**Temperature calibration and phylogenetically distinct distributions for
freshwater alkenones: Evidence from northern Alaskan lakes**

William M. Longo^{*1}, Susanna Theroux², Anne E. Giblin³, Yinsui Zheng^{1,4},
James T. Dillon¹, Yongsong Huang^{*1}

¹Department of Earth, Environmental and Planetary Sciences, Brown University,
Providence, RI, 02912

²Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley CA, 94720

³Ecosystems Center, Marine Biological Laboratory, Woods Hole, MA, 02543

⁴Josephine Bay Paul Center, Marine Biological Laboratory, Woods Hole, MA, 02543

* Corresponding Authors:

William M. Longo

E-mail: william_longo@brown.edu

Tel: +1 401 863 3659

Fax: +1 401 863 2058

Yongsong Huang

E-mail: yongsong_huang@brown.edu

Tel: +1 401 863 3822

Fax: +1 401 863 2058

Abstract

Alkenones are a class of unsaturated long-chain ketone biomarkers that have been used to reconstruct sea surface temperature and, more recently, continental temperature, by way of alkenone unsaturation indices (e.g. U_{37}^K and $U_{37}^{K'}$). Alkenones are frequently found in brackish and saline lakes, however species effects confound temperature reconstructions when multiple alkenone-producing species with different temperature responses are present. Interestingly, available genetic data indicate that numerous freshwater lakes host a distinct phylotype of alkenone-producing haptophyte algae (the Group I or Greenland phylotype), providing evidence that species effects may be diminished in freshwater lakes. These findings encourage further investigation of alkenone paleotemperature proxies in freshwater systems. Here, we investigated lakes from northern Alaska (n=35) and show that alkenones commonly occurred in freshwater lakes, where they featured distinct distributions, characterized by dominant $C_{37:4}$ alkenones and a series of tri-unsaturated alkenone isomers. The distributions were characteristic of Group I-type alkenone distributions previously identified in Greenland and North America. Our analysis of suspended particulate matter from Toolik Lake (68° 38'N, 149° 36'W) yielded the first *in situ* freshwater U_{37}^K calibration ($U_{37}^K=0.021*T-0.68$; $r^2=0.85$; $n=52$; $RMSE=\pm 1.37$ °C). We explored the environmental significance of the tri-unsaturated isomers using our northern Alaskan lakes dataset in conjunction with new data from haptophyte cultures and Canadian surface sediments. Our results show that these temperature-sensitive isomers are biomarkers for the Group I phylotype and indicators of multiple-species effects. Together, these findings highlight freshwater lakes as valuable

47 targets for continental alkenone-based paleotemperature reconstructions and demonstrate
48 the significance of the recently discovered tri-unsaturated isomers.

49

50 **Key Words:** U_{37}^K , haptophyte phylogeny, temperature calibration, tri-unsaturated
51 isomers, chemotaxonomy, freshwater lakes, continental paleoclimate

1. Introduction

Continental paleoclimate reconstructions are necessary for evaluating climate models and quantifying natural climate variability in the context of anthropogenically-forced climate change (Braconnot et al., 2014). Lakes offer high-quality sedimentary archives that can be used for climate reconstruction, however relatively few quantitative lacustrine temperature proxies exist. Transfer functions based on biological proxies are used for temperature reconstruction, however these proxies can be limited by confounding variables, poor analogs, and changes in species responses through time (e.g. Juggins, 2013). Organic geochemical temperature proxies have potential as another means for reconstructing continental temperature, yet many remain in development and require additional calibration and modern study (Castaneda and Shouten, 2011).

Long chain alkenones (LCAs) are a class of C₃₅-C₄₂ methyl and ethyl ketones found globally in lacustrine sediments and they hold potential as quantitative organic geochemical temperature proxies. LCAs were first described in marine sediments over three decades ago (De Leeuw et al., 1980) and it was later found that the degree of LCA unsaturation was well correlated with the temperature of the water in which LCAs are produced (Brassell et al., 1986). The U_{37}^K and $U_{37}^{K'}$ indices were developed to provide metrics of LCA unsaturation (Brassell et al., 1986; Prahl and Wakeham, 1987), calibrated as proxies for sea surface temperature (SST; Conte et al., 2006; Muller et al., 1998; Prahl et al., 1988) and applied widely for SST reconstructions at a variety of sites and timescales (Herbert, 2003). LCAs have since been discovered in numerous lakes around

the world (e.g. Chu et al., 2005; Cranwell, 1985; D’Andrea and Huang, 2005; Pearson et al., 2008; Thiel et al., 1997; Toney et al., 2010; Volkman et al., 1988; Zink et al., 2001; refs. in Castaneda and Shouten, 2011), where they are produced by several species of haptophyte algae (Theroux et al., 2010). Many lacustrine U_{37}^K and $U_{37}^{K'}$ temperature calibrations exist (Chu et al., 2005; Sun et al., 2007; D’Andrea et al., 2011; 2012; Nakamura et al., 2014; Theroux et al., 2013; Toney et al., 2010; 2012; Zink et al., 2001); however the effects of haptophyte species diversity on LCA distributions have prevented the broad application of LCA-based temperature reconstructions in lakes.

These “multi-species effects” have been well documented in saline lakes (Coolen et al., 2004; Theroux et al., 2010; Toney et al., 2012) and have overprinted U_{37}^K and $U_{37}^{K'}$ temperature estimates in sedimentary records (Randlett et al., 2014; Wang et al., 2015). *Ruttnera lamellosa* (formerly *Chrysotila lamellosa*), *Isochrysis galbana* and related species are the common alkenone-producing haptophyte algae in brackish waters and saline lakes. These species belong to a broad and relatively diverse phylogenetic clade (Group II; Bendif et al., 2013; Theroux et al., 2010) that produces LCAs with a range of distributions, and possesses a wide range of salinity tolerances and temperature sensitivities (Nakamura et al., 2014; Ono et al., 2012; Sun et al., 2007; Theroux et al., 2013; Toney et al., 2012). The phenotypic diversity of the brackish- and saline lake-dwelling Group II haptophytes has lead to species effects on LCA distributions and temperature calibrations in many saline lakes.

Alkenone-producing haptophytes from a number of freshwater and oligohaline environments belong to a separate phylogenetic clade referred to as the Greenland phylotype (D'Andrea et al., 2006) or Group I (Theroux et al., 2010). Though first discovered in a set of oligohaline lakes in southwestern Greenland (D'Andrea et al., 2006), the Group I phylotype was the sole Isochrysidales haptophyte clade detected in freshwater lakes in northeastern Canada (Theroux et al., 2010), northern Alaska (Crump et al., 2012), Japan (McColl et al., 2013), and Norway (D'Andrea et al., 2016). Extensive 18S rRNA gene sequencing in French Lakes confirmed the presence of another freshwater Isochrysidales haptophyte lineage – the EV clade (Simon et al., 2013) – that branches with Group I according to our phylogenetic analysis. While there is genetic diversity within the Group I phylotype (Theroux et al., 2010), we show that the salinity tolerance for the phylotype appears to be limited to the freshwater to oligohaline range. Furthermore, these LCA-containing fresh and oligohaline lakes appear to be devoid of the Group II alkenone-producers that are common in saline and brackish environments (Theroux et al., 2010). Indeed, ancient sedimentary DNA analyses in oligohaline BrayaSø, southwestern Greenland, indicated that the Group I phylotype was the sole alkenone-producer for the past 6,000 years, allowing for a paleotemperature reconstruction without species effects (D'Andrea et al., 2006; 2011). Although a limited number of lakes have been examined, the most parsimonious hypothesis from existing genetic and biomarker data is that Group I haptophytes are the common freshwater-dwelling alkenone-producers.

LCAs commonly found in freshwaters and presumably produced by the Group I haptophytes may elude species effects, since evidence suggests that their distributions are coherent across lakes and geographic regions (D'Andrea et al., 2006; D'Andrea et al., 2016; Longo et al., 2013). However, little is known about the environmental controls on these LCA distributions and their relationships to temperature. D'Andrea et al. (2005; 2006) and Theroux et al. (2010) described the *in situ* distributions of alkenones produced by the Group I phylotype, but these studies employed gas chromatography (GC) stationary phases that could not separate the full distribution of LCAs. The improved separation of LCAs (Longo et al., 2013) revealed that the Group I haptophytes produced a novel distribution of 16 LCAs, including four tri-unsaturated isomers with $\Delta^{14,21,28}$ double bond positioning (Fig. S.1; Dillon et al., 2016; Longo et al., 2013). This fully resolved LCA distribution has never been investigated or calibrated, despite the fact that genetic and geochemical data suggest it may commonly occur (Theroux et al., 2010). Further studies and calibrations are needed to verify the temperature response of the Group I phylotype and evaluate whether or not it differs regionally or between oligohaline and freshwater environments.

The main objective of this study was to calibrate and describe the taxonomic and environmental controls on freshwater LCAs using surface sediments and SPM from northern Alaskan lakes. These LCAs have a highly conserved distribution, which we show is distinctive of the Group I phylotype based on new culture data and our review of published genetic and geochemical data. Importantly, our results provide evidence that freshwater lakes that host this LCA distribution are devoid of species effects. We

142 performed an *in situ* temperature calibration using SPM from Toolik Lake (68° 38'N,
143 149° 36'W), a system where the Group I phylotype has been identified as the only known
144 alkenone-producer (Crump et al., 2012), and compared our calibration to those from
145 other lakes using revised U_{37}^K indices and bootstrapping methods, which account for the
146 tri-unsaturated isomers and rigorously treat the error in calibration coefficients. We also
147 explore the environmental significance of the recently discovered tri-unsaturated isomers,
148 with particular focus on their utility as temperature indicators and chemotaxonomic
149 indicators for Group I haptophytes.

2. Samples and Study Area

The North Slope of Alaska occurs between 68°N and 71°N and comprises three physiographic regions—the Brooks Range Mountains, Arctic Foothills and Arctic Coastal Plain—that together contain tens of thousands of lakes (Hobbie and Huryn, 2013). We collected surface sediments from a series of lakes (n=35) on a transect across the Arctic Foothills and Arctic Coastal Plain (Fig. 1). The majority of the lakes occur at the southeastern end of the transect and receive annual environmental monitoring by the Arctic Long Term Ecological Research program (ARC LTER). These lakes experience mean annual air temperatures (MAAT) of approximately -12 °C with low winter temperatures of -35 °C and high summer temperatures of 20 °C (Hobbie and Huryn, 2013). Mean annual precipitation for the region is 292 mm yr⁻¹ and mean annual evapotranspiration is 127 mm yr⁻¹ (Dery et al., 2005). The northwestern limb of the transect incorporates five lakes (IKP-003; INI-004; MC7916; Teshekpuk; UMI-004) that are part of the Circum-Arctic Lakes Observation Network (CALON; Arctic Observing Network). All study lakes occur between elevations of 0m and 800m above sea level and comprise large gradients in morphometric, physical and chemical properties.

More intensive environmental monitoring and sampling of SPM for the *in situ* temperature calibration were carried out in Toolik Lake (68° 38'N, 149° 36'W), a relatively large (surface area = 1.5 Km²) and deep (maximum depth = 25m) dimictic, multi-basin kettle lake. Toolik Lake is ice free from mid-June through September, reaching maximum summer surface water temperatures of approximately 18 °C or less in

173 July (ARC LTER database). It exists in a hydrologically open basin formed in moraine
174 and outwash deposits of the Itkillik II glaciation 12 to 14 Ka (Hamilton and Porter, 1975).
175 The ARC LTER has been monitoring the physical, chemical and biological parameters of
176 Toolik Lake since the 1970s.

3. Methods

3.1. Sample collection and preparation

Toolik Lake SPM samples were obtained from 1.8 to 3.0 L water samples, collected with a Van Dorn sampler, at various depths and on nine individual sampling days between June 15 and July 20, 2013. Accompanying water column measurements of temperature, pH, conductivity, and chlorophyll fluorescence were taken with a Hach™ Hydrolab water column sensor array. Water samples were kept in the dark at *in situ* temperatures and filtered within two hours of collection by vacuum filtration through ashed (450 °C; 8 hrs) Whatman GF/F filters (0.7 µM). The filtrate was discarded and filters were immediately frozen (-20 °C) for later analysis of SPM.

All northern Alaskan surface sediments were subsampled from short cores collected by pole or gravity coring devices and sectioned in the field. The top 1.0 cm of sediment was subsampled for ARC LTER lakes. CALON surface sediments were subsampled from the top 5.0 cm of sediment. Efforts were made to collect samples from the deep hole in all lakes, however partial ice cover made this impossible at some sites, in which cases cores were taken from below the average thermocline depth as determined from ARC LTER data. Sediments were sampled into Whirl-Pak™ collection bags and frozen at -20 °C for later analysis. Additionally, 11 freeze-dried surface sediments from Canadian lakes (Toney et al., 2011) were selected for re-extraction and reanalysis using GC methods that separate tri-unsaturated alkenone isomers (Longo et al, 2013).

200

201 Four Haptophyte cultures were grown and extracted for LCA analysis, including
202 *Emiliania huxleyi* (Van 556; isolated from the North Atlantic off the Norwegian coast),
203 *Ruttenra lamellosa* (CCMP1307; isolated from Keystone Salt Marsh, Whidbey Island,
204 Washington, USA), *Isochrysis galbana* (CCMP1323; isolated from Port Erin, Isle of
205 Man, North Atlantic Ocean), and *Tisochrysis lutea* (CCMP463; isolated from Little
206 Water Cay, Turks and Caicos Islands, British West Indies). All cultures were grown at 9
207 and 20°C in 0.2 µm filter-sterilized seawater with a salinity of 33 psu. Cultures were
208 amended with f/2 nutrients and subjected to a 16:8 light:dark cycle without aeration.
209 Cultures were harvested for alkenone analysis from stationary phase by filtration through
210 a 47-mm glass fiber filter (Whatman GF/A, 0.45 µm) and immediately frozen at -20 °C.

211

212 All LCAs were extracted and purified for analysis following established methods (e.g.
213 D'Andrea and Huang, 2005; Longo et al., 2013; Toney et al., 2010). Briefly, filters were
214 freeze-dried and extracted by sonication with dichloromethane:methanol (9:1 v/v). The
215 extracts were saponified and purified by column chromatography with silica gel before
216 analysis by gas chromatography-flame ionization detection (GC-FID). Culture samples
217 did not require saponification. Surface sediments were freeze-dried and extracted with a
218 Dionex™ accelerated solvent extraction (ASE) system at 120 °C and 1200 psi. The
219 extracts were separated into acid and neutral fractions by flash column chromatography
220 with Supelco Supelclean LC-NH2 (45 µm, 60 Å). The neutral fractions were further
221 separated into alkane, ketone and polar fractions by flash column chromatography using

silica gel (40–63 μm , 60 $\text{\AA}$). The ketone fraction was saponified, then purified again by silica gel column chromatography before analysis by GC-FID.

3.2. Analytical methods

LCAs from SPM and surface sediment samples were analyzed with an Agilent 6890N GC system equipped with a flame ionization detector (FID) and an Agilent VF-200ms capillary column (60 m x 250 μm x 0.10 μm). The GC method utilized pulsed splitless injection (20 psi at 320 $^{\circ}\text{C}$), a 36 cm s^{-1} column flow rate, and the following oven program: initial temperature of 60 $^{\circ}\text{C}$ (hold 1 min), ramp 20 $^{\circ}\text{C}/\text{min}$ to 255 $^{\circ}\text{C}$, ramp 3 $^{\circ}\text{C}/\text{min}$ to 320 $^{\circ}\text{C}$ (hold 10 min). LCA peak identification was accomplished by retention time comparison with an alkenone standard containing the 16 LCAs considered in this study and by GC-MS analysis, performed using an Agilent 6890N GC system coupled to an Agilent 5973N quadrupole mass spectrometer. The GC conditions used for GC-MS analysis were the same as those used for GC-FID. The MS conditions were the following: ionization energy, 70 eV; scan range, 40-600 m/z . LCA quantification was accomplished using 18-pentatriacontanone as an internal standard. Analysis of the carbon and nitrogen content of the surface sediments was accomplished using a ThermoQuest NC2100 elemental analyzer. Based on strong correlation between C and N in sediments ($r^2=0.88$), low C:N ratios (<16), and absence of carbonate minerals in smear slide analyses, we assume total carbon content is equal to total organic carbon content for the surface sediments of the northern Alaskan lakes.

3.3. Ratios of isomeric ketones and unsaturation indices

LCAs from northern Alaska and Greenland comprise 4 groups of structural homologs ($C_{37}Me$, $C_{38}Et$, $C_{38}Me$, $C_{39}Et$), each of which contain tri-unsaturated LCA isomers (Fig. S.1). We devised the following ratios of isomeric ketones (RIK) indices for each structural group, to examine the environmental significance of the isomers (eqs. 1–4).

$$RIK_{37} = \frac{C_{37:3a}}{C_{37:3a} + C_{37:3b}} \quad (1)$$

$$RIK_{38E} = \frac{C_{38:3a}Et}{C_{38:3a}Et + C_{38:3b}Et} \quad (2)$$

$$RIK_{38M} = \frac{C_{38:3a}Me}{C_{38:3a}Me + C_{38:3b}Me} \quad (3)$$

$$RIK_{39E} = \frac{C_{39:3a}Et}{C_{39:3a}Et + C_{39:3b}Et} \quad (4)$$

In each RIK index, ‘a’ is the common $\Delta^{7,14,21}$ isomer, and ‘b’ is the novel $\Delta^{14,21,28}$ isomer. RIK values of 1.0 indicate the presence of only the $\Delta^{7,14,21}$ homolog, whereas a value of 0.0 indicates the presence of only the isomeric $\Delta^{14,21,28}$ homolog. The presence of the tri-unsaturated isomers requires that distinctions be made between different U_{37}^K indices. The U_{37}^K index as described by Brassell et al. (1986; eq. 5) and modified and calibrated by Prahl and Wakeham (1987) and Muller et al. (1998), was initially applied to marine samples, which do not contain tri-unsaturated isomers according to a limited amount of environmental and culture data (Longo et al., 2013).

$$U_{37}^K = U_{37a}^K = \frac{C_{37:2} - C_{37:4}}{C_{37:2} + C_{37:3a} + C_{37:4}} \quad (5)$$

267

268 An alternate form of the U_{37}^K index was used by D'Andrea et al (2011) for calibration and
269 temperature reconstruction. This form of the index included both the $C_{37:3a}$ and $C_{37:3b}$
270 isomers, as these compounds were unresolved using traditional GC methods (eq. 6).

271

272
$$U_{37}^K = U_{37ab}^K = \frac{C_{37:2} - C_{37:4}}{C_{37:2} + (C_{37:3a} + C_{37:3b}) + C_{37:4}} \quad (6)$$

273

274 Here we distinguish between these alternate forms as U_{37a}^K (eq. 5) and U_{37ab}^K (eq. 6). We
275 also consider several other U^K and $U^{K'}$ indices for LCAs with different carbon numbers
276 and ketone positions and with one or both of the tri-unsaturated isomers. The indices take
277 the form of the U_{37}^K or $U_{37}^{K'}$ index, but apply to the other structural groups of LCAs. In
278 each U^K and $U^{K'}$ index, x refers to the carbon number (C_{37} , C_{38} , or C_{39}) and carbonyl
279 position (Me or Et) combination (e.g. $C_{38}Et$), and y refers to the tri-unsaturated isomers
280 used (a , b , or $a + b$; eqs. 7, 8).

281

282
$$U_{xy}^K = \frac{X_2 - X_4}{X_2 + X_y + X_4} \quad (7)$$

283
$$U_{xy}^{K'} = \frac{X_2}{X_2 + X_y} \quad (8)$$

284

285 Fractional Abundances (FA) of LCAs were computed as the concentration of each LCA
286 divided by the summed concentrations of all 16 LCAs.

287

288 **3.4. Environmental data**

289

290 All physical and chemical parameters not previously mentioned were measured by the
291 ARC LTER lakes monitoring program (ARC LTER Database) or CALON.
292 Morphometric data for ARC LTER lakes including surface area (SA) and mean depth
293 (mZ) were provided by the Toolik Field Station Office of Geographical Information
294 Systems. Conductivity (COND), pH, dissolved organic carbon (DOC), alkalinity (ALK),
295 and chlorophyll (Chla) are the averages of one or more measurements taken during the
296 ice-free season in 2012 at 1m depth. $[\text{SO}_4^{2-}]$ and $[\text{Cl}^-]$ are the averages of one or more
297 measurements taken during the ice-free season of 2005 at 1m depth.

298

299 Measurements of lake water properties (pH, COND, DOC, ALK, $[\text{SO}_4^{2-}]$, $[\text{Cl}^-]$, Chla)
300 showed significant variation between sites, while mean values within sites were relatively
301 consistent across sampling years (ARC LTER Database), providing evidence that our
302 single year measurements are representative of modern lake properties. By contrast, lake
303 water temperatures fluctuated considerably during the spring/early summer alkenone
304 bloom season, therefore attempts were not made to calibrate surface sediment LCAs to
305 point sampled water temperatures. The sampling transect occurs on a 4 - 6 °C gradient in
306 surface air temperature (NCEP/NCAR reanalysis), however LCA-containing lakes were
307 concentrated at the southeastern end of the transect; therefore attempts were not made to
308 calibrate surface sediment LCAs to surface air temperatures for lack of LCAs at the
309 northern end of the transect.

310

311 **3.5. Phylogenetic tree construction and haptophyte salinity review**

312

313 For the 18S rRNA gene phylogenetic analyses, SSU sequences were aligned with SILVA
314 version 121 reference sequences (Pruesse et al., 2007) and environmental sequences from
315 Simon et al. (2013) and manually curated in ARB (Ludwig et al., 2004). Gaps were
316 removed from the alignment resulting in an alignment with 1167 positions. RAxML 8.0.0
317 (Stamatakis, 2014) was used to generate a maximum likelihood tree with the GTRCAT
318 model of nucleotide substitution and 500 bootstrapped replicates. An accompanying
319 review of lacustrine haptophyte phylogeny and salinity was also undertaken. Lakes with
320 i) reported salinities and ii) reported 18S rRNA gene sequences from Isochrysidales
321 haptophytes were included in the review, which compiled observed salinities and salinity
322 ranges for various haptophytes identified by their 18s rRNA phylogeny. Salinity ranges
323 for a selection of published haptophyte culture experiments were also included.

324

325 **3.6. Statistical analyses**

326

327 All statistical analyses were carried out with the statistical software R 3.0.0, using the
328 boot (Canty and Ripley, 2015; Davidson and Hinkley, 1997), car (Fox and Weisberg,
329 2011) and vegan (Oksanen et al., 2013) packages. Before carrying out principal
330 components analysis (PCA) on lake properties and alkenone parameters, attempts were
331 made to transform data to normal distributions and detrended correspondence analysis
332 was used to ensure that LCA indices, concentrations and environmental data varied
333 linearly (Leps and Smilauer, 2003). For the surface sediment dataset, 29 of the 35 lakes
334 were included in original PCA analyses. The CALON lakes were not included because

environmental data were missing or collected in a different manner than for ARC LTER lakes. Perch Lake was also removed from the analysis as an outlier, because it featured anomalously high alkenone concentrations ($>4\sigma$ above mean concentrations) and disproportionately influenced the analysis. More than 90% of all environmental data were available for the 29-lake PCA dataset. The missing environmental data were replaced with mean values of the normalized data (Leps and Smilauer, 2003). A second PCA was carried out on a subset of 20 lakes from the original PCA dataset, in which LCAs were present and quantified.

In order to compare the U_{37}^K calibrations from this study to previously published calibrations, we performed ordinary non-parametric bootstrapped linear regressions of several published U_{37}^K vs. temperature datasets. We employed different U_{37}^K indices for calibration comparisons based on the premise that the BrayaSø calibration (D'Andrea et al., 2011) did not separate the isomers and therefore utilized U_{37ab}^K , whereas all published Group II calibrations take the form U_{37a}^K , since they likely did not include $C_{37:3b}$. Our criteria for selecting calibrations for comparison was to only use published lacustrine *in situ* U_{37}^K calibrations or published U_{37}^K calibrations from lacustrine haptophyte cultures. For each bootstrapped regression the raw U_{37}^K vs. temperature data were sampled with replacement 5000 times and used to formulate 5000 individual linear least squares regressions. The resulting regressions had normally distributed slope and intercept coefficients, allowing for the computation of bootstrapped confidence intervals for slope and intercept coefficients and proper comparison between calibrations.

4. Results

4.1. Salinity and lacustrine haptophyte phylogeny

Reviewed genetic and geochemical data indicated that Group I and Group II haptophytes occupy different salinity ranges. Group II haptophytes were found in lakes with a broad salinity range including the oligohaline Clear and Pyramid Lakes in North Dakota, and the hyperhaline Great Salt Lake, Utah (Fig. 2). Despite their apparently large range of salinity tolerances, Group II haptophytes have not been identified in freshwater lakes. By contrast, the Group I haptophytes appear to be restricted to a lower salinity range, having been found in freshwater lakes in Alaska and Canada, and in oligohaline lakes in Greenland and central North America (Fig. 2). While the Group I and Group II haptophytes have overlapping salinity ranges (notably in the oligohaline range of 0.5 - 5 psu), only the Group I haptophytes appear to be present in freshwaters (0 - 0.5 psu). Furthermore, new 18S rRNA gene phylogenetic analyses demonstrated that members of the recently described Isochrysidales 'EV clade' found in French freshwater lakes branched with the Group I haptophytes (Fig. S.2; Simon et al., 2013). Concurrently, sequences identified by Simon et al. (2013) from brackish or saline environments, branched with Group II (Fig. S.2), corroborating the observed relationships between salinity and the phylogeny of haptophyte algae.

4.2. Alkenone occurrence, abundance and distribution in northern Alaskan lakes

4.2.1. Surface sediments

LCAs were detected in 66% of all surface sediments analyzed and in 22 of the 34 freshwater lake sediments in the transect (Tables 1, S.1). C_{37} concentrations in LCA-containing surface sediments ranged from 1.23 - 177 $\mu\text{g gTOC}^{-1}$ with an average C_{37} concentration of 20.7 $\mu\text{g gTOC}^{-1}$ (Table 1). Total LCA concentrations were as high as 357 $\mu\text{g gTOC}^{-1}$. All LCA-containing surface sediments (n=23) featured a $C_{37:4}$ dominant profile and all with the exception of the single saline lake (IKP-003) included the tri-unsaturated isomers (Fig. 3). Among the nine lakes where all 16 LCAs could be quantified, mean % $C_{37:4}$ was 25.3 ± 3.3 , mean C_{37}/C_{38} ratios were 0.88 ± 0.16 and mean Me/Et ratios were 1.37 ± 0.28 . Mean RIK_{37} values for all lakes that contained the $C_{37:3b}$ isomer (n=22) were 0.54 ± 0.024 and ranged from 0.51 to 0.60 (Table 1).

Principal components analysis (PCA) of the northern Alaskan lakes dataset revealed the main modes of variability differentiating lakes. Principal components 1, 2 and 3 described 45%, 19% and 13% of the variability in the environmental data, respectively. The first principal component (PC1) was highly negatively correlated with pH, alkalinity and conductivity and served to separate the lakes by gradients in these properties. The second principal component (PC2) separated lakes primarily by mean depth (mZ). Individual lakes plotted as circles scaled to C_{37} LCA concentrations revealed that in general, lakes with relatively high concentrations of LCAs plotted in the top left quadrant of the biplot and/or were negatively (positively) correlated with PC1 (PC2). This

indicated that LCA occurrence and abundance were associated with higher conductivity, pH, alkalinity and mean depth (Fig. 4a; Table S.2).

A second PCA included only the lakes where LCAs were present and allowed for plotting LCA abundance and distribution parameters *a posteriori* as vectors in the ordination space to reveal their correlations with lake properties (Fig. 4b; Table S.2). PC1, PC2 and PC3 explained 44%, 17% and 14% of the variability in the environmental data, respectively. The main modes of variability from the original PCA were retained, as PC1 was highly correlated with pH, conductivity, alkalinity and $[\text{SO}_4^{2-}]$ and PC2 was correlated with mean depth. C_{37} LCA concentration and the $\text{C}_{37:4}/\text{C}_{38:4}\text{Et}$ parameter plotted in the top right quadrant of the biplot, indicating their positive correlations with both PC1 and PC2. ($\text{C}_{37:4}/\text{C}_{38:4}\text{Et}$ was used instead of $\text{C}_{37}/\text{C}_{38}$ or Me/Et, because it could be measured for all samples in the PCA dataset). U_{37}^{K} and RIK_{37} showed positive relationships with DOC and C:N ratios, with each of these parameters plotting in the lower left quadrant of the biplot (Fig. 4b; Table S.2).

4.2.2. Toolik Lake SPM

SPM samples from Toolik Lake had LCA concentrations that ranged from 11.3 - 467 $\mu\text{g L}^{-1}$ (Table S.3). LCAs were first detected in the water column when the lake had 90% ice cover and they persisted throughout the sampling period, which lasted until the mid-summer. The highest LCA concentrations by a factor of four occurred during the spring isothermal mixing period (IM; n=16), however LCAs were consistently detected when

the lake was partially ice covered (PIC; n=17) and thermally stratified (TS; n=19; Fig. 5). Because isothermal mixing occurred when water column temperatures were between 3 and 8 °C, the highest concentrations of LCAs were observed in the lower half of the *in situ* temperature range (Fig. 5).

All LCA-containing samples from Toolik Lake SPM (n=52) had C_{37:4} dominant profiles and all contained the full suite of tri-unsaturated alkenone isomers (including the C_{37:3b}Me, C_{38:3b}Et, C_{38:3b}Me, C_{39:3b}Et compounds; Fig. 3). SPM samples had mean %C_{37:4} values of 30.0 ± 3.8, mean C₃₇/C₃₈ ratios of 1.22 ± 0.089 and Me/Et ratios of 1.76 ± 0.17. %C_{37:4} had a comparable mean and standard deviation to the surface sediments, however C₃₇/C₃₈ and Me/Et ratios differed between SPM and surface sediments. Mean C₃₇/C₃₈ and Me/Et ratios from Toolik Lake surface sediments were more than 2σ and 1σ less than mean values from SPM, respectively. C₃₇/C₃₈ was significantly positively correlated with LCA concentrations (p<0.01; r²=0.23), and Me/Et ratios were significantly positively correlated with LCA concentrations (p<0.01; r²=0.27) and negatively correlated with sampling date (p<0.01; r²=0.26).

4.3. Alkenone temperature dependence in Toolik Lake

Fractional abundances of 14 LCAs were significantly correlated with water column temperatures based on parametric linear least squares regressions (Fig. 6). To control for non-normal data, non-parametric methods were also employed and indicated that fractional abundances of 12 LCAs had significant temperature effects (p<0.01,

Spearman's rank). Altogether, fractional abundances of 11 LCAs ($C_{37:4}$, $C_{37:3a}$, $C_{37:3b}$, $C_{37:2}$, $C_{38:3a}Et$, $C_{38:2}Et$, $C_{38:4}Me$, $C_{38:3b}Me$, $C_{38:2}Me$, $C_{39:3a}Et$, $C_{39:2}Et$) were significantly correlated with temperature based on both parametric and non-parametric methods ($p < 0.01$; $r^2 > 0.20$).

We provided the first lacustrine U_{37}^K calibration for northern Alaska and the first *in situ* calibration from a freshwater lake. The calibration ($U_{37ab}^K = 0.021 \cdot T - 0.68$; $r^2 = 0.85$; $n = 52$; Fig. 7) had relatively low error (RMSE = ± 1.37 °C) and performed best in the temperature range of 3 - 12 °C, which spans the typical range of spring/early summer water temperatures in Toolik Lake (ARC LTER Database). $U_{37}^{K'}$, U_{37a}^K , U_{37b}^K and additional U^K indices were also calibrated to *in situ* temperature with relatively low error (Table 2). Comparison of the Toolik Lake U_{37ab}^K calibration with the calibration of D'Andrea et al. (2011) revealed that the calibrations have comparable slope coefficients, based on the overlap of the bootstrapped confidence intervals at the 95% and 68% levels, however the intercepts were significantly different (Figs. 8; S.3).

Interestingly, the Toolik U_{37a}^K calibration was comparable in both slope and intercept (based on overlap of the 95% confidence intervals) to the *in situ* calibration of Toney et al. (2012) from Lake George, ND, which originated from distantly related alkenone producers (Figs. 8; S.3). The Toolik U_{37a}^K calibration was also similar in slope to the *P. paradoxa* calibration of Theroux et al. (2013), however this calibration was forced to fit a linear regression for comparison and as a result, showed large bootstrapped confidence intervals. The *P. paradoxa* calibration fits a polynomial regression somewhat better ($r^2 =$

0.98) than a linear regression ($r^2 = 0.91$; Theroux et al., 2013). The slope for the Sun et al. (2007) calibration of *R. lamellosa* and the intercepts from both the *R. lamellosa* and *P. paradoxa* calibrations differed significantly from the Toolik U_{37a}^K calibration (Figs. 8; S.3).

All RIK indices were significantly correlated with *in situ* temperature (Fig. 9; $p < 0.01$, $r^2 \geq 0.20$; Pearson's squared correlation coefficient). RIK_{37} showed a modest correlation, with little overall trend or variability in the index. However, RIK_{38E} was well correlated with *in situ* temperature ($p < 0.01$; $r^2 = 0.76$) and the index varied by more than 15% over a 12 °C temperature range. The methyl isomer ratios (RIK_{37} and RIK_{38M}) were both negatively correlated with temperature, whereas the ethyl isomer ratios (RIK_{38E} and RIK_{39E}) were positively and more strongly correlated with temperature.

4.4. Effects of haptophyte phylogeny and salinity on LCA distributions

LCAs in SPM samples from Toolik Lake were attributed to a Group I haptophyte producer based on published genetic data (Crump et al., 2012), and all contained Group I-type distributions including dominant $C_{37:4}$ and the tri-unsaturated isomers (Fig. 3). New analyses of one Group III strain (*E. huxleyi* Van 556) and three Group II strains (*I. galbana* CCMP1323, *R. lamellosa* CCMP1307, and *T. lutea* CCMP463) revealed the absence of $C_{37:3b}$ isomers despite abundant $C_{37:3a}$ in all of the cultures. Furthermore, the only $\Delta^{14,21,28}$ isomers found in these samples were trace amounts of $C_{38:3b}Et$ in the *I. galbana* and *R. lamellosa* cultures (Fig. S.4, Table S.4). Together these data suggested

that the tri-unsaturated isomers and particularly C_{37:3b}, are specific to the Group I haptophytes.

The LCA-containing surface sediments from northern Alaskan freshwater lakes (n=22) featured LCA distributions analogous to Group I-type distributions from Toolik Lake and BrayaSø (Fig. 3; Table 1). By contrast, the saline lake from northern Alaska (IKP-003) did not contain the C_{37:3b} isomer, or C₃₈Me LCAs (Fig. 3; Table S.1). Interior Canada lake sediments (Toney et al., 2011) reanalyzed for this study included four oligohaline lakes and seven brackish to hyperhaline lakes. All four oligohaline lakes contained the $\Delta^{14,21,28}$ tri-unsaturated isomers and in contrast, the seven lakes with higher salinity had no detectable C_{37:3b} LCAs (Fig. S.4, Table S.4). Together, these 34 new analyses of LCA-containing surface sediments from northern Alaska and Canada indicated that freshwater lakes invariably featured Group I-type LCA distributions, oligohaline lakes featured both Group I and Group-II type distributions, and mesohaline to hyperhaline lakes featured Group II-type distributions.

We also considered an expanded LCA dataset (n=124) including samples of any type that have been analyzed with methods that allow for the separation of the tri-unsaturated alkenone isomers. These included LCA measurements from all SPM and surface sediments from this study, all samples described by Longo et al. (2013), all new culture samples presented in this study, Chinese surface sediment data from Zhao et al. (2014), reanalyzed Canadian surface sediments from Toney et al. (2011), and *P. paradoxa* culture data from Theroux et al. (2013). Ternary diagrams illustrated the key differences

in C_{37} LCA distributions that separated the diverse samples by phylotype and salinity. We considered the C_{37} LCAs as a three component mixture consisting of: i) the $C_{37:2} + C_{37:3a}$ LCAs; ii) the $C_{37:4}$ LCAs; and iii) the $C_{37:3b}$ LCAs. In relation to phylotype, these three components roughly represent end members for Group III marine distributions, Group II $C_{37:4}$ dominant distributions, and Group I distributions, respectively. As expected, the Group III samples (including samples from two *E. Huxleyi* strains) plotted in the upper vertex of the ternary diagram indicating 100% of the C_{37} LCAs were either $C_{37:2}$ or $C_{37:3a}$ (Fig. 10a). Known Group II samples (including all Group II haptophyte culture data and Lake George surface sediment) plotted along the left axis reflecting the presence of $C_{37:4}$ in varying relative abundances. Known Group I samples (including surface sediment from BrayaSø, Greenland; surface sediment from Toolik Lake and Toolik Lake SPM) plotted towards the bottom axis and clustered closer to the center of the diagram.

Most surface sediments from interior Canada and China (Zhao et al., 2014) plotted along the left axis of the ternary diagram suggesting a Group II producer. By contrast, all freshwater northern Alaskan lake sediments plotted with the Group I cluster and the single saline lake sample from northern Alaska (IKP-003) plotted along the left axis with Group II samples. Three Surface sediments from interior Canadian lakes plotted between the Group I cluster and left axis suggesting mixed sources (Fig. 10a). Replotting the data with salinity designations (Venice system, 1959) corroborated our species-salinity hypothesis. Oligohaline to hyperhaline lakes plotted along the left axis and freshwater lakes plotted with the Group I cluster. Three oligohaline surface sediments from interior

539 Canada plotted between the Group I cluster and left axis indicating mixed Group I and
540 Group II producers (Fig. 10b).

5. Discussion

5.1. Genetic and geochemical evidence for a distinct freshwater alkenone producer

A growing number of studies have combined LCA analysis with DNA sequencing to identify relationships between lacustrine haptophyte phylogeny and LCA distributions (Coolen et al., 2004; D'Andrea et al., 2006; Jaraula et al., 2012; Randlett et al., 2014; Theroux et al., 2010). These efforts have lead to the classification of lacustrine Isochrysidales haptophytes into two broad phylogenetic groups including the Group I (or Greenland) phylotype and the Group II phylotype (D'Andrea et al., 2006; Theroux et al., 2010). The Group I and Group II phylotypes feature different distributions (Longo et al., 2013; Theroux et al., 2010) and temperature sensitivities (D'Andrea et al., 2011; Nakamura et al., 2014; Ono et al., 2012; Sun et al., 2007; Theroux et al., 2013; Toney et al., 2012); plotting the available genetic data by salinity indicates that these phylotypes occupy different salinity ranges as well, with Group I haptophytes occurring in fresh to oligohaline lakes (0 - 5 psu) and Group II haptophytes occurring in oligohaline to hyperhaline lakes (>0.5 psu; Fig. 2). The mechanism restricting phylotypes to specific salinity ranges is unknown, but may be related to different evolutionary transitions of the phylotypes from marine to continental environments (Simon et al., 2013). Interestingly, both Group I and Group II phylotypes are found in varied geographic regions and across continents, suggesting that dispersal is not an important factor in determining the presence or absence of different phylotypes (Theroux et al., 2010).

564 Recently, a study by Simon et al. (2013) identified a new lineage of freshwater dwelling
565 Isochrysidales haptophytes – the EV clade. Simon et al. (2013) carried out a phylogenetic
566 analysis that did not include the Group I BrayaSø water sequence (Theroux et al., 2010),
567 which is the only known Group I sequence of comparable length to the EV sequences.
568 We combined 18S rRNA sequences from Simon et al. (2013) with the published
569 representative sequences from Group I (the BrayaSø water sequence) and Group II to
570 compute a new phylogenetic tree (Fig S.2). Our analysis demonstrated that members of
571 the freshwater EV clade branched with the Group I haptophytes and concurrently, that
572 other sequences identified by Simon et al. (2013) from brackish or saline environments,
573 branched with Group II. Sequences from Simon et al. (2013) therefore provide new
574 evidence that the Group I and Group II phylotypes fall within the proposed salinity
575 ranges.

577 **5.2. Environmental controls on freshwater alkenones from northern Alaskan lakes**

579 *5.2.1. Surface sediments*

581 LCAs produced by Group I haptophytes occur in freshwater lakes, however little work
582 has been done to identify the environmental controls on the occurrence, abundance and
583 distribution of these compounds. In northern Alaska, LCAs were detected in 22 of 34
584 freshwater lakes analyzed (Tables 1, S.1), indicating that LCA occurrence was relatively
585 frequent compared with surface sediment surveys from other regions (e.g. Chu et al.,
586 2005; Pearson et al., 2008; Toney et al., 2010; Zink et al., 2001). These data suggest

587 alkenone-producing haptophytes can thrive in freshwaters under certain conditions. LCA
588 concentrations were comparable to those reported from English and German freshwater
589 lakes (Cranwell, 1985; Zink et al., 2001) and Spanish saline lakes (Pearson et al., 2008),
590 albeit several orders of magnitude lower than concentrations reported from oligohaline
591 lakes in Greenland (D'Andrea and Huang, 2005), Lake Qing Hai, China (Li et al., 1996)
592 and Lake Van, Turkey (Theil et al., 1997; Table 1). Lakes with extremely low LCA
593 concentrations had the most challenging sample matrices for analysis and LCAs were
594 often completely unresolvable without a saponification reaction to remove co-eluting wax
595 esters and alkyl-esters.

596
597 LCA distributions were consistent among the northern Alaskan freshwater lakes with
598 C_{37:4} dominant distributions and occurrence of the tri-unsaturated isomers in all samples
599 (Fig. 3). Despite substantial gradients in lake properties, the variability in LCA
600 distribution parameters (C₃₇/C₃₈, Me/Et and %C_{37:4}) was significantly lower in the
601 freshwater northern Alaskan lakes than in saline lakes from China (Chu et al., 2005; Liu
602 et al., 2011) and North America (Toney et al., 2010), demonstrating the conserved nature
603 of the Group I-type LCA distributions. RIK₃₇ values were calculated in all samples and
604 provided a new metric for describing LCA distributions. RIK₃₇ values demonstrated that
605 relative abundances of the C_{37:3b} isomers were relatively consistent among freshwater
606 sites (Table 1). These findings, combined with the presence of C₃₈Me LCAs in all
607 freshwater lakes, suggested the presence of a common or highly related Group I alkenone
608 producer.

609

In order to further investigate the environmental controls on LCA occurrence, abundance and distribution, we performed principal components analysis (PCA) using morphometric, physical and chemical properties of the study lakes as well as LCA distribution parameters and concentrations from the surface sediments. Before considering the LCA distributions, we first performed a PCA using only lake properties in order to describe the general variability differentiating the lakes (Fig. 4a; Table S.2). The PCA demonstrated that lakes with relatively high pH, alkalinity and conductivity were associated with the highest C₃₇ concentrations. This suggests that even in freshwaters with low solute concentrations, water chemistry affects LCA occurrence and abundance. [SO₄²⁻] and alkalinity varied together ($r^2 = 0.80$) making it difficult to elucidate individual controls of [SO₄²⁻] and [CO₃²⁻] on LCAs, as has been done in saline lakes (Chu et al., 2005; Pearson et al., 2008; Toney et al., 2010). Mean depth appeared to play an independent and significant role in determining LCA occurrence and abundance, with deeper lakes featuring higher C₃₇ concentrations. Lake depth is potentially important for a proposed vegetative stage in the haptophyte life cycle (Billard 1994; Toney et al., 2010, 2012), which may occur in the benthos. Notably, lakes that plotted in the lower right quadrant were nearly completely devoid of LCAs. These lakes had relatively low pH values, high C:N ratios and high concentrations of DOC and Chla, indicating that lakes with more dystrophic characteristics may be inhospitable to freshwater alkenone-producing haptophytes.

We performed a second PCA with LCA abundance and distribution parameters plotted *a posteriori* as vectors in the ordination space (Fig. 4b; Table S.2). The second PCA

demonstrated that C₃₇Me alkenones are favored over C₃₈Et alkenones in lakes with high LCA concentrations and relatively high pH, conductivity, alkalinity, [SO₄²⁻] and mean depth. Conte et al. (1998) found that in culture, *E. huxleyi* and *G. oceanica* produced higher proportions of C₃₈Et LCAs after entering logarithmic growth phase. Potentially, the LCA distributions in northern Alaskan lake sediments reflect haptophyte bloom dynamics, such that deep alkaline lakes with longer and larger haptophyte blooms, are characterized by higher sediment LCA concentrations and higher proportions of the C₃₇ LCAs. The PCA also revealed a strong relationship between U₃₇^K, DOC and C:N ratios. Depending on catchment and biome characteristics, high-latitude lakes can be subject to large inputs of terrestrial organic matter-derived DOC with high C:N ratios (Aitkenhead and McDowell, 2003). In the northern Alaskan transect, lakes with high DOC also tended to be more shallow and affected by high light extinction coefficients (ARC LTER Database). These properties would result in more efficient warming of the photic zone in the early summer (Wetzel, 2001) and thus are consistent with higher U₃₇^K values.

5.2.2. Toolik Lake SPM

LCAs were relatively abundant in the water column of Toolik Lake with concentrations comparable to those from Lake George, ND (Toney et al., 2010), and much greater than concentrations from Chinese saline lakes (Liu et al., 2011). The 4-fold higher LCA concentrations in Toolik Lake during the spring isothermal mixing period (Fig. 5) were in agreement with the inverse exponential relationship between water column LCA concentrations and lake stratification described by Toney et al. (2010). Since the LCA-

656 producing haptophytes in Toolik Lake and Lake George branch into different phylotypes
657 (Crump et al., 2012; Theroux et al., 2010; Toney et al., 2012), these results suggest that
658 early season haptophyte/LCA blooms are general characteristics of lacustrine haptophyte
659 ecology that transcend species differences. D'Andrea et al. (2011) found that LCA fluxes
660 to the sediment of Lake BrayaSø, Greenland were highest in early summer (late June and
661 early July), and therefore are consistent with the spring/early summer LCA production
662 identified in SPM from Toolik Lake and Lake George. Results from all three sites
663 indicate that LCAs record spring/early summer lake water temperatures.

664
665 LCA distributions from Toolik Lake SPM were relatively consistent and characteristic of
666 previously described Group I-type LCA distributions (Fig. 3; Longo et al., 2013). C_{37}/C_{38}
667 and Me/Et ratios differed between SPM and surface sediments, a feature that has been
668 observed in other lakes (Liu et al., 2011). The Toolik Lake surface sediment C_{37}/C_{38} ratio
669 was more than 2σ below the average values from Toolik Lake SPM. In marine species,
670 C_{38} Et homologs are preferentially produced after haptophytes transition from exponential
671 growth phase to late logarithmic or stationary phase (Conte et al., 1998). SPM samples
672 primarily incorporate LCAs during the haptophyte bloom period, which is potentially
673 reflected in the higher C_{37}/C_{38} ratios. Decreases in C_{37}/C_{38} and Me/Et ratios throughout
674 the Toolik Lake SPM sampling period further indicated that bloom dynamics affect LCA
675 distributions, with 'bloom' conditions during spring isothermal mixing favoring
676 production of the C_{37} Me homologs.

678 5.3. Temperature calibration for Group I freshwater alkenones

679
680 *5.3.1. Fractional abundance and unsaturation indices in Toolik Lake*
681

682 The GC separation of LCAs employed in this study resolves the full class of known C₃₇ -
683 C₃₉ LCAs. The separation and detection of all LCAs allowed for the treatment of LCAs
684 as fractional abundances, as is customarily done in studies of other biomarkers for
685 temperature calibration and chemotaxonomic purposes (e.g. Kim et al., 2010; Peterse et
686 al., 2012; Rampen et al., 2010; 2012). The significant correlation of several LCAs with *in*
687 *situ* temperature suggested temperature is a dominant variable controlling LCA
688 production in Toolik Lake (Fig. 6). Fractional abundances of tetra-unsaturated LCAs
689 were either negatively or un- correlated with temperature, whereas all di-unsaturated
690 LCAs were positively correlated with temperature. Interestingly, tri-unsaturated LCAs
691 were both positively and negatively correlated with temperature and in the case of C₃₈Me
692 and C₃₉Et LCAs, the sign of the correlation differed between the isomers (Fig 6).
693 Interestingly, C_{38:3b}Et showed no relationship with temperature, despite the fact that this
694 homolog makes up a significant fraction of the LCAs in any given sample (8.5-11%).
695 Environmental data do not point to a clear cause for the differential temperature
696 responses of individual LCAs, which may result from aspects of LCA biosynthesis
697 beyond the scope of this study.

698
699 C_{37:4} is the most abundant homolog in the northern Alaskan lakes and its fractional
700 abundance shows the largest absolute change, providing rationale for the use of U^K
701 indices over U^{K'} indices for temperature calibration and reconstruction. Additionally, the

C₃₇ homologs had the strongest correlations with temperature when considered as a structural group. This suggested that for Group-I type LCA distributions, variations of the U₃₇^K index are more appropriate for temperature calibration than U₃₈^K indices or other subsets of LCAs. Regardless, several unsaturation indices could be calibrated to *in situ* temperature with relatively low error and notably, the U₃₇^{K''} index (Zheng et al., 2016) performed as well as U₃₇^K (Table 2). While U₃₇^{K''} was shown to predict temperature better for Group II haptophytes and U₃₇^{K'} was confirmed as the strongest predictor for Group III (Zheng et al., 2016), our data indicate that both U₃₇^K and U₃₇^{K''} are robust predictors of temperature for the Group I haptophytes. As a whole, SPM data supported that LCA unsaturation is well correlated with temperature and they illustrate the differential temperature responses of various tri-unsaturated LCAs.

5.3.2. Toolik Lake *in situ* U₃₇^K calibration

Site- and species-specific U₃₇^K temperature calibrations are needed for quantitative LCA paleotemperature reconstruction from lakes (Castaneda and Shouten, 2011). The Toolik Lake calibration presented here (Fig. 7) is the first *in situ* U₃₇^K calibration for northern Alaska and the first *in situ* calibration from a freshwater lake. The error of the calibration (RMSE = ±1.37 °C) is smaller than that of the global core-top SST calibration (Muller et al., 1998) and comparable to or smaller than other methods for quantitative lacustrine temperature calibration (e.g. Larocque et al., 2001; Loomis et al., 2012; Pearson et al., 2011). Importantly, the only Isochrysidales 18S rRNA gene sequence detected in the water column of Toolik Lake shows 100% similarity to Group I sequences from Lake

BrayaSø water in Greenland (D'Andrea et al., 2006; Crump et al., 2012; Theroux et al., 2010). Therefore the Toolik Lake calibration expands LCA temperature reconstructions into freshwater lakes of a new geographic region, while improving upon the Group I calibration from Greenland, which mixes samples from Lake BrayaSø with those of unknown haptophyte precursors (D'Andrea et al., 2011; Zink et al., 2001).

The Toolik Lake calibration spanned 8 discrete sampling days that occurred during various stages in the seasonal stratification status of the lake and the SPM samples were taken from several depths throughout the photic zone. The comprehensive ranges in sample depth and sampling date provided strong evidence that the U_{37}^K temperature response is robust to changes in mixing dynamics, light levels and other environmental changes associated with the transition from spring to summer. Studies have shown that salinity changes can alter LCA distributions, particularly involving the $C_{37:4}$ LCA (Liu et al., 2011). The Toolik Lake SPM dataset includes samples from a narrow range in lake conductivities ($<100 \mu S cm^{-1}$), effectively ruling out salinity as a confounding variable in the temperature calibration. While the calibration performs well in the temperature range of 3 to 12 °C, there is more variability in the U_{37}^K index at higher temperatures, with a notable potential outlier at 12.4 °C. This sample was taken from the top of the thermocline on the warmest sampling day of the calibration; its anomalously warm U_{37}^K value may have originated from a later-season bloom of haptophytes acclimated to a warmer water column.

The separation of the tri-unsaturated isomers calls for revised comparisons between lacustrine U_{37}^K temperature calibrations. Two forms of the U_{37}^K index (U_{37ab}^K and U_{37a}^K ; section 3.3) have been applied interchangeably, however comparisons between different forms of the index are problematic because they include different compounds. Furthermore, several published U_{37}^K calibrations feature non-normal data and show major differences in their temperature ranges and variances, precluding comparisons based on parametric statistics. We addressed these issues by performing bootstrapped regressions on published U_{37}^K calibration data and comparing them to bootstrapped regressions of U_{37ab}^K and U_{37a}^K from Toolik Lake.

Our analysis demonstrated that the Toolik Lake and Lake BrayaSø (D'Andrea et al., 2011) U_{37ab}^K calibrations had comparable slope coefficients, which was consistent with the identification of the Group I alkenone-producers in both lakes. Interestingly, the intercepts were significantly different (Figs. 8, S.3). While the lakes occur at similar latitudes and biomes, they differ by several orders of magnitude in their conductivity. This suggests solutes may play a role in determining calibration intercept, whereas they do not appear to affect the slope. While the intercepts are needed for absolute temperature reconstruction, they are not necessary for the calculation of temperature anomalies (assuming proxy stationarity), which are frequently used for temperature reconstruction. A recent U_{37}^K synthesis by D'Andrea et al. (2016) used a statistical analysis to compare calibrations and argue for phylogenetic control over U_{37}^K temperature sensitivity. The Toolik Lake calibration shows general agreement with these findings, however suggests a

wider range of Group I U_{37}^K temperature sensitivities that allows for smaller slope coefficients.

We used the U_{37a}^K index for comparison of the Toolik Lake calibration to various Group II calibrations, because it excludes $C_{37:3b}$, which appears to be universally absent in Group II samples. The slope for the *R. lamellosa* calibration and the intercepts for both the *R. lamellosa* and *P. paradoxa* calibrations differed significantly from the Toolik U_{37a}^K calibration, which was expected based on their different LCA producers (Figs. 8; S.3). By contrast, the Toolik Lake and Lake George U_{37a}^K calibrations were comparable in both slope and intercept (Figs. 8; S.3) despite their distantly related precursors (Theroux et al., 2010; Toney et al., 2012). The similarity of these calibrations suggests that in some cases LCA temperature responses can be shared across phylotype, and it underscores the importance of comparing analogous U^K indices for interpretation of LCA temperature calibrations.

5.3.3. *RIK indices as temperature indicators*

To our knowledge, all previous LCA-based temperature calibrations and reconstructions have relied on unsaturation as the fundamental property governing LCA temperature dependence. The SPM dataset from Toolik Lake provided the opportunity to evaluate double bond position as a separate and independent parameter for LCA temperature dependence. Tri-unsaturated isomers within each structural group of LCAs differ by double bond position, with the normal isomer featuring $\Delta^{7,14,21}$ double bond positioning

and the new isomer featuring $\Delta^{14,21,28}$ positioning (Fig. S.1; Dillon et al., 2016; Longo et al., 2013). RIK indices (defined in section 3.3) provide a metric of the abundance of isomers relative to each other. Significant correlations of all RIK indices with temperature provided evidence that in addition to unsaturation, LCA double bond position responds to *in situ* temperature (Fig. 9).

The strong correlation between $\text{RIK}_{38\text{E}}$ and *in situ* temperature highlights its potential utility as a temperature proxy. $\text{RIK}_{38\text{E}}$ may hold benefits over the U^{K} indices in environments where the preservation potential of LCAs is low. $\text{RIK}_{38\text{E}}$ is composed of LCAs with identical numbers of double bonds, therefore non-selective biotic degradation of LCAs should affect the isomers equally, maintaining the integrity of the $\text{RIK}_{38\text{E}}$ index. Rontani et al. (2008) found evidence for non-selective LCA degradation in incubations of *Emiliana huxleyi* with bacterial communities, however the authors also found evidence of selective, site-specific degradation favoring attack on the Δ^7 double bond. These selective biodegradative pathways would affect both the U^{K} and RIK indices. Abiotic degradation pathways for LCAs have been less studied, however photolysis of LCAs was found to be negligible (Rontani et al., 1997). Autoxidation appeared to be selective for the Δ^{21} double bond (Rontani et al., 2013) and therefore would also affect both the RIK and U^{K} indices.

5.4. Chemotaxonomic biomarkers for Group I haptophytes

5.4.1. Evidence from haptophyte cultures and surface sediments

815

816 The $\Delta^{14,21,28}$ tri-unsaturated isomers appear to be specific to Group I haptophytes based on
817 a limited dataset (D'andrea et al., 2016; Longo et al., 2013) and we presented new culture
818 data that corroborate this hypothesis. Four new cultures presented in this study, in
819 addition to *P. paradoxa* culture data (Theroux et al., 2013) and additional *I. galbana* and
820 *E. huxleyi* culture data (Longo et al., 2013) demonstrated that while certain Group II
821 haptophytes produced $C_{38:3b}$ Et isomers, none of the Group II or Group III species
822 produced the $C_{37:3b}$ isomers (Fig. S.4; Table S.4). These culture experiments spanned a
823 significant temperature range and included organisms isolated from estuaries, coastal seas
824 and from the Atlantic and Pacific Ocean basins; therefore they provide substantial
825 evidence that $C_{37:3b}$ LCAs are not produced by Group II and III haptophytes and thus are
826 biomarkers for the Group I phylotype.

827

828 Surface sediments from northern Alaska, interior Canada (Toney et al., 2011) and China
829 (Zhao et al., 2014) spanned a salinity range of 0 - 146 psu and they provided an
830 opportunity to test whether or not the tri-unsaturated isomers followed predictable trends
831 based on inferred relationships between salinity and haptophyte phylogeny (Fig. 2). The
832 freshwater lakes exclusively hosted Group I-type LCA distributions, including the $C_{37:3b}$
833 LCAs. As predicted by genetic data, oligohaline lakes hosted both Group I- and Group II-
834 type distributions, whereas mesohaline to hyperhaline lakes exclusively hosted Group II-
835 type LCA distributions with no detectable $C_{37:3b}$ LCAs. Recent surface sediment data
836 from NW China corroborated these findings as the $C_{37:3b}$ isomers were detected in fresh
837 and oligohaline lakes, Shuanghu and Sailimu, but were absent from lakes with elevated

838 salinity (Song et al., 2016). While it is difficult to tease apart direct effects of salinity vs.
839 phylogeny on the relative abundance of C_{37:3b}, our data indicated no correlation between
840 salinity and RIK₃₇ (Fig. S.5). This suggests that changes in the relative abundance of
841 C_{37:3b} do not represent a direct physiological response to changing salinity, but rather
842 represent lake-to-lake differences in the relative abundances of Group I vs. Group II
843 haptophytes. Future studies are needed to clarify the environmental controls on C_{37:3b}
844 production.

845
846 The compilation of 124 discrete LCA samples measured with separation of the tri-
847 unsaturated isomers revealed clear patterns between phylogeny, salinity and LCA
848 distributions. Theroux et al. (2010) used combined genetic and geochemical analyses to
849 show that LCA distributions were not adequate for taxonomic identification of
850 haptophyte algae, however those analyses were carried out before the tri-unsaturated
851 isomers were separated and discovered. The new separation method employed here
852 (Longo et al., 2013) revealed clear and diagnostic links between haptophyte phylogeny
853 and LCA distributions, particularly in relation to the presence/absence and abundance of
854 the tri-unsaturated isomers (Fig. 10). Our results indicate that LCA distributions can be
855 used to differentiate broad phylogenetic groups, which is adequate for evaluating species
856 effects in freshwater or oligohaline lakes where Group I haptophytes are dominant.
857 Oligohaline to hyperhaline lakes hosting Group II haptophytes show higher variability in
858 LCA distributions making it more difficult to access species effects for these
859 distributions.

860

5.4.2. RIK_{37} as a multi-species indicator

The distinction of $C_{37:3b}$ as a Group I biomarker is useful for confirming the presence of Group I haptophytes in lakes where only one haptophyte species exists, however it cannot alone indicate relative contributions of Group I and Group II species within a lake. Here we propose to use the RIK_{37} index as a taxonomic indicator for multi-species effects (Table 3). Among Group I-type samples from BrayaSø, Toolik Lake and the northern Alaskan surface sediments, the RIK_{37} index changed only modestly (0.51 to 0.60) across a large gradient in temperature and lake properties. By contrast, RIK_{37} for all known Group II and Group III samples is 1.0. Therefore, lakes with significant LCA contributions from multiple phylotypes (e.g. Group I and Group II) should exhibit RIK_{37} values between 0.60 and 1.0. We hypothesize that lakes with RIK_{37} values of 0.51 to 0.60 host Group I haptophytes, exclusively (Table 3). The three oligohaline lake samples from interior Canada with RIK_{37} values between 0.60 and 1.0 therefore, likely contain both Group I and Group II haptophytes, which would be consistent with their oligohaline classification and positioning on the ternary diagram (Table 3; Fig. 10b). These surface sediment, SPM and culture samples provide a framework for interpreting RIK_{37} values as multi-species indicators in sedimentary records.

The demonstrated taxonomic significance of $C_{37:3b}$ encourages investigation of the biosynthetic origin of the $\Delta^{14,21,28}$ isomers. An independent biosynthetic pathway involving desaturation of the Δ^{28} double bond prior to the Δ^7 double bond was proposed to explain the occurrence of the tri-unsaturated isomers (Brassell, 2014). A recent study

884 reporting the carbon and hydrogen isotopic composition of tri-unsaturated isomers
885 isolated from Lake BrayaSø sediment suggests the biosynthesis of the $\Delta^{14,21,28}$ isomers
886 may be more distinct (Dillon et al., 2016). The apparent taxonomic specificity of the
887 isomers could therefore result from the evolution of a different $\Delta^{14,21,28}$ biosynthetic
888 pathway after the Group I haptophyte lineage split from the Group II/III
889 prymnesiophytes. The origin of these isomers cannot be fully assessed, however, until the
890 Group I haptophyte producers of these compounds are isolated in pure culture.

6. Conclusions

LCAAs from lacustrine environments show promise as continental paleotemperature proxies, however species effects, which are frequently observed in saline lakes, have hindered their broad application. Here, we present a comprehensive review and analysis of genetic and geochemical data that suggests freshwater lakes are an exception. Freshwater lakes host Group I haptophytes and appear to be devoid of Group II haptophytes, encouraging their further investigation as important sites for LCA paleotemperature reconstruction.

Our investigation of surface sediments and SPM demonstrated the temperature dependence and environmental controls on northern Alaskan freshwater LCAs. Occurrence and abundance of LCAs were positively correlated with pH, alkalinity, conductivity and mean depth and their distributions were relatively consistent across large gradients in several lake properties. We provided the first fresh water *in situ* alkenone calibration ($U_{37}^K = 0.021 \cdot T - 0.68$; $r^2 = 0.85$; $n = 52$; $RMSE = \pm 1.37$ °C), which improves the temperature range of the previous Group I haptophyte calibration. We used bootstrapping analyses to compare U_{37}^K calibrations and show that when the tri-unsaturated alkenone isomers are considered separately, the Toolik freshwater calibration is comparable to calibrations from Group I haptophytes and a $C_{37:4}$ dominant Group II haptophyte.

913 Our analyses demonstrated the paleoclimatic and taxonomic significance of tri-
914 unsaturated alkenone isomers, which are produced by Group I haptophytes and globally
915 distributed in high latitude freshwater lakes. The isomer ratio, RIK_{38E} was well correlated
916 with *in situ* temperature, whereas RIK_{37} was a taxonomic indicator for the Group I
917 phylotype. We demonstrated that RIK_{37} can be used as a multiple species indicator with
918 RIK_{37} values of 0.51 to 0.60 indicating a Group I haptophyte producer, whereas values
919 between 0.60 and 1.0 are indicative of mixed phylotypes. Importantly, this ratio can be
920 used to test for the presence of species effects in sedimentary records and may be used to
921 determine percent contributions from Group I and Group II haptophytes, guiding
922 alkenone calibration selection for paleotemperature reconstructions.

Acknowledgements

This work was funded by NSF grants PLR-1503846 and DEB-1026843, and a GSA Graduate Student Research Grant to W. Longo. The project would not have been possible without the support and limnological monitoring of the Arctic Long Term Ecological Research program. We thank J. Stuckey and the Toolik Field Station GIS office for providing morphometric data on the Arctic LTER lakes. We Thank B. Gaglioti and the Circum-Arctic Lakes Observation Network (CALON) for providing sediments and environmental data and F. S. Hu for providing surface sediment from Perch Lake. We also thank J. Toney for previous work on the interior Canada Lake sediments and for providing samples for new analyses. A. Carter, D. White and W. Daniels provided field assistance and J. Crowther and R. Taroza provided laboratory assistance. B. Fox-Kemper provided advice on statistical methods and L. Amaral provided advice on haptophyte phylogenetics.

938 **References**

- 939 Aitkenhead-Peterson, J., McDowell, W., Neff, J., Stuart, E., & Robert, L., 2003. Sources,
940 production, and regulation of allochthonous dissolved organic matter inputs to
941 surface waters. *Aquatic Ecosystems: Interactivity of Dissolved Organic Matter*;
942 Findlay, SEG, Sinsabaugh, RL, Eds, 25-70.
- 943 ARC LTER database. Available from <http://ecosystems.mbl.edu/ARC/datacatalog.html>.
- 944 Bendif, E.M., Probert, I., Schroeder, D.C., de Vargas, C., 2013. On the description of
945 *Tisochrysis lutea* gen. nov. sp. nov. and *Isochrysis nuda* sp. nov. in the
946 Isochrysidales, and the transfer of *Dicrateria* to the Prymnesiales (Haptophyta). *J.*
947 *Appl. Phycol.* **25**, 1763–1776.
- 948 Billard G. (1994) Life cycles. In *The Haptophyte Algae* (eds. J. C. Green and B. S. C.
949 Leadbeatter). *The Systematics Association Special Volume* **51**, pp. 167–186.
- 950 Braconnot, P., Harrison, S.P., Kageyama, M., Bartlein, P.J., Masson-delmotte, V., Abe-
951 ouchi, A., Otto-bliesner, B., Zhao, Y., 2012. Evaluation of climate models using
952 palaeoclimatic data. *Nat. Clim. Chang.* **2**, 417–424.
- 953 Brassell, S.C., 2014. Climatic influences on the Paleogene evolution of alkenones.
954 *Paleoceanography* **29**, 10.1002/2013PA002576.
- 955 Brassell, S.C., Eglinton, G., Marlowe, I.T., Pflaumann, U., Sarnthein, M., 1986.
956 Molecular stratigraphy: a new tool for climatic assessment. *Nature* **320**, 129–133.
- 957 Burr, J. M. (1995). *Lake trout studies in the AYK Region, and burbot index of abundance*
958 *in Galbraith Lake, 1994*. Alaska Department of Fish and Game, Division of Sport
959 Fish, Research and Technical Services.
- 960 Canty, A., & Ripley, B. (2015). Package ‘boot’.
- 961 Castañeda, I.S., Schouten, S., 2011. A review of molecular organic proxies for examining
962 modern and ancient lacustrine environments. *Quat. Sci. Rev.* **30**, 2851–2891.
- 963 Chu, G., Sun, Q., Li, S., Zheng, M., Jia, X., Lu, C., Liu, J., Liu, T., 2005. Long-chain
964 alkenone distributions and temperature dependence in lacustrine surface
965 sediments from China. *Geochim. Cosmochim. Acta* **69**, 4985–5003.
- 966 Conte, M.H., Sicre, M.-A., Rühlemann, C., Weber, J.C., Schulte, S., Schulz-Bull, D.,
967 Blanz, T., 2006. Global temperature calibration of the alkenone unsaturation
968 index (UK'37) in surface waters and comparison with surface sediments.
969 *Geochemistry, Geophys. Geosystems* **7**, 10.1029/2005GC001054.

- 970 Conte, M.H., Thompson, A., Lesley, D., Harris, R.P., 1998. Genetic and Physiological
971 Influences on the Alkenone/Alkenoate Versus Growth Temperature Relationship
972 in *Emiliania huxleyi* and *Gephyrocapsa Oceanica*. *Geochim. Cosmochim. Acta* **62**,
973 51–68.
- 974 Coolen, M.J.L., Muyzer, G., Rijpstra, W.I.C., Schouten, S., Volkman, J.K., Sinninghe
975 Damsté, J.S., 2004. Combined DNA and lipid analyses of sediments reveal
976 changes in Holocene haptophyte and diatom populations in an Antarctic lake.
977 *Earth Planet. Sci. Lett.* **223**, 225–239.
- 978 Cranwell, P.A., 1985. Long-chain unsaturated ketones in recent lacustrine sediments.
979 *Geochim. Cosmochim. Acta* **49**, 1545–1551.
- 980 Crump, B.C., Amaral-Zettler, L. a, Kling, G.W., 2012. Microbial diversity in arctic
981 freshwaters is structured by inoculation of microbes from soils. *ISME J.* **6**, 1629–
982 39.
- 983 D’Andrea, W.J., Huang, Y., 2005. Long chain alkenones in Greenland lake sediments:
984 Low $\delta^{13}\text{C}$ values and exceptional abundance. *Org. Geochem.* **36**, 1234–1241.
- 985 D’Andrea, W.J., Huang, Y., Fritz, S.C., Anderson, N.J., 2011. Abrupt Holocene climate
986 change as an important factor for human migration in West Greenland. *Proc.*
987 *Natl. Acad. Sci. U. S. A.* **108**, 9765–9.
- 988 D’Andrea, W.J., Lage, M., Martiny, J.B.H., Laatsch, A.D., Amaral-Zettler, L. a., Sogin,
989 M.L., Huang, Y., 2006. Alkenone producers inferred from well-preserved 18S
990 rDNA in Greenland lake sediments. *J. Geophys. Res.* **111**, G03013.
- 991 D’Andrea, W.J., Theroux, S., Bradley, R.S., Huang, X., 2016. Does phylogeny control
992 U_{37}^{K} -temperature sensitivity? Implications for lacustrine alkenone
993 paleothermometry. *Geochim. Cosmochim. Acta* **175**, 168–180.
- 994 D’Andrea, W.J., Vaillencourt, D. a., Balascio, N.L., Werner, a., Roof, S.R., Retelle, M.,
995 Bradley, R.S., 2012. Mild Little Ice Age and unprecedented recent warmth in an
996 1800 year lake sediment record from Svalbard. *Geology* **40**, 1007–1010.
- 997 Davison, A. C., & Hinkley, D. V., 1997. *Bootstrap methods and their application (Vol.*
998 *1)*. Cambridge university press.
- 999 De Leeuw, J.W., van der Meer, F.W., Rijpstra, W.I.C., Schenck, P.A., 1980. On the
1000 occurrence and structural identification of long chain unsaturated ketones and
1001 hydrocarbons in sediments. In: Douglas, A.G., Maxwell, J.R. (Eds.), *Advances in*
1002 *Organic Geochemistry 1979*. Pergamon Press, Oxford, pp. 211–217.
- 1003 Déry, S. J., Stieglitz, M., Rennermalm, Å. K., & Wood, E. F., 2005. The Water Budget of
1004 the Kuparuk River Basin, Alaska. *Journal of Hydrometeorology* **6**, 633–655.

- 1005 Dillon, J. T., Longo, W. M., Zhang, Y., Torozo, R., & Huang, Y., 2016. Identification of
1006 double- bond positions in isomeric alkenones from a lacustrine haptophyte. *Rapid*
1007 *Commun. Mass Spectrom.*, **30**, 112-118.
- 1008 Fox, J., Weisberg, S., 2012. *Bootstrapping Regression Models in R*, 1–17. Available from
1009 [http://socserv.socsci.mcmaster.ca/jfox/Books/Companion/appendix/Appendix-](http://socserv.socsci.mcmaster.ca/jfox/Books/Companion/appendix/Appendix-Bootstrapping.pdf)
1010 [Bootstrapping.pdf](http://socserv.socsci.mcmaster.ca/jfox/Books/Companion/appendix/Appendix-Bootstrapping.pdf)
- 1011 Hamilton, T. D., & Porter, S. C., 1975. Itkillik glaciation in the Brooks Range, northern
1012 Alaska. *Quaternary Research* **5**, 471-497.
- 1013 Herbert, T.D., 2003. Alkenone Paleotemperature Determinations, in: Holland, H.D.,
1014 Turekian, K.K. (Eds.), *Treatise in Marine Geochemistry. The Ocean and Marine*
1015 *Geochemistry*. El Sevier Ltd., pp. 391–432.
- 1016 Huryn, A., & Hobbie, J., 2012. *Land of extremes: a natural history of the Arctic North*
1017 *Slope of Alaska*. University of Alaska Press.
- 1018 Jaraula, C.M.B., Brassell, S.C., Morgan-Kiss, R.M., Doran, P.T., Kenig, F., 2010. Origin
1019 and tentative identification of tri to pentaunsaturated ketones in sediments from
1020 Lake Fryxell, East Antarctica. *Org. Geochem.* **41**, 386–397.
- 1021 Juggins, S., 2013. Quantitative reconstructions in palaeolimnology: New paradigm or
1022 sick science? *Quat. Sci. Rev.* **64**, 20–32.
- 1023 Kim, J.H., van der Meer, J., Schouten, S., Helmke, P., Willmott, V., Sangiorgi, F., Koç,
1024 N., Hopmans, E.C., Damsté, J.S.S., 2010. New indices and calibrations derived
1025 from the distribution of crenarchaeal isoprenoid tetraether lipids: Implications for
1026 past sea surface temperature reconstructions. *Geochim. Cosmochim. Acta* **74**,
1027 4639–4654.
- 1028 Larocque, I., Hall, R. I., & Grahn, E., 2001. Chironomids as indicators of climate change:
1029 a 100- lake training set from a subarctic region of northern Sweden (Lapland).
1030 *Journal of Paleolimnology* **26**, 307-322.
- 1031 Lepš, J., & Šmilauer, P., 2003. *Multivariate analysis of ecological data using CANOCO*.
1032 Cambridge university press.
- 1033 Li, J., Philp, R.P., Pu, F., Allen, J., 1996. Long-chain alkenones in Qinghai Lake
1034 sediments. *Geochim. Cosmochim. Acta* **60**, 235–241.
- 1035 Liu, W., Liu, Z., Wang, H., He, Y., Wang, Z., Xu, L., 2011. Salinity control on long-
1036 chain alkenone distributions in lake surface waters and sediments of the northern
1037 Qinghai-Tibetan Plateau, China. *Geochim. Cosmochim. Acta* **75**, 1693–1703.

- 1038 Longo, W.M., Dillon, J.T., Tarozo, R., Salacup, J.M., Huang, Y., 2013. Unprecedented
1039 separation of long chain alkenones from gas chromatography with a
1040 poly(trifluoropropylmethylsiloxane) stationary phase. *Org. Geochem.* **65**, 94–102.
- 1041 Loomis, S.E., Russell, J.M., Ladd, B., Street-Perrott, F.A., Sinninghe Damsté, J.S., 2012.
1042 Calibration and application of the branched GDGT temperature proxy on East
1043 African lake sediments. *Earth Planet. Sci. Lett.* **357-358**, 277–288.
- 1044 Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A.,
1045 Lai, T., Steppi, S., Jobb, G., Forster, W., Brettske, I., Gerber, S., Ginhart, A.W.,
1046 Gross, O., Grumann, S., Hermann, S., Jost, R., Konig, A., Liss, T., Lussmann, R.,
1047 May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N.,
1048 Vilbig, A., Lenke, M., Ludwig, T., Bode, A., Schleifer, K.H., 2004. ARB: a
1049 software environment for sequence data. *Nucl. acids res.* **32**, 1363-1371.
- 1050 M'boule, D., Chivall, D., Sinke-Schoen, D., Damsté, J. S. S., Schouten, S., & van der
1051 Meer, M. T., 2014. Salinity dependent hydrogen isotope fractionation in
1052 alkenones produced by coastal and open ocean haptophyte algae. *Geochim.*
1053 *Cosmochim. Acta* **130**, 126-135.
- 1054 McColl, J. L., Couto, J., Bendle, J. A., Henderson, A. C., Seki, O., Phoenix, V. R., &
1055 Toney, J. L. 2013, December. 18S rDNA analysis of alkenone-producing
1056 haptophyte (s) preserved in surface sediments of Lake Toyoni, Japan. In *AGU*
1057 *Fall Meeting Abstracts* (Vol. 1, p. 2084).
- 1058 Müller, P.J., Kirst, G., Ruhland, G., Von Storch, I., Rosell-Melé, A., 1998. Calibration of
1059 the alkenone paleotemperature index UK'37 based on core-tops from the eastern
1060 South Atlantic and the global ocean (60°N-60°S). *Geochim. Cosmochim. Acta* **62**,
1061 1757–1772.
- 1062 Nakamura, H., Sawada, K., Araie, H., Suzuki, I., Shiraiwa, Y., 2014. Long chain alkenes,
1063 alkenones and alkenoates produced by the haptophyte alga *Chrysotila lamellosa*
1064 CCMP1307 isolated from a salt marsh. *Org. Geochem.* **66**, 90–97.
- 1065 Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B.,
1066 Simpson, G.L., Solymos, P., Henry, Stevens, M.H.H., Wagner, H., 2015. *Package*
1067 *'vegan'*. *Community ecology package*, version, 2-2.
- 1068 Ono, M., Sawada, K., Shiraiwa, Y., Kubota, M., 2012. Changes in alkenone and
1069 alkenoate distributions during acclimatization to salinity change in *Isochrysis*
1070 *galbana*: Implication for alkenone-based paleosalinity and paleothermometry.
1071 *Geochem. J.* **46**, 235–247.

- 1072 Pearson, E.J., Juggins, S., Farrimond, P., 2008. Distribution and significance of long-
1073 chain alkenones as salinity and temperature indicators in Spanish saline lake
1074 sediments. *Geochim. Cosmochim. Acta* **72**, 4035–4046.
- 1075 Pearson, E.J., Juggins, S., Talbot, H.M., Weckström, J., Rosén, P., Ryves, D.B., Roberts,
1076 S.J., Schmidt, R., 2011. A lacustrine GDGT-temperature calibration from the
1077 Scandinavian Arctic to Antarctic: Renewed potential for the application of
1078 GDGT-paleothermometry in lakes. *Geochim. Cosmochim. Acta* **75**, 6225–6238.
- 1079 Peterse, F., van der Meer, J., Schouten, S., Weijers, J.W.H., Fierer, N., Jackson, R.B.,
1080 Kim, J.-H., Sinninghe Damsté, J.S., 2012. Revised calibration of the MBT–CBT
1081 paleotemperature proxy based on branched tetraether membrane lipids in surface
1082 soils. *Geochim. Cosmochim. Acta* **96**, 215–229.
- 1083 Prahl, F.G., Muehlhausen, L.A., Zahnle, D.L., 1988. Further evaluation of long-chain
1084 alkenones as indicators of paleoceanographic conditions. *Geochim. Cosmochim.*
1085 *Acta* **52**, 2303–2310.
- 1086 Prahl, F.G., Wakeham, S.G., 1987. Calibration of unsaturation patterns in long-chain
1087 ketone compositions for palaeotemperature assessment. *Nature* **330**, 367–369.
- 1088 Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W.G., Peplies, J., Glöckner,
1089 F.O., 2007. SILVA: a comprehensive online resource for quality checked and
1090 aligned ribosomal RNA sequence data compatible with ARB. *Nucl. Acids Res.* **35**,
1091 7188–7196
- 1092 Rampen, S.W., Abbas, B.A., Schouten, S., Damsté, J.S.S., 2010. A comprehensive study
1093 of sterols in marine diatoms (Bacillariophyta): Implications for their use as
1094 tracers for diatom productivity. *Limnol. Oceanogr.* **55**, 91–105.
- 1095 Rampen, S.W., Willmott, V., Kim, J.-H., Uliana, E., Mollenhauer, G., Schefuß, E.,
1096 Sinninghe Damsté, J.S., Schouten, S., 2012. Long chain 1,13- and 1,15-diols as a
1097 potential proxy for palaeotemperature reconstruction. *Geochim. Cosmochim. Acta*
1098 **84**, 204–216.
- 1099 Randlett, M.-È., Coolen, M.J.L., Stockhecke, M., Pickarski, N., Litt, T., Balkema, C.,
1100 Kwiecien, O., Tomonaga, Y., Wehrli, B., Schubert, C.J., 2014. Alkenone
1101 distribution in Lake Van sediment over the last 270 ka: influence of temperature
1102 and haptophyte species composition. *Quat. Sci. Rev.* **104**, 53–62.
- 1103 Ripley, B., 2007. boot: Bootstrap R (S-Plus) Functions (Canty). R package version, 1-2.
1104 Available from <http://astrostatistics.psu.edu/su07/R/html/boot/html/00Index.html>

- 1105 Rontani, J.-F., Harji, R., Guasco, S., Prahl, F.G., Volkman, J.K., Bhosle, N.B., Bonin, P.,
1106 2008. Degradation of alkenones by aerobic heterotrophic bacteria: Selective or
1107 not? *Org. Geochem.* **39**, 34–51.
- 1108 Rontani, J.-F., Volkman, J.K., Prahl, F.G., Wakeham, S.G., 2013. Biotic and abiotic
1109 degradation of alkenones and implications for U37K' paleoproxy applications: A
1110 review. *Org. Geochem.* **59**, 95–113.
- 1111 Rontani, J.F., Cuny, P., Grossi, V., Beker, B., 1997. Stability of long-chain alkenones in
1112 senescing cells of *Emiliana huxleyi*: Effect of photochemical and aerobic
1113 microbial degradation on the alkenone unsaturation ratio (Uk37)). *Org. Geochem.*
1114 **26**, 503–509.
- 1115 Simon, M., López-García, P., Moreira, D., Jardillier, L., 2013. New haptophyte lineages
1116 and multiple independent colonizations of freshwater ecosystems. *Environ.*
1117 *Microbiol. Rep.* **5**, 322–32.
- 1118 Song, M., Zhou, A., He, Y., Zhao, C., Wu, J., Zhao, Y., Liu, W., & Liu, Z., 2016.
1119 Environmental controls on long-chain alkenone occurrence and compositional
1120 patterns in lacustrine sediments, northwestern China. *Org. Geochem.* **91**, 43–53.
- 1121 Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-
1122 analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313.
- 1123 Sun, Q., Chu, G., Liu, G., Li, S., Wang, X., 2007. Calibration of alkenone unsaturation
1124 index with growth temperature for a lacustrine species, *Chrysotila lamellosa*
1125 (Haptophyceae). *Org. Geochem.* **38**, 1226–1234.
- 1126 Theroux, S., D'Andrea, W.J., Toney, J., Amaral-Zettler, L., Huang, Y., 2010.
1127 Phylogenetic diversity and evolutionary relatedness of alkenone-producing
1128 haptophyte algae in lakes: Implications for continental paleotemperature
1129 reconstructions, *Earth Planet. Sci. Lett.* **300**, 311–320.
- 1130 Theroux, S., Toney, J., Amaral-Zettler, L., Huang, Y., 2013. Production and temperature
1131 sensitivity of long chain alkenones in the cultured haptophyte *Pseudoisochrysis*
1132 *paradoxa*. *Org. Geochem.* **62**, 68–73.
- 1133 Thiel, V., Jenisch, A., Landmann, G., Reimer, A., Michaelis, W., 1997. Unusual
1134 distributions of long-chain alkenones and tetrahymanol from the highly alkaline
1135 Lake Van, Turkey. *Geochim. Cosmochim. Acta* **61**, 2053–2064.
- 1136 Toney, J.L., Huang, Y., Fritz, S.C., Baker, P.A., Grimm, E., Nyren, P., 2010. Climatic
1137 and environmental controls on the occurrence and distributions of long chain
1138 alkenones in lakes of the interior United States. *Geochim. Cosmochim. Acta* **74**,
1139 1563–1578.

- 1140 Toney, J.L., Leavitt, P.R., Huang, Y., 2011. Alkenones are common in prairie lakes of
1141 interior Canada. *Org. Geochem.* **42**, 707–712.
- 1142 Toney, J.L., Theroux, S., Andersen, R.A., Coleman, A., Amaral-Zettler, L., Huang, Y.,
1143 2012. Culturing of the first 37:4 predominant lacustrine haptophyte: Geochemical,
1144 biochemical, and genetic implications. *Geochim. Cosmochim. Acta* **78**, 51–64.
- 1145 Venice system, 1959. The Venice System for the Classification. *Arch. Oceanogr. Limnol.*
1146 **11**, 243–248.
- 1147 Volkman, J.K., Burton, H.R., Everitt, D.A., Allen, D.I., 1988. Pigment and lipid
1148 compositions of algal and bacterial communities in Ace Lake, Vestfold Hills,
1149 Antarctica, in: Ferris, J.M., Burton, H.R., Johnstone, G.W., Bayly, I.A.E. (Eds.),
1150 Biology of the Vestfold Hills, Antarctica SE - 5, *Developments in Hydrobiology*.
1151 Springer Netherlands, pp. 41–57.
- 1152 Wang, Z., Liu, Z., Zhang, F., Fu, M., An, Z., 2015. A new approach for reconstructing
1153 Holocene temperatures from a multi-species long chain alkenone record from
1154 Lake Qinghai on the northeastern Tibetan Plateau. *Org. Geochem.* **88**, 50–58.
- 1155 Wetzel, R. G. (2001). *Limnology: lake and river ecosystems*. Gulf Professional
1156 Publishing.
- 1157 Zhao, J., An, C., Longo, W.M., Dillon, J.T., Zhao, Y., Shi, C., Chen, Y., Huang, Y.,
1158 2014. Occurrence of extended chain length C41 and C42 alkenones in hypersaline
1159 lakes. *Org. Geochem.* **75**, 48–53.
- 1160 Zheng, Y., Huang, Y., Andersen, R. A., & Amaral-Zettler, L. A., 2016. Excluding the di-
1161 unsaturated alkenone in the U^{K}_{37} index strengthens temperature correlation for the
1162 common lacustrine and brackish-water haptophytes. *Geochim. Cosmochim. Acta*
1163 **175**, 36–46.
- 1164 Zink, K.G., Leythaeuser, D., Melkonian, M., Schwark, L., 2001. Temperature
1165 dependency of long-chain alkenone distributions in Recent to fossil limnic
1166 sediments and in lake waters. *Geochim. Cosmochim. Acta* **65**, 253–265.

1167

1168 Figure Captions

1169 **Fig. 1.** Map of study region. Numbers correspond to lake names from table 1. Green filled circles represent
1170 freshwater lakes in which LCAs were detected, blue filled circles represent saline lakes (n=1) in which
1171 LCAs were detected, and circles with an “X” represent lakes in which LCAs were not detected.

1172 **Fig. 2.** Salinity ranges for alkenone-producing haptophyte algae by species and phylogeny. Shaded regions
1173 represent salinity ranges where haptophytes from specific phylogenetic groups are commonly found in
1174 nature. Black bars represent the salinity ranges for haptophyte species based on experimental culture
1175 conditions, or ranges for operational taxon units based on 18S rRNA analysis of environmental samples
1176 from multiple lakes. Filled black circles represent reported salinities of lakes in which specific phylogenetic
1177 clades are dominant. (1, Crump et al., 2012; 2, Simon et al., 2013; 3, Theroux et al., 2010; 4, D’Andrea et
1178 al., 2006; 5, Jaraula et al., 2010; 6, Toney et al., 2012; 7, Randlett et al., 2014; 8, Sun et al., 2007; 9,
1179 M’boule et al., 2014).

1180 **Fig. 3.** Average fractional abundances of LCAs in (a) Toolik Lake SPM (n=52) and (b) surface sediments
1181 in which all 16 LCAs were resolved (n=9). Partial chromatograms of LCAs in Toolik Lake SPM from (c)
1182 12.3 °C and (d) 3.1 °C water samples are shown on the left. Partial chromatograms of LCAs from surface
1183 sediments in (e) Lake E5 and (f) Lake IKP-003 (the only saline lake in the transect) are shown on the right.

1184 **Fig. 4.** (a) PCA biplot constructed from morphometric, physical and chemical properties of northern
1185 Alaskan lakes. Individual lakes are plotted as circles or “X” symbols. Circles scale to surface sediment C_{37}
1186 concentrations ($\mu\text{g gTOC}^{-1}$) and “X” symbols represent lakes in which LCAs were not detected. (b) PCA
1187 biplot constructed from the same lakes properties, but excluding lakes where LCAs were not detected. C_{37}
1188 concentrations ($\mu\text{g gTOC}^{-1}$), and LCA distribution parameters are plotted as vectors in the ordination space
1189 based on linear correlations with the principal components.

1190 **Fig. 5.** Boxplots of LCA concentrations in Toolik Lake SPM binned by (a) seasonal lake stratification
1191 status and (b) observed water column temperatures.

1192 **Fig. 6.** Fractional abundances of individual LCAs in Toolik Lake SPM plotted vs. observed lake water
1193 temperature. Individual LCAs correspond to the double bond numbers shown across the top and the chain
1194 length/carbonyl positions shown along the right margin. All LCAs with significant temperature effects
1195 ($p < 0.01$; Pearson’s correlation coefficient) are plotted with linear least squares regression lines.

1196 **Fig. 7.** U_{37ab}^K vs. observed lake water temperatures with linear least squares regression line shown in gray.
1197 U_{37ab}^K temperature residuals are plotted below and the 1σ range is indicated with dashed lines.

1198 **Fig. 8.** Coefficient plots for bootstrapped linear regressions of U_{37ab}^K and U_{37a}^K vs. temperature based on
1199 current and previously published calibrations. Descriptions and references for the calibration datasets are
1200 shown in the left margin and the phylotypes of the alkenone-producers associated with each calibration are
1201 show in the right margin. Vertical lines represent coefficient values, horizontal black bars denote the 95%
1202 bootstrap confidence intervals and white boxes denote the 68% bootstrap confidence intervals for the
1203 regression coefficients.

1204 **Fig. 9.** RIK indices from Toolik Lake SPM plotted vs. observed lake water temperature. All RIK indices
1205 are shown with a common y-range for comparison. All correlations are significant ($p < 0.01$; Pearson’s
1206 correlation coefficient and Spearman’s Rank).

1207 **Fig. 10.** Ternary diagrams showing relative proportions of C_{37} LCAs in various samples. C_{37} distributions
1208 separate by phylogenetic groups (a) and by lake salinity (b). Sources and references for various LCA
1209 datasets are described in the text. NAK, northern Alaskan lakes; CAN, interior Canada lakes; CHN,
1210 northwestern Chinese lakes.

Table 1. Northern Alaskan lakes analyzed in this study, with lake properties and LCA concentrations and distribution parameters.

Map Number	Lake Name	Lat.	Long.	C ₃₇ Total (ug/gTOC)	U ₃₇ ^K	C _{37:4} /C _{38:4} Et ^a	RIK ₃₇	Mean Depth (m)	Surface Area (km ²)	pH	Conductivity (μS/cm)	Alkalinity (μEq/L)	SO ₄ ²⁻ (μM)	Cl ⁻ (μM)	DOC (μM)	C:N	Chla (μg/L)
01	E1	68.626	-149.554	24.11	-0.55	1.01	0.53	3.09	0.029	7.76	89.8	873	11.43	3.50	368	10.3	1.69
02	E5	68.643	-149.458	20.67	-0.53	1.15	0.52	6.33	0.109	6.84	12.9	104	2.51	5.31	487	9.9	4.13
03	Fog 1	68.684	-149.079	28.90	-0.59	1.64	0.56	8.37	0.035	8.15	121.4	1230	ND	ND	345	8.9	0.71
04	Fog 2	68.679	-149.089	21.72	-0.59	1.31	0.53	7.79	0.056	8.04	129.1	1344	45.82	5.55	296	7.5	0.40
05	Galbraith	68.460	-149.420	4.17	-0.48	1.49	0.60	4.2 ^c	4.1 ^c	8.22	165.2	ND	ND	ND	ND	14.7	ND
06	I1	68.569	-149.588	6.92	-0.57	1.03	0.54	3.91	0.246	7.13	27.5	263	2.52	3.01	395	8.6	0.85
07	I2	68.571	-149.566	21.72	-0.56	1.39	0.51	7.27	0.093	6.90	24.7	213	3.04	3.14	544	8.4	0.77
08	I3	68.575	-149.581	4.29	-0.48	0.69	0.56	1.84	0.034	7.03	24.2	226	2.02	4.94	487	10.6	1.45
09	I4	68.580	-149.583	3.57	-0.49	0.69	0.54	2.96	0.083	7.02	22.6	211	1.49	2.84	505	10.3	1.75
10	I5	68.587	-149.590	6.33	-0.48	1.02	0.52	3.87	0.209	6.96	23.4	213	2.09	3.15	483	9.4	1.21
11	I6	68.597	-149.593	4.33	-0.53	0.76	0.55	5.66	0.178	7.34	39.4	372	3.13	3.38	463	8.8	1.07
12	I6HW	68.581	-149.619	8.41	-0.55	0.93	0.51	3.63	0.044	7.67	92.9	939	7.27	3.45	409	9.7	0.63
13	I7	68.601	-149.593	1.23	-0.54	0.84	0.56	3.7	0.029	7.36	41.0	394	3.32	3.31	451	10.2	0.96
14	IKP-003	70.793	-154.517	19.16	-0.54	1.52	1.00	2.8	ND	7.30	4596.0	ND	ND	ND	ND	12.4	ND
15	INI-004	70.010	-153.155	9.95	-0.45	1.12	0.58	8	ND	6.20	58.0	ND	ND	ND	ND	9.6	ND
16	Iswamp	68.611	-149.597	3.46	-0.47	0.66	0.58	2.35	0.019	7.16	46.1	456	4.23	2.57	481	10.6	0.63
17	N1	68.640	-149.604	12.87	-0.56	1.24	0.52	4.81	0.047	8.05	111.9	1163	16.36	4.03	417	8.6	1.19
18	N2	68.641	-149.622	40.86	-0.62	1.28	0.54	4.99	0.016	8.02	154.8	1606	23.89	2.97	440	8.2	0.98
19	NE9B	68.653	-149.597	2.86	-0.55	1.09	0.52	2.94	0.003	7.97	215.8	1847	88.35	3.71	452	8.9	1.09
20	Perch	68.941	-150.500	176.65	-0.55	1.60	0.55	4.52	0.162	7.75	52.4	405	ND	ND	445	10.8	0.84
21	S11	68.631	-149.648	3.67	-0.62	0.93	0.53	3.3	0.004	8.22	149.8	1976	ND	ND	413	9.2	1.29
22	S6	68.630	-149.639	39.18	-0.63	1.42	0.52	2.97	0.012	8.48	159.2	1962	62.61	3.86	350	6.9	1.54
23	Toolik	68.632	-149.602	11.22	-0.54	1.29	0.54	7.35	1.488	7.61	75.9	709	10.02	4.30	416	8.3	1.22
24	Dimple	68.955	-150.195	-	-	-	-	2.15	0.100	6.81	18.3	137	ND	ND	845	15.3	2.29

25	E6	68.644	-149.439	-	-	-	-	1.62	0.019	6.77	13.6	110	1.36	4.85	616	8.7	2.65
26	Fog 4	68.680	-149.072	-	-	-	-	2.08	0.018	7.63	90.8	858	40.60	8.03	474	8.9	3.89
27	Horn	68.958	-150.310	-	-	-	-	0.71	0.360	7.04	22.9	187	ND	ND	868	11.8	1.92
28	I8	68.610	-149.582	-	-	-	-	2.73	0.182	7.66	72.3	659	8.96	3.25	529	10.8	1.90
29	Luna	68.962	-150.207	-	-	-	-	1.28	0.050	7.16 ^b	18.43 ^b	ND	ND	ND	ND	15.1	ND
30	MC7916	70.299	-151.463	-	-	-	-	2.1	ND	6.70	123.0	ND	ND	ND	ND	11.3	ND
31	North	69.325	-150.943	-	-	-	-	1	0.332	7.24 ^b	49.20 ^b	ND	ND	ND	ND	11.0	ND
32	S7	68.630	-149.643	-	-	-	-	0.82	0.010	8.22	183.7	2329	82.64	3.19	382	9.2	1.02
33	Shasta	68.938	-150.349	-	-	-	-	1.26	0.118	6.56	11.6	77	ND	ND	865	12.1	3.00
34	Teshekpuk	70.606	-153.687	-	-	-	-	5	2750	6.80	268.0	ND	ND	ND	ND	10.1	ND
35	UMI-004	69.350	-152.292	-	-	-	-	2.4	ND	6.40	270.0	ND	ND	ND	ND	9.1	ND

^a C_{37:4}/C_{38:4}Et is reported instead of C₃₇/C₃₈ because C₃₇/C₃₈ could not be not be calculated for all sediment samples due to co-eluting compounds.

^b 2011 measurements.

^c Data from Burr (1995).

^d ND = No data available.

Table 2. Regression statistics for various temperature calibrations from Toolik Lake SPM.

Index	Slope	Intercept	r^2	RMSE (°C)
U_{37ab}^K	0.021	-0.68	0.85	1.37
U_{37a}^K	0.021	-0.79	0.84	1.43
U_{37b}^K	0.021	-0.82	0.84	1.40
$U_{37ab}^{K'}$	0.0042	0.067	0.77	1.77
$U_{37a}^{K'}$	0.0068	0.12	0.78	1.73
$U_{37b}^{K'}$	0.0068	0.14	0.76	1.85
$U_{37ab}^{K''}$	0.016	0.29	0.85	1.35
$U_{37a}^{K''}$	0.013	0.17	0.84	1.42
$U_{37b}^{K''}$	0.013	0.15	0.85	1.35
U_{38Etab}^K	0.012	-0.54	0.74	1.95
U_{38Eta}^K	0.019	-0.78	0.82	1.55
U_{38Etb}^K	0.010	-0.58	0.68	2.23
U_{38Meab}^K	0.017	-0.35	0.77	1.79
U_{38Mea}^K	0.018	-0.38	0.77	1.78
U_{38Meb}^K	0.032	-0.69	0.81	1.59
U_{39Etab}^K	0.015	-0.36	0.51	3.21
U_{39Eta}^K	0.018	-0.42	0.55	2.98
U_{39Etb}^K	0.019	-0.51	0.58	2.79

Table 3. RIK₃₇ measurements for several environmental and culture samples, with RIK₃₇-inferred haptophyte producers.

Samples	n	Reported Salinity ^a	RIK ₃₇ (2σ)	RIK ₃₇ -inferred phylotype	Reference
BrayaSø Sediment	1	0.9	0.57	Group I	Longo et al., 2013 D'Andrea et al., 2011
Toolik Lake SPM	52	Fresh	0.53 (0.012)	Group I	This Study
NAK freshwater lakes	22	Fresh	0.54 (0.048)	Group I	This Study
Lake IKP-003	1	3.5 ^b	1	Group II	This Study
Lenore Lake	1	3	0.56	Group I	Toney et al., 2011
Fishing Lake	1	2.2	0.61	Group I + Group II	Toney et al., 2011
Shannon Lake	1	2.4	0.69	Group I + Group II	Toney et al., 2011
Humboldt Lake	1	1.5	0.78	Group I + Group II	Toney et al., 2011
CAN saline lakes	7	5.6 - 52.7	1	Group II	Toney et al., 2011
CHN saline lakes	11	0.7 - 146.3	1	Group II	Zhao et al., 2014
Lake George	1	9.7	1	Group II	Longo et al., 2013; Toney et al., 2010
Group II cultures	5	-	1	Group II	This study; Longo et al., 2013; Theroux et al., 2013
Group III cultures	2	-	1	-	This study; Longo et al, 2013
Combined ocean sediment sample	1	-	1	-	Longo et al., 2013

^a Salinities reported are from the epilimnion or surface waters with units reported in references.

^b Salinity calculated from conductivity using standard conversion.

^c NAK, Northern Alaska; CAN, interior Canada; CHN, northwestern China

Figure 1

FIGURES

Figure 2

- Toolik Lake, AK, USA¹
- Etang des Valleees, FR²
- Lake Annecy, FR²
- Upper Murray lake, Canada³
 - BrayaSo, Greenland⁴
 - LimnaeaSo, Greenland⁴
 - HundeSo, Greenland⁴

OTU 1³

OTU 5³

Group I

(Greenland Phylotype/EV clade)

- Clear Lake, ND, USA³
- Pyramid Lake, NV, USA³
 - Lake Fryxell, Antarctica⁵
 - Skoal Lake, ND, USA³
 - Medicine Lake, SD, USA³
 - Lake George, ND, USA³
 - Hap-A⁶
 - Hap-B⁶

Group II

- Lake Van, Turkey⁷
- Lake Van, Turkey⁷
- Lake Xiarinur, *Chrysotila lamellosa*⁸
- Great Salt Lake, USA³

OTU 7³

OTU 8³

*Isochrysis galbana*⁹

Group III

*Emiliana huxleyi*⁹

Salinity (psu)

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

U_{37}^K

U_{37}^K Temp
Residuals ($^{\circ}\text{C}$)

Figure 8

Figure 9

Ratio of Isomeric Ketones

Observed Temperature (°C)

Figure 10