Clinical microbiome data science

PRESENTER INFO:

Tuomas Borman   ✉ tuomas.v.borman@utu.fi

Leo Lahti   ✉ leo.lahti@utu.fi

wellcome connecting science

Microbiome Interactions in Health and Disease (Virtual Conference)
13–15 October 2021

CONTRIBUTORS:

Felix G.M. Ernst, Sudarshan A. Shetty, Ruizhu Huang, Domenick J. Braccia, Héctor Corrada Bravo, The miaverse collective

SUOMEN AKATEMIA
FINLANDS AKADEMI · ACADEMY OF FINLAND

UNIVERSITY OF TURKU

MOTIVATION

• Multi-omics has become more common in microbiome research, because it can support the analysis of microbiome dysbiosis and facilitate the discovery of biomarkers.
• Multi-omics integrate data from multiple sources, for example, microbial abundance data with biomolecular profiling data from blood samples.
• Despite the need, there have not been well-suitable solutions for this kind of analysis that integrates multiple assays.
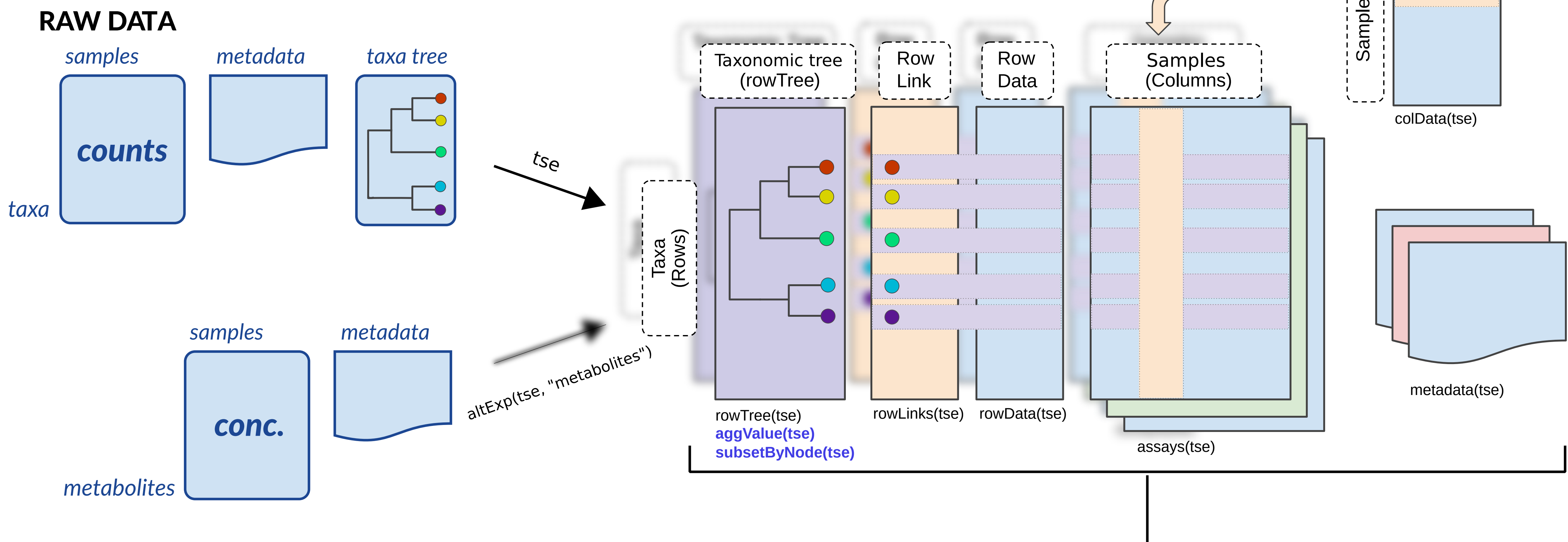
PROPOSAL

• Emerging analysis framework called *miaverse* (MIcrobiome Analysis uniVERSE) utilizes a common, standardized data container, which enables highly optimized integration of multi-assay microbiome profiling data from clinical studies.
• The *miaverse* contains methods for data analysis and visualization together with comprehensive online documentation.
• Common, shared data container enables usage of tools of other disciplines.
• The *miaverse* is a collaborative open-source project, and stable version is already available via Bioconductor.
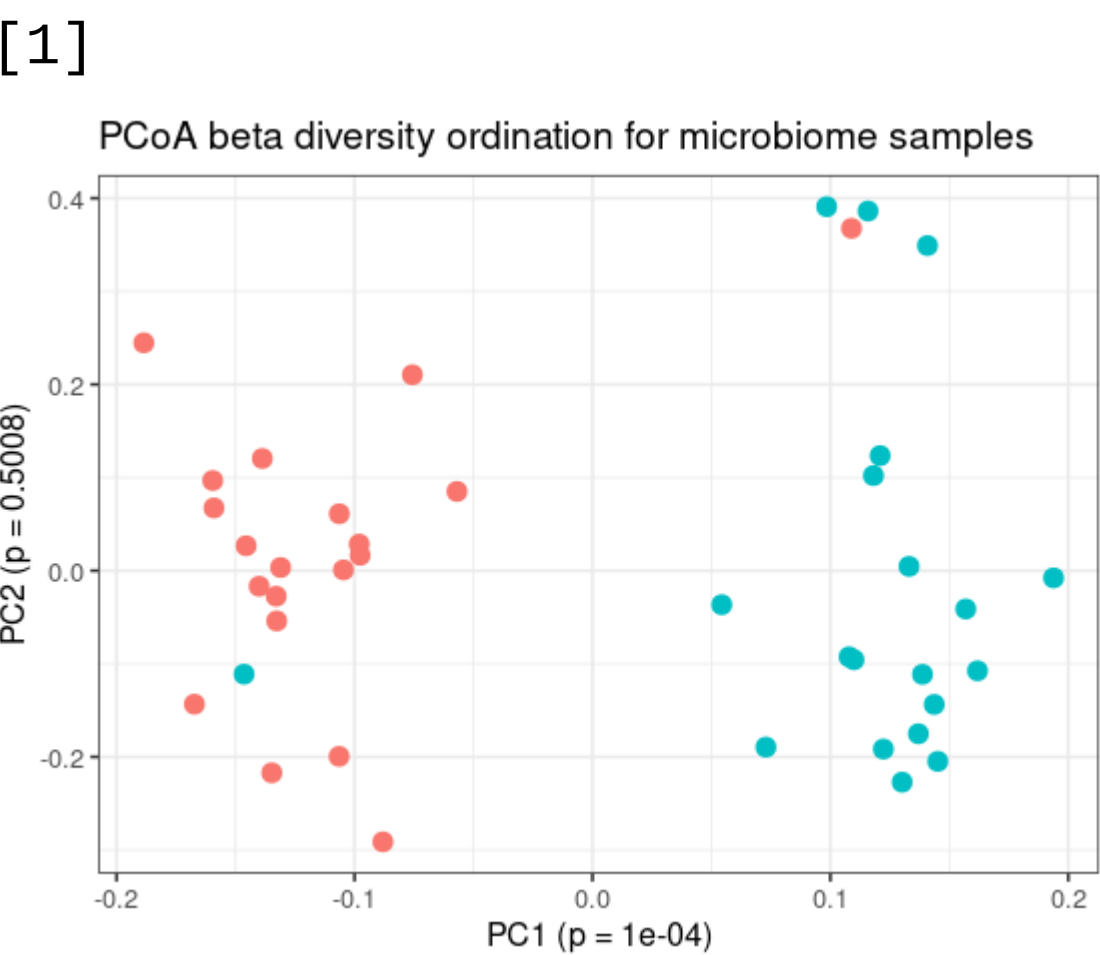
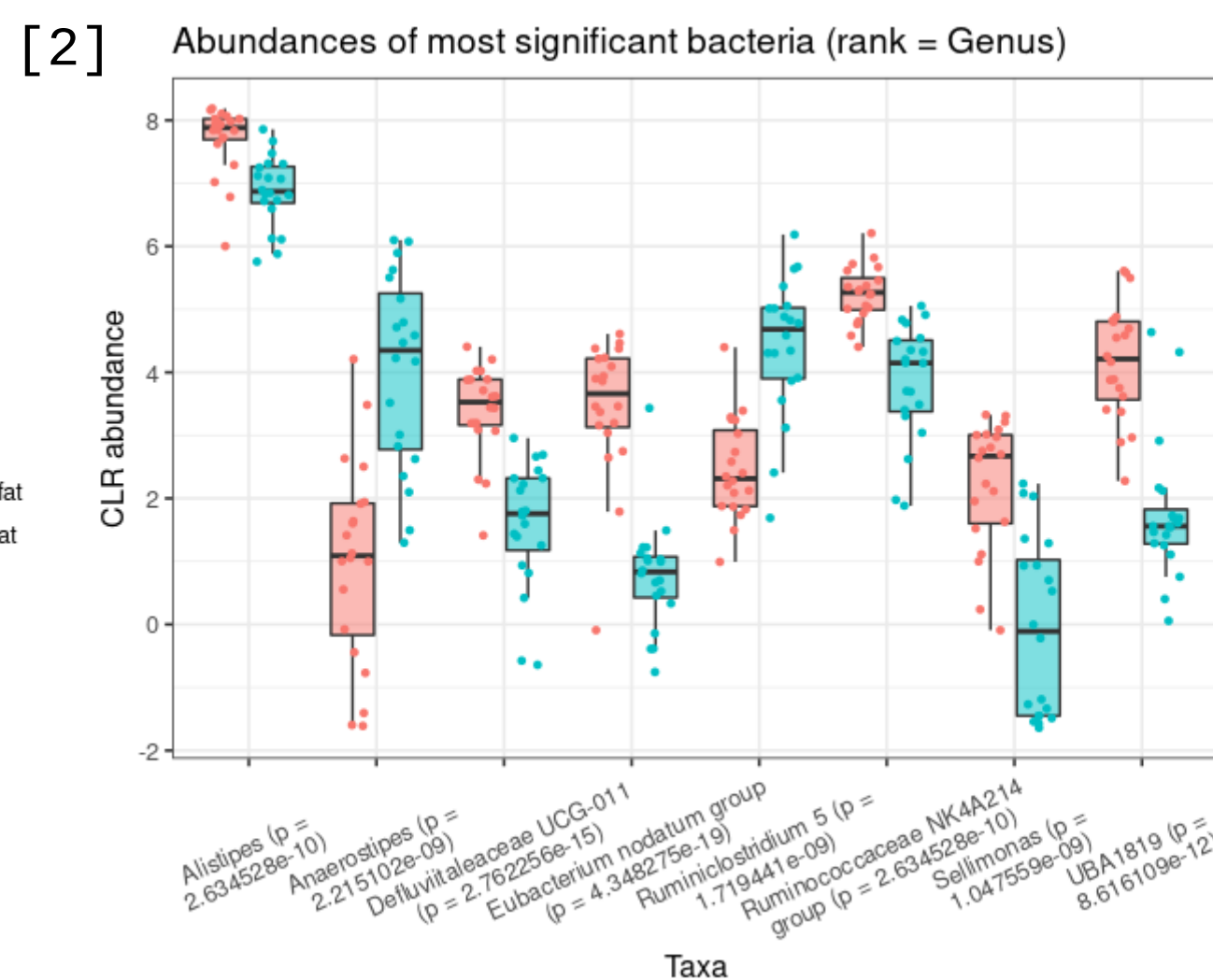## PROPOSED WORKFLOW: *TreeSE* & *miaverse*

Import Data

This workflow starts by importing bacterial abundance data and metabolite concentration data from stool samples into the *TreeSE* data container.

RAW DATA

samples · metadata · taxa tree

counts

taxa

tse

samples · metadata

conc.

metabolites

altExp(tse, "metabolites")

The *TreeSummarizedExperiment* object

The *TreeSE* object is uniquely positioned to support optimized multi-assay data access, manipulation, and visualization.

Taxonomic tree (rowTree) · Row Link · Row Data · Samples (Columns)

Taxa (Rows)

Samples

colData(tse)

rowTree(tse)
aggValue(tse)
subsetByNode(tse)

rowLinks(tse) · rowData(tse)

assays(tse)

metadata(tse)

The *miaverse* pipeline

[1]

PCoA beta diversity ordination for microbiome samples

PC1 (p = 1e-04)
PC2 (p = 0.5008)

Diet
● High-fat
● Low-fat

[2]

Abundances of most significant bacteria (rank = Genus)

CLR abundance

Taxa

Diet
High-fat
Low-fat

[3] Correlations between bacteria and metabolites (statistically significants marked with X)


```
[1]
mia::transformSamples(tse)
vegan::vegdist(x)
ecodist::pco(x)
ggplot2::ggplot(x)

[2]
mia::agglomerateByRank(tse)
mia::meltAssay(tse)
ggplot2::ggplot(x)

[3]
mia::subsetByPrevalentTaxa(tse)
microbiome::associate(bacteria, metabolites)
pheatmap::pheatmap(x)
```