

Report Number: DOE-MBARI-04765

Title of Report: Final Report: Connecting genomic capabilities to physiology and response:
Systems biology of the widespread alga *Micromonas*

Author: Alexandra Z. Worden, Ph.D.

Abstract: Increased stratification, less mixing and reduced nutrient concentrations in marine surface waters are predicted under a number of climate-change scenarios. These conditions are considered favorable for tiny photosynthetic algae (picophytoplankton), shaping their role in mediating future CO₂ conditions. One possibility is that picophytoplankton such as *Micromonas* that have broad geographical ranges will more successfully adapt to changing environmental conditions. However, their capacity to thrive under the multi-factorial impacts of low pH, low nutrients, increasing temperature and changes in community composition is not known. Here, we developed the dual-*Micromonas* model system, which entailed generating optimized genomic information for two *Micromonas* species and developing a high-performance chemostat system in which both CO₂ and nutrients could be consistently manipulated. This system is now fully operational. Project results are available in several publications will others are still in the analysis phase. Overall, our results show that *Micromonas* primary production will likely decrease under predicted future climate conditions. Furthermore, our studies on *Micromonas* provide new insights to the land plant ancestor, including the discovery of conserved signaling mechanisms (known to be essential to plant development) as well as the discovery of widespread chemical-sensing molecular switches. Collectively, this research highlights *Micromonas* as an important new model green alga for understanding plant gene networks and evolution as well as for investigating perturbation effects on marine primary production.

Date: 30 September 2014 (modified 16 Feb 2015)

Details of the Award

Title of Project: Connecting genomic capabilities to physiology and response: Systems biology of the widespread alga *Micromonas*

Institution: Monterey Bay Aquarium Research Institute

Identification Number: DE-SC0004765, #ER64993

Program Manager: Joseph R. Gruber

Date Award Received: 8/23/10

Date of Report: 9/30/2014; **Period Covered:** 07/01/2010 – 06/30/2014 (with update of publications 16 Feb 2015)

Principal Investigators: Alexandra Z. Worden (MBARI)

Co-PIs: Stephen Callister, Richard Smith (Pacific Northwest National Labs)

Joshua Stuart (University of California Santa Cruz)

Project Goals: Develop a systems biology approach to the study of the widespread marine green alga *Micromonas* and to use this new model organism to investigate gene function and pathways and the consequences of environmental perturbations related to climate change on primary production.

Executive Summary: Over the last four years this grant has supported a major initiative to develop a unicellular algal model system that can be used to attain molecular-level mechanistic information relevant to algal and plant biology. Developing this system has entailed both technological and genomic advancements. The experimental work performed explores the intersection of higher CO₂ conditions and nutrient replete and deplete conditions. The results show that under future CO₂ levels an increase in primary production is not likely. These results contrast with the common hypotheses that primary production will increase (because CO₂ should no longer be locally limiting in bloom conditions etc.). During the course of this research we have also generated new knowledge on the ancestor of land plants, and shown that *Micromonas* has advantages as a model system for understanding gene function of 'land plant' genes because of the low gene family expansion relative to multicellular land plants. Moreover, the massive population sizes that can be achieved in the bioreactors (chemostats) developed in this project make the statistical power of data coming from *Micromonas* outstrip that of many other plant-like organisms, especially when it comes to understanding the effects of environmental perturbations on physiology and molecular responses. Finally through this research, we have associated several unknown function genes with known biosynthesis pathways. The overall project outcomes have significantly advanced understanding of the molecular-underpinnings of algal responses and how algae related to land plants will likely respond under future environmental scenarios.

Project Activities: The above accomplishments were made possible by innovative work on the technical side, the genomic side, and in developing interpretive tools for linking proteomic and transcriptomic data. On the technical side we refined a pre-existing high performance chemostat that was developed at MBARI. While this system was attractive because of the real-time measurements possible, it could not be implemented in controlled biological studies or with consistency for exploring life in a high CO₂ world. However we overcame these issues by engineering system modifications. The primary model alga in use is *Micromonas*, or more specifically, two isolates that we demonstrated are different species, and in fact show greater evolutionary distances than observed between *Oryza* and *Maize*. *Micromonas* is a widespread picoeukaryote ($\leq 2 \mu\text{m}$) that is well represented in culture. *Micromonas* belongs to the prasinophytes - unicellular algae that form a sister group to land plants alongside distantly related model green algae like *Chlamydomonas*. Prasinophytes possess characteristics thought to reflect components

of the ancestral alga that gave rise to plants and are important in marine environments from the tropics to poles. The work performed here built on the existence of sequenced genomes of *Micromonas pusilla* and *Micromonas* sp. RCC299. Our first major task for this project was to develop evidence based approaches to genome annotation. We specifically aimed to provide finished genomes (in which there are no gaps in the sequence – these are rare) and evidence-based gene model sets. The latter has been achieved by generating biologically sensitive high-throughput LC/MS proteomes for the two species (in collaboration with Pacific North West National Labs) as well as state-of-the-art transcriptional data (pair-ended directional Illumina RNA-seq, in collaboration with the JGI, along with genome gap closing). We also constructed an evidence based schema for model prioritization. The new evidence-based model sets developed in my lab, and discoveries they enabled are part of two manuscripts that will be submitted in autumn 2014 (van Baren et al. in prep., Waltman et al. in prep.). Notably, in van Baren et al, we perform comparative studies not only between the species in the original DOE proposal but also an Arctic species of *Micromonas* for which we generated transcriptome assemblies. In 2009 *Micromonas* was shown to be on the rise in the Arctic hence we feel that understanding its biology will facilitate climate change research as well as comparative biology/systems biology level knowledge of these algae.

In 2013 and 2014 we made major strides towards project goals. Among other successes, we mastered CO₂ controls in the chemostat system (and although the grant has now ended we will be performing another “high CO₂ world” experiment in January 2015 pending funding). In addition to 2013 publications in the ISME Journal and Nature our results were presented at five International Congresses, as well as the JGI Annual Meeting, the Genome Sciences Annual Contractors Meeting in Bethesda Maryland, and several invited seminars (in 2013). The two 2013 publications involved implementation of analysis approaches developed as part of this project, but in an alga that is closely related to *Micromonas*, and another that belongs to a different eukaryotic superfamily. In 2014 two publications came out in PNAS and one in the ISME J. The major findings of these papers are briefly outlined below.

Over the last several years we have focused on a light-sensing signaling protein that is essential to shade avoidance and development in land plants. Because of these important roles, phytochrome proteins have been well-studied in land plants. However were thought to have limited distributions in other eukaryotes, whereby they were present in some fungi and heterokonts, including diatoms (19), but absent from other taxa including green algae. Indeed, of the six genome-sequenced marine prasinophytes (all of which belong to the class Mamiellophyceae) only *Micromonas pusilla* contains this protein, and related model chlorophyte algae (such as *Chlamydomonas*) lack it. Through transcriptome sequencing we found that phytochromes are actually present in the majority of prasinophyte clades (6 of the 7), as well as glaucophytes and cryptophytes. Further, our analyses show that plant phytochromes have distinct origins from those in cyanobacteria, although previously cyanobacteria were hypothesized to be the ancestral source (through endosymbiotic gene transfer). Surprisingly, some of these newly detected phytochromes have spectral properties very different from plants, detecting light wavelengths that are attenuated less rapidly in aquatic environments than the red/far red light perceived by plant phytochromes. Still, the overall signaling mechanism appears to be shared with plants, with translocation to the nucleus (presumably for transcription factor interactions) during the day.

We have also advanced knowledge on vitamin biosynthesis, particularly vitamin B₁ (thiamine). Indeed, our view of vitamin control of algal blooms has been revolutionized by using the combination of: (i) pathway gap analysis (recently transformed by availability of genome-level sequence data from many taxa), (ii) the discovery of widespread chemical-sensing molecular switches (TPP riboswitches) in *Micromonas* and other algae, and (iii) hypothesis testing in *Micromonas* and other algae. In this study we recognized that many marine algae (and some marine bacteria) have nearly complete vitamin B₁ biosynthesis pathways with just one or two known enzymes missing. Typically absence of one or a few ‘essential’ enzymes is considered evidence for genome reduction in marine microbes; we hypothesized instead that at the respective pathway step different enzymes are used in place of those known from model taxa (e.g. *Saccharomyces cerevisiae*, *Escherichia coli* and *Arabidopsis thaliana*). Our experiments showed that algae missing the known enzyme for synthesizing one of the two thiamine precursor moieties could grow in the absence of exogenous thiamine supplies, although previously considered dependent on external supplies. Our sequence surveys also revealed riboswitches are widespread in eukaryotes,

primarily affiliated with genes of unknown function (which we have now linked to thiamine biosynthesis). This contrasts with the idea that riboswitches are rare in eukaryotes.

Several other aspects of this grant are still in progress although the grant term has ended. Because of the nature of our findings on life in a high CO₂ world we are reproducing experimental results multiple times and further testing how nutrient limitation influences our findings. We are also working to understand the role of repetitive introns in *Micromonas* – or at minimum how they are propagated in the genome and relate to diversification of these algae.

Comparison of goals and achievements:

Research Objective I: a model systems approach to widespread marine algae: To this end we investigated diel patterns in cell biology, expression and translation in *Micromonas*. This has been tremendously valuable because it allows us to separate what is the result of changes related to normal synchronized growth and perturbations for which it is only feasible to sample at lower time resolution (e.g. daily). RNA-seq data has been generated as has proteomic data from other conditions – all of this has now being analyzed against the new genome build and model sets. For some datasets statistical analyses and interpretation are still underway.

Research Objective II: an environmentally relevant model systems approach: This has been accomplished – and required intensive work to optimize CO₂ manipulation alongside appropriate culturing conditions and other factors. The technological aspects are complete – and could be made available to the broader community.

Research Objective III: a step towards systems biology: Pathway models have been constructed and compared between the two *Micromonas* as well as land plants (*Arabidopsis*) and the green alga *Chlamydomonas*. Optimally there would be continued iteration with RNA-seq, proteomics, homology based analyses and analyses from land plants as well as *Chlamydomonas*.

Research Objective IV: the transcriptional/translational divide: This objective was undertaken by Dr. Peter Waltman under the guidance of Co-PI Stuart. Peter developed approaches and a manuscript on findings. It is clear that data with higher time resolution is needed to sensitively address the link between transcription and translation. Very interesting results on this are also provided in Duanmu, Bachy et al. PNAS 2014.

Several factors have influenced progress on this grant. One factor was the need for maternity leave by the PI (two children in the duration of the grant) and by a lead postdoc. While this in-and-of-itself did not slow progress significantly, it did make it more difficult to identify and implement contingency plans when another partner in the project fell far short of expectations and effectively did not help with the technological aspects that defined their role in the project.

Other issues arose when a robot jammed at the sequencing facility, causing us to lose the transcriptional data for a major chemostat experiment addressing the intersection of nutrient limitation and CO₂ availability. Moreover, when we repeated the experiment all the proteomic samples were destroyed due to polymer contamination at the mass spec facility. This combination of problems at external facilities has been challenging to remedy – especially given the massive amounts of people time that go into running the bioreactor (i.e., chemostat) experiments successfully (90 sequential days minimum). Nevertheless, we do plan to perform the major experiment that was compromised once more – so that we can have paired RNA-seq and proteomic data, but have not yet determined how to cover salary for the postdocs who would be needed to make this happen (the talent is available in house, but allocated to other projects since the end of this grant).

Products Delivered:

Published or in review manuscripts (IF, journal impact factor)

1. Ševčíková T, Horák A, Klimeš V, Zbráňková V, Demir-Hilton E, Sudek S, Jenkins J, Schmutz J, Přibyl P, Fousek J, Vlček C, Lang BF, Oborník M, Worden AZ & M Eliáš (in review). Updating algal evolutionary relationships through plastid genome sequencing: did alveolate plastids emerge through endosymbiosis of an ochrophyte?
2. Simmons MP, Bachy C, Sudek S, van Baren MJ, Ares M & AZ Worden (in revision). Intron invasions trace algal speciation and reveal nearly identical Arctic and Antarctic *Micromonas* populations. *Molecular Biology and Evolution*. IF 14.308
3. Duanmu[‡] D, Bachy[‡] C, Sudek S, Wong CH, Jimenez V, Rockwell NC, Martin SS, Ngan CY, Reistetter E, van Baren MJ, Price DC, Wei CL, Reyes-Prieto A, Lagarias JC & AZ Worden (2014). Marine algae and land plants share conserved phytochrome signaling systems. *Proceedings of the National Academy of Sciences USA*. doi: 10.1073/pnas.1416751111. [‡]Co-first authors. IF 9.809
4. McRose[‡] D, Guo[‡] J, Monier[‡] A, Sudek S, Wilken S, Yan S, Mock T, Archibald JM, Begley TP, Reyes-Prieto A & AZ Worden (2014). Alternatives to Vitamin B₁ uptake revealed with discovery of riboswitches in multiple marine unicellular eukaryotes. *The ISME Journal* doi:10.1038/ismej.2014.146. [‡]Co-first authors. IF 9.267
5. Rockwell, NC, Duanmu D, Martin SS, Bachy C, Price DC, Bhattacharya D, Worden AZ & JC Lagarias (2014). Eukaryotic algal phytochromes span the visible spectrum. *Proceedings of the National Academy of Sciences USA* Vol. 11:3871-3876. IF 9.809

Please also see <http://www.pnas.org/content/early/2015/02/06/1501871112.extract>

for a correction of the printer's faulty modification of grant attribution.

6. Monier A, Sudek S, Fast N & AZ Worden (2013). Gene invasion in distant eukaryotic lineages: discovery of mutually exclusive genetic elements reveal marine biodiversity. *The ISME Journal* Vol. 7:1764–1774. IF 9.267
7. Worden AZ^{*}, Janouskovec J, McRose D, Engman A, Welsh RM, Malfatti S, Tringe SG & PJ Keeling (2012). Global distribution of a wild alga revealed by targeted metagenomics. *Current Biology* Vol. 22:R675-R677. *Corresponding author. IF 10.227

This publication uses tools and pipelines developed for the analysis to be performed in the DOE systems biology grant and therefore acknowledges the grant since the work would not have been possible without these tools.

In preparation manuscripts (others are anticipated, those listed here should be submitted within the 2015 year)

Guo J, Sudek L, Reistetter E, Wilken S, Purvine S, Ansong C, Callister S, Limardo AJ, Klimov D & AZ Worden. Algal proteome remodeling in response to phosphate limitation (tentative title, in prep). For submission to TBD journal.

van Baren MJ, Bachy C, Reistetter E, Sudek S, Yu H, Purvine SO, Ansong C, Callister SJ, Wei C-L, Wong C-H, Grimwood J, Schmutz J & AZ Worden (in prep). Prasinophyte genomes reveal marine diversity and ancestral characteristics of plants despite small size. For submission to Genome Biology.

Cuvelier ML, van Baren MJ, Ortiz A & AZ Worden (in prep). Photoacclimation and adaptation in the picoeukaryotic prasinophyte *Micromonas*. For submission to Plant Biology.

Conference posters and talks

Invited (Worden, 2013-2014 only)

- 15th International Symposium on Microbial Ecology, Discovery of a photosensory signaling protein widespread in algae. Seoul, KOR, 2015. With post-doc Dr. Charles Bachy.
- Keynote, Congress of the International Union of Microbiological Societies, Phytoplankton diversity and the global climate. Montreal, CAN, 2015. With post-doc Dr. Susanne Wilken.
- Joint Aquatic Sciences Meeting (ASLO/PSA/SFS/SWS), Nitrogen utilization in photosynthetic picoeukaryotes. Portland, OR, 2015.
- European Molecular Biology Organization (EMBO) Conference on Comparative Genomics of Eukaryotic Microorganisms, Environmental genomics in pursuit of evolutionary & ecological dynamics. Sant Feliu de Guixols, Spain, 2014.
- Keynote, Annual Meeting of the Environmental Mutagenesis and Genomics Society. Ocean 'Omics: Understanding phytoplankton and the global carbon cycle. Monterey, CA, 2014.
- International Congress of Protistology XIV, Putting meta'omics to work in phytoplankton ecology. Vancouver, BC CAN, 2014.
- Stazione Zoologica di Napoli, Putting picoeukaryote genomes to work in microbial ecology. Napoli, IT, 2014.
- Integrated Microbial Diversity Meeting, Putting picoeukaryote genomes to work. Whistler, BC CAN, 2014.
- JGI Genomic Technologies Workshop/Genomics of Energy and the Environment Meeting, Gaining insight to marine photosynthesis and the plant lineage using omic technologies. Walnut Creek, CA, 2014.
- DOE Genomic Science Meeting, Connecting genomes to physiology & response in marine photosynthetic eukaryotes. Bethesda, MD, 2014.

General (all project years)

2014 Bachy C & AZ Worden. Invited talk: Discovery of a photosensory signaling protein widespread in algae. 15th International Symposium on Microbial Ecology. Seoul, Korea, 24-29 Aug.

2014 Guo J, Wilken S, Sudek L, Reistetter E, Purvine S, Ansong C, Callister S, Limardo AJ, Klimov D & AZ Worden. Poster: Responses of the widespread green alga *Micromonas* to elevated CO₂ at different levels of phosphorus availability. CIFAR Integrated Microbial Biodiversity Annual Meeting, Liblice, Czech Republic, 25-29 June.

2014 Jimenez V, Wei C-L, Ramamoorthy G, Bachy C, Wong C-H & AZ Worden. Poster: Response of the marine picoeukaryote *Micromonas* to nitrogen deprivation. CIFAR Integrated Microbial Biodiversity Program Meeting, June 25-29, 2014, Liblice, Czech Republic.

2014 Bachy C, Jiménez V, Sudek S, Lagarias JC & AZ Worden. Poster: Phytochromes in widespread marine algae reveal origins of plant signaling proteins. ASLO/AGU Ocean Sciences Meeting. Honolulu, HI, 23-28 Feb.

2014 Jimenez V & AZ Worden. Talk: Nitrogen metabolism in green algae: a case study using *Micromonas*. ASLO/AGU Ocean Sciences Meeting, Honolulu, HI, 23-28 Feb.

2014 Bachy C, Duanmu D, Sudek S, Jimenez V, van Baren MJ, Wong C-H, Rockwell NC, Martin SS, Ngan CY, Reistetter E, Wei C-L, Reyes-Prieto A, Callister S, Lagarias JC & AZ Worden. Poster: Phytochromes in widespread photosynthetic algae reveal origins of plant signaling proteins. Genomic Science Contractors-Grantees Meeting XII. Arlington, VA, 9-12 Feb.

2014 Guo J, Wilken S, Sudek L, Reistetter E, Purvine S, Ansong C, Callister S, Limardo AJ, Klimov D, Kolber Z & AZ Worden. Poster: Responses of the widespread green alga *Micromonas* to elevated CO₂ at different levels of phosphorus availability. Genomic Science Contractors-Grantees Meeting XII. Arlington, VA, 9-12 Feb.

2014 Waltman P, van Baren MJ, Reistetter E, Purvine S, Ansong C, Callister S, Worden AZ & J Stuart. Examining the post-transcriptional program governing the metabolic proteome of *Micromonas pusilla*. Genomic Science Contractors-Grantees Meeting XII. Arlington, VA, 9-12 Feb.

2014 Lagarias JC, Duanmu D, Bachy C, Rockwell NC, Jimenez V, Wong C-H, Sudek S, Martin SS, Ngan CY, Reistetter E, Price DC, van Baren MJ, Bhattacharya D, Reyes-Prieto A, Wei C-L & AZ Worden. Talk: Bilin signaling in the green lineage. 23rd Western Photosynthesis Conference. Asilomar, CA, 2-5 Jan.

2013 Simmons MP & AZ Worden. Invited talk: Population-specific algal transcripts and qPCR pinpoint novel marine ecotypes and biogeographic expression of unknown function genes. International Congress of Protistology XIV, Vancouver B.C. CAN 28 Jul - 2 Aug.

2013 Jimenez V, Wei CL, Wong CH, Ngan CY & AZ Worden. Poster: Nitrogen metabolism in plants and green algae: a case study using *Micromonas*. Integrated Microbial Diversity Meeting. Whistler, B.C. CAN 14-18 May.

2013 Jimenez V, Wei CL, Wong CH, Ngan CY & AZ Worden. Poster: Nitrogen metabolism in plants and green algae: a case study using *Micromonas*. Genomics of Energy and the Environment Meeting. Walnut Creek, CA 26-29 Mar.

2013 van Baren MJ, Reistetter E, Wong CH, Ngan CY, Purvine P, Ansong C, Grimwood J, Schmutz J, Callister S, Wei CL & AZ Worden. Evidence based approaches to the *Micromonas* genomes provide insights to marine algae and the plant lineage. Genomics of Energy and the Environment Meeting. Walnut Creek, CA 26-29 Mar.

2013 van Baren MJ, Waltman P, Sam D, Reistetter E, Purvine S, Ansong C, Callister S, Stuart J & AZ Worden. Poster: Systems biology of marine green algae: insights from the dual *Micromonas* model system. DOE Genomic Science Meeting. Bethesda, MD, 24-27 Feb.

2012 McRose D & AZ Worden. Talk: Elucidating the role of vitamins in phytoplankton ecology. ISME14, Copenhagen, DK 19-24 Aug.

2012 van Baren MJ, Waltman P, Sam D, Reistetter E, Purvine S, Ansong C, Callister C, J Stuart & AZ Worden. Computational and experimental approaches to systems biology of marine green algae. Genomic Science Awardee Meeting X. Bethesda, MD, 26-29 Feb.

2012 McRose D, Yan S, Reistetter E & AZ Worden. Vitamin biosynthesis and regulation in marine eukaryotic algae. Genomic Science Awardee Meeting X. Bethesda, MD, 26-29 Feb.

2011 Simmons MP & AZ Worden. Distribution of novel introns in the marine eukaryotic green alga *Micromonas*. Poster, 111th General Meeting American Society of Microbiology, New Orleans, LA 21-24 May.

2011 McRose D & AZ Worden. Growth and gene expression responses of two *Micromonas* species to thiamine deprivation suggests disparate metabolic capabilities. Poster, 111th General Meeting American Society of Microbiology, New Orleans, LA 21-24 May.

Other products delivered

Evidence-based gene model sets for *Micromonas* sp. RCC299 and *Micromonas pusilla* CCMP1545. These new gene calls have been provided to JGI and integrated in their browser system.

Ph.D. granted to Melinda P. Simmons, UC Santa Cruz

Masters granted to Darcy McRose, Stanford University
