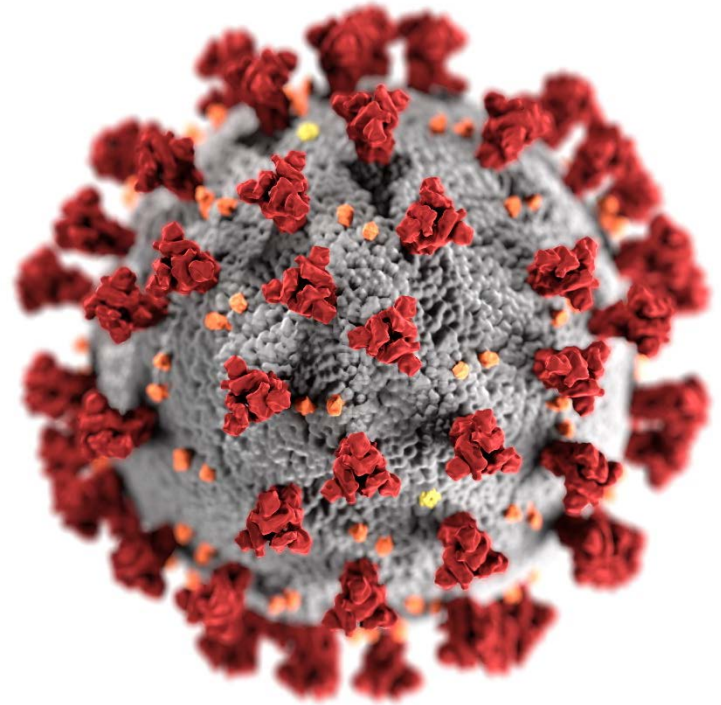


# Walking through Nextstrain trees

## COVID-19 Genomic Epidemiology Toolkit: Module 3.4

Krisandra Allen, MPH, MB(ASCP)  
Molecular Epidemiologist  
Washington State Department of Health



[cdc.gov/coronavirus](https://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

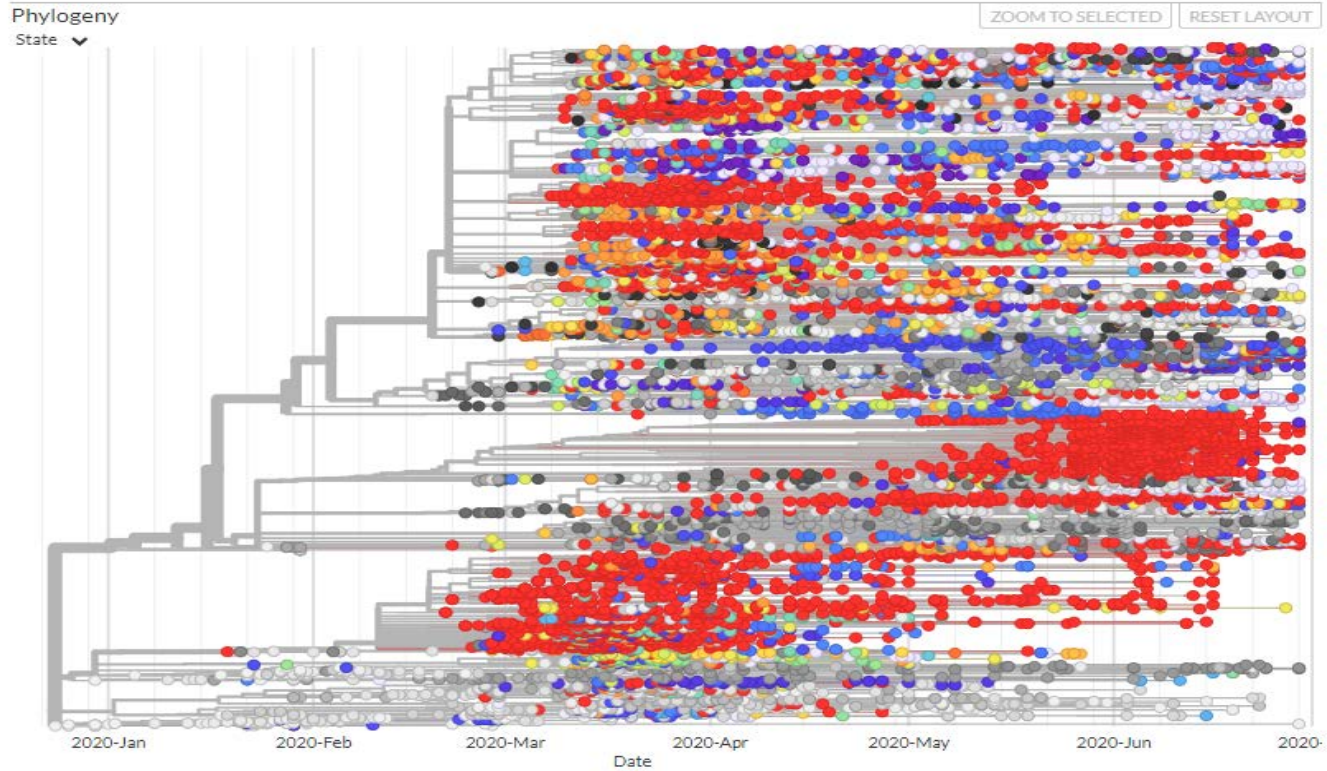
## Part 3: Implementation

- 3.1 Getting started with Nextstrain
- 3.2 Getting started with MicrobeTrace
- 3.3 Phylogenetics with USHER
- 3.4 Walking through Nextstrain trees**

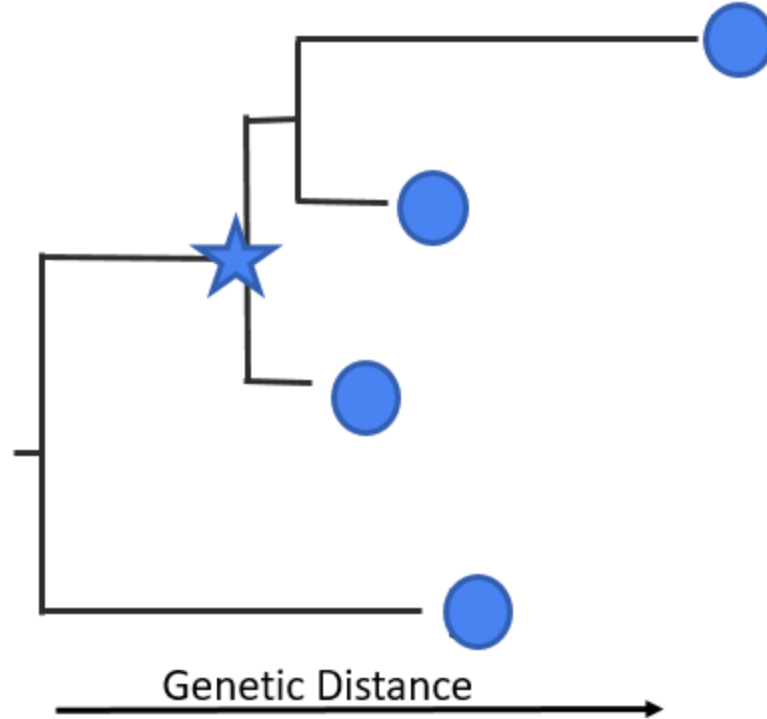
# What is Nextstrain?

- 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



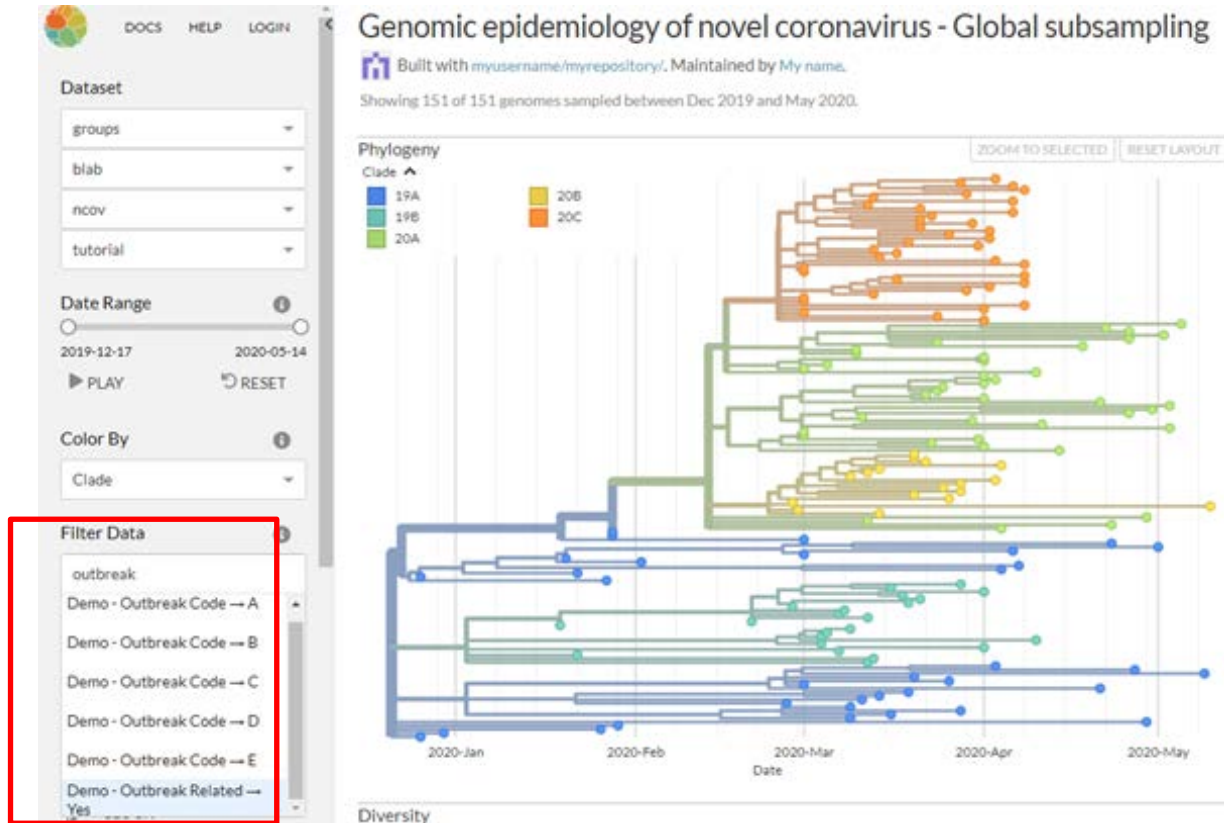
# Tree Terminology



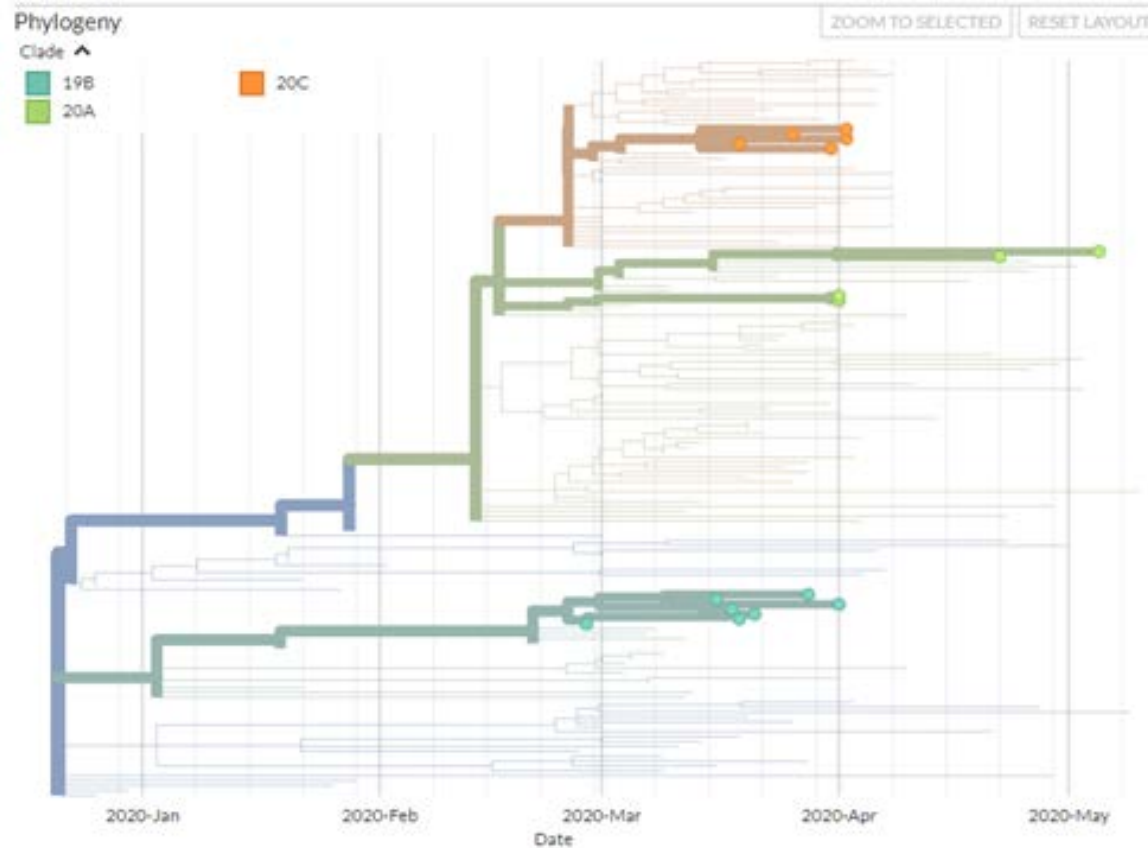
# Live - Demo

- 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

Showing 7 of 151 genomes sampled between Feb 2020 and Apr 2020. Filtered to { D (5)  , E (2)   } ∩

Yes (16)  .

Phylogeny

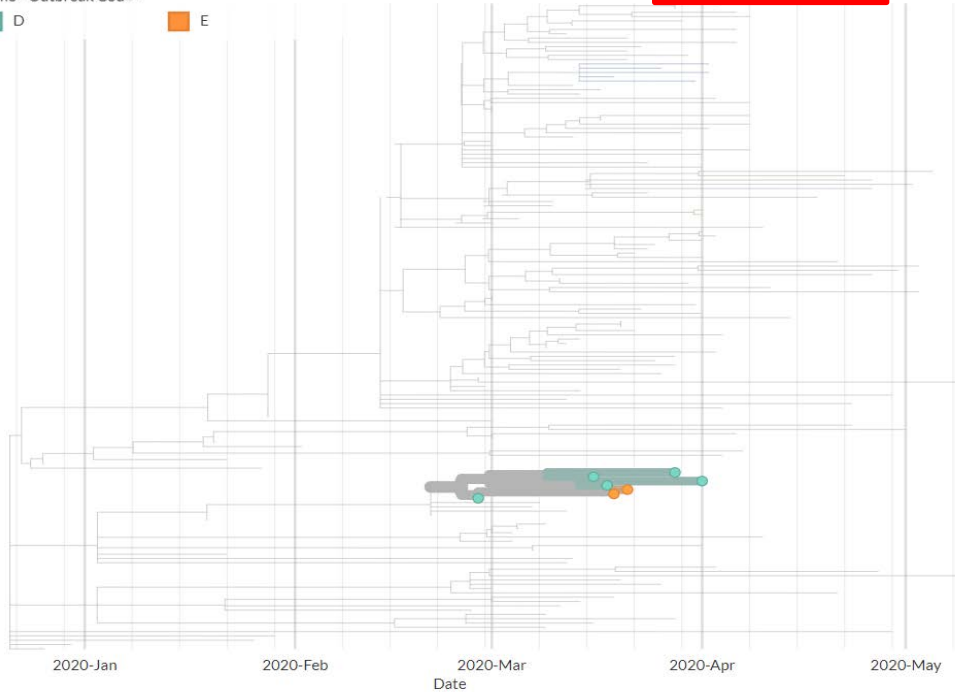
Demo - Outbreak Cod<sup>▲</sup>

 D

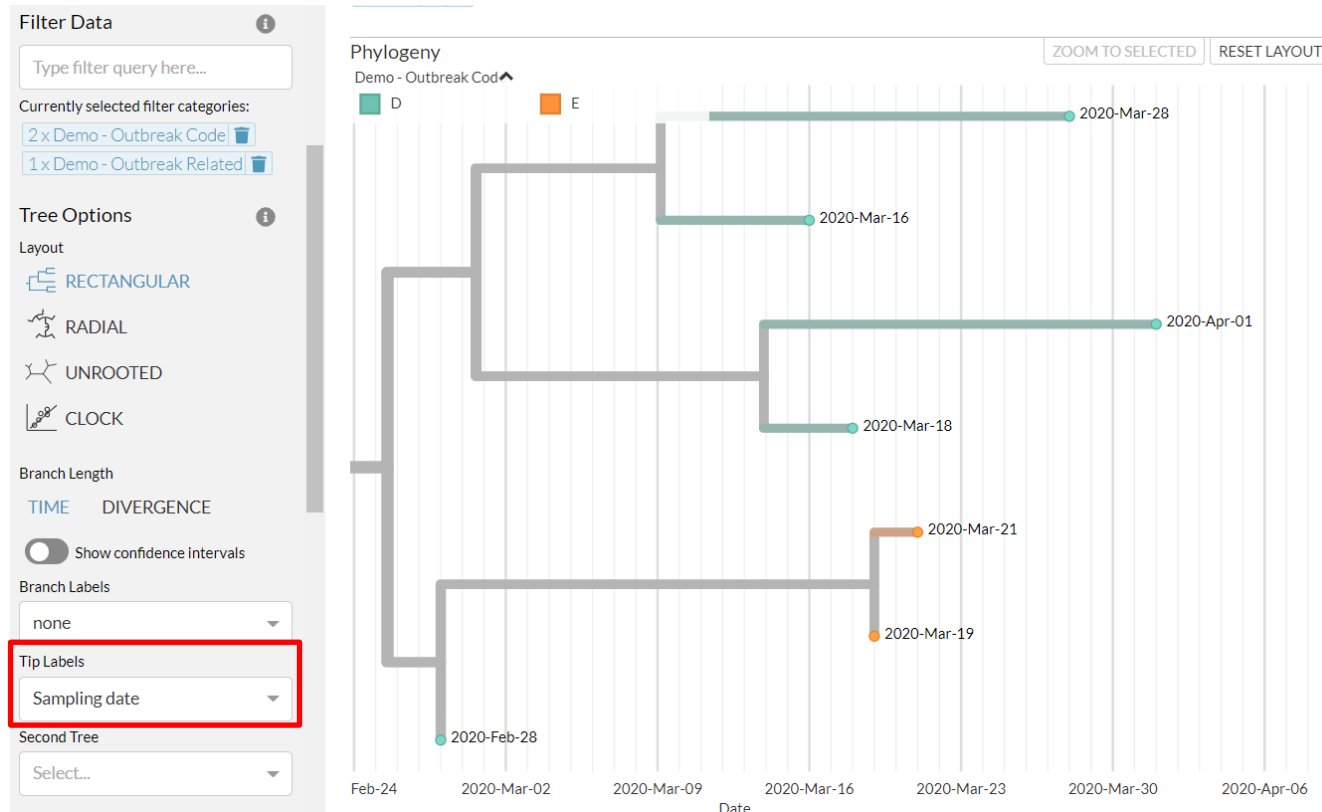
 E

ZOOM TO SELECTED

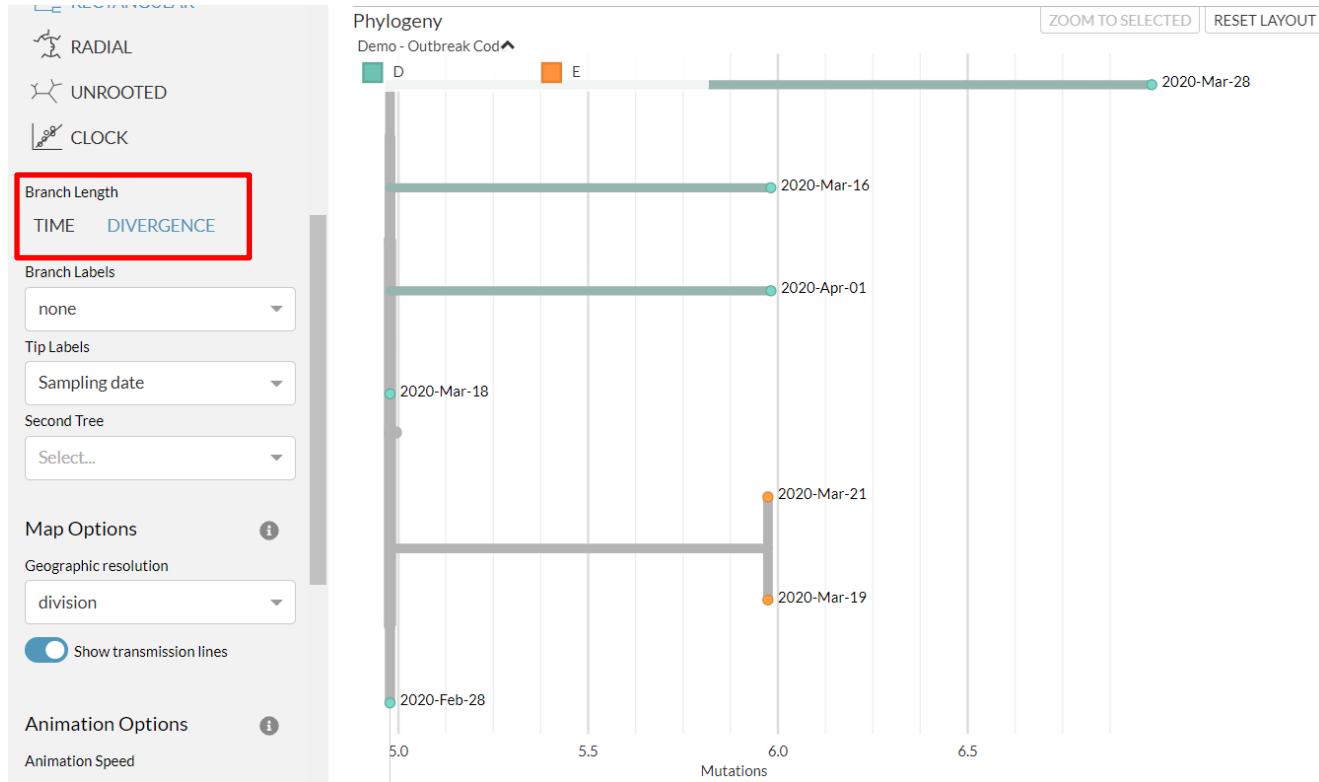
RESET LAYOUT



# 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

Data updated 2021-01-05 • [DOWNLOAD DATA](#) • Auspice v2.21.0  
Nextstrain: Hadfield et al, *Bioinformatics* (2018)

## Download data:

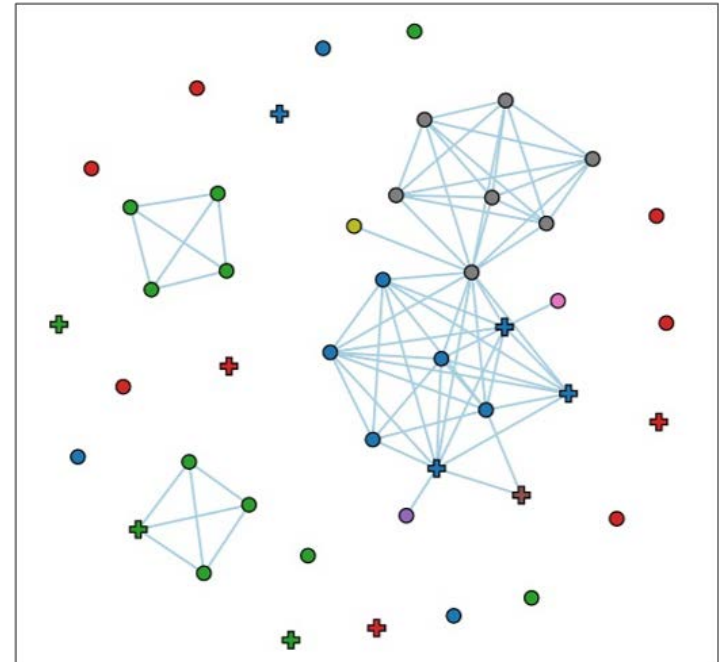
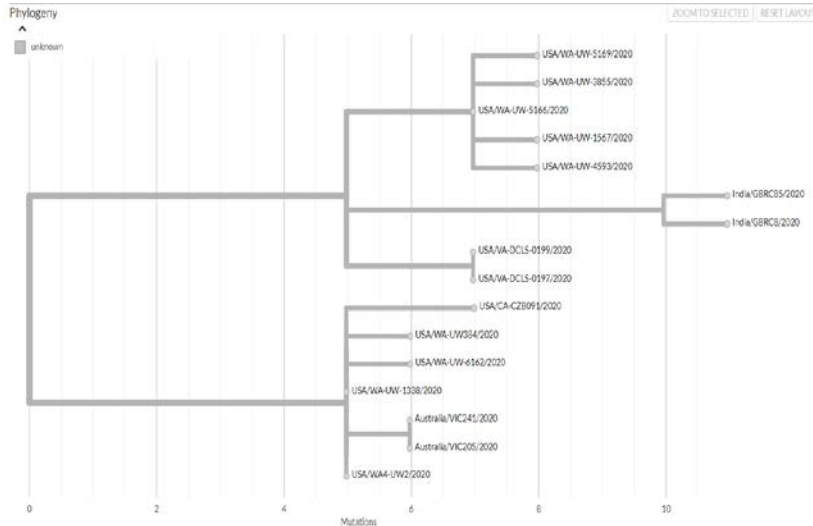
Downloaded data represents the currently displayed view. By zooming the tree, changing the branch-length metric, applying filters etc, the downloaded data will change accordingly.

Currently 16/151 tips are displayed and will be downloaded.

 TREE (NEWICK)	<i>Phylogenetic tree in Newick format with branch lengths in units of divergence.</i>
 TREE (NEXUS)	<i>Phylogeny in Nexus format with branch lengths in units of divergence. Colorings are included as annotations.</i>
 METADATA (TSV)	<i>Per-sample metadata (n = 16).</i>
 AUTHOR METADATA (TSV)	<i>Metadata for 16 samples grouped by their 16 authors.</i>
 GENETIC DIVERSITY DATA (TSV)	<i>The data behind the diversity panel showing normalised shannon entropy per codon.</i>
 SCREENSHOT (SVG)	<i>Screenshot of the current nextstrain display in SVG format; CC-BY licensed.</i>

# Interoperability with Auspice.us and MicrobeTrace

- Exported filtered data viewed in Auspice.us or MicrobeTrace for different perspectives



<https://auspice.us/>

# 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

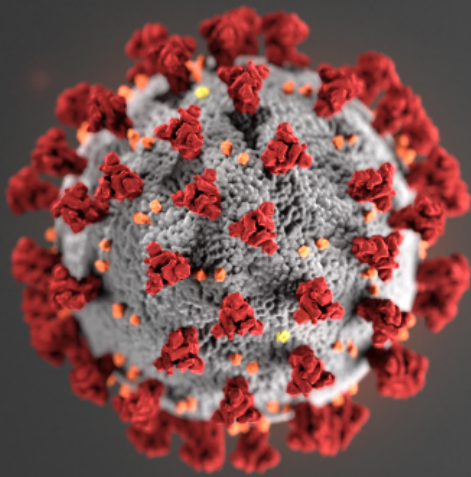
- NextStrain Tutorial
  - <https://nextstrain.org/groups/blab/ncov/tutorial>
- Metadata File
  - Module 3.4 - demo\_metadata.tsv



# 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](https://go.usa.gov/xAbMw)





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

