Robust Foundations for Clinical Host-Microbiome Data Analysis

PhD dissertation work from 2019-2023

Morgan Essex, 6 June 2023

Experimental & Clinical Research Center

MAX DELBRÜCK CENTER FOR MOLECULAR MEDICINE IN THE HELMHOLTZ ASSOCIATION

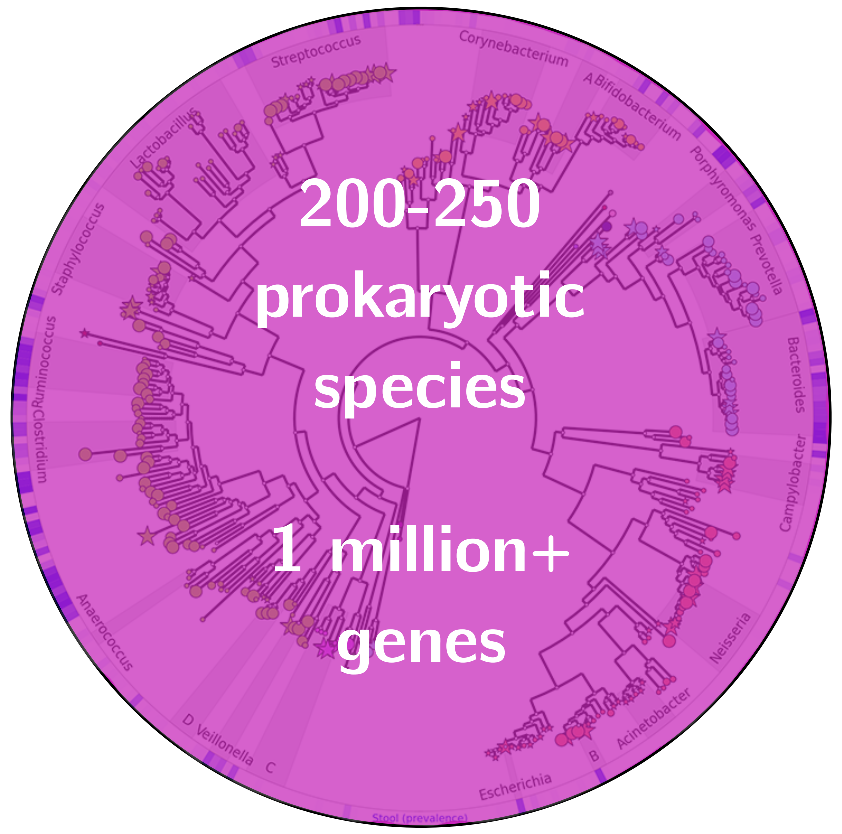
CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN

The human gut microbiome

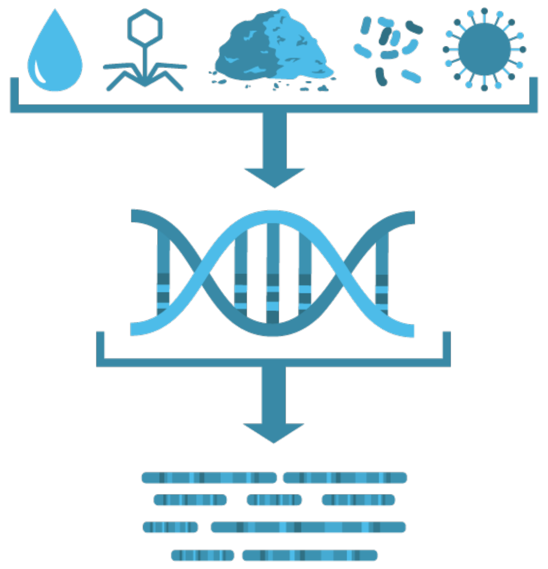
- Sequenced stool samples used as a proxy for the colonic microenvironment
- Culture-independent and reference-free approaches have uncovered enormous prokaryotic diversity
- "**Microbial organ**" with major roles in host physiology, metabolism, immunity
- Composition varies greatly even between healthy individuals

Statistical challenges ⟷ Personalized medicine

Human Genome
23,000 genes



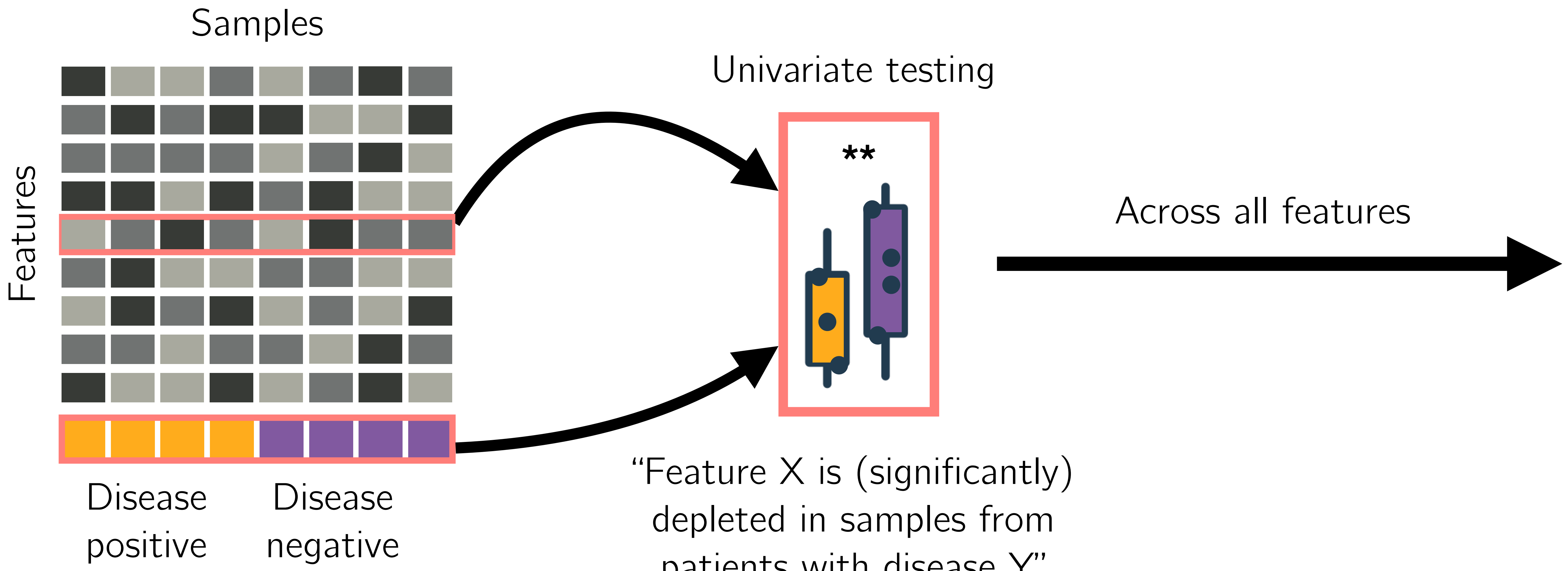
Metagenomics



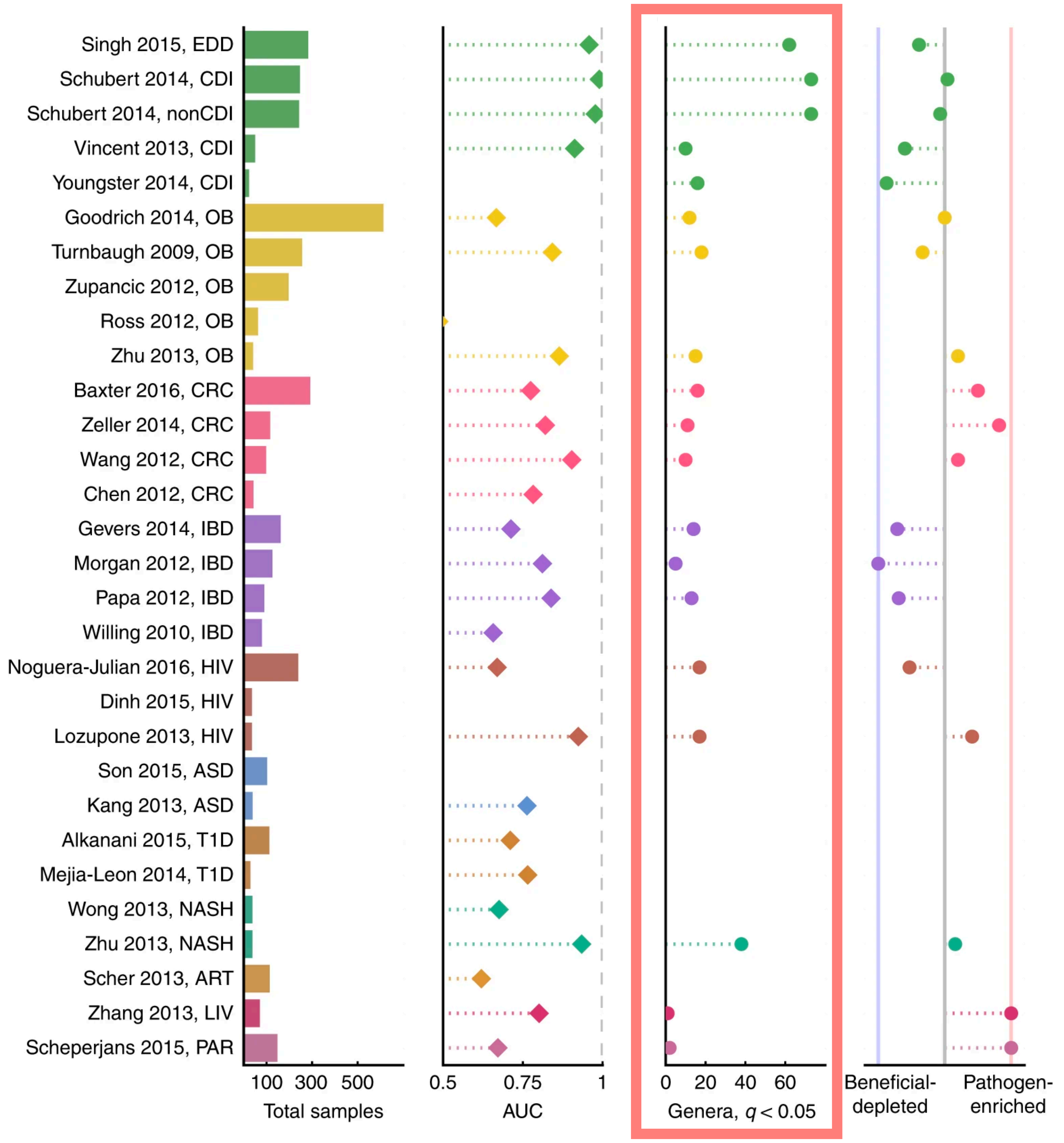
One way to analyze microbiome data

- Metagenomic sequencing produces data tables of per-sample microbial feature counts
- **Differential abundance** (DA) analysis is the process of grouping samples and statistically comparing them feature-by-feature (univariate)

Duvallet, C. *et al.* Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nature Communications* (2017)



"Feature X is (significantly) depleted in samples from patients with disease Y"

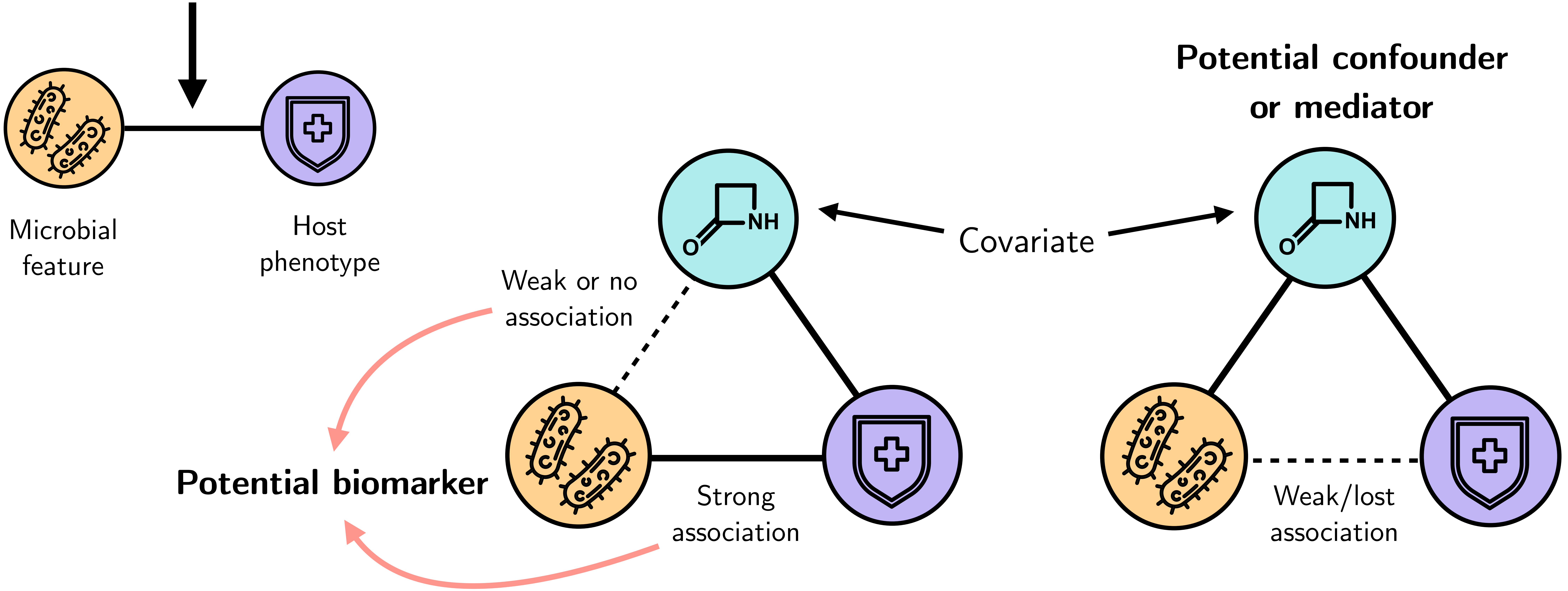


FDR-corrected!

“Correlation does not equal causation”

Metagenome-wide **association** studies

Goal: generate more precise, informed **hypotheses**



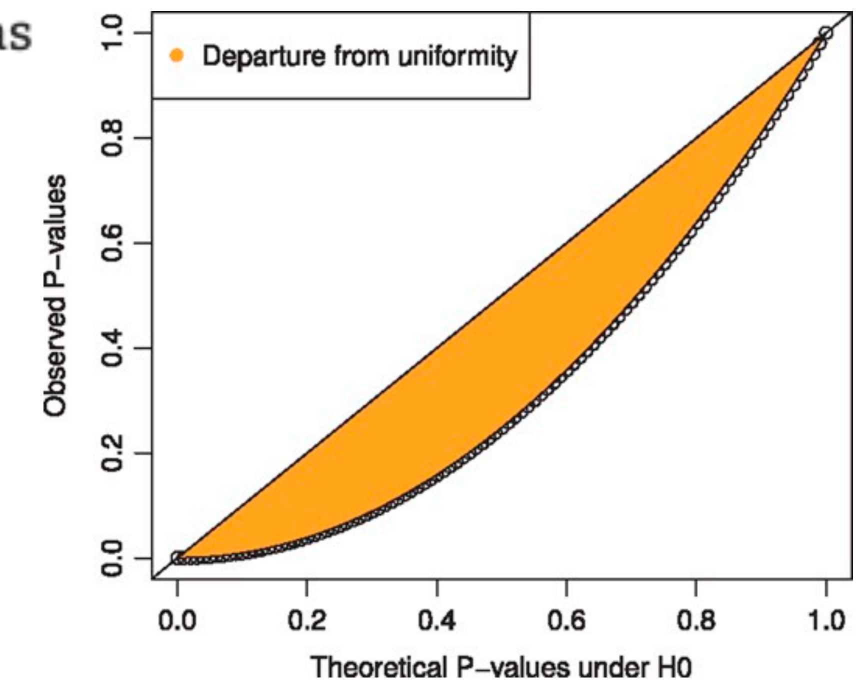
Back to empirical first principles (= basic propositions or assumptions)

- Sequenced microbiome data are **incredibly variable** from feature-to-feature, sample-to-sample, and dataset-to-dataset, for both **biological and technical reasons**
- Why do we use the negative binomial?
  - Count data are binomial
  - Sequenced data are overdispersed
- Shared **skepticism** regarding the use of **parametric methods** for simulation and testing of **taxonomic microbiome profiles**

A broken promise: microbiome differential abundance methods do not control the false discovery rate

Stijn Hawinkel, Federico Mattiello, Luc Bijnens and Olivier Thas

Briefings in Bioinformatics (2017) https://doi.org/10.1093/bib/bbx104



Sequence count data are poorly fit by the negative binomial distribution

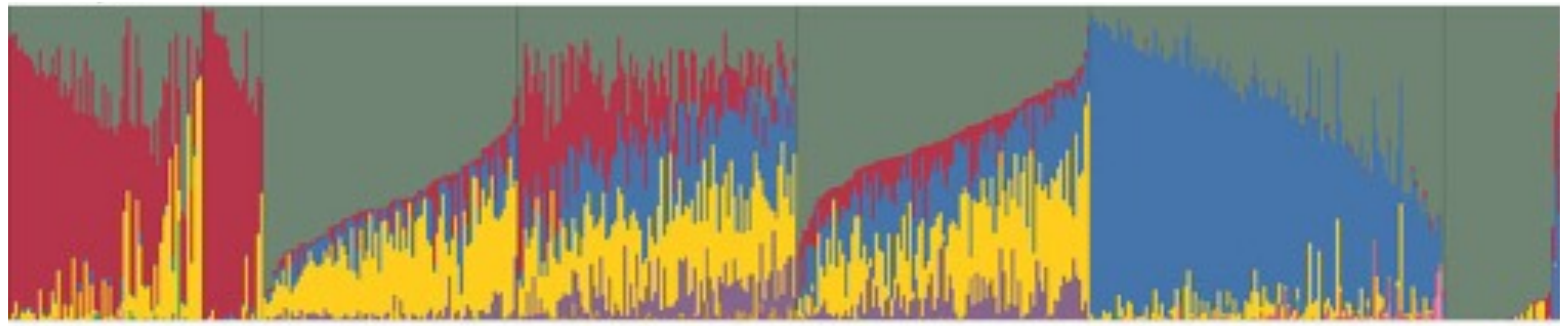
Stijn Hawinkel[1*], J. C. W. Rayner[2,5], Luc Bijnens[3,4], Olivier Thas[1,4,5]

Consequences

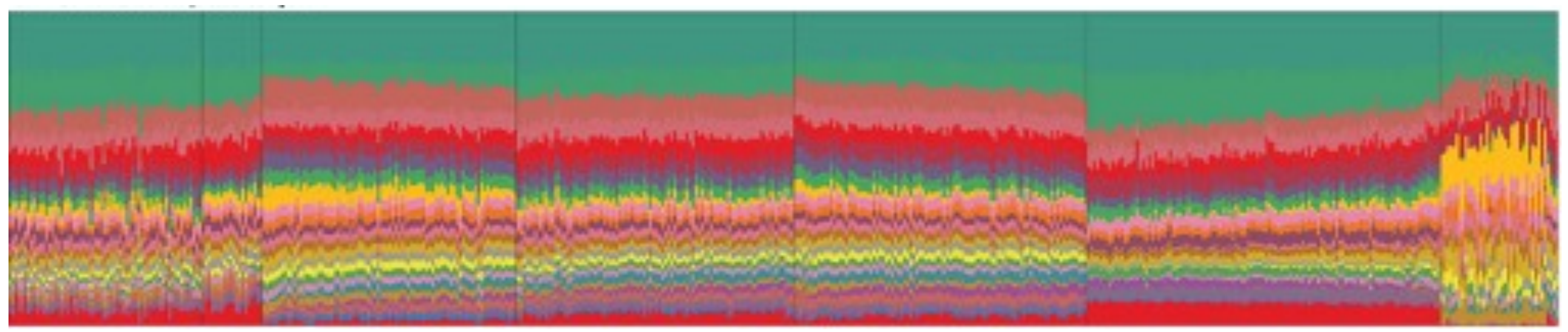
From the results from the previous section it is clear that many features in sequence count data do not follow the NB distribution. It has been shown before that the performance of statistical tests that rely on the NB distribution deteriorates when applied to realistically simulated data

PLoS One (2020) https://doi.org/10.1371/journal.pone.0224909

Phyla per sample and body site



Metabolic pathways per sample and body site



Nature (2012) https://doi.org/10.1371/journal.pone.0224909

Empirically-validated data for benchmarks

PLoS Comp. Bio (2014) https://doi.org/10.1371/journal.pcbi.1003531

Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible

Paul J. McMurdie, Susan Holmes*

BMC Genomics (2016) https://doi.org/10.1186/s12864-016-2386-y

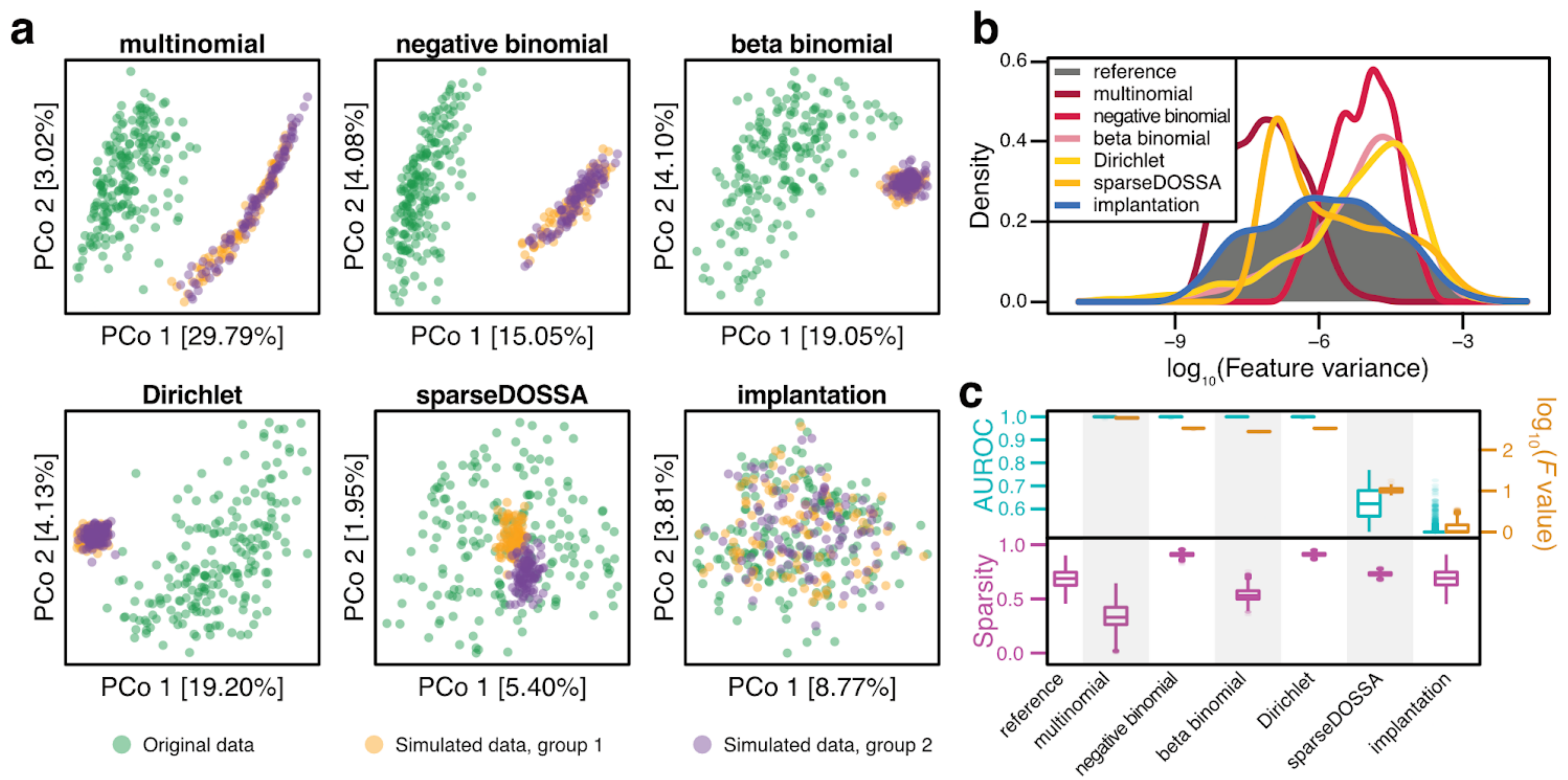
Statistical evaluation of methods for identification of differentially abundant genes in comparative metagenomics

Viktor Jonsson*, Tobias Österlund, Olle Nerman and Erik Kristiansson*

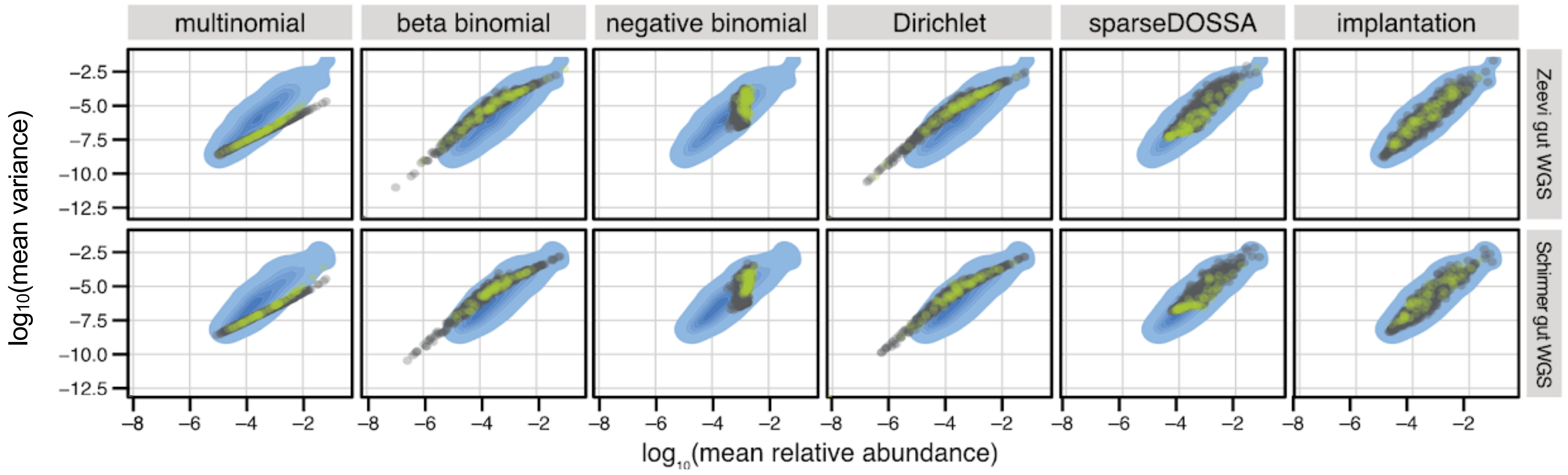
Microbiome (2017) https://doi.org/10.1186/s40168-017-0237-y

Normalization and microbial differential abundance strategies depend upon data characteristics

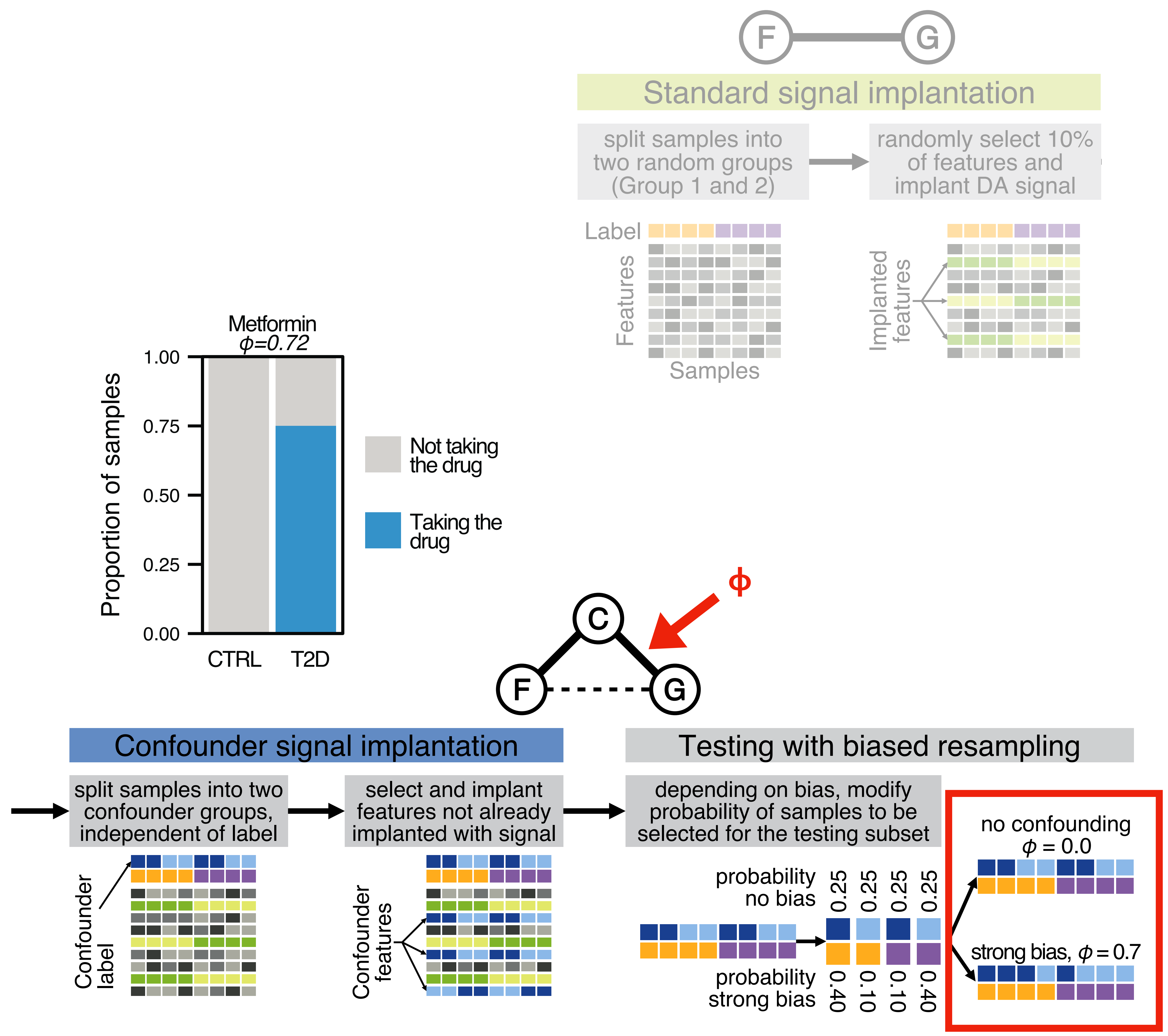
Sophie Weiss[1], Zhenjiang Zech Xu[2], Shyamal Peddada[3], Amnon Amir[2], Kyle Bittinger[4], Antonio Gonzalez[2], Catherine Lozupone[5], Jesse R. Zaneveld[6], Yoshiki Vázquez-Baeza[7], Amanda Birmingham[8], Embriette R. Hyde[2] and Rob Knight[2,7,9*]

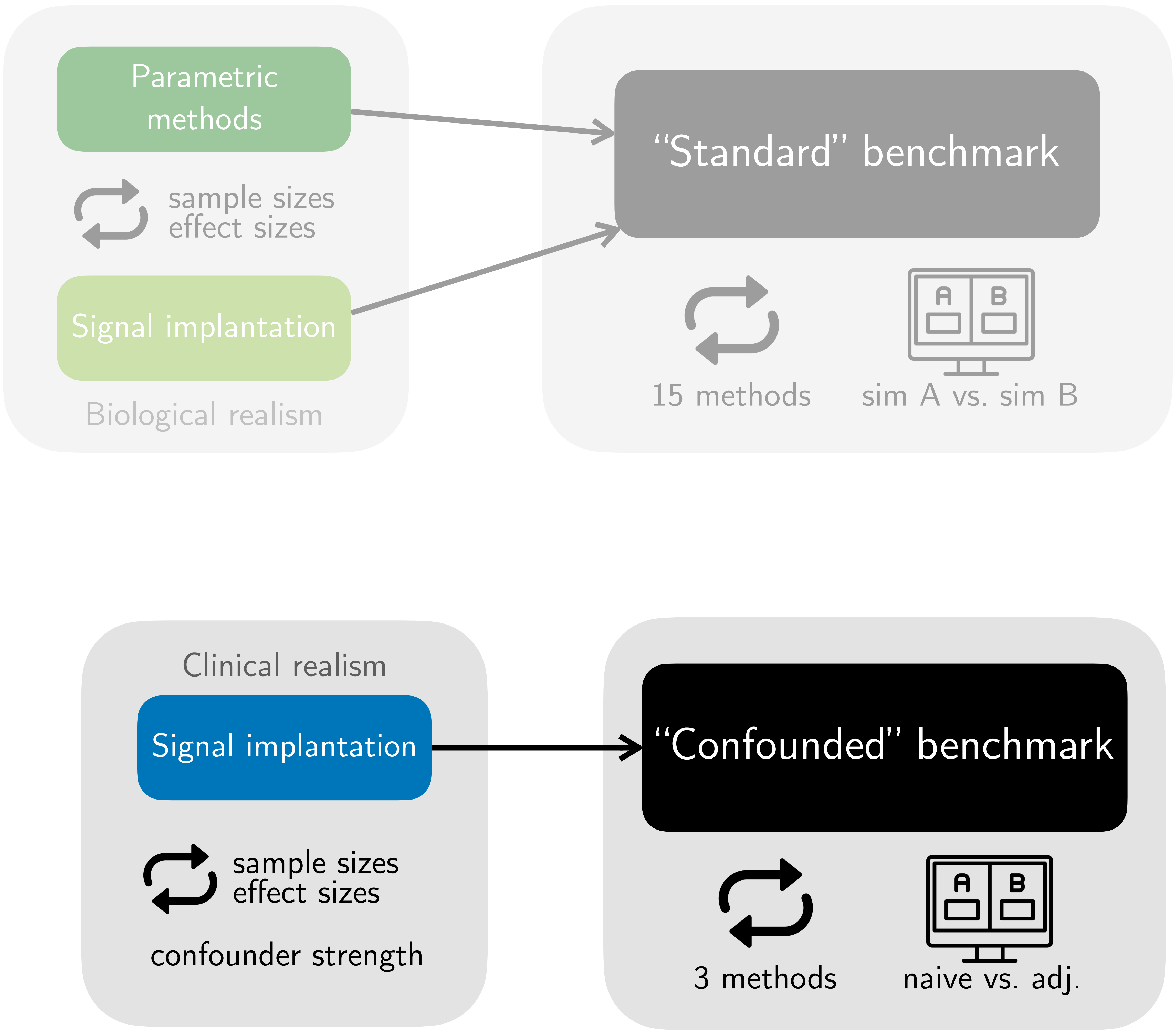


- These studies never compared their simulations to the input data!
- Feature variance, sparsity, and mean-variance relationships are drastically different



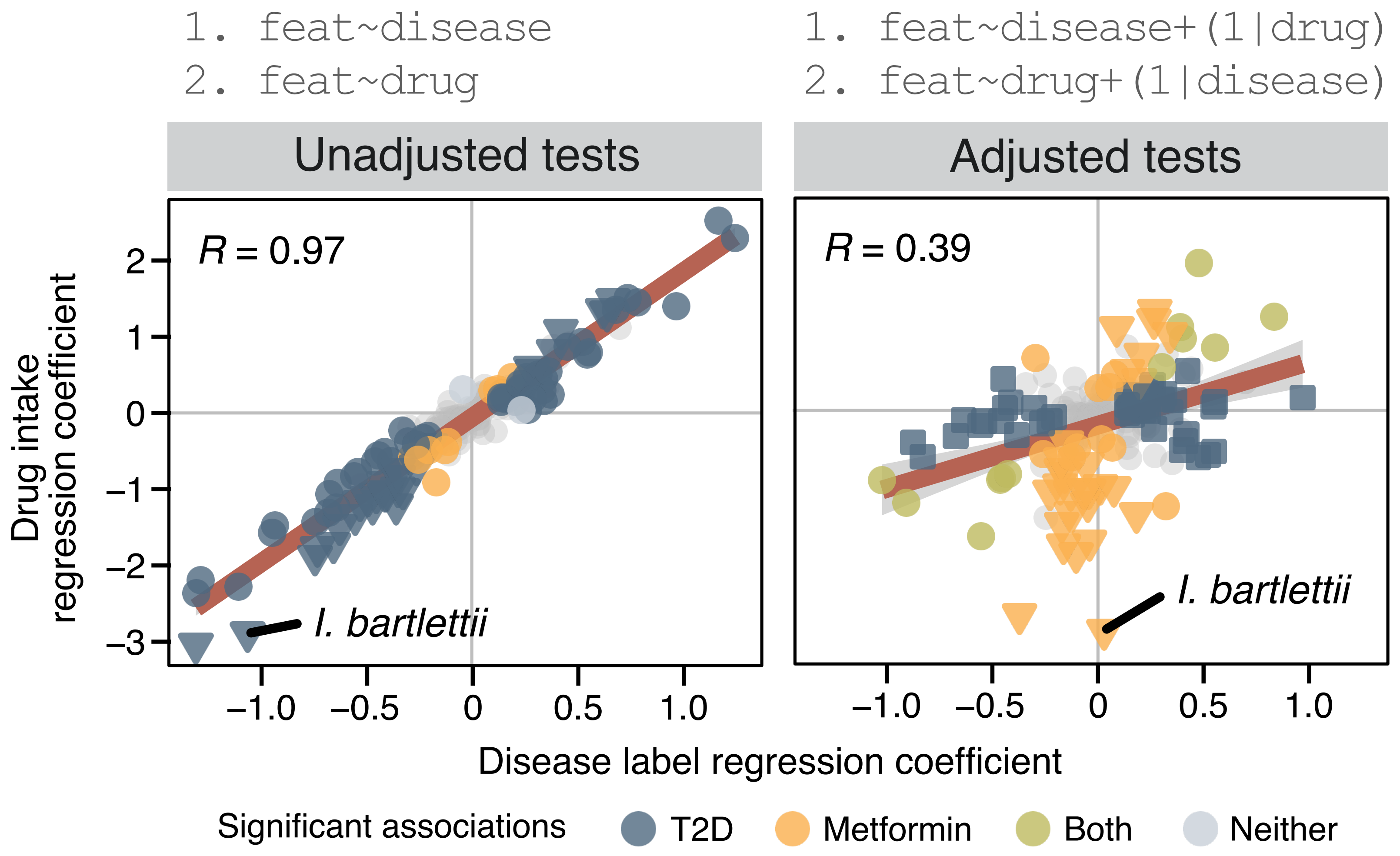
Building simulations and benchmarks for confounding DA - Results

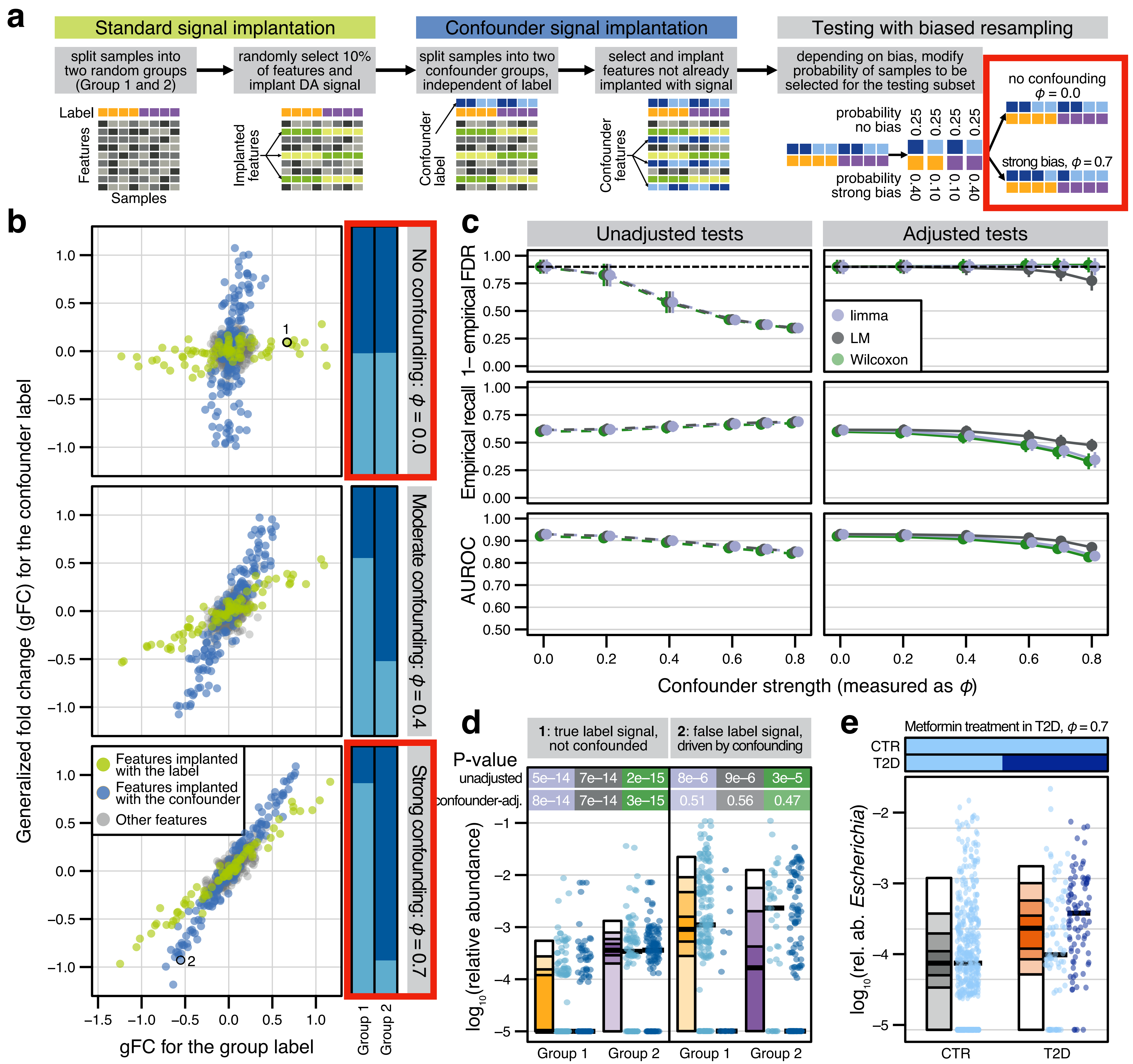




Linear models to diagnose & control confounding

- Phi is a measure of group (im)balance
- Confounding can be
  - visualized as an **'overlap' of signals**
  - quantified with **e.g. phi coefficients**
- Unchecked sources of variation (confounding) are the biggest reason why small datasets fail to reproduce





Final thoughts on differential abundance

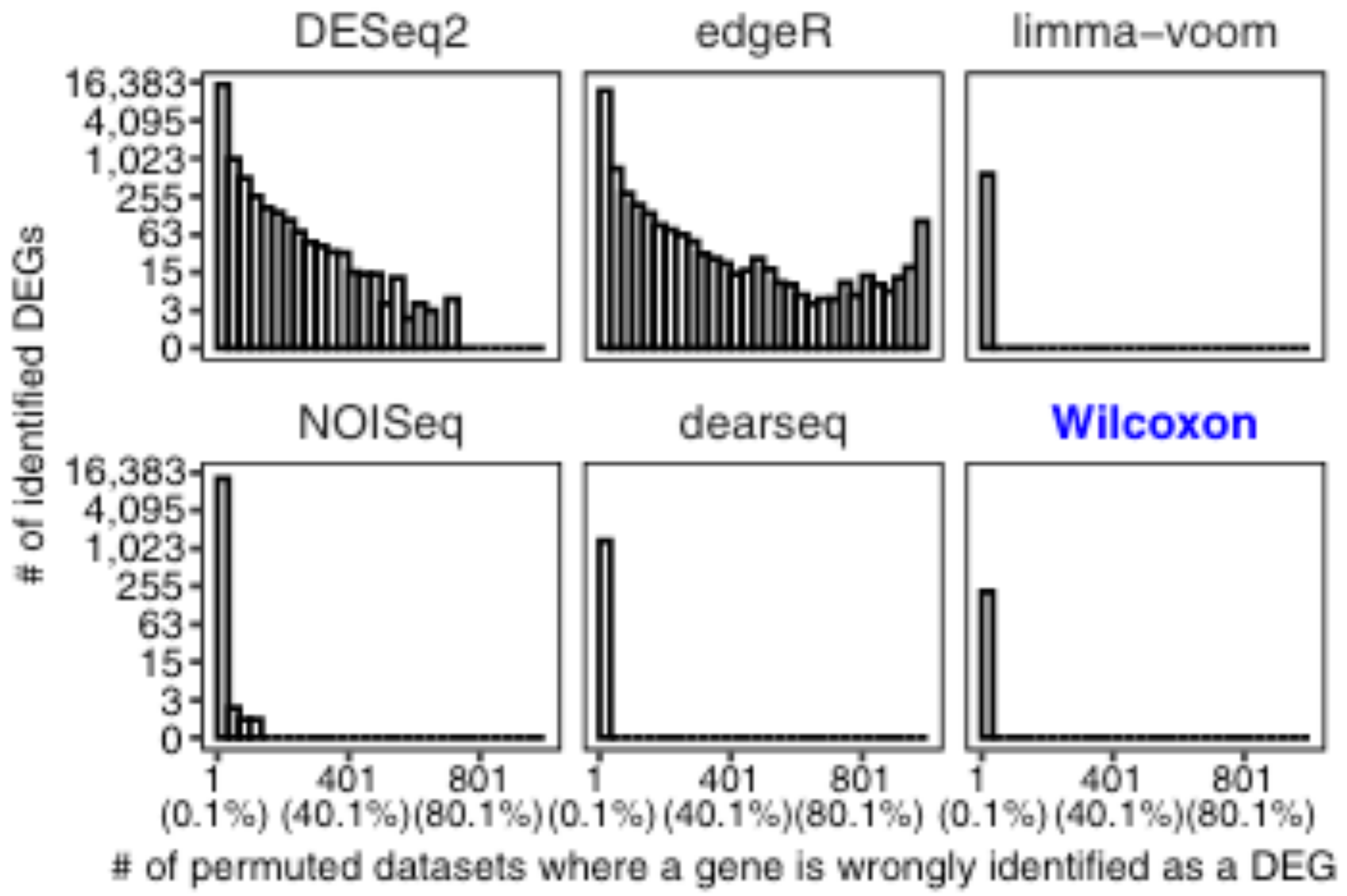
- RNAseq methods (*DESeq2, edgeR*) were developed and benchmarked with **very few replicates**
- Parametric methods are **not suitable** for large studies, or **human-associated microbiota** studies (esp. **clinical**)
- **Taxonomic profiles are unique** from functional profiles, cross-sections probably behave differently than **longitudinal data**

Genome Biology (2022) https://doi.org/10.1186/s13059-022-02648-4

Exaggerated false positives by popular differential expression methods when analyzing human population samples

Yumei Li[1†], Xinzhou Ge[2†], Fanglue Peng[3], Wei Li[1*] and Jingyi Jessica Li[2,4,5,6,7*]

In conclusion, **when the per-condition sample size is less than 8, parametric methods may be used** because their power advantage may outweigh their possibly exaggerated false positives [...] for large-sample-size data, the **Wilcoxon rank-sum test** is our recommended choice for its **solid FDR control and good power.**



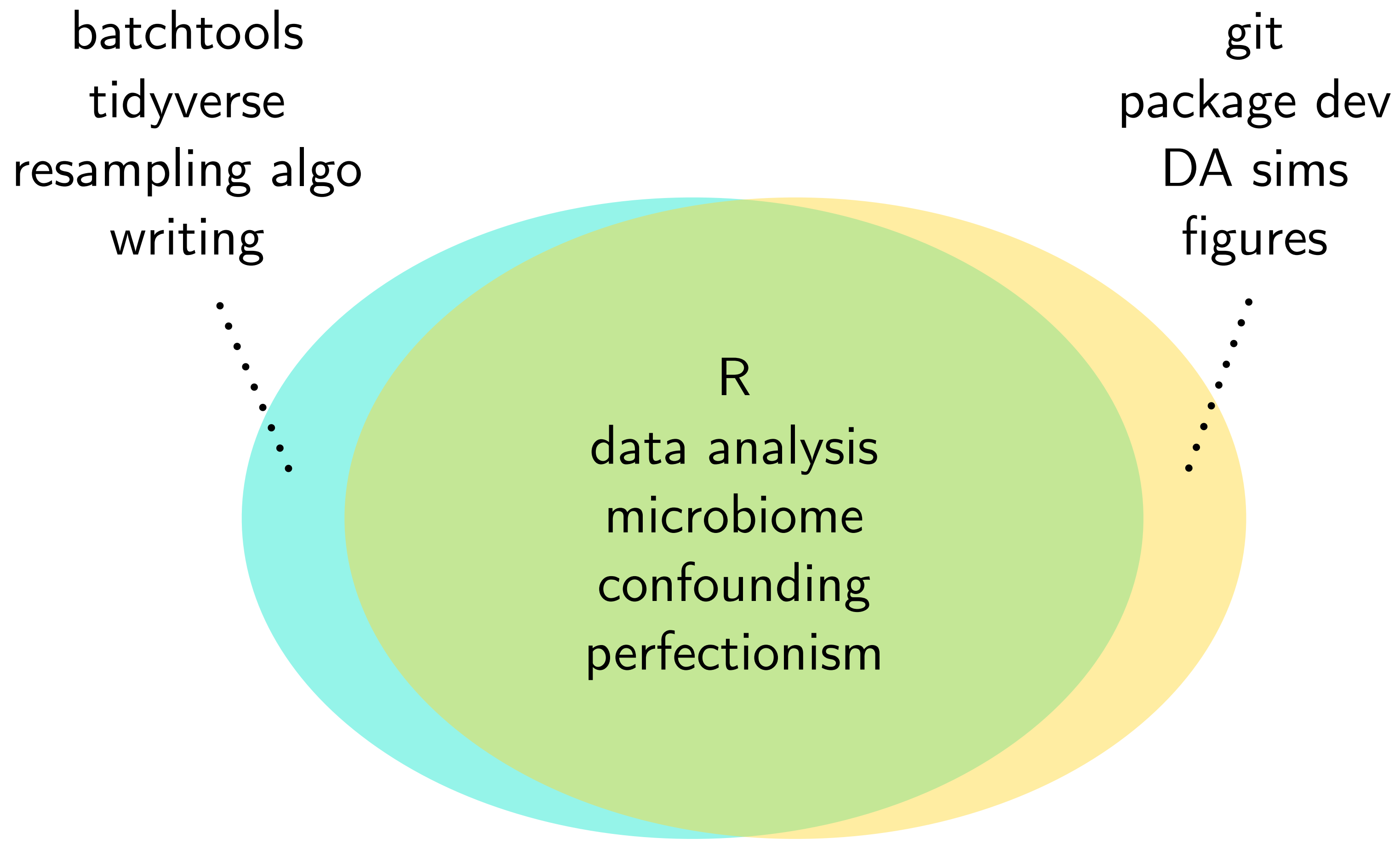
Overview of projects, papers, roles

Differential abundance (DA)

Evaluation of microbiome association models under realistic and confounded conditions

Jakob Wirbel, Morgan Essex, Sofia Kirke Forslund, Georg Zeller
doi: https://doi.org/10.1101/2022.05.09.491139

- Collaboration with EMBL scientists



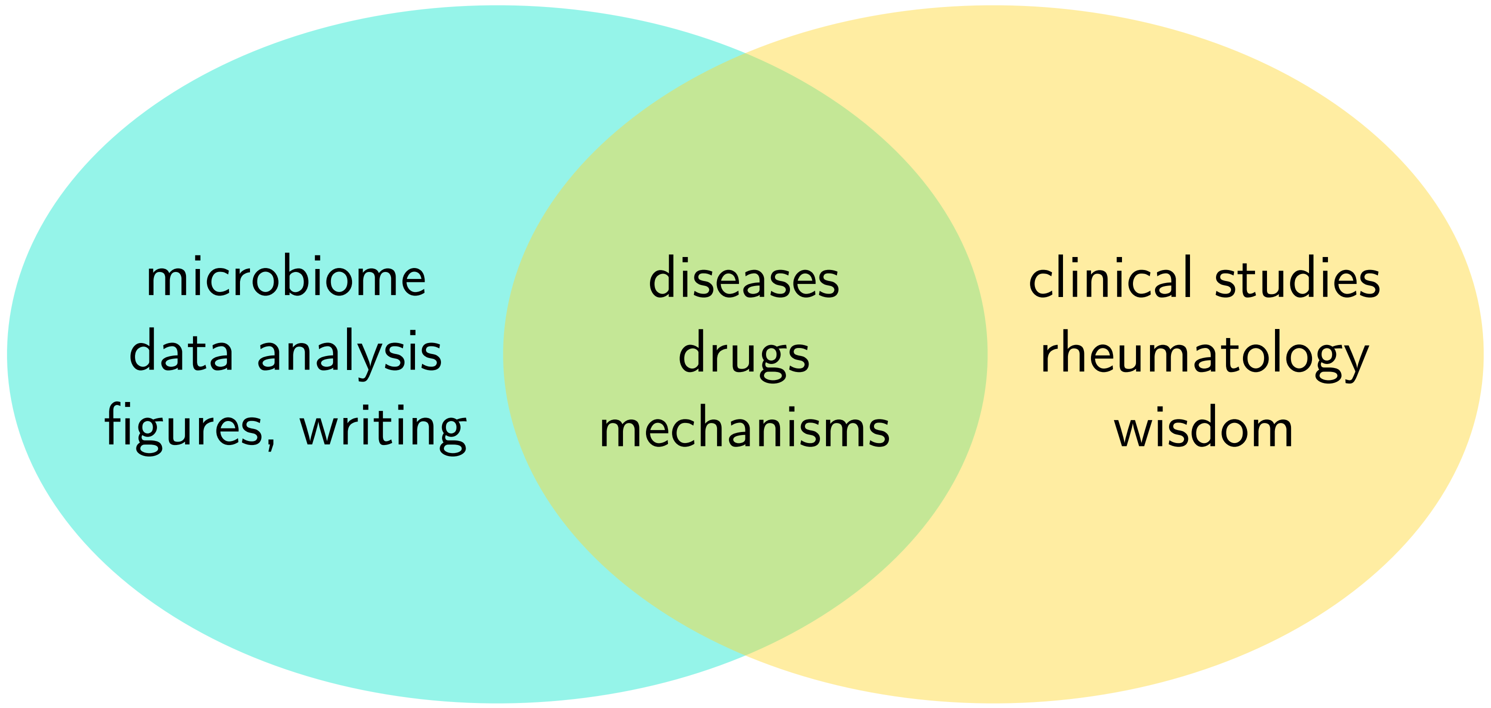
i.e. GESPIC

Chronic inflammatory disease

Spondyloarthritis, acute anterior uveitis, and Crohn's disease have both shared and distinct gut microbiota

Morgan Essex, Valeria Rios Rodriguez, Judith Rademacher, Fabian Proft, Ulrike Löber, Lajos Marko, Uwe Pleyer, Till Strowig, Jérémy Marchand, Jennifer A. Kirwan, Britta Siegmund, Sofia Kirke Forslund, Denis Poddubnyy
doi: https://doi.org/10.1101/2022.05.13.22275044

- Collaboration with Charité clinicians



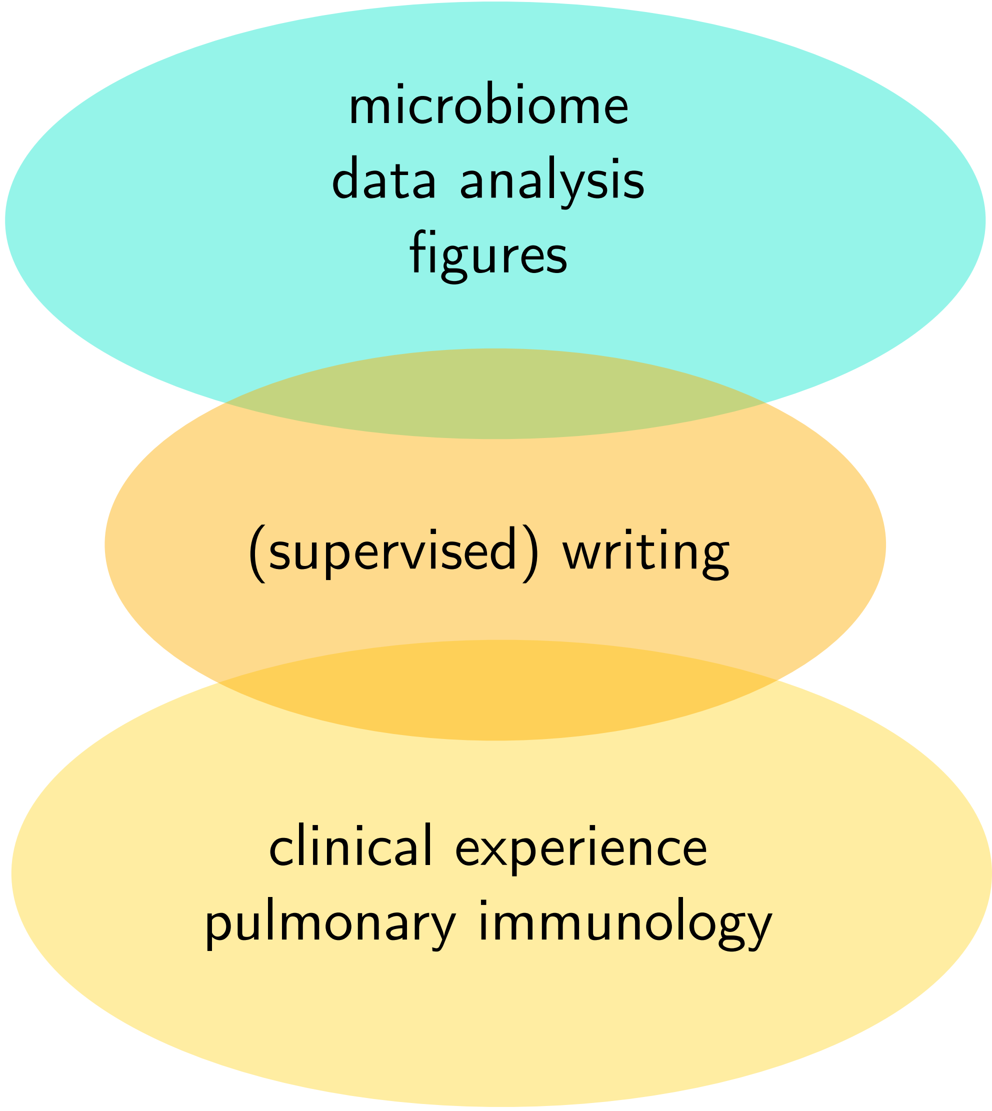
i.e. COVID

Acute inflammatory disease

Gut microbiota dysbiosis is associated with altered tryptophan metabolism and dysregulated inflammatory response in severe COVID-19

Morgan Essex, Belén Millet Pascual-Leone, Ulrike Löber, Mathias Kuhring, Bowen Zhang, Ulrike Bruening, Raphaela Fritsche-Guenther, Marta Krzanowski, Facundo Fiocca Vernengo, Sophia Brumhard, Ivo Röwekamp, Agata Anna Bielecka, Till Robin Lesker, Emanuel Wyler, Markus Landthaler, Andrej Mantei, Christian Meisel, Sandra Caesar, Charlotte Thiebeault, Victor Corman, Lajos Marko, Norbert Suttorp, Till Strowig, Florian Kurth, Leif E. Sander, Yang Li, Jennifer A. Kirwan, Sofia K. Forslund, Bastian Opitz
doi: https://doi.org/10.1101/2022.12.02.518860

- Collaboration with Charité clinicians



Essential tool suite for microbiome data science