours (médiane de 5 jours) (**Junejoa Y et al., 2020**).

I.6 Transmission :

Le SARS-COV-2 pénètre dans le corps par l'air expiré ou par les toux ou les éternuements. Les voyages internationaux des personnes infectées ont principalement causé la propagation de la maladie à travers les nations. Les personnes infectées portent une charge virale importante au niveau des voies respiratoires supérieures, pendant qu'elles menaient ses activités quotidiennes habituelles, ce qui a permis la propagation du virus et l'infection d'un nombre important de personnes (**Elayadeth-Meethal M et al., 2020**).

L'expansion pandémique du SARS-COV -2 est principalement due à sa transmission par des porteurs asymptomatiques (personnes qui ne présentent aucun symptôme lorsqu'ils sont porteurs du virus). Dans une étude menée à Wuhan, auprès de 277 personnes confirmées par le COVID-19, 200 d'entre eux n'avaient jamais visité le marché de Wuhan et n'étaient pas en contact étroit avec des patients gravement malades (**Elayadeth-Meethal M et al., 2020**).

CHAPITRE I : GENERALITE SUR COVID-19

Ce qui renforce l'idée que la transmission de ce virus est possible d'humain à humain, et le rend une grande menace pandémique.

Les études du mécanisme d'introduction du virus dans les cellules ont montré que l'enzyme de conversion de l'angiotensine cellulaire 2 (ACE2) est le récepteur du SARS-COOV-2, et qui est identique ou similaire chez différentes espèces d'animaux tels que les porcs, les furets, les chats, les orangs-outans, les singes (**Golzar Hossain Md et al., 2020**). En outre, une étude récente a prédit que les chiens pourraient être des hôtes secondaires au cours de l'évolution du SARS-COV-2 de la chauve-souris à l'homme. (**Golzar Hossain Md et al., 2020**). Par conséquent, il est possible de répandre le SARS-COV-2 par l'intermédiaire d'animaux domestiques. Il y a maintenant beaucoup de rapports de SARS-COV-2 cas positifs chez les chiens, les chats, les tigres, les lions et les visons. Des données expérimentales ont montré que les furets et les chats infectés par le virus SARS-COV-2 sont très susceptibles à transmettre le virus directement ou indirectement par des gouttelettes ou des voies aéroportées (**Golzar Hossain Md et al., 2020**).

I.7 Contamination, infection cellulaire et cycle de réplication

Le virus pénètre dans les cellules humaines via les récepteurs situés à la surface des cellules, et qui sont uniques pour chaque virus et cela détermine la sensibilité de divers organes dans le corps humain pour une attaque virale.

Un fois la liaison au récepteur est assurée, la protéine d'enveloppe virale se fusionne avec la membrane plasmique (la membrane externe des cellules humaines) permettant l'entrée du virus dans les cellules. Après la fusion il y'a le largage de la nucléocapside dans le cytosol de la cellule hôte, la machinerie cellulaire traduit le gène de la réplicase en deux polyprotéines (pp1a et pp1ab) clivées en nombreuses protéines indispensables au cycle viral (notamment deux protéases virales et une ARN-polymerase ARN-dépendant) s'assemblant en un large complexe de transcription et de réplication. Ce complexe permet d'une part de reproduire l'ARN viral et d'autre part, par le biais de la formation de petits brins d'ARN anti-sens appelés ARN sous-génomiques, la production de protéines de structure des nouveaux virions. Finalement les brins d'ARN synthétisés sont combinés avec la protéine N pour former la nucléocapside et l'assemblage avec les glycoprotéines d'enveloppe permet le bourgeonnement de nouvelles particules virales entraînant la maladie COVID-19 chez l'individu infecté (**Bonnya Vet al., 2020**).

CHAPITRE I : GENERALITE SUR COVID-19

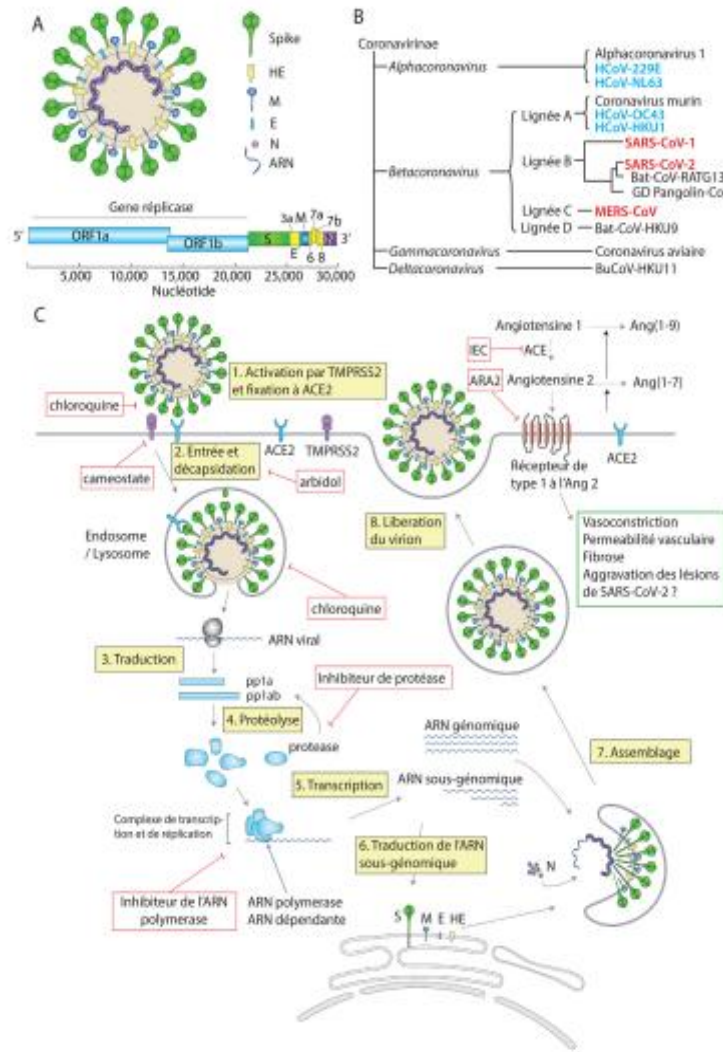


Figure 2 : Phylogénie, structure et répléation du SARS-CoV-2. A. Structure virale (Bonnya Vet al., 2020).

I.8 Réponse de l'hôte et considérations pour le développement de vaccins

Pour contenir les épidémies de COVID-19, la prévention de l'importation et l'exportation de personnes contagieuses est une intervention essentielle. Il est également important de clarifier les patients infectés par le SARS-COV-2 afin d'initier un suivi de contact efficace et réussi de la source d'infection, et par la suite les patients contagieux pourraient être mis en quarantaine pour éviter la propagation de la maladie (Elayadeth-Meethal M et al., 2020).

CHAPITRE I : GENERALITE SUR COVID-19

À l'heure actuelle, il n'existe pas de traitement antiviral spécifique pour le COVID-19, ni pour les autres Coronavirus, et les principaux traitements essayés sont de soutien. Telle que l'IFN recombinant avec le Ribavirine qui n'a que des effets limités contre l'infection par le Coronavirus. Après les épidémies du SARS-COV et du MERS-COV, de gros efforts ont été consacrés au développement de nouveaux antiviraux ciblant les protéases, les polymérases et les protéines d'entrée, cependant, aucune d'entre eux ne s'est avérée efficace dans les essais cliniques. Aussi le plasma et les anticorps obtenus à partir des patients convalescents ont été proposés pour un éventuel traitement contre ces virus (**Chen Y et al., 2020**).

Ce traitement consiste sur la transfusion de plasma de convalescence compatible ABO aux patients éligibles, les patients convalescents doivent montrer un état de guérison, ce dernier est définie comme suivant ; un statu afébrile pendant en moins 3 jours, soulagement des symptômes respiratoires, résultat négatif pour l'acide nucléique du SARS-COV-2 pour deux tests de RT-PCR consécutifs, Les donneurs doivent aussi être séronégatifs pour l'anti- VHB, le VHC et le VIH, et séropositifs pour l'anti - SARS - COV - 2. Pour la transfusion, les patients reçoivent au moins un cycle de transfusion plasmatique de convalescence compatible ABO (200 mL pour chaque cycle). Chaque transfusion est administrée sur une période de 30 minutes. (**Chen Y et al., 2020**).

Par ailleurs, les recherches thérapeutiques tentent plus de découvrir divers stratégies vaccinales, étant donné que la réponse immunitaire est identique dans le SARS, le MERS et le COVID-19, ceci donne la possibilité de concevoir une immunisation à réaction croisée contre les trois virus coronavirus humains (SARS-COV, MERS-COV et SARS-COV-2) , comme le vaccin à base d'acide désoxyribonucléique (ADN), le vaccin à vecteur viral, le vaccin sous-unitaire, un autre utilisant des particules de type virus vivant, et un vaccin à virus atténué et un vaccin à virus entier irradié au formaldéhyde/gamma sont les vaccins les plus promoteurs pour les COVID-19. Ces derniers peuvent cibler l'attachement de la protéine S qui contient un domaine de liaison au récepteur RBD qui est considéré comme un bon antigène (**Wang JT et al., 2020**).

I.9 Les clades du SARS-COV 2 et mutations

Actuellement, les études des séquences des différentes souches du virus SARS-CoV2 sont rassemblées dans la base de données GISAID qui est à la base une plateforme spécialisée dans l'étude du virus influenza. Les définitions des clades dans le GISAID sont fondées sur la distribution statistique des distances du génome dans les grappes phylogénétiques (Han A et

CHAPITRE I : GENERALITE SUR COVID-19

al., 2019), puis sur la fusion de lignées plus petites en clades principaux en fonction de variantes de marqueurs partagées. On compte les clades suivants : Clade S (contenant 439 génomes), Clade G (404), Clade V (108) et autres appelés Clades symbolisé par O (424). Le clade S contient la plupart des génomes américains. (Rodriguez-Morales J et al., 2020).

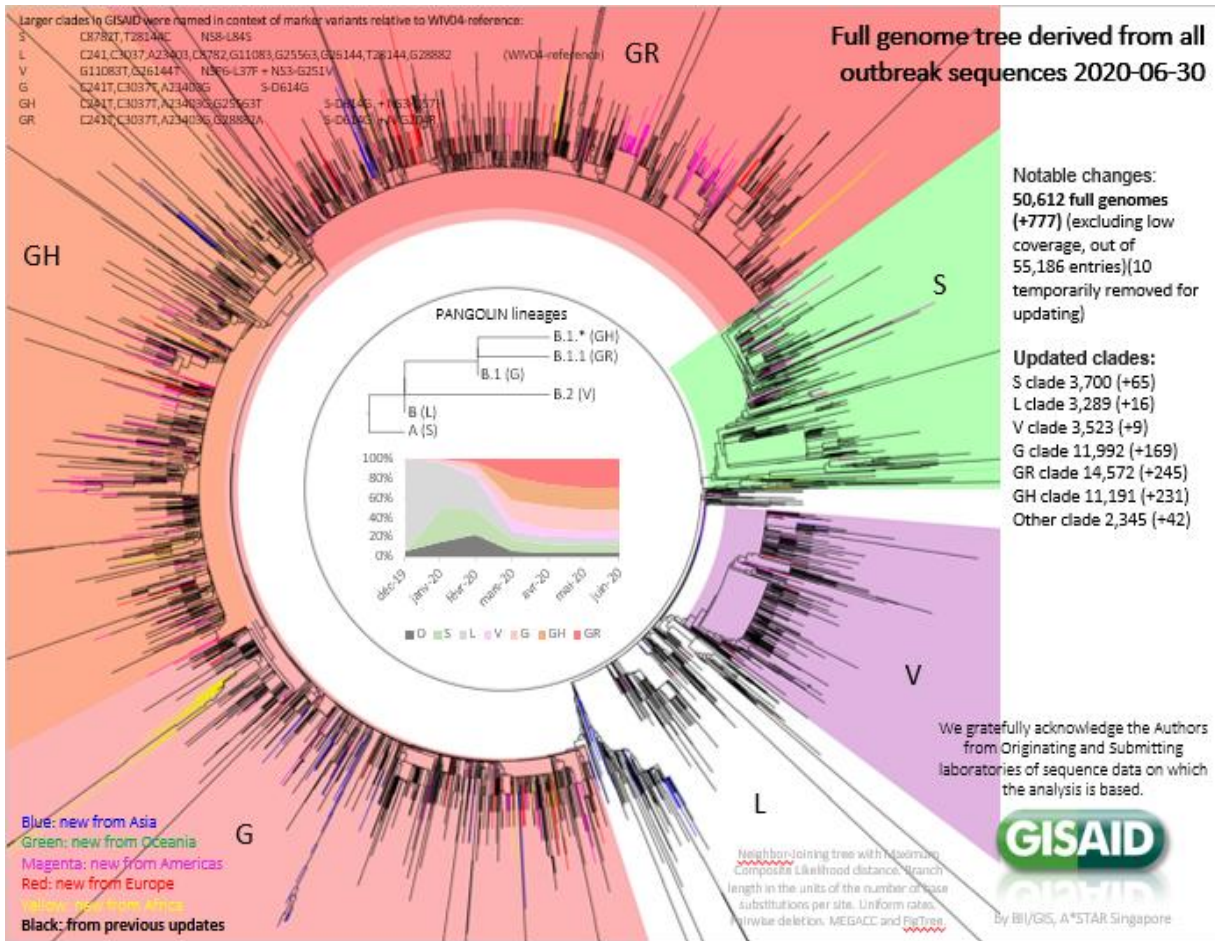


Figure 3 : les différents clades du covid-19 (GISAID., 2020).

En outre, pour la classification des clades, les analyses des génomes de la population génétique du SARS-COV-2 ont suggéré que ces virus ont évolué en deux types majeurs (désignés L et S), qui sont bien définis par deux SNP différents qui montrent presque liaison complète entre les souches virales séquencées à ce jour. Apparemment, le type L était plus répandu dans les premiers stades de l'épidémie à Wuhan, en Chine, et sa fréquence a diminué début janvier 2020. Le type S, ce type qui est le plus âgé, moins agressif et qui a subi l'évolution la plus rapide, pourraient avoir augmenté dans les fréquences due à une pression sélective relativement plus faible (Rodriguez-Morales J et al., 2020).

CHAPITRE I : GENERALITE SUR COVID-19

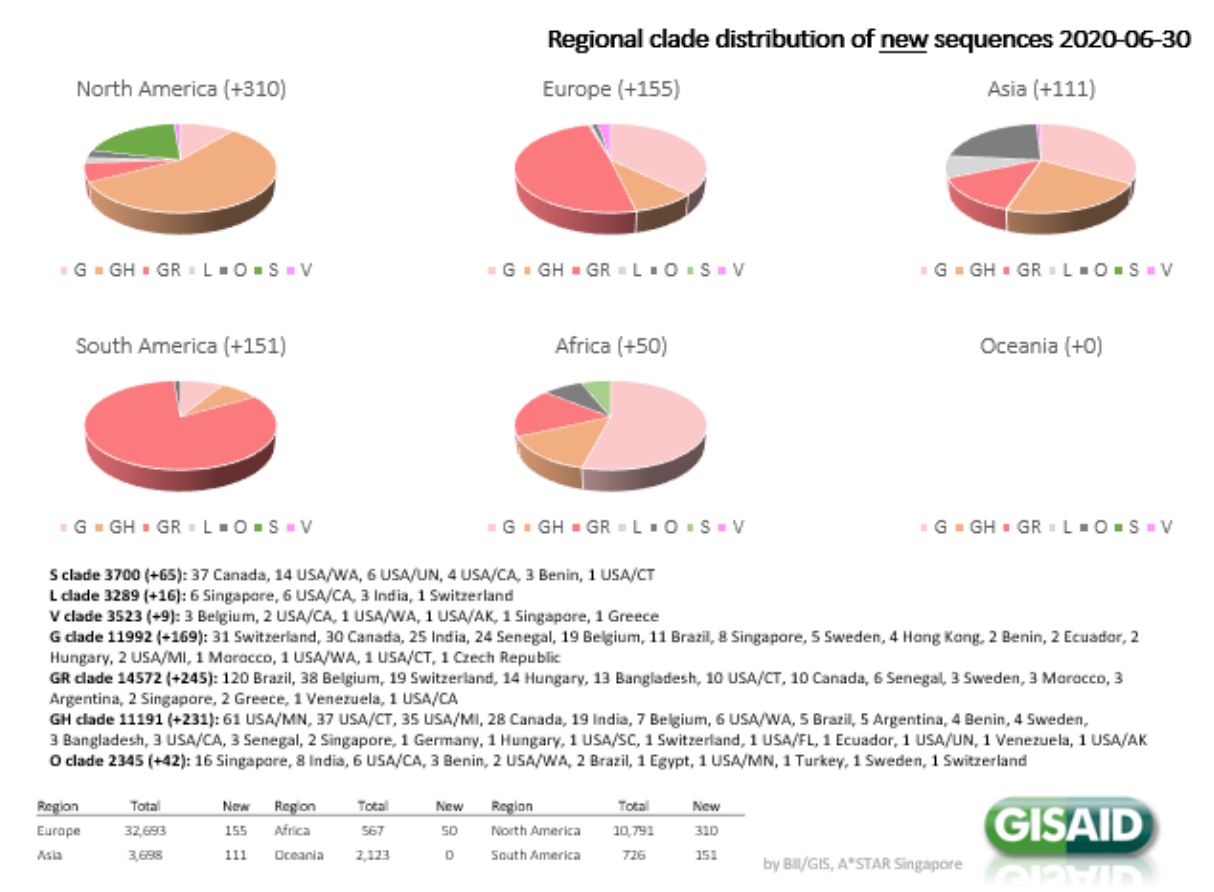


Figure 4 : la distribution des clades du covid-19 au monde (GISAID., 2020).

Le génome du SARS-COV-2 a la capacité d’obtenir des mutations rapides (Sayantan L et al., 2020). Chaque réplication du génome viral entraîne un taux de mutation à l’ARN du virus qui résulte plusieurs versions légèrement différentes du génome initial. Cela crée une population du virus aux divers génomes, connue sous le nom de quasi-espèce. Ce dernier peut contribuer à des résultats cliniques différents entre patients. Le fait qu’il existe plusieurs versions des virus rendent également la catégorisation de ces derniers difficiles (Johns Hopkins et al., 2020).

Toutefois, il n’est pas clair si les mutations génétiques du SARS-COV-2 influencent sa transmissibilité et/ou sa virulence, qui pourrait expliquer des décès plus élevés dans certains pays (Eaaswarkhanth M et al., 2020).

CHAPITRE

II

CHAPITRE II : BIO-INFORMATIQUE ET PHYLOGENIE

Chapitre II : BIOINFORMATIQUE ET PHYLOGÉNIE

Les progrès en biologie conduisent vers le développement de plusieurs nouvelles méthodes en bio-informatique, citons parmi ces méthodes :

- La comparaison et la prédiction dans l'analyse des structures macromoléculaires à partir de l'année 1950 ;
- Le séquençage depuis les années 1970 qui englobe l'alignement des séquences et la recherche des similarités dans les banques de données,
- Méthodes d'annotation et de classification fonctionnelle sur les génomes à partir des années 1990 ;
- Les premières cartes génétiques du génome humain par J.Weissenbach et D.Cohen entre l'année 1992 et l'année 1996 ;
- Les analyses multivariées sur les transcriptomes depuis 1997 ;
- L'analyse de graphes sur les interactomes presque à partir des années 2000 ;
- Le séquençage du génome humain annoncé en 14 Avril 2003 (**Thieffry D., 2002**).

Toutes ses avancées jouent un rôle important dans l'émergence du nouveau secteur scientifique, connu aujourd'hui par la « Bio-informatique », mais l'utilisation du terme est documentée pour la première fois en 1970 dans une publication de Paulien Hogeweg et Ben Hesper (Université d'Utrecht, Pays-Bas) en référence à l'étude des processus d'information dans les systèmes biotiques. La bio-informatique c'est tout un monde qui se développe jour après jour et avance très rapidement, les dernières années connaissent la découverte de plusieurs nouvelles applications conduisant vers une révolution à la biologie, l'ADN est la molécule centrale du cercle qui comprend tous types de travaux dans ce domaine, puisqu'il est le support de l'information génétique, donc il faut détailler pour expliquer comment porte-t-il cette information, et comment l'exprimer par autres moyens que l'ADN (**Maarouf C., 2015**).

II.1 Les banques de données biologiques

Les premières banques de données biologiques sont apparues au début des années 80 sous l'initiative de quelques équipes de recherches. Leur principale mission est de rendre publiques les séquences qui ont été déterminées. Les données biologiques stockées dans ces banques sont des séquences primaires d'ADN, d'ARN et de protéines. Les données peuvent être soumises et consultées par l'intermédiaire du Web. Ces séquences sont obtenues de plusieurs manières différentes. Il y a celles isolées à partir d'une cellule, déduites à partir de la séquence

CHAPITRE II : BIO-INFORMATIQUE ET PHYLOGENIE

nucléique par simple traduction (cas des séquences d'ARN ou protéines) ou encore par génie génétique (**Deriham TE., 2017**).

II.2 Le séquençage d'ADN :

Le séquençage d'ADN, c'est-à-dire le processus qui consiste à déterminer la séquence de nucléotides qui composent un fragment d'ADN, a entraîné une révolution de la recherche en biologie moléculaire. Au cours des dernières années, le développement de nouvelles techniques de biologie moléculaire a rendu possible le séquençage à grande échelle. Les améliorations qu'ont subies ces techniques ont beaucoup changé les méthodes de travail dans plusieurs domaines de recherche (**Gagnon J., 2004**).

II.3 Alignement multiple de séquences

L'alignement multiple des séquences d'ADN ou de protéines est une des techniques les plus utilisées dans l'analyse de séquence. Il est considéré parmi les problèmes les plus difficiles en bio-informatique. L'alignement multiple de séquences (Multiple Sequence Alignment) est une tâche cruciale et très importante en biologie moléculaire. Multiple Séquence Alignement offre aux biologistes un moyen pour analyser des séquences d'ADN ou de protéines et de déterminer par la suite leur degré d'homologie ou de divergence. Multiple Séquence Alignement est utilisé dans la construction des arbres phylogénétiques et identifier les motifs dans des familles de protéines, ceci permet de prédire leur aspect structurel et fonctionnel (**Deriham T E., 2017**).

II.4 Phylogénie

La phylogénie moléculaire est une discipline qui connaît un essor grandissant étant donné l'avancement spectaculaire des techniques de la biologie moléculaire et du génie génétique que l'on peut appeler maintenant biotechnologies moléculaires. Ces techniques ont permis un nombre incalculable de données biomoléculaires telles que les séquences des différents gènes et protéines (**Djekoun A et Hamidechi M. A., 2019**).

La phylogénie permet d'étudier les espèces végétales, animales et microbiennes, sur les deux plans phénotypique et génotypique, afin de les classer en fonction de leurs ressemblances et en fonction de leurs structures géniques (liens de parenté). La phylogénie étudie, en fait, les relations de parenté entre les individus et représente sous forme d'arbre le résultat de ces relations (**Djekoun A et Hamid chi M. A., 2019**).

CHAPITRE II : BIO-INFORMATIQUE ET PHYLOGENIE

II.5 Définition et terminologie

La phylogénie est un domaine d'étude associé au développement des êtres vivants. Dans un arbre phylogénétique (ou un X-arbre), l'analyse des caractères utilise le principe de l'hérédité incluant toutes les transformations aléatoires (inversion, substitution, insertion ou délétion). Du point de vue mathématique, un arbre phylogénétique est une structure de données spécifique. Ce dernier est constitué d'objets fondamentaux suivants :

- **Les sommets** : Internes et externes, appelés aussi feuilles, ils représentent les espèces vivantes ou éteintes. Les sommets internes représentent des espèces ancestrales. Des informations liées aux unités évolutives sont nécessaires pour l'induction de l'arbre.
- **Les branches ou les liens entre les nœuds** : Ce sont les segments de dépendance entre les sommets déterminant les liens de parenté selon une trace ascendante envers un ancêtre commun.
- **La racine** : Elle représente l'ancêtre commun de toutes les espèces dans l'arbre phylogénétique (Nancybadran., 2019).

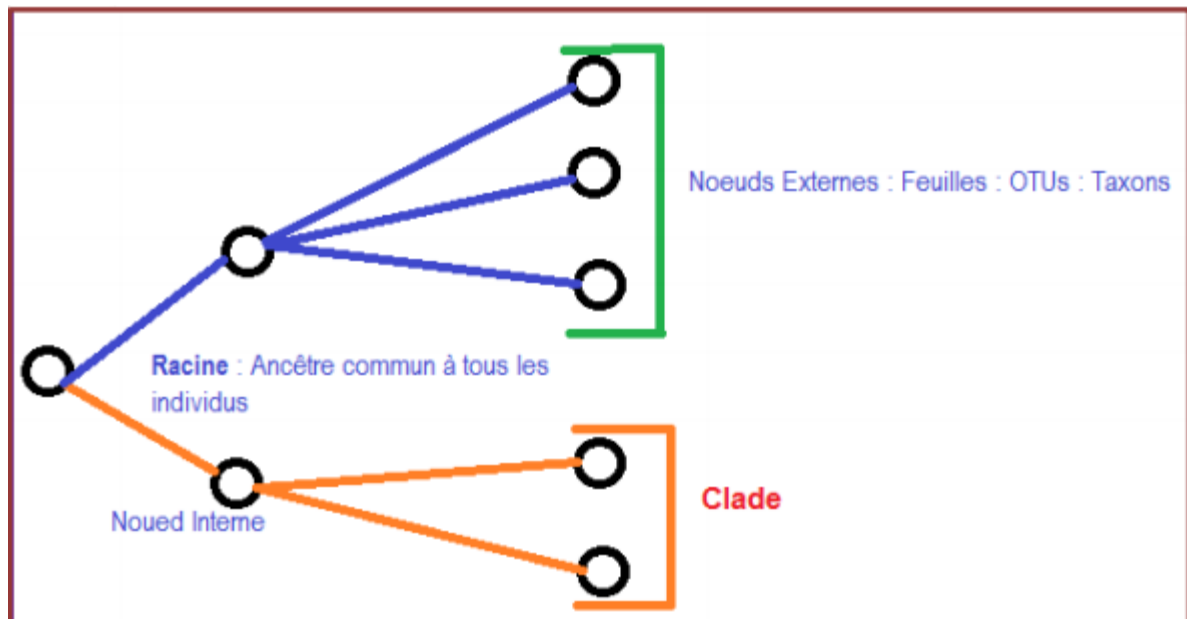


Figure 5 : les objets fondamentaux d'un arbre phylogénétique (Nancybadran., 2019).

II.6 Construction d'arbre phylogénétique

Lors d'une reconstruction d'arbres phylogénétiques, la première étape consiste à mettre en correspondance les sites des séquences de manière à pouvoir comparer ce qui est comparable. Cette étape est nommée "alignement". Une fois les séquences alignées, une méthode de reconstruction d'arbres phylogénétiques peut être appliquée pour obtenir l'arbre

CHAPITRE II : BIO-INFORMATIQUE ET PHYLOGENIE

qui reflète le mieux les données. Il existe actuellement au moins deux approches de reconstruction phylogénétiques : les méthodes basées sur les distances et celles basées sur les caractères (**Deriham TE., 2017**).

Deuxième partie :

CHAPITRE I

CHAPITRE I : MATERIEL ET METHODS

Chapitre I : Matériels et méthodes

I.1 Introduction

Afin d'étudier la croissance de diversité des lignes d'ascendances, nous allons comparer les génomes de 11 pays (Algérie, Allemande, Australie, Angleterre Espagne, France, Italie, KSA, Kuwait, Suisse, USA) pour retracer leur histoire évolutive, en identifiant également les mutations qui pourraient rendre le coronavirus plus dangereux ou plus contagieux à partir de comparaison des séquences de différents lignages des deux clades G et O.

I.2 Bases de données et programmes

I.2.A GisaID

Base de données publique de l'Initiative mondiale sur le partage de tous les Données sur la grippe (GISAID). GisaID a été initialement développé pour le partage de données génomiques sur la grippe, mais base de données SARSCoV-2 complète, dynamique et constamment mise à jour (Lucy van Dorpa et al., 2020)

I.2.B Mega 7

Le logiciel MEGA (Molecular Evolutionary Genetics Analysis) est une application de bureau conçue pour l'analyse des séquences de gènes homologues provenant de familles multigéniques ou de différentes espèces l'accent sur l'inférence des relations évolutives et des modèles de l'évolution de l'ADN et des protéines. En plus des outils d'analyse statistique des données, MEGA fournit de nombreuses facilités pratiques pour l'assemblage de données de séquence des fichiers ou des référentiels web, et il inclut des outils de présentation visuelle des résultats obtenus dans les arbres phylogénétiques interactifs et de matrices de distance évolutives. (Sudhir et al., 2008)

I.2.C Gblocks

Gblocks est un programme informatique qui élimine les positions mal alignées et les régions divergentes d'un alignement des séquences d'ADN ou de protéines. Ces positions peuvent ne pas être homologues ou avoir été saturées par de multiples substitutions et il est commandé de les éliminer avant l'analyse phylogénétique. Gblocks sélectionne les blocs de la même manière que ce qui se fait habituellement à la main, mais en suivant un ensemble de

CHAPITRE I : MATERIEL ET METHODS

conditions reproductibles. Les blocs sélectionnés doivent satisfaire à certaines exigences en ce qui concerne l'absence de grands segments de positions non conservées, l'absence ou la faible densité des positions d'écart et la conservation élevée des positions de flanc, ce qui rend l'alignement final plus approprié pour l'analyse phylogénétique. Gblocks produit plusieurs fichiers pour visualiser les blocs sélectionnés. L'utilisation d'un programme tel que Gblocks réduit la nécessité d'éditer manuellement plusieurs alignements, rend possible l'automatisation de l'analyse phylogénétique des grands ensembles de données et, enfin, facilite la reproduction des alignements et l'analyse phylogénétique subséquente par d'autres chercheurs. (Talavera, G et al., 2007).

I.2.D GenieGen

GENIEGEN est un logiciel permettant de faire des analyses de séquences d'ADN et d'ARN et de protéines afin de découvrir l'expression de l'information génétique, les liens entre génotype et phénotype, le polymorphisme des gènes, les familles multigéniques, les prévisions en génétique humaine. Ce logiciel est donc une banque de données sur de nombreuses séquences nucléiques et peptidiques que l'on peut modifier en chargeant de nouvelles séquences dans sa banque de données personnelles. (Marzin-Janvier, P, 2015)

I.3 Les étapes d'analyse

I.3.A Téléchargement et organisation des données

Dans ce travail, nous avons utilisé 298 séquences de génome de SRAS-CoV-2 humains de 11 pays (Algérie, Allemande, Australie, Angleterre, Espagne, France, Italie, KSA, Kuwait, Suisse, USA), dont nous avons pris une séquence de chaque lignage de chaque clade qui ont été téléchargées à partir de la base de données GISAID depuis le début de la pandémie jusqu'à la fin du mois d'août sous format Fasta (Figure 6). C'est également signalé qu'un tableau présentant les mutations des séquences téléchargées dans la section des résultats.

Afin de faciliter l'interprétation des résultats, nous avons renommé l'ID des séquences téléchargées en utilisant une nomenclature bien spécifique et objectif qui est comme suivant :

PAY.CLADE.LIGNAGE-N° DE SEQUENCE

Exemple : ALG.G.H.B.1-1, AU.L.A.2-1...

CHAPITRE I : MATERIEL ET METHODS

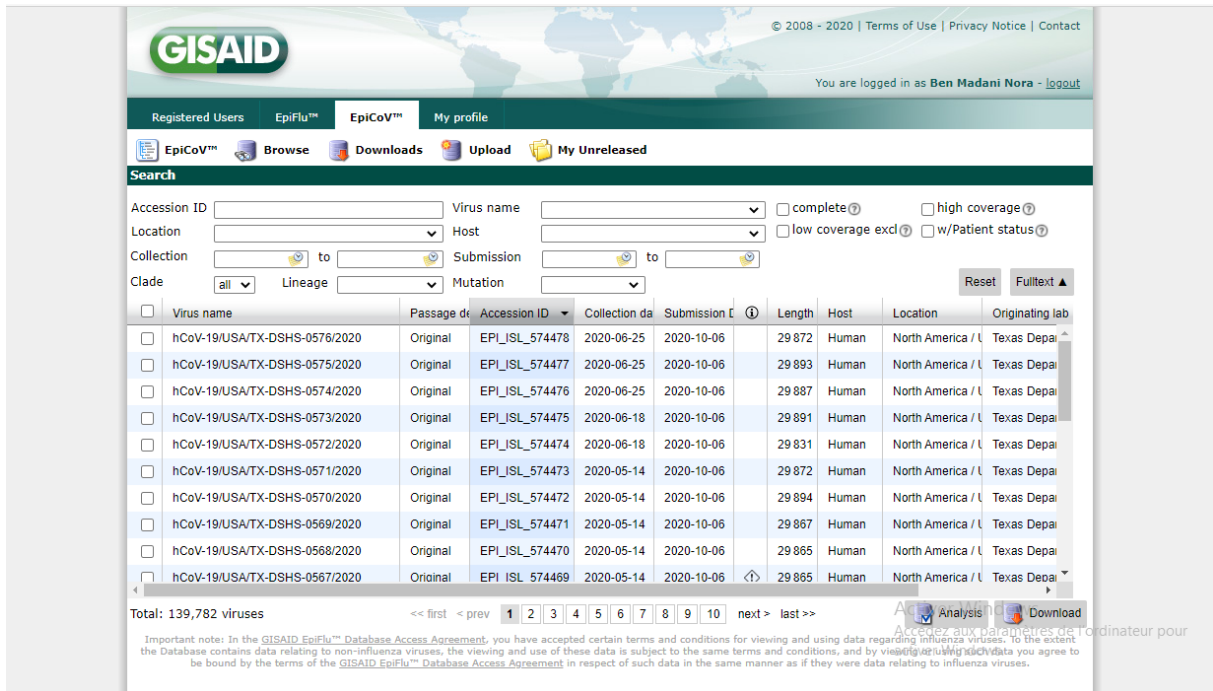
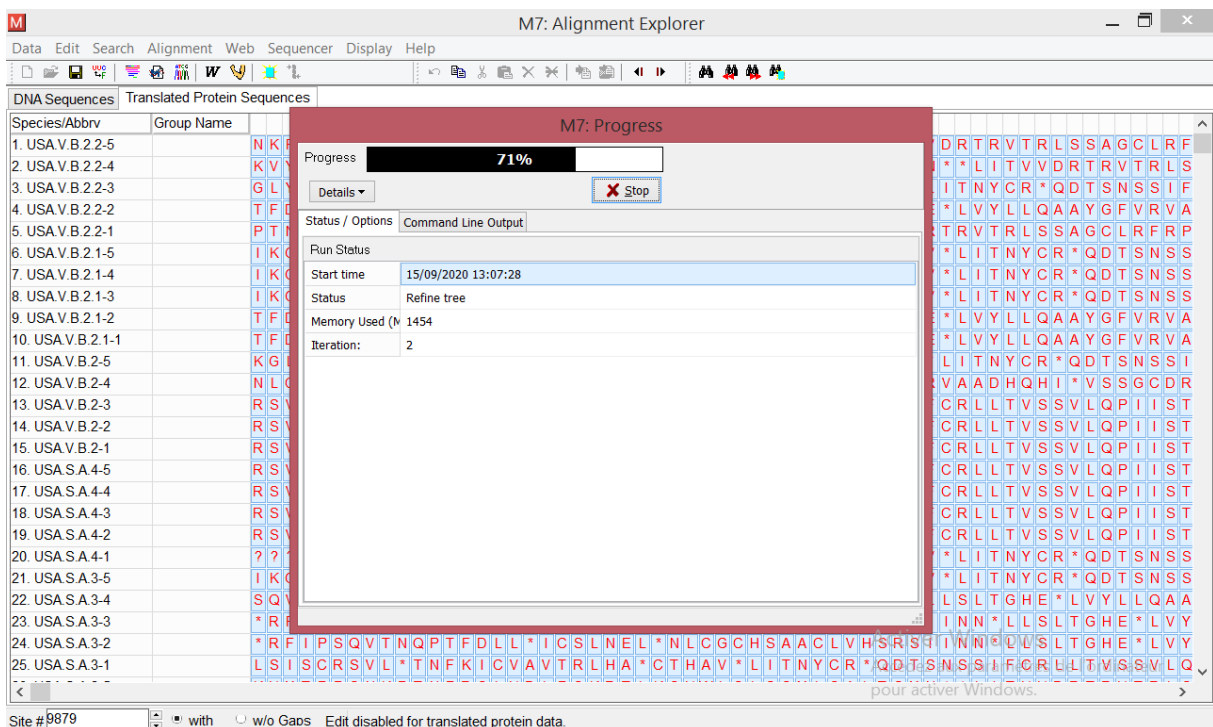


Figure 6 : téléchargement des données

I.3.B Alignement et Filtration des données :

Les assemblages qui ont une longueur moyenne de 29 831 pb ont été alignés à l'aide de logiciel Mega 7.0.26 (Figure 7).



CHAPITRE I : MATERIEL ET METHODS

FIGURE 7 : Alignement des séquences

La modification et le filtrage des séquences alignés a été assuré à l'aide de l'outil Gblocks (Figure 8).

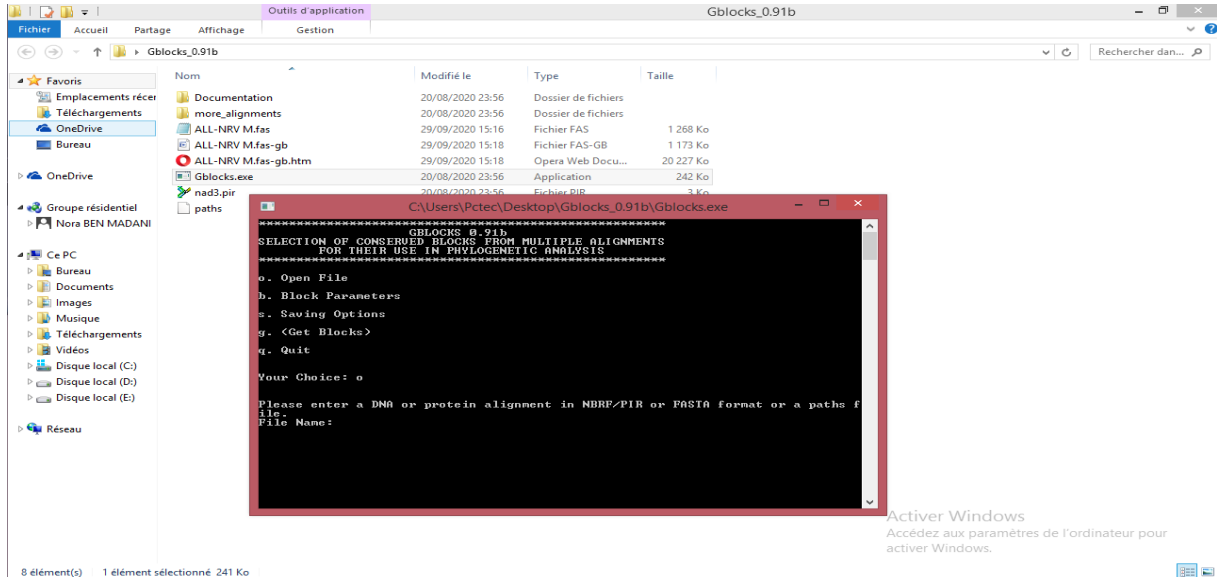
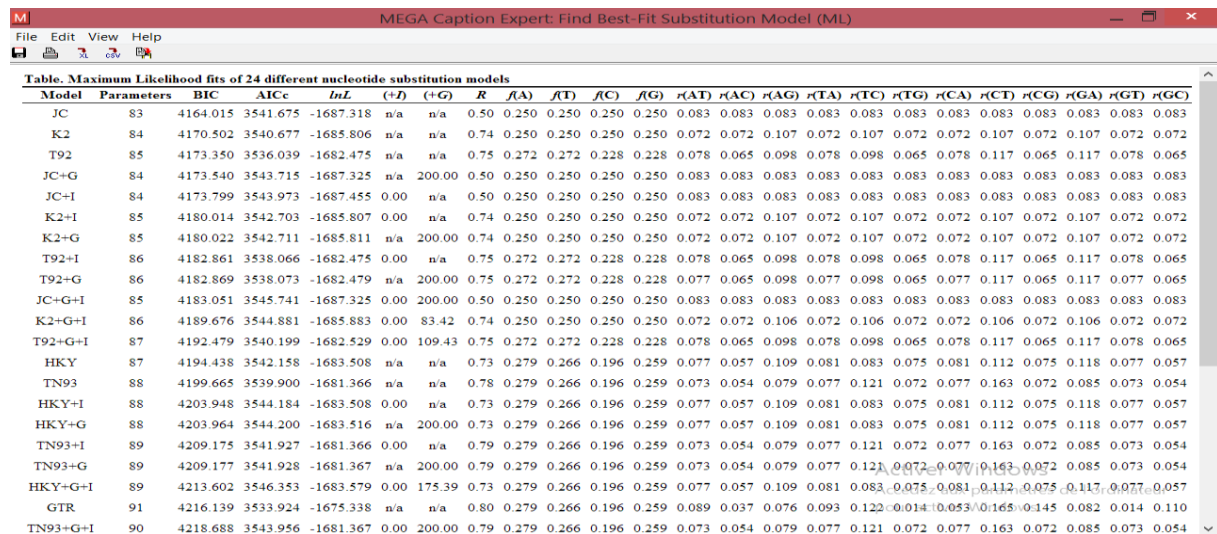


FIGURE 8 : Modification et filtrage des séquences alignés

I.3.C Construction de l'arbre phylogénétique

L'arbre phylogénétique de notre étude a été réalisé à l'aide des deux versions de logiciel Mega (Mega 7.0.26 et Mega X) après la détermination du meilleur caractère de comparaison entre les séquences alignées en utilisant l'option calculer le Models (Figure 9 et 10).



Model	Parameters	BIC	AICc	lnL	(+D)	(+G)	R	A	T	C	G	AT	AC	AG	TA	TC	TG	CA	CT	CG	GA	GT	GC
JC	83	4164.015	3541.675	-1687.318	n/a	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
K2	84	4170.502	3540.677	-1685.806	n/a	n/a	0.74	0.250	0.250	0.250	0.250	0.072	0.107	0.072	0.107	0.072	0.107	0.072	0.107	0.072	0.107	0.072	0.107
T92	85	4173.350	3536.039	-1682.475	n/a	n/a	0.75	0.272	0.272	0.228	0.228	0.078	0.065	0.098	0.078	0.098	0.065	0.078	0.117	0.065	0.117	0.078	0.065
JC+G	84	4173.540	3543.715	-1687.325	n/a	200.00	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC+I	84	4173.799	3543.973	-1687.455	0.00	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
K2+I	85	4180.014	3542.703	-1685.807	0.00	n/a	0.74	0.250	0.250	0.250	0.250	0.072	0.107	0.072	0.107	0.072	0.107	0.072	0.107	0.072	0.107	0.072	0.107
K2+G	85	4180.022	3542.711	-1685.811	n/a	200.00	0.74	0.250	0.250	0.250	0.250	0.072	0.107	0.072	0.107	0.072	0.107	0.072	0.107	0.072	0.107	0.072	0.107
T92+I	86	4182.861	3538.066	-1682.475	0.00	n/a	0.75	0.272	0.272	0.228	0.228	0.078	0.065	0.098	0.078	0.098	0.065	0.078	0.117	0.065	0.117	0.078	0.065
T92+G	86	4182.869	3538.073	-1682.479	n/a	200.00	0.75	0.272	0.272	0.228	0.228	0.077	0.065	0.098	0.077	0.098	0.065	0.077	0.117	0.065	0.117	0.077	0.065
JC+G+I	85	4183.051	3545.741	-1687.325	0.00	200.00	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
K2+G+I	86	4189.676	3544.881	-1685.883	0.00	83.42	0.74	0.250	0.250	0.250	0.250	0.072	0.106	0.072	0.106	0.072	0.106	0.072	0.106	0.072	0.106	0.072	0.106
T92+G+I	87	4192.479	3540.199	-1682.529	0.00	109.43	0.75	0.272	0.272	0.228	0.228	0.078	0.065	0.098	0.078	0.098	0.065	0.078	0.117	0.065	0.117	0.078	0.065
HKY	87	4194.438	3542.158	-1683.508	n/a	n/a	0.73	0.279	0.266	0.196	0.259	0.077	0.057	0.109	0.081	0.083	0.075	0.081	0.112	0.075	0.118	0.077	0.057
TN93	88	4199.665	3539.900	-1681.366	n/a	n/a	0.78	0.279	0.266	0.196	0.259	0.073	0.054	0.079	0.077	0.121	0.072	0.077	0.163	0.072	0.085	0.073	0.054
HKY+I	88	4203.948	3544.184	-1683.508	0.00	n/a	0.73	0.279	0.266	0.196	0.259	0.077	0.057	0.109	0.081	0.083	0.075	0.081	0.112	0.075	0.118	0.077	0.057
HKY+G	88	4203.964	3544.200	-1683.516	n/a	200.00	0.73	0.279	0.266	0.196	0.259	0.077	0.057	0.109	0.081	0.083	0.075	0.081	0.112	0.075	0.118	0.077	0.057
TN93+I	89	4209.175	3541.927	-1681.366	0.00	n/a	0.79	0.279	0.266	0.196	0.259	0.073	0.054	0.079	0.077	0.121	0.072	0.077	0.163	0.072	0.085	0.073	0.054
TN93+G	89	4209.177	3541.928	-1681.367	n/a	200.00	0.79	0.279	0.266	0.196	0.259	0.073	0.054	0.079	0.077	0.121	0.072	0.077	0.163	0.072	0.085	0.073	0.054
HKY+G+I	89	4213.602	3546.353	-1683.579	0.00	175.39	0.73	0.279	0.266	0.196	0.259	0.077	0.057	0.109	0.081	0.083	0.075	0.081	0.112	0.075	0.117	0.077	0.057
GTR	91	4216.139	3533.924	-1675.338	n/a	n/a	0.80	0.279	0.266	0.196	0.259	0.089	0.037	0.076	0.093	0.122	0.014	0.053	0.165	0.145	0.082	0.014	0.110
TN93+G+I	90	4218.688	3543.956	-1681.367	0.00	200.00	0.79	0.279	0.266	0.196	0.259	0.073	0.054	0.079	0.077	0.121	0.072	0.077	0.163	0.072	0.085	0.073	0.054

FIGURE 9 : le tableau représentant du meilleur caractère comparatif

CHAPITRE I : MATERIEL ET METHODS

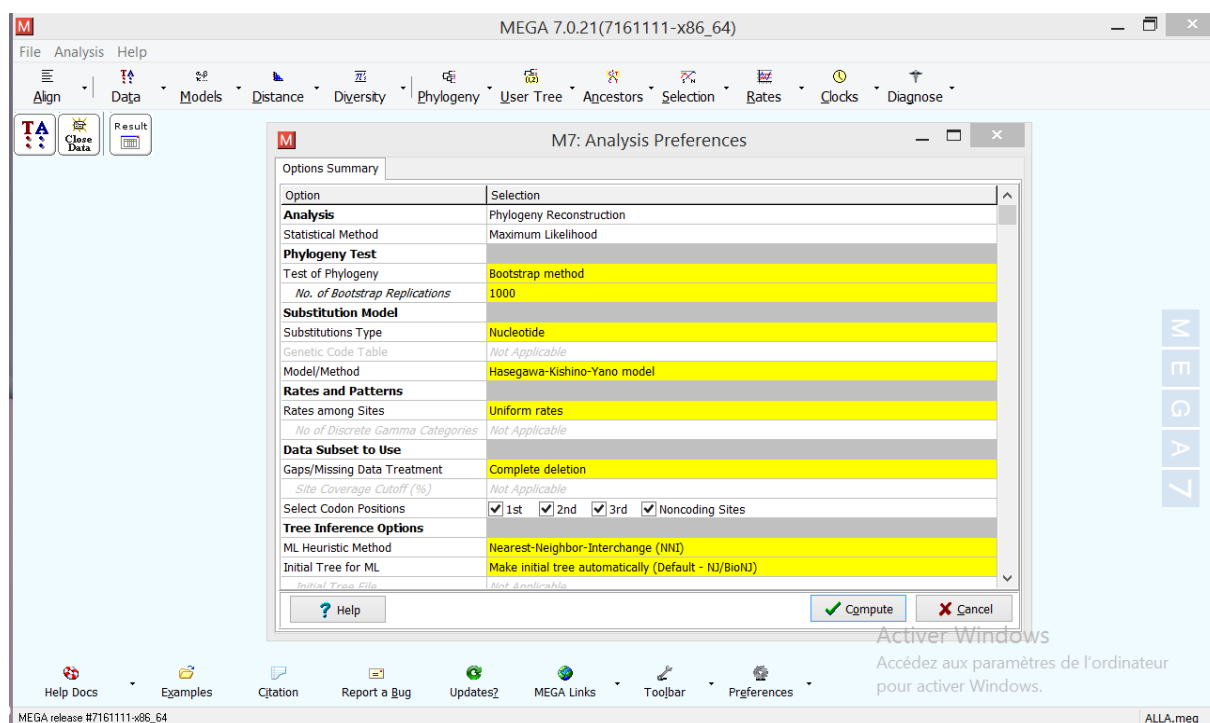


FIGURE 10 : Construction de l'arbre phylogénétique

Vue la quantité des données utilisés qui demandent des ordinateurs très performants pour l'analyse nous avons devisé les données en groupes dont on a construit des arbres phylogénétiques pour ces derniers en trois étapes en prenant les séquences les plus proches résultant de chaque étape pour les lancer sur un arbre phylogénétique final.

Les groupes de données sont choisis selon la situation géographique et la taille des données des pays étudiés :

- Groupe 1 : USA ;
 - Groupe 2 : ANG (Angleterre) ;
 - Groupe 3 : AU (Australie, Kuwait, KSA, Algérie) ;
 - Groupe 4 : ALL (Allemande, Italie, Suisse) ;
 - Groupe 5 : FE (France, Espagne).
1. La première étape : Construction de l'arbre phylogénétique des 5 groupes suscités ;
 2. Deuxième étape : Construction de l'arbre phylogénétique des séquences obtenus de la première partie du groupe 1 avec le groupe 3 et les trois groupes 2,4 et 5 ;
 3. Troisième étape : Construction de l'arbre phylogénétique finale regroupant tous les pays étudiés.

CHAPITRE I : MATERIEL ET METHODS

Les analyses phylogénétiques ont pris entre 24 h à 144h selon les données pour chaque analyse.

I.3.D Comparaison des séquences

Afin de déterminer et identifier les points de mutations affectant l'ensemble des séquences étudiés, nous avons utilisés logiciel de comparaison entre séquences GenieGen (Génie Génétique). Commenant par un alignement multiple des séquences comparés du clade G et clade O qui ont eu des résultats très significatifs dans l'étude statistique de 12 pays (Australie, Angleterre, Espagne, France, Italie, Jordanie, KSA, Kuwait, Norvège, Qatar, Saudia, Suisse) récente de Triqui.Ch et al. (Groupe I), en presse suivis par une comparaison manuelle des bases azotés. (Figure 11 et 12). Et pour des résultats plus fiables nous avons comparés 15 autres pays (Afrique de Sud, Allemande, Brésilille, Canada, Chine, India, Iran, Japon, Libanie, Maroc, Russie, Suède, Tunisie, Turquie, USA) (Groupe II).

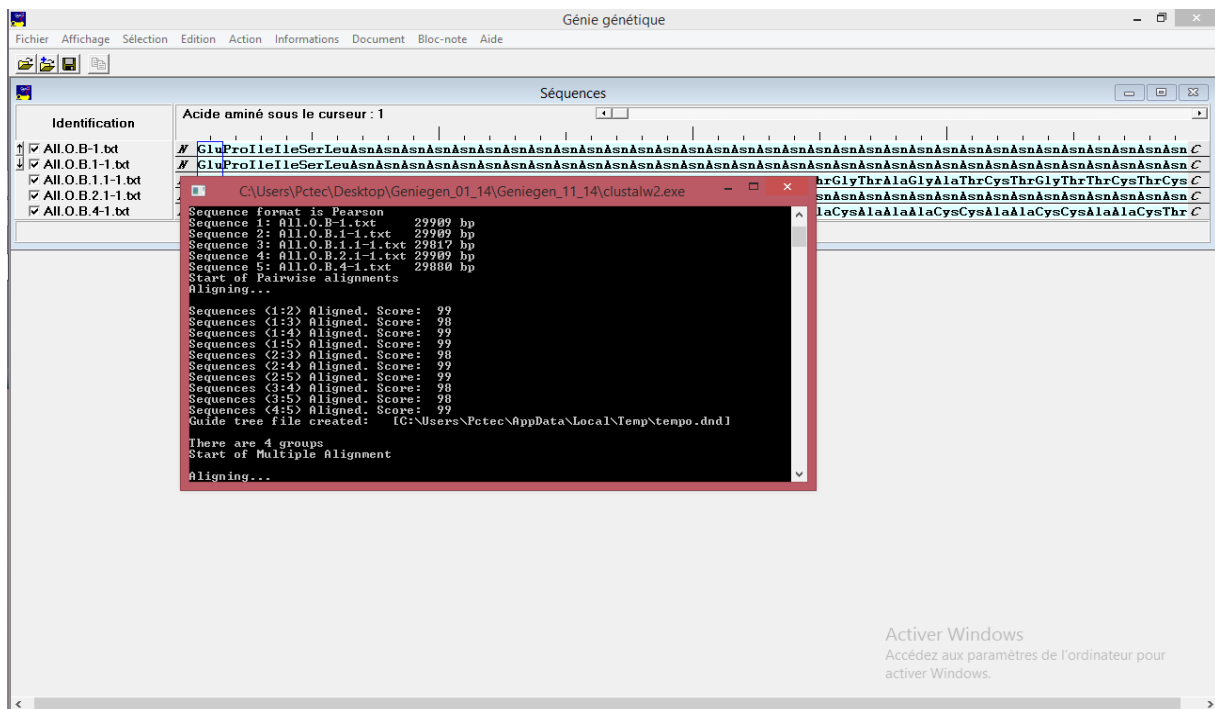


FIGURE 11 : Alignement des séquences à comparer

CHAPITRE I : MATERIEL ET METHODS

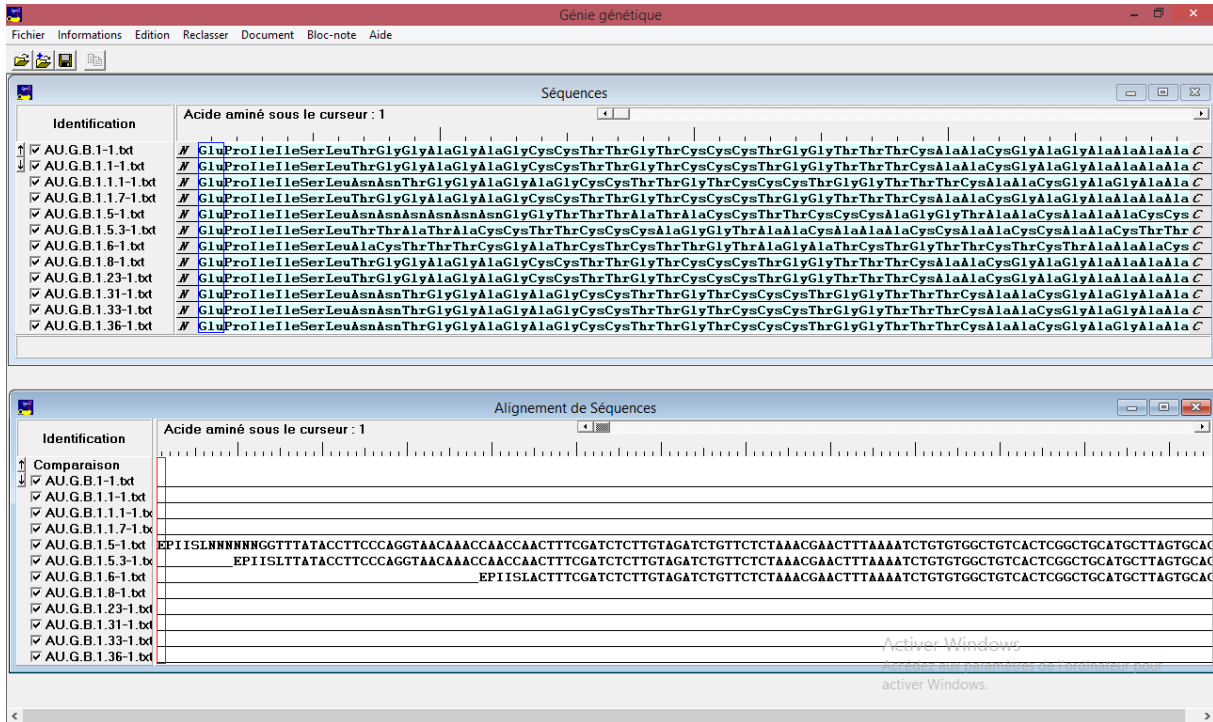


FIGURE 12 : Comparaison des séquences

Notez que logiciel GenieGen a permet l’obtention un tableau informatif des compositions des poly protéines démontré dans la section résultats et discussion.

CHAPITRE II : RESULTATS ET I NTERPRETATION

Chapitre II : Résultats et interprétations

L'analyse phylogénétique et la comparaison de séquences ont été réalisé afin de déterminer la diversité et l'évolution du SARS-COV 2 et pour identifier les différents points de mutations suite à l'étude statistique de Triqui.Ch., et al, en presse.

II.1 Phylogénie du covid-19 :

Les résultats sont exprimés sous forme d'arbres phylogénétique pour chaque groupe. Chaque arbre phylogénétique de nos résultats obtenus comprend un ensemble des sous-classes correspondant aux différents lignages des pays étudiés.

II.1.A Résultats de la première étape :

- L'arbre phylogénétique du groupe 1 (USA) :

CHAPITRE II : RESULTATS ET I NTERPRETATION

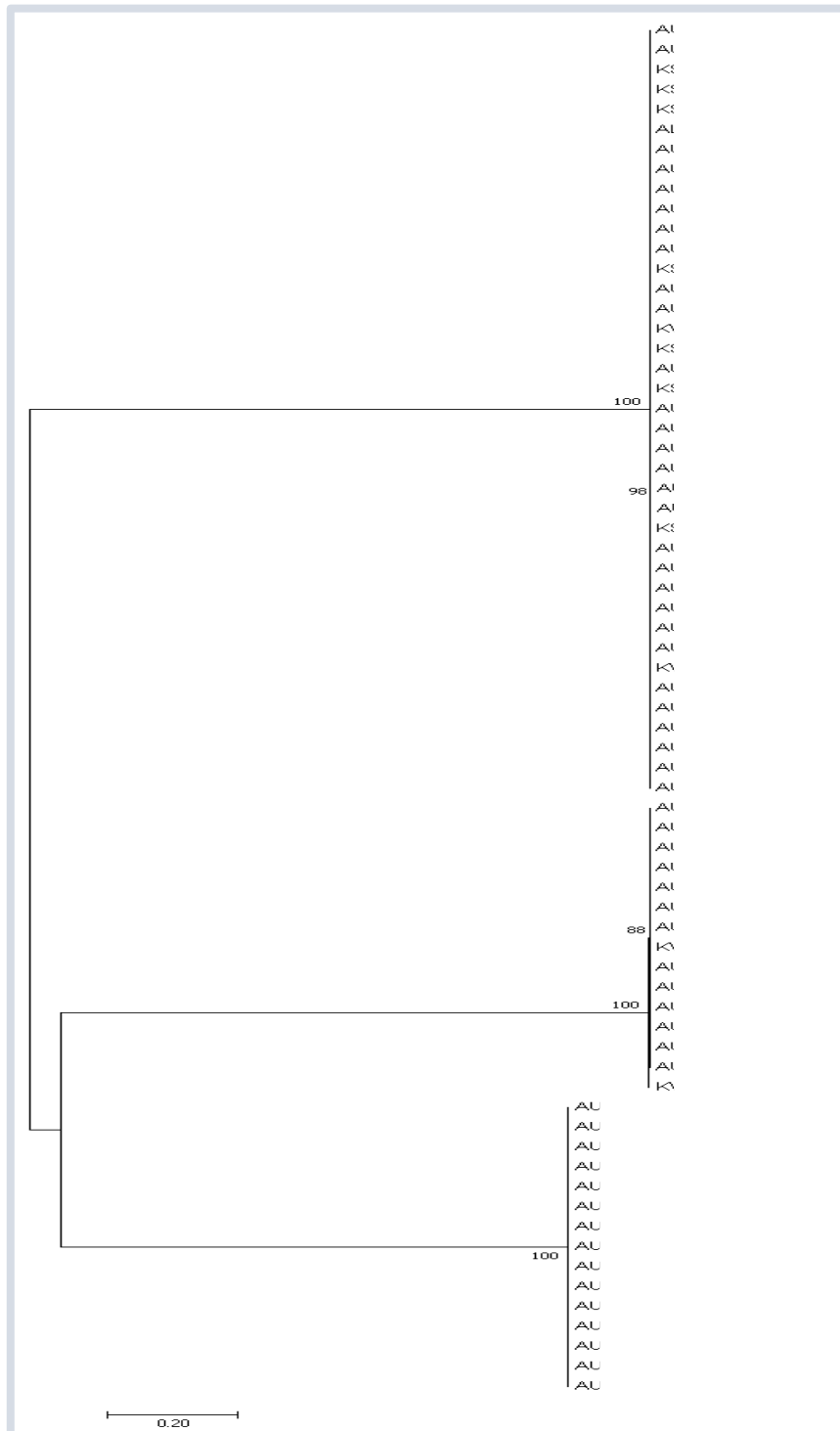


FIGURE 15 : L'arbre phylogénétique du troisième groupe (Au).

L'arbre phylogénétique du troisième groupe AU qui présente les 71 séquences alignées de l'Algérie, l'Australie, Kuwait et KSA est également divisé en classes et sous-classes regroupant les différents lignages des clades de différents pays.

- **L'arbre phylogénétique du groupe 4 (FE) :**

CHAPITRE II : RESULTATS ET I NTERPRETATION

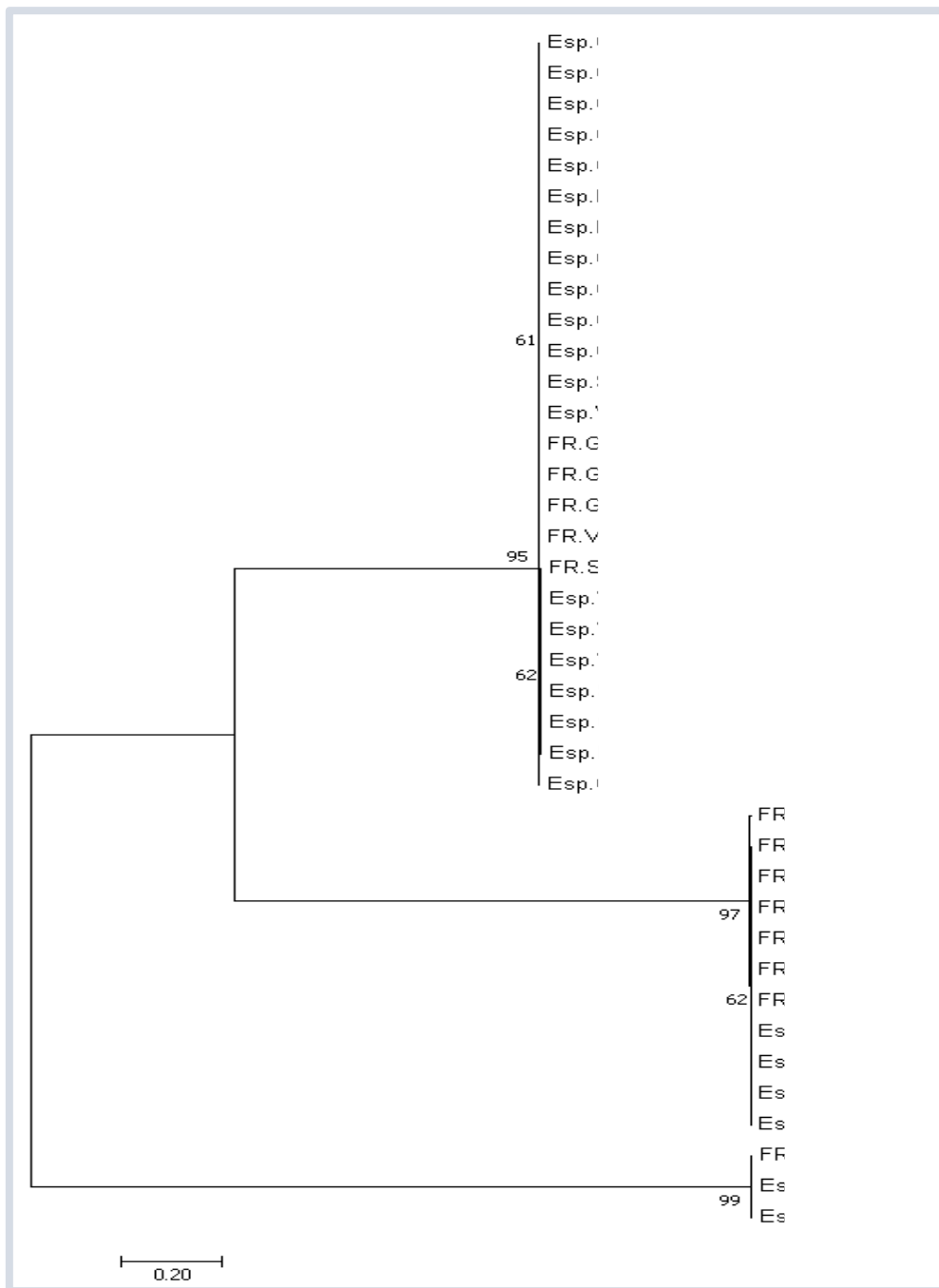


FIGURE 16 : L'arbre phylogénétique du quatrième groupe (FE).

La France et l'Espagne sont regroupés dans le même groupe FE résultant l'arbre phylogénétique en dessus qui est divisé en deux classes tout comme les trois premiers groupes. Ces deux classes présentent des sous classes montrant une grande diversité de ce groupe.

- **L'arbre phylogénétique du groupe 5 (ALL) :**

CHAPITRE II : RESULTATS ET I NTERPRETATION

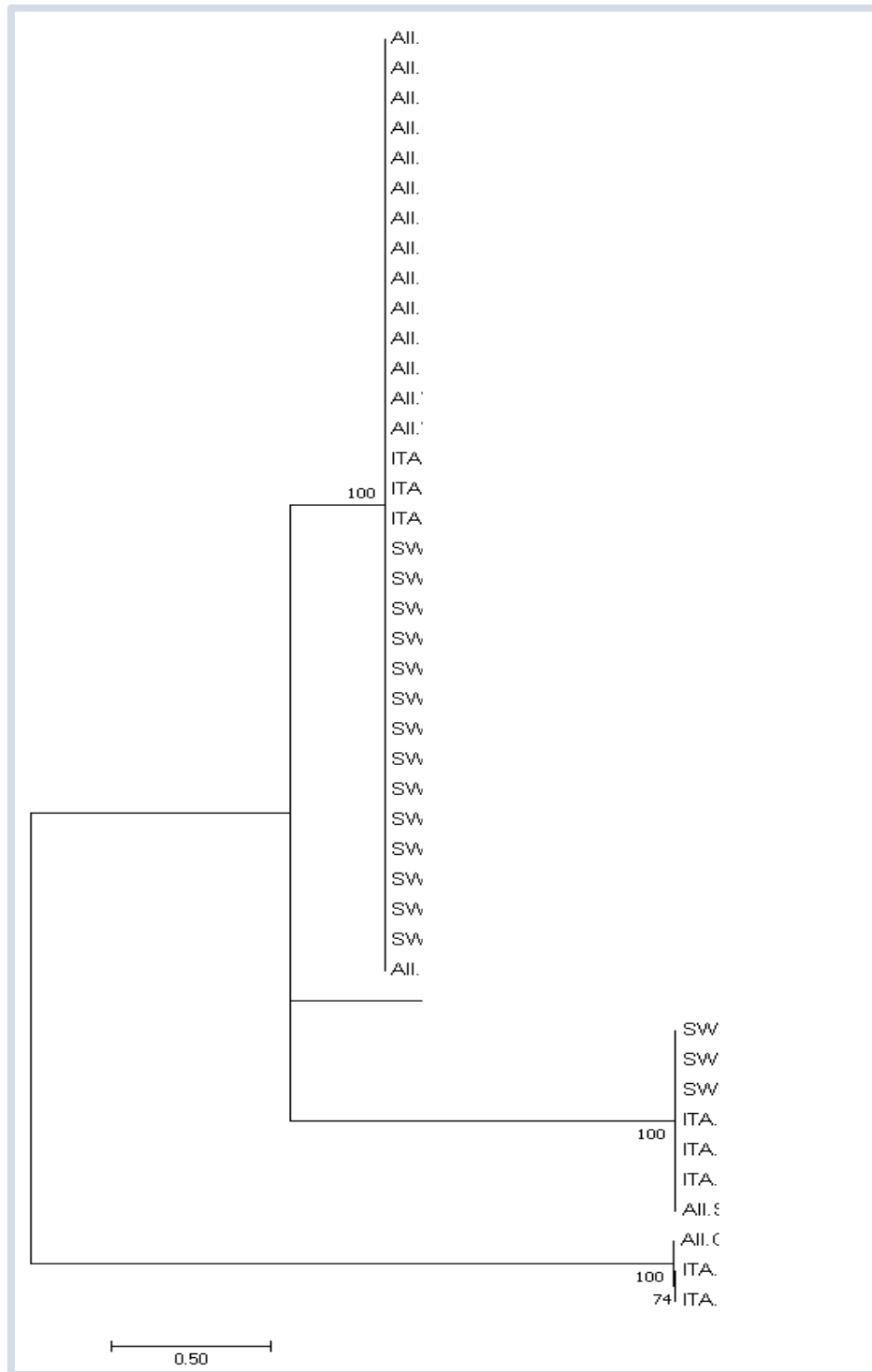


FIGURE 17 : L'arbre phylogénétique du cinquième groupe (ALL).

Les 43 séquences alignées du 5-ème groupe ALL (Allemande, Italie, Suisse) présentent sur l'arbre phylogénétique de la Figure 17 deux classes subdivisé en sous-classes qui montrent aussi une diversité de regroupement des séquences.

II.1.B Résultats de la deuxième étape :

CHAPITRE II : RESULTATS ET INTERPRETATION

Cette étape est complémentaire de l'étape précédente, elle comprend la construction de l'arbre phylogénétique des séquences obtenus de la première partie du groupe 1 avec le groupe 3 et les trois groupes 2,4 et 5.

- L'arbre phylogénétique des trois groupes 2,4 et 5 :

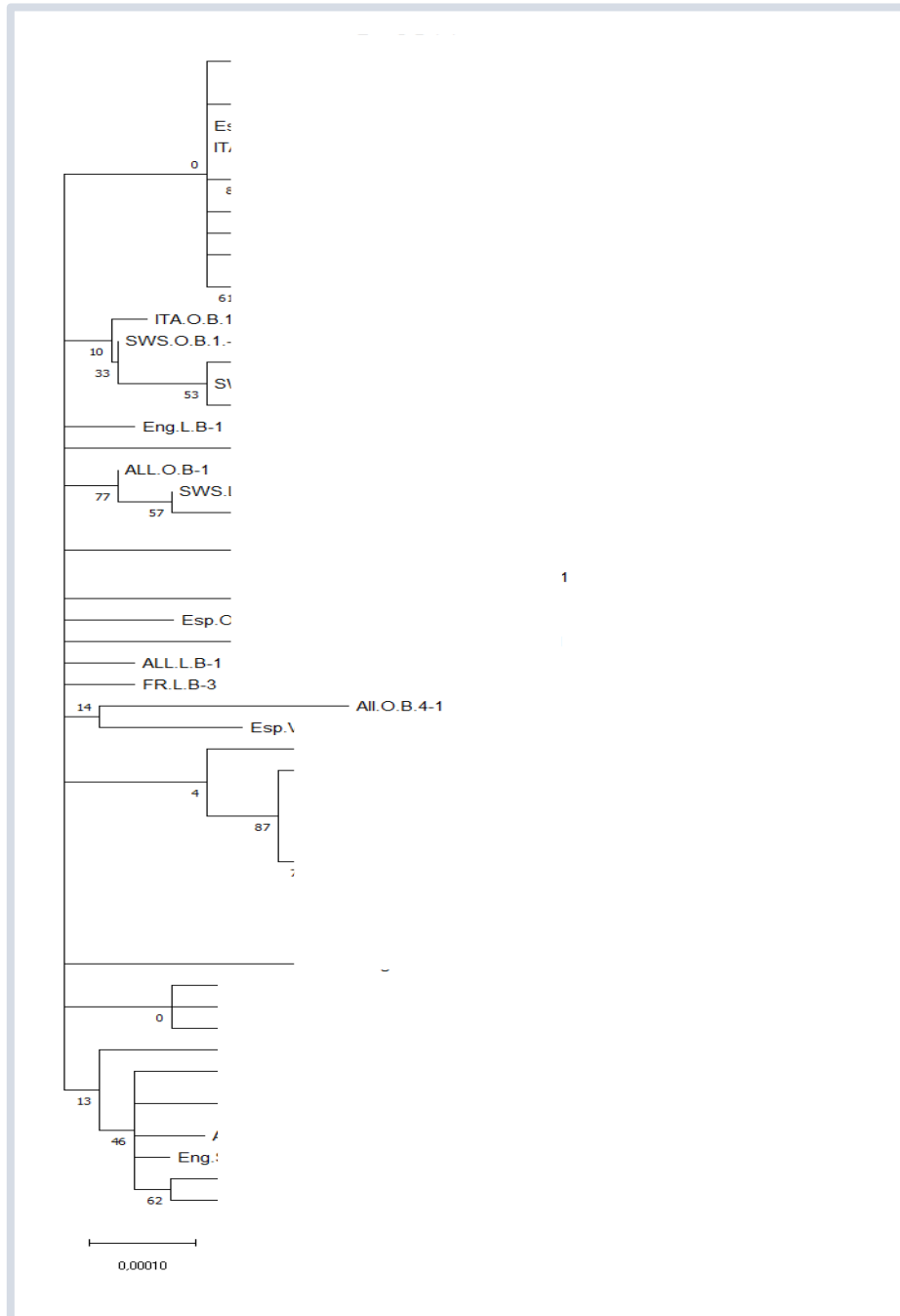


FIGURE 18 : L'arbre phylogénétique des trois groupes 2(ANG),4(FE) et 5(ALL).

Sur l'arbre phylogénétique qui présente les pays étudiés de l'Europe résultant de la première étape de l'étude phylogénétique (Figure 18), nous observons les différentes classes

CHAPITRE II : RESULTATS ET I NTERPRETATION

qui sont des fois présentées sous forme de nœuds simples et des fois divisées en sous-classes et regroupant les différents lignages sans aucun effet ni de la région, ni du clade, ni de lignages.

- L'arbre phylogénétique du groupe 1 avec le groupe 3 :

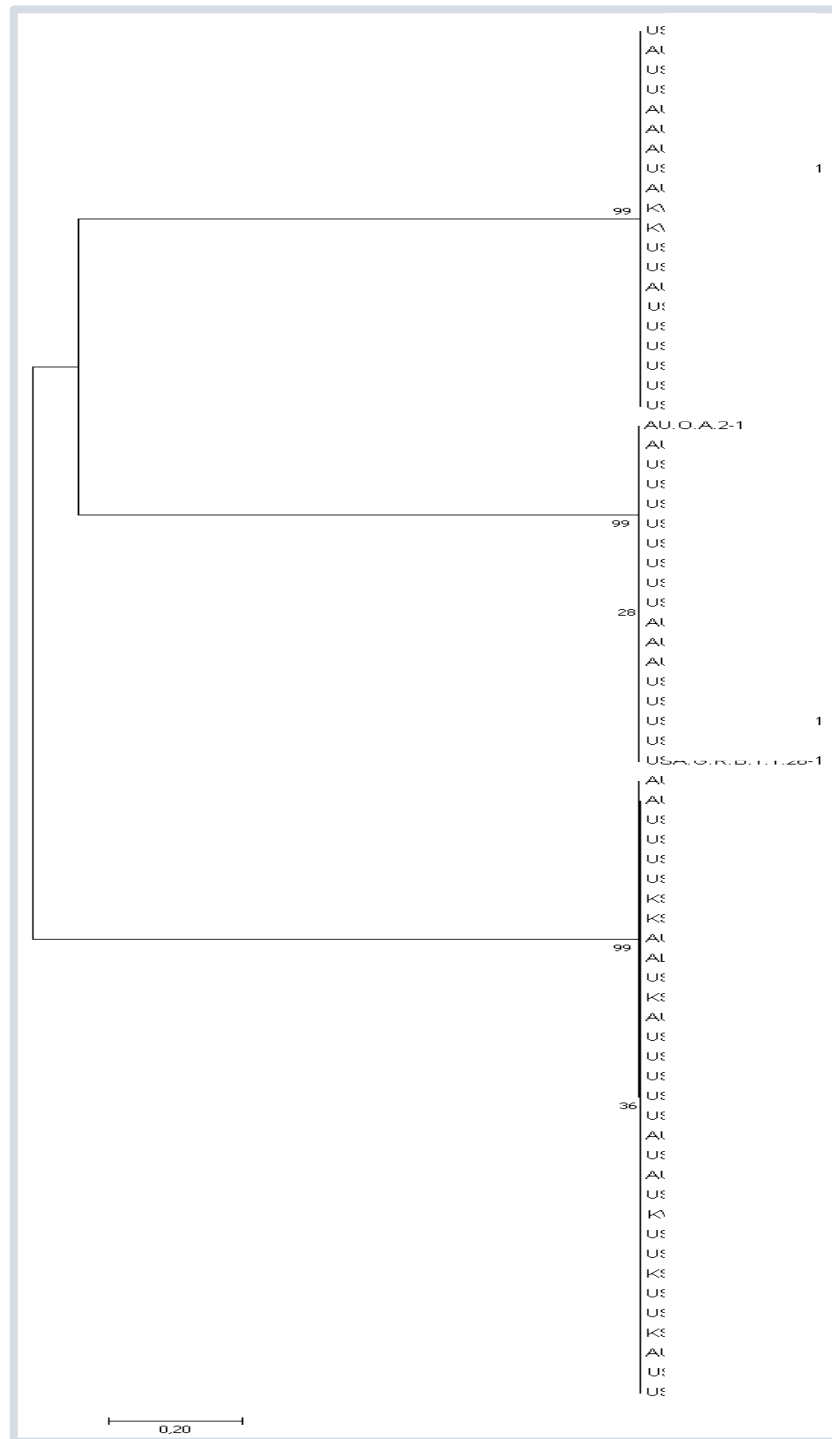


FIGURE 19 : L'arbre phylogénétique du groupe 1 (USA) avec le groupe 3 (AU).

CHAPITRE II : RESULTATS ET I NTERPRETATION

L'arbre phylogénétique de la Figure 19 ressemble aux arbres de la première étape sauf que les séquences présentées sur cet arbre sont obtenues de la première étape et qui regroupent les trois groupes 2,4 et 5. Il est divisé en deux classes divisées par la suite en sous-classes et présentant les différentes séquences de différents lignages.

II.1.C Résultat de la dernière étape :

Cette étape est l'étape final de l'étude phylogénétique de ce travail, tous les pays sont regroupés sur un seul arbre phylogénétique.

CHAPITRE II : RESULTATS ET I NTERPRETATION

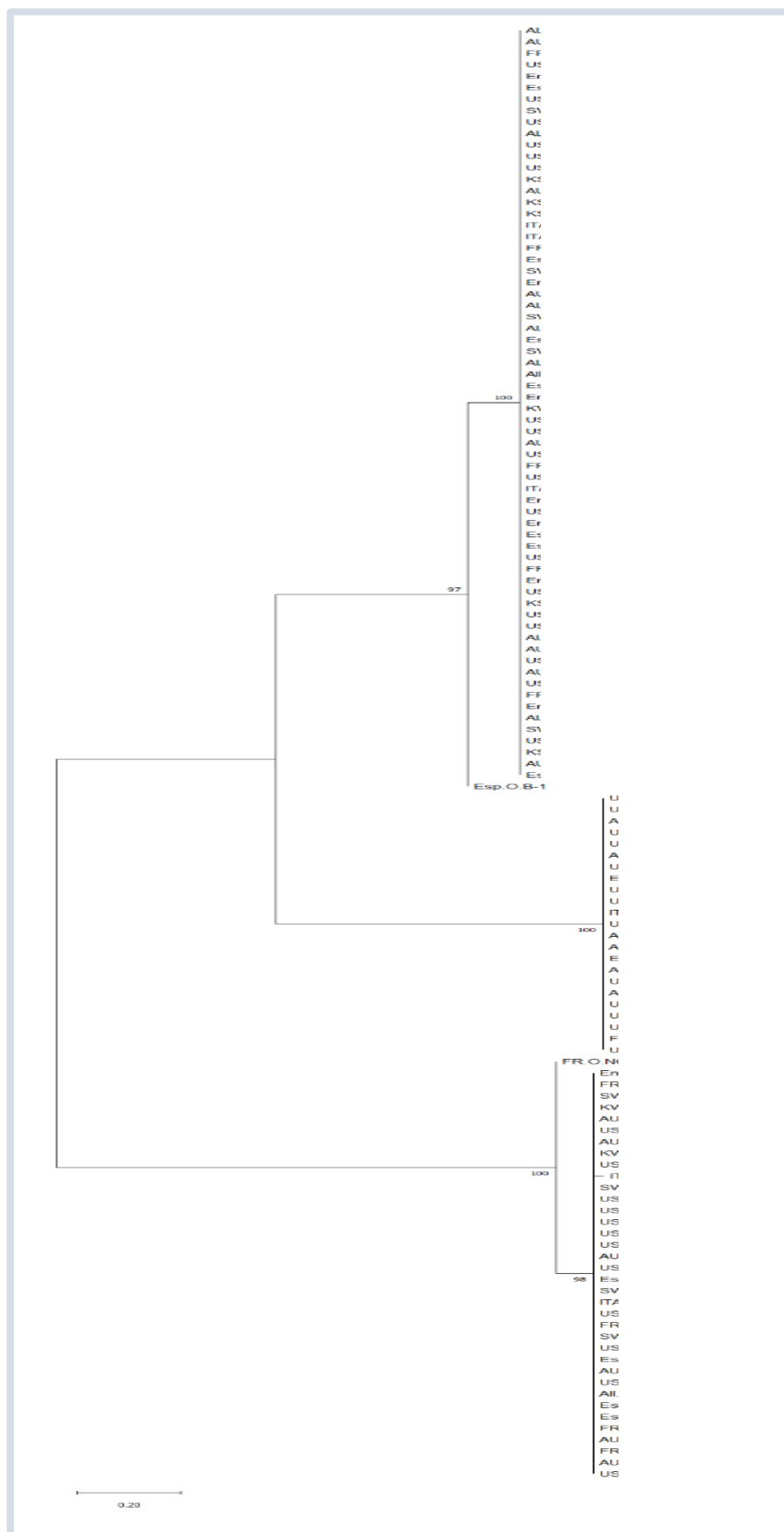


FIGURE 20 : L'arbre phylogénétique final

CHAPITRE II : RESULTATS ET I NTERPRETATION

L'arbre phylogénétique en dessus est l'arbre phylogénétique final qui regroupe le résultat des étapes précédentes et tous les pays étudiés, il est divisé en deux classes et ces dernières sont divisées également en sous-classes regroupant les différentes séquences de lignages des pays.

Nous avons remarqué sur les arbres phylogénétiques de cette étude qu'au niveau de chaque sous-classe ne représentent en aucun cas la région, ni le lignage, ni le clade ce qui reflète la très grande diversité et donc la grande mutabilité de ce virus.

II.2 Identifications des mutations :

Après l'étude phylogénétique et pour compléter notre étude, nous avons identifié toutes les mutations présentes sur les séquences utilisées pour ce présent travail comme il a été signalé dans la section matériel et méthodes. Le tableau suivant montre le résultat de cette identification.

Tableau 01: les différentes mutations des séquences étudiées à la construction d'arbre phylogénétique.

Pays	Mutations	Lignage (Clade)
Allemagne		B.1.1(Gr), B.1.1.1(Gr), B.1.15(Gr) B.1.1(Gr), B.1.1.1(Gr), B.1.1(O) B(L), B.3(L) B.2(O), B.2.1(O) B.2.1(O) B.2.1(O) B.1(O) B.4(O) B(O), B.1(O), B.1.1(O) B.1(G), B.1.5(G), B.1.8(G) B.1(Gh) B.1.1(Gh), B.1.1.1(Gh) B.1.1(O) A(S)
Australie		A.2(O), B.1(O), B.1.1(O),

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B.6(O) A.2(S), B.1(S), B.1.1(S), B.1.1(G) B.1(V)</p> <p>B.1(G), B.1.113(G) B.1(Gh), B.1.36(Gh), B.1.113(Gh)</p> <p>B.1(G), B.1.1.25(G)</p> <p>B.1.1(G), B.1.1.1(G), B.1.11(G), B.1.1.7(G), B.1.1.21(G), B.1.1.25(G), B.1.1.28(G), B.1.1.33(G), B.1.5(G), B.1.5.3(G) B.1.6(G), B.1.8(G), B.1.11(G), B.1.23(G), B.1.31(G), B.1.36(G), B.1.75(G), B.1.77(G), B.1.80(G), B.1.83(G), B.1. 88(G), B.1.113(G) B.1.108(Gh), B.1.110(Gh), B.1.111(Gh), B.1.113(Gh)</p> <p>B.1.1.25(Gr) A.1(S)</p> <p>B(G), B.1(G), B.1.1(G), B.1.1.1(G), B.1.1.25(G), B.1.1.28(G), B.1.1.33(G), B.1.1.7(G) B.1(Gr), B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.19(Gr), B.1.1.21(Gr), B.1.1.23(Gr), B.1.1.25(Gr), B.1.1.28(Gr), B.1.1.33(Gr), B.1.1.7(Gr), B.1.5(Gr) B.1.1(O) B.2(V)</p> <p>B(G), B.1(G), B.1.1(G), B.1.1.1(G), B.1.1.21(G), B.1.1.25(G), B.1.1.28(G), B.1.1.33(G), B.1.1.7(G) B.1(Gr), B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.19(Gr), B.1.1.21(Gr), B.1.1.23(Gr), B.1.1.25(Gr), B.1.1.28(Gr), B.1.1.33(Gr), B.1.1.7(Gr), B.1.5(Gr) B.1.1(O)</p> <p>B.1.1(G), B.1.1.33(G) B.1.1.33(Gr)</p>
--	--	---

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B.1(G), B.1.104(G), B.1.11(G), B.1.88(G)</p> <p>B.1(G), B.1.1(G), B.1.113(G), B.1.3(G), B.1.31(G), B.1.36(G) B(Gh), B.1(Gh), B.1.1(Gh), B.1.108(Gh), B.1.110(Gh), B.1.111(Gh), B.1.113(Gh), B.1.120(Gh), B.1.121(Gh), B.1.13(Gh), B.1.2(Gh), B.1.22(Gh), B.1.25(Gh), B.1.26(Gh), B.1.29(Gh), B.1.3(Gh), B.1.35(Gh), B.1.36(Gh), B.1.39(Gh), B.1.42(Gh),B.1.43(Gh), B.1.5(Gh)</p> <p>B.1(O), B.2(O), B.2.1(O), B.2.2, B.2.4(O), B.2.6(O), B.2.6(O)</p> <p>B.1(V), B.1.1(V), B.2(V), B.2.1(V), B.2.2(V), B.2.4(V), B.2.6(V), B.2.6(V)</p> <p>B.1(G), B.1.1(G) B.1(Gh), B.1.1(Gh), B.1.2(Gh) B.1.1(O), B.4(O)</p> <p>A(O), A.1(O), A.2(O), A.3(O), B.1(O), B.1.1(O) A(S), A.1(S), A.1.3(S), A.2(S), A.3(S), A.5(S), B.1(S), B.1.1(S), B.2(S), B.3(S), B.6(S)</p> <p>B.1(G), B.1.1(G), B.1.3(G), B.1.31(G) B(Gh), B.1(Gh), B.1.1(Gh), B.1.108(Gh), B.1.121(Gh), B.1.2(Gh), B.1.25(Gh), B.1.26(Gh), B.1.29(Gh), B.1.3(Gh), B.1.31(Gh), B.1.39(Gh), B.42(Gh), B.1.43(Gh), B.1.5(Gh) B.1(O)</p> <p>B.1.1.25(G) B(L), B.21(L), B.9(L) B(O), B.9(O)</p> <p>B.2.1(O) B.1.1(S) B.2.1(V)</p> <p>B.2.1(O)</p>
--	--	---

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B.1.1(S) B.2.1(V)</p> <p>B.6(O)</p> <p>B.1(G), B.1.1.1(G) B.1.1.1(Gr)</p> <p>B.1(G), B.1.1(G), B.1.1.1(G) B.1(Gh) B.1.1(Gr), B.1.1.1(Gr), B.1.1.25(Gr) A(O), B(O), B.1.1(O), B.2(O), B.2.1(O), B.2.2(O), B.2.4(O), B.2.6(O), B.3(O), B.4(O), B.5(O), B.6(O), B.9(O) A(S), A.1(S), A.2(S) B.1(V), B.1.1(V), B.2(V), B.2.1(V), B.2.2(V), B.2.4(V), B.2.6(V)</p> <p>B.1.1(G), B.1.104(G) B.6(O)</p> <p>B(G), B.1(G), B.1.1(G), B.1.1.1(G), B.1.1.10(G), B.1.1.21(G), B.1.1.25(G), B.1.1.28(G), B.1.1.33(G), B.1.1.7(G), B.1.104(G), B.1.11(G), B.1.113(G), B.1.13(G), B.1.23(G), B.1.3(G), B.1.31(G), B.1.36(G), B.1.5(G), B.1.5.3(G), B.1.6(G), B.1.75(G), B.1.77(G), B.1.8(G), B.1.80(G), B.1.83(G), B.1.88(G) B(Gh), B.1(Gh), B.1.1(Gh), B.1.108(Gh), B.1.110(Gh), B.1.111(Gh), B.1.113(Gh), B.1.120(Gh), B.1.121(Gh), B.1.13(Gh), B.1.2(Gh), B.1.22(Gh), B.1.25(Gh), B.1.26(Gh), B.1.29(Gh), B.1.3(Gh), B.1.31(Gh), B.1.36(Gh), B.1.39(Gh), B.1.42(Gh), B.1.43(Gh) B.1(Gr), B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.19(Gr), B.1.1.21(Gr), B.1.1.23(Gr), B.1.1.25(Gr), B.1.1.28(Gr), B.1.1.33(Gr), B.1.1.7(Gr), B.1.5(Gr) B.1(O), B.1.1(O), B.1.22(O), B.4(O) B.1.1(S)</p>
--	--	---

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B.1.1(Gr), B.1.1.25(Gr) A.1(O) A.1(S), A.1.3(S), B.1.1(S)</p> <p>A.1(O), A.1(S), A.1.3(S), B.1.1(S) B.2(S)</p>
<p>Angleterre</p>		<p>B.1(G), B.1.5(G), B.1.79(G) B.1(Gh), B.1.36(Gh) B.1.1(Gr), B.1.1.1(Gr) B(L) B.1.79(O) B.2(V)</p> <p>B.1(G), B.1.5(G) B.1.1(Gr) B.3(L) B.2.1(V), B.2.6(V)</p> <p>B.1.1(Gr) B.2(V)</p> <p>B.1(G), B.1.99(G) B.1.1(Gr)</p> <p>B.1.1(Gr), B.1.1.3(Gr)</p> <p>B.1.1(Gr)</p> <p>B(G), B.1(G), B.1.1(G), B.1.1.1(G), B.1.1.4(G), B.1.104(G), B.1.105(G), B.1.106(G), B.1.11(G), B.1.11(G), B.1.13(G), B.1.34(G), B.1.35(G), B.1.5(G), B.1.5.10(G), B.1.5.2(G), B.1.5.3(G), B.1.5.6(G), B.1.5.9(G), B.1.6(G), B.1.72(G), B.1.75(G), B.1.76(G), B.1.77(G), B.1.79(G), B.1.8(G), B.1.81(G), B.1.82(G), B.1.86(G), B.1.88(G), B.1.89(G), B.1.93(G), B.1.97(G), B.1.98(G), B.1.99(G), B.2.1(G) B.1(Gh), B.1.1(Gh), B.1.110(Gh), B.1.111(Gh), B.1.113(Gh), B.1.117(Gh), B.1.12(Gh), B.1.13(Gh), B.1.2(Gh), B.1.22(Gh), B.1.3(Gh), B.1.36(Gh), B.1.39(Gh), B.1.42(Gh), B.1.5(Gh), B.1.66, B.1.9(Gh)</p>

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B.1(Gr), B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.13(Gr), B.1.1.15(Gr), B.1.1.2(Gr), B.1.1.23(Gr), B.1.1.25(Gr), B.1.1.3(Gr), B.1.1.30(Gr), B.1.1.32 B.1.1.33(Gr), B.1.1.4(Gr), B.1.1. 5(Gr), B.1.1.7(Gr), B.1.5(Gr), B.2.1(Gr) B(O), B.1(O), B.1.1(O), B.1.1.1(O), B.1.1.5(O), B.1.11(O), B.1.13(O), B.1.36(O), B.1.5(O), B.1.5.9(O), B.1.79(O), B.2.1(O), B.4(O), B.6(O), B.9(O) B.1(V), B.2.1(V)</p> <p>B.1.36(Gh) B.2.1(V)</p> <p>B.1(G), B.1.5(G) B.1(Gh), B.1.113(Gh) B.1.1(Gr) B.6(O) A.2(S)</p> <p>B.1(G), B.1.5(G), B.1.79(G), B.1.99(G) B.1(Gh), B.1.113(Gh), B.1.36(Gh) B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr) B(L), B.1(L), B.16(L) B.2.1(V), B.2.2(V)</p> <p>B.1(G), B.1.1(G), B.1.1.1(G), B.1.1.4(G), B.1.5(G) B.1(Gh) B.1(Gr), B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.13(Gr), B.1.1.15(Gr), B.1.1.2(Gr), B.1.1.23(Gr), B.1.1.25(Gr), B.1.1.3(Gr), B.1.1.30(Gr), B.1.1.32 B.1.1.33(Gr), B.1.1.4(Gr), B.1.1.5(Gr), B.1.1.7(Gr), B.1.5(Gr), B.2.1(Gr) B.1.1(L), B.15(L) B.1(O), B.1.1(O), B.1.1.1(O), B.1.1.5(O), B.2(O) B.2(V), B.2.1(V)</p> <p>B.1(G), B.1.1(G), B.1.1.1(G), B.1.1.4(G)</p>
--	--	--

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B.1(Gr), B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.13(Gr), B.1.1.15(Gr), B.1.1.2(Gr), B.1.1.23(Gr), B.1.1.25(Gr), B.1.1.3(Gr), B.1.1.30(Gr), B.1.1.32 B.1.1.33(Gr), B.1.1.4(Gr), B.1.1.5(Gr), B.1.1.7(Gr), B.1.5(Gr), B.2.1(Gr) B.1(O), B.1.1(O), B.1.1.1(O), B.1.1.5(O), B.2.1(O) B.2(V), B.2.1(V)</p> <p>B.1(G) B.1.1(Gr), B.1.1.33(Gr)</p> <p>B.1(G), B.1.1(G), B.1.104(G), B.1.11(G), B.1.5(G), B.1.81(G), B.1.88(G) B.1.1(Gr), B.1.1.1(Gr) B.1.11(O) B.1(V)</p> <p>B.1(G), B.1.1(G) B.1(Gh), B.1.1(Gh), B.1.110(Gh), B.1.111(Gh), B.1.113(Gh), B.1.117(Gh), B.1.12(Gh), B.1.13(Gh), B.1.2(Gh), B.1.22(Gh), B.1.3(Gh), B.1.30(Gr), B.1.36(Gh), B.1.39(Gh), B.1.42(Gh), B.1.5(Gh), B.1.66, B.1.9(Gh) B.1.1(Gr) B.1(O), B.1.36(O), B.2(O), B.9(O) B.2(V), B.2.6(V)</p> <p>B.1(O), B.1.1(O), B.1.1.1(O), B.2(O), B.2.1(O), B.2.2(O), B.2.4(O), B.2.5(O), B.2.6(O) B.1(V), B.1.1(V), B.1.1.10(V), B.2(V), B.2.1(V), B.2.2(V), B.2.4(V), B.2.6(V), B.2.9(V)</p> <p>B.1.2(Gh)</p> <p>A(S), A.1(S), A.2(S), A.3(S), A.5(S), B.1.1(S), B.2.1(S) B.1.1(V)</p> <p>B.1(G), B.1.89(G) B.1(Gh), B.1.1(Gh), B.1.117(Gh), B.1.12(Gh), B.1.2(Gh), B.1.3(Gh),</p>
--	--	---

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B.1.30(Gh), B.1.39(Gh), B.1.42(Gh), B.1.5(Gh), B.1.66(Gh) B.1.1(Gr) B.1(O) B.2.1(V)</p> <p>B(L), B.1(L), B.1.1(L), B.10(L), B.11(L), B.15(L), B.16(L), B.2(L), B.2.1(L), B.21(L), B.9(L) B(O), B.1(O), B.1.1(O), B.10(O), B.15(O), B.2(O), B.2.1(O), B.21(O), B.9(O)</p> <p>B.1(O), B.2.1(O) B.1(V), B.1.1(V), B.2(V), B.2.1(V)</p> <p>B.1.1.1(Gr) B.1(O), B.1.1.1(O), B.2(O), B.2.1(O) B.1(V), B.1.1(V), B.2.1(V)</p> <p>B.1.5(G) B.1.1(Gr) B.6(O)</p> <p>B.1(G), B.1.1.1(G) B.1.1(Gr), B.1.1.1(Gr) B.1(O), B.1.1.1(O)</p> <p>B.1(G), B.1.1(G), B.1.104(G), B.1.105(G), B.1.11(G), B.1.5(G), B.1.79(G), B.2.1(G) B.1(Gh), B.1.117(Gh), B.1.12(Gh), B.13(Gh), B.1.22(Gh), B.1.30(Gh), B.1.36(Gh) B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.3(Gr) B.1(O), B.1.1(O), B.1.1.1(O), B.1.1.5(O), B.2(O), B.2.1(O), B.2.2(O), B.2.4(O), B.2.5(O), B.3(O), B.4(O), B.5(O), B.6(O), B.9(O) B.1(V), B.1.1(V), B.1.1.10(V), B.2(V), B.2.1(V), B.2.2(V), B.2.4(V), B.2.6(V), B.2.9(V)</p> <p>B.1(G), B.1.104(G), B.1.11(G), B.1.5(G) B.1.1(Gr), B.1.1.1(Gr) B.6(O) B.2.4(V)</p>
--	--	--

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B(G), B.1(G), B.1.1(G), B.1.1.1(G), B.1.1.4(G), B.1.102(G), B.1.104(G), B.1.105(G), B.1.106(G), B.1.11(G), B.1.13(G), B.1.34(G), B.1.35(G), B.1.5(G), B.1.5.10(G), B.1.5.2(G), B.1.5.3(G), B.1.5.6(G), B.1.5.9(G), B.1.6(G), B.1.72(G), B.1.75(G), B.1.76(G), B.1.77(G), B.1.79(G), B.1.8(G), B.1.81(G), B.1.82(G), B.1.86(G), B.1.88(G), B.1.89(G), B.1.93(G), B.1.97(G), B.1.98(G), B.1.99(G) B.1(Gh), B.1.1(Gh), B.1.111(Gh), B.1.113(Gh), B.1.117(Gh), B.1.12(Gh), B.1.13(Gh), B.1.2(Gh), B.1.22(Gh), B.1.3(Gh), B.1.36(Gh), B.1.39(Gh), B.1.42(Gh), B.1.5(Gh), B.1.66, B.1.9(Gh) B.1(Gr), B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.13(Gr), B.1.1.15(Gr), B.1.1.2(Gr), B.1.1.23(Gr), B.1.1.25(Gr), B.1.1.3(Gr), B.1.1.30(Gr), B.1.1.32 B.1.1.33(Gr), B.1.1.4(Gr), B.1.1.5(Gr), B.1.1.7(Gr), B.1.5(Gr), B.2.1(Gr) B.1(O), B.1.1(O), B.1.1.1(O), B.1.1.5(O), B.1.11(O), B.1.36(O), B.1.5(O), B.1.5.9(O), B.1.79(O)</p> <p>B.1.1(Gr) B.2.2(O) A.1(S), B.1.1(S) B.2.2(V)</p> <p>A.1(S), B.1.1(S)</p>
<p>Italie</p>		<p>B.1(G), B.1.1(G), B.1.107(G), B.1.5(G), B.1.5.5(G), B.1.69(G) B.1(Gh) B.1.1(Gr), B.1.1.1(Gr) B.1(O), B.1.1(O)</p> <p>B.1.1(G) B.1.1(Gh), B.1.1.1(Gh) B.1.1(Gr), B.1.1.1(Gr) B.1.1(O)</p>

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B.1.1(G) B.1.1(Gr), B.1.1.1(Gr) B.1.1(O)</p> <p>B.1(Gh)</p> <p>B.2(O) B.2(V)</p> <p>B.1(Gh)</p> <p>B.1.1.1(Gr)</p> <p>B.1(G) B.1(O), B.1.1(O) B.2(V)</p> <p>B.1(G), B.1.1(G), B.1.107(G), B.1.5(G), B.1.5.5(G) B.1(Gh) B.1.1(Gr), B.1.1.1(Gr) B.1(O)</p>
<p>France</p>	<p style="text-align: center;">—</p>	<p>B.1(G), B.1.11(G), B.1.5(G) B.1(Gh) B.1(Gr), B.1.1(Gr), B.1.1.7(Gr) B(O), B.1(O), B.1.1(O)</p> <p>B.6(O)</p> <p>B.1(G) B.1(Gr), B.1.1(Gr), B.1.1.7(Gr) B.1.1(O)</p> <p>B.1.1(Gr), B.1.1.7(Gr) B.1.1(O)</p> <p>B.1.11(G)</p> <p>B.1(Gh) B.1(O)</p> <p>B.2(O) B.2(V), B.2.1(V), B.2.6(V)</p> <p>A.2(S)</p> <p>B.1(Gh)</p> <p>B(L) B(O)</p> <p>B.2.1(V)</p> <p>B.2.1(V)</p>

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B.6(O)</p> <p>B.1(G) B.1(Gh) B.1.1(Gr) B(O), B.6(O) B.2(V), B.2.1(V), B.2.6(V)</p> <p>B.6(O)</p> <p>B.1(G), B.1.11(G), B.1.5(G) B.1(Gh) B.1(Gr), B.1.1(Gr), B.1.1.7(Gr) B.1(O)</p>
<p>Suisse</p>		<p>B.1(G)</p> <p>B.1(G)</p> <p>B.1.36(Gh)</p> <p>B(G), B.1(G), B.1.1(G), B.1.1.1(G), B.1.1.8(G), B.1.107(G), B.1.11(G), B.1.5(G), B.1.74(G), B.1.8(G), B.1.94(G) B.1(Gh), B.1.12(Gh), B.1.3(Gh), B.1.36(Gh), B.1.9(Gh) B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.33(Gr), B.1.1.6(Gr), B.1.1.7(Gr) B.1(O), B.1.1(O), B.1.1.1(O)</p> <p>B.1.1(Gr)</p> <p>B(G), B.1(G), B.1.5(G) B.1(Gh) B.1(O)</p> <p>B.1(G), B.1.1(G) B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.33(Gr), B.1.1.6(Gr), B.1.1.7(Gr) B.1.1(O), B.1.1.1(O)</p> <p>B.1.1(G) B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.33(Gr), B.1.1.6(Gr), B.1.1.7(Gr) B.1.1(O), B.1.1.1(O)</p> <p>B.1.1(Gr), B.1.1.36(Gr)</p> <p>B.1.11(G)</p>

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B.1(G)</p> <p>B.1(G) B.1.1(O), B.2(G) B.2(V), B.2.1(V)</p> <p>A.5(O) A.2(S)</p> <p>B.1(G) B.1(Gh), B.1.12(Gh), B.1.3(Gh)</p> <p>B(L), B.10(L)</p> <p>B.1(O) B.2.1(V)</p> <p>B.1(O) B.2.1(O)</p> <p>B.1.1(G), B.1.1.1(G) B.1.1.1(Gr)</p> <p>B.1(G), B.1.5(G) B.1(Gh) B.1.1(Gr), B.1.1.10(Gr) B.2(O) B.2(V), B.2.1(V)</p> <p>B.1.1(Gr)</p> <p>B(G), B.1(G), B.1.1(G), B.1.1.1(G), B.1.1.8(G), B.1.107(G), B.1.11(G), B.1.5(G), B.1.74(G), B.1.8(G), B.1.94(G)</p>
<p>Espagne</p>		<p>B.1.79(G) B.1.1(Gh) A.2</p> <p>B.1(G), B.1.79(G) A.2(S)</p> <p>A.2(S)</p> <p>B.1.5(G)</p> <p>B(G), B.1(G), B.1.1(G), B.1.1.10(G), B.1.102(G), B.1.5(G), B.1.5.3(G), B.1.74(G), B.1.77(G), B.1.79(G), B.1.84(G) B.1(Gh)</p>

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B.1.1(Gr), B.1.1.1(Gr) A(O), B.1(O), B.1.5(O) A.5(S) A.5(V)</p> <p>B.1.1(Gr)</p> <p>B.1(G)</p> <p>B.1.1(G), B.1.5(G) B.1.1(Gr), B.1.1.1(Gr)</p> <p>B.1.1(G), B.1.5(G) B.1.1(Gr), B.1.1.1(Gr)</p> <p>B.1(G)</p> <p>B.1(Gh)</p> <p>B.2(V) A.5(V), B.2(V), B.2.1(V)</p> <p>A(O), A.2(O), A.5(O), B.1.1(O) A(S), A.2(S), A.5(S), B.1.1(S), B.2(S) A.5(V)</p> <p>B.1(Gh) A.2(S)</p> <p>B.1.1(Gr) B(L) B(O), B.2(O)</p> <p>B.2.1(V)</p> <p>B.2(V)</p> <p>B.1.1(Gr), B.1.1.1(Gr)</p> <p>B.1(G), B.1.5(G) B.1(Gh) B.1.1(Gr) B(O), B.1(O), B.2(O) A.2(S), A.5(S) A.5(V), B.2(V), B.2.1(V)</p> <p>A.2(S)</p> <p>B.1(G), B.1.1(G), B.1.1.10(G), B.1.102(G), B.1.5(G), B.1.5.3(G), B.1.74(G), B.1.77(G), B.1.79(G) B.1(Gh) B.1.1(Gr), B.1.1.1(Gr)</p>
--	--	--

CHAPITRE II : RESULTATS ET I NTERPRETATION

USA		<p>B.1(O), B.1.5(O)</p> <p>B.1.79(G) B.1(Gh) B.1.1(Gr) A.1(S)</p> <p>B.1(G), B.1.5(G) B.1(Gh)</p> <p>B.1(G) B.1(Gh)</p> <p>B.1(Gh) A.1(S) A.1(O)</p> <p>B.1(Gh), B.1.42(Gh) A.1(S)</p> <p>A.1(O) A.1(S)</p> <p>B.2(V)</p> <p>A(G), B(G), B.1(G), B.1.1(G), B.1.1.33(G), B.1.102(G), B.1.103(G), B.1.104(G), B.1.106(G), B.1.107(G), B.1.11(G), B.1.110(G), B.1.123(G), B.1.124(G), B.1.19(G), B.1.2(G), B.1.21(G), B.1.26(G), B.1.33(G), B.1.35(G), B.1.43(G), B.1.5(G), B.1.5.3(G), B.1.5.8(G), B.1.68(G), B.1.74(G), B.1.77(G), B.1.79(G), B.1.8(G), B.1.83(G), B.1.85(G), B.1.88(G), B.1.89(G), B.1.96(G), B.6(G), B(Gh), B.1(Gh), B.1.1(Gh), B.1.108(Gh), B.1.11(Gh), B.1.110(Gh), B.1.111(Gh), B.1.112(Gh), B.1.113(Gh), B.1.114(Gh), B.1.116(Gh), B.1.119(Gh), B1.123(Gh), B.1.124(Gh), B.1.13(Gh), B.1.2(Gh), B.1.2(Gh), B.1.2(Gh), B.1.21(Gh), B.1.22(Gh), B.1.26(Gh), B.1.26(Gh), B.1.29(Gh), B.1.3(Gh), B.1.35(Gh), B.1.36(Gh), B.1.37(Gh), B.1.38(Gh), B.1.41(Gh), B.1.42(Gh), B.1.43(Gh), B.1.44(Gh), B.1.5(Gh),</p>

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B.1.9(Gh), B.1.98(Gh) B.1(Gr), B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.24(Gr), B.1.1.26(Gr), B.1.1.28(Gr), B.1.1.32(Gr), B.1.1.33(Gr), B.1.1.4(Gr) A.1(O), A.3(O), B(O), B.1(O), B.1.1(O), B.1.124(O), B.1.2(O), B.1.26(O), B.1.3(O), B.1.43(O), B.1.5(O), B.3(O) A(S), A.1(S), B.1(S)</p> <p>B.1.26(Gh)</p> <p>B.1(Gh), B.1.2(Gh) B(L), B.6(L) A.2(O), B.6(O) A.1(S), A.2(S), B.1.1(S)</p> <p>B(G), B.1.1(G), B.1.102(G), B.1.107(G), B.1.11(G), B.1.5(G), B.1.74(G), B.1.79(G), B.1.89(G) B.1(Gh), B.1.113(Gh), B.1.116(Gh), B.1.124(Gh), B.1.2(Gh), B.1.3(Gh) B(L), B.17(L), B.3(L) A.3(O)</p> <p>B.1(G), B.1.1(G), B.1.1.33(G), B.1.21(G) B.1(Gh), B.1.2(Gh) B.1(Gr), B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.24(Gr), B.1.1.26(Gr), B.1.1.28(Gr), B.1.1.32(Gr), B.1.1.33(Gr), B.1.1.4(Gr) B.1.1(O) A.1(S), A.3(S), B.1.1(S)</p> <p>B.1(G), B.1.1(G), B.1.1.33(G), B.1.21(G) B.1.3(Gh) B.1(Gr), B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.24(Gr), B.1.1.26(Gr), B.1.1.28(Gr), B.1.1.32(Gr), B.1.1.33(Gr), B.1.1.4(Gr) B.1.1(O) B.1.1(S)</p> <p>B.1(G), B.1.1(G), B.1.1.33(G) B.1(Gr), B.1.1(Gr), B.1.1.10(Gr), B.1.1.33(Gr)</p>
--	--	--

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B.1(G), B.1.104(G), B.1.11(G), B.1.5(G) B.1.11(Gh), B.12(Gh) B(L) A.3(S)</p> <p>B.1(G), B.1.1(G), B.1.103(G), B.1.123(G), B.1.2(G), B.1.21(G), B.1.3(G), B.1.43(G) B(Gh), B.1(Gh), B.1.1(Gh), B.1.108(Gh), B.1.11(Gh), B.1.110(Gh), B.1.111(Gh), B.1.112(Gh), B.1.113(Gh), B.1.114(Gh), B.1.116(Gh), B.1.119(Gh), B.1.123(Gh), B.1.124(Gh), B.1.13(Gh), B.1.2(Gh), B.1.35(Gh), B.1.36(Gh), B.1.37(Gh), B.1.38(Gh), B.1.41(Gh), B.1.42(Gh), B.1.43(Gh), B.1.44(Gh), B.1.5(Gh), B.1.9(Gh), B.1.98(Gh) B.1.1(Gr) A.1(O), B(O), B.1(O), B.1.1(O), B.1.124(O), B.2(O), B.1.26(O), B.1.3(O), B.1.43(O), B.14(O), B.3(O) A(S), A.1(S), B.1(S), B.1.1(S) B.2(V)</p> <p>B.1.1(O), B.2(O), B.2.1(O), B.2.2(O) B.1(V), B.1.1(V), B.2(V), B.2(V), B.2.1(V), B.2.10(V), B.2.2(V), B.2.4(V), B.2.6(V)</p> <p>B.1(G), B.1.1(G), B.1.2(G) B.1(Gh), B.1.1(Gh), B.1.2(Gh), B.1.2(Gh), B.1.35(Gh), B.1.44(Gh), B.1.5(Gh), B.1.98(Gh) B.1(O), B.1.2(O) A.1(S)</p> <p>A(G), B.1(G), B.1.5(G) B.1(Gh), B.1.2(Gh) B.1.1.26(Gh) A(O), A.1(O), A.2(O), A.3(O), B(O), B.1(O), B.1.1(O) A(S), A.1(S), A.1.1(S), A.2(S), A.3(S), A.4(S), A.5(S), B(S), B.1(S), B.1.1(S), B.1.5(S), B.2(S)</p>
--	--	--

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>B(G), B.1(G), B.1.1(G), B.1.123(G), B.1.2(G), B.1.21(G), B.1.26(G), B.1.3(G), b.1.43(G)</p> <p>B(Gh), B.1(Gh), B.1.1(Gh), B.1.108(Gh), B.1.11(Gh), B.1.114(Gh), B.1.116(Gh), B.1.116(Gh), B.1.119(Gh), B.1.123(Gh), B.1.13(Gh), B.1.2(Gh), B.1.3(Gh), B.1.35(Gh), B.1.36(Gh), B.1.38(Gh), B.1.41(Gh), B.1.42(Gh), B.1.43(Gh), B.1.44(Gh), B.1.5(Gh), B.1.98(Gh)</p> <p>B.1.1(Gr)</p> <p>A.1(O), B(O), B.1(O), B.1.123(O), B.1.2(O), B.1.26(O), B.1.3(O), B.1.43(O), B.14(O)</p> <p>A(S), A.3(S), B.1(S)</p> <p>B.1(Gh), B.1.2(Gh)</p> <p>B(L)</p> <p>B(O)</p> <p>A.5(O)</p> <p>A.5(O)</p> <p>B.1(V), B.2.1(V)</p> <p>B.1(Gh)</p> <p>B.2.1(O)</p> <p>B.2.1(V)</p> <p>B.6(L)</p> <p>B.6(O)</p> <p>B.1.21(Gh)</p> <p>B.1.1.1(Gr)</p> <p>B.1(G), B.1.102(G), B.1.103(G), B.1.2(G), B.1.5(G)</p> <p>B.1(Gh), B.1.1(Gh), B.1.110(Gh), B.1.116(Gh), B.1.119(Gh), B.1.123(Gh), B.1.124(Gh), B.1.2(Gh), B.1.21(Gh), B.1.26(Gh), B.1.3(Gh), B.1.38(Gh), B.1.41(Gh), B.1.43(Gh)</p> <p>B.1.1(Gr), B.1.1.26(Gr)</p> <p>A.1(O), B(O), B.1(O), B.1.1(O), B.1.21(O), B.1.3(O), B.17(O), B.1.19(O), B.2(O), B.2.1(O), B.3(O), B.4(O), B.5(O), B.6(O)</p>
--	--	--

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>A(S), A.1(S), A.3(S) B.1(V), B.1.1(V), B.2(V), B.2.1(V), B.2.10(V), B.2.2(V), B.2.4(V), B.2.6(V)</p> <p>B.1.104(G), B.1.5(G) B.1(Gh), B.1.2(Gh) B.6(L) B.6(O) B.1.1(S)</p> <p>B(G), B.1(G), B.1.1(G), B.1.1.33(G), B.1.102(G), B.1.103(G), B.1.104(G), B.1.106(G), B.1.107(G), B.1.11(G), B.1.110(G), B.1.123(G), B.1.124(G), B.1.19(G), B.1.2(G), B.1.21(G), B.1.26(G), B.1.3(G), B.1.35(G), B.1.43(G), B.1.5(G), B.1.5.3(G), B.1.5.8(G), B.1.68(G), B.1.74(G), B.1.77(G), B.1.79(G), B.1.8(G), B.1.83(G), B.1.85(G), B.1.88(G), B.1.89(G), B.1.96(G), B.6(G) B(Gh), B.1(Gh), B.1.1(Gh), B.1.108(Gh), B.1.11(Gh), B.1.110(Gh), B.1.111(Gh), B.1.112(Gh), B.1.113(Gh), B.1.114(Gh), B.1.116(Gh), B.1.119(Gh), B.1.123(Gh), B.1.124(Gh), B.1.13(Gh), B.1.2(Gh), B.1.21(Gh), B.1.22(Gh), B.1.26(Gh), B.1.29(Gh), B.1.3(Gh), B.1.35(Gh), B.1.36(Gh), B.1.37(Gh), B.1.38(Gh), B.1.41(Gh), B.1.42(Gh), B.1.43(Gh), B.1.44(Gh), B.1.5(Gh), B.1.9(Gh), B.1.98(Gh) B.1(Gr), B.1.1(Gr), B.1.1.1(Gr), B.1.1.10(Gr), B.1.1.24(Gr), B.1.1.26(Gr), B.1.1.28(Gr), B.1.1.32(Gr), B.1.1.33(Gr), B.1.1.4(Gr) A.1(O), B(O), B.1(O), B.1.123(O), B.1.2(O), B.1.26(O), B.1.3(O), B.1.3(O), B.1.43(O), B.14(O) A(S), A.1(S), B.1(S), B.1.1(S)</p> <p>B.1(G), B.1.2(G) B.1.1(Gr), B.1.1.26(Gr)</p>
--	--	---

CHAPITRE II : RESULTATS ET I NTERPRETATION

		<p>A.1(O), B(O), B.1(O), B.1.1(O) A.1(S), A.1.1(S), B(S), B.1(S), B.1.1(S)</p> <p>B.1(G) B.1.1(Gh)</p> <p>A.1(O), B(O), B.1(O), B.1.1(O) A(S), A.1(S), A.1.1(S), B(S), B.1(S), B.1.1(S)</p>
<p>KSA</p>	<p>—</p>	<p>B.1(G), B.1.1(G), B.1.113(G), B.1.36(G) A(Gh), B.1(Gh), B.1.113(Gh), B.1.36(Gh), B.1.77(Gh), B.1.9(Gh) B.1.1(Gr) B.1(O), B.1.1(O), B.1.113(O), B.1.36(O)</p> <p>B.6(O)</p> <p>B.1(G), B.1.113(G) B.1(Gh), B.1.113(Gh) B.1(O), B.1.113(O)</p> <p>B.1.1(G), B.1.36(G) B.1.1(Gr) B.1.1(O)</p> <p>B.1.1(G), B.1.36(G) B.1.1(Gr) B.1.1(O)</p> <p>B.1.1(G), B.1.36(G) A(Gh), B.1(Gh), B.1.113(Gh), B.1.36(Gh), B.1.77(Gh), B.1.9(Gh) B.1(O), B.1.1(O), B.1.113(O), B.1.36(O) A(S)</p> <p>B.1.36(Gh)</p> <p>A(Gh) A(O) A(S)</p> <p>B.1.1(G) B.1(G)</p> <p>B.6(O)</p> <p>B.1.36(Gh) B.6(O)</p> <p>B.6(O)</p>

CHAPITRE II : RESULTATS ET I NTERPRETATION

		B.1(G), B.1.1(G), B.1.113(G), B.1.36(G) B.1(Gh), B.1.113(Gh), B.1.36(Gh), B.1.77(Gh), B.1.9(Gh) B.1.1(Gr) B.1.1(O) A(S) A(S)
--	--	--

Kuwait		B.1.1(GR) B.1(O) B.1.1(Gr) B.1.1(Gr) B.4(O) B.1.1(Gr) B.1(O)
Algérie		B.1(Gh) B.1(Gh) B.1(Gh) B.1(Gh)

Le tableau montre les différentes mutations présentent sur l'ensemble des lignages des pays étudiés, nous remarquons sur le tableau que certaines mutations sont en commun entre les pays tels que les mutations Spike_D614G, N R203K et N G204R qui sont présentes dans tous les pays, et d'autres mutation qui sont présentes dans certains pays et certains non, d'autres aussi sont en commun entre lignages et clades.

II.3 Comparaison des séquences et identifications des mutations :

La comparaison manuelle des séquences étudiés en utilisant logiciel GenieGen a permet de déterminer et identifier 1209 positions de mutations dont 283 sont des mutations fréquentes qui sont comme suivant :

II.3.A Comparaison du Groupe I des pays :

Les résultats de comparaison du premier groupe (Groupe I) des pays sont présentés dans le tableau suivant :

CHAPITRE II : RESULTATS ET I NTERPRETATION

Tableau 02 : les différentes mutations résultantes de la comparaison des séquences du Groupe I des pays.

Position	Clade G	Clade O	Type de mutation	Nombre de répétition	Région touchée
	+	+	Substitution T/C	8 (O) / 2 (G)	ORF1ab
	-	+	Substitution T/C	10	ORF1ab
	-	+	Substitution T/G	9	ORF1ab
	+	+	Substitution C/T	10 (O) / 3 (G)	ORF1ab
	+	+	Substitution G/A	3 (O) / 9 (G)	ORF1ab
	-	+	Substitution A/G	4	Protéine S
	+	+	Substitution G/A	15 (O) / 4 (G)	Protéine S
	+	+	Substitution A/G	11 (O) / 1 (G)	Protéine S
	+	-	Substitution G/A	4	Protéine S
	-	+	Substitution T/G	8	ORF3a
	+	-	Substitution T/G	4	ORF3a
	+	+	Substitution T/G	4 (O) / 3 (G)	ORF3a
	-	+	Substitution C/G	2	Protéine M
	+	-	Substitution C/A	2	Protéine M
	-	+	Substitution C/G	2	ORF8
	-	+	Substitution C/T	6	ORF8
	-	+	Substitution C/T	2	ORF8
	-	+	Substitution C/T	5	Protéine N
	-	+	Substitution C/T	3	Protéine N
	-	+	Substitution C/T	3	Protéine N
	+	+	Substitution A/G	6 (O) / 2 (G)	Protéine N
	+	+	Substitution A/G	5 (O) / 2 (G)	Protéine N
	+	+	Substitution C/G	5 (O) / 2 (G)	Protéine N
	+	+	Substitution A/G	1 (O) / 3 (G)	Protéine N
	+	+	Substitution A/G	1 (O) / 3 (G)	Protéine N
	+	+	Substitution C/G	1 (O) / 3 (G)	Protéine N

Le tableau en dessus présente les mutations identifiées à travers la comparaison manuelle des séquences des deux clades G et O pour le Groupe I des pays ou on observe la présence des mutations sur le clade O et d'autres mutations sur le clade G et d'autres aussi qui

CHAPITRE II : RESULTATS ET I NTERPRETATION

sont en commun entre les deux clades et toutes ces mutations ont des fréquentes variantes touchant de différentes protéines du virus étudié.

II.3.B Comparaison du Groupe II des pays :

Les résultats de comparaison du deuxième groupe (Groupe II) des pays sont présentés dans le tableau suivant :

Tableau 03: les différentes mutations résultantes de la comparaison des séquences du Groupe II des pays.

Position	Clade G	Clade O	Type de mutation	Nombre de répétition	Région touchée
	-	+	Substitution C/T	6	ORF1ab
	-	+	Substitution T/C	7	ORF1ab
	+	+	Substitution C/T	17 (O) / 4 (G)	ORF1ab
	-	+	Substitution C/T	8	ORF1ab
	-	+	Substitution T/G	11	ORF1ab
	-	+	Substitution T/C	4	ORF1ab
	+	+	Substitution C/T	10 (O) / 6 (G)	ORF1ab
	-	+	Substitution A/G	4	Protéine S
	-	+	Substitution A/G	4	Protéine S
	+	-	Substitution G/A	16 (O) / 4 (G)	Protéine S
	-	+	Substitution C/G	2	Protéine M
	-	+	Substitution C/T	3	ORF8
	-	+	Substitution C/T	3	Protéine N
	+	+	Substitution A/G	6 (O) / 2 (G)	Protéine N
	+	+	Substitution A/G	5 (O) / 2 (G)	Protéine N
	+	+	Substitution C/G	5 (O) / 2 (G)	Protéine N

Sur le tableau 3, nous observons les mutations déterminées de la comparaison manuelle des séquences des deux clades G et O pour le Groupe II des pays, il y a ceux qui sont présente que sur le clade O d'autres que pour le clade G et d'autres aussi qui sont en commun entre les deux clades et quelques mutations qui sont en communs entre les deux Groupe I et II.

CHAPITRE II : RESULTATS ET I NTERPRETATION

Parmi les résultats de comparaisons obtenus, nous avons identifiés les 283 positions de mutations présentées sur les deux tableaux précédents et 926 positions de mutations rares, ce qui peut montrer le taux élevé des mutations que ce virus peut subir.

A partir de ces résultats et en prenant en considération les résultats statistiques de l'étude de **Triqui.Ch et al., en presse** nous pouvons suggérer que les mutations présentes sur le clade O peuvent être la cause de diminution de virulence du virus. Et les différences de résultats entre les deux groupes de pays nous permettent de conclure la grande diversité de ce virus.

CONCLUSION GENERALE

Conclusion

Dans notre étude nous nous sommes focalisés sur l'étude phylogénétique et comparaison du COVID-19 qui a eu une large couverture médiatique depuis son émergence en Décembre 2019, plusieurs laboratoires ont consacré un temps et un effort significatif pour but de caractériser le virus par multiple techniques basée sur le génome et pour avoir une vraie approche au génome de SARS-COV 2.

Presque 300 séquences de différents pays de différents continents obtenus de la base de données public GISAID on était analysés pour réaliser une résolution claire et net sur la diversité et l'évolution de virus ainsi les mutations et leur position dans son génome.

Le premier chapitre a été consacré pour une synthèse bibliographique sur les généralités de ce nouveau virus, où nous avons donné une définition générale du SARS-COV 2, son histoire et ses différentes classes.

Au niveau du chapitre II nous avons parlé de la phylogénie et la bio-informatique et ses différents aspects.

Dans le troisième chapitre, nous avons présenté le matériel et les différents logiciels utilisé. Nous avons abordé aussi les étapes de l'analyse phylogénétique et la comparaison des séquences.

Dans le chapitre IV nous avons exposé et discuté les résultats des calculs effectuer dans ce travail.

L'analyse des résultats nous as permis de conclure que :

- La distribution du virus au niveau mondial ne présente en aucun cas au niveau de chaque sous-classe ni la région, ni le lignage, ni le clade ce qui nous conduit à conclure la très grande diversité et donc la grande mutabilité de ce virus.
- Le taux de mutations élevé et les mutations présentes sur le clade O qui peuvent nous permettre de suggérer qui sont la cause de diminution de virulence du virus.

Perspectives :

Ce travail reste une initiation à d'autre études, nous envisageons de faire une modélisation moléculaire afin de mieux comprendre ce virus et déterminer les différentes mutations responsables de la diminution de la virulence.

REFERENCES BIBLIOGRAPHIQUE

Références bibliographie

- Badran, N. (2018). Un nouvel algorithme bioinformatique pour retrouver la relation entre la phylogénie et la phylogéographie des espèces.
- Bonny, V., Maillard, A., Mousseaux, C., Plaçais, L., & Richier, Q. (2020). COVID-19: Physiopathologie d'une maladie à plusieurs visages. *La Revue de Médecine Interne*.
- Bonny, V., Maillard, A., Mousseaux, C., Plaçais, L., & Richier, Q. (2020). COVID-19: Physiopathologie d'une maladie à plusieurs visages. *La Revue de Médecine Interne*.
- Chen, Y., Liu, Q., & Guo, D. (2020). Emerging coronaviruses: genome structure, replication, and pathogenesis. *Journal of medical virology*, 92(4), 418-423.
- DERIHAM, T. E. (2017). *Utilisation des arbres phylogénétiques dans l'alignement de séquence* (Doctoral dissertation, FACULTE: Mathématiques et Informatique-- UNIVERSITE MOHAMED BOUDIAF-M'SILA).
- Eaaswarkhanth, M., Al Madhoun, A., & Al-Mulla, F. (2020). Could the D614 G substitution in the SARS-CoV-2 spike (S) protein be associated with higher COVID-19 mortality?. *International Journal of Infectious Diseases*.
- Elayadeth-Meethal, M. COVID-19 and SARS-CoV-2: Molecular Genetics Perspectives.
- Elayadeth-Meethal, M. COVID-19 and SARS-CoV-2: Molecular Genetics Perspectives.
- Gagnon, J. (2004). Création d'outils pour l'automatisation d'analyses phylogénétiques de génomes d'organites.
- GISAID. Genomic epidemiology of novel coronavirus. Available at <https://nextstrain.org/ncov> 2020 Last accessed 9th March 2020
- Hossain, M. G., Akter, S., & Saha, S. (2020). SARS-CoV-2 host diversity: An update of natural infections and experimental evidences. *Journal of Microbiology, Immunology and Infection*.
- Hossain, M. G., Akter, S., & Saha, S. (2020). SARS-CoV-2 host diversity: An update of natural infections and experimental evidences. *Journal of Microbiology, Immunology and Infection*.

REFERENCES BIBLIOGRAPHIQUE

- Jaimes, J. A., André, N. M., Chappie, J. S., Millet, J. K., & Whittaker, G. R. (2020). Phylogenetic analysis and structural modeling of SARS-CoV-2 spike protein reveals an evolutionary distinct and proteolytically-sensitive activation loop. *Journal of molecular biology*.
- Jaimes, J. A., André, N. M., Chappie, J. S., Millet, J. K., & Whittaker, G. R. (2020). Phylogenetic analysis and structural modeling of SARS-CoV-2 spike protein reveals an evolutionary distinct and proteolytically-sensitive activation loop. *Journal of molecular biology*.
- Junejo, Y., Ozaslan, M., Safdar, M., Khailany, R. A., Rehman, S., Yousaf, W., & Khan, M. A. (2020). Novel SARS-CoV-2/COVID-19: Origin, pathogenesis, genes and genetic variations, immune responses and phylogenetic analysis. *Gene reports*, 20, 100752.
- Kumar, S., Nei, M., Dudley, J., & Tamura, K. (2008). MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in bioinformatics*, 9(4), 299-306.
- Laha, S., Chakraborty, J., Das, S., Manna, S. K., Biswas, S., & Chatterjee, R. (2020). Characterizations of SARS-CoV-2 mutational profile, spike protein stability and viral transmission. *Infection, Genetics and Evolution*, 85, 104445.
- Liu, Q., Zhao, S., Shi, C. M., Song, S., Zhu, S., Su, Y., ... & Chen, H. (2020). Population Genetics of SARS-CoV-2: Disentangling Effects of Sampling Bias and Infection Clusters. *Genomics, proteomics & bioinformatics*.
- MAAROUF, C. *Alignement et recherche des séquences génétiques* (Doctoral dissertation).
- Marzin-Janvier, P. (2015). Étayer la conception expérimentale par des environnements informatiques: études en génétique. *RDST. Recherches en didactique des sciences et des technologies*, (12), 87-112.
- Millett, P., & Rutten, P. (2020). COVID-19, SARS-CoV-2, and Export Controls. *Health security*, 18(4), 329-334.
- OMS. Organisation mondiale de santé. Available at <https://www.who.int/fr/> 2020.
- Philippon, A. (2020). covid 19: du microbiote à l'azithromycine, mythes et réalités. *Bulletin de l'Académie Vétérinaire de France*;
- Pr DJEKOUN A. Pr HAMIDECHI M. A. (2019). Cours de phylogénie moléculaire

REFERENCES BIBLIOGRAPHIQUE

- Rodriguez-Morales, A. J., Rodriguez-Morales, A. G., Méndez, C. A., & Hernández-Botero, S. (2020). Tracing new clinical manifestations in patients with covid-19 in chile and its potential relationship with the SARS-CoV-2 divergence. *Current Tropical Medicine Reports*, 1-4.
- Saha, I., Ghosh, N., Maity, D., Sharma, N., Sarkar, J. P., & Mitra, K. (2020). Genome-wide analysis of Indian SARS-CoV-2 genomes for the identification of genetic mutation and SNP. *Infection, Genetics and Evolution*, 85, 104457.
- Suzuki, Y. (2020). Methods for making multiple alignment of genomic sequences for severe acute respiratory syndrome coronavirus 2. *Meta Gene*, 100785.
- Talavera, G., & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic biology*, 56(4), 564-577.
- Thieffry, D., & De Jong, H. (2002). Modélisation, analyse et simulation des réseaux génétiques-Bio-informatique (3). *médecine/sciences*, 18(4), 492-502.
- Van Dorp, L., Acman, M., Richard, D., Shaw, L. P., Ford, C. E., Ormond, L., ... & Ortiz, A. T. (2020). Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. *Infection, Genetics and Evolution*, 104351.
- Wang, J. T., Lin, Y. Y., Chang, S. Y., Yeh, S. H., Hu, B. H., Chen, P. J., & Chang, S. C. (2020). The role of phylogenetic analysis in clarifying the infection source of a COVID-19 patient. *Journal of Infection*, 81(1), 147-178.

Résumé

Le SARS-COV 2, nommé covid-19, un virus qui fait partie d'un grand groupe de virus connu depuis les années 1960s appelées les coronavirus, a déclenché une pandémie globale en Décembre 2019 qui a eu un impact majeur sur la santé public notamment l'économie des pays les plus touchées, l'infection par ce dernier développe chez les patients une toux, la fièvre et une dyspnée, cette phase d'invasion viral est suivi d'une réaction immunitaire inadaptée marquée par l'aggravation du symptomatologie respiratoire. Il est connu que les chauves-souris sont le réservoir naturel des coronavirus mais jusqu'à présent l'origine de SARSS-COV 2 et sa capacité de se transmettre à l'humain reste jusqu'à présent ambiguë, de ce fait nous avons réalisé une étude phylogénétique entre 11 pays afin de suivre la diversité et l'évolution de ce virus et par la suite une comparaison des séquences des deux clades G et O du Covid-19 a été établit pour déterminer les positions de mutations que ce virus a subis au cours de son évolution, En visant à la fin d'envisager l'impact des mutations sur le comportement du virus et son effet sur les cellules hôtes.

Mots clés : SARS-COV 2, arbre phylogénétique, mutations

Anglais

SARS-cov 2, named covid-19, a virus which belongs to a big group known since the 1960s as the coronaviruses, has declared a global pandemic in December 2019 which had a major impact on the public health as well as the most affected country's economic, the infection by this virus can develops in paients a caught, fever and a dyspnea, this viral infection phase is followed by a dysregulated immune reaction marquedd by the aggravation of the respiratory symptoms. It is known that bats are the natural reservoir of coronaviruses but up until now the origine of SARS-Cov 2 and its capacity to transmet to humains is still ambiguous, for that matter we have realised a phylogenetic study between 11 countrys to track the virus's diversity and its evolution, furthermore a comparison of sequences of two clads G and O was established to determine the position of the mutitaions which this virus had during its evolution course, at the we aim to highlight the impact of the mutations on the virus's behaviour and it's effect on the host cells.

Keywords: SARS-COV 2, phylogenetic tree, mutations

Arabe

المسمى SARS-COV 2. فيروس ينتمي لمجموعة كبيرة من الفيروسات معروفة منذ سنوات ١٩٦٠ تدعى ، بدأ وباء عالمي في ديسمبر ٢٠١٩ الذي كان له تأثير كبير على الصحة العامة و مع ذلك إقتصاد الدول المتأثرة بهذا الفيروس، الإصابة بالفيروس تظهر عند المرضى السعال، حمى و ضيق التنفس، يلي هذه العدوى الفيروسية رد فعل مناعي غير نوعي متمثل في تقاوم الأعراض في الجهاز التنفسي. معروف عن الخفافيش أنها الخزان الطبيعي لفيروسات الكورونا لكن حتى يومنا هذا لا يزال مصدر SARS-COV 2. و قدرته على الإنتقالية إلى الإنسان غامض، و من هذا قمنا بدراسة SARS-COV 2 بين ١١ بلداً لمتابعة تنوع و تطور هذا الفيروس و بالتالي تم إنشاء مقارنة بين السلاسل الجينية التابعة لقسم G و قسم O لفيروس SARS-COV 2 لتحديد مواقع الطفرات التي تعرض لها الفيروس أثناء مرحلة تطوره، الهدف في الأخير تسليط الضوء على تأثير الطفرات على سلوك الفيروس و تأثير هذا الأخير على الخلايا المضيفة.

الكلمات الرئيسية: SARS-COV 2، شجرة الفيتولوجينيك، الطفرات