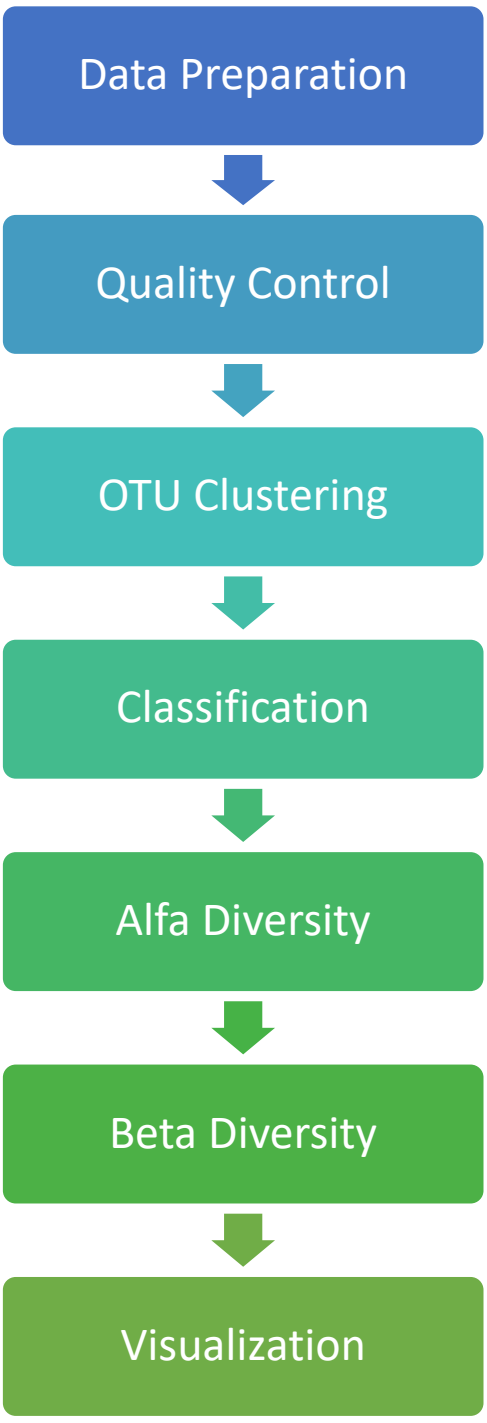
Student: Jan Izquierdo Ramos

PRACTICAL – EXIT TICKET

We applied this workflow in the practical. Fill it with the format file and software used in each step:



GALAXY TOOLS: List of Dataset Pairs

GALAXY TOOLS: Screen.seqs

GALAXY TOOLS: Cluster.split

GALAXY TOOLS: Classify.otu

GALAXY TOOLS: Rarefaction.single

GALAXY TOOLS: Dist.shared

GALAXY TOOLS: Make.biom

Describe the experimental design of the data set used in the practical session

Regularly testing mice feces to determine if a weight change alters the microbiota in their stomach
This is to analyze the variation in gut microbiota and its effect on the host.

How does the duplicate sequences removal works?

As microbiome samples include many repetitions of an organism we need to detect these sequences that are very similar and count each organism and how many repeats it has

Why is the OTU clustering one of the main steps on the analysis? What would you expect (in terms of diversity) if we apply an identity threshold < 97%?

Because OTUs cluster similar variations of 16s rDNA markers, where clusters represent taxonomic units.

The percentage is 97% because its the % needed to discriminate at species level

Explain the importance of rarefaction curves. If you have a sample which show a rarefaction curve which has not started to level off, how would you solve this problem? Give at least two possible solutions

These curves help estimate species diversity and asses sampling availability, when the curve is not fully leveled, it indicates incomplete sampling, missing biodiversity.

To solve the problem one option would be to increase the number of samples and sample size, the other option would be to find and use data from other studies that have researched similar samples with similar motives that the one of our study.