Important steps in the bioinformatics of metabarcoding data. The steps are listed in the order often used in, for instance, the dada2 package for R. The order of some of the steps might differ between pipelines. Common pipelines not mentioned in the table but still incorporate the same steps include qiime2 (https://qiime2.org/), mothur (https://github.com/mothur/mothur/), mjolnir3 (https://github.com/metabarpark/MJOLNIR3/), Obitools (https://git.metabarcoding.org/obitools/obitools4/) and Lotus2 (https://lotus2.earlham.ac.uk/), and more. The software and functions covered in the course are in bold italics. Functions in R are listed as package::function().

STEP	PURPOSE	SOFTWARE / FUNCTION
QC SEQUENCING RESULTS	Assess sequencing output quality	- fastQC, mulitQC
DEMULTIPLEXING	Assign reads to samples based on barcodes	- ***cutadapt*** - qiime demux, bcl2fastq, ++
QUALITY CONTROL, FILTERING, TRIMMING	Remove low-quality reads, trim adapters/primers	- ***dada2::filterAndTrim()*** - cutadapt,
DEREPLICATION	Collapse identical reads to reduce redundancy	- ***dada2::derepFastq()*** - vsearch --derep_fulllength
DENOISING / OTU CLUSTERING	Remove sequencing errors (called ASVs in DADA2), clustering into OTUs	- ***dada2::learnErrors(), dada2::dada()*** - ***vsearch --cluster_size*** - ***swarm***
CHIMERA REMOVAL	Remove chimeric sequences formed during PCR	- ***dada2::removeBimeraDenovo()*** - ***vsearch --uchime_denovo***
OTU TABLE CONSTRUCTION	Create count matrix assigning reads to OTUs/ASVs per sample	- ***dada2::makeSequenceTable()*** - vsearch --usearch_global --otutabout
OTU TABLE CURATION		- ***lulu::lulu()*** - ***mumu***
TAXONOMIC ASSIGNMENT	Assign taxonomy to ASVs or OTUs	- ***dada2::assignTaxonomy()*** - dada2::addSpecies(), - vsearch --usearch_global
REMOVAL OF NON-TARGETS	Exclude sequences from unwanted taxa (e.g., chloroplasts, host DNA, contaminants, mock data)	- Filter with bash scripts, python scripts or remove in excel ☺ - phyloseq::subset_taxa() - phyloseq::prune_taxa()
NORMALIZATION OR RAREFACTION	Normalize read counts across samples or rarefy to even depth	- vegan::rrarefey - phyloseq::rarefy_even_depth(),
DOWNSTREAM ANALYSIS AND PLOTTING	Alpha/beta diversity, ordination, visualizations	- Loads of R packages PCAtools, phyloseq, microbiome, vegan, ggplot2, metacoder