**S2 Protocol**

1. check\_id\_map.py (use original 454 fna/qual files and metadata file)

2. split\_libraries.py (removes short and low quality sequences, adds barcodes names) 🡪 seqs.fna

3. identify\_chimeric\_seqs.py (input seqs.fna and 454.fna [to ID chimeras based on parent sequences] files) 🡪 chimeras.txt

4. pick\_otus.py (input seqs.fna to pick otus denovo, without reference file) 🡪 seqs\_otus.txt

5. pick\_rep\_set.py (input seqs\_otus.txt; this step is usually done after denovo otu picking and *not ref otu picking*) 🡪 rep\_set.fna

6. assign\_taxonomy (input rep\_set.fna; need reference database [fasta or fna] and taxonomy id [txt] files to assign. Default BLAST method is used for taxonomy assignment) 🡪 rep\_set\_tax\_assignment.txt

7. make\_otu\_table.py (input rep\_set\_tax\_assign.txt. This otu table includes taxonomic assignments) 🡪 otu\_table.biom

8. filter\_otus\_from\_otu\_table.py (input otu\_table.biom. Singleton removal; this filters sequences appearing < 3 times if –n 3 is passed) 🡪 otu\_table\_no\_singletons.biom

9. filter\_otus\_from\_otu\_table.py (input otu\_table\_no\_singletons.biom and chimeras.txt [step 4]. Removes chimeras)

🡪 otu\_table\_no\_chimeras.biom

10. add\_metadata.py (input otu\_table\_no\_chimeras.biom and a txt file containing information to add (i.e. accession number for taxonomy assignments) 🡪 otu\_table\_metadata.biom

11. split\_otu\_table.py (input otu\_table\_metadata.biom and metadata file [step 1] to choose the category to split .biom file)

🡪 otu\_split\_metadata.biom

12. split\_otu\_table\_by\_taxonomy (input otu\_table\_metadata.biom to split specifically by taxonomic level. Adding –n 3 will split by phylum 🡪 output otu\_table\_Metazoans.biom

13. print\_biom\_table\_summary.py (input otu\_table\_Metazoans.biom, generates statistics in command window for copy to text file)

14. convert\_biome.py (input any biom file [steps 9,10,11,12], will convert files into tab delimitated file read by excel) 🡪 otu\_table.txt

15. align\_seqs.py (input rep\_set.fna [step 5] to use PYNAST method to align the representative sequences) 🡪 rep\_set\_aligned.fasta

16. filter\_aignment.py (input rep\_set\_aligned.fasta file to filter alignment) 🡪 rep\_set\_aligned\_pfiltered.fasta

17. make\_phylogeny.py (input rep\_set\_aligned\_pfiltered.fasta to make tree file needed to run some alpha/beta diversity analyses)

🡪 rep\_phylo.tre

18. summarize\_taxa\_through\_plots.py (input otu\_table\_Metazoans.biom [step 12] and metadata file [step 1] to generate plots based on a category specified in the metadata file. Add –c and name of category to specify separation by that category [i.e. mesh size, location]

🡪 per\_study\_otu\_tables.biom