

Large-scale microbiome profiling studies: opportunities and challenges

Presenter: Leo Lahti, University of Turku, Finland & VIB/KU Leuven, Belgium. Contact: <http://www.iki.fi/Leo.Lahti>
In collaboration with Prof Willem M de Vos (Helsinki/Wageningen) & Prof Jeroen Raes (VIB/KU Leuven)

Data availability & standardization

- Increasing sample sizes & statistical power
- Comprehensive population coverage & subject stratification
- Individuality in the larger context

HITChip Atlas:

- > 10,000 samples
- > 5,000 subjects
- > 1,000 phylotypes
- versatile geographical coverage
- some metabolomic profiles etc.
- comprehensive metadata
- intervention studies & long follow-ups (10+ years)
- reproducible & reliable down to 0.1% relative abundance; deep phylotype level profiling
- open data: Lahti et al. Nature Comm. 2014

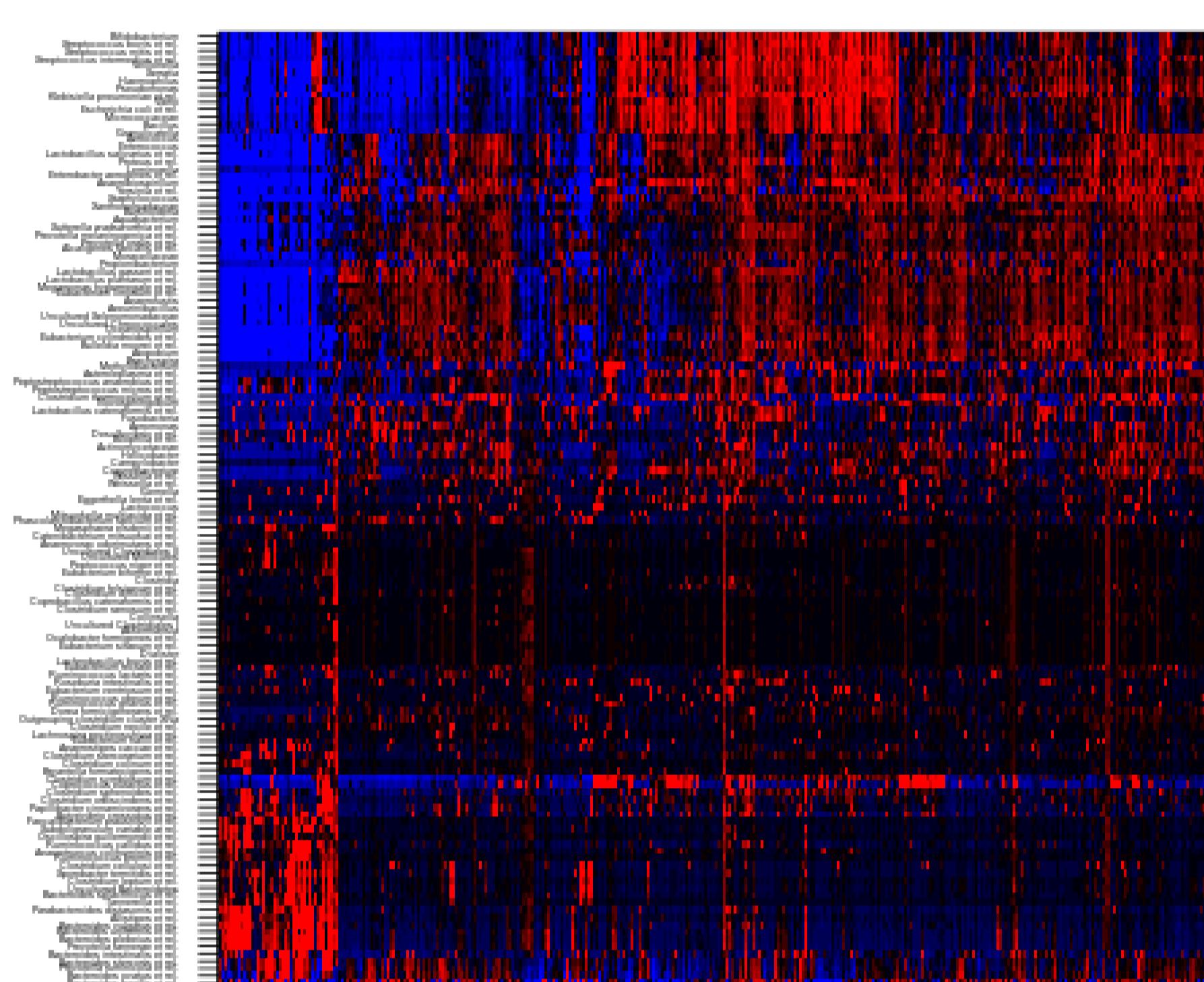


Flemish Gut Flora:

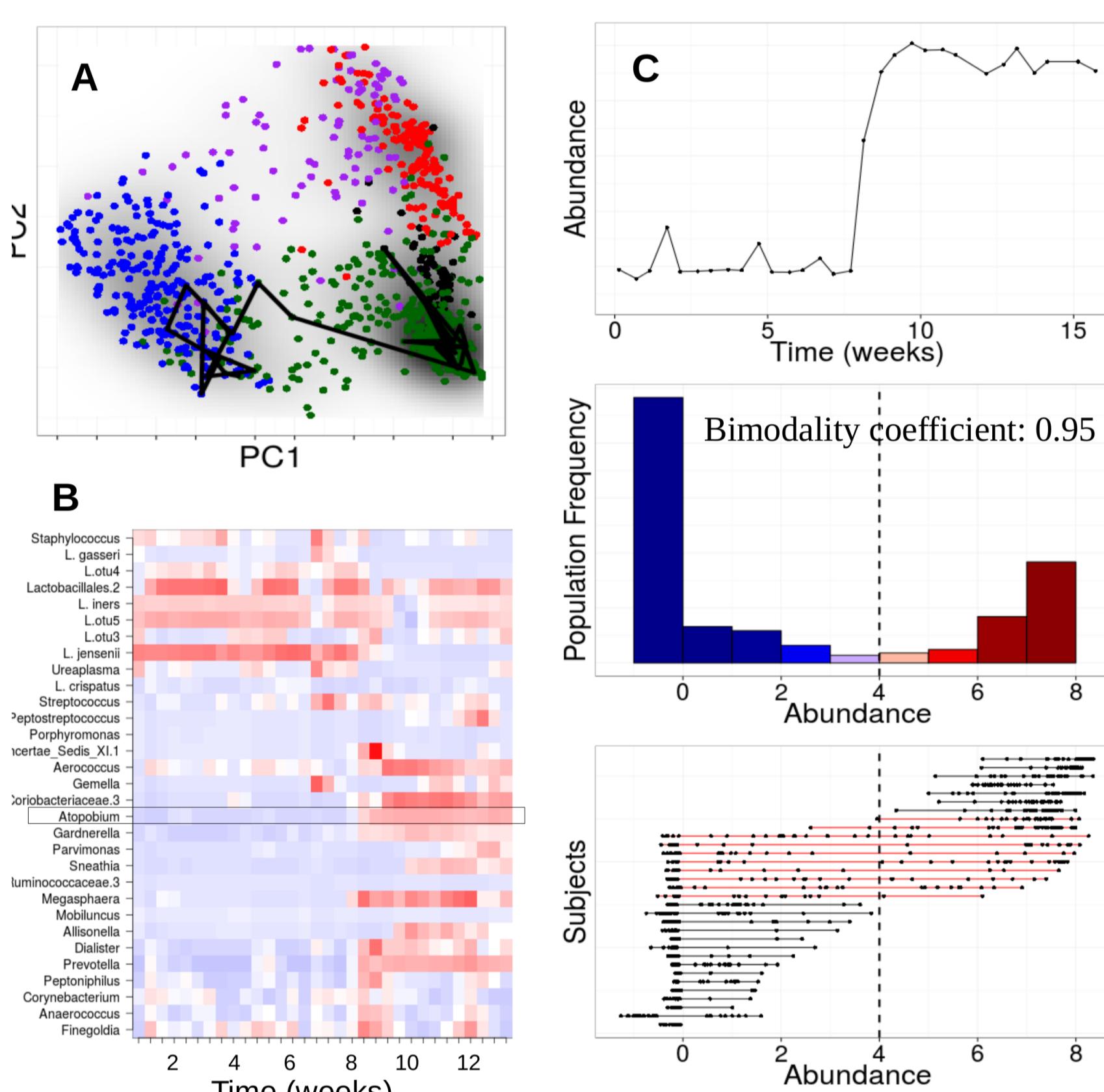
- > 5,000 samples
- > 1,000 subjects
- > 7,000 OTUs
- normal Flemish population
- metagenomic profiling (subset)
- clinical metadata & questionnaires
- longitudinal analyses
- standardized sample collection & processing
- open data: Falony et al. Science 2016



Cross-sectional vs. longitudinal analyses



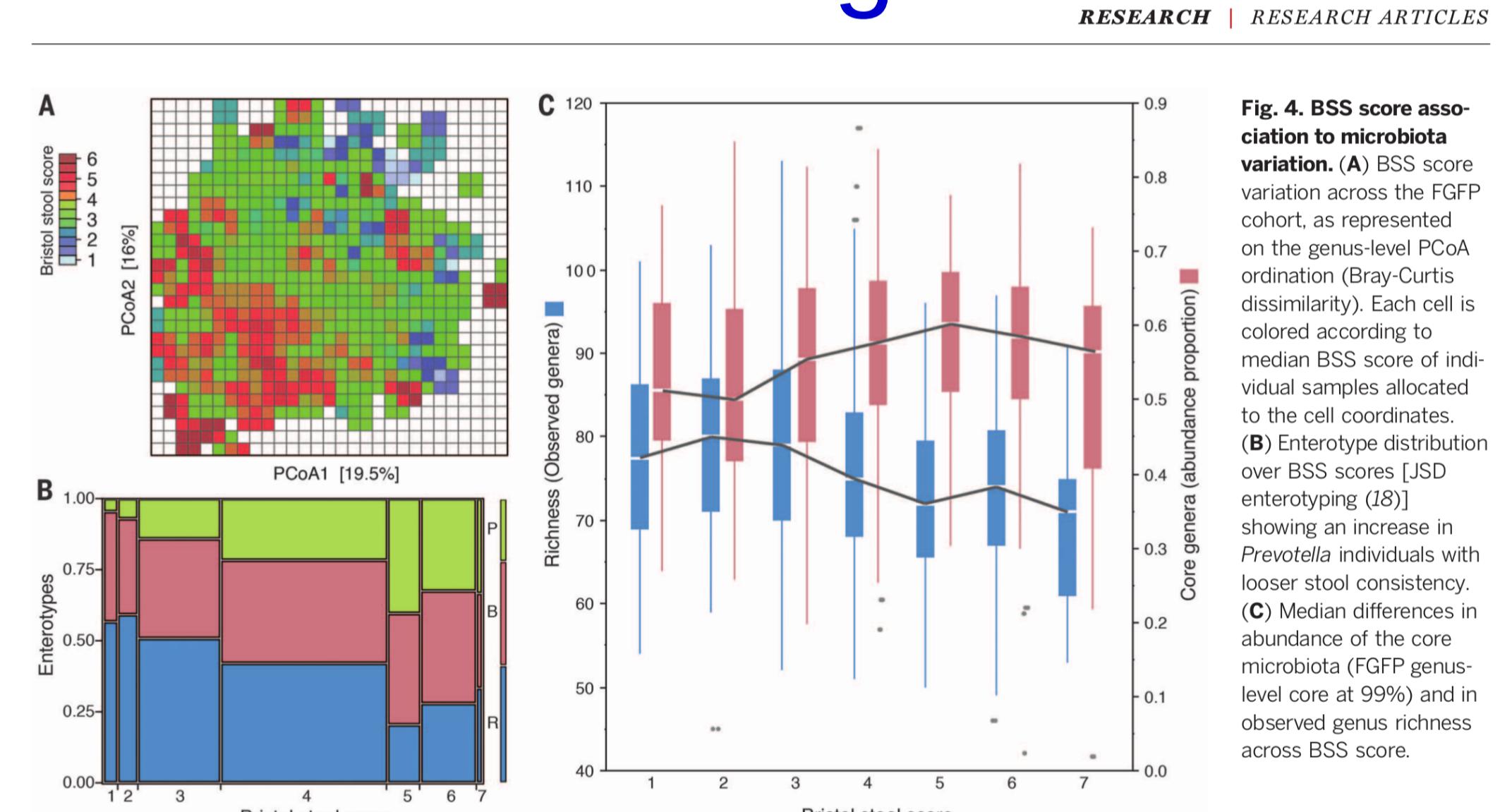
HITChip Atlas: 10,000+ samples;
Data: Lahti et al. 2014.



Vaginal microbiota community dynamics;
Data: Gajer et al. 2012 Figure: Faust et al. 2015

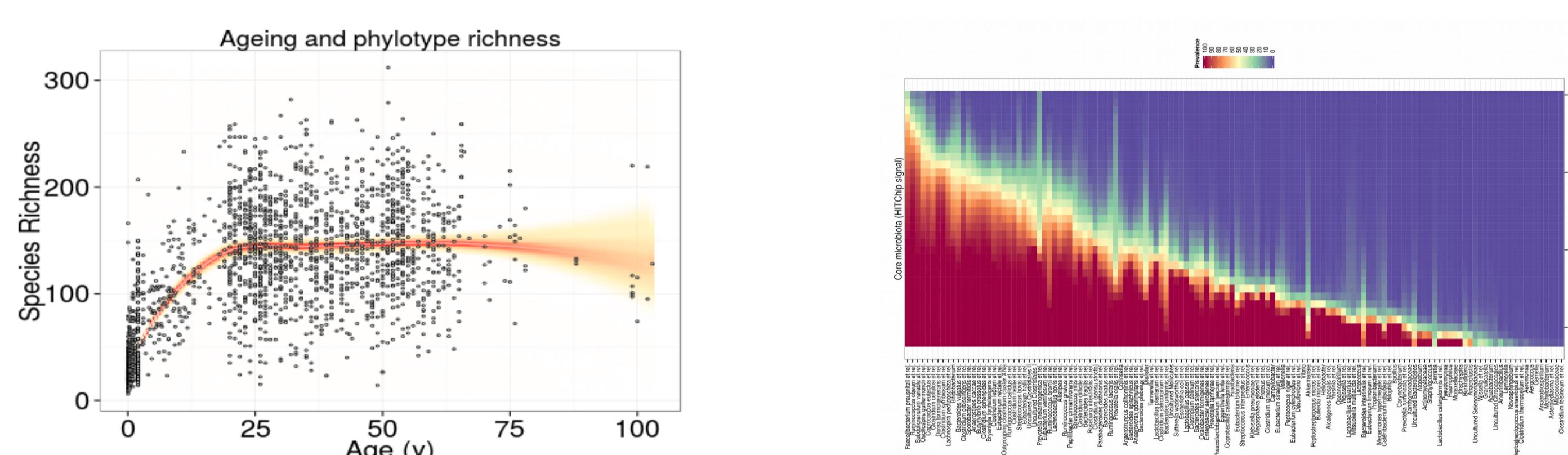
Custom preprocessing,
analysis & visualization tools
<http://microbiome.github.io>

Confounding factors



Stool consistency is a major confounder (Falony et al. 2016)

High individuality



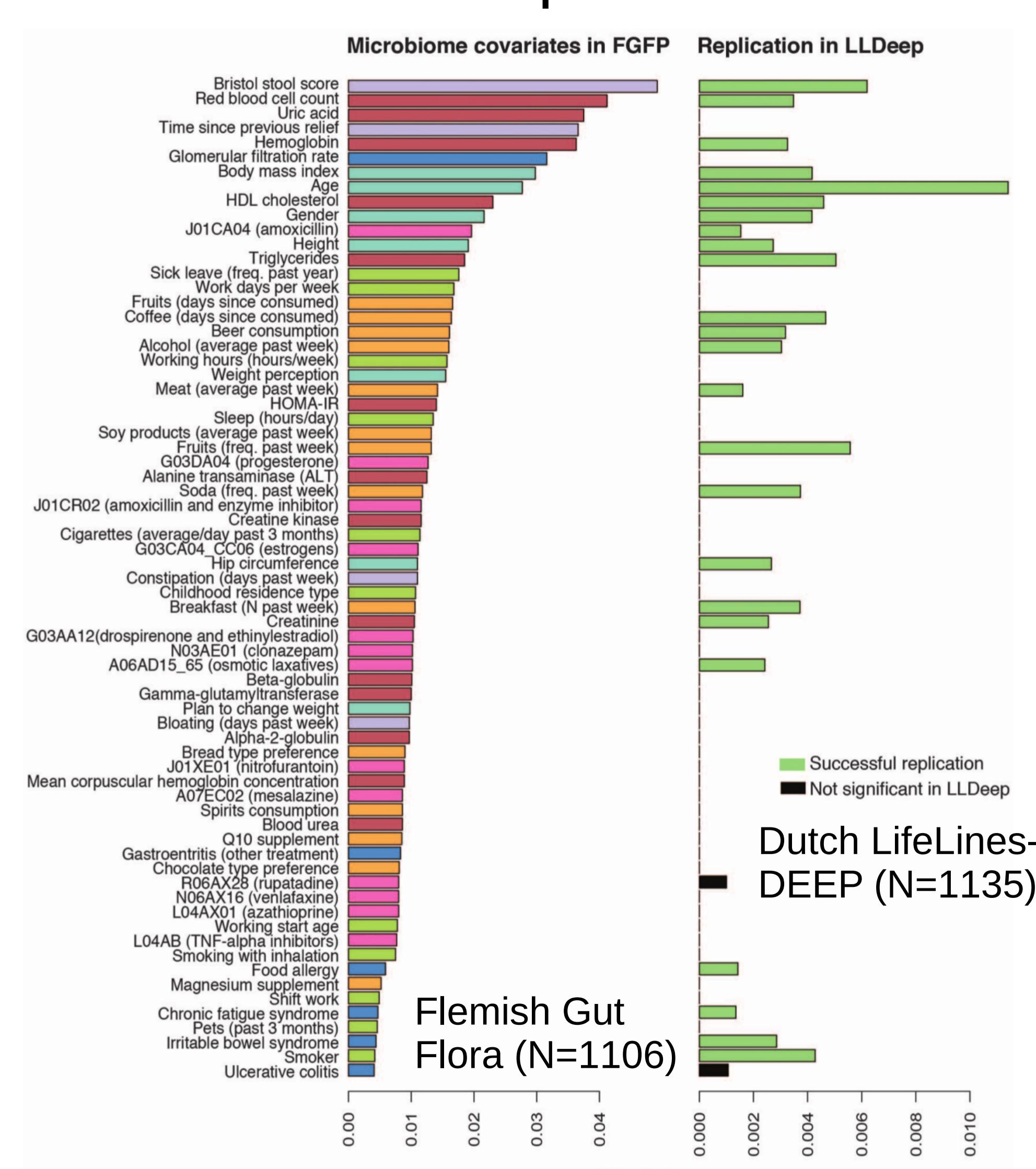
Core microbiota: only few species prevalent at a high abundance
(Data: HITChip Atlas, Lahti et al. 2014)

Complexity

Total explained variation 16.4%:

Proposed disease marker genera associated to host covariates and medication - inclusion in study design is essential! (Flemish Gut Flora; Falony et al. 2016.)

92% replication rate !



References

Population-level analysis of gut microbiome variation. Falony et al. *Science* 352(6285):560-4, 2016

Metagenomics meets time series analysis: unraveling microbial community dynamics. Faust, Lahti et al. *Current Opinion in Microbiology* 15:56-66 2015

Improved taxonomic assignment of human intestinal 16S rRNA sequences by a dedicated reference database Ritari et al. *BMC Genomics* 16:1056, 2015.

Fat, Fiber and Cancer Risk in African, Americans and Rural Africans. O'Keefe et al. *Nat. Comm.* 6:6342, 2015

Impact of diet and individual variation on intestinal microbiota composition and fermentation products in obese men. Salonen et al. *ISME Journal* 8:2218-30, 2014

Tipping elements in the human intestinal ecosystem. Lahti et al. *Nat. Comm.* 5:4344, 2014

Binning metagenomic contigs by coverage and composition (CONCOCT) Alneberg et al. *Nat. Methods* 11:1144-6, 2014.

