European Molecular Biology Laboratory

Evaluation of microbiome association models under realistic and confounded conditions

Jakob Wirbel[1]†, Morgan Essex[23]†, Sofia K. Forslund[23], Georg Zeller[1]

1) Structural and Computational Biology Unit, EMBL Heidelberg, Germany; 2) Charité Universitäts-Medizin Berlin, Germany;
3) Max-Delbrück Center for Molecular Medicine, Berlin, Germany; † both authors contributed equally to this work

jakob.wirbel@embl.de

Scan here to see the bioRxiv preprint

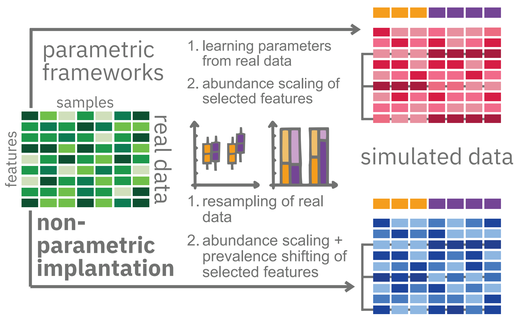
EMBL

Introduction

- Testing for differential abundance is a fundamental task in metagenome-wide association studies, yet **there is no consensus on statistical methodology**
- Benchmarks typically use reshuffled **real data**, which lacks a ground truth or **simulated data** generated by parametric methods, for which the resemblance to real metagenomic data is not clear
- **Technical or biological confounders** further hamper reliability and reproducibility of findings in clinical applications of differential abundance testing, yet have thus far been largely ignored in benchmarks

Methods and Study Design

- We developed a **novel simulation framework** mimicking case-control study designs that implants signals into **real metagenomic data**



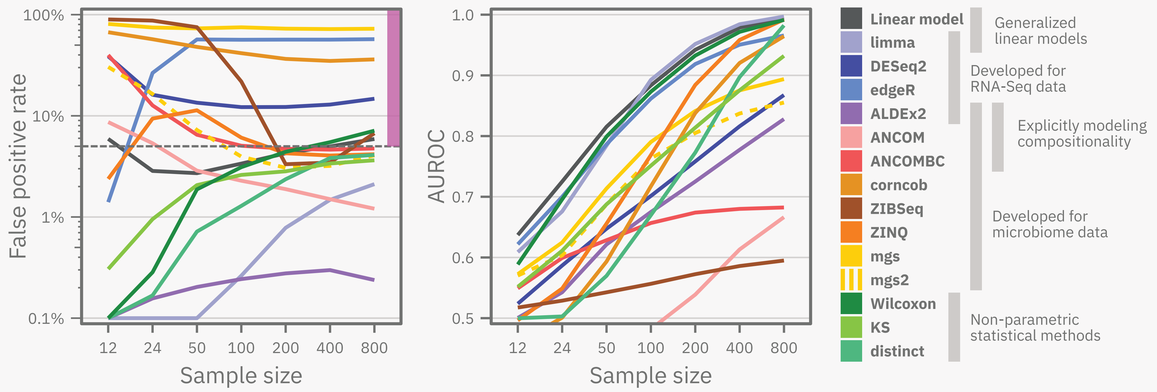
- We expanded our framework to simulate known patterns of confounding, mimicking both narrow (e.g. from medication) or broad (resembling study effects, for example) confounding
- Our code is open source to facilitate community benchmark development for other study designs, data types, or biomes

SIMBA (simulation of microbiome data with biological accuracy)

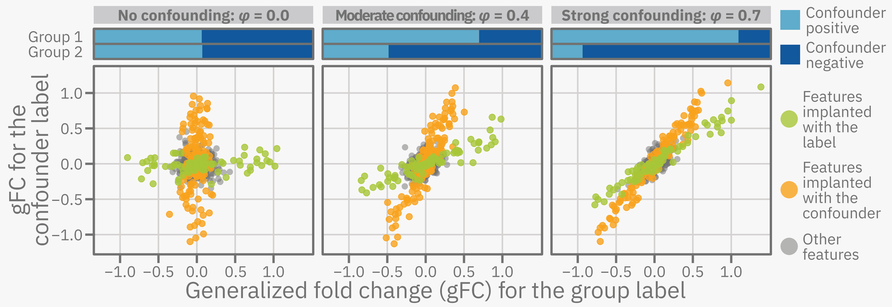
BAMBI (benchmarking analysis of microbiome inference methods)

Main Results

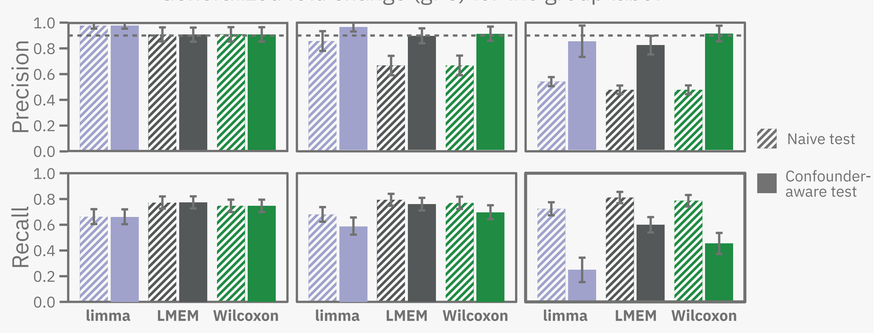
Most differential abundance testing methods applied in metagenome-wide association studies **fail to control the false positive rate or offer limited sensitivity** to detect biomarkers



Confounding factors (such as medication) can **exacerbate these issues.**

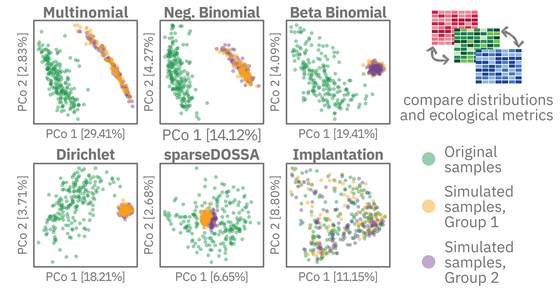


Linear mixed-effect models (LMEM) and the blocked Wilcoxon test can be adjusted for confounding covariates and **should be be preferred over other statistical methods** for the analysis of microbiome data.

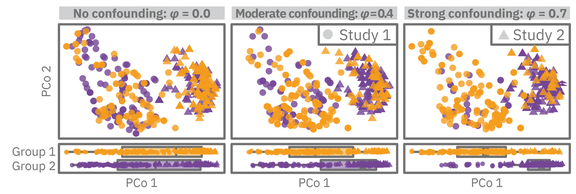


Other Results and Implications

- **Signal implantation**, but not other parametric simulations (Weiss et al. *Microbiome* 2017, Hawinkel et al. *Brief. Bioinform.* 2019, Ma et al. *bioRxiv* 2021), **can reproduce key characteristics of real metagenomic data**, critical for robust benchmarking



- Biased resampling produces simulated case-control groups aligned with a secondary confounding variable to a pre-calibrated degree (below: study effects/broad confounding, left: simulated medication/narrow confounding)



- **Linear mixed-effect models** offer the most flexibility and scalability in their implementation and **are capable of disentangling drug- and disease-associated microbial features** in real clinical data (Forslund et al. *Nature* 2021)