

# MGnifyR: bridging the gap between microbiome data resource and methods

Tuomas Borman, Ben Allen, Leo Lahti  
✉ [tuomas.v.borman@utu.fi](mailto:tuomas.v.borman@utu.fi) [in](#) [G](#) TuomasBorman



## MGnify: microbiome data resource

- **Generated from publicly available sequences in the European Nucleotide Archive (ENA)**
  - Standardized pipelines
- **Data types (number of analyses in 8 Sep 2024)**
  - Amplicon sequencing, ribosomal RNA (495,889)
  - Amplicon sequencing, non-ribosomal RNA, "barcoding" (2,050)
  - Whole genome shotgun sequencing (40,094)
  - Whole genome shotgun sequencing, assembled (60,113)
  - Whole genome shotgun sequencing, long-read (2)
  - Whole transcriptome sequencing (2,581)
- **Experiments**
  - Taxonomy annotations
  - Functional annotations
  - Genome catalogues

Explore the MGnify database



<https://www.ebi.ac.uk/metagenomics>

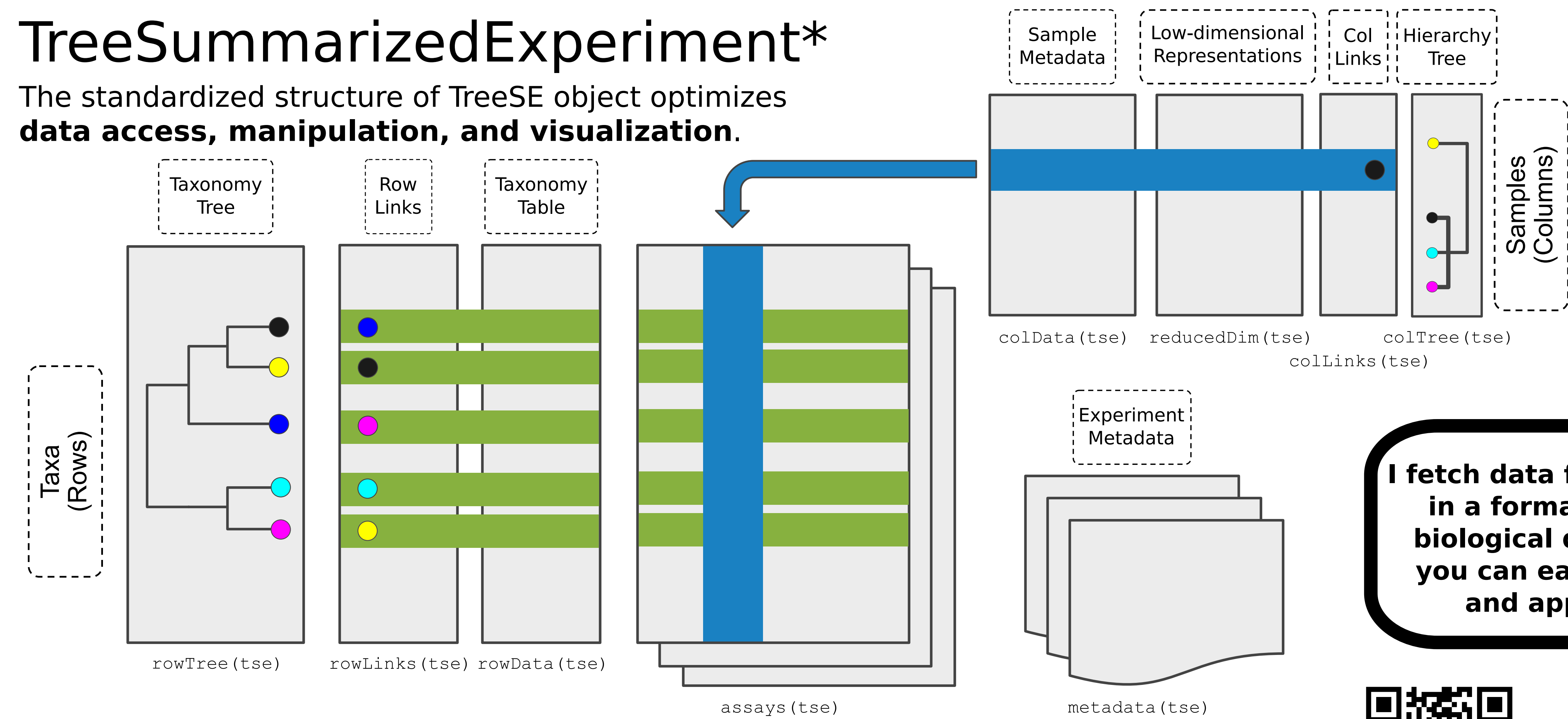
# MGnify

## MGnifyR

- **R/Bioconductor package for accessing MGnify data**
- **You can access raw data or abundance tables**
  - Abundance tables are fetched as TreeSummarizedExperiment objects, ready for direct use with Bioconductor methods

### TreeSummarizedExperiment\*

The standardized structure of TreeSE object optimizes **data access, manipulation, and visualization.**



I fetch data from the database in a format optimized for biological data analysis, so you can easily work with it and apply methods!



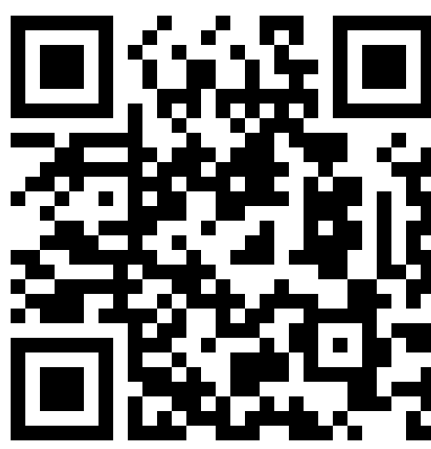
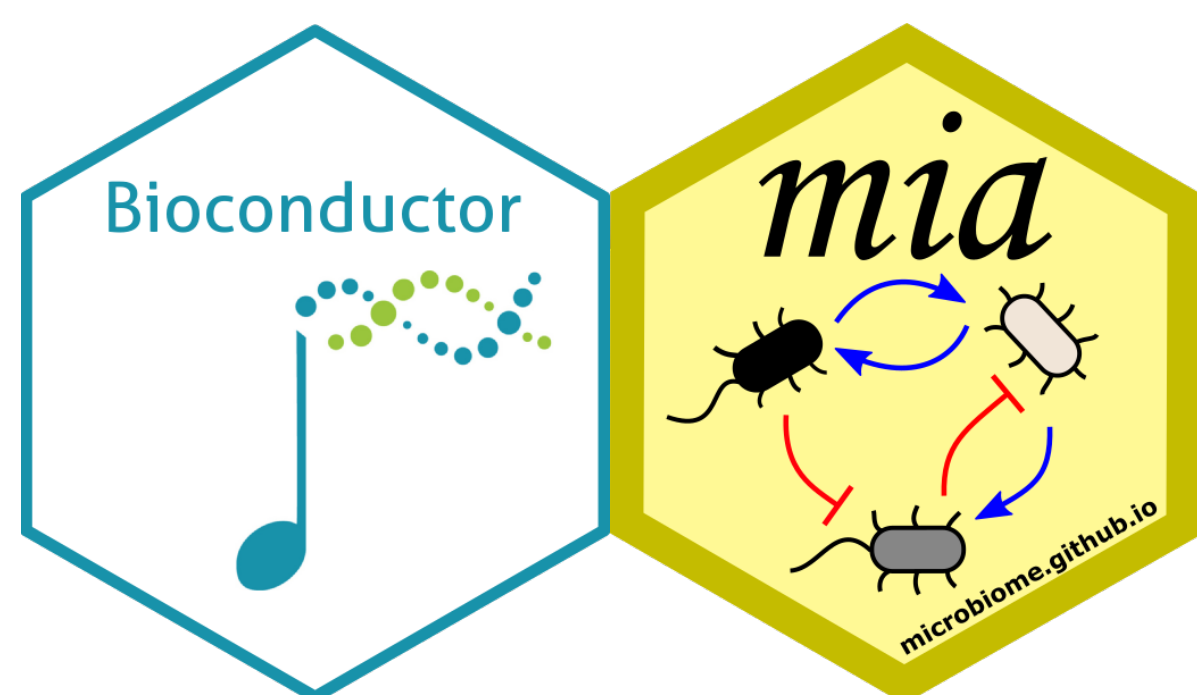
See MGnifyR tutorial

<https://ebi-metagenomics.github.io/MGnifyR>

## Bioconductor

- **World's largest bioinformatics project**
  - >2,300 R packages
  - Open-source
  - Peer-reviewed and tested
- **Cutting-edge methods for computational biology**
  - **mia (Microbiome Analysis)** offers tools for microbiome data science

Check Orchestrating Microbiome Analysis (OMA) online book for more information on microbiome data science in Bioconductor!



<https://microbiome.github.io/OMA>

