**Okinawa Institute of Science and Technology
Graduate University**

Thesis submitted for the degree
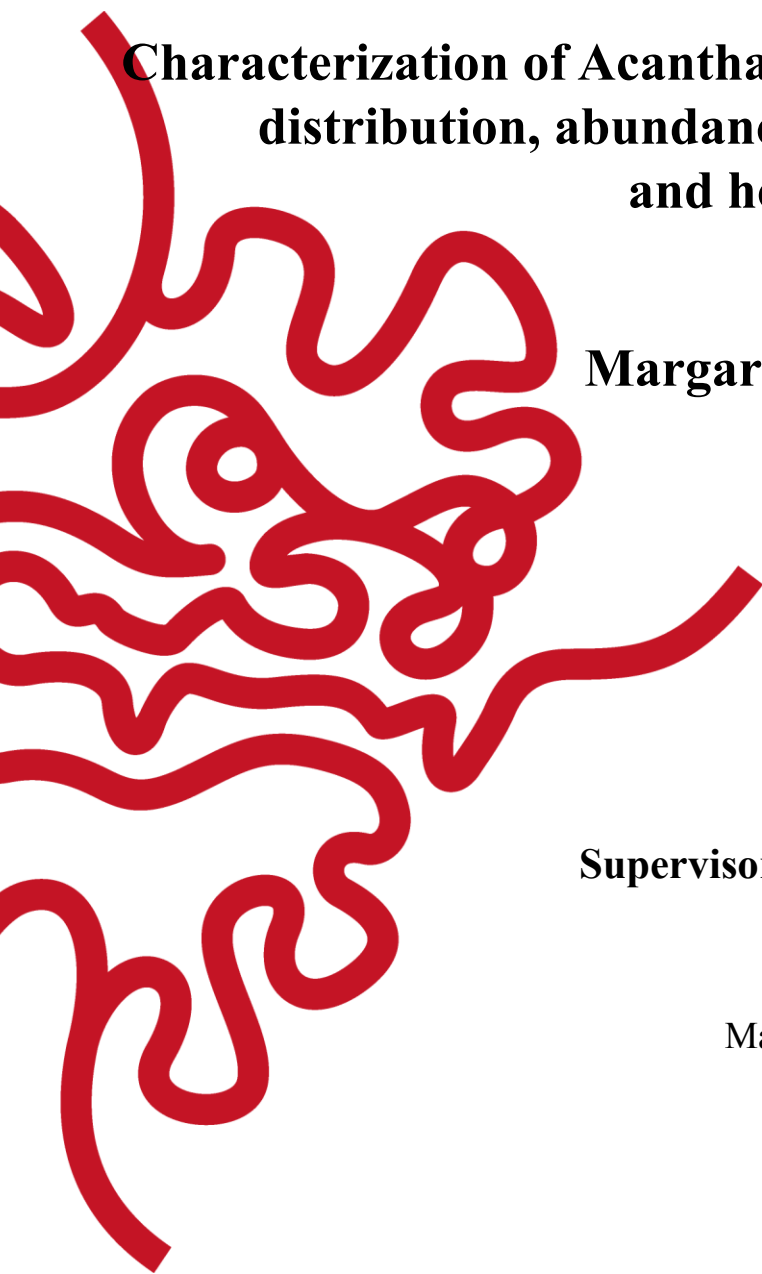
Doctor of Philosophy

Characterization of Acantharea-*Phaeocystis* photosymbioses: distribution, abundance, specificity, maintenance and host-control

by

Margaret Mars Brisbin

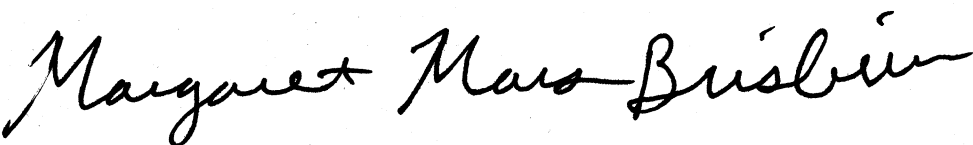**Supervisor: Satoshi Mitarai**

May 20, 2020

Declaration of Original and Sole Authorship

I, Margaret Mars Brisbin, declare that this thesis entitled "Characterization of Acantharea-Phaeocystis photosymbioses: distribution, abundance, specificity, maintenance and host-control" and the data presented in it are original and my own work.

I confirm that:

- This work was done solely while a candidate for the research degree at the Okinawa Institute of Science and Technology Graduate University, Japan.
- No part of this work has previously been submitted for a degree at this or any other university.
- References to the work of others have been clearly attributed. Quotations from the work of others have been clearly indicated and attributed to them.
- In cases where others have contributed to part of this work, such contribution has been clearly acknowledged and distinguished from my own work.
- None of this work has been previously published elsewhere, with the exception of the following: Chapter Two is published as "Intra-host Symbiont Diversity and Extended Symbiont Maintenance in Photosymbiotic Acantharea (Clade F)" in Frontiers in Microbiology (2018). (If the work of any co-authors appears in this thesis, authorization such as a release or signed waiver from all affected co-authors must be obtained prior to publishing the thesis. If so, attach copies of this authorization to your initial and final submitted versions, as a separate document for retention by the Graduate School, and indicate on this page that such authorization has been obtained).

Signature:

Date: February 13th, 2020

Thesis Abstract

Microbial eukaryotes (protists) are important contributors to marine biogeochemistry and play essential roles as both producers and consumers in marine ecosystems. Among protists, mixotrophs—those that use both heterotrophy and autotrophy to meet their energy requirements—are especially important to primary production in low-nutrient regions. Acantharian protists (clades E & F) accomplish mixotrophy by hosting *Phaeocystis spp.* as algal endosymbionts and are extremely abundant in subtropical low-nutrient regions where they form productivity hotspots. Despite their ecological importance, acantharians remain understudied due to their structural fragility and inability to survive in culture. In order to overcome these challenges and illuminate key aspects of acantharian biology and ecology—including distribution, abundance, and specificity and specialization of symbioses—single-cell RNA sequencing methods were developed for acantharians and used alongside environmental metabarcode sequencing and high-throughput, in-situ imaging. Major findings from this thesis were that i) acantharian cell (> 250 µm) concentrations decrease with depth, which correlates to patterns in relative sequence abundances for acantharian clades with known morphologies but not for those lacking known morphology, and that ii) while individual acantharians simultaneously harbor multiple symbiont species, intra-host symbiont communities do not match environmental communities, providing evidence for multiple uptake events but against continuous symbiont turnover, and that iii) photosynthesis genes are upregulated in symbiotic *Phaeocystis*, reflecting enhanced productivity in symbiosis, but DNA replication and cell-cycle genes are downregulated, demonstrating that hosts suppress symbiont cell division. Moreover, storage carbohydrate and lipid biosynthesis and metabolism genes are downregulated in symbiotic *Phaeocystis*, suggesting fixed carbon is relinquished to acantharian hosts. Gene expression patterns indicate that symbiotic *Phaeocystis* is not nutrient limited and likely benefits from host-supplied ammonium and urea, thus providing evidence for nutrient transfer between hosts and symbionts. Interestingly, genes associated with protein kinase signaling pathways that promote cell proliferation are downregulated in symbiotic *Phaeocystis*. Deactivation of these genes may prevent symbionts from overgrowing hosts and therefore represents a key component of maintaining the symbiosis. This research contributes new insights into the ecologically relevant photosymbioses between Acantharea and *Phaeocystis* and illustrates the benefits of combining single-cell sequencing and imaging technologies to illuminate important microbial relationships in marine ecosystems.

Acknowledgments

First, I want to acknowledge the support and guidance that my thesis advisor, Satoshi Mitarai, has provided throughout the last five years. Satoshi has given me an incredible gift by trusting me, allowing me to make independent research decisions, and respecting choices I made. As a result, I feel especially prepared to embark on a research career following my time at OIST. I also owe so much to Sasha Mikheyev, who introduced me to next-generation sequencing and data analysis, patiently tutored me in bioinformatics and inspired me to keep learning; I am extremely grateful for his continued mentorship. Mary Grossmann reintroduced me to the Acantharea at a pivotal moment, and ultimately inspired this entire thesis; her continued input has been invaluable. This thesis was further inspired by the works of Dave Caron and Johan Decelle, both of whom have offered generous advice and allowed me to visit their labs to learn from them and discuss acantharian biology. I must also thank past and present members of the OIST Marine Biophysics Unit and the OIST Ecology and Evolution Unit—especially Patricia Wepfer, Angela Ares Pita, Jo Tan, Mandy Tin, Lijun Qiu, Miguel Grau-Lopez, Carmen Emborski, and Valentin Churavy—who each played an essential role in the completion of this thesis by engaging in scientific discussion, reading and editing proposal and manuscript drafts, troubleshooting software installation and data analysis issues, helping with protocol optimization, preparing samples for sequencing, rubber ducking, giving feedback, and much more.

I would also like to thank the captain and crew of the R/V *Mirai* for their assistance and support in sample collection. Hiroyuki Yamamoto, Hiromi Watanabe, Dhugal Lindsay, Mary Grossmann and Yuko Hasagawa were instrumental in organizing and facilitating cruise sampling. Dhugal Lindsay, Andrew Carroll, Mehul Sangekar and Otis Brunner deployed the imaging system used to collect data for this thesis and Otis Brunner additionally assisted in filtering seawater samples. I am extremely grateful to Hiroki Goto for graciously providing sequencing advice and to the OIST sequencing section for performing sequencing. I would further like to acknowledge the OIST imaging section for providing microscopy training and maintaining instruments. This thesis was funded by the Okinawa Institute of Science and Technology Graduate University and the Japan Society for the Promotion of Science.

Lastly, and most importantly, I need to thank my whole family for their unwavering love and encouragement, but especially Kathleen Cummings, Jamie Brisbin, and Gnasher Brisbin. My mom, Kathleen Cummings, instilled an early love of science in me and fostered my curiosity and wonder for the natural world with weekly visits to the New York Aquarium, the Bronx Zoo, The American Museum of Natural History, and the Queens Science Museum (among others), despite being a single mother. She always made sure that I knew women can be scientists, too, and let me dream big. My husband, Jamie Brisbin, has given me confidence by believing in me, keeps

me grounded, always reminds me that there is more to life than work, and helps me enjoy it. And Gnasher, the sweetest and most intelligent and loving of all canine companions, makes sure that I take breaks for walks and sunshine and constantly provides reasons to smile.

Table of Contents

List of Figures and Tables

Figures

Tables

Thesis Introduction

Protists—microbial, unicellular eukaryotes—are important contributors to marine

biogeochemistry and play essential roles as both producers and consumers in marine ecosystems

(Sherr et al., 2007). Among protists, mixotrophs—or those that use both heterotrophy and

autotrophy to satisfy their energy requirements—are especially important to primary production

in low-nutrient regions where nutrient availability otherwise limits primary production (Mitra et

al., 2014; Worden et al., 2015). Acantharian protists in molecular clades E and F (Arthracanthida

and Symphyacanthida) accomplish mixotrophy by hosting predominantly *Phaeocystis* species as

algal endosymbionts (Decelle et al., 2012a; Mars Brisbin et al., 2018). These photosymbiotic

holobionts, encompassing hosts and multiple algal partners, are abundant in subtropical,

low-nutrient regions and create localized primary production hotspots (Michaels, 1991; Caron et

al., 1995). Despite their ecological importance, acantharians remain understudied due to their

structural fragility and inability to survive in culture (Michaels, 1988; Decelle et al., 2012a), and

as a result, major gaps exist in our understanding of acantharian biology. For example, when and

where photosymbiotic acantharians reproduce is only hypothesized (Decelle et al., 2013), as is

the fate of their symbionts when they reproduce (Decelle et al., 2012a; Mars Brisbin et al., 2018).

This thesis, therefore, combines multiple culture-free approaches with non-destructive sampling

in order to illuminate key aspects of acantharian biology and ecology: Chapter One evaluates

acantharian abundance and vertical distribution in the western North Pacific and considers results

in the context of acantharian life-history traits, Chapter Two investigates intracellular symbiont

diversity in acantharians and examines questions regarding symbiont uptake and turnover, and

Chapter Three investigates changes in symbiont gene expression compared to free-living cells and proposes mechanisms for host control and nutrient exchange within the symbiosis.

Typical acantharian morphologies include exquisite star-shaped skeletons, composed of strontium sulfate, that are embedded within amoeboid cells. While beautiful, acantharian structures are so delicate that they are often crushed or broken in plankton nets. Sampling acantharians is further complicated by the chemical composition of their skeletons: strontium sulfate dissolves in seawater and many traditional fixatives, including formalin (Michaels, 1988). As a result, acantharian contributions to plankton net and sediment trap samples are often underestimated (Michaels, 1988; Michaels et al., 1995). In fact, our current appreciation for acantharian abundance and primary production in low-nutrient surface waters rests almost completely on a series of studies performed in the late 1980's and early to mid 90's that used targeted methods to quantify acantharian abundances in the North Atlantic Subtropical Gyre, the North Pacific Subtropical Gyre, and the equatorial Pacific (Michaels, 1988, 1991; Caron et al., 1995; Michaels et al., 1995; Stoecker et al., 1996). Since then, acantharian contributions to plankton communities have primarily been documented using metabarcoding—amplifying and sequencing a small variable region within a conserved gene from a mixed, environmental sample to determine which organisms were present in the sample and their approximate relative abundances. This methodology has many advantages: it is easy, relatively inexpensive, and very fast. Unfortunately, because the resulting data are compositional (i.e. the abundance measured for a single group is inherently affected by the abundances of all other taxa) , they can never offer insight into actual abundances of different organisms (Gloor et al., 2017). As meta 'omics methods have become increasingly popular, the more tedious work of carefully sampling and

counting cells has fallen by the wayside and quantifying acantharians has not been revisited in the last decades, nor have new regions been surveyed.

Chapter One of this thesis takes a modern approach to quantify acantharian abundances in the western North Pacific; a high-throughput, in-situ plankton imaging system was deployed to image acantharians in their natural state and determine cell concentrations at an extremely fine vertical resolution. Metabarcoding was performed alongside imaging to allow for comparisons between results from the two methods and to ultimately provide deeper insights into acantharian biology, particularly in regards to morphology and life-history. In-situ plankton imaging has already revealed high abundances of larger photosymbiotic protists around the globe, but the smaller size of most acantharians has so far precluded using this method to fully evaluate acantharian abundances (Dennett et al., 2002; Biard et al., 2016). The imaging system used to collect data for this thesis is able to image cells with diameters greater than 250 μm, which aligns well with typical acantharian sizes. Results demonstrated that current acantharian cell concentrations in surface waters of the western North Pacific are quite similar to the concentrations previously reported from the eastern N. Pacific, equatorial Pacific, and N. Atlantic, indicating sustained acantharian abundances over space and time. Furthermore, imaging results correlated well with metabarcoding results for acantharians with known morphologies. In contrast, metabarcoding results for basal acantharian clades defined only by environmental sequences and without any known morphology did not correlate with abundances calculated from imaging. These results suggest that basal acantharian clades may not possess characteristic star-shaped skeletons, supporting predictions based on evolutionary relationships (Decelle et al., 2012c), or they may be too small to be caught on camera. Moreover, the observed variation in

cell size with depth aligns well with current hypotheses regarding asymbiotic acantharian cell-cycles: adult cells or cysts sink quickly to release reproductive swarmers at depth and juveniles slowly rise to the surface as they grow in size (Decelle et al., 2013). Previously, photosymbiotic acantharians were hypothesized to complete this cycle in the photic zone since they cannot form cysts and rely heavily on photosynthesis (Decelle et al., 2013), but results presented in Chapter One suggest photosymbiotic acantharians may also make the journey to the deep sea to reproduce. Lastly, the in-situ, non-destructive imaging allowed for a new acantharian behavior to be documented, which may represent a previously undescribed predation strategy among acantharians.

Photosymbioses, like those between acantharians and their algal symbionts, are generally seen as mutualistic, based on the assumption that symbionts benefit from nitrogen and phosphorus supplied by hosts while simultaneously providing hosts with photosynthetically derived fixed organic carbon. It is unknown, however, whether *Phaeocystis* cells symbiotic to acantharians retain reproductive capacity, nor if they can escape or are ever released from symbiosis, thus calling into question whether symbionts can truly benefit from the relationship (Decelle, 2013). Photosymbiotic hosts require more symbionts to meet their metabolic needs as they increase in size, and, indeed, larger acantharians have more symbionts than smaller acantharians (Michaels, 1991). At the same time, hosts must manage symbiont populations to ensure symbionts do not overgrow them. Hosts can limit symbiont population size by systematically digesting or releasing symbionts, or by preventing symbiont cell division (Boettcher et al., 1996; Titlyanov et al., 1996; Fishman et al., 2008). Hosts that manage symbiont populations by digesting and releasing symbionts can afford to allow symbionts to divide within

host tissue, but if hosts regulate symbiont populations by preventing symbiont division, they must collect additional symbionts from the environment to meet their metabolic needs. Chapter Two explores these possibilities by assessing intra-host symbiont diversity within individual acantharian hosts through single-holobiont metabarcoding and by using fluorescent confocal microscopy to assess whether symbionts are systematically digested. Results demonstrated that acantharians simultaneously host multiple strains, species, and genera of symbionts, indicating that rather than starting with one symbiont that multiplies as the host grows, hosts continually recruit new symbionts from the environment. Furthermore, fluorescent imaging revealed that symbionts are not found within phagolysosomes and that lysosomes are not concentrated near symbionts, demonstrating that hosts do not regularly digest symbionts. Together, these results suggest that symbionts are maintained within hosts, but are probably not dividing, providing new evidence against mutualism in this relationship.

Given that symbionts are maintained and not digested by hosts (Chapter Two), it follows that hosts should manipulate symbiont cell division to regulate symbiont population size. Chapter Three of this thesis uses transcriptome analyses to investigate molecular mechanisms involved in host control of symbionts and nutrient transfer between acantharians and *Phaeocystis*, and further explores the question of mutualism versus exploitation in acantharian photosymbioses. Moreover, endosymbiosis was fundamental to the evolution of current cellular complexity in extant eukaryotes (Archibald, 2015). Elucidating molecular processes involved in photosymbioses, which can represent early intermediary stages of chloroplast acquisition, will, therefore, lead to a better understanding of steps involved in the evolution of secondary and tertiary chloroplasts among photosynthetic eukaryotes (Keeling, 2004; Archibald, 2015).

Since acantharians collected from the same place can host distinct symbiont communities, as was shown in Chapter Two, it was necessary to perform single-holobiont RNA extractions in order to compare symbiotic gene expression to appropriate reference transcriptomes (as opposed to pooling holobionts). Unfortunately, the small size of acantharians, combined with the chemical makeup of their skeletons—divalent cations like those formed by strontium can interfere with nucleic acid extraction kit chemistry—meant that established methods for single-cell RNA extractions were not effective. Instead, methods for high-throughput single-cell sequencing were adapted and optimized for use with single acantharians (Trombetta et al., 2014). Once working, these methods eventually allowed for single-holobiont transcriptomes to be sequenced from sixteen individual acantharians.

Symbiotic gene expression within individual acantharian holobionts was compared to expression measured in three biological replicates each of two symbiont species grown in standard culture conditions: *Phaeocystis cordata* and *Phaeocystis jahnii*. Overall, symbiosis-associated gene expression differed slightly between the two *Phaeocystis* species, but there were several striking similarities. First, photosynthesis genes were significantly upregulated in symbiosis for both species, confirming that symbionts actively photosynthesize *in hospite*. Second, genes involved in DNA replication and cell-cycle progression were significantly downregulated in symbiosis for both species, indicating that cell division is inhibited *in hospite*. These results are consequential because they suggest a mechanism for symbiont population control in Acantharea-*Phaeocystis* photosymbioses—symbionts are prevented from dividing within host cells. Nuclear encoded chloroplast division genes were expressed at similar levels in symbiosis as in free-living cells, indicating that chloroplast division continues even though the

cell-cycle is inhibited. This is consistent with observations of altered phenotypes in symbiotic

Phaeocystis—symbiotic cells are larger than free-living cells and have many additional

chloroplasts (Febvre and Febvre-Chevalier, 1979; Decelle et al., 2019). Expression patterns of

nitrogen metabolism genes differed for the two species in symbiosis, but in both cases, results

point to hosts providing ammonium and urea, and neither species seems to be nitrogen or

phosphorus limited in symbiosis. Instead of limiting symbiont access to nutrients to suppress cell

division, hosts may manipulate cell signaling pathways that transduce extracellular signals to the

nucleus and influence cell-proliferation. By keeping symbionts well-fed and utilizing

cell-signaling to maintain symbiont population sizes, hosts simultaneously maximize

photosynthetic output and prevent symbiont overgrowth. While it cannot be determined from

these data if the transformation of symbionts is reversible, it seems unlikely that symbionts can

recover from such extreme remodeling, which would make the symbiosis an evolutionary dead

end for *Phaeocystis*, and, therefore, makes mutualism improbable.

Chapter One: Acantharian abundance and vertical distribution illuminated through paired high-throughput imaging and sequencing

Abstract

Acantharians are important contributors to surface primary production and to carbon flux to the deep sea, but are often underestimated because their delicate structures are destroyed by plankton nets or dissolved by preservatives. As a result, relatively little is known about acantharian biology, especially regarding their life-cycles. This study takes a paired approach, bringing together high-throughput, in-situ imaging and high-throughput sequencing to investigate acantharian abundance, vertical distribution, and life-history in the western North Pacific. Observed concentrations of acantharian cells correlated well with sequence abundances from acantharians with known, recognizable morphologies, but not to sequences from those without known morphology (basal environmental clades). These results suggest basal clades may lack characteristic star-shaped skeletons or are much smaller than known acantharians. The decreased size-range of acantharians imaged at depth supports current hypotheses regarding asymbiotic acantharian life cycles: cysts or vegetative cells release reproductive swarmer cells at depth and juvenile cells grow as they ascend towards the surface. Moreover, sequencing data present the possibility that photosymbiotic acantharians also reproduce at depth, like their asymbiotic, encysting relatives, which is counter to previous hypotheses. Finally, in-situ imaging captured a new acantharian behavior that may be a previously undescribed predation strategy.

1.1. Introduction

Acantharians are important contributors to primary production in surface waters and to carbon flux to the deep sea (Michaels, 1991; Michaels et al., 1995; Decelle et al., 2013; Belcher et al., 2018). Acantharians in molecular clades E and F (Arthracanthida and Symphyacanthida), which include the majority of described acantharian species, host algal endosymbionts from the haptophyte genus *Phaeocystis* (Decelle et al., 2012a; Mars Brisbin et al., 2018) that exhibit elevated photosynthetic efficiency when living symbiotically (Decelle et al., 2019). The acantharian skeleton is composed of strontium sulfate, the densest known organic biomineral, causing acantharians to sink quickly after death (Decelle et al., 2013). In addition, their amoeboid cell structure with sticky cellular extensions (pseudopodia) predisposes sinking acantharians to form aggregates, further enhancing sinking rate (Gutierrez-Rodriguez et al., 2019). The biogeochemical significance of acantharians has been historically underestimated, however, because traditional sampling methods often miss acantharians; plankton nets destroy delicate acantharian cell structures (Michaels, 1988) and common preservatives dissolve acantharian skeletons (Bernstein et al., 1992). DNA sequencing surveys have revealed that acantharians account for large numbers of sequences from the water column and sediment traps in diverse ecosystems, including tropical and subtropical regions (Fontanez et al., 2015; Hu et al., 2018), polar regions (Martin et al., 2010; Decelle et al., 2013), and productive temperate coastal regions (Countway et al., 2010; Gutierrez-Rodriguez et al., 2019). However, the relationship between DNA sequence abundance and acantharian biomass or flux is not clear and is complicated by acantharians being multinucleated and having multiple life stages, including encystment and reproductive swarmer production.

DNA metabarcoding—sequencing a region of the small-subunit ribosomal RNA gene for an entire community—has been extensively applied to estimating microbial community structure (e.g. de Vargas et al., 2015; Pernice et al., 2016). While this method has undoubtedly revolutionized our understanding of microbial diversity in different ecosystems, it has several significant limitations. First, metabarcoding and other meta 'omics produce compositional data, meaning that the abundance of any single group is inherently influenced by the abundance of other groups (Gloor et al., 2017). This issue is further complicated by the varying nucleus and gene-copy numbers among protist groups—some organisms, like dinoflagellates, have many copies of ribosomal RNA genes and may be more represented in sequencing datasets as a result (Gong et al., 2013; Gong and Marchetti, 2019). Gene-copy and nucleus number is especially problematic because it precludes the possibility of extrapolating absolute abundances from sequences and total cell counts—if the cell abundance to sequence abundance ratio was consistent, absolute abundance could be determined by multiplying the relative sequence abundance with the total cell count in a sample (Gong and Marchetti, 2019). The second major limitation of DNA metabarcoding is that DNA can persist after a cell dies and, therefore, does not reflect metabolic state. As a result, it is unknowable whether DNA sequences derive from actively metabolizing cells, dormant cells or cysts, reproductive cells, or dead cells and detritus (Torti et al., 2015). This is particularly relevant in evaluating acantharian abundances and relative contributions to biogeochemical cycles since acantharian vegetative cells, reproductive cells, and cysts will differentially contribute to photosynthesis, grazing/predation, and carbon flux (Decelle et al., 2013).

While not yet as widely adopted as molecular methods, high-throughput, in-situ imaging systems are being used to quantitatively assess abundances of marine microbes and other components of the plankton (Dennett et al., 2002; Grossmann et al., 2015; Biard et al., 2016). Such imaging systems can drastically improve the spatial resolution of sampling and process much larger volumes of water than can be included in DNA surveys. Furthermore, imaging cells where they naturally occur and in their native orientation can reveal previously undescribed behaviors and associations (Möller et al., 2012; Greer et al., 2013; Peacock et al., 2014). Analyzing data from high-throughput imaging, however, is still challenging; processing images and creating training sets for use with machine learning algorithms requires expertise in plankton taxonomy and is time-intensive (Orenstein et al., 2015). The taxonomic resolution attainable with a particular imaging system depends on image size and quality, but will almost always be less than is possible with molecular methods. Furthermore, taxonomic resolution will vary for different taxonomic groups and will be higher for those with more defined morphological features and lower for organisms, like flagellates, that lack identifying features (Sieracki et al., 2010). Finally, a single imaging system cannot image the entire size-range of marine plankton, necessitating multiple systems to holistically characterize plankton communities (Lombard et al., 2019). Vegetative acantharian cells, with their characteristic star-shaped skeletons, are particularly amenable to imaging surveys (Biard et al., 2016), but distinguishing acantharian reproductive cells or cysts with high-throughput, in-situ imaging may not be possible.

DNA metabarcoding and high-throughput, in-situ imaging both have benefits and drawbacks as methods for assessing plankton abundance and community structure. By applying these methods together, this study aimed to better characterize acantharian abundance,

water-column distribution, and life-history. The BellaMare In-Situ Ichthyoplankton Imaging

System (ISIIS) small-imager/area-scanner (Cowen and Guigand, 2008;

www.planktonimaging.com/smaller-imagers) was deployed at four sites along the Ryukyu

Archipelago in the western North Pacific. Replicate water samples for DNA sequencing were

also collected from the surface, deep chlorophyll maximum, middle water column, and about 10

m above the seafloor from each site where imaging was performed and from 10 additional sites

along the Ryukyu Archipelago. Water samples were sequentially size-fractionated in an effort to

separate acantharian vegetative cells from reproductive swarmers. Relative abundance of

acantharian sequences in the larger size fraction were compared to cell counts from imaging

profiles to assess the relationship between acantharian relative sequence abundance and cell

abundance. Metabarcoding results were further analyzed to evaluate the taxonomic distribution

of acantharians by depth in the western North Pacific and results from size fractionation were

considered in the context of hypothesized acantharian life-cycles.

1.2. Materials and Methods

1.2.1. Sampling locations

Water samples for DNA sequencing were collected from 14 sites spanning the length of

the Ryukyu Archipelago during the Japan Agency for Marine-Earth Science and Technology

(JAMSTEC) MR17-03C cruise from May 29 to June 13, 2017 (Figure 1.1). The JAMSTEC

DEEP TOW 6KCTD system, a towable frame outfitted with several imaging systems and a

Conductivity-Temperature-Depth (CTD) sensor, was additionally deployed to a maximum of

1,000 m at four of the sampling sites (3, 10, 15, and 17), to take vertical profiles of plankton

images (Figure 1.2).
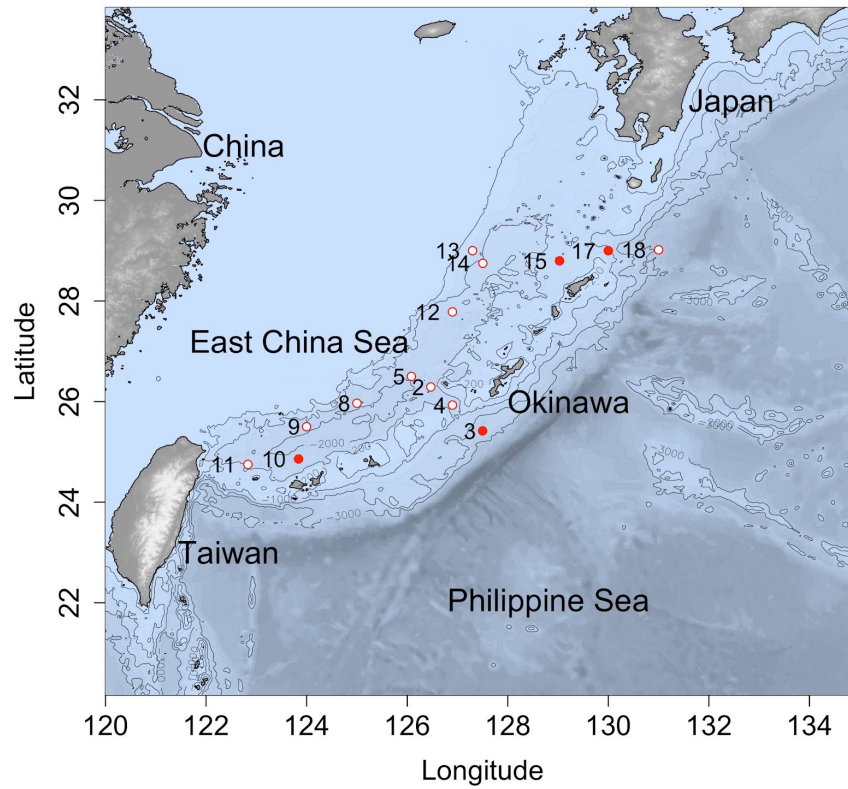
Figure 1.1. Map of sampling station locations. Stations where water samples for DNA analysis were collected and high-throughput imaging was performed are marked with closed red circles. Stations, where only water samples for DNA analysis were collected, are marked with open red circles.

A.

B.
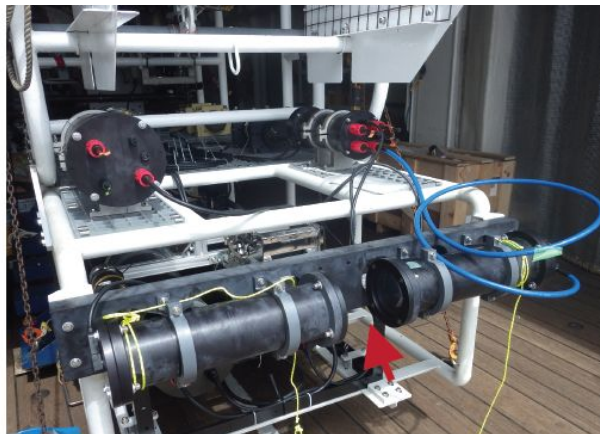
Figure 1.2. Photos of the JAMSTEC DEEP TOW 6KCTD (A) and BellaMare ISIIS small imager/area scanner attached to the DEEP TOW (B). Red arrows indicate the position of the ISIIS system on the DEEP TOW (A) and the imaging area of the ISIIS system (B). At each sampling station, the DEEP TOW was lowered straight down through the water column, to a maximum depth of 1000 m, before being towed at 1000 m and then continuing to be towed as it was raised back up through the water column. Only images from downward casts were used for this study, since forward motion of the frame prevented plankton from being properly imaged by the ISIIS.

1.2.2. Image acquisition and processing

An ISIIS small imager/area-scanner (BellaMare, San Diego, CA) was attached to the

DEEP TOW to collect vertical profiles of plankton images (Figure 1.2). The camera system was

set to image organisms > 250 μm, which aligns well with the size of vegetative acantharian cells.

The ISIIS camera was programmed to take 1 photo per second coinciding with an LED flash.

Each photo imaged 0.39 L (st. 3 and 10) or 0.35 L (st. 15 and 17) parcels of water in 2448 x 2050

pixel resolution, with each pixel being 22.5 x 22.5 μm. Because the ISIIS camera was attached to

the back of the DEEP TOW, only photos taken during the down-cast were considered in this

study because the forward motion of the DEEP TOW during the up-cast could interfere with

plankton moving naturally through the imaging area of the camera. A total of 4,010 photos were

taken during the down-cast at station 3; 3,639 at station 10; 3,056 at station 15; and 2,453 at

station 17, so that 13,158 photos were included in the study—an equivalent of 4,931.5 liters of

seawater. Down-cast photos were manually viewed by a single researcher and Regions of

Interest (ROIs) containing characteristically star-shaped acantharian vegetative cells were

cropped and saved. The ISIIS internal clock was calibrated to match that of a Sea-Bird SBE 9

CTD (Sea-Bird Scientific, Bellevue, WA) mounted to the DEEP TOW so that CTD data could

be used to determine the depth at which each image was taken. Concentrations of acantharian

cells per liter were determined by normalizing acantharian cell counts to the total number of

images taken for 10 m bins and correlation between cell concentration and depth was evaluated.

ROI image area was used as a proxy for cell size, allowing for comparisons in cell-size range

between sites and depths. Figure 1.3 illustrates the morphological diversity and size range of

acantharians imaged in this study. Acantharian ROIs and all raw images used in the study are

archived on Zenodo (https://doi.org/10.5281/zenodo.3605400).

273 µm

2583 µm
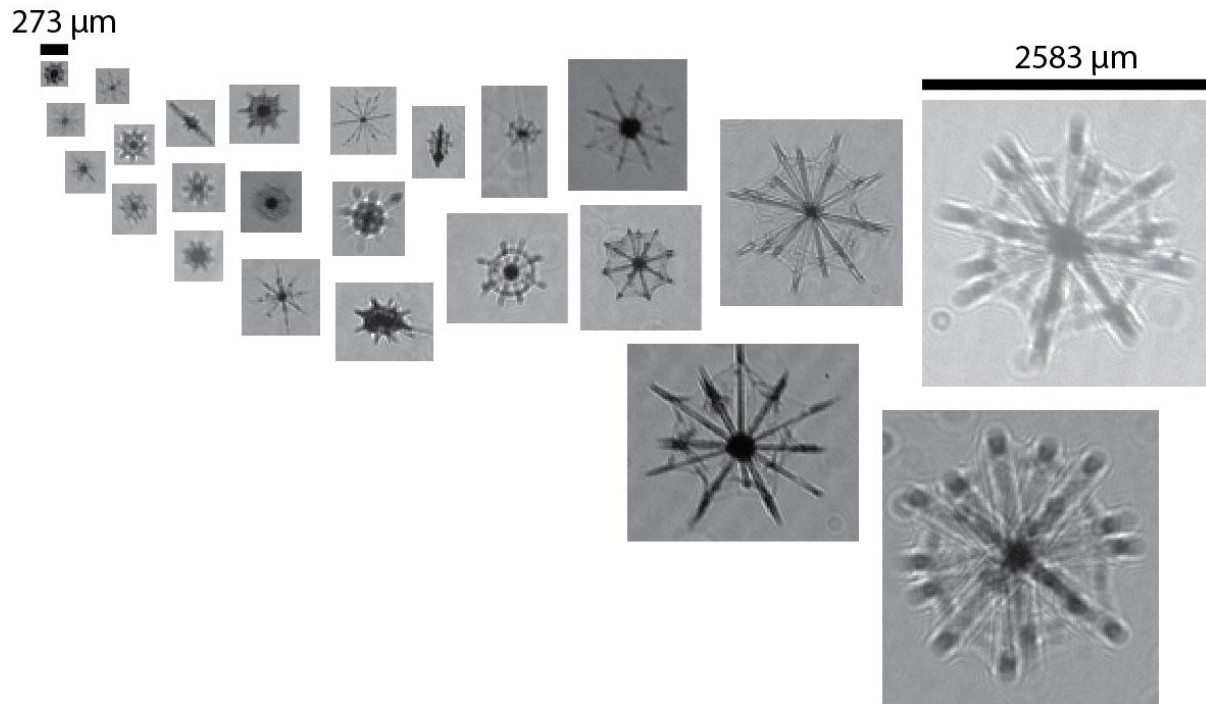
Figure 1.3. Acantharians imaged in this study, illustrating the morphological diversity and size range of imaged cells.

1.2.3. Water sampling and DNA extraction

A Niskin bottle rosette with 30 bottles (10 L) and fitted with a CTD probe (Sea-Bird SBE 911plus) was deployed at each cruise station to collect water from the deep chlorophyll maximum (DCM), the middle water column (mid), and approximately 10–20 m above the seafloor (bottom) (Table 1.1). Surface seawater was collected by bucket alongside the research vessel. Two replicates of 4.5 liters (surface) from separate bucket casts or 5 liters from separate Niskin bottles (DCM, mid, bottom) were sequentially filtered under a gentle vacuum through 10.0-µm and 0.2-µm pore-size polytetrafluoroethylene (PTFE) filters (Millipore, Burlington, MA). Sequential size filtering was implemented in order to separate vegetative acantharian cells and cysts from reproductive swarmer cells (< 5 µm, Decelle et al., 2012a), although complete

separation is probably not possible. Filters were flash-frozen in liquid nitrogen and stored at −80°C.

DNA was extracted from PTFE filters (n = 224, two replicates of two filter pore-sizes at four depths from 14 stations) following manufacturer's protocols for the DNeasy PowerWater Kit (Qiagen, The Netherlands) including the optional heating step for 10 min at 65°C to fully lyse cells. Sequencing libraries were prepared following the Illumina 16S Metagenomic Sequencing Library Preparation manual, but with universal eukaryotic primers for the V4 region of the eukaryotic 18S rRNA gene (F: CCAGCASCYGCGGTAATTCC, Stoeck et al., 2010; R: ACTTTCGTTCTTGATYR, Mars Brisbin et al., 2018) and 58°C annealing temperature in the initial PCR. Amplicon libraries were sequenced by the Okinawa Institute of Science and Technology (OIST) DNA Sequencing Section on the Illumina MiSeq platform with 2 x 300-bp v3 chemistry. Amplification and sequencing were successful for 211 samples, with at least one replicate for each sample type.

Table 1.1. Coordinates, sampling depths, and total depth for all sampling stations.

Station	Longitude (°E)	Latitude (°N)	"DCM" Depth (m)	"Mid" Depth (m)	"Bottom" Depth (m)	Site Depth (m)
2	126.468	26.290	92	700	1066	1080
3	127.500	25.416	72	1000	2385	2407
4	126.900	25.928	58	700	1834	1857
5	126.084	26.501	82	700	1900	1922
8	124.994	25.942	88	700	1671	1681
9	124.012	25.502	74	700	1890	1914
10	123.837	24.857	85	700	1515	1530
11	122.840	24.760	50	700	1217	1223
12	126.903	27.785	80	700	1024	1033
13	127.339	29.003	80	700	834	846

14	127.501	28.747	75	700	1013	1025
15	129.029	28.792	67	700	776	783
17	129.570	28.957	90	700	772	779
18	130.904	28.981	100	1500	2957	2981

1.2.4. Sequence analysis

Sequence data from each of four flow-cells were denoised separately using the Divisive Amplicon Denoising Algorithm (Callahan et al., 2016a) through the DADA2 plug-in for QIIME 2 (Bolyen et al., 2019). Denoised Amplicon Sequence Variant (ASV) tables were merged before taxonomy was assigned to ASVs with a naive Bayes classifier trained on the Protist Ribosomal Reference (PR^2) database (Guillou et al., 2013) using the QIIME 2 feature-classifier plug-in (Bokulich et al., 2018). Results were imported into the R statistical environment (R Core Team, 2018) for further processing with the Bioconductor package phyloseq (McMurdie and Holmes, 2013). Full protist communities (including all eukaryotic ASVs, except those classified as Metazoa) were analyzed first to evaluate to what degree overall community composition varied by sampling depth and by filter pore-size at each depth. Sequences classified as Acantharea were further analyzed separately to determine (i) if patterns by depth and filter pore-size for acantharians reflected overall community patterns, (ii) how much acantharian sequences contributed to the total number of sequences from each depth, (iii) how the relative abundance of different acantharian clades varied by depth, and (iv) if the acantharian contribution to total sequence numbers correlated to cell concentrations determined from imaging data. The data and code necessary to reproduce all statistical analyses are available on GitHub

(https://github.com/maggimars/Acanth_ImageSeq), including an interactive online document:

https://maggimars.github.io/Acanth_ImageSeq/Acanth_ImageSeq_Analysis.html.

1.3. Results

1.3.1. Sequencing results

Overall, 31.5 million sequencing reads were generated for this study, with

34,631–421,992 sequencing reads per sample (mean = 144,604). All raw sequence data is

available from the NCBI Sequencing Read Archive, accession number PRJNA546472.

Following denoising, 16.8 million sequences remained and 1.1 million were classified as

Acantharea. We identified 1,053 unique acantharian ASVs in our dataset, out of a total of 22,656

unique ASVs.

In Principal Coordinates Analyses (PCoA) of Aitchison distances between samples based

on full protist community compositions, samples clustered by depth first, with clear separation of

surface and DCM samples from mid and bottom water samples on the primary axis; DCM and

surface samples further separated from each other on the secondary axis (Figure 1.4A). When

full protist communities were analyzed for each depth separately, surface and DCM samples

separated by filter-pore size on the primary axis and mid and bottom water samples separated by

filter-pore size on the secondary axis (Figure 1.5A), but these results were not found to be

statistically significant with Permutational Analyses of Variance (PERMANOVA).

Notwithstanding, the clear sample clustering by filter pore-size for each depth suggests that

size-fractionation was moderately successful. It remains likely, however, that larger cells were

broken or otherwise squeezed through the larger pore-size filter to be captured on the lower

filter, and that some smaller cells were stuck and retained on the larger pore-size filter.

When only ASVs classified as Acantharea were included in PCoA based on Aitchison

distances, samples also clustered by depth, but the overall pattern was distinct from that seen

when full protist communities were analyzed. Acantharian communities varied more in mid and

bottom water samples than full protist communities did (Figure 1.4B). Furthermore, while the

full protist communities clustered separately by filter pore-size in each depth layer (Figure 1.5A),

this was not true for acantharian communities, which did not separate by filter pore-size at any

depth (Figure 1.5B).
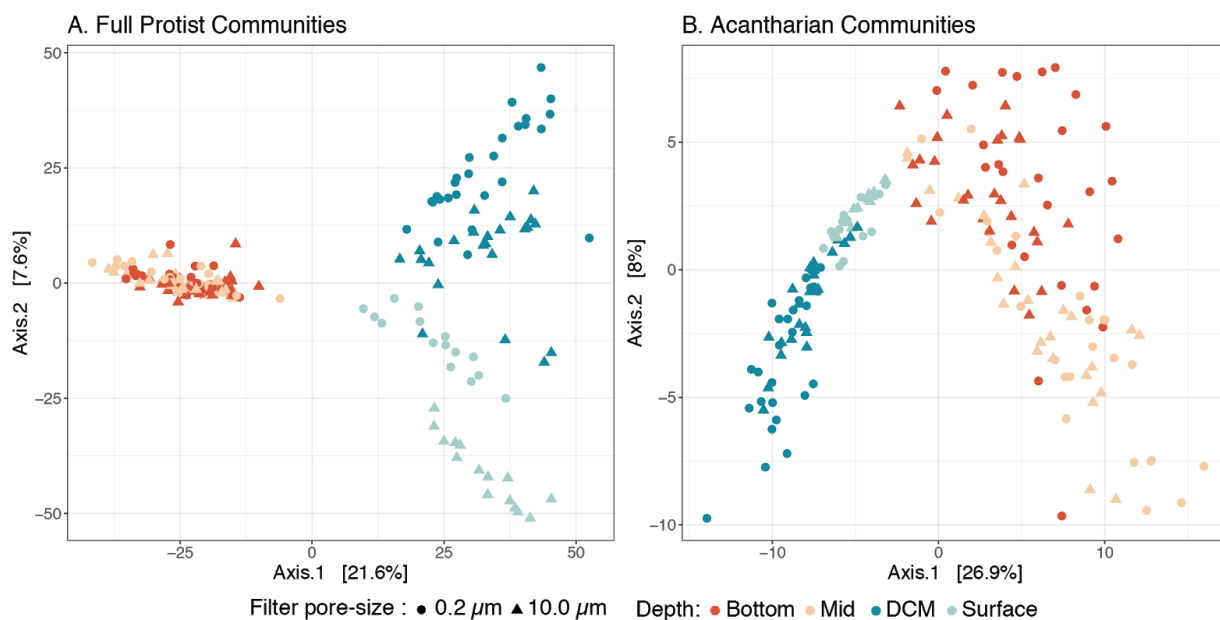
Figure 1.4. Principal coordinates analysis of Aitchison distance between full protist community compositions (A) and acantharian community compositions (B). Full protist communities (A) include all denoised sequences that were classified as Eukaryota, but not Metazoa. Acantharian communities (B) include all denoised sequences classified as Acantharea in the 4th taxonomic level of the PR^2 database (i.e. class). Color indicates the depth layer from which samples were collected and shape reflects the filter pore-size used to collect samples in μm. Full protist communities form three main clusters by depth—surface, deep chlorophyll maximum (DCM), and mid/bottom (A)—whereas the acantharian communities are more varied in the mid and near-bottom waters (B).

A. Full Protist Communities

B. Acantharian Communities

Filter pore-size: ● 10.0 *μm* ● 0.2 *μm*
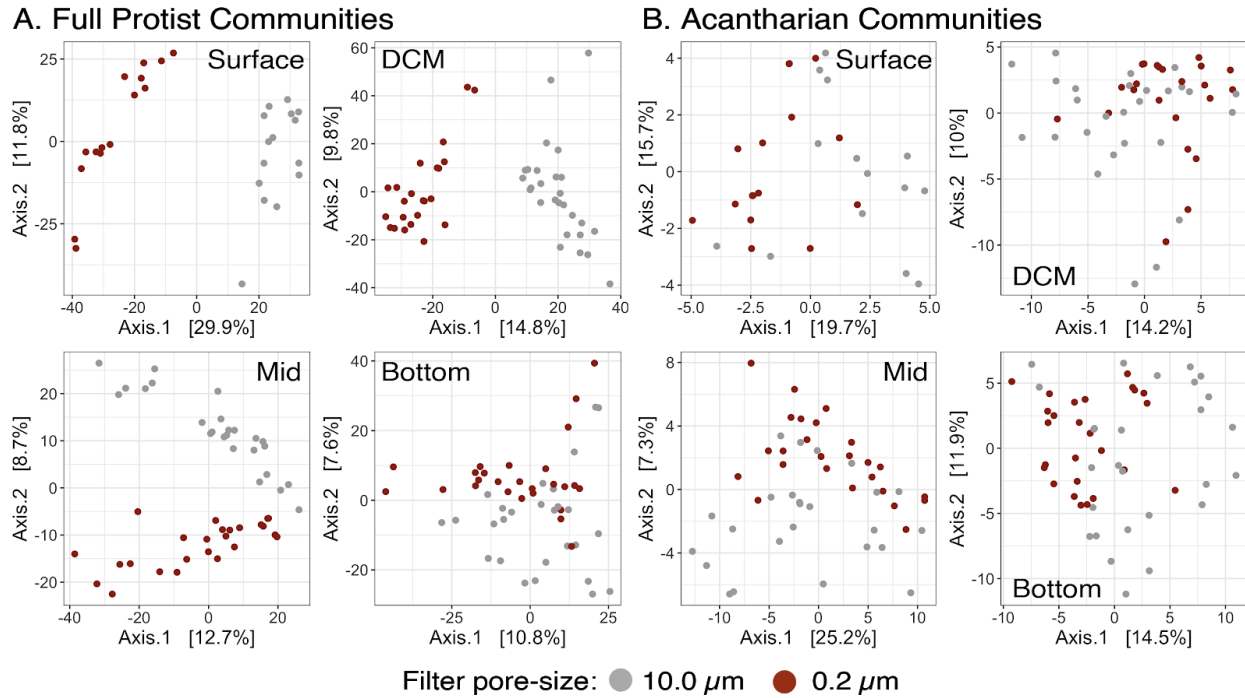
Figure 1.5. Principal coordinates analysis of Aitchison distances between full protist community compositions (A) and acantharian community compositions (B) in different depth layers. Separating samples by depth layer allows for better resolution of the effect of filter pore-size on community composition. Each panel represented results for a given depth layer and point color reflects the filter pore-size used to collect samples in μm. Full protist communities (A) include all denoised sequences that were classified as Eukaryota, but not Metazoa. Acantharian communities (B) include all denoised sequences classified as Acantharea in the 4th taxonomic level of the PR[2] database (i.e. taxonomic class). Full protist communities cluster by filter pore-size along the primary axis at the surface and DCM and along the secondary axis in the mid water-column (A); acantharian communities do not cluster by filter pore-size at any sampling depth (B).

At the surface, Arthracanthida and Symphyacanthida acantharians made up almost the entire acantharian community in both large and small size-fraction samples at every sampling station (Figure 1.6). Arthracanthida and Symphyacanthida are the most recently diverging acantharian clades (molecular clades E and F); acantharians belonging to these clades are photosymbiotic and have robust skeletons that are sometimes ornamented with elaborate appendages (Decelle et al., 2012c). In the DCM, the contribution of sequences deriving from

Chaunacanthida acantharians increased, as did sequences deriving from Acantharea-Group-II (Figure 1.6). The Chaunacanthinda clade diverged earlier than both Arthracanthida and Symphyacanthida clades. Chaunacanthida acantharians are generally asymbiotic, have less developed skeletons, and are capable of encystment (Decelle et al., 2012c, 2013). The Acantharea-Group-II is one of several basal clades that are defined entirely by sequences recovered from environmental samples and have no known morphology (Decelle et al., 2012c). In the mid and near-bottom water, the majority of the acantharian sequences derived from another basal environmental clade, the Acantharea-Group-I (Figure 1.6).
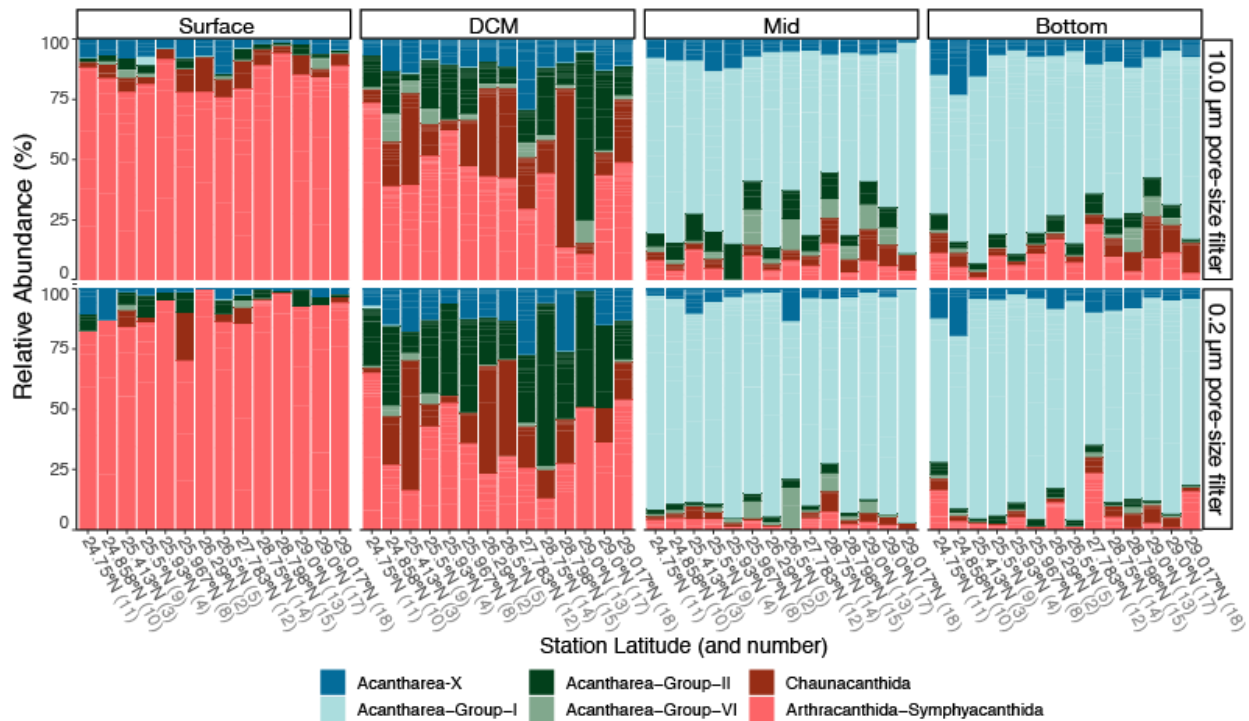
Figure 1.6. Relative abundance of acantharian groups in size-fractionated samples from four depths in the western North Pacific. Sampling stations are ordered on the x-axis from south to north and the plot is faceted by sampling depth (columns) and filter pore-size in μm (rows). When replicates were available for a particle station/depth/filter combination, replicates were collapsed and represented with a single stacked bar. Stacked bars are divided to represent the contribution of major acantharian clades to acantharian communities: dark blue (Acantharea-X) represents sequences that were not classified past the class level (Acantharea). The acantharian communities at the surface are dominated by Symphyacanthida and Arthracanthida acantharians, while at the DCM, Symphyacanthida and Arthracanthida

acantharians still make up a large proportion of reads but communities are more diverse. Acantharian communities are dominated by the basal environmental clade Acantharea-Group-I in the mid and near-bottom waters.

       In order to evaluate the relationship between depth and acantharian sequence abundance, the percent contribution of acantharian ASVs to all sequences was calculated for each sample and a linear regression model was fit to the percentages with depth as the independent variable. The linear model fit to acantharian sequence percentages in all samples demonstrated a significant positive correlation to depth ($R^2 = 0.25$, $p < 0.001$, Figure 1.7A). To facilitate comparisons between sequencing and imaging data, linear models were also fit individually to acantharian sequence percentages from stations where imaging profiles were taken. Increasing acantharian sequence percentage with depth was apparent for three of the four stations with imaging profiles, but model results were only significant for station 17 ($R^2 = 0.73$, $p < 0.01$, Figure 1.7B).
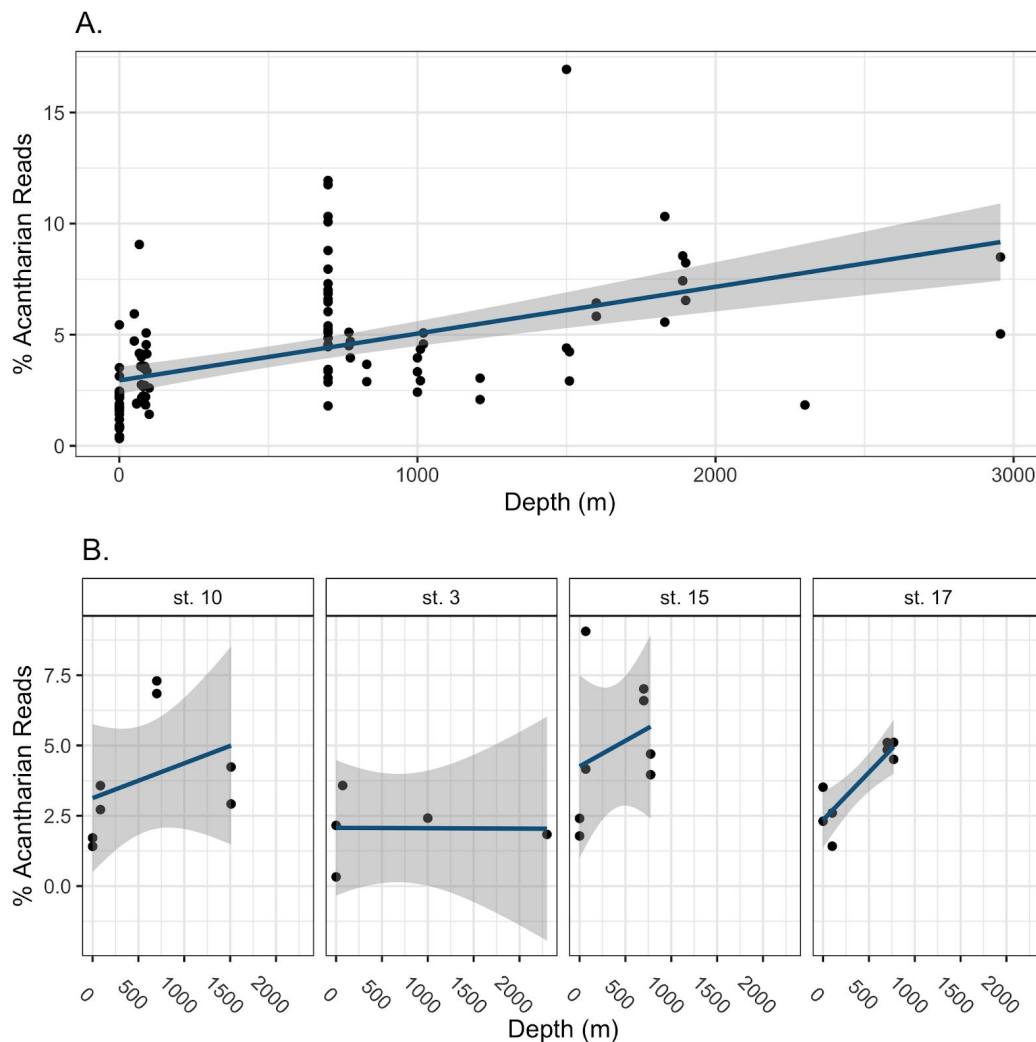
Figure 1.7. Percentage of sequences deriving from acantharians in water samples from all stations (A) and the four stations where plankton imaging was performed (B). Sequence percentages refer to the proportion of reads designated as acantharian out of all denoised sequences for a given sample, including metazoan sequences and sequences without a taxonomic classification. Linear models were fit to the data and are represented by blue lines with the 95% confidence intervals shaded grey. The linear model results for all samples (A) are statistically significant ($p < 0.001$) with an R^2 of 0.25. The linear model results for individual stations (B) were only statistically significant for station 17 ($p < 0.01$), which had an R^2 of 0.76. Overall, the contribution of acantharian sequences was smallest in the surface waters and tended to increase with depth sampled.

Only acantharians in the Chaunacanthida, Arthracanthida, and Symphyacanthida clades are definitively known to possess the characteristic star-shape used to identify acantharian

vegetative cells in images from this study. Therefore, the sequence percentages of

Chaunacanthida, Arthracanthida, and Symphyacanthida were further considered separately.

When only sequence percentages for these clades with known morphologies were included in a

linear model, sequence abundance was significantly negatively correlated with depth ($R^2 = 0.17$,

$p < 0.001$, Figure 1.8A)—opposite to the relationship when all clades were analyzed together.

The same trend was apparent for all stations with imaging profiles when evaluated

individually—percentage decreased with depth—but model results were only statistically

significant for station 10, with $R^2 = 0.65$ and $p < 0.01$ (Figure 1.8B).
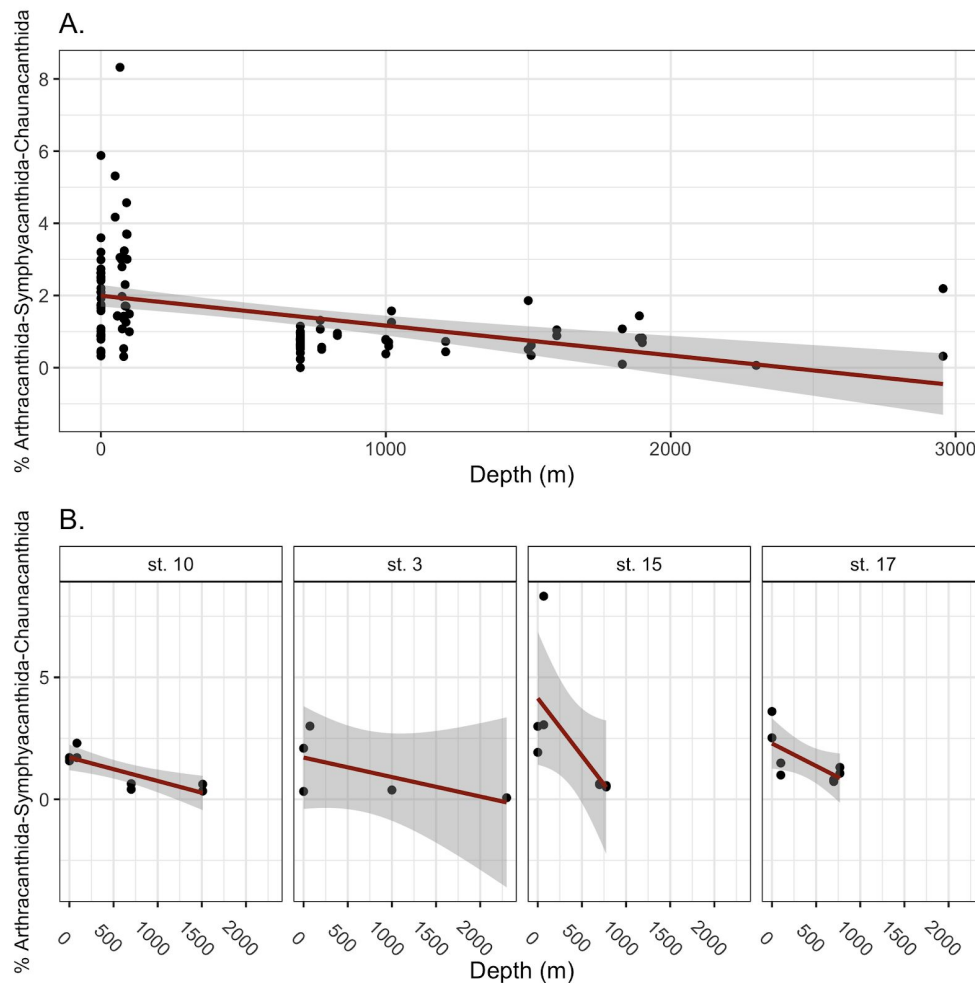
Figure 1.8. Percentage of sequences deriving from Symphyacanthida, Arthracanthida, and Chaunacanthida acantharians in water samples from all stations (A) and the four

stations where imaging profiles were performed (B). Percentages refer to the proportion of sequences designated as Symphyacanthida, Arthracanthida, or Chaunacanthida acantharians out of all denoised sequences for a given sample, including metazoan sequences and sequences without a taxonomic classification. Linear models were fit to the data and are represented by red lines with 95% confidence intervals shaded grey. The linear model results for all samples (A) are statistically significant ($p < 0.001$) with an R^2 of 0.17. The linear models for individual stations (B) were not statistically significant (R^2 = 0.27–0.65). Overall, the contribution of Symphyacanthida, Arthracanthida, and Chaunacanthida acantharians to the whole community was larger in surface waters and at the DCM and tended to be smaller in the mid and near-bottom waters.

1.3.2. Imaging results

Overall, 1,235 acantharian ROIs were identified in the study and the vast majority of these were found in images taken close to the sea surface (Figure 1.9A). When linear models were fit to log-transformed cell concentrations calculated from images, with depth as the independent variable, cell concentrations were significantly negatively correlated with depth at each station (R^2 = 0.63–0.78, $p < 0.001$, Figure 1.9A). These results match the relationship between Chaunacanthida, Arthracanthida and Symphyacanthida sequence percentages with depth but not the relationship between all acantharian sequence percentages with depth. To directly evaluate how well imaging results correlated with sequencing results, we averaged acantharian cell concentrations in each depth layer (surface: 0–50 m, DCM: 50–150 m, mid: 150–700 m, deep/bottom: > 700 m) and compared these values to Chaunacanthida, Arthracanthida and Symphyacanthida sequence percentages in samples from corresponding stations and depths. Averaged cell concentrations significantly positively correlated with Chaunacanthida, Arthracanthida and Symphyacanthida sequence abundance (R^2 = 0.33, $p < 0.05$) following exclusion of two outlying data points with exceptionally high sequence abundance or imaged cell concentration (Figure 1.10). Acantharian cell size ranged widely in the surface

waters and near the DCM, whereas the range was more constrained and cell size was generally

smaller in deeper water (Figure 1.9B). Interestingly, many acantharians were observed with long

pseudopodial extensions terminating in drop-shaped structures (Figure 1.11). This morphology

has not previously been observed in acantharians, probably due to damage caused by plankton

nets or other handling effects.
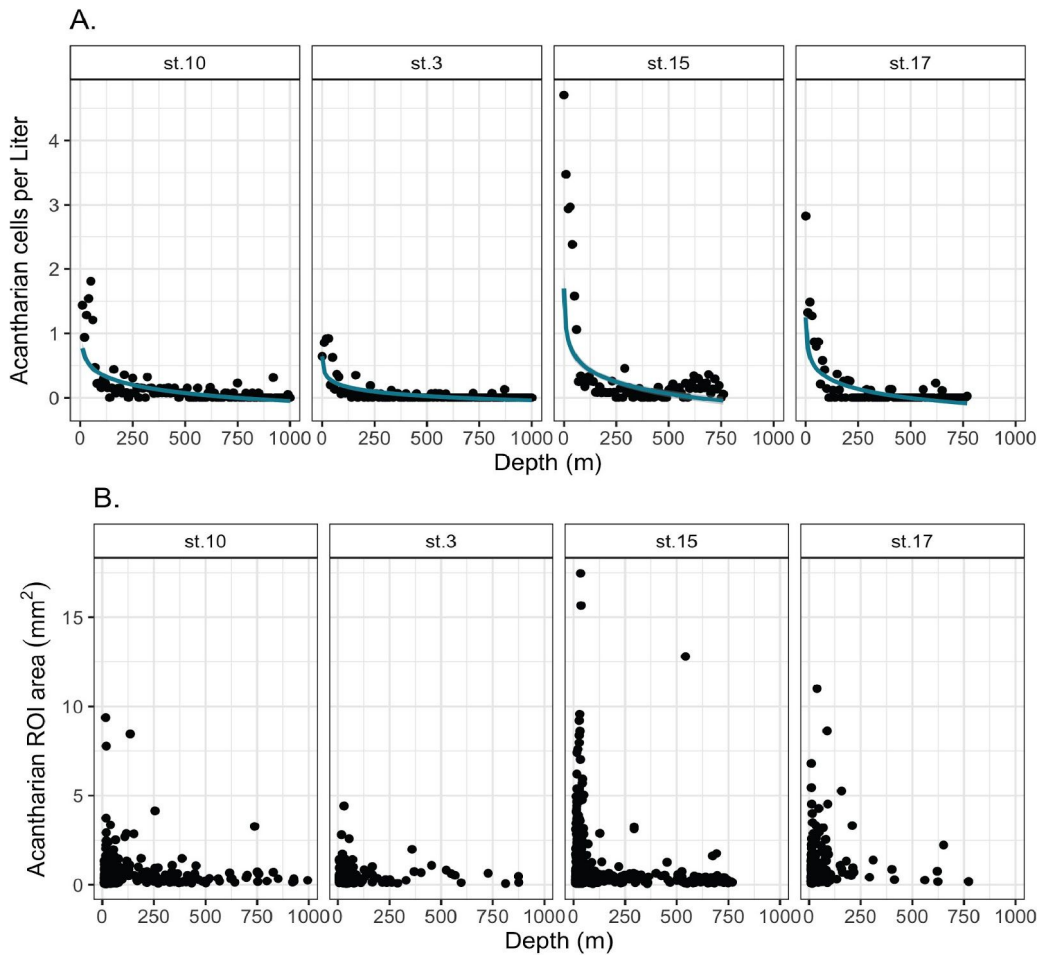
Figure 1.9. Concentrations of visible acantharian cells per Liter observed in vertical imaging profiles (A) and size distributions of acantharian Regions of Interest (ROIs) (B). Acantharian cell concentrations were determined by dividing the number of observed cells for a 10 m section of the vertical profile by the volume of water imaged in that section. Linear models were fit to log-transformed data and are plotted in turquoise with 95% confidence intervals shaded grey (visible for station 15). Linear models for each station (A) were statistically significant ($p < 0.001$) with st. 10 R^2 = 0.64; st. 3 R^2 = 0.63; st. 15 R^2 = 0.66; st. 17. R^2 = 0.78. The highest concentrations of visible acantharian cells were always observed close to the sea surface and decreased sharply with depth. Acantharian ROIs were cropped so that the edges of

the rectangular photos aligned with the outward reaches of cellular extensions in each direction. The pixel dimensions of each ROI image were converted to microns and then the area of ROI images was calculated and plotted against the depth at which it was imaged (B). ROIs exhibit a large size range in surface waters but larger ROIs become less common in deeper waters.
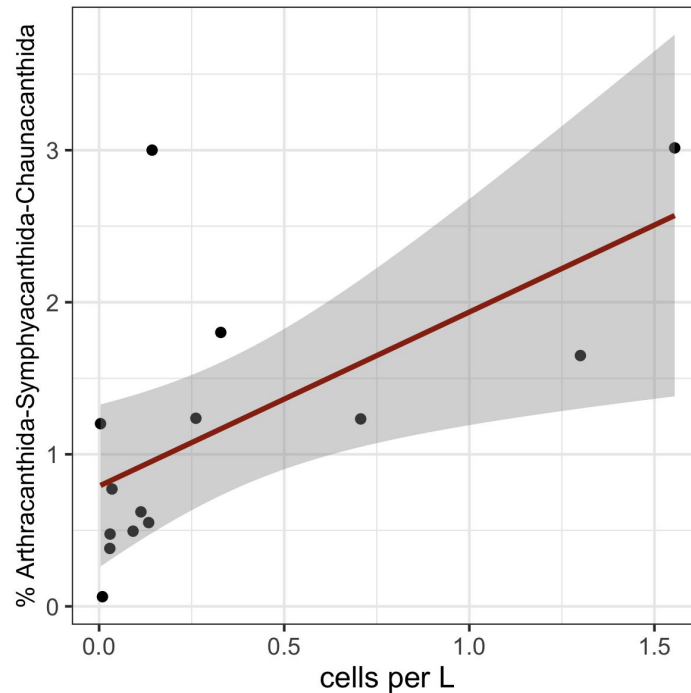
Figure 1.10. Linear regression of acantharian cells per liter averaged for surface, DCM, mid, and bottom depth layers against percentage sequence contributions of Arthracanthida, Symphyacanthida, and Chaunacanthida acantharians at stations 3, 10, 15 and 17. Acantharian cell concentration determined by high-throughput, in-situ imaging significantly correlates with percent sequence contribution of acantharians with known morphologies (R^2 = 0.33, *p* < 0.05).
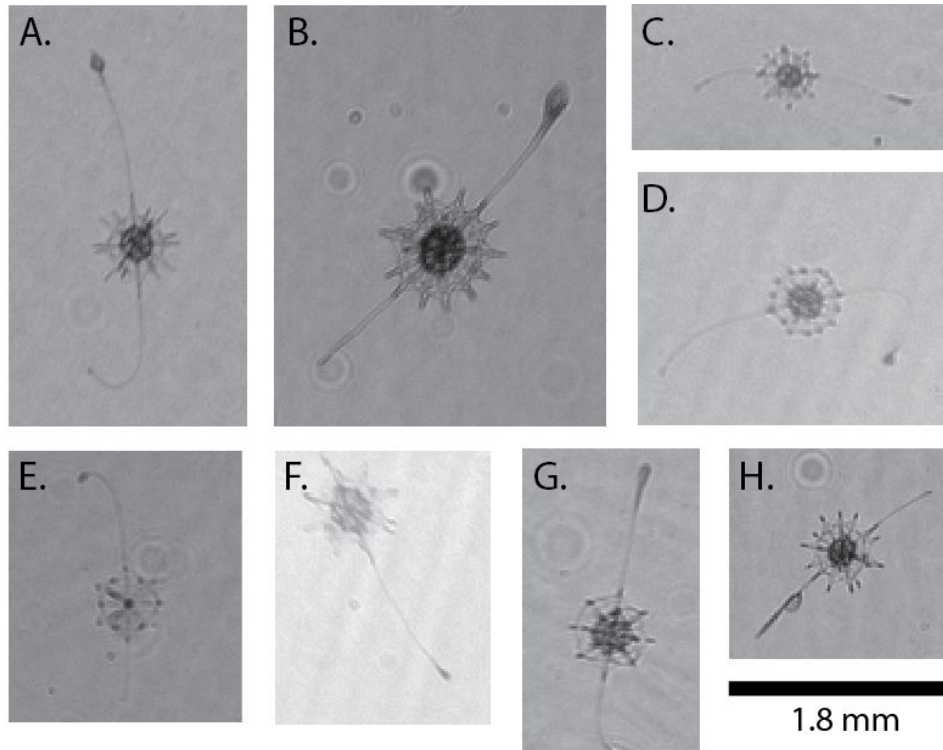
Figure 1.11. In-situ imaging reveals apparent acantharian predation behavior.
Acantharians imaged in this study were observed with long pseudopodial extensions ending in a drop-like shape. This behavior has not been seen previously, probably because the extensions are damaged in net-collected samples. Images were taken at 13.8 m, st. 15 (A); 16.4 m, st. 15 (B, C, E); 27.5 m, st. 15 (D); 31.2 m, st. 15 (F); 57.2 m, st. 17 (G); 13.8 m, st. 17 (H). Image aspect ratios are unaltered and the scale bar below panel H is accurate for all panels. Image orientations are likewise unchanged, with the top of images being toward the sea surface.

1.4. Discussion

Acantharians are important contributors to primary production throughout the global ocean, but detailed studies on absolute abundance and fine-scale distribution have been hindered by specific acantharian traits, such as their fragile cell structures and skeletons that dissolve in common fixatives. In addition, the smaller size of acantharians compared to other Rhizaria has precluded their full inclusion in quantitative imaging surveys (Biard et al., 2016; Biard and Ohman, 2019). As a result, advances in our understanding of acantharian biology and ecology have come primarily from molecular studies. This study took a paired approach and combined

molecular survey methods with high-throughput, in-situ imaging to better evaluate acantharian abundance and distribution. We found that vegetative acantharian cells were concentrated in the uppermost water column, but were sporadically present throughout the water column, including at the deepest depths images were taken (1,000 m, Figure 1.9). The concentrations of acantharian cells determined from imaging data correlated with the contribution of Arthracanthida, Symphyacanthida, and Chaunacanthida acantharians to sequences recovered from the same depth and location (Figure 1.8). In contrast, the percentages of sequences from all acantharians, including those belonging to undescribed environmental clades, increased with depth (Figure 1.7). Together, these results can provide information about the distribution and abundance of different clades of acantharians, the morphology of undescribed environmental clades, and the life-cycles of acantharians.

1.4.1. Acantharian abundance and distribution in the western N. Pacific

Maximum acantharian abundances of 0.9–4.7 cells per L were observed in this study by using an in-situ camera system capable of imaging organisms with diameters greater than 250 μm. Maximum abundances were observed in the upper euphotic zone at each station: 0–10 m depth at stations 15 and 17, 20–30 m depth at station 3, and 40–50 m depth at station 10 (Figure 1.9A). These results are consistent with previous studies that carefully preserved and counted acantharians collected by high-volume plankton pump or with Niskin bottles (Michaels, 1991; Michaels et al., 1995). Michaels et al. (1995) observed near-surface acantharia maxima in the subtropical North Atlantic with maximum abundances ranging from 5.5–18 cells per L (mean 1.2 cells with > 100 μm diameter per L in Niskin samples; mean 2.5 cells per L in pumped samples). Similarly, Michaels (1991) recorded 0.1–4 acantharian cells per L (> 100 μm

diameter) in the surface mixed layer of the eastern North Pacific Subtropical Gyre. Using high-throughput, in-situ imaging, Biard and Ohman (2016) likewise found high concentrations of acantharians near the sea surface in the California Current Ecosystem, although they identified acantharians only with diameter > 600 µm. Compared to other high-throughput imaging studies (Biard et al., 2016; Biard and Ohman, 2019), the acantharian abundances we recorded are closest to the cell abundances reported when cells were previously counted by microscopy (Michaels, 1991; Michaels et al., 1995). While the cell abundances we measured are likely still an underestimate since many acantharians are smaller than 250 µm (Michaels, 1991; Michaels et al., 1995), the results from our high-throughput imaging allow for a more quantitative estimate of acantharian abundance and vertical distribution than analyzing sequencing data alone.

1.4.2. Basal environmental clades of Acantharea

The relative abundance of sequences classified as Chaunacanthida, Symphyacanthida and Arthracanthida decreased as sampling depth increased (Figure 1.8), which correlated with the acantharian abundances determined from imaging data (Figures 1.9, 1.10). In contrast, the relative abundance of all sequences classified as Acantharea at the class level increased with depth and did not correlate with imaging results. The additional acantharian sequences in communities from deeper water were primarily classified as Acantharea-Group-I (Figure 1.6), which is basal to Chaunacanthida, Arthracanthida and Symphyacanthida, and has no known morphology (Decelle et al., 2012c; Decelle and Not, 2015). Similarly, the total acantharian contribution to clone libraries from coastal waters near California increased with depth (Schnetzer et al., 2011) and environmental sequences from basal acantharian clades I–III have been recovered from deep waters throughout the global ocean (López-García et al., 2001;

Edgcomb et al., 2002; Countway et al., 2007; Not et al., 2007; Terrado et al., 2009; Gilg et al., 2010; Quaiser et al., 2011; Decelle et al., 2013). Here, the discrepancy between the depth-related increase in sequence abundance for Acantharea-Group-I and the coinciding decline in cells imaged with characteristic acantharian morphologies may provide new evidence regarding the morphology of basal environmental acantharian clades.

The acantharian skeleton is a central feature to their morphological classification; the most recently diverged clades (Arthacanthida and Symphyacanthida) have spicules of varying lengths—some with elaborate appendages and apophyses—that are fused in a robust central junction, whereas earlier diverging clades (e.g. Chaunacanthida) have simpler spicules of equal length that either cross the central region of the cell or form loosely-fused central junctions (Decelle et al., 2012c). This evolutionary trajectory—from less to more developed skeletons—suggests that the earliest diverging acantharian clades (i.e. basal environmental clades I–III) may have only rudimentary skeletal structures, or may lack the quintessential acantharian skeleton altogether (Decelle et al., 2012c). The decreased observance of recognizable acantharian cells with depth coinciding with an increased sequence abundance for Group I acantharians suggests that they may, indeed, lack traditionally recognized acantharian morphologies. Alternatively, Group I acantharians may simply be too small to be imaged with the ISIIS small-imager used in this study. Ultimately, the morphology of the basal environmental acantharian clades can only be definitively resolved with single-cell sequencing of deep-sea isolates coupled with microscopy (Sieracki et al., 2019). However, these results demonstrate that small cells lacking symmetrical strontium sulfate skeletons should be considered for sequencing in studies seeking to determine the morphology of the earliest diverging acantharian clades.

The high relative abundance of sequences deriving from clade I acantharians in deeper waters (mid and bottom water samples) and the almost complete absence of clade I sequences in shallow water (surface and DCM, Figure 1.6) indicate that these acantharians occupy a different ecological niche than the better studied, later-diverging acantharians with known morphology. Specifically, clade I acantharians seem to only inhabit mesophotic (i.e. the twilight zone) and aphotic zones, meaning they must not rely on photosymbiosis, as all Arthracanthida and Symphyacanthida acantharians (Decelle et al., 2012a) and some Holacanthida acantharians (Decelle et al., 2012b) do. Instead, these acantharians are most likely pure heterotrophs and their small proposed size means that they might graze bacteria or other small flagellates. However, they may also take advantage of larger particles, such as marine snow, as a food source; particle association would contribute to their high relative abundance in the > 10 μm size fraction (Figure 1.6). The results, therefore, indicate that clade I acantharians have both distinct morphology and ecology from later-diverging acantharians.

1.4.3. Acantharian life cycles

Knowledge regarding acantharian life cycles remains relatively limited because a full acantharian life cycle has not yet been observed. However, cyst formation and swarmer release from cysts and vegetative cells have been observed in laboratory settings (Decelle et al., 2012a, 2013). Swarmers are small (2–3 μm) biflagellated cells with unknown ploidy; it is assumed that they are reproductive cells and fuse to form juveniles, but this has never been witnessed (Decelle et al., 2012a). So far, cyst formation has only been observed for earlier diverging acantharian lineages and has not been observed for Arthracanthida or Symphyacanthida acantharians. Acantharians that form cysts shed most of their spicules before cyst formation, suggesting that

acantharians in later diverging clades with more robust and elaborate skeletons cannot form cysts because of the fixed central junctions in their skeletons. In addition, cysts recovered from sediment traps have only ever been found to belong to earlier diverging clades, based on phylogenetic analysis, but not to environmental clades I or II (Decelle et al., 2013). As a result, current hypotheses propose that acantharians in earlier diverging clades, including Chaunacanthida, form cysts as a means for ballast, allowing them to sink to deep water where they release swarmer cells, whereas acantharians in the later diverging Arthracanthida and Symphyacanthida clades complete their life cycle in the photic zone, since they cannot form cysts and need to acquire photosymbionts at the start of each generation (Martin et al., 2010; Decelle et al., 2013).

The imaging results demonstrating decreased cell size and abundance below the surface mixed layer (Figure 1.9B) are consistent with the idea that many acantharians sink to release swarmers and the new juveniles grow in size as they make their way up towards the surface. Adult cells sink quickly, whether for reproduction or as detritus (Gutierrez-Rodriguez et al., 2019), and both vegetative cells and cysts dissociate after releasing swarmers (Decelle et al., 2012a, 2013), making them less likely to be caught on camera. A large number of swarmers released at depth would potentially produce many small juvenile cells that gradually increase in size as they slowly ascend. These smaller, more abundant juveniles would be much more likely to be imaged than the rarer, faster-moving adults. Alternatively, decreased cell-size range at depth could also reflect decreased nutritional resources available in deeper waters or constitutively smaller sized species being more common below the surface mixed layer. However, the sporadic presence of large cells in very deep water is evidence that they at least

occasionally reach deep water from the surface or that cells can reach larger sizes at depth (Figure 1.9B).

By combining metabarcoding of size-fractionated samples with imaging, further insight into acantharian life cycles can be gained. In principle, vegetative cells and cysts should have been retained on the upper filter with larger pore-size and the swarmer cells should have passed through the upper filter and been retained on the lower filter with smaller pore-size. A disparity in the contribution of one clade to sequences from the two size-fractions could, therefore, indicate that vegetative cells or reproductive cells belonging to that clade are more or less abundant at a particular depth. Such size separation can never be perfect—and it may be especially problematic with delicate cells like acantharians—but Principal Coordinate Analysis (PCoA) of Aitchison distances between entire plankton communities showed clear segregation by filter type at each depth (Figure 1.5A), suggesting size-fractionation was relatively successful. In contrast, PCoA for acantharian communities in the same samples did not show segregation by filter type (Figure 1.5B), thus providing some evidence that swarmers and vegetative cells or cysts coexist at the depths sampled. Given that acantharians belonging to the Chaunacanthida clade are among those that form cysts and are believed to sink before releasing swarmers, Chaunacanthida would be expected to be more abundant than non-cysting Arthracanthida and Symphyacanthida in both size fractions at depth. Interestingly, Arthracanthida and Symphyacanthida sequences were recovered from both large and small size-fraction samples from mid and near-bottom water at every sampling station and had similar relative abundances to Chaunacanthida sequences (Figure 1.6). This lack of differentiation in deep water sequence abundances of Arthracanthida-Symphyacanthida and Chaunacanthida does not support the idea

that the later diverging clades only reproduce in the photic zone. An alternative hypothesis might be that some Arthracanthida and Symphyacanthida acantharians also sink to deep water to reproduce but do so as vegetative cells, aided by their robust skeletons and fine buoyancy control (Febvre and Febvre-Chevalier, 2001), rather than in the form of cysts (Michaels et al., 1995). It cannot be ruled out that the DNA recovered from deep waters could be extracellular or derive from detrital matter, but the alternative hypothesis is further supported by the occasional observation of large acantharian cells at depth (Figure 1.9B, Biard and Ohman, 2019).

1.4.4. Acantharian behavior revealed by in-situ imaging

Being notoriously delicate and sticky, acantharians are often broken or clumped when collected by plankton net. As a result, their fine structure is usually damaged even when they do survive collection, which can preclude behavioral observations. In-situ imaging is especially useful in such cases, as it allows for the observation of natural orientation and behaviors that could not otherwise be seen. Acantharians are known to be active predators: microscopy of SCUBA collected acantharians revealed ciliates, diatoms, and dinoflagellates as acantharian prey items (Swanberg and Caron, 1991) and results from 18S sequence analysis of single acantharians included copepod, diatom, and dinoflagellate sequences (Mars Brisbin et al., 2018). However, actual predatory strategies of acantharians are unknown. In this study, we repeatedly observed acantharians that had long pseudopodial extensions terminating in drop-shaped structures. This morphology/behavior has not been previously reported and we hypothesize that the extensions may represent a fishing apparatus that allow acantharians to lure and capture prey. However, since we did not observe prey items stuck to the droplets, it remains possible that these structures may be involved in other processes (e.g. reproduction, buoyancy, or locomotion).

1.4.5. Prospects for future automatic classification of acantharians

The images produced in this study were annotated manually, which represents a major barrier in high-throughput imaging studies; manual image annotation is a large time commitment and requires experience identifying plankton groups. The ultimate goal for high-throughput imaging is to have annotated training sets that are extensive and comprehensive enough to provide highly accurate automatic image classification using machine learning algorithms. Currently, several instrument- and location-specific learning sets are available in the public realm: e.g. the WHOI-plankton dataset, which includes 3.4 million annotated images in 70 classes that were taken with the Imaging Flow Cytobot in Martha's Vineyard (Orenstein et al., 2015), and the PlanktonSet-1.0, which includes 30,336 images in 121 classes that were taken with the ISIIS line-scan imager in the Straits of Florida (Rodrigues et al., 2018). These datasets could be used for transfer learning to improve classification accuracy and efficiency, but since they do not include very many acantharian images and the plankton size range for both data sets excludes the majority of acantharians imaged in this study (Rodrigues et al., 2018), they alone would not have allowed for accurate classification of acantharians in our dataset. Therefore, by contributing over a thousand new annotated acantharian images, including cells ranging in diameter from 250–2500 μm, this study will facilitate accurate automatic acantharian classification in future datasets acquired with the ISIIS small-imager and other imaging systems that capture a similar size-range of organisms. Accurate automatic classification will eventually allow for larger studies of acantharian abundance and distribution, including expanded geographic and temporal scales, and thus a deeper understanding of acantharian contributions to biogeochemical processes in the ocean.

1.4.6. Conclusions

High-throughput imaging used in this study showed that acantharians are abundant in the surface waters of the western North Pacific and have similar concentrations as have been reported in the eastern North Pacific and the North Atlantic where cells were manually counted with microscopy. Similar to previous studies, vegetative acantharian cells were concentrated very close to the sea surface and decreased in abundance with depth, but were still sometimes observed at depths approaching 1000 m. Imaging data correlated with sequence abundances from acantharian clades with known and easily recognizable morphologies, but were in contrast to sequence abundances from acantharian environmental clade I, whose morphology is not known. This discrepancy suggests that basal environmental clades, such as clade I, may have morphologies distinct from other acantharians, and may lack characteristic star-shaped strontium sulfate skeletons. The high abundance relative of clade I acantharian sequences in deep water and their near total absence in surface waters suggest a distinct ecological niche for deep water acantharian populations, likely as grazers, as opposed to these sequences being representative of surface export. The size distribution of imaged acantharians is consistent with current hypotheses about acantharian life-cycles: size range decreases with depth, supporting the idea of reproduction at depth by way of small swarmer cells, followed by the ascension, and growth, of juveniles into surface waters. However, the similar relative abundance of different acantharian clades in small and large size fractions at depth suggests that later diverging clades (i.e. Arthracanthida and Symphyacanthida) may also reproduce at depth, which is counter to previous hypotheses. By pairing high-throughput sequencing with high-throughput, in-situ imaging, this study advances our understanding of acantharian biology but also highlights how much is still

unknown. Future work will benefit from the annotated images generated for this study, but

should consider further pairing imaging with RNA sequencing or single-cell genomics.

Chapter Two: Specificity and maintenance in acantharian photosymbioses

This chapter is published as:

Abstract

Photosymbiotic protists contribute to surface primary production in low-nutrient, open-ocean ecosystems and constitute model systems for studying plastid acquisition via endosymbiosis. Little is known, however, about host-symbiont dynamics in these important relationships, and whether these symbioses constitute mutualisms is debated. In this study, single-cell sequencing methods and advanced fluorescent microscopy were both applied to investigate host-symbiont dynamics in clade F acantharians, a major group of photosymbiotic protists in oligotrophic subtropical gyres. The 18S rRNA gene from single acantharian hosts and environmental samples was sequenced to assess intra-host symbiont diversity and to determine whether intra-host symbiont community composition directly reflects the available symbiont community in the surrounding environment. Results demonstrate that clade F acantharians simultaneously host multiple species from the haptophyte genera *Phaeocystis* and *Chrysochromulina,* indicating that symbiont uptake occurs more than once. The intra-host symbiont community composition was distinct from the external free-living symbiont community, suggesting that these acantharians

maintain symbionts for extended periods of time. After selectively staining digestive organelles, fluorescent confocal microscopy showed that symbionts were not being systematically digested, which is consistent with extended symbiont maintenance within hosts. Extended maintenance within hosts may benefit symbionts through protection from grazing or viral lysis, and therefore could enhance dispersal, but only if symbionts retain reproductive capacity and are ever released or escape from hosts.

2.1. Introduction

Photosymbiosis, a nutritional symbiosis where a heterotroph hosts photosynthetic endosymbionts, substantially increases surface primary production in oligotrophic marine ecosystems (Not et al., 2016; Beinart, 2019). Photosymbiosis is also believed to have led to the evolution of eukaryotic oxygenic photosynthesis and the eventual emergence of diverse photosynthetic eukaryotes, with many evolving from eukaryote-eukaryote secondary and tertiary endosymbiosis (Keeling, 2004). Eukaryote-eukaryote photosymbioses continue to be extremely common among marine protists and contribute significantly to the productivity of oligotrophic open-ocean ecosystems (Decelle et al., 2015). Nonetheless, little is known about host-symbiont dynamics, such as host-symbiont specificity or host mechanisms for symbiont recognition, uptake, and maintenance. Additionally, photosymbioses have traditionally been considered mutualisms under the assumption that hosts provide nitrogen to symbionts and symbionts provide organic carbon to hosts in return (Garcia and Gerardo, 2014). Whether these relationships are truly mutualistic or are instead cases of symbiont exploitation has been increasingly questioned in recent years and there is mounting evidence that exploitation is the rule rather than the exception (Keeling and McCutcheon, 2017). Determining the nature of

photosymbioses is particularly interesting, as it could provide insight into how the relationships

evolve and persist.

The Acantharea belong to the Rhizaria, a supergroup of amoeboid protists that includes

many photosymbiotic lineages (Burki and Keeling, 2014). Photosymbiotic acantharians are often

the most abundant photosymbiotic Rhizaria in oligotrophic surface waters (Michaels et al.,

1995), where they contribute significantly to primary production (Caron et al., 1995). The

majority of acantharian species (clades E and F) host algal symbionts in the Haptophyte genus

Phaeocystis (Decelle et al., 2012a). *Phaeocystis* is a globally distributed genus with species that

present multiple phenotypes—solitary, flagellated, and colonial—and sometimes form harmful

algal blooms (Schoemann et al., 2005). Despite the ecological significance of both partners, this

symbiosis remains largely unstudied. There is some evidence, however, suggesting that this

relationship is not a case of mutualism and symbionts are instead exploited (Decelle, 2013).

When photosymbiotic protists are cultured in high-light and low-prey conditions, as

found in oligotrophic surface waters, hosts benefit from an increased growth-rate, but symbiont

growth-rate can be suppressed and their photosynthetic efficiency is sometimes decreased

compared to free-living symbionts (Lowe et al., 2016). Therefore, some algal symbionts may

actually experience restricted nitrogen availability *in hospite*, rather than receiving supplemental

nitrogen from hosts, and thus do not benefit from symbiosis as assumed (Lowe et al., 2016).

Estimated populations of free-living *Phaeocystis* in oligotrophic conditions (Moon-van der Staay

et al., 2000) are much larger than possible symbiotic populations if estimated from acantharian

abundance and symbiont load (Michaels, 1991). The difference in population size between

symbiotic and free-living *Phaeocystis* suggests higher growth-rates in free-living than symbiotic

populations, and provides some evidence that the relationship is not beneficial to *Phaeocystis*

symbionts (Decelle, 2013).

Increased growth rate is not the only means by which symbionts can benefit from the

relationship: decreased predation, viral attack, or competition *in hospite* may allow symbionts to

benefit from enhanced dispersal and future reproduction, assuming reproductively viable

symbionts are ever released from hosts (Douglas, 2010; Garcia and Gerardo, 2014). Reproducing

symbionts are known to be released from some photosymbiotic protists: *Chlorella* escapes from

Paramecium hosts and establishes free-living populations when low-light inhibits host benefit

(Lowe et al., 2016) and dinoflagellate symbionts of colonial radiolarians establish free-living

populations when isolated from hosts (Probert et al., 2014). Some photosymbiotic forams,

however, digest all of their symbionts prior to gametogenesis (Takagi et al., 2016). It is currently

unknown whether symbiotic *Phaeocystis* retains reproductive capacity, but symbiotic cells have

not yet been cultured from hosts (Decelle et al., 2012a). It is possible that phenotypic changes

observed in symbiotic *Phaeocystis*—additional plastids and an enlarged central vacuole (Febvre

and Febvre-Chevalier, 1979; Decelle et al., 2012a)—are evidence that symbionts are incapable

of cell division, which would make the relationship an ecological and evolutionary dead-end for

Phaeocystis and preclude the possibility for mutualism (Decelle et al., 2012a; Decelle, 2013).

The number of symbionts observed in individual acantharians increases with host size

(Michaels, 1991), thus requiring that symbionts reproduce *in hospite*, that acantharians recruit

new symbionts, or possibly both. If acantharians recruit one or a few symbionts early in

development and then support a reproducing symbiont community, the intra-host symbiont

community would exhibit low diversity and may be divergent from the free-living environmental

community. If symbionts divide within hosts, hosts must exert an alternative form of population

control, potentially by shedding (mutualism) (Boettcher et al., 1996; Fishman et al., 2008) or

digesting (exploitation) excess symbionts (Titlyanov et al., 1996). Conversely, if acantharians

recruit new symbionts, the intra-host symbiont community is likely to be more diverse and

representative of the available free-living symbiont community in the surrounding waters.

Low-diversity intra-host symbiont communities would, therefore, suggest that symbionts

maintain reproductive capacity and allows for possible symbiont benefit, whereas high diversity

communities may be interpreted as further evidence against mutualism. However, neither

intra-host symbiont diversity, nor the relationship between symbiont identity and environmental

availability of potential symbionts have been investigated in photosymbiotic acantharians.

In this study, single-cell Next Generation Sequencing (NGS) was used to illuminate

intra-host symbiont diversity in individual acantharians collected from 7 sampling sites along the

Ryukyu Archipelago, spanning more than 1,000 km in the East China Sea (ECS), and near

Catalina Island (California, USA). NGS was further applied to evaluate the environmental

availability of symbionts where acantharians were sampled in the ECS. Intra-host symbiont

diversity was compared to intra-host symbiont population size by enumerating symbionts with

fluorescent confocal microscopy in a subset of acantharians prior to nucleic acid extraction.

Additional acantharians were collected and imaged after selectively staining lysosomes and

phagolysosomes in order to observe their proximity to symbionts and to determine if symbionts

are systematically digested. This study provides new evidence to the mutualism-exploitation

debate relative to Acantharea-*Phaeocystis* symbioses by investigating intra-host symbiont

diversity and by assessing host-symbiont specificity in the context of environmental symbiont availability.

2.2. Materials and Methods

2.2.1. Individual acantharian sampling

Single acantharians were collected from coastal water near Catalina Island (California, U.S.A.) and from 7 sampling sites along the Ryukyu Archipelago, including coastal water near Okinawa Island (Okinawa, Japan) and from 6 cruise stations visited during the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) MR17-03C cruise to the ECS aboard the R/V *Mirai* in May and June 2017 (Figure 2.1, Table 1.1). Okinawa Island and Catalina Island plankton samples were collected by pulling a Rigo Simple 20 cm diameter, 100-mm-mesh plankton net or a SEA-GEAR 12" diameter, 163-µm-mesh plankton net, respectively, along the sea surface approximately 5 m behind a small craft at its lowest speed. Aboard the R/V *Mirai*, plankton samples were collected by passing unfiltered seawater pumped from the sea surface through a 100-µm-mesh, hand-held plankton net (Rigo). Plankton samples were observed under a dissecting microscope and individual acantharians were transferred by glass micropipette to clean Petri-dishes. Acantharians were rinsed with 0.2-µm-filtered seawater several times, until all visible contaminants were removed, and then cells were incubated for 0.5–2 hr to allow for additional self-cleaning. Acantharians collected aboard the R/V *Mirai* and those from near Okinawa Island were imaged with inverted light microscopy (Zeiss Primovert, Olympus CKX53, Figures S2.1 and S2.2). Several acantharians collected near Okinawa Island were also imaged with laser confocal microscopy (described below). Each acantharian was transferred to a

maximum recovery PCR tube (Axygen) and successful transfer was confirmed by microscopy

before adding 30 µL of RLT-plus cell-lysis buffer to each tube (Qiagen). Immediately following

buffer addition, samples were flash-frozen with liquid nitrogen and stored at -80°C until later

processing in the lab.
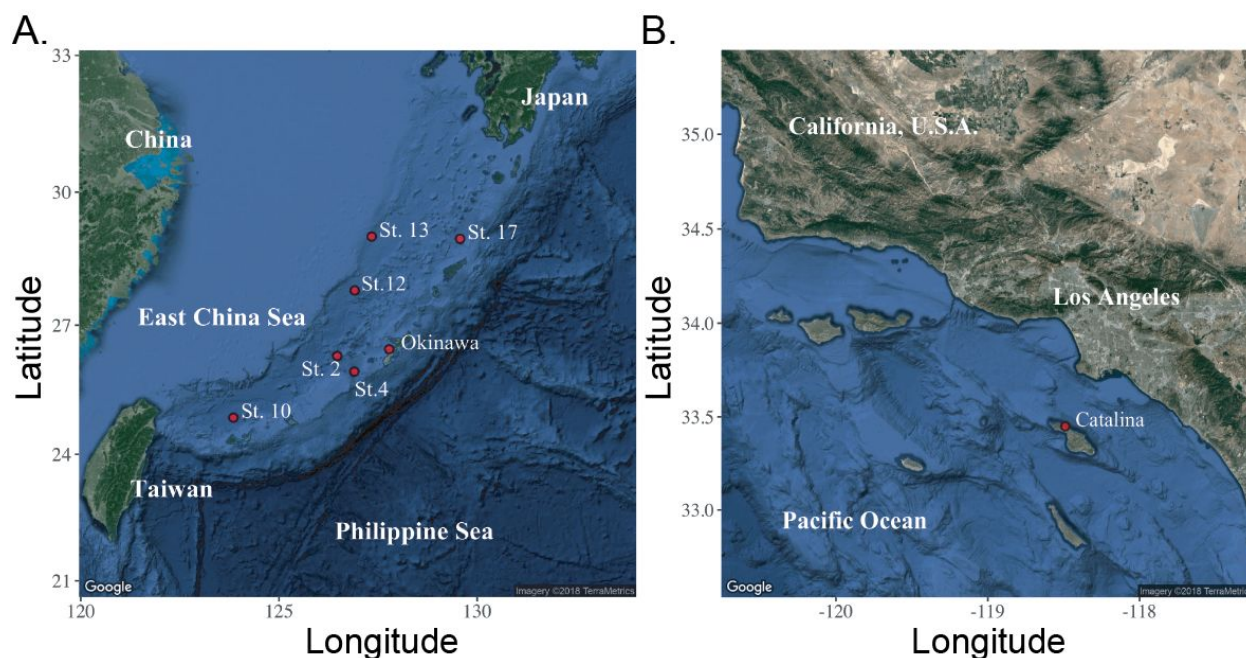
Figure 2.1. Sampling sites along the Ryukyu Archipelago in the East China Sea (ECS) (A) and near Catalina Island, California, U.S.A. (B). (A) ECS Stations 2, 4, 10, 12, 13, and 17 were sampled during the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) MR17-03C cruise in May and June 2017. Samples were collected from the Okinawa Island (Okinawa, Japan) sampling site in April, May, and December 2017. (B) Additional samples were collected near the University of Southern California's Wrigley Institute for Environmental Studies on Catalina Island, California, U.S.A. in May 2017.

2.2.2. Environmental sampling

Seawater samples were collected at each ECS cruise station visited by the JAMSTEC

MR17-03C cruise where acantharians were also isolated. Two replicates of 4.5 L of seawater

were collected from the sea surface by bucket and sequentially filtered under a gentle vacuum

through 10.0 µm and 0.2 µm pore-size Polytetrafluoroethylene (PTFE) filters (Millipore) to size

fractionate plankton and separate free-living *Phaeocystis* (< 10 μm) from acantharian hosts (> 10 μm). Filters were flash-frozen in liquid nitrogen onboard and stored at −80°C until processing onshore.

2.2.3. RNA extraction from individual acantharian hosts

RNA extractions from single acantharians (n = 42, 1–14 per site) were accomplished by modifying the methods of Trombetta et al. (2015). Samples were thawed over ice, vortexed twice (10 s, speed 7, Vortex-Genie 2), and then incubated at room temperature for 5 min to fully lyse cells. Agencourt RNAClean XP magnetic beads (Beckman Coulter) were added to each sample at a 2.2:1 V:V ratio and fully mixed by pipette prior to a 30-min incubation in order to bind all RNA to the magnetic beads. After two 80% ethanol washes, RNA was eluted from the beads in 11 μL of a custom elution buffer (10.72 μL nuclease-free water, 0.28 μL RNAase inhibitor) and 10.5 μL of eluted RNA was further processed following the single-cell protocol for the SMART-seq v4 Ultra Low Input Kit (Clonetech) with 18 cycles in the primary PCR. The resulting cDNA from each sample was quality checked with the Bioanalyzer High Sensitivity DNA Assay (Agilent) and quantified with the Qubit dsDNA High Sensitivity Assay (Qubit 3.0, ThermoFisher).

2.2.4. DNA extraction from environmental samples

Environmental DNA was extracted from the 0.2-μm pore-size PTFE filters (n = 12, 2 replicates from 6 stations) following manufacturer protocols for the Qiagen AllPrep DNA/RNA Mini Kit with limited modifications. Half of each PTFE filter was submerged in RLT-plus cell-lysis buffer (Qiagen) with garnet beads in a 2-mL tube (MoBio/Qiagen). Samples were

heated for 10 min at 65°C and then vortexed at maximum speed (Vortex-Genie 2) for 5 min with the MoBio/Qiagen vortex adapter to fully lyse cells. After cell lysis, DNA extraction proceeded without further modifications. Extracted DNA was quantified with the Qubit dsDNA High Sensitivity Assay on a Qubit 3.0 instrument (ThermoFisher).

2.2.5. Library preparation and sequencing

Library preparation for acantharian cDNA samples and environmental DNA samples followed procedures described in the Illumina 16S Metagenomic Sequencing Library Preparation manual, modified only to include universal eukaryotic primers for the v4 region of the eukaryotic 18S rRNA gene (Stoeck et al., 2010) and amplicon PCR conditions most appropriate for these primers. The forward primer, TAReuk454FWD1 (CCAGCASCYGCGGTAATTCC, Stoeck et al., 2010), was used unmodified. The reverse primer, TAReuk454REV3 (ACTTTCGTTCTTGATYRA, Stoeck et al., 2010) was reported to not amplify the *Phaeocystis* 18S gene in *in silico* PCRs (Tanabe et al., 2016). Further investigation revealed a mismatch between the final 3' adenine in the primer sequence and the *Phaeocystis* 18S gene sequence. Although we found that the original primers do amplify the v4 region of the *Phaeocystis* 18S gene in *de facto* PCRs with DNA extracted from *Phaeocystis* cultures, the mismatch could create bias against *Phaeocystis* sequences in more diverse samples. We, therefore, eliminated the final 3' "A" in the TAReuk454REV3 sequence and used a new reverse primer, TAReuk454REV3.1 (ACTTTCGTTCTTGATYR). The optimum annealing temperature for the Illumina-adapted primers was determined by performing temperature gradient PCRs (53–65°C, 0.5°C steps) and the annealing step in the amplicon PCR was set at 58°C thereafter. Following the second, indexing PCR and final product purification, amplicon libraries were quantified with the Qubit

dsDNA High Sensitivity Assay (Qubit 3.0, ThermoFisher) and the amplicon size was determined with the Bioanalyzer High Sensitivity DNA Assay (Agilent). Amplicon libraries were then submitted to the Okinawa Institute of Science and Technology (OIST) sequencing center for 300x300-bp paired-end sequencing on the Illumina MiSeq sequencing platform with v3 chemistry. Raw sequences generated for this study were submitted to the European Nucleotide Archive under the study accession number PRJEB24538.

2.2.6. Amplicon sequence analysis and annotation

Demultiplexed paired-end sequences returned from the OIST sequencing center were imported to QIIME 2 (v2017.11) software (Bolyen et al., 2019). The Divisive Amplicon Denoising Algorithm (DADA) was implemented with the DADA2 plug-in for QIIME 2 to perform quality filtering and chimera removal and to construct a feature table consisting of read abundance per amplicon Sequence Variant (SV) by sample (Callahan et al., 2016c). DADA2 models the amplicon sequencing error in order to identify unique SVs and infers sample composition more accurately than traditional Operational Taxonomic Unit (OTU) picking methods that identify representative sequences from clusters of sequences based on a % similarity cut-off (Callahan et al., 2016c). Taxonomy was assigned to SVs in the feature table with a naive Bayes classifier trained on SILVA 18S 97% representative sequences and consensus taxonomy (Release 128, (Quast et al., 2013) using the QIIME 2 feature-classifier plug-in (Bokulich et al., 2018). The SV feature table was split into two separate feature tables, one acantharian and one environmental, before both feature tables were extracted from QIIME 2 and imported into the R statistical environment (R Core Team, 2018) for further analysis with the R package phyloseq (McMurdie and Holmes, 2013).

Prevalence filtering was applied to both the acantharian and environmental feature tables

with phyloseq in order to remove low-prevalence (< 5%) SVs and decrease the chance of data

artifacts affecting the analysis (Callahan et al., 2016b). Prevalence filtering effectively eliminated

most sequences from known Rhizaria prey (i.e. metazoans and diatoms (Swanberg and Caron,

1991)) and parasites (i.e. alveolates (Bråte et al., 2012)) from the acantharian feature table. Of

the remaining SVs in the acantharian table, 26 were classified as Rhizaria and 21 as

Prymnesiophyceae, which includes *Phaeocystis*. In addition, 3 remaining sequences were

classified as Holozoa, 2 as Stramenopiles, 2 as Apusomonadidae, 2 as Ancyromonadida, 1 as

Chloroplastida, and 5 were not classified. When the unclassified sequences were queried against

the GenBank non-redundant nucleotide (nr/nt) database, the tops hits for all 5 sequences were

Acantharea (BLASTn 2.8.0+, 07/1/2018, Camacho et al., 2009). Unlike Holozoa (metazoans) or

Stramenopiles (diatoms), which are acantharian prey (Swanberg and Caron, 1991),

Apusomonadidae, Ancyromonidae, and Chloroplastida have not previously been found in

association with acantharians. As a result, we considered only the 21 Prymnesiophyceae

sequences as symbiotic SVs. In the environmental feature table, there were 187

Prymnesiophyceae sequences remaining after prevalence filtering, but since it is not possible to

know which of these can be acantharian symbionts, we further filtered the environmental feature

table to only include the 21 symbiotic SVs also found in acantharian samples.

The 21 symbiotic SVs were further classified by building a phylogenetic tree. An initial

BLAST query against the GenBank nr/nt database indicated that the symbiont SVs belong to the

Haptophyte genera *Phaeocystis* and *Chrysochromulina.* Likewise, a SILVA SSU sequence

search (03/23/2018, Quast et al., 2013) classified 18 of the sequences as class Prymnesiophycae,

order Phaeocystales or Prymnesiales, and genus *Phaeocystis* or *Chrysochromulina,* when

classified to genus level. The remaining 3 symbiotic SVs were designated "unclassified" in the

SILVA sequence search. Reference 18S rRNA gene sequences were downloaded from GenBank

(Benson et al., 2012) for the 5 Haptophyte orders (Pavlovales, Phaeocystales, Prymnesiales,

Isochrysidales, and Coccolithales (Medlin and Edvardsen, 2007)) and were included in a

Multiple Sequence Comparison by Log Expectation (MUSCLE) v3.8.31 (Edgar, 2004)

alignment along with the 21 symbiotic SVs. A Bayesian phylogenetic tree was then built from

the alignment using MrBayes v3.2.7 with the number of nucleotide substitution types (nst) set to

6 (Ronquist and Huelsenbeck, 2003) (Figure 2.2). A phylogenetic tree was built for the 5

dominant acantharian host SVs following the same methods (Figure 2.3) with representative

sequences for acantharians in clades E and F (Decelle et al., 2012c), also downloaded from

GenBank.
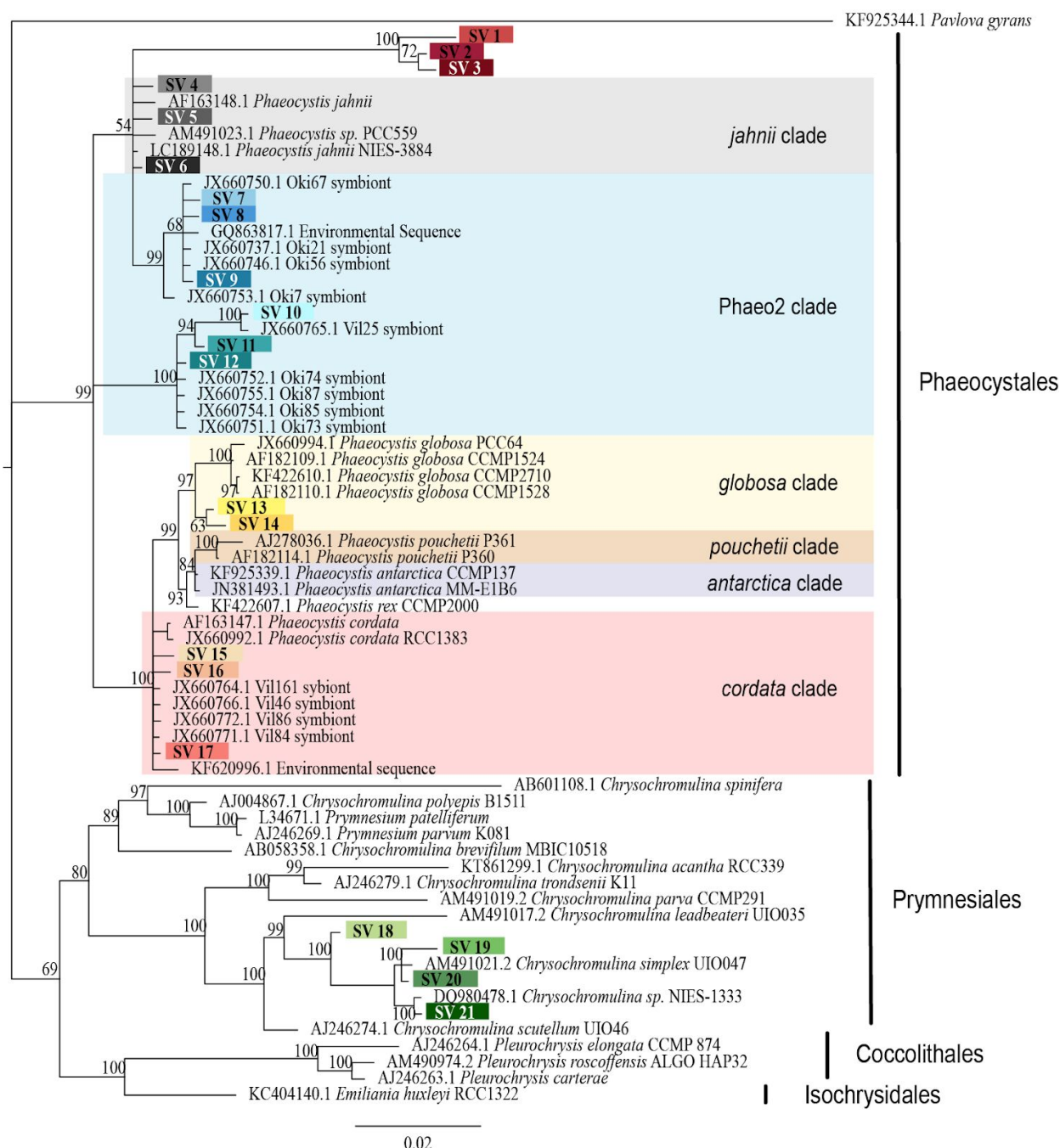
Figure 2.2. Phylogenetic placement of symbiotic Sequence Variants (SVs) identified from acantharian hosts. The phylogenetic tree was built from a MUSCLE v3.8.31 alignment of 21 symbiotic SVs and GenBank reference sequences representing the 5 haptophyte orders with MrBayes v3.2.7. The symbiotic SVs are highlighted to match the legend in Figure 2.4. *Phaeocystis* clades are color-coded and include the Phaeo2 clade, which is an uncultured clade identified from symbiotic sequences from acantharians collected near Okinawa (Decelle et al. 2012a). Values associated with nodes are posterior probabilities as a percent after 100,000

generations (average standard deviation of split frequencies = 0.019). The scale bar indicates 0.02 changes expected per site.
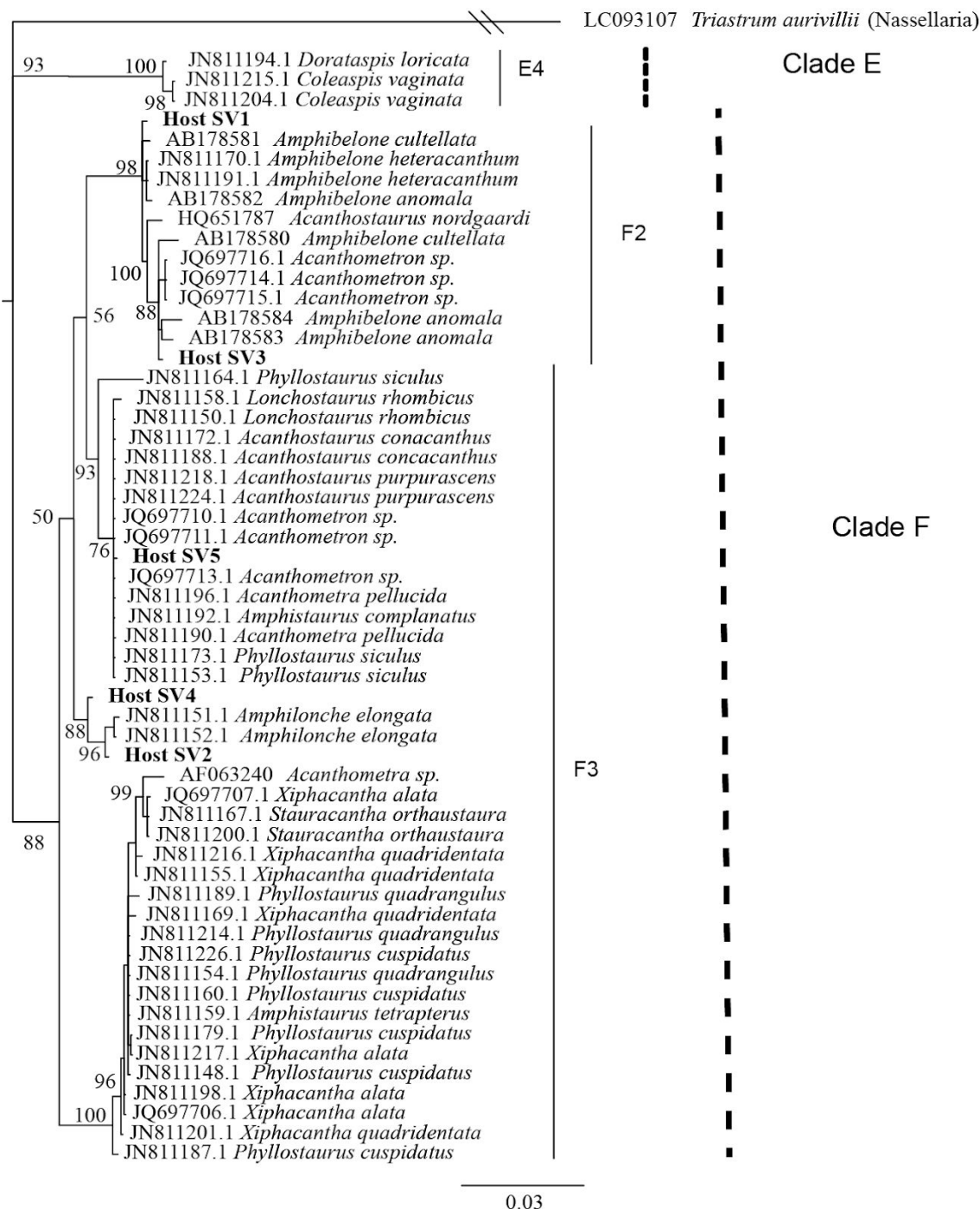
Figure 2.3. Phylogenetic placement of acantharian host Sequence Variants (SVs). The phylogenetic tree was built from a MUSCLE v3.8.31 alignment of the 5 dominant host SVs and GenBank reference sequences representing clades E and F acantharians and the nasellarian radiolarian *Triastrum aurivillii* (outgroup). The tree was built with MrBayes v3.2.7. Values associated with nodes are posterior probabilities as a percent after 40,000 generations (average

standard deviation of split frequencies = 0.01). The scale bar indicates 0.03 changes expected per site.

2.2.7. Statistical analyses

Bray-Curtis distances between symbiont community compositions in acantharian and environmental samples were computed from relative abundances of symbiotic SVs in the filtered feature tables with the R package phyloseq (McMurdie and Holmes, 2013). Bray-Curtis distances were used to perform Principal Coordinate Analyses (PCoA) within the phyloseq package, and PCoA plots were rendered with the R package ggplot2 (Wickham, 2010). Permutational Multivariate Analyses of Variance (PERMANOVA) with 999 permutations were performed with the adonis function in the R package vegan (Oksanen et al., 2019) to determine whether clustering observed in the ordination plots was statistically significant and to discern which covariates were deterministic of symbiont community composition. Specifically, adonis PERMANOVA were performed by location and by host SV on a Bray-Curtis distance matrix including all acantharian samples as well as by sample type (acantharian or environmental) for a Bray-Curtis distance matrix including the environmental samples and acantharian samples collected at environmental sampling locations. Pairwise PERMANOVA by location and by host SV were also performed on the Bray-Curtis distance matrix including all acantharian samples with the beta-group-significance function in the QIIME 2 diversity plugin. Differences were considered statistically significant when the *p*-value was < 0.05. Intermediate data files and data analysis pipelines are available at https://github.com/maggimars/AcanthareaPhotosymbiosis and https://maggimars.github.io/AcanthareaPhotosymbiosis/Analysis.html.

2.2.8. Fluorescent confocal microscopy

In order to observe and enumerate symbionts within hosts for which symbiont communities were also evaluated, acantharians collected near Okinawa in April (Oki.3A, Oki.4A) and May (Oki.3, Oki.7, Oki.10, Oki.11) were imaged without staining using an inverted laser scanning confocal microscope (Zeiss LSM 780) prior to RNA extraction. A z-stack of chlorophyll autofluorescence (ex633 nm, em670 nm) and halogen light images was assembled for each host to compare the number of visually discernible symbionts to the number of symbiotic SVs identified. In order to evaluate possible host digestion of symbionts, additional acantharian samples were collected in December 2017 (n = 3) and stained with LysoTracker Green DND-26 (ThermoFisher), a fluorescent dye that selectively stains acidic compartments (i.e. digestive organelles) within cells, including lysosomes and phagolysosomes. The LysoTracker dye was diluted from the 1-µM stock solution to a 100-nM working solution in 0.2-µm-filtered seawater and each sample was incubated in 100 µL working solution in the dark for 2 hr before imaging. Z-stacks for these samples were assembled by imaging with 3 channels: red for autofluorescence from symbiont chlorophyll (ex633 nm, em670 nm), green for LysoTracker-stained host lysosomes and phagolysosomes (ex488 nm, em514 nm), and grey for halogen light imaging. FIJI Image-J software (Schindelin et al., 2012) was used to adjust image brightness, merge color channels, and create 3-D projections for all imaged host cells.

2.3. Results

2.3.1. Intra-host symbiont diversity in individual acantharians

In order to determine if and to what extent intracellular symbiont diversity exists in acantharians, single-cell RNA extractions were performed and 18S rRNA gene amplicons were sequenced from 42 individual acantharians. A total of 6,154,808 sequences were generated from acantharian samples, with 28,828–241,809 sequences per sample. After quality filtering and feature table construction, 3,294,093 total sequences remained (5,579–136,339 per sample). Within each acantharian sample, 72–99% of sequences were classified as Rhizaria and 1–17% of sequences were classified as Prymnesiophyceae and therefore designated as deriving from symbionts (Figure 2.4). Twenty-one symbiotic SVs were identified from within the acantharian samples, and each acantharian host contained 4–12 unique symbiotic SVs (mean = 8, standard deviation = 2) (Figure 2.5A). Phylogenetic analysis determined that symbiotic SVs belonged to four *Phaeocystis* clades: the *globosa* clade, the *cordata* clade, the *jahnii* clade, which has not previously been reported as a symbiont, and the uncultured Phaeo2 clade, which was discovered by Decelle et al. (2012a) as an acantharian symbiont near Okinawa (Figure 2.3). Additionally, four symbiotic SVs belonged to the genus *Chrysochromulina*, which had not previously been identified as a symbiont in clade E or F acantharians. The majority of symbiotic SVs in acantharians collected from the ECS belonged to the *Phaeocystis* clades *cordata*, *jahnii* and Phaeo2, and only 3 of these samples contained SVs belonging to *Chrysochromulina*. The opposite pattern was observed in samples collected near Catalina Island: the majority of symbiotic SVs in these samples were *Chrysochromulina spp.* However, all three samples from

Catalina Island hosted symbionts from the Phaeo2 *Phaeocystis* clade, which had previously only

been found in hosts collected near Okinawa Island (Decelle et al., 2012a). These results

demonstrate that acantharians simultaneously host multiple symbiont species and that

Phaeocystis-hosting acantharians can also host *Chrysochromulina spp*.
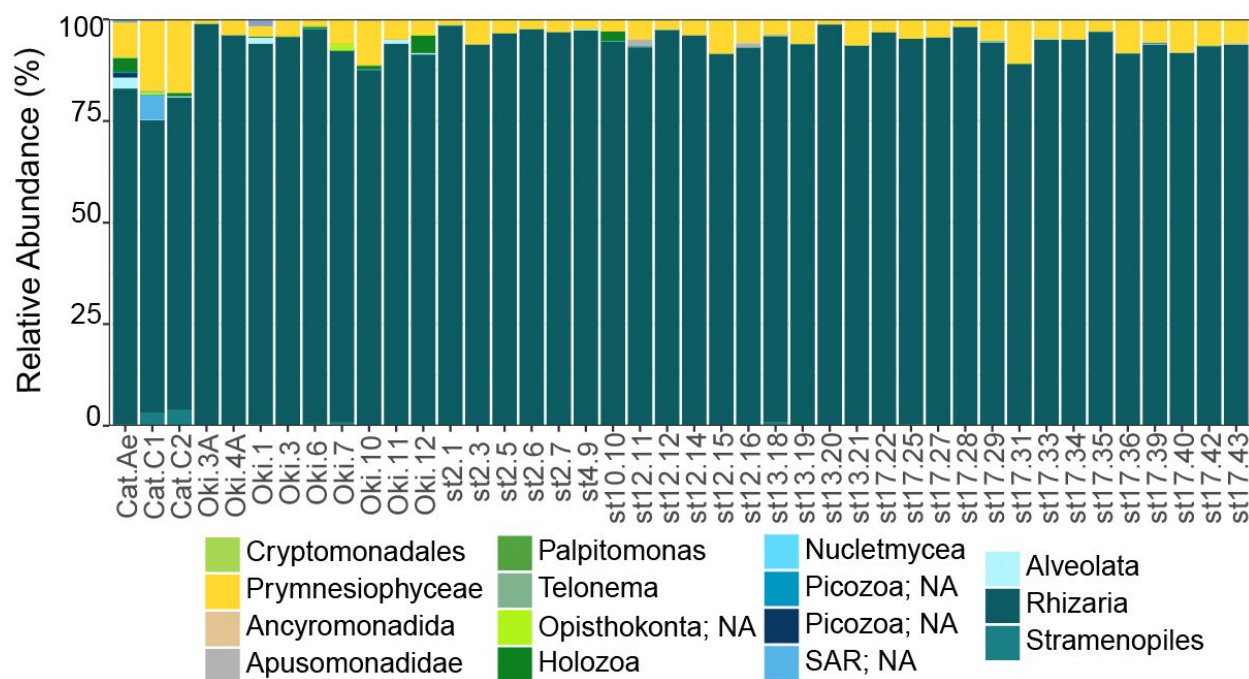
Figure 2.4. Relative abundance of Sequence Variants (SVs) in single acantharian holobionts. Each bar represents a single acantharian host and is labeled by collection location (st#: ECS cruise station number, Oki: Okinawa Island, Cat: Catalina Island) and sample ID. Relative abundance includes all SVs remaining after initial prevalence filtering. Within each host, 72–99% of sequences were classified as Rhizaria and 1–17% of sequences were classified as Prymnesiophyceae.
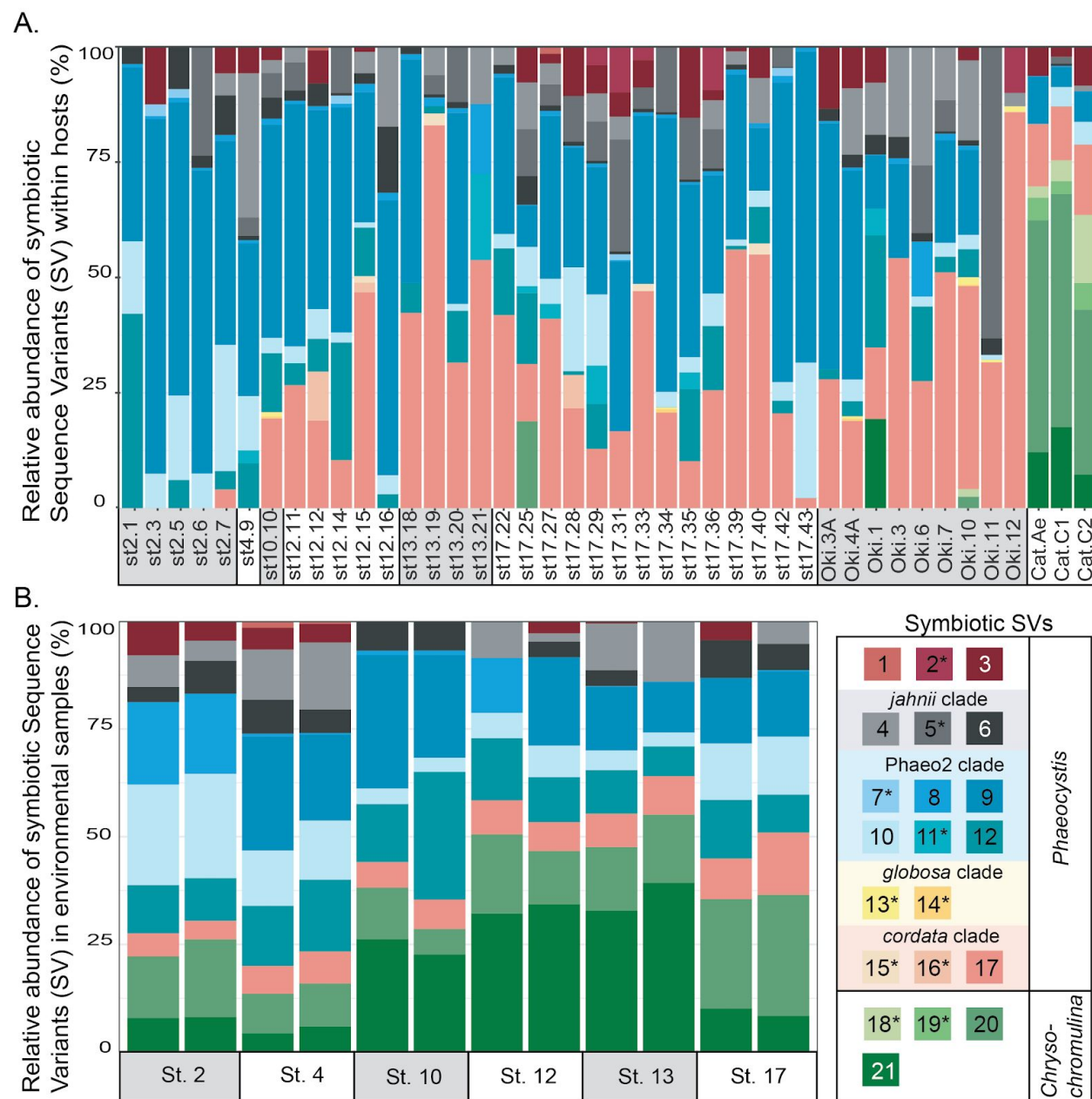
Figure 2.5. Relative abundance of symbiotic Sequence Variants (SVs) in individual acantharian hosts (A) and environmental samples (B). (A) Each bar represents a single acantharian host and is labeled by collection location (st.#: ECS cruise station number, Oki: Okinawa Island, Cat: Catalina Island) and sample ID. Individual acantharians contained 4–12 symbiotic SVs (mean = 8, standard deviation = 2). **(B)** Each bar represents an environmental replicate from ECS cruise stations. Symbiotic SVs are colored by clade: green is *Chrysochromulina*, blue is *Phaeocystis* clade Phaeo2, pink is *Phaeocystis cordata*, grey/black is *Phaeocystis jahnii*, and purple is *Phaeocystis*, but not placed with known clades. Ten of the symbiotic SVs were not found in environmental samples and are marked by an asterisk in the legend.

We also investigated the effect of host type and collection location on symbiont community composition within individual acantharians. Acantharian samples were each dominated by 1 of 5 unique Rhizarian SVs, all of which belonged to Acantharea clade F and represented 3 genera: *Amphibelone* (Host SVs 1 and 3), *Amphilonche* (Host SVs 2 and 4), and *Acanthometra* (Host SV5) (Figure 2.3). Host SV did not appear to determine the symbiont community in a Principal Coordinate Analysis (PCoA) plot based on Bray-Curtis distances between acantharian symbiont community compositions, but collection location did appear to play a role and the Catalina Island samples formed a distinct cluster in the PCoA plot (Figure 2.6). Statistical testing showed that collection location had a significant effect on symbiont community ($p = 0.001$, $R^2 = 0.47$) and in pairwise comparisons, only two comparisons—st. 17 to st. 12 and st. 13 to Okinawa—were not significantly different. Host SV also significantly affected symbiont community ($p = 0.009$, $R^2 = 0.19$), but pairwise comparisons showed that the significance was driven solely by the symbiont community associated with host SV 5, which was only found near Catalina Island, indicating that host effect was confounded by location.
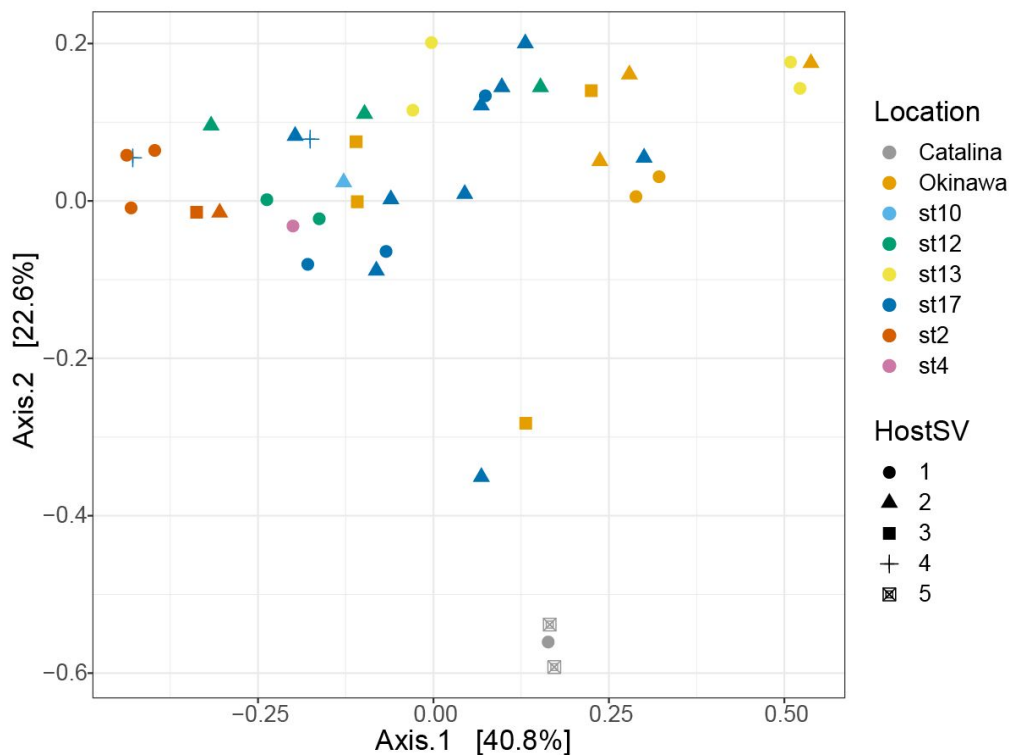
Figure 2.6. Principal Coordinates Analysis of Bray-Curtis distances between symbiont communities within individual acantharian hosts. Point shape corresponds to host sequence variant (SV) and the color corresponds to the collection location of the host acantharians (n = 42, 1–14 per site). The symbiont communities associated with acantharians collected near Catalina Island and from ECS station 2 form clusters, while communities from the other locations do not cluster separately. A PERMANOVA by location (excluding st. 4 and st. 10 due to insufficient sample size) on the Bray-Curtis distances between symbiont communities confirmed that collection location had a significant effect on symbiont community (p = 0.001, R^2 = 0.47 after 999 permutations). Host SV also had a significant effect (p = 0.009, R^2 = 0.19), but only SV 5 was significantly different when pairwise PERMANOVA comparisons were performed.

2.3.2. Comparison of intra-host and free-living symbiont communities

To determine whether the relative abundance of symbionts within hosts is a reflection of the relative abundance of available symbionts in surrounding water, we compared symbiont communities within acantharians collected from cruise stations in the ECS to environmental samples taken at the same time and place. 1,852,276 sequences were generated from environmental samples with 93,291–246,564 reads per sample. After quality filtering and feature

table construction, 645,325 sequences remained, with 35,691–89,163 sequences per sample. A

PCoA plot based on Bray-Curtis distances between the entire community for each environmental

sample confirmed that replicates from each location were more similar to each other than to

replicates from other locations (Figure 2.7). The environmental feature table was subset to

include only symbiotic SVs identified in acantharian samples, which comprised 1–3% of

sequences in environmental samples. Only 11 of 21 symbiotic SVs identified from acantharian

samples were also found in environmental samples (SVs 2, 5, 7, 11, 13, 14, 15, 16, 18, 19 were

missing) (Figure 2.5B). *Chrysochromulina* SVs 20 and 21, which were rare in the acantharian

samples, were among the most abundant symbiotic SVs in the environment, while *Phaeocystis*

cordata SV 17, which was one of the most abundant SVs in acantharian samples, was relatively

rare in the environmental samples (Figure 2.5A, B). Intra-host and environmental symbiont

communities clustered separately in a PCoA based on Bray-Curtis distances between samples

(Figure 2.8) and the observed difference between community compositions in the two sample

types was statistically significant ($p = 0.001$, $R^2 = 0.33$). These results indicate that the intra-host

community composition of symbiotic SVs is distinct from the relative availability of symbionts

in the surrounding environment.
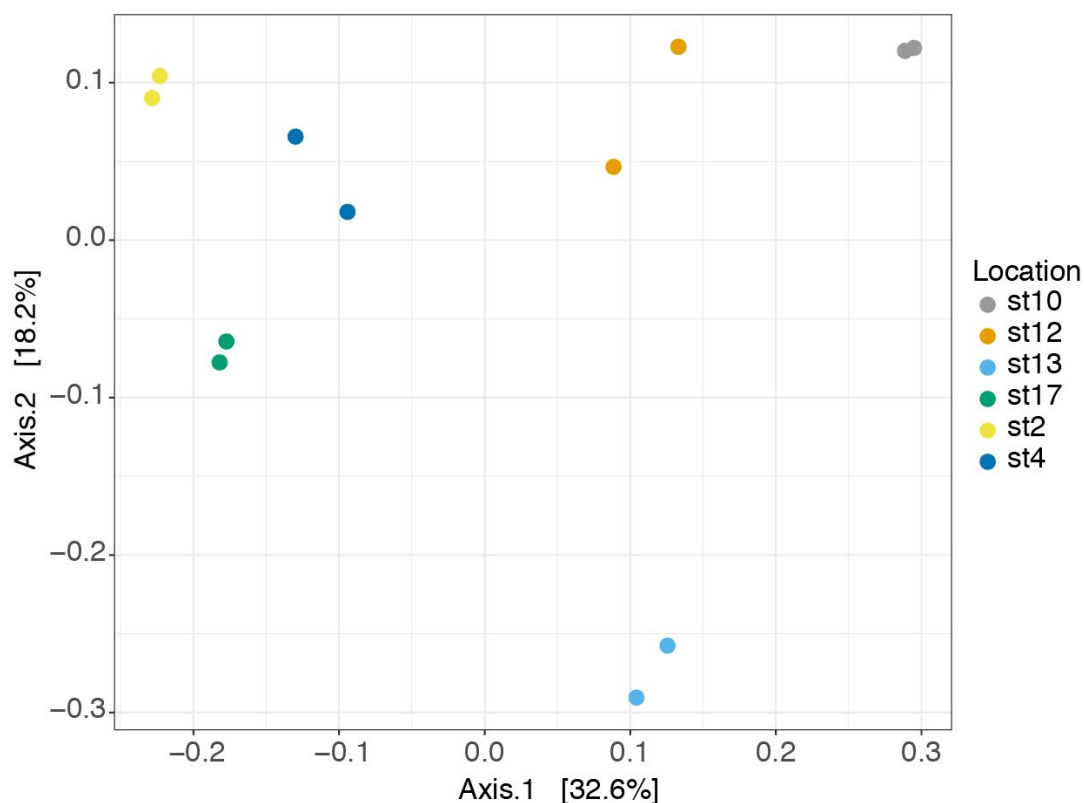
Figure 2.7. Principal Coordinates Analysis of Bray-Curtis distances between microbial Eukaryote communities at East China Sea (ECS) cruise stations. Point color corresponds to the 6 ECS cruise stations where surface water samples were collected when collecting acantharians. Bray-Curtis distances between microbial Eukaryote communities in ECS cruise station samples were calculated from entire microbial Eukaryote communities in the greater than 0.2 µm and less than 10.0 µm size fraction. Biological replicates of environmental microbial Eukaryote communities in each location are more similar to each other than to replicates from other locations.
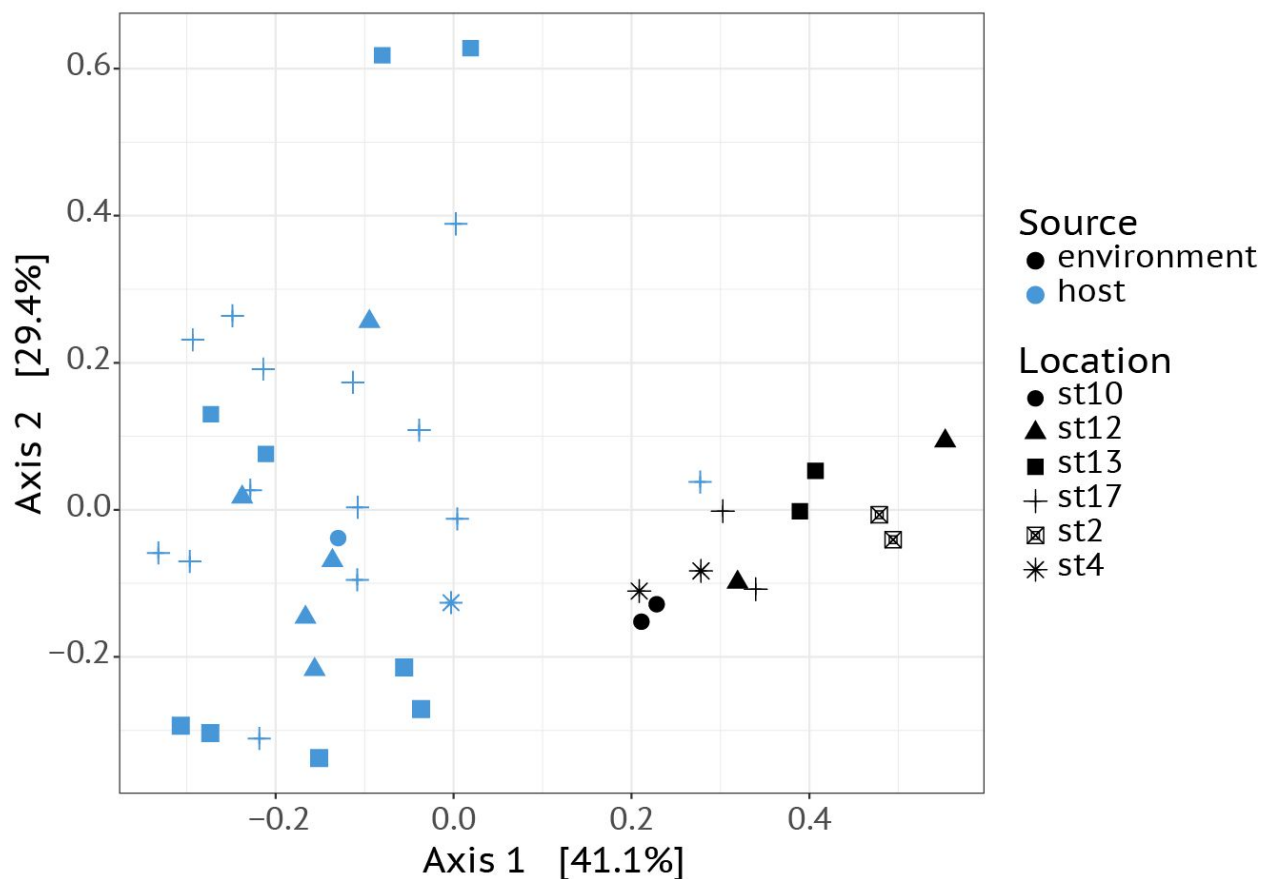
Figure 2.8. Principal Coordinates Analysis of Bray-Curtis distances between host-associated symbiont communities and free-living symbiont communities. The point color indicates sample type and point shapes represent sampling locations. Host-associated and environmental communities form two distinct clusters and the difference in community composition between the two sample types is statistically significant (PERMANOVA, *p* < 0.001, R^2 = 0.33, 999 permutations).

2.3.3. Visualization of host-associated symbionts and host digestive-organelles

By imaging symbiont chlorophyll autofluorescence with laser confocal microscopy, it was possible to clearly enumerate photosynthetic symbionts within hosts (Figure 2.9). The resulting images showed there were more individual symbiont cells within acantharians collected near Okinawa in April and May than there were unique symbiont SVs identified from those hosts (Table 2.1). Symbionts with the free-living phenotype (twin parietal plastids, cell diameter < 5

µm) and symbionts with the symbiotic phenotype (more than 2 plastids, enlarged central vacuole, cell diameter > 5 µm) were observed inside the same host cells (Figure 2.10A, B). By selectively staining host digestive-organelles with a fluorescent dye, host lysosomes and phagolysosomes and their proximity to symbionts were visualized (Figure 2.10C–E). Lysosomes were observed in both host exoplasm and endoplasm, while symbionts were only observed in host endoplasm. No symbionts were enclosed in phagolysosomes and lysosomes did not appear to be concentrated near symbiont cells.
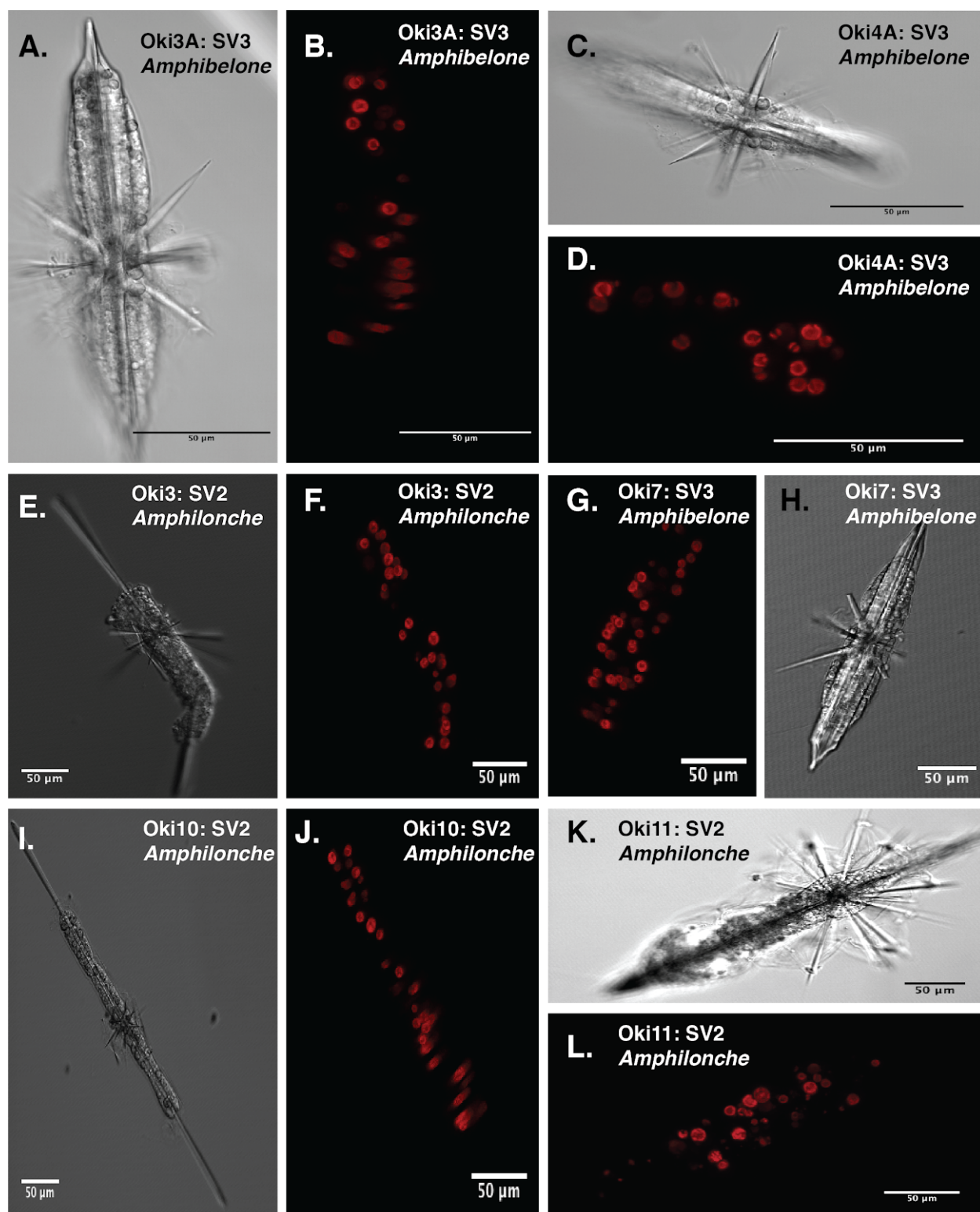
Figure 2.9. Laser confocal microscopy of acantharians collected near Okinawa in April and May 2017. Acantharians were imaged with a Zeiss LSM780 inverted laser scanning confocal microscope. Halogen light images (A, C, E, H, I, K) are single optical slices and fluorescent images (B, D, F, G J, L) are maximum projections of z-stacks spanning the entire

host, with symbiont chlorophyll autofluorescence colored red. Scale bars are 50 µm in all panels. Images are labeled by sample ID, host SV, and associated host genus.

Table 2.1. Number of acantharian symbionts visible by microscopy and unique symbiotic SVs identified per host. Symbionts within 6 acantharians collected in April (Oki.3A and Oki.4A) and May (Oki.3, Oki.7, Oki.10, Oki.11) from coastal waters near Okinawa, Japan were enumerated by visualizing chlorophyll autofluorescence with laser confocal microscopy. More symbiont cells were observed in each of the imaged hosts than unique SVs were detected from analyzing amplicon sequence data.

	Symbionts Visible in Micrograph	**Symbiotic SVs per Host**
Oki.3A	20	6
Oki.4A	17	10
Oki.3	29	10
Oki.7	38	8
Oki.10	28	12
Oki.11	24	5
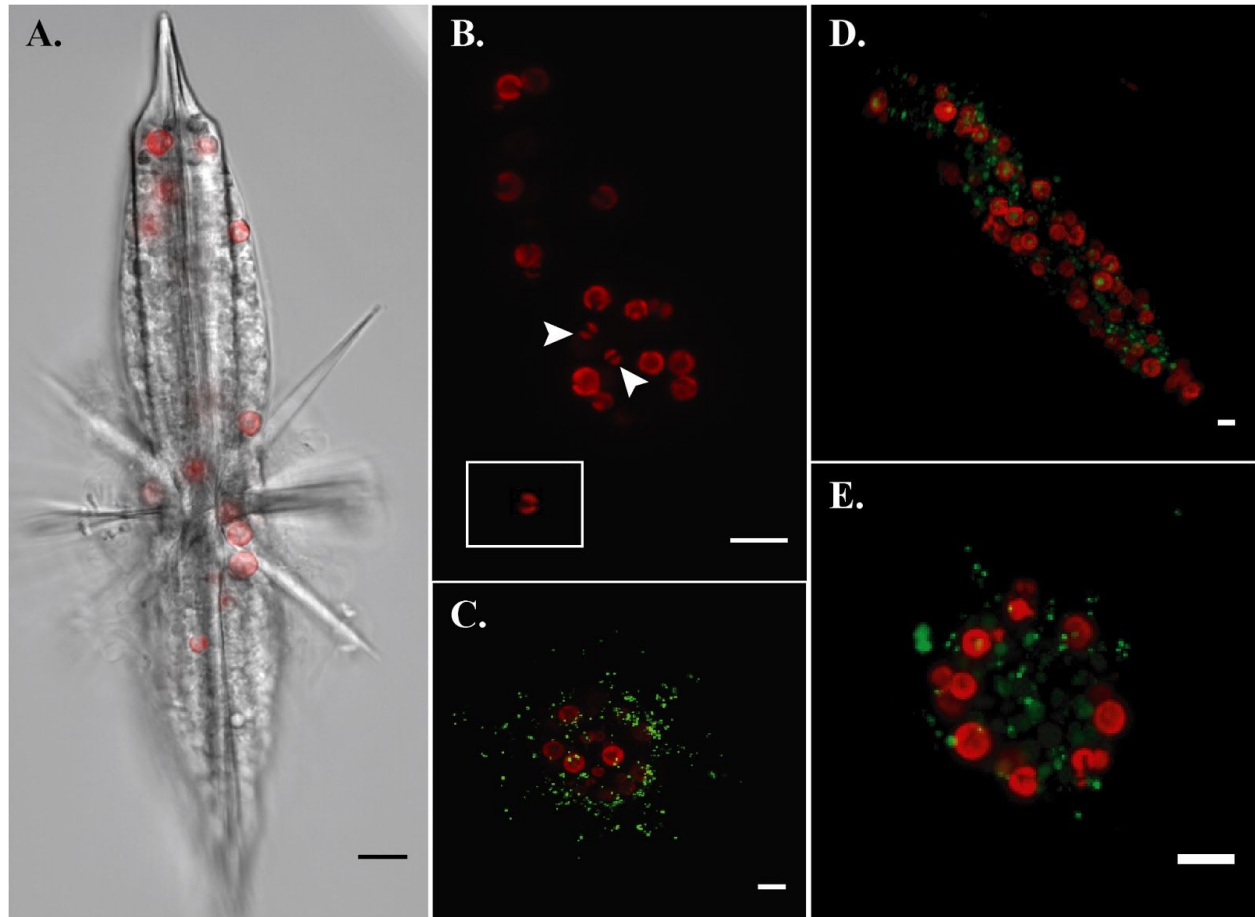
Figure 2.10. Fluorescent confocal microscopy of acantharians and their symbionts. (A) Single optical slice displaying autofluorescence of symbiont chlorophyll (red) and halogen light imaging of acantharian sample Oki.3A. (**B**) Maximum projection of a z-stack spanning the entire host sample Oki.4A, which contains 17 visible symbionts and 10 symbiotic sequence variants. Arrows indicate symbionts presenting the free-living phenotype—smaller cell diameter and two elongate, parietal chloroplasts—which is also visible in the inset image of a *Phaeocystis globosa* CCMP1528 cell in culture that was imaged following the same methods. Red fluorescence is chlorophyll autofluorescence. (**C–E**) Maximum projections of z-stacks spanning entire hosts collected near Okinawa in December 2017. Red fluorescence is from symbiont chlorophyll autofluorescence. Green fluorescent staining is LysoTracker Green, which selectively binds to low-pH digestive-organelles, including lysosomes and phagolysosomes. Symbionts are not held in phagolysosomes and lysosomes are not concentrated around symbionts, indicating that symbionts are not actively being digested. Scale bars are 10 μm in all panels.

2.4. Discussion

Acantharians are globally abundant but are especially so in low-nutrient subtropical gyres where they contribute to primary production by harboring intra-cellular algal symbionts (Michaels, 1991; Michaels et al., 1995). Despite their ecological importance, host-symbiont dynamics in acantharian photosymbioses remain largely unstudied. Decelle et al. (2012a) discovered that Acantharea-*Phaeocystis* symbioses are flexible in regard to symbiont species and hypothesized that the relationship is more akin to enslavement than mutualism (Decelle et al., 2012a; Decelle, 2013). If symbionts are manipulated to the extent that they are unable to reproduce, hosts will need to continuously recruit symbionts and should host diverse symbiont communities reflecting the relative availability of different symbionts. The amplicon sequencing results show that acantharians do simultaneously host multiple species of *Phaeocystis,* as well as *Chrysochromulina spp.*, but we found that the host-associated symbiont community does not simply mirror the free-living community. The microscopy results show that individual hosts harbor symbionts exhibiting the free-living phenotype as well as the symbiotic phenotype and demonstrate that symbionts are not being systematically digested. Together, these results suggest that hosts recruit symbionts more than once, but that they maintain recruited symbionts.

As acantharians increase in size, the number of symbionts they host also increases (Michaels, 1991). To accomplish this, acantharians could recruit microalgal partners early in development and nurture reproducing symbiont populations. Alternatively, acantharians could recruit more symbionts as they grow, which would likely lead to multiple species of symbionts coexisting within individual hosts. All 42 acantharians collected in this study hosted multiple species of *Phaeocystis* symbionts and several also hosted *Chrysochromulina spp.*, indicating that

acantharians recruit symbionts more than once. Acantharians in this study also contained more individual symbiont cells than unique Sequence Variants, which could mean that hosts recruit multiple symbiont cells from free-living populations during uptake events, that recruited symbionts divide within hosts, or both. The observation of symbionts with the free-living phenotype alongside symbionts with the symbiotic phenotype within single hosts can be seen as evidence that hosts continue to recruit new symbionts (Febvre and Febvre-Chevalier, 1979; Figure 2.10). Symbionts exhibiting the free-living phenotype within hosts could also represent recently divided cells rather than recently engulfed cells, but structural changes associated with mitotic division have not been observed in acantharian symbionts so far (Febvre and Febvre-Chevalier, 1979). The results presented here cannot exclude the possibility that symbionts reproduce *in hospite,* and symbionts are indeed known to reproduce within larger photosymbiotic rhizarians (Takagi et al., 2016), but continued symbiont uptake is evidence against intra-host symbiont division.

Since acantharians in this study simultaneously hosted multiple *Phaeocystis* and *Chrysochromulina* species, the intra-host symbiont community may be expected to reflect environmental symbiont availability. However, the relative abundance of symbiotic SVs within hosts was distinct from the relative abundance of these SVs in the surrounding environment. Additionally, acantharians sometimes hosted *Phaeocystis* genotypes (e.g. *P. jahnii* SV 5, *P. globosa* SVs 13 & 14) that were not detected in environmental samples collected from the same place. These results suggest acantharians may maintain symbionts for extended periods of time, beyond that required for external populations to respond to changing environmental conditions. This is feasible since *Phaeocystis* generation times can be as brief as 6.6 h and vary by species in

different light and temperature conditions (Jahnke, 1989), while acantharians likely survive for at least one month (Suzuki and Not, 2015). Intra-host symbiont communities may, therefore, be cumulative representations of all encountered environmental symbiont communities, rather than just a snapshot of the current community.

The observed differences between acantharian symbiont communities and the availability of symbionts could also indicate that hosts selectively uptake symbionts, especially since there were many more prymnesiophyte SVs found in the environment than were identified as symbionts. *Chrysochromulina* SVs were dominant in each acantharian sample collected near Catalina Island, but were only found in three of the acantharians collected from ECS cruise stations and from near Okinawa Island, despite being well-represented in all cruise station environmental samples. *Chrysochromulina spp.* may, therefore, make better partners in the Catalina Island environment compared to the Ryukyu Archipelago. *Chrysochromulina spp.* may also have been much more abundant in the waters near Catalina Island and therefore more likely to become symbionts, but we did not analyze the environmental community there. If acantharians select and concentrate environmentally rare symbionts, it increases the likelihood that our environmental sampling missed those symbionts and provides an alternative explanation as to why some symbiotic SVs were not observed in any environmental samples. It is also possible that the differences observed between intra-host and environmental symbiont communities could derive from different nucleic acid extraction methods used. However, lack of symbiont digestion by hosts provides additional support for extended symbiont maintenance *in hospite* (Figure 2.10)*,* even if recruitment is highly selective.

This study demonstrates for the first time that intra-cellular symbiont diversity exists in clade F acantharian photosymbioses. Acantharians in the clade B genus, *Acanthochiasma*, also host multiple symbiont types, including *Chrysochromulina spp.* and several dinoflagellate genera (Decelle et al., 2012b). Radiolarian and foraminiferan (Rhizaria) hosts also host dinoflagellates and prasinophytes or prymnesiophytes (Gast and Caron, 2001). Theoretically, simultaneously hosting multiple symbiont species or genotypes is ineffective and should negatively impact host fitness since different symbionts would compete for space and resources within hosts (Douglas, 1998). Planktonic hosts can be transported long distances and may, therefore, experience larger environmental gradients on shorter time-scales than stationary, benthic photosymbiotic hosts. This could make hosting a diverse community of symbionts more effective for planktonic hosts, especially if some symbionts perform better under different conditions. Indeed, different *Phaeocystis* species have different light and temperature optima (Jahnke, 1989) and different strains of a single species have varying abilities to utilize different nitrogen sources (Wang et al., 2011). Drifting buoys in the Global Drifter Program (Lumpkin et al., 2013) passing our sampling sites in spring, the season we sampled, traveled an average of 3.66° latitude in 30 days (n = 42), the estimated minimum survival time of acantharians (Suzuki and Not, 2015), suggesting they can travel at least this far in their lifetime. While the associated mean temperature gradient of 1.93°C is smaller than experimental temperature gradients shown to differentially influence *Phaeocystis spp.*, changes in day length and irradiance may affect photosynthetic output of *Phaeocystis* strains differently (Jahnke, 1989).

The possible fitness benefit for symbionts associated with acantharians remains enigmatic, but the evidence we present for extended symbiont maintenance allows that

Phaeocystis could glean some advantage from the symbiosis. Acantharian symbionts were not enclosed in phagolysosomes in host cells imaged in this study, which could be due to symbionts escaping phagosomes (Jamwal et al., 2016) or a failure of host lysosomes to fuse with symbiont-containing phagosomes (Hohman et al., 1982; Sibley et al., 1985). Electron microscopy suggests that there is a host membrane surrounding acantharian symbionts (Febvre and Febvre-Chevalier, 1979), so it is more probable that lysosomes do not fuse with symbiont-containing phagosomes and they are instead symbiosomes or long-term perialgal vacuoles. Symbiont signaling prevents lysosomes from fusing with symbiosomes housing *Chlorella* symbionts in *Paramecium* (Kodama et al., 2011) and *Hydra* (Hohman et al., 1982) hosts. Complex signaling between hosts and symbionts leading to development of symbiosomes and symbiont differentiation is suggestive of coevolution (Hinde and Trautman, 2001). If *Phaeocystis* also actively prevents lysosomes from fusing with phagosomes, then *Phaeocystis* has adapted to avoid digestion and perhaps to promote a stable symbiosis. Although our results suggest that acantharians maintain symbionts, we cannot rule out that they digest symbionts when stressed or before releasing reproductive swarmers. Acantharian swarmers, which are believed to be gametes, do not contain symbionts and although asexual reproduction has been observed in one group of clade B acantharians, it has not been reported for photosymbiotic acantharians in clades E or F (Decelle and Not, 2015). It is not yet known whether symbionts are digested or released during swarmer production (Decelle and Not, 2015), but nassellarian and spumellarian radiolarians release viable symbionts when producing swarmers (Yuasa and Takahashi, 2016). Additionally, it remains an open question whether released symbionts are reproductively competent outside the host (Decelle et al., 2012a).

The results presented here demonstrate that clade F acantharians simultaneously associate with multiple *Phaeocystis* and *Chrysochromulina* species, providing further evidence that the Acantharea-*Phaeocystis* photosymbiosis is relatively flexible. The results further suggest that symbionts escape host digestion for extended periods, but whether symbionts are capable of reproducing *in hospite* or after release and whether they benefit from the relationship is still undetermined. Until acantharians can be maintained for prolonged periods under laboratory conditions, it will remain challenging to elucidate many aspects of the host-symbiont dynamics in this system. LysoTracker dyes can be utilized to track symbiosome conditions for as long as hosts survive and species- or genotype-specific fluorescent probes may be used to investigate whether different species or genotypes are differentially transformed into the symbiotic phenotype or are compartmentalized within the host. If symbionts with a single 18S rRNA gene Sequence Variant (or a unique combination of SVs) are co-localized, it would provide evidence for *in hospite* reproduction. Efforts to culture symbionts from *Phaeocystis*-hosting acantharians should continue. If successful, it would demonstrate unequivocally that symbionts maintain reproductive capacity and cultured symbionts would be an invaluable resource for comparative genomics and transcriptomics. Differential gene expression analysis could then be utilized to investigate physiological shifts in symbiotic *Phaeocystis* or *Chrysochromulina* within hosts compared to free-living cells of symbiotic strains and could illuminate mechanisms of host-symbiont interaction. Further investigation into whether symbionts benefit from this relationship will be important to understanding host-symbiont dynamics in this and other protistan photosymbioses.

2.5. Supplementary Material
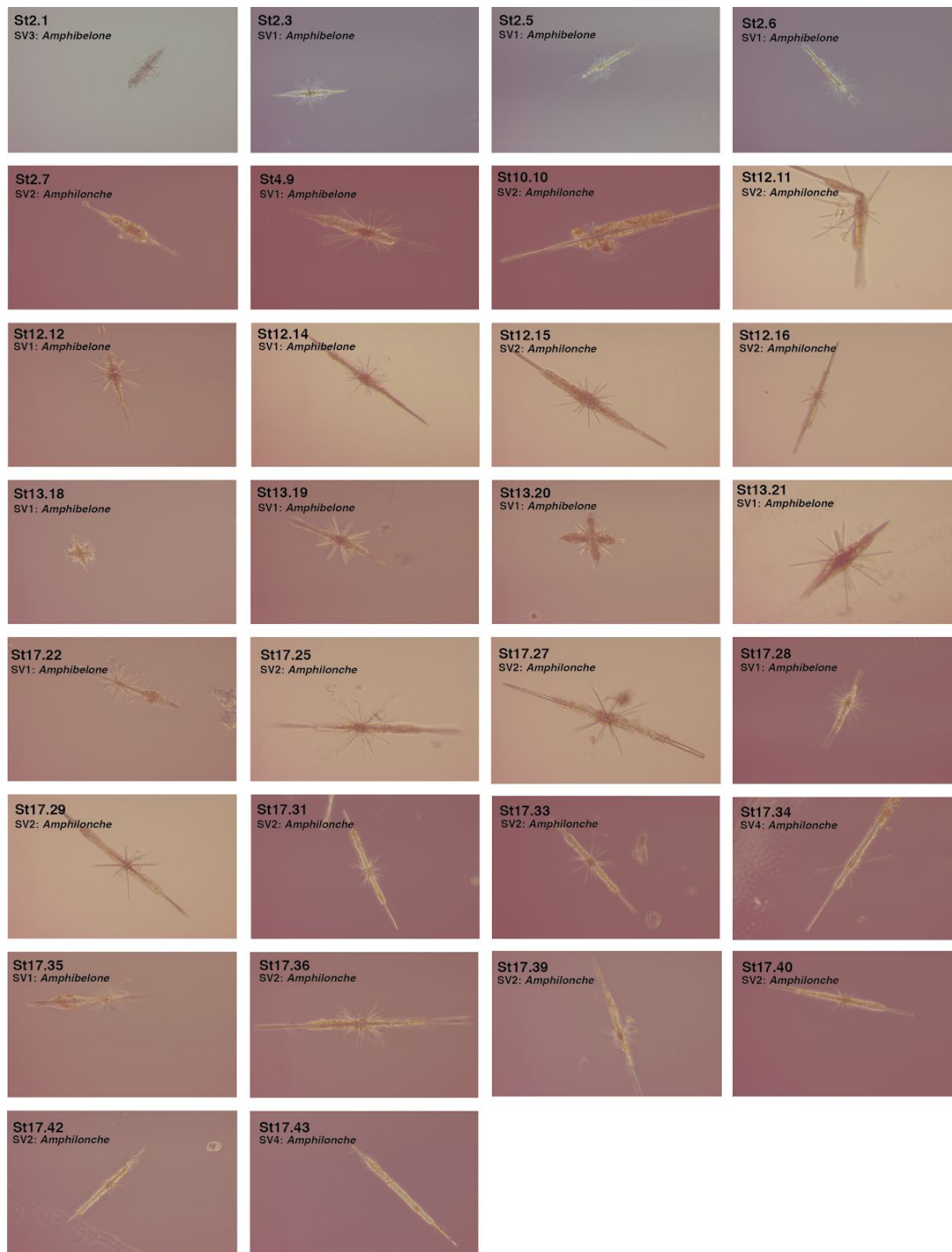
Figure S2.1. Light microscopy images of acantharians collected from ECS cruise stations in May and June 2017. Samples are labeled by the station and sample number. The Host SV

and the associated Acantharea genus are indicated below sample names. All samples were imaged at 200x on a Zeiss PrimoVert inverted light microscope. Images are not cropped or otherwise distorted.
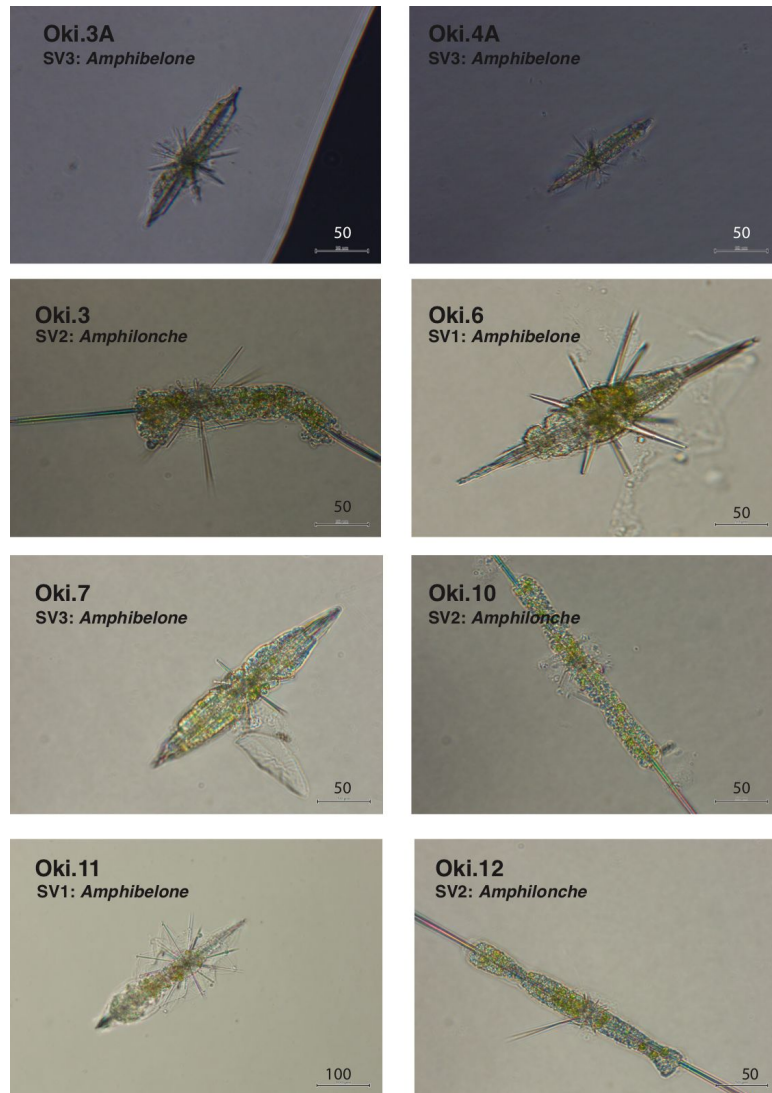
Figure S2.2. Light microscopy images of acantharians collected near Okinawa in April (Oki.3A and Oki.4a) and May 2017 (Oki.3, 6, 7, 10, 11 & 12). The Host SV and associated Acantharea genus are indicated below sample names. Acantharians were imaged with an Olympus CKX53 inverted light microscope. Scale bars are in micrometers.

Chapter Three: Symbiont maintenance and host-control in Acantharea-*Phaeocystis* photosymbioses revealed through single-holobiont transcriptomics

Abstract

Photosymbioses, like those between acantharians (Rhizaria) and *Phaeocystis* (Haptophyta), contribute significantly to primary production in low nutrient ecosystems. These relationships are generally assumed to be mutualistic, with hosts providing nitrogen and phosphorus and symbionts providing organic carbon in return. It is unknown, however, whether or not *Phaeocystis* symbionts are released or can escape from hosts or if they maintain reproductive capacity, calling into question whether the relationship can really be mutualistic. Furthermore, almost nothing is known about mechanisms for host control or nutrient transfer in rhizarian photosymbioses. This study aimed to elucidate cellular pathways important to symbiosis for *Phaeocystis* by comparing gene expression in free-living, cultured *Phaeocystis* with gene expression in symbiotic *Phaeocystis* within individual acantharian holobionts. The results provide evidence for strong host manipulation of symbionts at the molecular level. Key genes in the mitogen activated protein kinase (MAPK) signaling pathway are downregulated in symbiotic *Phaeocystis*, potentially leading to the downregulation of DNA replication and cell-cycle progression genes that was measured. While the cell-cycle is arrested in symbiotic *Phaeocystis*, chloroplast division genes continue to be expressed, causing the proliferation of plastids and enhancing photosynthesis. Symbionts forgo storage molecule biosynthesis and seem to export

small organic nitrogenous compounds to the host, while the host supplies urea and ammonium to fuel photosynthesis. By keeping symbionts flush with nutrients and utilizing a cell signaling pathway to manage symbiont population size, acantharian hosts exert fine control over symbionts and ensure that symbiont carbon fixation is maintained. These findings demonstrate strong interactions between hosts and symbionts in this system.

3.1. Introduction

Photosymbiosis—where photosynthesizing cells reside inside heterotrophic hosts—is ubiquitous among marine microbial eukaryotes and is especially common among Rhizaria (Foraminifera, Radiolaria, and Acantharea) (Not et al., 2016). These relationships make significant contributions to primary production in low nutrient ecosystems; photosymbiotic Rhizaria can contribute up to 4 orders of magnitude the primary production as occurs in an equivalent volume of oligotrophic seawater (Caron et al., 1995). Furthermore, because hosts often possess dense biomineralized skeletons, they can facilitate carbon sequestration from surface waters to the deep sea (Michaels, 1991; Michaels et al., 1995; Martin et al., 2010; Gutierrez-Rodriguez et al., 2019). Nutrient coupling between hosts and symbionts, such that symbionts relinquish fixed organic carbon to hosts and hosts provide inorganic nutrients (eg. N, P), is generally assumed, leading to photosymbioses being considered mutualistic. It is difficult, however, to demonstrate whether such symbioses confer advantages on symbionts, particularly because most Rhizaria cannot be cultured or even maintained in laboratory conditions. But, as more genetic and observational data become available, it seems increasingly likely that at least some of these relationships, specifically among acantharians, represent a more one-sided

arrangement where symbionts are being exploited (Decelle et al., 2012a; Decelle, 2013; Mars

Brisbin et al., 2018).

Photosymbiotic acantharians host algal symbionts from the haptophyte genus

Phaeocystis, but the dominant *Phaeocystis* species in symbiosis varies geographically (Decelle et

al., 2012a) and individual acantharians simultaneously host multiple strains and species of

Phaeocystis (Mars Brisbin et al., 2018). The inclusion of multiple symbiont strains inside hosts

indicates that hosts incorporate symbionts throughout their life-span, which could suggest

continued symbiont turn-over by host digestion and would not benefit symbionts. However,

microscopic observations demonstrate that symbionts are not systematically digested by hosts

(Mars Brisbin et al., 2018), which opens the question as to how hosts manage symbiont

populations. *Phaeocystis* symbionts undergo extensive structural remodeling *in hospite*;

symbiotic cells become enlarged with increased plastid abundance and altered plastid

morphology (Febvre and Febvre-Chevalier, 1979; Decelle et al., 2019). Cytokinesis may,

therefore, be blocked in symbiotic cells, causing them to accumulate chloroplasts and increase

photosynthetic output while simultaneously preventing symbionts from over-growing hosts. If

cytokinesis is indeed blocked in symbionts, whether the transformation is reversible is a key

question in determining if the relationship is mutually beneficial. For symbionts to benefit from

the relationship, they must reproduce following release or escape from hosts. Symbionts isolated

directly from hosts have not yet been successfully brought into culture, suggesting that this

symbiosis may be an evolutionary dead-end for symbionts (Decelle, 2013). As such, this

relationship may represent an early intermediary step between photosymbiosis and a more

permanent incorporation of photosynthetic machinery, making the elucidation of involved

mechanisms particularly interesting.

Comparable, but less extensive and reversible, symbiont remodeling has been observed in

other photosymbioses, including dinoflagellate symbionts of corals (Blank, 1987; Gates et al.,

1995), radiolarians, and foraminifera (Gast and Caron, 1996, 2001; Shaked and de Vargas, 2006;

Probert et al., 2014). Specifically, *Symbiodinium* dinoflagellates lose their flagella, transition

from a peanut to a coccoid shape, increase photosynthetic rate, and release more photosynthate

when living in cnidarian tissue than when free-living (Blank, 1987; Gates et al., 1995). Likewise,

Brandtodinium and *Pelagodinium* dinoflagellates lose their flagella and undergo similar

morphological transformations when living as symbionts in Polycystine radiolarians or in

foraminifera (Gast and Caron, 1996, 2001; Shaked and de Vargas, 2006; Probert et al., 2014).

Based on the phenotypic transformations observed in photosymbioses, hosts are likely

manipulating symbionts to match their needs, either directly through biochemical signaling or

indirectly by controlling the microenvironment surrounding the symbiont. Nutrient limitation is

suspected to be used in order to limit symbiont population growth (Xiang et al., 2020), but a

homogenate of cnidarian host tissue can also induce phenotypic changes in cultured

Symbiodinium, suggesting hosts may deploy chemical signaling as well (Gates et al., 1995;

Koike et al., 2004). Furthermore, experiments have shown that both free amino acids (Gates et

al., 1995) and lectins (Koike et al., 2004) induce at least some of the phenotypic changes

observed in symbiotic *Symbiodinium*. Host-homogenate studies with Rhizaria are not feasible

due to their small size and the difficulty in separating symbionts from host tissue, especially for

acantharians, but investigating gene expression in symbiosis presents an alternative method for

probing the mechanisms involved in symbiont manipulation among microbial eukaryotes (Balzano et al., 2015; Liu et al., 2019).

The inclusion of multiple symbiont strains and species in individual acantharian hosts complicates comparisons between symbiont gene-expression in pooled acantharian samples to free-living symbiont cells, since biological replicates of pooled hosts may contain more or less of a particular symbiont strain. As a result, transcriptomic evaluation of single holobionts is preferred, but is technically challenging. In this study, mRNA was extracted and sequenced from 16 individual acantharians and gene expression in symbiotic *Phaeocystis* was compared with free-living *Phaeocystis* by further sequencing mRNA from biological replicates of two *Phaeocystis* symbiont species available in culture (*P. cordata* and *P. jahnii*). Differences in symbiotic expression patterns between the two species are expected, particularly because *P. jahnii* produces mucilaginous, multicellular colonies and *P. cordata* does not (Zingone et al., 1999). However, by focusing on shared changes between the two species while in symbiosis, it should facilitate interpretation of results and bolster confidence in the conclusions. Specifically, this study aims to elucidate molecular mechanisms for host control of symbiotic *Phaeocystis* populations and nutrient transfer between hosts and symbionts.

3.2. Materials and Methods

3.2.1. Acantharian collection

Individual acantharians (n = 16) were collected from 5 sampling stations (n = 1–7 per station) in the Okinawa Trough (East China Sea) in May and June 2017 during the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) MR17-03C cruise (Figure 2.1A).

Plankton samples were collected by passing unfiltered seawater pumped from the sea surface through a 100-µm-mesh, hand-held plankton net (Rigo) and were observed under a dissecting microscope. Individual acantharians were transferred by glass micropipette to clean Petri-dishes, rinsed with 0.2-µm-filtered seawater several times, until all visible contaminants were removed, and incubated for 0.5–2 hr for additional self-cleaning. Each acantharian was imaged with inverted light microscopy (Zeiss Primovert, Figure S2.1) before being transferred to a maximum recovery PCR tube (Axygen) with the smallest possible volume of accompanying seawater. Transfer success was confirmed by microscopy before adding 30 µL of RLT-plus cell-lysis buffer (Qiagen) and immediate flash-freezing with liquid nitrogen. Samples were stored at −80°C until RNA extraction.

3.2.2. *Phaeocystis* culture conditions

Three biological replicates each of *Phaeocystis cordata* CCMP3104 and *Phaeocystis jahnii* CCMP2496 were prepared by inoculating 45 mL of sterile L1 media with 1 mL stock culture in 50-mL Erlenmeyer flasks. Replicates were placed on a gently rotating twist mixer in a plant growth chamber with cool white fluorescent lamps (CLE-305, TOMY) set to 22°C with light level 4 and a 12:12 day:night ratio. A HOBO temperature and light logger (Onset) was kept in the growth chamber during culturing: the daytime temperature was 21°C with about 1900–2000 Lux ($\sim$30 µmol m^{-2} s^{-1}) light intensity, and the nighttime temperature was 22°C. Positions of replicates were rotated daily to prevent position in the chamber from systematically affecting replicates. Four days after initiating cultures, 1 mL of each replicate was transferred to 45 mL of sterile L1 media in a clean 50-mL flask. The culture conditions were then repeated,

allowing for acclimation to the culture conditions. Algal cells were harvested for RNA extraction on the fourth day of the second culture round by filtering entire culture volumes through polytetrafluoroethylene (PTFE) filters (0.45-µm pore-size, Millipore) under gentle vacuum. Filters were immediately flash frozen in liquid nitrogen and stored at −80°C until RNA extraction.

3.2.3. RNA extraction and sequencing library preparation

RNA extractions from single acantharian holobionts were accomplished by modifying methods of Trombetta et al. (2014). Samples were thawed over ice, vortexed twice (10 s, speed 7, Vortex-Genie 2), and then incubated at room temperature for 5 min to fully lyse cells. Agencourt RNAClean XP magnetic beads (Beckman Coulter) were added to each sample at a 2.2:1 V:V ratio and fully mixed by pipette prior to a 30-min incubation to bind all RNA to the magnetic beads. After two 80% ethanol washes, RNA was eluted from the beads in 11 µL of a custom elution buffer (10.72 µL nuclease-free water, 0.28 µL RNAase inhibitor) and 10.5 µL of eluted RNA were directly processed following the single-cell protocol for the SMART-seq v4 Ultra Low Input Kit (Clonetech/Takara) with 18 cycles in the primary PCR.

RNA was extracted from *Phaeocystis* cells collected on PTFE filters by following the manufacturer's protocols for the MoBio PowerWater RNA extraction kit (Qiagen) including the optional initial heating step. Extracts were diluted so that 10 ng of RNA were used from each sample as input with the SMART-seq v4 Ultra Low Input RNA Kit. Each *Phaeocystis* sample was additionally supplemented with 2 µL of a 1:10,000 dilution of External RNA Controls Consortium (ERCC) spike-in mix one (Ambion) as an internal quality control. ERCC spike-in

was only added to *Phaeocystis* libraries, and not single holobiont libraries, because RNA concentrations in single-holobiont libraries were extremely low and the spike-in would, therefore, take up too much of the sequencing resources for those samples. The SMART-seq total RNA protocol was then followed with 12 cycles in the primary PCR.

The quality and concentration of cDNA from acantharian holobionts and *Phaeocystis* cultures were assessed before continuing with the manufacturer's protocols for the Nextera XT DNA Library Prep Kit (Illumina). Sixteen acantharian holobiont and six *Phaeocystis* culture cDNA libraries were submitted to the Okinawa Institute of Science and Technology DNA Sequencing Section. Eight acantharian libraries were pooled for 2x150bp paired-end sequencing on four lanes of an Illumina HiSeq4000 flow-cell and the second eight acantharian libraries and the six *Phaeocystis* libraries were pooled and sequenced on four lanes of an additional HiSeq4000 flow-cell.

3.2.4. *Phaeocystis* reference transcriptome assembly and annotation

Adapters were trimmed from sequencing reads and low-quality reads were filtered and discarded with trimmomatic software (Bolger et al., 2014). Reads mapping to ERCC reference sequences were removed to prevent internal standard sequences from being included in the transcriptome assemblies. The pre-processed *Phaeocystis cordata* and *P. jahnii* libraries were then assembled into *de novo* transcriptomes using the Trinity software (v2.8.4, Grabherr et al., 2011) and these assemblies were used as references for read mapping and counting prior to differential gene expression testing. The assemblies were deduplicated by removing reads with 95% similarity using CD-HIT-EST (Fu et al., 2012). While the SMART-seq kit employs poly-A

priming to target eukaryotic mRNA and reduce the amount of ribosomal and bacterial RNA present in sequencing libraries, some RNA deriving from these sources remain, especially in low-input samples. Sequences of bacterial origin were removed from transcriptomes by performing a blastn query against the NCBI nr-nt nucleotide database (downloaded March 2018, ncbi-blast v2.6.0+, Camacho et al., 2009) and parsing results to identify and remove bacterial contigs. The final assemblies were assessed for completeness by determining how many eukaryotic and protistan Benchmarking Universal Single-Copy Orthologs (BUSCO v3, Simão et al., 2015) were present in the transcriptomes with the HMMER3 software (Eddy, 2010). BUSCOs provide a method to quantitatively assess the quality of a transcriptome in terms of gene content: more complete transcriptomes contain more full-length BUSCOs, which are well-conserved genes appearing only once in complete, representative genomes for each BUSCO group. The *P. cordata* CCMP3104 BUSCO scores were compared with scores for the *P. cordata* RCC1383 transcriptome that was assembled as part of the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP) (Keeling et al., 2014). The *P. jahnii* transcriptome had not been previously sequenced and hence an assembly was not available for comparison.

The *Phaeocystis* transcriptomes were annotated using two different databases and functional annotation methods: Pfam (Finn et al., 2014) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2016). Pfam annotation was accomplished with the dammit software (Scott 2018), which wraps Transdecoder to translate transcriptome contigs to the longest possible amino acid sequence (Haas et al., 2016) and HMMER to assign Pfam protein homologs to sequences (Eddy, 2010). After discarding annotations with e-values greater

than 1E-5, the Pfam annotation with the lowest e-value was selected for each contig. The Pfam annotations were matched to corresponding Gene Ontology (GO) terms using the Gene Ontology Consortium's Pfam2GO mapping (geneontology.org/external2go/pfam2go, version 07/14/2018, Mitchell et al., 2015). KEGG Orthology (KO) annotation was performed using the GhostKOALA tool with the Transdecoder translated amino acid sequences (kegg.jp/ghostkoala, 05/21/2019, Kanehisa et al., 2016). The KEGG Orthology numbers were then used to access the KEGG API (kegg.jp/kegg/rest/keggapi.html, July 2019) and determine the KEGG pathways to which transcriptome contigs belonged. Additional annotations for genes of interest were acquired by blasting the *Phaeocystis* transcriptomes against the *Emiliania huxleyi* reference genome (Read et al., 2013).

3.2.5. Differential gene expression analysis

Quality filtered sequences from each of the sixteen acantharians were mapped to one or both of the reference transcriptomes, depending on symbiont presence in holobiont samples based on rRNA gene sequence analysis (Figure 3.1), but holobiont reads were only allowed to map to one reference transcriptome to avoid double counting (n = 12 acantharian holobionts with reads mapping to *P. cordata*, and n = 13 for *P. jahnii*). Sequences from *Phaeocystis* culture replicates were mapped to corresponding reference transcriptomes. All read mapping and quantifying steps were accomplished with the Salmon software (Patro et al., 2017). Counts for each sample were imported into the R statistical environment with tximport (Soneson et al., 2015). Differential gene expression testing between symbiotic and free-living replicates was performed with the DESeq function in the Bioconductor package DESeq2 (Love et al. 2014).

Genes that were differentially expressed with a False Discovery Rate (FDR) adjusted *p*-value (padj) less than 0.05 were considered statistically significant and included in gene set enrichment testing.
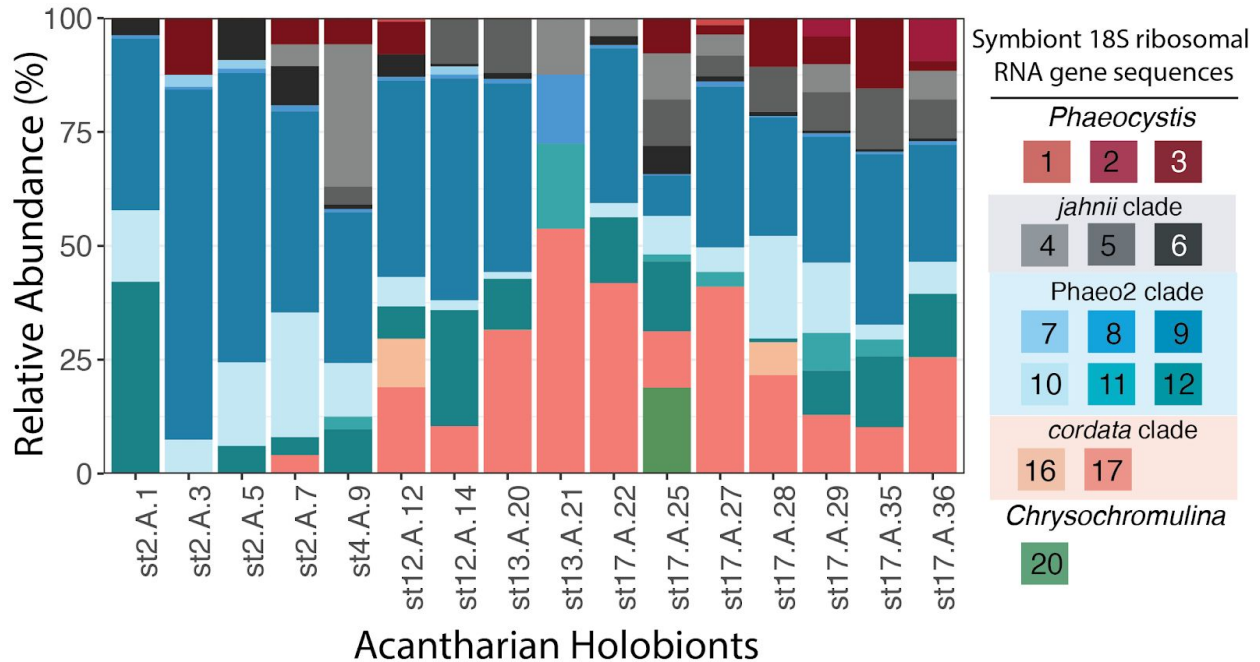
Figure 3.1. Intra-host symbiont communities in acantharian holobionts for which RNA-sequencing libraries were prepared. Symbiont community composition is based on 18S rRNA gene amplicon sequence variants. Each color represents an individual sequence variant. Sequence variants represented by similar colors clustered together on a phylogenetic tree with sequences from the main *Phaeocystis* clades, as indicated in the legend (Figure 2.2). Maroon sequences did not directly cluster with any cultured *Phaeocystis* strains or environmental *Phaeocystis* sequences from GenBank, but were most closely related to *Phaeocystis jahnii*.

3.2.6. Gene set enrichment testing

GO term enrichment among significantly up- and downregulated genes was determined with a hypergeometric test in the R package GOstats (Falcon and Gentleman, 2007). GOstats accommodates user-defined GO annotations, which are necessary when studying non-model organisms like *Phaeocystis*. Enrichment of KEGG pathways among up- and downregulated genes was tested for by applying linear model analysis with the kegga function from the

Bioconductor package edgeR (Robinson et al., 2010). GO terms and KEGG pathways were considered significantly enriched when the statistical test resulted in a *p*-value less than 0.05.

3.3. Results

3.3.1. Sequencing and quality control

Sequencing for this project produced over 2.9 billion read pairs with 78–350 million read pairs per sample. The sequencing data are available from the NCBI Sequence Read Archive (SRA) with accession number PRJNA603434. Following quality filtering with Trimmomatic, 2.5 billion read pairs remained, with 62–305 million read pairs per sample. Reads mapping to ERCC reference sequences (Cronin et al., 2004) were counted with RSEM software for each sample (Li and Dewey, 2011) and counts were further analyzed in the R statistical environment (R Core Team, 2018). To assess the relationship between ERCC sequence read counts and their original concentrations in the ERCC standard, the log of the observed FPKM (Fragments Per Kilobase per Million reads) was plotted against the log of the initial concentrations for each standard sequence. Linear regressions were fit for each sample and R^2 values ranged from 0.883–0.926 (Figure 3.2). The strong correlation between observed FPKM and initial concentration for ERCC sequences indicates minimal bias was introduced during PCR amplification, library preparation, and sequencing.
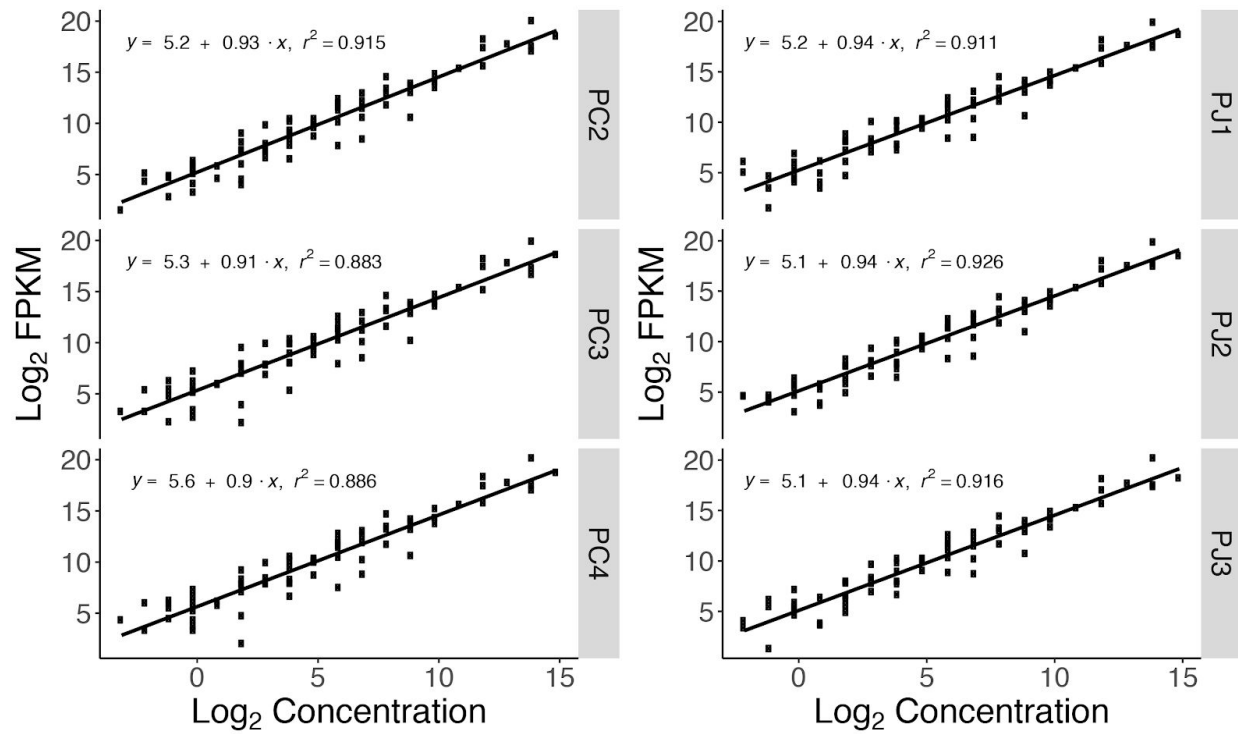
Figure 3.2. Correlation between External RNA Controls Consortium (ERCC) standard sequence initial concentrations and sequence counts from *Phaeocystis cordata* and *Phaeocystis jahnii* culture replicates. Quality-filtered sequences for each culture replicate (*P. cordata* replicates: PC2–4, *P. jahnii* replicates: PJ1–3) were mapped to ERCC standard sequences and read counts were determined with RSEM software. Fragments Per Kilobase per Million reads (FPKM) were log transformed and plotted against the log transformed initial ERCC sequence concentrations. A linear model was fit for each culture replicate in the R statistical environment and the resulting regression line, equation and R-squared are displayed on each plot panel.

3.3.2. Transcriptome assembly and annotation

The *P. cordata* CCMP3104 transcriptome assembled for this study is made up of 41,629 contigs and had an N50 of 989 bp; the *P. jahnii* CCMP2498 transcriptome is made up of 90,009 contigs and had an N50 of 1,351 bp (Table 3.1). The *P. cordata* transcriptome includes complete sequences for 74% of eukaryotic BUSCOs, whereas the *P. cordata* (RCC1383) transcriptome that was assembled for the MMETSP included complete sequences for only 46.5% of eukaryotic

BUSCOs. The *P. jahnii* transcriptome includes complete sequences for 81.2% of eukaryotic

BUSCOs (Figure 3.3). The *P. cordata* and *P. jahnii* transcriptomes had similar annotation rates,

with 41.5 and 35.1%, respectively, of contigs being assigned high confidence Pfam annotations.

Similarly, 18.5% and 20.3% of *P. cordata* and *P. jahnii* contigs were assigned KEGG orthology

annotations (Table 3.1).

Table 3.1. Assembly and annotation summary statistics for *Phaeocystis cordata* (CCMP3104) and *Phaeocystis jahnii* (CCMP2496) transcriptomes.

	P. cordata CCMP3104	*P. jahnii* CCMP2496
Number of contigs	41,629	90,009
Sum bp	49,775,870	95,573,289
min contig length	201	192
max contig length	24,502	26,543
med contig length	1,027	782
mean contig length	1,196	1,062
N50	989	1,351
GC%	67.1	67.7
With Pfam annotation	13,548 (41.5%)	31,580 (35.1%)
With GO annotation	7,119 (21.8%)	16,878 (18.8%)
With KO annotation	7,720 (18.5%)	18,233 (20.3%)
With Kegg Pathway	4,618 (11.1%)	11,071 (12.3%)
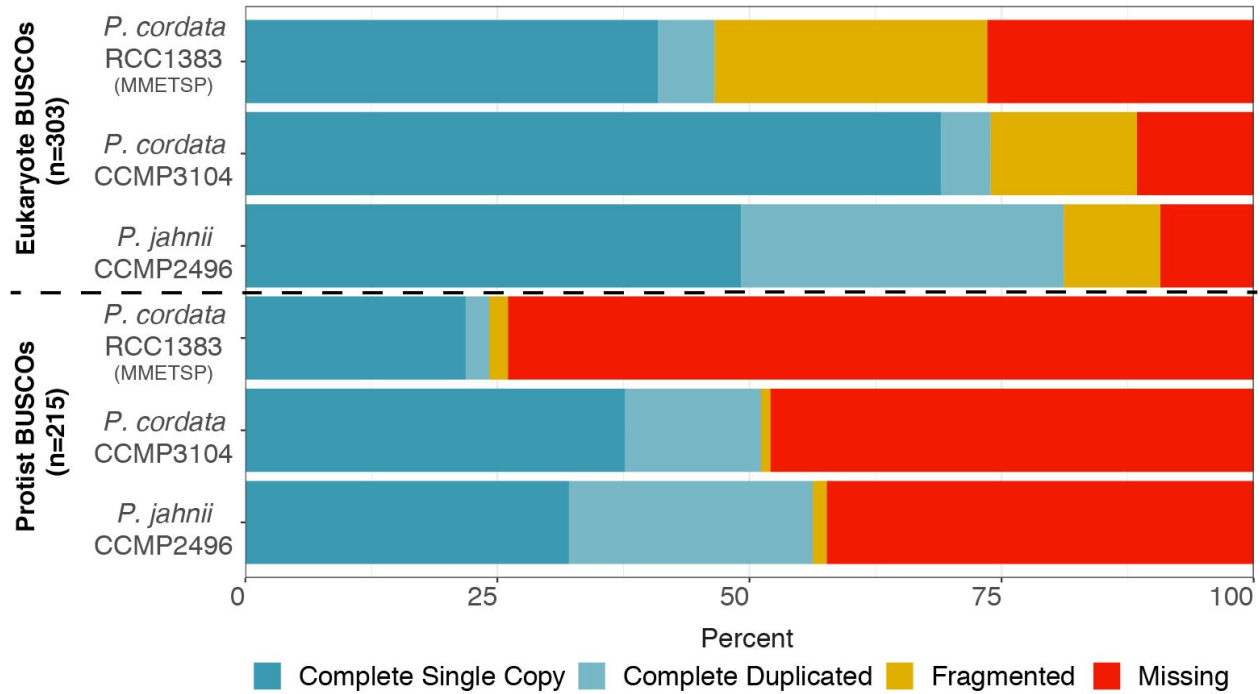
Figure 3.3. Percent of Eukaryotic and Protistan Benchmarking Universal Single-Copy Orthologs (BUSCOs) present in *Phaeocystis* transcriptomes. The *P. cordata* RCC1383 transcriptome was sequenced as part of the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP). The *P. cordata* CCMP3104 and *P. jahnii* CCMP2496 transcriptomes were sequenced for this study and are more complete than the previously assembled *P. cordata* transcriptome.

3.3.3. Differential expression in free-living and symbiotic *Phaeocystis*

Symbiotic and free-living samples for both *Phaeocystis* species clustered separately in principal component analyses (PCA) of variance stabilized gene expression results (Figure 3.4). In both cases, symbiotic and free-living samples separated along the primary axis and symbiotic samples further spread along the secondary axis, indicating that there is wider variation in gene expression in symbiotic than cultured *Phaeocystis*. For *Phaeocystis cordata*, 589 genes (1.4%) were significantly upregulated in symbiosis and 5211 genes (12.5%) were significantly downregulated in symbiosis ($p < 0.05$). For *Phaeocystis jahnii*, 132 (0.15%) genes were

significantly upregulated in symbiosis and 4545 (5.05%) were downregulated in symbiosis ($p <$ 0.05).
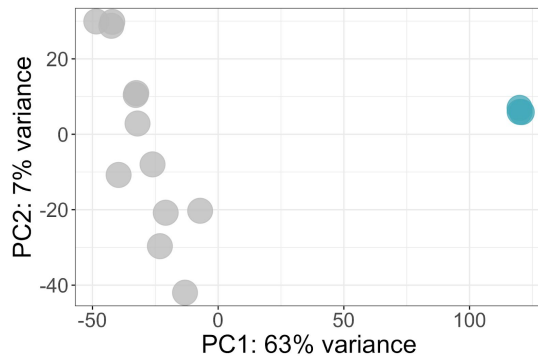
A. *P. cordata* **B.** *P. jahnii*
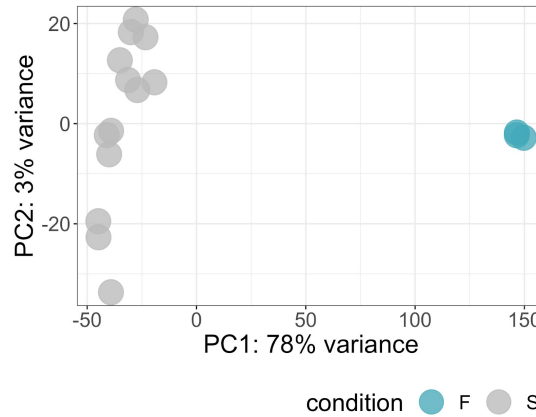
Figure 3.4. Principal component analysis of variance stabilized gene expression in symbiotic and free-living *Phaeocystis cordata* (A) and *Phaeocystis jahnii* (B). Free-living (F) *Phaeocystis* is blue and symbiotic (S) *Phaeocystis* is grey in both panels. PCA included 12 holobiont samples with symbiotic *Phaeocystis cordata* and 13 holobionts with symbiotic *Phaeocystis jahnii*. Symbiotic and free-living samples separated along the primary axis for both species. Symbiotic samples further separated along the secondary axis for both species, indicating more variation in gene expression between symbiotic cells compared to cultured cells.

3.3.4. Gene set enrichment analysis

For *P. cordata*, 139 Biological Process (BP) GO terms were enriched among genes upregulated in symbiosis (Table S3.1) and 31 BP GO terms were enriched among genes downregulated in symbiosis (Table S3.2). For *P. jahnii*, 90 BP GO terms were enriched among upregulated genes (Table S3.3) and 130 BP GO terms were enriched among downregulated genes (Table S3.4). GO terms were further categorized and visualized with REVIGO software (Supek et al., 2011) and GO terms enriched in up- and downregulated genes were plotted as treemaps for each species (Figure 3.5). Of the GO terms enriched among upregulated genes in symbiosis, 62 were shared between both species, whereas only eight GO terms enriched among

downregulated genes in symbiosis were shared between the two species. Enriched GO terms that were common to both species were categorized and visualized with REVIGO as a nonmetric multidimensional scaling (NMDS) plot (Figure 3.6). Notably, photosynthesis GO terms —including GO:Photosynthesis (count = 5, p = 0.001 for *P. cordata* and count = 2, p = 0.04 for *P. jahnii*), GO:Photosynthetic electron transport chain (count = 4, p = 0 for *P. cordata* and count = 1, p = 0.03 for *P. jahnii)*, and GO:Photosynthesis light reaction (count = 5, p = 0 for *P. cordata*)—are enriched among upregulated genes in symbiosis in the two symbiont species (Tables S3.1 and 3.3). Also of note, the DNA replication GO term is enriched among genes in symbiosis for both species (count = 15, p = 0.004 for *P. cordata* and count = 10, p = 0.03 for *P. jahnii*).

Consistent with the lower annotation rates for KEGG terms, less KEGG pathways were enriched among up- and downregulated genes in symbiotic *Phaeocystis* than were GO terms. For *P. cordata*, ten KEGG pathways were enriched among genes upregulated in symbiosis and sixteen pathways were enriched among downregulated genes. For *P. jahnii*, seven pathways were enriched among genes upregulated in symbiosis and ten pathways were enriched among downregulated genes. The two species shared four pathways enriched among upregulated genes and one pathway enriched among downregulated genes (Figure 3.7). Importantly, the photosynthesis KEGG pathway was enriched among upregulated genes in symbiosis for both species.

A. GO terms enriched among genes upregulated in symbiotic *P. cordata*

B. GO terms enriched among genes downregulated in symbiotic *P. cordata*

C. GO terms enriched among genes upregulated in symbiotic *P. jahnii*

D. GO terms enriched among genes downregulated in symbiotic *P. jahnii*

Figure 3.5. Treemaps for GO terms that were enriched among upregulated (A) and downregulated (B) genes in symbiotic *Phaeocystis cordata* and upregulated (C) and downregulated (D) genes in symbiotic *Phaeocystis jahnii*. REVIGO software was used to categorize GO terms and find a representative subset of terms with a clustering algorithm based

on semantic similarity measures. Each color depicts a cluster within one panel, and rectangles depict representative GO terms in that cluster.
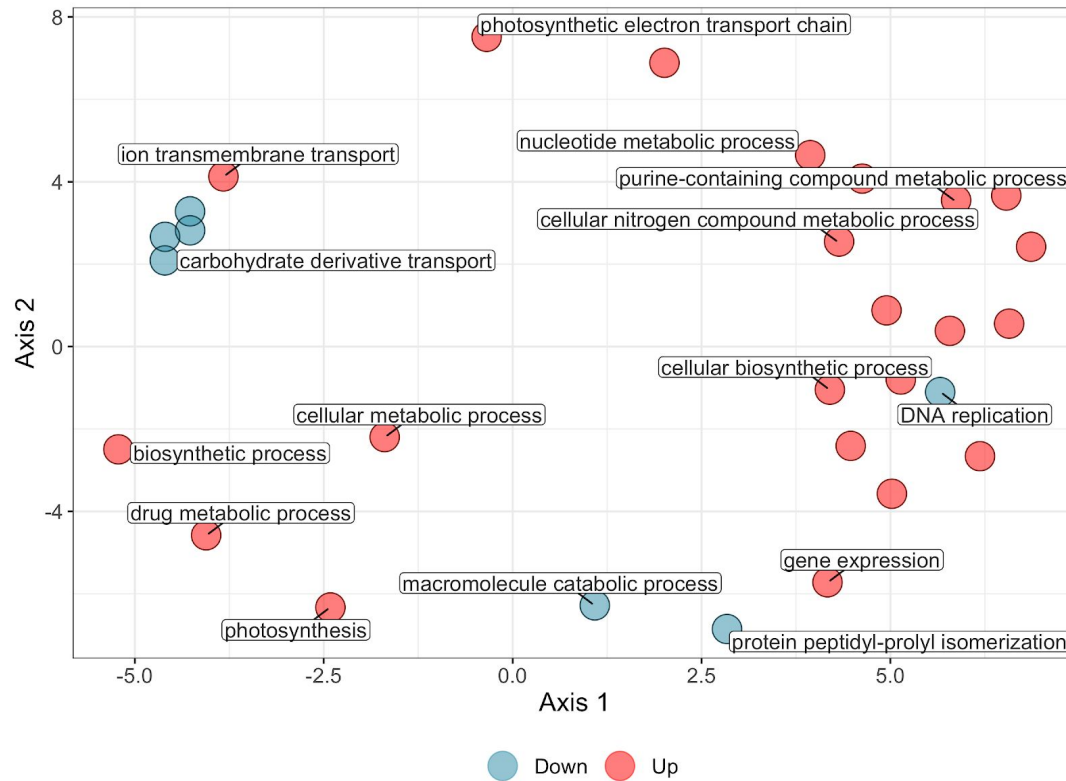
Figure 3.6. Multidimensional scaling plot of semantic similarities between non-redundant Gene Ontology (GO) terms over-represented in significantly up- and downregulated gene sets from symbiotic *Phaeocystis cordata* and *Phaeocystis jahnii*. The analysis was performed using the REVIGO online tool with allowed similarity set to 0.7 (to remove redundant GO terms) and the SimRel metric selected to calculate similarities. GO terms that were overrepresented in upregulated gene sets are colored red and GO terms overrepresented in downregulated gene sets are colored blue. GO terms with similar meanings are clustered closer together on the plot.

	Phaeocystis cordata	*Phaeocystis jahnii*
KEGG pathways enriched among upregulated genes	**Ribosome** (40/193) **Glycolysis / Gluconogenesis** (13/98) Porphyrin and chlorophyll metabolism (6/43) **Nicotinate and nicotinamide metabolism** (3/17) **Photosynthesis** (9/28) **Oxidative phosphorylation** (7/75) Ascorbate and aldarate metabolism (2/7) Human T-cell leukemia virus infection (2/8) Measles (1/1)	**Ribosome** (10/77) **Glycolysis / Gluconogenesis** (6/36) **Oxidative phosphorylation** (4/35) **Photosynthesis** (5/13) MAPK signaling pathway in yeast (2/6) Citrate cycle (TCA cycle) (2/9) Quorum sensing (2/4)
KEGG pathways enriched among downregulated genes	Chlorocyclohexane and chlorobenzene degradation (3/5) Galactose metabolism (14/45) Ras signaling pathway (15/51) Glycosaminoglycan degradation (15/45) Pyrimidine metabolism (11/28) Peroxisome (12/33) Folate biosyntheiss (7/18) Glutathione metabolism (13/34) DNA replication (17/42) Biotin metabolism (3/5) Fructose and mannose metabolism (9/22) Antifolate resistance (6/15) Oxidative phosphorylation (24/75) **N-glycan biosynthesis** (4/7) Taste transduction (4/8) Glycosaminoglycan biosynthesis (4/8)	Ribosome (46/77) Splicosome (29/54) **N-glycan biosynthesis** (12/19) Protein processing in the ER (26/47) Mucin type O-glycan biosynthesis (4/4) Ribosome biogenesis (25/44) Basal transcription factors (5/6) **Nicotinate and nicotinamide metabolism** (8/9)

Figure 3.7. Comparison of KEGG pathways enriched among up- and downregulated genes in symbiotic *Phaeocystis cordata* and *Phaeocystis jahnii*. Pathways that are upregulated or downregulated in both species are highlighted in green and pathways that are enriched in opposite gene sets in the two species are highlighted in red. Values following pathway names indicate the number of genes in that pathway that were significantly up- or downregulated over the number of genes in that pathway that were annotated in the transcriptome.

3.3.5. Expression of genes and pathways of interest

In addition to testing for significant changes in global expression, expression of genes and pathways of specific interest based on previous reports in the literature were analyzed. An important *Phaeocystis* characteristic is its role in the global sulfur cycle as a major producer of

DMS (Liss et al., 1994). Decelle et al. (2012a) found that acantharian holobionts produce more DMSP and DMS than cultured *Phaeocystis cordata*. So far, there are only two genes that have been experimentally proven to be involved in DMSP and DMS production in algae: *DSYB* catalyzes a key step in DMSP biosynthesis (Curson et al., 2018) and *Alma* (Alcolombri et al., 2015) catalyzes the breakdown of DMSP to DMS and acrylate. While orthologs for both of these genes were found in the transcriptomes for *P. cordata* and *P. jahnii*, they were not expressed at higher rates in symbiotic *Phaeocystis* when expressed at all (Figure S3.1), which opens the possibility that the host may be producing the additional DMSP/DMS, as has been demonstrated in coral symbioses (Raina et al., 2013).

Genes involved in chloroplast division are of particular interest since a drastic increase in chloroplast number is among the most striking characteristics of symbiotic *Phaeocystis*. Nuclear encoded chloroplast division genes were, therefore, identified within the *P. cordata* and *P. jahnii* transcriptomes and expression levels of these genes were compared between the free-living and symbiotic state for both species (Chen et al., 2018). *FtsZ* genes encode tubulin-like cytoskeletal GTPases that assemble into the Z ring (GTP-dependent self-assembly), which marks the constriction point during chloroplast division. The Min system regulates where the Z ring assembles on the chloroplast, and once it is assembled, dynamin-related proteins (encoded by *DRP5B*) and PD rings (encoded by *PDR1*) constrict the plastid membranes and ultimately divide the plastid (McFadden, 2014). Nuclear encoded chloroplast division genes were expressed at similar levels in free-living and symbiotic *Phaeocystis*, demonstrating that chloroplast division continues in symbiotic cells (Figure S3.2), despite the cell-cycle and DNA replication KEGG pathways—including several genes for cyclin proteins which are key for the advancement to

mitosis (Nigg, 1995)—being downregulated in symbiosis for both symbiont species (Figures S3.6, S3.7).

Genes related to nitrogen and phosphorus metabolism are of interest because these nutrients are limiting where photosymbiotic acantharians are most common. Heterotrophic hosts are widely believed to provide nitrogen and phosphorus to photosynthetic symbionts, and, in some cases, to use access to these nutrients as a means to control symbiont population size (Xiang et al., 2020). Determining whether symbiotic *Phaeocystis* cells are experiencing N or P limitation will provide evidence for or against hosts providing these nutrients to symbionts. Genes orthologous to those in nitrogen and phosphorus metabolism gene clusters annotated for the *Emiliania huxleyi* reference genome (Read et al., 2013) were identified in the *P. cordata* and *P. jahnii* transcriptomes and their expression in symbiotic and free-living cells were compared. Genes associated with phosphorus limitation, such as those for alkaline phosphatases (*PhoA, Ehap1*) and the synthesis of phosphorus-free membrane lipids (*SQD1*), were not expressed at significantly higher levels in symbiosis than in the free-living cells growing in nutrient-replete culture conditions (Figure S3.3). Importantly, the alkaline phosphatase genes were only expressed at a detectable level within a single symbiotic replicate, while these genes were expressed in all culture replicates of both symbiont species (Figure S3.3).

Patterns of expression for nitrogen metabolism genes were not fully congruent for the two symbiont species, and differences may be attributable to *P. jahnii* forming colonies in culture while *P. cordata* did not. The nitrate and nitrite transporter gene (*Nrt*) was significantly downregulated in symbiotic *P. cordata* (log2FC = −8.0, padj = 0.001) as was the gene encoding assimilatory nitrite reductase (*NirA*) (log2FC = −7.4, padj = 0.04). Genes encoding the

ammonium assimilating GS-GOGAT (Glutamine Synthetase and Glutamine Oxoglutarate Aminotransferase) enzymes, on the other hand, were significantly upregulated (log2FC = 10.8, padj = 0.0 and log2FC = 6.2, padj = 0.001, respectively) in symbiotic *P. cordata* (Figure S3.5A). In addition, the gene for carbamoyl-phosphate synthase (*CPS1*), which is involved in ammonium assimilation for arginine biosynthesis, was significantly upregulated in symbiotic *P. cordata* (log2FC = 8.5, padj = 0.0). In contrast, the nitrate and nitrite transporter gene (*Nrt*) was upregulated in symbiotic *P. jahnii* (log2FC = 5.6 , padj = 0.02) and the *GS* gene was downregulated (log2FC = −5.1, padj = 0.01), whereas *GOGAT* and carbamoyl-phosphate synthase genes were not differentially expressed (Figure S3.5B). Despite these contrasting expression patterns for important nitrogen metabolism genes in the two symbiont species, several similarities remain. For instance, neither urea nor ammonium transporter genes were differentially expressed in symbiotic *P. cordata* or *P. jahnii* (Figure S3.4). In addition, urea cycle genes were expressed at similar or higher levels in symbiotic cells compared to free-living cells for both species (Figure S3.4).

3.4. Discussion

Photosymbioses are ubiquitous throughout the global ocean and are especially important to primary production in low nutrient ecosystems. The Acantharea-*Phaeocystis* photosymbiosis is an ecologically relevant system, due to its abundance in oligotrophic surface waters, and is additionally of interest in an evolutionary context. While photosymbioses are often considered mutualistic, there is mounting evidence that symbionts may not benefit in some cases, such as in Acantharea-*Phaeocystis* associations, where symbionts are strongly manipulated by the host and

may be transitioning toward full integration as organelles. This study, therefore, attempted to decipher mechanisms of host control and nutrient transfer in Acantharea-*Phaeocystis* photosymbioses by comparing gene expression between symbiotic *Phaeocystis* and free-living *Phaeocystis* in culture conditions. The results demonstrate that photosynthesis genes are upregulated in both symbiont species that were included in the study, demonstrating that symbionts are ramping up photosynthesis *in hospite* at the molecular level. Key genes for the progression of the cell cycle are downregulated in both species, indicating that cell division is inhibited, but chloroplast division genes continue to be expressed at similar levels as in free-living cells, allowing for the proliferation of photosynthetic machinery. While expression of some nutrient assimilation pathways differs between the two species, neither species seems to be nutrient limited based on gene expression patterns. As a result, nutrient independent host inhibition of the cell-cycle must be taking place, thus suggesting a high level of symbiont integration. The downregulation of cell signaling pathways that influence cell proliferation while cells are in symbiosis presents potential mechanisms for symbiont manipulation by hosts.

3.4.1. Differences in morphology and gene expression patterns between *P. cordata* and *P. jahnii*

Comparisons between results from *P. cordata* and *P. jahnii* are somewhat complicated due to the distinct cell morphologies of the two species in culture; *P. cordata* CCMP3104 cultures include only solitary flagellated cells that are actively swimming, whereas *P. jahnii* CCMP2496 cultures include colonial and solitary cells, some of which are flagellated and actively swimming, but not all. Differences in gene expression related to these different

morphological phenotypes are evident, for example, in the enrichment of the KEGG pathway for Mucin type O-glycan biosynthesis among downregulated genes in symbiotic *P. jahnii* but not *P. cordata* (Figure 3.7). This result likely reflects a decrease in expression of genes associated with building the colony matrix while living symbiotically, as such polysaccharides have previously been indicated as components of *Phaeocystis* colonies (Janse et al., 1996). In addition, cellular respiration is enriched among genes downregulated in symbiotic *P. cordata* but not symbiotic *P. jahnii* (Figure 3.5), which is likely influenced by the cessation of active flagellar motility in symbiotic *P. cordata*, since flagellar locomotion is energy intensive (Ronkin, 1959). Nevertheless, major similarities in differential expression between symbiotic and free-living cells of the two species were apparent and are the primary focus in the interpretation of results.

3.4.2 Photosynthesis, cell cycle progression, and chloroplast division in symbiosis

The most striking similarities in gene expression results between *P. cordata* and *P. jahnii* include the upregulation of photosynthesis and the downregulation of DNA replication and progression through the cell cycle (Figures 3.6, S3.6, S3.7). Dinoflagellate symbionts in cnidarian hosts increase photosynthetic rate (Phipps and Pardy, 1982) and this enhancement of photosynthesis is considered a major feature of photosymbioses. Among Rhizaria, Decelle et al. (2019) reported higher photosynthetic efficiency for *Phaeocystis* residing in acantharian hosts than for free-living cells. The fact that photosynthesis genes are upregulated in both symbiotic *P. cordata* and *P. jahnii* in this study confirms that acantharians enhance symbiont photosynthetic rate at the molecular level. Takagi et al. (2019) proposed that a larger increase in photosynthesis in rhizarian than cnidarian symbionts when compared to their free-living counterparts, regardless

of ambient inorganic nutrient availability, indicates a higher level of host-symbiont interaction. Enhanced photosynthesis in symbiotic *Phaeocystis*, even when compared with free-living cells cultured in nutrient replete conditions, suggests strong interactions between symbionts and hosts. In this case, the extensive morphological remodeling—namely the proliferation of chloroplasts—in symbiotic *Phaeocystis* compared to free-living cells may be responsible for the increased photosynthetic output in the symbiotic state.

Symbiotic *Phaeocystis* cells have increased cell volume with the additional space primarily filled by multiple additional chloroplasts. Free-living *Phaeocystis* cells have only 2–4 chloroplasts, whereas symbiotic *Phaeocystis* can have many more chloroplasts, with some symbionts reported to contain up to 31 chloroplasts (Decelle et al., 2019). The increase in symbiont cell size accompanied by the multiplication of chloroplasts suggests that hosts are manipulating symbionts to prevent cell division and maximize photosynthetic output. Here, the downregulation of genes encoding genes for the major protein complexes responsible for DNA replication (e.g. DNA polymerase components, DNA clamp, DNA clamp loader) and proteins that promote progression through the cell cycle (e.g. cyclins, cyclin-dependent kinases, anaphase promoting complex) in both symbiont species supports the notion that hosts are preventing symbionts from dividing *in hospite* (Figures S3.6, S3.7). These results are remarkable because they suggest a mechanism for maintenance of the relationship. Preventing symbiont overgrowth presents a major dilemma for organisms hosting endosymbionts, which many hosts overcome by systematically digesting (Titlyanov et al., 1996) or releasing symbionts (Boettcher et al., 1996; Fishman et al., 2008). Previous results indicate that acantharians do not systematically digest symbionts and that they maintain symbionts long enough for external symbiont availability to

shift compared to internal symbiont populations (Mars Brisbin et al., 2018). While those results pointed to sustained symbiont maintenance without symbiont digestion, the current study shows that symbiont cell division is indeed inhibited within acantharian hosts.

Unlike plants, whose cells often have many chloroplasts that divide independently from the plant cell, algal chloroplast and cell division are typically synchronized (Sumiya, 2018). In symbiotic *Phaeocystis,* chloroplast and cell division have become decoupled, with chloroplasts continuing to divide and accumulate even though the cell-cycle is arrested. Chloroplast division in single-celled algae is generally initiated by the expression of nuclear-encoded chloroplast division genes during S-phase of the algal cell cycle, leading to the hypothesis that if algal cells are stuck in S-phase, chloroplasts may continue to divide and accumulate. Indeed, when *Cyanidioschyzon merolae* (a unicellular red alga) was artificially prevented from progressing from S-phase, chloroplasts continued to divide (Itoh et al., 1996), thus indicating that chloroplast division can be sustained when algal cells themselves do not divide. In this study, nuclear-encoded chloroplast division genes—*FtsZ, PDR1, DRP5B*—were not differentially expressed in symbiotic and free-living *Phaeocystis* (Figure S3.2), despite cell-cycle genes being significantly downregulated in symbiotic cells (Figure S3.7). These results suggest that symbiotic *Phaeocystis* may be arrested in S-phase, allowing chloroplasts to continue dividing and accumulating and ultimately keeping symbiont population density constrained while simultaneously increasing symbiont photosynthetic output.

3.4.3. Nutrient transfer between hosts and symbionts

Nutrient transfer between hosts and symbionts is considered central to photosymbioses, especially in low-nutrient regions where nutrient availability otherwise limits phytoplankton production. Symbiotic algae are expected to benefit from host-supplied nitrogen and phosphorus procured through prey capture and, in return, to supplement hosts with an additional energy source by sharing photosynthetically derived organic carbon. Understanding which compounds are transferred between hosts and symbionts is, therefore, a principal aim in studying these symbioses. Identifying whether and in what form hosts provide nitrogen and phosphorus to symbionts from gene expression data is challenging, though, because gene expression does not perfectly correlate to protein translation and is further removed from enzymatic activity and metabolic rates (Waldbauer et al., 2012). The matter is further complicated when symbionts are compared to free-living cells in nutrient replete conditions. A more informative comparison in this regard might be between symbiotic *Phaeocystis* and *Phaeocystis* living in naturally occurring communities in oligotrophic conditions, which could be accomplished in the future by preparing metatranscriptomes from size-fractionated seawater samples (to remove hosts). Notwithstanding, the results presented here can help decipher patterns of nutrient exchange between symbiotic *Phaeocystis* and acantharian hosts while keeping these limitations in mind.

Overexpression of the nitrate/nitrite transporter gene (*Nrt*) is a major indicator of nitrogen limitation in marine algae (Alipanah et al., 2015; Sanz-Luque et al., 2015; Xiang et al., 2020). In symbiotic *P. cordata*, *Nrt* was significantly downregulated, suggesting that symbiotic cells were not nitrogen limited. Furthermore, the assimilatory nitrate reduction gene (*NirA*) was

downregulated and ammonia assimilation genes—carbamoyl phosphate synthetase (*CPS1*), glutamine and glutamate synthetases (*GS-GOGAT*)—were upregulated in symbiosis (Figure S3.5A). These findings mirror those reported for another rhizarian photosymbiosis, between the Polycystine radiolarian *Thalassicolla nucleata* and its dinoflagellate symbiont, *Brandtodinium sp.* (Liu et al., 2019), where host-supplied ammonium was indicated as the main nitrogen source for symbionts. Ammonium transporter genes were expressed at low levels in symbiotic *P. cordata* and a urea transporter gene was also expressed at similar levels as in free-living cells, along with several urea metabolism genes, including argininosuccinic acid synthetase and argininosuccinate lyase (Figure S3.6A). *Phaeocystis*, like other haptophytes, can use urea as a nitrogen source by catalyzing its breakdown to ammonium and carbon dioxide, which would also help meet an increased carbon dioxide demand resulting from elevated photosynthetic rates in symbiosis (Lindehoff et al., 2011; Wang et al., 2011; Syrett and Leftley, 2016). Heterotrophic protists excrete both urea and ammonia as nitrogenous waste (Weatherby, 1929; Dolan, 1997), thus making both urea and ammonium potential nitrogen sources for symbiotic *Phaeocystis*.

Interestingly, the expression patterns for nitrogen metabolism genes were different in *P. jahnii* compared to *P. cordata*, which probably reflects the formation of colonies by free-living *P. jahnii* in culture. Unlike in symbiotic *P. cordata*, *Nrt* was upregulated in symbiotic *P. jahnii* and while genes for nitrate reductase, carbamoyl phosphate synthetase, and glutamate synthetase were not differentially expressed, the gene for glutamine synthetase was downregulated in symbiosis (Figures S3.6B, S3.7B). Ammonium transporter and urea transporter genes were expressed at similar levels in symbiotic *P. jahnii* as in *P. cordata* (Figure S3.4) and genes for argininosuccinate lyase (urea cycle) were expressed at similar levels in symbiotic and free-living

P. jahnii. *Phaeocystis* colonial matrices include amino group-containing compounds—potentially amino sugars or amino acids—that cells may later access as additional nitrogen sources (Hamm et al., 1999). If colonial cells rely more heavily on these alternative nitrogen sources than solitary cells, it could explain the differing expression patterns for nitrogen metabolism genes between symbiotic and free-living, colonial *Phaeocystis jahnii* compared to between symbiotic and free-living, non-colonial *P. cordata*. However, since cellular nitrogen compound biosynthesis, peptide biosynthesis and translation GO terms were still enriched among genes upregulated in symbiotic *P. jahnii*, like in *P. cordata*, is appears that *P. jahnii* is also not nitrogen limited within hosts.

Chloroplasts within symbiotic *Phaeocystis* cells have a higher N:P ratio than in free-living cells (Decelle et al., 2019), which could be considered evidence that symbiotic *Phaeocystis* is phosphorus limited. Alternatively, denser thylakoids and increased photosynthetic rate in symbiotic *Phaeocystis* could also skew the N:P ratio towards N (Decelle et al., 2019), as thylakoid membranes are primarily protein and lipid (only 9% phospholipid) (Andersson and Barber, 1994). Alkaline phosphatase, which catalyzes the hydrolysis of organic phosphorus compounds and releases inorganic phosphorus, is an indicator of P-limitation in marine phytoplankton (Annis and Cook, 2002; Alipanah et al., 2018; Li et al., 2018). Symbiotic *P. cordata* and *P. jahnii* did not express alkaline phosphatase genes (*PhoaA, Ehap1*) at detectable levels (except for in one holobiont sample), whereas free-living cells in nutrient replete media expressed both alkaline phosphatase genes in all replicates (Figure S3.5). Likewise, *P. cordata* and *P. jahnii* orthologs to *Emiliania huxleyi* genes associated with P-limitation (Read et al., 2013) were either expressed at similar or lower levels in symbiosis compared to in free-living

cells. Together, these results suggest that *Phaeocystis* symbionts are not nutrient-limited, especially since symbiotic gene expression was compared to expression in nutrient replete conditions. Therefore, Nitrogen and Phosphorus procured through host predation is probably available to the symbionts *in hospite*, since there are vanishingly low inorganic nutrient concentrations in the surface waters where holobionts were collected (Hirose and Kamiya, 2003).

The carbon fixation (the Calvin-Benson cycle) pathway and glycolysis/gluconeogenesis pathway were both upregulated in symbiotic *Phaeocystis*, indicating a potential increase in organic carbon production in symbiosis (Figures 3.7, S3.8). Excess fixed organic carbon can either be converted to storage molecules or transferred to the host. Eukaryotic phytoplankton in oligotrophic regions preferentially store organic carbon as triacylglycerols (TAG), as these lipid molecules are built from only C, O, and H—they do not use N or P, which are scarce—and they are more energy rich than carbohydrates (Becker et al., 2018). Haptophytes in the oligotrophic North Pacific Subtropical Gyre store energy as TAGs throughout the day and then use them as energy sources at night (Becker et al., 2018). Key genes for TAG biosynthesis, including those for acyl-CoA binding proteins and acyltransferases that are involved in the production of TAG precursors (e.g. phosphatidic acid) and diacylglycerol transferase, which performs the last step in TAG synthesis, were all significantly downregulated in both symbiotic *P. cordata* and *P. jahnii* (Tables S3.5, S3.6). Lipases specifically involved in TAG consumption were likewise all significantly downregulated in both symbiont species (Tables S3.5, S3.6). Symbiotic *Phaeocystis* does not appear to be using the carbon storage pathway most commonly used by free-living haptophytes in the environment from which holobionts were collected. Fixed carbon can alternatively be stored as starch, but starch synthesis (glycogen synthase, *GYS*) was significantly

downregulated in both species (log2FC = −6.9, padj = 0.02 for *P. cordata*; log2FC = −7.0, padj = 0.003 for *P. jahnii*). These results suggest that symbiotic cells may forego carbon storage and instead relinquish surplus photosynthate to hosts.

While reduced biosynthesis of storage molecules implies that symbionts export carbon to hosts, the form in which symbionts share organic carbon is not completely clear. Sugar metabolism (galactose, fructose, mannose) and polysaccharide biosynthesis (glycosaminoglycan, N-glycan, mucin type O-glycan) were downregulated in symbionts (Figure 3.7) and the GO term for carbohydrate derivative transport was enriched among downregulated genes for both species (count = 15, *p* = 0.01 for *P. cordata*, count = 17, *p* = 0.004 for *P. jahnii*). In contrast, multiple GO terms associated with the biosynthesis and transport of small organic nitrogen compounds, such as amides and peptides, were enriched among genes upregulated in symbiosis for both symbiont species (Tables S3.1, S3.3). Rather than sharing carbohydrates with hosts, it seems possible that symbionts may instead share small organic nitrogen compounds, especially since nitrogen is not limiting in symbiosis. Similarly, comparative transcriptomics with a dinoflagellate symbiotic to radiolarians also suggested that symbionts export organic nitrogen compounds rather than carbohydrates (Liu et al., 2019). It is particularly interesting that both *Phaeocystis* species in this study downregulated pathways for the biosynthesis of exopolysaccharides while in symbiosis, since *Phaeocystis*' predisposition towards exuding these compounds could have increased its attractiveness as a symbiont. Free-living colonial *Phaeocystis jahnii* uses these compounds to build the colonial matrix (Hamm, 2000) and solitary free-living *Phaeocystis cordata* and *jahnii* often excrete these compounds through exocytosis (Quesada et al., 2006). However, *Phaeocystis* also stores organic nitrogen compounds in the

colonial matrix (Solomon et al., 2003), indicating that adaptations for colonial living may still contribute to *Phaeocystis*' suitability as a symbiont if these compounds are also shared with hosts.

3.4.4. Host control of symbiont populations through cell-signaling pathways

Nutrient limitation is used to regulate symbiont population size in several photosymbiotic systems (Xiang et al., 2020). Nitrogen starved symbiotic dinoflagellates divide less frequently but maintain photosynthetic output (Xiang et al., 2020). Similarly, some dinoflagellates also maintain photosynthetic rates and increase in size but do not divide under P-limitation (Li et al., 2016). However, exogenous P or N enrichment releases symbionts from nutrient limitation, thus making nutrient limitation a relatively inefficient mechanism for controlling symbiont population density (Annis and Cook, 2002; Béraud et al., 2013). In addition, many phytoplankton species, including several diatom and dinoflagellate species, decrease photosynthetic output when N- or P-limited (Béraud et al., 2013; Alipanah et al., 2018). The data presented here suggest that symbiotic *Phaeocystis* cells are not nutrient limited, necessitating an alternative, nutrient-independent mechanism for symbiont population control by acantharian hosts.

One way hosts could slow or stop symbiont cell division without limiting access to nutrients might be to manipulate a cell-signaling pathway that regulates cell division. The mitogen activated protein kinase (MAPK) signaling pathway is involved in signal transduction from specific extracellular stimuli to targets in the cytoplasm and nucleus and affects cell proliferation, differentiation, and survival (Morrison, 2012). Several modules are included within the MAPK signaling pathway, including the extracellular-signal-related kinase (ERK) cascade,

which is involved in responding to mitogens, hormones, and other chemical signals, whereas the

other MAPK modules are generally activated by stress or environmental conditions. When

activated, the ERK cascade can stimulate cell proliferation, but can also induce differentiation,

development, cell survival, migration, or apoptosis depending on signal strength and duration,

intracellular conditions, and other signaling pathways (Shaul and Seger, 2007). The MAPK

signaling pathway, therefore, represents a potential nutrient-independent mechanism for host

cells to mediate symbiont cell division and differentiation *in hospite*.

In symbiotic *P. cordata,* genes encoding ERKs and genes encoding Ras proteins, which

initiate ERK cascades, were significantly downregulated compared to in free-living cells (Figure

S3.9A). The decreased expression of key genes in this pathway could indicate that the pathway is

suppressed in symbiosis, preventing progression of the cell cycle. Indeed, downregulation of

ERKs has a clear anti-proliferation effect in model systems (e.g. yeast and human cell lines)

(Fisher et al., 2001; Bae et al., 2013). Furthermore, ERK expression can be downregulated in

response to chemical signaling as well as mechanical signaling due to high cell-density (Bae et

al., 2013; Gérard and Goldbeter, 2014; VanHook, 2019). Acantharian hosts may be limiting

symbiont cell division by suppressing ERK expression by way of a not yet identified chemical

signal, by keeping symbionts in very tight quarters (Febvre and Febvre-Chevalier, 1979), or by a

combination of these tactics.

Slightly different expression patterns for MAPK genes were observed in symbiotic *P.*

jahnii compared to symbiotic *P. cordata.* Genes for ERKs and Ras proteins were not

significantly differentially expressed in symbiosis, but genes for RSK2 (ribosomal protein S6

kinase A6), which is activated by ERKs and induces cell-proliferation (Vaidyanathan et al.,

2007), were significantly downregulated in symbiotic *P. jahnii* (Figure S3.8B). This difference

could, again, arise from gene expression specific to colony formation in *P. jahnii*. The MAPK

pathway is downregulated in colonial *P. globosa* compared to solitary cells (Mars Brisbin and

Mitarai, 2019). If the MAPK pathway is similarly downregulated in colonial *P. jahnii* compared

to solitary cells, it could mask the downregulation of key genes in the pathway in symbiotic *P.

jahnii*. Alternatively, different regulatory mechanisms may be more important to different

Phaeocystis species in symbiosis. For example, genes for AKT serine-threonine kinases, which

also regulate cell proliferation (Marte and Downward, 1997), are significantly downregulated in

symbiotic *P. jahnii* (Figure S3.9B).

3.4.5. The question of mutualism in Acantharea-*Phaeocystis* photosymbioses

It is clear that photosymbiotic acantharians benefit from hosting *Phaeocystis* symbionts;

not only are photosymbiotic acantharians highly abundant in oligotrophic surface waters, but

they also outnumber asymbiotic acantharians in the same regions (Caron et al., 1995). In

addition, the energetic benefits gained from symbiont photosynthate are further evidenced by the

ability of photosymbiotic acantharians to produce more robust skeletons than asymbiotic

acantharians (Anderson, 1996; Decelle and Not, 2015). The results from this study indicate that

Phaeocystis symbionts are maintained in nutrient replete conditions within hosts. Acantharians

are, therefore, upholding their end of the bargain, in a sense—they are supplying nutrients in

return for fixed organic carbon. However, symbiont cell division is blocked, which leads to a

gross morphological transformation of symbiont cells. It remains possible that symbionts could

resume cell-cycle progression if they were released or escaped from hosts, but the extent of the

morphological transformation may very well make it irreversible and, therefore, permanently

prevents successful mitosis and cytokinesis. While the results presented here cannot

unambiguously determine whether Acantharea-*Phaeocystis* symbioses are mutualistic, available

evidence seems to point towards symbiont exploitation.
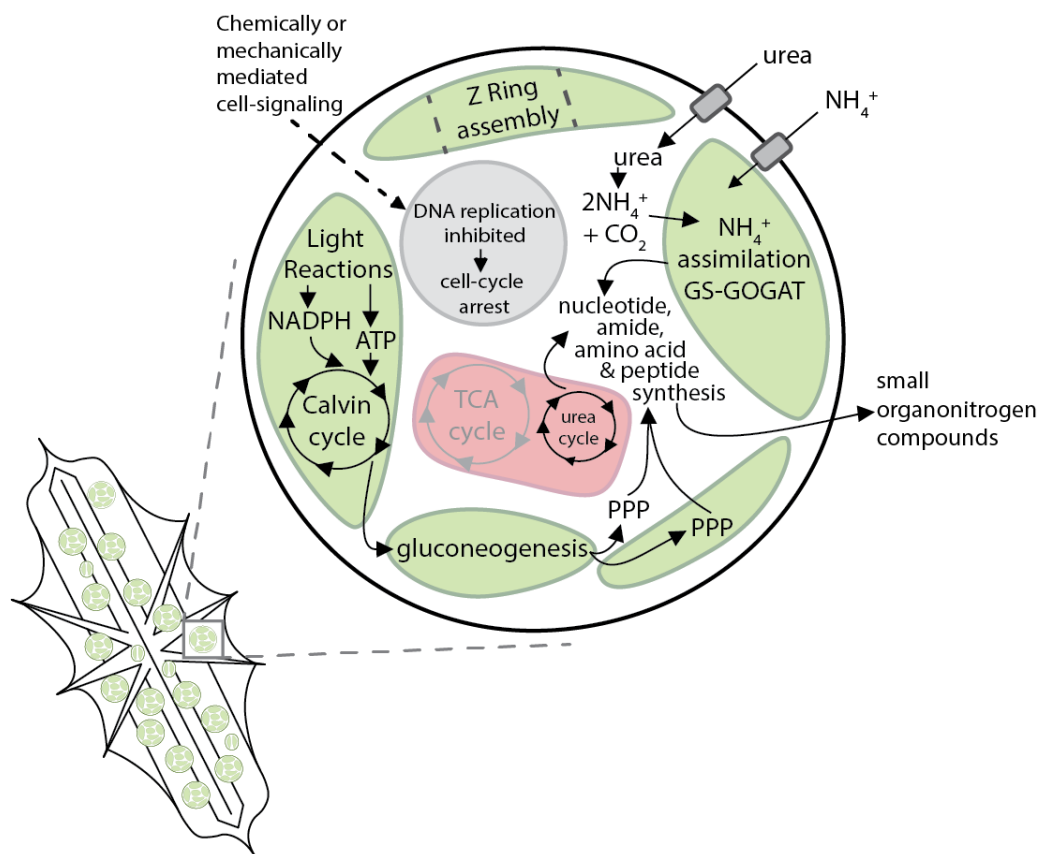
Figure 3.8. Conceptual model summarizing important processes in symbiotic *Phaeocystis cordata* cells inferred from gene expression data in this study. GS-GOGAT represents glutamine synthetase-glutamate synthetase enzymes, which assimilate ammonium; TCA stands for the tricarboxylic acid cycle (a.k.a. the Krebs cycle or the citric acid cycle); PPP stands for the pentose phosphate pathway, which occurs both in chloroplasts and in the cytoplasm. Shapes are color coded to represent organelles and cell structures as follows: chloroplasts are filled with green, the mitochondria is filled with red, the cell nucleus is filled with light grey, and transmembrane transporters are filled with dark grey. Solid black arrows indicate processes that are upregulated in symbiosis and solid grey arrows indicate processes that are downregulated in symbiosis.

3.4.6. Conclusions

Despite some differences in differential gene expression patterns in symbiotic *P. cordata* and *P. jahnii*, the results generally paint a coherent picture: DNA replication and progression through the cell cycle is inhibited in symbiotic *Phaeocystis* while chloroplasts continue to divide, photosynthetic rate is enhanced, and cells increase in size (Figure 3.8). While in symbiosis, *Phaeocystis* appears to benefit from host-supplied urea and ammonium as nitrogen sources, and instead of producing storage carbohydrates or lipids, produces and exports small organic nitrogen compounds to the host (Figure 3.8). Key genes in the MAPK signaling pathway are significantly downregulated in symbiosis for both species, indicating that hosts may be suppressing this important pathway to limit symbiont proliferation. By providing symbionts with ample nutrients and utilizing cell-signaling pathways rather than nutrient limitation to manage symbiont population sizes, acantharians have finer control over symbionts and ensure that symbiont photosynthetic rates are maintained. The results from this study demonstrate strong interactions between hosts and symbionts in Acantharea-*Phaeocystis* relationships, suggesting an exceptionally high level of symbiont integration. Whether symbionts ever regain reproductive capacity is still unknown, leaving open the possibility for mutualism in this relationship, but available evidence suggests hosts are probably exploiting symbionts. Continued study of such strong interactions in a horizontally transferred symbiosis may allow insight into mechanisms involved in the earliest stages of secondary or tertiary plastid acquisition among protists.

3.5. Supplementary Material

Table S3.1. GO terms enriched among significantly upregulated genes in symbiotic *P. cordata*.
https://maggimars.github.io/AcanthSymbiontDGE/Pcordata/enriched_up_GO.html

Table S3.2. GO terms enriched among significantly downregulated genes in symbiotic *P. cordata*.
https://maggimars.github.io/AcanthSymbiontDGE/Pcordata/enriched_down_GO.html

Table S3.3 GO terms enriched among significantly upregulated genes in symbiotic *P. jahnii*.
https://maggimars.github.io/AcanthSymbiontDGE/Pjahnii/enriched_up_GO.html

Table S3.4. GO terms enriched among significantly downregulated genes in symbiotic *P. jahnii*.
https://maggimars.github.io/AcanthSymbiontDGE/Pjahnii/enriched_down_GO.html

Table S3.5. Significantly differentially expressed genes associated with TAG biosynthesis in *P. cordata*

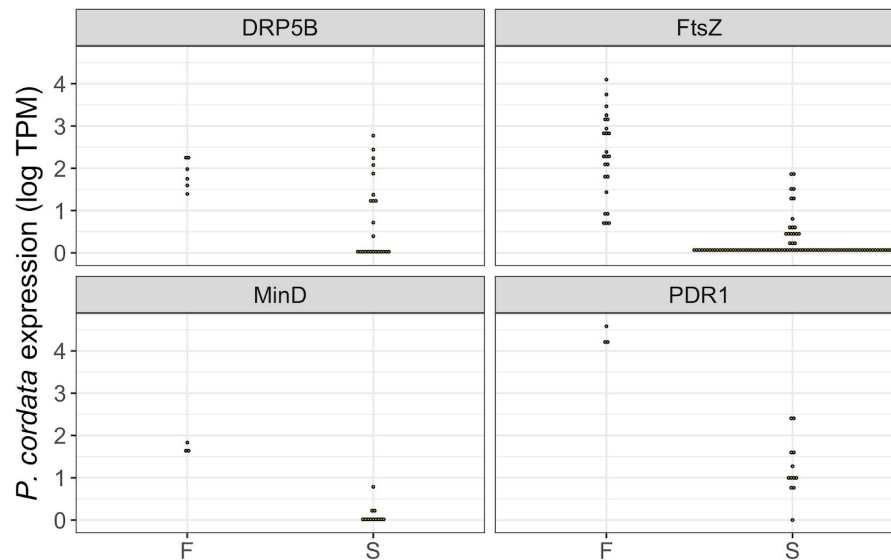| Protein	Pfam ID	Number of orthologs in transcriptome	Transcriptome contig IDs	Log fold change range in symbiosis
Acyl transferase	PF01553	4	TRINITY_DN12384_c1_g1_i1, TRINITY_DN16956_c0_g1_i1, TRINITY_DN198_c0_g1_i3, TRINITY_DN579_c1_g1_i4	−6.1 to −5.4
Lipase 3	PF01764	5	TRINITY_DN10360_c0_g1_i1, TRINITY_DN10733_c0_g1_i1, TRINITY_DN35072_c0_g1_i1, TRINITY_DN35118_c0_g1_i1, TRINITY_DN4009_c0_g1_i1	−8.1 to −5.1
Acyl-coa binding protein	PF00887	1	TRINITY_DN2559_c0_g2_i1	−5.0
DGAT	PF03982	2	TRINITY_DN1390_c0_g1_i5, TRINITY_DN3220_c0_g1_i1	−6.7 to −5.5

Table S3.6. Significantly differentially expressed genes associated with TAG biosynthesis in *P. jahnii*

Protein	Pfam ID	Number of orthologs in transcriptome	Transcriptome contig IDs	Log fold change range in symbiosis
Acyl transferase	PF01553	4	TRINITY_DN3743_c0_g1_i1, TRINITY_DN6295_c0_g1_i2, TRINITY_DN7813_c0_g1_i1, TRINITY_DN9200_c0_g1_i1	−6.1 to −5.4
Lipase 3	PF01764	3	TRINITY_DN1182_c0_g1_i2, TRINITY_DN19456_c0_g1_i1, TRINITY_DN21074_c0_g1_i1	−8.1 to −5.1
Acyl-coa binding protein	PF00887	4	TRINITY_DN2863_c0_g1_i1, TRINITY_DN3909_c0_g1_i4, TRINITY_DN5482_c0_g1_i3, TRINITY_DN9118_c0_g1_i2	−8.5 to −5.8
DGAT	PF03982	3	TRINITY_DN220_c0_g1_i1, TRINITY_DN3630_c0_g1_i1, TRINITY_DN4330_c0_g1_i1	−6.6 to −5.1

A. *P. cordata*

B. *P. jahnii*

Figure S3.1. Expression of *Alma* and *DSYB* orthologs in symbiotic (S) and free-living (F) *Phaeocystis cordata* (A) and *Phaeocystis jahnii* (B). Orthologs were identified by blasting DSYB protein sequences from the MMETSP database (Curson et al., 2018) and Alma protein sequences from Alcolombri et al. (2015) against the final transcriptome assemblies for *P. cordata* and *P. jahnii*. Plot facets are labeled with the protein name and the organism that the sequence originated from. TPM refers to Transcripts Per kilobase Million and normalizes for gene length and library size, and allows for comparisons in gene expression between different samples.
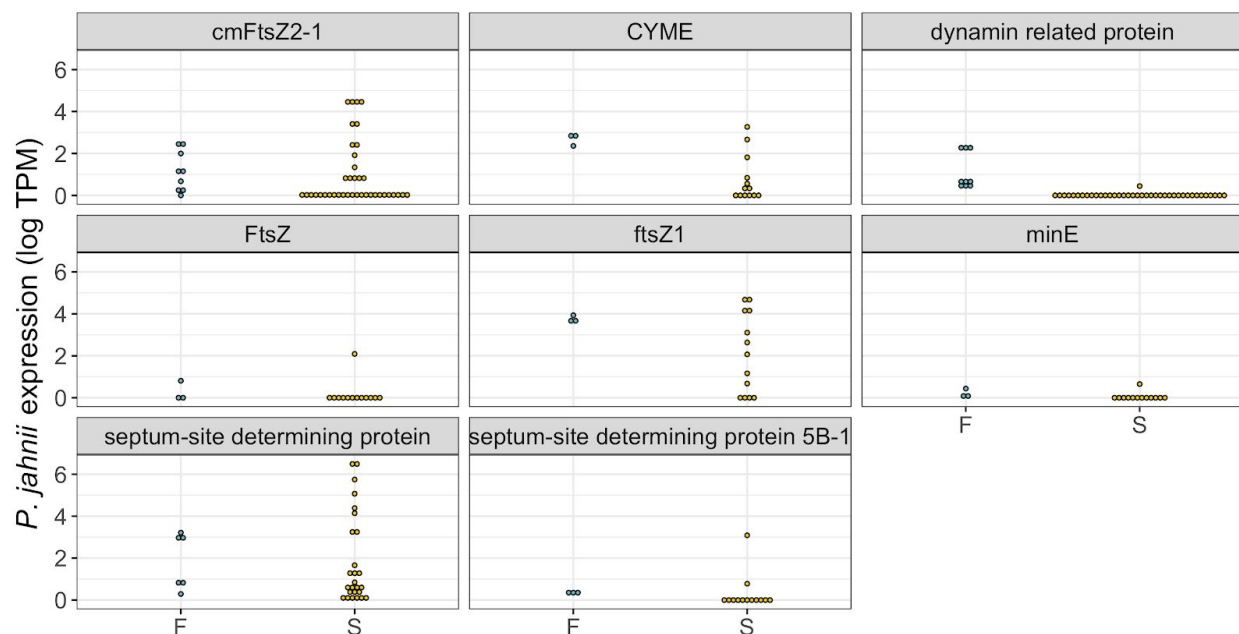
A. *P. cordata*

B. *P. jahnii*

Figure S3.2. Expression of nuclear encoded chloroplast division genes in symbiotic (S) and free-living (F) *Phaeocystis cordata* (A) and *Phaeocystis jahnii* (B). Nuclear encoded chloroplast division genes were found in the *P. cordata* and *P. jahnii* transcriptomes by blasting the gene sequences from Chen et al. (2018) against the two transcriptomes. Many of these genes were present as more than one gene copy in the transcriptomes, which is apparent where there are more than three data points for the free-living replicates. Chloroplast division genes were generally expressed at similar rates in symbiosis as in free-living cells, but not all isoforms were expressed in every holobiont sample.

A. *P. cordata*
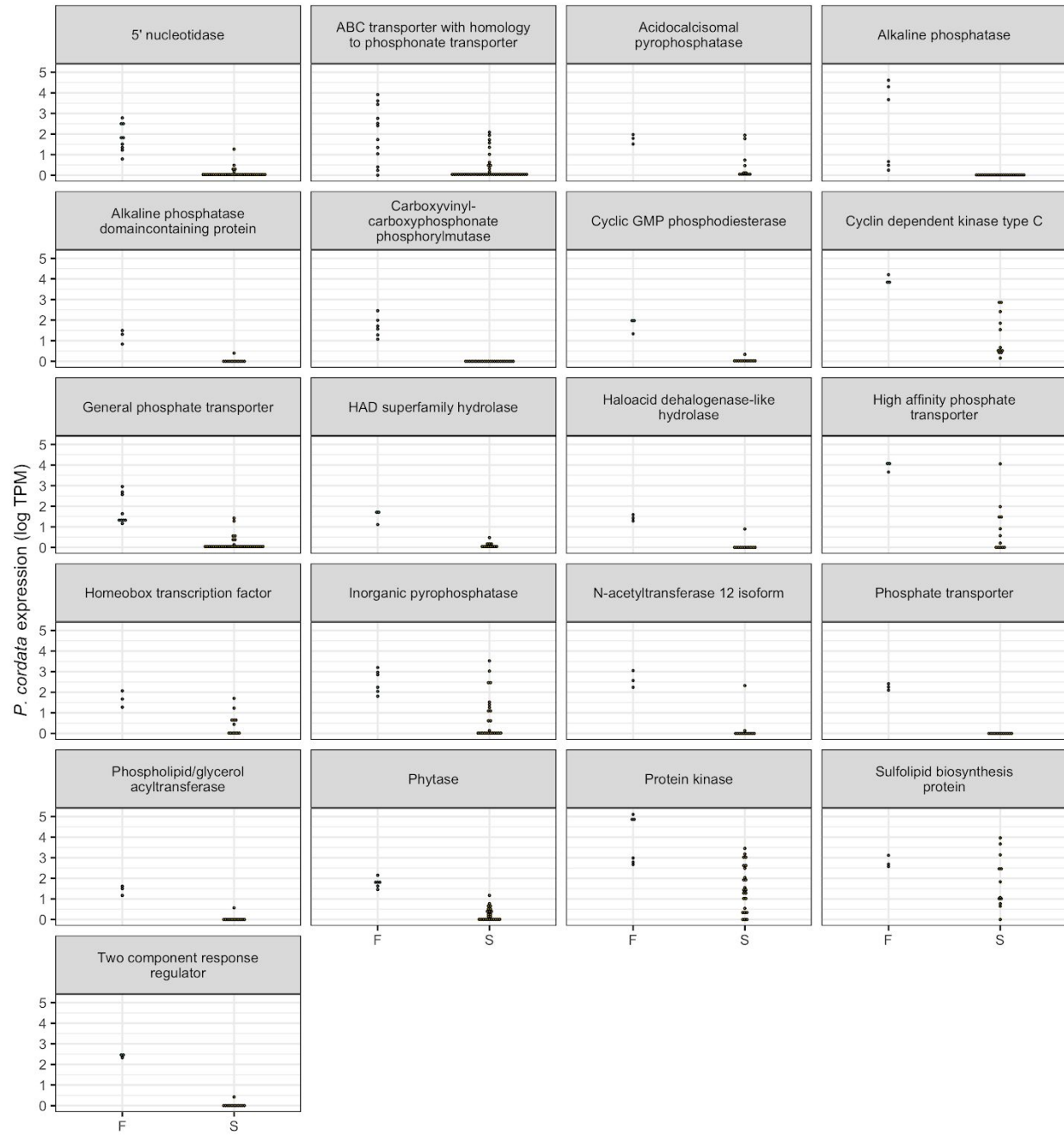
B. *P. jahnii*

Figure S3.3. Expression of Phosphorus-limitation associated genes in symbiotic (S) and free-living (F) *Phaeocystis cordata* **(A) and** *Phaeocystis jahnii* **(B).** Genes associated with phosphorus limitation were identified in the *P. cordata* and *P. jahnii* transcriptomes by blasting protein sequences from the phosphorus limitation cluster in the *Emiliania huxleyi* genome against both transcriptomes (Read et al., 2013). Many genes were present as more than one gene copy, which is apparent where there are more than three data points for the free-living replicates.

A. *P. cordata*

B. *P. jahnii*

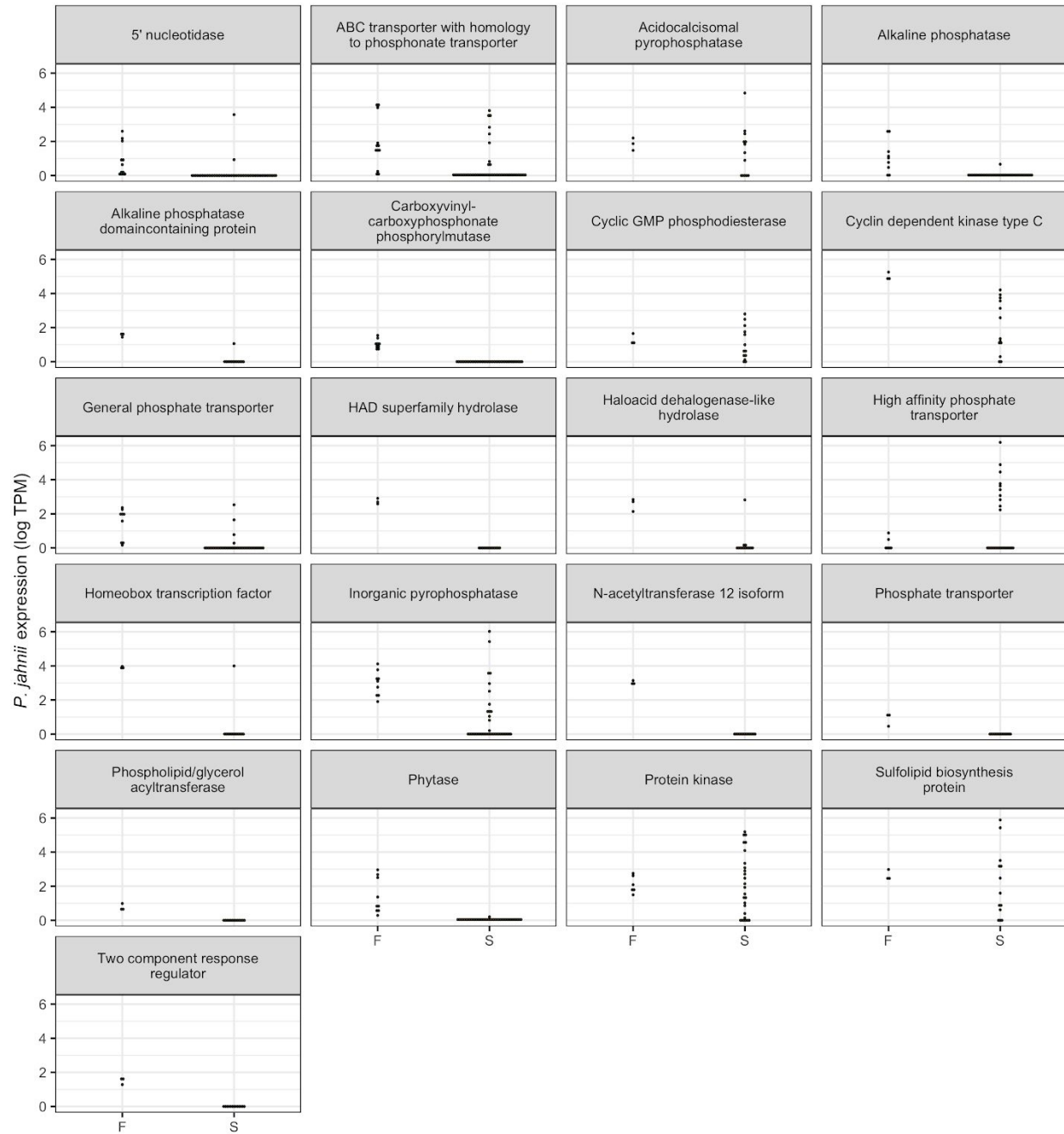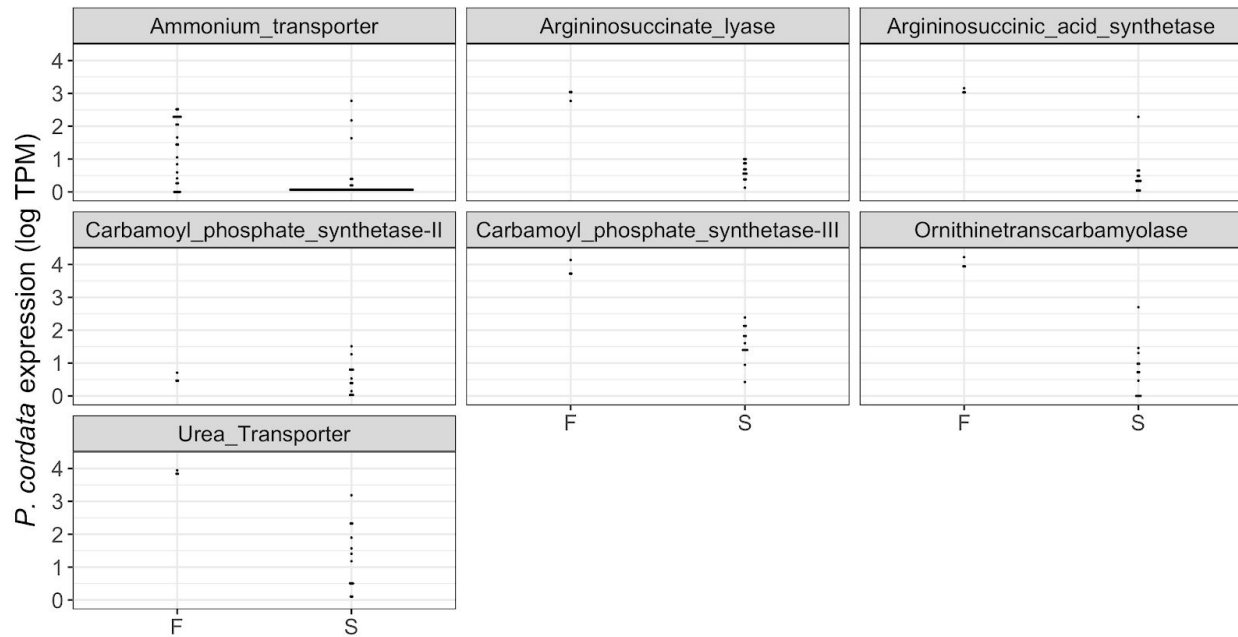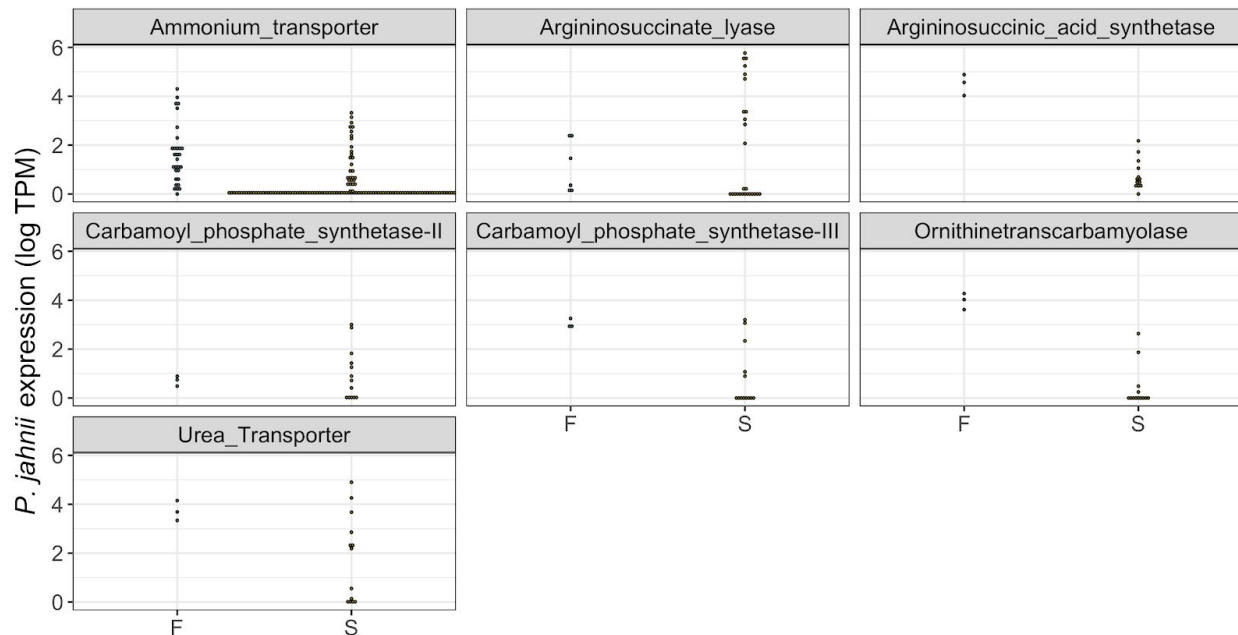**Figure S3.4. Expression of Nitrogen metabolism genes in symbiotic (S) and free-living (F)** *Phaeocystis cordata* **(A) and** *Phaeocystis jahnii* **(B).** Genes associated with nitrogen metabolism were identified in the *P. cordata* and *P. jahnii* transcriptomes by blasting the protein sequences from the nitrogen metabolism cluster in the *Emiliania huxleyi* genome against the two transcriptomes (Read et al., 2013). Many of these genes were present as more than one gene copy in the transcriptomes, which is apparent where there are more than three data points

for the free-living replicates. A urea transporter gene was expressed at similar levels in symbiotic and free-living cells for both species and ammonium transporter genes were also expressed at similar levels in the two sample types for *P. jahnii,* but were less consistently expressed by symbiotic *P. cordata*.

A. *P. cordata*

B. *P. jahnii*

Key: Gene for protein is significantly **down**regulated in symbiosis | Gene for protein is significantly **up**regulated in symbiosis | Gene for protein is not differentially expressed
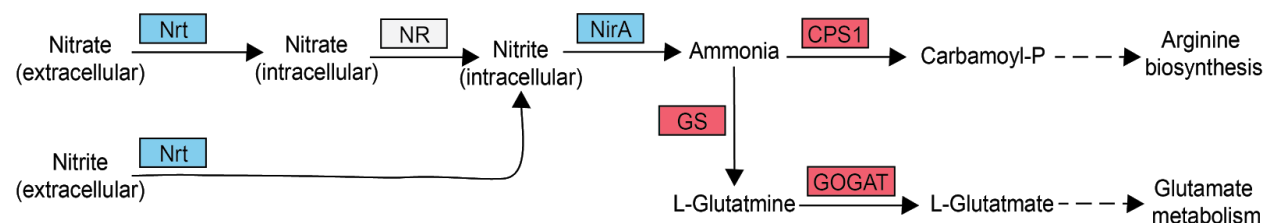
Figure S3.5. Nitrogen metabolism genes significantly up- and downregulated in symbiotic *P. cordata* (A) and *P. jahnii* (B). Proteins for which genes were significantly upregulated in symbiosis (log2FC > 1 and padj ≤ 0.05) are shaded red and proteins for which genes were significantly downregulated in symbiosis (log2FC < −1 and padj ≤ 0.05) are shaded blue. The genes for proteins in grey shaded boxes were not significantly differentially expressed in symbiosis.

A. *P. cordata*

B. *P. jahnii*

Figure S3.6. DNA replication genes significantly downregulated in symbiotic *P. cordata* (A) and *P. jahnii* (B). Proteins for which genes were significantly downregulated in symbiosis (log2FC < −1 and padj ≤ 0.05) are shaded blue. No genes for DNA replication complex subunits were significantly downregulated in symbiosis for either species. Genes for subunits shaded grey were not significantly differentially expressed in symbiosis. If genes for the protein were not annotated in the transcriptome, the subunit is white.

A. *P. cordata*

B. *P. jahnii*
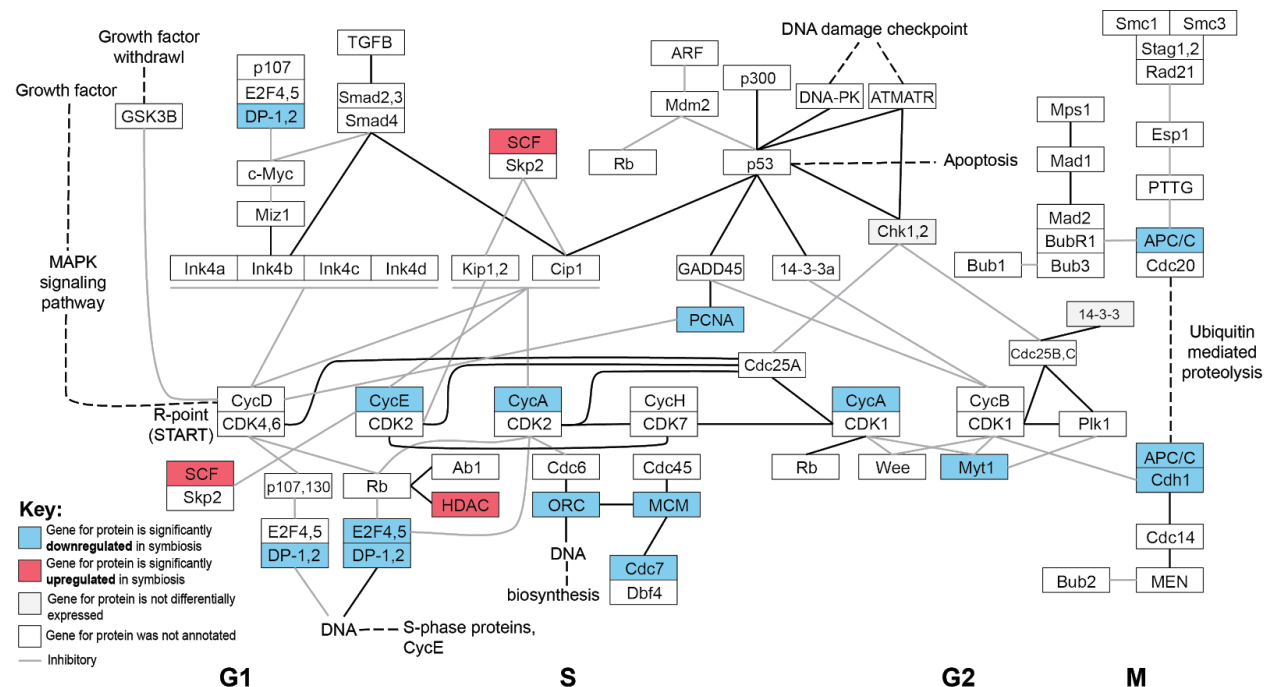
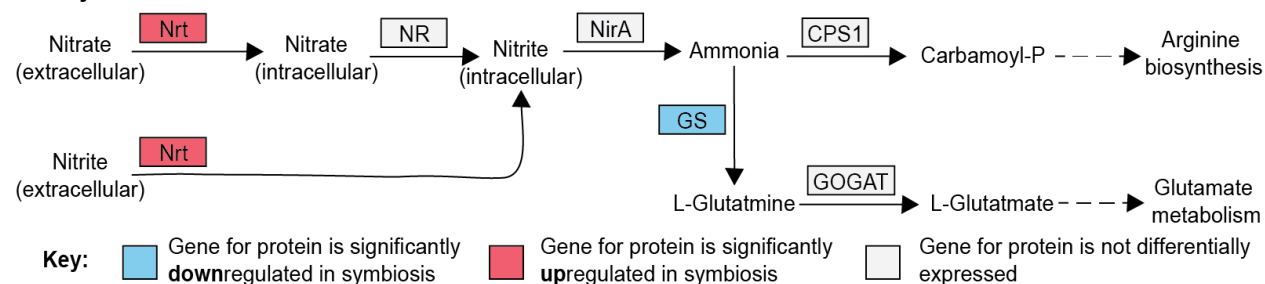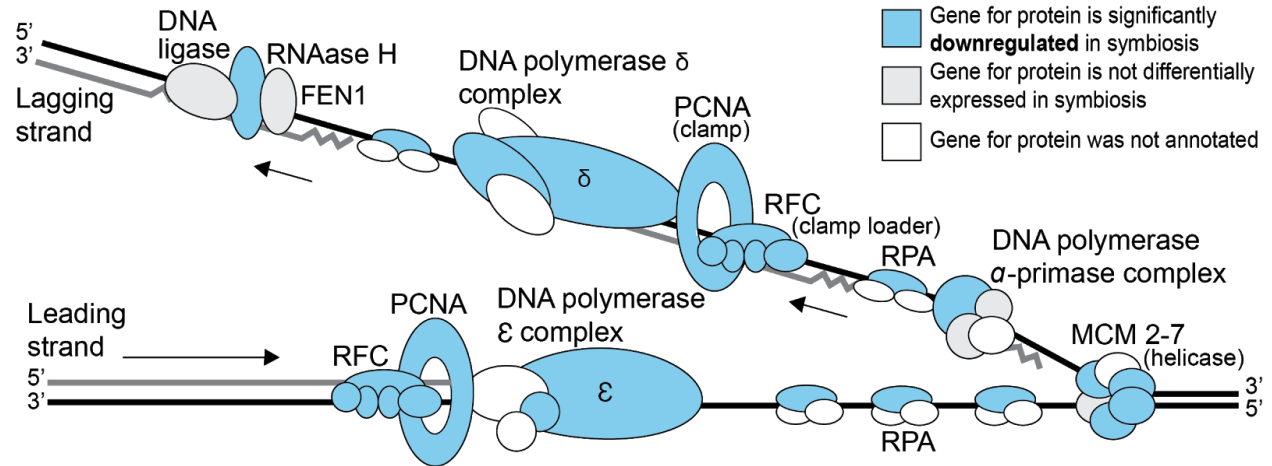**Figure S3.7. Genes in the Cell Cycle KEGG pathway significantly up- and downregulated in symbiotic *P. cordata* (A) and *P. jahnii* (B).** Proteins for which genes were significantly upregulated in symbiosis (log2FC > 1 and padj ≤ 0.05) are shaded red and proteins for which genes were significantly downregulated in symbiosis (log2FC < −1 and padj ≤ 0.05) are shaded blue. Genes for proteins shaded grey were not significantly differentially expressed in symbiosis.

If genes for the protein were not annotated in the transcriptome, it is unshaded. Black lines indicate activation interactions and grey lines indicate inhibitory interactions. The majority of genes in the pathway are significantly downregulated for both species. Upregulated genes in symbiotic *P. jahnii* are involved in progression from G1 to S phase, but no genes involved in progression past S phase are upregulated in symbiosis.

A. *P. cordata*
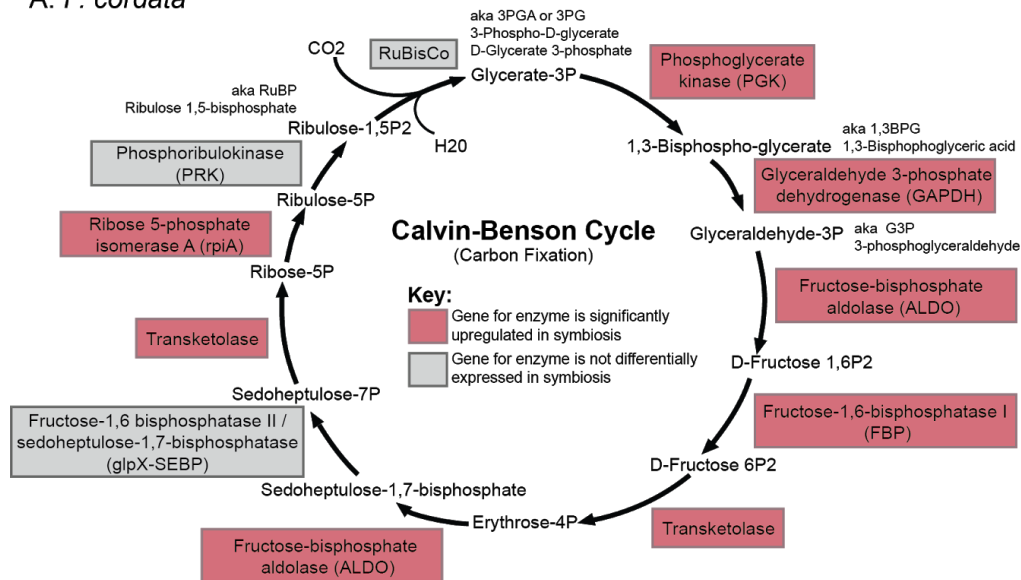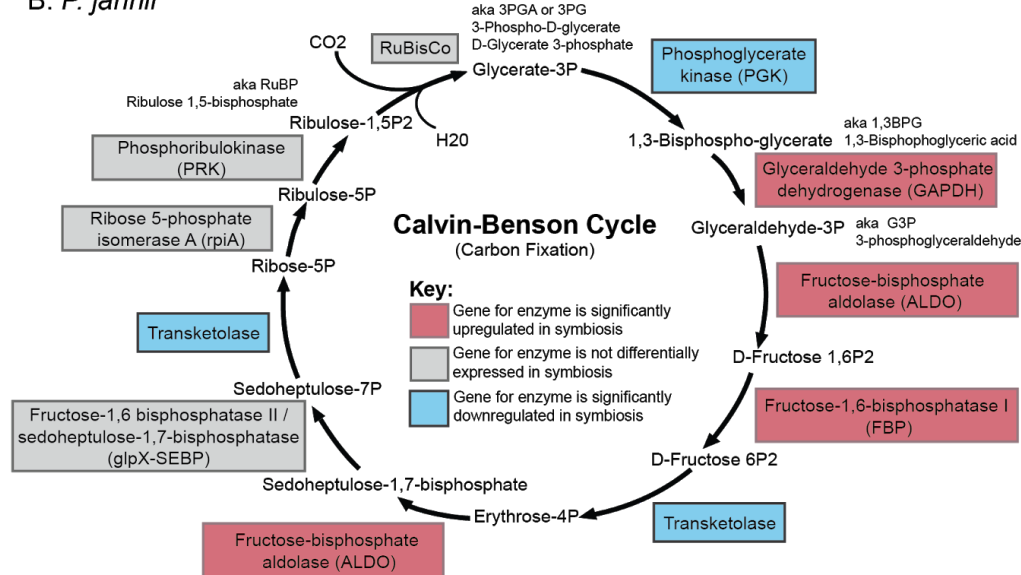
B. *P. jahnii*

Figure S3.8. Calvin-Benson cycle (carbon fixation) genes significantly up- and downregulated in symbiotic *P. cordata* (A) and *P. jahnii* (B). Enzymes for which genes were significantly upregulated in symbiosis (log2FC > 1 and padj ≤ 0.05) are shaded red and enzymes for which genes were significantly downregulated in symbiosis (log2FC < −1 and padj

≤ 0.05) are shaded blue. The genes for enzymes in grey shaded boxes were not significantly differentially expressed in symbiosis.
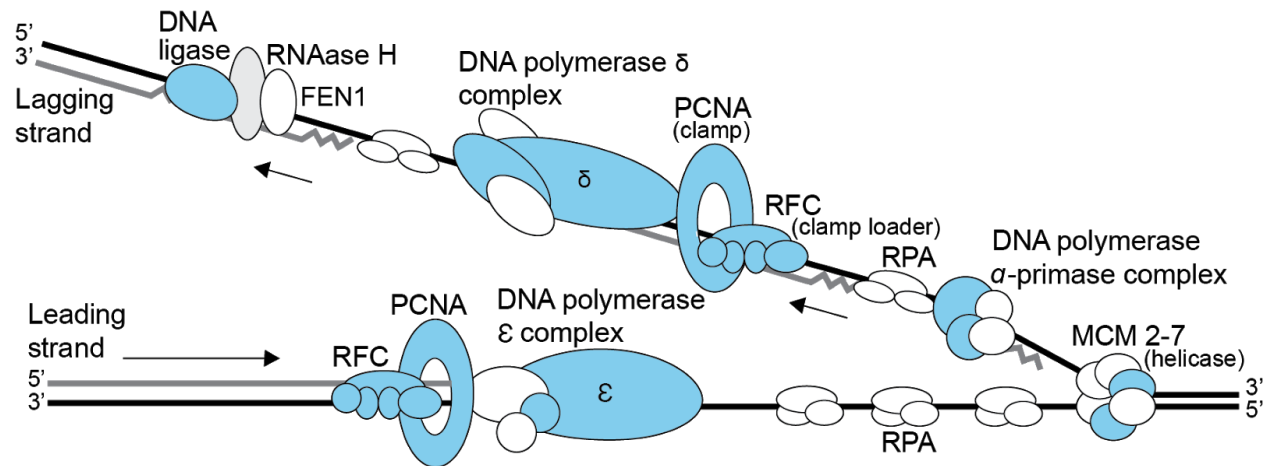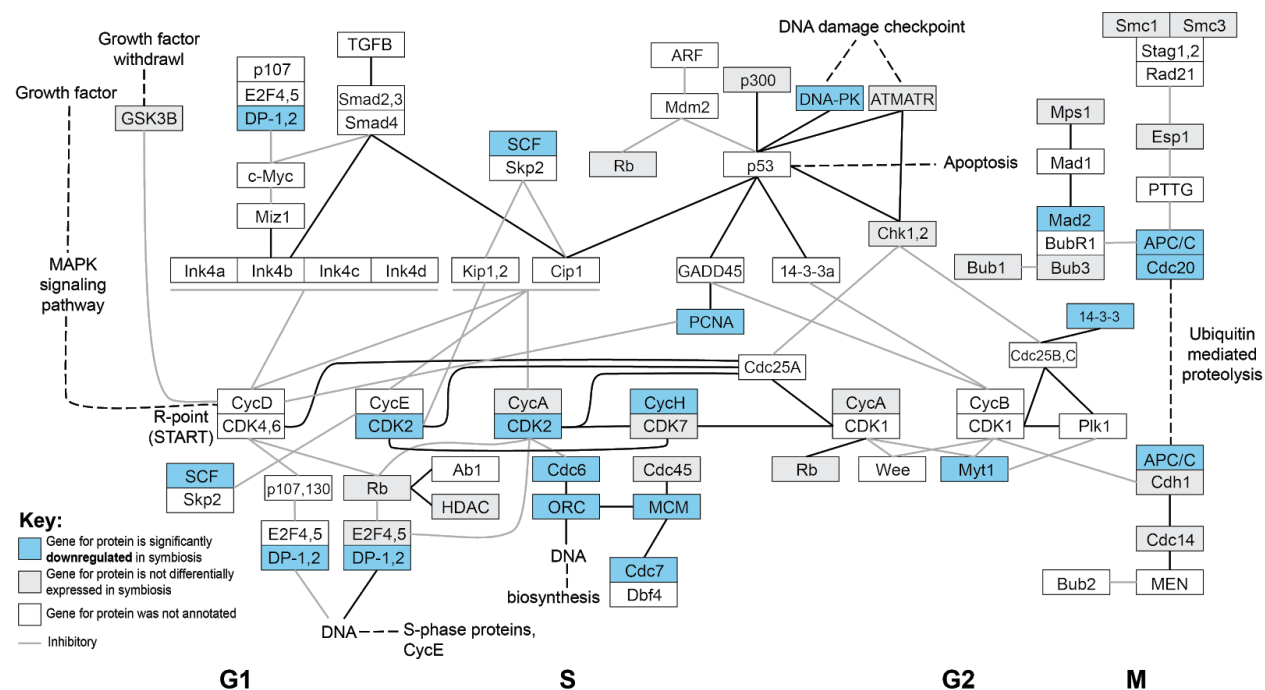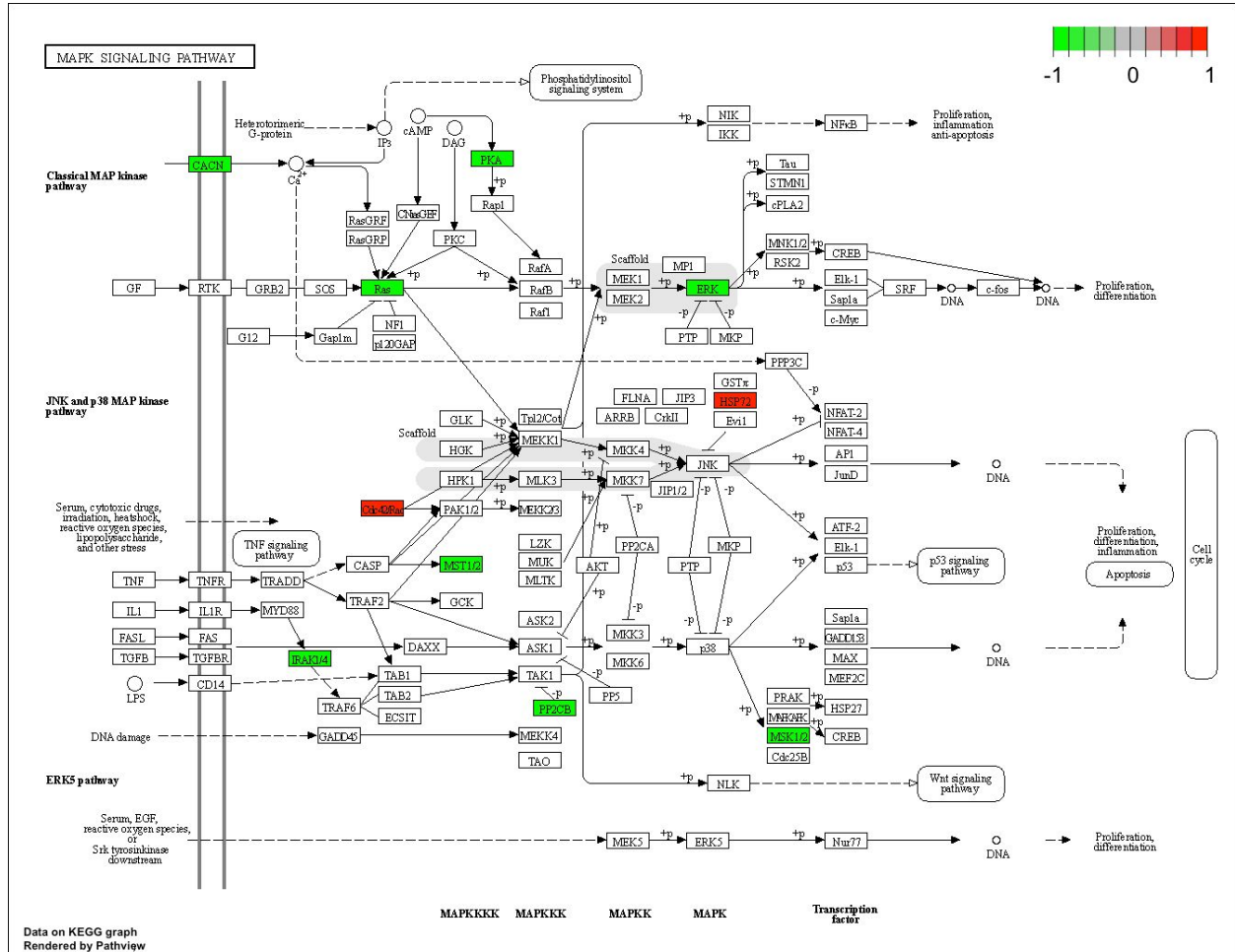
A. *P. cordata*
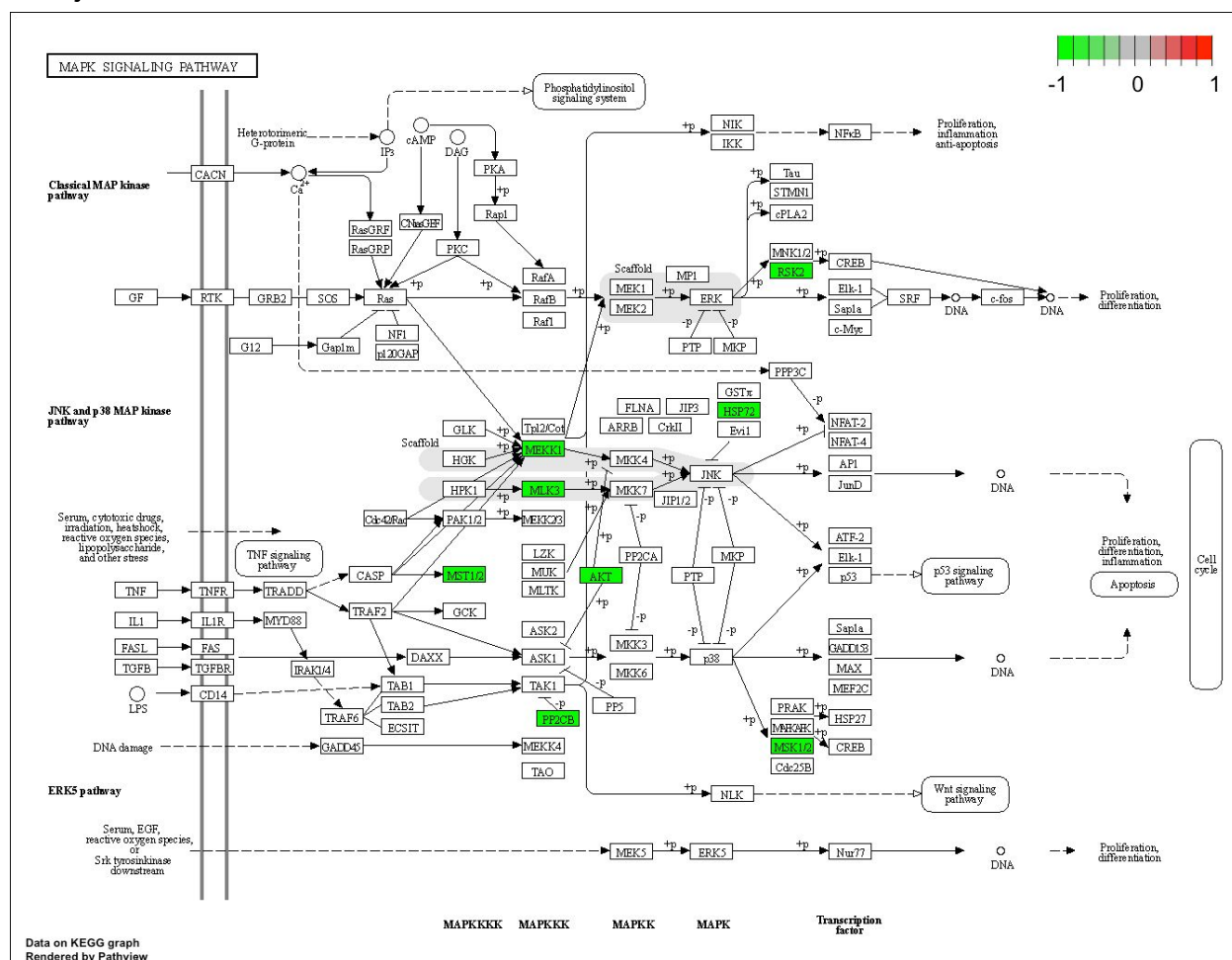
B. *P. jahnii*

Figure S3.8. Mitogen Activated Protein Kinase (MAPK) Signaling Pathway KEGG graphs with genes significantly up- and downregulated in symbiotic *P. cordata* (A) and *P. jahnii* (B). Genes significantly upregulated are colored red and genes significantly downregulated are colored green. Color intensity maxes out at a log fold change of -1 or 1, but log fold changes for genes in the plot may be higher (upregulated) or lower (downregulated) than these values. Genes represented by open boxes are not significantly differentially expressed or are not annotated in the transcriptome. Key genes in the chemical MAP kinase pathway are significantly downregulated in both species: Ras and ERK in *P. cordata* and RSK2 in *P. jahnii.* In addition, key genes in the JNK and p38 kinase pathway (MEKK and AKT) are significantly downregulated in symbiotic *P. jahnii*.

Thesis Conclusions

This thesis aimed to characterize key ecological and biological traits of Acantharea-*Phaeocystis* photosymbioses—relationships that have been historically understudied, despite being of high ecological and evolutionary significance. By pairing non-destructive, in-situ observations with environmental DNA sampling, high acantharian abundances were measured in the low-nutrient surface waters of the western North Pacific, demonstrating that acantharian abundances have been consistent across time and space. Moreover, results from these methodologies provided new evidence regarding acantharian vertical distributions that challenges current hypotheses about photosymbiotic acantharian life cycles by presenting the possibility that, like asymbiotic acantharians, photosymbiotic acantharians may also migrate to deep water reproduce. By taking advantage of advances in culture-independent, single-cell sequencing, individual acantharians were revealed to harbor diverse symbiont communities, thus demonstrating that acantharians recruit symbionts on more than one occasion. These results suggest that symbionts do not divide within hosts and that hosts may be exploiting symbionts. By further analyzing symbiont gene expression within individual hosts, it was shown that symbiont photosynthesis is enhanced *in hospite* and that symbiont cell division is indeed repressed. Together, these findings represent major advances in understanding both the current ecological significance of Acantharea-*Phaeocystis* photosymbioses and their functioning at the molecular level.

Earlier studies established that photosymbiotic acantharians were very abundant in the low-nutrient surface waters of the North Atlantic Subtropical Gyre (NASG), the eastern North

Pacific Subtropical Gyre (NPSG), and the equatorial Pacific (Michaels, 1988, 1991; Caron et al.,

1995; Michaels et al., 1995; Stoecker et al., 1996). These studies also demonstrated that

sampling methodology severely affected recorded abundances and that studies using traditional

plankton sampling techniques underestimated acantharian abundance. The results presented here

showed that acantharians were present in the low nutrient surface waters at the western edge of

the NPSG at comparable abundances as were reported for the eastern NPSG, NASG and

equatorial Pacific several decades ago, demonstrating that their abundance has been sustained in

time and space and confirming that acantharians continue to be major contributors to primary

production in the global ocean. These results are significant in the context of global climate

change—photosymbiotic acantharians can be important to carbon sequestration and their

skeletons are less susceptible to increased dissolution with rising pCO_2 than other

photosymbiotic hosts (Hallock, 2000; Fox et al., 2020)—and continued targeted monitoring will

help delineate how acantharian relative and absolute abundance is affected by changing ocean

chemistry and temperatures. Importantly, the images annotated for this thesis will facilitate

continued monitoring of acantharian abundance using high-throughput imaging.

Many aspects of acantharian biology are still only hypothesized, including their life

cycles and even their morphology for some clades. The in-situ imaging performed for this thesis

was able to provide additional evidence for and against existing hypotheses regarding

acantharian biology. First, opposing trends in imaged cell abundance and relative abundance of

sequencing reads from basal clades with unknown morphology supports hypotheses based on

evolutionary relationships that acantharians in these basal clades do not possess the characteristic

star-shaped skeletons usually used to identify acantharians. Second, the lack of differentiation in

sequence abundances deriving from asymbiotic versus symbiotic acantharians in deep waters challenges the idea that only asymbiotic clades sink into deep water to reproduce. Better understanding of acantharian life-cycles and morphologies is important to interpreting acantharian contributions to biogeochemical cycles. For instance, if basal clades do not biomineralize, sequences deriving from these clades should be excluded if researchers are relating environmental sequencing data with strontium and barium budgets (Bernstein et al., 1992). Furthermore, if photosymbiotic acantharians do in fact use buoyancy control to sink and release reproductive cells at depth, they actively transport surface production below the photic zone, thus increasing their contribution to carbon sequestration. Higher resolution camera systems that can image smaller cells and can better discriminate between different types of acantharians will be useful in more thoroughly characterizing these aspects of acantharian biology in the future.

It was previously discovered that acantharians host different species of *Phaeocystis* as their dominant symbiont depending on the location from which they are sampled (Decelle et al., 2012a), but the work presented in this thesis was the first to investigate intra-acantharian symbiont community compositions and to compare them to environmental availability of symbionts. Results revealed that individual acantharians host multiple symbiont strains, species, and genera, showing that hosts must collect new symbionts more than once. The intra-host symbiont communities did not directly reflect environmental communities, which could mean that hosts were selective in their symbiont uptake or that hosts maintain symbionts for extended periods of time—at least long enough for environmental communities to shift outside of hosts. Several lines of evidence—including the presence of symbionts within hosts that were not

present in the waters where hosts were sampled and confocal fluorescence microscopy demonstrating hosts do not systematically digest symbionts—point towards extended symbiont maintenance. The combination of multiple symbiont uptake events and extended maintenance suggest that symbionts are not dividing within hosts. Hosts need to keep recruiting new symbionts as their metabolic needs increase because symbionts are not dividing, and since symbionts are not dividing, hosts do not need to digest symbionts to prevent overgrowth (and digesting symbionts would mean they have to recruit more often). Results from differential gene expression analysis confirmed that symbiont cell division is inhibited within hosts.

Differential gene expression between *Phaeocystis* symbionts within individual hosts and free-living *Phaeocystis* showed that symbionts are actively photosynthesizing in hosts and rather than producing storage lipids or carbohydrates, symbionts seem to be exporting excess fixed carbon as small organonitrogen compounds. In return, hosts do seem to maintain symbionts in nutrient replete conditions; symbiont gene expression revealed no evidence for nitrogen or phosphorus limitation. Instead of limiting access to nutrients, it appears that hosts stall symbiont cell cycles by manipulating the MAPK signaling pathway, either through chemical or mechanical signaling. By influencing symbionts in this way, hosts have much finer control over symbiont populations and simultaneously ensure that symbionts maintain photosynthetic output. The data presented in this thesis cannot determine whether symbiont transformation is reversible, but it seems improbable that symbionts could overcome the extreme remodeling that occurs within hosts, which would rule out mutualism in this relationship. Overall, these results show that Acantharea-*Phaeocystis* photosymbioses are rather unique among better-studied symbioses, such as Cnidaria-dinoflagellate relationships, in which hosts limit nutrient availability to slow

symbiont division but as symbionts nevertheless continue to divide, hosts digest or expel excess

symbionts (Titlyanov et al., 1996; Xiang et al., 2020). Moreover, the more complex interactions

between acantharians and *Phaeocystis* indicate a higher level of symbiont integration and could

be evidence that the relationship is advancing towards a more complete integration of algal cells

as plastids. For example, the diatoms that are now obligate, permanent tertiary endosymbionts to

a group of dinoflagellates (i.e. dinotoms) exhibit similar traits to symbiotic *Phaeocystis*: they

have lost their distinctive cell wall, motility, and ability to divide mitotically, but have retained

their full nucleus, voluminous cytoplasm, and independent mitochondria (Imanian et al., 2012).

Considering the uniqueness of Acantharea-*Phaeocystis* photosymbioses along with the

high level of symbiont integration that was demonstrated in this thesis, further study of the

system is clearly warranted. Looking forward, follow-up work should further assess seasonal

cycles and long-term changes in photosymbiotic acantharian abundances and distributions and

should consider adaptations in photosymbiotic acantharians that allow them to select and

manipulate *Phaeocystis* symbionts. This thesis necessarily focused on differential gene

expression in symbionts because a comparison for hosts does not yet exist. As sequencing

technologies continue to advance, comparative genomics and transcriptomics between symbiotic

and asymbiotic acantharians will help uncover whether any gene transfer has occurred from

symbionts to photosymbiotic hosts and to what extent hosts have evolved adaptations specific for

symbiosis. In conclusion, this thesis showed that acantharians are abundant and important

contributors to primary production in the western North Pacific and demonstrated that

photosymbioses between acantharian hosts and *Phaeocystis* symbionts are unique,

highly-integrated relationships, whose further study can improve understanding of chloroplast

evolution.

References

Alcolombri, U., Ben-Dor, S., Feldmesser, E., Levin, Y., Tawfik, D. S., and Vardi, A. (2015). Identification of the algal dimethyl sulfide–releasing enzyme: A missing link in the marine sulfur cycle. *Science* 348, 1466–1469.

Alipanah, L., Rohloff, J., Winge, P., Bones, A. M., and Brembu, T. (2015). Whole-cell response to nitrogen deprivation in the diatom Phaeodactylum tricornutum. *J. Exp. Bot.* 66, 6281–6296.

Alipanah, L., Winge, P., Rohloff, J., Najafi, J., Brembu, T., and Bones, A. M. (2018). Molecular adaptations to phosphorus deprivation and comparison with nitrogen deprivation responses in the diatom Phaeodactylum tricornutum. *PLoS One* 13, e0193335.

Anderson, O. R. (1996). The Physiological Ecology of Planktonic Sarcodines with Applications to Paleoecology: Patterns in Space and Time. *J. Eukaryot. Microbiol.* 43, 261–274.

Andersson, B., and Barber, J. (1994). "Composition, Organization, and Dynamics of Thylakoid Membranes," in *Advances in Molecular and Cell Biology*, eds. E. E. Bittar and J. Barber (Elsevier), 1–53.

Annis, E. R., and Cook, C. B. (2002). Alkaline phosphatase activity in symbiotic dinoflagellates (zooxanthellae) as a biological indicator of environmental phosphate exposure. *Mar. Ecol. Prog. Ser.* 245, 11–20.

Archibald, J. M. (2015). Endosymbiosis and Eukaryotic Cell Evolution. *Curr. Biol.* 25, R911–21.

Bae, S. Y., Kim, G. D., Jeon, J.-E., Shin, J., and Lee, S. K. (2013). Anti-proliferative effect of (19Z)-halichondramide, a novel marine macrolide isolated from the sponge Chondrosia corticata, is associated with G2/M cell cycle arrest and suppression of mTOR signaling in human lung cancer cells. *Toxicol. In Vitro* 27, 694–699.

Balzano, S., Corre, E., Decelle, J., Sierra, R., Wincker, P., Da Silva, C., et al. (2015). Transcriptome analyses to investigate symbiotic relationships between marine protists. *Front. Microbiol.* 6, 98.

Becker, K. W., Collins, J. R., Durham, B. P., Groussman, R. D., White, A. E., Fredricks, H. F., et al. (2018). Daily changes in phytoplankton lipidomes reveal mechanisms of energy storage in the open ocean. *Nat. Commun.* 9, 5179.

Beinart, R. A. (2019). The Significance of Microbial Symbionts in Ecosystem Processes. *mSystems* 4. doi:10.1128/mSystems.00127-19.

Belcher, A., Manno, C., Thorpe, S., and Tarling, G. (2018). Acantharian cysts: high flux occurrence in the bathypelagic zone of the Scotia Sea, Southern Ocean. *Mar. Biol.* 165, 117.

Benson, D. A., Karsch-Mizrachi, I., Clark, K., Lipman, D. J., Ostell, J., and Sayers, E. W. (2012). GenBank. *Nucleic Acids Res.* 40, D48–53.

Béraud, E., Gevaert, F., Rottier, C., and Ferrier-Pagès, C. (2013). The response of the scleractinian coral Turbinaria reniformis to thermal stress depends on the nitrogen status of the coral holobiont. *J. Exp. Biol.* 216, 2665–2674.

Bernstein, R. E., Byrne, R. H., Betzer, P. R., and Greco, A. M. (1992). Morphologies and transformations of celestite in seawater: The role of acantharians in strontium and barium geochemistry. *Geochim. Cosmochim. Acta* 56, 3273–3279.

Biard, T., and Ohman, M. D. (2019). Vertical niche definition of test-bearing protists (Rhizaria) into the twilight zone revealed by in situ imaging. *bioRxiv*, 573410. doi:10.1101/573410.

Biard, T., Stemmann, L., Picheral, M., Mayot, N., Vandromme, P., Hauss, H., et al. (2016). In situ imaging reveals the biomass of giant protists in the global ocean. *Nature* 532, 504–507.

Blank, R. J. (1987). Cell architecture of the dinoflagellate Symbiodinium sp. inhabiting the Hawaiian stony coral Montipora verrucosa. *Mar. Biol.* 94, 143–155.

Boettcher, K. J., Ruby, E. G., and McFall-Ngai, M. J. (1996). Bioluminescence in the symbiotic squid Euprymna scolopes is controlled by a daily biological rhythm. *Journal of Comparative Physiology A* 179, 65–73.

Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., et al. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 6, 90.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* doi:10.1038/s41587-019-0209-9.

Bråte, J., Krabberød, A. K., Dolven, J. K., Ose, R. F., Kristensen, T., Bjørklund, K. R., et al. (2012). Radiolaria associated with large diversity of marine alveolates. *Protist* 163, 767–777.

Burki, F., and Keeling, P. J. (2014). Rhizaria. *Curr. Biol.* 24, R103–7.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016a). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat.*

Methods 13, 581–583.

Callahan, B. J., Sankaran, K., Fukuyama, J. A., McMurdie, P. J., and Holmes, S. P. (2016b). Bioconductor Workflow for Microbiome Data Analysis: from raw reads to community analyses. *F1000Res.* 5, 1492.

Callahan, B., McMurdie, P., Rosen, M., Han, A., Johnson, A., and Holmes, S. (2016c). DADA2:High-resulution sample inference from illumina amplicon data. *Nat. Methods* 13.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421.

Caron, D. A., Michaels, A. F., Swanberg, N. R., and Howse, F. A. (1995). Primary productivity by symbiont-bearing planktonic sarcodines (Acantharia, Radiolaria, Foraminifera) in surface waters near Bermuda. *J. Plankton Res.* 17, 103–129.

Chen, C., MacCready, J. S., Ducat, D. C., and Osteryoung, K. W. (2018). The Molecular Machinery of Chloroplast Division. *Plant Physiol.* 176, 138–151.

Countway, P. D., Caron, D. A., Gast, R. J., and Savai, P. (2007). Comparison of protistan diversity in deep (2500 m) vs euphotic zone assemblages in the Sargasso Sea and Gulf Stream (N. Atlantic). *Environ. Microbiol.* 9, 1219–1232.

Countway, P. D., Vigil, P. D., Schnetzer, A., Moorthi, S. D., and Caron, D. A. (2010). Seasonal analysis of protistan community structure and diversity at the USC Microbial Observatory (San Pedro Channel, North Pacific Ocean). *Limnol. Oceanogr.* 55, 2381–2396.

Cronin, M., Ghosh, K., Sistare, F., Quackenbush, J., Vilker, V., and O'Connell, C. (2004). Universal RNA reference materials for gene expression. *Clin. Chem.* 50, 1464–1471.

Curson, A., Williams, B., Pinchbeck, B., Sims, L., Martínez, A., Rivera, P., et al. (2018). DSYB catalyses the key step of dimethylsulfoniopropionate biosynthesis in many phytoplankton, Nat. Microbiol., 3, 430--439.

Decelle, J. (2013). New perspectives on the functioning and evolution of photosymbiosis in plankton. *Communicative & Integrative Biology* 6, e24560. doi:10.4161/cib.24560.

Decelle, J., Colin, S., and Foster, R. A. (2015). "Photosymbiosis in Marine Planktonic Protists," in *Marine Protists: Diversity and Dynamics*, eds. S. Ohtsuka, T. Suzaki, T. Horiguchi, N. Suzuki, and F. Not (Tokyo: Springer Japan), 465–500.

Decelle, J., Martin, P., Paborstava, K., Pond, D. W., Tarling, G., Mahé, F., et al. (2013). Diversity, ecology and biogeochemistry of cyst-forming acantharia (radiolaria) in the oceans. *PLoS One* 8, e53598.

Decelle, J., and Not, F. (2015). Acantharia. *eLS* 59, 1–10.

Decelle, J., Probert, I., Bittner, L., Desdevises, Y., Colin, S., de Vargas, C., et al. (2012a). An original mode of symbiosis in open ocean plankton. *Proc. Natl. Acad. Sci. U. S. A.* 109, 18000–18005.

Decelle, J., Siano, R., Probert, I., Poirier, C., and Not, F. (2012b). Multiple microalgal partners in symbiosis with the acantharian Acanthochiasma sp. (Radiolaria). *Symbiosis* 58, 233–244.

Decelle, J., Stryhanyuk, H., Gallet, B., Veronesi, G., Schmidt, M., Balzano, S., et al. (2019). Algal Remodeling in a Ubiquitous Planktonic Photosymbiosis. *Curr. Biol.* 29, 968–978.e4.

Decelle, J., Suzuki, N., Mahé, F., de Vargas, C., and Not, F. (2012c). Molecular phylogeny and morphological evolution of the Acantharia (Radiolaria). *Protist* 163, 435–450.

Dennett, M. R., Caron, D. A., Michaels, A. F., Gallager, S. M., and Davis, C. S. (2002). Video plankton recorder reveals high abundances of colonial Radiolaria in surface waters of the central North Pacific. *J. Plankton Res.* 24, 797–805.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., et al. (2015). Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. *Science* 348, 1261605.

Dolan, J. R. (1997). Phosphorus and ammonia excretion by planktonic protists. *Mar. Geol.* 139, 109–122.

Douglas, A. E. (1998). Host benefit and the evolution of specialization in symbiosis. *Heredity* 81, 599–603.

Douglas, A. E. (2010). *The Symbiotic Habit*. Princeton University Press.

Eddy, S. (2010). HMMER3: a new generation of sequence homology search software. *URL: http://hmmer. janelia. Org*.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.

Edgcomb, V. P., Kysela, D. T., Teske, A., de Vera Gomez, A., and Sogin, M. L. (2002). Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment. *Proc. Natl. Acad. Sci. U. S. A.* 99, 7658–7662.

Falcon, S., and Gentleman, R. (2007). Using GOstats to test gene lists for GO term association. *Bioinformatics* 23, 257–258.

Febvre, J., and Febvre-Chevalier, C. (1979). Ultrastructural study of zooxanthellae of three species of Acantharia (Protozoa: Actinopoda), with details of their taxonomic position in the prymnesiales (Prymnesiophyceae, Hibberd, 1976). *Journal of the Marine Biological Association of the United Kingdom* 59, 215–226. doi:10.1017/s0025315400046294.

Febvre, J., and Febvre-Chevalier, C. (2001). Acantharia. *eLS* 32, 211.

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., et al. (2014). Pfam: the protein families database. *Nucleic Acids Res.* 42, D222–30.

Fisher, M., Liu, B., Glennon, P. E., Southgate, K. M., Sale, E. M., Sale, G. J., et al. (2001). Downregulation of the ERK 1 and 2 mitogen activated protein kinases using antisense oligonucleotides inhibits proliferation of porcine vascular smooth muscle cells. *Atherosclerosis* 156, 289–295.

Fishman, Y., Zlotkin, E., and Sher, D. (2008). Expulsion of Symbiotic Algae during Feeding by the Green Hydra – a Mechanism for Regulating Symbiont Density? *PLoS ONE* 3, e2603. doi:10.1371/journal.pone.0002603.

Fontanez, K. M., Eppley, J. M., Samo, T. J., Karl, D. M., and DeLong, E. F. (2015). Microbial community structure and function on sinking particles in the North Pacific Subtropical Gyre. *Front. Microbiol.* 6, 469.

Fox, L., Stukins, S., Hill, T., and Miller, C. G. (2020). Quantifying the Effect of Anthropogenic Climate Change on Calcifying Plankton. *Sci. Rep.* 10, 1620.

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152.

Garcia, J. R., and Gerardo, N. M. (2014). The symbiont side of symbiosis: do microbes really benefit? *Frontiers in Microbiology* 5. doi:10.3389/fmicb.2014.00510.

Gast, R. J., and Caron, D. A. (1996). Molecular phylogeny of symbiotic dinoflagellates from planktonic foraminifera and radiolaria. *Mol. Biol. Evol.* 13, 1192–1197.

Gast, R. J., and Caron, D. A. (2001). Photosymbiotic associations in planktonic foraminifera and radiolaria. *Hydrobiologia* 461, 1–7.

Gates, R. D., Hoegh-Guldberg, O., McFall-Ngai, M. J., Bil, K. Y., and Muscatine, L. (1995). Free amino acids exhibit anthozoan "host factor" activity: they induce the release of photosynthate from symbiotic dinoflagellates in vitro. *Proc. Natl. Acad. Sci. U. S. A.* 92, 7430–7434.

Gérard, C., and Goldbeter, A. (2014). The balance between cell cycle arrest and cell proliferation: control by the extracellular matrix and by contact inhibition. *Interface Focus* 4, 20130075.

Gilg, I. C., Amaral-Zettler, L. A., Countway, P. D., Moorthi, S., Schnetzer, A., and Caron, D. A. (2010). Phylogenetic affiliations of mesopelagic acantharia and acantharian-like environmental 18S rRNA genes off the southern California coast. *Protist* 161, 197–211.

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., and Egozcue, J. J. (2017). Microbiome

Datasets Are Compositional: And This Is Not Optional. *Front. Microbiol.* 8, 2224.

Gong, J., Dong, J., Liu, X., and Massana, R. (2013). Extremely high copy numbers and polymorphisms of the rDNA operon estimated from single cell analysis of oligotrich and peritrich ciliates. *Protist* 164, 369–379.

Gong, W., and Marchetti, A. (2019). Estimation of 18S gene copy number in marine eukaryotic plankton using a next-generation sequencing approach. *Frontiers in Marine Science*. Available at: https://pdfs.semanticscholar.org/4589/e20d3cbd666d8fd0b35a383f87d6cf4d4e05.pdf.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652.

Greer, A. T., Cowen, R. K., Guigand, C. M., McManus, M. A., Sevadjian, J. C., and Timmerman, A. H. V. (2013). Relationships between phytoplankton thin layers and the fine-scale vertical distributions of two trophic levels of zooplankton. *J. Plankton Res.* 35, 939–956.

Grossmann, M. M., Gallager, S. M., and Mitarai, S. (2015). Continuous monitoring of near-bottom mesoplankton communities in the East China Sea during a series of typhoons. *J. Oceanogr.* 71, 115–124.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., et al. (2013). The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 41, D597–604.

Gutierrez-Rodriguez, A., Stukel, M. R., Lopes Dos Santos, A., Biard, T., Scharek, R., Vaulot, D., et al. (2019). High contribution of Rhizaria (Radiolaria) to vertical export in the California Current Ecosystem revealed by DNA metabarcoding. *ISME J.* 13, 964–976.

Haas, B., Papanicolaou, A., and Others (2016). TransDecoder (find coding regions within transcripts). *Google Scholar*.

Hallock, P. (2000). Symbiont-Bearing Foraminifera: Harbingers of Global Change? *Micropaleontology* 46, 95–104.

Hamm, C. E. (2000). Architecture, ecology and biogeochemistry of Phaeocystis colonies. *J. Sea Res.* 43, 307–315.

Hamm, C. E., Simson, D. A., Merkel, R., and Smetacek, V. (1999). Colonies of Phaeocystis globosa are protected by a thin but tough skin. *Mar. Ecol. Prog. Ser.* 187, 101–111.

Hinde, R., and Trautman, D. A. (2001). Symbiosomes. *Symbiosis Mech. Model Syst.*, 207–220.

Hirose, K., and Kamiya, H. (2003). Vertical Nutrient Distributions in the Western North Pacific

Ocean: Simple Model for Estimating Nutrient Upwelling, Export Flux and Consumption Rates. *J. Oceanogr.* 59, 149–161.

Hohman, T. C., McNeil, P. L., and Muscatine, L. (1982). Phagosome-lysosome fusion inhibited by algal symbionts of Hydra viridis. *J. Cell Biol.* 94, 56–63.

Hu, S. K., Connell, P. E., Mesrop, L. Y., and Caron, D. A. (2018). A Hard Day's Night: Diel Shifts in Microbial Eukaryotic Activity in the North Pacific Subtropical Gyre. doi:10.3389/fmars.2018.00351.

Imanian, B., Pombert, J.-F., Dorrell, R. G., Burki, F., and Keeling, P. J. (2012). Tertiary endosymbiosis in two dinotoms has generated little change in the mitochondrial genomes of their dinoflagellate hosts and diatom endosymbionts. *PLoS One* 7, e43763.

Itoh, R., Takahashi, H., Toda, K., Kuroiwa, H., and Kuroiwa, T. (1996). Aphidicolin uncouples the chloroplast division cycle from the mitotic cycle in the unicellular red alga Cyanidioschyzon merolae. *Eur. J. Cell Biol.* 71, 303–310.

Jahnke, J. (1989). The light and temperature dependence of growth rate and elemental composition of Phaeocystis globosa scherffel and P. Pouchetii (HAR.) Lagerh. in batch cultures. *Neth. J. Sea Res.* 23, 15–21.

Jamwal, S. V., Mehrotra, P., Singh, A., Siddiqui, Z., Basu, A., and Rao, K. V. S. (2016). Mycobacterial escape from macrophage phagosomes to the cytoplasm represents an alternate adaptation mechanism. *Sci. Rep.* 6, 23089.

Janse, I., van Rijssel, M., Gottschal, J. C., Lancelot, C., and Gieskes, W. W. C. (1996). Carbohydrates in the North Sea during spring blooms of Phaeocystis: a specific fingerprint. *Aquat. Microb. Ecol.* 10, 97–103.

Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *J. Mol. Biol.* 428, 726–731.

Keeling, P. J. (2004). Diversity and evolutionary history of plastids and their hosts. *Am. J. Bot.* 91, 1481–1493.

Keeling, P. J., Burki, F., Wilcox, H. M., Allam, B., Allen, E. E., Amaral-Zettler, L. A., et al. (2014). The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol.* 12, e1001889.

Keeling, P. J., and McCutcheon, J. P. (2017). Endosymbiosis: The feeling is not mutual. *J. Theor. Biol.* 434, 75–79.

Kodama, Y., Inouye, I., and Fujishima, M. (2011). Symbiotic Chlorella vulgaris of the ciliate Paramecium bursaria plays an important role in maintaining perialgal vacuole membrane

functions. *Protist* 162, 288–303.

Koike, K., Jimbo, M., Sakai, R., Kaeriyama, M., Muramoto, K., Ogata, T., et al. (2004). Octocoral chemical signaling selects and controls dinoflagellate symbionts. *Biol. Bull.* 207, 80–86.

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12, 323.

Li, M., Shi, X., Guo, C., and Lin, S. (2016). Phosphorus Deficiency Inhibits Cell Division But Not Growth in the Dinoflagellate Amphidinium carterae. *Front. Microbiol.* 7, 826.

Lindehoff, E., Granéli, E., and Glibert, P. M. (2011). Nitrogen uptake kinetics of Prymnesium parvum (Haptophyte). *Harmful Algae* 12, 70–76.

Liss, P. S., Malin, G., Turner, S. M., and Holligan, P. M. (1994). Dimethyl sulphide and Phaeocystis: A review. *J. Mar. Syst.* 5, 41–53.

Li, T., Guo, C., Zhang, Y., Wang, C., Lin, X., and Lin, S. (2018). Identification and Expression Analysis of an Atypical Alkaline Phosphatase in Emiliania huxleyi. *Front. Microbiol.* 9, 2156.

Liu, Z., Mesrop, L. Y., Hu, S. K., and Caron, D. A. (2019). Transcriptome of Thalassicolla nucleata holobiont reveals details of a radiolarian symbiotic relationship. *Frontiers in Marine Science*. Available at: https://pdfs.semanticscholar.org/62de/73521d55b6183ac41c32ae8d71324648efb8.pdf.

Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemmann, L., et al. (2019). Globally Consistent Quantitative Observations of Planktonic Ecosystems. *Front. Mar. Sci.* 6, 1705.

López-García, P., Rodríguez-Valera, F., Pedrós-Alió, C., and Moreira, D. (2001). Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* 409, 603–607.

Lowe, C. D., Minter, E. J., Cameron, D. D., and Brockhurst, M. A. (2016). Shining a Light on Exploitative Host Control in a Photosynthetic Endosymbiosis. *Curr. Biol.* 26, 207–211.

Lumpkin, R., Grodsky, S. A., Centurioni, L., Rio, M.-H., Carton, J. A., and Lee, D. (2013). Removing Spurious Low-Frequency Variability in Drifter Velocities. *Journal of Atmospheric and Oceanic Technology* 30, 353–360. doi:10.1175/jtech-d-12-00139.1.

Mars Brisbin, M., Mesrop, L. Y., Grossmann, M. M., and Mitarai, S. (2018). Intra-host Symbiont Diversity and Extended Symbiont Maintenance in Photosymbiotic Acantharea (Clade F). *Front. Microbiol.* 9, 1998.

Mars Brisbin, M., and Mitarai, S. (2019). Differential Gene Expression Supports a Resource-Intensive, Defensive Role for Colony Production in the Bloom-Forming

Haptophyte, Phaeocystis globosa. *J. Eukaryot. Microbiol.* 66, 788–801.

Marte, B. M., and Downward, J. (1997). PKB/Akt: connecting phosphoinositide 3-kinase to cell survival and beyond. *Trends Biochem. Sci.* 22, 355–358.

Martin, P., Allen, J. T., Cooper, M. J., Johns, D. G., Lampitt, R. S., Sanders, R., et al. (2010). Sedimentation of acantharian cysts in the Iceland Basin: Strontium as a ballast for deep ocean particle flux, and implications for acantharian reproductive strategies. *Limnol. Oceanogr.* 55, 604–614.

McFadden, G. I. (2014). Origin and evolution of plastids and photosynthesis in eukaryotes. *Cold Spring Harb. Perspect. Biol.* 6, a016105.

McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8, e61217.

Medlin, L., and Edvardsen, B. (2007). Molecular systematics of Haptophyta. *Unravelling the algae*, 183–196. doi:10.1201/9780849379901.ch10.

Michaels, A. F. (1988). Vertical distribution and abundance of Acantharia and their symbionts. *Mar. Biol.* 97, 559–569.

Michaels, A. F. (1991). Acantharian abundance and symbiont productivity at the VERTEX seasonal station. *J. Plankton Res.* 13, 399–418.

Michaels, A. F., Caron, D. A., Swanberg, N. R., Howse, F. A., and Michaels, C. M. (1995). Planktonic sarcodines (Acantharia, Radiolaria, Foraminifera) in surface waters near Bermuda: abundance, biomass and vertical flux. *J. Plankton Res.* 17, 131–163.

Mitchell, A., Chang, H.-Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., et al. (2015). The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res.* 43, D213–21.

Mitra, A., Flynn, K. J., Burkholder, J. M., Berge, T., Calbet, A., Raven, J. A., et al. (2014). The role of mixotrophic protists in the biological carbon pump. *Biogeosciences* 11, 995–1005. doi:10.5194/bg-11-995-2014.

Möller, K. O., John, M. S., Temming, A., Floeter, J., Sell, A. F., Herrmann, J.-P., et al. (2012). Marine snow, zooplankton and thin layers: indications of a trophic link from small-scale sampling with the Video Plankton Recorder. *Mar. Ecol. Prog. Ser.* 468, 57–69.

Moon-van der Staay, S. Y., van der Staay, G. W. M., Guillou, L., Vaulot, D., Claustre, H., and Medlin, L. K. (2000). Abundance and diversity of prymnesiophytes in the picoplankton coμmunity from the equatorial Pacific Ocean inferred from 18S rDNA sequences. *Limnol. Oceanogr.* 45, 98–109.

Morrison, D. K. (2012). MAP kinase pathways. *Cold Spring Harb. Perspect. Biol.* 4.

doi:10.1101/cshperspect.a011254.

Nigg, E. A. (1995). Cyclin-dependent protein kinases: key regulators of the eukaryotic cell cycle. *Bioessays* 17, 471–480.

Not, F., Gausling, R., Azam, F., Heidelberg, J. F., and Worden, A. Z. (2007). Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environ. Microbiol.* 9, 1233–1252.

Not, F., Probert, I., Gerikas Ribeiro, C., Crenn, K., Guillou, L., Jeanthon, C., et al. (2016). "Photosymbiosis in Marine Pelagic Environments," in *The Marine Microbiome: An Untapped Source of Biodiversity and Biotechnological Potential*, eds. L. J. Stal and M. S. Cretoiu (Cham: Springer International Publishing), 305–332.

Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019). vegan: Community Ecology Package. R package version 2.5-4. Available at: https://CRAN.R-project.org/package=vegan.

Orenstein, E. C., Beijbom, O., Peacock, E. E., and Sosik, H. M. (2015). WHOI-Plankton- A Large Scale Fine Grained Visual Recognition Benchmark Dataset for Plankton Classification. *arXiv [cs.CV]*. Available at: http://arxiv.org/abs/1510.00745.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419.

Peacock, E. E., Olson, R. J., and Sosik, H. M. (2014). Parasitic infection of the diatom Guinardia delicatula, a recurrent and ecologically important phenomenon on the New England Shelf. *Mar. Ecol. Prog. Ser.* 503, 1–10.

Pernice, M. C., Giner, C. R., Logares, R., Perera-Bel, J., Acinas, S. G., Duarte, C. M., et al. (2016). Large variability of bathypelagic microbial eukaryotic communities across the world's oceans. *ISME J.* 10, 945–958.

Phipps, D. W., and Pardy, R. L. (1982). HOST ENHANCEMENT OF SYMBIONT PHOTOSYNTHESIS IN THE HYDRA-ALGAE SYMBIOSIS. *Biol. Bull.* 162, 83–94.

Probert, I., Siano, R., Poirier, C., Decelle, J., Biard, T., Tuji, A., et al. (2014). Brandtodinium gen. nov. and B. nutricula comb. Nov. (Dinophyceae), a dinoflagellate commonly found in symbiosis with polycystine radiolarians. *J. Phycol.* 50, 388–399.

Quaiser, A., Zivanovic, Y., Moreira, D., and López-García, P. (2011). Comparative metagenomics of bathypelagic plankton and bottom sediment from the Sea of Marmara. *ISME J.* 5, 285–304.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Glo, F. O., et al. (2013). The SILVA ribosomal RNA gene database project : improved data processing and web-based tools. 41,

590–596.

Quesada, I., Chin, W.-C., and Verdugo, P. (2006). Mechanisms of signal transduction in photo-stimulated secretion in Phaeocystis globosa. *FEBS Lett.* 580, 2201–2206.

Raina, J.-B., Tapiolas, D. M., Forêt, S., Lutz, A., Abrego, D., Ceh, J., et al. (2013). DMSP biosynthesis by an animal and its role in coral thermal stress response. *Nature* 502, 677–680.

R Core Team (2018). R: A language and environment for statistical computing. Available at: https://www.R-project.org/.

Read, B. A., Kegel, J., Klute, M. J., Kuo, A., Lefebvre, S. C., Maumus, F., et al. (2013). Pan genome of the phytoplankton Emiliania underpins its global distribution. *Nature* 499, 209–213.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.

Rodrigues, F. C. M., Hirata, N. S. T., Abello, A. A., Leandro, T., La Cruz, D., Lopes, R. M., et al. (2018). Evaluation of Transfer Learning Scenarios in Plankton Image Classification. in *VISIGRAPP (5: VISAPP)*, 359–366.

Ronkin, R. R. (1959). MOTILITY AND POWER DISSIPATION IN FLAGELLATED CELLS, ESPECIALLY CHLAMYDOMONAS. *Biol. Bull.* 116, 285–293.

Ronquist, F., and Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.

Sanz-Luque, E., Chamizo-Ampudia, A., Llamas, A., Galvan, A., and Fernandez, E. (2015). Understanding nitrate assimilation and its regulation in microalgae. *Front. Plant Sci.* 6, 899.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682.

Schnetzer, A., Moorthi, S. D., Countway, P. D., Gast, R. J., Gilg, I. C., and Caron, D. A. (2011). Depth matters: Microbial eukaryote diversity and community structure in the eastern North Pacific revealed through environmental gene libraries. *Deep Sea Res. Part I* 58, 16–26.

Schoemann, V., Becquevort, S., Stefels, J., Rousseau, V., and Lancelot, C. (2005). Phaeocystis blooms in the global ocean and their controlling mechanisms: a review. *J. Sea Res.* 53, 43–66.

Shaked, Y., and de Vargas, C. (2006). Pelagic photosymbiosis: rDNA assessment of diversity and evolution of dinoflagellate symbionts and planktonic foraminiferal hosts. *Mar. Ecol.*

Prog. Ser. 325, 59–71.

Shaul, Y. D., and Seger, R. (2007). The MEK/ERK cascade: from signaling specificity to diverse functions. *Biochim. Biophys. Acta* 1773, 1213–1226.

Sherr, B. F., Sherr, E. B., Caron, D. A., Vaulot, D., and Worden, A. Z. (2007). Oceanic Protists. *Oceanography* 20, 130–134.

Sibley, L. D., Weidner, E., and Krahenbuhl, J. L. (1985). Phagosome acidification blocked by intracellular Toxoplasma gondii. *Nature* 315, 416–419.

Sieracki, M. E., Benfield, M., Hanson, A., Davis, C., Pilskaln, C. H., Checkley, D., et al. (2010). Optical plankton imaging and analysis systems for ocean observation. *Proceedings of ocean Obs* 9, 21–25.

Sieracki, M. E., Poulton, N. J., Jaillon, O., Wincker, P., de Vargas, C., Rubinat-Ripoll, L., et al. (2019). Single cell genomics yields a wide diversity of small planktonic protists across major ocean ecosystems. *Sci. Rep.* 9, 6025.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212.

Solomon, C. M., Lessard, E. J., Keil, R. G., and Foy, M. S. (2003). Characterization of extracellular polymers of Phaeocystis globosa and P. antarctica. *Mar. Ecol. Prog. Ser.* 250, 81–89.

Soneson, C., Love, M. I., and Robinson, M. D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res.* 4, 1521.

Stoecker, D. K., Gustafson, D. E., and Verity, P. G. (1996). Micro-and mesoprotozooplankton at 140* W in the equatorial Pacific: heterotrophs and mixotrophs. *Aquat. Microb. Ecol.* 10, 273–282.

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D. M., Breiner, H.-W., et al. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* 19 Suppl 1, 21–31.

Sumiya, N. (2018). Mechanism of coordination between cell and chloroplast division in unicellular algae. *PLANT MORPHOLOGY* 30, 83–89.

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6, e21800.

Suzuki, N., and Not, F. (2015). "Biology and Ecology of Radiolaria," in *Marine Protists: Diversity and Dynamics*, eds. S. Ohtsuka, T. Suzaki, T. Horiguchi, N. Suzuki, and F. Not

(Tokyo: Springer Japan), 179–222.

Swanberg, N. R., and Caron, D. A. (1991). Patterns of sarcodine feeding in epipelagic oceanic plankton. *J. Plankton Res.* 13, 287–312.

Syrett, P. J., and Leftley, J. W. (2016). Nitrate and urea assimilation by algae. *Perspectives in experimental biology* 2, 221–234.

Takagi, H., Kimoto, K., Fujiki, T., Kurasawa, A., Moriya, K., and Hirano, H. (2016). Ontogenetic dynamics of photosymbiosis in cultured planktic foraminifers revealed by fast repetition rate fluorometry. *Mar. Micropaleontol.* 122, 44–52.

Takagi, H., Kimoto, K., Fujiki, T., Saito, H., Schmidt, C., Kucera, M., et al. (2019). Characterizing photosymbiosis in modern planktonic foraminifera. *Biogeosciences* 16, 3377–3396.

Tanabe, A. S., Nagai, S., Hida, K., Yasuike, M., Fujiwara, A., Nakamura, Y., et al. (2016). Comparative study of the validity of three regions of the 18S-rRNA gene for massively parallel sequencing-based monitoring of the planktonic eukaryote community. *Mol. Ecol. Resour.* 16, 402–414.

Terrado, R., Vincent, W. F., and Lovejoy, C. (2009). Mesopelagic protists: diversity and succession in a coastal Arctic ecosystem. *Aquat. Microb. Ecol.* Available at: https://www.int-res.com/abstracts/ame/v56/n1/p25-39/.

Titlyanov, E. A., Titlyanova, T. V., Leletkin, V. A., Tsukahara, J., van Woesik, R., and Yamazato, K. (1996). Degradation of zooxanthellae and regulation of their density in hermatypic corals. *Marine Ecology Progress Series* 139, 167–178. doi:10.3354/meps139167.

Torti, A., Lever, M. A., and Jørgensen, B. B. (2015). Origin, dynamics, and implications of extracellular DNA pools in marine sediments. *Mar. Genomics* 24 Pt 3, 185–196.

Trombetta, J. J., Gennert, D., Lu, D., Satija, R., Shalek, A. K., and Regev, A. (2014). Preparation of Single‑Cell RNA‑Seq Libraries for Next Generation Sequencing. *Curr. Protoc. Mol. Biol.* 107. doi:10.1002/0471142727.mb0422s107.

Vaidyanathan, H., Opoku-Ansah, J., Pastorino, S., Renganathan, H., Matter, M. L., and Ramos, J. W. (2007). ERK MAP kinase is targeted to RSK2 by the phosphoprotein PEA-15. *Proc. Natl. Acad. Sci. U. S. A.* 104, 19837–19842.

VanHook, A. M. (2019). Putting the squeeze on ERK signaling. *Sci. Signal.* 12. doi:10.1126/scisignal.aaw6857.

Waldbauer, J. R., Rodrigue, S., Coleman, M. L., and Chisholm, S. W. (2012). Transcriptome and proteome dynamics of a light-dark synchronized bacterial cell cycle. *PLoS One* 7, e43432.

Wang, X., Wang, Y., and Smith, W. O. (2011). The role of nitrogen on the growth and colony development of Phaeocystis globosa (Prymnesiophyceae). *Eur. J. Phycol.* 46, 305–314.

Weatherby, J. H. (1929). Excretion of Nitrogenous Substances in Protozoa. *Physiol. Zool.* 2, 375–394.

Wickham, H. (2010). ggplot2: elegant graphics for data analysis. *J. Stat. Softw.* Available at: https://www.researchgate.net/profile/Virgilio_Gomez_Rubio/publication/315785562_Book_review_ggplot2_-_Elegant_Graphics_for_Data_Analysis_2nd_Edition/links/58e4b7f8aca272d62977aaab/Book-review-ggplot2-Elegant-Graphics-for-Data-Analysis-2nd-Edition.pdf.

Worden, A. Z., Follows, M. J., Giovannoni, S. J., Wilken, S., Zimmerman, A. E., and Keeling, P. J. (2015). Environmental science. Rethinking the marine carbon cycle: factoring in the multifarious lifestyles of microbes. *Science* 347, 1257594.

Xiang, T., Lehnert, E., Jinkerson, R. E., Clowez, S., Kim, R. G., DeNofrio, J. C., et al. (2020). Symbiont population control by host-symbiont metabolic interaction in Symbiodiniaceae-cnidarian associations. *Nat. Commun.* 11, 108.

Yuasa, T., and Takahashi, O. (2016). Light and electron microscopic observations of the reproductive swarmer cells of nassellarian and spumellarian polycystines (Radiolaria). *Eur. J. Protistol.* 54, 19–32.

Zingone, A., Chrétiennot-Dinet, M.-J., Lange, M., and Medlin, L. (1999). MORPHOLOGICAL AND GENETIC CHARACTERIZATION OF PHAEOCYSTIS CORDATA AND P. JAHNII (PRYMNESIOPHYCEAE), TWO NEW SPECIES FROM THE MEDITERRANEAN SEA. *Journal of Phycology* 35, 1322–1337. doi:10.1046/j.1529-8817.1999.3561322.x.