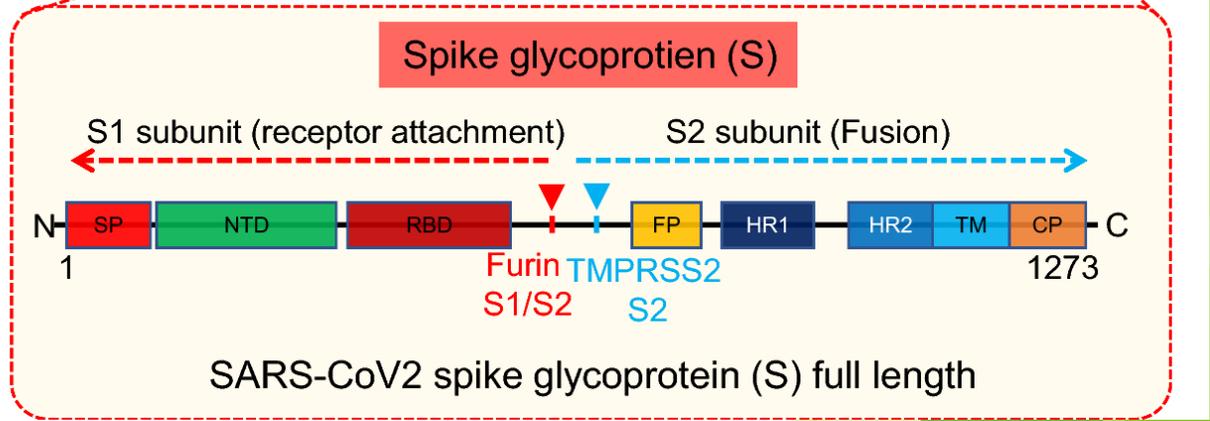
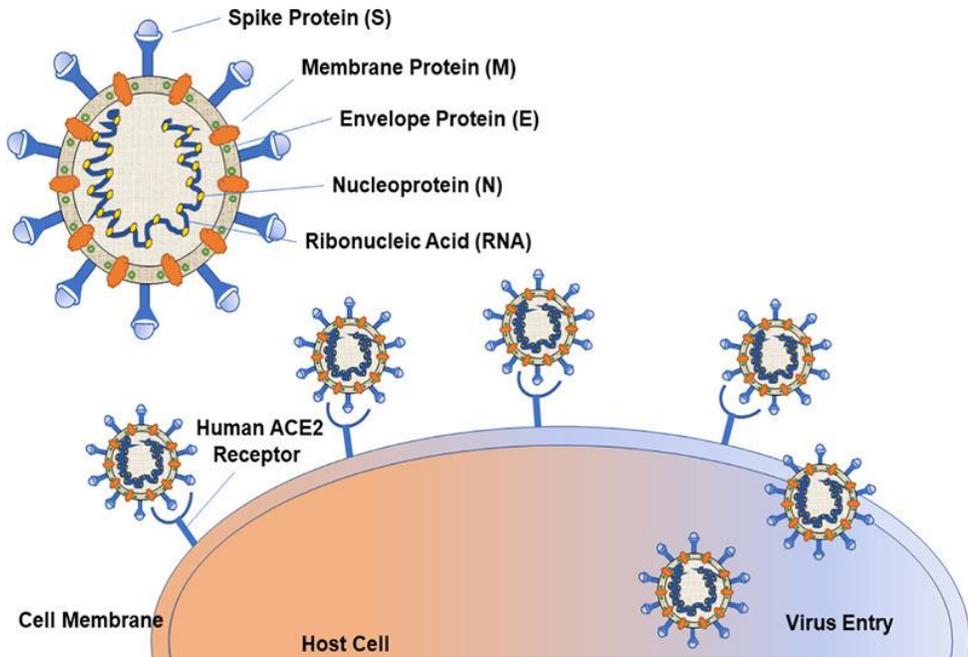
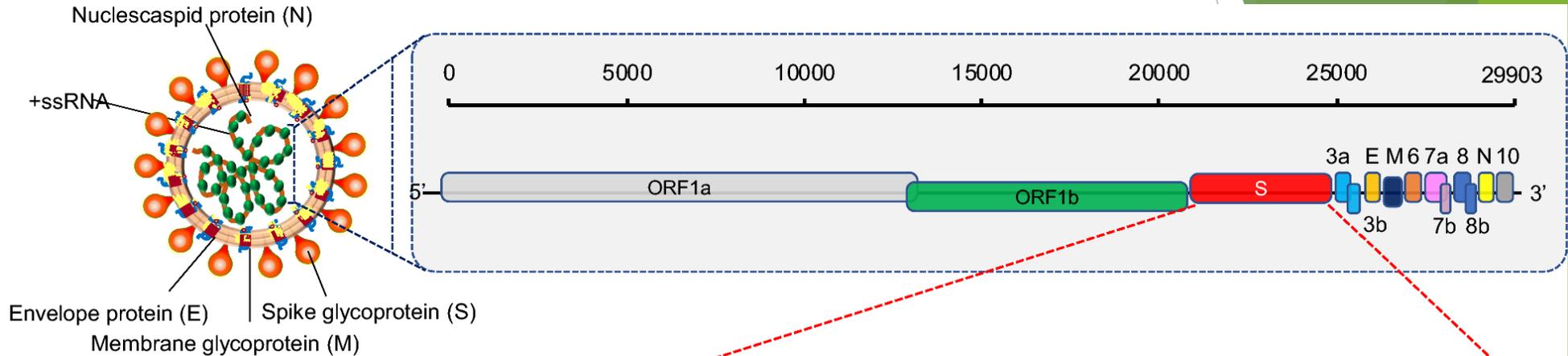


# Analyse de recombinaison et de transferts horizontaux de gènes chez SARS-CoV-2

Stéphane Samson et Vladimir Makarenkov  
Université du Québec à Montréal

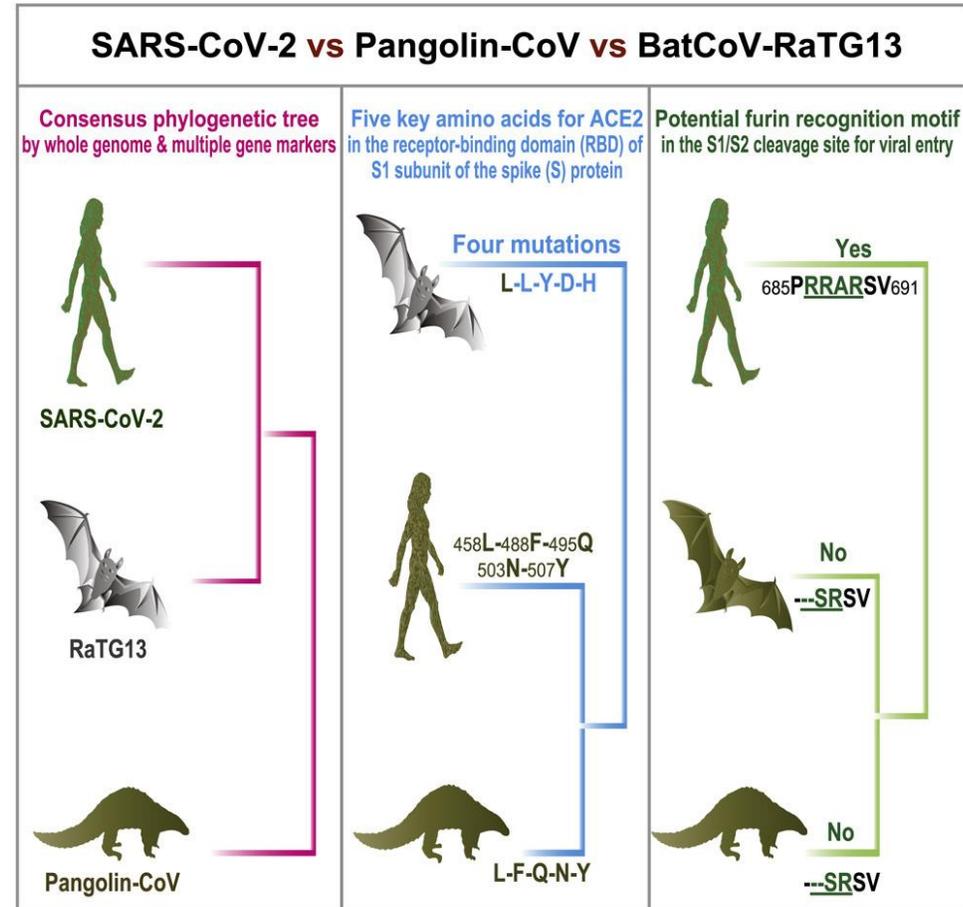
30 janvier 2024

# Génome de SARS-CoV-2

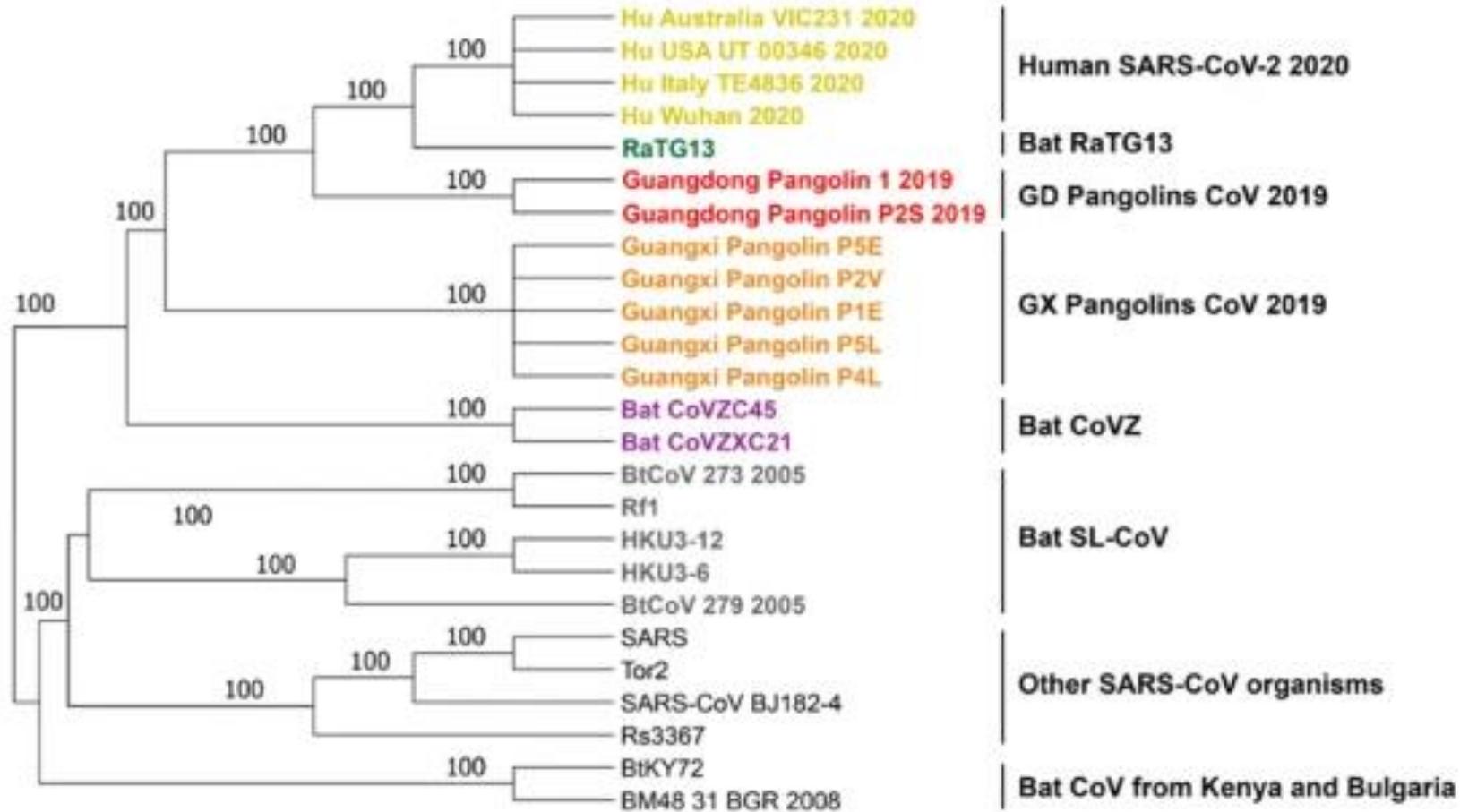


# Dans la littérature

- ▶ Plusieurs études ont suggéré qu'un proche parent du SARS-CoV-2 est le RaTG13 CoV trouvé dans *Rhinolophus affinis*
  - ▶ Réservoir naturel
- ▶ Des analyses récentes du génome du SARS-CoV-2 indiquent des fortes similarités avec certains génomes de pangolins malais
  - ▶ Guangdong vs Guangxi

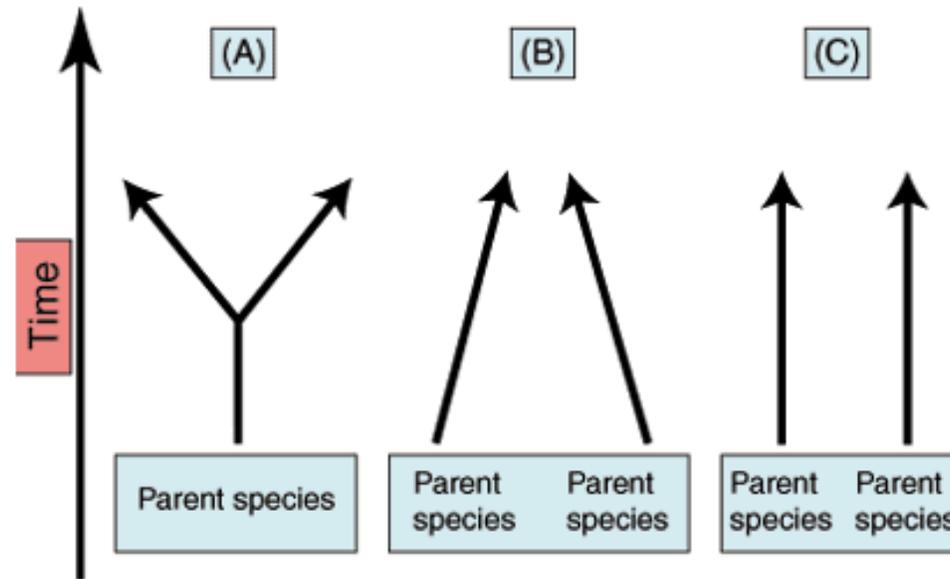


# Arbre phylogénétique d'espèces de SARS-CoV-2



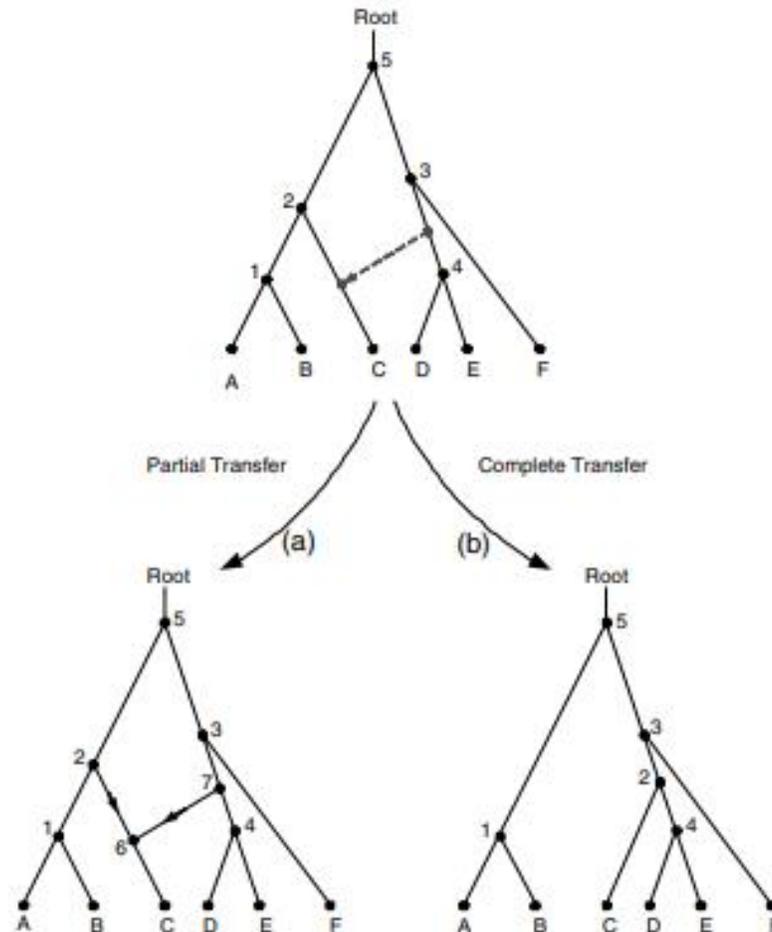
# 3 hypothèses évolutives

- ▶ Évolution parallèle
  - ▶ Pression évolutive commune mène à des traits communs
- ▶ Évolution divergente
  - ▶ GD Pangolin CoV et SARS-CoV-2 aurait des ancêtres communs, d'où leur similarité
- ▶ Recombinaison



# Recombinaison partielle et complète

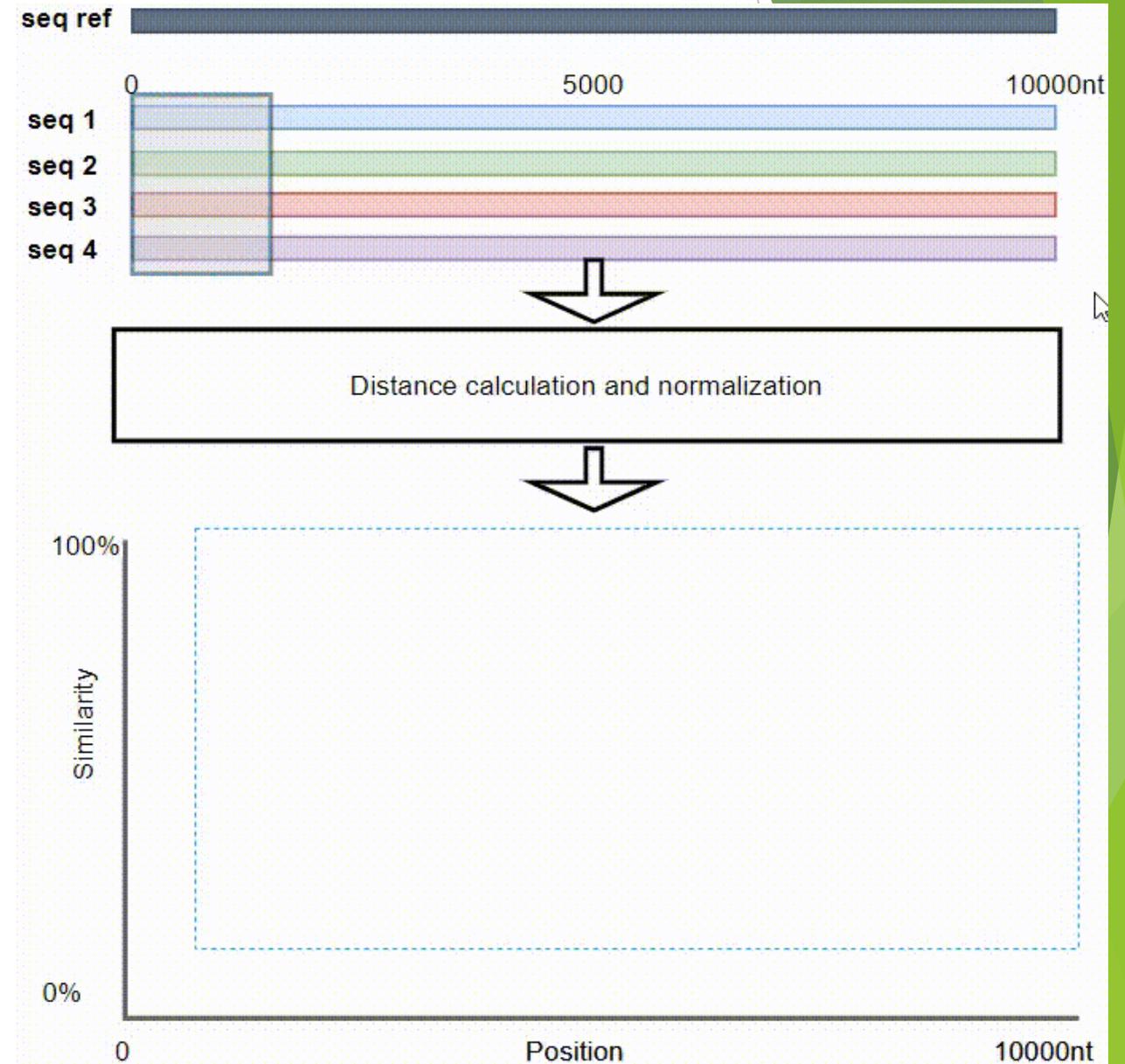
- ▶ Modèle de transfert complet
  - ▶ Incorporer le gène en entier dans le génome récipient
  - ▶ Recombinaison intergénique
- ▶ Modèle de transfert partiel
  - ▶ Formation de gènes mosaïques
  - ▶ Recombinaison intragénique



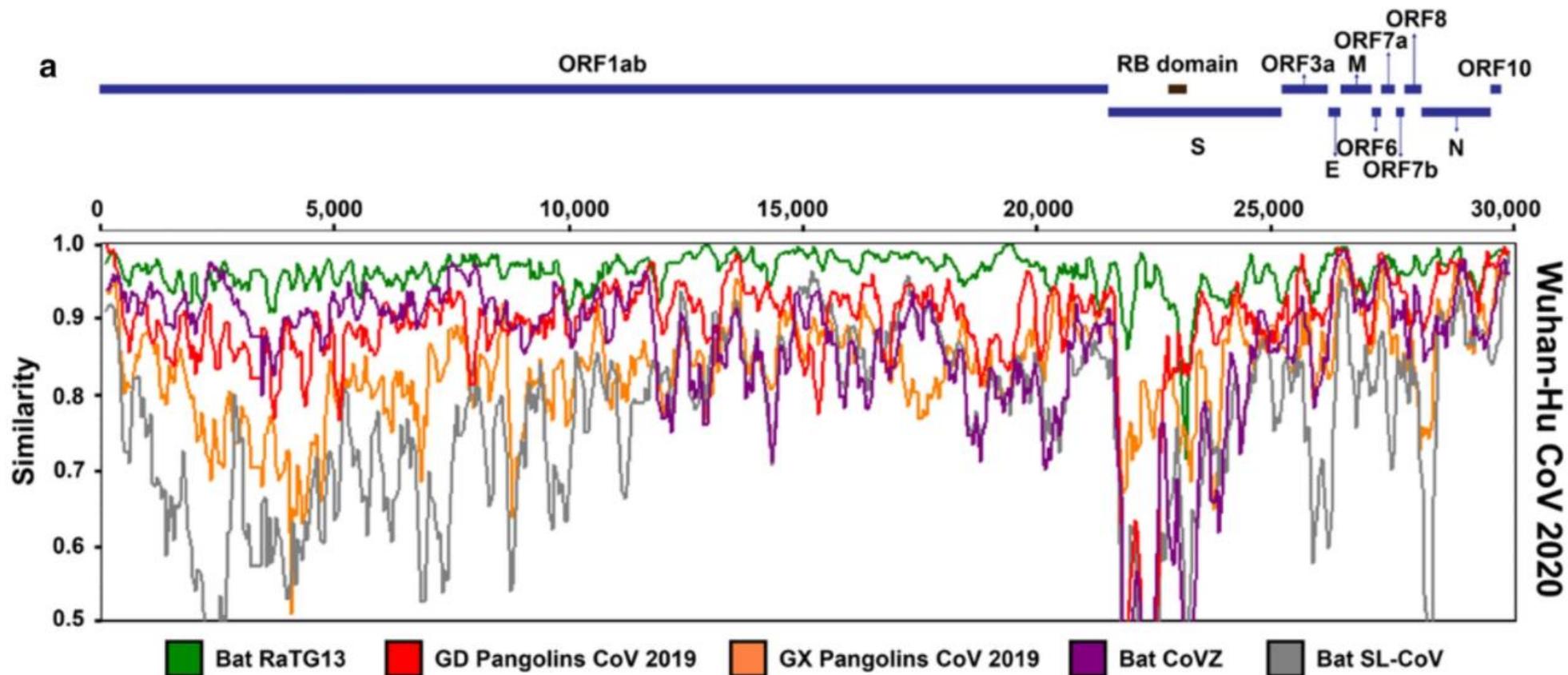
Deux modèles évolutifs assumant un HGT partiel (A) ou complet (B). Dans le premier cas, une partie du gène est incorporée dans le génome récipient et l'arbre est transformé en réseau dirigé, alors que dans le second, le gène donneur est acquis dans son intégralité par le génome receveur et l'arbre d'espèces va être modifié

# Logiciel SimPlot

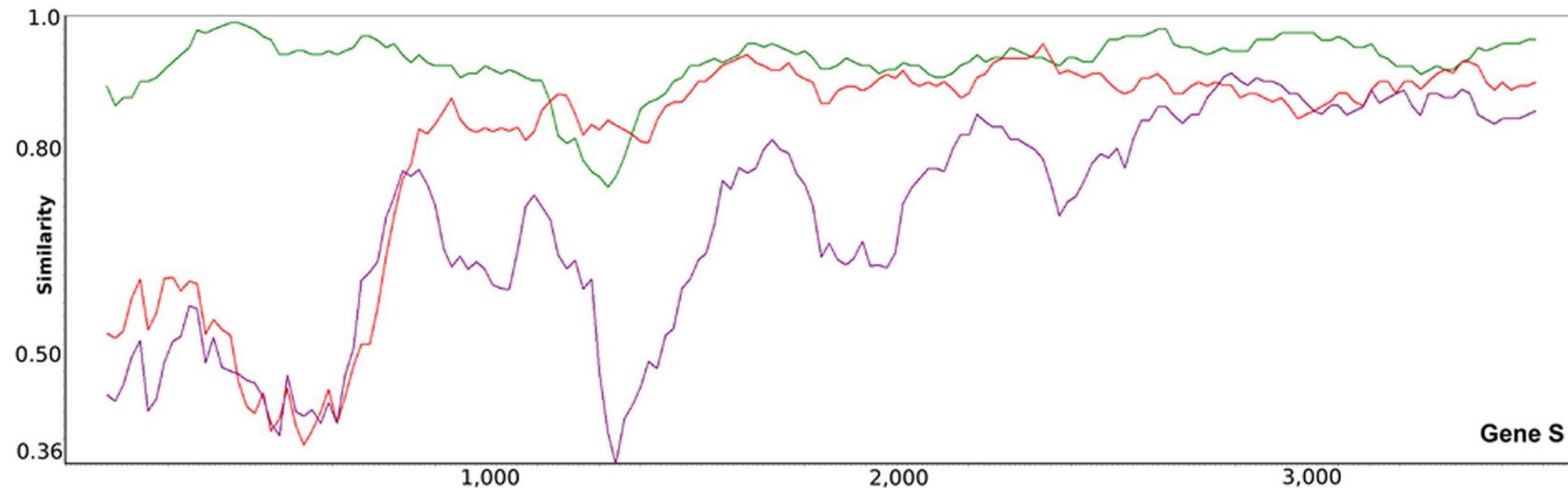
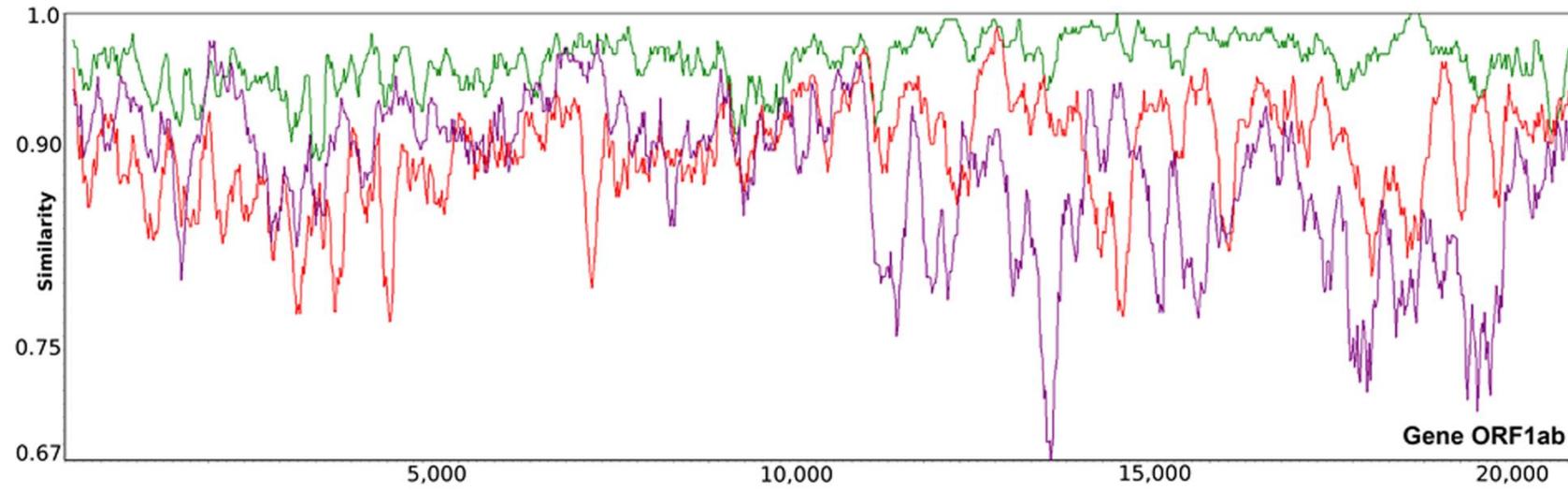
- ▶ Analyse de variation de similarité
  - ▶ Fenêtre et pas
  - ▶ Modèles des distances



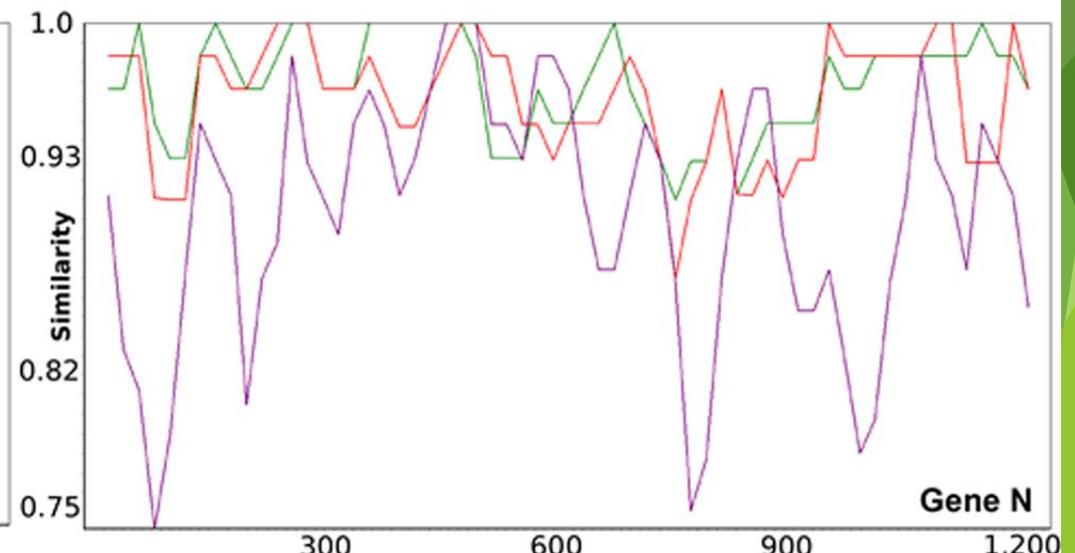
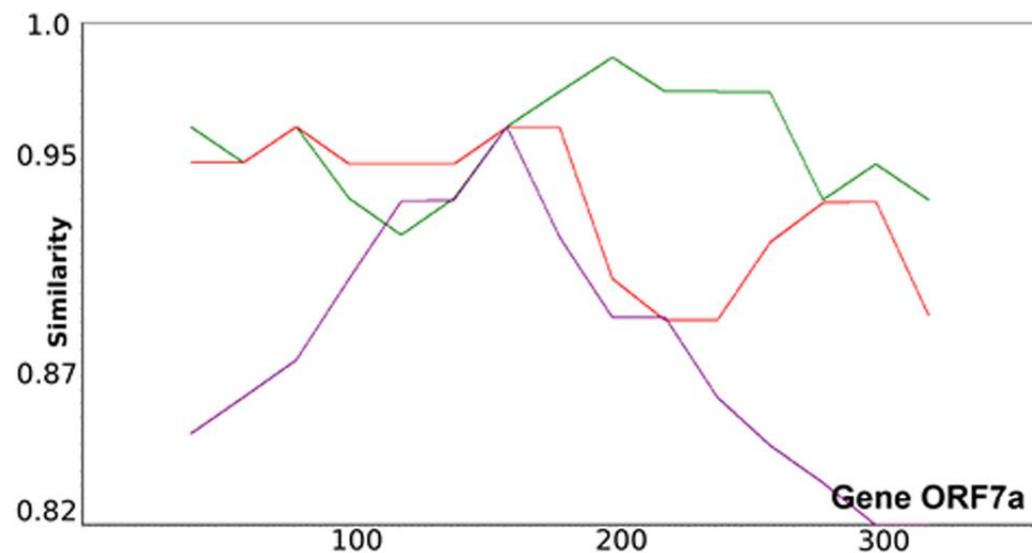
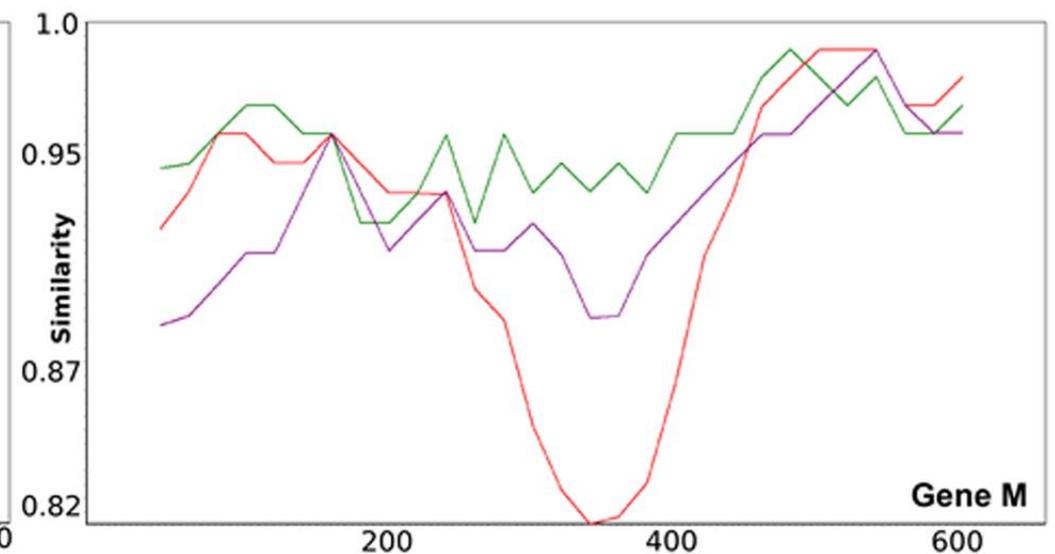
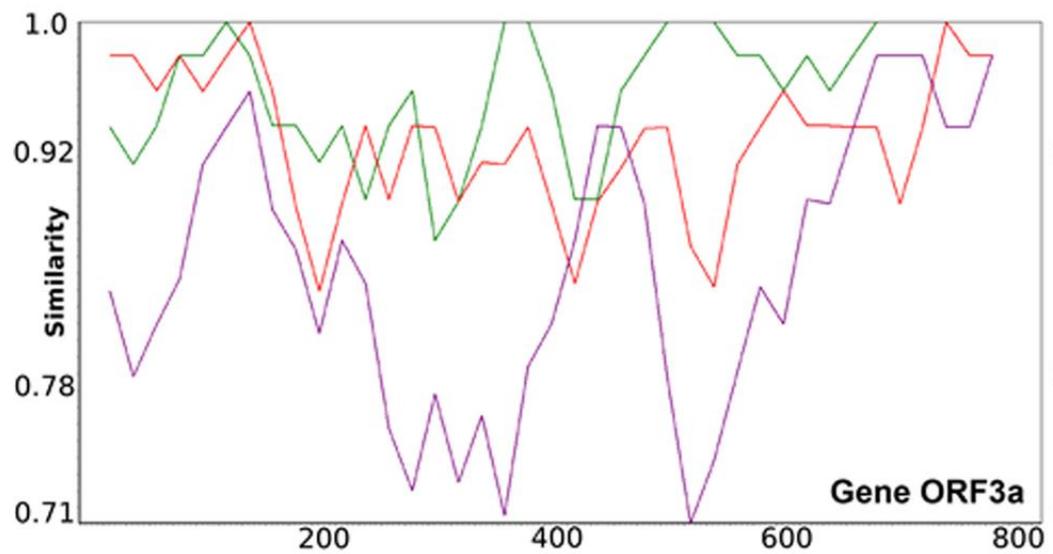
# Analyse de similarité des génomes de betacoronavirus faite par SimPlot



# Analyse de similarité des gènes ORF1ab et S de betacoronavirus faite par SimPlot



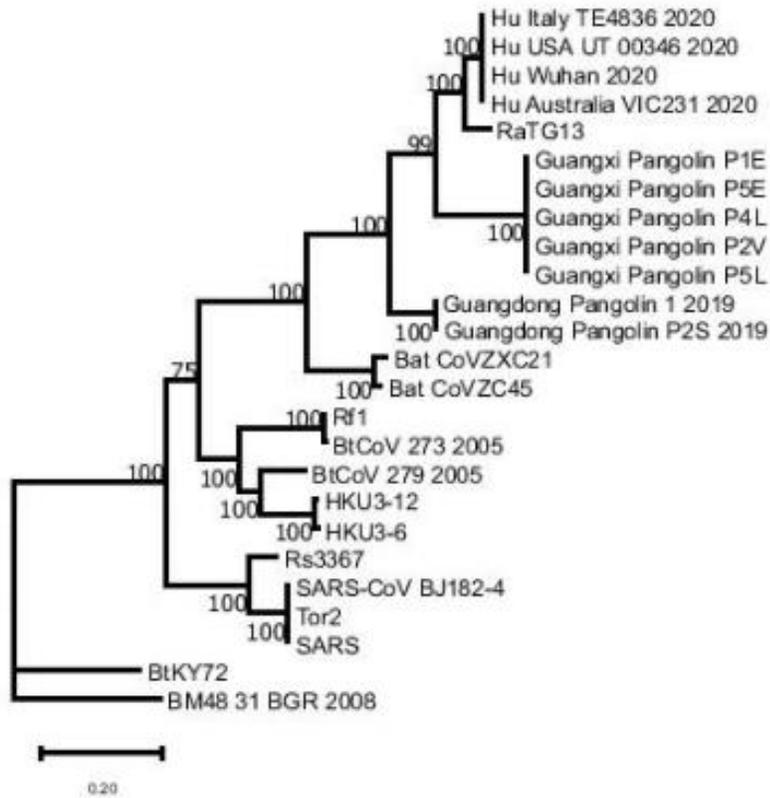
# Analyse de similarité des gènes ORF3a, M, ORF7a et N de betacoronavirus faite par SimPlot



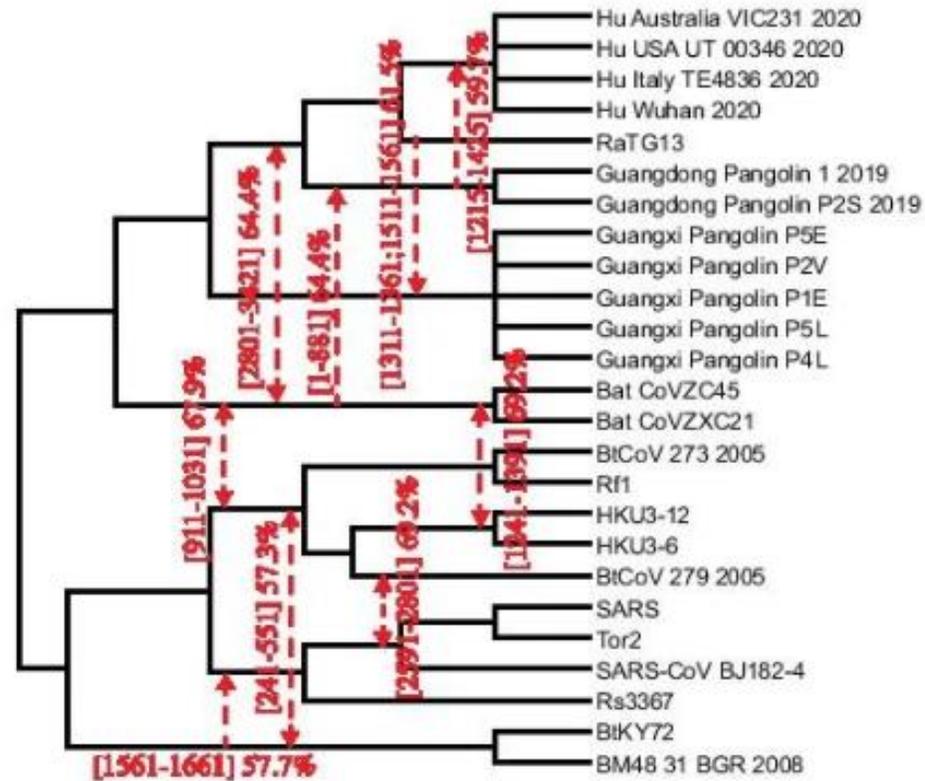
# Détection de transferts horizontaux de gènes pour le gène S

**b**

Tree for gene S

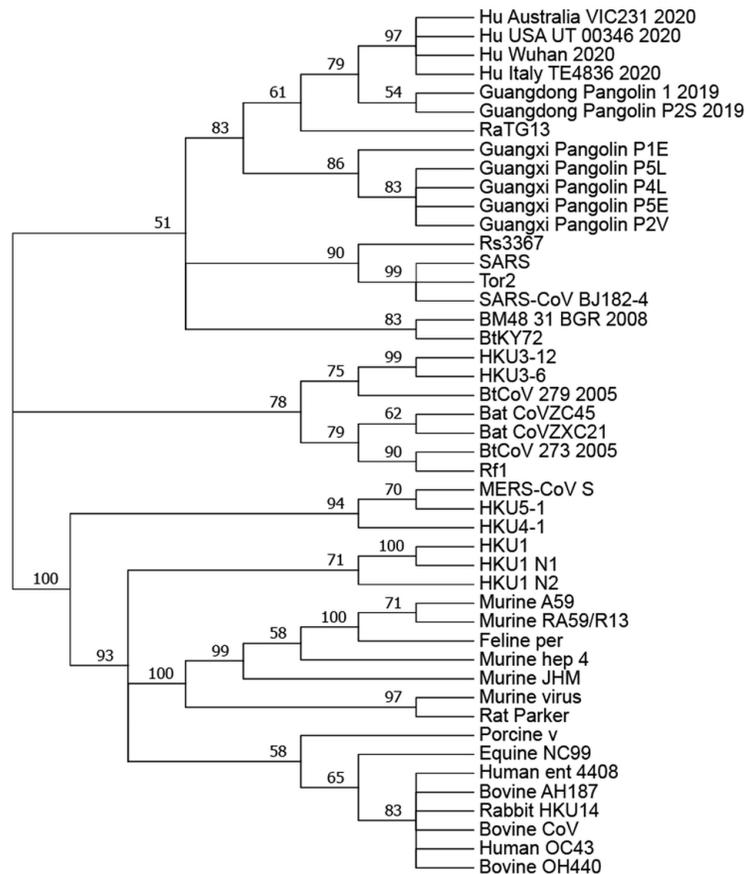


Putative gene transfer and recombination events for gene S

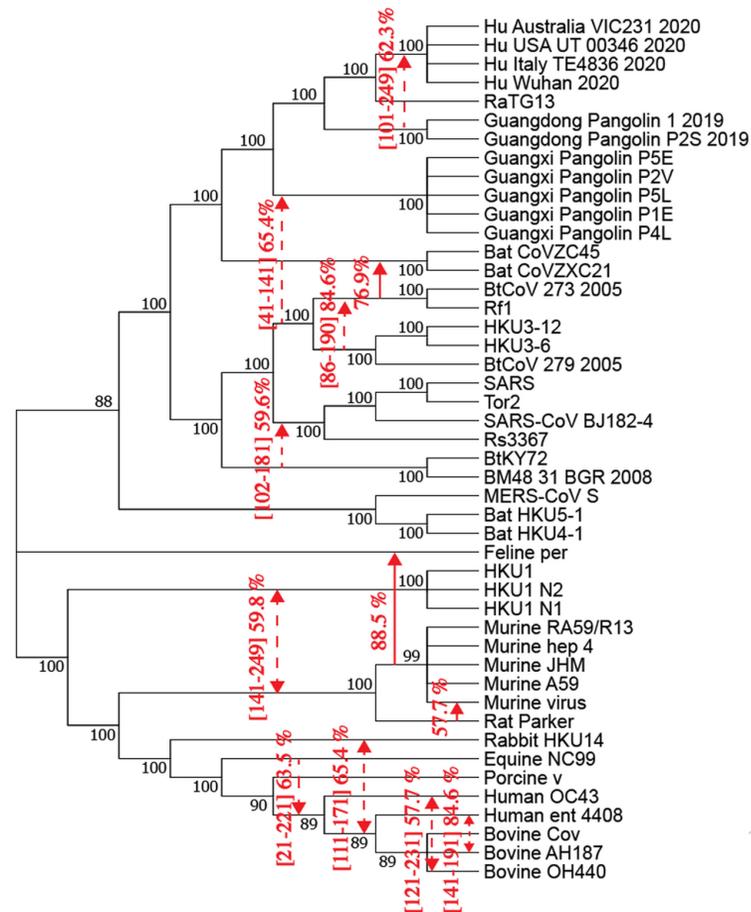


# Détection des transferts horizontaux de gènes pour le RB domain

**a** Tree for RB domain for 46 Betacoronavirus organisms

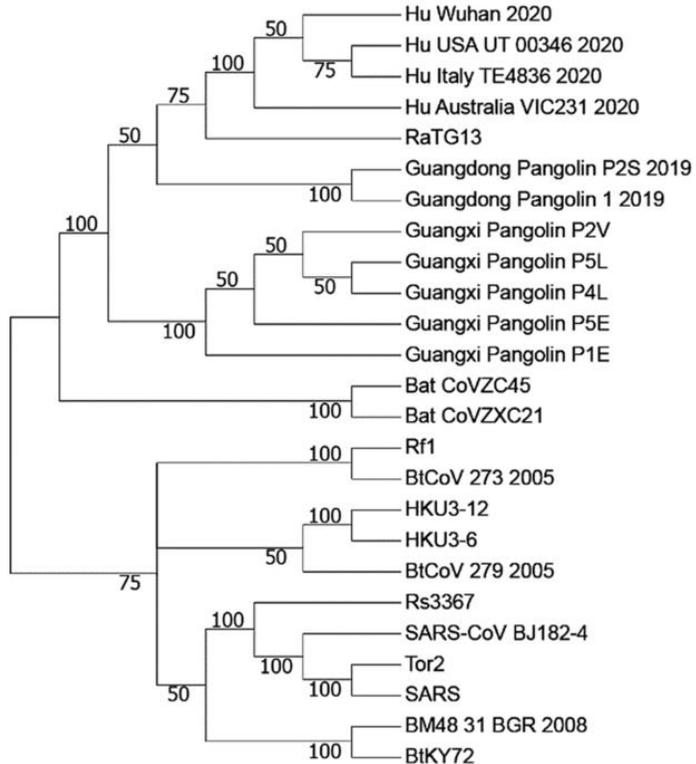


**b** Putative gene transfer and recombination events for RB domain for 46 Betacoronavirus organisms

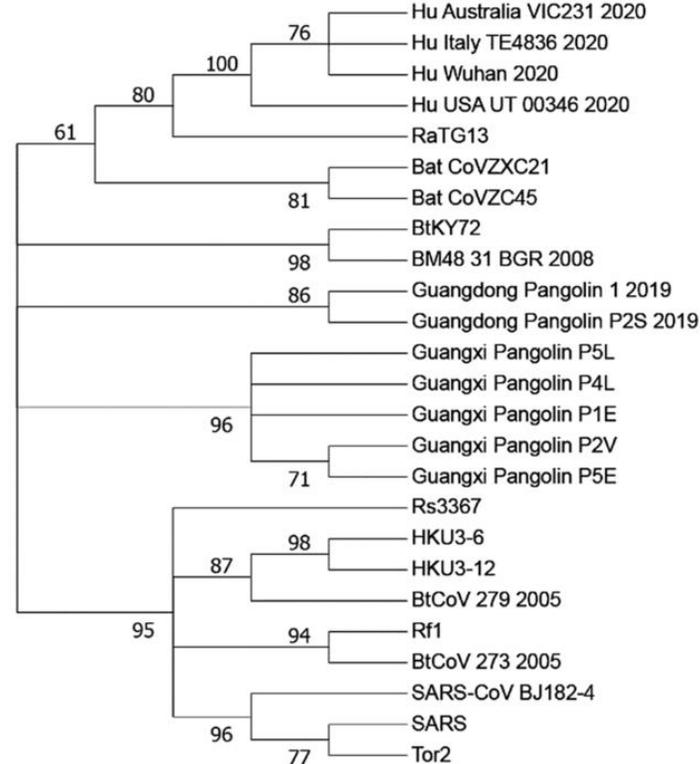


# Trois modèles d'évolution des gènes de SARS-CoV-2

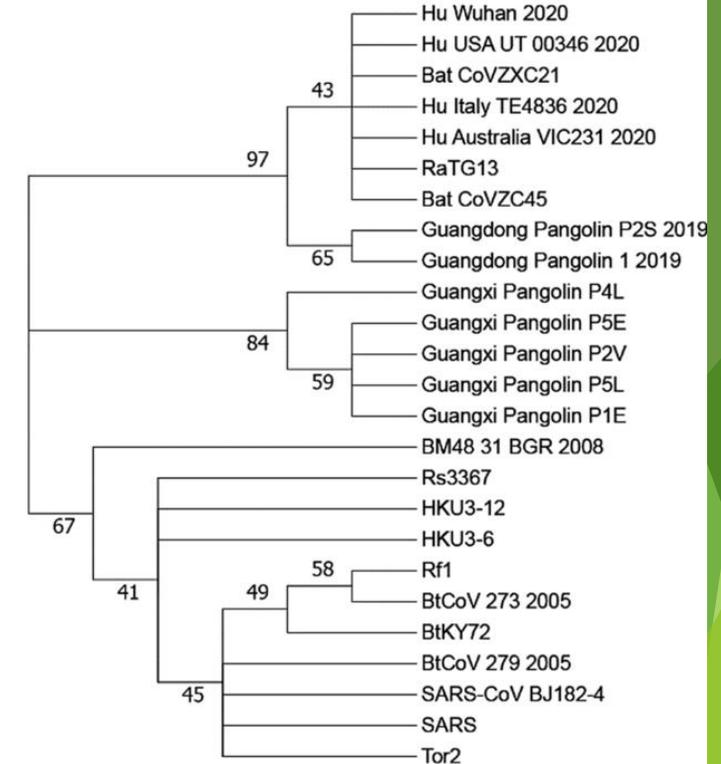
**a** Consensus tree for genes ORF1ab, S, RB domain, N



**b** Consensus tree for genes ORF3a, E, M, ORF6, ORF7a, ORF7b, ORF8



**c** Consensus (gene) tree for gene ORF10



# Analyse de recombinaison par $\Phi$ -test

Region	$\Phi$ -test result (p-value)	Recombination detected (yes/no)	Window size
Gene ORF1ab	$2.23 \times 10^{-3}$	Yes	150
RB domain	Too short	–	
Gene S	$8.60 \times 10^{-3}$	Yes	200
Gene ORF3a	$1.13 \times 10^{-5}$	Yes	100
Gene E	0.56	No	200
Gene M	0.226	No	100
Gene ORF6	Too short	–	
Gene ORF7a	0.00155	Yes	200
Gene ORF7b	Too short	–	
Gene ORF8	0.0453	Yes	200
Gene N	0.024	Yes	50
Gene ORF10	Too short	–	
Whole genomes	0.0156	Yes	200

On rejette l'hypothèse nulle ( $H_0$ ) d'absence de recombinaison selon la corrélation généalogique entre les sites adjacents et distants.

- Cette corrélation est invariable si l'on effectue des permutations de sites dans des séquences non-recombinées puisqu'elles partagent une histoire évolutive.
- Chez les séquences mosaïques, les sites adjacents sont plus corrélés que les sites distants.

# SimPlot++

The screenshot displays the SimPlot++ application window. The title bar reads "SimPlot++". The main menu bar includes icons for "Group Page", "SimPlot", "BootScan", "Find Sites", and "Network". On the left sidebar, there are buttons for "File Browser", "ReOpen", and "Save Groups".

The central area is titled "List of current groups (at least 2 groups needed)". It contains a list of five groups, each with a colored circle, a text input field, and a sequence count:

- SARS-CoV-2** (Red circle): 3 sequences
- RatG13** (Purple circle): 1 sequence
- GD Pangolin** (Pink circle): 2 sequences
- GX Pangolin** (Blue circle): 6 sequences

Below this list is a large red arrow pointing upwards with the text "Groupes".

To the right of the group list are two buttons: "New group" and "Delete group".

The right-hand panel is titled "SARS-CoV-2" and contains a list of three sequences:

- BetaCoV/Wuhan/IVDC-HB-04/2020
- BetaCoV/Wuhan/WIV04/2019
- BetaCoV/Wuhan-Hu-1/2019

A large red arrow points upwards from the text "Séquences du groupe SARS-CoV-2" towards the sequence list.

Below the SARS-CoV-2 panel is a section titled "Ungrouped Sequences" containing a list of ten sequences:

- bat-SL-CoVZC45
- bat-SL-CoVZXC21
- BtCoV/273/2005
- BtCoV/279/2005
- BtCoV/BM48-31/BGR/2008
- BtKY72
- HKU3-13
- HKU3-6
- PC4\_13
- Rf1
- Rs3367
- Tor2

A large red arrow points from the text "Séquences non-groupées" towards the "Ungrouped Sequences" list.

# SimPlot ++



Group Page



SimPlot



BootScan



Find Sites



Network

- ▶ SimPlot: Analyse de similarité par fenêtre coulissante
- ▶ BootScan: détection de variation de support de bootstrap pour les clades d'un arbre phylogénique par fenêtre coulissante
  - ▶ Bootstrap
- ▶ Find Sites: Identification des sites informatifs
- ▶ Network: visualisation sous forme de réseau des résultats de SimPlot

# Références

- ▶ Kumar, M., Al Khodor, S. Pathophysiology and treatment strategies for COVID-19. *J Transl Med* **18**, 353 (2020). <https://doi.org/10.1186/s12967-020-02520-8>
- ▶ Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol.* 1999 Jan;73(1):152-60.
- ▶ Makarenkov, V., Mazouze, B., Rabusseau, G. et al. Horizontal gene transfer and recombination analysis of SARS-CoV-2 genes helps discover its close relatives and shed light on its origin. *BMC Ecol Evo* 21, 5 (2021). <https://doi.org/10.1186/s12862-020-01732-2>
- ▶ Tao Zhang, Qunfu Wu, Zhigang Zhang, Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak, *Current Biology*, Volume 30, Issue 7, 2020, Pages 1346-1351.e2,
- ▶ Trevor C Bruen, Hervé Philippe, David Bryant, A Simple and Robust Statistical Test for Detecting the Presence of Recombination, *Genetics*, Volume 172, Issue 4, 1 April 2006, Pages 2665–2681, <https://doi.org/10.1534/genetics.105.048975>