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

Sanger Sequencing | Next-Generation Sequencing (NGS)*

* Also called: high-throughput sequencing, or whole-genome sequencing

Moore’s Law

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
How it started

How it’s going

Seven years of building NGS capacity in state public health laboratories

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
State/local level:

- Identify clusters

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

State/local level:

- Identify clusters
- Investigate outbreaks
- ...

“Genetic sequencing found facility-specific clustering of viral genomes from HCP and residents’ specimens, suggesting intrafacility transmission.”

State/local level:

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

“The data reveal over 80 introductions into the Boston area....”

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.”
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