

TP : Responsabilité collective, Immunité, Génétique et propagation d'une épidémie

Activités avec Tableur, Anagène/Génigène, Rastop, Edumodèle ...

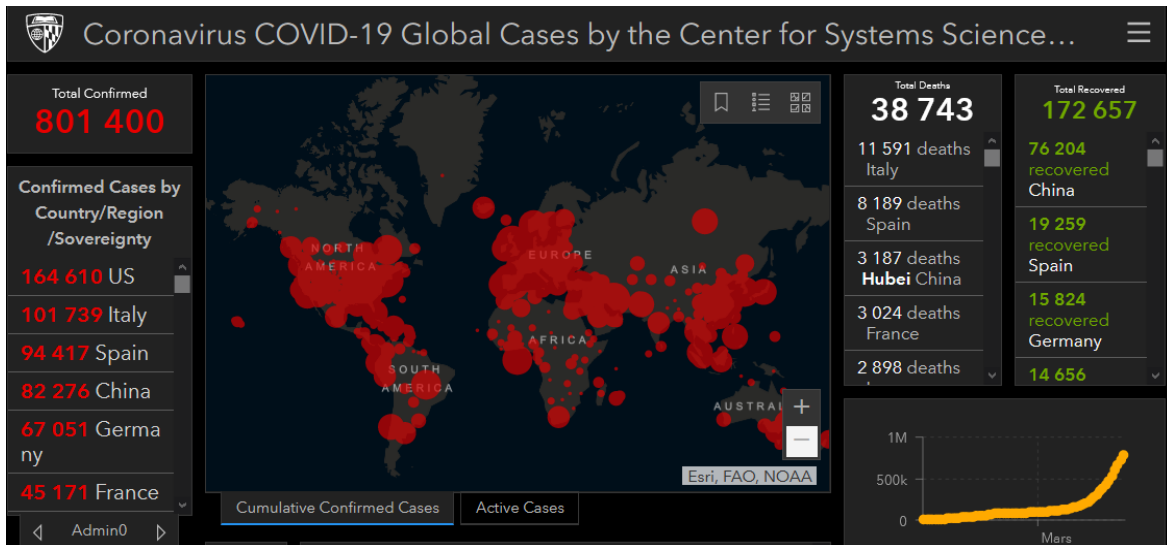
Capacités	Compétences travaillées
<p>– Recenser, extraire et organiser des informations pour identifier :</p> <ul style="list-style-type: none"> * l'origine multigénique de certaines pathologies ; * l'influence de facteurs environnementaux. <p>– Identifier, dans le cas d'une maladie à causalité multifactorielle, les principes, les intérêts et les limites de l'épidémiologie (descriptive ou analytique) et de ses méthodes (étude de cohortes et témoins).</p> <p>– Mener une analyse statistique simple sur des données de santé ; cartographier une pathologie en la visualisant sur un système d'information géoscientifique,</p> <p>– Appréhender de manière critique les conditions de validité d'affirmations lues ou entendues concernant la responsabilité d'un gène ou d'un facteur de l'environnement dans le développement d'une maladie.</p> <p>– Savoir expliciter ses comportements face à un risque de santé pour exercer sa responsabilité individuelle ou collective.</p> <p>– Modéliser et calculer le taux de couverture vaccinale efficace pour un vaccin (par exemple : rougeole).</p> <p>– Prendre conscience que la vaccination est une démarche dans laquelle le bénéfice collectif est très largement supérieur au risque vaccinal individuel.</p>	<p>Pratiquer des démarches scientifiques</p> <ul style="list-style-type: none"> – Formuler et résoudre une question ou un problème scientifique. – Concevoir et mettre en oeuvre des stratégies de résolution. – Observer, questionner, formuler une hypothèse, en déduire ses conséquences testables ou vérifiables, expérimenter, raisonner avec rigueur, modéliser, argumenter. – Interpréter des résultats et en tirer des conclusions. – Comprendre le lien entre les phénomènes naturels et le langage mathématique. <p>Pratiquer des langages</p> <ul style="list-style-type: none"> – Utiliser des outils numériques. – Conduire une recherche d'informations sur internet en lien avec une question ou un problème scientifique, en choisissant des mots-clés pertinents, et en évaluant la fiabilité des sources et la validité des résultats. – Utiliser des logiciels d'acquisition, de simulation et de traitement de données. <p>Adopter un comportement éthique et responsable</p> <ul style="list-style-type: none"> – Identifier les impacts (bénéfiques et nuisances) des activités humaines sur l'environnement à différentes échelles. – Fonder ses choix de comportement responsable vis-à-vis de sa santé ou de l'environnement en prenant en compte des arguments scientifiques. – Comprendre les responsabilités individuelle et collective en matière de préservation des ressources de la planète (biodiversité, ressources minérales et ressources énergétiques) et de santé. – Participer à l'élaboration de règles de sécurité et les appliquer au laboratoire et sur le terrain.

<p>L'évolution permet de comprendre des phénomènes biologiques ayant une importance médicale.</p> <p>L'évolution rapide des organismes microbiens nécessite d'adapter les stratégies prophylactiques, les vaccins et les antibiotiques.</p>	<p>Comparer les valeurs fournies par un modèle à des données réelles afin de tester sa validité.</p>
---	--

Mise en situation ou contextualisation:

Vidéo: [extrait de la bande annonce du film Contagion et présentation de l'activité](#)

Cartographie de la Pandémie



Source: <http://edumed.unice.fr/fr/contents/news/homepage/pandemie-de-covid-19#.XoM7QnLgoUA>

Vidéo de l'INSERM: La réponse immunitaire adaptative (rappel) -

Comment montrer à partir d'un modèle épidémiologique l'influence des facteurs environnementaux et génétiques dans le cas d'une épidémie ? Comment mettre en place des stratégies sanitaires et calculer un taux de couverture vaccinale efficace avec le R0 ?

Vous organiserez votre réponse selon une démarche de votre choix intégrant des données issues des documents (toutes les données moléculaires par exemple ne sont pas à étudier) et les connaissances complémentaires nécessaires.

- expliquez les moyens de lutte contre une pandémie
- comparer les modèles proposés et en dégager les principales observations ;
- montrer comment le coronavirus est passé de l'animal à l'homme et quels sont les moyens de lutte

L'objectif est de représenter par un modèle, des données réelles afin de tester sa validité et d'étudier les particularités du programme génétique du virus et son évolution.

Document 1: Le taux de reproduction de base de la maladie ou R0

Les spécialistes semblent s'accorder sur le fait que chaque malade infecterait entre 2 et 3 personnes (ce qu'on appelle le "taux de reproduction de base" de la maladie, ou R0).

D'autre part, certains mettent en garde sur une sous-estimation possible du nombre de cas. Ainsi une étude de chercheurs du centre des maladies infectieuses de l'Imperial college de Londres *"estime qu'environ les deux-tiers des cas de Covid-19 sortis de Chine sont restés indétectés au niveau mondial, avec pour résultat potentiel des chaînes multiples non-détectées de transmission humaine hors de Chine"*. Cette hypothèse pourrait expliquer l'apparition de foyers multiples en Italie.

Une des préoccupations pourrait être les cas asymptomatiques, avec des patients infectés ne présentant aucun signe de la maladie. La durée d'incubation, estimée selon les études entre deux et 10 jours, a conduit à fixer à 14 jours la période d'observation ou de quarantaine pour les cas suspects et les personnes rapatriées.

Selon Kermack & McKendrick : $R_0 = \rho c d$,

ρ étant la probabilité de transmission, c le nombre de contacts par unité de t et d la durée de la période contagieuse, d'où le temps de doublement : $T_d = d \ln(2) / R_0 - 1$.

Pour les médecins britanniques de l'Imperial College de Londres à l'origine de la première estimation, cela signifie qu'*"il faudrait bloquer plus de 60 % des contaminations par des mesures de contrôle pour maîtriser l'épidémie."* En l'absence de vaccin et de traitement, ces mesures se résument à la détection des malades et à leur isolement.

Ce chiffre reste toutefois une moyenne, la variabilité entre les personnes semble être très importante. En clair, quelques malades vont infecter un grand nombre de personnes tandis que la majorité n'en contaminera aucune.

A titre de comparaison, **la grippe** touche 2 à 6 millions de personnes en France chaque année. Le taux d'incidence des consultations pour syndrome grippal en France, pour la seule semaine 07 de 2020 (du 10 au 16 février 2020), est estimé à 296 pour 100 000 habitants, soit près de 200 000 nouveaux cas. Le nombre attribuable à la grippe est ainsi estimé à près de 150 000 nouveaux cas en une semaine sur le territoire français.

Taux de contagion des différents virus (en partie contrôlé):

- H1N1 : 1,5 - Grippe saisonnière: 1,3 à 2,5 - Sras : 3 <ul style="list-style-type: none">• SRAS (2003-2004): 2-4• Mers (2012): 2-5• Covid-19: 2,5	- VIH : 3,5 - Hépatite B: 5 - Varicelle 8,5 - Rougeole: 12 à 18 - Rotavirus (gastro-entérite): 10 à 18 personnes. - Ebola : 2,5
---	---

Sources : OMS

Document 2: Virulence: La virulence est la capacité du virus à se multiplier dans l'organisme.

Pour le Covid-19, une personne sur 5 environ pour laquelle l'infection a été confirmée développe des symptômes graves. La mortalité avoisinerait les 2.3 % (8 % chez les 70-79 ans et 14,8 % chez les plus de 80 ans). Elle était de 9,6 % pour le SRAS et de 34 % pour le Mers.

En théorie, il suffit de diviser le nombre de morts par le nombre de cas rapportés pour établir ce chiffre. Seulement, en pratique, cette estimation est complexe. D'abord parce qu'on ne meurt pas directement d'une infection virale mais des conséquences de la défaillance d'un ou plusieurs organes qu'elle provoque. Ce qui arrive d'autant plus souvent que la personne est fragile, âgée ou déjà malade (diabètes, cirrhoses, hypertension artérielle ou maladies coronariennes). Ensuite, dans la mesure où chaque personne réagit différemment à l'infection, on peine à estimer le nombre de personnes qui sont malades sans le savoir, soit parce qu'elles ne développent aucun symptôme, soit parce que ceux-ci sont modérés et trop peu spécifiques pour être associés à une maladie en particulier.

Document 3: La modélisation

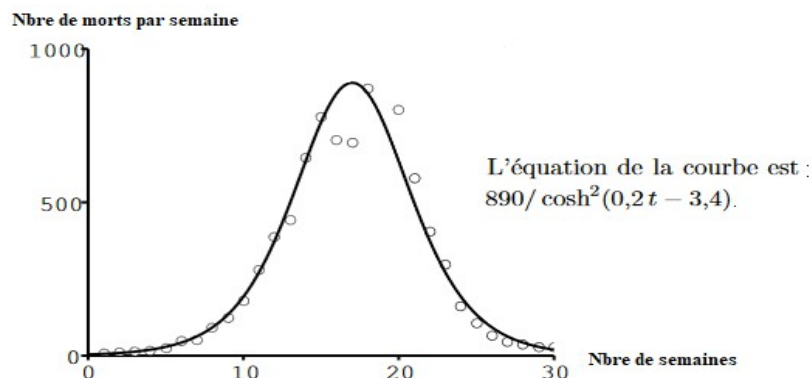
a) Le principe d'une épidémie.

Il y a un démarrage avec un premier cas. Ensuite celui-ci va en générer d'autres. Plus il y en a, plus la contamination s'accélère. L'épidémie s'arrête naturellement lorsque:

- le virus détruit rapidement son hôte: Ebola, MERS-COV30, Marburg ..qui n'a donc pas ou peu de temps pour le reproduire
- le vivier de personnes saines sera remplacé par suffisamment des personnes immunisées (il est possible d'être immunisé soit par l'effet d'un vaccin soit en ayant contracté le virus).

b) Evolution théorique et modélisation à partir des données de terrain:

Une épidémie (ou une pandémie) peut être étudiée par modélisation utilisant des équations.

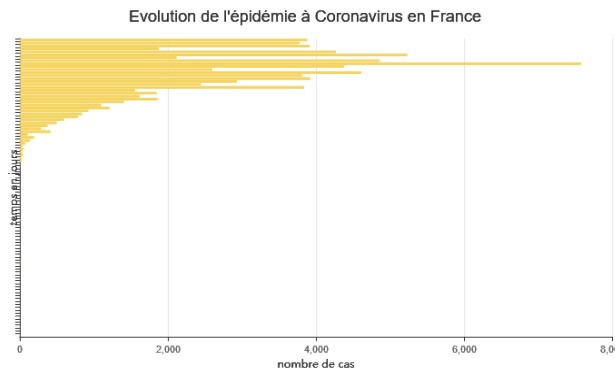


Nombre hebdomadaire de décès dus à la peste à Bombay entre le 17 décembre 1905 et le 21 juillet 1906.

d'après le Modèle de Kermack et McKendrick (1907)

A partir des données fournies ([lien vers le fichier tableur](#) sur le site data.europa.eu).

- Réaliser graphiquement l'évolution de la maladie en France et la comparer avec des valeurs théoriques de la propagation d'une épidémie (peste de bombay, grippe saisonnière...)



- après l'analyse des résultats, PROPOSER DES HYPOTHESES sur l'évolution de la pandémie à coronavirus

c) Modélisation "SIR": Activité facultative pour SVT - Mathématiques (Terminale)

Le modèle de base est appelé SIR, où S désigne, au sein de la population concernée, les individus Sains (ou Susceptibles d'être infectés), I désigne ceux qui sont Infectés et R ceux qui sont Rétablis (*Recovered* en anglais) et ne peuvent plus être infectés, sous l'hypothèse, liée à ce premier modèle, qu'un individu guéri est définitivement immunisé. L'effectif de chacune de ces populations est évidemment variable dans le temps, modélisable de ce fait par une fonction de la variable indépendante t, le temps : S(t), I(t) et R(t). Si, au cours de la propagation de l'épidémie, l'effectif P de la population totale peut être considéré constant, on écrit :

$$S(t) + I(t) + R(t) = P$$

Ce système peut être représenté graphiquement par un ensemble de trois compartiments connectés par des flux d'individus qui passent de l'un à l'autre :

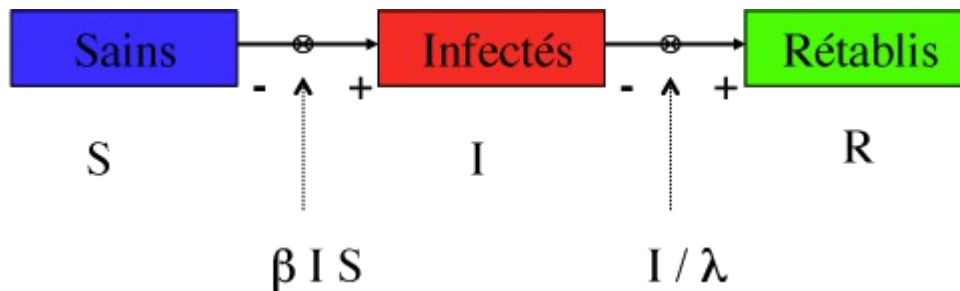


Diagramme des flux d'individus dans le modèle SIR.

L'épidémie se propage par les contacts entre les individus infectés et les individus sains. Le nombre de ces contacts est proportionnel à S et à I, effectifs respectifs des populations d'individus sains et infectés. Les malades guérissent en moyenne au bout d'un temps λ ; ils sont alors immunisés et ne peuvent plus, ni infecter d'autres personnes, ni être réinfectés.

À chaque compartiment est associée une variable d'état : S, I et R. Il s'agit maintenant d'écrire un système d'équations différentielles qui relie la dérivée des fonctions, $dS(t)/dt$, $dI(t)/dt$ et $dR(t)/dt$, aux fonctions elles-mêmes, S(t), I(t) et R(t). Les valeurs de S, I et R sont toujours positives et sans dimension.

La propagation de l'épidémie résulte des contacts contaminants entre les personnes infectées et les personnes saines. L'effectif des infectés s'accroît en fonction du nombre de contacts contaminants entre les individus infectés et les individus sains. Ce nombre est proportionnel à l'effectif de la population infectée et à l'effectif de la population saine, et donc au produit de ces deux effectifs I S.

On peut donc écrire :

$$dI(t)/dt = \beta I S$$

$\beta I S$ est la mesure instantanée (comptée en individus par unité de temps) du flux, qui varie dans le temps, des individus qui passent du compartiment « Sains » au compartiment « Infectés ». Le paramètre $\beta > 0$, appelé taux d'incidence, pourrait s'écrire sous la forme d'un produit de deux paramètres. La valeur du premier rend compte de la proportion des contacts effectifs entre une personne saine et une personne infectée, parmi tous ceux possibles pendant un intervalle de temps dt, soit I(t) S(t) (ce paramètre est relié à la densité de population, par exemple). La valeur du second paramètre rend compte de la probabilité qu'un tel contact transmette la maladie de la personne infectée à la personne saine (probabilité dépendante de la virulence de l'agent infectieux).

L'effectif de la population saine diminue symétriquement : $dS(t)/dt = -\beta I S$

Si un individu reste en moyenne malade λ jours, I / λ est, à tout instant, la mesure du flux d'individus qui guérissent, quittent donc le compartiment « Infectés », viennent s'accumuler dans le compartiment « Rétablis » et ne peuvent plus être contaminés, car immunisés. I / λ a pour dimension [temps⁻¹] ; ici son unité est jour⁻¹. L'équation différentielle qui régit l'évolution de l'effectif des infectés s'écrit : $dI(t)/dt = \beta I S - I / \lambda$

I / λ est la valeur instantanée du flux entre « Infectés » et « Rétablis ». L'effectif des rétablis s'accroît ainsi symétriquement du même flux de valeur I / λ des personnes guéries :

$$dR(t)/dt = I / \lambda$$

Le modèle s'écrit donc :

$$dS(t)/dt = -\beta I S$$

$$dI(t)/dt = \beta I S - I / \lambda$$

$$dR(t)/dt = I / \lambda$$

En additionnant ces trois équations membre à membre, on obtient

$$dS/dt + dI/dt + dR/dt = -\beta I S + \beta I S - I / \lambda + I / \lambda = 0$$

ce qui est conforme au fait que l'effectif de la population P est constant, donc de dérivée nulle :

$$dS/dt + dI/dt + dR/dt = d(S + I + R) / dt = dP/dt = 0$$

Le comportement des variables $I(t)$, $S(t)$ et $R(t)$ de ce modèle peut être obtenu par simulation, c'est-à-dire à travers la résolution numérique des équations différentielles associées. Dans le modèle simulé, les variables S et I ont été normalisées : leur valeur est comprise entre 0 et 1 et représente une fraction de la population totale. Pour toute simulation, il est nécessaire de fixer :

1. la valeur des conditions initiales, c'est-à-dire les valeurs de I , de S et de R au temps $t = 0$, soit $I(0)$, $S(0)$ et $R(0)$, notées I_0 , S_0 et R_0 .
Dans ce premier scénario, aucun individu n'est immunisé au début de la période épidémique : $R_0 = 0$. Et puisqu'à tout instant $I + S + R = P$, il suffit de fixer S_0 , effectif initial de la population d'individus infectés, ici une fraction de la population P , soit une valeur comprise entre 0 et 1.
2. les valeurs des paramètres, ici β et λ .

Ce modèle est en effet particulièrement simple, car il véhicule de très nombreuses hypothèses simplificatrices : pas de mortalité, qu'elle soit liée ou non à la maladie, brassage uniforme de la population, etc.

Mais il est possible de relâcher certaines de ces hypothèses, en les remplaçant par d'autres, moins simplificatrices, et de modifier le modèle en conséquence. Ainsi, un terme $-\mu I$ permet de prendre en compte la mortalité liée à la maladie, notée μ , dans l'équation différentielle qui régit l'évolution de l'effectif des infectés :

$$dI(t)/dt = \beta I S - I / \lambda - \mu I$$

$\mu > 0$ est un paramètre dont la valeur est proportionnelle à la virulence de l'agent contaminant.

Avec l'adjonction de ce terme $-\mu I < 0$, la population totale P ne peut plus rester constante. Elle diminue sous l'effet de cette mortalité. La simulation permet d'explorer l'impact de cette adjonction, selon les valeurs du paramètre μ .

- modifiez la valeur de μ , explicitiez le changement de comportement que vous attendez

- confrontez cette attente au résultat de la simulation.

- pourquoi l'effectif de la population des individus sains se stabilise-t-il à un niveau d'autant plus élevé que le taux de mortalité μ est élevé ?

Modèle SIR étendu. Une période d'incubation est introduite : pendant un laps de temps moyen v , les individus contaminés ne sont pas encore contagieux ; ils forment la population d'effectif C . Un taux de mortalité μ a également été introduit.

Le système devient :

$$dS/dt = -\beta I S$$

$$dC/dt = \beta I S - C / v$$

$$dI/dt = C / v - I / \lambda - \mu I$$

$$dR/dt = I / \lambda$$

La complexité du modèle augmente ; il est de plus en plus difficile de prédire le comportement du système sans en lancer la simulation.

Source: La propagation d'un agent infectieux au sein d'une population est un phénomène dynamique : source <https://interstices.info/>

Document 4: Le modèle en application

Dans ce modèle, deux phénomènes simples sont pris en compte :

- Les gens infectés vont infecter les gens sains.
- Les gens infectés vont progressivement guérir.

Pour cela, 3 paramètres sont utilisés :

- La durée λ (D) de la maladie, pendant laquelle on est contagieux ou durée moyenne d'infection
- Le nombre moyen C de contacts que l'on a chaque jour avec d'autres gens.
- La probabilité β (P) qu'un contact entre un infecté et un sain conduise à une transmission du virus appelé taux de transmission.

Prenons une personne infectée, chaque jour elle va croiser C personnes, qu'il contaminera avec une probabilité P. Et cela se produira pendant chacun des D jours que durera sa maladie.

Le nombre total de personne qu'il contaminera sera donc le produit de ces trois termes, que l'on note traditionnellement R_0 avec $R_0 = C \times P \times D$

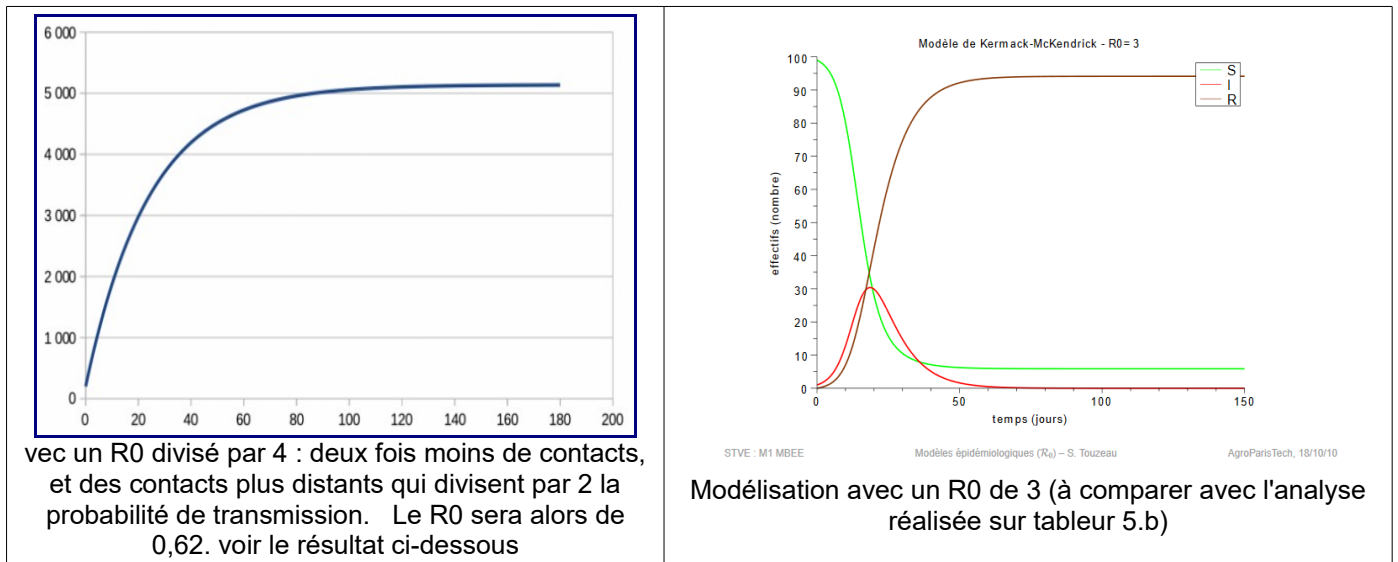
On appelle ce paramètre le taux de reproduction, il a une influence déterminante sur le devenir de l'épidémie.

S'il vaut disons 2 : chaque infecté contaminera 2 personnes, qui elles-même contamineront 2 personnes, qui elles-même contamineront 2 personnes etc. On a une réaction en chaîne, le nombre de malades augmente de façon exponentielle, l'épidémie explose (voir le graphique du document 9.a).

Maintenant si ce coefficient est inférieur à 1 : chaque infecté refilera la maladie à moins d'une personne, donc le nombre net de malade diminuera et progressivement l'épidémie s'éteindra.

Cette valeur peut être intrinsèque à la maladie si aucune action n'est entamée pour la contraindre avec différents niveaux des comportements (gestes barrière, hygiène, limitation des regroupements, confinement).

Prenons un R_0 de 2,5. On peut l'obtenir en disant que la maladie dure 10 jours, et que chaque jour on a 50 contacts avec une probabilité de transmission de 0,5%, c'est donc le produit des deux qui est important. En 6 mois, la majeure partie de la population aura été contaminé.



Source: d'après "Épidémie, nuage radioactif et distanciation sociale – Par David Louapre" [Science étonnante, David Louapre](#), 12-03-2020

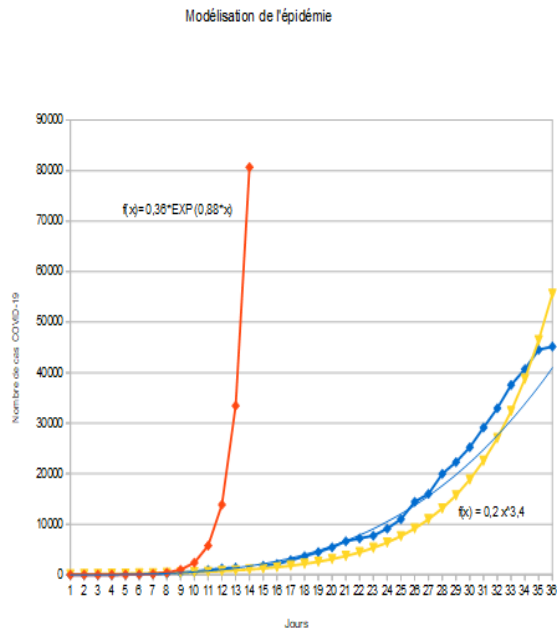
Sources: *Modeles épidémiologiques (R0)*, Suzanne TOUZEAU -INRA

Document 5: analyse de 3 simulations épidémiologiques:

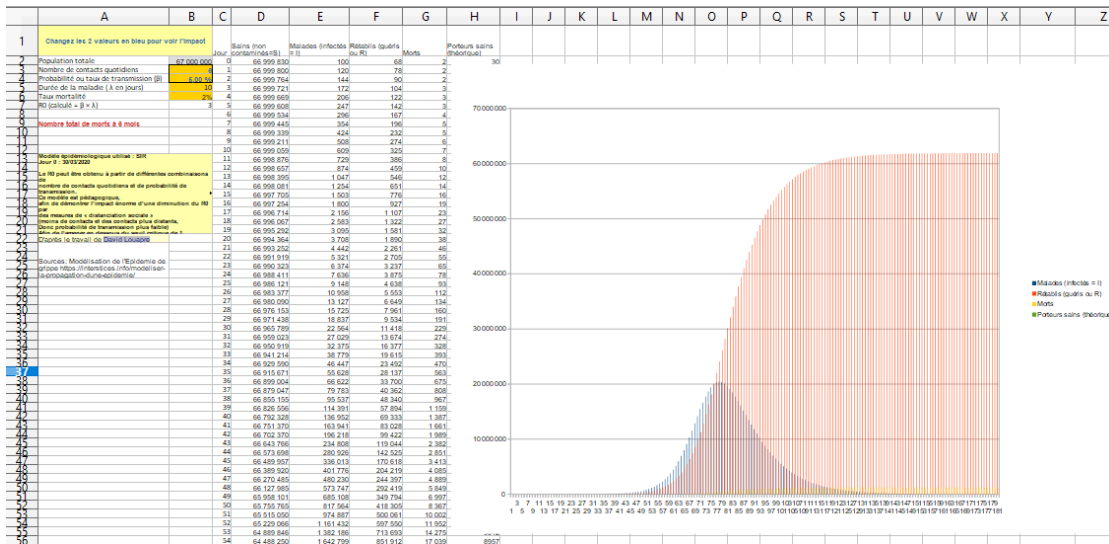
- croissance exponentielle (peu ou pas contrôlée) en rouge (5.a) construite sur les premiers résultats de l'épidémie
- nombre de cas diagnostiqués après application des stades 3 et 4 en bleu (5.a)
- nombres de malades estimés par le modèle SIR et estimation du pic épidémiologique (5.b) en jaune (R_0 de 3, ici avec nombre de contacts 6, taux de transmission 5%)

Doc 5.a: données mises à jour à télécharger sur <https://www.data.gouv.fr/fr/datasets/coronavirus-covid19-evolution-par-pays-et-dans-le-monde-maj-quotidienne/> ou sur Data Europa: <https://data.europa.eu/euodp/fr/data/dataset/covid-19-coronavirus-data>

1	jour	date	nombre de cas	simulation (Exponentielle)	Nombre de malades - Modél. SIR
2	1	24/02/20	1	1	100
3	2	25/02/20	2	2	120
4	3	26/02/20	5	5	144
5	4	27/02/20	11	12	172
6	5	28/02/20	38	29	206
7	6	29/02/20	73	71	247
8	7	01/03/20	178	170	296
9	8	02/03/20	285	411	354
10	9	03/03/20	423	991	424
11	10	04/03/20	613	2388	508
12	11	05/03/20	949	5758	609
13	12	06/03/20	1248	13882	729
14	13	07/03/20	1425	33468	874
15	14	08/03/20		80888	1047
16	15	09/03/20	1784	194831	1254
17	16	10/03/20	2124	468998	1503
18	17	11/03/20	2876	1130701	1800
19	18	12/03/20	3861	2726008	2168
20	19	13/03/20	4500	6572131	2583
21	20	14/03/20	5437	15844750	3095
22	21	15/03/20	6633	38200102	3708
23	22	16/03/20	7200		4442
24	23	17/03/20	7730		5321
25	24	18/03/20	9134		6374
26	25	19/03/20	10995		7638
27	26	20/03/20	14459		9148
28	27	21/03/20	18018		10958
29	28	22/03/20	19980		13127
30	29	23/03/20	22300		15725
31	30	24/03/20	25233		18837
32	31	25/03/20	29155		22564
33	32	26/03/20	32954		27029
34	33	27/03/20	37575		32375
35	34	28/03/20	40704		38779
36	35	29/03/20	44550		46447
37	36	30/03/20	45171		55628
38					
39					



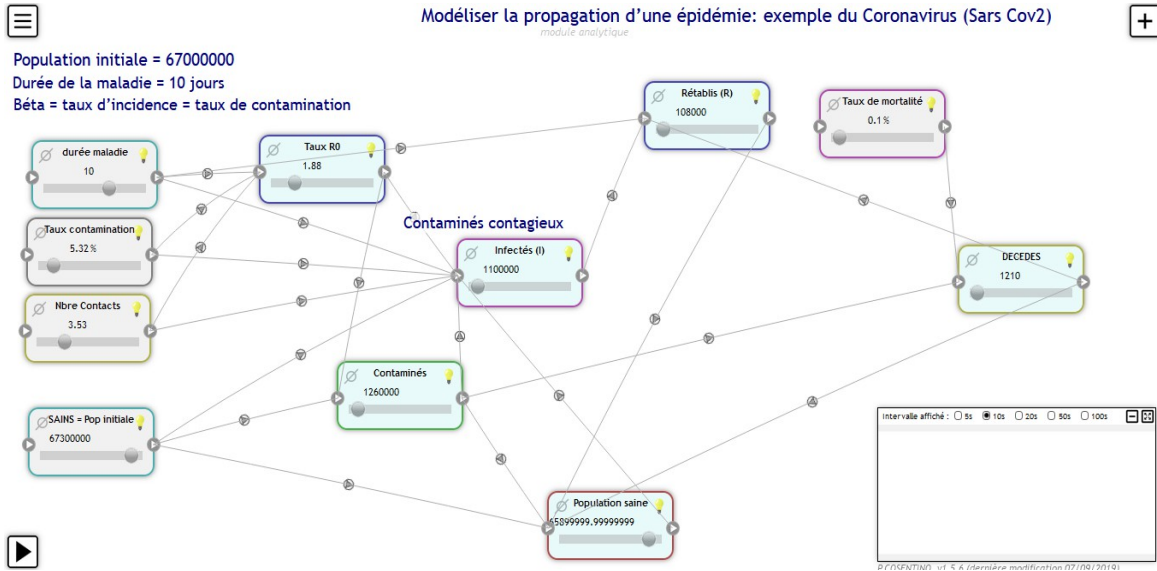
5.b



Questions et "guide": A partir de l'étude des documents 4 et 5

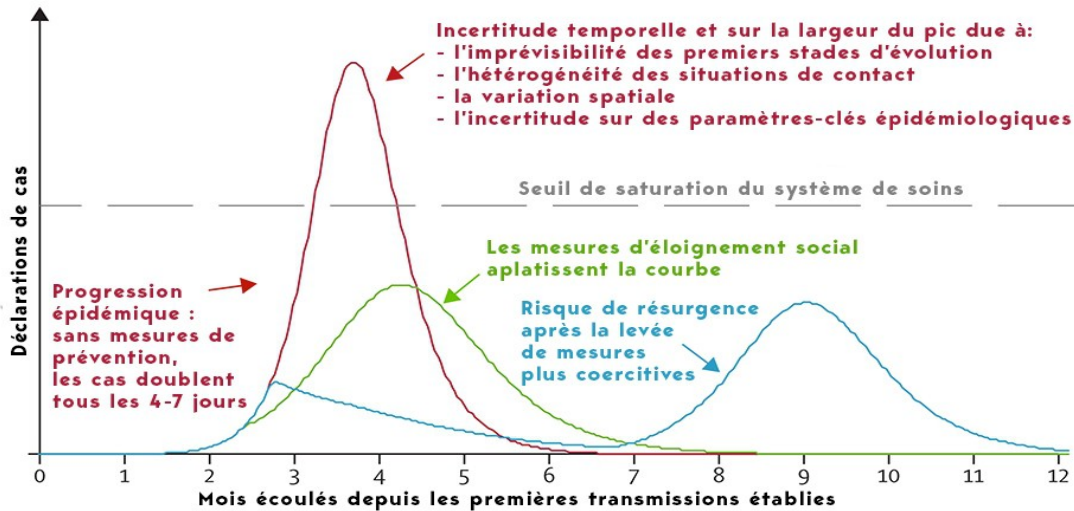
- charger le fichier "coronavirus_epidemie_sir.xls" ou ".ods" de votre dossier
- Modifier les paramètres, Nombre de contacts quotidiens, Probabilité ou taux de transmission (β), Durée de la maladie (λ en jours), Taux mortalité, qu'observez vous sur le graphique ?
- Quel est l'objectif des mesures de distanciation sociale ?
- Montrer l'existence d'un seuil
- Guide:** Avant de fixer des valeurs pour les conditions initiales et les paramètres et visualiser l'évolution des variables S, I, R ..
- explicitez le comportement que vous attendez
- confrontez votre attente aux résultats de la simulation;
- cherchez à comprendre les éventuelles différences.
- Aviez-vous prévu que les trois variables atteindraient un équilibre (où leur dérivée est nulle)?
- Êtes-vous surpris par l'existence d'un pic de l'effectif des infectés ?
- Êtes-vous surpris par le comportement logistique de l'effectif des rétablis ?
- Observez l'impact des changements des valeurs des paramètres sur le comportement du système.
- La simulation est-elle en accord avec votre compréhension du phénomène ?
- Si non, continuez jusqu'à ce que ce soit le cas, ou bien remettez le modèle en cause...

c) Edumodèle: fiche technique dans votre dossier



Document 6: stratégie de prévention : freiner au maximum la propagation du virus

<https://www.univ-rennes1.fr/actualites/coronavirus-reperes-et-point-sur-les-connaissances-avec-le-pr-matthieu-revest-infectiologue>



Source : [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)30567-5/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30567-5/fulltext)

Question: comparer vos résultats à ce document de référence

- Ce document de référence est-il en accord avec vos résultats précédents, expliquez pourquoi ?
- pourquoi y a-t-il un risque de résurgence ?
- Comparer ces données avec un modèle appliquant la vaccination

Document 7: Population proactive ou comment agir face à la maladie

Il suffit qu'une proportion donnée de la population acquière une immunité protectrice, par une infection ou par vaccination, pour qu'elle serve de tampon au reste de la population. Ce principe, appelé immunité grégaire, repose sur le fait qu'en diminuant le nombre d'individus infectés et infectieux, on diminue significativement la probabilité qu'un individu sensible soit infecté.

Si p est la proportion de la population qu'il faut immuniser pour bloquer le déclenchement d'une épidémie $p > 1 - 1/R_0$. Ainsi, pour la grippe $p = 50-75\%$, pour la rougeole $p = 93-95\%$.

Selon Kermack & McKendrick $R_0 = \rho c d$, ρ étant la probabilité de transmission, c le nombre de contacts par unité de t et d la durée de la période contagieuse. Les autorités peuvent ainsi agir sur R_0 par ses trois composantes :

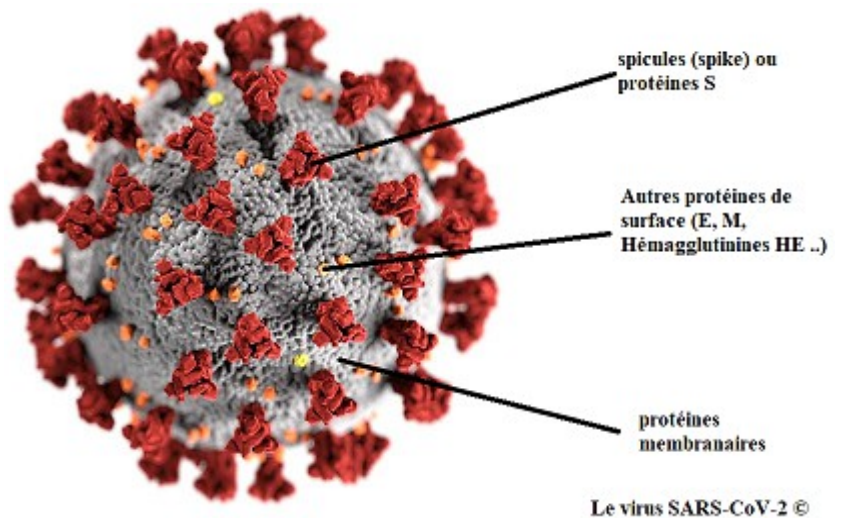
- diminution de la probabilité de transmission par des mesures d'hygiène et de prophylaxie (utilisation d'antiviraux curatifs et préventifs, vaccins, lavage des mains, masques de protection) ;
- diminution du taux de contacts par des incitations à diminuer la vie sociale (augmentation de la « distance sociale » par fermeture des crèches, des écoles, l'isolement des malades, la suppression des rassemblements publics, la réduction des mouvements de population, etc.) ;
- diminution de la durée de contagiosité (utilisation d'antiviraux, de médicaments à visée symptomatique).

Il existe depuis longtemps des campagnes pour sensibiliser la population, comme la Journée mondiale de lutte contre le SIDA, ou pour inciter à la vaccination.

Document 8: Génomes, origine et épidémiologie des nouveaux Coronavirus

Le coronavirus humain 229E (HCoV-229E) est responsable de rhumes mais peut entraîner de graves complications respiratoires chez les personnes âgées ou atteintes d'une maladie chronique. Les coronavirus sont des virus enveloppés avec un génome à ARN positif simple brin. Trois protéines virales sont ancrées dans l'enveloppe virale : la protéine spike (S), la protéine de membrane (M) et la protéine d'enveloppe (E). Les protéines M et E sont impliquées dans l'assemblage viral et la sécrétion.

La protéine du pic d'enveloppe (S) assure la liaison aux récepteurs et la fusion membranaire. Elle est cruciale pour déterminer le tropisme (réaction d'orientation ou d'attraction) vers l'hôte et la capacité de transmission.



On compare la séquence partielle de la protéine S et de ses récepteurs ACE2 de plusieurs souches de virus COV2. La protéine ACE2 est susceptible d'intervenir dans des mécanismes de résistance.

Données moléculaires : [séquences proposées](#) (source Accés ENS Lyon)

a) À partir des neuf échantillons de patients analysés, huit séquences ARN complètes ont été obtenues :

Vous analyserez les séquences.edi fournies ou à télécharger avec Génigen ou Anagène

- Séquences nucléiques de personnes originaires de Chine. (Génomes complets.Sars-CoV-2-China-Adn.edi)

La première capture d'écran montre l'interface 'Séquences' d'un logiciel. Elle contient une liste de séquences avec des colonnes pour l'identification, la base sous le curseur, et des boutons pour naviguer dans les séquences. Les séquences listées incluent MG772933.1-Bat-SA, MT049951.1-China, MT093631.2-China, MT121215.1-China, MT123290.1-China, MT123291.2-China, MT123292.2-China, MT123293.2-China, MT135041.1-China, MT135042.1-China, MT135043.1-China, MT135044.1-China, et NC_045512.2-China. Une légende indique 'Première séquence cochée (éditable) : MG772933.1-Bat-SARS-like'.

La deuxième capture d'écran montre l'interface 'Alignement de Séquences'. Elle présente une vue de séquences alignées avec des lignes de séquences et des colonnes de bases nucléiques (A, C, G, T) alignées. Les séquences alignées incluent MT123291.2_C, MT123293.2_C, MT121215.1_C, NC_045512.2_C, MT123290.1_C, MT049951.1_C, MT135041.1_C, MT135043.1_C, MT135042.1_C, MT135044.1_C, MT123292.2_C, et MG772933.1-Bat.

- Séquences protéiques de personnes originaires de Chine. (Sars-CoV-2-China-Pro.edi)

Séquences	MT123291_2_-	MT123293_2_-	MT121215_1_-	MT093631_2_-	NC_045512_2_-	MT123290_1_-	MT049951_1_-	MT135041_1_-	MT135043_1_-	MT135042_1_-	MT135044_1_-	MT123292_2_-	MG772933.1-B
MT123291_2_-	0	16	19	18	17	19	22	21	22	21	21	22	3658
MT123293_2_-		0	20	19	18	20	23	22	23	22	22	23	3659
MT121215_1_-			0	16	15	17	20	19	20	19	19	20	3656
MT093631_2_-				0	14	16	19	18	19	18	18	19	3655

Alignement de Séquences : données concernant les séquences sélectionnées

Longueur totale après alignement : 29876
 Longueur commune (en éliminant les extrémités où toutes les séquences ne sont pas représentées) : 29873
 Nombre de positions identiques : 26203
 Pourcentage de positions identiques (*): 87,71 %
 Nombre de positions différentes : 3670
 Pourcentage de positions différentes (*): 12,29 %

Séquence de Référence : MT123291_2_- _China.adn

Séquence MT123293_2_- _China.a comparée à la séquence de référence
 Nombre de positions différentes : 16
 Pourcentage de positions différentes : 0,05 %

Séquence MT121215_1_- _China.a comparée à la séquence de référence
 Nombre de positions différentes : 19
 Pourcentage de positions différentes : 0,06 %

Séquence MT093631_2_- _China.a comparée à la séquence de référence
 Nombre de positions différentes : 18
 Pourcentage de positions différentes : 0,06 %

Calcul des...
 différences
 ressemblances

Choix techniques

...en
 nombres
 pourcentages

b) Comparaisons des régions de fixation au récepteur ACE2

Des études antérieures ont montrées que les différents coronavirus (SARS-CoV29, MERS-CoV.30 ..) se lient sur un récepteur appelé ACE2.

Fichiers : sars-CoV2, Sars-CoV1, RaTG13 et Pangolin: Domaine de fixation au récepteur ACE.edi

Affichage des séquences

Spike-CoV-2	0	CysProPheGlyGluUaIPheAsn
Spike-CoV-1	0	CysProPheGlyGluUaIPheAsn
Spike-RaTG13	0	CysProPheGlyGluUaIPheAsn
Spike-Pangolin	0	CysThrPheGlyGluUaIPheAsn

Sélection : 0/4 lignes

Comparaison simple

Traitement	0	Comparaison simple de sé
Spike-CoV-2	0	CysProPheGlyGluUaIPheAsn
Spike-CoV-1	0	- - - - -
Spike-RaTG13	0	- - - - -
Spike-Pangolin	0	- Thr- - - - -

Sélection : 5/5 lignes

Informations

soit 11,0 % de différence

le signe - représente les identités

Spike-Pangolin

Séquence peptidique

longueur : 181 acides aminés

-> 6 a.a. différents de la séquence de référence Spike-CoV-2,

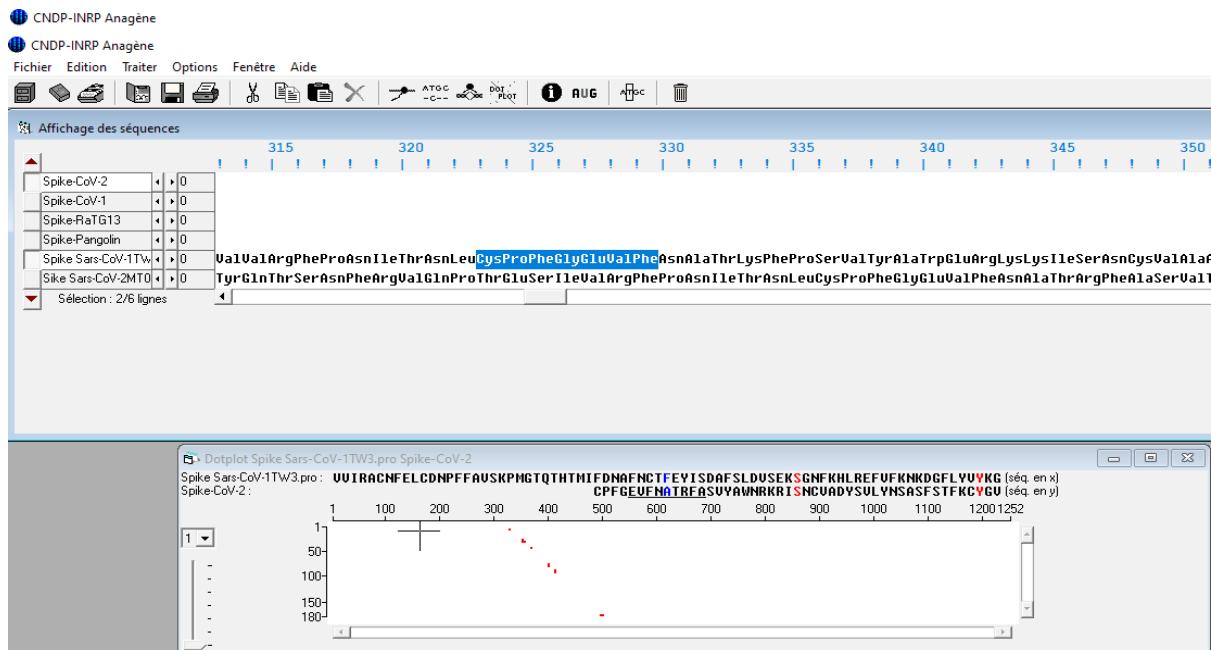
soit 96,7 % d'identité

Commentaires :

Comparaison simple de séquences peptidiques

- Fichier des "Spike" (Spicules) Sars-Cov-1 et Sars-CoV-2 Complets avec la zone de fixation au récepteur ACE2.
 Séquences protéiques: Spike Sars-CoV2-CoV1-Complets.edi

- Visualisation (option dot plot) de la zone de fixation au récepteur ACE2 dans les "Spike" de Sars-CoV-1 et Sars-CoV-2



c) Représentation actualisée des relations de parenté

- En tant que virus à ARN typique, le taux d'évolution moyen des coronavirus est d'environ 10 à 4 substitutions de nucléotides par site et par an, avec des mutations survenant au cours de chaque cycle de réplication. Vos résultats sur l'étude de la population chinoise permet elle de le vérifier ?

Comparaison des spicules (protéines des enveloppes) pour établir une phylogénie (une parenté / origine) :

HKUI: coronavirus humain de Hong Kong (2005)

OC43: coronavirus du rhume humain

1TW3: coronavirus du SARS-COV29

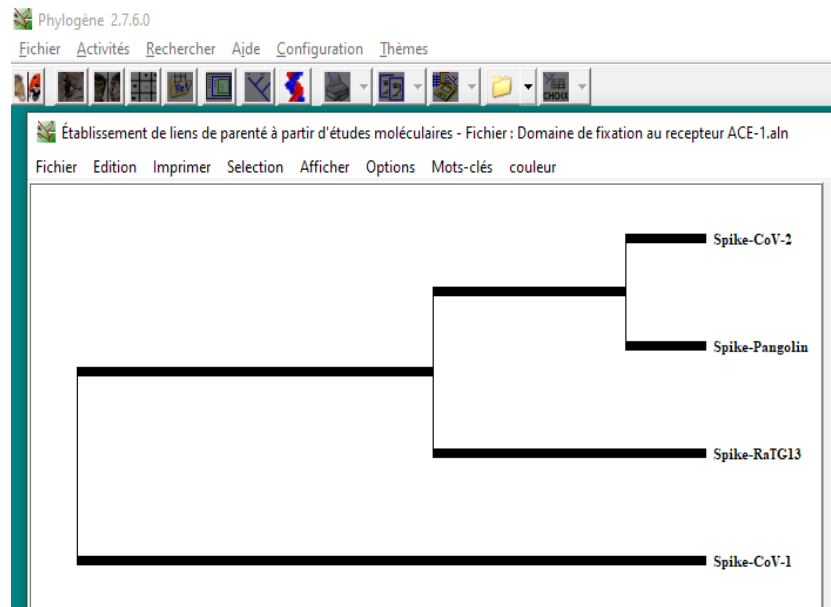
RaTG13: coronavirus de la chauve-souris

NL53: coronavirus de la civette et de la chauve-souris présent chez l'homme (infection respiratoire) depuis près de 1000 ans

229E : coronavirus trouvé chez l'homme (rhume saisonnier) et la chauve-souris

d) Arbres et matrices des distances avec Phylogène: [séquences pour Phylogène](#)

- Séquences nucléiques alignées des coronavirus Sars-CoV-1, Sars-CoV-2 (Covid-19) et MERS chez l'homme, ainsi que de la chauve-souris RaTG13 et du pangolin (Génomes complets.Sars-Cov-2.aln)



- Séquences protéiques des "Spike" chez 16 individus (Coronavirus humains HCov-pro.aln) afin de suivre les mutations spontanées

- Fichier de différents coronavirus. Séquences protéiques alignées (Sars-CoV-2-Surface protéine.aln)

- Fichier de la région de fixation au récepteur ACE2 pour sars-CoV2, Sars-CoV1, RaTG13 et Pangolin.

Les analyses suggèrent que RaTG13 et SARS-CoV2 partageaient un ancêtre commun il y a 25 à 65 ans (estimation faite à partir de la différence de nucléotides entre les virus et le taux présumé de mutation chez d'autres coronavirus). Il aura ainsi fallu **plusieurs décennies pour que les virus RaTG13 et SARS-CoV2 divergent à partir de cet ancêtre commun.**

Les analyses des séquences de SARS-CoV2 isolé de différentes personnes infectées montre une différence d'au plus quelques nucléotides. **Cela suggère que la transmission chez l'Homme s'est faite très récemment** puisque plus un virus circule dans une population humaine, plus il a le temps d'accumuler des mutations. **Cela suggère aussi que le SARS-CoV2 n'a pas eu besoin de muter pour s'adapter et se propager.**

- **Quelles sont les comparaisons moléculaires qui confirment des informations ?**

Document 9: La protéine S

Généralement, la protéine en pointe, ou spicule (spike), des coronavirus est fonctionnellement divisée domaine S1 (en particulier les positions 318 à 510 du SARS-CoV), responsable de la liaison aux récepteurs, et le domaine S2, responsable de la fusion de la membrane cellulaire

La protéine 2019-nCoV S2 a montré environ 93% d'identité de séquence avec bat-SL-CoVZC45 et bat-SL-CoVZXC21 des chauve-souris - beaucoup plus élevé que celui du domaine S1, qui n'avait que 68% d'identité avec ces virus dérivés de chauves-souris.

La variation en acides aminés dans la protéine S des coronavirus

CoV 2019 et SARS-CoV possèdent environ 50 acides aminés conservés dans S1, tandis que la plupart des virus dérivés de chauves-souris présentaient des différences de mutation .

Cependant, malgré l'importance des chauves-souris, plusieurs faits suggèrent qu'un autre animal agit comme hôte intermédiaire entre les chauves-souris et les humains.

- **Proposer des analyses et des résultats de comparaisons moléculaires qui validerait le fait que le nCoV 2019 à l'origine de l'épidémie de Wuhan pourrait également être initialement hébergé par des chauves-souris et aurait pu être transmis à l'homme via des animaux sauvages actuellement inconnus vendus au marché des fruits de mer de Huanan. .**

- **Montrer que la similitude structurelle entre les domaines de liaison aux récepteurs du SARS-CoV et du 2019-nCoV suggère que 2019-nCoV pourrait utiliser ACE2 comme récepteur, malgré la présence de mutations d'acides aminés dans le domaine de liaison au récepteur 2019-nCoV.**

source: Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding , Prof Roujian Lu, MSc *, Xiang Zhao, MD *, Juan Li, PhD *, Peihua Niu, PhD *, Bo Yang, MSc *, Honglong Wu, MSc *, et al. Published: January 30, 2020 - [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)30251-8/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30251-8/fulltext)

Document 10: inhibiteurs thérapeutiques

La protéine S s'assemble en trimères à la surface des virions et joue un rôle-clé dans l'entrée du virus dans sa cellule-cible. Elle est constituée de deux domaines, le domaine S1 responsable de la liaison du virus à son récepteur et le domaine S2 responsable de la fusion de l'enveloppe virale avec une membrane cellulaire. La fusion est activée par des **protéases cellulaires** par clivage de la protéine S. cette protéase indispensable à l'infection par le virus serait la « transmembrane protéase, sérine 2 » (TMPRSS2). Cette enzyme est impliquée dans de nombreux processus physiologiques, mais aussi pathologiques. D'où l'idée des biologistes de tester le camostat, un inhibiteur connu des TMPRSS2 (source : [https://www.cell.com/cell/fulltext/S0092-8674\(20\)30229-4](https://www.cell.com/cell/fulltext/S0092-8674(20)30229-4)).

Fichiers pdb des molécules impliquées:

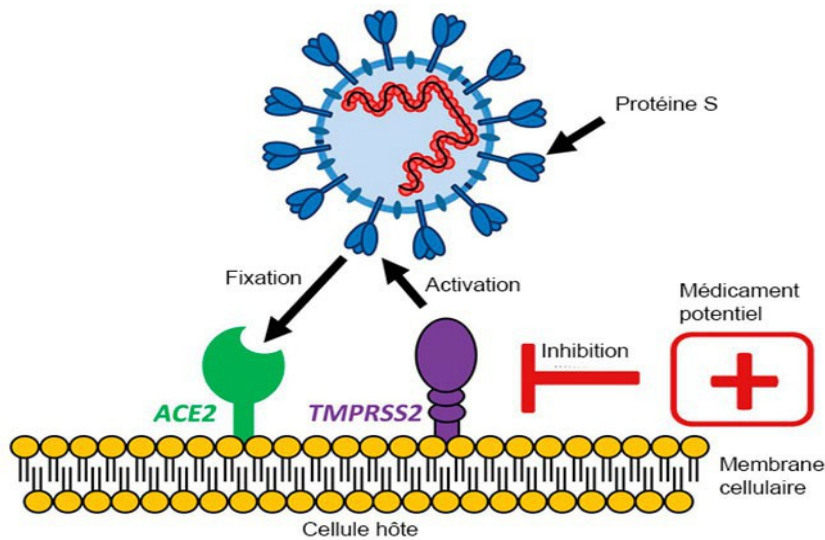
6lu7.pdb : The crystal structure of COVID-19 main protease in complex with an inhibitor N3.

6ACG : Trypsin-cleaved and low pH-treated SARS-CoV spike glycoprotein and ACE2 complex, ACE2-bound conformation 1

6ACJ : Trypsin-cleaved and low pH-treated SARS-CoV spike glycoprotein and ACE2 complex, ACE2-bound conformation 2

6ACK : Trypsin-cleaved and low pH-treated SARS-CoV spike glycoprotein and ACE2 complex, ACE2-bound conformation 3

2AJF : Structure of SARS coronavirus spike receptor-binding domain complexed with its receptor



Vaccins et traitements contre le Covid-19 : les principales pistes, Pour la sciences,Loïc Mangin| 26 mars 2020

Le mécanisme d'infection d'une cellule par le SARS-CoV-2. La protéine S (*en bleu*) reconnaît le récepteur cellulaire ACE2 (*en vert*) lorsqu'il a été activé par la protéine TMPRSS2 (*en violet*). Cette dernière est une cible potentielle pour de possibles médicaments (*en rouge*). © S. Pöhlmann

D'autres molécules peuvent inhiber la fixation du virus sur ses récepteurs ACE2

site de fixation de divers composés pharmaceutiques ou naturels	acides aminés impliqués	
site actif du Covid 19 (liaison avec des ponts hydrogène) sur le récepteur cellulaire	THR24, THR26, PHE140, ASN142, GLY143, CYS145, HIS163, HIS164, GLU166, HIS172	
zone de fixation de la peptidase du coronavirus	LYS5, ALA7, THR25, HIS41 , MET49, TYR54, VAL125, TYR126, GLY127, PHE140, LEU141, ASN142, GLY143, SER144, CYS145, HIS163, HIS164, MET165, GLU166, LEU167, PRO168, HIS172, ASP187, ARG188, GLN189, GLN192, ALA198, LYS236, TYR237, GLN273	
Nelfinavir (antiviral)	Glu166, Gln189, and Gln192	
Lopinavir (antiviral)	Glu166, Arg188, and Gln189	
Quercetine (molécule issue de certains végétaux comme les capres, piment, sureau noir ...)	His164, Glu166, Asp187, Gln192, Thr190	
Rutine (capre olive noire ...)	Tyr 54,Cys 145, Asn 142, Phe 140, Glu 166, Gln 192, Thr 190, Asp 187, His 164	
Apiine (issu du persil et du céleri)	Ser 144, Leu 141, Cys 145, Thr 26, Thr 190, Phe 140, Asn 142, Leu 141,His 41, Arg 188, Gln 189, Met 165, Pro 168	

source: Potential Inhibitor of COVID-19 Main Protease (Mpro) from Several Medicinal Plant Compounds by Molecular Docking Study - Siti Khaerunnisa 1,*, Hendra Kurniawan 2,3*, Rizki Awaluddin 4, Suhartati Suhartati5, Soetjipto

Questions:

- A partir des [données moléculaires en 3D sélectionnées et analysées sous Rastop](#) montrez comment la cible de ces inhibiteurs à été évaluée ?
- quelles sont les cibles moléculaires recherchées par les biologistes pour se prémunir du coronavirus ?
- comparer les acides aminés impliqués dans la fixation du Cov19 et proposez la ou les meilleures substances thérapeutiques pharmaceutiques ou naturelles pour bloquer la fixation du virus
- visualisez sous Rastop ces sites de fixation en mettant en évidence les acides aminés impliqués