Bootstrapping SARS-CoV-2 Genomic Capacity for Public Health Surveillance and Outbreak Response

SARS-CoV-2 Sequencing for Public Health Emergency Response, Epidemiology, and Surveillance

DUNCAN MACCANNELL PHD

ADVANCED MOLECULAR DETECTION (AMD) PROGRAM
CENTERS FOR DISEASE CONTROL AND PREVENTION, ATLANTA

@dmaccannell

dmaccannell@cdc.gov

Why Sequence SARS-CoV-2 at scale?

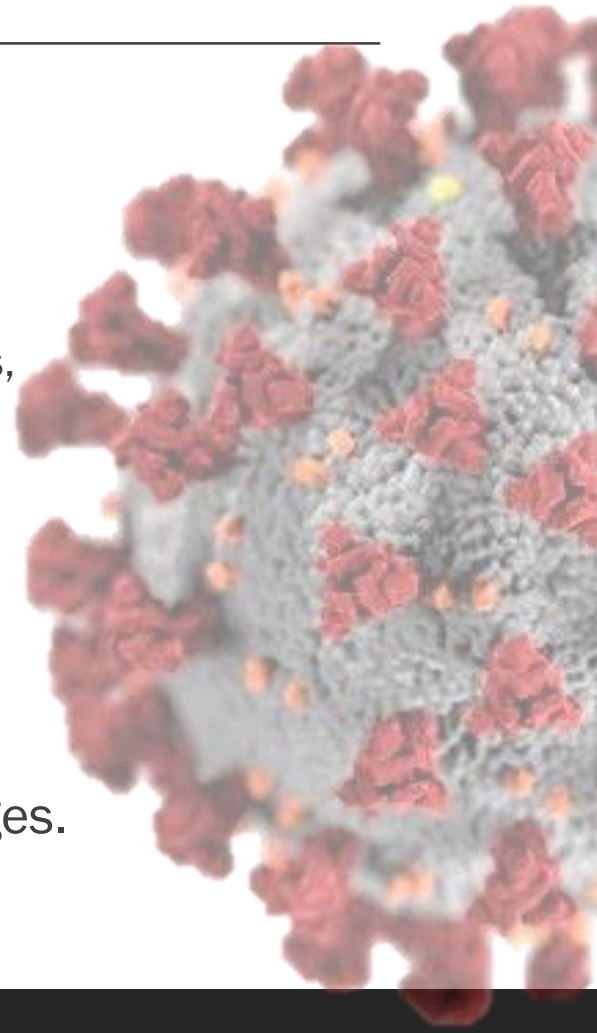
Monitor viral diversity over time:

- Understand transmission dynamics (locally, regionally, nationally)
- Help inform public health responses, including contact tracing and overall containment/mitigation strategies
- Identify the emergence of clinically-important variants: unusual virulence/phenotypes, potential disruption of diagnostic, antigenic or therapeutic targets.

Establish national and regional baselines:

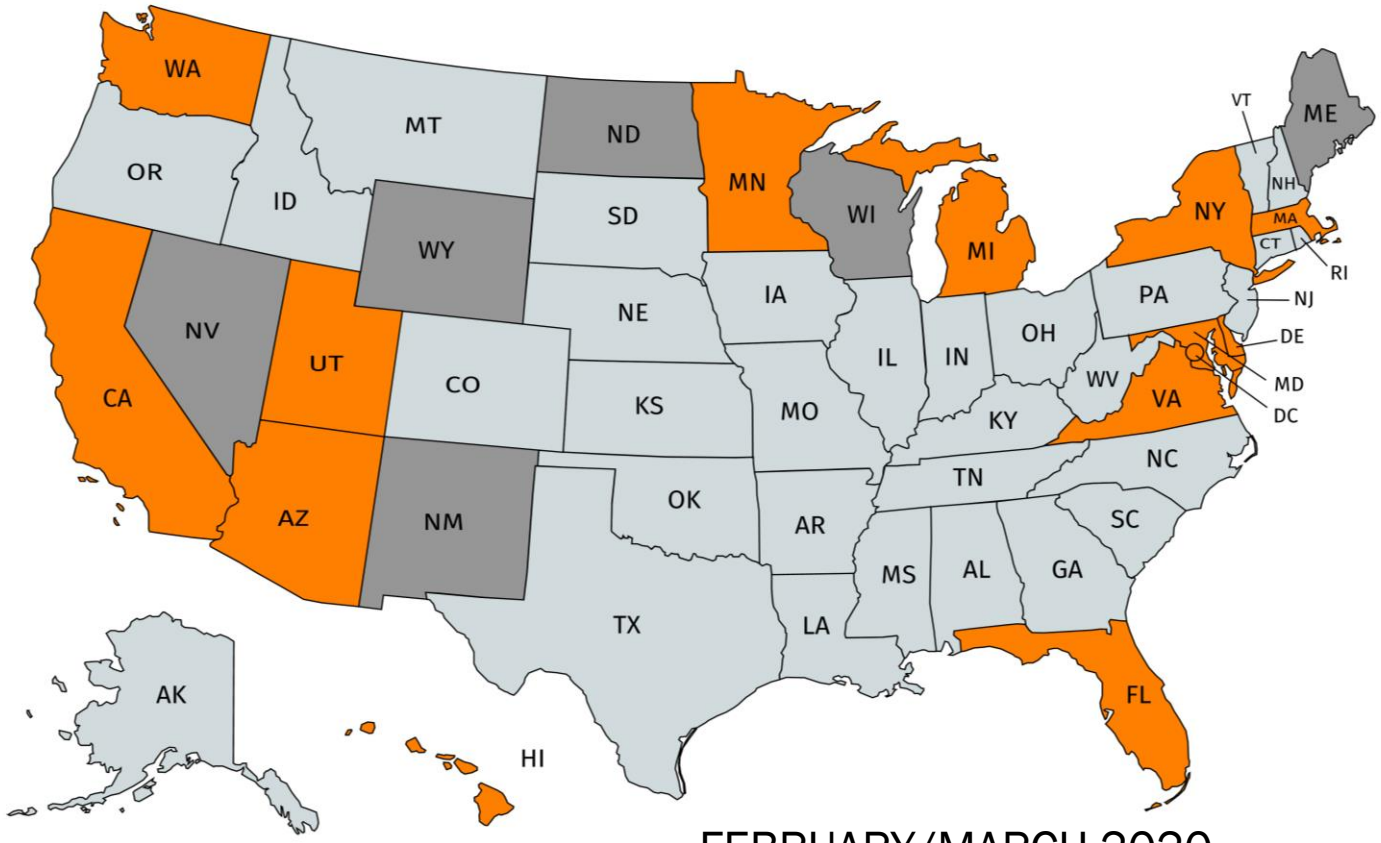
- Sustainable and longitudinal data collection
- Balance oversampling and undersampling
- Compare regions with different timelines and response strategies
- Inform public health action on flare ups, after initial pandemic wave.

Foster new collaboration and innovation to meet current and future challenges.

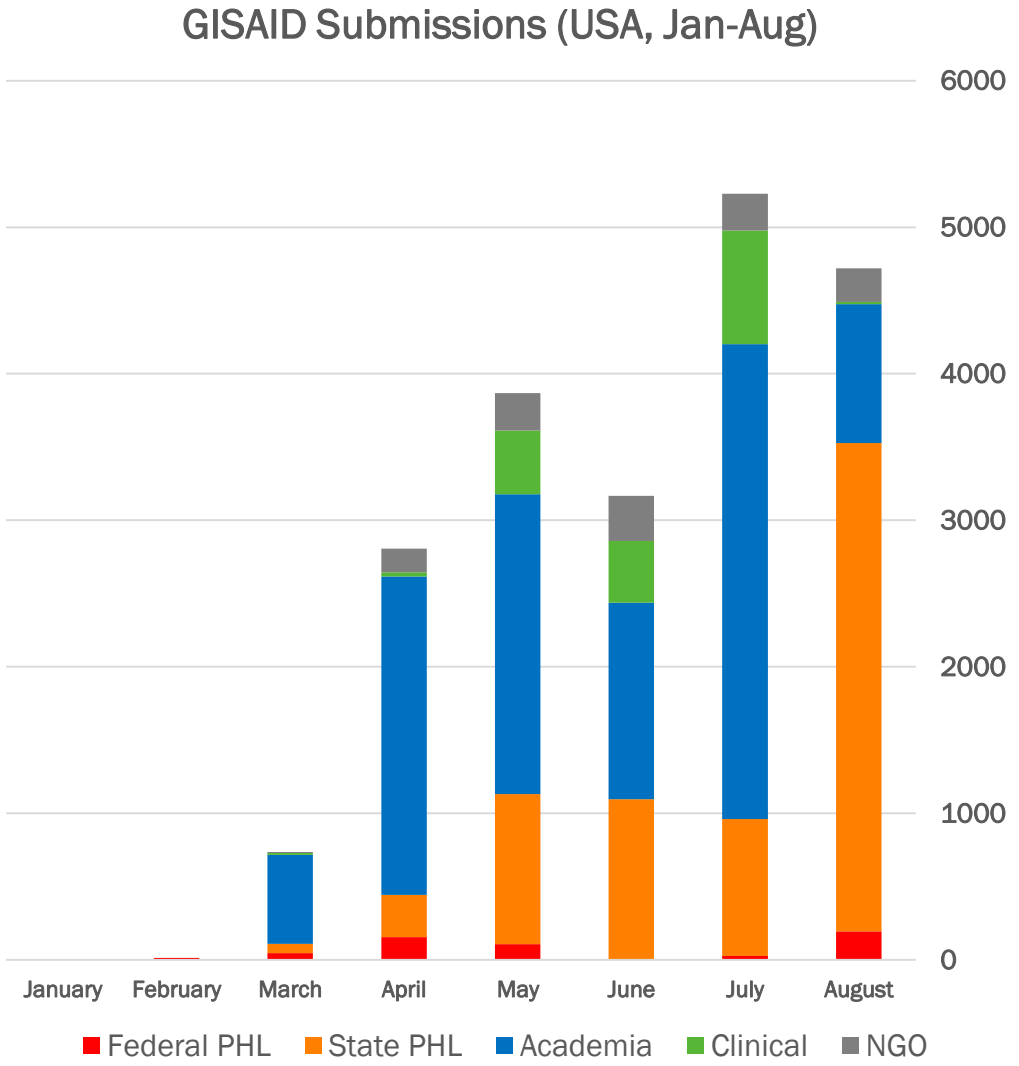


THE EARLY DAYS

- JAN-FEB REFERENCE/CONFIRMATORY TESTING AT CDC
- CDC SEQUENCING AND RELEASING ALL POSITIVES

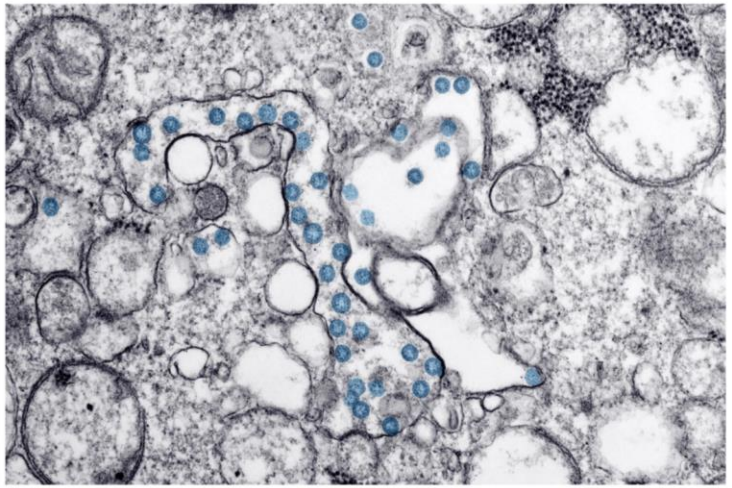


FEBRUARY/MARCH 2020



Labs Across U.S. Join Federal Initiative to Study Coronavirus Genome

The project, announced by the C.D.C., will help trace patterns of transmission, investigate outbreaks and map how the virus is evolving, which can affect a cure.



Virus particles, marked in blue, that were taken from an early Covid-19 patient in the U.S. C.D.C./E.P.A., via Shutterstock

By Sheri Fink

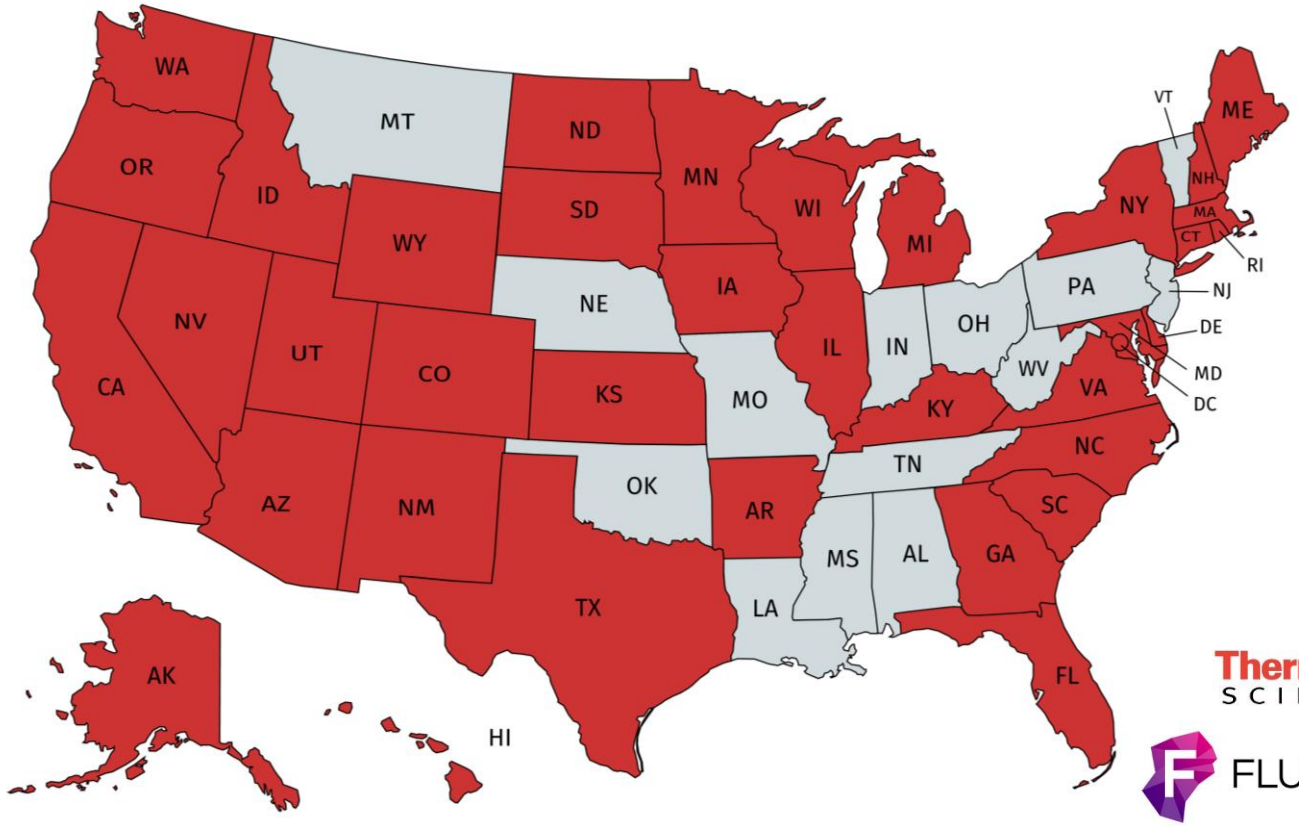
Published April 30, 2020 Updated May 6, 2020

30

The Centers for Disease Control and Prevention on Thursday announced a national initiative to speed research into how the coronavirus was spreading around the country, bringing together at least 75 public health, academic and commercial institutions studying its genome.

https://www.nytimes.com/2020/04/30/science/coronavirus-genetic-sequencing.html

Who makes up the SPHERES Consortium?

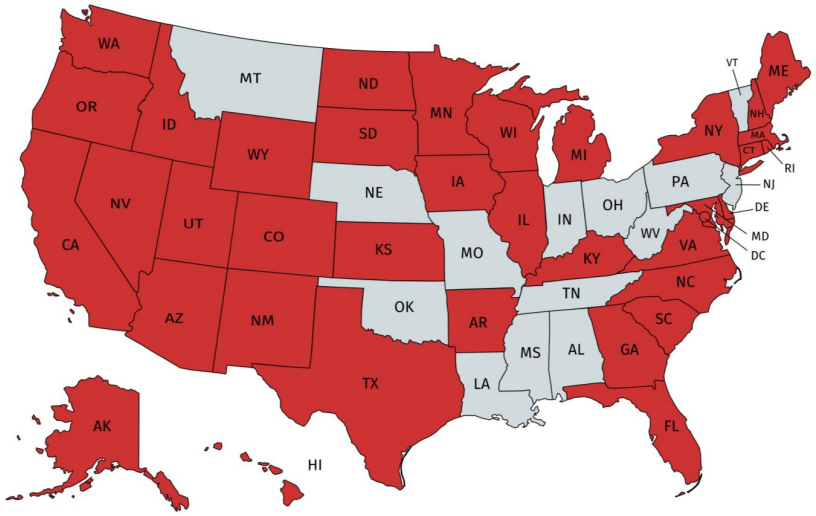


37 State Public Health Departments
20 County Public Health Departments
CDC, NCBI, NIAID, FDA, NIST, APHL
and…

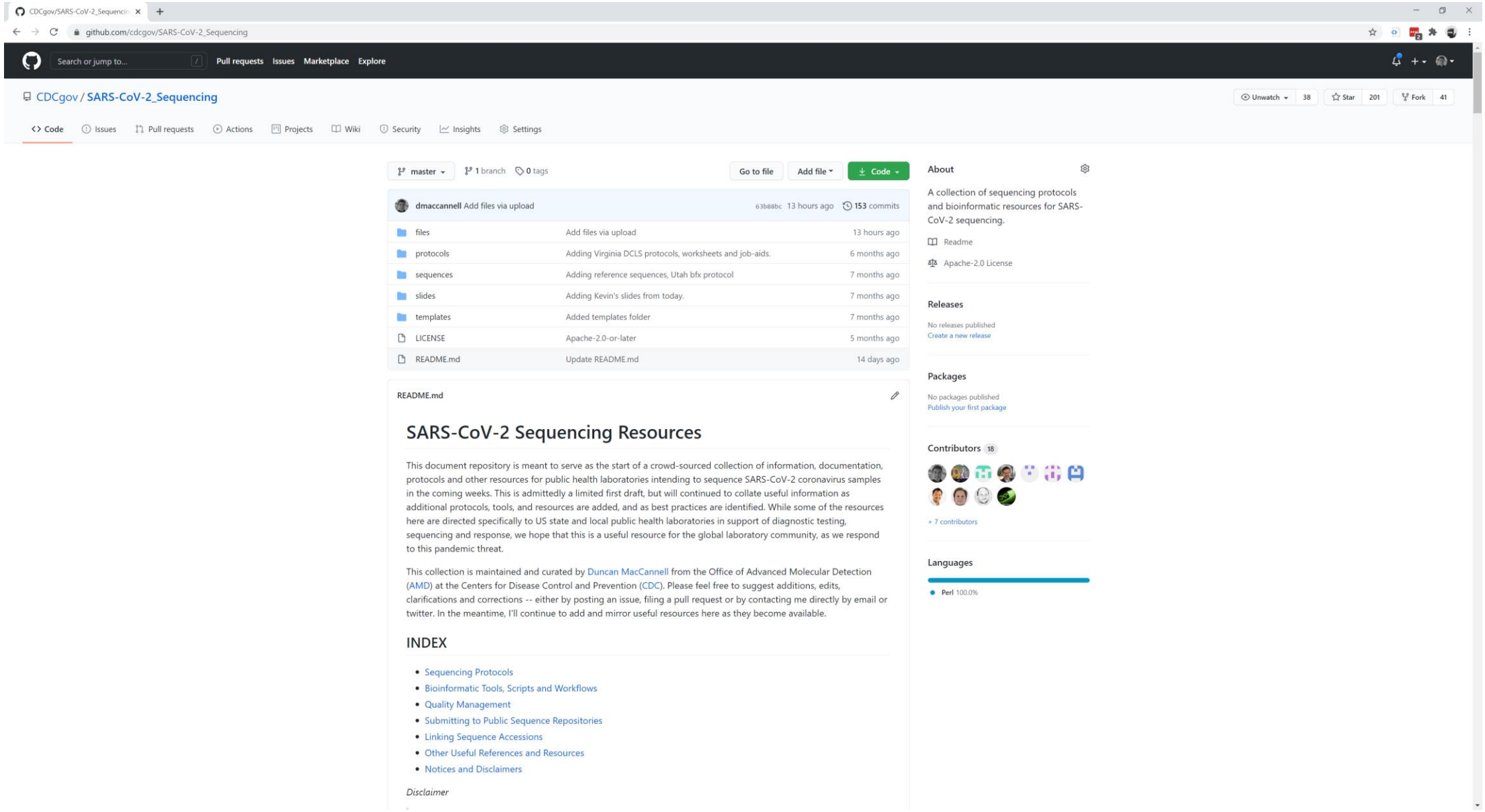
639 SCIENTISTS
150+ ORGANIZATIONS

SARS-CoV-2 SPHERES – Pathogen Genomics

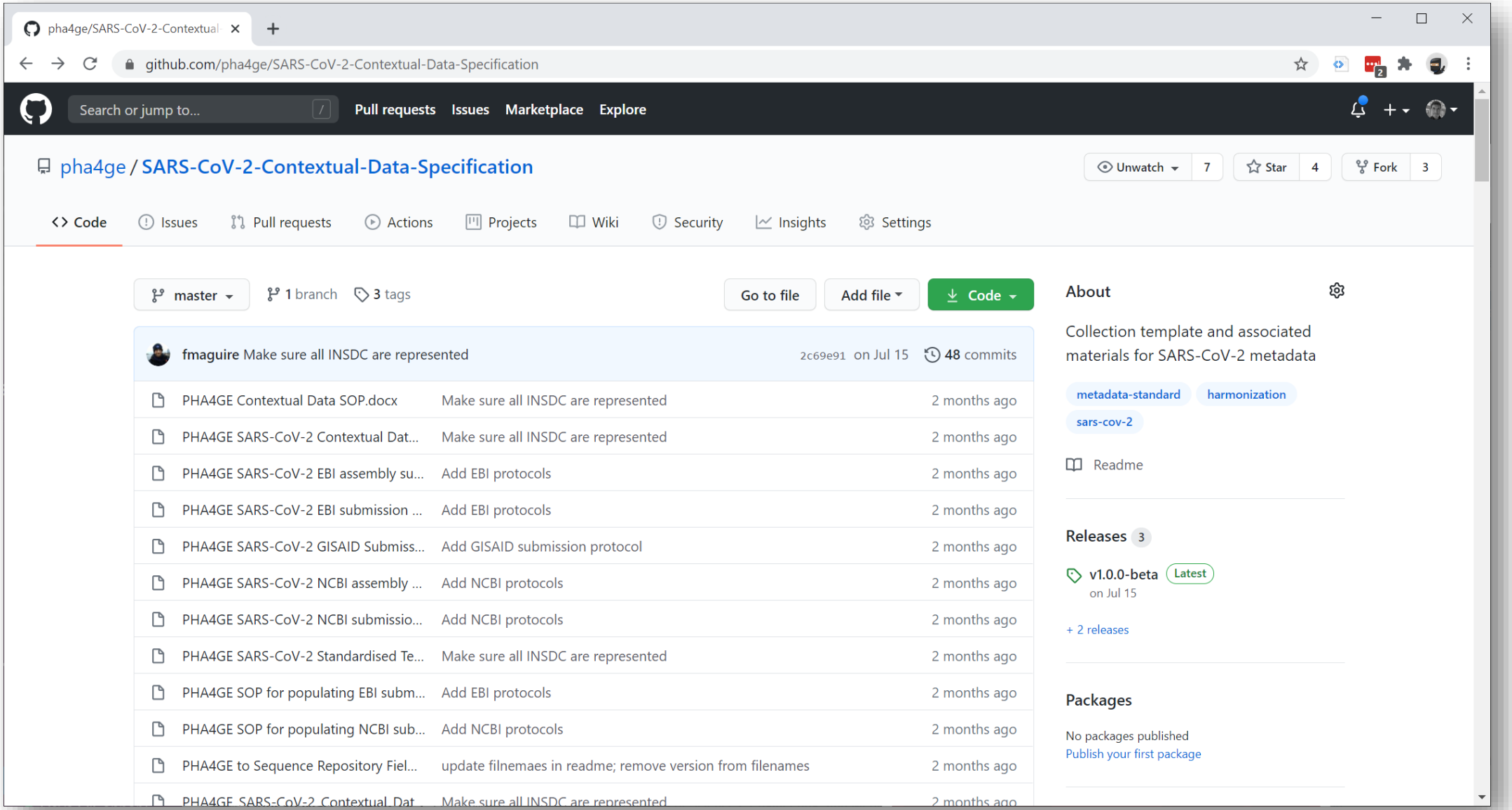
- **GOAL:** Maximize quality, quantity and usefulness of SARS-CoV-2 sequence data; drive open data for public health and basic research.
- Builds upon 5+ years of AMD genomics/bioinformatics investments across the public health system.
- Coordinates across dozens of individual sequencing and genomics projects from across the country.
- Engages new partners from across government, public health, academia, NGO, clinical labs & private sector.
- Align with international sequencing efforts (COG-UK, CanCOGeN, AusTrakka, Africa CDC, PHA4GE).



https://go.usa.gov/xvMmZ



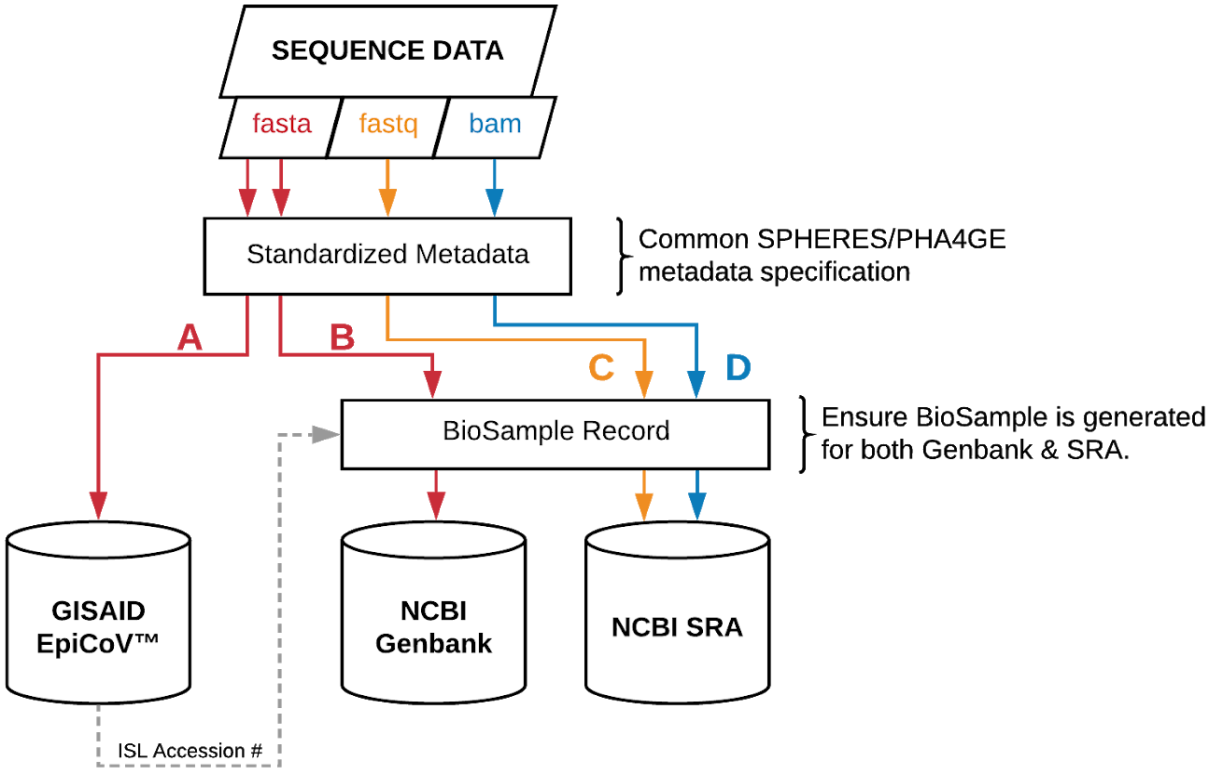
https://github.com/cdcgov/SARS-CoV-2_Sequencing

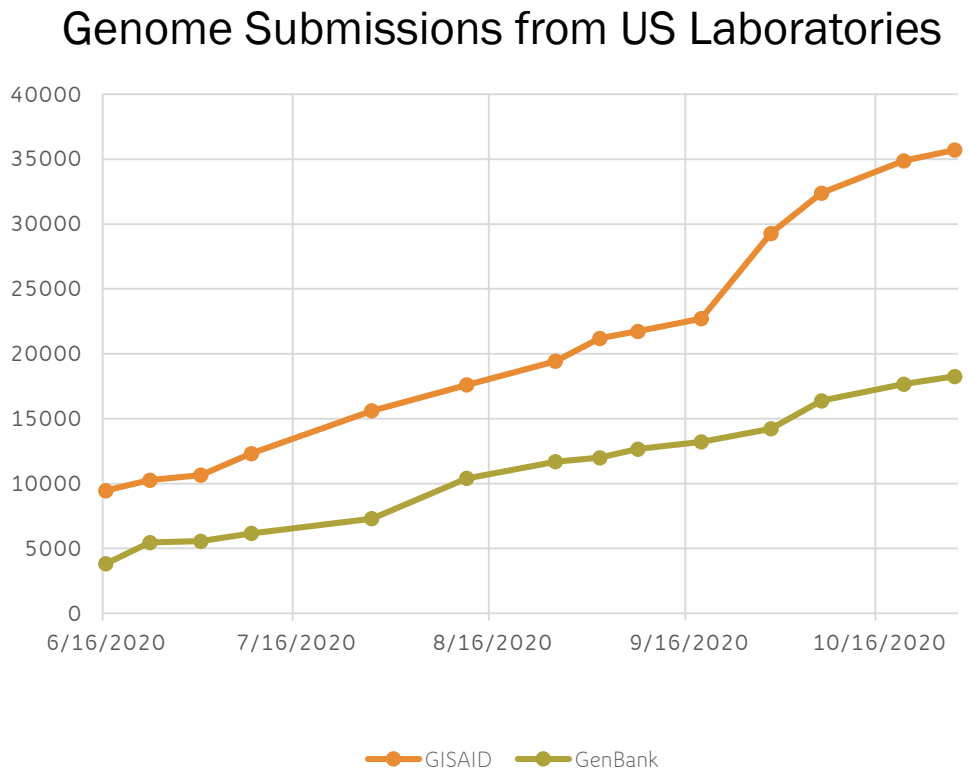


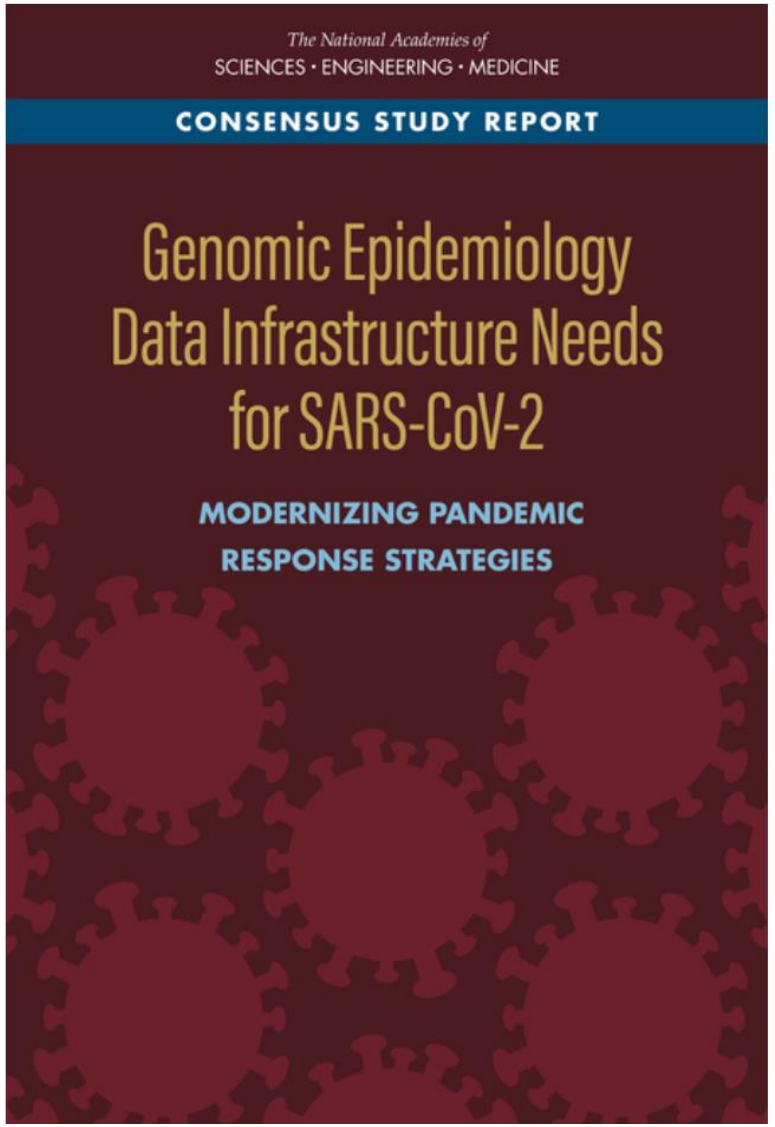
https://github.com/pha4ge/SARS-CoV-2-Contextual-Data-Specification

Protocols, Zoom, Slack, Software, Repositories.

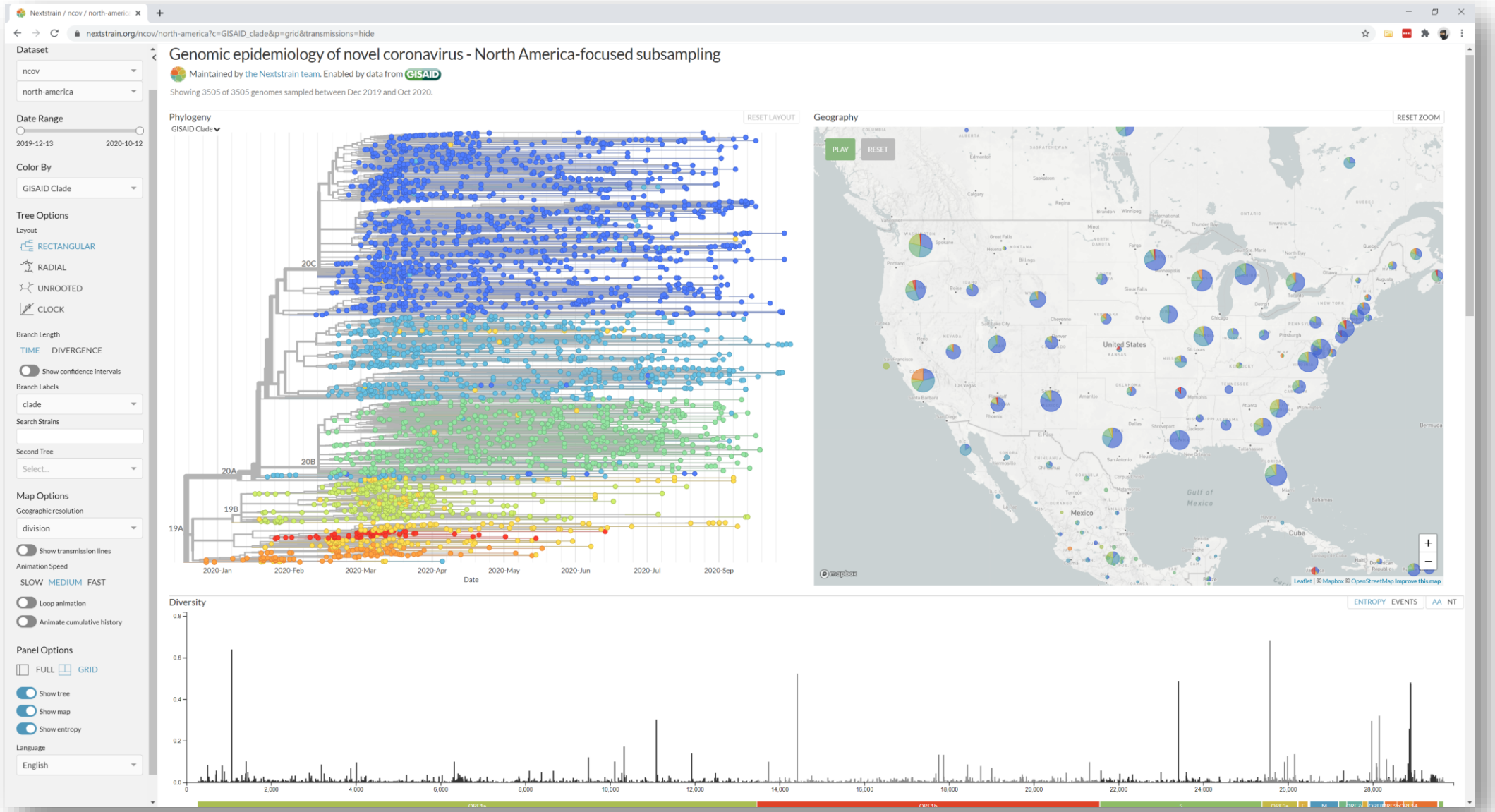
SARS-CoV-2 Submissions (GISAID + NCBI)






The National Academies of
SCIENCES • ENGINEERING • MEDICINE
CONSENSUS STUDY REPORT
Genomic Epidemiology Data Infrastructure Needs for SARS-CoV-2
MODERNIZING PANDEMIC RESPONSE STRATEGIES

"patchy"



<0.5% of total US SARS-CoV-2 genomes have been sequenced.

USA (10/20/2020): 34,874 sequences in GISAID EpiCoV | NCBI (INSDC) >> 17,671 in GenBank – ??? in SRA

How the coronavirus's genetic co

washingtonpost.com/graphics/2020/health/coronavirus-genetic-code/

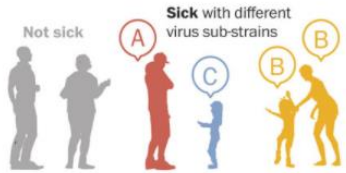
View site information

Sections

The Washington Post
Democracy Dies in Darkness

fms2@cdc.gov

Coronavirus Latest news U.S. map World map FAQ Vaccine tracker Coronavirus Living Extraordinary People



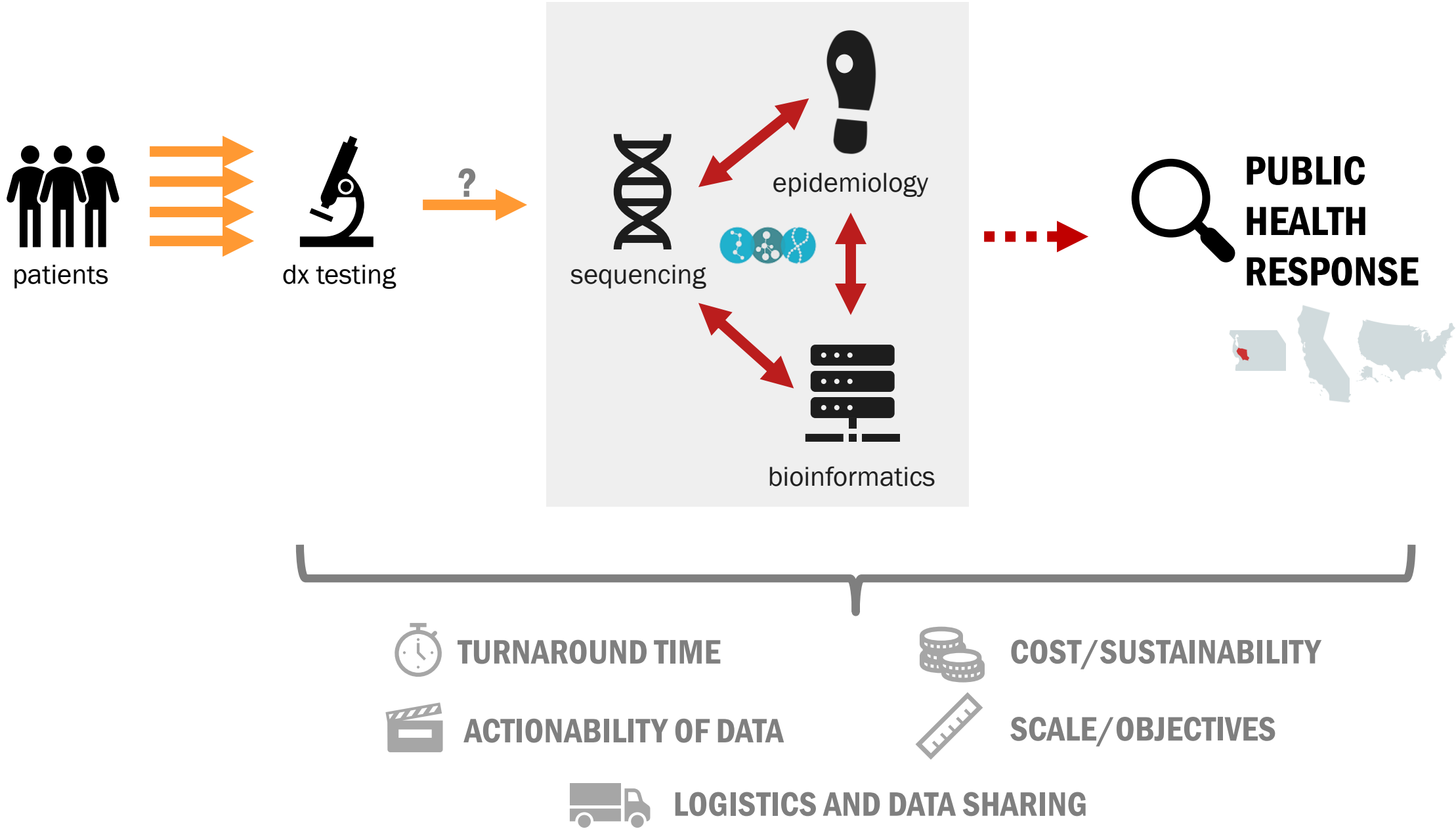
Scientists have a powerful new tool for controlling the coronavirus: Its own genetic code.

The United States, home to the world's biggest outbreak, has failed to tap the technique's full potential.

Scroll to continue

https://www.washingtonpost.com/graphics/2020/health/coronavirus-genetic-code/

>> So how do we reach our full potential?



Epi-backed genomic sequence analysis within an actionable timeframe of <24 to 48hrs?

Comment on this paper

Rapid implementation of real-time SARS-CoV-2 sequencing to investigate healthcare-associated COVID-19 infections

Luke W Meredith, William L Hamilton, Ben Warne, Charlotte J Houldcroft, Myra Hosmillo, Aminu Jahun, Martin D Curran, Surendra Parmar, Laura Caller, Sarah L Caddy, Fahad A Khokhar, Anna Yakovleva, Grant R Hall, Theresa Feltwell, Sally N Forret, Sushmita Sridhar, Michael p Weekes, Stephen Baker, Nicholas Brown, Elinor Moore, Theodore Gouliouris, Ashley Popay, Iain Roddick, Mark Reacher, Sharon Peacock, Gordon Dougan, M. Estee Torok, Ian Goodfellow

doi: https://doi.org/10.1101/2020.05.08.20095687

Abstract

Background The burden and impact of healthcare-associated COVID-19 infections is unknown. We aimed to examine the utility of rapid sequencing of SARS-CoV-2 combined with detailed epidemiological analysis to investigate healthcare-associated COVID-19 infections and to inform infection control measures. Methods We set up rapid viral sequencing of SARS-CoV-2 from PCR-positive diagnostic samples using nanopore sequencing, enabling sample-to-sequence in less than 24 hours. We established a rapid review and reporting system with integration of genomic and epidemiological data to investigate suspected cases of healthcare-associated COVID-19. Results Between 13 March and 24 April 2020 we collected clinical data and samples from 5191 COVID-19 patients in the East of England. We sequenced 1000 samples, producing 747 complete viral genomes. We conducted combined epidemiological and genomic analysis of 299 patients at our hospital and identified 26 genomic clusters involving 114 patients. 66 cases (57.9%) had a strong epidemiological link and 15 cases (13.2%) had a plausible epidemiological link. These results were fed back clinical, infection control and hospital management teams, resulting in infection control interventions and informing patient safety reporting. Conclusions We established real-time genomic surveillance of SARS-CoV-2 in a UK hospital and demonstrated the benefit of combined genomic and epidemiological analysis for the investigation of healthcare-associated COVID-19 infections. This approach enabled us to detect cryptic transmission events and identify opportunities to target infection control interventions to reduce further healthcare-associated infections.

https://www.medrxiv.org/content/10.1101/2020.05.08.20095687v1

Three areas of sequencing priority and focus:

1. **Coordinated sequencing through SARS-CoV-2 SPHERES network.**
   - Directed studies and deep-dives:
     - Research funding opportunities: Broad Agency Announcements, other possible mechanisms.
   - Public-private partnership + shared interest = tremendous opportunities for applied public health research, innovation and engagement.
   - Initial passive surveillance using existing sequencing efforts.

2. **S3 + Baseline: structured sequencing for national surveillance.**
   - State Strain Surveillance: Sequencing of PHL specimens shipped to CDC, based on existing seasonal flu model.
   - SPHERES Baseline: Contracted sequencing model; synergistic, flexibility, ability to target sequencing.

3. **Flexible support for state/local outbreak investigations and surveillance.**
   - Cluster investigations, local and regional transmission patterns, vulnerable populations, reinfections, institutes of higher education, extended healthcare, workplace, etc.

A New Model: Positives and Negatives

POSITIVES.

1. Close, active **engagement** between STLT public health, academia and private sector. New expertise/collaboration.
2. Testbed for new methods, bioinformatic tools, standards and **NIAID/NCBI/CDC** resources.
3. Rapid **information sharing** – common problems, best practices, unusual findings, new ideas.
4. More options for **direct access** to samples, populations. Incredibly flexible. Scope.
5. **Participation** has not tailed off: participants on weekly Zoom: 85 to 120.

NEGATIVES.

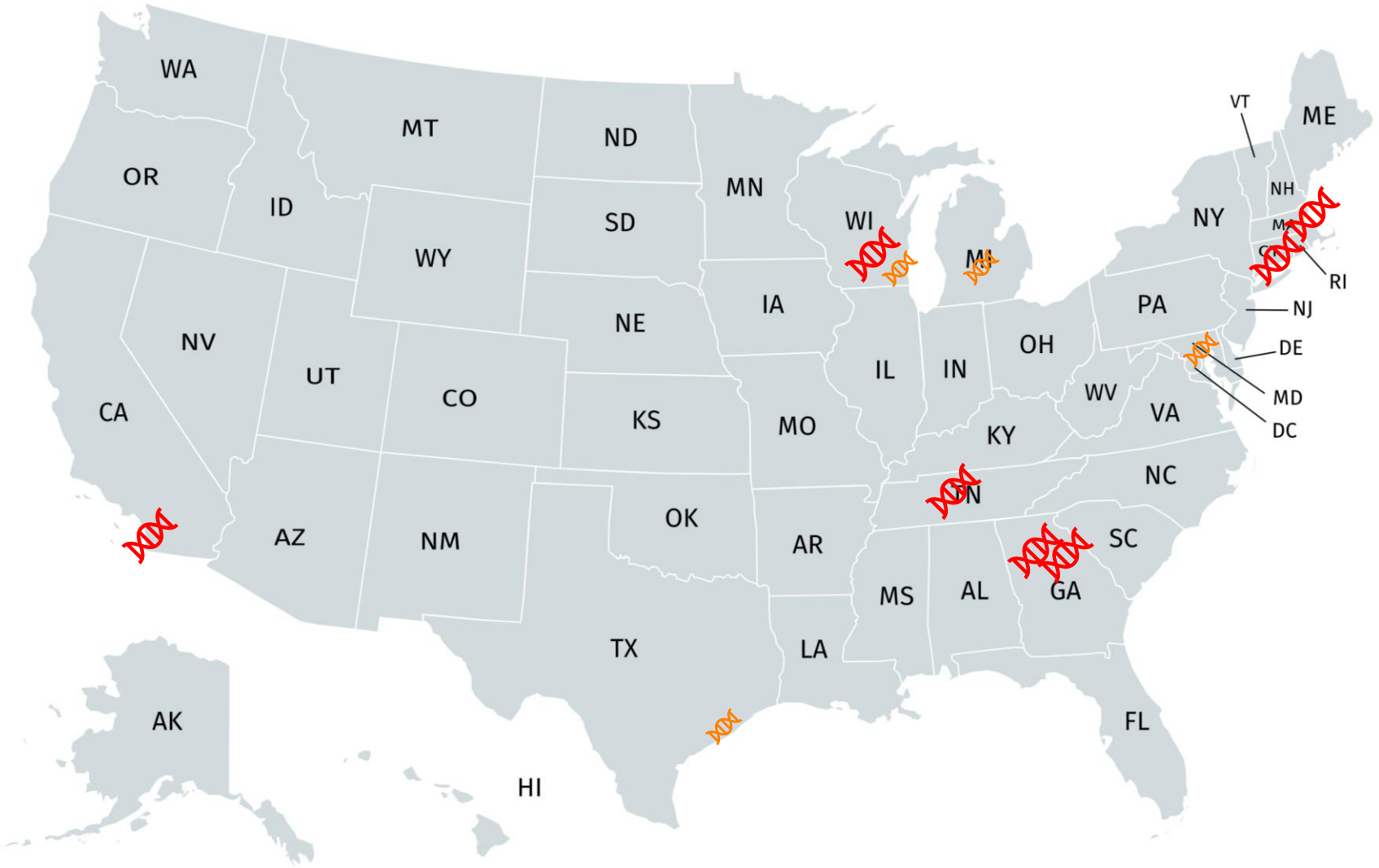
1. Many **competing interests and priorities**, different levels of experience with public health.
2. Research vs. non-research – different philosophies and **funding eligibility**.
3. Uneven participation – not all states, slowly ramping epi engagement
4. Reproducibility and sustainability of this model?

CDC Coronavirus Funding to Jurisdictions (April 23, 2020)

The data provided is covered in the HHS Announces CARES Act Funding Distribution to States and Localities in Support of COVID-19 Response news release.

Jurisdiction	Coronavirus Preparedness and Response Supplemental Appropriations Act						Coronavirus Aid, Relief, and Economic Security Act	
	Crisis CoAg Award #1 Complete	Crisis CoAg Award #2 Complete	Crisis CoAg Award #3 Complete	Epi & Lab Capacity Award Complete	Emerging Infections Program Award Complete	Tribal Awards (in progress)	Epi & Lab Capacity Reopen America (in progress)	Tribal Awar (in progres
Alabama		$8,148,798.90					$9,054,221.00	
Alaska		$4,902,840.00					$5,447,600.00	
American Samoa		$370,246.50	$740,493				$411,385.00	
Arizona	$500,000.00	$11,201,871.60	$4,520,040				$12,446,524.00	
Arkansas		$6,205,347.00					$6,894,830.00	
California	$3,500,000.00	$37,706,709.60	$18,525,727	$1,000,000.00	$2,600,000.00		$41,896,344.00	
Colorado		$9,331,323.30	$3,589,716	$500,000.00	$2,600,000.00		$10,368,137.00	
Connecticut		$7,058,270.70	$2,251,727	$500,000.00	$2,600,000.00		$7,842,523.00	
Delaware		$4,567,500.00					$5,075,000.00	

BAAs



- 7 awards this cycle
- Supporting work in 11 sites.
- <u>Themes</u>: IHE, capacity development, genomic epidemiology, surveillance and outbreak investigations, host/pathogen interface, clinical risk factors.

Three areas of sequencing priority and focus:

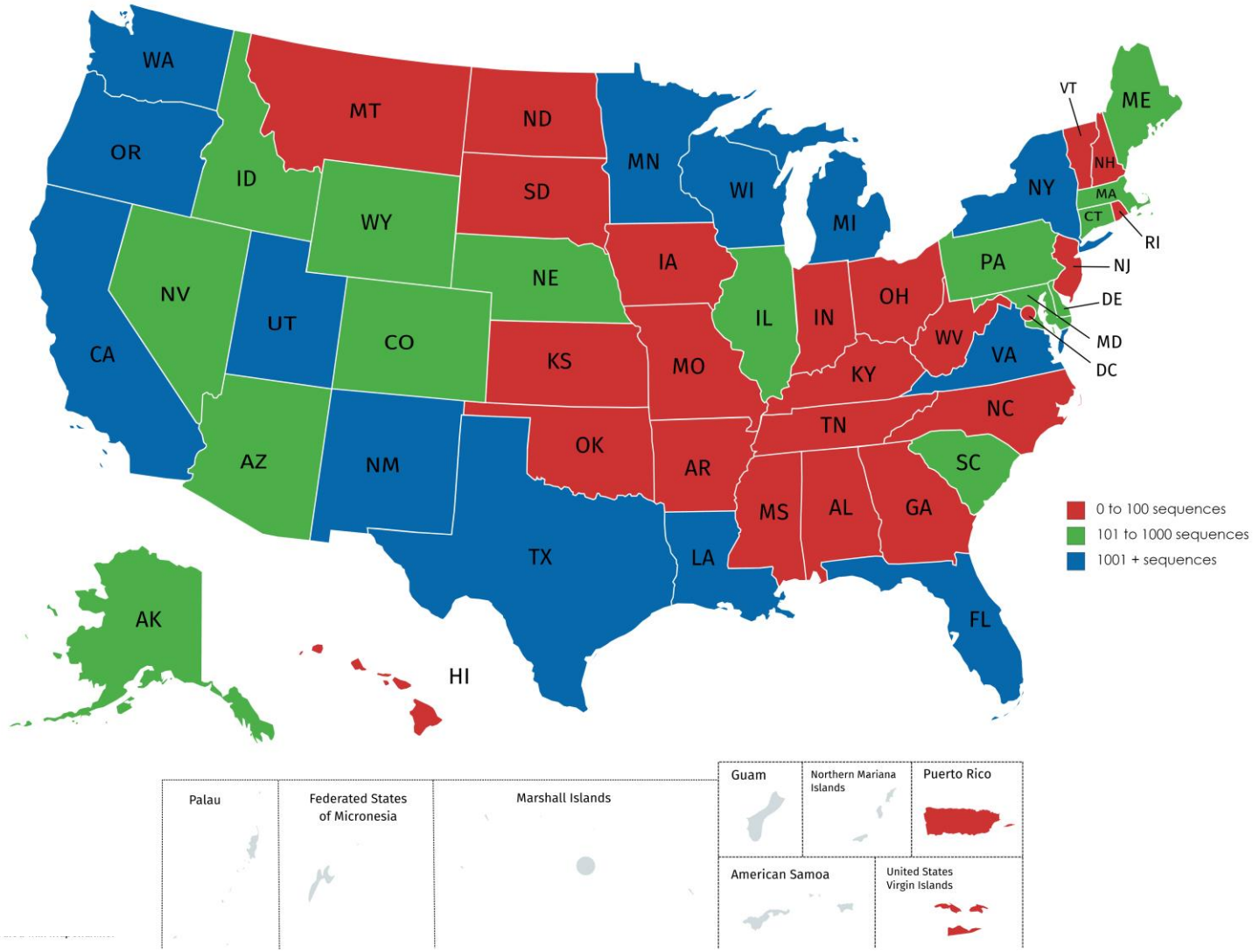
1. Coordinated sequencing through SARS-CoV-2 SPHERES network.
   - Directed studies and deep-dives:
     - Research funding opportunities: Broad Agency Announcements, other possible mechanisms.
   - Public-private partnership + shared interest = tremendous opportunities for applied public health research, innovation and engagement.
   - Initial passive surveillance using existing sequencing efforts.

2. S3 + Baseline: structured sequencing for national surveillance.
   - State Strain Surveillance: Sequencing of PHL specimens shipped to CDC, based on existing seasonal flu model.
   - SPHERES Baseline: Contracted sequencing model; synergistic, flexibility, ability to target sequencing.

3. Flexible support for state/local outbreak investigations and surveillance.
   - Cluster investigations, local and regional transmission patterns, vulnerable populations, reinfections, institutes of higher education, extended healthcare, workplace, etc.

S3 + BASELINE

TEN STATES WITH ≤10 SEQUENCES



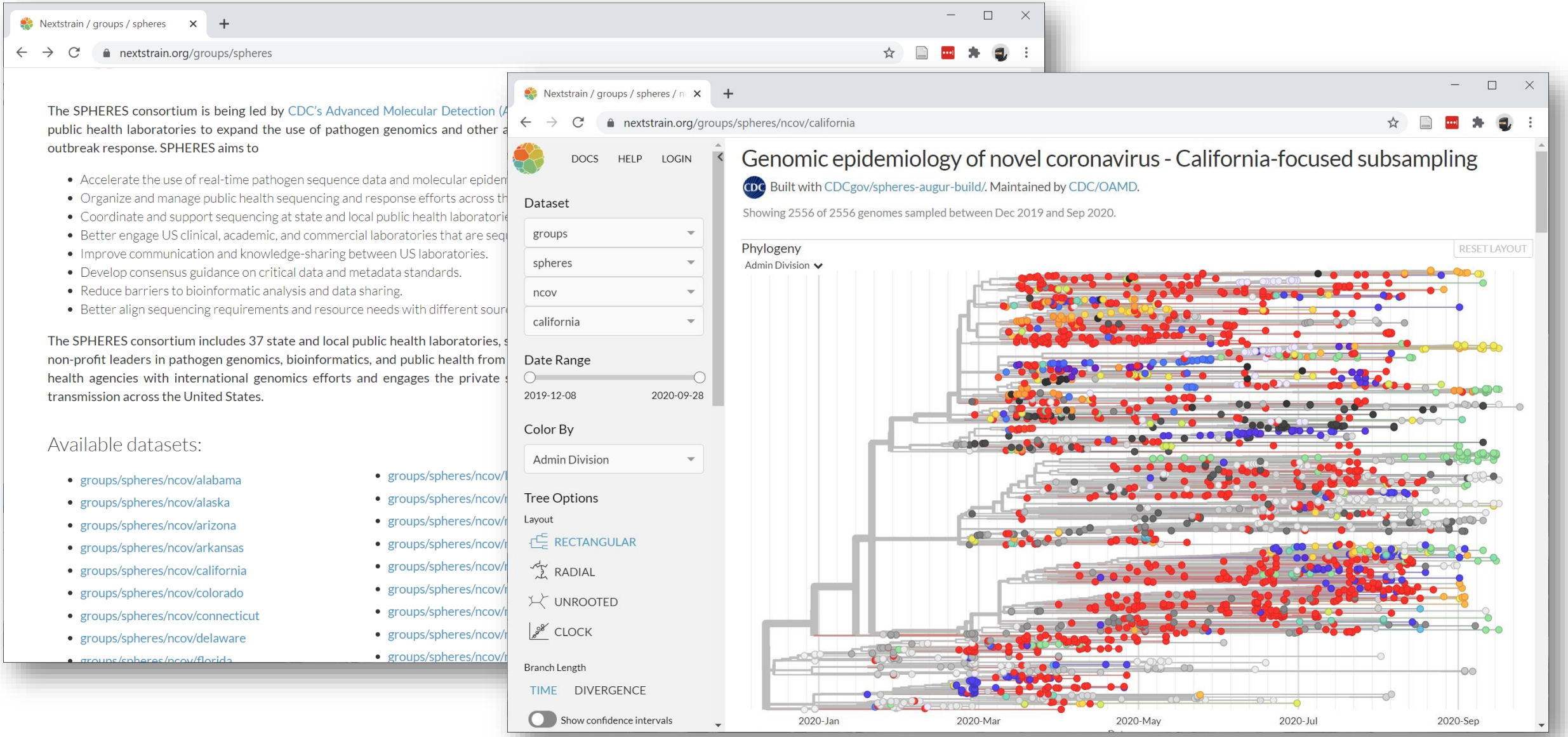
State Strain Surveillance

- 10 random positive samples/month/state
- 500+ sequences/month, 6,500/year

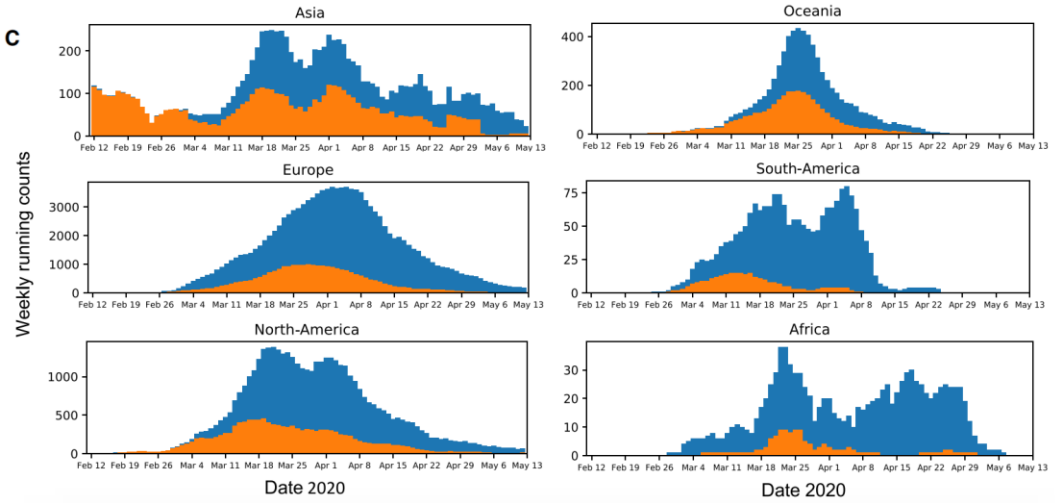
SPHERES Baseline

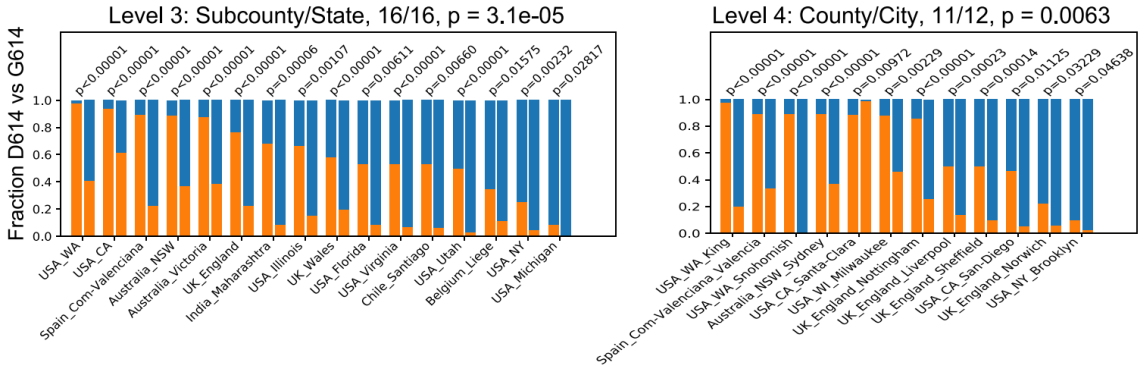
- Working with large diagnostic labs to
  - Retrospectively sequence a ton of samples
  - Prospectively enhance and supplement S3

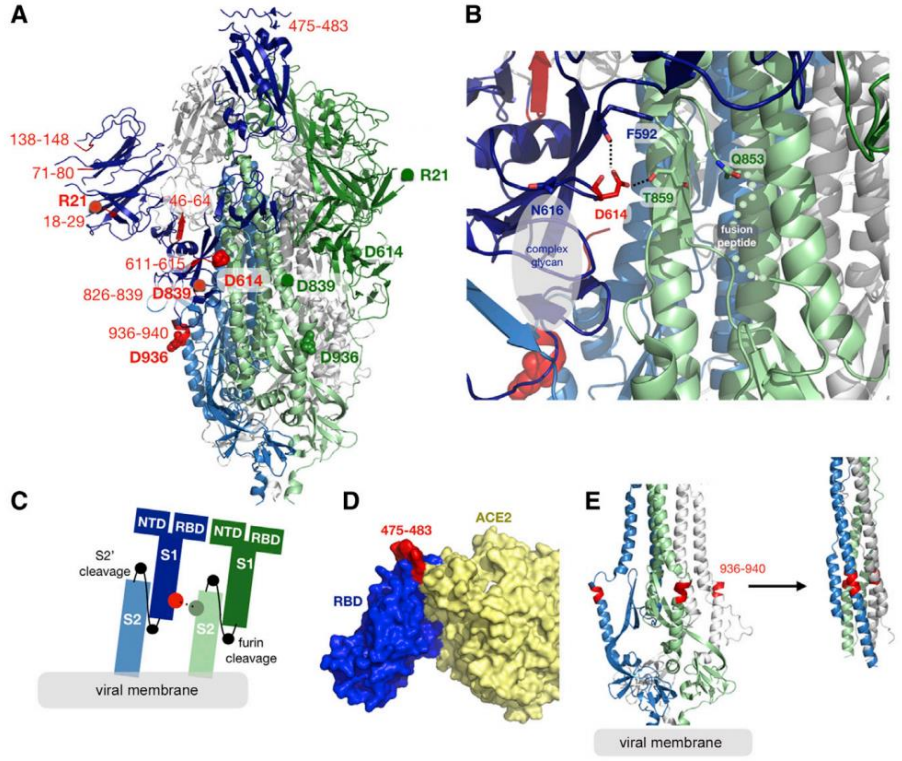
GISAID submissions as of 10/28/2020



State-specific public builds, updated weekly.
50 states + DC + PR (currently)



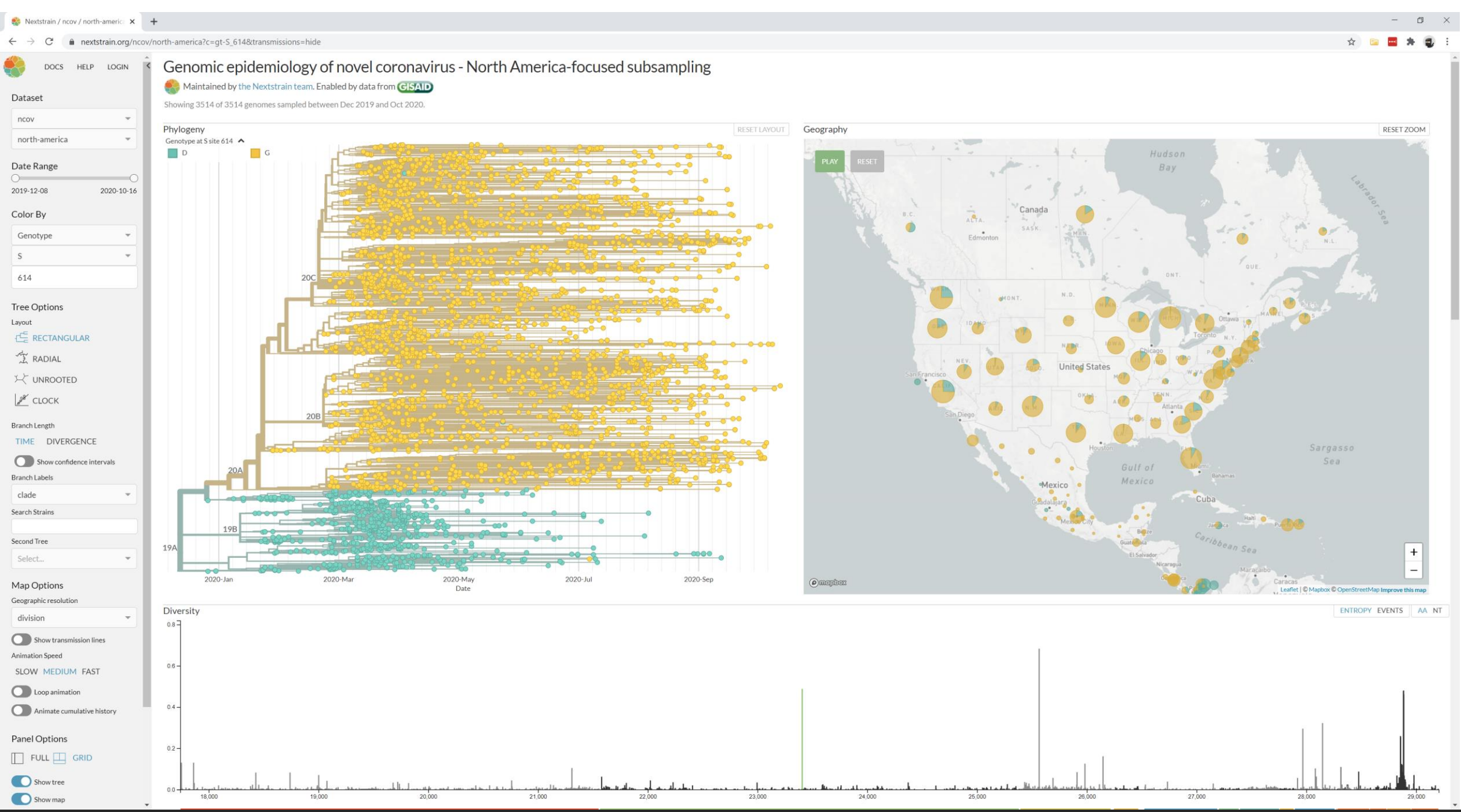




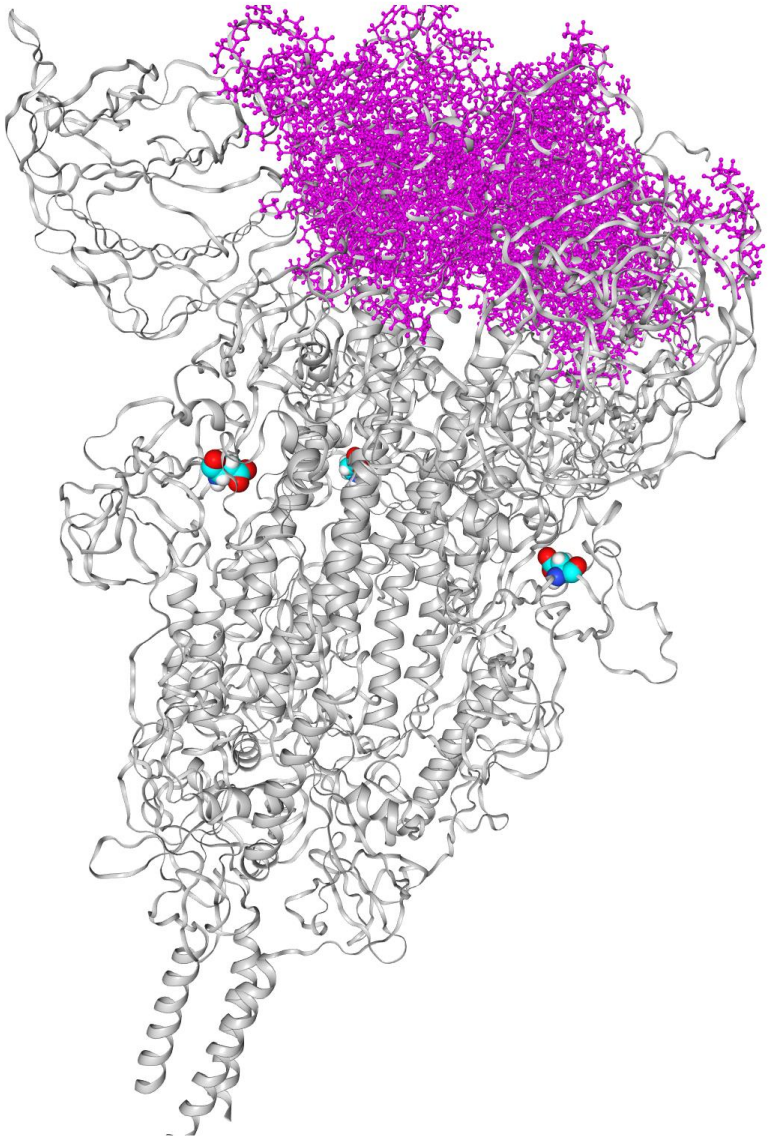
D614G first observed in a sequence from China Jan 24.

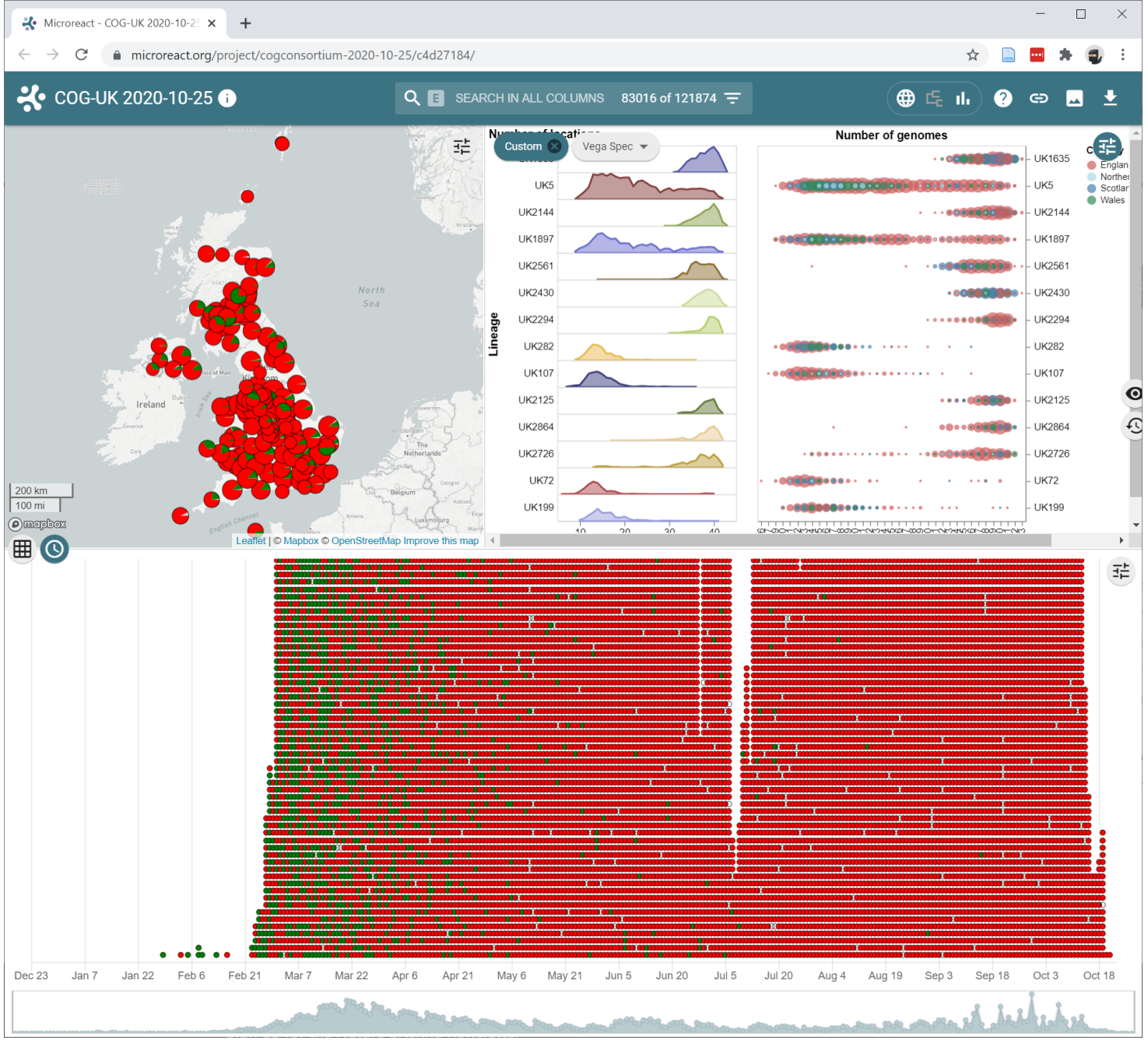
23403 A->G

Korber Cell 2020 https://doi.org/10.1016/j.cell.2020.06.043



https://nextstrain.org/ncov/north-america?c=gt-S_614&transmissions=hide





https://cov3d.ibbr.umd.edu/MutViewer/QOP60983

npj | Vaccines

www.nature.com/npjvaccines

BRIEF COMMUNICATION OPEN

Check for updates

Experimental and in silico evidence suggests vaccines are unlikely to be affected by D614G mutation in SARS-CoV-2 spike protein

Alexander J. McAuley[1], Michael J. Kuiper[2], Peter A. Durr[1], Matthew P. Bruce[1], Jennifer Barr[1], Shawn Todd[1], Gough G. Au[1], Kim Blasdell[1], Mary Tachedjian[1], Sue Lowther[1], Glenn A. Marsh[1], Sarah Edwards[1], Timothy Poole[1], Rachel Layton[1], Sarah-Jane Riddell[1], Trevor W. Drew[1], Julian D. Druce[3], Trevor R. F. Smith[4], Kate E. Broderick[4] and S. S. Vasan[1,5]✉

The 'D614G' mutation (Aspartate-to-Glycine change at position 614) of the SARS-CoV-2 spike protein has been speculated to adversely affect the efficacy of most vaccines and countermeasures that target this glycoprotein, necessitating frequent vaccine matching. Virus neutralisation assays were performed using sera from ferrets which received two doses of the INO-4800 COVID-19 vaccine, and Australian virus isolates (VIC01, SA01 and VIC31) which either possess or lack this mutation but are otherwise comparable. Through this approach, supported by biomolecular modelling of this mutation and the commonly-associated P314L mutation in the RNA-dependent RNA polymerase, we have shown that there is no experimental evidence to support this speculation. We additionally demonstrate that the putative elastase cleavage site introduced by the D614G mutation is unlikely to be accessible to proteases.

npj Vaccines (2020)5:96; https://doi.org/10.1038/s41541-020-00246-8

INTRODUCTION

COVID-19 vaccine candidates primarily target the trimeric 'spike' (S) glycoprotein, as this factor enables binding to the 'angiotensin-converting enzyme 2' (ACE2) host surface receptors and facilitates virus entry into the cells[1]. Over the last few months, an Aspartate-to-Glycine amino acid change has arisen at position 614 of the S protein (resulting from a single A-to-G nucleotide change at position 23,403 in the Wuhan-Hu-1 reference genome), with G614 variants accounting for 75% of published genome sequences worldwide as of 1 July 2020. This mutation has resulted in a number of articles and preprints postulating that isolates containing this 'D614G' mutation have a structural advantage[2], including as a better substrate to the S1 furin cleavage domain[3], and are

(which are D614) and 'VIC31' (which is G614), were used in standard virus neutralisation assays preformed at the Australian Centre for Disease Preparedness, as described under Methods.

Previous studies in rodents with INO-4800 have demonstrated the induction of humoural and cellular immune responses targeting SARS-CoV-2 spike protein[8]. In this study, ferrets were shown to have developed SARS-CoV-2 neutralising antibody responses following vaccination with INO-4800, demonstrating that ferrets are an appropriate animal to model COVID-19 vaccine immunogenicity, and that this DNA vaccine stimulates an effective B cell response. The overall median $\log_2$ neutralisation titre against the three virus isolates combined was 6.32 (range 4.32 to 8.32). Comparison of the titres by virus isolate (SA01, VIC01, and VIC31)

nature COMMUNICATIONS

ARTICLE

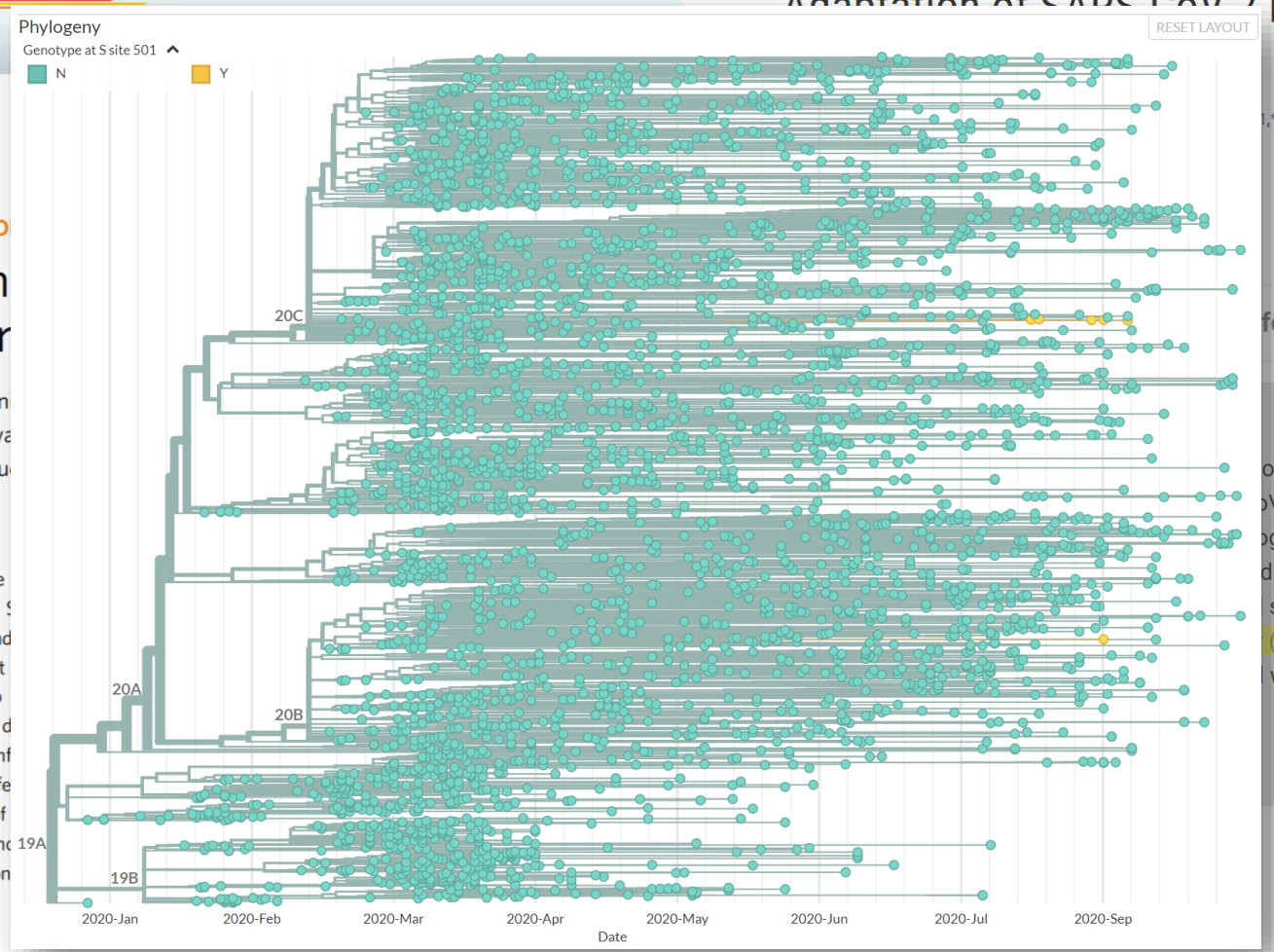
https://doi.org/10.1038/s41467-020-17367-2

SARS-CoV-2 is tran
the air between ferr

Mathilde Richard[1], Adinda Kok[1], Denn
Nisreen M. A. Okba[1], Martje Fentener va
Marion P. G. Koopmans[1], Ron A. M. Fou

SARS-CoV-2, a coronavirus that emerged in late
information about the modes of transmission of S
apply appropriate infection control measures and
SARS-CoV-2 is transmitted efficiently via direct
droplets and/or aerosols) between ferrets, 1 to
respectively. The pattern of virus shedding in the d
is similar to that of the inoculated ferrets and inf
animals, showing that ferrets are productively infe
experimental evidence of robust transmission of
implementation of community-level social distanc
countries in the world and informing decisions on
settings.

N501T



RESEARCH ARTICLE

Adaptation of SARS-CoV-2 in BALB/c mice for testing

[1,*], Hang Fan[1,*], Yong-Qiang Deng[1,*], Yanxiao Wang[2], Yue ...

fo & Metrics eLetters PDF

op medical interventions to treat severe acute
V-2) infections, high on the list are informative
genesis. Gu *et al.* developed a mouse model in
d could cause an inflammatory response and
strain in the mouse appeared to be dependent
(N501Y), within the receptor-binding domain of
was used to study neutralizing antibodies and a

N501Y

https://www.nature.com/articles/s41467-020-17367-2

https://science.sciencemag.org/content/369/6511/1603

Three areas of sequencing priority and focus:

1. Coordinated sequencing through SARS-CoV-2 SPHERES network.
   - Directed studies and deep-dives:
     - Research funding opportunities: Broad Agency Announcements, other possible mechanisms.
   - Public-private partnership + shared interest = tremendous opportunities for applied public health research, innovation and engagement.
   - Initial passive surveillance using existing sequencing efforts.

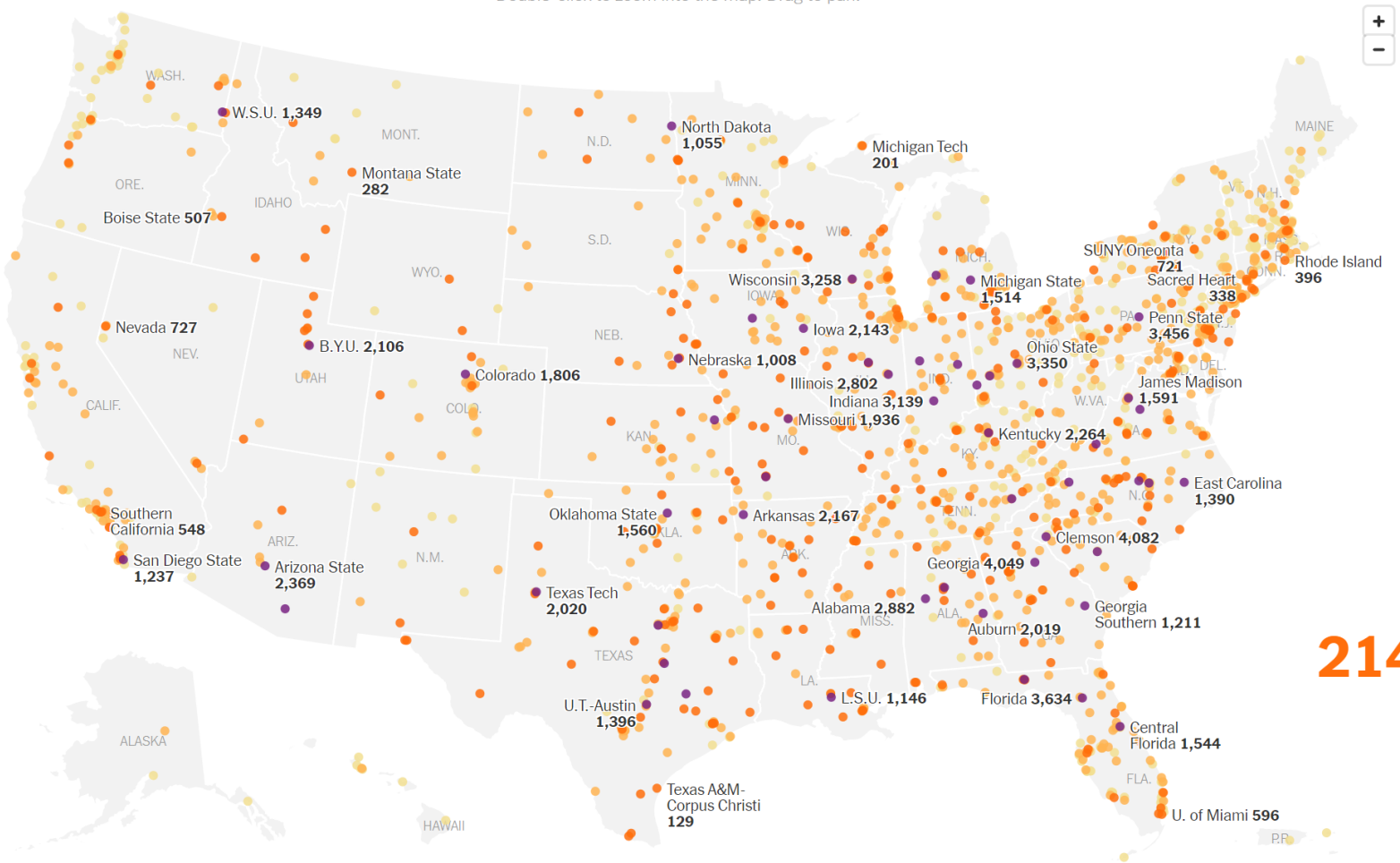
2. S3 + Baseline: structured sequencing for national surveillance.
   - State Strain Surveillance: Sequencing of PHL specimens shipped to CDC, based on existing seasonal flu model.
   - SPHERES Baseline: Contracted sequencing model; synergistic, flexibility, ability to target sequencing.

3. Flexible support for state/local outbreak investigations and surveillance.
   - Cluster investigations, local and regional transmission patterns, vulnerable populations, reinfections, institutes of higher education, extended healthcare, workplace, etc.

Colleges with coronavirus cases since the pandemic began

● 1,000 or more cases ● 100-999 cases ● 10-99 cases ● Fewer than 10 cases

Double-click to zoom into the map. Drag to pan.



Note: Data is as recent as Oct. 22. Only schools with known cases are shown.

By The New York Times Updated Oct. 22, 2020

214,000+ Cases | **1,600+** Colleges

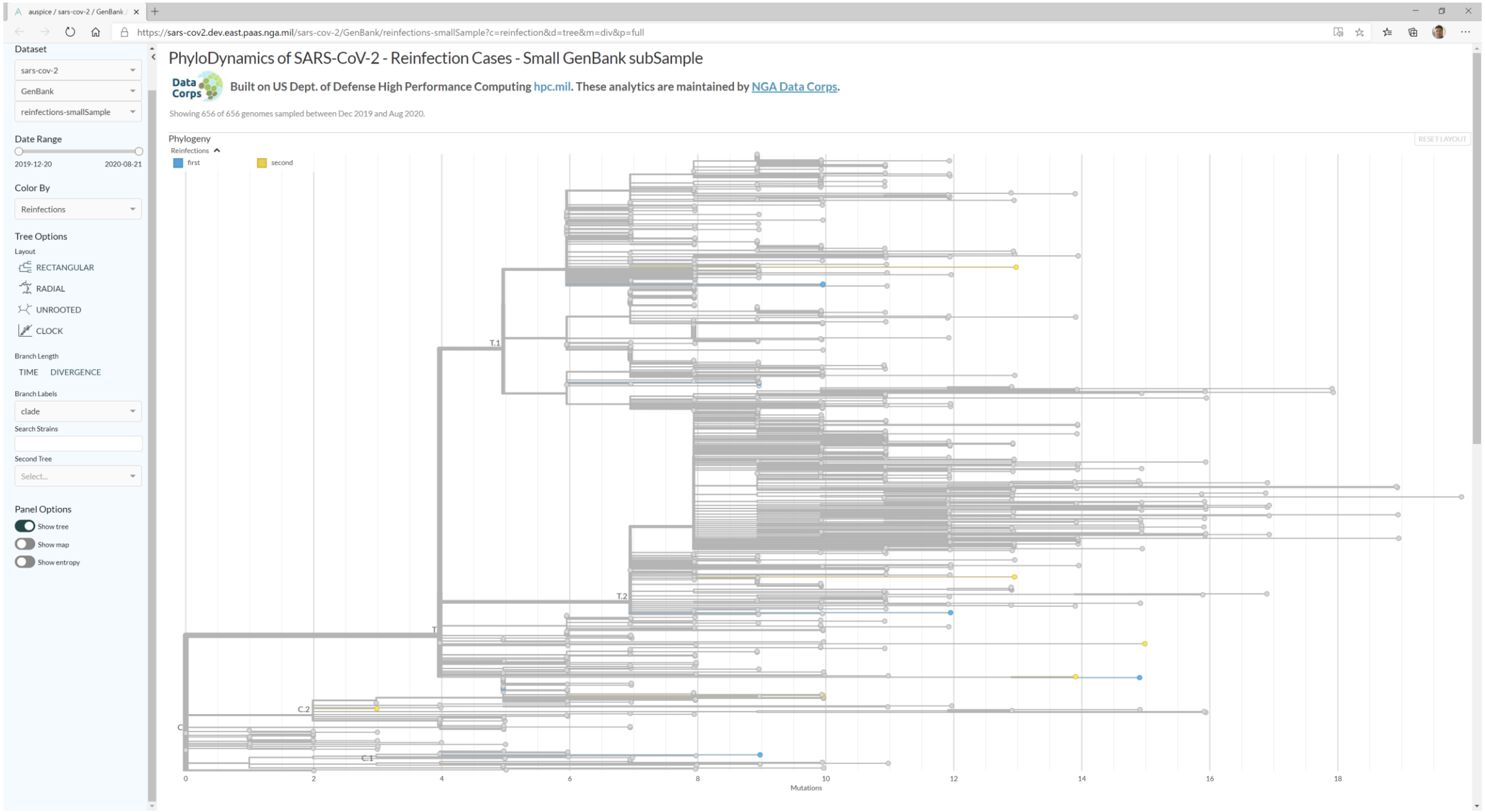
https://www.nytimes.com/interactive/2020/us/covid-college-cases-tracker.html

Suspected SARS-CoV-2 Reinfections

- SPHERES data discussion.
  - How real are these?
    - Intrahost variability, minor population variants.
    - Functional importance: PCR detection vs. clinical infectivity. (HK vs NV CR)
    - Reinfection and normal clearance by an immune host?
    - “Long-haul” patients.
  - How frequently have we seen these types of cases? How best to confirm?
  - How might we optimize SPHERES to help detect these (rare?) events?

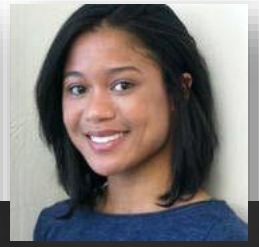
- Investigative Criteria for Suspected Cases of SARS-CoV-2 Reinfection (ICR)
  - https://www.cdc.gov/coronavirus/2019-ncov/php/invest-criteria.html

- Common Investigation Protocol for Investigating Suspected SARS-CoV-2 Reinfection
  - https://www.cdc.gov/coronavirus/2019-ncov/php/reinfection.html


auspice / sars-cov-2 / GenBank /
https://sars-cov2.dev.east.paas.nga.mil/sars-cov-2/GenBank/reinfections-smallSample?c=reinfection&d=tree&m=div&p=full
Dataset
sars-cov-2
GenBank
reinfections-smallSample
Date Range
2019-12-20
2020-08-21
Color By
Reinfections
Tree Options
Layout
RECTANGULAR
RADIAL
UNROOTED
CLOCK
Branch Length
TIME
DIVERGENCE
Branch Labels
clade
Search Strains
Second Tree
Select...
Panel Options
Show tree
Show map
Show entropy
PhyloDynamics of SARS-CoV-2 - Reinfection Cases - Small GenBank subSample
Data Corps
Built on US Dept. of Defense High Performance Computing hpc.mil. These analytics are maintained by NGA Data Corps.
Showing 656 of 656 genomes sampled between Dec 2019 and Aug 2020.
Phylogeny
RESET LAYOUT
Reinfections
first
second
T.1
T.2
T
C.2
C
C.1
0
2
4
6
8
10
12
14
16
18
Mutations

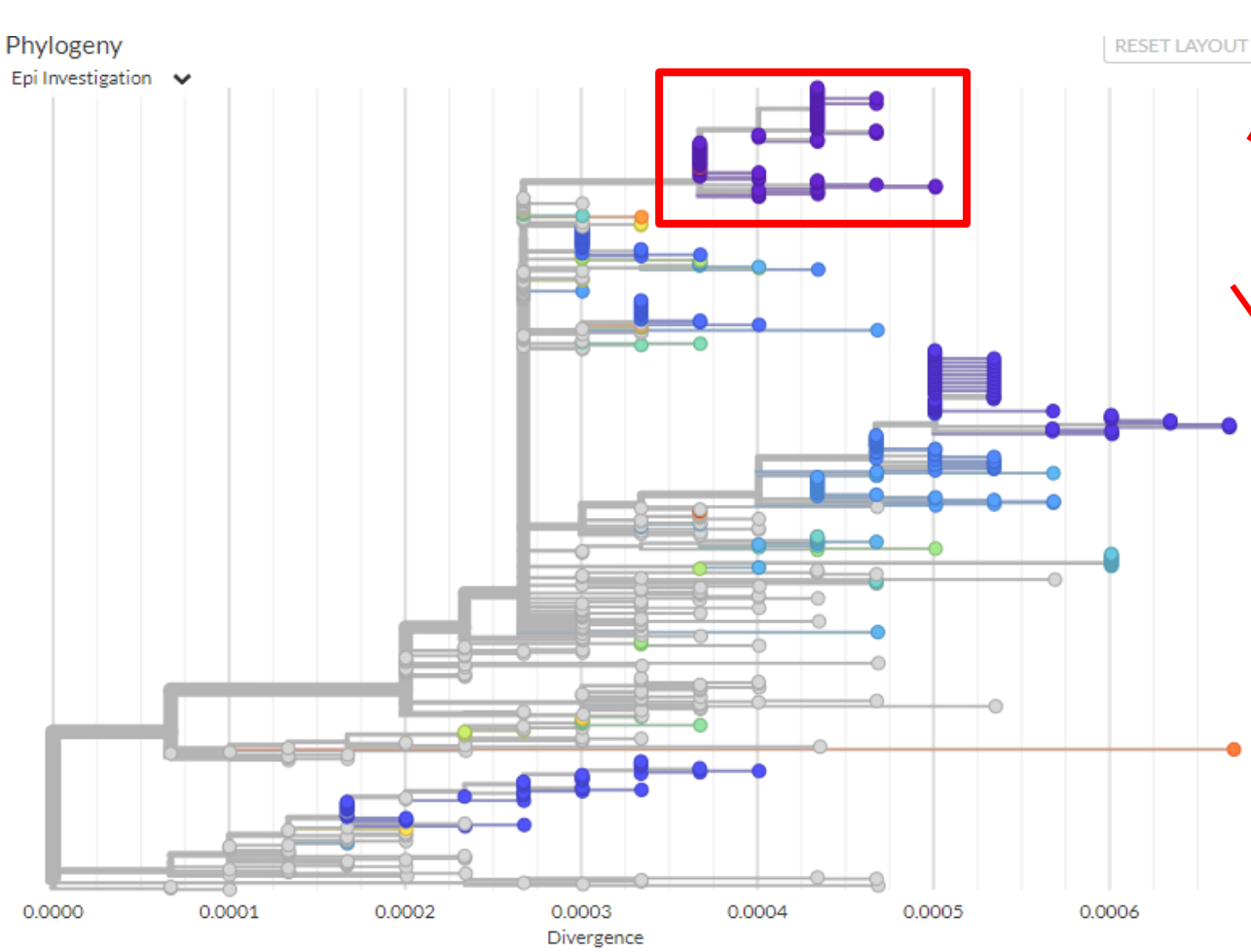
COVID Molecular Epidemiology Toolkit

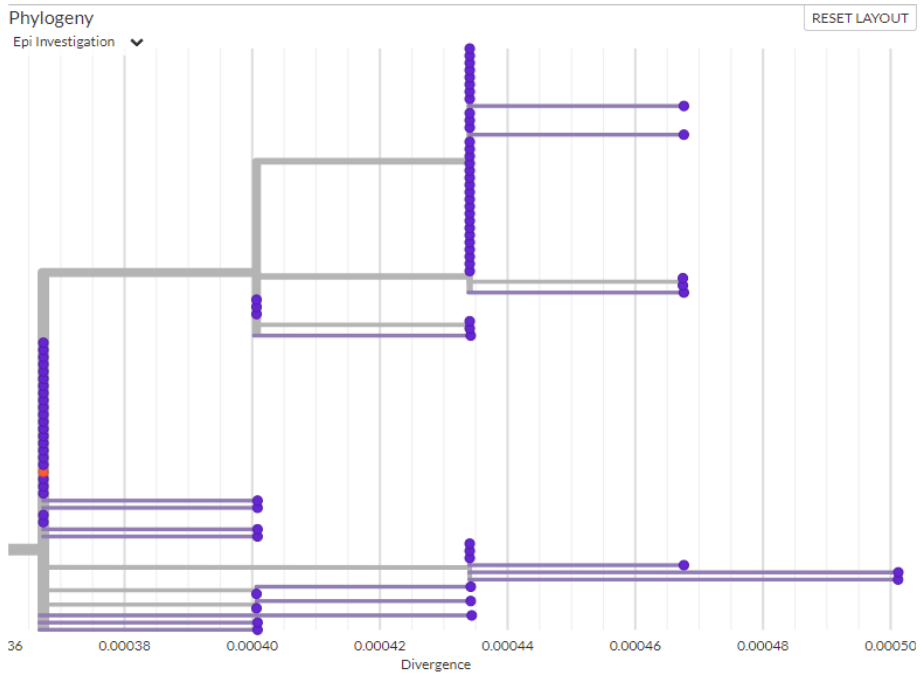
- Modular training for public health epidemiologists, driven by real-world case studies from across US.
- Working with state, local and academic laboratories across SPHERES to develop curriculum.
- Modules on community transmission, LTC/SNF, general considerations/limitations of data.
- Launching virtually soon!



Nancy Chow (NCEZID/DFWED)
Michael Weigand (NCIRD/DBD)

A long-term healthcare facility

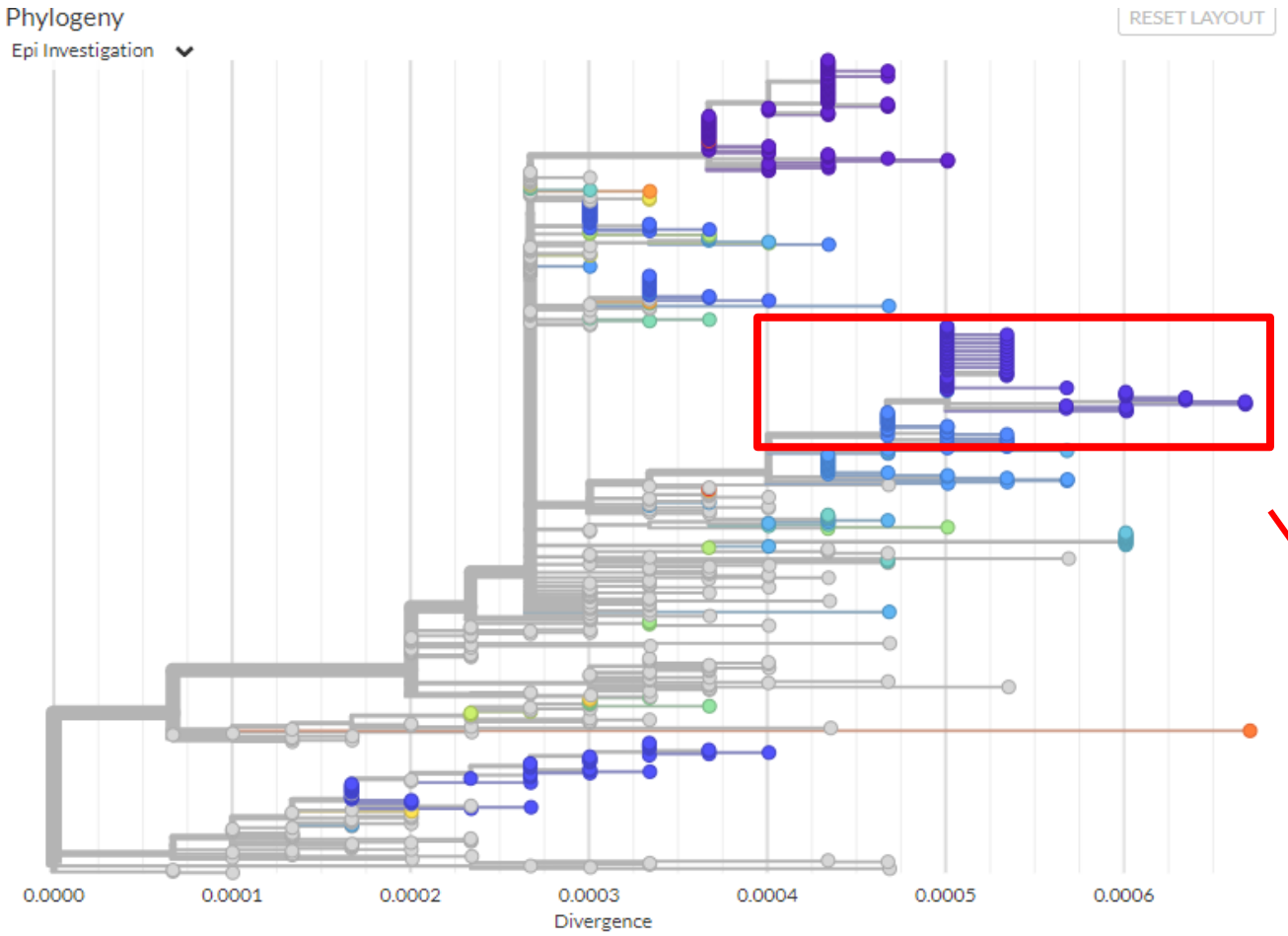




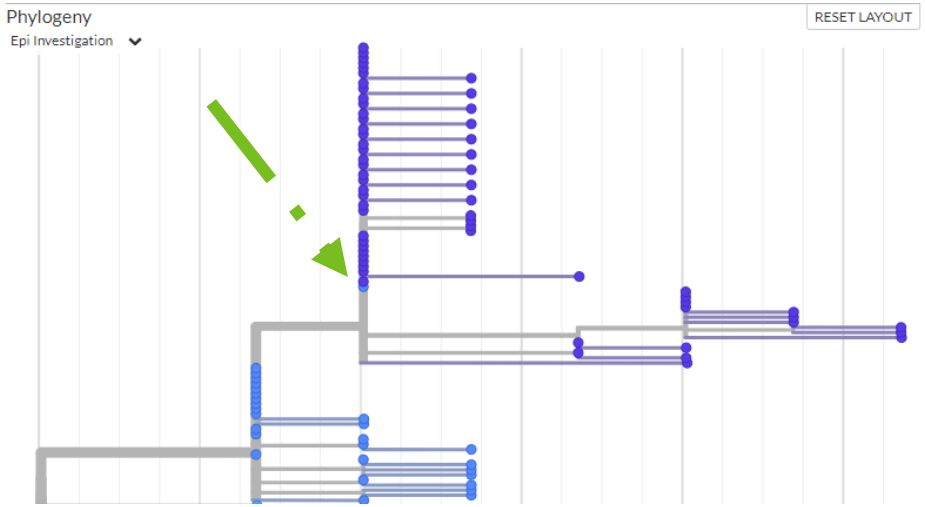
DEPARTMENT OF HEALTH

Thanks, Sean Wang, Ruth Lynfield and Sarah Vedder!

Outbreaks in LTC and correctional facility: linked through household contact?

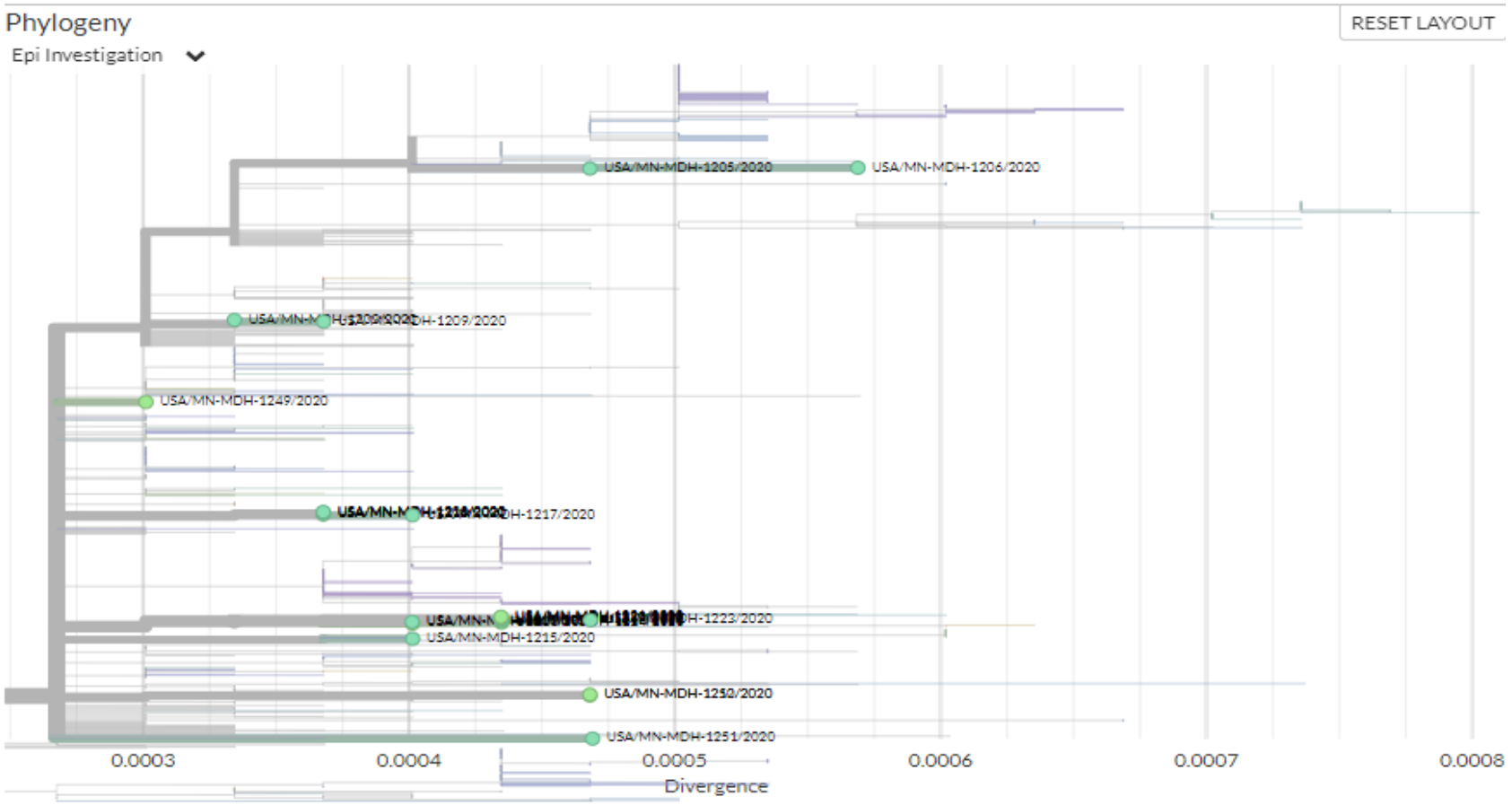


Correction facility



Long-term healthcare facility

Meat processing plants with multiple introductions.



An outdoor concert at a southwest Minnesota

SPHERES

Participatory Model
Pooled Capacity/Expertise
Sampling/Coverage Challenges

~50,000 - 75,000 sequences/yr

CDC S3 + “Baseline”

Based on existing influenza surveillance architecture.
Goal: establish consistent national baseline.
Approach: ~10-20 random positives/state/month->CDC
Important value for context, national trending.

6,250-13,500 sequences/yr + 20-30k more

CDC/SPHL Response

Sequencing/Bioinformatics/Epi Support
Centralized Capacity – CDC/State/Partners
Typically focused: cluster, setting, response.
Some efforts towards comprehensive capture.
Flexible and scalable capacity.
State Consortia: AZ, CA, MN, Others?

20,000+ (overlaps with SPHERES)

GISAID

National Center for Biotechnology Information
NCBI

Consensus Sequences
Filtered Raw Reads
Reference-aligned BAMs

Consistent metadata.

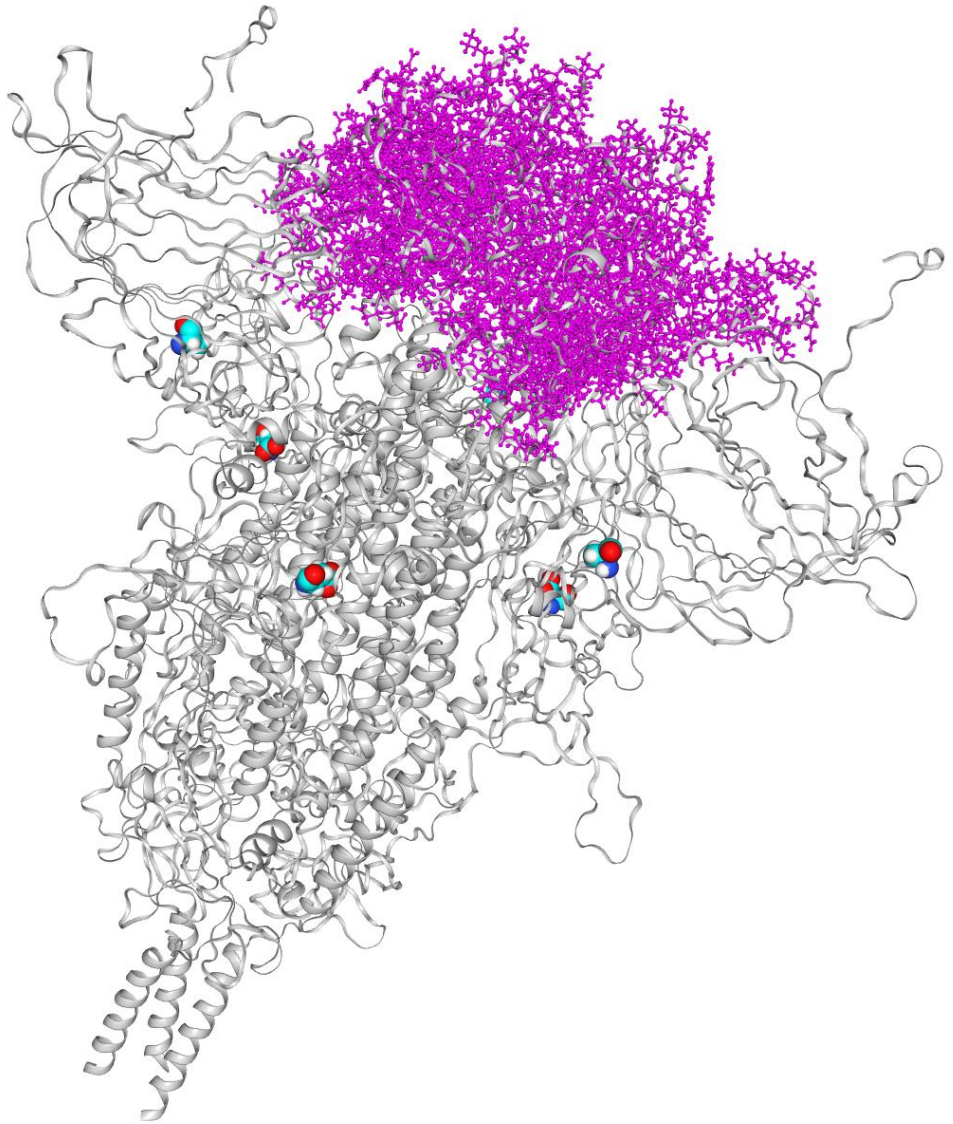
SARS-CoV-2 SPHERES: Takeaways

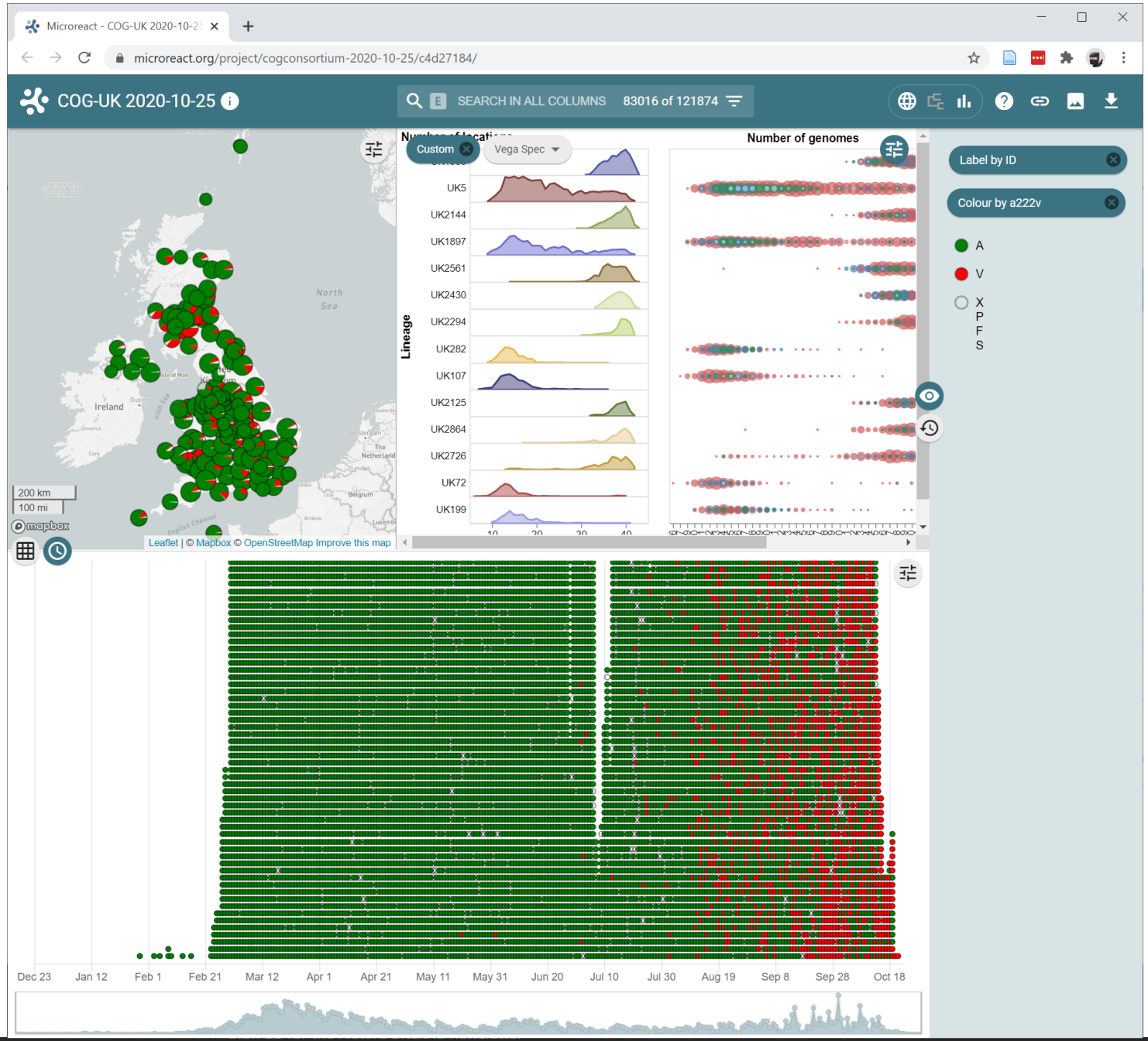
1. SPHERES represents a new model for public-private partnership in genomic epidemiology and infectious disease surveillance.
   - Engages academia, clinical and commercial partners to help generate useful public health data and coordinates the output.
   - By leveraging AMD investments, a national-scale sequencing effort stood up within weeks, and is already returning useful data for response.
2. Supplements traditional, laboratory-based structured surveillance platforms with the speed, flexibility and grassroots innovation of academia and the private sector.
3. SPHERES model can be applied rapidly across a range of different pathogens of concern, both domestically and globally.

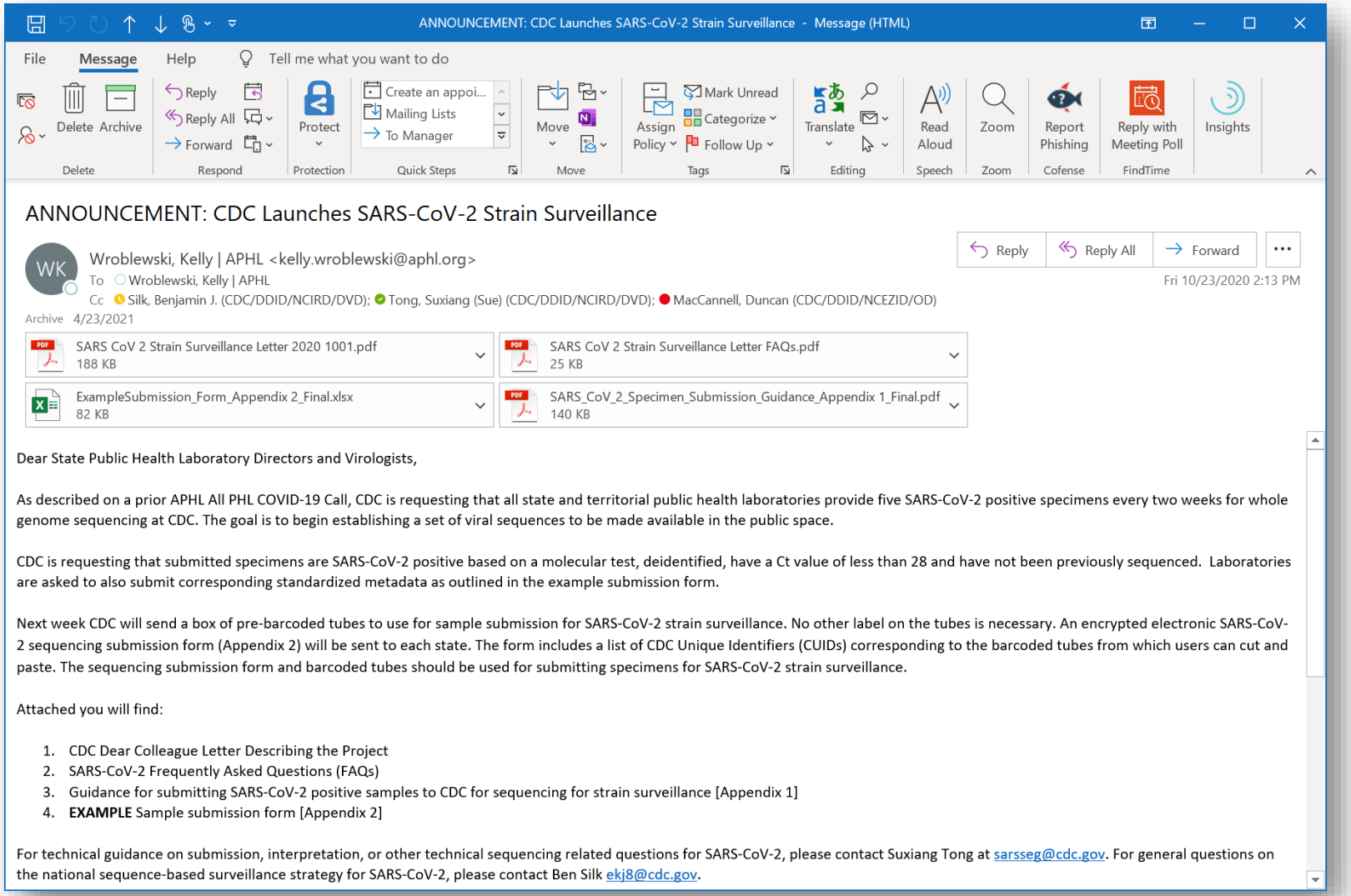
@dmaccannell
dmaccannell@cdc.gov

cdc.gov/amd
@cdc_amd

The findings and conclusions in this presentation are those of the author and do not necessarily represent the official position of the Centers for Disease Control and Prevention. Use of trade names is for identification only and does not imply endorsement by the Centers for Disease Control and Prevention or by the U.S. Department of Health and Human Services.




Microreact - COG-UK 2020-10-25
microreact.org/project/cogconsortium-2020-10-25/c4d27184/
COG-UK 2020-10-25
SEARCH IN ALL COLUMNS
83016 of 121874
Custom
Vega Spec
Number of genomes
Lineage
UK5
UK2144
UK1897
UK2561
UK2430
UK2294
UK282
UK107
UK2125
UK2864
UK2726
UK72
UK199
Label by ID
Colour by a222v
A
V
X
P
F
S
Ireland
North Sea
200 km
100 mi
Leaflet | © Mapbox © OpenStreetMap Improve this map
Dec 23
Jan 12
Feb 1
Feb 21
Mar 12
Apr 1
Apr 21
May 11
May 31
Jun 20
Jul 10
Jul 30
Aug 19
Sep 8
Sep 28
Oct 18



ANNOUNCEMENT: CDC Launches SARS-CoV-2 Strain Surveillance - Message (HTML)

File Message Help Tell me what you want to do

Delete Archive | Reply Reply All Forward | Protect | Create an appoi... Mailing Lists To Manager | Move | Assign Policy Mark Unread Categorize Follow Up | Translate | Read Aloud | Zoom | Report Phishing | Reply with Meeting Poll | Insights

Delete Respond Protection Quick Steps Move Tags Editing Speech Zoom Cofense FindTime

ANNOUNCEMENT: CDC Launches SARS-CoV-2 Strain Surveillance

WK Wroblewski, Kelly | APHL <kelly.wroblewski@aphl.org>

To Wroblewski, Kelly | APHL

Cc Silk, Benjamin J. (CDC/DDID/NCIRD/DVD); Tong, Suxiang (Sue) (CDC/DDID/NCIRD/DVD); MacCannell, Duncan (CDC/DDID/NCEZID/OD)

Archive 4/23/2021

Reply Reply All Forward ...

Fri 10/23/2020 2:13 PM

SARS CoV 2 Strain Surveillance Letter 2020 1001.pdf 188 KB

SARS CoV 2 Strain Surveillance Letter FAQs.pdf 25 KB

ExampleSubmission_Form_Appendix 2_Final.xlsx 82 KB

SARS_CoV_2_Specimen_Submission_Guidance_Appendix 1_Final.pdf 140 KB

Dear State Public Health Laboratory Directors and Virologists,

As described on a prior APHL All PHL COVID-19 Call, CDC is requesting that all state and territorial public health laboratories provide five SARS-CoV-2 positive specimens every two weeks for whole genome sequencing at CDC. The goal is to begin establishing a set of viral sequences to be made available in the public space.

CDC is requesting that submitted specimens are SARS-CoV-2 positive based on a molecular test, deidentified, have a Ct value of less than 28 and have not been previously sequenced. Laboratories are asked to also submit corresponding standardized metadata as outlined in the example submission form.

Next week CDC will send a box of pre-barcoded tubes to use for sample submission for SARS-CoV-2 strain surveillance. No other label on the tubes is necessary. An encrypted electronic SARS-CoV-2 sequencing submission form (Appendix 2) will be sent to each state. The form includes a list of CDC Unique Identifiers (CUIDs) corresponding to the barcoded tubes from which users can cut and paste. The sequencing submission form and barcoded tubes should be used for submitting specimens for SARS-CoV-2 strain surveillance.

Attached you will find:

1. CDC Dear Colleague Letter Describing the Project
2. SARS-CoV-2 Frequently Asked Questions (FAQs)
3. Guidance for submitting SARS-CoV-2 positive samples to CDC for sequencing for strain surveillance [Appendix 1]
4. **EXAMPLE** Sample submission form [Appendix 2]

For technical guidance on submission, interpretation, or other technical sequencing related questions for SARS-CoV-2, please contact Suxiang Tong at sarsseg@cdc.gov. For general questions on the national sequence-based surveillance strategy for SARS-CoV-2, please contact Ben Silk ekj8@cdc.gov.

- Five (5) random, current SARS-CoV-2 positive samples per state every two weeks. (~6500+/yr)
- Not previously sequenced. Ct <28. Standardized data collection and submission.
- Centralized sequencing and bioinformatic analysis.