From flask to field: tracking the drivers of phytoplankton physiological ecology across marine ecosystems

Sonya Dyhrman
Department of Earth and Environmental Sciences
Lamont-Doherty Earth Observatory
Columbia University
The ocean makes our planet livable
The ocean acts as a buffer for CO$_2$ in the atmosphere

Between 1800 - 1994, ocean has absorbed ~120 petagrams of CO$_2$
Oceanic sink accounts for ~48% of fossil-fuel emissions

Sabine et al. (2004) *Science*
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
  - Account for roughly half of global primary production

- *make the planet habitable*

Image courtesy C-MORE
Marine phytoplankton are highly diverse
Phytoplankton underpin ocean ecosystem function

Haptophytes

Diatoms

N₂ Fixers

Chisholm 2000 Nature
Phytoplankton play a profound role in the earth system

Half of global primary production
Seasonal chlorophyll distributions in the sea - highlights the global significance of phytoplankton
Sampling microbes across marine ecosystems
Tracking physiological ecology: from the flask to the field

**Culture-based experiments**
Species-specific responses to well-controlled environment

**Limitations:**
- Species must be in culture
- Time consuming
- Extrapolations to the field

**Field-based studies**
Assess whole community dynamics in a natural environment

**Limitations:**
- Not species-specific

Micro/Mesocosm

‘Omic-enabled advances allowing to query cells in their environment in a species-specific way
Challenges and opportunities in microbial oceanography

- Long standing challenges:
  - Populations are dilute
  - Few species-specific assays
  - Few genome or transcriptome sequences

- New opportunities
  - Novel concentration and detection strategies
  - Increases in whole genome sequences
  - Increases in transcriptomes for eukaryotic taxa

Increasingly able to use ‘omic and ‘metaomic approaches!
Leveraging ‘omic data to study marine microbes

**Taxonomic Diversity:** Who is there?

**Metabolic capacity:** What are the molecular underpinnings of resource metabolism?

**Metabolic plasticity:** How are those pathways regulated and expressed *in situ*?
How did we get here?

- **1990s**
  - First marine bacterial WGS
  - Fosmid cloning of community DNA

- **2000s**
  - BAC libraries of community DNA
  - Sequencing and WGS assembly of whole community DNA
  - First marine microbial eukaryote WGS

- **2010s**
  - Bacterial community RNA sequencing
  - Bacterial community proteomics and metabolomics
  - SAGs
  - MMETSP - marine microbial eukaryote transcriptomes
  - Eukaryotic community RNA sequencing
Sequencing advances opens new ’omic approaches
Sequencing advances opens new ’omic approaches
Sequencing advances opens new ’omic approaches
Sequencing advances opens new ’omic approaches

Cost per Raw Megabase of DNA Sequence

Thalassiosira pseudonana

Image credit: DOE Joint Genome Institute

The marine centric diatom *Thalassiosira pseudonana* was chosen as the first eukaryotic marine phytoplankton for whole genome sequencing because this species has served as a model for diatom physiology studies, the genus *Thalassiosira* is cosmopolitan throughout the world’s oceans, and the genome is relatively small at 34 mega base pairs.
Sequencing advances opens new ’omic approaches
Sequencing advances opens new ’omic approaches

Figure from Read et al., 2013
Ongoing challenges

We need to grow the inventory of species and genes

We have too few samples, and too few contiguous datasets

Processing loses ecological context

Lots of “dark matter”

Still too expensive to generate and store data
• From genome to biome: Tracking the metabolism and microbiome of a keystone N$_2$ fixer

![Image of N$_2$ fixer]

Genome - enabled

• Co-existing in a sea of competition: Leveraging transcriptome data to track the physiological ecology of phytoplankton from key groups

![Image of phytoplankton]

Transcriptome - enabled
Thank you

Gwenn Hennon  Mónica Rouco  Sheean Haley  Kyle Frischkorn

Matt Harke  Maria Hernandez  Harriet Alexander
Core questions

Metabolic traits and trade-offs

• What phosphorus is bioavailable?
• What are the biogeochemical constraints on N$_2$ fixation?

Host microbiome interactions

• Who is there? Microbiome diversity
• What are they doing? Microbiome functional diversity, holobiont physiology and controls on N$_2$ fixation
Nitrogen-fixing marine cyanobacteria

- **Symbionts**
  - UNCYN-A
  - *Richelia*

- **Free-living**
  - *Crococosphaera*
  - *Trichodesmium*
    - *Trichodesmium contortum*
    - *Trichodesmium erythraeum*
    - *Trichodesmium tenue*
    - *Trichodesmium thiebautii*
    - *Trichodesmium spiralis*
    - *Trichodesmium hildebrandtii*
Trichodesmium: critical to ecosystem function
Trichodesmium: critical to ecosystem function

Responsible for approximately 50% of marine N₂ fixation

Fe?  
P?
Phosphorus pools in the open ocean

Particulate phosphorus

Dissolved inorganic phosphorus

DIP

Dissolved organic phosphorus

DOP

Phosphonate

Esters (C-O-P)

HMWDOM

Phosphonate (C-P)

Kolowith et al. 2001 Limnol. Oceanogr.

Lomas et al. (2010) Biogeosciences
Trichodesmium erythraeum IMS101 genome page

Marine filamentous cyanobacteria of the genus Trichodesmium play a major role in the tropical and subtropical oceans both as primary producers and suppliers of "new" nitrogen through their ability to fix atmospheric dinitrogen (N2) (Capone et al., 1997). They are simple undifferentiated filamentous forms that divide in a simple plane. They may occur as single filaments but more commonly as macroscopic colonial aggregates (0.5-2 mm) containing many filaments. The colonies vary in color from yellowish-brown to deep red because they contain phycoerythrin as their primary light harvesting pigment. Trichodesmium spp. are planktonic and use their buoyancy to the possession of gas vacuoles. The genus currently contains five species characterized from natural populations (Carpenter et al., 1993; Janson et al., 1995). The species are distinguished by cell and colony morphology, pigmentation and buoyancy.

Questions about this project, contact: Paul Richardson, DOE Joint Genome Institute © 2002.
Phosphorus metabolic traits and trade-offs

- Phosphonate
  - C-P Lyase (Fe co-factor)

- Ester
  - *phoX* type alkaline phosphatase (Ca Fe co-factor)
  - *phoA* type alkaline phosphatase (Zn co-factor)

- Phosphite
  - *ptxD* gene cluster

Comparative genomics: phosphorus traits and trade offs

Other $N_2$ fixing cyanobacteria genomes do not encode the same pathways for phosphorus metabolism - less available substrates, but less metal requirement

Tracking genomic potential with expression studies

Culture cells

↓

Harvest and preserve samples

↓

Transcriptome
Proteome

+P  -P  RF  RF
Collect samples
Remove rRNA
Submit samples to sequencing center
Sequence using RNAseq
Stringent mapping to IMS101 genome
Removed reads with ambiguous mapping
Target 30 million 100 bp paired-end reads

Differential expression evaluated with EdgeR
Proteome analysis pipeline

Collect samples → Extract protein → Generate peptide → LC-MS/MS → ID MS against gene models

Relative differences of spectral counts with Scaffold
How does *Trichdoesmium* respond to low phosphorus?
Transcriptome - proteome coordination
Transcripts turnover more quickly than proteins
How does *Trichdoesmium* respond to low phosphorus?

*phoX* transcripts

*PhoX* proteins
Tracking genomic potential with expression studies

Culture cells

Harvest and preserve samples

qRT-PCR of phoX Activity

P supply

Growth rate

25%  15%  10%
Calibrating gene expression to growth and \( \text{N}_2 \) fixation

Orchard and Dyhrman unpublished
Calibrating gene expression to N$_2$ fixation

Orchard et al. in prep
Sampling different P regimes

- High P
- Low P
Measurements of quantitative gene expression for *Trichodesmium sp.*
Gene expression increases at low phosphorus

$r^2=0.86$ (TDP)

$r^2=0.75$ (DIP)

Low P supply
North Atlantic

High P supply
South Pacific

Orchard and Dyhrman unpublished
Calibrating gene expression to $N_2$ fixation

Orchard and Dyhrman unpublished
Data predicts that P supply limits N$_2$ fixation in the western N. Atlantic

Orchard and Dyhrman unpublished
Constraints on *Trichodesmium* N$_2$ fixation

Molecular patterns corroborate predictive models in the western north Atlantic

*phoX* - P regulated ester metabolism (Orchard et al. 2009 *Environ. Micro.*)
*idiA* - Fe regulated iron metabolism (Chappell et al. 2013 *ISME J.*)
*rnpB* - reference gene
*nifH* - N$_2$ - fixation

(Moore et al. 2004)
High DIP/DOP
Low Fe
(n=3)

Low DIP/DOP
High Fe
(n=12)
Collect samples
Remove rRNA
Submit samples to sequencing center
Sequence using RNAseq
Target 30 - 60 million 100 bp paired-end reads

N. Atlantic  N. Pacific

Relative abundance

Clade I  Clade III  Clade IV  Unclassified

Genome strain


Stringent mapping to custom database
Target 30 - 60 million 100 bp paired-end reads
Collect samples

Remove rRNA

Submit samples to sequencing center

Sequence using RNAseq

Stringent mapping to custom database

Curated, custom database: *Trichodesmium* metagenome sequences clustered into orthologous groups

Removed reads with ambiguous mapping

Target 30 - 60 million 100 bp paired-end reads
The reality.....
Metatranscriptome analysis

**RNA extraction**

**Bacterial mRNA enrichment:**
- Euk RNA removal - MICROBEnrich kit (Ambion)
- Bacterial rRNA removal - Ribo-Zero (Epicentre)

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

**Output:**
- .fastq

**Preparation of reference metagenome:**
- Extraction of *Trichodesmium*-only scaffolds from metagenome data

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

**Sequence processing:**
- Sequence quality - FASTQC (.fastq)
- Trimming - Trimmomatic (.fastq)

**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
Biomarkers consistent with model prediction

Fe-limited P-limited

Biomarkers confirm model predictions.
Significant differences in global transcription between the North Pacific and the North Atlantic

Ruoco et al. (2018) ISMEJ
Transcription patterns indicate metalloenzyme trade-offs and geochemical controls.

Ruoco et al. (2018) *ISMEJ*
Summary - Metabolic traits and trade-offs

What phosphorus forms are bioavailable?
What are the biogeochemical constraints on N\textsubscript{2} fixation?

- **Genome**: *Trichodesmium* genome suggests bioavailability of phosphonate, ester, and phosphite

- **Marker transcripts**: *Trichodesmium phoX* expression levels suggest that supply of bioavailable P is low in the western N. Atlantic, which could constraint N\textsubscript{2} fixation

- **Metatranscriptome**: Predicted biogeochemical drivers of N\textsubscript{2} fixation are reflected in *Trichodesmium* transcriptional signals including likely metalloenzyme switching
Core questions

Metabolic traits and trade-offs

• What phosphorus is bioavailable?
• What are the biogeochemical constraints on N₂ fixation?

Host microbiome interactions

• Who is there? Microbiome diversity
• What are they doing? Microbiome functional diversity, holobiont physiology and controls on N₂ fixation
Modeling N\textsubscript{2} fixation is still a challenge

- Models do not balance the N cycle in the ocean or recapitulate patterns well.
- Assays of nitrogen fixation are technically difficult = variability.
- Information on distribution over time and with depth is still patchy.
- Geochemistry is not necessarily a good predictor of distribution or N\textsubscript{2} fixation.

Olson et al. 2015 *DSR II*
Host microbiome interactions

- *Trichodesmium* colonies harbor epibionts in cultures and field populations (Ruoco et al. 2016 EM)

- Quorum sensing communication molecules (acylated homoserine lactones - AHL) detected in colonies (Van Mooy et al. 2012 ISME J)

- Addition of AHLs to field colonies changes activity independent of geochemistry (Van Mooy et al. 2012 ISME J)

Image courtesy Tracy Mincer
Epibiont diversity

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

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

Data visualization and statistical analyses:
R (vegan package – Oksanen et al. 2016)
- Dissimilarity matrix
- Visualization: PCOA
- Mantel tests
- PERMANOVA (adonis function)

OTU table (.csv)

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)

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

- .fasta
- .count_table
- Green gene database
Colony composition by region

Colonies are not likely species specific, the raft morphology is more diverse except in the S. Pacific

16S amplicon sequencing indicates that *Trichodesmium* colonies harbor diverse epibionts distinct from common water column bacteria, and those found on sinking particles.
Microbiome communities significantly differ by ocean basin, and with colony morphology, except for the S. Pacific. Communities are distinct from the water column, and sinking particles.
What drives community assembly?

- **Niche?** What type of *Trichodesmium*, physiological ecology in the colony, environment..

- **Lottery?** Random selection of potential copiotrophs, role of taxonomic v. functional group uncertain...

- Working to examine the *Trichodesmium* holobiont with metagenomics/metatranscriptomics and “germ-free” *Trichodesmium*
Initial trial of “germ-free” *Trichodesmium* ran into problems - phase two scheduled for spring 2019.
Metabolic potential in the *Trichodesmium* holobiont

Each sample:
- 20 gigabytes
- ~120 million reads
- >200,000 protein coding genes
Pool reads and assemble

Bin by read coverage and tetranucleotide frequency, assess bin completeness & length

Locate ORFs, translate into proteins

Annotate proteins and assign taxonomy

Cluster binned scaffolds into orthologous groups

IDBA-UD

MaxBin

Prodigal

MCL

Metagenome Pipeline

Eel Pond mRNAseq Protocol (Titus Brown)

Map reads from each sample to de novo assembly

Assign taxonomy and annotate (and maybe sum read counts by species/group)

Find genes/OGs with differential expression between samples

Find genes/OGs with key expr. patterns

Eel Pond mRNAseq Protocol (Titus Brown)

DIAMOND-BLAST (vs NR)

MEGAN

KEGG

RSEM

Thaben and Westermark, 2014

EDGE (variance stabilize)

TMM ( NOISeq in R)

TPM

Normalise expression between different samples

ASC (no reps)

EdgeR (reps)
Composition of the holobiont

Nearly complete (65-90%) genome bins were reconstructed from a merged assembly and results are consistent with 16S data

Frischkorn et al. (2017) ISMEJ
Epibiont genome bins are detected at all stations, but the relative abundance varies.
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
Variable distribution of functional pathways among epibionts

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

- Phosphonate, heme and siderophore functions are enriched relative to water column microbes in the Sargasso Sea.
Comparing metabolic potential in the holobiont

Metagenomes → Orthologous group analysis → Epibionts v. *Trichodesmium*

Image courtesy Tracy Mincer
Epibionts confer the majority of the metabolic potential

Orthologous (OG) group analysis suggests that epibionts confer the vast majority of metabolic functions to the holobiont.

Epibiont metabolism expands Fe and P functions for the holobiont.
Epibionts can produce organic iron complexes that likely modulate iron in the holobiont microenvironment.
Uptake and metabolism of reduced phosphorus forms

Phosphorus metabolisms are present in both Trichodesmium and the microbiome.

C-P Lyase
PhoX
PhoA

Phosphite uptake
Phosphonates are produced at high rates in the holobiont - hot spot for reduced phosphorus cycling. Is it *Trichodesmium* or the epibionts?

**References**

- Van Mooy et al. (2015) *Science*
Phosphonate production is a shared metabolism

Phosphonate biosynthesis

Phosphonate produced by both *Trichodesmium* and the epibionts – at least in this environment
Microbial cross talk within the *Trichodesmium* holobiont

Tracking the interactome… with metatranscriptome profiling after addition of NO and QS molecules

QS and cell signaling
Joint metagenome and metatranscriptome analysis

- 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 phosphorus reduction and cycling in this environment?
Van Mooy et al. (2015) *Science*
Physiological ecology of *Trichodesmium* and its microbiome in the western tropical South Pacific

Moutin et al. 2017, *Biogeosci*
Phosphonates in the oligotrophic ocean
- 25% DOP is in the C-P bond class
- Recalcitrant
- Mysterious origins
- *Trichodesmium* colonies are hotspots

Metagenomic evidence of P reduction

*Trichodesmium* phosphonate biosynthesis gene cassette

McGrath et al. 2013
Dyhrman et al., 2006
Van Mooy et al., 2015
Genes are expressed with P reduction

Measure gene expression

Metatranscriptome reads mapped to the P reduction gene cassette

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
- Interactions & ecology?

\[ \text{ppm} \rightarrow \text{ppd} \rightarrow 2\text{-AEP-TA} \rightarrow \text{Me-T} \rightarrow \text{Cy-T} \rightarrow \text{GT} \]
Evidence for use of reduced phosphorus compounds in *Trichodesmium* and the microbiome

Query for genes:
- Phosphonate C-P lyase (*phnJ*)
- Phosphite dehydrogenase (*ptxD*)

Genes detected:
- *phnJ*
- *ptxD*

Genes expressed:

- Ability to access reduced P is enriched in *Trichodesmium* consortia

Frischkorn et al. (2018) *Biogeoscience*

A cryptic P currency?
Summary - Host microbiome interactions

What is the role of the microbiome in *Trichodesmium* physiological ecology?

- **16S community amplicon sequencing**: Colonies harbor diverse epibionts distinct from water column, that are dynamically curated across gradients in the environment.

- **Metagenomics**: Epibionts confer substantial metabolic potential which likely underpins *Trichodesmium* fitness.

- **Metatranscriptomics**: Distinct physiology and interactions between *Trichodesmium* and its microbiome.
Core questions

Metabolic traits and trade-offs

- What phosphorus is bioavailable?
- What are the biogeochemical constraints on $N_2$ fixation?

Host microbiome interactions

- Who is there? Microbiome diversity
- What are they doing? Microbiome functional diversity, holobiont physiology and controls on $N_2$ fixation
A strategy for assessing the physiological impact of the microbiome on *Trichodesmium* N₂ fixation

Microbiome members can communicate using a mechanism called quorum sensing.
A strategy for assessing the physiological impact of the microbiome on *Trichodesmium* N$_2$ fixation

Microbiome members can communicate using a mechanism called quorum sensing.
Microbiome members can communicate using a mechanism called quorum sensing.

AHLs found in colonies

Van Mooy et al., 2012

***Trichodesmium can’t “hear” QS molecules***

AHLs can be used to selectively modify microbiome behavior

A strategy for assessing the physiological impact of the microbiome on *Trichodesmium* N$_2$ fixation
Selectively modifying the *Trichodesmium* microbiome with AHLs in the North Atlantic

- Isolate *Trichodesmium*
- Incubate with a cocktail of AHL molecules

Does AHL addition influence microbiome gene expression?
Microbiome modulates N$_2$ fixation

The microbiome can alter Trichodesmium N$_2$ fixation.

Isolate *Trichodesmium*

Frischkorn et al. (2018) *L&O Letters*
Can the microbiome modulate host N₂ fixation?

- **Metagenome**: *Trichodesmium* can’t “hear” AHLS, while the microbiome can

- **Metatranscriptome**: AHL addition induces known quorum sensing responses and shifts host N₂ fixation

- Understanding the mechanisms that underlying interactions will help models and predictions of *Trichodesmium* distribution and N₂ fixation
16S community amplicon sequencing: Colonies harbor diverse epibionts distinct from water column, that are dynamically curated across gradients in the environment.

Metagenomics: Epibionts confer substantial metabolic potential which likely underpins *Trichodesmium* fitness.

Metatranscriptomics: Distinct physiology and interactions between *Trichodesmium* and its microbiome.
Conclusions

Genome-enabled approaches are providing new tools to trace the activities and physiological ecology of *Trichodesmium* and its microbiome.

Lessons learned: There is value in comparisons/time-series and having coincident datasets (Multi-omics)