



benchdamic: benchmarking of differential abundance methods on microbiome data

RATIONALE

Issue:

Many approaches have been proposed for Differential Abundance (DA) analysis in metagenomics, it is widely recognised that the perfect method does not exist.



Solution:

A careful exploratory data analysis is necessary to address methodological choices. We benchmarked methods from bulk RNA-seq, metagenomics, and single-cell RNA-seq on metagenomics data.



Application:

benchdamic is a new R/Bioconductor package that provides a computational framework to guide researchers in the selection of the method that best fits their data.



IMPLEMENTATION

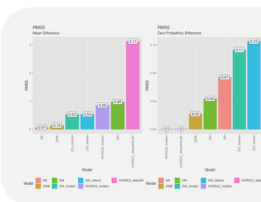
Goodness of fit

Question: Which are the parametric distributions that are able to fit both the proportion of zeros and the counts in your data?

Input: phyloseq or TreeSummarizedExperiment objects

Working function: FitModels

Output: plotRMSE, plotMD



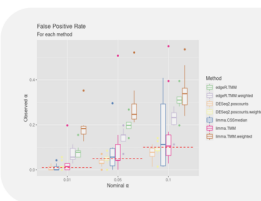
Type I Error Control

Question: Which are the DA methods that are able to control the number of false positives in your data?

Input: phyloseq or TreeSummarizedExperiment objects

Working function: createMocks, runMocks, createTIEC

Output: plotFPR, plotFDR, plotKS, plotQQ, plotLogP



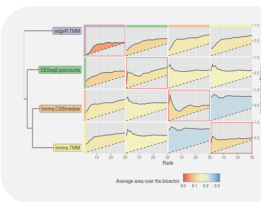
Concordance

Question: If multiple DA methods are run on the same data, would they be concordant?

Input: phyloseq or TreeSummarizedExperiment objects

Working function: createSplits, runSplits, createConcordance

Output: plotConcordance



Enrichment

Question: If some prior knowledge about your experiment is available, would the findings be coherent with that knowledge?

Input: phyloseq or TreeSummarizedExperiment objects

Working function: runDA, createEnrichment, createPositives

Output: plotEnrichment, plotContingency, plotMutualFindings, plotPositives

