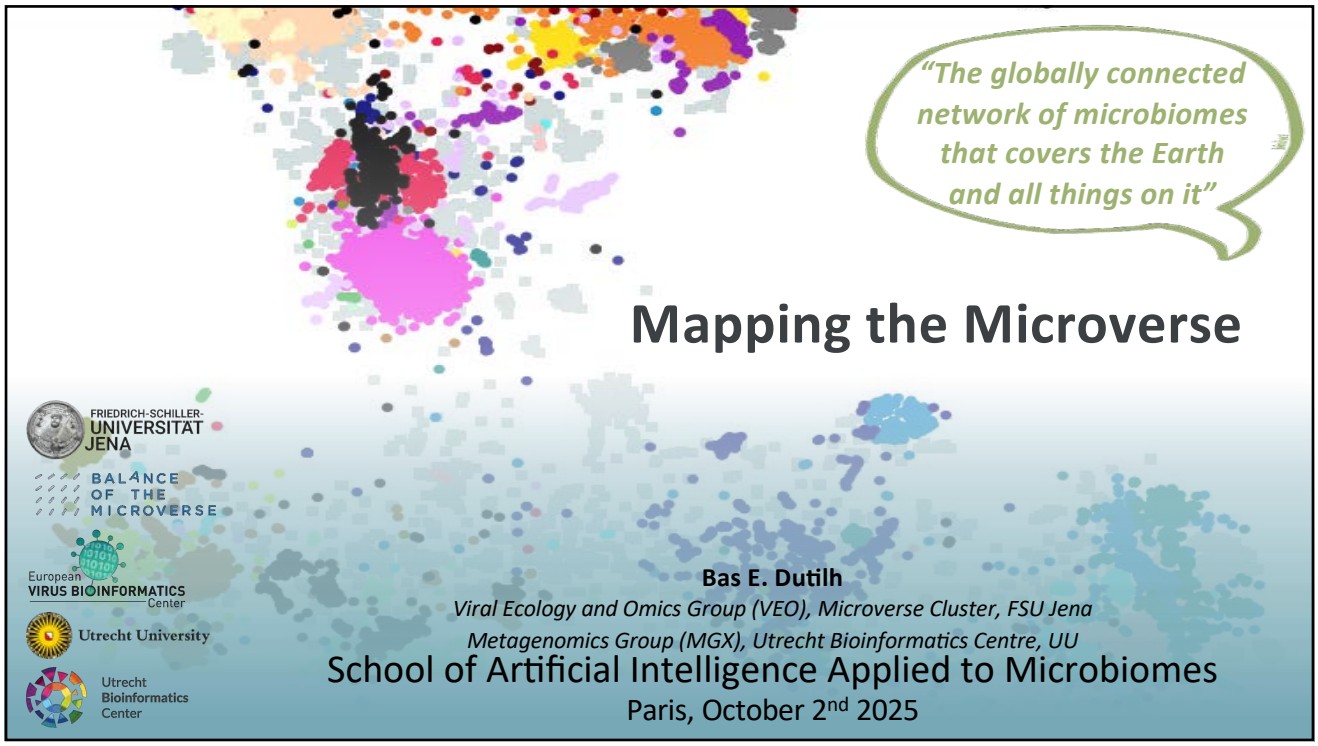
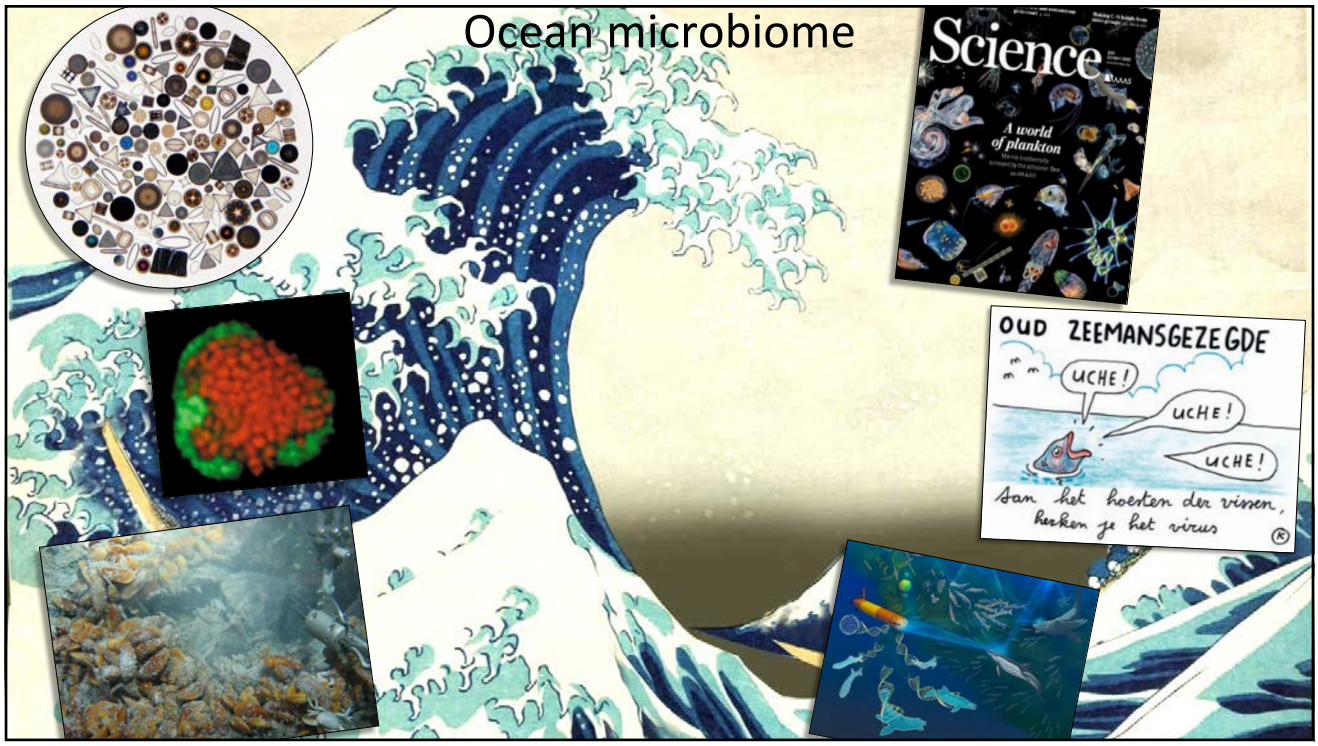
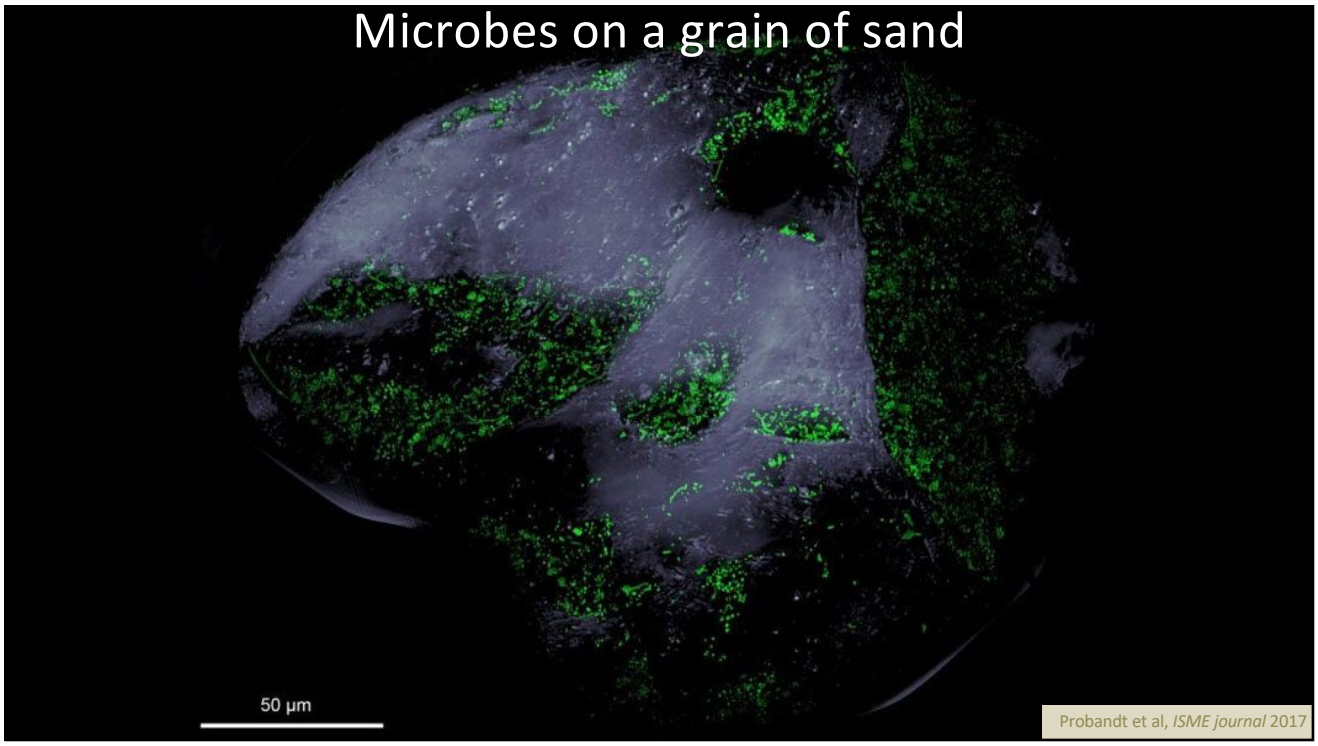

"The globally connected network of microbiomes that covers the Earth and all things on it"
Mapping the Microverse
FRIEDRICH-SCHILLER-UNIVERSITÄT JENA
BALANCE OF THE MICROVERSE
European VIRUS BIOINFORMATICS Center
Utrecht University
Utrecht Bioinformatics Center
Bas E. Dutilh
Viral Ecology and Omics Group (VEO), Microverse Cluster, FSU Jena
Metagenomics Group (MGX), Utrecht Bioinformatics Centre, UU
School of Artificial Intelligence Applied to Microbiomes
Paris, October 2nd 2025

1


Ocean microbiome
Science
A world of plankton
OUD ZEEMANSGEZEGDE
UCHE!
UCHE!
UCHE!
Aan het hoesten der vissen, herken je het virus

2

Microbes on a grain of sand

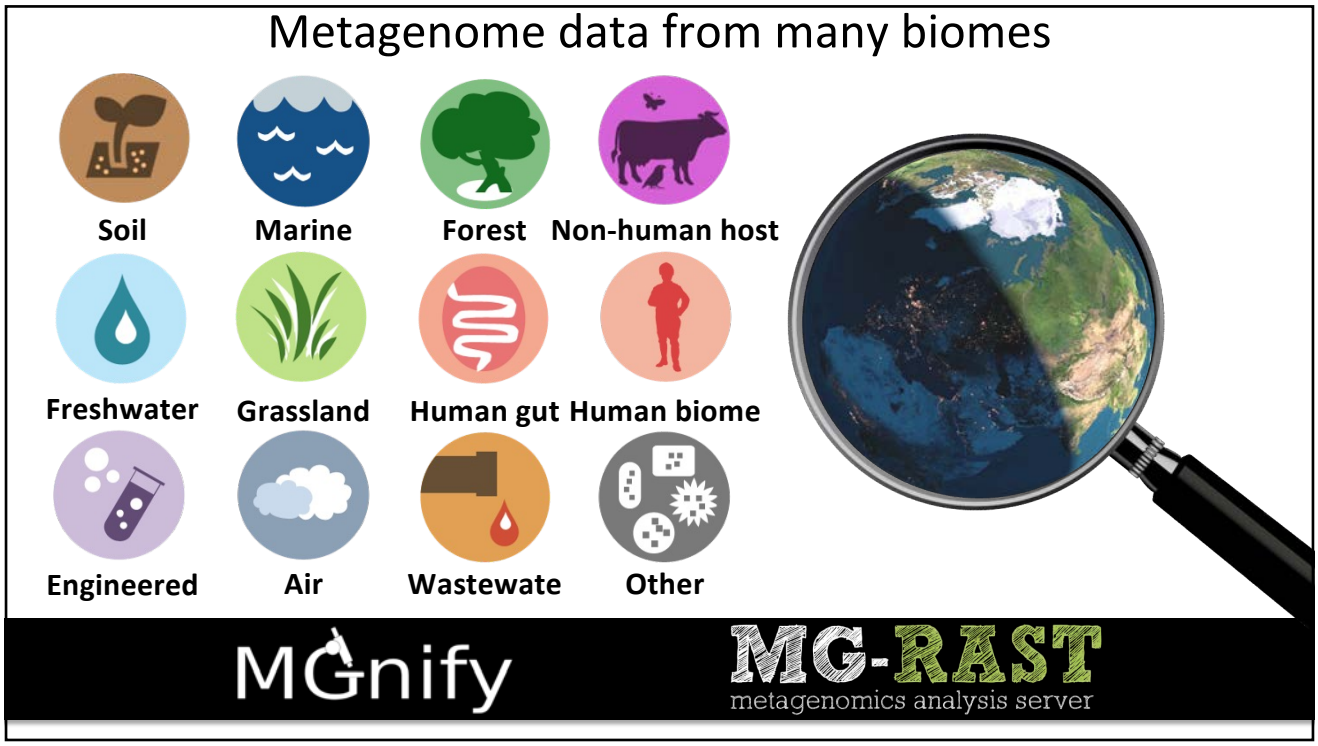


Probandt et al, *ISME journal* 2017

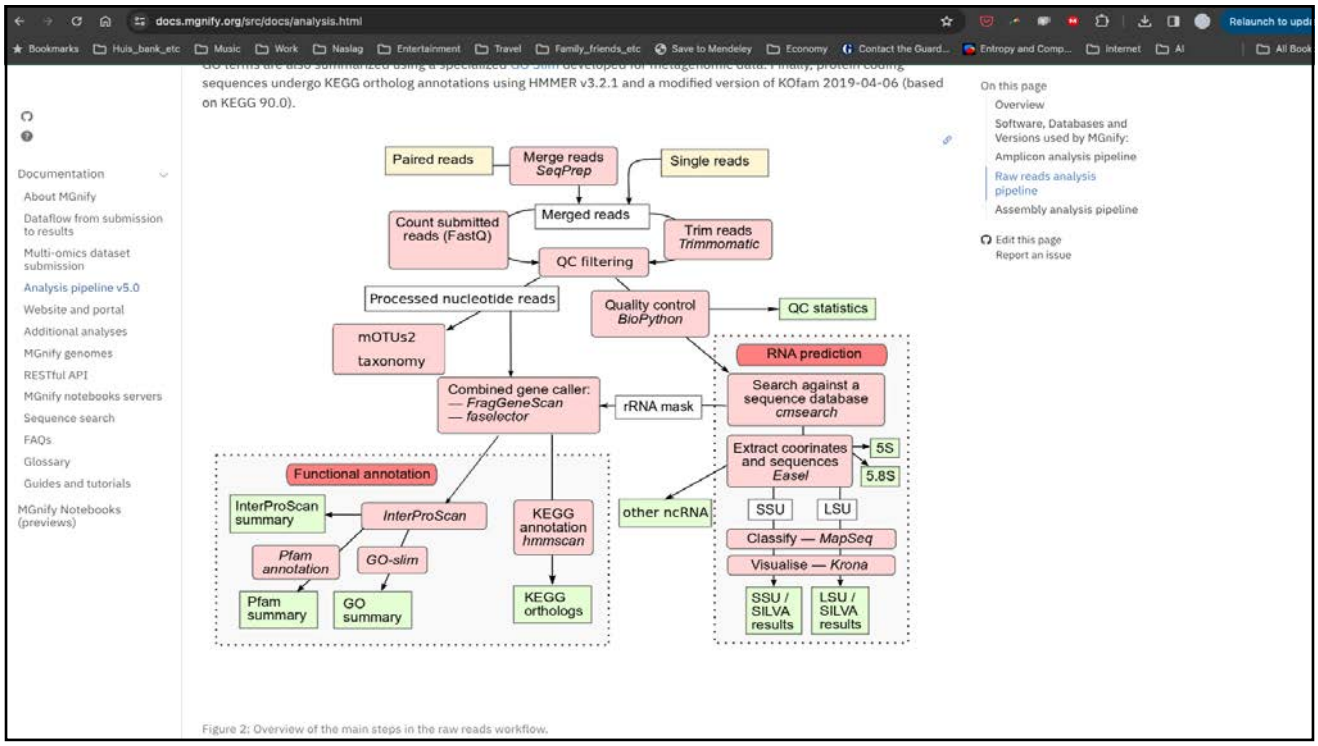
3

Human microbiome

4

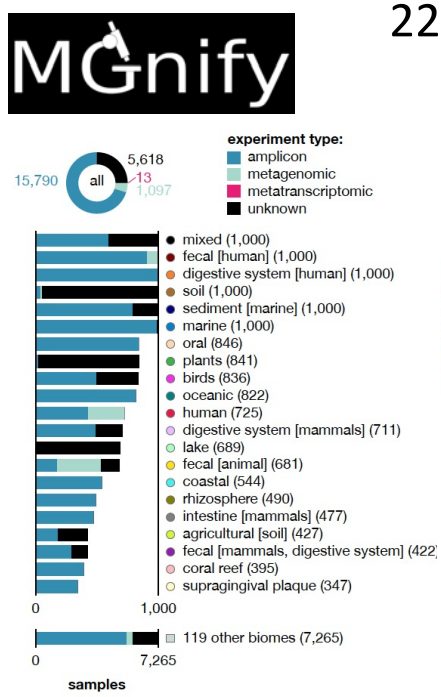

Metagenome data from many biomes
Soil
Marine
Forest
Non-human host
Freshwater
Grassland
Human gut
Human biome
Engineered
Air
Wastewate
Other
MGnify
MG-RAST
metagenomics analysis server

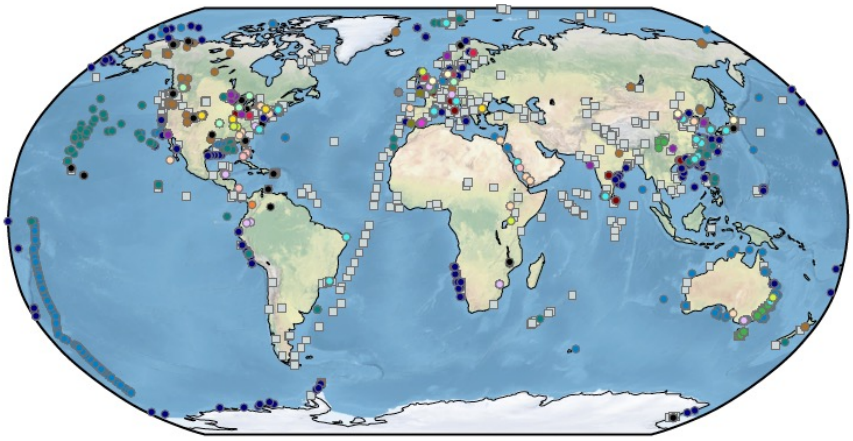
5


docs.mgnify.org/src/docs/analysis.html
sequences undergo KEGG ortholog annotations using HMMER v3.2.1 and a modified version of KOfam 2019-04-06 (based on KEGG 90.0).
Documentation
About MGnify
Dataflow from submission to results
Multi-omics dataset submission
Analysis pipeline v5.0
Website and portal
Additional analyses
MGnify genomes
RESTful API
MGnify notebooks servers
Sequence search
FAQs
Glossary
Guides and tutorials
MGnify Notebooks (previews)
On this page
Overview
Software, Databases and Versions used by MGnify:
Amplicon analysis pipeline
Raw reads analysis pipeline
Assembly analysis pipeline
Edit this page
Report an issue
Paired reads
Merge reads SeqPrep
Single reads
Merged reads
Count submitted reads (FastQ)
Trim reads Trimmomatic
QC filtering
Processed nucleotide reads
Quality control BioPython
QC statistics
mOTUs2 taxonomy
RNA prediction
Combined gene caller: — FragGeneScan — faselector
rRNA mask
Search against a sequence database cmsearch
Extract coorinates and sequences Easel
5S
5.8S
Functional annotation
InterProScan summary
InterProScan
KEGG annotation hmmscan
other ncRNA
SSU
LSU
Pfam annotation
GO-slim
Classify — MapSeq
Visualise — Krona
Pfam summary
GO summary
KEGG orthologs
SSU / SILVA results
LSU / SILVA results
Figure 2: Overview of the main steps in the raw reads workflow.

6

22,518 metagenomes from 140 biomes

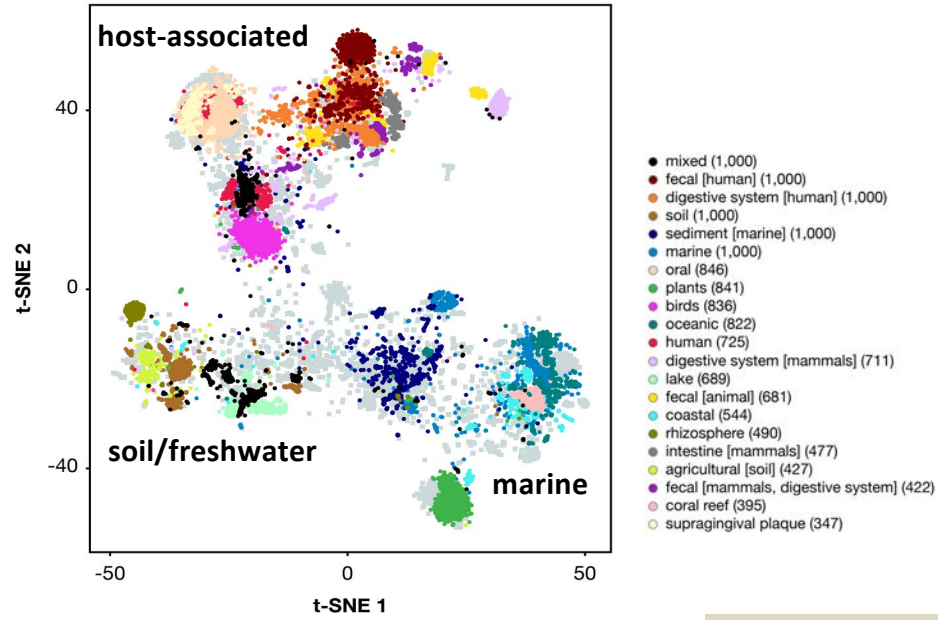




Von Meijenfeldt, Hogeweg & Dutilh, *Nature Ecology & Evolution* 2023

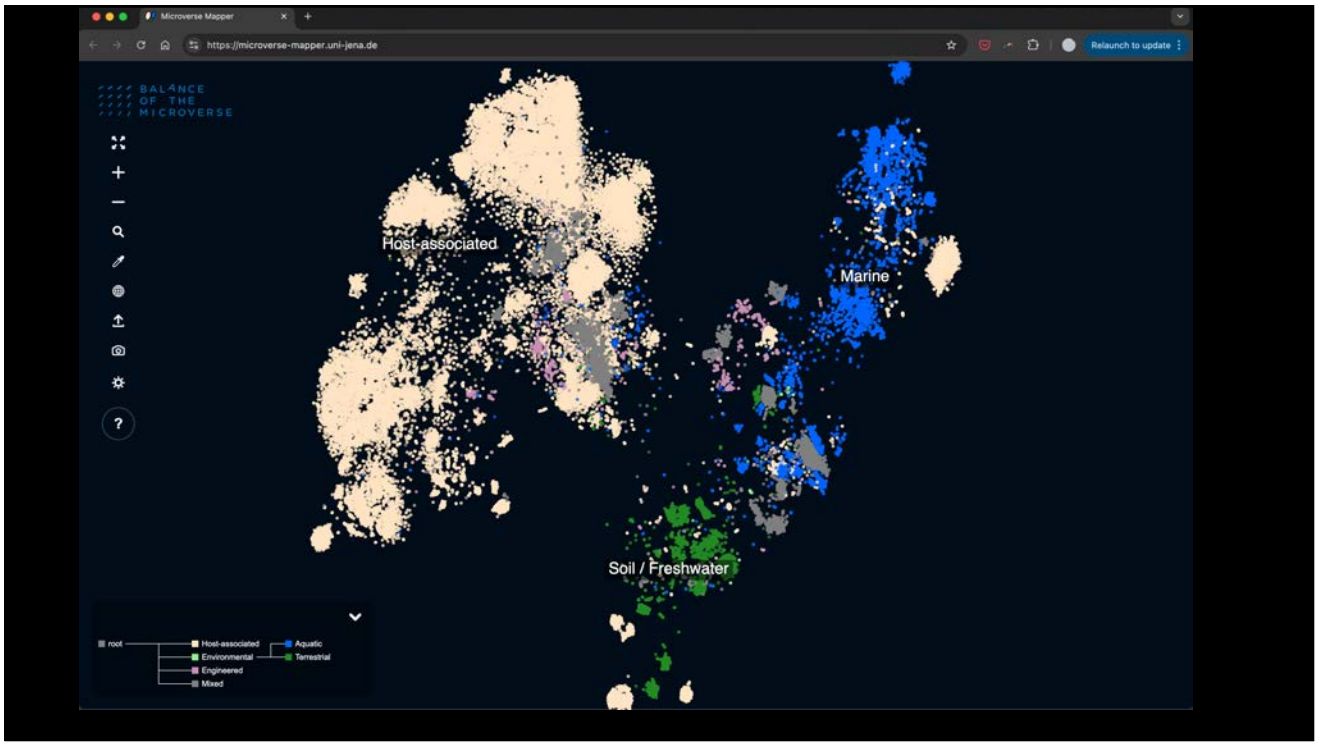
7

Map of the Microverse



Von Meijenfeldt, Hogeweg & Dutilh, *Nature Ecology & Evolution* 2023

8



9

Assignments

https://microverse-mapper.uni-jena.de/

1. Please take a look at the Microverse Mapper website.
   a) Read the Help page.
2. *Roseobacter*:
   a) In how many microbiome samples was *Roseobacter* found? *
   b) From how many different biomes were these samples derived?
   c) In which biome is *Roseobacter* typically found?
   d) What are the 10 most abundant other genera in marine samples containing at least 100 *Roseobacter* reads?
   e) Which of these are generalists and which are specialists?

* Cutoff: 100 reads

10

diagram, you can reduce the selection to only the marine samples. Next, you can go to the "Barchart" tab and select the "genus" rank to find the top ten genera (note that they are reverse sorted in the color legend):

1. *Desulfobacula*
2. *Amylibacter*
3. *Candidatus* Actinomarina
4. *Roseobacter*
5. *Glaciecola*
6. *Acinetobacter*
7. *Ascidiacei-habitans*
8. *Sulfitobacter*
9. *Algicola*
10. *Roseovarius*

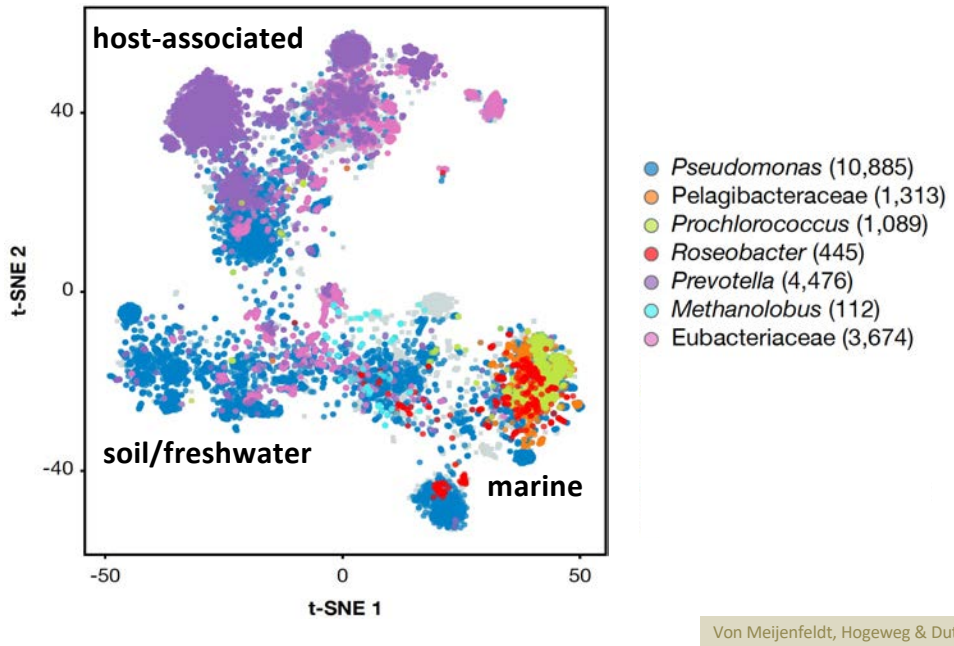
Searching these again with the "Smart Search" function (100 read cutoff) yields the following numbers of samples in the "Biome" tab & if you download the csv file you can find the biomes (scroll down to the bottom of the pie chart color legend):

1. 41 & 4
2. 985 & 26
3. 2194 & 30
4. 156 & 11
5. 458 & 24
6. 9742 & 102
7. 573 & 18
8. 334 & 16
9. 34 & 6
10. 245 & 19

Based on these results, *Desulfobacula* is the most specialist and *Acinetobacter* the most generalist genus.

11

Map of the Microverse



Von Meijenfeldt, Hogeweg & Dutilh, *Nature Ecology & Evolution* 2023

12

Generalist ⇔ Specialist

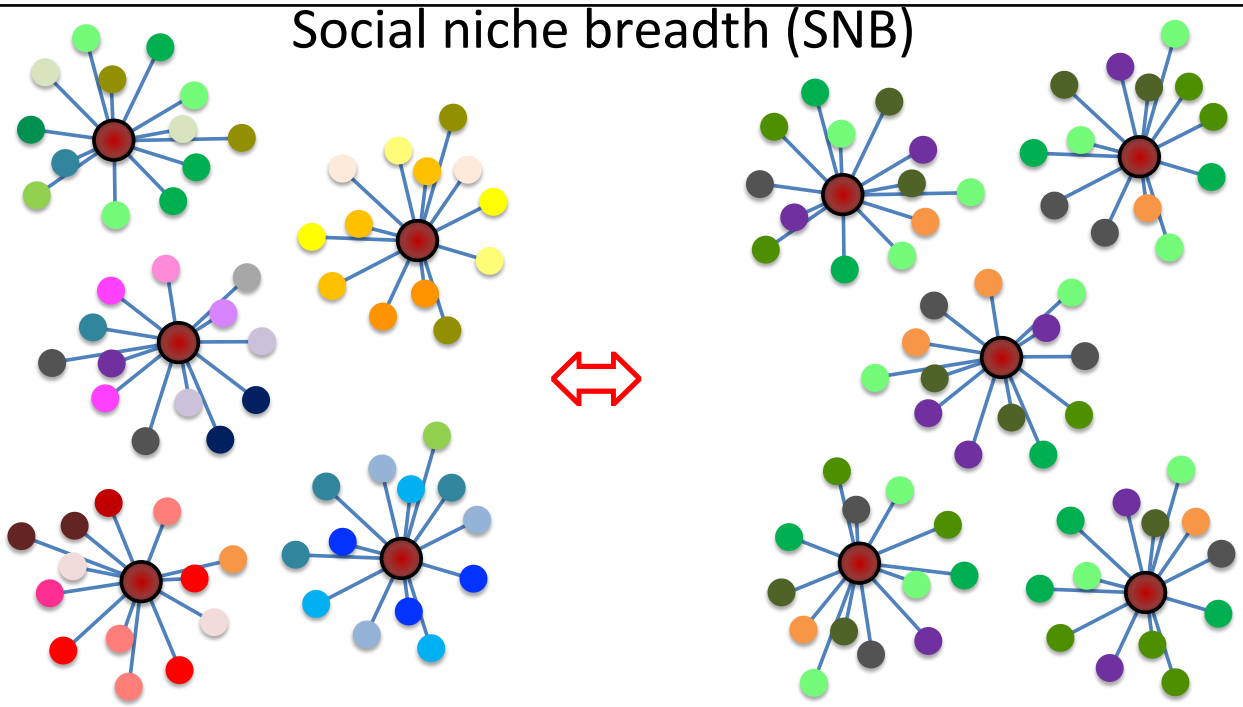
13

Assignments

3. Microverse Mapper
   a) Which oral samples stand out?
   b) Can you figure out why?
   c) What does this tell you about the data?
4. How many different types of bioreactors are there?

14

Social niche breadth (SNB)



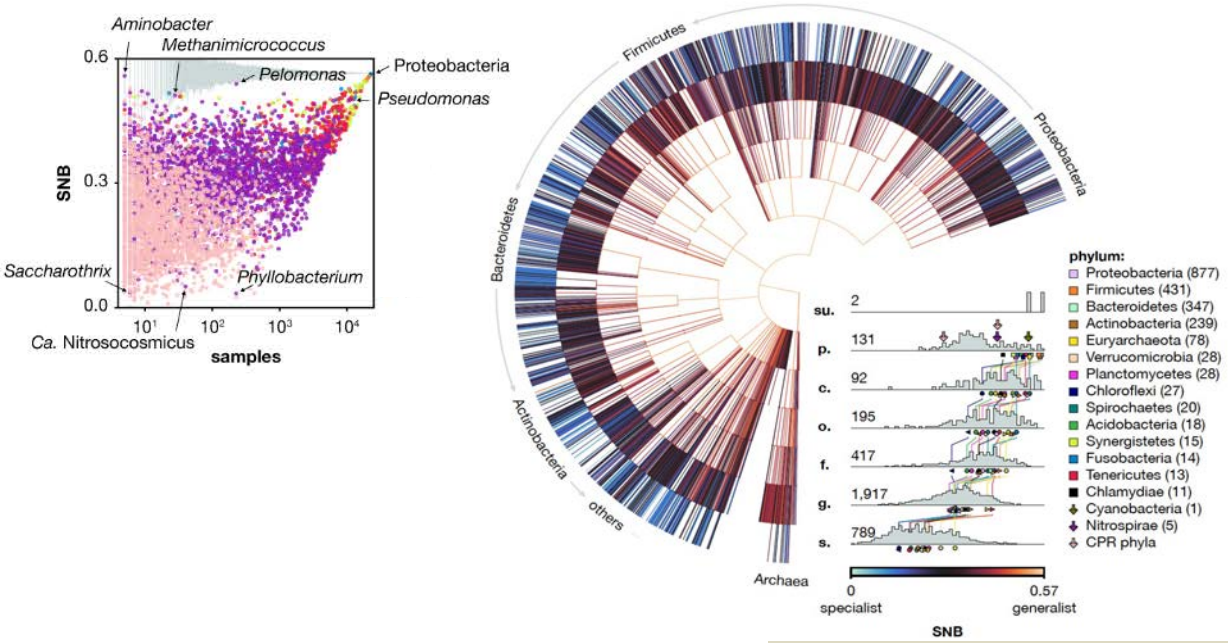
15

Assignment

5. *Roseobacter* again
   a) What is the taxonomic lineage of *Roseobacter*?
      - Genus: *Roseobacter*
      - Family:
      - Order:
      - Class:
      - Phylum:
   b) In how many samples/biomes are each of these found?

16

SNB at different taxonomic ranks



Von Meijenfeldt, Hogeweg & Dutilh, *Nature Ecology & Evolution* 2023

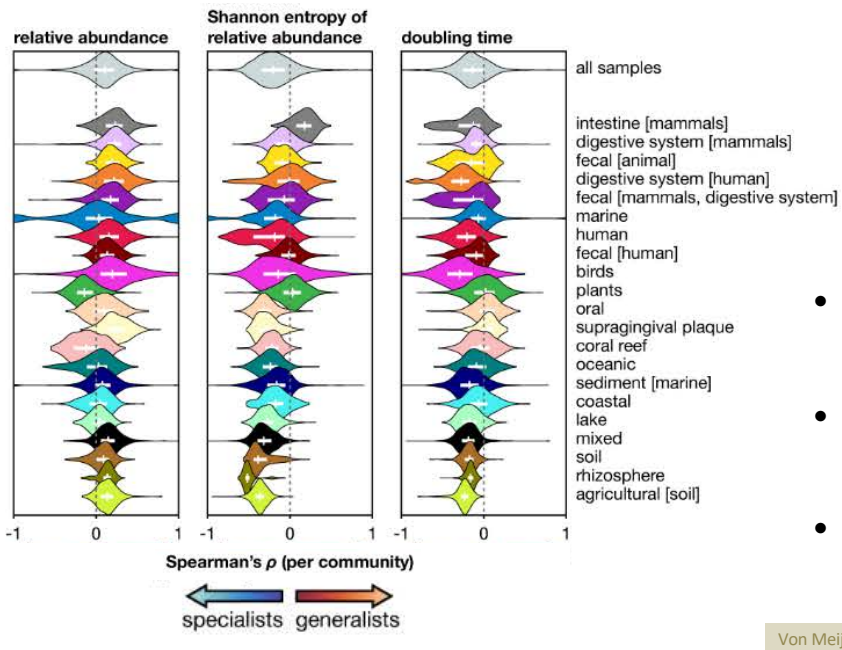
17

Questions

6. What's the difference?
   a) How do generalists become so successful?
   b) How do specialists prevent getting outcompeted?

18

Generalists: Jacks of all traits?

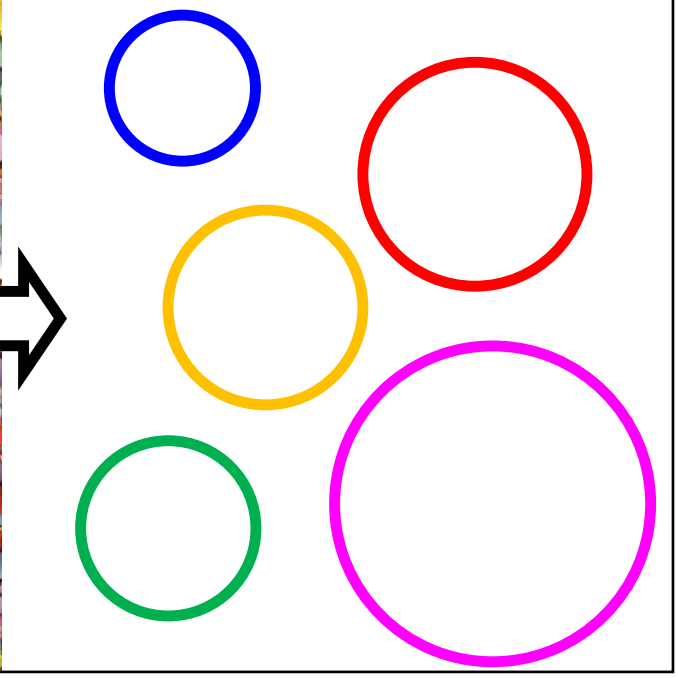


- Generalists often locally outcompete specialists
- Generalists are variable across communities
- Generalists are fast growing opportunists

Von Meijenfeldt, Hogeweg & Dutilh, *Nature Ecology & Evolution* 2023

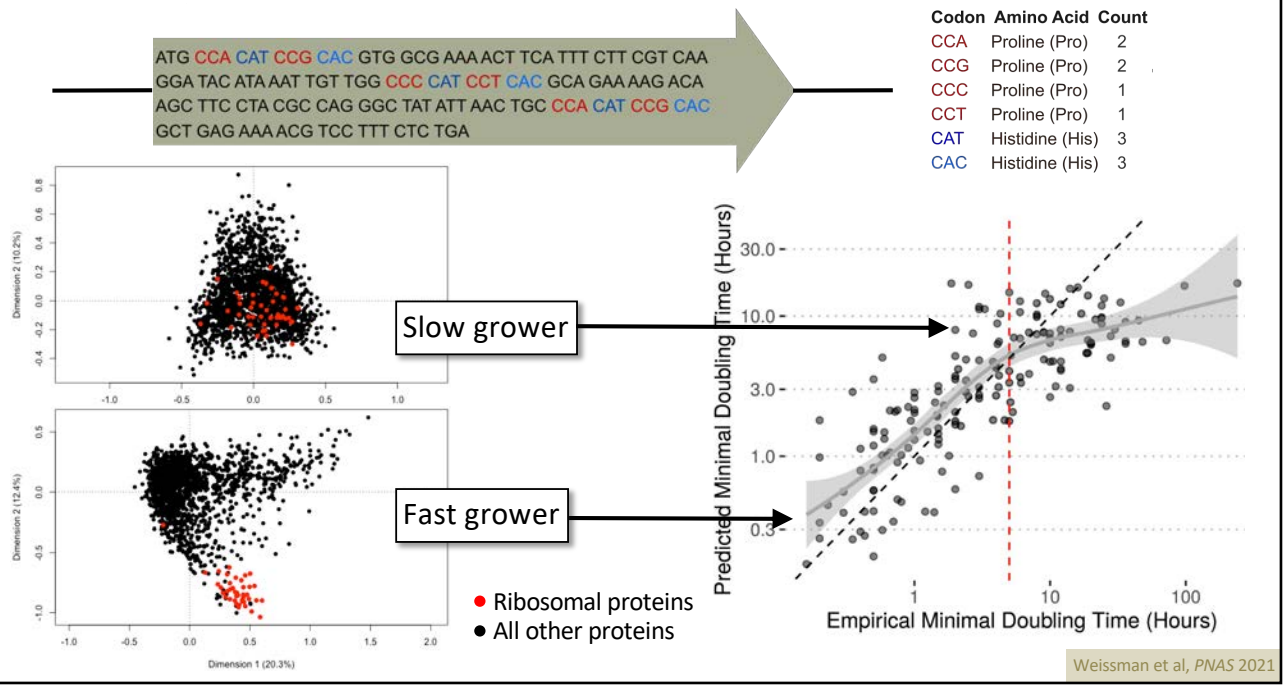
19

Metagenomics 2.0: genomes from metagenomes



20

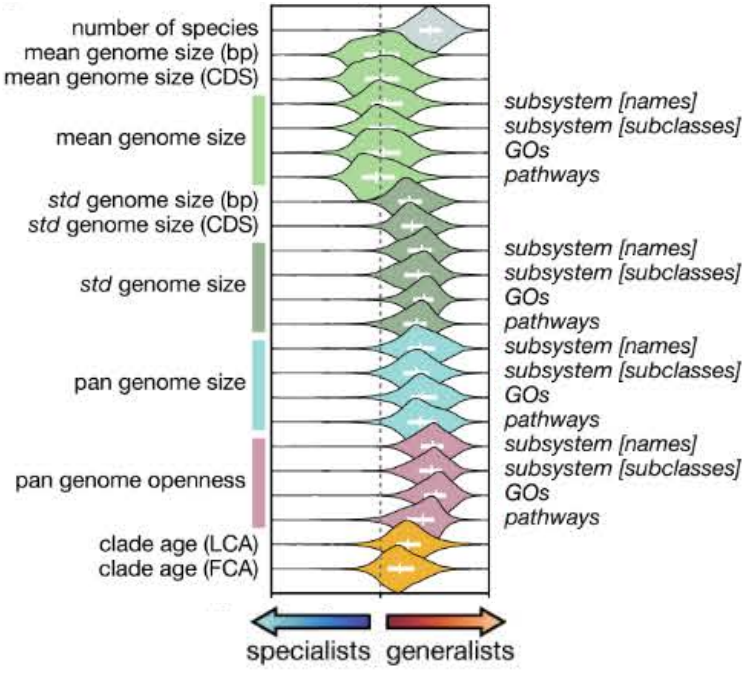
Growth rate predictions from codon usage bias



Weissman et al, *PNAS* 2021

21

Overall, generalists do not have larger genomes than specialists



- Generalists and specialists have similar genome sizes
- Generalist genera vary in size and gene content
  - Larger pan genome
  - Open pan genome
- Specialist species have more similar genomes than generalists
- Generalists are older than specialists

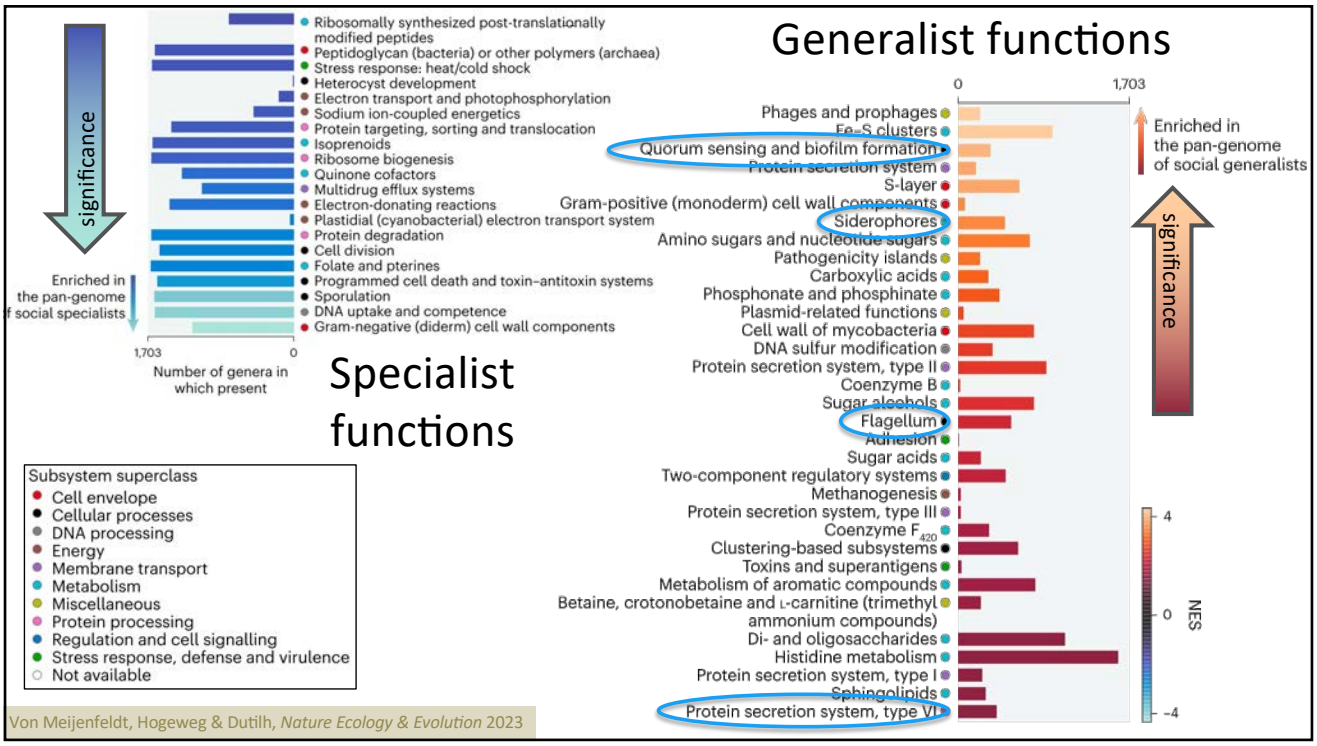
Von Meijenfeldt, Hogeweg & Dutilh, *Nature Ecology & Evolution* 2023

22

Assignments

7. Get yourself some microbiome datasets
   a) Either your own data, or search for your favorite keyword in MGnify
      - https://www.ebi.ac.uk/metagenomics
   b) Convert them to BIOM format
      - http://biom-format.org/documentation/biom_format.html
8. Upload them to the Microverse Mapper
   a) Do your samples fall where you expect them on the map?
   b) Select the 100 closest other samples. From how many different studies are they derived?
   c) Can you find the associated publications?

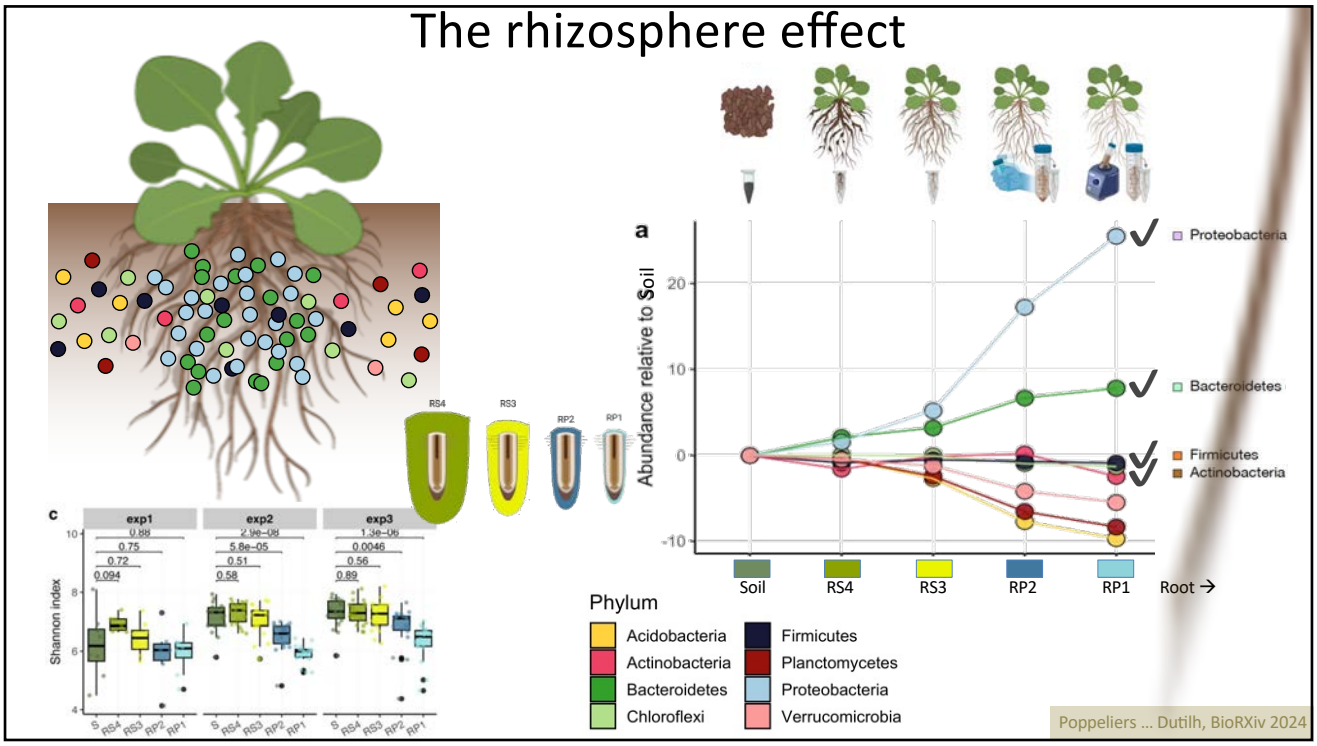
23



Von Meijenfeldt, Hogeweg & Dutilh, *Nature Ecology & Evolution* 2023

24

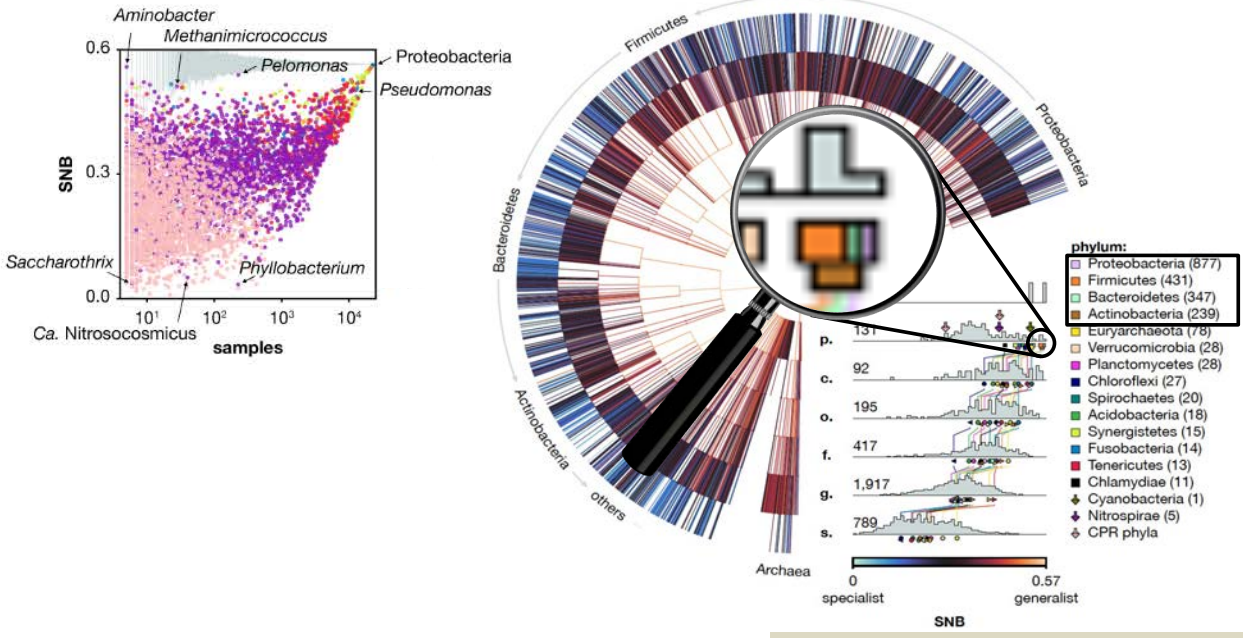
The rhizosphere effect



Poppeliers … Dutilh, BioRXiv 2024

25

SNB at different taxonomic ranks

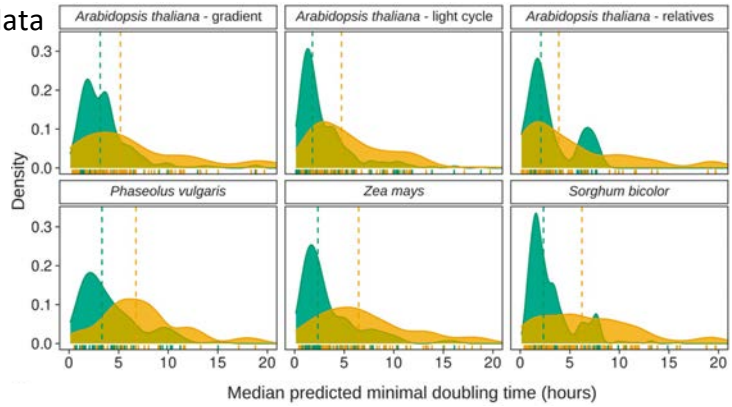


Von Meijenfeldt, Hogeweg & Dutilh, *Nature Ecology & Evolution* 2023

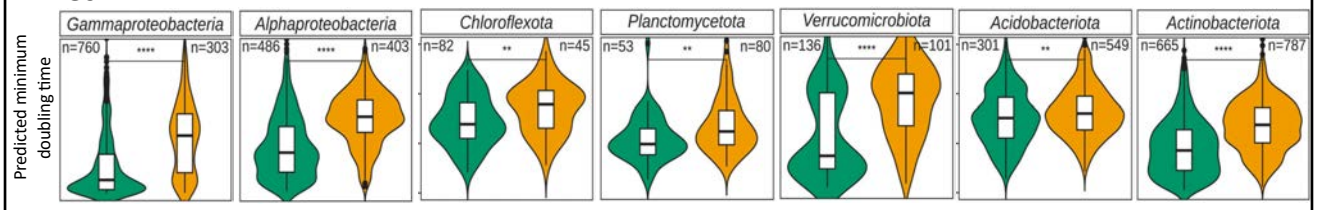
26

In different plants and bacteria

Microbiome data



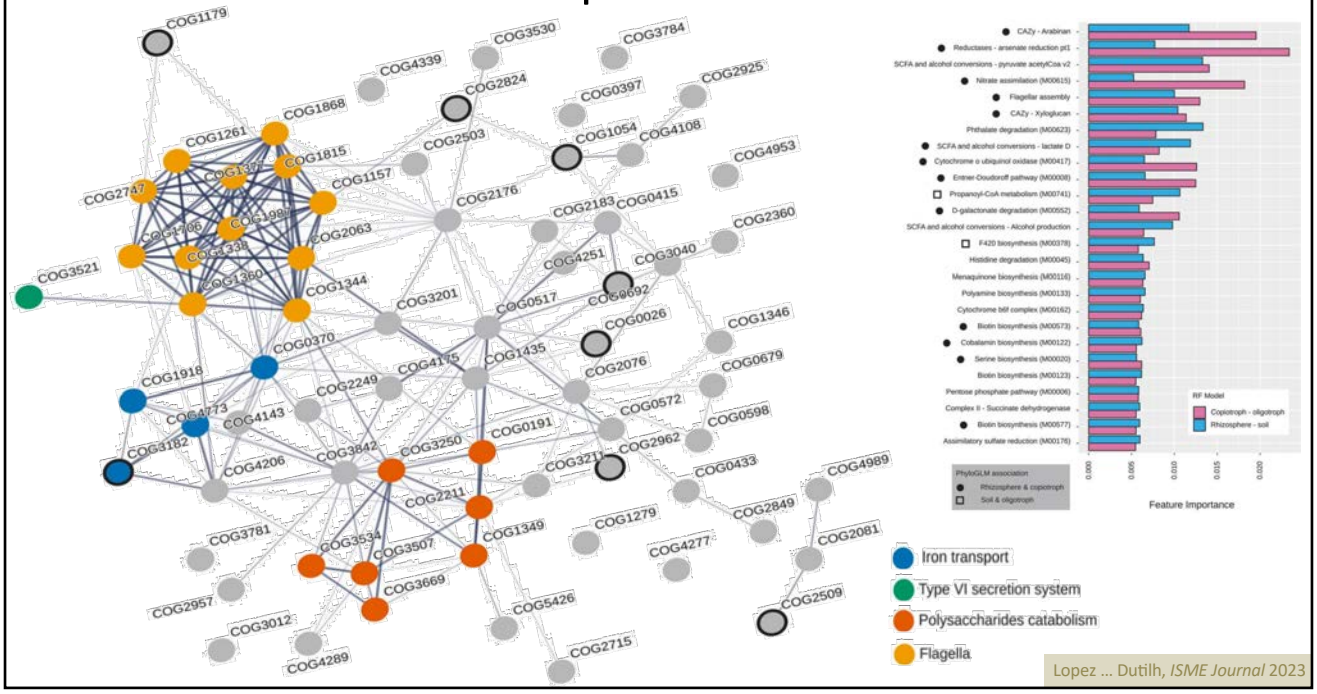
MAGs



Lopez … Dutilh, *ISME Journal* 2023

27

Functions of rhizosphere-associated bacteria



Lopez … Dutilh, *ISME Journal* 2023

28

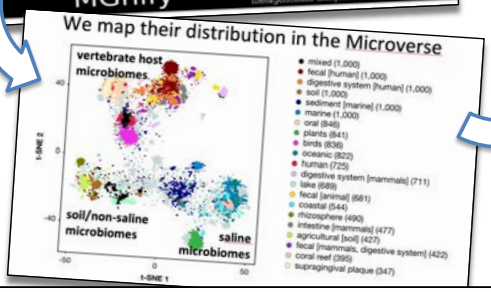
Conclusions

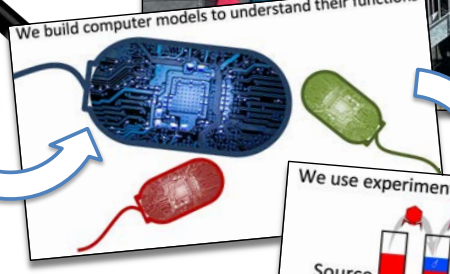
1. Bacteria are differentially distributed across the Microverse
2. Wide and narrow niche breadths reflect different eco-evolutionary strategies
   - Generalists are fast growing opportunists with flexible genomes
   - Specialists are lower in abundance but ecologically and evolutionarily stable
3. Rich nutrient sources such as the plant root attract fast growers with competitive functions

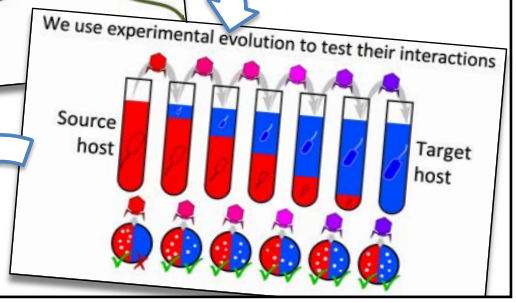
29

Research lines of the

Viral Ecology and Omics Group







For projects email b.e.dutilh@uni-jena.de
Experience with Python/R/bash is recommended

30

Thank you

Funded by the German Government through the Alexander von Humboldt Foundation.

Funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme grant agreement No. 865694: DiversiPHI.

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2051 – Project-ID 390713860.

Funded by the Netherlands Organization for Scientific Research (NWO) Vidi grant 864.14.004.

@ bedutilh.bsky.social

Utrecht Bioinformatics Center

b.e.dutilh@uni-jena.de

BAL4NCE OF THE MICROVERSE

31