

**Title**

Clinical microbiome data science

**Authors**

Tuomas Borman, Leo Lahti

**Keywords**

microbiome, data integration, metagenomics, multi-omics, data science

**Abstract**

Because of the complex and high dimensional nature of microbiome profiling data, machine learning and other computational approaches have become an instrumental part of the researcher's toolkit in this field. There is an increasing need to develop robust and reproducible methods to integrate and analyse taxonomic, functional, and clinical data across multiple sources, such as microbial abundances in the gut with biomolecular profiling data from blood samples.

This kind of integrative multi-omic approaches can support the analysis of microbiome dysbiosis and facilitate the discovery of novel biomarkers for health and disease. The currently available solutions for multi-assay microbiome data have severe limitations in terms of scalability and the ability to incorporate different types of complementary data sources in a single reproducible workflow.

Emerging analysis ecosystem 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* is a collaborative open-source project.

We have developed open data science methods for data analysis and visualization, with a particularly designed support for multi-assay data integration and analysis. Stable versions are available via the peer-reviewed Bioconductor network for open research software, together with comprehensive online documentation.

The emerging framework fulfills the need for open and reproducible analysis of multi-assay data in microbiome studies. We anticipate that the framework has the potential to be widely adopted by microbiome researchers. This is further facilitated by the tight links to related application domains, such as single cell sequencing, where closely related open-source techniques are now being developed.