Example data formats for the sDIV project *sEcoEvo: Biodiversity Dynamics – The Nexus Between Space & Time.*

We are providing example data for all the different data types we are interested in gathering for communities.

- Time calibrated trees (with branch lengths even if unresolved)

  An example tree in Newick format is in the file **example.newick**. Please note that trees can include both sampled and unsampled lineages (the example tree includes 7 lineages that were not sampled in the local community). The structure of the tree looks like this:

- Abundances

  We provide an example site-by-abundance file including 3 sites and the 5 sampled species (B, D, E, F, & K) in the file **example_abundances.txt**. This is a simple tab-delimited text file with rows corresponding to sites, and columns corresponding to species.

```
##      B D E F K
## site1 5 5 5 5 5
## site2 5 5 5 5 5
## site3 5 5 5 5 5
```

- Per location per taxon sequence data (including identical sequences).

Example sequence data is shown in the **fastqs/** directory, with one fastq file per site, including all sequences for all individuals sequenced at that site. We will use these sequences to calculate nucleotide diversity per species per site.

- Sample/species/site mapping file

We include one additional example file which is a mapping of sample names to species ID and sampling site called **example_pops.txt**. This is a tab delimited text file with an optional header where each row contains 3 items, the sample id (which should match the sample name in the fastq files), the species ID (which should match the ID in the newick tree), and the sample site (which should match the site name in the abundances file).