HoloFoodR: bridging the gap between data retrieval and data analysis in multi-omics

Artur Sannikov[2], Tuomas Borman[1], Kati Hanhineva[2], Leo Lahti[1]

[1]Department of Computing, University of Turku, [2]Department of Life Technologies, University of Turku
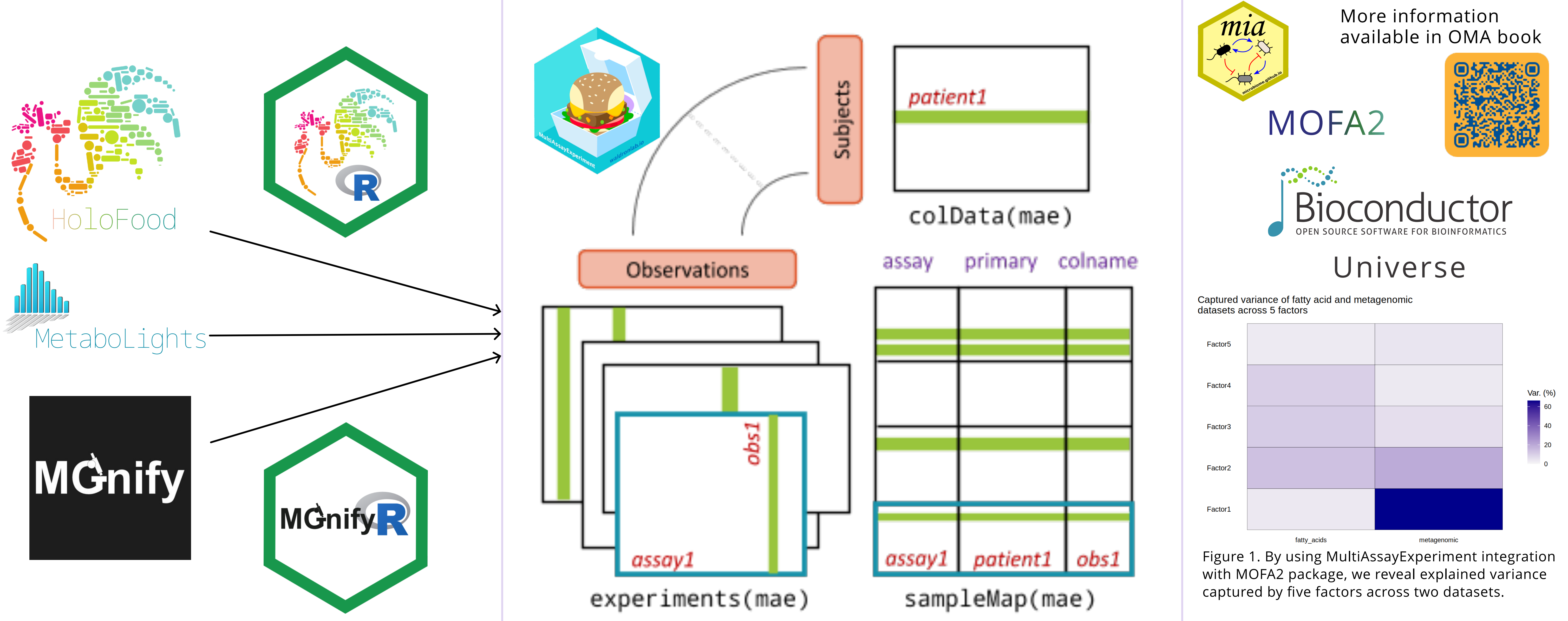
[2]arsann@utu.fi 0000-0001-7765-123X

Introduction

Holo-omic is a novel research field[1] that integrates multi-omic data from the host organism and its microbiome. Created by the HoloFood consortium, the openly available HoloFood database[2] grants access to the holo-omic data for over 2000 chickens and salmons. This database aims to study the impact of food additives on physiological and biomolecular parameters of farmed animals. However, access to such a wealthy database is complex and requires a profound understanding of the interrelations between various omic datasets, creating a gap between data availability and analytical frameworks.

Check out HoloFoodR on Bioconductor

Framework

More information available in OMA book

MOFA2

Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS

Universe

Figure 1. By using MultiAssayExperiment integration with MOFA2 package, we reveal explained variance captured by five factors across two datasets.

Access and Retrieve
complex multi-omic data on hosts and microbiome from HoloFood, the largest open holo-omicdatabase

Seamlessly integrate HoloFood data with metagenomic data from MGnify and metabolomic data from MetaboLights

Manage
data in MultiAssayExperiment, a Bioconductor container that unifies diverse biological datasets to accelerate discovery in multi-omics

MultiAssayExperiment simplifies access to more than 2300 R/Bioconductor packages by reducing data handling complexity and facilitating robust, validated analysis

Discover
new knowledge about the impact of food additives on physiological processes in chickens and salmons to improve

Via the integration with MOFA2 we compute the captured variance of five factors across two datasets

Results and conclusions

We have developed HoloFoodR, an R/Bioconductor package to bridge the gap between data, computing environments, and data analysis frameworks. Our package leverages the capabilities of multi-omic data containers that provide systematic means to store omics and associated metadata. We demonstrate our package with a case study in which we retrieve salmon data from HoloFood, integrate it with MGnify and perform common omic analysis steps, such as normalization, visualization of diversity, and machine learning. In particular, we opted to leverage the integration of MOFA2[3] package with multi-omic data containers to compute captured variance across metagenomic and fatty acid datasets. In summary, HoloFoodR facilitates scientific discovery and enhances the efficiency of holo-omic data analysis, supporting the identification of novel biological relations.

References

1. Limborg, Morten T. et al. Applied Hologenomics: Feasibility and Potential in Aquaculture. Trends in Biotechnology 2018;**36**:252–264.
2. https://www.holofooddata.org
3. Argelaguet, Ricard et al. MOFA+: A Statistical Framework for Comprehensive Integration of Multi-Modal Single-Cell Data. Genome Biology 2020;**21**:1–17.

Acknowledgements

FindingPheno

J&AE
JANE AND AATOS
ERKKO FOUNDATION

Research Council of Finland

UNIVERSITY OF TURKU