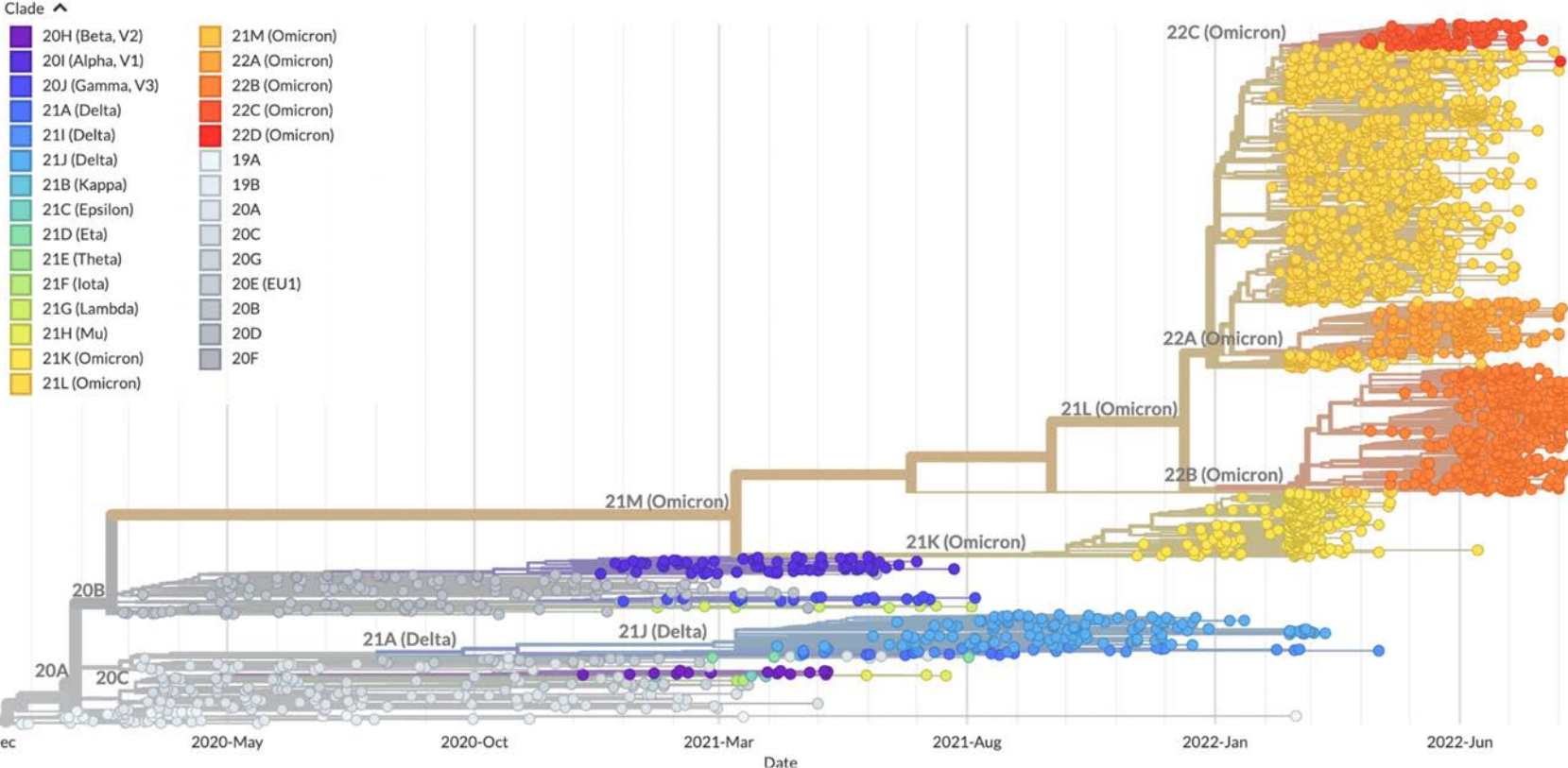


High-res wastewater sequencing uncovers early, cryptic SARS-CoV-2 variant spread

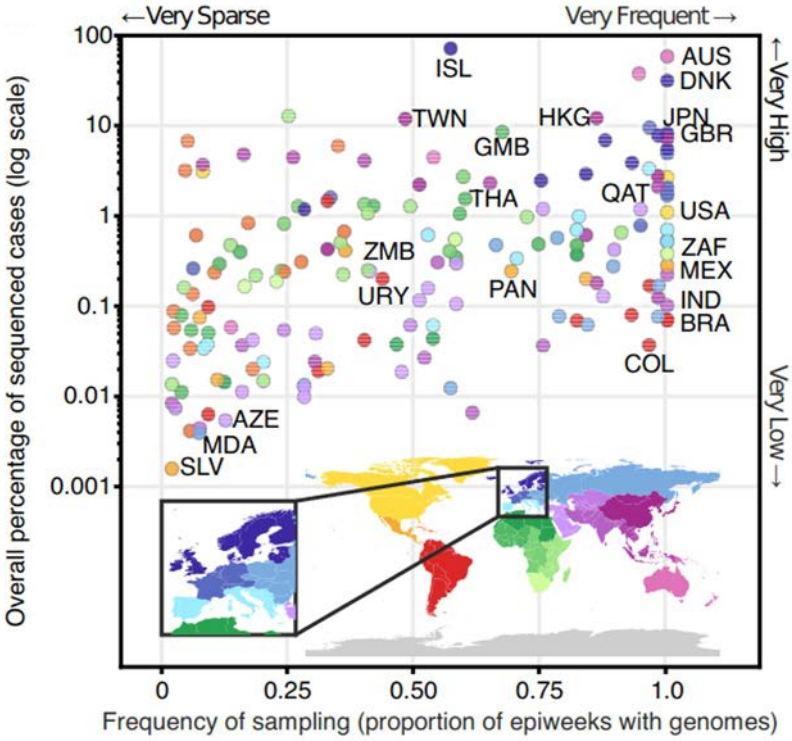
Smruthi Karthikeyan, PhD



SARS-CoV-2 continues to evolve

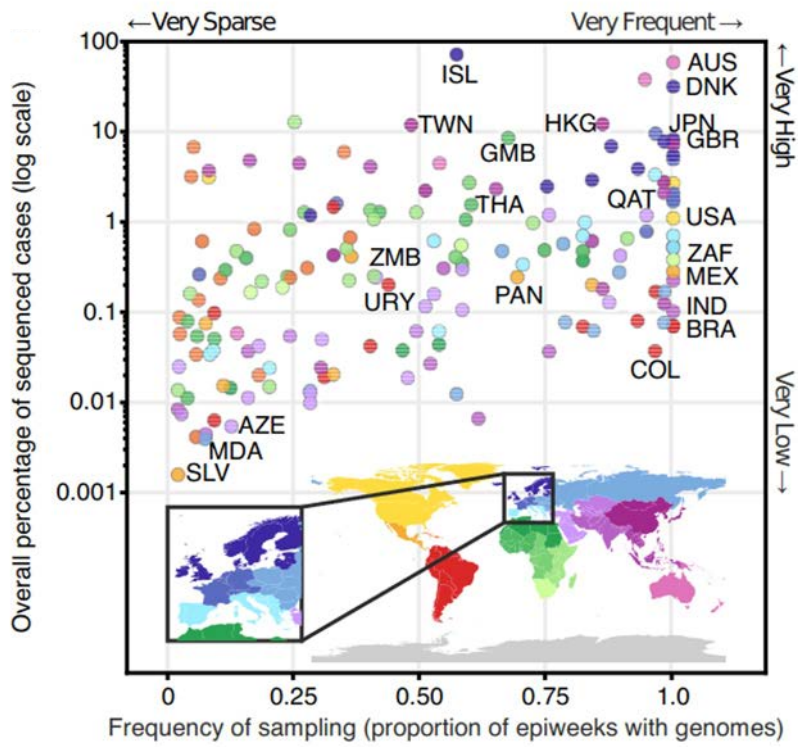


Clinical sampling blind spots

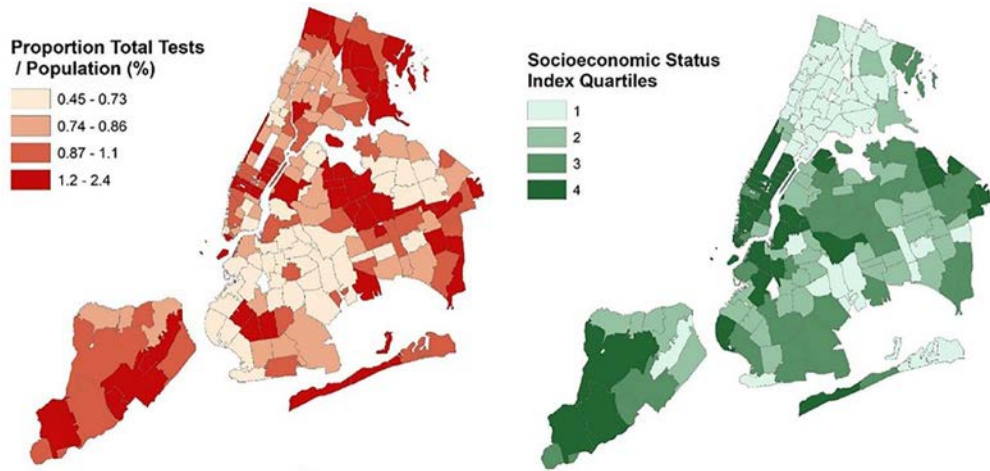


Brito, Semenova, Dudas et al., 2021

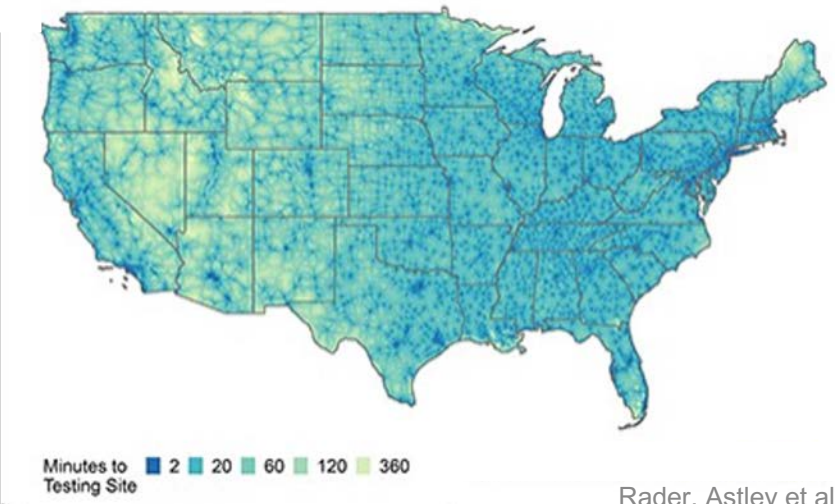
Clinical sampling blind spots



Brito, Semenova, Dudas et al., 2021

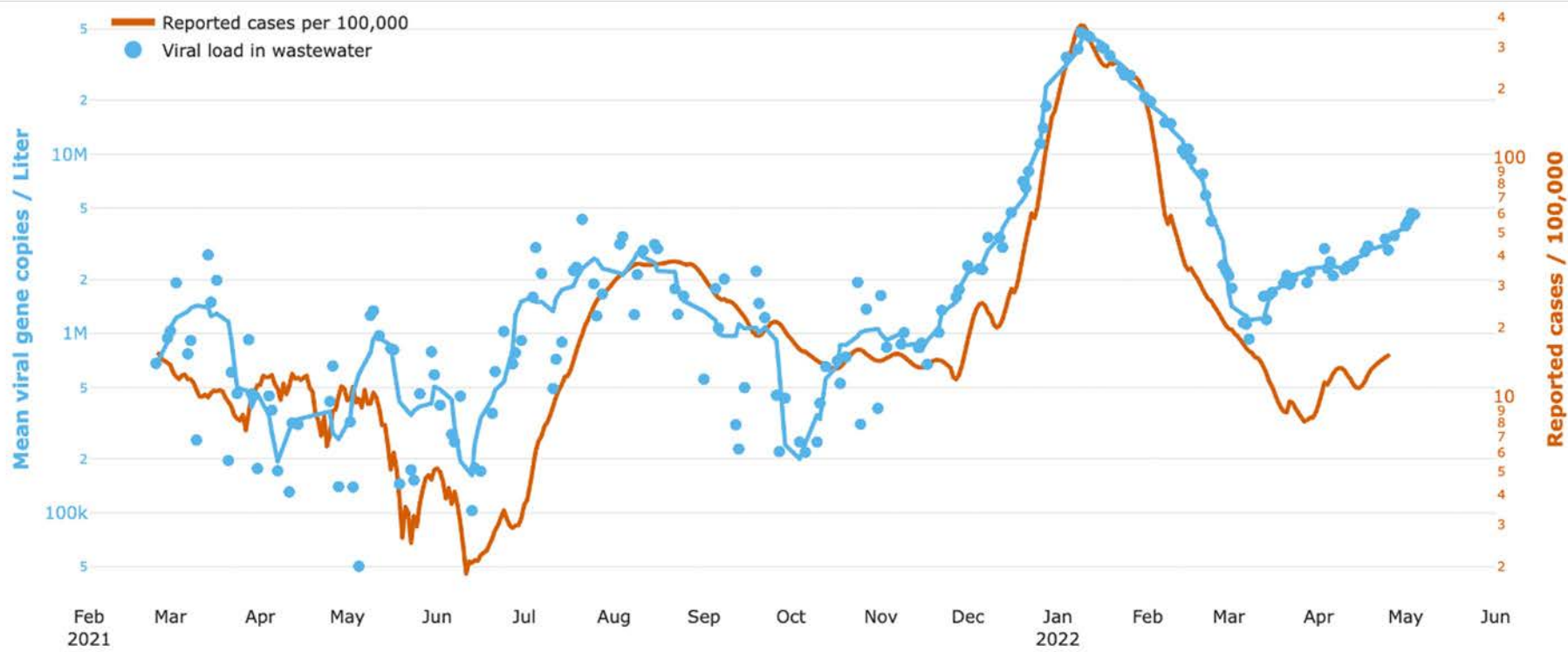


Lieberman-Cribbin et al., 2020



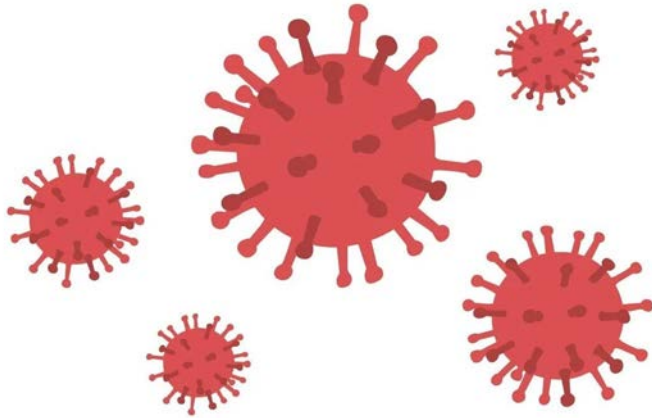
Rader, Astley et al., 2020

Wastewater as a promising alternative



Large-scale wastewater surveillance

BOTTLENECKS



Viral RNA concentration step

Expensive

Laborious

Long turnaround time

Typically only very few samples can be processed in a day

Poor quality sequence data

High-throughput wastewater SARS-CoV-2 viral detection pipeline

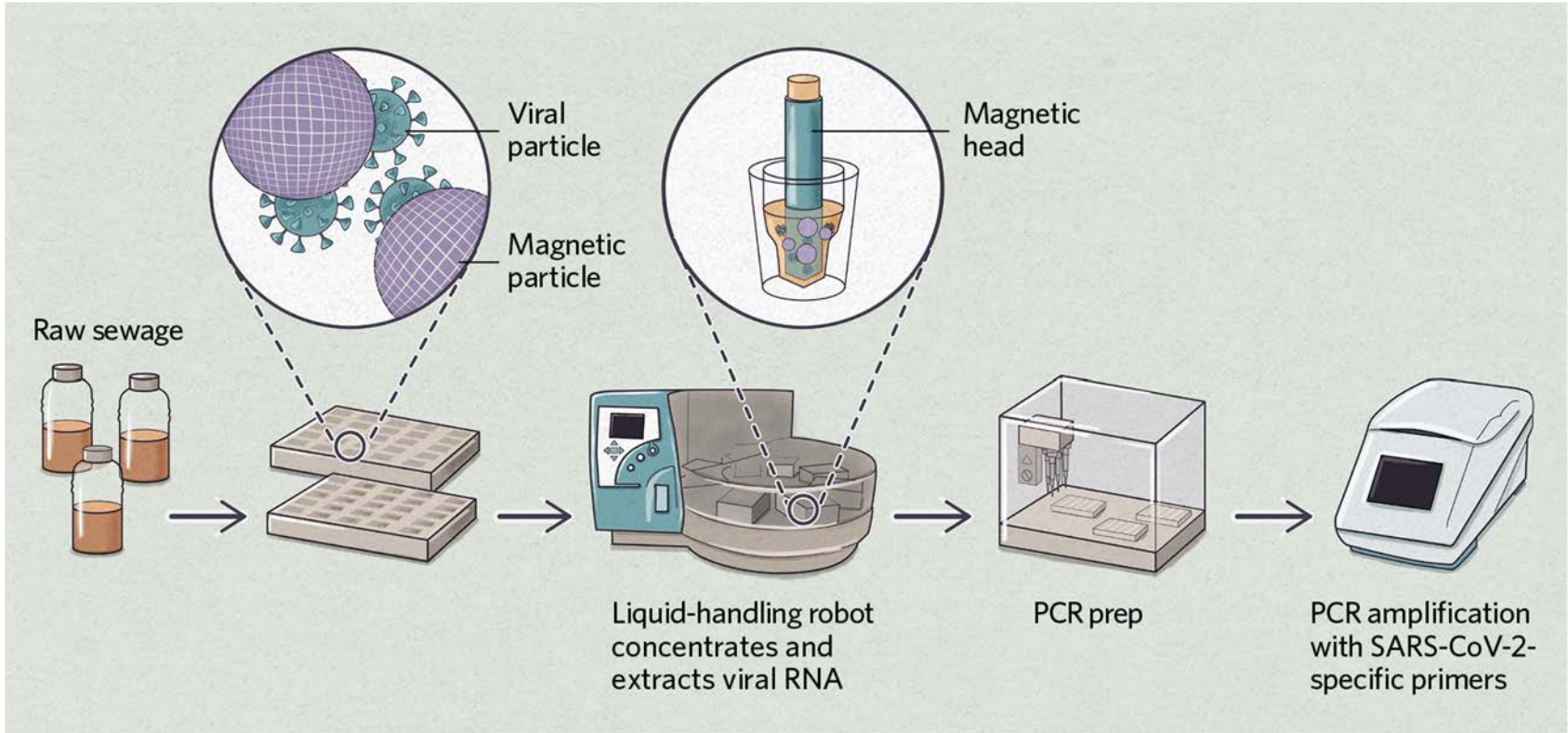


Image courtesy: The Scientist

Karthikeyan et al. mSystems 2021

Wastewater sampling: Campus + County

Sewage testing on UCSD campus

November 2020-Present

RTL-Return to Learn
program

Samples processed: **45,492**

<https://returntolearn.ucsd.edu/dashboard/index.html>

San Diego school districts

October 2020-Present

SASEA-Safer at school early
alert system

Samples processed: **5911**

35 elementary schools/child
care centers

<https://saseasystem.org/>

San Diego County

July 2020-Present

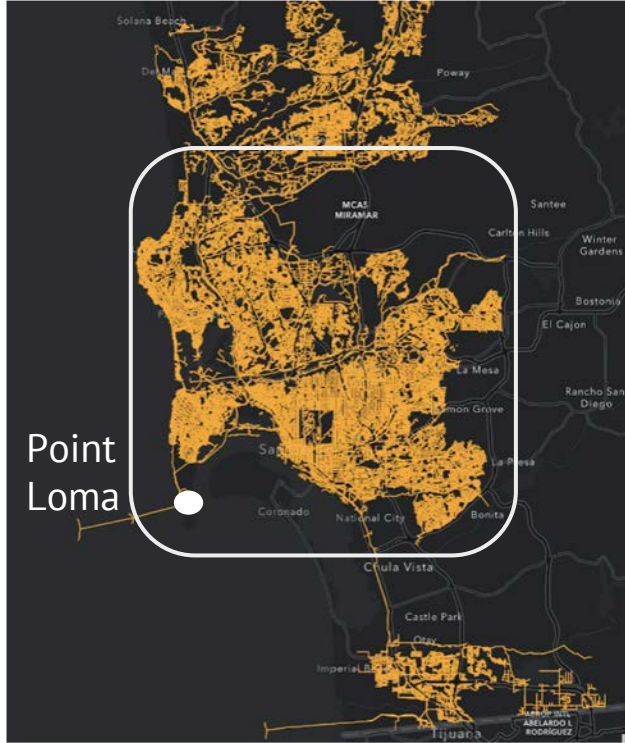
Point Loma, Encina,
South Bay

Samples processed: **1010**

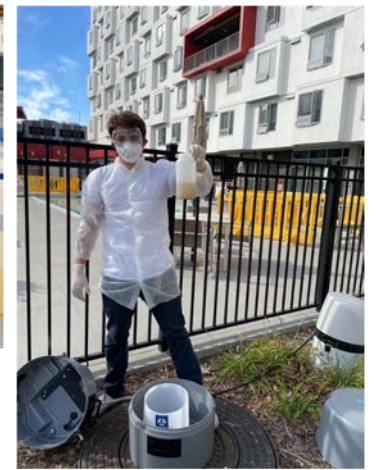
Total Catchment area :
~2.9 million residents

Total: 52,413

County-level wastewater surveillance



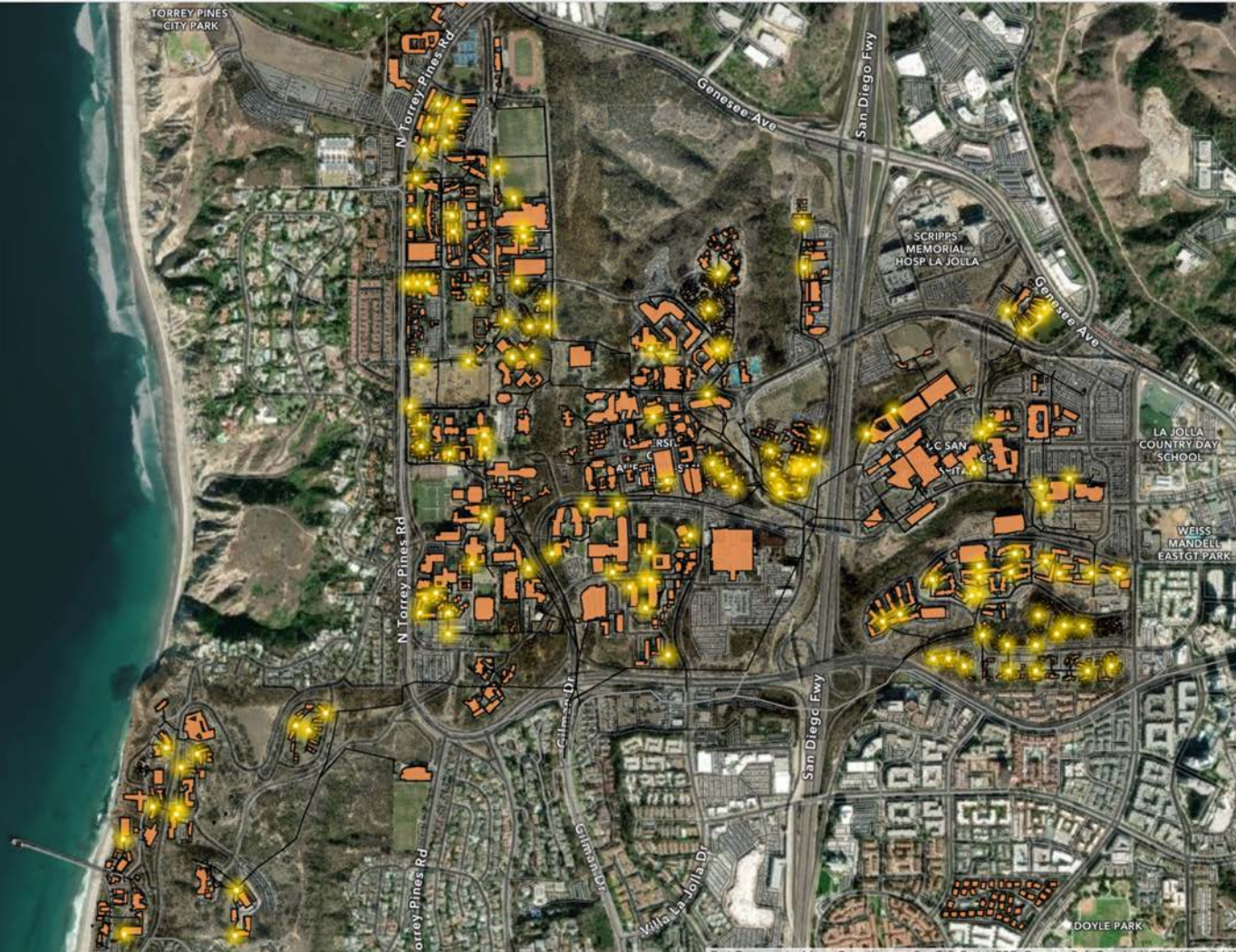
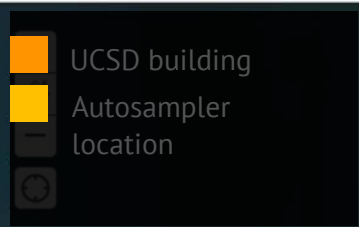
Point Loma: catchment 2.2 million



131 active autosamplers covering ~360 residence and non-residence buildings

Fall 2021: 18,000 residents in campus housing

AUTOSAMPLER LOCATIONS ON CAMPUS





Wastewater Sample Collection Form

Asset UCSD ID *

C2M003

Scan Sample Barcode *

AS_090

Scan Autotester Barcode

AS090MAR09

Is a Sample Being Collected Today? *

Yes

No

What was the Sampling Method?

Composite

Grab

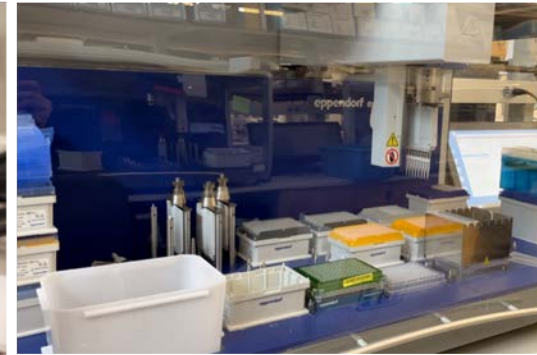
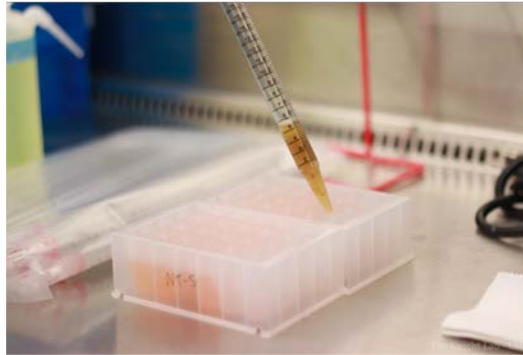
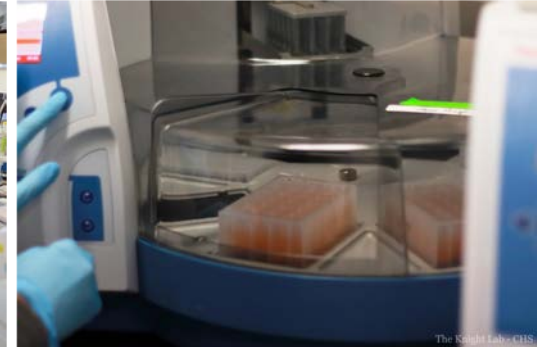
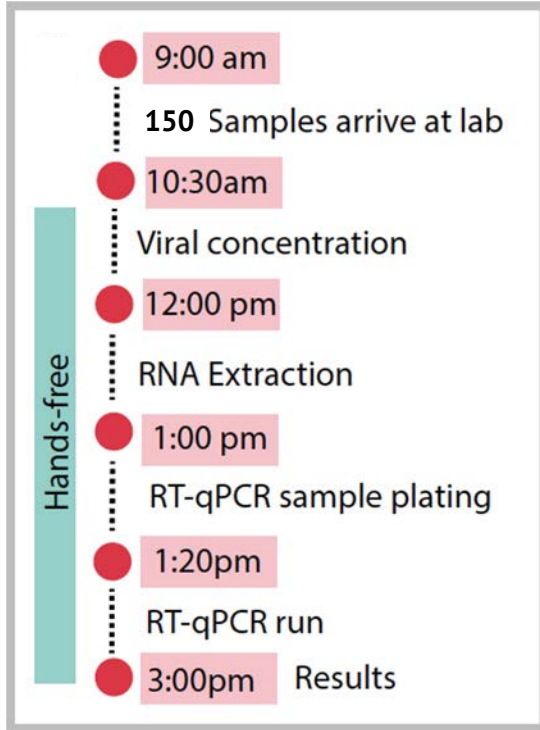


Sample collection and data logging

Autosamplers linked to manhole and can be read by phone app directly

Samples scanned in by field staff as and when they get picked up

A day in the life of a sewage sample

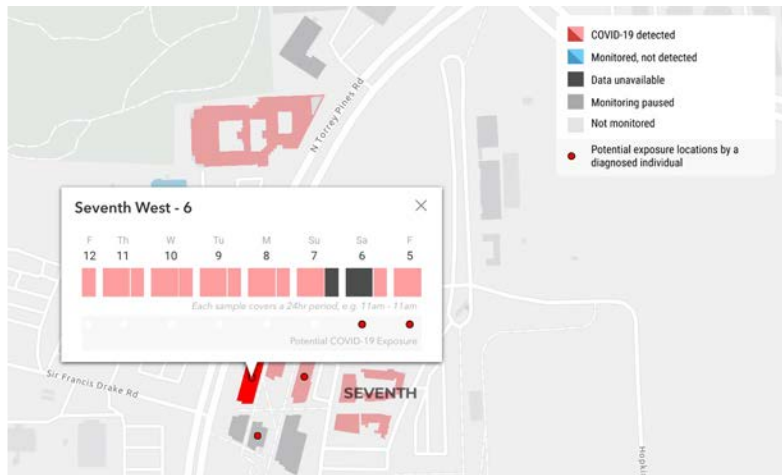


Normalization : PMMoV (fecal indicator)
Recovery: SARS-CoV-2 spike-ins from BSL-3 lab

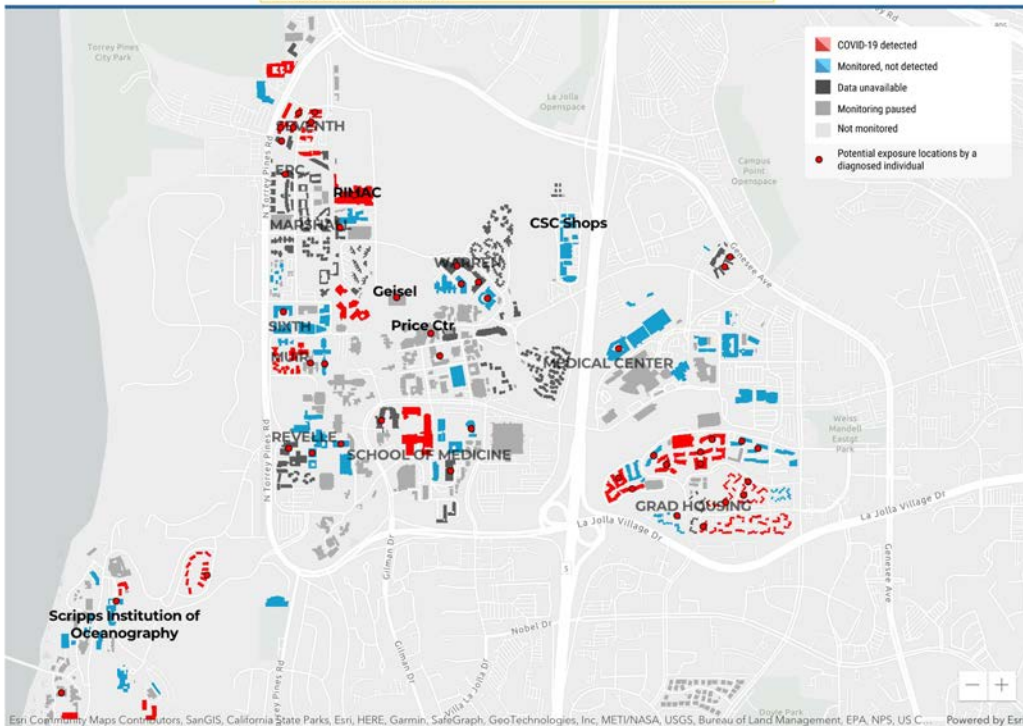
Automated wastewater viral concentration and extraction

[dx.doi.org/10.17504/protocols.io.bshvnb66](https://doi.org/10.17504/protocols.io.bshvnb66)

- Multiple buildings can feed wastewater into one sample
- COVID-19 positive individuals self-reported to be present at potential exposure locations
- Where can I get a COVID-19 Test?



Public dashboard for wastewater



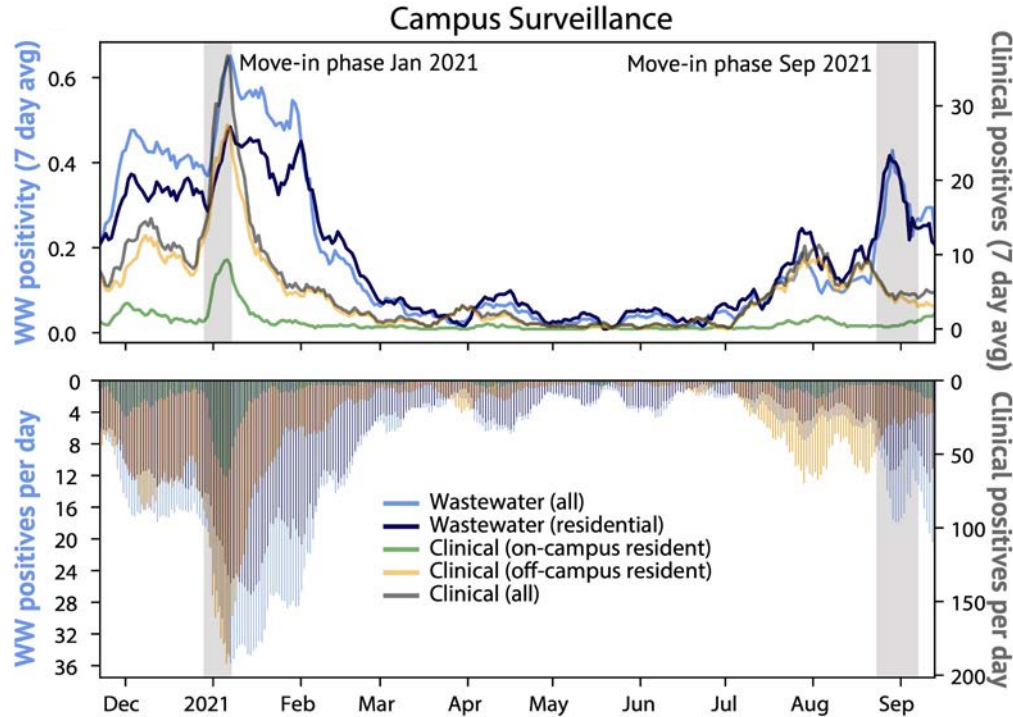
Esri, Community Maps Contributors, SanGIS, California State Parks, Esri, HERE, Garmin, SafeGraph, GeoTechnologies, Inc, METI/NASA, USGS, Bureau of Land Management, EPA, NPS, US Coast Guard, NOAA, and the National Oceanic and Atmospheric Administration. © 2022 Esri. All rights reserved. Powered by Esri.

Additional Information Layers

Potential COVID-19 Exposure Locations COVID-19 Test Kit Vending Machines

	Residential	Non-residential	Total
Positivity Rate: 8/12/22	71.43%	21.43%	44.16%
Seven-day Average	62.62%	17.35%	40.98%

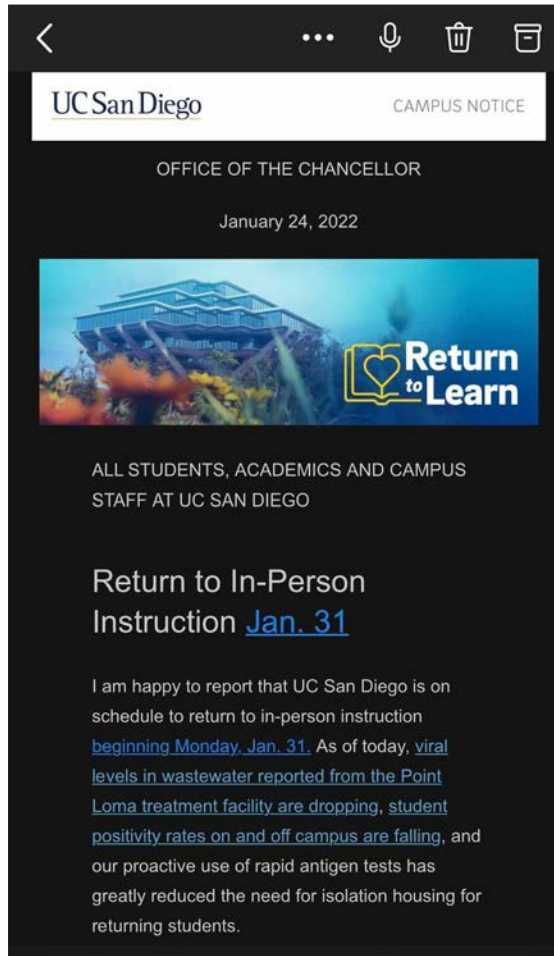
Early detection of 85% of cases on campus



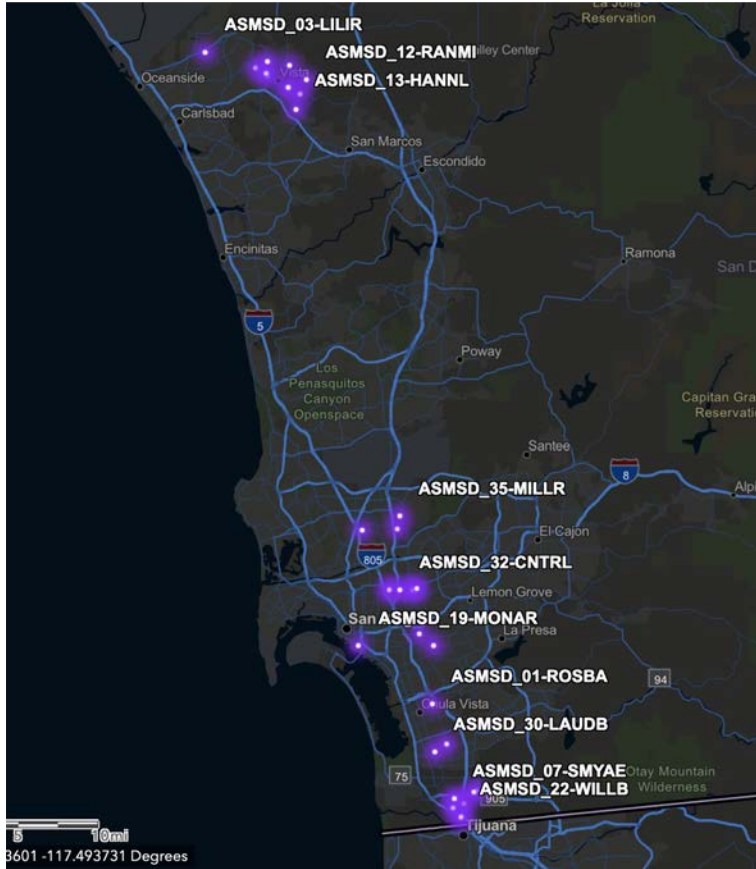
Self-administered test via vending machines

Pilot stage: All on-campus residents
mandated to test weekly

Broader applications



The image is a screenshot of an email notification from UC San Diego. At the top, there is a navigation bar with a back arrow, three dots, a microphone icon, a trash can icon, and a document icon. Below this, the email header includes the UC San Diego logo and the text "CAMPUS NOTICE". The sender is identified as the "OFFICE OF THE CHANCELLOR" and the date is "January 24, 2022". A banner image shows a modern building with a blue sky and the "Return to Learn" logo, which features a yellow heart shape with a book icon inside. The main body of the email is in white text on a dark background. It addresses "ALL STUDENTS, ACADEMICS AND CAMPUS STAFF AT UC SAN DIEGO" and announces the "Return to In-Person Instruction" starting on [Jan. 31](#). The text continues: "I am happy to report that UC San Diego is on schedule to return to in-person instruction [beginning Monday, Jan. 31](#). As of today, [viral levels in wastewater reported from the Point Loma treatment facility are dropping](#), [student positivity rates on and off campus are falling](#), and our proactive use of rapid antigen tests has greatly reduced the need for isolation housing for returning students."



<https://saseasystem.org/>

SD School Districts

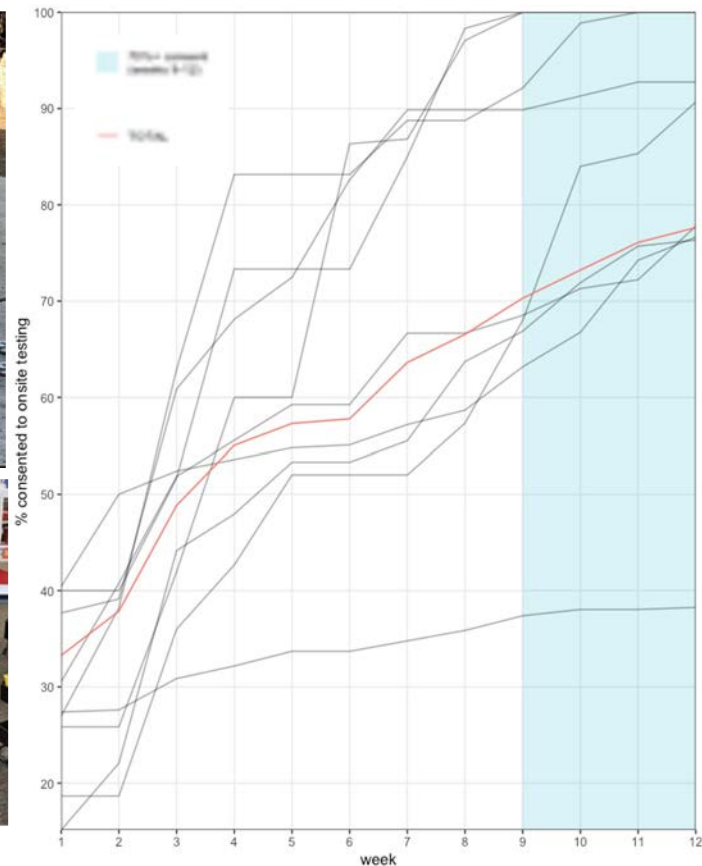
35 sites were selected in ZIP Codes with rates of COVID-19 above the county median and high levels of social vulnerability according to the California Healthy Places Index



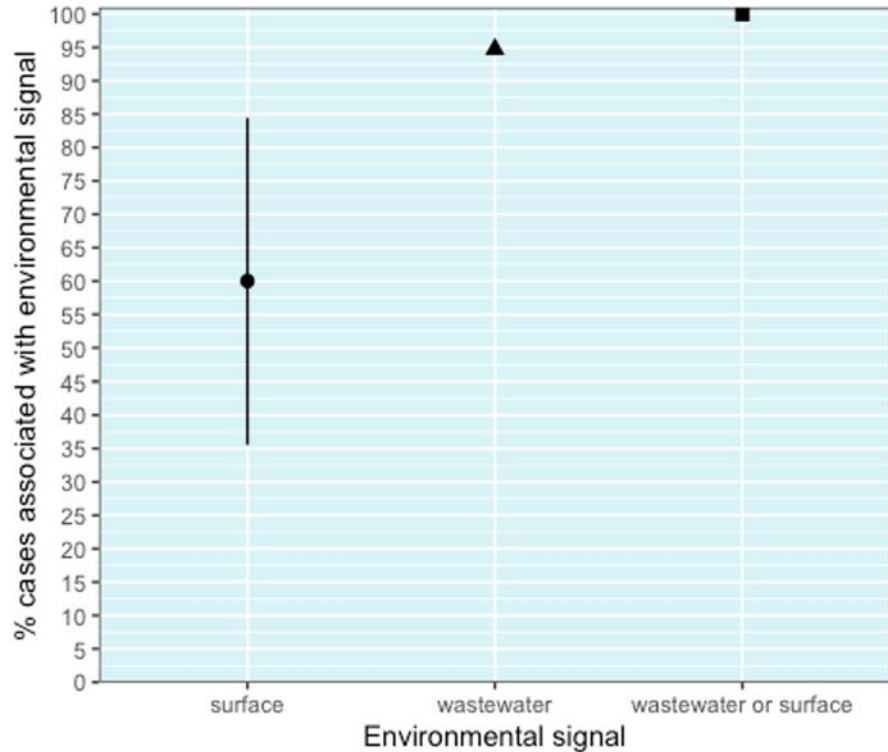
Daily Wastewater monitoring

Daily Surface monitoring

Weekly Diagnostic testing



On-site testing consent rates increased through the study (students + staff)



95% of identified cases were associated with a positive wastewater sample

Fielding-Miller et al.
<https://doi.org/10.1101/2021.10.19.21265226>

Tracking viral variants in wastewater via whole genome seq

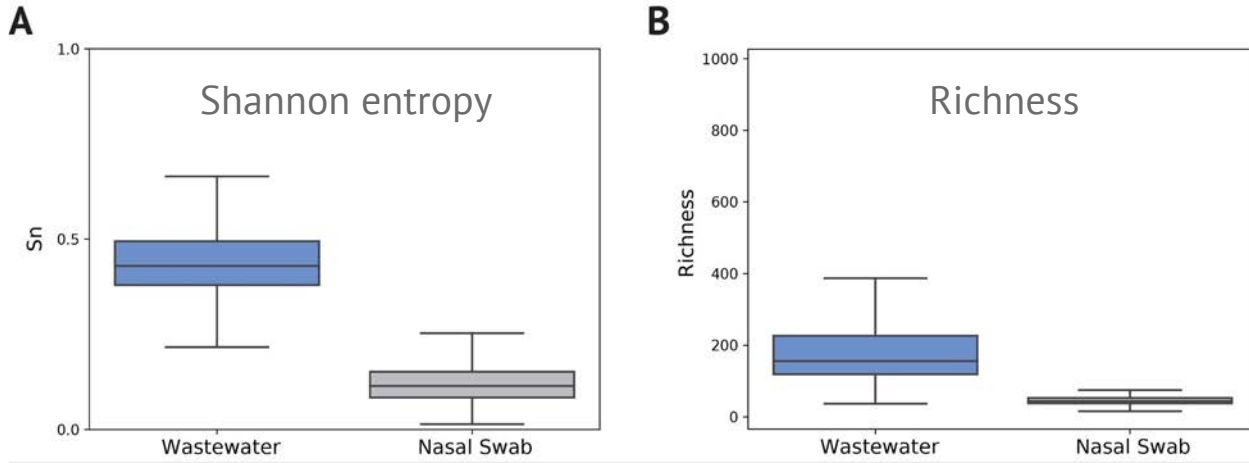
PCR-based wastewater surveillance cannot identify epidemiological transmission links or monitor lineages in the population

Biggest Challenge: obtain good quality genome sequences from environmental samples and accurately discern lineages

All wastewater positives sequenced-miniaturized sequencing approach (1/12th volume) to enable sequencing of >1500 samples a week

Median SARS-CoV-2 Genome Coverage (WW): 96.053%

Higher viral diversity observed in wastewater

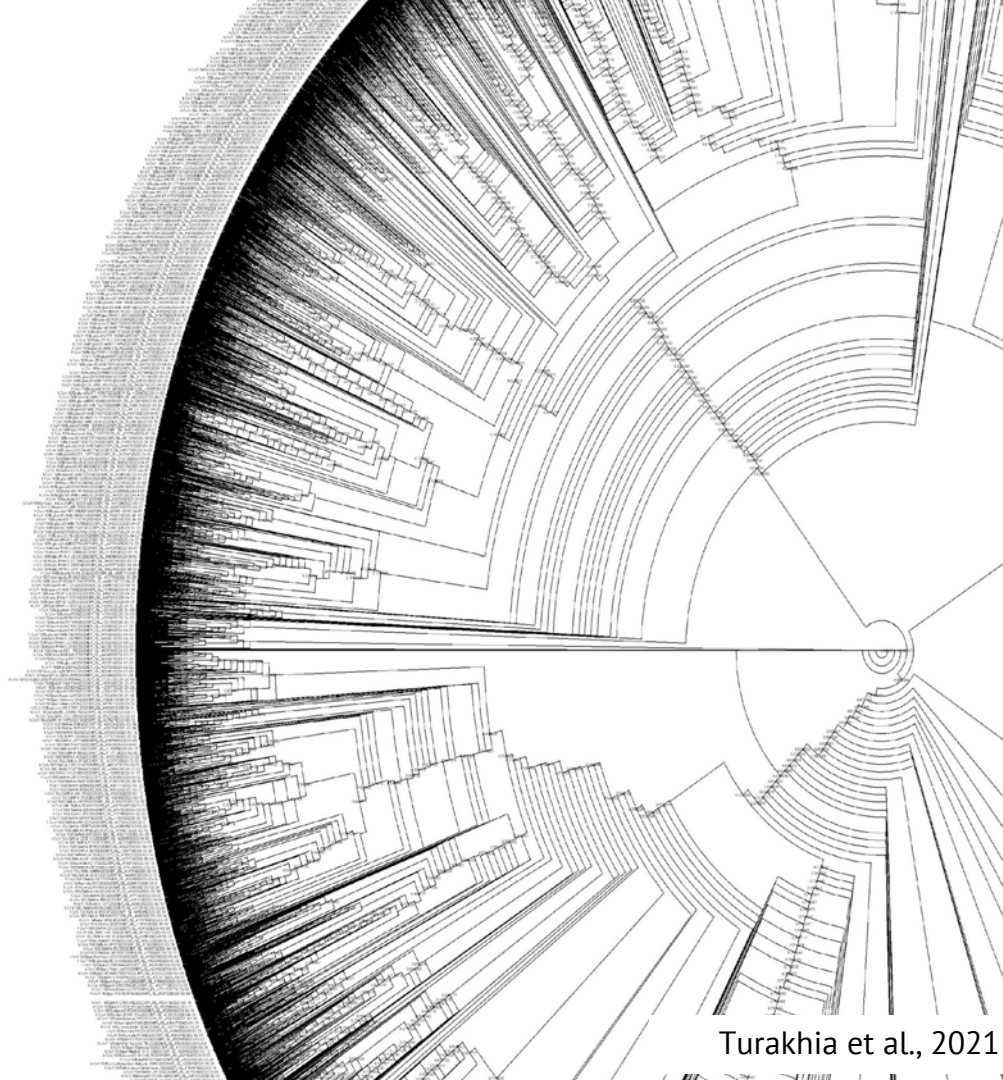


Uncertainty associated with randomly sampling an allele

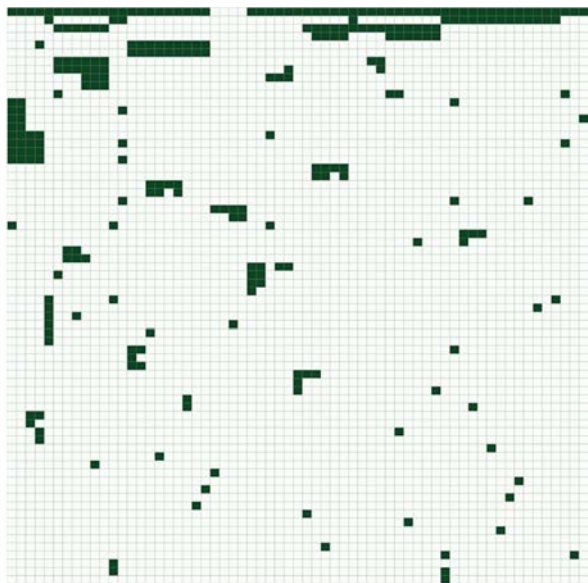
Number of single nucleotide variant (SNV) sites

Existing lineage calling methods require a single consensus sequence to perform assignment

Resolving multiple strains in wastewater

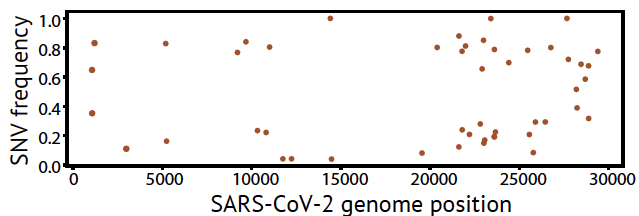


lineage defining mutations



Haplotype

Detection of Single Nucleotide Variants



$$\hat{x} = \underset{\sum_{x=1}^N \mu_i > 0}{\operatorname{argmin}} \|A^T x - b\|_{1W}, \quad \text{where} \quad \|\mu\|_{1W} = \sum_{i=1}^N d_i |\mu_i|$$

Freyja: tool to estimate relative abundance of SARS-CoV-2 lineages from sequencing of mixed-lineage samples

- SNV frequency estimation
- **Depth-weighted de-mixing**


nature

[Explore content](#) ▾ [About the journal](#) ▾ [Publish with us](#) ▾

[nature](#) > [articles](#) > [article](#)

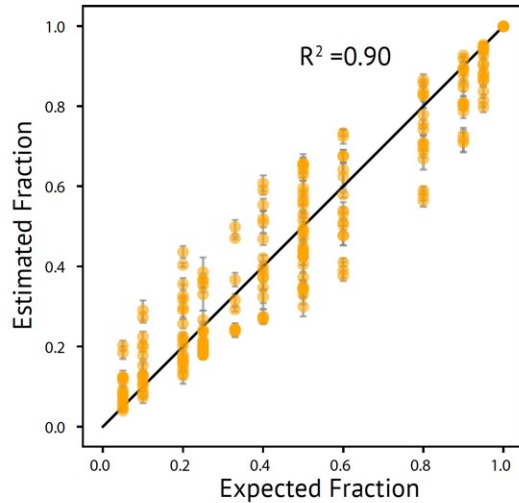
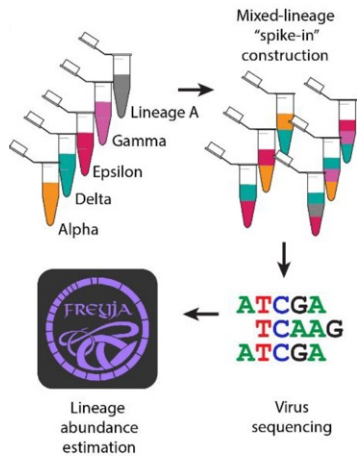
Article | [Open Access](#) | [Published: 07 July 2022](#)

Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission

[Smruthi Karthikeyan](#), [Joshua I. Levy](#), [Peter De Hoff](#), [Greg Humphrey](#), [Amanda Birmingham](#), [Kristen Jepsen](#), [Sawyer Farmer](#), [Helena M. Tubb](#), [Tommy Valles](#), [Caitlin E. Tribelhorn](#), [Rebecca Tsai](#), [Stefan Aigner](#), [Shashank Sathe](#), [Niema Moshiri](#), [Benjamin Henson](#), [Adam M. Mark](#), [Abbas Hakim](#), [Nathan A. Baer](#), [Tom Barber](#), [Pedro Belda-Ferre](#), [Marisol Chacón](#), [Willi Cheung](#), [Evelyn S. Cresini](#), [Emily R. Eisner](#), ... [Rob Knight](#)  [+ Show authors](#)

Nature **609**, 101–108 (2022) | [Cite this article](#)

104k Accesses | **8** Citations | **1095** Altmetric | [Metrics](#)

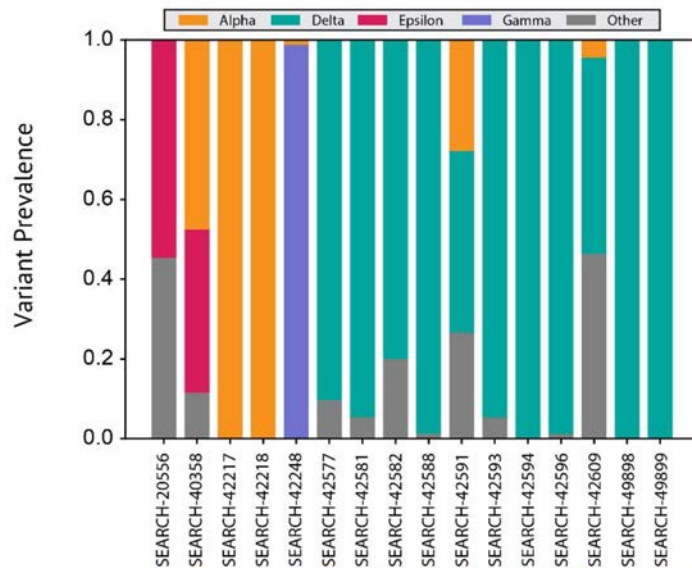


Validation

Strain mixtures (n=96) isolated from mammalian cell culture infected with one of five strains combined at different proportions and sequenced in triplicate to evaluate accuracy of Freyja

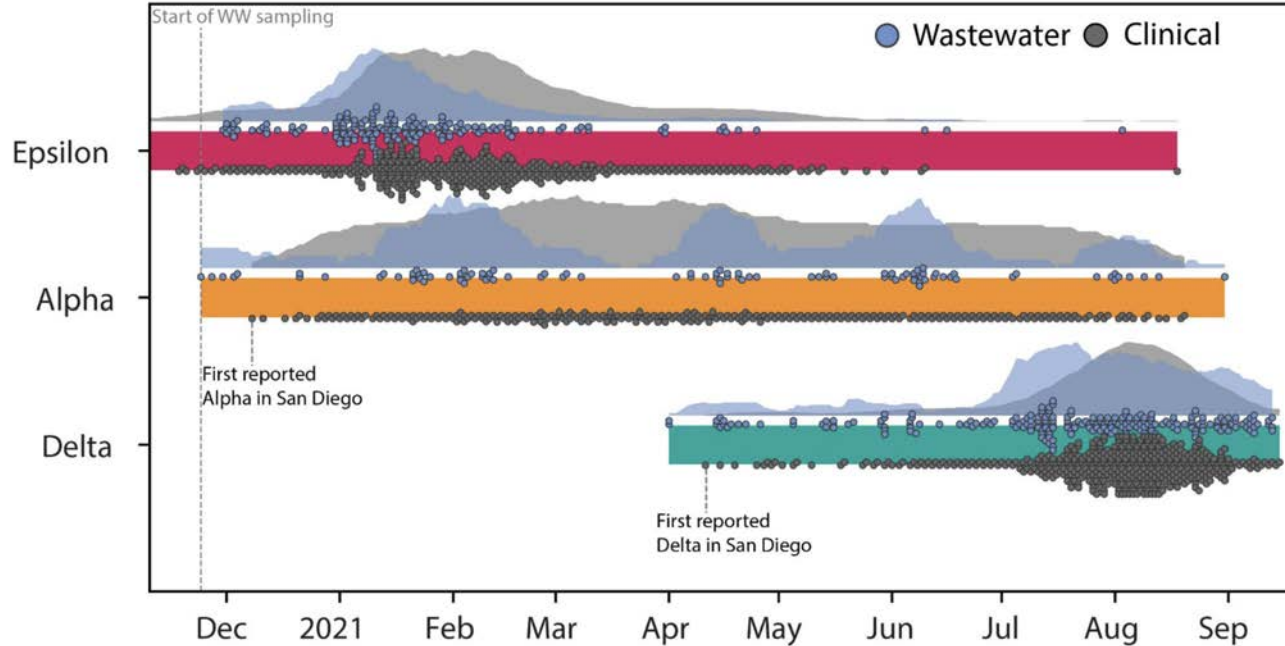
Does it work on Wastewater?

		Sequencing Result		
		+	-	
Overall VOC qPCR Result	+	True Positives 55	False Positive 0	Positive Predictive Value 100%
	-	False Negative 1	True Negative 80	Negative Predictive Value 99%
				Overall Rate of Agreement
				99%



Mutation	Variants	Mutation	Variants
L452R	Delta, Epsilon, Kappa	K417T	Gamma
P681R	Delta, Kappa	K417N	Beta
E484Q	Kappa	DelHV69/70	Alpha
DelY144	Alpha	N501Y	Alpha, Beta, Gamma

San Diego County Surveillance

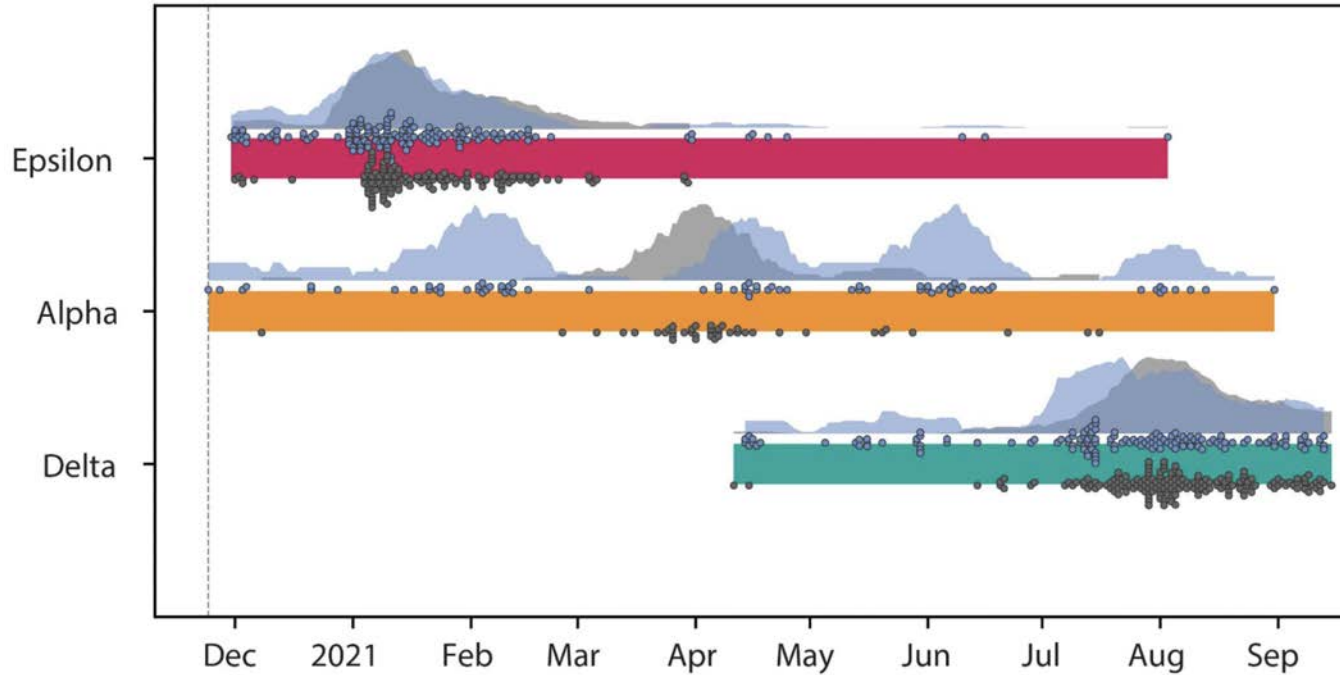


Wastewater samples show earlier appearance of VOCs

Data from 31,149 nasal swab sequences from SD county and 734 wastewater sequences

Timeline and epidemiological curves for VOC detection in county samples

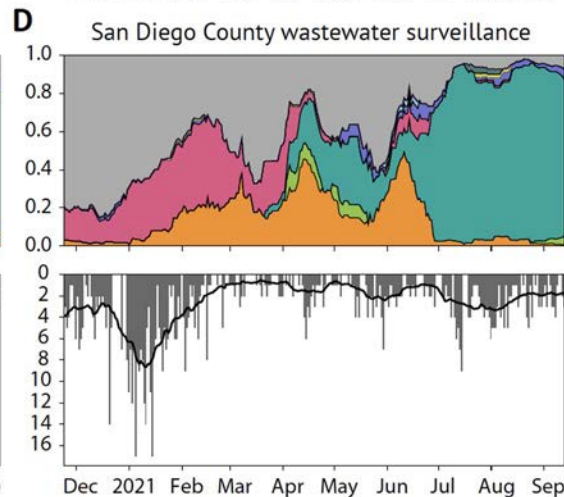
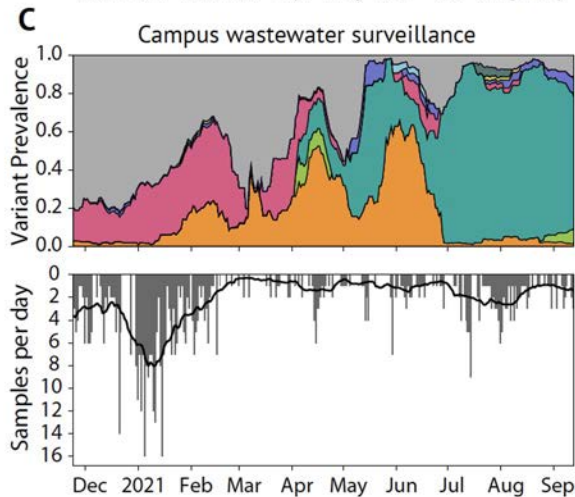
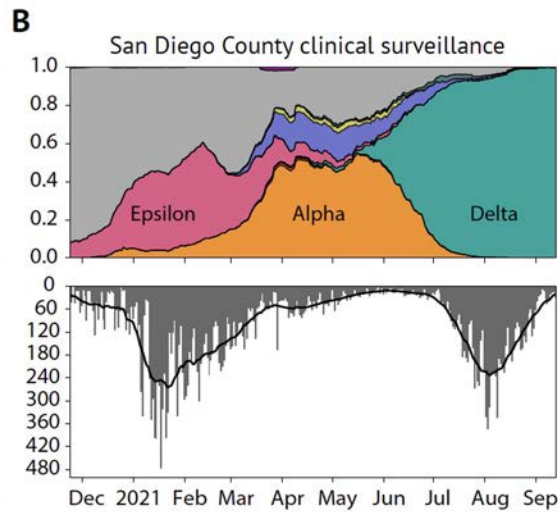
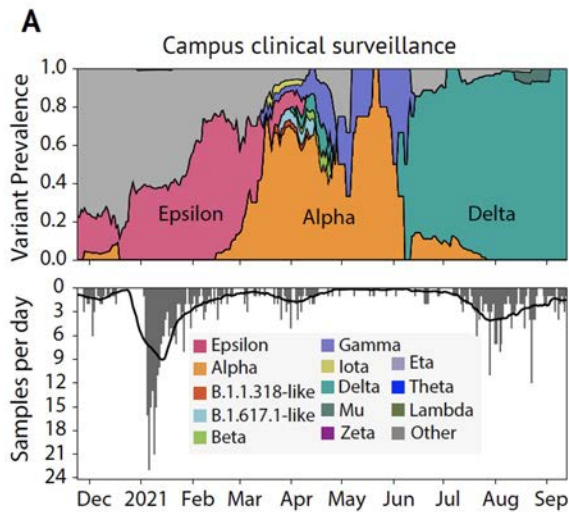
UCSD Campus Surveillance



Timeline and epidemiological curves for VOC detection in the campus samples

UCSD Campus data

Extended cryptic virus spread during early Alpha and Delta waves



WW identifies early, cryptic transmission

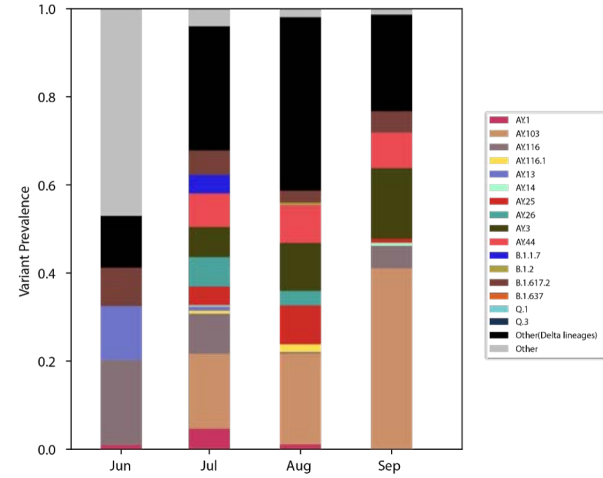
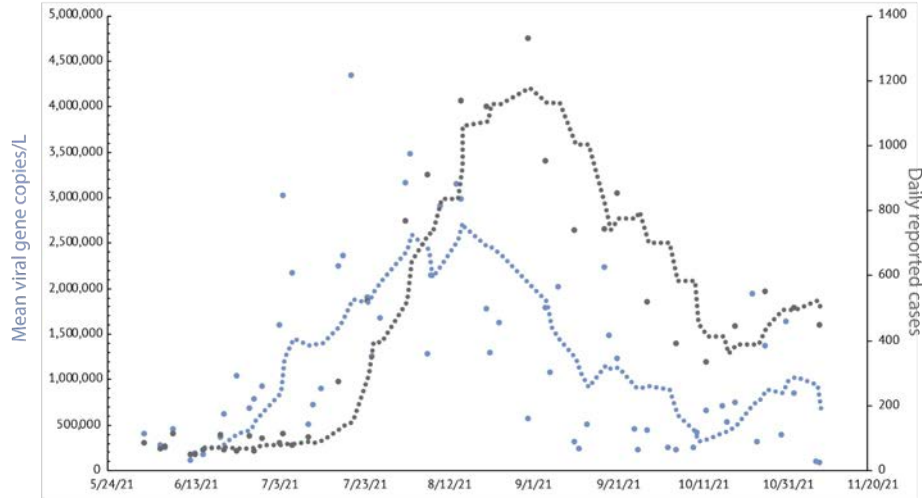
Early identification of emerging strains in WW (even with low abundance in nasal swabs)

“Mu” detected in WW before classification as VOI

Increases likelihood of identifying missing links/areas where transmission may have happened

Rare lineages observed in wastewater

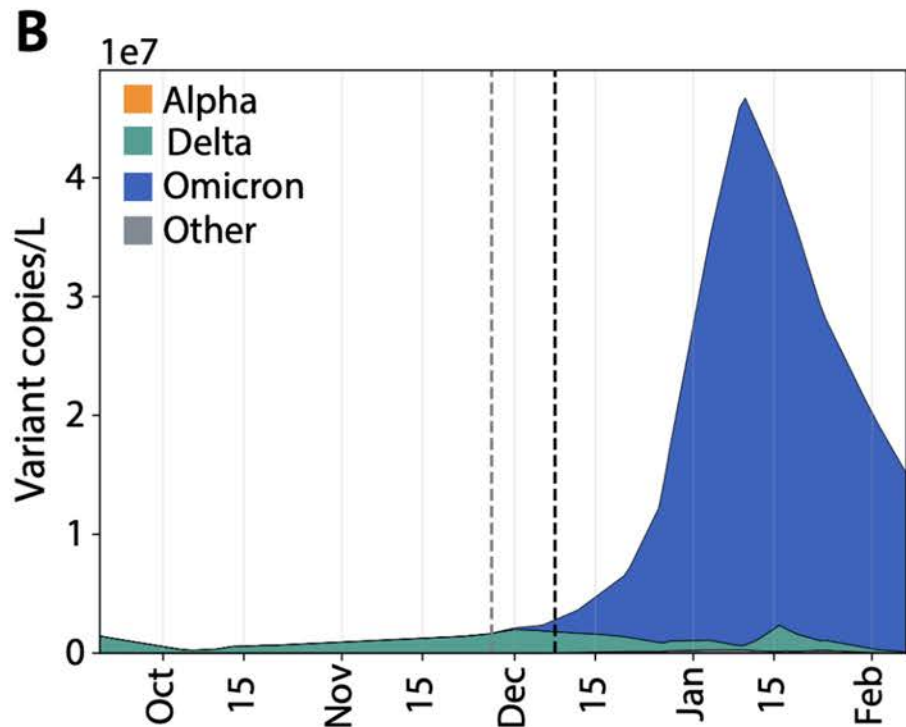
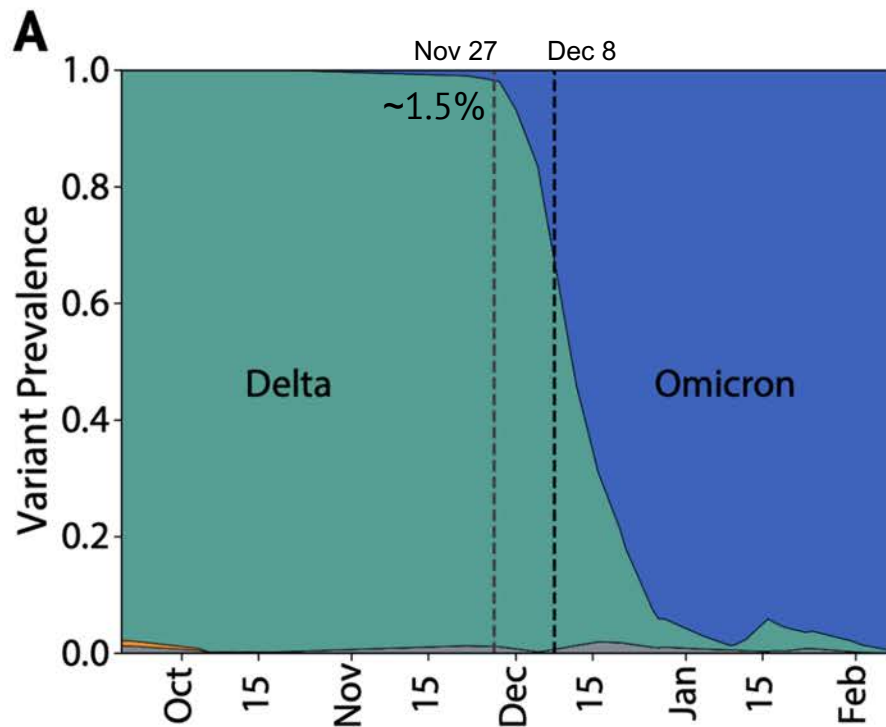
Point Loma

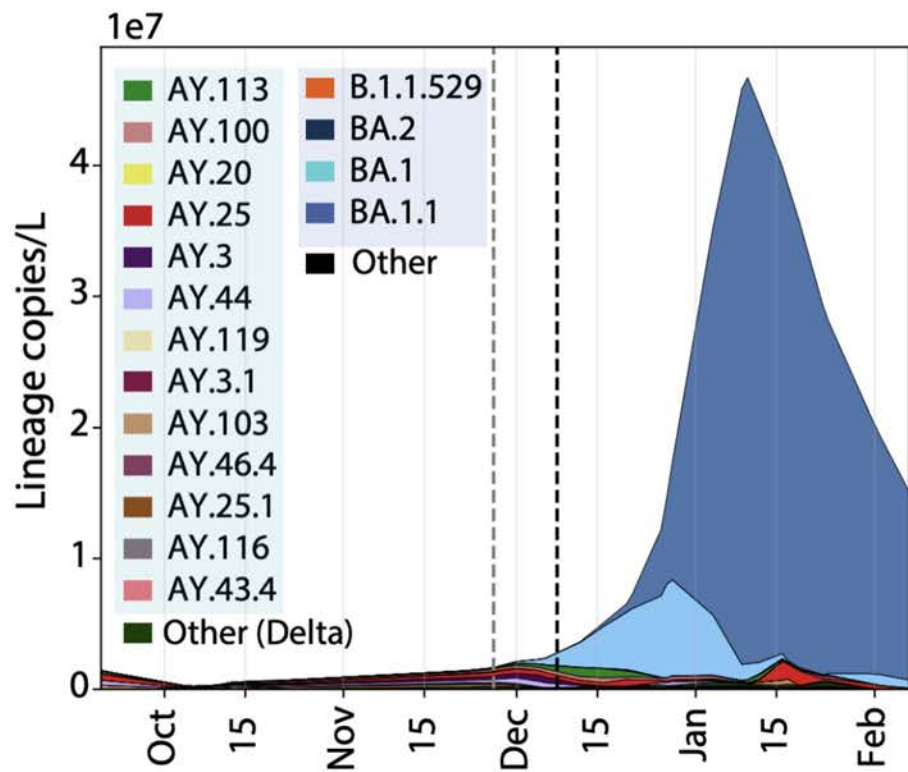
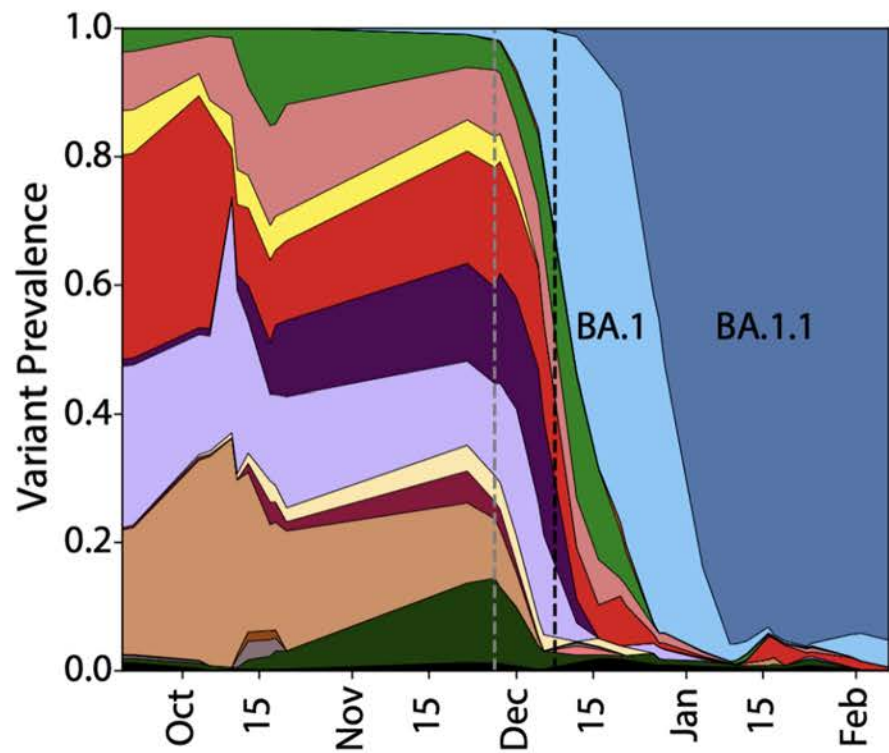


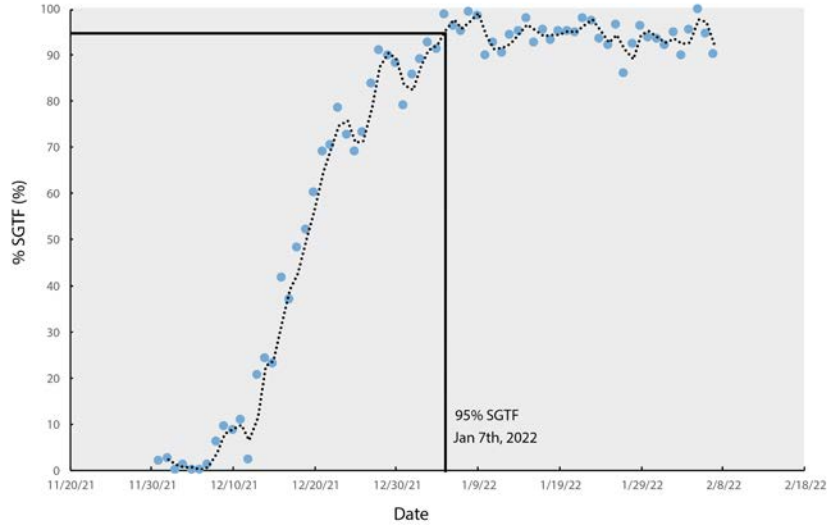
The rise of the Delta variant during Summer 2021

First detection of delta in WW: 2 weeks ahead of clinical

The Omicron wave in San Diego

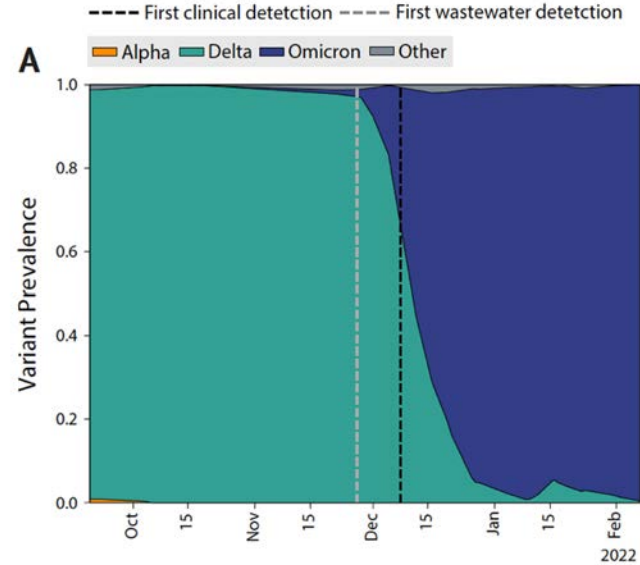






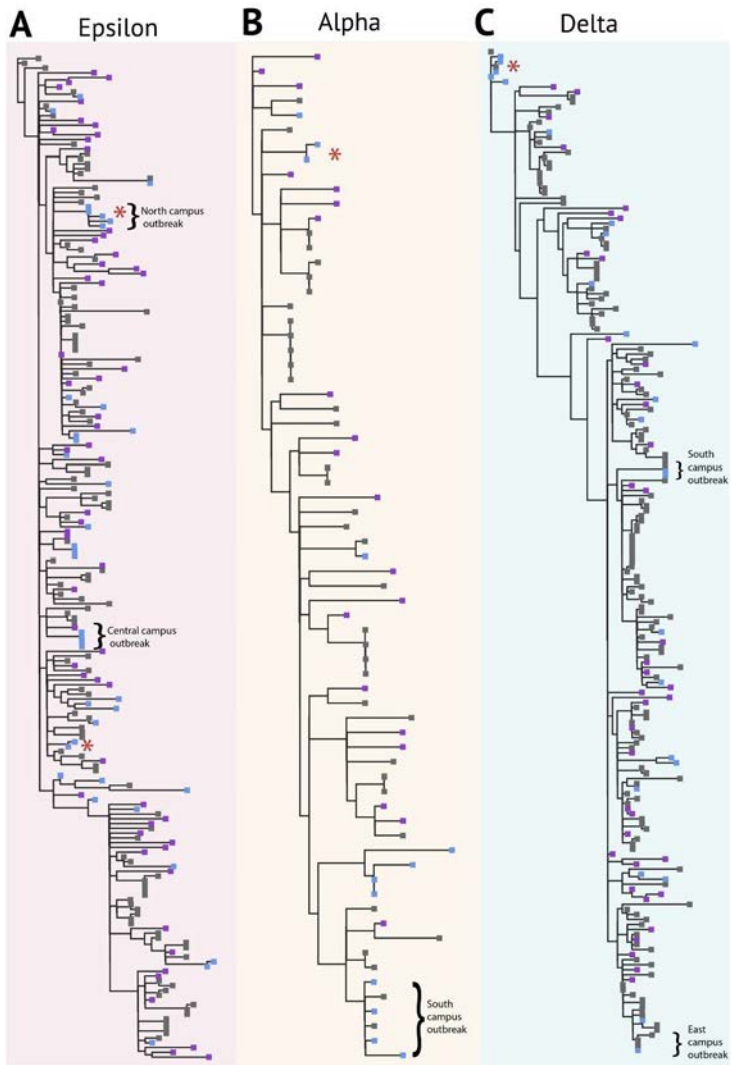
Clinical data

<https://searchcovid.info/dashboards/omicron-estimates/>



WW data

Trends captured earlier in wastewater (~10d)



Campus transmission networks with consensus genomes

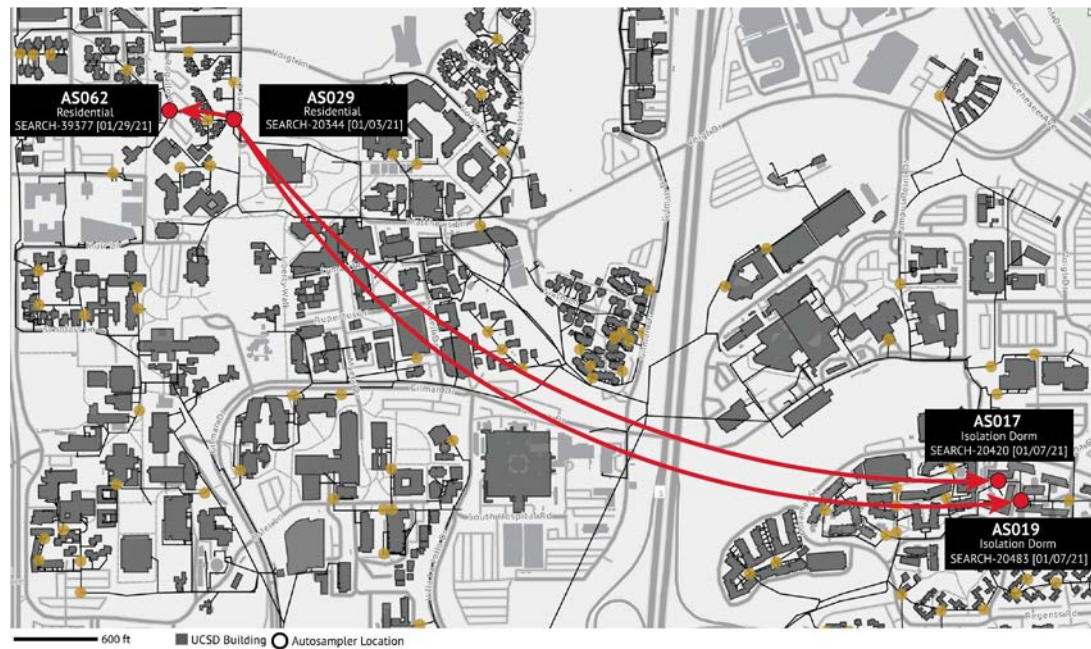
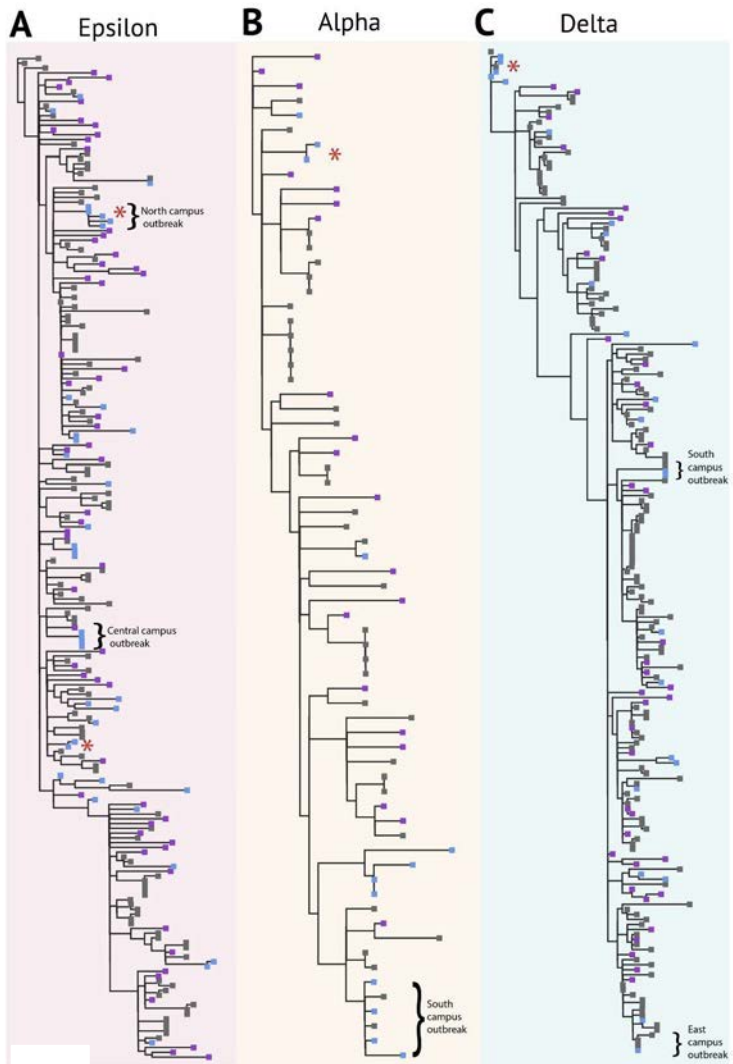
High resolution sampling on UCSD capture individual infections, clusters of SARS-CoV-2 spread

Freyja indicates sample is composed of a single lineage

Can use standard consensus sequence calling methods (iVar etc.)

■ UCSD Wastewater ■ UCSD Clinical ■ United States

Data de-identified in accordance with our IRB (approval #210699, #200477)

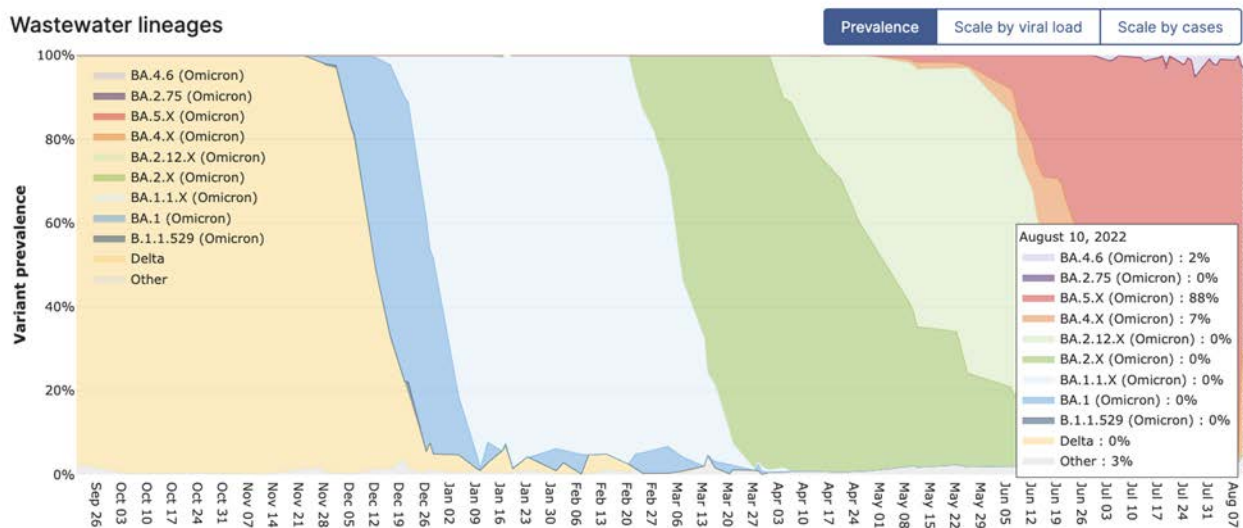


Real-time wastewater sequence tracking

Rapid turnaround: 26h sample-sequence data



Wastewater lineages



<https://searchcovid.info/dashboards/wastewater-surveillance/>

Hands-on prep

60-90min



Run

Library -result
21-25h



Online status tracking

Instruments

GNXS-0463

220811_WW
Analysis Completed

Ion Torrent™ GX5™ Chip
21DAHK00734241GX5, 4 lanes consumed

Library Templating Sequencing Analysis Plugins

Run
220811_WW
Serial
2619121070463
Started
2022-08-11 15:28

Operator
GenexusAdmin
Completed
2022-08-12 13:31

[Details](#)

PQ Status Not Verified

Last init 2022-08-04 15:16

Results

Run stats

SMS-C01-2 Insight Waterwater Research Assay_L1mga

Loading	Total Reads	Total Reads	Raw Read Accuracy
93.4%	5,776	32,428,541	98.45%

Loading 93.4% Empty Wells 8.61%

Enrichment 99.79% Unenriched 0.30%

Library 99.93% Control 0.07%

Final Reads 53.48% Filtered out 46.52%

Final Reads 32,428,541

Loading Density

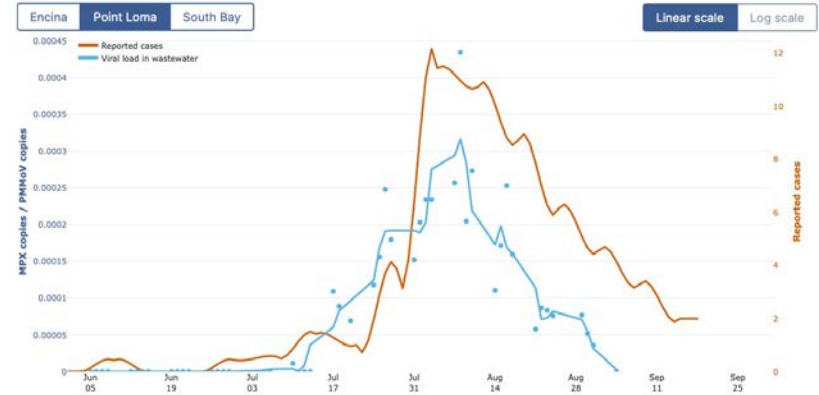
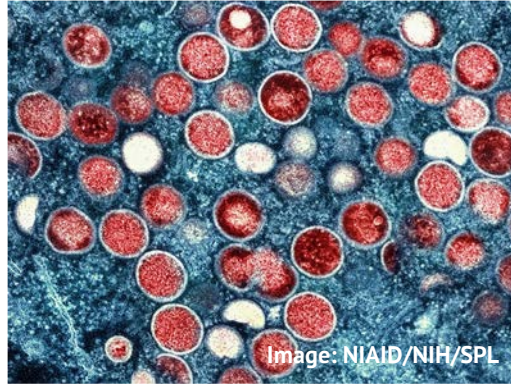
Run Samples

Sample Name	Nucleic Acid Type	Barcode	Total Reads	Mean Read Length	≠ Q20 Bases	Mapped Reads	On Target Reads	Mean Depth	Uniformity	Read Length Histogram
ENCHA_A_A_21	RNA	IonRun_21_001	4,205,480	186	475,894,291	2,734,254	98.80%	23,549	93.80%	
IonRun_21_001	RNA	IonRun_21_001	156	175,725,687	1,757,023	13,888%	95,020	97.62%		
PL_A_18_22_1	RNA	IonRun_21_001	129	493,366,492	1,860,887	70.12%	23,389	98.91%		
PL_A_7_20_1	RNA	IonRun_21_001	175	4,374,288	2,278,176	36.86%	30,126	94.20%		
PL_A_8_22_1	RNA	IonRun_21_001	181	761,402,447	1,977,268	98.20%	23,768	91.70%		
PL_A_9_22_1	RNA	IonRun_21_001	189	493,963,119	1,671,264	98.70%	22,860	99.20%		
SB_A_7_20_1	RNA	IonRun_21_001	176	468,768,264	1,617,470	96.86%	14,912	96.02%		
SB_A_8_22_1	RNA	IonRun_21_001	183	493,762,478	1,979,441	97.90%	22,338	99.8%		

Current and future extensions

Monkeypox Surveillance

WW surveillance
at county-level

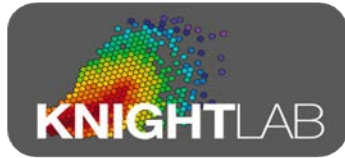


<https://searchcovid.info/dashboards/wastewater-surveillance-mpx/>

High-throughput RNA-Seq

Untargeted longitudinal metatranscriptomics of
WW data

Thank you



Knight Lab WW processing team:

Helena Tubb, Tommy Valles, Rebecca Tsai,
Caitlin Tribelhorn, Sawyer Farmer, Greg
Humphrey, Madison Ambre, Clare Lawless

UCSD RTL Team

Expedited COVID Identification Environment CLIA (EXCITE) Lab at UCSD

Institute of Genomic Medicine, CCBB (UCSD)
Facilities Management, HDH, EHS

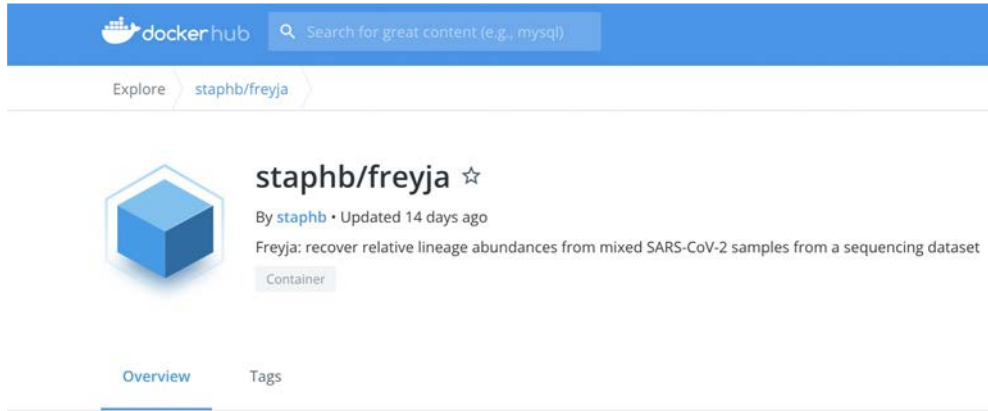


Funding



Code deployment and usage

BIOCONDA[®]



The screenshot shows the Docker Hub interface for the container image 'staphb/freyja'. At the top, there is a search bar with the text 'Search for great content (e.g., mysql)'. Below the search bar, the breadcrumb 'Explore > staphb/freyja' is visible. The main content area features a blue cube icon representing the container image. To the right of the icon, the text reads 'staphb/freyja ☆', 'By staphb • Updated 14 days ago', and 'Freyja: recover relative lineage abundances from mixed SARS-CoV-2 samples from a sequencing dataset'. A 'Container' tag is displayed below the description. At the bottom of the page, there are two tabs: 'Overview' (which is selected) and 'Tags'.

