



Contents lists available at ScienceDirect

Organic Geochemistry

journal homepage: www.elsevier.com/locate/orggeochem

Systematic chemotaxonomic profiling and novel paleotemperature indices based on alkenones and alkenoates: Potential for disentangling mixed species input

Yinsui Zheng^{a,b}, Patrick Heng^a, Maureen H. Conte^{b,c}, Richard S. Vachula^{a,d}, Yongsong Huang^{a,†}^a Department of Earth, Environmental and Planetary Sciences, Brown University, Providence, RI 02912, USA^b Ecosystems Center, Marine Biological Laboratory, Woods Hole, MA 02543, USA^c Bermuda Institute of Ocean Sciences, St Georges, Bermuda^d Institute at Brown for Environment and Society, Brown University, Providence, RI 02912, USA

article info

Article history:

Received 14 November 2018

Received in revised form 17 December 2018

Accepted 20 December 2018

Available online 21 December 2018

Keywords:

Alkenones

Alkenoates

Haptophytes

Culture

Chemotaxonomy

Temperature calibration

abstract

The unsaturation indices (U_{37}^K , U_{37}^{Kl}) of long chain alkenones are powerful paleotemperature proxies and have been widely applied for sea surface temperature (SST) reconstructions in the past three decades. However, these indices encounter major difficulties in systems harboring different alkenone-producing haptophyte species, such as saline lakes and marginal ocean environments. All haptophytes produce C_{37} alkenones, but different species often display large differences in temperature calibrations and may bloom in different seasons, hindering the use of U_{37}^K and U_{37}^{Kl} indices for reliable paleotemperature reconstructions in mixed systems. To overcome these problems, we have recently reported a new analytical method that allows comprehensive separation of up to 32 alkenones, alkenoates and their double bond positional isomers in culture and sediment samples. Here we report a systematic analysis of alkenones and alkenoates from six haptophyte cultures growing at a wide range of temperatures (4–25 °C). Together with a compilation of 230 previously published culture data sets, we present here systematic calibrations of temperature-sensitive indices based on all alkenone and alkenoate homologues (including isomers). Using this dataset, we extract systematic chemotaxonomic criteria for differentiating individual haptophyte species and demonstrate such chemotaxonomic features can be encoded into a machine learning model for reliable species identifications. Specifically, we show that temperature calibrations based on C_{38} methyl ketones and C_{39} ethyl ketones are potentially useful for disentangling mixed inputs in estuarine systems where Group III (*E. huxleyi*) and Group II alkenones mix, and that C_{36} ethyl alkenoate isomeric ratios display minimal species heterogeneity and are potentially more suited for reconstructing temperatures in mixed systems with different Group II haptophytes. Using the culture data as base profiles, we construct a mathematical model for estimating percentage inputs from alkenones of different Isochrysidales groups in mixed systems, with potential implications for inferring past salinity changes. Overall, the results from this study demonstrate important new applications of alkenone and alkenoate biomarkers in paleoclimate and paleoenvironmental research.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

Long-chain alkenones (LCAs) in marine sediments have been widely and successfully used to reconstruct sea surface temperatures for ~40 years (e.g., Brassell et al., 1986; Herbert, 2001; Conte et al., 2006; Bendle and Rosell-Melé, 2007; Lawrence et al., 2009; Ho et al., 2012; Tzanova and Herbert, 2015), and have recently been applied to lacustrine systems (e.g., Zink et al.,

2001; D'Andrea et al., 2011, 2012). The remarkable diagenetic stability of LCAs in sediments (e.g., Grimalt et al., 2000; Huguet et al., 2009; Rontani et al., 2013), combined with well-demonstrated temperature correlation between LCA unsaturation ratios (U_{37}^{Kl}) from laboratory cultures and core top sediments spanning large latitudinal ranges (Conte et al., 2006), have made alkenones one of most successful paleo-thermometers ever developed.

LCAs are produced by various algal species in the Isochrysidales, an order of Haptophyceae. In the open ocean, the most common producers of LCAs are *Emiliania huxleyi* and *Gephyrocapsa oceanica*, whereas in estuaries and saline lakes the most common species are

[†] Corresponding author.

E-mail address: yongsong_huang@brown.edu (Y. Huang).

Isochrysis galbana and *Ruttenra lamellosa* (Chu et al., 2005; Zheng et al., 2016a, 2016b). In freshwater and oligosaline lakes, a new species termed “the Greenland haptophyte” appears to be the primary producer of LCAs in southwestern Greenland lakes (D’Andrea and Huang, 2005; D’Andrea et al., 2006, 2011), and similar lakes worldwide (Crump et al., 2012; Longo et al., 2013; Longo et al., 2016, 2018). Based on 18S rRNA marker gene studies, Theroux et al. (2010) have classified alkenone-producing haptophytes into three groups: Group I includes the Greenland lake species, Group II includes the saline/brackish water species *I. galbana*, *R. lamellosa* and *Tisochrysis lutea*, and Group III includes the open ocean *E. huxleyi* and *G. oceanica*. Optimal salinity ranges for the three groups of the alkenone-producing haptophytes are different, but the group distributions may overlap in intermediate salinity ranges or systems with significant salinity oscillation (Longo et al., 2016, 2018).

The unsaturation ratios of LCAs in various forms have been demonstrated to strongly correlate with temperature (Marlowe et al., 1984a; Prahl et al., 1988). In addition to the well-known unsaturation ratios U_{37}^K and U_{37}^K , a new index has recently been proposed for Group II haptophytes that includes the $C_{37:4}$ alkenone but excludes $C_{37:2}$ [$U_{37}^{Kv} = C_{37:3}/(C_{37:3} + C_{37:4})$] and shows greater linearity against temperature in culture samples (Zheng et al., 2016b; Table 2). Temperature calibrations differ among different haptophyte species. While the unsaturation ratios U_{37}^K and U_{37}^K provide excellent paleotemperature reconstructions for open ocean sediments, applying the unsaturation ratios U_{37}^K , U_{37}^K or U_{37}^{Kv} to temperature reconstruction in coastal settings and saline lakes encounters considerable difficulties. In such settings it is common that multiple different haptophytes with differing temperature calibrations co-exist.

In particular, marginal ocean environments (MOEs) such as estuaries exhibit large salinity gradients and/or fluctuations: LCAs in MOEs can derive from both open ocean and brackish water haptophytes (Mercer et al., 2005; Kaiser et al., 2017). Ironically, MOEs often have high sedimentation rates and high organic contents, and thus would be ideal for high-resolution paleotemperature reconstructions using LCAs. However, at present it is difficult or perhaps impossible to use U_{37}^K and U_{37}^K for reconstructing SST in such environments due to the variable contributions of different alkenone-producing groups.

One potential solution for the mixed species problem is to study the full suite of LCAs and long chain alkenoates (LCEs). In mixed systems, the temperature signal could be extracted if one particular species makes temperature-sensitive LCA or LCE pairs that are not produced by other species, or if for certain indices the two mixing species display the same temperature relationship. Most of the previous paleoclimate studies have focused on C_{37} methyl alkenones (Herbert, 2001; Conte et al., 2006 and References therein) in part because conventional gas chromatographic (GC) methods do not adequately separate the individual alkenones and alkenoates. In fact, partial and even full co-elution occurs frequently, making it difficult to reliably and consistently determine the unsaturation patterns of non- C_{37} LCA homologues and LCEs. Additionally, the relationship between temperature and alkenoate abundance and distributions also shows substantial genetic and physiological variability among alkenone synthesizers, making interpretations ambiguous (Conte et al., 1994, 1995, 1998). Therefore, even though the unsaturation ratios of C_{38} methyl/ethyl alkenones (U_{38Me}^K , U_{38Et}^K) and those of LCEs are long known to be temperature sensitive, these unsaturation ratios have gained little application in paleotemperature reconstructions (Conte and Eglinton, 1993; Conte et al., 1998; Grossi et al., 2000; Nakamura et al., 2014; Rontani et al., 2004; Rosell-Melé et al., 1994).

Building upon the previously reported separation using a VF-200 ms column (Longo et al., 2013), Zheng et al. (2017) reported an optimized GC method that fully resolves all the alkenone and alkenoate compounds including their isomers. This method provides new opportunities to investigate the chemotaxonomic fingerprints of different groups of LCA-producing haptophytes in unprecedented detail. The main objectives of this study were to: (1) to systematically profile key alkenone-producing haptophyte species for all alkenones and alkenoates in order to identify prominent chemotaxonomic differences between species, and how such profiles link to genetic lineages; (2) to use the chemotaxonomic differences revealed in our data, in combination with published species-specific alkenone/alkenoate profiles, to develop an open-ended mathematical model that will allow quantitative partitioning of different species inputs and whose skill improves as more culture data become available; and (3) to culture key species of Isochrysidales over a wide range of temperatures (including at 4 °C to observe low temperature response) to develop novel temperature proxies and to compare temperature sensitivities of different proxies. Specifically, we paid particular attention to temperature sensitivity of LCA/LCE isomeric ratios and unsaturation ratios previously not reported due to the limitation of traditional analytical methods.

2. Materials and methods

2.1. Culture experiments

2.1.1. Origin of the cultured alkenone-producing strains

Four Group II Isochrysidales strains and two Group III Isochrysidales strains were cultured in this study (Table 1): (1) a recently isolated Isochrysidales strain from Lake George (denoted as LG strain in this paper) and identified as *Ruttenra lamellosa* (Zheng et al., 2016b). The LG strain was collected from Lake George, North Dakota (46.74°N, 99.49°W), USA in June 2009 and isolated into pure culture in April 2011; (2) *R. lamellosa* Strain CCMP1307, *I. galbana* strain CCMP1323, *T. lutea* strain CCMP463 and *E. huxleyi* strain CCMP 2758 from National Center for Marine Algae and Microbiota (NCMA, formerly CCMP); (3) *E. huxleyi* strain Van556 from the North West Pacific Culture Collection (NEPCC). Because a pure culture of Group I Isochrysidales is currently unavailable, the LCA profile for Group I was derived from a re-analysis of surface sediment samples from Braya Sø, Southwestern Greenland (D’Andrea and Huang, 2005; Longo et al., 2013; Zheng et al., 2017).

2.1.2. Culture conditions for the alkenone-producing strains

The culture conditions and results for LG strain have been reported previously in Zheng et al. (2016b). We acclimated the LG strain culture at 4 °C, 10 °C, 15 °C, 20 °C and 25 °C in 0.2 lμm filter-sterilized seawater adjusted to a salinity of 10 g/kg amended with f/2 nutrients in 25 mL glass tubes with full spectrum lighting on a 12:12 light:dark cycle for two weeks prior to the start of the culture experiment. We then grew all cultures in triplicate volumes of 500 mL in 1 L flasks under the same conditions as acclimation without aeration. We monitored cell concentrations for five weeks using haemocytometer counts to determine if the cultures were in the log phase of growth. We monitored incubator temperatures every day and averaged the recorded temperatures that were then used for calculating the temperature calibrations. These actual measured temperatures were 5 °C, 9 °C, 17 °C, 20 °C and 25 °C respectively.

Subsamples (1 mL each) were collected every day and preserved with 50 lμl HistoChoice fixative (Amresco, Solon, OH) to ascertain the cell densities of the cultures. For alkenone and

Table 1
Haptophyte species selected for culturing.

Species	Strain	Collection center	Location where isolated	References
"Lake George"	LG strain	BU/MBL [*]	A	Zheng et al. (2016b)
<i>R. lamellosa</i>	CCMP1307	NCMA	B	Nakamura et al. (2014)
<i>I. galbana</i>	CCMP1323	NCMA	C	Versteegh et al. (2001)
<i>T. lutea</i>	CCMP463	NCMA	D	Nakamura et al. (2016)
<i>E. huxleyi</i>	Van 556	NEPCC	E	Conte et al. (1998)
<i>E. huxleyi</i>	CCMP2758	NCMA	F	Prahl and Wakeham (1987)

A: Lake George, North Dakota, USA (48.06°N, 100.44°W)

B: Keystone Salt Marsh, Washington, USA, North Pacific (48.22°N, 122.64°W).

C: Marine Biological Station, Isle of Man, North Atlantic (54.08°N, 4.77°W).

D: Turks and Caicos Islands, British West Indies, North Atlantic (22.05°N, 72.21°W).

E: Northeast Pacific (49.41°N, 144.41°W).

F: Northeast Pacific (50.1833°N, 144.95°W).

^{*} BU refers to Brown University and MBL refers to Marine Biological Laboratory.

alkenoate characterizations, we harvested 200 mL volumes of the cultures when they were at the late log to early stationary growth phase. We acclimated the strains CCMP1307, CCMP1323, CCMP463, Van556 and CCMP2758 to the following temperatures in 0.2 μ m filter-sterilized seawater: 4 °C, 9 °C, 14 °C, 20 °C and 25 °C respectively. The seawater with a salinity of 33 g/kg was amended with f/2 nutrients in 25 mL glass tubes. The acclimation cultures were irradiated with full spectrum lighting on a 16:8 light/dark cycle for 2 weeks prior to the start of the culture experiments. We grew all cultures in 200 mL solution in 500 mL Pyrex flasks under the same conditions as acclimation.

Cultured samples were harvested at the late log to early stationary growth phase by monitoring the optical density at 750 nm (OD750) using a plate reader (Thermo Scientific). At the time of sample collection, cultures were filtered through a 47 mm diameter glass fiber filter (Whatman GF/A, 0.45 μ m) for lipid analysis. The filters were wrapped with aluminum foil, and immediately frozen at 20 °C.

2.2. Published culture and core-top sediment samples

In order to maximize the robustness of our statistical analyses, we also include the published culture data from Nakamura et al. (2014, 2016) and Conte et al. (1998), as well as published in situ samples from Longo et al. (2016) and from Conte and Eglinton (1993). Only the culture samples that were quantified for both LCEs (designated as C₃₆OMe for C₃₆ methyl alkenoates, C₃₆OEt for C₃₆ ethyl alkenoates) and LCAs (C₃₇Me, C₃₈Et, C₃₈Me, C₃₉Et, C₃₉Me and C₄₀Et, where the numbers in subscripts designate carbon numbers, Me and Et represent methyl and ethyl ketones, respectively) are included in this study. Nakamura et al. (2014, 2016) used a VF-200 ms gas chromatographic column for analysis, which provides sufficient compound separation. Conte et al. (1998) analyzed LCAs using CPSil5CB column but used cool on-column injection (rather than the common split-splitless injection), which also provide compound separations sufficient for inclusion in our analysis here (excluding ratios of alkenone or alkenoate isomers that cannot be resolved using CPSil5CB). Note that Conte et al. (1998) trans-esterified the culture samples and combined both methyl alkenoates and ethyl alkenoates, which also reduces the likelihood of coelution between alkenones and alkenoates.

Several culture data sets were not included in our analyses. We did not include the culture data from Araie et al. (2018) in this paper because of our concerns regarding the chromatographic resolution between isomers and between alkenoates and alkenones. In our experience, when tetra-unsaturated alkenones are present, it is difficult to accurately quantify C₃₆ ethyl alkenoates, regardless of the GC column used. Helium as the carrier gas also leads to significantly lower chromatographic resolution for alkenones and

alkenoates than using hydrogen as a carrier gas. For example, Araie et al. (2018) found C₃₆ ethyl tri-unsaturated alkenoate isomeric ratios decline with rising culture temperatures, opposite to our observations (see Section 3.3). We also did not include the culture data set of Conte et al. (1995) in our analyses due to concerns about the excessive LCA compositional variability (which were not observed in later studies) among the various *E. huxleyi* and *G. oceanica* strains grown under the same temperature. We have also omitted the LCA dataset from the mesocosm experiments in the Raunefjorden in southern Norway (Conte et al., 1994, mesocosm experimental details given in Egge and Heimdal, 1994). Unusually high percentages of tetraunsaturated alkenones relative to open ocean *E. huxleyi* strains in the induced blooms in the mesocosm experiments and in nearby fjords (Conte et al., 1998, 1994) may have originated from complexity in Isochrysidales species and genotypes.

2.3. Sediment samples

In order to test our model for disentangling inputs from different groups of Isochrysidales, three Black Sea sediment samples from a gravity core in the Unit I and Unit II (Xu et al., 2001) were analyzed. This core was collected from the central western basin in April 1988 during Leg 1 of the R/V Knorr cruise 134. The giant gravity core (GGC-18, core length 2.55 m) was collected at 42.87°N, 31.39°E at a water depth of 2091 m. This GGC-18 Black Sea core was collected at the same site as the GGC-19 Black Sea core analyzed by Xu et al. (2001).

2.4. Analysis of alkenones by GC-FID

All samples were processed using established methods for cultured algae (Toney et al., 2012; Theroux et al., 2013; Zheng et al., 2016b) and sediments (e.g., D'Andrea and Huang, 2005). The culture filter samples were freeze-dried overnight and extracted using 3 × 30 min bursts of sonication in 40 mL of dichloromethane (DCM). The freeze-dried sediment samples were extracted with a Dionex accelerated solvent extraction (ASE 200) system at 120 °C and 1200 psi. After extraction, lipids were separated into different fractions using a Supelco Supelclean flash column (LC-NH₂, 40–63 mm). The neutral fractions were eluted with DCM/isopropyl alcohol (2:1, v/v), and were subsequently separated into alkane, ketone/ester and alcohol fractions using hexane, DCM and methanol (MeOH), respectively, using silica gel columns (SiO₂, 40–63 mm). The alkenones were quantified by GC-FID using an Agilent 7890 N Series instrument fitted with a Restak-200 GC column (105 m × 250 μ m × 0.25 μ m) (Zheng et al., 2017). With the new GC method, we were able to completely resolve all the alkenoate and alkenone peaks at the baseline levels.

2.5. Principal component analyses (PCA)

Utilizing the PCA functions from MATLAB, we executed principal component analysis on the fractional abundances up to thirty-two alkenones and alkenoates quantified in the culture samples from this study and in culture samples from published data (Conte et al., 1998; Longo et al., 2016; Nakamura et al., 2016, 2014). The LCA profile for Group I is derived from re-analysis of surface sediment samples from Braya Sø, Southwestern Greenland (D'Andrea and Huang, 2005), using the GC method of Zheng et al. (2017).

2.6. Machine learning classification model and multi-source mixing model

The classification model was built using Python and the Scikit-learn module (supplementary information), which integrates a wide range of state-of-the-art machine learning algorithms for medium-scale supervised and unsupervised problems (Pedregosa et al., 2012). We used the K nearest neighbors (KNeighbors) classifier, a canonical example of a nonparametric machine-learning algorithm, which averages the categorical species labels (R. lamellosa, I. galbana, T. lutea, E. huxleyi, G. oceanica and Group I (“Greenland haptophyte”) of the K nearest neighbors in the training set (normalized percentage of the 32 alkenoates and alkenones).

We also developed a multi-source mixing model to estimate the percentage inputs from each group of haptophytes in sediment samples in reference to a model we previously developed for differentiating plant leaf wax sources (Gao et al., 2011). In our model, the normalized alkenone distributions of the 230 samples from six different species are known (Supplementary Table S1) while the normalized alkenone distributions for Black Sea sediments are unknown (Supplementary Table S2). Note that the Black Sea sediment samples were saponified to eliminate esters. Therefore, in this model, we only use the distribution of the 24 alkenones but not alkenoates. There are six end members (species E. huxleyi, G. oceanica, R. lamellosa, I. galbana, T. lutea and Group I) in total. The average alkenone distributions for each end member species are obtained from Supplementary Table S1 (excluding alkenoates) by averaging over all the samples of each specific species to form the A matrix of coefficients. We analyzed three Black Sea sediment samples to estimate the fraction of different haptophyte groups using this multi-source model.

The linear system is defined as $Ax = B$ where matrix A is a 25×6 matrix of the coefficients (the average distribution of alkenones for each species). B is a 25×1 column vector (the alkenone distribution of Black Sea sediment samples) and the variable vector X is a 6×1 column vector (the fraction inputs from five species/endmember).

$$\begin{array}{ccc}
 2O & \delta\%K1; 1b & 13 \\
 6B & \delta\%K2; 1b & C7 \\
 6B & \delta\%K2; 6b & C7 \\
 6B & \delta\%K23; 1b & C7 \\
 A \frac{1}{4} & \delta\%K23; 6b & C7 \\
 & & \delta 1b \\
 6B & & C7 \\
 6B & \delta\%K24; 1b & C7 \\
 6B & \delta\%K24; 6b & C7 \\
 4C & & A5 \\
 1 & 1 &
 \end{array}$$

$$B \frac{1}{4} \frac{1}{3} \frac{1}{3} K1; sedib \quad \delta\%K2; sedib \quad \delta\%K24; sedib \quad 1]$$

$$X \frac{1}{4} \frac{1}{3} f1 \quad f2 \quad f3 \quad f4 \quad f5 \quad f6]$$

where %K1, 1 is the first ketone percentage of the species 1 haptophyte and f1 is the percentage of haptophyte inputs. The Python code of this multi-source mode is also provided in the supplementary data.

3. Results

3.1. Chemotaxonomic differences for LCAs and LCEs in Group I, Group II and Group III Isochrysidales

In order to increase our sample size and ensure a more statistically robust comparison, we have combined our own culture data with previously published culture results (Conte et al., 1998; Nakamura et al., 2014, 2016) to identify general characteristics of alkenones and alkenoates in different haptophyte groups. These published culture results contain relatively complete profiles of the alkenones and alkenoates we use in this study. Although the GC column used by Nakamura et al. (2016) cannot separate the $C_{36:2}OEt$ from $C_{37:4}Me$, their culture strain T. lutea does not produce $C_{37:4}Me$ at the growth temperatures ($>15^\circ C$) used in that study; hence the dataset is included for this study. Note that the alkenoate data reported in Conte et al. (1998) are methyl alkenoates because all the ethyl alkenoates were transformed to methyl alkenoates by trans-esterification.

Fig. 1 presents the GC chromatograms of LCEs ($C_{36}OMe$ and $C_{36}OEt$) and LCAs ($C_{37}Me$, $C_{38}Et$, $C_{38}Me$, $C_{39}Et$, $C_{39}Me$, $C_{40}Et$) produced by Group I (surface sediment from Braya Sø, southwestern Greenland; D'Andrea and Huang, 2005), and four Group II and two Group III strains cultured at $9^\circ C$ in this study. GC chromatograms for other growth temperatures are given in Supplementary Figs. S1–6. Even though pure Group I strains have not

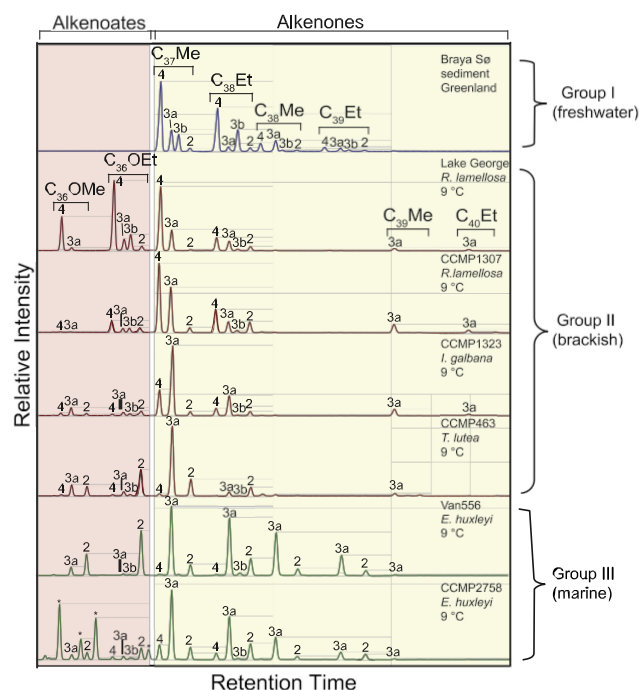


Fig. 1. GC-FID chromatograms showing typical distributions of long chain

alkenones and long chain alkenoates in Group I, Group II and Group III samples. The Group I alkenones (blue) are represented using Braya Sø surface sediment sample, since currently a pure culture for Group I haptophyte is unavailable. Group II alkenones and alkenoates (red) are represented by Group II strains LG R. lamellosa, CCMP1307 R. lamellosa, CCMP1323 I. galbana and CCMP463 T. lutea respectively. Group III alkenones and alkenoates are represented by Group III haptophytes Van556 and CCMP2857 E. huxleyi. Alkenones are designated with C_nMe or C_nEt , where n is the carbon number, and Me and Et represent methyl and ethyl alkenones respectively. Similarly, alkenoates are designated with C_nOMe or C_nOEt where -OMe and -OEt are methyl and ethyl esters respectively. Numbers 4, 3, and 2 on top of individual peaks denote the number of double bonds; 3a and 3b are tri-unsaturated double bond positional isomers. Peaks labeled with an asterisk are unknown co-eluting compounds. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

yet been isolated from the environment, DNA analyses show that Group I haptophytes are common in relatively high pH, freshwater lakes (Longo et al., 2016), and display consistent alkenone profiles (Longo et al., 2018).

Group I haptophytes do not produce alkenoates whereas both Group II and Group III haptophytes produce variable amounts of alkenoates (C_{36} alkenoic acid methyl and ethyl esters, designated as $C_{36}OMe$ and $C_{36}OEt$, respectively). A characteristic feature of Group I haptophytes is that they produce abundant tri-unsaturated alkenone isomers with $D^{14,21,28}$ double bond positions (Dillon et al., 2016; Longo et al., 2016, 2013). In contrast, Group II and Group III strains produce only small amounts of $C_{38:3b}Et$ alkenone isomer ($D^{14,21,28}$), and do not produce $C_{37:3b}Me$ alkenone isomer ($D^{14,21,28}$) (Longo et al., 2013; Zheng et al., 2017). As a result, the ratio of RIK_{38E} ($RIK_{38E} = C_{38:3a}Et / (C_{38:3a}Et + C_{38:3b}Et)$) is much smaller for Group I samples (0.241 ± 0.047) compared to that for Group II and Group III samples (0.984 ± 0.045) grown at a wide range of temperature (Fig. 1, Supplementary Table S1). In addition to the $C_{38:3b}Et$ alkenone isomers, both Group II and Group III produce $C_{36:3b}OEt$ alkenoate isomers.

Group II haptophytes generally do not produce $C_{38}Me$ or $C_{39}Et$ alkenones (Lopez et al., 2005; Theroux et al., 2013; Zheng et al., 2017). The $>C_{37}$ alkenones from Group II haptophytes are $C_{38}Et$, $C_{39}Me$ and $C_{40}Et$, respectively (Fig. 1). In contrast, both $C_{38}Me$ and $C_{39}Et$ alkenones are commonly produced by Group I and Group III haptophytes (e.g., Conte et al., 1998; Nakamura et al., 2014; Longo et al., 2016), although $C_{38}Me$ and $C_{39}Et$ in Group I contain tri-unsaturated isomers that are generally absent in Group III haptophytes.

Group II species tend to biosynthesize more tetra-unsaturated alkenones relative to Group III haptophytes under the same environmental temperatures (Fig. 1, Supplementary Figs. S1–S4). A notable exception is *T. lutea*, whose $C_{37:4}Me$ production is small, similar to *E. huxleyi*. The differences seen here are consistent with the relatively high proportions of tetra-unsaturated alkenones, especially $C_{37:4}$ homologue in estuary/lacustrine environments (e.g., Cranwell, 1985; Volkman et al., 1988; Chu et al., 2005; Toney et al., 2010) relative to open ocean regions, which are dominated by the cosmopolitan species *E. huxleyi* (Fig. 1). The Group III marine species also do not produce tetra-unsaturated alkenoates (Fig. 1, Supplementary Figs. S5 and S6).

All haptophytes also produce small amounts of longer chain alkenones, $C_{39}Me$ and $C_{40}Et$. Group II haptophytes produce both $C_{39}Me$ and $C_{40}Et$ while Group III haptophytes only make $C_{39}Me$ but no $C_{40}Et$ alkenones. We did not detect any $C_{39}Me$ or $C_{40}Et$ alkenones in Group I samples (Fig. 1).

In general, the unsaturation pattern of alkenones is similar to that of alkenoates. For example, in Group II species where tetra-unsaturated homologues (e.g., $C_{37:4}Me$) are abundant, the corresponding tetra-unsaturated alkenoates are also abundant (e.g., $C_{36:4}OMe$ and $C_{36:4}OEt$; Fig. 1). This is generally true for most sets of alkenones and alkenoates except $C_{39}Me$ and $C_{40}Et$, which tend to always have dominant tri-unsaturated compounds but no tetra-unsaturated alkenones even at low temperatures. The culture data suggest that there are systematic differences in chemotaxonomic features of alkenone-producing haptophytes that may be helpful in identifying the groups of alkenone-producing haptophytes in situ.

The three alkenone-producing species in the family of Isochrysidaceae cultured here (Group II), *I. galbana*, *R. lamellosa* and *T. lutea*, have alkenone/alkenoate profiles that are similar to each other although there are still differences (Fig. 1, Supplementary Figs. S1–4). For example, tetra-unsaturated alkenones/alkenoates are dominant for *R. lamellosa* strains at growth temperature <14 °C (Supplementary Figs. S1 and S2; Supplementary Table S1), as has been found in previous studies (Marlowe et al., 1984a,

1984b; Rontani et al., 2004; Sun et al., 2007; Nakamura et al., 2014; Zheng et al., 2016a, 2016b). The *I. galbana* strain also makes significant amounts of tetra-unsaturated alkenone/alkenoates, but in general 10–30% lower than *R. lamellosa* at any given temperature. In comparison, the *T. lutea* strain only makes trace amounts of tetra-unsaturated alkenone and alkenoates even when grown at low temperatures (<9 °C) (Fig. 1, Supplementary Fig. S4). In fact, previous studies have found no tetra-unsaturated alkenones in *T. lutea* cultures grown at relatively high growth temperature (15–35 °C) (Marlowe et al., 1984a; Nakamura et al., 2016). Our study is the first to report the existence of small amounts of tetra-unsaturated alkenones/alkenoates for *T. lutea* strains when cultured at lower temperatures (<14 °C; Supplementary Table S1).

3.2. Alkenoate unsaturation index: U_{36Me}^E and U_{36Et}^E

In a study of *T. lutea* cultures using the VF-200 ms GC column, Nakamura et al. (2016) defined two alkenoate unsaturation indices (U_{36Me}^E and U_{36Et}^E) based on the relative abundance of di-, tri- and tetra-unsaturated alkenoates:

$$U_{36Me}^E = \frac{1}{4} \frac{C_{36:4}OMe + C_{36:2}OMe}{C_{36:4}OMe + C_{36:3}OMe + C_{36:3b}OMe + C_{36:2}OMe} \quad \delta 2b$$

$$U_{36Et}^E = \frac{1}{4} \frac{C_{36:4}OEt + C_{36:2}OEt}{C_{36:4}OEt + C_{36:3}OEt + C_{36:3b}OEt + C_{36:2}OEt} \quad \delta 3b$$

These equations are analogous to the definition of U_{37}^K for alkenones (Brassell et al., 1986). Fig. 2 shows the linear regressions of U_{36Me}^E (left) and U_{36Et}^E (right) for our six cultured strains plotted against growth temperature. The U_{36Me}^E and U_{36Et}^E indices for Group II and Group III strains are significantly correlated with temperature ($R^2 > 0.89$). We observe an unusual pattern, however, for *T. lutea* strain CCMP463, which has a weak correlation with a coefficient of determination R^2 value of only 0.331 and 0.247, respectively, for U_{36Me}^E and U_{36Et}^E indices. Close examination reveals an unusual reduction in these unsaturation index values with increasing temperature when growth temperature is <14 °C, opposite to the trends at higher temperatures (Fig. 2). Notably, previous culture experiments only reported data for *T. lutea* at temperatures above 15 °C (Nakamura et al., 2016), and hence did not identify such unusual behavior at lower temperatures.

In our previous study (Zheng et al., 2016b), we found that the U_{37}^{K0} , which excludes di-unsaturated alkenones, is more linearly correlated with growth temperature for the Group II species, whereas U_{37}^{K3} , which includes only di- and tri-unsaturated alkenones, is more linearly correlated with growth temperature for the Group III marine strains (Table 3). For comparison purposes, we have calculated the correlation coefficients of all three indices for alkenoates in our cultured strains (Table 3). Consistent with our previous observations, the double prime and no prime alkenone and alkenoate indices have higher R^2 values for Group II species while the single prime index has higher R^2 values for Group III species, as found in Zheng et al. (2016b).

3.3. Isomeric ratios of alkenoates and alkenones

All Group II and Group III strains produce both C_{36} methyl and ethyl alkenoates. Interestingly, only the tri-unsaturated ethyl C_{36} alkenoate (but not C_{36} methyl alkenoate) has an isomeric alkenoate with different double bond positions (Zheng et al., 2017; Fig. 1). These isomers were evaluated for their response to growth temperature. Our data also reveal major temperature responses for alkenoate and alkenone isomer ratios (Fig. 3). Araie et al. (2018) has proposed a new index RIA_{38} , the same format as the RIE_{36E} .

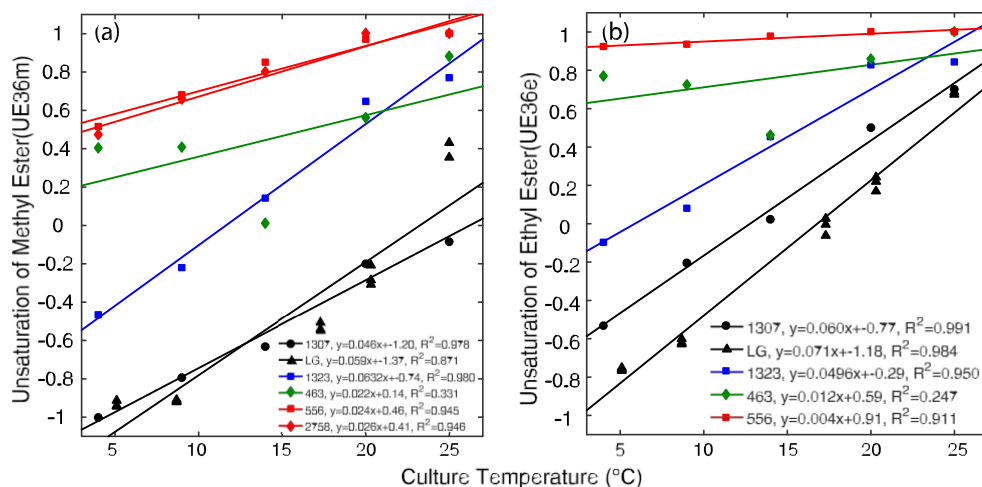


Fig. 2. Unsaturation index calibrations of C_{36} methyl and ethyl alkenoates for six haptophyte strains. (Black circle = LG strain; Black triangle = CCMP1307 R. lamellosa; blue square = CCMP715 I. galbana; Green diamond = T. lutea; Red square = CCMP2758 E. huxleyi; Red diamond = Van556 E. huxleyi). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2

Various unsaturation and isomeric indices of long chain alkenones and alkenoates discussed in this paper. Me = methyl alkenone, Et = ethyl alkenone, OMe = methyl alkenoate, OEt = ethyl alkenoate.

Index	Equation	Reference
U_{36Me}^E	$(36:2OMe - 36:4OMe)/(36:2OMe + 36:3OMe + 36:4OMe)$	Nakamura et al. (2016)
U_{36Me}^{E2}	$36:2OMe/(36:2OMe + 36:3OMe)$	Nakamura et al. (2016)
U_{36Me}^{E3}	$36:3OMe/(36:3OMe + 36:4OMe)$	This study
U_{36Et}^E	$(36:2OEt - 36:4OEt)/(36:2OEt + 36:3OEt + 36:4OEt)$	Nakamura et al. (2016)
U_{36Et}^{E2}	$36:2OEt/(36:2OEt + 36:3OEt)$	Nakamura et al. (2016)
U_{36Et}^{E3}	$36:3OEt/(36:3OEt + 36:4OEt)$	This study
U_{37}^K	$(37:2Me - 37:4Me)/(37:2Me + 37:3Me + 37:4Me)$	Brassell et al. (1986)
U_{37}^{K1}	$37:2Me/(37:2Me + 37:3Me)$	Prahl and Wakeham (1987)
U_{37}^{K2}	$37:3Me/(37:3Me + 37:4Me)$	Zheng et al. (2016b)
U_{38Et}^K	$(38:2Et - 38:4Et)/(38:2Et + 38:3Et + 38:4Et)$	Conte et al. (1998)
U_{38Et}^{K1}	$38:2Et/(38:2Et + 38:3Et)$	Conte et al. (1998)
U_{38Et}^{K2}	$38:3Et/(38:3Et + 38:4Et)$	Zheng et al. (2016b)
U_{38Me}^K	$(38:2Me - 38:4Me)/(38:2Me + 38:3Me + 38:4Me)$	Conte and Eglinton (1993)
U_{38Me}^{K1}	$38:2Me/(38:2Me + 38:3Me)$	Conte and Eglinton (1993)
U_{38Et}^{K2}	$38:3Me/(38:3Me + 38:4Me)$	This study
U_{39Et}^K	$(39:2Et - 39:4Et)/(39:2Et + 39:3Et + 39:4Et)$	This study
U_{39Et}^{K1}	$39:2Et/(39:2Et + 39:3Et)$	This study
U_{39Et}^{K2}	$39:3Et/(39:3Et + 39:4Et)$	This study
RIE_{36E}	$36:3aOEt/(36:3aOEt + 36:3bOEt)$	Araie et al. (2018)
RIK_{38E}	$38:3aEt/(38:3aEt + 38:3bEt)$	Longo et al. (2016)
RK2	$37:2Me/38:2Et$	This study
R3b	$37:3bMe/38:3bEt$	This study

(Ratio of Isomer Ester for C_{36} ethyl ester) index we used in this text:

$$RIE_{36E} = \frac{C_{36:3a}OEt}{C_{36:3a}OEt + C_{36:3b}OEt} \quad (4)$$

Temperature correlations for isomeric ratios in our cultured species are presented in Fig. 3. E. huxleyi Van556 strain and T. lutea CCMP463 strain make very small amounts of the $C_{36:3}OEt$ isomer. As a result, the absolute values of RIE_{36E} for E. huxleyi and T. lutea are high at all culture temperatures and ranged from 0.65 to 0.7 and from 0.73 to 0.88, respectively. In contrast, the values of RIE_{36E} for R. lamellosa and I. galbana are low at cold temperatures and increase with temperature, ranging from 0.3 to 0.8. In addition, all the culture species (Group II and Group III species) produce

abundant double bond positional alkenoate isomers while only Group I produce abundant alkenone isomers. $C_{36:3b}OEt$ ($D^{14,21,28}$) refers to the isomer of the regular $C_{36:3a}$ ethyl alkenoate ($D^{7,14,21}$).

RIE_{36E} is significantly correlated with growth temperature for most of the strains, with the exception of the T. lutea strain displaying a zigzag pattern with a relatively low R^2 value of 0.404 (Fig. 3a). The RIE_{36E} -temperature linear correlation coefficients for R. lamellosa LG strain, CCMP1307 strain and I. galbana CCMP1323 strain are 0.978, 0.975 and 0.891, respectively. One interesting observation from our data is that our three cultured Group II haptophytes (R. lamellosa and I. galbana strains) display similar temperature correlations (Fig. 3b), in contrast to the relatively large discrepancies for various unsaturation indices (Fig. 2).

The double bond positional isomer for $C_{38:3}$ ethyl ketones ($C_{38:3b}Et$) is produced by all three groups of haptophytes. However,

Table 3
Summary of the linear regression calibrations for various temperature indices in culture experiments carried out in this study. The full strain identification codes are given in Table 1.

Strains	Index	Slope	Intercept	R ²	p-Value
1307	RIE _{36E}	0.019	0.226	0.987	p < 0.05
1323	RIE _{36E}	0.014	0.349	0.944	p < 0.05
2758	RIE _{36E}	—	—	—	—
463	RIE _{36E}	0.007	0.735	0.636	ns
556	RIE _{36E}	0.003	0.707	0.995	ns
LG	RIE _{36E}	0.025	0.241	0.989	p < 0.05
1307	RIK _{38E}	0.000	0.840	0.010	ns
1323	RIK _{38E}	0.003	0.909	0.602	ns
2758	RIK _{38E}	0.000	1.000	0.000	ns
463	RIK _{38E}	0.000	1.000	0.000	ns
556	RIK _{38E}	0.000	1.000	0.000	ns
LG	RIK _{38E}	0.005	0.808	0.718	p < 0.05
1307	U ^{E1} _{36Et}	0.028	0.228	0.986	p < 0.05
1323	U ^{E1} _{36Et}	0.024	0.509	0.998	p < 0.05
2758	U ^{E1} _{36Et}	—	—	—	—
463	U ^{E1} _{36Et}	0.006	0.845	0.528	ns
556	U ^{E1} _{36Et}	—	—	—	—
LG	U ^{E1} _{36Et}	0.036	0.004	0.997	p < 0.05
1307	U ^{E1} _{36Me}	0.021	0.012	0.939	p < 0.05
1323	U ^{E1} _{36Me}	0.030	0.372	0.979	p < 0.05
2758	U ^{E1} _{36Me}	—	—	—	—
463	U ^{E1} _{36Me}	0.004	0.890	0.678	ns
556	U ^{E1} _{36Me}	—	—	—	—
LG	U ^{E1} _{36Me}	0.045	0.254	0.972	p < 0.05
1307	U ^{E2} _{36Et}	0.025	0.114	0.985	p < 0.05
1323	U ^{E2} _{36Et}	0.030	0.156	0.969	p < 0.05
2758	U ^{E2} _{36Et}	—	—	—	—
463	U ^{E2} _{36Et}	0.009	0.667	0.464	ns
556	U ^{E2} _{36Et}	0.005	0.899	0.980	p < 0.05
LG	U ^{E2} _{36Et}	0.028	0.073	0.970	p < 0.05
1307	U ^{E2} _{36Me}	0.034	0.365	0.955	p < 0.05
1323	U ^{E2} _{36Me}	0.037	0.142	0.997	p < 0.05
2758	U ^{E2} _{36Me}	0.026	0.409	0.973	p < 0.05
463	U ^{E2} _{36Me}	0.018	0.235	0.546	ns
556	U ^{E2} _{36Me}	0.024	0.464	0.972	p < 0.05
LG	U ^{E2} _{36Me}	0.012	0.081	0.650	p < 0.05
1307	U ^E _{36Et}	0.060	0.767	0.996	p < 0.05
1323	U ^E _{36Et}	0.050	0.292	0.975	p < 0.05
2758	U ^E _{36Et}	—	—	—	—
463	U ^E _{36Et}	0.001	0.690	0.046	ns
556	U ^E _{36Et}	0.005	0.899	0.980	p < 0.05
LG	U ^E _{36Et}	0.071	1.185	0.992	p < 0.05
1307	U ^E _{36Me}	0.044	1.196	0.996	p < 0.05
1323	U ^E _{36Me}	0.063	0.735	0.990	p < 0.05
2758	U ^E _{36Me}	0.026	0.409	0.973	p < 0.05
463	U ^E _{36Me}	0.022	0.143	0.575	ns
556	U ^E _{36Me}	0.024	0.464	0.972	p < 0.05
LG	U ^E _{36Me}	0.059	1.372	0.933	p < 0.05
1307	U ^{K1} ₃₇	0.032	0.136	0.994	p < 0.05
1323	U ^{K1} ₃₇	0.015	0.629	0.914	p < 0.05
2758	U ^{K1} ₃₇	0.013	0.636	0.974	p < 0.05
463	U ^{K1} ₃₇	0.001	0.965	0.583	ns
556	U ^{K1} ₃₇	0.001	0.968	0.619	ns
LG	U ^{K1} ₃₇	0.038	0.020	0.986	p < 0.05
1307	U ^{K1} _{38Et}	0.028	0.242	0.989	p < 0.05
1323	U ^{K1} _{38Et}	0.014	0.635	0.973	p < 0.05
2758	U ^{K1} _{38Et}	0.011	0.776	0.870	ns
463	U ^{K1} _{38Et}	0.000	1.000	0.000	ns
556	U ^{K1} _{38Et}	0.003	0.939	0.840	ns
LG	U ^{K1} _{38Et}	0.029	0.248	0.994	p < 0.05
1307	U ^{K1} _{38Me}	—	—	0.000	—
1323	U ^{K1} _{38Me}	—	—	—	—
2758	U ^{K1} _{38Me}	0.004	0.917	0.693	ns

Table 3 (continued)

Strains	Index	Slope	Intercept	R ²	p-Value
463	U _{38Me} ^{K⁰}	0.000	1.000	0.000	ns
556	U _{38Me} ^{K⁰}	0.002	0.957	0.693	ns
LG	U _{38Me} ^{K⁰}	0.032	0.391	0.975	p < 0.05
1307	U ₃₇ ^{K¹}	0.007	0.050	0.893	p < 0.05
1323	U ₃₇ ^{K¹}	0.005	0.040	0.844	ns
2758	U ₃₇ ^{K¹}	0.031	0.067	0.977	p < 0.05
463	U ₃₇ ^{K¹}	0.018	0.045	0.588	ns
556	U ₃₇ ^{K¹}	0.030	0.035	0.991	p < 0.05
LG	U ₃₇ ^{K¹}	0.008	0.011	0.788	p < 0.05
1307	U _{38Et} ^{K¹}	0.018	0.274	0.950	p < 0.05
1323	U _{38Et} ^{K¹}	0.020	0.014	0.820	ns
2758	U _{38Et} ^{K¹}	0.032	0.018	0.988	p < 0.05
463	U _{38Et} ^{K¹}	0.010	0.523	0.397	ns
556	U _{38Et} ^{K¹}	0.030	0.049	0.986	p < 0.05
LG	U _{38Et} ^{K¹}	0.019	0.118	0.927	p < 0.05
1307	U _{38Me} ^{K¹}	—	—	0.000	—
1323	U _{38Me} ^{K¹}	—	—	—	—
2758	U _{38Me} ^{K¹}	0.029	0.090	0.985	p < 0.05
463	U _{38Me} ^{K¹}	0.009	0.207	0.325	ns
556	U _{38Me} ^{K¹}	0.032	0.084	0.978	p < 0.05
LG	U _{38Me} ^{K¹}	0.006	0.040	0.954	p < 0.05
1307	U ₃₇ ^{K²}	0.042	0.871	0.994	p < 0.05
1323	U ₃₇ ^{K²}	0.021	0.339	0.932	p < 0.05
2758	U ₃₇ ^{K²}	0.046	0.447	0.984	p < 0.05
463	U ₃₇ ^{K²}	0.019	0.012	0.590	ns
556	U ₃₇ ^{K²}	0.031	0.006	0.992	p < 0.05
LG	U ₃₇ ^{K²}	0.047	1.076	0.992	p < 0.05
1307	U _{38Et} ^{K²}	0.053	0.580	0.987	p < 0.05
1323	U _{38Et} ^{K²}	0.036	0.363	0.936	p < 0.05
2758	U _{38Et} ^{K²}	0.043	0.206	0.982	p < 0.05
463	U _{38Et} ^{K²}	0.010	0.523	0.397	ns
556	U _{38Et} ^{K²}	0.033	0.010	0.985	p < 0.05
LG	U _{38Et} ^{K²}	0.053	0.714	0.989	p < 0.05
1307	U _{38Me} ^{K²}	—	—	0.000	—
1323	U _{38Me} ^{K²}	—	—	—	—
2758	U _{38Me} ^{K²}	0.029	0.090	0.985	p < 0.05
463	U _{38Me} ^{K²}	0.009	0.207	0.325	ns
556	U _{38Me} ^{K²}	0.032	0.084	0.978	p < 0.05
LG	U _{38Me} ^{K²}	0.037	0.649	0.985	p < 0.05

ns: not significant.

C_{38:3} methyl ketones do not contain the corresponding isomers for Group II and III species. C_{37:3b}Me isomers, C_{38:3b}Me isomers as well as C_{39:3b}Et isomers are unique to Group I haptophytes. To illustrate the isomeric ratios of C_{38:3b}Et to temperature, we calculate RIK_{38E} with

$$RIK_{38E} = \frac{C_{38:3a}Et}{C_{38:3b}Et} \quad (5)$$

Supplementary Fig. S7 shows the RIK_{38E} value of the culture samples. Due to the low abundance of the C_{38:3b}Et (especially for Group III species), the RIK_{38E} values of culture samples are all >0.75 and the correlation of RIK_{38E} with growth temperature is relatively weak, hence this index is potentially not very useful for paleotemperature reconstructions.

3.4. Alkenone unsaturation indices

With baseline separation for all alkenones and isomers, we can compare all unsaturation indices for all C₃₇, C₃₈ and C₃₉ alkenone homologues (U₃₇^K, U₃₇^{K¹}, U₃₇^{K²}, U_{38Et}^K, U_{38Et}^{K¹}, U_{38Et}^{K²}, U_{38Me}^K, U_{38Me}^{K¹}, U_{38Me}^{K²}).

U_{39Et}^K, U_{39Et}^{K¹}, U_{39Et}^{K²}) with respect to their correlation with growth temperature (Tables 2 and 3, Supplementary Fig. S8). Similar to the ester unsaturation index (i.e., U_{36Me}^E and U_{36Et}^E), the alkenone unsaturation of C₃₇, C₃₈ and C₃₉ (i.e., U_{37Me}^K and U_{38Et}^K) is linearly correlated with growth temperature in all the studied strains except for the T. lutea strain CCMP463. The R² values for R. lamellosa, I. galbana and E. huxleyi strains are all >0.88. In contrast, the U₃₇^K and U₃₈^K temperature calibrations of CCMP463 T. lutea show a weak correlation against growth temperature with R² values of only 0.349 and 0.158, respectively. Similar to alkenoate unsaturation ratios plotted in Fig. 2, T. lutea alkenone unsaturation ratios also show a turning point at around 15 °C, with the lower temperature range showing opposite trends relative to higher temperature range (Supplementary Fig. S8). The strong correlation between U^K proxies and growth temperature confirm the linearity of the U₃₇^K and U₃₈^K calibrations in previous culture studies (Conte et al., 1998; Nakamura et al., 2014; Zheng et al., 2016a, 2016b). In addition, even though the concentration of C₃₉ alkenones is lower than those of C₃₇ and C₃₈ alkenones, the unsaturation of C₃₉ is also linearly correlated to temperature change.

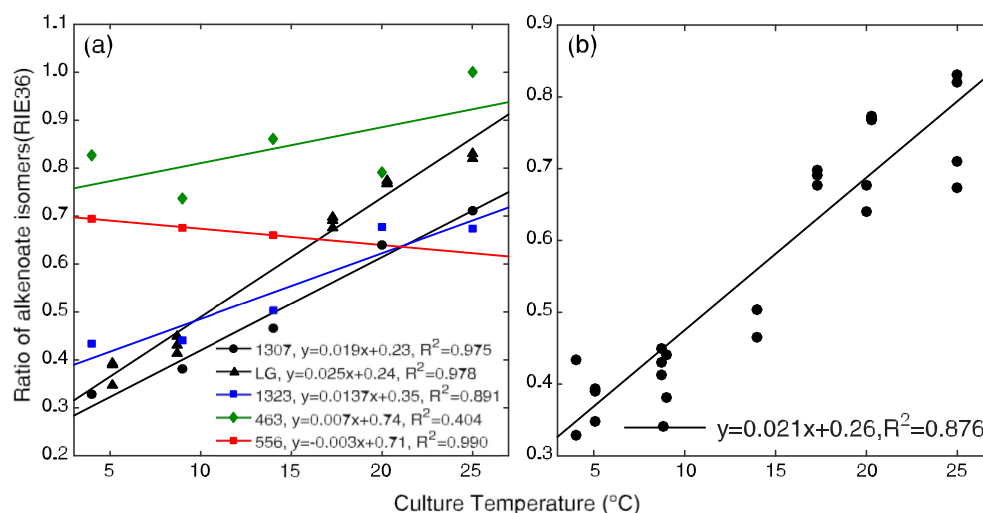


Fig. 3. RIE_{36E} index (ratio of C_{36:3a}OEt and C_{36:3b}OEt) for Group II and Group III haptophytes vs growth temperature: (a) RIE_{36E} calibrations for different cultured strains; (b) Combined RIE_{36E} calibration for the three Group II species excluding *T. lutea*.

3.5. Alkenones and alkenoates in *T. lutea* and comparison with Group III species and other Group II species

Tisochrysis represents one of the three genera of Isochrysidaceae (Group II, Theroux et al., 2010; with the other two genera being Isochrysis and Ruttnera), based on a recent phylogenetic classification (Bendif et al., 2013). Tisochrysis lutea was previously referred to as Isochrysis aff. galbana clone T-iso (also “Tahiti Isochrysis”) (Patterson et al., 1994). Nakamura et al. (2016) conducted the first systematic culture experiments for two strains of *T. lutea* (CCMP463, NIES-2590) at temperatures ranging from 10, 15, 20, 25, 30, 35 or 40 °C. The study suggested that *T. lutea* is adapted to relatively high environmental temperatures and grows poorly at 10 °C (as well as 40 °C).

The same study also reported that *T. lutea* does not produce tetra-unsaturated alkenones and alkenoates. In contrast to previous observations, however, we successfully grew *T. lutea* at temperatures as low as 4 °C. In contrast to the findings of Nakamura et al. (2016), we did find that this alga produces small amounts of tetra-unsaturated alkenones and alkenoates at low temperatures (Fig. 1, Supplementary Fig. S4). Notably, however, that the C_{37:4} methyl ketone is chromatographically difficult to separate from the C_{36:2} alkenoate using the VF-200 GC column, but such separation is complete using the RTX-200 column employed in the current study (Fig. 1). We found the C_{37:4} methyl ketone in both 4 °C and 9 °C culture samples (Supplementary Fig. S4; Supplementary Table S1).

The alkenone distributions of *T. lutea* differ significantly from *I. galbana* and *R. lamellosa* in that the %37:4 is low at low growth temperatures (Fig. 1, Supplementary Figs. S1–S4). In that sense, *T. lutea* is more similar to the Group III haptophytes than to other Group II species. One other feature also points to the greater similarity of *T. lutea* to Group III species: the relative abundances of C₃₇ and C₃₈ alkenones. For comparison, we computed the RK2 ratio ($RK2 = C_{37:2}Me/C_{38:2}Et$) to illustrate the relative abundance of di-unsaturated C₃₇ methyl and C₃₈ ethyl alkenones (we chose the ratio of C_{37:2}Me and C_{38:2}Et because di-unsaturated C₃₇Me and C₃₈Et are produced regardless of strain types or temperature range) (Fig. 4). Once again, RK2 values of *T. lutea* are more similar to *E. huxleyi* and *G. oceanica* with relatively high values which increase rapidly with rising temperatures. In contrast, RK2 values in Group II (and Group I) species are relatively low and insensitive to temperature changes (Fig. 4).

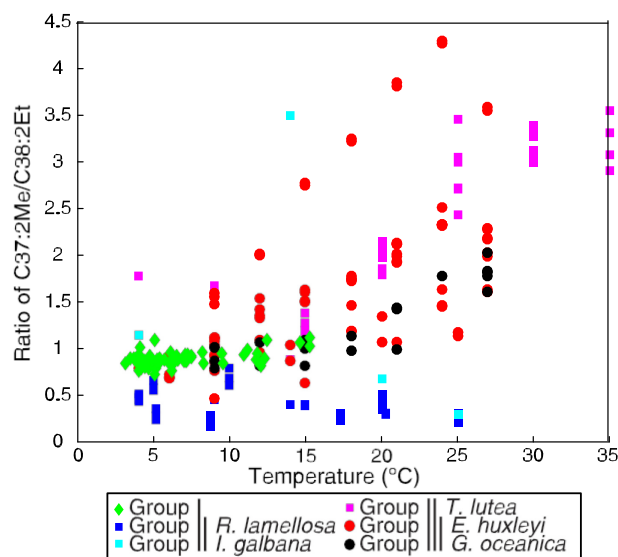


Fig. 4. Regression of RK2 index values (ratio of C_{37:2}Me and C_{38:2}Et) against temperatures among different species of haptophytes. Green diamonds represent Group I alkenones from Alaskan lakes (Longo et al., 2016). Squares represent Group II alkenones, with blue representing *R. lamellosa*, pink *T. lutea*, and cyan *I. galbana*. Circles represent Group III alkenones, with red denoting *E. huxleyi*, and black *G. oceanica*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.6. Principal component analysis (PCA)

To statistically evaluate alkenone and alkenoate distribution variations, PCA was performed on the distributions of alkenoates (C₃₆OMe and C₃₆OEt) and alkenones (C₃₇Me, C₃₈Et, C₃₈Me, C₃₉Et, C₃₉Me, C₄₀Et) from 230 culture and sediment data (sediment data for Group I since no pure culture is available) listed in Sections 2.1 and 2.2. The distributions of the alkenones and alkenoates are the normalized percentage of each compound relative to the total abundance of all alkenones and alkenoates for each sample (Supplementary Table S1).

Fig. 5 shows the scores of the alkenones/alkenoates and scores of the different samples. The three haptophyte Groups are in three non-overlapped regions on the PCA plot, indicating they have distinct chemotaxonomic signals, which can be elucidated with the

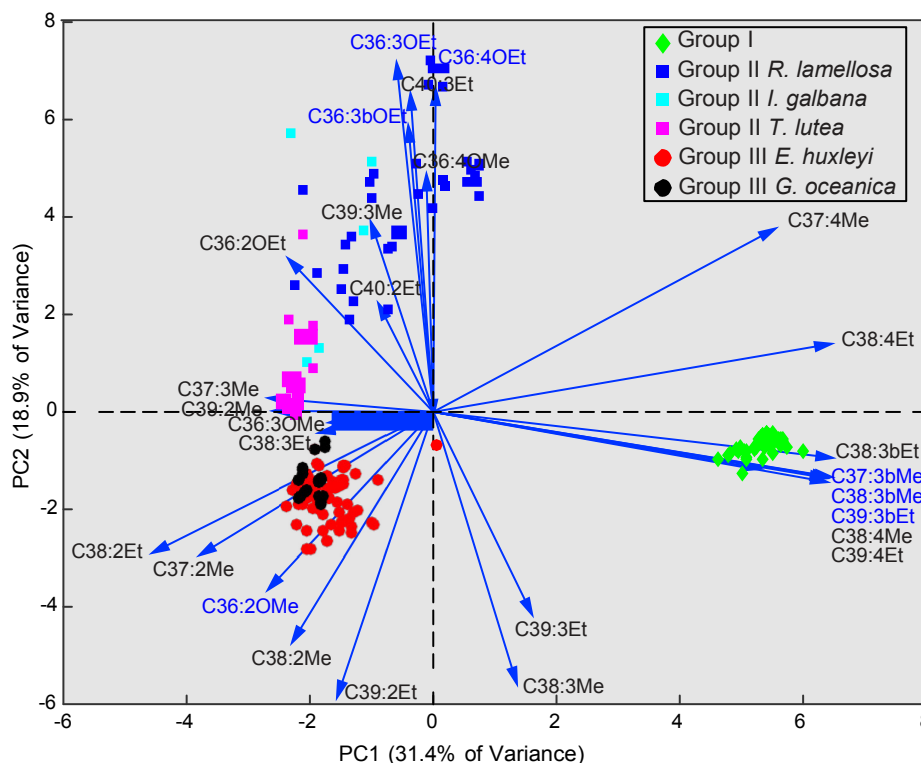


Fig. 5. Principal component analysis based on the standardized abundances (fractional total) of up to eight alkenoates and 24 alkenones from 230 culture and sediment samples used in this study. Plot shows the scores of the alkenones/alkenoates and scores of the different samples on PC1 (31.4% of variance) and PC2 (18.9% of variance). The scores of compounds are enlarged by eight times to have similar scale as the scores of the samples. The symbol and color designations are identical to those used in Fig. 4 for Group I, II and III haptophytes.

first and second principal components. Principal component 1 (31.4% of the total variance) primarily reflects the degree of unsaturation of the alkenones and alkenoates as well as alkenone isomers. These features distinguish the Group I from Group II and Group III, consistent with the chemotaxonomic features we described above. Principal component 2 (18.9% of the total variance) reflects that the fractional abundances of the $C_{36}OEt$, $C_{40}Et$, $C_{38}Me$ and $C_{39}Et$. $C_{38}Me$ and $C_{39}Et$ score negatively, whereas those of $C_{36}OEt$ and $C_{40}Et$ score positively. As shown in Fig. 1, Group II species synthesize $C_{36}OEt$ and $C_{40}Et$ but generally do not produce $C_{38}Me$ and $C_{39}Et$. The fractional abundance of alkenones and alkenoates can also be used to differentiate the Group I, Group II and Group III haptophyte strains, with Group I producing no alkenoates and Group II making alkenoates with higher degrees of unsaturation under the same growth temperatures.

3.7. Classification model using machine learning

The data above indicate LCAs and LCEs from different groups of alkenone-producing haptophytes have characteristically different distributions, as well as different temperature responses/calibrations. However, it is not always easy to visually distinguish the producer groups. To test if it is possible to accurately and objectively differentiate different groups of haptophytes, we developed a machine-learning classification algorithm with a k nearest neighbor classifier to build a classification model. Machine learning is a process of learning a set of rules from instances (examples in a training set), or more generally speaking, creating a classifier that can be used to generalize from new instances. Therefore, we need a training dataset that can train the algorithm to learn how to classify the alkenone/alkenoate assemblages. The full dataset of 230 cultures and sediment alkenone data for Group I haptophyte was

shuffled randomly. Seventy percent of the dataset was used to train the data, while the other 30 percent was used to test the accuracy of the model. The calculated accuracy for recognizing specific alkenone-producing haptophyte groups is 98%, which indicates this learning model is robust for identifying these groups when only one group of alkenone producer is present at a study site (Supplementary Fig. S9).

3.8. Quantifying the relative percentage inputs from different groups of alkenone producers

Marginal ocean environments such as estuaries can be complicated by the mixing of alkenones derived from open ocean Group III haptophytes (e.g., *E. huxleyi* and *G. oceanica*) and Group II species from brackish water environments. To determine the relative percentage inputs of different haptophyte groups at a study site with mixed species inputs, we built a multi-source model using the multivariate linear regression approach developed by Gao et al. (2011) for n-alkyl lipids. Our model employs the multivariate linear regression in the form of $Ax = B$, where A is a matrix of the average distribution of alkenones for five species (Lake George *R. lamellosa*, CCMP1307, CCMP1323, CCMP463, VAN556; we excluded the shorter chain alkenone producing species CCMP2758 here), B is a column vector of the alkenone distribution in the sediment sample, and variable vector x is the unknown fractional inputs of each of the five species/endmembers.

We applied this multi-source model to three Black Sea sediment samples from a gravity core GGC-18 (see Material and methods). Core GGC-18 was collected as one of the multicore sets with core GGC-19 that was analyzed by Xu et al. (2001). We estimated the sediment unit of the GGC-18 core according to the age model of the GGC-19 core based on our sampling depths (Xu et al., 2001).

For the purpose of demonstrating the model here, we use the averaged data from all growth temperatures (4 °C, 9 °C, 14 °C, 20 °C and 25 °C) from our four cultured species (Lake George R. lamellosa, CCMP1307, CCMP1323, CCMP463) to represent the Group II LCA and LCE profile and the data of VAN556 to represent Group III distributions. Fig. 6 shows the calculated fractional input of haptophyte Groups II and III from three sediment horizons: 2 (Unit I), 12 cm (Unit I), and 40 cm (Units II). Alkenones in both samples in Unit I are 100% derived from Group III, whereas the 40 cm sample from Unit II contain 41% Group II and 59% Group III alkenones.

4. Discussion

4.1. Quantitative assessment of alkenone inputs from different groups

The changes in the Group II and Group III species in the Black Sea suggested by our model (Fig. 6a) are generally consistent with the ancient DNA results reported by Coolen et al. (2009). Specifically, our model results suggest that both Group II and Group III species existed in Unit II sections, whereas Group III species are dominant in Unit I of the Black Sea. Similarly, genetic data indicate that Group III *E. huxleyi* haptophyte co-existed with Group II *Isochrysis*-related species in the Unit II section, but Unit I is dominated by Group III species. The changes in the input groups are readily visible from a visual inspection of the GC chromatograms (Fig. 6b). The most distinctive feature of the 40–42 cm Unit II sample is the much lower relative abundance of C_{38} methyl to C_{38} ethyl alkenones, which is consistent with the presence of Group II alkenones (Fig. 1).

4.2. Potential application of alkenoate unsaturation indices for paleotemperature reconstructions

Alkenone-producing haptophyte algae also make alkenoates that display changing degree of unsaturation in response to temperature changes (Marlowe et al., 1984a, 1984b; Brassell et al., 1986; Conte et al., 1998). Alkenoates are particularly abundant in the cold regions in the ocean (Conte and Eglinton, 1993; Rosell-Melé et al., 1994) and in some strains of *E. huxleyi* that inhabit high latitude coastal regions (Conte et al., 1994). In our culture samples of four different haptophyte species including *R. lamellosa*, *I. galbana*, *T. lutea* (Group II) and *E. huxleyi* strains Van556 and CCMP2758 (Group III), we found a strong increase in alkenoate/alkenone ratios (sum of alkenoates divided by the sum of all $C_{37:4}$ alkenones) as temperature decreases (Supplementary Fig. S10; Supplementary Table S3). Previous studies by Prahl et al. (1988) and Conte et al. (1992) developed the indices EE/K37 and AA₃₆ that combine both alkenones and alkenoates, to quantify relative differences in alkenoate vs alkenone concentrations. However, the application of these indices may be limited because of differential degradation rates for alkenones and alkenoates during diagenesis (Teece et al., 1998) and because alkenoate production shows significant genetic variability and is sensitive to algal physiological status (Conte et al., 1995, 1998). Since we can now fully resolve alkenoates from alkenones (Zheng et al., 2017), unsaturation indices based solely on alkenoates represent a new avenue of important applications for paleotemperature reconstructions, especially in cold regions where relative abundances of alkenoates increase at the expense of alkenones (Supplementary Fig. S9).

4.2.1. Alkenoate calibrations for estimating temperature

Calibration of alkenoate unsaturation indices of *T. lutea* species against growth temperature was previously carried out using data from a VF-200 ms GC column (Nakamura et al., 2016). However,

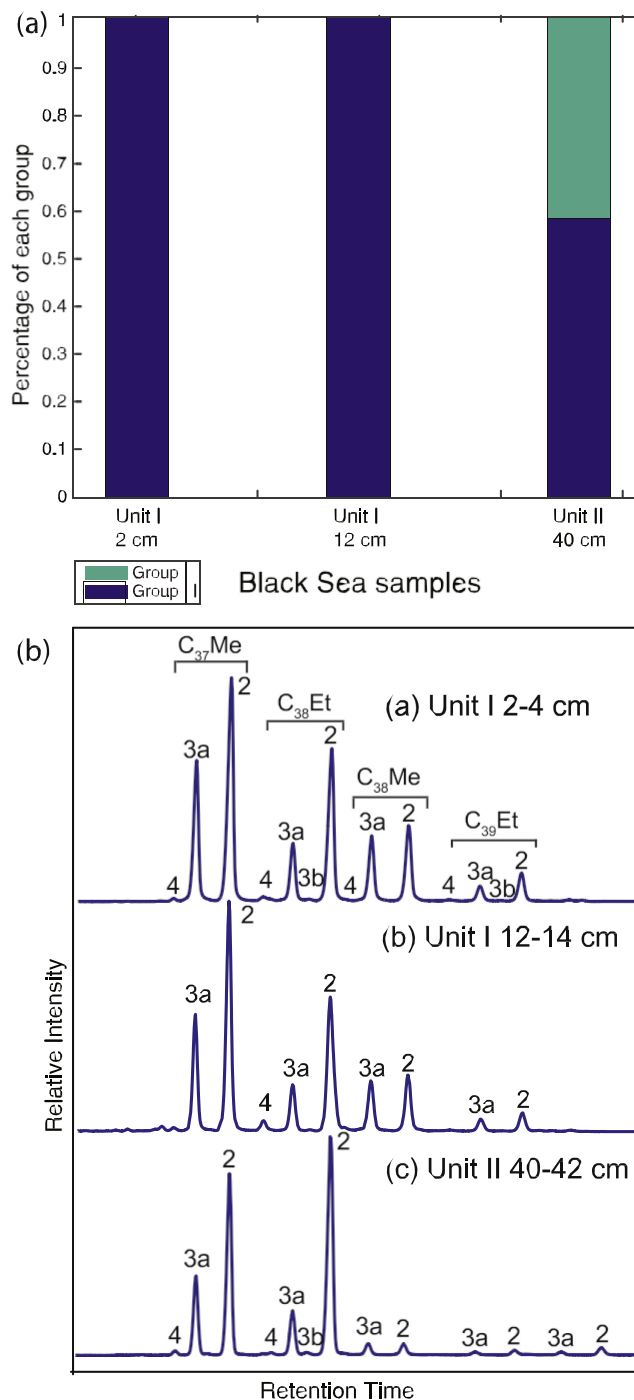


Fig. 6. Stacked bar plot showing the predicted percentage of the haptophyte groups at the Black Sea site based on the multi-source model (a) and the gas chromatographic profiles of alkenones in the corresponding black sea sediment samples (b).

this GC column cannot resolve $C_{36:2}$ OEt alkenoate from $C_{37:4}$ methyl ketone (Zheng et al., 2017). Since *T. lutea* does not produce the $C_{37:4}$ alkenone at growth temperatures >15 °C, VF-200 ms should have provided accurate data for calibration in Nakamura et al. (2016). However, our culture experiments show that *T. lutea* does produce small amounts of $C_{37:4}$ alkenone at low temperatures, and more importantly, all cultures of Group II and III species produce variable amounts of $C_{37:4}$ (in some cases $C_{37:4}$ is dominant at low growth temperatures) (Fig. 1, Supplementary Fig. S1–6), hence the $C_{36:2}$ OEt alkenoate must be fully resolved from the $C_{37:4}$ methyl ketone to accurately calibrate unsaturation indices

of LCAs and LCEs. In this study, we report the alkenoate unsaturation indices using the RTX-200 GC column which fully resolves $C_{36:2}OEt$ alkenoates from the $C_{37:4}$ methyl ketone (Zheng et al., 2017) for all six cultured Group II and Group III species (note Group I species do not produce alkenoates; Fig. 2).

The two *E. huxleyi* strains, CCMP2758 and VAN556, yield very similar temperature calibrations for U_{36Me}^E and U_{36Et}^E (Fig. 2). This is in contrast to the relatively large differences in temperature calibrations for C_{37} methyl ketones (Supplementary Fig. S8). Because alkenoates increase in abundance as temperature decreases, it is possible that U_{36Et}^E and/or U_{36Me}^E may provide accurate paleotemperature reconstructions in cold, high latitude regions such as the Nordic Seas. More culture experiments are needed to demonstrate if the observed similarity in alkenoate unsaturation indices is consistent across different *E. huxleyi* strains.

4.2.2. Alkenoate isomeric ratios

In this study, we also investigated the temperature sensitivity of the alkenoate double bond positional isomers. This is the first time that full resolution of the alkenoates isomers ($C_{36:3}OEt$ and $C_{36:3b}OEt$) has been achieved for temperature calibrations: the previous study by Araie et al. (2018) observed significant coelution of these two isomers. Longo et al. (2016) found that alkenone isomer ratios RIK_{38E} , RIK_{38M} , RIK_{39E} in the Greenland haptophytes (Group I) are significantly correlated with in situ temperatures with R^2 values up to 0.76, but the RIK_{37} value based on the isomeric ratios of $C_{37:3b}$ and $C_{37:3a}$ isomers is insensitive to temperature and can serve as a conservative ratio for diagnosing and quantifying species mixings. However, Group I Greenland haptophytes do not produce alkenoates, hence the ester isomer ratios are unique only to Group II and III haptophytes.

Our culture results reveal that the $C_{36:3}$ alkenoate isomer ratios (RIE_{36}) in Group II haptophytes (excluding *T. lutea* which will be discussed separately in Section 4.4) display stronger correlations with temperature (Fig. 3a; R^2 values 0.891–0.978) than the alkenone isomer ratios RIK_{38E} , RIK_{38M} , RIK_{39E} observed in suspended particulate matter samples in a series of Alaskan lakes (Longo et al., 2016). More importantly, our three cultures of *I. galbana* and *R. lamellosa* show quite consistent calibrations when all data points are combined ($y = 0.021x + 0.26$, $R^2 = 0.876$; Fig. 3b), in contrast to alkenoate or alkenone unsaturation ratios which display large interspecies differences in temperature relationships (Fig. 2, Supplementary Fig. S8). The application of alkenone unsaturation indices in brackish lacustrine systems or coastal brackish regions has encountered considerable difficulties in the past, partly due to the large variance in calibrations of *R. lamellosa* and *I. galbana* species vs temperature (e.g., Toney et al., 2012). Here we find that the RIE_{36E} ratios of *R. lamellosa* and that of *I. galbana* are more similar to each other than their corresponding alkenone and alkenoate unsaturation indices. Therefore, in environmental settings with mixed Group II (*I. galbana* and *R. lamellosa*) haptophytes, such as in many saline lakes (Theroux et al., 2010; Toney et al., 2012; Randlett et al., 2014), it is possible that RIE_{36} is a superior index to the unsaturation index for paleotemperature reconstruction that minimizes the impact of interspecies variations in temperature relationships. Notably, *T. lutea* displays large differences in temperature calibrations relative to *I. galbana* and *R. lamellosa*, hence the presence of *T. lutea* could be problematic. Fortunately, *T. lutea* has been found in only a few (mostly tropical) coastal regions (Nakamura et al., 2016) and has not yet been detected in lacustrine systems.

In addition to the possible inter-species consistency, the RIE_{36E} index is likely more stable than unsaturation ratios based on alkenoates or alkenones with different numbers of double bonds. For example, the $C_{37:3}$ alkenone has been found to degrade faster

than its $C_{37:2}$ homologue in certain environments, causing diagenetic distortion in U_{37}^K ratios (Gong and Hollander, 1999; Rontani et al., 2013). Since the isomers of C_{36} ethyl alkenoates contain the same numbers of double bonds, it is possible that diagenesis would be less selective in their degradation and cause diagenetic bias in RIE_{36E} ratios in sediment cores.

The Group III *E. huxleyi* culture shows an unusual behavior in its RIE_{36E} value. It varies in a very narrow range around 0.69 and is slightly negatively correlated with temperature. Unfortunately, we only have three samples grown at temperatures below 15 °C with sufficient esters to permit computation of this ratio. At higher temperatures, alkenoate concentrations are too low for accurate quantification of this ratio. However, this relatively consistent value within the small range of variation, if confirmed by more culture studies, could suggest that the RIE_{36E} value may be used as a species indicator, i.e., RIE_{36E} values around 0.69 indicate pure *E. huxleyi* inputs, analogous to the RIK_{37} ratio for Group I haptophytes (Longo et al., 2016). This is particularly relevant for colder regions of the ocean where there could be alkenone and alkenoate production from mixed haptophyte species: RIE_{36E} values significantly less than 0.69 signify production from Group II haptophytes.

4.3. Disentangling species effects in settings with mixed Group II and III inputs

Many studies have shown that mixing of different alkenone-producing haptophyte species at one site is common in brackish waters, saline lakes, estuaries and other coastal settings (Schulz et al., 2000; Coolen et al., 2009; Toney et al., 2012; Longo et al., 2016; Zheng et al., 2016b). In estuarine settings, C_{37} alkenones are often produced by both Group II and Group III species with variable proportions over time. Due to the distinctive U_{37}^K and U_{37}^K calibrations of different species, it is generally impossible to obtain quantitative paleotemperature reconstructions in such settings (Schulz et al., 2000; Coolen et al., 2009). Mixing might also extend to high latitude oceans such as the Nordic Sea (Zheng et al., 2016b), where U_{37}^K and U_{37}^K indices correlate poorly with sea surface temperatures (Rosell-Melé et al., 1998; Zheng et al., 2016b). Other explanations for the mismatching of alkenone-inferred temperatures and SST include: (1) different *E. huxleyi* genotypes co-existing in high latitude regions with different alkenone temperature responses, as indicated by mesocosm experiments in the Norwegian fjords (Conte et al., 1994); and (2) bias in seasonal production of alkenones towards the warmer seasons in high latitude North Atlantic (Tierney and Tingley, 2018). However, U_{37}^{KI} , rather than U_{37}^K or U_{37}^K , displays the strongest linear temperature regression with SST in the Nordic Sea – a characteristic of Group II haptophytes, suggesting that there could be substantial Group II haptophyte production of alkenones (Zheng et al., 2016b).

Because Group II haptophytes generally do not produce C_{38} methyl alkenones and C_{39} ethyl alkenones (Fig. 1; Lopez et al., 2005; Zheng et al., 2017), it is possible that we can take advantage of unsaturation ratios of C_{38} methyl and C_{39} ethyl alkenones in Group III haptophytes (*E. huxleyi* and *G. oceanica*) to reconstruct paleotemperatures without interference from Group II haptophytes. The greatly improved separation of C_{38} methyl and ethyl alkenones using the relatively polar GC columns (Longo et al., 2013; Zheng et al., 2017) allows for more accurate determination of unsaturation ratios for C_{38} methyl alkenones.

In order to apply the unsaturation index of C_{38} methyl ketones for paleotemperature reconstruction, we developed U_{38Me}^K linear calibration against temperature (Fig. 7) from both culture samples (including seven *E. huxleyi* strains and one *G. oceanica* strain) as well as ocean water column particulate samples (collected from

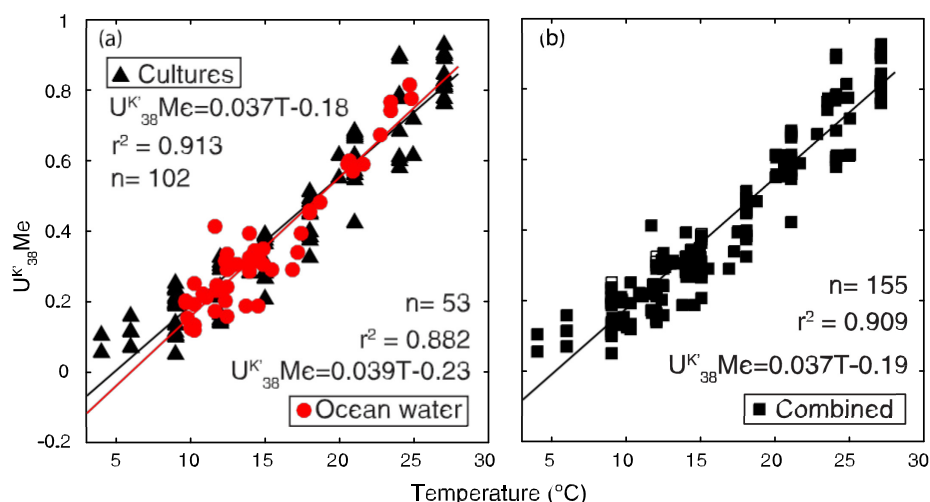


Fig. 7. Unsaturations calibrations of C_{38} methyl alkenones (U_{38Me}^{Kl}) based on cultured Group III marine strains (a) and in situ water samples from North Atlantic Ocean (b). The black triangles are all culture data from two *E. huxleyi* strains in this study and five *E. huxleyi* and one *G. oceanica* strains reported in Conte et al. (1998). The red circles represent suspended particulate matter reported in Conte and Eglinton (1993). Note that the U_{38Me}^{Kl} calibration vs temperature was previously only reported for a subset of suspended particulate samples ($>16^\circ\text{C}$). U_{38Me}^{Kl} calibration for the culture samples was reported individually for the five *E. huxleyi* and one *G. oceanica* strains. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the North Atlantic Ocean spanning a latitudinal range of 24°N to 62°N ; Conte and Eglinton, 1993). Surprisingly, we found that the combined culture and water column data are very consistent in their temperature regressions. Specifically, U_{38Me}^{Kl} calibrations of culture samples and in situ water column samples are almost identical with slope values of 0.037 and 0.039 respectively (Fig. 7a). Individually, the culture-based U_{38Me}^{Kl} is significantly positively correlated with temperature ($p < 0.01$; $R^2 = 0.91$) and the in situ U_{38Me}^{Kl} value is also strongly correlated with temperature ($p < 0.01$; $R^2 = 0.88$). We further tested whether the slopes for the two calibrations are statistically distinguishable or not using the student t-test. The calculated test statistic ($t = \frac{\bar{a} - \bar{b}}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$) is 0.0497, which

is smaller than the test critical value of 1.9758 ($\alpha = 0.05$, $df = n_1 + n_2 - 4 = 102 + 53 - 4 = 151$) and indicating that the slopes of the culture calibration and the in situ calibration of U_{38Me}^{Kl} are not statistically different. In Fig. 7b, we combine the culture samples and in situ water samples together to build a more general U_{38Me}^{Kl} . The slope of the combined U_{38Me}^{Kl} calibration (0.037) is slightly higher than the U_{38Me}^{Kl} calibration against the U_{37}^{Kl} -inferred temperature (0.032) derived by Herbert et al. (1998) using three sediment cores along the California margin. The samples we used were from all over the North Atlantic and from eight different *E. huxleyi* cultured strains while the samples used in Herbert et al. (1998) were downcore sediment samples from one small region, which may contribute to the small difference of the U_{38Me}^{Kl} calibrations.

In addition to $C_{38}\text{Me}$, $C_{39}\text{Et}$ alkenones are also unique to Group III marine species and may be used as a temperature proxy at sites with both Group II and III haptophytes. However, due to the analytical difficulty of measuring the relatively low abundance of $C_{39}\text{Et}$ compared to $C_{38}\text{Me}$ alkenones, we only have sufficient data for $C_{39}\text{Et}$ from culture samples. The culture-based U_{39Et}^{Kl} is $U_{39Et}^{Kl} = 0.038 \times T - 0.0197$ ($R^2 = 0.88$) (culture data from Conte et al. (1998) are included here) (Supplementary Fig. S11). It is very interesting to note that the calibration of U_{39Et}^{Kl} is extremely similar to U_{38Me}^{Kl} , suggesting a similar biosynthetic process leading to different degrees of unsaturation between $C_{38}\text{Me}$ and $C_{39}\text{Et}$ alkenones, as previously suggested by Conte et al. (1998).

Using alkenone and alkenoate offsets to disentangle mixed inputs has recently been proposed for Group I and II haptophyte systems in oligosaline lakes where seasonal or variable term mixing can occur (Longo et al., 2018). R3b, the ratio of $C_{37:3b}\text{Me}$ and $C_{38:3b}\text{Et}$ alkenones is found to correlate with environmental temperatures in Group I haptophytes (Longo et al., 2018). Because R3b consists of compounds that are only produced by Group I haptophytes, its use potentially eliminates Group II haptophyte influences. Our present culture study presents an alternative de-mixing approach for Group I and II contributions as Group I does not produce alkenoates whereas Group II does. Therefore, alkenoate unsaturation indices might be useful to disentangle Group I and Group II alkenones for paleotemperature reconstructions, especially at low temperatures. The identification of Group I haptophytes in environmental samples is relatively straightforward, as the $C_{37:3b}$ double bond positional isomer has only been observed in Group I haptophytes.

4.4. The unusual alkenone and alkenoate response to temperature of *T. lutea*

Based on the *T. lutea* culture experiment data of Nakamura et al. (2016), alkenones and alkenoates differ from those of other Isochrysidales species in following ways: (1) no tetraunsaturated alkenones and alkenoates are produced; (2) $C_{37:3}$ alkenones persist even at temperatures as high as 35°C , potentially allowing application of unsaturation indices for paleotemperature reconstructions at significantly higher temperatures; and (3) there are trace amounts of C_{38} methyl and C_{39} ethyl alkenones produced in *T. lutea*, in contrast to the general absence of these two compounds in *I. galbana* and *R. lamellosa* Group II haptophytes (Fig. 1). Nevertheless, the ratios of $C_{38}\text{Me}/C_{38}\text{Et}$ and $C_{39}\text{Et}/C_{38}\text{Me}$ alkenones in *T. lutea* are much lower than those in Group III (*E. huxleyi* and *G. oceanica*) haptophytes; (4) high ratios of $C_{37:2}\text{Me}$ over $C_{38:2}\text{Et}$ alkenones compared to Group II and I species (Fig. 4).

Our study has revealed additional unusual features of alkenones and alkenoates in *T. lutea* as compared to other Group II species, despite *T. lutea* being phylogenetically a Group II haptophyte (Bendif et al., 2013). First, *T. lutea* is in fact capable of growing at temperature as low as 4°C . Second, at growth temperatures of 4°C

and 9 °C, *T. lutea* makes small amounts of tetra-unsaturated alkenones and alkenoates (Supplementary Fig. S4; Supplementary Table S1), even though, as found by Nakamura et al. (2016), tetra-unsaturated homologues are absent at growth temperatures >15 °C. Third, there is a very clear reversal of temperature response for various unsaturation indices of alkenones and alkenoates at 4 °C and 9 °C: in effect, the degrees of unsaturation increase with temperature (Fig. 2, Supplementary Fig. S8). This trend is opposite to the common understanding of how alkenones produced by various haptophytes adjust the degrees of unsaturation for alkenone and alkenoates in response to temperature change (Conte et al., 1998; Prahl et al., 1988). Fourth, the ratio of $C_{37:2}Me/C_{38:2}Et$ and its temperature response is more similar to other Group III haptophytes than to Group II haptophytes (Fig. 4). Fifth, the alkenone and alkenoate distributions in *T. lutea* appear to assume intermediate values between Group III and Group II haptophytes (Figs. 4 and 5). On the PCA plot (Fig. 5), the *T. lutea* samples (pink squares; data include our results and those from Nakamura et al. (2016)) plot very close to the Group III *G. oceanica* samples (black circles), indicating a statistical similarity. We acknowledge that we only have limited culture data, and further replication in the future will be particularly useful for verifying our findings.

These observations suggest that *T. lutea* may represent an evolutionary intermediary between Group II and Group III haptophytes, although phylogenetically it is classified as one of the genera of Isochrysidaceae (Bendif et al., 2013). Some of the characteristics of alkenones and alkenoates of *T. lutea* are similar to *E. huxleyi*, e.g., low abundance of tetra-unsaturated homologues and high RK2 ratios, whereas others, such as the relatively low abundance of C_{38} methyl and C_{39} ethyl alkenones (Supplementary Fig. S4) are more akin to *I. galbana* and *R. lamellosa*. It is also interesting to note that at lower growth temperatures (<15 °C), the unsaturation indices trend towards the absolute values for *E. huxleyi*, which tends to have higher values (Fig. 2, Supplementary Fig. S8). Moreover, the phylogenetic tree of 40 strains covering Group II and Group III haptophytes shown in Fig. 1 of Nakamura et al. (2016) also suggests that *T. lutea* is more closely related to the common ancestor of Group II and Group III compared to *R. lamellosa* and *I. galbana*. Since Group III *G. oceanica* and Group II *T. lutea* were both isolated from tropic regions with high environmental temperatures (Conte et al., 1998; Nakamura et al., 2016), one likely scenario is that *T. lutea* may have evolved from an ancestor of *G. oceanica* adapted to coastal brackish waters. Evolutionary rates may have accelerated due to the high temperatures in the tropical seas (Brown, 2014).

5. Conclusions

Our culture experiments of six haptophyte species at different environmental temperatures, combined with previously published culture and field data, reveal several new taxonomical features of alkenone and alkenoate production and potentially expand the application of these biomarkers to environments where haptophyte species co-occur.

We have extended the culture-based temperature calibrations for unsaturation indices of all alkenones and alkenoates of different chain lengths with methyl and ethyl substitutions. We have also developed new temperature calibrations based on isomeric ratios of $C_{36:3}$ ethyl alkenoates (RIE_{36E}). Unlike alkenone unsaturation ratios, the RIE_{36E} index additionally displays major interspecies similarities among three different Group II haptophyte species (two *R. lamellosa* and one *I. galbana*) which may prove advantageous over various unsaturation ratios for paleotemperature reconstructions in environments where there are multiple Group II species. In addition, because the RIE_{36E} index is based on ratios

of biomarkers containing the same number of double bonds (all with three double bonds), any potential diagenetic influence arising from differential degradation rates may be minimized.

We have developed culture and field-based temperature calibrations for unsaturation ratios of the C_{38} methyl alkenone and C_{39} ethyl alkenone in *E. huxleyi*. Because these two compounds are generally absent in the Group II haptophytes, these unsaturation indices should be useful for paleotemperature reconstructions in brackish estuarine environments where both *E. huxleyi*, *I. galbana* and *R. lamellosa* may co-occur.

The alkenone and alkenoate distributions in *T. lutea* appear to be intermediate between Group II and Group III haptophytes, despite its phylogenetic classification in Group II. Combined with phylogenetic groupings, our data suggest a potential direct evolutionary relationship between *G. oceanica* and *T. lutea*.

We show that C_{36} alkenoate methyl and ethyl unsaturation indices display a strong linear relationship with temperature. Alkenoates have been reported to increase in abundance relative to alkenones as temperature decreases and we observe the same phenomenon in our cultures (Supplementary Fig. S3). Therefore, alkenoate-based unsaturation indices have potential for paleotemperature reconstructions in high latitude cold ocean waters. In addition, the alkenoate based temperature proxies can be applied in systems with mixed inputs from Group I and Group II haptophytes, since Group I species do not produce alkenoates.

Different groups of alkenone-producing haptophytes make alkenones and alkenoate distributions with sufficient chemotaxonomical difference permitting conclusive identification using a machine learning model. In systems with mixed inputs, we show it is possible to use a linear combination model to estimate percentage inputs from different groups.

Acknowledgments

This work was supported by the United States National Science Foundation awards to Y.H. (EAR-1122749, PLR-1503846, EAR-1502455; EAR-1762431). We would like to thank Fred Jackson and Christopher Claussen for their assistance during culture growth experiments. We are grateful for the constructive comments from three anonymous reviewers and Dr. John Volkman which helped improved this manuscript.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.orggeochem.2018.12.008>.

Associate Editor—Ann Pearson

References

- Araie, H., Nakamura, H., Toney, J.L., Haig, H.A., Planck, J., Shiratori, T., Leavitt, P.R., Seki, O., Ishida, K., Sawada, K., Suzuki, I., Shiraiwa, Y., 2018. Novel alkenone-producing strains of genus *Isochrysis* (Haptophyta) isolated from Canadian saline lakes show temperature sensitivity of alkenones and alkenoates. *Organic Geochemistry* 121, 89–103.
- Bendif, E.M., Probert, I., Schroeder, D.C., de Vargas, C., 2013. On the description of *Tisochrysis lutea* gen. nov. sp. nov. and *Isochrysis nuda* sp. nov. in the Isochrysidales, and the transfer of *Dicrateria* to the Prymnesiales (Haptophyta). *Journal of Applied Phycology* 25, 1763–1776.
- Bendle, J., Rosell-Melé, A., 2007. High-resolution alkenone sea surface temperature variability on the North Icelandic Shelf: implications for Nordic Seas palaeoclimatic development during the Holocene. *The Holocene* 17, 9–24.
- Brassell, S.C., Eglinton, G., Marlowe, I.T., Pflaumann, U., Sarnthein, M., 1986. Molecular stratigraphy: a new tool for climatic assessment. *Nature* 320, 129–133.
- Brown, J.H., 2014. Why are there so many species in the tropics? *Journal of Biogeography* 41, 8–22.

- Chu, G., Sun, Q., Li, S., Zheng, M., Jia, X., Lu, C., Liu, J., Liu, T., 2005. Long-chain alkenone distributions and temperature dependence in lacustrine surface sediments from China. *Geochimica et Cosmochimica Acta* 69, 4985–5003.
- Conte, M.H., Eglinton, G., 1993. Alkenone and alkenoate distributions within the euphotic zone of the eastern North Atlantic: correlation with production temperature. *Deep Sea Research Part I: Oceanographic Research Papers* 40, 1935–1961.
- Conte, M.H., Eglinton, G., Madureira, L.A.S., 1992. Long-chain alkenones and alkyl alkenoates as palaeotemperature indicators: their production, flux and early sedimentary diagenesis in the Eastern North Atlantic. *Organic Geochemistry* 19, 287–298.
- Conte, M.H., Thompson, A., Eglinton, G., 1994. Primary production of lipid biomarker compounds by *Emiliania huxleyi* - results from an experimental mesocosm study in Fjords of southwestern Norway. *Sarsia*, 319–331.
- Conte, M.H., Thompson, A., Eglinton, G., Green, J.C., 1995. Lipid biomarker diversity in the coccolithophorid *Emiliania huxleyi* (Prymnesiophyceae) and the related species *Gephyrocapsa oceanica*. *Journal of Phycology* 31, 272–282.
- Conte, M.H., Thompson, A., Leslay, D., Harris, R.P., 1998. Genetic and physiological influences on the alkenone/alkenoate versus growth temperature relationship in *Emiliania huxleyi* and *Gephyrocapsa oceanica*. *Geochimica et Cosmochimica Acta* 62, 51–68.
- Conte, M.H., Sicre, M.-A., Rühlemann, C., Weber, J.C., Schulte, S., Schulz-Bull, D., Blanz, T., 2006. Global temperature calibration of the alkenone unsaturation index (U_{37K}) in surface waters and comparison with surface sediments. *Geochimica et Cosmochimica Acta* 70, 1029–1044. <https://doi.org/10.1016/j.gca.2005.09.014>.
- Coolen, M.J.L., Saenz, J.P., Giosan, L., Trowbridge, N.Y., Dimitrov, P., Dimitrov, D., Eglinton, T.I., 2009. DNA and lipid molecular stratigraphic records of haptophyte succession in the Black Sea during the Holocene. *Earth and Planetary Science Letters* 284, 610–621.
- Cranwell, P.A., 1985. Long-chain unsaturated ketones in recent lacustrine sediments. *Geochimica et Cosmochimica Acta* 49, 1545–1551.
- Crump, B.C., Amaral-Zettler, L.A., Kling, G.W., 2012. Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *The ISME Journal* 6, 1629–1639.
- D'Andrea, W.J., Huang, Y., 2005. Long chain alkenones in Greenland lake sediments: low $\delta^{13}\text{C}$ values and exceptional abundance. *Organic Geochemistry* 36, 1234–1241.
- D'Andrea, W.J., Huang, Y., Fritz, S.C., Anderson, N.J., 2011. Abrupt Holocene climate change as an important factor for human migration in West Greenland. *Proceedings of the National Academy of Sciences of the United States of America* 108, 9765–9769.
- D'Andrea, W.J., Lage, M., Martiny, J.B.H., Laatsch, A.D., Amaral-Zettler, L.A., Sogin, M. L., Huang, Y., 2006. Alkenone producers inferred from well-preserved 18S rDNA in Greenland lake sediments. *Journal of Geophysical Research* 111, G03013.
- D'Andrea, W.J., Vaillencourt, D.A., Balascio, N.L., Werner, A., Roof, S.R., Retelle, M., Bradley, R.S., 2012. Mild Little Ice Age and unprecedented recent warmth in an 1800 year lake sediment record from Svalbard. *Geology* 40, 1007–1010.
- Dillon, J.T., Longo, W.M., Zhang, Y., Toranzo, R., Huang, Y., 2016. Identification of double-bond positions in isomeric alkenones from a lacustrine haptophyte. *Rapid Communications in Mass Spectrometry* 30, 112–118.
- Edge, J., Heimdal, R.B., 1994. Blooms of phytoplankton including *Emiliania huxleyi* (Haptophyta). Effects of nutrient supply in different N: P ratios. *Sarsia* 79, 333–348.
- Gao, L., Hou, J., Toney, J., MacDonald, D., Huang, Y., 2011. Mathematical modeling of the aquatic macrophyte inputs of mid-chain n-alkyl lipids to lake sediments: implications for interpreting compound specific hydrogen isotopic records. *Geochimica et Cosmochimica Acta* 75, 3781–3791.
- Gong, C., Hollander, D.J., 1999. Evidence for differential degradation of alkenones under contrasting bottom water oxygen conditions: Implication for paleotemperature reconstruction. *Geochimica et Cosmochimica Acta* 63, 405–411.
- Grimalt, J.O., Rullkötter, J., Sicre, M.-A., Summons, R., Farrington, J., Harvey, H.R., Goñi, M., Sawada, K., 2000. Modifications of the C₃₇ alkenone and alkenoate composition in the water column and sediment: possible implications for sea surface temperature estimates in paleoceanography. *Geochimica et Cosmochimica Acta* 64, 1029–1044. [https://doi.org/10.1016/S0016-7169\(00\)00053-3](https://doi.org/10.1016/S0016-7169(00)00053-3).
- Grossi, V., Raphael, D., Aubert, C., Rontani, J.F., 2000. The effect of growth temperature on the long-chain alkenes composition in the marine coccolithophorid *Emiliania huxleyi*. *Phytochemistry* 54, 393–399.
- Herbert, T.D., 2001. Review of alkenone calibrations (culture, water column, and sediments). *Geochimica et Cosmochimica Acta* 65, 1029–1044. [https://doi.org/10.1016/S0016-7169\(01\)00055-5](https://doi.org/10.1016/S0016-7169(01)00055-5).
- Herbert, T.D., Schuffert, J.D., Thomas, D., Lange, C., Weinheimer, A., Peleio-Alampay, A., Herguera, J.C., 1998. Depth and seasonality of alkenone production along the California margin inferred from a core top transect. *Paleoceanography* 13, 263–271.
- Ho, S.L., Mollenhauer, G., Lamy, F., Martínez-García, A., Mohtadi, M., Gersonde, R., Hebbeln, D., Nunez-Ricardo, S., Rosell-Melé, A., Tiedemann, R., 2012. Sea surface temperature variability in the Pacific sector of the Southern Ocean over the past 700 kyr. *Paleoceanography* 27, 1–15.
- Huguet, C., Kim, J.H., de Lange, G.J., Sinninghe Damsté, J.S., Schouten, S., 2009. Effects of long term oxic degradation on the U₃₇^K, TEX₈₆ and BIT organic proxies. *Organic Geochemistry* 40, 1188–1194.
- Kaiser, J., van der Meer, M.T.J., Arz, H.W., 2017. Long-chain alkenones in Baltic Sea surface sediments: new insights. *Organic Geochemistry* 112, 93–104.
- Lawrence, K.T., Herbert, T.D., Brown, C.M., Raymo, M.E., Haywood, A.M., 2009. High-amplitude variations in North Atlantic sea surface temperature during the early Pliocene warm period. *Paleoceanography* 24, 1–15.
- Longo, W.M., Dillon, J.T., Tarozo, R., Salacup, J.M., Huang, Y., 2013. Unprecedented separation of long chain alkenones from gas chromatography with a poly (trifluoropropylmethylsiloxane) stationary phase. *Organic Geochemistry* 65, 94–102.
- Longo, W.M., Theroux, S., Giblin, A.E., Zheng, Y., James, T., Huang, Y., 2016. Temperature calibration and phylogenetically distinct distributions for freshwater alkenones: evidence from northern Alaskan lakes. *Geochimica et Cosmochimica Acta* 180, 177–196.
- Longo, W.M., Huang, Y., Yao, Y., Zhao, J., Giblin, A.E., Wang, X., Zech, R., Haberzettl, T., Jardillier, L., Toney, J.L., Liu, Z., Krivonogov, S., Chu, G., D'Andrea, W.J.D., Harada, N., Nagashima, K., Sato, M., Yonenobu, H., Yamada, K., Gotanda, K., Shinozuka, Y., 2018. Widespread occurrence of distinct alkenones from Group I haptophytes in freshwater lakes: implications for paleotemperature and paleoenvironmental reconstructions. *Earth and Planetary Science Letters* 492, 239–250.
- Lopez, J.F., de Oteyza, T.G., Teixidor, P., Grimalt, J.O., 2005. Long chain alkenones in hypersaline and marine coastal microbial mats. *Organic Geochemistry* 36, 861–872.
- Marlowe, I.T., Brassell, S.C., Eglinton, G., Green, J.C., 1984a. Long chain unsaturated ketones and esters in living algae and marine sediments. *Organic Geochemistry* 6, 135–141.
- Marlowe, I.T., Green, J.C., Neal, A.C., Brassell, S.C., Eglinton, G., Course, P.A., 1984b. Long chain (n-C₃₇–C₃₉) alkenones in the Prymnesiophyceae. Distribution of alkenones and other lipids and their taxonomic significance. *British Phycological Journal* 19, 203–216.
- Mercer, J.L., Zhao, M., Colman, S.M., 2005. Seasonal variations of alkenones and UK37 in the Chesapeake Bay water column. *Estuarine, Coastal and Shelf Science* 63, 675–682.
- Nakamura, H., Sawada, K., Araie, H., Suzuki, I., Shiraiwa, Y., 2014. Long chain alkenes, alkenones and alkenoates produced by the haptophyte alga *Chrysotila lamellosa* CCMP1307 isolated from a salt marsh. *Organic Geochemistry* 66, 90–97.
- Nakamura, H., Sawada, K., Araie, H., Shiratori, T., Ishida, K., Suzuki, I., Shiraiwa, Y., 2016. Composition of long chain alkenones and alkenoates as a function of growth temperature in marine haptophyte *Tisochrysis lutea*. *Organic Geochemistry* 99, 78–89.
- Patterson, G.W., Wikfors, G.H., Gladu, P.K., Chitwood, D.J., 1994. Sterols and alkenones of *Isochrysis*. *Science* 35, 1233–1236.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2012. Scikit-learn: machine learning in python. *Journal of Machine Learning Research* 12, 2825–2830.
- Prahl, F.G., Wakeham, S.G., 1987. Calibration of unsaturation patterns in long-chain ketone compositions for paleotemperature assessment. *Nature* 330, 367–369.
- Prahl, F.G., Muehlhausen, L.A., Zahnle, D.L., 1988. Further evaluation of long-chain alkenones as indicators of paleoceanographic conditions. *Geochimica et Cosmochimica Acta* 52, 2303–2310.
- Randlett, M.-É., Coolen, M.J.L., Stockhecke, M., Pickarski, N., Litt, T., Balkema, C., Kwieciec, O., Tomonaga, Y., Wehrli, B., Schubert, C.J., 2014. Alkenone distribution in Lake Van sediment over the last 270 ka: influence of temperature and haptophyte species composition. *Quaternary Science Reviews* 104, 53–62.
- Rontani, J.-F., Beker, B., Volkman, J.K., 2004. Long-chain alkenones and related compounds in the benthic haptophyte *Chrysotila lamellosa* Anand HAP 17. *Phytochemistry* 65, 117–126.
- Rontani, J.-F., Volkman, J.K., Prahl, F.G., Wakeham, S.G., 2013. Biotic and abiotic degradation of alkenones and implications for paleoproxy applications: a review. *Organic Geochemistry* 59, 95–113.
- Rosell-Melé, A., Carter, J., Eglinton, G., 1994. Distributions of long-chain alkenones and alkyl alkenoates in marine surface sediments from the North East Atlantic. *Organic Geochemistry* 22, 501–509.
- Rosell-Melé, A., Weinelt, M., Sarntinoranont, M., Koç, N., Jansen, E., 1998. Variability of the Arctic front during the last climatic cycle: application of a novel molecular proxy. *Terra Nova* 10, 86–89.
- Schulz, H.-M., Schöner, A., Emeis, K.-C., 2000. Long-chain alkenone patterns in the Baltic sea—an ocean-freshwater transition. *Geochimica et Cosmochimica Acta* 64, 469–477.
- Sun, Q., Chu, G., Liu, G., Li, S., Wang, X., 2007. Calibration of alkenone unsaturation index with growth temperature for a lacustrine species, *Chrysotila lamellosa* (Haptophyceae). *Organic Geochemistry* 38, 1226–1234.
- Teece, M.A., Getliff, J.M., Leftley, J.W., Parkes, R.J., Maxwell, J.R., 1998. Microbial degradation of the marine prymnesiophyte *Emiliania huxleyi* under oxic and anoxic conditions as a model for early diagenesis: long chain alkenones, alkenones and alkyl alkenoates. *Organic Geochemistry* 29, 863–880.
- Tierney, J.E., Tingley, M.P., 2018. BAYSPLINE: a new calibration for the alkenone paleothermometer. *Paleoceanography and Paleoclimatology* 33, 281–301.
- Theroux, S., Toney, J., Amaral-Zettler, L., Huang, Y., 2013. Production and temperature sensitivity of long chain alkenones in the cultured haptophyte *Pseudoisochrysis paradoxa*. *Organic Geochemistry* 62, 68–73.
- Theroux, S., D'Andrea, W.J., Toney, J.L., Amaral-Zettler, L.A., Huang, Y., 2010. Phylogenetic diversity and evolutionary relatedness of alkenone-producing haptophyte algae in lakes: implications for continental paleotemperature reconstructions. *Earth and Planetary Science Letters* 300, 311–320.

- Toney, J.L., Huang, Y., Fritz, S.C., Baker, P.A., Grimm, E., Nyren, P., 2010. Climatic and environmental controls on the occurrence and distributions of long chain alkenones in lakes of the interior United States. *Geochimica et Cosmochimica Acta* 74, 1563–1578.
- Toney, J.L., Theroux, S., Andersen, R.A., Coleman, A., Amaral-Zettler, L.A., Huang, Y., 2012. Culturing of the first 37:4 predominant lacustrine haptophyte: geochemical, biochemical, and genetic implications. *Geochimica et Cosmochimica Acta* 78, 51–64.
- Tzanova, A., Herbert, T.D., 2015. Regional and global significance of Pliocene sea surface temperatures from the Gulf of Cadiz (Site U1387) and the Mediterranean. *Global and Planetary Change* 133, 371–377.
- Versteegh, G.J.M., Riegman, R., Leeuw, J.W. De, Jansen, J.H.F.F., 2001. U_{37}^{Kl} values for *Isochrysis galbana* as a function of culture temperature, light intensity and nutrient concentrations. *Organic Geochemistry* 32, 785–794.
- Volkman, J.K., Burton, H.R., Everitt, D.A., Allen, D.I., 1988. Pigment and lipid compositions of algal and bacterial communities in Ace Lake, Vestfold Hills, Antarctica. *Hydrobiologia* 165, 41–57.
- Xu, L., Reddy, C.M., Farrington, J.W., Frysinger, G.S., Gaines, R.B., Johnson, C.G., Nelson, R.K., Eglinton, T.I., 2001. Identification of a novel alkenone in Black Sea sediments. *Organic Geochemistry* 32, 633–645.
- Zheng, Y., Tarozo, R., Huang, Y., 2017. Optimizing chromatographic resolution for simultaneous quantification of long chain alkenones, alkenoates and their double bond positional isomers. *Organic Geochemistry* 111, 136–143.
- Zheng, Y., Dillon, J.T., Zhang, Y., Huang, Y., 2016a. Discovery of alkenones with variable methylene-interrupted double bonds: implications for the biosynthetic pathway. *Journal of Phycology* 1050, 1037–1050.
- Zheng, Y., Huang, Y., Andersen, R.A., Amaral-Zettler, L.A., 2016b. Excluding the di-unsaturated alkenone in the U_{37}^{Kl} index strengthens temperature correlation for the common lacustrine and brackish-water haptophytes. *Geochimica et Cosmochimica Acta* 175, 36–46.
- Zink, K.-G., Leythaeuser, D., Melkonian, M., Schwark, L., 2001. Temperature dependency of long-chain alkenone distributions in recent to fossil limnic sediments and in lake waters. *Geochimica et Cosmochimica Acta* 65, 253–265.