Metabarcoding Pipeline Building

An introduction with hands on exercises

CUSO – DPEE Activity

Gerhard Thallinger, PhD, Graz University of Technology Rachel Korn, PhD, Université de Fribourg Magdalena Steiner, PhD, Agroscope Wädenswil







Software

- mothur
- QIIME
- DADA2
- Stand-alone tools (e.g., UNOISE, USEARCH...)
- And many more!

mothur

- Command line program
- OTU (& ESV/ASV)
- Tutorials
- Integrates various bioinformatic & analysis tools
- Actively developed



QIIME



- Command line program + GUI
- OTU (& ASV)
- Tutorials
- Integrates various bioinformatic & analysis tools (including mothur & DADA2)
- Actively developed

Caporaso et al. 2010

DADA2



- "Divisive Amplicon Denoising Algorithm"
- R/Bioconductor based
- ASV
- Tutorials
- Actively developed

Callahan et al. 2016

mothur



- Command line program
- OTU (& ASV)
- Tutorials
- Integrates various bioinformatic & analysis tools
- Actively developed

Schloss et al. 2009

General Workflow I

Sample Splitting

Sequence Merging

Sequence Filtering

Dereplicate Sequences

Sequence Alignment

- Barcode based splitting of sequence file(s) (depends on sequencing provider), remove primers
- Build "contigs" based on overlapping forward and reverse read, trim reads
- Filter sequences based on "Ns", base quality, overall read quality, read length
- Remove duplicate sequences

Align contigs to reference

Kozich et al. 2013

General Workflow II



Filter Alignment

Precluster Sequences

Remove Chimeras

Distance Calculation

Clustering

- Remove columns with gaps in the alignment
- Precluster sequences to reduce noise and create ASVs/ ESVs
- Identify and remove chimeras
- Calculate pairwise distances for clustering

Cluster sequences at a given dissimilarity to build OTUs

Kozich et al. 2013

General Workflow III



Create OTU Table

Extract Representative Sequences

Taxonomic Classification

Merge OTU Table with Classification

Downstream Analysis

- Create OTU table per sample
- Extract a representative sequence per OUT cluster

- Classify the representative sequence
- Merge OTU table with classification

• Further analysis (alpha, beta, gamma diversity, ...)

Kozich et al. 2013

mothur Walkthrough

- Use the RStudio instance in your browser
- Switch to the terminal tab
- Open the file with the mothur commands
- Execute command(s) with
 <ctrl><alt><enter>

