**Nextstrain Tutorial**

In this tutorial we are going to generate a [nextstrain.org](http://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/andersonbrito/nextstrain\_course](https://github.com/andersonbrito/nextstrain_course)

(3) **Decompress** the downloaded .zip file and **copy** its content into your workdir.

(4) Now, your workdir should look as follows:

workdir/

├── config/

│ ├── auspice\_config.json

│ ├── colors.tsv

│ ├── dropped\_strains.txt

│ ├── lat\_longs.tsv

│ └── reference.gb

├── data/

│ ├── metadata.tsv

│ └── sequences.fasta

├── LICENSE

├── metadata.xlsx

├── README.md

├── Snakefile

└── tutorial\_nextstrain.pdf

(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](https://www.ncbi.nlm.nih.gov/nuccore/JX669468)

(8) **Open** this .gb file on a text editor, and **edit** it as follows, so that it looks like the example file inspected at step (7):

1. **Replace** ‘mat\_peptide’ by ‘CDS’
2. **Replace** ‘product’ by ‘gene’
3. **Rename** long gene names in the ‘gene’ fields by standard gene acronyms (see 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:

1. ‘Add your project title here’
2. ‘Add your name here’
3. ‘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’):

1. ‘addHereYourProjectName\_tree.json’
2. ‘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