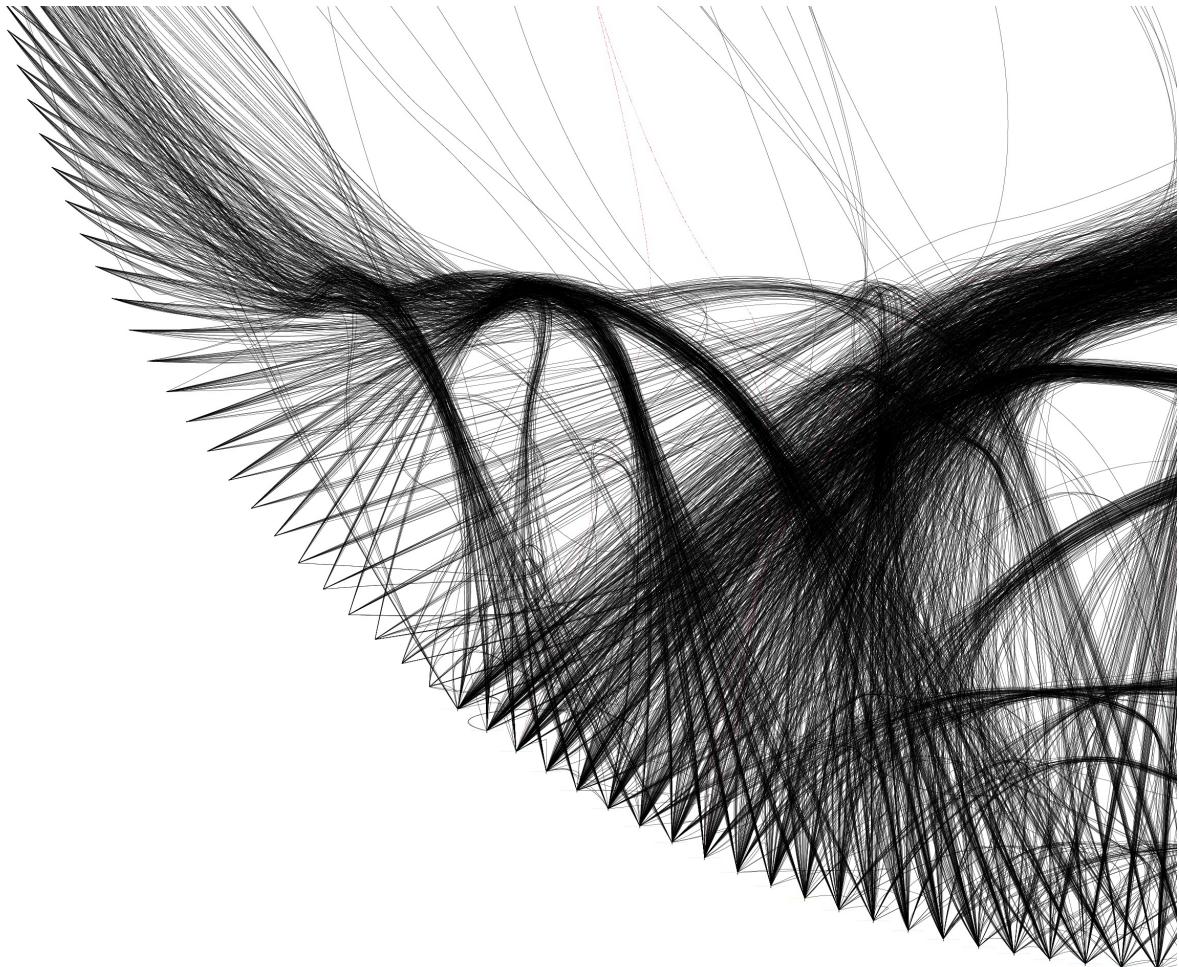


# NETWORKS

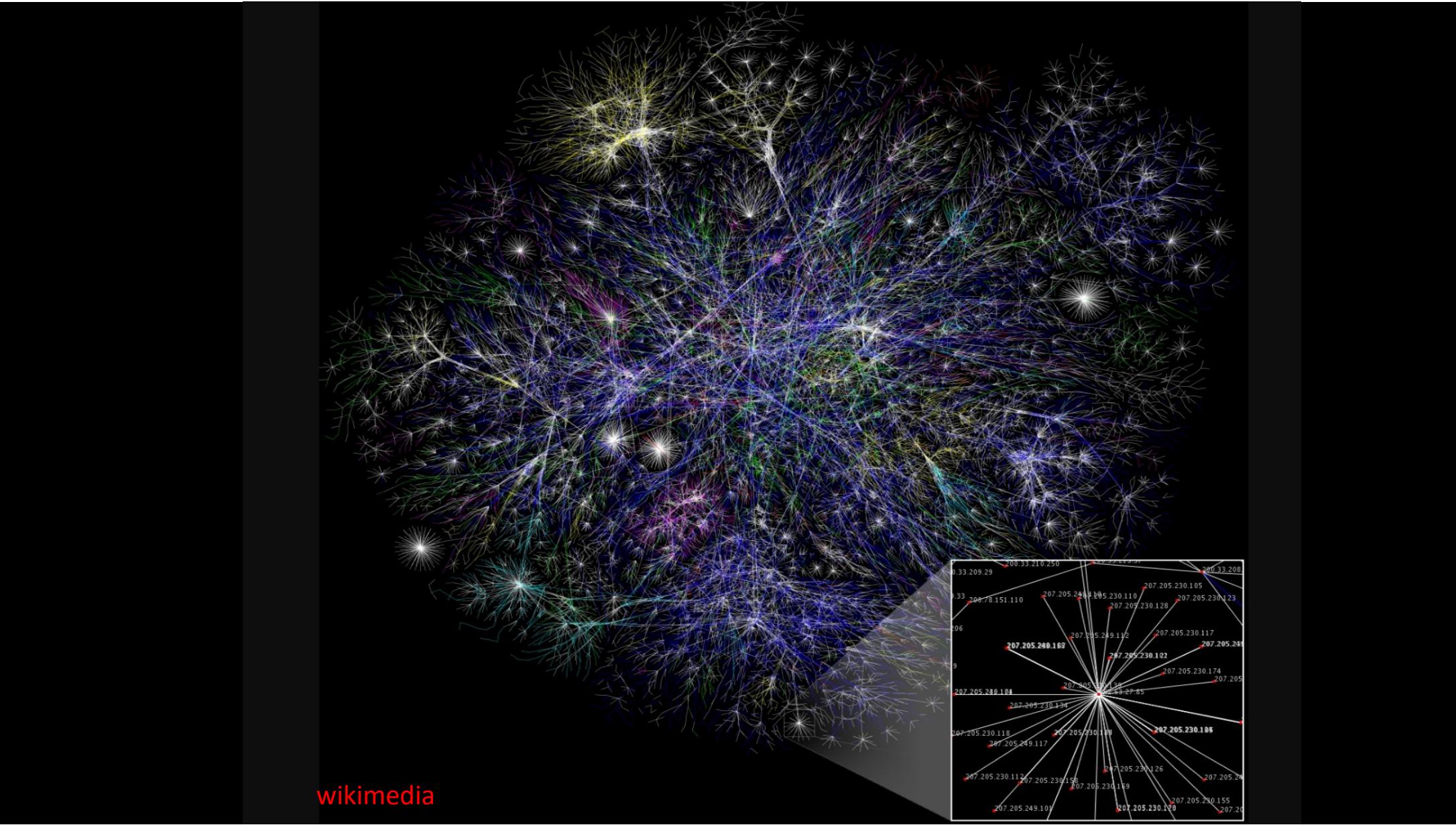
Anders K. Krabberød, University of Oslo

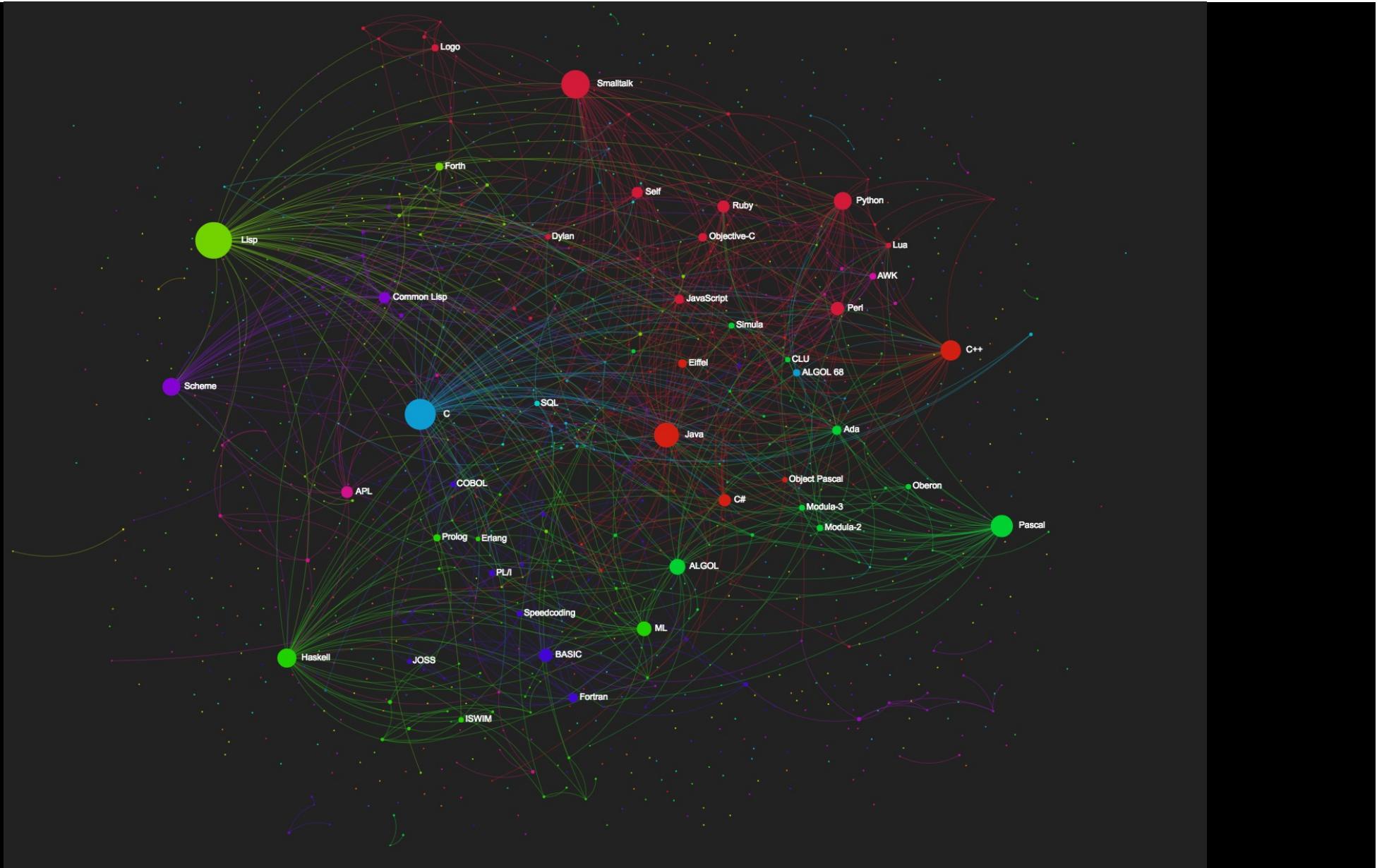


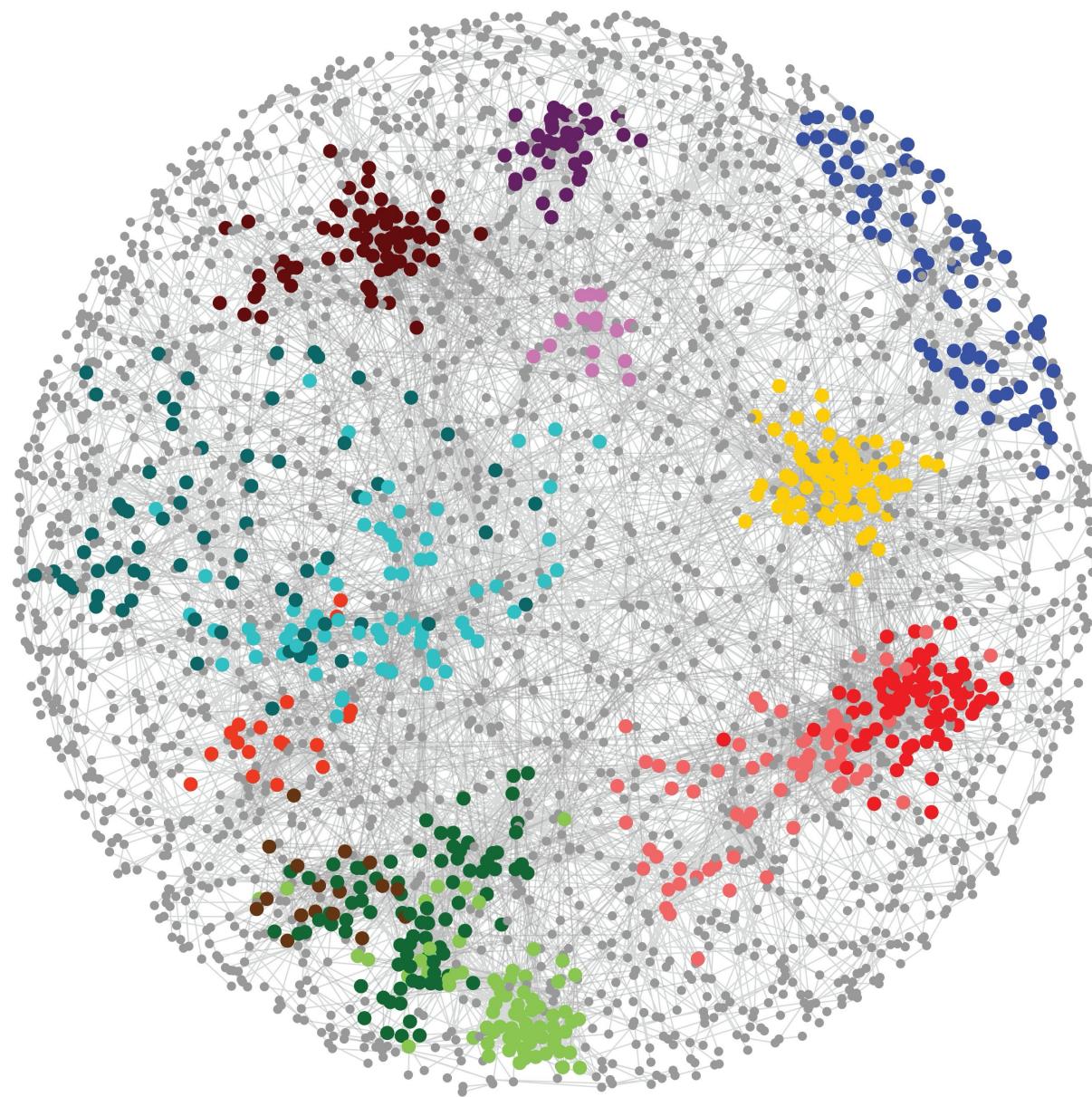
- BackGround

## Why Networks?

- Networks are ideal for exploring relationships and associations in data
- Network topology reveals patterns of interaction and potential ecological relationships between microbial taxa
- Visualize and analyze high-dimensional, complex microbial datasets
- examples (of non-microbial networks):
  - *“Six degrees of separation”*
  - Kevin Bacon game
  - Internet server architecture

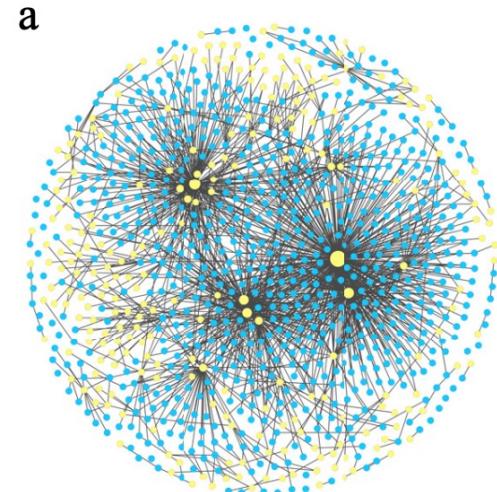






# Ecological association networks

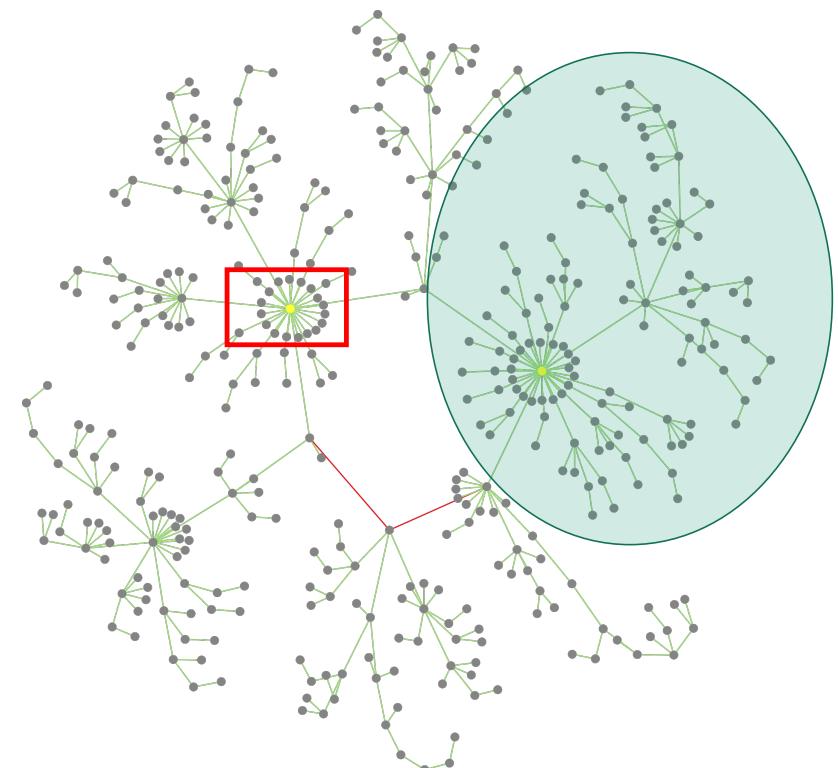
- Co-occurrence of OTUs/ASVs can suggest:
  - Potential ecological interactions (e.g. , host–symbiont, predator–prey)
  - Shared environmental responses
  - Similar niche preferences
- Identify key taxa (hubs) and groups of co-associated species (modules)



Williams et al. 2020

## Network Features – Some terminology

- ▶ **Hubs** are highly connected nodes, often key species in the network.
- ▶ **Modules**: sets of nodes that are more linked to each other than the rest of the network.



Krabberød et al. (2017)



# Spatial vs Temporal Network Inference

## **Spatial:**

- Identify OTUs that co-occur across similar habitats
- Enables analysis over large geographic scales
- Misses temporal dynamics and community turnover

## **Temporal:**

- Track community changes at a single location over time
- Resolution depends on the sampling frequency
- Rapid turnover: days to weeks
- Seasonal shifts: months to years
- Captures dynamic interactions and succession



# Time series of BBMO

- 10 years of monthly sampling (120 samples)
- Blanes Bay Marine Observatory
- DNA from two-size fractions
  - *Pico* (0.2-3um)
  - *nano* (3-20um)
- Microbial eukaryotes and bacteria (V4)
- Illumina data processed with DADA2
- 18 environmental parameters

Krabberød et al. *Environmental Microbiome* (2022) 17:22  
<https://doi.org/10.1186/s40793-022-00417-1>

Environmental Microbiome

RESEARCH ARTICLE

Open Access

Long-term patterns of an interconnected core marine microbiota

Anders K. Krabberød<sup>1\*</sup>, Ina M. Deutschmann<sup>2</sup>, Marit F. M. Bjørbaekmo<sup>1</sup>, Vanessa Balagué<sup>2</sup>, Caterina R. Giner<sup>2</sup>, Isabel Ferrera<sup>2,3</sup>, Esther Garcés<sup>2</sup>, Ramon Massana<sup>2</sup>, Josep M. Gasol<sup>2,4</sup> and Ramiro Logares<sup>1,2\*</sup> 

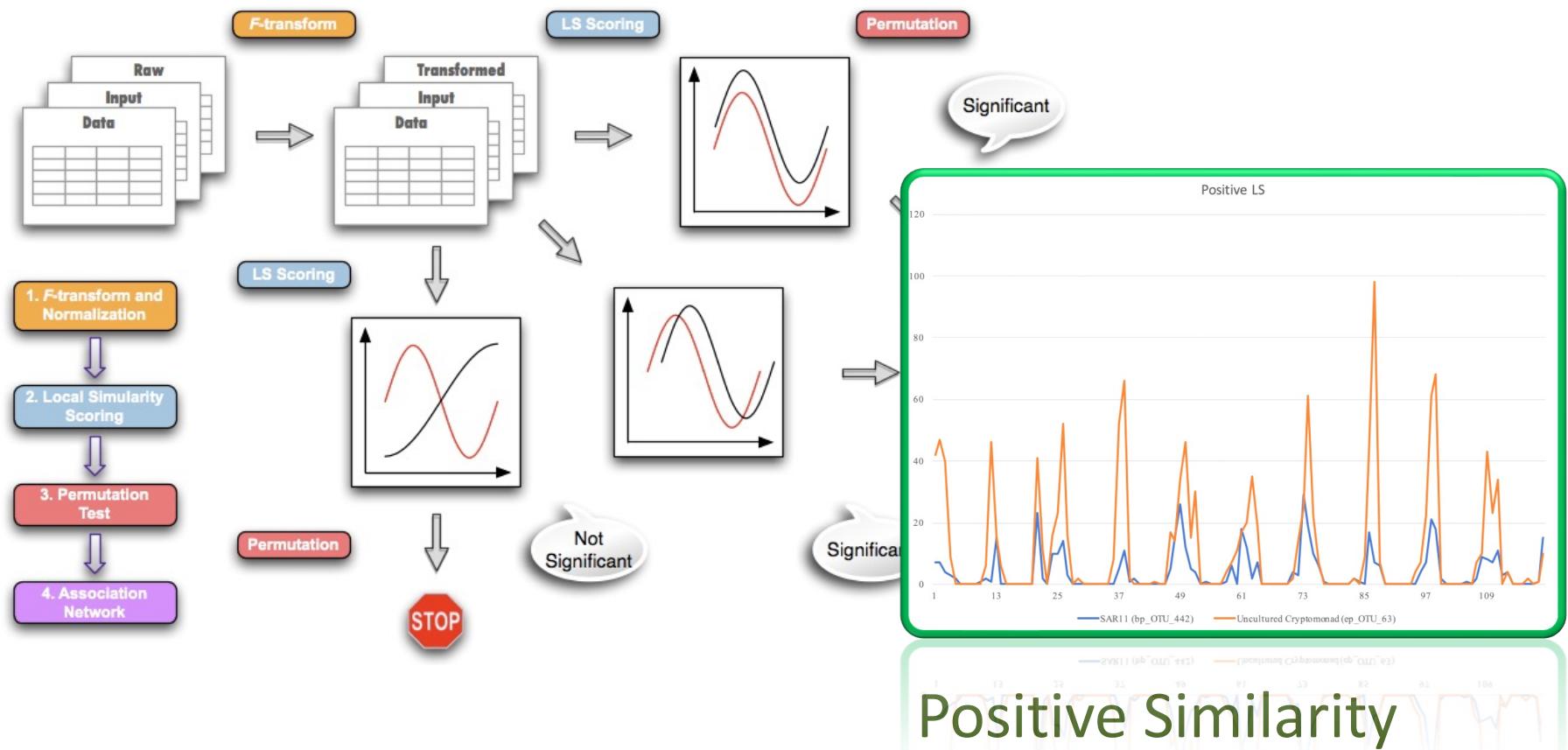


## Similartity analysis

- **ELSA** – Extended Local Similarity Analysis
  - Ruan et al (2006) ;Xia et al (2011)
- Designed for time series
- Detects (local) similarity between potentially shifted or lagged time series
- It can be used to calculate lagged response where the maximum allowed time delays/shifts can be specified.
  - However, we didn't use this feature since we had samples for each month, which was too long to give a meaningful lagged response.

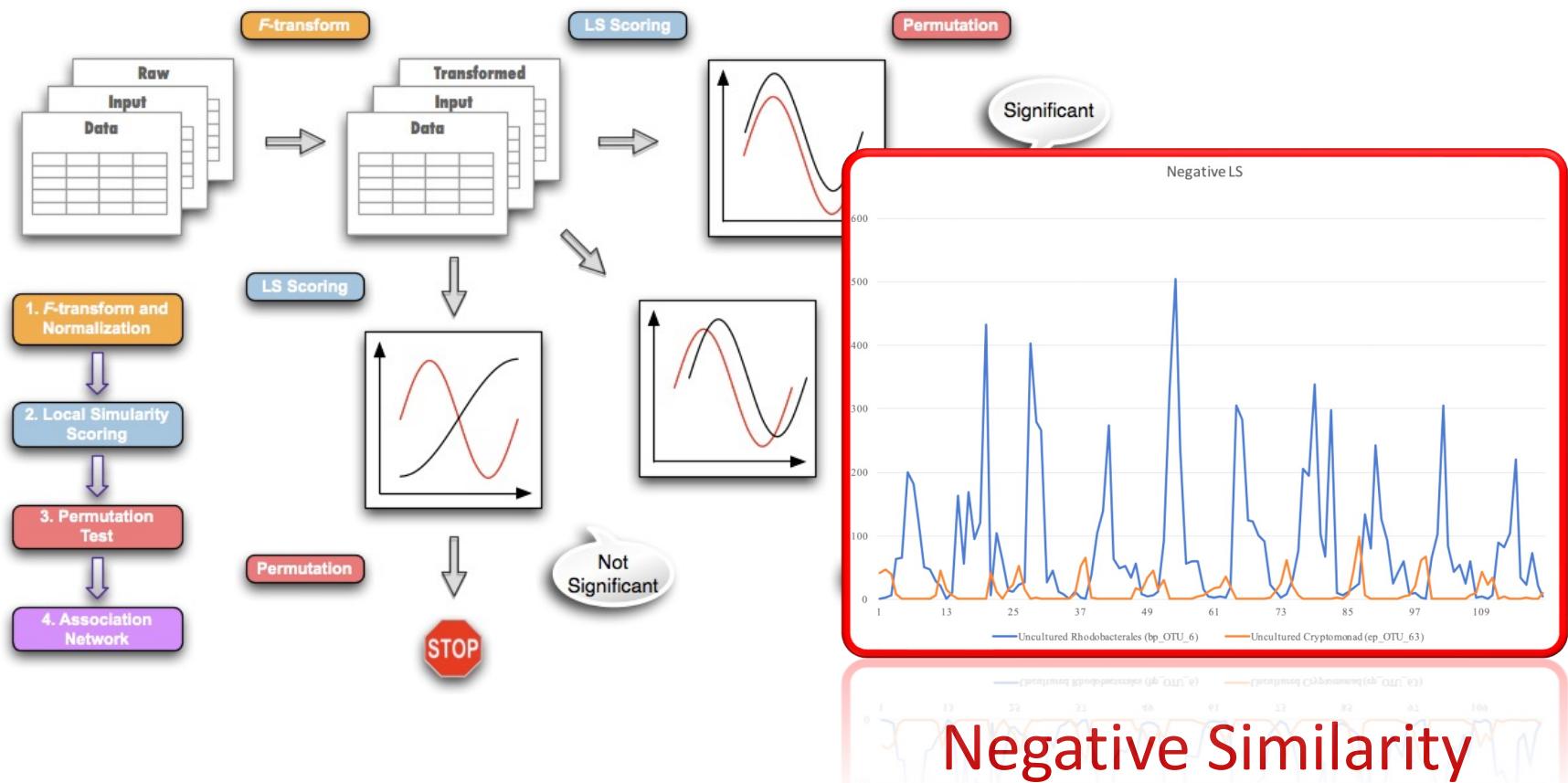
# Local Similarity Analysis

Ruan et. al 2006, Xia et al. 2011, 2013



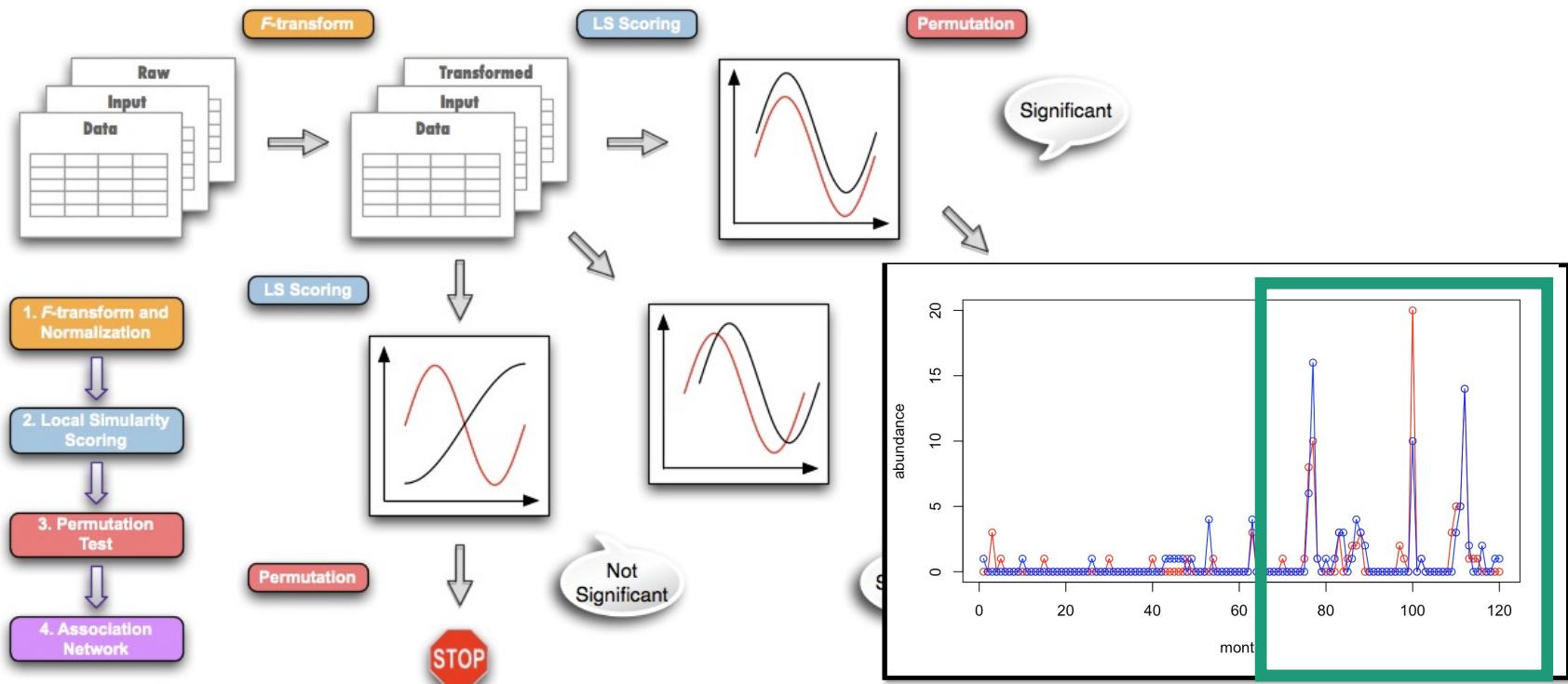
# Local Similarity Analysis

Ruan et. al 2006, Xia et al. 2011, 2013



# Local Similarity Analysis

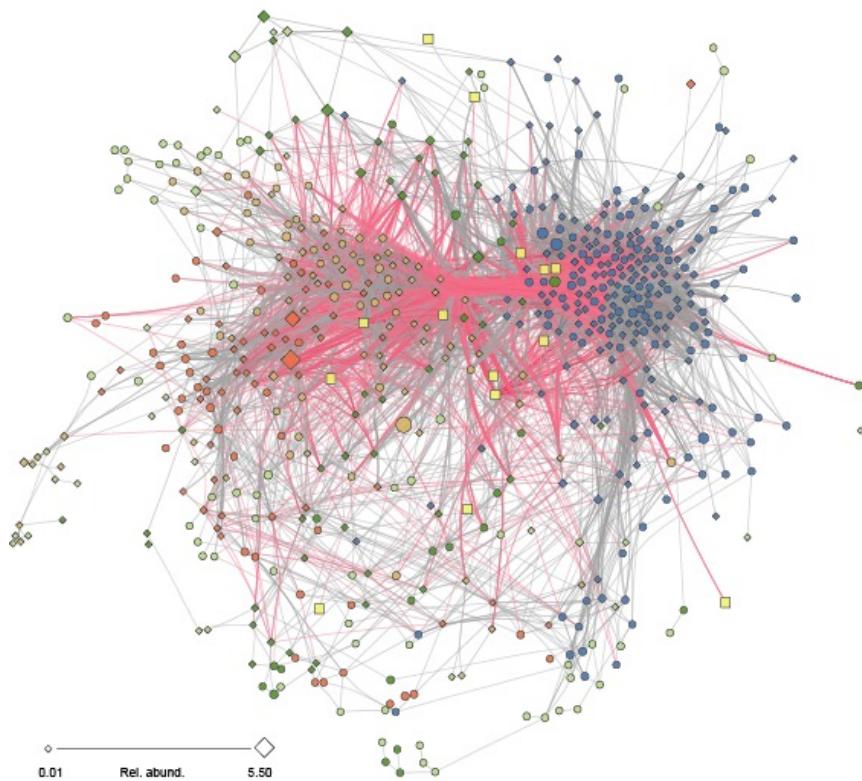
Ruan et. al 2006, Xia et al. 2011, 2013



Local similarity

## The resident microbiota

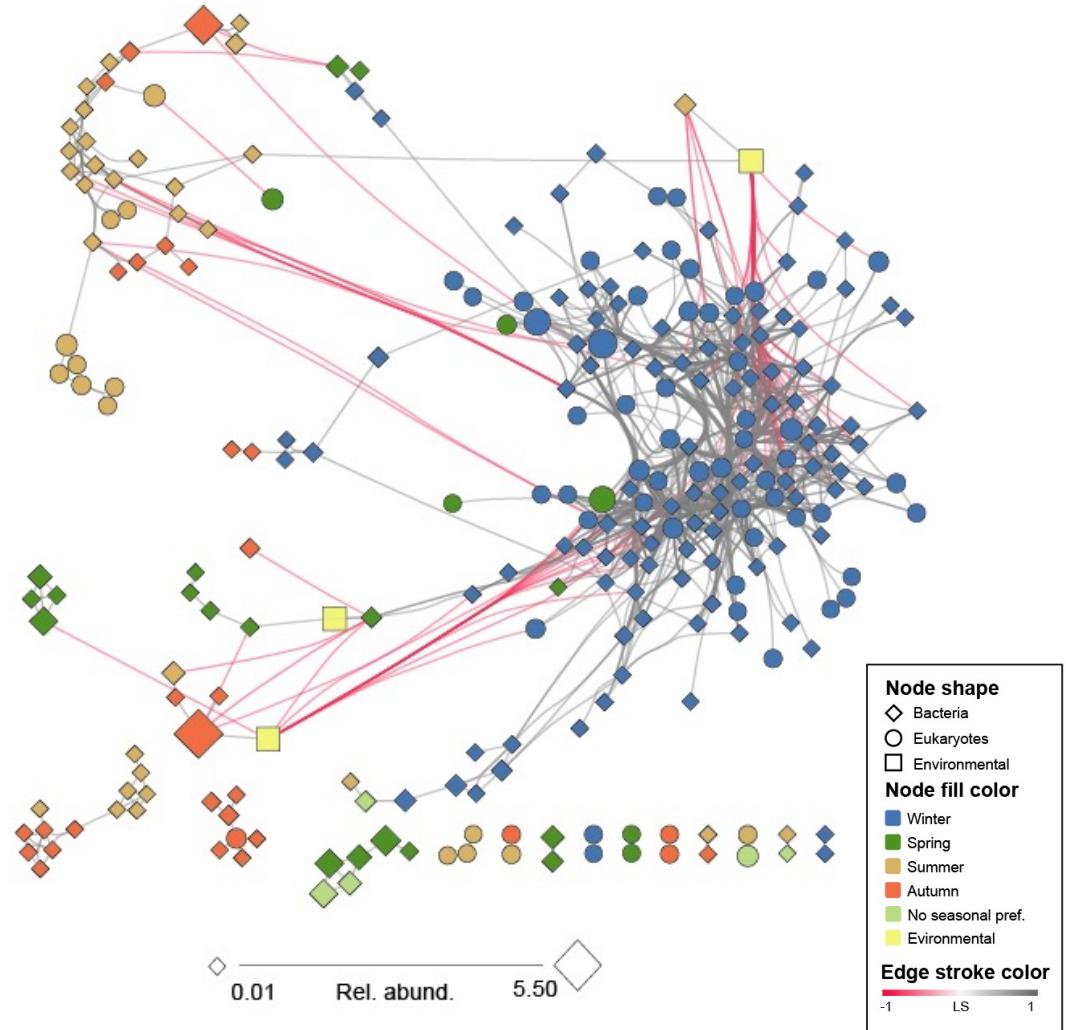
Species in more than 30% of the time



- **709 ASVs and 14 environmental parameters**
- **eLSA Without time lags**
- After removing non-significant edges ( $P&Q<0.05$ )
  - 36631 associations (edges)
  - 29663 positive (81%) and
  - 6968 negative (19%)

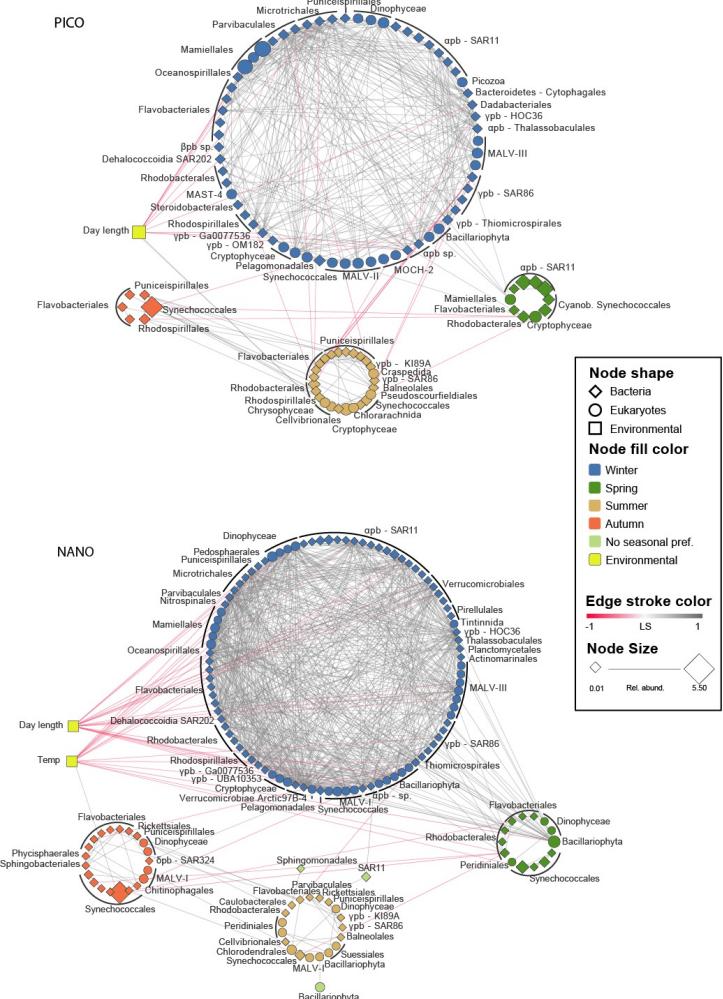
## Core Interactome

- ▶ Filtered and kept the strong associations:
- ▶ Local similarity score  $|LS| > 0.7$ , Spearman correlation  $|\rho| > 0.7$ , Bonferroni adjusted  $p < 0.0001$
- ▶ There are more bacteria species (67.9%) than eukaryotic species
- ▶ Almost all associations are positive (95 %)
- ▶ One part is highly interconnected These are ASVs that are more common during the winter season (as calculated by IndVal).



# Seasonality

- ▶ Winter community has the highest proportion of taxa (ASVs) in both size fractions
- ▶ Winter/spring communities has a higher number of association and are tighter knit (clustering coefficient 0.452), than summer/autumn (cc 0.305)
- ▶ There is overlapping associations between winter and summer communities.
- ▶ Summer and autumn are much looser structured



# Interactions?

- ▶ Some of the associations (edges) are due to two organisms reacting to the same environmental variable.
- ▶ What about direct biotic interactions (symbiosis, parasitism, predation) ?
- ▶ We spent some effort removing edges in the network that are “indirect” -> caused by two OTUs that co-vary with the same environmental variable
- ▶ EnDED, Deutschmann (2021)

Deutschmann et al. *Microbiome* (2021) 9:232  
<https://doi.org/10.1186/s40168-021-01141-7>

Microbiome

METHODOLOGY

Open Access

Disentangling environmental effects in microbial association networks



Ina Maria Deutschmann<sup>1\*</sup>, Gipsi Lima-Mendez<sup>2</sup>, Anders K. Krabberød<sup>3</sup>, Jeroen Raes<sup>4,5</sup>, Sergio M. Vallina<sup>6</sup>, Karoline Faust<sup>5†</sup> and Ramiro Logares<sup>1††</sup>

## Validation

- ▶ Correlation does not imply causation
- ▶ Need tools (and/or experiments) to validate associations as interaction
- ▶ We have built database of protist interactions



# PIDA – Protist interaction database

- We assembled a database based on known interactions from existing literature and integrated it with other databases
- PIDA: Protist interaction database
  - Github: <https://github.com/ramalok/PIDA>
  - The ISME (2019) doi.org/10.1038/s41396-019-0542-5
- over 2500 entries from 537 publications
- ca 700 species of host and predators
- ca 400 species of symbionts and prey

[nature](#) > [the isme journal](#) > [articles](#) > [article](#)

Article | [Open Access](#) | Published: 04 November 2019

## The planktonic protist interactome: where do we stand after a century of research?

Marit F. Markussen Bjørbækmo, [Andreas Evenstad](#), [Line Lieblein Røsæg](#), [Anders K. Krabberød](#)✉ & [Ramiro Logares](#)✉

*The ISME Journal* 14, 544–559 (2020) | [Cite this article](#)

6465 Accesses | 67 Citations | 29 Altmetric | [Metrics](#)



Marit Bjorbækmo



Andreas Evenstad



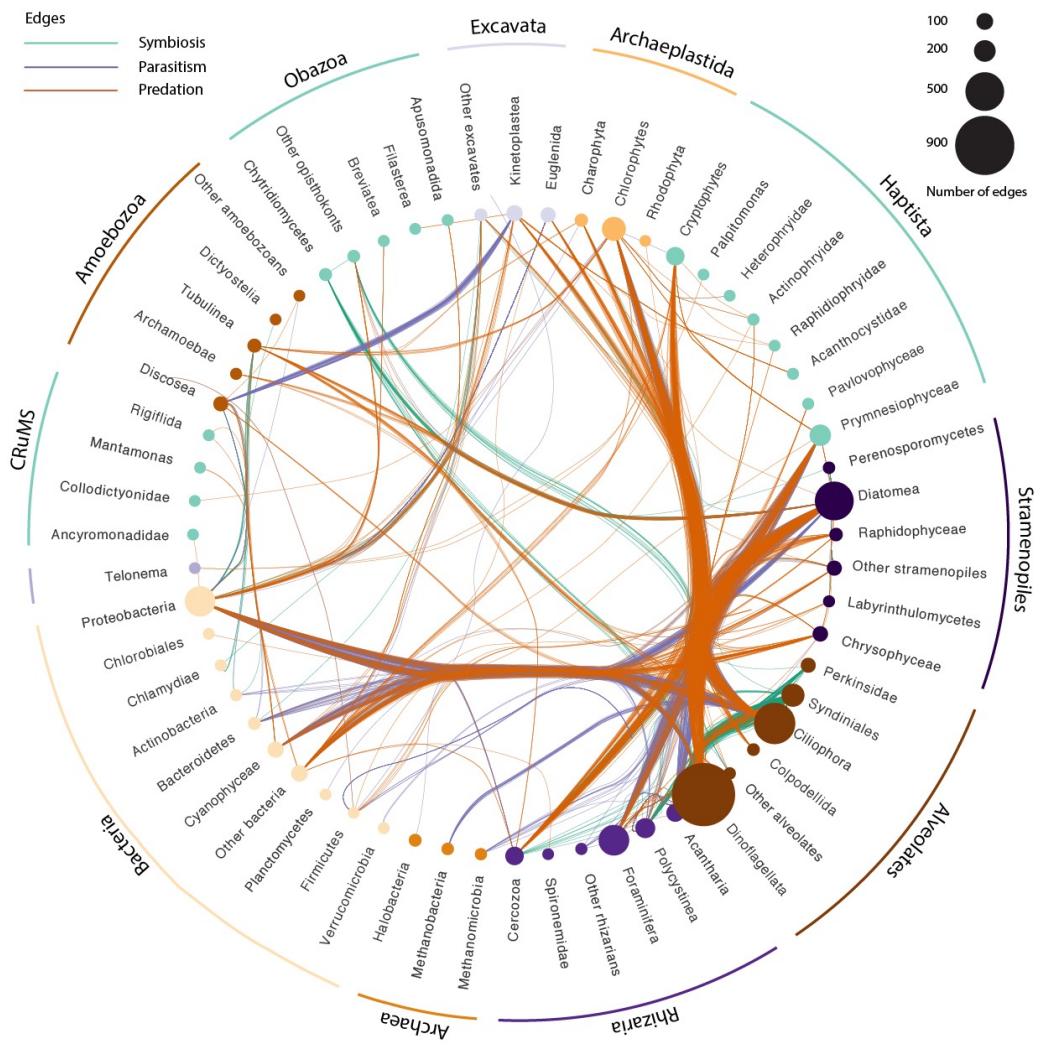
Ramiro Logares



Line Røsæg

# PIDA - summary

- The sizes of the nodes are proportional to the number of registered interactions.
- Diatoms, dinoflagellates, and ciliates have the highest number of entries in the database.
- Annotate the network from Blanes Bay with interaction information from PIDA
- But, there were few matches...



# Soil Depth Study

- Sunil Mundra, Håvard Kauserud et al.

- Purpose:

- Investigate protist, fungus and bacteria at different depths in soil
- Soil depth represents a strong physiochemical gradient that greatly affects
- Test whether depth-dependent variation in soil chemistry affects the distribution and co-occurrence patterns of microbial communities.



FEMS Microbiology Ecology, 97, 2021, fiab022  
doi: 10.1093/femsec/fiab022  
Advance Access Publication Date: 6 February 2021  
Research Article

Downloaded from https://academic.oup.com/femsec/article/97/



UiO :

# Soil Depth Study

- Four depths at 60 sites:
  - Litter-humus-fermentation (LHF), Mineral 1, M2, M3, and M4
- Illumina data from three markers
  - 16S V9 for bacteria
  - 18 V4 for eukaryotes
  - ITS2 for fungi
  - Vsearch, 97% for 16S and ITS, 98% for 18S
- 1540 bacterial OTUs
- 4388 fungal (3461 ITS-based; 927 18S-based)
- 2025 micro-eukaryotic



FEMS Microbiology Ecology, 97, 2021, fiab022  
doi: 10.1093/femsec/fiab022  
Advance Access Publication Date: 6 February 2021  
Research Article

Downloaded from https://academic.oup.com/femsec/article/97/

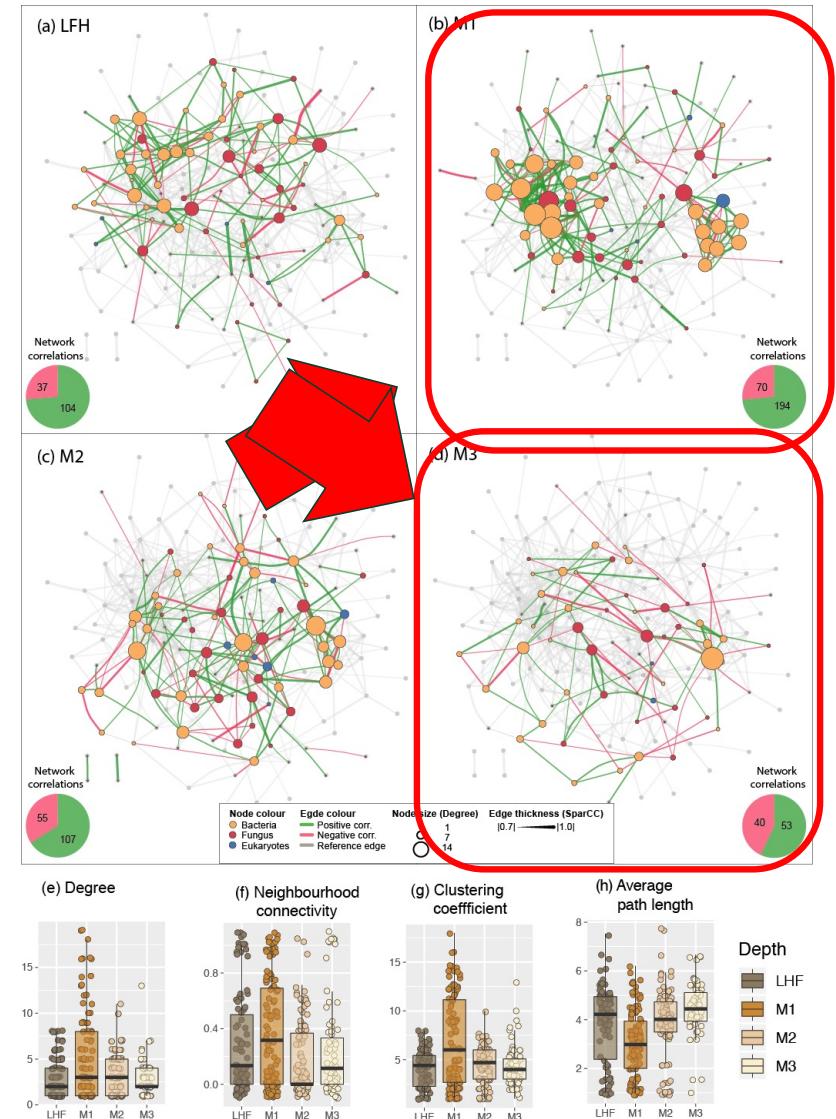


# SparCC

- A network algorithm for inferring correlation in sparsely composed data
- SparCC (Friedman and Alm, 2012) infers correlations among OTUs by utilizing log-ratio transformation to eliminate the effect of the total number of read counts, while imposing sparsity of correlations among otu.

# Soil Depth Study

- SparCC networks on four depths
  - LFH, M1, M2, M3, M4
- The network has the highest density for the M1 depth (i.e., the first mineral layer)
- At the deepest level, the proportion of negative occurrences was the highest



# Some Resource

- ▶ SparCC: a python module for computing correlations in compositional data (16S, metagenomics, etc'). <https://bitbucket.org/yonatanf/sparcc>
- ▶ iGraph package for R:
  - ▶ <https://www.r-bloggers.com/an-example-of-social-network-analysis-with-r-using-package-igraph/>
- ▶ The code repository for Schmidt et. al 2017 “A Family of Interaction-Adjusted Indices of Community Similarity” doi:10.1038/ismej.2016.139
  - ▶ [https://github.com/defleury/Schmidt\\_et\\_al\\_2016\\_community\\_similarity/](https://github.com/defleury/Schmidt_et_al_2016_community_similarity/)
- ▶ eLSA: <https://bitbucket.org/charade/elsa/wiki/Home>
- ▶ fast eLSA: <http://www.cmde.science.ubc.ca/hallam/fastLSA/>
- ▶ Tutorial for WGCNA:  
<https://labs.genetics.ucla.edu/horvath/CoexpressionNetwork/Rpackages/WGCNA/Tutorials/index.html>