

## Supplementary methods

Reads containing Illumina Universal Adapters or PhiX were discarded using bbdduk version 37.55 (BBTools, <https://jgi.doe.gov/data-and-tools/bbtools>) and the remaining reads were demultiplexed using je version 1.2<sup>1</sup>. Indexes, heterogeneity spacers and primers were trimmed with cutadapt version 1.14<sup>2</sup> and the paired end reads were subsequently quality trimmed and merged using bbmerge version 37.55<sup>3</sup>. The merged contigs were quality filtered using default values in Quantitative Insights Into Microbial Ecology (QIIME) version 1.9.1<sup>4</sup>. Closed reference operational taxonomic unit (OTU) mapping to the Silva database<sup>5</sup> (version 128, reference OTUs pre-clustered at 97% sequence similarity) was performed using SortMeRNA version 2.0<sup>6</sup> through QIIME. OTUs with a number of sequences <0.005% of the total number of mapped sequences were discarded as recommended<sup>7</sup>. The samples were rarefied (subsampling) to an OTU count of 10768 per sample, and all further analyses were performed on this rarefied dataset.

1. Girardot C, Scholtalbers J, Sauer S, et al. Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. *BMC Bioinformatics* 2016; 17: 419. DOI: 10.1186/s12859-016-1284-2.
2. Martin M. *CUTADAPT removes adapter sequences from high-throughput sequencing reads*. 2011.
3. Bushnell B, Rood J and Singer E. BBMerge - Accurate paired shotgun read merging via overlap. *PLoS One* 2017; 12: e0185056. DOI: 10.1371/journal.pone.0185056.
4. Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010; 7: 335-336. DOI: 10.1038/nmeth.f.303.
5. Quast C, Pruesse E, Yilmaz P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2013; 41: D590-596. DOI: 10.1093/nar/gks1219.
6. Kopylova E, Noe L and Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 2012; 28: 3211-3217. DOI: 10.1093/bioinformatics/bts611.
7. Bokulich NA, Subramanian S, Faith JJ, et al. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods* 2013; 10: 57-59. DOI: 10.1038/nmeth.2276.