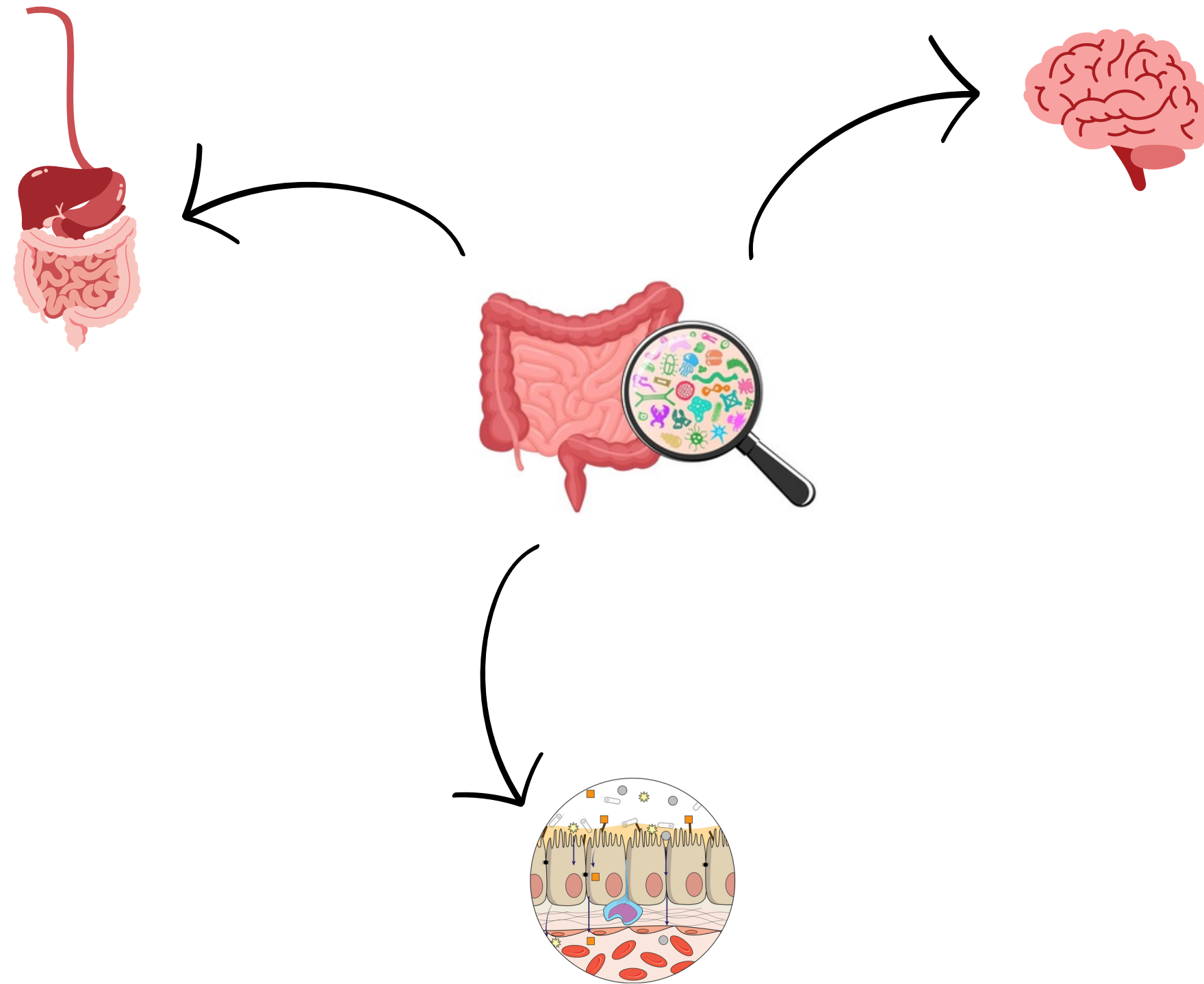


Identification of bacterial strains in gut microbiota: tools comparison and application

Lune Angevin, Josipa Lipovac, Riccardo Vicedomini, Pierre Peterlongo and Emeline Roux



Gut microbiota impacts the entire body

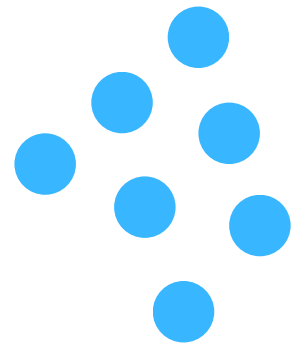


Adhak and Khan, 2019

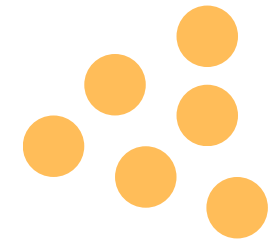
Borre et al., 2014

The gut microbiota influences the host through metabolites production

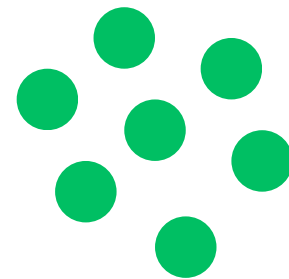
Tryptophan derivatives



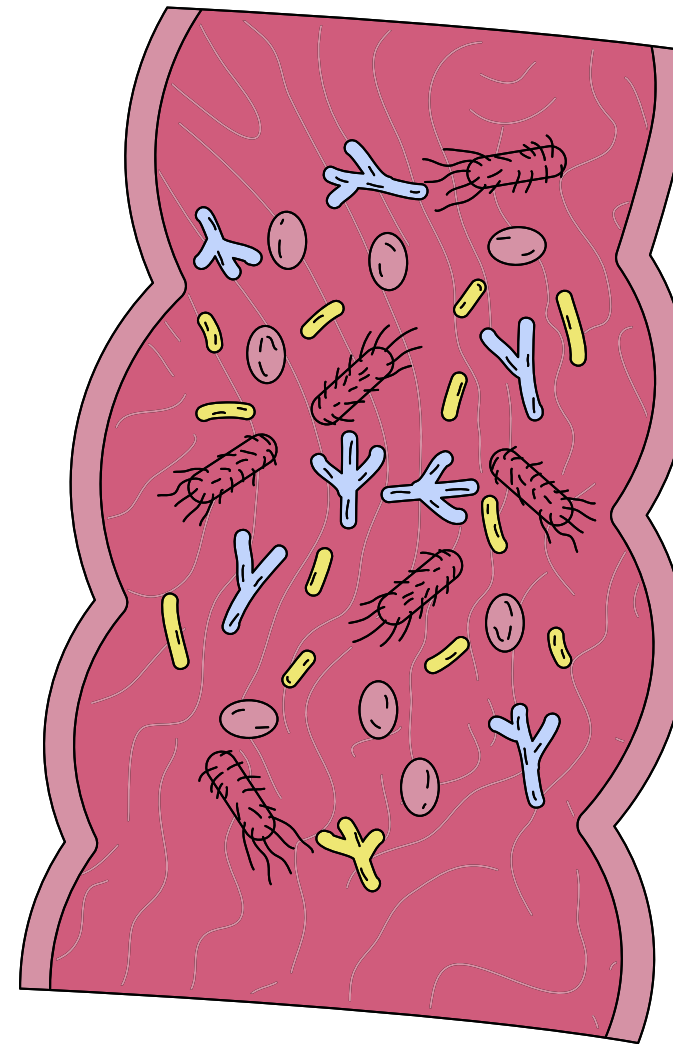
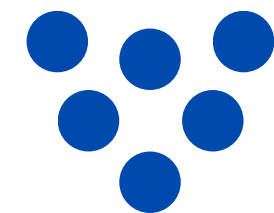
Neurotransmitters



Polyamines



Short-chain fatty acids



Adhak and Khan, 2019

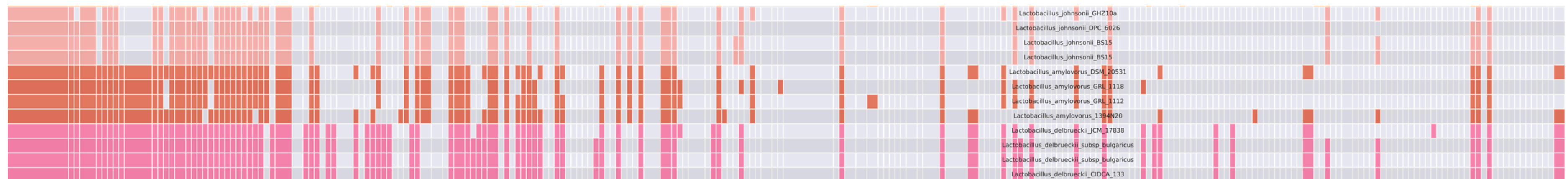
Borre et al., 2014

Actual gut microbiota characterisation: 16S sequencing

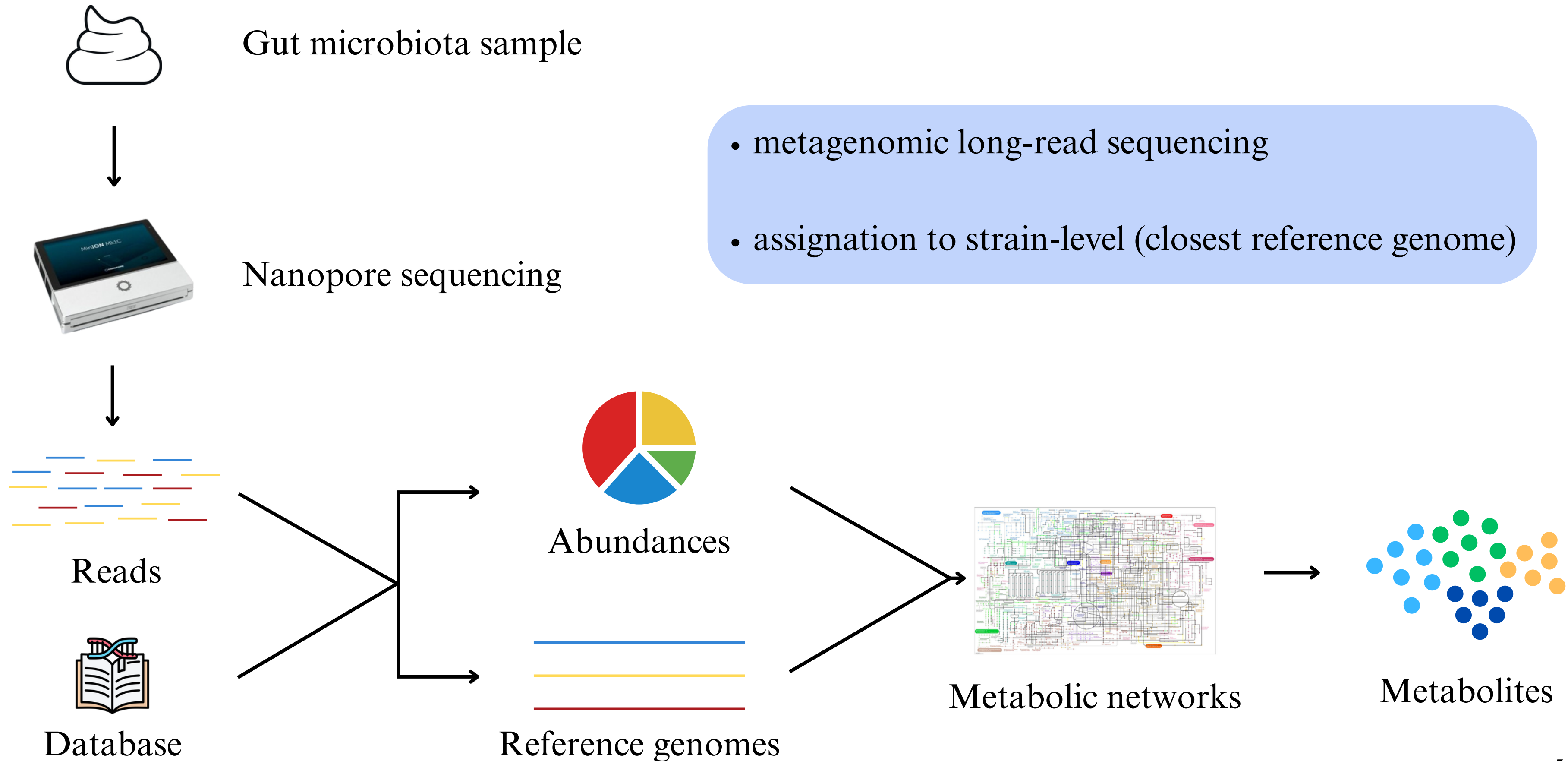
16S sequencing allows to go to genus level

Limiting ourselves to genus level means losing a lot of information in terms of functionalities

Composition of metabolic networks for three species of the genus *Lactobacillus*



Objective of my PhD: characterize gut microbiota and predict metabolites production



Objective of my PhD: characterize gut microbiota and predict metabolites production



Gut microbiota sample



Nanopore sequencing

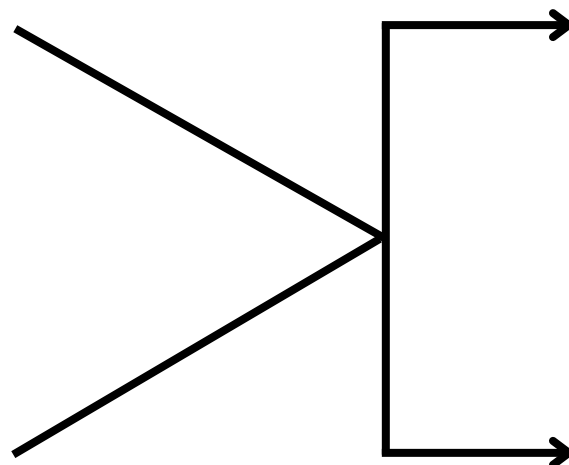
- metagenomic long-read sequencing
- assignation to strain-level (closest reference genome)



Reads



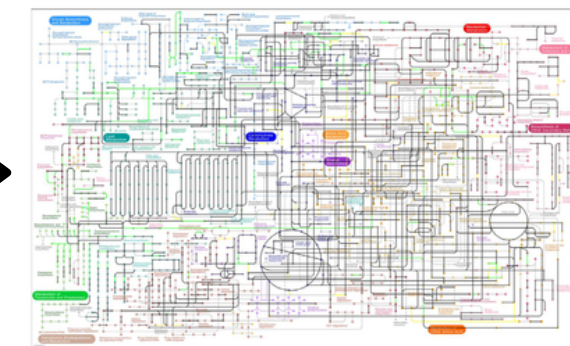
Database



Abundances



Reference genomes



Metabolic networks



Metabolites

State-of-the-art taxonomic assignment tools to species and strain-level for long reads

Species-level

Kraken2 (Wood et al., 2019)

Sylph (Shaw and Yu, 2024)

Strain-level

ORI (Siekaniec et al., 2021)



MADRe (Lipovac et al., 2025, preprint)

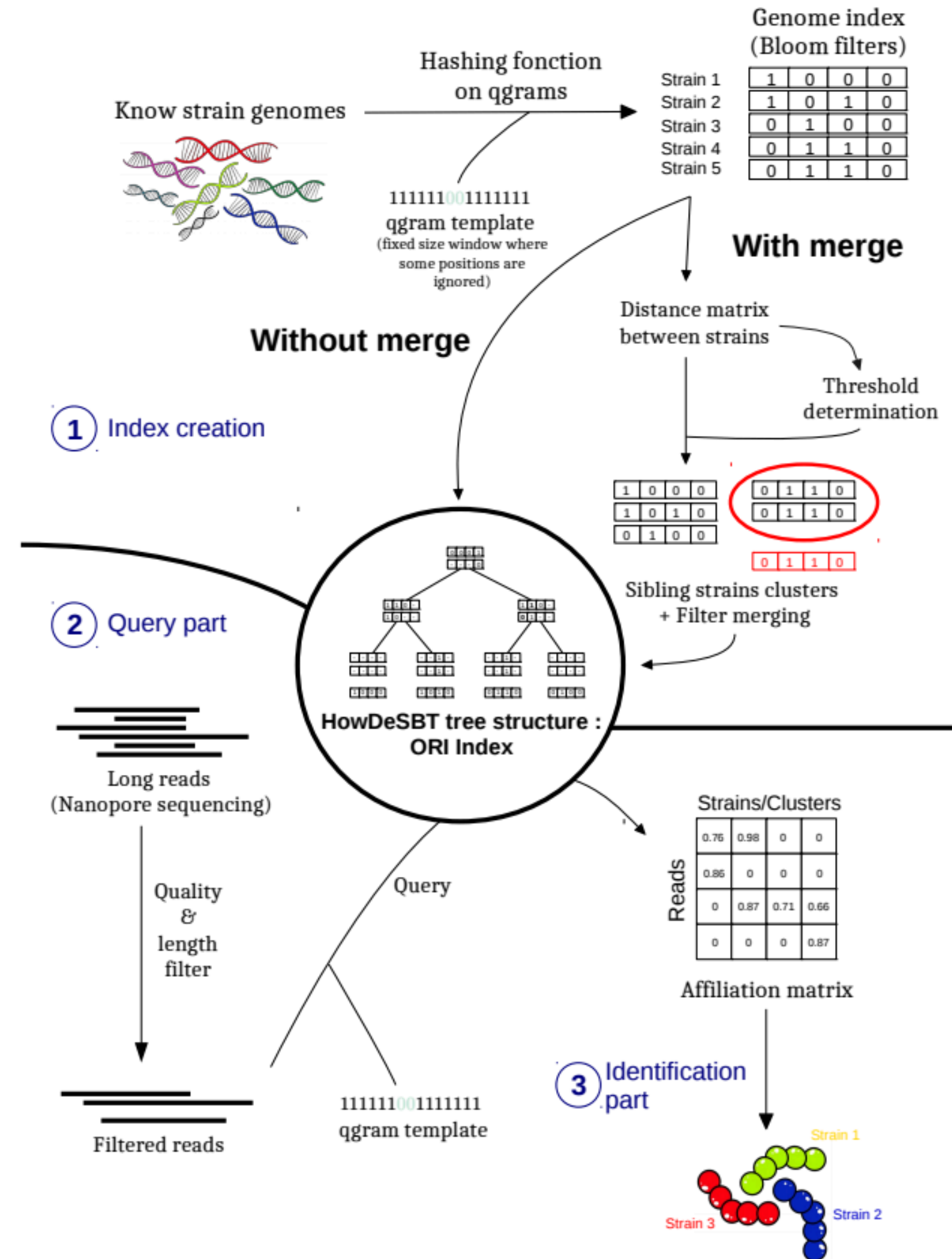


MORA (Zheng et al., 2024)

ORI¹: taxonomic assignment to strain-level

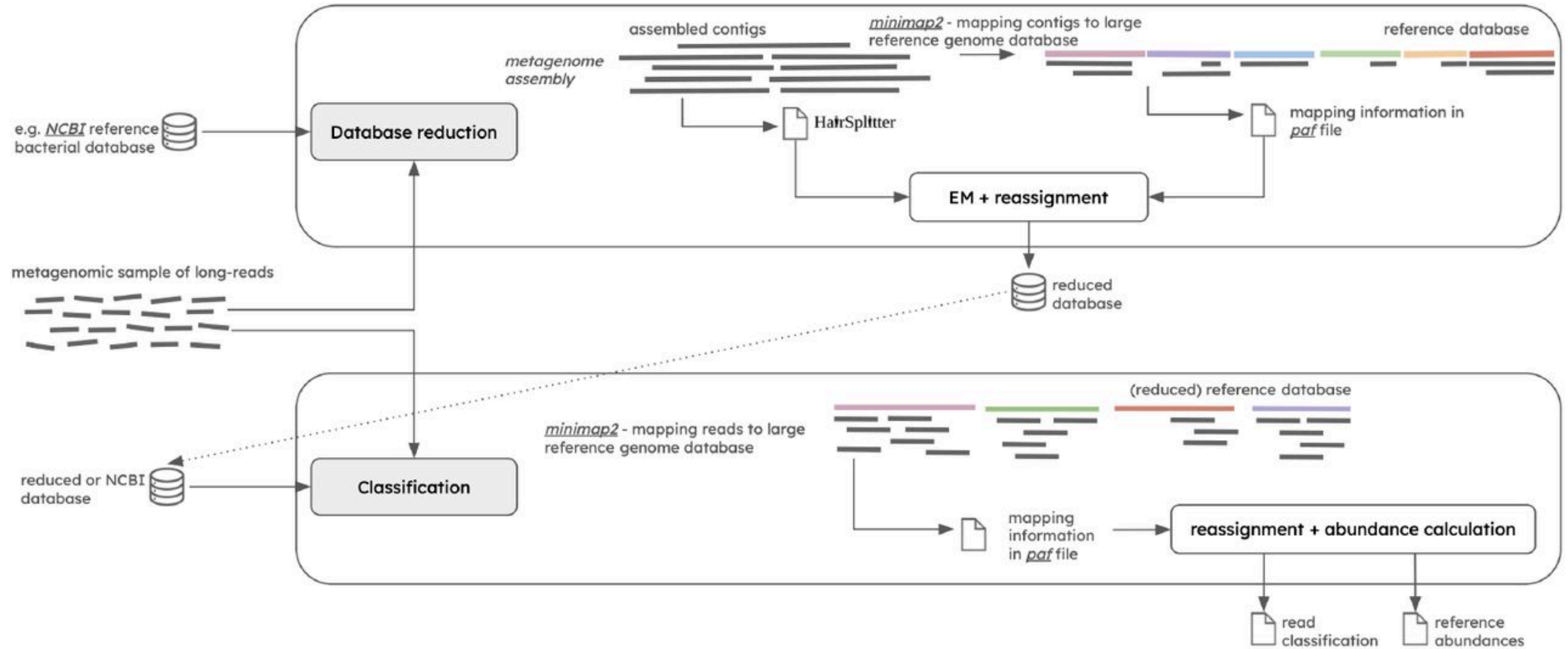
To run after a first step
of taxonomic assignment at species-level

Works with k-mers and spaced seeds patterns



¹Siekaniec et al., 2021

MADRe²: taxonomic assignment to strain-level



How to compare these tools?

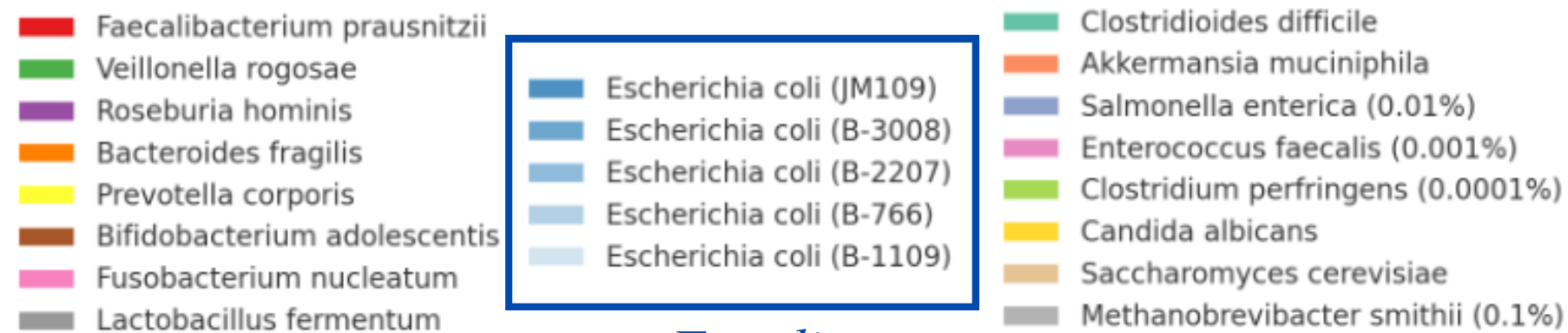
Mock community = artificial assembly of microorganisms in known and controlled proportions

Mock community ZymoBIOMICS GUT³

“Human Gut Microbiota”

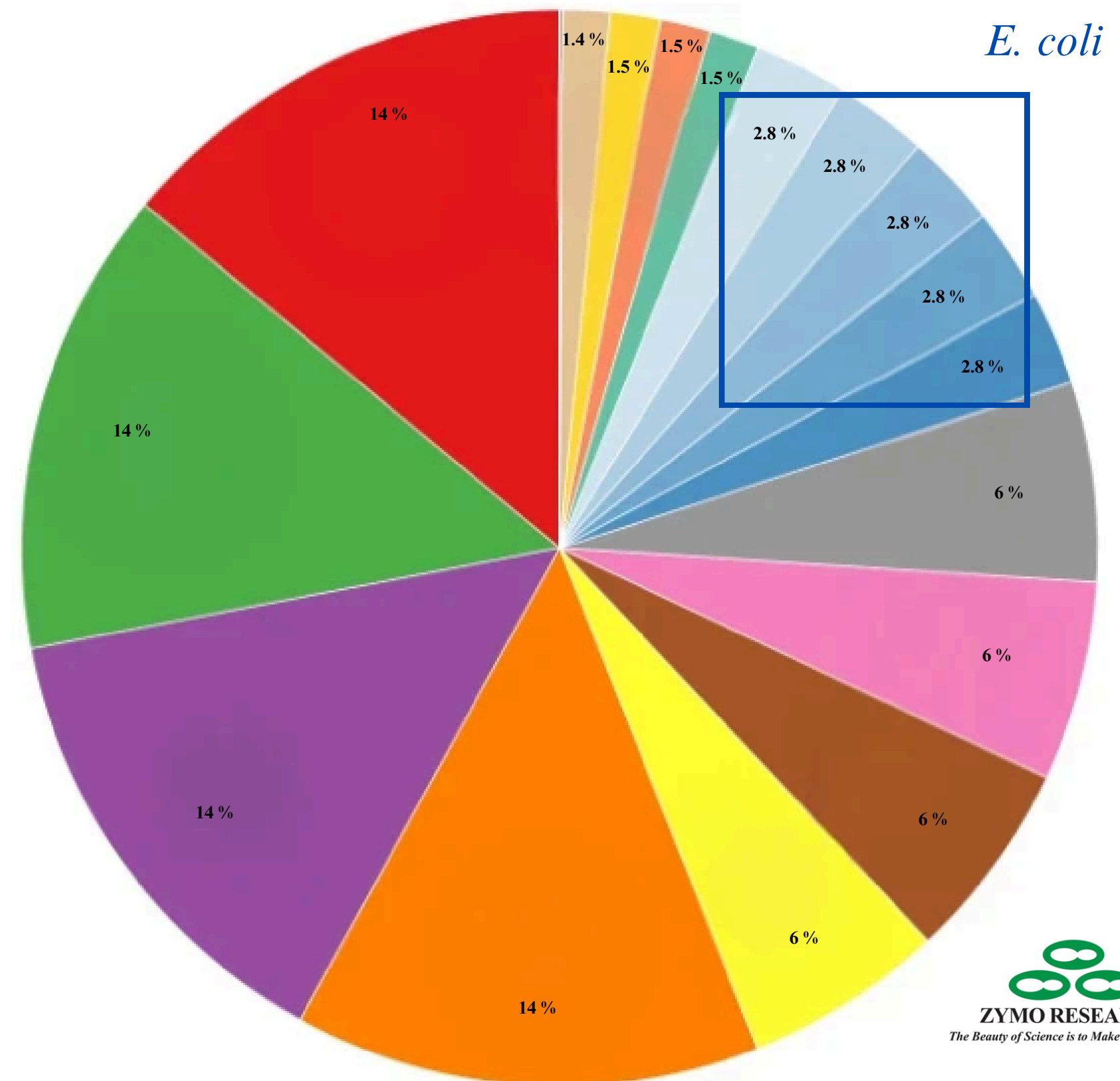
→ 2 yeasts + 1 archaea

→ 18 bacteria including 5 strains from *E. coli*



E. coli

Nanopore R10.4 sequencing reads publicly available
(most recent method)

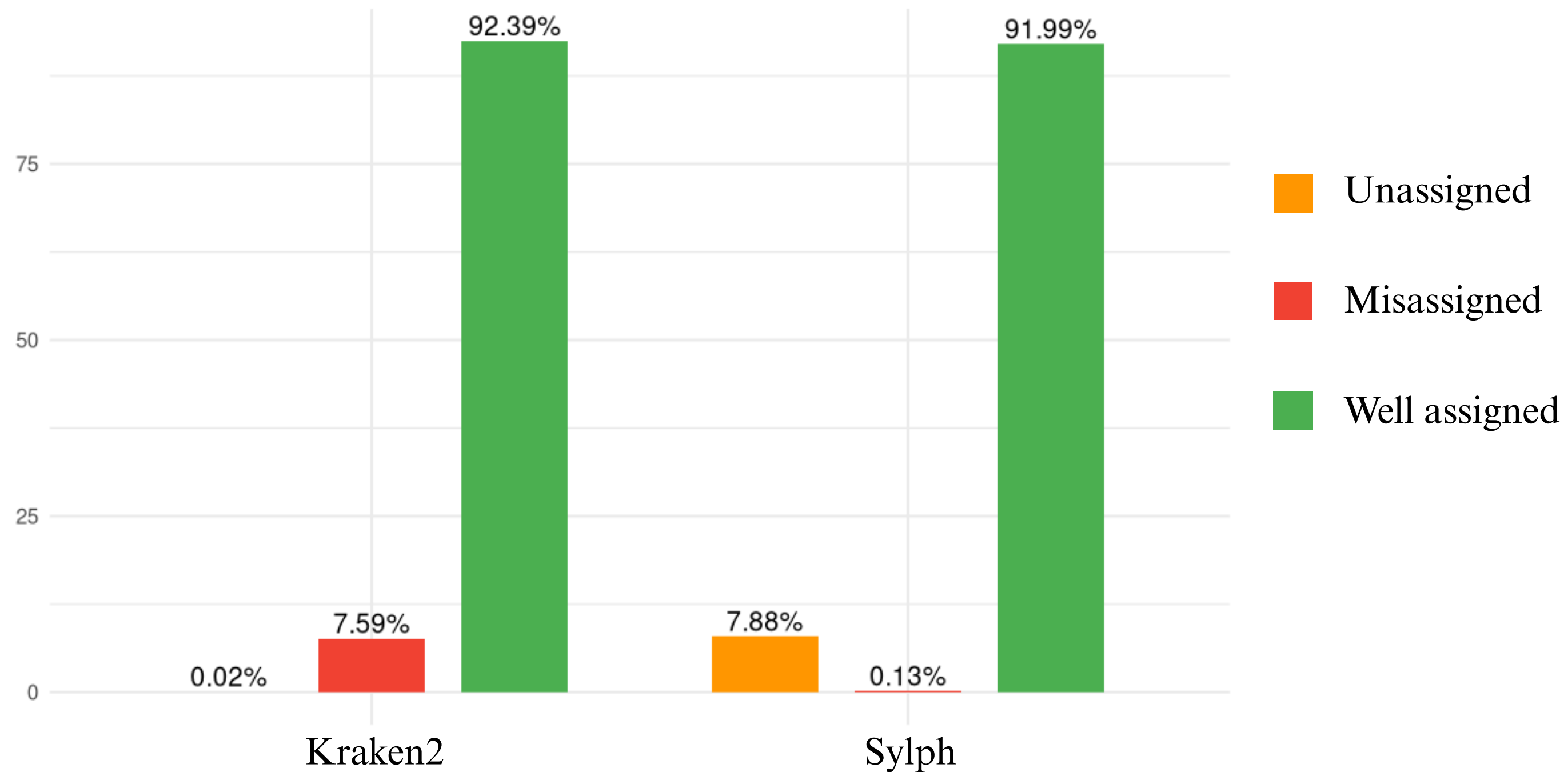


³<https://zymoresearch.eu/products/zymbiomics-gut-microbiome-standard>

⁴<https://github.com/Kirk3gaard/MicroBench?tab=readme-ov-file#zymogut>

Taxonomic assignment at species-level: Kraken2⁵ vs Sylph⁶

Comparison of reads assignment performances by Kraken2 and Sylph

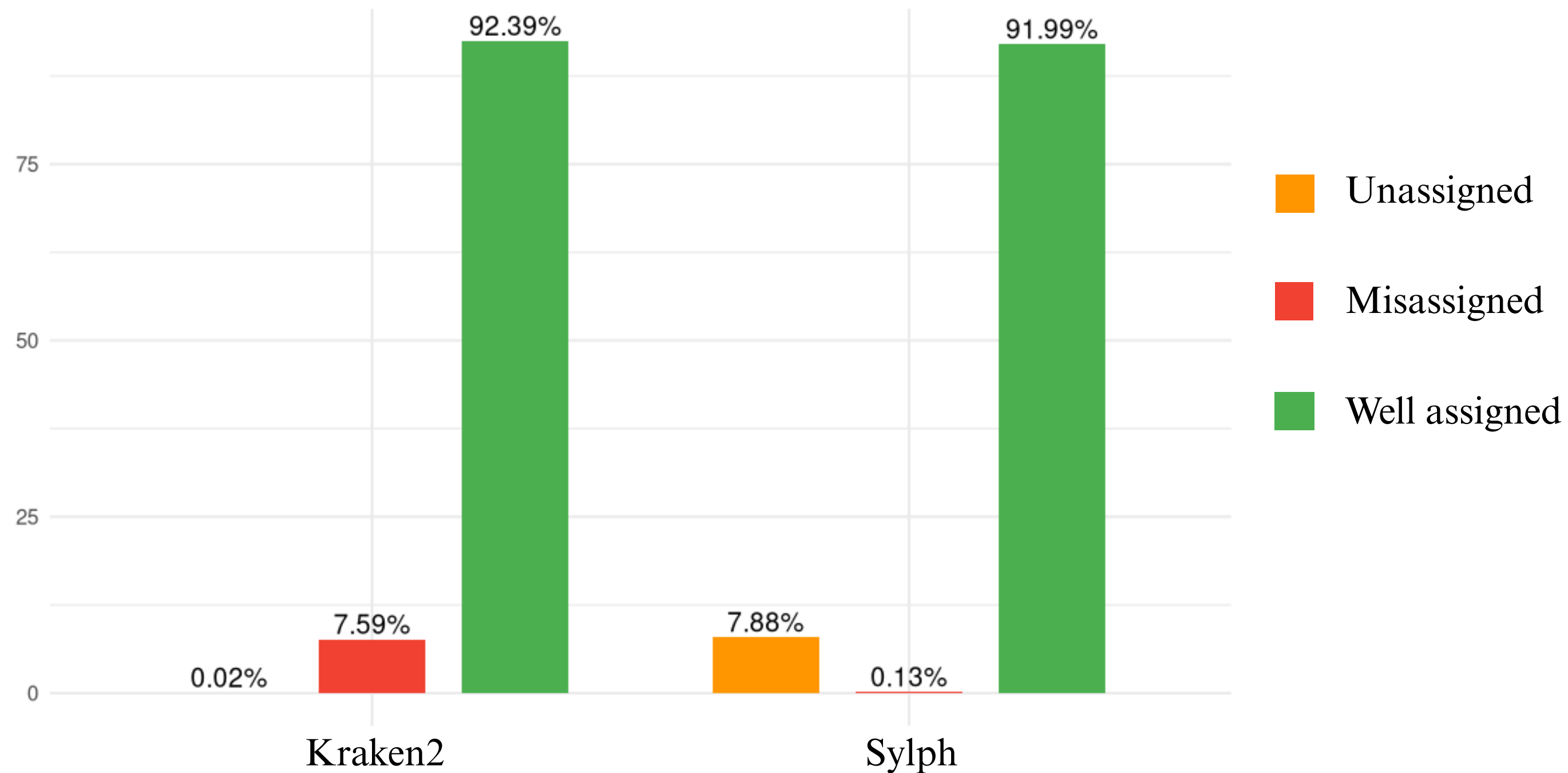


⁵Wood et al., 2019

⁶Shaw and Yu, 2024

Taxonomic assignment at species-level: Kraken2 vs Sylph

Comparison of reads assignment performances by Kraken2 and Sylph



⁵Wood et al., 2019

⁶Shaw and Yu, 2024



Choosing Sylph for an assignment at the species level

Taxonomic assignment at strain-level: ORI¹, preliminary evaluation

	ORI results
Other species	9 strains / 13
<i>Escherichia coli</i>	1 strain / 5
False positives	5%

Unidentified species

→ abundances under 0.1%

→ *Bifidobacterium adolescentis*

For *E. coli*: queried only 1% of reads

Parameters to optimize for complex microbiota

¹Siekaniec et al., 2021

Taxonomic assignment at strain-level: MADRe², preliminary evaluation

	MADRe
Other species	10 strains / 13
<i>Escherichia coli</i>	5 strains / 5 + 62 other strains
False positives	88.5%

Unidentified species
→ abundances under 0.01%
→ *Fusobacterium nucleatum*

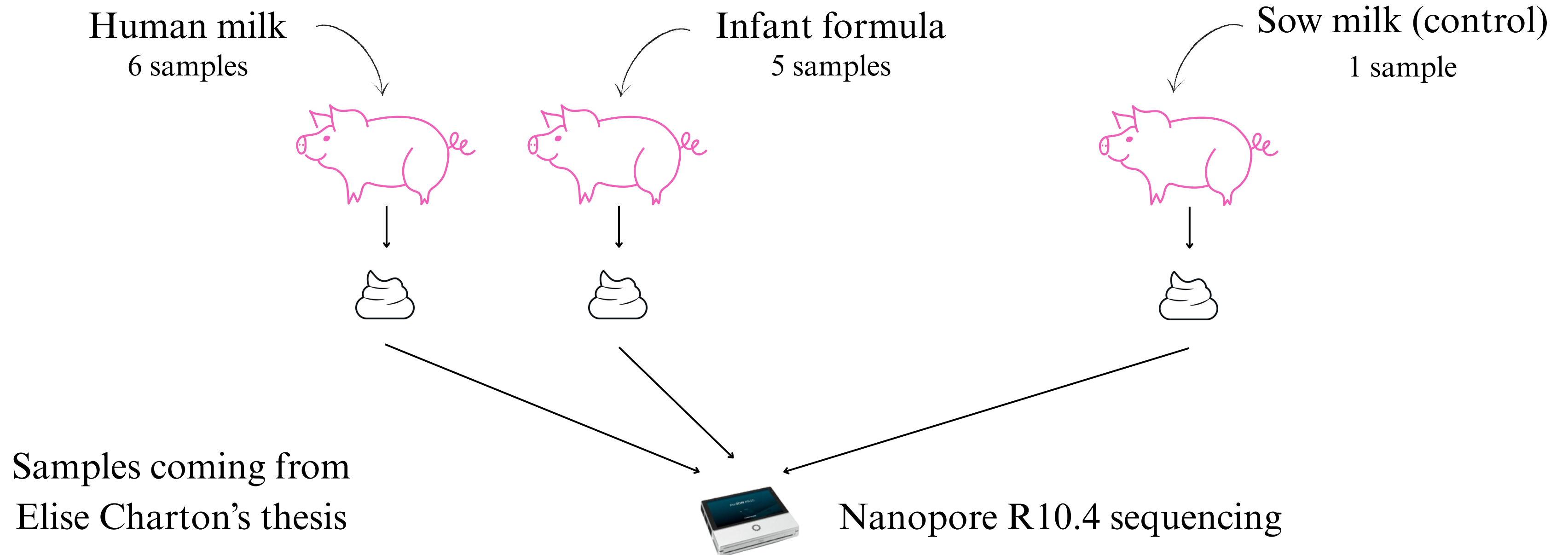
Still under development
Possible threshold choice

²Lipovac et al., 2025

Application: piglets gut microbiota

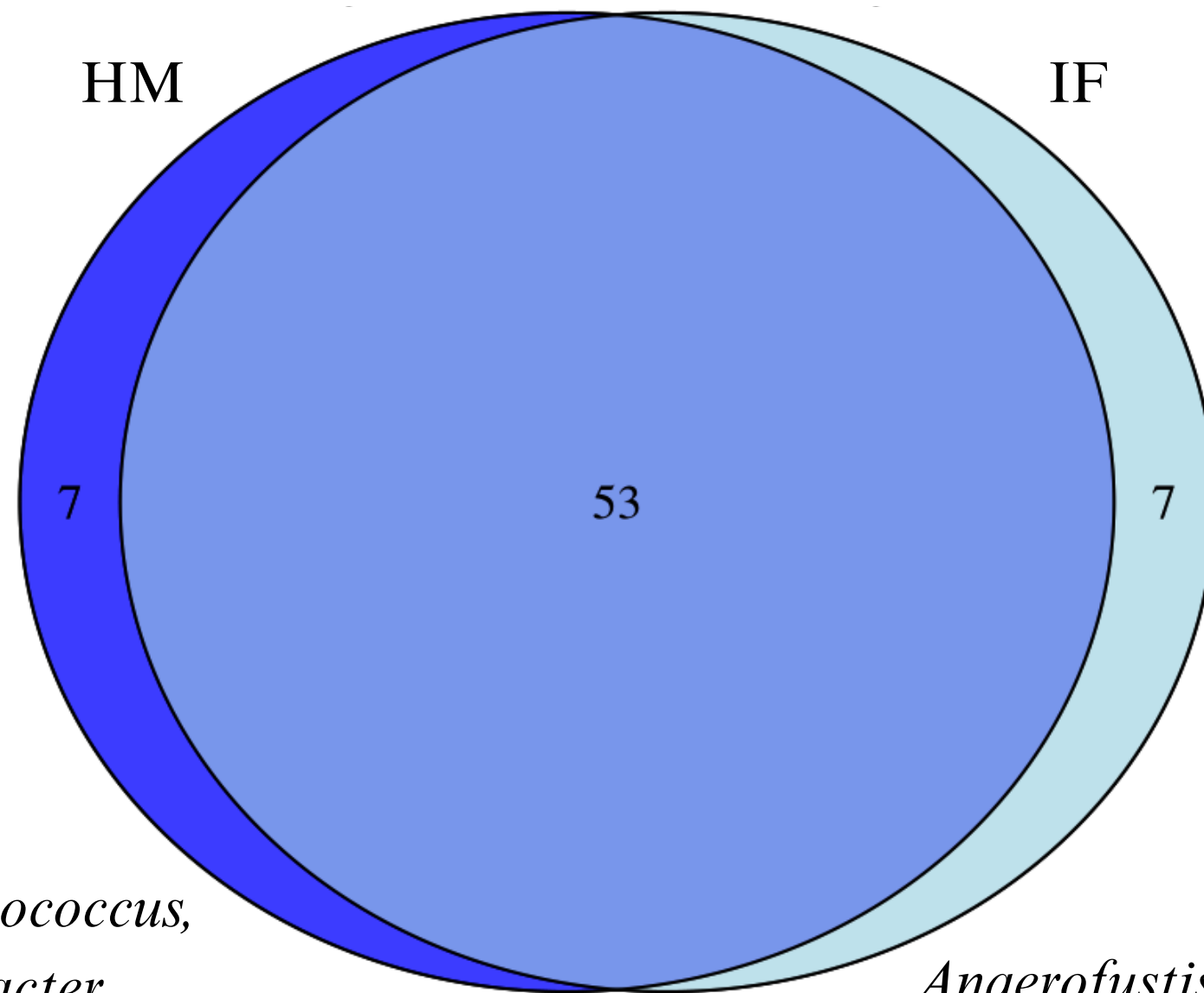


What are the differences in the microbiota composition and producible metabolites in newborns depending on their diet?



Taxonomic assignment at species-level with Sylph⁶

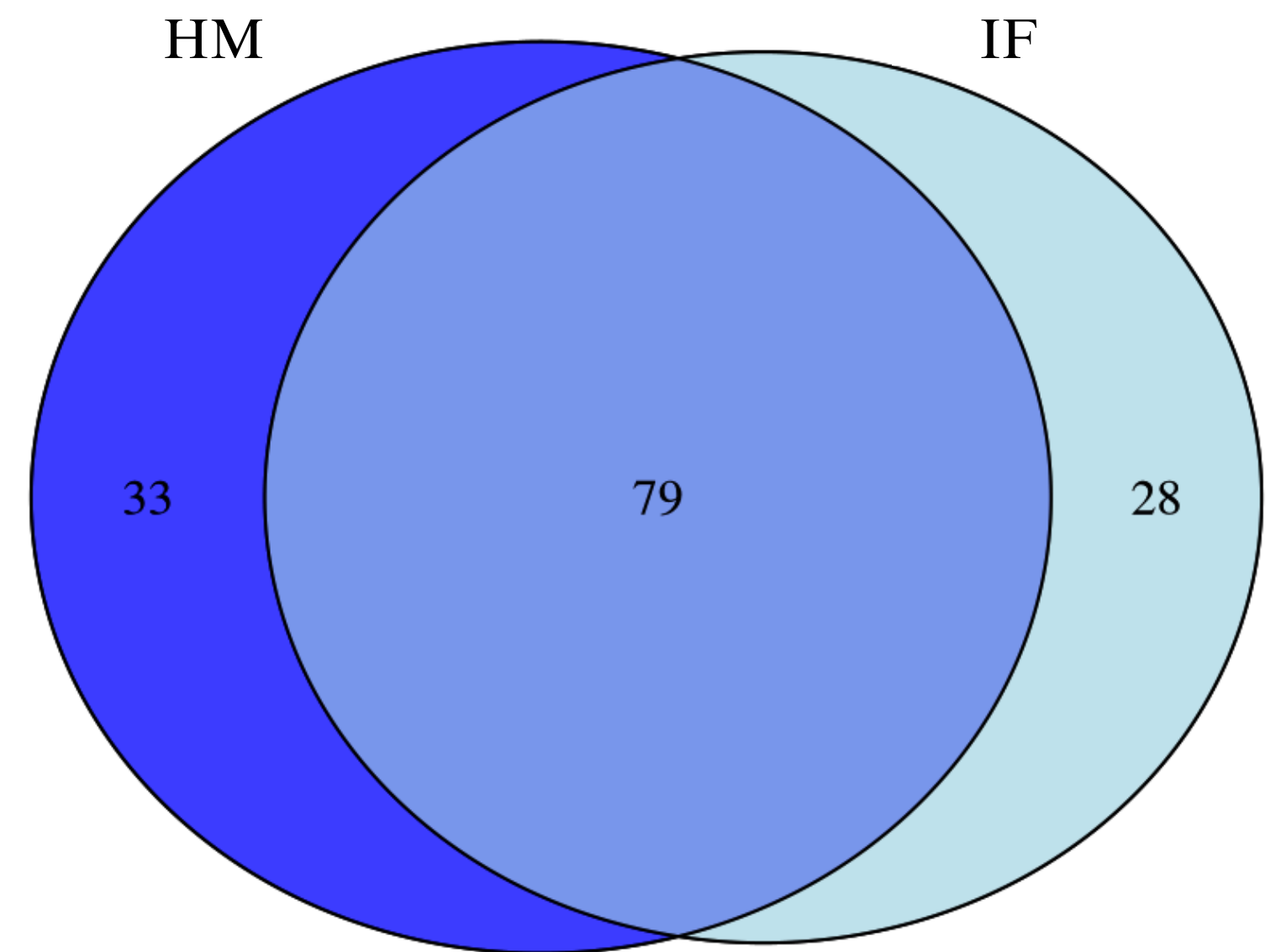
Venn diagram of composition at genus-level



Acidaminococcus,
Acinetobacter,
Collinsella, Helicobacter,
Propionibacterium, Pseudomonas,
Sodaliphilus

Anaerofustis, Berryella,
Clostridioides, Lactococcus,
Romboutsia, Trueperella,
Veillonella

Venn diagram of composition at species-level



⁶Shaw and Yu, 2024

Conclusion and perspectives

Taxonomic assignment at species-level : Sylph

Taxonomic assignement at strain-level : ORI / MADRe / MORA / ... ?

Characterisation of piglets gut microbiota depending on the diet

Reconstruction of metabolic networks and prediction metabolites production depending on the diet

Improve infant formulas

Thanks to the teams

Thanks to my supervisors

Emeline Roux

Pierre Peterlongo

Riccardo Vicedomini



Thanks to Josipa Lipovac



Inserm

INRAE

Inria

