

# Recombinaisons génomiques chez les coronavirus

*Des chauves-souris et des hommes ... et des pangolins ?*

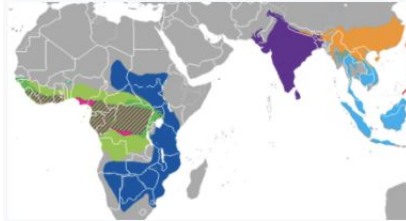
# Pangolins

- Mode de vie
- Aire géographique
- Migration
- Contacts avec les chauves-souris
- Contacts avec l'humain

## Répartition géographique de différentes espèces de pangolin

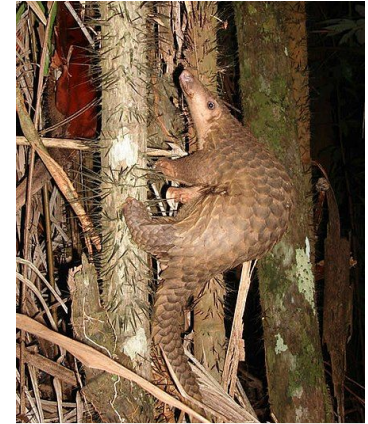
### Genera

- *Manis*
- *Phataginus*
- *Smutsia*



### Species ranges

-  *Manis crassicaudata*
-  *Manis pentadactyla*
-  *Manis javanica*
-  *Manis culionensis*
-  *Phataginus tricuspis*
-  *Phataginus tetradactyla*
-  *Smutsia gigantea*
-  *Smutsia temminckii*



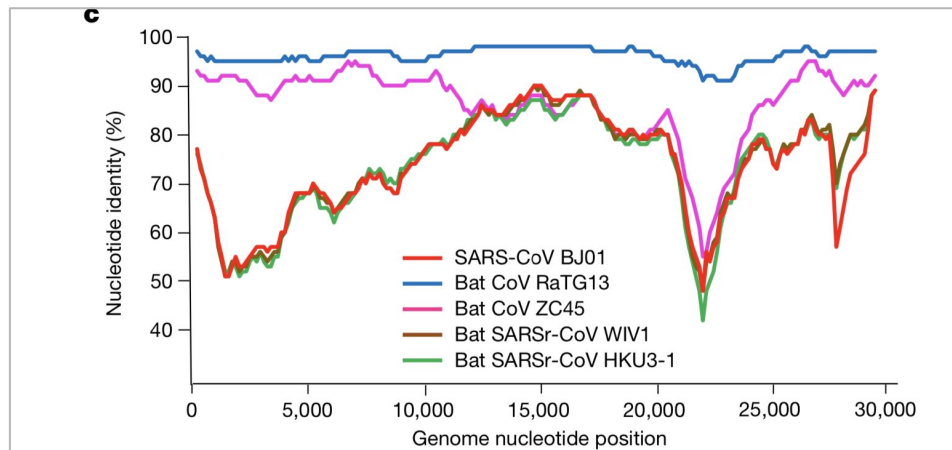
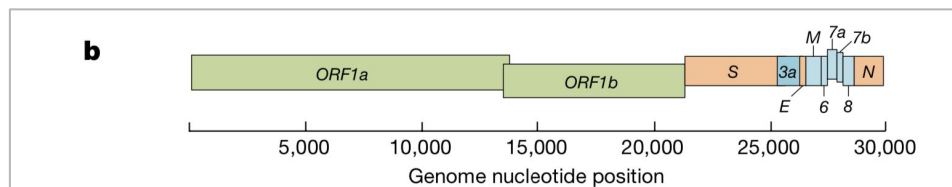
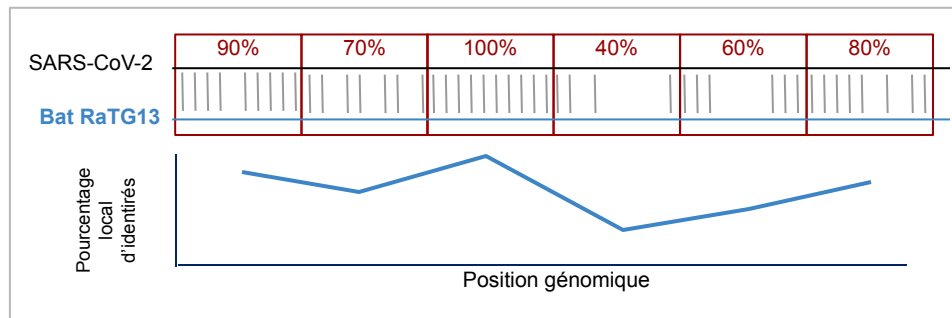
*Manis javanica*  
(Pangolin malais)

### Aire de répartition de *Manis javanica*



# Profils de positions identiques (PPI)

- Haut: principe de calcul du PPI
  - alignement d'une paire de séquences
  - découpage de la séquence en "fenêtres"
  - calcul du pourcentage local d'identité de chaque fenêtre
  - dessin du profil de positions identiques (PPI)
- Milieu : positions des gènes de SARS-CoV-2 sur le génome
- Bas: PPI de quelques génomes de coronavirus sur celui de SARS-CoV-2
- Commentaires dans la diapo suivante



# Une origine probable: la chauve-souris

- 3 février 2020: publication du génome complet de SARS-CoV-2
- Recherche de virus similaires dans les bases de données de séquence
  - Les virus les plus proches sont des virus de chauves-souris (Bat CoV ZC45)
- Dans le même article, les auteurs décrivent un nouveau génome de chauve-souris: **RaTG13**
  - A ce jour la souche virale la plus proche de SARS-CoV-2 connue
- Figures du bas: profil de positions identiques (PPI), expliqué ci-après.

## Article

### A pneumonia outbreak associated with a new coronavirus of probable bat origin

<https://doi.org/10.1038/s41586-020-2012-7>

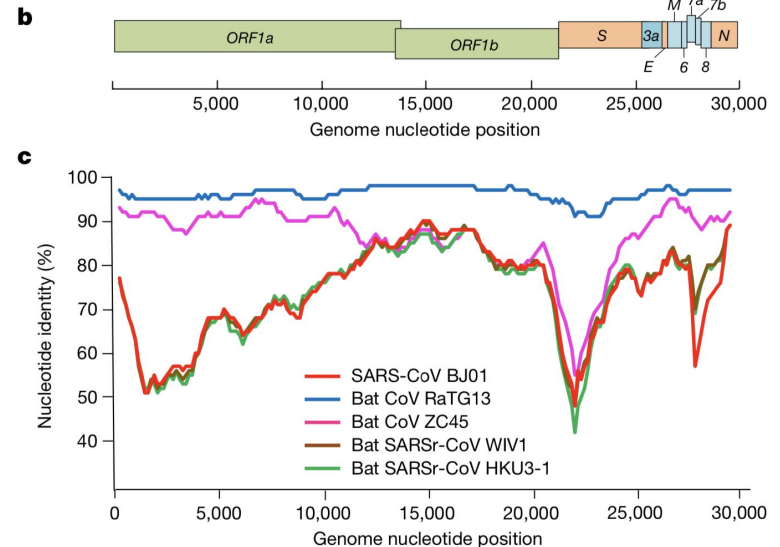
Received: 20 January 2020

Accepted: 29 January 2020

Published online: 3 February 2020

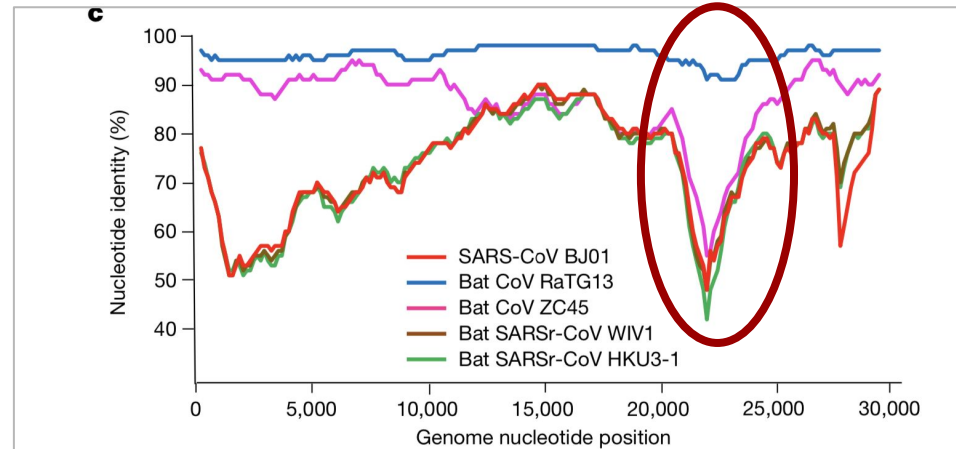
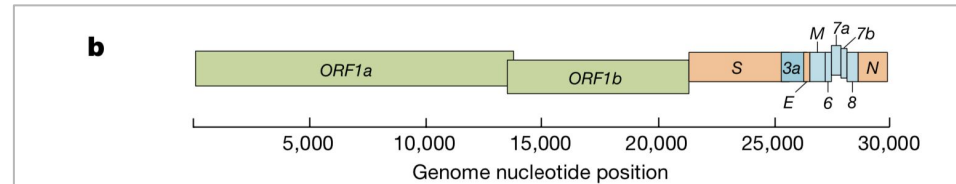
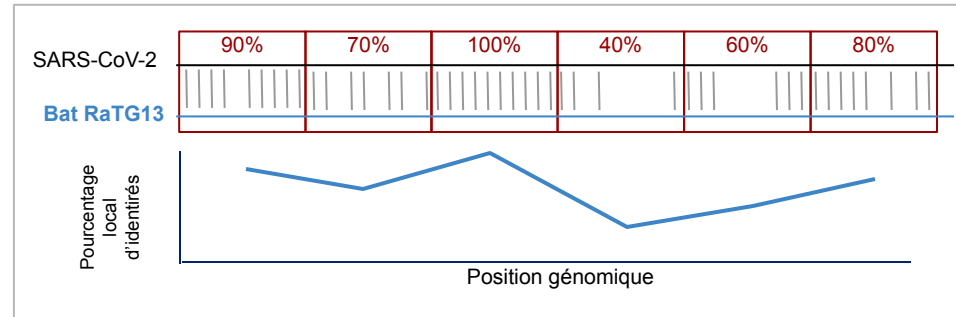
Open access

Peng Zhou<sup>1,5</sup>, Xing-Lou Yang<sup>1,5</sup>, Xian-Guang Wang<sup>2,5</sup>, Ben Hu<sup>1</sup>, Lei Zhang<sup>1</sup>, Wei Zhang<sup>1</sup>, Hao-Rui Si<sup>1,5</sup>, Yan Zhu<sup>1</sup>, Bei Li<sup>1</sup>, Chao-Lin Huang<sup>2</sup>, Hui-Dong Chen<sup>2</sup>, Jing Chen<sup>1,3</sup>, Yun Luo<sup>1,3</sup>, Hua Guo<sup>1,3</sup>, Ren-Di Jiang<sup>1,3</sup>, Mei-Qin Liu<sup>1,3</sup>, Ying Chen<sup>1,3</sup>, Xu-Rui Shen<sup>1,3</sup>, Xi Wang<sup>1,3</sup>, Xiao-Shuang Zheng<sup>1,3</sup>, Kai Zhao<sup>1,3</sup>, Qian-Jiao Chen<sup>1</sup>, Fei Deng<sup>1</sup>, Lin-Lin Liu<sup>4</sup>, Bing Yan<sup>1</sup>, Fa-Xian Zhan<sup>4</sup>, Yan-Yi Wang<sup>1</sup>, Geng-Fu Xiao<sup>1</sup> & Zheng-Li Shi<sup>1,5\*</sup>



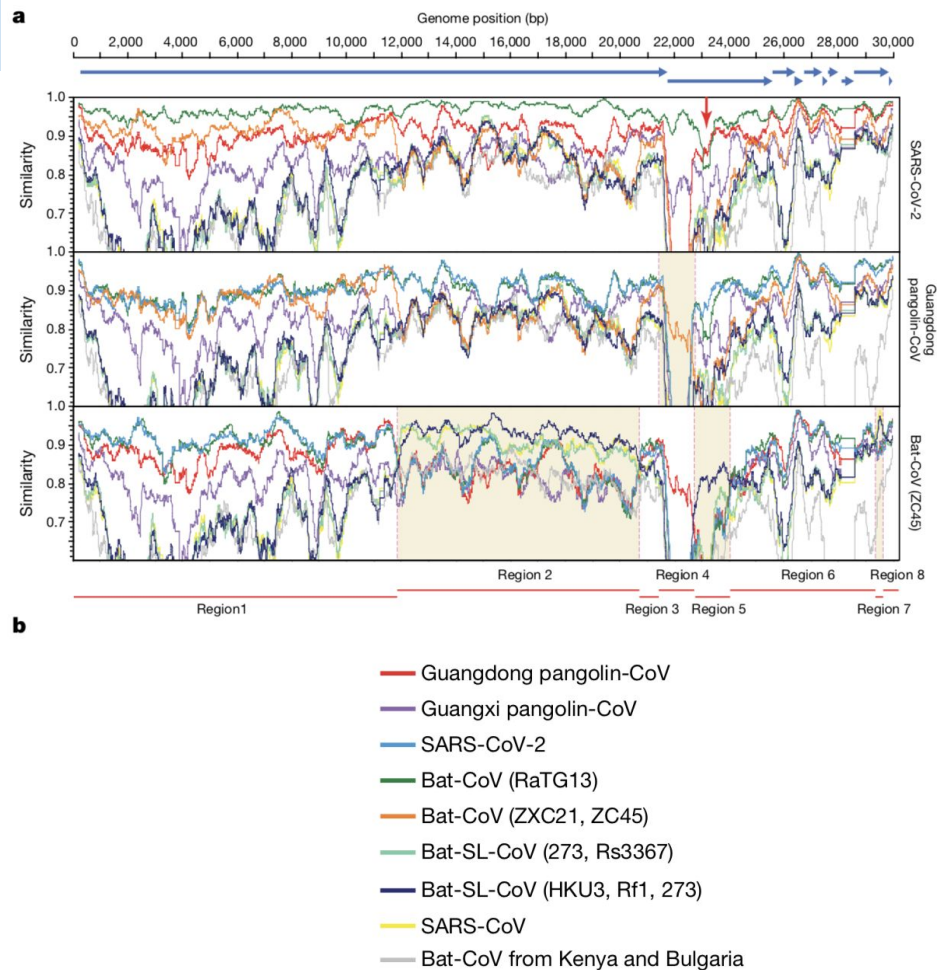
# Profils de positions identiques (PPI) de génomes de coronavirus

- RaTG13 (virus de chauve-souris) est le génome le plus proche de SARS-CoV-2
- On a identifié d'autres virus de chauve-souris relativement proches de SARS-CoV-2 (Cov ZC45)
- Les virus SARS-CoV humains (pandémie 2002-2003) sont moins proches
- Pour chaque espèce, on observe des fluctuations le long du profil de PPI
- Entre 22.000 et 25.000 : chute brutale des PPI
  - Chez RaTG13 : on passe de ~96% à 89%
  - Chez tous les autres coronavirus : effondrement du pourcentage de positions identiques
  - Cette région correspond au domaine de liaison au récepteur (Receptor Binding Domain, RBD), qui est crucial pour permettre au virus d'entrer dans la cellule.



## ... et des pangolins ?

- Lam et collègues comparent le génome de SARS-CoV-2 à des génomes de virus isolés à partir de pangolins.
- Ces virus sont globalement plus éloignés de SARS-CoV-2 que ceux de chauve-souris.
- Cependant, on observe une identité plus élevée dans la région particulière où les PPI des autres coronavirus s'affaissent.
- Ceci suggère la possibilité d'une recombinaison entre des virus de chauve-souris et de pangolin.



# Et si les pangolins n'y étaient pour rien ?

L'hypothèse du pangolin est fortement remise en cause pour plusieurs raisons.

- Les génomes des virus de pangolin les plus proches dont on dispose sont plus éloignés de SARS-CoV-2 que ceux des génomes de chauves-souris.
- La région où la similarité est la plus forte correspondent au gène S, qui code pour la protéine spicule. Or la protéine spicule de ces virus de pangolin n'est pas capable d'adhérer au récepteur de cellules humaines.

franceinfo:

vidéos

radio

jt

magazines

DIRECTV

DIRECT RADIO



## le billet sciences

Anne-Laure Barral

◆ / replay\_radio / Le billet sciences

### Covid-19 : et si le pangolin n'y était pour rien ?

Les chercheurs ont peut-être accusé un peu trop vite le pangolin d'être à l'origine de la pandémie qui a débuté à Wuhan en Chine fin décembre 2019. Ce petit mammifère à écailles est de plus en plus disculpé par des études scientifiques.

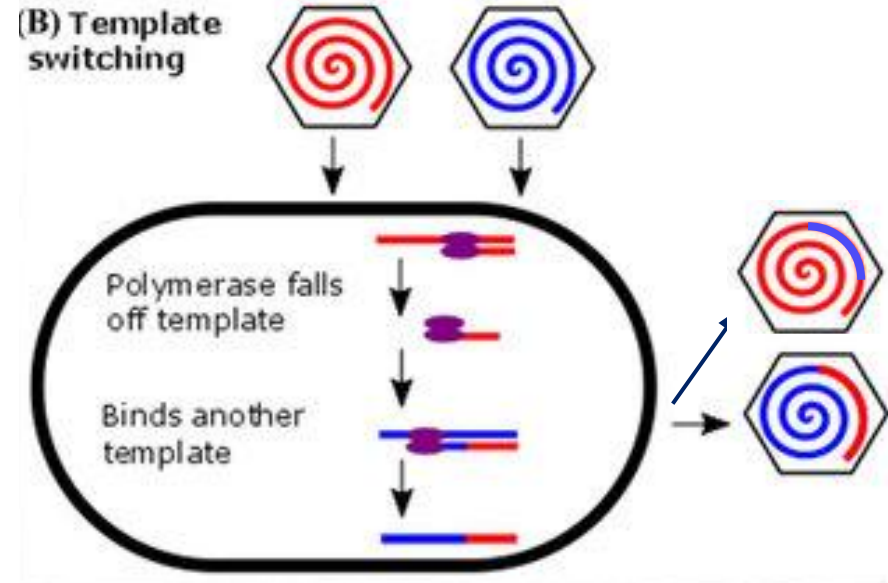




*Détecter les recombinaisons par comparaison de génomes*

# Recombinaisons chez les coronavirus

- Une chauve-souris peut se retrouver infectée par plusieurs coronavirus en même temps.
- Pendant la réplication des coronavirus, il arrive que la polymérase de l'ARN "saute" d'un virus à l'autre.
- Ceci donne naissance à un virus "chimérique", dont le génome est composé de fragments d'origines différentes.
- Ceci complique l'analyse de la phylogénie des virus, car différents fragments génomiques résultent d'histoires évolutives différentes.



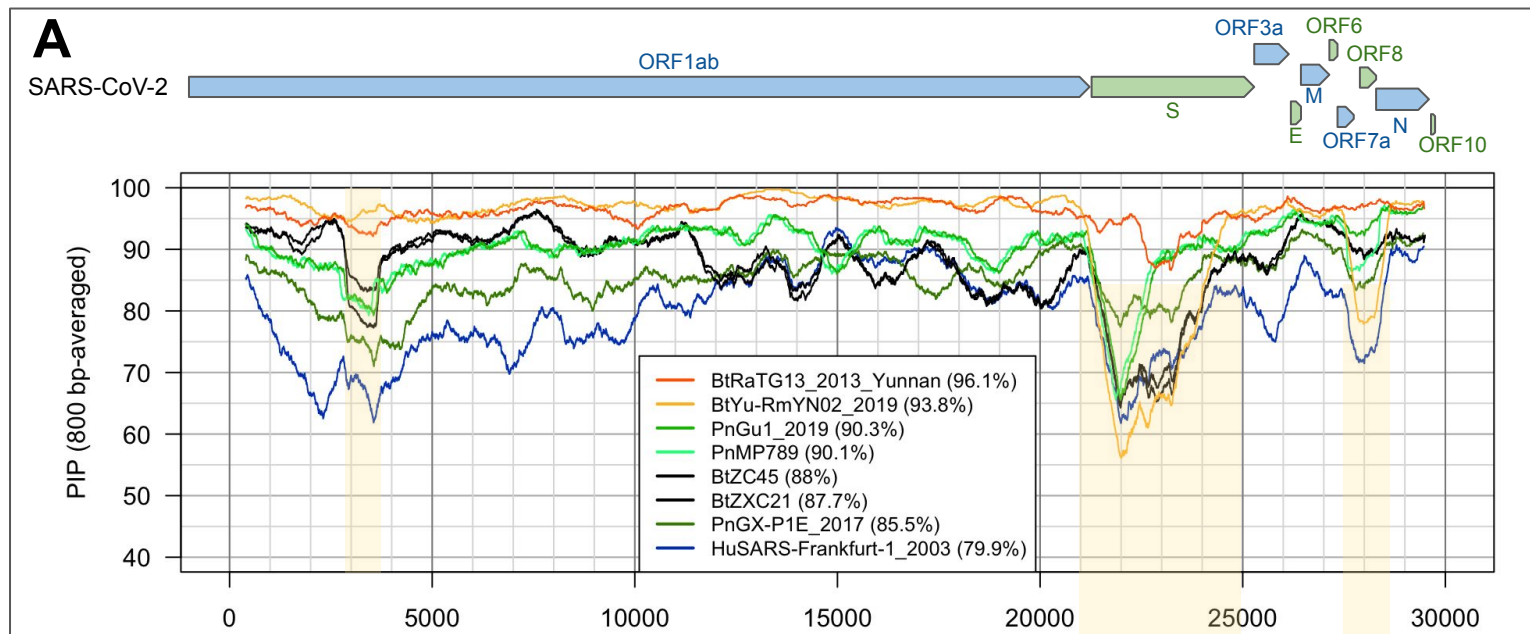
Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus

Ben Hu ,et al Plos path : November 30, 2017

J. Dennehy, Evolutionary ecology of virus emergence: Virus emergence, 2016, Annals of the New York Academy of Science

# Détection des recombinaisons génomiques

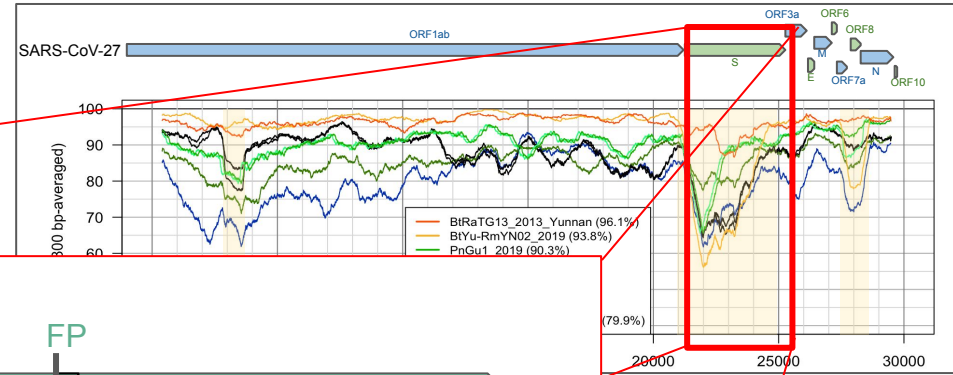
Les chutes brutales d'identité sur les profils PPI (fond jaune) dénotent des régions résultant vraisemblablement de recombinaisons.



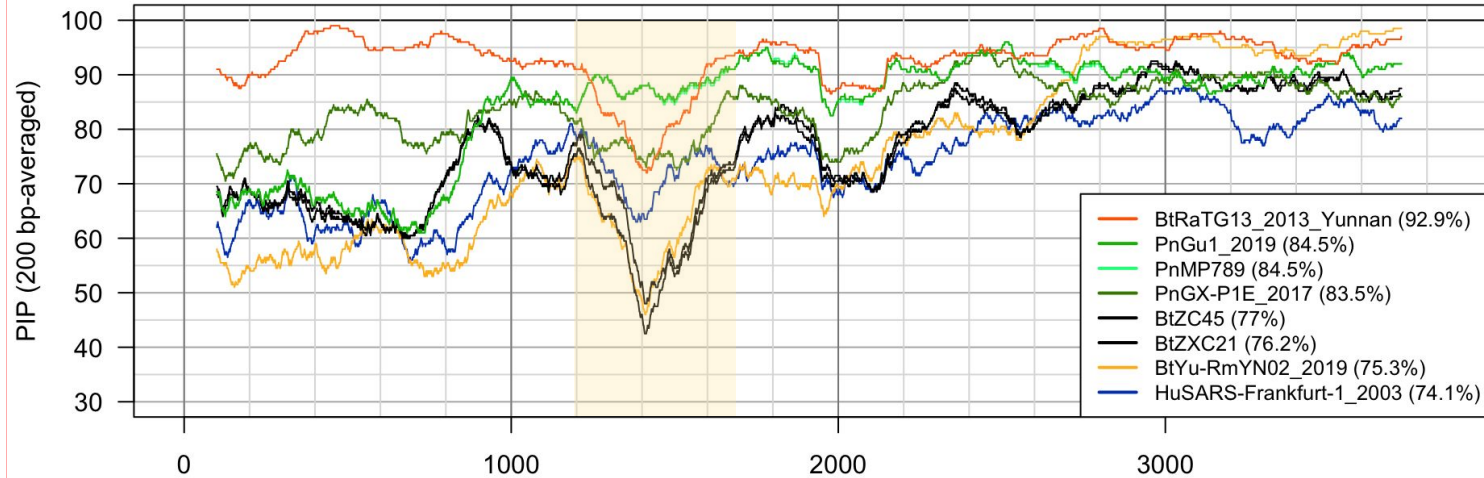
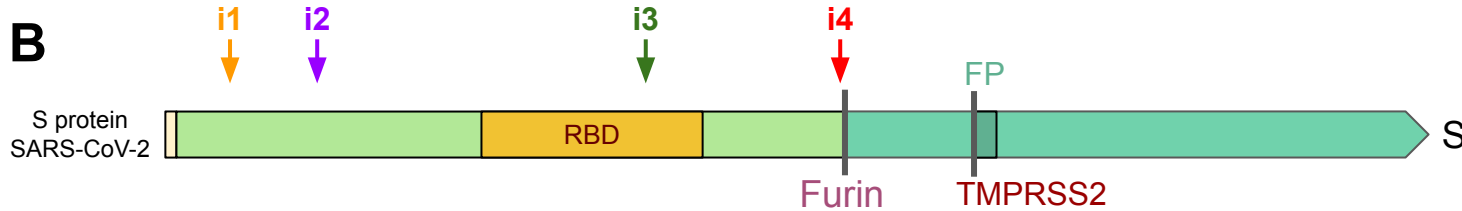
- Sallard, E., Halloy, J., Casane, D., van Helden, J. & Decroly, É. 2020. **Retrouver les origines du SARS-CoV-2 dans les phylogénies de coronavirus.** *Med Sci (Paris)* 36: 783–796.
- English version : Erwan Sallard, José Halloy, Didier Casane, Etienne Decroly, Jacques van Helden. **Tracing the origins of SARS-CoV-2 in coronavirus phylogenies.** [hal-02891455](https://hal.archives-ouvertes.fr/hal-02891455)

# Comparaison entre coronavirus - gène S

Profils de pourcentages de positions identiques (PPI) entre régions génomiques du gène S de différents coronavirus et SARS-CoV-2 (la référence à 100%).



**B**



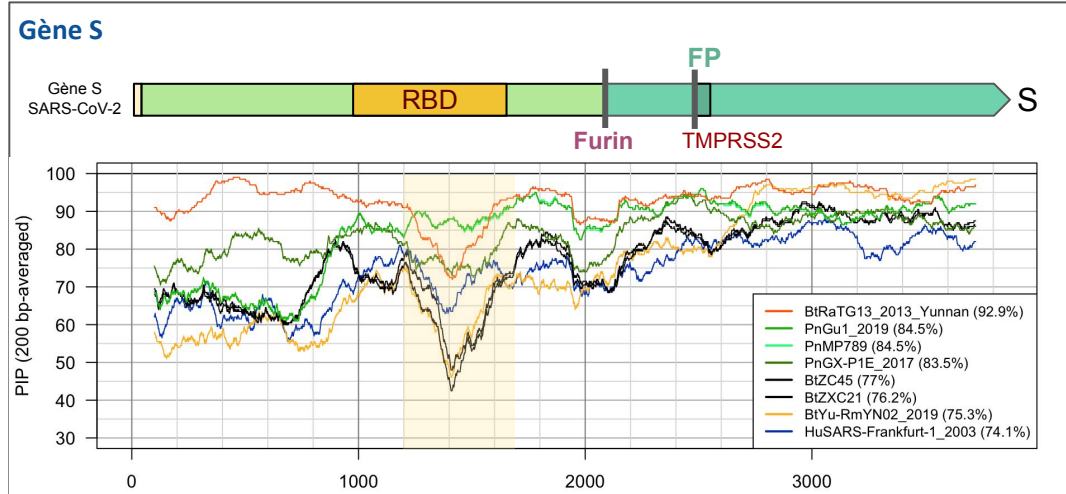
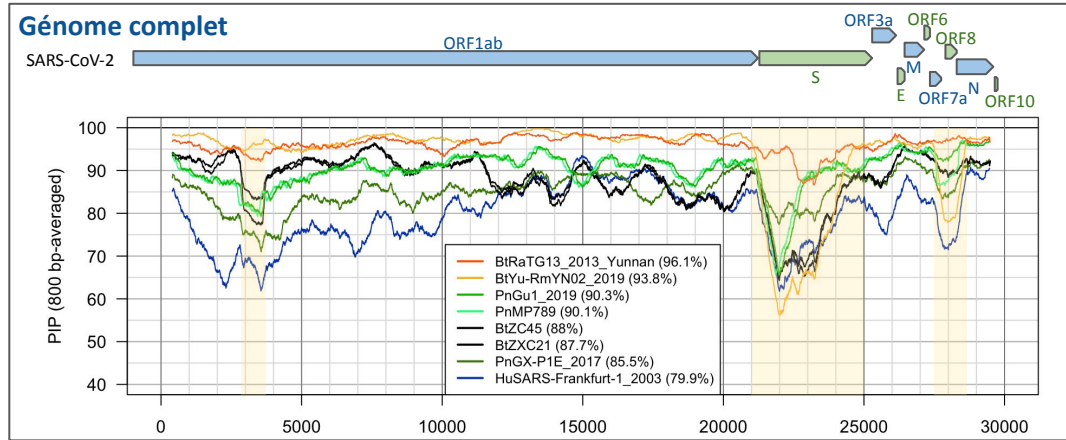
# Recombinaisons génomiques dans les génomes de coronavirus

## Le profil génomique

- régions ayant vraisemblablement fait l'objet de recombinaisons (fond jaune).

## Profil PPI du gène spicule (S)

- **S:** spicule
- **RBD:** receptor binding domain
- Dans la région du RBD forte baisse des PPI
- Le RBD est en évolution rapide, pourquoi ?
  - Immunogène → forte pression sélective en faveur de variations qui permettent d'échapper à l'immunité
  - Spécificité d'espèce → modifications permettent de changer d'hôte



# Un métagénome de virus de chauve-souris (*RmYN02*)

## Current Biology

### A Novel Bat Coronavirus Closely Related to SARS-CoV-2 Contains Natural Insertions at the S1/S2 Cleavage Site of the Spike Protein

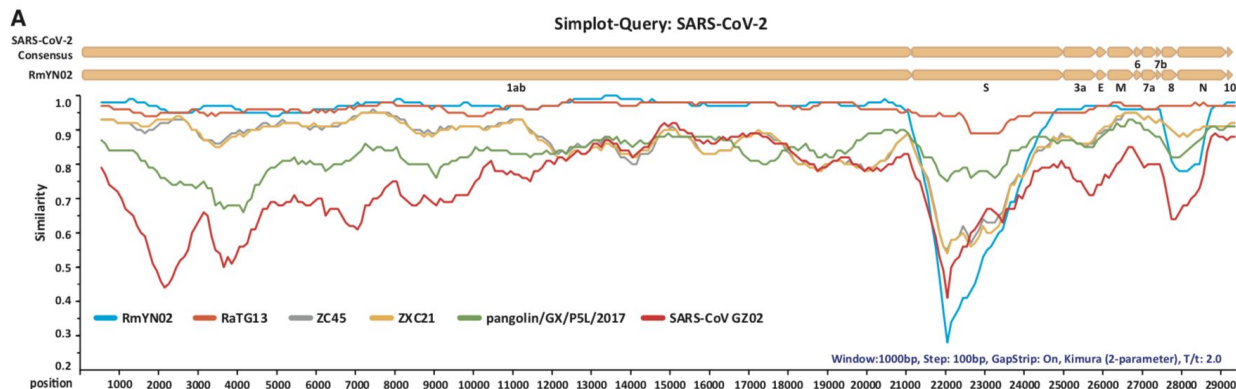
#### Highlights

- Metagenomic analysis identified a novel coronavirus, RmYN02, from *R. malayanus*

#### Authors

Hong Zhou, Xing Chen, Tao Hu, ..., Alice C. Hughes, Yuhai Bi, Weifeng Shi

- Hong Zhou et coll. (2019)
- Séquence métagénomique d'un coronavirus de chauve-souris assemblée à partir de 11 échantillons de chauves-souris de Yunnan
- Génome très proche de SARS-CoV-2 **sauf** dans la région du gène spicule (S).
  - ce virus serait donc un recombinant de RaTG13
  - oui mais** ce métagénome est reconstruit à partir de 11 échantillons différents. La recombinaison pourrait donc être un artéfact
- Les auteurs soulignent aussi la présence d'une insertion similaire au site furine de SARS-CoV-2.
  - Oui mais** ce scénario est peu vraisemblable, car il aurait nécessité quatre événements évolutifs (une insertion et trois délétions).



**H**

Polybasic cleavage site

SARS-CoV-2 Numbering	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693
Consensus SARS-CoV-2	G	A	G	I	C	A	S	Y	Q	T	Q	T	N	S	P	R	R	A	R	S	V	A	S	Q	S	I	I
RmYN02	G	A	G	V	C	A	S	Y	-	-	-	-	N	S	P	-	A	A	R	-	V	G	T	N	S	I	I
RaTG13	G	A	G	I	C	A	S	Y	Q	T	Q	T	N	S	-	-	-	R	S	V	A	S	Q	S	I	I	
ZC45	G	A	G	I	C	A	S	Y	H	T	A	S	I	L	-	-	-	R	S	T	S	Q	K	A	I	V	
ZXC21	G	A	G	I	C	A	S	Y	H	T	A	S	I	L	-	-	-	R	S	T	G	Q	K	A	I	V	
pangolin/MP789/2019	G	A	G	I	C	A	S	Y	Q	T	Q	T	N	S	-	-	-	R	S	V	S	S	X	A	I		
pangolin/GX/P5L/2017	G	A	G	I	C	A	S	Y	H	S	M	S	S	F	-	-	-	R	S	V	N	Q	R	S	I	I	
SARS-CoV GZ02	G	A	G	I	C	A	S	Y	H	T	V	S	L	L	-	-	-	R	S	T	S	Q	K	S	I	V	
RmYN01	G	A	G	I	C	A	S	Y	H	T	A	S	L	L	-	-	-	R	N	T	G	Q	K	S	I	V	

O-linked glycan residues