Processing amoA amplicons from sequencing data

Dylan Bodington

2023-11-10

Table of Contents

16S rRNA gene amplicon sequences were trimmed with TrimGalore[1](#ref-krueger2021), a wrapper using CutAdapt[2](#ref-EJ200) for trimming of sequencing primers and adapters and FastQC[3](#ref-andrews2012) for quality control. Trimmed reads were merged and assigned to amplicon sequence variants (ASVs) using dada2[4](#ref-callahan2016), with quality filtering and chimera removal using the default settings (Table S7). Taxonomy was assigned to ASVs in dada2 using the Silva rRNA database (SSU Ref NR 99 v138.1)[5](#ref-quast2012). Community diversity (Table S8) and relative abundance (Table S9) of all sample replicates and fractions were calculated using phyloseq[6](#ref-mcmurdie2013).

1. Krueger, F., James, F., Ewels, P., Afyounian, E. & Schuster-Boeckler, B. TrimGalore. (2021). doi:<https://doi.org/10.5281/zenodo.5127899>

2. Martin, M. [Cutadapt removes adapter sequences from high-throughput sequencing reads](https://doi.org/10.14806/ej.17.1.200). *EMBnet.journal* **17,** 10–12 (2011).

3. Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C. & Wingett, S. FastQC. (2012).

4. Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A. & Holmes, S. P. [DADA2: High-resolution sample inference from Illumina amplicon data](https://doi.org/10.1038/nmeth.3869). *Nature Methods* **13,** 581–583 (2016).

5. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. & Glöckner, F. O. [The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools](https://doi.org/10.1093/nar/gks1219). *Nucleic Acids Research* **41,** D590–D596 (2012).

6. McMurdie, P. J. & Holmes, S. [Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data](https://doi.org/10.1371/journal.pone.0061217). *PloS One* **8,** e61217 (2013).