Nextstrain Tutorial

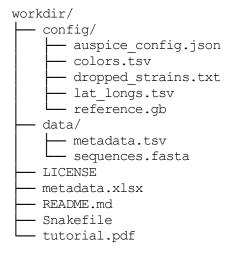
Yale school of public health

In this tutorial we are going to generate a nextstrain.org community build. To do so, we will run a bioinformatics pipeline (augur), which will perform sequence alignment, phylogenetic reconstruction, infer ancestral states (for location and mutations), and finally combine such results to be visualized using a tool named auspice.

- (1) Create a folder in your directory of preference: it will be your working directory (workdir).
- (2) Access the repository below and download its content as a .zip file:

github.com/grubaughlab/nextstrain course

- (3) **Decompress** the downloaded .zip file and copy its content into your workdir.
- (4) Now, your workdir should look as follows:



- (5) Open the file metadata.xlsx. This spreadsheet contains three tabs, named after some of the files in workdir, they are: metadata, lat longs, and colors.
- (6) Copy and paste the data from each tab in the corresponding files, replacing their existing data.
- (7) The current reference.gb file is included in config/ only as an example. Inspect the format of this file using a text editor, and check the fields 'CDS' and 'gene' in this annotation file. Access the link below, and download this new reference genome in .gb format (click on Send to > File > Format:Genbank > Create File).

ncbi.nlm.nih.gov/nuccore/JX669468

- (8) Open this .gb file on a text editor, and edit it as follows, keeping all the edited elements aligned, so that it looks like the file inspected at step (7):
 - a. Replace 'mat_peptide' by 'CDS'
 - b. Replace 'product' by 'gene'
 - c. Rename long gene names at 'gene' fields by standard gene acronyms (check the literature)
- (9) **Delete** the existing reference.gb; move the file created in the previous step to workdir/config, and; rename it as your new reference.gb file.
- (10) Open the file auspice_config.json using a text editor and edit the lines showing the following content:
 - a. 'Add your project title here'
 - b. 'Add your name here'
 - c. 'Add your github link repository here'
- (11) Open the file Snakefile using a text editor, and edit the following lines as appropriate, keeping their suffixes ('tree' and 'meta'):
 - a. 'addHereYourProjectName_tree.json'
 - b. 'addHereYourProjectName_meta.json'
- (12) Having done all the edits above, access your workdir in the Terminal, and run augur using the command below:

snakemake export

- (13) Watch for potential error messages: they may indicate issues with file formatting in previous steps.
- (14) If no issues were flagged, now run auspice to visualize the results using the command below:

auspice view

(15) **Open** a web browser and **access** the results using the link generated by the command above, which may look like the url below:

http://localhost:4000