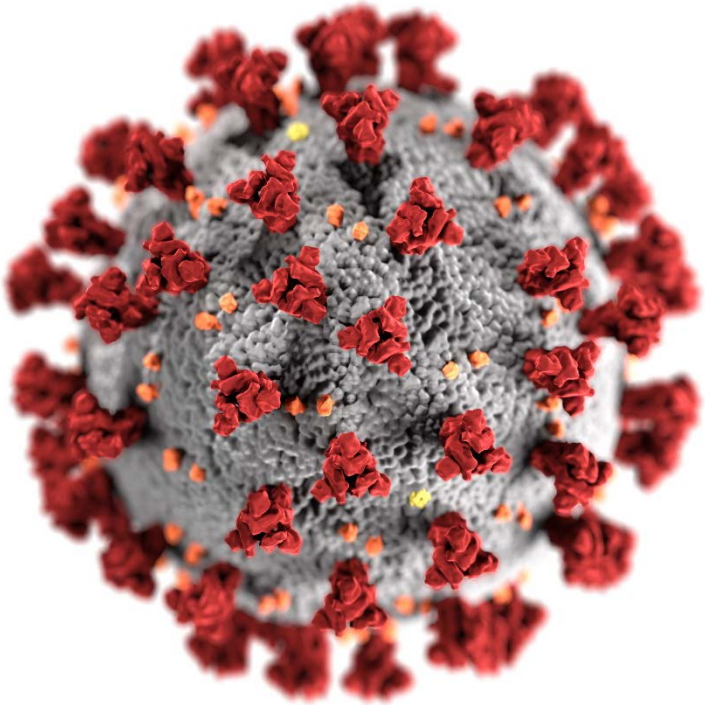
# COVID-19 Genomic Epidemiology Toolkit

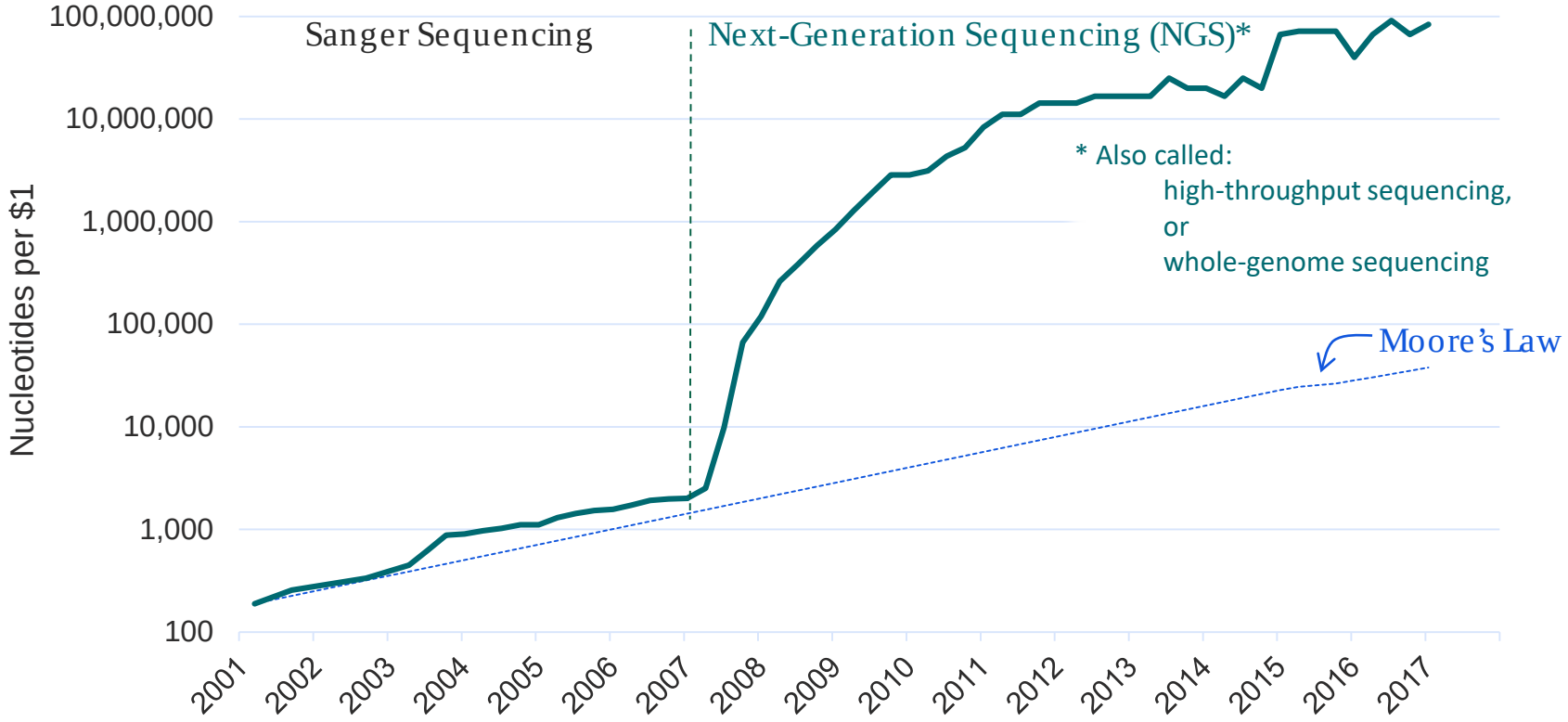
## Welcome and Overview



Gregory L. Armstrong, MD
Director, Office of Advanced Molecular Detection
Center for Disease Control and Prevention

cdc.gov/coronavirus

# Trends in Sequencing Output




Adapted from NHGRI (https://www.genome.gov/sequencingcosts/)

# Microbial genomics now in wide use in US Public Health

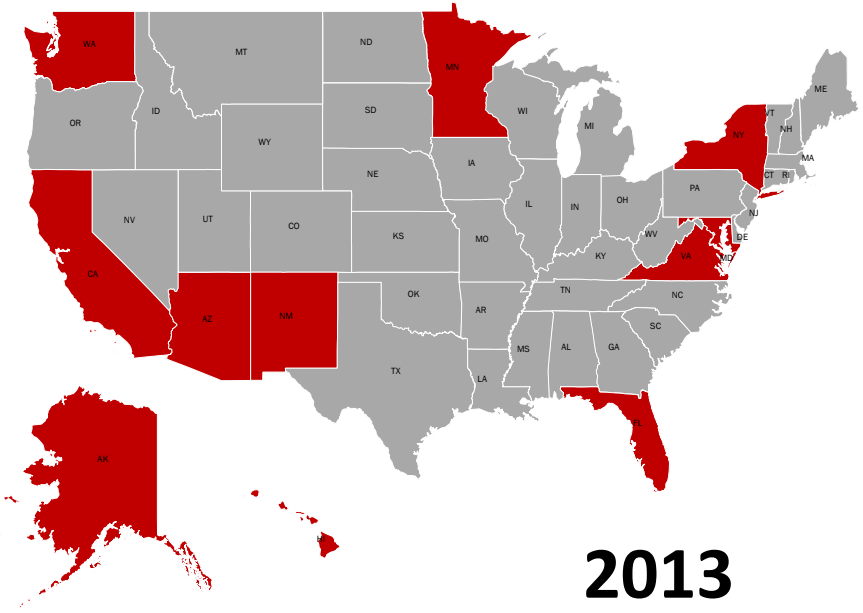
- Food safety: identify and investigate outbreaks
- Antimicrobial resistant organisms
  - infer resistance
  - better understand transmission
- Influenza: strain surveillance
- Tuberculosis: identify and investigate clusters
- Group A *Streptococcus:* investigate outbreaks in healthcare settings
- Malaria: monitor resistance and other clinically important traits
- Rabies: track variants
- … and many more

NEJM 2019; 381:2569-80.

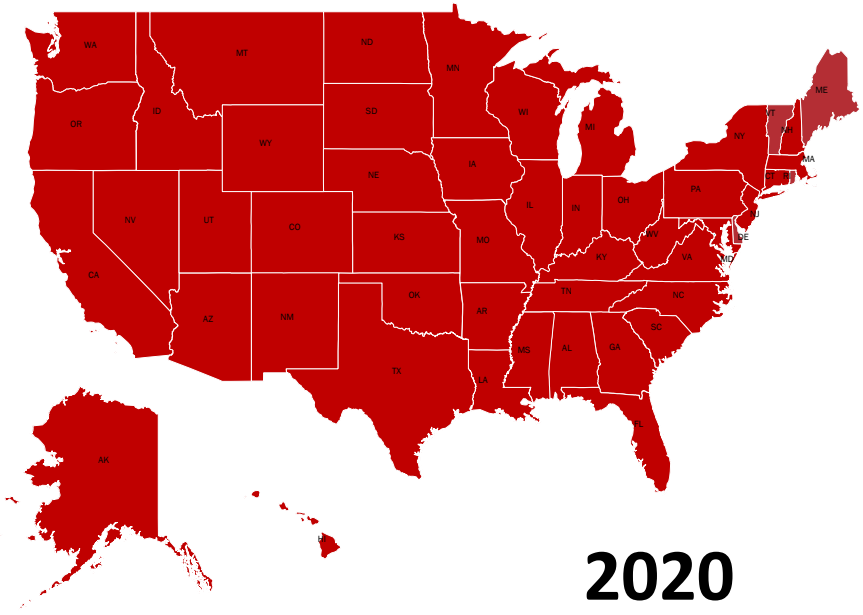
# How it started

# How it's going

Seven years of building NGS capacity in state public health laboratories



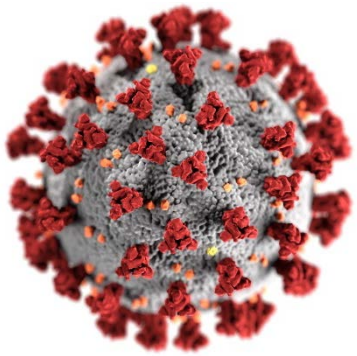

2013




2020

www.aphl.org/aboutAPHL/publications/Documents/ID_NGSSurveyReport_52015.pdf

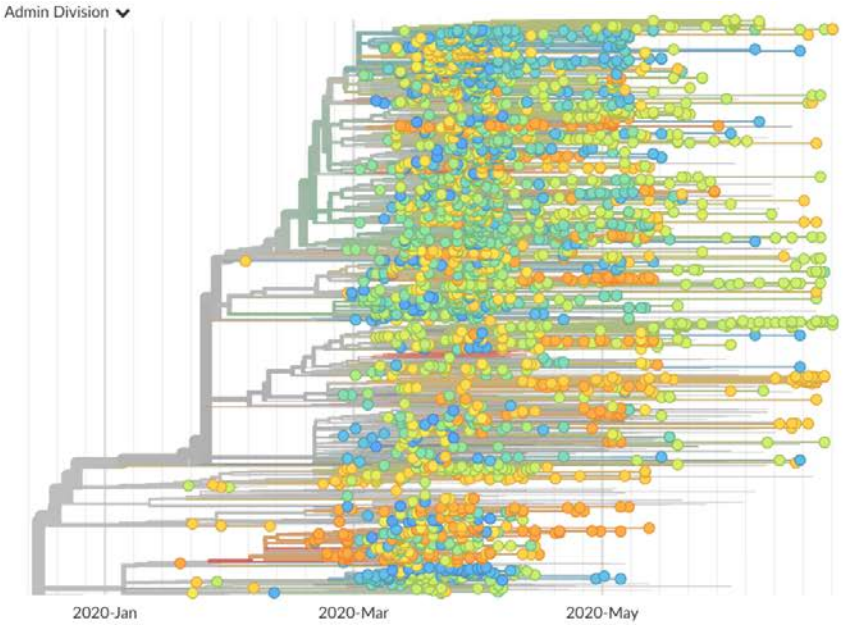
# Rationale for sequencing of SARS-CoV-2

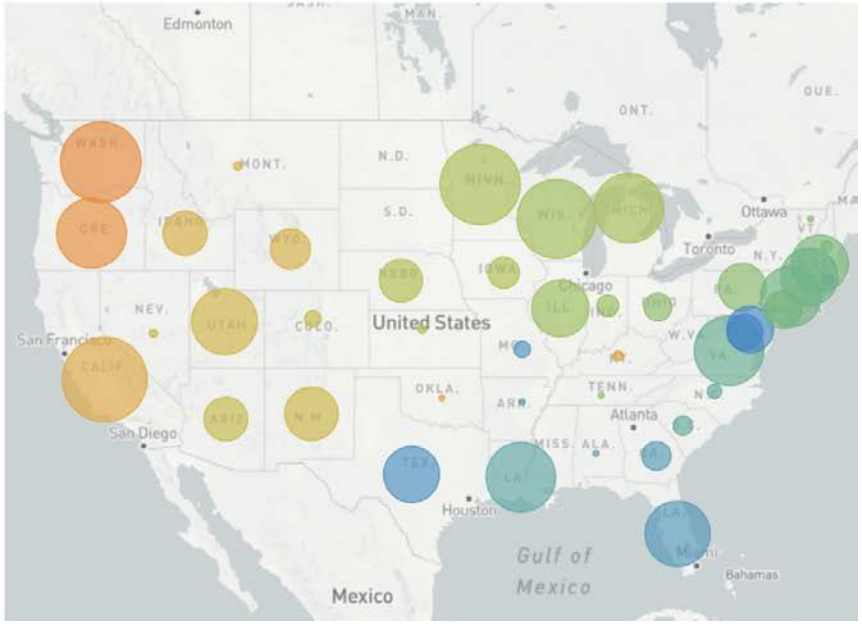


- National level
  - strain surveillance
  - help guide diagnostics, vaccine and therapeutic development
- State/Local level
  - supplement control efforts
  - better understand epidemiology

# National level:

- Monitor emergence of important new strains
- Monitor trends after interventions such as vaccination




Nextstrain.org

# State/local level:

- Identify clusters
- …

*"… Our data unexpectedly uncovered a sustained outbreak in six care facilities within the region. …"*

medRxiv preprint doi: https://doi.org/10.1101/2020.09.28.20201475; this version posted November 16, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

## Large scale sequencing of SARS-CoV-2 genomes from one region allows detailed epidemiology and enables local outbreak management

Andrew J. Page[1,*,#], Alison E. Mather[1,4,#], Thanh Le-Viet[1], Emma J. Meader[3], Nabil-Fareed Alikhan[1], Gemma L. Kay[1], Leonardo de Oliveira Martins[1], Alp Aydin[1], Dave J. Baker[1], Alexander J. Trotter[1,4], Steven Rudder[1], Ana P. Tedim[1,2], Anastasia Kolyva[1,3], Rachael Stanley[3], Maria Diaz[1], Will Potter[3], Claire Stuart[3], Lizzie Meadows[1], Andrew Bell[1], Ana

Source: Page A, Mather A, e-Viet T, et al. Large scale sequencing of SARS-CoV-2 genomes from one region allows detailed epidemiology and enables local outbreak management. medRxiv 2020.09.28.20201475; DOI: https://doi.org/10.1101/2020.09.28.20201475

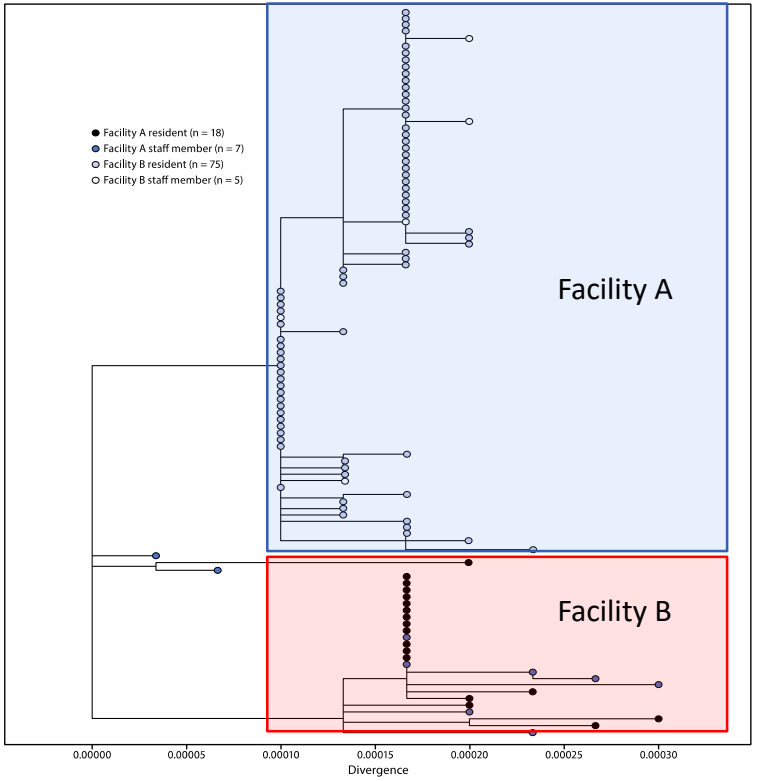
# State/local level:

- Identify clusters
- Investigate outbreaks
- …

*"Genetic sequencing found facility-specific clustering of viral genomes from HCP and residents' specimens, suggesting intrafacility transmission."*



**FIGURE 2. Phylogenetic trees showing genetic distance between available* SARS-CoV-2 virus specimens collected from health care personnel (HCP) and residents at facility A† and facility B§— Minnesota, April–June 2020**

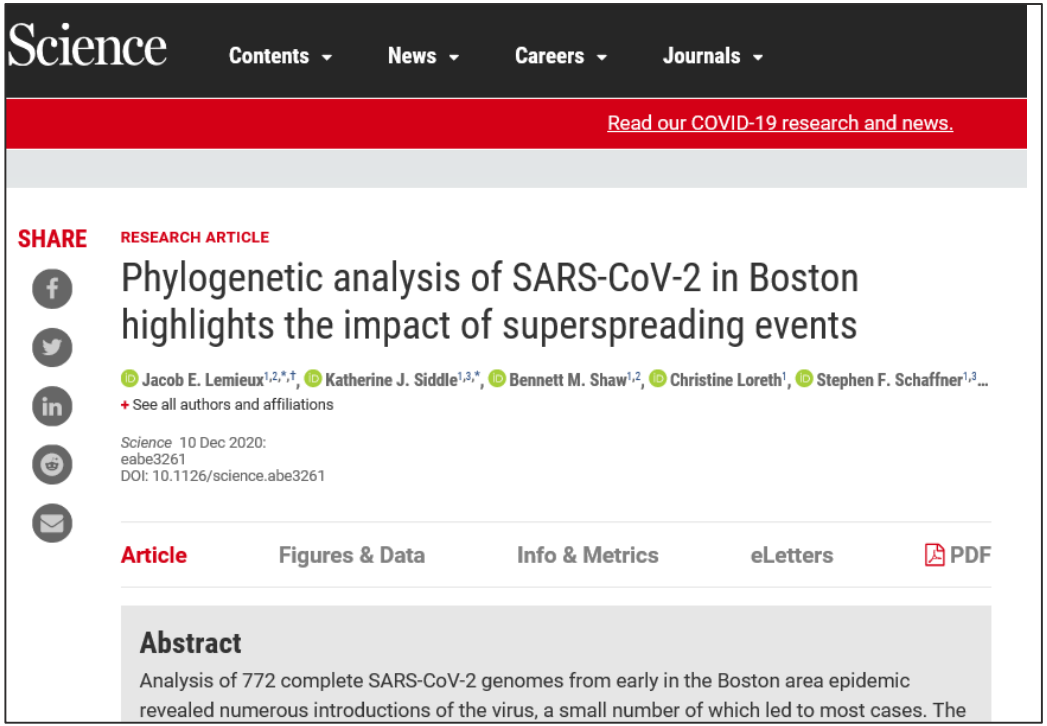



Source: Taylor J, Carter RJ, Lehnertz N, et al. Serial Testing for SARS-CoV-2 and Virus Whole Genome Sequencing Inform Infection Risk at Two Skilled Nursing Facilities with COVID-19 Outbreaks — Minnesota, April–June 2020. MMWR Morb Mortal Wkly Rep 2020;69:1288–1295. DOI: http://dx.doi.org/10.15585/mmwr.mm6937a3

# State/local level:

- Identify clusters
- Investigate outbreaks
- Improve understanding of local epidemiology
- …

*"The data reveal over 80 introductions into the Boston area…."*



Science Contents News Careers Journals

Read our COVID-19 research and news.

SHARE

RESEARCH ARTICLE

Phylogenetic analysis of SARS-CoV-2 in Boston highlights the impact of superspreading events

Jacob E. Lemieux[1,2,*,†], Katherine J. Siddle[1,3,*], Bennett M. Shaw[1,2], Christine Loreth[1], Stephen F. Schaffner[1,3]…

+ See all authors and affiliations

*Science* 10 Dec 2020:
eabe3261
DOI: 10.1126/science.abe3261

Article Figures & Data Info & Metrics eLetters PDF

**Abstract**

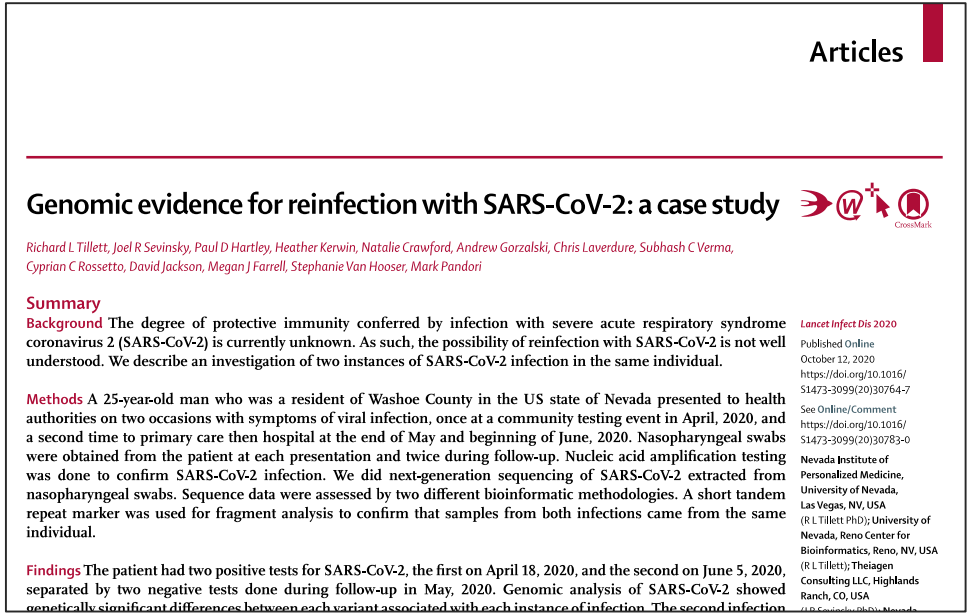
Analysis of 772 complete SARS-CoV-2 genomes from early in the Boston area epidemic revealed numerous introductions of the virus, a small number of which led to most cases. The

Source: Lemieux J, Siddle K, Shaw B, et al. Phylogenetic analysis of SARS-CoV-2 in the Boston area highlights the role of recurrent importation and superspreading events. Science 2020 10 Dec DOI: 10.1126/science.abe3261

# State/local level:

- Identify clusters
- Investigate outbreaks
- Improve understanding of local epidemiology
- … and many other applications

*"These findings suggest that the patient was infected by SARS-CoV-2 on two separate occasions by a genetically distinct visus."*



Articles

## Genomic evidence for reinfection with SARS-CoV-2: a case study

*Richard L Tillett, Joel R Sevinsky, Paul D Hartley, Heather Kerwin, Natalie Crawford, Andrew Gorzalski, Chris Laverdure, Subhash C Verma, Cyprian C Rossetto, David Jackson, Megan J Farrell, Stephanie Van Hooser, Mark Pandori*

**Summary**

**Background** The degree of protective immunity conferred by infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is currently unknown. As such, the possibility of reinfection with SARS-CoV-2 is not well understood. We describe an investigation of two instances of SARS-CoV-2 infection in the same individual.

**Methods** A 25-year-old man who was a resident of Washoe County in the US state of Nevada presented to health authorities on two occasions with symptoms of viral infection, once at a community testing event in April, 2020, and a second time to primary care then hospital at the end of May and beginning of June, 2020. Nasopharyngeal swabs were obtained from the patient at each presentation and twice during follow-up. Nucleic acid amplification testing was done to confirm SARS-CoV-2 infection. We did next-generation sequencing of SARS-CoV-2 extracted from nasopharyngeal swabs. Sequence data were assessed by two different bioinformatic methodologies. A short tandem repeat marker was used for fragment analysis to confirm that samples from both infections came from the same individual.

**Findings** The patient had two positive tests for SARS-CoV-2, the first on April 18, 2020, and the second on June 5, 2020, separated by two negative tests done during follow-up in May, 2020. Genomic analysis of SARS-CoV-2 showed genetically significant differences between each variant associated with each instance of infection. The second infection

*Lancet Infect Dis 2020*

Published Online
October 12, 2020
https://doi.org/10.1016/S1473-3099(20)30764-7

See Online/Comment
https://doi.org/10.1016/S1473-3099(20)30783-0

**Nevada Institute of Personalized Medicine, University of Nevada, Las Vegas, NV, USA** (R L Tillett PhD); **University of Nevada, Reno Center for Bioinformatics, Reno, NV, USA** (R L Tillett); **Theiagen Consulting LLC, Highlands Ranch, CO, USA**

Lancet Infect Dis 2020 Oct 12;S1473-3099(20)30764-7. doi: 10.1016/S1473-3099(20)30764-7.

# The Molecular Epidemiology Toolkit

- CDC designed this toolkit to help epidemiologists at state, local, tribal, and territorial levels respond to COVID-19.
- The toolkit may also be useful for public health laboratory scientists, bioinformaticians, and administrators.
- Additional resources are included to encourage broader application of genomic epidemiology to controlling COVID-19.

# Modular toolkit format

## Part 1: Introduction

1.1 What is genomic epidemiology?

1.2 The SARS-CoV-2 genome

1.3 How to read phylogenetic trees

## Part 2: Case Studies

2.1 SARS-CoV-2 sequencing in Arizona

2.2 Healthcare cluster transmission

2.3 Community Transmission

## Part 3: Implementation

3.1 Getting started with Nextstrain

3.2 Getting started with MicrobeTrace

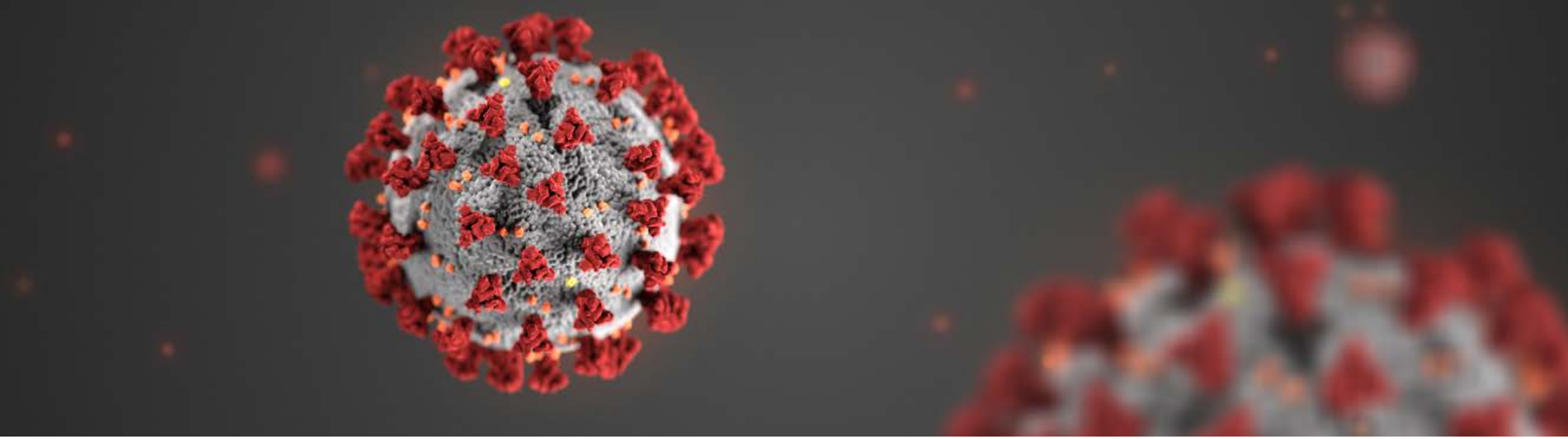
3.3 Linking epidemiologic data

# Acknowledgements


- The Translational Genomics Research Institute (TGEN)
  - Hayley Yaglom
- Minnesota Department of Health
  - Nick Lehnertz
  - Ruth Lynfield
  - Sean Wang
- Centers for Disease Control and Prevention
  - Nancy Chow
  - Luciann Draper
  - Marta Gwinn
  - Shatavia Morrison
  - Michael Weigand

# Get Email Updates

To receive updates on new modules as they are released, fill out the email subscription field on the toolkit page: go.usa.gov/xAbMw



For more information, contact CDC
1-800-CDC-INFO (232-4636)
TTY: 1-888-232-6348 www.cdc.gov

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.