

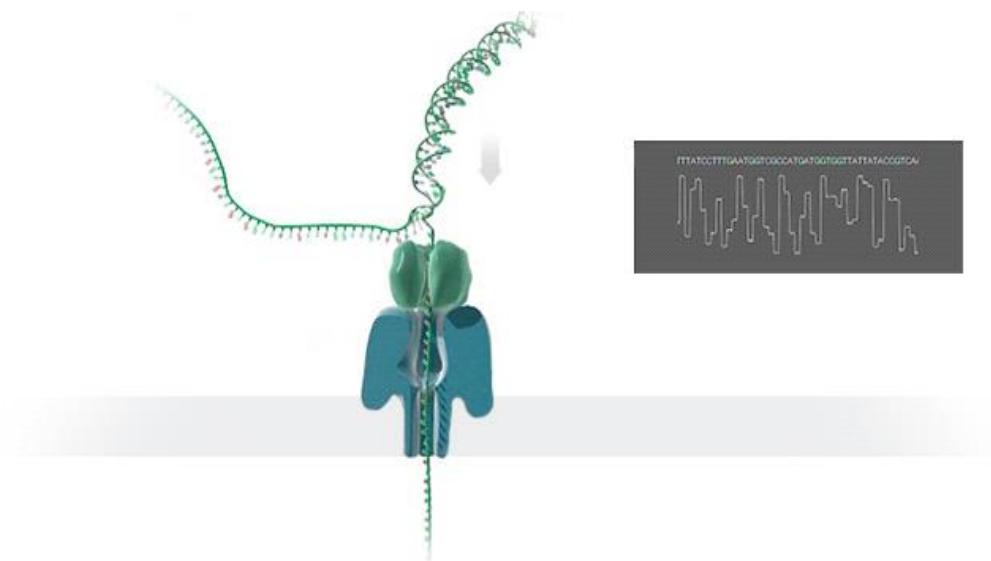


INSTITUTO
NACIONAL DE
SALUD

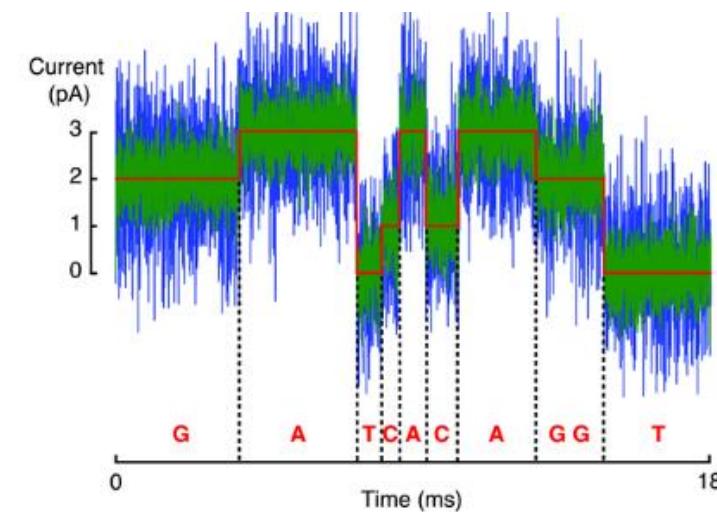
ENSAMBLAJE CON ARTIC Y ANÁLISIS DE VARIANTES

Instituto Nacional de Salud
Bogotá

Archivos fast5 y llamado de bases: datos de señal eléctrica



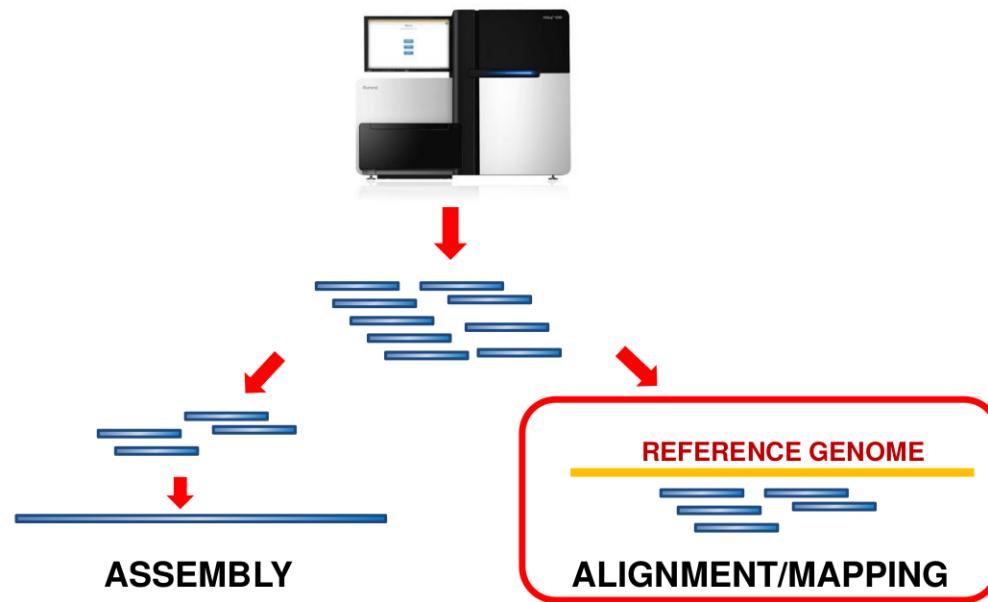
<https://nanoporetech.com/how-it-works>



<https://konradpaszkiewicz.files.wordpress.com/2014/04>



Ensamblaje por mapeo

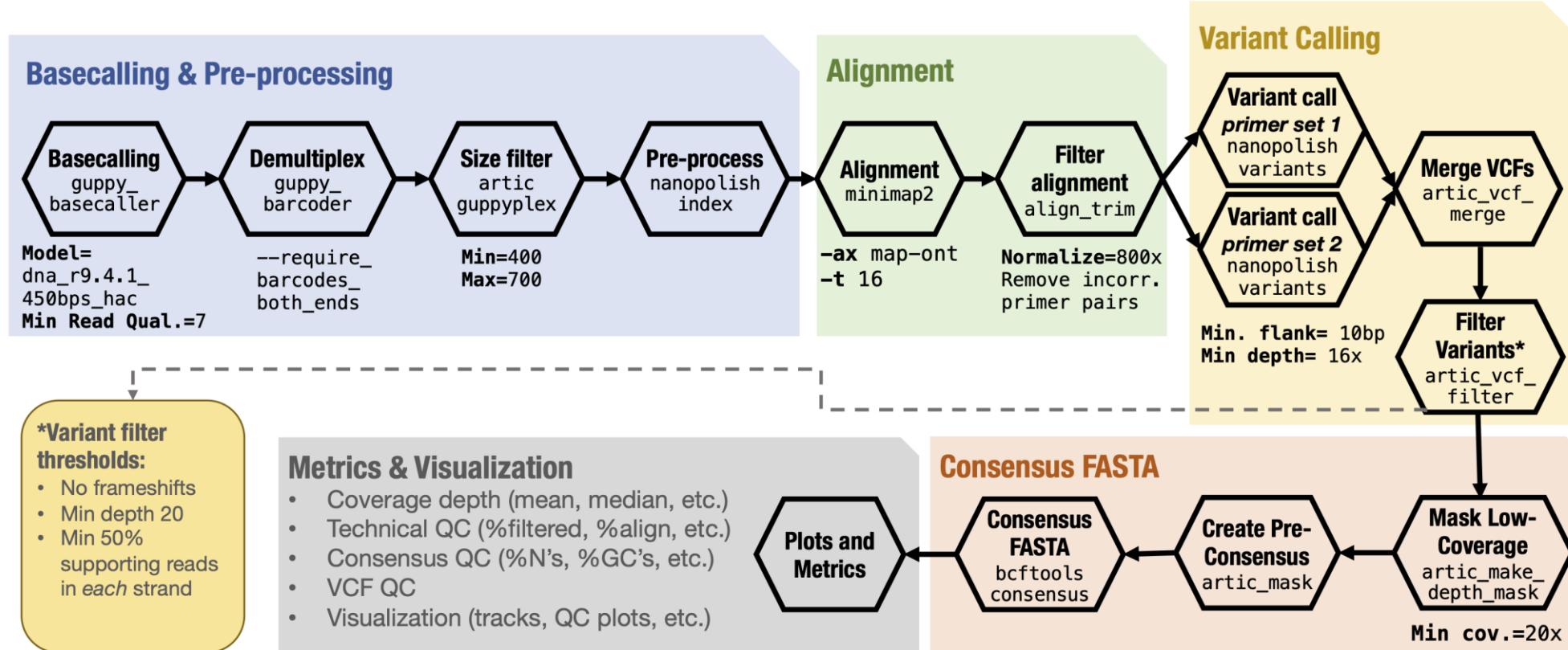


Ensamblaje por mapeo



SARS-CoV2 ARTIC Analysis Pipeline

(Nanopolish)



Reference Genome: Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1 (GenBank MN908947.3)

Software versions: guppy-GPU v3.4.4; minimap2 v2.17; samtools v1.9; bcftools v1.9; bedtools v2.27.0; python v3.6; nanopolish v0.13.1; muscle v3.8.31

ARTIC pipeline cloned from: <https://github.com/artic-network/fieldbioinformatics>

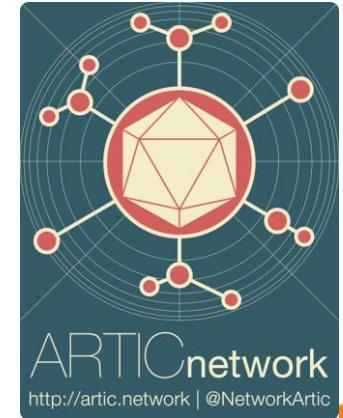


artic 1. Filtro por tamaño de lecturas

- artic guppyplex --min-length 400 --max-length 700 --directory
 - Remoción de lecturas químéricas y filtro por output_directory/barcode03 --prefix run_name

■ Directorio donde se guardan las secuencias filtradas

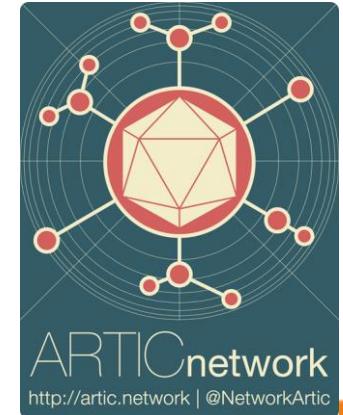
■ Prefijo con el que se guardan los archivos de salida



artic 2. Ensamblaje (minimap2 y Nanopolish)

```
artic minion --normalise 200 --threads 4 --scheme-directory ~/artic-ncov2019/primer_schemes --  
read-file run_name_${barcode}.fastq --fast5 directory path_to_fast5 --sequencing-summary  
path_to_sequencing_summary.txt nCoV-2019/V3 samplename
```

- Directorio donde están los datos crudos fast5
- Prefijo con el que se guardamos los archivos de salida en el paso de control de calidad
- Sequencing summary de Guppy. Único para todos los barcodes.
- Nombre de la muestra (Usualmente el barcode que se analiza)



¿Qué hacemos luego de obtener una secuencia?

ATTCCACGTAGGAATGTGGCAACTTACAAGCTGAAAATGTAACAGGACTCT
TGTAGTAAGGTAATCACTGGGTTACATCCTACACAGGCACCTACACAC06CA
ACTAAAATTCAAAAATGAAAGGTTATGTGTTGACATACTGGCATACTAAGG
TATAGAAGACTCATCTCATGATGGGTTTAAAATGAATTATCAAGTTAATG
AACATGTTTATCACCCCGCGAAGAAGCTATAAGACATGTACGTGCATGGATTG
GTCGAGGGGTGTCATGCTACTAGAGAAGCTGTTGGTACCAATTACCTTAC
TTTCTACAGGTGTTAACCTAGTTGCTGTACCTACAGGTTATGTTGATAACAC
ACAGATTTTCCAGAGTTAGTGCTAAACCACCGCCTGGAGATCAATTAAAC
CCACTTATGTACAAAGGACTTCCTTCCAATGTAGTGCCTATAAAAGATTGTAC
AGTGACACACTAAAAATCTCTCTGACAGAGTCGTATTGTCTTATGGCAC
GAGTTGACATCTATGAAGTATTGTGAAAATAGGACCTGAGCGCACCTGTT
GATAGACGTGCCACATGCTTTCCACTGCTTCAGACACTTATGCCTGTTGGC
ATTGGATTTGATTACGTCTATAATCCGTTATGATTGATGTTCAACAATGGG
GGTAACCTACAAAGCAACCAGATCTGTATTGTCAAGTCCATGGTAATGCAC
AGTTGTGATGCAATCATGACTAGGTGTCTAGCTGCCACGAGTGCTTGTAA
GAUTGGACTATTGAATATCCTATAATTGGTGTAGACTGAAGGATTAATGCGG
AAGGTTCAACACATGGTTGTTAAAGCTGCATTATTAGCAGACAAATTCCCAG
GACATTGGTAACCCTAAAGCTATTAAGTGTGTACCTCAAGCTGATGTAGAAT
TATGATGCACAGCCTGTAGTGACAAAGCTTATAAAATAGAAGAATTATTCT
GCCACACATTCTGACAAATTCACAGATGGTGTATGCCTATTGGATTGCA
AGATATCCTGCTAATTCCATTGTTGTAGATTGACACTAGAGTGCTATCTA

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

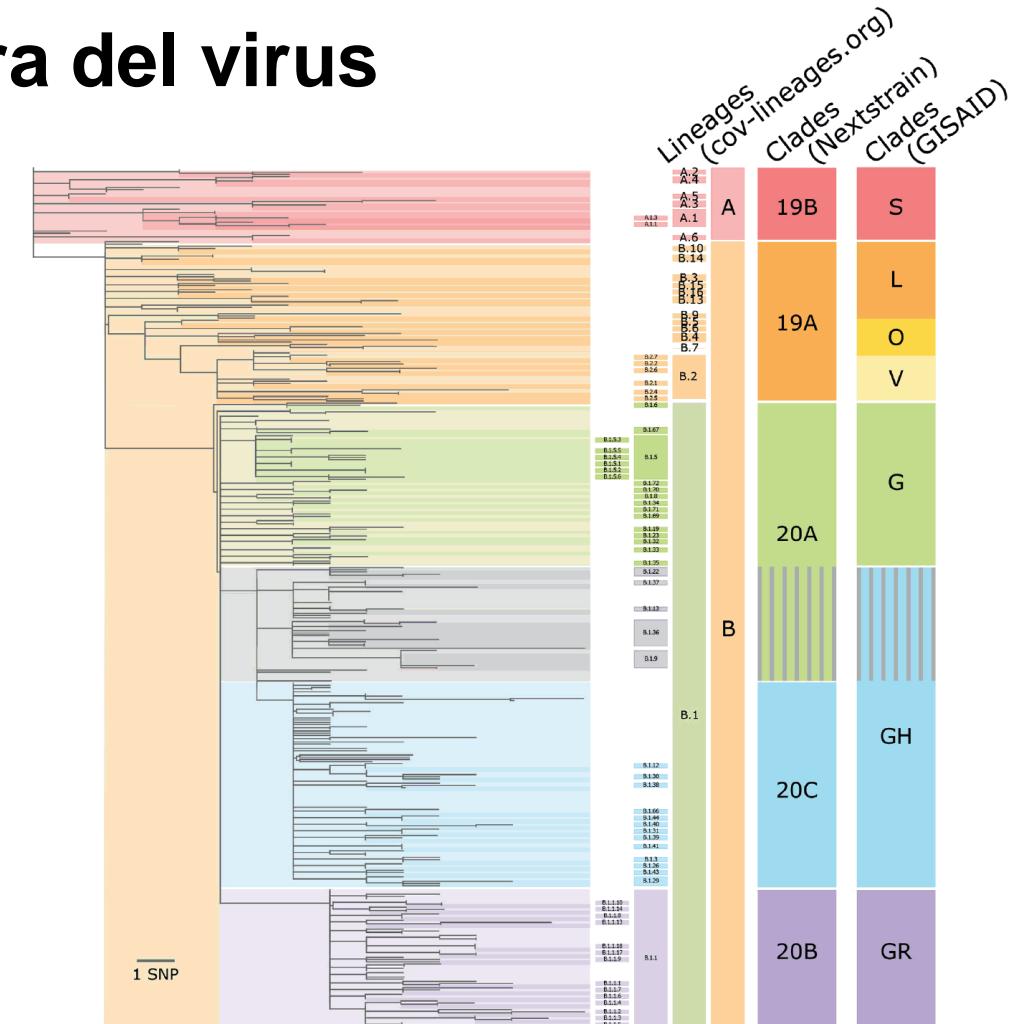
Retry Failed Sequences Reset entries Upload another file

File name	Sequence name	Lineage
— FAILED (Click warning icon for more info) 1 sequence		
 all_consensus.fasta	barcode15/ARTIC/nanopolish MN908947.3	
— ANALYSED (Click tick icon for more info) 5 sequences 		
 all_consensus.fasta	barcode09/ARTIC/nanopolish MN908947.3	B.1.1.28   
 all_consensus.fasta	barcode86/ARTIC/nanopolish MN908947.3	P.1   
 all_consensus.fasta	barcode37/ARTIC/nanopolish MN908947.3	B.1.1.348   
 all_consensus.fasta	barcode61/ARTIC/nanopolish MN908947.3	B.1.621   
 all_consensus.fasta	barcode79/ARTIC/nanopolish MN908947.3	B.1.625   



Nomenclatura del virus

- La nomenclatura de linajes se propuso a partir de análisis de filogenia por máxima verosimilitud.
- Hay varios sistemas de nomenclatura

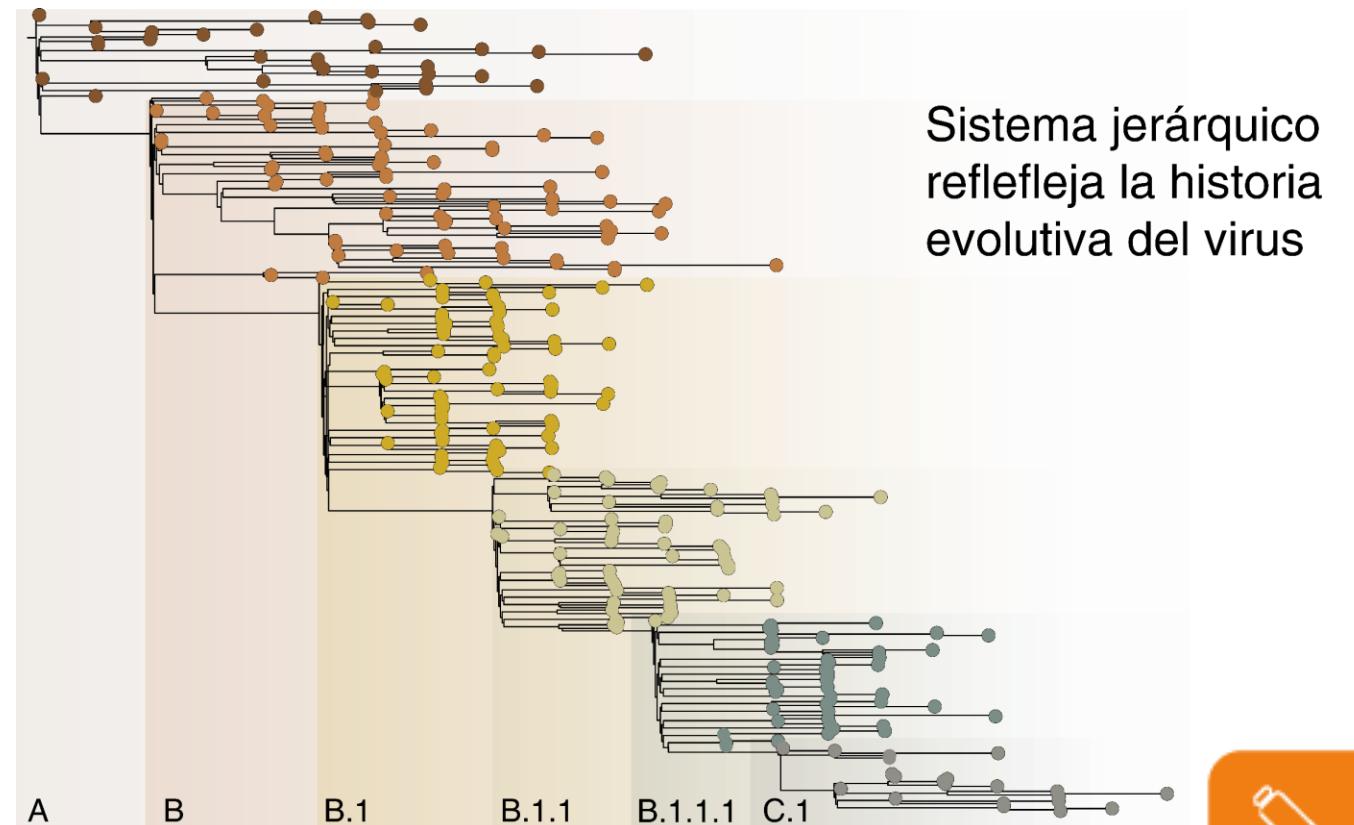


<https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.32.2001410?crawler=true>



Nomenclatura de linajes

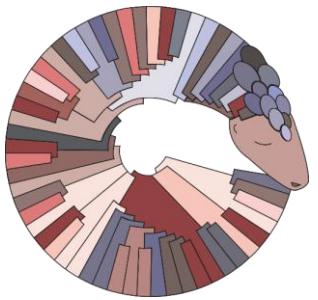
- La nomenclatura tiene máximo tres niveles.
- Sublinajes más allá se les asigna un alias.
- Ejemplo:
 - C.1 es un alias de B.1.1.1.1
 - P.1 es un alias de B.1.1.28.1



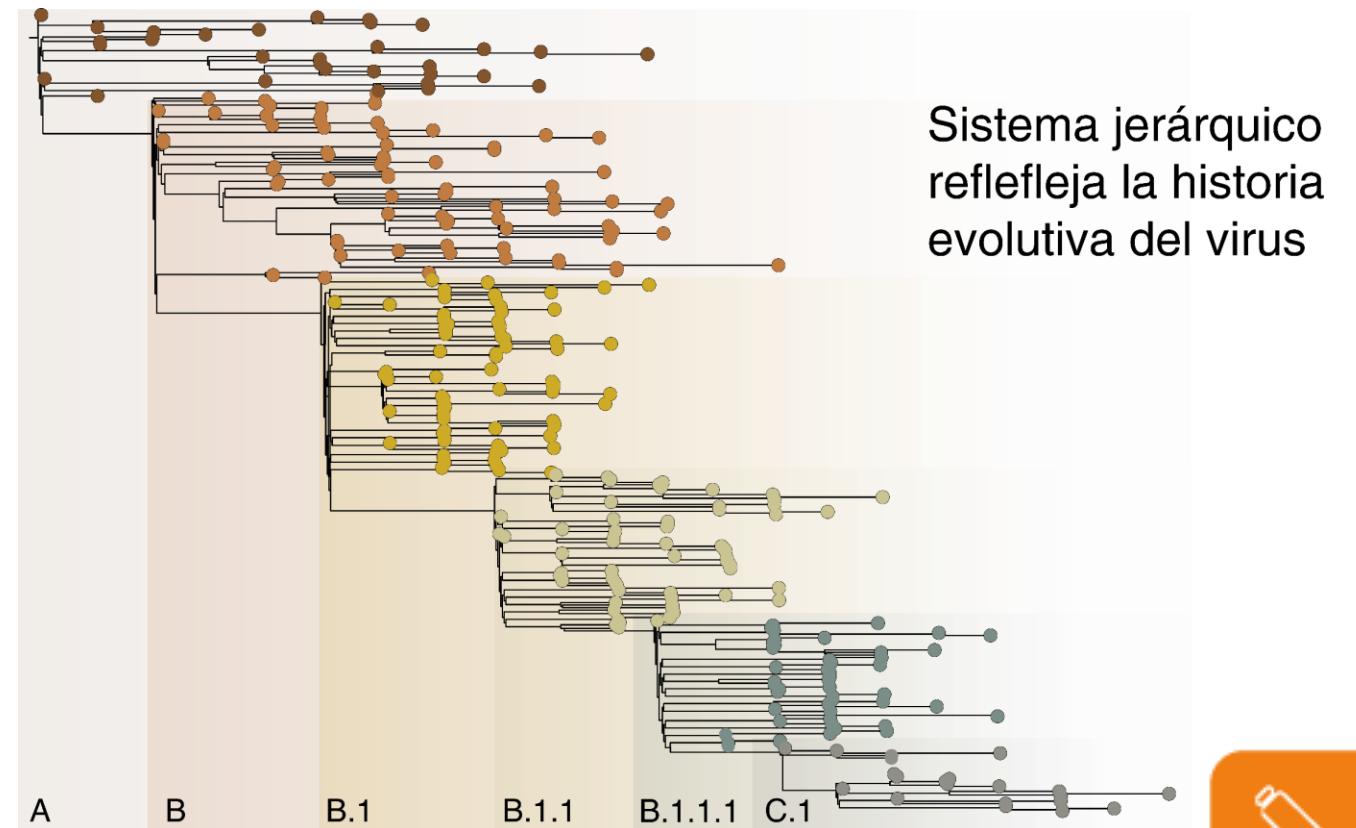
<https://www.pango.network/how-does-the-system-work/what-are-pango-lineages/>



Clasificación de linajes



UShER



<https://www.pango.network/how-does-the-system-work/what-are-pango-lineages/>



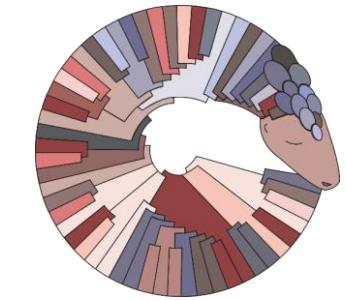
Salida de Pangolin (.csv)

taxon	lineage	conflict	ambiguity_s	scorpio_call	scorpio_sup	scorpio_conf	version	pangolin_ve	pangoLEARN
barcode09/A	B.1.1.28	0.0	0,57				PLEARN-v1.23.1.11	9/08/2021	
barcode86/A	P.1	0.0	0,97	Gamma (P.1-like)	0,875	0.000000	PLEARN-v1.23.1.11	9/08/2021	
barcode37/A	B.1.1.348	0.0	0,9				PLEARN-v1.23.1.11	9/08/2021	
barcode61/A	B.1.621	0.0	1	B.1.621-like	1	0.000000	PLEARN-v1.23.1.11	9/08/2021	
barcode79/A	B.1.625	0.0	1				PLEARN-v1.23.1.11	9/08/2021	
barcode15/A	None						PANGO-v1.23.1.11	9/08/2021	

<https://cov-lineages.org/resources/pangolin/output.html>

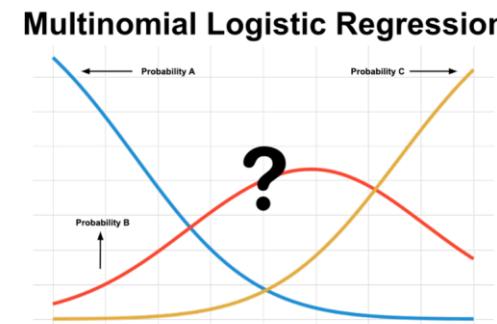
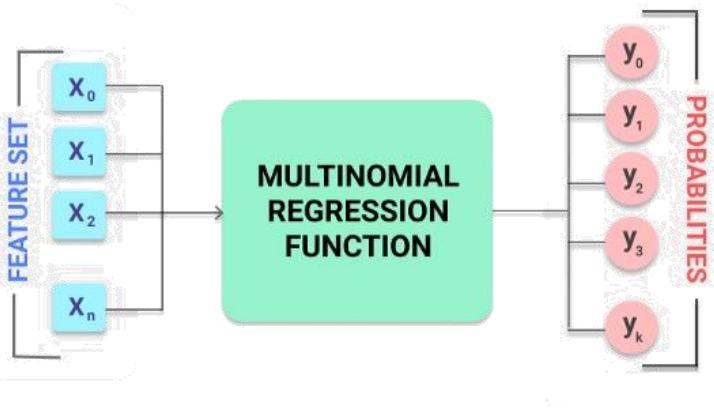


¿Cómo asigna linajes Pangolin?



- PangoLEARN 1 – Regresión Logística Multinomial

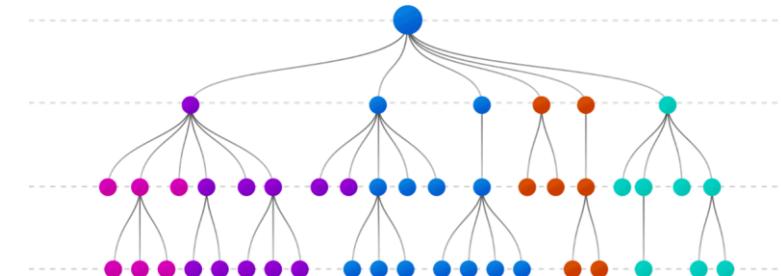
Asigna una probabilidad a la asignación de cada posible linaje para clasificar las secuencias.



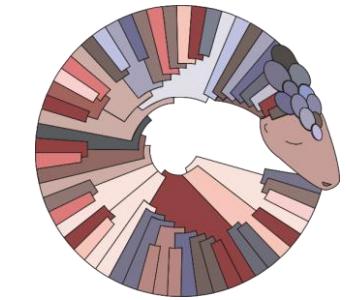
- **PangoLEARN Actual – Árboles de decisión**

Decisiones condicionales en una estructura jerárquica.

La información viene de filogenias, por lo que la estructura jerárquica replica este comportamiento.

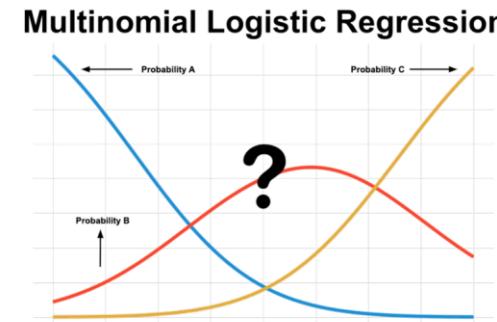
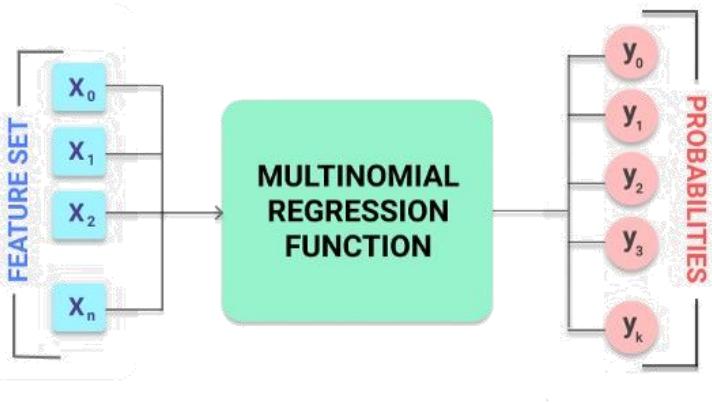


¿Cómo asigna linajes Pangolin?

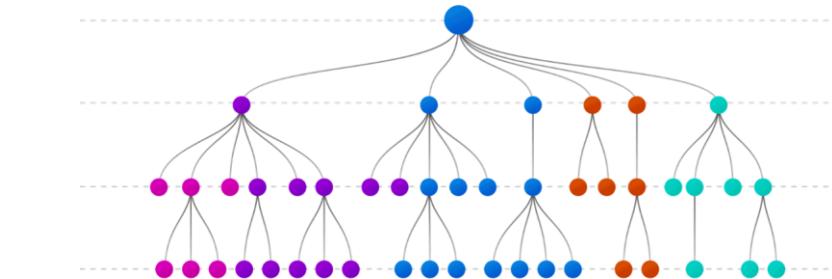


- PangoLEARN 1 – Regresión Logística Multinomial

Asigna una probabilidad a la asignación de cada posible linaje para clasificar las secuencias.



El modelo cambia cada vez que aparecen nuevos linajes



Scorpio – Apoyo a la asignación de VOCs/VOIs



- Asignación de linajes basada en llamado de variantes (SNPs, INDELs)
- Usada para VOCs, VOIs, y linajes difíciles de asignar

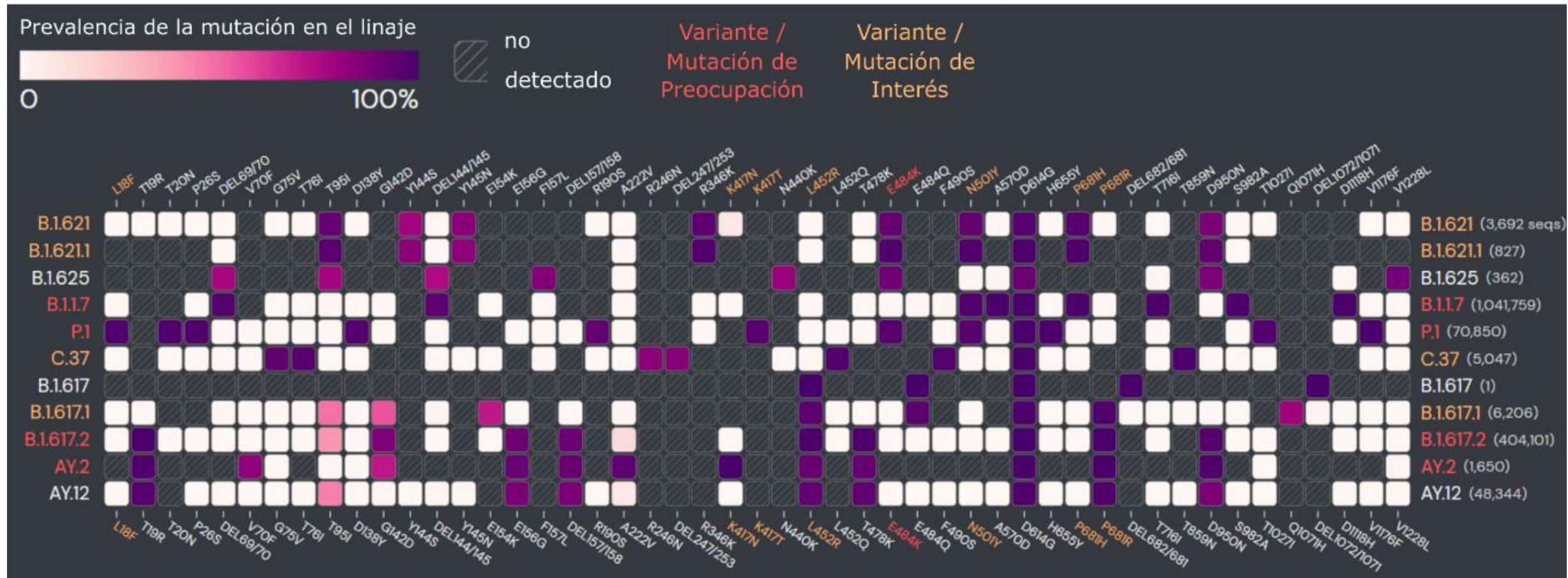


Salida de Pangolin con Scorpio (.csv)

Dos	taxon	lineage	conflict	ambiguity_s	scorpio_call	scorpio_sup	scorpio_conf	version	pangolin_ve	pangoLEARN
lina	barcode09/AB.1.1.28		0.0		0,57			PLEARN-v1.23.1.11	9/08/2021	
sec	barcode86/AP.1		0.0		0,97 Gamma (P.1-like)		0,875 0.000000	PLEARN-v1.23.1.11	9/08/2021	
con	barcode37/AB.1.1.348		0.0		0,9			PLEARN-v1.23.1.11	9/08/2021	
	barcode61/AB.1.621		0.0		1 B.1.621-like		1 0.000000	PLEARN-v1.23.1.11	9/08/2021	
	barcode79/AB.1.625		0.0		1			PLEARN-v1.23.1.11	9/08/2021	
	barcode15/A None							PANGO-v1.2 3.1.11	9/08/2021	

<https://cov-lineages.org/resources/pangolin/output.html>





<https://outbreak.info/compare-lineages>



Mutaciones con consecuencias

New Results

[View current version of this article](#)

[Follow this preprint](#)

SARS-CoV-2 spike P681R mutation enhances and accelerates viral fusion

The N501Y spike substitution enhances SARS-CoV-2 transmission

Yang Liu,^{1,2,*} Jianying Liu,^{2,3,4,*} Kenneth S. Plante,^{2,3,4,*} Jessica A. Plante,^{2,3,4} Xuping Xie,¹ Xianwen Zhang,¹ Zhiqiang Ku,⁵ Zhiqiang An,⁵ Dionna Scharton,^{2,3,4} Craig Schindewolf,^{2,3,4} Vineet D. Menachery,^{2,3,4} Pei-Yong Shi,^{1,2,3,#} and Scott C. Weaver^{2,3,4,6,#}

a,

► Author information ► Copyright and License information ► Disclaimer

PDF [409 KB]

SARS-CoV-2 spike E484K mutation reduces antibody neutralisation

Sonia Jangra • Chengjin Ye • Raveen Rathnasinghe • Daniel Stadlbauer • Personalized Virology Initiative study group
Florian Krammer • et al. Show all authors

Open Access • Published: 01 June 2021 • DOI: [https://doi.org/10.1016/S0960-2567\(21\)00068-9](https://doi.org/10.1016/S0960-2567(21)00068-9)

Short Article

SARS-CoV-2 spike L452R variant evades cellular immunity and increases infectivity

Chihiro Motozono¹, Mako Toyoda^{1,22}, Jiri Zahradnik^{2,22}, Akatsuki Saito^{3,4,5,22}, Hesham Nasser^{6,7},
22, Toong Seng Tan¹, Isaac Ngare¹, Izumi Kimura⁸, Keiya Uriu⁸, Yusuke Kosugi⁸, Yuan Yue⁶, Ryo
Shimizu⁶, Junpei Ito⁸, Shiho Torii^{9,10,11}, Akiko Yonekawa¹², Nobuyuki Shimono¹², Yoji Nagasaki¹³,
Rumi Minami¹⁴ ... Kei Sato^{8,20,21,23}



SPIKE: mutaciones de interés											
VOC / VOI	Linaje	R346K*	K417*	L452*	S477N ⁺	T478K ^{**}	E484K*	F490S*	N501Y ⁺	H655Y*	P681**
Alpha	B.1.1.7							X		P681H	
Gamma	P.1; P.1.1; P.1.2			K417T		X		X	X		
Delta	B.1.617.2; AY.3			L452R	X				P681R		
Delta	AY.1; AY.2			K417N	L452R	X				P681R	
Epsilon	B.1.427; B.1.429			L452R							
Iota	B.1.526			L452R	X	X					
Lambda	C.37			L452Q			X				
B.1.621	B.1.621	X				X	X		P681H		

* Posible evasión/atenución respuesta inmune

+ Evidencia mejor unión Spike-ACE2

** Sitio Clivado Furinio S1/S2

Review Article | Published: 01 June 2021

SARS-CoV-2 variants, spike mutations and immune escape

William T. Harvey, Alessandro M. Carabelli, Ben Jackson, Ravindra K. Gupta, Emma C. Tho Harrison, Catherine Ludden, Richard Reeve, Andrew Rambaut, COVID-19 Genomics UK (Consortium, Sharon J. Peacock & David L. Robertson✉

Nature Reviews Microbiology **19**, 409–424 (2021) | Cite this article

98k Accesses | 1 Citations | 878 Altmetric | Metrics



<https://outbreak.info/compare-lineages>



VOI

WHO label	Pango lineage*	GISAID clade	Nextstrain clade	Additional amino acid changes monitored ^o	Earliest documented samples	Date of designation
Alpha	B.1.1.7 [#]	GRY	20I (V1)	+S:484K +S:452R	United Kingdom, Sep-2020	18-Dec-2020
Beta	B.1.351	GH/501Y.V2	20H (V2)	+S:L18F	South Africa, May-2020	18-Dec-2020
Gamma	P.1	GR/501Y.V3	20J (V3)	+S:681H	Brazil, Nov-2020	11-Jan-2021
Delta	B.1.617.2 [§]	G/478K.V1	21A	+S:417N	India, Oct-2020	VOI: 4-Apr-2021 VOC: 11-May-2021

VOC

WHO label	Pango lineage*	GISAID clade	Nextstrain clade	Earliest documented samples	Date of designation
Eta	B.1.525	G/484K.V3	21D	Multiple countries, Dec-2020	17-Mar-2021
Iota	B.1.526	GH/253G.V1	21F	United States of America, Nov-2020	24-Mar-2021
Kappa	B.1.617.1	G/452R.V3	21B	India, Oct-2020	4-Apr-2021
Lambda	C.37	GR/452Q.V1	21G	Peru, Dec-2020	14-Jun-2021
Mu	B.1.621	GH	21H	Columbia, Jan-2021	30-Aug-2021

<https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>



Llamado de SNPs, INDELs, sustituciones aa

Nextclade

Clade assignment, mutation calling, and sequence quality checks

Simple
No installation or setup - drop a file and see the results

Private
No remote processing - sequence data never leaves your computer

Mutation Calling
Find differences of your sequences relative to the reference in standard numbering

Clade Assignment
Find out which Nextstrain clades your samples are from

Phylogenetic Placement
See where on the SARS-CoV-2 tree your sequences fall

Quality Control
Check your data against multiple QC metrics

SARS-CoV-2 (Wuhan-Hu...)

Simple mode Advanced mode

Sequences required

Drag & Drop a file here

or



<https://clades.nextstrain.org/>

Revisar mutaciones, efecto en primers, frameshifts

Sequence name	QC	Clade	Mut.	non-ACGTN	Ns	Gaps	Ins.	Gene ORF3a
barcode09/ARTIC/nanopolish MN908947.3	N M P C F S	20J (Gamma, V3)	17	0	8974	9	4	
barcode15/ARTIC/nanopolish MN908947.3	N M P C F S	20A	19	0	13387	8	0	
barcode86/ARTIC/nanopolish MN908947.3	N M P C F S	20J (Gamma, V3)	36	0	976	9	4	
barcode37/ARTIC/nanopolish MN908947.3	N M P C F S	20B	24	0	2840	0	0	
barcode61/ARTIC/nanopolish MN908947.3	N M P C F S	21H	48	0	122	4	3	Gene "ORF3a" is missing
barcode79/ARTIC/nanopolish MN908947.3	N M P C F S	20A	30	0	123	23	0	

A	C	T	G	R	K	S	I	U
seqName	insertions	aaSubstitution	aaDeletions	missing	nonACGTNs	pcrPrimerCharAlignmentScore	al	
barcode09/AF28263:AACA	N:R203K,N:G2	ORF1a:S3675-	1-664,733,1001,1624-2205,2	ChinaCDC_N_				89601
barcode15/AF27897-27901	M:I82T,N:Q41S:Y144-		1-342,705-965,1313-1595,1	Pasteur_IP2_1				89595
barcode86/AF28263:AACA	N:P80R,N:R20	ORF1a:S3675-	1-54,9539,10738-11022,144	ChinaCDC_N_				89525
barcode37/AF2T,T11078C,G	N:S2Y,N:R203K,N:G204R,N:		1-54,3903,5621-5894,7672-	ChinaCDC_N_				89613
barcode61/AF21992:ACT	N:T205I,N:Q241K,N:D402Y,		1-54,19484,29837-29903	ChinaCDC_N_				89491
barcode79/AF21765-21770,2	M:I82T,N:M2:ORF1a:S3675-		1-54,24981-24982,29837-29	ChinaCDC_N_				89491

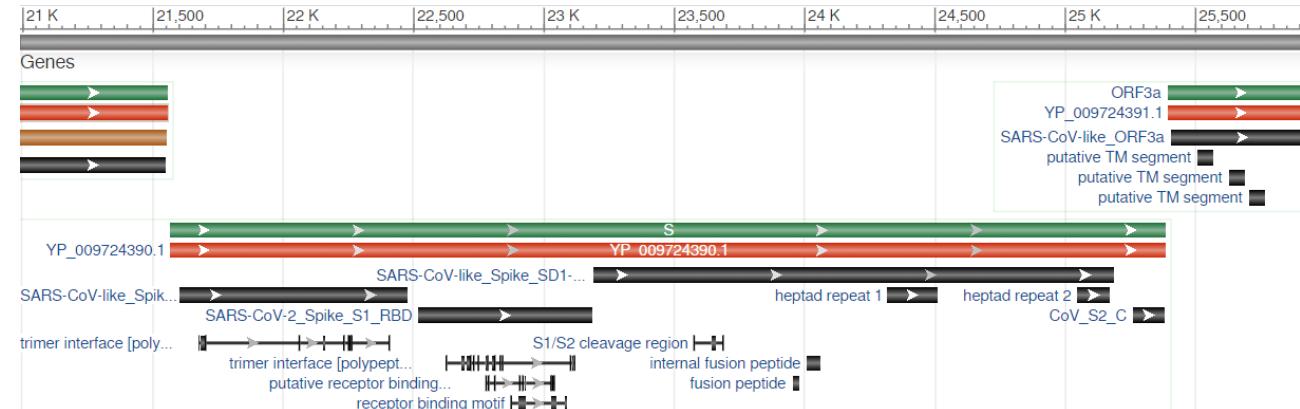
<https://clades.nextstrain.org/>



Vigilar sitios de interés en Spike y otros ORF

Mutaciones en RBD spike afectan la interacción con el receptor ACE2

Cambios en otros ORF y cambios en marco de lectura podrían tener efectos novedosos



<https://www.ncbi.nlm.nih.gov/gene/43740568>

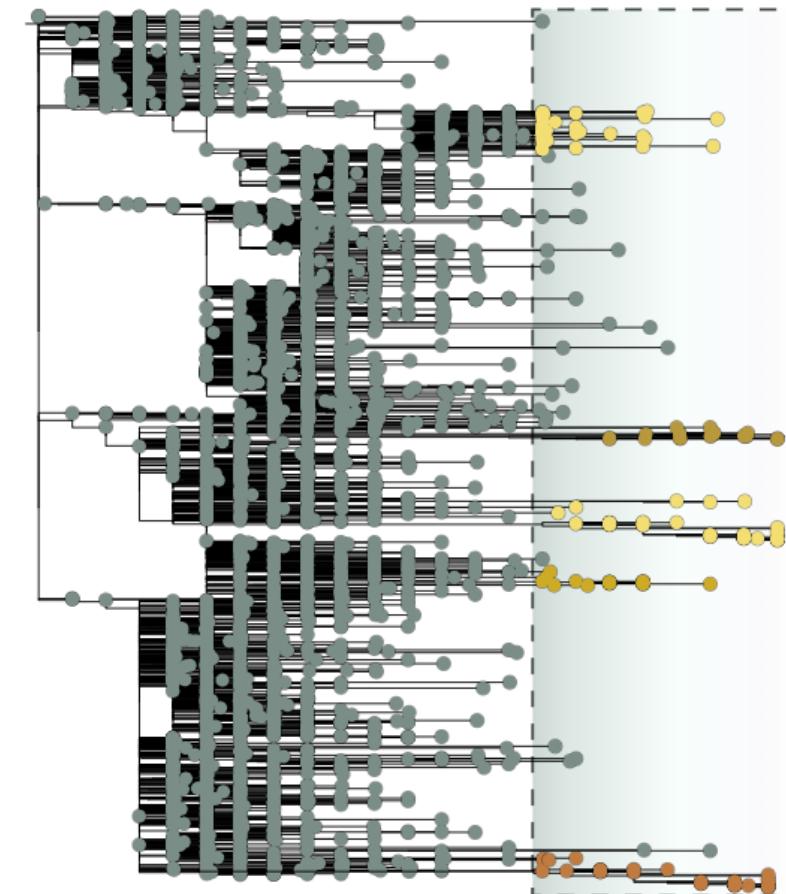


Capturar la frontera emergente de la pandemia

Vigilar las mutaciones que emergen en linajes que circulan actualmente. ¿Podrían tener consecuencia en su fenotipo?

Encontrar posibles linajes nuevos

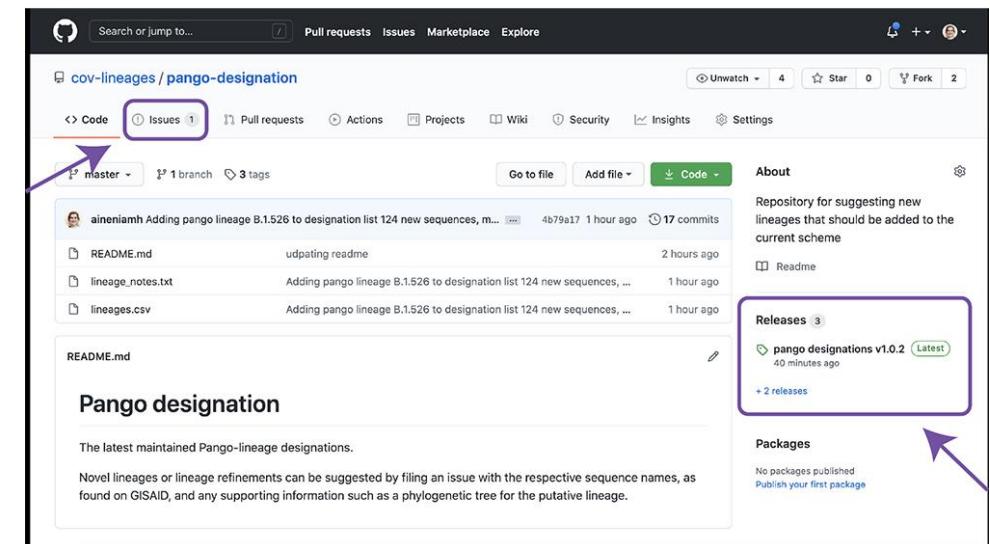
Vigilar el surgimiento de linajes y variantes de interés o preocupación importados



Reportando nuevos linajes

Linajes:

- Grupo de infección con una ancestría común
- Relevancia epidemiológica, nueva ubicación, rápido incremento de casos, nuevos fenotipos



<https://github.com/cov-lineages/pango-designation/issues>



Más herramientas



<https://nextstrain.org/ncov/gisaid/global>

covSPECTRUM

Enabled by data from **GISAID**

<https://cov-spectrum.ethz.ch/explore/World/AllSamples/AllTimes>



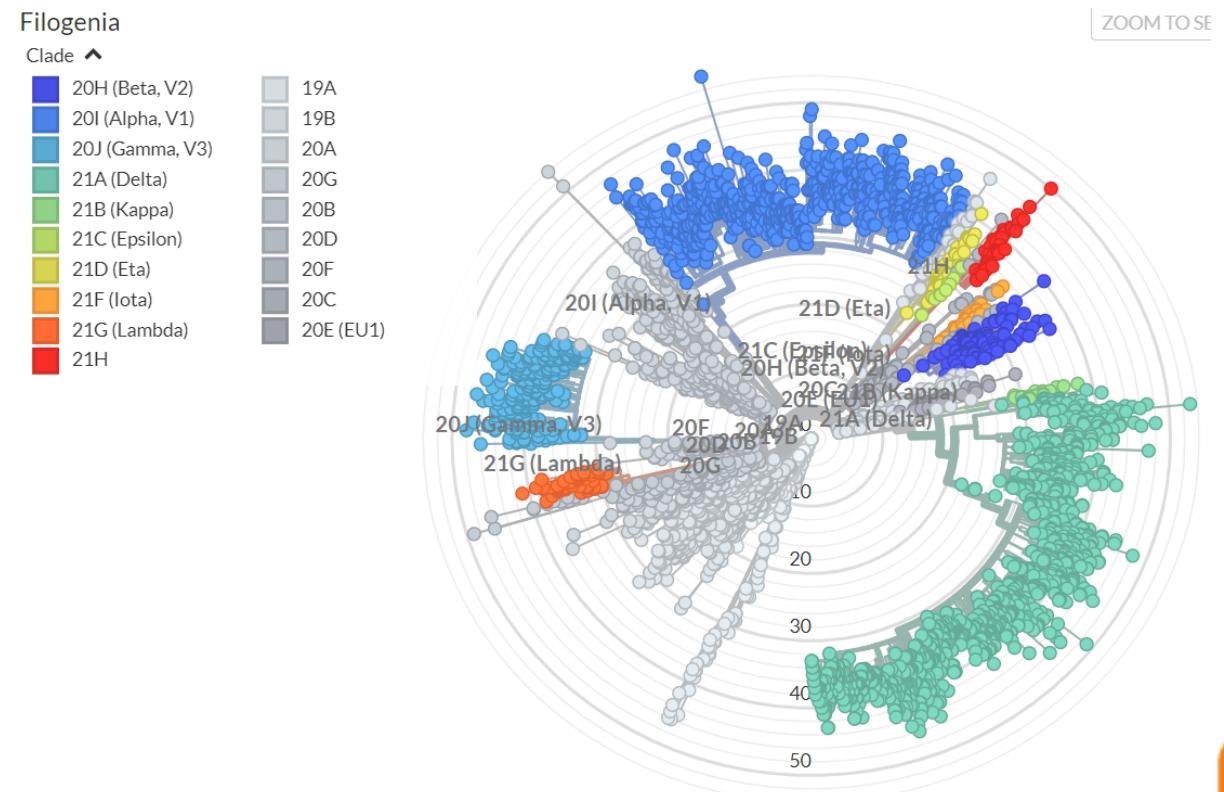
Capturar la frontera emergente de la pandemia

Mutaciones emergentes

Linajes nuevos

Surgimiento de variantes de interés o preocupación

Importación de VOCs y VOIs



<https://nextstrain.org/ncov/gisaid/global>



INNS



Investiga



Coordina



Vigila



Observa



Produce



Capacita