UNIVERSITY OF CALIFORNIA, SAN DIEGO

Physiological, ultrastructure, and preliminary genomic characterization of the marine protist *Thraustochytrium* sp. LLF1b (class Labyrinthulomycetes)

A Thesis submitted in partial satisfaction of the requirements for the degree

Master of Science

in

Biology

by

Alexandra Rose Chun

Committee in charge:

Eric E. Allen, Chair
Mark Hildebrand
Roy Kaustuv

2017

The Thesis of Alexandra Rose Chun is approved and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____
Chair

University of California, San Diego

2017

TABLE OF CONTENTS

Signature Page……. iii

Table of Contents……. iv

List of Figures……. v

List of Tables……... vi

List of Supplemental Figures……..... vii

Acknowledgements…….. viii

Abstract of the Thesis…….. ix

Introduction…….. 1

Materials and Methods……. 7

Results and Discussion……..... 23

Appendix

Figures...……... 35

Tables...…… 42

Supplementary Figures...……. 43

Works Cited……. 48

LIST OF FIGURES

Figure 1. SEM of *Thraustochytrium* sp. LLF1b sample on a glass coverslip. The image shows the diversity in size for vegetative cells. …….. 35

Figure 2. SEM of *Thraustochytrium* sp. LLF1b sample on a glass coverslip. (A)Vegetative cell associated with an unknown substrate that is possibly forming an ectoplasmic net (the tendril-like projection that seems to be emanating from the cell membrane). (B) Vegetative cell……... 36

Figure 3. SEM of *Thraustochytrium* sp. LLF1b sample (A) on a glass coverslip showing the variability in cell surface topography and B) on a filter showing a bumpy cell surface texture and polyp-like protrusions clearly attached to the surface of the larger vegetative cell……... 37

Figure 4. SEM of *Thraustochytrium* sp. LLF1b sample on a glass cover slip. (A) and zoospore next to a vegetative cell. Zoospore appears oblong and has two flagella, one being a tinsel flagellum and the other a whiplash flagellum. (B) The tinsel flagellum with hairs appearing to be flush…….. 38

Figure 5. SEM of *Thraustochytrium* sp. LLF1b sample on a filter showing cells at different stages. A zoospore with hairs branching off the tinsel flagellum is seen in the center of the image. A young vegetative cell with a smooth surface can be seen directly to the right of the zoospore..... 39

Figure 6. Effect of supplemental carbon source on the growth of Thraustochytrid isolate LLF1b. Optical density (A) at 600 nm and wet weight of biomass (B) was used to measure cell growth over a period of five days with the initial inoculation at zero hours. PB = Plain Burja…..... 40

Figure 7. Growth of *C. cryptica* growth when co-cultured with Thraustochytrid isolate LLF1b. Viable cell counts were measured using a MUSE Cell Analyzer every few days for 3 weeks with the initial inoculation (equal to $4x10^4$ viable cells/ml) at zero hours. Cc, F/2 = *C. cryptica* in……..... 41

LIST OF TABLES

Table 1. Percent of total fatty acid (TFA) content for *Thraustochytrium* sp. LLF1b as determined through FAME analysis performed by Marco Allemann, University of California, San Diego, CA.…… 42

LIST OF SUPPLEMENTAL FIGURES

Figure 1S. A bird's eye view of the set up for the LLF1b and *C. cryptica* co culture experiment. Flasks were arranged so that corresponding replicates in each culture condition were exposed to the same light level. Circles inside the rectangle represent the flasks placed on top of…… 43

Figure 2S. Estimation of LLF1b genome size by fluorescent staining of cell nuclei and flow cytometry, performed by Nurcan Vardar, Hildebrand Lab, SIO. A) Standard curve of costained cells of known genome size. B) Determination of genome size based on standard curve…… 44

Figure 3S. Read length histogram of PacBio generated sequence data used as input for assembly of the LLF1b genome……..... 45

Figure 4S. Heterotrophic growth of *C. cryptica* in the dark with glucose or glycerol as a sole carbon source. Viable cell counts of *C. cryptica* cultures were measured using a MUSE Cell Analyzer once a week for two weeks then several weeks after. Gly = glycerol supplemented F/2…… 46

Figure 5S. Initial phylogenetic tree using full length 18S rRNA gene sequences of the locally isolated Labyrinthulomycetes from the Allen Lab. The tree includes reference 18S rRNA genes from *Thraustochytrium kinnei* and *Aplanochytrium kerguelense* along with that of and uncultured……. 47

ACKNOWLEDGEMENTS

I would like to thank Malcolm Wood from the Center for Cell Ultrastructure and Morphology at The Scripps Research Institute, Dr. Hildebrand, Olga Gaidarenko and Nurcan Vardar from the Hildebrand Lab at SIO, and Marco Allemann from the Allen Lab at SIO for their valued technical assistance.

I would also like to thank Dr. Allen, Marco Allemann, and Jessica Blanton, from the Allen Lab at SIO for their expertise, guidance, and encouragement during my time in the Allen Lab.

This work was supported by the National Science Foundation (MCB-1149552, PI Eric E. Allen).

ABSTRACT OF THE THESIS

Physiological, ultrastructure, and preliminary genomic characterization of the marine protist *Thraustochytrium* sp. LLF1b (class Labyrinthulomycetes)

by

Alexandra Rose Chun

Master of Science in Biology

University of California, San Diego, 2017

Professor Eric E. Allen, Chair

Single-celled marine protists belonging to the genus *Thraustochytrium* (class Labyrinthulomycetes) are of biotechnological interest due to their high production of polyunsaturated fatty acids (PUFAs). They produce particularly high amounts of omega-3 PUFAs making them an attractive source of omega-3 PUFAs as opposed to oil sourced from fish. Preliminary genomic information, laboratory growth on various carbon sources, effect on diatom (*Cyclotella cryptica*) growth in co-culture, and ultrastructural characteristics of the marine protist *Thraustochytrium* sp. LLF1b is presented in this work.

INTRODUCTION

Omega-3 fatty acids, a type of polyunsaturated fatty acid, are vital for human health and play key roles ranging from proper neural development to decreasing risk of cardiovascular disease (Gupta *et al*, 2012). Numerous studies suggest diverse health benefits of a diet high in omega-3 fatty acids, including decreasing mortality caused by coronary heart disease (Kris-Etherton *et al*, 2008) to decreasing inflammatory effects in rheumatoid arthritis (Ruxton *et al*, 2004). The human body cannot make sufficient amounts of omega-3 fatty acids on its own and must instead get them from dietary intake. The omega-3 fatty acid alpha linolenic acid (ALA; 18:3n-3) is an essential fatty acid that is a precursor for eicosapentaenoic acid (EPA; 20:5n-3) and docosahexaenoic acid (DHA; 22:6n-3), two omega-3 fatty acids that have been repeatedly implicated as being beneficial for human health (Ruxton *et al*, 2004). For example, increased intake of omega-3 fatty acids may improve cardiovascular health by increasing the amount of DHA and EPA in the cell membranes leading to positive membrane physicochemical changes and membrane protein signaling (Lee *et al*, 2009). Humans can endogenously convert ingested plant-derived ALA into EPA and then into DHA, however this conversion is an inefficient pathway and only a mere 5% of ingested ALA ends up as DHA in the human body (Greene *et al*, 2013). This inefficiency makes it necessary for humans to obtain the required amounts of omega-3 fatty acids directly from their diet. In light of the potential health benefits of a diet high in omega-3 fatty acids, health organizations around the world have recommended that people consume a minimum specified amount of oily fish, referring to fish with

high levels of DHA and EPA such as salmon, mackerel or herring, to increase omega-3 fatty acid intake. This would increase the amount of omega-3 fatty acids available for use by the body since humans cannot efficiently convert ALA into DHA and EPA. For example, in 2002 the American Heart Association recommended healthy individuals eat oily fish two or more times a week (Green *et al*, 2013). Recently, the Food Standards Agency in the UK increased its recommended amount of oily fish eaten from one to two servings per week to one to four servings per week (Ruxton *et al*, 2004).

Fish oil supplements have been used increasingly as an additional source of omega-3 fatty acid in the human diet. According to National Center for Complementary and Integrative Health (NCCIH) the percent of self-reporting adults in the U.S. that used fish oil supplements increased 3%, from 4.8% to 7.8%, between 2007 and 2012 (Clarke *et al*, 2015). According to the Fisheries and Aquaculture Department (FAO) 20.8 million tonnes (14.6%) of the total global fish harvest were used for production of fishmeal and fish oil in 2008 (compared to nearly 81% of the total global fish harvest used for human consumption that same year) (FAO, 2010). However, there has been increasing concern over the decrease in global fish populations since global fish catches have leveled off in the 1990s and started decreasing in 1995 (Green *et al*, 2013). The sustainability issues inherent to the current fishing trends has placed focus on the development of alternative sources of omega-3 fatty acids, including aquaculture, bioengineering of plant sources, and algal oils (Lee *et al*, 2009). Sourcing omega-3 fatty acids from non-fish sources also eliminates the

problem of ‘fishy’ odor as well as possible contaminants and pollutants, such as mercury, that inherently come with fish-derived oil supplements.

Microbially-produced oils are an attractive source of omega-3 fatty acids that are both sustainable and contain fewer pollutants compared to fish oil. Single-celled stramenopilian protists of the Class Labyrinthulomycete (SAR supergroup, Heterokonta superphylum) are ubiquitous in the marine environment and play important ecological roles in the degradation of detritus and the recycling of carbon and nutrients by virtue of their extensive repertoire of hydrolytic enzymes and subsequent predation by other organisms in the marine foodweb, e.g. mesozooplankton (Raghukumar and Damare, 2011; Singh *et al*, 2014). In addition to their ecological importance in aquatic ecosystems, Labyrinthulomycetes have attracted strong biotechnological interest due to their production of high quantities of omega-3 PUFAs in the form of triacylglycerides (TAGs), specifically DHA (Raghukumar, 2008). Broadly, the Labyrinthulomycetes consist of three orders, Labyrinthulida, Amphitremida, and Thraustochytrida, representing over 15 recognized genera (Pan *et al*, 2016). They have cells that range in size from 2-20 µm with cell walls containing sulfated polysaccharides, produce an ectoplasmic net (EN) (with exception of the thraustochytrid genus *Althonia)*, and produce heterokont biflagellate zoospores (Raghukumar, 2002, Raghukumar and Damare, 2011). The EN, an extension of the plasma membrane produced by an organelle on the cell membrane known as the bothrosome, or sagenogenetosome, seems to be significant in the ecological role of Labyrinthuolomycetes. The EN contains hydrolytic enzymes that breakdown organic

matter as well as attaches the cells to the substrate (Raghukumar, 2002). It has been reported that EN is not present when cells are grown in rich media but are produced when grown in nutrient poor conditions, further suggesting the role of the EN in nutrient absorption (Jain *et al*., 2007). Labyrinthulomycetes are even associated with specific substrates. For example, *Labyrinthula* species are commonly associated with sea grasses and fallen mangrove leaves. Labyrinthulomycetes have also been found in substantial numbers in the water column where they may use detritus or marine snow for growth (Raghukumar, 2002, Raghukumar and Damare, 2011). Indeed, Bochdansky *et al.* (2017) found that in terms of biomass Labyrinthulomycetes, along with fungi, surpassed all other groups (eukaryotic and prokaryotic) found on marine snow collected from the bathypelagic region of the ocean.

Thraustochytrids, which includes the genus *Thraustochytrium,* is one group within the Labyrinthulomycetes (Honda *et al*., 1999, and Raghukumar, 2002). Thraustochytrids consist of unicellular spherical or nearly spherical cells that generally produce EN from a single bothrosome and only reproduce through the production of biflagellate zoospores (Honda *et al*. 1999, and Raghukumar and Damare, 2011). Thraustochytrids seem to grow optimally at temperatures between 25 and 30℃ and within a pH range of 5 to 8. Additionally, thraustochytrids do not require light for growth (Raghukumar, 2008). Thraustochytrids have been found ubiquitously in marine environments and require Na^+ for growth, seeming to prefer salinity levels between 20 and 34% (Raghukumar, 2002). Some Thraustochytrids have been found to produce 30-50% of their total fatty acids as DHA predominantly in the form of TAG

lipid bodies but also as membrane-associated phospholipids (Armenta and Valentine, 2013; Raghukumar, 2008). The amount of omega-3 fatty acids produced varies greatly between species and culture conditions. For example, *Thraustochytrium* sp. ONC-T18 isolated by Burja *et al.* (2006) produced more than 4 grams per liter of DHA when grown in media supplemented with glucose; which is about five times the amount of DHA produced by another *Thraustochytrium* species, *T. aureum*, grown under similar conditions. Stored lipids may have several different roles in thraustochytrids. It was reported by Jain *et al.* (2007) that while many lipid bodies were observed in vegetative cells of a thraustochytrid isolate, fewer or no lipid bodies remained in vegetative cells that had changed into motile amoebae. Additionally, they observed that the amount of lipids in the cells decreased under starvation conditions. It was also observed that thraustochytrid cells with stained lipid bodies when moved from a nutrient rich to nutrient poor media produced fluorescent EN. This suggests that lipids stored inside thraustochytrid cells act as an energy source which is depleted as needed by the cell and as a source of fatty acids when the cells need to generate additional membrane material (Jain *et al*., 2007).

While there has been extensive research on Labyrinthulomycetes, some of their basic biology, including complete genomes, is not adequately understood (Raghukumar, 2008). Currently, only two full Labyrinthulomycete genomes have been published in the NCBI genome database thus far (Ji *et al*, 2015 and Liu *et al*, 2016). The first, published in 2015, is from a *Schizochytrium* species (Ji *et al*, 2015). The second, published in 2016, is from an *Aurantiochytrium* species (Liu *et al*, 2016).

Despite these genomes being sequenced and publicly available, no in-depth analysis of the genetic or metabolic potential of these genomes have been reported in the literature. Consequently, very little is known about the biology of the Labyrinthulids at the genetic level. Importantly, there is currently no representative genome available for *Thraustochytrium* species.

Thraustochytrium sp. LLF1b was isolated using a feather baiting technique with decaying algae (*Ulva* sp.) collected in close proximity to the Scripps Institution of Oceanography, La Jolla, CA, as the inoculation source (Shulse CN, unpublished). The "LLF1b" strain designation corresponds to its isolation procedure as follows: "LL" = Lettuce Leaf, "F" = feather bait, "1b" = isolate 1b. The goal of this thesis is to provide better understanding of the model Thraustochytrid strain *Thraustochytrium* sp. LLF1b in terms of its physiology, ultrastructure, and genomic characteristics. New insights into the biology of this representative *Thraustochytrium* strain will optimistically enhance this species' value as a candidate for application in large scale oil production or as a supplement in feed stock for aquaculture or agriculture. This will also increase what is known about the genus *Thraustochytrium* as a whole by providing preliminary genomic information on this underrepresented group.

MATERIALS AND METHODS

Cultivation

Continuous liquid cultures of LLF1b were maintained from 7.5.16 to 10.13.16 and then from 10.19.16 to 6.22.17. Subcultures were made every 4-10 days using inoculum in a 1:1000 dilution with the media and incubated at room temperature at 200 rpm. Burja liquid media was used; containing filtered seawater, 2 g/l peptone, 2 g/l yeast extract, 5 g/l glucose (based on the media used in Burja *et al*. 2006). 1% glycerol was used in the media recipe 7.12.16 onwards. The Burja base (seawater, peptone, and yeast extract) was autoclaved at 121 ℃ for 20 minutes prior to the addition of glucose and glycerol, autoclaved separately, and intermittent antibiotics, filter sterilized using a 0.22 μm filter. The antibiotics Penicillin and Streptomycin (final concentration of 300 μg/ml Penicillin, and final concentration 500 μg/ml Streptomycin) were added to the media continuously between 7.5.16 and 10.13.16 then sporadically from 10.19.16 to present day as needed if bacterial contamination of the cultures was suspected visually or via 16s PCR. Each reaction in the 16S rRNA gene PCRs contained 10 μl Hot Start Taq 2x Master Mix from New England BioLabs, 1 μl 785r (reverse primer with sequence 5'- GAC TAC HVG GGT ATC TAA TCC-3'), 1 μl 341f (forward primer with sequence 5'- CCT ACG GGN GGC WGC AG-3'), 7 μl ultra-pure PCR water, and 1 μl boil prep of sample (boil preps are 10 μl of 1xTE buffer and 10 μl liquid culture, heated to 95 ℃ for 5 minutes). The 16s PCRs are run to the program: 95 ℃ for 2 minutes, 29 cycles of (95 ℃ for 30 seconds, 55 ℃ for 30 seconds, 68 ℃ for 1 minute), 68 ℃ for 5 minutes, then held at 4 ℃. Control reactions

included *E. coli* MG1655 gDNA (positive control), LLF1b sample (negative control), and 1xTE buffer (negative control) in lieu of the 1 µl boil prep of samples. PCR products were run through a 1% agarose gel at 100V for 30-40 minutes until the loading dye was positioned halfway through the length of the gel. Glycerol stocks were routinely made with 0.5 ml 50% glycerol and 0.5 ml culture, inverted to mix, then stored at -80℃.

Scanning Electron Microscopy

A sample from a one-week old (late-exponential phase) culture of LLF1b was delivered to Dr. Malcolm R. Wood at the Center for Cell Ultrastructure and Morphology at The Scripps Research Institute who prepared samples of LLF1b for examination under the SEM. First, 2.27% glutaraldehyde in seawater was used to fix the cells overnight at 4 ℃. Fixed cells were then processed in one of two ways; 1) placed onto glass coverslips coated with polylysine or 2) filtered onto Whatman Nuclepore PC 13mm, 0.2 µm filters. Following filtration, a second filter of the same type was placed on top of the original filter with the fixed cells in between them. The stack of the two filters were then mounted between two nylon O-rings. Following these initial divergent steps all samples (on both the glass coverslips and filters) were washed for 30 minutes in ice-cold seawater. Next, all samples were treated for 2 hours in a solution of 1% osmium tetroxide in 0.1 m sodium cacodylate buffer then washed again for 1 hour in distilled water. The samples were then dehydrated in a graded ethanol series for critical point drying. Due to the size constraints of the critical point drier the samples were dehydrated in two batches with glass coverslips and filters

dehydrated separately. Following dehydration, the two filters in the O-ring mounted stacks were separated. Sample coated filters were kept while the second upper filters were discarded. The glass coverslips and the sample coated filters were then mounted onto stubs using carbon tape. To help increase electrical conductivity silver paint was applied around the circumference of the stubs. All samples were then kept overnight in a vacuum desiccator. Finally, a layer of iridium 7.4 nm thick was sputter coated onto all samples prior to examination under the SEM.

Lipid Characterization

For fatty acid methyl ester (FAME) analysis, LLF1b cell pellets were collected by centrifugation and placed in a Labconco Freeze Dry System for several hours to lyophilize the cell biomass. A portion of each cell pellet, approximately 1 cm by 1.5 cm in size (~10 mg), was broken off, crushed into a fine powder using a metal spatula, and placed into small vials. Fresh methanol/sulfuric acid solution (approximately 5-10% H_2SO_4 in methanol) was added to one third the volume of each vial. Whole cell methanolysis and transesterification of methyl ester derivatives was accomplished by capping vials and incubation in a heat block at 90°C for 90 minutes. The vials were then removed from the heat block and left to cool for several minutes before adding one third the volume of each vial with hexane. The remaining third of each vial was filled with 10% NaCl (in water). Vials were then inverted to mix and left for several minutes for the phases to settle. The top (organic) phase was drawn off in each vial and placed in separate secondary vials. Additional hexane was then added to the original vials, inverted, and left to separate for a second time. The top phase was

drawn off each a second time and added to their respective secondary vials. These secondary vials containing the hexane and fatty acid methyl esters were then stored at -80°C prior to analysis using gas chromatography-mass spectrometry as described previously (Allen *et al*., 1999).

Growth Characterization

A growth curve experiment was done to determine if LLF1b had a preferred supplemental carbon source when grown in a common base liquid media. Four culture conditions were used with supplemental carbon sources normalized by molarity: 1) Burja medium (2 g/l peptone, 2 g/l yeast extract in seawater) with no supplemental carbon source; 2) Burja supplemented with 27.75 mM glucose; 3) Burja supplemented with 27.75 mM glycerol; and 4) Burja supplemented with 27.75 mM sodium acetate. All culture conditions were represented by 20 ml cultures in triplicate inoculated with 100 µl (1:200) inoculum from the same 5-day starter culture. All cultures were incubated at room temperature (23°C) at 200 rpm. Every day following inoculation at approximately the same time the OD_{600} and wet weight of the samples were taken over a period of 5 days. To take the OD_{600} first a pipette was used to rinse the pellicle (biofilm) that develops on the walls of the flasks. If the pellicle was strongly adhered then the flasks were swirled to dislodge the pellicle and resuspended by pipetting. 3 ml of each culture was placed into small test tubes which were then vortexed vigorously for 5 seconds on high in an attempt to homogenize the samples. It should be noted that small particulates of various sizes still remained despite vortexing. The OD_{600} of any given sample was measured immediately after that sample was vortexed. When

cultures became too dense to measure without dilution they were diluted using plain Burja media. By the fourth and fifth days of the experiment all cultures were diluted 1:3 and 1:6 with Burja media, respectively (with exception of the cultures in Burja supplemented with sodium acetate, which yielded no measurable growth for the duration of the experiment). To measure wet weight the biomass contained in each sample was spun down in pre-weighed 2 ml centrifuge tubes at 13.2k rpm for 2 minutes. The supernatant was removed first by inverting the tubes over a beaker then by using a micropipette to carefully remove any remaining liquid. Centrifugation and decanting was done in two rounds per 3 ml culture. During this step on the second day of the experiment in one sample (Burja supplemented with 27.75 mM glycerol, flask 2) a significant amount of the pellet was lost. Consequently, this data point was omitted from the final wet weight growth curve and the average wet weight for the second day for this culture condition is representative of the two other cultures under the same culture conditions. Lastly, the weight of tubes containing cell pellets were measured and the weight of each corresponding tube when empty was subtracted to get the weight of each cell pellet. As noted above, the wet weight of all cultures in Burja supplemented with sodium acetate were not taken due to the complete absence of growth under these conditions.

Co-culture with the Diatom *Cyclotella cryptica*

Cyclotella cryptica and *Thalassiosira pseudonana* cultures were obtained from Olga Gaidarenko, Hildebrand Lab, SIO. Continuous cultures were kept in F/2 + 13 mg/l Na_2SiO_3 in filtered seawater (Guillard and Ryther, 1962). F/2 components and

Na_2SiO_3 stock were filter sterilized using a 0.22 μm filter prior to their addition to autoclaved filtered seawater. Cultures were incubated at room temperature (23℃) on a stationary glass plate over a light bar set to a 12-12 day-night cycle. Light levels read as 55 microeinsteins directly over the light source and 45 microeinsteins towards the edge of the glass plate farther away from the light source with a gradient of values between those two measured values. Subcultures were made every few days when the cultures were sufficiently dense for the concentration of cells in the subculture to equal $1x10^5$ viable cells/ml for *C. cryptica* and $1x10^6$ viable cells/ml for *T. pseudonana.* Viable cell counts were taken using a MUSE Cell Analyzer (Millipore Sigma, Billerica, MA) for cell count and viability. Unfortunately, *T. pseudonana* did not grow consistently under these laboratory conditions. Consequently, only *C. cryptica* was used for co-cultivation studies.

Prior to the co-culture experiment, a small experiment was run to check the heterotrophic capabilities of *C. cryptica* grown in the dark with media containing glycerol. For this experiment *C. cryptica* was grow in F/2 supplemented either with 1% glycerol or 1% glucose as the control. To make the media, autoclaved glycerol or filter sterilized glucose was added to prepared F/2 media described above. The inoculum for both conditions was taken from the same original *C. cryptica* culture so that the starting concentration was $1x10^5$ viable cells/ml. Inoculated flasks were then completely wrapped in aluminum foil to force heterotrophic growth and incubated without shaking at room temperature (23℃). Flasks were swirled vigorously prior to sampling for measurement of viable cell counts with the MUSE. Measurements were

only taken once a week for two weeks to limit light exposed to the cultures. One final measurement was taken about 5 weeks after inoculation. It should be noted that only one culture for each condition was used in this experiment.

The co-culture experiment included four distinct culture conditions: 1) *C. cryptica* in F/2; 2) LLf1b and *C. cryptica* in F/2; 3) LLF1b and *C. cryptica* in F/2 supplemented with 1% glycerol; and 4) LLF1b in F/2 (negative control). Three independent cultures were inoculated for each condition. All cultures containing LLF1b were inoculated from the same initial LLF1b culture in a 1:100 dilution to the total volume of the new culture. All cultures containing *C. cryptica* were inoculated from the same initial *C. cryptica* culture to a final concentration of $4x10^4$ viable cells/ml. All cultures were placed on the glass plate/light bar set up with replicates lined up perpendicularly to the light bar so that the light levels would be consistent between replicates of different culture conditions (Figure 1S). Viable cell counts were measured using the MUSE every 3 days for 3 weeks. Cultures were vortexed vigorously prior to sampling and immediately before measuring with the MUSE. Cultures that contained LLF1b were examined under a light microscope prior to measuring. Samples that had dense LLF1b growth were diluted 1:2 with sterile F/2 media prior to examination and measurement of viable cell counts. Samples that contained large aggregates of cells were not measured to avoid potentially clogging the MUSE. This happened at the 165-hour time point with one *C. cryptica* and LLF1b culture grown in F/2 media supplemented with glycerol, consequently the average

viable cells/ml for this culture condition and time point is only representative of two independent cultures rather than three.

Phylogenetic Placement and Preliminary Genomic Characterization

Prior to sequencing of the LLF1b genome, an estimate of the genome size was determined (Figure 2S). A 50 ml culture of LLF1b was harvested in mid-exponential phase. The cell pellet was resuspended in 5 ml methanol and stored at -20°C overnight. Cells were then washed twice with 1x PBS and stained with the fluorescent DNA binding dye SYBR Green. Stained samples were run on the BD Influx cell sorter to estimate genome size relative to a set reference genomes used to establish a standard curve, including *Escherichia coli* (4.65 Mbp), *Saccharomyces cerevisiae* (12.29 Mbp), and *Thalassiosira pseudonana* (34.5 Mbp). Genome size estimate were performed by Nurcan Vardar, Hildebrand Lab, SIO.

To determine phylogenetic placement of LLF1b and other locally collected Labyrinthulomycetes isolates, the 18S rRNA gene of axenic cultures were cloned and sequenced. First, phenol chloroform extractions were performed on cell pellets from 16 Labyrinthulomycetes isolates archived in the Allen Lab culture collection. This collection included two Schizochytrium strains, *Schizochytrium* sp. CNK019 and S31, *Aurantiochytrium limacinum,* and thirteen locally isolated Labyrinthulomycetes. Since the isolates grow to different densities in liquid culture the amount of biomass used in each extraction was normalized according to pellet size so that genomic DNA was extracted from approximately the same amount of biomass per strain. For sample preparation and cell lysis, cell pellets were resuspended in Buffer A (50 mM Tris, 50

mM EDTA, pH 7.5). Proteinase K and SDS was added to the Buffer A cell suspension to final concentrations of approximately 330 µl/ml Proteinase K and 3% SDS. Tubes were inverted to mix and incubated in a 55°C water bath then inverted to mix every 10 minutes for 1 hour. The samples were cooled briefly on ice before one volume of phenol:chloroform:isoamyl alcohol (PCI; 25/24/1, v/v) equal to the volume of the sample was added to each tube and inverted to mix until the solution was homogenized. Tubes were then centrifuged for 5 minutes at 10,000 rpm. From these original tubes the top aqueous phase was drawn off and ejected into secondary tubes containing an equal volume plus a little excess of chloroform each. The secondary tubes containing PCI/chloroform was inverted to mix and then centrifuged for 2 minutes at 10,000 rpm. From these secondary tubes the top phase was drawn off and ejected into tertiary tubes. The genomic DNA was precipitated with the addition of 1/10 volume 3M sodium acetate and one volume of cold isopropanol roughly equal to the volume of the sample. Tertiary tubes were inverted to mix until the wispy white threads of genomic DNA visibly precipitated. The tertiary tubes were then centrifuged for 2 minutes at 13.2k rpm, and the supernatant was removed. A wash step was done in which 70% ethanol was added to each of the tertiary tubes and centrifuged at 13,200 rpm for 4 minutes to firmly secure the pellets to the side of the tubes. The supernatant was carefully removed using a micropipette and the genomic DNA pellets were allowed to air dry for 8 minutes for the ethanol to evaporate, being careful to not let the pellets dry out completely. 100 µl of 1x TE buffer was added to the tertiary tubes containing the rinsed genomic DNA pellets and were then placed in the

refrigerator overnight for the genomic DNA pellets to resuspend. Following the overnight resuspension RNase was added to each sample (final concentration 100 µg/ml) and heated to 55℃ for 30 minutes. Following the RNase treatment, the genomic DNA was precipitated a second time as described previously however a volume of 100% ethanol equal to two times the volume of the sample was used instead of isopropanol. Samples were chilled on ice for 20-30 minutes and then centrifuged for 10 minutes at 1200 rpm at 4℃. The supernatant was removed and the wash step repeated (with centrifugation changed to 1200 rpm at 4℃ for 5 minutes). Finally, the genomic DNA pellets were resuspended overnight in 100 µl of 1xTE buffer and stored at -20℃.

18S rRNA gene PCR was performed using the extracted genomic DNA from the 16 Labyrinthulomycete isolates in our culture collection. Each PCR tube contained the following: 20 µl Q5 Hot Start 2x Master Mix High Fidelity DNA Polymerase from New England BioLabs, 2 µl 10µM EukA (forward primer: 5'-AAC CTG GTT GAT CCT GCC AGT-3'), 2 µl 10 µM Euk B (reverse primer: 5'-TGA TCC TTC TGC AGG TTC ACC TAC-3'), 14 µl ultra-pure PCR water, and 2 µl genomic DNA (approximately 10 - 50 ng). The thermal cycle program was as follows: 98℃ for 30 seconds, 30 cycles of [98℃ for 10 seconds, 72℃ for 30 seconds, 72℃ for 1 minute], 72℃ for 2 minutes, and held at 4℃. Several PCRs had to be performed using various dilutions of DNA for some isolates. Controls reactions performed in every round of PCR included *E. coli* MG1655 genomic DNA (negative control), 1x TE buffer (negative control), and fish gut genomic DNA (positive control). PCR products were

confirmed by running 5 µl of the 18S PCR products (diluted with the same volume loading dye) through a 1% agarose gel at 90 Volts for 1 hour or until the loading dye bands are half way through the length of the gel. PCR products were purified using a ZYMO kit according to manufacturer directions, however since PCR products were made as 40 µl samples, 200 µl of DNA Binding Buffer was used.

TSS chemically competent *E. coli* DH5α cells were made to use for cloning the 18S rRNA gene sequences from all 16 Labyrinthulid isolates. LB broth was inoculated with stock DH5α culture in 1:100 dilution and incubated at 30℃ at 200 rpm. After 3 hours and 40 minutes (OD_{600} = 0.37) 1 ml portions were aliquoted to 2 ml centrifuge tubes and centrifuged for 1 minute at 13,200 rpm to pellet the cells. Supernatant was removed and cell pellets were chilled on ice for 20 minutes. 100 µl of 1xTSS buffer was added to each tube and pellets were gently resuspended. The newly made TSS DH5α competent cells were stored at -80℃ until needed.

Cloning of the 18S rRNA gene for each of the 16 isolates was done in separate small batches. All steps for cloning, unless otherwise stated, were the same in each batch. First, 5 µl of purified PCR product (18S rRNA gene amplicon) was combined with 5 µl of 2x Taq master mix from New England BioLabs and incubated at 68℃ for 15 minutes for addition of A overhangs to amplicons. The reagents for the vector reaction were kept on ice until use. Each vector reaction contained 7.5 µl sterile water, 1 µl salt solution, 0.5 µl PCR 4 TOPO vector, and 1 µl prepared PCR product. Two control reactions were included with 1) vector only reaction (1 µl sterile water added in lieu of insert) and 2) insert only reaction (0.5 µl sterile water added in lieu of

vector). Assembled reactions were left at room temperature for 30 minutes. Following incubation, 2 µl of the vector reaction was added to 100 µl of TSS DH5α competent cells and pipetted gently to mix. The cell/vector transformation mixture was then placed on ice for 20 minutes. Samples were then heat shocked in a 42℃ water bath for 30 seconds then immediately placed back on ice for 2 minutes. At this point, 250 µl of SOC medium was added to each transformation reaction and samples were shaken at 30℃, 200 rpm for 1 hour. Following recovery and outgrowth, transformed cells were spun down for 1 minute at 13.2k rpm and the supernatant was removed. Cells were resuspended in 100 µl of SOC per sample and plated onto LB-Kanamycin plates coated in X-gal/IPTG. All plates contained 50 µl/ml Kanamycin. Plates were incubated at 30℃ for 24 hours then, due to extremely small colony sizes, plates were incubated for an additional 24 hours at room temperature. Following incubation, four transformant colonies per isolate were picked using sterile toothpicks and streaked onto LB+Kan grid plates and incubated at room temperature for 2 days. Cell lysates were made concurrently with the grid plates by swirling the toothpicks in 50 µl 1xTE buffer in 96 well plates immediately after a given toothpick was used to streak the LB+Kan grid plate. Cell lysates were then heated to 95 ℃ for 5 minutes. PCR of the cell lysates were used to further screen for successful transformants. Each reaction contained 10 µl Hot Start Taq 2x Master Mix from New England BioLabs, 1 µl EukA, 1 µl EukB, 7 µl ultra-pure PCR water, and 1 µl clone lysate. Controls used were *E. coli* MG1655 gDNA, 1x TE buffer, and fish gut genomic DNA added in lieu of clone lysate. The following program was used: 95℃ for 1 minute, 29 cycles of (95℃ for 30

seconds, 54℃ for 40 seconds, 68℃ for 2 minutes), 68℃ for 5 minutes, hold at 4℃. Successful clones were identified as having a single band (1.8 kb in length) when illuminated under UV after being run through a 1% agarose gel at 100V for 30-40 minutes. Following screening, one successful clone per isolate was sampled from the LB+Kan grid plates and T-streaked onto LB+Kan plates and incubated at room temperature for 2-5 days. For the majority of the cloned isolates, the T-streaks were then used to inoculate overnight cultures using 10 ml of LB and shaken at 30℃, 200 rpm. Plasmid minipreps were made the next day following inoculations using the ZYMO Research Plasmid miniprep kit according to manufacturer directions. Success of the minipreps was checked via Qubit. A plasmid sample for each of the 16 isolates was then submitted for Sanger sequencing using four internal primers to determine full length 18S rRNA gene sequences: FA2 (5'-GTC TGG TGC CAG CAG CCG CG-3', from Mo *et al*. 2002), FA2Rev (5'-CGC GGC TGC TGG CAC CAG AC-3'), LabyA (5'-GGG ATC GAA GAT GAT TAG-3', from Stokes *et al*. 2002), and LabyARev (5'-CTA ATC ATC TTC GAT CCC-3', from Collado-Mercado *et al*. 2010). Raw sequence data were assembled and validated using Sequencher DNA Sequence Analysis Software. A phylogenetic tree was constructed in Geneious using Muscle alignment and the FastTree algorithm. Reference 18S rRNA genes from *Thraustochytrium kinnei* (gi│345105398│gb│HQ228964.1│) and *Aplanochytrium kerguelense* (gi│5509882│dbj│AB022103.1│) were used along with that of an uncultured Heterokont clone (gi│95115859│gb│DQ504337.1│) as the outgroup.

In attempts to sequence the genome of LLF1b high molecular weight genomic DNA was extracted from LLF1b via phenol chloroform extraction according to the previously outlined protocol (up to the point at which the gDNA was precipitated) except for the following changes: First a pellet from a dense 250 ml LLF1b culture was resuspended in 10 ml of Buffer A, from which 9 ml was used for the extraction. The final concentration of Proteinase K added to the Buffer A/cell mixture was increased approximately 1.75x. All steps involving inversion to mix was done extremely gently. All centrifugation steps were done at 4000 rpm for 25 minutes. All steps involving pipetting of genomic DNA was done using wide bore pipette tips. To precipitate the genomic DNA after extraction with phenol and chloroform sodium acetate and cold isopropanol was added by slowly pouring down the side of the tertiary tube. This tube was slowly inverted until wispy threads formed and the sample had a jelly like appearance. Here, the precipitated DNA was spooled onto two glass pipettes that had been sealed at the tip, rather than centrifuged to pellet the genomic DNA. The spooled genomic DNA was allowed to air dry for 5-10 minutes before being placed in 1x TE buffer for 10 minutes. The spooled genomic DNA was then gently shaken off the glass pipettes. Samples were placed in the refrigerator for 3 days to resuspend followed by RNase treatment with a fresh stock of RNase. Treated samples were then transferred to secondary tubes containing equal volumes of chloroform, gently mixed by inversion and centrifuged at 4,000 rpm for 25 minutes. The top phase in each tube was drawn off and ejected into tertiary falcon tubes in which the genomic DNA was precipitated and spooled as described above. The

spooled genomic DNA was then washed in 70% ethanol and dried as previously. The spooled genomic DNA samples were then resuspended and refrigerated in Buffer A overnight. LLF1b genomic DNA samples were then checked for bacterial contaminants via PCR using bacterial 16S rRNA gene primers 341F and 785R as outlined previously. Control reactions contained *E. coli* MG1655 DNA (positive control), *A. limacinum* DNA (negative control), and Buffer A buffer (negative control). Prior to sequencing, spooled LLF1b genomic DNA was precipitated, centrifuged at 10,000 for 15 minutes to pellet, washed, and resuspended again in a small volume of 1xTE buffer. To resuspend the genomic DNA, which did not resuspend after 3 days at 4℃, samples were gently heated in a 40℃ water bath and gently agitated until fully resuspended. Relative purity of the LLF1b genomic DNA was evaluated using the Nanodrop and showed 260/280 = 1.92 with no 230 nm contamination. To estimate the length of LLF1b genomic DNA, a sample was run through a 6% agarose gel at 100 V for 80 minutes, until the loading dye was at the bottom of the length of the gel. When visualized half of the ladder had run off the gel; however, the genomic DNA presented a bright smear above the 10kb band of the ladder, indicating that the spooled LLF1b genomic DNA sample was of high molecular weight and thus suitable for PacBio library construction. The concentration of DNA in the spooled LLF1b sample was quantified via Qubit and found to be 188 ng/µl. An aliquot of the spooled LLF1b gDNA sample was then submitted to the Institute for Genomic Medicine DNA sequencing facility (School of Medicine, UC San Diego) for library construction (20 kb Blue Pippin size selected library) and single

molecule DNA sequencing using the Pacific Biosciences (PacBio) RS Single Molecule Real-Time (SMRT) sequencing platform. SMRT sequencing is a third generation sequencing technology that allows long-read sequencing. Sequencing of LLF1b genomic data was performed on two separate SMRT cells employing the P5/C3 chemistry. Post-filter subread statistics for each SMRT cell were as follows: SMRT 1 produced 57,934 reads totaling 229,169,746 bp (mean length=3,955 kb; N50=6,403 bp); SMRT 2 produced 77,731 reads totaling 368,785,108 bp (mean length=4,744 bp; N50=7,497 bp). Figure 3S shows the read length histogram of the combined data set that served as input for assembly (91,913 total reads; 409,491,897 total bases).

A preliminary draft genome of LLF1b was assembled using the Canu single-molecule sequence assembly package (Koren *et al*., 2017). Assembly was performed using default parameters with an assigned genome size of 25 Mbp and correctedErrorRate stringency parameter of 0.035 on the Allen Lab's *teal* workstation (data directory: 'LLf1b_combined_0.035_teal_110916'). Assembly statistics were assessed using the Quast quality assessment tool (Gurevich *et al*., 2013). *Ab initio* gene prediction was performed using the Maker annotation pipeline (Cantarel *et al*., 2008).

RESULTS AND DISCUSSION

Scanning Electron Microscopy

SEM was used to examine the ultrastructure characteristics of LLF1b. The size of vegetative cells of LLF1b appears to greatly vary (Figure 1). While all vegetative cells visible in Figure 1 appear to be spherical they range in size from 8 μm in diameter at the largest cell to approximately 2 μm in diameter at the smallest clearly visible cell. Like other *Thraustochytrium* species, LLF1b may produce ectoplasmic nets. These extensions of the cells' plasma membranes have been previously described as rhizoid in appearance (Iwata *et al*., 2017). In Figure 2A a vegetative cell appears to have such a projection starting to emanate from its membrane and is possibly beginning to produce an ectoplasmic net. However, no extensive ectoplasmic nets were observed in LLF1b samples. This may be because the samples were cultivated in a rich media and thraustochytrids usually do not form ectoplasmic nets in rich liquid media (Raghukumar, 2002). The ectoplasmic net is produced by the bothrosome within an invagination of the plasma membrane that is connected to the endoplasmic reticulum inside the cell (Iwata *et al*. 2017). A possible bothrosome may be seen in Figure 2B as a circularly shaped pit in the surface of the cell.

The samples examined under SEM showed varied cell surface topography (Figure 3A). One or more large depressions were seen in the surface of many cells. It is unlikely these large depressions represent bothrosomes as multiple such features can be seen on a single cell. Thraustochytrids reportedly produce their ectoplasmic nets from a single bothrosome (Iwata *et al*., 2017). The surface of the cells was generally

smooth or bumpy, with small protrusions covering the cell surface and with or without larger irregularly shaped protrusions from the cell surface that resemble polyps or membrane vesicles. Several of these vesicle-like protrusions are clearly seen attached to the surface of the cell in Figure 3B, along the left edge of the vegetative cell in the center of the image. Small bumpy protrusions can also be clearly seen in Figure 2B providing some insight into the cell wall structure. Similar observations were made by Gareth Jones and Harrison (1974) who noted that the surface of young vegetative *Thraustochytrium striatum* cells was completely smooth while that of older cells appeared to be blistered due to old cell wall material beginning to break off the cell. While these blisters on *T. striatum* cells do not exactly match the bumpy texture seen on LLF1b in Figure 3A and 3B they are very similar in appearance. As such, this bumpy texture may also suggest increased age of a given vegetative cell in LLF1b.

Thraustochytrium reproduce by producing biflagellate zoospores. Numerous zoospores were observed in the LLF1b cultures (Figure 4). The body of the zoospores appeared to be oblong, approximately 3.2 µm x 2.3 µm at the longest and widest points, with a tinsel flagellum coated with hairs and a shorter, naked whiplash flagellum (Figure 4A). The hairs on the tinsel flagellum can be seen flush against the shaft of the flagellum (Figure 4B). These biflagellate zoospores and corresponding flagella have been previously reported in *Thraustochytrium.* These zoospores are formed inside the mature cell (a zoosporangia) after which they are released through a partial or complete breakup of the zoosporangia cell wall. Following zoospore release a proliferation body remains which may develop into another vegetative cell and

subsequent zoosporangia (Bongiorni *et al*., 2005). In *Schizochytrium aggregatum* zoospore settlement has been observed in five stages. First, the zoospore attaches to a surface, after which the cell body starts to become more spherical. The two flagella cease movement and retract one at a time back into the cell. At this point the cell body is round and does not have flagella or an ectoplasmic net. Lastly, the cells produce ectoplasmic nets (Iwata *et al*., 2017). Figure 5 shows several examples of LLF1b cells in various stages of this cycle. First, a zoospore can be seen in the center of the image. This zoospore is approximately 2 µm wide (length of zoospore cannot be measured accurately as it appears to be positioned at an angle) and has the two characteristic flagella described previously. However, the tinsel flagellum of this zoospore has hairs branching off the shaft of the tinsel flagellum, rather than flush against the shaft as seen in Figure 4. A young vegetative cell with a smooth surface can be seen directly to the right of the zoospore. The surface of the young vegetative cell and the zoospore are very similar in terms of their smooth texture. This young vegetative cell may have settled recently as it is round and lacks both flagella and an ectoplasmic net (lack of ectoplasmic net may also be due to the rich culture conditions as described above). Two presumably older vegetative cells can be seen along the left edge of the image with the small bumpy protrusions covering the cell surface as described previously in relation to Figure 2. Numerous small spherical bodies 1 µm or smaller in diameter can be seen on the filters (Figure 3B and Figure 5) but are not present on the glass cover slips. This may possibly be an artifact from the filtering done on the samples or it may be that these small spherical bodies were present in all samples, both glass coverslips

and filters, but were washed off of the exposed glass coverslips while they remained secured in between the stacks of filters during sample preparation. Some of these bodies have web-like tendrils radiating from them. Some of these bodies may be proliferation bodies or simply cellular debris from zoosporangia following zoospore release.

Lipid Characterization

FAME analysis of LLF1b indicates that it produces a large portion of its total fatty acid content as omega-3 polyunsaturated fatty acids (Table 1). LLF1b was found to produce the omega-3 polyunsaturated fatty acids EPA and DHA. While EPA only made up 4.80% of total fatty acids, DHA comprised 32.14% of total fatty acids. LLF1b also produced omega-6 polyunsaturated fatty acids, which made up 6.01% of total fatty acids. Altogether, polyunsaturated fatty acids (both omega-3 and omega-6) comprise 42.95% of total fatty acids in LLF1b. This high DHA production is comparable to that of other reported *Thraustochytrium* isolates and according to Armenta and Valentine (2013) it is not uncommon for thraustochytrids to produce widely unequal amounts of DHA and EPA, when both are produced, as is seen in LLF1b. This data underscores the potential of LLF1b and *Thraustochytrium* in general as an attractive source of single-cell omega-3 polyunsaturated fatty acids.

Growth Characterization

When grown in Burja media supplemented with different additional carbon sources, the *Thraustochytrid* isolate LLF1b shows little preference for the type of

supplemental carbon source added to the media (Figure 6). Burja, glycerol supplemented Burja and glucose supplemented Burja culture conditions produced cell densities that reached an optical density (600 nm) around 0.8 after about 100 hours of growth (Figure 6A). This was the maximum level of growth for the plain Burja and Glucose supplemented cultures. The plain Burja and glucose supplemented Burja cultures showed decreased cell densities from 100 hours to about 120 hours of growth. The only significant difference between the three growth conditions is that growth in the glycerol supplemented media was still increasing at 120 hours, albeit at a decreasing rate, resulting in a final average OD_{600} of 0.94 at 120 hours compared to that of the plain Burja and glucose supplemented Burja, 0.67 and 0.62, respectively (as determined by one-way ANOVA weighted-means analysis, $P<0.01$). Wet weight also shows similar trends between the plain Burja, glucose supplemented Burja, and glycerol supplemented Burja with all three cultures increasing in mg biomass per ml of culture for the duration of the experiment and ending with average wet weights of 11.67 mg/ml, 16.20 mg/ml, and 17.93 mg/ml, respectively (Figure 6B). The only significant difference was found between plain Burja and glycerol supplemented Burja with $P<0.05$. This may be due to the fact that the base media used, Burja, is already a fairly rich media and permits LLF1b growth in all culture conditions regardless of supplemental carbon source. However, sodium acetate supplemented media showed no growth with all three independent cultures measuring an OD_{600} of zero for the entire duration of the experiment. It is predicted that pH changes upon addition of

sodium acetate resulted in unfavorable growth conditions. Future experiments will require adequate buffering of the media to maintain circumneutral pH.

Previous studies have reported that *Schizochytrium limacinum* (a Labyrinthulid within the same family as LLF1b) grows well and produces high amounts of DHA in media with glucose, fructose, or glycerol as carbon sources (Raghukumar, 2008). Additionally, *Thraustochytrid* sp. ONC-T18 isolated by Burja *et al.* (2006) produced increasing amounts of biomass (g/l) when grown in a base media (10 g/l yeast extract, and 6 g/l sea salt) supplemented with increasing glucose concentrations up to a certain point. In light of these prior results we wanted to determine how LLF1b would respond to variable carbon sources. LLF1b growth in plain Burja was expected to be significantly lower than that of the cultures cultivated in Burja media supplemented with either 27.75 mM glucose or 27.75mM glycerol. However, as seen in Figure 6A and 6B, LLF1b growth in plain Burja and Burja supplemented with glycerol or with glucose is markedly similar with exception of the significant differences found between glycerol supplemented Burja and plain (OD_{600} and wet weights) or glucose supplemented burja (OD_{600} only) at 120 hours of growth. The base media (plain Burja) used in this experiment is already a fairly rich media and may be permitting the growth of LLF1b regardless of the presence of supplemental carbon sources in the media. It should be noted that since LLF1b is a different species than *Thraustochytrium* sp. ONC-T18, LLF1b may simply have different preferred carbon sources, which may have enabled it to grow equally well in the glucose or glycerol supplemented and unsupplemented Burja media. To more thoroughly examine the

effect of available carbon sources on LLF1b growth the growth experiment can be repeated using a minimal media as the base. This base media could have little to no yeast extract or peptone so that LLF1b growth (or lack thereof) must rely on the supplemental carbon source for extended periods of growth. This will help to clarify to what extent the LLF1b growth seen in this experiment (Figure 6) under all the culture conditions is due to utilization of the yeast extract/peptone in the base media instead of the supplemental carbon sources. This would also clarify if LLF1b does indeed grow better with glycerol as a supplemental carbon source, as these data may hint. This would be valuable to know as growth on low cost substrates, such as glycerol, is important to reduce cost of large scale production.

Co-culture with Diatom *Cyclotella cryptica*

In this co-culture experiment we wanted to investigate whether or not the presence of LLF1b in culture together with *C. cryptica* would enhance growth of a photosynthetic partner strain. Prior to the co-culture experiment a preliminary experiment was done to investigate whether *C. cryptica* would grow in the dark heterotrophically on glycerol (Figure 4S). The concentration of viable *C. cryptica* cells/ml in glycerol supplemented F/2 media did not exceed $1.37x10^5$ cells/ml for the duration of the experiment, while growth in the glucose supplemented F/2 media control reached a peak of $1.36x10^6$ viable cells/ml after 14 days. This indicated that *C. cryptica* is not likely to use glycerol heterotrophically for growth in the dark. Here, we assumed that *C. cryptica* would not use glycerol heterotrophically when exposed to

light as well. As such, glycerol was added to one of the co-culture conditions (*C. cryptica* and LLF1b in glycerol supplemented F/2 media). The reasoning was that if LLF1b relied on waste products from or dead or dying diatoms in co-culture then LLF1b would need a supplemental carbon source for initial growth while *C. cryptica* growth was in the lag phase.

In the co-culture experiment growth of *C. cryptica* is adversely affected when co-cultured with LLF1b (Figure 7). When cultured alone in F/2 media *C. cryptica* displays a typical sigmoidal growth pattern with a lag phase for the first 165 hours of the experiment followed by a growth phase between 165 and 380 hours that started to plateau from 380 hours to the end of the experiment. The maximum concentration of viable cells/ml for this culture condition was $4.54x10^5$ cells/ml at the end of the experiment (553 hours). *C. cryptica* shows no active growth when co-cultured with LLF1b in F/2 media supplemented with glycerol. For this culture condition the average concentration of viable cells/ml was $4.08x10^4$ cells/ml towards the beginning of the experiment around 45 hours, after which the concentration of viable cells consistently decreased, until it became undetectable by the MUSE after 450 hours (undiluted sample contained less than 10 particles/µl). *C. cryptica* growth when co-cultured with LLF1b in F/2 media was highly variable between the three independent cultures. The concentration of viable cells/ml increased slightly from $4.00x10^4$ cells/ml at initial inoculation (zero hours) to $5.63x10^4$ cells/ml around 240 hours. While viable cell concentration decreased after 240 hours, there was a sudden spike at 553 hours reaching $7.40x10^4$ cells/ml. However, as this data point has a standard deviation of +/-

$10.10x10^4$ cells/ml it is likely inaccurate. *C. cryptica* growth when cultured alone in F/2 media was found to be statistically different from growth when co-cultured with LLF1b in glycerol supplemented F/2 media at 305 hours with $P<0.05$. *C. cryptica* growth from 380 hours onward in both co-culture conditions was statistically different (as determined by one-way ANOVA weighted-means analysis with $P<0.01$). There was no significant difference found between *C. cryptica* growth when co-cultured with LLF1b in F/2 media or in glycerol supplemented F/2 media at any time in the experiment. There was no *C. cryptica* growth detected in the LLF1b in F/2 control as expected and the zero values for this culture condition were not included in statistical analysis. While concentration of LLF1b in the co-culture conditions was not measured quantitatively during this experiment it is worthwhile to note that there was discernible LLF1b growth in the glycerol supplemented F/2 media and to a lesser extent the plain F/2 media from 45 hours onwards. There was no discernible LLF1b growth when LLF1b was cultured alone in F/2 media.

Since Labyrinthulomycetes have been suggested to play a role in nutrient cycling (Raghukumar and Damare, 2011) growth of *C. cryptica* when co-cultured with LLF1b was expected to be higher than that when cultured alone. It was hypothesized that LLF1b would degrade waste products produced by the diatoms, or even dead or dying diatoms themselves, thereby keeping co-cultures a cleaner and more suitable environment for phototrophic growth. However, as seen in Figure 7, this was not the case, and growth of *C. cryptica* when cultured alone was significantly higher than growth in the two co-culture conditions. Since LLF1b did not seem to grow in F/2

media alone but did grow when co-cultured with *C. cryptica* it may be possible that LLF1b could have been utilizing *C. cryptica* cells or exudates as substrates for growth in both co-culture conditions, thereby inhibiting *C. cryptica* growth. Further investigation is required to elucidate the ecological role and potential applications of Labyrinthulomycetes in polyculture.

Phylogenetic Placement and Preliminary Genome Characterization

To determine phylogenetic placement a small initial phylogenetic tree was constructed using the full length 18S rRNA gene sequences of the locally isolated Labyrinthulomycetes from the Allen Lab (Figure 5S). All local isolates grouped with *Thraustochytrium kinnei* which suggests that these local isolates, including LLF1b, are of the genus *Thraustochytrium.* This preliminary observation was further investigated by aligning our Labyrinthulomycetes 18S rRNA gene sequences with those of the Labyrinthulomycetes reference tree described in Pan *et al.* (2016). This comprehensive Labyrinthulomycetes reference tree consists of 1,181 18S rRNA gene sequences and reveals nearly all groups identified in this class. When aligned with this reference tree our Labyrinthulomycetes 18S rRNA gene sequences grouped squarely within the *Thraustochytrium kinnei* group. This analysis confirmed our initial results suggesting LLF1b is a member of the genus *Thraustochytrium.*

Based on an experimental genome size estimation experiment, LLF1b has an estimated genome size of approximately 25 Mb (Figure 2S). PacBio sequencing of high-molecular weight genomic DNA generated a total of 409.5 Mbp of sequence data

(91,913 reads), equating to approximately 16.4-fold genome coverage. Canu assembly of the PacBio sequence data generated a total of 598 contigs with a total length of 26,456,353 bp (largest contig=504,556 bp; N50=57,727 bp; percent G+C=51.43%). The data generated as part of this thesis resulted in a modest 16.4-fold coverage of the LLF1b genome, thus not sufficient to allow robust assembly which typically mandates at least 40-fold coverage using PacBio data alone. Despite these considerations, preliminary *ab initio* gene prediction produced 14,537 gene models, a number consistent with genomes of comparable size. While compelling to consider further analyses with the data available, the decision was made to not pursue additional genome characterization experiments until additional sequence data is generated to improve fidelity and confidence in the analysis of the draft LLF1b genome. Future efforts will prioritize generating additional PacBio sequence data to allow improved assembly however this work is beyond the scope of this thesis. Once generated, this data will allow accurate assessment of the genetic and metabolic potential of *Thraustochytrium* sp. LLF1b, analysis of horizontal gene transfer events, and comparative genomic analysis against other Stramenopile genomes.

Summary

This thesis has provided new insights into the model Labyrinthulomycete *Thraustochytrium* sp. LLF1b in terms of its physiology, ultrastructure, and genomic characteristics. The results presented here, along with continued genomic investigations in the Allen Laboratory, have revealed fundamental new information

about members of the *Thraustochytrium* genus in general while specifically enhancing LLF1b's candidacy for applications in omega-3 PUFA single-cell oil production and potential as a feed supplement for agriculture or aquaculture.

APPENDIX

Figures

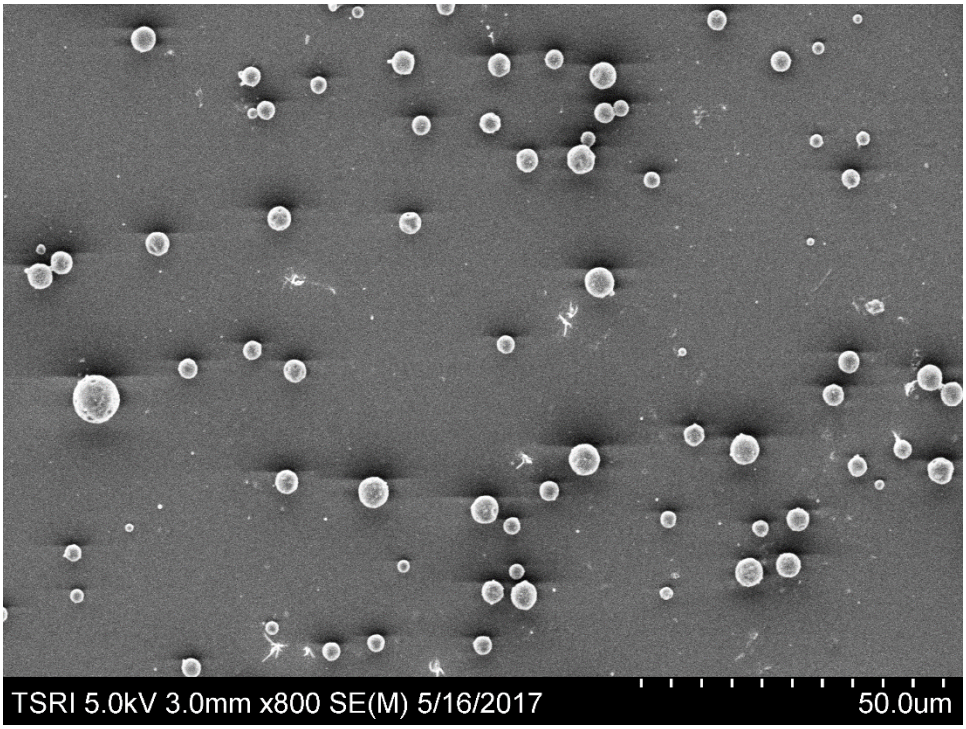


Figure 1. SEM of *Thraustochytrium* sp. LLF1b sample on a glass coverslip. The image shows the diversity in size of vegetative cells.

A.

B.

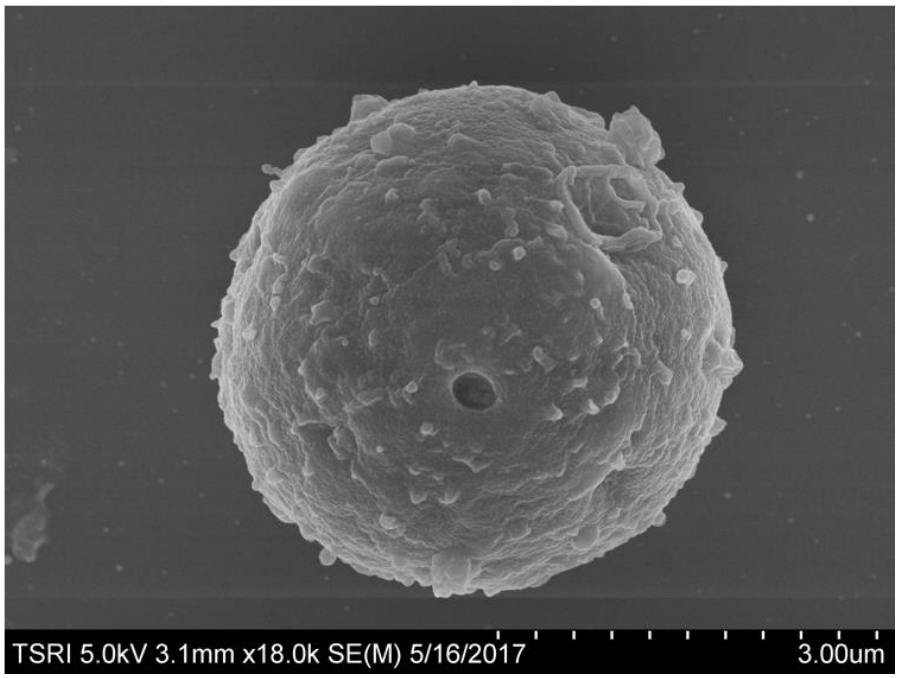


Figure 2. SEM of *Thraustochytrium* sp. LLF1b sample on a glass coverslip. (A)Vegetative cell associated with an unknown substrate that is possibly forming an ectoplasmic net (the tendril-like projection that seems to be emanating from the cell membrane). (B) Vegetative cell with possible bothrosome, the organelle responsible for ectoplasmic net production within an invagination of the plasma membrane, located towards the center of the image.

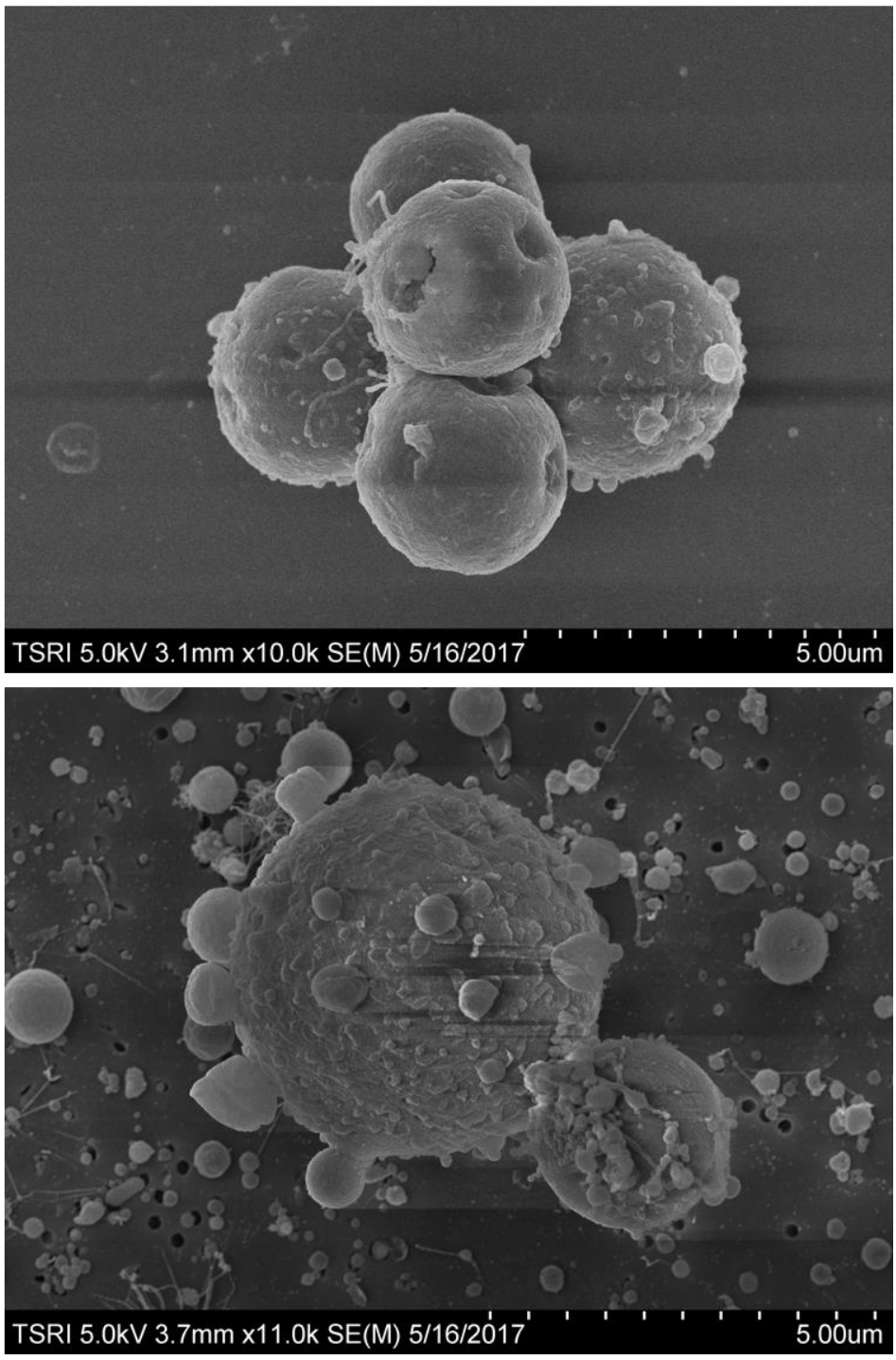


Figure 3. SEM of *Thraustochytrium* sp. LLF1b sample (A) on a glass coverslip showing the variability in cell surface topography and B) on a filter showing a bumpy cell surface texture and polyp-like protrusions clearly attached to the surface of the larger vegetative cell.

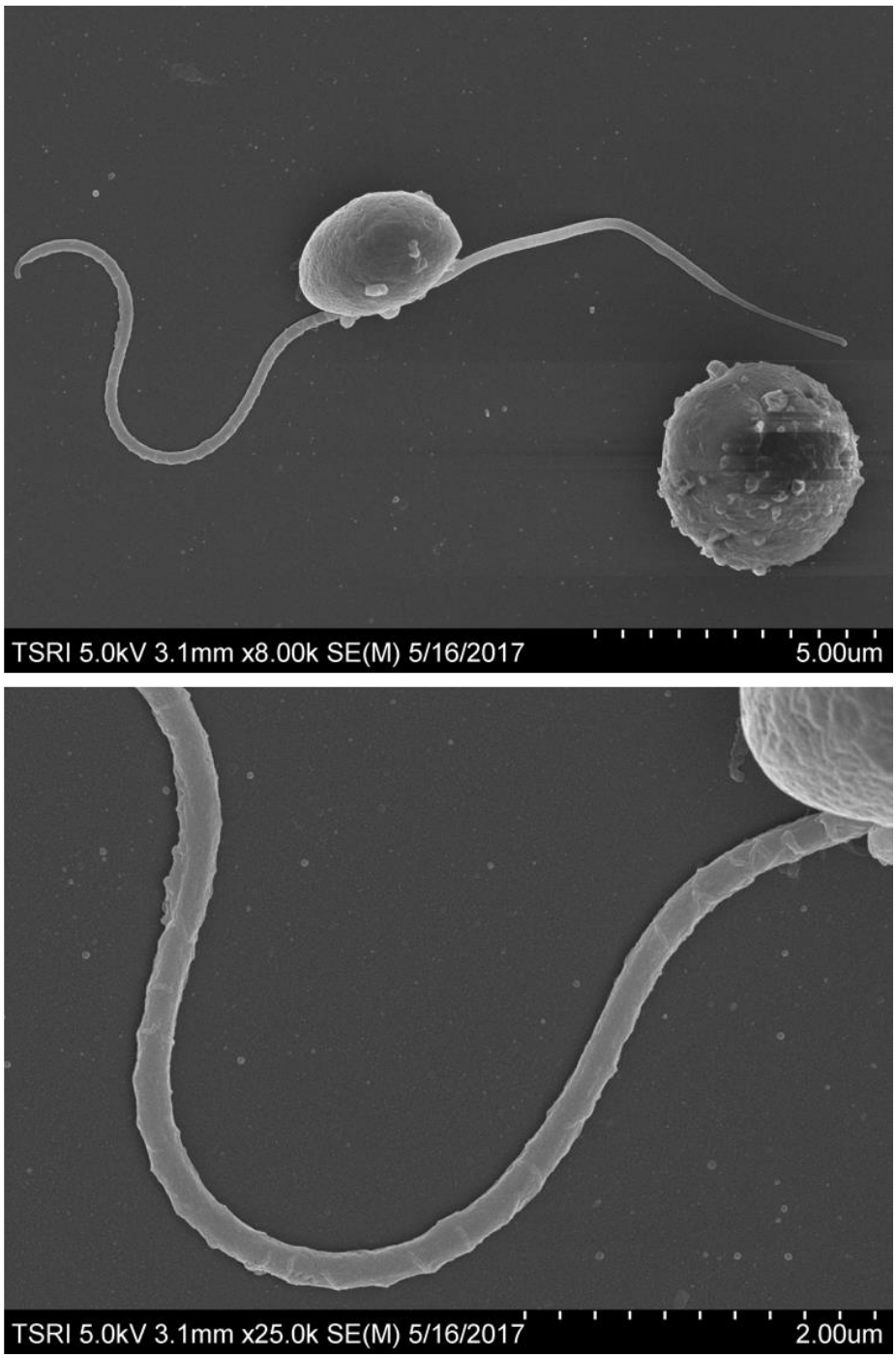


Figure 4. SEM of *Thraustochytrium* sp. LLF1b sample on a glass cover slip. (A) and zoospore next to a vegetative cell. Zoospore appears oblong and has two flagella, one being a tinsel flagellum and the other a whiplash flagellum. (B) The tinsel flagellum with hairs appearing to be flush against the shaft of the flagellum.

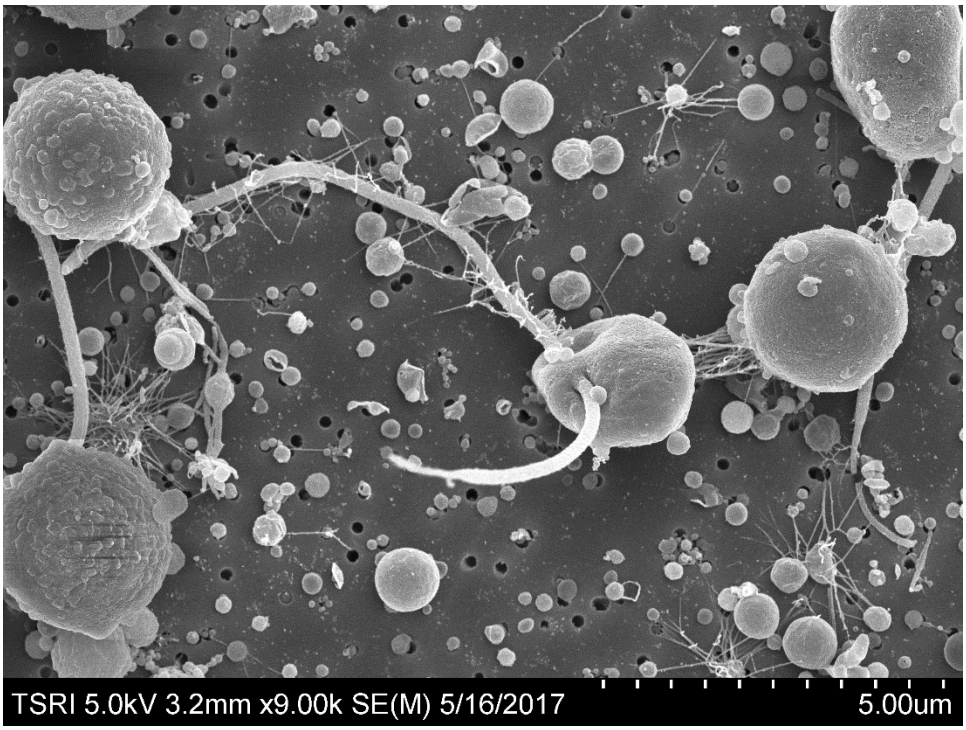


Figure 5. SEM of *Thraustochytrium* sp. LLF1b sample on a filter showing cells at different stages. A zoospore with hairs branching off the tinsel flagellum is seen in the center of the image. A young vegetative cell with a smooth surface can be seen directly to the right of the zoospore. Two bumpy and, presumably, older vegetative cells can be seen along the left edge of the image.

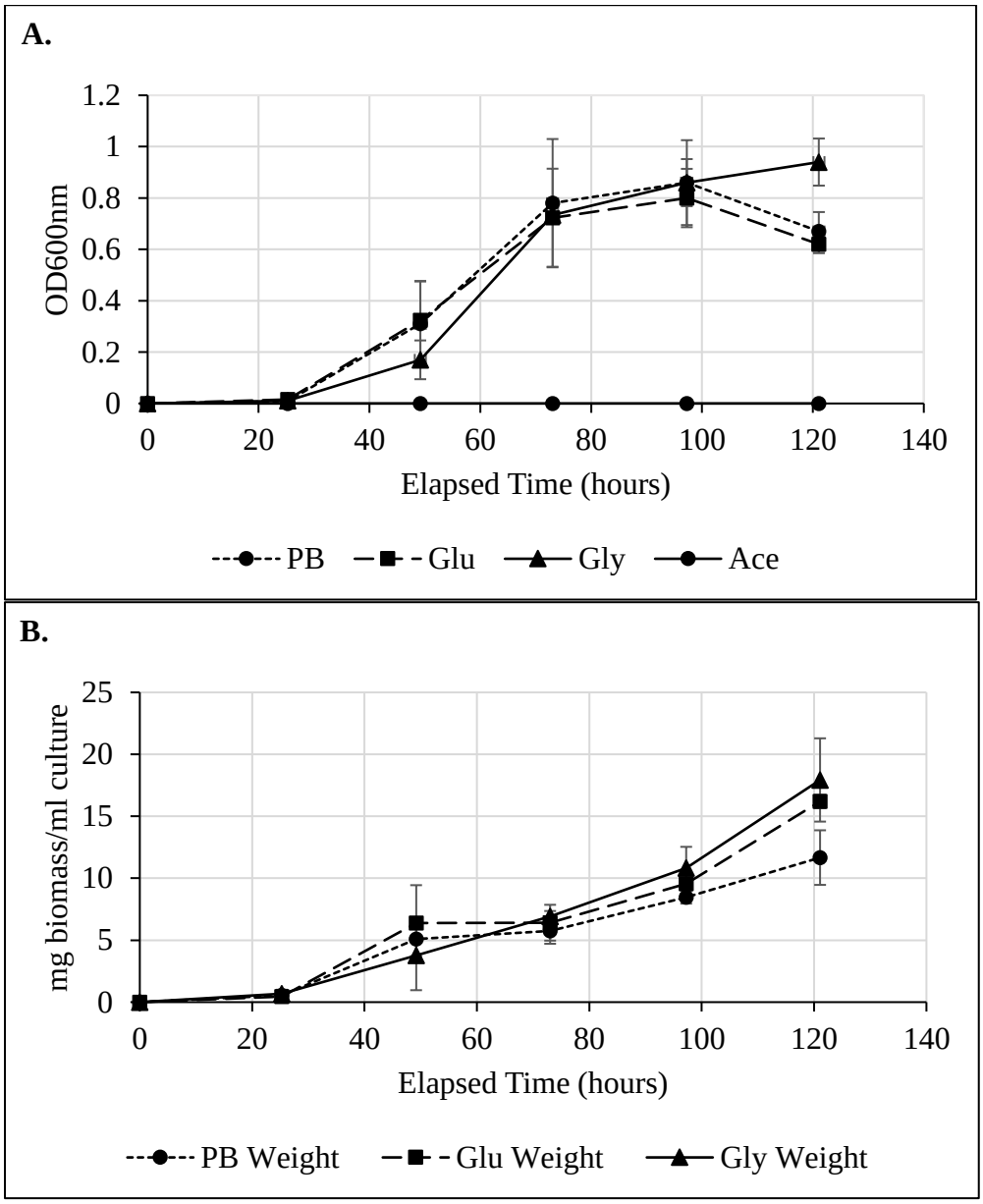


Figure 6. Effect of supplemental carbon source on the growth of Thraustochytrid isolate LLF1b. Optical density (A) at 600 nm and wet weight of biomass (B) was used to measure cell growth over a period of five days with the initial inoculation at zero hours. PB = Plain Burja media with no supplemental carbon source, Glu = glucose supplemented Burja media, Gly = glycerol supplemented Burja media, and Ace = sodium acetate supplemented media. Supplemental carbon source normalized by molarity. Data points are an average of three independent cultures, unless otherwise stated (see materials and methods). Error bars represent standard deviations of triplicate cultures.

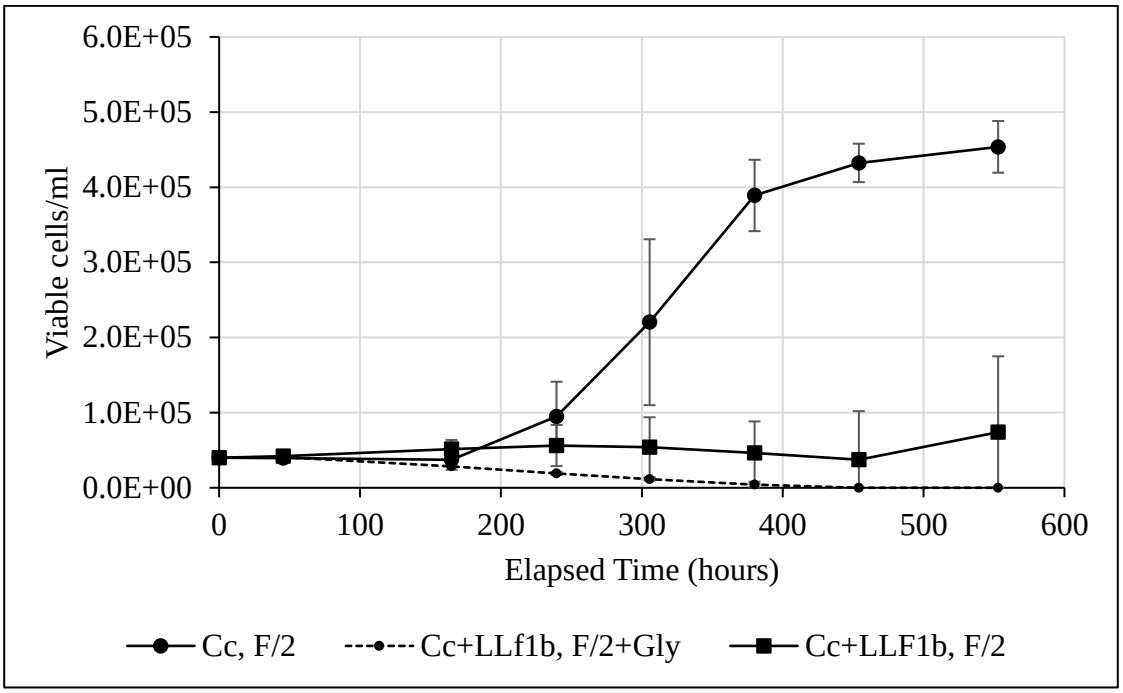


Figure 7. Growth of *C. cryptica* growth when co-cultured with Thraustochytrid isolate LLF1b. Viable cell counts were measured using a MUSE Cell Analyzer every few days for 3 weeks with the initial inoculation (equal to $4x10^4$ viable cells/ml) at zero hours. Cc, F/2 = *C. cryptica* in plain F/2 media, Cc+LLF1b, F/2+Gly = *C. cryptica* and LLF1b in F/2 media supplemented with glycerol, Cc+LLF1b, and F/2 = *C. cryptica* and LLF1b in F/2 media. Data points are an average of three independent cultures. Error bars represent the standard deviation.

Tables

Table 1. Percent of total fatty acid (TFA) content for *Thraustochytrium* sp. LLF1b as determined through FAME analysis performed by Marco Allemann, University of California, San Diego, CA.

Fatty Acid	%TFA
12:0	0.78663
14:0	6.15306
15:0	3.19317
16:0	36.6675
16:1n9	0.54513
17:0	0.5146
18:0	1.94986
18:1n9 (9)	7.24177
18:2n6 (9,12)	0.96592
20:4n6 (5,8,11,14)	2.47233
20:5n3 (5,8,11,14,17)†	4.79569
22:5n6 (4,7,10,13,16)	2.57001
22:6n3 (4,7,10,13,16,19)*	32.1443

†Eicosapentaenoic acid (EPA)

*Docosahexaenoic acid (DHA)

Supplemental Figures

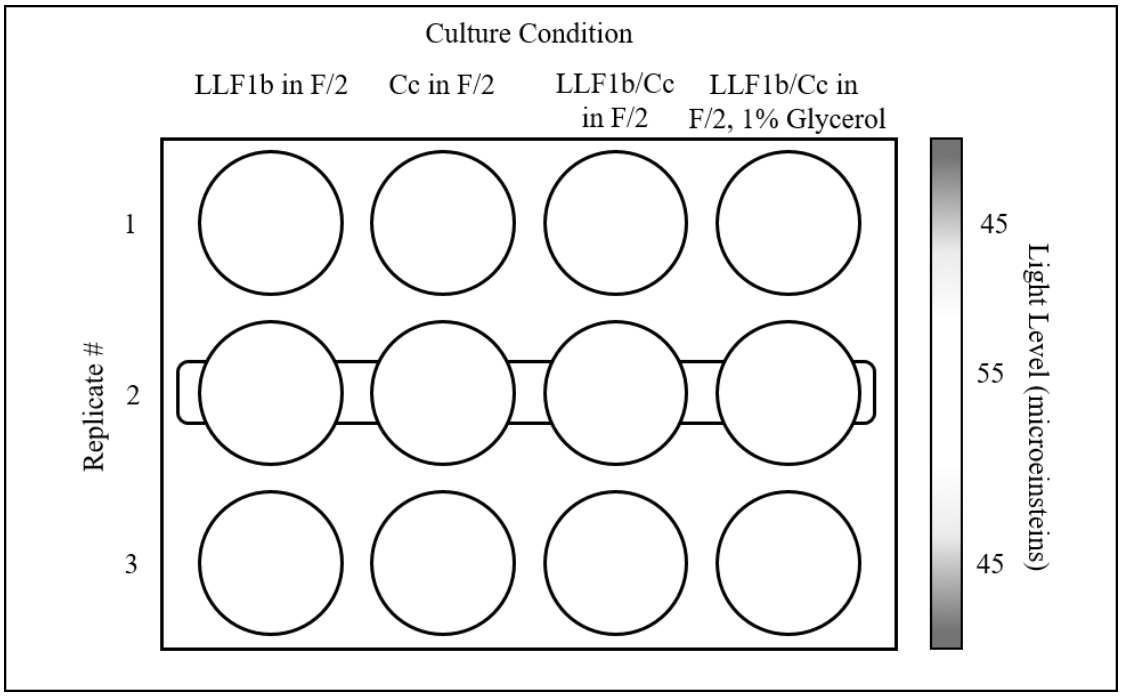
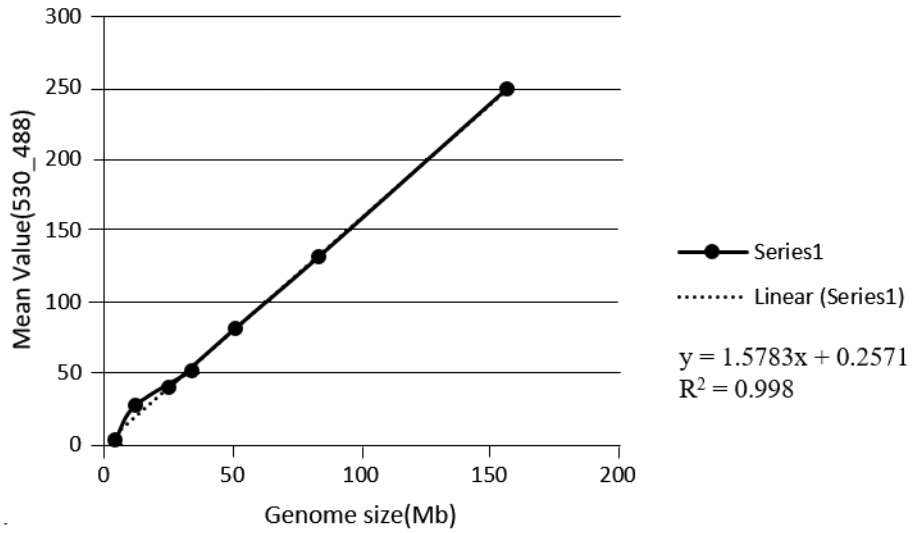


Figure 1S. A bird's eye view of the set up for the LLF1b and *C. cryptica* co-culture experiment. Flasks were arranged so that corresponding replicates in each culture condition were exposed to the same light level. Circles inside the rectangle represent the flasks placed on top of the glass plate. The rounded rectangle in the center (behind the circles) represents the light bar underneath the glass plate. The gradient bar is meant to illustrate the gradient of light levels from 55 microeinsteins at the center of the plate to 45 microeinsteins towards the edge of the plate but does not hold any quantitative meaning.

A.



B.

	Genome size (Mb)	Mean Value (530_488)
E-coli	4.6	3.65
Yeast	12.29	27.18
Tp	34.5	51.70
LLF1B	25.63	40.22
Schizochytrium sp. (ATCC20888)	102.43	161.93

Figure 2S. Estimation of LLF1b genome size by fluorescent staining of cell nuclei and flow cytometry, performed by Nurcan Vardar, Hildebrand Lab, SIO. A) Standard curve of costained cells of known genome size. B) Determination of genome size based on standard curve florescence values.

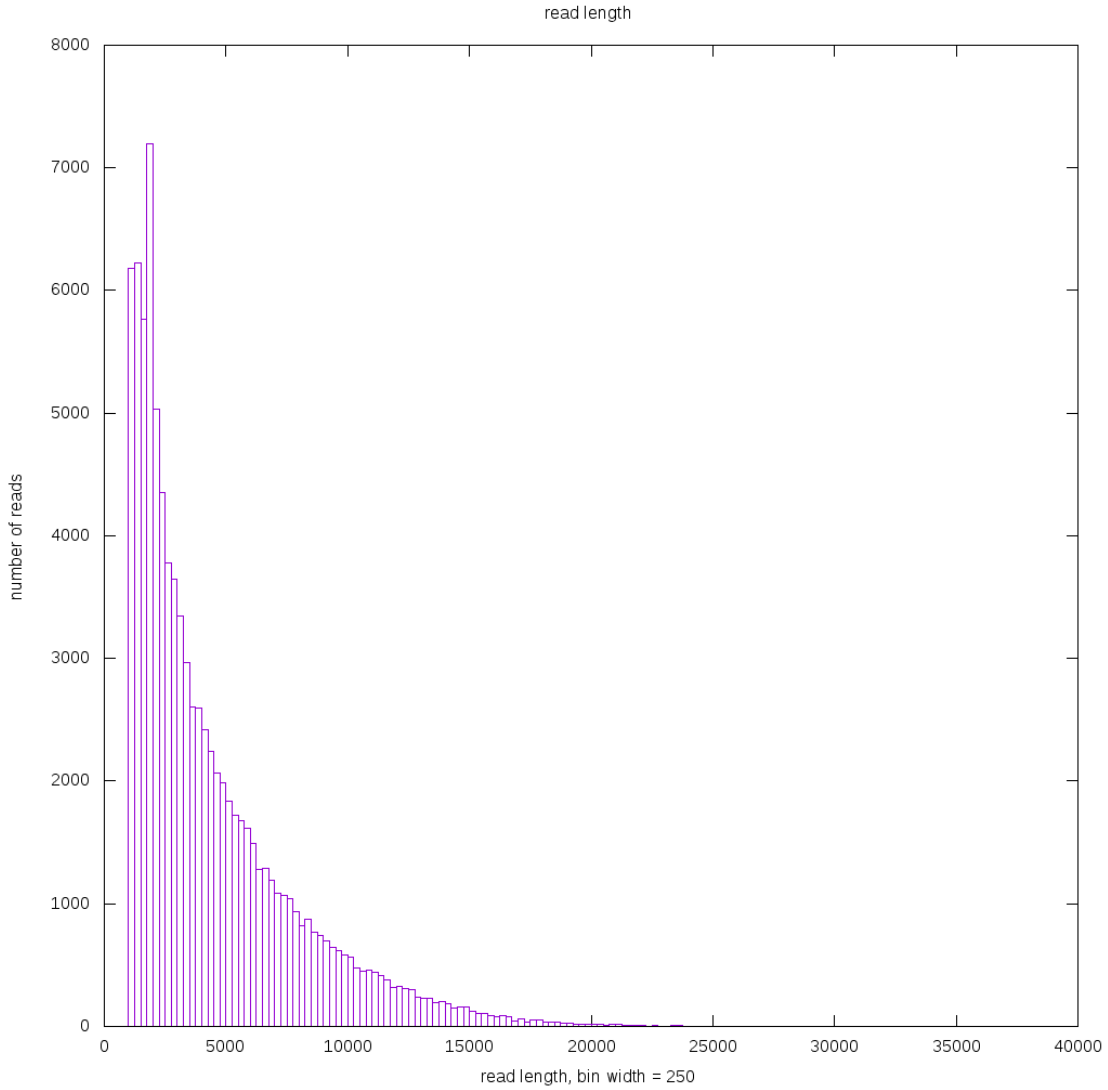


Figure 3S. Read length histogram of PacBio generated sequence data used as input for assembly of the LLF1b genome.

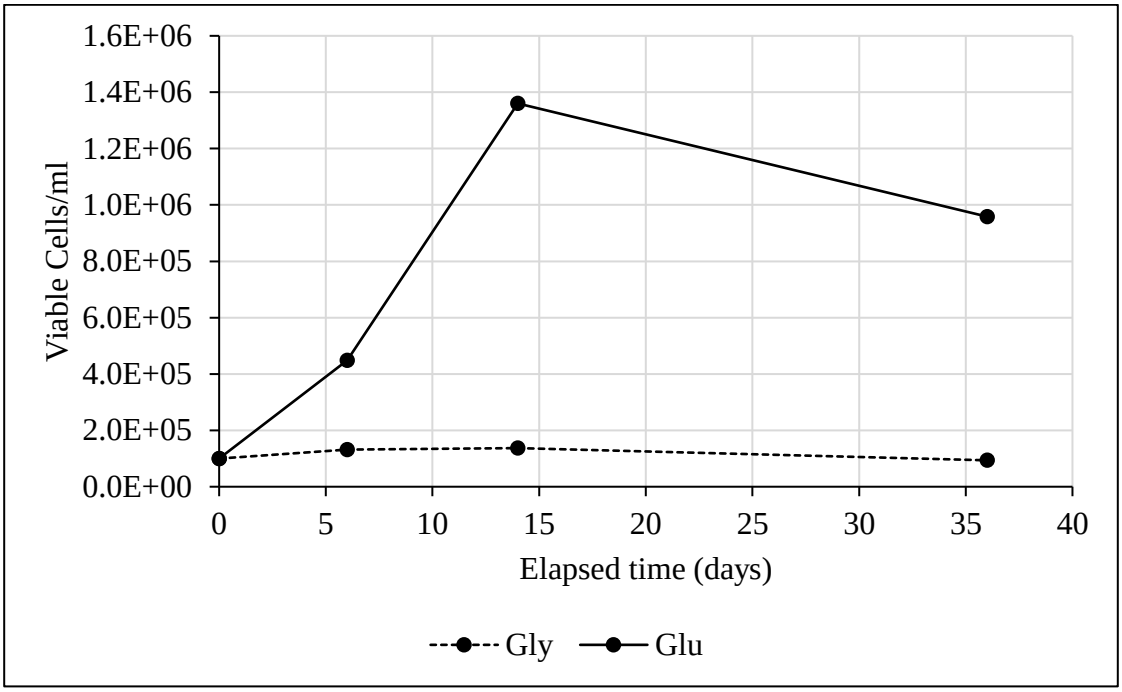


Figure 4S. Heterotrophic growth of *C. cryptica* in the dark with glucose or glycerol as a sole carbon source. Viable cell counts of *C. cryptica* cultures were measured using a MUSE Cell Analyzer once a week for two weeks then several weeks after. Gly = glycerol supplemented F/2 media, and Glu = glucose supplemented F/2 media. Each data point is representative of a single culture.

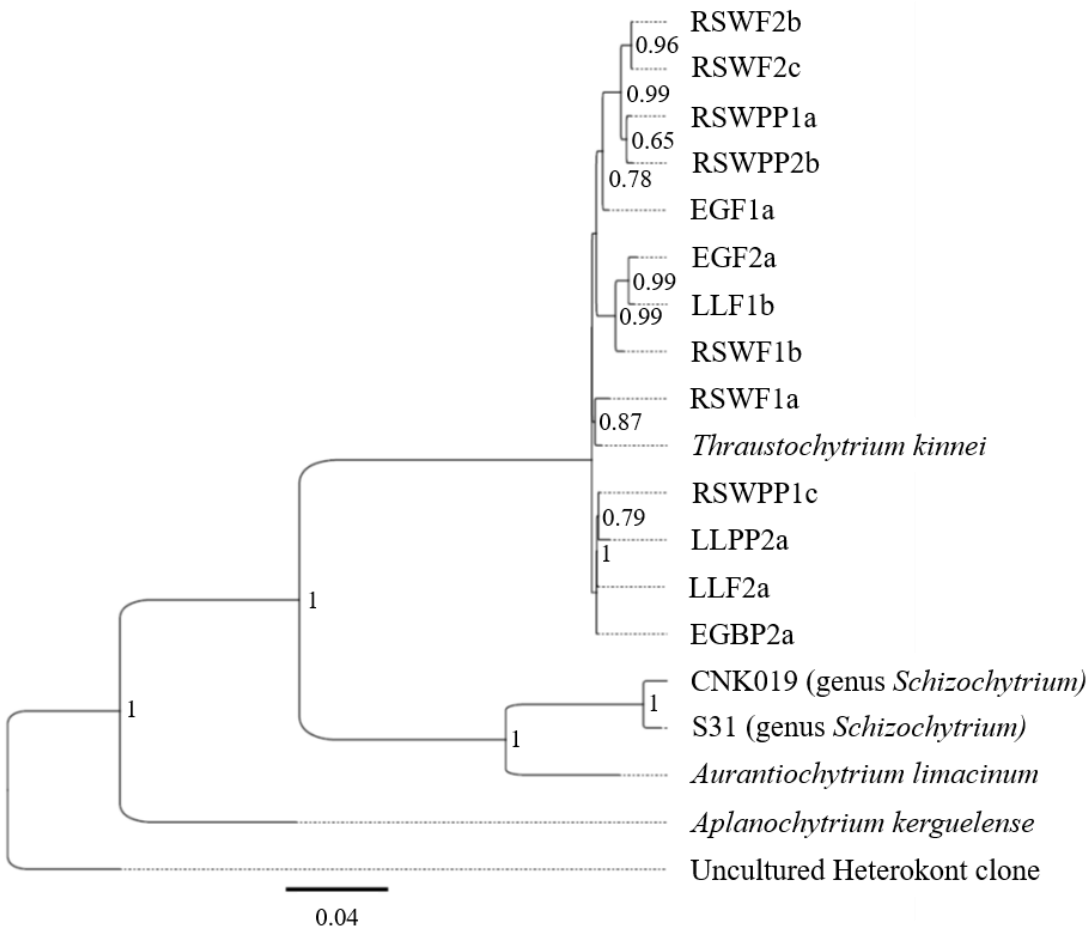


Figure 5S. Initial phylogenetic tree using full length 18S rRNA gene sequences of the locally isolated Labyrinthulomycetes from the Allen Lab. The tree includes reference 18S rRNA genes from *Thraustochytrium kinnei* and *Aplanochytrium kerguelense* along with that of and uncultured Heterokont clone as the outgroup.

WORKS CITED

Allen, E.E., Facciotti, D., Bartlett, D. H. “Monounsaturated but not polyunsaturated fatty acids are required for growth of the deep-sea bacterium Photobacterium profundum SS9 at high pressure and low temperature.” *Applied and Environmental Microbiology,* 65.4 (1999): 1710-1720.

Armenta, R.E., Valentine, M.C. “Single-Cell Oils as a Source of Omega-3 Fatty Acids: An Overview of Recent Advances.” *J Am Oil Chem Soc.* 90 (2013): 167-182.

Bochdansky, A.B., Clouse, M.A., Herndl, G.J. “Eukaryotic microbes, principally fungi and labyrinthulomycetes, dominate biomass on bathypelagic marine snow.” *The ISME Journal*. 11 (2017): 362-373.

Bongiorni, L., Jain, R., Raghukumar, S., Aggarwal, R.K. “*Thraustochytrium gaertnerium* sp. nov.: a New Thraustochytrid Stramenopilan Protist from Mangroves of Goa, India.” *Protist.* 156 (2005): 303-315.

Burja, A. M., Radianingtyas, H., Windust, A., Barrow, C.J. "Isolation and characterization of polyunsaturated fatty acid producing Thraustochytrium species: screening of strains and optimization of omega-3 production" *Appl Microbiol Biotechnol* 72 (2006): 1161-1169

Cantarel, B.L., Korf, I., Robb, S.M.C., Parra, G., Ross, E., Moore, B., Holt, C., Sanchez Alvarado, A., Yanell, M. “MAKER: an easy-to-use annotation pipeline designed for emerging model genomes.” *Genome Res* 18 (2008): 188-196.

Clarke, T.C., Black, L.I., Stussman, B.J., Barnes, P.M., Nahin, RL. “Trends in the use of complementary health approaches among adults: United States, 2002–2012”. *National health statistics reports*; no 79. Hyattsville, MD: National Center for Health Statistics. 2015.

Collado-Mercado, E., Radway, J.C., Collier, J.L. “Novel uncultivated labyrinthulomycetes revealed by 18S rDNA sequences from seawater and sediment samples.” *Aquat Microb Ecol.* 58 (2010): 215-228.

FAO. “World Review of Fisheries and Aquaculture; Part 1.” *The State of World Fisheries and Aquaculture 2010.* 2010

Gareth Jones, E.B., Harrison, J.L. “Zoospore Discharge in *Thraustochytrium striatum*.” *Trans. Br. Mycol. Soc.* 62.2 (1974): 283-288.

Greene, J., Ashburn, S. M., Razzouk, L., Smith, D. A. “Fish Oils, Coronary Heart Disease, and the Environment.” *Am J Public Health,* 103, 9, (2013): 1568-1576.

Guillard, R.L., Ryther, J.H. “Studies of Marine Planktonic Diatoms: I. Cyclotella Nana Hustedt, and Detonula Confervacea (Cleve) Gran.” *Canadian Journal of Microbiology.* 8.2 (1962): 229-239

Gupta, A., Barrow, C.J., Puri, M. “Omega-3 biotechnology: Thraustochytrids as a novel source of omega-3 oils.” *Biotechnology Advances,* (2012) doi:10.1016/j.biotechadv.2012.02.014

Gurevich A., Saveliev, V., Vyahhi, N., Tesler G. “QUAST: quality assessment tool for genome assemblies.” *Bioinformatics* 29 (2013): 1071-1075.

Honda, D., Yokochi, T., Nakahara, T., Raghukumar, S., Nakagiri, A., Schaumann, K., Higashihara, T. “Molecular Phylogeny of Labyrinthulids and Thraustochytrids Based on the Sequencing of 18S Ribosomal RNA Gene.” *J. Eukaryot. Microbiol.* 46.6 (1999): 637-647.

Iwata, I., Kimura, K., Tomaru, Y., Motomura, T., Koike, K., Koike, K., Honda, D. “Bothrosome Formation in *Schizochytrium aggregatum* (Labyrinthulomycetes, Stramenopiles) during Zoospore Settlement.” *Protist.* 168 (2017): 206-219.

Jain, R., Raghukumar, S., Sambaiah, K., Kumon, Y., Nakahara, T. “Docosahexaenoic acid accumulation in thraustochytrids: search for the rationale.” *Mar Biol.* 151 (2007): 1657-1664.

Ji, X.J., Mo, K.Q., Ren, L.J., Li, G.L., Huang, J.Z., Huang, H. “Genome Sequence of *Schizochytrium* Sp. CCTCC M209059, an Effective Producer of Docosahexaenoic Acid-Rich Lipids.” *Genome Announcements* 3.4 (2015): e00819–15. *PMC*. Web. 7 Feb. 2017.

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N. H., Phillippy, AM. “Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation.” *Genome Res* 27 (2017): 722-736.

Kris-Etherton, P.M., Harris, W.S., Appel, L.J. “Fish consumption, fish oil, omega-3 fatty acids, and cardiovascular disease.” Circulation. 2002;106(21):2747–2757.

Lee J.H., O’Keefe J.H., Lavie C.J., Harris, W.S. Omega-3 fatty acids: cardiovascular benefits, sources and sustainability. Nature Reviews. Cardiology. 2009;6(12):753–758.

Liu, B., Ertesvag, H., Aasen, I.M., Vadstein, O., Brautaset, T., Heggeset, T.M.B. “Draft Genome Sequence of the Docosahexaenoic Acid Producing Thraustochytrid

Aurantiochytrium Sp. T66." *Genomics Data* 8 (2016): 115–116. *PMC*. Web. 7 Feb. 2017.

Mo, C., Douek, J., Rinkevich, B. "Development of a PCR strategy for thraustochytrid identification based on 18S rDNA sequence." *Marine Biology.* 140 (2002): 883-889.

Pan, J.W., del Campo, J., Keeling, P.J. "Reference Tree and Environmental Sequence Diversity of Labyrinthulomycetes." *Journal of Eukaryotic Microbiology.* (2016)

Raghukumar, S. "Ecology of the marine protists, the Labyrinthulomycetes (Thraustochytrids and Labyrinthulids)." *Europ. J. Protistol.* 38 (2002): 127-145.

Raghukumar, S. "Thraustochytrid Marine Protists: Production of PUFAs and Other Emerging Technologies." *Mar Biotechnol.* 10, (2008): 631-640.

Raghukumar, S., Damare, V.S. "Increasing evidence for the important role of Labyrinthulomycetes in marine ecosystems." *Botanica Marina.* 54 (2011): 3-11.

Ruxton, C.S., Reed, S.C., Simpson, M.A., Millington, K. J., "The health benefits of omega-3 polyunsaturated fatty acids: a review of the evidence". *Journal of Human Nutrition and Dietetics*, 17, (2004): 449–459.

Stokes, N.A., Ragone Calvo, L.M., Reece, K.S., Burreson, E.M. "Molecular diagnostics, field validation, and phylogenetic analysis of Quahog Parasite Unknown (QPX), a pathogen of the hard clam *Mercenaria mercenaria." Dis Aquat Org.* 52 (2002): 233-247.