Example data formats for the sDIV project *sEcoEvo: Biodiversity Dynamics – The Nexus Between Space & Time*
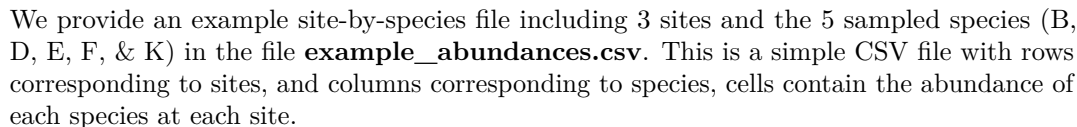
Here we provide example data for all the different data types we are interested in gathering for communities.

- 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-species file including 3 sites and the 5 sampled species (B, D, E, F, & K) in the file **example_abundances.csv**. This is a simple CSV file with rows corresponding to sites, and columns corresponding to species, cells contain the abundance of each species at each site.

	B	D	E	F	K
site1	0	8	2	5	7
site2	8	0	8	1	2
site3	0	0	3	8	15

- 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

    Sequence data need to be matched to species and sites. We include an example file called **example_pops.csv** which does this. This is CSV 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).

sample_ID	species	site
B_0_1	B	site1
B_1_1	B	site1

sample_ID	species	site
B_2_1	B	site1
B_3_1	B	site1
B_4_1	B	site1
D_0_1	D	site1

- Site metadata file

    We need to know a few things about sites, most importantly their goegraphic locations, and hopefully also something about sampling methodology such as area, effort, sampling technique. The included example file **site_metadata.csv** refers to arthropod pitfall trapping as an example.

site	lon	lat	method	duration
site1	-155.5134	19.23092	pitfall	5 days
site2	-155.5167	19.23250	pitfall	5 days
site3	-155.5155	19.23053	pitfall	5 days