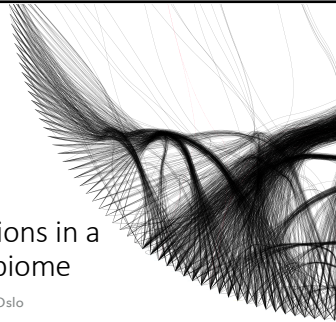


Long-term interactions in a core marine microbiome

Anders K. Krabberød, University of Oslo
a.k.krabberod@ibv.uio.no

In collaboration with ICM, Spain (Ramiro Longres et al.)



Outline

- Present a marine microbial ecosystem
- Look at the difference between userach clustering (OTUs) and DADA2 (ASVs) on amplicon data
- Give a brief introduction to ecological association networks
- Results for association networks for a long-term times series



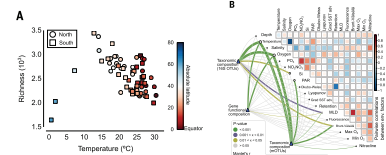
Why study marine microbes?

- Oceans covers 70% of the earth's surface
- It is populated by a huge amount of microbes (micro-eukaryotes and prokaryotes)
- Phytoplankton: 50% Primary production of the earth (Field et al. 1998)
- Diversity studies have given insight into important questions about biogeography, community compositions in different environments etc.



Abiotic factors

- Some abiotic parameters are important for microbe community dynamics, e.g. temperature
- But the abiotic factors does not give the full picture of the variation and seasonal dynamics in marine microbial ecosystems.
- How the microbes **interact** is poorly studied

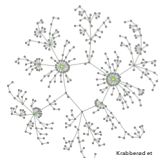


Sunigawa et al, Science (2015)



Ecological association networks

- Networks can help us find patterns behind small- and large-scale ecological and evolutionary processes
- Moving from from identification of parts (species classification, diversity, community composition, geographical distribution etc.) to connections of parts.



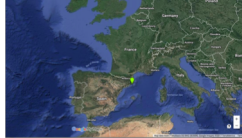
Categories of data for networks

- Spatial:
 - Can cover large geographical areas.
 - Miss out on seasonal variation and turnover in community composition
- Temporal
 - Smaller geographic coverage, with multiple data points for the same location. Allows following microbial communities, and their changes, in time.
 - Different scales
 - Turn-over for microbial communities can be in the span of days or weeks
 - In addition there are seasonal patterns of community assembly



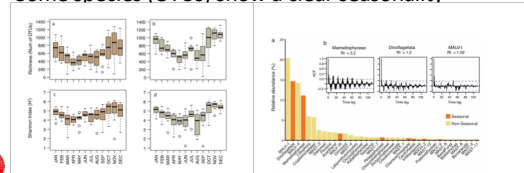
Blanes Bay Marine Observatory

- ▶ 10 years of monthly sampling (120 samples)
- ▶ DNA from two size fractions
 - ▶ *Pico* (0.2-3µm)
 - ▶ *nano* (3-20µm)
- ▶ Microbial eukaryotes and bacteria
- ▶ 18 physico-chemical parameters, chlorophyll and cell abundance



Diversity and seasonality

- ▶ Giner et al. (2018, Molecular Ecology)
- ▶ There is a seasonal pattern in the eukaryotic community
- ▶ Some species (OTUs) show a clear seasonality



Amplicon data

- ▶ Plankton was filtered according to size:
 - ▶ Picoplankton (<3 µm) and nanoplankton (3-20 µm)
- ▶ Illumina amplicons from 18S (V4) and 16S (V3)
 - ▶ 16 Millions sequences total
- ▶ Two types of clustering:
 - ▶ 1) clustering based on 99% sequence similarity with userach-> Operational Taxonomic Units (**OTU**)
 - ▶ 2) clustering without imposing arbitrary similarity thresholds -> Amplicon sequence variants (**ASV**)



ASV vs. OTU

- ▶ OTUs
 - ▶ Group sequences based on a predefined similarity value. We used 99% (did the same for bacteria and eukaryotes)
 - ▶ (eg Edgar et al. 2013)
- ▶ ASVs
 - ▶ With this method it is possible to correct for sequencing errors, and distinguish sequence variants differing by as little as one nucleotide
 - ▶ (e.g. Callahan et al. ISME 2017)



Microbiome

- ▶ Total number of OTUs and ASVs

	OTU	ASV
Total OTUs/ASVs	52 216	45 482
Eukaryotes (18S)	36 020	28 876
Prokaryotes (16S)	16 196	16 606

- ▶ OTUs and ASVs in >10% of the samples after subsampling

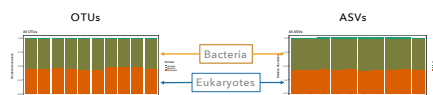
	OTU	ASV
Total OTUs/ASVs	5 637	2987
Eukaryotes (18S)	2 825	1366
Prokaryotes (16S)	2 812	1621

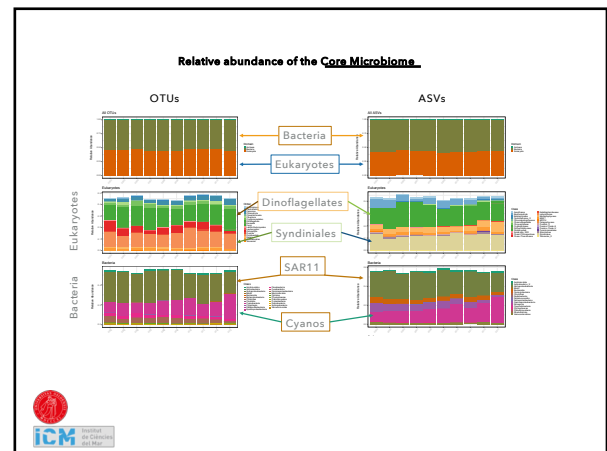
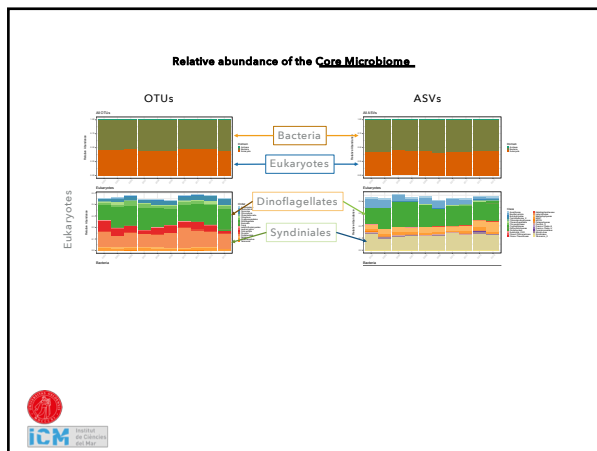
- ▶ “Core Microbiome”; OTUs and ASVs >30% of the samples

	OTU	ASV
Total OTUs/ASVs	1 953	744
Eukaryotes (18S)	888	356
Prokaryotes (16S)	1 067	388



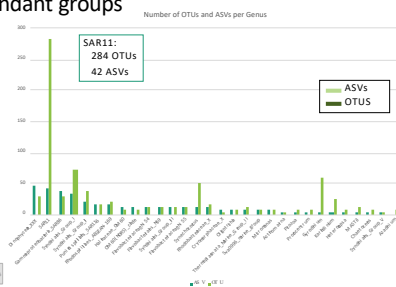
Relative abundance of the Core Microbiome





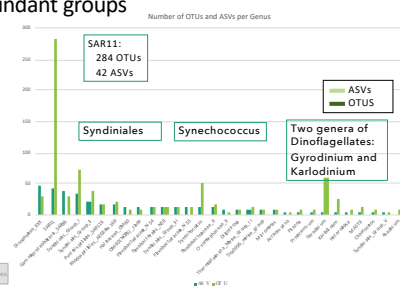
OTU vs ASV

- High number of OTUs compared to ASV for very abundant groups



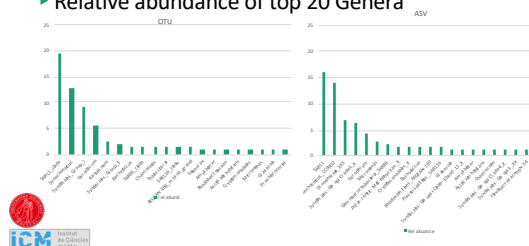
OTU vs ASV

- High number of OTUs compared to ASV for very abundant groups



Core microbiome

- We defined the core microbiome as those that are present at least 30% of the time
- Relative abundance of top 20 Genera



Ecological association networks

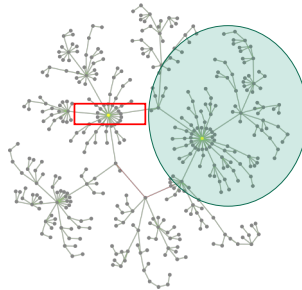
- Associations/co-occurrence between microbial OTUs/ASVs can reveal:
 - similar response to environmental variation
 - similar niche preference
 - direct interactions between organisms, e.g. host/symbiont, prey/predator
- key species (hubs) and assemblages of interacting species (modules)



NETWORKS

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)



NETWORKS

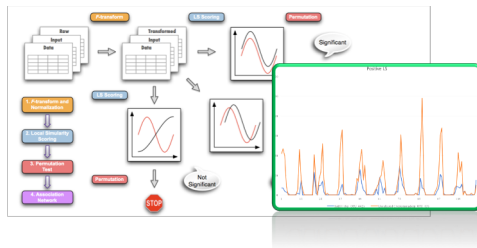
Similarity analysis

- **ELSA** – Extended Local Similarity Analysis
 - Ruan et al (2006) ;Xia et al (2011)
- Designed for time series
- Can incorporate time delays (the maximum allowed time delay has to be specified)
- Multiple co-occurrences and co-exclusions in the same community



NETWORKS

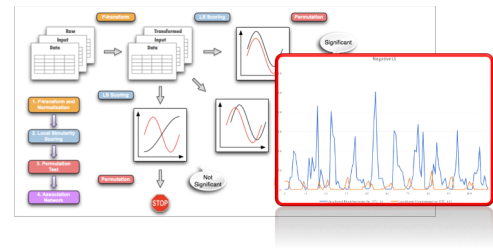
Local Similarity Analysis



NETWORKS

Local Similarity Analysis

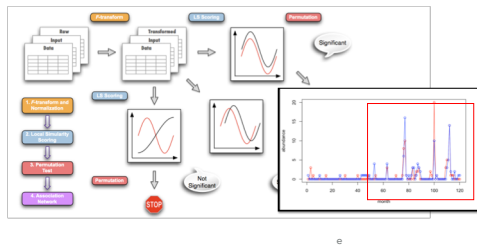
Ruan et. al 2006; Xia et al. 2011, 2013



NETWORKS

Local Similarity Analysis

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



NETWORKS

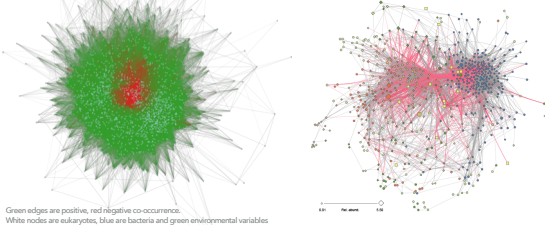
- 1955 OTUs and 15 environmental parameters
- After removing non-significant edges ($P \& Q < 0.05$)
 - 90566 associations (edges)
 - 86540 positive (95.5%) and
 - 4026 negative (4.5%)



Green edges are positive, red negative co-occurrence.
White nodes are eukaryotes, blue are bacteria and green environmental variables

NETWORKS

- 1955 OTUs and 15 environmental parameters
- After removing non-significant edges ($P \& Q < 0.05$)
 - 90566 associations (edges)
 - 86540 positive (95.5%) and
 - 4026 negative (4.5%)
- 837 ASV and 14 environmental parameters
- After removing non-significant edges ($P \& Q < 0.05$)
 - 36631 associations (edges)
 - 29663 positive (81%) and
 - 6968 negative (19%)

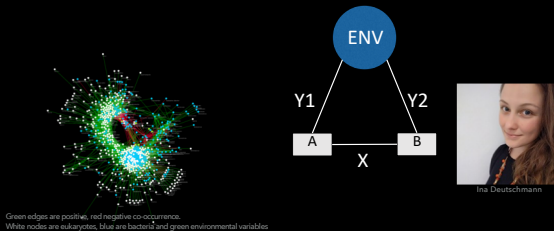


Interactions?

- 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 need methods to remove unwanted associations from the network.

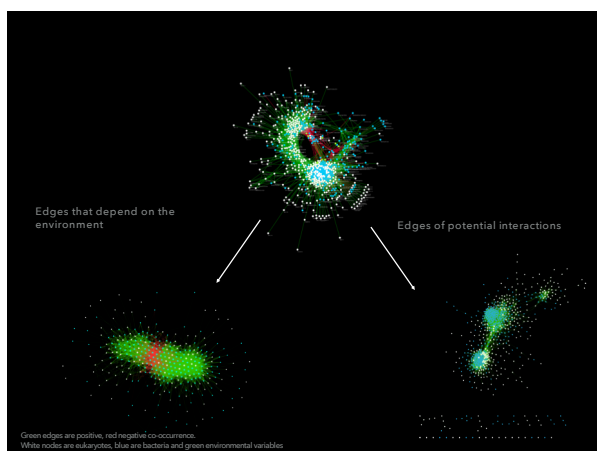
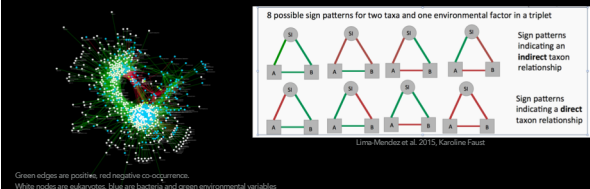
Pruning indirect edges

- Remove environment-driven co-occurrence
- Identification of triplets where the environmental factor is stronger than OTU-OTU interactions



Pruning indirect edges

- Remove environment-driven co-occurrence
- Identification of triplets where the environmental factor is stronger than OTU-OTU interactions



NETWORKS

Applying the program on the “core microbiome”

- 709 ASV and 14 environmental parameters ->
- After removing non-significant edges ($P \& Q < 0.05$)
 - 36631 edges
 - 29663 positive (81%) and 6968 negative (19%)



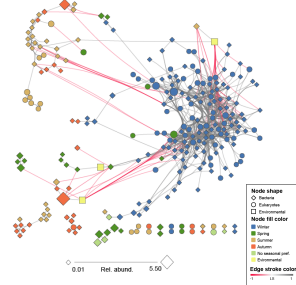
4960 edges were found to be indirect using a combination of 4 methods, and removed

Then we increased the cut-off for local similarity and Spearman correlation to >0.7 to look at the strongest association.

NETWORKS

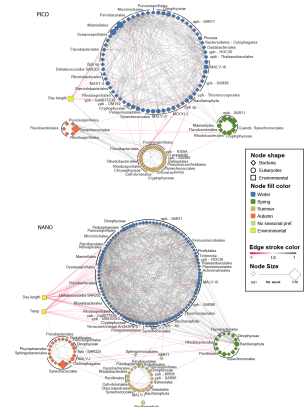
Core Interactome?

- ▶ The "Core interactome".
 - ▶ 182 bacteria ASVs,
 - ▶ 77 eukaryotic ASVs
- ▶ Almost all associations are positive (95.1 %)
- ▶ One part is highly interconnected. These are ASV 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



NETWORKS

Validation of associations

- ▶ Correlation does not imply causation
- ▶ Need tools to validate associations as interaction
- ▶ We have built database of confirmed protist interactions. I.e. from studies using cultures or microcopy

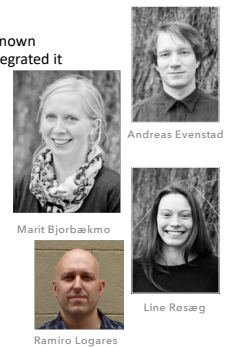


NETWORKS

PIDA – Protist interaction database

- ▶ We have assembled a database based on known interactions from existing literature and integrated it with other databases
- ▶ PIDA: Protist interaction database
 - ▶ Github: <https://github.com/ramirolog/PIDA>
 - ▶ BioRxiv: doi.org/10.1101/587352
- ▶ over 2500 entries from 537 publications
- ▶ ca 700 species of host and predators
- ▶ ca 400 species of symbionts and prey

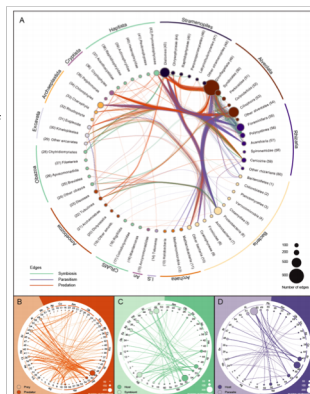
The planktonic protist interactome: emerging trends after a century of research
 Mark F. Roederer, Ramiro Logares*, Andreas Ekenstam* and Line Læhre Rosæg*, Indira N. Rørdam*, and Sørensen Sørensen*
 *Corresponding authors: ramiro.logares@uio.no, line.laehre@uio.no, andreas.ekenstam@uio.no, mark.f.roederer@uio.no, indira.n.rordam@uio.no
 The BioRxiv preprint server is not certified by peer review
 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted BioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



PIDA

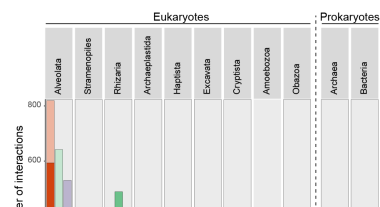
PIDA

- ▶ Size of the nodes are proportional to the number of registered interactions.
- ▶ Diatoms, dinoflagellates, and ciliates have the highest numbers of entries in the database.



PIDA

Interaction by Groups



PIDA

Next step

- ▶ Annotate the network from Blanes Bay with interaction information from PIDA
- ▶ Can help with:
 - ▶ Validate associations as true interactions
 - ▶ Generation of hypotheses on novel interactions



Conclusion

- Long time series and co-occurrence networks can help us understand important associations and interactions in microbial ecosystem
- We are trying to separate co-occurrence due to environmental parameters from biotic interactions.
- Databases can help to validate and identify interactions.

Acknowledgments

- | | |
|---|--|
| <ul style="list-style-type: none"> • <u>@University of Oslo:</u> • Marit Bjorbækmo • Line Røsæg • Russell Orr • Andreas Evenstad | <ul style="list-style-type: none"> • <u>@ICM CSIC:</u> • Ramiro Logares • Ina M. Deutschmann • Caterina R. Giner • Fran Latorre • Lidia Montiel • Ramon Massana • Ramunas Stepanauskas @ Bigelow |
|---|--|