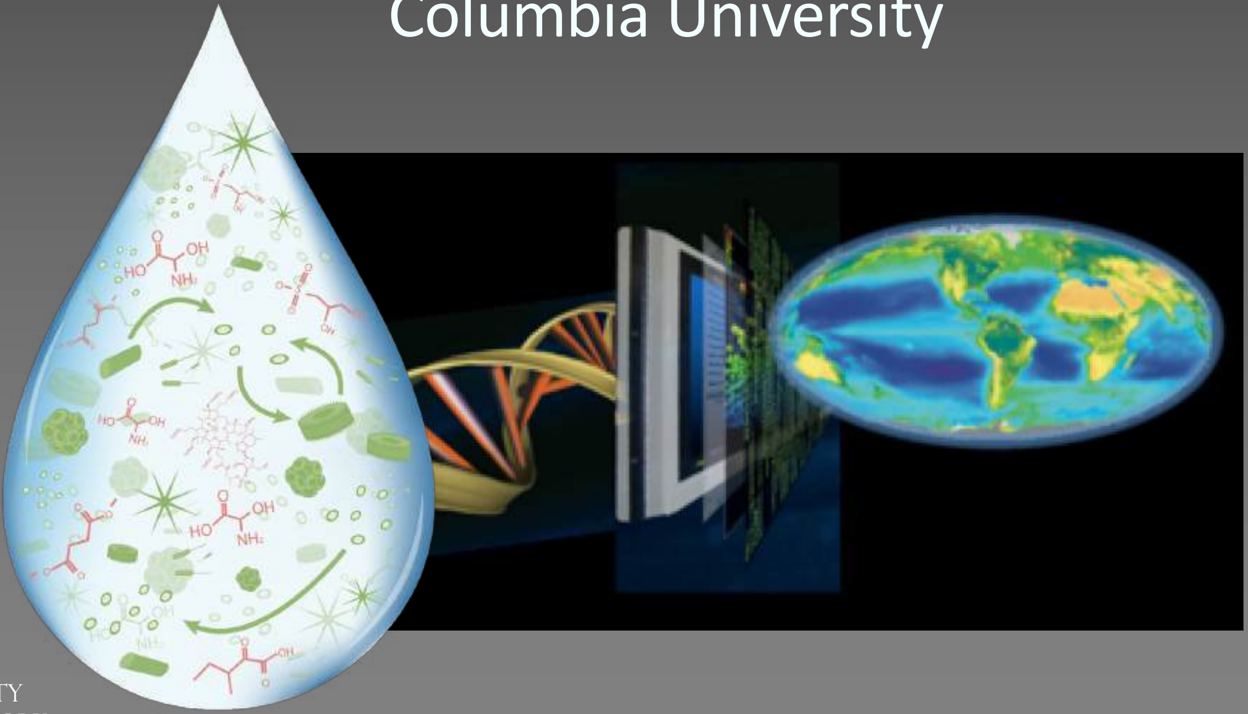
Ocean genomics: microbial physiological ecology across marine ecosystems

Sonya Dyhrman

Professor, Department of Earth and Environmental Sciences

Senior Staff, Lamont-Doherty Earth Observatory

Columbia University



COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

LAMONT-DOHERTY
EARTH OBSERVATORY

Sonya Dyhrman

High School

Grad school/Postdoc

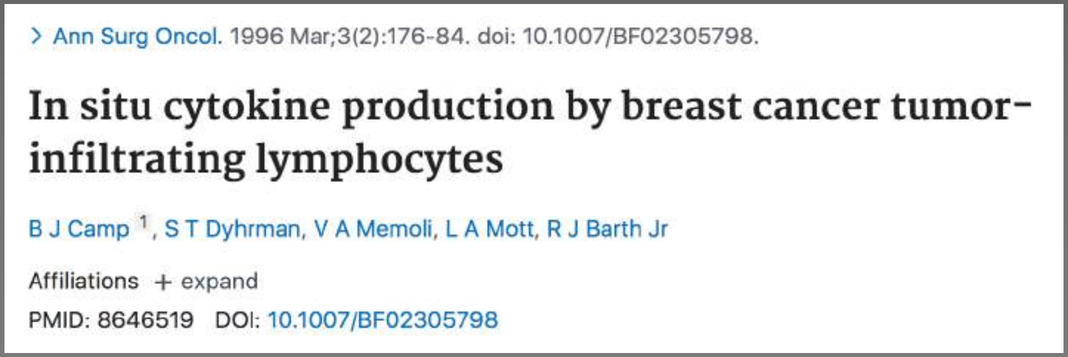
Professor

I have my dream job!

Sonya Dyhrman

New Hampshire
Dartmouth College

NIH National Institutes of Health
Turning Discovery Into Health



> Ann Surg Oncol. 1996 Mar;3(2):176-84. doi: 10.1007/BF02305798.

In situ cytokine production by breast cancer tumor-infiltrating lymphocytes

B J Camp [1], S T Dyhrman, V A Memoli, L A Mott, R J Barth Jr

Affiliations + expand

PMID: 8646519 DOI: 10.1007/BF02305798

Sonya Dyhrman

California
PhD Scripps - 1999

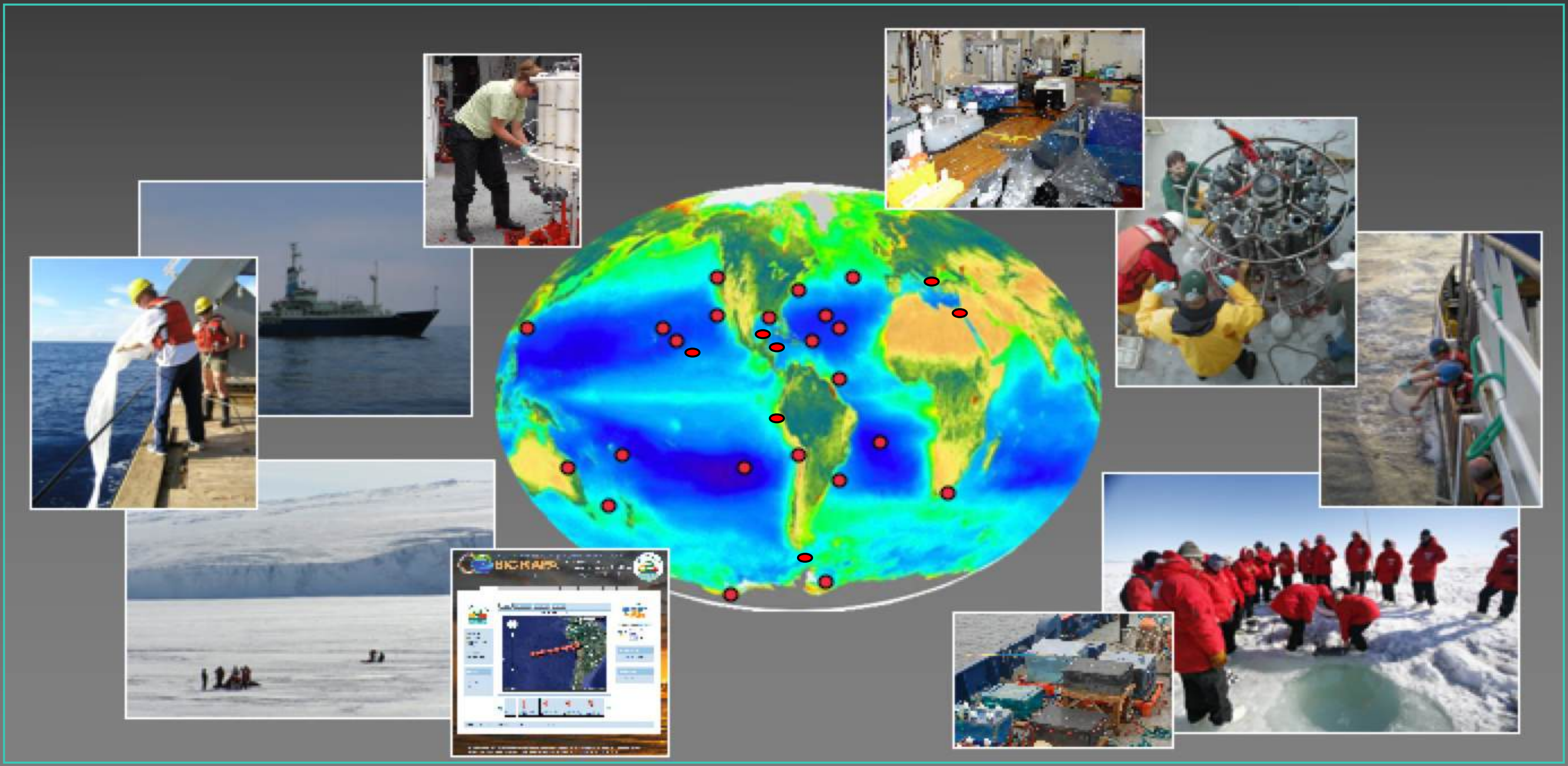
Massachusetts
Postdoc WHOI –
2000-2002

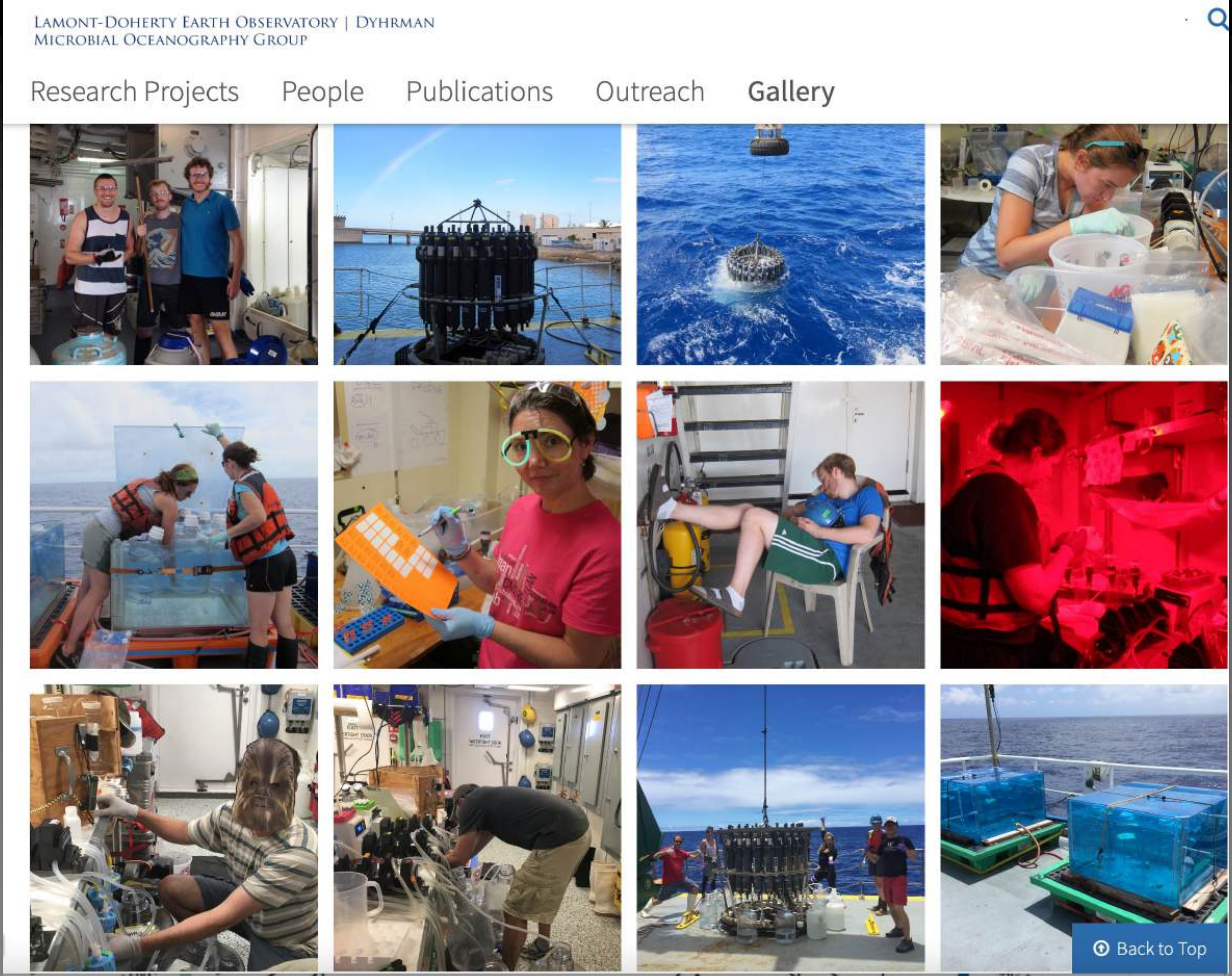
Scientific Staff WHOI
2002 - 2013

New York: Faculty Columbia University - 2013

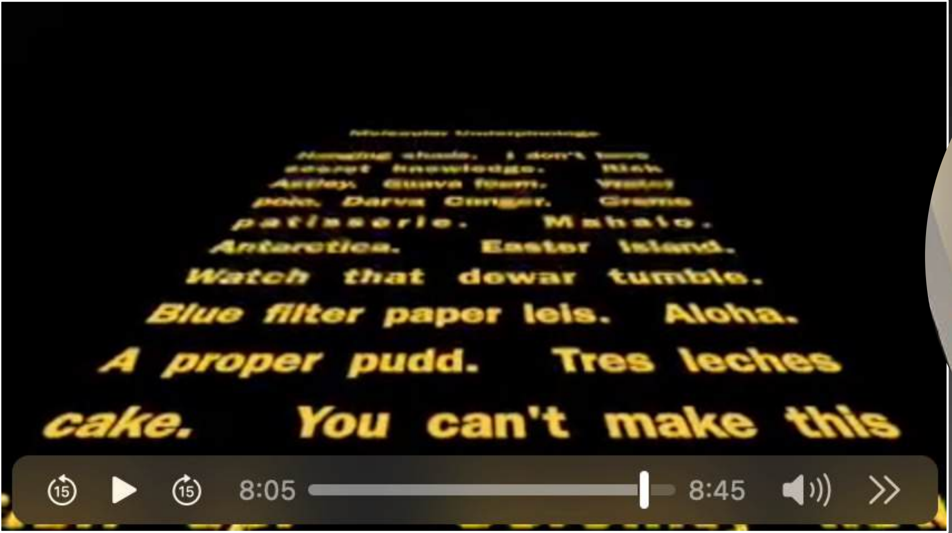
Columbia runs RV Langseth

Field work bonus – microbes are everywhere!




LAMONT-DOHERTY EARTH OBSERVATORY | DYHRMAN MICROBIAL OCEANOGRAPHY GROUP
Research Projects
People
Publications
Outreach
Gallery
Back to Top

Sonya – random (fun?) facts



[illegible] secret knowledge. Rick Astley. Guava foam. Water polo. Darva Conger. Creme patisserie. Mahalo. Antarctica. Easter Island. Watch that dewar tumble. Blue filter paper leis. Aloha. A proper pudd. Tres leches cake. You can't make this

C-CoMP at COP27

Dr. Sonya Dyhrman

The Key Role of Microbes in the Surface Ocean

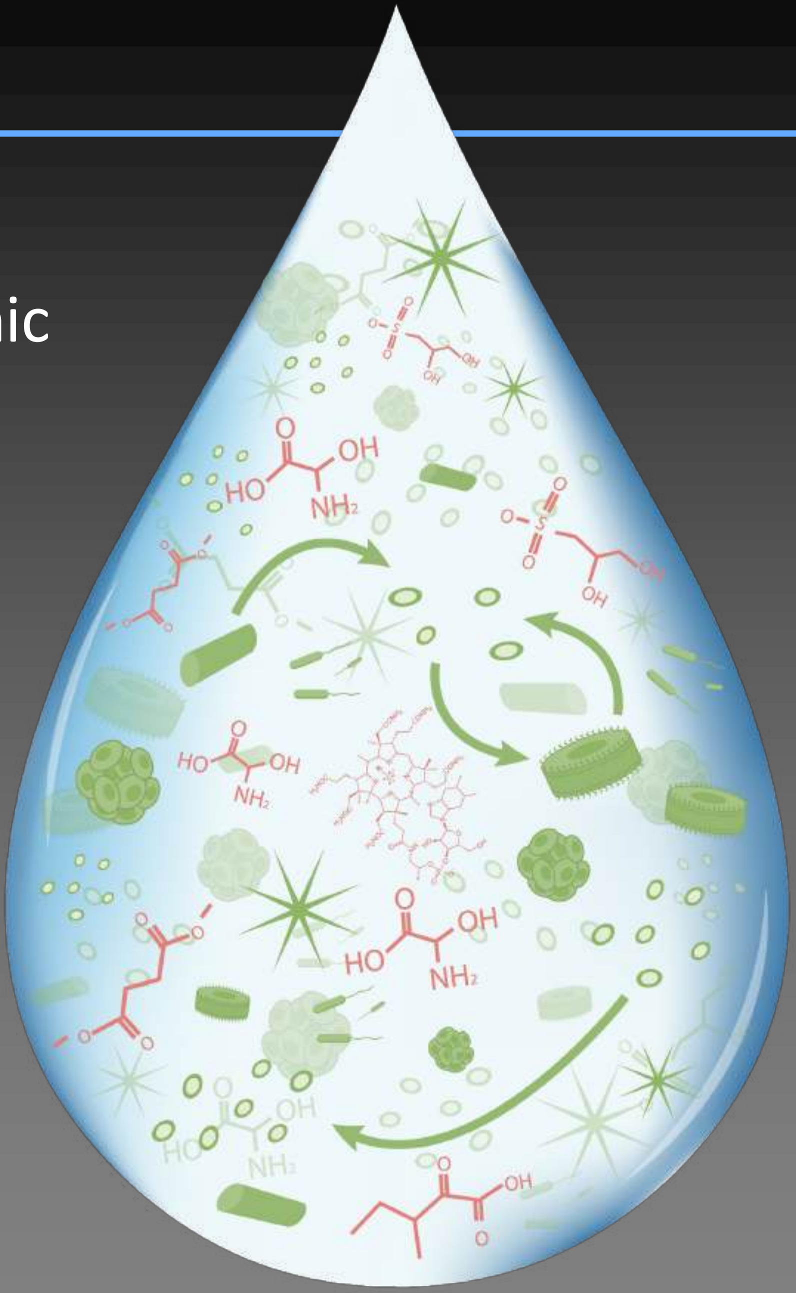
November 16th, 2022

Key points for today

- Highlight how we ask and answer questions in the ocean with genomic tools

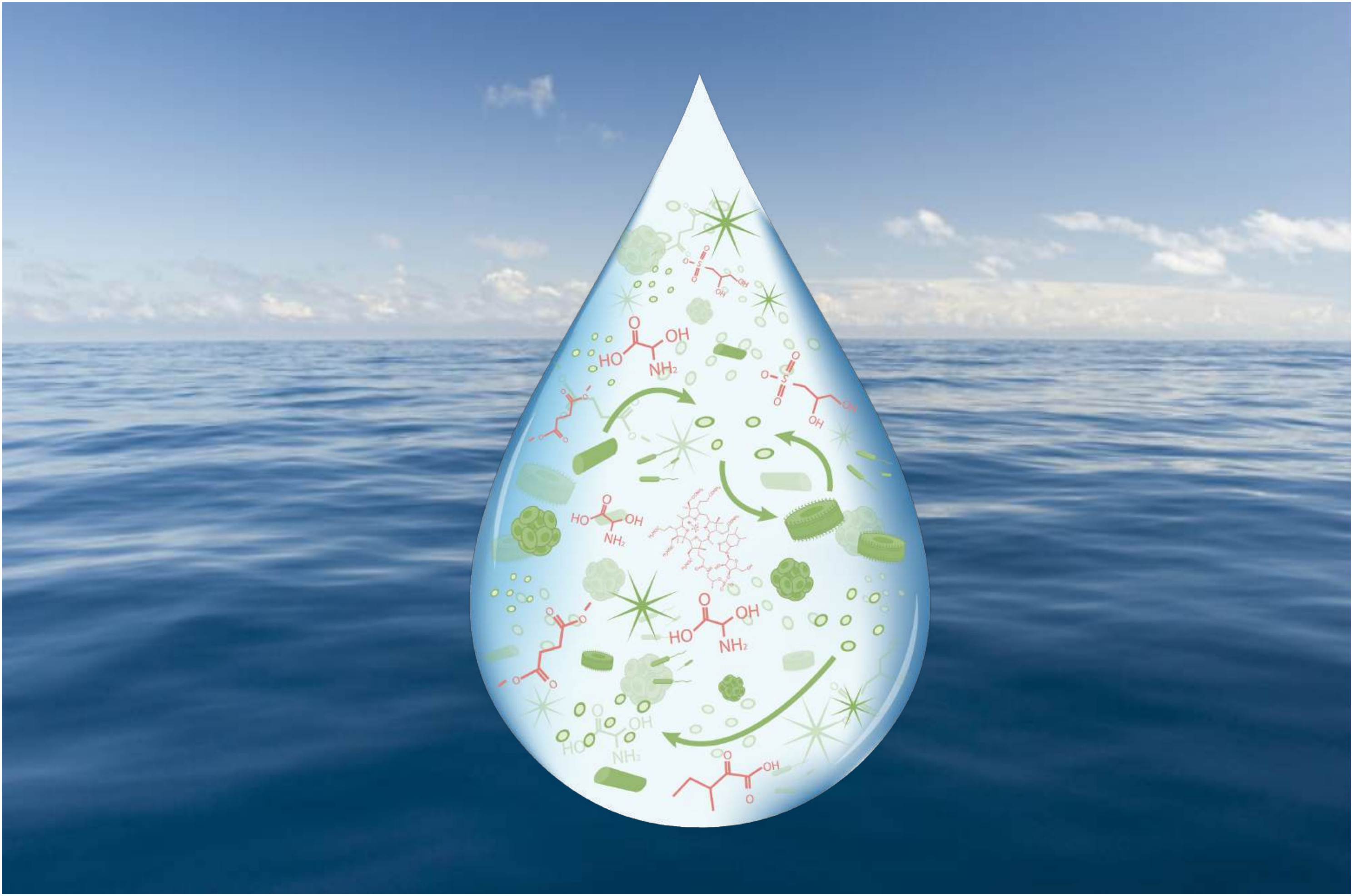
- Eukaryote physiological ecology
    - *Metatranscriptomes*
    - *Pan genomics*
- Microbial interactions
    - *16S Amplicons*
    - *Metagenomes*
    - *Metatranscriptomes*

- *Share my enthusiasm for marine microbes and 'omics*



The ocean makes our planet habitable

The vast unseen microbial populations play a critical role in ocean ecosystem function and the climate system

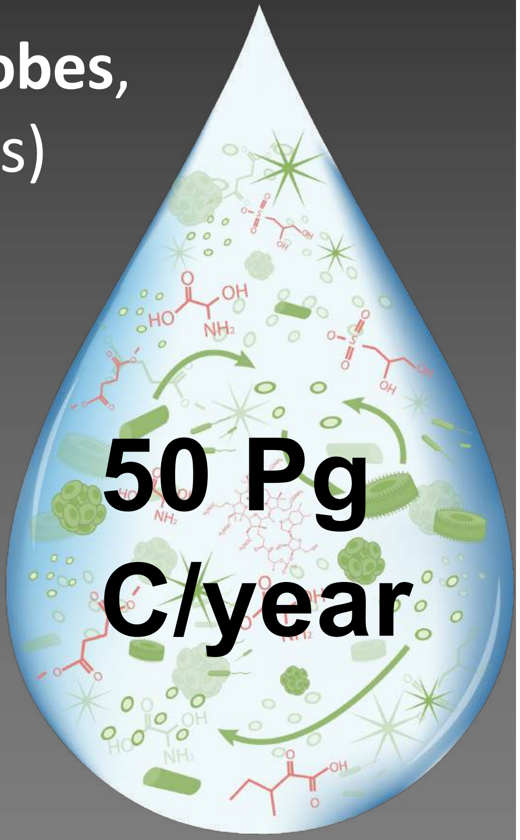


A vast chemical-microbe network underpins marine food webs and biogeochemical cycling

① In every 1mL of seawater, there are **10^6 microbes**, (*phytoplankton*, archaea, bacteria and protists) and **10^{10} viruses**

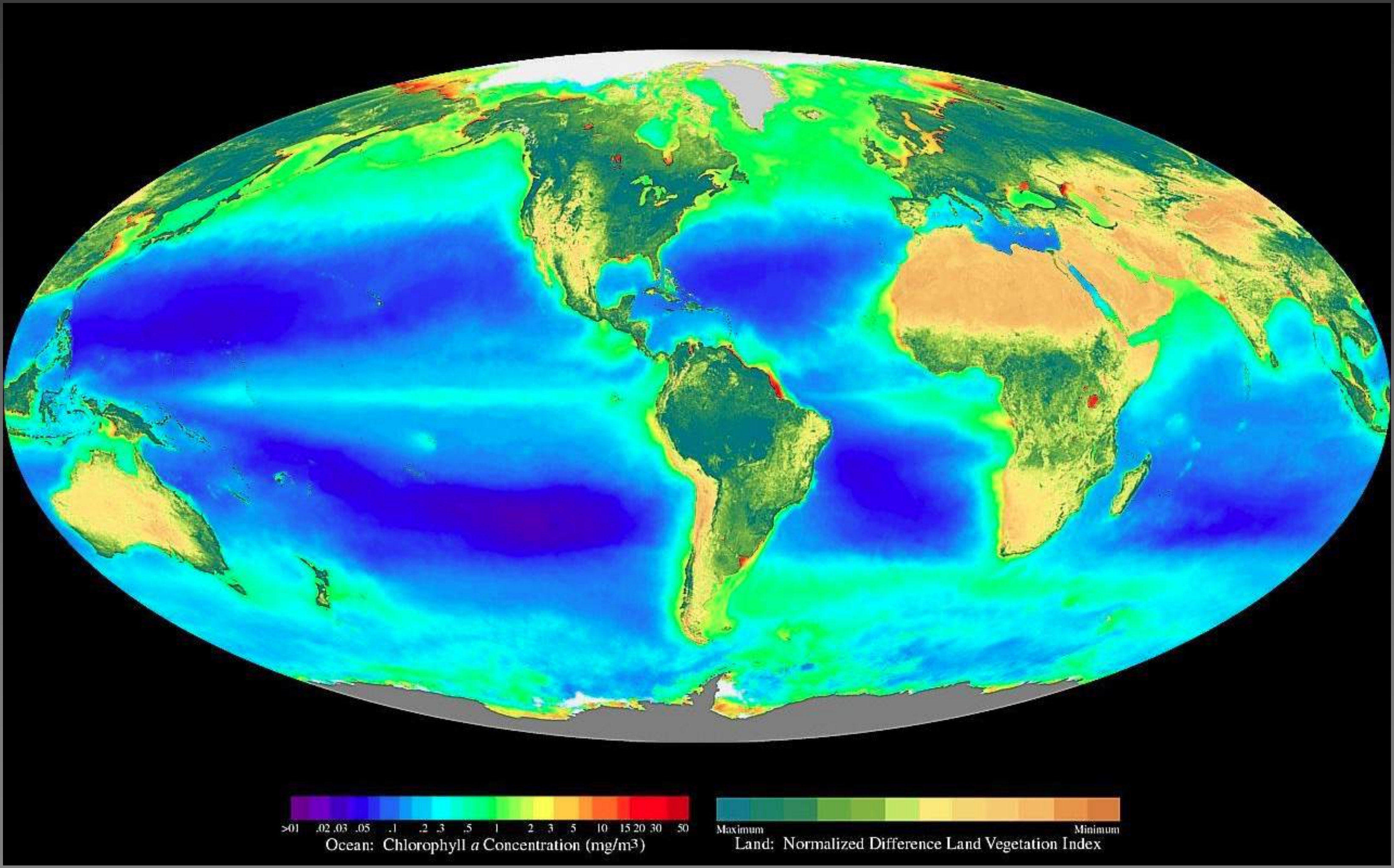
② There are **~100,000 different molecules** of varying concentrations (**femtomolar** to nanomolar) acting as substrates and signals

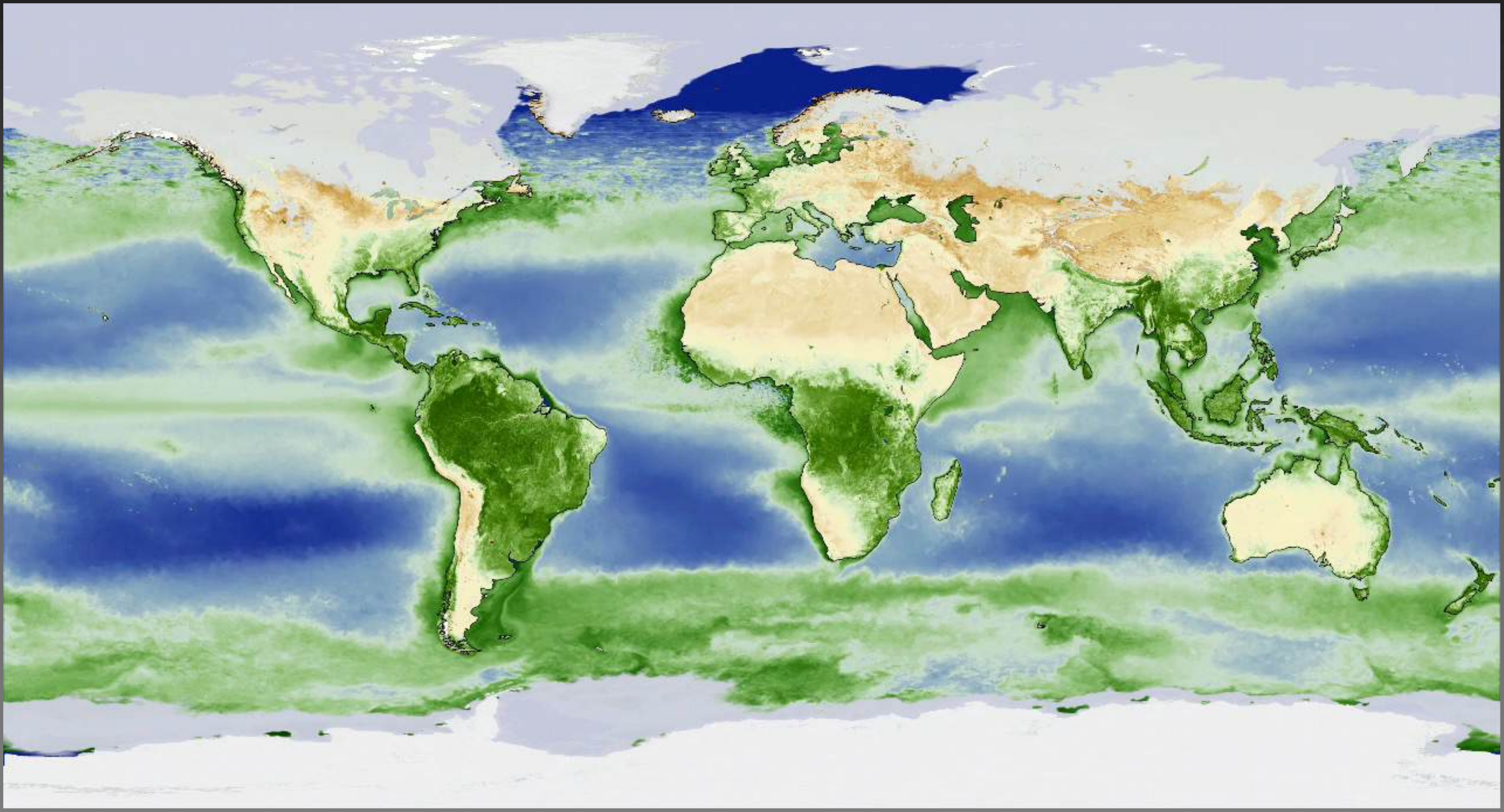
③ There are **myriad activities and interactions** within this network- many of which have not been resolved



Phytoplankton play a profound role in the earth system

Phytoplankton drive half of global primary production





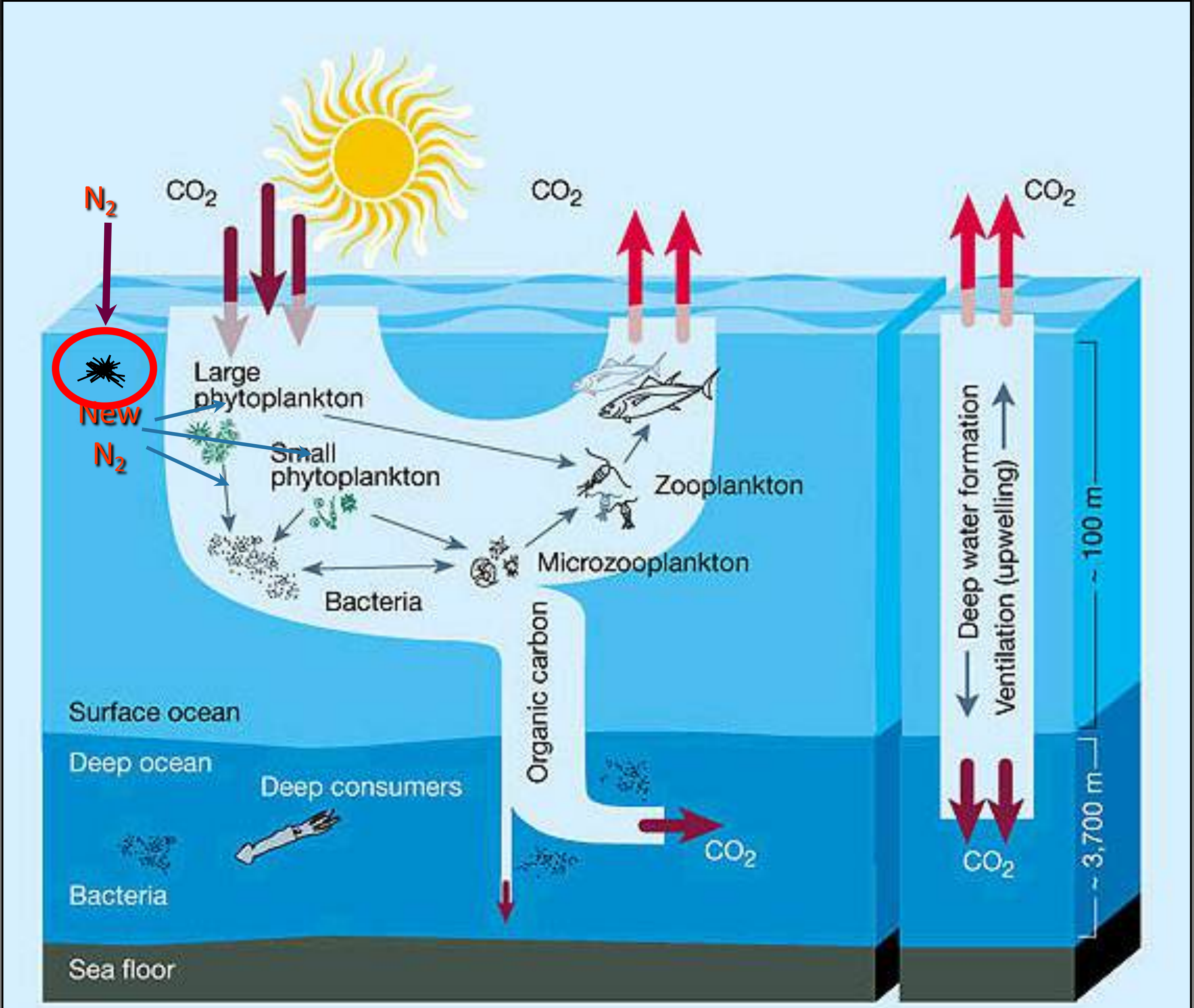
https://svs.gsfc.nasa.gov/30709

Phytoplankton underpin ocean ecosystem function

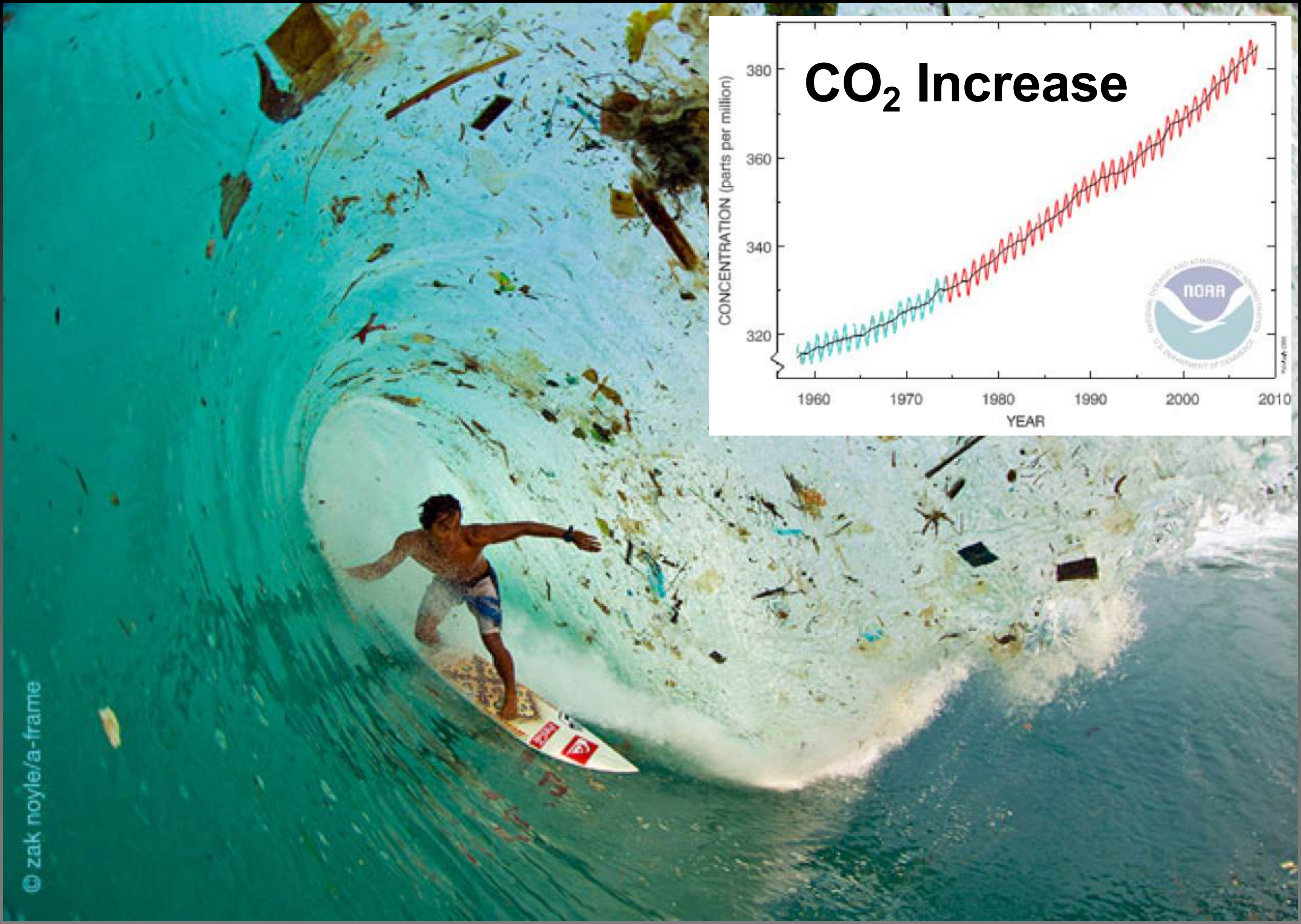
Haptophytes

Diatoms

N_2 Fixers



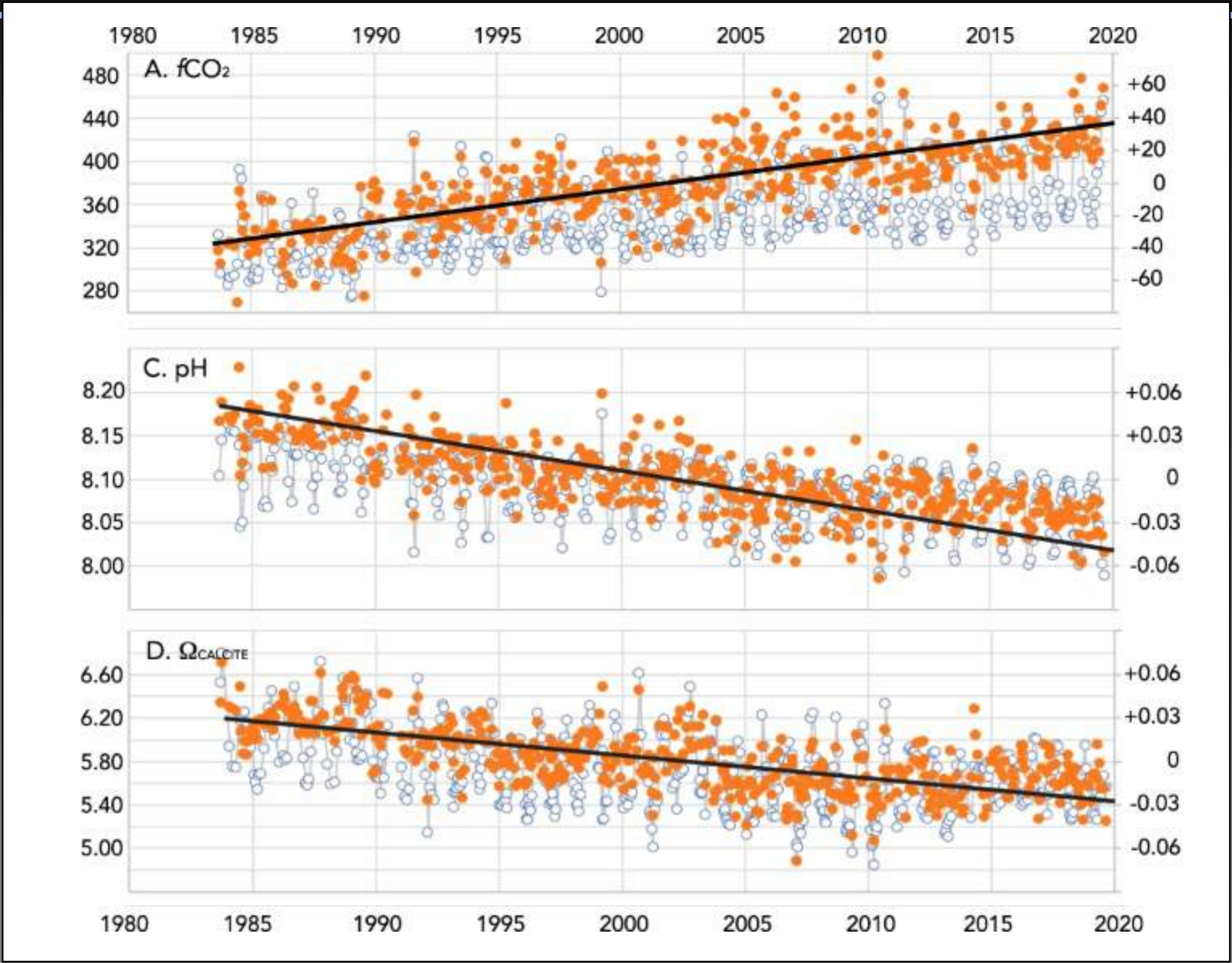
Chisholm 2000 *Nature*



© zak noyle/a-frame

Ocean conditions are changing and this will change the distribution and activities of microbes

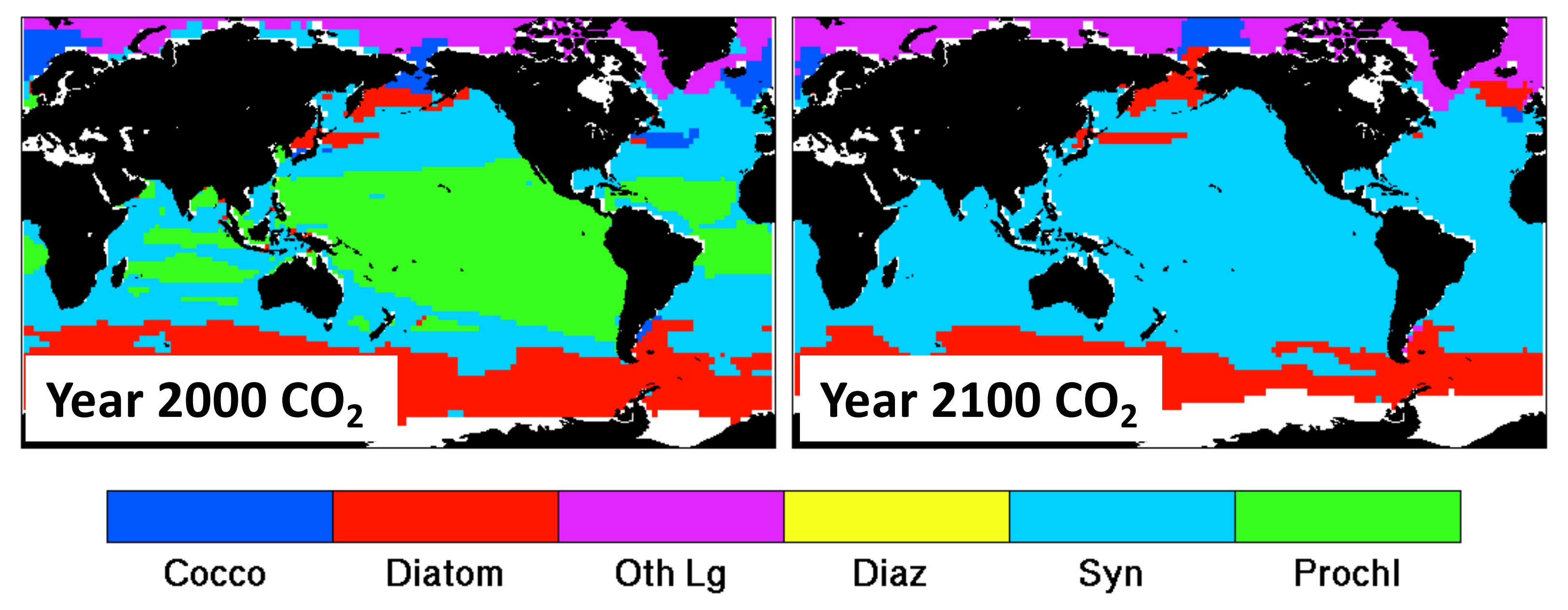
Changes in temperature CO_2, pH, etc.



Bates and Johnson (2020) *Communications Earth Env.*

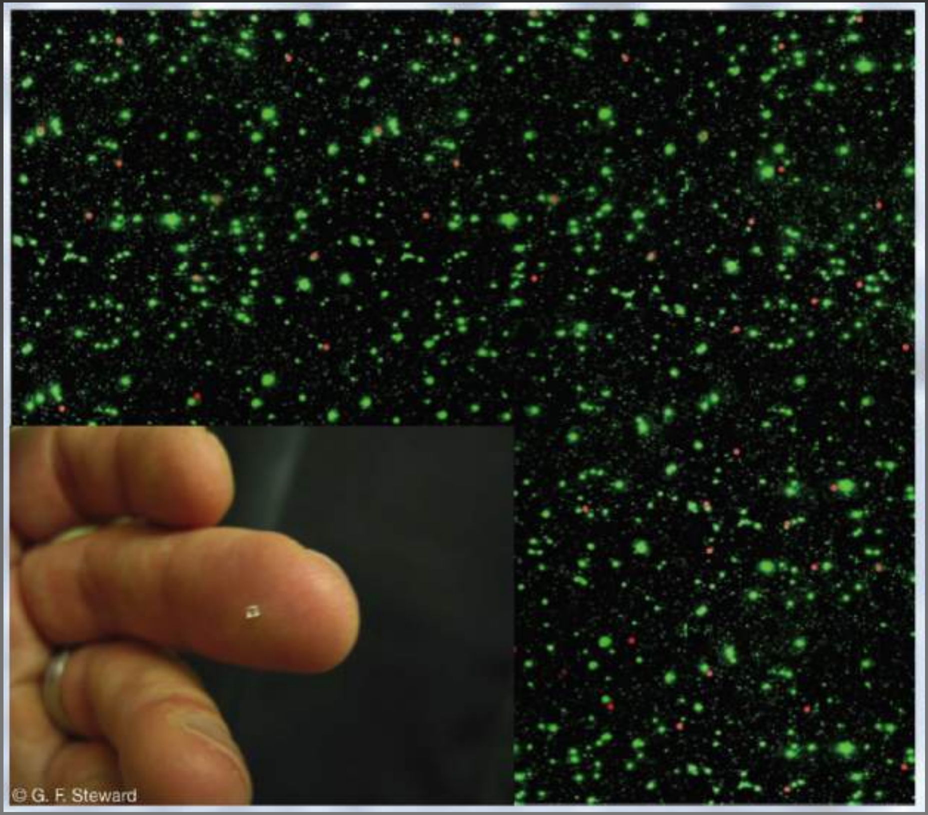
Models predict future shifts in microbial communities

Dominant Functional Group



Dutkiewicz et al. (2015) *Nature Clim. Change.*

What are the rules that govern the chemical – microbe network?



Who is there?
What are they doing?

If we learn the rules for how the system operates – then we have a better chance of predicting future ocean responses

'Omic-enabled advances allowing us to query cells in their environment in a species-specific way

Thank you Dyhrman Lab family past and present


Sheean Haley


Hanna Anderson


Jill Paquette


Erin Maybach


Amanda Ellis


Maxfield Palmer


Henry Holm


Annika Gomez

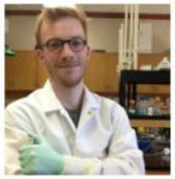

Gwenn Hennon


Mónica Rouco


Matt Harke


Maria Hernández-Limón


Harriet Alexander


Kyle Frischkorn

SIMONS FOUNDATION

WSL

NSF

C-CoMP

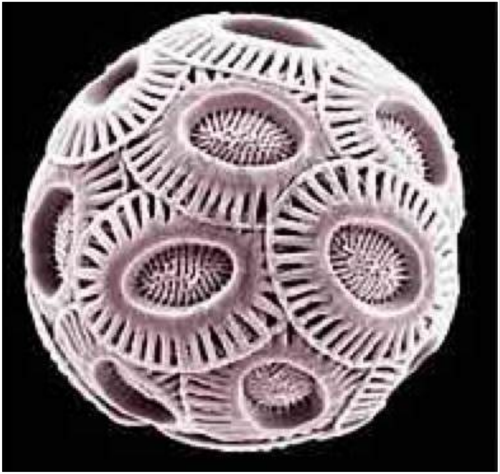
THE G. UNGER VETLESEN FOUNDATION

NOAA

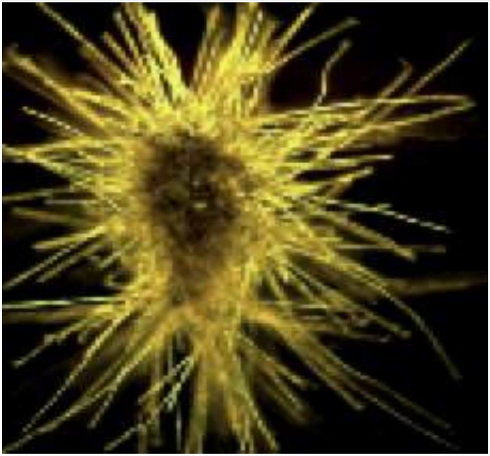
PAUL M. ANGELL FAMILY FOUNDATION

Special acknowledgement to WHOI and Harriet Alexander for the graphics

Vignettes



- Co-existing in a sea of competition: tracing drivers of phytoplankton physiological ecology in the open ocean



- Microbiome interactions: Tracking the microbiome of a keystone N_2 fixer

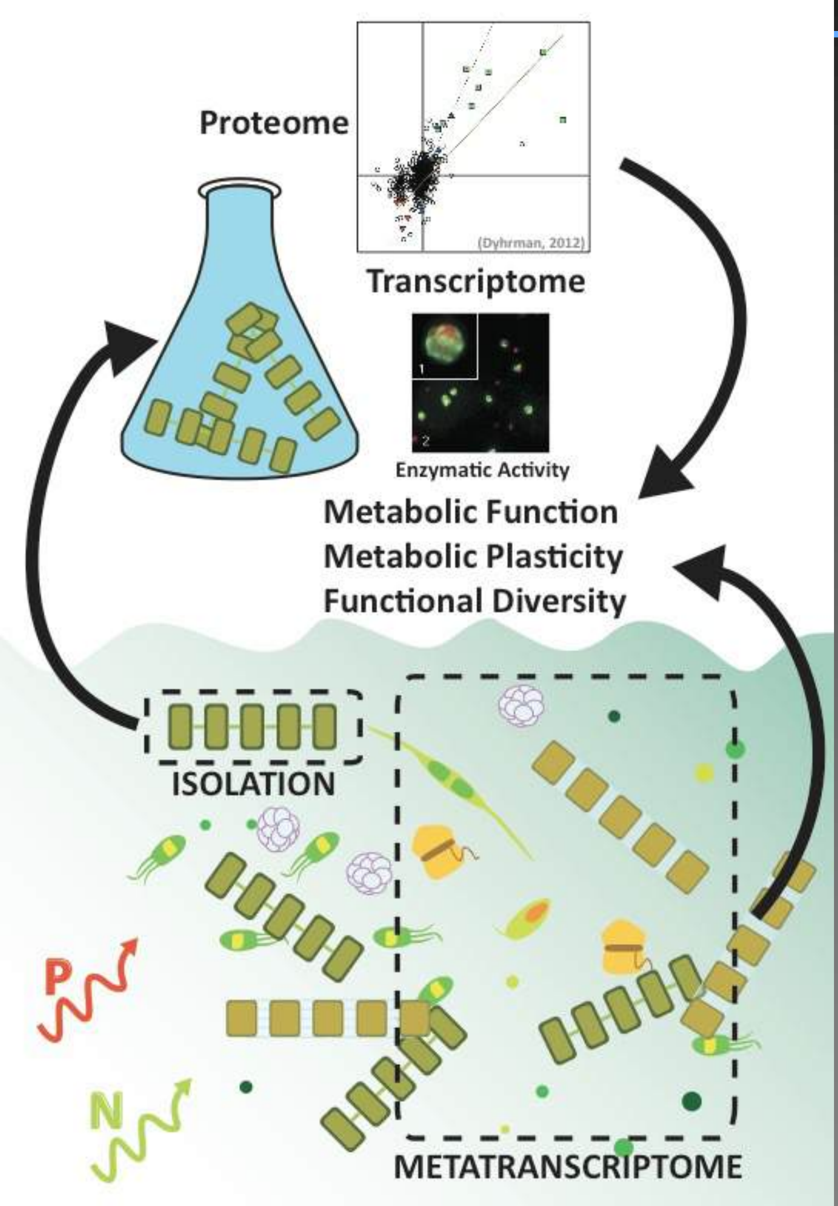
Leveraging metatranscriptome data to study the chemical – microbe network

- Cultures offer controlled conditions – link genes to functions and study their regulation

- Field samples offer context

- Metatranscriptomes can be species-specific

- Metatranscriptomes offer high resolution tracking of finely tuned responses



Microeukaryote metratranscriptomes with RNA-Seq

3x20L

Filter
(>5µm
.2µm)

Extract RNA
(and standards)

polyA Selected

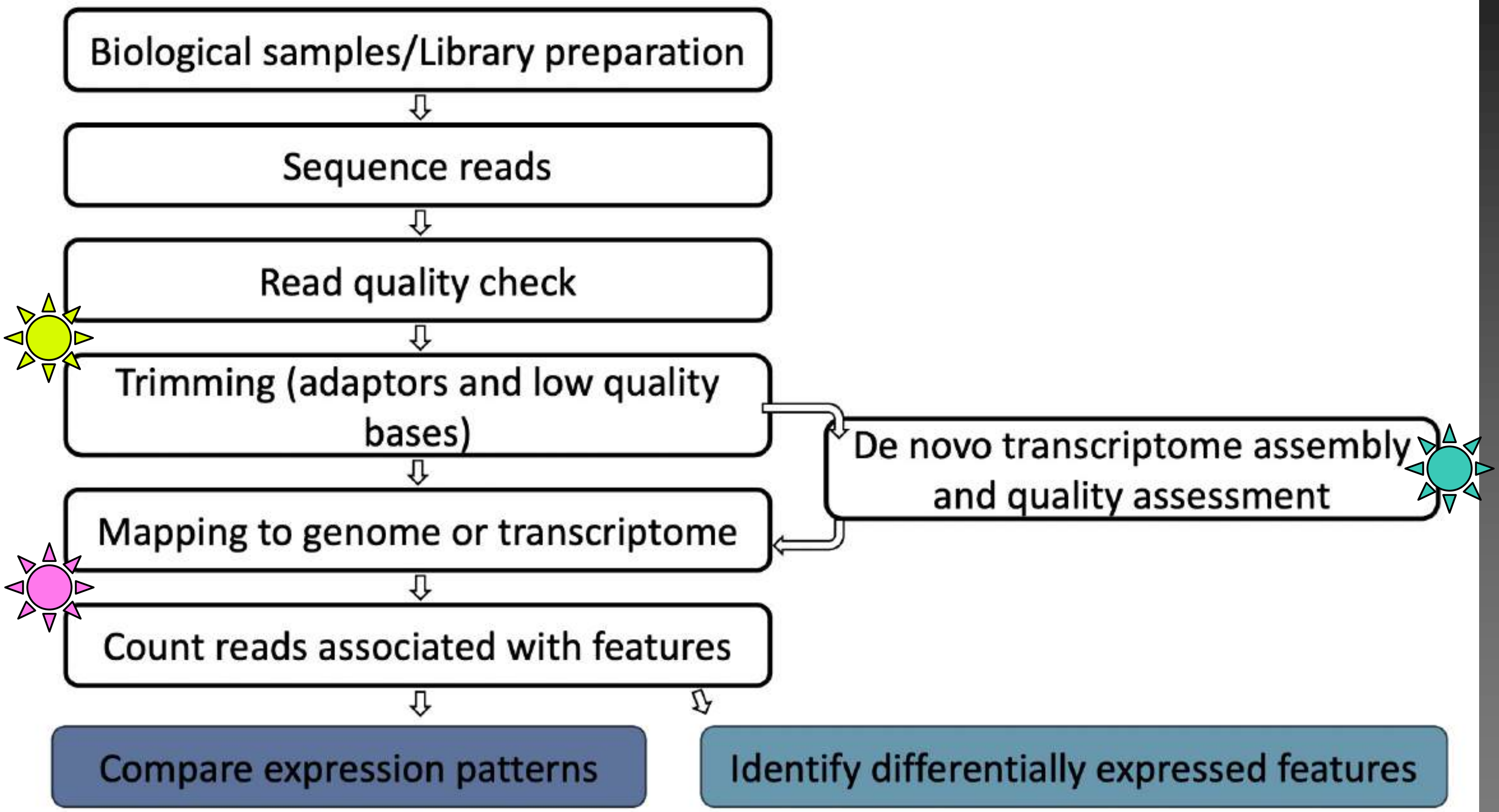
De novo assembly
Reference mapping

Unselected

rRNA
depletion

De novo assembly
Reference mapping

Gene expression analysis



Biological samples/Library preparation

⇩

Sequence reads

⇩

Read quality check

⇩

Trimming (adaptors and low quality bases)

⇩

De novo transcriptome assembly and quality assessment

Mapping to genome or transcriptome

⇩

Count reads associated with features

⇩

Compare expression patterns

Identify differentially expressed features

Quality Control
- Trimmomatic
- FastQC

Removal of Sequences
- riboPicker
- BBMap

Functional Annotation
- BLAST
- DIAMOND
- HMMer

- KEGG
- KOG
- GO
- PFam

Assembly/Reference
- **Trinity**
- TransABYSS
- rnaSPADES
- MEGAHIT

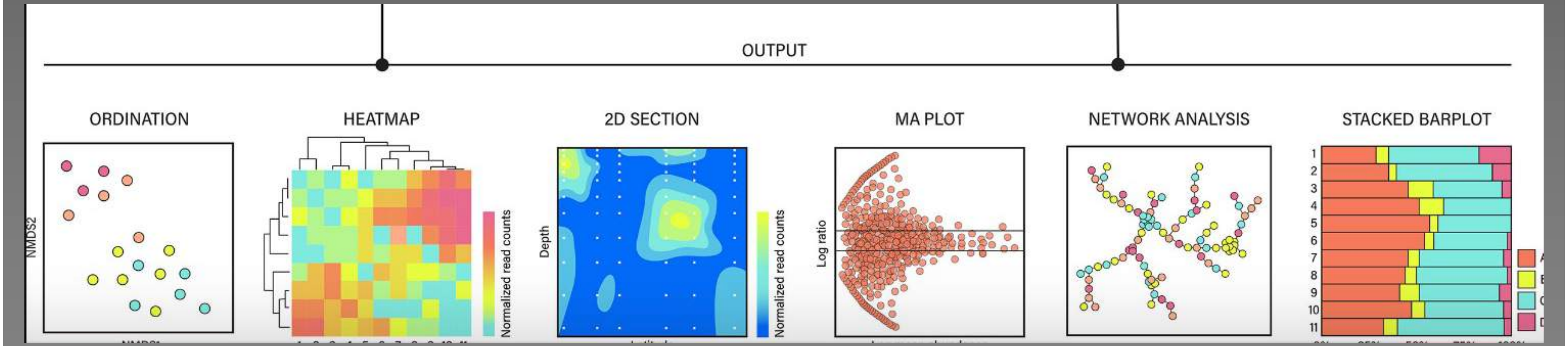
Protein Prediction and Clustering
- TransDecoder
- OrthoFinder

Taxonomic Annotation
- BLAST
- DIAMOND
- MMSeq2

- MMETSP
- EukProt
- EukZoo
- PhyloDB
- MarFERReT (2024)

Read Mapping/Counting
- **Salmon**
- Kallisto
- Bowtie2
- RSEM
- BWA etc/HTSeq

Normalization & DE
- TPM
- Variance stabilization
- EdgeR
- **DESeq2**
- **WGCNA**



After Cohen et al. (2020) *Front. Mar. Science*

Sequencing and analytical workflows are constantly improving

Harriet Alexander @nekton4plankton · Apr 27

So check it out!

GitHub here: github.com/AlexanderLabWH...
Documentation here: eukrhythmic.readthedocs.io/en/latest/
jEUKbox here: eukrhythmic.readthedocs.io/en/latest/

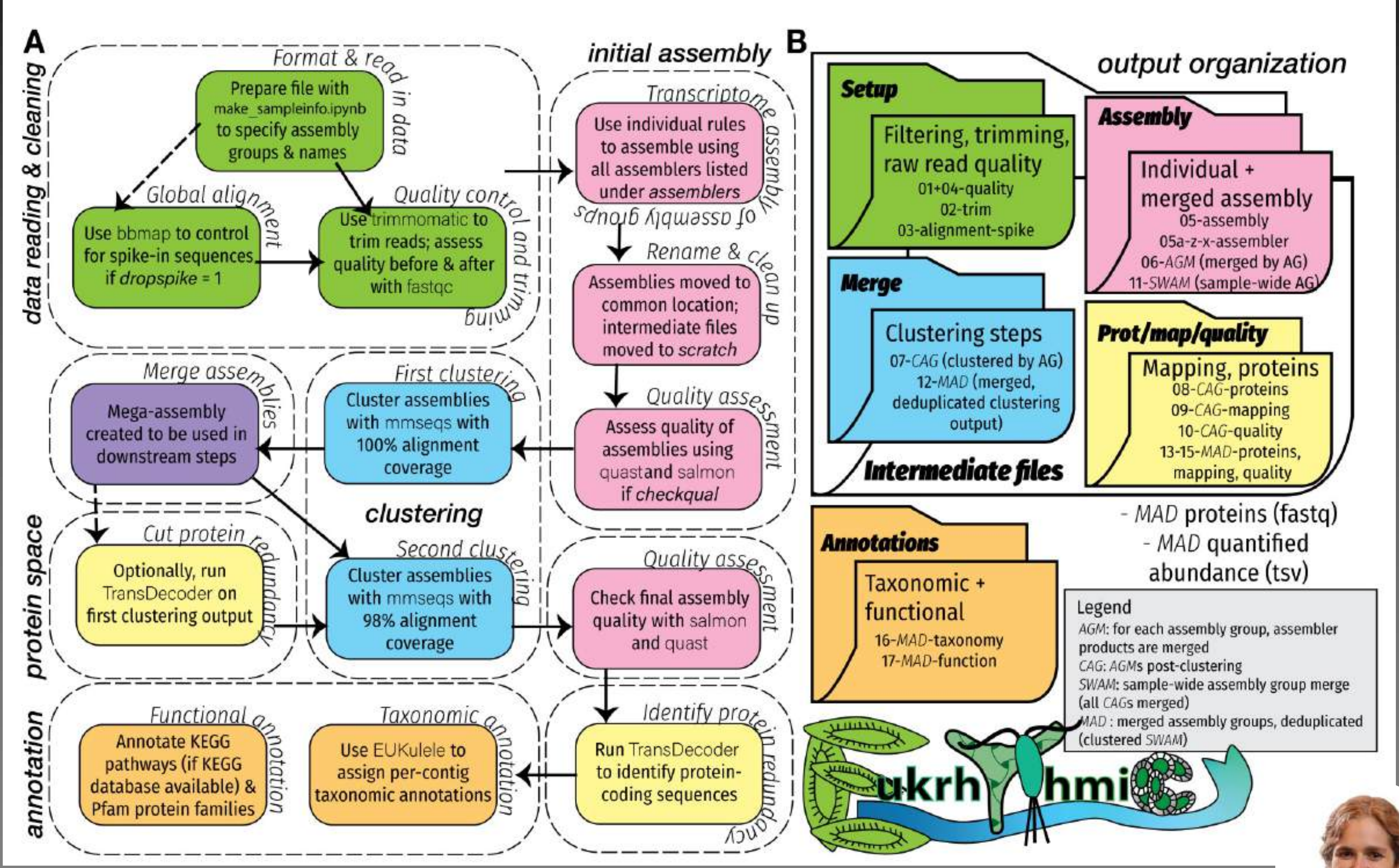
Feedback / comments welcome!

github.com

GitHub - AlexanderLabWHOI/eukrhythmic: A metatranscriptomic pipe...

A metatranscriptomic pipeline optimized for the study of microeukaryotes. - GitHub - AlexanderLabWHOI/eukrhythmic: A ...

Sequencing and analytical workflows are constantly improving



A

data reading & cleaning

Format & read in data

Prepare file with make_sampleinfo.ipynb to specify assembly groups & names

Global alignment

Use bbmap to control for spike-in sequences if *dropspike* = 1

Quality control and trimming

Use trimmomatic to trim reads; assess quality before & after with fastqc

initial assembly

Transcriptome assembly of assembly groups

Use individual rules to assemble using all assemblers listed under *assemblers*

Rename & clean up

Assemblies moved to common location; intermediate files moved to *scratch*

Quality assessment

Assess quality of assemblies using quastand salmon if *checkqual*

clustering

First clustering

Cluster assemblies with mmseqs with 100% alignment coverage

Merge assemblies

Mega-assembly created to be used in downstream steps

protein space

Cut protein redundancy

Optionally, run TransDecoder on first clustering output

Second clustering

Cluster assemblies with mmseqs with 98% alignment coverage

Quality assessment

Check final assembly quality with salmon and quast

annotation

Identify protein redundancy

Run TransDecoder to identify protein-coding sequences

Taxonomic annotation

Use EUKulele to assign per-contig taxonomic annotations

Functional annotation

Annotate KEGG pathways (if KEGG database available) & Pfam protein families

B

output organization

Setup

Filtering, trimming, raw read quality
01+04-quality
02-trim
03-alignment-spike

Assembly

Individual + merged assembly
05-assembly
05a-z-x-assembler
06-*AGM* (merged by AG)
11-*SWAM* (sample-wide AG)

Merge

Clustering steps
07-*CAG* (clustered by AG)
12-*MAD* (merged, deduplicated clustering output)

Prot/map/quality

Mapping, proteins
08-*CAG*-proteins
09-*CAG*-mapping
10-*CAG*-quality
13-15-*MAD*-proteins, mapping, quality

Intermediate files

Annotations

Taxonomic + functional
16-*MAD*-taxonomy
17-*MAD*-function

- *MAD* proteins (fastq)
- *MAD* quantified abundance (tsv)

Legend
AGM: for each assembly group, assembler products are merged
CAG: *AGM*s post-clustering
SWAM: sample-wide assembly group merge (all *CAG*s merged)
MAD : merged assembly groups, deduplicated (clustered *SWAM*)

Eukrhythmic

Krinos et al. 2023 *BMC Bioinformatics*

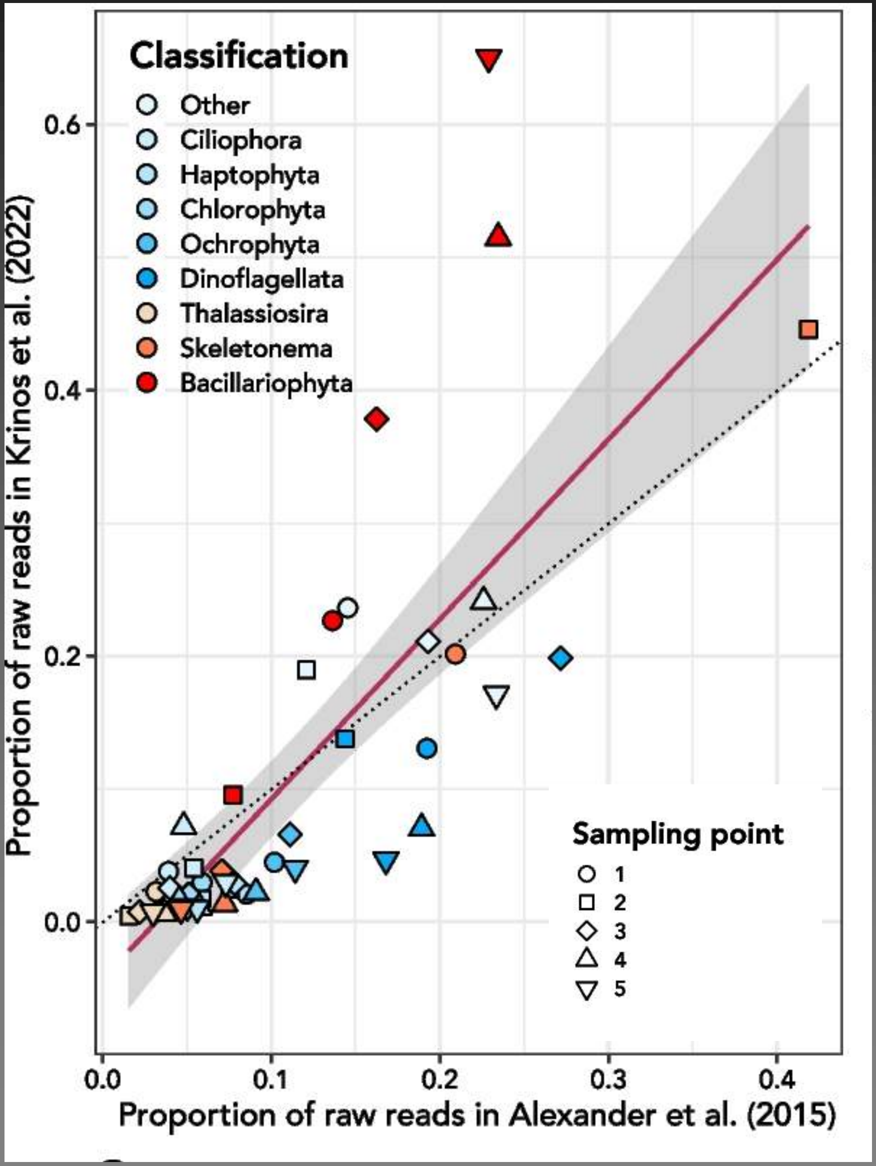
Metatranscriptome workflow comparison effort

- Assembly captures a broader swath of reads in high abundance – highly diverse groups (red)

- Reference mapping captures a broader swath of reads in low abundance groups (blue)

- Equally good at recapitulating patterns when references are from local isolates (orange and yellow)

De novo Assembly



Reference mapping

Krikinos et al. 2023 *BMC Bioinformatics*

Eukaryote - Metatranscriptome intercomparison effort

- Large diameter filter sampled with CLIO (June 2023)

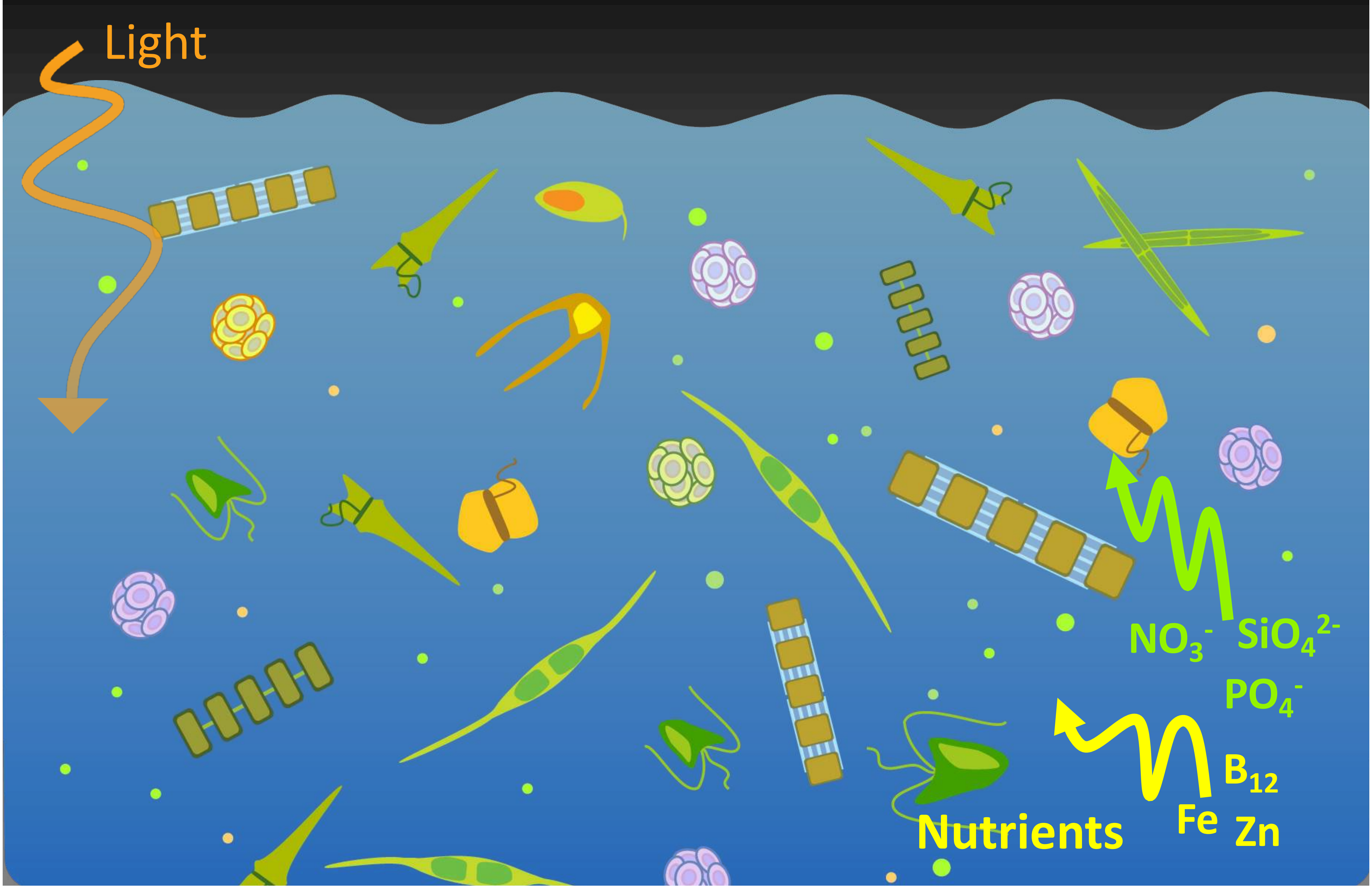
- 2x filter pieces to ~ 7 groups
  - Extraction
  - Library prep
  - Sequencing depth and read length

- Fall 2024 hackathon:
  - Workflow intercomparison
    - One sample with different workflows
    - All samples with one workflow

- Will not answer all questions – but it will be a valuable starting point for building intercomparable datasets

Thank you N. Cohen, H. Alexander, S. Hu and A. Marchetti

Complex community dynamics driven in part by resources



Tracking phytoplankton physiological ecology

Ecosystem function and biogeochemistry

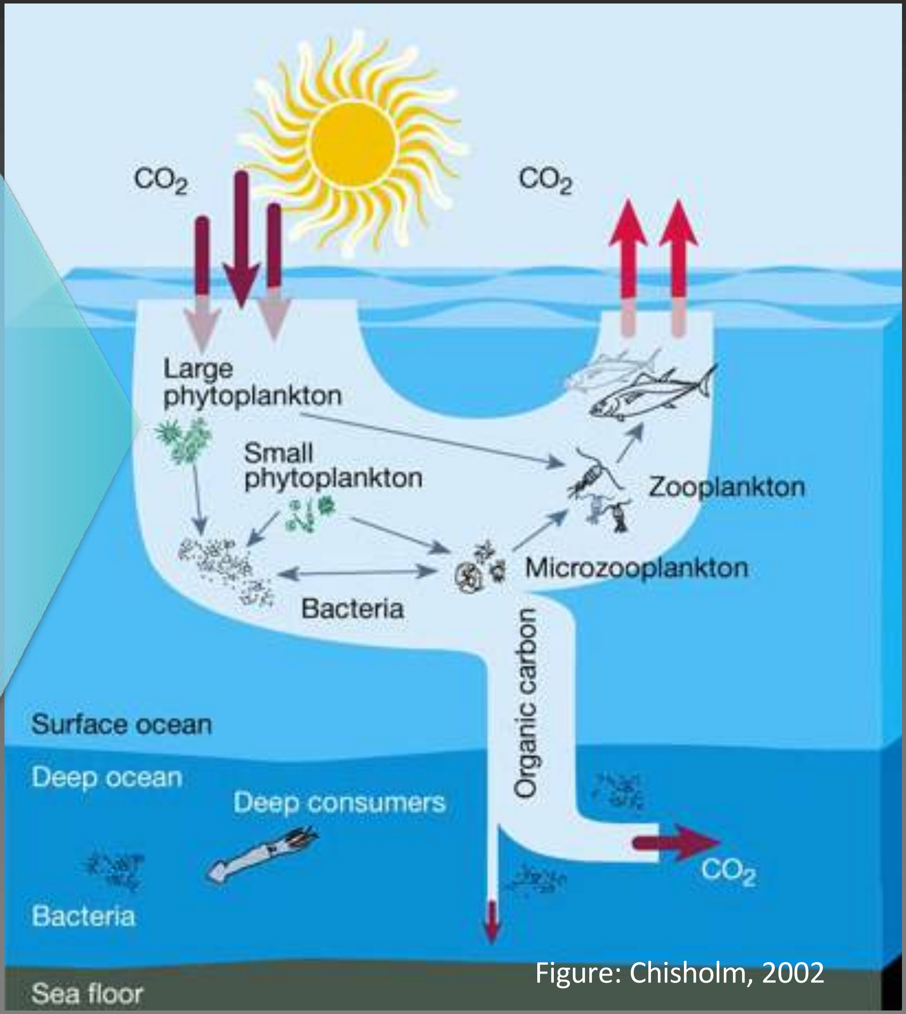
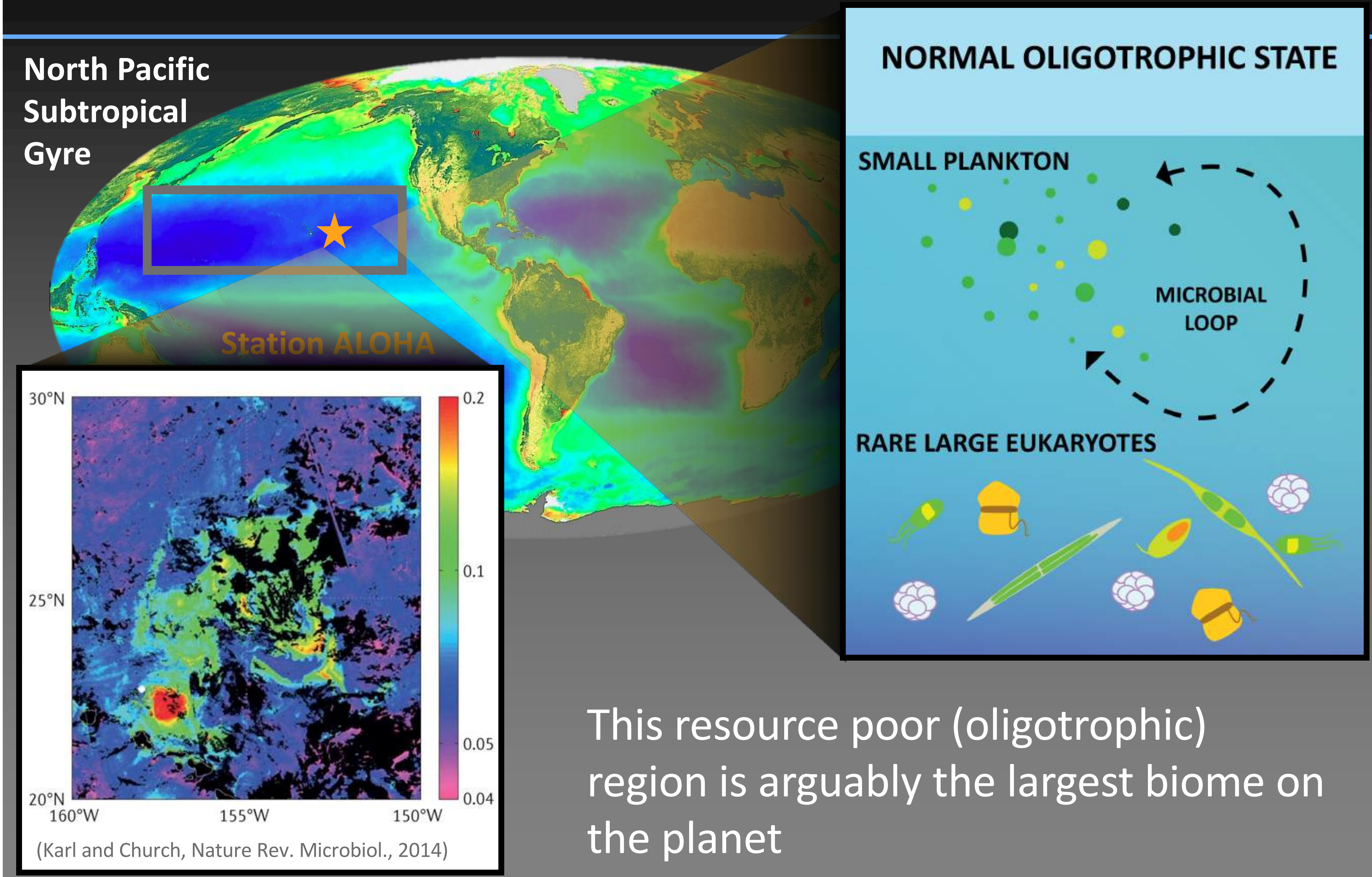


Figure: Chisholm, 2002

How do resources drive phytoplankton distributions and activities?

North Pacific Subtropical Gyre

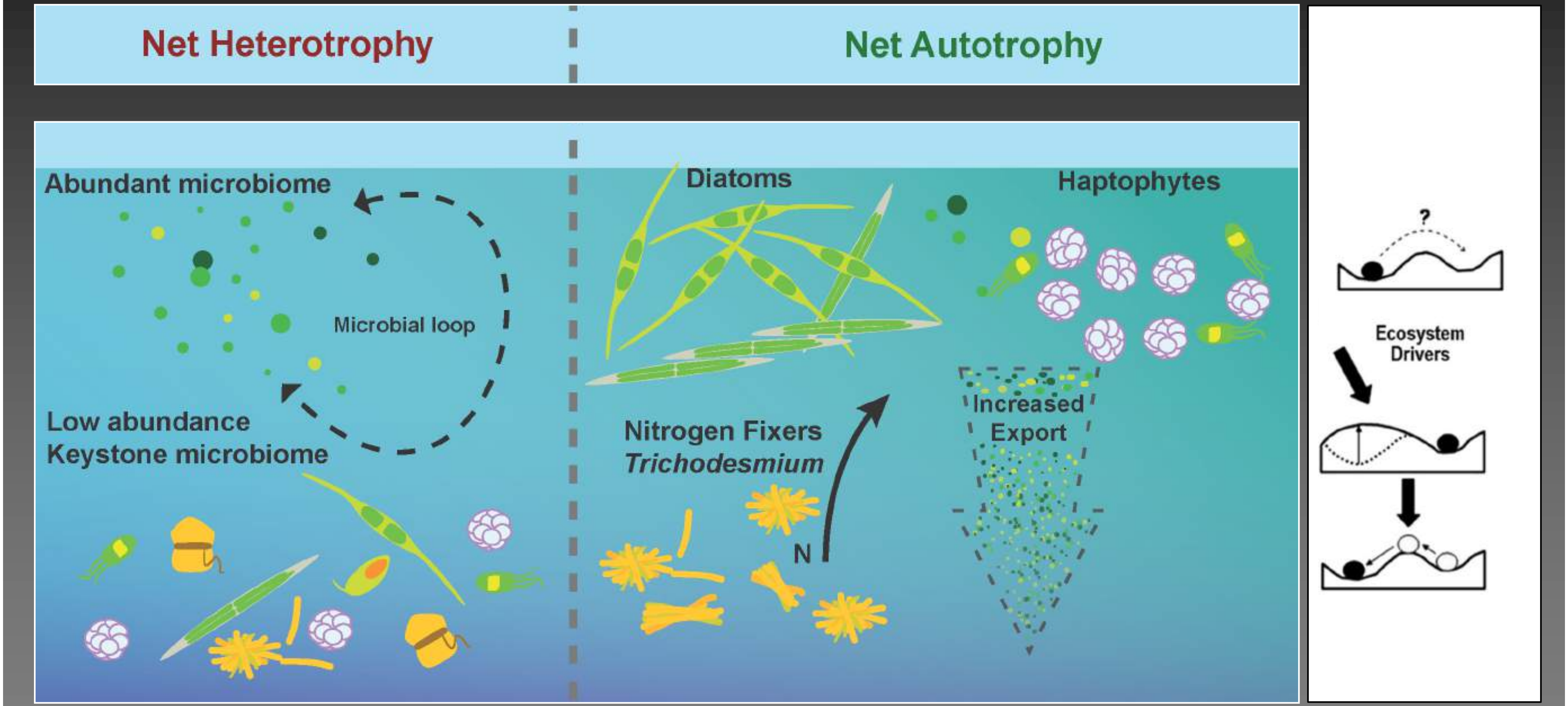


(Karl and Church, Nature Rev. Microbiol., 2014)

This resource poor (oligotrophic) region is arguably the largest biome on the planet

Image credits: SeaWIFS Global Chlorophyll

Blooms of keystone species can shift ecosystem state and carbon flow



Nutrient input, or other forcings

Core questions… adventures in eukaryote metatranscriptomics

Limitation

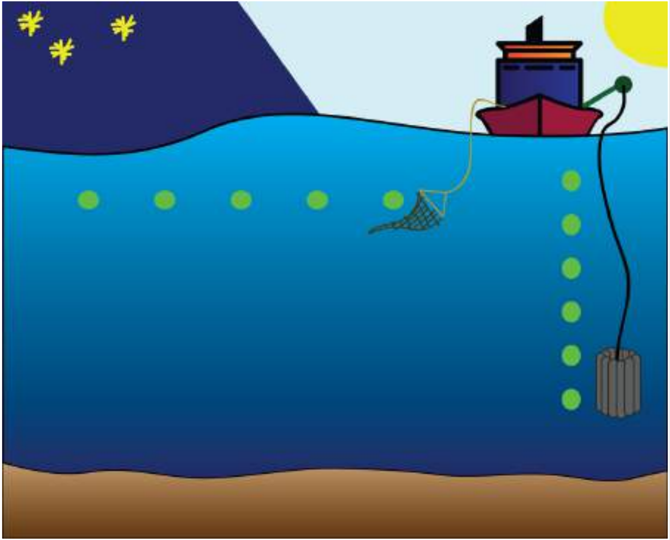
- What limits carbon cycling? – tracing physiological ecology in an expanding oligotrophic ocean with metatranscriptomics

Competition

- A day in the life – tracing competitive interactions with multi-’omics

HOE:DYLAN

HOE:MAHALO



* *Emiliania huxleyi:* a cosmopolitan, globally significant species

- Calcification - critical role in global carbon cycle and strongly linked to climate driven ocean acidification
- Source of paleoproxies for climate reconstructions
- Form dense blooms, drivers largely unknown
- First eukaryotic marine phytoplankton to have multiple strains/ecotypes sequenced

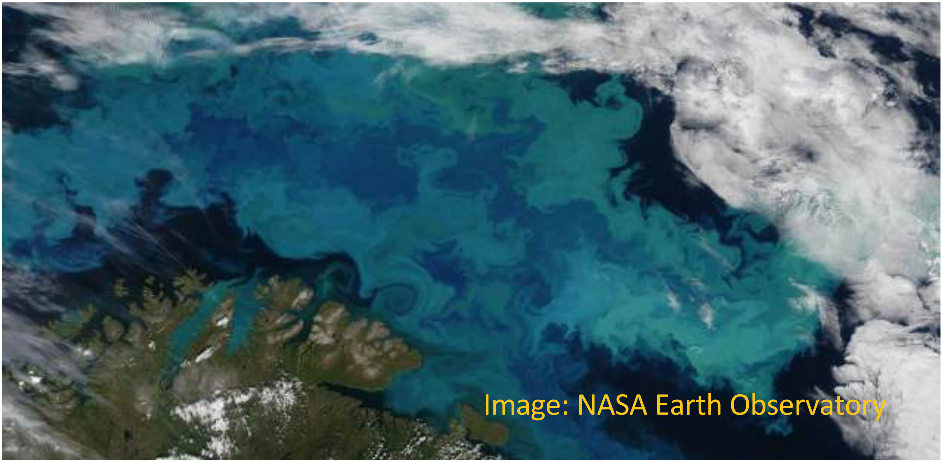
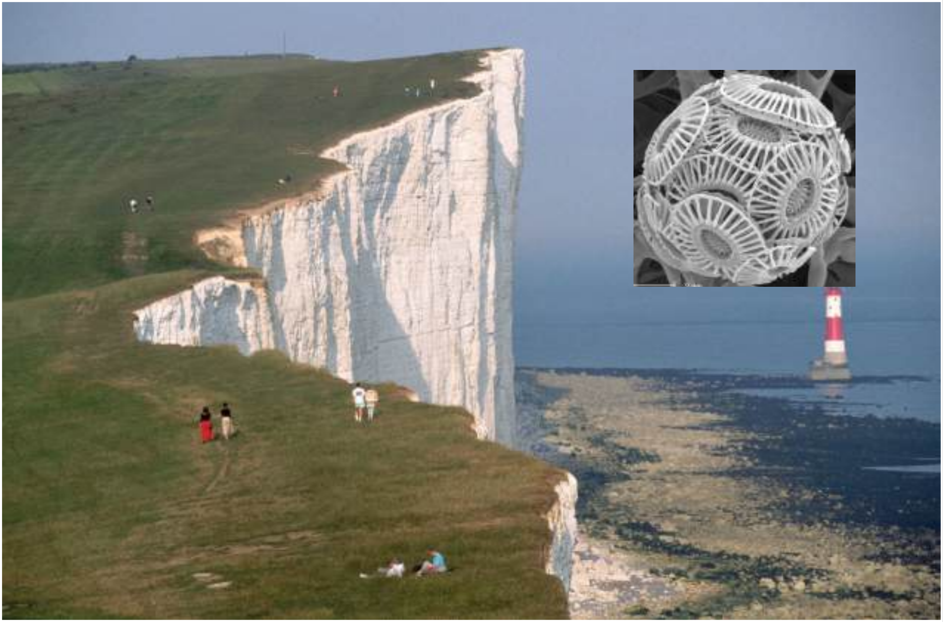
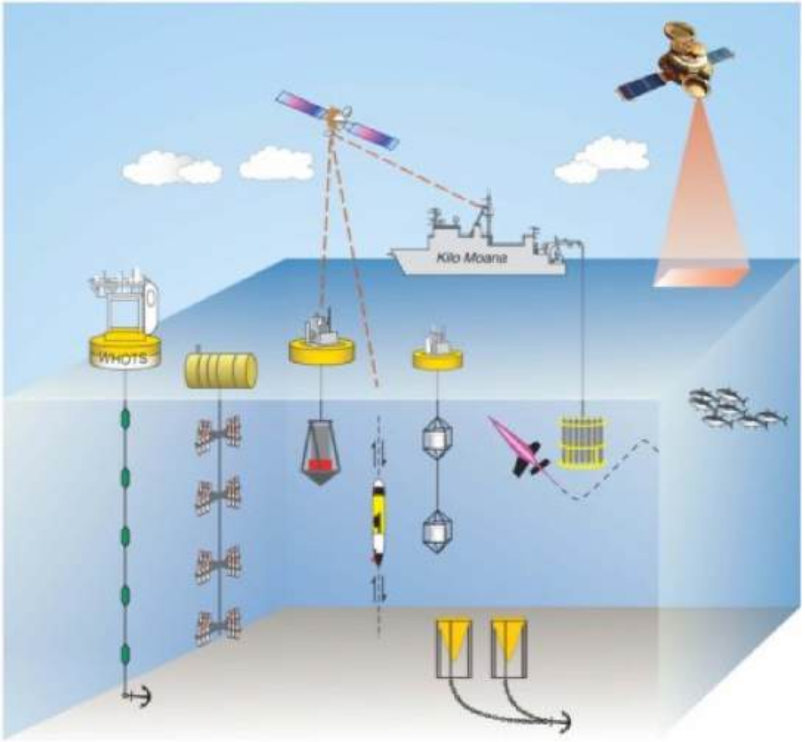


Image: NASA Earth Observatory



**Gephyrocapsa* (Bendif et al., 2023; Filatov et al., 2021)

Hawaii Ocean Experiment: Dynamics of Light and Nutrients



Wilson et al. 2015 *GBC*

David M. Karl

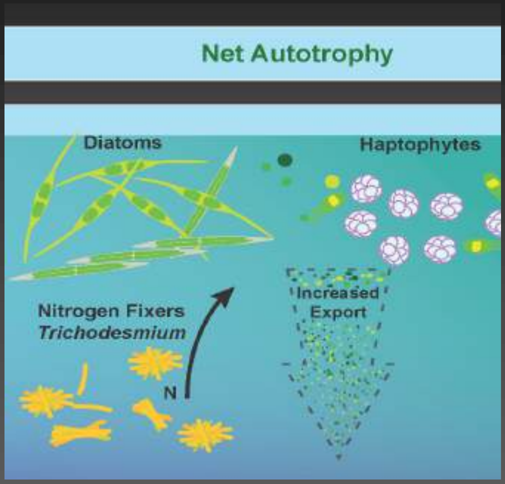
center for microbial oceanography: research and education

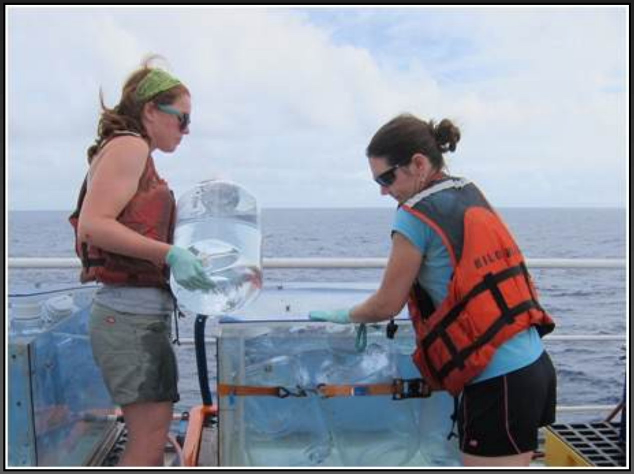
C·MORE *linking genomes to biomes*

(HOE-DYLAN)

HOE:DYLAN: Water sampling

HOE:DYLAN: Semi-factorial resource additions done 2 times





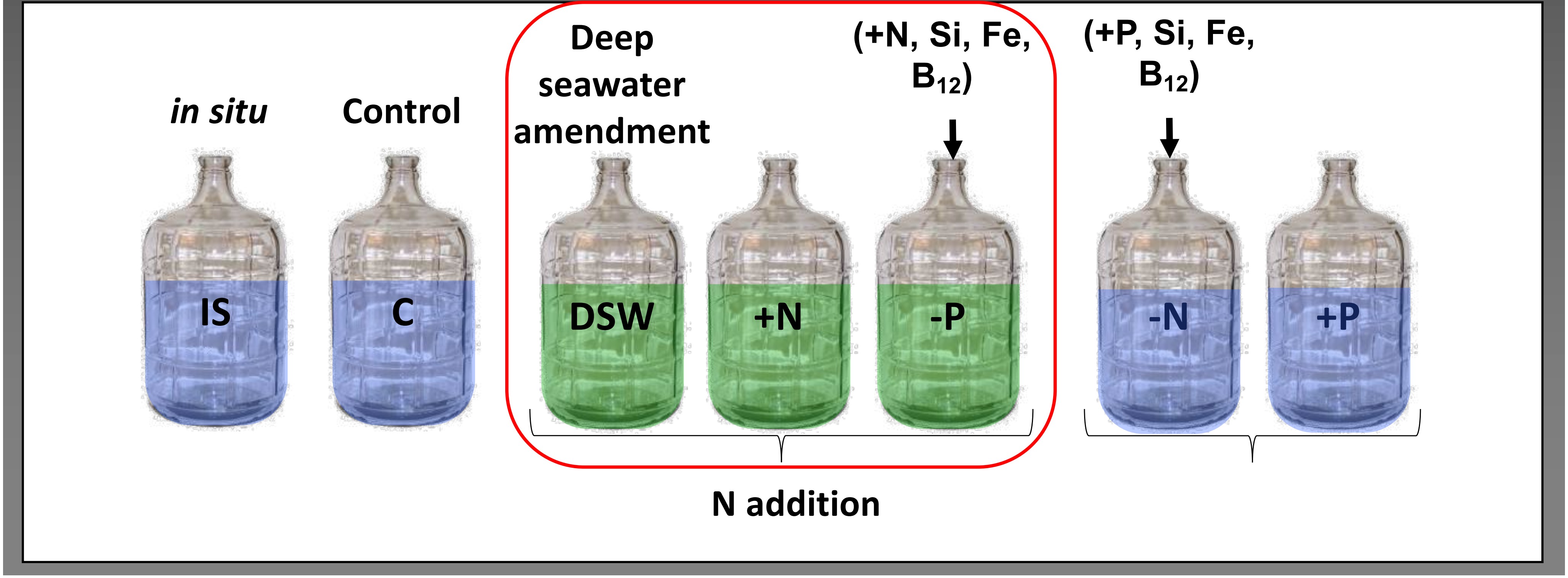
Nitrate/nitrite*: ~40 μM

Phosphate*: ~2.6 μM

Silicate*: ~100 μM

Fe: ~0.6 nmol/kg

B_{12}: <0.1 pM



in situ

Control

Deep seawater amendment

(+N, Si, Fe, B_{12})

(+P, Si, Fe, B_{12})

IS

C

DSW

+N

-P

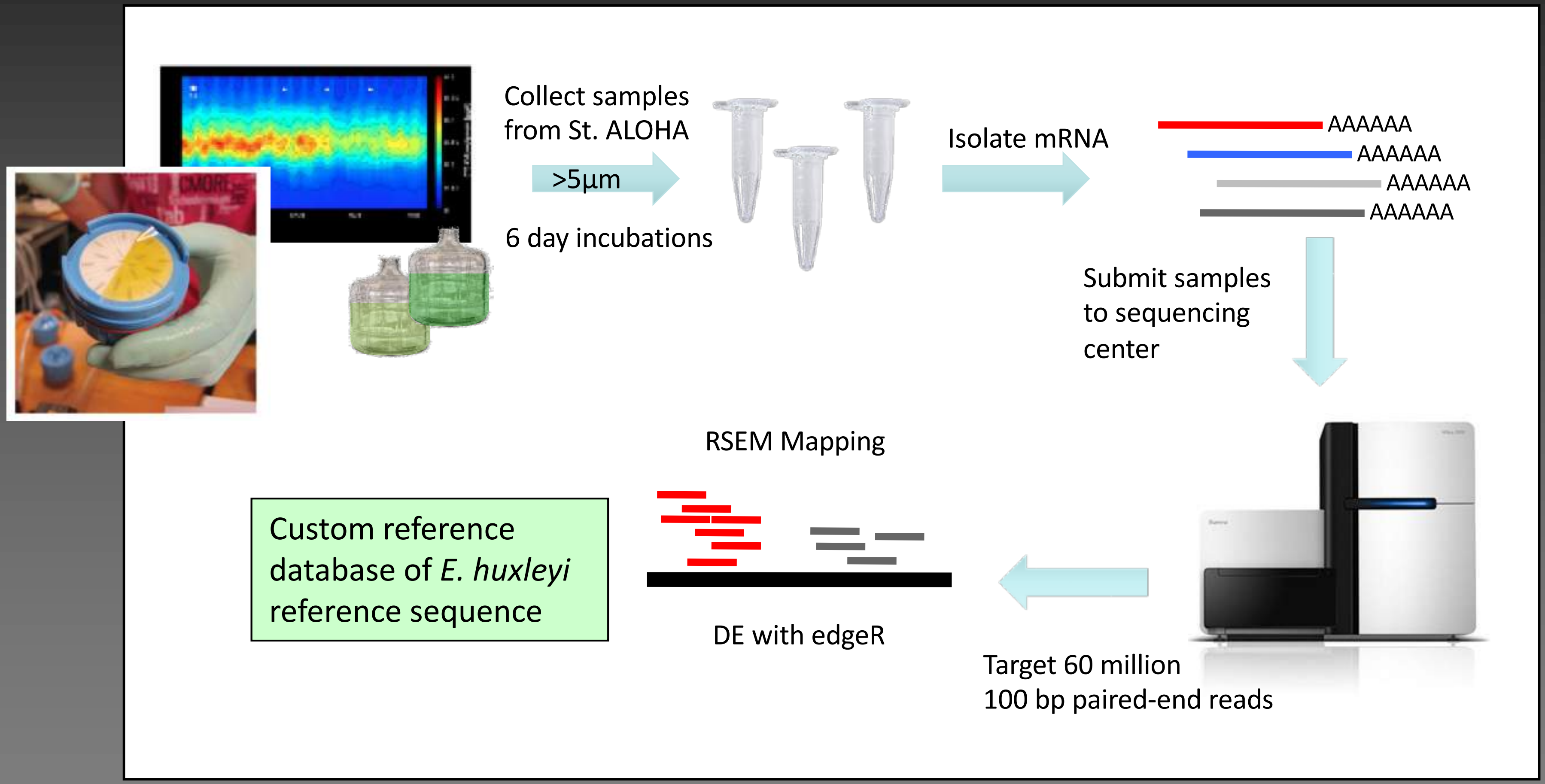
-N

+P

N addition


DN-1
C-1

Sampling and workflow



Collect samples
from St. ALOHA

>5µm

6 day incubations

Isolate mRNA

AAAAAA
AAAAAA
AAAAAA
AAAAAA

Submit samples
to sequencing
center

Target 60 million
100 bp paired-end reads

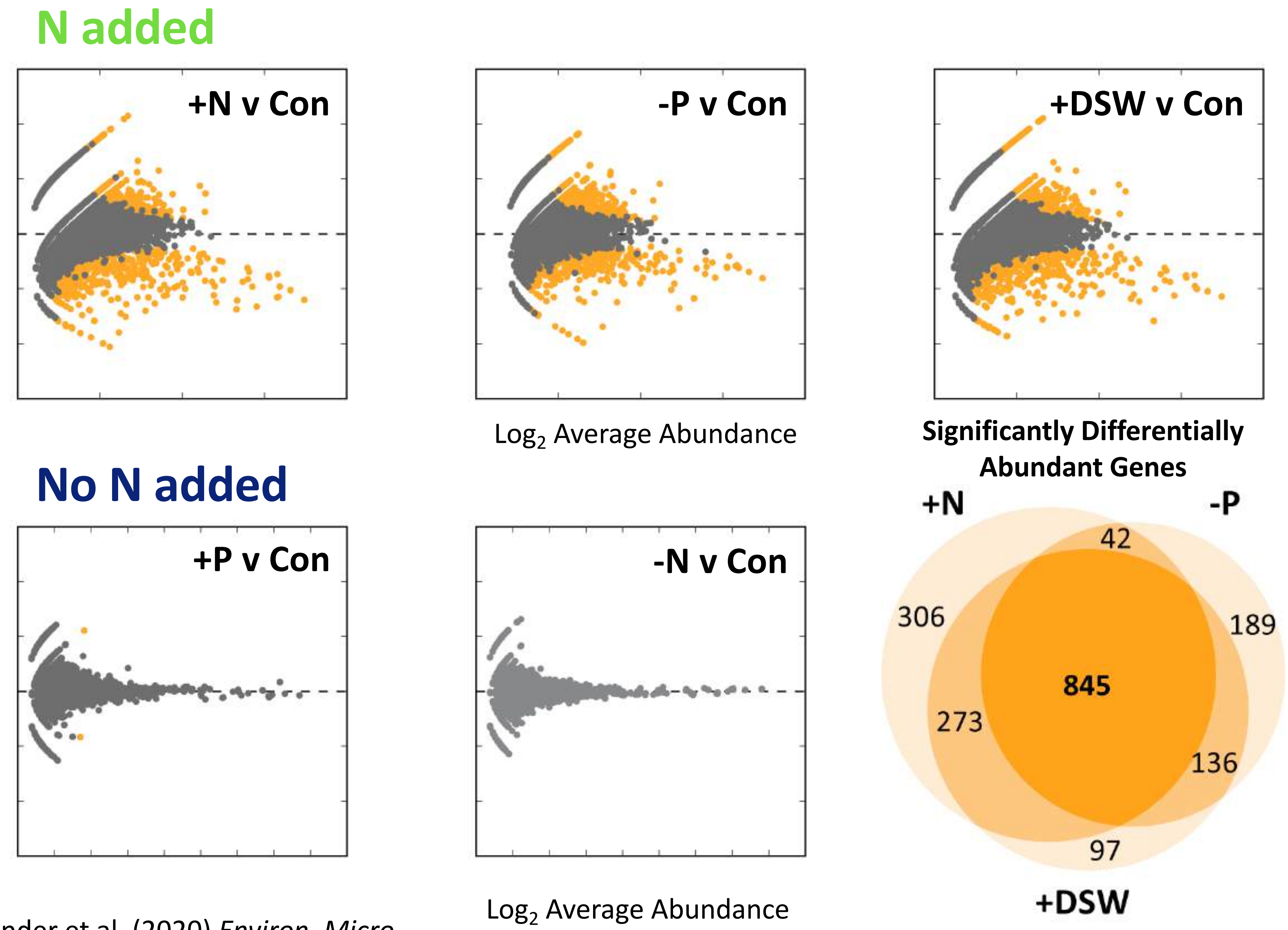
RSEM Mapping

DE with edgeR

Custom reference
database of *E. huxleyi*
reference sequence

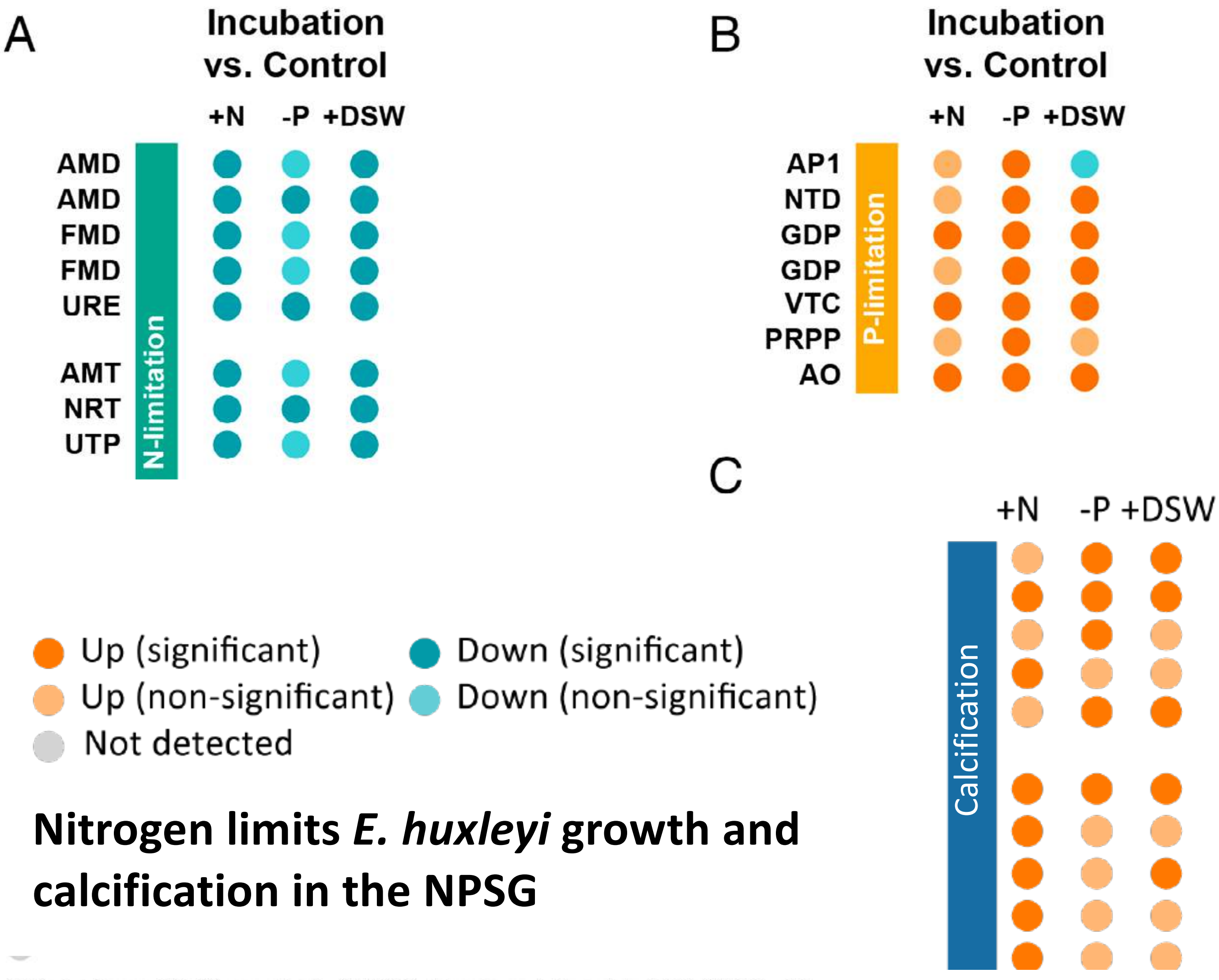
E. huxleyi physiological response suggests N control

Log$_2$ Fold Change



Alexander et al. (2020) *Environ. Micro*.

Metatranscriptome identifies physiological patterns

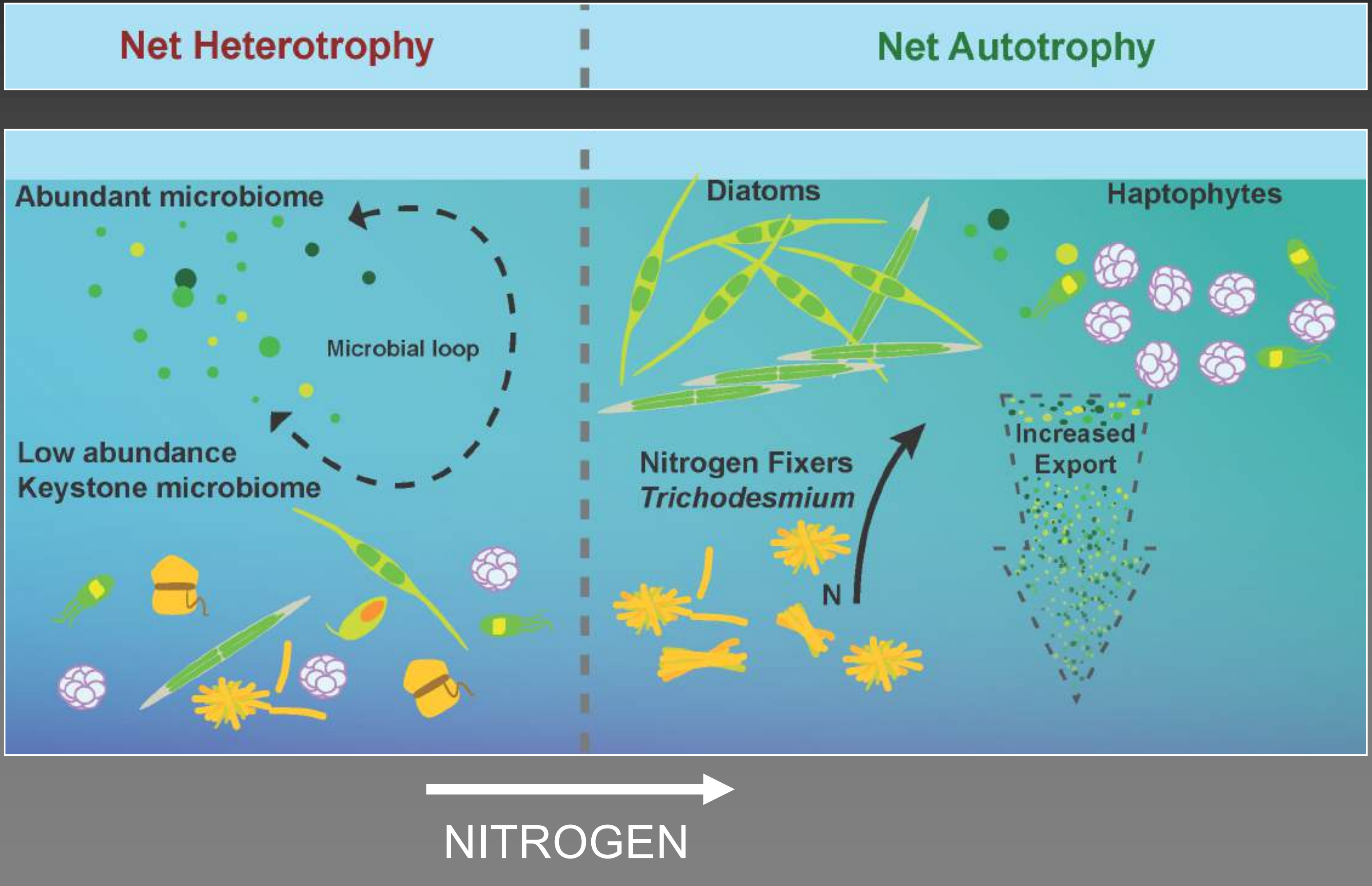


Nitrogen limits *E. huxleyi* growth and calcification in the NPSG

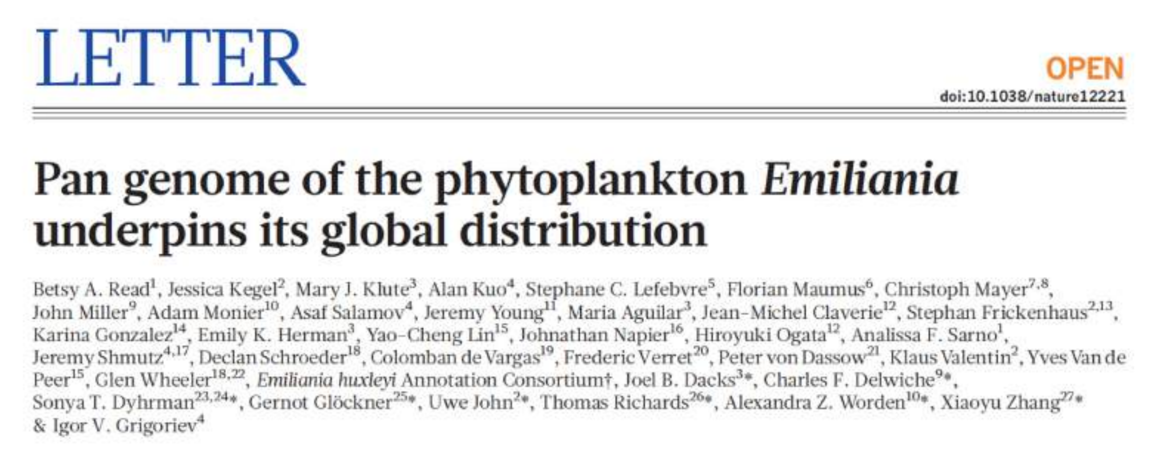
* Data from McKew et al. (2015) *Environ Microbiol* 17:4050–62.

Alexander et al. (2020) *Environ. Micro.*

Nitrogen is a driver of *E. huxleyi* dynamics



Emiliania huxleyi has a broad distribution of diverse ecotypes

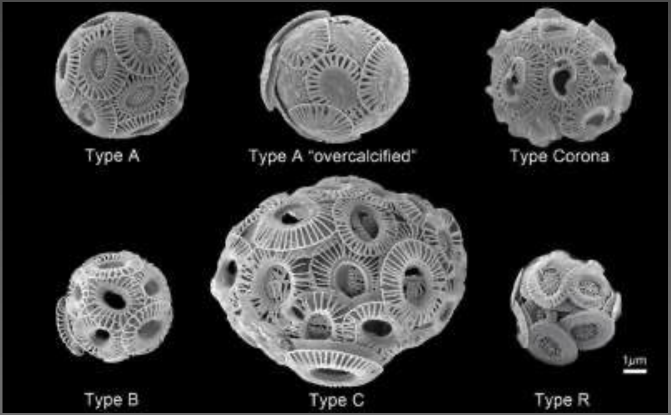

LETTER

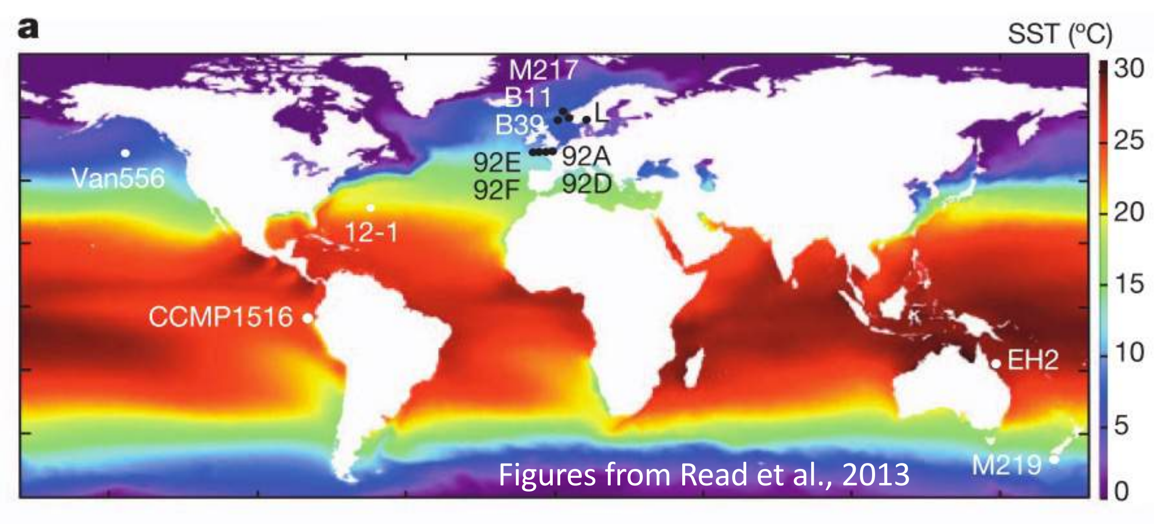
OPEN

doi:10.1038/nature12221

Pan genome of the phytoplankton *Emiliania* underpins its global distribution

Betsy A. Read[1], Jessica Kegel[2], Mary J. Klute[3], Alan Kuo[4], Stephane C. Lefebvre[5], Florian Maumus[6], Christoph Mayer[7,8], John Miller[9], Adam Monier[10], Asaf Salamov[4], Jeremy Young[11], Maria Aguilar[3], Jean-Michel Claverie[12], Stephan Frickenhaus[2,13], Karina Gonzalez[14], Emily K. Herman[3], Yao-Cheng Lin[15], Johnathan Napier[16], Hiroyuki Ogata[12], Analissa F. Sarno[1], Jeremy Shmutz[4,17], Declan Schroeder[18], Colomban de Vargas[19], Frederic Verret[20], Peter von Dassow[21], Klaus Valentin[2], Yves Van de Peer[15], Glen Wheeler[18,22], *Emiliania huxleyi* Annotation Consortium†, Joel B. Dacks[3]*, Charles F. Delwiche[9]*, Sonya T. Dyhrman[23,24]*, Gernot Glöckner[25]*, Uwe John[2]*, Thomas Richards[26]*, Alexandra Z. Worden[10]*, Xiaoyu Zhang[27]* & Igor V. Grigoriev[4]





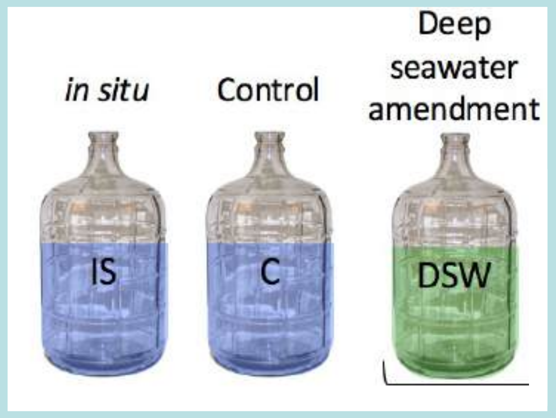
Figures from Read et al., 2013

Cultured strains are highly diverse - isolated from a broad temperature range and displaying considerable physiological diversity

Baas Becking hypothesis

“Everything is every where, but the environment selects”

Does the variable component of the pan genome underpin which strain dominates a given environment?



Sequences from *E. huxleyi* strains

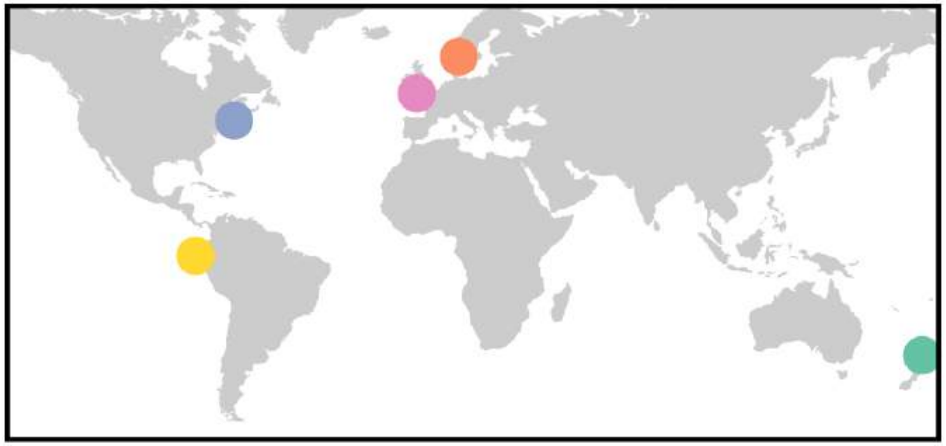
- ● CCMP1516
- ● CCMP379
- ● CCMP374
- ● CCMP370
- ● PLYM219

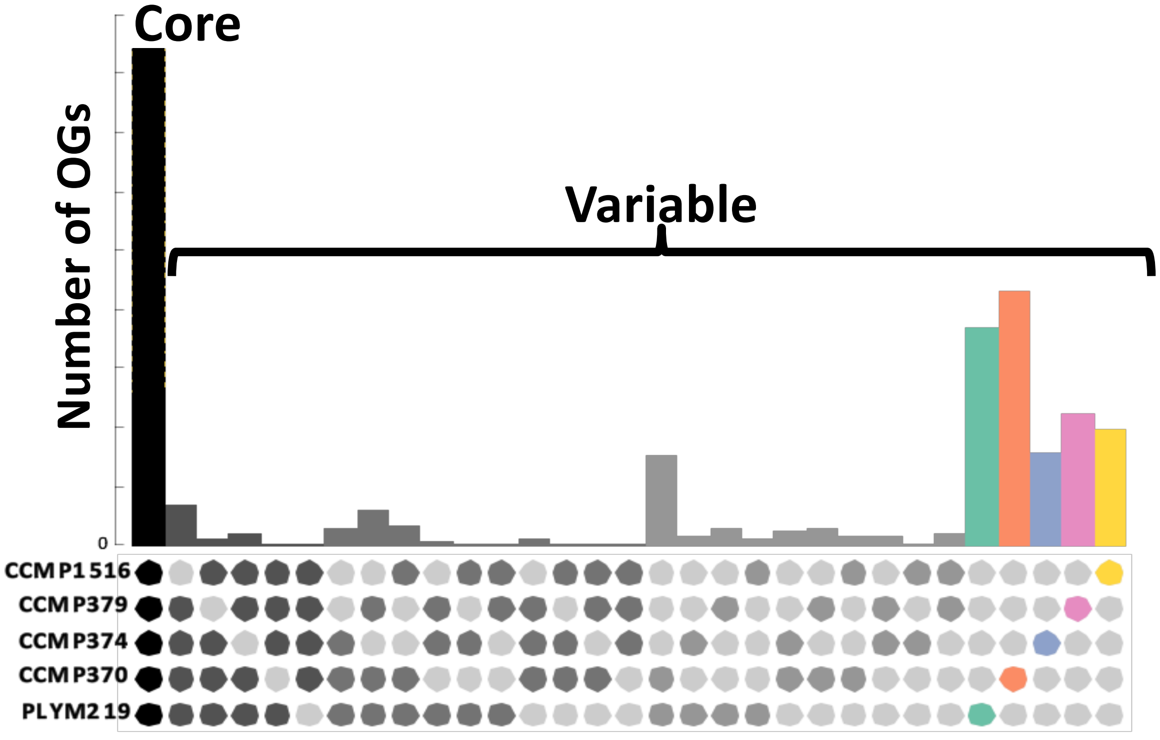
↓

Cluster into orthologous groups to identify core and variable OGs
(OrthoMCL)

↓

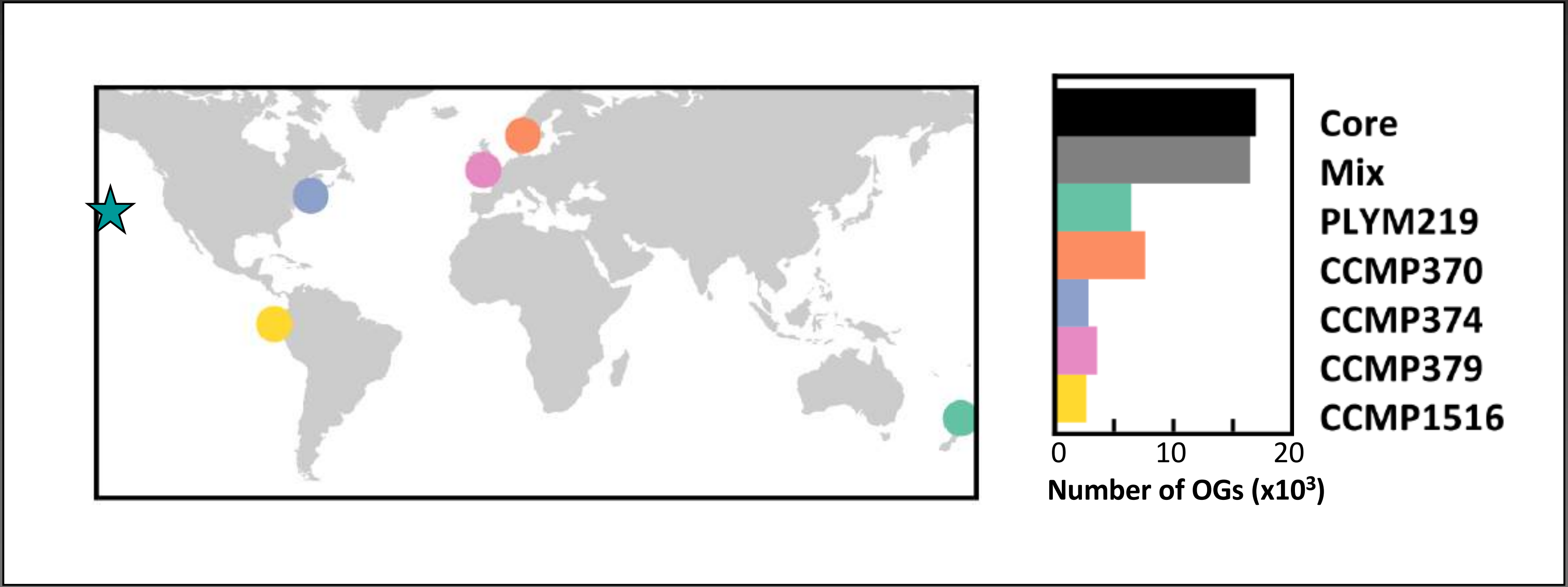
Map field data and examine shifts in expression
(RSEM, edgeR)





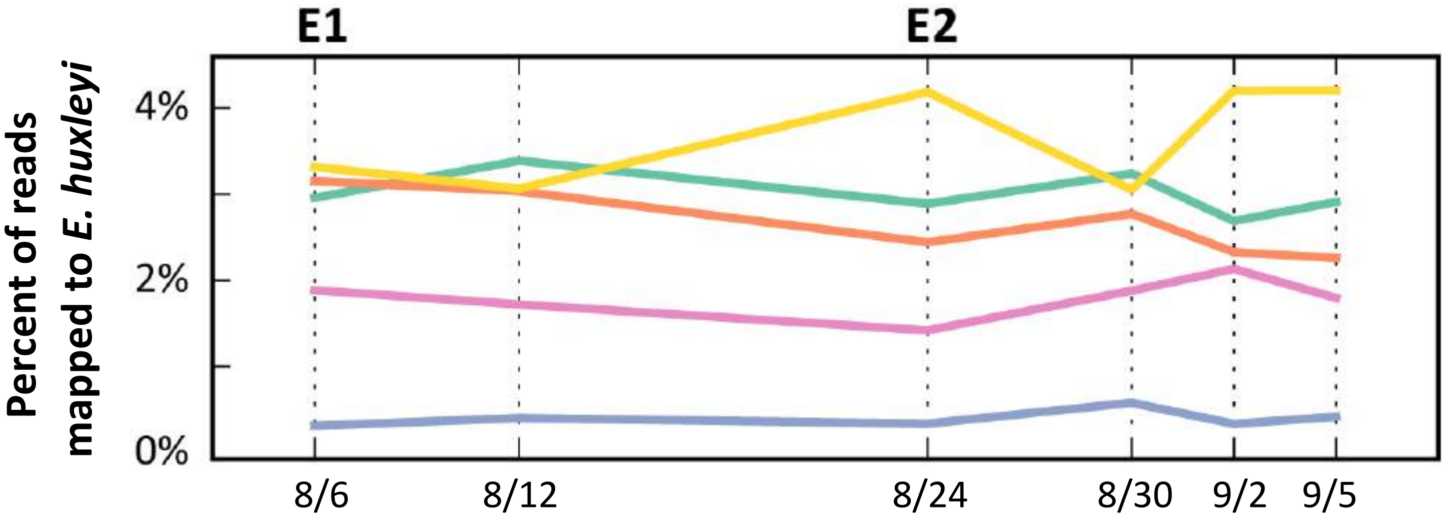
Alexander et al. unpub.

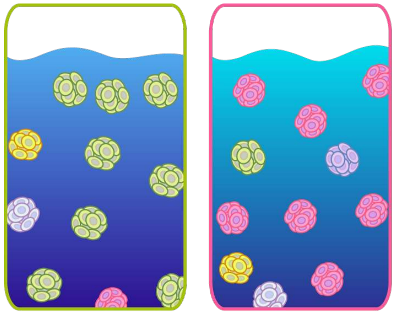
Tracking orthologous groups (OG) at St. ALOHA

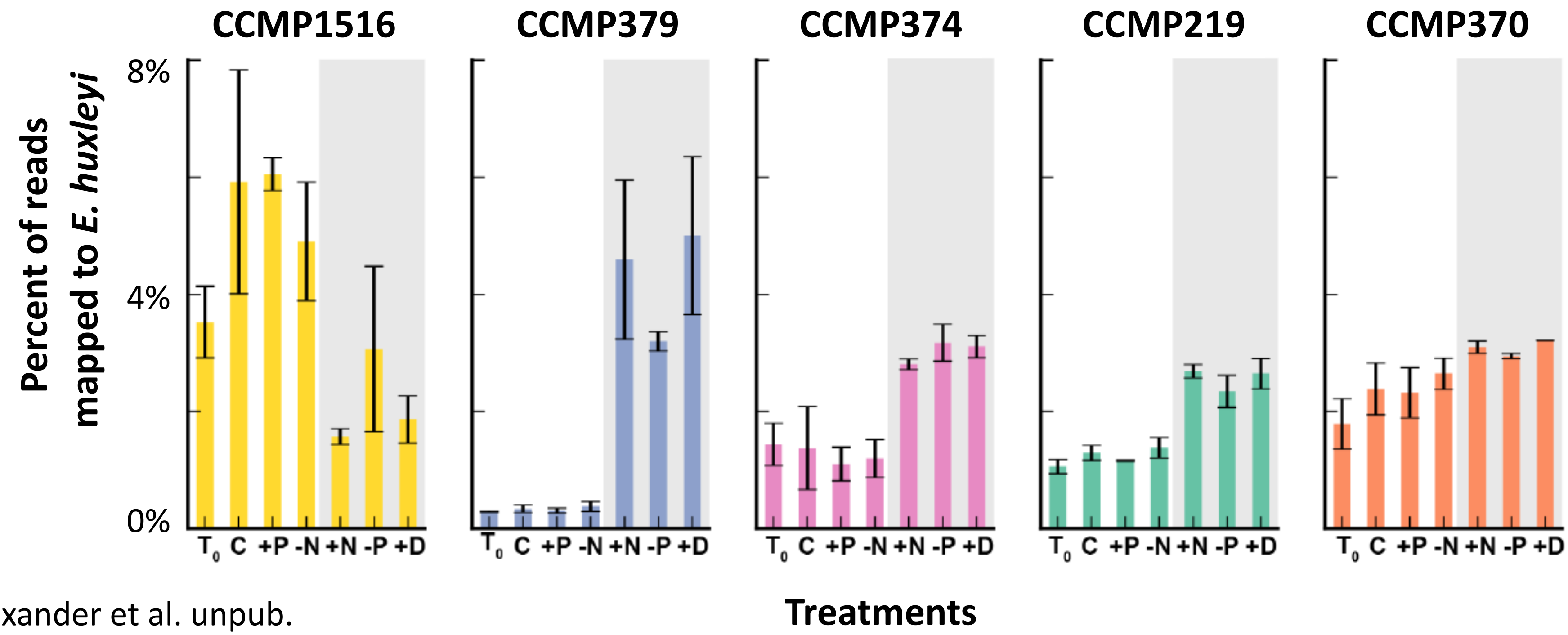


Use diagnostic OGs to track strain distribution

Strain distribution *in situ* and in the incubations

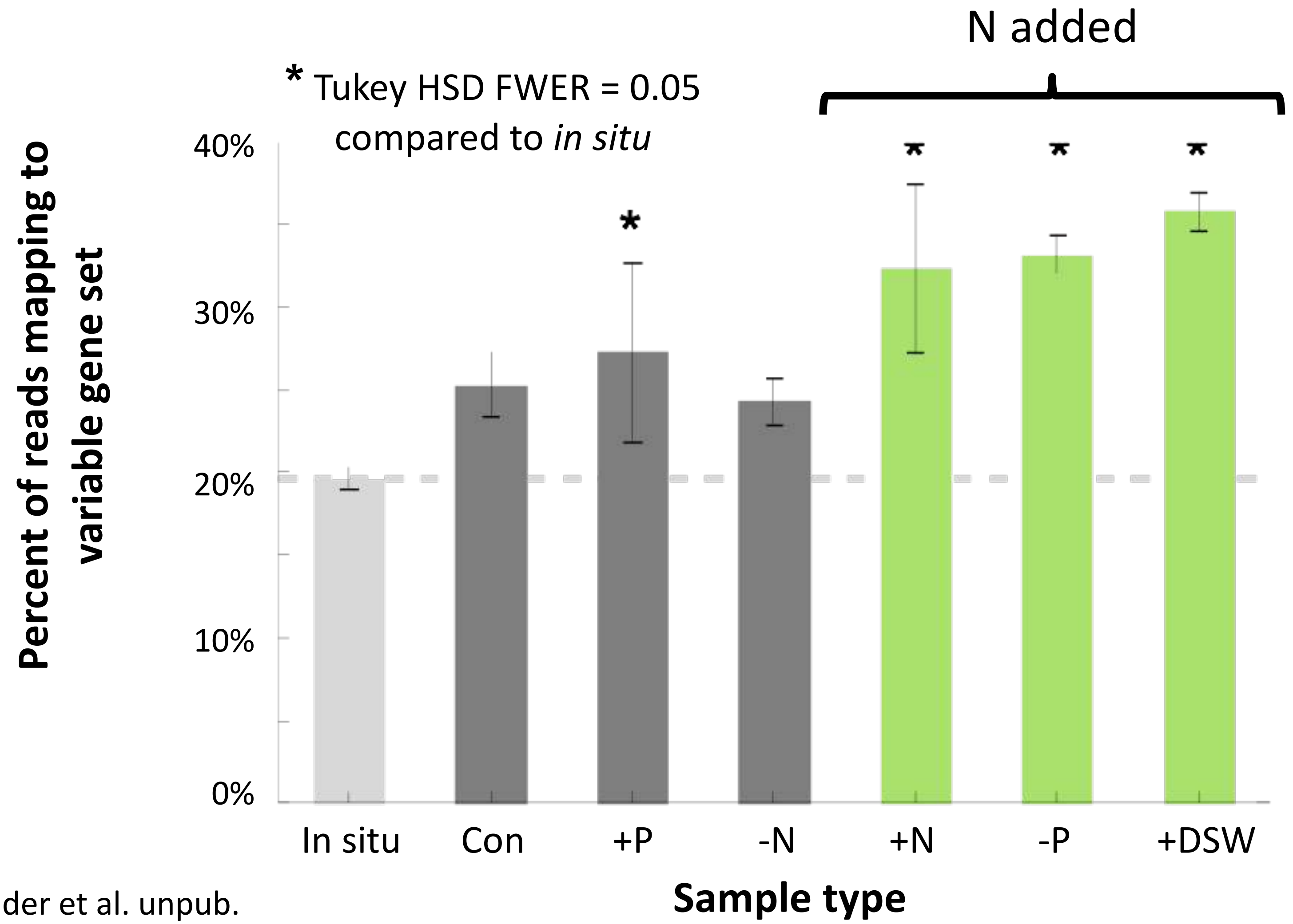






Alexander et al. unpub.

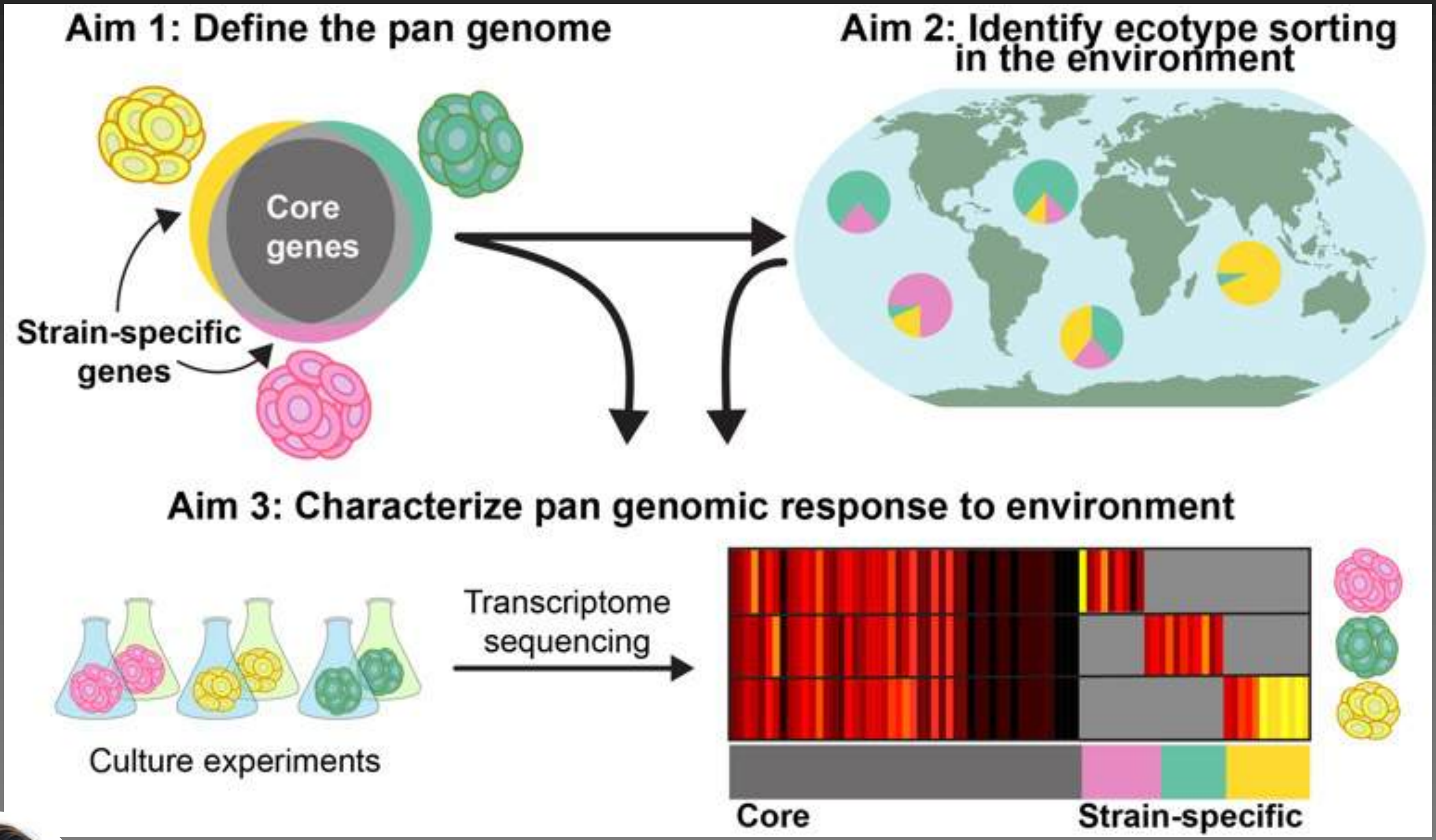
Expression of the variable portion of the pan genome



Alexander et al. unpub.

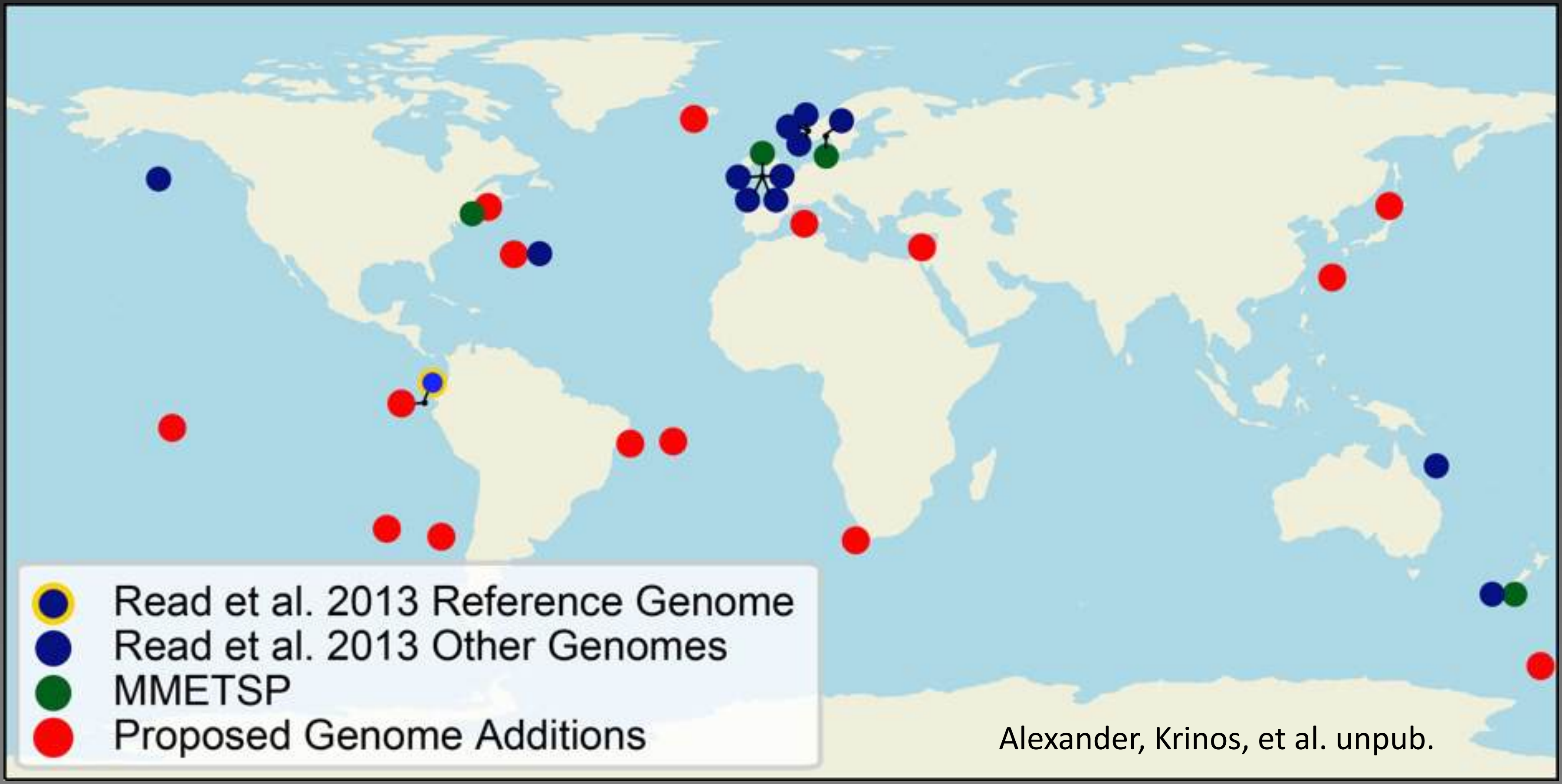
Expression of variable gene set significantly increased following N-addition

Ongoing work to define the pan genome



Harriet Alexander
https://alexanderlabwhoi.github.io/

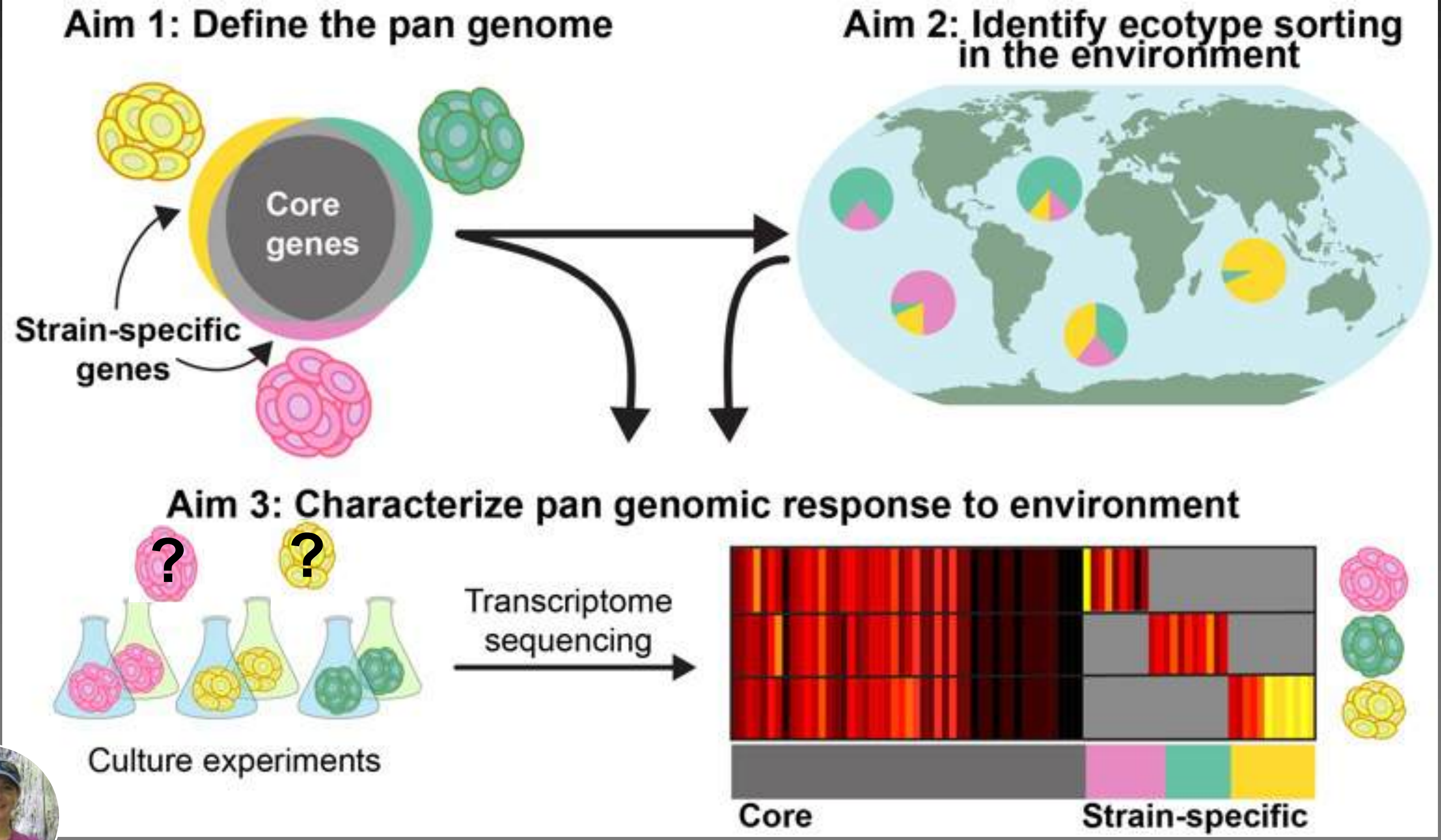
Expanding the pan genome



Alexander, Krinos, et al. unpub.

Many new whole genome sequences being finalized (Nanopore/Illumina) with RNASeq libraries to support gene prediction

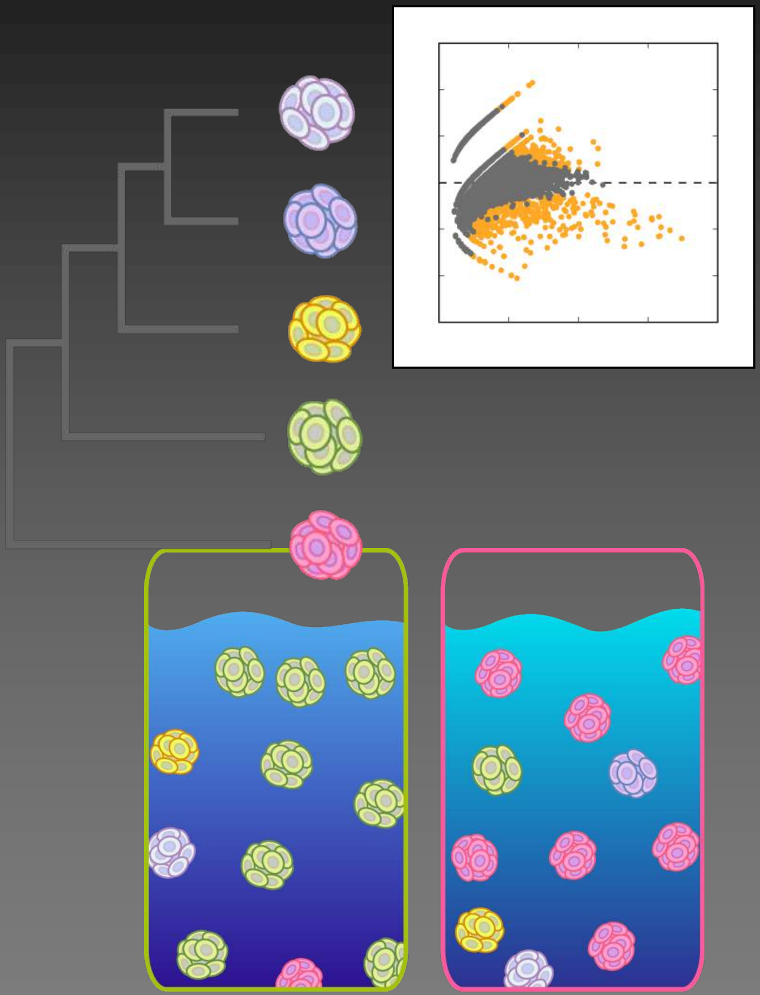
Ongoing work on to identify how environment influences ecotype distribution paired with competition studies



Haley et al. unpub.

Summary

- **Metatranscriptome RNA-Seq**:
- Species-specific analyses are possible even for low abundance keystone populations
- Nitrogen limits *E. huxleyi* activity, and influences calcification
- Variable genes in the pan genome may influence ecotype strain sorting.
- **Nanopore/Illumina WGS:**
- Expands a pan genomic framework for ecotype studies



Core questions… adventures in eukaryote metatranscriptomics

Limitation

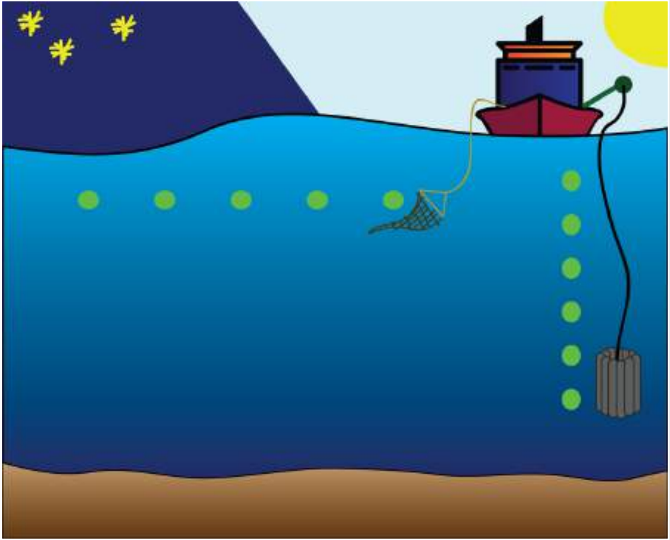
- **Nitrogen limits carbon export and cycling in a keystone haptophyte**
- **The environment *may* select**

Competition

- A day in the life – tracing competitive interactions with multi-’omics

HOE:DYLAN

HOE:MAHALO



Tracking phytoplankton physiological ecology

Ecosystem function and biogeochemistry

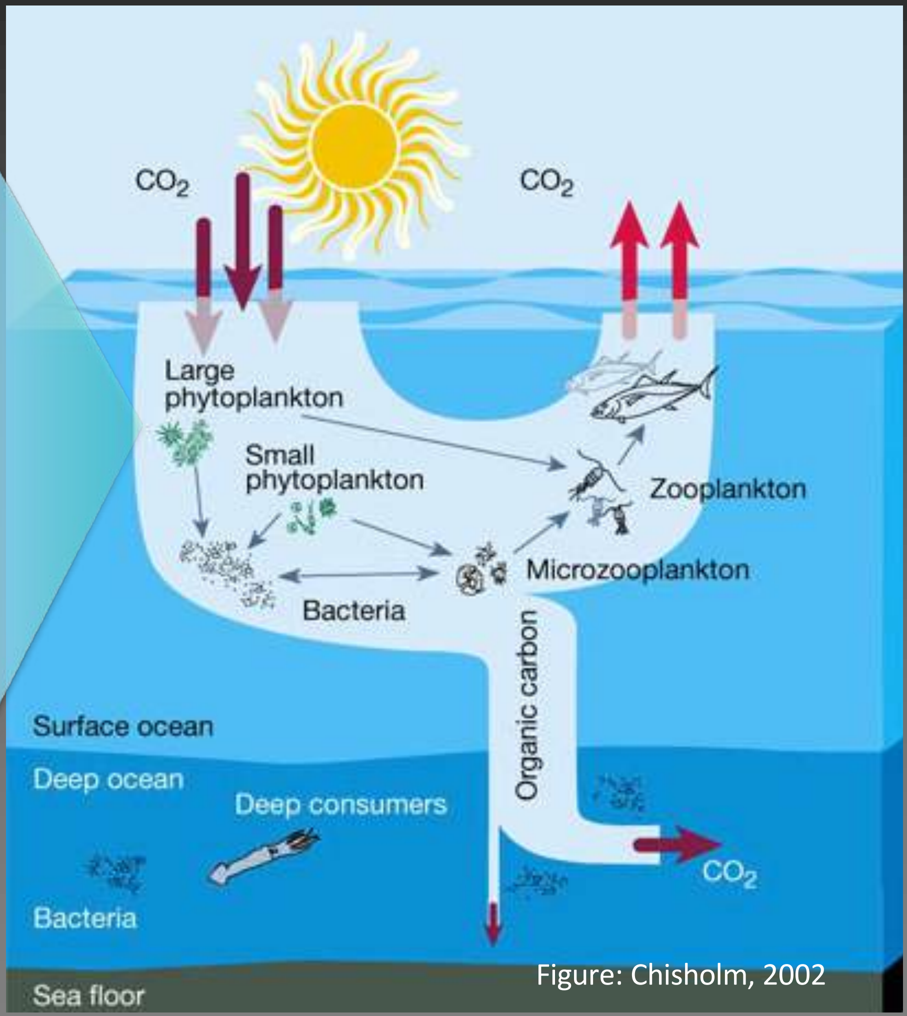


Figure: Chisholm, 2002

How do resources drive phytoplankton distributions and activities?

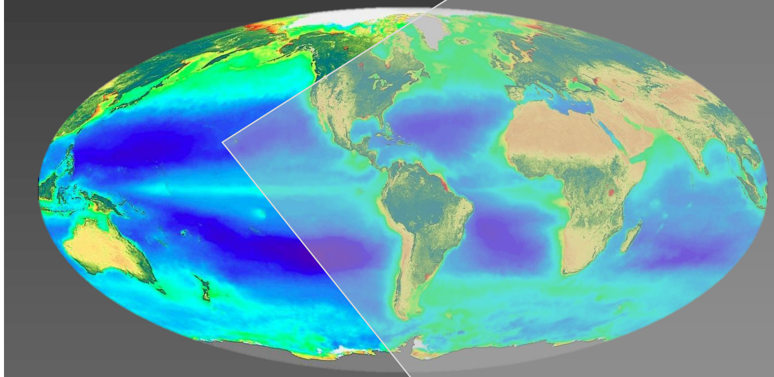
Hawaii Ocean Experiment: Diel time-series (MAHALO)

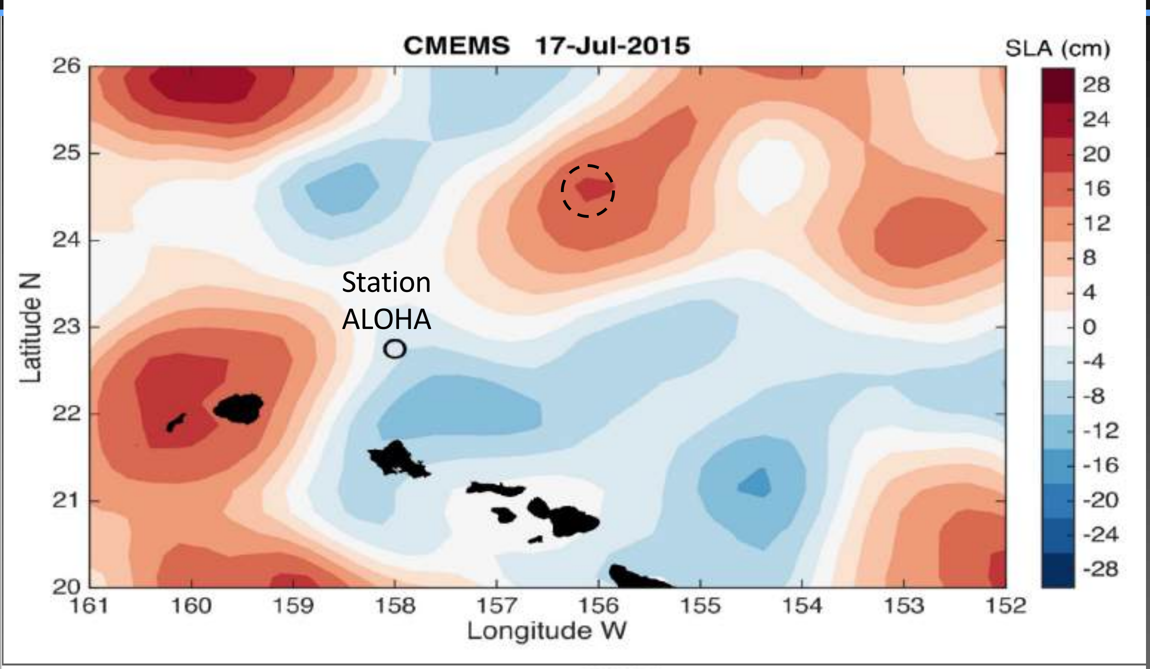
Figure 1 R/V Kilo Moana conducting one of the diel study CTD casts in 2015 (photo: Tara Clemente, taken from the R/V Ka'imikai-O-Kanaloa)

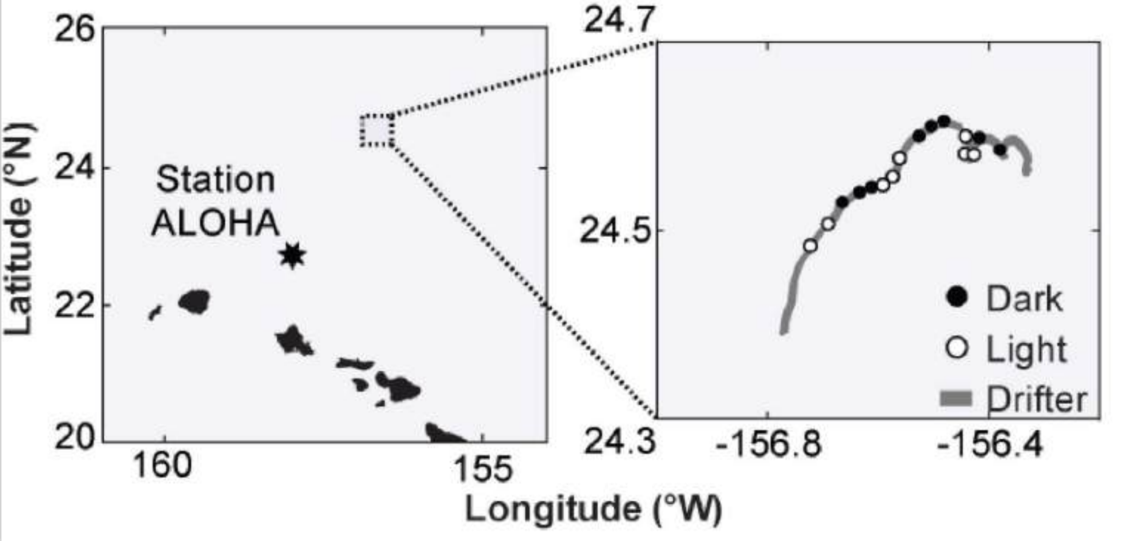
SCOPE
SIMONS COLLABORATION ON OCEAN PROCESSES AND ECOLOGY
Inputs
Microbes
Models
Outputs
SCOPE.SOEST.HAWAII.EDU

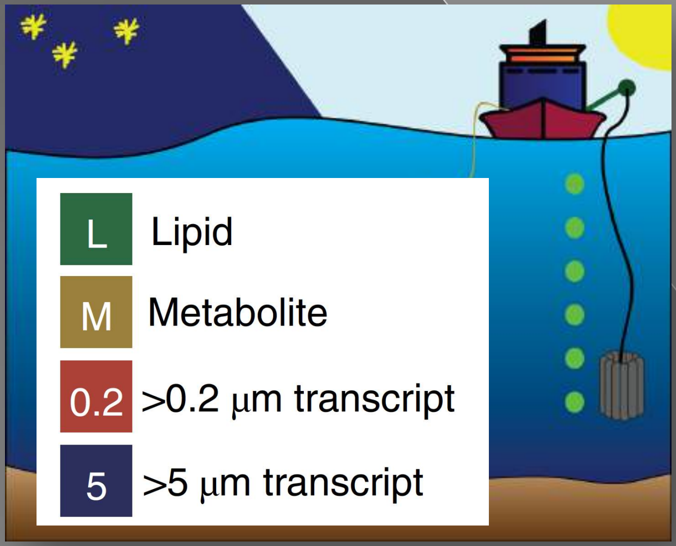
MAHALO: Special thanks to Daniel Muratore and Joshua Weitz for synthesis efforts

Sampling light driven responses in the NPSG

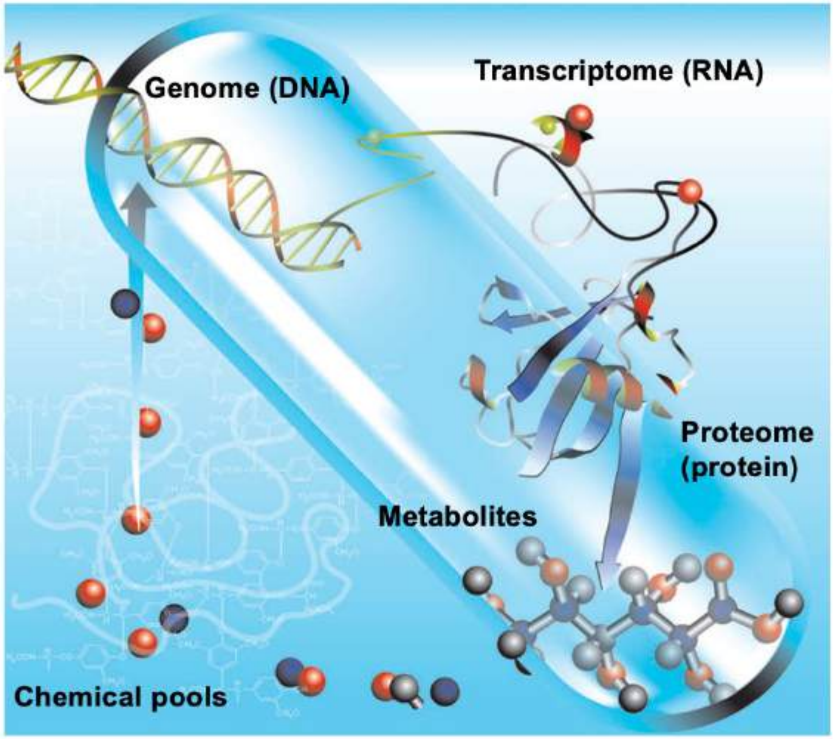








Jointly applying new multi ‘omic techniques



Kujawinski et al. (2011) *Annual Review Marine Science*

L Lipid

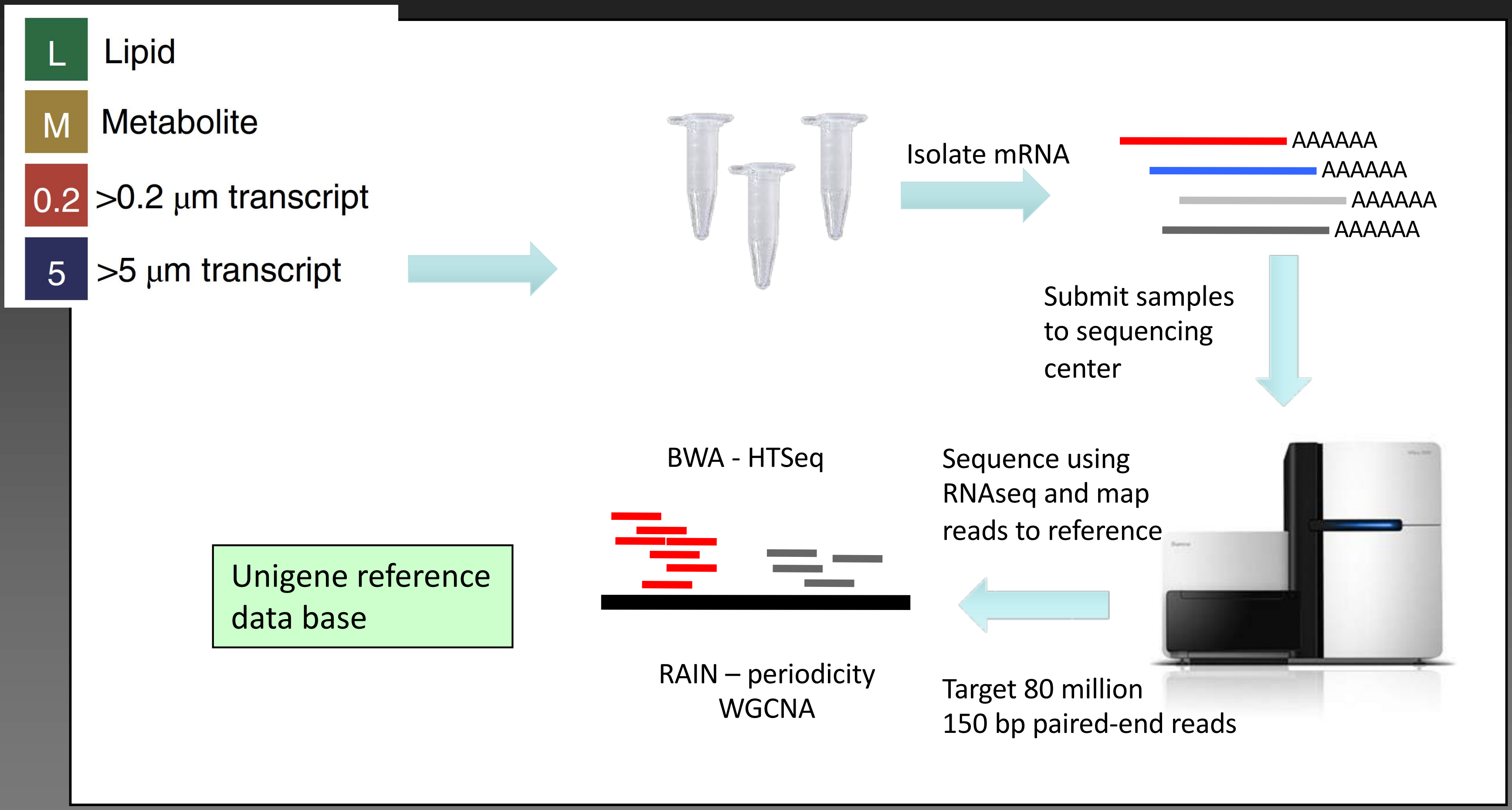
M Metabolite

0.2 >0.2 µm transcript

5 >5 µm transcript

Sampling every 4hrs for many days ….

Sampling and pipeline



L Lipid

M Metabolite

0.2 >0.2 µm transcript

5 >5 µm transcript

Isolate mRNA

Submit samples to sequencing center

Sequence using RNAseq and map reads to reference

Target 80 million 150 bp paired-end reads

BWA - HTSeq

RAIN – periodicity
WGCNA

Unigene reference data base

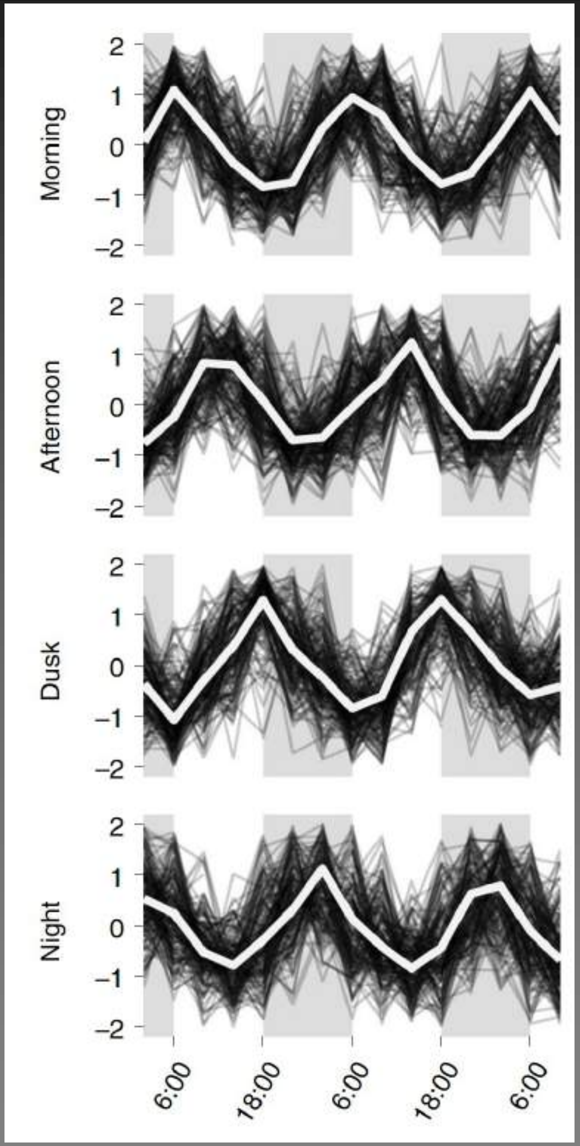
https://omictools.com/rhythmicity-analysis-incorporating-nonparametric-methods-tool

https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-9-559

Visualizing diel patterns

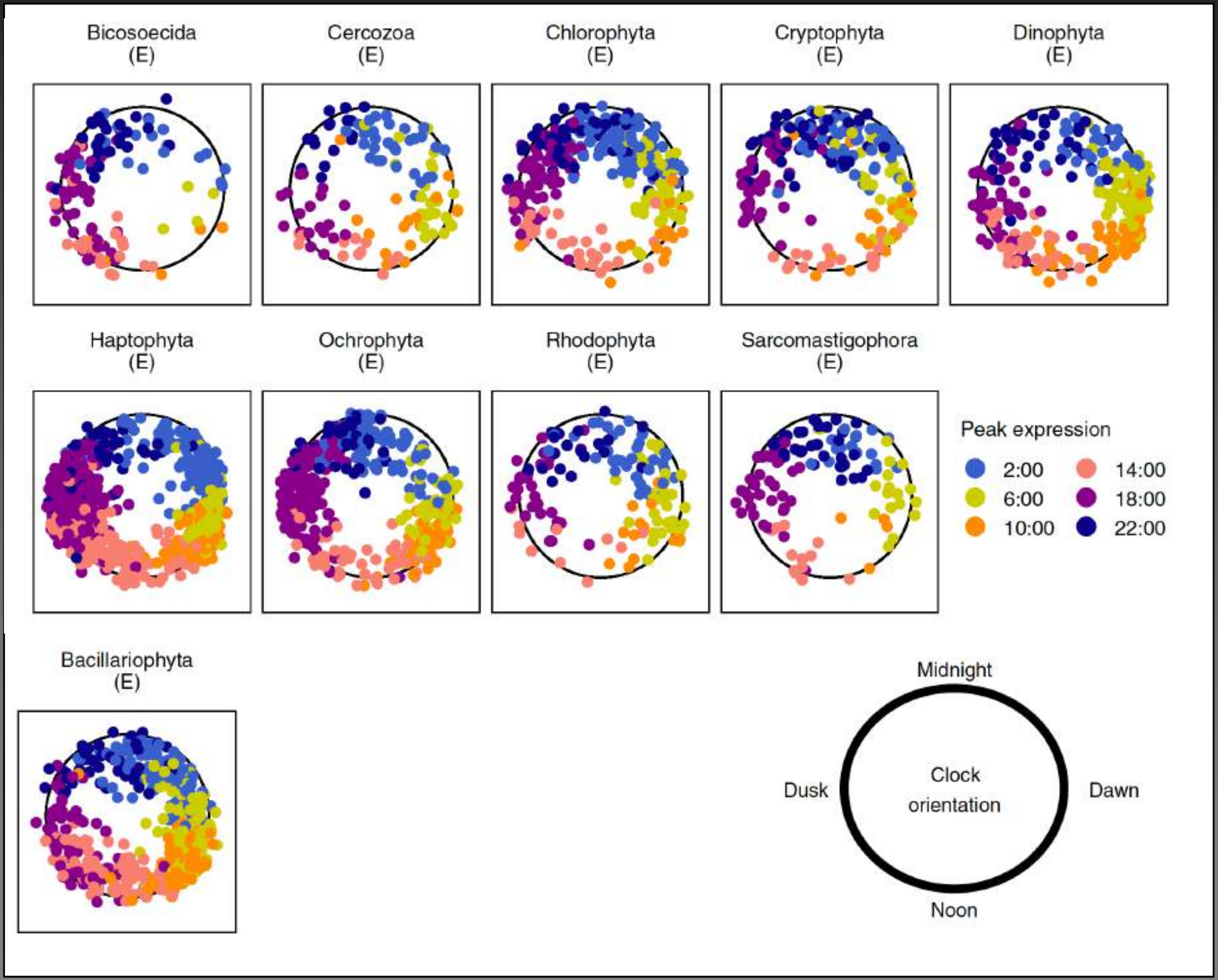
- L Lipid
- M Metabolite
- 0.2 >0.2 µm transcript
- 5 >5 µm transcript

~65,000 analyte time series



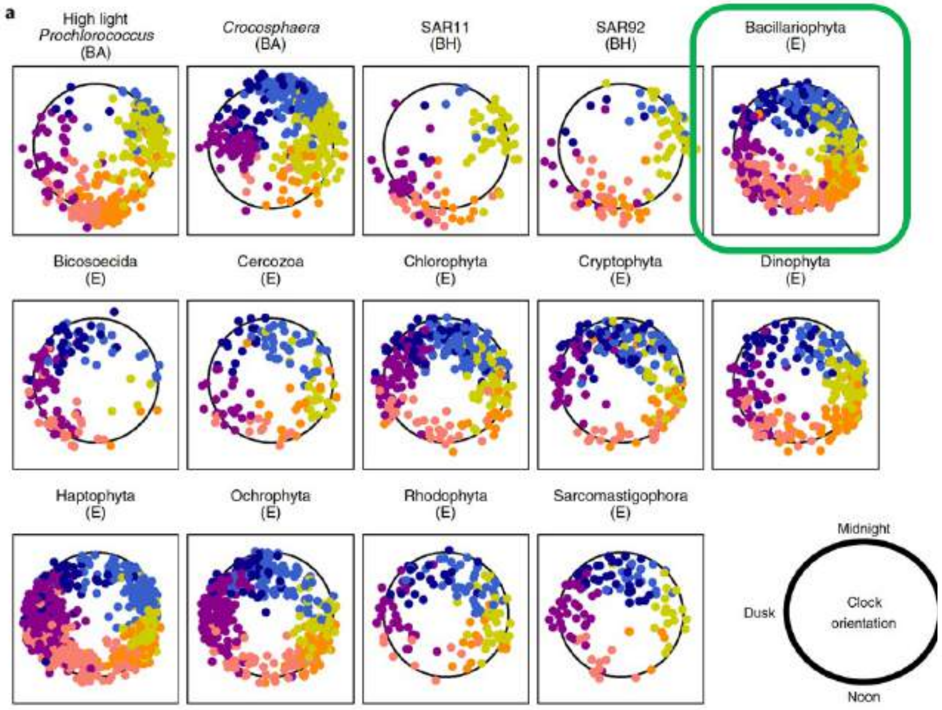
Muratore, Boysen, Harke et al. (2022) *Nature Ecol Evo*.

Different taxa have distinct diel transcription patterns

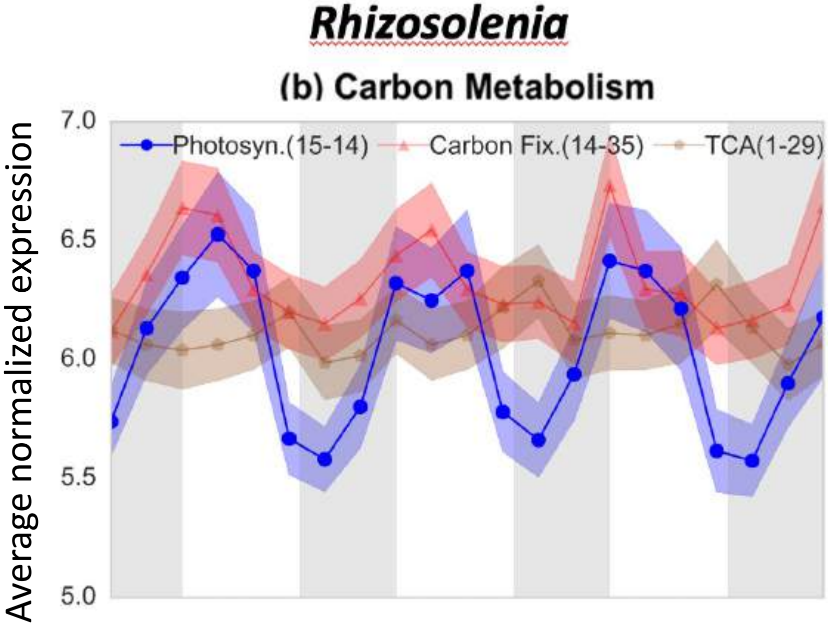


Muratore, Boysen, Harke et al. (2022) *Nature Ecol Evo*.

Carbon fixation and respiration are highly coordinated



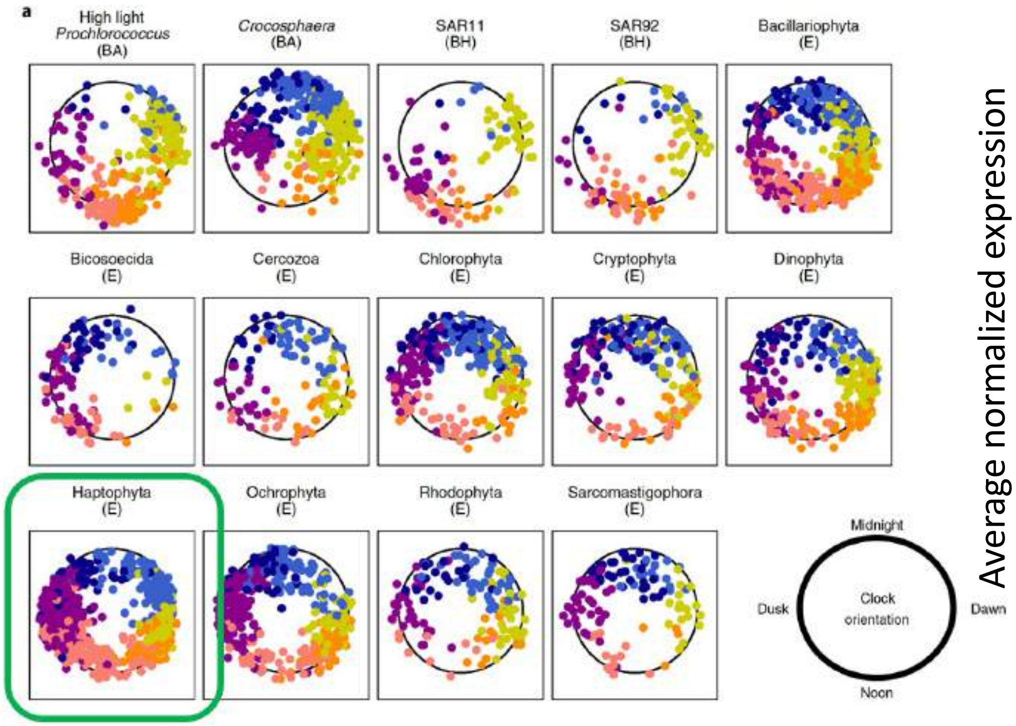
Muratore, Boysen, Harke et al. (2022) *Nature Ecol. Evol.*


Rhizosolenia

(b) Carbon Metabolism

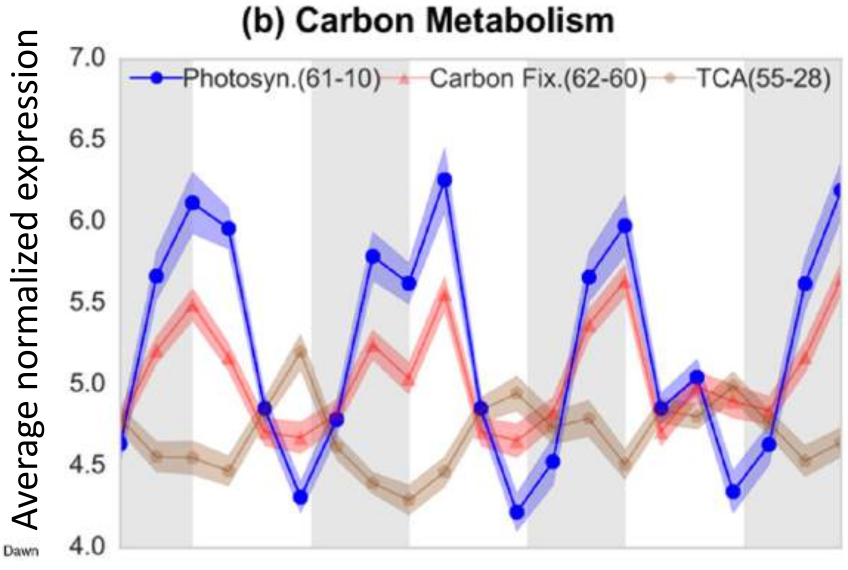

Harke et al. 2019 *ISME J*

Carbon fixation and respiration are highly coordinated



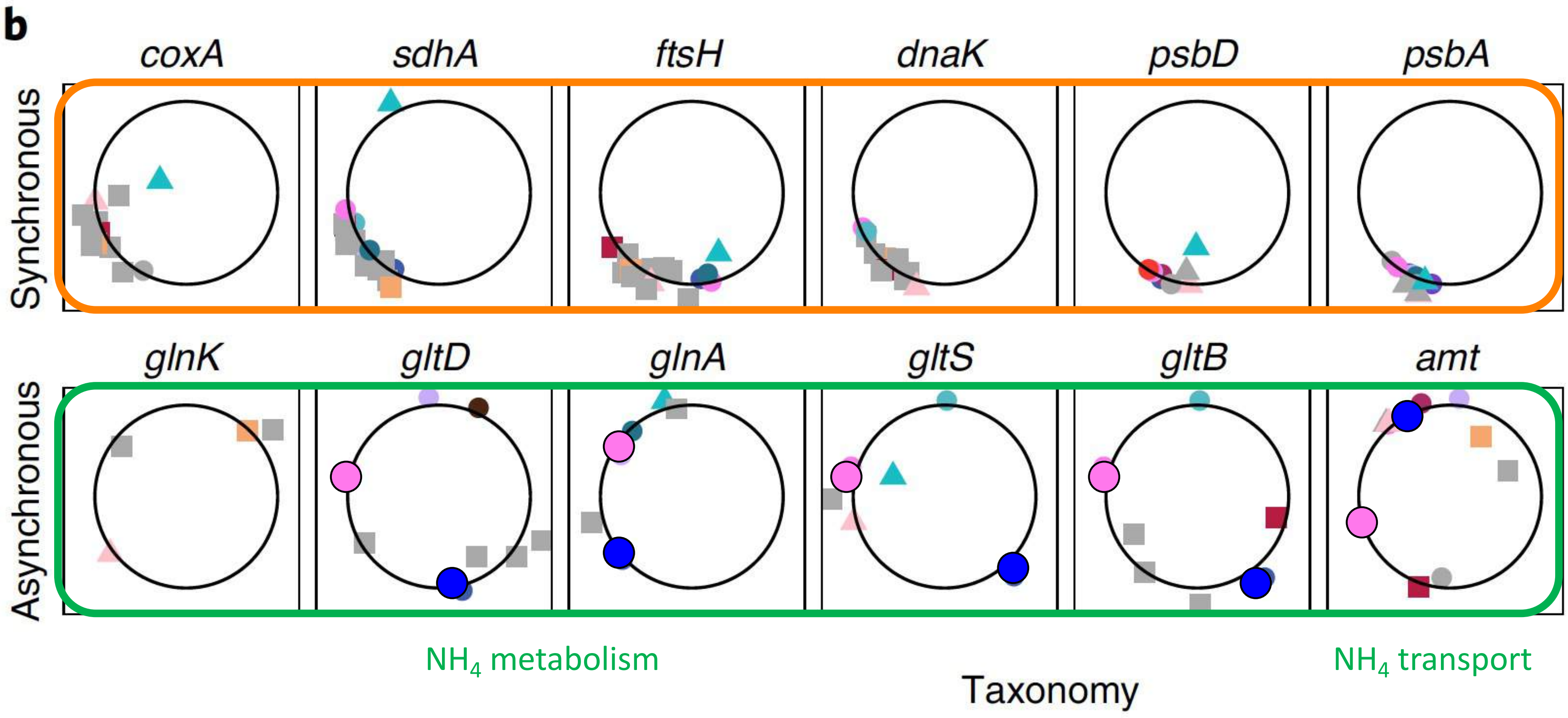
Muratore, Boysen, Harke et al. (2022) *Nature Ecol. Evol.*

Emiliania



Limón-Hernandez et al. (2020) *Environ. Micro.*

N metabolism functions are expressed asynchronously

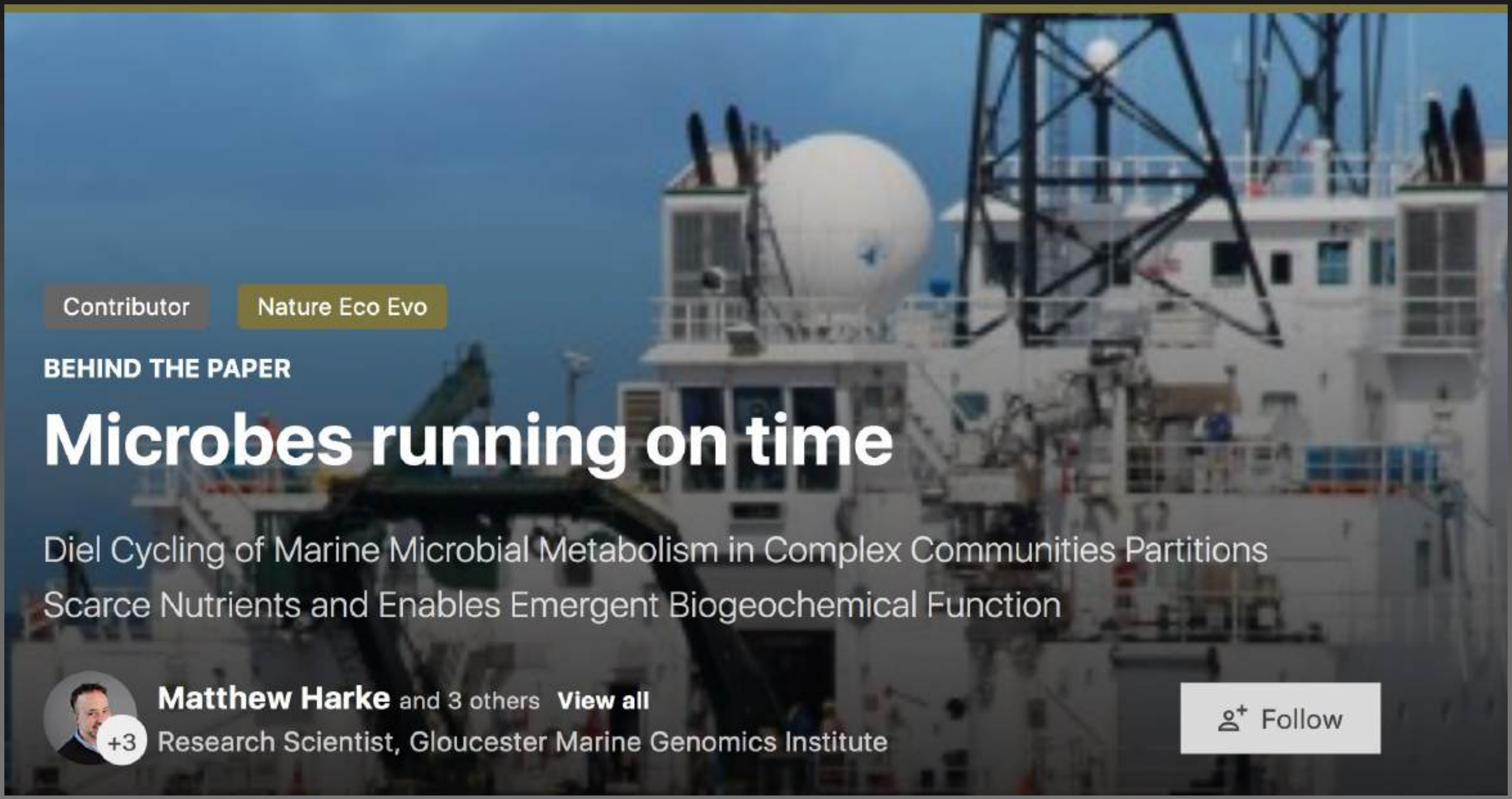


NH$_4$ metabolism

NH$_4$ transport

Asynchronous transcription of N uptake and metabolism between microeukaryotes and between taxonomic domains suggest competition for limiting resource is driving temporal niche partitioning

Muratore, Boysen, Harke et al. (2022) *Nature Ecol Evo*.



BEHIND THE PAPER

Microbes running on time

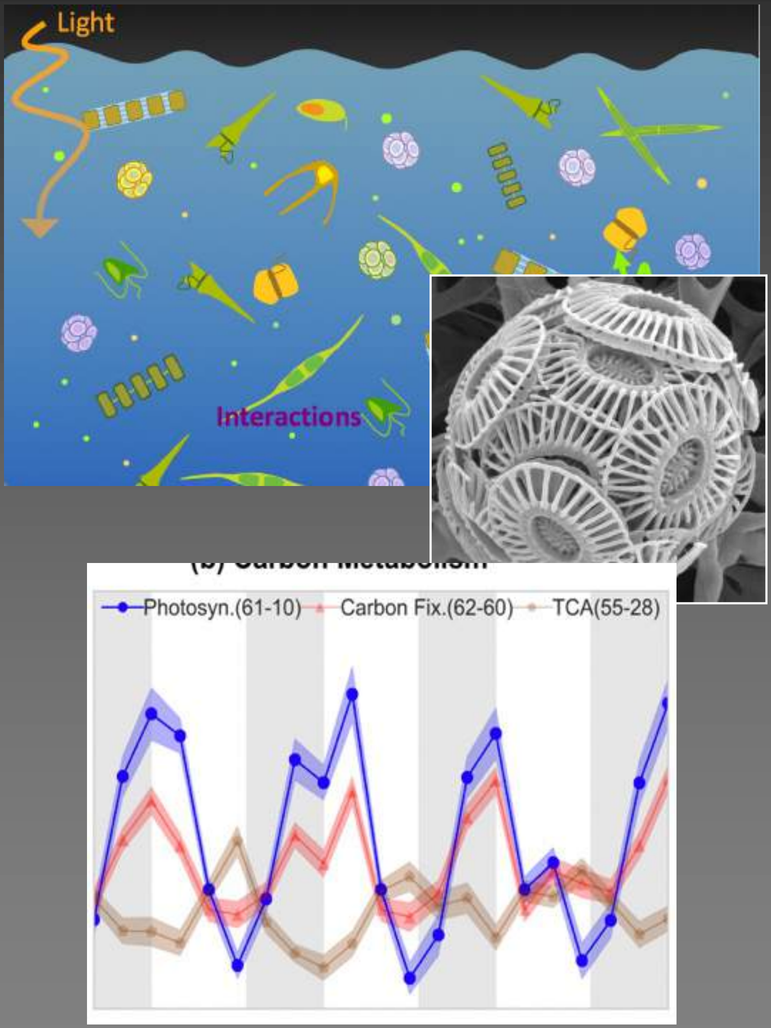
Diel Cycling of Marine Microbial Metabolism in Complex Communities Partitions Scarce Nutrients and Enables Emergent Biogeochemical Function

Matthew Harke and 3 others **View all**
Research Scientist, Gloucester Marine Genomics Institute

Muratore, Boysen, Harke et al. (2022) *Nature Ecol Evo*.

Summary

- **Metatranscriptome RNA-Seq**:
- Light drives synchronous metabolic cascades which underpin carbon cycling in the NPSG
- Light drives asynchronous metabolism of nitrogen – a limiting resource in the NPSG
- Temporal niche partitioning of N uptake and metabolism may decrease competition for limiting resources



Core questions… adventures in eukaryote metatranscriptomics

Limitation

- **Nitrogen limits carbon export and cycling in a keystone hapotophyte**

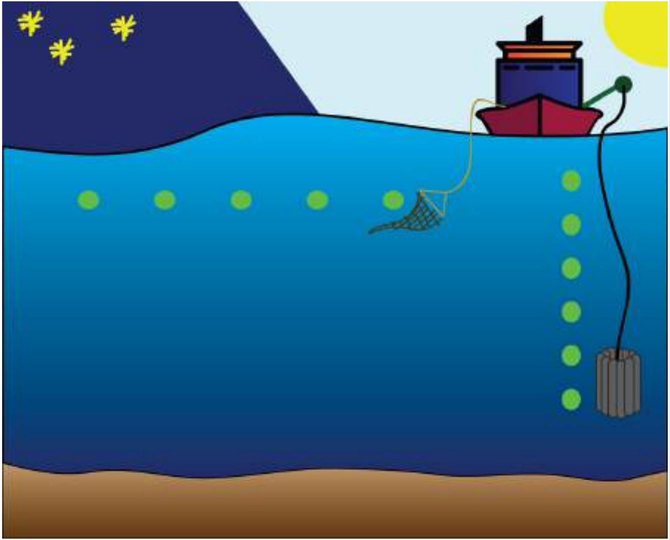
- **The environment *may* select**

Competition

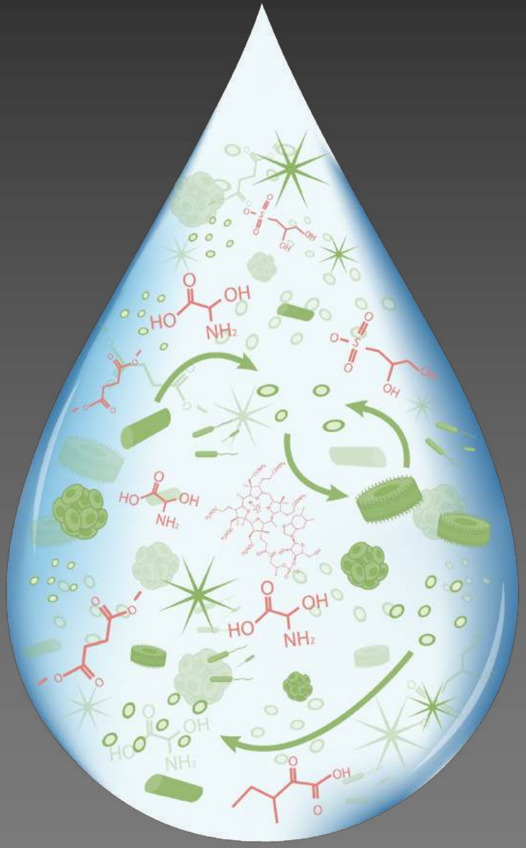
- **Temporal niche partitioning supports a complex community in a resource poor environment**

HOE:DYLAN

HOE:MAHALO



A vast chemical-microbe network underpins marine food webs and biogeochemical cycling

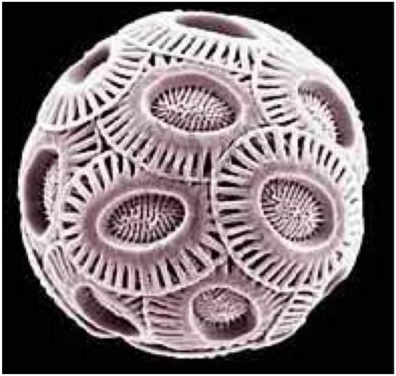


If we learn the rules for how the system operates – then we have a better chance of predicting future ocean responses

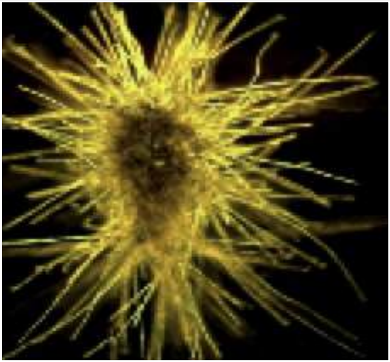
‘Omic-enabled advances allowing to query cells in their environment in a species-specific way

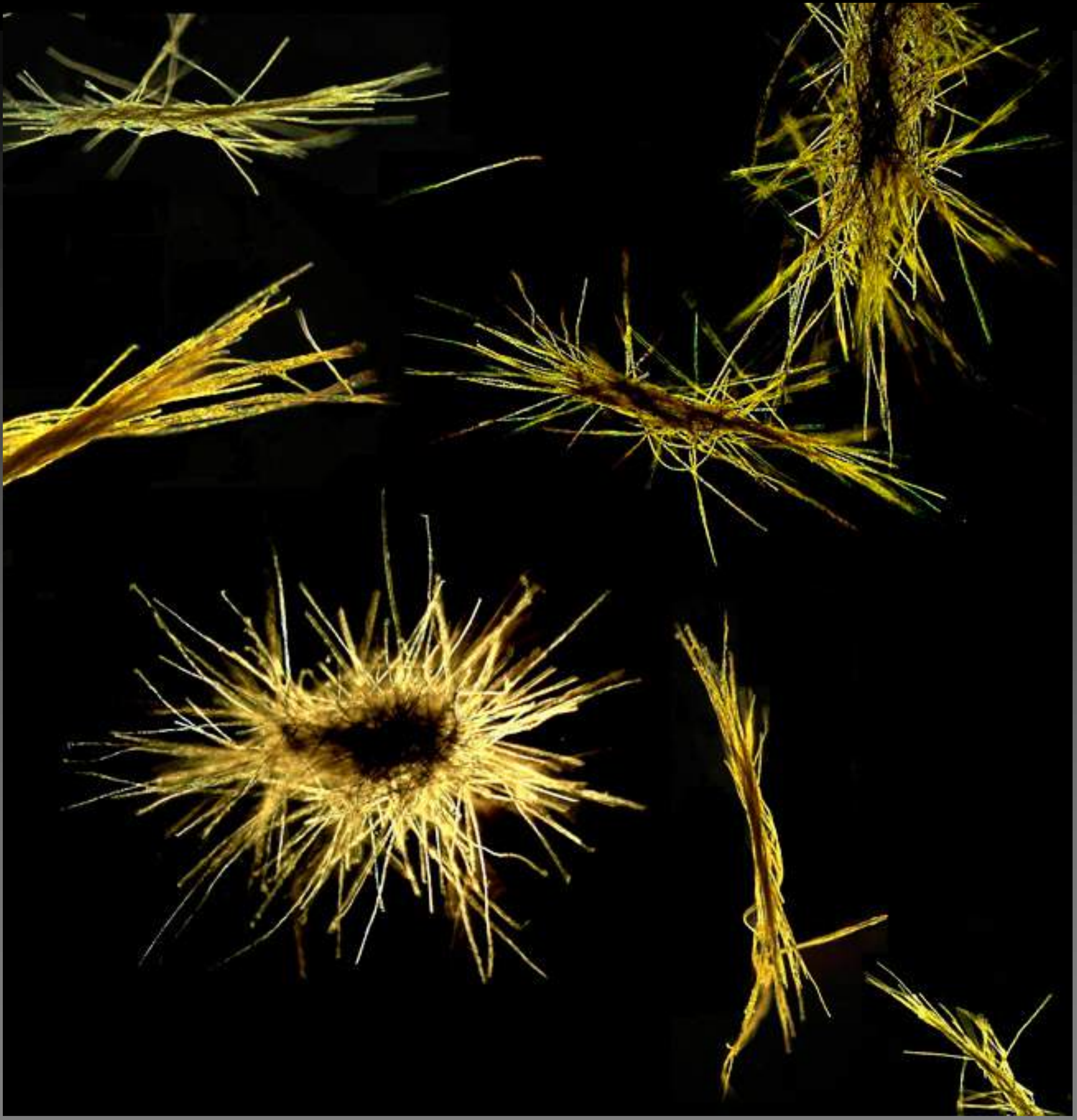
Vignettes

- Co-existing in a sea of competition: tracing drivers of phytoplankton physiological ecology in the open ocean



- Microbiome interactions: Tracking the microbiome of a keystone N_2 fixer

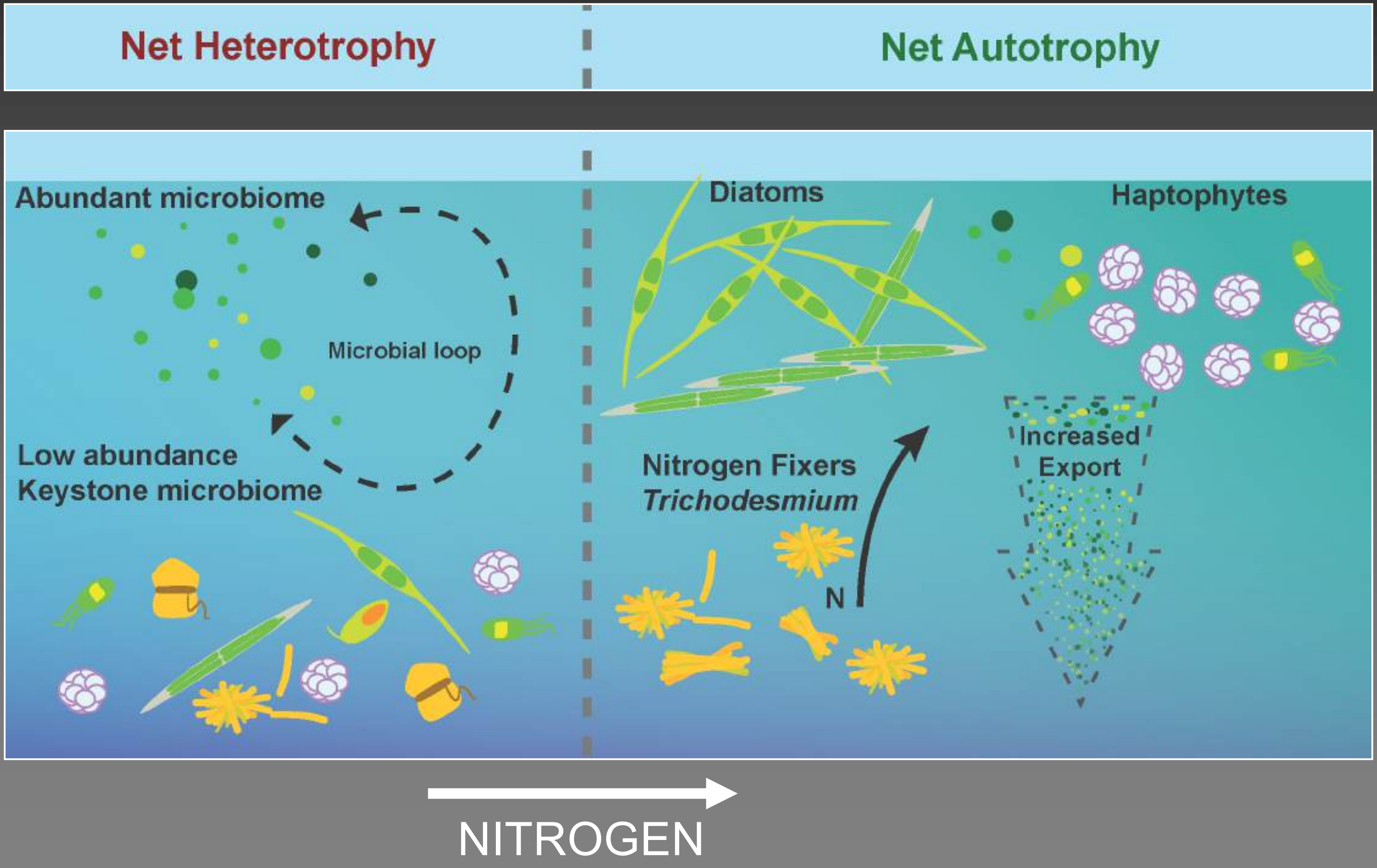




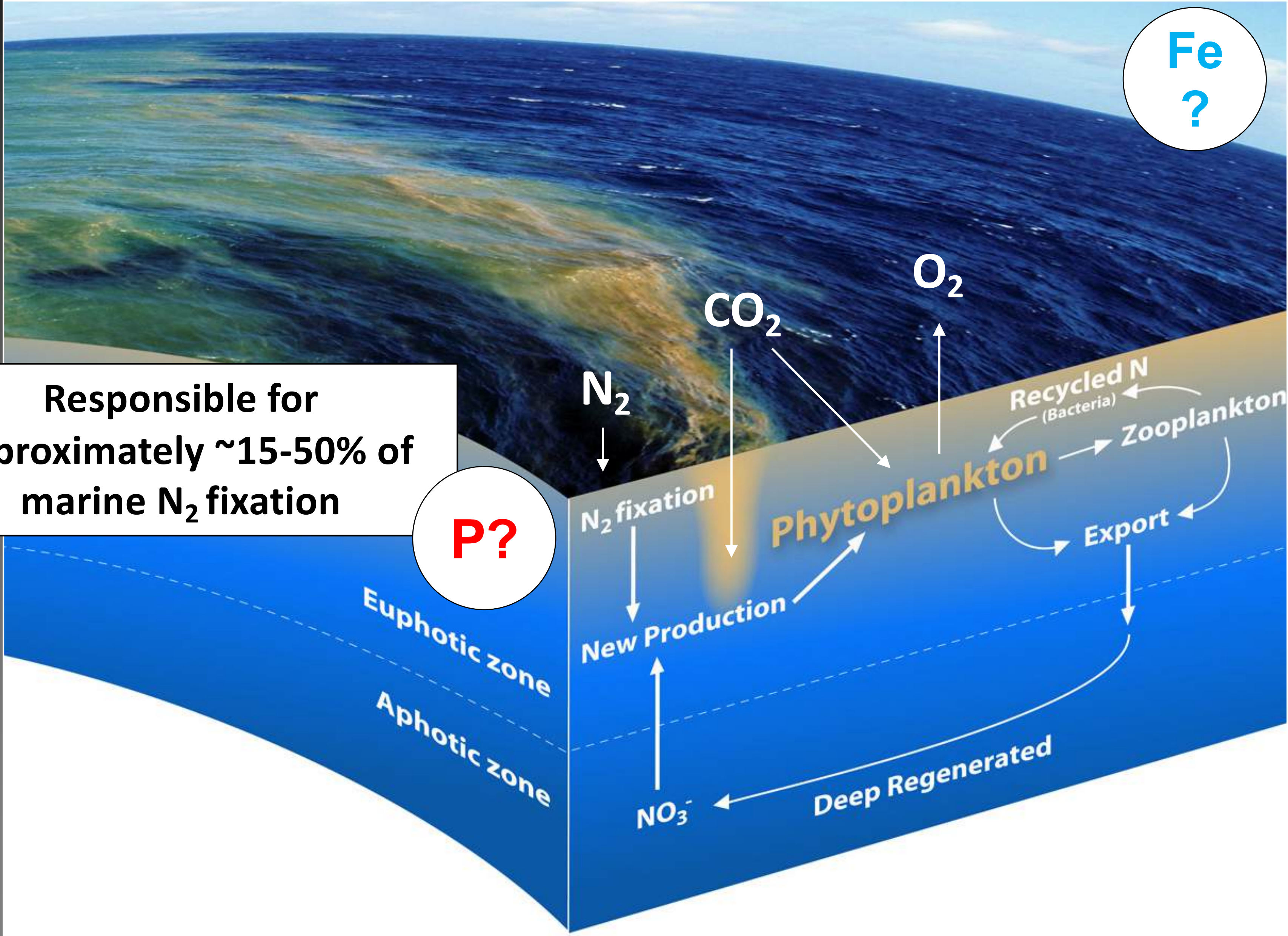
Trichodesmium: critical to ecosystem function

Photo: Chris Wade
Tricho. micrograph: WHOI

Nitrogen inputs can shift ecosystem state and carbon flow

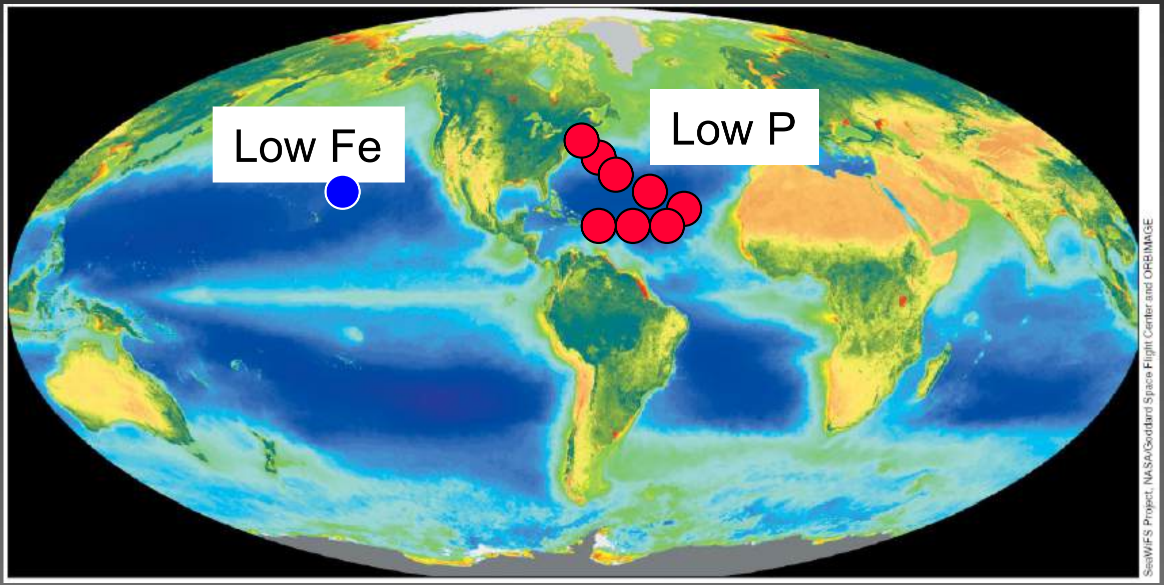


Trichodesmium: critical to ecosystem function



Responsible for approximately ~15-50% of marine N$_2$ fixation

Predicting the role of Fe and P in *Trichodesmium* distribution and N_2 fixation



SeaWiFS Project, NASA/Goddard Space Flight Center and ORBIMAGE

North Atlantic

P-limited?

North Pacific

Fe-limited?

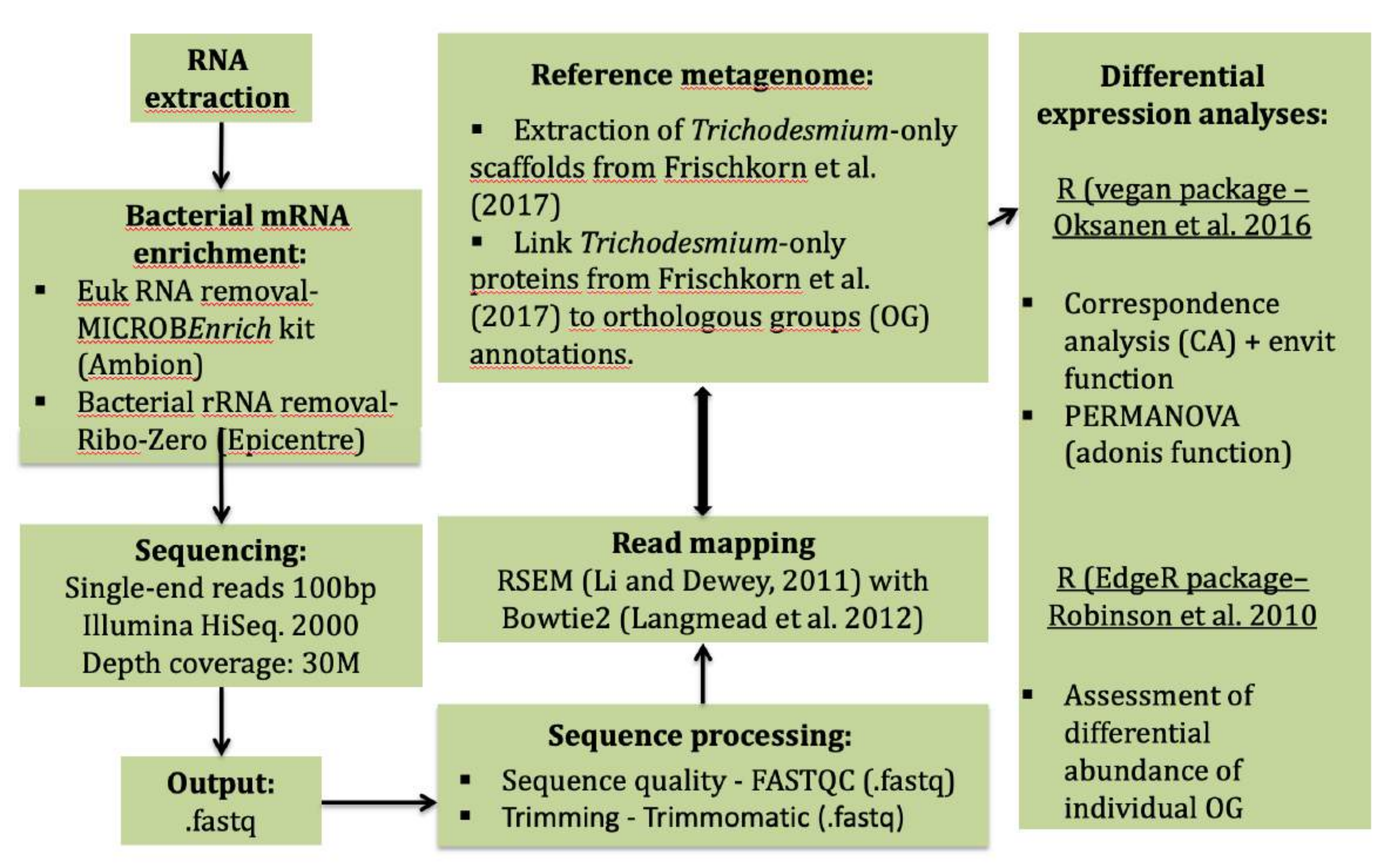
2000s

- Protein biomarkers: Dyhrman et al. (2002) *L&O*
- qRT-PCR of transcript biomarkers: Dyhrman et al. (2006) *Nature*

2010s

- Metagenomics and Metatranscriptomics: Rouco et al. (2018) *ISME J*.

Metatranscriptome workflow – reference mapping



RNA extraction

Bacterial mRNA enrichment:

- Euk RNA removal- MICROB*Enrich* kit (Ambion)
- Bacterial rRNA removal- Ribo-Zero (Epicentre)

Sequencing:
Single-end reads 100bp
Illumina HiSeq. 2000
Depth coverage: 30M

Output:
.fastq

Sequence processing:

- Sequence quality - FASTQC (.fastq)
- Trimming - Trimmomatic (.fastq)

Read mapping
RSEM (Li and Dewey, 2011) with Bowtie2 (Langmead et al. 2012)

Reference metagenome:

- Extraction of *Trichodesmium*-only scaffolds from Frischkorn et al. (2017)
- Link *Trichodesmium*-only proteins from Frischkorn et al. (2017) to orthologous groups (OG) annotations.

Differential expression analyses:

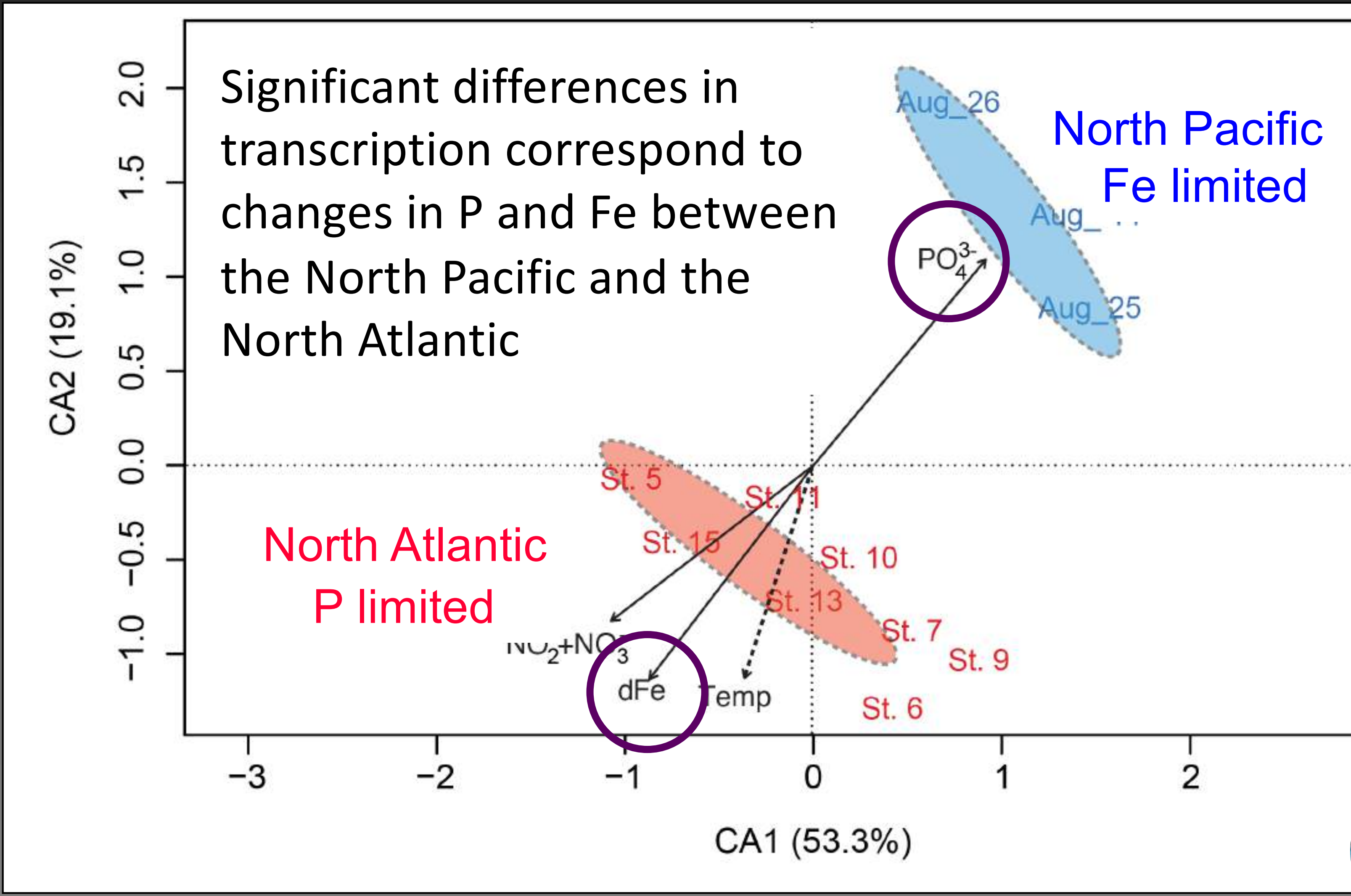
R (vegan package – Oksanen et al. 2016

- Correspondence analysis (CA) + envit function
- PERMANOVA (adonis function)

R (EdgeR package– Robinson et al. 2010

- Assessment of differential abundance of individual OG

Trichodesmium transcriptional patterns between the North Pacific and the North Atlantic



Significant differences in transcription correspond to changes in P and Fe between the North Pacific and the North Atlantic

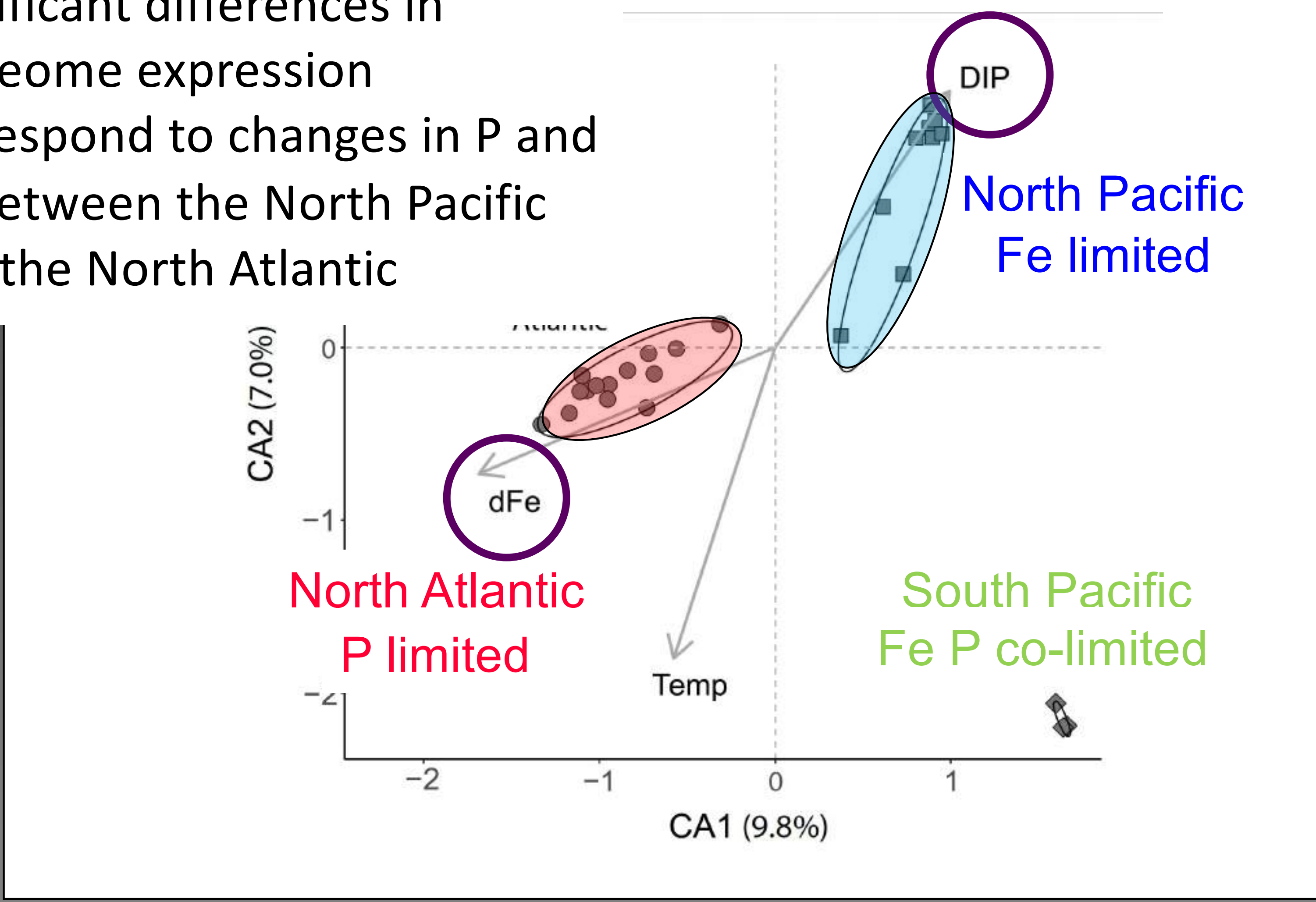
North Pacific
Fe limited

North Atlantic
P limited

Ruoco et al. (2018) *ISME J*

Trichodesmium metaproteome patterns between regions

Significant differences in proteome expression correspond to changes in P and Fe between the North Pacific and the North Atlantic

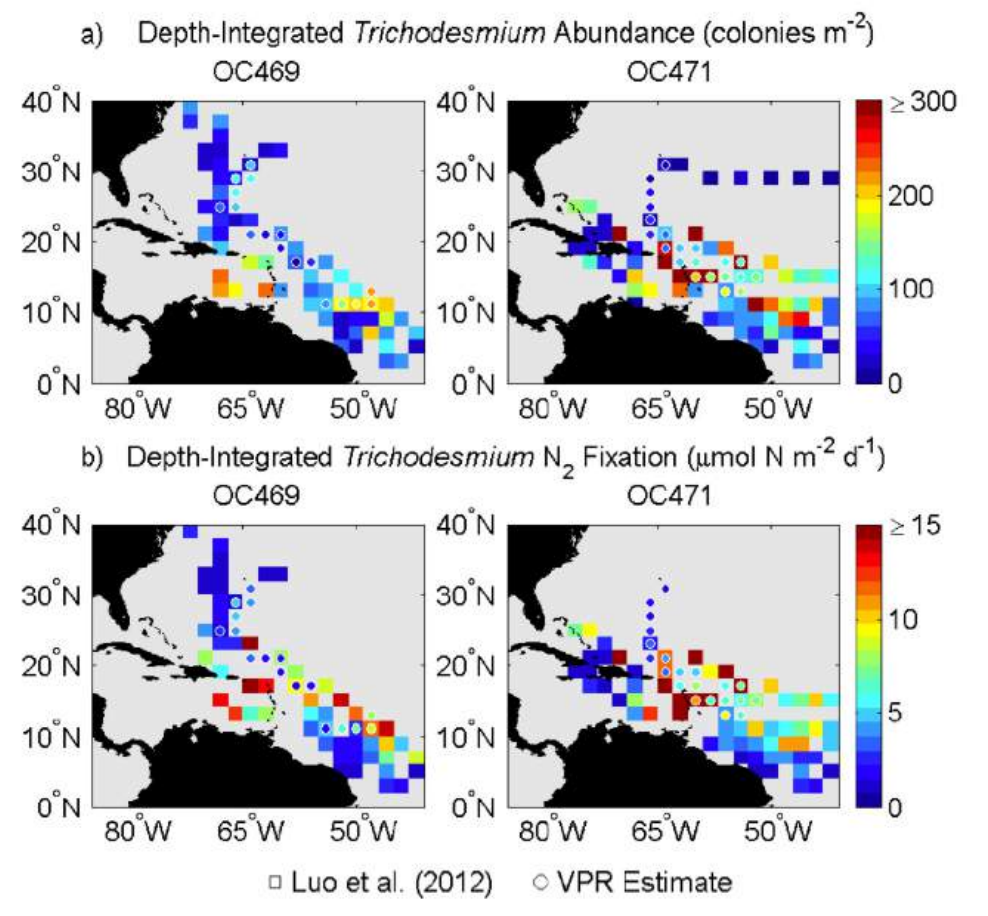


Anderson et al. (in review) *ISME J*

Modeling N_2 fixation is still a challenge

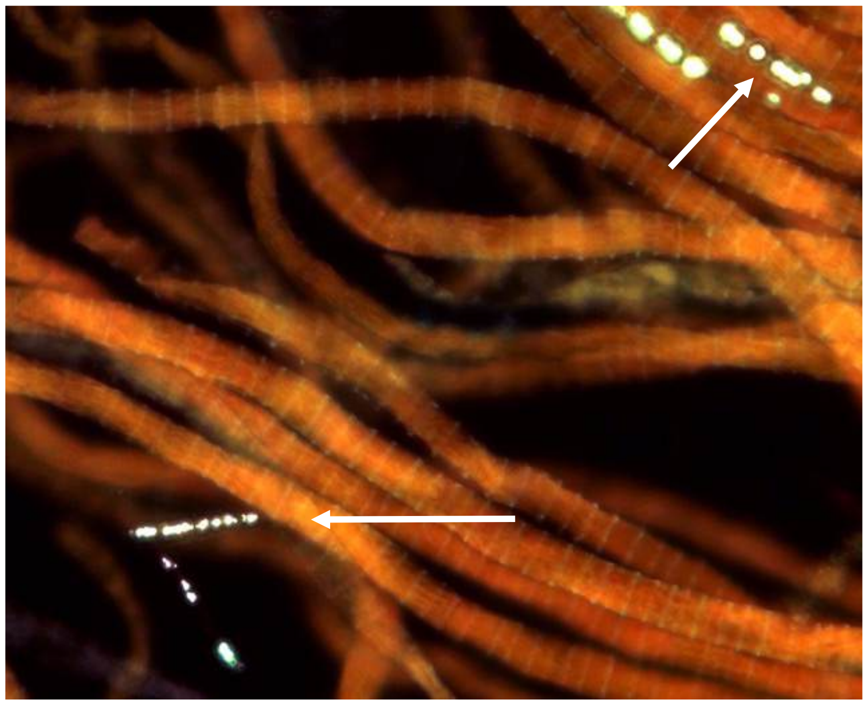
- Models do not balance the N cycle in the ocean or recapitulate patterns well

- **Geochemistry alone is not a good predictor of distribution or N_2 fixation**



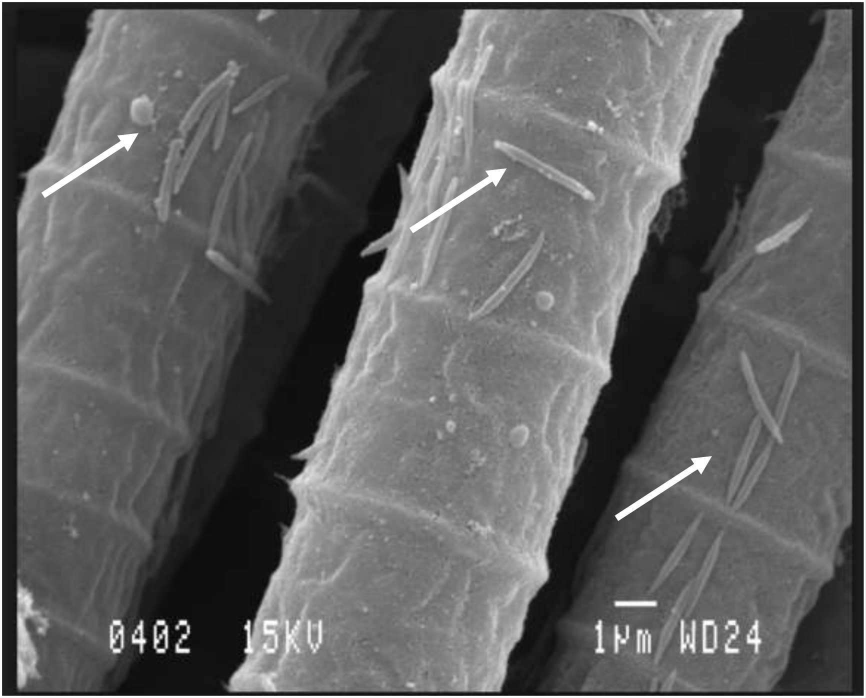
Olson et al. 2015 *DSR* II

Trichodesmium hosts a community of epibiotic microbes



Trichodesmium* and *Plectonema

Dyhrman et. al. (2002) *L&O*



***Trichodesmium* and Epibionts**

Image courtesy Tracy Mincer

Trichodesmium colonies harbor other microbes with distinct physiology.

Core questions about the *Trichodesmium* microbiome

Who is there?

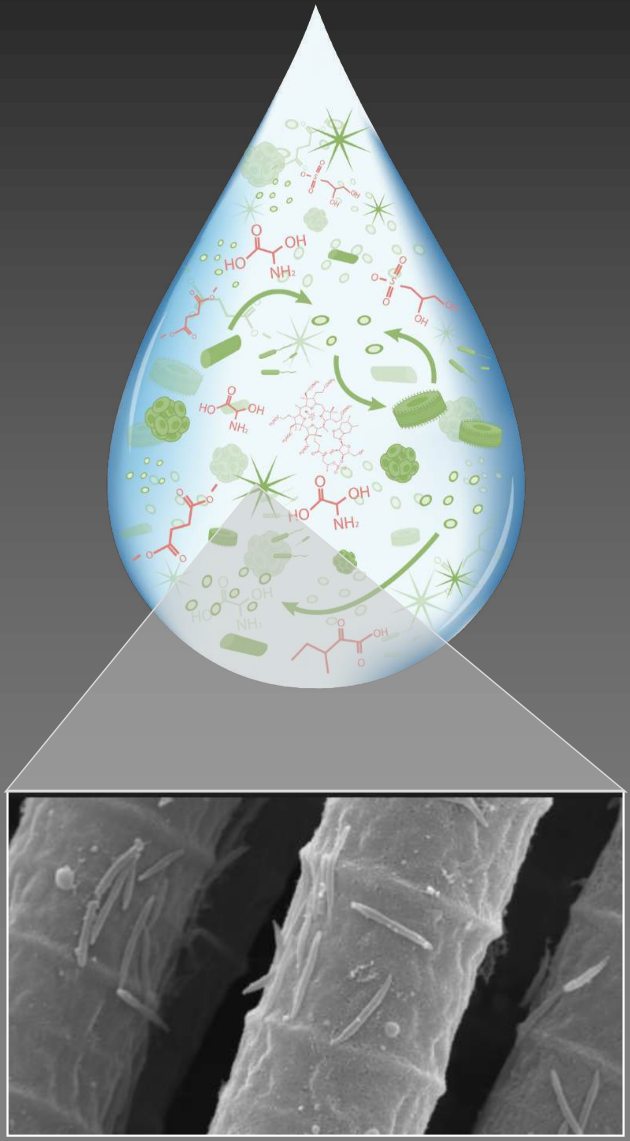
- Microbiome diversity
- *16S amplicon sequencing*

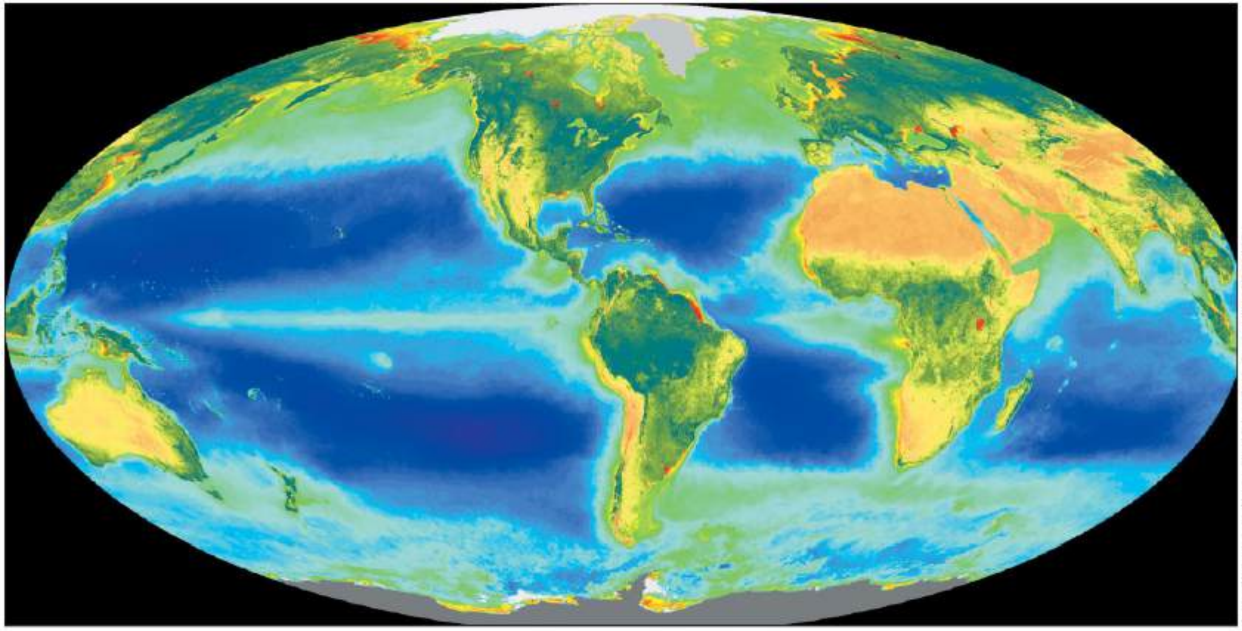
What are they doing?

- Microbiome functional diversity
- *Metagenome assembled genomes*

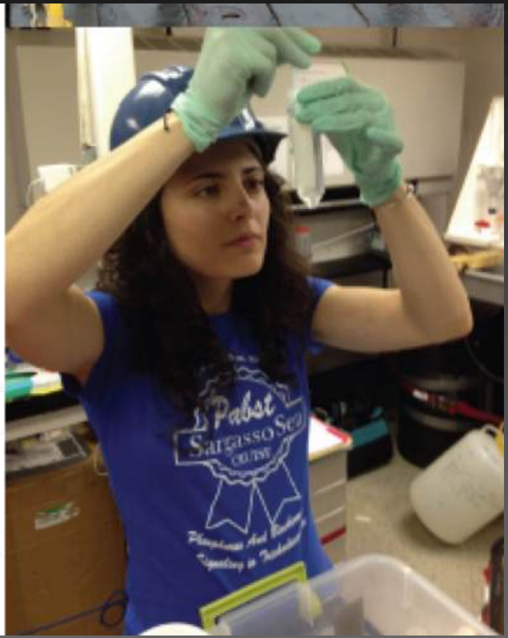
Are they interacting?

- Coordinated diel signaling
- *Metatranscriptomics*

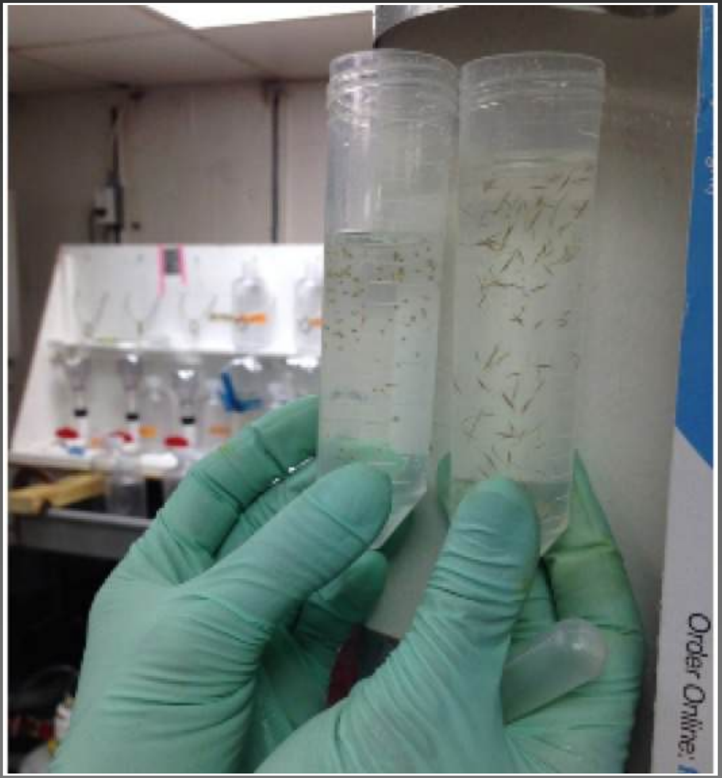




SeaWiFS Project, NASA/Goddard Space Flight Center and ORBIMAGE



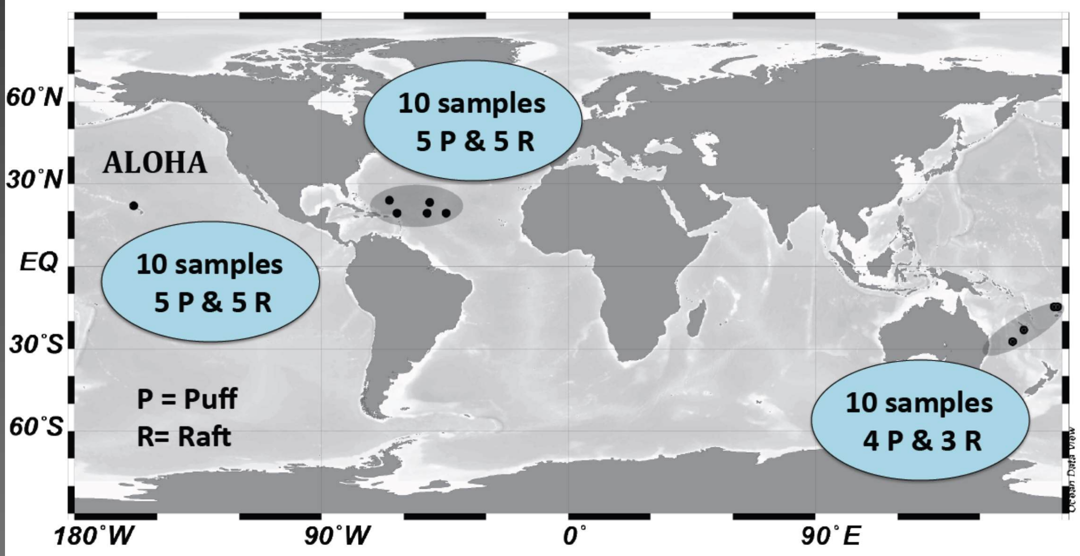
Trichodesmium microbiome diversity



Trichodesmium colonies are picked and rinsed from net tows for 16S sequencing to remove all but the tightly associated epibionts.

Epibiont diversity

EXPERIMENTAL APPROACH



Collection of colonies from 3 ocean basins (~top 25m)

↓

DNA extraction

↓

Sequencing: V4 region of 16S rRNA gene **Miseq (2x150 bp)** →

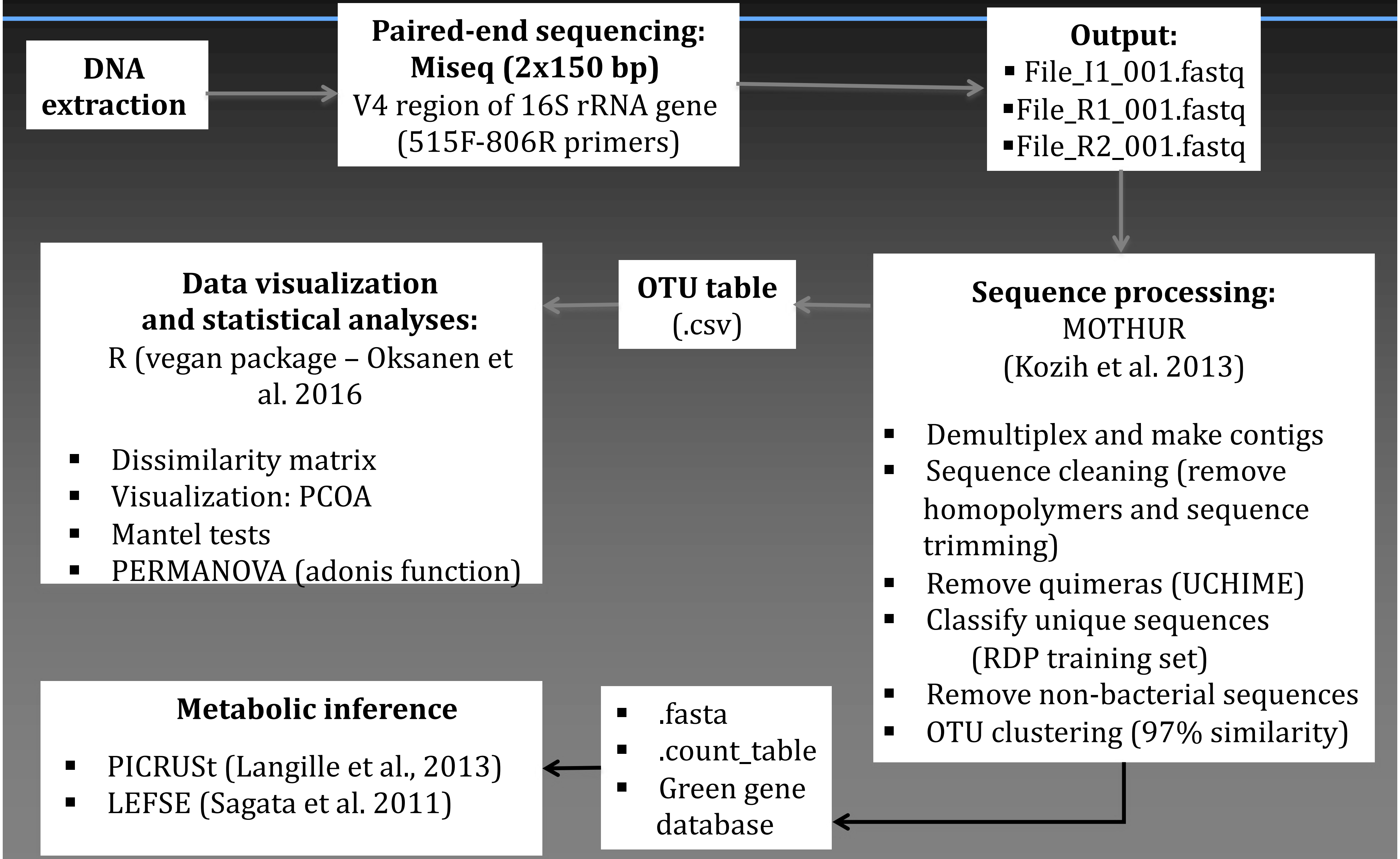
Mothur v. 1. 34. 0 (936,749 reads)

- ***Trichodesmium***: ~67% of reads
- **Epibionts**: ~32% of reads

What groups co-occur with *Trichodesmium*?

Are epibiont communities distinct as a function of colony morphology or environment?

16S rDNA analyses



DNA extraction

Paired-end sequencing:
Miseq (2x150 bp)
V4 region of 16S rRNA gene
(515F-806R primers)

Output:
- File_I1_001.fastq
- File_R1_001.fastq
- File_R2_001.fastq

Sequence processing:
MOTHUR
(Kozih et al. 2013)

- Demultiplex and make contigs
- Sequence cleaning (remove homopolymers and sequence trimming)
- Remove quimeras (UCHIME)
- Classify unique sequences (RDP training set)
- Remove non-bacterial sequences
- OTU clustering (97% similarity)

OTU table
(.csv)

Data visualization and statistical analyses:
R (vegan package – Oksanen et al. 2016

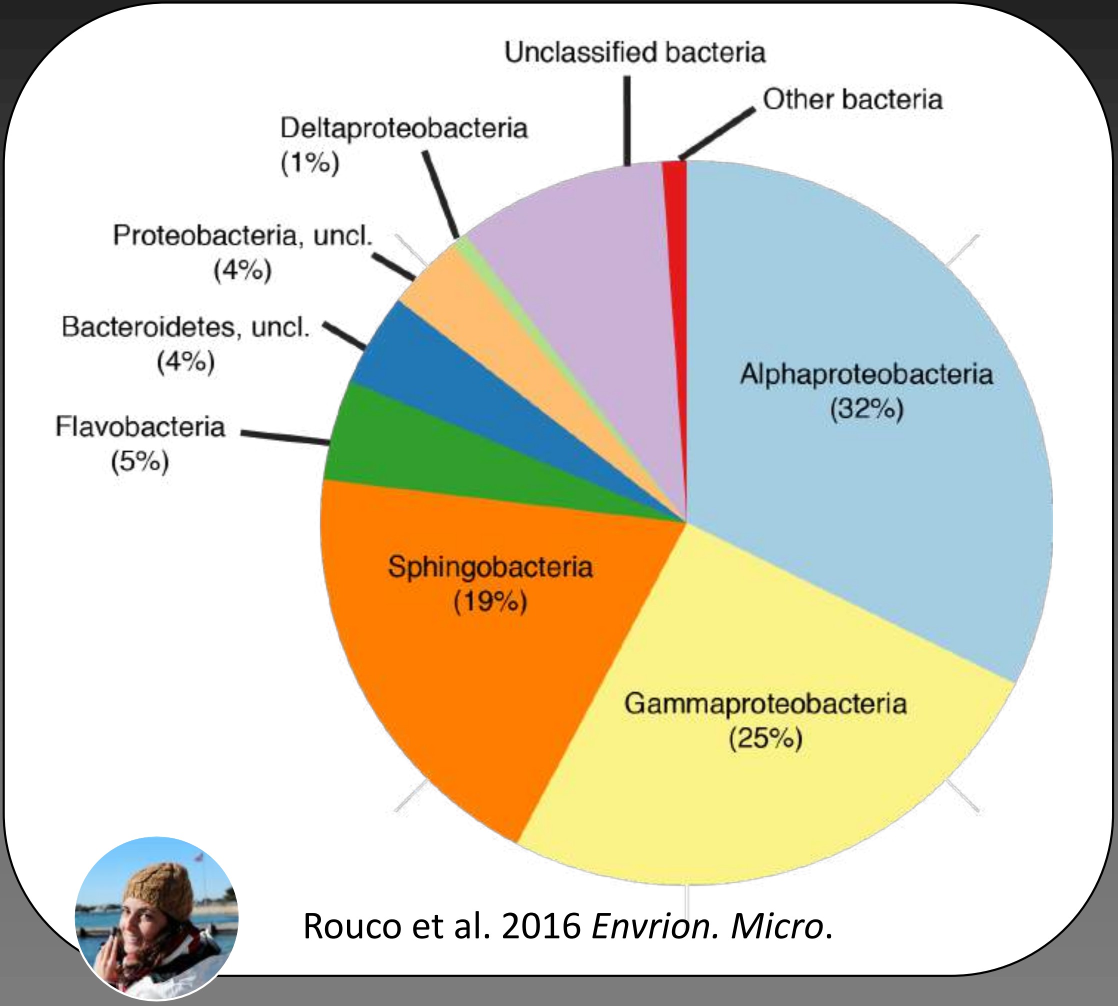
- Dissimilarity matrix
- Visualization: PCOA
- Mantel tests
- PERMANOVA (adonis function)

- .fasta
- .count_table
- Green gene database

Metabolic inference

- PICRUSt (Langille et al., 2013)
- LEFSE (Sagata et al. 2011)

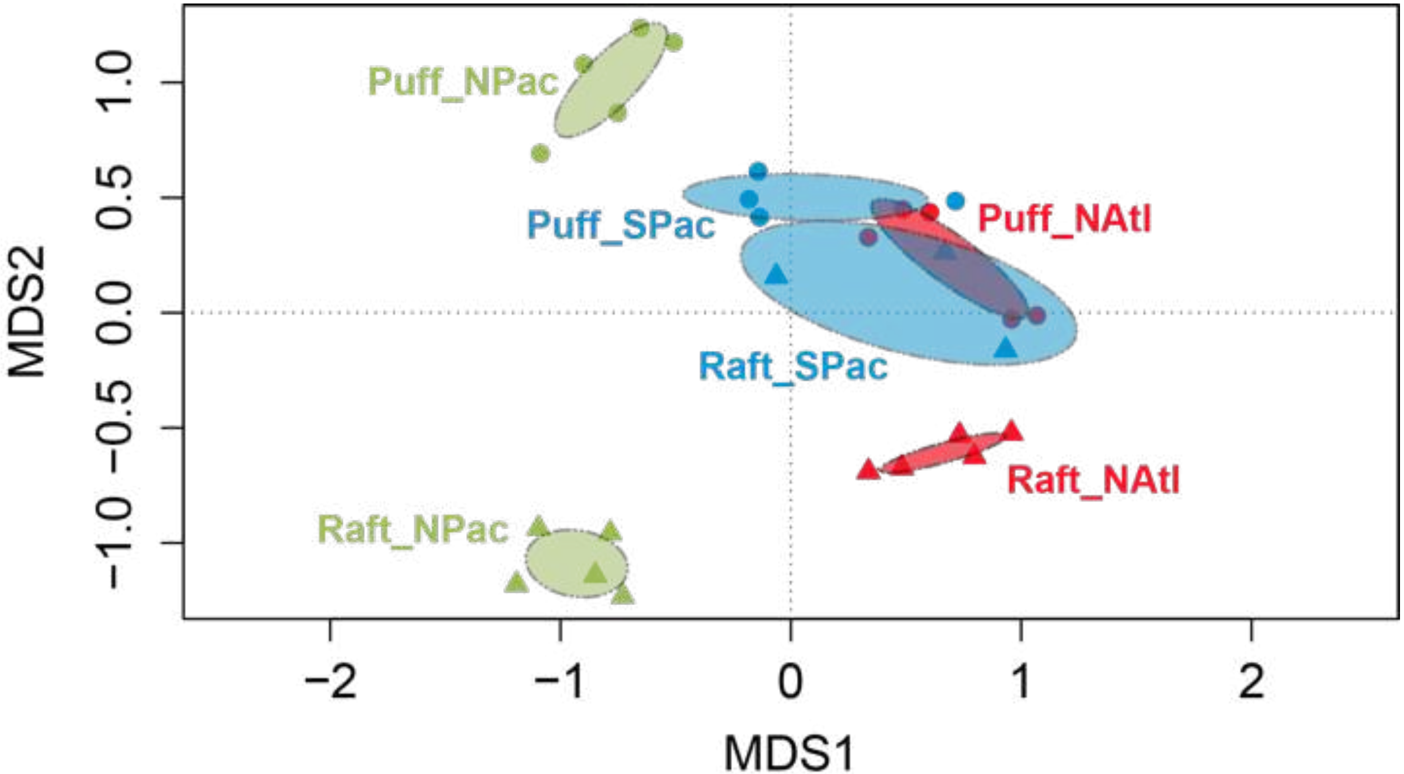
Average epibiont community



Rouco et al. 2016 *Envrion. Micro*.

16S amplicon sequencing indicates that *Trichodesmium* colonies harbor diverse epibionts distinct from common water column bacteria, and those found on sinking particles.

Microbiome community diversity (16S)



Rouco et al. (2016) *Envrion. Micro*.

Microbiome communities significantly differ by region and colony morphology, except for the S. Pacific where the *Trichodesmium* composition of rafts and puffs were not significantly different

Summary

- **Need to think about *Trichodesmium* as a holobiont**

- *What drives community assembly?*
  - *Unique niche?*
    - Type of *Trichodesmium*
    - Physiological of the colony
    - Environment
  - *Lottery?*
    - Random selection of potential copiotrophs with certain gene functions?
- *Stability over time?*

Core questions about the *Trichodesmium* microbiome

Who is there?

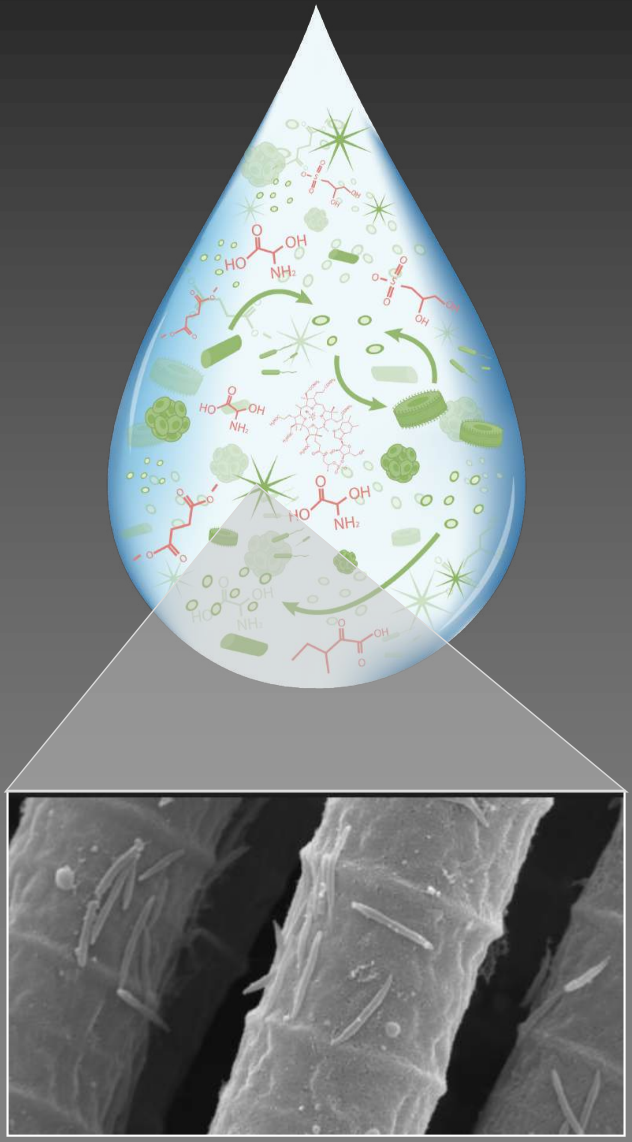
- **Diverse community, distinct from water column**

What are they doing?

- Microbiome functional diversity

Are they interacting?

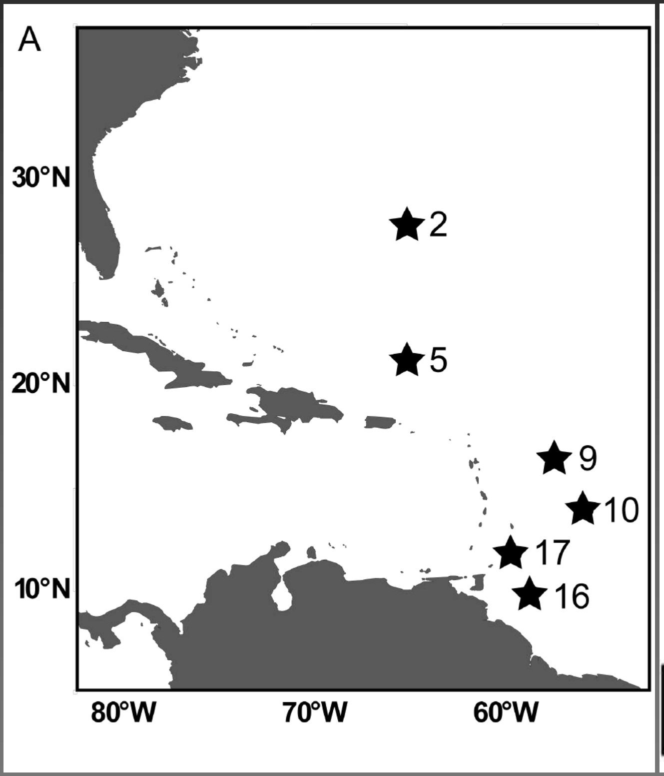
- Diel patterns in the consortia

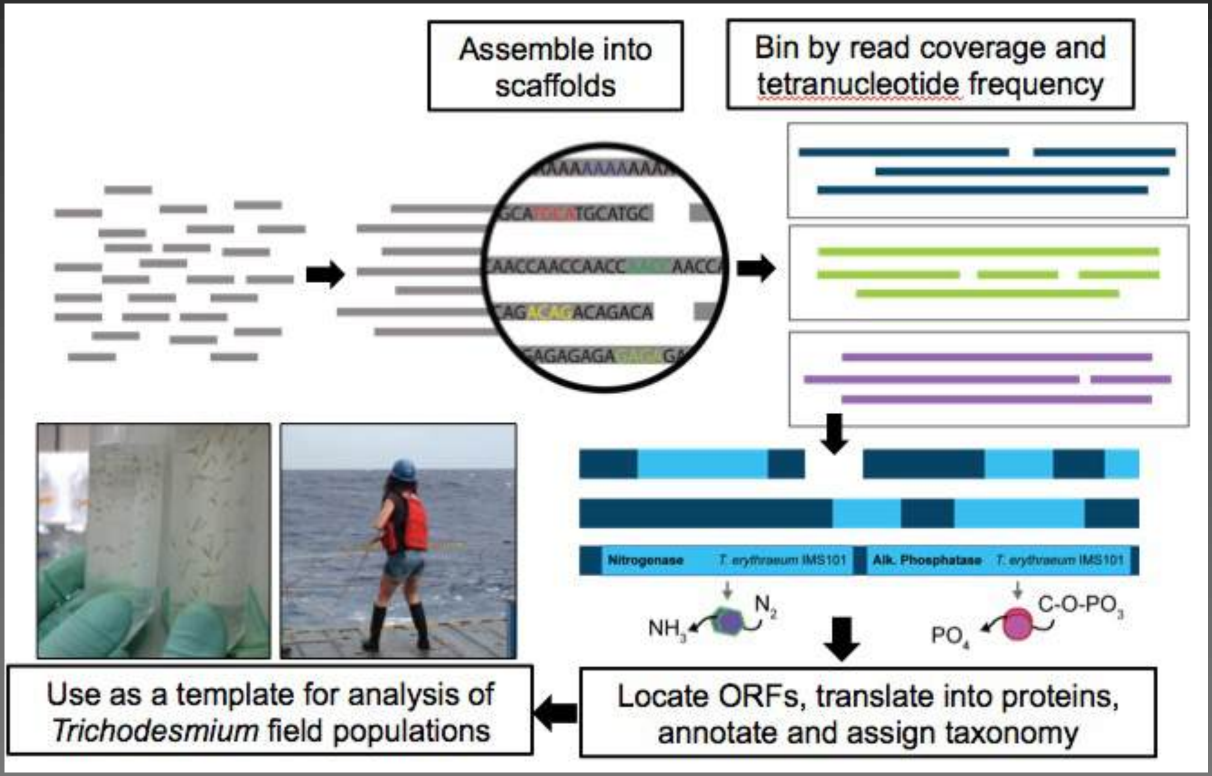


Metagenome derived metabolic potential in the *Trichodesmium* holobiont

IDBA-UD

MaxBin2



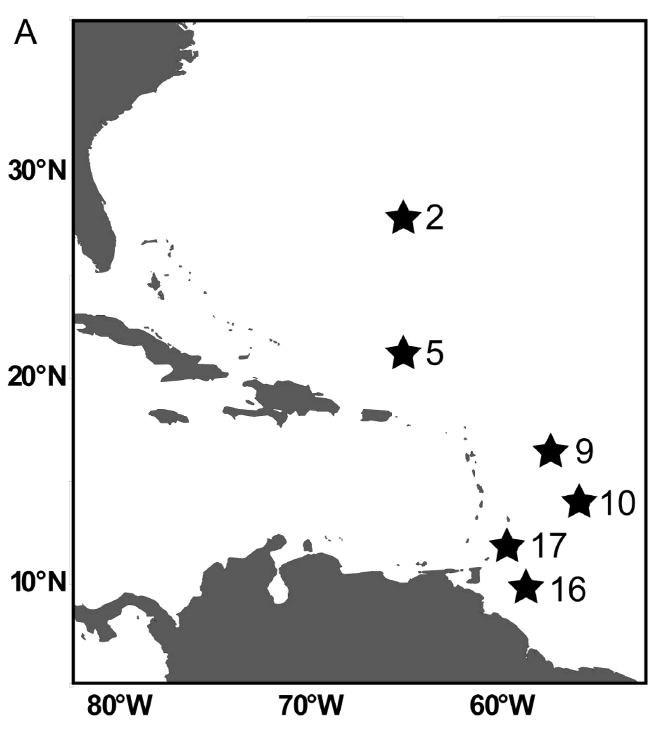


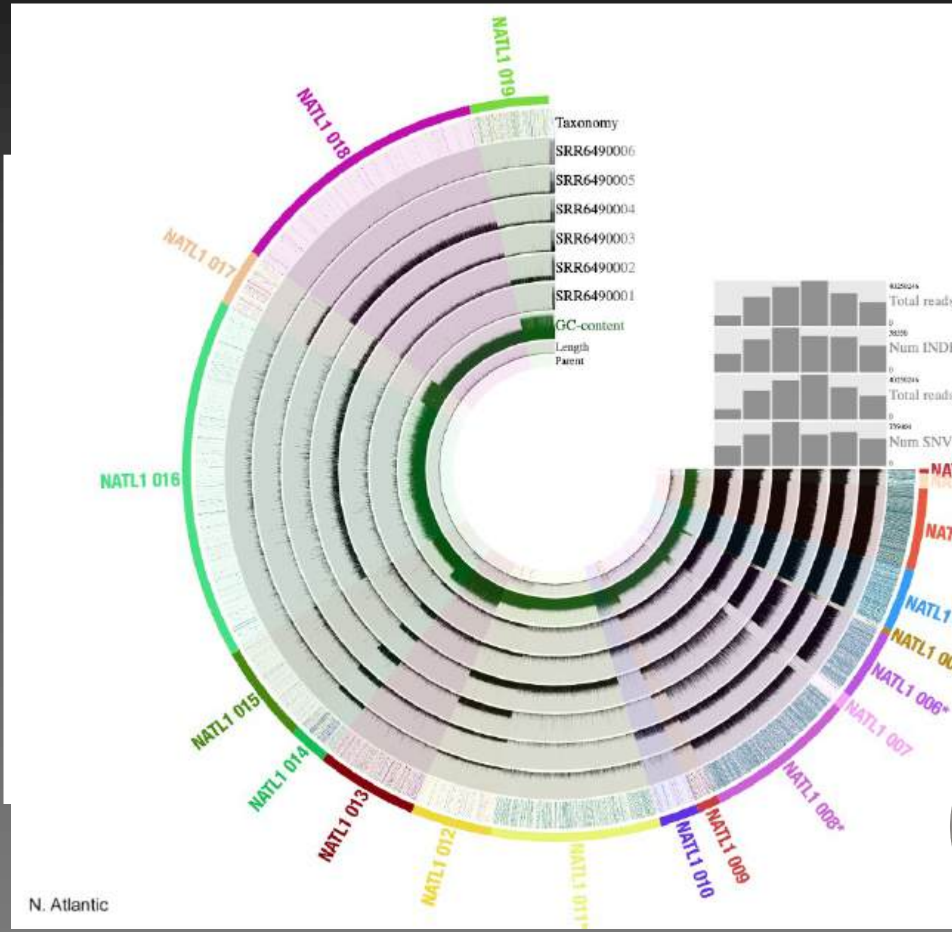
Prodigal

DIAMOND-BLAST
MEGAN KEGG

Trichodesmium colonies were isolated for metagenome sequencing along a phosphorus gradient in the western north Atlantic.

Composition of the holobiont



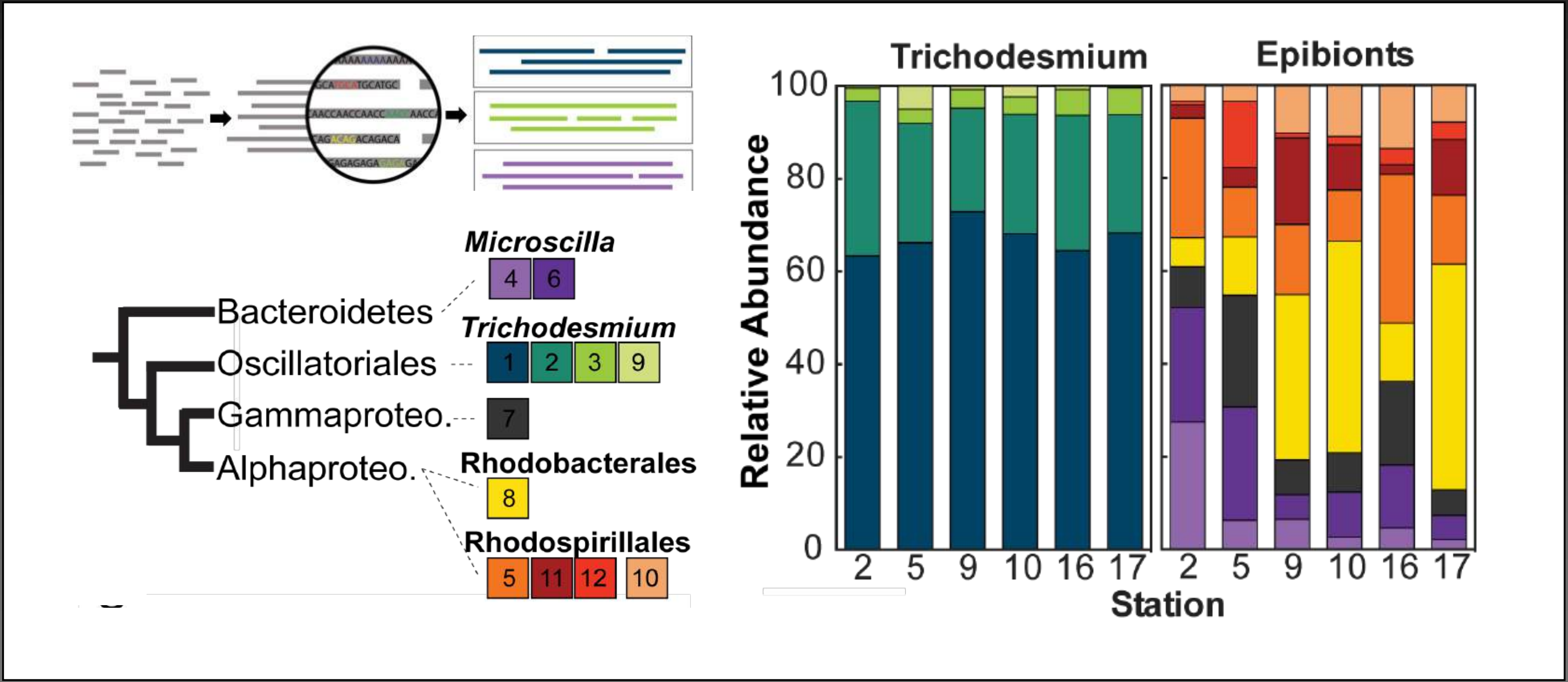


Nearly complete (~90%) genome bins were assembled and results are consistent with 16S data
= Metagenome assembled genome or MAG

Frischkorn et al. (2017) *ISMEJ*

Assembly graphic – anvio.org

Relative distribution of MAGs with read mapping

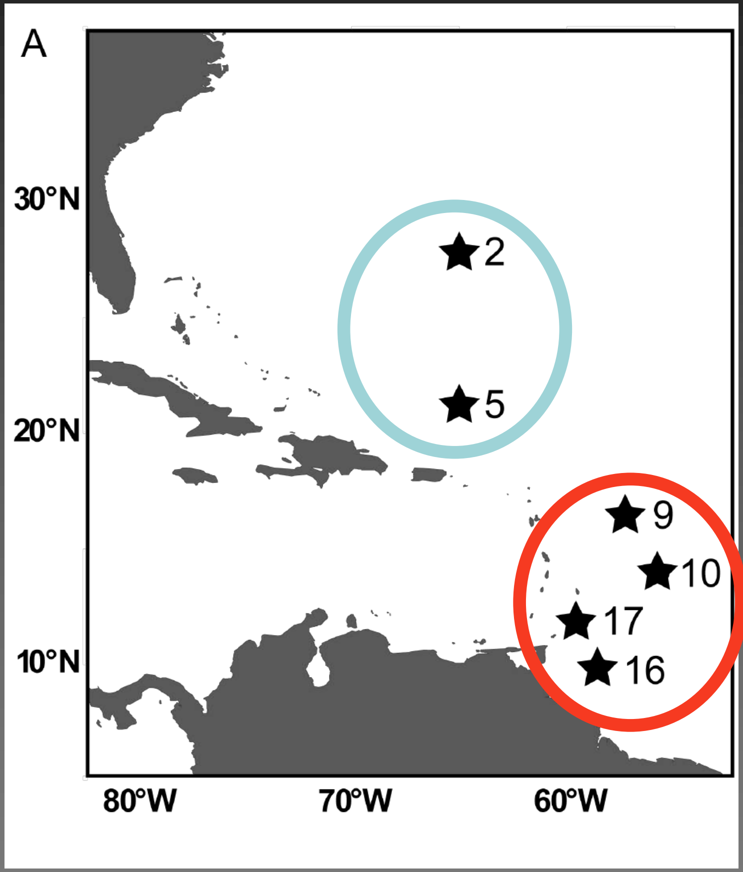


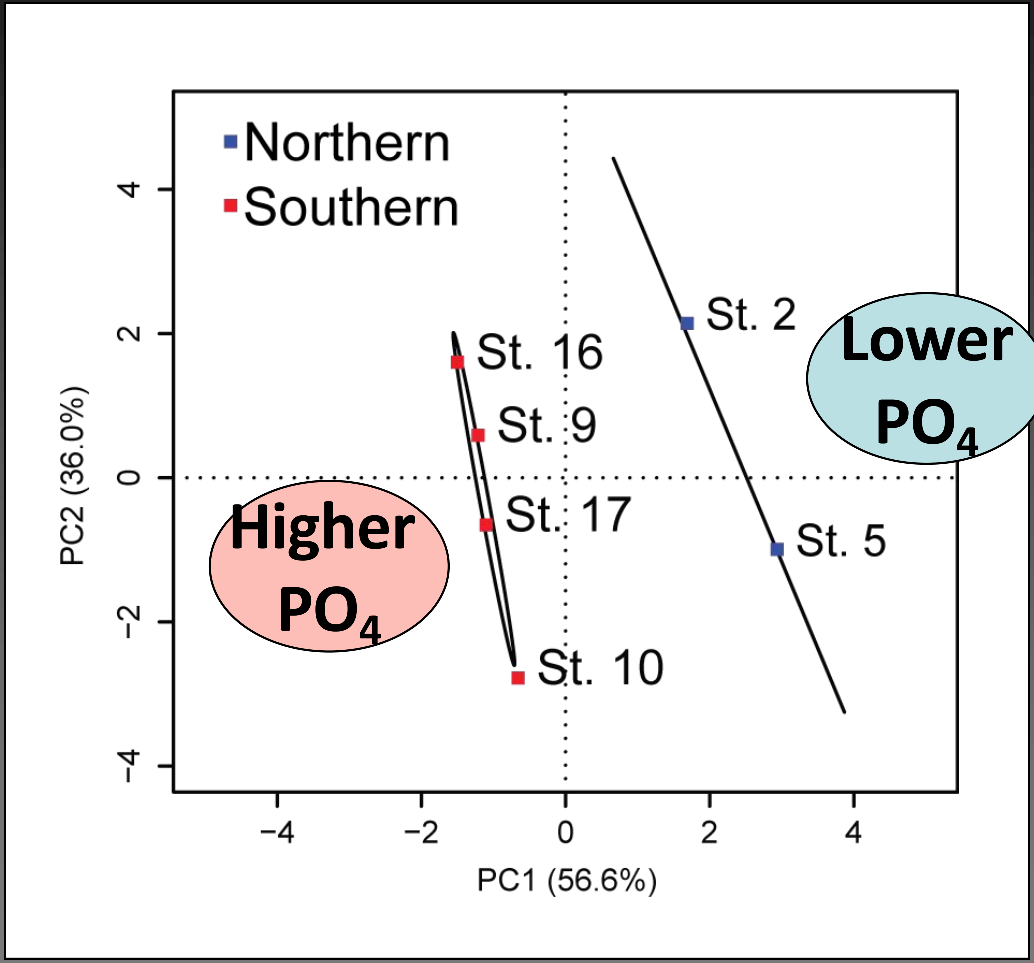
Frischkorn et al. (2017) *ISMEJ*

Relative proportion of *Trichodesmium* bins consistent across stations

Epibiont genome bins are detected at all stations, but the relative abundance varies

Significant differences across transect



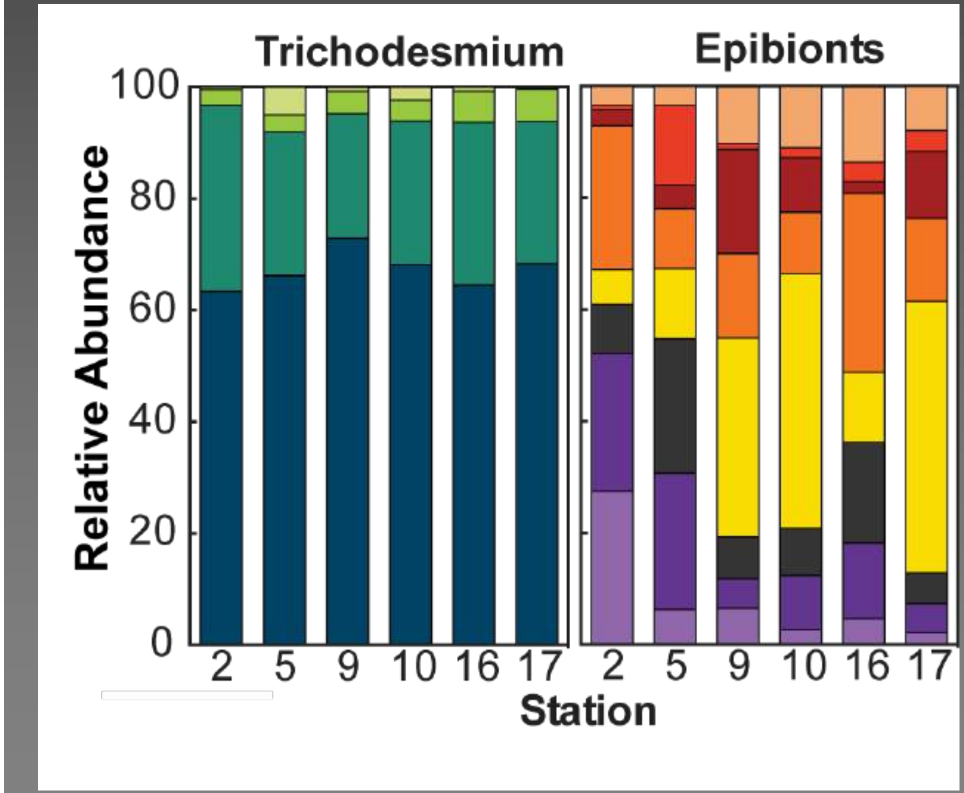


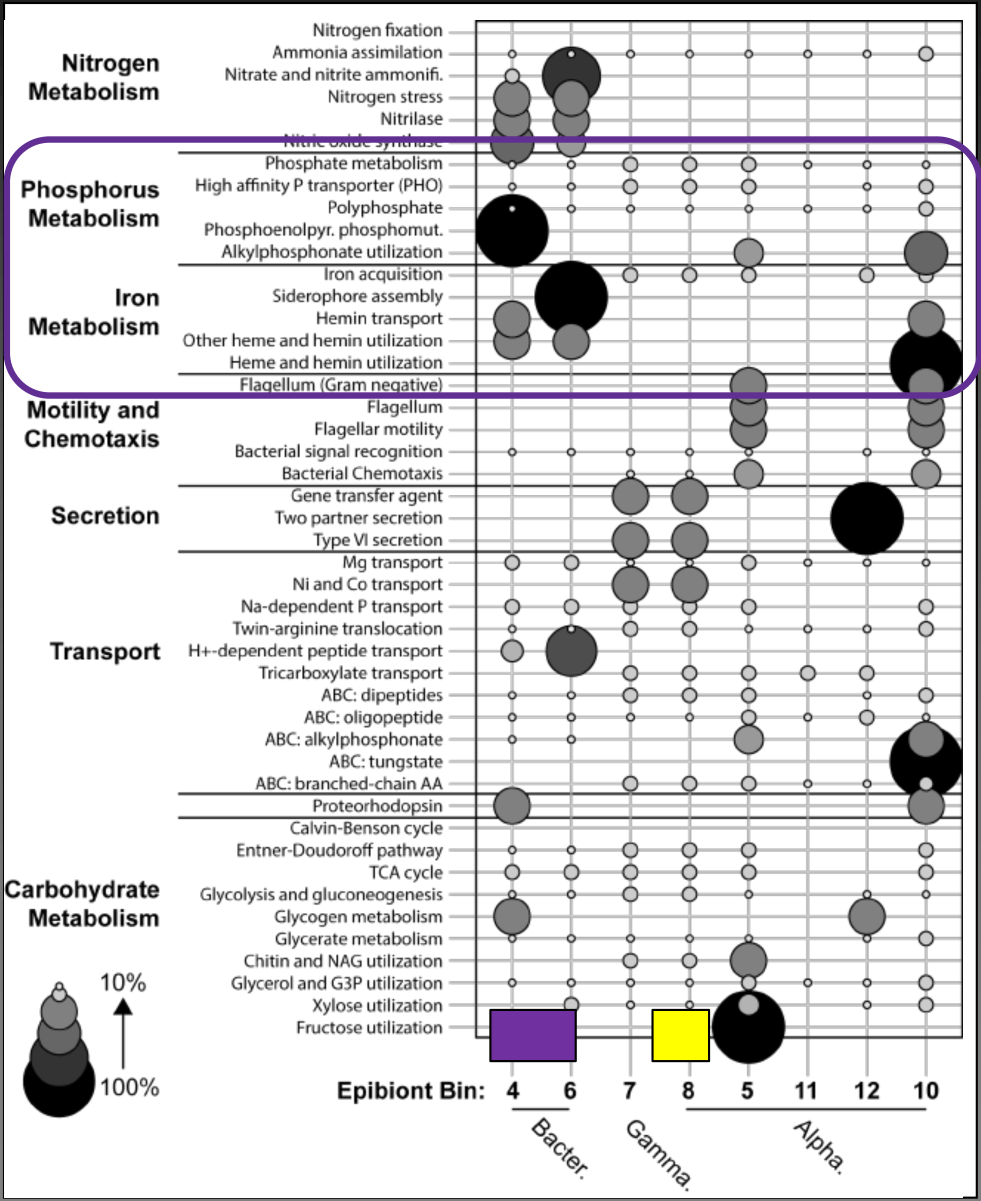
Frischkorn et al. (2017) *ISME J*

Significant difference in epibiont relative abundance with total dissolved phosphorus

Variable distribution of functional pathways among epibionts

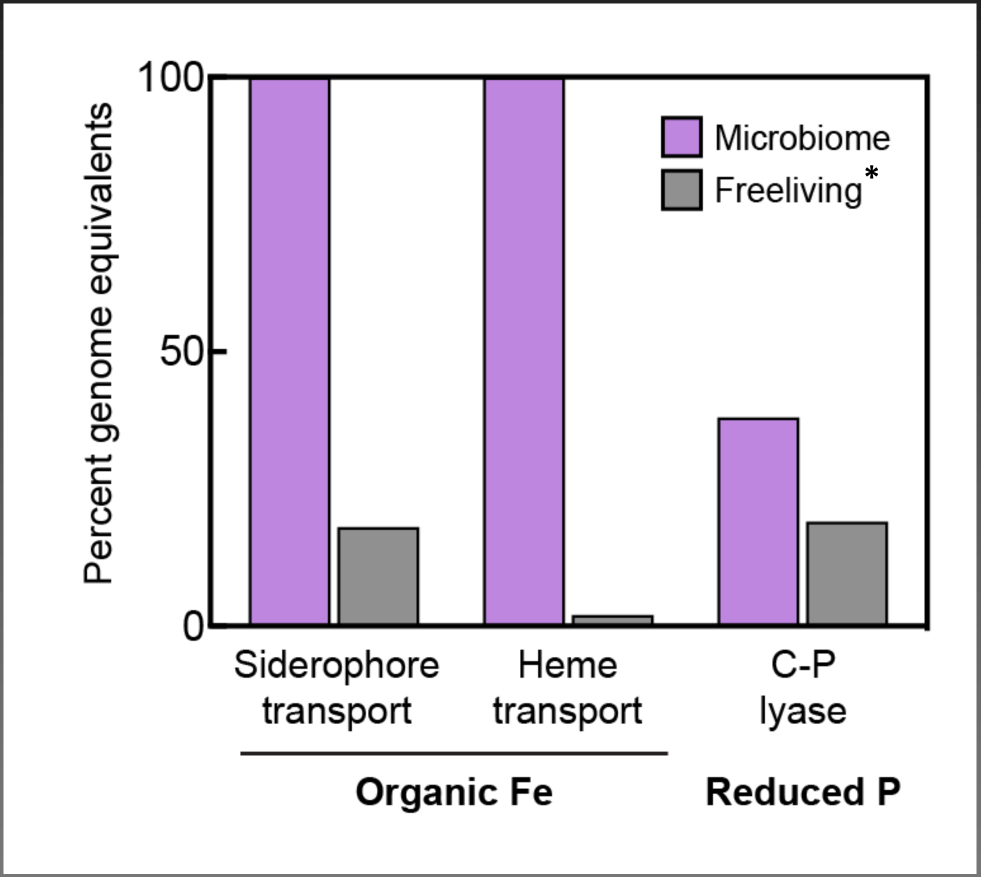
- Differential pathway enrichment consistent with a microbiome that is modulated as a function of environment





Frischkorn et al. (2017) *ISMEJ*

Epibionts enrich holobiont for key functions



Frischkorn et al. (2017) *ISME J*

Phosphonate, heme and siderophore functions are enriched in the epibionts relative to water column microbes

Comparing metabolic potential in the holobiont

Metagenomes

↓

Orthologous group analysis

↙ ↘

Epibionts v. *Trichodesmium*

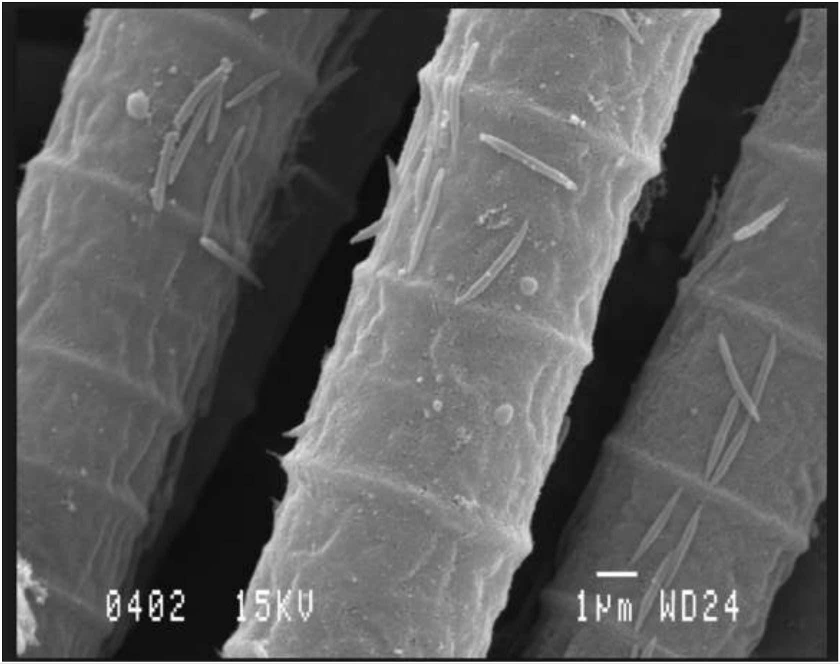
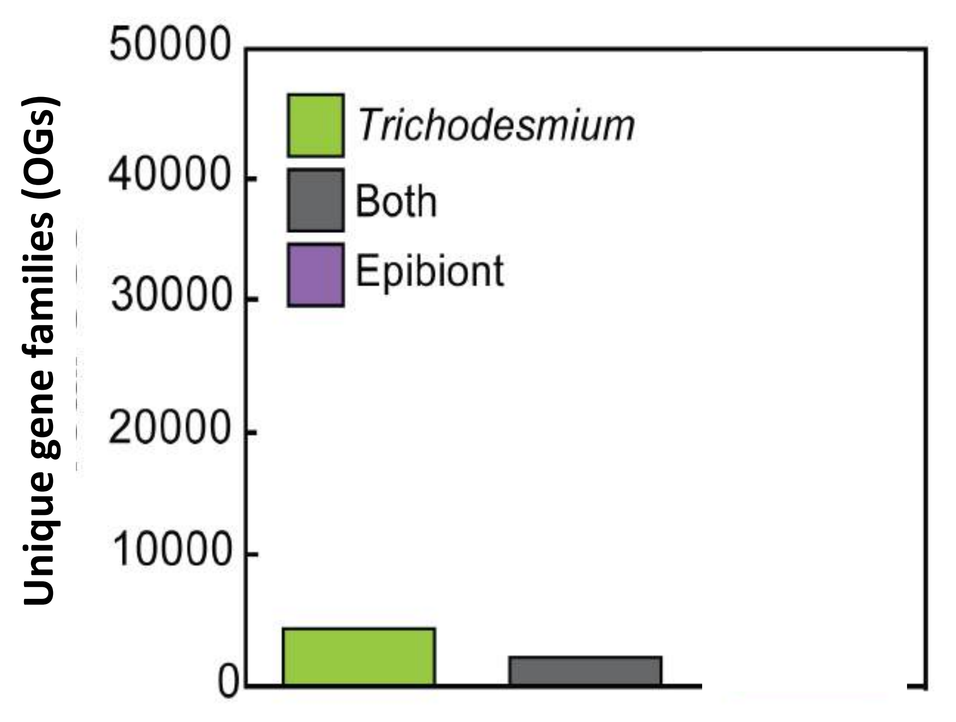


Image courtesy Tracy Mincer

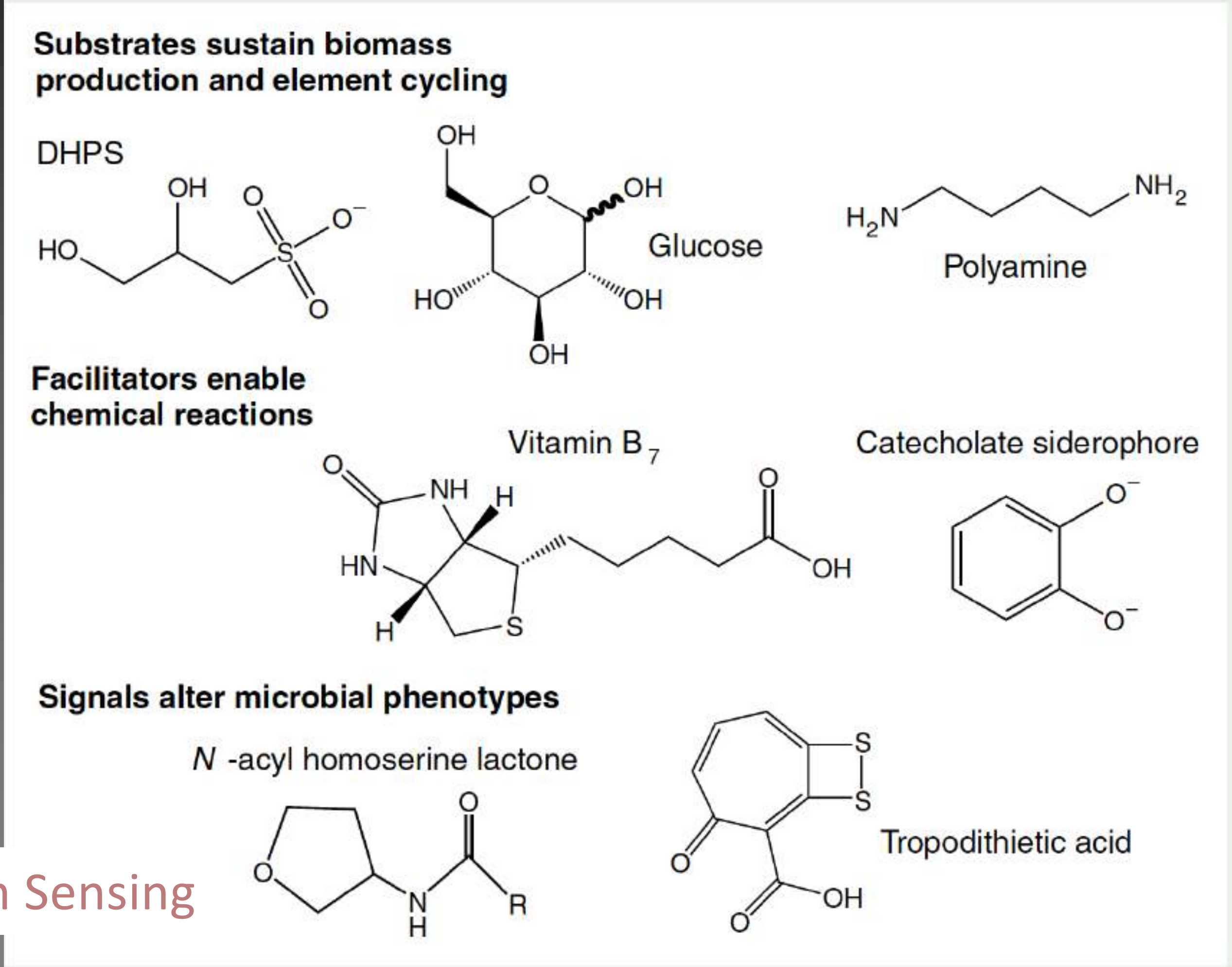
Epibionts significant metabolic potential



Frischkorn et al. (2017) *ISME J*

Orthologous (OG) group analysis suggests that epibionts confer the vast majority (>90% of OGs) of metabolic *functions* to the holobiont.

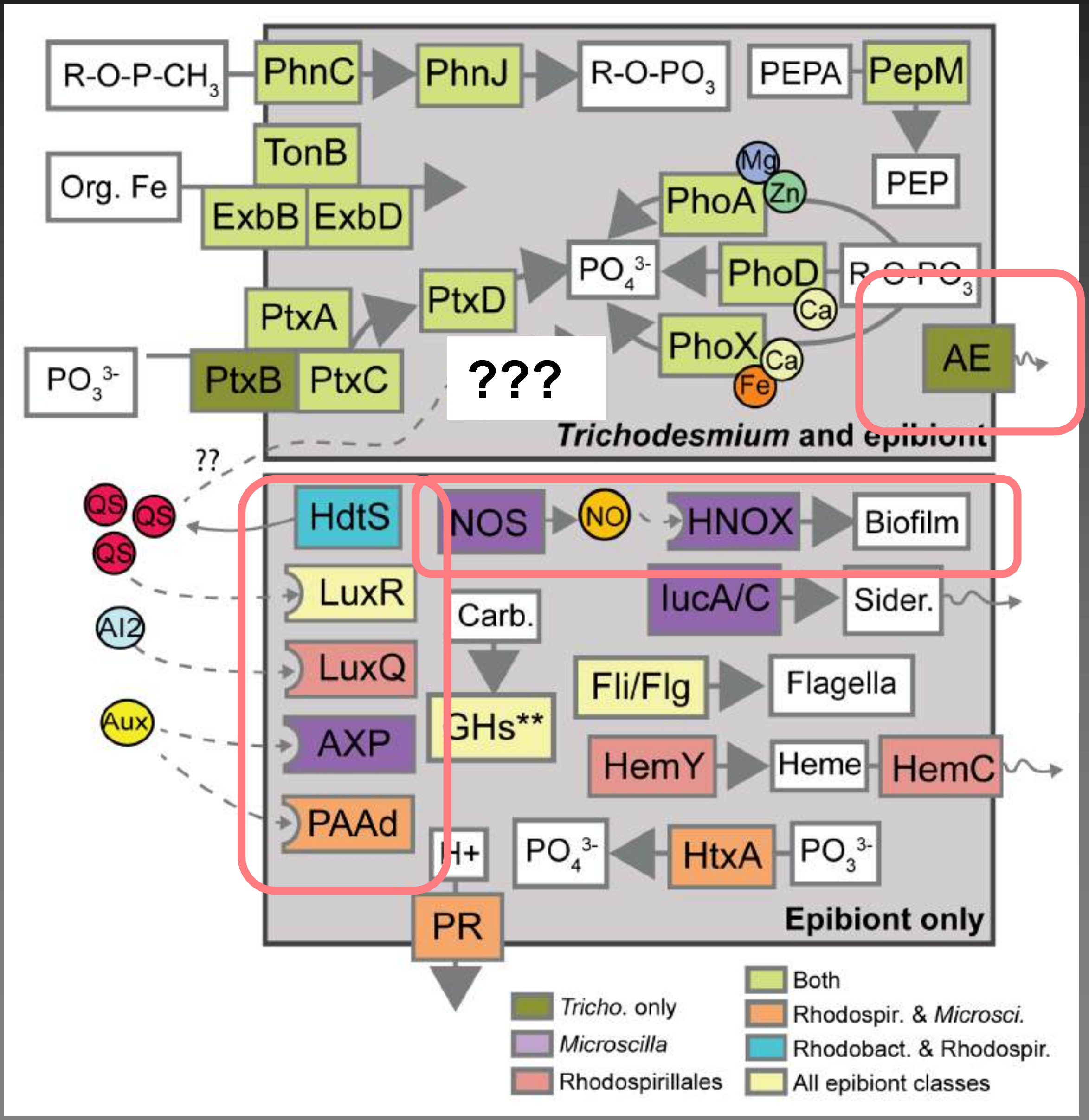
Chemical-microbe network within the *Trichodesmium* holobiont



Quorum Sensing

Moran et al. (2022) *Nature Micro.*

Microbial cross talk within the *Trichodesmium* holobiont



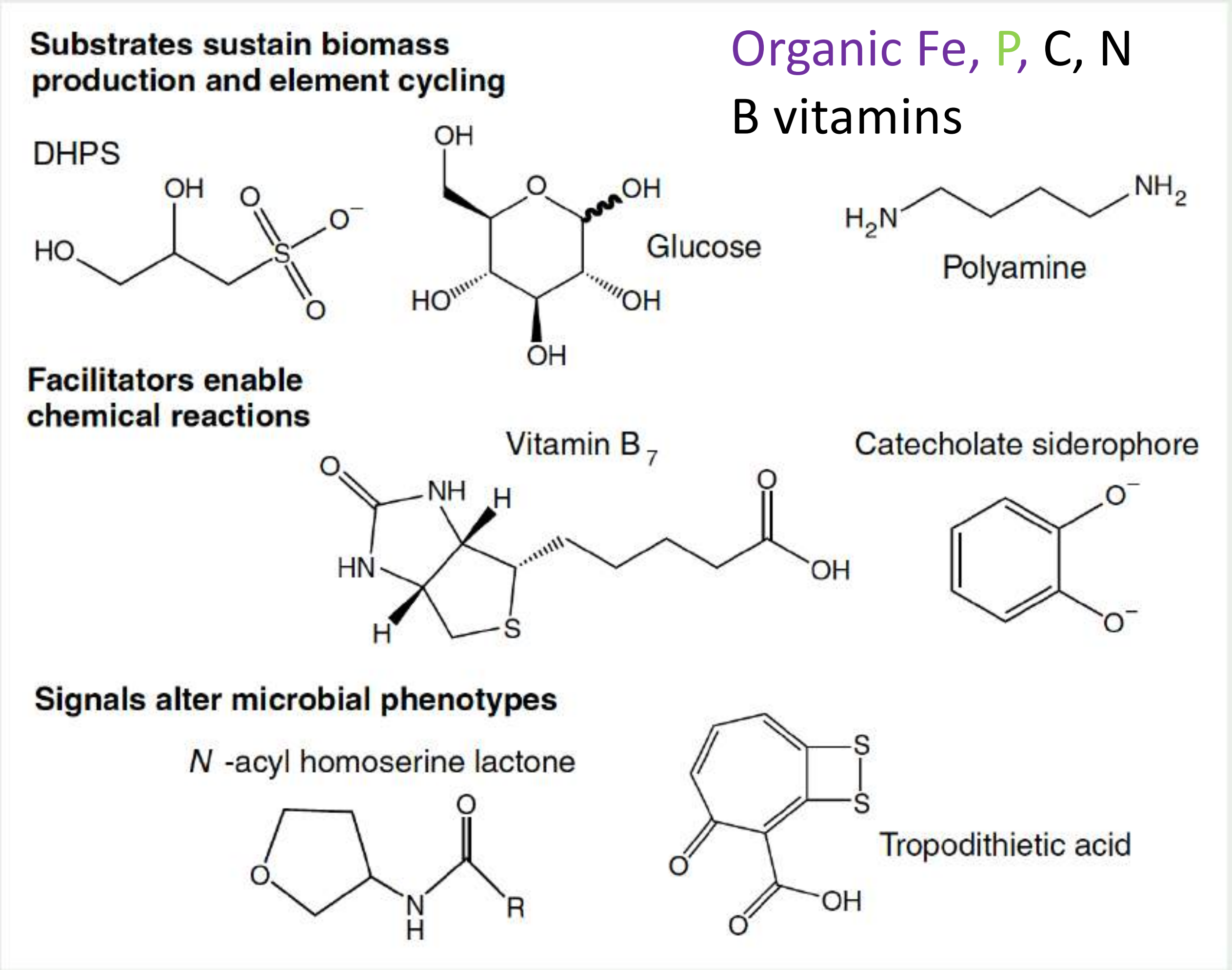
Auxin sensing and response

NO production

AHL QS Pathways are isolated to microbiome

QS and cell signaling

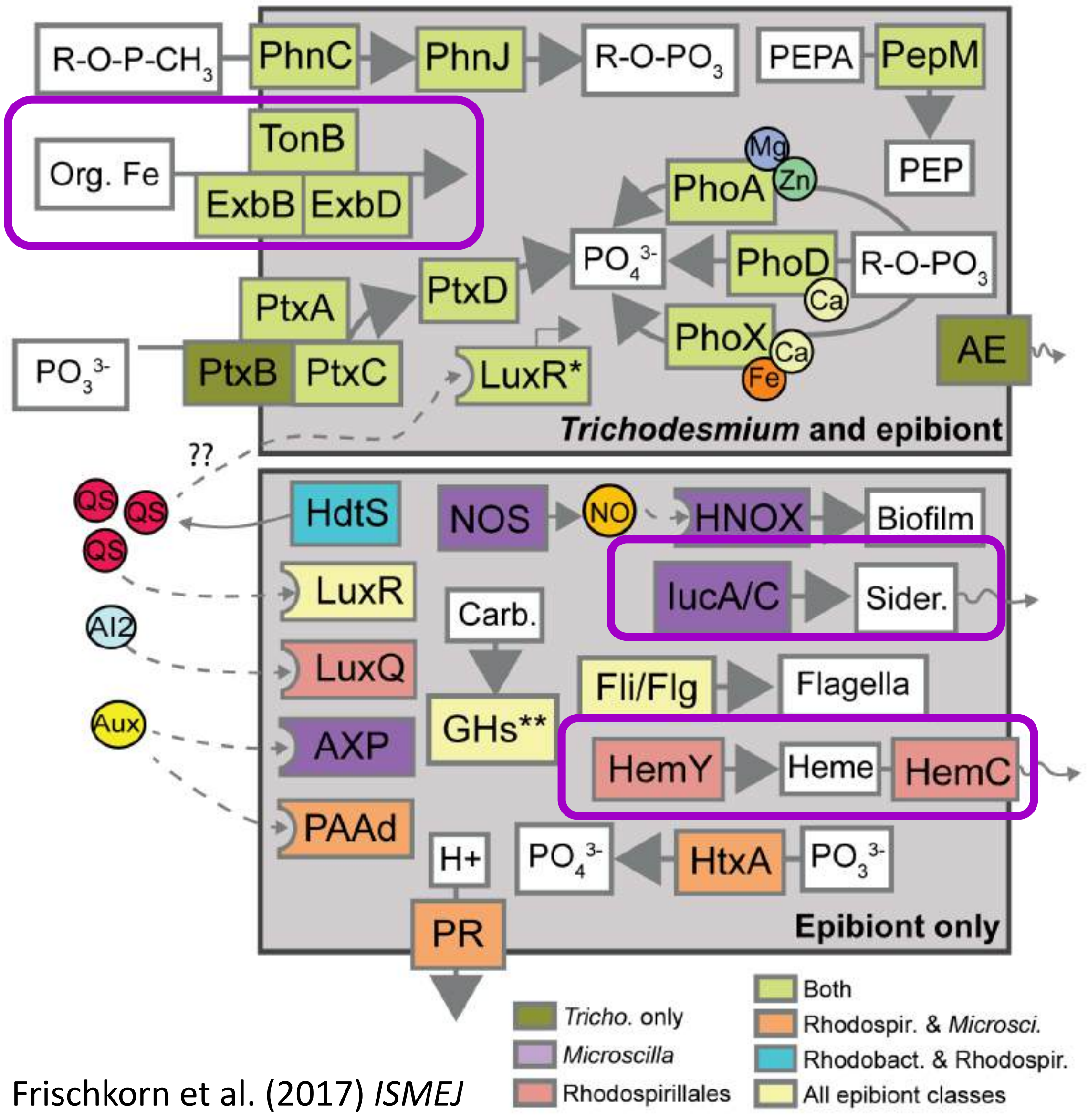
Chemical-microbe network within the *Trichodesmium* holobiont



Organic Fe, P, C, N
B vitamins

Moran et al. (2022) *Nature Micro.*

Metabolic partitioning within the *Trichodesmium* holobiont

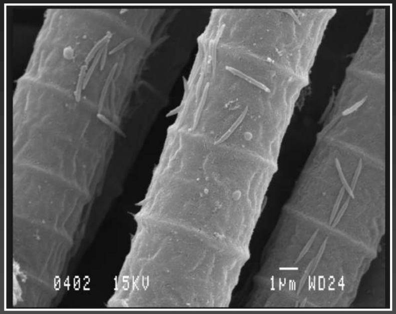


Frischkorn et al. (2017) *ISMEJ*

Organic Iron

Epibionts can produce organic iron complexes that likely modulate iron in the holobiont microenvironment

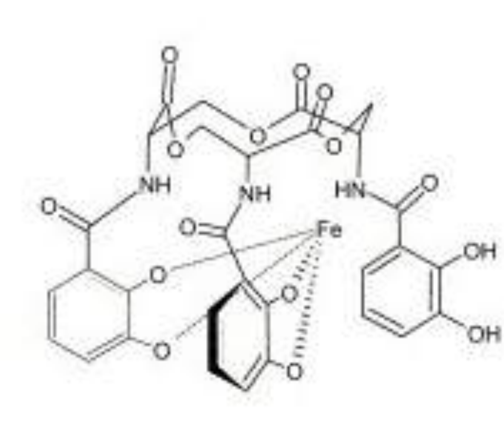
Cultured epibionts produce siderophores in low iron media



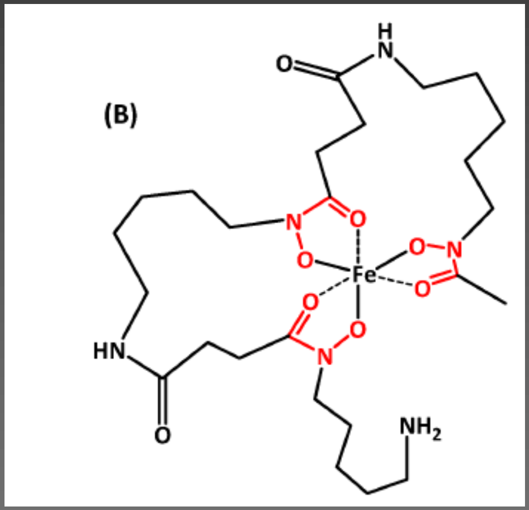
Isolate epibionts

LC-MS

Aerobactin



Desferrioxamine B

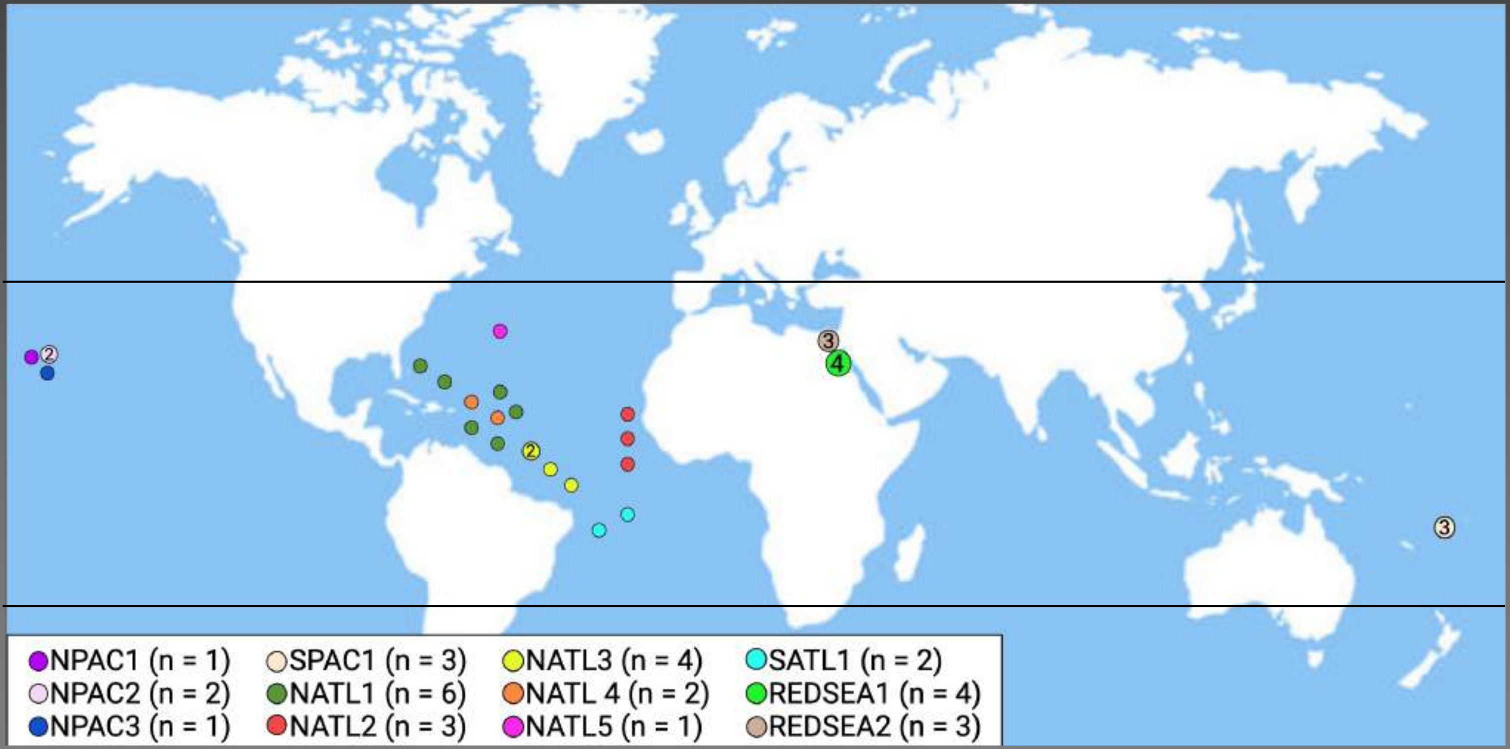


All six epibionts produce siderophores including those related to aerobactin and desferrioxamine consistent with MAG predictions and iron experiments.

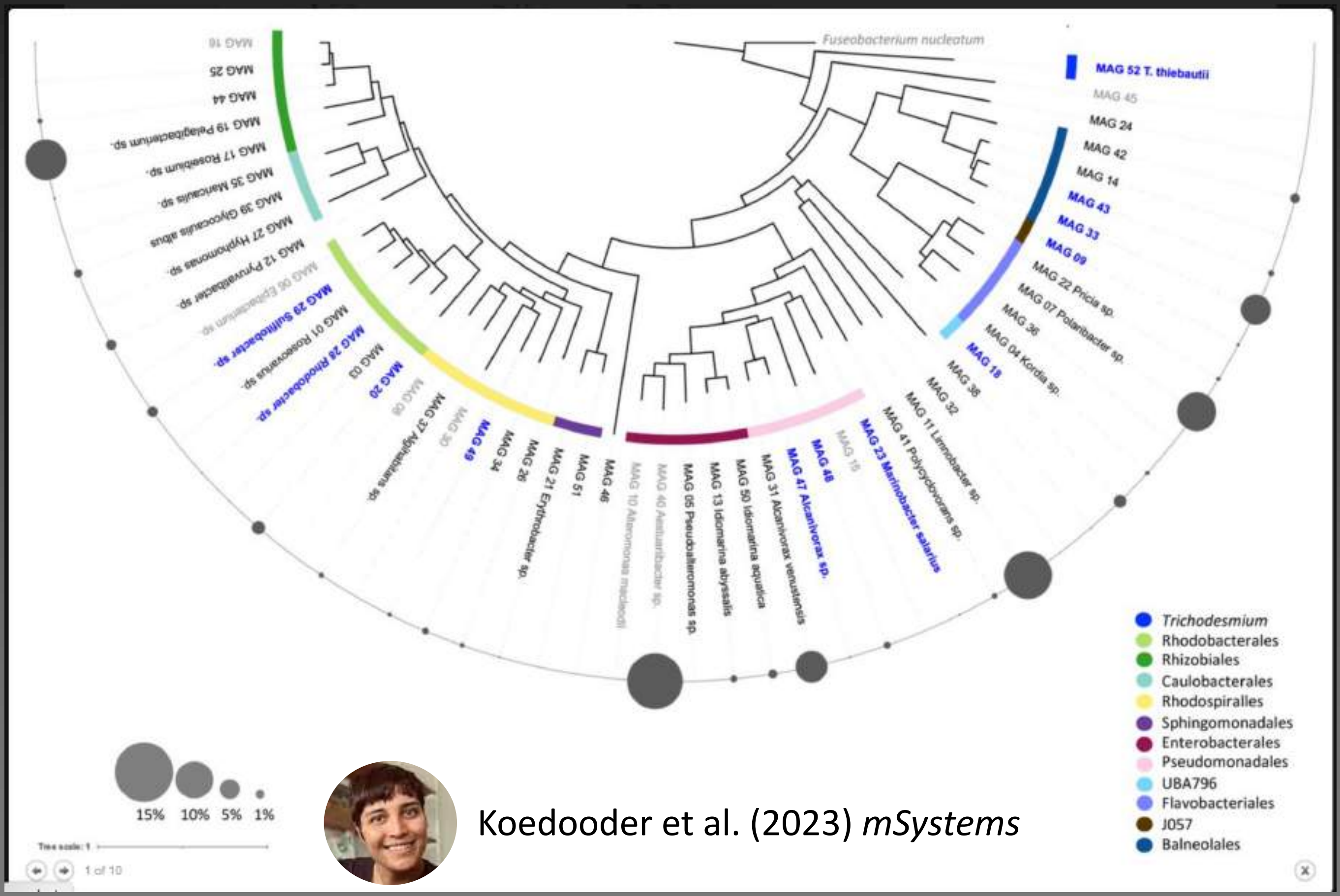
O. Baars et al. unpublished

More metagenome sequencing expands the analysis

What is the biogeography of *Trichodesmium* and epibiont traits as a function of changing ocean ecosystems?



Epibiont distribution in Red Sea samples



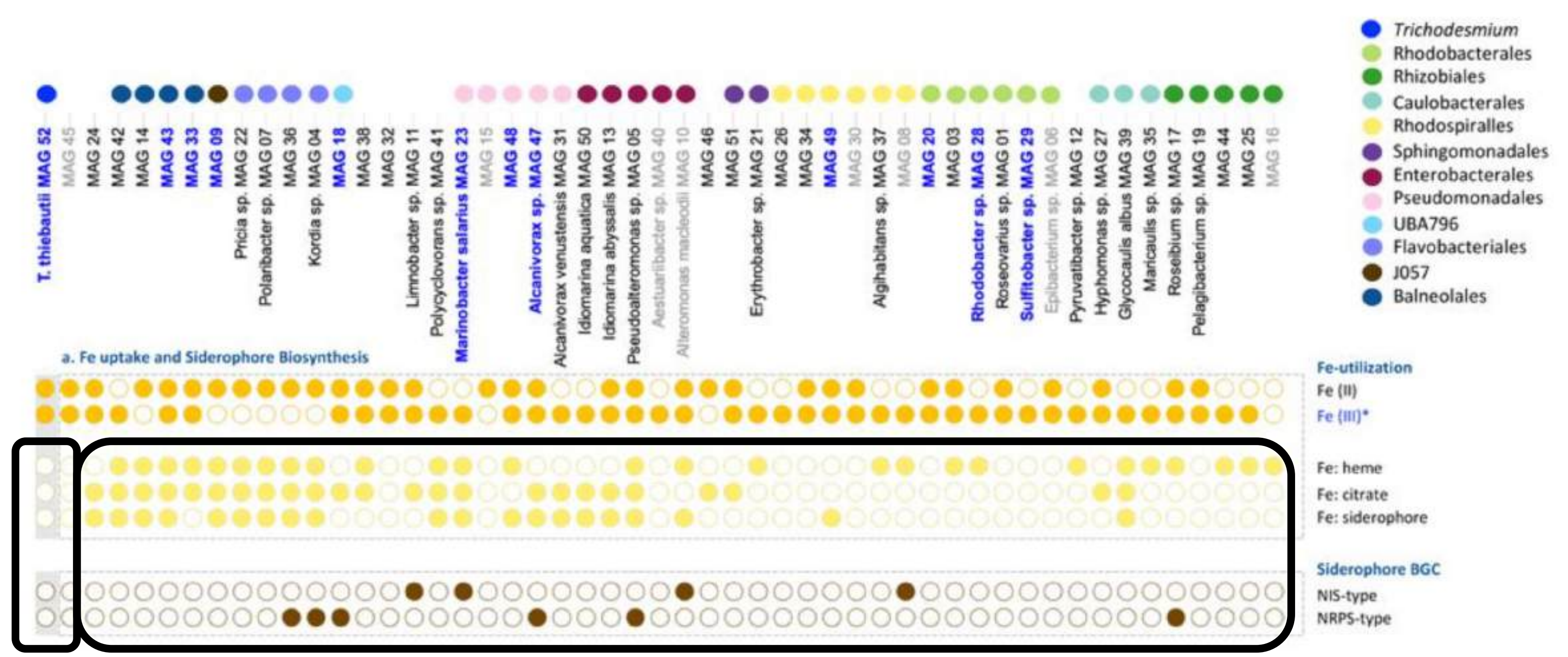
Koedooder et al. (2023) *mSystems*

43 >98% complete MAG

Broad similarities with North Atlantic samples

Ongoing effort in collaboration with Yeala Shaked

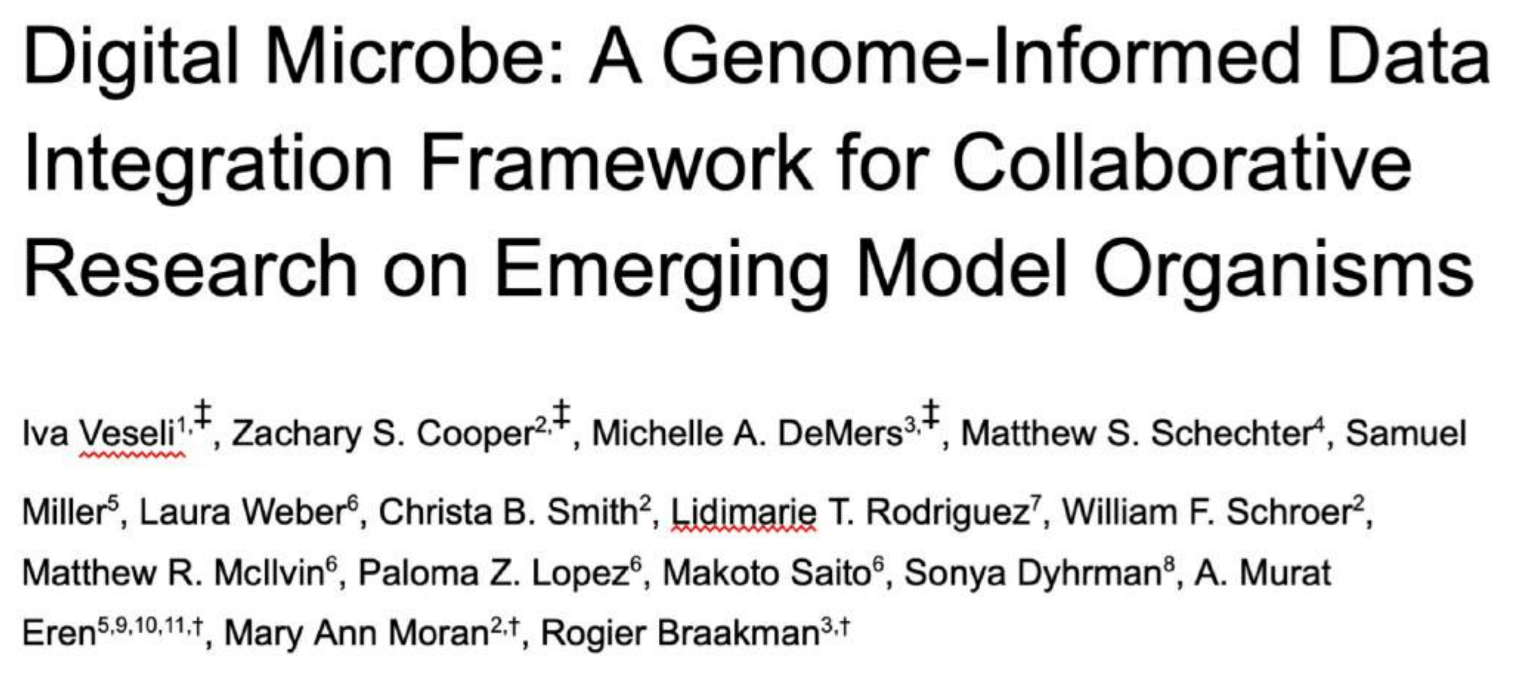
Red Sea epibionts carry enriched Fe traits relative to *Trichodesmium*



Koedooder et al. (2023) *mSystems*

User need for genome/pan genome data integration

Common frameworks and accessible data types facilitate team-science.



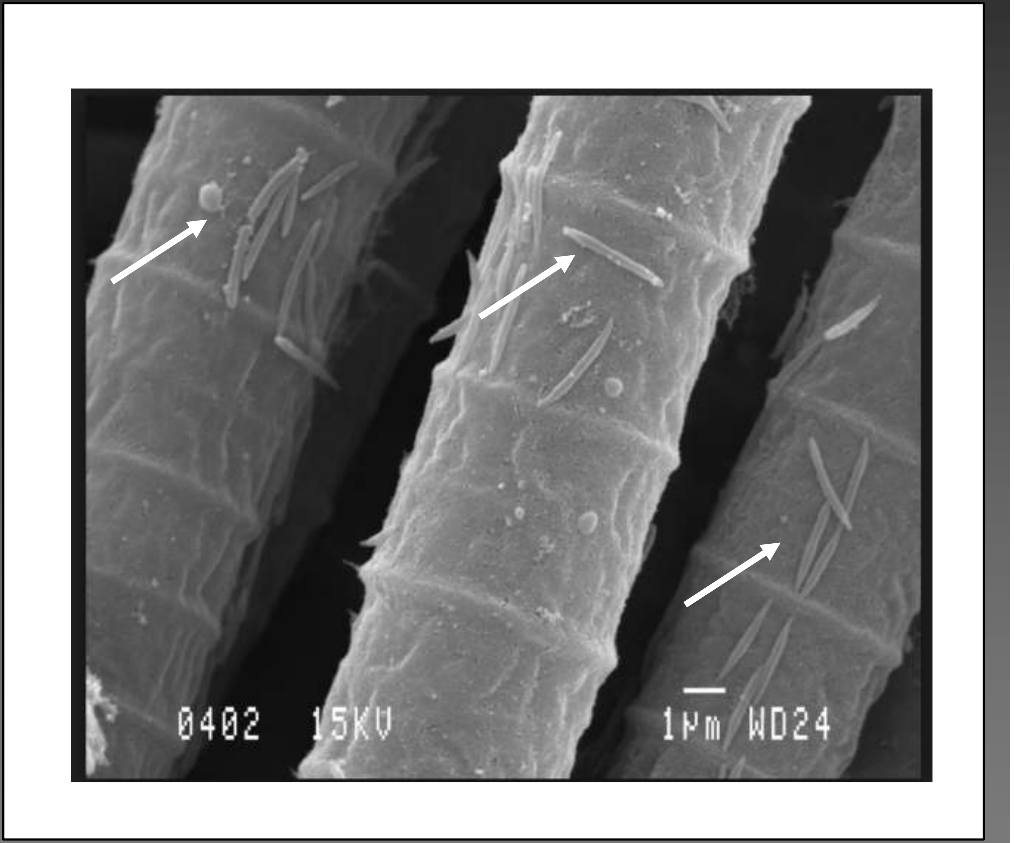
Digital Microbe: A Genome-Informed Data Integration Framework for Collaborative Research on Emerging Model Organisms

Iva Veseli[1,‡], Zachary S. Cooper[2,‡], Michelle A. DeMers[3,‡], Matthew S. Schechter[4], Samuel Miller[5], Laura Weber[6], Christa B. Smith[2], Lidimarie T. Rodriguez[7], William F. Schroer[2], Matthew R. McIlvin[6], Paloma Z. Lopez[6], Makoto Saito[6], Sonya Dyhrman[8], A. Murat Eren[5,9,10,11,†], Mary Ann Moran[2,†], Rogier Braakman[3,†]

Veseli et al. (2024) *Scientific Data*

Summary

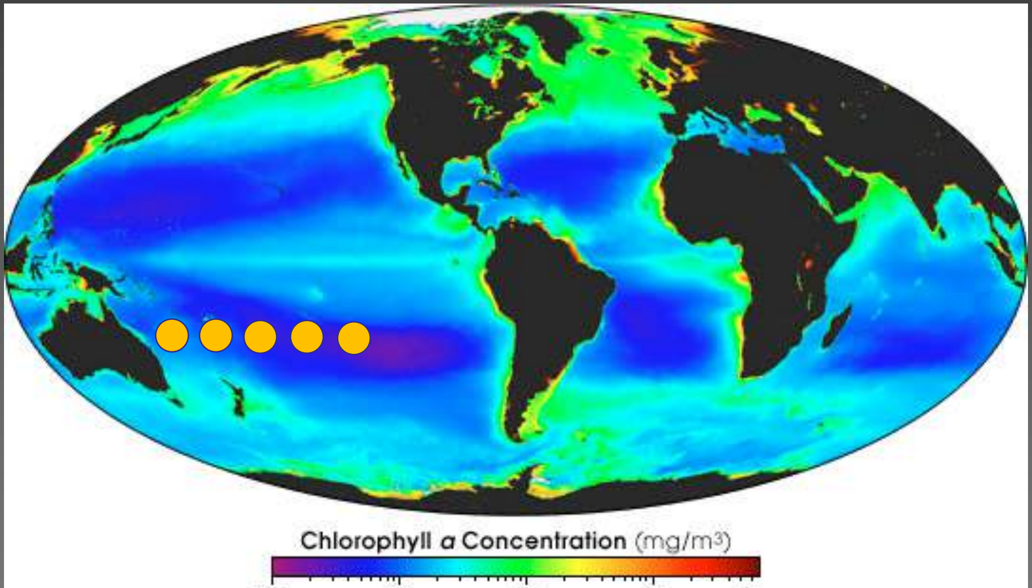
- *Trichodesmium* epibionts are functionally diverse, and these functions suggest complex interactions within holobiont (substrates, signals)
- Common workflows and accessible data types help facilitate collaborative-science.
- Now we are looking at how those metagenomic functions are expressed…..



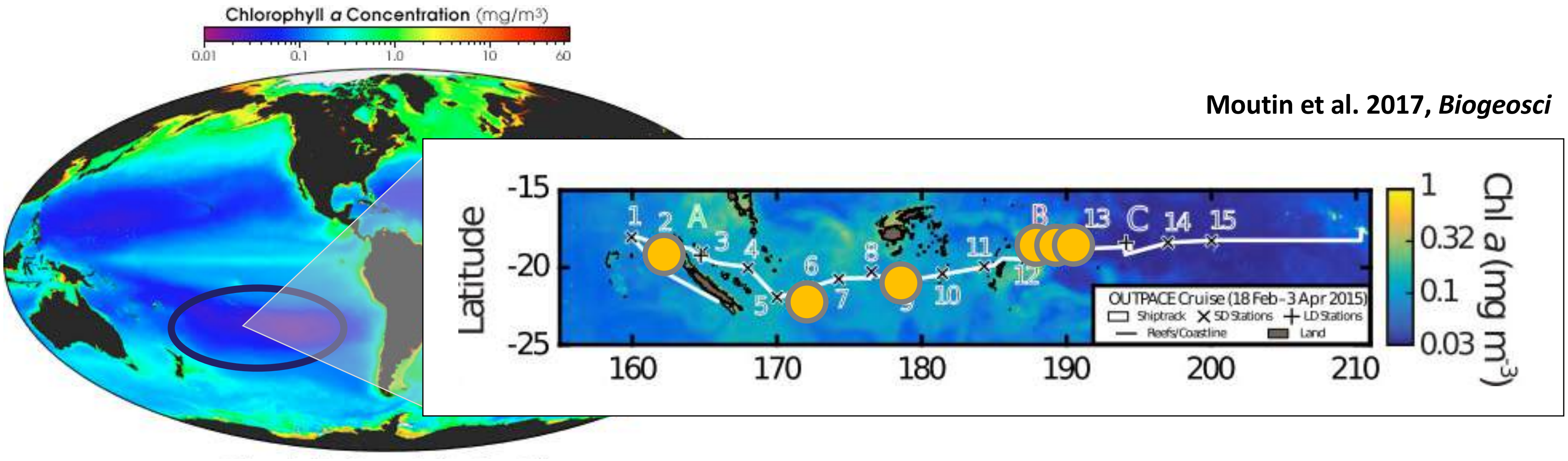
Coordinated analysis of phosphorus cycling

- South Pacific is undersampled and the dynamics of the *Trichodesmium* holobiont are not well understood
- Unique opportunity to sample metagenome, metatranscriptome, and key activities.
- Is there evidence of holobiont phosphorus reduction and cycling in this environment?

OUTPACE



Physiological ecology of *Trichodesmium* and its microbiome in the western tropical South Pacific

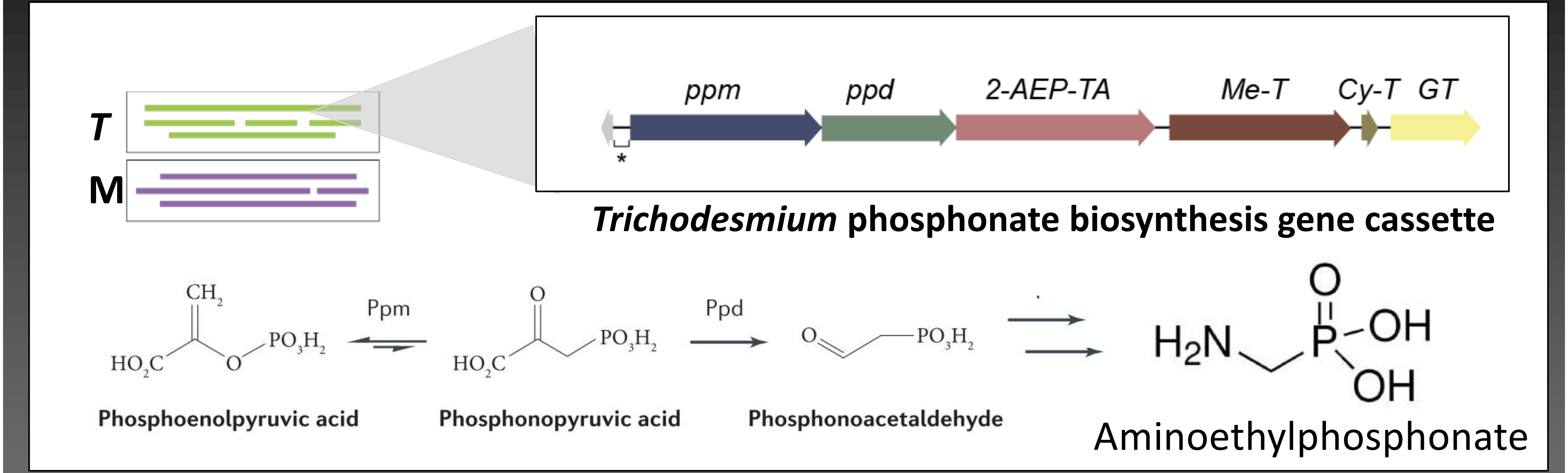


Moutin et al. 2017, *Biogeosci*

● ***Trichodesmium***
- ***Metagenome***
- ***Metatranscriptome***
- ***P reduction***

Metatranscriptomes

Metagenomic evidence of P reduction



***Trichodesmium* phosphonate biosynthesis gene cassette**

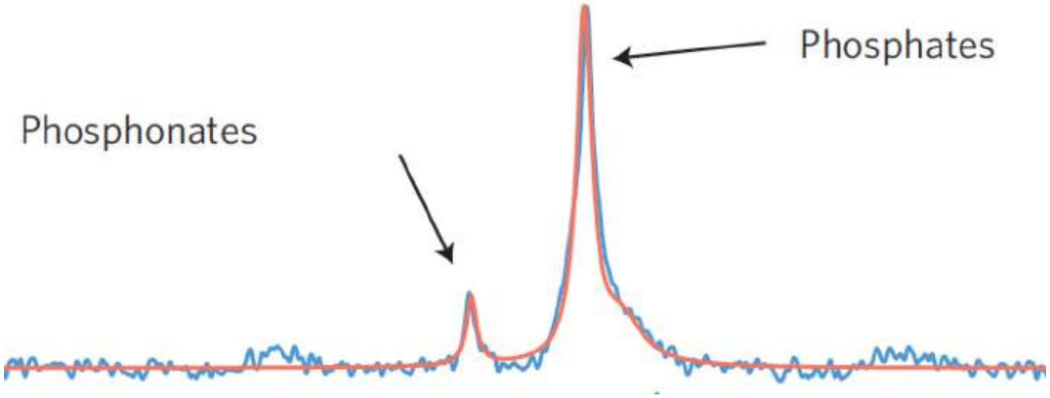
Complete gene set for phosphonate biosynthesis assembled in *Trichodesmium* genome bin (MAG).

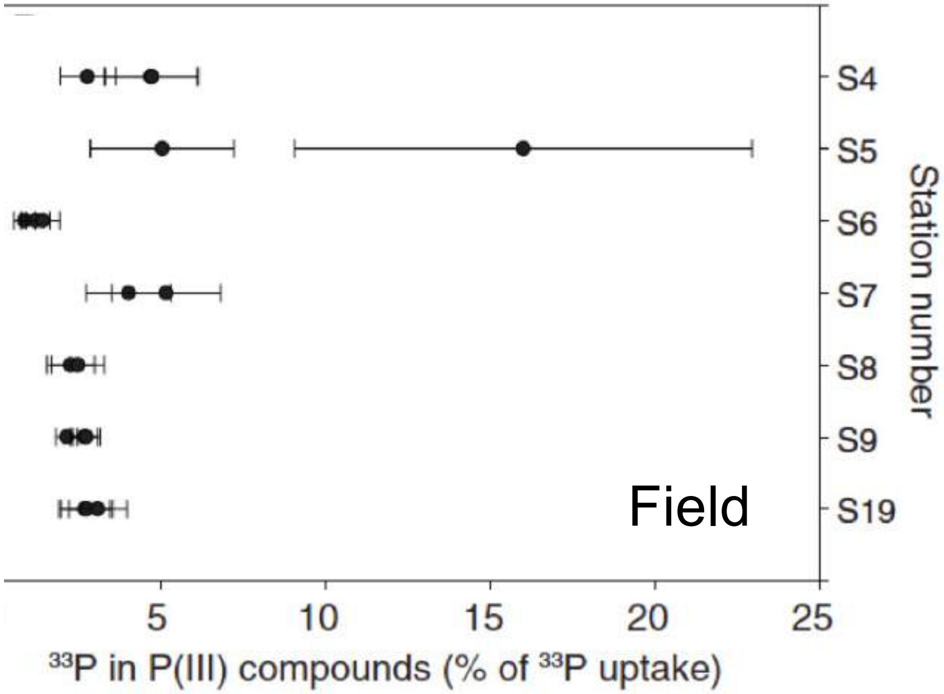
Frischkorn et al. (2018) *Biogeoscience*

Answers to enduring mysteries… who makes C-P compounds?

Phosphonate (C-P) biosynthesis

Enrichment Culture




Field

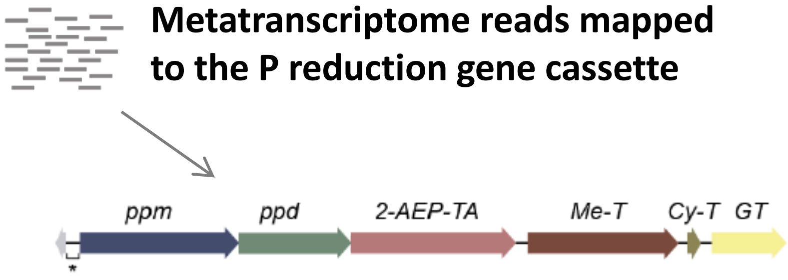
Dyhrman et al. (2009) *Nature Geo.*

Van Mooy et al. (2015) *Science*

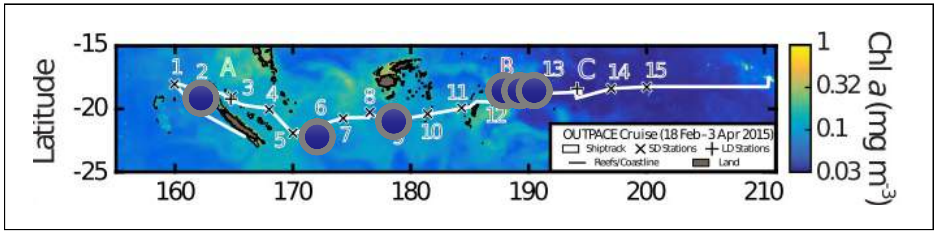
Phosphonates are produced at high rates in the holobiont - metagenomic signal suggests this is related to *Trichodesmium* production

Genes are expressed with P reduction

Measure gene expression

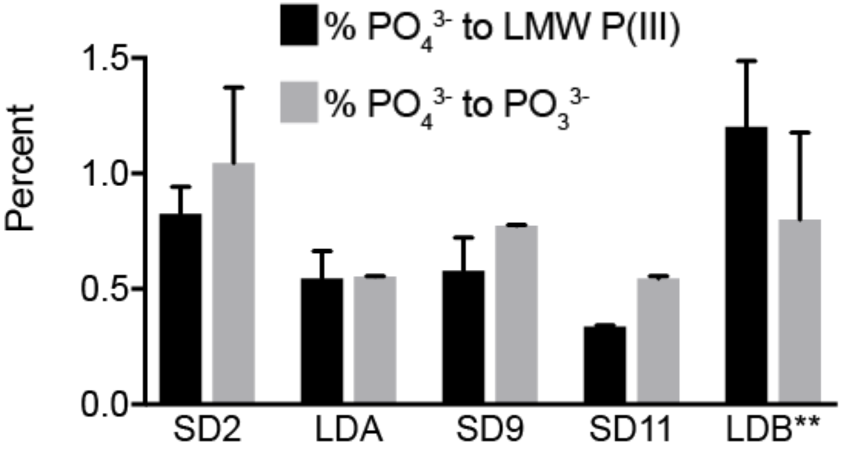

Metatranscriptome reads mapped to the P reduction gene cassette

Frischkorn et al. (2018) *Biogeoscience*



Measure phosphate reduction

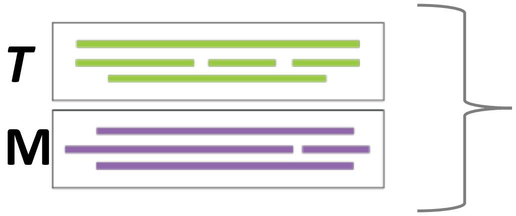
Percentage of radiolabeled phosphate taken up and reduced by *Trichodesmium* colonies



*methylphosphonate, phosphonoacetylaldehyde, or 2-aminoethylphosphonate

- ✓ Genes detected
- ✓ Genes expressed
- ✓ Activity measured
- ? P cycling

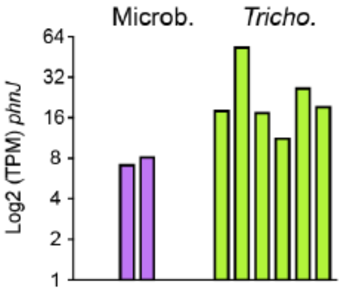
Evidence for metabolism of reduced phosphorus compounds in *Trichodesmium* and the microbiome

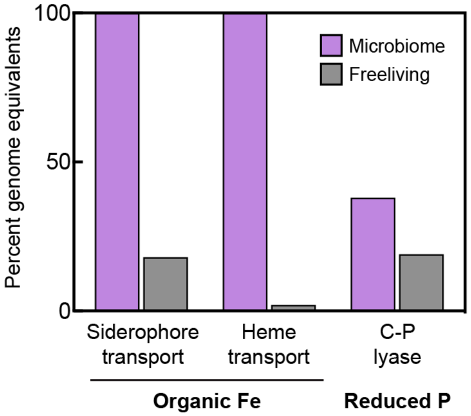


Query for C-P metabolism: Phosphonate C-P lyase (*phnJ*)

✓ **Genes detected**

✓ **Genes expressed**

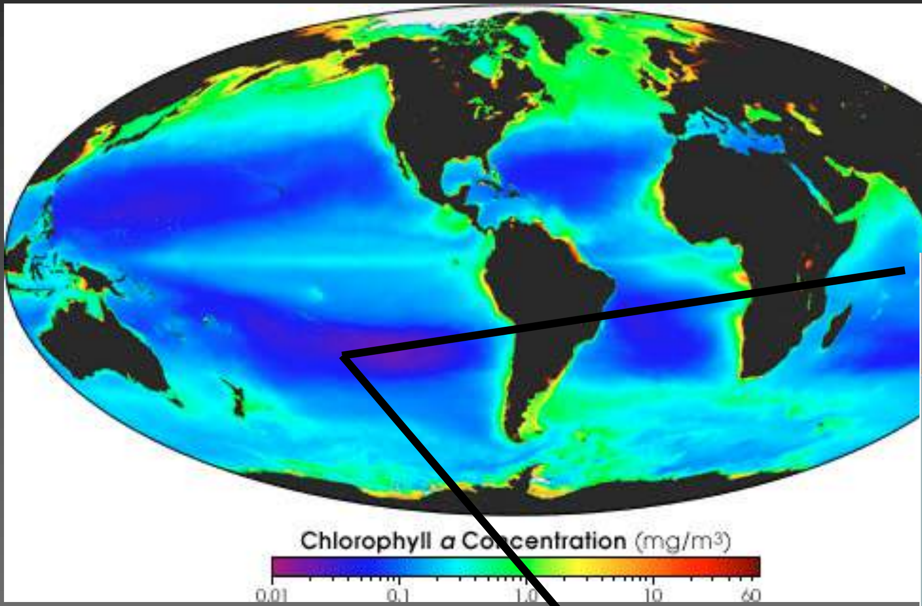


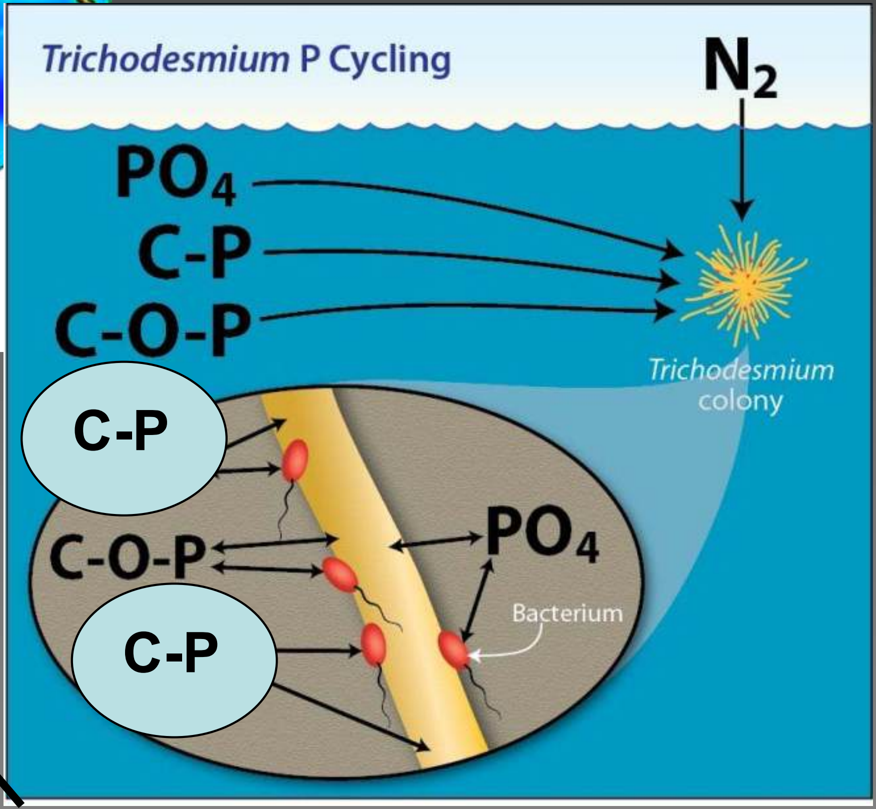


Ability to metabolize reduced P is enriched in *Trichodesmium* microbiome

Frischkorn et al. (2018) *Biogeoscience*

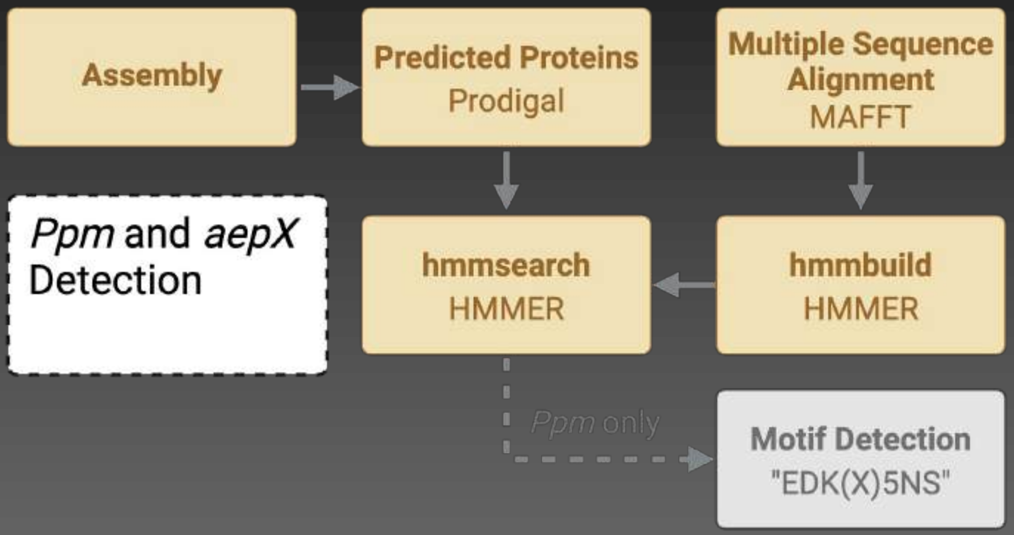
Novel P currencies used in holobiont….





Mitigate competition
for scarce resource?

Ongoing analysis of phosphonate biosynthesis and uptake biogeography



Ppm and *aepX* Detection

Assembly

Predicted Proteins
Prodigal

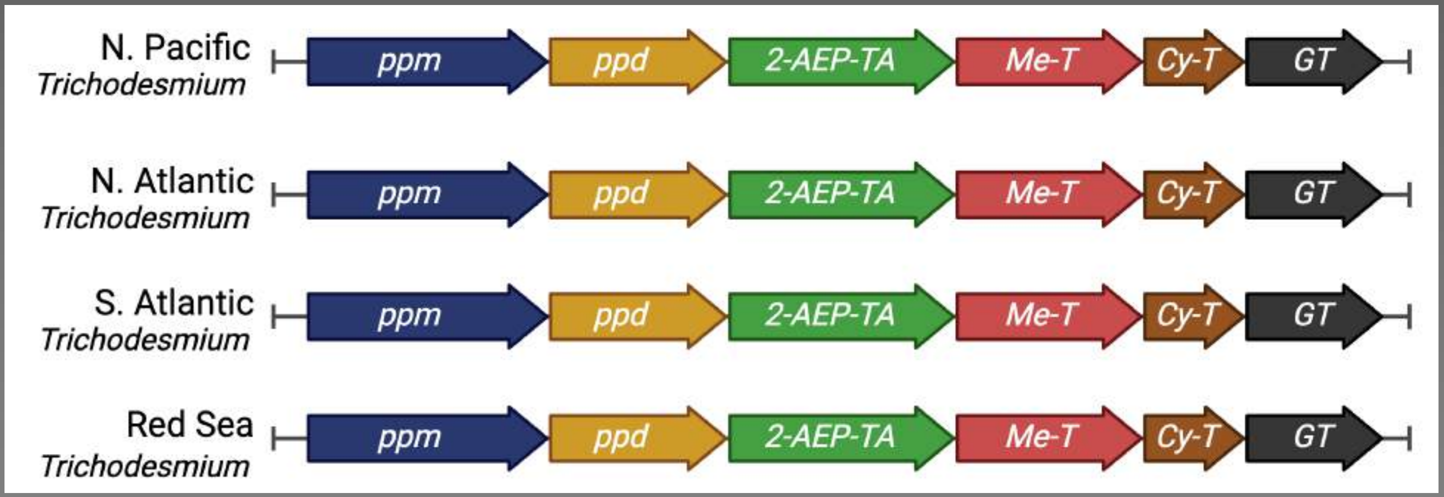
Multiple Sequence Alignment
MAFFT

hmmsearch
HMMER

hmmbuild
HMMER

Ppm only

Motif Detection
"EDK(X)5NS"



Maybach et al. unpubl.

Summary

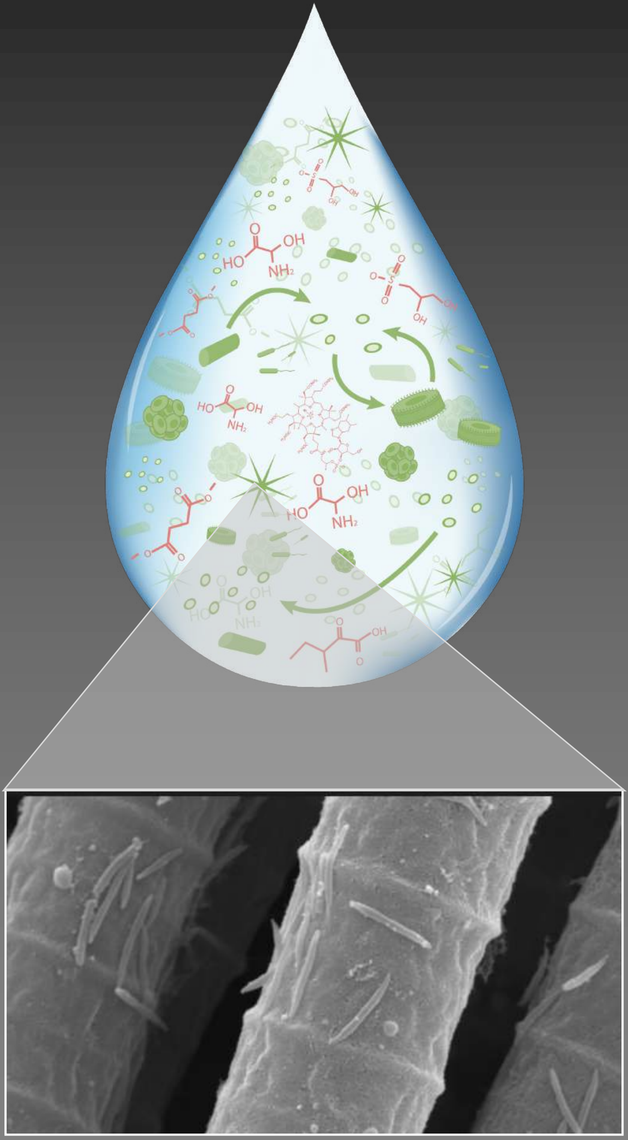
- Need to think about *Trichodesmium* as a holobiont with a broader metabolic potential than just *Trichodesmium* alone.

- Microbiome is distinct and varies as a function of environment

- Holobiont is a potential hot-spot of reduced phosphorus cycling.

- *Consistency of epibiont diversity and functional capacity across environments?*
    - Metagenomes and metatranscriptomes from different environments

Core questions about the *Trichodesmium* microbiome



Who is there?

- **Diverse community, distinct from water column**

What are they doing?

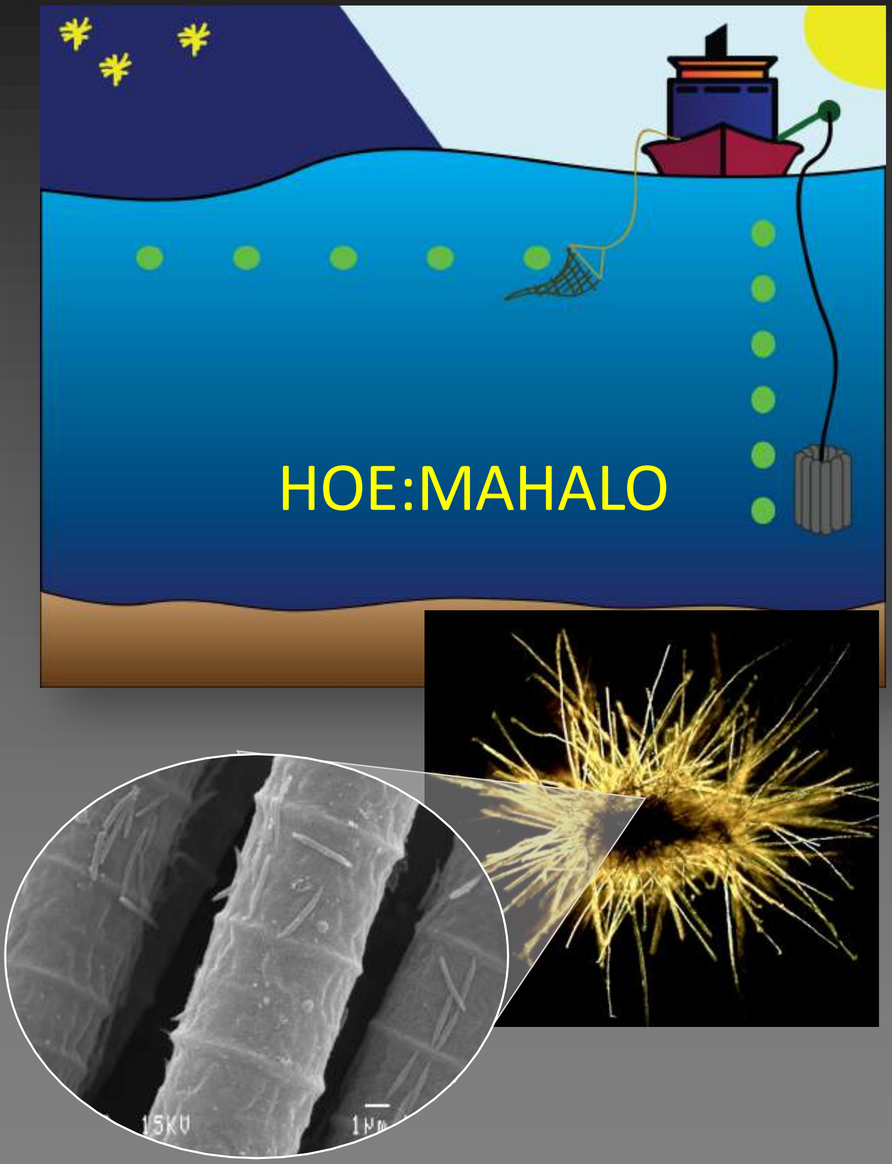
- **Microbiome contributes functional diversity to community**

Are they interacting?

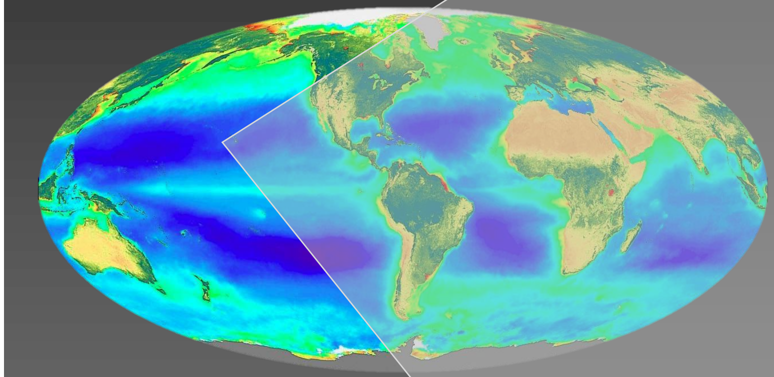
- Diel patterns in the holobiont

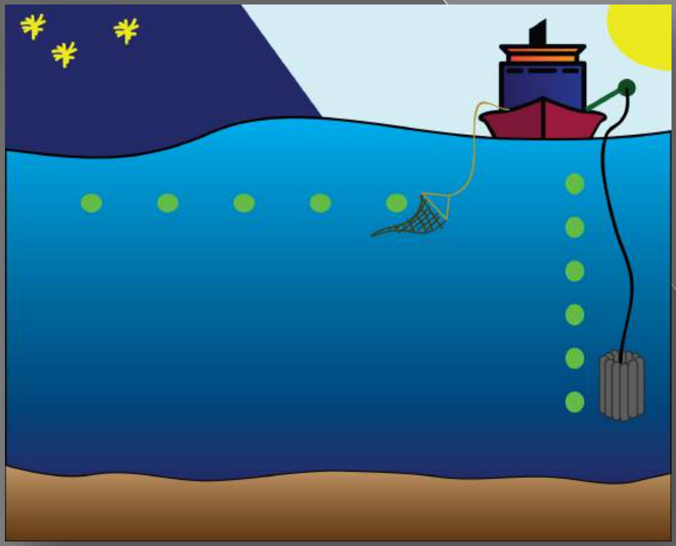
Using light to study interactions

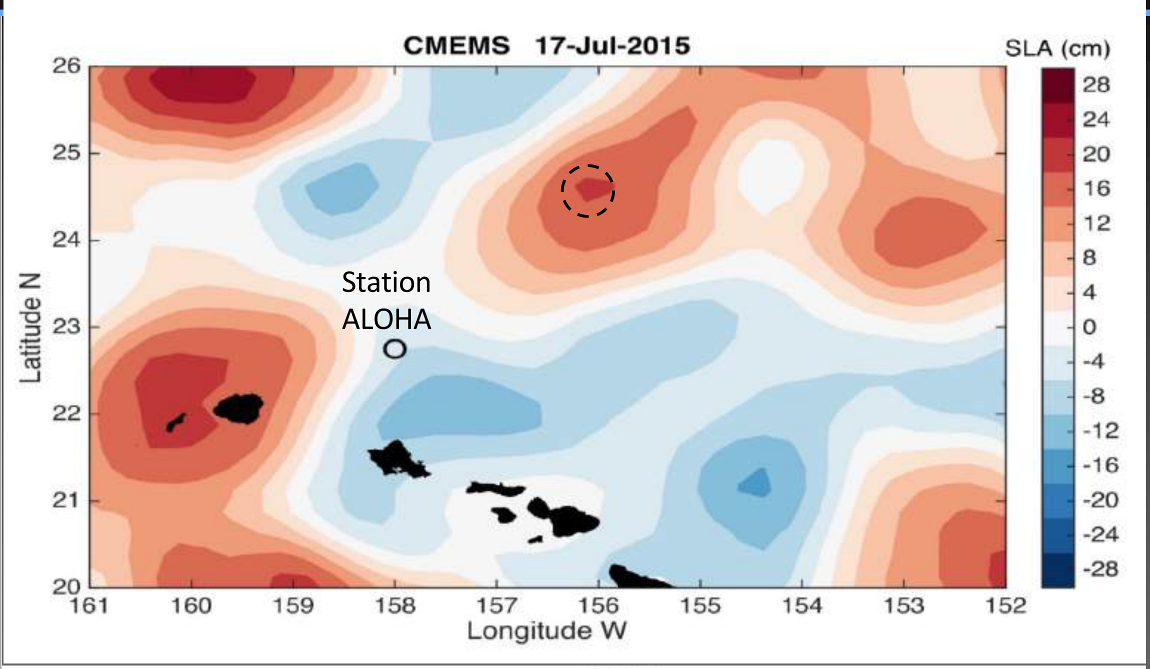
- Light drives diel changes in *Trichodesmium* CO_2 and N_2 fixation
- Do epibionts have diel responses, and what might they tell us about interactions between the host and microbiome?

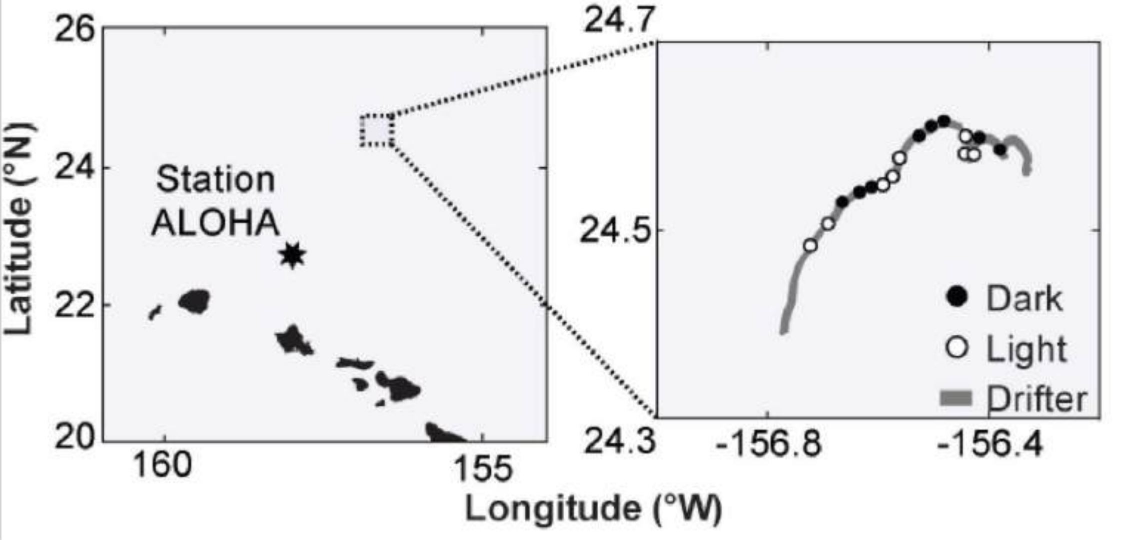


Sampling light driven responses in the NPSG

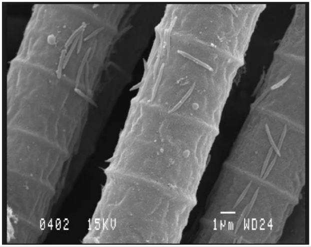








Coordinated expression dynamics in host and microbiome



Metatranscriptomes

↓

Orthologous group analysis

↓

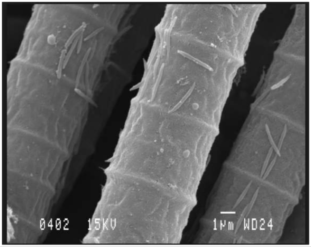
Periodicity: RAIN

Co-expression: WGCNA

https://omictools.com/rhythmicity-analysis-incorporating-nonparametric-methods-tool

https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-9-559

Coordinated expression dynamics in host and microbiome



Metatranscriptomes

↓

Orthologous group analysis

↓

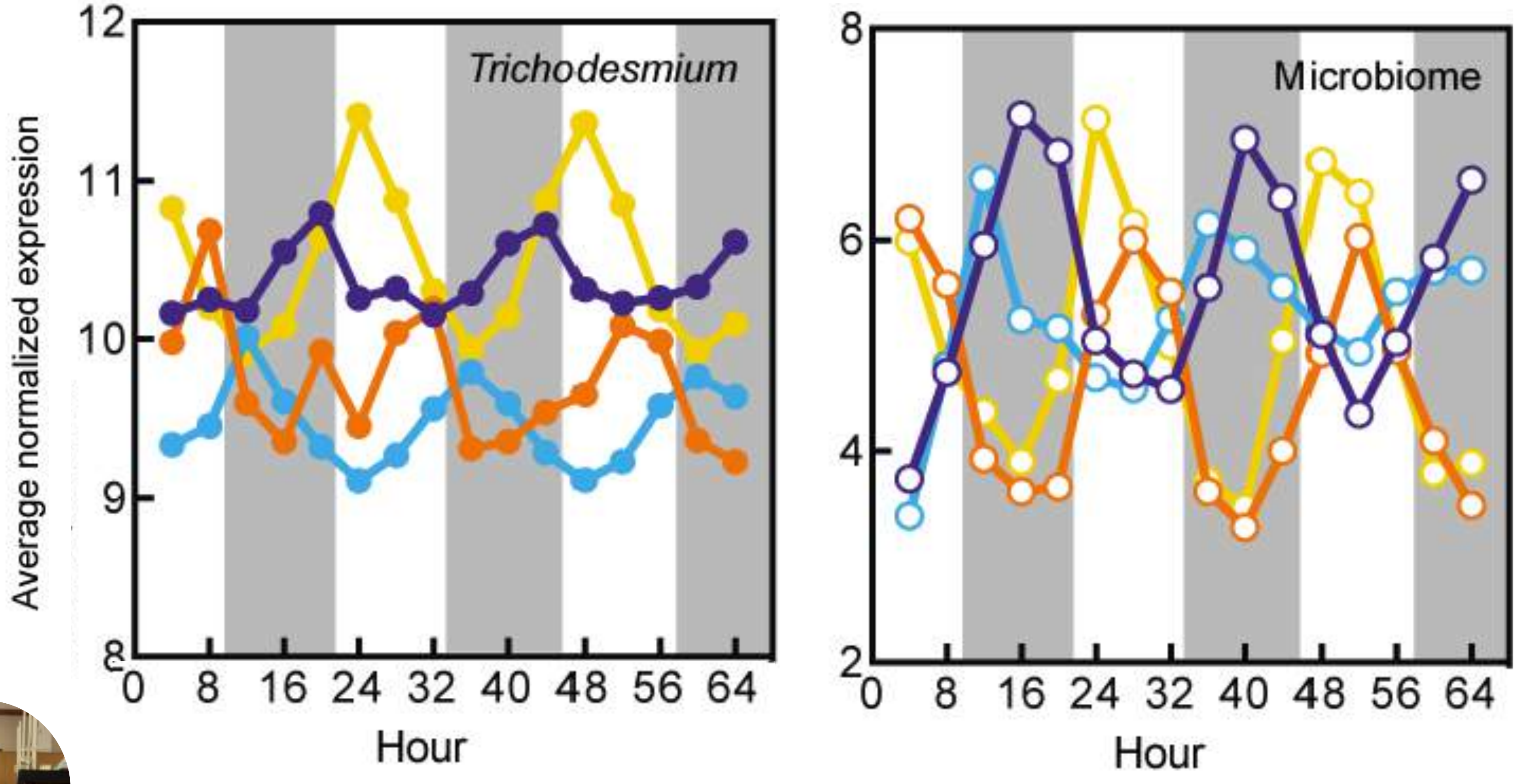
Periodicity: RAIN

Co-expression: WGCNA

https://omictools.com/rhythmicity-analysis-incorporating-nonparametric-methods-tool

https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-9-559

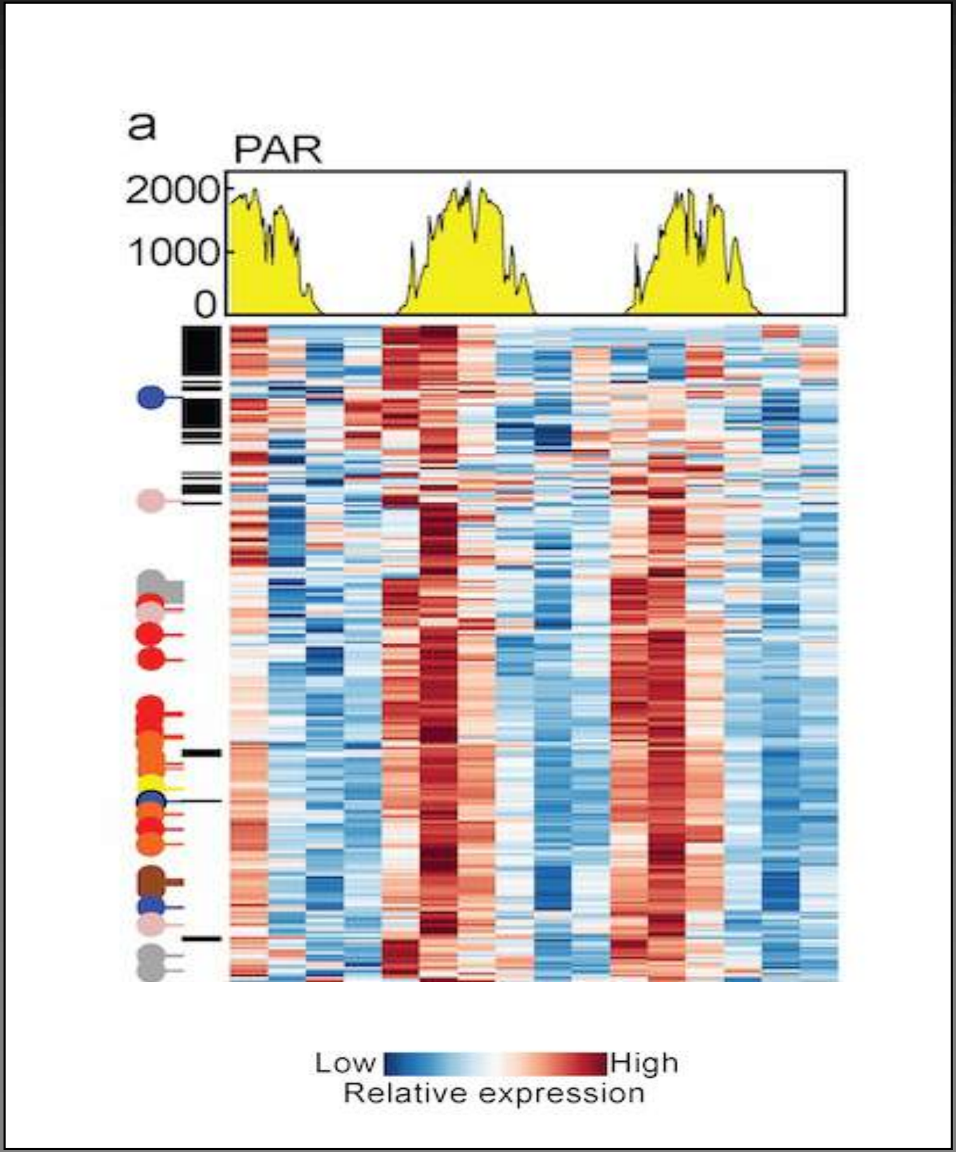
Diel modulation of transcripts in *Trichodesmium*



Frischkorn et al. (2018) *ISMEJ*

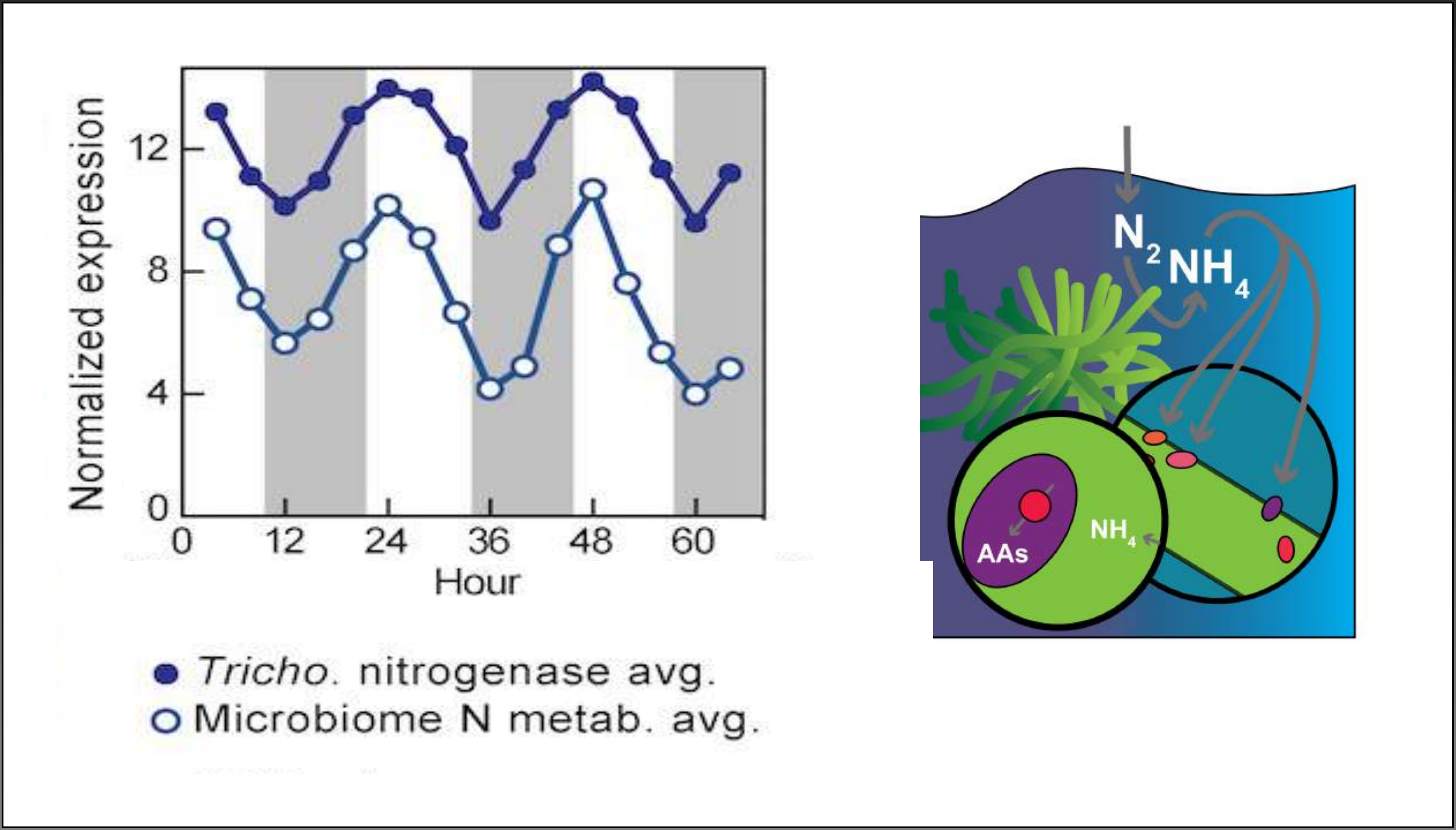
WGCNA co-expression network

Module example: Significantly coordinated expression patterns between *Trichodesmium* and microbiome.



Frischkorn et al. (2018) *ISMEJ*

Diel modulation of transcripts suggest coordination: Nitrogen



Frischkorn et al. (2018) *ISMEJ*

Summary

- The microbiome and *Trichodesmium* are interacting!
  - Nitrogen
  - Carbon
  - Vitamin B12, P, Fe etc.

- Heterotrophic epibionts have striking diel signals that likely underpin signaling, substrate exchanges and other activities in the holobiont.

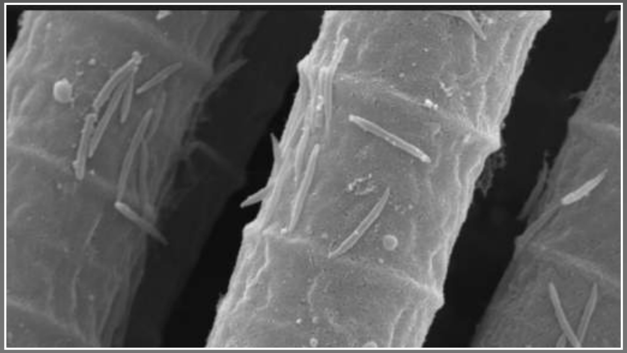
Core questions about the *Trichodesmium* microbiome

Who is there?

- **Diverse community, distinct from water column**

What are they doing?

- **Microbiome contributes functional diversity to community**



Host → Microbiome

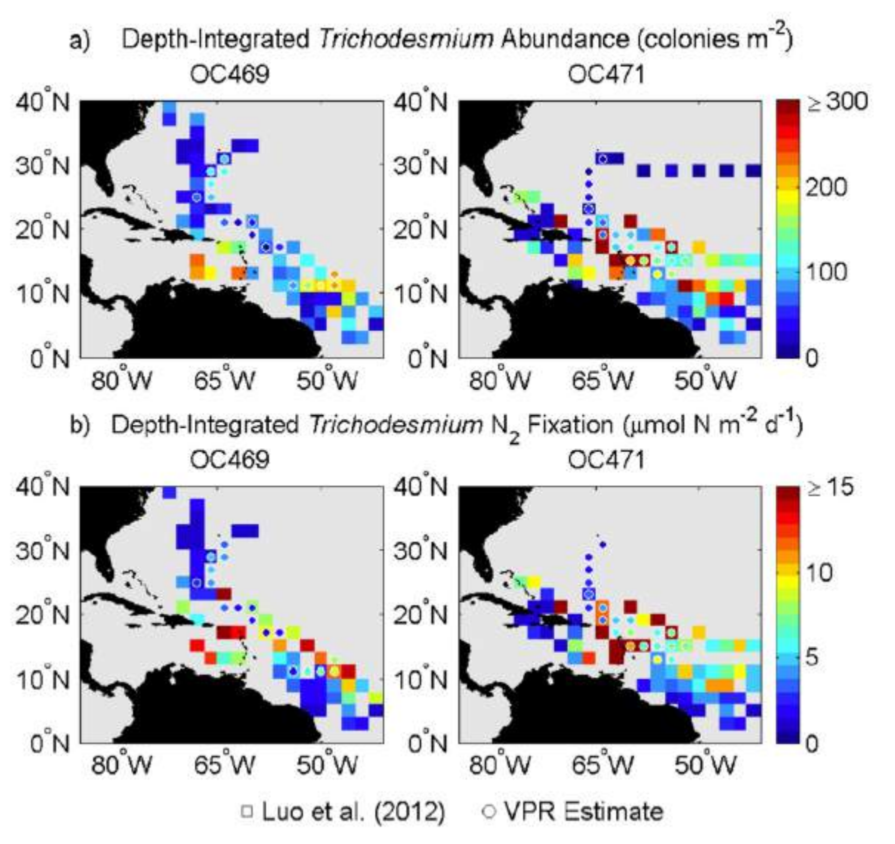
Microbiome → Host

Are they interacting?

- **Yes! Significant co-expression of host-microbiome genes**

Can the microbiome influence host N_2 fixation?

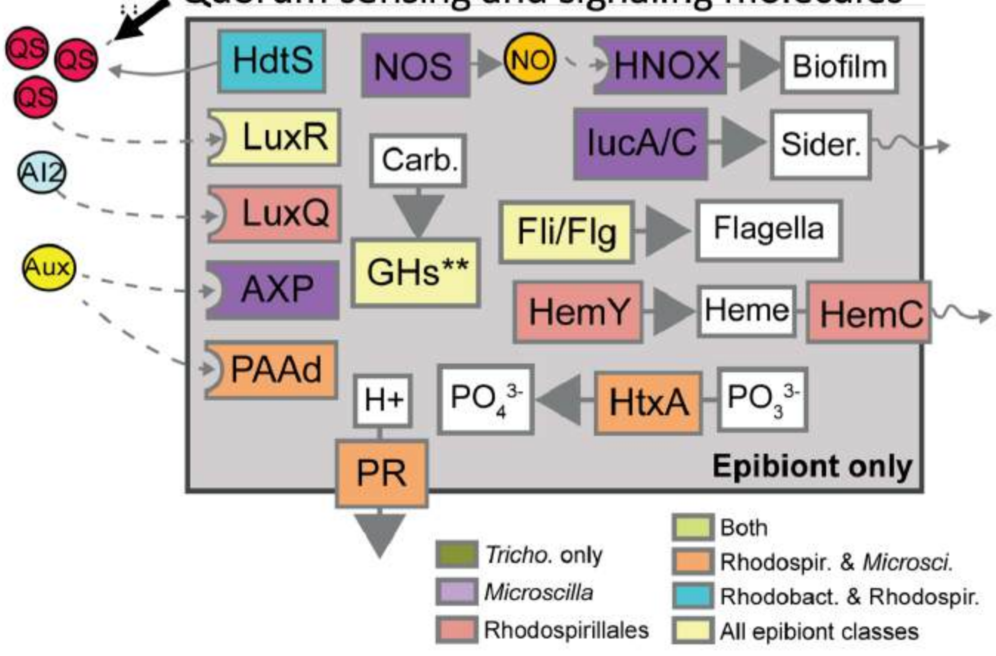
Maybe the models are missing the role of the microbiome!



Olson et al. 2015 *DSR* II

Selective manipulation of the microbiome…

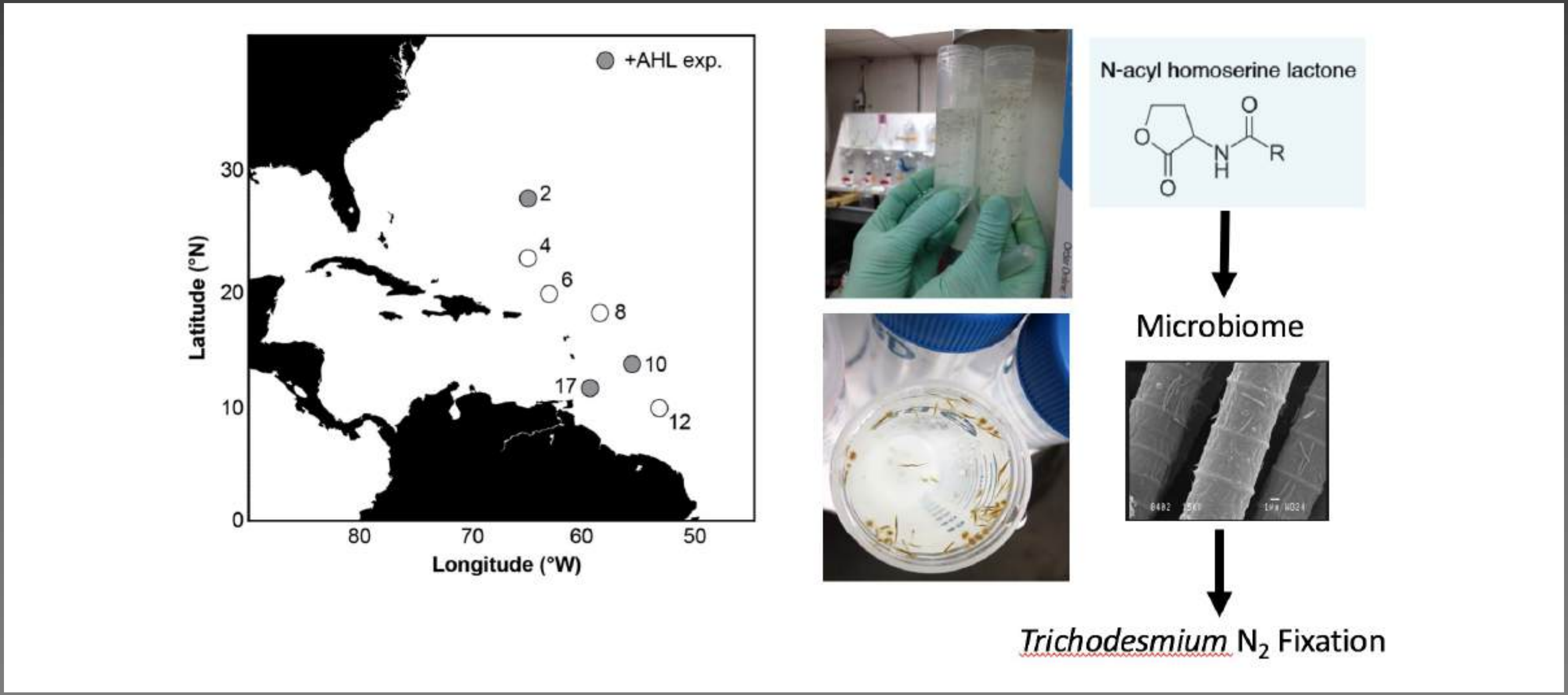
- Quorum sensing communication molecules (acylated homoserine lactones – AHL) detected in colonies (Van Mooy et al. 2012 *ISME J*)
- Genes for AHL biosynthesis and AHL receptors detected in epibionts not *Trichodesmium*



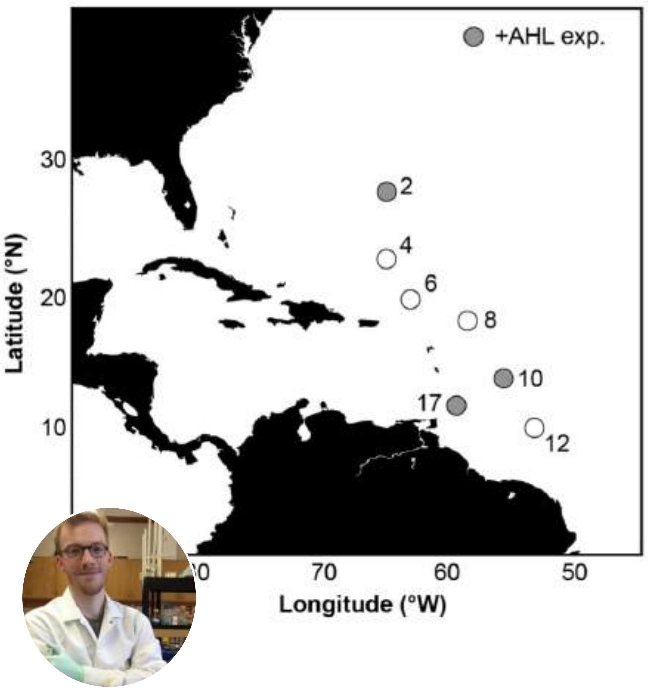
Quorum sensing and signaling molecules

Frischkorn et al. (2017) *ISMEJ*

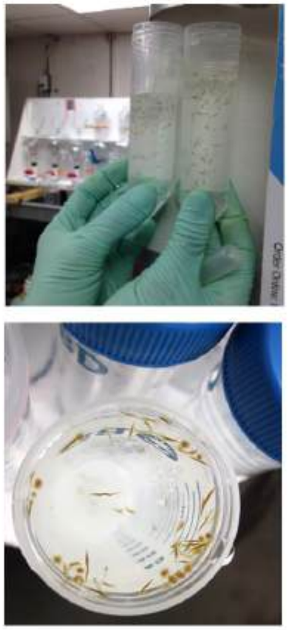
Do microbiome infochemicals influence *Trichodesmium* phenotype?

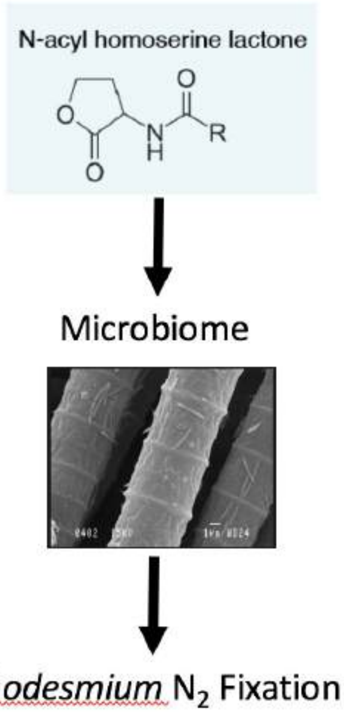


AHL addition modulates *Trichodesmium* N_2 fixation



Frischkorn et al. (2018) *L&O Letters*

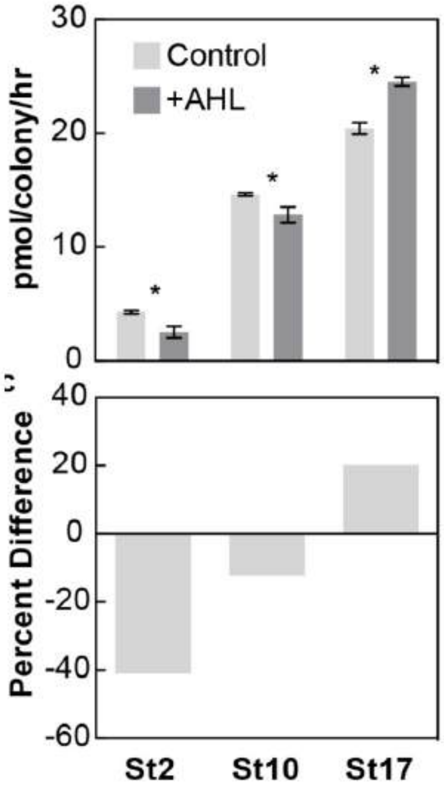




N-acyl homoserine lactone

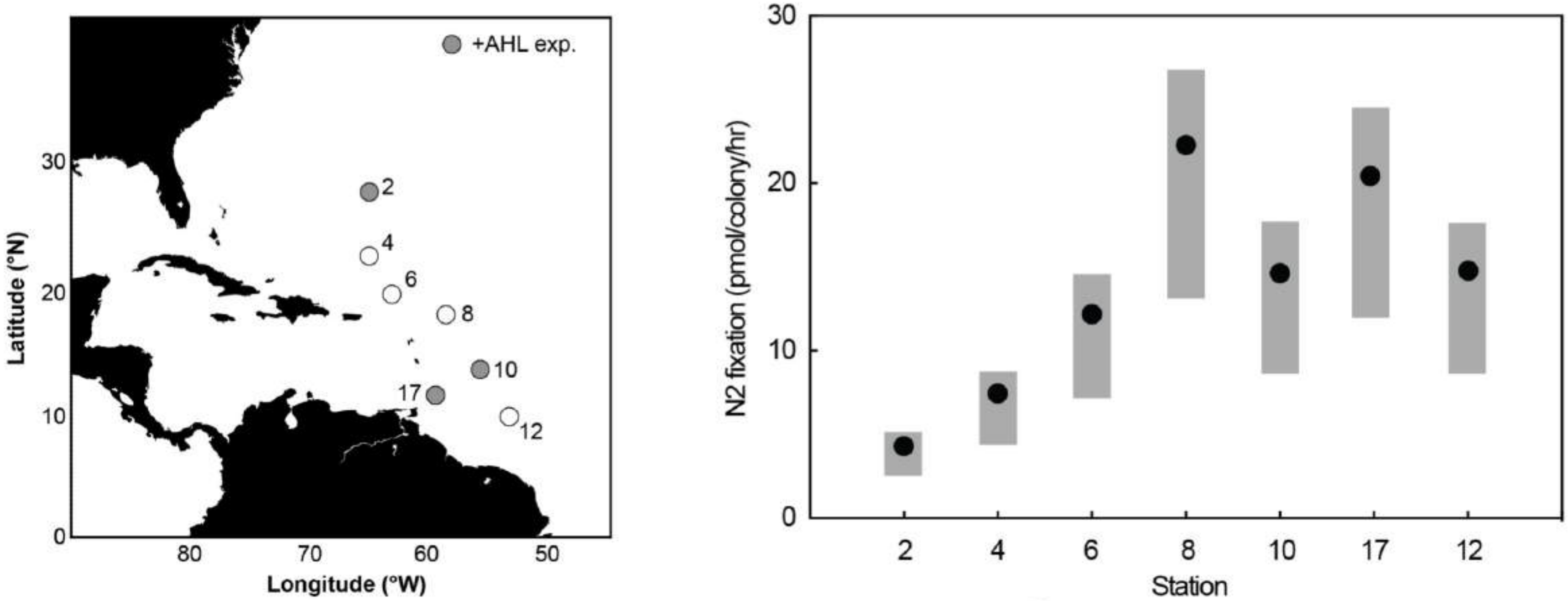
Microbiome

Trichodesmium N_2 Fixation



Microbiome can modulate host N_2 fixation rates up to 40%, in orthogonal directions, and independently of known drivers (P, Fe, light, temperature)

Microbiome activities modulate N_2 fixation in host

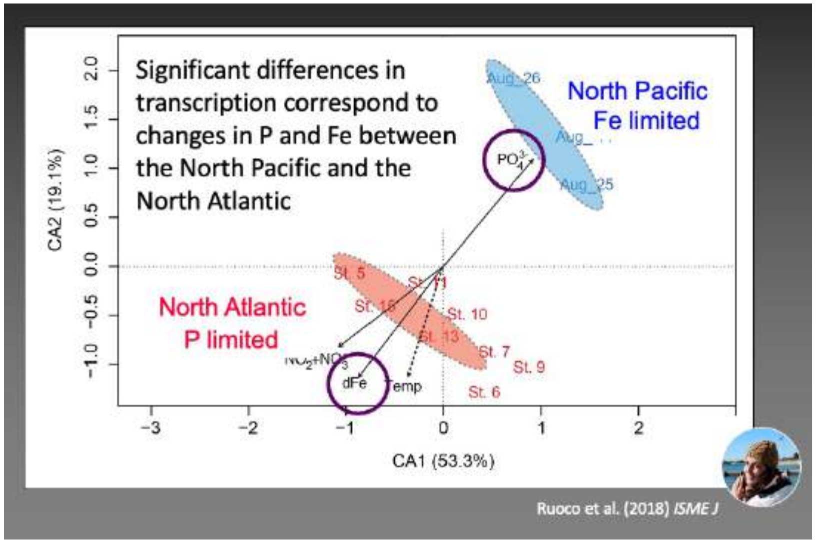


Frischkorn et al. (2018) *L&O Letters*

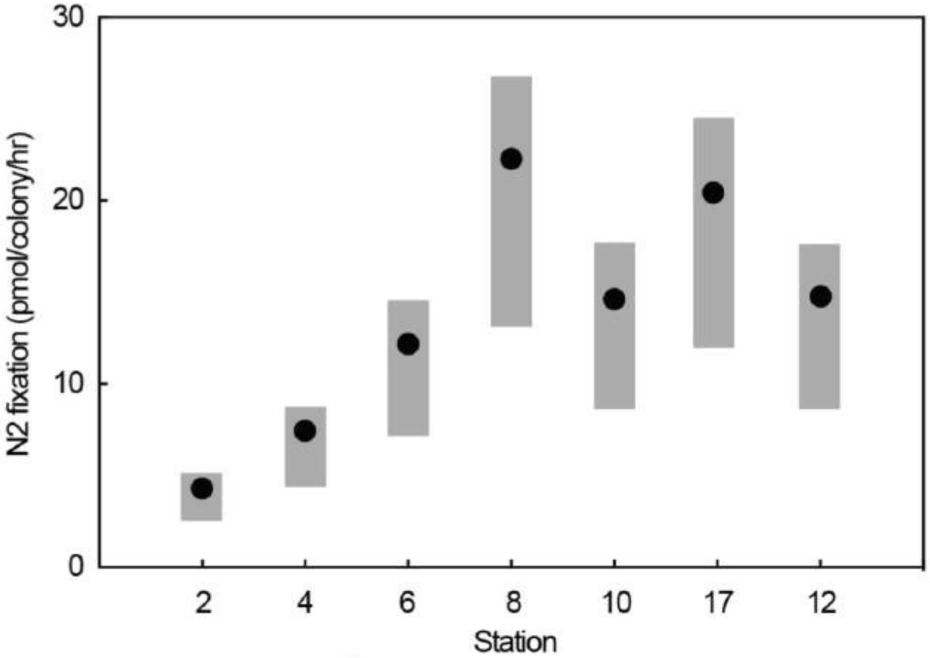
Biological interactions are a previously unrecognized driver of N_2 fixation

Need to revise modeling efforts

Microbiome activities modulate N_2 fixation in host


Significant differences in transcription correspond to changes in P and Fe between the North Pacific and the North Atlantic

Ruoco et al. (2018) *ISME J*



Frischkorn et al. (2018) *L&O Letters*

Biological interactions are a previously unrecognized driver of N_2 fixation
Need to revise modeling efforts to include the geochemistry (e.g. Fe and P) and the biology (interactions)!

Summary

- The microbiome and *Trichodesmium* are interacting!
- Epibionts can influence N_2 fixation, but the full mechanisms that underpin this still need to be identified.

Core questions about the *Trichodesmium* microbiome

Who is there?

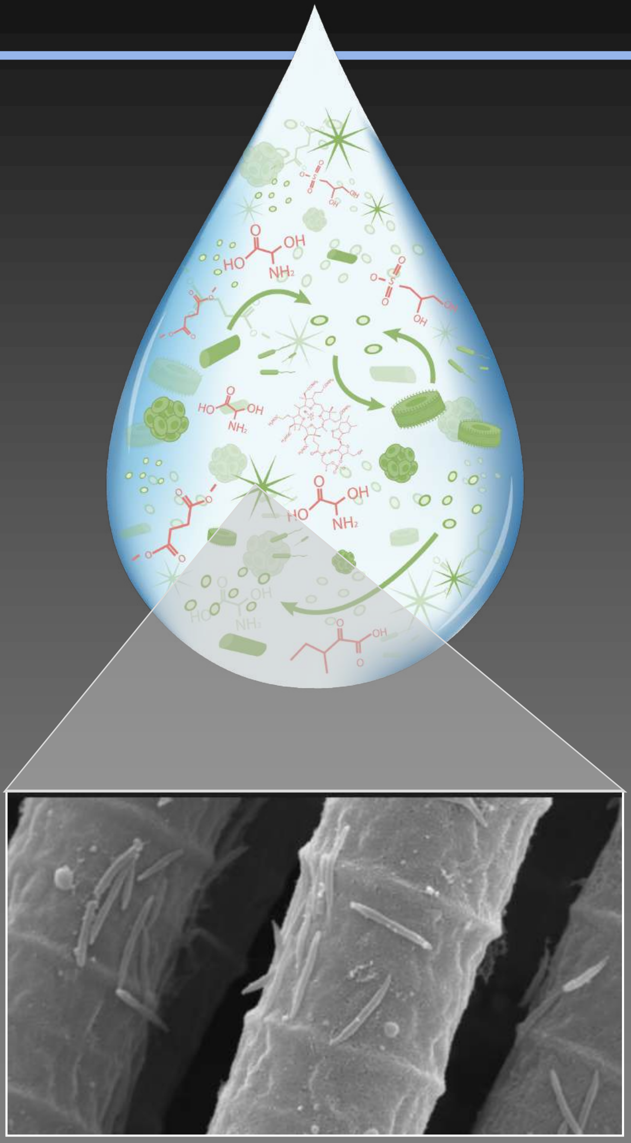
- **Diverse community, distinct from water column**

What are they doing?

- **Microbiome contributes functional diversity to community**

Are they interacting?

- **Yes! Significant co-expression of host-microbiome genes**



Microbiome can influence host N_2 fixation!

Summary

- **16S community amplicon sequencing**: Colonies harbor diverse epibiont microbiomes
- **Metagenomics**: Epibionts confer substantial metabolic potential, and are dynamically curated across gradients in the environment
- **Metatranscriptomics**: Novel P cycling currencies may avoid competition
- Apparent interactions between *Trichodesmium* and its microbiome

Microbiome plays a critical role in *Trichodesmium* physiological ecology and biogeochemical cycling

Enjoy the rest of the workshop!

'Omics approaches are providing new insights into the chemical-microbe network that drive ocean ecosystem structure and function