

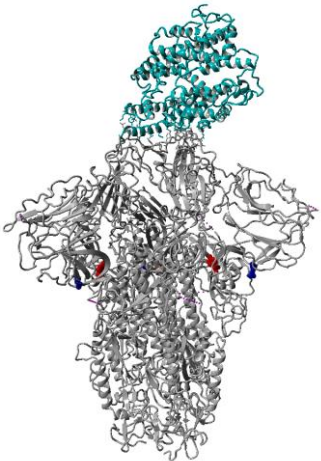
Real-time Surveillance of Pathogen Evolution and Global Transmission

Dr. Sebastian Maurer-Stroh, ED (D) BII

Bioinformatics Institute, A*STAR

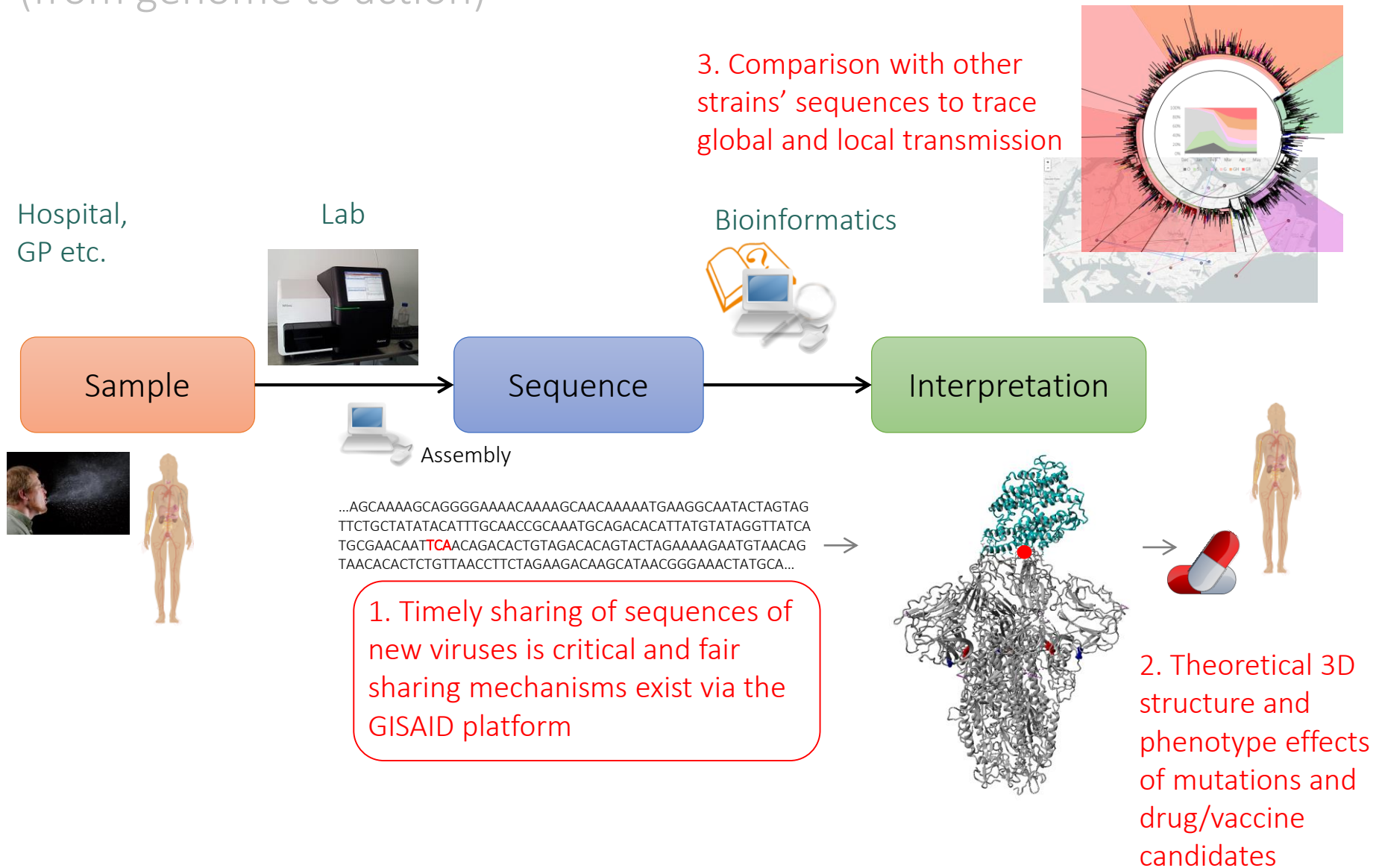
ID Labs, A*STAR

GISAID



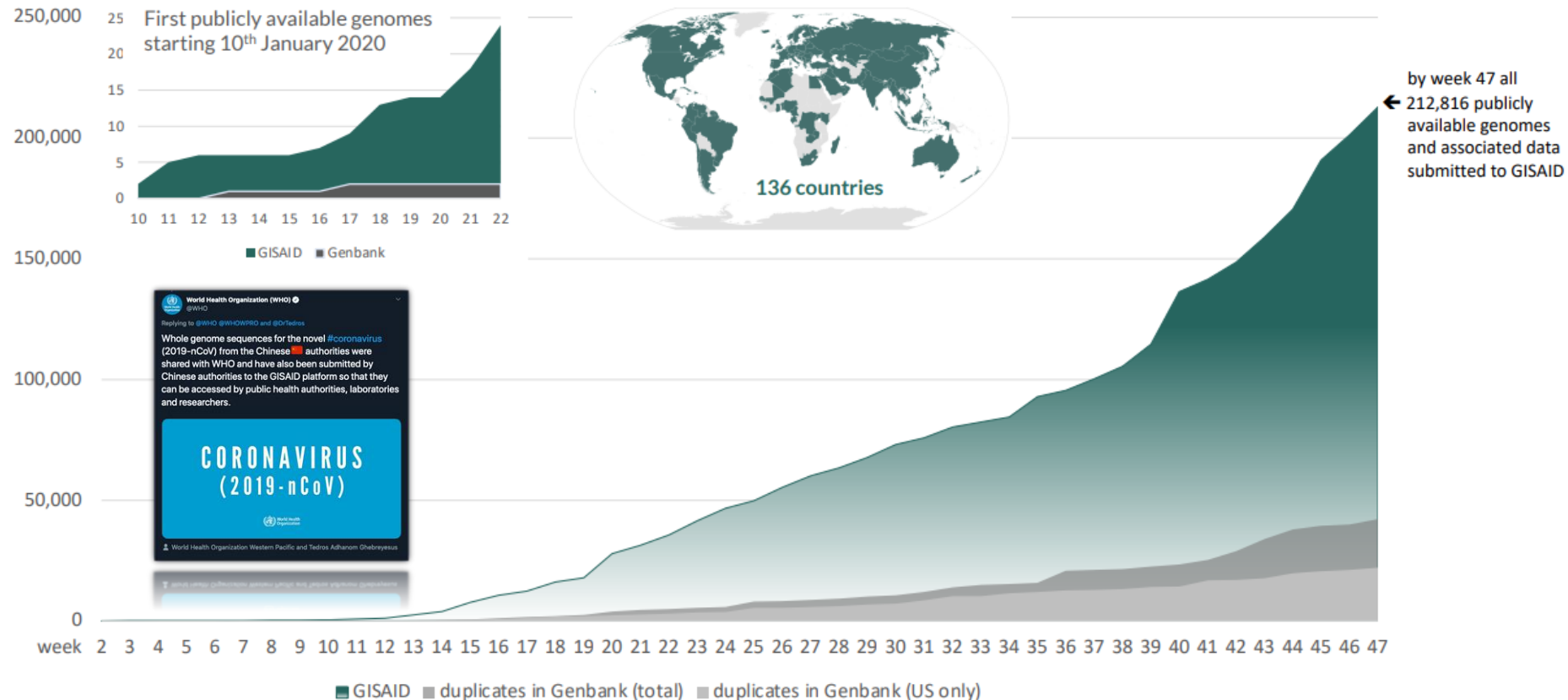
Real-time Surveillance of Pathogen Evolution and Global Transmission

(from genome to action)



Real-time data sharing is not achieved by governmental Regulations

... it is incentivized by the confidence in transparent sharing mechanisms

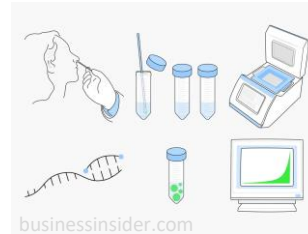


“GISAID was a well-oiled machine when the Coronavirus hit”
BBC and PRI the World

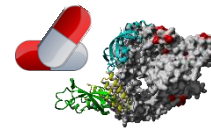


Genome use for first response

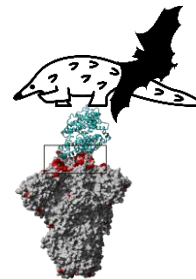
- How can you detect it?
- How can you treat it?
- Where did it come from?



Enabling first PCR and antigen-based diagnostic kits



Repurposed and new drugs, e.g. Remdesivir, mAbs, Vaccine candidates

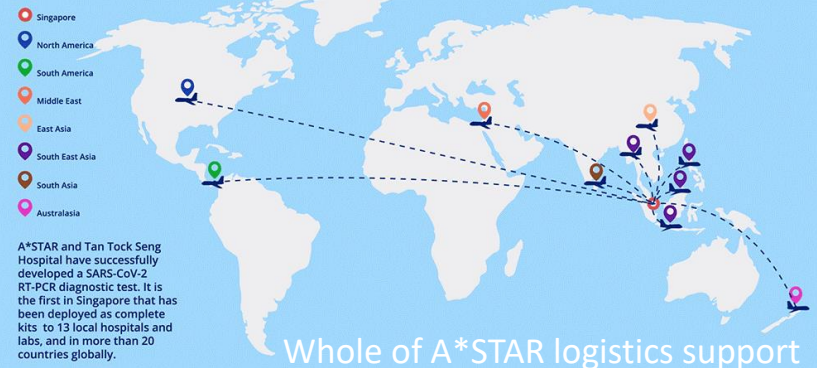


Likely animal source, understanding animal to human jump

Team Fortitude – an ongoing collaboration from SARS to flu to MERS to ZIKA to COVID

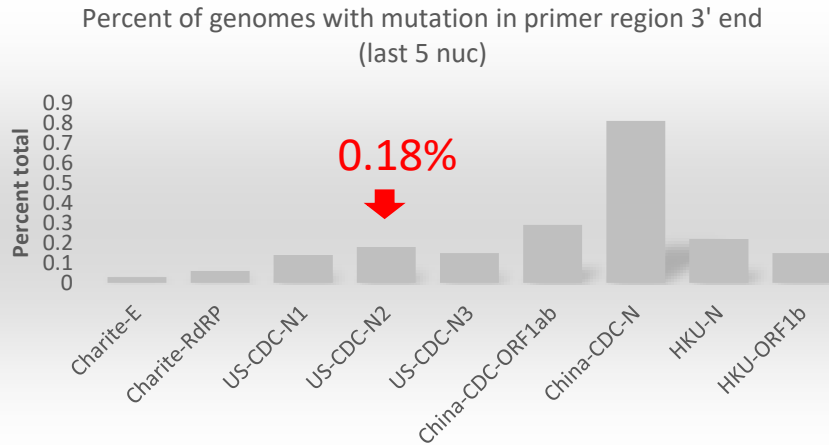


From Singapore to the world: Where Fortitude diagnostic kits have been deployed to fight COVID-19

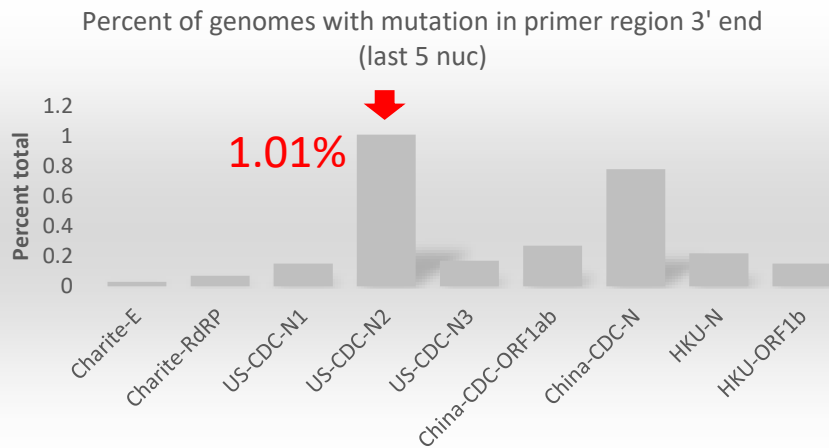


Real-time Surveillance of Primer Diagnostic Kit Mutations

Aug 7th



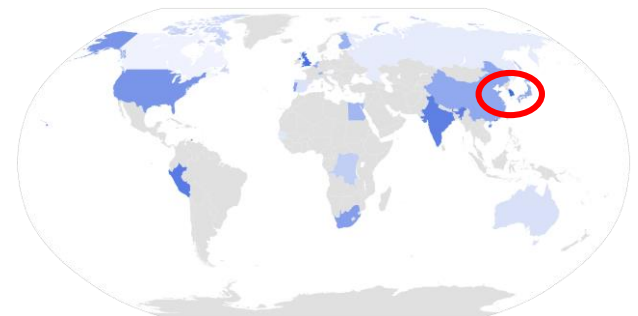
Aug 21st



PCR-based (RNA amplification)

Target regions need to be specific (not conserved with other viruses) but also evolutionary stable

Mutations can occur and primers/probes may need to be updated



Sources of primer sequences (partially updated in meantime):

<https://www.who.int/docs/default-source/coronaviruse/protocol-v2-1.pdf>
<https://www.who.int/docs/default-source/coronaviruse/peiris-protocol-16-1-20.pdf>
http://ivdc.chinacdc.cn/kyjz/202001/t20200121_211337.html
<https://www.who.int/docs/default-source/coronaviruse/uscdcr-pcr-panel-primer-probes.pdf>

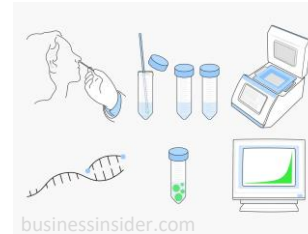
by BII/GIS, A*STAR Singapore



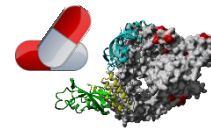
Genome use for first response



- How can you detect it?
- How can you treat it?



Enabling first PCR and antigen-based diagnostic kits



Repurposed and new drugs, e.g. Remdesivir, mAbs, Vaccine candidates

The NEW ENGLAND JOURNAL of MEDICINE

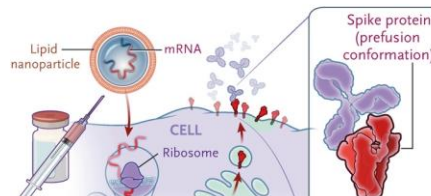
RESEARCH SUMMARY

Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine

F.P. Polack, et al. DOI: 10.1056/NEJMoa2034577

CLINICAL PROBLEM

Safe and effective vaccines to prevent severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection and Covid-19 are urgently needed. No vaccines that protect against betacoronaviruses are currently available, and mRNA-based vaccines have not been widely tested.

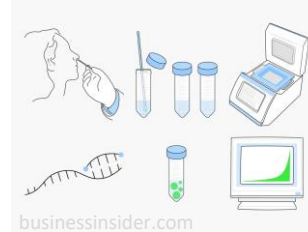


The Food and Drug Administration (FDA) announced Dec 11th it has issued an emergency use authorization (EUA) for the Pfizer-BioNTech COVID-19 vaccine, allowing shipments to begin and marking a significant milestone in the battle against the virus.

investment of resources. The development of BNT162b2 was initiated on January 10, 2020, when the SARS-CoV-2 genetic sequence was released by the Chinese Center for Disease Control and Prevention and disseminated globally by the GISAID (Global Initiative on Sharing All Influenza Data) initiative.

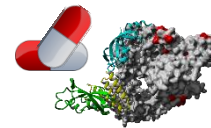
Genome use for first response

- How can you detect it?



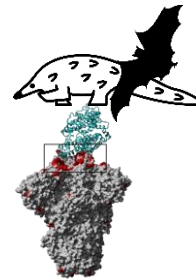
Enabling first PCR and antigen-based diagnostic kits

- How can you treat it?



Repurposed and new drugs, e.g. Remdesivir, mAbs, Vaccine candidates

- Where did it come from?

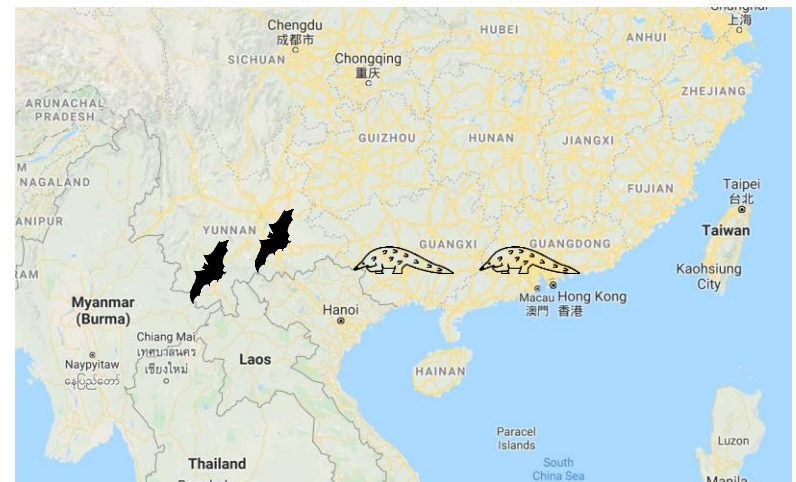


Likely animal source, understanding animal to human jump

Genome can point to zoonotic source

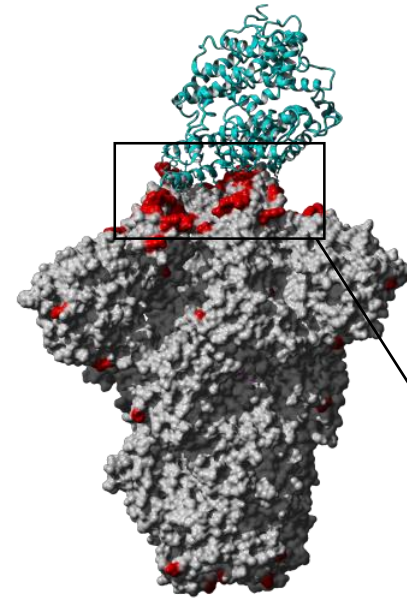


Light Orange ... previous bat CoVs
 Orange ... previous closest bat precursor (Yunnan 2013)
 Red ... new bat CoVs (Yunnan 2019)
 Light blue ... hCoV-19 2019-2020
 Green ... pangolin CoV (Southern China 2019)
 Blue ... SARS CoV



Southern China 2019: precursors in 2 species

How can a virus jump from animals to humans?



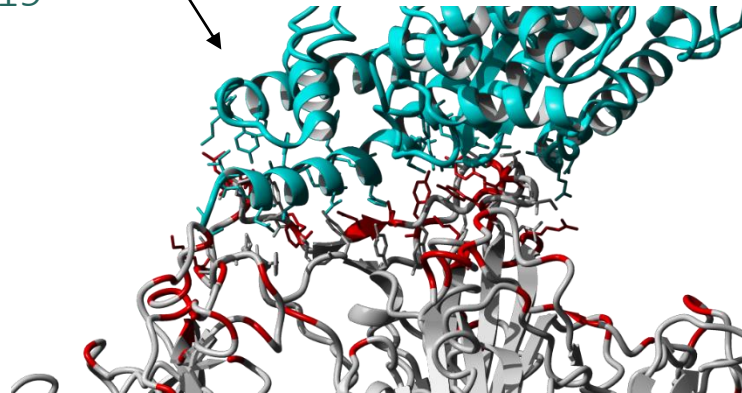
Cyan ... ACE2 human host receptor

Gray ... CoV spike glycoprotein

Red ... mutations between bat precursor
RaTG13 vs human outbreak WIV04 CoV

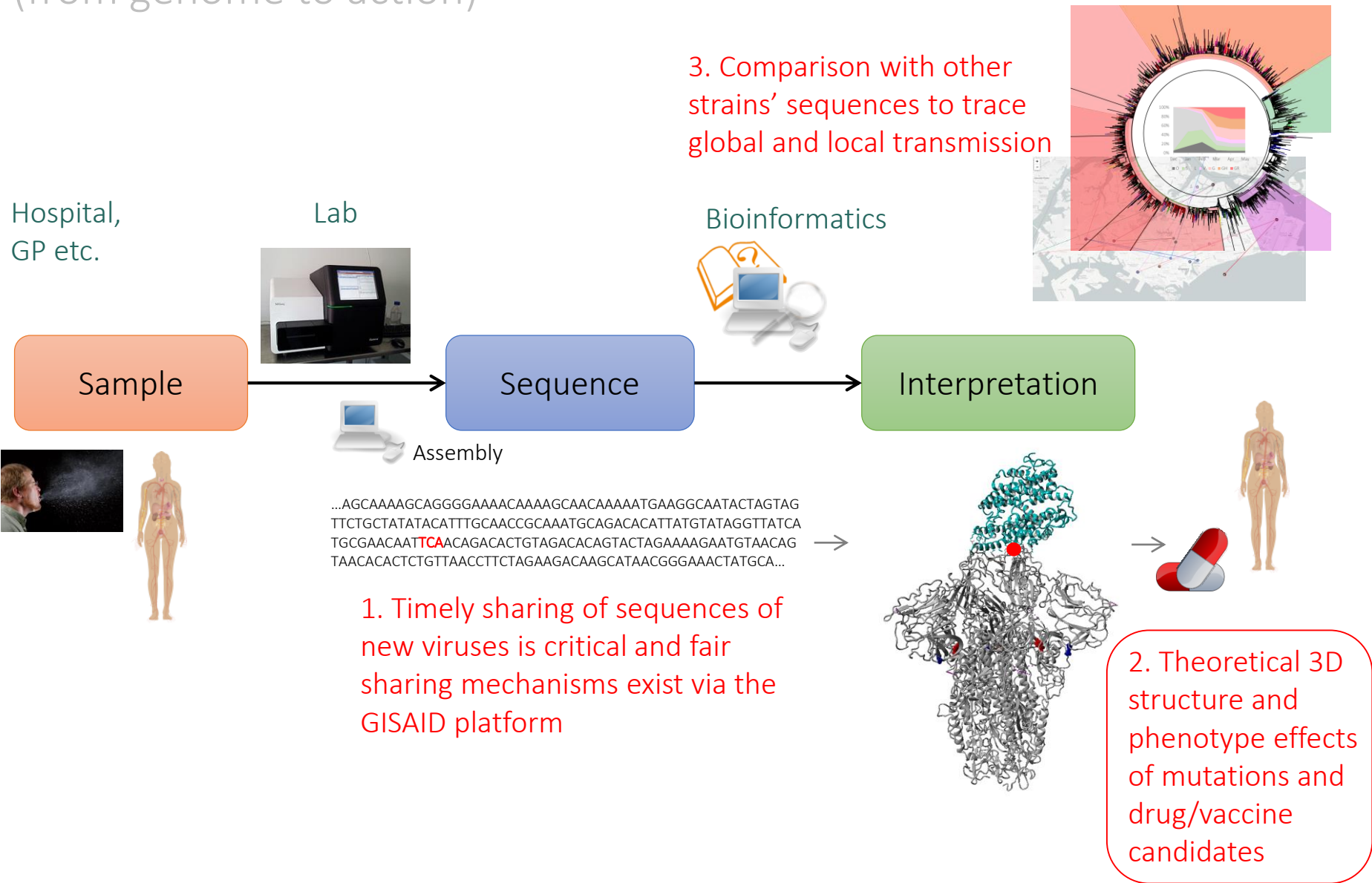
bat RaTG13 vs hCoV-19

- Surface proteins are 98% identical
- Bat precursor differences in receptor binding interface indicative of changes that allowed host switch



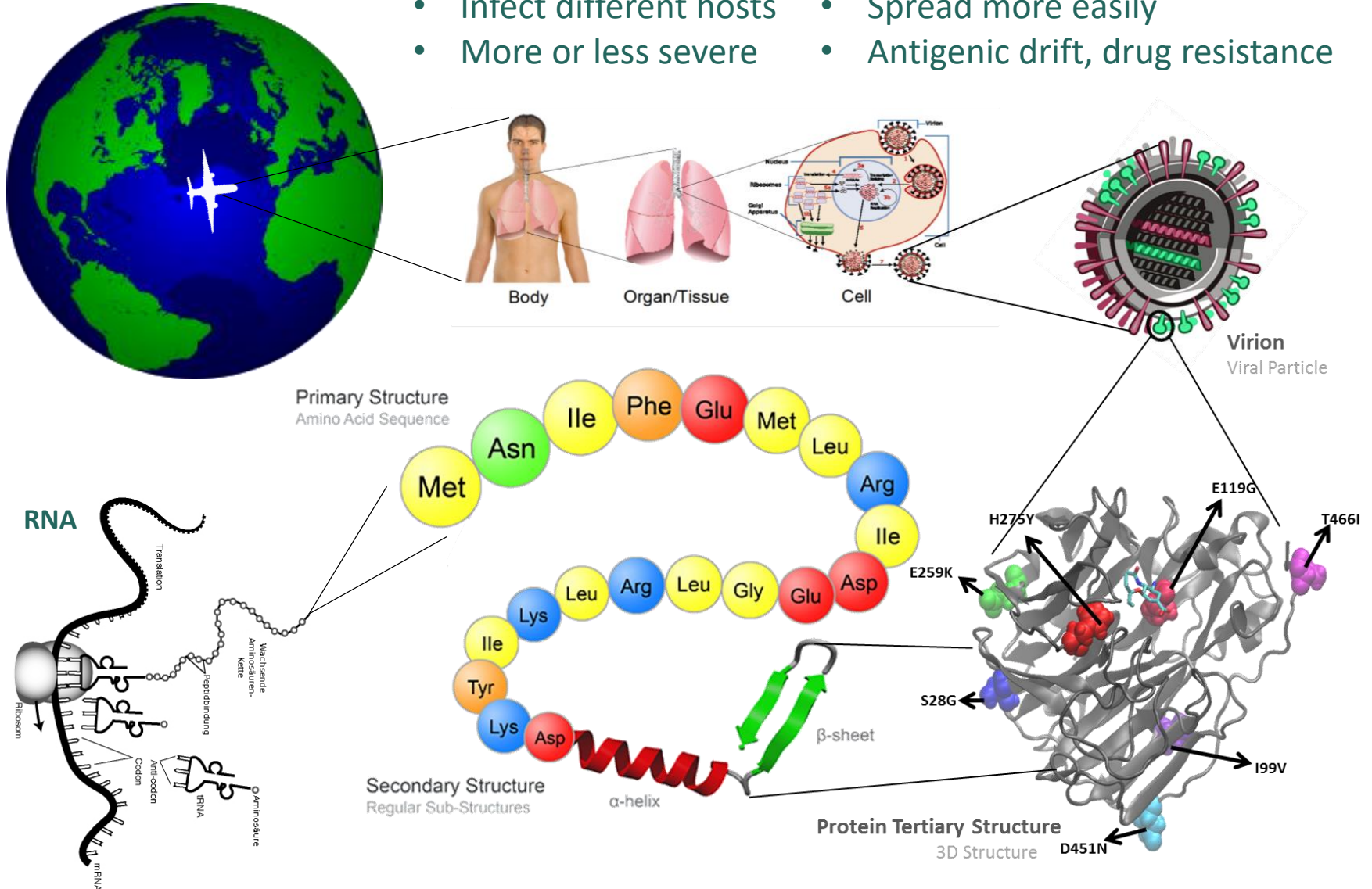
Real-time Surveillance of Pathogen Evolution and Global Transmission

(from genome to action)



From the **sequence** and **structure** we can partially deduce important **properties** of the virus

- Infect different hosts
- Spread more easily
- More or less severe
- Antigenic drift, drug resistance



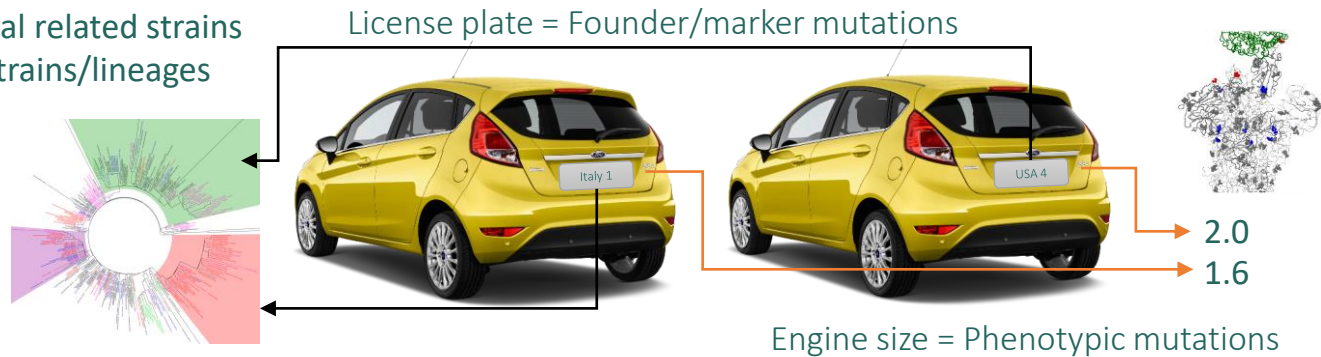
The virus is mutating into different strains and what does this mean?

All viruses naturally mutate resulting in many different but **highly similar variants** (also called **strains**) which most often do not cause any change in severity but this can be used to track movement of the virus. **Virus mutations can also be seen as changes creating a barcode of a virus allowing its identification.**

Strain = 1 virus barcode

Lineage = several related strains

Clade = many strains/lineages



Mutant virus strains are highly similar and can be imagined like this:

- same car model, same car color, **different license plate = no change in severity** but can be used to track virus (clusters, imports, transmission)
- same car model, same car color, **different engine size = possible change in severity**

Estimated frequency:

>99.99%

<0.01%

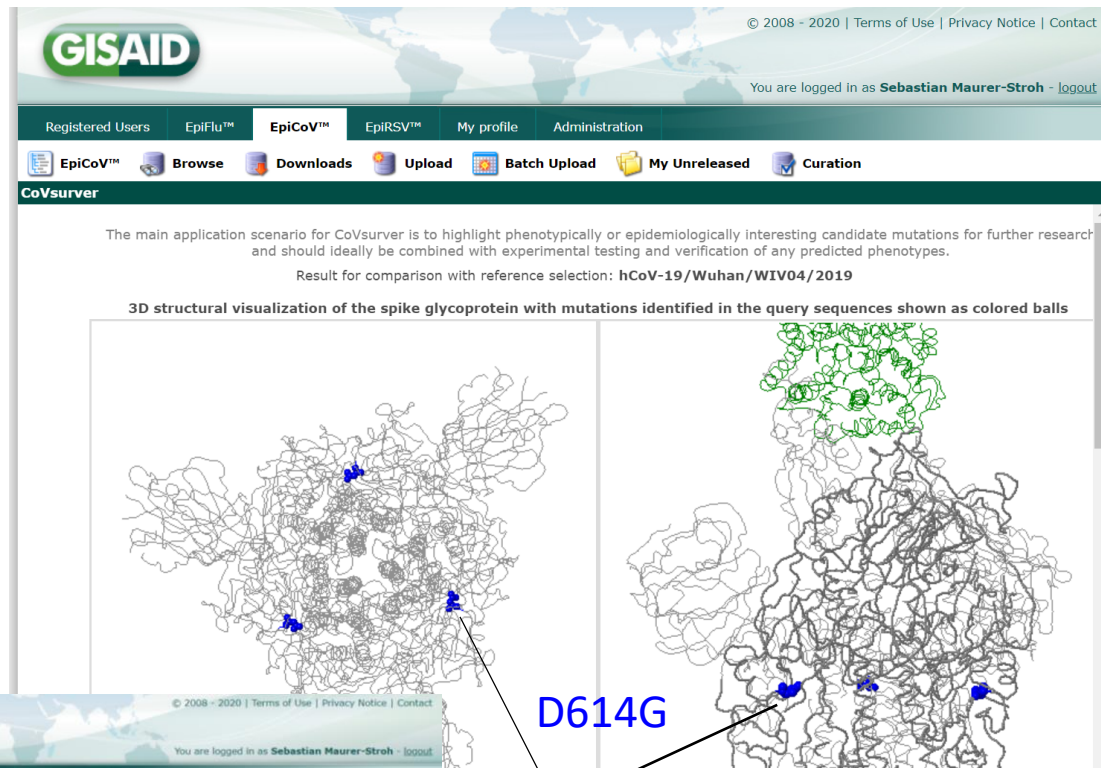
So far no mutation linked to severity identified, S-D614G not linked to severity

CoVsurver:

New analysis tool released on GISAID developed by BII A*STAR

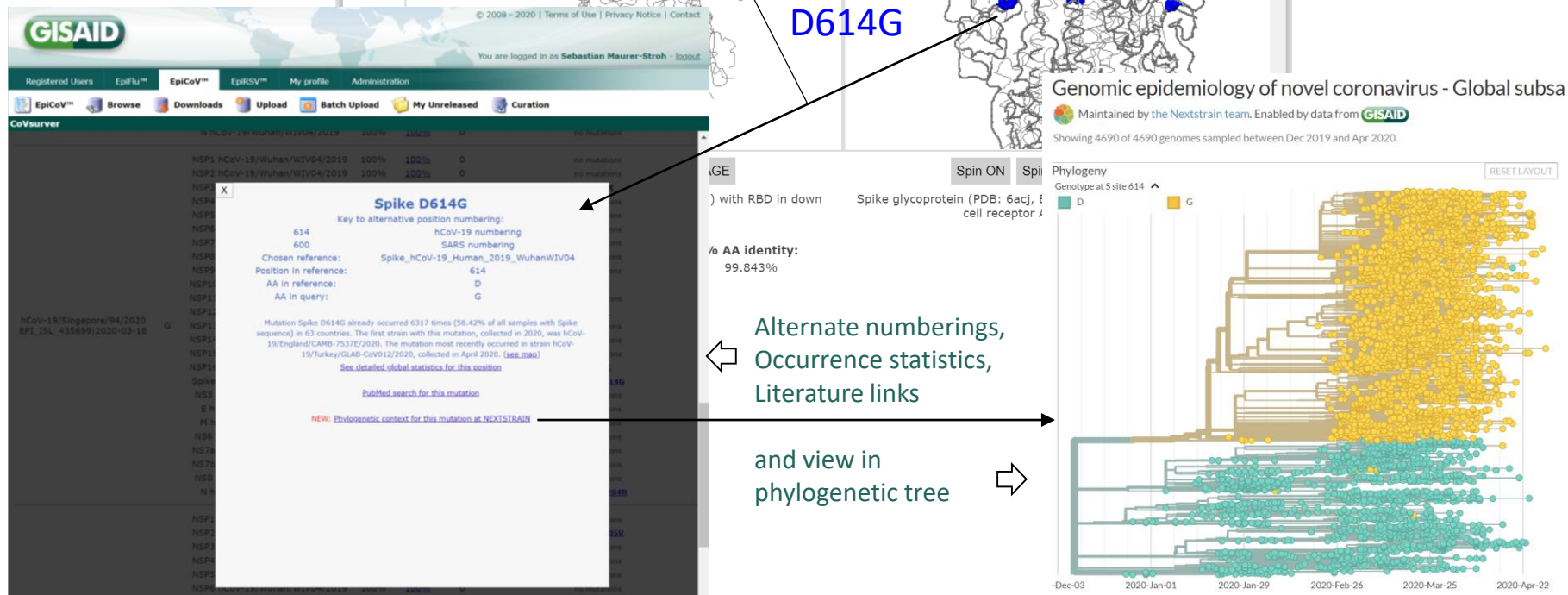
Capabilities:

Identification and interpretation of mutations including 3D structure and database with curated literature links for 166 mutation reports



Example D614G:

- Structural position not near receptor binding site, not in RBD antigenic sites, **at oligomer interface (can indirectly affect stability of complex)**
- Phylogenetic distribution points to founder-like effect and is accompanied with multiple other mutations

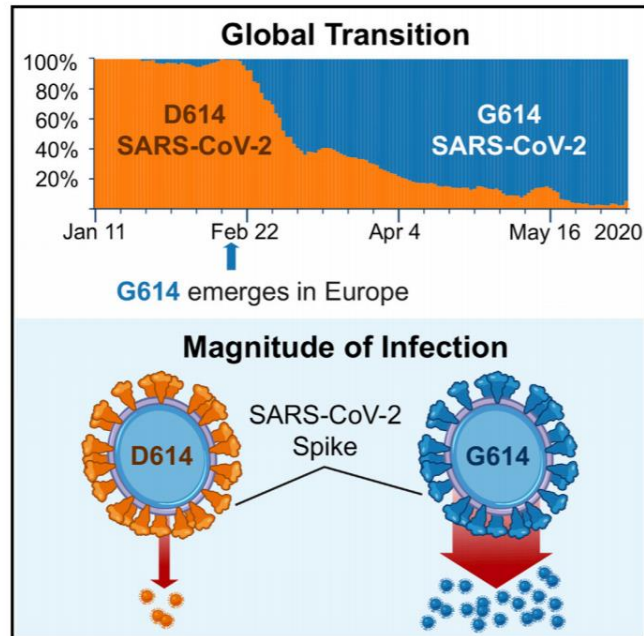


Spike-D614G: License plate or engine size mutation?

Cell

Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus

Graphical Abstract



Authors

Bette Korber, Will M. Fischer, Sandrasegaram Gnanakaran, ..., Celia C. LaBranche, Erica O. Saphire, David C. Montefiori

Correspondence

btk@lanl.gov

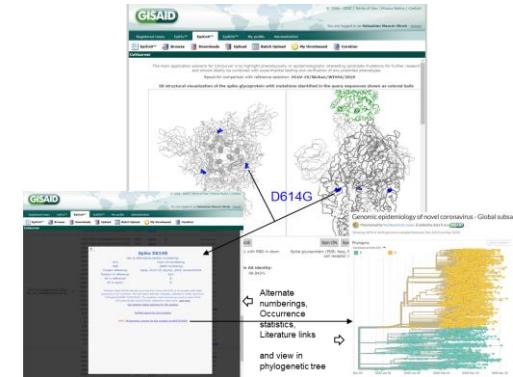
In Brief

Korber et al. present evidence that there are now more SARS-CoV-2 viruses circulating in the human population globally that have the G614 form of the Spike protein versus the D614 form that was originally identified from the first human cases in Wuhan, China. Follow-up studies show that patients infected with G614 shed more viral nucleic acid compared with those with D614, and G614-bearing viruses show significantly higher infectious titers *in vitro* than their D614 counterparts.

Article

CoVsurver:

New analysis tool released on GISAID developed by BII A*STAR



Evidence for/against founder effect (license plate):

- Variant present in 3+ major Chinese cities in Feb but did not cause big outbreak
- Europe where big increase occurred was not prepared to contain virus
- Spain is example where different virus was more successful (first import wins)
- Some regions where 614G outperformed the D variant (COG-UK paper)

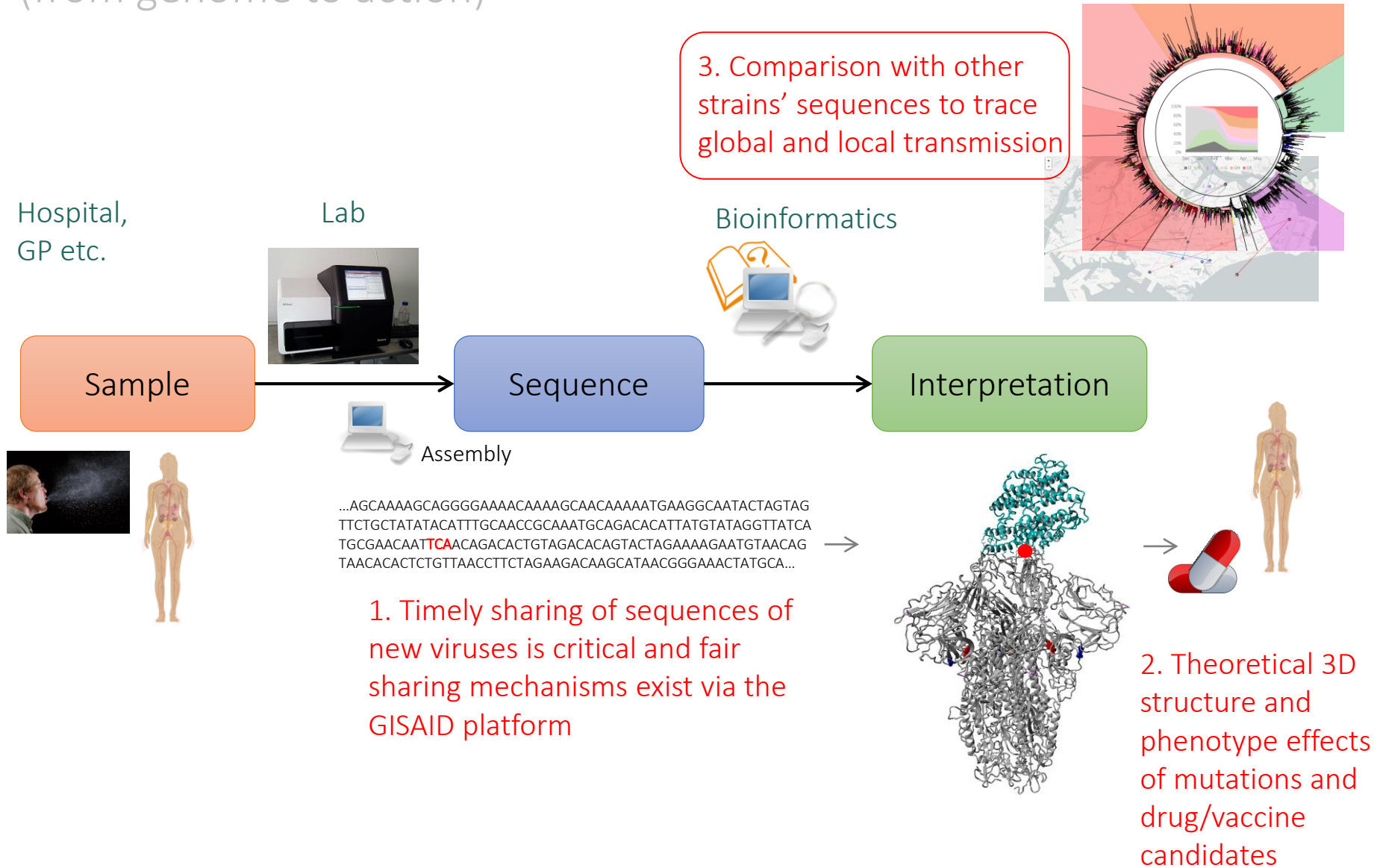
Evidence for/against phenotype (engine size):

- Increased infectivity in cells (higher titres *in vitro*)
- Increased spike stability (other paper)
- Hamster transmission (other paper)
- No clinical correlation with severity (also confirmed in SG data)

Verdict: Probably a bit of both... effect on cell infectivity/titre plausible, increased transmissibility claim partially disputed, consensus that there is **no correlation with clinical severity**

Real-time Surveillance of Pathogen Evolution and Global Transmission

(from genome to action)



January: Early outbreak genomes showed very low diversity, only a handful of mutations over 30,000 bases

2019-12-24	2019-12-30	2019-12-30	2019-12-30	2019-12-30	2019-12-30	2019-12-30	2019-12-30	2019-12-30	2019-12-30	2019-12-30	2019-12-30	2019-12-30	2019-12-30	2019-12-30	2020-01-01	2020-01-08	2020-01-13	2020-01-14	2020-01-15	2020-01-15	2020-01-15	2020-01-15	2020-01-16	2020-01-17	2020-01-17	2020-01-18
Wuhan/	Wuhan/	Wuhan/	Wuhan/	Wuhan/	Wuhan/	Wuhan/	Wuhan/	Wuhan/	Wuhan/	Wuhan/	Wuhan/	Wuhan/	Wuhan/	Wuhan/	Wuhan/	Nontha	Nontha	Guang	Guang	Guang	Guang	Zhejiang	Guang	Zhejiang	Guang	
IPBCA	HBCD	IPBCA	IPBCA	IPBCA	IPBCA	IPBCA	IPBCA	IPBCA	IPBCA	IPBCA	IPBCA	IPBCA	IPBCA	IPBCA	IPBCA	bur	bur	dong	dong	dong	dong	g/WZ	dong	g/WZ	dong	
MS-	C-HB-	MS-	MS-	MS-	MS-	MS-	MS-	MS-	MS-	MS-	MS-	MS-	MS-	MS-	MS-	bur	bur	0SF01	0SF01	0SF01	0SF01	g/WZ	0SF02	g/WZ	0SF04	
WH-01	01	WH-02	WH-03	WH-04	WH-04	WH-04	WH-04	WH-04	WH-04	WH-04	WH-04	WH-04	WH-04	WH-04	WH-04	bur	bur	2	3	4	5	01	8	02	0	

2019-12-24 Wuhan/IPBCAMS-WH-01

2019-12-30 Wuhan/HBCDC-HB-01

2019-12-30 Wuhan/IPBCAMS-WH-02

2019-12-30 Wuhan/IPBCAMS-WH-03

2019-12-30 Wuhan/IPBCAMS-WH-04

2019-12-30 Wuhan/IVDC-HB-01

2019-12-30 Wuhan/IVDC-HB-05

2019-12-30 Wuhan/WIV02

2019-12-30 Wuhan/WIV04

2019-12-30 Wuhan/WIV05

2019-12-30 Wuhan/WIV06

2019-12-30 Wuhan/WIV07

2020-01-01 Wuhan/IVDC-HB-04

2020-01-08 Nonthaburi/61

2020-01-13 Nonthaburi/74

2020-01-14 Guangdong/20SF012

2020-01-15 Guangdong/20SF013

2020-01-15 Guangdong/20SF014

2020-01-15 Guangdong/20SF025

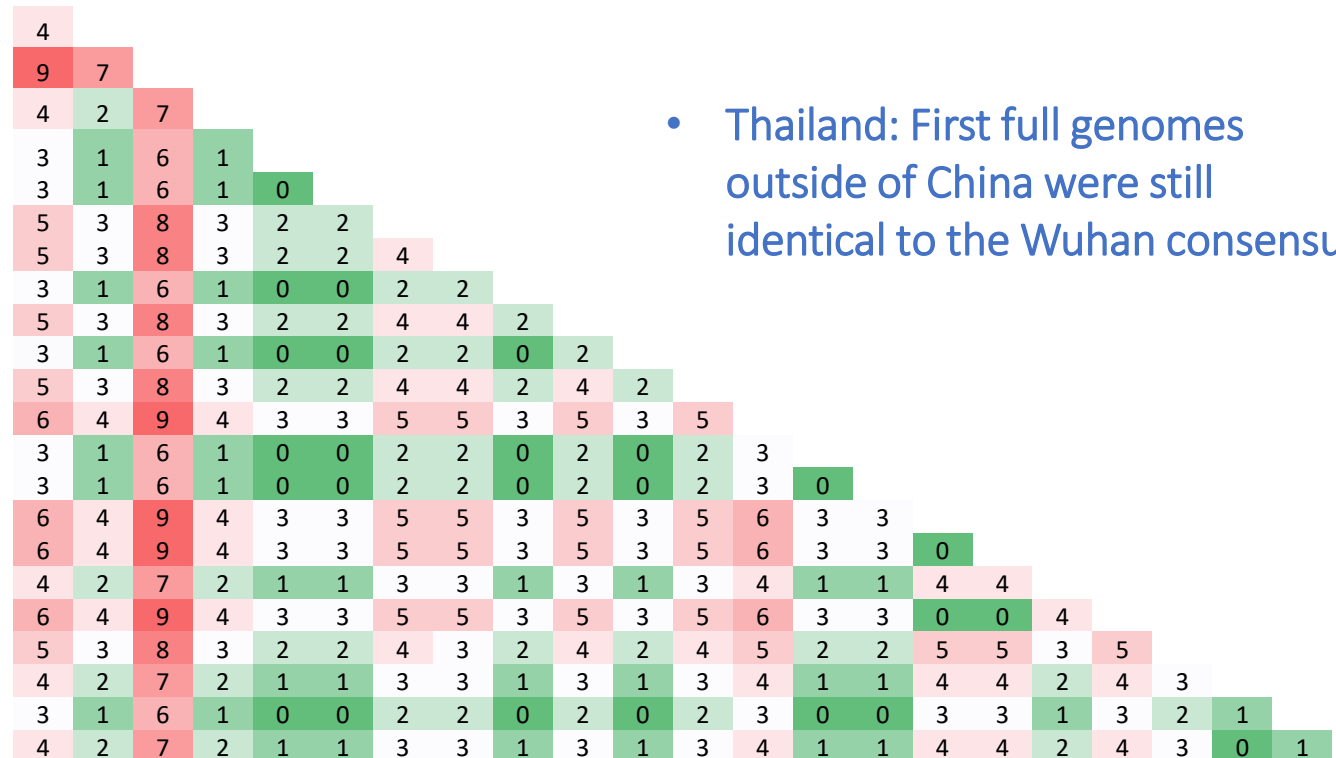
2020-01-16 Zhejiang/WZ-01

2020-01-17 Guangdong/20SF028

2020-01-17 Zhejiang/WZ-02

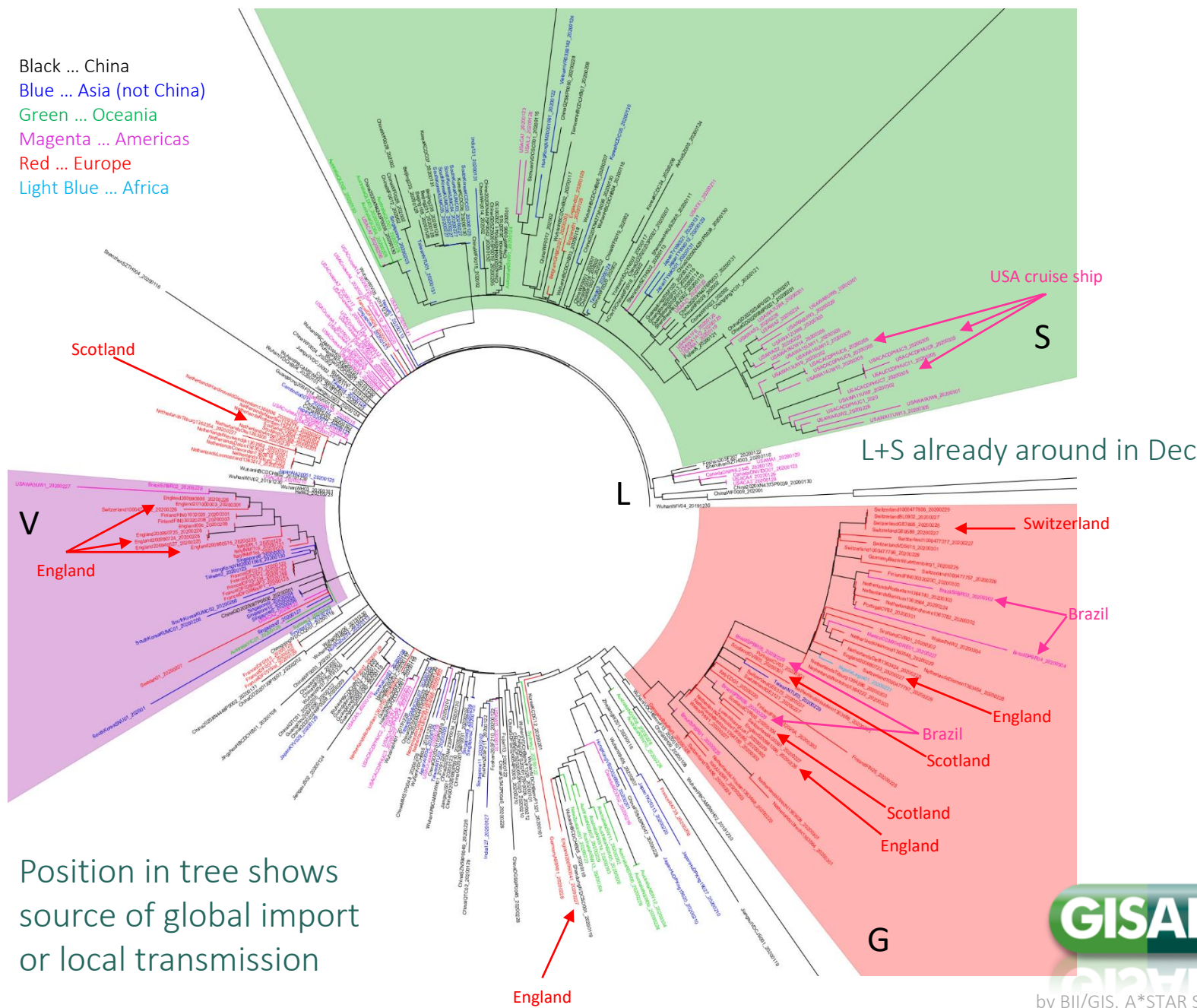
2020-01-18 Guangdong/20SF040

- Thailand: First full genomes outside of China were still identical to the Wuhan consensus



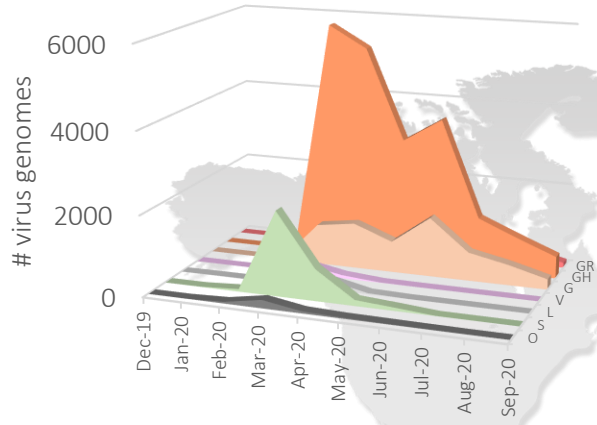
Nucleotide (base) differences among early outbreak strains

Feb/March: Split into genetic groups (clades)

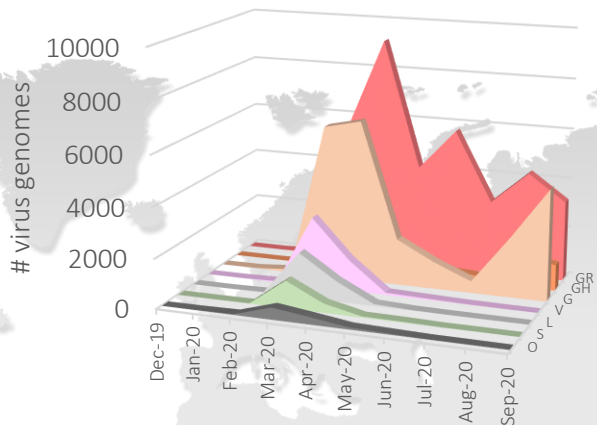


October: Temporal and regional distribution of clades

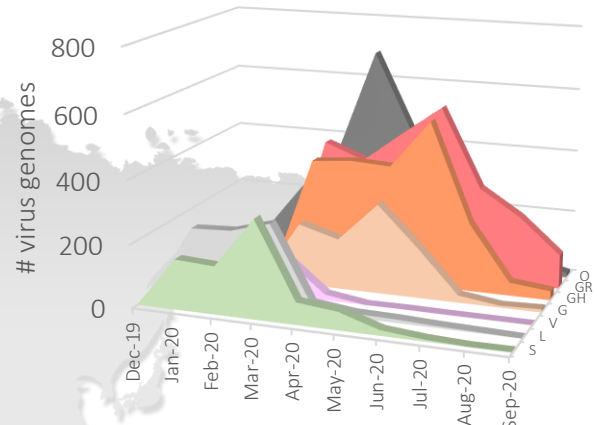
North America



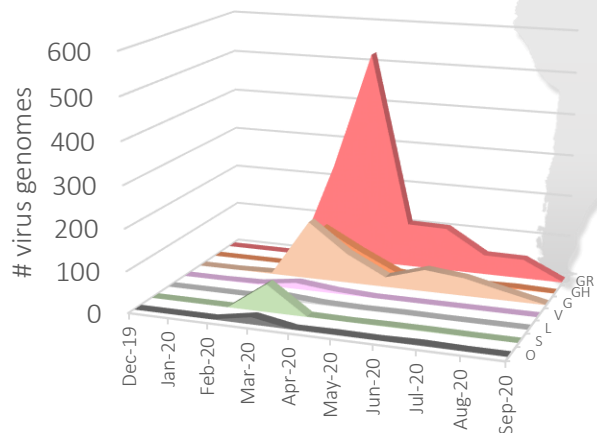
Europe



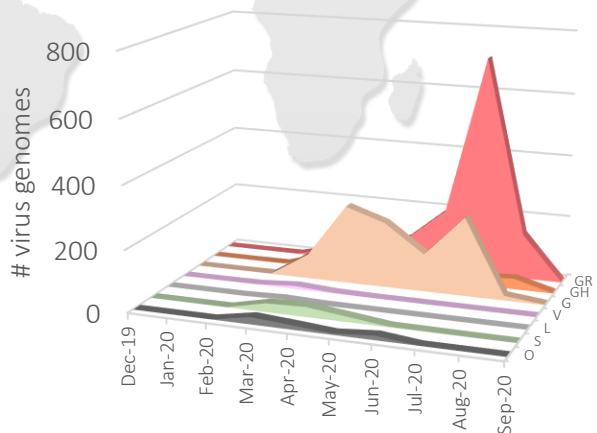
Asia



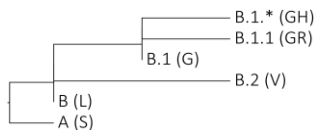
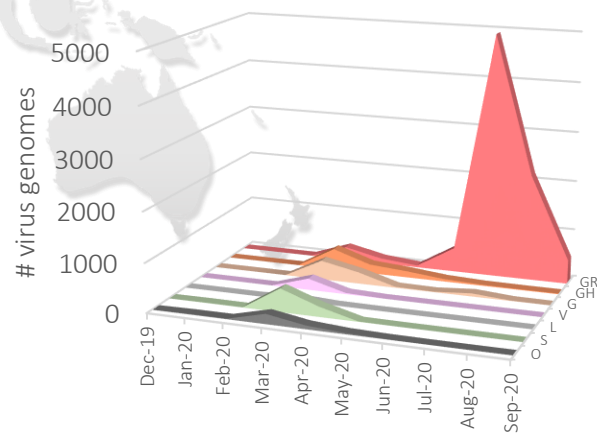
South America



Africa



Oceania



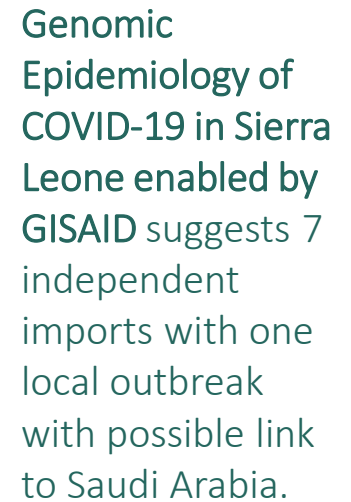
GISAID clades and PANGOLIN lineages

O S L V G GH GR

by BII/GIS, A*STAR Singapore

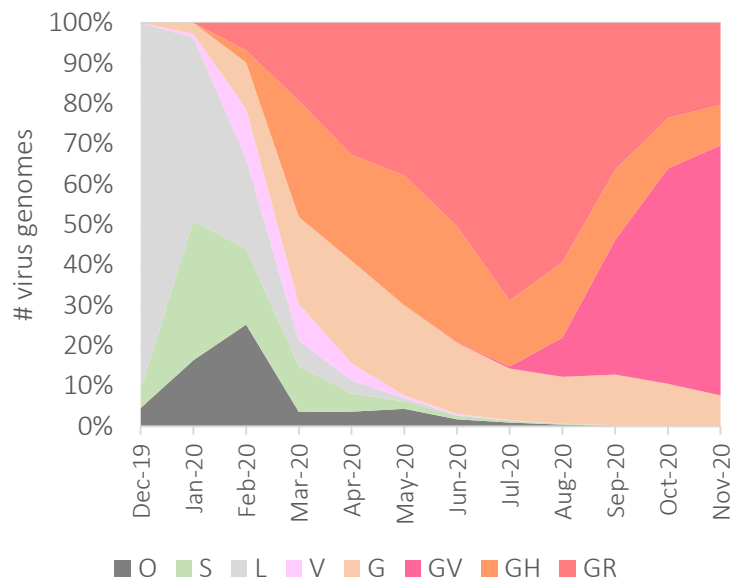


Genetic markers for contact tracing



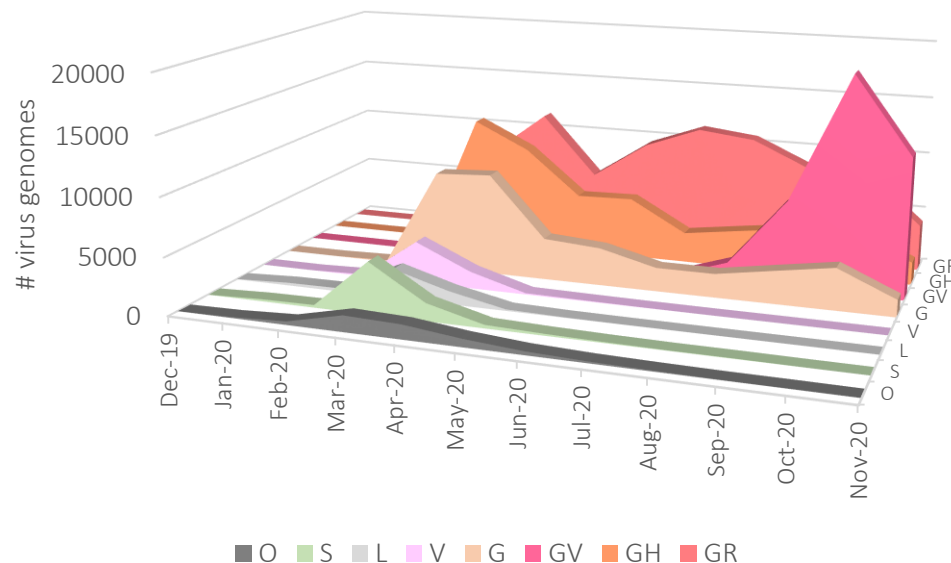
by BII/GIS, A*STAR Singapore

Global



Global to local

Global

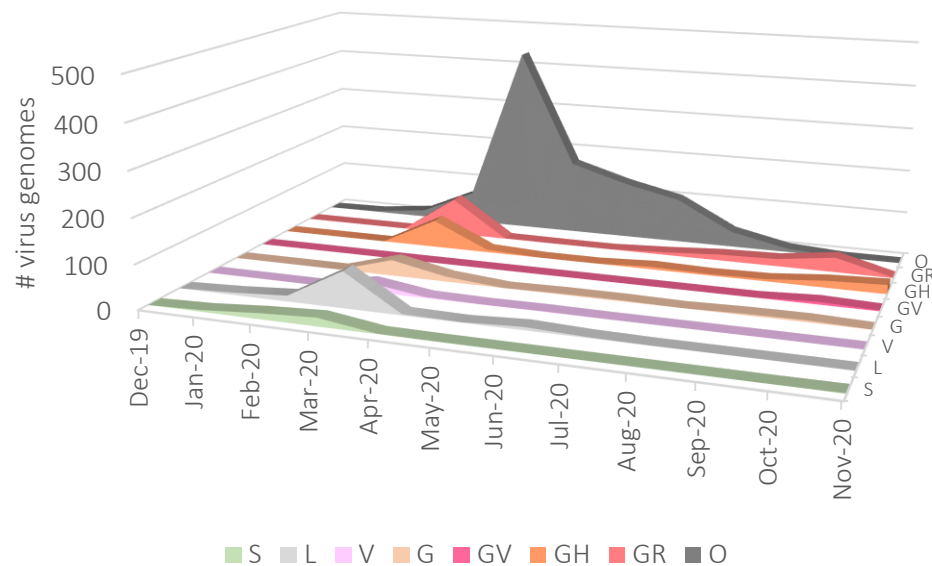


GV (B.1.177) hCoV-19 Spain CT-ISCIII-2013597 2020
 GR (B.1.1.1) hCoV-19 England 20168037604 2020 EPI IS
 GH (B.1.*) hCoV-19 Canada ON-PHL-8751 2020 EPI ISL 41834
 G (B.1) hCoV-19 Germany BY-ChVir-929 2020 EPI ISL 406862 2020
 V (B.2) hCoV-19 Italy LAZ-INMI-SPL1 2020 EPI ISL 412974 2020-01-;
 L (B) hCoV-19 Wuhan WIV04 2019 EPI ISL 402124 2019-12-30
 S (A) hCoV-19 Guangdong 20SF012 2020 EPI ISL 403932 2020-01-1.

H

0.000050

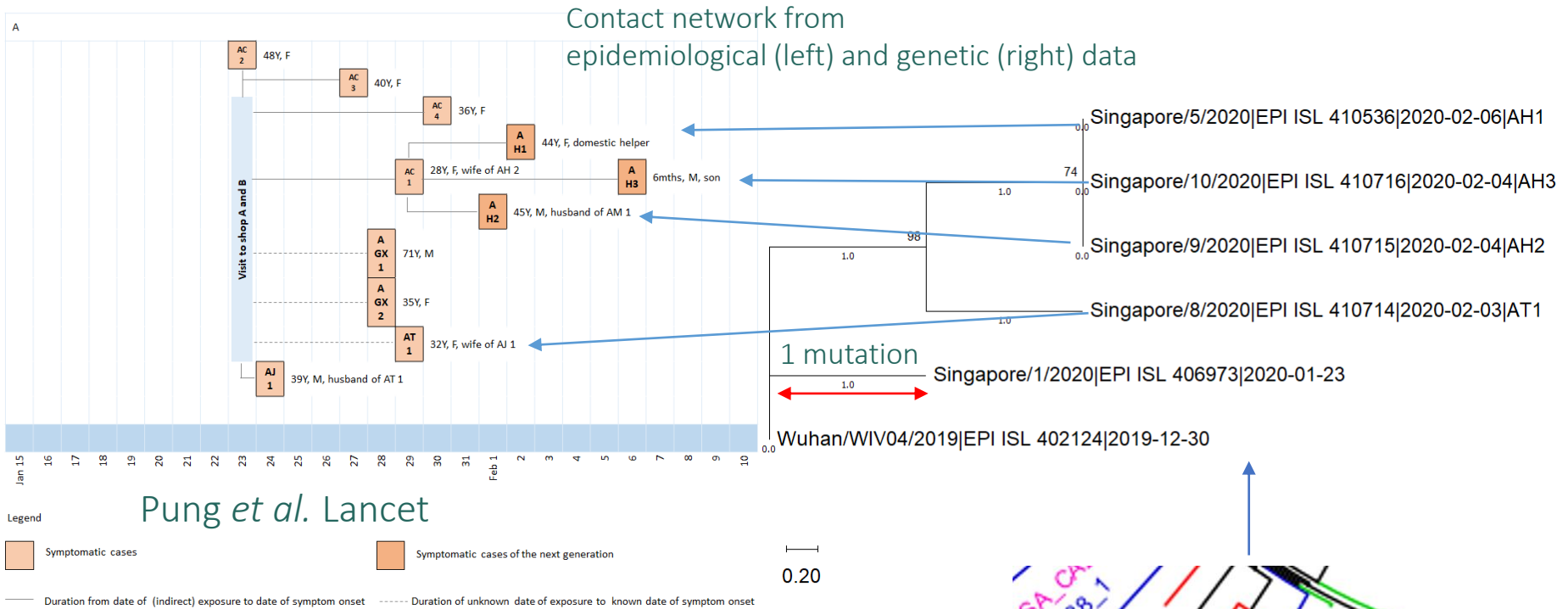
Singapore



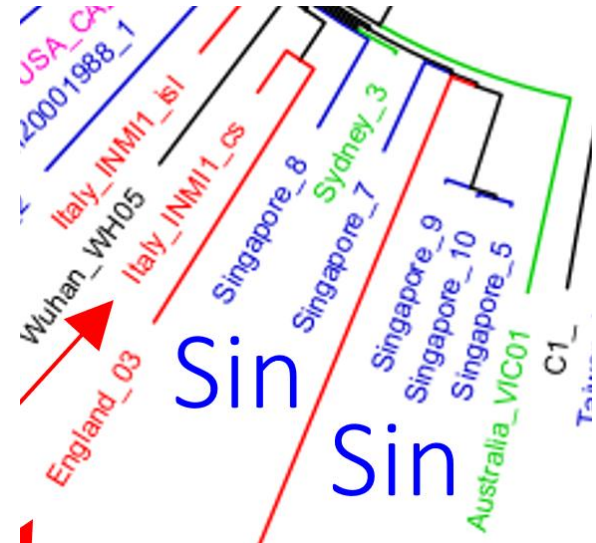
by BII/GIS, A*STAR Singapore



Genetic markers for contact tracing



Not all local sequences can be connected easily but linkage can be significant if they are identical or share unique mutations or deletions (SG 382bp deletion: Young *et al.*, Lancet). Full genomes needed for interpretation.

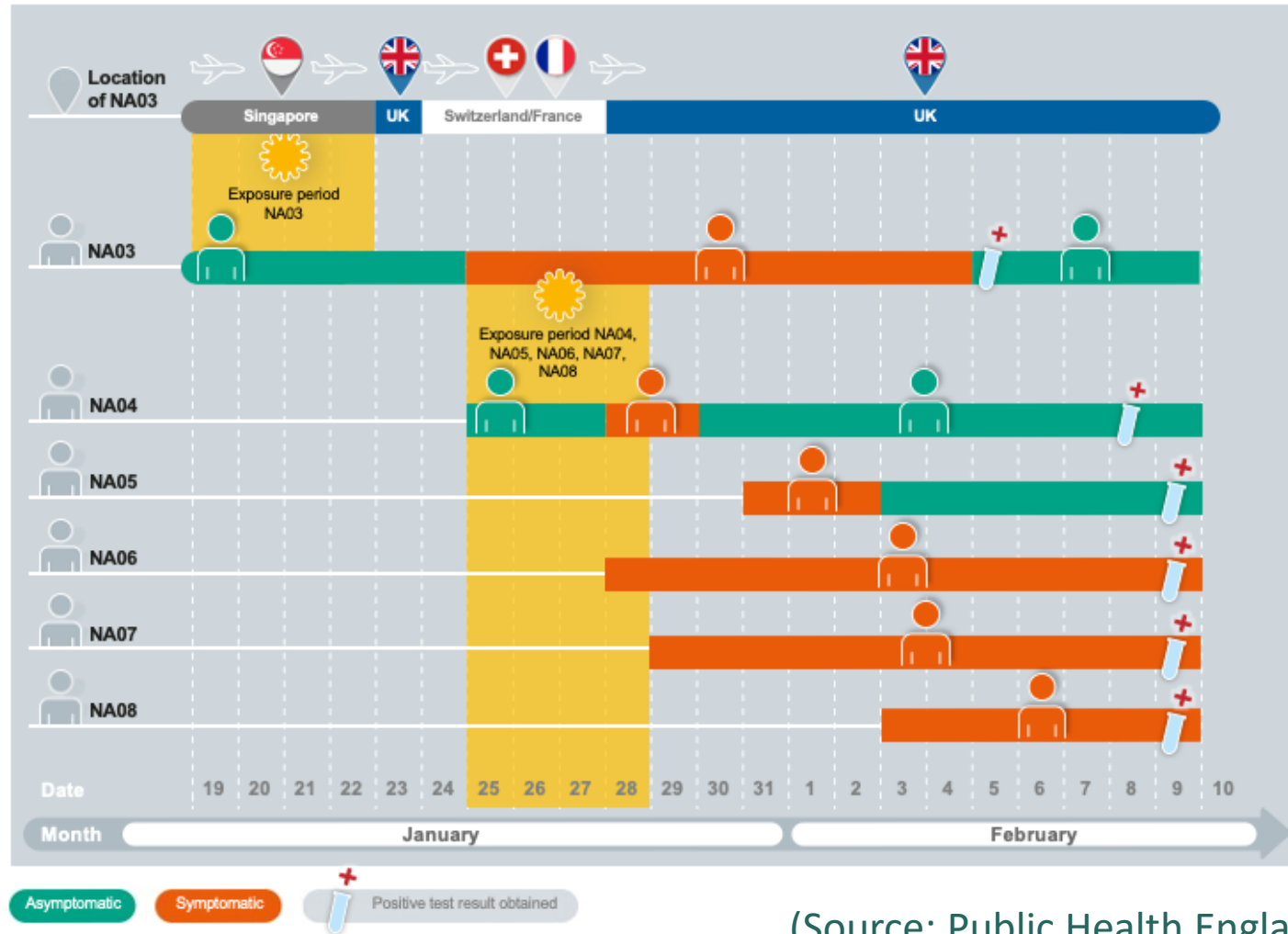


Genetic markers for contact tracing

One variant has a unique 3bp deletion in nsp2 that allows to follow a detailed transmission chain and spread of a single variant

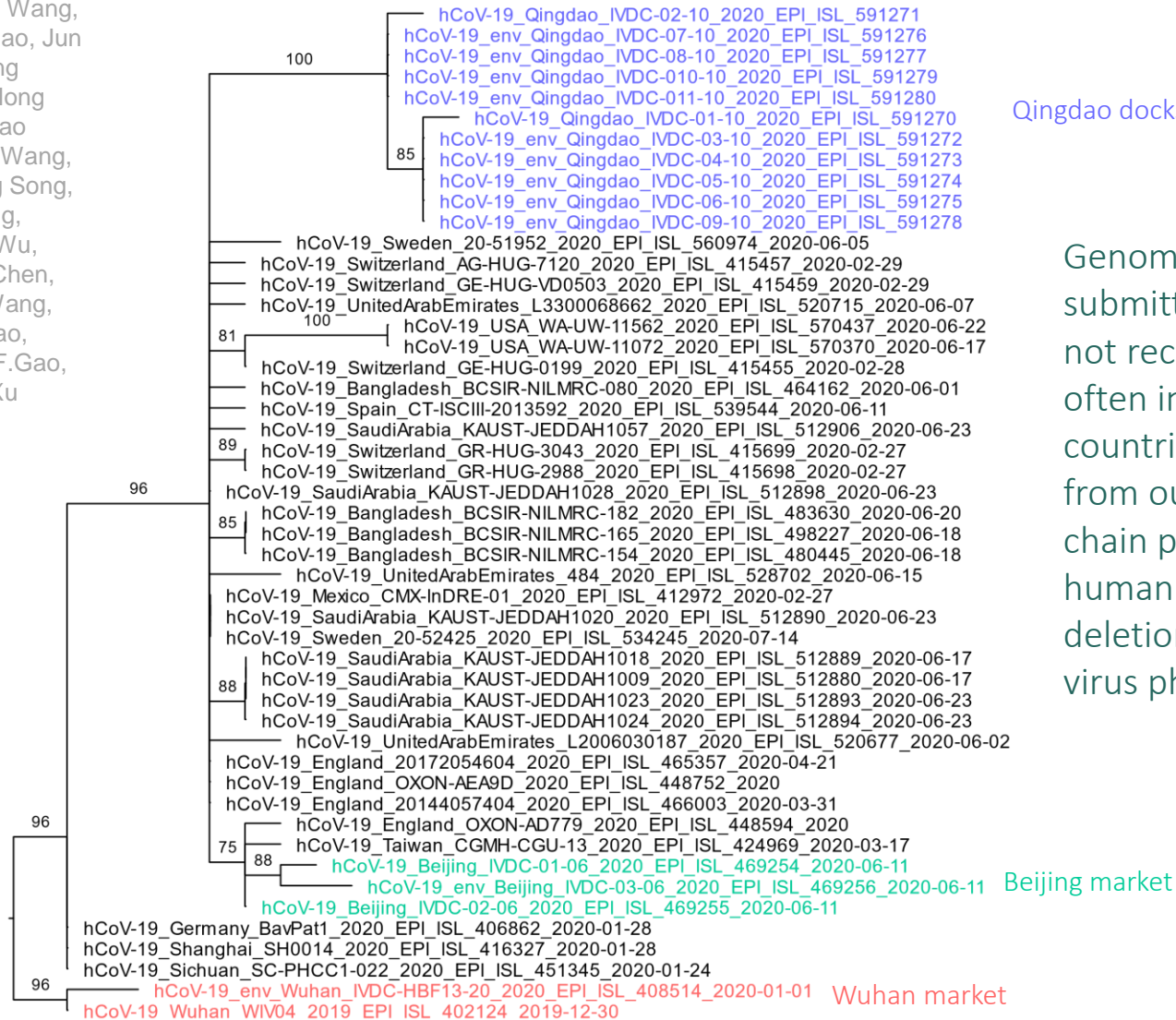
COVID-19 Transmission chain UK cluster 2 diagram

07.05.2020



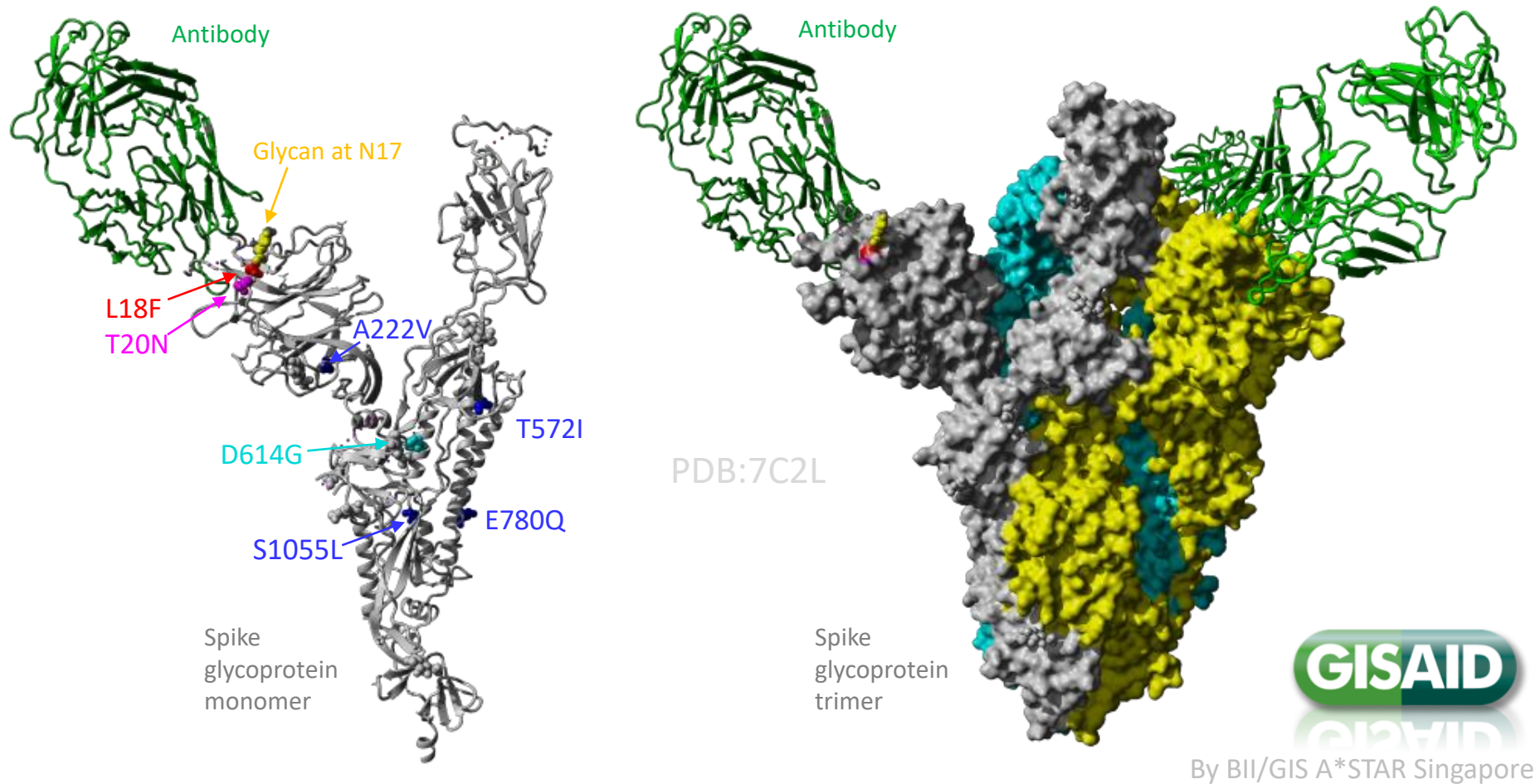
(Source: Public Health England)

Huilai Ma,
Zhaoguo Wang,
Xiang Zhao, Jun
Han, Yong
Zhang, Hong
Wang, Cao
Chen, Ji Wang,
Jingdong Song,
Yao Meng,
Yuchao Wu,
Zhixiao Chen,
Dayan Wang,
Ruqin Gao,
George F. Gao,
Wenbo Xu



Genomes from Qingdao submitted to GISAID. Genotype not recently seen in China but often in Europe and other countries. Samples taken from outer packaging of cold chain products match the human samples (share unique deletion marker not linked to virus phenotype).

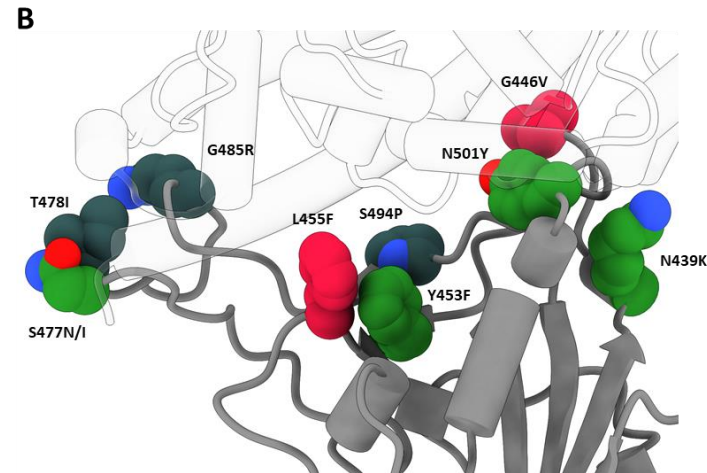
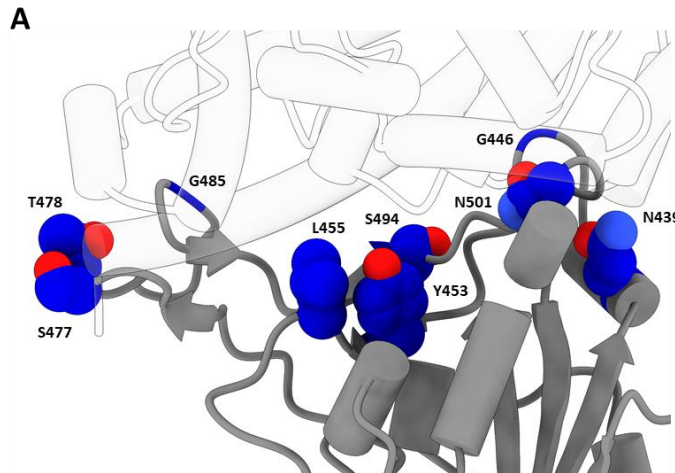
Mutations and potential vaccine/immunity escape: reinfection cases



2 of 4 reinfection cases have mutations possibly interfering with the structural conformation of glycosylation sites in a region that is also broadly recognized by antibodies which would provide a hypothetical mechanism for immune escape potentially contributing to permitting second infection. However, this doesn't apply to all cases and many other factors could play a role too. Importantly, these mutations are rare and occur sporadically without causing large clusters so far.

Real-time Surveillance of receptor binding pocket mutations

MS in preparation



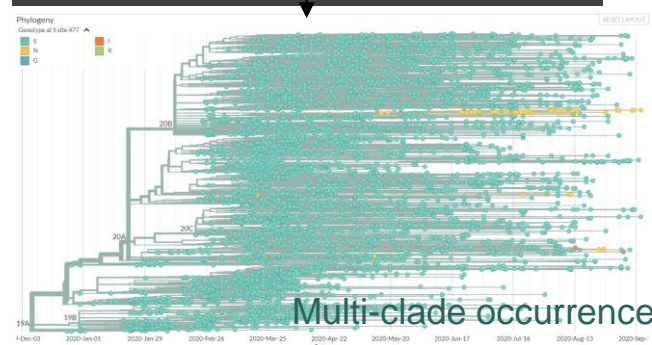
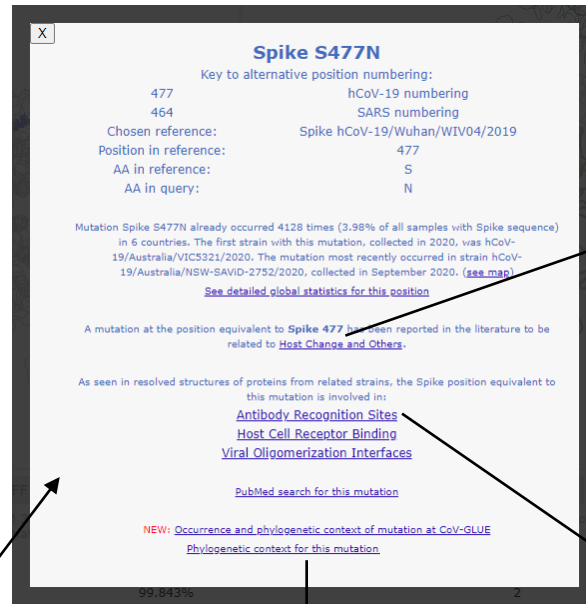
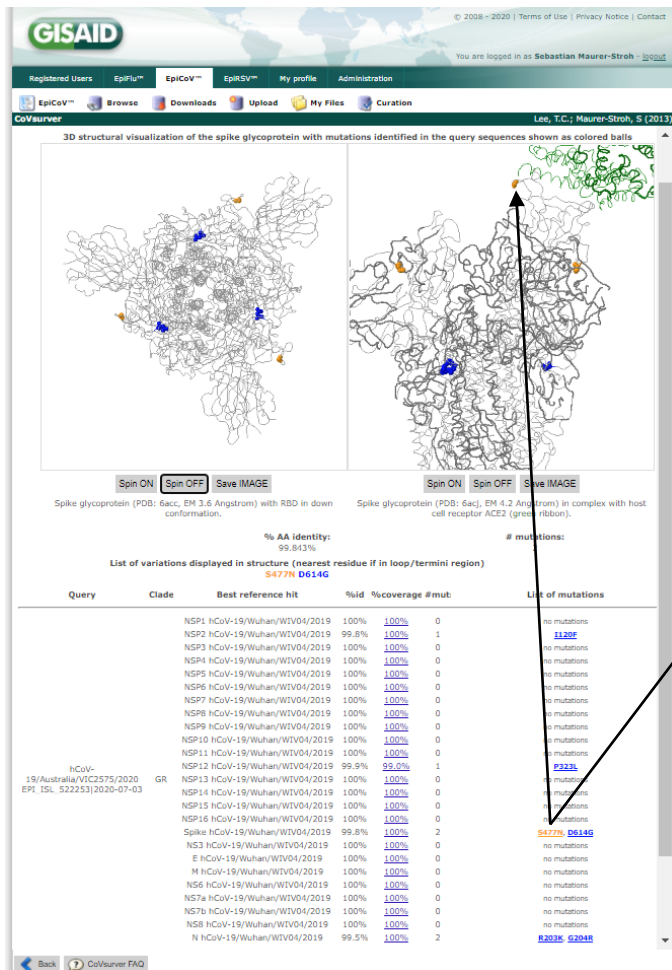
AA	Frequency	Location counts
S477N	12115	8427 Australia/VIC , 1300 England, 1032 Denmark, 425 Australia/NSW, 303 Switzerland, 153 France, 92 Belgium, 80 Netherlands, 66 Scotland, 44 Norway, 44 Australia/QLD, 37 Wales, 27 Australia/WA, 18 Sweden, 11 Northern Ireland, 9 USA/FL, 8 Australia/SAP, 5 Australia/ACT, 4 CotedIvoire, 4 Germany, 4 Slovakia, 3 New Zealand, 2 Australia/TAS, 2 Tunisia, 2 Australia/NT, 2 Spain, 1 Hong Kong, 1 Lebanon, 1 Canada, 1 Thailand, 1 Gibraltar, 1 USA/MA, 1 Singapore, 1 USA/CA, 1 USA/WA, 1 Japan, 1 India
N439K	3344	1277 England, 772 Denmark, 726 Scotland, 251 Wales, 140 Ireland, 47 Switzerland, 35 Northern Ireland, 26 Czech Republic, 16 France, 12 Norway, 11 Netherlands, 8 Faroe Islands, 5 Germany, 3 Italy, 3 Belgium, 3 USA/IL, 2 Australia/SAP, 1 Sweden, 1 New Zealand, 1 USA/WI, 1 Australia/NSW, 1 Singapore, 1 Australia/WA, 1 Romania
Y453F	343	329 Denmark, 6 Netherlands, 3 Switzerland, 2 South Africa, 1 Faroe Islands, 1 Russia, 1 USA/UT
N501Y	220	140 Wales, 37 Australia/VIC, 33 England, 5 USA/MA, 1 USA/OR, 1 USA/NY, 1 USA/IL, 1 Denmark, 1 Brazil
T478I	116	113 England, 1 Northern Ireland, 1 South Africa, 1 Spain
S477I	100	70 England, 11 USA/NM, 5 Wales, 3 USA/CA, 3 Scotland, 1 South Africa, 1 Singapore, 1 Indonesia, 1 USA/SC, 1 India, 1 Colombia, 1 Luxembourg, 1 Australia/VIC
S494P	89	34 England, 26 USA/CA, 7 USA/MA, 4 USA/MI, 3 Scotland, 2 Sweden, 2 India, 2 Denmark, 1 CotedIvoire, 1 Australia/VIC, 1 France, 1 Wales, 1 Singapore, 1 Nigeria, 1 USA/UT, 1 USA/MD, 1 Spain
G485R	37	37 Australia/VIC
L455F	36	17 Australia/VIC, 7 England, 3 Scotland, 2 South Africa, 2 South Korea, 2 USA/CA, 1 Italy, 1 USA/WA, 1 USA/MO
G446V	31	9 England, 5 Australia/VIC, 2 USA/VA, 2 USA/TX, 2 South Korea, 1 USA/MN, 1 Italy, 1 Scotland, 1 Australia/NSW, 1 Wales, 1 USA/CA, 1 USA/WA, 1 Israel, 1 Finland, 1 Portugal, 1 Denmark

Collaboration with DSO and VIDRL and Uni Melb



by BII/GIS, A*STAR/Singapore

CoVsurver tool to analyse mutations –spike S477N



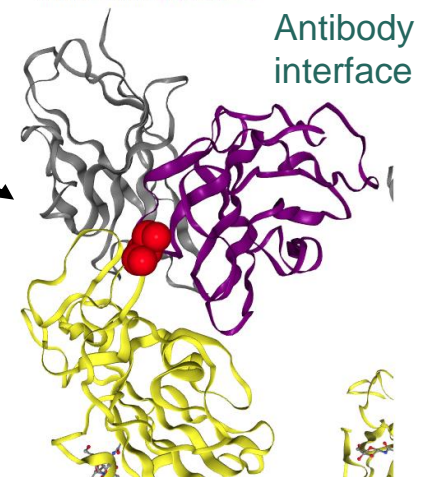
Protein: Spike
 Coronavirus type: Yeast SARS-CoV-2 (2019)
 Mutation (as in paper): S477N
 neutral AA: S
 neg. eff. AA: N
 Effect: Host Change

Receptor binding

Comment:
 In a deep mutational scanning experiment that expresses Spike RBD in a yeast-display platform, S477N mildly increases the binding to ACE2 (apparent dissociation constant delta-log10 value: 0.06)

[Literature reference](#)
 (Mutation S477N in the paper is at an equivalent position of the mutation in your query)

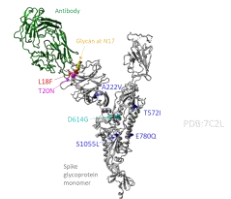
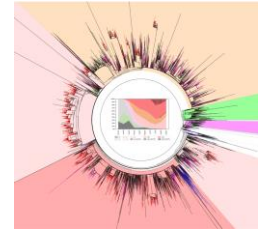
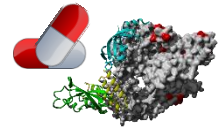
Screenshot
 Antibody interaction: The mutation position (red atoms) corresponds to position 477 on viral chain E (yellow backbone) of protein entry 6xcn, originating from Severe acute respiratory syndrome coronavirus 2 and with a label of Spike glycoprotein. The mutation is within 6 Å from antibody chain F (purple backbone).



Real-time Surveillance of Pathogen Evolution and Global Transmission

(from genome to action)

1. Development of first diagnostics kits and refinement through ongoing surveillance for mutations
2. Identification of potential drug and vaccine targets on hCoV-19 through repurposing
3. Genomic epidemiology of hCoV-19, allows analysis of the exportation and importation events of viruses between countries, contact-tracing in countries, or identification of transmission chains
4. Evidence that the virus has not drifted to significant strain difference, with in particular the cell receptor binding pocket being followed closely
5. Identification of animal precursors of hCoV-19 (in bats and pangolins)



We gratefully acknowledge the Authors from Originating and Submitting laboratories of sequence and meta data.

Supported by
BII/GIS, A*STAR Singapore



BII – CoVsurver

Raphael

Ashar

Louis, Roland

Taipang+IT

BII/GIS – GISAID, MOH

Suma, Meera, Swaine (GIS)

Fernanda

Yi Hong, Shilpa

Niranjan (GIS)

Shawn (IMCB)

BII – drug-related modelling

Hao, Peter, Roland, Chandra

Ken

Sam

COVID Interns (Jan-Dec 2020)

Rishabh Anand

NJC

Shi Shu Yuan

NUS High

Daniel See Rui En

Hwa Chong Institution

Shannon Low Hui Ching

Imperial College London

Teo Yu Xing

NUS

Zhu Xingyu

NUS

Matthew Wong Kai Xian

University of St Andrews

Nur Shaheen Binte Zainudin

King's College London

Vijayakumar Ragavi

University of Cambridge

Erin Chung

Boston University

Yeo Keng Hoang Winston

University of Waterloo

Yang Yiqing

NUS

Lo Hui Zhen

Monash University

**GIS, SigN/ID labs, IMCB, EDDC, DxD, Duke-NUS,
DSO, NUS, NPHL, NCID, MOH, TTSH, NUH, SGH
...and many more!**

THANK YOU!

Funding by A*STAR, NMRC, WHO, CEPI