

# Clinical microbiome data science with *MultiAssayExperiment*

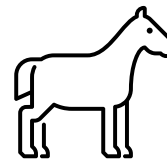
Tuomas Borman, Paulina Salminen, Leo Lahti

EuroBioC2022 in Heidelberg, Germany; 14<sup>th</sup> Sep 2022



UNIVERSITY  
OF TURKU



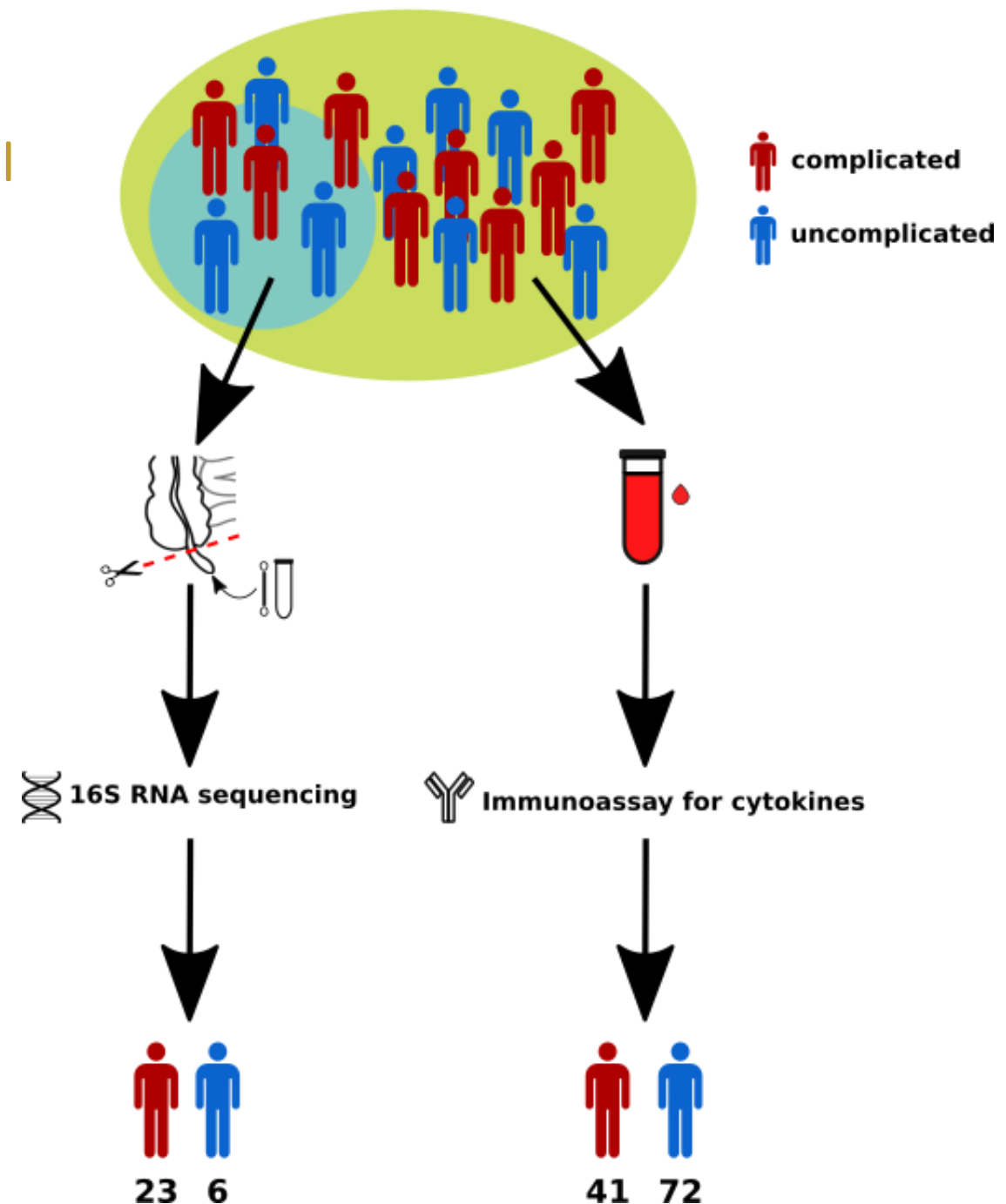


# Contents

1. Appendicitis case study (MAPPPAC)
- 2. *miaverse* (R/Bioc framework)**
3. Results of appendicitis study

## Microbiology APPendicitis ACuta (MAPPAC) trial

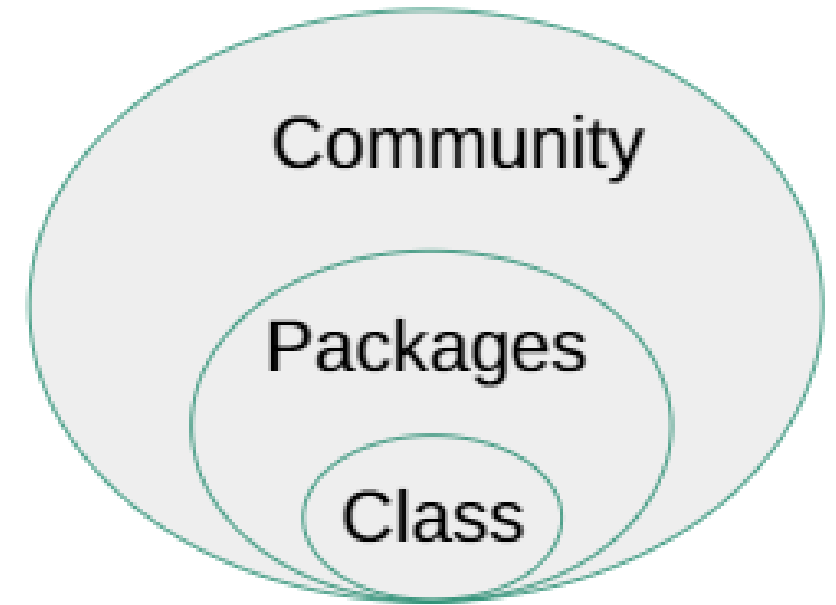
- Appendicitis
  - Inflammation of appendix
- Two clinically different types
  - **Complicated** → **appendectomy**
  - **Uncomplicated** → **antibiotics**
- **Do microbes of appendix explain the type of appendicitis?**



WHAT CAN  
WE DO WITH  
THIS DATA?!?

# Class

- In central role
- Standardized, optimized...
- The type of data sets demands on requirements
  - Clinical data, economics data...



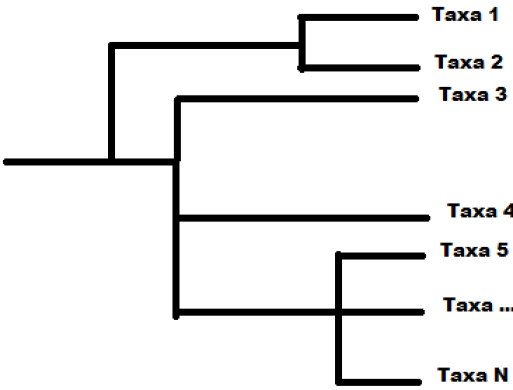
	A	B	C	D	E	F	G	H	I	J	K
1		Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6	Patient 7	Patient 8	Patient 9	Patient 10
2	Taxa 1	10	28	32	73	96	87	6	11	81	1
3	Taxa 2	41	67	42	52	40	92	59	92	98	35
4	Taxa 3	28	64	57	82	57	50	81	76	44	72
5	Taxa 4	53	76	1	42	91	61	33	99	61	4
6	Taxa 5	29	44	85	51	5	80	44	56	88	27
7	Taxa 6	48	72	84	75	48	78	57	43	96	20
8	Taxa 7	89	100	54	0	54	37	51	0	29	52
9	Taxa 8	63	87	8	0	35	48	68	0	11	33
10	Taxa 9	80	28	21	0	85	7	86	71	0	41
11	Taxa 10	53	89	87	0	19	59	89			
12	Taxa 11	53	83	41	0	24	52	26			
13	Taxa 12	84	63	97	49	1	96	83			
14	Taxa 13	90	35	66	68	5	5	56			
15	Taxa 14	41	53	75	66	33	11	59			
16	Taxa 15	8	19	9	29	49	96	70			
17	Taxa 16	64	57	41	24	27	92	43			
18	Taxa 17	59	88	72	94	16	17	25			
19	Taxa 18	8	93	8	14	30	43	79			

	A	B	C	D
1		Patient 1	Patient 2	Patient 3
2	Cytokine 1	95	71	66
3	Cytokine 2	97	31	19
4	Cytokine 3	64	71	70
5	Cytokine 4	78	9	75
6	Cytokine 5	96	64	51
7	Cytokine 6	6	96	28
8	Cytokine 7	81	77	20
9	Cytokine 8	61	61	19
10	Cytokine 9	78	94	99
11	Cytokine 10	3	29	7

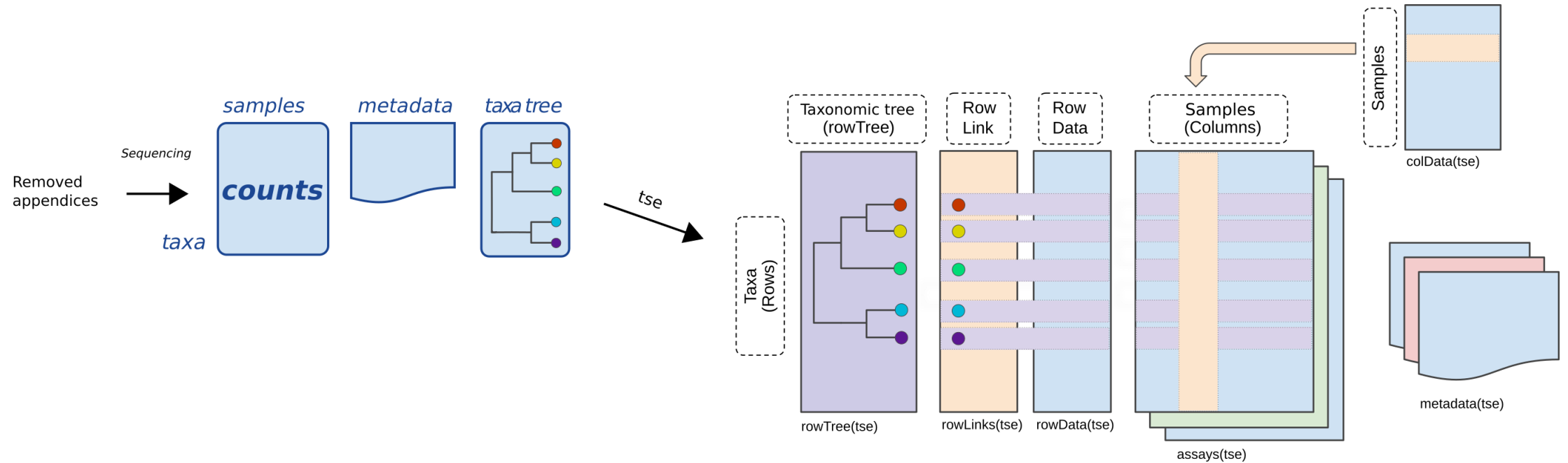
	A	B	C	D	E	F	G	H
1		Kingdom	Phylum	Class	Order	Family	Genus	Species
2	Taxa 1	Bacteria	Something	Something	Something	Something	Something	Something
3	Taxa 2	Bacteria	Something	Something	Something	Something	Something	Something
4	Taxa 3	Archaea	Something	Something	Something	Something	Something	Something
5	Taxa 4	Bacteria	Something	Something	Something	Something	Something	Something
6	Taxa 5	Bacteria	Something	Something	Something	Something	Something	Something
7	Taxa 6	Bacteria	Something	Something	Something	Something	Something	Something
8	Taxa 7	Bacteria	Something	Something	Something	Something	Something	Something
9	Taxa 8	Bacteria	Something	Something	Something	Something	Something	Something
10	Taxa 9	Bacteria	Something	Something	Something	Something	Something	Something
11	Taxa 10	Bacteria	Something	Something	Something	Something	Something	Something
12	Taxa 11	Bacteria	Something	Something	Something	Something	Something	Something
13	Taxa 12	Bacteria	Something	Something	Something	Something	Something	Something
14	Taxa 13	Bacteria	Something	Something	Something	Something	Something	Something

	A	B	C	D
1		Diagnosis	Sex	Age
2	Patient 1	Sick	M	20
3	Patient 2	Sick	M	44
4	Patient 3	Healthy	F	36
5	Patient 4	Sick	F	54
6	Patient 5	Healthy	F	34
7	Patient 6	Healthy	M	34
8	Patient 7	Sick	M	45
9	Patient 8	Sick	F	65
10	Patient 9	Healthy	M	34
11	Patient 10	Healthy	F	77

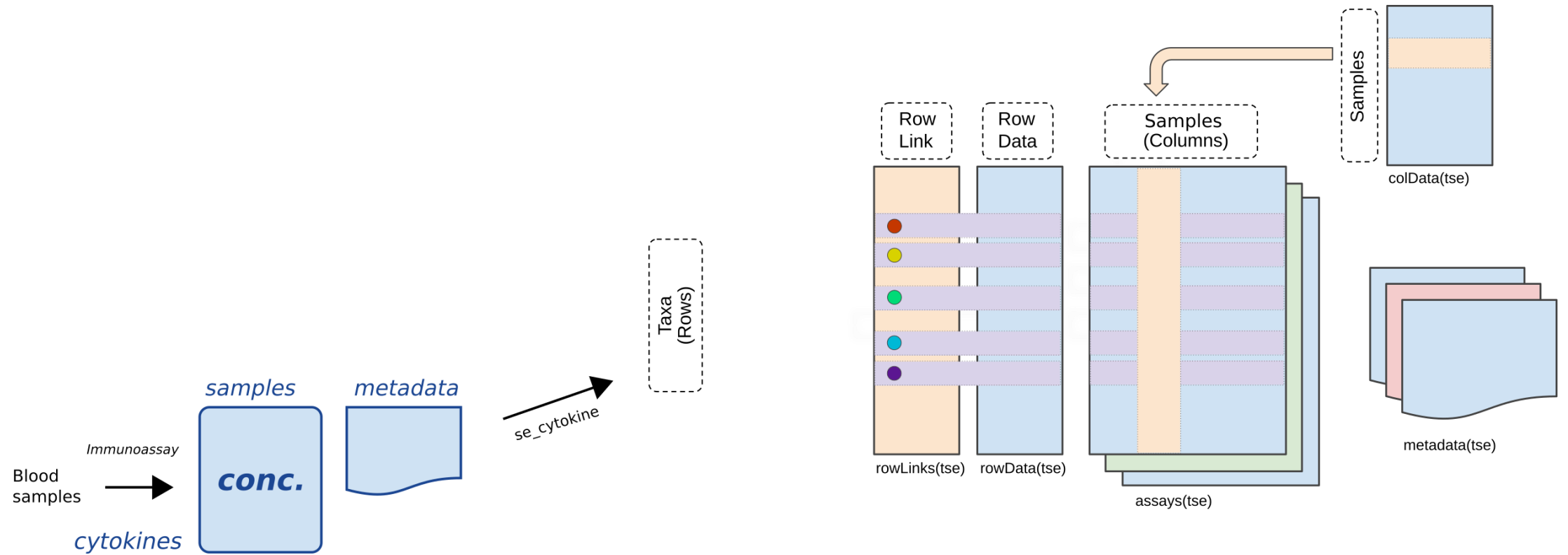
	K
9	Patient 10
2	78
65	16
27	2
78	85
94	37
55	79
24	41
55	40
0	10
2	51



# *(Tree)SummarizedExperiment (TreeSE)*

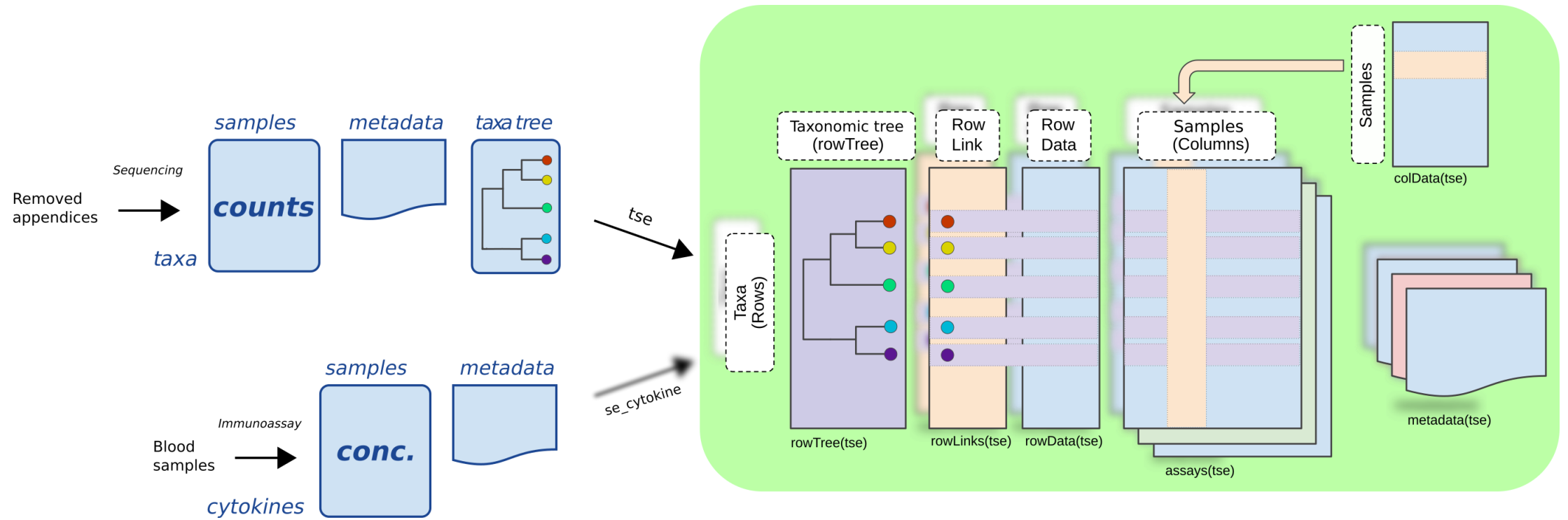


# *(Tree)SummarizedExperiment (TreeSE)*



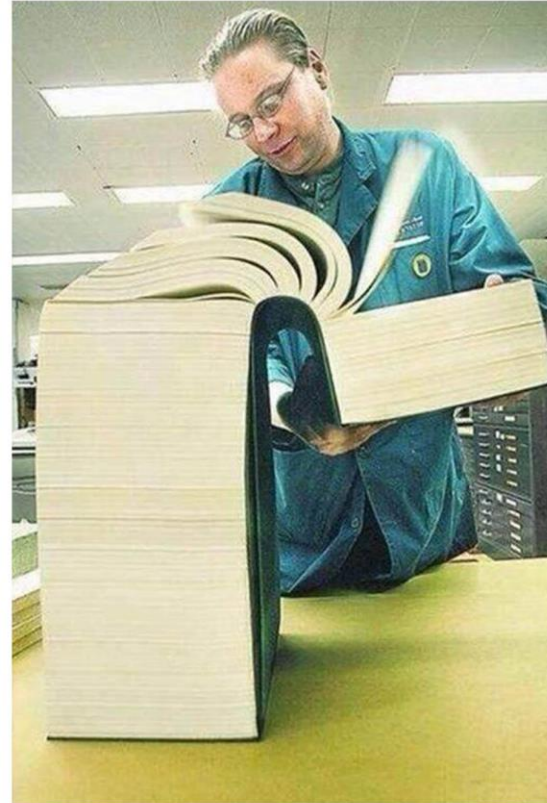


# MultiAssayExperiment (MAE)



# miaverse

**Associations  
between microbes  
and diseases**



**How microbes cause  
diseases?**



# miaverse

- R/Bioc framework for microbiome analytics
- Is based on *TreeSummarizedExperiment* class
  - Integrated to *SummarizedExperiment* ecosystem

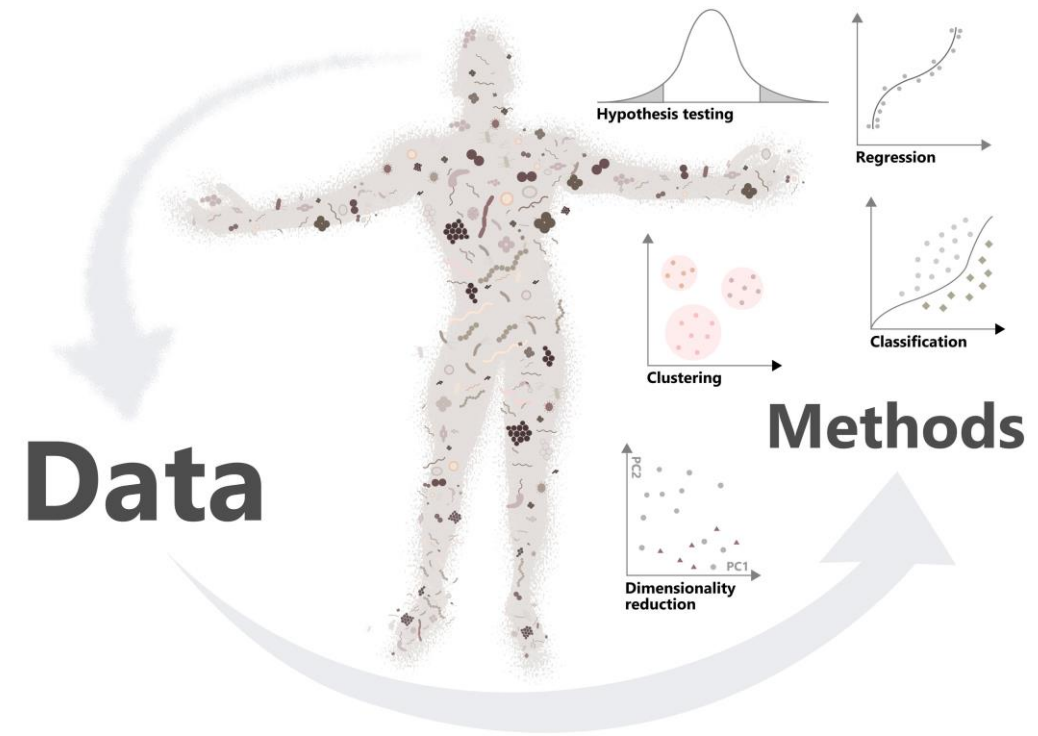
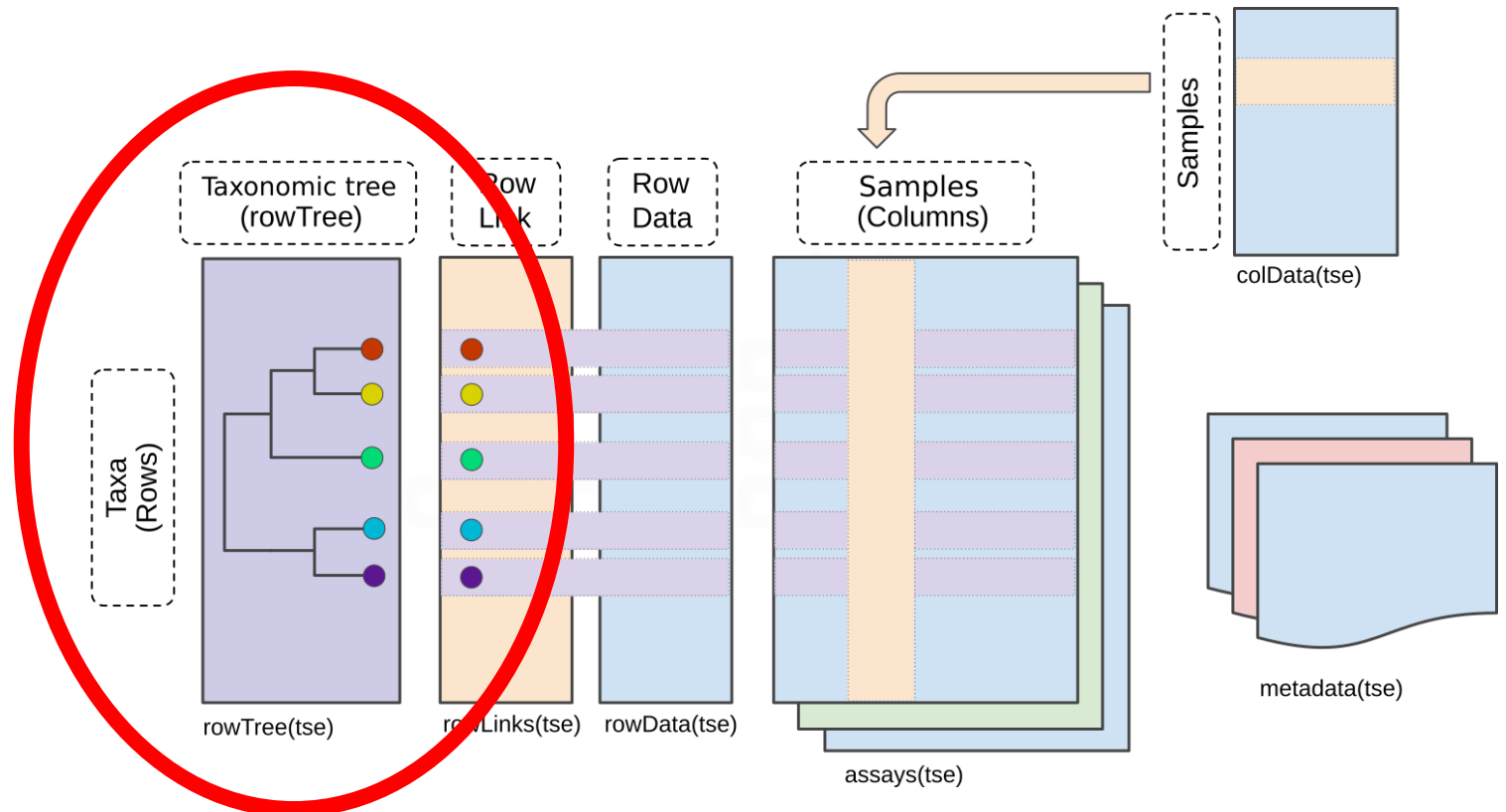


Figure source: Moreno-Indias et al. (2021) Statistical and Machine Learning Techniques in Human Microbiome Studies: Contemporary Challenges and Solutions. *Frontiers in Microbiology* 12:11.

# miaverse

*phyloseq* is  
a subset of  
*TreeSE*

- ***TreeSE* extends *SingleCellExperiment*** class by adding slots for row and column trees
  - Hierarchical data



# *miaverse*

- Methods **extend** available *SummarizedExperiment* toolkit

- Transformations
  - Log, Z, CLR...
- Diversity
  - Alpha and beta
- Subsetting based on
  - Prevalent or rare features
- Multi-assay analyses
  - Cross-correlation...
- And more...

```
mia::transformSamples()
```

```
miaViz::plotSeries()
```

```
miaViz::estimateDiversity()
```

```
miaViz::plotPrevalence()
```

```
mia::runRDA()
```

```
mia::testExperimentCrossAssociation()
```

# miaverse

- Consists of multiple packages
  - *mia* (analysis)
  - *miaViz* (visualization)
  - *miaTime* (time series analysis)...

- And tutorial book called **Orchestrating Microbiome Analysis (OMA)**

- Similar to Orchestrating Single-Cell Analysis (OSCA)

Bioconductor  
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS

Home » Bioconductor 3.15 » Software Packages » mia

mia

platforms all rank 294 / 2140 support 0 / 0 in Bioc 1 year  
build ok updated

DOI: [10.18129/B9.bioc.mia](https://doi.org/10.18129/B9.bioc.mia)

Microbiome ana

Bioconductor version: miaViz

mia implements tools  
SingleCellExperiment  
context of taxonomic  
as community indices

platforms all rank 536 / 2140 support 0 / 0 in Bioc 1 year  
build ok updated before release dependencies 136

DOI: [10.18129/B9.bioc.miaViz](https://doi.org/10.18129/B9.bioc.miaViz)

Orchestrating Microbiome Analysis

Welcome

I Introduction

1 Introduction

2 Microbiome Data

2.1 Data science framework

2.2 Data containers

2.3 Loading experimental microbi...

2.4 Demonstration data

Session Info

3 Packages

3.1 Package installation

3.2 Some available packages

Session Info

II Focus Topics

4 Data Manipulation

4.1 Tidying and subsetting

4.2 Merge data

## Orchestrating Microbiome Analysis

Authors: Leo Lahti [aut], Sudarshan Shetty [aut], Tuomas Borman [aut, cre], Felix GM Ernst [aut]

Version: 0.98.10

Modified: 2022-06-19

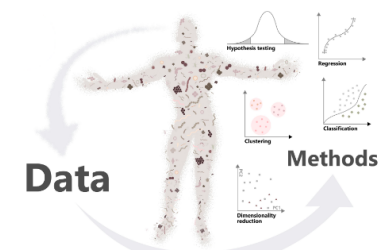
Compiled: 2022-09-06

Environment: R version 4.2.1 (2022-06-23), Bioconductor 3.15

License: CC BY-NC-SA 3.0 US

Copyright:

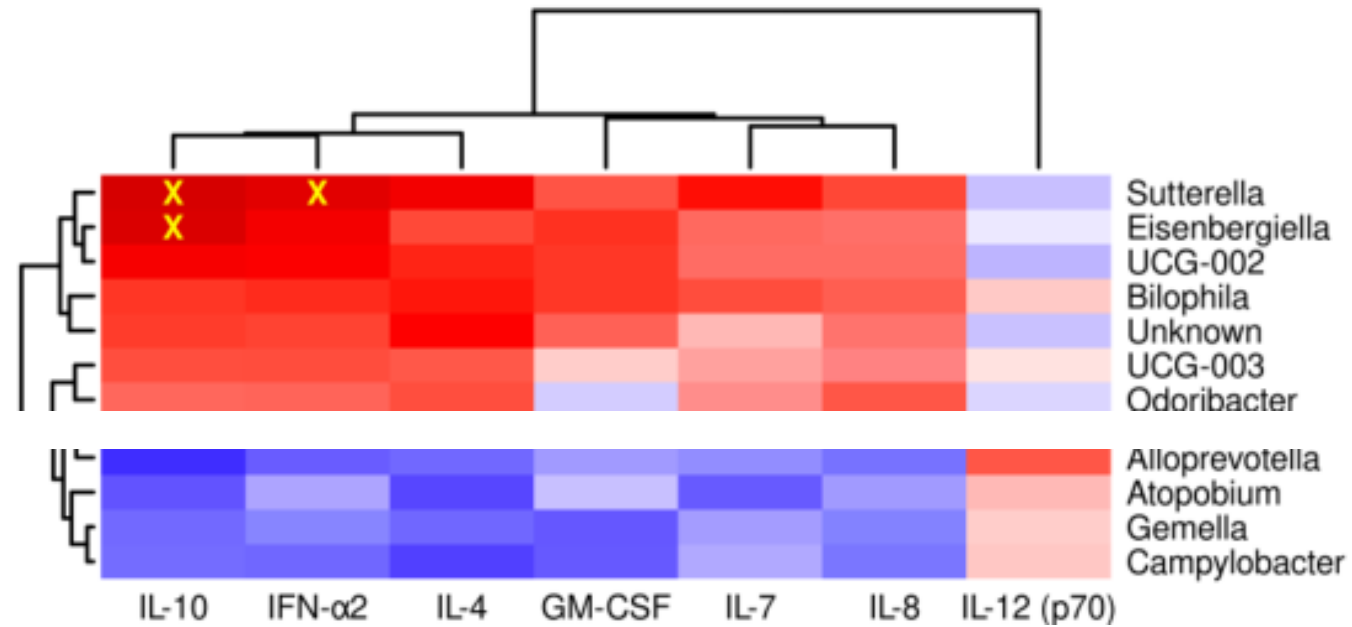
Source: <https://github.com/microbiome/OMA>



# Results of MAPPAC trial

## Main findings

1. **Cytokine** and **microbial** compositions **differ** between appendicitis types
2. Certain microbes **correlate** with certain cytokines
3. Appendicitis type **can be predicted** based on cytokines





# Summary

1. **MultiAssayExperiment** is suitable for multiomics
2. **miaverse** extends **SummarizedExperiment** ecosystem
3. **Appendicitis** can be divided into two different types

# Thank you for your time!

## MAPPAC

Sanja Vanhatalo,  
Eveliina Munukka,  
Tatu Han, Suvi Sippola,  
Sirpa Jalkanen, Juha  
Grönroos, Harri  
Marttila, Erkki Eerola,  
Saija Hurme, Antti J  
Hakanen, Paulina  
Salminen

## miaverse

Felix G.M. Ernst,  
Sudarshan A. Shetty,  
Tuomas Borman,  
Yagmur Simsek, Henrik  
Eckermann, Chandler  
Ross, Chouaib  
Benchraka, Rajesh  
Shigdel, Ruizhu Huang,  
Domenick J. Braccia,  
Héctor Corrada Bravo,  
Leo Lahti

**Slack:**  
#miaverse

**Project website:**  
[microbiome.github.io](https://microbiome.github.io)

