Wastewater tracking of pathogen dynamics

Provides more accurate estimates of population infection rates than clinical testing alone due to inherent limitations in testing resources and/or uptake ,especially in underserved communities

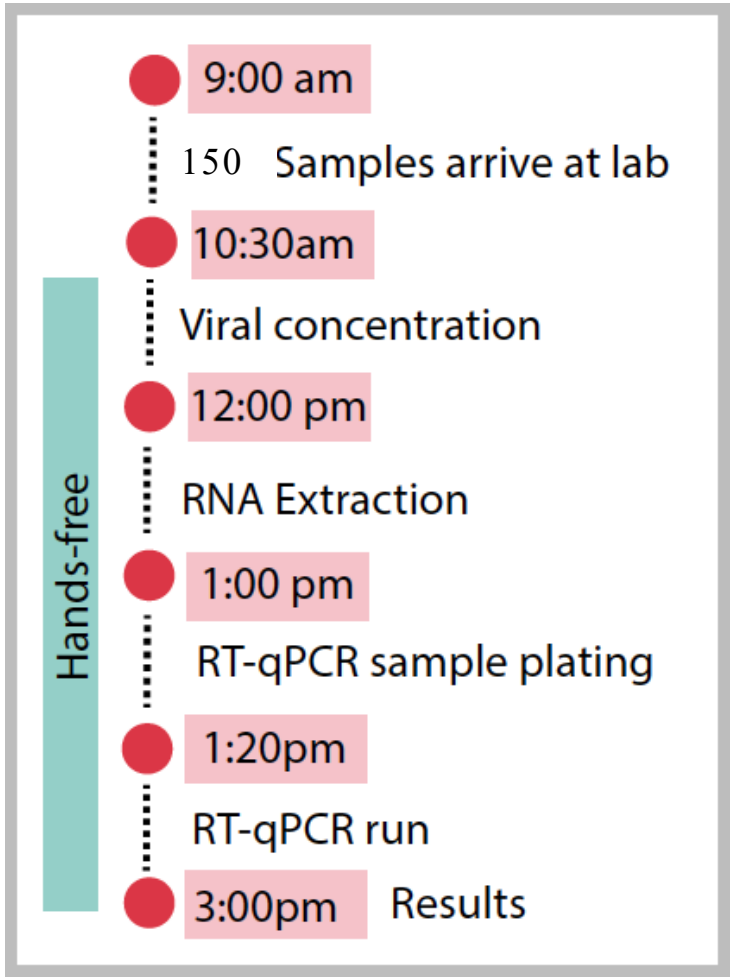
Rob Knight, Ph.D.

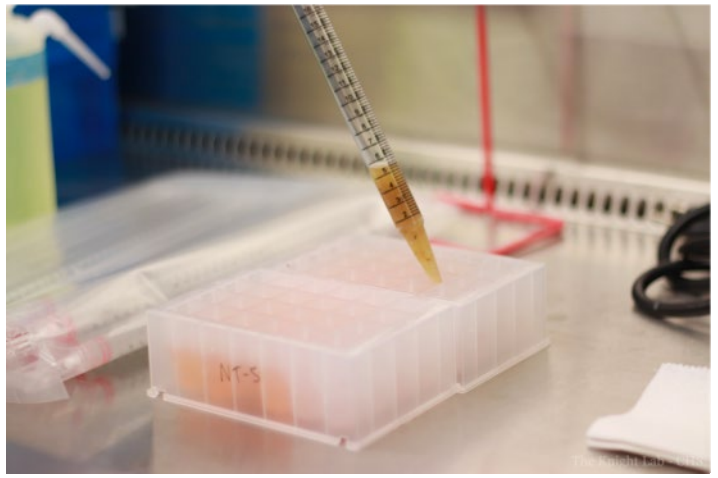
Wolfe Family Endowed Chair of Microbiome Research at Rady Children's Hospital of San Diego

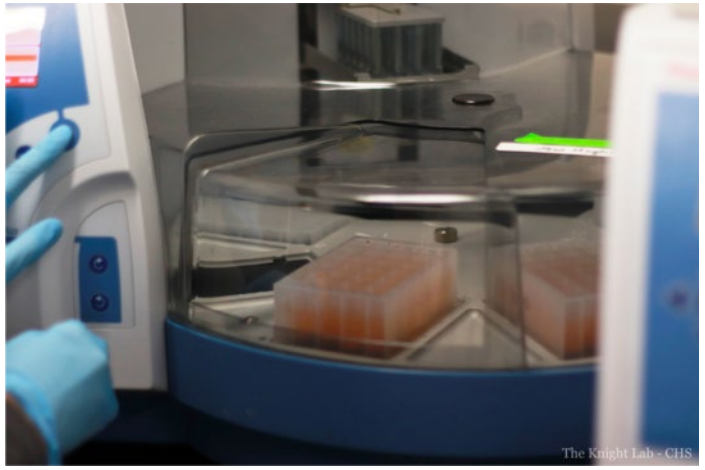
Director, Center for Microbiome Innovation, and Professor, Pediatrics, Bioengineering, CSE & HDSI at University of California San Diego

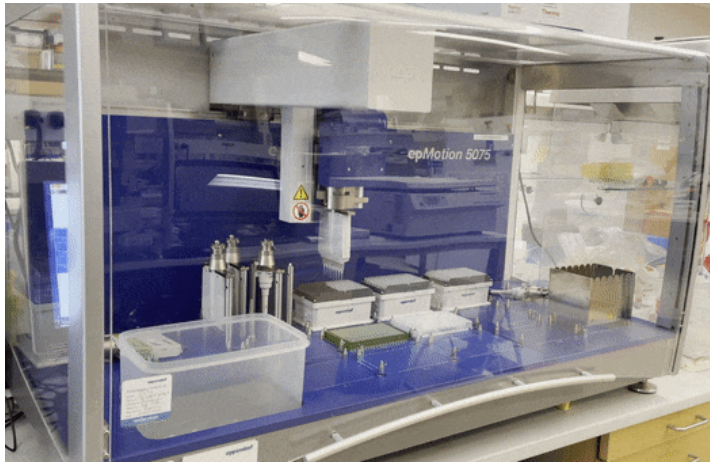
Antoine Dore, Nature

A day in the life of a sewage sample







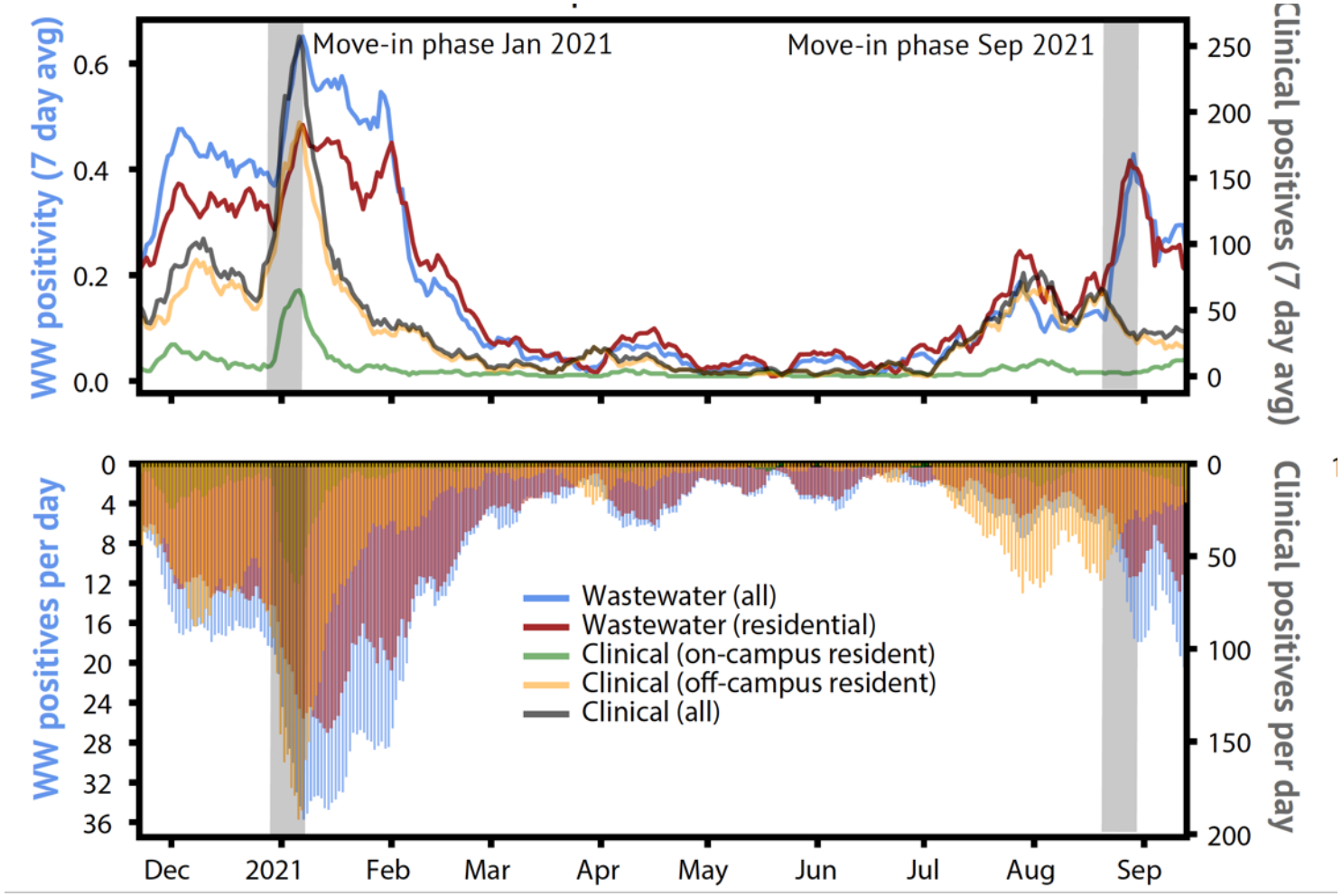


Normalization : PMMoV (fecal indicator)
Recovery: SARS-CoV-2 spike-ins from BSL-3 lab

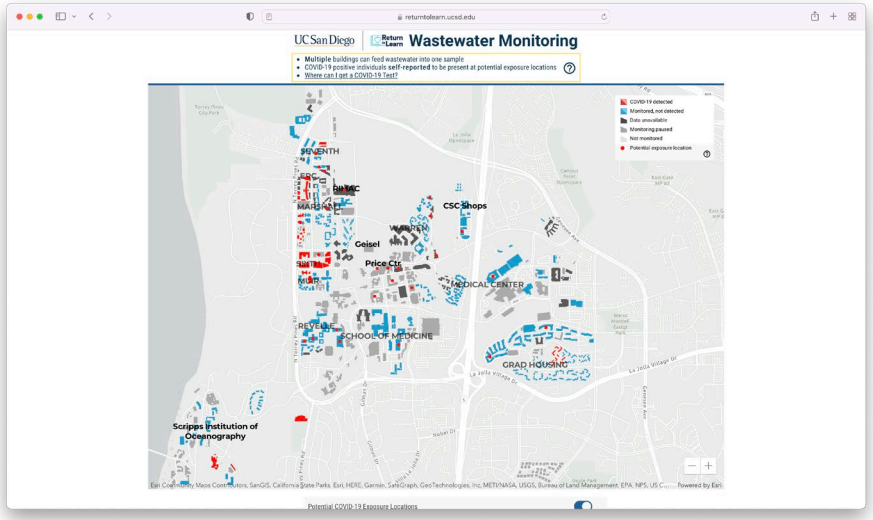
Automated wastewater viral concentration and extraction

dx.doi.org/10.17504/protocols.io.bshvnb66

Early detection of 85% of cases on campus



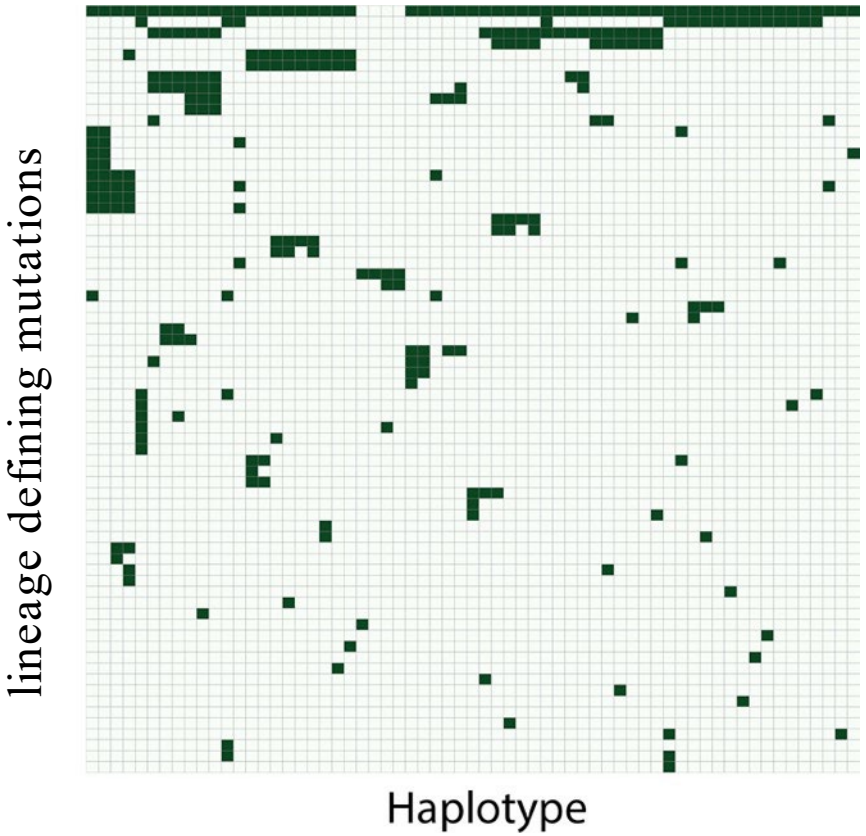
UCSD Campus Testing stats



Self-administered test via vending machines
Pilot stage: All on-campus residents mandated to test weekly

Karthikeyan et al.,https://doi.org/10.1128/mSystems.00793-21

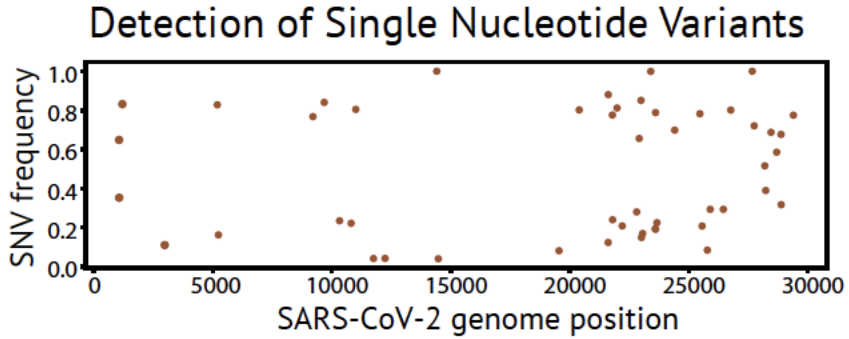
Resolving multiple strains in wastewater



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

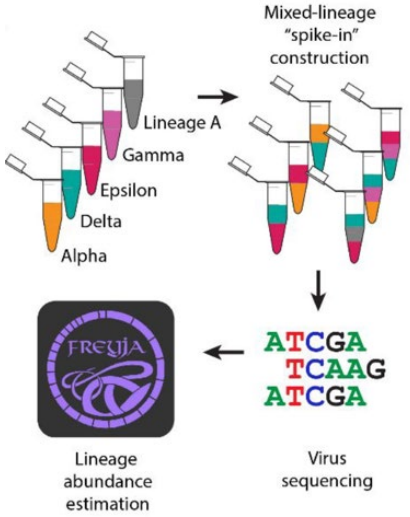
- SNV frequency estimation
- Depth-weighted de-mixing

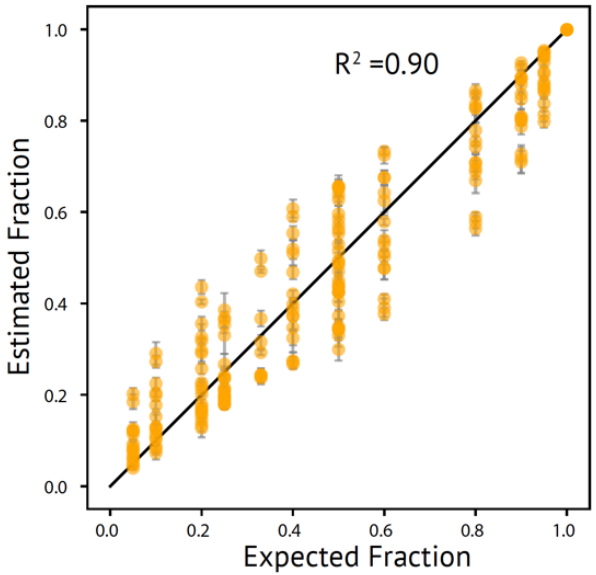
Josh Levy/Kristian Andersen, Scripps Research
Karthikeyan, Levy et al, 2022, Nature

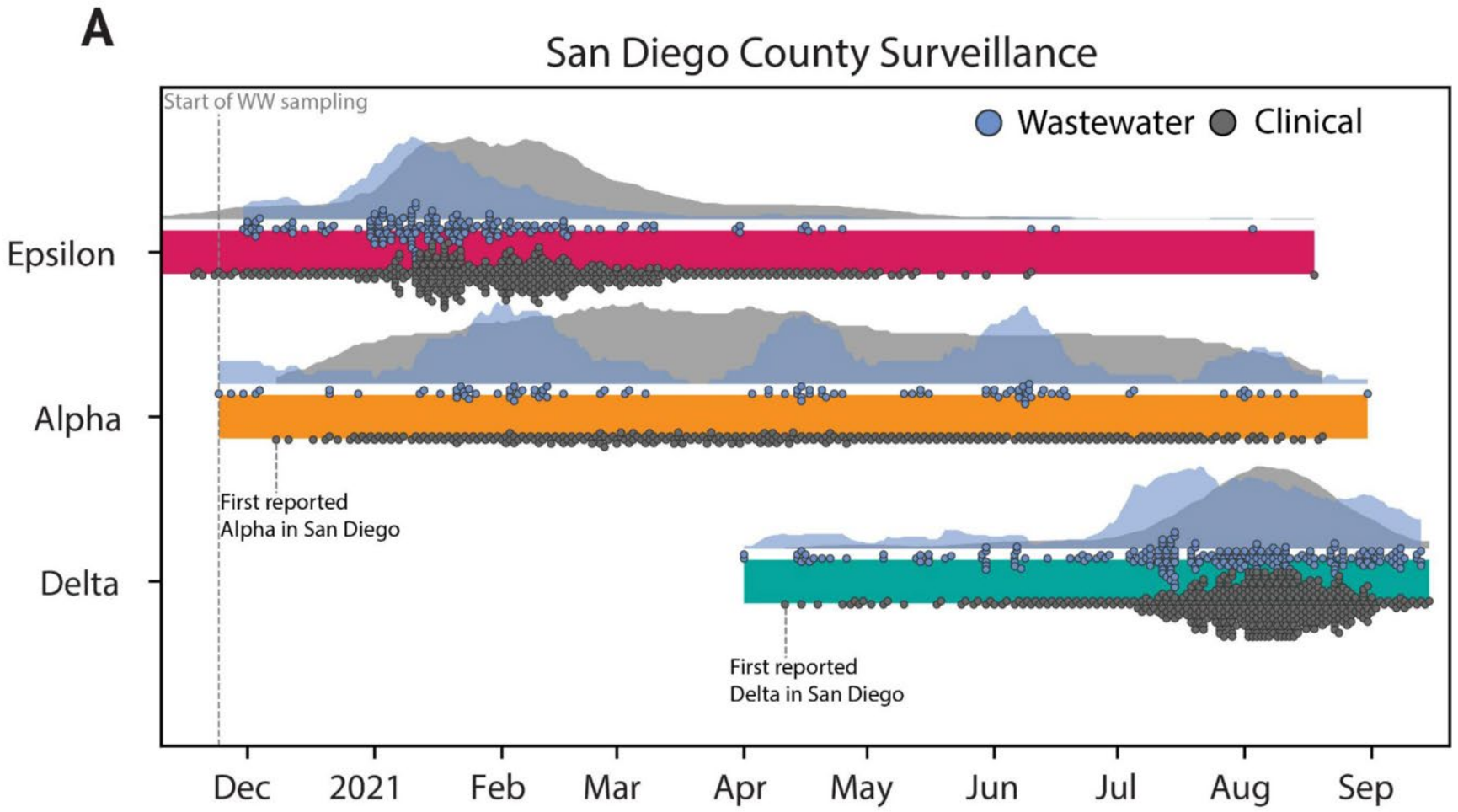


$$\hat{x} = \underset{\substack{x \geq 0 \\ \sum x = 1}}{\operatorname{argmin}} ||A^T x - b||_{1W}, \quad \text{where} \quad ||\mu||_{1W} = \sum_{i=1}^{N} d_i |\mu_i|$$

A: matrix containing the barcodes
b: frequency of mutations in sample
x: abundances of each lineage
d: sequencing depth at each position







Timeline and epidemiological curves for VOC detection in county samples

Wastewater samples show earlier appearance of VOCs

Data from 31,149 nasal swab sequences from SD county and 1200 wastewater sequences

Karthikeyan et al. 2022 Nature

Future directions/potential

- Expand monitoring (currently SARS-CoV-2 and HepA, have done mpox in past, dozens of other pathogens possible)
- Metagenomics and metatranscriptomics: allows total DNA or RNA characterization from wastewater, pathogen ID still challenging but can predict infection status from total community (collaboration with Charles Chiu, UCSF)
- Long read sequencing dramatically improves specificity at expense of sensitivity
- Target capture panels (Twist, Illumina etc.) show substantial potential as input to metagenomics/amplicon sequencing genomics methods for enrichment, still need qPCR or ddPCR to quantify

Acknowledgements

Knight Lab, especially Smruthi Karthikeyan (now at Caltech)

Freyja: Josh Levy, Andersen lab, Scripps

Modeling: Nancy Ronquillo, Prof. Tara Javidi, Machine-Integrated Computing, UCSD

GIS Workflow automation: Andrew Nguyen, Yijian Zong, Andrew Bartko, Ali Long, Cassidy Symons, Nitya Agarwal, Diana Henderson, UCSD

Clinical sample testing: Expedited COVID IdenTification Environment (EXCITE Lab) at UCSD

Sequencing: Andersen Lab (Scripps Research), Institute of Genomic Medicine (UCSD)

UCSD RTL Team: https://returntolearn.ucsd.edu/about/program-leadership/ UCSD Facilities

Management, HDH, EHS, Logistics

CMI

County of San Diego

SEARCH Alliance via SPHERES, CDC 75D30120C09795, NIH Pioneer 1DP1AT010885, NSF RAPID 2029069, San Diego County Health and Human Services Agency, Return to Learn