DIVERSITY AND ECOLOGY OF LACUSTRINE HAPTOPHYTE ALGAE: IMPLICATIONS FOR PALEOTHERMOMETRY

By:

SUSANNA M. THEROUX

B.S. Williams College, Williamstown, MA, 2005

M.Sc. Brown University, Providence RI, 2009

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in The Department of Geological Sciences at Brown University

PROVIDENCE, RHODE ISLAND

MAY 2013

This dissertation by SUSANNA M. THEROUX is accepted in its present form by The Department of Geological Sciences as satisfying the dissertation requirement for the degree of Doctor of Philosophy.

Date_____

_____

Yongsong Huang, Ph.D., Co-Advisor

Date_____

_____

Linda Amaral-Zettler, Ph.D., Co-Advisor

Recommended to the Graduate Council

Date_____

_____

Timothy D. Herbert, Ph.D., Reader

Date_____

_____

James M. Russell, Ph.D., Reader

Date_____

_____

Marco J.L. Coolen, Ph.D., Reader

Approved by the Graduate Council

Date_____

_____

Peter Weber, Dean of the Graduate School

CURRICULUM VITAE

SUSANNA M. THEROUX

EDUCATION

Ph.D.: Geological Sciences, Brown University, Providence, RI,
Brown-Marine Biological Laboratory in Woods Hole graduate program
(2012 expected)

M.Sc.: Geological Sciences, Brown University, Providence, RI,
Brown-Marine Biological Laboratory in Woods Hole graduate program, 2009

B.A.: Biology and Geology with honors, Williams College, Williamstown, MA, 2005

FELLOWSHIPS & AWARDS

- American Association of University Women (AAUW) Dissertation Fellowship, 2011
- Geological Society of America (GSA) Student Research Grant, 2010
- Brown-MBL Graduate Fellowship in Biological and Environmental Sciences, 2007-2009
- Professor Jack L. Strominger Graduate Fellowship, Brown University, 2008
- Freeman Foot Award (for thesis presentation), Williams College, 2005
- 1960s Scholar in Geosciences, Williams College, 2005
- Student Scholar Athlete, Williams College, 2005
- Dean's List, Williams College, 2001-2005

TEACHING EXPERIENCE

- Woods Hole Semester in Environmental Studies, 2011, student mentor
- Woods Hole Partnership Education Program, 2011, guest lecturer and mentor
- GEO 031: Paleoecology and the Fossil Record, Brown University, Teaching Assistant and Lab Instructor
- GEO 020: Introduction to Oceanography, Brown University, Teaching Assistant

PUBLICATIONS

Theroux, S., Toney, J.L., Amaral-Zettler, L., Huang, Y. *in prep*. Production and temperature sensitivity of long chain alkenones in cultures of *Pseudoisochrysis paradoxa*, nomen nudum. Geochimica et Cosmochimica Acta.

Theroux, S., Huang, Y., Amaral-Zettler. 2012. Comparative molecular microbial ecology of spring haptophyte blooms in an arctic oligosaline lake in Greenland. Frontiers in Microbiology. doi: 10.3389/fmicb.2012.00415.

Toney, J., **Theroux, S**., Andersen, R., Coleman, A., Amaral-Zettler, L., Huang, Y. 2012. Culturing of the first $_{37:4}$ predominant lacustrine haptophyte: Geochemical, biochemical, and genetic implications. Geochimica et Cosmochimica Acta 78, 51-64.

Theroux, S., D'Andrea, W.J., Toney, J., Amaral-Zettler, L.A., Huang, Y. 2010. Phylogenetic diversity and evolutionary relatedness of alkenone producing haptophyte algae in lakes: Implications for continental paleotemperature reconstructions. EPSL 300, 311-320.

Amaral-Zettler,L., Zettler, E., **Theroux, S**., Palacios, C., Aguilera, A., Amils, R. 2010. Microbial community structure across the tree of life in the extreme Río Tinto. ISME Journal. doi:10.1038/ismej.2010.101

Stoll, H. M., Ziveri, P., Shimizu, N., Conte, M., **Theroux, S**. 2007. Relationship between coccolith Sr/Ca ratios and coccolithophore production and export in the Arabian Sea and Sargasso Sea. Deep Sea Research II: 54, 581-600.

Theroux, S. 2005. Effects of nutrient limitation on the productivity of coccolithophore algae and the paleoclimatic implications. BA thesis, Williams College.

PRESENTATIONS & POSTERS

Theroux, S., Diversity and ecology of Alkenone-producing haptophyte algae. *Invited*. Lamont Doherty Earth Observatory Spring Lecture Series. April, 2012.

Theroux, S., Toney, J., Andersen, R., Bohn, R., Nyren, P., Amaral-Zettler, L., Huang, Y. Dynamics of a haptophyte bloom in Lake George, ND: Implications for alkenone-based temperature reconstructions. American Geophysical Union Fall Meeting, 2011.

Salacup, J.M., **Theroux, S.**, Herbert, T.H., and Prell, W.L. Time-series of water column alkenones and 18S rRNA confirm that $U^{k'}_{37}$ is a viable SST proxy in Narragansett Bay, RI. American Geophysical Union Fall Meeting, 2011.

Theroux, S., Amaral-Zettler, L., Huang, Y. Insights into haptophyte bloom dynamics in an arctic lake using massively parallel tag sequencing. International Society of Microbial Ecology 13, 2010.

Theroux, S., D'Andrea, W. J., Toney, J.L., Amaral-Zettler, L., Huang, Y. Relating Phylogeny to Alkenone Distributions in Lacustrine Alkenone-Producing Haptophytes: Implications for Continental Paleotemperature Reconstructions. American Geophysical Union Fall Meeting, 2008.

Toney, J.L., Fritz, S., Baker, P.A., Grimm, E.C., Nyren, P., **Theroux, S**., Huang, Y. Environmental and Climatic Control on the Occurrence and Abundance of Long Chain

Alkenones in Lakes of the Interior United States. American Geophysical Union Fall Meeting, 2008.

Zettler, E., Amils, R., **Theroux, S**., Palacios, C., Sogin, M., Amaral-Zettler, L. Protistan distribution in relation to spatial variations of extreme geochemical parameters in the Rio Tinto, Spain. Phycology Society of American Annual Meeting, 2007.

Theroux, S., Burke, A., Stoll, H., Shimuzu, N., Ziveri, P. Constraining nutrient effects in monsoon-driven productivity shifts during the Mediterranean sapropel events. Presentation, Geological Society of America Northeastern Section annual meeting, 2005.

EMPLOYMENT

- Research Assistant, 2005-2007
Marine Biological Laboratory, Woods Hole, MA
Josephine Bay Paul Center for Comparative Molecular Biology and Evolution,
Dr. Linda Amaral-Zettler Laboratory

- Summer Research Fellowship in Geosciences, 2004-2005
Williams College, Williamstown, MA
Dr. Heather Stoll Laboratory

- Summer Research Fellowship in Geosciences, 2003
Vrije Universiteit, Amsterdam, NL
Dr. Patrizia Ziveri Lab oratory

MEMBERSHIPS & CERTIFICATIONS

Algal Biomass Organization (ABO)
American Association of University Women (AAUW)
Earth Science Women's Network (ESWN)
American Geophysical Union (AGU)
Sigma XI Scientific Research Society
Geological Society of America (GSA)
International Society for Microbial Ecology (ISME)
Phycological Soceity of America (PSA)
Spanish Language Advanced Proficiency

FIELD EXPERIENCE

Lake George, North Dakota: lake sampling, algal culturing, 2011
Kangerlussuaq, Greenland: lake sampling, algal culturing, 2009
Kangerlussuaq, Greenland: lake sampling, coring expedition, algal culturing, 2008
Kangerlussuaq, Greenland: lake sampling, lake sediment coring, 2007
Tiputini Biodiversity Station, Parque Nacional Yasuni, Ecuador: 2004

Dedicated to my family

ACKNOWLEDGEMENTS

I would like to extend my greatest appreciation to my co-advisors, Linda Amaral-Zettler and Yongsong Huang. I would also like to thank the members of my thesis committee, Timothy Herbert, James Russell, Zoe Cardon and Marco Coolen. I would like to thank the Earth Systems History group at Brown University and the Josephine Bay Paul Center at the Marine Biological Laboratory. I would also like to thank my friends and family, especially Mom, Dad, Katie, Katie and Casper.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

SUSANNA M. THEROUX

BROWN UNIVERSITY, *Department of Geological Sciences*
MARINE BIOLOGICAL LABORATORY, *Josephine Bay Paul Center*

1. Introduction

In the face of climate change, the need for accurate climate forecasting relies heavily upon accurate climate hindcasting. Molecular fossils preserved in the geologic record present ideal opportunities for reconstructing ancient environments. However, a carpenter is only as good as his (her) tools and likewise, a paleoclimatologist is only as good as his (her) proxy. In this vein, this thesis seeks to improve our understanding of extant, alkenone-producing haptophyte algae in order to improve our application of the alkenone-based paleotemperature proxy.

1.1 Alkenone production in haptophyte algae

Haptophyte algae are a globally important group of photosynthetic microbes. Haptophytes are responsible for large-scale bloom events throughout the world's oceans, are credited as being the largest producers of calcium carbonate on the planet and are arguably the dominant carbon-fixers in the ocean, even out-pacing cyanobacteria (Liu *et al.,* 2009). Certain haptophyte species within the Class Isochrysidales (Division Haptophyta) produce long chain polyunsaturated alkenones (LCAs) (Marlowe *et al*., 1984). These diagenetically stable compounds have been found in sediments dating back to the lower Aptian (120.5 Ma) although the original haptophyte species responsible for these alkenones is unknown (Brassell *et al.,* 2004). In modern marine environments, alkenone production is dominated by species *Emiliania huxleyi* and its close relative *Gephyrocapsa oceanica*. In the lesser-studied coastal and lacustrine environments, *Isochrysis galbana* and *Chysotila lamellosa* are known alkenone-producers (Marlowe *et al*., 1984).

The physiological role of the alkenone molecule is still unknown. Originally believed to be a membrane lipid, this theory was dismissed upon discovery of the alkenone's trans (versus typical membrane *cis*) unsaturated bonds (Rechka and Maxwell, 1988). The observation that alkenones accumulated in cells under nutrient deprivation (Rontani *et al*., 2004) and are metabolized when cells are incubated in the dark (Eltgroth *et al*., 2005) suggested they are used as an energy storage molecule. Observation studies localized the alkenone lipid to cytoplasmic vesicles often associated with chloroplasts (Eltgroth *et al*., 2005). A proposed biosynthetic pathway theorized that alkenones are synthesized in the chloroplast from an acetyl- or propinoyl-SCoA precursor, exported to the cytoplasm, and progressively unsaturated by a series of desaturases (Rontani *et al*., 2006).

1.2 Alkenone-based paleothermometry

Only four species of Haptophyta have been systematically cultivated in pure culture to gauge alkenone production: predominant, and bloom-forming, marine haptophyte *Emiliania huxleyi* (Prahl *et al*., 1988; Prahl *et al*., 2003; Conte *et al*., 1998), close marine relative *Gephyrocapsa oceanica* (Sawada *et al*., 1996; Conte *et al*., 1998), and coastal/lacustrine species *Chrysotila lamellosa* (Marlowe *et al*., 1984; Sun *et al*., 2007) and *Isochrysis galbana* (Marlowe *et al*., 1984; Versteegh *et al*., 2001).

Haptophytes fortuitously record the temperature of their surrounding waters in the degree of unsaturation of their alkenones; as temperature decreases, the alkenones grow systematically more unsaturated. This relationship between water temperature

and alkenone unsaturation was first described by Brassell *et al*. (1986) and incorporated the relative concentrations of alkenones of 37 carbon atoms (C_{37}) with either two, three of four points of unsaturation ($C_{37:2}$, $C_{37:3}$, $C_{37:4}$, respectively). This relationship, the U^k_{37} alkenone-based paleotemperature proxy, was defined by Brassell *et al*. (1986) as:

$$U^k_{37} = \frac{[C37:2] - [C37:4]}{[C37:2] + [C37:3] + [C37:4]}$$

In environments with negligible $C_{37:4}$ concentrations, the U^k_{37} index was modified by Prahl and Wakeham (1987) to:

$$U^k_{37} = \frac{[C37:2]}{[C37:2] + [C37:3]}$$

The empirical definition of the linear relationship between the increasingly unsaturated compounds ($C_{37:3}$) with cooler water temperatures in cultures of *E. huxleyi* established the first calibration for the U^k_{37} proxy to calculate ancient sea surface temperatures (Prahl and Wakeham, 1987):

$$U^k_{37} = 0.040T - 0.11$$

This relationship was confirmed with a global coretop calibration by Muller *et al*. (1998):

$$U^k_{37} = 0.033T + 0.044$$

These U^k_{37} calibrations have been used successfully to reconstruct sea surface temperatures from sediment cores throughout the world oceans.

The desire to extend the use of the U^k_{37} paleotemperature proxy to the continents focused attention on lacustrine alkenone archives. In contrast to marine sediments, lake alkenones are characterized by an abundance of tetraunsaturated alkenones, indicating the presence of a different alkenone-producing haptophyte species or an environmental control dictating a distinct alkenone profile. Alkenones have been identified in lakes sediments around the world, including Antarctica, Europe, Greenland, South America, China and the western United States (Thiel *et al.,* 1997; Coolen *et al.,* 2004; D'Andrea and Huang, 2005; Liu *et al.,* 2008; Pearson *et al.,* 2008; Toney *et al.,* 2010; Zink *et al.,* 2001). Lacustrine alkenone unsaturation ratios correspond to seasonal surface water temperatures and mean annual air temperature (Chu *et al.,* 2005; D'Andrea and Huang, 2005; Sun *et al.,* 2007; Zink *et al.,* 2001) further confirming the potential utility of alkenone archives from lake sediments. However, because of the paucity of lacustrine U^k_{37} calibrations, efforts to reconstruct paleotemperatures from a suite of German lakes yielded erroneous alkenone-based paleotemperature estimates (Zink *et al.*, 2001). Clearly, a new method of identifying haptophytes species, and their U^k_{37} calibrations, is needed.

1.3 Molecular biological approaches to temperature reconstructions

For the past eight years, there has been a movement towards identifying haptophyte species via their most specific biomarker: DNA. A study of preserved DNA in the sediments of Ace Lake, Antarctica revealed the presence of multiple haptophyte species closely related to known coastal/lacustrine species *Isochrysis galbana* (Coolen *et al.,* 2004). Like Ace Lake, the sediments of a suite of oligosaline

lakes in southwestern Greenland were dominated by the $C_{37:4}$ alkenone (D'Andrea and Huang, 2005). However, unlike Ace Lake, the Greenland lake sediments also had the presence of the C_{38} methyl ketone, typically only seen in marine alkenone profiles (D'Andrea and Huang, 2005; Schulz *et al.,* 2000). Haptophyte 18S ribosomal RNA (rRNA) gene sequences revealed a novel clade of haptophyte algae, one distinct from both the marine and lacustrine haptophyte clades (D'Andrea *et al*. 2006). These preliminary studies set the stage for the global exploration of alkenone-producing haptophytes in lake environments. Using a multidisciplinary approach, this thesis aims to i) identify and classify novel species of alkenone-producing haptophytes, ii) calibrate their U^{k}_{37} alkenone-unsaturation indices and iii) determine how haptophyte bloom events impact the lacustrine alkenone sediment record.

1.4 Dissertation outline

This dissertation applies molecular biological techniques to answer enduring questions in the development of the alkenone-based paleothermometry in lake environments.

Chapter 2 is the first global survey of lacustrine sediments to couple alkenone profiles with haptophyte DNA sequences. We analyzed 18S rRNA gene sequences specifically for haptophytes to test the validity of inferring haptophyte species from alkenone lipid profiles alone. Our results showed that DNA sequences were the more specific identifier of haptophyte species and that lakes with identical alkenone profiles can have different haptophyte communities. Additionally, some lakes hosted multiple, distantly related haptophyte species, indicating that alkenone profiles may

be composite profiles with unknown contributions by either species. In this study we also sequenced paleo-DNA and found the conservation of alkenone-producing haptophytes in downcore sediments from Lake BrayaSø, Greenland. This result green-lights the use of a single U^k_{37} calibration to be applied when reconstructing paleotemperatures from the BrayaSø sediment record.

Chapter 3 is a high-throughput DNA sequencing survey of the microbial communities surrounding a haptophyte bloom. We analyzed water samples from both the 2007 and 2009 haptophyte bloom events in Lake BrayaSø, Greenland, and characterized the bacterial and eukaryotic microbial communities. We found a surprising lack of overlap in microbial taxa between the two bloom years. However, this study confirmed the presence of only one haptophyte species in Lake BrayaSø, a necessary condition for future application of a single U^k_{37} calibration. Using quantitative PCR, our results also confirmed that alkenone concentrations correspond to haptophyte cell concentrations in the water column. This study serves as a benchmark microbial survey in an arctic oligosaline lake, providing baseline knowledge of arctic microbial communities as they encounter rapid climate change.

Chapter 4 addressed a prerequisite for the further development of the lacustrine U^k_{37} proxy: the definition of species-specific U^k_{37} calibrations. In this study we grew haptophyte *Pseudoisochrysis paradoxa* in pure culture at different temperatures to define its alkenone unsaturation—temperature relationship. *Pseudoisochrysis paradoxa* is the closest relative to a novel species of haptophyte identified from Lake George, ND, and this study serves as an example of how *in situ* investigations of alkenone production by haptophytes requires a complimentary

dataset of reference haptophyte species that have been examined *in silico*. The cultures of *P. paradoxa* had dominant $C_{37:3}$ isomers and high concentrations of alkenones per cell at the lowest growth rates. The U^k_{37} calibration for *P. paradoxa* was most similar to the calibration for species *C. lamellosa* instead of *P. pardoxa's* closest relative, *Isochrysis galbana*.

Chapter 5 investigated the hypothesis that $C_{37:4}$ alkenone concentrations can be used as a proxy for lake water salinity. We sampled a suite of lakes on the Chinese-Tibetan plateau in the Lake Qinghai region, a region populated by lakes of various salinities containing alkenone profiles with variable $C_{37:4}$ concentrations. Our results supported previous findings that $C_{37:4}$ concentrations increased with decreasing salinity. However, instead of a single haptophyte species changing alkenone production with varying salinity, we found evidence to suggest that variations in $C_{37:4}$ alkenone concentrations are a reflection of changing haptophyte communities. We identified two haptophyte species in Lake Qinghai, each with a distinct alkenone signature. The fluctuation in abundance of these two haptophyte species explains the observed variations in $C_{37:4}$ concentrations. We conclude that environmentally-dictated variations in haptophyte communities are responsible for the observed relationship between $C_{37:4}$ alkenone concentrations with salinity.

In Chapter 6 we targeted the elusive $C_{37:4}$ dominant alkenone-producing haptophyte from Lake George, ND. Previous research (Theroux *et al.,* 2010; Toney *et al.,* 2012) had shown that enrichment cultures inoculated with Lake George sediments could reliably produce $C_{37:4}$ alkenones. Relying upon a size fractionation experiment, we successfully isolated the $C_{37:4}$-dominant haptophyte away from its $C_{37:3}$-dominant

relative, and grew this novel $C_{37:4}$-producing haptophyte in culture to determines its U^k_{37} calibration. The $C_{37:4}$-dominant haptophyte was isolated in the <3μm fraction, and microscopy revealed this haptophyte is indeed a picoplankter. This finding improves our understanding of the ecology of this novel, and long-mysterious haptophyte species and provides insight into how its physiology gives it a competitive advantage in lake systems.

Chapter 7 was our most aggressive investigation into a haptophyte bloom event and alkenone production in a multiple-haptophyte system. With this study, we analyzed the geochemical and biological profiles in Lake George, ND throughout the course of the seasonal bloom cycle. Using biweekly sampling, we generated alkenone concentration profiles coupled with DNA profiling of eukaryotic microbial communities. This is the first study to couple high-throughput DNA sequencing with alkenone analyses to decipher changes in haptophyte community composition and biomarker production throughout a seasonal cycle. Additionally, we isolated individual haptophytes from Lake George and grew these novel species at controlled temperatures to define their U^k_{37} calibrations. Our multi-disciplinary approach revealed that in Lake George, the two-haptophyte system, at times, produces a composite alkenone signature recorded in the sediment record.

References

Brassell SC, Eglinton G, Marlowe IT, Pflaumann U, Sarnthein M. (1986). Molecular stratigraphy - a new tool for climatic assessment. *Nature* **320**:129-133.

Brassell SC, Dumitrescu M. (2004). Recognition of alkenones in a lower Aptian porcellanite from the west-central Pacific. *Organic Geochemistry* **35**:181-188.

Chu GQ, Sun Q, Li SQ, Zheng MP, Jia XX, Lu CF, Liu JQ, Liu TS. (2005). Long-chain alkenone distributions and temperature dependence in lacustrine surface sediments from China. *Geochimica Et Cosmochimica Acta* **69**:4985-5003.

Conte MH, Thompson A, Lesley D, Harris RP. (1998). Genetic and physiological influences on the alkenone/alkenoate versus growth temperature relationship in *Emiliania huxleyi* and *Gephyrocapsa oceanica*. *Geochimica Et Cosmochimica Acta* **62**:51-68.

Coolen MJL, Muyzer G, Rijpstra WIC, Schouten S, Volkman JK, Sinninghe Damsté J. (2004). Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake. *Earth and Planetary Science Letters* **223**:225-239.

D'Andrea WJ, Huang Y. (2005). Long chain alkenones in Greenland lake sediments: Low δ13C values and exceptional abundance. *Organic Geochemistry* **36**:1234-1241.

D'Andrea WJ, Lage M, Martiny JBH, Laatsch AD, Amaral-Zettler LA, Sogin ML, Huang YS. (2006). Alkenone producers inferred from well-preserved 18S rDNA in Greenland lake sediments. *Journal of Geophysical Research-Biogeosciences* **111**:doi:10.1029/2005JG000121.

Eltgroth ML, Watwood RL, Wolfe GV. (2005). Production and cellular localization of neutral long-chain lipids in the haptophyte algae Isochrysis galbana and Emiliania huxleyi. *Journal of Phycology* **41**:1000-1009.

Liu H, Probert I, Uitz J, Claustre H, Aris-Brosou S, Frada M, Not F, de Vargas C. (2009). Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. *Proceedings of the National Academy of Sciences of the United States of America* **106**:12803-12808.

Liu WG, Liu ZH, Fu MY, An ZH. (2008). Distribution of the C37 tetra-unsaturated alkenone in Lake Qinghai, China: A potential lake salinity indicator. *Geochimica Et Cosmochimica Acta* **72**:988-997.

Marlowe IT, Green JC, Neal AC, Brassell SC, Eglinton G, Course PA. (1984). Long-Chain (**n**-C37-C39) alkenones in the prymnesiophyceae. Distribution of

alkenones and other lipids and their taxonomic significance. *British Phycological Journal* **19**:203-216.

Pearson EJ, Juggins S, Farrimond P. (2008). Distribution and significance of long-chain alkenones as salinity and temperature indicators in Spanish saline lake sediments. *Geochimica Et Cosmochimica Acta* **72**:4035-4046.

Prahl FG, Wolfe GV, Sparrow MA. (2003). Physiological impacts on alkenone paleothermometry. *Paleoceanography* **18**:doi:10.1029/2002PA000803.

Prahl FG, Muehlhausen LA, Zahnle DL. (1988). Further evaluation of long-chain alkenones as indicators of paleoceanographic conditions. *Geochimica Et Cosmochimica Acta* **52**:2303-2310.

Prahl FG, Wakeham SG. (1987). Calibration of unsaturation patterns in long-chain ketone compositions for paleotemperature assessment. *Nature* **330**:367-369.

Rechka JA, Maxwell JR. (1988). Characterisation of alkenone temperature indicators in sediments and organisms. *Organic Geochemistry* **13**:727-734.

Rontani JF, Beker B, Volkman JK. (2004). Long-chain alkenones and related compounds in the benthic haptophyte *Chrysotila lamellosa* Anand HAP 17. *Phytochemistry* **65**:117-126.

Rontani JF, Prahl FG, Volkman JK. (2006). Re-examination of the double bond positions in alkenones and derivatives: biosynthetic implications. *Journal of Phycology* **42**:800-813.

Sawada K, Handa N, Shiraiwa Y, Danbara A, Montani S. (1996). Long-chain alkenones and alkyl alkenoates in the coastal and pelagic sediments of the northwest North Pacific, with special reference to the reconstruction of Emiliania huxleyi and Gephyrocapsa oceanica ratios. *Organic Geochemistry* **24**:751 - 764.

Schulz H-M, Schöner A, Emeis K-C. (2000). Long-chain alkenone patterns in the Baltic sea--an ocean-freshwater transition. *Geochimica Et Cosmochimica Acta* **64**:469-477.

Sun Q, Chu GQ, Liu GX, Li S, Wang XH. (2007). Calibration of alkenone unsaturation index with growth temperature for a lacustrine species, *Chrysotila lamellosa* (Haptophyceae). *Organic Geochemistry* **38**:1226-1234.

Theroux SM, D'Andrea WJ, Toney JL, Amaral-Zettler LA, Huang Y. (2010). Phylogenetic diversity and evolutionary relatedness of alkenone producing haptophyte algae in lakes: Implications for continental paleotemperature reconstructions. *Earth and Planetary Science Letters* **300**:311-320.

Thiel V, Jenisch A, Landmann G, Reimer A, Michaelis W. (1997). Unusual distributions of long-chain alkenones and tetrahymanol from the highly alkaline Lake Van, Turkey. *Geochimica Et Cosmochimica Acta* **61**:2053-2064.

Toney JL, Huang Y, Fritz SC, Baker PA, Grimm E, Nyren P. (2010). Climatic and environmental controls on the occurrence and distributions of long chain alkenones in lakes in the interior United States. *Geochimica Et Cosmochimica Acta*

Versteegh GJM, Riegman R, de Leeuw JW, Jansen JHF. (2001). $U^{k'}_{37}$ values for *Isochrysis galbana* as a function of culture temperature, light intensity and nutrient concentrations. *Organic Geochemistry* **32**:785-794.

Zink KG, Leythaeuser D, Melkonian M, Schwark L. (2001). Temperature dependency of long-chain alkenone distributions in Recent to fossil limnic sediments and in lake waters. *Geochimica Et Cosmochimica Acta* **65**:253-265.

CHAPTER 2

PHYLOGENETIC DIVERSITY AND EVOLUTIONARY RELATEDNESS OF ALKENONE PRODUCING HAPTOPHYTE ALGAE IN LAKES: IMPLICATIONS FOR CONTINENTAL PALEOTEMPERATURE RECONSTRUCTIONS

SUSANNA THEROUX[1,2]

WILLIAM D'ANDREA[1,*]

JAIME TONEY[1,‡]

LINDA AMARAL-ZETTLER[1,2,3]

YONGSONG HUANG[1]

[1] BROWN UNIVERSITY, *Department of Geological Sciences*

[2] MARINE BIOLOGICAL LABORATORY, *Josephine Bay Paul Center*

[3] BROWN UNIVERSITY, *Department of Ecology and Evolutionary Biology*

[*] Present address: LAMONT-DOHERTY EARTH OBSERVATORY, *Biology and Paleo Environment*

[‡] Present address: UNIVERSITY OF GLASGOW, *School of Geographical and Earth Sciences*

Published in:
Earth and Planetary Science Letters 300, 311-320.

Abstract

Alkenones have been found in an increasing number of lakes around the world, making them a promising new tool for continental paleoclimate reconstruction. However, individual lakes may harbor different species of haptophyte algae with different sensitivities to temperature variations, thus presenting a significant challenge to the use of lacustrine alkenones for paleotemperature reconstructions. To explore the extent of lacustrine haptophyte diversity, we conducted the first comprehensive phylogenetic and geochemical survey of lacustrine alkenone producers. We sampled 15 alkenone-containing lake surface sediments from a variety of geographic locales and inferred identities of environmental sequences using 18S ribosomal RNA (rRNA) gene based phylogenies. For two lakes, BrayaSø in southwest Greenland and Tso Ur on the Tibetan Plateau, we also analyzed both surface and down-core sediments to characterize haptophyte populations through time. In parallel with phylogenetic analyses, we determined the alkenone distributions (including C_{37}/C_{38} ratios, and the presence/absence of C_{38} methyl ketones and C_{40} compounds) in all the samples. The resulting alkenone profiles from this study do not all align with traditional "marine" versus "coastal/lacustrine" alkenone profiles. Additionally, our genetic data indicate the presence of multiple haptophyte species from a single lake sediment sample; these composite haptophyte populations could not be discerned from the alkenone profiles alone. These results show that alkenone profiles are not a reliable way to assess the haptophyte algae in lakes and that DNA fingerprinting is a preferred approach for species identification. Although closely-related haptophyte species or subspecies may not warrant different temperature calibrations, our results emphasize the importance of genetic data for inferring haptophyte identities and eventually selecting alkenone-temperature calibrations for paleoclimate reconstructions.

1. Introduction

Marine haptophyte algae produce unsaturated long-chain alkenones (LCAs) that serve as a well-established paleotemperature proxy (U^{K}_{37} or $U^{K'}_{37}$) in marine sediments (Volkman *et al*., 1980; Marlowe *et al*., 1984; Brassell *et al*., 1986; Prahl and Wakeham, 1987; Conte *et al*., 2006). Alkenones are produced by a limited number of haptophyte algal species within the order Isochrysidales (Müller *et al*., 1998; Medlin *et al*., 2008), notably *Emiliania huxleyi* and *Gephyrocapsa oceanica* in the open ocean (Volkman *et al*., 1980; Marlowe *et al*., 1984), and *Isochrysis galbana* (Marlowe *et al*., 1984; Versteegh *et al*., 2001) and *Chrysotila lamellosa* in coastal regions (Volkman *et al*., 1980; Marlowe *et al*., 1984; Volkman *et al*., 1995; Conte *et al*., 1998; Rontani *et al*., 2004). *Emiliania huxleyi* and *G. oceanica* (Müller *et al*., 1998; Conte *et al*., 2006) are the dominant alkenone-producers in marine surface waters, enabling the application of a global $U^{K'}_{37}$ calibration to marine sea surface temperature (SST) reconstructions.

Long chain alkenones (LCAs) in lake sediments have great potential for continental climate reconstructions. LCAs have been increasingly reported in lakes around the world, including Antarctica, Europe, Greenland, South America, China and the western United States (Zink *et al*., 2001; Coolen *et al*., 2004; D'Andrea and Huang, 2005; Liu *et al*., 2008; Pearson *et al*., 2008; Toney *et al*., 2009). Lacustrine alkenone unsaturation ratios correspond to lake seasonal surface water temperatures and mean annual air temperature (Zink *et al*., 2001; Chu *et al*., 2005; D'Andrea and Huang, 2005; Sun *et al*., 2007). However, individual alkenone-containing lakes may require different calibrations, depending on the haptophyte species present (Prahl and

Wakeham, 1987; Prahl *et al*., 1988; Volkman *et al*., 1995; Versteegh *et al*., 2001; Zink *et al*., 2001; Rontani *et al*., 2004; Chu *et al*., 2005; Sun *et al*., 2007; D'Andrea, 2008; Liu *et al*., 2008). Previous studies have generated a variety of LCA profiles from different lakes (Liu *et al*., 2008; Pearson *et al*., 2008), with varying ratios of C_{37}/C_{38} alkenone, percent concentrations of tetraunsaturated C_{37} ketones (% $C_{37:4}$), and the presence/absence of C_{38} methyl ketones and C_{40} alkenones. These variable distributions suggest that alkenones from different lakes could be produced by a diversity of haptophyte species, and therefore a single global temperature calibration may not be viable.

Although alkenone signatures provide a simple way to compare alkenones from different lakes, they do not provide definitive taxonomic information at the species level. Recent studies identified haptophyte species using genetic methods and enabled the identification of novel haptophytes species without the need for cultivation. Work by Coolen *et al*. (2004) targeted the 18S rRNA molecule to identify the haptophyte responsible for alkenone production in ancient Ace Lake, Antarctic sediments. These authors identified preserved DNA of a haptophyte closely related to *I. galbana*. D'Andrea and Huang (2005) reported the occurrence of a novel bloom-forming haptophyte species in lakes of Southwestern Greenland with extremely high alkenone concentrations, characterized by an abundance of tetra-unsaturated alkenones and the presence of C_{38} methyl and ethyl ketones. Genetic work determined this Greenland haptophyte forms a distinct phylogenetic group from the marine and lacustrine phylotypes within the Isochrysidales (D'Andrea *et al*., 2006).

This paper aims to expand the phylogenetic dataset of lacustrine alkenone-producing haptophytes and compare 18S rRNA gene-based phylogenetic relationships with the observed alkenone lipid signatures. We studied lake sediments from Greenland, China, Northern Canada, and the continental United States. We hypothesized that different alkenone lipid signatures in lake sediments can be attributed to the genetic diversity of the haptophyte species, and that multiple alkenone-producing haptophyte species could be present in individual lakes. With this comprehensive survey we were able to identify a) presence/absence of biogeographical trends in lacustrine haptophyte distribution, b) relationships between alkenone lipid profiles and genetic rRNA-inferred phylogenies, and c) the accuracy and fidelity of the alkenone signature as a haptophyte identification tool. This dataset of haptophyte alkenone signatures and corresponding species classifications assesses the reliability of haptophyte species identification through alkenone signature alone.

2. Methods

2.1 *Sampling*

Previous studies have demonstrated that DNA of haptophyte algae can be readily extracted from lake sediments (Coolen *et al.,* 2004; D'Andrea *et al.,* 2006). We analyzed sediments from 15 lakes (**Figure 1, Table 1**) from the continental United States, Canada, Greenland, and Tibet. Lake surface sediment samples were collected by gravity core, split, and the upper 1 cm was used for analysis. Lakes Clear, Medicine, George, and Skoal sediments were freeze-dried before DNA

extraction. All other sediment samples were kept at 4 °C and in the dark until processing. Individual sediment samples were split for DNA and alkenone analysis.

For Lake BrayaSø water column DNA, we collected water on June 26, 2007 at 10 m depth, corresponding to the chlorophyll maximum that day. One liter was filtered using a 0.22 μm pore size Sterivex filter (Millipore, Bedford, MA, USA), treated with Puregene cell lysis buffer (Qiagen, Carlsbad, CA, USA) and kept cool and in the dark until freezing at -20 °C.

2.2 *DNA extraction*

We extracted DNA from 1 g of lake surface sediment for each lake we examined using MoBio UltraPure DNA Extraction Kit for sediments (Qiagen, Carlsbad, CA, USA) according to the manufacturer's instructions. For lakes BrayaSø and Tso Ur, down-core sediment samples were also extracted for DNA analysis. For water samples, Sterivex filters were extracted using a Puregene Cell Kit (Qiagen) according to the manufacturers instructions. Genomic DNA was polyethylene glycol (PEG) purified to remove proteins and other contaminants that inhibit PCR reactions: DNA was suspended in PEG at 4 °C overnight, centrifuged, and the pellet rinsed with ethanol (LaMontagne *et al.,* 2002). We quantified total extracted genomic DNA yields using a NanoDrop nucleic acid spectrophotometer (Thermo Scientific, Wilmington, DE).

2.3 *DNA amplification and sequencing*

We amplified genomic DNA using haptophyte-specific oligonucleotide (Simon *et al*., 2000; Coolen *et al*., 2004) primers targeting 18S rRNA coding regions. Forward and reverse primers corresponded to *E. coli* 18S rRNA positions 429 and 887, respectively (**Table 2**). Polymerase chain reactions (PCRs) were performed on an Eppendorf Gradient Thermocycler (Hamburg, Germany) with the following conditions after D'Andrea *et al*. (2006): 4 min initial denaturing at 96 °C, 35 cycles of denaturing for 30 s at 94 °C, followed by 40 s primer annealing at 55 °C and primer extension 40 s at 72 °C, with a final extension of 10 min at 72 °C. We devoted a separate PCR run for each lake to avoid cross contamination. We ran each sample at optimal dilution (1:10, 1:100, or 1:1000) for maximum product yield. PCR reactions were run in 50 μl volume using Promega GoTaq polymerase (Madison, WI, USA). Triplicate PCR products were pooled for each sample, and templates were purified using an Invitrogen PureLink Purification kit (Carlsbad, CA, USA) with the high-cut-off binding buffer to eliminate fragments < 200 bp. Purified PCR products were A-tailed and purified using the PureLink high cut-off kit. Cloning was performed using the Invitrogen TOP10 cloning kit with electrocompetent cells. Protocol was according to the manufacturer's instructions. One hundred clones were picked for each sample. Plasmid DNA was isolated using a RevPrep Orbit robotic template preparation instrument (Genomic Solutions, Ann Arbor, MI), and prepared templates were sequenced on an ABI 3730XL (Applied Biosystems, Foster City, CA) capillary sequencer using the BigDye protocol with universal M13 forward and reverse primers according to the manufacturer's instructions. All sequencing was performed at the

Marine Biological Laboratory W. M. Keck Ecological and Evolutionary Genetics Facility.

For additional PCR amplification of the complete Greenland Lake BrayaSø haptophyte 18S rRNA gene, we designed an internal primer and paired this with universal eukaryotic forward and reverse primers (Medlin *et al.,* 1988), in addition to the haptophyte-specific primers (**Table 2**). The PCR program was as follows: 2 min denaturing at 94 °C, 30 cycles of denaturing for 1 min at 94 °C, primer annealing at 45 °C for 1 min, and primer extension for 2 min at 72 °C. The final extension was for 10 min at 72 °C. DNA purification, cloning, and sequencing were as above.

2.4 *Bioinformatics and phylogenetic reconstructions*

A bioinformatics pipeline using the programs phred, cross-match, and phrap, translated chromatograms into base-calls and associated quality scores, removed vector sequences and assembled forward and reverse reads into full-length sequences for each of the cloned PCR amplicons (Ewing and Green, 1998; Ewing *et al.,* 1998; Lasek-Nesselquist *et al.,* 2008). Only sequences greater than 400 bp and with a complete forward and reverse primer were retained. Base-calls were verified and sequences were manually edited with the program Consed (Gordon *et al.,* 1998) for chromatogram viewing. We screened edited sequences for chimeras using the computer program Mallard (Ashelford *et al.,* 2006). The number of sequences recovered for each sample is listed in **Table 3.** For samples with only single read coverage, we verified the sequence by re-sequencing the template in both the forward and reverse directions. We identified closest relatives within the GenBank database

using BLAST (Altschul *et al*., 1997). Assembled sequences were aligned using the ARB software program v. 07.07.11 (Ludwig *et al*., 2004) against the April 2008 Silva 94 Ref database (Pruesse *et al*., 2007) using the FastAligner option followed by manual adjustment. Operational Taxonomic Units (OTUs) were obtained using the DOTUR program (Schloss and Handelsman, 2005) furthest neighbor algorithm setting at 95 % cut-off criterion. Representative sequences for each OTU were selected for tree construction. Full-length 18S rRNA gene contigs were assembled using the MacClade software package version 4 (Maddison and Maddison, 2000). We constructed a distance matrix by calculating differences/sequence length for aligned reference haptophyte sequences and representative OTU sequences, using the neighbor joining method as implemented in ARB.

A custom filter of 1,756 positions was constructed manually for both Bayesian analyses of 18S rRNA sequences. We selected *Cyclonexis annularis*, *Chrysoxys* sp., *Ochromonas danica, Odontella sinensis* and *Thraustochytrium multirudimentale* as outgroups for our Bayesian analyses after de Vargas *et al*. (2007). We subjected our datasets to a Bayesian analysis using MrBayes version 3.0b4 (Ronquist and Huelsenbeck, 2003) under the GTR model of substitution (Lanave *et al*., 1984; Tavaré, 1986; Rodriguez *et al*., 1990) considering invariants and a gamma-shaped distribution of the rates of distribution among sites. The chain length for our analysis was 1,000,000 generations with trees sampled every 100 generations using MCMC (Markov Chain Monte Carlo) analysis. The first 10,000 trees were discarded as burn-in for the tree topology and posterior probability. OTU representative sequences and full-length 18S rRNA gene sequences from reference taxa were analyzed to infer

OTU species' identities, and included a total of 92 taxa in our analysis. We further analyzed the full-length 18S rRNA molecule for the BrayaSø Greenland haptophyte and aligned this against a full-length haptophyte reference taxa dataset consisting of 220 taxa.

2.5 *Lipid extraction and analysis*

Alkenone extraction was after D'Andrea and Huang (2005). Alkenone samples came from the same sample used for DNA work. We freeze-dried and homogenized samples manually. We extracted samples with 9:1 Dichloromethane (DCM) : Methanol (MeOH) using an Accelerated Solvent Extractor ASE200 (Dionex, Sunnyvale, CA, USA). Extracts were separated into acid and neutral fractions using a solution of DCM:Isopropyl alcohol 2:1 (v/v). The neutral fraction was further separated into aliphatic (hexane elution), ketone (DCM), and alcohol (ethyl acetate: hexane 1:3) fractions using a flash silica gel column. The DCM fraction was analyzed using an Agilent 6890^{plus} Gas Chromatograph Flame Ionization Detector (GC-FID) for quantification. Organic compounds were verified by running on the GC-MS (Agilent 6890N Gas Chromatograph coupled to an Agilent 5973 Network Mass Selective Detector) with the same temperature program and chromatograms were compared to mass spectral data from previously reported standards and GC retention times (de Leeuw *et al.,* 1980; Marlowe *et al.,* 1984). LCA concentrations were determined from GC-FID analysis of the ketone fractions based on an internal C_{36} alkane standard.

3. Results

3.1 Phylogenetic relationships among alkenone-producing haptophytes

Ten out of fifteen lake sediments yielded positive amplification for haptophytes. Five lakes (Lake Qinghai, Brush Lake, Moon Lake, Mahoney Lake, and GaHai) yielded genomic DNA (DNA extracts >5 ng/μl; data not shown) but were negative for amplification using haptophyte-specific primers. Except Lake Qinghai, with a previously published alkenone profile (Li *et al.,* 1996), the lakes with no haptophyte DNA amplification also had below detection or convoluted alkenone profiles, indicative of low haptophyte cell numbers in the surface sediments. From the ten lake sediments with positive haptophyte DNA amplification, we identified eight OTUs at 95 % similarity from the partial 18S rRNA gene data. Individual lakes had multiple OTUs present, and the percent distribution for each lake is outlined in **Table 3a**.

We constructed a phylogeny using a representative sequence for each OTU along with previously published haptophyte 18S rRNA genes **(Fig. 2**). A posterior probability value of 1.0 supports the monophyly of the Isochrysidales order, in which all OTU sequences are grouped. The alkenone-producing haptophyte sequences clustered into three groups. Haptophyte Group I was comprised of representative sequences from OTUs 1, 2, 3, 4 and 5, and previously published sequences from the Greenland lakes (D'Andrea *et al.,* 2006), with strong posterior probability support (1.0). We also identified these Group I OTUs in lake sediments from BrayaSø (BS), Skoal (SK), Upper Murray (UM), Tso Ur downcore (TUd), Keluke (KE), and Pyramid (PY) lakes. Group II was comprised of OTU 6, OTU 7 and OTU 8

sequences, in addition to *C. lamellosa, I. litoralis*, *Dircrateria* sp. sequences, and sequences from Ace Lake, Antarctica (Coolen *et al.,* 2004). Sequences OTU 6, OTU 7 and OTU 8 appeared in Medicine (MD), George (LG), Skoal, Great Salt (GS), Pyramid, Keluke Hu and Tso Ur surface sediments (TUs). These sequences formed a sister group to marine species *E. huxleyi*, *G. oceanica*, and an unidentified marine coccoid haptophyte (U40924) comprise Group III.

A pairwise comparison among the eight OTU representative sequences and the Greenland haptophyte full-length sequence showed that OTU 1 and the Greenland Lake BrayaSø haptophyte sequence, as well as OTU 5 and the BrayaSø haptophyte sequence, were the most similar, with only 0.2 % differences (**Table 4a**). OTU 2 sequences were found in sediments from BrayaSø as well as Tibetan plateau Tso Ur, and OTU 5 sequences were found in BrayaSø, as well as Keluke Hu, Pyramid lake, and Upper Murray Lake. OTU 2 and OTU 6 representative sequences had the greatest distance, with 10.2 % difference. Notably, OTU 6 was found only in Tso Ur surface sediments and OTU 2 was found in Tso Ur ancient sediments.

According to the full-length 18S rRNA gene-based phylogenies (**Fig. 3**), the Greenland haptophyte (extracted from the BrayaSø water column) branched basal to a lineage of coastal/lacustrine haptophytes *I. galbana* and *C. lamellosa*, and shared common ancestry with marine species *E. huxleyi* and *G. oceanica* (posterior probability of 0.92). **Table 4b** shows the pairwise distances among full-length 18S rRNA sequences for the reference taxa and the BrayaSø water haptophyte. The BrayaSø haptophyte sequence shared the percent similarity with the *E. huxleyi* (AF184167) *and G. oceanica* (AJ246276) sequences, with only 1.9 % difference,

while it was the most divergent from *C. lamellosa* (AM490998) with 4.1 % difference.

3.2 Alkenone signatures of individual lake sediments

In order to numerically compare the alkenone distributions in different samples, we computed the relative abundances of C_{37} to C_{38} alkenones (Prahl *et al.,* 1988; Conte *et al.,* 1998; Chu *et al.,* 2005), the presence/absence of C_{38} methyl ketones (Conte *et al.,* 1994; Schulz *et al.,* 2000; Bendle *et al.,* 2005), the presence/absence of C_{40} compounds, and the percent concentration of $C_{37:4}$ relative to total C_{37} alkenones. Three exemplary alkenone signatures of individual lake sediments are presented in **Figure 4**, and corresponding C_{37}/C_{38} values are given in **Table 5b**. **Figure 4a** shows the alkenone gas chromatogram for Upper Murray lake surface sediments. These alkenones had a C_{37}/C_{38} ratio of 2.4, the presence of the C_{38} methyl ketone, and trace amounts of the C_{40} compound **(Table 5b)**. The presence of the C_{38} methyl ketone was similar to marine alkenone signatures from *E. huxleyi* and BrayaSø alkenone signatures (D'Andrea and Huang, 2005). Also like BrayaSø, the Upper Murray sediments had a high percentage of $C_{37:4}$ (63 %) and both lake had haptophyte sequences that branched with Group I (**Table 5**). **Figure 4b** shows the alkenone profile for Medicine Lake with a C_{37}/C_{38} ratio of 3.8, very low concentrations of C_{38} methyl ketones, and the presence of the C_{40} compound (**Table 5b**). The predominance of the ethyl ketone and the high C_{37}/C_{38} ratio was similar to *I. galbana* alkenone signatures, and accordingly, Medicine Lake haptophyte sequences branched with *I. galbana* sequences within Group II (**Table 5a**). **Figure 4c** depicts

the alkenone profile for Pyramid lake surface sediments, with a C_{37}/C_{38} ratio of 1.9, the presence of the methyl ketone and trace levels of C_{40} compounds (**Table 5b**). Pyramid lake sediments had a predominance of $C_{37:3}$ alkenones, and a percent $C_{37:4}$ of 22%, and DNA sequences branched with both Group I and Group II of the Isochrysidales.

The majority of the lake sediments analyzed in this study had a predominance of the tetraunsaturated C_{37} alkenone (**Table 5b**). Elevated percent concentrations of $C_{37:4}$ have been proposed as characteristics for lacustrine *Chrysotila* and *Isochrysis* species (Conte *et al.,* 1994), and have been suggested as a paleosalinity proxy, with low salinity corresponding to high percent $C_{37:4}$ (Liu *et al.,* 2008). Sediments from Great Salt Lake, Pyramid Lake, and Tso Ur had higher concentration of tri-unsaturated ketones. The lake sediments with the highest percentage of $C_{37:4}$ alkenones was Medicine lake, with an average salinity of 9, and the lake with the lowest percentage of $C_{37:4}$ alkenones was Pyramid Lake, NV, with a salinity of 5. Given that the salinity for the lakes in this study ranged from 0.04 to 270, our results therefore do not support a linear relationship between salinity and percent $C_{37:4}$ across multiple lakes and species.

3.3 Downcore alkenone and OTU distribution in BrayaSø and Tso Ur

Surface sediment alkenone signatures from both lakes BrayaSø and Tso Ur highly resembled ancient down-core sediment alkenone signatures from the lakes. BrayaSø alkenone signatures were characterized by a high tetra-unsaturated ($C_{37:4}$) alkenone concentration, the presence of the C_{38} methyl ketone, a C_{37}/C_{38} ratio of < 2,

and the absence of the C_{40} compounds in both surface and downcore sediments (D'Andrea and Huang, 2005). **Table 3b** shows the down-core variation in BrayaSø OTU sequence distribution. Spanning 20 cm to 100 cm depth in the core, or roughly 1,370 to 6,200 years BP, there was little change in the distribution of OTUs. As shown in **Table 3a**, the maximum percent dissimilarity among these OTU representative sequences was 6.5 % (OTU 4 and OTU 2), the minimum difference was 0.4 % (OTU 1 and OTU 5).

Tso Ur sediments were extracted at 0 cm and 20 cm core depth (age unknown). Tso Ur surface and ancient sediments were very similar with C_{37}/C_{38} values of < 2, the presence of C_{38} methyl ketones, and the presence of C_{40} compounds (**Table 5b**). Surprisingly, the OTU sequence distribution was very dissimilar, with sequences from Tso Ur surface sediments falling into Group II, and sequences from Tso Ur downcore sediments falling into Group I (**Table 3a**).

4. Discussion

4.1 OTU distribution and alkenone signatures

All haptophyte sequences in this study grouped within the order Isochrysidales. Previous tenet predicted distinct alkenone signatures according to marine or lacustrine species representation (Conte *et al.,* 1998; Prahl *et al.,* 1988; Chu *et al.,* 2005; Liu *et al.,* 2008). However, our results did not corroborate this pattern. Our results confirm our hypothesis that alkenone signatures alone do not adequately reflect the extent of haptophytes species diversity in various lakes: genetic data provided a more comprehensive measurement of the haptophyte species community.

Lakes Clear, Great Salt, George and Medicine contained haptophyte sequences that branched with haptophyte *Chrysotila* and *Isochrysis* species. Upper Murray lake also contained Group I representation. Pyramid and Skoal lakes yielded haptophyte sequences that branched into both Groups I and II, and TsoUr yielded Group I sequences in the ancient sediments and Group II sequences in the recent sediments.

We evaluated C_{37}/ C_{38} ratios and determined there was no clear pattern with phylogenetic placement. While lacustrine profiles often displayed a higher C_{37}/ C_{38} ratio than marine profiles, Conte *et al*. (1998) found an increase in C_{38} ethyl alkenones in comparison to C_{38} methyl alkenones during late logarithmic and stationary growth phase of *E. huxleyi*. We therefore conclude that C_{37}/ C_{38} ratios do not give an accurate indication of haptophyte species identity, and may reflect algal growth status at alkenone biosynthesis. The presence of the C_{38} methyl ketone was previously attributed to marine haptophyte presence (Schulz *et al*., 2000); with this dataset we show the presence of the C_{38} methyl ketone in multiple lakes with Group II haptophytes (e.g. Great Salt lake, Pyramid Lake), i.e. haptophytes more closely related to lacustrine *Isochrysis* and *Chrysotila* species. In all samples, the presence of the C_{38} methyl ketone was confirmed with mass spectral analysis, although the relative abundance of C_{38} methyl versus ethyl varied (data not shown). The C_{40} alkenone was found in trace amounts in four of our lakes. With the exception of ancient TsoUr sediments, the C_{40} compound was only found in sediments with haptophyte species branching among Group II. While not all Group II haptophytes appeared to produce C_{40} compounds, the presence of C_{40} may indicate the presence of a Group II haptophyte species. In the downcore sediments of TsoUr, the Group II

haptophytes (possible source for C_{40} alkenones) may be less abundant and were therefore undetected from older sediments.

One of the significant observations from this study was the occurrence of DNA from multiple haptophyte species in a single lake. Lakes Skoal and Pyramid, and Keluke Hu ancient sediments highlight the importance of verifying haptophyte species identities with genetic data, as opposed to alkenone signature alone. For these three sites, the alkenone signature may have been representative of a single haptophyte species, or the combined signature from two haptophyte species. It is also possible one or more of the haptophyte species does not contribute to the alkenone profile. Verifying the contributors to this alkenone profile is of great importance when selecting an appropriate $U_{37}^{K'}$ calibration for reconstructing water surface temperature. Likewise, both TsoUr recent and ancient sediments displayed very similar alkenone signatures. Upon DNA sequencing, it was revealed that sequences from two separated haptophyte groups were amplified from the each sample. These results could arise if 1) the haptophyte that dominates the alkenone signature (by producing large quantities of alkenones) is not the same as the haptophyte dominating DNA production or 2) environmental conditions, rather than phylogeny, determine alkenone lipid signatures. In contrast, Lake BrayaSø had both conserved alkenone signatures and haptophyte species representation over the past 6 ka, deeming this lake an excellent candidate for U_{37}^{K} based temperature reconstructions.

4.2. *Haptophyte diversity and genetic markers*

Sequence recovery varied among samples, and may be due to poor DNA preservation conditions *in situ*, or a poor cloning success rate. Our sequence-yield numbers therefore do not represent haptophyte DNA concentration, but number of successful sequences recovered. According to our partial 18S rRNA gene sequencing, no lake with > 5 sequences recovered contained a single dominant haptophyte species. The presence of multiple OTUs in single lakes suggested a significant degree of microheterogeneity in the lake haptophyte populations. Our haptophyte-specific primers targeted the most variable region of the haptophyte 18S rRNA gene. Incorporating a larger region of the 18S rRNA molecule enabled an improved support for the position of the novel Greenland Lake BrayaSø haptophyte species within the Isochrysidales (posterior probably increased from 0.66 to 0.92; **Figure 3**). Additional sequencing of the large subunit rRNA gene, plastid genes, or COI genes (Coolen *et al*., 2009) may further clarify poorly resolved haptophyte phylogenies.

4.3 *Geographical distribution and environmental controls*

The OTU distributions did not display a strong geographical signal. Most OTUs were represented in multiple lakes, e.g. OTUs 7 and 8 were present in more than half of the lakes analyzed. OTU 3 and OTU 4 only occurred in Lake BrayaSø and OTU 6 only occurred in TsoUr. Given the ubiquitous distribution of these haptophyte OTUs, haptophyte species' dispersal mechanisms must be able to achieve great distance. Whether by biotic or abiotic (winds, clouds) mechanisms, these haptophyte species may potentially encyst to survive transport. The haptophyte species *Prymnesium parvum* is known to form cysts (Pienaar, 1980), and given the

absence of haptophyte cells in the water column during the non-bloom periods in Greenland, these cells likely encyst or enter a vegetative stage in their life cycle. Marine aerosols can also transports organic compounds into the lake (Sicre *et al.,* 1990), but this would not explain the presence of haptophyte DNA, and therefore we assume the alkenones present in the lake are synthesized *in situ*.

The Greenland phylotype shared a common ancestor with the *Isochrysis*/*Chrysotila* group, which together formed a sister group with the marine haptophytes *E. huxleyi* and *G. oceanica*. The BrayaSø alkenone signatures were similar to traditional marine alkenone signatures (i.e. presence of C_{38Me}, absence of C_{40}, low C_{37}/C_{38}), with the exception that BrayaSø contained high tetraunsaturated alkenone concentrations. The conditions of the ancestral Lake BrayaSø species are unclear, as the lake was not formed by marine capture (Anderson and Bennike, 1997; D'Andrea and Huang, 2005) and there is no evidence for alkenone lipids in the first 2000 years of lake sediments. The full-length 18S rRNA sequence phylogeny showed that the BrayaSø haptophyte diverged before the lacustrine and coastal *I. galbana* and *C. lamellosa* species. According to this phylogeny, the last common ancestor to the BrayaSø haptophyte and *I. galbana* and *C. lamellosa* may have been marine. If we attribute the high tetraunsaturated alkenones in lake BrayaSø sediments to very cold temperature waters and/or low salinity, the BrayaSø alkenone signature has been conserved from its marine ancestor.

Lake BrayaSø formed approximately 8,000 years before present (BP) (McGowan *et al.,* 2003), and the occurrence of alkenones in the sediments began ~ 2,000 years later at a time of increased lake water salinity (McGowan *et al.,* 2003;

Anderson and Leng, 2004). Salinity has previously been suggested as a pre-requisite for alkenone-producing haptophyte populations (Rosell-Melé, 1998; Schulz *et al.*, 2000). Nine of ten lakes in our study were either oligosaline, brackish or alkaline and, notably, Upper Murray is a freshwater lake (**Table 1**). It is possible that the alkenone-producing haptophytes require a saline environment for survival, or perhaps the haptophyte communities are present but only produce alkenones when the water reaches a certain salinity threshold. Salinity gradient culture experiments using novel haptophyte species would help clarify this assertion.

5. Conclusions

Haptophyte communities in lakes can be precisely determined using genetic fingerprints preserved in lake sediments and collected from the water column during algal growing seasons. Alkenone profiles, however, are not a reliable way to determine the haptophyte species in lakes: similar alkenone profiles may be produced by different species, whereas different alkenone profiles may be derived from similar haptophyte species. We identified eight operational taxonomic units (OTUs) from a series of ten lakes, and generated a haptophyte phylogeny that further resolved the evolutionary relationships of alkenone-producing haptophytes. Although different OTUs do not necessarily imply differential alkenone-temperature sensitivity, the correct identification of haptophyte community members is critical for ultimately selecting the appropriate U^K_{37} calibration for both sea and lake surface temperature reconstructions. This study examined surface sediments in an order to capture the most-recent members of the haptophyte community; however, if there were any

degradation between the overlying water column and the sediments, these haptophyte members would not be identified. The best alkenone-temperature calibration for individual lacustrine haptophyte species will eventually require isolation and culturing of algae under laboratory-controlled conditions.

Our lacustrine haptophyte sequences branched within the Isochrysidales, as expected for alkenone-producers, and none of our representative sequences branched directly with marine species *E. huxleyi* and *G. oceanica*. However, traditional lacustrine vs. marine alkenone profile trends were not conserved in our lacustrine dataset: C_{37}/C_{38} ratios ranged from 0.9 to 3.8, therefore spanning both the expected marine and lake ratios, and C_{38} methyl ketones were almost ubiquitously present, previously associated with marine profiles. The most prominent feature of lacustrine alkenones relative to marine alkenones is the predominance of the tetraunsaturated C37 alkenones, which may have resulted from low salinity in lakes. Further research, preferably using algal cultures, will be needed to determine if low salinity conditions influence alkenone biosynthesis.

Sequencing the full-length 18S rRNA molecule sequence improved the resolution of the phylogenetic placement of the novel Greenland BrayaSø haptophyte species: the BrayaSø haptophyte shared a common ancestry with marine species *E. huxleyi* and *G. oceanica*, and branched basal to coastal/lacustrine species *C. lamellosa* and *I. galbana*. This result is not surprising given the similarity in alkenone signatures between the Greenland species and marine haptophytes, both with the presence of a C_{38} methyl ketone and the absence of a C_{40} ketone. The Greenland down-core samples have a well-conserved alkenone signature through time, although we detected five

separate OTUs in these samples. All OTUs branch into Group I, suggesting the species are closely-related, if not subspecies or single species with microheterogeneity in their copies of 18S rRNA genes. This result suggests that when species are closely-related, a single U_{37}^{K} calibration is adequate for paleoclimate reconstructions, which can be achieved by performing an *in situ* calibration of the alkenone unsaturation index (D'Andrea, 2008; Toney *et al.,* 2009), or culture growth under controlled conditions. A continued appraisal of haptophyte species distributions, and the environmental constraints on alkenone production in a global diversity of lakes will greatly enhance our understanding of haptophyte community success and environmental controls on alkenone biosynthesis.

Acknowledgements

We would like to thank Marcelo Alexandre, Rafael Tarozo, and the members of the W. M. Keck Ecological and Evolutionary Genetics Facility for lab support. We would also like to thank Ian Walker for contributing Mahoney Lake samples, and Xiangdong Yang for the Tibetan plateau samples. We are grateful for financial support from the US National Science Foundation grant OPP 0520718 to Y. Huang, and for graduate student support to S. Theroux provided by the Brown University Strominger Fellowship.

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., Lipman, D. J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389-3402.

Anderson, N. J., Bennike, O., 1997. Holocene lake sediments in West Greenland and their palaeoclimatic and palaeoecological implications Geology of Greenland Survey Bulletin 176, 89-94.

Anderson, N. J., Leng, M. J., 2004. Increased aridity during the early Holocene in West Greenland inferred from stable isotopes in laminated-lake sediments. Quaternary Science Reviews 23, 841-849.

Ashelford, K. E., Chuzhanova, N. A., Fry, J. C., Jones, A. J., Weightman, A. J., 2006. New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. Appl. Environ. Microbiol. 72, 5734-5741.

Bendle, J., Rosell-Mele, A., Ziveri, P., 2005. Variability of unusual distributions of alkenones in the surface waters of the Nordic seas. Paleoceanography 20. doi 10.1029/2004pa001025.

Besonen, M. R., Patridge, W., Bradley, R. S., Francus, P., Stoner, J. S., Abbott, M. B., 2008. A record of climate over the last millennium based on varved lake sediments from the Canadian High Arctic. Holocene 18, 169-180.

Brassell, S. C., Eglinton, G., Marlowe, I. T., Pflaumann, U., Sarnthein, M., 1986. Molecular stratigraphy - a new tool for climatic assessment. Nature 320, 129-133.

Chu, G. Q., Sun, Q., Li, S. Q., Zheng, M. P., Jia, X. X., Lu, C. F., Liu, J. Q., Liu, T. S., 2005. Long-chain alkenone distributions and temperature dependence in lacustrine surface sediments from China. Geochimica Et Cosmochimica Acta 69, 4985-5003.

Conte, M. H., Sicre, M. A., Ruhlemann, C., Weber, J. C., Schulte, S., Schulz-Bull, D., Blanz, T., 2006. Global temperature calibration of the alkenone unsaturation index $U^{k'}_{37}$ in surface waters and comparison with surface sediments. Geochemistry Geophysics Geosystems 7. doi:10.1029/2005gc001054.

Conte, M. H., Thompson, A., Lesley, D., Harris, R. P., 1998. Genetic and physiological influences on the alkenone/alkenoate versus growth temperature relationship in *Emiliania huxleyi* and *Gephyrocapsa oceanica*. Geochimica Et Cosmochimica Acta 62, 51-68.

Conte, M. H., Volkman, J. K., Eglinton, G. 1994. Lipid biomarkers of the Haptophyta. *In*: Green, J. C. & Leadbetter, B. S. C. (eds.) The Haptophyte Algae. Clarendon Press. 351-377.

Coolen, M. J. L., Muyzer, G., Rijpstra, W. I. C., Schouten, S., Volkman, J. K., Sinninghe Damsté J., 2004. Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake. Earth and Planetary Science Letters 223, 225-239.

Coolen, M. J. L., Saenz, J. P., Giosan, L., Trowbridge, N. Y., Dimitrov, P., Dimitrov, D., Eglinton, T. I., 2009. DNA and lipid molecular stratigraphic records of haptophyte succession in the Black Sea during the Holocene. Earth and Planetary Science Letters 284, 610-621.

D'Andrea, W. J. 2008. Development and application of lacustrine alkenone paleothermometry in southwestern Greenland. PhD dissertation. Brown University, Providence. 197.

D'Andrea, W. J., Huang, Y., 2005. Long chain alkenones in Greenland lake sediments: Low $\delta^{13}C$ values and exceptional abundance. Organic Geochemistry 36, 1234-1241.

D'Andrea, W. J., Lage, M., Martiny, J. B. H., Laatsch, A. D., Amaral-Zettler, L. A., Sogin, M. L., Huang, Y. S., 2006. Alkenone producers inferred from well-preserved 18S rDNA in Greenland lake sediments. Journal of Geophysical Research-Biogeosciences 111. doi: 10.1029/2005JG000121.

de Leeuw, J. W., van der Meer, F. W., Rijpstra, W. I. C., Schenck, P. A., 1980. On the occurrence and structural identification of long chain unsaturated ketones and hydrocarbons in sediments. Physics and Chemistry of the Earth 12, 211-217.

de Vargas, C., Aubry, M. P., Probert, I., Young, J. 2007. Origin and evolution of coccolithophores: From coastal hunters to oceanic farmers. *In*: Falkowski, P. & Knoll, A. H. (eds.) Evolution of Aquatic Photoautotrophs. Elsevier Academic, New York. 251-281.

Donovan, J. J., 1994. On the measurement of reactive mass fluxes in evaporative groundwater-source lakes. Symp. Sedimentology and geochemistry of modern and ancient saline lakes. SEPM Special Publication 50, 33-50.

Ewing, B., Green, P., 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 8, 186-194.

Ewing, B., Hillier, L., Wendl, M. C., Green, P., 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 8, 175-185.

Galat, D. L., Lider, E. L., Vigg, S., Robertson, S. R., 1981. Limnology of a large, deep, North American terminal lake, Pyramid Lake, Nevada, U.S.A. Hydrobiologia 81-82, 281-317.

Gordon, D., Abajian, C., Green, P., 1998. Consed: A graphical tool for sequence finishing. Genome Res. 8, 195-202.

LaMontagne, M. G., Michel, F. C., Holden, P. A., Reddy, C. A., 2002. Evaluation of extraction and purification methods for obtaining PCR-amplifiable DNA from compost for microbial community analysis. J. Microbiol. Methods 49, 255-264.

Lanave, C., Preparata, G., Saccone, C., Serio, G., 1984. A new method for calculating evolutionary substitution rates. J. Mol. Evol. 20, 86-93.

Lane, D. J. 1991. 16s/23s rRNA sequencing. *In*: Stackerbrandt, E. G., M. (ed.). Wiley, New York.115-175.

Lasek-Nesselquist, E., Bogomolni, A. L., Gast, R. J., Welch, D. M., Ellis, J. C., Sogin, M. L., Moore, M. J., 2008. Molecular characterization of *Giardia intestinalis* haplotypes in marine animals: variation and zoonotic potential. Dis. Aquat. Organ. 81, 39-51.

Li, J. G., Philp, R. P., Pu, F., Allen, J., 1996. Long-chain alkenones in Qinghai Lake sediments. Geochimica Et Cosmochimica Acta 60, 235-241.

Liu, W. G., Liu, Z. H., Fu, M. Y., An, Z. H., 2008. Distribution of the C_{37} tetra-unsaturated alkenone in Lake Qinghai, China: A potential lake salinity indicator. Geochimica Et Cosmochimica Acta 72, 988-997.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T., Steppi, S., Jobb, G., Forster, W., Brettske, I., Gerber, S., Ginhart, A. W., Gross, O., Grumann, S., Hermann, S., Jost, R., Konig, A., Liss, T., Lussmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A., Schleifer, K. H., 2004. ARB: a software environment for sequence data. Nucleic Acids Res. 32, 1363-1371.

Maddison, D. R., Maddison, W. P. (2000) MacClade version 4: Analysis of phylogeny and character evolution. Sunderland, Massachusetts, Sinauer Associates.

Marlowe, I. T., Green, J. C., Neal, A. C., Brassell, S. C., Eglinton, G., Course, P. A., 1984. Long-Chain (**n**-C_{37}-C_{39}) alkenones in the prymnesiophyceae. Distribution of alkenones and other lipids and their taxonomic significance. British Phycological Journal 19, 203-216.

McGowan, S., Ryves, D. B., Anderson, N. J., 2003. Holocene records of effective precipitation in West Greenland. Holocene 13, 239-249.

Medlin, L., Elwood, H., Stickel, S., Sogin, M., 1988. The characterization of enzymatically amplified eukaryotic 16S-like rRNA coding regions. Gene 71, 491-499.

Medlin, L. K., Saez, A. G., Young, J. R., 2008. A molecular clock for coccolithophores and implications for selectivity of phytoplankton extinctions across the K/T boundary. Marine Micropaleontology 67, 69-86.

Morgan, D. L., 1947. The Great Salt Lake. University of Utah Press, Salt Lake City.

Müller, P. J., Kirst, G., Ruhland, G., von Storch, I., Rosell-Mele, A., 1998. Calibration of the alkenone paleotemperature index $U^{k'}_{37}$ based on core-tops from the eastern South Atlantic and the global ocean (60°N-60 °S). Geochimica Et Cosmochimica Acta 62, 1757-1772.

Overmann, J., Beatty, J. T., Hall, K. J., Pfennig, N., Northcote, T. G., 1991. Characterization of a Dense, Purple Sulfur Bacterial Layer in a Meromictic Salt Lake. Limnol. Oceanogr. 36, 846-859.

Pearson, E. J., Juggins, S., Farrimond, P., 2008. Distribution and significance of long-chain alkenones as salinity and temperature indicators in Spanish saline lake sediments. Geochimica Et Cosmochimica Acta 72, 4035-4046.

Pienaar, R. N., 1980. Observations on the structure and composition of the cyst of *Prymensium* (Prymnesiophyceae). Proceedings of the Electron Microscopy Society of Southern Africa 10, 73-74.

Prahl, F. G., Muehlhausen, L. A., Zahnle, D. L., 1988. Further evaluation of long-chain alkenones as indicators of paleoceanographic conditions. Geochimica Et Cosmochimica Acta 52, 2303-2310.

Prahl, F. G., Wakeham, S. G., 1987. Calibration of unsaturation patterns in long-chain ketone compositions for paleotemperature assessment. Nature 330, 367-369.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W. G., Peplies, J., Glockner, F. O., 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. 35, 7188-7196.

Rodriguez, R., Oliver, J. L., Marin, A., Medina, J. R., 1990. The general stochastic model of nucleotide substitution. J. Theor. Biol. 142, 485-501.

Ronquist, F., Huelsenbeck, J. P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

Rontani, J. F., Beker, B., Volkman, J. K., 2004. Long-chain alkenones and related compounds in the benthic haptophyte *Chrysotila lamellosa* Anand HAP 17. Phytochemistry 65, 117-126.

Rosell-Melé, A., 1998. Interhemispheric appraisal of the value of alkenone indices as temperature and salinity proxies in high-latitude locations. Paleoceanography 13, 694-703.

Schloss, P. D., Handelsman, J., 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl. Environ. Microbiol. 71, 1501-1506.

Schulz, H.-M., Schöner, A., Emeis, K.-C., 2000. Long-chain alkenone patterns in the Baltic sea--an ocean-freshwater transition. Geochimica Et Cosmochimica Acta 64, 469-477.

Sicre, M. A., Marty, J. C., Saliot, A., 1990. n-Alkanes, fatty acid esters, and fatty acid salts in size fractionated aerosols collected over the Mediterranean Sea. J. Geophys. Res. 95, 3649-3657.

Simon, N., Campbell, L., Örnolfsdottir, E., Groben, R., Guillou, L., Lange, M., Medlin, L. K., 2000. Oligonucleotide probes for the identification of three algal groups by dot blot and fluorescent whole-cell hybridization. The Journal of Eukaryotic Microbiology 47, 76-84.

Sun, Q., Chu, G. Q., Liu, G. X., Li, S., Wang, X. H., 2007. Calibration of alkenone unsaturation index with growth temperature for a lacustrine species, *Chrysotila lamellosa* (Haptophyceae). Organic Geochemistry 38, 1226-1234.

Tavaré, S., 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Lectures on Mathematics in the Life Sciences 17, 57-86.

Toney, J. L., Huang, Y., Fritz, S. C., Baker, P. A., Grimm, E., Nyren, P., 2009. Climatic and environmental controls on the occurrence and distributions of long chain alkenones in lakes in the interior United States. Geochimica Et Cosmochimica Acta. 74, 10.1016/j.gca.2009.11.021

Valero-Garces, B. L., Kelts, K. R., 1995. A sedimentary facies model for perennial and meromictic saline lakes - Holocene Medicine Lake Basin, South-Dakota, USA. Journal of Paleolimnology 14, 123-149.

Versteegh, G. J. M., Riegman, R., de Leeuw, J. W., Jansen, J. H. F., 2001. $U^{k'}_{37}$ values for *Isochrysis galbana* as a function of culture temperature, light intensity and nutrient concentrations. Organic Geochemistry 32, 785-794.

Volkman, J. K., Barrett, S. M., Blackburn, S. I., Sikes, E. L., 1995. Alkenones in *Gephyrocapsa oceanica* - Implications for Studies of Paleoclimate. Geochimica Et Cosmochimica Acta 59, 513-520.

Volkman, J. K., Eglinton, G., Corner, E. D. S., Forsberg, T. E. V., 1980. Long-Chain Alkenes and Alkenones in the Marine Coccolithophorid *Emiliania-Huxleyi*. Phytochemistry 19, 2619-2622.

Yurkova, N., Rathgeber, C., Swiderski, J., Stackebrandt, E., Beatty, J. T., Hall, K. J., Yurkov, V., 2002. Diversity, distribution and physiology of the aerobic phototrophic bacteria in the mixolimnion of a meromictic lake. FEMS Microbiol. Ecol. 40, 191-204.

Zink, K. G., Leythaeuser, D., Melkonian, M., Schwark, L., 2001. Temperature dependency of long-chain alkenone distributions in Recent to fossil limnic sediments and in lake waters. Geochimica Et Cosmochimica Acta 65, 253-265.

Table 1. Lake locations, maximum depth and salinity.

Name	Location	Lat	Long	Max. depth (m)	Salinity (ppt)	Reference
BrayaSø	Greenland	66.99	-51.01	22.4	0.9-1.9	D'Andrea et al., 2006
Brush Lake	Montana, USA	48.60	-104.11	18	3	Donovan, 1994; Zink et al., 2001; Toney et al., 2009
Clear Lake	North Dakota, USA	47.86	-100.10	1.7	5	Toney et al., 2009
GaHai	Tibetan plateau, China	37.13	97.57	10.5	-	http://www.homepages.ucl.ac.uk/~ucfasdt/Tibet/index.htm
Great Salt Lake	Utah, USA	41.17	-112.58	10	270	Morgan, 1947
Keluke Hu	Tibetan plateau, China	37.28	96.89	8	-	http://www.homepages.ucl.ac.uk/~ucfasdt/Tibet/index.htm
Lake George	North Dakota, USA	46.74	-99.49	18.8	10	Toney et al., 2009
Lake Qinghai	Tibetan plateau, China	36.53	99.60	27	15	Li et al., 1996; Liu et al., 2008
Mahoney Lake	British Columbia, Canada	49.00	-119.00	17.8	13240	Overmann et al., 1991; Yurkova et al., 2002
Medicine Lake	South Dakota, USA	44.82	-97.35	8.3	9	Valero-Garces and Kelts, 1995; Toney et al., 2009
Moon Lake	North Dakota, USA	46.86	-98.16	8.2	3	Toney et al., 2009
Pyramid Lake	Nevada, USA	40.00	-119.50	103	5	Galat et al., 1981
Skoal Lake	North Dakota, USA	47.92	-101.47	1.8	8	Toney et al., 2009
Tso Ur	Tibetan plateau, China	31.48	91.52	8.4	-	http://www.homepages.ucl.ac.uk/~ucfasdt/Tibet/index.htm
Upper Murray	Elesmere Island, Canada	81.33	-69.50	82	0.04-0.06	Besonen et al., 2008

Table 2. Primers used for SSU rRNA gene amplification.

Universal eukaryotic primer	5'→3'	T_m (°C)	Reference
1391R	GAC GGG CGG TGT GTR CA	57	Lane, 1991
EukA 5'F	AC CTG GTT GAT CCT GCC AGT	57	Medlin et al., 1988
EukB 3'R	C TTC TGC AGG TTC ACC TAC	53	Medlin et al., 1988
Haptophyte-specific primers			
Prym429F	GCG CGT AAA TTG CCC GAA	65	Coolen et al., 2004
Prym887R (aka PRYM02)	GGA ATA CGA GTG CCC CTG AC	62	Simon et al., 2000
Greenland haptophyte-specific			
GL625R	G CTA AGA GTA GGC GCG GTC TTT GGG AGG	74	this work

Table 3. Percent OTU representation in lake sediment clone libraries. (A) Surface sediment distribution; (B) Greenland Lake BrayaSø downcore sediment relative OTU distribution. Approximate age given (Anderson and Leng, 2004). (+) denotes presence of OTU but a low confidence value for percentage.

A.

	Group I					Group II				
	OTU 1	OTU 2	OTU 3	OTU 4	OTU 5	OTU6	OTU 7	OTU 8	Sequences recovered	Symbol
BrayaSo, Greenland (total)	50%	1%	1%	2%	47%	-	-	-	187	BS
Clear lake, ND, USA	-	-	-	-	-	-	-	+	<5	CL
Great Salt Lake, UT, USA	-	-	-	-	-	-	+	-	<5	GS
Keluke Hu downcore, TP, CN	-	-	-	-	60%	-	20%	20%	5	KE
Lake George, ND, USA	-	-	-	-	-	-	53%	47%	62	LG
Medicine Lake, SD, USA	-	-	-	-	-	-	40%	60%	5	MD
Pyramid Lake, NV, USA	-	-	-	-	29%	-	57%	14%	14	PY
Skoal Lake, ND, USA	6%	-	-	-	-	-	63%	31%	16	SK
Tso Ur downcore, TP, CN	37%	63%	-	-	-	-	-	-	94	TUd
Tso Ur surface, TP, CN	-	-	-	-	-	4%	38%	58%	24	TUs
Upper Murray Lake, EI, CA	75%	-	-	-	25%	-	-	-	24	UM

B.

	OTU 1	OTU 2	OTU 3	OTU 4	OTU 5	Sequences recovered	Age (^{14}C yrs BP)
BrayaSo core 24 cm	53%	-	-	1%	46%	81	1184
BrayaSo core 40 cm	51%	1%	-	1%	47%	77	2039
BrayaSo core 50 cm	56%	-	-	-	44%	73	2574
BrayaSo core 70 cm	26%	-	3%	1%	71%	147	3643
BrayaSo core 88 cm	-	-	-	-	+	<5	4606
BrayaSo core 110 cm	45%	-	-	6%	49%	83	5783

Table 4: Uncorrected distance matrices of pairwise comparisons among sequences. (A) Representative OTU sequences and Greenland BrayaSø water sequence; (B) Reference taxa and Greenland BrayaSø water sequence. Numbers are % difference.

	OTU 1	BrayaSø water	OTU 2	OTU 3	OTU 4	OTU 5	OTU 6	OTU 7	OTU 8
OTU 1	-	0.2	4.8	2.8	2.6	0.4	6.7	4.3	5.2
BrayaSø water	-	-	4.6	2.6	2.4	0.2	6.5	4.1	5
OTU 2	-	-	-	2.4	6.5	4.3	10.2	8	9.1
OTU 3	-	-	-	-	4.5	2.4	8.7	6.3	7.1
OTU 4	-	-	-	-	-	2.2	8.4	6	6.9
OTU 5	-	-	-	-	-	-	6.3	3.9	4.8
OTU 6	-	-	-	-	-	-	-	3.2	5
OTU 7	-	-	-	-	-	-	-	-	2.2
OTU 8	-	-	-	-	-	-	-	-	-

Table 5. Alkenone trends among haptophyte algae. (A) Reference taxa; (B) lakes reported in this study.

	C_{37}/C_{38}	C_{38} MeK	C_{40}	Phylogenetic group	$U_{37}^{K'}$ calibration	% $C_{37:4}$ *
A. Coastal/Lacustrine						
Chysotila lamellosa (marine)	> 2	absent	present	Group II	Uk37' =0.01T+0.0033[a]	
					Uk37' = 0.0011T^2 -	
Chysotila lamellosa (lacustrine)	> 2	absent	present	Group II	0.0157T + 0.1057[b]	
Isochrysis galbana	> 3	absent	present	Group II	Uk37' = 0.00932 - 0.0413[c]	
Ace Lake, Antarctica	> 2	absent	present	Group II	NR[d]	
Lake BrayaSo, Greenland	1.3	present	absent	Group I	Uk37' = (T + 5.4)/155.2[e]	65
Marine						
Emiliania huxleyi	1-2	present	absent	Group III	Uk37' = 0.033T + 0.043[f]	
Gephyrocapsa oceanica	< 1	present	absent	Group III	Uk37' = 0.049T - 0.52[g]	
B. This work						
Clear lake, North Dakota	> 2	-	-	Group II		-
Great Salt Lake, Utah	1.9	present	trace?	Group II		31
Keluke Hu downcore, Tibetan Plateau	0.9	-	trace	Group I, II		-
Lake George, North Dakota	2.0	present	trace	Group II		55
Medicine Lake, South Dakota	3.8	absent	trace?	Group II		67
Pyramid Lake, Nevada	1.9	present	trace?	Group I, II		22
Skoal Lake, North Dakota	2.2	-	-	Group I, II		-
Tso Ur surface, Tibetan Plateau	1.7	present	trace	Group II		30
Tso Ur downcore, Tibetan Plateau	1.8	present	trace	Group I		27
Upper Murray Lake, Ellesmere Island	2.4	present	absent	Group I		63

* $\%C_{37:4} = C_{37:4}/(C_{37:2} + C_{37:3} + C_{37:4})$

a. Conte et al., 1994; Rontani et al., 2004

b. Sun et al., 2007

c. Marlowe et al., 1984; Conte et al., 1994; Versteegh et al., 2001

d. Coolen et al., 2004

e. D'Andrea and Huang, 2005; D'Andrea, 2008

f. Volkman et al., 1980; Chu et al., 2005; Liu et al., 2008; Prahl and Wakeman, 1987

g. Volkman, 1995

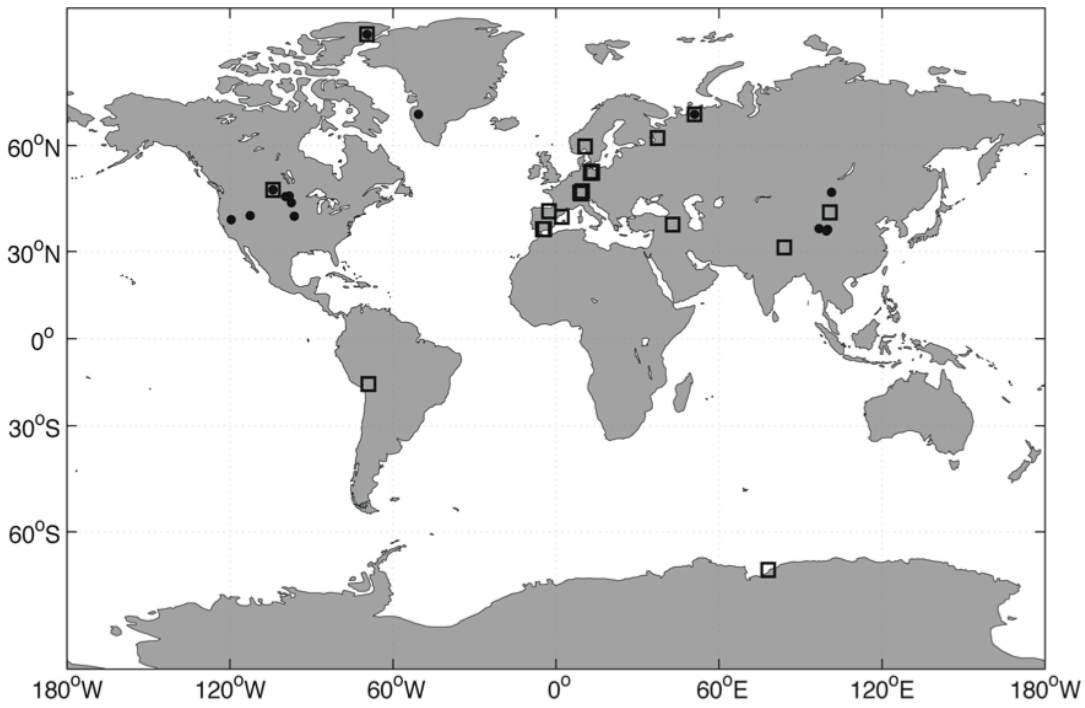

Figure 1. Map of alkenone-containing lakes. Squares denote previously reported lakes, circles denote lakes analyzed in this study.

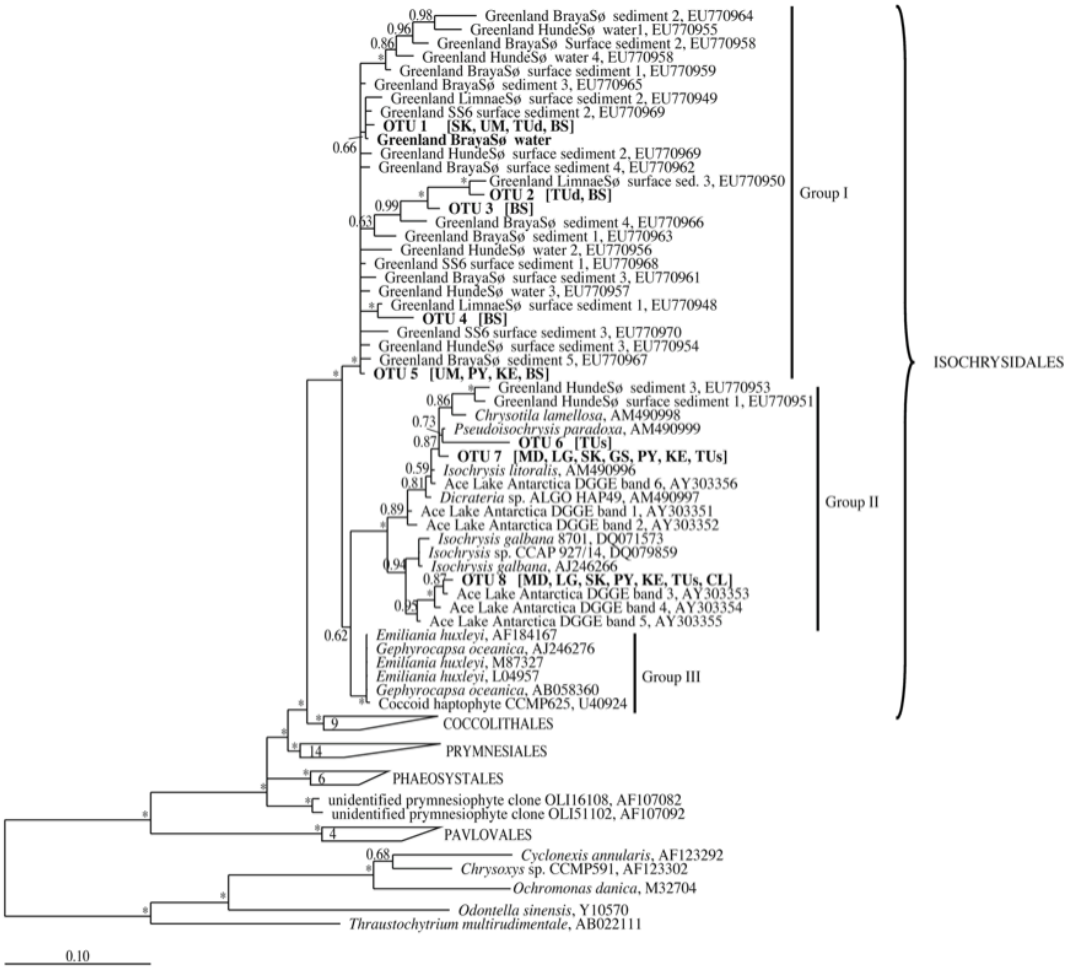

Figure 2. A consensus Bayesian phylogenetic tree depicting 18S rRNA gene-inferred relationships among haptophyte algae. An asterisk (*) indicates posterior probability values of 1.00; all other values as shown. Bold designates sequences from this study. Operational Taxonomic Units (OTUs) were defined at a 95 % similarity level. Order classification after de Vargas *et al.* (2007), with number of sequences per order as indicated. The evolutionary distance for the number of changes per site is represented by the scale bar. GenBank accession numbers follow species and sequence names. Lakes with given OTU representation shown in brackets; abbreviations as shown in Table 3.

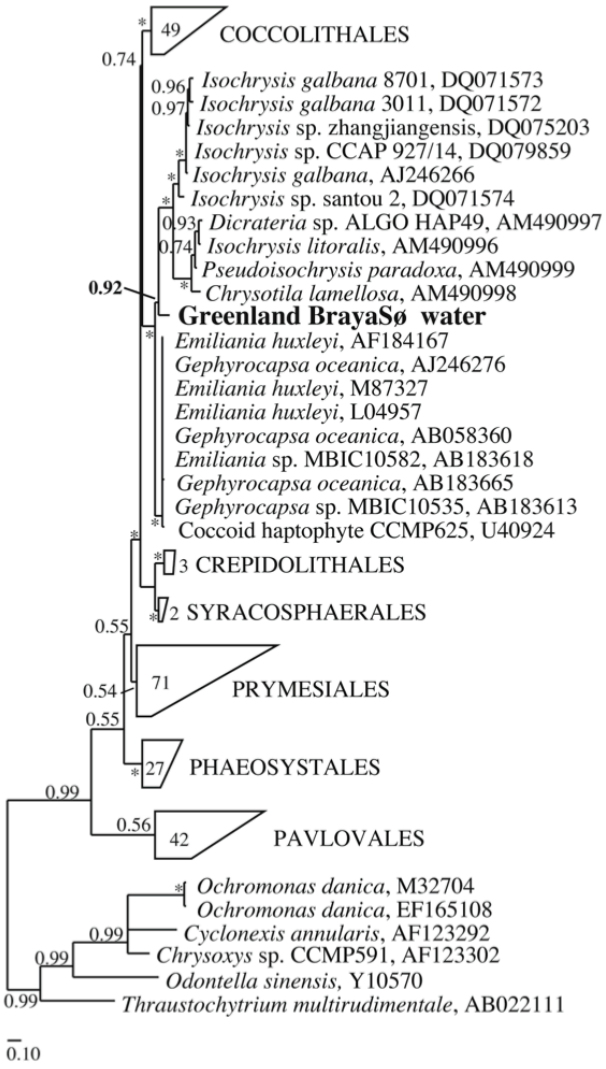

Figure 3. A consensus Bayesian phylogenetic tree depicting 18S rRNA gene-inferred relationships among haptophyte algae using only full-length sequences. Note improved support for the placement of the Greenland BrayaSø sequence. Operational Taxonomic Units (OTUs) defined at a 95 % similarity level. An asterisk (*) indicates posterior probability values of 1.00; all other values as shown. The evolutionary distance for the number of changes per site is represented by the scale bar. GenBank accession numbers follow species names and sequence names.

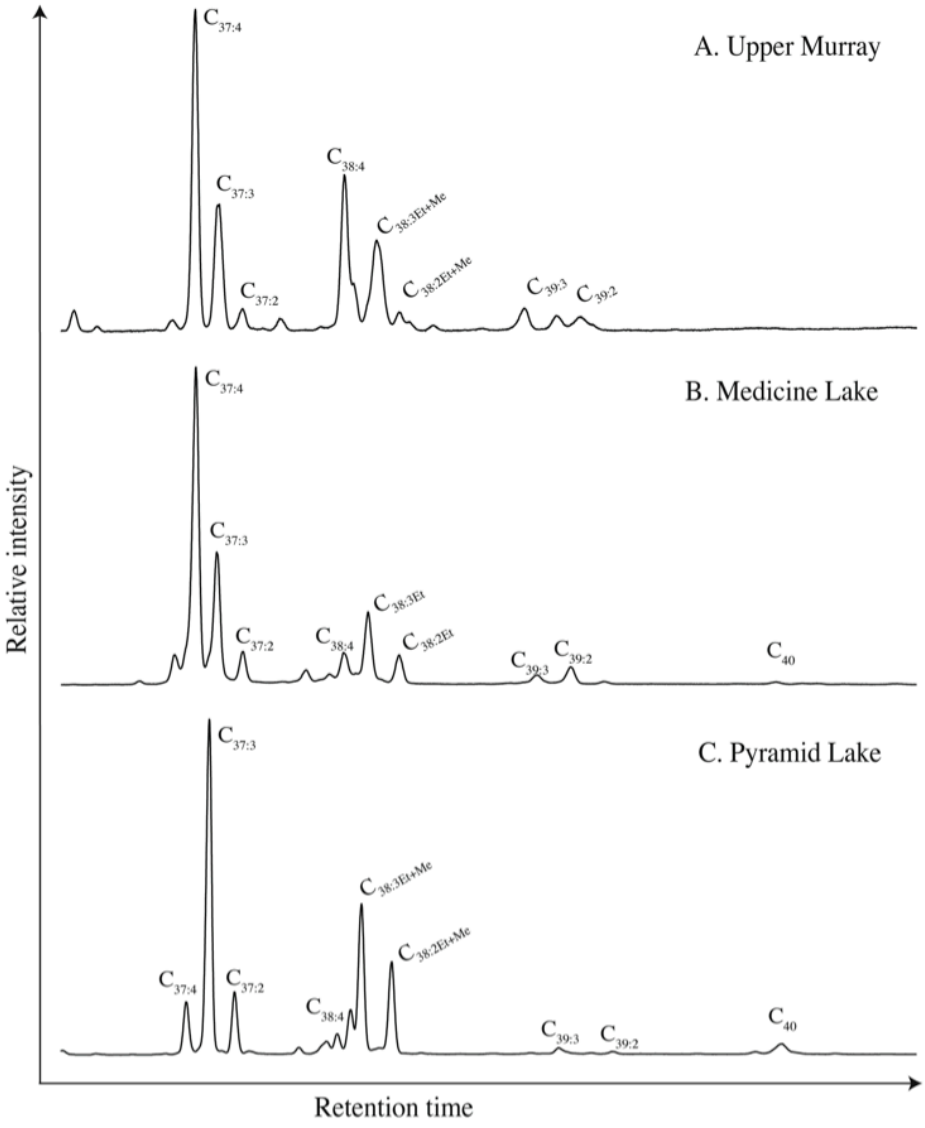

Figure 4. Lake sediment alkenone signatures. A) Upper Murray Lake Group I example, B) Medicine Lake Group II example, C) Pyramid lake with both Group I and Group II sequences.

CHAPTER 3

COMPARATIVE MOLECULAR MICROBIAL ECOLOGY OF A SPRING HAPTOPHYTE BLOOM IN A GREENLAND ARCTIC OLIGOSALINE LAKE

SUSANNA THEROUX[1,2]

LINDA AMARAL-ZETTLER[1,2,3]

YONGSONG HUANG[1]

[1] BROWN UNIVERSITY, *Department of Geological Sciences*

[2] MARINE BIOLOGICAL LABORATORY, *Josephine Bay Paul Center*

[3] BROWN UNIVERSITY, *Department of Ecology and Evolutionary Biology*

In review:

Frontiers in Microbiology

Abstract

The Arctic is highly sensitive to increasing global temperatures and is projected to experience dramatic ecological shifts in the next few decades. Oligosaline lakes are common in arctic regions where evaporation surpasses precipitation, however these extreme microbial communities are poorly characterized. Many oligosaline lakes, in contrast to freshwater ones, experience annual blooms of haptophyte algae that generate valuable alkenone biomarker records that can be used for paleoclimate reconstruction. These haptophyte algae are globally important, and globally distributed, aquatic phototrophs yet their presence in microbial molecular surveys is scarce. To target haptophytes in a molecular survey, we compared microbial community structure during two haptophyte bloom events in an arctic oligosaline lake, Lake BrayaSø in southwestern Greenland, using high-throughput pyrotag sequencing. Our comparison of two annual bloom events yielded surprisingly low taxon overlap, only 13% for bacterial and 26% for eukaryotic communities, which indicates significant annual variation in the underlying microbial populations. Both the bacterial and eukaryotic communities strongly resembled high-altitude and high-latitude freshwater environments. In spite of high alkenone concentrations in the water column, and corresponding high haptophyte rRNA gene copy numbers, haptophyte pyrotag sequences were not the most abundant eukaryotic tag, suggesting that sequencing biases obscured relative abundance data. With over 170 haptophyte tag sequences, we observed only one haptophyte algal Operational Taxonomic Unit, a prerequisite for accurate paleoclimate reconstruction from the lake sediments. Our

study is the first to examine microbial diversity in a Greenland lake using next generation sequencing and the first to target an extreme haptophyte bloom event. Our results provide a context for future explorations of aquatic ecology in the warming arctic.

1. Introduction

Oligosaline lakes (salinity 0.5-5 ppt) develop in polar regions near ice sheets where evaporation exceeds precipitation and provide a unique habitat apart from the more common glacially-derived freshwater lakes. These high latitude lakes serve as sensitive indicators of the ecosystem response to global climate change (Marchetto, 2004; Quayle *et al.,* 2002) as their low salinity reflects small changes in hydrological balance. In the past decade alone, southwestern Greenland has undergone marked warming, and major warming is predicted for the future (Bennike *et al.,* 2010). Microbial surveys targeting the 18S ribosomal RNA (rRNA) gene have revealed previously unknown diversity in microbial eukaryotes lineages such as cryptomonads, katablepharids, dinoflagellates and Perkinsea (Logares *et al.,* 2007; Slapeta, Moreira, and Lopez-Garcia, 2006; Shalchian-Tabrizi *et al.,* 2011). However, haptophyte algae have been largely absent from these studies, potentially the result of naturally low haptophyte abundances in the environments selected such as deep sea habitats or anoxic lakes (Stoeck *et al.,* 2009; Stoeck *et al.,* 2010; Edgcomb *et al.,* 2011; Pawlowski *et al.,* 2011; Shalchian-Tabrizi *et al.,* 2011). The GC-rich haptophyte genomes may also hinder amplification reactions that use universal primer sets (Moon-van der Staay *et al.,* 2001; Liu *et al.,* 2009; Stoeck *et al.,* 2010). In this study, we targeted the haptophyte-rich waters of an arctic oligosaline lake spring bloom event to shed light on the microbial diversity of these unique ecosystems.

Lake BrayaSø in southwestern Greenland experiences a seasonal haptophyte bloom approximately two weeks after ice off (D'Andrea *et al.,* 2011). These haptophyte blooms result in exceptional abundances of alkenones in BrayaSø

sediments (82 mg/g total organic carbon, D'Andrea *et al*., 2005) that provide the first quantitative temperature record for the past five thousand years for southwestern Greenland (D'Andrea *et al*., 2011). Only a few species of haptophyte algae, in the order Isochrysidales, produce alkenone lipids. These species and their alkenone lipids have been extensively studied in marine environments, where alkenones are preserved in marine sediments as a record of sea surface temperature back through time (Brassell et al, 1986; Conte et al, 2006; Marlowe *et al*., 1984; Prahl and Wakeham, 1987; Volkman et al, 1980; Muller et al, 1998). The endeavor to extend this alkenone-based proxy to the continents has resulted in pan-continental surveys of lake sediments and waters for alkenone-producing haptophyte algae. The use of haptophyte-specific primers targeting environmental DNA has revealed considerable diversity in lake-dwelling haptophytes (Coolen *et al*., 2004; D'Andrea *et al*., 2006; Theroux *et al*., 2010). However, these haptophytes are largely absent in molecular surveys targeting universal genes and microbial diversity studies of haptophyte blooms are nonexistent.

Nutrient loading and seasonal irradiance levels are known to trigger marine and estuarine haptophyte blooms (Tyrrell and Merico, 2004) that occur across latitudes in both cold and coastal regions (Brown and Yoder, 1994). Increasing global temperatures will result in longer ice-free periods in arctic lakes and an increase in runoff from thawing tundra catchments, undoubtedly affecting the annual haptophyte bloom events. Given the significance of haptophyte algae in aquatic ecosystems, their absence in previous datasets, and the desire to anticipate their future response to global climate change, the objectives of our study were two fold: (1) to sequence a

haptophyte-rich environment with a universal molecular approach; and (2) to produce a benchmark species survey for an arctic oligosaline lake during the spring bloom. Using high-throughput pyrotag sequencing, we targeted both bacterial and eukaryotic communities from two separate years to evaluate the consistency of the bloom-associated microbial populations. Our study provides an important baseline to contrast future BrayaSø microbial community change during its ice-free period in anticipation of a warmer Arctic possessing longer ice-free periods.

2. Materials and Methods

2.1 Site description

The Kangerlussuaq region of Southwestern Greenland lies at the head of the Søndre Strømfjord, 150 km from the ocean outlet, and has a series of saline lakes that provide ideal locations of paleoclimate reconstruction due to their anoxic hypolimnions and excellent preservation of lake sediments (**Figure 1**). Lake BrayaSø (66.99°N, -51.01°W) is a meromictic, oligosaline lake (salinity 2-3) that has an ice cap from September to late May. The surface area of Lake BrayaSø is approximately 72 hectares, with a maximum depth of 24 m. The dominant salts are $NaCO_3$, $NaHCO_3$ and $MgHCO_3$, which are delivered to the lakes via aeolian transport from nearby sand sheets and input from erosion within the lake catchments (Anderson *et al*., 2001). The bedrock in the region is granodioritic gneiss with occasional ultrabasic intrusions (Heggen *et al*., 2001). The climate is low-Arctic continental with > 500 mm/yr of precipitation and a mean annual temperature at Kangerlussuaq of -6°C (Heggen *et al*., 2001). The lake is calcium-depleted relative to fresh lakes of the region due to $CaCO_3$

precipitation, and the dominant cations are $Na^{+} > Mg^{2+} > K^{+} > Ca^{2+}$ (D'Andrea 2008). Total nitrogen is approximately 803 μg/L and total phosphorus is approximately 9 μg/L (Brutemark *et al.,* 2006). Dissolved organic carbon is approximately 90 mg/L (Anderson *et al*. 2009). Alkenone lipids are present in the sediments of Lake BrayaSø (D'Andrea *et al.,* 2005) and sediment trap data indicated the annual haptophyte bloom in the lake occurs approximately two weeks after ice-off (D'Andrea, 2008).

2.2 Water column

We selected samples for pyrosequencing based upon chlorophyll and alkenone concentrations in the water column, choosing the 10-m depth for 2007 and 4-m depth from 2009 (**Figure 2**). For 2007, 10-m corresponded to the peak in chlorophyll and alkenone concentrations. For 2009, 4-m depth was slightly above the chlorophyll maximum and coincided with the alkenone concentration peak (**Figure 1**). The June average temperature for each sampling day in 2007 and 2009 was 11.2°C and 9.9°C, respectively (**Figure 3**A). The average daily air temperature for June 2007 was 10.7°C and for June 2009 was 10°C (**Figure 3**B).

2.3 Water sampling

The bloom event occurred in BrayaSø, Greenland in June of 2007 and 2009. For both years, we analyzed a sample collected during the first week of the haptophyte bloom. We collected geochemical data using a YSI Sonde (Ohio, USA) equipped with probes to measure temperature, conductivity, dissolved oxygen, and chlorophyll a fluorescence. At 1-m intervals, we collected water with a Van Dorn water sampler

and preserved these samples for alkenone and genomic DNA analysis. For alkenone analysis, we filtered one liter of water onto a pre-combusted (550°C) GF/F 0.7μm, 47 mm glass filter, and kept it frozen until analysis. For DNA analysis, we filtered a separate liter of lake water onto a 0.2μm Sterivex™ filter (Millipore, Billerica, MA, USA), flooded the filter with Puregene lysis buffer (Qiagen, Valencia, CA, USA), and froze it at -20°C until processing. We selected samples for sequencing based on maximum alkenone concentrations.

2.4 Lipid analysis

Alkenone extraction was after D'Andrea and Huang (2005). Alkenone samples and DNA samples were sourced from the same water sample. We freeze-dried and homogenized samples manually. We extracted samples with 9:1 Dichloromethane (DCM):Methanol (MeOH) using an Accelerated Solvent Extractor ASE200 (Dionex, Sunnyvale, CA, USA). Extracts were separated into acid and neutral fractions using a solution of DCM, Isopropyl alcohol 2:1 (v/v). The neutral fraction was further separated into aliphatic (hexane elution), ketone (DCM), and alcohol (ethyl acetate:hexane 1:3) fractions using a flash silica gel column. The DCM fraction was analyzed using an Agilent 6890plus Gas Chromatograph Flame Ionization Detector (GC-FID) (Santa Clara, CA) for quantification. Chromatograms were compared to previously reported alkenone standards and their GC retention times (de Leeuw *et al.,* 1980; Marlowe *et al.,* 1984). Alkenone concentrations were determined from GC-FID analysis of the ketone fractions based on an internal C_{36} alkane standard.

2.5 DNA extraction

We extracted Sterivex™ filters using a Qiagen Puregene Cell Kit (Venlo, Netherlands) according to the manufacturer's instructions. Genomic DNA was polyethylene glycol (PEG) purified (LaMontagne *et al.,* 2002) to remove proteins and other contaminants that inhibit PCR reactions. DNA was suspended in PEG at 4°C overnight, centrifuged, and the pellet rinsed with ethanol. The DNA was resuspended in DNA hydration solution (Qiagen). We quantified total extracted genomic DNA yields using a NanoDrop nucleic acid spectrophotometer (Thermo Scientific, Wilmington, DE) to ensure they were RNA-free.

2.6 Quantitative polymerase chain reaction

Purified DNA extracts were also subjected to real-time quantitative polymerase chain reaction (qPCR) to gauge haptophyte cell concentrations with depth and ensure that the sample selected for sequencing was at the point of highest haptophyte cell concentration in the water column. We performed the qPCR reaction using 18S rRNA gene haptophyte-specific primers Prym-429F (5'-GCG CGT AAA TTG CCC GAA-3'; T_m = 65°C), and Prym-887R (5'-GGA ATA CGA GTG CCC CTG AC-3'; T_m = 62°C) (Coolen *et al.,* 2004; Simon *et al.,* 2000). These primers yield an amplicon that is approximately 463bp in size. These primers have previously been screened for specificity: the forward primer Prym-429F is 100% specific for Haptophyta order Prymnesiales and matched 93% of orders Coccosphaerales, Isochrysidales, Prymnesiales, the genus *Pleurochrysis*, as well as unclassified haptophytes (Coolen *et al.,* 2004). The reverse primer is specific to Prymnesiophyceae (Simon *et al.,* 2000).

We further confirmed primer specificity using the ARB probematch tool in SILVA ARB database v111 (Pruesse *et al*., 2007). This SSU reference database contains 739,633 high quality 16S/18S rRNA gene sequences. The Prym429F primer matched 72% of full-length 18S rRNA gene haptophyte sequences with two mismatches, and no non-haptophyte sequences. The Prym88R primer returned 99.3% of haptophyte sequences with one mismatch and no non-haptophyte sequences.

The qPCR reactions were run in triplicate, including a no-template control, on an Applied Biosystems StepOnePlus™ Real-Time PCR System (Foster City, California), using a SYBR Green I assay. We also ran a positive control of *Isochrysis galbana* DNA extracted from a culture with cell concentrations at $1.5x10^6$ cells/ml. The Cq for each sample had a deviation of less than 0.5. Each 20 µl reaction contained 7.2 µl of sterile water, 10 µl of KAPA SYBR® FAST Universal 2X qPCR Master Mix (Woburn, MA), 0.4 µl each of the forward and reverse primers (0.2µM) and 2 µl of template DNA. Template DNA ranged in concentration from 2-10ng/µl The qPCR cycling program was after Coolen *et al*. (2009) and consisted of 38 cycles of denaturation at 94°C for 30s, annealing at 62°C for 40s, primer extension at 72°C for 60s, a photo step of 80°C for 20s. We used between 10^1 and 10^6 copies (ten-fold dilution series) of linearized plasmids containing the 18S rRNA gene of *Isochrysis galbana* CCMP1323 as the external standard to calibrate the copy numbers of haptophyte RNA genes in the BrayaSø water samples. Our standard curve was established using four points of the diluted standard, with an R^2 value of >.999 and slope of –3.991. Our reaction efficiency was 78.1%. We used StepOne Software version 2.2 (Applied Biosystems) to analyze our results. These conditions are

reported in accordance with the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines (Bustin *et al.,* 2009).

2.7 Pyrosequencing

We performed genomic DNA amplifications using eukaryotic and bacterial-specific primers targeting the V9 (Amaral-Zettler *et al.,* 2009) or V6-V4 (Morrison and Sogin, in prep.) regions, respectively. Eukaryotic sequences were generated using a Genome Sequencer FLX (Roche, Switzerland) with the GS-LR70 long-read sequencer kit at the Marine Biological Laboratory Keck Sequencing Facility. Amplifications and sequencing for eukaryotic sequences were after Amaral Zettler *et al*. (2009). We sequenced the V6-V4 hypervariable region of the bacterial 16S rRNA gene using bacterial primers 515F and 1046R on a Roche GS FLX pyrosequencer using GS FLX Titanium Series reagents (Roche Diagnostics, Basel, Switzerland) following manufacturer's protocols. Sequences were trimmed and screened for quality after Huse *et al*. (2007). To assign taxonomy to the remaining quality-controlled tags, we used the Global Alignment for Sequence Taxonomy (GAST) algorithm (Huse *et al.,* 2008). Tag sequences were grouped into Operational Taxonomic Units using SLP-PWAL (refer to Huse *et al*. 2010), with bacterial sequences clustered at 3% and eukaryotic sequences clustered at 6%. Venn diagrams were constructed using BioVenn (Hulsen *et al.,* 2008). Bacterial diversity estimates were calculated using EstimateS v8.2.0 (Colwell 2005) and CatchAll (Bunge, 2011). All sequences have been deposited in the NCBI Sequence Read Archive (SRA) under the SRA number SRA059384, and are MIMARKS compliant (Yilmaz *et al.,* 2011).

3. Results and Discussion

3.1 Water column

Both 2007 and 2009 samples were selected from the first week of the two-week haptophyte bloom. In both years, secchi depth was 5 m indicating the photic zone terminated at approximately 10-12.5 m depth. The alkenone peak in 2007 corresponded to the oxycline at 10m depth, whereas the alkenone peak in 2009 was at the thermocline.

Quantitative PCR analysis confirmed that the water sample from peak alkenone depth corresponded to peak haptophyte cell numbers (**Table 1**, **Figure 2**). The structure of the water column between the two years was markedly different, with chlorophyll peaking at 10 m depth in 2007 and about 6.5m depth in 2009. Alkenone concentrations peaked at 15 μg/L in 2007 and about 59 μg/L in 2009. Correspondingly, rRNA gene copies peaked at 2146 copies/ml in 2007 and 9898 copies/ml in 2009 (**Figure 1**). This equates to approximate cellular alkenone concentrations of 6-7 ng per cell if the 18S rRNA gene copies occur singularly or 3-3.5 ng/cell if there are two copies of the 18S rRNA gene in these haptophytes. This is on par with previously observed cellular alkenone concentrations in lacustrine haptophytes of 0.009 to 2 pg/cell (Versteegh *et al*., 2001; Marlowe *et al*., 1984). Given that Lake BrayaSø has the highest sedimentary concentrations of alkenones ever reported (D'Andrea and Huang, 2005), the high cellular concentrations as estimated by our qPCR analysis is not surprising. Our results also agree with observations by Boere *et al*. (2011) that alkenone concentrations can serve as a proxy for haptophyte cell numbers.

3.2 Bacterial community diversity

A total of 6,409 bacterial OTUs were observed between 2007 and 2009 (**Figure 4**A). The two years had comparable OTU yields, 2883 from 2007 and 2727 from 2009, yet shared only 13% of their OTUs. This overlap in OTUs was surprisingly low, considering reports from other arctic lake surveys demonstrating up to 73% overlap in bacterial community membership (Crump *et al*., 2003). When singletons were ignored, this overlap increased to 44% (**Figure 4**B), indicating a third of the taxon differences came from the rarest members of the community. Our study revealed much greater bacterial diversity at the phylum level, 25 phyla, than previous studies from freshwater and oligosaline lakes on the Tibetan plateau that identified only 13 phyla (Liu *et al*., 2010). The estimates of alpha diversity of the bacterial community generated species richness estimates with overlapping confidence bounds (**Table 2**), demonstrating that our similar OTU yields reflected the similar alpha diversity or richness during the two years.

The Morisita-Horn index of similarity, an abundance-based distance measure of beta diversity, was 0.848 (complete overlap = 1) (**Table 2**) indicating the most abundant species were present in both 2007 and 2009. Of the most abundant bacterial OTUs (**Table 3**A), the notable difference between communities in 2007 and 2009 was the presence of the sulfur-oxidizing bacteria in 2009 (*Thiomicrospira*, *Sulfurovum*, *Sulfuricurvum*) and fewer Flavobacteria in 2009. This flavobacterial OTU matched environmental sequences from freshwater environments, including 100% sequence identity to a bacterium isolated during a spring phytoplankton bloom

in Lake Zurich (Eckert *et al.,* 2011) and bacteria from lakes on the Tibetan plateau (Zhang, R. and Liu, W.-T. unpublished, GenBank HM128691).

While Lake BrayaSø is oligosaline, its bacterial community resembled previously reported freshwater environments in addition to high altitude environments. The bacterial OTUs were dominated by Actinobacteria, which are known to occur ubiquitously in terrestrial and aquatic ecosystems (Embley and Stackebrandt, 1994) and can dominate lake epilimnia (Newton *et al.,* 2011). The most abundant actinobacterial OTU matched environmental sequences from Lake Taihu (China) and other freshwater lakes with 100% identity. The second most abundant phylum represented, the beta-proteobacteria, occurs more commonly in freshwater environments than marine (Nold *et al.,* 1998) and represents the most abundant bacteria in glacial meltwater communities (Cheng and Foght, 2007). Overall, the most abundant bacterial taxa (**Table 3**A) matched sequences from other freshwater environments, and resembled that of high-altitude lakes from the Tibetan plateau in the abundance of Actinobacteria, alpha- and beta- Proteobacteria (Xing *et al.,* 2009; Liu *et al.,* 2010). Lakes at high altitude experience similar environmental pressures as lakes at high latitude, including oligotrophy, low temperature, and high UV radiation in the surface waters; the similarity in their bacterial communities suggests these particular phyla can withstand harsh environmental conditions across latitudes.

Toolik Lake in Alaska experiences an increase in primary and bacterioplankton production in the first month of spring as melting snow increases organic matter and nutrient transport into the lake and allows for an increase in sunlight reaching the water column (Hobbie *et al.,* 1983; Crump *et al.,* 2003). A

similar trend is observed on the western shelf of the Antarctic peninsula, where seasonal melting dictates irradiance levels, mixed layer depth, and organic carbon availability (Montes-Hugo *et al*., 2010), with an increase in primary production resulting in an increase in bacterial production. Given the increasing global temperatures, we anticipate an increased supply of organic matter into Lake BrayaSø and thus an increase in bacterioplankton production. An increasing freshwater input as a result of melting arctic tundra may affect the local hydrological balance enough to freshen Lake BrayaSø and shift the microbial community further towards one of a more freshwater composition.

3.3 Eukaryotic community diversity

Previous work in Lake BrayaSø identified only 11 eukaryotic phyla (Brutemark *et al*., 2006); using high throughput sequencing we were able to identify 9 times more phyla, including picoplankton that were undetectable with the previous visual identification methods (**Table 3**B). A total of 97 eukaryotic OTUs were observed between 2007 and 2009, with an overlap between the two years of only 26% (**Figure 4**C, **Table 2**). When singletons were ignored, this number increased to 41% (**Figure 4**D). The eukaryotic community was dominated by diatom and ciliate OTUs in spite of the presence of a haptophyte "bloom". Diatoms are known to have high copy numbers of their 18S rRNA genes, which may be the cause of the high abundance of their OTUs (Zhu *et al*., 2005; Not *et al*. 2009), as are alveolates which range up to 9,000 copies per cell (Prescott, 1994). Haptophyte 18S rRNA gene copy numbers are estimated at 2-3 copies per cell (Zhu *et al*., 2005) and our qPCR analysis

using an *Isochrysis galbana* standard yielded approximate 18S rRNA gene copy number at 1 copy per cell (**Table 1**). Given the high ciliate and diatom tag sequences, these patterns in eukaryotic community structure likely reiterate a cautionary note on the interpretation of abundance data for 18S rRNA gene studies, although these concerns can be minimized when comparing intra-species abundances.

The eukaryotes present in BrayaSø were typical of freshwater meso- and eutrophic environments. Diatom-related OTUs were the most abundant tags we recovered in both 2007 and 2009 (**Table 3**B). The most abundant diatom OTU represented 31% of all eukaryotic tag sequences, and shared 100% sequence identity with araphid diatoms from fresh and brackish water. The second most abundant OTU in 2007 was assigned a ciliate taxonomy that matched environmental sequences from floodplain soil and an ephemeral pond to 96% and 95%, respectively. In contrast, the second most abundant tag in 2009 matched a metazoan, and shared 100% sequence identity with the copepod *Leptodiaptomus moorei* (GenBank AY339154). This metazoan was notably absent in the 2007 dataset. Other OTUs present in > 10 fold higher abundance in 2009 versus 2007 included *Chlamydomonas*, an unidentified environmental ciliate, an alveolate, and a chrysophyte (**Table 3**B). Of the most abundant eukaryotes, an unidentified ciliate had the greatest average GAST distance of 0.1668 (**Table 3**B) and is likely a novel species. The similarity of protistan lineages to other freshwater environments confirms the results from other studies increasingly showing the distinction between marine and freshwater communities (Logares *et al*., 2009).

Haptophytes comprised only 3% of the eukaryotic community in 2007, and 0.4% in 2009 (**Table 3**B). Despite high levels of recorded alkenones at the depth of sampling, we recovered only one type of haptophyte OTU in both years albeit in much greater abundance in 2007 than 2009 (**Table 3**B). The most abundant tag in this OTU cluster shared 100% identity with a previously sequenced 18S rRNA gene from the BrayaSø water column (GenBank HQ446272; (Theroux *et al.,* 2010)). We detected identical haptophyte V9 tag reads to this BrayaSø OTU from Toolik Lake, Alaska (Crump *et al.,* 2012) and Plum Island, Massachusetts (Amaral-Zettler, personal observation) but nowhere else in the VAMPS global database (http://vamps.mbl.edu). Singularity of the haptophyte population impacts the ability to use alkenone-derived paleoclimate records from a site; the presence of multiple haptophyte species during a bloom could jeopardize the reliability of the alkenone record. The occurrence of a single haptophyte OTU in both 2007 and 2009 samples is encouraging for the use of this environment as a paleoclimate archive; alkenone-based temperature reconstructions would therefore only require a single temperature calibration for the single haptophyte species present.

The alkenone concentrations at peak alkenone depths in the water column were 15 μg/L in 2007 and 59 μg/L in 2009. Our qPCR results (**Table 1**) indicated that there were approximately 1700 and 2900 haptophyte cells/ml in the 2007 and 2009 water samples, respectively. Our qPCR results also indicated that *Isochrysis galbana* has 1-2 copies of its 18 rRNA gene per cell. The occurrence of only 155 and 16 haptophyte tags in 2007 and 2009, respectively, suggested that the haptophyte DNA may have been dwarfed by greater copy number diatom and ciliate 18S rRNA

sequences. Empirical obstacles also may have resulted in the low haptophyte tag yield, including primer mismatches or difficulties amplifying GC-rich haptophyte DNA.

4. Conclusions

Arctic lakes will undoubtedly experience shifts in microbial populations with increasing annual temperatures, prolonged ice-free periods and thawing tundra catchments. This study is the first to examine the microbial community of an artic oligosaline lake using high-throughput sequencing, providing a deeper resolution of the microbial community structure in these rapidly changing arctic environments. Using high throughput sequencing, we were able to detect greater phylum-richness and new phyla previously unobserved in BrayaSø, the benefit of a molecular versus microscopy-based approach. Even though BrayaSø is an oligosaline lake, both the bacterial and eukaryotic communities resembled other high latitude and high altitude freshwater environments. The low overlap in microbial communities between the 2007 and 2009 samplings suggested large interannual variations in microbial species. However, the 2009 sample had fewer haptophyte tags but a greater abundance of other phototrophs, suggesting the functional overlap of the eukaryotic communities may be greater than the species overlap. Future studies examining microbial populations throughout the course of a spring bloom event will help resolve these temporal shifts in species abundances and functional roles.

This study is also the first to analyze a haptophyte bloom event using next generation sequencing. We generated fewer haptophyte pyrotag sequences than

expected given their alkenone biomarker concentrations in the water column. Our qPCR data confirmed that haptophyte cell numbers peaked with alkenone concentrations, and also showed that high-throughput tag sequences for haptophytes did not correspond well with qPCR counts. This result serves as a reminder that the interpretation of relative abundance data using a tag sequencing approach with eukaryotes must be done so cautiously and that complimentary, haptophyte-specific qPCR provides greater detail of cell abundances. Given the depth of DNA sequencing, and the generation of over 200 haptophyte tags, we are encouraged by the presence of a single haptophyte OTU in Lake BrayaSø, and maintain that this is a worthy location for temperature reconstruction using alkenone-based proxies. Future studies throughout the haptophyte bloom event in BrayaSø will resolve the temporal shifts in microbial communities and will help decipher the communities most susceptible to increasing arctic temperatures.

Acknowledgements

This project was supported by grants from the National Science Foundation (NSF) to Y. Huang (ARC-0402383), a Brown University SEED grant to Y. Huang and L. Amaral-Zettler, and an American Association of University Women American Fellowship to S. Theroux. We thank Melissa Paddock for her assistance with qPCR analyses.

References

Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM. (2009). A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS One* 4, e6372.

Anderson NJ, Brodersen KP. (2001). Determining the date of ice-melt for low Arctic lakes along Søndre Strømfjord, southern West Greenland. *Geology of Greenland Survey Bulletin* 189, 54-58.

Anderson NJ, D'Andrea W, Fritz SC. (2009). Holocene carbon burial by lakes in SW Greenland. *Global Change Biology* **15**:2590-2598.

Bennike O, Anderson NJ, McGowan S. (2010). Holocene palaeoecology of southwest Greenland inferred from macrofossils in sediments of an oligosaline lake. *ISME J* 43, 787-798.

Boere, A. C., W. I. C. Rijpstra, G. J. de Lange, E. Malinverno, J. S. Sinninghe Damsté, and M. J. L. Coolen. (2011). Exploring preserved fossil dinoflagellate and haptophyte DNA signatures to infer ecological and environmental changes during deposition of sapropel S1 in the eastern Mediterranean, *Paleoceanography*, 26, PA2204, doi, 10.1029/2010PA001948.

Brassell SC, Eglinton G, Marlowe IT, Pflaumann U, Sarnthein M. (1986). Molecular stratigraphy - a new tool for climatic assessment. *Nature* 320, 129-133.

Brown, C. and Yoder, J. (1994). Coccolithophorid blooms in the ocean. *J. Geophys. Res.* 99, 7467–7482.

Brutemark A, Rengefors K, Anderson NJ. (2006). An experimental investigation of phytoplankton nutrient limitation in two contrasting low arctic lakes. *Polar Biol* 29, 487-494.

Bunge J. (2011). Estimating the number of species with CatchAll. *Pac Symp Biocomput* 2011, 121-130.

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, and Wittwer CT. (2009) The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin. Chem.* 55, 611-622.

Cheng SM, Foght JM. (2007). Cultivation-independent and-dependent characterization of Bacteria resident beneath John Evans Glacier. *FEMS Microbial Ecol.* 59, 318-330.

Colwell, R. K. (2005). EstimateS: Statistical estimation of species richness and shared species from samples. Version 8.2.0. User's Guide and application published at: http://purl.oclc.org/estimates.

Conte MH, Sicre MA, Ruhlemann C, Weber JC, Schulte S, Schulz-Bull D, Blanz T. (2006). Global temperature calibration of the alkenone unsaturation index $U^{k'}_{37}$ in surface waters and comparison with surface sediments. *Geochem Geophy Geosy* 7, doi, 10.1029/2005GC001054.

Coolen MJL, Muyzer G, Rijpstra WIC, Schouten S, Volkman JK, Sinninghe Damsté J. (2004). Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake. *Earth Planet Sci Lett* 223, 225-239.

Coolen, MJL. Saenz, JP, Giosan, L, Trowbridge, NY, Dimitrov, P, Dimitrov, D, Eglinton,TI. (2009). DNA and lipid molecular stratigraphic records of haptophyte succession in the Black Sea during the Holocene, *Earth Planet Sci Lett* 284, 610-621.

Crump BC, Amaral-Zettler LA, Kling GW. (2012). Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *ISME J* doi, 10.1038/ismej.2012.9.

Crump BC, Kling GW, Bahr M, Hobbie JE. (2003). Bacterioplankton community shifts in an arctic lake correlate with seasonal changes in organic matter source. *Appl Environ Microbiol* 69, 2253-2268.

D'Andrea WJ, Huang Y. (2005). Long chain alkenones in Greenland lake sediments, Low $\delta^{13}C$ values and exceptional abundance. *Org Geochem* 36, 1234-1241.

D'Andrea WJ, Lage M, Martiny JBH, Laatsch AD, Amaral-Zettler LA, Sogin ML, Huang YS. (2006). Alkenone producers inferred from well-preserved 18S rDNA in Greenland lake sediments. *J Geo Res-Biogeo* 111, doi, 10.1029/2005JG000121.

D'Andrea WJ. (2008). Development and application of lacustrine alkenone paleothermometry in southwestern Greenland. PhD dissertation, 197.

de Leeuw JW, van der Meer FW, Rijpstra WIC, Schenck PA. (1980). On the occurrence and structural identification of long chain unsaturated ketones and hydrocarbons in sediments. *Phys Chem Earth* 12, 211-217.

Eckert EM, Posch MMST, Eugster B, Pernthaler J. (2011). Rapid successions affect microbial N-acetyl-glucosamine uptake patterns during a lacustrine spring

phytoplankton bloom. *Environ Microbiol* 14, 794-806.

Edgcomb V, Orsi W, Bunge J, *et al*. (2011) Protistan microbial observatory in the Cariaco Basin, Caribbean. I. Pyrosequencing vs Sanger insights into species richness. *ISME Journal* 5, 1344-1356.

Embley TM, Stackebrandt E. (1994). The molecular phylogeny and systematics of the actinomycetes. *Annu Rev Microbiol* 48, 257-289.

Haukka K, Kolmonen E, Hyder R, Hietala J, Vakkilainen K, Kairesalo T, Haario H, Sivonen K. (2006). Effect of nutrient loading on bacterioplankton community composition in lake mesocosms. *Microb Ecol* 51, 137-146.

Hulsen T, De Vlieg J, Alkema W. (2008). BioVenn–a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC genomics* 9, 488.

Huse SM, Mark Welch D, Morrison HG, Sogin ML. (2010). Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environ Microbiol* 1889-1898.

Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM. (2007). Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol* 8, R143.

Huse SM, Huber JA, Welch DM, Relman DA, Sogin ML. (2008). Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS genetics* 4, e1000255.

LaMontagne MG, Michel FC, Holden PA, Reddy CA. (2002). Evaluation of extraction and purification methods for obtaining PCR-amplifiable DNA from compost for microbial community analysis. *J Microbiol. Methods* 49, 255-264.

Liu H, Probert I, Uitz J, Claustre H, Aris-Brosou S, Frada M, Not F, de Vargas C. (2009). Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. *Proc Natl Acad Sci* 106(31), 12803-12808

Liu X, Yao T, Kang S, Jiao N, Zeng Y, Liu Y. (2010). Bacterial community of the largest oligosaline lake, Namco on the Tibetan Plateau. *Geomicrobiol J* 27, 669-682.

Logares, R., Shalchian-Tabrizi, K., Boltovskoy, A. & Rengefors, K. 2007. Extensive dinoflagellate phylogenies indicate infrequent marine–fresh- water transitions. *Mol. Phylogenet. Evol.* 45, 887–903.

Marchetto A, Mosello R, Rogora M, Manca M, Boggero A, Morabito G, Musazzi S, Tartari GA, Nocentini AM, Pugnetti A. (2004). The chemical and biological

response of two remote mountain lakes in the Southern Central Alps (Italy) to twenty years of changing physical and chemical climate. *J Limnol* 63, 77-89.

Marlowe IT, Green JC, Neal AC, Brassell SC, Eglinton G, Course PA. (1984). Long-Chain (n-C_{37}-C_{39}) alkenones in the prymnesiophyceae. Distribution of alkenones and other lipids and their taxonomic significance. *Br Phycol J* 19, 203-216.

Moon-van der Staay S.Y., De Wachter R., Vaulot D. (2001). Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* 409, 607– 610.

Müller PJ, Kirst G, Ruhland G, von Storch I, Rosell-Mele A. (1998). Calibration of the alkenone paleotemperature index $U^{k'}_{37}$ based on core-tops from the eastern South Atlantic and the global ocean (60°N-60°S). *Geochim Cosmochim Acta* 62, 1757-1772.

Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S. (2011). A guide to the natural history of freshwater lake bacteria. *Microbiol Mol Biol R* 75, 14-49.

Nold SC, Zwart G. (1998). Patterns and governing forces in aquatic microbial communities. *Aquat Ecol* 32, 17-35.

Not F, Latasa M, Scharek R, Viprey M, Karleskind P, *et al*. (2008) Phytoplankton diversity across the Indian Ocean: A focus on the picoplanktonic size fraction. *Deep-Sea Res. I* 55, 1456–1473.

Pawlowski J, Christen R, Lecroq B *et al*. (2011) Eukaryotic Richness in the Abyss, Insights from Pyrotag Sequencing. *PLoS ONE* 6, e18169, doi, 10.1371/journal.pone.0018169.

Prahl FG, Wakeham SG. (1987). Calibration of unsaturation patterns in long-chain ketone compositions for paleotemperature assessment. *Nature* 330, 367-369.

Prescott, D. M. 1994. The DNA of ciliated protozoa. *Microbiol. Rev*. 58, 233–267.

Presthus Heggen M, Birks HH, Anderson NJ. (2010). Long-term ecosystem dynamics of a small lake and its catchment in west Greenland. *Holocene* 20, 1207-1222.

Pruesse E, Quast C, Knittel K, Fuchs B, Ludwig W, Peplies J, Glöckner FO. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acid Res*. 35:7188

Quayle WC, Peck LS, Peat H, Ellis-Evans JC, Harrigan PR. (2002). Extreme responses to climate change in Antarctic lakes. *Science* 295, 645.

Simon N, Campbell L, Ornolfsdottir E, Groben R, Guillou L, Lange M. and Medlin LK. (2000). Oligonucleotide probes for the identification of three algal groups by dot blot and fluorescent whole-cell hybridization. *J. Eukaryot. Microbiol.* 47: 76-84.

Shalchian-Tabrizi K, Reier-Røberg K, Ree DK, Klaveness D, Bråte J (2011) Marine freshwater colonizations of Haptophytes inferred from phylogeny of environmental 18S rDNA sequences. *J Eukaryot. Microbiol.* 58, 315-318.

Slapeta, J., Moreira, D. & Lopez-Garcia, P. 2005. The extent of protist diversity insights from molecular ecology of freshwater eukaryotes. *Proc. R. Soc. Biol. Sci. Ser. B*, 272, 2073–2081.

Stoeck T, Behnke A, Christen R, Amaral-Zettler L, Rodriguez-Mora M, Chistoserdov A *et al.* (2009). Massively parallel tag sequencing reveals the complexity of anaerobic marine protistan communities. *BMC Biology* 7, 72.

Stoeck T, Bass D, Nebel M, Christen R, Jones MD, Breiner HW, Richards TA (2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* 19, 21- 31.

Straile D. (2005). Food webs in lakes—seasonal dynamics and the impact of climate variability. *Aquatic food webs, an ecosystem approach. Oxford Univ. Press, Oxford* 41-50.

Theroux SM, D'Andrea WJ, Toney JL, Amaral-Zettler LA, Huang Y. (2010). Phylogenetic diversity and evolutionary relatedness of alkenone producing haptophyte algae in lakes, Implications for continental paleotemperature reconstructions. *Earth Planet Sci Lett* 300, 311-320.

Tyrrell, T and Merico, A. (2004) "*Emiliania huxleyi*: Bloom observations and the conditions that induce them" in *Coccolithophores—from molecular processes to global impact*. Eds. H. R. Thierstein and J. R. Young (Springer), 75–97.

Volkman JK, Eglinton G, Corner EDS, Forsberg TEV. (1980). Long-chain alkenes and alkenones in the marine coccolithophorid *Emiliania-huxleyi*. *Phytochemistry* 19, 2619-2622.

Xing P, Hahn MW, Wu QL. (2009). Low taxon richness of bacterioplankton in high-altitude lakes of the eastern Tibetan Plateau, with a predominance of Bacteroidetes and *Synechococcus* spp. *Appl Environ Microbiol* 75, 7017-7025.

Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, Gilbert JA, Karsch-Mizrachi I, Johnston A, Cochrane G. (2011). Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nat Biotechnol* 29, 415-420.

Zhu F, Massana R, Not F, Marie D, Vaulot D (2005) Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbial Ecol*. 52, 79–92.

Table 1. BrayaSø water column C_{37} alkenone concentration and haptophyte rRNA gene copy number. Asterisks denote samples chosen for pyrosequencing.

Date	Water depth (m)	[C37] (µg/L)	Haptophyte rRNA gene copies/mL
6/20/07	5	4	649
6/20/07	10*	15	1,788
6/20/07	14	4	986
6/20/07	20	1	230
6/17/09	4*	59	2,969
Isochrysis galbana control			1,493

Table 2. Sequencing summary, OTU distributions, and diversity estimates. Upper and lower confidence bounds are 95%. Abbreviations for CatchAll: Estimate = estimated total species richness; Chao1 = Chao1 estimated total species richness; ACE = 'abundance-based coverage estimator' estimated total species richness.

	Bacteria		**Eukaryota**	
	2007	**2009**	**2007**	**2009**
Sequenced tags	16 533	17 034	1 554	2 795
Total Observed OTUs	3 682	3 526	59	63
Singletons	2 679	2 604	18	25
Shared OTUs	799		25	
Shared OTUs >1	72%		100%	
Jaccard	0.125		0.258	
Sorensen	0.222		0.41	
Morisita-Horn	0.848			
Bray-Curtis	0.527			
Richness estimates	**2007**	**2009**		
CatchAll Estimate	27 353 (- 8148, + 12 425)	18 727 (-1864, +2124)		
Chao	11 142 (- 838, + 945)	11 060 (-866, + 979)		
ACE	20 823 (- 2398, + 2689)	22 639 (-2749, + 3211)		

Table 3. Most abundant bacterial and eukaryotic OTUs. Bacterial OTUs defined at 97% similarity, eukaryote OTUs defined at 96% similarity. Asterisk (*) denotes OTUs with a relative abundance difference greater than one order of magnitude. GAST = average distance between OTU and Global Alignment for Sequence Taxonomy reference sequences (A) The 35 most abundant bacterial tags, ranked by total relative abundance. (B) The 15 most abundant eukaryote tags, ranked by total relative abundance.

A. Bacteria Taxonomy	2007	2009	Total	GAST
Bacteria; Actinobacteria; Actinobacteria; Actinomycetales; Sporichthyaceae	1660	1136	2796	0.0112
Bacteria; Actinobacteria; Actinobacteria; Actinomycetales; Microbacteriaceae; Limnoluna	1115	796	1911	0.0079
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Burkholderiaceae; Polynucleobacter	435	660	1095	0.0054
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Comamonadaceae	505	488	993	0.0066
Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales; SAR11; Pelagibacter	550	292	842	0.0198
Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Cryomorphaceae; Owenweeksia	429	322	751	0.0051
Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales; SAR11; Pelagibacter	390	308	698	0.0242
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Alcaligenaceae	174	470	644	0.009
Bacteria; Actinobacteria; Actinobacteria; Actinomycetales; Sporichthyaceae	360	268	628	0.0145
Bacteria; Chloroflexi	177	434	611	0.0121
Bacteria; Bacteroidetes; Sphingobacteria; Sphingobacteriales; Cyclobacteriaceae; Algoriphagus	338	191	529	0.0114
Bacteria; Actinobacteria; Actinobacteria; Acidimicrobiales; Acidimicrobiaceae	148	317	465	0.0046
Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Flavobacteriaceae; Flavobacterium	420*	25	445	0.0066
Bacteria; Actinobacteria; Actinobacteria; Actinomycetales; Sporichthyaceae	215	228	443	0.0122

Bacteria; Proteobacteria; Gammaproteobacteria; Thiotrichales; Piscirickettsiaceae; Thiomicrospira	5	432*	437	0.0334
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales; Helicobacteraceae; Sulfurovum	0	424*	424	0.0344
Bacteria; Actinobacteria; Actinobacteria; Acidimicrobiales; Iamiaceae; Iamia	254	145	399	0.0064
Bacteria; Actinobacteria; Actinobacteria; Actinomycetales; Microbacteriaceae	142	212	354	0.0086
Bacteria; Actinobacteria; Actinobacteria; Actinomycetales; Sporichthyaceae	190	110	300	0.0099
Bacteria; Actinobacteria; Actinobacteria; Acidimicrobiales; Acidimicrobiaceae	102	182	284	0.0058
Bacteria; Actinobacteria; Actinobacteria; Nitriliruptorales; Nitriliruptoraceae; Nitriliruptor	74	171	245	0.0174
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales; Helicobacteraceae; Sulfuricurvum	0	237*	237	0.0184
Bacteria; Firmicutes; Erysipelotrichi; Erysipelotrichales; Erysipelotrichaceae	90	135	225	0.1072
Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Flavobacteriaceae; Lutibacter	56	162	218	0.0102
Bacteria; Bacteroidetes; Sphingobacteria; Sphingobacteriales; Cytophagaceae; Adhaeribacter	116	99	215	0.0389
Bacteria; Bacteroidetes; Sphingobacteria; Sphingobacteriales	127	87	214	0.0219
Bacteria; Bacteroidetes; Sphingobacteria; Sphingobacteriales; Sphingobacteriaceae; Sphingobacteriaceae; Pedobacter	146	68	214	0.005
Bacteria; Firmicutes; Erysipelotrichi; Erysipelotrichales; Erysipelotrichaceae	99	107	206	0.0422
Bacteria; Proteobacteria; Betaproteobacteria; Methylophilales; Methylophilaceae	52	149	201	0.0074
Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales; Flavobacteriaceae; Flavobacterium	49	129	178	0.0267
Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales; Planctomycetaceae; Gemmata	9	163*	172	0.0058
Bacteria; Verrucomicrobia; Verrucomicrobiae; Verrucomicrobiales; Verrucomicrobiaceae	9	159*	168	0.0133
Bacteria; Actinobacteria; Actinobacteria; Actinomycetales; Sporichthyaceae	97	66	163	0.0421
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Burkholderiaceae; Polynucleobacter; cosmopolitanus	144*	14	158	0.0159
Bacteria;Planctomycetes;Planctomycetacia; Planctomycetales;Planctomycetaceae;Pirellula	71	70	141	0.0273

B. Eukaryote Taxonomy	2007	2009	Total	GAST
Eukaryota; stramenopiles; Bacillariophyta; Fragilariophyceae	1518	929	2447	0.0059
Eukaryota; Alveolata; Ciliophora	864	597	1461	0.0577
Eukaryota; stramenopile	640	262	902	0.1185
Eukaryota; Metazoa; Arthropoda	0	897*	897	0.0189
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Chlamydomonadales	10	501*	511	0.01
Eukaryota; Alveolata; Ciliophora	35	185	220	0.0189
Eukaryota; Alveolata; Ciliophora	1	181*	182	0.0141
Eukaryota; Haptophyceae	155	16	171	0.0179
Eukaryota; Katablepharidophyta; Katablepharidaceae	120	41	161	0.1036
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Chlamydomonadales	33	110	143	0.0071
Eukaryota; Alveolata; Ciliophora	70	60	130	0.1668
Eukaryota; Cryptophyta	99*	8	107	0.0107
Eukaryota; Alveolata	1	65*	66	0.0856
Eukaryota; stramenopiles; Chrysophyceae	0	47*	47	0.0539
Eukaryota; stramenopiles; Chrysophyceae	29	16	45	0.0287

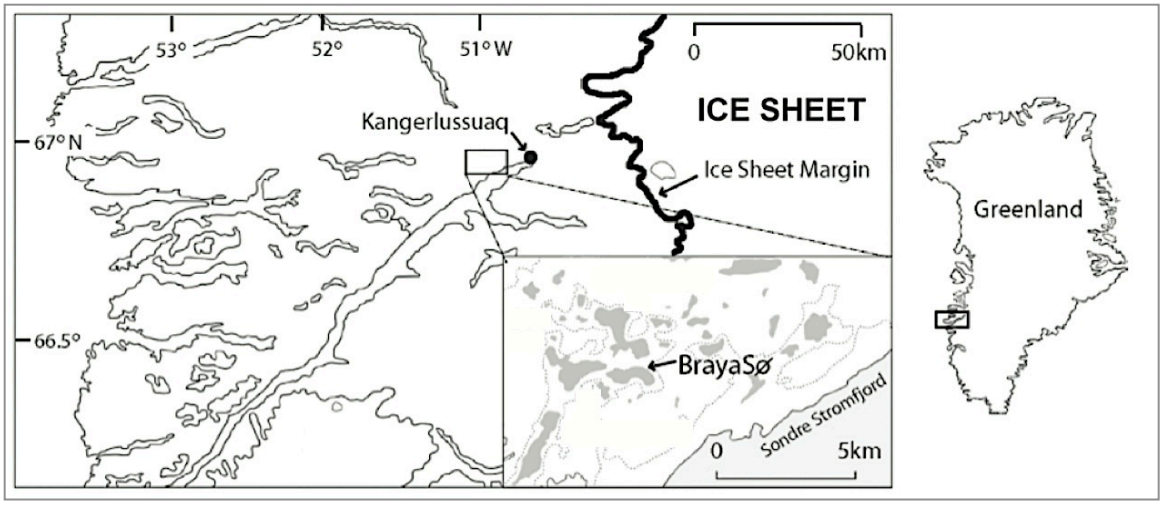

Figure 1. Site map showing Lake BrayaSø in the Kangerlussuaq region, Greenland. Modified from D'Andrea *et al.,* 2006.

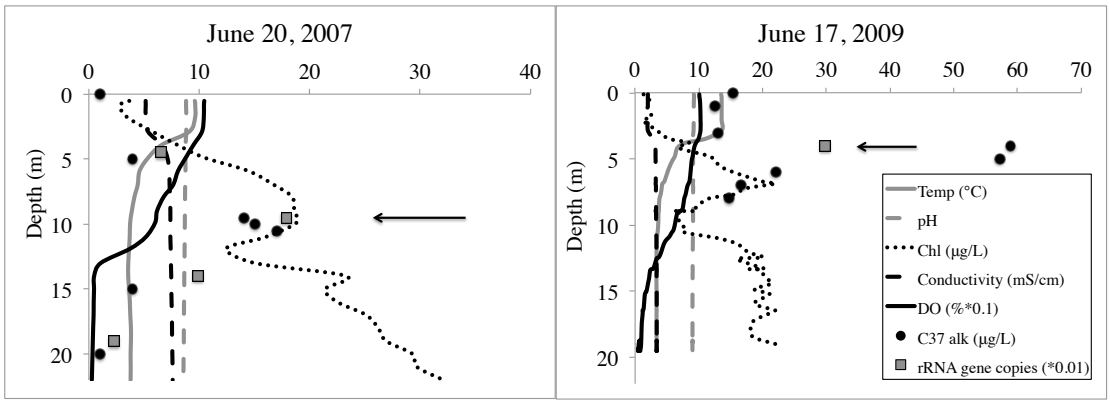

Figure 2. BrayaSø water column. Left panel shows 2007 sampling profile, and right panel shows 2009 sampling profile. Arrows denote sample depth for the samples we sequenced Dissolved oxygen (DO) is decreased one order of magnitude and haptophyte rRNA gene copy number is decreased two orders of magnitude to plot along the same axis.

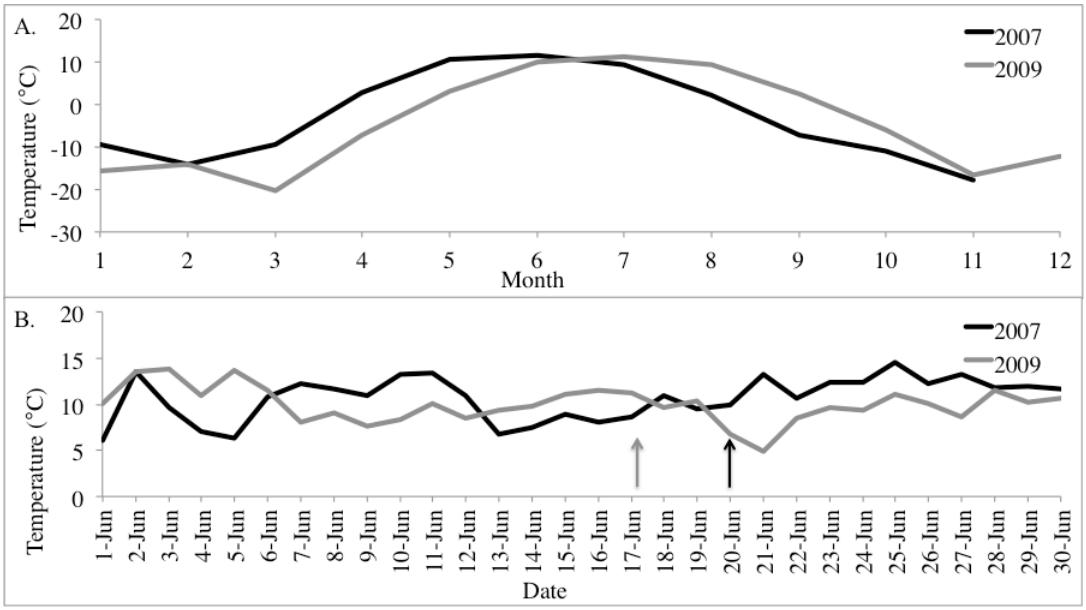

Figure 3. (A) June average air temperatures for 2007 and 2009. (B) Average monthly air temperatures for 2007 and 2009. Station data from Kangerlussuaq, Greenland (67.017° N, -50.700° W). http://www.ncdc.noaa.gov/oa/ncdc.html. Arrows denote sampling dates.

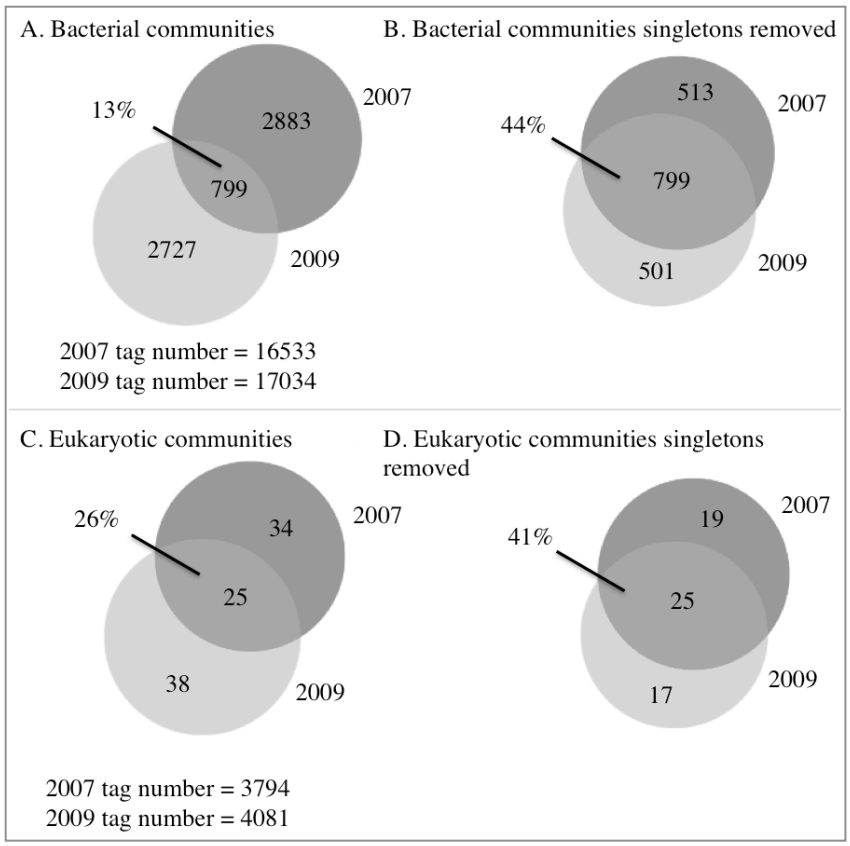

Figure 4. Venn diagrams of microbial communities. (A) Overlap between 2007 and 2009 bacterial communities, OTUs defined at 97% similarity (B) As in (A) with singletons removed (C) Overlap between 2007 and 2009 eukaryotic communities, OTUs defined at 94% similarity, (D) as in (C) with singletons removed.

CHAPTER 4

PRODUCTION AND TEMPERATURE SENSITIVITY OF LONG CHAIN ALKENONES IN CULTURES OF *PSEUDOISOCHRYSIS PARADOXA,* NOMEN NUDUM

SUSANNA THEROUX[1,2]

JAIME TONEY[1,‡]

LINDA AMARAL-ZETTLER[1,2,3]

YONGSONG HUANG[1]

[1] BROWN UNIVERSITY, *Department of Geological Sciences*

[2] MARINE BIOLOGICAL LABORATORY, *Josephine Bay Paul Center*

[3] BROWN UNIVERSITY, *Department of Ecology and Evolutionary Biology*

[‡] Present address: UNIVERSITY OF GLASGOW, *School of Geographical and Earth Sciences*

Prepared for:

Abstract

The alkenone unsaturation index (U^{k}_{37} and $U^{k'}_{37}$) serves as a critical tool for reconstructing paleotemperatures in marine environments. Lacustrine haptophyte species are distinct from their ubiquitous and well-studied marine counterparts, and the unknown genotypic effects on long chain alkenone production by lacustrine species have hindered the application of the U^{k}_{37} paleotemperature proxy to lake sediment records. Previous work in Lake George, North Dakota revealed the presence of two novel alkenone-producing haptophyte species, one that is closely related to the cultured organism *Pseudoisochrysis paradoxa* CCMP715. Like its close relative *Isochrysis galbana, P. paradoxa* is an alkenone producer, with a lipid fingerprint abundant in tetraunsaturated alkenones. We present here the first calibration of the U^{k}_{37} paleotemperature proxy for a culture of *P. paradoxa*. While the extent of its geographic distribution remains unclear, *P. paradoxa* undoubtedly contributes to sedimentary archives throughout the world's lacustrine regions, thus confirming the importance of culture-based and *in situ* U^{k}_{37} calibrations in lake environments.

1. Introduction

For over two decades, alkenone lipids have enabled the reconstruction of sea surface temperatures from marine sediments. Haptophyte algae in the group Isochrysidales are the exclusive producers of alkenone lipids, an extremely specific biomarker found throughout the world's pelagic, coastal, and lacustrine environments. The paleotemperature proxy U^{k}_{37} incorporates the relative abundance of the di-($C_{37:2}$), tri-($C_{37:3}$) and tetra-($C_{37:4}$) unsaturated alkenones as a proxy for water temperature (Brassell *et al.,* 1986), with greater degrees of unsaturation occurring at lower temperatures (Marlowe, 1984; Brassell *et al.,* 1986). A modified $U^{k'}_{37}$ (Prahl and Wakeham, 1987; Prahl *et al.,* 1988) has been widely applied to marine sediments where the tetraunsaturated alkenone is largely absent. The marine calibration of the $U^{k'}_{37}$ is used universally (Prahl and Wakeham, 1987; Muller *et al.,* 1988), the consequence of cosmopolitan *Emiliania huxleyi* and *Gephyrocapsa oceanica* being responsible for the majority of alkenone production in marine waters.

The recent discovery of widespread alkenone lipids distributions in lake sediments (Cranwell, 1985; Zink *et al.,* 2001; Chu *et al.,* 2005; D'Andrea and Huang, 2005; Pearson *et al.,* 2008; Toney *et al.,* 2010; Theroux *et al.,* 2010; Toney *et al.* 2011) has generated interest in the applicability of the U^{k}_{37} proxy to lake sediment archives. Alkenone lipids in lakes have been shown to reflect mean annual air temperature (Chu *et al.,* 2005) and *in situ* lake water temperature (Toney *et al.,* 2010; D'Andrea *et al.,* 2011). However, there is no universal U^{k}_{37} calibration for all lake environments, largely the result of the genetic diversity of the lacustrine haptophytes

(Theroux *et al*., 2010). The abundance of novel haptophyte species in these lake environments creates a new demand for *in situ* and culture-based U^k_{37} calibrations.

A genetic survey of Lake George, ND surface sediments revealed the presence of two haptophyte algal species responsible for alkenone production, one that is closely related (99% similar over 461bp of 18S rRNA gene sequence) to *Pseudoisochrysis paradoxa* (Theroux *et al*., 2010). Originally isolated from the brackish York River Estuary in the Chesapeake Bay, Virginia, USA, *P. paradoxa* has never been formally described (*nomen nudum*; Jordan *et al*., 2004). The 18S ribosomal RNA (rRNA) gene sequence for *P. paradoxa* CCMP715 (Genbank AM490999; Medlin *et al*., 2008) is 100% identical to that of the coastal/lacustrine haptophyte *Isochrysis galbana* (Genbank HM149541), highlighting the difficulty in distinguishing haptophyte species through 18S rRNA gene sequences alone. *Isochrysis galbana* is a known alkenone producer, and *P. paradoxa* was reported to have alkenones (Marlowe *et al*., 1984), although its U^k_{37} temperature calibration has never been determined.

The global extent of *P. paradoxa* populations is currently unknown; if it ecologically resembles its close relative, *I. galbana*, it can survive in a wide range of brackish environments and marine environments and contribute to the alkenone sediment records (Marlowe *et al*., 1990; Versteegh *et al*., 2001; Liu *et al*., 2009). Only a few species of haptophytes have been grown in culture to determine their U^k_{37} calibration: ubiquitous marine species *Emiliania huxleyi* (Prahl *et al*., 1988; Prahl *et al*., 2003; Conte *et al*., 1998) and *Gephyrocapsa oceanica* (Sawada *et al*., 1996; Conte *et al*., 1998), and lacustrine/brackish species *Isochrysis galbana* (Versteegh *et*

al., 2001) and *Chrysotila lamellosa* (Sun *et al*., 2007). Paleoclimate work in these brackish environments requires a better understanding of the ecology and controls on alkenone unsaturation of these haptophyte species. To this end, we grew *P. paradoxa* at a variety of temperatures to determine its alkenone unsaturation-temperature relationship.

2. **Methods**

2.1 Pseudoisochrysis paradoxa cultures

We purchased cultures of *Pseudoisochrysis paradoxa* (CCMP715, also known as CCAP 949/1, CCAPVA12, UTEX 1988) from the Provasoli-Guillard National Center for Marine Algae and Microbiota. We grew cultures in 0.2 μm filter-sterilized seawater amended with f/2 nutrients (Guillard, 1975) in full spectrum light on a 24:0 light:dark cycle. We verified the media was alkenone-free. We grew these batch cultures in triplicate volumes of 50 ml at 5, 10, 15, 21 and 24°C. We initiated cultures at equal cell concentrations of 8000 cells/ml using an inoculum from a culture previously acclimated to a given temperature. We monitored cell concentrations for three weeks using haemocytometer counts to ensure the cultures remained in the exponential growth phase. At the end of three weeks we harvested the cultures for alkenone analysis. One of the three 5°C cultures was discarded because it failed to grow.

2.2 Lipid analysis of cultures

We filtered culture material onto precombusted 47 mm glass microfiber filters glass fiber filters (Whatman, Piscataway, NJ) immediately froze at -20°C and then freeze-dried the filters overnight (Labconco, Kansas City, MO). We extracted the filters using three 20-minute bursts of sonication in dichloromethane (DCM), and ran the total lipid extracts on an Agilent 6890^{plus} Gas Chromatograph Flame Ionization Detector (GC-FID) for detection and quantification of alkenones using an internal C_{36} *n*-alkane standard and an external alkenone standard of known U^{k}_{37} -temperature value to ensure analytical precision (< 0.1°C). We used a Varian VF200 60m GC column (60m × 250μm × 0.10 μm) with the following temperature program: an initial temperature of 100°C (hold 1-min), ramp 20°C/min to 200°C (hold 1-min), ramp 4°C/min to 320°C (hold 0 min), hold at 320°C for 5-min.

3. Results

3.1 Cultures

Triplicate cultures of *P. paradoxa* behaved similarly within temperature regimes (**Table 1**). Growth rate and alkenone concentrations per cell displayed an inverse relationship (**Figure 1**), with growth rate highest at 21°C (54.30 divisions/day) and lowest at 5°C (0.01 div/day) and alkenone per cell concentrations were highest at 5°C (1.961 pg/cell) and lowest at 21°C (0.042 pg/cell). Final cell counts were highest at 21°C and lowest at 5°C (**Table 1**).

3.2 U^{k}_{37} *and* $U^{k'}_{37}$ *calibrations*

We calculated the U^{k}_{37} and $U^{k'}_{37}$ calibrations for the *P. paradoxa* cultures and plotted these versus growth temperature (**Figure 2**). The U^{k}_{37} calibration (U^{k}_{37} = 0.0226T-0.5149, R^2 = 0.905, RMSE = 0.05) was slightly more robust than the $U^{k'}_{37}$ calibration ($U^{k'}_{37}$ = 0.001T^2-0.0256T+0.1754, R^2 = 0.89, RMSE = 0.02), although only the U^{k}_{37} temperature calibration fit a linear regression. The $C_{37:4}$ alkenone comprised almost 40% of the C_{37} alkenones in the 5°C culture (**Figure 3**), although at all temperatures the $C_{37:3}$ was the dominant alkenone homolog (**Figure 4**). The U^{k}_{37} calibration for *P. paradoxa* had a similar slope to the *in situ* calibration from Lake George, ND (**Figure 5**; Toney *et al*., 2010), and clustered with other lake-based calibrations that were all distinct from the marine calibration. The $U^{k'}_{37}$ calibration was also distinct from previously reported $U^{k'}_{37}$ calibrations from other haptophyte cultures and environmental samples (**Figure 6**), and was best fit with a second order polynomial regression (**Figure 2**) whose y-intercept was closest to the marine U^{k}_{37} calibration.

4. Discussion

4.1 Growth stage and alkenone production

The *P. paradoxa* cultures demonstrated an enhanced alkenone production per cell at the slowest growth rates in the 5°C culture. This agrees with previous observations in cultures of *E. huxleyi* and *G. oceanica* that demonstrated highest alkenone concentrations per cell with lowest growth temperatures (Conte *et al*., 1998). However, low growth temperature often corresponds to low growth rates, so it is unclear if low growth temperature alone will result in enhanced alkenone

accumulation. Alkenones are presumed to serve as an energy storage molecule in haptophytes (Epstein *et al*., 2001; Eltgroth *et al*., 2005) and cellular production of alkenones increases during stationary growth phase and decreases after cultures are placed in the dark (Epstein *et al*., 2001; Eltgroth *et al*., 2005). The accumulation of alkenones at low temperatures and low growth rates may be the result of photosynthetic energy input exceeding the cell's capacity for growth and division (Roessler, 1990) as was observed at both low temperatures and low nutrient conditions. The *P. paradoxa* U^{k}_{37} calibration R^2 value of 0.905 supports the robustness of the alkenone temperature-dependent relationship. The coherency in *P. paradoxa* biological replicates further supports this linear relationship.

The affect of growth rate on haptophyte U^{k}_{37} and $U^{k'}_{37}$ is unclear (Conte *et al*., 1995; Epstein *et al*., 1998; Popp *et al*., 1998). In this study, we used a 24:0 light to dark regime to eliminate alkenone metabolism during darkness. In batch culture experiments, $U^{k'}_{37}$ decreases under nutrient stress and increases under prolonged darkness (Prahl *et al*., 2006), both conditions that may result in slower growth rates. Cultures of various strains of *E. huxleyi* grown in 12:12 or 0:24 light to dark regimes exhibited contrasting increases or decreases in $U^{k'}_{37}$ values depending on light regime (Epstein *et al*., 2001). These fluctuations in $U^{k'}_{37}$ were 0.013 to 0.029 units, a small amount, and therefore we do not believe the 24-hour light regime exerted a significant change in U^{k}_{37} values. Previous studies have shown that differences can exist between batch and continuous-culture methods (Popp *et al*., 1998) although it is debated which of these two methods more accurately replicates the natural environment. Growth

phase also has been shown to influence alkenone unsaturation in batch culture (Conte *et al*., 1998; Epstein *et al*., 1998), although continuous culture, and therefore constant growth state, imparted no change on U^{k}_{37} values (Popp *et al*., 1998). To control for this variation we harvested all cultures while in exponential growth phase.

4.2 Comparison to other species

Pseudoisochrysis paradoxa resembled other lacustrine haptophytes in the presence of the $C_{37:4}$ alkenone (Cranwell, 1985; Li *et al*., 1996; Zink *et al*., 2001). Similar to its close relative *I. galbana*, *P. paradoxa* cultures had predominant $C_{37:3}$ alkenones and the absence of the C_{38} methyl ketone (**Figure 3**). The abundance of the $C_{37:4}$ alkenone in the *P. paradoxa* cultures may explain the strong U^{k}_{37} calibration and weaker $U^{k'}_{37}$ calibration (**Figure 2**). The *P. paradoxa* U^{k}_{37} calibration matched previously reported lacustrine U^{k}_{37} calibrations (**Figure 5**), and had the same slope as the *in situ* calibration generated in Lake George, ND (Toney *et al*., 2010). This result was surprising as the Lake George *in situ* calibration was dominated by $C_{37:4}$ alkenones, whereas *P. paradoxa* cultures had abundant but not dominant $C_{37:4}$ alkenones. Their similarity in slope suggests the importance of incorporating the $C_{37:4}$ homolog into the calibration equation when its abundance is high (>20% of total C37 alkenones). The $U^{k'}_{37}$ calibration was best fit with a second-order polynomial regression, indicating a non-linear relationship between temperature and alkenone unsaturation. At lower temperatures, the *P. paradoxa* cultures actually had decreasing $U^{k'}_{37}$ values with increasing temperatures. This result indicates that the inclusion of the

tetraunsaturated alkenone is critical to the accurate reconstruction of temperature in cultures of *P. paradoxa*.

While *P. paradoxa* cultures had abundant $C_{37:4}$ alkenones, sediments and water samples from Lake George, ND (Toney *et al.,* 2010), as well as Lake BrayaSø in Greenland (D'Andrea and Huang, 2005; D'Andrea *et al.,* 2011) and Ace Lake, Antarctica (Coolen *et al.,* 2004) contain alkenone signatures with dominant $C_{37:4}$. Previously, percent $C_{37:4}$ ($C_{37:4}/C_{37:4}+C_{37:3}+C_{37:2}$) was proposed as a paleosalinity proxy (Schulz *et al.,* 2000; Roselle-Mele *et al.,* 2002) although the lack of correlation between salinity and percent $C_{37:4}$ in a global array of lake systems (Mercer *et al.,* 2005; Chu *et al.,* 2005; Toney *et al.,* 2010; Theroux *et al.,* 2010) suggests that the relationship between $C_{37:4}$ and salinity is a result of alkenone-producing species distributions across salinity tolerances. Percent $C_{37:4}$ can range from 8-53% in *C. lamellosa* (Rontani *et al.,* 2004) and 0-34% in *I. galbana* (Marlowe *et al.,* 1984), and up to 96% in a series of lakes in China (Chu *et al.,* 2005). In this study, percent $C_{37:4}$ ranged from 6% at 24°C to 40% at 5°C (**Table 1**, **Figure 4**). The alkenone concentrations per cell were within the range observed for *I. galbana* (0.0098-0.61 pg/cell; Versteegh *et al.,* 2001). The highest alkenone/cell concentrations observed for *P. paradoxa* (1.961 pg/cell) was close to the range observed in an *Isochrysis* sp. culture (1.8 pg/cell; Marlowe, 1984).

4.3 Application to the natural environment

The similarity of the *P. paradoxa* calibration to the *in situ* Lake George calibration further confirms the utility of *in situ* calibrations. More than one alkenone

producer occurs in Lake George waters (Theroux *et al*., 2010; Toney *et al*., 2012), and the dual methods of alkenone production by two distinct species likely explains the offset between the Lake George *in situ* U^{k}_{37} calibration and the *P. paradoxa* culture-based calibration. The Lake George alkenone fingerprint was dominated by the tetraunsaturated $C_{37:4}$ alkenone, unlike *P. paradoxa* cultures with abundant $C_{37:4}$ but dominant $C_{37:3}$ (Toney *et al*., 2010; Theroux *et al*. 2010; Toney *et al*., 2011). The presence of the $C_{37:4}$ homolog in both Lake George waters and the *P. paradoxa* cultures supports the use of the U^{k}_{37} proxy instead of the $U^{k'}_{37}$ index. Of note, alkenone production in Lake George occurs during a brief window of time following spring ice-off (Toney *et al*., 2010), and alkenone records likely reflect spring bloom water temperatures alone.

The *P. paradoxa* alkenone "fingerprint" closely resembled the alkenone signature of *I. galbana*, but their $U^{k'}_{37}$ calibrations are very different (**Figure 5**). Studies have shown that different geographically isolated strains of the same species of haptophyte, *E. huxleyi* and *G. oceanica*, have different modes of alkenone unsaturation with temperature (Conte *et al*., 1995; Conte *et al*., 2006). Therefore, it is no surprise that different but closely related haptophytes will also possess different U^{k}_{37} calibrations. The slopes of many U^{k}_{37} calibrations across haptophyte species are similar (**Figure 5**). This suggests that the temperature response of haptophytes is consistent, but the determining factor of the U^{k}_{37} calibration y-intercept, is still unknown. Reconstructing temperature from organisms with similar U^{k}_{37} calibration

slopes will prevent the reconstruction of absolute temperature but may still be able to reconstruct relative temperature change.

5. Conclusions

The application of the U^k_{37} paleotemperature proxy to lake sediments depends upon a robust U^k_{37} calibration. As evidenced by studies to date, lakes are not uniform in their alkenone-producing haptophyte populations, thereby complicating the use of a universal lake U^k_{37} proxy calibration. Our study is one of few to cultivate an individual haptophyte species at various temperatures to calibrate the U^k_{37} - temperature relationship. *Pseudoisochrysis paradoxa* has proven to be distinct in its mode of alkenone unsaturation, with a calibration equation that is significantly different from its close genetic relative *Isochrysis galbana*. The similarity between the Lake George *in situ* calibration and the *P. paradoxa* calibration reported in this study gives credence to the applicability of *in situ* U^k_{37} calibrations. As with all alkenone-based temperature reconstructions, it is important to verify the continuity of alkenone distributions throughout downcore sediments in order to apply a single calibration to the temperature reconstruction. We anticipate future studies comparing both *in situ* and culture-based U^k_{37} calibrations to further resolve species-specific modes of alkenone production in lake environments.

Acknowledgements

This work was supported by a National Science Foundation award to Y. Huang (EAR-1122749) and L. Amaral-Zettler (EAR-1124192), a Brown SEED fund to Y.

Huang and L. Amaral-Zettler, and an American Association of University Women dissertation fellowship to S. Theroux.

References

Brassell SC, Eglinton G. (1986). Molecular geochemical indicators in sediments. *Acs Symposium Series* **305**:10-32.

Brassell SC, Eglinton G, Marlowe IT, Pflaumann U, Sarnthein M. (1986). Molecular stratigraphy - a new tool for climatic assessment. *Nature* **320**:129-133.

Chu GQ, Sun Q, Li SQ, Zheng MP, Jia XX, Lu CF, Liu JQ, Liu TS. (2005). Long-chain alkenone distributions and temperature dependence in lacustrine surface sediments from China. *Geochimica Et Cosmochimica Acta* **69**:4985-5003.

Conte MH, Sicre MA, Ruhlemann C, Weber JC, Schulte S, Schulz-Bull D, Blanz T. (2006). Global temperature calibration of the alkenone unsaturation index $U^{k'}_{37}$ in surface waters and comparison with surface sediments. *Geochemistry Geophysics Geosystems* **7**:doi:10.1029/2005GC001054.

Conte MH, Thompson A, Lesley D, Harris RP. (1998). Genetic and physiological influences on the alkenone/alkenoate versus growth temperature relationship in *Emiliania huxleyi* and *Gephyrocapsa oceanica*. *Geochimica Et Cosmochimica Acta* **62**:51-68.

Conte MH, Thompson A, Eglinton G, Green JC. (1995). Lipid biomarker diversity in the coccolithophorid *Emiliania huxleyi* (Prymnesiophyceae) and the related species *Gephyrocapsa oceanica*.*Journal of Phycology* **31**:272–282.

Coolen MJL, Muyzer G, Rijpstra WIC, Schouten S, Volkman JK, Sinninghe Damsté JS. (2004). Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake. *Earth and Planetary Science Letters* **223**:225-239.

Cranwell PA. (1985). Long-Chain Unsaturated-Ketones in Recent Lacustrine Sediments. *Geochimica Et Cosmochimica Acta* **49**:1545-1551.

D'Andrea WJ, Huang Y. (2005). Long chain alkenones in Greenland lake sediments: Low δ13C values and exceptional abundance. *Organic Geochemistry* **36**:1234-1241.

D'Andrea WJ, Huang Y, Fritz SC, Anderson NJ. (2011). Abrupt Holocene climate

change as an important factor for human migration in West Greenland. *Proceedings of the National Academy of Sciences*

D'Andrea WJ, Huang Y, Fritz SC, Anderson NJ. (2011). Abrupt Holocene climate change as an important factor for human migration in West Greenland. *Proceedings of the National Academy of Sciences* **108**:9765.

Eltgroth ML, Watwood RL, Wolfe GV. (2005). Production and cellular localization of neutral long-chain lipids in the haptophyte algae *Isochrysis galbana* and *Emiliania huxleyi*. *Journal of Phycology* **41**:1000-1009.

Epstein BL, D'Hondt S, Hargraves PE. (2001). The possible metabolic role of C-37 alkenones in *Emiliania huxleyi*. *Organic Geochemistry* **32**:867-875.

Epstein BL, D'Hondt S, Quinn JG, Zhang JP, Hargraves PE. (1998). An effect of dissolved nutrient concentrations on alkenone-based temperature estimates. *Paleoceanography* **13**:122-126.

Guillard RRL (1975). Culture of phytoplankton for feeding marine invertebrates. In *Culture of marine invertebrate animals*. WL Smith, MH Chanley, eds. Plenum Press, New York, p. 26.

Jordan RW, Cros L, Young JR. (2004). A revised classification scheme for living haptophytes. *Micropaleontology* **50**:55-79.

Li JG, Philp RP, Pu F, Allen J. (1996). Long-chain alkenones in Qinghai Lake sediments. *Geochimica Et Cosmochimica Acta* **60**:235-241.

Liu Y, Yao T, Zhu L, Jiao N, Liu X, Zeng Y, Jiang H. (2009). Bacterial diversity of freshwater alpine lake Puma Yumco on the Tibetan Plateau. *Geomicrobiology Journal* **26**:131-145.

Marlowe IT, Brassell SC, Eglinton G, Green JC. (1990). Long-chain alkenones and alkyl alkenoates and the fossil coccolith record of marine sediments. *Chemical Geology* **88**:349 - 375.

Marlowe IT, Green JC, Neal AC, Brassell SC, Eglinton G, Course PA. (1984). Long-Chain (**n**-C37-C39) alkenones in the prymnesiophyceae. Distribution of alkenones and other lipids and their taxonomic significance. *British Phycological Journal* **19**:203-216.

Medlin LK, Saez AG, Young JR. (2008). A molecular clock for coccolithophores and implications for selectivity of phytoplankton extinctions across the K/T boundary. *Marine Micropaleontology* **67**:69-86.

Mercer JL, Zhao MX, Colman SM. (2005). Seasonal variations of alkenones and U-

37(K) in the Chesapeake Bay water column. *Estuarine Coastal and Shelf Science* **63**:675-682.

Müller PJ, Kirst G, Ruhland G, von Storch I, Rosell-Mele A. (1998). Calibration of the alkenone paleotemperature index $U^{k'}_{37}$ based on core-tops from the eastern South Atlantic and the global ocean (60°N-60°S). *Geochimica Et Cosmochimica Acta* **62**:1757-1772.

Pearson EJ, Juggins S, Farrimond P. (2008). Distribution and significance of long-chain alkenones as salinity and temperature indicators in Spanish saline lake sediments. *Geochimica Et Cosmochimica Acta* **72**:4035-4046.

Popp BN, Kenig F, Wakeham SG, Laws EA, Bidigare RR. (1998). Does Growth Rate Affect Ketone Unsaturation and Intracellular Carbon Isotopic Variability in Emiliania Huxleyi? *Paleoceanography* **13**:35-41.

Prahl FG, Wolfe GV, Sparrow MA. (2003). Physiological impacts on alkenone paleothermometry. *Paleoceanography* **18**:doi:10.1029/2002PA000803.

Prahl FG, Muehlhausen LA, Zahnle DL. (1988). Further evaluation of long-chain alkenones as indicators of paleoceanographic conditions. *Geochimica Et Cosmochimica Acta* **52**:2303-2310.

Prahl FG, Wakeham SG. (1987). Calibration of unsaturation patterns in long-chain ketone compositions for paleotemperature assessment. *Nature* **330**:367-369.

Prahl FG, Mix AC, Sparrow MA. (2006). Alkenone paleothermometry: Biological lessons from marine sediment records off western South America. *Geochimica et Cosmochimica Acta* **70**:101-117.

Roessler PG. (1990). Environmental control of glycerolipid metabolism in microalgae: commercial implications and future research directions. *Journal of Phycology* **26**:393–399.

Rontani JF, Beker B, Volkman JK. (2004). Long-chain alkenones and related compounds in the benthic haptophyte *Chrysotila lamellosa* Anand HAP 17. *Phytochemistry* **65**:117-126.

Rosell-Melé A, Jansen E, Weinelt M. (2002). Appraisal of a molecular approach to infer variations in surface ocean freshwater inputs into the North Atlantic during the last glacial. *Global and Planetary Change* **34**:143-152.

Sawada K, Handa N, Shiraiwa Y, Danbara A, Montani S. (1996). Long-chain alkenones and alkyl alkenoates in the coastal and pelagic sediments of the northwest North Pacific, with special reference to the reconstruction of *Emiliania huxleyi* and *Gephyrocapsa oceanica* ratios. *Organic Geochemistry*

24:751 - 764.

Schulz H-M, Schöner A, Emeis K-C. (2000). Long-chain alkenone patterns in the Baltic sea--an ocean-freshwater transition. *Geochimica Et Cosmochimica Acta* **64**:469-477.

Sun Q, Chu GQ, Liu GX, Li S, Wang XH. (2007). Calibration of alkenone unsaturation index with growth temperature for a lacustrine species, *Chrysotila lamellosa* (Haptophyceae). *Organic Geochemistry* **38**:1226-1234.

Theroux SM, D'Andrea WJ, Toney JL, Amaral-Zettler LA, Huang Y. (2010). Phylogenetic diversity and evolutionary relatedness of alkenone producing haptophyte algae in lakes: Implications for continental paleotemperature reconstructions. *Earth and Planetary Science Letters* **300**:311-320.

Toney JL, Huang Y, Fritz SC, Baker PA, Grimm E, Nyren P. (2010). Climatic and environmental controls on the occurrence and distributions of long chain alkenones in lakes in the interior United States. *Geochimica Et Cosmochimica Acta*

Toney JL, Theroux S, Andersen RA, Coleman A, Amaral-Zettle L, Huang Y. (2012). Culturing of the first 37:4 predominant lacustrine haptophyte: Geochemical, biochemical, and genetic implications. *Geochimica et Cosmochimica Acta* **78**:51 - 64.

Versteegh GJM, Riegman R, de Leeuw JW, Jansen JHF. (2001). $U^{k'}_{37}$ values for *Isochrysis galbana* as a function of culture temperature, light intensity and nutrient concentrations. *Organic Geochemistry* **32**:785-794.

Zink KG, Leythaeuser D, Melkonian M, Schwark L. (2001). Temperature dependency of long-chain alkenone distributions in recent to fossil limnic sediments and in lake waters. *Geochimica Et Cosmochimica Acta* **65**:253-265.

Table 1. Average growth rate, cell concentrations and alkenone concentration of *Pseudoisochrysis paradoxa*.

Table 1. Average growth rate, cell concentrations, and alkenone concentrations for *P. paradoxa* cultures. All cultures were grown in triplicate. Standard deviation (SD) as noted.

Temperature (°C)	Growth rate (div/day)	*SD*	Final cell concentration (cells/ml)	*SD*	Alkenones (g/L)	*SD*	Alkenones (pg/cell)	*SD*	Uk37	*SD*	%	C37:4	*SD*
5	0.01	*0.00*	10,000	*0.00*	20	*6*	1.961	*0.601036*	-0.350	*0.035*		39%	2%
10	21.99	*0.50*	4,053,333	*92,376*	373	*18*	0.092	*0.003440*	-0.290	*0.004*		31%	0.4%
15	46.70	*4.74*	8,600,000	*871,780*	960	*109*	0.112	*0.018491*	-0.237	*0.004*		25%	0.4%
21	54.30	*10.87*	10,000,000	*2,000,000*	406	*39*	0.042	*0.010556*	-0.075	*0.014*		13%	0.3%
24	14.45	*2.51*	2,666,667	*461,880*	532	*74*	0.202	*0.030216*	0.088	*0.019*		6%	1%

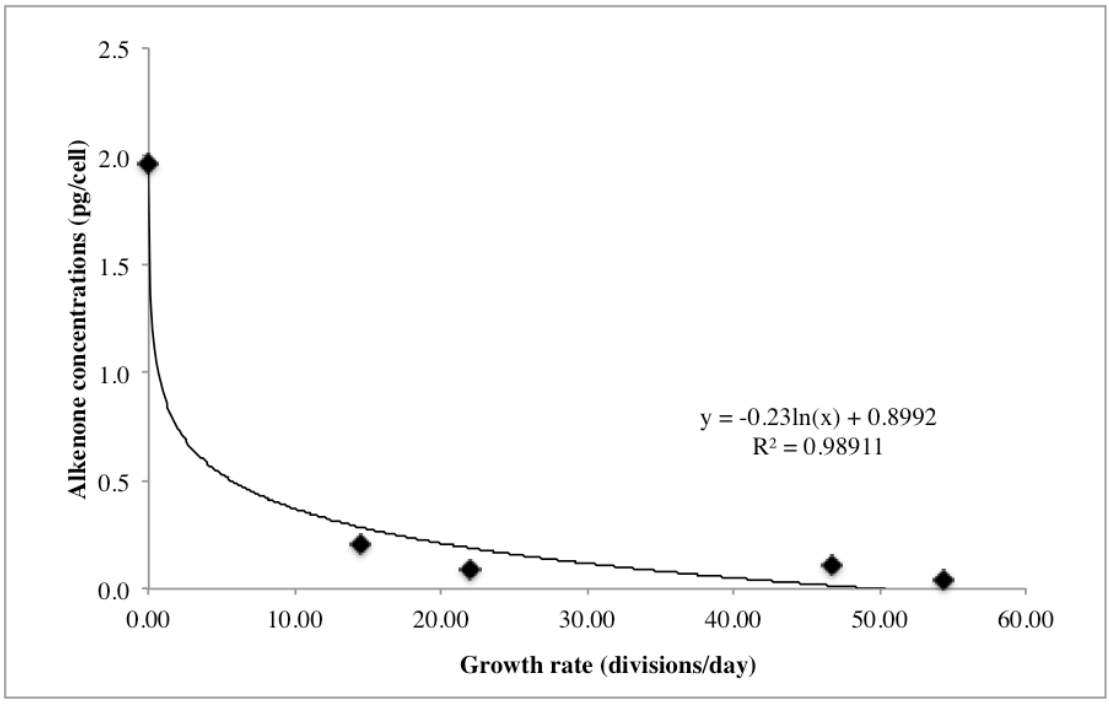

Figure 1. Alkenone cell concentration versus growth rate in *P. paradoxa* cultures.

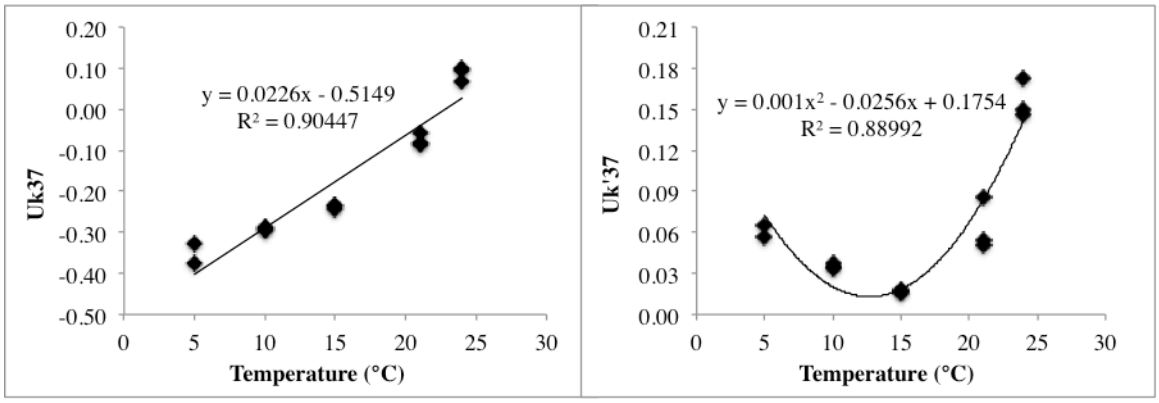

Figure 2. (Left panel) ***Pseudoisochrysis paradoxa*** $\mathbf{U^{k}_{37}}$ **calibration (Right panel)** ***Pseudoisochrysis paradoxa*** $\mathbf{U^{k'}_{37}}$ **calibration.**

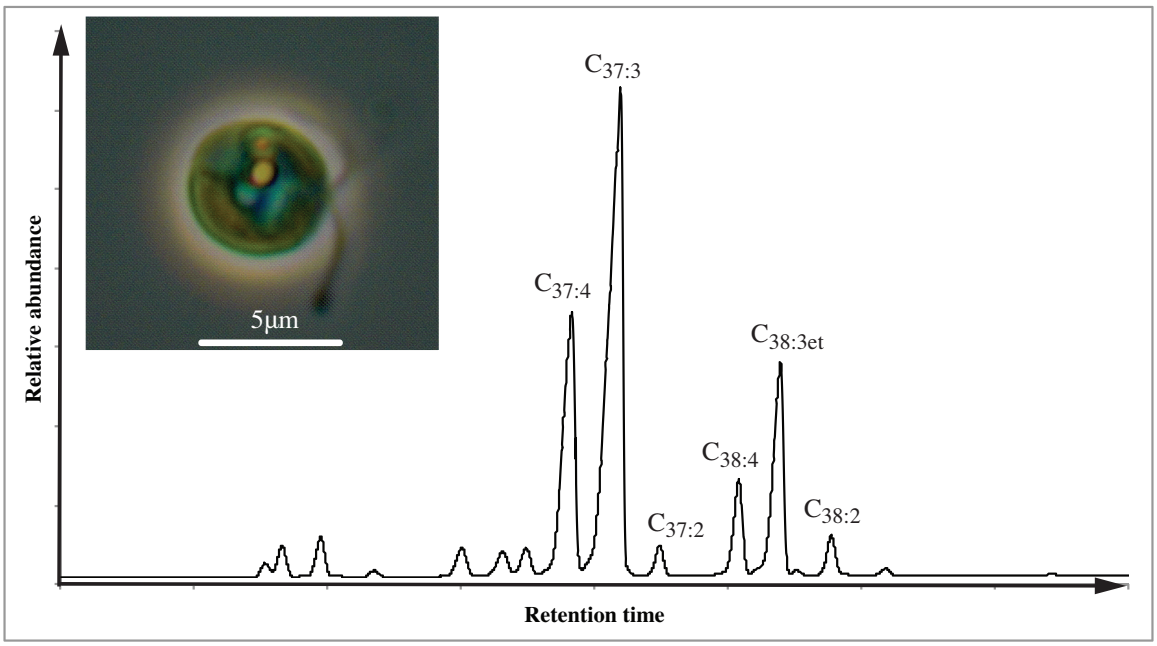

Figure 3. Gas chromatogram of *P. paradoxa* cultures at 10°C. Inset: Photomicrograph of *P. paradoxa* culture. Scale bar 5*µ*m.

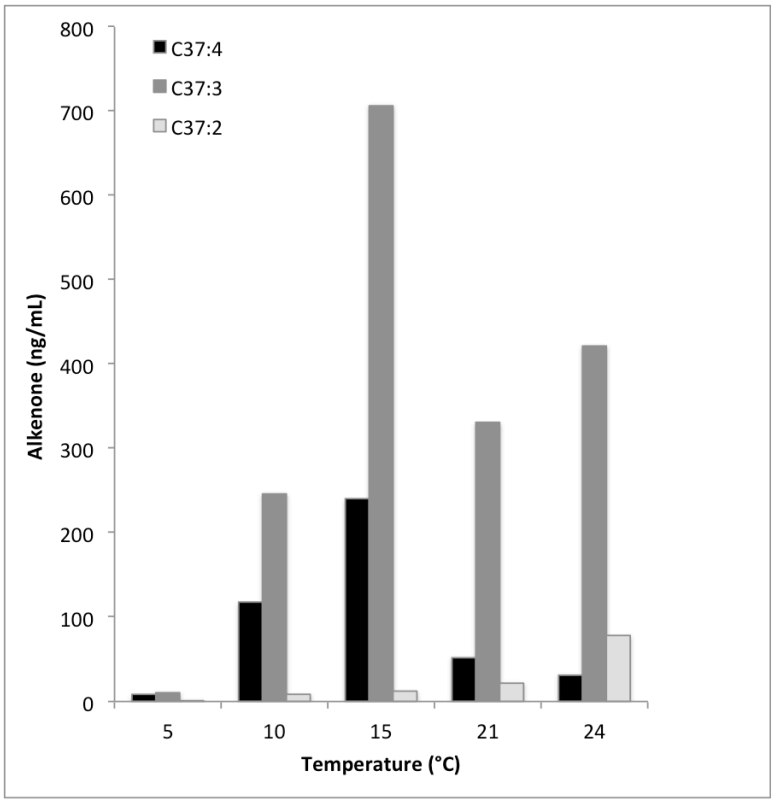

Figure 4. Average alkenone concentrations for *P. paradoxa* cultures at different temperatures.

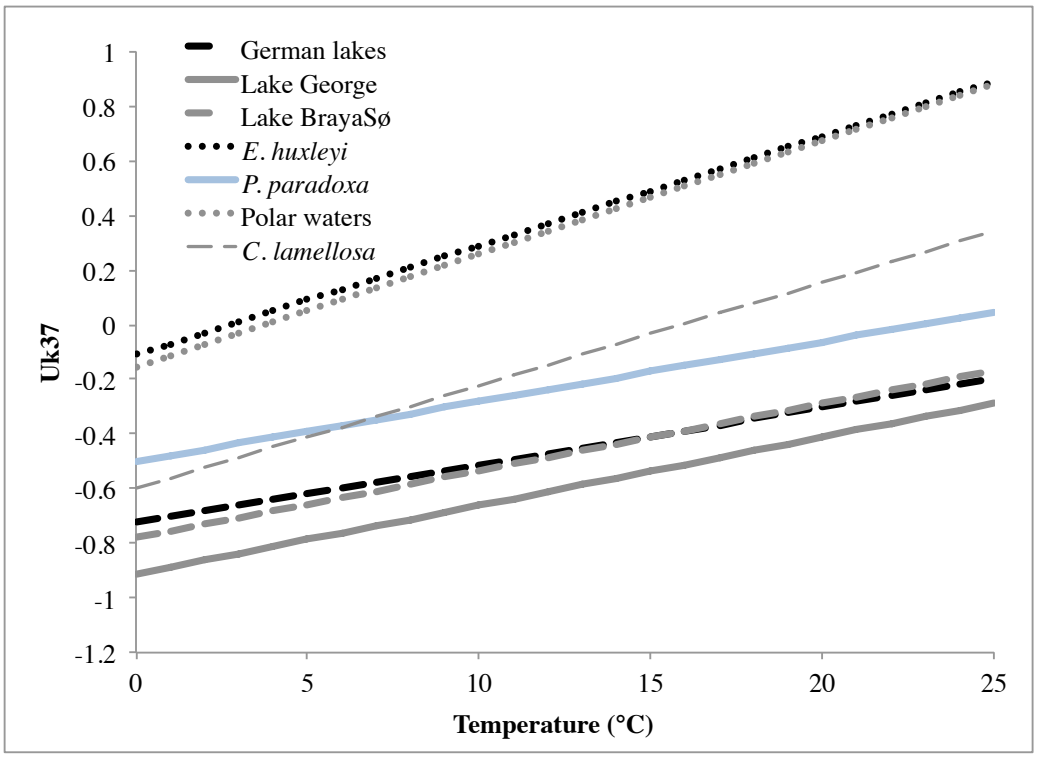

Figure 5. Comparison of *P. paradoxa* U^{k}_{37} calibration with other species and lake-based calibrations. References are as follows: German Lakes, Zink *et al.,* 2001; Lake George *in situ*, Toney *et al.,* 2010; Lake BrayaSø, D'Andrea *et al.,* 2011; *E. huxleyi*, Prahl *et al.,* 1988; P. paradoxa, this paper; Polar waters, Sikes and Volkman 1993; Lake George *in situ* 2011, Theroux, this work; *C. lamellosa*, Sun *et al.,* 2007.

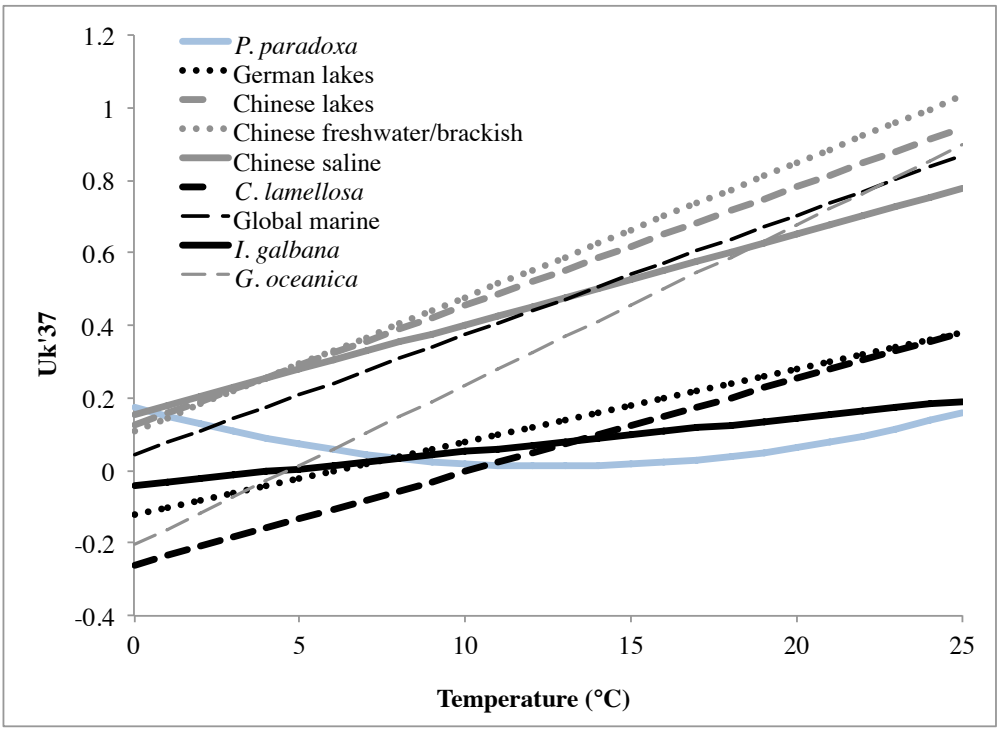

Figure 6. Comparison of *P. paradoxa* $U^{k'}_{37}$ calibration with other species and lake-based calibrations. References are as follows: *P. paradoxa*, this paper; German lakes, Zink *et al.,* 2001;Chinese lakes, Chu *et al.,* 2005; Chinese freshwater/brackish lakes, Chu *et al.,* 2005; Chinese saline, Chu *et al.,* 2005; *C. lamellosa*, Sun *et al.,* 2007; Global marine, Prahl and Wakeham 1987; *I. galbana*; Versteegh *et al.,* 2001; *G. oceanica*, Sawada *et al.,* 1996.

CHAPTER 5

HAPTOPHYTE DIVERSITY ALONG A SALINITY GRADIENT IN LAKES OF THE QINGHAI REGION

SUSANNA THEROUX[1,2]

ELIZABETH K. THOMAS[1]

LINDA AMARAL-ZETTLER[1,2,3]

YONGSONG HUANG[1]

[1] BROWN UNIVERSITY, *Department of Geological Sciences*

[2] MARINE BIOLOGICAL LABORATORY, *Josephine Bay Paul Center*

[3] BROWN UNIVERSITY, *Department of Ecology and Evolutionary Biology*

Prepared for:
Geochimica et Cosmochimica Acta

Abstract

The alkenone unsaturation index (U^{k}_{37} or $U^{k'}_{37}$) serves as a robust paleotemperature proxy in aquatic environments. In environments characterized by high $C_{37:4}$ alkenone concentrations, a $C_{37:4}$-based paleosalinity proxy has been proposed as a means to reconstruct ancient salinity levels from lake sediments. The Qinghai region on the Chinese-Tibetan plateau provides an ideal environment to test this proposed paleosalinity proxy given its abundance of alkenone-containing lakes, its range of lake salinities and a shared mean annual air temperature among the lakes. We sampled Lake Qinghai and three of its satellite lakes to gauge alkenone-producing haptophyte species across a salinity gradient. Our geochemical results agreed with previous research that demonstrated a relationship between $C_{37:4}$ alkenone concentrations and water salinity. However, our DNA sequencing analyses revealed the presence of multiple alkenone-producing haptophytes, with different alkenone profiles, in a single lake. Instead of a single haptophyte producing varying alkenone signatures, we propose that $C_{37:4}$ alkenones trend with lake salinity as a result of shifts in haptophyte community composition. This study is the first to identify the haptophyte species responsible for alkenone production in Lake Qinghai and provides an alternate explanation to the observed relationship between lacustrine $C_{37:4}$ concentrations and salinity.

1. Introduction

For over two decades, alkenone lipids have enabled the reconstruction of sea surface temperatures from marine sediments. Haptophyte algae in the group Isochrysidales are the exclusive producers of alkenone lipids, an extremely specific biomarker found throughout the world's pelagic, coastal, and lacustrine environments. The paleotemperature proxy U^{k}_{37} incorporates the relative abundance of the di-($C_{37:2}$), tri-($C_{37:3}$) and tetra-($C_{37:4}$) unsaturated alkenones as a proxy for water temperature (Brassell *et al.,* 1986), with greater degrees of unsaturation occurring at lower temperatures (Marlowe, 1984; Brassell *et al.,* 1986). A modified $U^{k'}_{37}$ (Prahl and Wakeham, 1987; Prahl *et al.,* 1988) has been widely applied to marine sediments where the tetraunsaturated alkenone is largely absent. The marine calibration of the $U^{k'}_{37}$ is used universally (Prahl and Wakeham, 1987; Muller *et al.,* 1988), the consequence of cosmopolitan *Emiliania huxleyi* and *Gephyrocapsa oceanica* being responsible for the majority of alkenone production in marine waters. In lacustrine waters, *Isochrysis galbana* and *Chrysotila lamellosa* (Marlowe *et al.,* 1984) are known alkenone-producers, and recent research has identified novel species of haptophytes that fall within the Isochrysidales and are therefore potential alkenone producers (Coolen *et al.,* 2004; D'Andrea *et al.,* 2006; Theroux *et al.,* 2010).

Dominant $C_{37:4}$ alkenone abundances ($\%_{37:4} > (C_{37:3}+C_{37:2})$) have been reported in many lacustrine settings (Cranwell, 1985; Volkman *et al.,* 1988; Li *et al.,*1996; Thiel *et al.,* 1997; Wang and Zheng, 1998; Sun *et al.,* 2004; Coolen *et al., 2*004; Chu et al, 2005; D'Andrea and Huang 2005; Liu *et al.,* 2006; Toney *et al.,* 2010). In contrast, abundant or dominant $C_{37:4}$ concentrations in marine environments are restricted to the North Atlantic

and Nordic Seas (Rosell-Mele, 1998; Sicre *et al*., 2002) and nearshore environments (Schulz *et al*., 2000; Chu *et al*., 2005; Mercer *et al*., 2005). The accumulation of more unsaturated alkenones has previously been observed when haptophytes are grown under nutrient deplete conditions, and therefore lower growth rates (Prahl *et al*., 2003; Prahl *et al*., 2006). However, given their abundance in low salinity waters, $C_{37:4}$ alkenones have been proposed as a paleosalinity proxy, with $C_{37:4}$ concentrations decreasing as salinity increases. Since first proposed by Rosell-Mele *et al*., (2002), numerous studies have attempted to confirm or refute this salinity proxy and the results are still inconclusive. On regional scales, the $\%C_{37:4}$ ($[C_{37:4}]$ / $[C_{37:2}]+[C_{37:3}]+[C_{37:4}]$) salinity relationship has been observed in northern high latitude oceans (Rosell-Mele, 1998; Sicre *et al*., 2002; Harada *et al*., 2003) and has been used to reconstruct paleo-salinity in the Sea of Okhotsk (Seki *et al*., 2005). Concentrations of $C_{37:4}$ fail to accurately reconstruct salinity levels from lakes in Greenland (D'Andrea and Huang 2005) and Chesapeake Bay (Mercer *et al*., 2005) and a linear relationship between salinity and $\%C_{37:4}$ was not found in a global survey of alkenone-containing lake sediments (Theroux *et al*., 2010) and a survey of lakes in the Northern Great Plains (Toney *et al*., 2010).

Temperature effects may overwrite patterns in $C_{37:4}$ concentrations with salinity (Prahl *et al*., 1988; Bendle *et al*., 2005), and a thorough examination of the paleosalinlity proxy is best reserved for a small geographical boundary with a large salinity gradient and a shared temperatures regime. The Lake Qinghai region on the Tibetan plateau provides an ideal setting for this type of study, with Lake Qinghai and its surrounding satellite lakes providing a salinity range between 1-76 ppt and a percent $C_{37:4}$ range between 10-46% (Liu et al., 2008). Two studies have examined the salinity-$\%C_{37:4}$

relationship in the waters and surface sediments of this region, reporting a consistent increase in $C_{37:4}$ concentrations with decreasing salinities (Liu *et al.,* 2008, 2011). These authors posit that a similar species of haptophyte may be responsible for alkenone production at these different sampling locations, and that salinity dictates the abundance of $C_{37:4}$ produced by the haptophyte. Based on their results, Liu *et al*. (2008) proposed a $C_{37:4}$-based paleosalinity proxy for application to lakes in the Qinghai region.

An examination of haptophyte community composition in the sediment paleorecord from Ace Lake, Antarctica revealed a clear shift in haptophyte species, and corresponding alkenone fingerprints, with changes in salinity (Coolen *et al*., 2004). With our study, we aim to verify if salinity or LCA-producing haptophyte species diversity is responsible for variations in alkenone signatures in a modern day lake environment. We examine a series of lakes on the Tibetan plateau that were previously the subject of alkenone-based paleosalinity studies (Chu *et al.,* 2005; Liu *et al*., 2008) and use DNA sequencing to identify the haptophyte species present in these environments. Our study provides the first genetic confirmation of the haptophyte species responsible for alkenone production in Lake Qinghai and the first genetic-based test of $C_{37:4}$ concentrations across a salinity gradient in modern day lake environments.

2. Methods

2.1 Site description

The Lake Qinghai region on the Qinghai-Tibetan plateau in China is characterized by a semi-arid/arid and cold continental climate (Liu *et al*., 2011). The high altitude of 2180-3200m maintains a cool mean summer air temperature of 11-14°C (Shi *et al*., 1958;

Henderson *et al*., 2003; Colman *et al*., 2007). A negative precipitation-evaporation rate maintains the elevated salinity of lakes in the region. Lake Qinghai (**Figure 1**) is the largest inland brackish lake in China, with an average salinity of 12 ppt. Nearby lakes Erhai and HaiyanWan have salinities of 0.97 ppt and 18, respectively (**Table 1**). Gahai lake is to the west, at the northeastern edge of the Qaidam Basin, with an elevated salinity of 76.4 ppt (**Table 1**).

Water sampling

We collected water samples in May of 2010. For each sample, we collected two liters of water using a Van Dorn water sampler and preserved one liter for DNA analyses and one liter for alkenone analyses. For alkenone analysis, we filtered one liter of water onto a pre-combusted (550°C) GF/F 0.7um, 47mm glass filter and kept it frozen at -20°C until freeze drying and extraction. For DNA analysis, we filtered one liter of water onto a 0.2um Sterivex™ filter (Millipore, Billerica, MA, USA), flooded the filter with Puregene lysis buffer (Qiagen, Valencia, CA, USA) and froze it at -20°C until processing.

Lakes Erhai and HaiyanWan were sampled at the surface (**Table 1**). Lake Gahai was sampled at about 2m depth. Lake Qinghai was sampled twice at location QH2, at 3m (QH2_3m) and at 6m depth (QH2_6m), and once at location QH12 at 8m depth (QH12_8m) (**Figure 2**).

2.2 Alkenone extraction and analysis

We extracted the alkenone filters using three 20-minute bursts of sonication in dichloromethane (DCM), and ran the total lipid extracts on an Agilent 6890plus Gas

Chromatograph Flame Ionization Detector (GC-FID) for detection and quantification of alkenones using an internal C36 n-alkane standard and an external alkenone standard of known U^{k}_{37}-temperature value to ensure analytical precision (< 0.1°C). We used a Varian VF200 60m GC column (60m × 250 μm × 0.10 μm) with the following temperature program: an initial temperature of 100°C (hold 1-min), ramp 20°C/min to 200°C (hold 1-min), ramp 4°C/min to 320°C (hold 0 min), hold at 320°C for 5-min.

2.3 DNA extraction and sequencing

We extracted Sterivex™ filters using a Puregene Cell Kit (Qiagen) according to the manufacturer's instructions. Genomic DNA was further purified using the Qiagen DNA Purification to remove proteins and other contaminants that inhibit PCR reactions: We quantified total extracted genomic DNA yields using a NanoDrop nucleic acid spectrophotometer (Thermo Scientific, Wilmington, DE).

We amplified genomic DNA using haptophyte-specific oligonucleotide (Coolen *et al*., 2004; Simon *et al*., 2000) primers targeting 18S rRNA coding regions. Forward and reverse primers corresponded to *Escherichia coli* 18S rRNA positions 429 and 887, respectively. Polymerase chain reactions (PCRs) were performed on an Eppendorf Gradient Thermocycler (Hamburg, Germany) with the following conditions after D'Andrea *et al*., (2006): 4 min initial denaturing at 96 °C, 35 cycles of denaturing for 30 s at 94 °C, followed by 40 s primer annealing at 55 °C and primer extension 40 s at 72 °C, with a final extension of 10 min at 72 °C. PCR reactions were run in 50 µl volume using Promega GoTaq polymerase (Madison, WI, USA). Triplicate PCR products were pooled for each sample, and templates were purified using an Invitrogen PureLink Purification

kit (Carlsbad, CA, USA) with the high cut-off binding buffer to eliminate fragments < 200 bp. Purified PCR products were A-tailed and purified using the PureLink high cut-off kit. Cloning was performed using the Invitrogen TOP10 cloning kit with electro-competent cells according to the manufacturer's instructions. One hundred clones were picked for each sample. Plasmid DNA was isolated using a RevPrep Orbit robotic template preparation instrument (Genomic Solutions, Ann Arbor, MI), and prepared templates were sequenced on an ABI 3730XL (Applied Biosystems, Foster City, CA) capillary sequencer using the BigDye protocol with universal M13 forward and reverse primers according to the manufacturer's instructions. All sequencing was performed at the Marine Biological Laboratory W. M. Keck Ecological and Evolutionary Genetics Facility. All sequences have been deposited in GenBank under accession numbers XXXXX-XXXXXX.

2.4 Bioinformatics and phylogenetic reconstructions

A bioinformatics pipeline using the programs phred, cross-match, and phrap, translated chromatograms into base-calls and associated quality scores, removed vector sequences and assembled forward and reverse reads into full-length sequences for each of the cloned PCR amplicons (Ewing and Green, 1998; Lasek- Nesselquist *et al.,* 2008). Only sequences greater than 400 bp and with a complete forward and reverse primer were retained. Base-calls were verified and sequences were manually edited with the program Consed (Gordon *et al.,* 1998) for chromatogram viewing. We screened edited sequences for chimeras using the computer program Bellerophon (Huber *et al.,* 2004) and the edited dataset contained 182 sequences. Assembled sequences were aligned using the ARB

software program v. 07.07.11 (Ludwig *et al*., 2004) against the July 2012 Silva 111 Ref database (Pruesse *et al*., 2007) using the FastAligner option followed by manual adjustment. Operational Taxonomic Units (OTUs) were obtained using UCLUST (Edgar, 2010) with a 97% cut-off criterion. OTU cutoffs of less than 97% failed to differentiate known reference taxa sequences. Consensus sequences for each OTU were selected for tree construction.

A custom filter of 429 positions was constructed manually for the Bayesian analyses of 18S rRNA sequences. We selected *Cyclonexis annularis, Chrysoxys sp*., *Ochromonas danica, Odontella sinensis* and *Thraustochytrium multirudimentale* as outgroups for our Bayesian analyses after de Vargas *et al*., (2007). We subjected our datasets to a Bayesian analysis using MrBayes version 3.0b4 (Ronquist and Huelsenbeck, 2003) under the GTR model of substitution (Lanave *et al*., 1984; Rodriguez *et al*., 1990; Tavaré, 1986) considering invariants and a gamma-shaped distribution of the rates of distribution among sites. The chain length for our analysis was 1,000,000 generations with trees sampled every 100 generations using MCMC (Markov Chain Monte Carlo) analysis. The first 10,000 trees were discarded as burn-in for the tree topology and posterior probability. OTU consensus sequences and full-length 18S rRNA gene sequences from reference taxa were analyzed to infer OTU species' identities, and included a total of 185 taxa in our analysis. To compare microbial OTU distributions across samples, we used PRIMER-E statistical software (Clarke and Gorley, 2006) to perform cluster analyses and non-metric multidimensional scaling analyses (NMDS).

3. Results

3.1 Alkenone profiles

Hypersaline lake Gahai had the highest concentrations of alkenone lipids (**Figure 4**A), an order of magnitude higher than the other lakes. The concentrations of C_{37} alkenones for each of the six lakes samples are listed in **Table 1.** Some of our water samples had alkenone lipid concentrations that were below detection and we therefore include reference alkenone data for these samples from previous studies (**Table 1**). These supplemental alkenone data allow us to infer characteristic alkenone profiles although we interpret this data with caution.

Lake Erhai, the freshest lake study, had water concentrations dominated by $C_{37:4}$ alkenones (Liu *et al.,* 2011), with $C_{37:4}$ comprising 92% of the C37 alkenones. Our lipid analysis did not detect alkenones in the Erhai water at the time of sampling. Lake Qinghai samples 2_3m and 2_6m both had salinities of 12.62, and Qinghai sample 12_8m had a salinity of 12.61. Based on their GPS coordinates, both QH site 2 and site 12 were compared to corresponding water sampling locations from the summer sampling by Liu *et al.,* (2011). Our results yielded alkenone signatures for Qinghai samples 2_3m and 2_6m were both slightly dominated by the $C_{37:4}$ alkenone, however the Liu *et al.,* (2011) data for this location report a $C_{37:3}$ dominant profile. Qinghai sample 12_8m did not yield detectable alkenones and those reported by Liu *et al.,* had a dominant $C_{37:4}$ profile. The Gahai lake sample was the most saline of the lakes sampled, with a salinity of 76.4, and this lake had an alkenone profile dominated by $C_{37:3}$, as expected for a saline lake.

3.2 Haptophyte DNA sequences

Our analyses yielded 182 haptophyte DNA sequences that passed quality screening, approximately 30 sequences per sample. These 182 haptophyte sequences were clustered at 97% similarity to yield 5 operational taxonomic units (OTUs). These 5 OTUs (China A-China E) were aligned to reference taxa to infer taxonomic identities (**Figure 3**). Lake Erhai, the freshest lake, yielded haptophytes that clustered with *Chrysochromulina* spp. and *Prymnesium* spp. Neither of these species is a known-alkenone producing haptophyte which corroborates our lipid analysis from this lake that failed to yield a discernable alkenone signature.

Lake Qinghai sample 2_3m yielded sequences from all five haptophyte OTUs. The Qinghai sample from the same site but at 6m depth (2_6m) yielded a *Chrysochromulina* spp. as well as a close relative of *C. lamellosa* (**Figure** 3). This result confirms the suspicion by Liu *et al*., (2008) that *C. lamellosa* was responsible for alkenone production in Lake Qinghai due to the elevated $C_{37:4}$ concentrations and lack of C38 methyl ketones. The same pattern in OTU presence/absence in QH2_3m was observed in Qinghai sample 12_8m and Qinghai 2_6m and in Gahai (**Table 2**), the most saline sample. HaiyanWan, another brackish lake, only yielded sequences from *Chrysochromulina* spp., although again we did not detect alkenones from this sample.

Our NMDS analysis of each lake's haptophyte community composition revealed the grouping of samples in accordance with $C_{37:4}$ alkenone concentrations (**Figure 4**). The lakes with intermediate and high $C_{37:4}$ concentration lakes, HaiyanWan, QH12_8m, and Erhai, were most similar in OTU compositions (**Figures 5**). By comparing this with our OTU distributions (**Table 2**), we can see that this similarity is shaped by the presence of

OTU D, the close relative of *Chrysochrumlina* spp. that was the dominant OTU in all three of these lake samples.

4. Discussion

4.1 Patterns in OTU distribution

Operational Taxonomic Unit representative sequence China A matches with 100% identity to OTU 5 from Theroux *et al.,* (2010), an OTU that was identified in Greenland lake BrayaSø sediments and associated with a haptophyte that makes predominant $C_{37:4}$ alkenone lipid signatures. OTU China A was discovered only in Qinghai sample QH2_3m, and only a single DNA sequence that matched this OTU was identified. Our alkenone analyses did reveal a dominant $C_{37:4}$ alkenone signature for sample QH2_3m. However, OTU China A was not the only OTU detected in QH2_3m; in fact, this sample yielded sequences from all 5 OTUs. The only other putative alkenone producer from sample QH2_3m is a close relative to another OTU from a Greenland lake (HundeSø) and a close relative of reference taxon *C. lamellosa*. *Chrysotila lamellosa* is a known lacustrine haptophyte species that has been isolated from Lake Xiarinur in China previously (Sun *et al.,* 2007). Culture of the Lake Xiarinur *C. lamellosa* had a dominant $C_{37:3}$ signature, and lacked a methyl ketone (ibid). Our water sample had a dominant $C_{37:4}$ alkenone signature and the presence of a methyl ketone, similar to the profile observed from lake BrayaSø (D'Andrea and Huang, 2005). Although both OTU China A and *C. lamellosa*-like OTU China B were detected in the QH2_3m sample, we believe OTU China A was responsible for the $C_{37:4}$-dominant alkenone fingerprint. A previous study in Lake Qinghai (Liu *et al., 2*011) reported a $C_{38:3}$ dominant alkenone profile near the location of Qinghai site 2 for the summer surface water sample (**Table 1**). This result

suggests that during our sampling of Qinghai site 2, the China A OTU was responsible for the majority of alkenone production but at other points during the year, the OTU China B haplotype may be producing a $C_{38:3}$ dominant alkenone signature.

Deeper in the water column at site QH2, at 6m depth, we only detected putative alkenone producing haptophyte, OTU China B, the close relative to *C. lamellosa* (**Figure 3**). This sample also had a dominant $C_{37:4}$ alkenone signature, suggesting OTU China A concentrations were below detection at 6m depth. From sample QH2_8m we detected only a non-alkenone producing haptophyte related to *Chrysochromulina* spp. (OTU China D), further indicating the change in haptophyte community with depth and the greater abundance of alkenone-producing haptophytes in the surface waters.

Lake Qinghai sample QH12_8m had sequences from OTU China B, as well as OTU China D (**Table 2**). Of the three Lake Qinghai samples, QH2_3m had DNA sequences from all five China OTUs, whereas samples QH2_6m and QH12_8m had the presence of OTUs China B and China D. Samples QH2_3m and QH2_6m had dominant $C_{37:4}$ alkenone profiles, and while alkenones were below detection with our analysis, previous alkenone reports from site QH12 reported a dominant $C_{37:4}$ alkenone signature. As mentioned, previous results from site QH2 reported a $C_{38:3}$ dominant alkenone signature, suggesting the potential seasonality in the dominance of alkenone producers at this site, with alternations between in OTU A (dominant $C_{37:4}$) and OTU B (dominant $C_{38:3}$).

The most saline site in this study, lake Gahai with a salinity of 76.4 ppt, yielded haptophyte sequences closely related to *C. lamellosa*-like OTU China B (**Table 2**), suggesting the plasticity of *C. lamellosa* in a variety of environmental salinities. Our

analysis revealed the dominance in the $C_{38:3}$ alkenone, as expected because of the both the high salinity and the presence of *C. lamellosa*. Unsurprisingly, our two lakes that did not yield putative alkenone producing haptophyte sequences, Erhai and HaiyanWan, both did not yield detectable alkenone signatures. Both of these lakes have alkenones present in their surface and downcore sediments, suggesting the seasonal presence of alkenone-producing haptophytes in the photic zone of these lakes.

4.2 Implications for paleoclimatology

In this study, we examined haptophyte species assemblages as a means to explain gradients in alkenone lipid signatures across salinity gradients. As expected, our least saline lake, Erhai, had the highest percentage of $C_{37:4}$ alkenones, comprising greater than 90% of total C37 alkenones (**Figure 4**B). Our brackish sites, QH2 and QH12, both had dominant $C_{37:4}$ alkenones, although previously reported alkenone signature from QH2 at a similar salinity (Liu *et al.,* 2011; **Table 1**) reported dominant $C_{37:3}$ alkenone. HaiyanWan, also a brackish lake, and Gahai, the most saline lake, had dominant $C_{38:3}$ alkenone profiles with $C_{37:4}$ percentages of about 43% and 39%, respectively (**Figure 4**B). These results were in agreement with previous studies from the Qinghai region demonstrating elevated $C_{37:4}$ concentrations in fresher waters (Liu et al., 2008; 2011).

Using the $C_{37:4}$ percentages from our analyses, we calculated the approximate salinities as proposed by Liu *et al.,* (2008), defined as $\%C_{37:4} = 53.4\ (\pm 7.8) - 1.73\ (\pm 0.45)$ x Salinity. Using this paleosalinity proxy, our approximated salinities based on $\%C_{37:4}$ were off by up to three orders of magnitude (**Table 3**). Given the conflicting alkenone lipid signatures reported in this study and by previous studies for the same lakes (and

salinities) in the Qinghai region (Liu *et al*., 2011), it is not surprising that the paleosalinity proxy as defined by Liu *et al*., (2008) failed to accurately approximate lake salinities. This highlights the significant intra-lake alkenone signature heterogeneity.

The lakes we sampled fall within the geographic bounds prescribed by Liu *et al*., (2008) and like previous studies in the region our samples demonstrated a relationship between sample salinity and $\%C_{37:4}$ alkenones (**Figure 7**). While previous studies did not have water temperature measurements for comparison, our results confirm the $C_{37:4}$-salinity relationship was not simply a result of a temperature effect, as there was no consistent trend between temperature and salinity (**Figure S1**). Liu *et al*. (2008) recommended only applying the paleosalinity proxy within a single lake system, an effort to avoid complications arising from multiple alkenone-producing haptophytes across disparate geographies. However, we now know that Lake Qinghai alone hosts five haptophyte species, and two distinct alkenone-producing haptophyte species, OTU A that is likely responsible for elevated $C_{37:4}$ concentrations, and OTU B responsible for $C_{38:3}$ dominant alkenone signatures. Our results suggest that the observed variations in $\%C_{37:4}$ concentrations with salinity are in fact a result of a multiple-haptophyte system. It remains to be seen if other environmental factors besides salinity will preferentially select for different haptophyte populations.

5. Conclusions

The abundance of $C_{37:4}$ alkenones is a hallmark of coastal and lacustrine environments. Concentrations of $C_{37:4}$ alkenones can be affected by temperature, cell physiology, salinity, or as examined here, haptophyte taxonomy. Like previous studies in

the Qinghai region, our results showed a relationship between $\%C_{37:4}$ concentrations and salinity. Haptophyte *C. lamellosa* was isolated from nearby Lake Xiarinur (Sun *et al.,* 2007) and the lipid profiles from Lake Qinghai suggested the presence of *C. lamellosa* (Liu *et al.,* 2011). Our results confirmed the presence of *C. lamellosa* (OTU China B) in Lake Qinghai, likely responsible for dominant $C_{37:3}$ alkenone signatures. We also identified another known alkenone-producer, OTU China A, responsible for predominant $C_{37:4}$ alkenone signatures (D'Andrea *et al.*, 2008; Theroux *et al.,* 2010). Because of the individual alkenone-lipid signatures specific to these two haptophyte species, we believe changing environmental conditions preferentially select for different haptophyte species resulting in the observed trends in $C_{37:4}$ variations with salinity.

The dominant Isochrysidales (alkenone-producing) OTU across the lakes in this study, OTU B *Chrysotila lamellosa*, was present in the brackish samples from Qinghai and the most saline samples from Gahai. We did not detect putative alkenone-producing haptophytes from our samples in Erhai and HaiyanWan, likely reflecting the seasonality of alkenone production in the water column of these lakes. Future studies that capture the alkenone-producing haptophytes from Erhai, HaiyanWan, and the other lakes in the Qinghai region will bolster the salinity-$C_{37:4}$ relationship and help to resolve trends in $C_{37:4}$ concentrations with haptophyte species. It will be important to verify that salinity alone, and not other environmental perturbations, are shaping the haptophyte community. Only then will the paleosalinity proxy prove valuable for application across a diversity of lake environments.

Acknowledgements

This work was supported by National Science Foundation awards to Y. Huang (EAR-0902805, EAR-1122749) and L. Amaral-Zettler (EAR-1124192), a Brown SEED fund to Y. Huang and L. Amaral-Zettler, and an American Association of University Women dissertation fellowship to S. Theroux.

References

Bendle J., Rosell-Mele´ A. and Ziveri P. (2005) Variability of unusual distributions of alkenones in the surface waters of the Nordic seas. Paleoceanography 20, PA2001. doi:10.1029/ 2004PA001025.

Brassell SC, Eglinton G, Marlowe IT, Pflaumann U, Sarnthein M. (1986). Molecular stratigraphy - a new tool for climatic assessment. *Nature* **320**:129-133.

Chu GQ, Sun Q, Li SQ, Zheng MP, Jia XX, Lu CF, Liu JQ, Liu TS. (2005). Long-chain alkenone distributions and temperature dependence in lacustrine surface sediments from China. *Geochimica Et Cosmochimica Acta* **69**:4985-5003.

Clarke KR, Gorley RN. (2006). PRIMER v6: User Manual/Tutorial.

Colman SM, Yu S-Y, An ZH, Shen J, Henderson ACG. Late Cenozoic climate changes in China's western interior: a review of research on Lake Qinghai and comparison with other records. *Quaternary Science Reviews* **26**:2281-2300.

Coolen MJL, Muyzer G, Rijpstra WIC, Schouten S, Volkman JK, Sinninghe Damsté J (2004). Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake. *Earth and Planetary Science Letters* **223**:225-239.

Cranwell PA. (1985). Long-Chain Unsaturated-Ketones in Recent Lacustrine Sediments. *Geochimica et Cosmochimica Acta* **49**:1545-1551.

D'Andrea WJ. (2008). Development and application of lacustrine alkenone paleothermometry in southwestern Greenland. *Geological sciences* **PhD**:197.

D'Andrea WJ, Huang Y. (2005). Long chain alkenones in Greenland lake sediments: Low δ13C values and exceptional abundance. *Organic Geochemistry* **36**:1234-1241.

D'Andrea WJ, Lage M, Martiny JBH, Laatsch AD, Amaral-Zettler LA, Sogin ML, Huang YS. (2006). Alkenone producers inferred from well-preserved 18S rDNA in Greenland lake sediments. *Journal of Geophysical Research-Biogeosciences* **111**:doi:10.1029/2005JG000121.

de Vargas C, MP Aubry, I Probert, J Young (2007) Origin and evolution of coccolithophores: From coastal hunters to oceanic farmers In *Evolution of Aquatic Photoautotrophs*. P Falkowski, AH Knoll, eds. Elsevier Academic, New York, p. 251.

Edgar RC. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**:2460-2461.

Ewing B, Green P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Research* **8**:186-194.

Gordon D, Abajian C, Green P. (1998). Consed: A graphical tool for sequence finishing. *Genome Research* **8**:195-202.

Henderson ACG, Holmes JA, Zhang JW, Leng MJ, Carvalho LR. A carbon- and oxygen-isotope record of recent environmental change from Qinghai Lake, NE TIbetan Plateau. *Chinese Science Bulletin* **48**:1463-1468.

Huber T, Faulkner G, Hugenholtz P. (2004). Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics Bioinformatics* **20**:2317-2319.

Lanave C, Preparata G, Saccone C, Serio G. (1984). A new method for calculating evolutionary substitution rates. *Journal of Molecular Evolution* **20**:86-93.

Lasek-Nesselquist E, Bogomolni AL, Gast RJ, Welch DM, Ellis JC, Sogin ML, Moore MJ. (2008). Molecular characterization of *Giardia intestinalis* haplotypes in marine animals: variation and zoonotic potential. *Diseases of Aquatic Organisms* **81**:39-51.

Li JG, Philp RP, Pu F, Allen J. (1996). Long-chain alkenones in Qinghai Lake sediments. *Geochimica Et Cosmochimica Acta* **60**:235-241.

Liu WG, Liu ZH, Fu MY, An ZH. (2008). Distribution of the C37 tetra-unsaturated alkenone in Lake Qinghai, China: A potential lake salinity indicator. *Geochimica Et Cosmochimica Acta* **72**:988-997.

Liu W, Liu Z, Wang H, He Y, Wang Z, Xu L. (2011). Salinity control on long-chain alkenone distributions in lake surface waters and sediments of the northern Qinghai-Tibetan Plateau, China. *Geochimica et Cosmochimica Acta* **75**:1693-1703.

Liu Z, Henderson ACG, Huang Y. (2006). Alkenone-based reconstruction of late-Holocene surface temperature and salinity changes in Lake Qinghai, China. *Geophys. Res. Lett.* **33**:L09707.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar Buchner A, Lai T, Steppi S, Jobb G, Forster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Jost R, Konig A, Liss T, Lussmann R, May M, Nonhoff B, Reichel B, Strehlow R, Stamatakis A, Stuckmann N, Vilbig A, Lenke M, Ludwig T, Bode A, Schleifer KH. (2004). ARB: a software environment for sequence data. *Nucleic Acids Research* **32**:1363-1371.

Marlowe IT, Green JC, Neal AC, Brassell SC, Eglinton G, Course PA. (1984). Long-Chain (**n**-C37-C39) alkenones in the prymnesiophyceae. Distribution of alkenones and other lipids and their taxonomic significance. *British Phycological Journal*

19:203-216.

Mercer JL, Zhao MX, Colman SM. (2005). Seasonal variations of alkenones and U-37(K) in the Chesapeake Bay water column. *Estuarine Coastal and Shelf Science* **63**:675-682.

Müller PJ, Kirst G, Ruhland G, von Storch I, Rosell-Mele A. (1998). Calibration of the alkenone paleotemperature index $U^{k'}_{37}$ based on core-tops from the eastern South Atlantic and the global ocean (60°N-60°S). *Geochimica Et Cosmochimica Acta* **62**:1757-1772.

Prahl FG, Muehlhausen LA, Zahnle DL. (1988). Further evaluation of long-chain alkenones as indicators of paleoceanographic conditions. *Geochimica Et Cosmochimica Acta* **52**:2303-2310.

Prahl FG, Wakeham SG. (1987). Calibration of unsaturation patterns in long-chain ketone compositions for paleotemperature assessment. *Nature* **330**:367-369.

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig WG, Peplies J, Glockner FO. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research* **35**:7188-7196.

Rodriguez R, Oliver JL, Marin A, Medina JR. (1990). The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology* **142**:485-501.

Ronquist F, Huelsenbeck JP. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics Bioinformatics* **19**:1572-1574.

Rosell-Melé A. (1998). Interhemispheric appraisal of the value of alkenone indices as temperature and salinity proxies in high-latitude locations. *Paleoceanography* **13**:694-703.

Rosell-Melé A, Jansen E, Weinelt M. (2002). Appraisal of a molecular approach to infer variations in surface ocean freshwater inputs into the North Atlantic during the last glacial. *Global and Planetary Change* **34**:143-152.

Schulz H-M, Schöner A, Emeis K-C. (2000). Long-chain alkenone patterns in the Baltic sea--an ocean-freshwater transition. *Geochimica Et Cosmochimica Acta* **64**:469-477.

Seki O., Kawamura K., Sakamoto T., Ikehara M., Nakatsuka T. and Wakatsuchi M. (2005) Decreased surface salinity in the Sea of Okhotsk during the last glacial period estimated from alkenones. *Geophys. Res. Lett.* **32**, L08710.

Shi YF, Lu MX, Li WZ. Physical geography with emphasis on geomorphology around Lake Qinghai. *Acta Geogr. Sin.* **24**:33-48.

Sicre MA, Marty JC, Saliot A. (1990). n-Alkanes, fatty acid esters, and fatty acid salts in size fractionated aerosols collected over the Mediterranean Sea. *J. Geophys. Res.* **95**:3649-3657.

Sicre M. A., Bard E., Ezat U. and Rostek F. (2002) Alkenone distributions in the North Atlantic and Nordic sea surface waters. *Geochem. Geophys. Geosyst.* 3. doi:10.1029/2001GC000159.

Simon N, Campbell L, Örnolfsdoˈttir E, Groben R, Guillou L, Lange M, Medlin LK. (2000). Oligonucleotide probes for the identification of three algal groups by dot blot and fluorescent whole-cell hybridization. *The Journal of Eukaryotic Microbiology* **47**:76-84.

Sun Q, Chu G, Li S, Lü C, Zheng M. (2004). Long-chain alkenones in sulfate lakes and its paleoclimatic implications. *Chinese Science Bulletin* **49**:2082-2086.

Sun Q, Chu GQ, Liu GX, Li S, Wang XH. (2007). Calibration of alkenone unsaturation index with growth temperature for a lacustrine species, *Chrysotila lamellosa* (Haptophyceae). *Organic Geochemistry* **38**:1226-1234.

Tavaré S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences* **17**:57-86.

Theroux SM, D'Andrea WJ, Toney JL, Amaral-Zettler LA, Huang Y. (2010). Phylogenetic diversity and evolutionary relatedness of alkenone producing haptophyte algae in lakes: Implications for continental paleotemperature reconstructions. *Earth and Planetary Science Letters* **300**:311-320.

Thiel V, Jenisch A, Landmann G, Reimer A, Michaelis W. (1997). Unusual distributions of long-chain alkenones and tetrahymanol from the highly alkaline Lake Van, Turkey. *Geochimica Et Cosmochimica Acta* **61**:2053-2064.

Toney JL, Huang Y, Fritz SC, Baker PA, Grimm E, Nyren P. (2010). Climatic and environmental controls on the occurrence and distributions of long chain alkenones in lakes in the interior United States. *Geochimica Et Cosmochimica Acta*

Volkman JK, Burton HR, Everitt DA, Allen DI. (1988). Pigment and lipid compositions of algal and bacterial communities in Ace Lake, Vestfold Hills, Antarctica. *Hydrobiologia* **165**:41-57.

Wang RL, Zheng MP. (1998). Occurrence and environmental significance of long-chain alkenones in Tibetan Zabuye salt lake, SW China. *International Journal of Salt Lake Research* **6**:281-302.

Table 1. Environmental characteristics of lakes in this study.

Table 1. Descriptions of water samples for the lakes in this study. Alkenone data from previous studies is provided when our analysis did not yeild dete

	Salinity (ppt)	Lat, Long	Elevation	Depth of sample (m)	Max depth @ site (m)	Temp (°C)	Cond. mS/cm	pH	DO %	C37:4	C37:3	C37:2	C37total (ng/L)	%C37 :4	Uk37	Uk'37	Reference
Erhai	0.97	36°32', 100°43'		0		13	1.89	9.1	76.2	nd	nd	nd	nd	nd	nd	nd	This study
Erhai	1.6									179.4	10.7	4.0	194.0	92.5	0.3	-0.9	Liu *et al.*, 2011 GCA
Qinghai 2_3m	12.62	36°39', 100°36'	3190	3.02	24.5	8.8	21.11	9.1	73.5	33.9	23.2	8.6	65.8	0.5	0.3	-0.4	This study
Qinghai 2_6m	12.62	36°39', 100°36'	3190	5.95	24.5	8.65	21.12	9.1	74.1	40.8	34.2	8.5	83.6	0.5	0.2	-0.4	This study
QH2*	15.8									50.7	107.9	12.0	170.5	29.7	-0.2	0.100	Liu *et al.*, 2011 GCA
Qinghai 12_8m	12.61	36°55', 99°52'	3195	8	19	9.8	21.06	9.1	69.1	nd	nd	nd	nd	nd	nd	nd	This study
QH12+	15.1									65.3	48.1	6.8	120.1	54.4	-0.5	0.1	Liu *et al.*, 2011 GCA
HaiyanWan	18.74	36°48', 100°45'	3183	0.15		7.77	30.39	9.1	70.8	nd	nd	nd	nd	nd	nd	nd	This study
HaiyanWan	25.9									13.5	15.9	2.1	31.6	42.9	0.1	-0.4	Liu *et al.*, 2011 GCA
GaHai	76.4	37°08', 97°33'	2848	1.9		15.2	104.7	8.3	79.2	863.1	1225.2	101.2	2189.4	0.4	0.1	-0.4	This study

*Qinghai 8_2 from Liu *et al.*, 2011
+Qinghai 8_7 from Liu *et al.*, 2011

Table 2. OTU distribution among the six samples by percentage. Each of the 5 OTUs are listed with their closest relative reference taxon (based on BLAST distance) and the number of sequences that was detected in each sample for the specific OTU.

OTU	OTU closest relative	Erhai	Qinghai 2_3m	Qinghai 2_6m	Qinghai 12_8m	Haiyan Wan	GaHai
China A	OTU 5/Greenland phylotype		2.9				
China B	Chrysotila lamellosa		2.9	28.1	3.2		82.8
China C	Prymnesium spp.	10	40.0				
China D	Chrysochromulina spp.	90	20.0	71.9	96.8	100.0	17.2
China E	Chrysochromulina spp.		34.3				
Total seqence number		30	35	32	31	25	29

Table 3. Salinity reconstruction using Liu *et al.* (2008) paleosalinity proxy based on $\%C_{37:4}$. Measured salinity for each lake is provided. Salinity reconstruction yielded values that were up to three orders of magnitude incorrect. Asterisk denotes when the data is from Liu *et al.,* 2011.

	%C37:4	Salinity Approximation	Measured Salinity
Erhai	179.40	3.39	1.60
Qinghai 2_3m	0.06	0.03	12.62
Qinghai 2_6m	0.08	0.03	12.62
Qinghai 2*	50.70	0.98	15.80
Qinghai 12*	65.30	1.26	15.10
HaiyanWan	13.50	0.29	25.90
GaHai	1.63	0.06	76.40

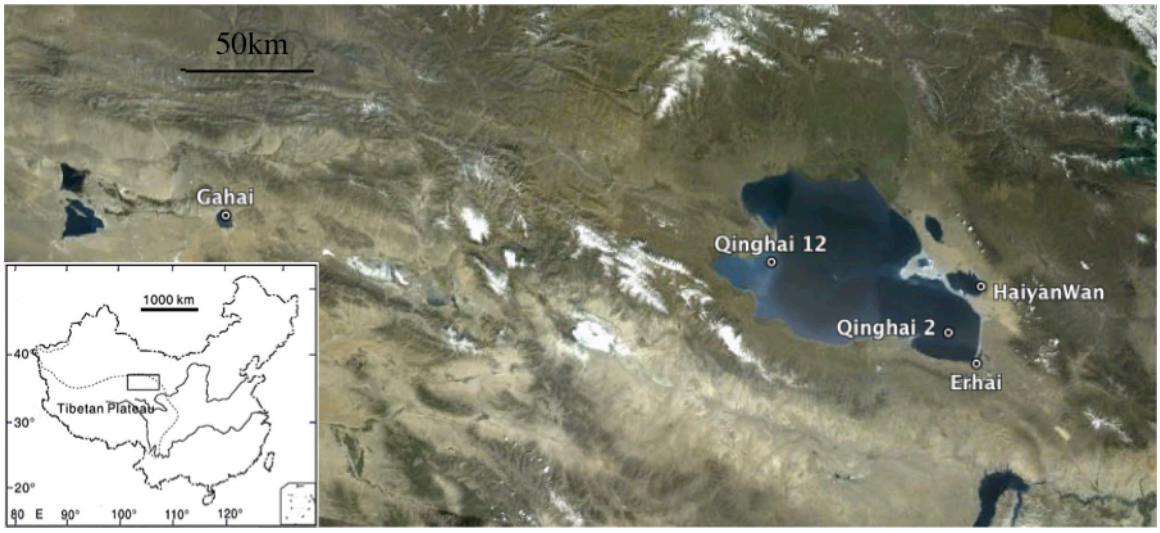

Figure 1. Map of lake locations in the Qinghai basin. Inset of Tibetan plateau with rectangle indicating the sampling locations (modified from Liu *et al., 2*011).

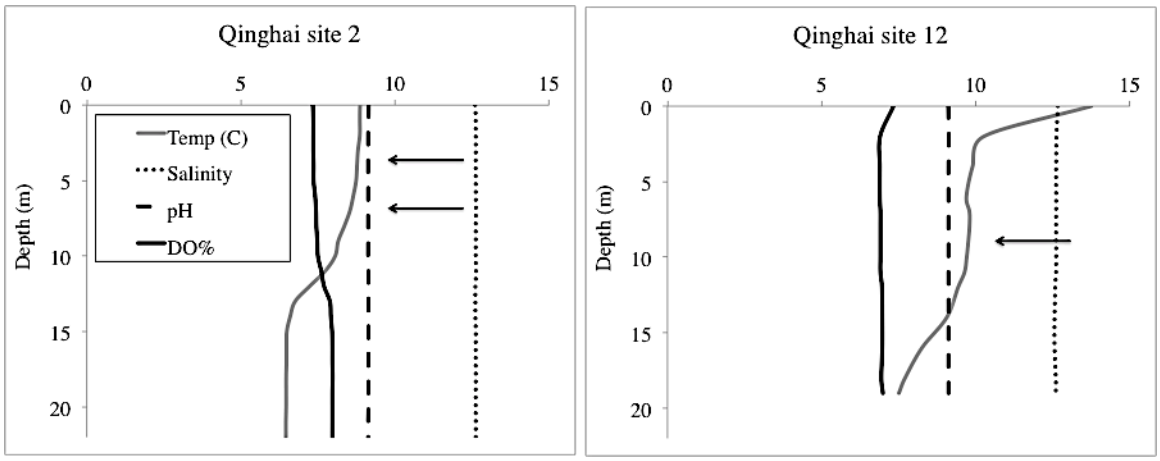

Figure 2. Water column profiles of sampling sites Qinghai 2 and Qinghai 12. Lakes were sampled on June 29, 2010. Arrows indicate location of samples QH2_3m, QH2_6m and QH12_8m.

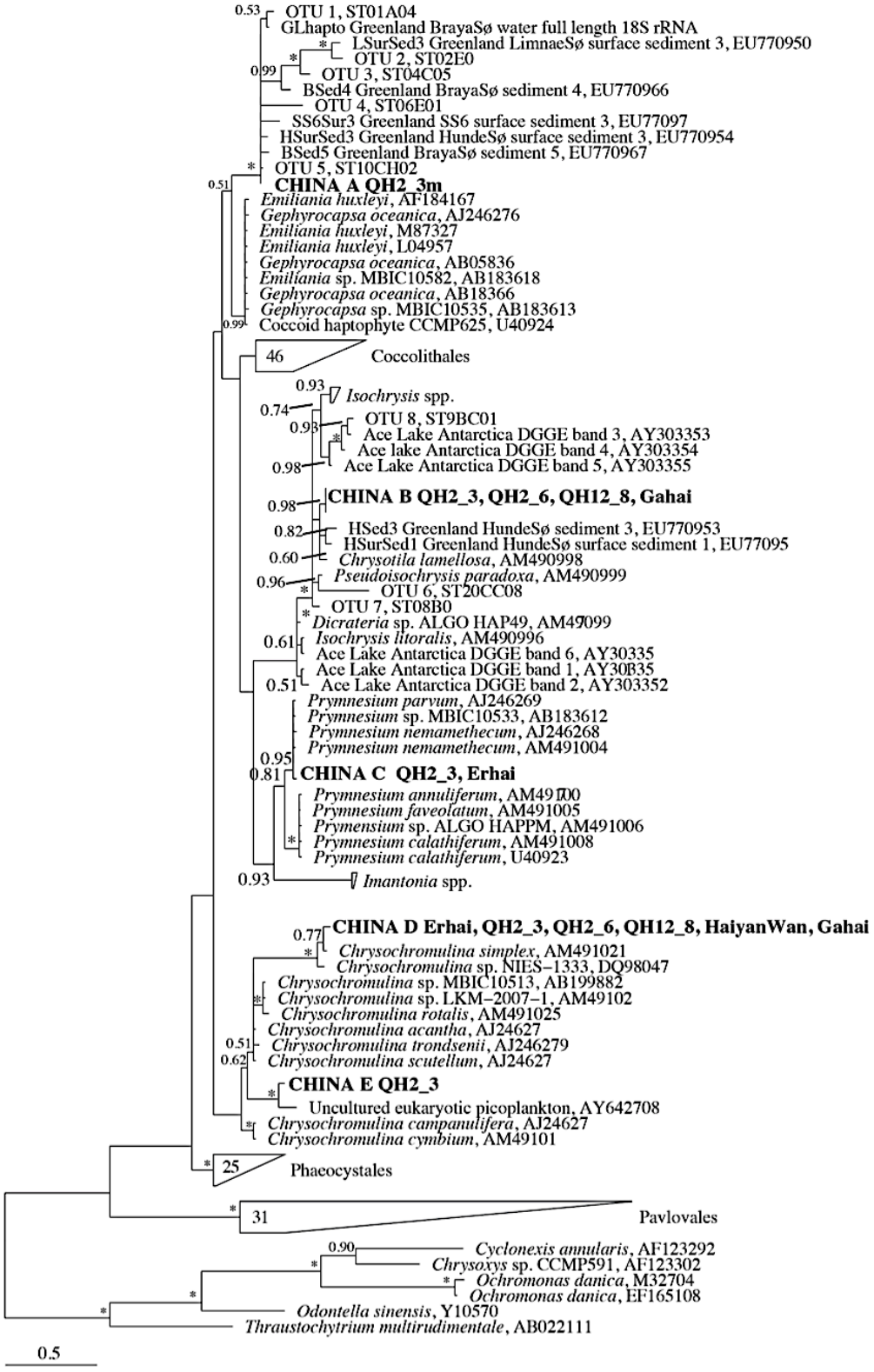
OTU 1, ST01A04
GLhapto Greenland BrayaSø water full length 18S rRNA
LSurSed3 Greenland LimnaeSø surface sediment 3, EU770950
OTU 2, ST02E0
OTU 3, ST04C05
BSed4 Greenland BrayaSø sediment 4, EU770966
OTU 4, ST06E01
SS6Sur3 Greenland SS6 surface sediment 3, EU77097
HSurSed3 Greenland HundeSø surface sediment 3, EU770954
BSed5 Greenland BrayaSø sediment 5, EU770967
OTU 5, ST10CH02
CHINA A QH2_3m
Emiliania huxleyi, AF184167
Gephyrocapsa oceanica, AJ246276
Emiliania huxleyi, M87327
Emiliania huxleyi, L04957
Gephyrocapsa oceanica, AB05836
Emiliania sp. MBIC10582, AB183618
Gephyrocapsa oceanica, AB18366
Gephyrocapsa sp. MBIC10535, AB183613
Coccoid haptophyte CCMP625, U40924
46
Coccolithales
Isochrysis spp.
OTU 8, ST9BC01
Ace Lake Antarctica DGGE band 3, AY303353
Ace lake Antarctica DGGE band 4, AY303354
Ace Lake Antarctica DGGE band 5, AY303355
CHINA B QH2_3, QH2_6, QH12_8, Gahai
HSed3 Greenland HundeSø sediment 3, EU770953
HSurSed1 Greenland HundeSø surface sediment 1, EU77095
Chrysotila lamellosa, AM490998
Pseudoisochrysis paradoxa, AM490999
OTU 6, ST20CC08
OTU 7, ST08B0
Dicrateria sp. ALGO HAP49, AM490
Isochrysis litoralis, AM490996
Ace Lake Antarctica DGGE band 6, AY30335
Ace Lake Antarctica DGGE band 1, AY30B35
Ace Lake Antarctica DGGE band 2, AY303352
Prymnesium parvum, AJ246269
Prymnesium sp. MBIC10533, AB183612
Prymnesium nemamethecum, AJ246268
Prymnesium nemamethecum, AM491004
CHINA C QH2_3, Erhai
Prymnesium annuliferum, AM491
Prymnesium faveolatum, AM491005
Prymensium sp. ALGO HAPPM, AM491006
Prymnesium calathiferum, AM491008
Prymnesium calathiferum, U40923
Imantonia spp.
CHINA D Erhai, QH2_3, QH2_6, QH12_8, HaiyanWan, Gahai
Chrysochromulina simplex, AM491021
Chrysochromulina sp. NIES−1333, DQ98047
Chrysochromulina sp. MBIC10513, AB199882
Chrysochromulina sp. LKM−2007−1, AM49102
Chrysochromulina rotalis, AM491025
Chrysochromulina acantha, AJ24627
Chrysochromulina trondsenii, AJ246279
Chrysochromulina scutellum, AJ24627
CHINA E QH2_3
Uncultured eukaryotic picoplankton, AY642708
Chrysochromulina campanulifera, AJ24627
Chrysochromulina cymbium, AM49101
25
Phaeocystales
31
Pavlovales
Cyclonexis annularis, AF123292
Chrysoxys sp. CCMP591, AF123302
Ochromonas danica, M32704
Ochromonas danica, EF165108
Odontella sinensis, Y10570
Thraustochytrium multirudimentale, AB022111
0.5

Figure 3. A consensus Bayesian phylogenetic tree depicting 18S rRNA gene-inferred relationships among haptophyte algae. An asterisk (*) indicates posterior probability values of 1.00; all other values as shown. Bold designates sequences from this study. Operational Taxonomic Units (OTUs) were defined at a 97% similarity level. Order classification after de Vargas *et al.* (2007), with number of sequences per order as indicated. The evolutionary distance for the number of changes per site is represented by the scale bar. GenBank accession numbers follow species and sequence names.

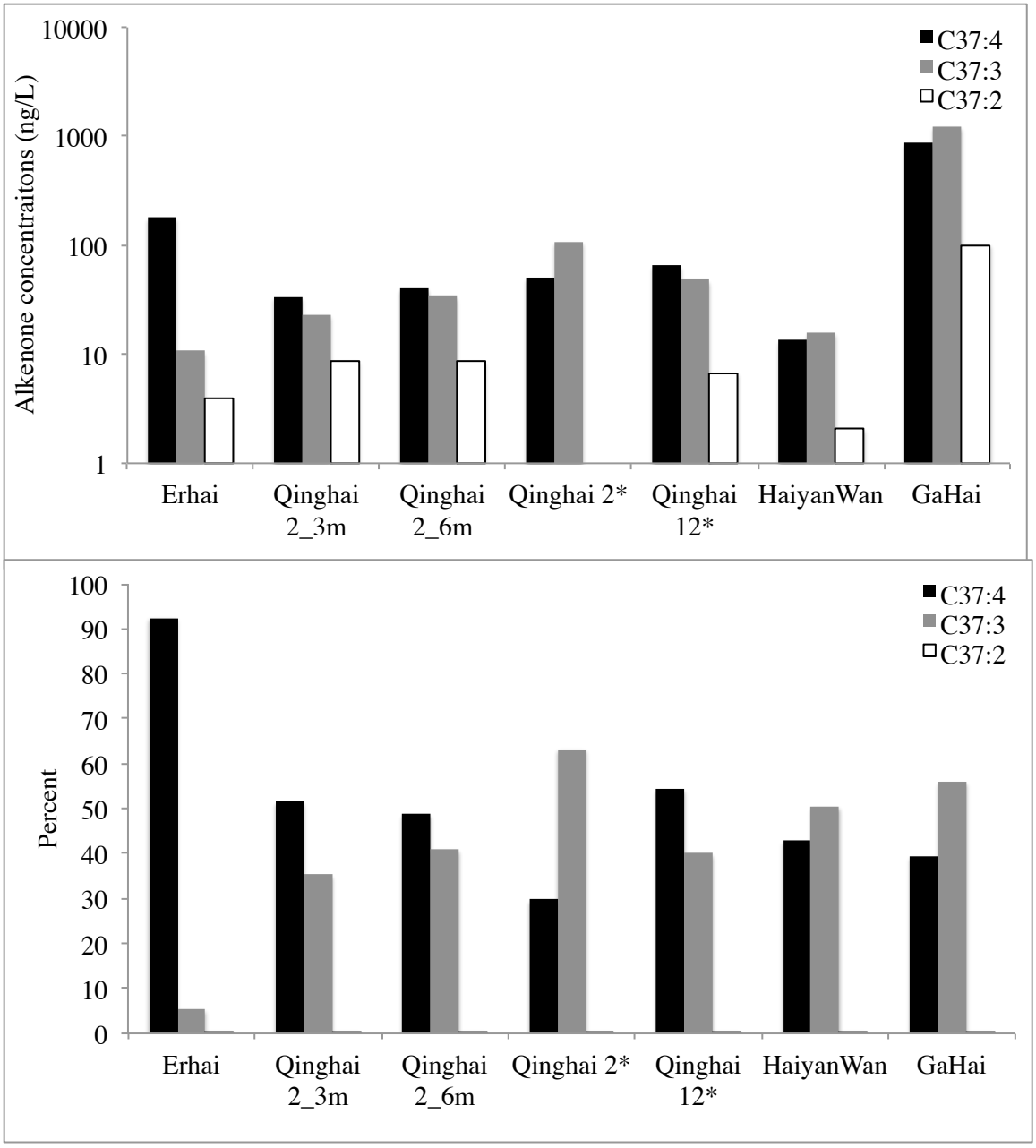

Figure 4. Concentrations and percent abundances of alkenone homologs for each lake in this study. A. Alkenone concentrations plotted on a log scale. B. Relative alkenone concentrations. Asterisk denotes when the data is from Liu *et al.,* 2011.

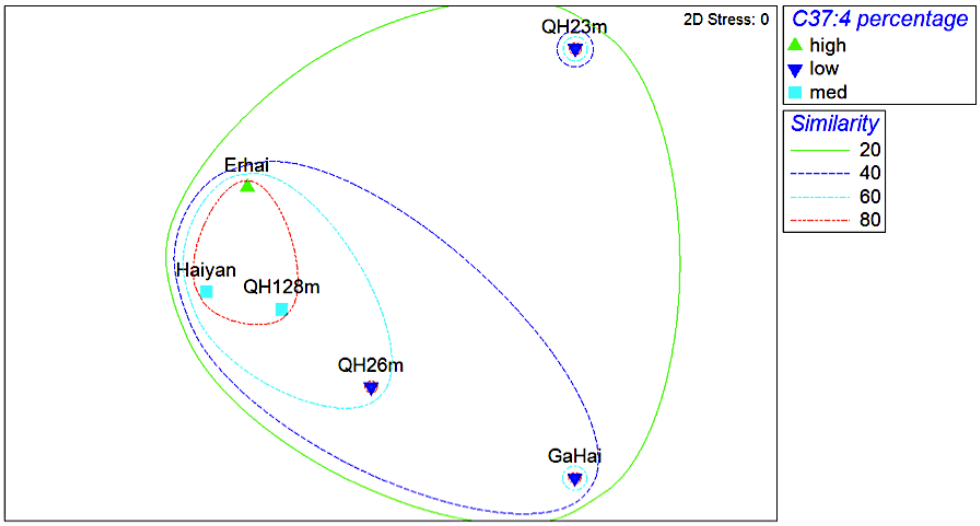

Figure 5. Non-metric multidimensional scaling plot of China lake samples and their Bray-Curtis similarity in haptophyte OTU distributions. OTU distributions were normalized by total and square root transformed. $C_{37:4}$ percent concentrations are defined as high (> 40%), medium (20-50%) and low (< 20%). Boundaries of percent similarity as noted.

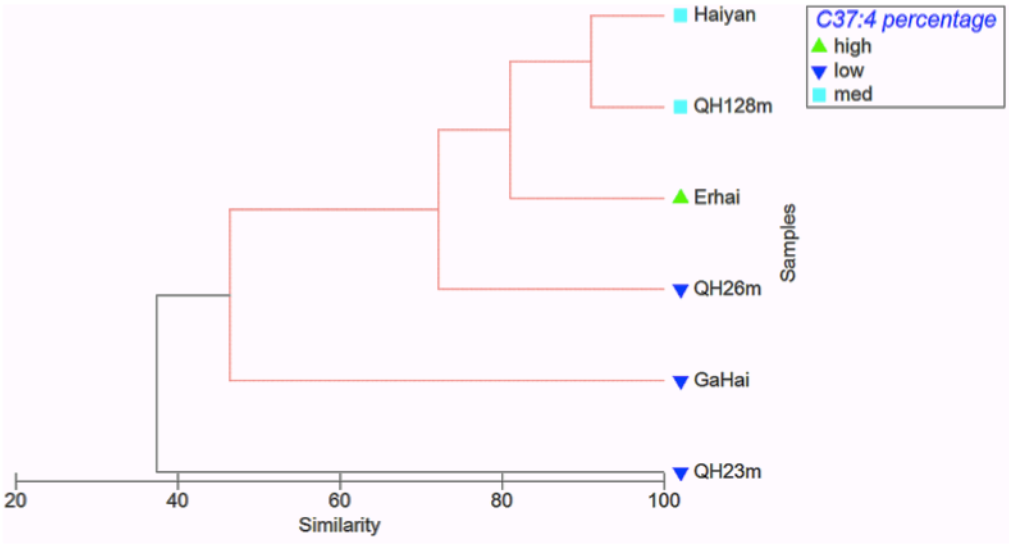

Figure 6. Hierarchical clustering dendrogram of haptophyte OTU distributions among lakes in this study based on Bray-Curtis similarity. OTU distributions were normalized by total and square root transformed. $C_{37:4}$ percent concentrations are defined as high (> 40%), medium (20-50%) and low (< 20%).

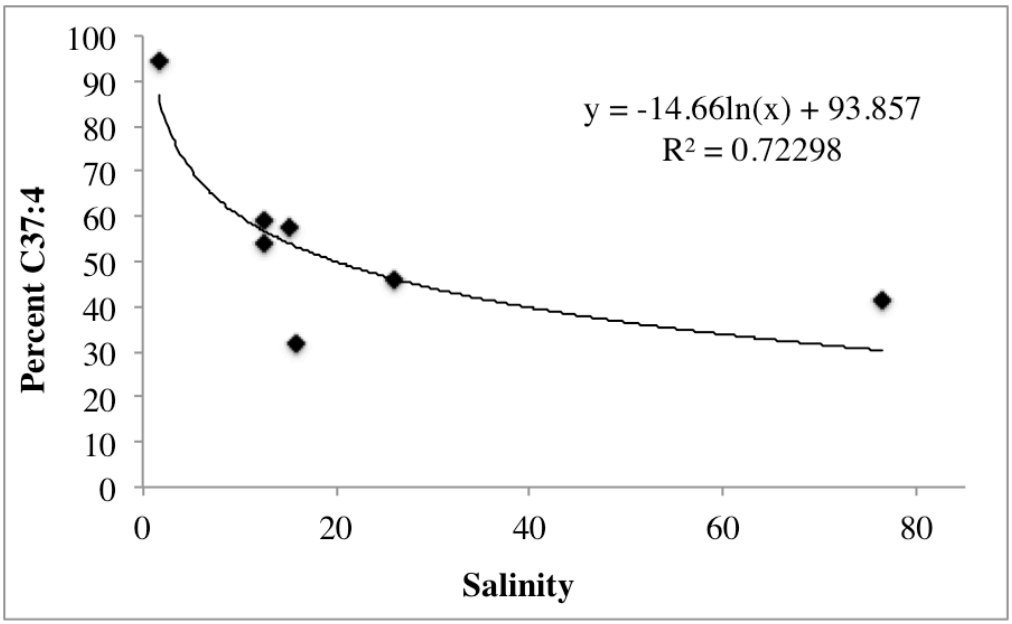

Figure 7. Relationship between lake water sample salinities and percent $C_{37:4}$ alkenones in lakes from this study.

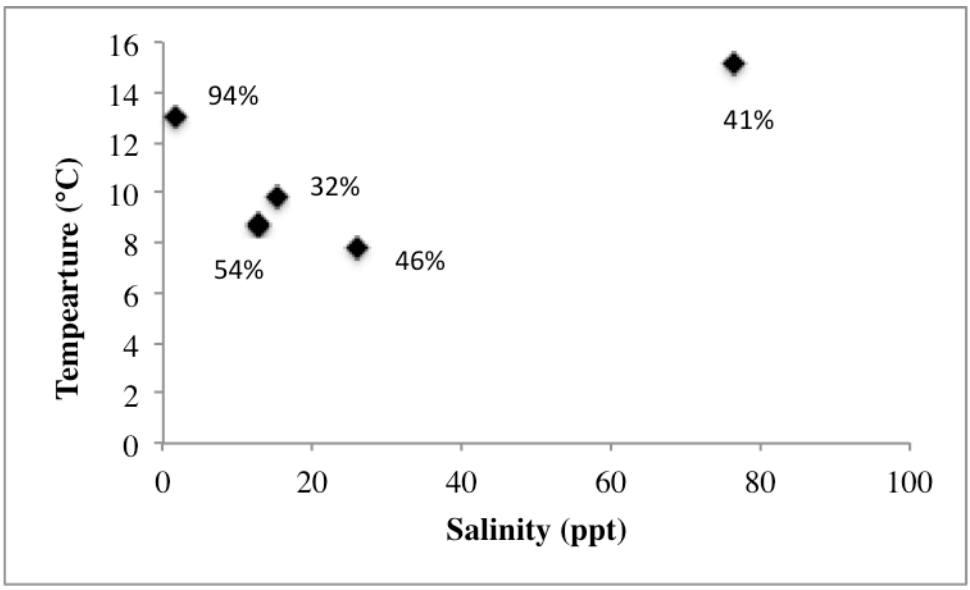

Figure S1. Relationship between lake salinity and water temperature for lakes in this study. $\%C_{37:4}$ concentrations as shown.

CHAPTER 6

A PICOPLANKTON HAPTOPHYTE IS RESPONSIBLE FOR DOMINANT $C_{37:4}$ ALKENONE PRODUCTION

SUSANNA THEROUX[1,2]

JAIME TONEY[1,‡]

ROBERT ANDERSEN[3]

YONGSONG HUANG[1]

LINDA AMARAL-ZETTLER[1,2,4]

[1] BROWN UNIVERSITY, *Department of Geological Sciences*

[2] MARINE BIOLOGICAL LABORATORY, *Josephine Bay Paul Center*

[3] UNIVERSITY OF WASHINGTON, *Friday Harbor Laboratories*

[4]BROWN UNIVERSITY, *Department of Ecology and Evolutionary Biology*

[‡] Present address: UNIVERSITY OF GLASGOW, *School of Geographical and Earth Sciences*

Prepared for:

Geochimica et Cosmochimica Acta

Abstract

Lacustrine alkenone records hold potential to be valuable sedimentary archives of continental paleotemperature. Uncertainties in alkenone biosynthesis by novel lake-dwelling haptophytes hamper the use of this proxy. All alkenone compounds are produced by haptophyte algae in the order Isochrysidales, a few species of which have been grown in culture to gauge alkenone production. The tetra-unsaturated alkenone is a signature compound found predominantly in lacustrine and coastal waters. Of all the cultured species, none has constitutively produced a dominant tetra-unsaturated alkenone ($C_{37:4}$) lipid signature. We recently maintained a dominant tetra-unsaturated alkenone producing haptophyte in enrichment culture. In an effort to isolate this organism, we size-fractionated the enrichment culture and grew these fractions at various temperatures to determine the alkenone-unsaturation index (U^k_{37}) calibration. The < 3μm subculture was the only alkenone-containing fraction, indicating this novel $C_{37:4}$-predominant haptophyte is a picoplankter. Our study is the first to successfully maintain the predominant $C_{37:4}$-producing haptophyte in isolation and to determine its alkenone unsaturation index (U^k_{37}) calibration. These results reveal the diverse ecology of alkenone-producing haptophytes and provide an explanation for the widespread occurrence of $C_{37:4}$-producing haptophytes in oligotrophic waters.

1. Introduction

The alkenone unsaturation index (U^{k}_{37} and $U^{k'}_{37}$) is a valuable tool in marine paleothermometry. Marine alkenones are produced predominantly by ubiquitous haptophyte species *Emiliania huxleyi* and its close relative *Gephyrocapsa oceanica*, allowing for a nearly global calibration of the $U^{k'}_{37}$ index based upon culture (Brassell *et al*., 1986; Prahl and Wakeham, 1987) and coretop calibrations (Muller *et al*., 1998). In the North Atlantic, and many coastal and lacustrine settings, the abundance of a tetraunsaturated alkenone ($C_{37:4}$) suggests the presence of a different alkenone-producing haptophyte species (Cranwell, 1985; Zink *et al*., 2001; Chu *et al*., 2005; D'Andrea and Huang, 2005; Pearson *et al*., 2008; Theroux *et al*., 2010; Toney *et al*., 2010). Genetic studies have confirmed the presence of novel haptophyte species in lacustrine settings (Coolen *et al*. 2004; D'Andrea *et al*., 2005; Theroux *et al*., 2010). Only coastal/lacustrine species *Isochrysis galbana* (Versteegh *et al*. 2001) and *Chrysotila lamellosa* (Rontani 2004; Sun *et al*., 2007) have been grown in culture to gauge their alkenone lipid production at different temperatures, a key step in empirically-deriving robust U^{k}_{37} calibrations. *Isochrysis galbana* and *C. lamellosa* both produce $C_{37:4}$ alkenones although they are not the dominant alkenone isomer.

Lake George, ND hosts an annual haptophyte bloom event in late spring (Toney *et al*., 2010) that is responsible for the majority of annual alkenone production. The alkenones present in the sediments of Lake George can reach concentrations of 313 μg/g TOC (Toney *et al*., 2010) and are characterized by the dominance of the $C_{37:4}$ alkenone over the tri and di-unsaturated alkenones ($C_{37:3}$ and $C_{37:2}$, respectively). Using sediment enrichment culturing, Toney *et al*. (2012) were able to successfully maintain the

haptophyte responsible for $C_{37:4}$ production (Hap-A) in enrichment culture, the first time this organism has ever been grown in a lab. The Hap-A species is a close relative of an uncultured haptophyte from Ace Lake, Antarctica that produces a $C_{37:4}$ dominant signature (Coolen *et al*., 2004, Theroux *et al*. 2010; Toney *et al*. 2012). Lake George also hosts a haptophyte species (Hap-B) closely related to *C. lamellosa* and *Pseudoisochrysis paradoxa* that does not produce the dominant $C_{37:4}$ alkenone (Theroux *et al*. 2010; Toney *et al*., 2012).

Only four haptophyte species have been successfully cultured to calibrate their U^k_{37} index, two marine species (*E. huxleyi and G. oceanica*) and two coastal/lacustrine species (*C. lamellosa, I. galbana*). Beyond these four species, the ecology of alkenone-producing haptophytes is unknown. Studies are mounting that suggest the novelty and unexplored diversity of lake-dwelling, alkenone-producing haptophytes (Coolen *et al*., 2004; D'Andrea *et al*., 2006; Theroux *et al*., 2010). Recent work targeting marine haptophytes suggests there is an unknown haptophyte contribution to marine primary productivity, largely present in the picoplankton fraction (Liu *et al*., 2009). Formally defined as cells between 0.2-2μm in size, an extended definition of picoeukaryotes as any cell that passes through a 3μm filter is often applied in field studies (Moon-van der Staay *et al*., 2001) and is the definition we use here.

In an effort to further isolate and characterize lipid production by Hap-A, we performed an enrichment culturing experiment after Toney *et al*. (2012) and used this source material to further isolate the $C_{37:4}$-producing haptophyte. We performed a size fractionation experiment to isolate Hap-A and performed a series of manipulation experiments to gauge its alkenone response to temperature variation. This study is the

first to successfully visualize the novel species responsible for $C_{37:4}$ production and to identify it as a picoplankter. Our contribution highlights the importance of molecular work to characterize haptophyte species that are previously unidentifiable given their small size and anonymous cell structure.

2. Methods

2.1 *Enrichment culture and size fractionation subcultures*

We initiated a sediment enrichment culture after Toney *et al*. (2012). We inoculated a glass beaker with 50 g Lake George surface sediment overlain with 2 liters of 0.2 μm filtered Lake George water amended with f/2 nutrients (Guillard, 1975). We placed this sediment enrichment culture in a 4°C incubator with a 24:0 light:dark light regime using a full-spectrum light. After four weeks, we subsampled the enrichment culture supernatant and separated this subculture into a < 3 μm, 3-5 μm, and a 5-10 μm fractions using Millipore Isopore cellulose nitrate filters (Billerica, Massachusetts). We filtered each size fraction onto a 44 mm GFF Whatman filter for alkenone analysis and kept the filter at -20°C until extraction. We detected alkenones in only the < 3μm micron fraction, and used this fraction for a temperature manipulation experiment.

We grew < 3 μm subcultures in glass vials in 0.2 μm filtered Lake George water amended with f/2 nutrients (Guillard, 1975). We grew the cultures in volumes of 50 ml at 5, 10, 15, 21 and 24°C under 24:0 light:dark regime using a full-spectrum light. All cultures were inoculated with approximately 8000 cells/ml. The cultures were sampled biweekly for cell counts and cell concentrations were determined using a haemocytometer to ensure the cultures were maintained in the logarithmic growth phase.

At the end of two months, we harvested the cultures for alkenone analysis by filtering onto a 44mm GFF Whatman filter and freezing the filter at -20°C until extraction. We also filtered a subsection of the < 3μm fraction onto a Sterivex (Millipore, Billerica, MA) filter and flooded the filter with Puregene cell lysis buffer (Qiagen, Hilden, Germany) to preserve for later DNA extraction and analysis. This Sterivex filter was also kept at -20°C until analysis.

2.2 Alkenone extraction and analysis

We extracted the alkenone filters using three 20-minute bursts of sonication in dichloromethane (DCM), and ran the total lipid extracts on an Agilent 6890plus Gas Chromatograph Flame Ionization Detector (GC-FID) for detection and quantification of alkenones using an internal C36 n-alkane standard and an external alkenone standard of known U^{k}_{37}-temperature value to ensure analytical precision (< 0.1°C). We used a Varian VF200 60m GC column (60 m × 250 μm × 0.10 μm) with the following temperature program: an initial temperature of 100°C (hold 1-min), ramp 20°C/min to 200°C (hold 1-min), ramp 4°C/min to 320°C (hold 0 min), hold at 320°C for 5-min. Analytical accuracy, tracked via the injection of laboratory $U^{k'}_{37}$ standard of known temperature, was ±0.042 $U^{k'}_{37}$ units (±0.1°C).

2.3 FISH experiment

We followed the protocol for Fluorescent *in situ* Hybridization (FISH) after Toney *et al*. (2012). We concentrated five milliliters of the 5°C < 3μm culture by centrifugation and resuspended the cell pellet in ~80% ethanol for three days at -20°C to

remove autofluorescence. We probed the cells using haptophyte-specific oligonucleotide probe PRYM02 (Simon *et al.,* 2000) targeting the 18S rRNA gene sequence. We used specially designed oligonucleotide probes to target haptophyte "A", the haptophyte present in the < 3 μm fraction. The sequence for this probe targeted basepairs 684 through 663 on the 18S rRNA molecule (5'-GCGCGTCCTTTTTCC-3'). We suspended cells hybridization buffer (18 μl 5M NaCl, 2ul 1M Tris–HCl pH 7.4, 1μl 1% SDS, 20μl 100% formamide, 1μl probe (0.2 nmol/μl), 58 μl distilled H2O) and the cells were incubated at 46°C for 2-h. We then centrifuged and resuspended the cells in wash buffer (4.3μl 5M NaCl, 2 μl 1M Tris, 1μl 1% SDS, 92.7μl distilled H2O) and incubated at 48 °C for 15 min. We centrifuged the cells and resuspended them in wash buffer, mounted this onto an agar coated slide, and dried at room temperature for 1-h. Finally, we mounted the cells with Citiflor:Vectashield (4:1 v/v) (London) to preserve probe fluorescence. Slides were viewed and photographed on a Zeiss Axioskop 2 MOT.

2.4 DNA extraction and sequencing

We extracted the DNA filters (Sterivex and Isopore) using a Puregene Cell Kit (Qiagen) according to the manufacturer's instructions. Genomic DNA was further purified using the Qiagen DNA Purification to remove proteins and other contaminants that inhibit PCR reactions. We quantified total extracted genomic DNA yields using a NanoDrop nucleic acid spectrophotometer (Thermo Scientific, Wilmington, DE).

For capillary sequencing, we amplified genomic DNA using haptophyte-specific oligonucleotide (Coolen *et al.,* 2004; Simon *et al.,* 2000) primers targeting 18S rRNA coding regions. Forward and reverse primers corresponded to *Escherichia coli* 18S rRNA

positions 429 and 887, respectively. Polymerase chain reactions (PCRs) were performed on an Eppendorf Gradient Thermocycler (Hamburg, Germany) with the following conditions after D'Andrea *et al*. (2006): 4 min initial denaturing at 96 °C, 35 cycles of denaturing for 30 s at 94 °C, followed by 40 s primer annealing at 55 °C and primer extension 40 s at 72 °C, with a final extension of 10 min at 72 °C. PCR reactions were run in 50 µl volume using Promega GoTaq polymerase (Madison, WI, USA). Triplicate PCR products were pooled for each sample, and templates were purified using an Invitrogen PureLink Purification kit (Carlsbad, CA, USA) with the high cut-off binding buffer to eliminate fragments < 200 bp. Purified PCR products were A-tailed and purified using the PureLink high cut-off kit. Cloning was performed using the Invitrogen TOP10 cloning kit with electro- competent cells. Protocol was according to the manufacturer's instructions. One hundred clones were picked for each sample. Plasmid DNA was isolated using a RevPrep Orbit robotic template preparation instrument (Genomic Solutions, Ann Arbor, MI), and prepared templates were sequenced on an ABI 3730XL (Applied Biosystems, Foster City, CA) capillary sequencer using the BigDye protocol with universal M13 forward and reverse primers according to the manufacturer's instructions. All sequencing was performed at the Marine Biological Laboratory W. M. Keck Ecological and Evolutionary Genetics Facility.

3. Results

3.1 Size fractionation experiment

The 5°C initial sediment enrichment culture was initially devoid of alkenones, but after six weeks the enrichment culture supernatant yielded a $C_{37:4}$ dominant alkenone

signature (**Figure 1**). This enrichment culture alkenone signature resembles those of Lake George sediments and previous sediment enrichment cultures (Toney *et al*. 2010; 2012), with a dominant $C_{37:4}$ isomer and the absence of a $C_{38:3}$ methyl ketone. We separated the supernatant into three size fractions: < 3 μm, 3-5 μm, and 5-10 μm fractions. Of the three size fractions, only the < 3 μm size fraction yielded an alkenone signature, and this alkenone signature resembled the signature of the sediment enrichment culture (**Figure 1**) and Lake George surface sediments.

3.2 Alkenone production in < 3μm culture

Culture of the < 3 μm fraction yielded dominant $C_{37:4}$ alkenones at 5°C, comprising about 68% of the total C37 alkenones (**Table 1**). This was on par with concentrations of $C_{37:4}$ that were observed in enrichment cultures (Toney *et al*., 2012) that ranged from 78% in the early stage of the experiment at 4°C to 47% at later stages during the experiment (deep culture, Toney *et al*., 2010). The abundance of the $C_{37:4}$ compound decreased as temperature increased, and 10, 15, 21 and 24°C cultures were all characterized by a dominant $C_{37:3}$ (**Figure 2**).

Alkenone concentrations per cell were similar to previous observations of alkenone concentrations in *Isochrysis galbana* (0.009-2pg/cell; Versteegh *et al*., 2001; Marlowe *et al*., 1984), primarily about 0.2-2pg/cell. However, the 5°C culture had much higher alkenone concentrations, up to 8pg/cell. None of the cultures reached stationary phase by the termination of the experiment, and cultures between 10 and 24°C had double rates between 2-3 days (**Table 1**). The 5°C culture had the slowest growth rates with doubling times of greater than 7 days. The highest growth rates were observed at

15°C and lowest at 5°C. The lowest alkenone per cell concentrations were observed at 15°C and the highest at 5°C (**Table 1**).

The alkenones in the < 3μm fraction demonstrated a response temperature, increasing in $C_{37:3}$ alkenone concentrations with warmer temperatures (**Figure 3**). The degree of unsaturation of alkenone lipids at the different temperatures yielded a U^{k}_{37} calibration of

$$\mathbf{U^{k}_{37} = 0.0373T - 0.7531}$$

where U^{k}_{37} is the concentration of $(C_{37:2}-C_{37:4})/(C_{37:2}+C_{37:3}+C_{37:4})$ and T is temperature in Celsius. This calibration had a root mean squared error value of 1.88 °C. The $\mathbf{U^{k'}_{37}}$ calibration did not demonstrate a response to temperature (**Figure 4**). The U^{k}_{37} calibration shared a similar slope to the U^{k}_{37} calibration for the *Chrysotila lamellosa* culture (Sun *et al*., 2007) and had a similar y-intercept to the German lakes U^{k}_{37} calibration (Zink *et al*., 2001). The *C. lamellosa* culture did not have a dominant $C_{37:4}$ alkenone, although $C_{37:4}$ alkenones were present. The Zink *et al*. (2001) calibration as well included data from lakes with dominant $C_{37:4}$ alkenones.

3.3 Confirmation of Hap-A identity in < 3µm culture

To verify that alkenones were being produced in culture by Hap-A and only Hap-A, we performed a FISH experiment using both a haptophyte-specific and a Hap-A specific probe. Using source material from the 5°C <3μm size fraction, the haptophyte specific probe PRYM02 only illuminated sub < 5μm sized cells (**Figure 5A**). These small cells were also illuminated with the Hap-A specific primer, confirming that the

only haptophytes in the < 3μm culture were Hap-A cells. To better observe the morphology of these cells, which is compromised when cells are preserved in ethanol as during the FISH protocol, we photographed the < 3μm fraction under DIC light (**Figure 6**). These cells are approximately 5μm and smaller, which can be expected if the cell grows after separation from the source material with the 3μm filter. The cells also appear to be engorged with lipid vacuoles.

4. Discussion

4.1 *Cellular alkenone concentrations*

The lowest growth rates and highest alkenone concentrations occurred at 5°C, in agreement with previous observations of *E. huxleyi* and *G. oceanica* cultures which had highest alkenone concentrations at lowest growth temperatures (Conte *et al.,* 1998). Alkenones are believed to be energy storage molecules in haptophytes (Epstein *et al.,* 2001; Eltgroth *et al.,* 2005) and alkenone concentrations per cell increased during stationary growth phase and decrease after cultures are placed in the dark (Epstein *et al.,* 2001; Eltgroth *et al.,* 2005). Our observed accumulation of alkenones at low temperatures and low growth rates may be the result of photosynthetic energy input exceeding the cell's capacity for growth and division (Roessler, 1990) as was observed at both low temperatures and low nutrient conditions (Conte *et al.,* 1998; Epstein *et al.,* 2001).

The lowest growth temperature is also expected to have the highest $C_{37:4}$ concentration as this is the premise for the U^{k}_{37} alkenone unsaturation index. The accumulation of more unsaturated alkenones has previously been observed when haptophytes were grown under nutrient deplete conditions, and therefore lower growth rates (Prahl *et al.,* 2003; Prahl *et al.,* 2006). This is in agreement with what we observed,

with the highest $C_{37:4}$ concentrations at the lowest growth rate at 5°C. The switch to $C_{37:3}$ dominated alkenone signature at warmer temperatures may be a reflection of growth rate and cell physiology, indicating that the $C_{37:4}$ dominance is dependent on cell physiological status. The conservation of the U^{k}_{37} index over the five temperature experiments suggests that the culture is responding systematically to variations in temperature.

4.2 U^{k}_{37} *index*

The U^{k}_{37} calibration for the < 3 μm culture had a y-intercept close to that of the German lakes and a very similar slope to the *C. lamellosa* calibration. The $U^{k'}_{37}$ index failed to capture the alkenone response to temperature, indicating the importance of incorporating the $C_{37:4}$ alkenone when calibrating the Hap-A culture <3 μm culture.

Alkenone signatures were $C_{37:4}$ dominant in the 5°C culture but $C_{37:3}$ alkenones were dominant at warmer temperatures (Figure 2). Our FISH results confirm that Hap-A was the single haptophyte in the Hap-A cultures, and this culture served as the inoculum for warmer cultures. We did not perform a FISH experiment to confirm that Hap-A was the only haptophyte in the warmer cultures, but the absence of a $C_{38:3}$ methyl ketone in the warmer cultures is confirmation that the cultures were not contaminated with Hap-B cells (Toney *et al*., 2012). Lake George enrichment cultures demonstrate a $C_{37:4}$-dominant alkenone signature at 20°C (Toney *et al*., 2012) which our pure cultures failed to reproduce. The higher growth rates afforded by warmer temperatures may have resulted

in lower $C_{37:4}$ alkenone production, as has previously been observed in fast growing *E. huxleyi* cultures (Prahl *et al*., 2003; Prahl *et al*., 2006).

4.3 Picohaptophytes

DNA sequencing revealed the dominant $C_{37:4}$-producing haptophyte from Lake George is a close relative of other dominant $C_{37:4}$-producing, uncultured haptophytes from Ace Lake, Antarctica (Theroux, *et al*., 2010; Coolen *et al*., 2004). Our discovery that the Lake George Hap-A is of very small size suggests an ecological context for the occurrence of these unique alkenone producers. Picoplankton are often found in oligotrophic waters, their high surface area to volume ratio an advantage in low nutrient conditions (Vaulot *et al*. 2008). Hap-A was isolated from anoxic Lake George surface sediments (40m depth) and blooms *in situ* upon spring thaw and in culture when exposed to light (Toney *et al*., 2010; Toney *et al*., 2012). Hap-A's high surface-area to volume ratio likely provides a competitive advantage to survive long periods of time out of the photic zone. Although some larger haptophytes can be mixotrophic (Brutemark and Graneli, 2011), the small size of Hap-A cells likely prevents it from being mixotrophic (Robert Andersen, personal communication).

Cyanobacteria compose a greater percentage of the picophototrophic microbial community, but the high cellular carbon concentrations and high growth rates of picoeukaryotes make them the largest fixers of carbon in the world's oceans (Li, 1996; Liu *et al*., 2009). Haptophytes account for up to 35% of the Atlantic Ocean picoeukaryote community (Not *et al*., 2005) and up to 25% of the global picophytoplankton community biomass (Cuvelier *et al*., 2010). Picoplankton haptophytes are known contributors to

primary production in ocean waters as indicated by microscopy, fluorescent *in situ* probing and pigment analyses (Not *et al*., 2005; Liu *et al*., 2009), but the majority of environmental haptophyte DNA sequences identify novel, uncultured species (Cuvelier *et al*., 2010). Picoplanktonic haptophytes can comprise 90% of the haptophyte community in marine waters (Cuvelier *et al*., 2010). Only six picoplanktonic haptophyte species are in culture, including species from genus *Phaeocystis* and *Chrysochromulina,* although none of these are alkenone-producers (Vaulot *et al*., 2008, Marlowe *et al*., 1984). Our paper provides the first description of an alkenone-producing pico-haptophytes.

Hap-A likely has a cyst hibernation stage that allows it to survive the long winters in Lake George (Toney *et al*., 2012). When haptophytes are incubated in the dark, they metabolize their alkenones beginning with the more-unsaturated compounds (Epstein *et al*., 2001). Similarly, when cells are grown in nutrient deplete conditions, they accumulate more unsaturated alkenones (Epstein *et al*., 2001; Eltgroth *et al*., 2005). The abundance of $C_{37:4}$ alkenones in Hap-A may serve as a competitive advantage, providing large storage of readily metabolized compounds. High concentrations of $C_{37:4}$ alkenones, in addition to a high surface area to volume ratio, may allow Hap-A cells to survive extended hibernation periods in anoxic waters, an advantage other haptophytes, including Lake George Hap-B, do not have.

5. Conclusions

Our isolation and cultivation of the haptophyte Hap-A finally revealed that this $C_{37:4}$-dominant haptophyte is picoplanktonic, a characteristic that has likely allowed for its success in low-nutrient environments. The cultures of Hap-A displayed the

characteristic $C_{37:4}$ dominant alkenone signature at 5°C but not at higher temperatures. Fluorescence probing and the absence of a C_{38} methyl ketone confirmed that the Hap-A culture was in fact pure of other haptophyte contaminants and $C_{37:3}$ dominance likely arises because of high growth rates. What we now know about the Lake George picohaptophyte has implications for the application of the U^k_{37} paleothermometry in lake environments where dominant $C_{37:4}$ alkenone signatures have been an enduring mystery. The isolation and description of this novel species will further improve our understanding of haptophyte diversity and ecology in lake environments. Our results finally shed light on the ecology of a haptophyte that spends nine months of its year in the dark.

Acknowledgements

This work was supported by a National Science Foundation award to Y. Huang (EAR-1122749) and L. Amaral-Zettler (EAR-1124192), a Brown SEED fund to Y. Huang and L. Amaral-Zettler, and an American Association of University Women dissertation fellowship to S. Theroux.

References

Brassell SC, Eglinton G, Marlowe IT, Pflaumann U, Sarnthein M. (1986). Molecular stratigraphy - a new tool for climatic assessment. Nature 320:129-133.

Brutemark A, Graneli E. (2011). Role of mixotrophy and light for growth and survival of the toxic haptophyte *Prymnesium parvum*. Harmful Algae 10: 388-394.

Chu GQ, Sun Q, Li SQ, Zheng MP, Jia XX, Lu CF, Liu JQ, Liu TS. (2005). Long-chain alkenone distributions and temperature dependence in lacustrine surface sediments from China. Geochimica Et Cosmochimica Acta 69:4985-5003.

Conte MH, Thompson A, Lesley D, Harris RP. (1998). Genetic and physiological influences on the alkenone/alkenoate versus growth temperature relationship in Emiliania huxleyi and Gephyrocapsa oceanica. Geochimica Et Cosmochimica Acta 62:51-68.

Coolen MJL, Muyzer G, Rijpstra WIC, Schouten S, Volkman JK, Sinninghe Damsté J. (2004). Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake. Earth and Planetary Science Letters 223:225-239.

Cranwell PA. (1985). Long-Chain Unsaturated-Ketones in Recent Lacustrine Sediments. Geochimica Et Cosmochimica Acta 49:1545-1551.

Cuvelier ML, Allen AE, Monier A, McCrow JP, Messie M, Tringe SG, Woyke T, Welsh RM, Ishoey T, Lee J-H, Binder BJ, DuPont CL, Latasa M, Guigand C, Buck KR, Hilton J, Thiagarajan M, Caler E, Read B, Lasken RS, Chavez FP, Worden AZ. (2010). Targeted metagenomics and ecology of globally important uncultured eukaryotic phytoplankton. Proceedings of the National Academy of Sciences 107:14679-14684.

Cuvelier ML, Allen AE, Monier A, McCrow JP, Messie M, Tringe SG, Woyke T, Welsh RM, Ishoey T, Lee J-H, Binder BJ, DuPont CL, Latasa M, Guigand C, Buck KR, Hilton J, Thiagarajan M, Caler E, Read B, Lasken RS, Chavez FP, Worden AZ. (2010). Targeted metagenomics and ecology of globally important uncultured eukaryotic phytoplankton. Proceedings of the National Academy of Sciences 107:14679-14684.

D'Andrea WJ, Huang Y. (2005). Long chain alkenones in Greenland lake sediments: Low δ13C values and exceptional abundance. Organic Geochemistry 36:1234-1241.

D'Andrea WJ, Lage M, Martiny JBH, Laatsch AD, Amaral-Zettler LA, Sogin ML, Huang YS. (2006). Alkenone producers inferred from well-preserved 18S rDNA in Greenland lake sediments. Journal of Geophysical Research-Biogeosciences 111:doi:10.1029/2005JG000121.

Eltgroth ML, Watwood RL, Wolfe GV. (2005). Production and cellular localization of neutral long-chain lipids in the haptophyte algae Isochrysis galbana and Emiliania huxleyi. Journal of Phycology 41:1000-1009.

Epstein BL, D'Hondt S, Hargraves PE. (2001). The possible metabolic role of C-37 alkenones in Emiliania huxleyi. Organic Geochemistry 32:867-875.

Guillard RRL (1975). Culture of phytoplankton for feeding marine invertebrates. In Culture of marine invertebrate animals. WL Smith, MH Chanley, eds. Plenum Press, New York, p. 26.

Li WKW. (1995). Composition of ultraphytoplankton in the central North Atlantic. Mar Ecol Prog Ser 122:1–8.

Liu H, Probert I, Uitz J, Claustre H, Aris-Brosou S, Frada M, Not F, de Vargas C. (2009). Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. Proceedings of the National Academy of Sciences of the United States of America 106:12803-12808.

Moon-van der Staay SY, van der Staay GWM, Guillou L, Vaulot D, Claustre H & Medlin LK (2000) Abundance and diversity of prymnesiophytes in the picoplankton community from the equatorial Pacific Ocean inferred from 18S rDNA sequences. Limnol Oceanogr 45: 98–109.

Marlowe IT, Green JC, Neal AC, Brassell SC, Eglinton G, Course PA. (1984). Long-Chain (n-C37-C39) alkenones in the prymnesiophyceae. Distribution of alkenones and other lipids and their taxonomic significance. British Phycological Journal 19:203-216.

Müller PJ, Kirst G, Ruhland G, von Storch I, Rosell-Mele A. (1998). Calibration of the alkenone paleotemperature index $U^{k'}_{37}$ based on core-tops from the eastern South Atlantic and the global ocean (60°N-60°S). Geochimica Et Cosmochimica Acta 62:1757-1772.

Not F, Massana R, Latasa M, Marie D, Colson C, Eikrem W, Pedros-Alios C, Vaulot D, Simon N. (2005). Late summer community composition and abundance of photosynthetic picoeukaryotes in Norwegian and Barents Seas. Limnology and Oceanography 50:1677-1686.

Pearson EJ, Juggins S, Farrimond P. (2008). Distribution and significance of long-chain alkenones as salinity and temperature indicators in Spanish saline lake sediments. Geochimica Et Cosmochimica Acta 72:4035-4046.

Prahl FG, Wolfe GV, Sparrow MA. (2003). Physiological impacts on alkenone paleothermometry. Paleoceanography 18:doi:10.1029/2002PA000803.

Prahl FG, Muehlhausen LA, Zahnle DL. (1988). Further evaluation of long-chain alkenones as indicators of paleoceanographic conditions. Geochimica Et Cosmochimica Acta 52:2303-2310.

Prahl FG, Wakeham SG. (1987). Calibration of unsaturation patterns in long-chain ketone compositions for paleotemperature assessment. Nature 330:367-369.

Prahl FG, Mix AC, Sparrow MA. (2006). Alkenone paleothermometry: Biological lessons from marine sediment records off western South America. Geochimica et Cosmochimica Acta 70:101-117.

Riebesell U, Zondervan I, Rost B, Tortell PD, Zeebe RE, Morel FMM. (2000). Reduced calcification of marine plankton in response to increased atmospheric. Nature 407:364-367.

Roessler PG. (1990). Environmental control of glycerolipid metabolism in microalgae: commercial implications and future research directions. Journal of Phycology 26:393–399.

Rontani JF, Beker B, Volkman JK. (2004). Long-chain alkenones and related compounds in the benthic haptophyte Chrysotila lamellosa Anand HAP 17. Phytochemistry 65:117-126.

Sikes EL, Volkman JK. (1993). Calibration of alkenone unsaturation ratios ($U^{k'}_{37}$) for paleotemperature estimation in cold polar waters. Geochimica et Cosmochimica Acta 57:1883-1889.

Sun Q, Chu GQ, Liu GX, Li S, Wang XH. (2007). Calibration of alkenone unsaturation index with growth temperature for a lacustrine species, Chrysotila lamellosa (Haptophyceae). Organic Geochemistry 38:1226-1234.

Theroux SM, D'Andrea WJ, Toney JL, Amaral-Zettler LA, Huang Y. (2010). Phylogenetic diversity and evolutionary relatedness of alkenone producing haptophyte algae in lakes: Implications for continental paleotemperature reconstructions. Earth and Planetary Science Letters 300:311-320.

Toney JL, Huang Y, Fritz SC, Baker PA, Grimm E, Nyren P. (2010). Climatic and environmental controls on the occurrence and distributions of long chain alkenones in lakes in the interior United States. Geochimica Et Cosmochimica Acta

Toney JL, Theroux S, Andersen RA, Coleman A, Amaral-Zettle L, Huang Y. (2012). Culturing of the first $_{37:4}$ predominant lacustrine haptophyte: Geochemical, biochemical, and genetic implications. Geochimica et Cosmochimica Acta 78:51 - 64.

Vaulot D, Eikrem W, Viprey M, Moreau H. (2008). The diversity of small eukaryotic phytoplankton (‚â§3¬†Œºm) in marine ecosystems. FEMS Microbiology Reviews 32:795-820.

Versteegh GJM, Riegman R, de Leeuw JW, Jansen JHF. (2001). $U^{k'}_{37}$ values for *Isochrysis galbana* as a function of culture temperature, light intensity and nutrient concentrations. Organic Geochemistry 32:785-794.

Zink KG, Leythaeuser D, Melkonian M, Schwark L. (2001). Temperature dependency of long-chain alkenone distributions in Recent to fossil limnic sediments and in lake waters. Geochimica Et Cosmochimica Acta 65:253-265.

Table 1. Alkenone abundance data for < 3μm culture experiment.

Temp. (°C)	Final cell concentration (cell/ml)	Growth rate (Doubling time)	C37:4 (ng/L)	C37:3 (ng/L)	C37:2 (ng/L)	C37total (ng/L)	% C37:4	C37alk/ cell (pg)	Uk37	Uk'37
5	60000	7.22	6929	2636	491	10056	68.9	0.17	-0.64	0.16
10	800000	3.16	7876	15465	0	23341	33.7	0.03	-0.34	0.00
15	5120000	2.25	5139	16337	2813	24288	21.2	0.00	-0.10	0.15
21	640000	3.32	2119	27009	1586	30713	6.9	0.05	-0.02	0.06
24	800000	3.161	958	18385	3612	22956	4.2	0.03	0.12	0.16

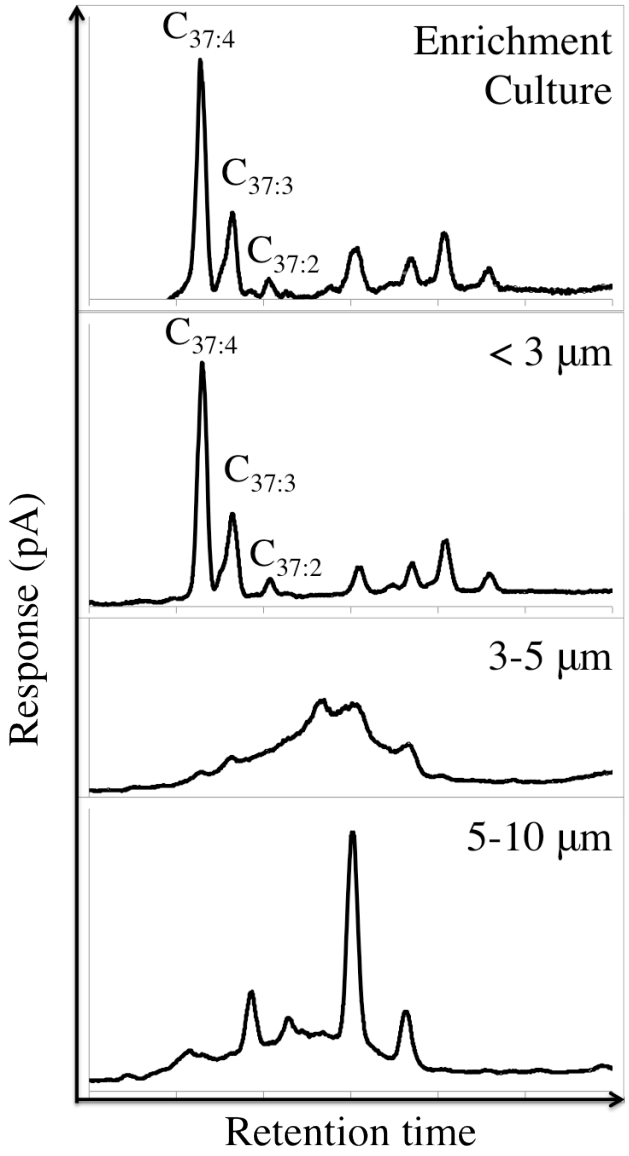

Figure 1. Gas chromatograms of culture size fractions. Enrichment culture was subdivided into < 3μm , 3-5μm, and a 5-10μm fractions. Enrichment culture chromatogram is at 6 weeks.

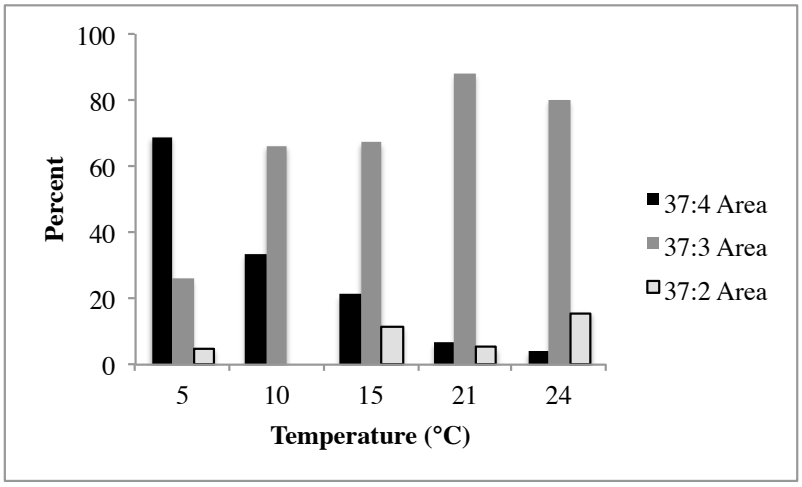

Figure 2. Percent abundances of < 3μm culture alkenones at all temperatures.

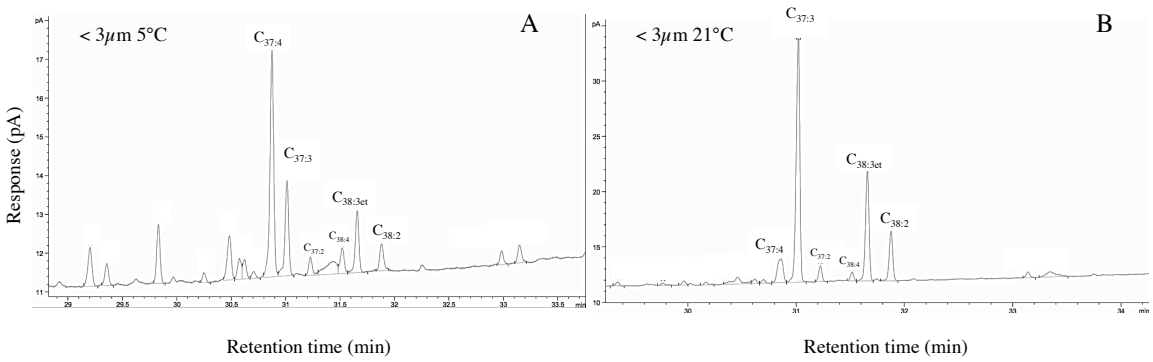

Figure 3. Abundance of C37 isomers in the < 3μm cultures. A. GC-FID chomatogram of < 3μm culture at 5°C culture. B. GC-FID chromatogram of <3μm culture at 21°C.

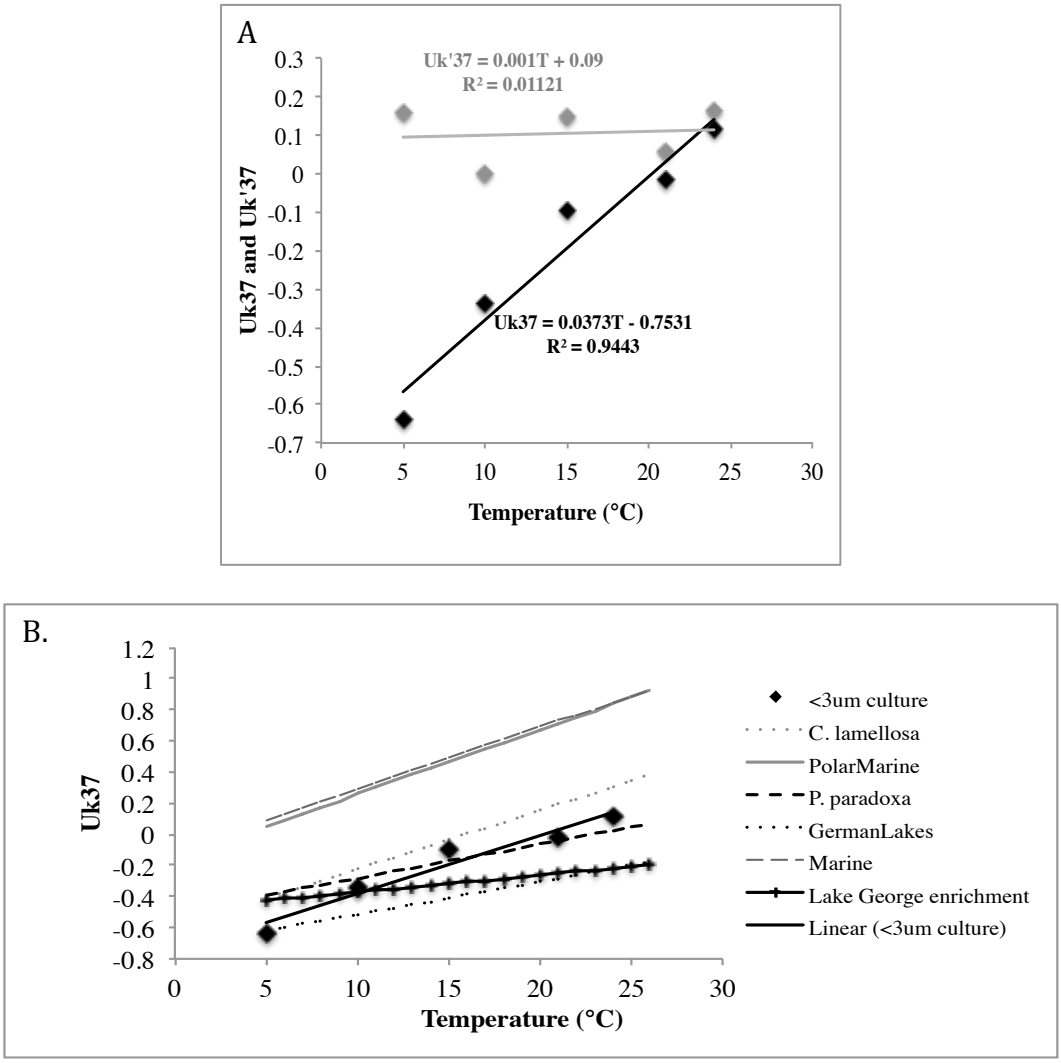

Figure 4. Alkenone unsaturation indices for the < 3μm fraction culture A. U^{k}_{37} and $U^{k'}_{37}$ calibrations. The RMSE for the < 3μm U^{k}_{37} calibration is 1.88°C. B. The U^{k}_{37} index for both the <3um culture and previously published calibrations. References for the calibrations are: *C. lamellosa* (Sun *et al.,* 2007), Marine Polar (Sikes and Volkman 1993), *P. paradoxa* (Theroux, Chapter 4), German lakes (Zink *et al.* 2001), Marine calibration (Prahl *et al.,* 1998), Lake George enrichment (Toney *et al.,* 2012).

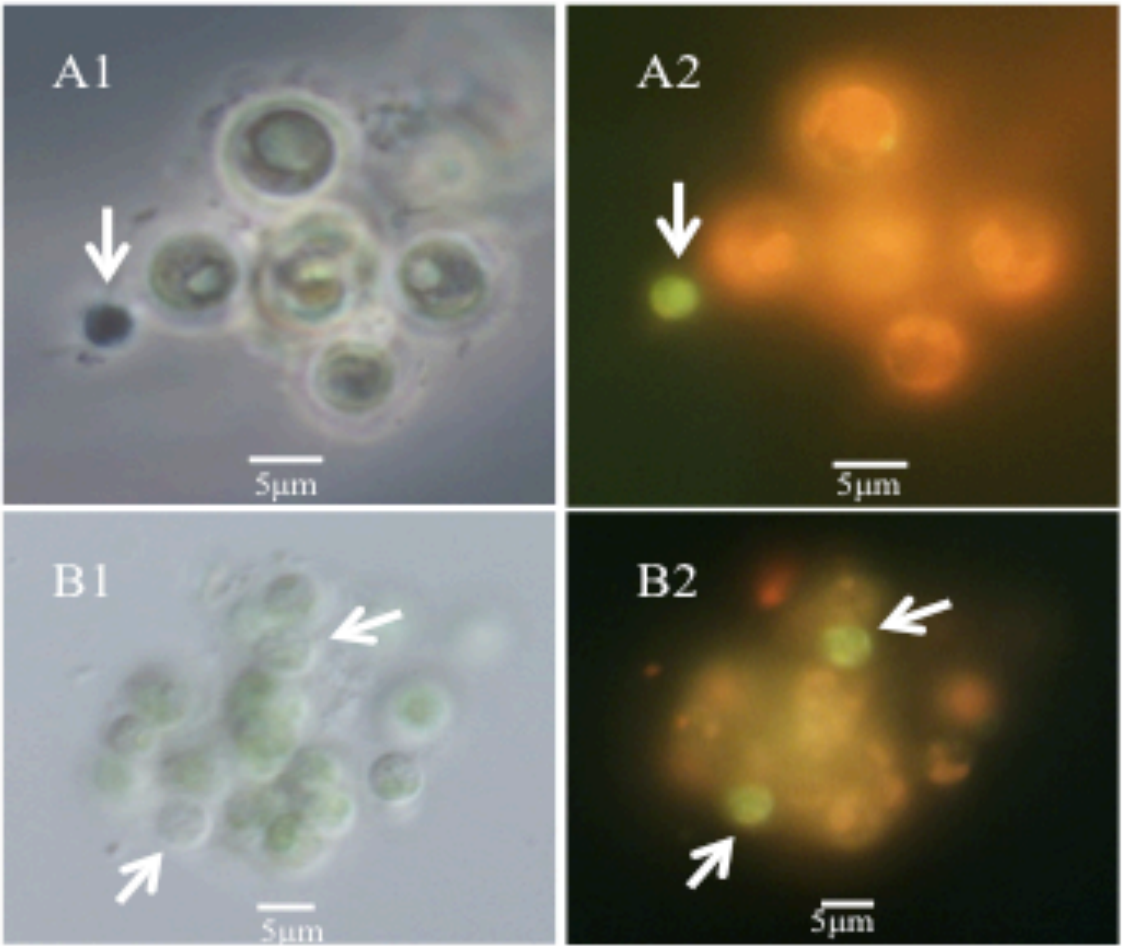

Figure 5. Lake George enrichment culture < 3*µ*m fraction hybridized with haptophyte and Lake George Hap-A specific probes. A1) Light micrograph of enrichment culture hybridized with haptophyte-specific oligonucleotide probe. A2) Same as A1 under blue light. Arrows denote fluorescing cells. B1) Light micrograph of enrichment culture with Lake George Hap-A specific probe. B2) Same as B1 with blue light. Arrows denote fluorescing cells.

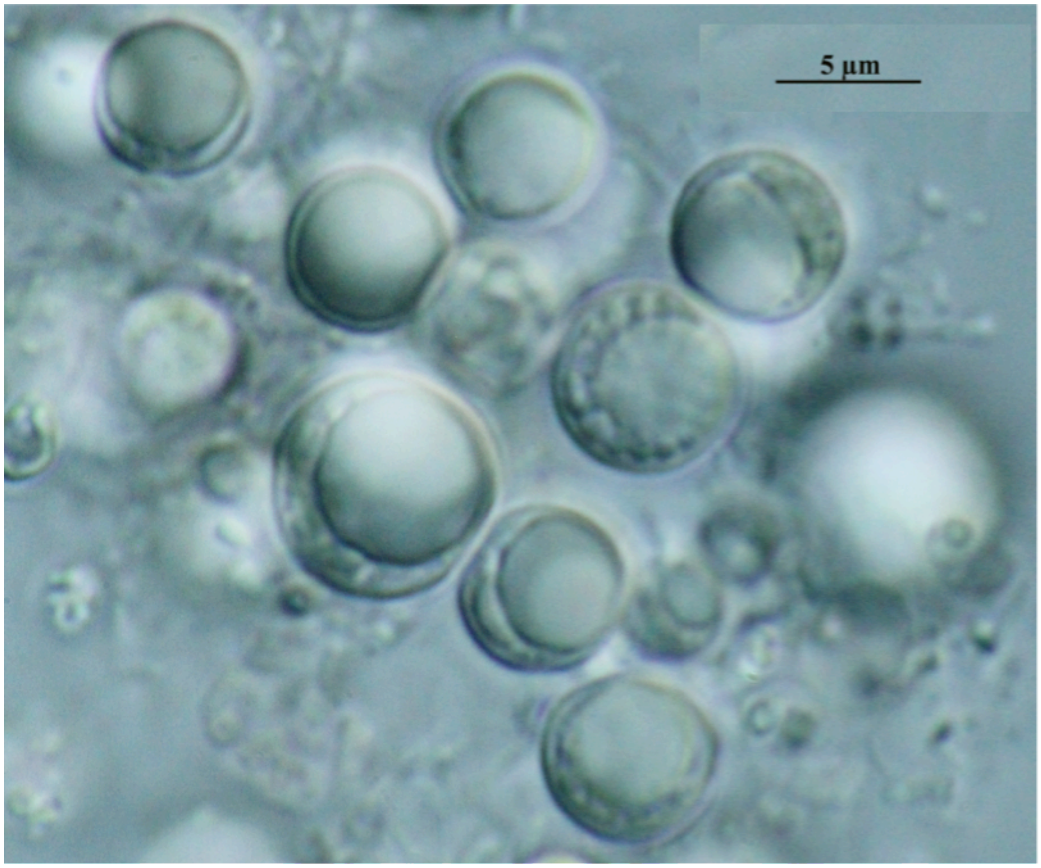

Figure 6. Photomicrograph of < 3μm culture. Cells appear to be engorged with lipid vacuoles.

CHAPTER 7

LATE BLOOMERS: *IN SITU* and *EX SITU* OBSERVATIONS OF HAPTOPHYTE SPECIES DISTRIBUTIONS THROUGHOUT THE COURSE OF A SEASONAL BLOOM

SUSANNA THEROUX[1,2]

JAIME L. TONEY[3]

ROBERT ANDERSEN[4]

PAUL NYREN[5]

RICK BOHN[5]

JEFFREY SALACUP[1]

YONGSONG HUANG[1]

LINDA AMARAL-ZETTLER[1,2,6]

[1]BROWN UNIVERSITY, *Department of Geological Sciences*

[2]MARINE BIOLOGICAL LABORATORY, *Josephine Bay Paul Center*

[3]UNIVERSITY OF GLASGOW, *School of Geographical and Earth Sciences*

[4]UNIVERSITY OF WASHINGTON, *Friday Harbor Laboratories*

[5]NORTH DAKOTA STATE UNIVERSITY, *Central Grasslands Research Extension Center*

[6]BROWN UNIVERSITY, *Department of Ecology and Evolutionary Biology*

Prepared for:

Geochimica et Cosmochimica Acta

Abstract

The ability to reconstruct paleotemperatures using the alkenone-based U^k_{37} temperature proxy is dependent upon the selection of a proper U^k_{37} calibration. When multiple haptophyte species occur in a single environment, the selection of an appropriate species-specific U^k_{37} calibration is unclear. Previous research revealed the presence of two haptophyte species in Lake George, ND and the contribution to the alkenone sediment record by each phylotype was unknown. We analyzed the haptophyte community during the spring bloom event in Lake George, ND to determine the haptophyte (or haptophytes) responsible for the seasonal alkenone production. We report the first study to monitor a lacustrine haptophyte bloom throughout the course of its seasonal cycle using a combined approach of DNA sequencing, quantitative PCR and geochemical analyses. Our results revealed a seasonal succession in haptophyte species, with an early and late alkenone peak dominated by different haptophyte species. Culture-based and *in situ* analyses revealed individual alkenone fingerprints for each haptophyte species, although at times the water column alkenones were composite signatures from both species. Our results established a suite of U^k_{37} calibrations that can be used in consortium to accurately reconstruct paleotemperatures from Lake George sediments. This multidisciplinary approach can serve as a guideline for future alkenone-based temperature reconstructions from multiple haptophyte environments.

1. Introduction

Since its establishment in 1986 (Brassell *et al.,* 1986), the alkenone-based paleotemperature proxy (U^{k}_{37} and $U^{k'}_{37}$) has been a popular and reliable tool for reconstructing sea surface temperatures from marine sediment records. The desire to extend the use of this paleotemperature proxy to lake sediments resulted in an increased interest in lacustrine alkenone production. The U^{k}_{37} proxy relies upon the degree of unsaturation of alkenone compounds produced exclusively by haptophyte algae in the order Isochrysidales. Marine alkenones are predominantly produced by cosmopolitan haptophyte species *Emiliania huxleyi* and its close relative *Gephyrocapsa oceanica*. Their predominance in marine waters allows for the application of a single, universal $U^{k'}_{37}$ calibration, one that was first developed empirically though culture work (Brassell, 1986; Prahl and Wakeham 1987) and later confirmed with a global core-top survey (Müller *et al*.,1998).

Like their marine counterparts, lacustrine alkenone unsaturation ratios correspond to lake seasonal surface water temperatures and mean annual air temperature (Chu *et al*., 2005; D'Andrea and Huang, 2005; Sun *et al*., 2007; Zink *et al*., 2001). However, the novelty and unexpected diversity of lacustrine haptophyte species has impeded the application of this proxy to lacustrine sediments. Only two species of coastal/lacustrine alkenone-producing haptophytes are available in culture, *Chrysotila lamellosa* and *Isochrysis galbana (*Versteegh *et al*. 2001; Rontani 2004; Sun *et al*., 2007). In lakes such as Lake George, ND, the presence of multiple haptophyte species complicates the selection of a U^{k}_{37} calibration (Theroux *et al*., 2010; Toney *et al*., 2012).

Traditional patterns in alkenone distribution differ between marine and lacustrine environments. Marine sediments typically have low concentrations of the $C_{37:4}$ alkenones and the presence of the $C_{38:3}$ methyl ketone while in contrast, lacustrine alkenone profiles are characterized by abundant tetraunsaturated alkenone isomers ($C_{37:4}$) and the absence of a $C_{38:3}$ methyl ketone (Chu *et al., 2*005; Liu *et al., 2*008; Pearson *et al., 2*008). Haptophyte species responsible for alkenone production in lakes were once inferred from alkenone lipid signatures, but DNA sequencing has revealed a more complex species history in many locations (Coolen *et al*., 2004; D'Andrea and Huang, 2005; Theroux *et al*., 2010). The discovery of a novel clade of haptophytes from a series of lakes in Greenland (D'Andrea *et al*., 2006; Theroux *et al., 2*010) highlights the potentially untapped diversity of haptophyte species in lakes. Lacustrine haptophytes experience bloom events (D'Andrea *et al*., 2011; Toney *et al*., 2010) that result in large inputs of alkenones to the lake sediment record. The environmental triggers of seasonal alkenone-producing haptophyte blooms are unknown, although nutrient loading and seasonal irradiance levels can cause marine bloom events (Tyrrell and Merico, 2004).

Alkenone production in Lake George, North Dakota is attributed to two haptophyte species whose DNA was recovered from the lake's surface sediments. Hap-A is a close relative of haptophytes from Ace Lake, Antarctica (Coolen *et al*., 2004) and cultured species *Isochrysis galbana* (Theroux *et al*., 2010). Hap-B is a close relative of cultured haptophytes *Pseudoisochrysis paradoxa* and *Chrysotila lamellosa* (Theroux *et al*., 2010). Lake George is situated in the Northern Great Plains and has a pristine alkenone sediment record that reflects the lake's environmental change over the past 8000 years (Toney, 2011). Climate reconstructions from tree rings and speleotherms in

this region are discontinuous or scarce (Toney, 2011), emphasizing the importance of a reliable temperature proxy.

The accurate estimation of paleotemperature from the Lake George sediment record requires an improved understanding of both the ecology of these novel haptophyte species and their modes of alkenone production. In particular, our study aimed to determine a) if both haptophyte species are responsible for alkenone production and b) if both species produce similar alkenone signatures. To address these questions, we sampled the 2011 Lake George haptophyte bloom event and analyzed the geochemical and biological artifacts of the bloom. We used high-throughput DNA sequencing, quantitative PCR and culture manipulations to gauge haptophyte succession and alkenone production, three approaches that yielded corroborating results. Our study provides the first comprehensive depiction of species succession during a haptophyte bloom and uncovers how the haptophyte seasonal cycle is recorded in the lake sedimentary record.

2. Methods

2.1 Site description

Lake George (46.74°N, 99.49°W) is situated in the Northern Great Plains in North Dakota (**Figure 1**). Lake George has a maximum depth of 60m and a salinity averaging 9.72 ppm, making it unique among the lakes in North Dakota (Fritz, 2011). Glacial retreat formed the lake over 12,000 years ago, and the lake's salinity is likely maintained by input from deep sources of saline groundwater (ibid). Lake George sits atop the Hell Creek, Fox Hills and Pierre formations of the Upper Cretaceous (Whitehead, 1996), and

sources its high sulfate concentrations (7370 mg/L) from the Pierre shale. The average pH of the lake waters is 8.98 (Toney *et al*., 2010).

2.2 Bloom sampling

In April 2011 we began sampling the waters of Lake George to gauge haptophyte cell and alkenone abundance throughout the course of the seasonal cycle. Biweekly sampling of Lake George continued until late July. We collected water using a VanDorn water sampler at 0, 5, and 10m depths. From each depth, we collected 1L each for DNA analysis, alkenone analysis, for geochemical analysis. For alkenone analysis, we filtered one liter of water onto a pre-combusted (550°C) GF/F 0.7μm, 47 mm glass filter, and kept it frozen until analysis. For DNA analysis, we filtered a separate liter of lake water onto a 0.2μm Sterivex^{TM} filter (Millipore, Billerica, MA, USA), flooded the filter with Puregene lysis buffer (Qiagen, Valencia, CA, USA), and froze it at -20°C until processing. Water collected for geochemical analysis was filtered through a 0.2 μm filter to remove microbes and particulate matter and stored at -20°C until processing.

2.3 Geochemical analysis

We used a YSI 85 meter to measure lake water temperature, pH, dissolved oxygen and conductivity *in situ*. We measure nitrate, nitrite and phosphate using a Westco Smartchem 200 Discrete Analyzer at Brown University. Trace metals were analyzed at the Barnstable County Department of Health and Environment Laboratory using Agilent ICP-MS EPA method 200.8 and a Perkin Elmer Flame Atomic Absorption Spectrometer (Ca, Mg, Na).

2.4 Mock-bloom event

We initiated a sediment enrichment culture after Toney *et al*. (2011). We inoculated a glass beaker with 50g Lake George surface sediment overlain with 2 liters of 0.2 μm filtered Lake George water amended with f/2 nutrients (Guillard 1975). We placed this sediment enrichment culture in a 4°C incubator with a 24:0 light:dark light regime using a full-spectrum light. We performed biweekly sampling of the enrichment culture supernatant to gauge cell numbers and collect water for DNA and alkenone analysis. We filtered 50 ml onto a 44 mm GFF Whatman filter for alkenone analysis and kept the filter at -20°C until extraction. We filtered 15 ml onto a 25 mm 0.2 μm Isopore (Millipore) cellulose nitrate filter for DNA analysis and flooded the filter with lysis buffer and stored at -20°C until extraction.

2.5 Culture studies to determine Hap-B U^{k}_{37} *calibration*

We achieved individual clonal cultures of Haptophyte B (#A17-903) through cell-picking with sterilized glass pipettes. We grew these clonal cultures in glass vials in 0.2µm filtered Lake George water amended with f/2 nutrients (Guillard 1975). We grew the cultures at 5,10,15,21 and 24°C under 24:0 light:dark regime using a full-spectrum light. The cultures were sampled biweekly for cell counts and cell concentrations were determined using a haemocytometer to ensure the cultures were maintained in the logarithmic growth phase. At the end of two months, we harvested the cultures for alkenone analysis.

2.6 Alkenone extraction and analysis

We extracted the alkenone filters using three 20-minute bursts of sonication in dichloromethane (DCM), and ran the total lipid extracts on an Agilent 6890plus Gas Chromatograph Flame Ionization Detector (GC-FID) for detection and quantification of alkenones using an internal C36 n-alkane standard and an external alkenone standard of known U^{k}_{37}-temperature value to ensure analytical precision (< 0.1°C). We used a Varian VF200 60m GC column (60m × 250μm × 0.10 μm) with the following temperature program: an initial temperature of 100°C (hold 1-min), ramp 20°C/min to 200°C (hold 1-min), ramp 4°C/min to 320°C (hold 0 min), hold at 320°C for 5-min.

Alkenones were extracted from the sediment and/or filters and isolated from co-eluting compounds before calculation of alkenone concentrations and ratios. Sediments and filters were freeze-dried before analysis. Samples were then powdered and extracted three times with dichloromethane on an Accelerated Solvent Extractor (Dionex, ASE200) at 150°C and 1500 psi to produce a total lipid extract (TLE). TLEs were then separated over silica gel using hexane, dichloromethane, and methanol to yield hydrocarbon, ketone (alkenone), and polar fractions, respectively.

Two quantification standards (n-C36 and n-C37 alkanes) were added to all samples before being injected from an autosampler into a 112°C CIS-PTV (cooled injection system-programmed temperature vaporizer) inlet operated in solvent vent mode. After the initial vent, the inlet was ramped at 12°C/min to 240°C, held isothermally for 5 minutes, ramped again at 12°C to 320°C, and held isothermally for 2 minutes before cryogenic cooling. A 60m, 0.32mm ID, 0.10um filmDB-1 with a 5m fused guard column (DB-1 duraguard) was used. The oven began at 90°C for 2 minutes, was ramped at 40°C/min to 255, at 1°C to 302, and at 10°C to 325 where it was held isothermally for 20

minutes. Hydrogen was used as a carrier gas. Analytical accuracy, tracked via the injection of laboratory $U_{37}^{k'}$ standard of known temperature, was ±0.042 $U_{37}^{k'}$ units (±0.1°C).

2.7 DNA extraction and sequencing

We extracted the DNA filters (Sterivex and Isopore) using a Puregene Cell Kit (Qiagen) according to the manufacturer's instructions. Genomic DNA was further purified using the Qiagen DNA Purification to remove proteins and other contaminants that inhibit PCR reactions: We quantified total extracted genomic DNA yields using a NanoDrop nucleic acid spectrophotometer (Thermo Scientific, Wilmington, DE).

For capillary sequencing, we amplified genomic DNA using haptophyte-specific oligonucleotide (Coolen *et al*., 2004; Simon *et al*., 2000) primers targeting 18S rRNA coding regions. Forward and reverse primers corresponded to *Escherichia coli* 18S rRNA positions 429 and 887, respectively. Polymerase chain reactions (PCRs) were performed on an Eppendorf Gradient Thermocycler (Hamburg, Germany) with the following conditions after D'Andrea *et al*. (2006): 4min initial denaturing at 96 °C, 35 cycles of denaturing for 30 s at 94 °C, followed by 40 s primer annealing at 55 °C and primer extension 40 s at 72 °C, with a final extension of 10 min at 72 °C. PCR reactions were run in 50µl volume using Promega GoTaq polymerase (Madison, WI, USA). Triplicate PCR products were pooled for each sample, and templates were purified using an Invitrogen PureLink Purification kit (Carlsbad, CA, USA) with the high cut-off binding buffer to eliminate fragments < 200 bp. Purified PCR products were A-tailed and purified using the PureLink high cut-off kit. Cloning was performed using the Invitrogen TOP10

cloning kit with electro- competent cells. Protocol was according to the manufacturer's instructions. One hundred clones were picked for each sample. Plasmid DNA was isolated using a RevPrep Orbit robotic template preparation instrument (Genomic Solutions, Ann Arbor, MI), and prepared templates were sequenced on an ABI 3730XL (Applied Biosystems, Foster City, CA) capillary sequencer using the BigDye protocol with universal M13 forward and reverse primers according to the manufacturer's instructions. All sequencing was performed at the Marine Biological Laboratory W. M. Keck Ecological and Evolutionary Genetics Facility.

For Ion Torrent (IT) sequencing, we performed genomic DNA amplifications using eukaryotic-specific primers targeting the 18S rRNA gene variable V9 region (1380F, 1389F, 1510R) Amaral-Zettler *et al.,* 2009). We then ligated both barcodes and adapters to these V9 amplicons. We sequenced the amplicons using an Ion Torrent Personal Genome Machine (PGM) with bidirectional sequencing at the Life Technologies facility.

2.8 Bioinformatics

Ion Torrent DNA sequences reads were trimmed and screened for quality after Huse *et al.* (2007), including the paired end information as a measure of accuracy. We clustered the sequences at 97% similarity using SLP-PWAN (Huse 2010) into Operational Taxonomic Units (OTU). For each OTU, we selected a representative sequence to assign taxonomy. To assign taxonomy to the remaining quality-controlled sequences, we used the Global Alignment for Sequence Taxonomy (GAST) algorithm

(Huse *et al.,* 2008). Our dataset consisted of 1,018,154 sequence reads distributed across 3,750 OTUs.

To compare microbial OTU distributions across samples, we used PRIMER-E statistical software (Clarke and Gorley, 2006) to perform cluster analyses and multidimensional scaling analyses (MDS). Before analysis, we randomly resampled the data so that all samples had the same number of sequence reads (11,078) equivalent to the lowest read sequence number of a single sample (after removing datasets from 5/5 10m and 5/17 10m with <2000 sequences). We also transformed the data to presence/absence counts and created a distance matrix using the Jaccard index of similarity. Using this same randomly resampled and transformed data matrix, we performed a canonical correspondence analysis using CANOCO 4.5 (Microcomputer Power, Ithaca, NY, USA) (ter Braak and Similauer, 2002) to compare sample microbial communities with environmental geochemical gradients. We used GraphPad Prism version 6 software (La Jolla, CA, www.graphpad.com) for performing linear regressions and regression comparisons.

2.9 Quantitative polymerase chain reaction

Purified DNA extracts were also subjected to real-time quantitative polymerase chain reaction (qPCR) to gauge haptophyte DNA concentrations with depth and ensure that the sample selected for sequencing was at the point of highest haptophyte cell concentration in the water column. We performed the qPCR reaction using 18S rRNA gene haptophyte specific primers Prym-429F and Prym-887R (Coolen *et al.,* 2004). The qPCR reactions were run on an Applied Biosystems StepOnePlus^{TM} Real-Time PCR System (Foster City,

California), using a SYBR Green I assay. Each 20 μl reaction contained 7.2 μl of sterile water, 10 μl of KAPA SYBR® FAST Universal 2X qPCR Master Mix (Woburn, MA), 0.4 μl each of the forward and reverse primers (0.2μM) and 2 μl of template DNA. The qPCR cycling program was after Coolen *et al.* (2009) and consisted of 38 cycles of denaturation at 94°C for 30s, annealing at 62°C for 40s, primer extension at 72°C for 60s, a photo step of 80°C for 20s. We used between 10^1 and 10^6 copies (ten-fold dilution series) of linearized plasmids containing 18S rDNA of *Isochrysis galbana* CCMP1323 as the external standard to calibrate the copy numbers of haptophyte DNA in the BrayaSø water samples.

3. Results

3.1 Geochemical results

As expected, the Lake George water samples exhibited a gradual warming over the course of the three-month sampling period (**Figure 2**B) from late April to late July. Salinity reached a slight plateau on the first half of the summer, and a second peak in the latter half of the seasonal cycle (**Figure 2**C). Coincident with the second escalation of salinity we see the decline in C_{37} alkenone concentrations in the water column (Day 45, **Figure 2**A). The C_{37} alkenone concentrations exhibit a bimodal peak, with a preliminary rise and fall in alkenone concentrations from days 1-22, and again from days 25-70 (**Figure 2**A).

The three sampling depths generally exhibited similar trends in temperature, salinity and alkenone concentrations. As expected, 10m depth samples had lower temperatures than the upper waters at 0m and 5m depths. Alkenone concentrations

peaked at 10m depth slightly delayed after the peak at 0m and 5m depths occurring on Day 37. C_{37} alkenone concentrations reached 2.5mg/L at the highest point on Day 37. Previous to Day 37, nutrient concentrations were low, and at Day 37 reached concentrations where they remained the remainder of our sampling (**Figure 3**). The lowest concentrations of nitrate, nitrite, and phosphate on Day 24 corresponded to the lowest alkenone concentrations. We will refer to the two distinct peaks in alkenone concentrations as defined by early Peak I from days 1-22 and then later Peak II from days 24-70. Peak I C_{37} alkenones reached concentrations of 1000ng/L and Peak II reached concentrations of 2500 ng/L (**Figure 2**A).

3.2 In situ *alkenone distributions and* U^k_{37} *calibration*

The first peak in alkenone concentrations, Peak I, was characterized by both dominant tetraunsaturated $C_{37:4}$ alkenones and dominant $C_{37:3}$ alkenone peaks (**Figure 4**). Peak II 5m and 10m depths had a dominant $C_{37:4}$ alkenone signature that changed to a dominant $C_{37:3}$ signature after the maximum alkenone production peak at day 37 (**Figure 4**). Peak II also saw the appearance of the $C_{38:3}$ methyl ketones (**Figure 4**). Peak I alkenones were at their maximum concentration at Day 10 (**Figure 2**A) and Day 10 alkenone signatures were $C_{37:3}$ dominant at all depths. Peak II alkenones were at their maximum concentration at depths 0m and 5m at Day 37 (**Figure 2**A), both characterized by dominant $C_{37:4}$. At 10m depth, alkenone concentrations were highest at Day 42 and were also characterized by a dominant $C_{37:4}$ signature.

The Lake George *in situ* U^k_{37} calibration was defined by the equation:

$$\mathbf{U^k_{37} = 0.0255T - 0.697}$$

This calibration had an R^2 value of 0.712 and a root mean squared error of 3.7°C (**Figure 5**A). The largest scatter in the U^k_{37} calibration occurred at about 15°C, corresponding to Peak II Days 37-45 (**Figure 5**A). The *in situ* calibration is most similar to a previously reported U^k_{37} calibration from Greenland's lake BrayaSø, a modified *in situ* calibration (D'Andrea *et al., 2*011) that was developed by combining an *in situ* calibration with a calibration from a suite of German lakes (Zink *et al., 2*001). Like Lake George, BrayaSø sediments are also characterized by dominant $C_{37:4}$ alkenones (**Figure 5C**).

3.3 In situ *haptophyte abundances: qPCR and IT sequencing*

All samples between April 26th and June 16th (Days 1-52) were analyzed for haptophyte cell abundances using both qPCR and Ion Torrent Sequencing. This date range spans both Peak I and Peak II, although it does not include the long tail of alkenone concentrations after Peak II. The qPCR results give an idea of total haptophyte cell abundance, as inferred from haptophyte rRNA gene copy numbers. The Ion Torrent results describe the entire eukaryotic microbial community during the haptophyte bloom, and we pay special attention to just the haptophyte Ion Torrent reads.

The qPCR probe targets all haptophyte algae, and like the alkenone data, the qPCR data exhibited a bi-modal distribution (**Figure 6**). In contrast to the alkenone data, Peak I had greater haptophyte rRNA gene copy numbers, suggesting greater concentrations of haptophytes cells during Peak I, reaching concentrations of almost 3000 cells/ml in the surface waters on Day 10, the apex of Peak I. As for Peak II, the 0m depth sample reached highest haptophyte cell numbers at Day 42, slightly after the 0m alkenone

peak. Curiously, 5m and 10m depth qPCR results did not exhibit the same trends as those in the 0m samples and fluctuated between a few hundred haptophyte cells/ml.

The Ion Torrent data allowed us to more closely examine the haptophyte distributions in Lake George. After OTU clustering, there were two Isochrysidales haptophyte OTUs. After generating V9 haptophyte sequences from cultures of Hap-B (Table S1), we were able to identify which OTU represented Hap-A and Hap-B sequences. Hap-A, the inferred dominant $C_{37:4}$-alkenone producing haptophyte (Chapter 6), and Hap-B, the C38:3 methyl producing haptophyte (Toney *et al*., 2012) showed bimodal distributions similar to both the alkenone and qPCR data (**Figure 4**). At 0m depth, Hap-A peaked during alkenone Peak I at Day 10, the same day as the qPCR peak and the alkenone peaks. At 5m and 10m depths, Hap-A was again the more dominant haptophyte during Peak I. At 0m depth, Peak II was dominated by Hap-B OTU sequences, and peaked at Day 42 in accordance with qPCR results, whereas 5m and 10m depth samples peaked at Day 45. At 10m depth, the Hap-A and Hap-B trends were not bimodal, and Hap-A sequences actually peaked at Day 37 followed by Hap-B becoming dominant.

Our randomly resampled Ion Torrent dataset yielded 8172 haptophyte reads across 56 OTUs. If singletons are ignored, the haptophyte yield was 11 OTUs. The haptophyte OTUs group into 5 phyla, Prymnesiales, Phaeocystales, Pavlovales, Isochrysidales and Coccolithales. The haptophyte phyla distributions are shown in **Figure 7.** The Isochrysidales and Pavlovales alternate in abundance, and Prymnesiales is abundant in the early season samples. Coccolithales and Phaeocystales only have one Ion

Torrent read each during the seasonal cycle. Of the 1230 Isochrysidales total sequences, Hap-A and Hap-B comprise 1201 of these sequences (97.6%).

A cluster analysis of the Lake George bloom Ion Torrent sequence data revealed a clear separation between pre- and peak/post-peak bloom samples (**Figure 8**). The earliest sample, April 26th at 0m depth (LG1_0) was a clear outlier from the other samples, possibly a result of its high *Chlamydomonas* concentrations (4386 reads, Table S2) and high copepod concentrations (4030 reads, Table S2), a combination unique to only this early surface sample. A canonical correspondence analysis revealed that 70% of the microbial assemblage variation among the samples could be explained by temperature and salinity (**Figure 9**). While percent $C_{37:4}$ was used as an explanatory variable, it is not an environmental variable shaping community composition but instead a variable that clearly fluctuates with the changing microbial community. As expected, Hap-B, the haptophyte that appeared 'late' in the bloom event, was closely associated with warmer temperatures as it plots along the temperature axis (**Figure 9**). There were no clear associations between Hap-A occurrence and environmental variabilities.

3.4 Culture studies and U^k_{37} *calibrations*

Hap-A was grown in culture to determine its U^k_{37} calibration and was identified as a picoplankton haptophyte (Chapter 6). Hap-B (Strain #903) was also isolated from Lake George waters and grown in pure culture at 5,10,15,21 and 24°C. Culture Hap-B alkenones demonstrated the characteristic increase in saturated alkenone isomers with increasing temperatures (**Figure 10**A). Hap-B also had a dominant $C_{37:3}$ alkenone signature and the presence of trace levels of $C_{38:3}$ methyl ketones (**Figure 10**A). The U^k_{37}

calibration for Hap-B (**Figure 5**) was distinct from the Hap-A culture at the lower temperatures although their calibrations overlapped at higher temperatures. The Hap-B calibration is defined by the equation

$$\mathbf{U^{k}_{37} = 0.0279T - 0.5889}$$

with an R^2 value of 0.95563 and an RMSE of 1.7°C. Neither Hap-A nor Hap-B calibrations match that of the *in situ* Lake George calibration, although the Hap-A calibration has a very similar slope to the *in situ* Lake George U^{k}_{37} calibration. The Hap-B calibration and the Hap-A calibrations are statistically similar and can be explained by a single curve, as depicted in **Figure 5B**. The equation for the composite culture U^{k}_{37} calibration is

$$\mathbf{U^{k}_{37} = 0.0326T - 0.671}$$

The Lake George *in situ* calibration from our study is with an R2 value of 0.92 and RMSE of 1.7°C. Hap-B displayed a notable change in morphology at different temperatures (**Figure 10**B), and the highest growth rates were at 15°C and the lowest growth rates at 5°C (**Table 1**). Alkenone per cell concentrations ranged from 0.011 pg/cell at 15°C to 3.299 pg/cell at 5°C.

3.5 Mock bloom study

To determine if we could recreate the Lake George bloom event *in silico*, we performed a mock bloom event by incubating Lake George surface sediments and monitoring the supernatant alkenone concentrations and haptophyte DNA over time. Like observed in the environmental samples, the mock bloom event was characterized by Hap-

A dominance in the early weeks and the appearance of Hap-B later in the bloom cycle (**Figure 11**). Unlike our field samples, the mock-bloom event alkenone signatures were all $C_{37:4}$ dominant, although the late-bloom samples had unresolvable alkenone peaks. The timing of the appearance of Hap-B in the mock bloom event was very similar to that observed in the field, with Hap-B DNA first detected at Day 42 (versus *in situ* Hap-B rises to dominance in the surface waters at Day 37).

4. Discussion

4.1 Peak I and Peak II haptophyte communities

Alkenone concentrations in the Lake George water column defined Peak I as spanning from Day 1-22 of the sampling effort, and a larger Peak II spanning from Days 24-70. This dual-bloom cycle was confirmed by qPCR analyses, however the qPCR analyses suggested that Peak I was actually a larger, more haptophyte-dense bloom event than Peak II. These two findings are not mutually exclusive, considering alkenone concentrations per cell vary and Peak I may have larger cell numbers with lower alkenone concentrations followed by Peak II with lower haptophyte cell numbers and higher alkenone cell concentrations. Alkenone concentrations vary with cell stage or nutrient concentrations, with cells accumulating alkenones at both lower growth rates (Popp *et al*., 1998; Epstein *et al*., 2001) and low nutrient conditions (Eltgroth *et al*., 2005). There may also be genetic predispositions of alkenone production, with different haptophyte strains synthesizing more or less alkenones under identical environmental and physiological conditions (Conte *et al*., 1998). However, it is important to remember that

the qPCR data reflect total haptophyte cell abundances, as opposed to alkenones that are unique to Isochrysidales.

In order to delineate total haptophyte cell numbers from alkenone-producing haptophytes, we refer to the Ion Torrent sequence read abundances. **Figure 7** displays the abundance and phylogenetic affiliation of haptophyte OTUs. We can see that like the qPCR results, the Ion Torrent results display a higher number of haptophyte sequences in Peak I, equaling 5663 sequence tags, and Peak II has only 2506 haptophyte sequence tags over twice as many days. Isochrysidales are almost equal between Peak I and Peak II, 293 and 301 sequence reads, respectively. These results suggest that haptophyte cell abundances are indeed higher during Peak I, as indicated by qPCR results and Ion Torrent sequencing results, however Isochrysidales cell numbers are almost equal during both peaks. The variation in alkenone concentrations, therefore, between Peak I and Peak II is likely a reflection of haptophyte taxonomy or cell physiology.

4.2 Environmental bloom conditions

The bimodal alkenone distribution depicts a preliminary haptophyte bloom while nutrient concentrations are low in the spring followed by a second haptophyte bloom coinciding with higher nutrient concentrations in the late spring/early summer. The onset of Peak II coincides with an increase and stabilization of nutrient concentrations in the water column and stable salinity values. Ion Torrent sequencing revealed the dominant haptophyte during Peak I is Hap-A and the dominant haptophyte during Peak II is Hap B. Our canonical correspondence analysis revealed the association of Hap-A with lower

salinities, and the association of Hap-B with higher temperatures and salinities, although this may just be the result of the Hap-B bloom occurring later in the bloom season.

We know from previous research (Chapter 5) that Hap-A is a picoplankter, indicating it may have a competitive advantage in low-nutrient conditions due to its higher surface-area to volume ratio. The early season Peak I water column, when the seasonal ice melt is still mixing throughout the lake's water column, may allow picoplanktonic Hap-A to outcompete larger, and more nutrient-dependent Hap-B. While temperature and salinity explained 70% of the total eukaryotic OTU variance, it remains to be seen if warmer temperatures and higher salinities are requirements for a Hap-B bloom.

4.3 Hap-A and Hap-B alkenone production

As suggested in a previous study, and confirmed in this one, Hap-B is a dominant $C_{37:3}$ alkenone producer, similar to its close relatives *Pseudoisochrysis paradoxa* and *Chrysotila lamellosa* (Toney *et al*., 2012; Theroux *et al*., 2010). DNA sequences placed Hap-A as a close relative of an uncultured haptophyte from Ace Lake, Antarctica (Coolen *et al*., 2004), also a presumed dominant $C_{37:4}$ alkenone producers. Our mock-bloom results suggest that when Hap-A is dominant, the alkenone signature is $C_{37:4}$ dominant (**Figure 11**). Hap-B, in contrast, produced only $C_{37:3}$-dominant signatures in culture (**Table 1**). Our *in situ* alkenone lipid signatures from Lake George did not reflect this clear separation between Hap-A dominance and Hap-B dominance—instead, during both Peak I and Peak II there were alternating $C_{37:4}$- and $C_{37:3}$-dominant alkenones. The culture studies indicate that Hap-A and Hap-B do have distinct alkenone lipid signatures at lower

temperatures, with Hap-A producing abundant $C_{37:4}$ signatures and Hap-B producing only $C_{37:3}$-dominant alkenone signatures.

As suggested from previous culture studies, Hap-A may lose its $C_{37:4}$ dominant alkenone signature when at higher growth rates as experienced in pure culture (Chapter 6) and at certain times in the bloom cycle. If we examine the Ion Torrent Hap-A and Hap-B abundances more closely (Figure 4), we see this pattern exhibited in the 0m depth samples. When Hap-A is at low growth rates early in the season, the alkenone signature is characterized by an abundant $C_{37:4}$ alkenone. When Hap-A reaches it peak cell concentrations, at Day 10, the alkenone signature is characterized by dominant $C_{38:3}$ alkenones. Again, when Hap-A concentrations drop at Day 22, the alkenone signature is $C_{37:4}$ dominant. This pattern is not maintained when Hap-B appears in the water column, at which point there is a mixture of competing alkenone fingerprints as indicated by abundant $C_{37:4}$ alkenones co-occurring with $C_{38:3}$ methyl ketones.

4.4 U^k_{37} *alkenone unsaturation indices*

Genotypic effects on alkenone production are evident in the individual alkenone signatures for Hap-A and Hap-B cultures. However, Hap-A and Hap-B U^k_{37} calibrations could be explained by a single composite calibration (**Figure 5**B). The higher U^k_{37} values derived from haptophytes cultures as opposed to the *in situ* measurements may be the result of artificially high growth rates experienced by these cells in culture. The *in situ* U^k_{37} calibration for Lake George is closely aligned with previous calibrations from lake BrayaSø, Greenland (D'Andrea *et al*., 2011) and from a series of German lakes (Zink *et al*., 2001).

The Hap-A and Hap-B cultures had similar growth rates, with doubling times average 3-4 days, with the exception of the 5°C Hap-B culture with very low growth rates and very high alkenone concentrations (**Table 1**). The cultures have similar U^k_{37} values for 15°C, which is curious considering this is the temperature at which the *in situ* U^k_{37} calibration has the highest degree of scatter (**Figure 5**A). This may indicate the samples collected on the 15°C days, which center around Days 37-55, may have been during high wind, high mixing conditions leading to transport of alkenones throughout the water column and erroneous U^k_{37} estimations. Indeed, Days 45 and 52 have overlapping temperature measurements for 0m and 5m depths, indicating mixing of the surface waters. Future *in situ* temperature estimates should use caution when interpreting alkenone concentrations from days with low stratification indices.

Both Ion Torrent sequencing and qPCR results showed much lower Isochrysidales cell concentrations at 5m and 10m depths in the water column. Ion Torrent Isochrysidales sequences averaged 48 reads in the 0m samples, 17 reads at 5m depth, and 12 reads at 10m depth. We recalculated the U^k_{37} for each individual depth (Figure S1) and 0m, 5m, and 10m samples all had an increasing U^k_{37} with increasing temperature, suggesting alkenones collected at depth still record a temperature response. Both qPCR and Ion Torrent data indicate a higher haptophyte cell concentration at 0m depth, and alkenone concentrations grade slightly from 0m > 5m > 10m, suggesting the majority of alkenones being produced in the Lake George water column are from 0m depth.

4.5 Implications for paleothermometry

The majority of Lake George downcore sediments are $C_{37:4}$ dominant, with a few exceptions when alkenone signatures have dominant $C_{37:3}$ and the presence of the $C_{38:3}$ methyl ketone, a hallmark of the Hap-B cultures (Toney, 2011; Toney *et al*., 2012). These $C_{37:3}$ dominant signatures may be the result of environmental conditions favoring Hap-B, such as warmer water temperatures, higher salinities or higher nutrient concentrations as suggested by our field observations.

Neither the Hap-A nor the Hap-B *ex situ* calibrations yielded an exact match to the Lake George *in situ* calibration, likely because pure culture conditions do not accurately replicate those experienced in the field. Nutrient cycling, variable temperature regimes, and predator/prey fluctuations are not captured in pure culture studies, and may be impacting the *in situ* alkenone signatures. With this in mind, we believe the *in situ* calibration most accurately captures the haptophyte response to water temperature in Lake George and the *in situ* calibration would be the most reliable metric for reconstructing temperatures from the Lake George alkenone record.

5. Conclusions

Our study of a seasonal haptophyte bloom in Lake George, ND provides new insights into lacustrine haptophyte physiology, ecology and bloom conditions. Throughout a seasonal cycle, the two species of haptophytes in Lake George alternate in abundance, as first indicated by *in situ* alkenone concentrations. DNA sequencing revealed that the bimodal alkenone distribution reflected alternating abundances of species Hap-A and Hap-B. Hap-A was dominant in the early bloom stage Peak I, and Hap-B was dominant during subsequent Peak II. As evidenced by both field observations

and our Hap-B culture manipulations, we have confirmed that Hap-B is responsible for dominant $C_{37:3}$ and $C_{38:3}$ methyl ketone production and Hap-A is the dominant tetraunsaturated alkenone producer. The dual peak in alkenone concentrations throughout the course of the seasonal cycle was again reflected in qPCR haptophyte gene copy numbers and the Hap-A to Hap-B succession was also observed in our mock-bloom enrichment culture experiment. These complimentary results demonstrate how our multi-disciplinary approach enabled a greater resolution of haptophyte cell abundances than was previously possible through geochemical analyses alone.

The environmental triggers for Hap-A and Hap-B dominance are still unclear. While Hap-B appears during warmer temperatures and higher salinities, it remains to be seen if these conditions are a prerequisite for its bloom. Efforts are currently underway to determine if, in a mock-bloom experiment, a warmer temperature incubation results in a preliminary bloom by Hap-B instead of Hap-A. Clearly, the interplay between the length of Peak I and Peak II, and the abundance of individual Hap-A and Hap-B, results in variable inputs of alkenones into the sediment record. It remains to be seen if Hap-A alkenones are preferentially exported to the Lake George sediments, given the characteristic $C_{37:4}$ dominant peak in the sediment record. Although futher research is needed, we believe the *in situ* U^k_{37} calibration provides an accurate representation of the relationship between alkenone-unsaturation and lake water temperature and therefor provides a means of reconstructing temperature from the Lake George sediment record.

Acknowledgements

This work was supported by a National Science Foundation award to Y. Huang (EAR-1122749) and L. Amaral-Zettler (EAR-1124192), a Brown SEED fund to Y. Huang and L. Amaral-Zettler, and an American Association of University Women dissertation fellowship to S. Theroux.

References

Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM. (2009). A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS One* **4**:e6372.

Brassell SC, Eglinton G, Marlowe IT, Pflaumann U, Sarnthein M. (1986). Molecular stratigraphy - a new tool for climatic assessment. *Nature* **320**:129-133.

Chu GQ, Sun Q, Li SQ, Zheng MP, Jia XX, Lu CF, Liu JQ, Liu TS. (2005). Long-chain alkenone distributions and temperature dependence in lacustrine surface sediments from China. *Geochimica Et Cosmochimica Acta* **69**:4985-5003.

Clarke KR, Gorley RN. (2006). PRIMER v6: User Manual/Tutorial.

Conte MH, Thompson A, Lesley D, Harris RP. (1998). Genetic and physiological influences on the alkenone/alkenoate versus growth temperature relationship in *Emiliania huxleyi* and *Gephyrocapsa oceanica*. *Geochimica Et Cosmochimica Acta* **62**:51-68.

Coolen MJL, Muyzer G, Rijpstra WIC, Schouten S, Volkman JK, Sinninghe Damsté J. (2004). Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake. *Earth and Planetary Science Letters* **223**:225-239.

Coolen MJL, Saenz JP, Giosan L, Trowbridge NY, Dimitrov P, Dimitrov D, Eglinton TI. (2009). DNA and lipid molecular stratigraphic records of haptophyte succession in the Black Sea during the Holocene. *Earth and Planetary Science Letters* **284**:610-621.

D'Andrea WJ, Huang Y. (2005). Long chain alkenones in Greenland lake sediments: Low δ13C values and exceptional abundance. *Organic Geochemistry* **36**:1234-1241.

D'Andrea WJ, Lage M, Martiny JBH, Laatsch AD, Amaral-Zettler LA, Sogin ML, Huang YS. (2006). Alkenone producers inferred from well-preserved 18S rDNA in Greenland lake sediments. *Journal of Geophysical Research-Biogeosciences* **111**:doi:10.1029/2005JG000121.

D'Andrea WJ, Huang Y, Fritz SC, Anderson NJ. (2011). Abrupt Holocene climate change as an important factor for human migration in West Greenland. *Proceedings of the National Academy of Sciences*

Eltgroth ML, Watwood RL, Wolfe GV. (2005). Production and cellular localization of neutral long-chain lipids in the haptophyte algae Isochrysis galbana and Emiliania huxleyi. *Journal of Phycology* **41**:1000-1009.

Epstein BL, D'Hondt S, Hargraves PE. (2001). The possible metabolic role of C-37 alkenones in Emiliania huxleyi. *Organic Geochemistry* **32**:867-875.

Fritz SC. (2011). Lake George and its geological history during the last ~10,000 years. *CGREC Annual Report*

Guillard RRL (1975) Culture of phytoplankton for feeding marine invertebrates. In *Culture of marine invertebrate animals*. WL Smith, MH Chanley, eds. Plenum Press, New York, p. 26.

Huse SM, Mark Welch D, Morrison HG, Sogin ML. (2010). Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environmental Microbiology* 1889-1898.

Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM. (2007). Accuracy and quality of massively parallel DNA pyrosequencing. *Genome biology* **8**:R143.

Huse SM, Huber JA, Welch DM, Relman DA, Sogin ML. (2008). Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS genetics* **4**:e1000255.

Liu WG, Liu ZH, Fu MY, An ZH. (2008). Distribution of the C37 tetra-unsaturated alkenone in Lake Qinghai, China: A potential lake salinity indicator. *Geochimica Et Cosmochimica Acta* **72**:988-997.

Müller PJ, Kirst G, Ruhland G, von Storch I, Rosell-Mele A. (1998). Calibration of the alkenone paleotemperature index $U^{k'}_{37}$ based on core-tops from the eastern South Atlantic and the global ocean (60°N-60°S). *Geochimica Et Cosmochimica Acta* **62**:1757-1772.

Pearson EJ, Juggins S, Farrimond P. (2008). Distribution and significance of long-chain alkenones as salinity and temperature indicators in Spanish saline lake sediments. *Geochimica Et Cosmochimica Acta* **72**:4035-4046.

Popp BN, Kenig F, Wakeham SG, Laws EA, Bidigare RR. (1998). Does Growth Rate Affect Ketone Unsaturation and Intracellular Carbon Isotopic Variability in Emiliania Huxleyi? *Paleoceanography* **13**:35-41.

Prahl FG, Wakeham SG. (1987). Calibration of unsaturation patterns in long-chain ketone compositions for paleotemperature assessment. *Nature* **330**:367-369.

Rontani JF, Prahl FG, Volkman JK. (2006). Re-examination of the double bond positions in alkenones and derivatives: biosynthetic implications. *Journal of Phycology* **42**:800-813.

Sikes EL, Volkman JK. (1993). Calibration of alkenone unsaturation ratios ($U^{k'}_{37}$) for paleotemperature estimation in cold polar waters. *Geochimica et Cosmochimica Acta* **57**:1883-1889.

Simon N, Campbell L, Örnolfsdottir E, Groben R, Guillou L, Lange M, Medlin LK. (2000). Oligonucleotide probes for the identification of three algal groups by dot blot and fluorescent whole-cell hybridization. *The Journal of Eukaryotic Microbiology* **47**:76-84.

Sun Q, Chu GQ, Liu GX, Li S, Wang XH. (2007). Calibration of alkenone unsaturation index with growth temperature for a lacustrine species, *Chrysotila lamellosa* (Haptophyceae). *Organic Geochemistry* **38**:1226-1234.

Ter Braak CJF, Smilauer P. (2002). CANOCO reference manual and CanoDraw for Windows user's guide: Software for canonical community ordination (version 4.5). Microcomputer Power, Ithaca, NY.

Theroux SM, D'Andrea WJ, Toney JL, Amaral-Zettler LA, Huang Y. (2010). Phylogenetic diversity and evolutionary relatedness of alkenone producing haptophyte algae in lakes: Implications for continental paleotemperature reconstructions. *Earth and Planetary Science Letters* **300**:311-320.

Toney JL, Huang Y, Fritz SC, Baker PA, Grimm E, Nyren P. (2010). Climatic and environmental controls on the occurrence and distributions of long chain alkenones in lakes in the interior United States. *Geochimica Et Cosmochimica Acta*

Toney, JL. (2011). Climatic and environmental controls on the occurrence and distributions of long-chain alkenones in lakes of the interior United States. PhD thesis, Brown University. 351pgs.

Toney JL, Theroux S, Andersen RA, Coleman A, Amaral-Zettle L, Huang Y. (2012). Culturing of the first 37:4 predominant lacustrine haptophyte: Geochemical, biochemical, and genetic implications. *Geochimica et Cosmochimica Acta* **78**:51 - 64.

Tyrrell T, A Merico (2004). Emiliania huxleyi: bloom observations and the conditions that induce them In *Coccolithophores-from molecular processes to global impact*. HR Thierstein, J Young, eds. Springer, p. 75.

Whitehead RL. (1996). Groundwater atlas of the United States - Segment 8. *USGS*

Zink KG, Leythaeuser D, Melkonian M, Schwark L. (2001). Temperature dependency of long-chain alkenone distributions in Recent to fossil limnic sediments and in lake waters. *Geochimica Et Cosmochimica Acta* **65**:253-265.

Table 1. Haptophyte Hap-B #903 culture cell counts and alkenone concentrations.

Temp. (°C)	Final cell concentration (cell/ml)	Growth rate (Doubling time)	C37:4 (ng/L)	C37:3 (ng/L)	C37:2 (ng/L)	C37total (ng/L)	% C37:4	C37alk/cell (pg)	Uk37	Uk'37
5	10000	71.44	16363	16375	251	32989	49.6	3.299	0.02	-0.49
10	640000	3.64	8831	13779	1515	24125	36.6	0.038	0.10	-0.30
15	1600000	3.009	3583	12855	1951	18390	19.5	0.011	0.13	-0.09
21	250000	4.632	1861	15067	884	17812	10.4	0.071	0.06	-0.05
24	310000	4.359	367	2649	640	3656	10.0	0.012	0.19	0.07

Figure 1. Map of North Dakota (inset) and Lake George, ND.

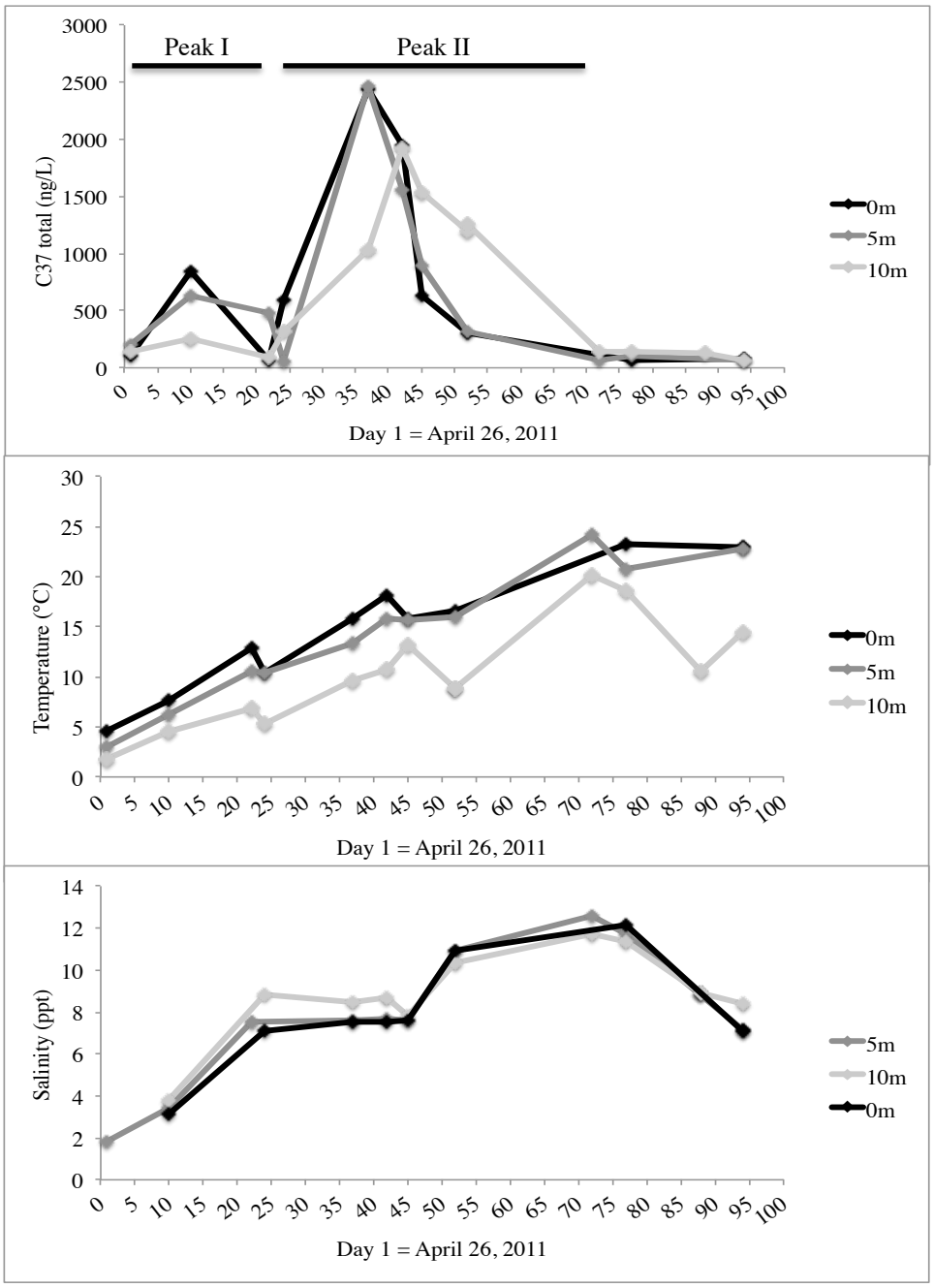

Figure 2. Lake George water column throughout the sampling period at 0m, 5m, and 10m depths. A. C_{37}-alkenone concentrations. B. Temperature C. Salinity. Peaks I and II as indicated.

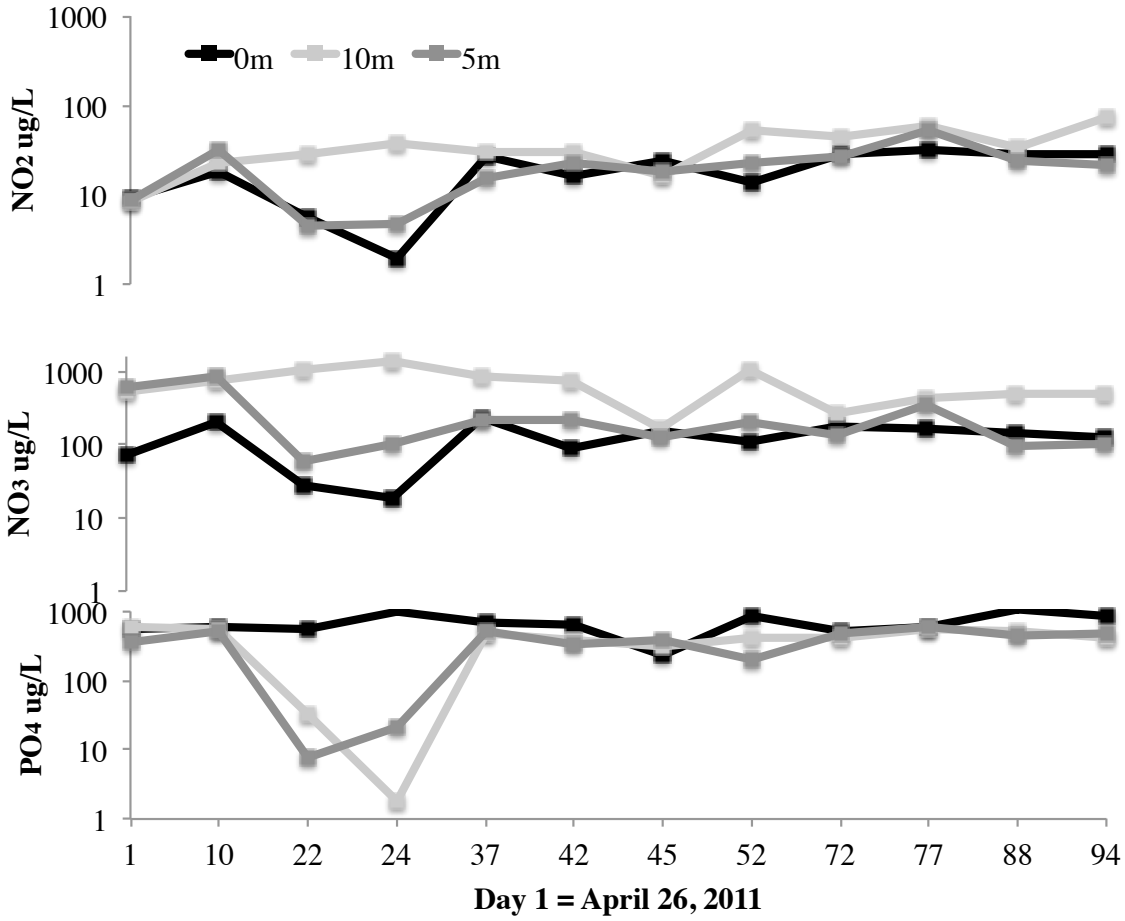

Figure 3. Nutrient concentrations during the Lake George seasonal cycle.

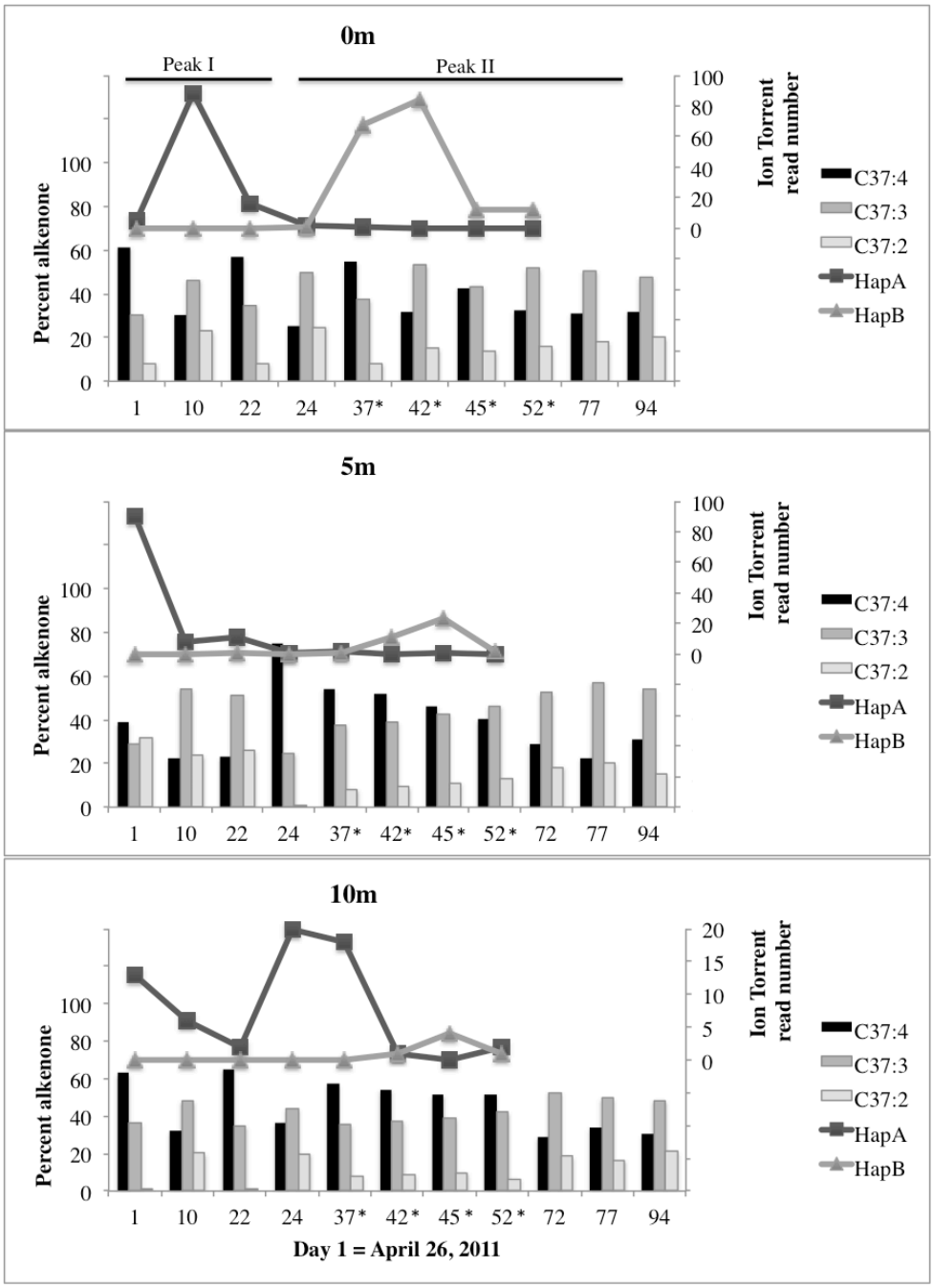

Figure 4. Percent concentrations of alkenone isomers and Ion Torrent sequence read numbers of Hap-A and Hap-B throughout the seasonal cycle at 0m, 5m, and 10m depths. Alkenone concentration Peaks I and II as indicated. Asterisk denotes a sample with $C_{38:3}$ methyl ketones present.

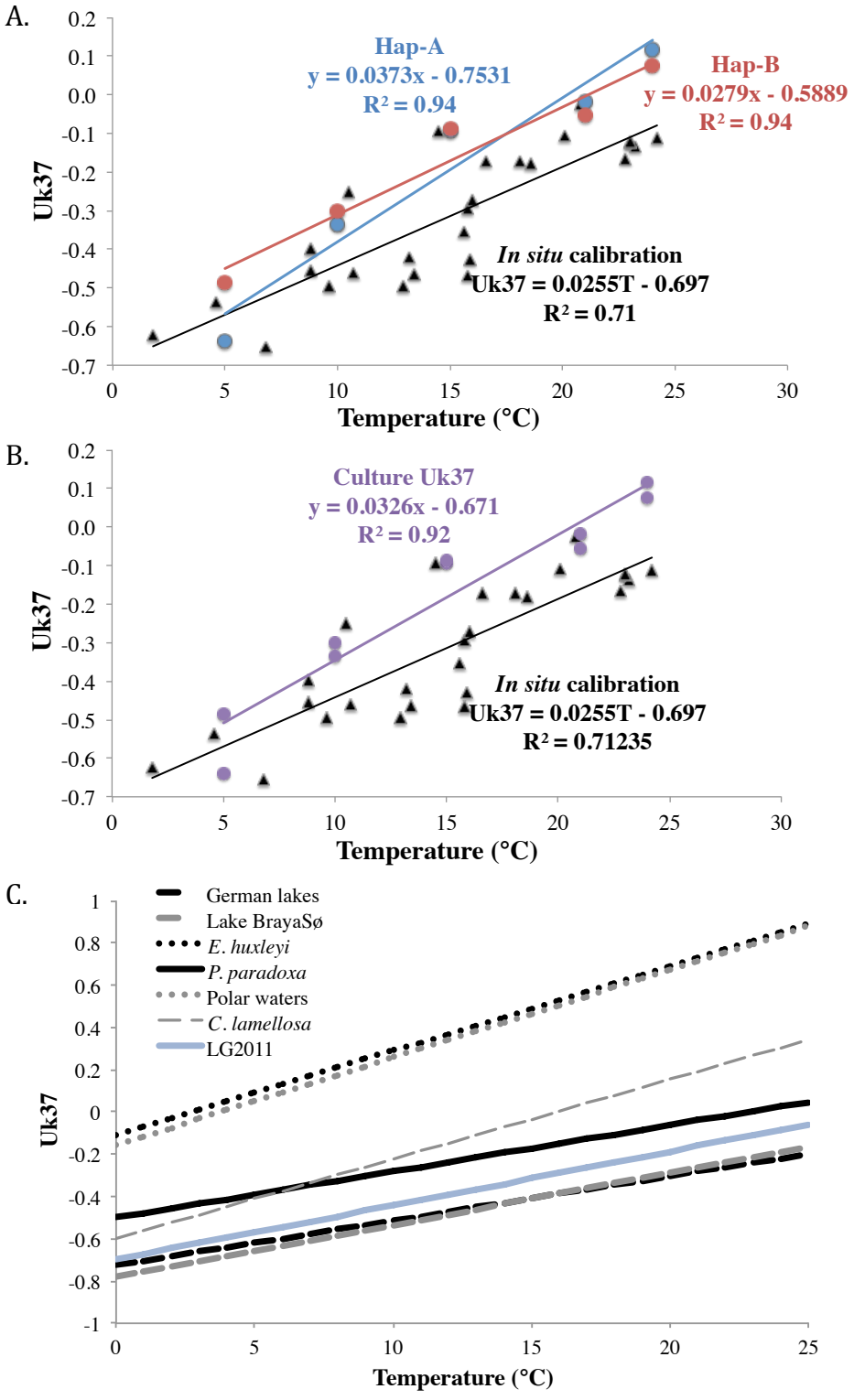

Figure 5. U^k_{37} calibrations for the Lake George field and culture alkenones. A. U^k_{37} calibrations for Lake George 2011 *in situ* calibration, Hap-B isolate #903, and Hap-

A <3μm culture. *In situ* calibration 2011 from this work. B. Composite "culture U^k_{37} " calibration from Hap-A and Hap-B calibrations plotted with *in situ* calibration. C. Lake George (LG2011) *in situ* calibration compared to previously published U^k_{37} calibrations. References are as follows: German Lakes, Zink *et al.,* 2001; Lake George *in situ*, Toney *et al.,* 2010; Lake BrayaSø, D'Andrea *et al.,* 2011; *E. huxleyi*, Prahl *et al.,* 1988; P. paradoxa, Chapter 4; Polar waters, Sikes and Volkman 1993; Lake George *in situ* 2011, Theroux, this work; *C. lamellosa*, Sun *et al.,* 2007.

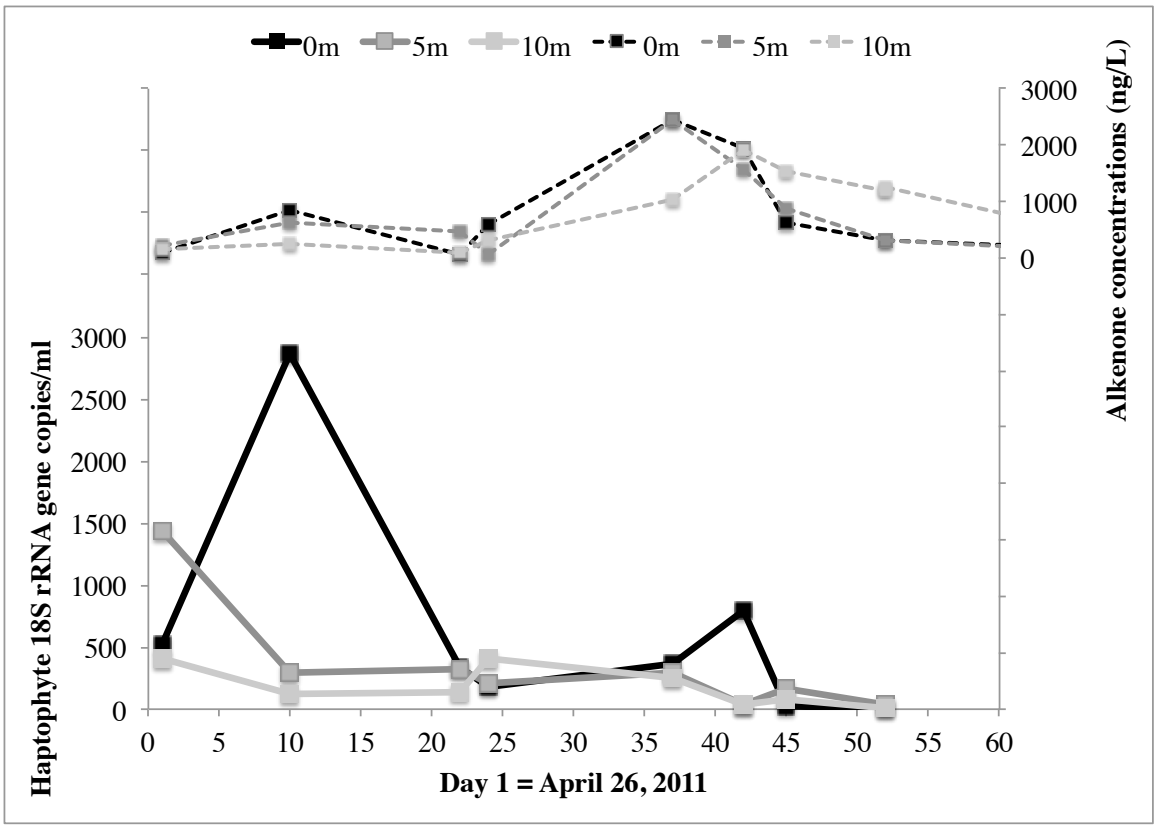

Figure 6. qPCR haptophyte gene copy number concentrations throughout the course of the Lake George seasonal cycle. Alkenone concentrations as a comparison.

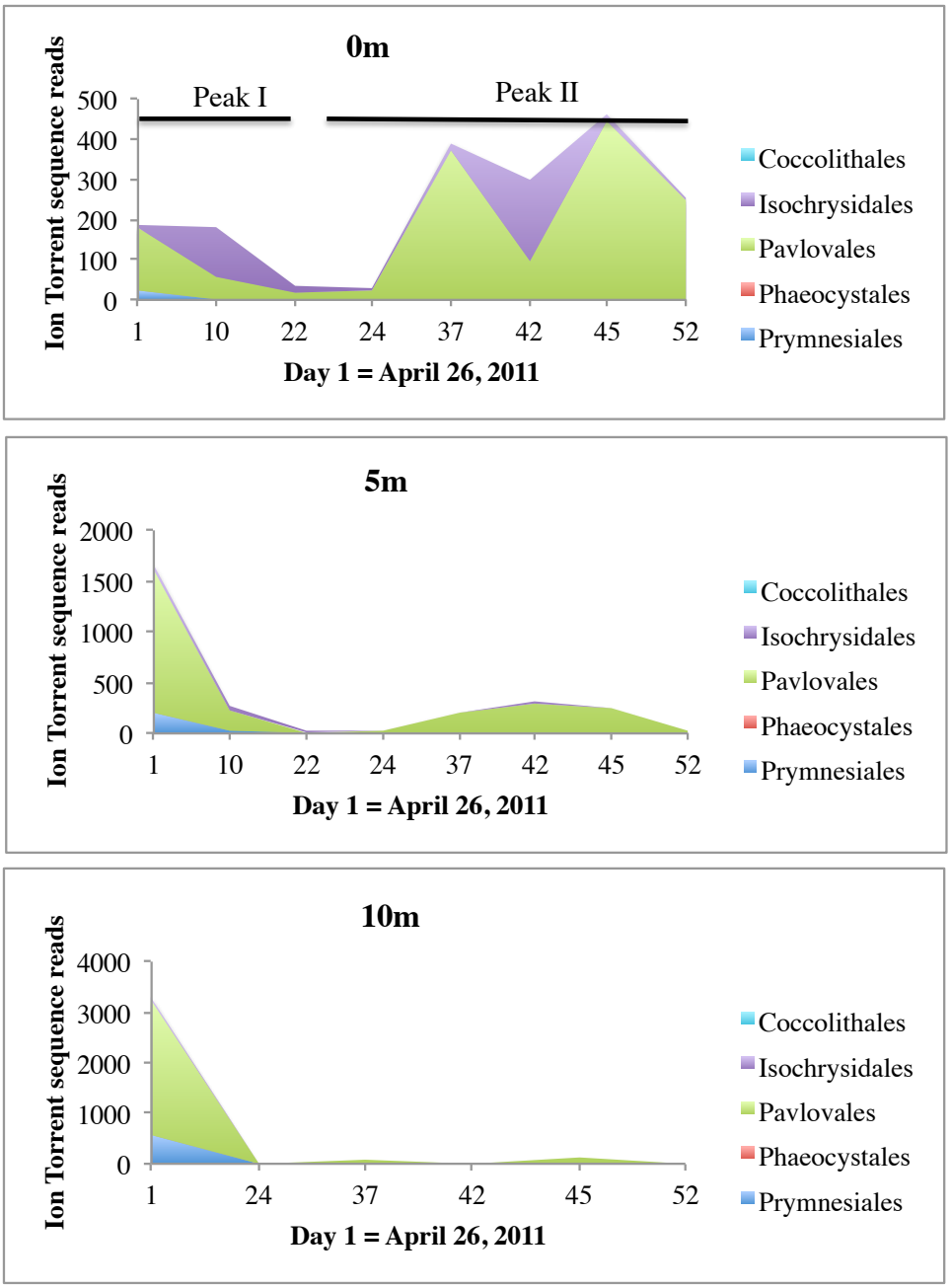

Figure 7. Haptophyte phyla abundances from Ion Torrent OTUs at 0m, 5m, and 10m depths. Peak I and Peak II as indicated. OTUs are defined at 97% similarity and the dataset was randomly resampled to yield equal OTU numbers per sample.

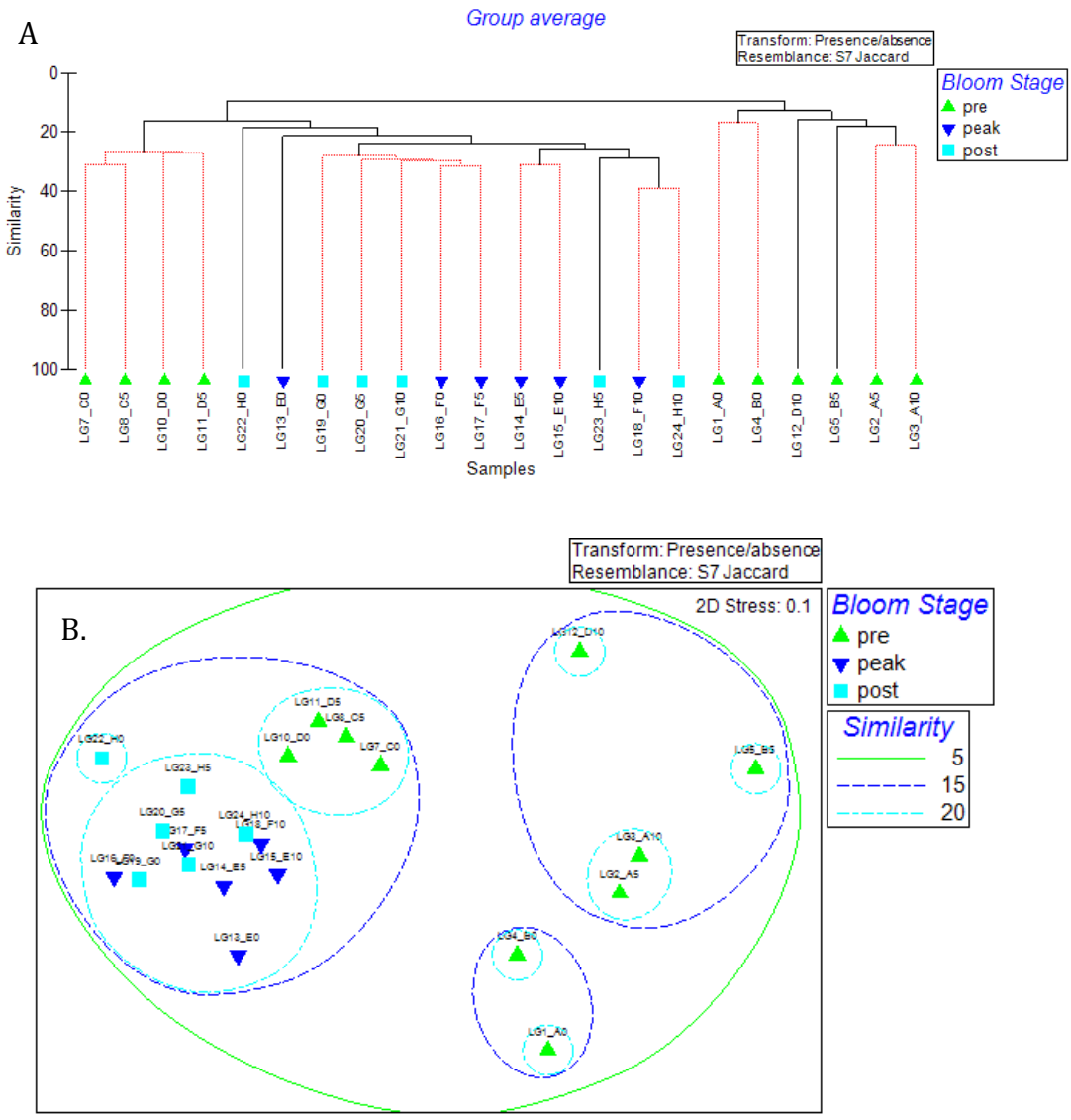

Figure 8. A. Cluster analysis of Lake George Ion Torrent microbial community similarity among samples. All sample OTU abundances were first transformed into presence/absence data and then Bray-Curtis similarities were computed before analysis. Bloom stage is overlain and defined as Pre-bloom (Days 1-22), Peak (Days 24-42), Post (Day 45-52). B. Multi-dimensional scaling analysis as in (A). All sample OTU abundances were first transformed into presence/absence data and then

Jaccard similarities were computed before analysis. Bloom stage is overlain and defined as Pre-bloom (Days 1-22), Peak (Days 24-42), Post (Day 45-52).

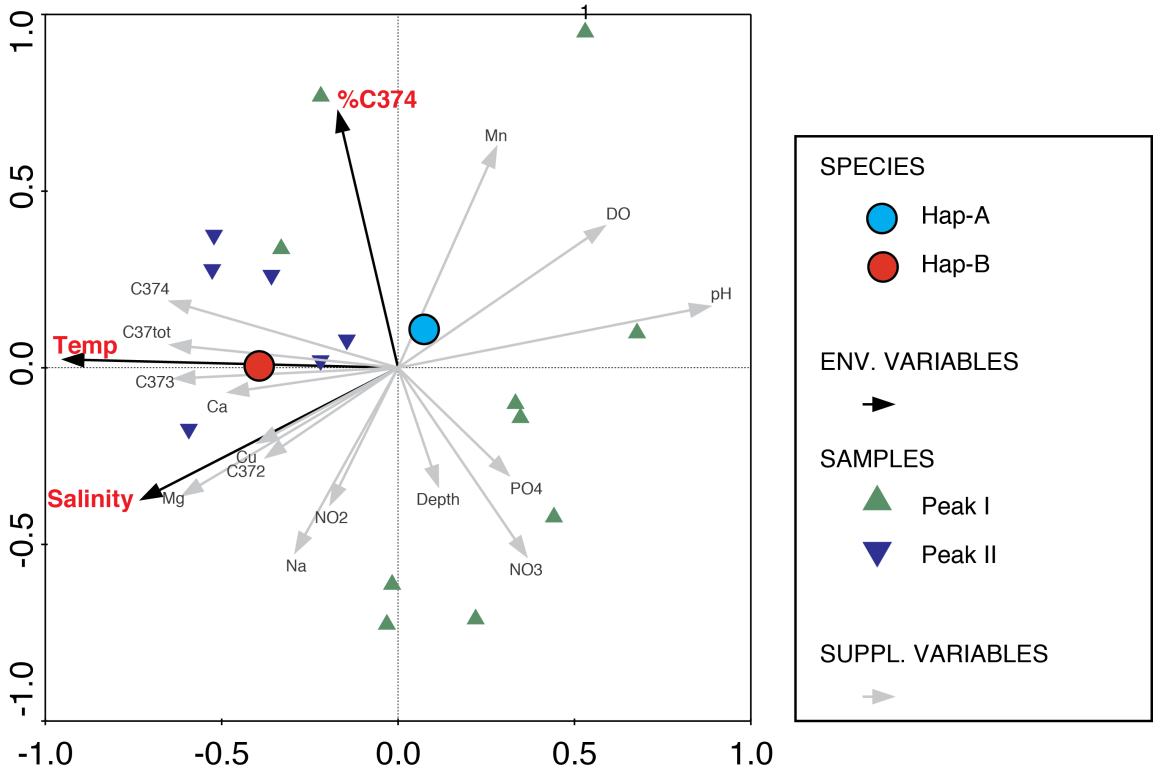

Figure 9. Canonical correspondence analysis of Lake George geochemical gradients and OTU abundances. All OTUs were randomly resampled, transformed to presence/absence and Jaccard distances were computed before analysis. Hap-A and Hap-B are indicated on the graph. Fist and second axis are defined by temperature and percent $C_{37:4}$, respectively, explaining 70% of the OTU variance. Salinity is the third explanatory variable and the rest of the geochemical data are supplemental variables. Peak I is defined as pre-bloom, and Peak II by peak-bloom and post-bloom.

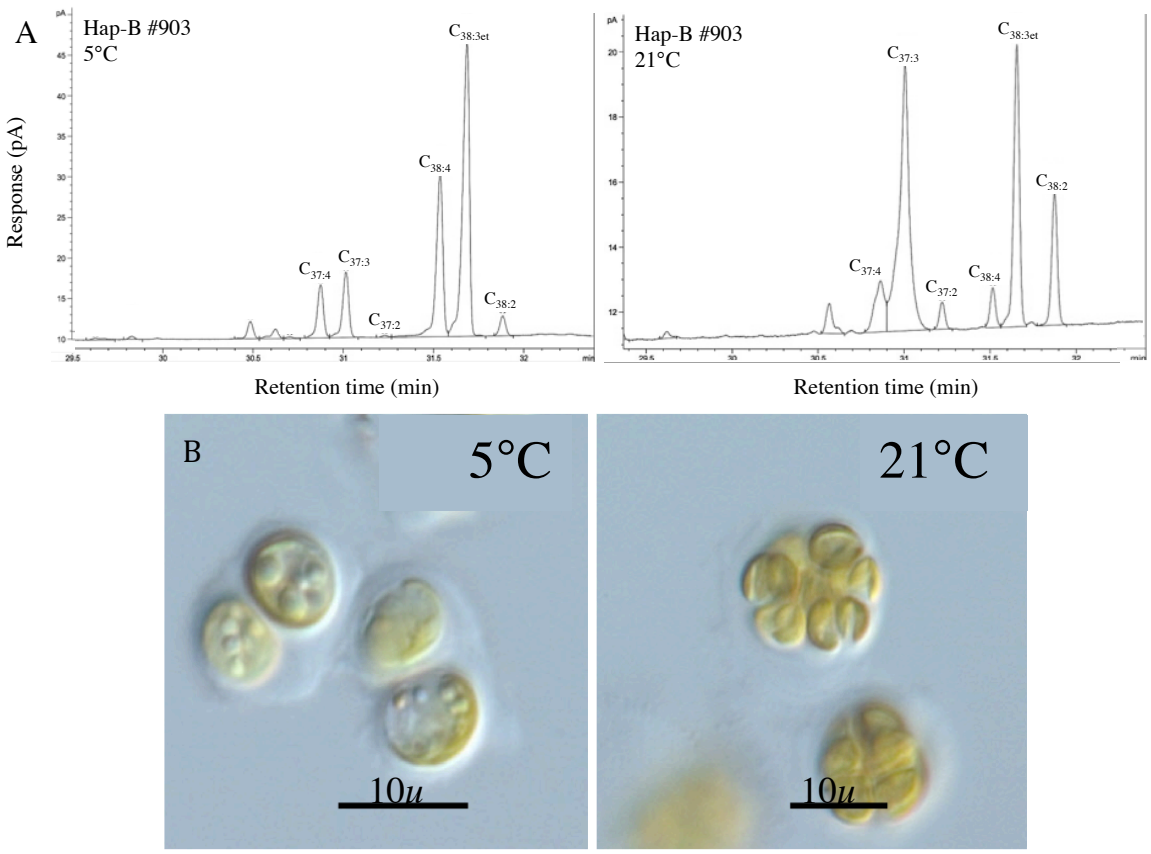

Figure 10. Hap-B strain #903 alkenones and morphology. A. Gas chromatogram of Hap-B #903 alkenone signature at 5°C and 21°C. B. Hap-B cells. Hap-B is shown at 5°C and 21°C to illustrate the cell morphology change with temperature.

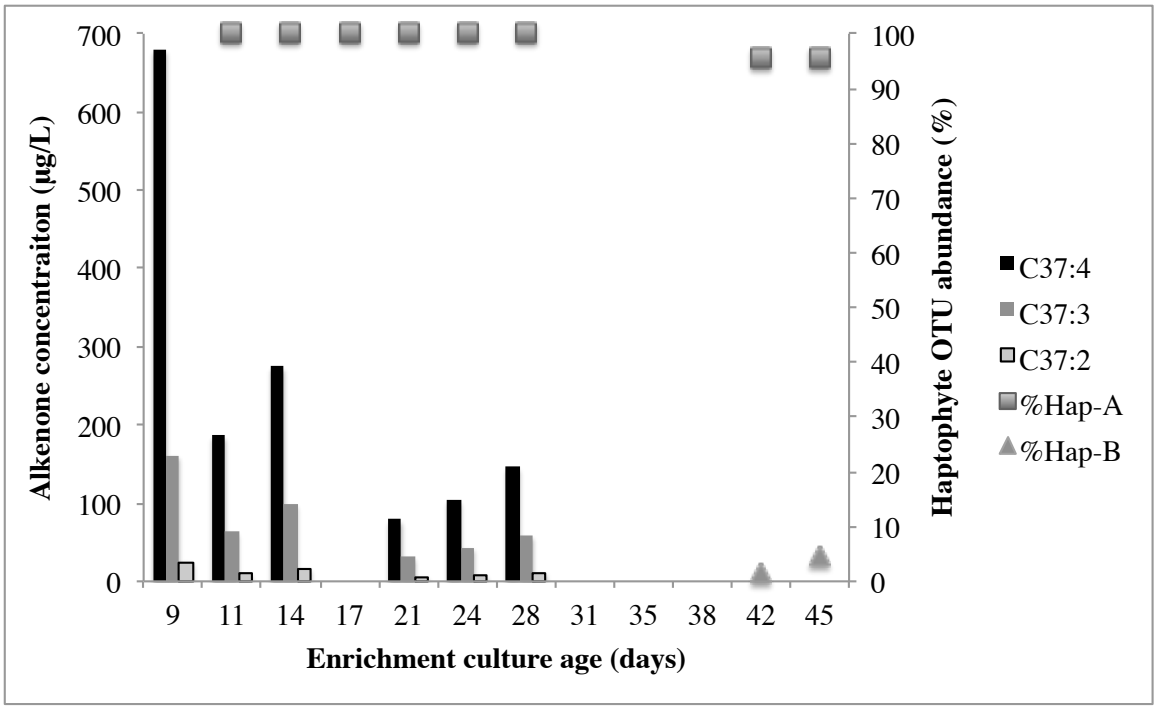

Figure 11. Mock bloom event alkenone signatures and haptophyte DNA sequence taxonomy.

Supplemental

Table S1. Hap-B culture 18S rRNA partial gene sequence. Generated using primers 429-Forward/EukB-Reverse.

>HapB
GCGCGTAAATTGCCCGAATCCTGACACATTGAGGTAGTGACAAGAAATAA
CAATACAGGGCTCTTCGAGTCTTGTAATTGGAATGAGTACAATTTACATC
TCTTCACGAGGATCAATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTA
ATTCCAGCTCCAATAGCGTATACTAAAGTTGTTGCAGTTAAAACGCTCGT
AGTCGGATTTCGGGGCGGGCCCGCCGGTCTGCCGATGGGTATGCACTGGC
GGGCGCGTCCTTCCTCCCGGAGACGGCTGCTACTCCTAACTGAGCGGTGG
CCGGAGACGGGATATTTACTTTGAAAAAATCAGAGTGTTTCAAGCAGGCA
GTCGCTCTTGCATGGATTAGCATGGGATAATGAAATAGGACTCTGGTGCT
ATTTTGTTGGTTTCGAGCACCGGAGTAATGATTAACAGGGACAGTCAGGG
GCACTCGTATTCCGCCGAGAGAGGTGAAATTCTCAGACCAGCGGAAGACG
AACCACTGCGAAAGCATTTGCCAGGGATGTTTTCACTGATCAAGAACGAA
AGTTAGGGGATCGAAGACGATCAGATACCGTCGTAGTCTTAACCATAAAC
CATGCCGACTAGGGATTGGAGGATGTTCCATTTGTGACTCCTTCAGCACC
TTTCGGGAAACTAAAGTCTTTGGGTTCCGGGGGGAGTATGGTCGCAAGGC
TGAAACTTAAAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCG
GCTTAATTTGACTCAACACGGGGAAACTTACCAGGTCCAGACATTGTGAG
GATTGACAGATTGAGAGCTCTTTCTTGATTCGATGGGTGGTGGTGCATGG
CCGTTCTTAGTTGGTGGAGTGATTTGTCTGGTTAATTCCGTTAACGAACG
AGACCGCAGCCTGCTAAATAGCGTGCCGAACCCTTTGTTGGGCGTCGCTT
CTTAGAGGGACAACTTGTCTTCAACAAGTGGAAGTTCGCGGCAATAACAG
GTCTGTGATGCCCTTAGATGTTCTGGGCCGCACGCGCGCTACACTGATGC
ATTCAGCGAGTCTATCACCTTGACCGAGAGGTCCGGGTAATCTTTTGAAA
TTGCATCGTGATGGGGATAGATTATTGCAACTATTAATCTTCAACGAGGA
ATTCCTAGTAAGCGTGTGTCATCAGCGCACGTTGATTACGTCCCTGCCCT
TTGTACACACCGCCCGTCGCTCCTACCGATTGAATGATCCGGTGAGGCCC
CCGGACTGCGGCTCCGTCGCTGGTTCTCCAACTTTGGGGTTGCGGGAAGC
TGTCCGAACCTTATCATTTAGAGGAAGGAGAAGTCGTAACAAGGTTTCCG
TAGGTGAACCTGCAGAGGATCA

Table S2. Most abundant eukaryote OTU Ion Torrent sequence reads. Hap-A and Hap-B abundances are included. OTUs were defined at 97% similarity. Raw data was randomly resampled to yield equal sequence read count for each sample (11,078). Peak I and Peak II as indicated.

	Peak I							Peak II														
OTU Taxonomy	**Day1 0m**	**Day1 5m**	**Day1 10m**	**Day1 0 0m**	**Day10 5m**	**Day22 0m**	**Day22 5m**	**Day24 0m**	**Day24 5m**	**Day24 10m**	**Day37 0m**	**Day37 5m**	**Day37 10m**	**Day4 2 0m**	**Day42 5m**	**Day42 10m**	**Day4 5 0m**	**Day4 5 5m**	**Day45 10m**	**Day52 0m**	**Day52 5m**	**Day52 10m**	**Total reads**
Eukarya; Arthropoda; Copepoda; Calanoida; Diaptomidae	4030	114	199	4601	95	7383	7639	6102	6948	61	1787	7644	6666	3512	6933	8874	5806	6949	7925	8018	9848	7896	119030
Eukarya; Cryptophyta	596	283	89	483	109	108	113	283	255	5	2836	1777	392	703	2192	78	3162	1880	1122	975	113	15	17569
Eukarya; Cercozoa; environmental samples	36	803	1444	94	3894	246	74	24	31	9926	65	35	203	35	28	147	6	16	27	3	11	201	17349
Eukarya; Viridiplantae; Chlorophyta; Chlamydomonas	4386	1159	802	3978	2661	93	29	22	15	117	6	0	17	4	2	4	10	6	6	7	0	2	13326
Eukarya; environmental samples	1	2	4	237	8	1188	311	1480	732	1	1198	92	8	5063	335	50	423	187	253	476	71	66	12186
Eukarya; stramenopiles; Bacillariophyta; Fragilariaceae	12	31	23	239	0	914	2259	2201	2228	0	8	7	134	0	0	18	1	0	7	0	0	2	8084
Eukarya; Alveolata; Apicomplexa; Coccidia; environmental samples	6	478	549	196	463	99	14	30	11	192	45	33	1979	0	0	660	4	1	1	2	17	552	5332
Eukarya; Haptophyceae; Pavlovales; Pavlovaceae; Rebecca	64	1401	2668	2	160	5	0	0	0	1	1	1	3	0	1	0	0	0	0	0	0	0	4307
Eukarya; Viridiplantae; Chlorophyta; Pyramimonas	1	3504	522	8	136	0	0	0	0	1	1	0	0	0	0	1	0	1	0	0	1	2	4178
Eukarya; Metazoa; Arthropoda; Crustacea; Decapoda	0	4	0	1	3	0	0	0	0	0	2	0	0	70	47	0	178	648	10	394	736	995	3088
Eukarya; Cercozoa; environmental samples	1	1	0	65	1	152	131	331	284	84	433	82	15	244	344	88	50	130	270	3	7	34	2750
Eukarya; stramenopiles; Chrysophyceae; Paraphysomonas	90	658	1212	24	0	10	13	16	22	0	10	10	88	1	0	65	0	1	3	0	1	73	2297
Eukarya; Haptophyceae; Pavlovales; Pavlovaceae; Rebecca	32	7	1	50	0	9	11	24	32	0	366	207	65	81	292	7	445	248	127	246	28	0	2278
Eukarya; stramenopiles; Ochromonadales; Uroglena	3	5	3	1	2	0	0	3	0	0	1416	471	193	0	0	17	0	1	0	0	1	4	2120
Eukarya; Viridiplantae; Streptophyta; Desmidiales	1	20	8	49	3	138	117	174	152	2	122	45	149	16	76	223	39	52	306	12	30	152	1886
Hap-A: Eukarya; Haptophyceae; Isochrysidales; Isochrysidaceae; Isochrysis	1	55	46	118	40	21	8	4	4	5	1	0	7	0	0	1	0	0	0	0	0	2	313
Hap-B: Eukarya; Haptophyceae; Isochrysidales; Isochrysidaceae; Isochrysis	0	0	0	0	0	0	1	0	0	0	20	2	2	199	12	2	15	10	4	3	3	1	274

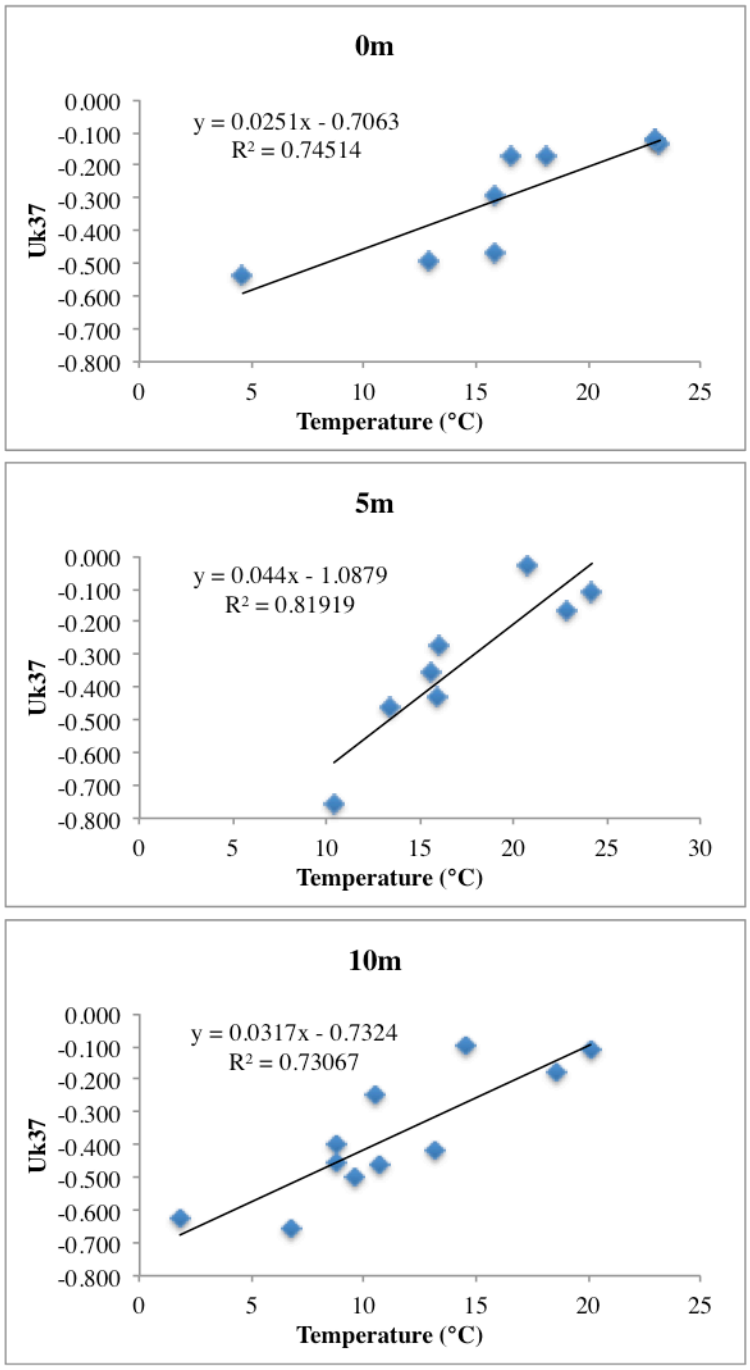

Figure S1. U^{k}_{37} calibration using only samples from 0m, 5m, and 10m depths.

CHAPTER 8

CONCLUSIONS

SUSANNA M. THEROUX

BROWN UNIVERSITY, *Department of Geological Sciences*
MARINE BIOLOGICAL LABORATORY, *Josephine Bay Paul Center*

1. Conclusions

Haptophyte algae have supplied paleoclimatologists with one of the most powerful methods to reconstruct ancient climate history. The discovery of novel alkenone-producing haptophyte species provides new challenges for this 25-year-old proxy. However, the allure of accurate, alkenone-based, continental paleotemperature reconstructions makes the endeavor worthwhile.

This thesis employs a heavily biological approach to understand the genetic and physiological determinants of alkenone sedimentary records. We have expanded the known species diversity of lacustrine alkenone-producing haptophytes and found that haptophyte species can have worldwide distributions, with single species being found in both lakes at the Arctic Circle and the Tibetan Plateau. We discovered that sedimentary records can be composites of multiple haptophyte species and that it is not always possible to decipher the haptophyte species identity based upon an alkenone fingerprint. We have also found that fluctuating concentrations of $C_{37:4}$ alkenones are likely a reflection of haptophyte community changes, as opposed to physiological changes by a single haptophyte species.

The isolation of novel haptophyte species enabled us to compare *in situ* U^k_{37} estimations with *ex situ* calibrations. We have now expanded the dataset of cultured haptophyte species with lab-based U^k_{37} calibrations, adding *Pseudoisochrysis paradoxa* and Lake George Hap-B to the list. Our isolation of Hap-A has been an achievement in endurance, and the discovery that this species it is a picoplankter and a dominant $C_{37:4}$ producer suggests it is specifically adapted for long winters and oligotrophic waters.

The monitoring of the Lake George haptophyte bloom event refined our understanding of alkenone production in Lake George and alkenone production by two novel haptophyte species. As suspected, Hap-A was present during low nutrient conditions in the early bloom season, producing predominantly $C_{37:4}$ alkenones and Hap-B was present in the late bloom season when nutrient concentrations were higher and waters were more stable. We were able to replicate this bloom event in an enrichment culture, monitoring haptophyte species over time, and just as observed in the field, Hap-A is the dominant species during the early stages of the bloom. Our culturing results determined the Hap-A and Hap-B U^{k}_{37} calibrations are statistically identical but distinct from the *in situ* calibration, and we believe the *in situ* calibration reflects the most accurate relationship between water temperature and alkenone unsaturation in Lake George.

At times, DNA sequencing makes the haptophyte community, and therefore the alkenone record, appear dauntingly complex. However, DNA sequencing is able to resolve complexities in haptophyte communities that are otherwise indecipherable or overlooked. This thesis stands in support of a multi-disciplinary approach to alkenone-based paleothermometry; the U^{k}_{37} index will be better for it.