Ocean genomics: phytoplankton physiological ecology across marine ecosystems

Sonya Dyhrman
Department of Earth and Environmental Sciences
Lamont-Doherty Earth Observatory
Columbia University
Sonya Dyhrman

PhD Scripps

Krumlov - 2016

Postdoc WHOI

Scientific Staff WHOI

Faculty Columbia University
Columbia runs RV Langseth
Sonya – random (fun?) facts
Understanding the vast network of microbes and chemicals in the ocean
High School

Postdoc

Professor
Evolving ‘omics and chemical methods allow new progress

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

Moran et al. (2022) *Nature Microbiology*
Key points for today

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

• Eukaryote physiological ecology
  • Metatranscriptomes

• Host –Microbiome 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 function
Marine Microbes - fundamental to ocean ecosystem function

- Marine microbes...
  - Produce and consume greenhouse gases
  - Supply the marine food web
  - Recycle organic matter
  - Account for roughly half of global primary production
  - *make the planet habitable*

**phyto-** = photosynthetic
**plankton** = floating organisms

Image courtesy C-MORE
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
Phytoplankton underpin ocean ecosystem function

Haptophytes

Diatoms

N₂ Fixers

Chisholm 2000 Nature
Ocean conditions are changing and this will change the distribution and activities of microbes.
The ocean acts as a buffer for CO$_2$ in the atmosphere

Oceanic reservoir is 38,000 petagrams of carbon

Sabine et al. (2004) Science
Changes in temperature $\text{CO}_2$, pH, etc.
Models predict future shifts in microbial communities

Pressing need to understand the rules that govern the ocean chemical – microbe network of interactions
The chemical – microbe network drives the fate of a lot of carbon.
What is the network of interacting microbes and chemicals?

What is the resilience of this network to perturbation?

Aspects of this network that predict regional and global patterns?
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.
Challenges and opportunities

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

- 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!
Leveraging metatranscriptome data to study the chemical – microbe network

- Cultures offer controlled conditions
- Field samples offer context
- Metatranscriptomes are species-specific
- Metatranscriptomes offer high resolution tracking of finely tuned responses
• Light and nutrients: defining the drivers of physiological ecology and competition in microeukaryotes

• Microbiome interactions: Tracking the microbiome of a keystone N₂ fixer
Thank you Dyhrman Lab family past and present

Gwenn Hennon  Mónica Rouco  Sheean Haley  Kyle Frischkorn

John Tracey  Hanna Anderson  Angela Zoumplis  Erin Maybach  Hannah Goldberg

Sheean Haley  Matt Harke  Maria Hernández-Limón  Harriet Alexander
Microeukaryote metatranscriptomes with RNA-Seq

3x20L

Filter
(>5μm)

Extract RNA

Selected
De novo assembly
Reference mapping

Unselected
De novo assembly
Reference mapping
Quality Control
- Trimmomatic
- FastQC

Assembly/Reference
- Trinity
- TransABYSS
- rnaSPADES
- MEGAHIT

Read Mapping
- Salmon
- BWA
- Bowtie2
- RSEM

Removal of Sequences
- riboPicker
- BBMap

Protein Prediction and Clustering
- TransDecoder
- OrthoFinder

Normalization & DE
- EdgeR
- DESeq2
- TPM
- TMM

Functional Annotation
- BLAST
- DIAMOND
- HMMer
- KEGG
- KOG
- GO
- PFam

Taxonomic Annotation
- BLAST
- DIAMOND
- MMETSP
- EukProt
- EukZoo
- PhyloDB

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.
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
Complex community dynamics driven in part by resources

Light

Nutrients

\( \text{Fe}, \text{Zn}, \text{B}, \text{NO}_3^- , \text{SiO}_4^{2-}, \text{PO}_4^- , \text{B}_{12}, \text{Fe}, \text{Zn} \)
Tracking phytoplankton physiological ecology

Ecosystem function and biogeochemistry

How do resources drive phytoplankton distributions and activities?
North Pacific Subtropical Gyre

Image credits: SeaWIFS Global Chlorophyll

(Karl and Church, Nature Rev. Microbiol., 2014)
Blooms of keystone species can shift ecosystem state and carbon flow.
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
Drivers of carbon flow in oligotrophic systems

North Pacific Subtropical Gyre

Station ALOHA

- Atmospheric CO$_2$
- Seawater carbonate buffer system: $\text{CO}_2 + \text{H}_2\text{O} \leftrightarrow \text{HCO}_3^- + \text{H}^+ \leftrightarrow \text{CO}_3^{2-} + \text{H}^+$
- $\text{CaCO}_3$ (in coccoliths)
- Coccolith
- Organic carbon
- Sinking flux

Coccolithophore cell
Hawaii Ocean Experiment: Dynamics of Light and Nutrients

Wilson et al. 2015 GBC

center for microbial oceanography: research and education
 linking genomes to omics

(HOE-DYLAN)
HOE: DYLAN: Water sampling
HOE: DYLAN: Semi-factorial resource additions

*In situ*  |  Control  |  Deep seawater amendment  |  (+N, Si, Fe, B$_{12}$)  |  (+P, Si, Fe, B$_{12}$)

- IS
- C
- DSW
- +N
- -P
- -N
- +P

**Concentrations:**
- **Nitrate/nitrite**: ~40 µM
- **Phosphate**: ~2.6 µM
- **Silicate**: ~100 µM
- **Fe**: ~0.6 nmol/kg
- **B$_{12}$**: <0.1 pM
Sampling and pipeline

1. Collect samples from St. ALOHA
2. Isolate mRNA
3. Submit samples to sequencing center
4. RSEM Mapping
5. DE with edgeR

- Custom reference database of *E. huxleyi* reference sequence
- Target 60 million 100 bp paired-end reads

>5 µm
6 day incubations
**E. huxleyi** physiological response suggests N control

- **N added**
  - +N v Con
  - -P v Con
  - +DSW v Con

- **No N added**
  - +P v Con
  - -N v Con

Significantly Differentially Abundant Genes

Log$_2$ Fold Change

Log$_2$ Average Abundance

Transcriptional response

Nitrogen

- N-enzymes
- N-transport

Phosphorus

- PO₄⁴⁻ Acquisition

Calcification

- Carbonate Acquisition
- Calcium Acquisition

Legend:

- Up (significant)
- Up (non-significant)
- Down (significant)
- Down (non-significant)
- Not detected

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

Nitrogen is a driver of *E. huxleyi* dynamics

**Net Heterotrophy**
- Abundant microbiome
- Low abundance
  - Keystone microbiome
- Microbial loop

**Net Autotrophy**
- Diatoms
- Nitrogen Fixers
  - *Trichodesmium*
- Haptophytes
  - Increased Export

**NITROGEN**
Emiliania huxleyi has many diverse isolates

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

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

Use diagnostic OGs to track strain distribution
Strain distribution *in situ* and in the incubations

**Percent of reads mapped to E. huxleyi**

- **CCMP1516**
  - T0: 8%
  - C: 8%
  - +P: 8%
  - -N: 8%
  - +N: 8%
  - -P: 8%
  - +D: 8%

- **CCMP379**
  - T0: 4%
  - C: 4%
  - +P: 4%
  - -N: 4%
  - +N: 4%
  - -P: 4%
  - +D: 4%

- **CCMP374**
  - T0: 0%
  - C: 0%
  - +P: 0%
  - -N: 0%
  - +N: 0%
  - -P: 0%
  - +D: 0%

- **CCMP219**
  - T0: 0%
  - C: 0%
  - +P: 0%
  - -N: 0%
  - +N: 0%
  - -P: 0%
  - +D: 0%

- **CCMP370**
  - T0: 0%
  - C: 0%
  - +P: 0%
  - -N: 0%
  - +N: 0%
  - -P: 0%
  - +D: 0%

**Treatments**

- T0: Control
- C: Control
- +P: Phosphate
- -N: Nitrogen
- +N: Nitrogen
- -P: Phosphate
- +D: Dicarboxylate

Alexander et al. unpub.
Expression of the variable portion of the pan-genome

Expression of variable gene set significantly increased following N-addition.

Alexander et al. unpub.
Future work to define the pan genome

Alexander et al. unpub.
Expanding the pan genome

~15 new whole genome sequences underway (Nanopore/Illumina)
RNASeq libraries to support gene prediction
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.
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
MAHALO: Special thanks to Daniel Muratore and Joshua Weitz for synthesis efforts
Sampling light driven responses in the NPSG
Sampling every 4hrs for many days ....
Sampling and pipeline

Isolate mRNA

Submit samples to sequencing center

Sequence using RNAseq and map reads to reference

Target 80 million 100 bp paired-end reads

RAIN – periodicity

WGCNA - network

BWA HTSeq

MMETSP ~400 transcriptomes

Genome references and other unigene databases

[MMETSP = Marine Microbial Eukaryotic Transcriptome Project]

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

Different taxa have distinct diel transcription patterns

Carbon fixation and respiration are highly coordinated

Carbon fixation and respiration are highly coordinated.


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*


Asynchronous transcription of N uptake and metabolism between microeukaryotes and between taxonomic domains suggest competition for limiting resource is driving temporal niche partitioning.
Tracing carbon flow through the network

Tracing carbon flow through the network

Multi-omics in the field: progress in resolving the microbial-chemical network

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 resource.
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