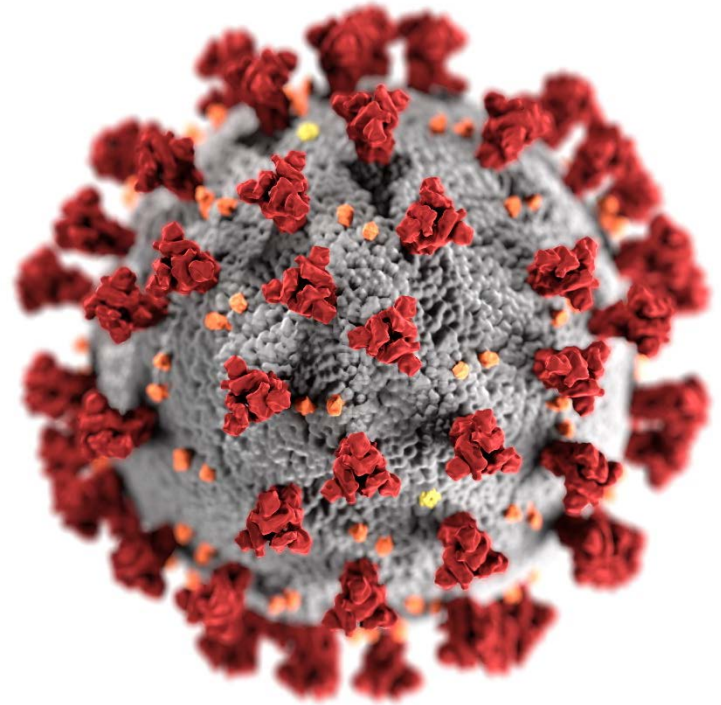# Walking through Nextstrain trees

**COVID-19 Genomic Epidemiology Toolkit:
Module 3.4**

Krisandra Allen, MPH, MB(ASCP)

Molecular Epidemiologist

Washington State Department of Health

cdc.gov/coronavirus

Toolkit map

Part 1: Introduction

1.1 What is genomic epidemiology?

1.2 The SARS-CoV-2 genome

1.3 How to read phylogenetic trees

1.4 Emerging variants of SARS-CoV-2

Part 2: Case Studies

2.1 SARS-CoV-2 sequencing in Arizona

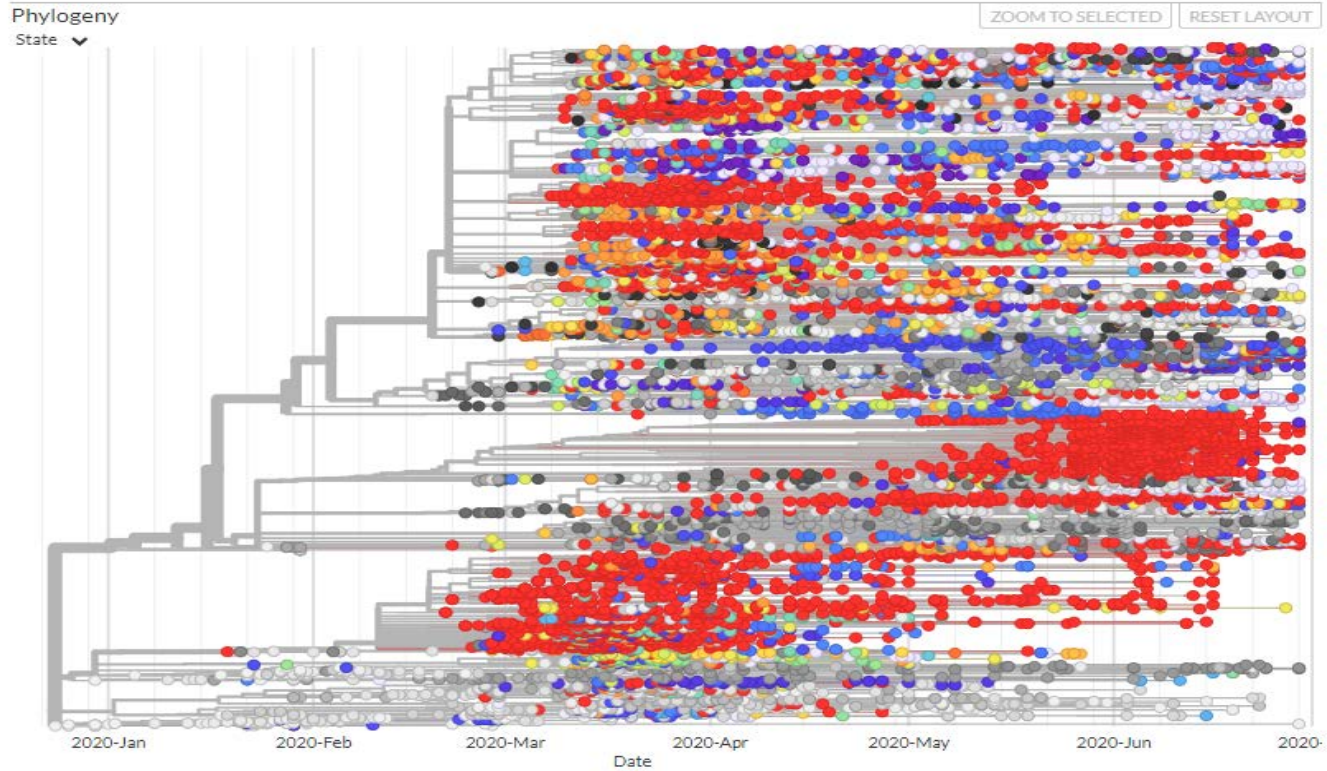2.2 Healthcare cluster transmission

2.3 Community transmission

Part 3: Implementation

3.1 Getting started with Nextstrain

3.2 Getting started with MicrobeTrace

3.3 Phylogenetics with UShER

**3.4 Walking through Nextstrain trees**

What is Nextstrain?

- Powerful analytics and interactive visualization tool (Module 3.1)
    - Designed to aid epidemiological understanding, improve outbreak response, and provide real-time snapshots of evolving pathogen populations

- Currently used to support SARS-CoV-2 genomic epidemiology at local, state, and national levels

- Many SARS-CoV-2 builds available to the public
    - https://nextstrain.org/sars-cov-2/
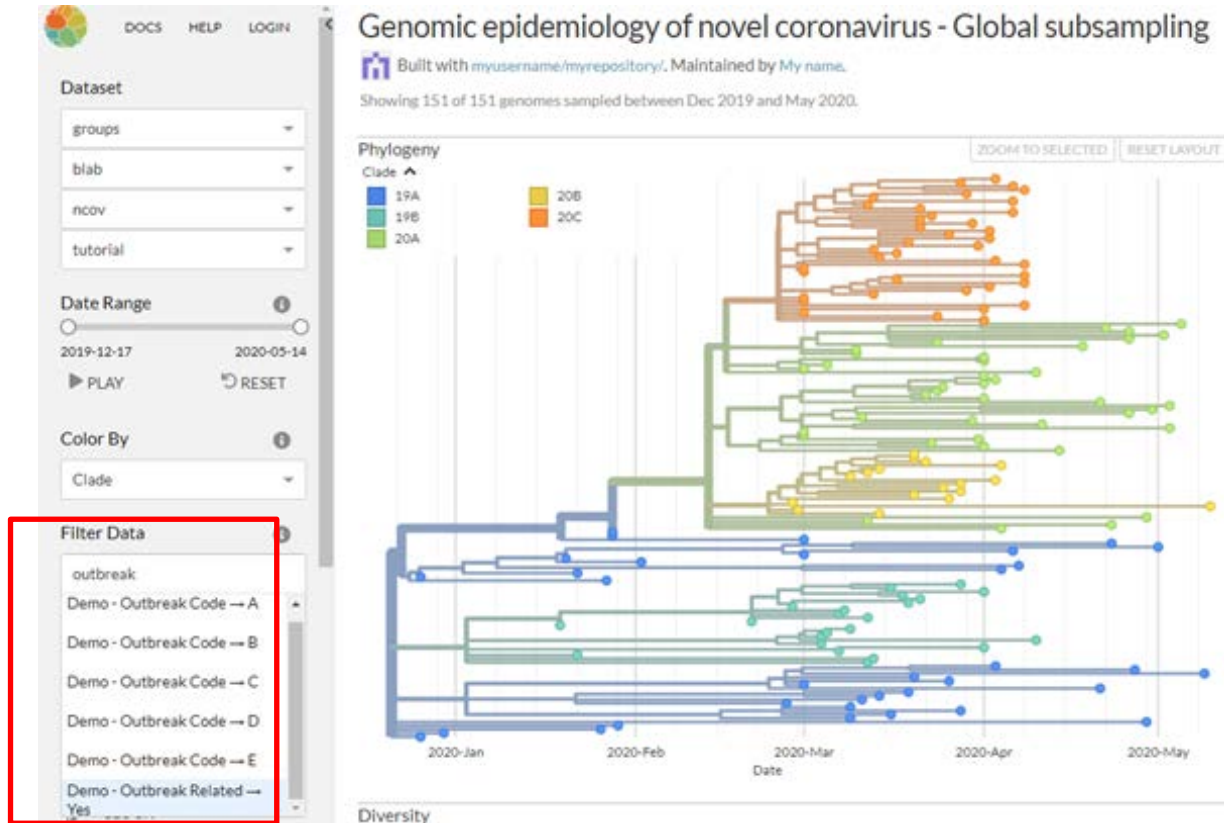   - https://nextstrain.org/groups/spheres

Complex phylogenetic trees

Tree Terminology

Genetic Distance
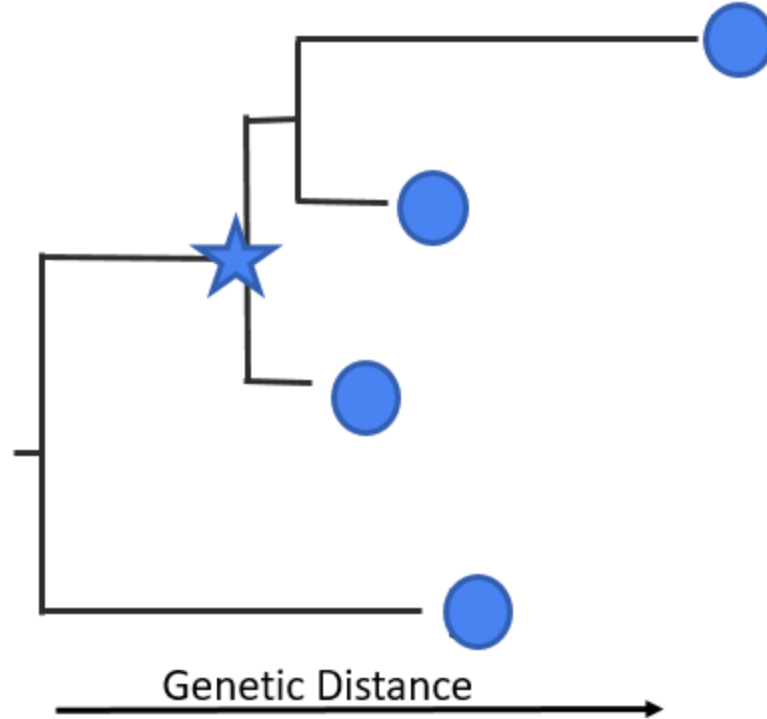
Live - Demo

- Tutorial MetaData File - demo_metadata.tsv (URL)
- Tutorial Newick File - https://nextstrain.org/groups/blab/ncov/tutorial

Locate Tips of Interest using the 'Filter' Function

Locate Tips of Interest using the 'Filter' Function
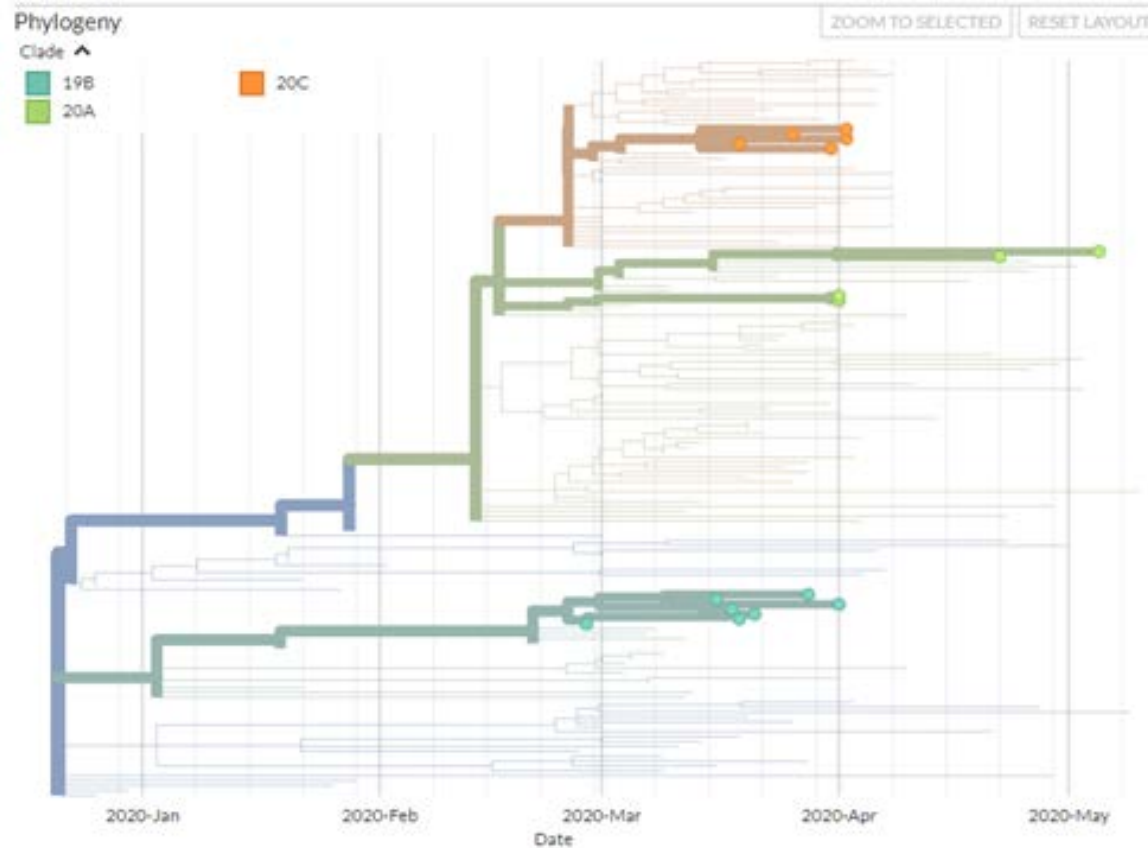
Locate Individual Outbreaks using 'Color By' Function

Advanced Functionality - Combining filters
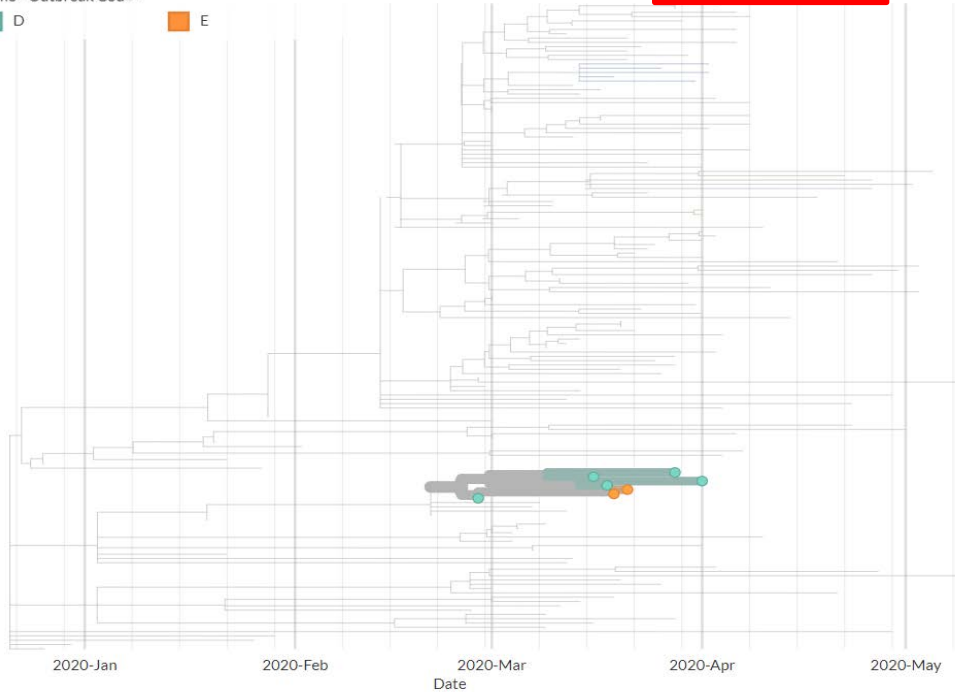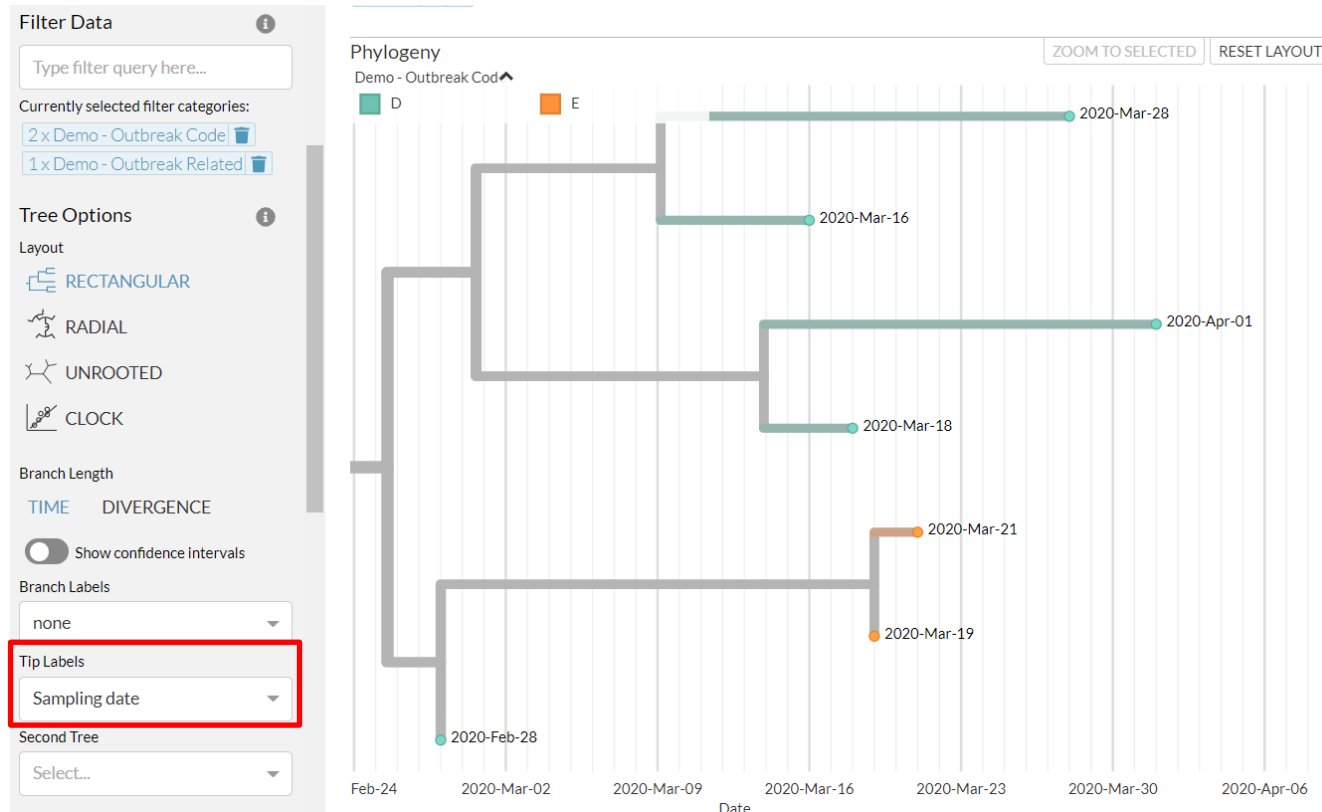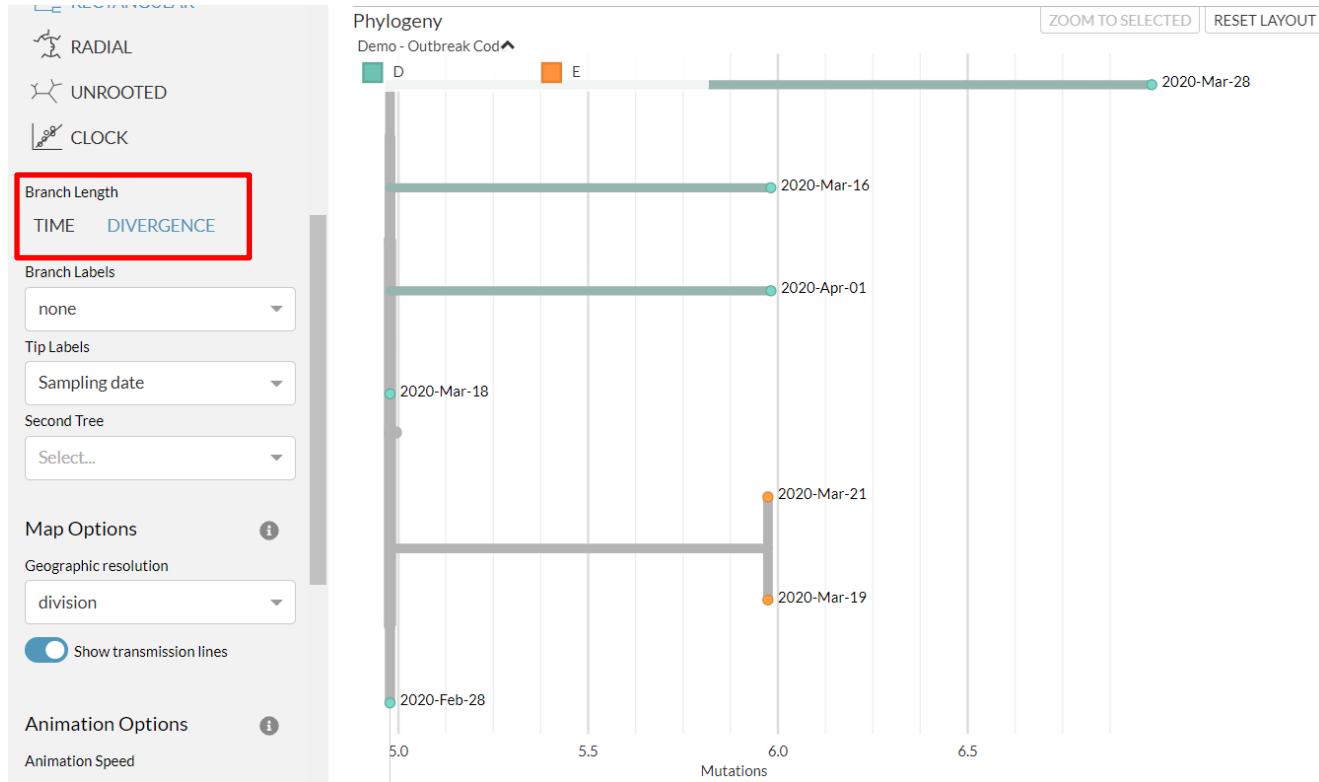
Advanced Functionality -
Changing Tip Labels on a Zoomed-in View

Advanced Functionality - Switching Between Time and Divergence Trees

Exporting filtered data

- After sequences of interest are selected on the tree, they can be exported for additional exploratory analyses in tools such as MicrobeTrace and auspice.us

Interoperability with Auspice.us and MicrobeTrace

- Exported filtered data viewed inAuspice.us or MicrobeTrace for different perspectives
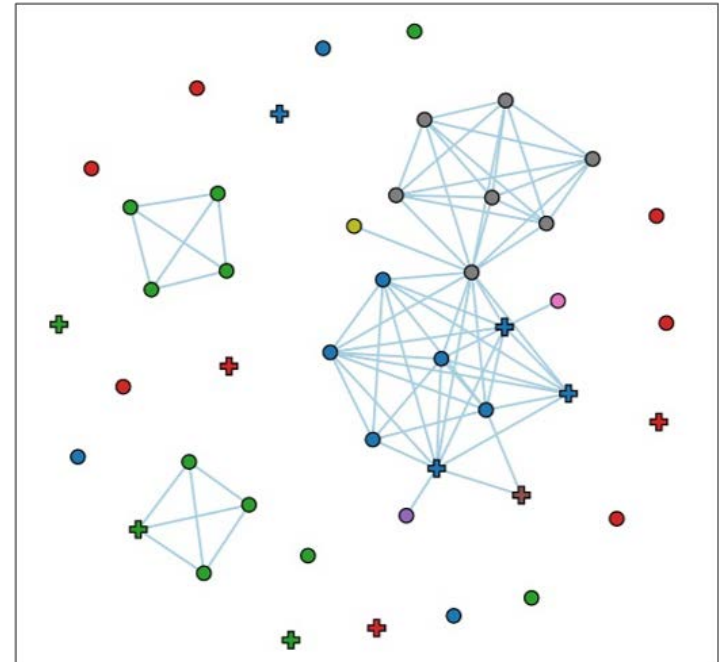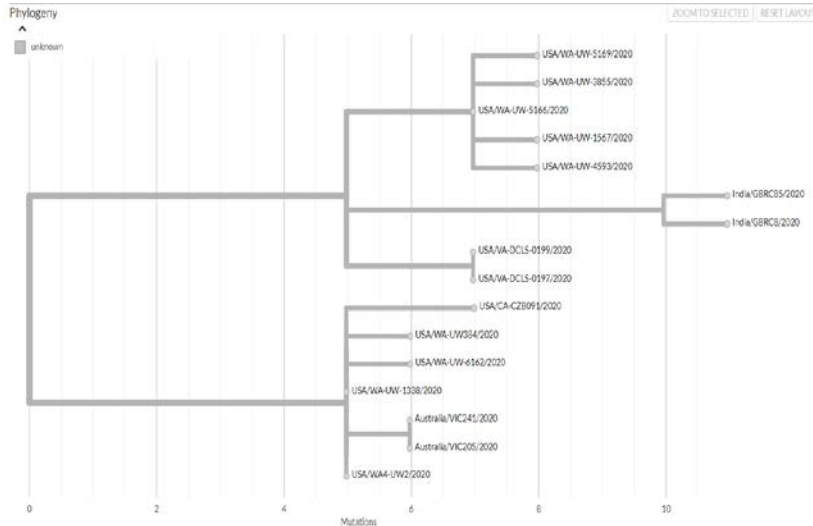
Summary

- NextStrain can assist exploration of genomic and epidemiologic data
  - Filtering, sorting, zooming, and overlaying

- Complete trees or filtered datasets can be exported
  - Analysis and visualization with other tools, such as MicrobeTrace and Auspice.us

Module 3.4 Hands-On Material

- NextStrain Tutorial

  - https://nextstrain.org/groups/blab/ncov/tutorial

- Metadata File

  - Module 3.4 - demo_metadata.tsv

Learn more

- Other modules in Part 3: Implementation
  - Getting Started with Nextstrain– Module 3.1
  - Getting Started with MicrobeTrace – Module 3.2
  - Phylogenetics with UShER – Module 3.3

- COVID-19 Genomic Epidemiology Toolkit
  - Find further reading
  - Subscribe to receive updates on new modules as they are released
  - go.usa.gov/xAbMw
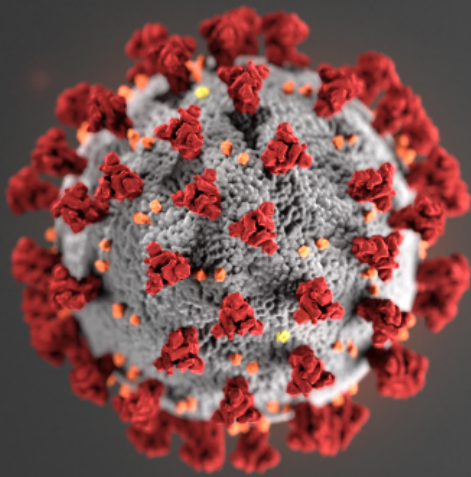
For more information, contact CDC
1-800-CDC-INFO (232-4636)
TTY: 1-888-232-6348   www.cdc.gov

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.

Resources

- Nextstrain Builds
  - https://nextstrain.org/groups/spheres