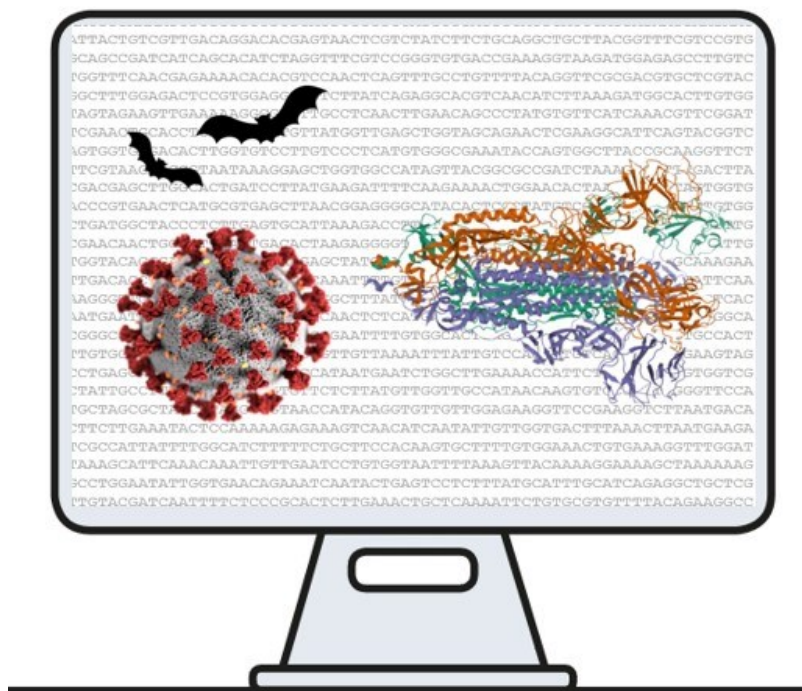


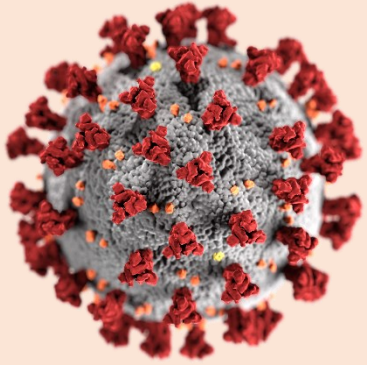
A la découverte du nouveau coronavirus et de la bioinformatique...

*Cet atelier permet de
... découvrir SARS-CoV-2 et des
banques de données & outils
bioinformatiques utilisés par les
chercheurs du monde entier!*



*...de comprendre l'importance
d'avoir librement accès à toutes
les données (**open data**)*

*...de mieux comprendre les démarches
scientifiques et les défis...*



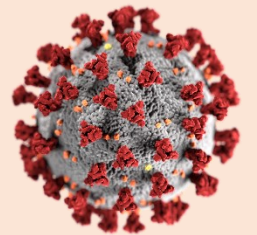
1 - Introduction

Biologie

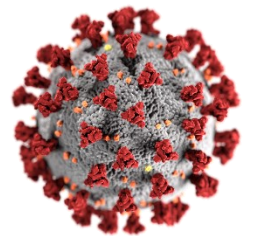
Des virus ?

Des virus qui infectent l'homme ?

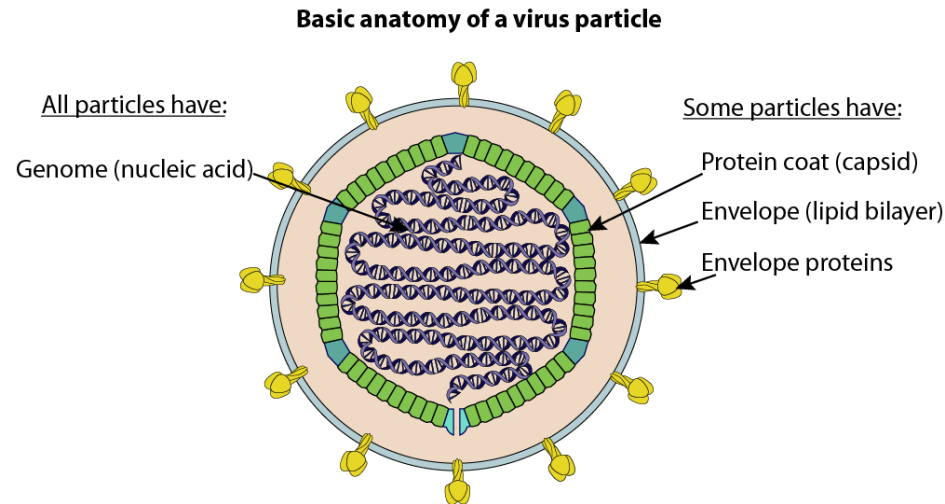
SARS-CoV-2 ?



C'est quoi un virus ?



Un virus est un agent parasite transmis par une particule microscopique qui contient du matériel génétique (ADN ou ARN). Un virus ne peut se répliquer qu'en pénétrant dans une cellule et en utilisant la machinerie cellulaire.



Il existe des virus qui infectent des animaux et d'autres qui infectent des végétaux. Des virus peuvent aussi infecter les bactéries: ce sont les bactériophages.

S'ils provoquent des maladies, les virus sont considérés comme pathogènes.

Des virus, des chiffres et nous...

Un article de 2011 publié dans Nature Microbiology estimait qu'il y avait plus d'un quintillion - un suivi de 30 zéros- de virus sur Terre.

"Ce sont les organismes les plus diversifiés de notre planète (...) et nous ne savons encore rien d'eux".

Chaque jour, des virus tombent du ciel : des dizaines de millions de bactéries et des milliards de virus sont déposés sur un mètre carré...

Nous touchons des centaines de millions de virus chaque jour.

8 % du génome humain est d'origine virale. Une protéine d'origine virale, la syncytine a joué un rôle clé dans l'évolution des mammifères, permettant l'apparition du placenta.

[Autres protéines humaines d'origine virale](#) (UniProtKB)

Un virus peut produire plus de 1000 copies de lui-même par jour

Les virus qui infectent les bactéries (bactériophages) jouent un rôle important pour en régulant les populations de bactéries nocives pour l'homme. Ils pourraient être une alternative aux antibiotiques dans certains cas (phagothérapie).

Le monde vivant ne pourrait pas exister sans les virus !

"Nous avalons plus d'un milliard de virus chaque fois que nous allons nager"

Les virus jouent un rôle clé dans la régulation du carbone dans l'océan.

Vidéo 1 (2:29)

C'est quoi un virus ?

Quel(s) rôle(s) jouent les virus dans l'environnement ?



Des virus qui infectent l'homme ?

850'000 virus susceptibles de passer du règne animal à l'homme...

6 grandes pandémies depuis le début du 20ème siècle...

Importance de la protection de l'environnement...(RTS, 30 nov 2020)

Plusieurs milliers de virus différents 'infectent' l'homme

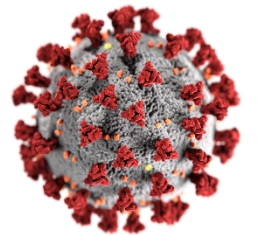
(on ne les a pas encore tous découverts....),

mais seulement ~ 160 sont responsables d'une pathologie et ont donc été bien étudiés

(grippe, SIDA, herpès, hépatite, verrue, ...)

[Le virome humain \(Scientific American\)](#)

"Biologists estimate that 380 trillion viruses are living on and inside your body right now—10 times the number of bacteria."



Il existe environ 160 virus 'pathogènes' connus pour infecter l'homme.
Les modes d'infection (Transmission) sont variés de même que les pathologies associées (Disease)

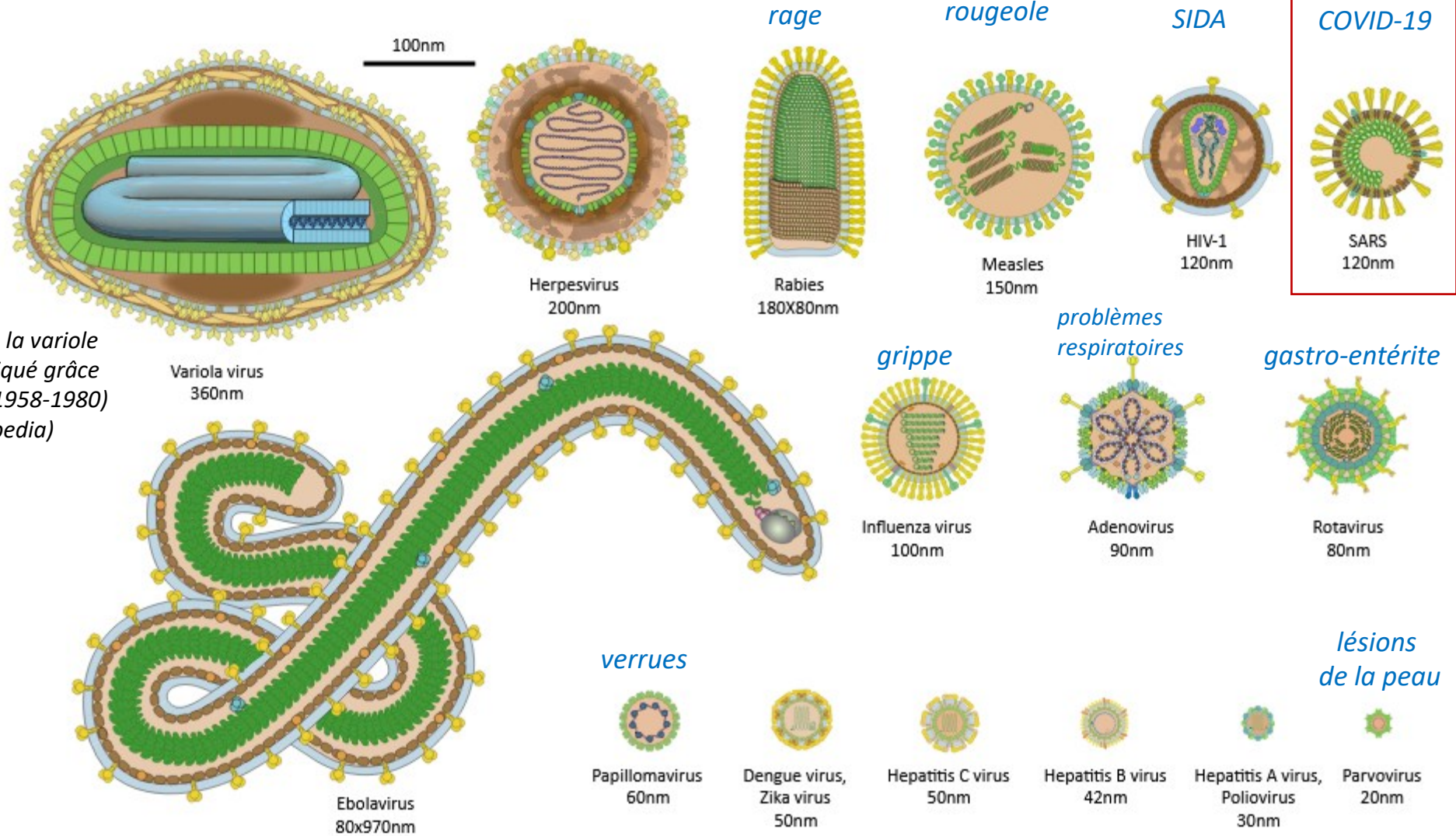
<https://viralzone.expasy.org/678>

 ViralZone search				
Human viruses and associated pathologies The table below displays the list of human viral pathogens, with transmission and general facts about associated pathologies. (See human viruses by Baltimore classification)				
Virus	Genus, Family	Host	Transmission	Disease
Adeno-associated virus	Dependovirus , Parvoviridae	Human, vertebrates	Respiratory	None
Aichi virus	Kobuvirus , Picornaviridae	Human	Fecal-oral	Gastroenteritis
Australian bat lyssavirus	Lyssavirus , Rhabdoviridae	Human, bats	Zoonosis, animal bite	Fatal encephalitis
BK polyomavirus	Polyomavirus , Polyomaviridae	Human	Respiratory fluids or urine	None
Banna virus	Seadornavirus , Reoviridae	Human, cattle, pig, mosquitoes	Zoonosis, arthropod bite	Encephalitis
Barmah forest virus	Alphavirus , Togaviridae	Human, marsupials, mosquitoes	Zoonosis, arthropod bite	Fever, joint pain
Bunyamwera virus	Orthobunyavirus , Bunyaviridae	Human, mosquitoes	Zoonosis, arthropod bite	Encephalitis
Bunyavirus La Crosse	Orthobunyavirus , Bunyaviridae	Human, deer, mosquitoes, tamias	Zoonosis, arthropod bite	Encephalitis
Bunyavirus snowshoe hare	Orthobunyavirus , Bunyaviridae	Human, rodents, mosquitoes	Zoonosis, arthropod bite	Encephalitis
Cercopithecine herpesvirus	Lymphocryptovirus , Herpesviridae	Human, monkeys	Zoonosis, animal bite	Encephalitis
Chandipura virus	Vesiculovirus , Rhabdoviridae	Human, sandflies	Zoonosis, arthropod bite	Encephalitis
Chikungunya virus	Alphavirus , Togaviridae	Human, monkeys, mosquitoes	Zoonosis, arthropod bite	Fever, joint pain

Des exemples de virus 'pathogènes' qui infectent l'homme, leur taille respective et les pathologies associées



Le virus de la variole
a été éradiqué grâce
au vaccin (1958-1980)
(Wikipedia)



Vidéo 2 (2:33)

Pourquoi certains virus rendent malades ?

D'où viennent les virus qui infectent l'homme ?

Quel est le rôle des chauve-souris ?

☰ YouTube^{CH} Search 🔍



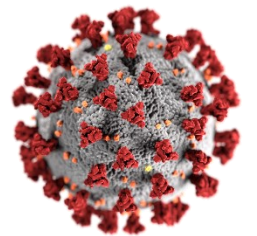
Pourquoi certains virus 'rendent malades' ?

Phillippe Le Mercier
Biologiste, Doctorat en virologie - Institut Pasteur
ViralZone, Groupe Swiss-Prot, Genève
SIB Swiss Institute of Bioinformatics

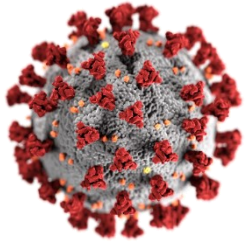


▶ ⏮ 🔊 0:00 / 2:33 CC ⚙️ 📺 🔍

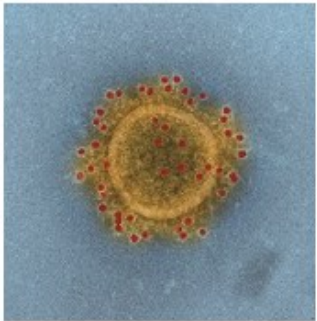
Les coronavirus et SARS-CoV-2



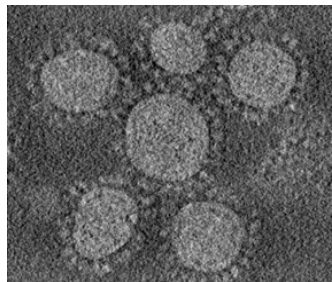
Les coronavirus sont une grande famille de virus (plus de 40 espèces), dont la plupart sont inoffensifs pour l'homme. Quatre types de coronavirus (OC43, 229E, NL63 et HKU1) sont endémiques et connus pour provoquer des rhumes; trois types de coronavirus peuvent causer de graves infections pulmonaires:



- Severe Acute Respiratory Syndrome-related Coronavirus 2
- **SARS-CoV-2**
- responsable de la pandémie de Coronavirus disease-19 (COVID-19)
- réservoir: chauve-souris (?)
- [Génome](#) (référence)



- Middle-East Respiratory Syndrome Coronavirus
- **MERS-CoV**
- responsable d'épidémies récurrentes depuis 2012
- réservoir: chameau, chauve-souris (?)
- [Génome](#) (référence)

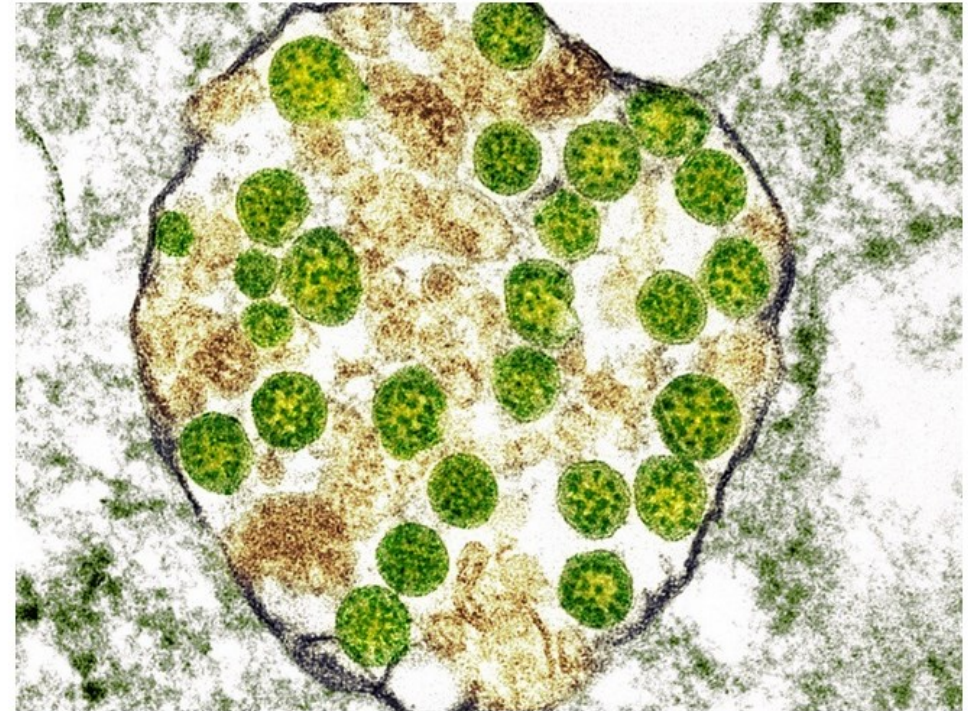
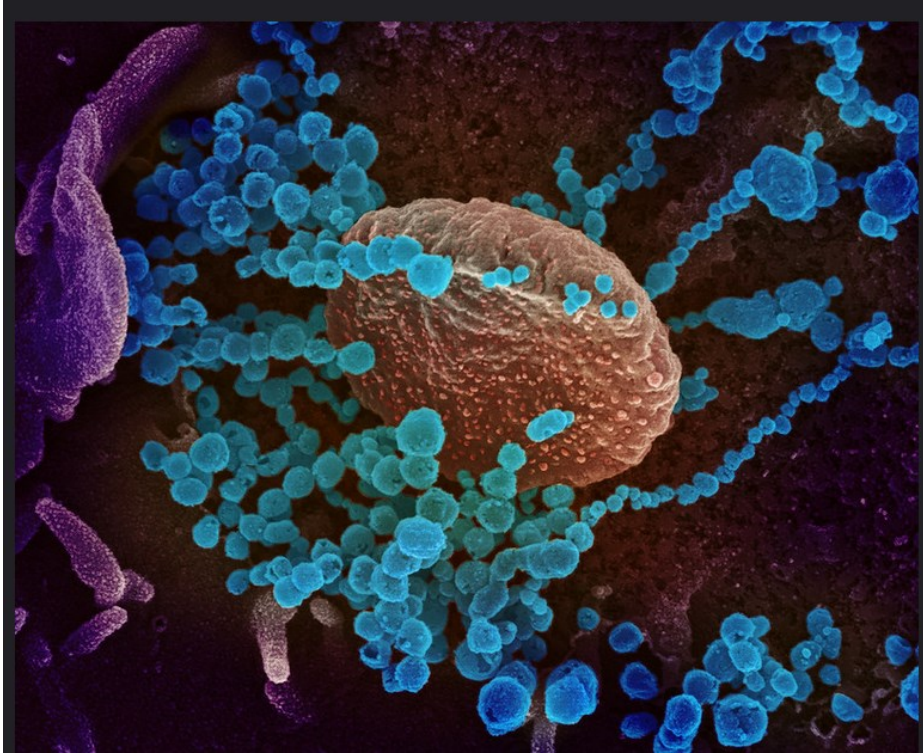


- Severe Acute Respiratory Syndrome-related Coronavirus
- **SARS-CoV**
- responsable d'une épidémie en 2003 qui a touché plus de 30 pays
- réservoir: chauve-souris; animal intermédiaire: civette
- [Génome](#) (référence)

Les estimations concernant l'apparition des premiers coronavirus varient considérablement, de 10 000 ans à 300 millions d'années.

[Plus d'info sur SARS-CoV-2](#)

Des images de SARS-CoV-2 à l'extérieur et à l'intérieur d'une cellule....



A cell infected with particles of SARS-CoV-2. Credit: Cynthia S. Goldsmith and Azaibi Tamin/CDC/SPL

<https://www.flickr.com/photos/niaid/49557550751/>

<https://www.nature.com/articles/d41586-020-00502-w>

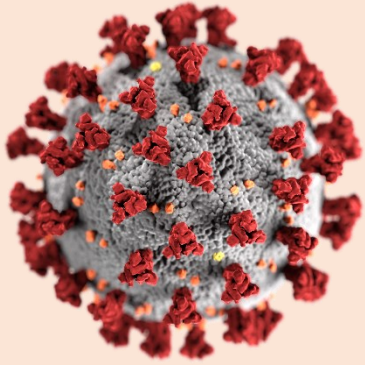
La toux d'un voisin qui envoie 10 particules virales pourrait suffire à déclencher une simple infection de la gorge. Mais si le voisin est plus proche et qu'il envoie 100 particules virales, le virus pourrait descendre jusqu'aux poumons et être plus difficile à éliminer...

Le virus peut se 'reproduire' dans la gorge (et être ainsi présent dans la salive) avant même que les symptômes ne se manifestent...

[Nature 581, 22-26 \(2020\)](#)

Note: le virus pourrait se 'reproduire' 100x toutes les 48 heures. Il 'échappe' en partie à notre système immunitaire, c'est pourquoi on est contagieux avant les symptômes ([source](#))

Note: le virus est détruit par les solutions hydroalcooliques 😊, ce n'est pas le cas de tous les virus !



2 - Le génome de SARS-CoV-2

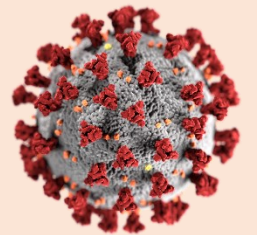
Bioinformatique

Banque de données de séquences ADN & ARN: [GenBank](#)

Outil de recherche de séquences dans les génomes: [BLAT@UCSC](#)

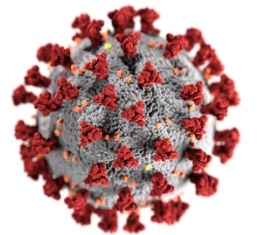
Biologie

[C'est quoi un génome?](#)



La séquence du génome de SARS-CoV-2

ATTAAAGGTTTATACCTTCCCAGGTAAACAAACCAACCAACTTTCGATCTCTTGTAGATCTGTTCTCTAA
CGAACTTTAAAACTGTGTGGCTGTCACTCGGCTGCAIGCTTAGTGCACCTCACGCAGTATAATTAATAA
TAATTACTGTCTGTTGACAGGACACGAGTAACTCGTCTATCTTCTGCAGGCTGCTTACGGTTTCGTCCGT
TTGCAGCCGATCATCAGCACATCTAGGTTTCGTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGT
CCTGGTTTCAACGAGAAAACACACGTCCAACCTCAGTTTGCCTGTTTTACAGGTTTCGCGACGTGCTCGTA
GTGGCTTTGGAGACTCCGTGGAGGAGGTCTTATCAGAGGCACGTCAACATCTTAAAGATGGCACTTGTG
CTTAGTAGAAGTTGAAAAAGGCGTTTTGCCTCAACTTGAACAGCCCTATGTGTTTCATCAAACGTTTCGGA



Le 10 janvier 2020, une équipe chinoise a soumis une première séquence du génome de SARS-CoV-2 à la banque de données GenBank.



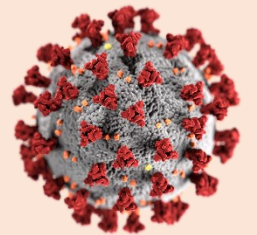
Zhang Yongzhen a choisi de partager la séquence du génome de SARS-CoV2: cette séquence est devenue 'La séquence de référence'.

La séquence du génome de SARS-CoV-2 est composée de 29'903 nucléotides*.

** le génome de SARS-CoV-2 est un ARN simple brin. Dans les banques de données comme GenBank, on représente ce type de génome sous forme de séquence ADN (T à la place de U).*

3- La spécificité du test de détection PCR

Spécificité des amorces (primers)



Séquencer le génome de SARS-CoV-2 a permis de rapidement mettre en place un **test PCR*** pour détecter la présence du virus dans des **frottis** naso-pharyngés (nez) ou oro-pharyngés (gorge).

Le test peut être 'positif' à partir d'une centaine de virus présents dans le frottis!

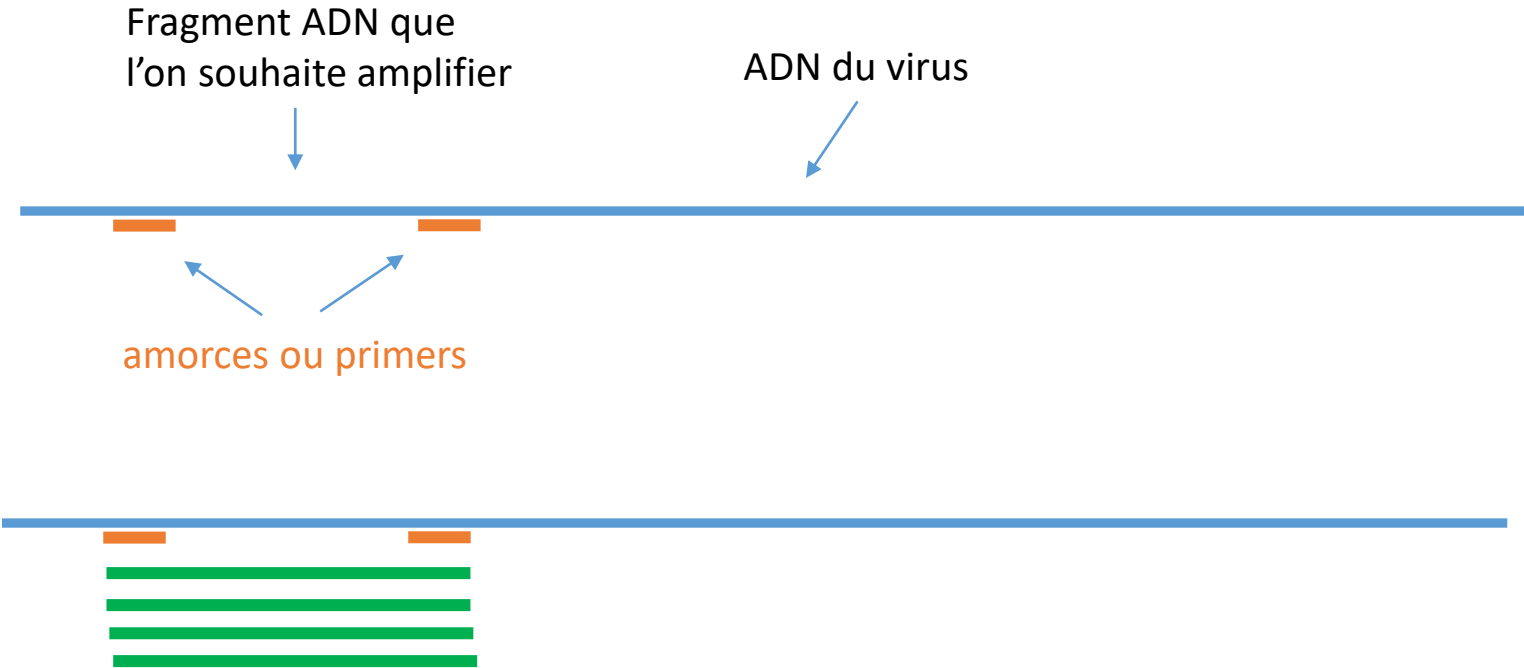


Photo d'illustration. • © JOE RAEDLE / GETTY IMAGES NORTH AMERICA / Getty Images via AFP

* Il s'agit formellement d'une RT-PCR (Reverse Transcription PCR: l'ARNm est d'abord transformé en DNA:

La PCR est une technique de laboratoire qui permet d'amplifier de façon sélective un fragment d'ADN, c'est-à-dire d'en produire des milliers de copies.

Deux petites séquences d'ADN d'environ 15 nucléotides appelées **amorces ou primers**, sont choisies avec soin.
Elles sont complémentaires (100% identique) au **brin d'ADN** et 'encadrent' le fragment à amplifier



Amplification du **fragment** grâce à plusieurs cycles de PCR

Si le virus est présent dans l'échantillon testé, les fragments d'ADN amplifiés seront visibles sur un gel par exemple.

Virus +	Virus -
	

Afin de valider le test, et selon les tests, plusieurs fragments (différentes régions dans le génome) sont amplifiés.

[Un peu de math...](#)
[PCR statistiques: VPN, VPP](#)

Primers (amorces) utilisés pour le test PCR

Voici des exemples de primers utilisés pour les tests PCR:

- Primer no 1 CTCGAACTGCACCTCATGG
- Primer no 2 GGCATACACTCGCTATGTC

! Petit casse-tête (géré par les [programmes bioinformatiques](#) qui permettent de sélectionner les bons primers ...)

Biologie (brins complémentaires)

```
CTCGAACTGCACCTCATGG (primer)
|||||
GAGCTTGACGTGGAGTACC (ADN/ARN)
```

Bioinformatique (alignement)

```
CTCGAACTGCACCTCATGG (primer)
|||||
CTCGAACTGCACCTCATGG (ADN/ARN)
```

Cette activité est basée sur le **point de vue bioinformatique...**

- 1a.** Combien de fois retrouvez-vous la séquence d'un primer dans le génome du virus (manuellement)?
- Chercher la séquence d'un primer dans le [génom du virus](#) (Utiliser Ctrl F ou commandF sur mac)

Région que l'on souhaite amplifier

ADN du virus

amorces

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

NCBI Reference Sequence: NC_045512.2

[GenBank](#) [Graphics](#)

>NC_045512.2 Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome

```
ATTAAAGGTTTATACCTTCCAGGTAACAAACCAACCAACTTTCGATCTCTTGTAGATCTGTTCTCTAAA
CGAACTTTAAAATCTGTGTGGCTGTCTCGGCTGCATGCTTAGTGCACTCACGCAGTATAATTAATAAC
TAATTACTGTCTGTGACAGGACAGAGTAACCTCGTCTATCTTCTGCAGGCTGCTTACGGTTTCGTCGGTG
TTGCAGCCGATCATCAGCACATCTAGGTTTCGTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTG
CCTGGTTTCAACGAGAAAAACACGTCCTCAACTCAGTTTGCTGTTTACAGGTTTCGCGACGTGCTCGTAC
GTGGCTTTGGAGACTCCGTGGAGGAGGCTTATCAGAGGCACGTCAACATCTTAAAGATGGCACTTGTGG
CTTAGTAGAAGTTGAAAAAGGCGTTTTGCCTCAACTTGAACAGCCCTATGTGTTTCATCAACGTTTCGGAT
CTCGAAGTGCACCTCATGGTTCATGTTATGTTGAGCTGTAGCAGAACTCGAAGGCATTGAGTACGGTC
GTAGTGGTGAGACACTTGGTGTCTTGTCCCTCATGTGGCGAAATACCAAGTGGCTTACCGCAAGGTTCT
TCTTCGTAAGAACCGTAATAAAGGAGCTGGTGGCCATAGTTACGGCGCCGATCTAAAGTCATTTGACTTA
GGCGACGAGCTTGGCACTGATCCTTATGAAGATTTTCAAGAAAACCTGGAACACTAAACATAGCAGTGGTG
TTACCCGTGAACCTCATGCGTGAGCTTAAACGAGGGGCATACACTCGCTATGTCGATAACAACCTTCTGTGG
CCCTGATGGCTACCTCTTGTAGTGCACTTAAAGACCTTCTAGCACGTGCTGGTAAAGCTTTCATGCACCTTG
TCCGAACAACTGGACTTTTATTGACACTAAGAGGGGTGTATACTGCTGCCGTGAACATGAGCATGAAATTG
CTTGGTACACGGAACGTTCTGAAAAGAGCTATGAATGACAGACCTTTTGAATTAATTTGGCAAAGAA
ATTTGACACCTTCAATGGGGAATGTCCAAATTTTGTATTTCCCTTAAATTCATAATCAAGACTATTCAA
CCAAGGGTTGAAAAGAAAAAGCTTGATGGCTTTATGGGTAGAATTCGATCTGCTATCCAGTTGCGTCAC
CAAATGAATGCAACCAATGTGCTTTCAACTCTCATGAAGTGTGATCATTGTGGTGAACCTTCATGGCA
```

➤ Primer no 1 CTCGAACTGCACCTCATGG

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

NCBI Reference Sequence: NC_045512.2

[GenBank](#) [Graphics](#)

>NC_045512.2 Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome

```
ATTAAAGGTTTATACCTTCCAGGTAACAAACCAACCAACTTTCGATCTCTTGTAGATCTGTTCTCTAAA
CGAACTTTAAAATCTGTGTGGCTGTCTCGGCTGCATGCTTAGTGCACTCACGCAGTATAATTAATAAC
TAATTACTGTCTGTGACAGGACAGAGTAACCTCGTCTATCTTCTGCAGGCTGCTTACGGTTTCGTCGGTG
TTGCAGCCGATCATCAGCACATCTAGGTTTCGTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTG
CCTGGTTTCAACGAGAAAAACACGTCCTCAACTCAGTTTGCTGTTTACAGGTTTCGCGACGTGCTCGTAC
GTGGCTTTGGAGACTCCGTGGAGGAGGCTTATCAGAGGCACGTCAACATCTTAAAGATGGCACTTGTGG
CTTAGTAGAAGTTGAAAAAGGCGTTTTGCCTCAACTTGAACAGCCCTATGTGTTTCATCAACGTTTCGGAT
GCTCGAAGTGCACCTCATGGTTCATGTTATGTTGAGCTGTAGCAGAACTCGAAGGCATTGAGTACGGTC
GTAGTGGTGAGACACTTGGTGTCTTGTCCCTCATGTGGCGAAATACCAAGTGGCTTACCGCAAGGTTCT
TCTTCGTAAGAACCGTAATAAAGGAGCTGGTGGCCATAGTTACGGCGCCGATCTAAAGTCATTTGACTTA
GGCGACGAGCTTGGCACTGATCCTTATGAAGATTTTCAAGAAAACCTGGAACACTAAACATAGCAGTGGTG
TTACCCGTGAACCTCATGCGTGAGCTTAAACGAGGGGCATACACTCGCTATGTCGATAACAACCTTCTGTGG
CCCTGATGGCTACCTCTTGTAGTGCACTTAAAGACCTTCTAGCACGTGCTGGTAAAGCTTTCATGCACCTTG
TCCGAACAACTGGACTTTTATTGACACTAAGAGGGGTGTATACTGCTGCCGTGAACATGAGCATGAAATTG
CTTGGTACACGGAACGTTCTGAAAAGAGCTATGAATGACAGACCTTTTGAATTAATTTGGCAAAGAA
ATTTGACACCTTCAATGGGGAATGTCCAAATTTTGTATTTCCCTTAAATTCATAATCAAGACTATTCAA
CCAAGGGTTGAAAAGAAAAAGCTTGATGGCTTTATGGGTAGAATTCGATCTGCTATCCAGTTGCGTCAC
CAAATGAATGCAACCAATGTGCTTTCAACTCTCATGAAGTGTGATCATTGTGGTGAACCTTCATGGCA
```

➤ Primer no 2 GGCATACACTCGCTATGTC

Primers (amorces) utilisés pour le test PCR

Voici des exemples de primers utilisés pour les tests PCR:

- Primer no 1: CTCGAACTGCACCTCATGG
- Primer no 2: GGCATACACTCGCTATGTC

1b. Combien de fois retrouvez-vous la séquence d'un primer dans le génome du virus (avec un outil bioinformatique) ?

- Utiliser [BLAT@UCSC](#)
- Copier/coller la séquence d'un primer
- Choisir 'genome' SARS-CoV-2
- Cliquer 'Submit'

2. Combien de fois retrouvez-vous la séquence d'un primer dans le génome humain? Le génome humain est composé de quelque 3 milliards de nucléotides...

- Utiliser [BLAT@UCSC](#)
- Choisir 'genome' human
- Copier/coller la séquence d'un primer
- Cliquer 'Submit'

1b. Combien de fois retrouvez-vous la séquence d'un primer dans le génome du virus (avec un outil bioinformatique [BLAT@UCSC](#)) ?

BLAT Search Genome

Genome: ☐ Search all
SARS-CoV-2

Assembly: Dec. 2013 (GRCh38/hg38)

Query type: BLAT's guess

Sort output: query,score

Output type: hyperlink

CTCGAACTGCACCTCATGG

☐ All Results (no minimum matches)

Submit I'm feeling lucky Clear

BLAT Search Results

Go back to [NC_045512v2:492-510](#) on the Genome Browser.

Custom track name: blat YourSeq

Custom track description: blat on YourSeq

Build a custom track with these results

ACTIONS	QUERY	SCORE	START	END	QSIZE	IDENTITY	CHROM	STRAND	START	END	SPAN
browser details	YourSeq	19	1	19	19	100.0%	NC_045512v2	+	492	510	19

Le primer est retrouvé une seule fois dans le génome du virus (100 % identique)

2. Combien de fois retrouvez-vous la séquence d'un primer dans le génome humain ([BLAT@UCSC](#)) ? Le génome humain est composé de quelque 3 milliards de nucléotides...

BLAT Search Genome

Genome: ☐ Search all

Human

Assembly:

Dec. 2013 (GRCh38/hg38)

Query type:

BLAT's guess

Sort output:

query,score

Output type:

hyperlink

CTCGAACTGCACCTCATGG

☐ All Results (no minimum matches)

Submit

I'm feeling lucky

Clear



Genomes

Genome Browser

Tools

Mirrors

Downloads

My Data

Projects

Help

About Us

Human (hg38) BLAT Results

Sorry, no matches found (with score at least 20)

Le primer n'est pas retrouvé
dans le génome humain
(3 milliards de lettres !)

Primers (amorces) utilisés pour le test PCR: spécificité

3. Essayer de taper une séquence au hasard (15 lettres): la retrouvez-vous dans le génome du virus SARS-CoV-2 ?

- Chercher la séquence de votre 'primer' dans le [génom du virus](#) (manuellement: utiliser Ctrl F ou commandF sur mac)
- Chercher la séquence de votre 'primer' dans le génome du virus avec [BLAT@UCSC](#)

4. Retrouvez-vous les primers dans le génome d'un autre coronavirus ?

- Primer no 1 CTCGAACTGCACCTCATGG
- Primer no 2 GGCATACACTCGCTATGTC

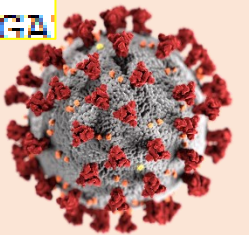
- Chercher la séquence d'un primer dans le [génom de SARS CoV](#) (2003) (Utiliser Ctrl F ou commandF sur mac)
- Illustration: alignement des séquences des 2 coronavirus dans la région amplifiée par les 2 primers:

NC_045512.2	SARS-CoV-2	419	GGCTTAGTAGAAGTTGAAAAAGGCGTTTTGCCTCAACTTGAACAGCCCTATGTGTTTCATC
AY274119.3	SARS-CoV	418	GGTCTAGTAGAGCTGGAAAAAGGCGTACTGCCCCAGCTTGAACAGCCCTATGTGTTTCATT
			** *****. * *****: **** *.*****
NC_045512.2	SARS-CoV-2	479	AAACGTTTCGGATGCTCGAACTGCACCTCATGGTCATGTTATGGTTGAGCTGGTAGCAGAA
AY274119.3	SARS-CoV	478	AAACGTTTCTGATGCCCTTAAGCACCAATCACGGCCACAAGGTCGTTGCTGGCTTGGTGGT
			***** ***** ****** ** :.*****
NC_045512.2	SARS-CoV-2	539	CTCGAAGGCATTTCAGTACGGTCGTAGTGGTGAGACACTTGGTGTCC
AY274119.3	SARS-CoV	538	ATGGACGGCATTTCAGTACGGTCGTAGCGGTATAACACTGGGAGTAC
			. * ..***** ***** ** :.***** ** :.*****

NC_045512.2	SARS-CoV-2	719	GATCCTTATGAAGATTTTCAAGAAAACCTGGAACACTAAACATAGCAGTGGTGTACCCGT
AY274119.3	SARS-CoV	718	GATCCCATTTGAAGATTATGAACAAAACCTGGAACACTAAGCATGGCAGTGGTGCACCTCCGT
			***** :.*****: * * *****.***.***** :. ****
NC_045512.2	SARS-CoV-2	779	GAACTCATGCGTGAGCTTAACGGAGGGGCATACACTCGCTATGTCGATAACAACCTTCTGT
AY274119.3	SARS-CoV	778	GAACTCACTCGTGAGCTCAATGGAGGCGCAGTCACTCGCTATGTCGACAACAATTTCTGT
			***** ***** * * * * * :***** ***** *****
NC_045512.2	SARS-CoV-2	839	GGCCCTGATGGCTACCCTCTTGAGTGCATTAAAGACCTTCTAGCACGTGCTGGTAAAGCT
AY274119.3	SARS-CoV	838	GGCCAGATGGGTACCCTCTTGATTGCATCAAAGATTTCTCGCACGCGCGGGCAAGTCA
			*****:***** ***** ***** ***** *****.***** * * * * * :

4 - Les génomes de SARS-CoV-2

```
ATTAAAGGTTTATACCTTCCCAGGTAACAAACCAACCAACTTTTCGATCTCTTGTAGATCTGTTCTCTAA
CGAACTTTTAAAAICTGTGTGGCTGTCACTCGGCTGCAITGCTTAGTGCACCTCACGCAGTATAATTAAATAA
ATTAAAGGTTTAAATTACTGTCGTTGACAGGACACGAGTAACTCGTCTATCTTCTGCAGGCTGCTTACGGTTTTCGTCCGT
CGAACTTTTATTGTCAGCCGATCATCAGCACATCTAGGTTTTCGTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGT
TAATATTAAACCTTGGTTTCAACGAGAGAAAACACACGTCCAACCTCAGTTTGCCTGTTTTTACAGGTTTCGCGACGTGCTCGTA
TTGCGAACTGTGGCTTTGGAGACTCCGTGGAGGAGGTCTTATCAGAGGCACGTCAACATCTTAAAGATGGCACTTGTG
CCTGTAAATTCTTAGTAGAAGTTGAAAAAGGCGTTTTTGCTCAACTTGAACAGCCCTATGTGTTTCATCAAACGTTTCGGA
GTGGTTGCGAACTTTAAAAICTGTGTGGCTGTCACTCGGCTGCAITGCTTAGTGCACCTCACGCAGTATAATTAAATAA
CTTAATTAAATTACTGTCGTTGACAGGACACGAGTAACTCGTCTATCTTCTGCAGGCTGCTTACGGTTTTCGTCCGT
GTGTTGCAGCCGATCATCAGCACATCTAGGTTTTCGTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGT
CTTCCTTGGTTTCAACGAGAGAAAACACACGTCCAACCTCAGTTTGCCTGTTTTTACAGGTTTCGCGACGTGCTCGTA
GTGGCTTTTGGAGACTCCGTGGAGGAGGTCTTATCAGAGGCACGTCAACATCTTAAAGATGGCACTTGTG
CTTAGTAGAAGTTGAAAAAGGCGTTTTTGCTCAACTTGAACAGCCCTATGTGTTTCATCAAACGTTTCGGA
```



Voici les génomes de virus (provenant de pays différents, à plusieurs mois d'intervalle):

[Séquence de SARS-CoV-2 Chine](#) (Wuhan, 10 janvier 2020) ; séquence de 'référence'

[Séquence de SARS-CoV-2 USA](#) (Minnesota, 21 août 2020)

[Séquence de SARS-CoV-2 Egypt](#) (7 octobre 2020)

[Séquence de SARS-CoV-2 B.1.1.7 UK](#) (8 Janvier 2021)

Comparer ces séquences (copier/coller 2 séquences dans l'outil [align@UniProt](#))

Cela peut prendre plusieurs minutes....

Trouver les différences (= les mutations) et en particulier celle située en position 23'403:

ceci est une mutation



NC_045512.2
MT911538.1

```
23401 GGATGTTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAACTTACTCCTACTTG
23347 GGGTGTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAACTTACTCCTACTTG
**_******
```

23460
23406

NC 045512.2	23281	TGATGCTGTCCGTGATCCACAGACACTTGAGATTCTTGACATTACACCATGTTCTTTTGG	23340
MT911538.1	23227	TGATGCTGTCCGTGATCCACAGACACTTGAGATTCTTGACATTACACCATGTTCTTTTGG	23286

NC 045512.2	23341	TGGTGTCTAGTGTTATAACACCAGGAACAAATACTTCTAACCAGGTTGCTGTTCTTTATCA	23400
MT911538.1	23287	TGGTGTCTAGTGTTATAACACCAGGAACAAATACTTCTAACCAGGTTGCTGTTCTTTATCA	23346

NC 045512.2	23401	GGATGTTAACTGCACAGAAGTCCCTGTTGCTATTTCATGCAGATCAACTTACTCCTACTTG	23460
MT911538.1	23347	GGGTGTTAACTGCACAGAAGTCCCTGTTGCTATTTCATGCAGATCAACTTACTCCTACTTG	23406
		,***	

ceci est une mutation



NC 045512.2
MT911538.1

23401	GGATGTTAACTGCACAGAAGTCCCTGTTGCTATTTCATGCAGATCAACTTACTCCTACTTG	23460
23347	GGGTGTTAACTGCACAGAAGTCCCTGTTGCTATTTCATGCAGATCAACTTACTCCTACTTG	23406
	,***	

NC 045512.2	23761	GACATCAGTAGATTGTACAATGTACATTTGTGGTGATTCAACTGAATGCAGCAATCTTTT	23820
MT911538.1	23707	GACATCAGTAGATTGTACAATGTACATTTGTGGTGATTCAACTGAATGCAGCAATCTTTT	23766

NC 045512.2	23821	GTTGCAATATGGCAGTTTTTGTACACAATTAAACCGTGCTTTAACTGGAATAGCTGTTGA	23880
MT911538.1	23767	GTTGCAATATGGCAGTTTTTGTACACAATTAAACCGTGCTTTAACTGGAATAGCTGTTGA	23826

NC 045512.2	23881	ACAAGACAAAAACACCCAAGAAGTTTTTGCACAAGTCAAACAAATTTACAAAACACCACC	23940
MT911538.1	23827	ACAAGACAAAAACACCCAAGAAGTTTTTGCACAAGTCAAACAAATTTACAAAACACCACC	23886

NC 045512.2	23941	AATTAAAGATTTTGGTGGTTTTAATTTTTTCACAAATATTACCAGATCCATCAAAACCAAG	24000
MT911538.1	23887	AATTAAAGATTTTGGTGGTTTTAATTTTTTCACAAATATTACCAGATCCATCAAAACCAAG	23946

NC 045512.2	24001	CAAGAGGTCATTTATTGAAGATCTACTTTTCAACAAAGTGACACTTGCAGATGCTGGCTT	24060
MT911538.1	23947	CAAGAGGTCATTTATTGAAGATCTACTTTTCAACAAAGTGACACTTGCAGATGCTGGCTT	24006

NC 045512.2	24061	CATCAAAACAATATGGTGATTGCCTTGGTGATATTGCTGCTAGAGACCTCATTTGTGCACA	24120
MT911538.1	24007	CATCAAAACAATATGGTGATTGCCTTGGTGATATTGCTGCTAGAGACCTCATTTGTGCACA	24066

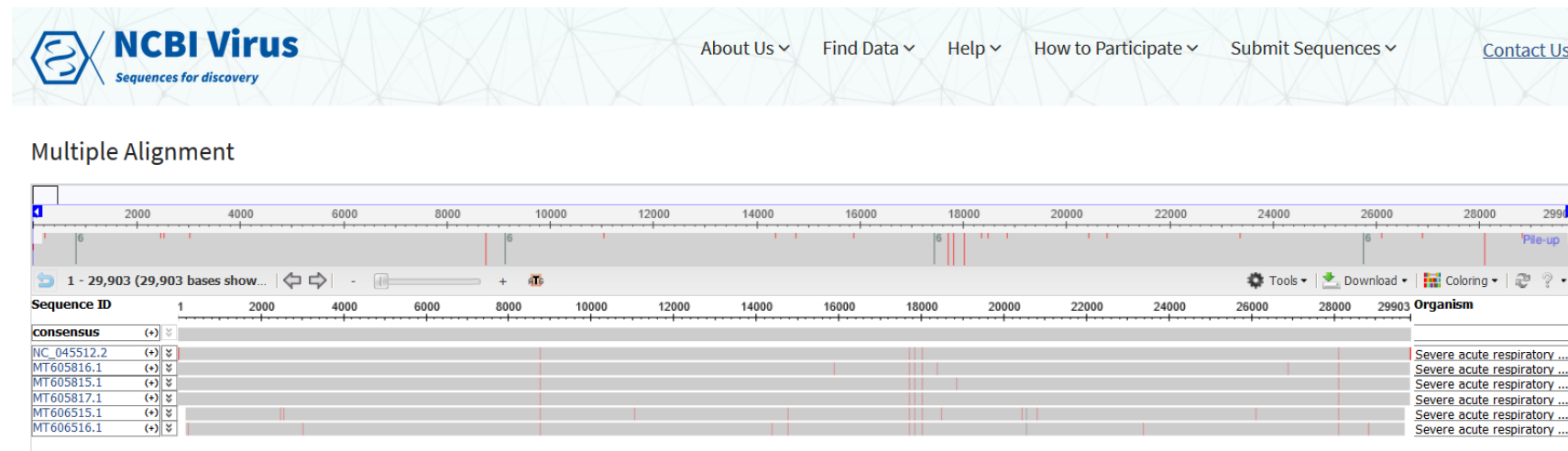
NC 045512.2	24121	AAAGTTTAAACGGCCTTACTGTTTTGCCACCTTTGCTCACAGATGAAATGATTGCTCAATA	24180
MT911538.1	24067	AAAGTTTAAACGGCCTTACTGTTTTGCCACCTTTGCTCACAGATGAAATGATTGCTCAATA	24126

N = le nucléotide n'a
pas pu être identifié

Plusieurs centaines de milliers de génomes de SARS-CoV-2 ont été séquencés dans différents pays. Plusieurs centres de recherche donnent librement accès à ces données. Et c'est essentiel !

Exemple (NCBI): [Severe acute respiratory syndrome coronavirus 2 data hub](#)

- Cliquer sur ce [lien](#) (attention, le site web est régulièrement mis à jour...)
- Ce site permet de comparer des génomes: les différences par rapport à la séquence du génome de référence (consensus) apparaissent sous forme de barres verticales
- Exemple: le génome du virus 'chinois' et plusieurs génomes de virus 'américains'



Expert - Pour avoir accès aux séquences : copier/coller le **numéro d'accèsion/Sequence ID** dans l'URL:
<https://www.ncbi.nlm.nih.gov/nuccore/MT911538.1?report=fasta>

Le variant anglaise (B.1.1.7; 501Y V1)

La 'lignée britannique' du SARS-CoV-2 (B.1.1.7) a été découverte le 8 décembre 2020:

- 17 'nouvelles' mutations sont présentes dans le génome du virus B.1.1.7

Comparaison des génomes de SARS-CoV-2:

- [génom de référence \(NC_045512.2\)](#)
- [génom du variant anglais B.1.1.7 \(LR991698\)](#)

[Align@UniProt](#) (cliquer sur 'similarity')

Les différences entre le génome de SARS-CoV-2 (génom de référence) et le génome du virus 'anglais B.1.1.7' apparaissent en clair dans cet alignement.

Ces différences correspondent aux 'nouvelles' mutations présentes dans le virus anglais.

ceci est une mutation



LR991698.2	23023	ATCATATGGTTTCCAACCCACTTATGGTGTGGTTACCAACCATACAGAGTAGTAGTACT	23082
NC_045512.2	23041	ATCATATGGTTTCCAACCCACTAATGGTGTGGTTACCAACCATACAGAGTAGTAGTACT	23100
		*****:*****	
LR991698.2	23083	TTCTTTTGAACCTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTT	23142
NC_045512.2	23101	TTCTTTTGAACCTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTT	23160

LR991698.2	23143	GGTTAAAAACAAATGTGTCAATTTCAACTTCAATGGTTTAAACAGGCACAGTTCTTAC	23202
NC_045512.2	23161	GGTTAAAAACAAATGTGTCAATTTCAACTTCAATGGTTTAAACAGGCACAGTTCTTAC	23220

LR991698.2	23203	TGAGTCTAACAAAAAGTTTCTGCCTTTCCAACAATTTGGCAGAGACATTGATGACACTAC	23262
NC_045512.2	23221	TGAGTCTAACAAAAAGTTTCTGCCTTTCCAACAATTTGGCAGAGACATTGCTGACACTAC	23280

LR991698.2	23263	TGATGCTGTCCGTGATCCACAGACACTTGAGATTCTTGACATTACACCATGTTCTTTTGG	23322
NC_045512.2	23281	TGATGCTGTCCGTGATCCACAGACACTTGAGATTCTTGACATTACACCATGTTCTTTTGG	23340

LR991698.2	23323	TGTCAGTGTATAACACCAGGAACAAATACTTCTAACAGGTTGCTGTTCTTTATCA	23382
NC_045512.2	23341	TGTCAGTGTATAACACCAGGAACAAATACTTCTAACAGGTTGCTGTTCTTTATCA	23400

LR991698.2	23383	GGTGTTAAGTGCACAGAAGTCCCTGTGCTATTTCATGCAGATCAACTTACTCCTACTTG	23442
NC_045512.2	23401	GGATGTTAAGTGCACAGAAGTCCCTGTGCTATTTCATGCAGATCAACTTACTCCTACTTG	23460
		** *****	
LR991698.2	23443	GCGTGTATTATCTACAGGTTCTAATGTTTTTCAAACCGTGCAGGCTGTTTAAAGGGGC	23502
NC_045512.2	23461	GCGTGTATTATCTACAGGTTCTAATGTTTTTCAAACCGTGCAGGCTGTTTAAAGGGGC	23520

LR991698.2	23503	TGAACATGTCAACAACCTCATATGTGTGACATACCCATTGGTGCAGGTATATGCGCTAG	23562
NC_045512.2	23521	TGAACATGTCAACAACCTCATATGTGTGACATACCCATTGGTGCAGGTATATGCGCTAG	23580

LR991698.2	23563	TTATCAGACTCAGACTAATTCTCATCGGCGGGCACGTAAGTGTAGCTAGTCAATCCATCAT	23622
NC_045512.2	23581	TTATCAGACTCAGACTAATTCTCATCGGCGGGCACGTAAGTGTAGCTAGTCAATCCATCAT	23640

LR991698.2	23623	TGCCTACACTATGTCACTTGGTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGC	23682
NC_045512.2	23641	TGCCTACACTATGTCACTTGGTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGC	23700

(! seule une toute petite partie de l'alignement est visible ici)

Résultat alignement génome de référence / génome B.1.1.7 (UK): [fichier txt](#)

« Le coronavirus mute relativement peu. En tout cas moins qu'une grippe, une gastro-entérite ou une hépatite. Mais il mute suffisamment pour qu'on arrive à le reconnaître et à identifier ses 'ancêtres'", note Didier Trono.

La génomique à la rescousse pour étudier l'évolution du coronavirus

<https://www.rts.ch/info/sciences-tech/11559187-la-genomique-a-la-rescousse-pour-etudier-levolution-du-coronavirus.html>

Pourquoi les coronavirus mutent-ils peu ?

Les coronavirus sont parmi les rares virus à ARN dotés d'un mécanisme de 'correction génomique' - qui empêche le virus d'accumuler des mutations.

Note: Les médicaments qui 'tuent' les virus (comme l'hépatite C) en induisant des mutations, ne marchent donc pas contre SARS-CoV-2!

[Nature **581**, 22-26 \(2020\)](#)

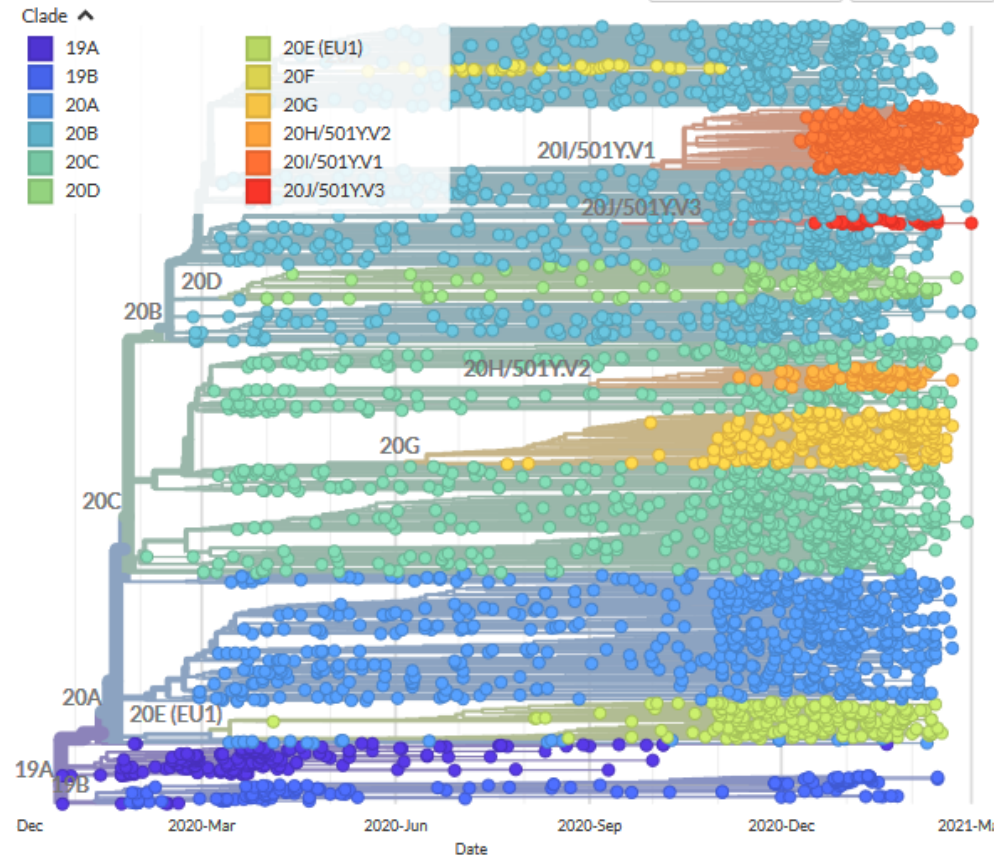
Genomic epidemiology of novel coronavirus - Global subsampling

Maintained by the [Nextstrain team](#). Enabled by data from [GISAID](#)

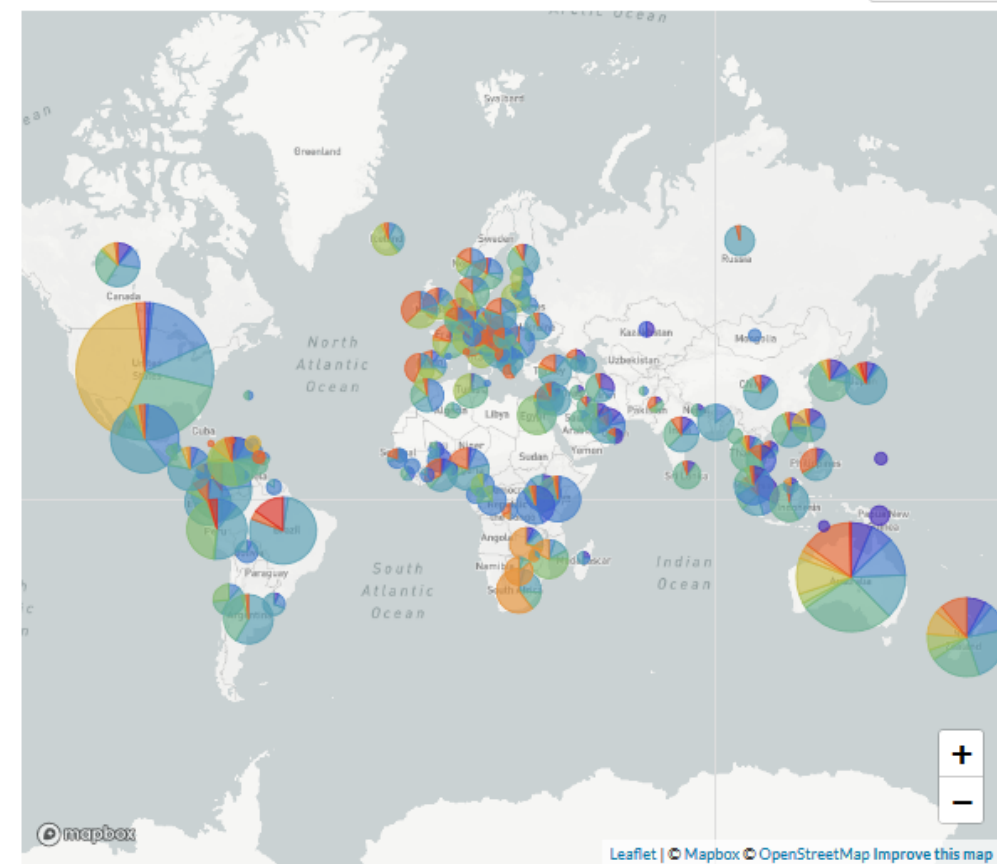
Showing 3960 of 3960 genomes sampled between Dec 2019 and Mar 2021.

Mars 2021

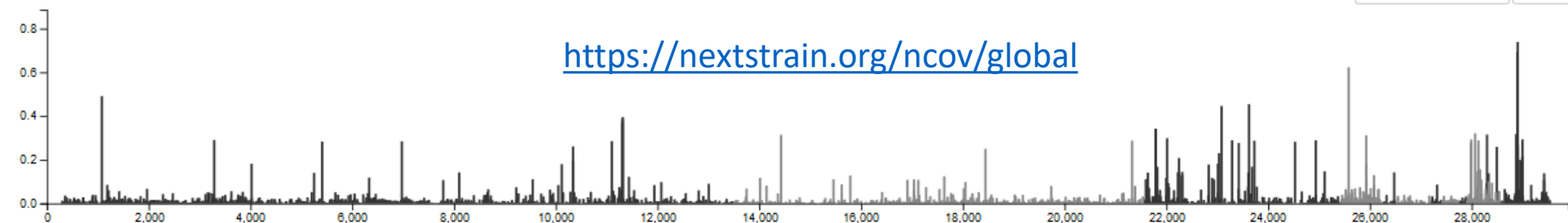
Phylogeny

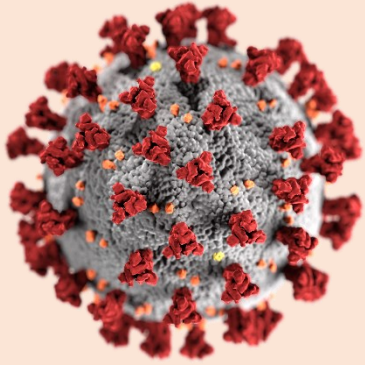


Geography



Diversity





5 - Les protéines de SARS-CoV-2

Bioinformatique

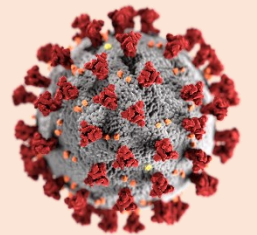
Banque de données de protéines: [UniProtKB/Swiss-Prot](#)

Outil de traduction ADN -> protéine: [translate @ ExPASy](#)

Banque de données de structures 3D: [PDB](#)

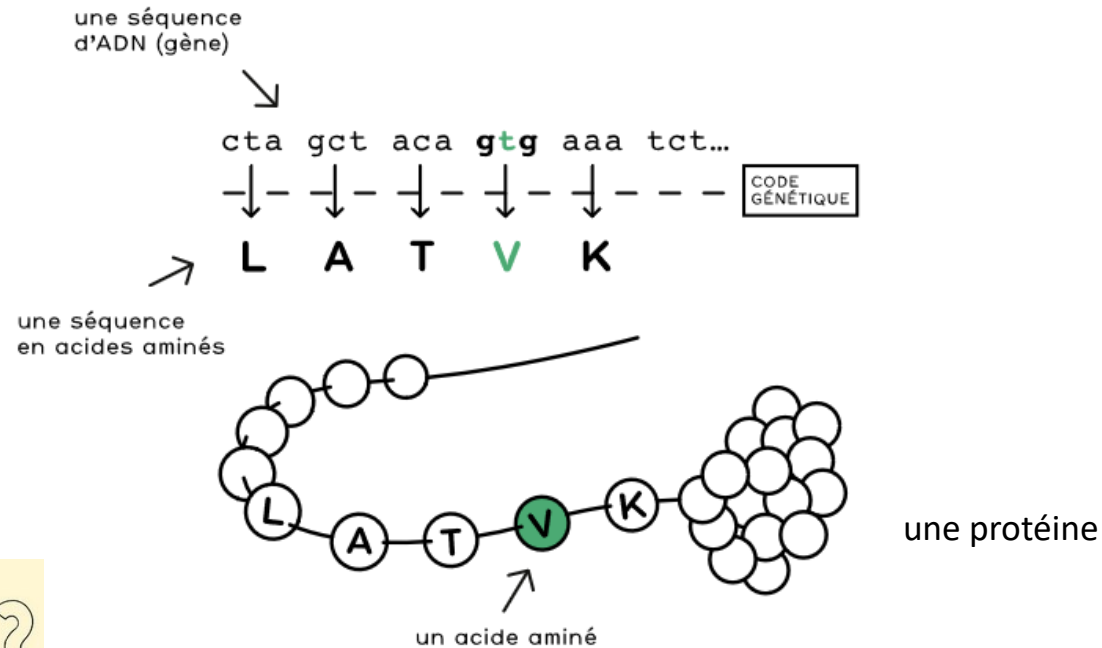
Biologie

[C'est quoi une protéine?](#)



L'analyse de la séquence du génome du virus a permis de trouver les séquences en acides aminés des protéines du virus. Environ 17 gènes codant pour des protéines ont été identifiés.

DES NUCLÉOTIDES AUX ACIDES AMINÉS...



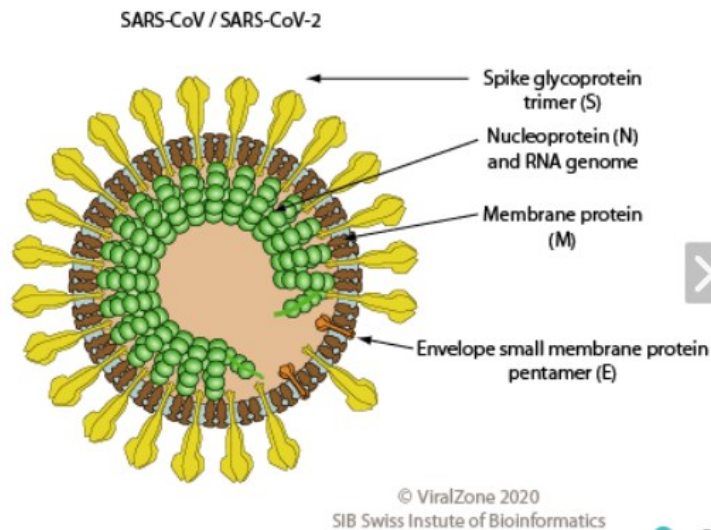
[Une protéine?](#)

[C'est quoi une protéine?](#)
www.precisionmed.ch

☐	Entry	Entry name		Protein names		Gene names	Length
☐	P0DTD1	R1AB_SARS2		Replicase polyprotein 1ab		rep 1a-1b	7,096
☐	P0DTC1	R1A_SARS2		Replicase polyprotein 1a			4,405
☐	P0DTC2	SPIKE_SARS2		Spike glycoprotein		S 2	1,273
☐	P0DTC9	NCAP_SARS2		Nucleoprotein		N	419
☐	P0DTD2	ORF9B_SARS2		ORF9b protein		9b	97
☐	P0DTC3	AP3A_SARS2		ORF3a protein		3a	275
☐	P0DTC7	NS7A_SARS2		ORF7a protein		7a	121
☐	P0DTC6	NS6_SARS2		ORF6 protein		6	61
☐	P0DTC8	NS8_SARS2		ORF8 protein		8	121
☐	P0DTC5	VME1_SARS2		Membrane protein		M	222
☐	P0DTC4	VEMP_SARS2		Envelope small membrane protein		E 4	75
☐	P0DTD8	NS7B_SARS2		ORF7b protein		7b	43
☐	P0DTF1	ORF3B_SARS2		ORF3b protein			22
☐	P0DTD3	ORF9C_SARS2		ORF9c protein		9c	73
☐	P0DTG0	ORF3D_SARS2		ORF3d protein			57
☐	P0DTG1	ORF3C_SARS2		ORF3c protein			41
☐	A0A663DJA2	A0A663DJA2_SARS2		ORF10 protein		ORF10 orf10	38

Les protéines de SARS-CoV-2 dans UniProtKB

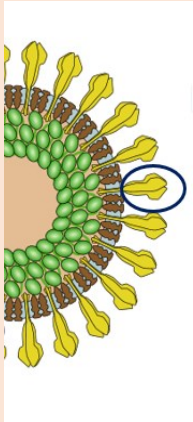
☐	Entry ▾	Entry name ▾		Protein names ▾	Gene names ▾	Length ▾
☐	P0DTD1	R1AB_SARS2		Replicase polyprotein 1ab	rep 1a-1b	7,096
☐	P0DTC1	R1A_SARS2		Replicase polyprotein 1a		4,405
☐	P0DTC2	SPIKE_SARS2		Spike glycoprotein	S 2	1,273
☐	P0DTC9	NCAP_SARS2		Nucleoprotein	N	419
				ORF9b protein	9b	97
				ORF3a protein	3a	275
				ORF7a protein	7a	121
				ORF6 protein	6	61
				ORF8 protein	8	121
				Membrane protein	M	222
				Envelope small membrane protein	E 4	75
				ORF7b protein	7b	43
☐	P0DTF1	ORF3B_SARS2		ORF3b protein		22
☐	P0DTD3	ORF9C_SARS2		ORF9c protein	9c	73
☐	P0DTG0	ORF3D_SARS2		ORF3d protein		57
☐	P0DTG1	ORF3C_SARS2		ORF3c protein		41
☐	A0A663DJA2	A0A663DJA2_SARS2		ORF10 protein	ORF10 orf10	38



Les protéines de SARS-CoV-2 dans UniProtKB

☐	Entry ▾	Entry name ▾		Protein names ▾	Gene names ▾	Length ▾
☐	P0DTD1	R1AB_SARS2		Replicase polyprotein 1ab	Proofreading exoribonuclease	096
☐	P0DTC1	R1A_SARS2		Replicase polyprotein 1a		4,405
☐	P0DTC2	SPIKE_SARS2		Spike glycoprotein	S 2	1,273
☐	P0DTC9	NCAP_SARS2		Nucleoprotein	N	419
☐	P0DTD2	ORF9B_SARS2		ORF9b protein	9b	97
☐	P0DTC3	AP3A_SARS2		ORF3a protein	3a	275
☐	P0DTC7	NS7A_SARS2		ORF7a protein	7a	121
☐	P0DTC6	NS6_SARS2		ORF6 protein	6	61
☐	P0DTC8	NS8_SARS2		ORF8 protein	8	121
☐	P0DTC5	VME1_SARS2		Membrane protein	M	222
☐	P0DTC4	VEMP_SARS2		Envelope small membrane protein	E 4	75
☐	P0DTD8	NS7B_SARS2		ORF7b protein	7b	43
☐	P0DTF1	ORF3B_SARS2		ORF3b protein		22
☐	P0DTD3	ORF9C_SARS2		ORF9c protein	9c	73
☐	P0DTG0	ORF3D_SARS2		ORF3d protein		57
☐	P0DTG1	ORF3C_SARS2		ORF3c protein		41
☐	A0A663DJA2	A0A663DJA2_SARS2		ORF10 protein	ORF10 orf10	38

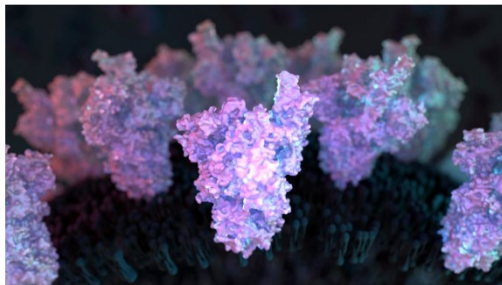
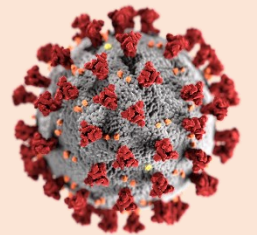
Les protéines de SARS-CoV-2 dans UniProtKB



6 - La protéine Spike de SARS-CoV-2

*Spike ou spicule en français

«Une seule protéine située à la surface du nouveau coronavirus est responsable de sa transmission chez l'humain et de la pandémie qui en a découlé: la protéine Spike. Poussés par les développements technologiques majeurs de ces dernières années, les scientifiques ont rapidement déterminé sa composition et sa structure 3D, ce qui a grandement aidé le développement des vaccins.» [Radio Canada](#)

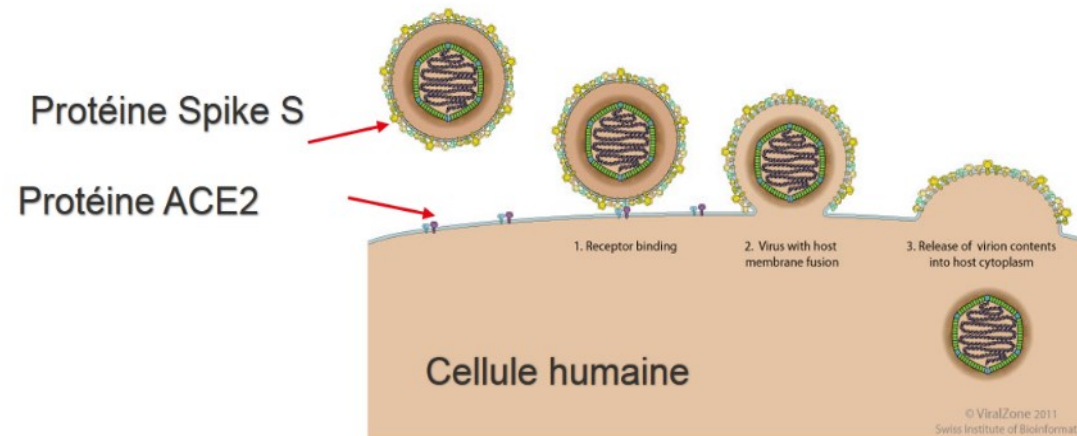
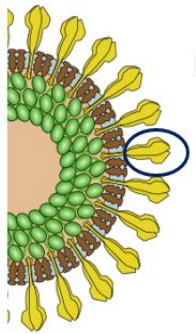


An illustration of the spike proteins that the SARS-CoV-2 virus uses to break into human cell membranes. Credit: Design Cells/Science Photo Library

[source](#)

Spike et ACE2

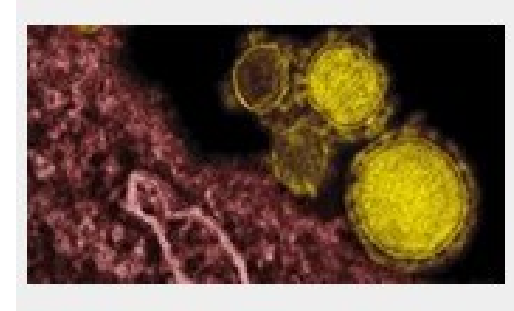
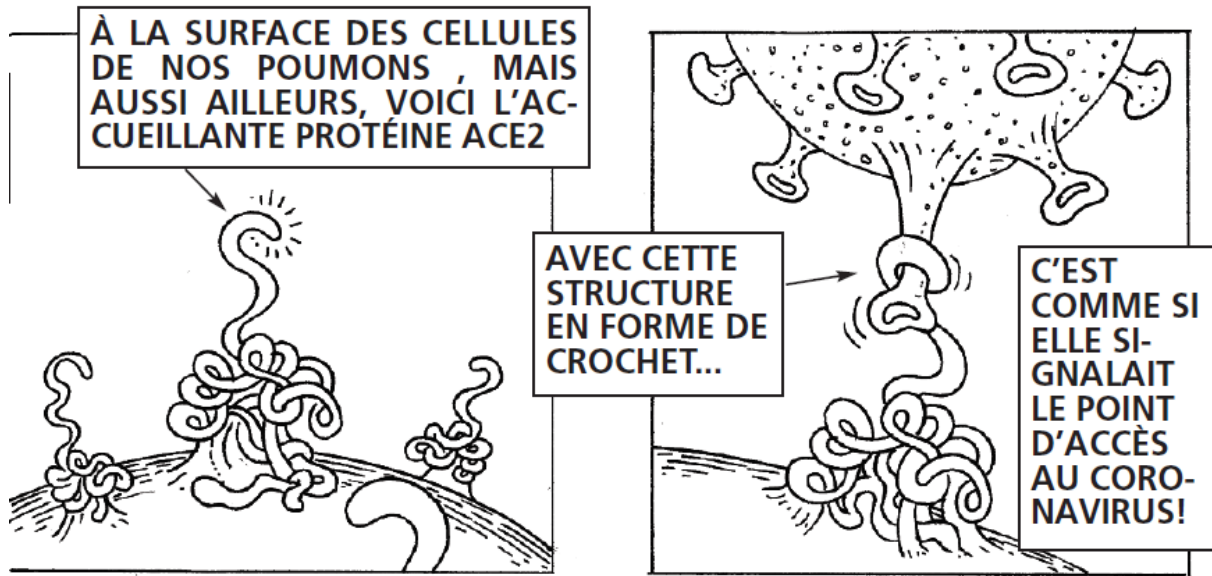
La protéine **Spike** permet au virus de pénétrer dans nos cellules, en interagissant entre autre avec une protéine humaine appelée **ACE2**, présente à la surface de certaines cellules humaines (nez, poumons, intestins, reins, vaisseaux sanguins, ...).



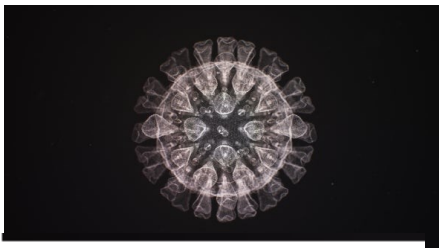
<https://viralzone.expasy.org/9096>

Note: d'autres protéines humaines sont impliquées dans ce processus: elles coupent Spike et induisent ainsi la production d'un 'peptide de fusion' (TMPRSS2, Furin)

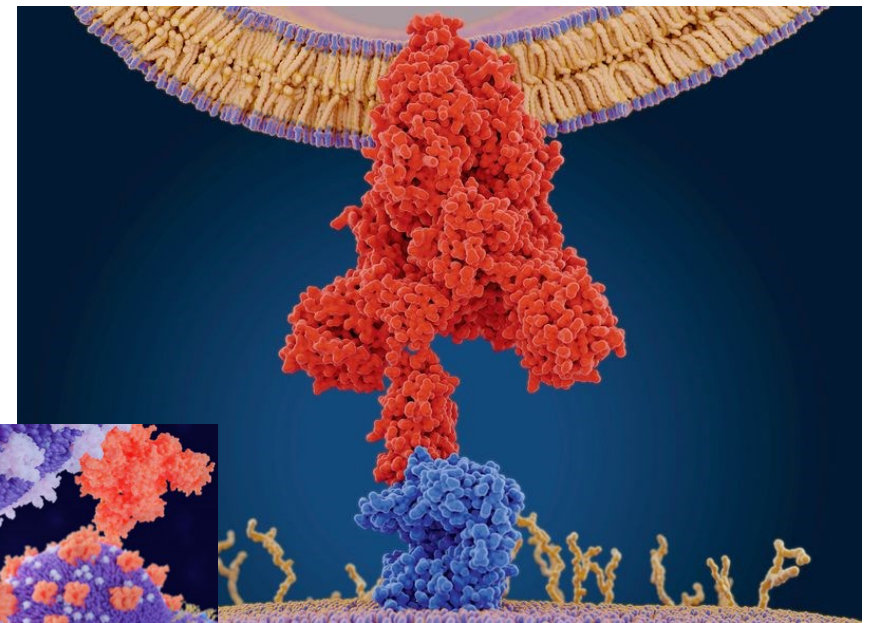
angiotensin-converting enzyme 2



[Protein Spotlight comics ACE2](#): Entrez donc !



Credit: Jordan Siemens *Getty Images*



[source](#)

Vidéo 3 (4:54)

Comment SARS-CoV-2 entre dans les cellules humaines ?

Pourquoi des symptômes aussi différents, comme par exemple la perte d'odorat ?

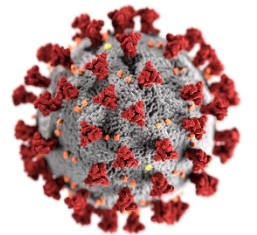
Pourquoi une infection avec SARS-CoV ou SARS-CoV-2 ne conduit pas aux mêmes symptômes ?

Pourquoi SARS-CoV n'a pas été à l'origine d'une pandémie en 2003 ?



Du génome aux protéines

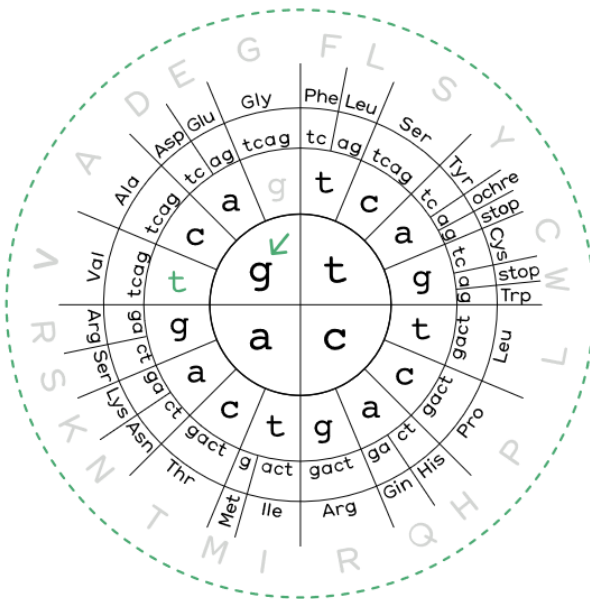
[Bad News Wrapped in Protein: Inside the Coronavirus Genome](#) (New-York Times)



Du génome aux protéines

Voici le début de la séquence du gène codant pour la protéine Spike:

```
atgtttgt ttttcttggt ttattgccac tagtctctag tcagtgtggt
aatcttactaa ccagaactca attaccccct gcatacacta attctttcac
acgtggtggt tattaccctg acaaagtgtt cagatcctca gttttacatt
caactcagga cttgttctta cctttctttt ccaatgttac ttggttccat
gctatacatg tctctgggac caatgggtact aagagggttg ataaccctgt
cctaccattt aatgatgggt tttattttgc ttccactgag aagtctaaca
taataagagg ctggatgttt ggtactactt tagattcgaa gaccagtc
ctacttattg ttaataacgc tactaatggt ...
```



Quels sont les premiers acides aminés de la protéine Spike?

atg ttt gtt ttt ctt gtt tta ttg ...
M F **V**

```
>Gene_Spike
atgttttgttttttttttttttttattgccaactagctctctagtcagtggtttaaattttacaacc
agaactcaatttaccocctgcatacactaatttttccacacgtggtgtttattaccctgcac
aaagttttcagatccctcagttttacattcaactcaggacttggtttacctttctttttcc
aatgttactttggttccatgctatcacatgtctctgggacaaatggtactaagaggtttgat
aacctgttccctaccatttaattgagtggtgttttttttgcctccactgagaaggttaacata
ataagaggtgtgatttttggtaactcttagatctcgaagaccagtcctcacttattgttt
aataacgctactaatgtgttataaagtctgtgaatttcaattttgtatgatccattt
ttgggtgttttttaccacaaaaacacaaagtgtggtgaaagtgcaggttcagagtttat
tctagtgcgaataattgcacattttgaatattgtctctcagcctttttctttaggacctgaa
ggaaaaacagggttaatttcaaaaatcttagggaatttggttttaagaattattgatggttat
tttaaaatattattcgaacacgccttattatttagtgcgtgatctccctcagggtttt
tcggcttttagaaccattggttagatttgcaataggtattacatcactaggtttcaaat
ttactttgctttacatagaagttatttgactcctggtgattctctctcaggttgacagct
ggtgctgcagcttattattggtggttatcttcaacctaggactttttctataaaatataat
gaaaatggaaacattacagatgctgtagactgtgacccctctcagaaacaaag
tgtacgttgaattcctcactgttagaaaaaggaatctatacaacttctcaacttagagtc
caaccaacagaactctattgttagatttcttaattacaaacttggtcccttttggtgaa
gtttttacagccacagatttgctctgttttatgcttggaacaggagaatcagcaac
tgtgttgctgattattctgtcctataaattccgcacatcttttcaacttttaagtgttat
ggagtgctcctactaaaataaagtactgtctgtttactaatgtctatgagatctatt
gtaattagaggtgatgaagtcagacaaatcgctccagggtgaaactggaagattgtgat
tataattataaattaccagatgattttacaggtgcgttatagcttggaatttcaaat
cttgattctaaaggtggtggttaattataattacctgtatagattgttaggaagctaat
ctcaaaccttttgagagagattttcaactgaaatctcactcagcggtagcacacctgt
aatggtgttgaaggttttaattgttactttcctttacaaatcatatggtttccaaacaaat
aatggtgttgggttaccacacatcacagagtagtagtactttctttgaaacttctacatgca
cagcaactgtttgtgacctaaaaagtctactaatttggttaaaaaacaaatgtgtcaat
ttcaacttcaatggttttaacaggcagcaggtgttcttactgagcttaacaaaaagtctgtg
cctttccaaacaaatttgccagagacattgtgacactactgatgtgctcgtgatccacag
caacttgagattcttgacattacacacattgttcttttggtggtgtaggtttataacacca
ggaacaaactcttaaccaggttgctgttctttatcaggatgttaactgcacagaagtc
cotgttgctattcatgcagatcaacttactcctacttggtggtttattctacaggttct
aatgtttttcaaacacgtgcaggtgttttaattaggggtgaacatgcaacaaactcatat
gagtgtagacataccatttgcaggttatatgcgttagttatcagaactcagactaattct
cctcggcgccgacgttagtgtagctcaatccatcttgccacactatgcaactgtgtg
gcagaaaaatcagttgcttactctaaactctatttgccatccccacaaatttactatt
agtgttaccacagaaattctaccaggtctctatgaccaagacatcagtagattgtacaatg
tacatttgggtgatttcaactgaaatgcagcaatcttttggtaaatggcaggtttttgt
acacaaataaacgtgctttaaactggaatagctgttgaacaaagcaaaaaaccccagaa
gtttttgcaacagtcacaaatattacaaaaacacacaaatataagatttttggtggtttt
aatttttcacaaatattaccagatccatcaaaacaaagcaagaggtcatttttgaagat
ctacttttcaacaaagtgcacactgcagatgctggtctcatcaacaaatattggtgattgc
cttggtgataattgctgctagagacctcatttggtgacaaaaagttaaaggccttactgtt
ttgcccactttgtctacagatgaaatgattgtcacaacactctcagctgttagcgggt
acaactcactctgttggtgacotttggtgcaggtgctgcatcaaaatcactttgctatg
caaatggcttataggttttaattggtatttgaggttacacagaatgttctctatgagaacca
aaattgattgcaacaaatataatagtgctattggcaaaatcaagactcactttcttcc
acagcaagtgcacttggaacactcaagatgtggtgcaacaaatgcacaaagctttaaac
acgcttgttaaaccaacttagctccaatttttggtgcaatttcaaggtgttttaaatgatac
ctttcacgtcttgcaaaagttgaggtgaaagtgcacaaatgtaggttgatcacagcaga
cttcaagtttgacagacatgtgactcaacaaatatttagagctgcagaaatcagagct
totgtaattctgtgctactaaaatgtcagagtggtgacttggaacatcaaaaagatt
gatttttggtgaaagggtctatcatcttatgtctcctcctcagtcagcactcaggtgtga
gtctctctgcatgacttatgtcctcgcacagaaagaaacttcacaaactgctcctgccc
attgtcatgatggaagacacactttcctcgtggaaggtgtctttgtttcaaatggcaca
cactggtttgtaacacaaaggaatttttatgaacacacaaatcattactacagacacaca
tttgtgtcgtgtaactgtgattgttaataggaattgtcaacacacagtttatgatcct
ttgcaacctgaattagactcattcaaggaggagtttagataaatttttaagaatcataca
tcaccagatgtgttttagtgacatctctggcatttaattgctcaggttgaacacatcaa
aagaaattgaccgcctcaatgaggttgccaagaatttaaatgaatctctcatgatctc
caagaacttggaagtatgagcagtatataaaatggccatggtacatttggtcaggttt
atagctggtctgattgcatagtaattggtgacaattatgctttgtctgtatgacaggttgc
tgtagttgtcgaagggtgtgttctcttggtggaatcctcgtgcaaaattgatgaagacgac
ctgagccagtgctcaaggagtcacaaattacattacacata
```

(1) Voici la [séquence complète du gène](#) codant pour la protéine Spike de SARS-CoV-2.

Utiliser l'outil [Translate@Expasy](#) pour traduire cette séquence de nucléotides en séquence d'acides aminés

La séquence en acides aminés de la protéine Spike se trouve dans le 'Frame 1'

Cette séquence, sous forme d'ARNm, est présente dans les vaccins mRNA (Pfizer et Moderna)

(2) Expert: 'Voir' le gène codant pour la protéine Spike dans le génome de SARS-CoV-2:

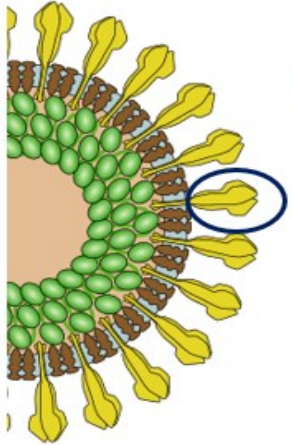
- Depuis: https://www.ncbi.nlm.nih.gov/nucore/NC_045512.2
- Chercher 'Spike' dans la page (Utiliser Ctrl F ou commandF sur mac)
- Cliquer sur 'gene'

gene

```
21563..25384
/gene="S"
/locus_tag="GU280_gp02"
/gene_synonym="spike glycoprotein"
/db_xref="GeneID:43740568"
```

```
21361 aatccaaatc agttgtcttc ctattcttta ttgacatga gtaaatttcc cottaataa
21421 aggggtactg ctgttatgtc tttaaaagaa ggtcaaatca atgatatgat tttatctctt
21481 cttagtaaa gtagacttat aattagagaa aacaacagag ttgttatctc tagtgatgtt
21541 ctgtttaaca actaaacgaa caatgtttgt tttctttgtt ttattgcaac tagtctctag
21601 tcaagtgttt aatcttcaaa ccagaactca attacccctc gcatacacta attctttcac
21661 acgtgtgttt tattacccctg acaaaagtgtt cagatccctca gttttacatt caactcagga
21721 ctgtgtctta cttttctttt ccaatgttac ttgggtccat gctatacatg totctgggac
21781 caatggtaac aagaggtttg ataacccctg cctaccattt aatgatgttg tttattttgc
21841 ttccactgag aagtcaaca taataagagg ctggattttt ggtactactt tagattcgaa
21901 gaccacgtcc ctactatttg ttaataacgc tactaaggtt gttattaaag tctgtgaatt
21961 tcaattttgt aatgatccat ttttggtgtt ttattaccac aaaaacacaa aaagtggat
```

La protéine Spike...et ses variants



Découvrez la [séquence en acides aminés](#) de la protéine Spike de SARS-CoV-2 dans la banque de données UniProtKB/Swiss-Prot.

Quelle est la longueur (nombre d'acides aminés) de la protéine Spike?

Certaines mutations dans la séquence du génome peuvent modifier la séquence en acide aminé de la protéine correspondante.

Exemple:

LR991698.2	23023	ATCATATGGTTTCCAACCCACTTATGGTGTG	GGTTACCAACCATAACAGAGTAGTAGTACT	23082
NC_045512.2	23041	ATCATATGGTTTCCAACCCACTAATGGTGTG	GGTTACCAACCATAACAGAGTAGTAGTACT	23100
		*****:*****		

Génome 'UK' (LR991698.2)

Protéine Spike UK

tat ggt gtt

Y **G** **V**

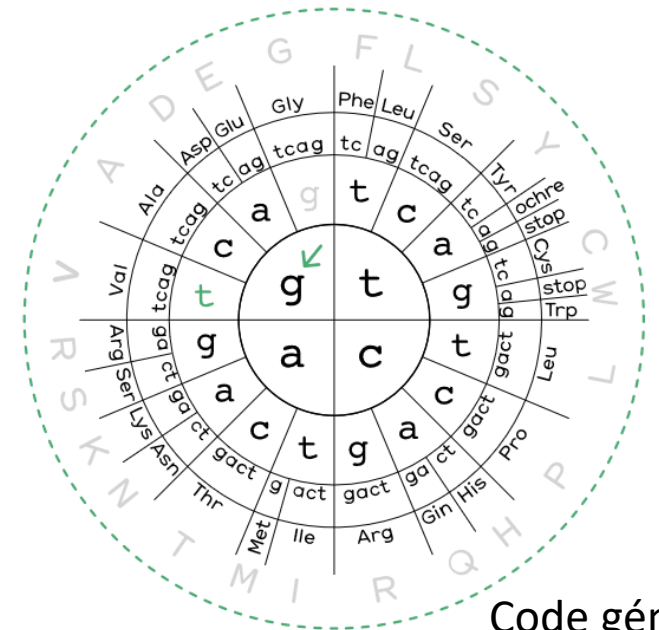
Génome référence (NC_045512.2)

Protéine Spike référence

aat ggt gtt

N **G** **V**

... 501 502 503 ...



La protéine Spike est composée de 1273 acides aminés.

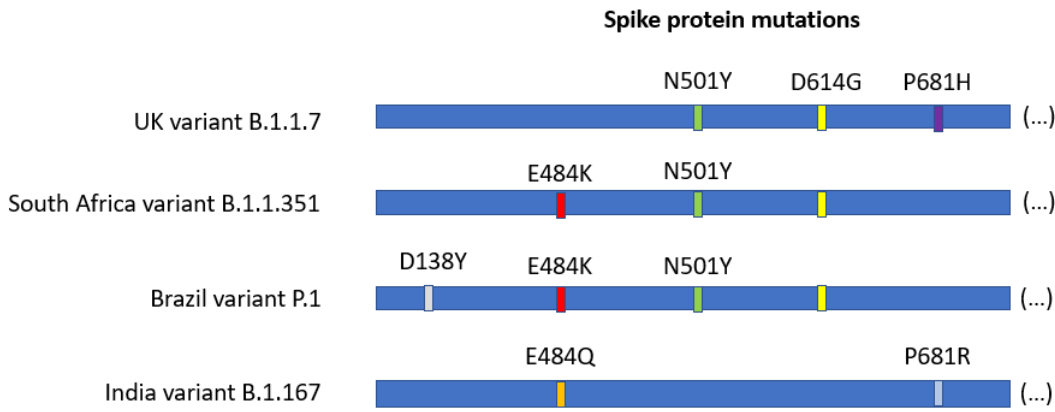
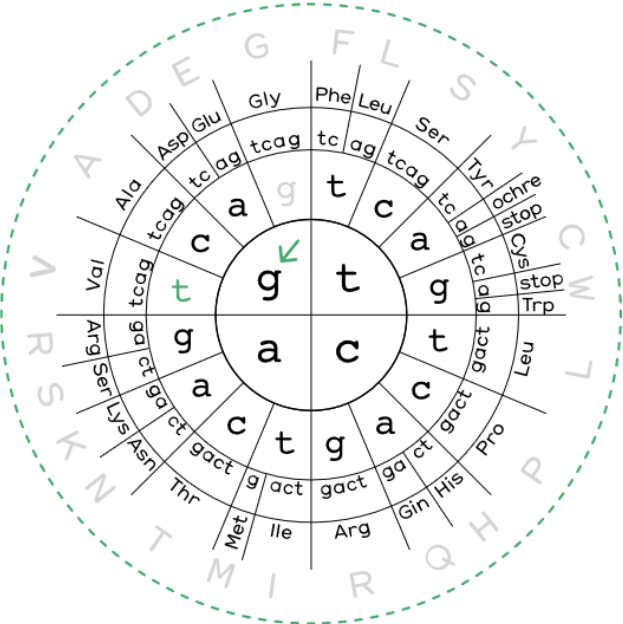
Cette mutation (ou variant) est en position 501: elle est appelée **N501Y**
...et le virus anglais est appelé B.1.1.7 (**501Y.V1**) ('Nelly')

Certaines mutations dans la séquence du génome
ne modifie pas la séquence en acide aminé de la protéine correspondante ☺

Exemple:

Génome référence (NC_045512.2) aat ggt gtt
Protéine Spike référence **N G V**
... 501 502 503 ...

Génome variant XX aa**c** ggt gtt
Protéine Spike variant XX ... **G V** ...



Code génétique
(‘redundant’, ‘dégénéré’)



“(...) No doubt you’ve heard about the novel coronavirus variants that are evolving around the world. **There now appear to be more than a dozen versions of SARS-CoV-2, which are of varying degrees of concern because some are linked to increased infectivity and lethality while others are not.** It’s easy to be overwhelmed by this diversity and to fear that we’ll never achieve herd immunity. Yet evidence is growing that these variants share similar combinations of mutations. This may not be the multifront war that many are dreading, with an infinite number of new viral versions.

(...) While **most mutations are one-offs that go extinct**, some establish new lineages that become more frequent as the virus succeeds in replicating and infecting many people.

(...) One pattern we see is called **convergent evolution**, where the same trait emerges in different independent lineages over time, usually as **they adapt to similar environments**.

(...) We must remember that the more infections there are, the more chance mutations will occur, and those that best help the virus to survive will proliferate. **This is why stopping new infections is key.**”

<http://covariants.org/>

[source](#)

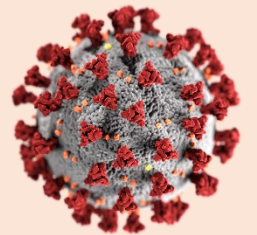
<http://outbreak.info>

Spike variants vs Wuhan-Hu1 (19A)

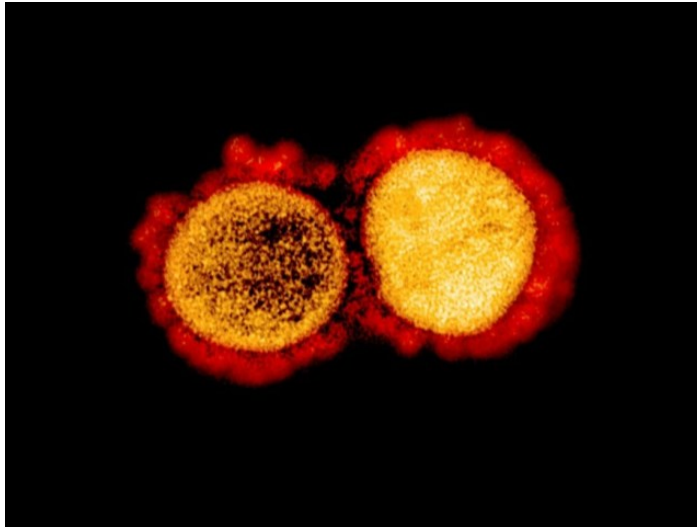
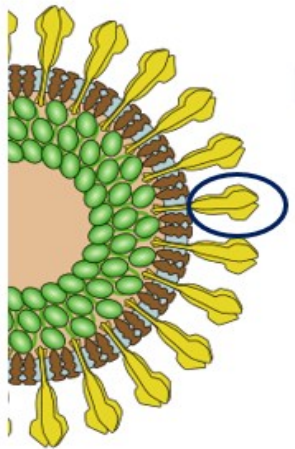
		Common name:	UK	SA	P1 Brazil	P2 Brazil	B.1.1.318	B.1.375	B.1.427	CAL20C	B.1.525	B.1.526	EU1	EU2	A.23.1	A.28	A.27
		PANGO name:	B.1.1.7	B.1.351	B.1.1.28.1	B.1.1.28.2				B.1.429			B.1.177	B.1.7		19B/501T	19B/501Y
		Nextstrain name:	20I/501Y.V1	20H/501Y.V2	20J/501Y.V3	20?				20C			20A	20A			
Neutralizing antibodies binding	13	Signal	L5F									F					
		NTD	S131						I	I							
			L18F	(F)	F												F
			T20N		N												
			P26S		S												
			Q52R								R						
			Del HV 69-70	Del			Del									Del	
			D80A	A													
			T95I				I					I					
			R102I												(I)		
			D138 Y		Y												
			del Y144	Del			Del										
		S1	W152C						C	C							
			F157L												L		
			R190S		S												(S)
			D215G	G													
			A222V										V				
			del LAL 242-244	(Del)													
	319	RBD	D253G									G					
ACE2 receptor and Neutralizing antibodies binding			V367F													F	
			K417N/T	N	N/T												
Cleavage S1/S2			L452R						R	R							R
			S477N									(N)		N			
			E484K	K	K	K	K				K	(K)					
			N501Y/T	Y	Y	Y										T	Y
Fusion activity; modulates tropism			A570D	D													
			Q613H												H		
			D614G	G	G	G	G	G	G	G	G	G	G	G			
			A653V														V
			H655Y		Y											Y	Y
			Q677H				H				H						(H)
			P681H/R	H											(R)		
			A701V		V							V					
			T716I	I													
			D796Y/H				H										Y
Fusion activity; modulates tropism			T859I			(I)											
			F888L								L						
		HR1	S982A	A													
			T1027I		I												
			D1118H	H													
		HR2	V1176F		F	(F)											
			I1183V														(V)
	1213		G1219V														V

7 - De la séquence à la structure 3D (protéine Spike)

[Bad News Wrapped in Protein: Inside the Coronavirus Genome](#) (New-York Times)

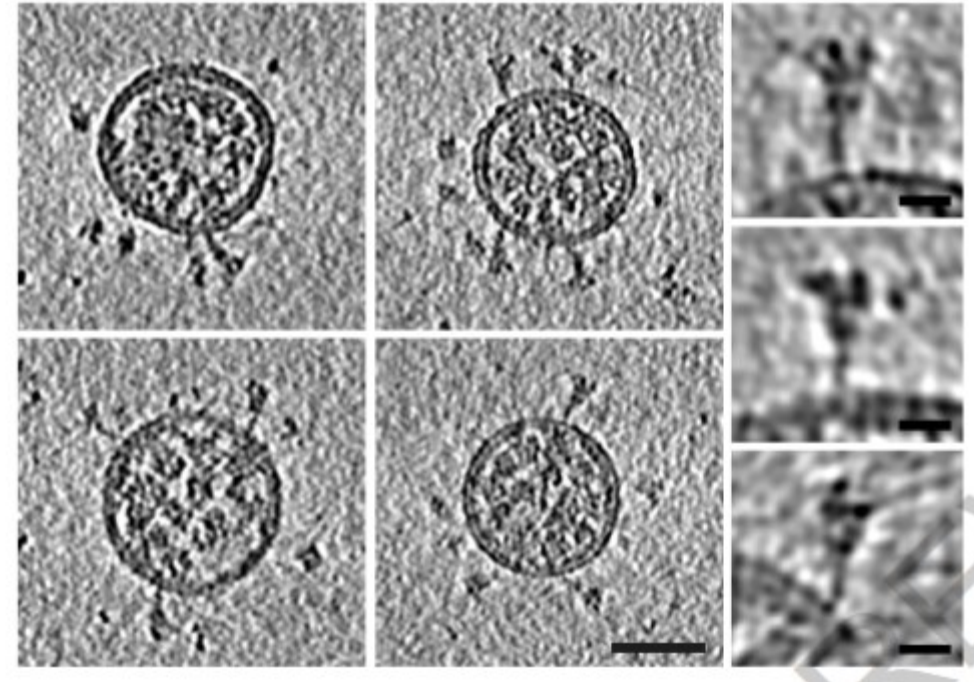


Images cryo-EM de la protéine Spike à la surface du virion



Une image du virus SARS-CoV-2 transmise par les Instituts nationaux de santé (NIH/NIAD) le 11 août 2020.

NIH NATIONAL INSTITUTE OF ALLERGY AND INFECTIOUS DISEASES/AFP/ARCHIVES - HANDOUT



https://www.nature.com/articles/s41586-020-2665-2_reference.pdf

Voici une représentation de la **structure 3D de la protéine Spike** et de sa séquence en acides aminés (bande de données PDB)

Depuis:

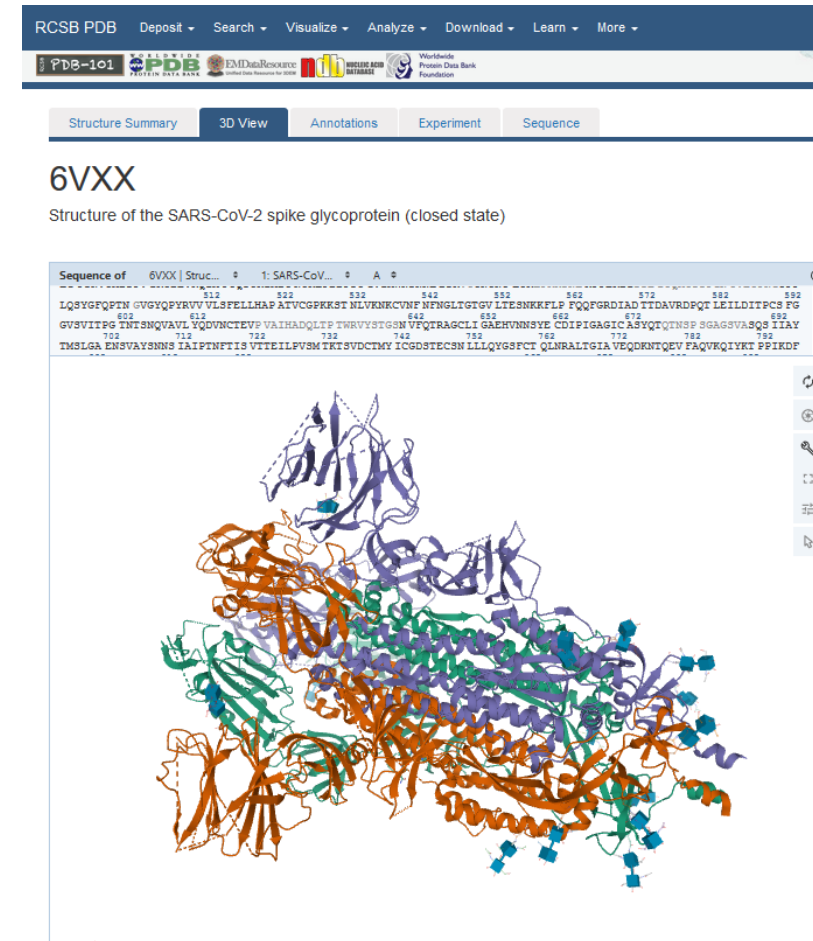
<https://www.rcsb.org/3d-view/6vxx>



Cliquer sur un acide aminé pour voir où il est localisé dans la structure 3D.

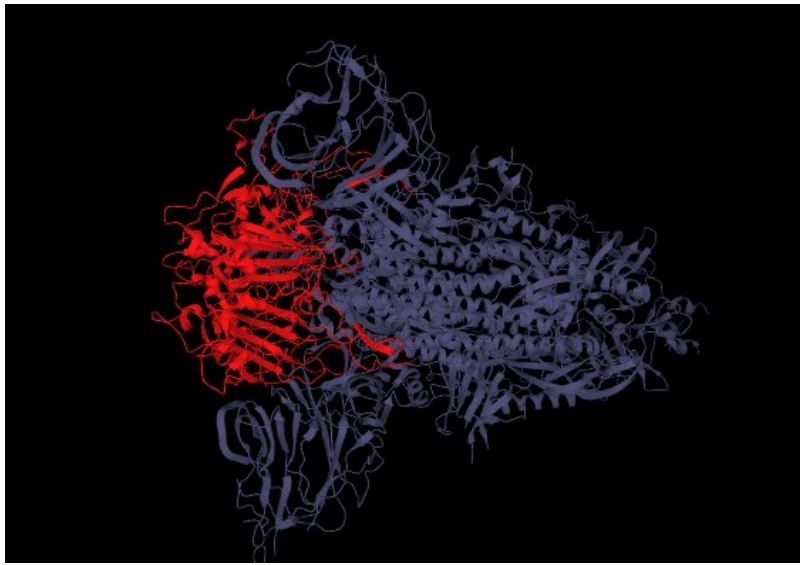
Expert: comparer les 2 formes de la protéine Spike

- 6vxx: Spike forme fermée
- 6vsb: Spike forme ouverte

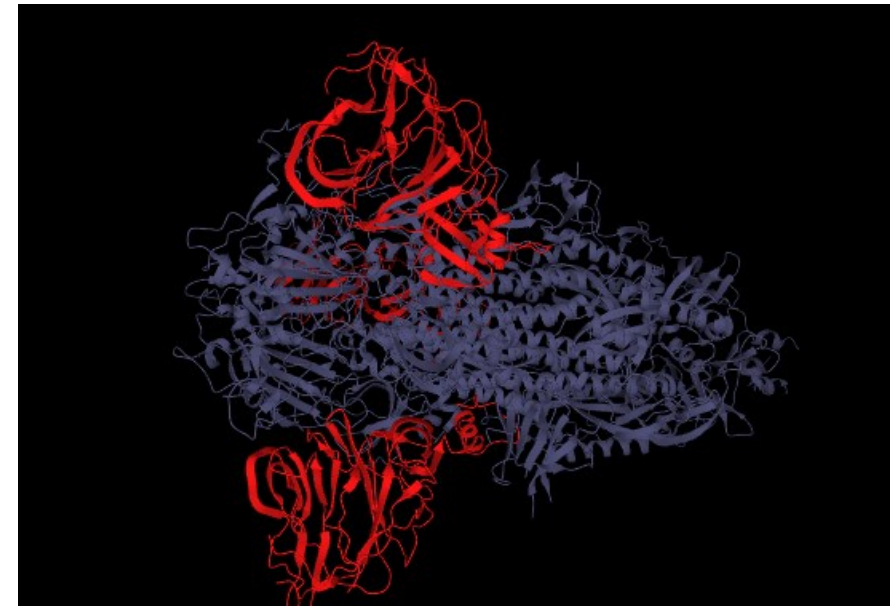
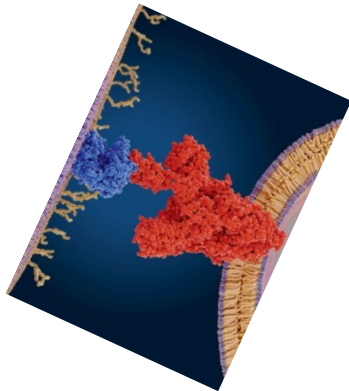


Note: les cubes bleus représentent des sucres

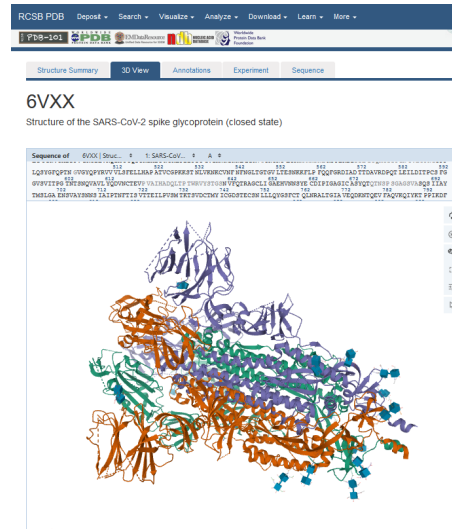
[Nature, août 2020](#)



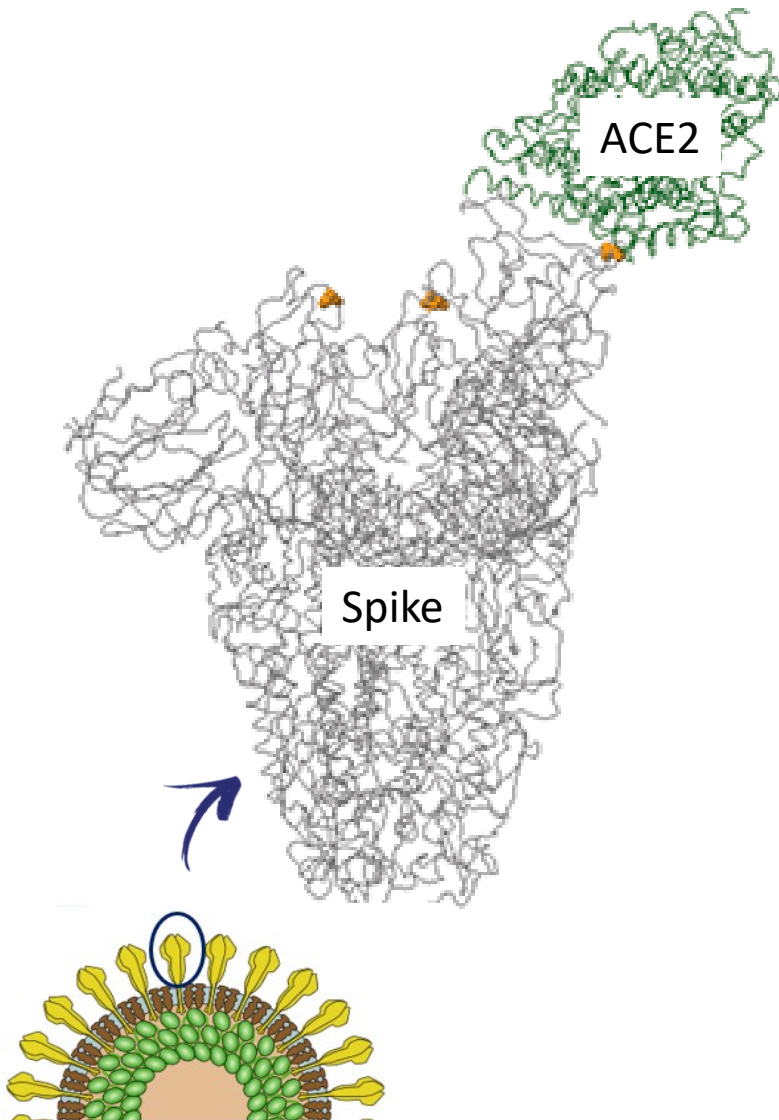
Région de la protéine Spike
impliquée
dans l'interaction avec les
cellules humaines
'infectivité'



Région de la protéine Spike
impliquée
dans l'interaction avec certains
anticorps
'réaction au vaccin'



<https://viralzone.expasy.org/9556>

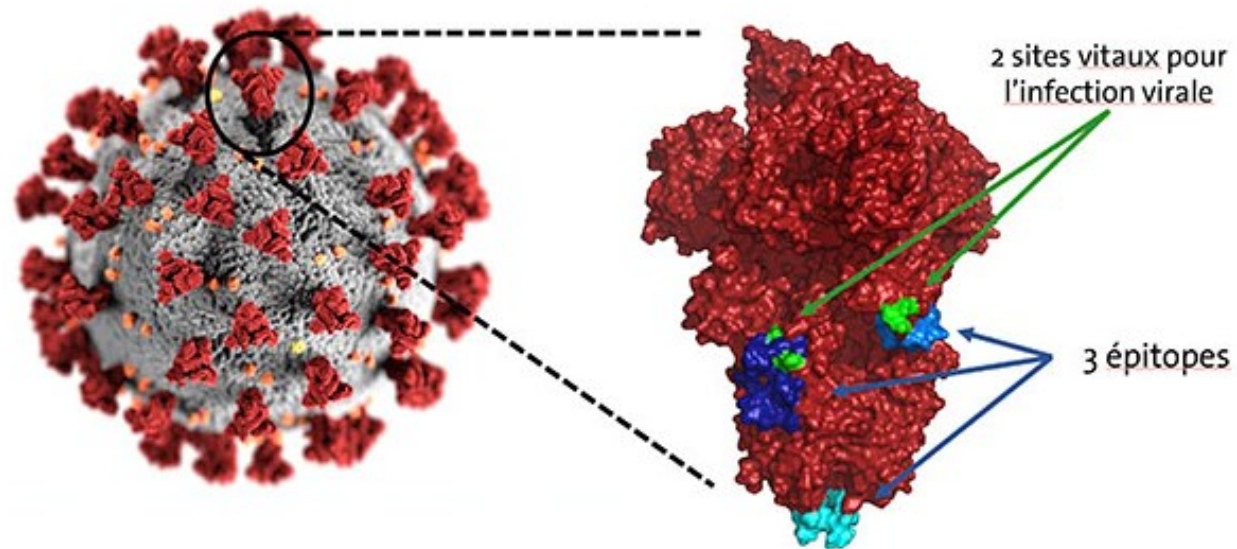


La guerre des clones....

La 'variante britannique' du SARS-CoV-2 (appelée B.1.1.7), est un virus parmi d'autres qui s'est développé particulièrement rapidement en Angleterre en décembre 2020.

- 17 mutations sont présentes dans le génome du virus B.1.1.7: elles entraînent toutes des changements d'acides aminés dans les protéines du virus, dont 9 mutations dans la protéine Spike.
 - La mutation **N501Y** de la protéine Spike (en jaune sur cette image) pourrait modifier les interactions de la protéine avec les récepteurs présents sur les cellules humaines (ACE2).

Des scientifiques de l'UNIGE et des HUG ont identifié les parties de l'enveloppe du SARS-Cov-2 qui sont le plus fréquemment visées par les anticorps. Elles représentent une cible potentielle pour le développement d'un vaccin.



Localisation des trois sites ciblés le plus fréquemment par les anticorps humains sur les clous (ou "spikes") du coronavirus responsable du COVID-19 @UNIGE

<https://www.unige.ch/communication/communiques/2020/deux-failles-potentielles-dans-la-cuirasse-du-coronavirus/>

Voici une représentation de la **structure 3D de la protéine Spike** et de sa séquence en acides aminés (bande de données PDB)

Depuis:

<https://www.rcsb.org/3d-view/6vxx>

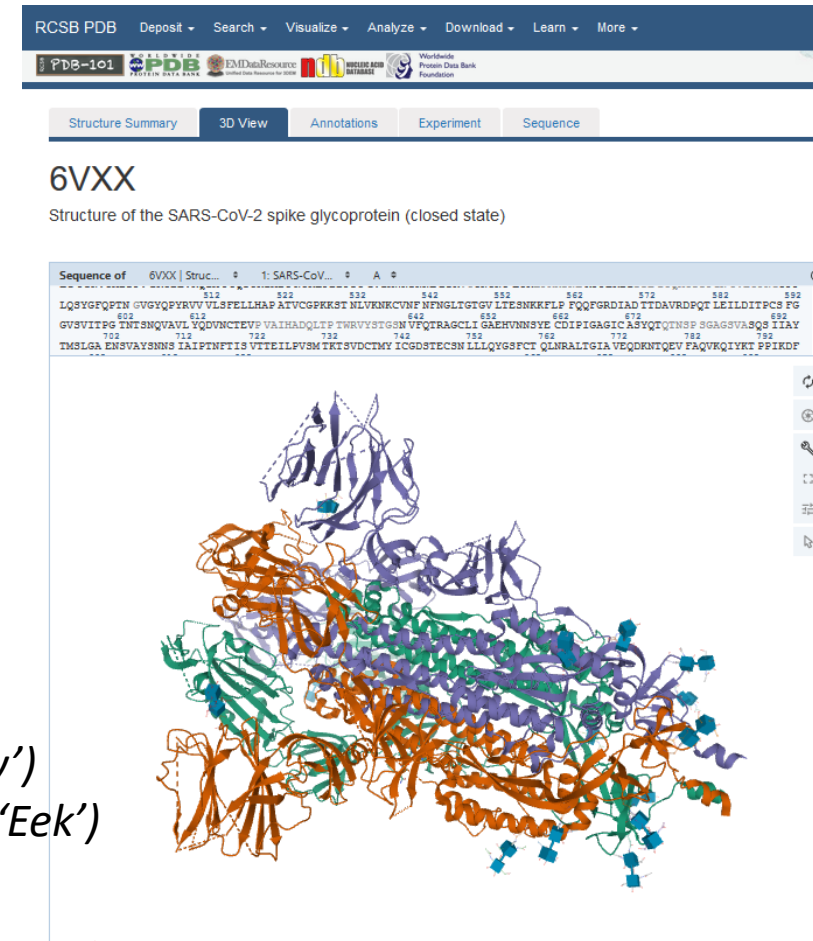
Cliquer sur un acide aminé pour voir où il est localisé dans la structure 3D.

Localiser l'acide aminé en position 614, une mutation très présente pendant l'été 2020 (*'Doug'*)

"Making sense of mutation: what **D614G** means for the COVID-19 pandemic remains unclear"

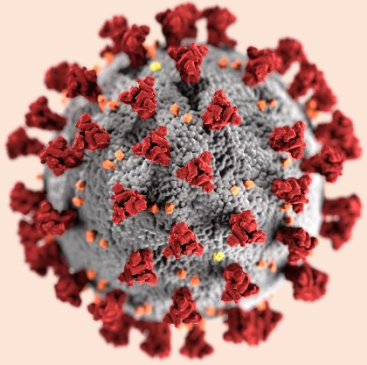
Localiser une mutation 'variant anglais': N501Y (*'Nelly'*)

Localiser une mutation 'variant sud africain' : E484K (*'Eek'*)



Note: les cubes bleus représentent des sucres

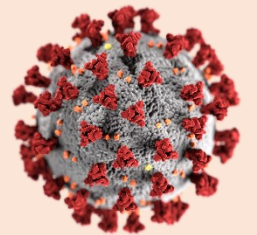
[Nature, août 2020](#)



8 - A la recherche de l'origine de SARS-CoV-2

Bioinformatique

Banque de données de taxonomie: [Taxonomy@NCBI](https://www.ncbi.nlm.nih.gov/taxonomy)



Les coronavirus infectent de nombreuses espèces de mammifères



SARS coronavirus (*homme*)
Civet coronavirus (*civette*)
Hedgehog coronavirus (*hérisson*)
Bat coronavirus (*chauve-souris*)
Rabbit coronavirus (*lapin*)
Camel coronavirus (*chameau*)
Dog coronavirus (*chien*)
Rat coronavirus (*rat*)
Bovine coronavirus (*vache*)
Equine coronavirus (*cheval*)
Yak coronavirus (*yak*)
Pangolin coronavirus (*pangolin*)
Porcine coronavirus (*cochon*)



Certains coronavirus provoquent également des bronchites chez les poulets et une maladie intestinale mortelle chez les porcs

*La classification des virus est
particulièrement complexe....
Il n'est pas toujours facile de s'y retrouver...*

Question à se poser pour étudier l'origine de SARS-CoV-2:

Quel coronavirus est le plus similaire à SARS-CoV-2 ?

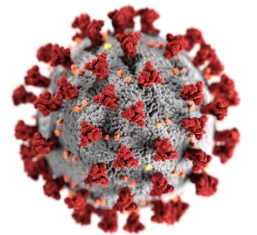
Bioinformatique

Outil de recherche de similarité: [BLAST](#)

Banque de données de protéines: [UniProtKB/Swiss-Prot](#)

Note: Il existe plusieurs milliers de coronavirus de chauve-souris....

20minutes: [Les 4 scénarios étudiés par les experts sur l'origine du virus](#)



La protéine Spike chez différents coronavirus (1)

Voici des séquences partielles de la protéine Spike de différents coronavirus infectant différentes espèces, différentes années.

Quelle séquence est la plus similaire à human SARS-CoV-2 ?

Quelle(s) hypothèse(s) pourriez-vous faire ?

Human SARS-CoV-2	(2020)	IRGDEVVRQIAPGQTGKIAD
Pangolin coronavirus	(2020)	VRGDEVVRQIAPGQTGRIAD
Human SARS-CoV	(2003)	VKGDDVRQIAPGQTGVIAD
Bat coronavirus	(2020)	ITGDEVVRQIAPGQTGKIAD



wikipedia



wikipedia

La protéine Spike chez différents coronavirus (2)

Voici des séquences partielles de la protéine Spike de différents coronavirus infectant différentes espèces, différentes années.

```
>Human_SARS2020
FSTFKCYGVSP TKLNDLCFTNVYADSFVIRGDEV RQIAPGQTGKIAD
>Civet_2003_coronavirus
FSTFKCYGVSA TKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIAD
>Pangolin_2020_coronavirus
FSTFKCYGVSP TKLNDLCFTNVYADSFVVRGDEV RQIAPGQTGRIAD
>Human_SARS2003
FSTFKCYGVSA TKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIAD
>Bat_2020_coronavirus
FSTFKCYGVSP TKLNDLCFTNVYADSFVITGDEV RQIAPGQTGKIAD
>Human_MERS2012
VNDF TCSQISPA AIASNCYSSLILDYFSYPLSMKSDLSVSSAGPISQ
>Bat_2007_coronavirus
VDEF SCNGISPDSI ARG CYSTLTVDYFAYPLSMKSYIRPGSAGNIPL
```

La protéine Spike chez différents coronavirus (2)

Comparer les séquences partielles de la protéine Spike de différents coronavirus, en faisant des alignements 2 à 2 avec la protéine Spike de SARS-CoV-2 (2020) à l'aide de [Align@UniProt](#).

Remplir le tableau suivant:

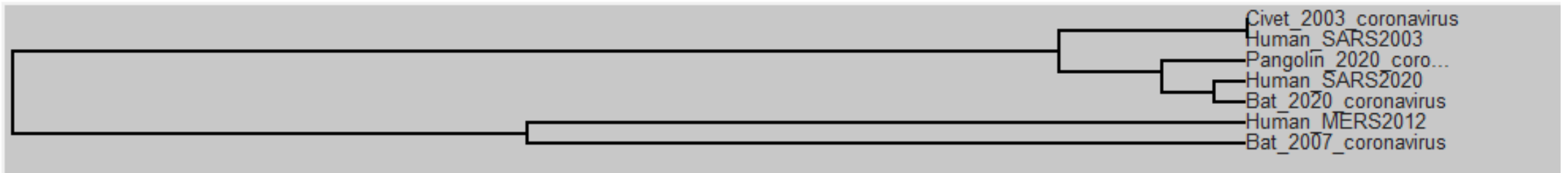
% identité	Human_SARS2020
Human_SARS2020	100
Human_SARS2003	
Civet_2003_coronavirus	
Pangolin_2020_coronavirus	
Bat_2020_coronavirus	
Human_MERS2012	
Bat_2007_coronavirus	

% identité	Human_SARS2020
Human_SARS2020	100
Human_SARS2003	82.2
Civet_2003_coronavirus	82.2
Pangolin_2020_coronavirus	86.7
Bat_2020_coronavirus	85.9
Human_MERS2012	18.1
Bat_2007_coronavirus	27.7

Réponse:

La protéine Spike chez différents coronavirus (2)

Tree



Attention: il s'agit de découvrir la démarche scientifique
et de formuler des hypothèses et en aucun cas des conclusions !
Des conclusions sur les chaînes de transmission sont impossibles à faire
avec si peu de données!

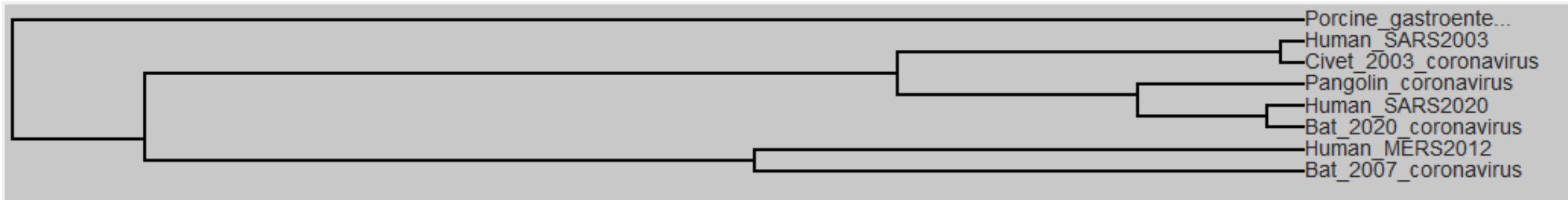
L'échantillonnage (présence ou absence d'une séquence (coronavirus
du chameau 2012)) et/ou des erreurs de séquençage peuvent
influencer de manière significative l'interprétation

La protéine Spike chez différents coronavirus (3)

Copier/coller les [séquences complètes*](#) des protéines Spike de différents coronavirus dans [Align@UniProt](#)

Voici une représentation très simplifiée des relations évolutives existant entre différents coronavirus infectant différentes espèces de mammifères

Tree



* http://education.expasy.org/cours/Outreach/BIOINFOCOVID/Spike_fasta_renamed.txt

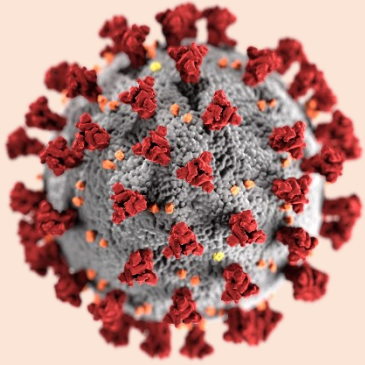
Quelle(s) chaîne(s) de transmission de SARS-CoV-2 serait envisageable selon ces résultats ?

homme -> homme	1
cochon -> homme	2
civette -> homme	3
chauve-souris -> homme	4
pangolin -> homme	5
chauve-souris -> pangolin -> chauve-souris -> homme	6

“Researchers examining coronaviruses and bats are used to working with large numbers.

A 2017 study of 12 333 bats from Latin America, Africa, and Asia found that almost 9% carried at least one of 91 distinct coronaviruses. The authors estimated that there are at least 3200 coronaviruses that infect bats. Moreover, there are over 1400 species of bat. Figuring out which ones are susceptible to which coronaviruses is no small task. The plethora of bat coronaviruses, coupled with the uncertainty about the role of an intermediary animal, makes it also tricky to know how to go about preventing a future spillover.”

[https://www.thelancet.com/journals/laninf/article/PIIS1473-3099\(20\)30641-1/fulltext](https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30641-1/fulltext)



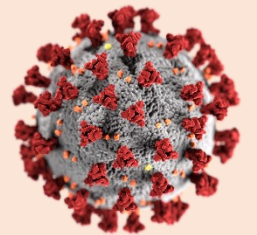
9 – A la recherche d'un traitement ...

Biologie:

C'est quoi un médicament?

Bioinformatique

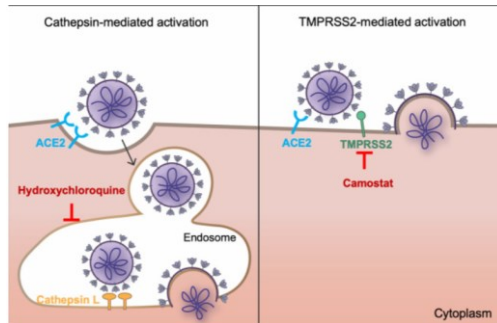
Banque de données de structure 3D: [PDB](#)





<https://viralzone.expasy.org/9078>

Un schéma du cycle d'infection du virus dans une cellule humaine et des **principaux traitements** en voie d'investigation: **vaccin** (neutralizing antibody) et **molécules médicament** ciblant différentes voies biologiques et protéines du coronavirus.



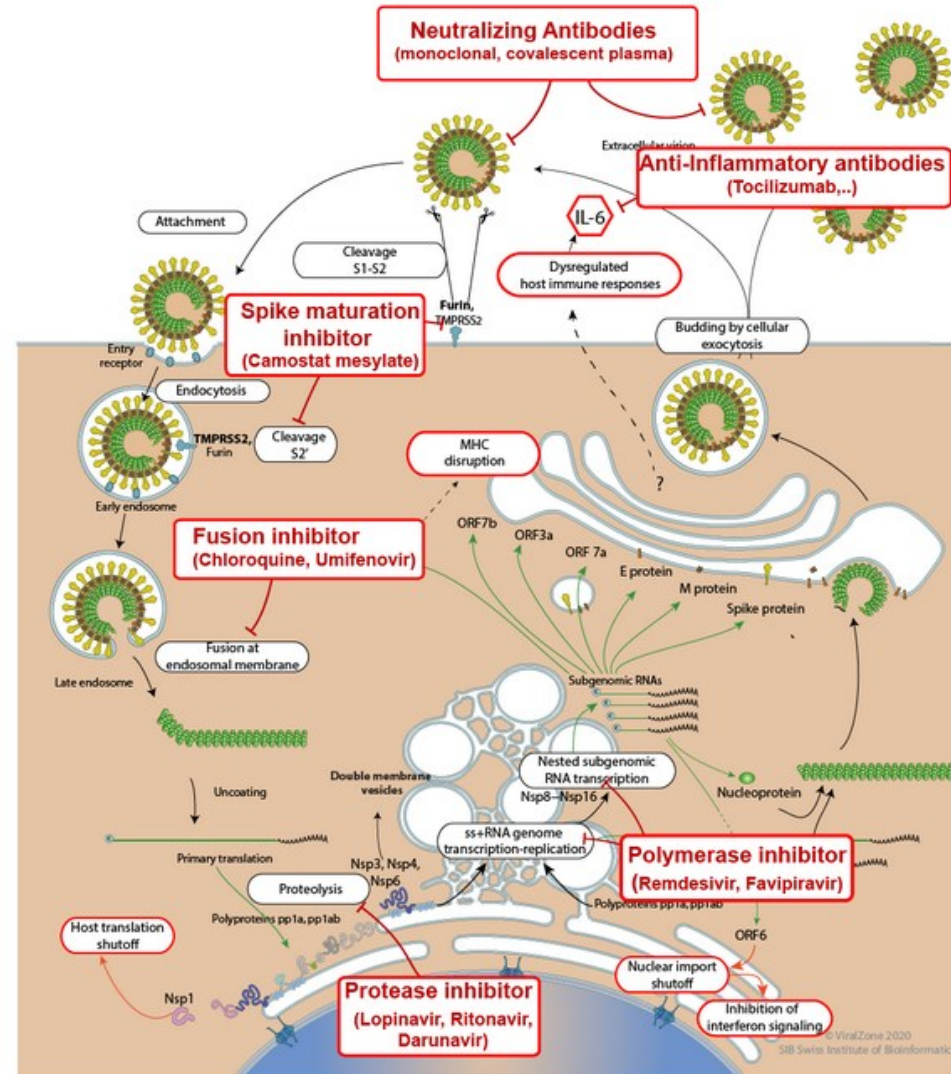
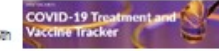
Le virus pourrait se passer de la 'voie endosome': ceci pourrait expliquer l'échec de la chloroquine...

<https://pubmed.ncbi.nlm.nih.gov/33465165/>

Antiviral drugs

There is no antiviral drug globally accepted for treatment, but many drugs are under investigations. About 21 trials of antivirals are under development (5th May 2020).
Antiviral drugs are difficult to find, because of the cell-parasitic nature of viruses. It has been long to get efficient drugs against HIV or Hepatitis C. COVID-19 vaccine may arrive before we get efficient antivirals.

External resource

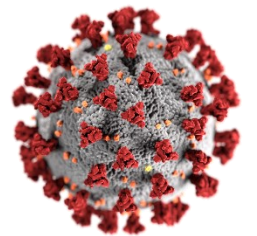


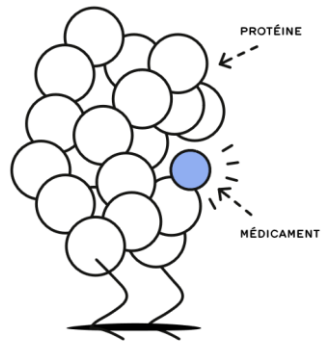
ANTIVIRAL DRUGS This table displays the main drugs under investigation

Médicaments ?

[C'est quoi un médicament?](#)

[Une revue sur la recherche de médicaments](#)





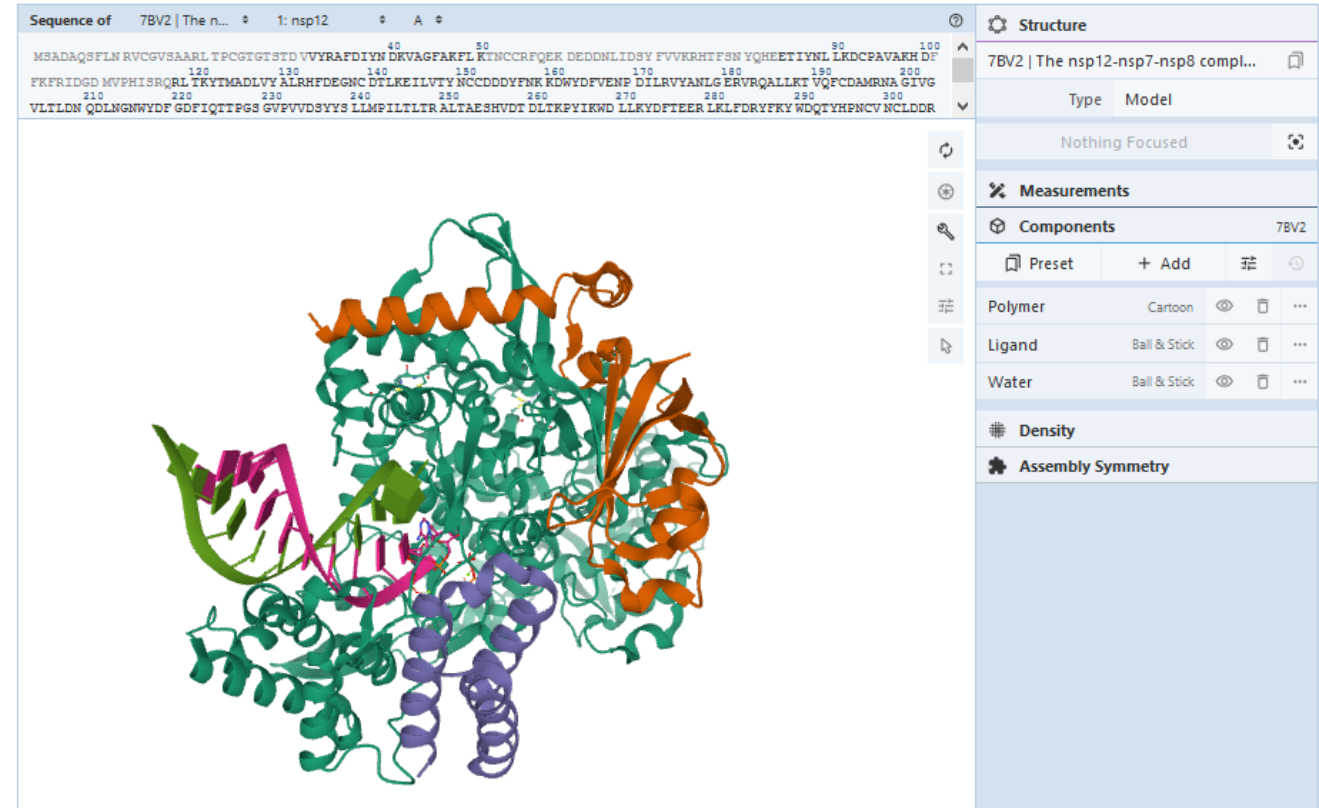
L'interaction protéine-médicament est un peu à l'image de l'interaction d'une clé avec une serrure: elle dépend pour beaucoup de la forme du médicament et de la protéine.

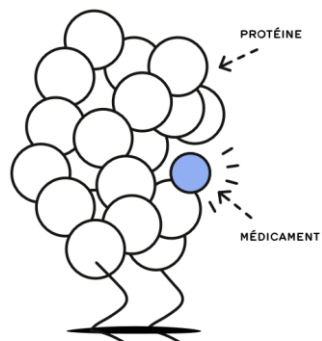
Structure 3D de la polymérase du coronavirus en présence du médicament Remdesivir ([Publication](#))

- Visualiser l'ARN en train d'être dupliqué (double brin) et le médicament
- Comment les chercheurs conçoivent-ils les médicaments de demain? [Atelier Drug Design](#)

7BV2

The nsp12-nsp7-nsp8 complex bound to the template-primer RNA and triphosphate form of Remdesivir(RTP)



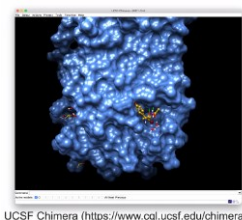


L'interaction protéine-médicament est un peu à l'image de l'interaction d'une clé avec une serrure: elle dépend pour beaucoup de la forme du médicament et de la protéine.

[Structure 3D de la protéase](#) du coronavirus en présence d'un inhibiteur potentiel ([Publication](#))

La même protéase avec un autre médicament (en cours de validation):

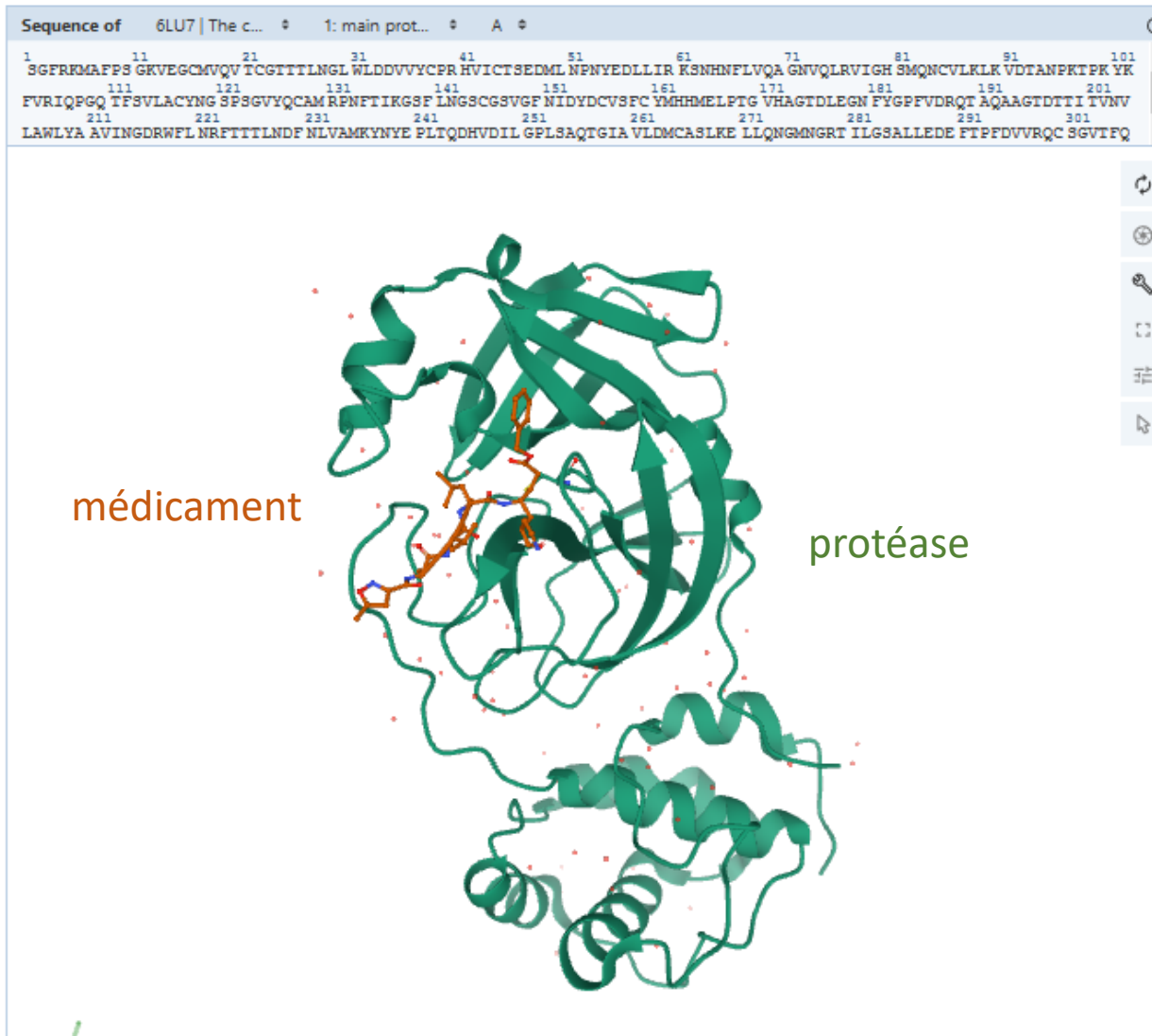
<https://www.rcsb.org/3d-view/5RG1/>



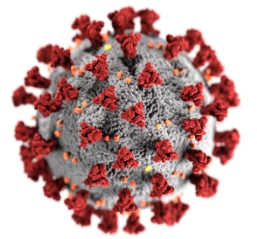
- Comment les chercheurs conçoivent-ils les médicaments de demain? [Atelier Drug Design](#)
- [La protéase dans UniProtKB/Swiss-Prot](#) (3C-like proteinase)

6LU7

The crystal structure of COVID-19 main protease in complex with an inhibitor N3



Vaccin ?



SARS-CoV-2 and Neutralizing Antibodies, 2020



<http://pdb101.rcsb.org/sci-art/goodsell-gallery/sars-cov-2-and-neutralizing-antibodies>

Acknowledgement: David S. Goodsell, RCSB Protein Data Bank and Springer Nature; doi: 10.2210/rcsb_pdb/goodsell-gallery-025

This painting shows a cross section through SARS-CoV-2 surrounded by blood plasma, with neutralizing antibodies in bright yellow. The painting was commissioned for the cover of a special COVID-19 issue of Nature, presented 20 August 2020.

It incorporates information from two cryoelectron microscopy studies that explore the shape and distribution of spikes and the nucleoprotein:

Yao H et al. (2020) Molecular architecture of the SARS-CoV-2 virus. bioRxiv preprint DOI: [10.1101/2020.07.08.192104](https://doi.org/10.1101/2020.07.08.192104)

Ke Z et al. (2020) Structures, conformations and distributions of SARS-CoV-2 spike protein trimers on intact virions. bioRxiv preprint DOI: [10.1101/2020.06.27.174979](https://doi.org/10.1101/2020.06.27.174979)

Vidéo 4 (2:26)

Vaccins ou médicaments anti-viraux : qu'est-ce qui est le plus facile à développer ?

Menu icon YouTube^{CH} Search



Swiss Institute of Bioinformatics

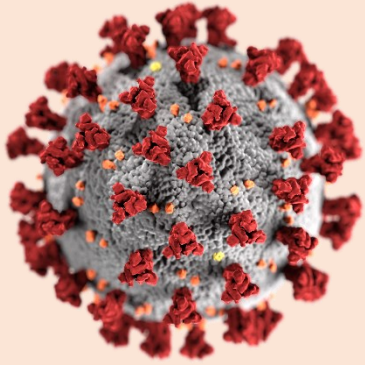


Vaccins ou médicaments anti-viraux:
qu'est-ce qui est le plus 'facile' à développer ?

Phillippe Le Mercier
Biologiste, Doctorat en virologie - Institut Pasteur
ViralZone, Groupe Swiss-Prot, Genève
SIB Swiss Institute of Bioinformatics



Play button icon 0:00 / 2:25 CC Settings Fullscreen

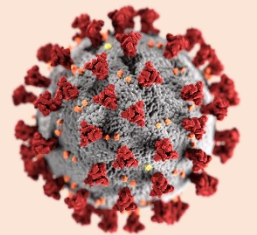


10 - Un exemple de recherche menée au SIB

Référence:

Christian Sigrist, Alan Bridge, Philippe Le Mercier

DOI:<https://doi.org/10.1016/j.antiviral.2020.104759> (pdf)



Le virus SARS-CoV-2 est un cousin très proche du virus SARS-CoV,
responsable de l'épidémie de 2003,
qui avait infecté plus de 8'000 personnes dans 30 pays différents.

Des chercheurs du SIB ont eu l'idée de **comparer** la séquence en acide aminé de la protéine Spike de différents coronavirus.

Ils ont comparé en particulier la protéine Spike du coronavirus SARS-CoV-2 avec la protéine Spike du coronavirus SARS-CoV, le virus qui avait infecté l'homme en 2003.

Comparer ces 2 séquences en acide aminé:

>Spike_SARS-CoV-2

RISNCSVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVVRQIAPGQTG

>Spike_SARS-CoV

KISNCSVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTG

Vous pouvez le faire manuellement ou à l'aide de l'outil bioinformatique [Align@UniProt](https://www.uniprot.org/align)

Combien de différences ?

Retrouvez-vous 3 acides aminés consécutifs 'RGD' dans une des séquences ? Laquelle ?

Comparer les séquences en acide aminé complètes des protéines Spike de SARS-CoV-2 et de SARS-CoV

Spike_SARS-CoV-2: [lien](#)

Spike_SARS-CoV: [lien](#)

Copier et coller les séquences dans [l'outil Align@Uniprot](#)

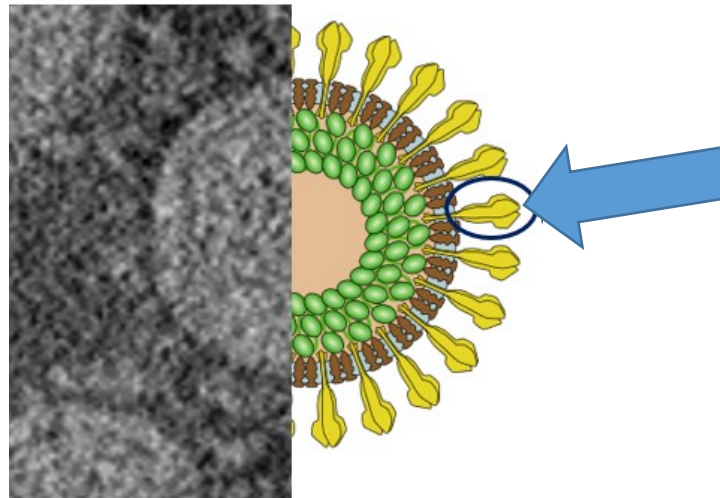
Chercher le motif RGD dans l'alignement (Ctrl F)

A quelle position dans la séquence de SARS-CoV-2 le retrouvez-vous ?

Ces chercheurs ont ainsi découvert la présence d'un **motif RGD** dans la séquence de la protéine de SARS-CoV-2, motif qui est absent dans la séquence de la protéine Spike de SARS-CoV.

Ce motif est très célèbre en virologie, car il permet aux virus d'entrer dans les cellules humaines en utilisant une autre porte d'entrée que ACE2, une autre protéine appelée **intégrine**.

Pour que le motif RGD puisse jouer un rôle dans l'interaction avec les cellules humaines, il doit se trouver à la surface de la protéine Spike



Est-ce que le motif RGD est localisé à la surface de la protéine Spike ?

Pour répondre à cette question, il faut étudier la structure 3D de la protéine Spike.

Voici une représentation de la **structure 3D de la protéine Spike** et de sa séquence en acides aminés (bande de données PDB):

RCSB PDB Deposit Search Visualize Analyze Download Learn More

PDB-101 PDB EMDatabank EMDataResource Worldwide Protein Data Bank

Structure Summary 3D View Annotations Experiment Sequence

6VXX

Structure of the SARS-CoV-2 spike glycoprotein (closed state)

Sequence of 6VXX | Struc... 1: SARS-CoV... A

LQSYGFQPTN GVGYPYRVV VLSPELLHAP ATVCGPKKST NLVNRKCVNF NFNGLTGTGV LTESNKKFLP FQQFGRDIAD TIDAVRDPQT LEILDITPCS PG
GVSVITPG TINSQVAVL YQDVNCTEVP VAIHADQLIP TWRVYSTGSH VFQTRAGCLI GAHVNSYE CDIPGAGIC ASYQTINSP SGAGSVASQS IIAV
TMSLGA ENSVAYSNNIS IAIPTNFTIS VTTEILPVSM TRTSVDCTMY ICGDSTECNS LLLQVGSFCT QLNRLALGIA VEQDRNTQEV FAQVQIYKT PPIKDF

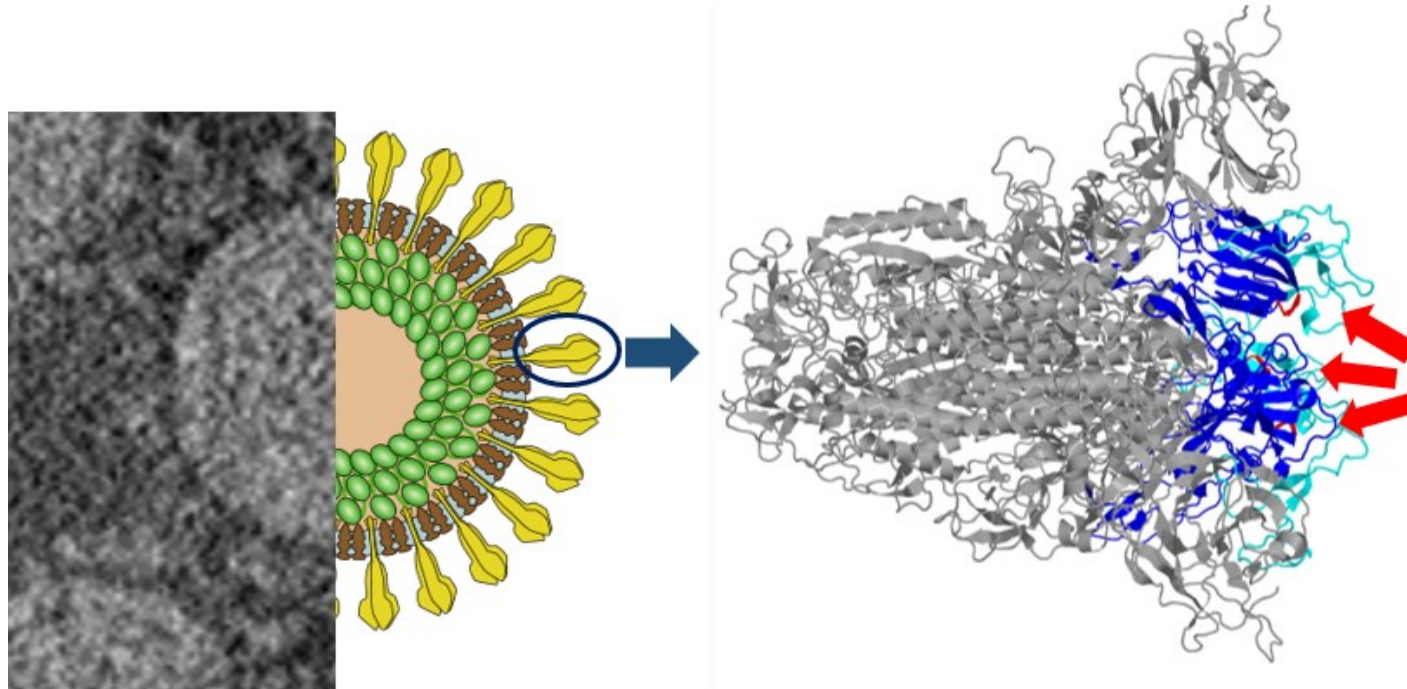
<https://www.rcsb.org/3d-view/6vxx>

Localiser le motif RGD dans la structure 3D de la protéine Spike

Truc: le motif RGD est en position 403-405 dans la séquence

[Nature, août 2020](#)

En utilisant les données connues de la structure 3D de la protéine Spike du SARS-CoV-2, les chercheurs du SIB ont montré que le motif RGD (en rouge) est localisé à la surface de la protéine Spike, proche de la région qui est impliquée dans l'interaction avec les protéines 'récepteur' humaines (en bleu).



En résumé: Spike, le motif RGD et les intégrines

Le motif RGD, retrouvé à la surface de la protéine Spike du virus SARS CoV-2, est connu pour favoriser l'interaction avec d'autres protéines humaines, appelées intégrines.

Ce motif a été retrouvé chez tous les virus SARS-CoV-2 qui ont été séquencés jusqu'à maintenant. Il est possible que le virus ait acquis ce motif au cours de son évolution et ait ainsi acquis la capacité de lier les intégrines pour favoriser son entrée dans les cellules de son hôte, mais cela reste à prouver.

Il existe actuellement peu de molécules antivirales efficaces contre le SARS-CoV-2. Les agents qui bloquent la liaison aux intégrines pourraient constituer une piste prometteuse. Les bloqueurs connus de la liaison aux intégrines comprennent entre autre l'anticorps natalizumab utilisé pour le traitement de la sclérose en plaques et la maladie de Crohn.



Antiviral Research

Volume 177, May 2020, 104759



A potential role for integrins in host cell entry by SARS-CoV-2

Christian JA Sigrist , Alan Bridge , Philippe Le Mercier  

Swiss-Prot Group, SIB Swiss Institute of Bioinformatics, Switzerland

Received 20 February 2020, Revised 26 February 2020, Accepted 28 February 2020, Available online 1 March 2020.



Swiss Institute of
Bioinformatics

<https://doi.org/10.1016/j.antiviral.2020.104759>

Grâce à cette information, des médecins espagnols ont continué à traiter un patient Covid-19 avec un médicament ciblant les intégrines



Multiple Sclerosis and Related Disorders

Volume 44, September 2020, 102250

Correspondence

Covid-19 in a patient with multiple sclerosis treated with natalizumab: May the blockade of integrins have a protective role?

Clara Aguirre ^a , Virginia Meca-Lallana ^a, Ana Barrios-Blandino ^b, Beatriz del Río ^a, Jose Vivancos ^c

[Show more](#) 

<https://doi.org/10.1016/j.msard.2020.102250> [Get rights and content](#)

<https://doi.org/10.1016/j.msard.2020.102250>

Grâce entre autre à cette information, des chercheurs sont à la recherche d'un médicament qui bloque l'interaction entre Spike et les intégrines.

The Integrin Binding Peptide, ATN-161, as a Novel Therapy for SARS-CoV-2 Infection

Brandon J Beddingfield ^{1 2}, Naoki Iwanaga ³, Prem P Chapagain ^{4 5}, Wenshu Zheng ⁶, Chad J Roy ^{1 2}, Tony Y Hu ⁶, Jay K Kolls ³, Gregory J Bix ^{7 8 9 10 11}

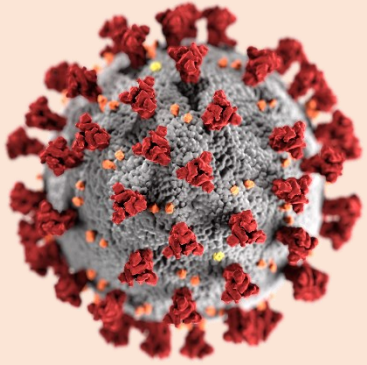
Affiliations + expand

PMID: 33102950 PMCID: [PMC7566794](#) DOI: [10.1016/j.jacbts.2020.10.003](#)

[Free PMC article](#)

Abstract

Many efforts to design and screen therapeutics for the current severe acute respiratory syndrome coronavirus (SARS-CoV-2) pandemic have focused on inhibiting viral host cell entry by disrupting ACE2 binding with the SARS-CoV-2 spike protein. This work focuses on the potential to inhibit SARS-CoV-2 entry through a hypothesized $\alpha 5 \beta 1$ integrin-based mechanism, and indicates that inhibiting the spike protein interaction with $\alpha 5 \beta 1$ integrin (+/- ACE2), and the interaction between $\alpha 5 \beta 1$ integrin and ACE2 using a novel molecule ATN-161 represents a promising approach to treat COVID-19.

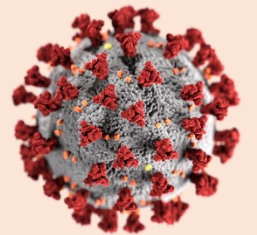


11 - SARS-CoV-2 et HIV

un exemple de fake news

Bioinformatique

Recherche de similarité et notion de 'hasard'



SARS-CoV-2: un virus provenant d'une manipulation humaine ?



- SARS-Cov-2 est très similaire à de nombreuses souches de coronavirus circulant dans la nature en Asie, avant et après la pandémie.
- Ces virus sont connus pour sauter d'une espèce à l'autre sans problème, donc rien de surprenant.
- L'analyse génétique de SARS-CoV-2 présente une organisation génomique et des protéines en tout point similaires aux autres virus sauvages.

Les thèses selon lesquelles une partie d'HIV a été insérée par l'homme dans le virus sont de mauvaises interprétation des résultats de recherche de similarité (BLAST)

Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag

Prashant Pradhan, Ashutosh Kumar Pandey, Akhilesh Mishra, Parul Gupta, Praveen Kumar Tripathi, Manoj Balakrishnan Menon, James Gomes, Perumal Vivekanandan, Bishwajit Kundu

doi: <https://doi.org/10.1101/2020.01.30.927871>

This article is a preprint and has not been certified by peer review [what does this mean?].

Abstract

Info/History

Metrics

 Preview PDF

Abstract

This paper has been withdrawn by its authors. They intend to revise it in response to comments received from the research community on their technical approach and their interpretation of the results. If you have any questions, please contact the corresponding author.

**FAKE NEWS
FACTS**

[Ce papier](#) a été démenti et retiré. Les auteurs ont utilisé des fragments de 6 à 10 acides aminés: il existe près de 1 millions de séquences d'HIV hypervariables, c'est inévitable qu'il y ait des similarités juste par hasard. Ce papier, même si il n'a jamais été publié, a fait le buzz...

Pour les experts (1):

[Emerg Microbes Infect.](#) 2020; 9(1): 378–381.

PMCID: PMC7033698

Published online 2020 Feb 14. doi: [10.1080/22221751.2020.1727299](https://doi.org/10.1080/22221751.2020.1727299)

PMID: [32056509](https://pubmed.ncbi.nlm.nih.gov/32056509/)

HIV-1 did not contribute to the 2019-nCoV genome

[Chuan Xiao](#),^{a,CONTACT} [Xiaojun Li](#),^b [Shuying Liu](#),^c [Yongming Sang](#),^d [Shou-Jiang Gao](#),^e and [Feng Gao](#)^{b,f}

► [Author information](#) ► [Article notes](#) ► [Copyright and License information](#) [Disclaimer](#)

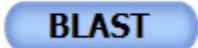
Voici des exemples de séquences d'acides aminés retrouvées dans les protéines de SARS-CoV-2 utilisées pour 'démontrer' que SARS-CoV-2 contient des morceaux de génome de HIV:

TNGTKR
HKNNKS
RSYLTPGDSSSG
QTNSPRRA

La 'E value' est une probabilité de trouver le même résultat par hasard. Plus cette valeur est petite (< 0) plus le match est 'validé'.

Faire un Blast 'protein' au NCBI: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>

Sélectionner Organism: "Human immunodeficiency virus (taxid:12721)"

Cliquer sur 

Créer une séquence aléatoire avec les mêmes 'lettres' ([edit sequence]) et refaire un BLAST.

Que pouvez-vous conclure?

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7033698/>

Pour les experts (2):

Voici un morceau du gène codant pour la protéine Spike:

AATGGTACTAAGAGGTTTGATAACCCTG

Faire un Blast 'nucleotide' au NCBI: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>.

Sélectionner Organism: "Human immunodeficiency virus (taxid:12721)"

Cliquer sur **BLAST**

La 'E value' est une probabilité de trouver le même résultat par hasard. Plus cette valeur est petite (< 0) plus le match est 'validé'.

Créer une séquence aléatoire avec les mêmes 'lettres' ([edit sequence]) et refaire un BLAST.

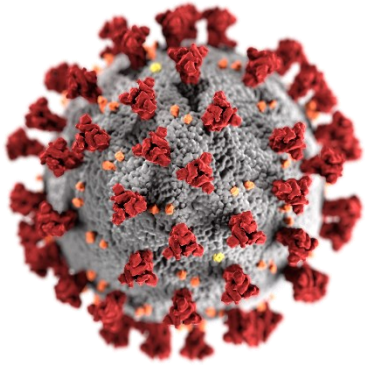
Que pouvez-vous conclure?

A HIV-1 isolate 19828.PPH11 from Netherlands envelope glycoprotein (env) gene, partial cds				
Sequence ID:	HQ644953.1	Length:	1143	Number of Matches: 1
Score	Expect	Identities	Gaps	Strand
38.3 bits(41)	7.5	25/28(89%)	0/28(0%)	Plus/Plus
Query	86	AATGGTACTAAGAGGTTTGATAACCCTG	113	
Sbjct	967	AATGGTACTAAAAGGTTAGATAACACTG	994	
B HIV-1 isolate patient B clone 16.3 from Netherlands envelope glycoprotein (env) gene, complete cds				
Sequence ID:	HQ386166.1	Length:	2580	Number of Matches: 1
Score	Expect	Identities	Gaps	Strand
39.2 bits(42)	2.1	27/31(87%)	0/31(0%)	Plus/Minus
Query	351	CCTAAAAGTTCCTTGTAACTGTATTATT	381	
Sbjct	2523	CCTAAAAGTTCCTTGTAACTATTCTATAATT	2493	

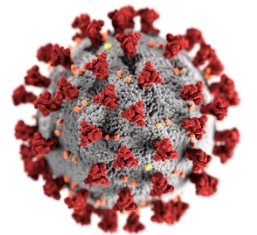
la séquence aléatoire. On peut en conclure que la similarité entre la séquence codante de la protéine S et le génome du VIH n'est pas significative. Les alignements ont été réalisés sur le site BLAST du NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

Figure 6. Recherche de similarités entre les séquences codant pour la protéine spike de CoV2 et le génome de HIV. A. Alignement le plus significatif entre la séquence codant pour la protéine S de SARS-CoV-2 (query) et le génome du VIH (subject). B. Contrôle négatif : alignement le plus significatif entre une séquence aléatoire, obtenue en mélangeant les nucléotides de la séquence précédente, et le génome du VIH. Noter la valeur du score *expect*, qui indique le nombre de faux-positifs attendus au hasard. Ce score présente pour les deux alignements des valeurs supérieures à 1, et est même plus élevé pour l'alignement de la séquence de CoV que pour

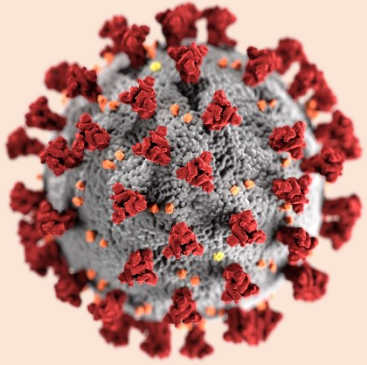
<https://www.medecinesciences.org/fr/articles/medsci/pdf/first/msc200195.pdf>



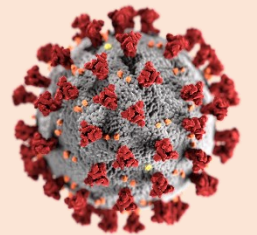
Cet atelier est issu d'une collaboration entre
L'éprouvette, laboratoire public de l'Université de Lausanne
(Service Culture et Médiation Scientifique)
et
le groupe Swiss-Prot - [SIB Swiss Institute of Bioinformatics](https://www.sib.swiss)



Les droits du contenu sont régis par [Creative Commons Attribution – Pas d'Utilisation Commerciale 4.0 International](https://creativecommons.org/licenses/by-nc/4.0/).



12- Dernières nouvelles



Cette séquence est la **séquence du gène codant pour la protéine Spike de SARS-CoV-2**.

Elle est présente dans les vaccins contre SARS-CoV-2 dits 'vaccins à ARN messager (ou ARNm)'. Ces vaccins contiennent une séquence ARNm composée au final de 4284 nucléotides, qui a été modifiée afin de permettre une production optimale de la protéine Spike dans les cellules.

20 années de recherche concentrées dans un vaccin...

"A 30 microgram dose turns out to actually contain 30 micrograms of RNA. (...) This consists of around 13,000 billion repetitions of the same 4284 characters. In addition, there is a clever lipid (fatty) packaging system that gets the mRNA into our cells.(...)."

- Un article (expert): «...[the source code of the BioNTech/Pfizer SARS-CoV-2 vaccine](#)»
- Un article (SimplyScience): [Contre le coronavirus, voici le vaccin à ARNm !](#)

```
>Gene_Spike
atgtttgtttttctgtttttatgtccactagtcctcagtcagtggttaattctacaacc
agaactcaattaccocctgcatacacaattcttccacacggtggtttattaccctgcac
aaagtgttcagatcctcagttttacattcaactcaggactgttcttacccttctttcc
aatgttacttgggtccatgatacagtcctcctgggaccaaaggttaagaggtttgat
aacctgtcctaccatttaagtgggtttattttgtccactgagaagttcaacata
ataagaggtggatttttggtactactttagattcgaagacagtcctactatttatt
aatacgcctactaatgtttattaaagtctgtgaattcaattttgtaagtccattt
ttgggtgtttattaccacaaaaaacaaaaggtggatggaaagtgagttcagagtttt
tctagtcgcaataattgcacttttgaatgtctcctcagccttttctatggacctgaa
ggaaaaacgggttaattcaaaaaatctagggaaatttggtttaagaatattgatggttat
ttaaaaatatttctaagcagcgcctataatttgaatggcgtgatcctcaggggtttt
tcgggttttagaaccttggttagtttgcacaaagttatcaactcactaggtttcaaaat
ttactgtcttacaagaagttatttgactcctgtggtgatttcttcaagttggacagct
gggtgtcagcttattatgtgggttatcttcaactcagagactttctataaaaaataat
gaaaaatggaaacttacagatgtctgagctgtgcaacttgacccctctcagaaacaaag
tgtactgtgaaatccttcaactgtgaaaaaaggaaatctatacaactctcaacttagagtc
caaccaacagaatctattgttagatttctaaattacaaacttgcccttttggtgaa
gtttttaacgccacagattgcatctgtttatgtgtggaacaggaaagaaatcagcaac
tgtgtgtcgtgatattctgtctataaattccgcacattttccactttaagtgttat
ggagtgctcctactaaataatgatctgtcttactaatgtctatcagagattcattt
gtaattagaggtgatgaagtgcagaaatcgcctcaggcgaactggaagattgctgat
tataattaaaaattaccagatgattttacaggtgcgttatagcttggaaattcaacaat
cttgattctaaaggttgggtgaattataattacctgtatagatttttaggaagtctaat
ctcaaaccttttgagagagatttcaactgaaatctacaggccgttagcacacctgt
aatgggttgaaaggttttaattgttaacttcttcaactcaatcatgttttcaacccaat
aatgggttggttaccacacatcagagatgattagtttcttttgaactctatcagca
ccagcaactgtttgtggacataaaagattcactaaatttgggttaaaacaaatgtcact
ttcaacttcaatggtttaacagccagcaggtgttcttactgagtcatacaaaaagtttctg
cctttcaacaatttggcagagacattgtgacactactgagctgtcgtgatccagag
acactgagattctgacattacacactgttcttttgggtgtgctgattataacacca
ggaaacaaatcttcaacacaggttgcgtgttcttcaaggtatttaactgcacagaagtc
cctgtgctattcagatcaacttactcacttggcgtgtttattctacaggttct
aatgttttcaaacacgtgcagcgtgtttaaaggcgtgaacatgcacaaactcatat
gagtgacataccactgggtcaggtatgtcgtgattatcagactcagactaattct
cctggcggcagcgtagtgatgactcaactcactgtgctacactatgtcacttgggt
gcagaaaaattcagttgttactctataaactctattgccaatacccaaaatttactatt
agtttaccacagaaattctaccagtgctatgaccaagacatcagtagattgtacaatg
tacatttgggtgattcaactgaatgcagcaactcttttggcgaattggcagtttttgg
acacaattaaacgcgtgtttaaactggaatagctgttgaacaaacaaaaacccaagaa
gttttgcacaaagtcacaaatattacaaacacccaactaaagatttttgggtgttt
aatttttcaacaattatcagagatcctcaaaacaaagcagaggtcatttattgaagat
ctacttttcaacaaagtgcacactcagagatgctggtctcacaacaaatgtgattgct
ctgtgtgattgtcgtgtagagacactcatttggcacaaaagtttaacggccttactgtt
ttggcacttttgcagagatgaaatgattgctcaatacacttctgcaactgttagcgggt
acaactcactctgggtggacctttgggtcaggtgtgctatcaaaatcactttgctatg
caaatggccttattgtttatgtgtattggagttacacagaattgtctctatgagaaacca
aaattgttgcacaaacttaattagtgctatttggcaaaattcaagactcactttctcc
acagcaagtgcacttggaaaaactcagagatgtggtcaacaaaatgcacaaagctttaaac
acgctgtttaaacaacttgcactcacttttgggtgcaatttcaaggtttttaaagtatac
ctttcagcttgcacaaagttagcgtgaaagtgcacaaattgattgtgattcagagcaga
cttcaaaagttgcagacatattgtgactcaacaataattagagctgcagaaatcagagct
ctctgtaactctgctgctactaaaaatgcagagtggtgacttggcaacataaaagagtt
gatttttgggaaggcctatcatcttatgtctctccctcagtcagcactcaggtgtga
gtctcttgcagtgacttatgtcctgcacaaagaaagaaacttcaacactgctcctgcc
atttgcagtggaagacacacttctcgtgaaggtgtctttgtttcaaatggcaca
cactggtttgaacacaaaggaatttttgaacacaaatcattactacagacaacaca
ttgtgtcgtgtaactgtgattgttgaataggaatttgcacacacagtttatgatcct
ttgcacactgaattagactcattcaaggaggaattagataaatttttaagaatcataca
tcaccaagattgttattagtgacatctcggcattatgtctcagttgttaaacattcaa
aaagaaattgcagcctcactaggttggccaaagatttaaatgaatctctcagatctc
caagaacttggaaagattgagcagtatataaaatggccatggatcatttggcaggttt
atagctggctgtgctcagatgaattgggtgacaaattgtcttggctgtatgacaggttc
gttagttgtcgaaggcgtgtgttcttgggtgactcgtgcacaaattgatgaagacgac
cttgagcagctgtcaaaaggagtcacattacattacacata
```

«The coronavirus is here to stay — here's what that means

A vaccine that is 90% effective at blocking transmission will need to reach at least 55% of the population to achieve temporary herd immunity as long as some social distancing measures — such as face masks and many people working from home — remain in place to keep transmission in check, according to a model developed by Alexandra Hogan at Imperial College London and her colleagues. (A vaccine would need to reach almost 67% of people to provide herd immunity if all social distancing measures were lifted.) But if the rate of transmission increases because of a new variant, or if a vaccine is less effective than 90% at blocking transmission, vaccine coverage will need to be greater to blunt circulation.»

https://www.nature.com/articles/d41586-021-00396-2?utm_source=Nature+Briefing&utm_campaign=6aea13811d-briefing-dy-20210216&utm_medium=email&utm_term=0_c9dfd39373-6aea13811d-45521262

2

Dernières nouvelles... Le variant anglais

La ‘lignée britannique’ du SARS-CoV-2 (B.1.1.7) a été découverte le 8 décembre 2020:

- 17 ‘nouvelles’ mutations sont présentes dans le génome du virus B.1.1.7

Comparaison des génomes de SARS-CoV-2:

- [génom de référence \(NC_045512.2\)](#)
- [génom du variant anglais B.1.1.7 \(LR991698\)](#)

(! seule une toute petite partie de l’alignement est visible ici)

Les différences entre le génome de SARS-CoV-2 (génom de référence) et le génome du virus ‘anglais B.1.1.7’ apparaissent en clair dans cet alignement.

Ces différences correspondent aux ‘nouvelles’ mutations présentes dans le virus anglais.

ceci est une mutation



LR991698.2	23023	ATCATATGGTTTCCAACCCACTTATGGTGTGGTTACCAACCATAACAGAGTAGTAGTACT	23082
NC_045512.2	23041	ATCATATGGTTTCCAACCCACTAATGGTGTGGTTACCAACCATAACAGAGTAGTAGTACT	23100
		*****:*****	
LR991698.2	23083	TTCTTTTGAACCTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTT	23142
NC_045512.2	23101	TTCTTTTGAACCTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTT	23160

LR991698.2	23143	GGTTAAAAACAAATGTGTCAATTTCAACTTCAATGGTTTAAACAGGCACA	23202
NC_045512.2	23161	GGTTAAAAACAAATGTGTCAATTTCAACTTCAATGGTTTAAACAGGCACA	23220

LR991698.2	23203	TGAGTCTAACAAAAAGTTTCTGCCTTTCCAACAATTTGGCAGAGACATTGATGACACTAC	23262
NC_045512.2	23221	TGAGTCTAACAAAAAGTTTCTGCCTTTCCAACAATTTGGCAGAGACATTGCTGACACTAC	23280

LR991698.2	23263	TGATGCTGTCCGTGATCCACAGACACTTGAGATTCTTGACATTACACCATGTTCTTTTGG	23322
NC_045512.2	23281	TGATGCTGTCCGTGATCCACAGACACTTGAGATTCTTGACATTACACCATGTTCTTTTGG	23340

LR991698.2	23323	GTGAGTGTATATAACACCAGGAACAAATACTTCTAACCAGGTGCTGTTCTTTATCA	23382
NC_045512.2	23341	GTGAGTGTATATAACACCAGGAACAAATACTTCTAACCAGGTGCTGTTCTTTATCA	23400

LR991698.2	23383	GGGTGTTAACTGCACAGAAGTCCCTGTTGCTATTTCATGCAGATCAACTACTCCTACTTG	23442
NC_045512.2	23401	GGGTGTTAACTGCACAGAAGTCCCTGTTGCTATTTCATGCAGATCAACTACTCCTACTTG	23460
		** *****	
LR991698.2	23443	GCGTGTTTATTCTACAGGTTCTAATGTTTTTCAAACAGTGCAGGCTGTTTAAATAGGGGC	23502
NC_045512.2	23461	GCGTGTTTATTCTACAGGTTCTAATGTTTTTCAAACAGTGCAGGCTGTTTAAATAGGGGC	23520

LR991698.2	23503	TGAACATGTCAACAACCTCATATGTGTGACATACCCATTGGTGCAGGTATATGCGCTAG	23562
NC_045512.2	23521	TGAACATGTCAACAACCTCATATGTGTGACATACCCATTGGTGCAGGTATATGCGCTAG	23580

LR991698.2	23563	TTATCAGACTCAGACTAATTCTCATCGGCGGGCAGTAGTGTAGCTAGTCAATCCATCAT	23622
NC_045512.2	23581	TTATCAGACTCAGACTAATTCTCATCGGCGGGCAGTAGTGTAGCTAGTCAATCCATCAT	23640

LR991698.2	23623	TGCCTACACTATGTCACTTGGTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGC	23682
NC_045512.2	23641	TGCCTACACTATGTCACTTGGTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGC	23700

2

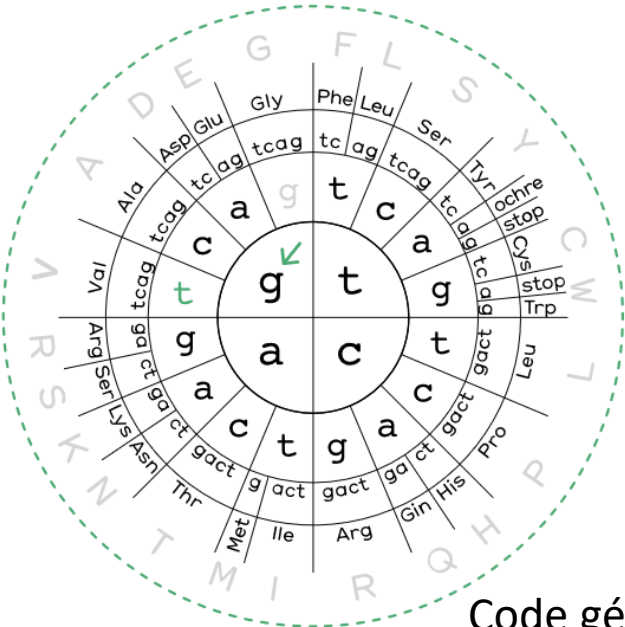
Dernières nouvelles... Le variant anglais

Certaines mutations dans la séquence du génome peuvent modifier la séquence en acide aminé de la protéine correspondante.
Exemple avec le gène codant pour la protéine Spike:

LR991698.2
NC_045512.2

23023 ATCATATGGTTTCCAACCCACTTATGGTGTGTTGGTTACCAACCATAACAGAGTAGTAGTACT 23082
23041 ATCATATGGTTTCCAACCCACTAATGGTGTGTTGGTTACCAACCATAACAGAGTAGTAGTACT 23100
*****:*****

Génome 'UK' (LR991698.2)	tat	ggt	ggt
Protéine Spike UK	Y	G	V
Génome référence (NC_045512.2)	aat	ggt	ggt
Protéine Spike référence	N	G	V
	... 501	502	503 ...



La protéine Spike est composée de 1273 acides aminés.
Cette mutation (ou variant) est en position 501: elle est appelée **N501Y**
...et le virus anglais est aussi appelé B.1.1.7 (**501Y.V1**)

Code génétique

Dernières nouvelles...

Le variant anglais

Comparaison des séquences en acide aminé de la [protéine Spike de référence](#) et de la protéine Spike du virus anglais (B.1.1.7) ([Accès aux séquences](#))

Les différences (changements d'acide aminé / mutations / variants) apparaissent en clair dans cet alignement.

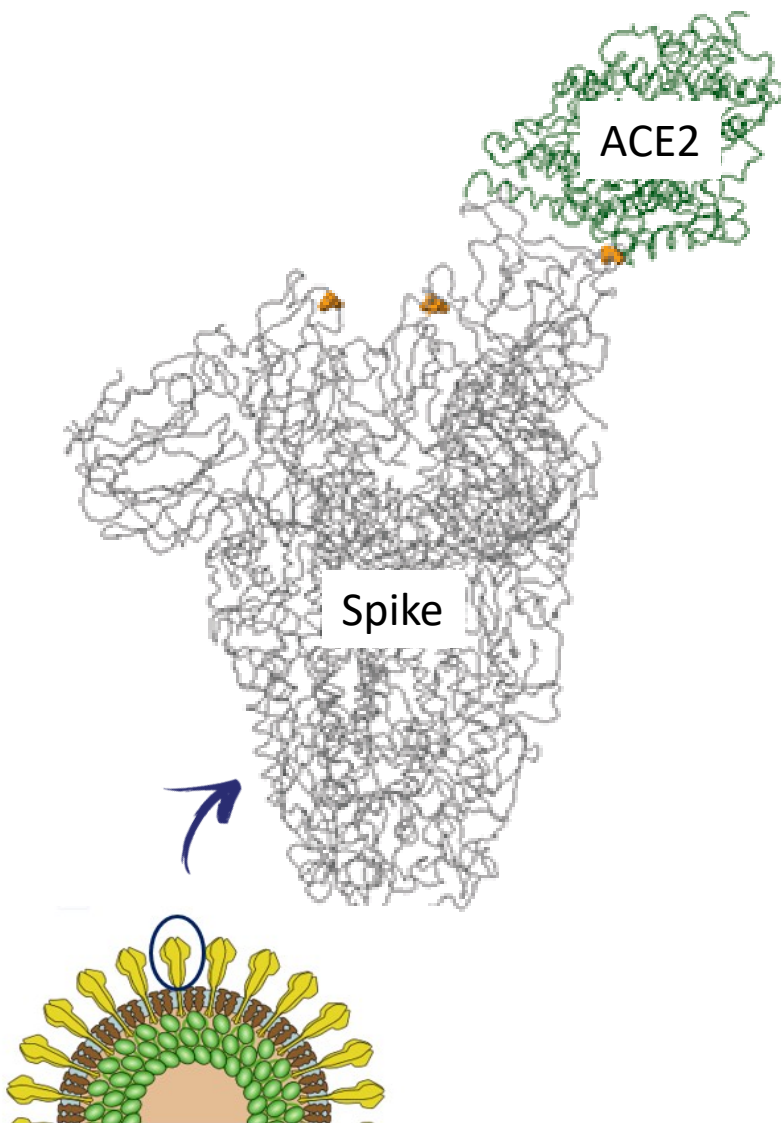
Strain	Accession	Length	MD5	Strain	Accession	Length	MD5
P0DTC2 SPIKE SARS2 Spike_UK_B.1.1.7	1	60	601	601	601	60	601
	1	60	601	601	601	60	601
P0DTC2 SPIKE SARS2 Spike_UK_B.1.1.7	61	120	661	661	661	120	720
	61	120	661	661	661	120	720
P0DTC2 SPIKE SARS2 Spike_UK_B.1.1.7	121	180	721	721	721	180	780
	121	180	721	721	721	180	780
P0DTC2 SPIKE SARS2 Spike_UK_B.1.1.7	181	240	781	781	781	240	840
	181	240	781	781	781	240	840
P0DTC2 SPIKE SARS2 Spike_UK_B.1.1.7	241	300	841	841	841	300	900
	241	300	841	841	841	300	900
P0DTC2 SPIKE SARS2 Spike_UK_B.1.1.7	301	360	901	901	901	360	960
	301	360	901	901	901	360	960
P0DTC2 SPIKE SARS2 Spike_UK_B.1.1.7	361	420	961	961	961	420	1020
	361	420	961	961	961	420	1020
P0DTC2 SPIKE SARS2 Spike_UK_B.1.1.7	421	480	1021	1021	1021	480	1080
	421	480	1021	1021	1021	480	1080
P0DTC2 SPIKE SARS2 Spike_UK_B.1.1.7	481	540	1081	1081	1081	540	1140
	481	540	1081	1081	1081	540	1140
P0DTC2 SPIKE SARS2 Spike_UK_B.1.1.7	541	600	1141	1141	1141	600	1200
	541	600	1141	1141	1141	600	1200
			1201	1201	1201		1260
			1201	1201	1201		1260
			1261	1261	1261		1273
			1261	1261	1261		1273

align@UniProt

Liste des 'variants Spike UK' selon UniProtKB: **N501Y**, P681H, T716I, S982A, D1118H

La guerre des clones....

La 'variante britannique' du SARS-CoV-2 (appelée B.1.1.7), est un virus parmi d'autres qui s'est développé particulièrement rapidement en Angleterre.



- 17 mutations sont présentes dans le génome du virus B.1.1.7: elles entraînent toutes des changements d'acides aminés dans les protéines du virus, dont 9 mutations dans la protéine Spike.
 - La mutation **N501Y** de la protéine Spike (en jaune sur cette image) pourrait modifier les interactions de la protéine avec les récepteurs présents sur les cellules humaines (ACE2).
- Le PDG de BioNTech, Uğur Şahin, a souligné que la protéine Spike codée par l'ARN messenger du vaccin était composée de 1270 acides aminés: 'seulement' 9 d'entre eux sont différents dans la variante britannique du virus.
- La propagation du variant B.1.1.7 a eu lieu pendant une période propice à la transmission des virus respiratoires (janvier, février). Il faut souligner que plus le virus SARS-CoV-2 circulera, plus il aura la possibilité de 'muter'.

3

Dernières nouvelles...

Le variant Afrique du Sud (501Y.V2; B 1.351)

Comparaison des séquences en acide aminé de la [protéine Spike de référence](#) et de la protéine Spike du virus d'Afrique du Sud (501Y.V2) ([Accès aux séquences](#))

Les différences (changements d'acide aminé / mutations / variants) apparaissent en clair dans cet alignement.

P0DTC2 SPIKE SARS2	1	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS	60	P0DTC2 SPIKE SARS2	601	GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTNRVYSTGSMVFQTRAGCLIGAEHVNSV	660
Spike_South_Africa_50..	1	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS	60	Spike_South_Africa_50..	601	GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTNRVYSTGSMVFQTRAGCLIGAEHVNSV	660
		*****				*****	
P0DTC2 SPIKE SARS2	61	NVTWFHAIHVSNGTGKREDNVPLFPNDGVYFASTEKSNIRGWIFGTTLDSTQSLIV	120	P0DTC2 SPIKE SARS2	661	ECDFIGAGICASYQQTNSPRRARSVASQSIIATMSLG*ENSVAYSNNISIAIPTNTI	720
Spike_South_Africa_50..	61	NVTWFHAIHVSNGTGKREDNVPLFPNDGVYFASTEKSNIRGWIFGTTLDSTQSLIV	120	Spike_South_Africa_50..	661	ECDFIGAGICASYQQTNSPRRARSVASQSIIATMSLG*ENSVAYSNNISIAIPTNTI	720
		*****				*****	
P0DTC2 SPIKE SARS2	121	NNAINVVIKVCPEQPCNDPFLGVYHKNKSKMESEFRVYSSANNCTFEVYSQPFIMDL	180	P0DTC2 SPIKE SARS2	721	SVTTEILPVSMKTSVDCMTMVICGDSCTCSNLLQYGSFCTQLNRALTGIAVEQDRNTQ	780
Spike_South_Africa_50..	121	NNAINVVIKVCPEQPCNDPFLGVYHKNKSKMESEFRVYSSANNCTFEVYSQPFIMDL	180	Spike_South_Africa_50..	721	SVTTEILPVSMKTSVDCMTMVICGDSCTCSNLLQYGSFCTQLNRALTGIAVEQDRNTQ	780
		*****				*****	
P0DTC2 SPIKE SARS2	181	GKQGNFKNLREFVFNIDGYFKIYSKHTPINLVRDLQGFSALEPLVDLPIGINITRFQT	240	P0DTC2 SPIKE SARS2	781	VFAQTEQIYKTPPIKDFGGNFSSQILPDPSKPSKSFIEDLLENKVTLDAGFIKQYGD	840
Spike_South_Africa_50..	181	GKQGNFKNLREFVFNIDGYFKIYSKHTPINLVRDLQGFSALEPLVDLPIGINITRFQT	240	Spike_South_Africa_50..	781	VFAQTEQIYKTPPIKDFGGNFSSQILPDPSKPSKSFIEDLLENKVTLDAGFIKQYGD	840
		*****				*****	
P0DTC2 SPIKE SARS2	241	LLALHRSYLTFGDSSSGWTAGAAAYVGYLQPRIFLLKYNENGTITDAVDCALDPLSETF	300	P0DTC2 SPIKE SARS2	841	LGDTAARDLCAQKFNGLTVLPPLTDEMIAQTTSALLAGTITSGWTFGAGALQIPFAM	900
Spike_South_Africa_50..	241	LLALHRSYLTFGDSSSGWTAGAAAYVGYLQPRIFLLKYNENGTITDAVDCALDPLSETF	300	Spike_South_Africa_50..	841	LGDTAARDLCAQKFNGLTVLPPLTDEMIAQTTSALLAGTITSGWTFGAGALQIPFAM	900
		*****				*****	
P0DTC2 SPIKE SARS2	301	CTLKSTVEKGIYQTSNFRVQPTESIVRFPNITNLCPGEVFNATRFASVYAMNKRISN	360	P0DTC2 SPIKE SARS2	901	QMAFRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALN	960
Spike_South_Africa_50..	301	CTLKSTVEKGIYQTSNFRVQPTESIVRFPNITNLCPGEVFNATRFASVYAMNKRISN	360	Spike_South_Africa_50..	901	QMAFRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALN	960
		*****				*****	
P0DTC2 SPIKE SARS2	361	CVADYSVLNYSASFSTFKCYGVSPTKLNLCFTNVYADSFVIRGDEVQIAPGQTGNIAD	420	P0DTC2 SPIKE SARS2	961	TLVKQLSSNFGAISSVINDILSRLDKVEAEVQIDRLITGRLOSLQTYVTQQLIRAAEIRA	1020
Spike_South_Africa_50..	361	CVADYSVLNYSASFSTFKCYGVSPTKLNLCFTNVYADSFVIRGDEVQIAPGQTGNIAD	420	Spike_South_Africa_50..	961	TLVKQLSSNFGAISSVINDILSRLDKVEAEVQIDRLITGRLOSLQTYVTQQLIRAAEIRA	1020
		*****				*****	
P0DTC2 SPIKE SARS2	421	YNYKLPDDFTCCVIAMNNSNLSKVGNNYLYRLFRKSNLKPFERDISTEIIYQAGSTPC	480	P0DTC2 SPIKE SARS2	1021	SANLAATRMSECVLGQSKRVDFCGRGYHLSFPQSAPHGVVFLHVTYVPAQEKNTTAPA	1080
Spike_South_Africa_50..	421	YNYKLPDDFTCCVIAMNNSNLSKVGNNYLYRLFRKSNLKPFERDISTEIIYQAGSTPC	480	Spike_South_Africa_50..	1021	SANLAATRMSECVLGQSKRVDFCGRGYHLSFPQSAPHGVVFLHVTYVPAQEKNTTAPA	1080
		*****				*****	
P0DTC2 SPIKE SARS2	481	NGVSGFNHFLQSYVQPTNGVQOPYRVVVLSEFLLHAPAVCGPKKSTNLVKNKCVN	540	P0DTC2 SPIKE SARS2	1081	ICHGDKAHFPREGVFSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVIGIVNNIVYDS	1140
Spike_South_Africa_50..	481	NGVSGFNHFLQSYVQPTNGVQOPYRVVVLSEFLLHAPAVCGPKKSTNLVKNKCVN	540	Spike_South_Africa_50..	1081	ICHGDKAHFPREGVFSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVIGIVNNIVYDS	1140
		*****				*****	
P0DTC2 SPIKE SARS2	541	FNENGLTGTGVLTSNKRFLFPQQFGRDIDTDAVRDPQTLEILDITPCSGGGVSVITF	600	P0DTC2 SPIKE SARS2	1141	LQPELDSFKELDKYPKNHTSPDVLDGDISGINASVNIQREIDRLNEVAKNINESLID	1200
Spike_South_Africa_50..	541	FNENGLTGTGVLTSNKRFLFPQQFGRDIDTDAVRDPQTLEILDITPCSGGGVSVITF	600	Spike_South_Africa_50..	1141	LQPELDSFKELDKYPKNHTSPDVLDGDISGINASVNIQREIDRLNEVAKNINESLID	1200
		*****				*****	
P0DTC2 SPIKE SARS2	1201	QELGKYEQYIKWPIWLGFIAGLIAIVMTIIMLCMTSCCCLGKCCSCGSCCKFDEDD	1260	P0DTC2 SPIKE SARS2	1201	QELGKYEQYIKWPIWLGFIAGLIAIVMTIIMLCMTSCCCLGKCCSCGSCCKFDEDD	1260
Spike_South_Africa_50..	1201	QELGKYEQYIKWPIWLGFIAGLIAIVMTIIMLCMTSCCCLGKCCSCGSCCKFDEDD	1260	Spike_South_Africa_50..	1201	QELGKYEQYIKWPIWLGFIAGLIAIVMTIIMLCMTSCCCLGKCCSCGSCCKFDEDD	1260
		*****				*****	
P0DTC2 SPIKE SARS2	1261	SEPVLRGKVLHWT	1273	P0DTC2 SPIKE SARS2	1261	SEPVLRGKVLHWT	1273
Spike_South_Africa_50..	1261	SEPVLRGKVLHWT	1273	Spike_South_Africa_50..	1261	SEPVLRGKVLHWT	1273
		*****				*****	

[align@UniProt](#)

Liste des 'variants Spike du virus d'Afrique du Sud' selon [UniProtKB](#): L18F, D80A, D215G, R246I, K417N, **E484K**, **N501Y**, A570D, A701V

Publication : <https://www.medrxiv.org/content/10.1101/2020.12.21.20248640v1>

"Eight lineage-defining mutations in the spike protein, including three at important residues in the receptor-binding domain (K417N, E484K and N501Y) that may have functional significance."

Génome: <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA694014>

3

Dernières nouvelles...

Le variant Afrique du Sud (501Y.V2; B 1.351)

Comparaison des séquences en acide aminé de la [protéine Spike de référence](#) et de la protéine Spike du virus d'Afrique du Sud (501Y.V2) ([Accès aux séquences](#))

Les différences (changements d'acide aminé / mutations / variants) apparaissent en clair dans cet alignement.

P0DTC2 SPIKE SARS2	1	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS	60	P0DTC2 SPIKE SARS2	601	GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTNRVYSTGNSVFPQTRAGCLIGAEHVNSY	660
Spike_South_Africa_50..	1	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS	60	Spike_South_Africa_50..	601	GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTNRVYSTGNSVFPQTRAGCLIGAEHVNSY	660
		*****				*****	
P0DTC2 SPIKE SARS2	61	NVTWFHAIHVSGTNGTKREFNPVLPFNDGVYFASTEKSNIRGWIFGTTLDSTQSLIV	120	P0DTC2 SPIKE SARS2	661	ECDFIPGAGICASYQTQNSPRRARSVASQSLIAITMSLG*ENSVAYSNNNSIAIPTNFTI	720
Spike_South_Africa_50..	61	NVTWFHAIHVSGTNGTKREFNPVLPFNDGVYFASTEKSNIRGWIFGTTLDSTQSLIV	120	Spike_South_Africa_50..	661	ECDFIPGAGICASYQTQNSPRRARSVASQSLIAITMSLG*ENSVAYSNNNSIAIPTNFTI	720
		*****				*****	
P0DTC2 SPIKE SARS2	121	NNAINVVIKVCPEQPCNDPFLGVYHKNKSKMESEFRVSSANNCTFEVYSQPFIMDL	180	P0DTC2 SPIKE SARS2	721	SVTTEILPVSMKTSVDCMTMVICGDSCTCSNLLQYGSFCTQLNRALTGIAVEQDRNTQE	780
Spike_South_Africa_50..	121	NNAINVVIKVCPEQPCNDPFLGVYHKNKSKMESEFRVSSANNCTFEVYSQPFIMDL	180	Spike_South_Africa_50..	721	SVTTEILPVSMKTSVDCMTMVICGDSCTCSNLLQYGSFCTQLNRALTGIAVEQDRNTQE	780
		*****				*****	
P0DTC2 SPIKE SARS2	181	GKQGNFKNREFVFNKIDGYFKIYSKHTPINLVR*LPQGSFALEPLVDLPIGINITREFQT	240	P0DTC2 SPIKE SARS2	781	VFAQTEQIYKTPPIKDFGGNFSSQILPDPSKPSKSFIEDLLENKVLADAGFIKQYGDG	840
Spike_South_Africa_50..	181	GKQGNFKNREFVFNKIDGYFKIYSKHTPINLVR*LPQGSFALEPLVDLPIGINITREFQT	240	Spike_South_Africa_50..	781	VFAQTEQIYKTPPIKDFGGNFSSQILPDPSKPSKSFIEDLLENKVLADAGFIKQYGDG	840
		*****				*****	
P0DTC2 SPIKE SARS2	241	LLALHRSYLTFGDSSSGWTAGAAAYVGYLQPRITFLKYNENGTITDAVDCALDPLSETK	300	P0DTC2 SPIKE SARS2	841	LGDTAARDLCAQKFNGLTVLPPLTDEMIAQTTSALLAGTITSGWTFGAGALQIPFAM	900
Spike_South_Africa_50..	241	LLALHRSYLTFGDSSSGWTAGAAAYVGYLQPRITFLKYNENGTITDAVDCALDPLSETK	300	Spike_South_Africa_50..	841	LGDTAARDLCAQKFNGLTVLPPLTDEMIAQTTSALLAGTITSGWTFGAGALQIPFAM	900
		*****				*****	
P0DTC2 SPIKE SARS2	301	CTLKSTVEKGIYQTSNFRVQPTESIVRFPNITNLCPGGEVFNATRFASVYAMNKRISN	360	P0DTC2 SPIKE SARS2	961	TLVKQLSSNFGAISSVINDILSRDKVEAEVQIDRLITGRLOSLQTVTCQLIRAAEIRA	1020
Spike_South_Africa_50..	301	CTLKSTVEKGIYQTSNFRVQPTESIVRFPNITNLCPGGEVFNATRFASVYAMNKRISN	360	Spike_South_Africa_50..	961	TLVKQLSSNFGAISSVINDILSRDKVEAEVQIDRLITGRLOSLQTVTCQLIRAAEIRA	1020
		*****				*****	
P0DTC2 SPIKE SARS2	361	CVADYSVLNYSASFSTFKCYGVSPTKLNLCFTNVYADSFVIRGDEVQIAPQGTG*IA	420	P0DTC2 SPIKE SARS2	1021	SANLAATFMSECVLGQSKRVDFCGRGYHLSFPQAPHGVLHVTVYPAQEKNTTAPA	1080
Spike_South_Africa_50..	361	CVADYSVLNYSASFSTFKCYGVSPTKLNLCFTNVYADSFVIRGDEVQIAPQGTG*IA	420	Spike_South_Africa_50..	1021	SANLAATFMSECVLGQSKRVDFCGRGYHLSFPQAPHGVLHVTVYPAQEKNTTAPA	1080
		*****				*****	
P0DTC2 SPIKE SARS2	421	YNYKLDDFTCCVIAMNNSNLDKVGNGVNYLYRLFRKSNLKPFFERDISTEIIYQAGSTPC	480	P0DTC2 SPIKE SARS2	1081	ICHDKRAHFPREGVFSNGTHWFVTQRNFYEPQIITTDNTFVSG	
Spike_South_Africa_50..	421	YNYKLDDFTCCVIAMNNSNLDKVGNGVNYLYRLFRKSNLKPFFERDISTEIIYQAGSTPC	480	Spike_South_Africa_50..	1081	ICHDKRAHFPREGVFSNGTHWFVTQRNFYEPQIITTDNTFVSG	
		*****				*****	
P0DTC2 SPIKE SARS2	481	NGVSGFNFVFLQSYV*EQPTYG*QOPYRVVVLSEFLLHAPAVCGPKKSTNLVKNKCVN	540	P0DTC2 SPIKE SARS2	1141	LQPELDSFKEELDKYFNKHTSPDVLGDISGINASVNIQREII	
Spike_South_Africa_50..	481	NGVSGFNFVFLQSYV*EQPTYG*QOPYRVVVLSEFLLHAPAVCGPKKSTNLVKNKCVN	540	Spike_South_Africa_50..	1141	LQPELDSFKEELDKYFNKHTSPDVLGDISGINASVNIQREII	
		*****				*****	
P0DTC2 SPIKE SARS2	541	FNENGLTGTGVLTSNKRFLFPQQFGRDIDTDAVRDPQTLEILDITPCSFGGVSVITF	600	P0DTC2 SPIKE SARS2	1201	QELGKYEQYIKNPWYIWLGTAGLIAIVMTIMLCMTSCCSC	
Spike_South_Africa_50..	541	FNENGLTGTGVLTSNKRFLFPQQFGRDIDTDAVRDPQTLEILDITPCSFGGVSVITF	600	Spike_South_Africa_50..	1201	QELGKYEQYIKNPWYIWLGTAGLIAIVMTIMLCMTSCCSC	
		*****				*****	
P0DTC2 SPIKE SARS2				P0DTC2 SPIKE SARS2	1261	SEPVLEKGVKLHWT	
Spike_South_Africa_50..				Spike_South_Africa_50..	1261	SEPVLEKGVKLHWT	

[align@UniProt](#)

Cette mutation semble spécifique au variant sud africain

Liste des 'variants Spike du virus d'Afrique du Sud' selon [UniProtKB](#): L18F, D80A, D215G, R246I, K417N, **E484K**, **N501Y**, A570D, A701V

• Publication : <https://www.medrxiv.org/content/10.1101/2020.12.21.20248640v1>

"Eight lineage-defining mutations in the spike protein, including three at important residues in the receptor-binding domain (K417N, E484K and N501Y) that may have functional significance."

• "On a une mutation dans la souche dite sud-africaine, qui fait que le vaccin d'AstraZeneca ne marche pas, Et c'est pour ça qu'il a été retiré de la distribution en Afrique du Sud", souligne Didier Trono (RTS, mars 2021)

Génome: <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA694014>

Cette mutation est aussi retrouvée dans le variant anglais

Les mutations Spike du variant brésilien P1:

L18F, T20N, P26S, D138Y, R190S, **K417(T ou N)**, **E484K**, **N501Y**, D614G, H655Y, T1027I, V1176F

(Philippe Le Mercier, UniProtKB)

12 nonsynonymous mutations including K417T, E484K, and N501Y in the receptor-binding domain of the spike protein (variant P.1 for Pangolin variant 1).

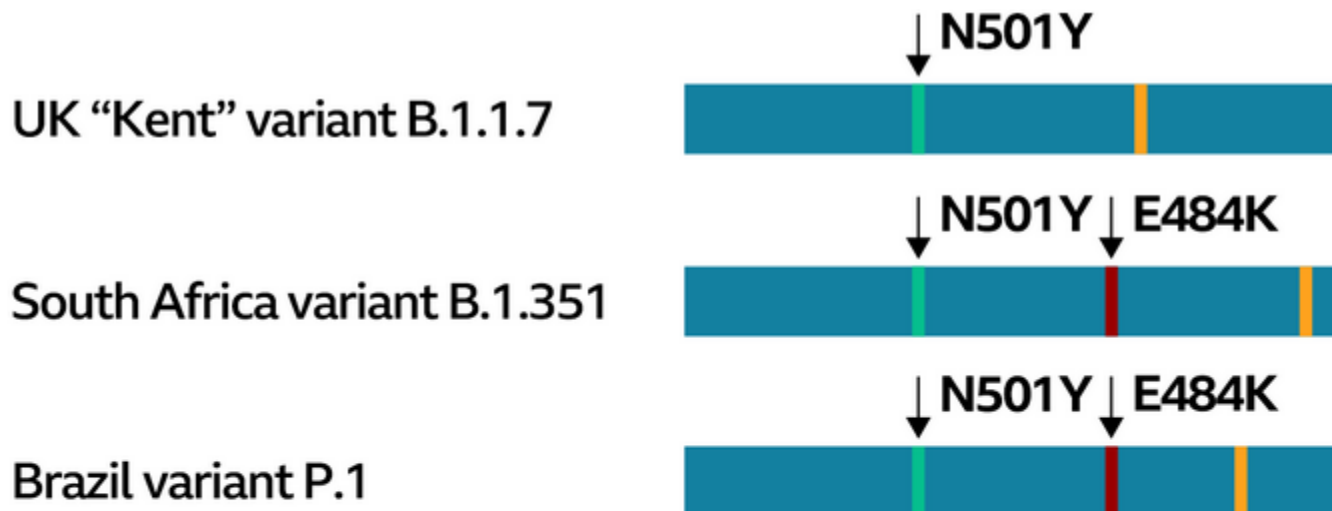
Abstract: Cases of SARS-CoV-2 infection in Manaus, Brazil, resurged in late 2020, despite high levels of previous infection there. Through genome sequencing of viruses sampled in Manaus between November 2020 and January 2021, we identified the emergence and circulation of a novel SARS-CoV-2 variant of concern, lineage P.1, that acquired 17 mutations, including a trio in the spike protein (**K417T, E484K and N501Y**) associated with increased binding to the human ACE2 receptor. Molecular clock analysis shows that P.1 emergence occurred around early November 2020 and was preceded by a period of faster

https://wwwnc.cdc.gov/eid/article/27/4/21-0138_article

4

Dernières nouvelles...

Le variant Brésilien (P1, P.1 ou P.1.1.28.1)



K417T/N, E484K, and N501Y

N501Y mutation seen in UK, South Africa and Brazil variants may help the virus spread more easily.

E484K mutation seen in South Africa, Brazil and some UK variants may affect the antibody response.

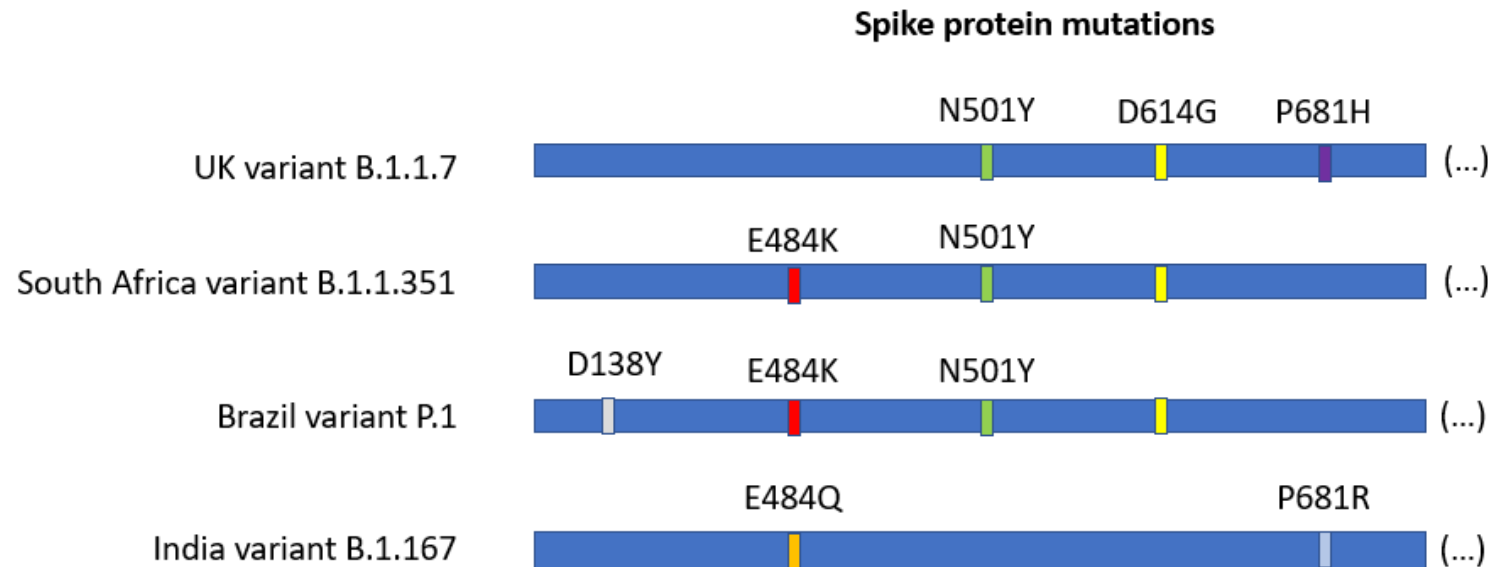
*Ces 2 mutations sont également présentes dans le '**variant américain**'!*

<https://www.bbc.com/news/health-55659820>

5

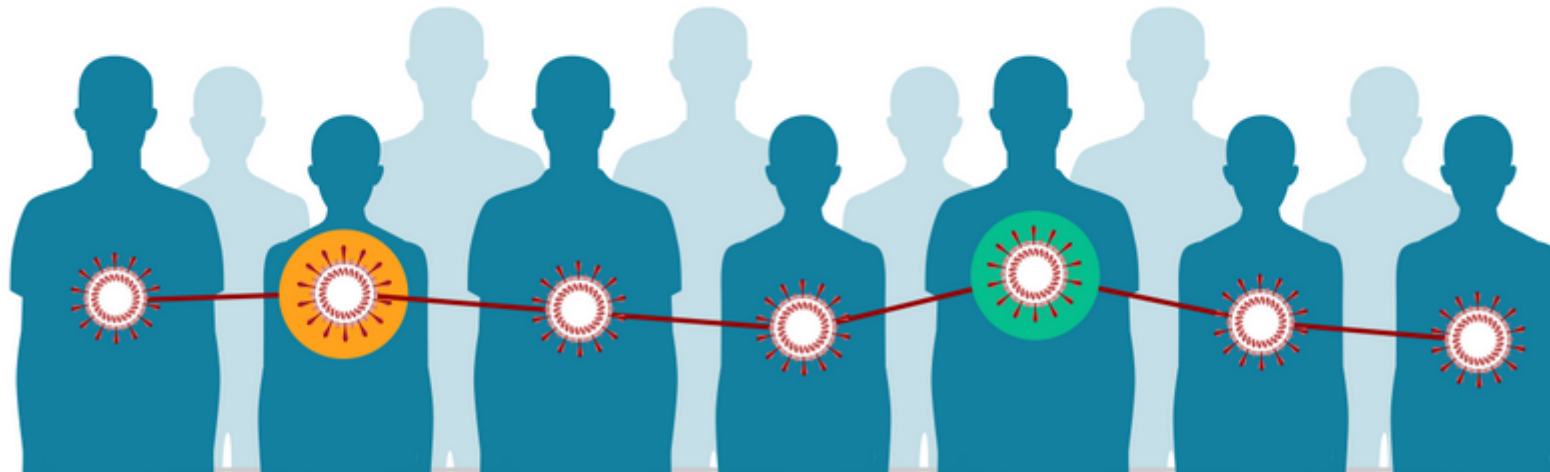
Dernières nouvelles... Le variant indien (B.1.617)

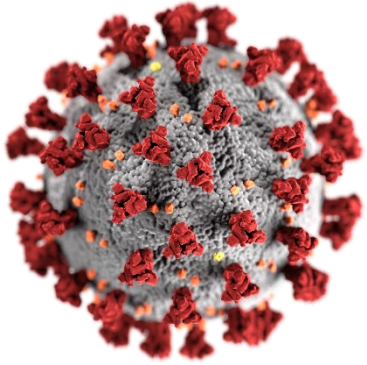
- Plusieurs variants circulent en Inde.
- Récemment, le variant B.1.617 est devenu prédominant.
- B.1.617: 8 mutations dans la protéine Spike, dont L452R, E484Q, P681R (voir le [tableau dans Viralzone](#))
- B.1.617: la mutation N501Y n'est pas présente



1 High numbers of cases increase risk of mutations

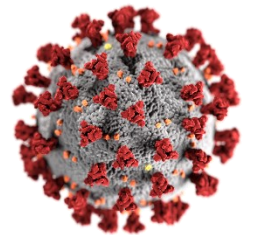
The more a virus spreads, the more chance it has to mutate.
Thousands of small changes have been seen in coronavirus so far - most with little impact.





Un outil bioinformatique pour identifier les ‘variants’

PANGO lineages



Analyse d'un nouveau génome avec 'PANGO lineages'

Choisir une séquence:
[NCBI COVID-19 hub](#)

[illegible]

Exemple: <https://www.ncbi.nlm.nih.gov/nuccore/MW922897.1?report=fasta>

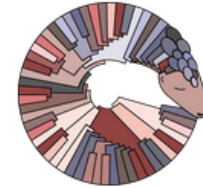
Copier la séquence (format Fasta) dans un fichier.fasta:

http://www.education.expasy.org/cours/Outreach/Coronavirus/Test_Pango1.fasta

Drag and drop fasta file

Select fasta file to upload

Pangolin COVID-19 Lineage Assigner



**Phylogenetic Assignment of Named
Global Outbreak LINEages**

You can upload one or more sequences by dragging and dropping a (multi)fasta file or clicking "Select fasta file to upload" and selecting a (multi)fasta file.

This Web Application assigns lineages to COVID-19 sequences based on the methodology described in this [article](#)

The software to assign lineages based on the algorithm that was developed by [Áine O'Toole](#), [Verity Hill](#), [JT McCrone](#), [Emily Scher](#) and [Andrew Rambaut](#).

The source code can be found [here](#)

Recommended browsers   or 

<https://pangolin.cog-uk.io/>

http://www.education.expasy.org/cours/Outreach/Coronavirus/Test_Pango1.fasta

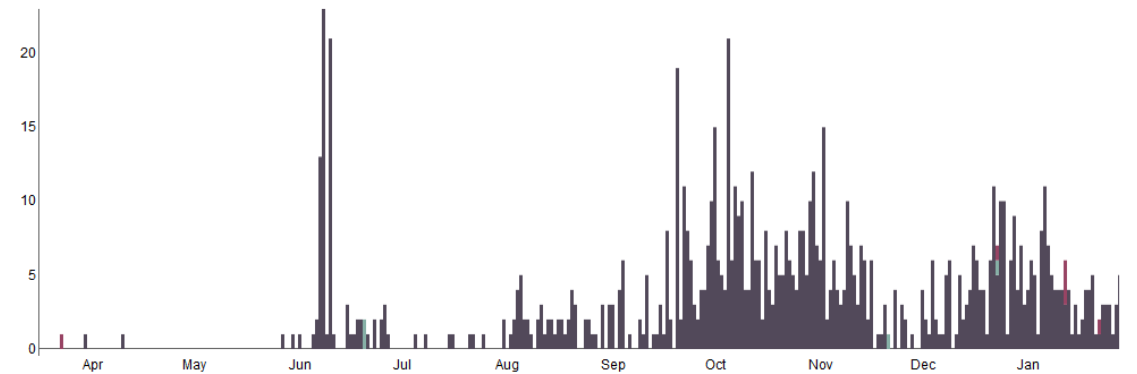
 Reset entries Upload another file Help			
File name	Sequence name	Lineage	Assignment probability
— ANALYSED (Click tick icon for more info) 1 sequence 			
✓ Test_Pango1.fasta	MW922897	B.1.396   	1.0

Lineage B.1.396

[GO TO PARENT LINEAGE: B.1](#)



Distribution dans le monde



Distribution au cours de ces derniers mois

http://www.education.expasy.org/cours/Outreach/Coronavirus/Test_Pango2_UK.fasta



Reset entries

Upload another file

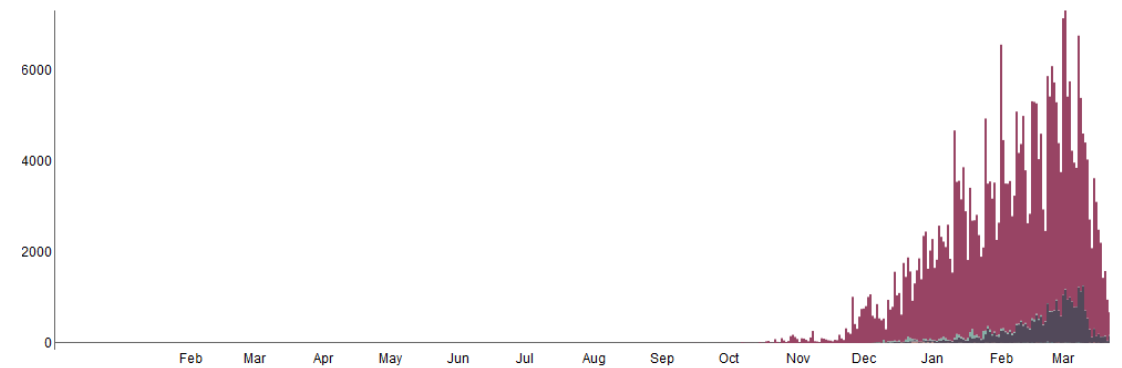
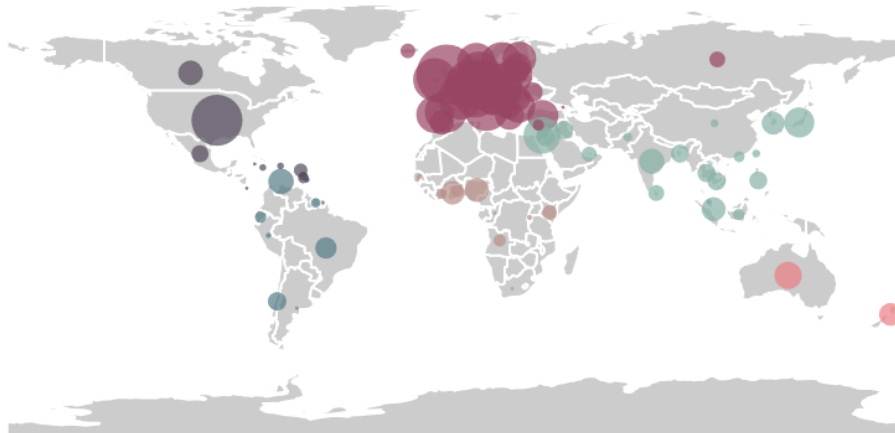
Help

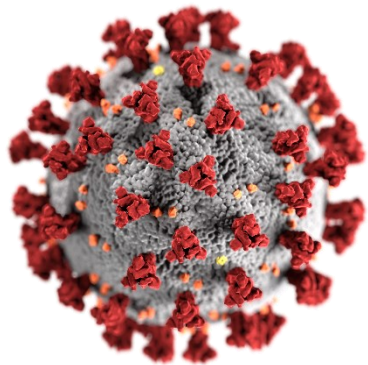
Le variant 'anglais'

File name	Sequence name	Lineage	Assignment probability
— ANALYSED (Click tick icon for more info) 1 sequence 			
✓ Test_pango2_UK.fasta	LR991698.2 Severe acute respiratory syndrome coronavirus 2 genome assembly, complete genome: monopartite	B.1.1.7   	1.0

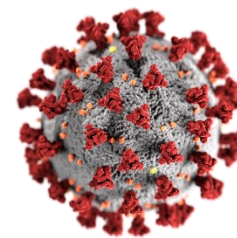
Lineage B.1.1.7

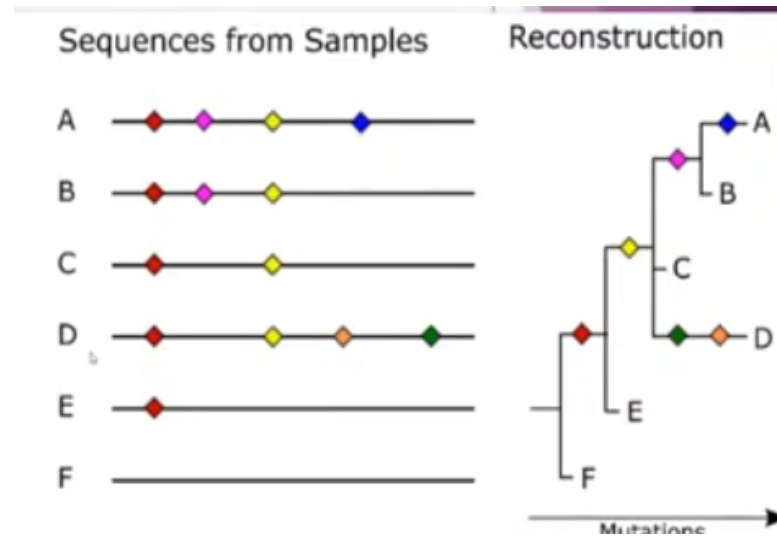
GO TO PARENT LINEAGE: B.1.1





Divers





ADN (référence)

ggt ttc caa ccc act aat ggt gtt ggt tac

ADN (P4)

ggt ttc caa ccc act aat ggt gtt ggt tac

ADN (P1)

ggt ttc caa ccc act **t**at ggt gtt ggt tac

ADN (P3)

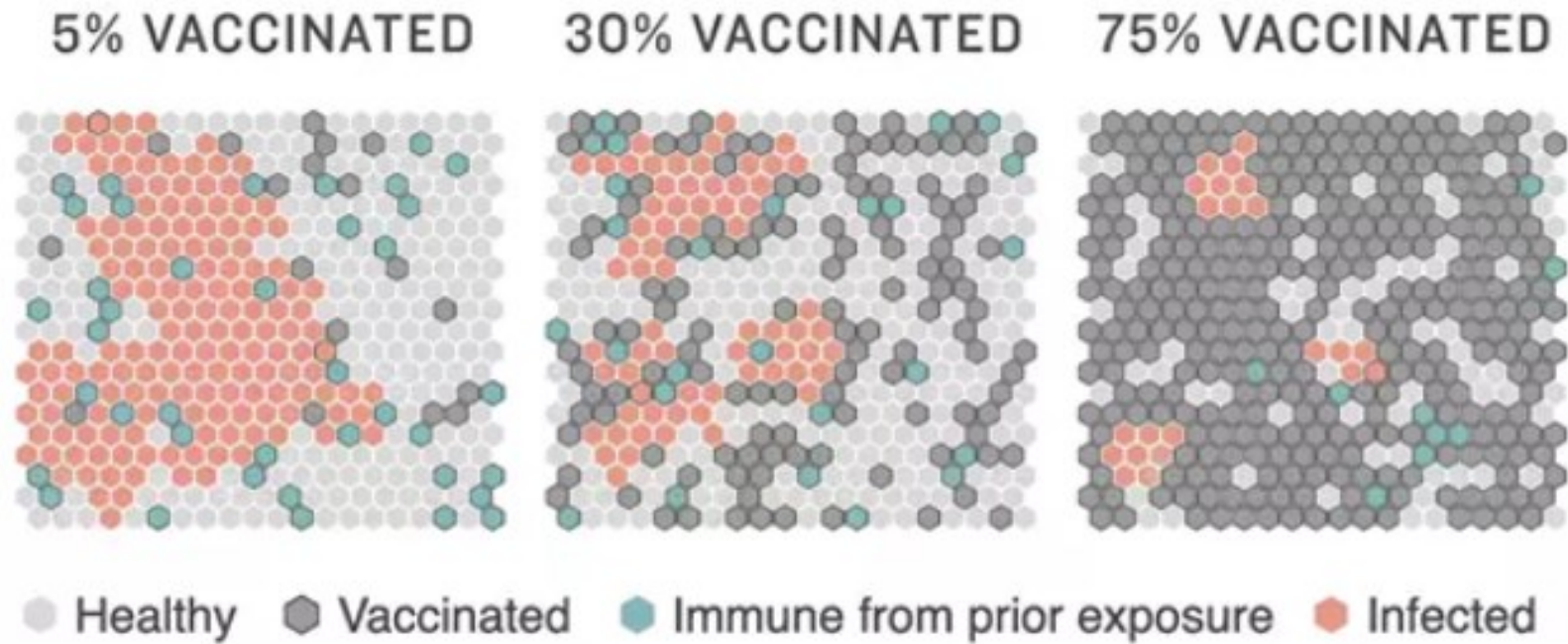
gg**c** ttc caa ccc acc **t**at ggt gtt ggt tac

ADN (P2)

ggt ttc caa ccc act aac ggt gtt ggt tac

ADN (P5)

gg**c** ttc caa ccc act aac ggt gtt ggt tac ?????



https://www.npr.org/sections/health-shots/2021/02/18/967462483/how-herd-immunity-works-and-what-stands-in-its-way?utm_medium=social&utm_term=nprnews&utm_campaign=npr&utm_source=facebook.com&fbclid=IwAR1AHyixs5okXMsBJLNUYffpIPTZjy0OJ8fuKVbYw1CPHzzoqEA-qrojQ&t=1617949276488

- Un site pour découvrir les virus: <https://viropourtous.ch/>
- Une vidéo qui explique l'évolution du virus et les variants: <https://www.youtube.com/watch?v=B8UEZ9cfgz4>
- Un article 'Pourquoi et à quelle vitesse les virus mutant-ils? ([National Geographic](#)): «La compréhension des dynamiques évolutives du SARS-CoV-2 est cruciale pour garantir que les traitements et les vaccins ne soient pas distancés par le virus. »

About the Astra Zenica vaccine: “Some of that [confusion](#) stems from an urgent need to act quickly on the basis of messy, incomplete and capricious real-world data. “

En Inde: La pandémie de COVID-19 a perturbé la prévention et le traitement d'une multitude de maladies (dont la tuberculose). L'année dernière, alors que les pays étaient en état d'urgence, les campagnes de vaccination de masse contre la rougeole, la polio, la méningite et d'autres maladies se sont arrêtées, laissant des millions d'enfants exposés à un risque accru de maladies mortelles évitables. ([source](#))

- Le virus du Covid-19 se réplique 5 à 10 fois plus vite que celui du SRAS de 2003 et produit trois fois plus d'agents pathogènes. Il pourrait se répliquer jusqu'à 100 fois en 48h.
- Malgré cela, il induit une réponse immunitaire et inflammatoire beaucoup plus faible.
- La haute charge virale et cette faible réponse expliqueraient pourquoi autant de personnes très contagieuses propagent le virus sans le savoir. ([source](#))

