AN EVOLUTIONARY PERSPECTIVE ON SELECTING HIGH-LIPID-ACCUMULATING DIATOMS (BACILLARIOPHYTA): LITERATURE REVIEW, NEW DATA, AND FUTURE PROSPECTS

by

FRANCIS JOSEPH FIELDS IV

B.A., University of Colorado Boulder, 2012

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Masters of Arts

Department of Ecology and Evolutionary Biology

2014

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

This thesis entitled:
An Evolutionary Perspective on Selecting High-Lipid-Accumulating Diatoms (Bacillariophyta): Literature Review, New Data, and Future Prospects
written by Francis Joseph Fields IV
has been approved for the Department of Ecology and Evolutionary Biology

J. Patrick Kociolek

Barbara Demmig-Adams

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Fields, Francis Joseph (M.A., Ecology and Evolutionary Biology)

An Evolutionary Perspective on Selecting High-Lipid-Accumulating Diatoms (Bacillariophyta): Literature Review, New Data, and Future Prospects

Thesis directed by Professor J. Patrick Kociolek

ABSTRACT

Lipid-producing microalgae are a feedstock for commercial products such as nutritional supplements, aquatic animal feed, and biofuels. Unlike most algal phyla, the diatoms (Bacillariophyta) characteristically produce storage lipids throughout their entire lifecycle. This research compiled previously published data regarding diatom lipid accumulation under nutrient-replete and nutrient-deplete conditions as well as generating novel lipid and growth data from ten species of diatoms within lineages that have been historically under-examined. In this study, lipids were extracted via chloroform-methanol and quantified as percent dry weight, μg/mL, and pg/100 μm^3 and then analyzed for a phylogenetic signal by comparing the variability between lineages to the variability within lineages for each metric. These ten taxa were then paired with the data gathered from the literature and examined for a phylogenetic signal using previously described methods. In the first analysis, there was greater variability between than within lineages during stationary growth when using percent dry weight as a metric; the Biddulphiophycidae lineage accumulated a significantly lower amount of lipids than other lineages. In the second analysis, a statistically significant phylogenetic signal was detected for nutrient-deplete (i.e. stationary) growth experiments when examining the genus-level phylogeny ($P = 0.013$).

CONTENTS

TABLES

Table

FIGURES

Figure

1. INTRODUCTION

Diatoms (Bacillariophyta) are the most diverse and widely distributed group of unicellular microalgae with about 64,000 - 100,000 species in nearly 1,200 genera and 45 orders (Falkowski, 2004; Fourtainer & Kociolek, 1999; Norton et al., 1996; Round et al., 1990). The diversity of diatoms is most easily observed through the unique morphology of their ornate siliceous cell walls. Dissolved silica is taken up from the surrounding environment and deposited to form the majority of the cell wall, a process that requires less energy than the formation of an organic cell wall from, e.g., cellulose (Round et al. 1990). The Bacillariophyta are composed primarily of photosynthetic species found abundantly in aquatic ecosystems and estimated to contribute 25% of total global primary productivity (Scala & Bowler, 2001). The principal photosynthetic pigments of diatoms are chlorophyll a, chlorophyll c, and the accessory pigment fucoxanthin, the latter of which absorbs wavelengths in the green region of the solar spectrum poorly absorbed by chlorophylls and is as effective at stimulating photosynthesis as chlorophyll a (Jorgensen, 1977). In addition to absorbing a wide range of wavelengths, diatoms possess the most efficient ribulose-1,5-bisphosphate carboxylase oxygenase (Rubisco) enzyme among autotrophs (Giordano et al., 2005). Together, these characteristics may have allowed diatoms to evolve into the most ecologically diverse and efficient photoautotrophic microalgae.

The main photosynthates produced by diatoms are chrysolaminarin and lipids, with neutral lipids such as triacylglycerides (TAGs) making up 20-60% of total lipid content (Chen, 2012; Scholz & Liebezeit, 2013). The efficient accumulation of high-energy lipids

consequentially makes diatoms a main food source for many aquatic heterotrophs (Chen, 2012) as well as a key candidate for the commercial production of valuable products, such as feed for organisms in aquaculture, vitamin and nutritional supplements, and feedstock for the production of biofuels, such as biodiesel (Graham et al., 2011).

Biofuels, liquid or gaseous fuels used predominately for transportation and electricity (Demirbas, 2007), are derived from recently living biomass as a renewable source of energy and mitigate the accumulation of atmospheric carbon dioxide (Sheehan et al., 1998). The goals of domestic energy security and global environmental stability have resulted in an increased production of renewable fuels (Demirbas, 2007). While no single biofuel may replace fossil fuels, biodiesel, derived from neutral lipids, can make a significant impact on commercial transportation demands (Sheehan et al., 1998). Currently, soybean crop is the main feedstock for biodiesel production in the United States (Demirbas, 2007) and is cultivated by traditional agricultural means, requiring large amounts of fresh water, fertilizers (produced with large fossil-fuel inputs), and land area (Demirbas, 2007; Sheehan et al., 1998). The many estimates of land area required to replace petrol-diesel with biodiesel from conventional crop-based feedstocks conclude that there is not enough arable land for both food and fuel production (Chisti, 2007; Demirbas, 2007; Demirbas & Demirbas, 2011; Graham et al., 2001; Sheehan et al., 1998).

From 1978 - 1996, the U.S. Department of Energy's (DOE) Office of Fuels Development allocated resources to researching the possibility of developing transportation fuels from algae (Sheehan et al., 1998). This program, referred to as the Aquatic Species Program, was focused on identifying algal species capable of accumulating high volumes of lipids as a feedstock for biodiesel production (Sheehan et al., 1998). Over

the 18 years of this program, the National Renewable Energy Laboratory and subcontracted researchers collected approximately 3,000 algal strains and analyzed about 300 species for lipid accumulation, growth rates, and/or metabolic responses to stress (Sheehan et al., 1998). Environmental stressors (e.g. nutrient depletion, intense light, high or low temperature) were identified as mechanisms to arrest cell division (especially vegetative mitosis) and increase the accumulation of lipids, thus creating a trade-off between total algal biomass produced and total lipid content per cell. Most efforts were focused on species in the Chlorophyta ("green" algae) and Bacillariophyta due to their ability to accumulate neutral lipids. However, unlike diatoms, only a small percentage of Chlorophyta species can accumulate neutral lipids and these taxa may require an environmental stress trigger to do so (Chen, 2012; Hu et al., 2008; Sheehan et al., 1998).

In 2010 the U.S. DOE's Offices of Biomass and Energy Efficiency and Renewable Energy funded an additional follow up study, providing a roadmap for the algal industry, identifying its strengths and weaknesses and outlining the necessary research that must still be conducted (US DOE, 2010).

In these two reports, the DOE makes a compelling case to support research on, and development of, microalgae as a commercial feedstock, specifically for biofuel production. Reported advantages of microalgae include their abilities to produce a large amount of biomass per land area, minimize competition for arable land used in conventional agriculture, utilize abundant water resources like sea water, act as a sink for atmospheric carbon dioxide, and produce valuable industrial co-products (US DOE, 2010). However, the vast diversity of algal physiology creates a unique challenge in understanding which species to select. Species selection is the most influential factor in commercial production,

as it will determine environmental requirements, growth rates, lipid accumulation, and the availability of additional co-products (Chisti, 2007; Demirbas & Demirbas, 2011; Sheehan et al., 1998). Therefore, the DOE's 2010 roadmap first called for a deeper understanding of algal biology with the goal of being able to quickly analyze a large diversity of species for characteristics desirable to commercial cultivators (US DOE, 2010).

While numerous morphological and molecular studies have addressed the diversity and evolutionary relationships of the diatoms (Ashworth et al., 2013; Round et al., 1990; Sims et al., 2006; Theriot et al., 2010), information on phenotypes important for commercial application (i.e. lipid content, growth rate) is scarce. Table 1 summarizes the available data on diatom lipid accumulation, as percent lipid of algal dry weight in nutrient-replete and nutrient-deplete conditions, for 62 diatom species in 22 genera and 9 orders. Quantitative information on lipid content is available for less than 2% of all diatom genera after 30 years of research (Figure 1). Detailing the lipid-accumulation capacity of all known diatom species to understand which groups are most productive would be an impractical endeavor.

The concept of a phylogenetic signal, i.e., the tendency of related species to resemble each other (Blomberg et al., 2003), could help expedite the quantification of lipids in diatoms. If diatom lipid accumulation were to adhere to the phylogenetic signal theory, diatom species closely related to each other should accumulate similar amounts of lipids. If this were true, only a few species from each lineage would need to be assayed for lipid content to allow the lipid content for the remaining species in that lineage to be inferred. In essence, this approach would provide an evolutionary-based guide to the most desirable lineages. With such a guide in hand, commercial cultivators could examine the local

Table 1 Average lipid content per dry weight (% lipid per dry weight) for 62 diatom species and their corresponding orders reported in the literature; n.a. = no data available

Order	**Species**	**Nutrient Replete**	**Nutrient Deplete**	**References[a]**
Achnanthales		23.98%	30.24%	
	Achnanthes brevipes	19.60%	n.a.	13
	Achnanthes delicatula hauckiana	29.80%	n.a.	13
	Achnanthes exigua	25.60%	n.a.	13
	Cocconeis peltoides	20.90%	n.a.	13
	Cocconeis scutellum	n.a.	30.24%	2
Bacillariales		22.41%	42.20%	
	Cylindrotheca closterium	20.30%		13
	Cylindrotheca sp	16.00%	36.00%	2, 3
	Hantzschia DI-160	n.a.	66.00%	14
	Hantzschia DI-60	29.00%	53.00%	14
	Nitzschia closterium	26.80%	n.a.	10, 11
	Nitzschia constricta	29.70%	n.a.	13
	Nitzschia dissipata	n.a.	46.00%	14
	Nitzschia epithemoides	22.30%	n.a.	13
	Nitzschia frustulum	24.75%	n.a.	10, 11
	Nitzschia grossestriata	n.a.	41.49%	2
	Nitzschia longissima	21.00%	n.a.	10
	Nitzschia ovalis	11.80%	n.a.	10
	Nitzschia panduriformis	n.a.	38.99%	2
	Nitzschia sp	n.a.	46.00%	3
	Nitzschia TR-114	n.a.	28.10%	14
	Tryblionella naviularis	24.20%	n.a.	13
Chaetocerotales		32.77%	37.38%	
	Chaetoceros calcitrans	39.80%	n.a.	12
	Chaetoceros muelleri	25.73%	37.38%	2, 9, 12, 14
Fragilariales		10.20%	n.a.	
	Fragilaria sp	10.20%	n.a.	10
Melosirales		18.50%	33.02%	
	Melosira nummuloides	n.a.	33.02%	2
	Melosira sp	18.50%	n.a.	11
Naviculales		22.85%	37.61%	
	Caloneis platycephala	n.a.	38.39%	2
	Diploneis didyma	25.30%	n.a.	13
	Diploneis littoralis	22.90%	n.a.	13
	Gyrosigma littorale	30.10%	n.a.	13
	Gyrosigma peisoneis	22.40%	n.a.	13
	Gyrosigma spenceri	34.50%	n.a.	13

Order	Species	Nutrient Replete	Nutrient Deplete	Reference[a]
	Navicula acceptata	21.80%	42.70%	14
	Navicula digito-radiata	18.90%	n.a.	13
	Navicula forcipata	19.70%	n.a.	13
	Navicula gregaria	17.30%	n.a.	13
	Navicula jeffreyi	6.60%	n.a.	8
	Navicula lyra	n.a.	37.59%	2
	Navicula pelliculosa	31.95%	33.80%	4, 10
	Navicula perminuta	23.80%	n.a.	13
	Navicula phyllepta	23.10%	n.a.	13
	Navicula pseudotenelloides	n.a.	42.50%	14
	Navicula salinicola	25.50%	n.a.	13
	Navicula saprophila	26.00%	44.00%	14
	Navicula sp	13.75%	n.a.	8, 11
	Phaeodactylum tricornutum	17.64%	26.93%	1, 3, 5, 10, 12, 14
	Pinnularia ambigua	19.40%	n.a.	13
	Pleurosigma angulatum	33.50%	n.a.	13
	Seminavis gracilenta	n.a.	34.99%	2
Thalassiophysales		24.29%	35.79%	
	Amphora angusta	21.40%	n.a.	13
	Amphora arenaria	24.70%	n.a.	13
	Amphora bigibba	n.a.	34.30%	2
	Amphora exigua	25.10%	39.42%	2, 10, 13
	Amphora graeffii	24.40%	n.a.	13
	Amphora hyaline	22.10%	33.65%	14
	Amphora sp	28.03%	n.a.	7, 10
Thalassiosirales		18.32%	34.40%	
	Cyclotella cryptica	14.10%	41.57%	14
	Skeletonema costatum	21.10%	37.62%	2, 12
	Skeletonema sp	17.55%	n.a.	8, 12
	Thalassiosira pseudonana	16.83%	n.a.	8, 10, 12
	Thalassiosira weisflogii	22.00%	24.00%	6
Triceratiales		11.00%	n.a.	
	Odontella aurita	11.00%	n.a.	10

[a]Key to references: 1 Ceron Garcia et al. (2000); 2 Chen (2012); 3 Chisti (2007); 4 Coombs et al. (1967); 5 Gatenby et al. (2003); 6 Griffiths et al. (2009); 7 De la Pena (2007); 8 Mansour et al. (2005); 9 Nagle & Lemke (1990); 10 Orcutt & Patterson (1975); 11 Renaud et al. (1994); 12 Rodolfi et al. (2008); 13 Scholz & Liebezeit (2013); 14 Sheehan et al. (1998)

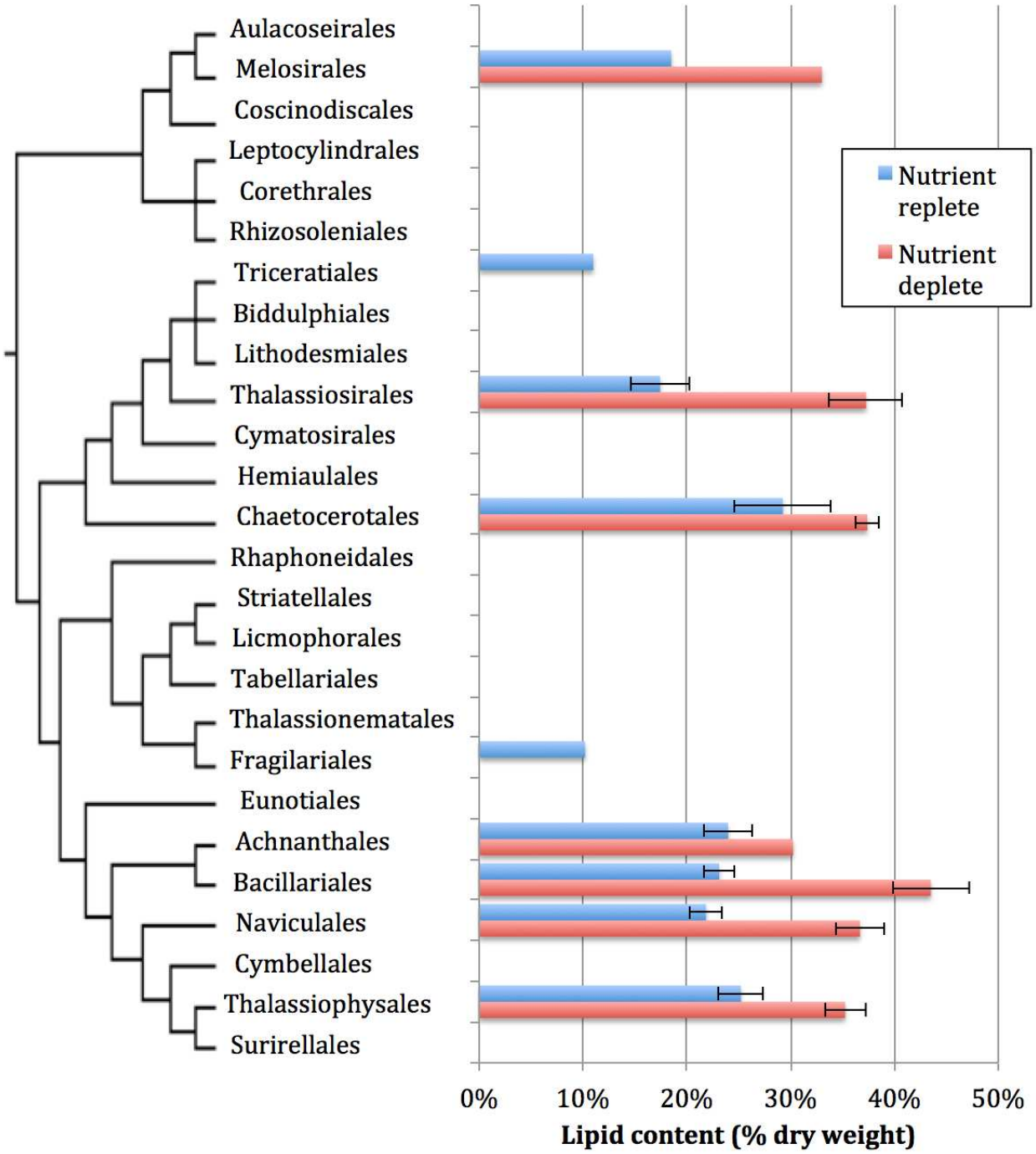

Figure 1. Average lipid content per dry weight from literature sources in nutrient-replete and nutrient-deplete conditions for various groups of diatoms. Data are grouped by diatom order and compared within the context of a cladogram of the Bacillariophyta derived from Sims et al. (2006); for the majority of orders, no data are available. Error bars indicate standard error of the mean

microbial community and use the guide to select species from the highest producing lineage(s). Additionally, research could be expanded towards lineages that contain species with naturally elevated lipid content to maximize the efforts of additional molecular work such as genetic engineering.

The present research is the first to test the viability of an evolutionary-based approach to analyze a diverse group of algae (e.g., the diatoms) for lipid content based on a combination of data available in the literature as well as new data on lineages, for which little or no published lipid data are available. Additional lipid and growth data were generated from 10 species of diatoms within three lineages under-reported in the literature; lipid contents were expressed on various reference bases as percent dry weight and mass per volume (as μg/mL, and pg/100 μm^3). A balanced nested ANOVA design using nine species was conducted to compare the variability between lineages; the 10^{th} species was used to test the robustness of any conclusions drawn from the ANOVA results. The latter data were then combined with the data from the literature and mapped onto diatom phylogenies generated from 18S ribosomal RNA sequences gathered from GenBank to test for a phylogenetic signal at the subclass, ordinal, and genus level.

2. MATERIALS AND METHODS

2.1 Sample collection and taxon selection

Taxa targeted for collection were from lineages underrepresented in the literature (i.e., species within the subclasses Coscinodiscophycidae, Biddulphiophycidae, and Fragilariophycidae). Ecological and environmental descriptions from Round et al. (1990) were used to select sampling sites in Southern California and Colorado to locate these taxa. A light microscope (Olympus BX51) was used to examine the collected samples and identify which of the desired taxa were present (and available for isolation and culturing).

2.2 Media formulation, isolation, and species identification

Two growth media were used for diatom cultures, i.e., WC medium (Guillard, 1975) and artificial seawater prepared from Instant Ocean® sea salt. Artificial seawater was supplemented with the same amount of silica, nitrogen, phosphorus, and potassium as WC medium and prepared with a conductivity of 40 mS cm^{-1} and pH 8.10 (confirmed with a YSI Environmental model 556, Yellow Springs Instrument Company, Yellow Springs, Ohio). For both types of media, the standard amount of silica was doubled to 56.84 mg L^{-1} $Na_2SiO_3 \bullet 9H_2O$.

Single cells were isolated through serial dilution via micropipette into 16x125-mm test tubes containing 10-mL of the appropriate media. Non-axenic monocultures were allowed to grow for 2-3 weeks until successful isolates could be transferred to 125-mL

Erlenmeyer flasks. The resulting cultures (Table 2) were maintained in triplicate on open shelving at 24^{0}C under lighting with a 12:12 light:dark cycle at a photon flux density of 50 μmol m^{-2} s^{-1}; cultures were transferred into new media every two weeks prior to experimentation. Culture subsamples were cleaned with nitric acid to remove organic matter, mounted onto glass slides with naphrax, and the diatom cell walls were observed through a light microscope to identify species (Hoppenrath et al., 2009; Krammer & Lange-Bertalot, 2008; Round et al., 1990; Tomas, 1997); resulting slides and cleaned material are available in the Kociolek collection at the University of Colorado Boulder.

2.3 Experimental setup

Lipid content was quantified for each species during both logarithmic and stationary growth. Cultures for logarithmic growth analyses were grown in silica-replete media until growth had been logarithmic for at least 2 days and then harvested at the end of the light cycle (total growth time of 4-7 days). In contrast, the cultures for stationary growth analysis were not harvested during logarithmic growth but instead transferred into a silica-deplete medium to arrest cell division and induce a stationary phase; cultures were harvested at the end of the light cycle after growth had been stationary for 3 days (total growth time of 7-15 days).

2.4 Culturing trials and biomass quantification

Trials examining either logarithmic or stationary phases were carried out for each

Table 2 Species isolated for experimentation in the present study and their origins							
Species	Date collected	Location	GPS	Ecology	Media	JPK accession #	Slide #
Actinoptychus senarius	May 24, 2013	Ocean Beach San Diego, CA	32°45'24.1"N 117°14'52.1"W	Planktonic	Marine	9856	560; 76-77
Amphitetras antediluviana	March 24, 2013	Doheny State Beach Dana Point, CA	33°27'38.8"N 117°41'20.5"W	Epiphytic	Marine	9847	560; 58-59
Aulacoseira ambigua	June 3, 2013	Varsity Lake Boulder, CO	40°00'36.0"N 105°16'28.2"W	Planktonic	Fresh	5700	417; 96-100
Biddulphia alternans	May 24, 2013	Ocean Beach San Diego, CA	32°45'24.1"N 117°14'52.1"W	Episammic	Marine	9856	560; 76-77
Diatoma tenuis	May 9, 2013	Boulder Creek Boulder, CO	40°00'41.4"N 105°15'40.8"W	Planktonic	Fresh	9586	557; 85-86
Fragilariforma nitzschioides	May 9, 2013	Boulder Creek Boulder, CO	40°00'41.4"N 105°15'40.8"W	Planktonic	Fresh	9586	557; 85-86
Melosira nummuloides	March 25, 2013	Ocean Beach San Diego, CA	32°45'24.1"N 117°14'52.1"W	Planktonic	Marine	9854	560; 72-73
Odontella aurita	March 25, 2013	Ocean Beach San Diego, CA	32°45'24.1"N 117°14'52.1"W	Planktonic	Marine	9851	560; 66-67
Pleurosira laevis	March 15, 2013	Plum Creek Denver, CO	39°42'51.6"N 104°57'05.0"W	Epilithic	Fresh	9857	560; 78-79
Synedra acus	May 9, 2013	Boulder Creek Boulder, CO	40°00'41.4"N 105°15'40.8"W	Epiphytic	Fresh	9586	557; 85-86

species in triplicate within an environmentally controlled growth chamber (VWR diurnal growth chamber model 2015, Van Waters & Rodgers, Inc., Radnor, PA; 12:12h light:dark cycle, 20^{0}C, photon flux density of 130 umol m^{-2} s^{-1}). 250 mL of silica-replete media were inoculated with 2 mL of the isolated monoculture in 250-mL glass Erlenmeyer flasks, stirred constantly, and bubbled with hydrated CO_2 at atmospheric concentrations (Figure 2a). Every 24-48, hours a 2-mL subsample was removed to estimate cell density (cells/mL) via cell counts (Palmer-Maloney slide, 100-uL well, average counts of 3-6 random optical fields) and optical density measurements (Eppendorf Biophotometer, OD 600, Eppendorf, Hamburg, Germany). Specific growth rate, which is the number of individuals produced by each individual in the population over some unit of time (μ; d^{-1}), was calculated using the formula $\mu = 1/t \times \ln N_t/N_0$ where N_0 is initial cell density and N_t is cell density after t days of growth (Wang et al. 2009). Once a culture was ready to be harvested, the final volume was measured and the whole culture was filtered onto a pre-massed fiberglass filter (1-um porosity) (Figure 2b), allowed to dry, and massed to quantify total dry biomass (Figure 2c).

2.5 Lipid Extraction

Total lipids were extracted from dried biomass in a Soxhlet extractor (125-mL bottom flask) with 80 mL of chloroform-methanol (2:1; v/v) as a solvent (methods modified from Folch et al., 1957 and Laurens et al., 2012) (Figure 3a). The Soxhlet extractor was allowed to run for a minimum of 16 hours to insure that all lipids were dissolved into solution. The remaining solution was processed through a Kuderna-Danish concentrator to evaporate most of the chloroform-methanol solvent (Figure 3b) and the resulting lipid

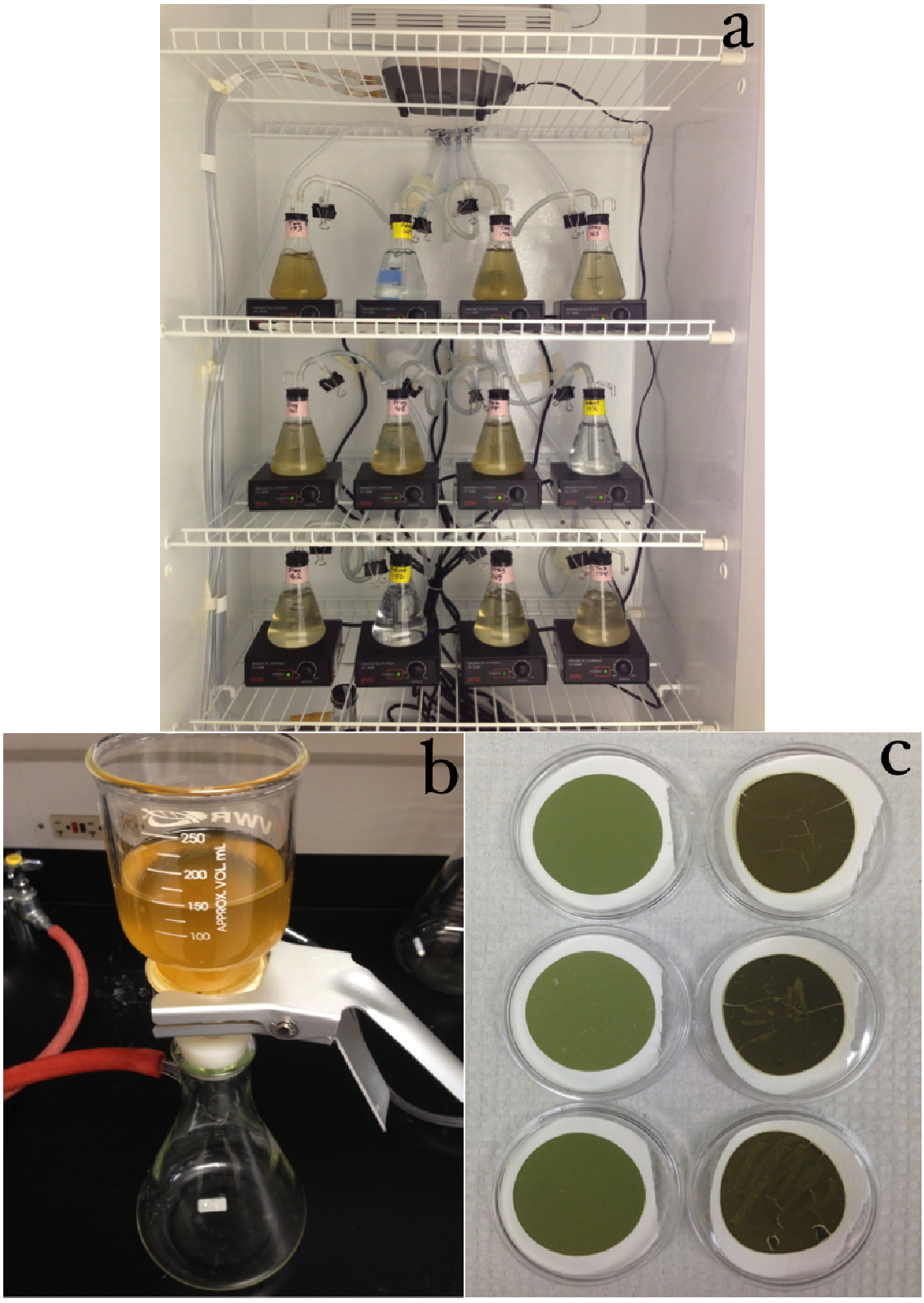

Figure 2. Experimental procedures including: (a) growth chamber with stir plates and aeration system; (b) cultures being filtered; and (c) dried biomass of *O. aurita* on a fiberglass filter from logarithmic (left column) and stationary (right column) trials

extracts were transferred into pre-massed vials and allowed to air dry (Figure 3c). Lipids were washed with water and then chloroform to remove any water-soluble compounds, such as salts, that persisted through the extraction and condensation processes. Vials were again allowed to air dry until only the dry total lipid extract remained to be quantified.

Extracted lipids were expressed as different metrics, i.e., percent dry weight (dry lipid mass/dry biomass * 100; % dry weight), and mass per volume (expressed as either µg of lipid per culture volume [dry lipid mass/final culture volume; µg/mL], or as pg of lipid per individual algal cell volume [dry lipid mass/cell volume x 100; pg/100µm^3]).

2.6 Calculating average cell volume

Volumetric formulas previously suggested for each genus (Hillebrand et al., 1999) were used to calculate cell volume. A light microscope equipped with an ocular micrometer was used to measure cellular dimensions for each species; dimensions of ten cells were averaged and used in the final volumetric calculation.

2.7 Phylogenetic comparative methods

Lipid data generated via the present study were divided into three corresponding subclasses. ANOVAs were used to test for significant differences between species within each subclass and TukeyHSDs were then used to generate *P* values describing those differences in RStudio (v0.98.501) for each metric. A 2-way ANOVA was then used to compare lipid content between logarithmic and stationary growth.

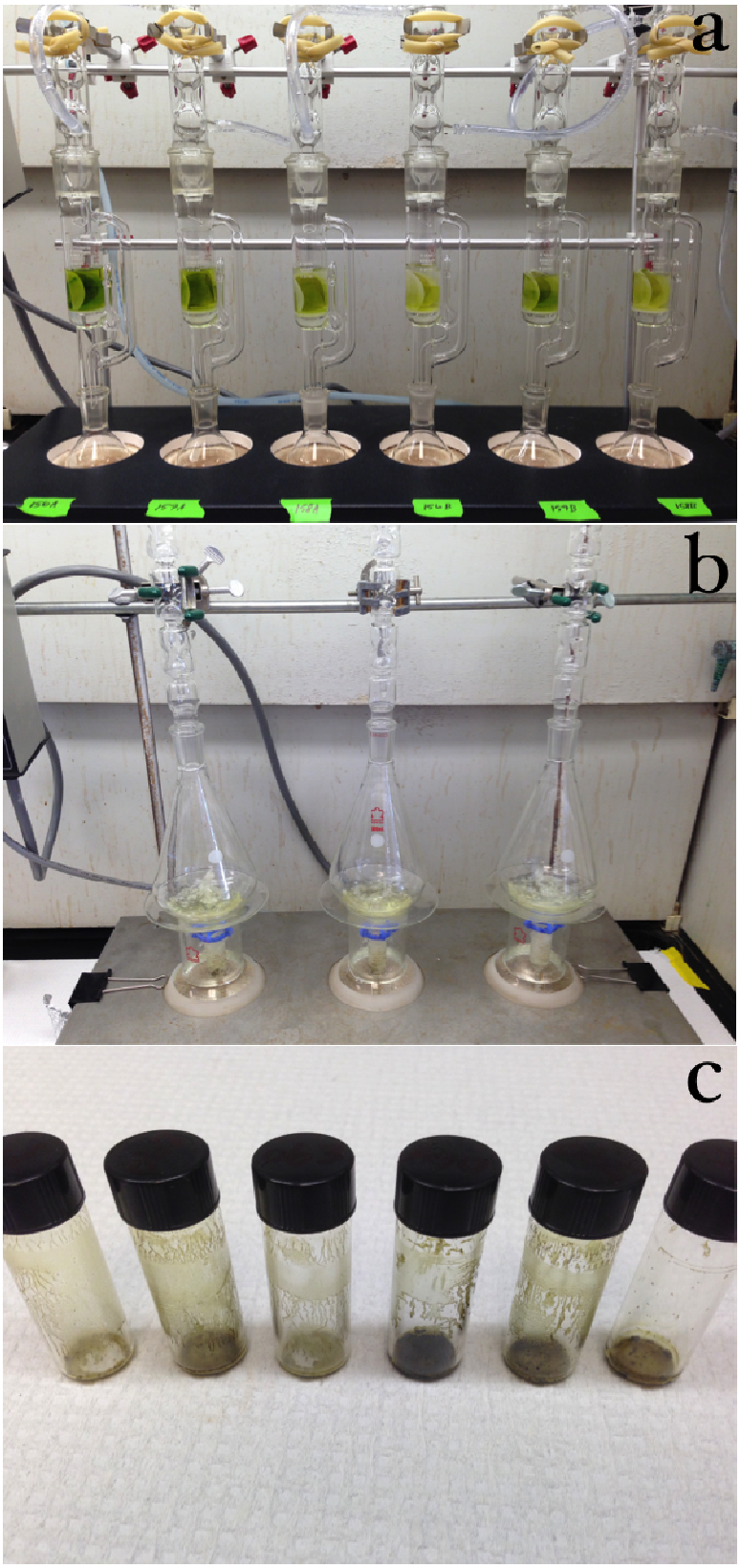

Figure 3. Experimental procedures, including: (a) Soxhlet-extractor, (b) Kuderna-Danish concentrator, and (c) dried extracted lipids

A balanced nested ANOVA design was used to assess whether a phylogenetic signal existed for lipid content by lineage (i.e., whether lipid content varied more strongly between lineages versus within lineages) for a group of nine of the ten species. A t-test was then used to test for consistency by comparing the tenth species to the other three species within that same lineage.

Data generated in the present study were then combined with data from the literature to assess whether a phylogenetic signal was present across the entire diatom phylogeny, as described by Blomberg et al. (2003). The combined data were analyzed on three different taxonomic levels, i.e., subclass, order, and genus. Phylogenetic trees were generated from GenBank 18S ribosomal RNA sequences (listed in the results), aligned in Geneious (v5.6.6, Biomatters, Auckland, New Zealand), rooted in SeaView (v4.4.3, Gouy et al., 2010), and used to calculate the significance of a phylogenetic signal and the *K* statistic in RStudio using the phylosignal command from the picante package (v1.6-2) with 100,000 randomizations (Blomberg et al., 2003).

Two analyses were conducted for each tree, one with lipid data generated from nutrient-replete (i.e. logarithmic growth) experiments and the other with data from nutrient-deplete (i.e. stationary growth) experiments. For each of the three phylogenies, the combined lipid data were sorted into their appropriate taxonomic group (i.e. subclass, order, or genus) and paired with the corresponding branch on the phylogeny.

The phylosignal test in RStudio was used to assess whether there was a significant phylogenetic signal at each taxonomic level by calculating a *P* value based on the variance of independent contrasts relative to tip shuffling randomizations (Blomberg et al., 2003). The phylosignal test also calculated the *K* statistic used to quantify the amount of

phylogenetic signal on a scale of zero to infinity and allowing comparisons across differing traits and trees (Blomberg et al., 2003).

3. RESULTS

3.1 Coscinodiscophycidae

Examination with a light microscope resulted in the identification of *Aulacoseira ambigua* (Grunow) Simonsen*, Melosira nummuloides* (Agardh) Greville in Hooker*,* and A*ctinoptychus senarius* (Ehrenberg) Ehrenberg as the three taxa examined within the Coscinodiscophycidae (see images in Figure 4a-c).

As shown in Figure 5, *A. ambigua and M. nummuloides* displayed similar growth patterns in both the logarithmic and stationary trials while *A. senarius* grew slower and not as dense. *Melosira nummuloides* exhibited the highest specific growth rate, followed by *A. ambigua*, and then *A. senarius* (Table 3). During both trials, *M. nummuloides* and *A. senarius* produced more biomass than *A. ambigua* (Table 3). *Aulacoseira ambigua* had the smallest cell volume of 277 μm^3, while *M. nummuloides* had the largest with 5,491 μm^3; *A. senarius* had a volume of 1,775 μm^3 (Table 4).

Table 5 shows statistical outcomes of the resulting TukeyHSD analyses performed for all three metrics within the Coscinodiscophycidae in both logarithmic and stationary trials. During logarithmic growth, lipid content, as percent dry weight, of *M. nummuloides* was significantly lower than that of either *A. ambigua* or *A. senarius* ($P = 0.027$; $P = 0.012$; Figure 6a). Lipid mass per culture volume or cell volume (as µg/mL and pg/100μm^3, respectively) of *A. senarius* was significantly higher than that of *A. ambigua* ($P = 0.002$; $P = 0.008$; Figure 6b) and *M. nummuloides* ($P = 0.007$; $P = 0.006$; Figure 6c).

During stationary growth, there was no significant difference between species in

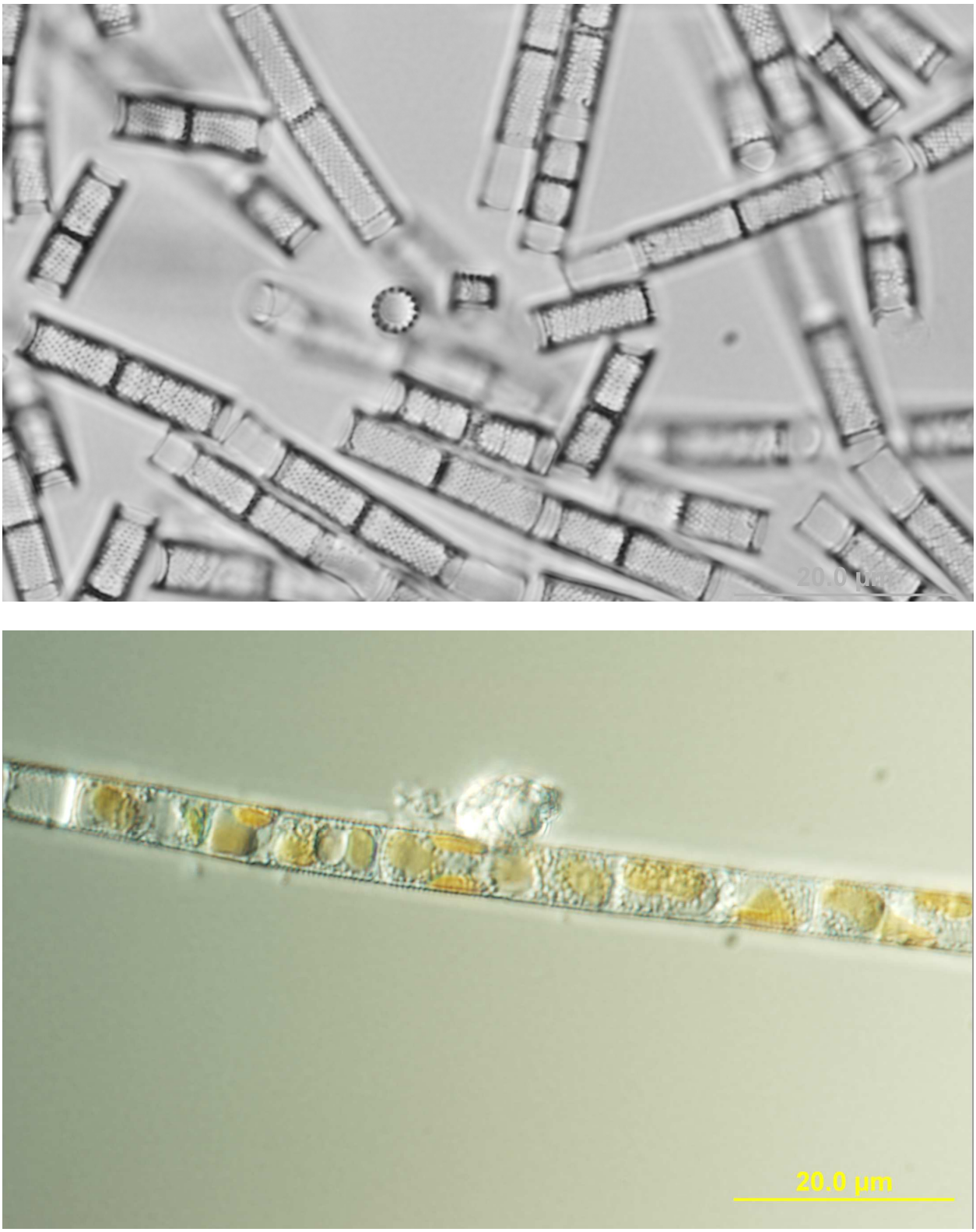

Figure 4a. *Aulacoseira ambigua* viewed with a light microscope under a 100x objective with oil; scale bar show is 20 µm in both the cleaned sample (top) and live culture (bottom)

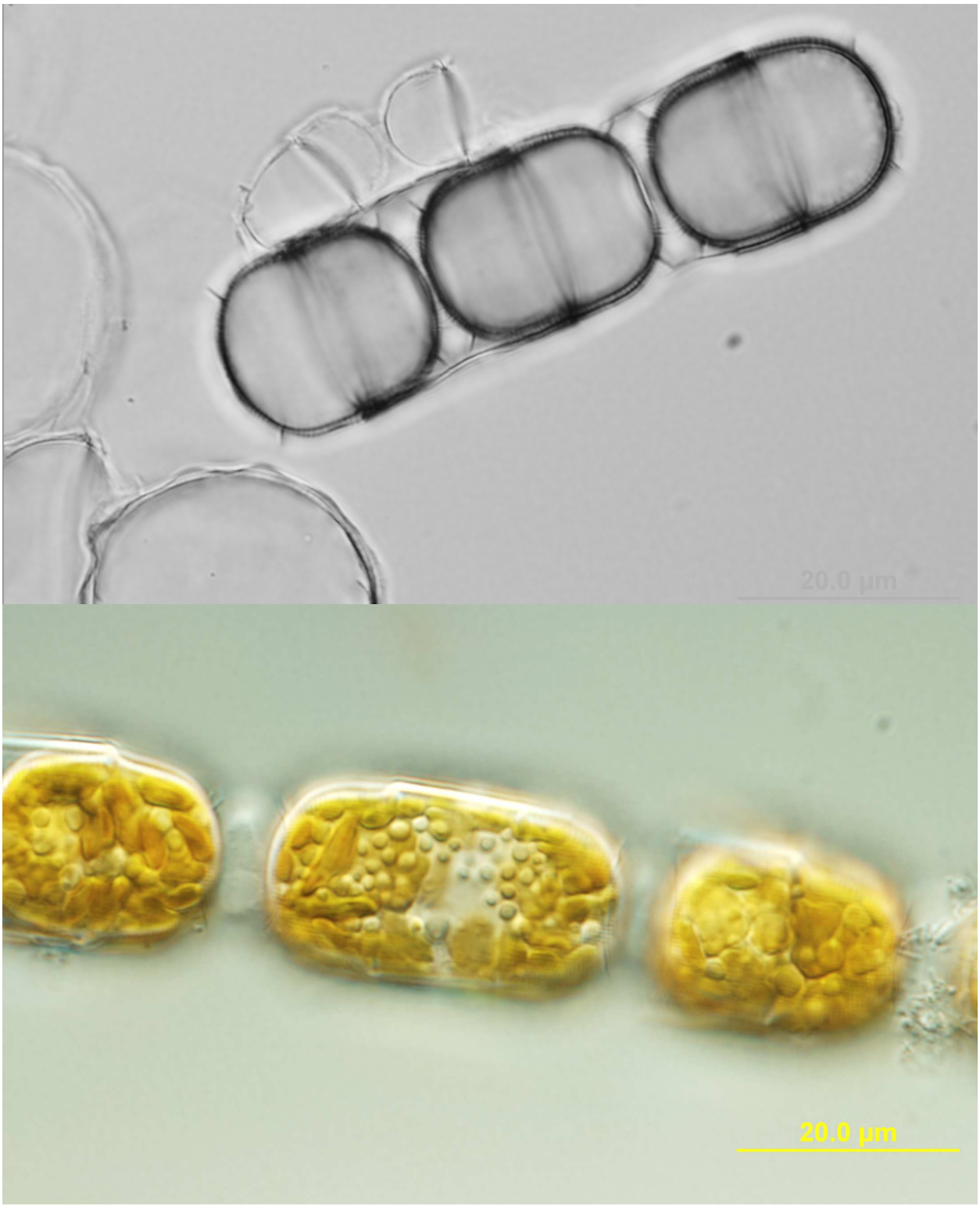

Figure 4b. *Melosira nummuloides* viewed with a light microscope under a 100x objective with oil; scale bar show is 20 µm in both the cleaned sample (top) and live culture (bottom)

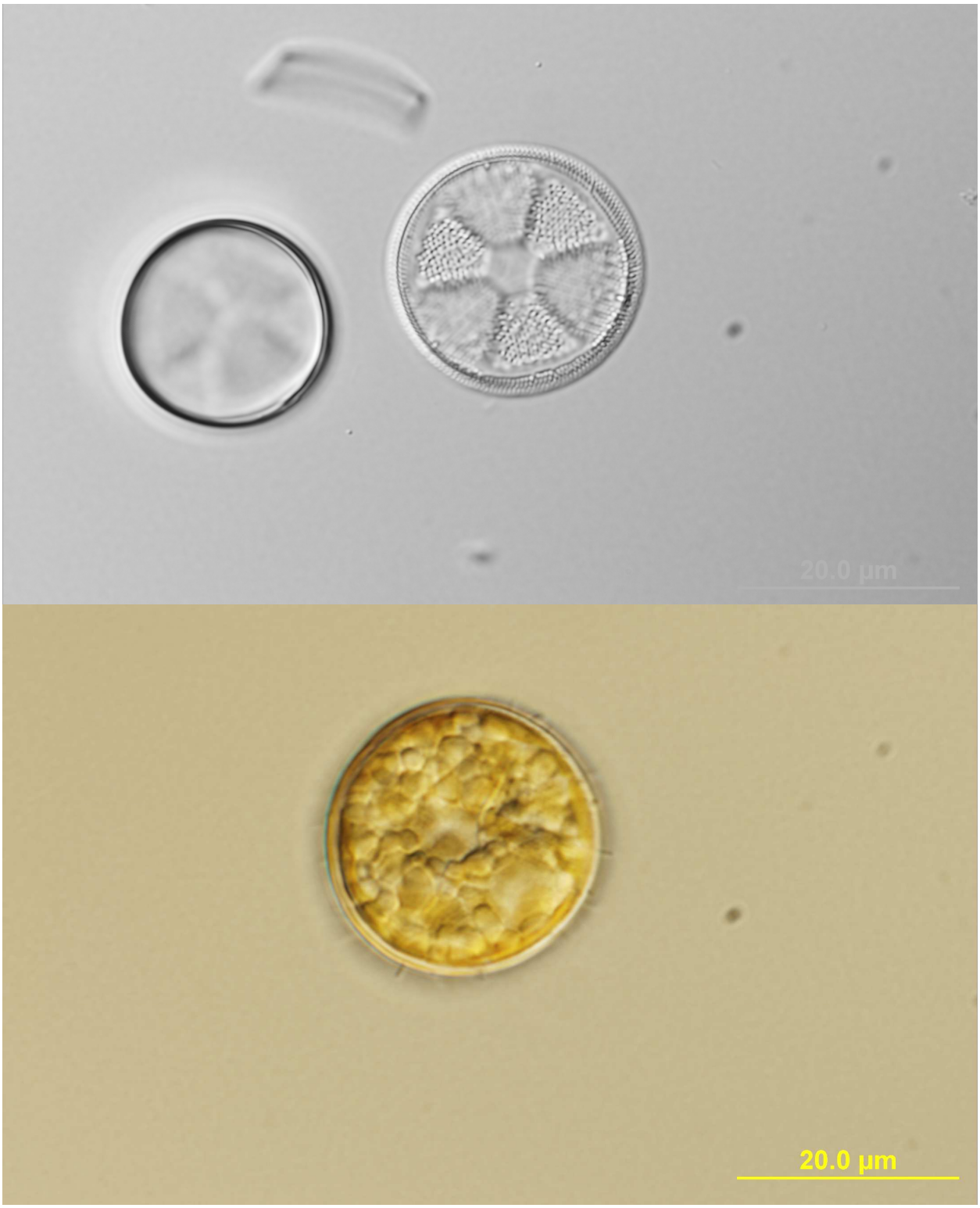

Figure 4c. *Actinoptychus senarius* viewed with a light microscope under a 100x objective with oil; scale bar show is 20 µm in both the cleaned sample (top) and live culture (bottom)

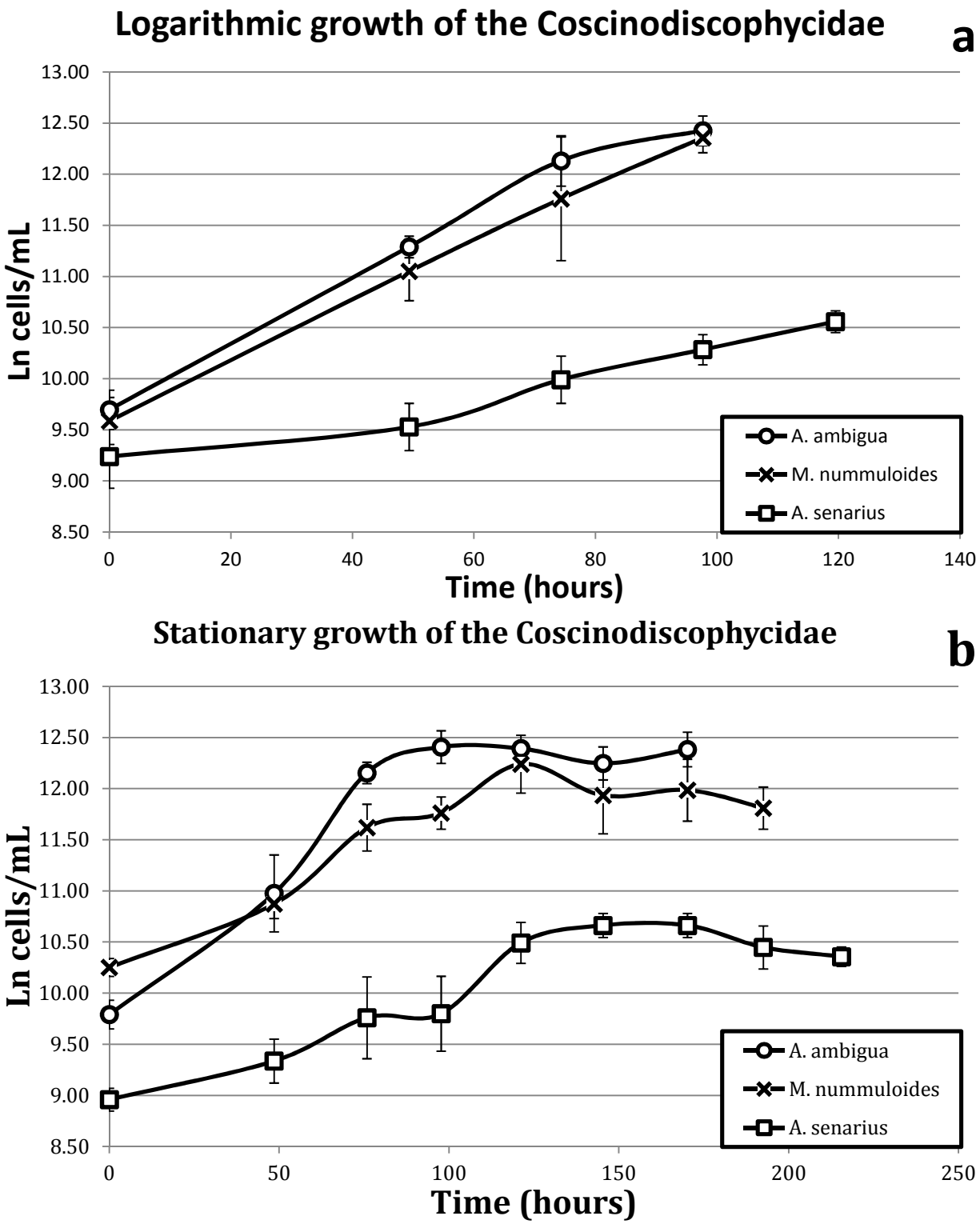

Figure 5. (a) Logarithmic and (b) stationary growth trials for the Coscinodiscophycidae; points are averages of the natural log of cell number per milliliter from three replicates and error bars denote the standard error

Table 3 Growth data for the Coscinodiscophycidae

		Logarithmic trial			Stationary trial		
Species	Specific growth rate (μ max; d^{-1})	Trial length (h)	Final density (cells/mL)	Dry biomass (mg)	Trial length (h)	Final density (cells/mL)	Dry biomass (mg)
A. ambigua	0.85	102	254070 ± 33,947	14.96 ± 0.47	174	245409 ± 39,084	15.00 ± 0.21
M. nummuloides	0.94	98	237469 ± 32,253	35.2 ± 2.61	193	140749 ± 30,984	39.70 ± 3.59
A. senarius	0.57	120	38977 ± 4,331	34.90 ± 0.90	216	31759 ± 2,887	37.37 ± 0.52

Table 4 Cell volume, total extracted lipid mass, and lipid content on various bases for the Coscinodiscophycidae									
		Logarithmic				Stationary			
Species	Cell volume (μm^3)	Lipid mass (mg)	% dry weight	µg/mL	pg/100 μm^3	Lipid mass (mg)	% dry weight	µg/mL	pg/100 μm^3
A. ambigua	277 ± 19	2.93 ± 0.17	19.70 ± 1.70	13.06 ± 0.79	19.00 ± 1.67	3.53 ± 0.06	23.57 ± 0.72	15.48 ± 0.05	24.11 ± 4.27
M. nummuloides	5491 ± 1,337	3.73 ± 0.54	10.52 ± 0.87	16.47 ± 2.58	1.28 ± 0.17	8.90 ± 1.53	22.09 ± 1.92	39.55 ± 6.88	5.30 ± 0.81
A. senarius	1775 ± 217	7.46 ± 0.67	21.52 ± 2.50	32.00 ± 2.87	47.50 ± 7.22	8.70 ± 0.72	23.33 ± 2.20	38.45 ± 3.22	68.46 ± 3.02

Table 5 *P* values for pairwise statistical comparisons by TukeyHSDs within the Coscinodiscophycidae; * denotes values less than *P* = 0.05						
	Logarithmic			Stationary		
Species comparison	% dry weight	µg/mL	pg/100 μm^3	% dry weight	µg/mL	pg/100 µm3
A. ambigua - M. nummuloides	0.027*	0.570	0.059	0.823	0.019*	0.011*
M. nummuloides - A. senarius	0.012*	0.007*	$6.4x10^{-4}$*	0.872	0.982	$1.5x10^{-5}$*
A. senarius - A. ambigua	0.767	0.002*	0.007*	0.994	0.023*	$1.2x10^{-4}$*

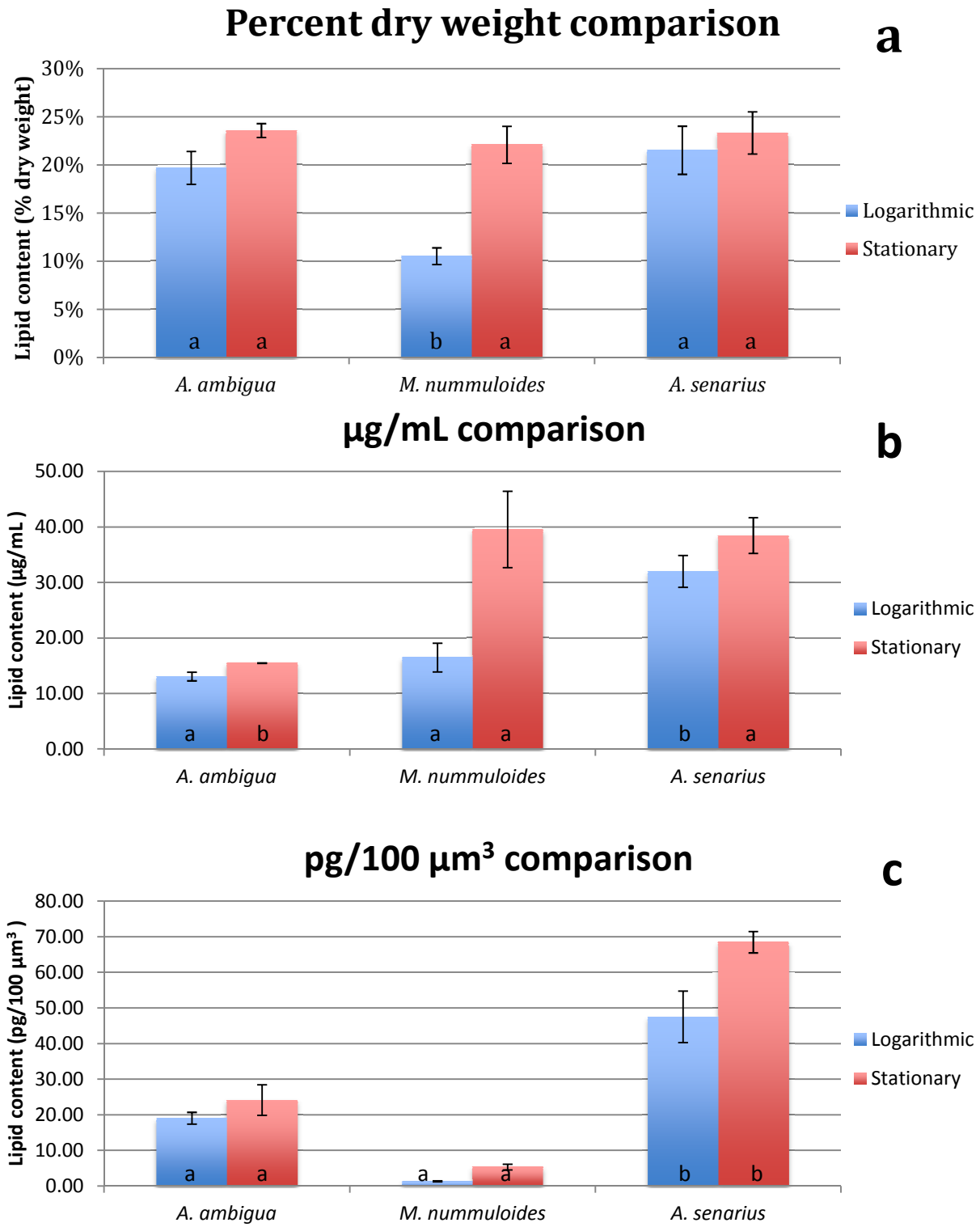

Figure 6. Total lipid accumulation during the logarithmic trial (blue bars) and stationary trial (red bars) quantified as (a) percent dry weight, (b) lipid mass per culture volume in µg/mL, and (c) lipid mass per cell volume in pg/100µm^3; Lowercase letters indicate significant differences within a trial (same letter = no significant difference; different letter = significant difference)

lipid content expressed as percent dry weight (P = 0.816; Figure 6a). In contrast, lipid mass per culture volume (μg/mL) was significantly lower in *A. ambigua* than in *M. nummuloides* and *A. senarius* (P = 0.019; P = 0.023; Figure 6b). Furthermore, the three species also differed significantly from each other in lipid mass per algal cell volume (pg/100um^3), with *A. senarius* yielding the highest value and *M. nummuloides* yielding the lowest (Figure 6c).

3.2 Biddulphiophycidae

Examination with a light microscope resulted in the identification of *Odntella aurita* var. minima (Grunow) De Toni, *Pleurosira laevis* (Ehrenberg) Compère, *Amphitetras antediluviana* Ehrenberg, and *Biddulphia alternans* (J.W. Bailey) Van Heurck as the four taxa examined within the Biddulphiophycidae (see images in Figure 7a-d).

As shown in Figure 8, growth patterns in the logarithmic trial were similar between all four species. In contrast, during the Stationary phase, *O. aurita* exhibited extended growth despite being in silica-deplete media. *Pleurosira laevis* exhibited the highest specific growth rate, followed by *O. aurita*, *A. antediluviana,* and then *B. alternans* (Table 6). Biomass production was consistent between species during both trials, except for *O. aurita* that produced twice as much biomass in the stationary trial (Table 6). *Odontella aurita* and *B. alternans* had the smallest cell volumes of 2,490 and 2,971 μm^3, respectively, while *P. laevis* had a volume of 28,894 μm^3 and *A. antediluviana* had the largest volume of 44,748 μm^3 (Table 7).

Table 8 shows the statistical outcomes of the resulting TukeyHSD analyses performed for all three metrics within the Biddulphiophycidae in both the logarithmic and

Figure 7a. Odontell aurita viewed with a light microscope under a 100x objective with oil; scale bar show is 20 µm in both the cleaned sample (top) and live culture (bottom)

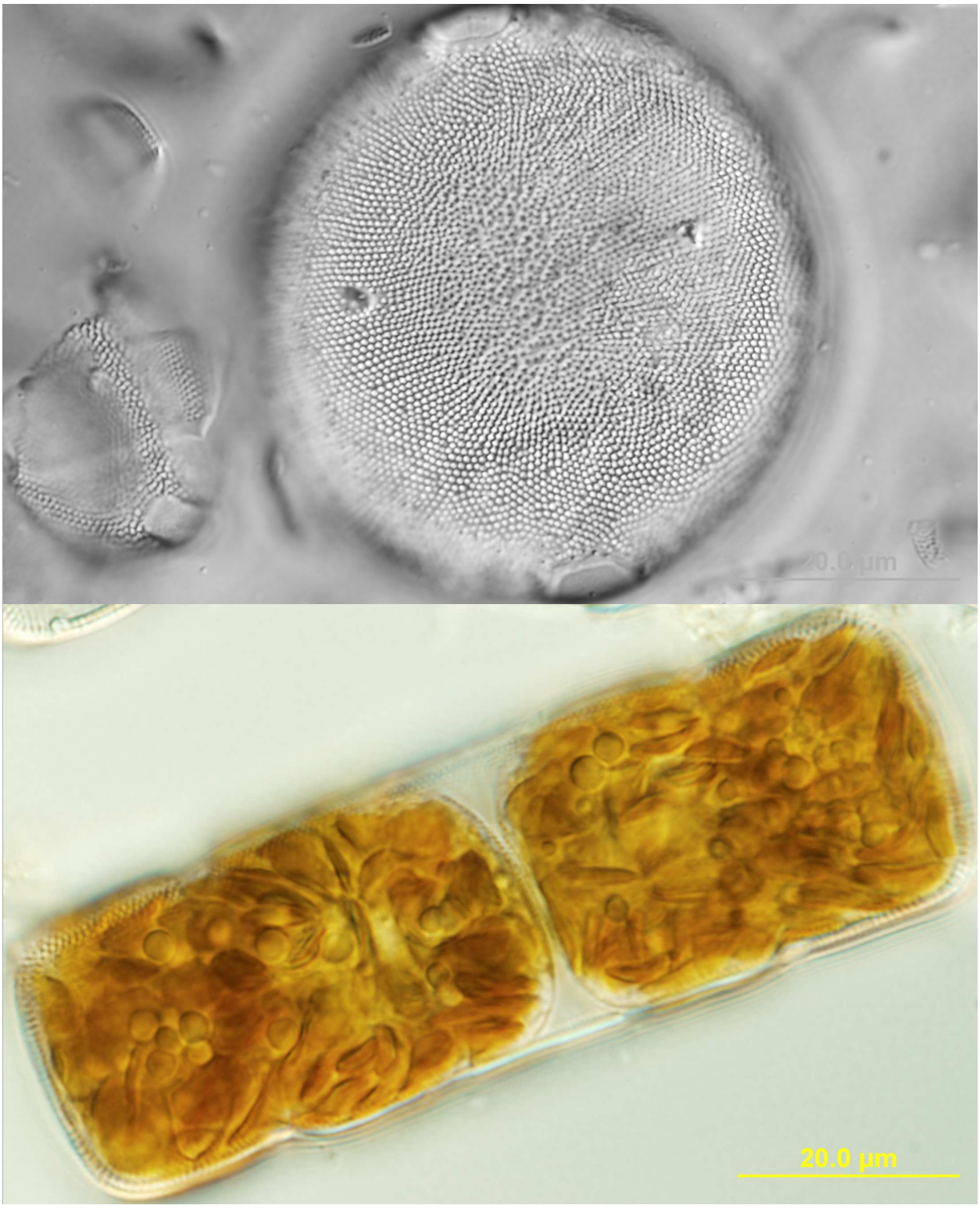

Figure 7b. *Pleurosira laevis* viewed with a light microscope under a 100x objective with oil; scale bar show is 20 µm in both the cleaned sample (top) and live culture (bottom)

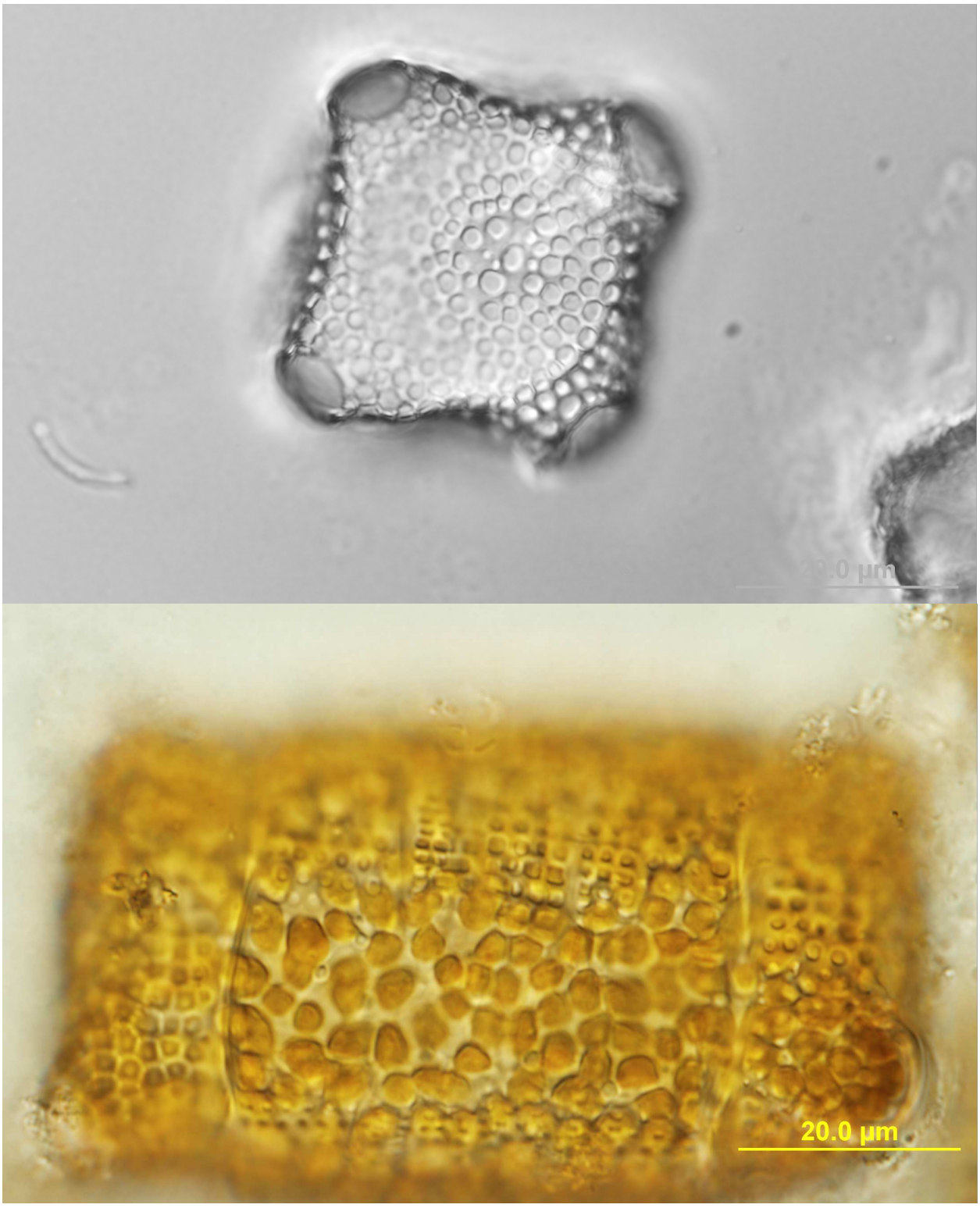

Figure 7c. *Amphitetras antediluviana* viewed with a light microscope under a 100x objective with oil; scale bar show is 20 µm in both the cleaned sample (top) and live culture (bottom)

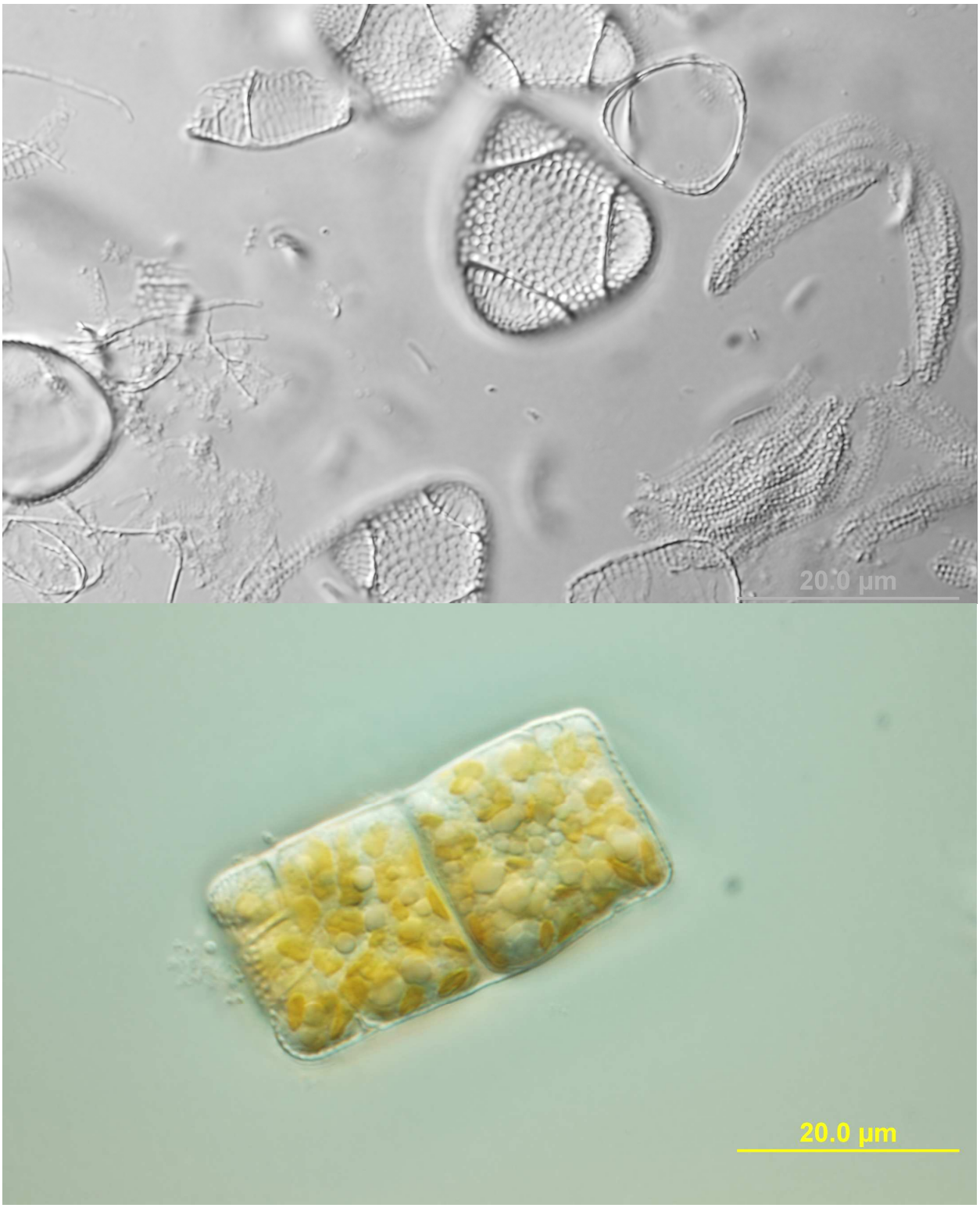

Figure 7d. *Biddulphia alternans* viewed with a light microscope under a 100x objective with oil; scale bar show is 20 µm in both the cleaned sample (top) and live culture (bottom)

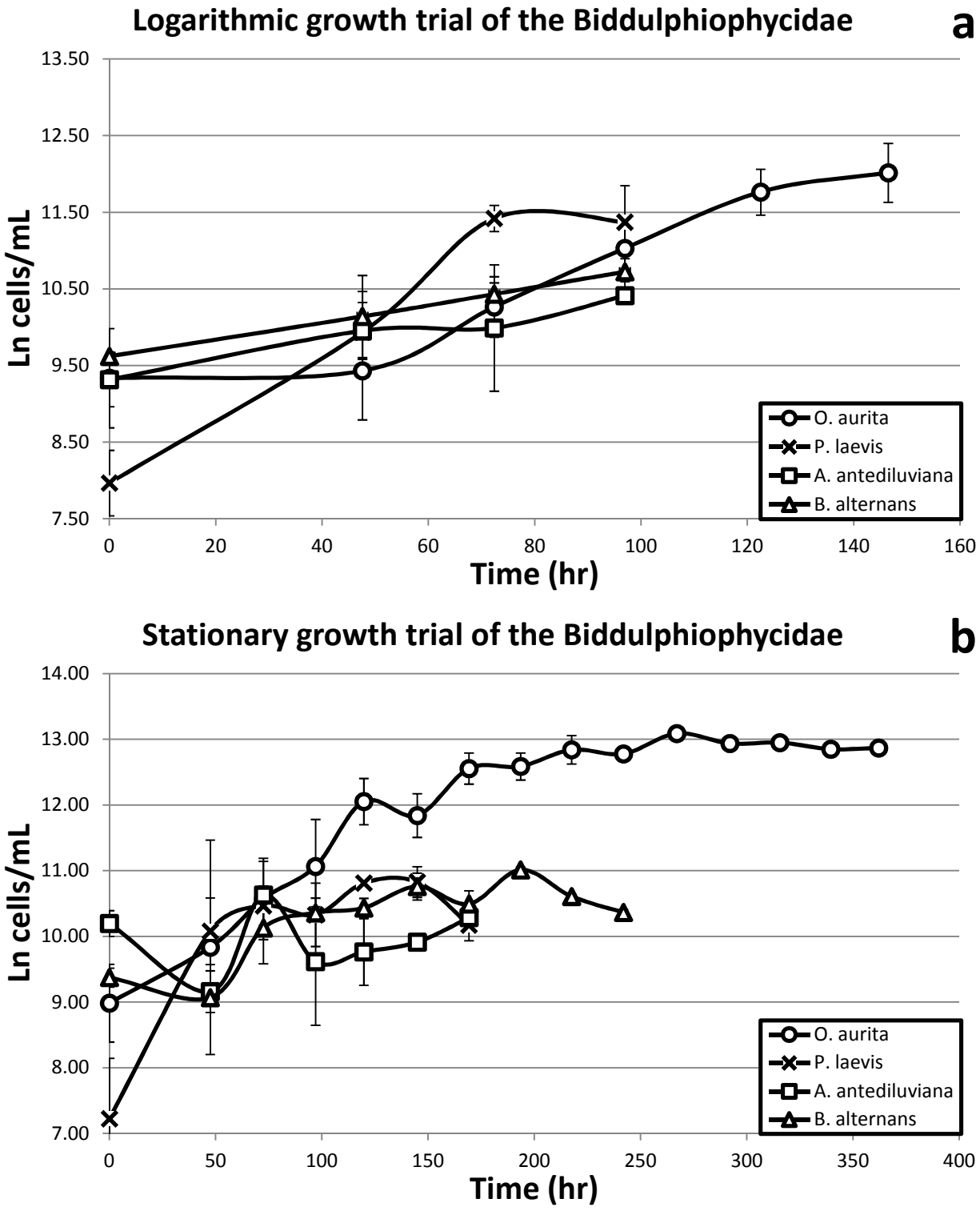

Figure 8. (a) Logarithmic and (b) stationary growth trials for the Coscinodiscophycidae; points are averages of the natural log of cell number per milliliter from three replicates and error bars denote the standard error

Table 6 Growth data for the Biddulphiophycidae

		Logarithmic trial			Stationary trial		
Species	Specific growth rate (μ max; d^{-1})	Trial length (h)	Final density (cells/mL)	Dry biomass (mg)	Trial length (h)	Final density (cells/mL)	Dry biomass (mg)
O. aurita	0.98	147	147,246 ± 48,225	40.90 ± 3.51	363	441,737 ± 57,671	90.86 ± 1.46
A. antediluviana	0.75	99	33,364 ± 1790	33.63 ± 1.90	170	29,420 ± 900	31.03 ± 0.89
P. laevis	1.65	96	104,525 ± 35,994	36.23 ± 0.87	170	10,566 ± 27,609	31.67 ± 7.72
B. alternans	0.60	96	49,082 ± 14,436	39.46 ± 2.26	241	31,759 ± 1,444	39.47 ± 2.26

Table 7 Cell volume, total extracted lipid mass, and lipid content on various bases for the Biddulphiophycidae									
		Logarithmic				Stationary			
Species	Cell volume (μm^3)	Lipid mass (mg)	% dry weight	µg/mL	pg/100 μm^3	Lipid mass (mg)	% dry weight	µg/mL	pg/100 μm^3
O. aurita	2,490 ± 378	2.13 ± 0.18	5.32 ± 0.75	9.65 ± 0.84	8.28 ± 6.03	12.47 ± 0.32	13.73 ± 0.39	54.93 ± 1.55	5.17 ± 0.69
A. antediuluviana	44,748 ± 4,710	2.27 ± 0.18	6.76 ± 0.51	9.44 ± 0.73	0.63 ± 0.03	2.50 ± 0.1	8.07 ± 0.45	11.24 ± 0.41	0.86 ± 0.06
P. laevis	28,894 ± 3,331	3.5 ± 0.11	9.68 ± 0.54	15.29 0.60	0.78 ± 0.40	2.90 ± 0.75	9.06 ± 0.43	12.89 ± 3.39	1.59 ± 0.24
B. alternans	2,971 ± 517	2.16 ± 0.32	5.45 ± 0.59	9.37 ± 1.36	7.34 ± 2.03	4.70 ± 0.40	11.04 ± 0.80	20.55 ± 1.71	21.74 ± 1.23

Table 8 *P* values for pairwise statistical comparisons by TukeyHSDs within the Biddulphiophycidae; * denotes values less than *P* = 0.05						
	Logarithmic			Stationary		
Species comparison	% dry weight	µg/mL	pg/100 μm^3	% dry weight	µg/mL	pg/100 µm3
B. alternans – A. antediluviana	0.463	0.999	0.487	0.020*	0.050*	$1.0x10^{-7}$*
O. aurita – A. antediluviana	0.389	0.998	0.385	$3.6x10^{-4}$*	$1.9x10^{-6}$*	0.011*
P. laevis – A. antediluviana	0.037*	0.009*	0.999	0.597	0.939	0.885
O. aurita – B. alternans	0.998	0.996	0.996	0.033*	$1.2x10^{-5}$*	$1.0x10^{-6}$*
P. laevis – B. alternans	0.005*	0.009*	0.504	0.121	0.112	$2.0x10^{-7}$*
P. laevis – O. aurita	0.004*	0.011*	0.400	0.001*	$2.5x10^{-6}$*	0.031*

stationary trials. During logarithmic growth, the lipid content of *P. laevis* was significantly higher than *O. aurita*, *A. antediluviana,* and *B. alternans* when comparing percent dry weight (P = 0.004, P = 0.037, P = 0.005; Figure 9a) and lipid mass per culture volume (µg/mL) (P = 0.011, P = 0.009, P = 0.009; Figure 9b). There was no significant difference between any species when comparing lipid mass per cell volume (pg/100 μm^3) (Figure 9c).

During stationary growth, the lipid content of *O. aurita* was significantly higher than *P. laevis*, *A. antediluviana,* and *B. alternans* when comparing percent dry weight (P = 0.001, $P = 3.6 \times 10^{-4}$, P = 0.033; Figure 10a) and *B. alternans* was significantly different from *A. antediluviana* (P = 0.020; Figure 10a). The same pattern arose when comparing lipid mass per culture volume (µg/mL), in which *O. aurita* was significantly higher than *P. laevis*, *A. antediluviana,* and *B. alternans* ($P = 2.5 \times 10^{-6}$, $P = 1.9 \times 10^{-6}$, $P = 1.2 \times 10^{-5}$; Figure 9b) and *B. alternans* was significantly different from *A. antediluviana* (P = 0.050; Figure 9b). When comparing lipid mass per cell volume (pg/100 μm^3), *B. alternans* was significantly higher *O. aurita*, *P. laevis*, and *A. antediluviana* ($P = 1.0 \times 10^{-6}$, $P = 2.0 \times 10^{-7}$, $P = 1.0 \times 10^{-7}$; Figure 9c); *O. aurita* was significantly higher than *P. laevis* and *A. antediluviana* (P = 0.031, P = 0.011; Figure 9c).

3.3 Fragilariophycidae

Examination with a light microscope resulted in the identification of *Diatoma tenuis* Agardh, *Fragilariforma nitzschioides* (Grunow) Lange-Bertalot, and *Synedra acus* Ehrenberg as the three taxa examined within the Fragilariophycidae (see images in Figure 10a-c). Comparing the growth patterns shown in Figure 11, *F. nitzschioides* and *S. acus* were most

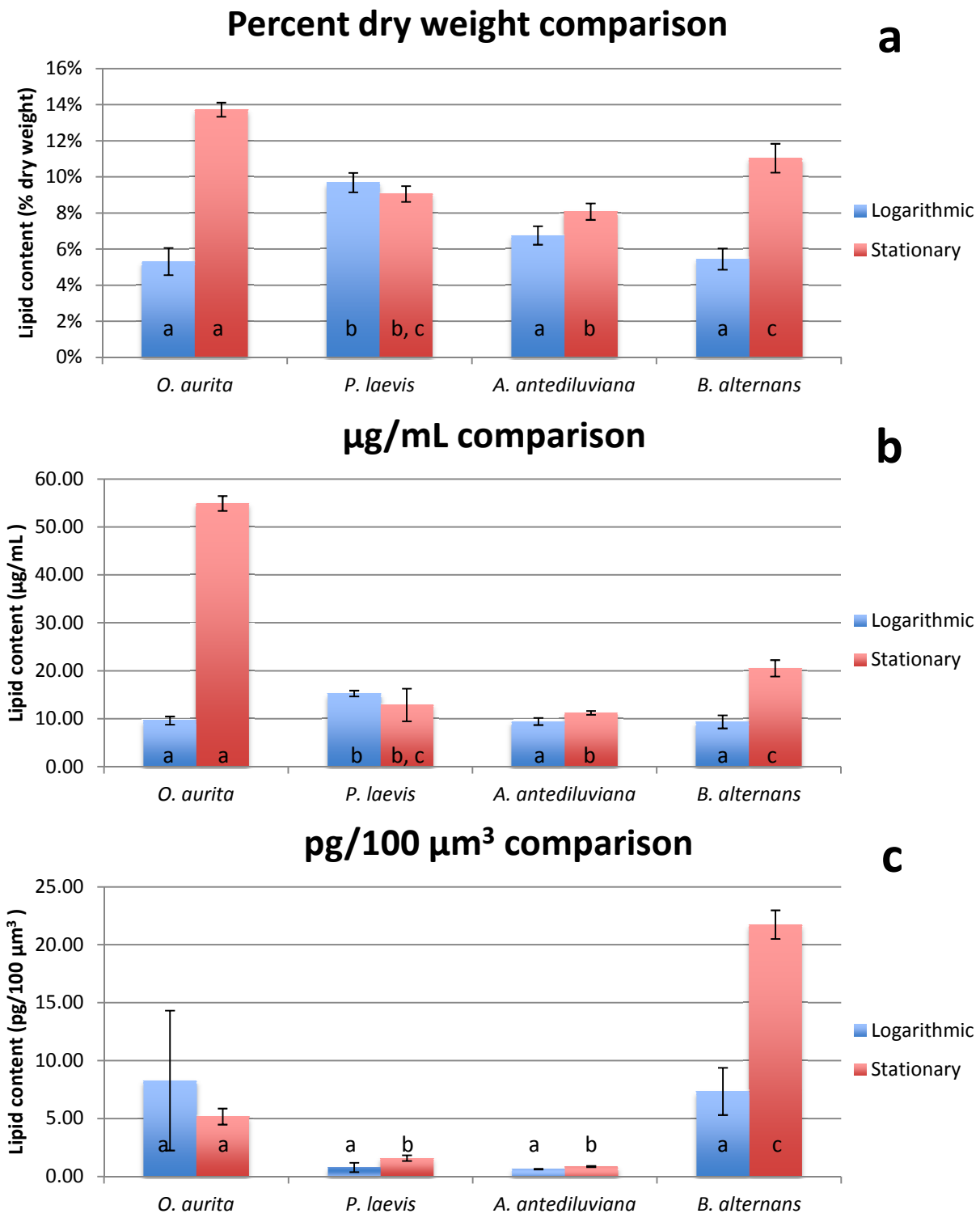

Figure 9. Total lipid accumulation during the logarithmic trial (blue bars) and stationary trial (red bars) quantified as (a) percent dry weight, (b) lipid mass per culture volume in µg/mL, and (c) lipid mass per cell volume in pg/100µm^3; Lowercase letters indicate significant differences within a trial (same letter = no significant difference; different letter = significant difference)

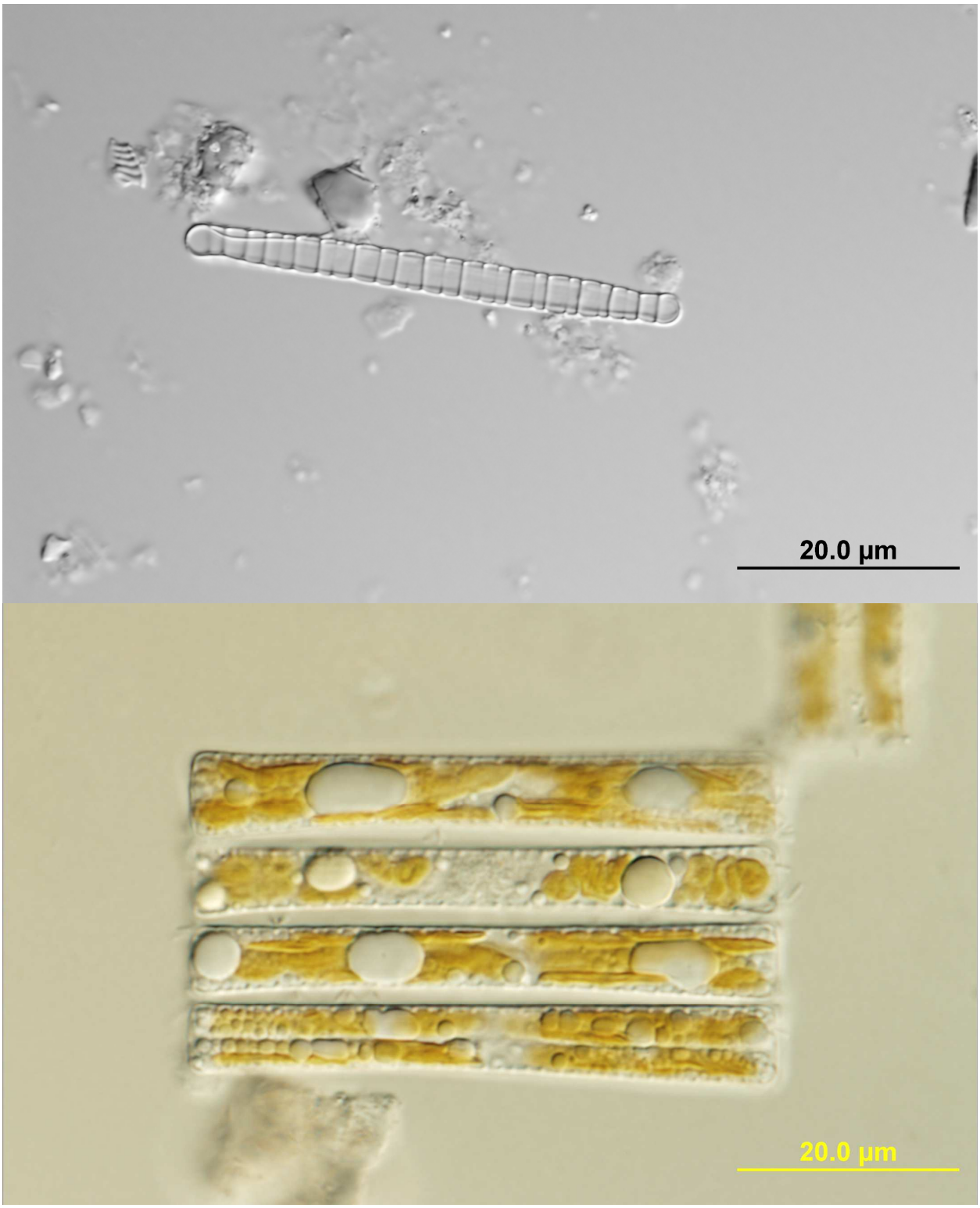

Figure 10a. *Diatom tenuis* viewed with a light microscope under a 100x objective with oil; scale bar show is 20 µm in both the cleaned sample (top) and live culture (bottom)

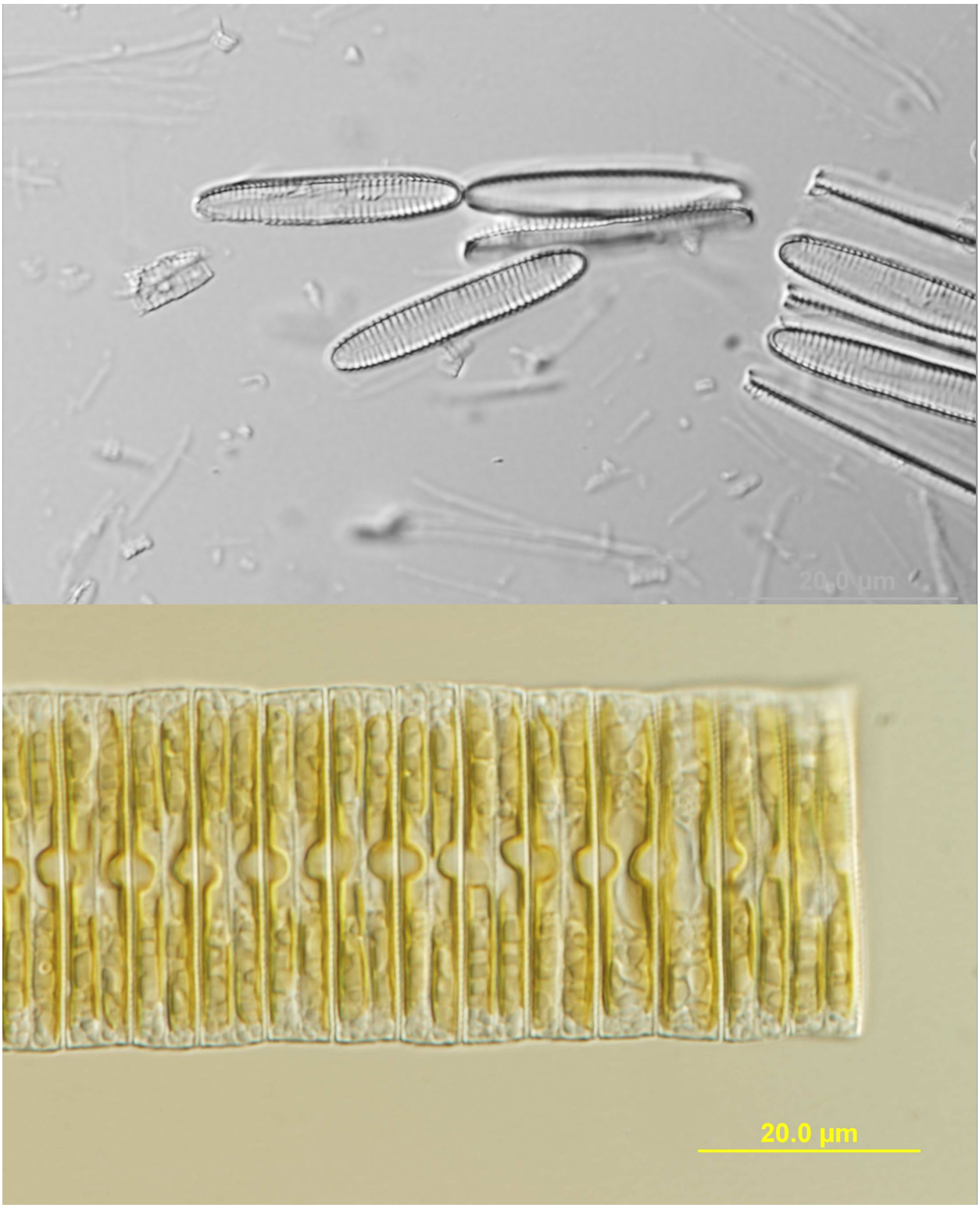

Figure 10b. *Fragilariforma nitzschioides* viewed with a light microscope under a 100x objective with oil; scale bar show is 20 µm in both the cleaned sample (top) and live culture (bottom)

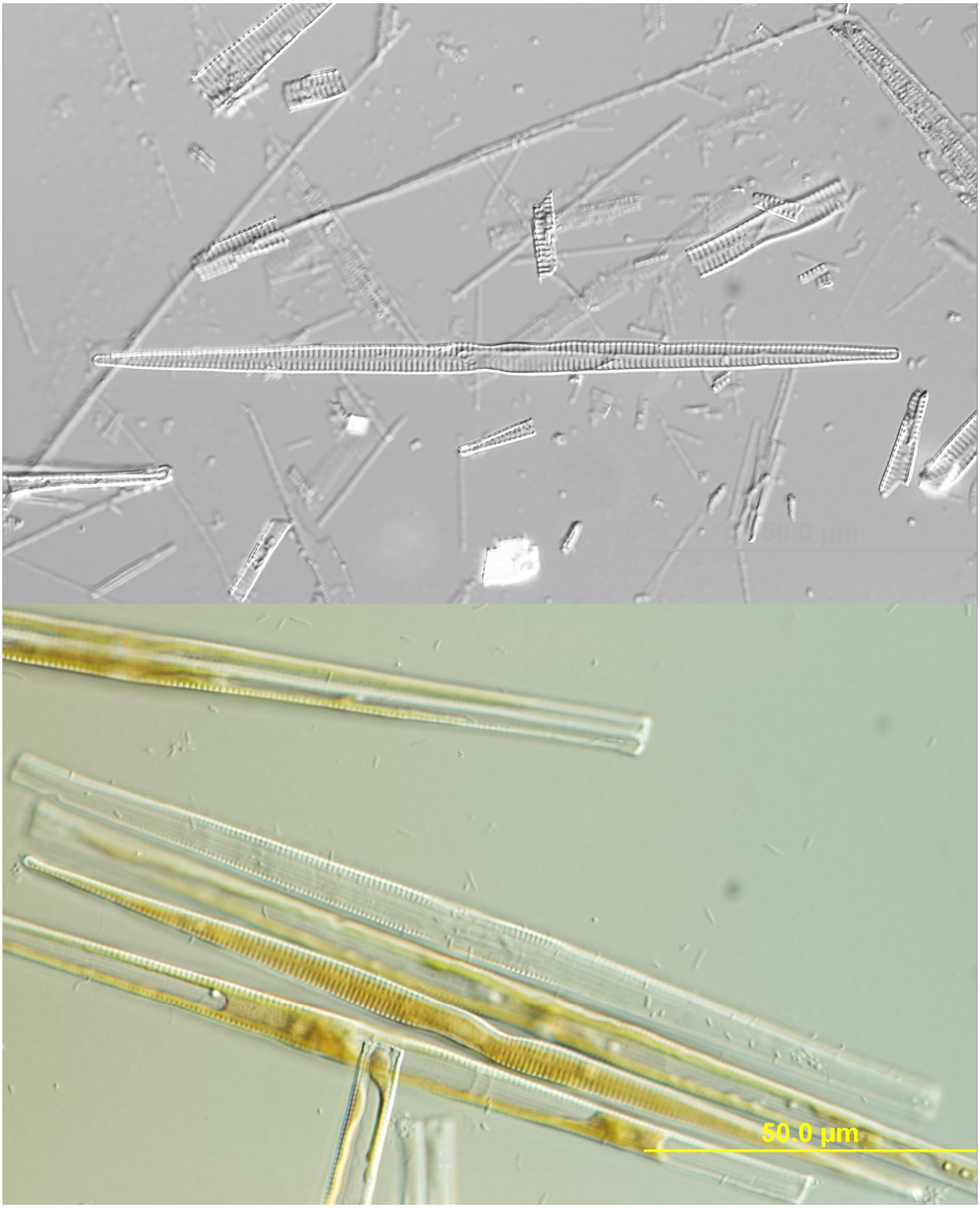

Figure 10c. *Synedra acus* viewed with a light microscope under a 60x objective with oil; scale bar show is 50 µm in both the cleaned sample (top) and live culture (bottom)

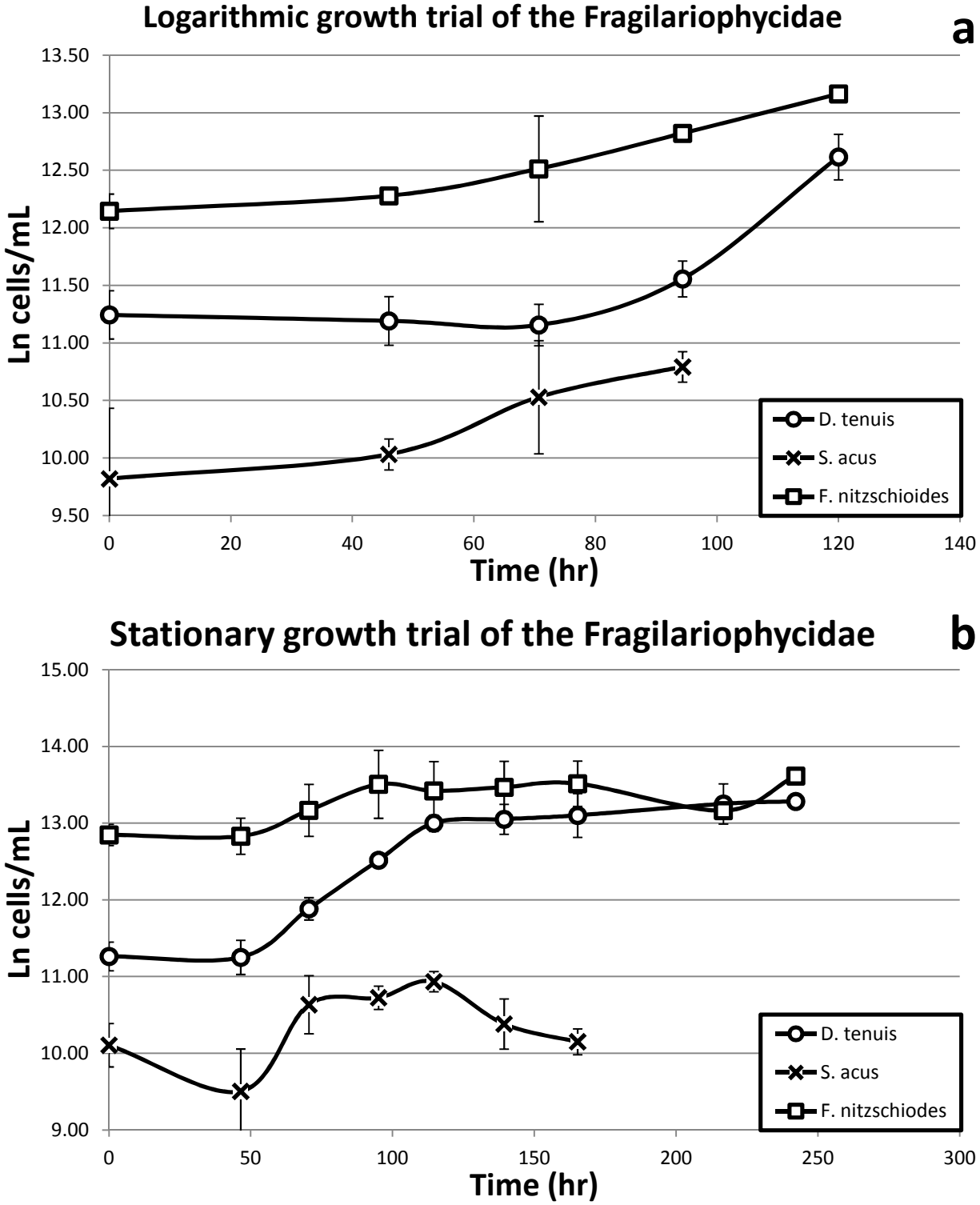

Figure 11. (a) Logarithmic and (b) stationary growth trials for the Coscinodiscophycidae; points are averages of the natural log of cell number per milliliter from three replicates and error bars denote the standard error

similar in the logarithmic trial while *D. tenuis* and *F. nitzschioides* were most similar in the stationary trial. *S. acus* exhibited the highest specific growth rate, followed by *D. tenuis,* and then *F. nitzschiodies* (Table 9)*. Fragilariforma nitzschioides* and *S. acus* produced a similar amount of biomass during the logarithmic trial, while *D. tenuis* produced the least. However, in the stationary trial *D. tenuis* produced the most biomass followed by *F. nitzschioides* and then *S. acus* (Table 9). Cell volume increased from 360 μm^3 in *F. nitzschioides* to 592 μm^3 in *D. tenuis* and 3,310 μm^3 in *S. acus* (Table 10).

Table 11 shows the statistical outcomes of the resulting TukeyHSD analyses performed for all three metrics within the Fragilariophycidae for both the logarithmic and stationary trials. During logarithmic growth, percent lipid per dry weight in *F. nitzschioides* was significantly higher than in *D. tenuis* (P = 0.065; Figure 12a), with no significant difference in lipid content of *S. acus* versus either of the latter species (Figure 12a). Lipid mass per culture volume (µg/mL) was not significantly different between any of the species (Figure 12b). Lipid mass per culture cell volume (pg/100 μm^3) was significantly lower in *D. tenuis* compared to *F. nitzschioides* and *S. acus* (P = 0.001, $P = 5.7 \times 10^{-4}$; Figure 12c).

During stationary growth, percent lipid per dry weight was significantly higher in *D. tenuis* than in *F. nitzschioides* (P = 0.035; Figure 12a), while lipid content expressed in *S. acus* was not significantly different from that of the either of the former species (Figure 12a). Lipid mass per culture volume (µg/mL) was significantly higher in *D. tenuis* versus the other species, and was lowest in *S. acus* (Figure 12b). Lipid mass per algal cell volume (pg/100 μm^3) was not significantly different between the species (Figure 12c).

Table 9 Growth data for the Fragilariophycidae

		Logarithmic trial			Stationary trial		
Species	Specific growth rate (μ max; d^{-1})	Trial length (h)	Final density (cells/mL)	Dry biomass (mg)	Trial length (h)	Final density (cells/mL)	Dry biomass (mg)
D. tenuis	0.39	120	312, 488 ± 54,798	13.70 ± 1.65	242	587, 477 ± 8,461	42.17 ± 3.41
F. nitzschioides	0.34	120	522,777 ± 26,446	17.60 ± 1.00	242	818,120 ± 10,215	35.20 ± 1.96
S. acus	0.57	91	49,447 ± 6,268	17.33 ± 1.26	164	37,533 ± 11,816	16.13 ± 1.52

Table 10 Cell volume, total extracted lipid mass, and lipid content on various bases for the Fragilariophycidae									
		Logarithmic				Stationary			
Species	Cell volume (μm^3)	Lipid mass (mg)	% dry weight	µg/mL	pg/100 μm^3	Lipid mass (mg)	% dry weight	µg/mL	pg/100 μm^3
D. tenuis	592 ± 51	2.17 ± 0.30	15.76 ± 0.51	9.53 ± 1.34	5.26 ± 0.41	9.50 ± 0.69	22.56 ± 0.18	41.28 ± 2.85	11.85 ± 0.65
F. nitzschioides	360 ± 33	3.57 ± 0.26	20.25 ± 0.80	15.82 ± 1.38	8.38 ± 0.33	6.83 ± 0.43	19.41 ± 0.50	30.08 ± 2.28	10.21 ± 0.73
S. acus	3,310 ± 417	3.40 ± 0.52	19.39 ± 1.67	14.72 ± 2.11	8.96 ± 0.25	3.46 ± 0.41	21.42 ± 1.01	15.01 ± 1.61	14.01 ± 3.18

Table 11 *P* values for pairwise statistical comparisons by TukeyHSDs within the Fragilariophycidae; * denotes values less than *P* = 0.05						
	Logarithmic			Stationary		
Species comparison	% dry weight	µg/mL	pg/100 μm^3	% dry weight	µg/mL	pg/100 μm^3
F. nitzschioides – D. tenuis	0.065*	0.078	0.001*	0.035*	0.032*	0.821
S. acus – D. tenuis	0.130	0.144	$5.7x10^{-4}$*	0.487	$4.7x10^{-4}$*	0.720
S. acus – F. nitzschioides	0.852	0.884	0.480	0.159	0.009*	0.397

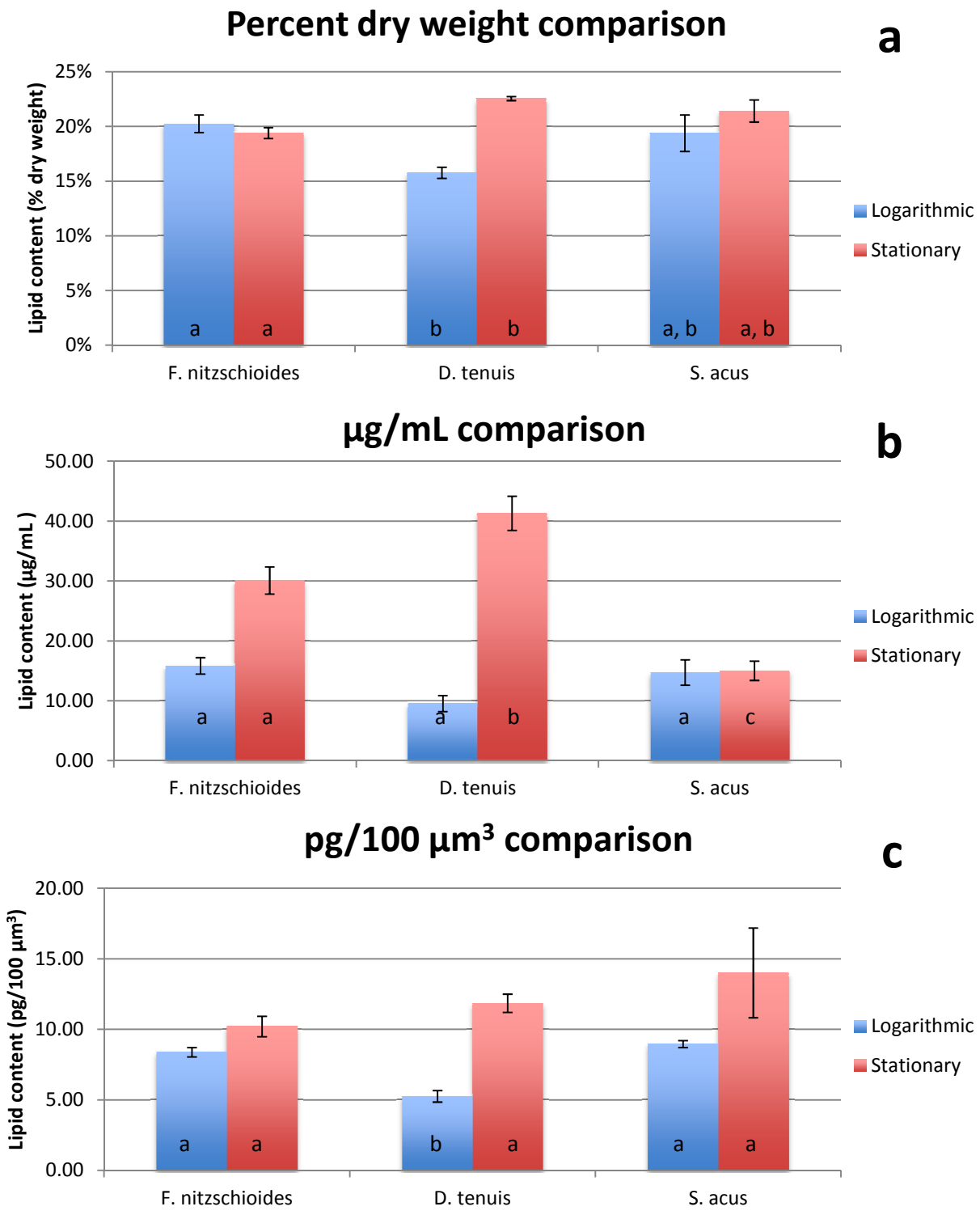

Figure 12. Total lipid accumulation during the logarithmic trial (blue bars) and stationary trial (red bars) quantified as (a) percent dry weight, (b) lipid mass per culture volume in µg/mL, and (c) lipid mass per cell volume in pg/100µm^3; Lowercase letters indicate significant differences within a trial (same letter = no significant difference; different letter = significant difference)

3.4 Comparison between logarithmic and stationary trials

Percent lipid per dry weight in *M. nummuloides, O. aurita,* and *D. tenuis* was significantly higher during stationary versus logarithmic growth (Figure 13a, Table 12). A similar pattern emerged for lipid mass per culture volume (μg/mL), with *F. nitzschioides* also exhibiting significantly higher values during station growth (Figure 13b, Table 12). *Actinoptychus senarius* and *B. alternans* were the only two species that exhibited a significantly greater lipid mass per cell volume (pg/100 μm^3) in the stationary phase versus logarithmic phase (Figure 13c, Table 12).

3.5 Phylogenetic comparison using a balanced nested ANOVA

A balanced nested ANOVA was used to compare within-lineage variability to between-lineage variability. Three species within each lineage were used in this analysis; *O. aurita, P. laevis,* and *A. antediluviana* were selected as the three taxa from the Biddulphiophycidae lineage to be included due to their close phylogenetic relationship (Ashworth et al., 2013). *Biddulphia alternans* was also used to test the robustness of the conclusions drawn from the nested ANOVA.

As shown in Table 13, percent lipid per dry weight during stationary growth was the only parameter responding as predicted, i.e., showing a greater variability between lineages ($P = 5.1 \times 10^{-4}$) than within lineages ($P = 0.020$). Percent lipid per dry weight was significantly lower in the Biddulphiophycidae lineage than in the Coscinodiscophycidae and the Fragilariophycidae lineages ($P = 1.0 \times 10^{-8}$; $P = 1.0 \times 10^{-8}$; Figure 13a). All other

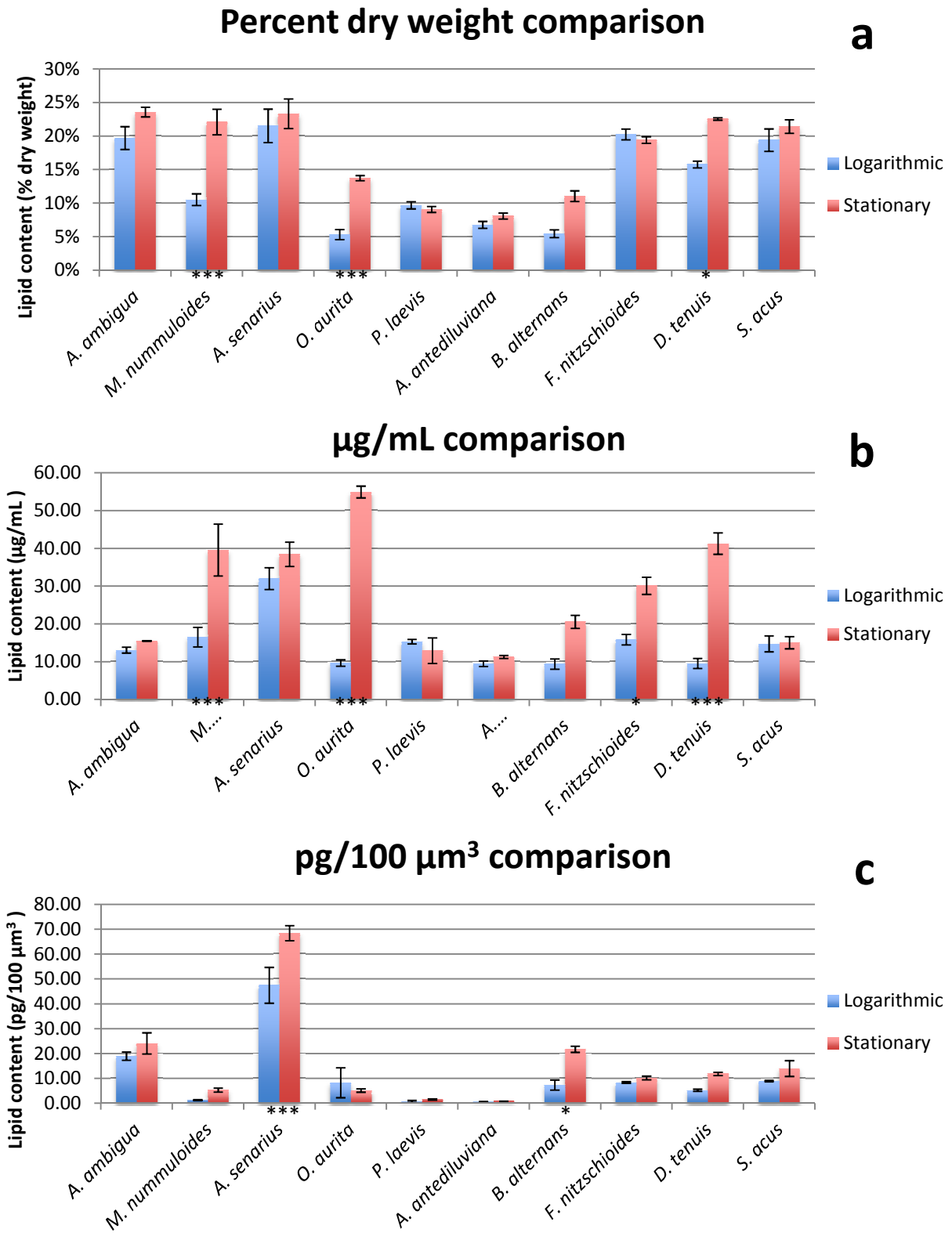

Figure 13. Total lipid accumulation during the logarithmic trial (blue bars) and stationary trial (red bars) quantified as (a) percent dry weight, (b) µg/mL, and (c) pg/100µm^3; * indicate significance level (* < 0.05, ** < 0.01, *** < 0.001)

Table 12 *P* values from the resulting statistical comparisons between logarithmic and stationary growth; * denotes values less than *P* = 0.05			
Species	% dry weight	µg/mL	pg/100 µm^3
A. ambigua	0.674	0.999	0.997
M. nummuloides	2.3x10^{-6}*	7.5x10^{-6}*	0.999
A. senarius	0.999	0.928	2.2x10^{-4}*
O. aurita	0.001*	1.0x10^{-8}*	0.999
P. laevis	1.000	0.999	1.000
A. antediluviana	0.999	1.000	1.000
B. alternans	0.116	0.169	0.039*
F. nitzschioides	1.000	0.019*	1.000
D. tenuis	0.018*	1.0x10^{-8}*	0.953
S. acus	0.999	1.000	0.997

Table 13 *P* values from the resulting nested ANOVA; * denotes where the variability between lineages is greater than the variability within lineages						
	Logarithmic			Stationary		
Lineage	% dry weight	µg/mL	pg/100 µm^3	% dry weight	µg/mL	pg/100 µm^3
Between	0.023	0.268	0.270	5.1x10^{-4}*	0.948	0.215
Within	1.0x10^{-4}	4.5x10^{-6}	8.4x10^{-7}	0.020	5.7x10^{-4}	5.2x10^{-12}

parameters exhibited a greater variability within lineages than between lineages (Table 13).

3.6 Comparing *B. alternans* to the Biddulphiophycidae

Biddulphia alternans was compared to the combined group of *O. aurita, P. laevis,* and *A. antediluviana* to further assess the above result that there was less variability within than between lineages in percent lipid per dry weight during stationary growth. Average lipid content expressed as percent dry weight of *B. alternans* compared to the combined average of *O. aurita, P. laevis,* and *A. antediluviana* during stationary growth was indeed not significantly different ($P = 0.550$; Figure 14).

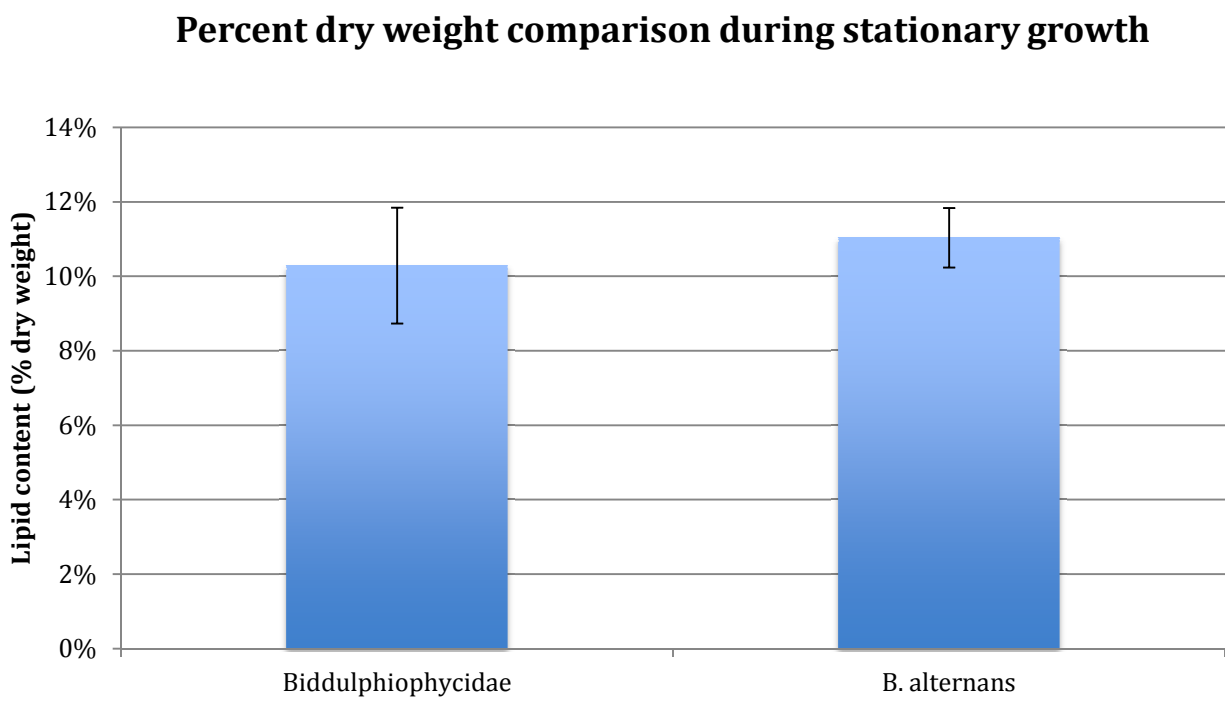

Figure 14. The combined average lipid accumulation of *O. aurita, P. laevis,* and *A. antediluviana* (left column) compared to *B. alternans* (right column) reported as percent dry weight during stationary growth; there is no significant difference between the means

3.7 Phylogenetic analysis of all data

Data on percent lipid per dry weight as generated in the study (Table 4, 7, 10) were paired with the data published in the literature (Table 1) to test for a phylogenetic signal across the entire diatom phylogeny (Blomberg et al., 2003). The combined data were analyzed on three different taxonomic levels, i.e., subclass (Table 14), order (Table 15), and genus (Table 16). A phylogenetic tree was assembled for each taxonomic level using 18S ribosomal RNA sequences from GenBank (Figure 15, 16, 17).

Table 17 shows the resulting *P* and *K* values from the phylosignal test run in RStudio. A statistically significant phylogenetic signal was only detected when examining the genus-level phylogeny with lipid data generated during nutrient-deplete (i.e. stationary growth) experiments ($P = 0.013$; Table 17). It should be noted that due to gaps in the data on the genus-level (Table 16), the genera *Caloneis* and *Seminavis* removed from the phylogenetic tree for the nutrient-replete analysis and the genera *Achnanthes, Diploneis, Fragilaria, Gyrosigma, Pinnularia, Pleurosigma,* and *Tryblionella* were removed from the nutrient-deplete analysis.

The genus-level analysis of nutrient-deplete data yielded the highest *K* value, meaning that it had the strongest phylogenetic signal out of all the analyses, further supporting the results from the randomization test (Table 17).

Table 14 Branch-tip data (% lipid per dry weight) for the subclass level analysis			
Subclass	Nutrient replete	Nutrient deplete	GenBank accession
Bacillariophycidae	22.91	38.58	HQ912589
Biddulphiophycidae	7.12	10.47	HQ912564
Chaetocerotophycidae	29.25	37.38	HQ912558
Coscinodiscophycidae	17.37	24.00	HQ912606
Fragilariophycidae	17.64	21.13	HQ912622
Thalassiosirophycidae	17.43	37.26	HQ912555

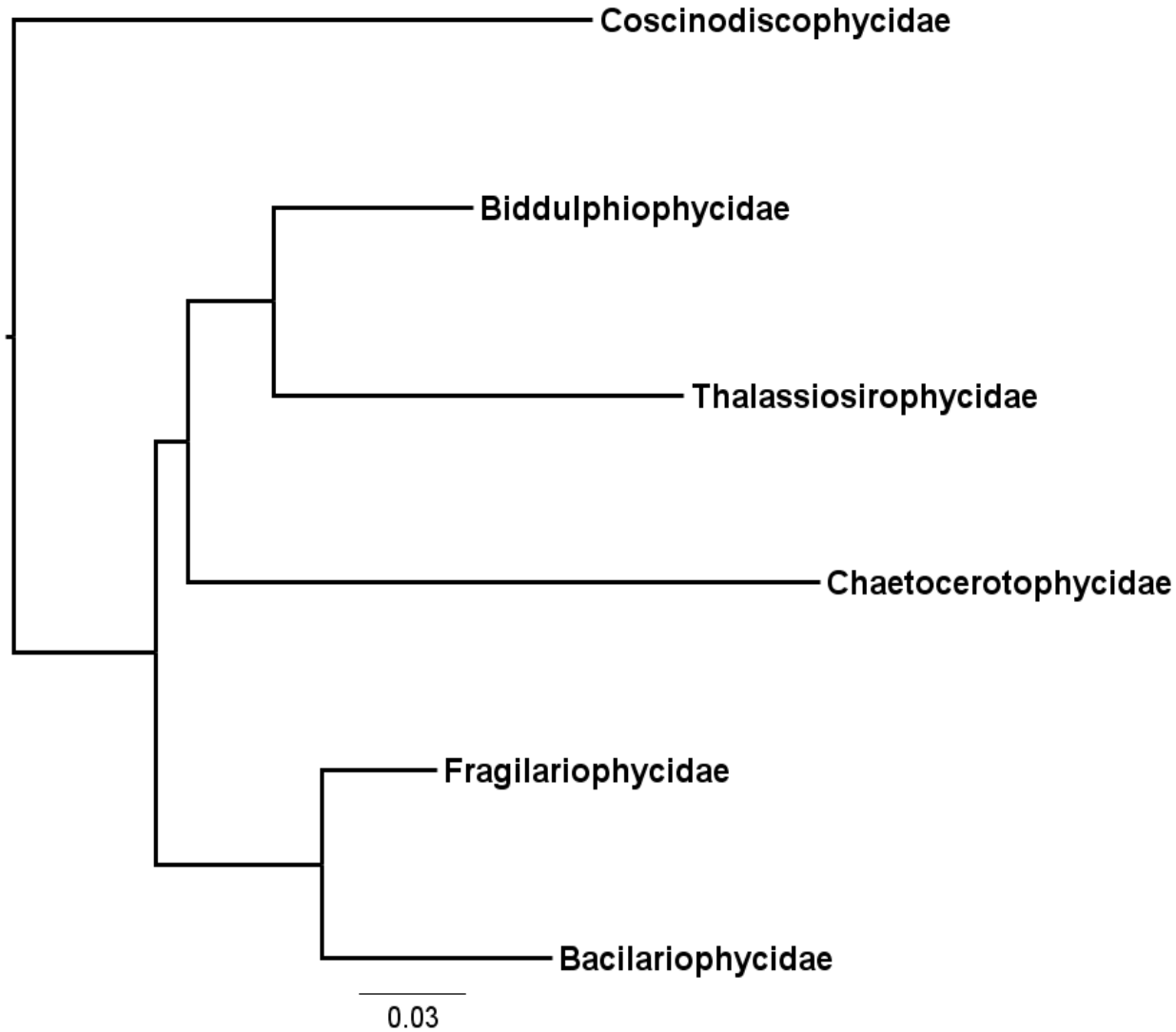

Figure 15. Maximum likelihood phylogeny of the diatom subclasses inferred from 18S rRNA sequences taken from GenBank. This tree was used to test for a phylogenetic signal among the diatom subclasses by pairing the data in Table 14 to the corresponding branch-tip

Table 15 Branch-tip data (% lipid per dry weight) for the order level analysis			
Order	Nutrient replete	Nutrient deplete	GenBank accession
Achnanthales	23.98	30.24	HQ912594
Aulacoseirales	19.70	23.57	HQ912606
Bacillariales	23.12	43.51	HQ912589
Biddulphiales	5.45	11.04	HQ912677
Chaetocerotales	29.25	37.38	HQ912558
Coscinodiscales	21.52	23.33	HQ912640
Fragilariales	17.64	21.13	HQ912622
Melosirales	12.52	24.82	HQ912566
Naviculales	21.85	36.68	HQ912603
Thalassiophysales	25.21	35.26	HQ912602
Thalassiosirales	17.43	37.26	HQ912555
Triceratiales	7.63	10.29	HQ912564

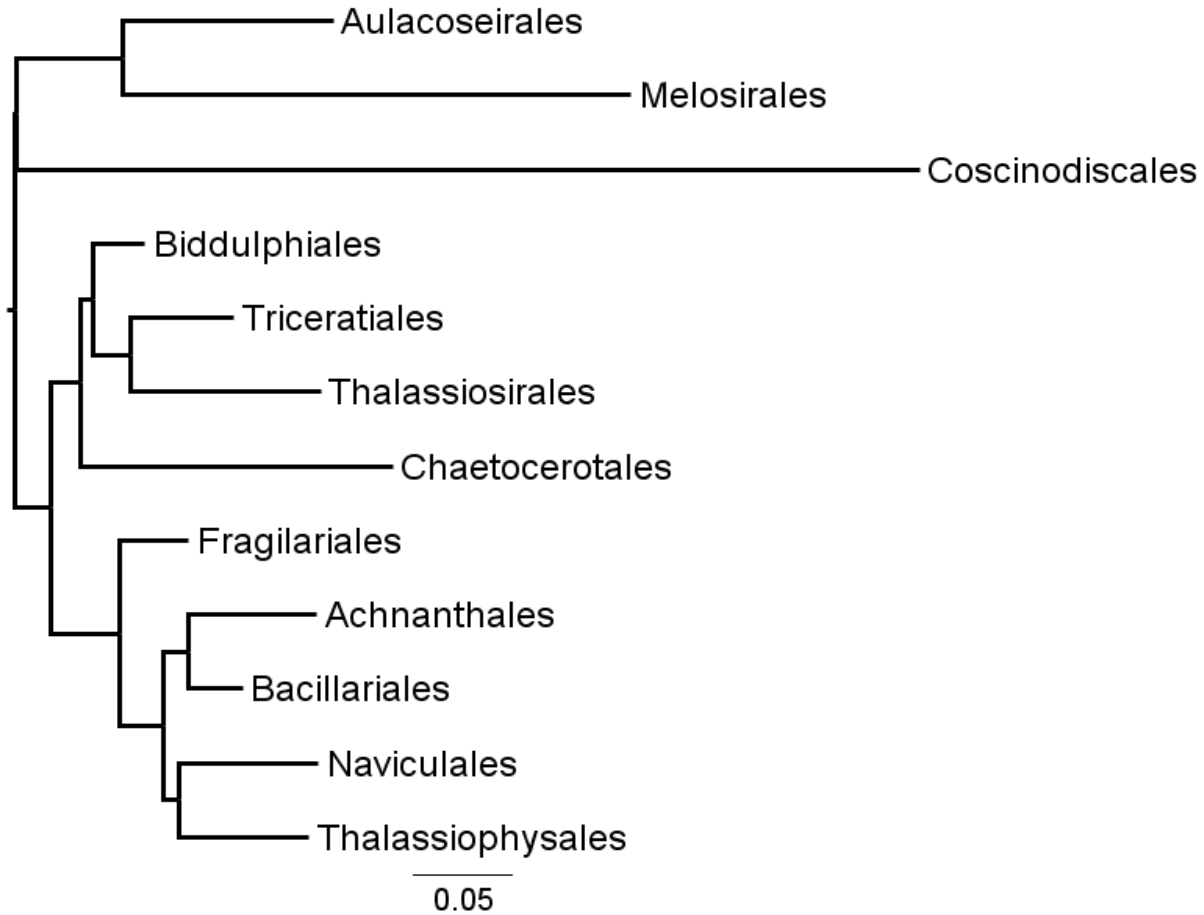

Figure 16. Maximum likelihood phylogeny of the diatom orders inferred from 18S rRNA sequences taken from GenBank. This tree was used to test for a phylogenetic signal among the diatom orders by pairing the data in Table 15 to the corresponding branch-tip

Table 16 Branch-tip data (% lipid per dry weight) for the genus level analysis; n.a. = no data available			
Genus	Nutrient replete	Nutrient deplete	GenBank ID
Achnanthes	25.00	n.a.	AJ535151
Actinoptychus	21.52	23.33	AJ535182
Amphitetras	6.76	8.07	JQ315688
Amphora	24.29	35.79	AB754837
Aulacoseira	19.70	23.57	X85404
Biddulphia	5.45	11.04	JX401229
Caloneis	n.a.	38.39	AM501954
Chaetoceros	32.77	37.38	AY625896
Cocconeis	20.90	30.24	AM502010
Cyclotella	14.10	41.57	DQ514855
Cylindrotheca	18.15	36.00	KC899347
Diatoma	15.76	22.56	AM497731
Diploneis	24.10	n.a.	HQ912597
Fragilaria	10.20	n.a.	AM712616
Fragilariforma	20.25	19.41	AJ535137
Gyrosigma	29.00	n.a.	HQ912598
Hantzschia	29.00	59.50	HQ912404
Melosira	14.51	27.56	X85402
Navicula	20.76	40.12	AY485484
Nitzschia	22.73	40.12	KF959653
Odontella	8.16	13.73	EU818943
Phaeodactylum	17.64	26.93	AJ269501
Pinnularia	19.40	n.a.	JN418576
Pleurosigma	33.50	n.a.	AY485514
Pleurosira	9.68	9.06	AJ535188
Seminavis	n.a.	34.99	FJ624252
Skeletonema	19.33	37.62	JN676163
Synedra	19.39	21.42	AM497723
Thalassiosira	19.42	24.00	HM991698
Tryblionella	24.20	n.a.	HQ912600

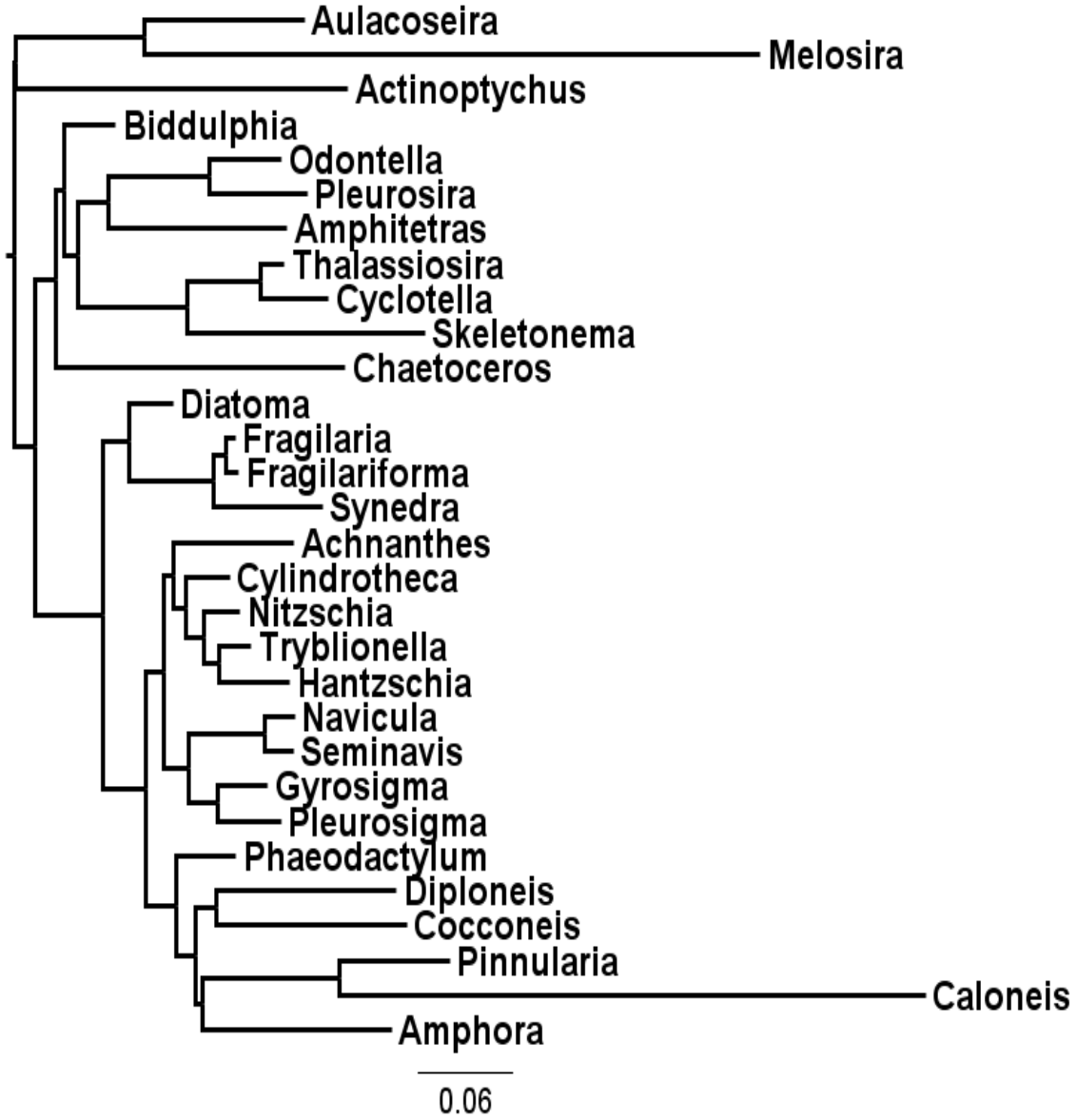

Figure 17. Maximum likelihood phylogeny of the diatom genera inferred from 18S rRNA sequences taken from GenBank. This tree was used to test for a phylogenetic signal among the diatom genera by pairing the data in Table 16 to the corresponding branch-tip

Table 17 *P* and *K* values resulting from the phylogenetic signal tests; * indicate a significant phylogenetic signal				
	Logarithmic		Stationary	
	P	*K*	*P*	*K*
Subclass	0.453	0.814	0.783	0.600
Order	0.209	0.795	0.717	0.604
Genus	0.113	0.641	0.013*	0.962

4. DISCUSSION

4.1 Comparison between growth stages and the impact of the reference basis used to express lipid

In the present study, lipid content of ten species of diatoms from three lineages under-reported in the literature was quantified during exponential growth and stationary growth; the latter induced though silica deprivation. Among the ten diatom species examined, cultures contained 5-22% of their dry weight as lipids in the exponential growth phase (i.e. nutrient-replete) and 8-24% in the stationary growth phase (i.e. nutrient-deplete).

Griffiths (2009) states that lipid content reported under silica-deplete conditions were on average 168% greater than cultures grown in silica-replete media. The majority of species examined in the present study seem to contradict this claim; only three species showed a significantly greater increase in lipid content as percent dry weight in the stationary phase, i.e., *M. nummuloides* (210% increase), *O. aurita* (258% increase), and *D. tenuis* (143% increase). These findings may be the result of differences in silicification, that is, in turn, dependent on cell size and the availability of silica in the medium (Theriot, 1987). Theriot (1987) observed that larger cells of the diatom *Stephanodiscus* were more heavily silicified than smaller ones, and that cells in waters with a high silica:phosphorus (Si:P) ratios were more heavily silicified than similarly-sized cells in waters with lower Si:P ratios. These features would cause the dry weight of larger cells, such as *A. antediluviana,* to be over-proportionally greater and the resulting lipid content per dry weight to be lower compared to smaller cells, such as *A. ambigua*. Additionally, varying levels of silicification

dependent on silica concentrations would shift the computed percent lipid per dry weight values down during exponential growth (i.e. silica-replete) and up during stationary growth (i.e. silica-deplete) – even without any actual differences in lipid accumulation between these two growth stages.

As observed during certain trials in this study, expressing lipid content per dry weight may yield misleading results due to inclusion of dead cells, extracellular compounds, and microbial contaminants in the calculation of total dry biomass. Inclusion of the latter components will increase the dry biomass without an effect on the quantity of lipids produced, thus resulting in an erroneously low lipid content per dry weight. Species such as *P. laevis* and *S. acus* exhibited a large amount of cell death at the end of the stationary phase (Figures 8b & 11b), and lipid content per dry weight computed for these species is thus likely underestimated. To gain a better understanding of lipid accumulation in relation to algal biology, alternative metrics such as lipid mass per culture volume or per algal cell volume should yield more meaningful estimates.

4.6 The relationship between specific growth rate and lipid content

Sheehan et al. (1998) report that a high algal biomass productivity and high lipid accumulation are mutually exclusive. The taxa examined in the present study revealed an inverse correlation between lipid content and μ max, further supporting the findings reported in Sheehan et al. (1998). For example, *A. senarius* had the highest lipid content per dry weight during exponential growth (21.52% dry weight) and a low μ max of 0.57 d^{-1}. In contrast, *P. laevis* had a high μ max of 1.65 d^{-1} and low lipid content per dry weight (9.68%

dry weight) during exponential growth. Therefore, if implemented into a commercial facility, *A. senarius* may produce a higher portion of lipids but suffer due to low biomass productivity, whereas *P. laevis* may produce a large amount of biomass but only a small portion would contain lipids. In efforts to maximize lipid production, a balance of high lipid content and a high algal growth rate is sought after by the commercial industry (Sheehan et al., 1998; Rodolfi et al., 2008). *Aulacoseira ambigua* exhibited the best apparent compromise between these two traits, with a relatively high lipid content of 19.7% dry weight and a reasonably high growth rate (μ max) of 0.85 d^{-1} during exponential growth. Therefore, we would recommend further analyses into this genus and other taxa within the Aulacoseirales to identify highly productive species for commercial production.

4.2 Phylogenetic comparison (balanced nested ANOVA)

The focus of the present study was to compare the variability of diatom lipid accumulation within lineages to the variability between lineages. Greater between-lineage variability versus within-lineage variability provides a phylogenetic signal in support of the hypothesis that lipid accumulation is a heritable, evolutionary trait.

Overall, there was no clear significant evidence to conclude that differences observed between species during exponential growth were due to their evolutionary differences. However, lipid accumulation per dry weight during stationary growth exhibited greater variability between versus within lineages, with the Biddulphiophycidae lineage accumulating significantly lower amounts of lipids compared to the

Coscinodiscophycidae and Fragilariophycidae lineages, which is evidence for a possible phylogenetic signal.

Due to the fact that cell size varies within these lineages, likely resulting in differing silicification rates (as discussed above), as well as the inclusion of non-lipid-producing biomass (e.g., dead algal cells, bacterial contaminants), there are unavoidable inaccuracies when quantifying lipids as percent dry weight. Since neither lipid content per culture volume or per cell volume during exponential growth nor lipid content during stationary growth differed between lineages (more than within lineages), there is not enough evidence to conclude that lipid accumulation in diatoms is a heritable, evolutionary trait on this scale.

4.5 Phylogenetic analysis of all data

When tested for a phylogenetic signal using the methods outlined by Blomberg et al. (2003), the combined data from this study and others showed a statistically significant phylogenetic signal only when examining the genus-level phylogeny with lipid data (percent lipid per dry weight) generated during nutrient-deplete (i.e. stationary growth) experiments. Blomberg et al. (2003) state that, in order to have a robust test of a phylogeny, a minimum of twenty taxa must be present in the analysis. Here, less than twenty taxa for the subclass and order-level phylogenies and only 23 taxa were included in the genus-level analysis, which is likely why there was only a signal detected in the genus-level analysis. Although it will never be possible to have twenty taxa at the subclass-level since there are only eleven subclasses (Round et al., 1990), we decided to include the subclass phylogeny

in this study to compare this approach to the nest ANOVA analysis, which was also conducted on the subclass-level. As more data from diverse taxa are generated in the following years, they can be added to this dataset and lead to a more concrete analysis on the order and genus-level. Figure 18 shows the data generated in this study combined with the data from the literature and organized by order. The lipid content within the orders corresponds to either nutrient-replete or nutrient-deplete experiments and is plotted next to a cladogram adapted from Sims et al. (2006), as was Figure 1. Figure 18 shows which diatom orders have been analyzed and which ones future research should focus on to gain a more complete understanding of the potential lipid accumulation across the diatom phylogeny.

4.4 Evaluating former hypotheses and directing future research

Graham et al. (2011) proposes that the shape of a diatom cell can predict the potential lipid accumulation due to the effect cell shape has on the surface-to-volume ratio of a cell. They suggest that cylindrical cells will accumulate more lipids than fusiform (i.e., spindle-shaped) cells because a cylinder shape more efficiently contains a large volume. In the present study, the taxa examined had a variety of shapes and the reported lipid content would suggest that there is no correlation between cell shape and lipid accumulation potential. Additionally, cell volumes from the present study and volumes reported by Scholz & Liebezeit (2012) were plotted against percent lipid content per dry weight and no correlation was found, suggesting that neither cell volume nor cell shape has an effect on potential lipid accumulation.

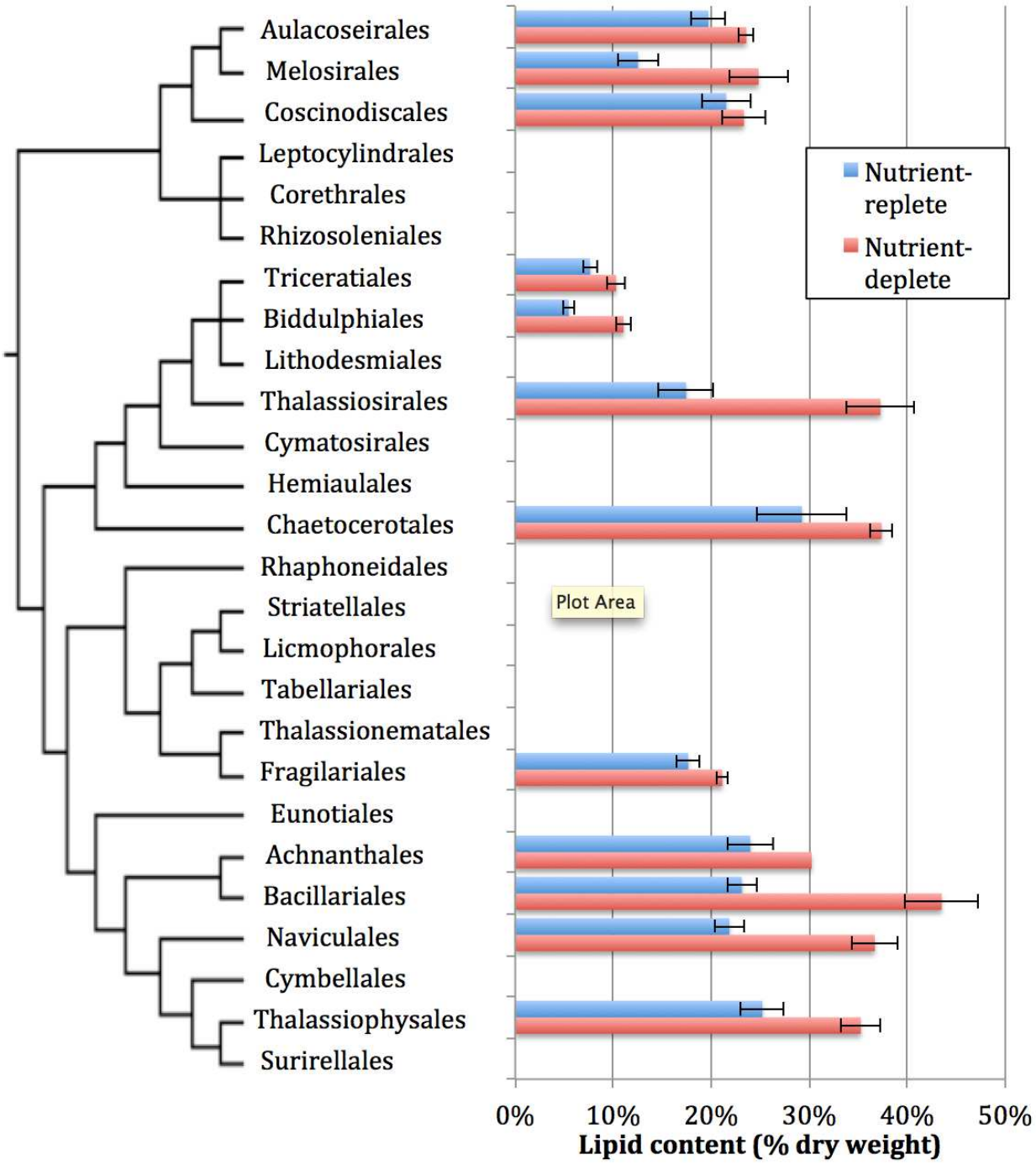

Figure 18. Average lipid content (expressed as percent dry weight) of diatoms, from this study and published work, in nutrient-replete and nutrient-deplete conditions. Data are grouped by order and compared within the context of a cladogram of the Bacillariophyta derived from Sims et al. (2006); the majority of orders have no data. Error bars indicate standard error of the mean

Chen (2012) states that it is difficult to make generalizations about the production and storage of lipids by diatoms in response to environmental shifts due to species-specific factors. Results from the present study suggest that certain lineages have a low variability between species (for lipid accumulation on certain reference bases) and a species' lipid response to an environmental shift could therefore be predicted based on the lineage to which that the species belongs.

There was low variability within the Fragilariophycidae lineage – except for lipid accumulation per culture volume during stationary growth (Figure 12). This means that it is possible to predict the response of an unknown species belonging to the Fragilariophycidae (more specifically, the Fragilariales) under most circumstances based on the current understanding of this lineage since there is little variability between the observed species. Additionally, as seen in the results comparing *B. alternans* to the Biddulphiophycidae, there was low variability of lipid content within this lineage during stationary growth. Although *O. aurita* did not respond to the silica-deplete medium in the same fashion as the other taxa, the former species continued to grow for an extended period (while the other taxa in this group moved immediately into a stationary phase) and still accumulated a similar amount of lipids (as percent lipid per dry weight) as these taxa did by the end of the stationary phase (Figure 9a). However, this pattern was not observed for lipid content per culture volume as *O. aurita* grew to a much higher density than the other taxa, and its lipid content per culture volume was thus significantly higher during stationary phase.

These findings indicate that a large variety of species from each major lineage (i.e. subclass or order) would need to be assayed to gain a further understanding of potential

lipid accumulation during stationary growth, assuming there is consistently low variability within lineages. Some species are difficult to culture in a laboratory, as observed during this research and others (Sheehan et al., 1998), and it will require an extensive sampling effort to collect the diverse taxa necessary to complete this endeavor.

5. CONCLUSIONS

Being able to deciphering a phylogenetic signal is dependent upon the reference basis used to quantify lipid content, the phylogenetic scale of analysis, and the growth phase the taxa are in when lipids are extracted. In the present study, results from both nested ANOVA and genus-level phylosignal test (Blomberg et al., 2003) revealed a phylogenetic signal during stationary growth for lipid content per dry weight. The fact that no signal could be detected during exponential growth or for lipid content per culture volume or cell volume indicates that there is currently insufficient evidence for a phylogenetic signal in diatom lipid accumulation. Currently, enough data to examine diatoms on a large scale only exists for lipid content per dry weight, a metric with inherent flaws. Additional metrics of lipid quantification, such as lipid content per culture volume or per cell volume (μg/mL and pg/100 μm^3), should be reported in future studies to provide more insight into the biological mechanisms behind lipid accumulation and to eliminate the cofounding effect of physiological differences between diverse species.

Together, the above factors control the variability observed within diatom lineages. Variability of lipid accumulation within higher lineages (e.g. subclass) seems to be irregular; some lineages display more variability between taxa than others. Data detailing lineages with little variability between species can be used to estimate the lipid content of similar species in these lineages. However, in order to decisively identify and make use of a phylogenetic signal for species selection, analyses may need to be conducted on a smaller taxonomic scale (e.g. within a genus), such as to eliminate the effect of variability between genera observed within some lineages.

Diatoms are an extremely diverse group in which certain lineages have adapted to niches within aquatic environments. Commercial operators should consider the lineage of origin of a species as well as multiple metrics that quantify physiology, lipid content, and growth rate for selecting a species for cultivation, such as to gain a better understanding of the biological parameters than govern lipid accumulation in microalgal species.

6. REFERENCES

Ashworth MP, Nakov T, Theriot EC (2013) Revisiting Ross and Sims (1971): toward a molecular phylogeny of the Biddulphiaceae and Eupodiscaceae (Bacillariophyceae). Journal of Phycology 49:1207-1222

Blomberg SP, Garland T, Ives AR (2003) Testing for phylogenetic signal in comparative data: behavioral traits and more labile. Evolution 57(4):717-745

Ceron Garcia MC, Garcia Camacho F, Sanches Miron A, Fernandex Sevilla JM, Chisti Y, Molina Grima E (2000) Journal of Microbiology and Biotechnology 16(5):689-694

Chen YC (2012) The biomass and total lipid content and composition of twelve species of marine diatoms cultured under various environments. Food Chemisty 131:211-219

Chisti Y (2007) Biodiesel from microalgae. Biotechnology Advances 25:294-306

Coombs J, Darley WM, Holm-Hansen O, Volcani BE (1967) Studies on the biochemistry and fine structure of silica shell formation in Diatoms. Chemical composition of *Navicula pelliculosa* during silicon-starvation synchrony. Plant Physiology 42:1601-1606

De la Pena MR (2007) Cell growth and nutritive value of the tropical benthic diatoms, *Amphora* sp., at varying levels of nutrient and light intensity, and different culture locations. Journal of Applied Phycology 19:647-655

Demirbas A (2007) Progress and recent trends in biofuels. Progress in Energy and Combustion Science 33:1-18

Demirbas A, Demirbas MF (2011) Importance of algae oil as a source of biodiesel. Energy Conversion and Management 52:163-170

Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJR (2004) The evolution of modern eukaryotic phytoplankton. Science 305(5682):354-360

Folch J, Lees M, Stanley GHS (1957) A simple method for the isolation and purification of total lipids from animal tissues. Journal of Biological Chemistry 226:497-509

Fourtainer E, Kociolek JP (1999) Catalogue of the diatom genera. Diatom Research 14(1):1-190

Gatenby CM, Orcutt DM, Kreeger DA, Parker BC, Jones VA, Neves RJ (2003) Biochemical composition of three algal species proposed as food for captive freshwater mussels. Journal of Applied Phycology 15:1-11

Graham JM, Graham LE, Zulkifly SB, Pfleger BF, Hoover SW, Yoshitani J (2011) Freshwater diatoms as a source of lipids for biofuels. Journal of Industrial Microbiology and Biotechnology 39(3):419-428

Griffiths MJ, Harrison STL (2009) Lipid productivity as a key characteristic for choosing algal species for biodiesel production. Journal of Applied Phycology 21:493-507

Giordana M, Beardall J, Raven JA (2005) CO_2 Concentration mechanisms in algae: mechanisms, environmental modulation, and evolution. Annual Review of Plant Biology 56:99-131

Gouy M, Guindon S, Gascuel O (2010) SeaView version 4 : a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular Biology and Evoluton 27(2):221-224

Guillard RRL (1972) Culture of phytoplankton for feeding marine invertebrates. Culture of Marine Invertebrate Animals. Plenum Publishers, New York: 29-60

Hillebrand H, Durselen CD, Kirschtel D, Pollingher U, Zohary T (1999) Biovolume calculation for pelagic and benthic microalgae. Journal of Phycology 35:403-424

Hoppenrath M, Elbrachter M, Drebes G (2009) Marine phytoplankton: selected microphytoplankton species from the North Sea around Helgoland and Sylt. E Schweizerbart Science Publishers, Stuttgart

Hu Q, Sommerfeld M, Jarvis E, Ghirardi M, Posewitz M, Seibert M, Darzins A (2008) Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. The Plant Journal 54:621-639

Jorgensen EG (1977) Photosynthesis. In: Werner D (ed) The biology of diatoms, Botanical Monographs Volume 13. University of California Press, Berkeley and Los Angeles

Krammer K, Lange-Bertalot H (2008) Bacillariophyceae. 3. Teil: Centrales, Fragilariaceae, Eunotiaceae. In Ettle, H, Gerloff J, Heynig H, Mollenhauer D (ed) Susswasser flora von Mitteleuropa, Band 2/3. Spektrum Akademischer Verlag, Heidelberg

Laurens LML, Quinn M, Wychen SV, Templeton DW, Wolfrum EJ (2012) Accurate and reliable quantification of total microalgal fuel potential as fatty acid methyl esters by *in situ* transesterification. Analytical and Bioanalytical Chemistry 403:167-178

Mansour MP, Frampton DMF, Nichols PD, Volkman JK, Blackburn SI (2005) Lipid and fatty acid yield of nine stationary-phase microalgae: Applications and unusual C_{24}-C_{28} polyunsaturated fatty acids. Journal of Applied Phycology 17:287-300

Nagle N, Lemke P (1990) Production of methyl ester fuel from microalgae. Applied Biochemistry and Biotechnology 24-25(1):355-361

Norton TA, Melkonian M, Andersen RA (1996) Algal biodiversity. Phycologia 35(4):308-326

Orcutt DM, Patterson GW (1975) Sterol, fatty acid and elemental composition of diatoms grown in chemically define media. Comparative Biochemistry and Physiology 50B:579-583

Renaud SM, Parry DL, Thinh LV (1994) Microalgae for use in tropical aquaculture I: Gross chemical and fatty acid of twelve species of microalgae from the Northern Territory, Australia. Journal of Applied Phycology 6:337-345

Rodolfi L, Zittelli GC, Bassi N, Padovani G, Biondi N, Bionini G, Tredici MR (2008) Mircoalgae for oil: strain selection, induction of lipid synthesis and outoor mass cultivation in a low-cost photobioreactor. Biotechnology and Bioengineering 102(1):100-112

Round FE, Crawford RM, Mann DG (1990) The diatoms: biology and morphology of the genera. Cambridge University Press, Cambridge

Scala S, Bowler C (2001) Molecular insights into the novel aspects of diatom biology. Cellular and Molecular Life Sciences 58:1666-1673

Scholz B, Liebezeit G (2013) Biochemical characterization and fatty acid profiles of 25 benthic marine diatoms isolated from the Solthorn tidal flat (southern North Sea). Journal of Applied Phycology 25:453-465

Sheehan J, Dunahay T, Benemann J, Roessler P (1998) A look back at the U.S. Department of Energy's Aquatic Species Program: Biodiesel from Algae. Close-Out report. National Renewable Energy Lab, Department of Energy, Golden, Colorado, U.S.A. Report number NREL/TP-580-24190, dated July 1998

Sims PA, Mann DG, Medlin LK (2006) Evolution of the diatoms: insights from fossil, biological and molecular data. Phycologia 45(4):361-402

Theriot E (1987) Principal component analysis and taxonomic interpretation of environmentally related variation in silicification in *Stephanodiscus* (Bacillariophyceae). British Phycological Journal 22:359-373

Theriot EC, Ashworth M, Ruck E, Nakov T, Jansen RK (2010) A preliminary multigene phylogeny of the diatoms (Bacillariophyta): challenges for future research. Plant Ecology and Evolution 143(3):278-296

Tomas, CR (1997) Identifying marine phytoplankton. Academic Press, Boston

US DOE (Department of Energy) (2010) National algal biofuels technology roadmap. US Department of Energy, Office of Energy Efficiency and Renewable Energy, Biomass Program

Wang C, Kong H, He S, Zheng X, Li C (2009) The inverse correlation between growth rate and cell carbohydrate content of *Microcystis aeuginosa*. Journal of Applied Phycology 22:105-107