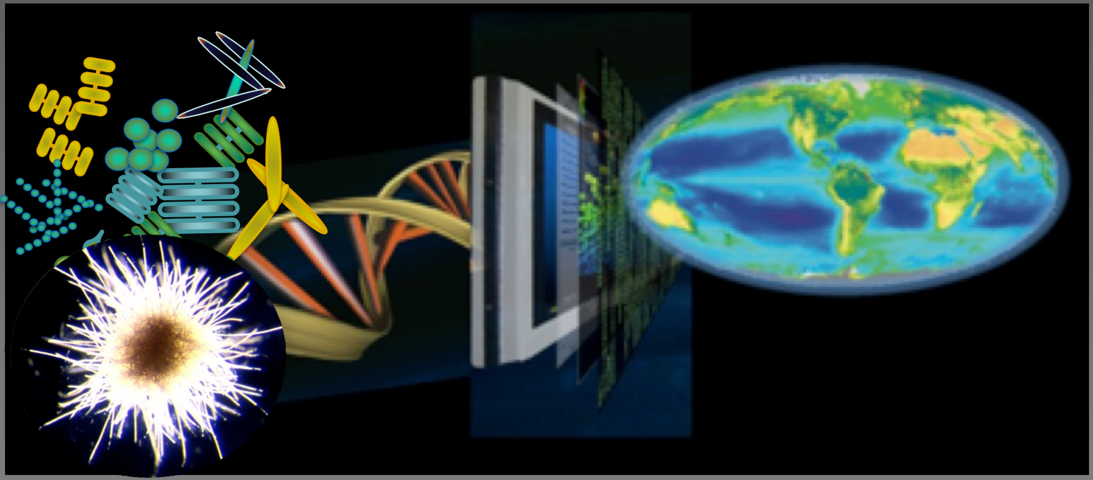
Ocean genomics: microbial physiological ecology across marine ecosystems

Sonya Dyhrman
Professor, Department of Earth and Environmental Sciences
Senior Staff, Lamont-Doherty Earth Observatory
Columbia University



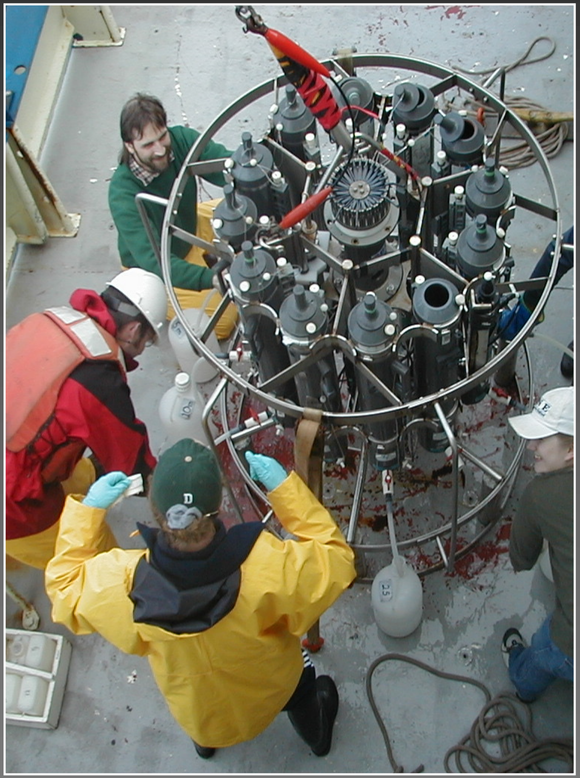
COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

LAMONT-DOHERTY
EARTH OBSERVATORY

Sonya Dyhrman

High School

Grad school/Postdoc



Professor

I have my dream job

Sonya Dyhrman

Krumlov - 2016

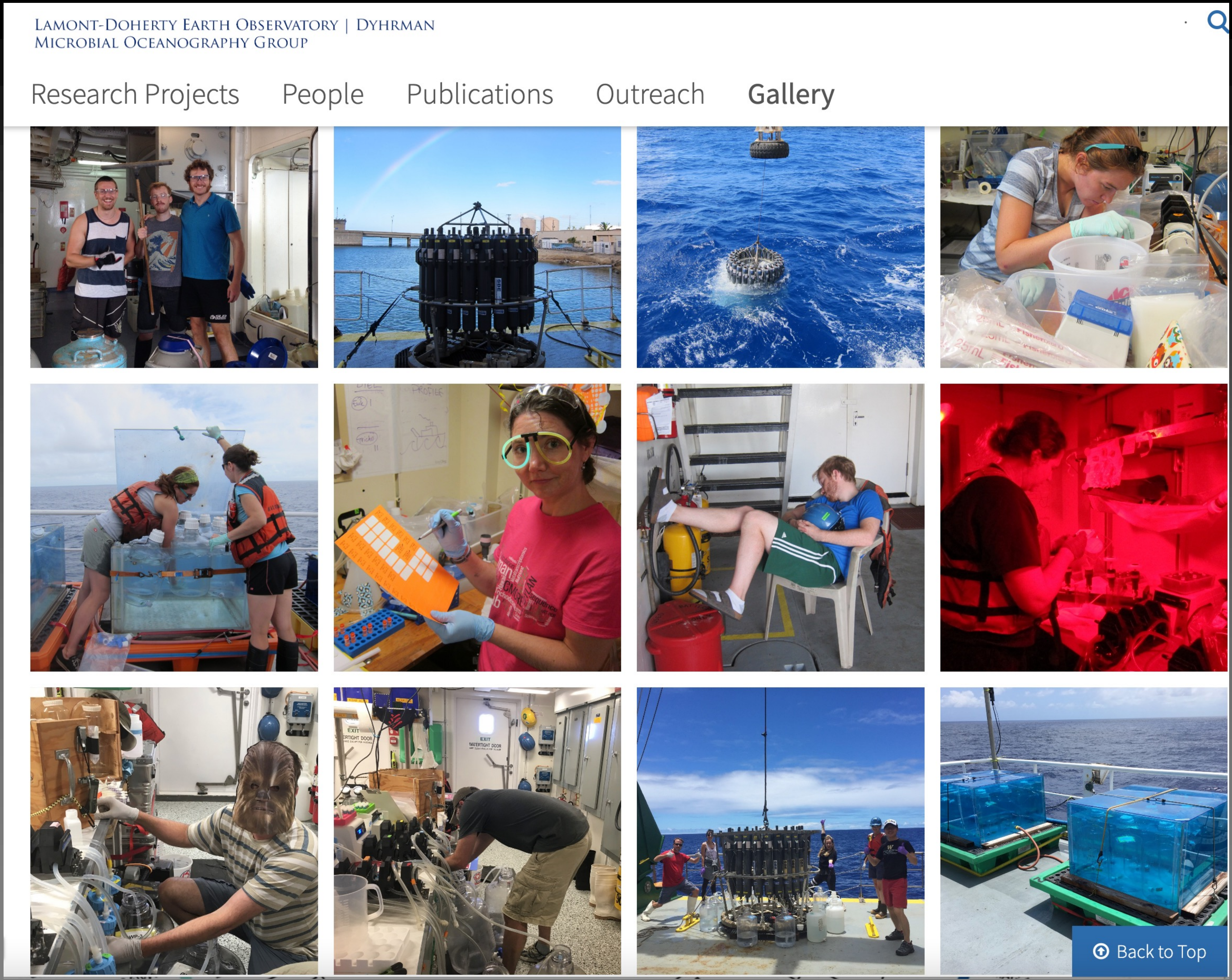
California
PhD Scripps

Massachusetts
Postdoc WHOI

Scientific Staff WHOI

New York: Faculty Columbia University

Columbia runs RV Langseth


Lamont-Doherty Earth Observatory | Dyhrman Microbial Oceanography Group
Research Projects
People
Publications
Outreach
Gallery
Back to Top

Sonya – random (fun?) facts

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

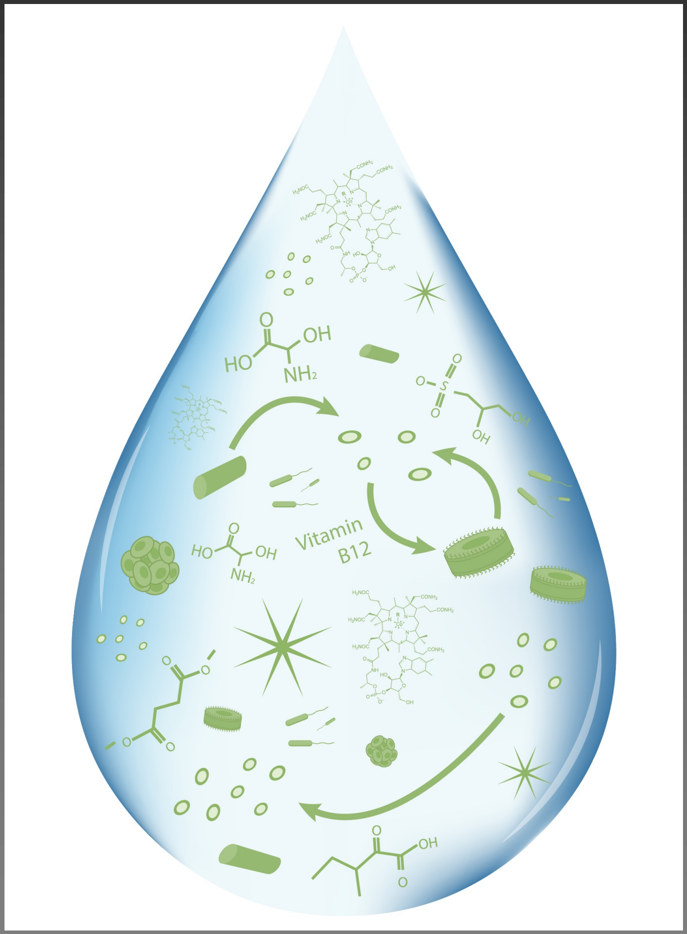
November 16th, 2022

Key points for today

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

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

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



Graphic: WHOI

The ocean makes our planet habitable

The vast unseen microbial populations play a critical role in ocean function

Marine Microbes - fundamental to ocean ecosystem function

- Marine microbes…
  - Produce and consume green house gases
  - Supply the marine food web
  - Recycle organic matter
- *make the planet habitable*

phyto- = photosynthetic
-plankton = floating organisms

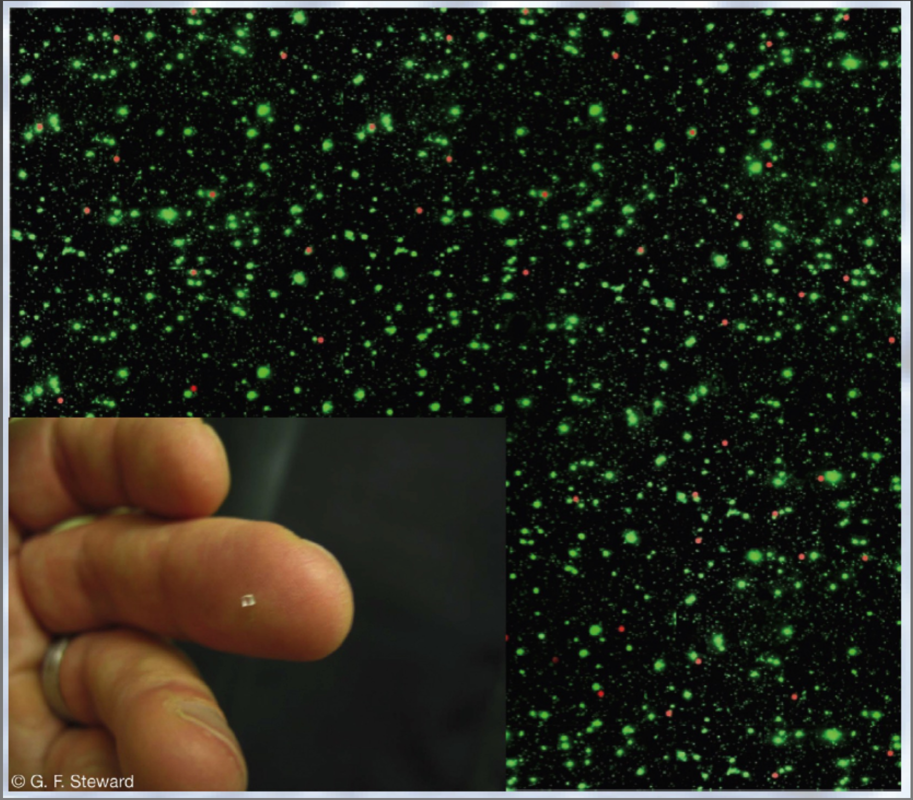
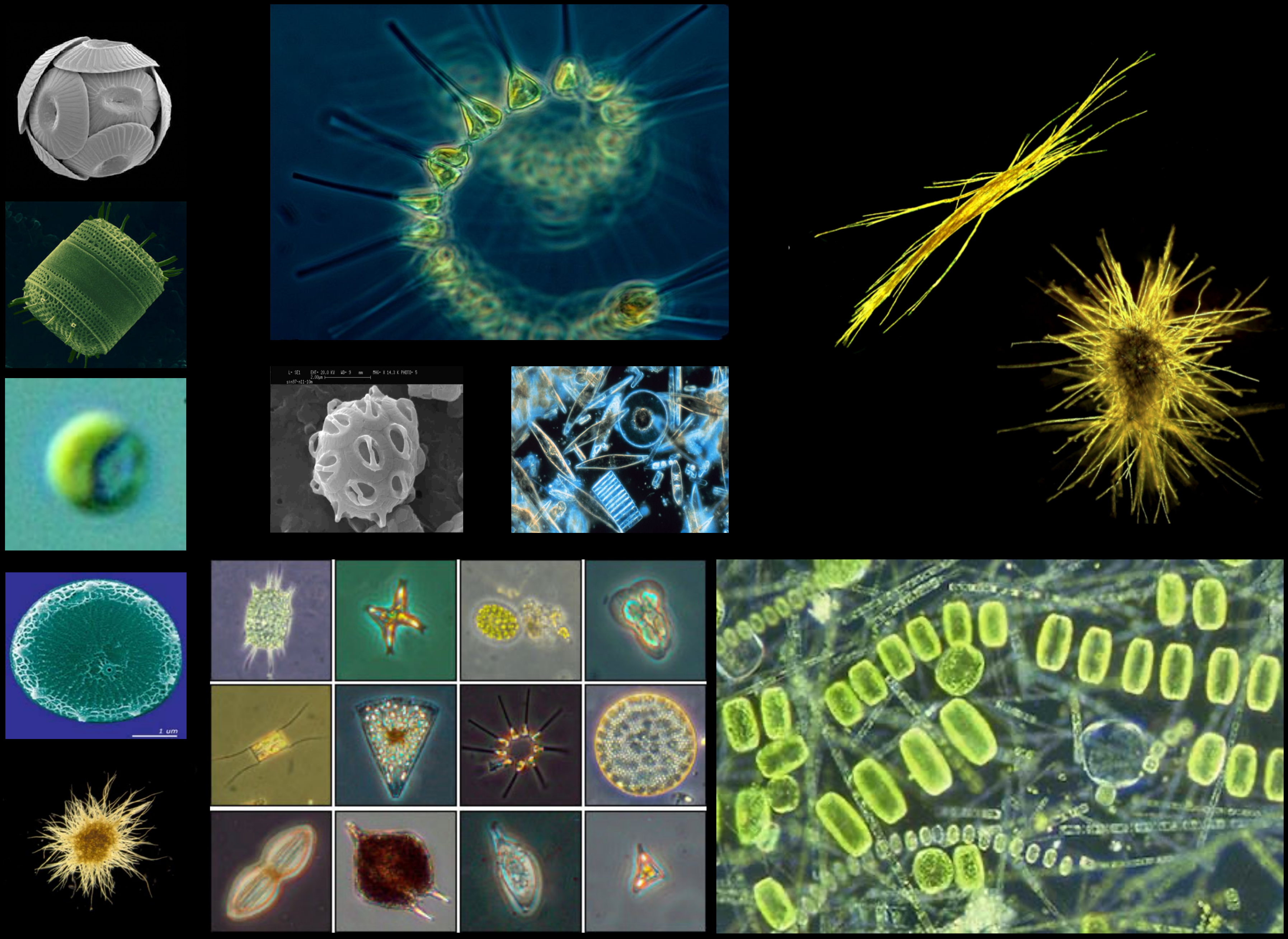
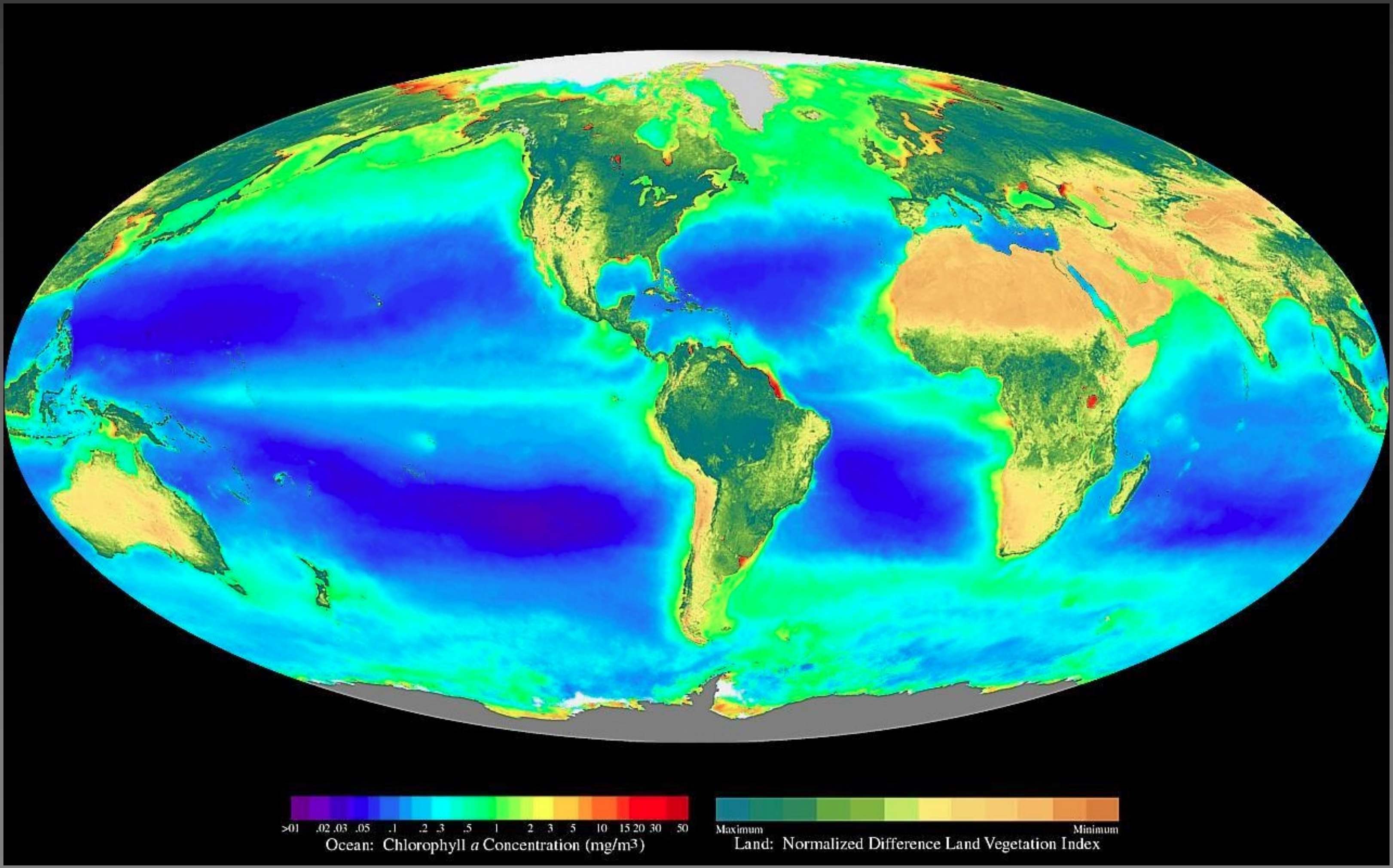


Image courtesy C-MORE

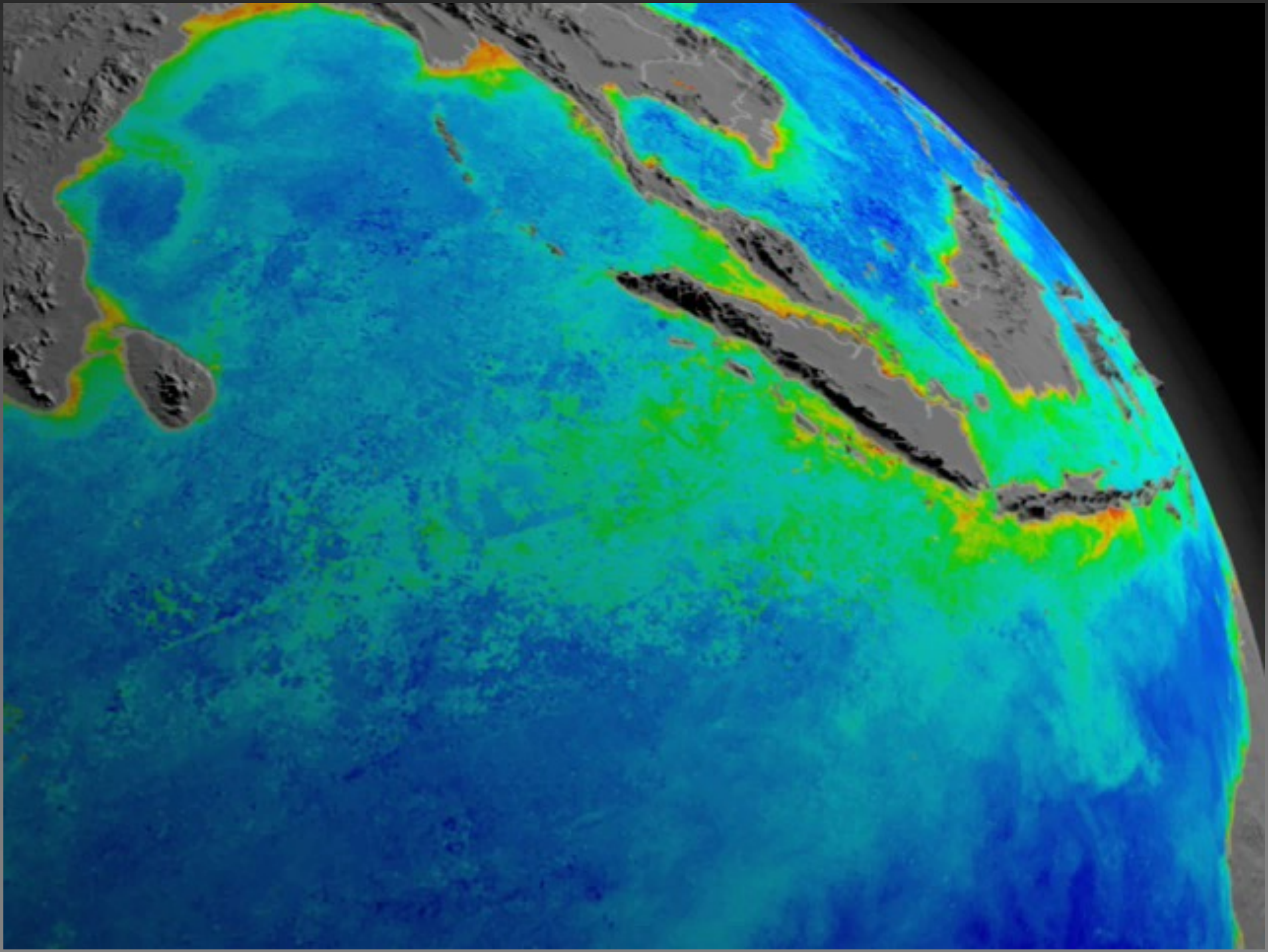

L= SE1 EHT= 20.0 KV WD= 9 mm MAG= X 14.3 K PHOTO= 5
2.00µm
sin97-n11-10m
1 um

Phytoplankton play a profound role in the earth system

Phytoplankton drive half of global primary production

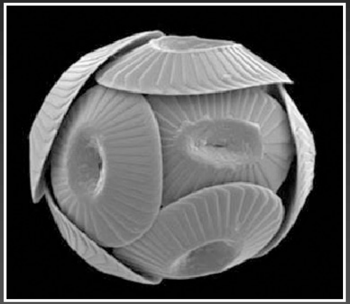


Seasonal chlorophyll distributions in the sea - highlights the global significance of phytoplankton



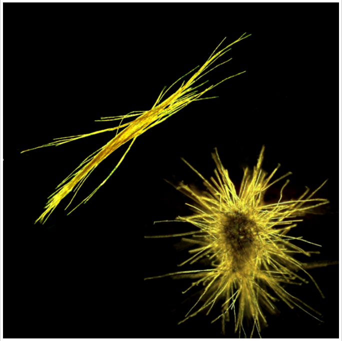
Phytoplankton underpin ocean ecosystem function

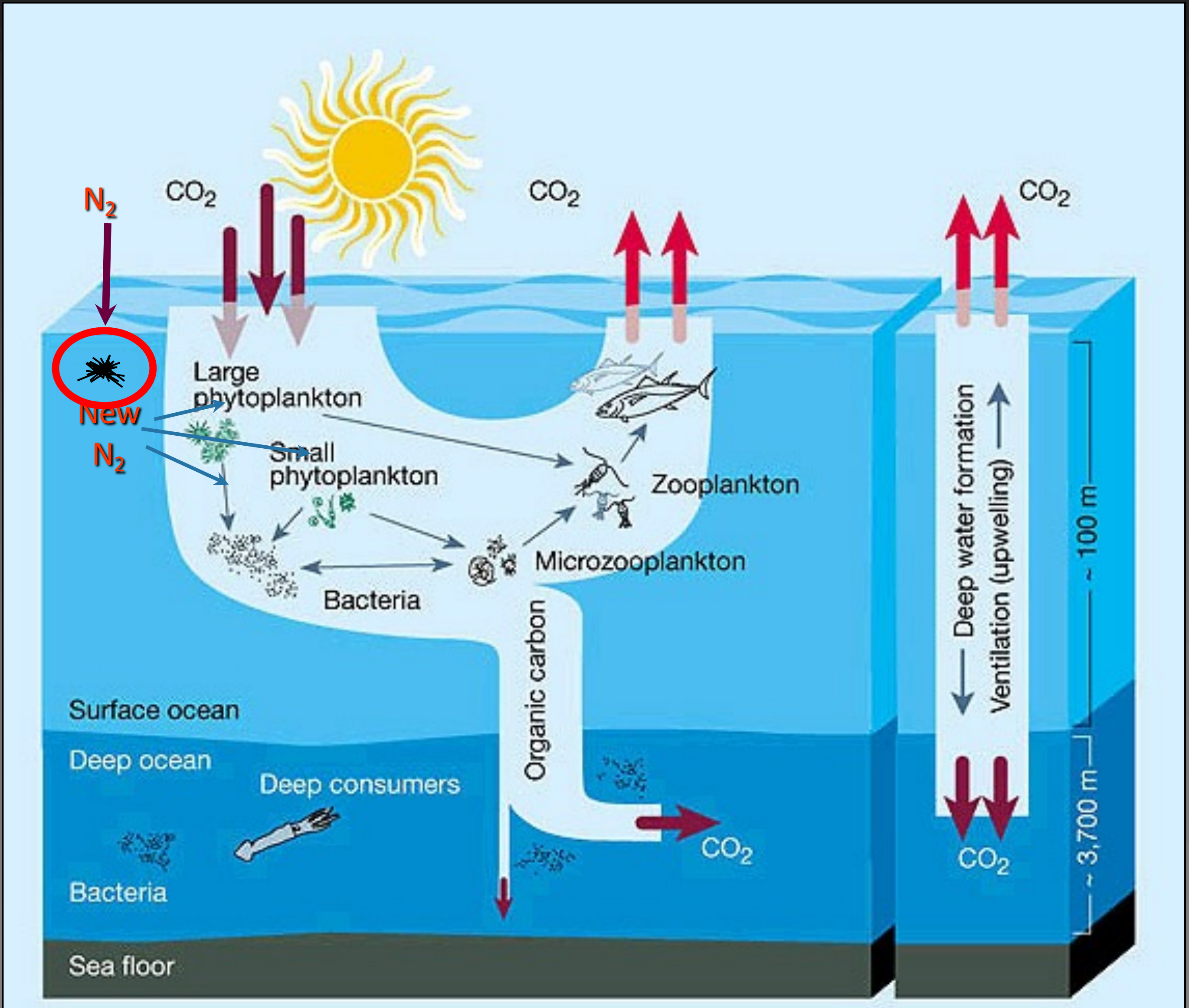
Haptophytes



Diatoms

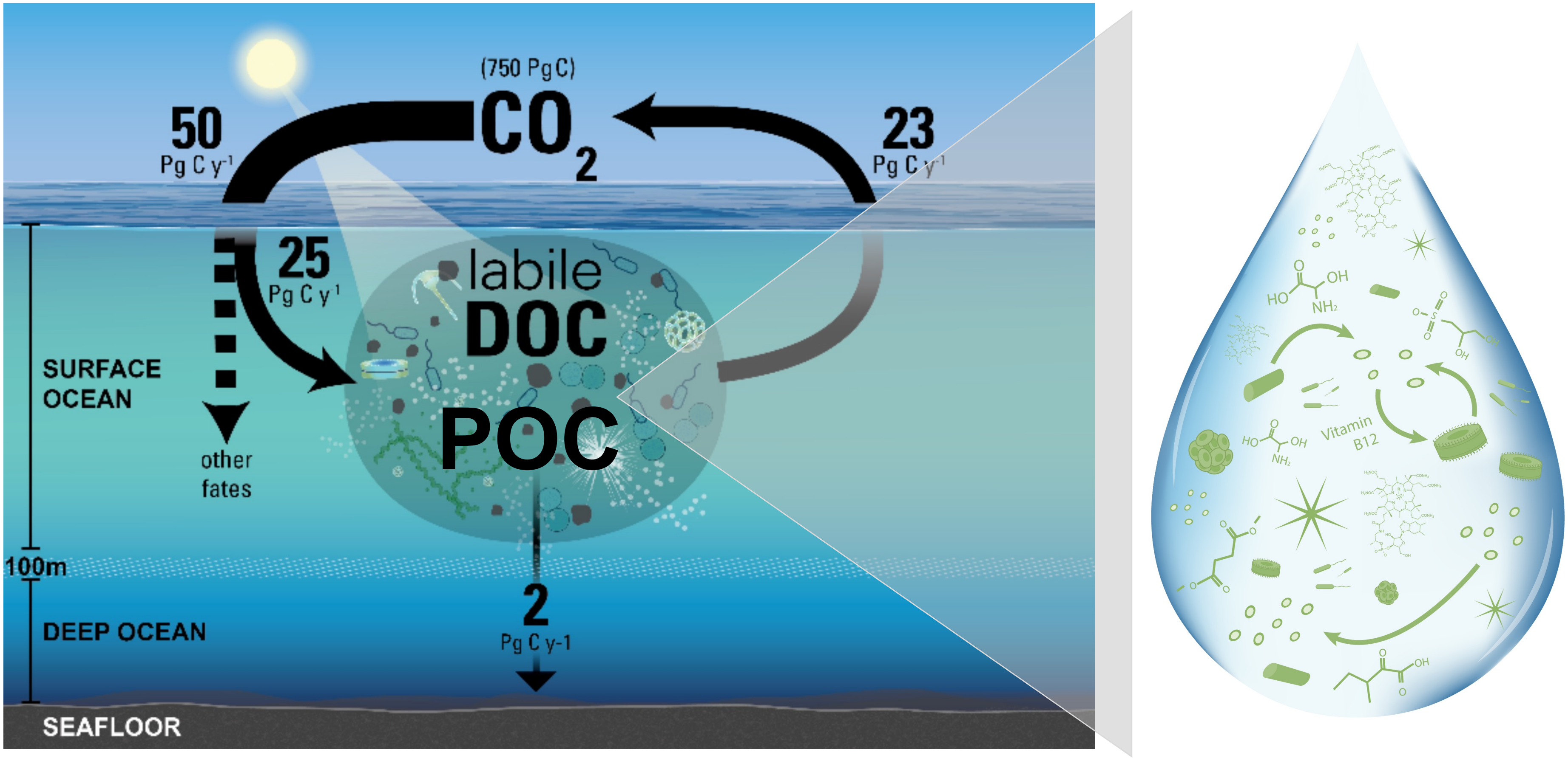
N_2 Fixers



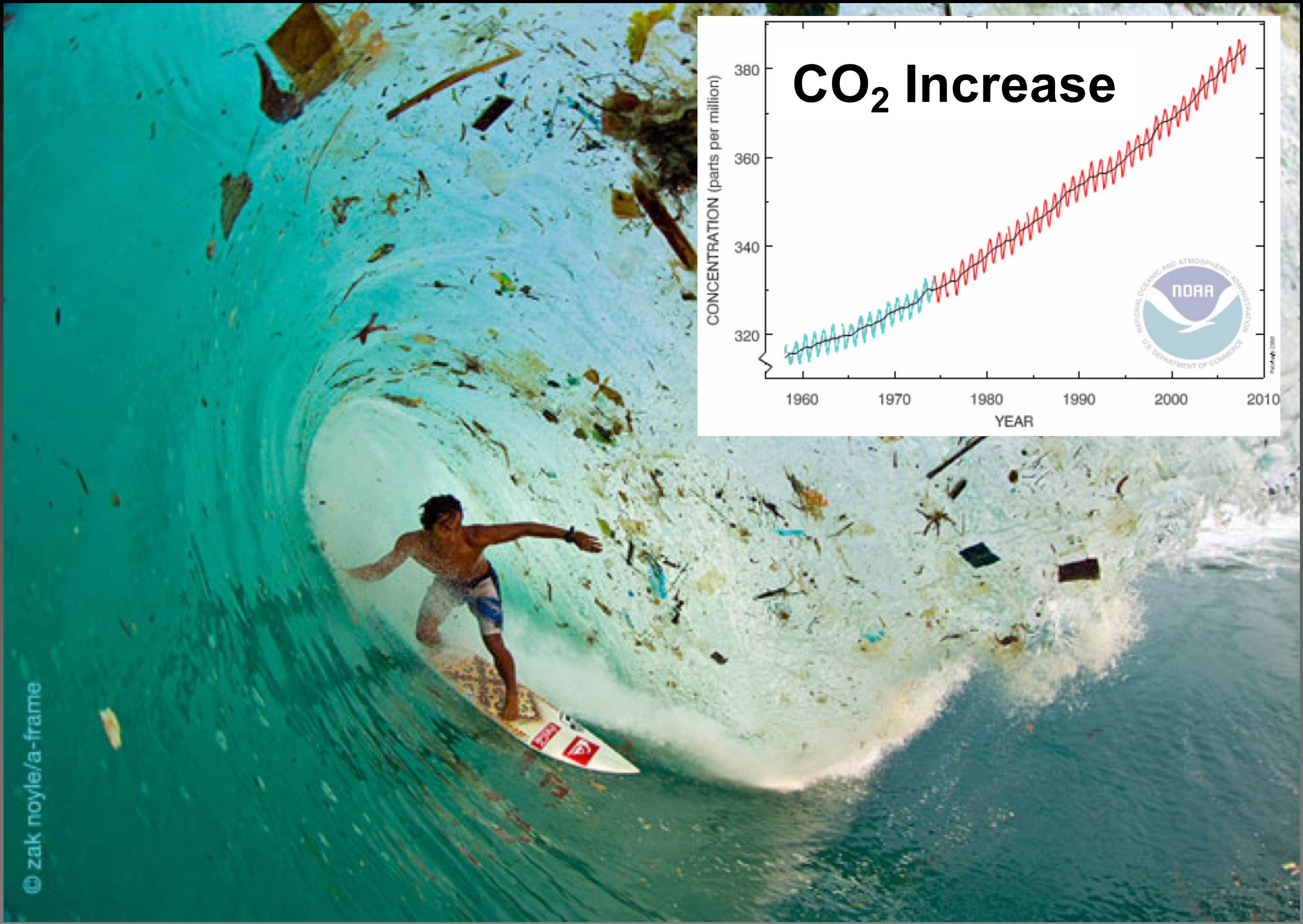


Chisholm 2000 *Nature*

The chemical – microbe network drives the fate of *a lot* of carbon



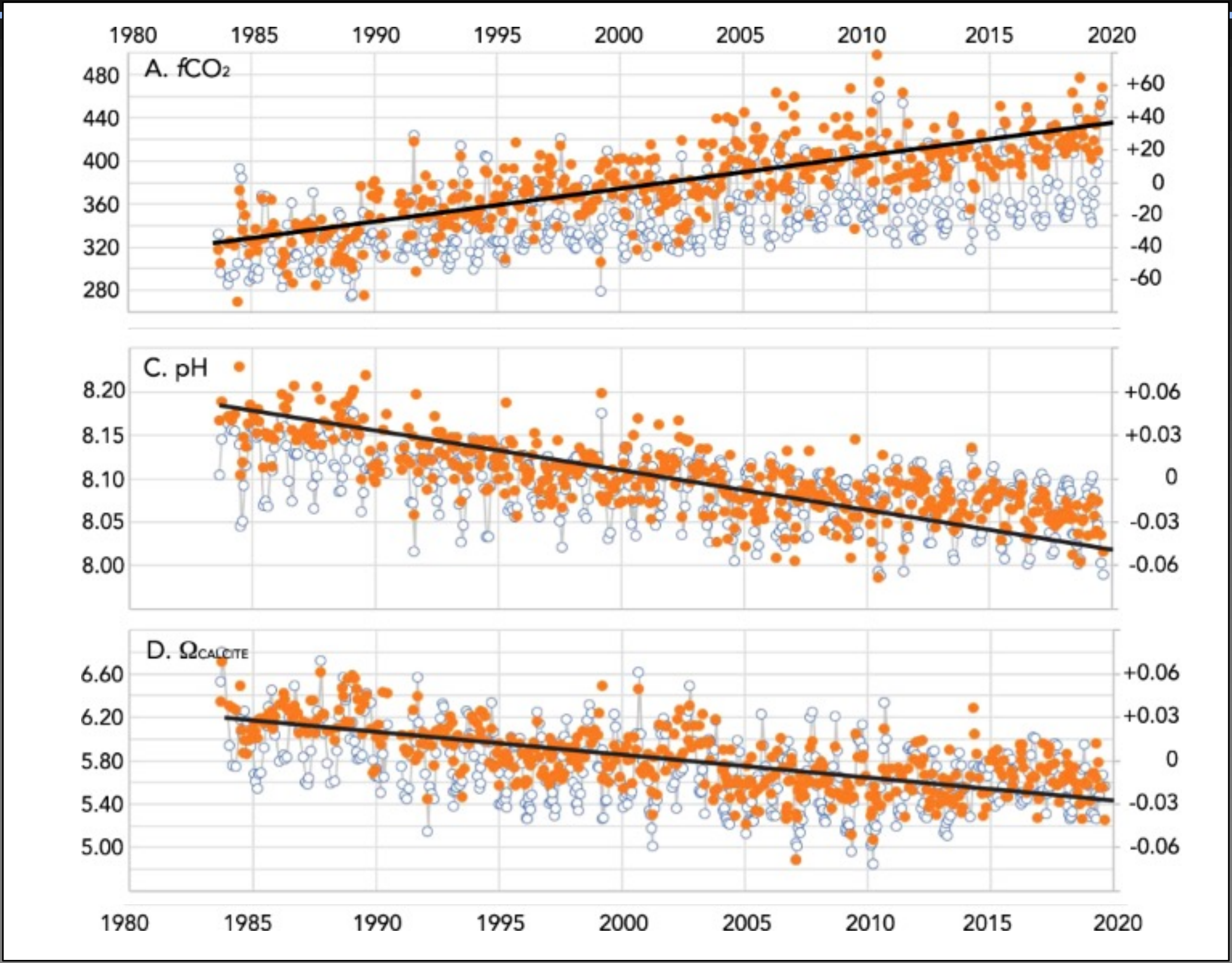
Graphics: WHOI



© zak noyle/a-frame

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

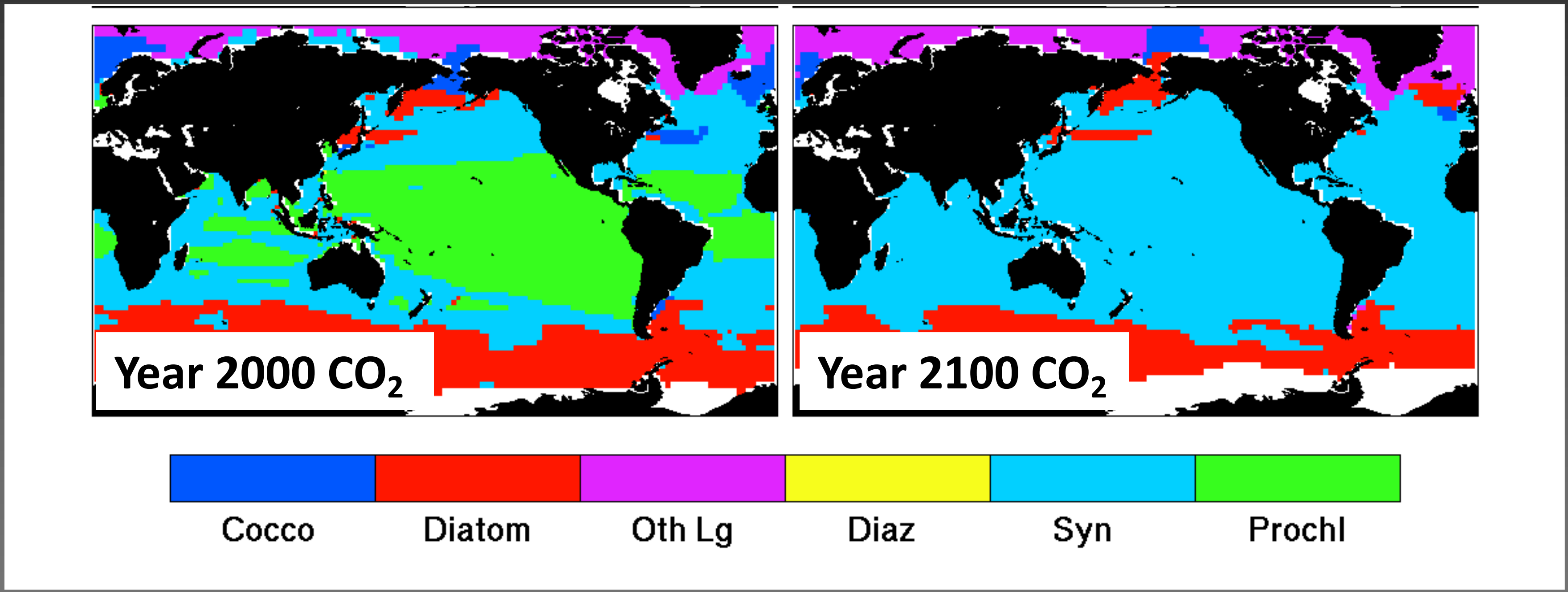
Changes in temperature CO_2, pH, etc.



Bates and Johnson 2020 *Communications Earth Env.*

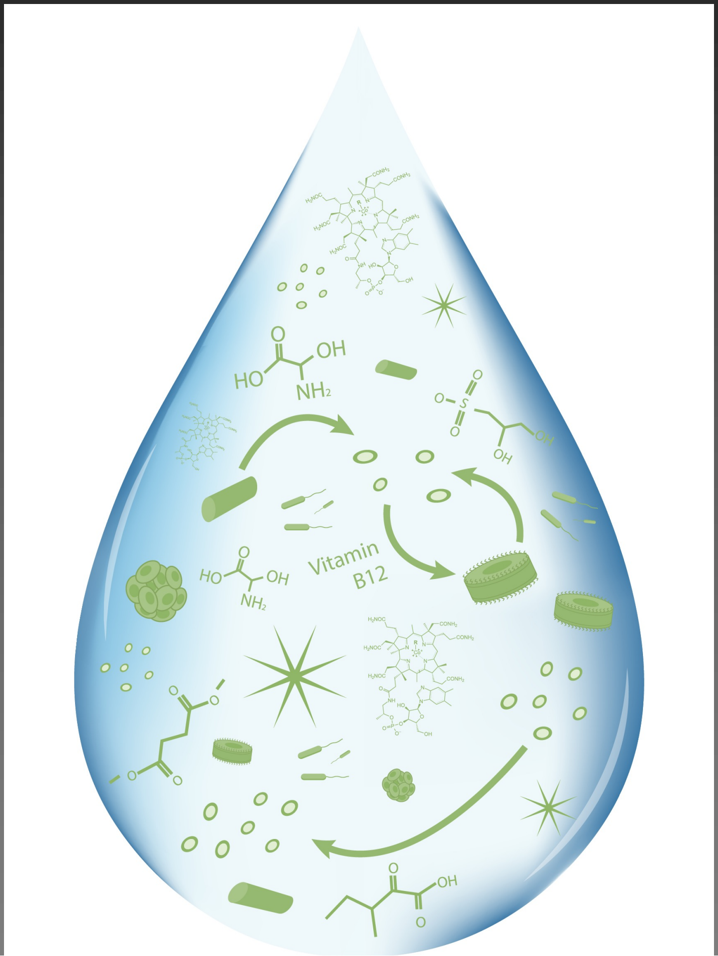
Models predict future shifts in microbial communities

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 to query cells in their environment in a species-specific way

Graphic: WHOI

Challenges and opportunities

- Long standing challenges:
  - Populations are dilute and they move around!
  - Salt!
  - Some keystone groups not in culture
  - Few genomes or transcriptome sequences (particularly for the eukaryotes)
- New opportunities:
  - Increases in genomes/transcriptomes for eukaryotic taxa
  - *New array of bioinformatics and sequencing tools*

A decade of meta ‘omic and multi ‘omic approaches!

100 Diatom Genomes

Diatoms are the most species-rich algal class with estimated ≥ 100,000 species. They contribute to 20% of global carbon fixation and oxygen production. However, with genomic information from only about 10 diatoms, our ability to harness their unique biology is very limited. Consequently, this project will sequence 100 diatom species for providing unique insights into their roles as key players for capturing carbon dioxide and as the foundation of diverse aquatic food webs. These insights will also be critical for advancing diatom-based biotechnology and synthetic biology platforms.

Proposer: Thomas Mock, University of East Anglia (UK)
Proposal: 100 Diatom Genomes Project

Scanning electron micrographs of diverse diatoms. (Credits: Diana Sarno, Marina Montresor, Nicole Poulsen, Gerhard Dieckmann)

Challenges and opportunities

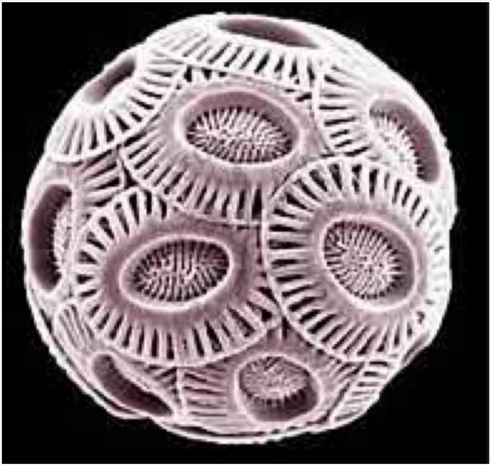
- Long standing challenges:
  - Populations are dilute and they move around!
  - Some keystone groups not in culture
  - Few genomes or transcriptome sequences for eukaryotic taxa

- New opportunities
  - Increases in genomes/transcriptomes for eukaryotic taxa
  - *New array of bioinformatics and sequencing tools*

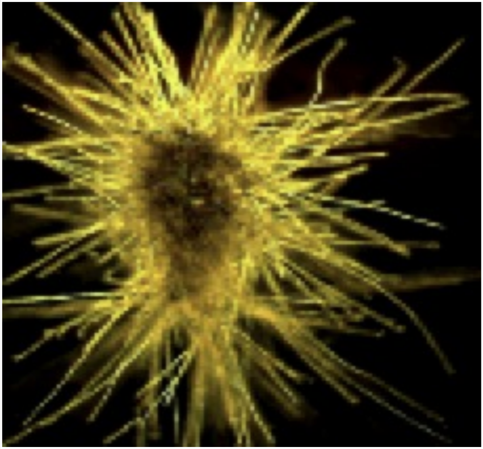
A decade of meta ‘omic and multi ‘omic approaches!

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

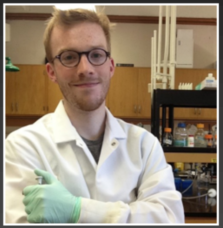


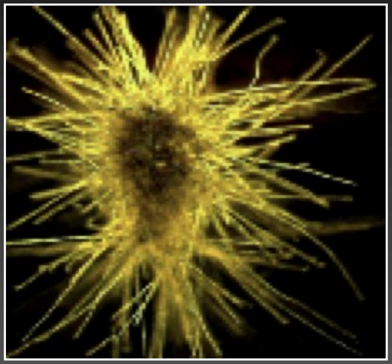
Thank you Dyhrman Lab family past and present


Gwenn Hennon


Mónica Rouco


Sheean Haley


Kyle Frischkorn




Sheean Haley


Hanna Anderson


Angela Zoumplis

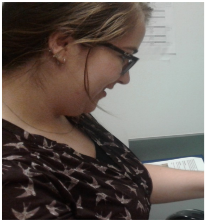

Erin Maybach


Hannah Goldberg

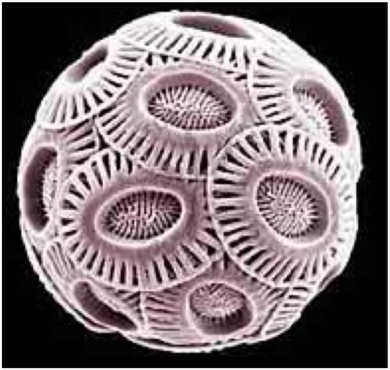

Maxfield Palmer


Sheean Haley


Matt Harke


Maria Hernández-Limón


Harriet Alexander



Acknowledgements

- Dyhrman Lab and many extraordinary collaborators

- NSF
- DOE
- Simons Foundation
- World Surf League
- Angell Family Foundation

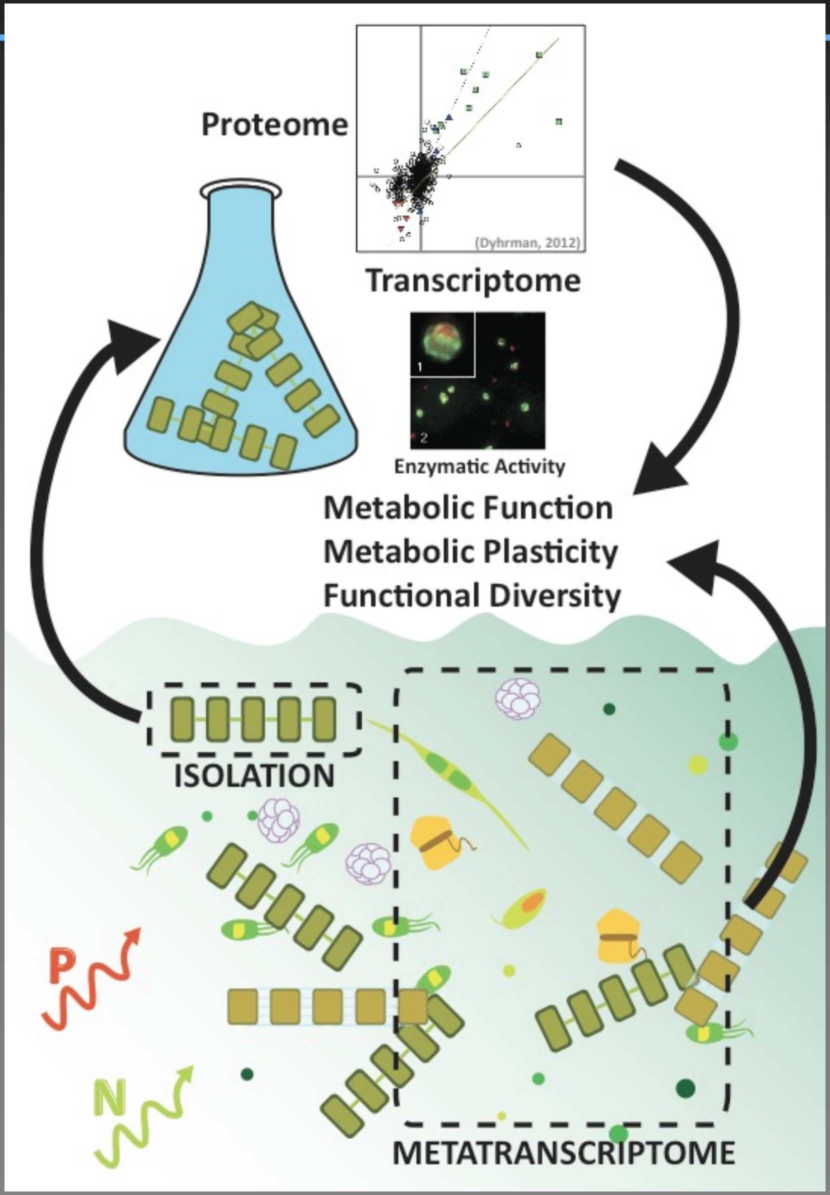
Leveraging metatranscriptome data to study the chemical – microbe network

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

- Field samples offer context

- Metatranscriptomes are species-specific

- Metatranscriptomes offer high resolution tracking of finely tuned responses



Graphic: H. Alexander

Microeukaryote metratranscriptomes with RNA-Seq

3x20L

Filter
(>5µm)

Extract RNA
(and standards)

Selected

De novo assembly
Reference mapping

Unselected

rRNA
depletion

De novo assembly
Reference mapping

Quality Control
- Trimmomatic
- FastQC

Removal of Sequences
- riboPicker
- BBMap

Functional Annotation
- BLAST
- DIAMOND
- HMMer

- KEGG
- KOG
- GO
- PFam

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

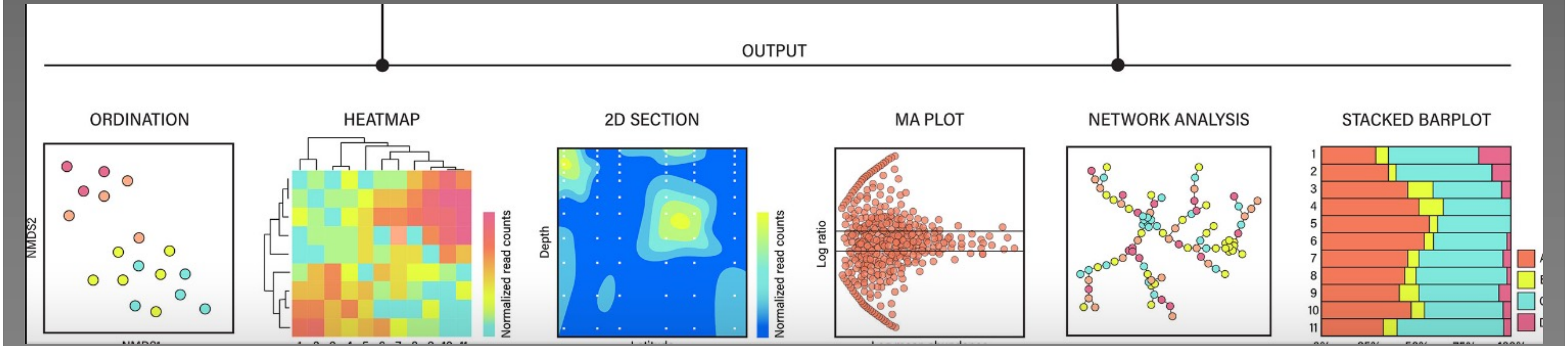
Protein Prediction and Clustering
- TransDecoder
- OrthoFinder

Read Mapping
- Salmon
- Kallisto
- Bowtie2
- RSEM

Normalization & DE
- EdgeR
- DESeq2
- TPM
- Variance stablization

Taxonomic Annotation
- BLAST
- DIAMOND
- MMSeq2

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



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/

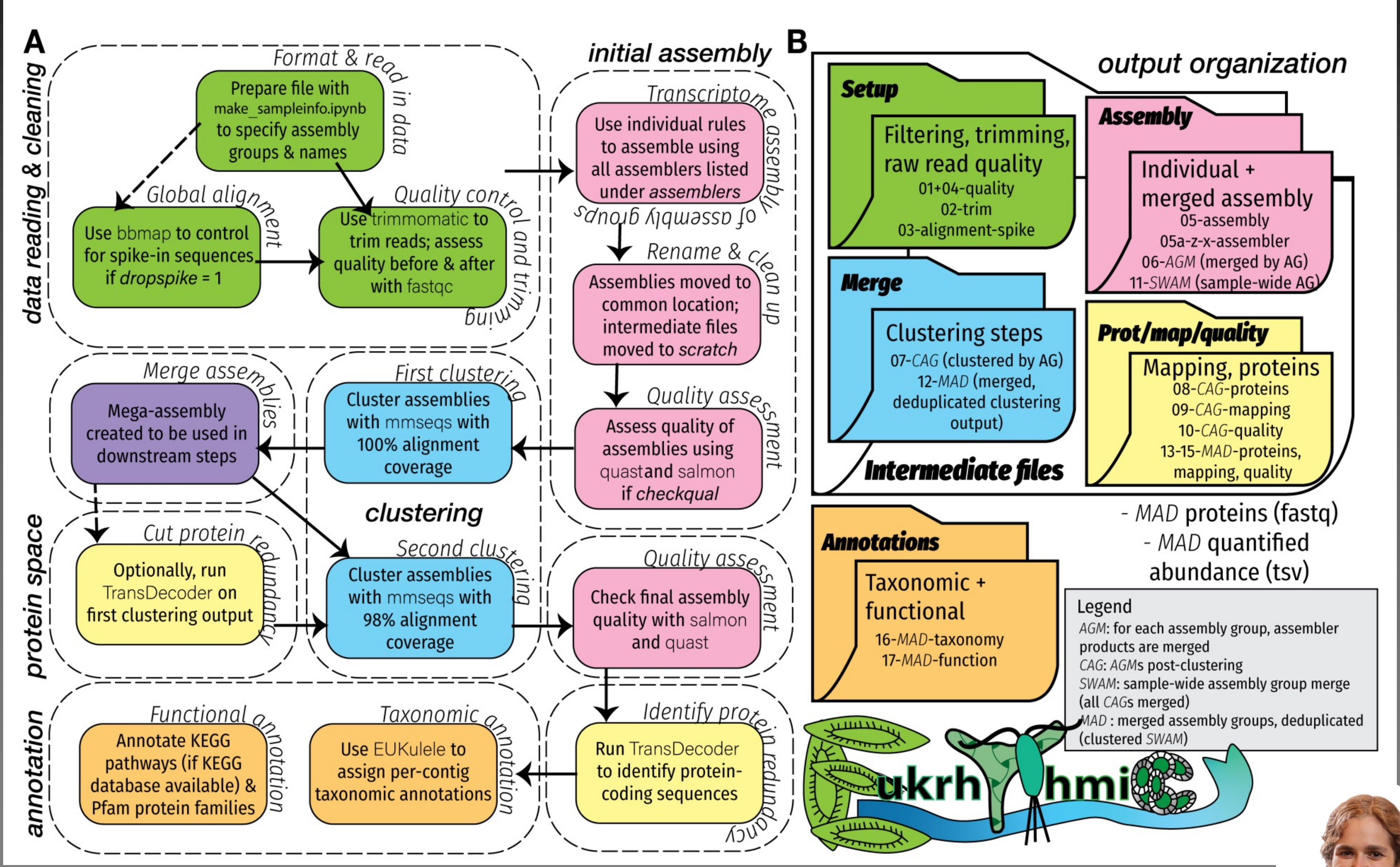
Feedback / comments welcome!

github.com

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

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

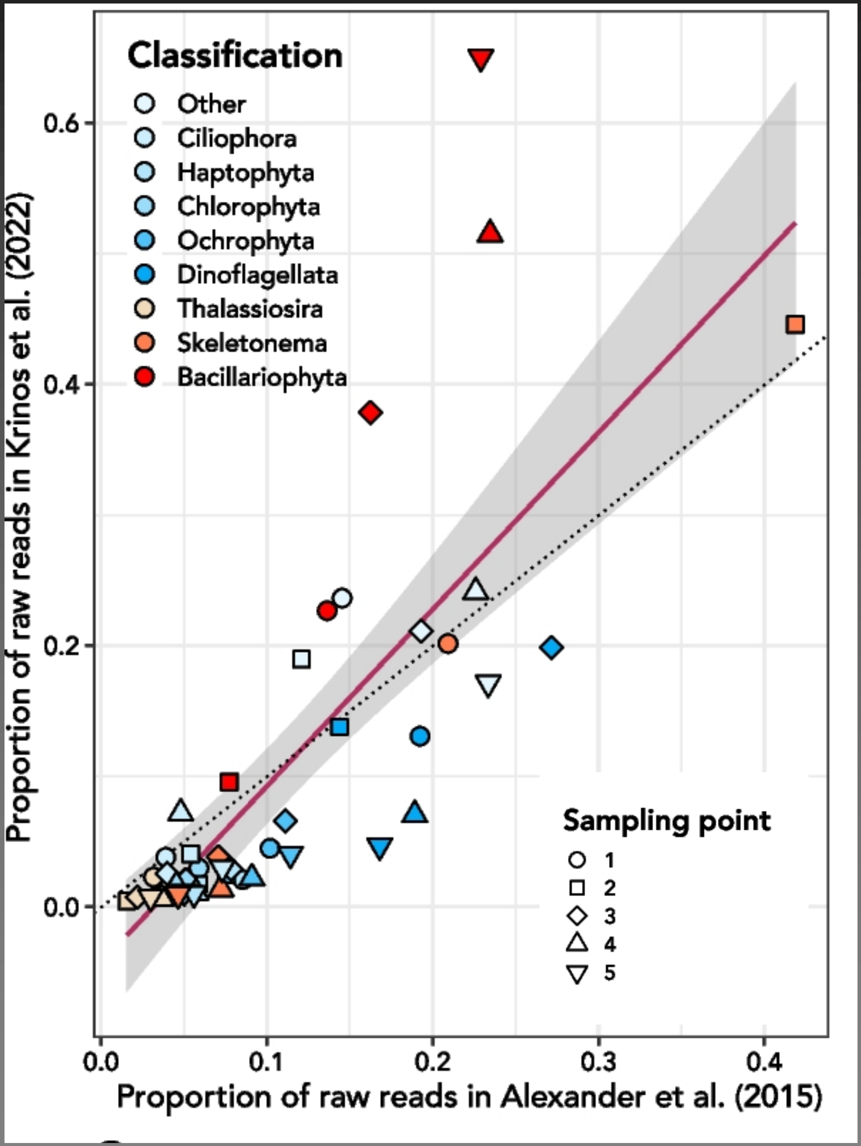
Sequencing and analytical workflows are constantly improving



Krikinos et al. 2023 *BMC Bioinformatics*

Metatranscriptome workflow comparison effort

- Assembly captures a swath of reads in high abundance – highly diverse groups (red)
- Reference mapping captures a swath of reads in low abundance groups (blue)
- Generally equally good at recapitulating patterns when references are from local isolates (orange and yellow)



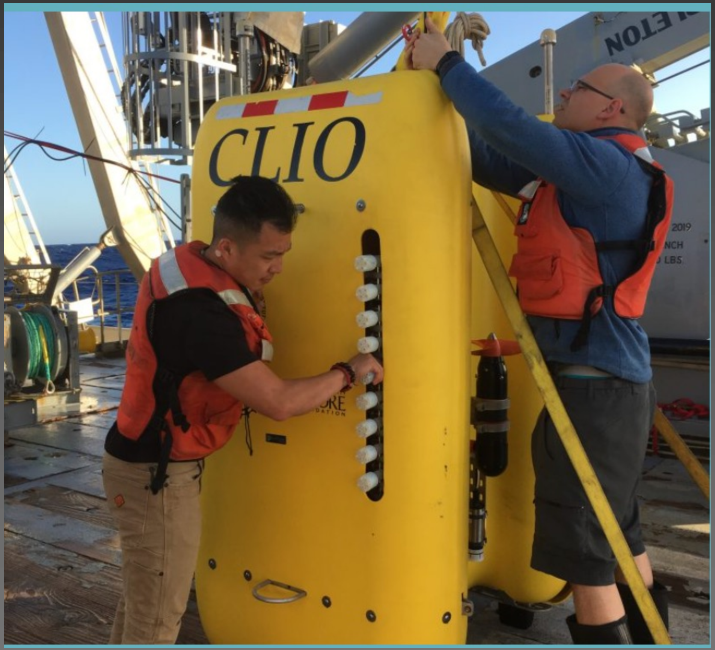
De novo Assembly

Reference mapping

Krikinos et al. 2023 *BMC Bioinformatics*

Metatranscriptome intercalibration effort – phase 1

- Large diameter filter sampled with CLIO (June 2023)
- 2x filter pieces to ~ 7 groups
- Evaluate:
  - Extraction
  - Library prep
  - Workflow intercomparison



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

Thank you N. Cohen and H. Alexander
for leadership and M. Saito for samples

Core questions… adventures in eukaryote metatranscriptomics

Limitation

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

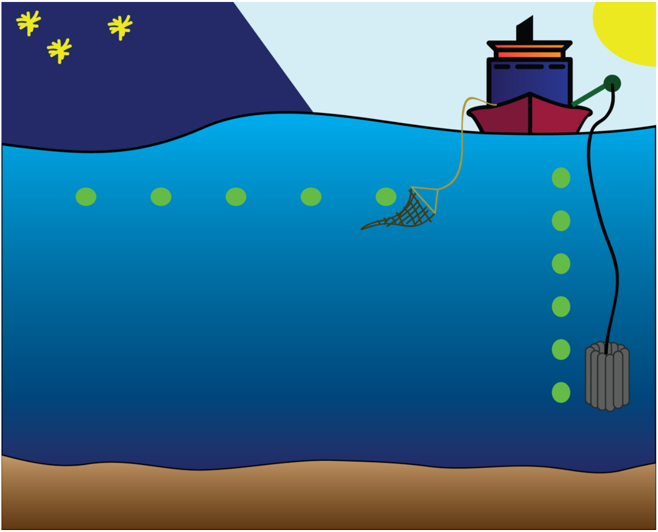
Competition

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

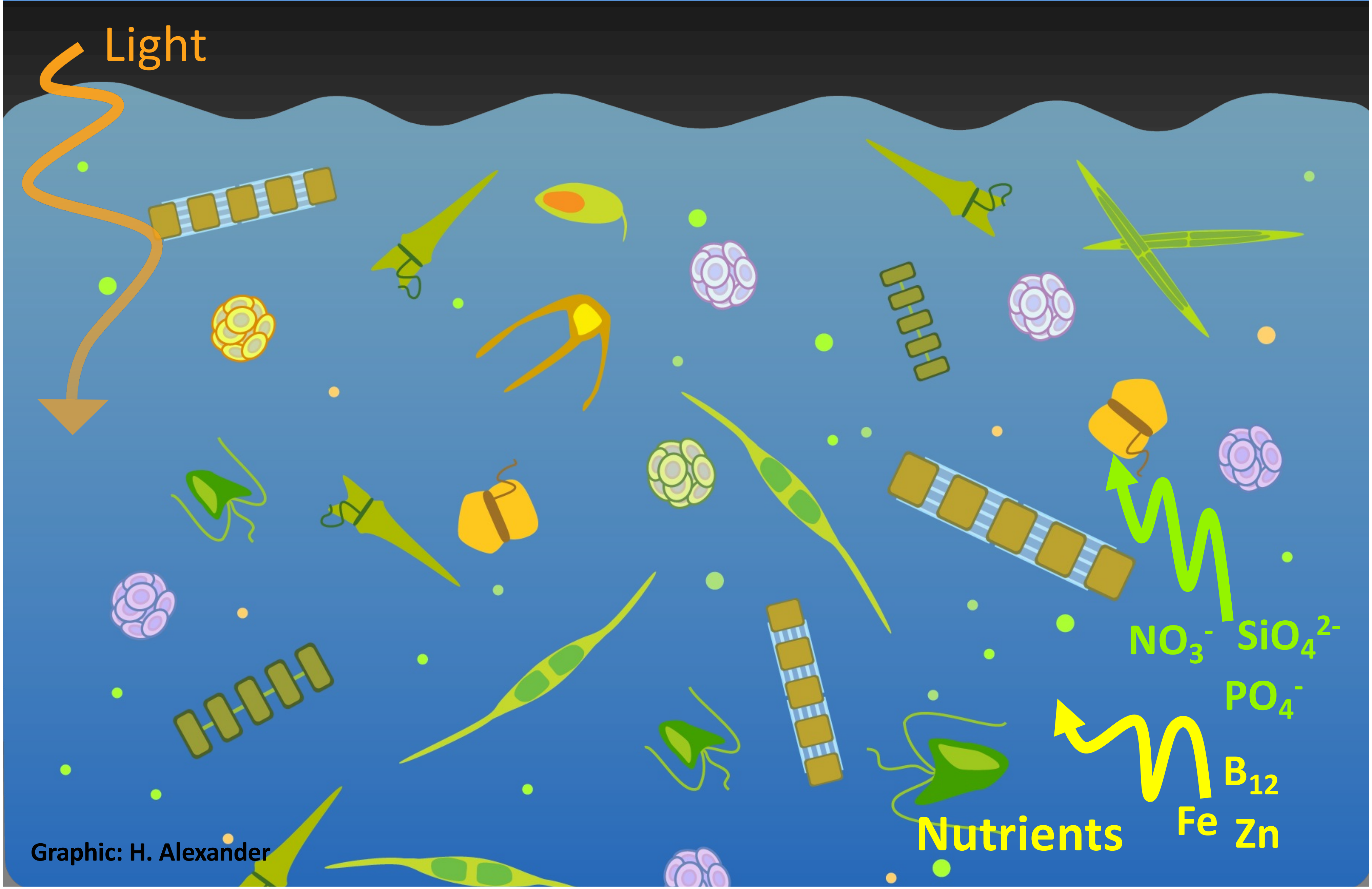
HOE:DYLAN

H. Alexander

HOE:MAHALO



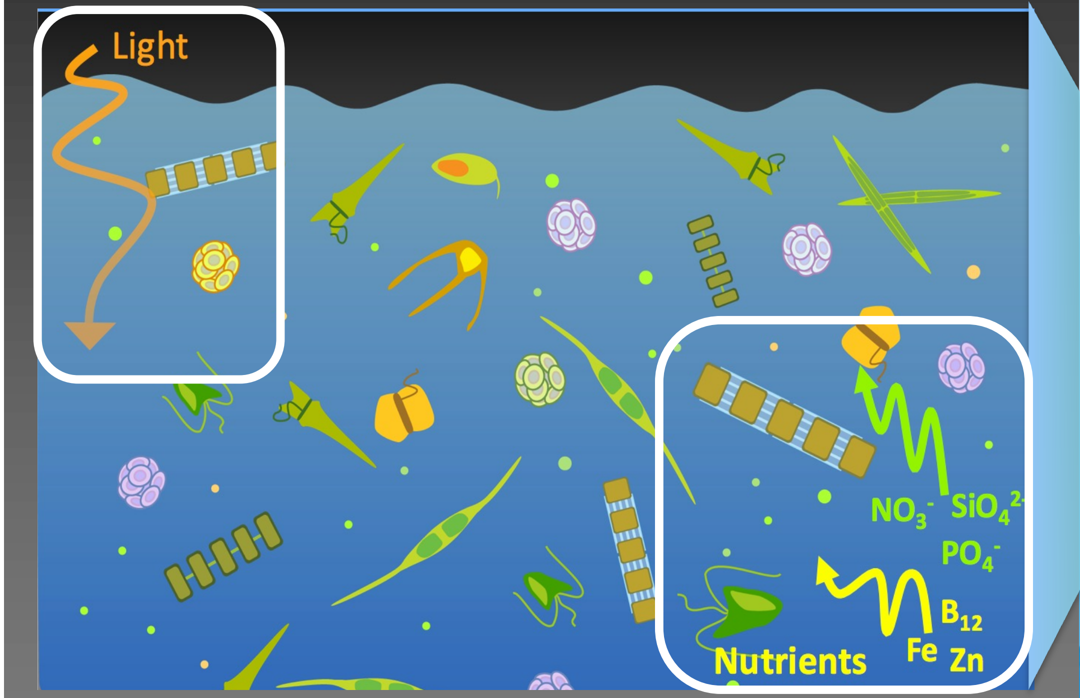
Complex community dynamics driven in part by resources



Graphic: H. Alexander

Tracking phytoplankton physiological ecology

Ecosystem function and biogeochemistry



Graphic: H. Alexander

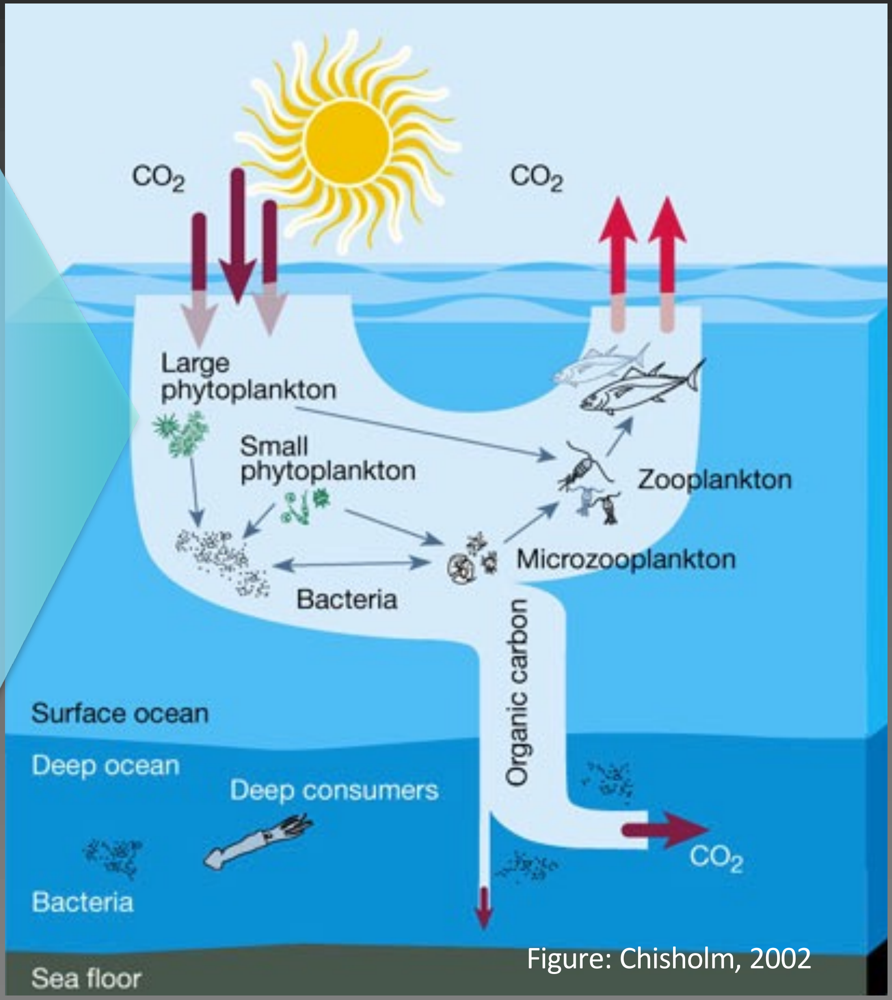
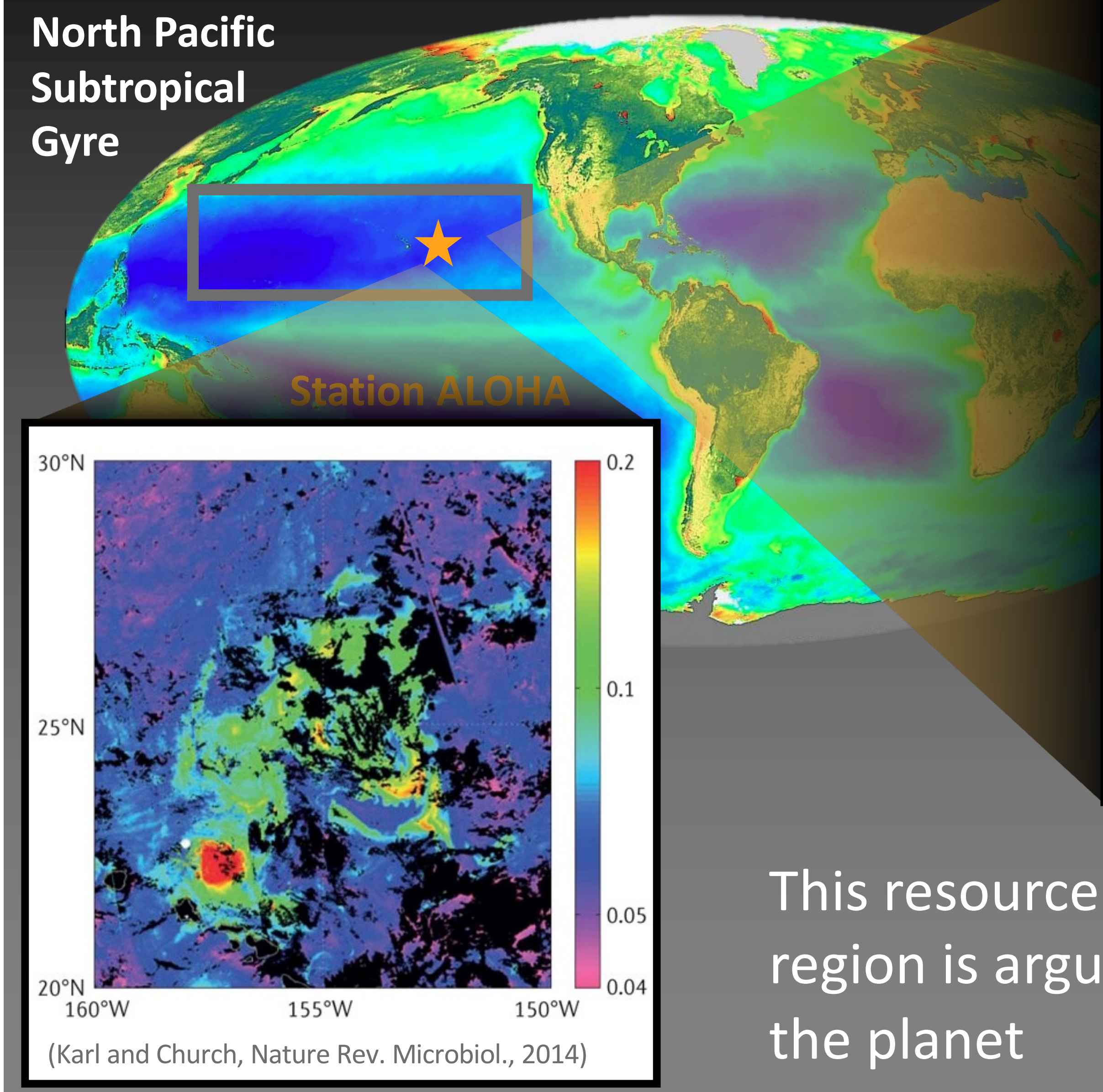


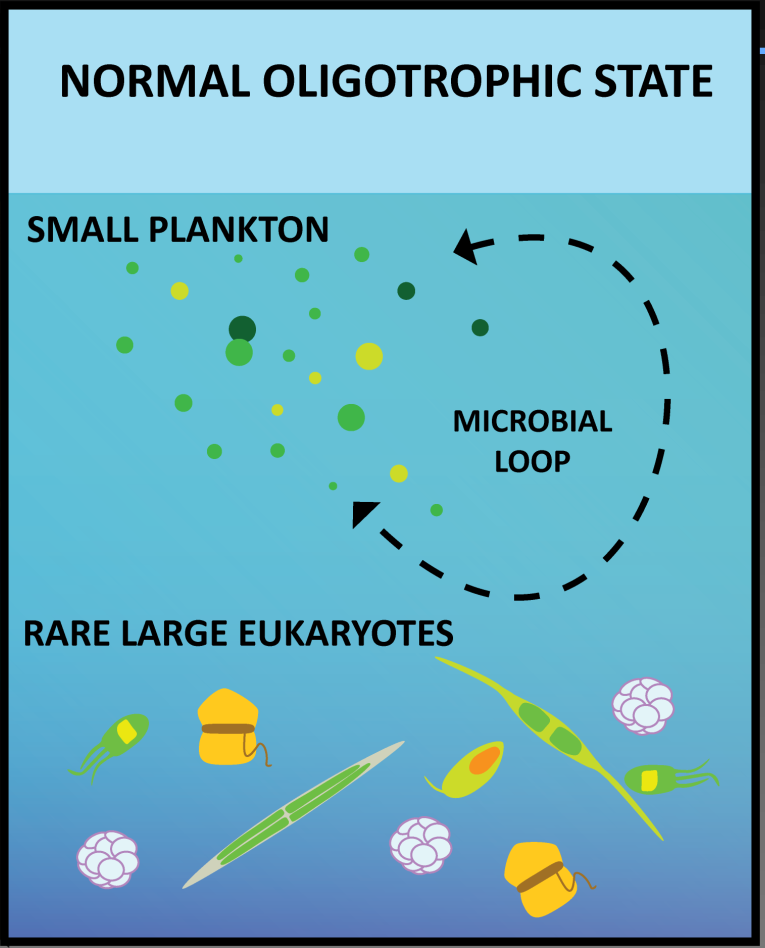
Figure: Chisholm, 2002

How do resources drive phytoplankton distributions and activities?

North Pacific Subtropical Gyre



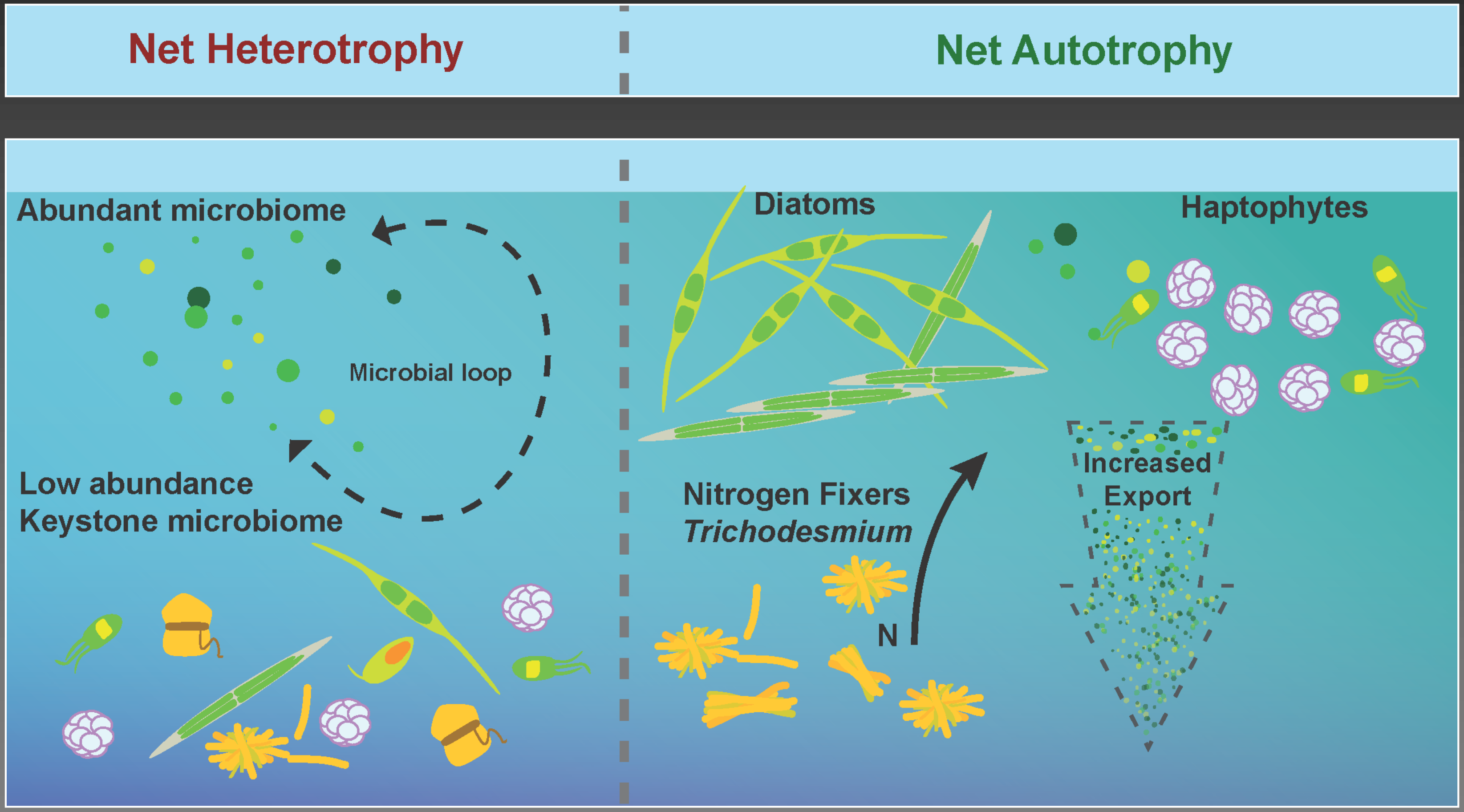
(Karl and Church, Nature Rev. Microbiol., 2014)



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

Image credits: SeaWIFS Global Chlorophyll, H. Alexander

Blooms of keystone species can shift ecosystem state and carbon flow



Net Heterotrophy

Net Autotrophy

Abundant microbiome

Microbial loop

Low abundance
Keystone microbiome

Diatoms

Haptophytes

Nitrogen Fixers
Trichodesmium

N

Increased
Export

Graphic: H. Alexander

Nutrient input, or other forcings

* *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 sequenced, identifying pan genome

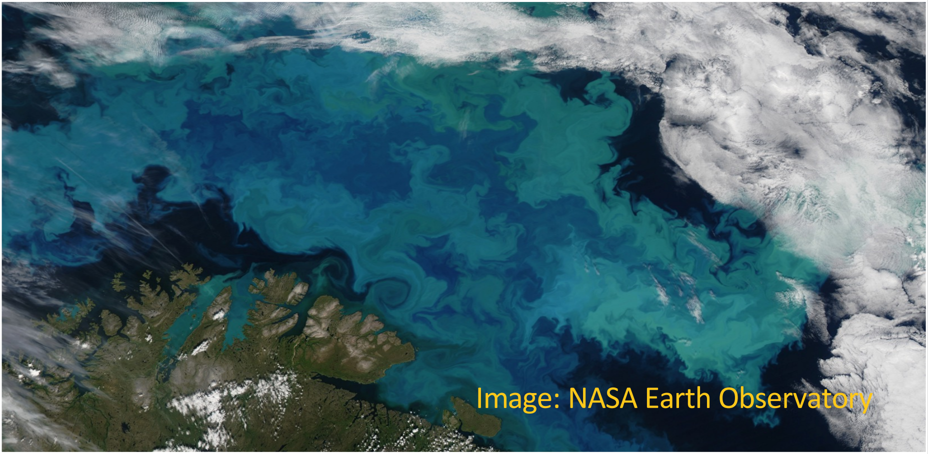
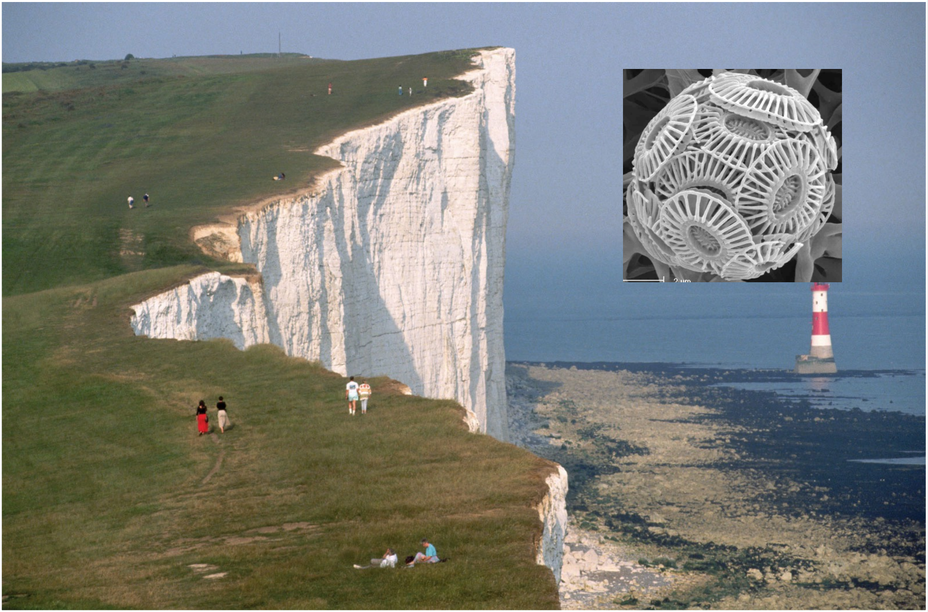
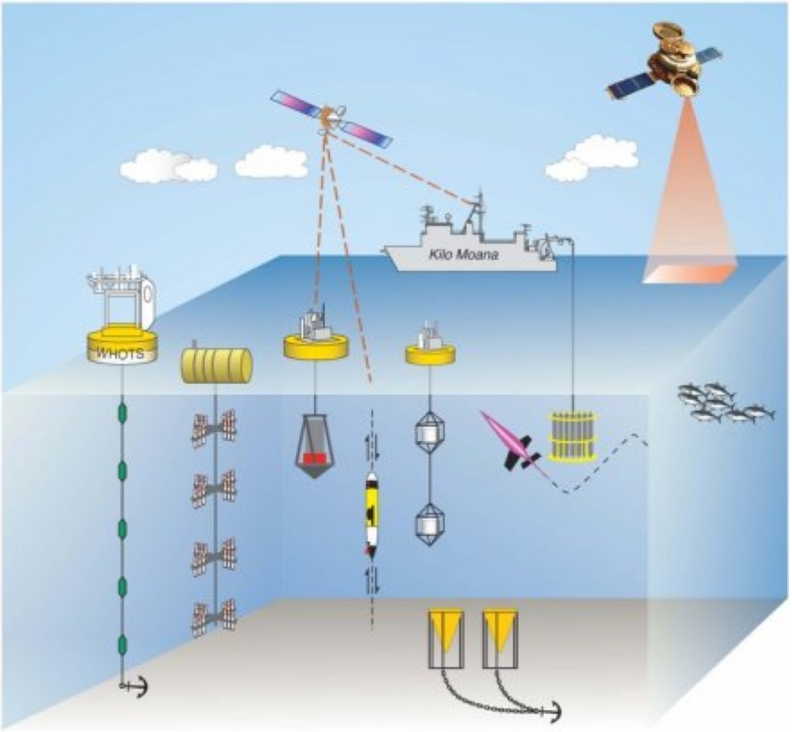


Image: NASA Earth Observatory



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

Hawaii Ocean Experiment: Dynamics of Light and Nutrients



Wilson et al. 2015 *GBC*

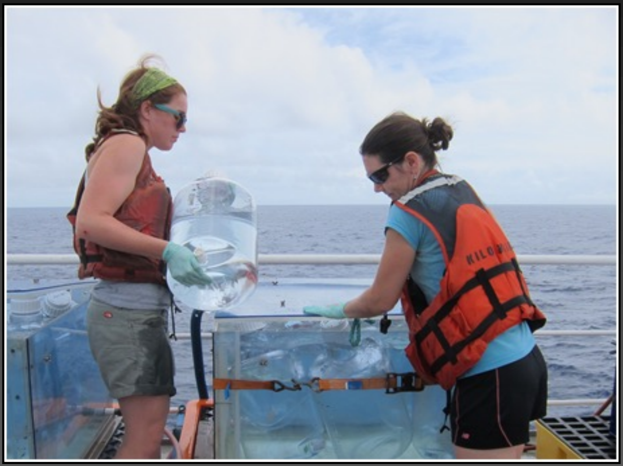
center for microbial oceanography: research and education

C·MORE *linking genomes to biomes*

(HOE-DYLAN)

HOE:DYLAN: Water sampling

HOE:DYLAN: Semi-factorial resource additions



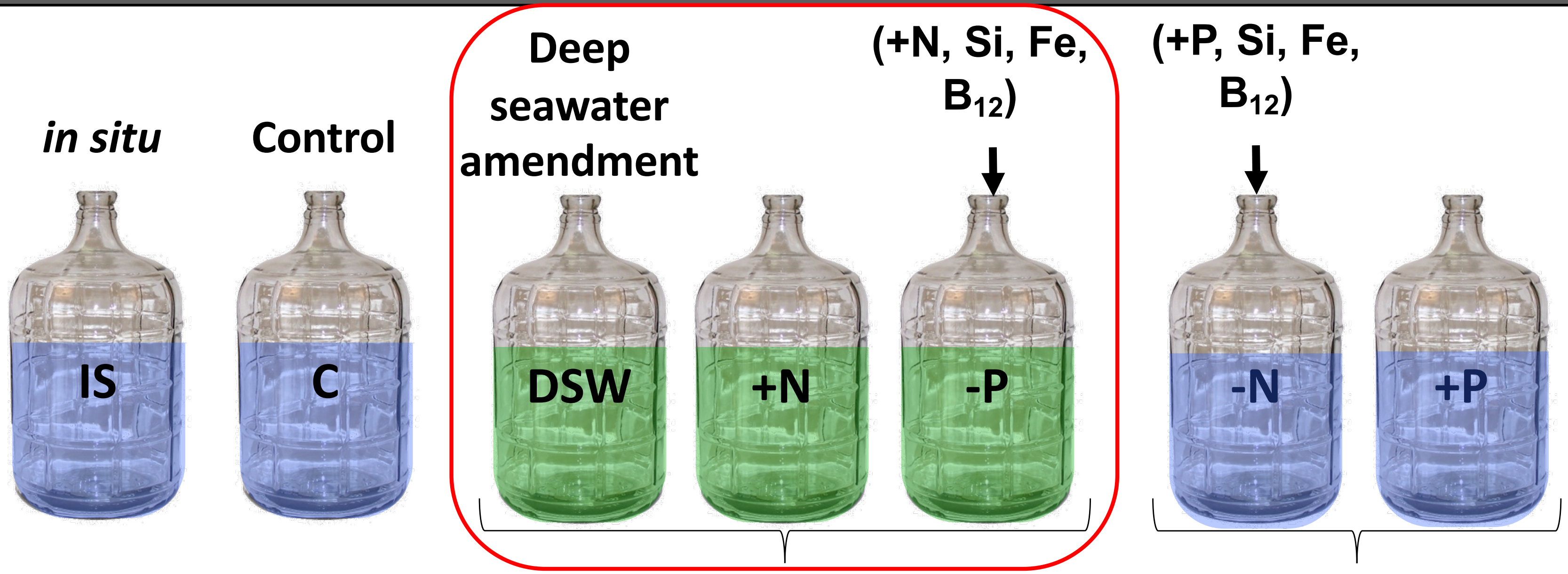
Nitrate/nitrite*: ~40 μM

Phosphate*: ~2.6 μM

Silicate*: ~100 μM

Fe: ~0.6 nmol/kg

B_{12}: <0.1 pM



in situ

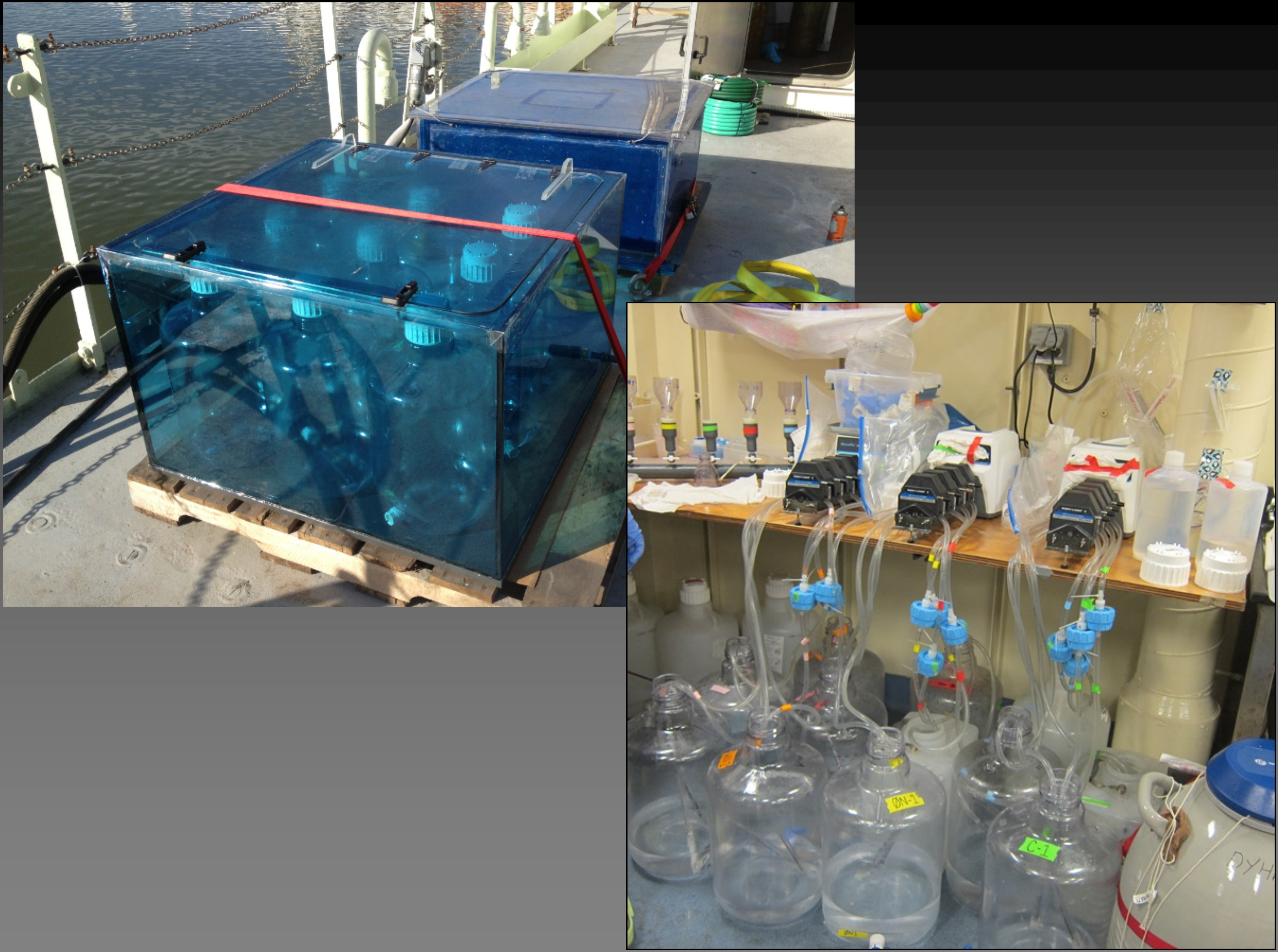
Control

Deep seawater amendment

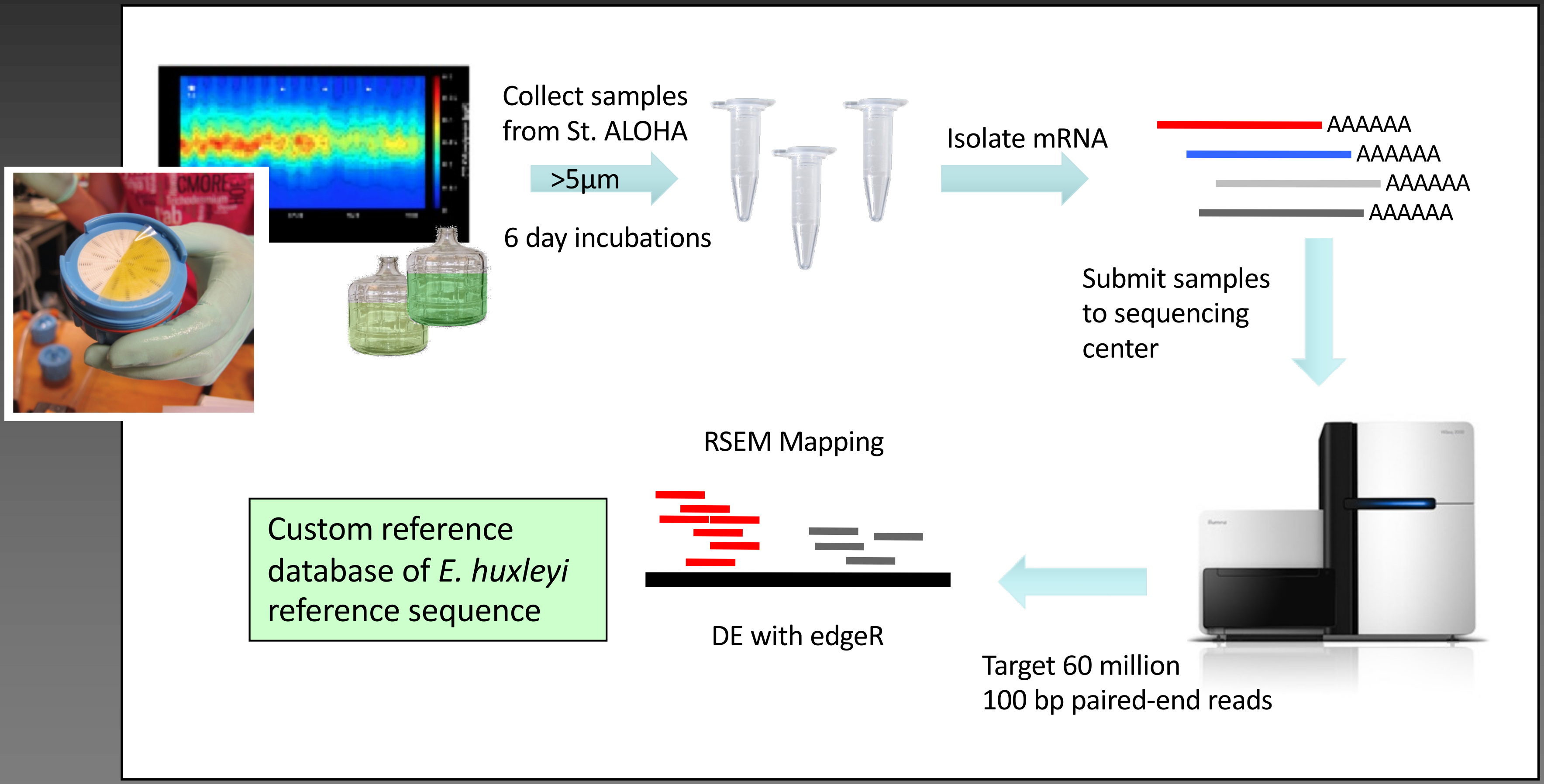
(+N, Si, Fe, B_{12})

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

N addition


0N-1
C-1

Sampling and pipeline



Collect samples from St. ALOHA

>5µm

6 day incubations

Isolate mRNA

Submit samples to sequencing center

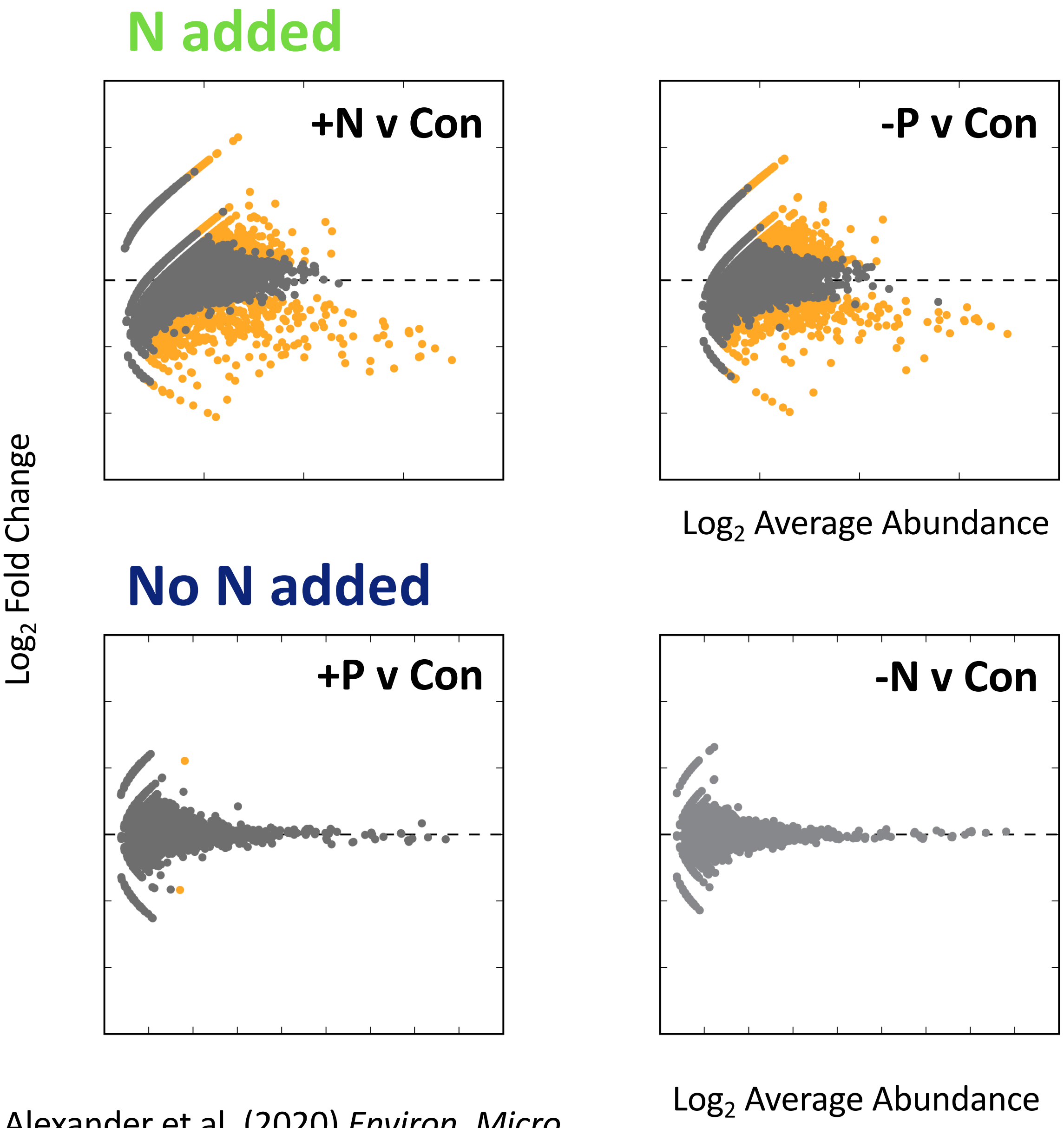
Target 60 million
100 bp paired-end reads

RSEM Mapping

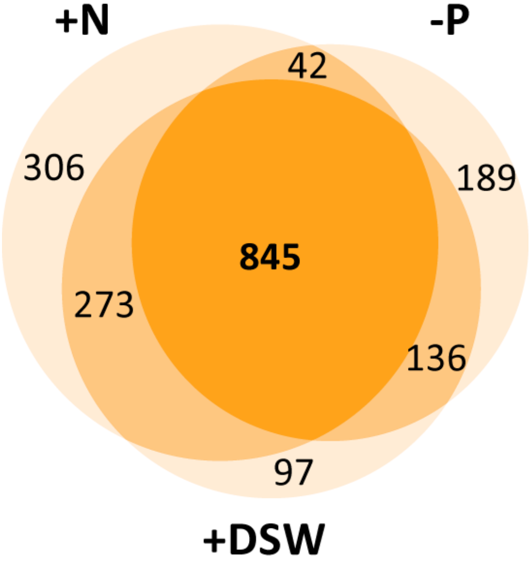
DE with edgeR

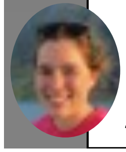
Custom reference database of *E. huxleyi* reference sequence

E. huxleyi physiological response suggests N control



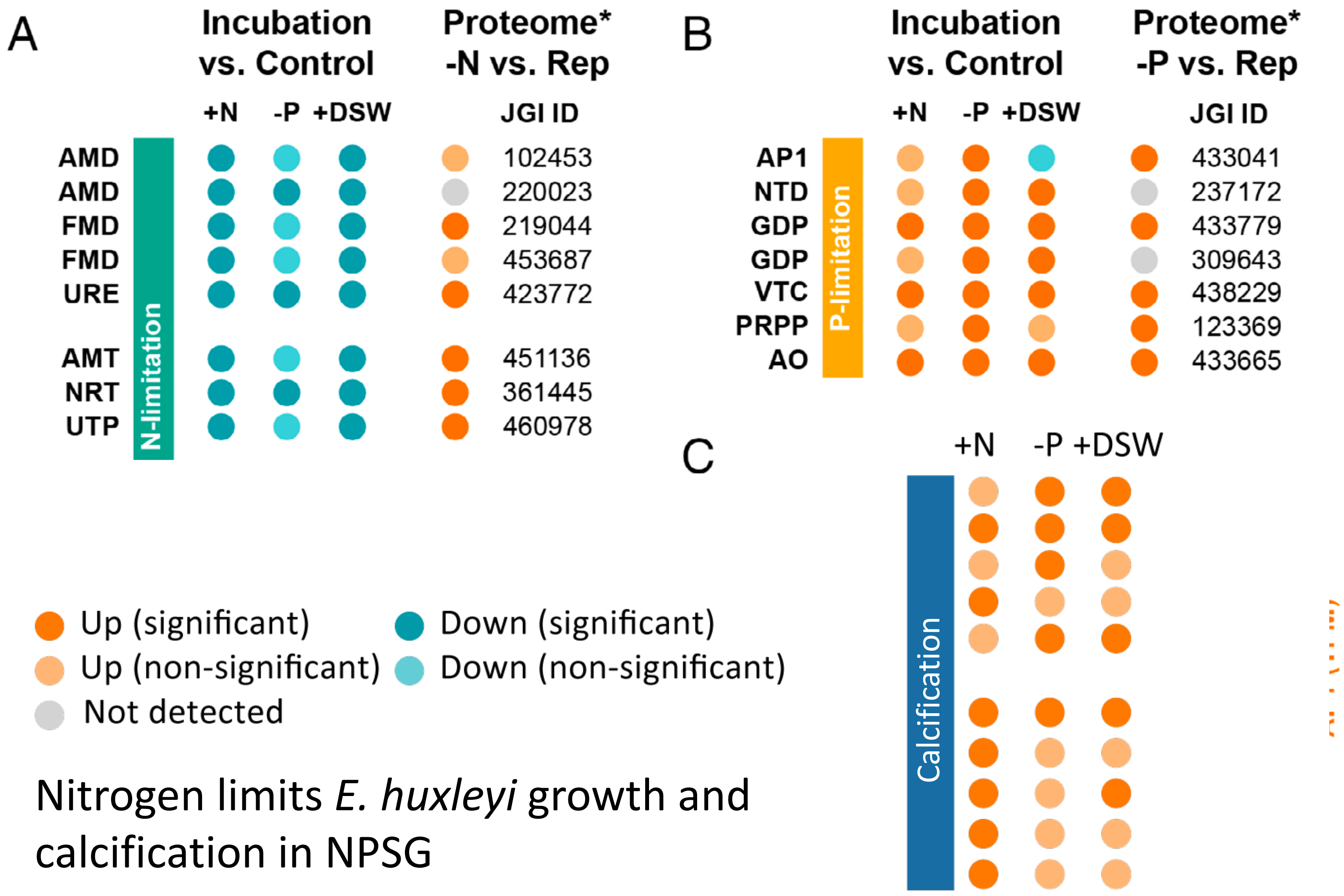
Significantly Differentially Abundant Genes





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

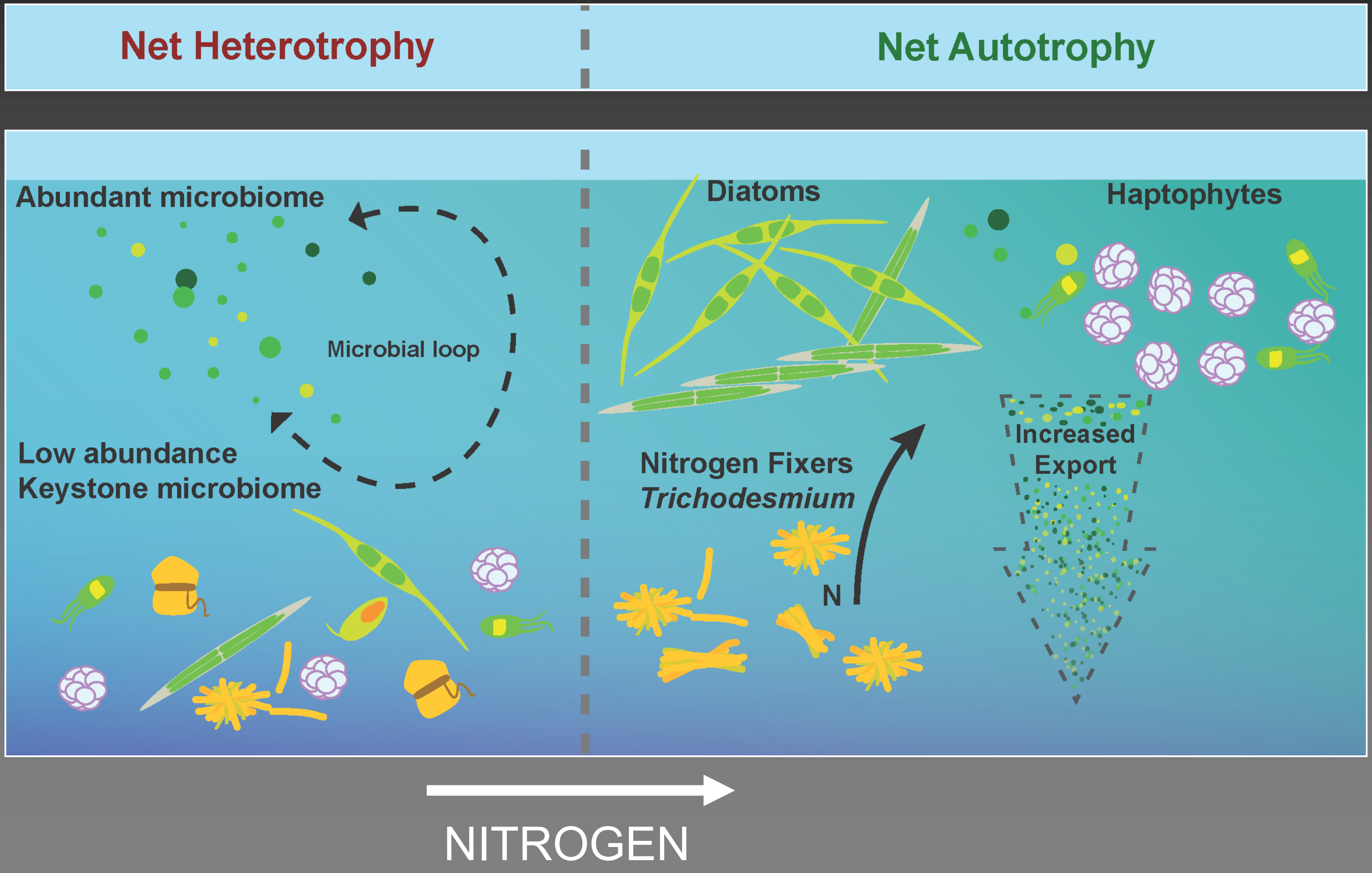
In situ transcriptome responses track with culture proteome



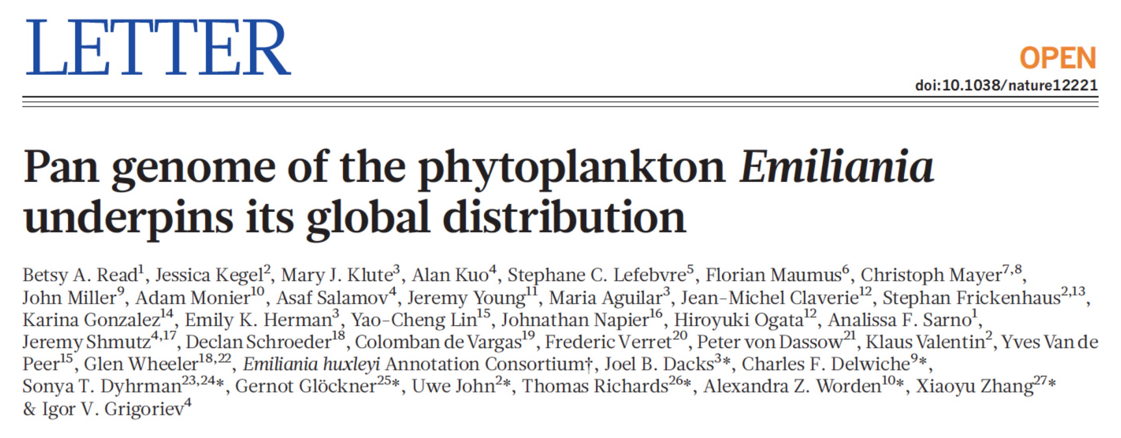
Nitrogen limits *E. huxleyi* growth and calcification in NPSG

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

Nitrogen is a driver of *E. huxleyi* dynamics



Emiliania huxleyi has many diverse isolates


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]

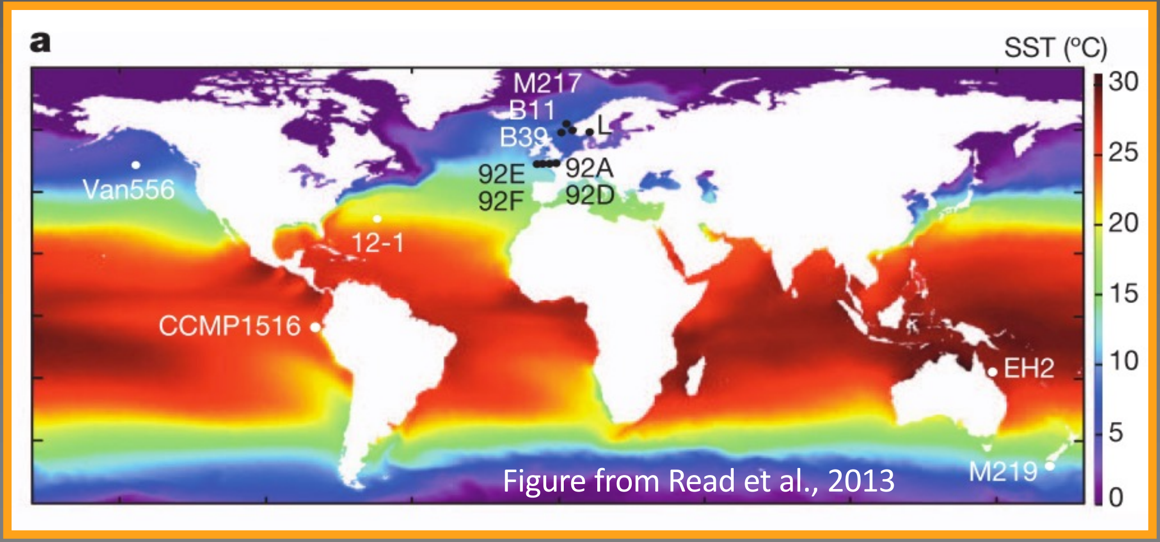


Figure from Read et al., 2013

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

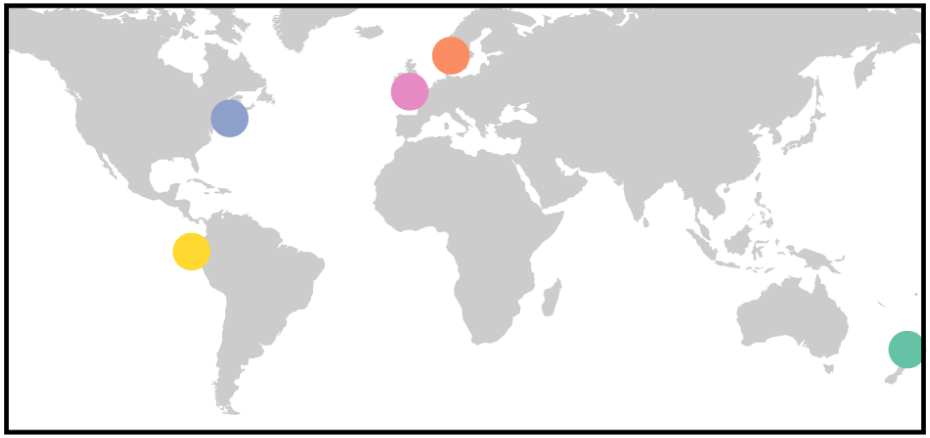
Baas Becking hypothesis

“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

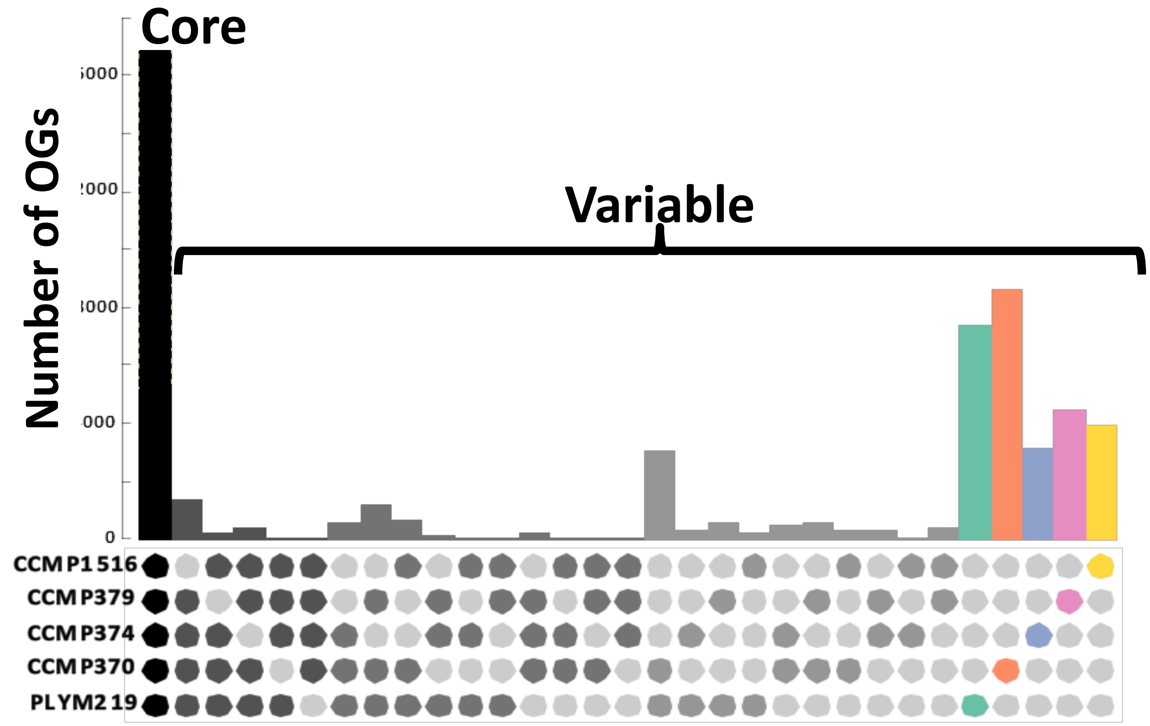


↓

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

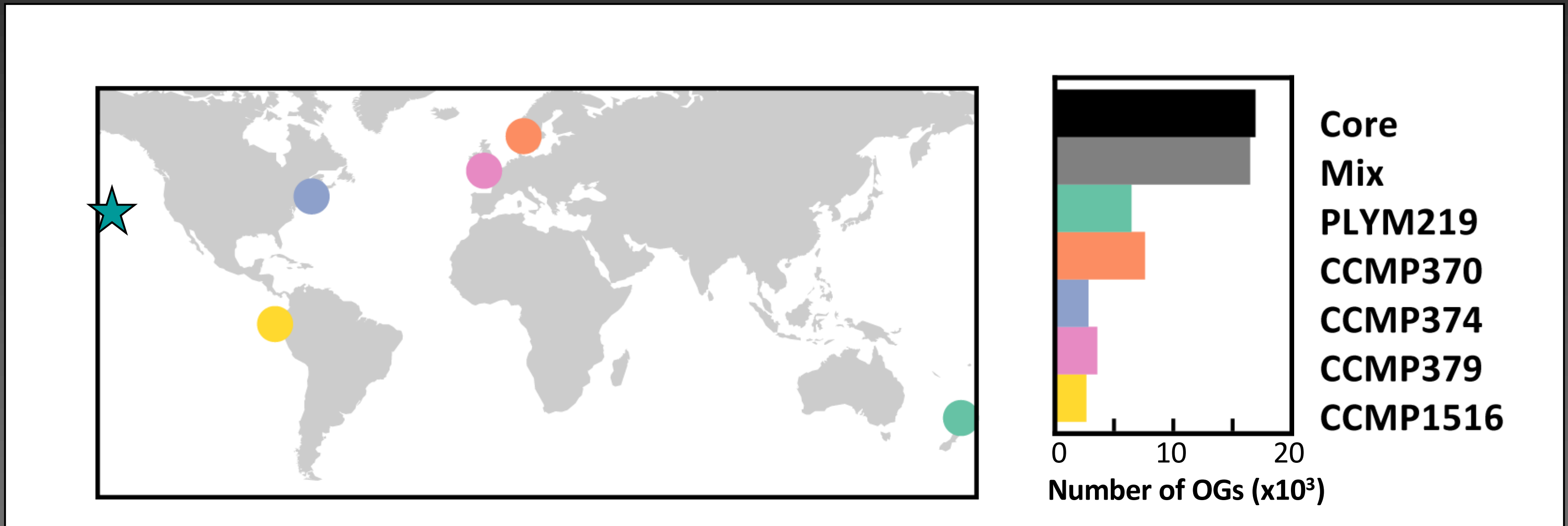
↓

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



Alexander et al. unpub.

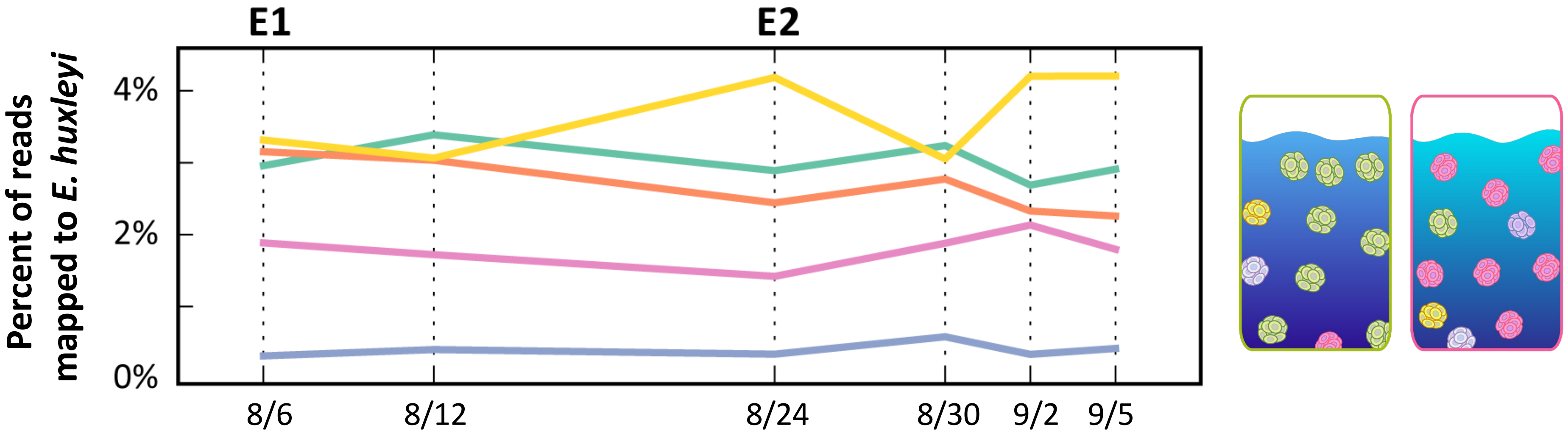
Tracking orthologous groups at St. ALOHA

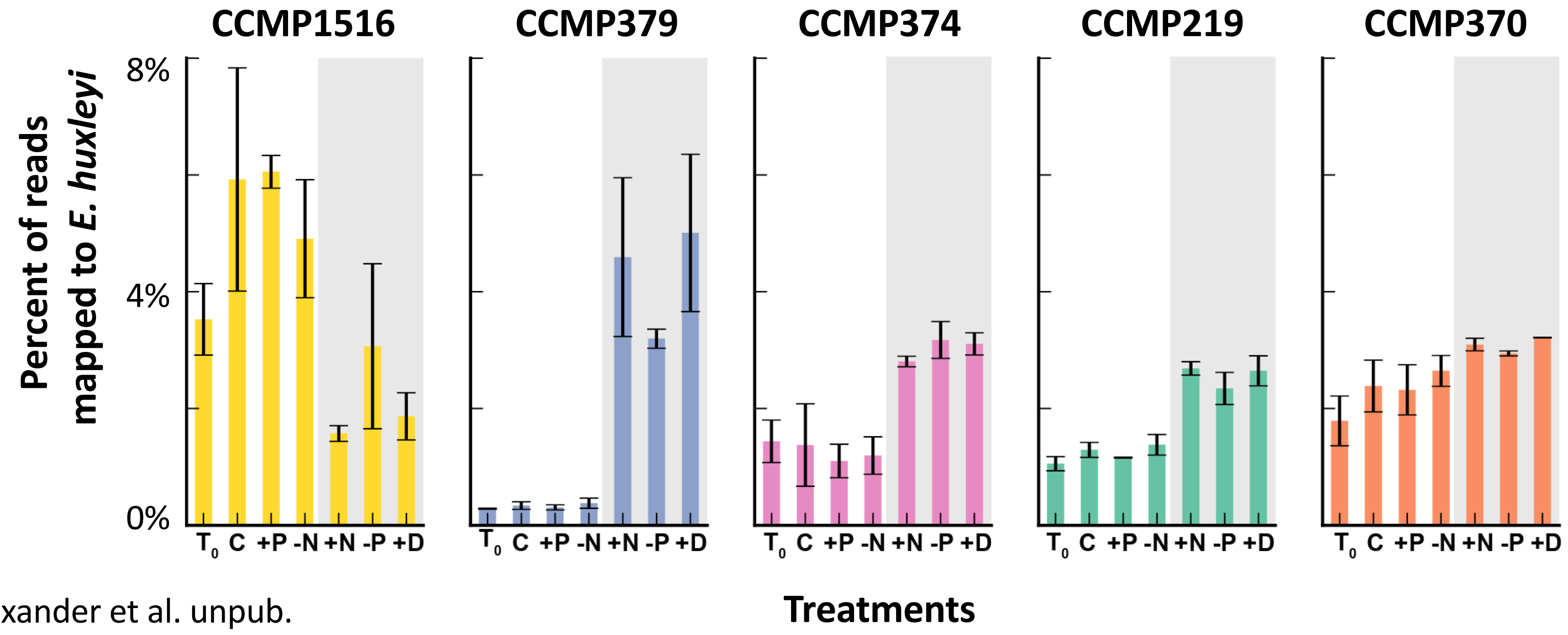


Alexander et al. unpub.

Use diagnostic OGs to track strain distribution

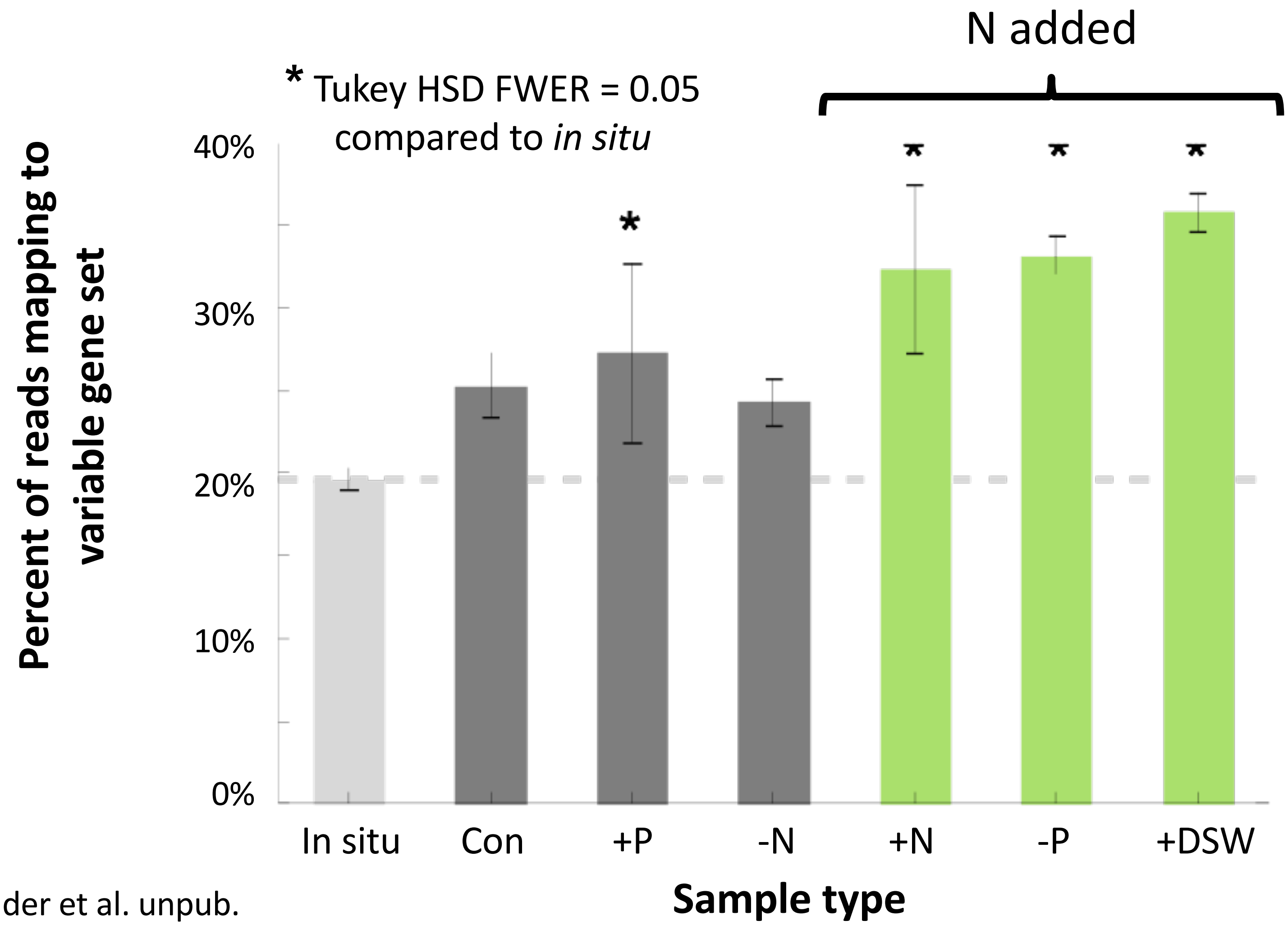
Strain distribution *in situ* and in the incubations





Alexander et al. unpub.

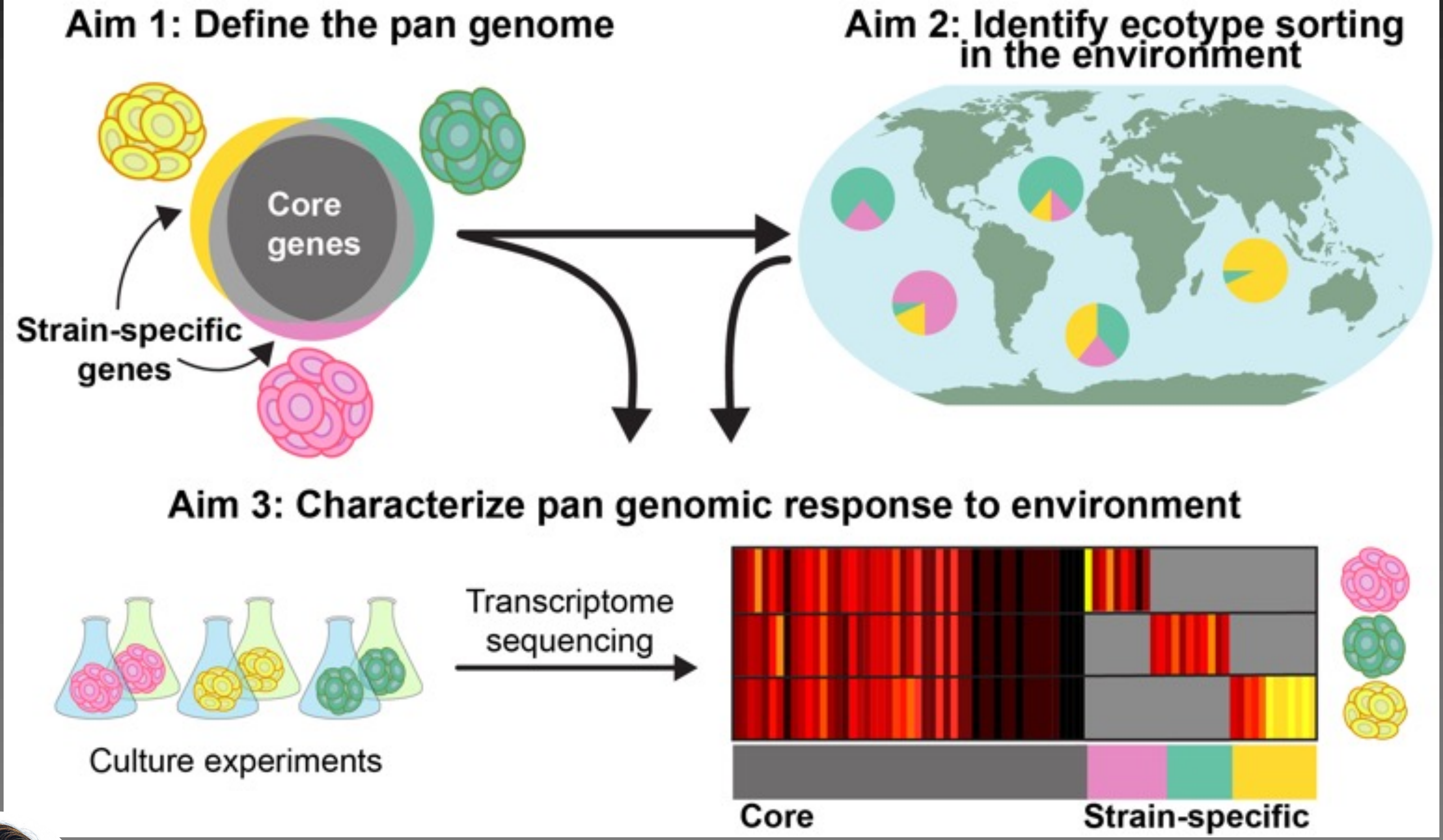
Expression of the variable portion of the pan-genome



Alexander et al. unpub.

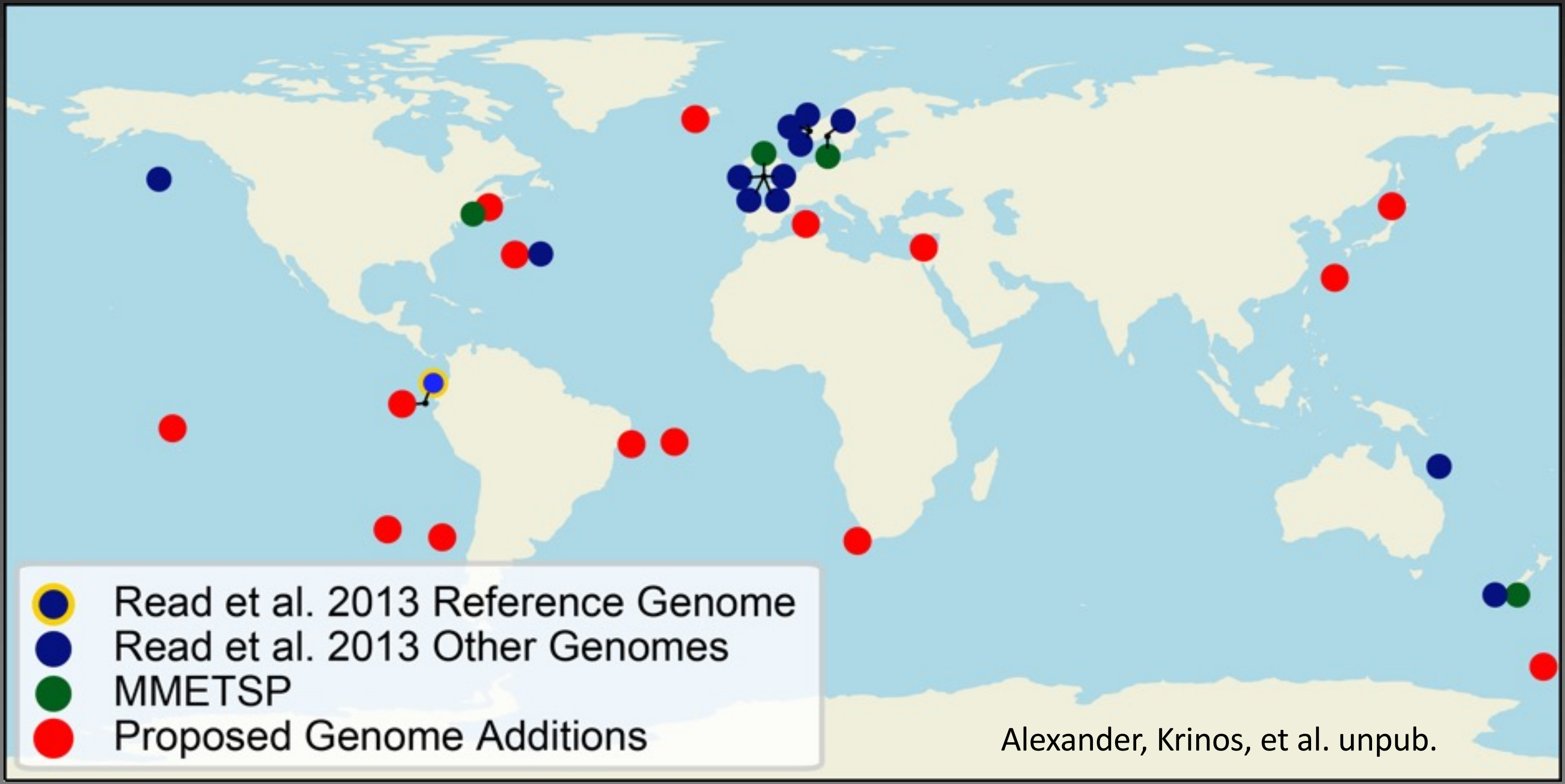
Expression of variable gene set significantly increased following N-addition

Ongoing work to define the pan genome



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

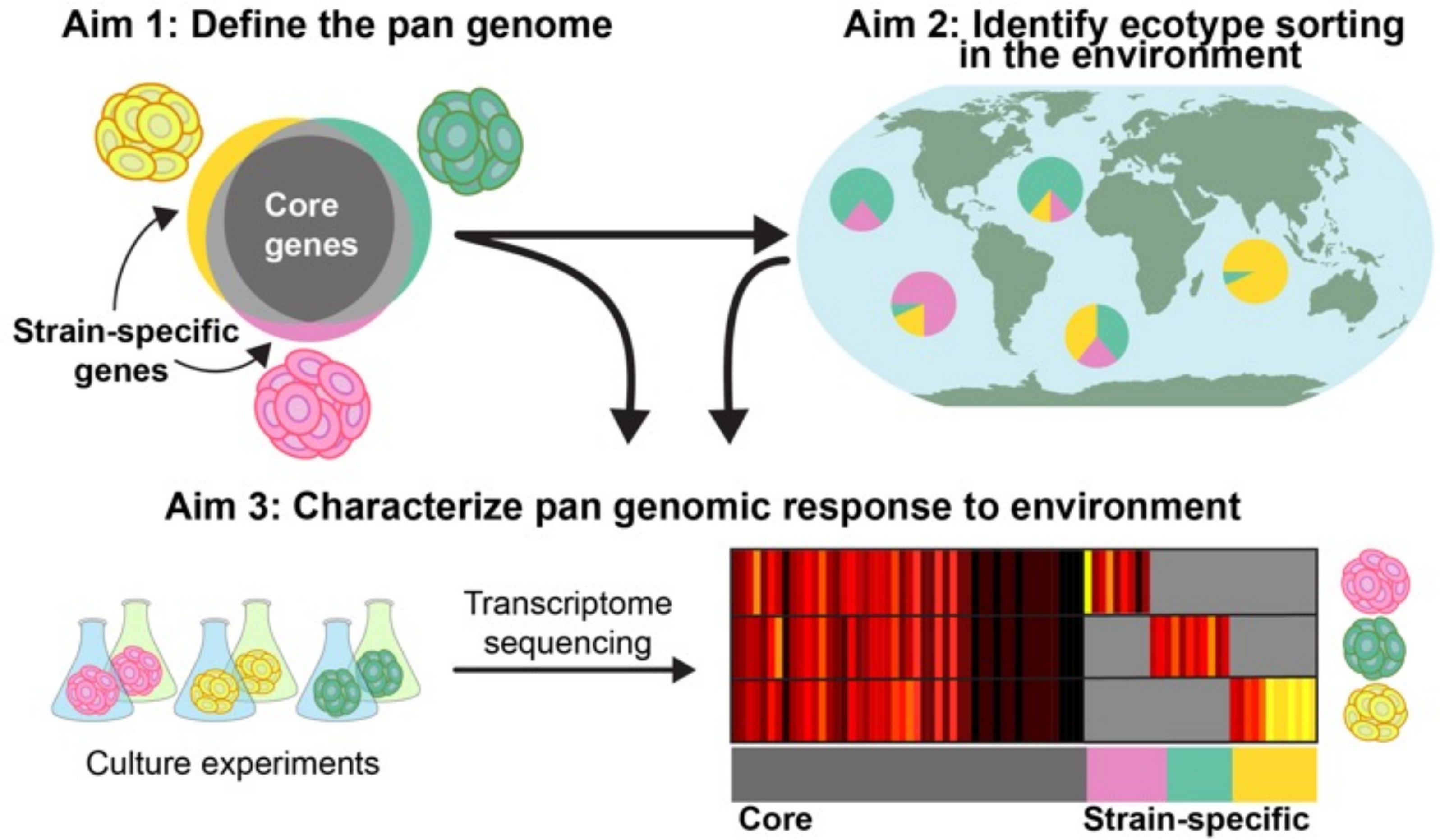
Expanding the pan genome



Alexander, Krinos, et al. unpub.

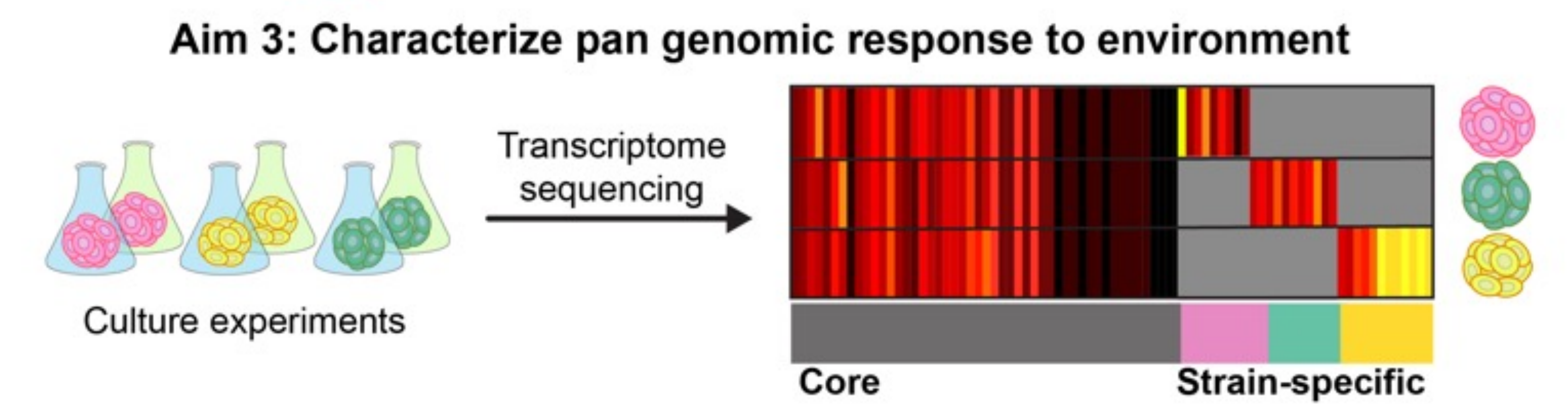
~15 new whole genome sequences being finalized (Nanopore/Illumina) and RNASeq libraries to support gene prediction

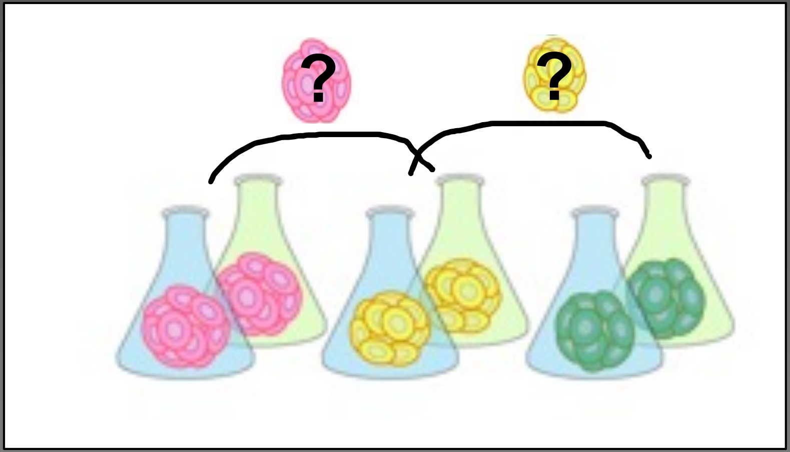
Future work to define the pan genome



Alexander et al. unpub.

How do genomic traits underpin competition outcomes?

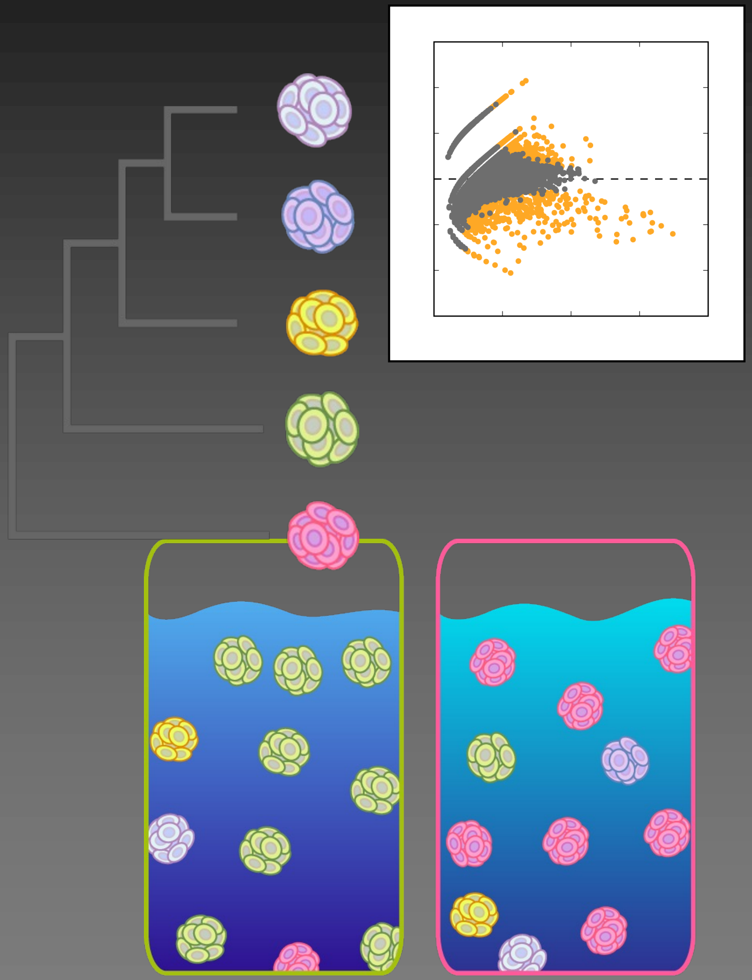




Haley, Palmer et al. unpub.

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.
- Move towards predictive responses of the chemical –microbe network in a changing ocean



Graphics: H. Alexander

Core questions… adventures in eukaryote metatranscriptomics

Limitation

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

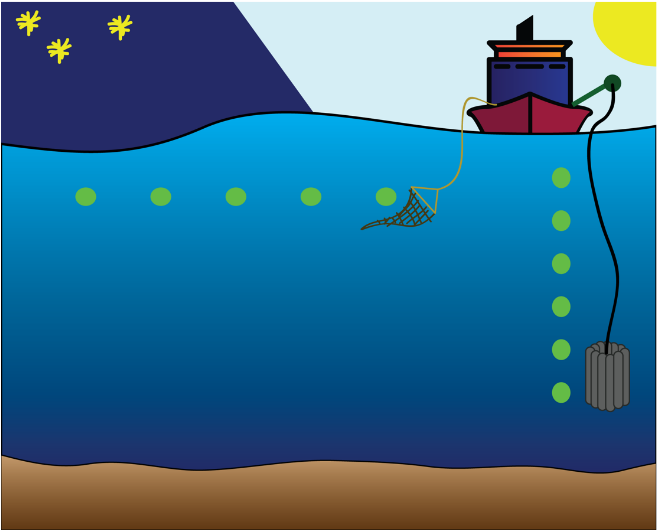
Competition

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

HOE:DYLAN

H. Alexander

HOE:MAHALO

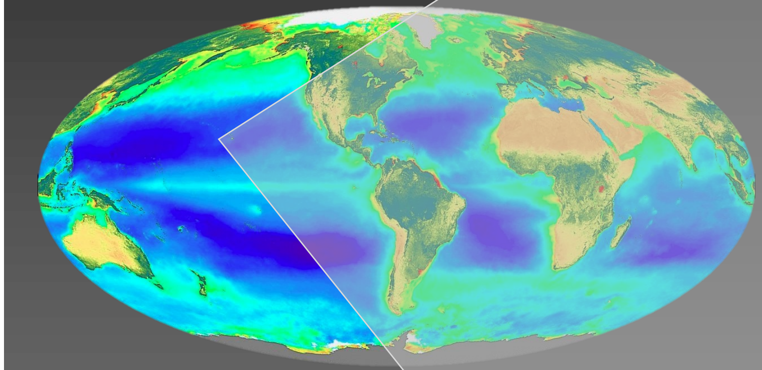


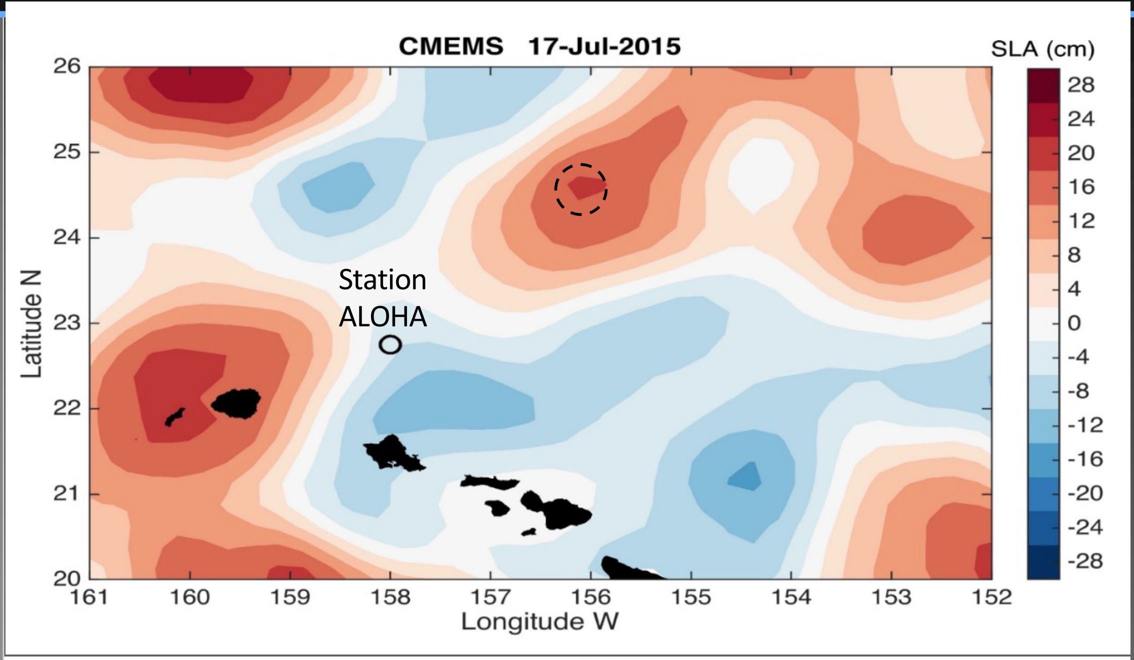
Hawaii Ocean Experiment: Dynamics of Light (MAHALO)

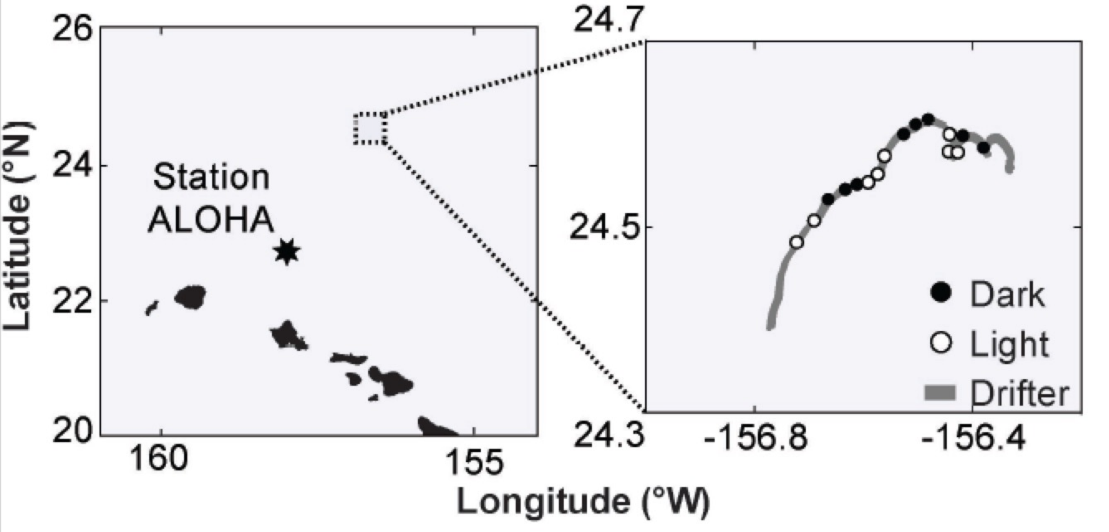
Figure 1 R/V Kilo Moana conducting one of the diel study CTD casts in 2015 (photo: Tara Clemente, taken from the R/V Ka'imikai-O-Kanaloa)

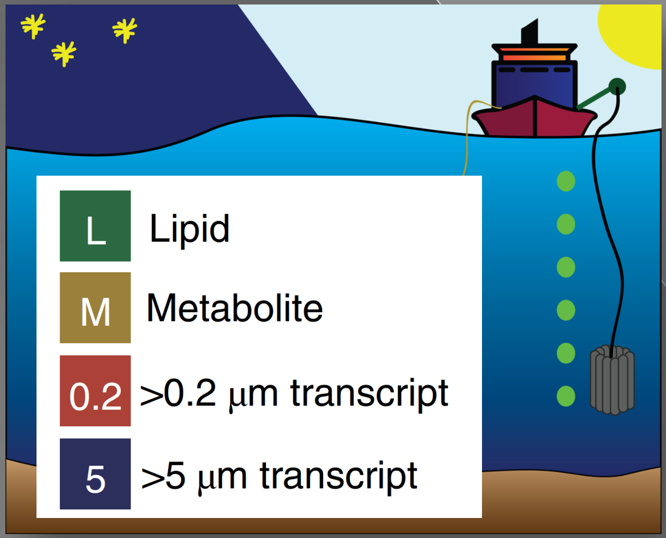
MAHALO: Special thanks to Daniel Muratore and Joshua Weitz for synthesis efforts

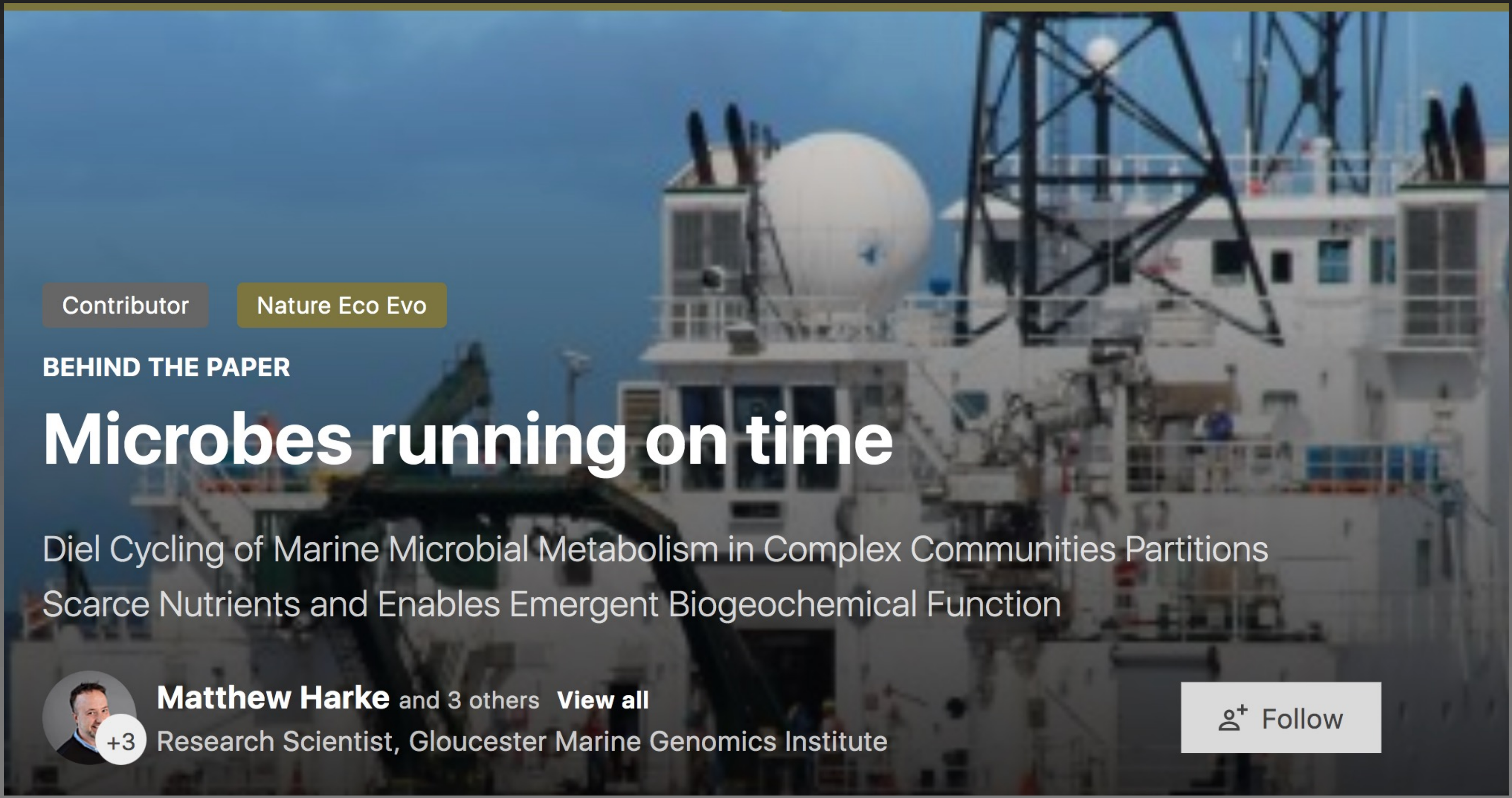
Sampling light driven responses in the NPSG











BEHIND THE PAPER

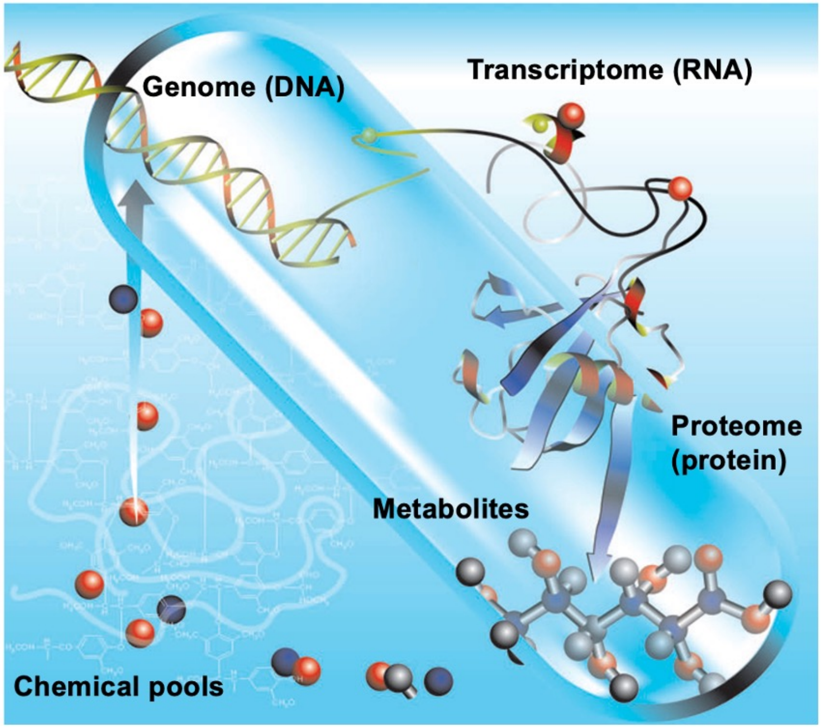
Microbes running on time

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

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

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

Jointly applying new multi ‘omic techniques

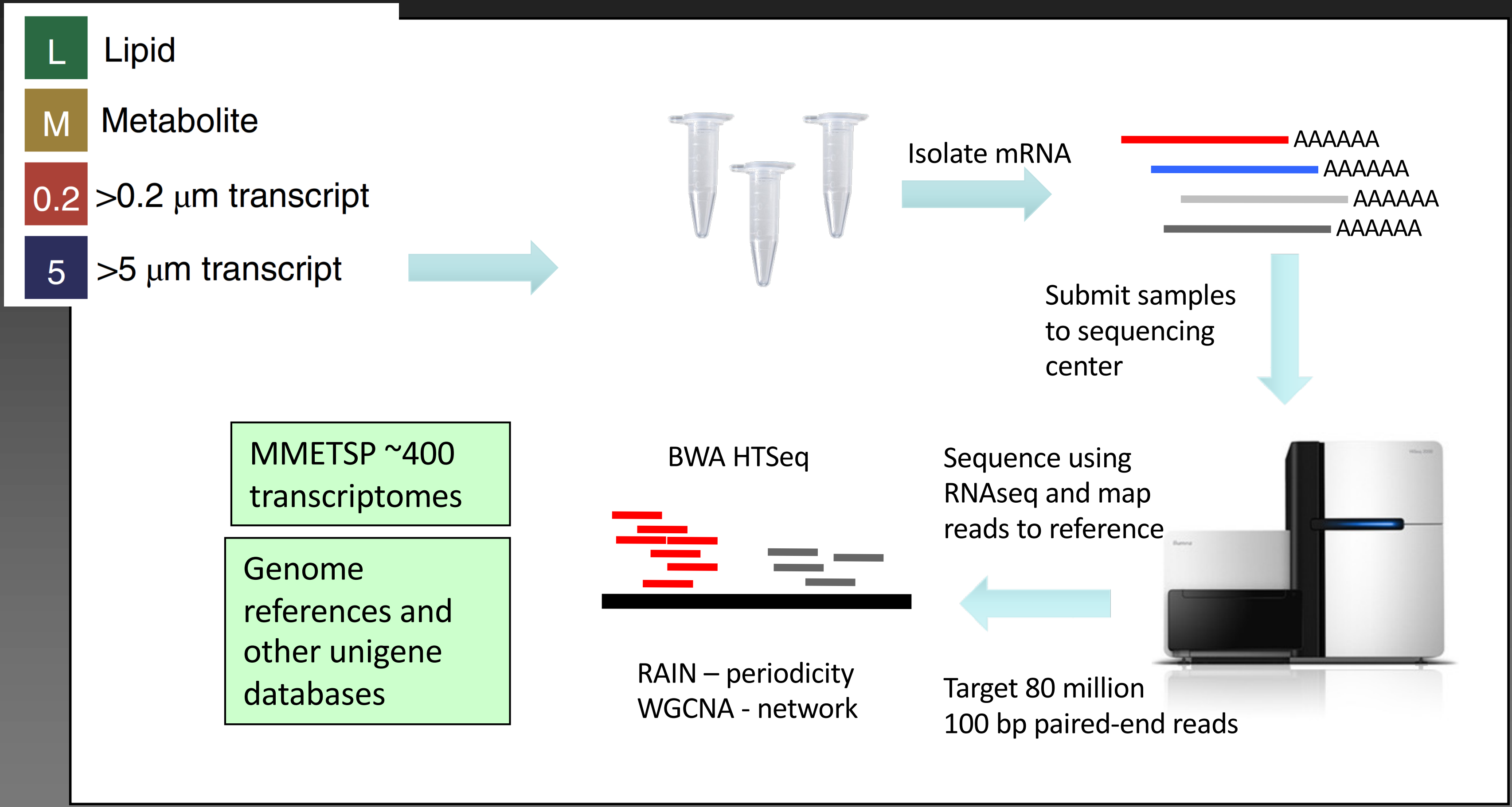


Kujawinski et al. (2011) *Annual Review Marine Science*

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
100 bp paired-end reads

BWA HTSeq

RAIN – periodicity
WGCNA - network

MMETSP ~400 transcriptomes

Genome references and other unigene databases

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

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

MMETSP = Marine Microbial Eukaryotic Transcriptome Project

Visualizing diel patterns

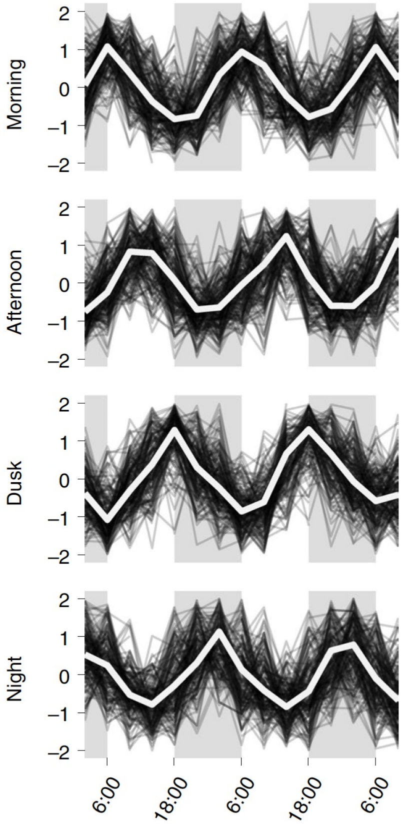
L Lipid

M Metabolite

0.2 >0.2 µm transcript

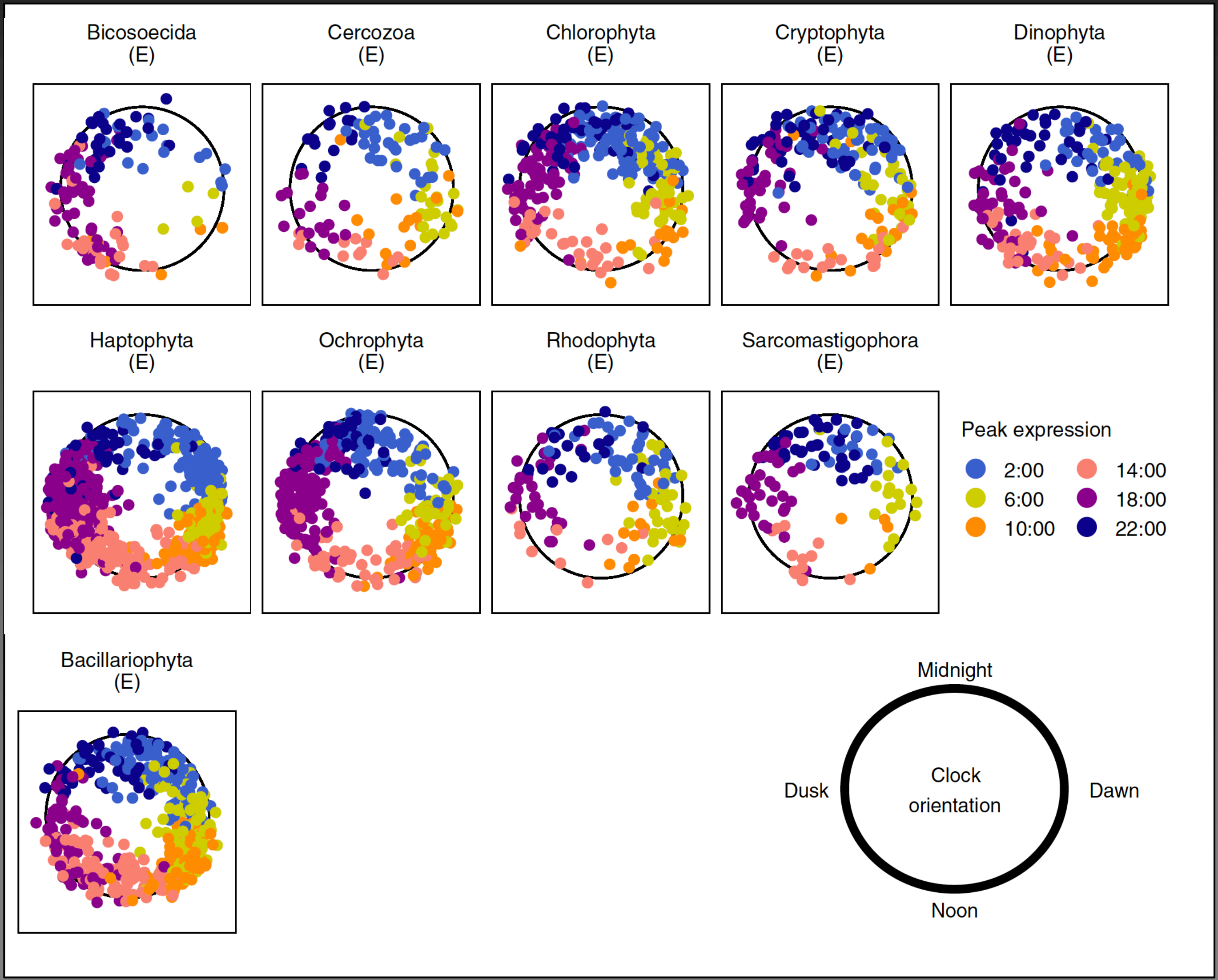
5 >5 µm transcript

~65,000 analyte
time series



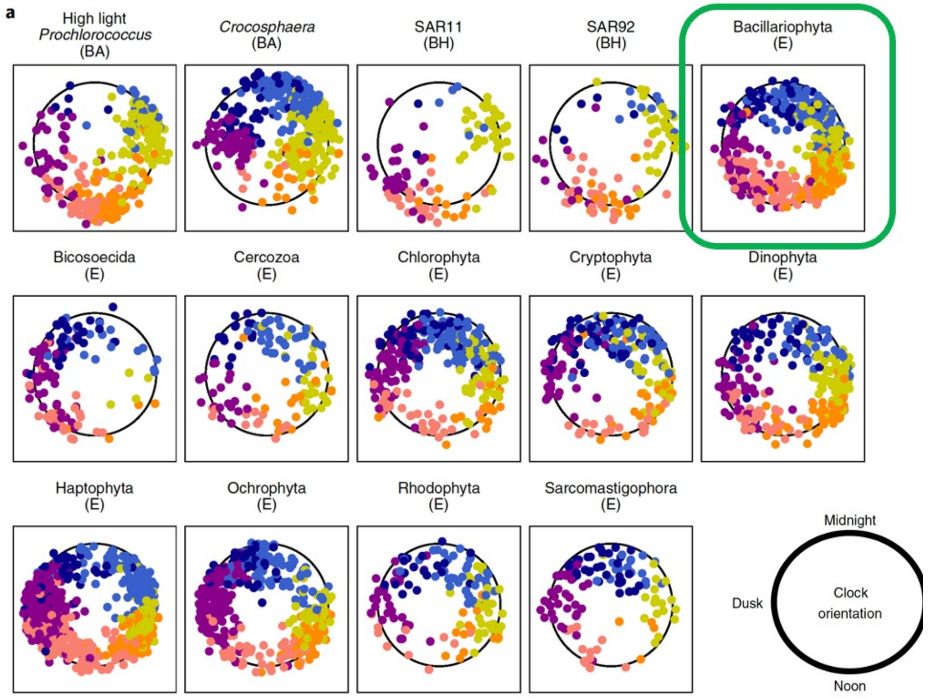
Muratore, Boysen, Harke et al. (2022) *Nature Ecol Evo*.

Different taxa have distinct diel transcription patterns

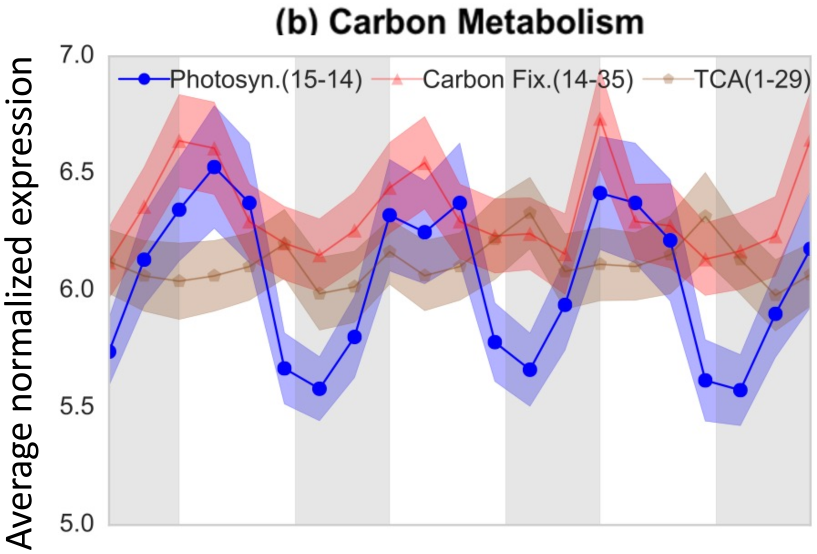


Muratore, Boysen, Harke et al. (2022) *Nature Ecol Evo*.

Carbon fixation and respiration are highly coordinated



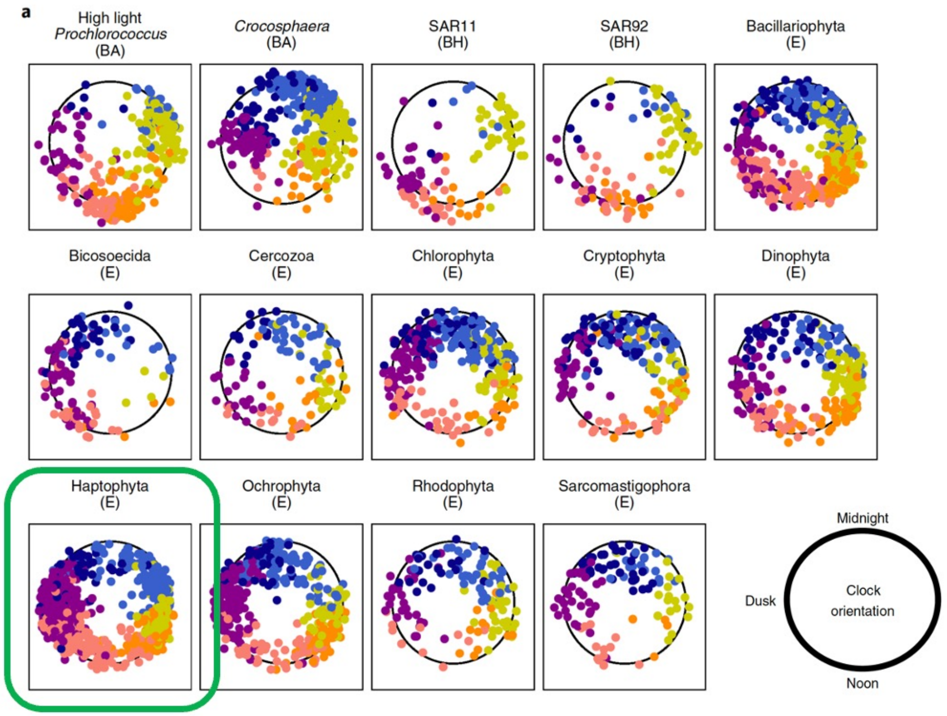
Rhizosolenia



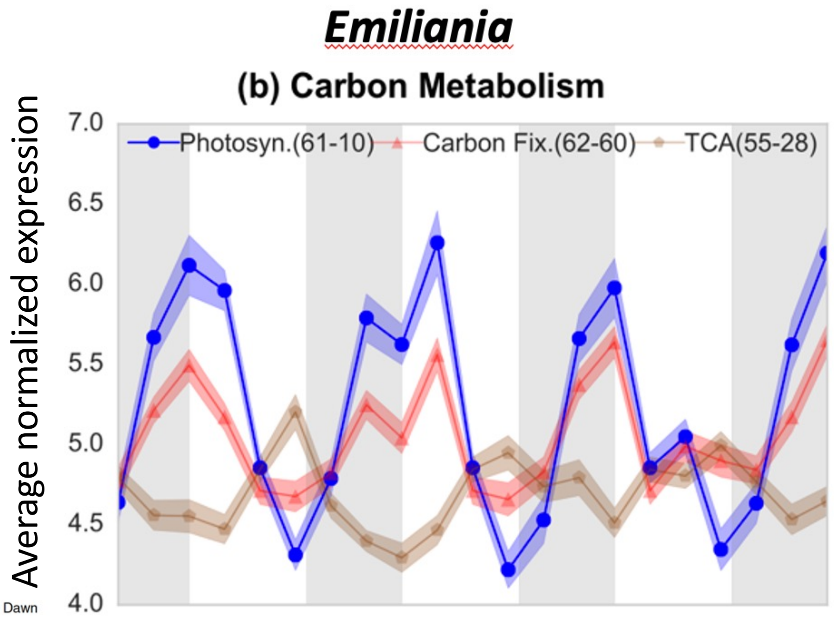
Muratore, Boysen, Harke et al. (2022) *Nature Ecol. Evol.*

Harke et al. 2019 *ISME J*

Carbon fixation and respiration are highly coordinated

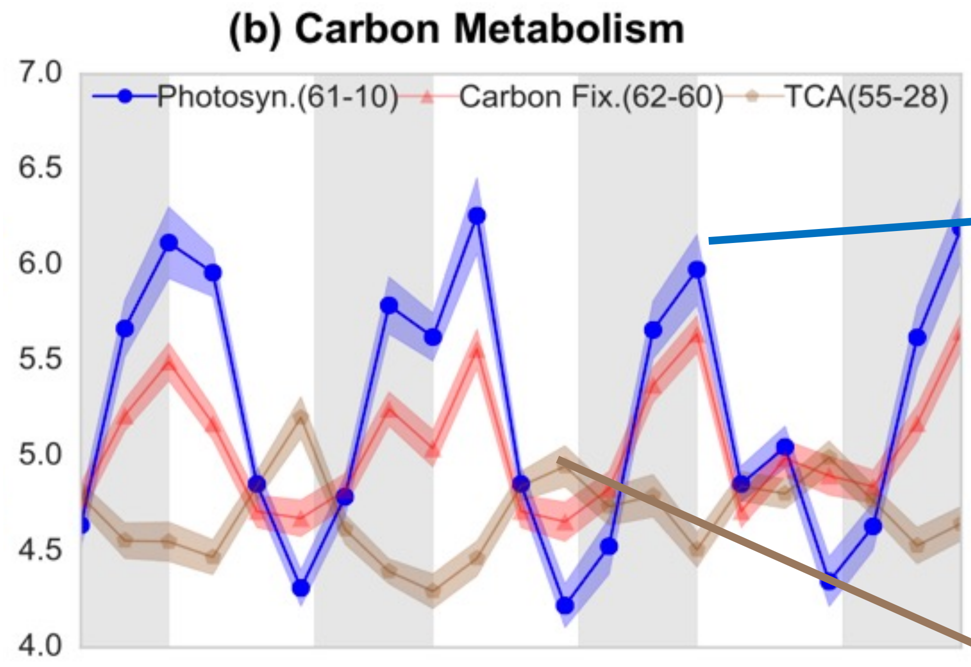


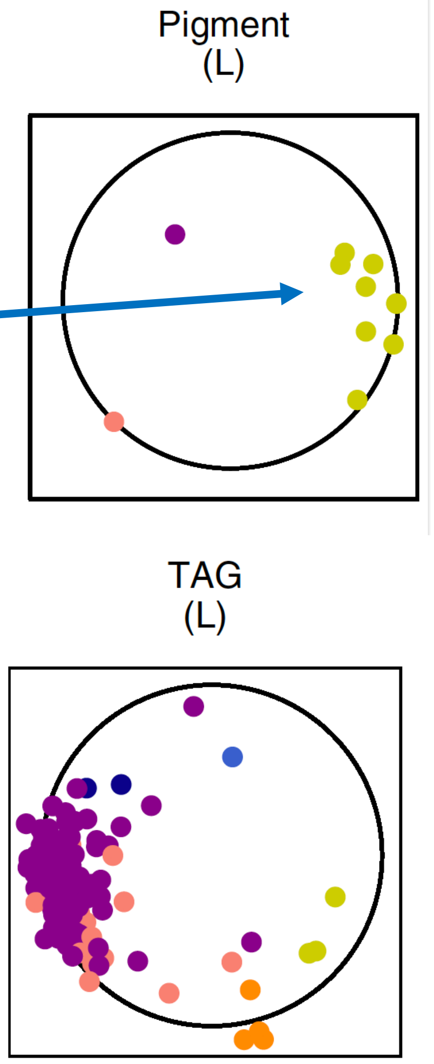
Muratore, Boysen, Harke et al. (2022) *Nature Ecol. Evol.*

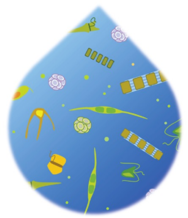


Limón-Hernandez et al. (2020) *Environ. Micro.*

Carbon metabolism synchronized across all phytoplankton







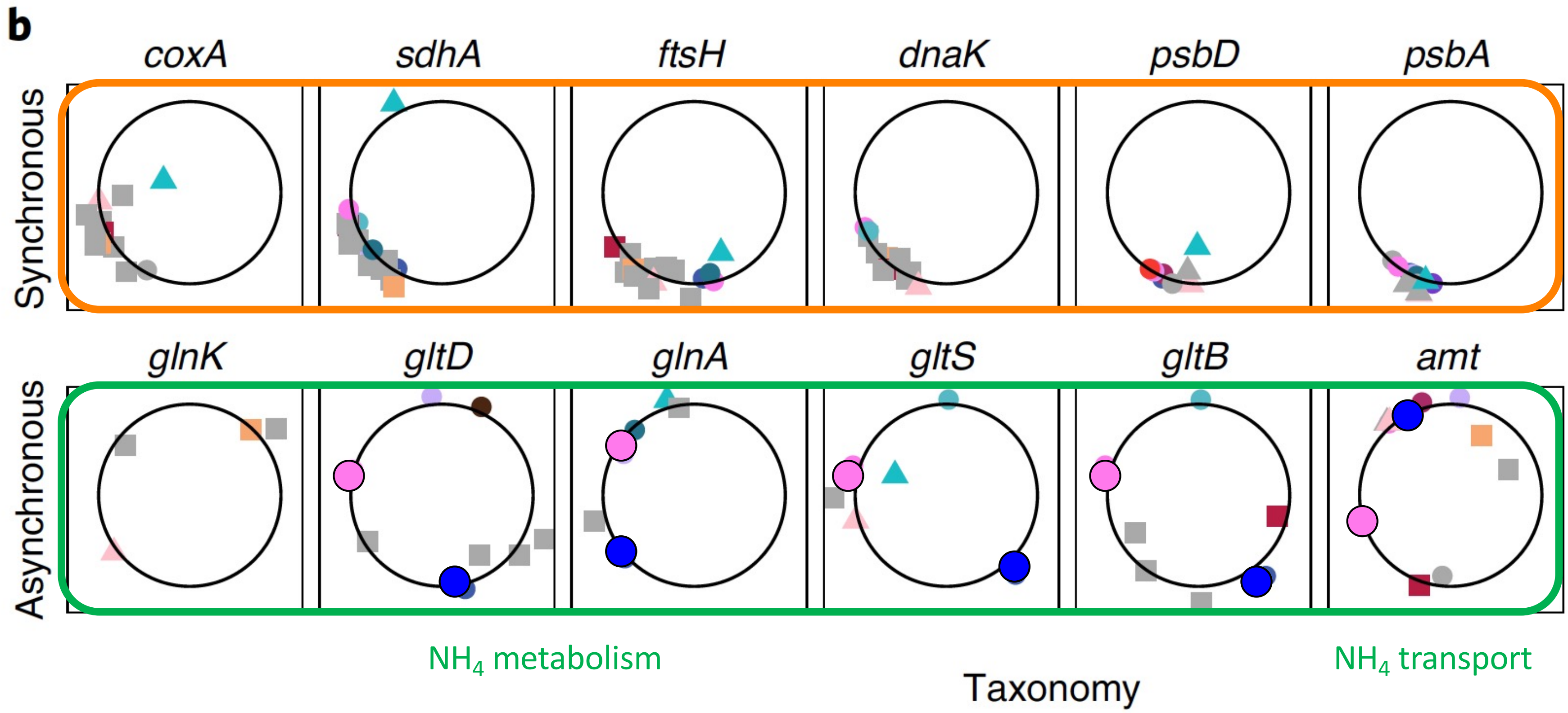
All photosynthesis, carbon fixation, and TCA cycle processes are synchronized across bacteria and eukaryotes!

Becker, Harke et al. (2021) *ISME J*

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

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

N metabolism functions are expressed asynchronously



NH_4 metabolism

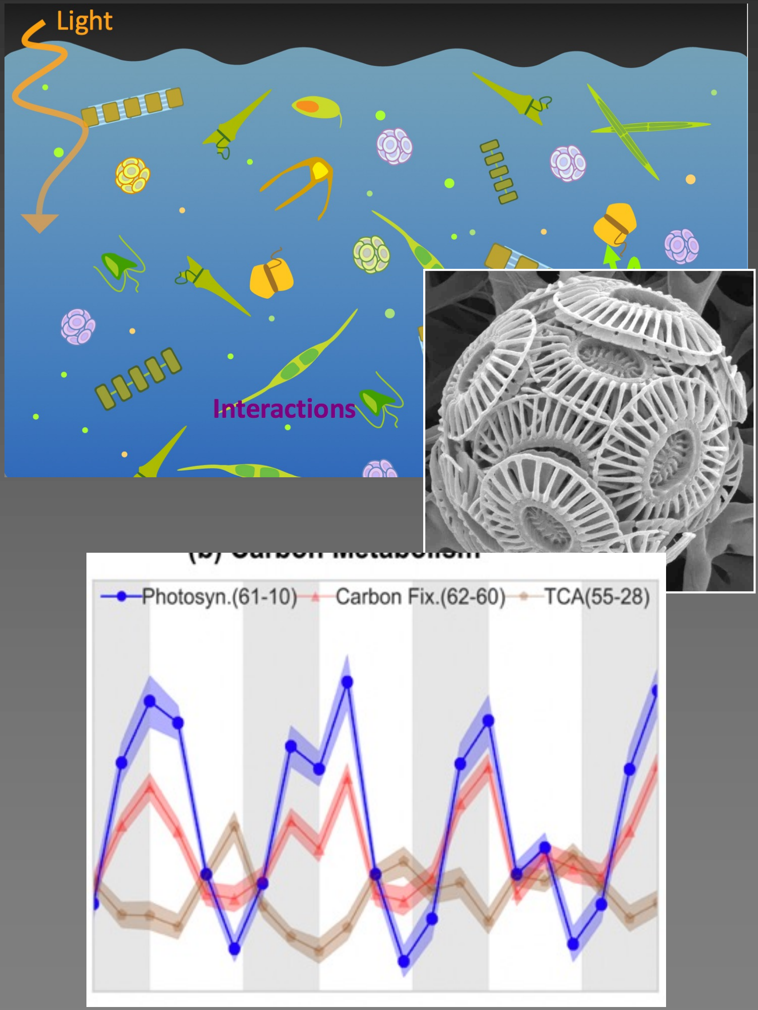
NH_4 transport

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*.

Summary

- **Metatranscriptomes**:
- 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 selects**

Competition

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

HOE:DYLAN

HOE:MAHALO