Noelaerhabdaceae coccolithophores as recorders of ancient atmospheric CO_2

Samuel Rice Phelps

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2020

Abstract

Noelaerhabdaceae coccolithophores as recorders of ancient atmospheric CO_2

Samuel Rice Phelps

Understanding the relationship between carbon dioxide and global climate is critical for predicting the severity of future climate change resulting from ongoing anthropogenic CO_2 emissions. By the year 2100, atmospheric CO_2 concentrations may reach close to 1000 ppm. The impacts of such a massive perturbation to the carbon cycle will last for hundreds of thousands of years. Estimates for the magnitude of warming range between ~3 and ~12°C, depending upon the timescale considered and actual CO_2 trajectory. To better understand the likely impacts of the Anthropocene, we must examine how CO_2 and climate have varied in the past, particularly during warm periods in Earth history with CO_2 levels higher than modern. Studies of this nature require the use of geological archives to reconstruct past environmental conditions.

The carbon isotope fractionation of marine algae recorded in alkenone biomarkers (ε_p) is one of the primary tools for reconstructing past atmospheric CO_2 variations. These molecules are produced by the Noelaerhabdaceae, a family of calcifying coccolithophorid algae that has flourished in the open ocean for tens of millions of years. Measurements of alkenone ε_p have offered critical insight into the evolution of Earth's atmosphere. However, CO_2 reconstructions using alkenone ε_p in the Pleistocene seldom agree with the atmospheric CO_2 variations known from ice cores. The reasons for this incongruity are not well understood. Recent studies have provided a deeper understanding of the complex mechanisms affecting alkenone carbon isotope fractionation, suggesting that the conventional framework for interpreting these records in terms

of CO_2 may require reconsideration. Quantitatively understanding the link between alkenone ε_p and CO_2 can provide new insights to the history of atmospheric CO_2 and global climate on Earth.

Here I use an empirical approach to explore relationships between physiology, environment, and alkenone carbon isotope fractionation in order to quantify the influence of CO_2 on alkenone ε_p. I begin with both new data and a meta-analysis of previous controlled laboratory studies on one of the dominant extant alkenone-producing algae. These experiments identify and quantify the significant roles of CO_2, irradiance, and cell size on alkenone ε_p. I calibrate multiple linear regression models from these culture data and find that when irradiance is included, growth rate is an insignificant contributor to ε_p variability. I test this model in core-top sediments and in Pleistocene deep-sea sediment sequences when atmospheric CO_2 concentrations are known from ice-core gas bubble analysis. With paired measurements of coccolith size and alkenone carbon isotope ratios, I find that although there is no spatial relationship between ε_p and CO_2 in the modern ocean, alkenone ε_p can be modeled reasonably well in core-top sediments, suggesting that the culture-based model is faithfully capturing the sensitivity of ε_p to irradiance, CO_2 and cell size in the modern ocean. Using Pleistocene alkenone ε_p and coccolith size records, I identify locations best-suited for paleobarometry, finding that tropical sites with stable surface hydrography and low variability in incident irradiance are sites where CO_2 and cell size dominate the ε_p signal during the late Pleistocene. I apply this model to a new deep-sea sediment record to reconstruct CO_2 changes over the last ~20 million years. This record exhibits a drop in atmospheric pCO_2 of ~600 μatm over this time period, and suggests that the global cooling and C_4 grassland expansion in the Miocene were likely caused by pCO_2 decline. Together, these results present a new way to understand alkenone ε_p variations through time and quantify past CO_2 changes.

Table of Contents

List of Figures

Acknowledgements

Completing this dissertation would not have been possible without the ever-present support, kindness, and generosity of so many. First, to my steadfast advisor, Pratigya Polissar. I feel incredibly fortunate to have been able to learn under your tutelage. I have felt more intellectual growth from our collaborations than I ever could have imagined. My experience in graduate school and the challenging and rewarding process of carrying out this dissertation was possible because of your guidance and encouragement. Thank you for allowing tangential but passionate and exciting conversations to take over an entire afternoon. I am extremely grateful for your always-open door, your flexibility, thoughtfulness, and support.

To Bärbel Hönisch and Peter deMenocal. You are both incredible scientists and exceptional role models. I hope I can have a career even half as successful and productive as you two. Thank you for your unwavering support and sage guidance. You have taught me to pay attention to the details, but to not lose sight of the big picture.

To Heather Stoll and Clara Bolton, who have been inspiring and wonderful collaborators throughout the entirety of this work. Thank you for welcoming me into your scientific lives and homes. Thank you both for your thoughtful and critical insights, and for helping me develop as a scientist throughout the evolution of this project. This all began with a proposal to collaborate in 2013, and I am very thankful that it took off.

To my cohort and friends at LDEO who made starting and enduring the journey to a PhD all the more enjoyable, especially Laura, Kyle, Frankie, Olivia, Maayan, Logan, Lorelei, Jesse, Ale, Allison, Mike, and Nick. Thank you to the other world-class instructors I was privileged to spend hours with, especially Mo Raymo, Jerry McManus, Bob Anderson, and Wally Broecker. To the Organic Geochemistry group, past and present, who I have been fortunate to call my

colleagues: Billy D'Andrea, Kevin Uno, Lorelei Curtin, Nicole deRoberts, Jon Nichols, Nick O'Mara, Helen Habicht, Rachel Lupien, Tammo Reichgelt, Hannah Rabinowitz, Gerry Rustic, Yoni Goldsmith, Cassy Rose, Nick Balascio, Francesco Muschitiello. Thank you for broadening my horizons every Friday with stimulating scientific conversations and entertaining banter. To Wei Huang and Nichole Anest, thank you for all your support in the isotope and core laboratories. Thank you to everyone on the CUSP team for helping enrich my academic, personal, and professional experiences as a member of the Columbia community.

I owe a sincere gratitude to my funding sources: The National Science Foundation GRFP, the French Embassy in the U.S., the International Drilling Program, Sigma Xi, the Geological Society of America, the World Surf League, the Chevron Student Initiative Fund at LDEO, and the Lamont Climate Center. Without these generous institutions, none of this would be possible.

To my family and friends, who have encouraged my curiosity and pursuits, and always kept things lighthearted and supported me during the rougher days of this journey. Mom and Dad, thank you for all you have done to put me on this path.

And to Colby, who has been with me through all of the ups and downs along this ride. You have been steadfast in your support and love. You have helped me keep my wits and my sanity. Thank you for your hours of proofreading, your patience on this long and often bumpy road, and your endless encouragement. You have been a guiding light throughout this process, and I could not have done it without you.

Finally, I am grateful to the dozens of careful and talented scientists whose labors have provided me with a fascinating topic of study.

Chapter 1: Introduction

1.1 Atmospheric CO_2 in the Earth System

The concentration of CO_2 in the atmosphere has increased by ~100 parts per million (ppm) since direct measurements began in the 1950s (Keeling, 1960; Dlugokencky and Tans, 2020). Annual anthropogenic emissions release around 10 gigatons of carbon (GtC) per year (Le Quéré et al., 2018), equivalent to ~5 ppm per year if all emissions were to remain in the atmosphere (Clark, 1982). However, only about 40% of the ~2300 GtC emitted since the preindustrial era have done so; the remaining 60% are absorbed by the ocean and biosphere (Le Quéré et al., 2018). Although CO_2 is a trace constituent of the atmosphere (~0.004%), it strongly absorbs infrared radiation that the Earth emits (Tyndall, 1860). This energy is re-radiated from the atmosphere, warming Earth's surface and maintaining surface temperatures within the habitable range, partially through water vapor feedbacks. The scientific community has known of the radiative nature of gaseous CO_2 for over 150 years. In 1856, experiments by Eunice Foote suggested that higher concentrations of CO_2 gas in closed chambers can increase the temperature of those systems (Foote, 1856). Towards the end of the 1800s, Svante Arrhenius provided the first calculations of surface temperature change associated with atmospheric CO_2 fluctuations (Arrhenius, 1897). One key question facing society and our planet is just how much Earth's surface temperature will increase given current and projected emissions.

Projections for atmospheric CO_2 concentrations by the year 2100 range between ~420 and >1000 ppm, depending on the balance between emissions and the success of emissions reductions and CO_2 removal initiatives (van Vuuren et al., 2011). A key metric for evaluating the effects of CO_2 change is climate sensitivity—defined as the mean surface temperature response to the radiative forcing from doubling atmospheric CO_2 concentrations above preindustrial levels

(Rohling et al., 2012). One approach to this question quantifies the temperature change once short-term feedbacks, which include the response of clouds, aerosols, water vapor, and sea ice, among others, respond to the +3.7 W m^{-2} radiative effect of a CO_2 doubling (Myhre et al., 1998). These adjustments typically happen on the timescale of decades to centuries (Rohling et al., 2012; Previdi et al., 2013). An early study by Charney et al. (1979) using nascent global climate models estimated an equilibrium climate sensitivity of 3 ± 1.5 K per CO_2 doubling, and decades of work have not narrowed the range of estimates significantly (Rohling et al., 2012). Whether climate sensitivity is closer to 1.5 K per doubling or to 4.5 K per doubling has profound implications for the trajectory of our planet in the coming centuries.

A perturbation in the carbon cycle cascades through the rest of the ocean-atmosphere-biosphere system directly through mass transfer and indirectly through climate feedbacks. The "slow" feedbacks, such as the melting of terrestrial ice sheets and large-scale changes in the distribution of vegetation, can cause substantial additional warming (Lunt et al., 2010), while also dramatically affecting sea level and terrestrial environments (Hansen et al., 2013; Sage, 2020). Because anthropogenic CO_2 will remain in the Earth system for thousands of years, and because the effects of resulting climate change will persist even longer (Archer et al., 2009), it is prudent to consider Earth system sensitivity (Previdi et al., 2013). Earth system sensitivity includes the temperature response of these slower feedbacks to a doubling of CO_2, and is higher than the Charney, or "fast-feedback" sensitivity (Lunt et al., 2010; Previdi et al., 2013; Royer, 2016).

It is therefore imperative to understand the natural variability in atmospheric CO_2 as well as the relationship between CO_2 and global climate at temperatures and CO_2 levels similar to those predicted for the next several decades. The instrumental record is helpful for understanding

the current radiative forcing by CO_2 and the global climate system response, but predicting the magnitude of future warming requires understanding CO_2 and climate variability during times in Earth history when temperature and CO_2 were higher than today. Determining thresholds in the climate system—such as ice sheet and ecosystem stability—to CO_2 perturbations requires investigation of the paleoclimate record. Identifying these tipping points is critical for understanding the full impact of committed and ongoing CO_2 emissions. To this end, we must use geological archives that record past temperature and CO_2 variability to improve our understanding of the response of global climate to CO_2 perturbations.

1.2 Reconstructing ancient atmospheric CO_2

A handful of methods exist for reconstructing past variations in atmospheric CO_2. Gas bubbles trapped in polar ice cores are the most precise archives of past CO_2 variability. A continuous record exists for the last ~800 kyr (Petit et al., 1999; Bereiter et al., 2015), with some snapshots between ~1 to 2 Ma (Higgins et al., 2015; Yan et al., 2019), though all CO_2 conditions were lower than modern. Estimates beyond this period require proxy methods that relate CO_2 to geochemical or biological signatures preserved in sedimentary archives. Prominent approaches include the stable carbon isotope ($\delta^{13}C$) ratios and arrangement of stomata in fossil leaves, the carbon isotope ratios of calcium carbonate nodules formed in soils, the boron isotope ratio of marine carbonates (primarily planktonic foraminifera), and the stable carbon isotope ratios of algal organic matter (Beerling and Royer, 2011). Together, these methods have revealed a close relationship between atmospheric CO_2 and global climate across a range of timescales including the glacial cycles of the Quaternary (Petit et al., 1999), the Pliocene (Martínez-Botí et al., 2015), and the early Cenozoic (Anagnostou et al., 2016). However, gaps in our understanding still

remain, in part due to varied proxy precision and accuracy. One key interval is the Late Miocene (~10 – 6 Ma), where globally-resolved sea surface temperature records document cooling on the order of 3 to 10°C (Herbert et al., 2016), but many CO_2 records show little change (Zhang et al., 2013; Sosdian et al., 2018). In this dissertation I focus on the algal CO_2 proxy and quantitatively evaluate the sensitivity of photosynthetic carbon isotope discrimination to changes in CO_2. Below I briefly describe the history and theoretical framework of the proxy that serve as a foundation for this dissertation.

1.3 The stable carbon isotope signatures of algal organic matter

1.3.1 Theoretical framework

Early measurements of algal organic matter (Sackett et al., 1965; Degens et al., 1968a; Degens et al., 1968b; Deuser et al., 1968; Wong and Sackett, 1978; Rau et al., 1982; Rau et al., 1992) revealed non-random relationships between the carbon isotope ratio of particulate organic carbon ($\delta^{13}C_{POC}$) and the environmental conditions under which that biomass was synthesized. Concurrent measurements of $\delta^{13}C_{POC}$ and aqueous CO_2 concentrations ($[CO_2]_{aq}$) in surface oceans demonstrated an inverse correlation between $\delta^{13}C_{POC}$ and $[CO_2]_{aq}$: higher $[CO_2]_{aq}$ was associated with more negative $\delta^{13}C_{POC}$ (Rau et al., 1989; Rau et al., 1992).

The framework for the algal-CO_2 proxy is developed from theory describing photosynthetic fractionation in plants, which are also isotopically depleted relative to ambient CO_2 (Craig, 1953). Traditional models (Vogel, 1980; O'Leary, 1981; Farquhar et al., 1982) frame carbon isotope fractionation during photosynthesis in two steps: (1) the diffusion of CO_2 into the cell (-4.4‰ in air, -0.7‰ in solution (O'Leary, 1984)), and (2) enzymatic carboxylation of CO_2 into an organic compound, typically by Ribulose bisphosphate carboxylase/oxygenase

(RubisCO), which has a canonical kinetic isotope fraction of ~25‰. The realized $\delta^{13}C_{POC}$ is a function of these two fractionating steps, where the fractionation by RubisCO is more strongly expressed when the rate of CO_2 supply to the site of carbon fixation exceeds the rate of carboxylation. In the converse, when the rate of CO_2 fixation is high relative to the rate of CO_2 diffusion, the fraction of CO_2 converted into organic carbon is also high, and the expression of RubisCO's ^{12}C preference is reduced. The degree of isotopic separation between organic matter and the CO_2 source is referred to as ε_p, which is defined by Freeman and Hayes (1992) as:

$$\varepsilon_p \equiv [(\delta^{13}C_{CO2aq} + 1000) / (\delta^{13}C_{POC} + 1000) - 1] \times 1000 \quad \text{[Eq. 1]}$$

Early works employed the carbon isotope ratios of primary photosynthetic products (e.g. geoporphyrins) and found that ε_p and global climate generally covaried over the Phanerozoic, and that ε_p may record past CO_2 variations (Popp et al., 1989). However, additional investigations identified algal physiology and growth environment as important influences on ε_p.

Several culture and field studies have found a correlation between smaller ε_p and higher algal productivity in addition to a relationship with $[CO_{2(aq)}]$ (Cifuentes et al., 1988; Fry and Wainwright, 1991; Laws et al., 1995). Francois et al. (1993) adopted the plant $\delta^{13}C$ model of Farquhar et al. (1982) to explain variations in ε_p measured in particulate organic carbon collected from the surface ocean across a large environmental gradient in the South Atlantic. Importantly, Francois et al.'s model incorporated the role of cell membrane permeability in determining CO_2 influx and formalized the effect of carbon fixation rate on algal $\delta^{13}C$, noting that higher carboxylation rates or lower permeabilities should produce lower ε_p values at the same $[CO_{2(aq)}]$. Rau et al. (1996) synthesized the modeling efforts of many earlier studies (Rau et al., 1992; Francois et al., 1993; Goericke et al., 1994) and provided one of the foundational mechanistic models of carbon isotope fractionation in algae. The model assumed diffusive supply of CO_2 and

quantified the dependence of ε_p on the physiological factors listed above, also including kinetic considerations of the CO_2 concentration at the cell surface based on the diffusivity of CO_2 in seawater and the conversion of HCO_3^- to CO_2 (Riebesell et al., 1993; Rau et al., 1996). In a set of nitrate-limited chemostat experiments with two diatoms and a coccolithophore, Popp et al. (1998b) found that the slope of the relationship between ε_p and the ratio of growth rate to CO_2 is proportional to the volume to surface area ratio of the alga. Their experiments produced an empirical framework to quantify the scaling factors between cellular carbon demand and diffusive CO_2 supply. In this work I reevaluate these relationships with a larger culture dataset and examine their applicability to the marine environment.

1.3.2 The alkenone paleobarometer

Because photosynthetic products in marine sediments are sourced from the entire algal community with varying physiology and growth parameters, subsequent works targeted alkenones, a specific class of organic biomarkers. These di- and tri-unsaturated ketones with 37 to 39 carbon atoms are found in most marine environments. In pelagic environments, alkenones are produced specifically by the coccolithophore family Noelaerhabdaceae. Modern alkenone producers include *Gephyrocapsa spp*. and *Emiliania huxleyi*, a cosmopolitan algae that is likely the most abundant coccolithophore in the modern ocean (Winter et al., 2014). The calcium carbonate platelets, called coccoliths, that these algae secrete to shield their cell bodies are also preserved in sediments. Fossil coccolithophores are identified at the species level based on the morphometry of their coccoliths. Because alkenones have a unique biological source, they have been used for targeted algal paleobarometry since the 1990s. Advancements in coccolithophore micropaleontology have since identified ways to reconstruct the cell sizes of ancient alkenone-

producing algae (Henderiks and Pagani, 2007), but proxies for growth rate remain elusive. Because of these difficulties, early applications of the alkenone CO_2 proxy used a general hyperbolic equation to relate ε_p to CO_2 (Jasper et al., 1994):

$$\varepsilon_p = \varepsilon_f - b/CO_2 \quad \text{[Eq. 2]}$$

where ε_f is the fractionation by RubisCO (typically 25‰), CO_2 is the concentration of aqueous CO_2 (μmol kg^{-1}), and *b* accounts for all non-CO_2 effects on carbon isotope fractionation, including cell size, growth rate, carbon acquisition pathways, among others. Early calibration work that measured ε_p in suspended particulate organic matter in the modern ocean found a strong relationship between *b* and in-situ phosphate concentrations (Bidigare et al., 1997). Applications of the alkenone paleobarometer have typically used modern phosphate concentrations at the core location, and the modern *b*-PO_4 calibration to interpret temporal variations in ε_p in terms of CO_2. The empirical coefficient *b* describes the slope of the relationship between ε_p and $1/CO_2$, and in the conventional framework *b* is directly proportional to cell radius and growth rate. Apart from some recent advancements attempting to quantify algal physiology more discretely, Equation 2, termed the "conventional model" or the "diffusive model" is the basis for alkenone paleobarometry.

Alkenone ε_p records from the Cenozoic show a dramatic decline that reflects global climate trends over the last ~40 Myr: ε_p is substantially higher during the hothouse of the Paleogene compared to the icehouse of the Pleistocene (Fig 1.1). However, in detail, ε_p variations show less correspondence with CO_2 than expected over the Pleistocene, and records from different locations spanning the same time interval have yielded markedly different results (Badger et al., 2019; Zhang et al., 2019). Recent works have suggested that the conventional diffusive model does not properly quantify the scaling between CO_2 and ε_p (Badger et al., 2019;

Stoll et al., 2019). Others have recently argued that carbon acquisition and carbon isotope fractionation are far more complicated processes than their descriptions in the conventional model (Rost et al., 2003; Holtz et al., 2017; McClelland et al., 2017; Wilkes and Pearson, 2019). As the cellular mechanics driving carbon isotope fractionation in algae are still being discovered, here I take an empirical approach to define relationships from culture experiments and test their efficacy for understanding ε_p variations through time.

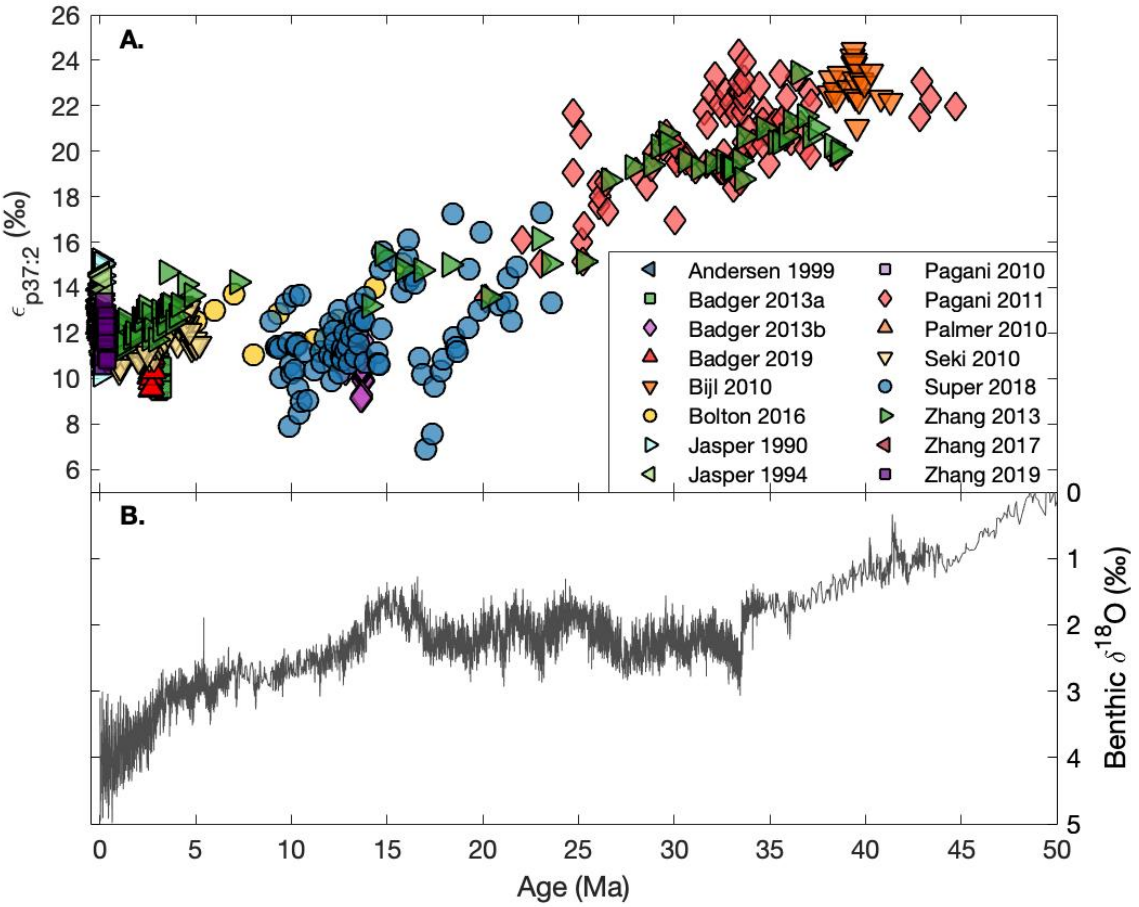

Figure 1. Cenozoic alkenone ε_p records. (A) Existing records of carbon isotope fractionation in $C_{37:2}$ alkenones, biomarkers produced by the coccolithophore family Noelaerhabdaceae. The data show a large decrease from ~23‰ in the Paleogene to ~13‰ in the Pleistocene. Legend entries show the first author's last name and the year of publication. The data of Pagani et al., (1999) and Pagani et al. (2005) are not shown because the planktic $\delta^{18}O$ values used to calculate SST are diagenetically altered. (B) Benthic $\delta^{18}O$ stack reflecting the cooling of the deep ocean and the growth of land ice from ~50 Ma to the Pleistocene. Shown is a 5-point smoothing of the original record of Zachos et al. (2008).

1.4 Dissertation outline and objectives

This dissertation focuses on improving our understanding of the processes that determine the stable carbon isotope signatures of alkenone biomarkers, and evaluating their utility as records of past environmental conditions. My approach is informed by theoretical considerations, but the implementation is largely empirical. I build upon decades of careful culture, marine sampling, and sediment studies that have documented some of the fundamental processes that control carbon isotope fractionation. This dissertation is comprised of four chapters that have been written as manuscripts for publication. In Chapters 2 through 4, I use new analyses in conjunction with data from the literature to provide an updated, standardized, and holistic view of alkenone carbon isotope fractionation in cultures, and spatiotemporally in the modern and Pleistocene ocean. In Chapter 5, I apply the method developed in Chapter 2 to reconstruct atmospheric CO_2 over the last ~20 million years from a core site in the Western Equatorial Pacific. Chapter 2 was conducted at LDEO in collaboration with Dr. Gwenn Hennon, Dr. Sonya Dyhrman, María Hernández Limón, and Olivia Williamson, and a manuscript is currently under review at *Geochimica et Cosmochimica Acta*. Chapters 3 through 5 were part of a collaboration with Dr. Heather Stoll, and they include work completed at the Centre Européen de Recherche et d'Enseignement des Géosciences de l'Environnement (CEREGE, Aix-en-Provence, France) under the guidance of Dr. Clara Bolton and Dr. Luc Beaufort.

I first examine the fundamental controls on organic carbon isotope fractionation in the Noelaerhabdaceae through laboratory culture experiments. **Chapter 2** builds off a recent paper by Stoll et al. (2019), who presented the first multiple linear regression model for predicting algal ε_p from growth rate, cell radius, and natural log transformations of CO_2 and light intensity. This chapter reports results from cultures of the alkenone producer *Gephyrocapsa oceanica*

under manipulated CO_2 concentrations at variable pH and finds that the ε_p response of *G. oceanica* is similar to existing studies with variable CO_2 at constant pH. I compile all available literature data using strict quality control and a data standardization approach that accounted for temporal differences in culture harvesting that can introduce systematic offsets in estimated cell size between and within studies. Using this compilation and my new data, I provide an updated set of equations to model ε_p as a function of irradiance, growth rate, cell radius, and CO_2, with low residual unexplained variance. The regression model identifies growth rate as an insignificant contributor to ε_p variability across the full spectrum of growth conditions, which will be important in Chapters 3 through 5. Chapter 2 confirms that alkenone ε_p does indeed record ambient CO_2 information, but that irradiance and cell radius are as important predictors of ε_p and are required to explain the full range of ε_p observed in cultures, the modern ocean, and the geologic record. This chapter is the groundwork for the remainder of the dissertation, as it provides a means for quantitatively relating ε_p to its physiological and environmental controls.

Chapter 3 pivots to the natural environment to test whether the relationships observed in culture experiments are apparent in the modern and Late Holocene ocean. I use new measurements of alkenone $\delta^{13}C$ and coccolith size from a suite of core-top sediments, and contextualize these data by compiling all available alkenone $\delta^{13}C$ measurements from core-top and water column particulate samples, as well as studies of coccolith length and coccolithophore abundance to estimate cell size from the literature. I take advantage of existing field measurements of growth rates of alkenone producing coccolithophores and estimate growth rates globally via a simple nutrient limitation model.

This study is the broadest assessment of the response of ε_p in the global calibration set to the three canonical variables in the conventional model: CO_2, growth rate, and cell size. In

contrast to the culture observations in Chapter 2, the data from the marine environment show no spatial relationship between ε_p and CO_2. However, larger ε_p is found at higher irradiance relative to nutrient concentrations, mirroring culture observations. I use the sediment and particulate data to examine how well the conventional model (Eq. 2) and the culture-based multiple linear regression models explain ε_p variability in the modern ocean. The conventional model prescribes a positive correlation between growth rate and *b*, but the available data show these two variables have no relationship. Furthermore, I show with an expanded dataset that the relationship between *b* and phosphate that has been used to calibrate the conventional model arises because of an underlying correlation between *b* and $[CO_{2(aq)}]$. Together these findings suggest that the traditional approach does not accurately describe carbon isotope fractionation in these algae and that use of the *b* parameter in paleo-CO_2 studies leads to substantial errors in both absolute and relative CO_2 values.

I then test whether the multiple linear regression models developed in Chapter 2 can provide a way forward for interpreting ε_p in the ocean. While the model does not predict ε_p in particulate samples, it performs reasonably well when applied to core-top sediments. This difference is attributed to the difference in irradiance conditions at a surface ocean particulate sampling station compared to the conditions integrated by the alkenone flux out of the mixed layer. These results demonstrate that ε_p variations in sediments can be explained using the quantitative relationships derived from cultures in Chapter 2.

Because irradiance is one of the input variables determining carbon isotope fractionation, modeling ε_p with the regressions derived from cultures depends on the absolute value of irradiance used. This presents a challenge for paleo studies because there is no robust proxy for past irradiance. I examined this issue in **Chapter 4** by analyzing two new late-Pleistocene ε_p and

coccolith size records, as well as all published late-Pleistocene ε_p and available coccolith size data near core sites with alkenone ε_p records. My culture-based model shows ε_p is moderately sensitive to CO_2, and given the surface ocean aqueous CO_2 variability on glacial/interglacial timescales of ~3 μmol kg^{-1}, the resulting ε_p change should be small (~1.5‰) if irradiance is stable. I found that at some locations, it is possible to model ε_p given the measured coccolith size history and the known CO_2 history from ice-core records. Locations with low intra-annual variability in surface irradiance, mixed layer depth, and turbidity—all of which affect irradiance at the depth of alkenone production—are the best targets for paleobarometry. This chapter proposes that we can understand ε_p variations through time if irradiance conditions are demonstrably stable.

Chapter 5 uses the inferences from Chapter 4 to interpret a new record of alkenone ε_p and coccolith size from a core site in the Western Equatorial Pacific. This location appears to have relatively stable irradiance conditions in the Pleistocene, permitting the use of the multiple linear regression model to interpret ε_p and coccolith size changes in terms of CO_2. I reconstructed changes in surface ocean (and by inference, atmospheric) pCO_2 over the last ~20 million years and found a ~600 μatm decrease from the Middle Miocene (~16 Ma) to the Pleistocene. Using estimates of global temperatures for the Middle and Late Miocene, I estimate Earth system sensitivity of ~5 K per CO_2 doubling, in agreement with estimates for other warm periods in the Cenozoic. The timing of CO_2 decline corresponds with the expansion of C_4 grasslands across the tropics and subtropics, suggesting a proximal role for CO_2 in this ecological revolution in support of the original hypotheses for the emergence of these ecosystems (Cerling et al., 1997).

This dissertation addresses some of the apparent inconsistencies of the conventional alkenone paleobarometer that were brought to light in 2019 (Badger et al., 2019; Zhang et al.,

2019; Stoll et al., 2019). It identifies some of the factors responsible for the observed ε_p variability in space and time. I argue that the conventional diffusive model does not describe alkenone carbon isotope fractionation in cultures or in the natural environment, suggesting CO_2 records using this model require reevaluation. The culture-based multiple linear regression model presents a way forward, but even the calibration and application are still in their infancy. Additional culture experiments at untested combinations of CO_2, irradiance, and cell size can help to refine the calibration equations. Development of paleoirradiance proxies can help to constrain the largest source of uncertainty in the application of the proxy. Overall, this dissertation presents a new view on the alkenone paleobarometer and Miocene CO_2, and provides a roadmap for future studies to test and refine this proxy.

1.5 References

Anagnostou E., John E. H., Edgar K. M., Foster G. L., Ridgwell A. J., Inglis G. N., Pancost R. D., Lunt D. J. and Pearson P. N. (2016) Changing atmospheric CO2 concentration was the primary driver of early Cenozoic climate. *Nature* **533**, 380–384.

Archer D., Eby M., Brovkin V., Ridgwell A. J., Cao L., Mikolajewicz U., Caldeira K., Matsumoto K., Munhoven G., Montenegro A. and Tokos K. (2009) Atmospheric Lifetime of Fossil Fuel Carbon Dioxide. *Annu. Rev. Earth Planet. Sci.* **37**, 117–134.

Arrhenius S. (1897) On the Influence of Carbonic Acid in the Air upon the Temperature of the Earth. *Publ. Astron. Soc. Pacific* **9**, 14.

Badger M. P. S., Foster G. L., Chalk T. B., Gibbs S. J., Badger M. P. S., Pancost R. D., Schmidt D. N., Sexton P. F., Mackensen A., Bown P. R. and Pälike H. (2019) Insensitivity of alkenone carbon isotopes to atmospheric CO2 at low to moderate CO2 levels. *Clim. Past* **15**, 539–554.

Beerling D. J. and Royer D. L. (2011) Convergent Cenozoic CO2 history. *Nat. Geosci.* **4**, 418–420.

Bereiter B., Eggleston S., Schmitt J., Nehrbass-Ahles C., Stocker T. F., Fischer H., Kipfstuhl S. and Chappellaz J. (2015) Revision of the EPICA Dome C CO2 record from 800 to 600-kyr before present. *Geophys. Res. Lett.* **42**, 542–549.

Bidigare R. R., Fluegge A., Freeman K. H., Hanson K. L., Hayes J. M., Hollander D., Jasper J. P., King L. L., Laws E. A., Milder J., Millero F. J., Pancost R., Popp B. N., Steinberg P. A. and Wakeham S. G. (1997) Consistent fractionation of 13C in nature and in the laboratory: Growth-rate effects in some haptophyte algae. *Global Biogeochem. Cycles* **11**, 279–292.

Cerling T. E., Harris J. M., Macfadden B. J., Leakey M. G., Quadek J., Eisenmann V., Ehleringer J. R., Quade J., Eisenmann V. and Ehleringer J. R. (1997) Global vegetation change through the Miocene/Pliocene boundary. *Nature* **389**, 153–158.

Charney J. G., Arakawa A., Baker D. J., Bolin B., Dickinson R. E., Goody R. M., Leith C. E., Stommel H. M. and Wunsch C. L. (1979) *Carbon Dioxide and Climate: a Scientific Assessment*., Washington, D.C.

Cifuentes A. L. A., Sharp J. H., Fogel M. L., Cifuentes L. A. and Fogel L. (1988) Stable carbon and nitrogen in the isotope biogeochemistry Delaware estuary. **33**, 1102–1115.

Clark W. C. ed. (1982) *Carbon dioxide review*. Vol. 303., Clarendon Press, Oxford, UK.

Craig H. (1953) The geochemistry of the stable carbon isotopes. *Geochim. Cosmochim. Acta* **3**, 53–92.

Degens E. T., Behrendt M., Gotthardt B. and Reppmann E. (1968a) Metabolic fractionation of

carbon isotopes in marine plankton-II. Data on samples collected off the coasts of Peru and Ecuador. *Deep. Res.* **15**, 11–20.

Degens E. T., Guillard R. R. L., Sackett W. M. and Hellebust J. A. (1968b) Metabolic fractionation of carbon isotopes in marine plankton—I. Temperature and respiration experiments. *Deep Sea Res. Oceanogr. Abstr.* **15**, 1–9.

Deuser W. G., Degens E. T. and Guillard R. R. . (1968) Carbon isotope relationships between plankton and sea water. *Geochim. Cosmochim. Acta* **32**, 657–660.

Dlugokencky E. and Tans P. P. (2020) Trends in atmospheric carbon dioxide, National Oceanic & Atmospheric Administration, Earth System Research Laboratory. *NOAA/ESRL*. Available at: https://www.esrl.noaa.gov/gmd/ccgg/trends/gl_data.html [Accessed January 9, 2020].

Farquhar G. D., O'Leary M. H. and Berry J. A. (1982) On the Relationship Between Carbon Isotope Discrimination and the Intercellular Carbon Dioxide Concentration in Leaves. *Funct. Plant Biol.* **9**, 121.

Foote E. N. (1856) ART. XXXI.--Circumstances affecting the Heat of the Sun's Rays;: (Read before the American Association, August 23d, 1856.). *Am. J. Sci. Arts* **22**, 382.

Francois R., Altabet M. a., Goericke R., McCorkle D. C., Brunet C. and Poisson A. (1993) Changes in the δ13C of surface water particulate organic matter across the subtropical convergence in the SW Indian Ocean. *Global Biogeochem. Cycles* **7**, 627.

Freeman K. H. and Hayes J. M. (1992) Fractionation of carbon isotopes by phytoplankton and estimates of ancient CO2 levels. *Global Biogeochem. Cycles* **6**, 185–198.

Fry B. and Wainwright S. C. (1991) Diatom sources of 13C-rich carbon in marine food webs. *Mar. Ecol. Prog. Ser.* **76**, 149–157.

Goericke R., Montoya J. P. and Fry B. (1994) Physiology of isotopic fractionation in algae and cyanobacteria. In *Stable isotopes in ecology and environmental science* (eds. K. Lajtha and R. H. Michener). Methods in ecology. Blackwell Scientific Publications, Oxford; Boston. pp. 187–221.

Hansen J., Sato M., Russell G. and Kharecha P. (2013) Climate sensitivity, sea level and atmospheric carbon dioxide. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **371**, 20120294.

Henderiks J. and Pagani M. (2007) Refining ancient carbon dioxide estimates: Significance of coccolithophore cell size for alkenone-based pCO2 records. *Paleoceanography* **22**, 1–12.

Herbert T. D., Lawrence K. T., Tzanova A., Peterson L. C., Caballero-Gill R. and Kelly C. S. (2016) Late Miocene global cooling and the rise of modern ecosystems. *Nat. Geosci.* **9**, 843–847.

Higgins J. A., Kurbatov A. V, Spaulding N. E., Brook E., Introne D. S., Chimiak L. M., Yan Y., Mayewski P. A. and Bender M. L. (2015) Atmospheric composition 1 million years ago from blue ice in the Allan Hills, Antarctica. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 6887–91.

Holtz L. M., Wolf-Gladrow D. and Thoms S. (2017) Stable carbon isotope signals in particulate organic and inorganic carbon of coccolithophores – A numerical model study for Emiliania huxleyi. *J. Theor. Biol.* **420**, 117–127.

Jasper J. P., Hayes J. M., Mix A. C. and Prahl F. G. (1994) Photosynthetic fractionation of 13C and concentrations of dissolved CO2 in the central equatorial Pacific during the last 255,000 years. *Paleoceanography* **9**, 781–798.

Keeling C. (1960) The Concentration and Isotopic Abundances of Carbon Dioxide in the Atmosphere. *Tellus*.

Laws E. A., Popp B. N., Bidigare R. R., Kennicutt M. C. and Macko S. A. (1995) Dependence of phytoplankton carbon isotopic composition on growth rate and [CO2]aq: Theoretical considerations and experimental results. *Geochim. Cosmochim. Acta* **59**, 1131–1138.

Lunt D. J., Haywood A. M., Schmidt G. A., Salzmann U., Valdes P. J. and Dowsett H. J. (2010) Earth system sensitivity inferred from Pliocene modelling and data. *Nat. Geosci.* **3**, 60–64.

Martínez-Botí M. a., Foster G. L., Chalk T. B., Rohling E. J., Sexton P. F., Lunt D. J., Pancost R. D., Badger M. P. S. and Schmidt D. N. (2015) Plio-Pleistocene climate sensitivity evaluated using high-resolution CO2 records. *Nature* **518**, 49–54.

McClelland H. L. O., Bruggeman J., Hermoso M. and Rickaby R. E. M. (2017) The origin of carbon isotope vital effects in coccolith calcite. *Nat. Commun.* **8**, 14511.

Myhre G., Highwood E. J., Shine K. P. and Stordal F. (1998) New estimates of radiative forcing due to well mixed greenhouse gases. *Geophys. Res. Lett.* **25**, 2715–2718.

O'Leary M. H. (1981) Carbon isotope fractionation in plants. *Phytochemistry* **20**, 553–567.

O'Leary M. H. (1984) Measurement of the isotope fractionation associated with diffusion of carbon dioxide in aqueous solution. *J. Phys. Chem.* **88**, 823–825.

Pagani M., Arthur M. A. and Freeman K. H. (1999) Miocene evolution of atmospheric carbon dioxide. *Paleoceanography* **14**, 273–292.

Pagani M., Zachos J. C., Freeman K. H., Tipple B. and Bohaty S. (2005) Marked Decline in Atmospheric Carbon Dioxide Concentrations During the Paleogene. *Science (80-.).* **309**, 600–603.

Petit R. J., Raynaud D., Basile I., Chappellaz J., Ritz C., Delmotte M., Legrand M., Lorius C. and Pe L. (1999) Climate and atmospheric history of the past 420,000 years from the

Vostok ice core, Antarctica. *Nature* **399**, 429–413.

Popp B. N., Laws E. A., Bidigare R. R., Dore J. E., Hanson K. L. and Wakeham S. G. (1998) Effect of phytoplankton cell geometry on carbon isotopic fractionation. *Geochim. Cosmochim. Acta* **62**, 69–77.

Popp B. N., Takigiku R., Hayes J., Louda J. and Baker E. W. (1989) The post-paleozoic chronology and mechanism of 13C depletion in primary marine organic matter.pdf. *Am. J. Sci.* **289**, 436–454.

Previdi M., Liepert B. G., Peteet D., Hansen J., Beerling D. J., Broccoli A. J., Frolking S., Galloway J. N., Heimann M., Le Quéré C., Levitus S. and Ramaswamy V. (2013) Climate sensitivity in the Anthropocene. *Q. J. R. Meteorol. Soc.* **139**, 1121–1131.

Le Quéré C., Andrew R. M., Friedlingstein P., Sitch S., Hauck J., Pongratz J., Pickers P. A., Korsbakken J. I., Peters G. P., Canadell J. G., Arneth A., Arora V. K., Barbero L., Bastos A., Bopp L., Chevallier F., Chini L. P., Ciais P., Doney S. C., Gkritzalis T., Goll D. S., Harris I., Haverd V., Hoffman F. M., Hoppema M., Houghton R. A., Hurtt G., Ilyina T., Jain A. K., Johannessen T., Jones C. D., Kato E., Keeling R. F., Goldewijk K. K., Landschützer P., Lefèvre N., Lienert S., Liu Z., Lombardozzi D., Metzl N., Munro D. R., Nabel J. E. M. S., Nakaoka S., Neill C., Olsen A., Ono T., Patra P., Peregon A., Peters W., Peylin P., Pfeil B., Pierrot D., Poulter B., Rehder G., Resplandy L., Robertson E., Rocher M., Rödenbeck C., Schuster U., Schwinger J., Séférian R., Skjelvan I., Steinhoff T., Sutton A., Tans P. P., Tian H., Tilbrook B., Tubiello F. N., van der Laan-Luijkx I. T., van der Werf G. R., Viovy N., Walker A. P., Wiltshire A. J., Wright R., Zaehle S. and Zheng B. (2018) Global Carbon Budget 2018. *Earth Syst. Sci. Data* **10**, 2141–2194.

Rau G. H., Riebesell U. and Wolf-Gladrow D. (1996) A model of photosynthetic 13C fractionation by marine phytoplankton based on diffusive molecular CO2 uptake. *Mar. Ecol. Prog. Ser.* **133**, 275–285.

Rau G. H., Sweeney R. . and Kaplan I. R. (1982) Plankton 13C:12C ratio changes with latitude: differences between northern and southern oceans. *Deep Sea Res. Part A. Oceanogr. Res. Pap.* **29**, 1035–1039.

Rau G. H., Takahashi T., Desmarais D. J., Repeta D. J. and Martin J. H. (1992) The Relationship between d13C of Organic-Matter and [CO2(aq)] in Ocean Surface-Water: Data from a JGOFS Site in the Northeast Atlantic-Ocean and a Model. *Geochim. Cosmochim. Acta* **56**, 1413–1419.

Rau G. H., Takahashi T. and Des Marais D. J. (1989) Latitudinal variations in plankton δ13C: Implications for CO2 and productivity in past oceans. *Nature* **341**, 516–518.

Riebesell U., Wolf-Gladrow D. A. and Smetacek V. (1993) Carbon dioxide limitation of marine phytoplankton growth rates. *Nature* **361**, 249–251.

Rohling E. J., Sluijs A., Dijkstra H. A., Köhler P., van de Wal R. S. W., von der Heydt A. S., Beerling D. J., Berger A., Bijl P. K., Crucifix M., DeConto R. M., Drijfhout S. S., Fedorov A. V., Foster G. L., Ganopolski A., Hansen J., Hönisch B., Hooghiemstra H., Huber M., Huybers P., Knutti R., Lea D. W., Lourens L. J., Lunt D., Masson-Demotte V., Medina-Elizalde M., Otto-Bliesner B. L., Pagani M., Pälike H., Renssen H., Royer D. L., Siddall M., Valdes P., Zachos J. C. and Zeebe R. E. (2012) Making sense of palaeoclimate sensitivity. *Nature* **491**, 683–691.

Rost B., Riebesell U., Burkhardt S. and Sültemeyer D. (2003) Carbon acquisition of bloom-forming marine phytoplankton. *Limnol. Oceanogr.* **48**, 55–67.

Royer D. L. (2016) Climate Sensitivity in the Geologic Past. *Annu. Rev. Earth Planet. Sci.* **44**, 277–293.

Sackett W. M., Eckelmann W. R., Bender M. L. and Be A. W. H. (1965) Temperature Dependence of Carbon Isotope Composition in Marine Plankton and Sediments. *Science (80-.).* **148**, 235–237.

Sage R. F. (2020) Global change biology: A primer. *Glob. Chang. Biol.* **26**, 3–30.

Sosdian S. M., Greenop R., Hain M. P., Foster G. L., Pearson P. N. and Lear C. H. (2018) Constraining the evolution of Neogene ocean carbonate chemistry using the boron isotope pH proxy. *Earth Planet. Sci. Lett.* **498**, 362–376.

Stoll H. M., Guitian J., Hernandez-Almeida I., Mejia L. M., Phelps S. R., Polissar P. J., Rosenthal Y., Zhang H. and Ziveri P. (2019) Upregulation of phytoplankton carbon concentrating mechanisms during low CO2 glacial periods and implications for the phytoplankton pCO2 proxy. *Quat. Sci. Rev.* **208**, 1–20.

Tyndall J. (1860) VII. Note on the transmission of radiant heat through gaseous bodies. *Proc. R. Soc. London* **10**, 37–39.

Vogel J. C. (1980) Fractionation of the Carbon Isotopes During Photosynthesis. In *Fractionation of the Carbon Isotopes During Photosynthesis* Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 5–29.

van Vuuren D. P., Edmonds J., Kainuma M., Riahi K., Thomson A., Hibbard K., Hurtt G. C., Kram T., Krey V., Lamarque J.-F., Masui T., Meinshausen M., Nakicenovic N., Smith S. J. and Rose S. K. (2011) The representative concentration pathways: an overview. *Clim. Change* **109**, 5–31.

Wilkes E. B. and Pearson A. (2019) A general model for carbon isotopes in red-lineage phytoplankton: Interplay between unidirectional processes and fractionation by RubisCO. *Geochim. Cosmochim. Acta* **265**, 163–181.

Winter A., Henderiks J., Beaufort L., Rickaby R. E. M. and Brown C. W. (2014) Poleward

expansion of the coccolithophore Emiliania huxleyi. *J. Plankton Res.* **36**, 316–325.

Wong W. W. and Sackett W. M. (1978) Fractionation of stable carbon isotopes by marine phytoplankton. *Geochim. Cosmochim. Acta* **42**, 1809–1815.

Yan Y., Bender M. L., Brook E. J., Clifford H. M., Kemeny P. C., Kurbatov A. V., Mackay S., Mayewski P. A., Ng J., Severinghaus J. P. and Higgins J. A. (2019) Two-million-year-old snapshots of atmospheric gases from Antarctic ice. *Nature* **574**, 663–666.

Zachos J. C., Dickens G. R. and Zeebe R. E. (2008) An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* **451**, 279–283.

Zhang Y. G., Pagani M., Liu Z., Bohaty S. M. and DeConto R. (2013) A 40-million-year history of atmospheric CO2. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **371**, 20130096.

Zhang Y. G., Pearson A., Benthien A., Dong L., Huybers P., Liu X. and Pagani M. (2019) Refining the alkenone-pCO2 method I: Lessons from the Quaternary glacial cycles. *Geochim. Cosmochim. Acta* **260**, 177–191.

Chapter 2: Carbon isotope fractionation in the Noelaerhabdaceae and a critical evaluation of the alkenone paleobarometer

2.1 Abstract

The carbon isotope fractionation (ε_p) recorded in algal organic matter, and specifically the long-chain alkenones produced by the coccolithophorid family Noelaerhabdaceae, has long been used as a means for reconstructing past atmospheric CO_2 levels. The canonical proxy method linearly relates ε_p to changes in cellular carbon demand relative to diffusive CO_2 supply, with larger ε_p values occurring at lower values of demand relative to supply (abundant CO_2). While the physiological and environmental controls on algal ε_p have been a topic of study for decades, the response of *Gephyrocapsa oceanica*, one of the two dominant alkenone producers in the modern ocean and of the last few million years, has not been closely studied. Here we subject *G. oceanica* to various CO_2 levels by increasing the partial pressure of CO_2 in the culture headspace, as opposed to increasing dissolved inorganic carbon (DIC) and alkalinity concentrations at a constant pH. We note no substantial change in algal physiology, but observe an increase in ε_p as carbon demand relative to CO_2 supply decreases, consistent with the response in DIC manipulations. We conduct a careful compilation of existing Noelaerhabdaceae ε_p data and show that the ε_p response is better described by a nonlinear function of carbon demand relative to diffusive CO_2 supply. A further meta-analysis of these data shows that the slope of the ε_p response in individual treatments is dependent on the range of carbon demand relative to CO_2 supply in the culture treatment, and a simple linear function relating ε_p to $1/CO_2$ should not be used to estimate paleoatmospheric CO_2 levels. We develop a multiple linear regression model to predict ε_p from key physiological and environmental variables, finding that both photoperiod

duration and light intensity are critical parameters, and that including these variables allows us to model algal ε_p with high accuracy and precision. We recommend that while alkenone ε_p does indeed record a signature of the environmental CO_2, light intensity and photoperiod must be accounted for when interpreting past changes in alkenone ε_p.

2.2 Introduction

Anthropogenic CO_2 emissions are perturbing the global carbon cycle more rapidly than at any time in the last 66 million years (Zeebe et al., 2016). One key component of the natural carbon cycle is the coccolithophorid algae, which modify dissolved inorganic carbon concentrations through photosynthetic carbon fixation and the biogenic production of calcium carbonate plates (Balch et al., 1992). These eukaryotic algae are abundant and widely distributed throughout the global oceans today (Geitzenauer et al., 1976; McIntyre and Bé, 2003; Hagino and Young, 2015), and in Cenozoic sediments (Gartner, 1969; Burky, 1971). Certain coccolithophores in the family Noelaerhabdaceae—such as *Emiliania huxleyi* and *Gephyrocapsa oceanica* (Young et al., 2003)—are well known for the unique long-chain unsaturated ketone molecules they produce (Marlowe et al., 1984). These lipid biomarkers, called alkenones, have been found in sediments as old as 120 Ma (Brassell et al., 2004) and can be easily isolated from sediments and other co-occurring organic compounds. Because alkenones have a specific algal source, the carbon isotope ratio of these molecules ($\delta^{13}C_{Mk37:2}$) has been employed as a paleobarometer for estimating past atmospheric CO_2 concentrations (Jasper and Hayes, 1990; Jasper et al., 1994; Pagani et al., 1999).

The carbon stable isotope discrimination recorded in primary photosynthetic matter has been used for decades to reconstruct ancient atmospheric CO_2 levels (Jasper and Hayes, 1990;

Freeman and Hayes, 1992). The carboxylating enzyme Ribulose 1, 5-biphosphate carboxylase/oxygenase (RubisCO) in plants and algae prefers ^{12}C to ^{13}C by ~11-30‰, depending on the specific form of RubisCO employed by the organism (Goericke et al., 1994; Falkowski and Raven, 2007; Boller et al., 2011; Boller et al., 2015). This kinetic preference makes photosynthetic particulate organic carbon ($\delta^{13}C_{POC}$) isotopically light relative to the ambient aqueous CO_2 ($\delta^{13}C_{CO2aq}$). The carbon isotope fractionation is termed ε_p and is calculated using the following equation:

$$\varepsilon_p = 1000 \times [(1000 + \delta^{13}C_{CO2aq}) / (1000 + \delta^{13}C_{POC}) - 1] \quad \text{[Eq. 1]}$$

The magnitude of carbon isotope fractionation (ε_p) has been found to be inversely related to the ambient $CO_{2(aq)}$ concentration: as $[CO_{2(aq)}]$ declines, so too does the diffusive influx of CO_2 and the amount of CO_2 inside the cell, reducing the apparent carbon isotope fractionation of organic matter with respect to the external CO_2 source (Rau et al., 1989; Rau et al., 1992). In addition to $[CO_{2(aq)}]$, controlled culture experiments have identified other environmental and physiological factors that influence algal $\delta^{13}C$, including algal growth rate, species, and growth conditions, such as chemostat or batch culturing, or irradiance or nutrient limitation (Thompson and Calvert, 1995; Laws et al., 1995; Burkhardt et al., 1999; Rost et al., 2002). However, the fundamental assumptions on which the alkenone-CO_2 proxy are built have recently been called into question (McClelland et al., 2017; Badger et al., 2019; Stoll et al., 2019).

The original framework for interpreting alkenone $\delta^{13}C$ assumes that the algae producing these molecules only acquire CO_2 through diffusion, and most alkenone ε_p records published to date have been analyzed this way (Pagani et al., 1999; Pagani et al., 2011; Zhang et al., 2013; Zhang et al., 2017; Badger et al., 2019; Zhang et al., 2019). However, more recent work has shown these algae employ various carbon acquisition mechanisms (such as HCO_3^- uptake and

conversion to CO_2) to supplement the diffusive supply of CO_2 to RubisCO, violating this key assumption (Rost et al., 2006; Bach et al., 2013; Bolton and Stoll, 2013; Isensee et al., 2014; Kottmeier et al., 2014; Stoll et al., 2019). Quantitatively understanding the factors other than [$CO_{2(aq)}$] that affect carbon isotope fractionation, as well as how carbon acquisition strategies vary with growth conditions, can help improve the accuracy of algal paleobarometry.

Here we evaluate the alkenone paleobarometer using new and published culture experiments of alkenone-producing coccolithophores. We conduct careful standardization of several published datasets in our compilation to robustly examine the relationship between carbon isotope fractionation in these algae and the environmental and physiological factors that have been proposed to influence algal $\delta^{13}C$. We begin with a background that highlights two structural issues with the diffusive model.

2.3 Background

2.3.1 Carbon isotope fractionation by RubisCO and the basis for paleobarometry

The carbon isotope ratio of primary organic matter has the potential to record information about algal physiology and the ambient environment because of the fractionation by RubisCO (ε_f). RubisCO's fractionation of ^{12}C to ^{13}C is ~11-30‰, depending on the form of RubisCO expressed (Roeske and O'Leary, 1985; Boller et al., 2011; Boller et al., 2015). Types of RubisCO fall into four different families: forms I, II, III, and IV. While similar in function, they have different structures and varying catalytic and kinetic parameters (Badger et al., 1998; Tcherkez et al., 2006; Tabita et al., 2007). Form I is found in higher plants and nearly all eukaryotic algae, and is further subdivided into forms IA through ID. However, most eukaryotic

algae—including alkenone-producing coccolithophores—express Form ID, while most higher plants and green algae possess Form IB.

Because the kinetic fractionation of Form ID RubisCO from a coccolithophore or diatom had not been measured until this decade, paleobarometry studies using algal ε_p generally adopted an ε_f of 25‰ as the maximum possible fractionation. The 25‰ value is based on canonical ε_f measurements from higher plants, which yielded fractionation factors of ~20-29‰ (O'Leary, 1984; Roeske and O'Leary, 1985; Goericke et al., 1994). The applicability of this value to Form ID was supported by nitrate-limited chemostat cultures of several eukaryotic algae where regressions of ε_p against carbon demand relative to supply had convergent intercepts (i.e. infinite CO_2 supply and maximum fractionation) of ~25‰ (Bidigare et al., 1997; Popp et al., 1998b).

However, recent *in vitro* measurements of Form ID RubisCO extracted from the coccolithophore *E. huxleyi* and diatom *Skeletonema costatum* yielded ε_f = 11.1‰ and 18.5‰, respectively (Boller et al., 2011; Boller et al., 2015), substantially lower than the commonly applied ε_f = 25‰ of Form IB. Taken at face value, these smaller RubisCO fractionations suggest other mechanisms must be responsible for sourcing isotopically light carbon to the site of carboxylation and elevating ε_p above the maximum limit of fractionation by RubisCO, ~11‰ in *E. huxleyi*. Culture experiments with the dinoflagellate *Alexandrium tamarense* demonstrate *in vivo* ε_f values of ~27‰ (Wilkes et al., 2017), similar to the chemostat results of Bidigare et al. (1997), yet this organism expresses Form II RubisCO, with an estimated *in vitro* ε_f range of 18 to 23‰ (Roeske and O'Leary, 1985). Recent modeling work by Wilkes and Pearson (2019) suggests that the difference between *in vivo* and *in vitro* ε_f estimates may result from the activity of a photocatalytic carbon acquisition pathway. They propose a mechanism by which CO_2 is hydroxylated to HCO_3^- during transport across the thylakoid membrane under high photon flux, a

process that imparts a significant kinetic ^{13}C depletion. In this work we examine the relationship between ε_p and ε_f empirically through the diffusive model and new multilinear regression models.

2.3.2 Non-diffusive CO_2 supply

The models of carbon isotope fractionation described up to this point assume that all carbon used for photosynthesis and growth is derived from the passive diffusion of CO_2 to the site of carboxylation. While many have argued against this (Keller and Morel, 1999; Laws et al., 2002; Bolton and Stoll, 2013; Holtz et al., 2015a; Holtz et al., 2017; McClelland et al., 2017), it remains central to the theory behind the alkenone paleobarometer. At least three factors render the assumption of solely diffusive supply unrealistic. First, photorespiration at low CO_2:O_2 ratios reduces the efficiency of RubisCO. RubisCO is a carboxylase-oxygenase enzyme and will oxidize organic carbon through photorespiration when the CO_2:O_2 ratio is low (Andersson, 2008). CO_2 and O_2 are similar in structure and polarity, making it difficult for the active sites of RubisCO to distinguish between these two molecules (Tcherkez et al., 2006). Photorespiration can be counteracted by elevating CO_2 concentrations in the vicinity of RubisCO (Badger et al., 1998). Second, aqueous CO_2 levels in many regions of the ocean are suboptimal for RubisCO efficiency. In the modern ocean, $[CO_{2(aq)}]$ in the surface ocean ranges from ~10 to ~25 μmol kg^{-1}, with most of the tropics <15 μmol kg^{-1} (Olsen et al., 2016). In RubisCO isolated from marine eukaryotes, the CO_2 half-saturation constants—the concentration at which half the maximum catalytic rate is achieved—are between 10 and 50 μM CO_2 (Young et al., 2016; Heureux et al., 2017); natural CO_2 concentrations in many oceanographic settings are therefore suboptimal for RubisCO to act as a carboxylase. Third, distinct from the CO_2 molecule that is fixed into sugar,

RubisCO requires an additional CO_2 molecule to carbamylate and activate the RubisCO enzyme, priming it to accept a CO_2 at the active site for fixation (Lorimer and Miziorko, 1980; Andersson, 2008). Low CO_2 supply from diffusion can lead to inefficient activation of the RubisCO enzyme and suboptimal cellular fixation rates. These factors indicate that at modern aqueous (and atmospheric) CO_2 levels, diffusive supply to RubisCO is suboptimal and, in some cases, insufficient for fixation.

The requirements of RubisCO and its slow catalysis have promoted the development of carbon concentrating mechanisms (CCMs) in many algae (Raven, 1997; Badger et al., 1998). These CCMs—any chemical or biophysical process that enhances CO_2 concentrations at the site of carboxylation—include bicarbonate pumps and enzyme-facilitated conversion to CO_2 that take advantage of the abundant ambient DIC (Raven and Beardall, 2003; Reinfelder, 2011); the localization of RubisCO in organelles with low permeability that prevent the diffusive loss of CO_2 (Badger et al., 1998; Hopkinson et al., 2011); or the use of intracellular pH gradients to modify the equilibrium concentrations of CO_2 inside the cell (Bach et al., 2013; Holtz et al., 2015a). CCMs are fundamental components of the algal photosynthetic machinery, and are widespread across algal taxa (Giordano et al., 2005). While specific CCMs in coccolithophores are still being uncovered, there is experimental evidence for CO_2-dependent CCM activity in this algal group (Rost et al., 2006; Bach et al., 2013; Isensee et al., 2014; Kottmeier et al., 2014). Such activity violates the assumptions of the diffusive model and will change the relationships between ε_p and τ predicted from this model (Laws et al., 2002; Freeman and Pagani, 2005). One goal of this study is to clarify the presence of any CCM activity through its effects on the τ-ε_p relationship.

2.3.3 Summary and present work

Here we examine carbon isotope fractionation in new and existing culture experiments of alkenone-producing algae to understand the controls on fractionation and its relationship to aqueous CO_2 concentrations. We evaluate the results within the diffusive CO_2 supply model of Rau et al. (1996), but also in light of more recent studies suggesting that CCMs are active in coccolithophore algae (Badger et al., 1998; Bach et al., 2013; Bolton and Stoll, 2013; Heureux et al., 2017; McClelland et al., 2017; Stoll et al., 2019), and that fractionation by form ID RubisCO may be as low as 11‰. Taking an approach similar to Stoll et al. (2019), we present multiple linear regression models to cast ε_p as a function of the variables that past studies and this analysis identify as important predictors of ε_p.

2.4 Methods

2.4.1 New *G. oceanica* cultures

2.4.1.1 Culture set-up

We grew non-axenic batch cultures of *Gephyrocapsa oceanica* (Roscoff Culture Collection RCC1303; isolated from the Bay of Biscay in 1999) in triplicate at five different CO_2 partial pressures, with target pCO_2 of 200, 400, 600, 800, and 1000 µatm. Cultures were conducted simultaneously, and the stock strain was pre-acclimated to each CO_2 level for ~5 generations prior to the experimental treatment. CO_2 concentrations were set in 2.5-L polycarbonate culture vessels (filled with one liter of media) by continuously aerating the headspace of the bottle with gas from tanks with ~79% N_2, 21% O_2, and either 200, 400, 600, 800, or 1000 ppmv CO_2 (acquired from TechAir, NY, USA), following the experimental design of Hennon et al. (2019). Culture media were not bubbled because this has been shown to have

adverse effects on the physiology of algal cultures (Juhl and Latz, 2002). Instead, gas streams of pre-mixed CO_2 partial pressures were directed through sterilized glass pipets and positioned just above the media surface to break the boundary layer, thereby the $[CO_{2(aq)}]$ of each vessel was set by diffusion from the headspace. The gas stream tubing was held in place above the headspace-media interface by wrapping the tubing in cheesecloth and wedging the cheesecloth between the tubing and the neck of the polycarbonate vessel. Cells were grown in 0.2µm-filtered natural seawater collected from Avery Point, CT, USA. Nutrients were based on the L1 recipe (Guillard and Hargraves, 1993), but reduced as follows: L1/15 phosphate, L1/10 nitrate, and L1/25 vitamins and trace metals. The salinity was 32.53, and cultures were incubated at 18°C. Light intensity was measured using a QSPL-2100 light meter (Biospherical Instruments Inc., San Diego, CA, USA), and was delivered in a 14:10 photoperiod (light on from 0800 to 2200 hours). Light intensity in the culture chamber varied slightly depending on flask location, but averaged 99 ± 14 µmol photons m^{-2} s^{-1} (mean and one standard deviation of all 15 replicates). We calculate total daily irradiance as the product of the light intensity (µmol m^{-2} s^{-1}) and the number of seconds of illumination, and convert the units mol m^{-2} d^{-1}.

2.4.1.2 Cell physiology

Cells grew exponentially for five days and were harvested in the late morning to early afternoon on day five. This sampling duration was chosen to ensure sufficient material for geochemical and physiological analyses, while preventing the cells from encountering nutrient limitation and entering stationary phase. Cell density in each culture vessel was monitored daily using relative fluorescence, measured with an AquaFlash fluorometer (Turner Designs, San Jose, CA, USA), as a proxy for cell concentration. 1-mL samples were also collected daily, preserved

in paraformaldehyde (1.5%), flash frozen in liquid nitrogen, and stored at -80C. Preserved samples were subsequently thawed and measured by flow cytometry on a Guava EasyCyte Mini cytometer (Millipore, Darmstadt, Germany) within 6 months. Cell counts from the day of harvest were determined in technical triplicate by microscopy on a hemocytometer. Cellular growth rates (μ, day^{-1}) were calculated from the slope of the relationship between ln(cell concentration) and time in days after inoculation.

Relative changes in cell volume between treatments and replicates were constrained using forward scatter estimates from the Guava flow cytometer. Using 1 µm-diameter reference beads (Fluoresbrite, YG) to calibrate the flow cytometer response, we compared the forward scatter from aliquots of each culture to the reference beads and across the range of treatments to examine relative differences. We calculated a response ratio by dividing the natural logarithm of the forward scatter of each sample aliquot to the forward scatter of the reference beads. To convert this relative response ratio to a cell volume, we compared all treatments to a "baseline" value. We chose the 400-µatm treatment because it is closest to natural conditions for these cells (mean response factor = 0.75; n = 3). We then prescribed a mean cell diameter of 5.2 µm for the 400-µatm treatment group—equivalent to a mean cell volume of 73.6 μm^3—and calculated changes in cell volume relative to this baseline. This prescribed cell diameter is based on eight micrographs of our culture stock of this strain as well as measurements of cell diameter of this strain from other batch cultures (Aloisi, 2015; Faucher et al., 2017). We estimated the mean cell volume from each flow cytometry measurement by dividing the response factor (forward scatter of sample relative to standard bead) by the "baseline" response factor, then multiplying by the prescribed "baseline" cell volume. We used the standard deviation of the 400-µatm response factor (1 s.d. = 0.07, n = 3) as the uncertainty on our estimate, and propagated this into a cell

diameter uncertainty of ± 0.15 µm. Cell size estimates were corrected for differences in time of sampling between treatments using the approach of Aloisi (2015) (see Eq. 5). The cell size at time of sampling and the time-normalized sizes can be found in Appendix A, Table S1.

2.4.1.3 Carbonate chemistry

Carbonate chemistry was monitored using measurements of pH and total alkalinity. The pH of the media in each culture vessel was measured daily using the m-cresol purple method (Dickson et al., 2007) using a Shimadzu UV-1800 spectrophotometer (Kyoto, Japan), calibrated by pH certified reference materials from the Dickson Lab, (Scripps Institution of Oceanography, La Jolla, CA, USA). Total alkalinity was assayed by titration at the start and end of the experiment using a closed cell titration with a Metrohm Titrando autotitrator (Herisau, Switzerland), calibrated by certified reference materials from the Dickson Lab (Scripps Institution of Oceanography, La Jolla, CA, USA). Alkalinity and pH measurements were made at standard temperature (25°C), and the speciation of the carbonate system was calculated at 18°C using CO2Sys_v2.1 (Lewis and Wallace, 1998). pH is reported on the total scale; we used K_1 and K_2 CO_2 constants of Lueker et al. (2000) with KSO_4 from Dickson, and total boron from Uppström (1974).

12-mL samples for dissolved inorganic carbon concentrations and $\delta^{13}C_{DIC}$ were taken from each culture vessel at the beginning (t_0) and end (t_f) of the experiments. Samples were sterile filtered through a 0.2 µm polycarbonate filter in a laminar flow hood to remove bacterial contamination from the culture vessels, sealed without headspace in glass vials with septa (model E2852, EA Consumables, Pennsauken, NJ, USA), and shipped to the Stable Isotope Facility at UC Santa Cruz (California, USA) for analysis within 6 months. Each t_0 and t_f sample was

analyzed 1 to 8 times. The average precision (average one standard deviation of all replicates) was 100 μmol kg^{-1} for DIC and 0.21‰ for $\delta^{13}C_{DIC}$. We calculated $\delta^{13}C_{CO2aq}$ using Equation 2 of Rau et al. (1996). Because cells grew exponentially, the vast majority of the total biomass—and its average carbon isotope ratio—was generated during the last two cell divisions, which is much closer to the t_f measurement than t_0 measurement. We therefore used the t_f measurements of carbonate chemistry and $\delta^{13}C_{DIC}$ rather than t_0 or an average between t_f and t_0.

2.4.1.4 Lipid extraction and stable carbon isotope measurements

Approximately 150 mL of culture medium from each treatment was gently vacuum-filtered onto pre-combusted glass fiber filters. Free lipids from these filters were extracted via Dionex ASE 350 using a dichloromethane and methanol mixture (9:1 v/v, respectively) at 100°C and 10.3 kPa (1500 psi). The total lipid extract (TLE) was spiked with a general recovery standard containing stearyl stearate, then evaporated to dryness under a stream of purified N_2 gas. Carbon isotope ratios of alkenone compounds were measured using a Thermo TraceGC coupled to an Isolink device and a Thermo DeltaV mass spectrometer. Samples were diluted in toluene and injected into a programmable temperature vaporizing inlet at 90°C, held isothermal for 1.5 minutes, then heated to 325°C with a splitless time of 5 minutes, followed by a cleaning phase with a 50:1 split ratio at 350°C for 5 minutes. The GC oven temperature started at 90°C, was held isothermal for 1.5 minutes, ramped 25°C/min to 250°C, then ramped 1°C/min to 313°C, followed by a ramp of 10°C/min to 320°C, and held isothermal for 20 minutes. Water was removed from the analyte stream using an in-line Nafion membrane. Carbon isotope ratios were translated from the laboratory reference gas scale to the VPDB scale using standard *n*-alkane mixtures of known isotope composition (Mix A5 and Mix B4; produced by A. Schimmelman,

Indiana University), and the uncertainty in our isotope measurements was calculated using the MATLAB routine of Polissar and D'Andrea (2014).

Because the cultures were contaminated with cotton cheesecloth fibers during sampling, we could not reliably measure bulk particulate carbon or its stable isotope composition. Instead, we estimated cellular $\delta^{13}C_{POC}$ using the alkenone carbon isotope measurements ($\delta^{13}C_{Mk37:2}$). We pooled the paired $\delta^{13}C_{POC}$ and $\delta^{13}C_{Mk37:2}$ measurements of Popp et al. (1998a), Riebesell et al. (2000) and Wilkes et al. (2018), and, following the methods of Popp et al. (1998a) after Laws (1997), calculate the weighted-mean biomass-lipid isotopic offset ($\varepsilon_{bio/alk}$). The weighting function uses the uncertainties of the $\delta^{13}C_{POC}$ and $\delta^{13}C_{Mk37:2}$ measurements and gives greater weight to values with smaller uncertainties. The standard deviation of this estimated $\varepsilon_{bio/alk}$ is calculated from the inverse square-root of the sum of the inverse of the variances (see Popp et al. (1998a), their Table 2). Our new calculations with all available data yielded an $\varepsilon_{bio/alk}$ value of 4.44‰ ± 0.05‰, which we applied to our samples to calculate ε_p. This approach is directly analogous to sediment-based paleobarometry studies using alkenone ε_p, and by accounting for this biosynthetic fractionating step, our new measurements can be compared to the existing culture data measuring $\delta^{13}C_{POC}$.

2.4.2 Statistical analyses

All calculations and plotting were carried out in MATLAB R2019a. We conducted simple multivariate linear regressions using the *fitlm* function. The two-parameter exponential fit was generated using the *fit* function in the Curve Fitting toolbox with *fittype* set to '*exp2*'. Confidence intervals of coefficients of linear regressions were calculated using the *coefCI* function.

2.4.3 Culture compilation and standardization

2.4.3.1 Dataset requirements and calculating carbon demand and supply

To evaluate the dependence of $\delta^{13}C_{POC}$ on cellular carbon demand and diffusive CO_2 supply, constraints on the following parameters are necessary: cell size (surface area, volume), particulate organic carbon per cell, cellular growth rate, aqueous CO_2 concentration, the $\delta^{13}C_{DIC}$ of the culture media, and the $\delta^{13}C$ of organic matter ($\delta^{13}C_{POC}$) or of alkenones ($\delta^{13}C_{Mk37:2}$). We examined carbon isotope fractionation with respect to carbon demand relative to carbon supply using the non-dimensional variable τ introduced by McClelland et al. (2017) (Eq. 2). We slightly modified their formulation to explicitly include POC $cell^{-1}$ and cellular surface area instead of cellular carbon density, in keeping with the traditional definitions of carbon demand and diffusive CO_2 supply, though our expressions are mathematically equivalent for spherical cells. We recast τ as:

$$\tau = \frac{POC \times \mu_i}{SA \times P_C \times C_e} \quad \text{[Eq. 2]}$$

where POC is the organic carbon content of the cell (mol C); μ_i is the instantaneous growth rate of the cell ($seconds^{-1}$); SA is the surface area over which CO_2 diffuses into the cell (m^2), P_C is the permeability of this cell membrane to diffusion of aqueous CO_2 ($m\ s^{-1}$; here we use $2e^{-5}$), and C_e is the aqueous CO_2 concentration ($mol\ m^{-3}$; which requires an estimate or assumption about the density of the culture media, here assumed to be 1023 $kg\ m^{-3}$).

Instantaneous growth rate (μ_i) is calculated by:

$$\mu_i = \frac{\mu \times (L + D)}{L - D \times r} \quad \text{[Eq. 3]}$$

where μ is the specific growth rate (d^{-1}), L and D are the lengths of the light period and the dark period (hours), and r is a coefficient for the effect of respiratory carbon loss in the dark. Here we

take *r* to be 0.15, following Laws and Bannister (1980) and previous studies (e.g. Riebesell et al., 2000; Rost et al., 2002).

2.4.3.2 Data compilation

We compiled data from nine additional studies (Bidigare et al., 1997; Popp et al., 1998b; Popp et al., 1998a; Riebesell et al., 2000; Rost et al., 2002; Rickaby et al., 2010; Moolna and Rickaby, 2012; McClelland et al., 2017; Wilkes et al., 2018). Although the vast majority of the coccolithophores examined here are calcifying strains, we restricted our analysis to the carbon isotope fractionation in organic matter as it pertains to the alkenone paleobarometer; analysis of the stable isotope fractionation in calcite will be discussed in a future manuscript and while those data may complement our results, they do not affect our conclusions as presented. Below we detail our approach, including the requirements for inclusion in the compiled dataset.

We included only studies where all components of τ—CO_2, POC or cell size, and growth rate—were measured or could be estimated with confidence. We did not include the *G. oceanica* culture data of Rickaby et al. (2010); recent batch culture experiments of *G. oceanica* under similar culture conditions (McClelland et al., 2017) yielded dramatically different $\delta^{13}C_{POC}$ results, and McClelland et al. (2017) also excluded the *G. oceanica* data of Rickaby et al. (2010) in their recent analysis and interpretation. We also did not include the batch culture experiments of Moolna and Rickaby (2012) because we could not reconcile the reported "end" carbonate system data with the reported $[CO_{2(aq)}]$; calculating $[CO_{2(aq)}]$ using the reported DIC and alkalinity resulted in estimates an order of magnitude higher than the reported $[CO_{2(aq)}]$. Furthermore, in our statistical analysis, we did not include the maximum ε_p measurement of Bidigare et al. (1997) (24.9‰) because it was achieved at extremely high $[CO_{2(aq)}]$ (274 μM) and

exerted exorbitant leverage on the regression coefficients. It is an important observation and demonstrates that ε_p ~25‰ can indeed be achieved in culture, but this CO_2 treatment is >9 standard deviations away from the mean [$CO_{2(aq)}$] of the dataset, and is unreasonably high for the natural habitat of coccolithophores. Because one of the goals of this study was to develop quantitative relationships that could be applied to marine settings, we removed this outlying CO_2 treatment so a single data point did not have overwhelming influence on the regression coefficients.

In this dataset, ε_p ranges from ~7 to ~23‰, consistent with $\varepsilon_{p37:2}$ values observed across the Cenozoic (Pagani, 2014). Calculated τ values span ~2 orders of magnitude (~0.08 to ~6 units; absolute values dependent on the choice of membrane permeability because the membrane permeability is a linear multiplier on the denominator of τ; here we calculate τ using a value of $2e^{-5}$ m s^{-1} for all culture samples).

2.4.3.3 Approach to constraining τ

Few datasets report enough information to calculate τ directly. However, this is also the case in sediment studies because input variables, such as POC and cell counts of the alkenone-producing population to calculate POC $cell^{-1}$, are not quantitatively preserved. Where the components of τ are not fully reported, we approximate certain values using relationships between cell size and POC. We justify and explain our rationale below.

2.4.3.3.1 Particulate organic carbon and cell size

There is a robust relationship across a range of species between algal cell volume and particulate organic carbon (Verity et al., 1992; Montagnes et al., 1993; Popp et al., 1998b). These

studies span a variety of algal species and correspondingly many orders of magnitude of algal biovolume, and found that the relationship between POC $cell^{-1}$ and cell volume followed a power law with the exponent very close to one. More targeted work on coccolithophores (Aloisi, 2015), and one study of alkenone-producing algae specifically (McClelland et al., 2017), found a constant organic carbon density of ~15 and 18.5 fmol C per μm^3, respectively. We adopt a value of 17.5 fmol C μm^{-3} ± 0.75 fmol C μm^{-3} to incorporate this range of measured values. Volume and POC are therefore related by:

$$POC = 17.5 \pm 0.75 \text{ fmol C } \mu m^{-3} \times V \ \mu m^3 \quad \text{[Eq. 4]}$$

where POC is the particulate organic carbon content in femtomole C, and V is the cell volume. The cell geometry of alkenone-producing coccolithophores (the family Noelaerhabdaceae) is very nearly spherical, and most studies to date have elected to assume a spherical geometry for these organisms (Popp et al., 1998b; Henderiks and Pagani, 2007; Henderiks, 2008; McClelland et al., 2016; Bolton et al., 2016; Stoll et al., 2019). In treatments where POC $cell^{-1}$ is not reported, we estimate POC using the constant proportionality between POC and biovolume described above (Eq. 4). Where cell radius is not reported but POC $cell^{-1}$ is reported, we estimate cell volume from the same transfer function and calculate the surface area assuming a spherical cellular geometry.

2.4.3.3.2 Cell growth during a light:dark cycle

As emphasized by Aloisi (2015), the time at which samples were collected is an important consideration when comparing results across experiments and treatments because cell growth and division is synchronized with the light:dark cycle (Müller et al., 2008). Under continuous light there is no synchronization of cell division, so sampling the culture media at any

point does not introduce a bias, and draws from the probability distribution of the different growth phases of these unicellular algae. However, cells grown under a discontinuous photoperiod add biomass and increase in size throughout the light period. The time at which the sample is collected can therefore bias the estimate of cell diameter or POC cell^{-1}. A sample taken in the morning would yield lower POC cell^{-1} and smaller cell diameters than samples harvested at the end of the photoperiod, as cells have been fixing carbon and expanding in volume during the illumination phase. To account for differences in sampling time, we use the method outlined in Appendix A1 of Aloisi (2015), which normalizes POC cell^{-1} to a given time in the photoperiod. For our analysis, we normalized to the middle of the photoperiod. POC at a given sample time under a light:dark cycle is calculated as:

$$\mathrm{POC(t)} = \frac{\mathrm{L} \times \mathrm{POC(S_T)}}{\mathrm{L+S_T}} \times \left(1 + \frac{\mathrm{t}}{\mathrm{L}}\right) \qquad \text{[Eq. 5]}$$

where POC is the particulate organic carbon (pg C), *t* is the normalization time in hours after the start of the light period, *L* is the length of the light period in hours, and S_T is the sample collection time in hours after the start of the light period.

2.4.3.3.3 Membrane permeability

In order to evaluate whether the diffusive model of carbon isotope fractionation is supported by the Noelaerhabdaceae culture data, we require an estimate of the cell membrane permeability to CO_2 diffusion (P_C). Cell membrane permeability to CO_2 has not yet been reported for the Noelaerhabdaceae. In their model, Rau et al. (1996) use a value of 1×10^{-4} m s^{-1}. In culture experiments of four diatom strains, Hopkinson et al. (2011) calculated cell membrane

CO_2 permeabilities using membrane inlet mass spectrometry and the cellular uptake of isotopically-labeled DIC, yielding a range from 1.5×10^{-4} to 5.6×10^{-4} m s^{-1}. Membrane permeability also appears in the denominator of the τ calculation. Our primary goal is to compare the Rau et al. (1996) model to the experimental data in the framework of τ, the cellular carbon demand relative to diffusive CO_2 supply. We are guided by a broad visual fit between the diffusive model and the experimental data on a plot of ε_p vs. τ, and choose a permeability of 2×10^{-5} m s^{-1}. This value allows us to show the Rau et al. (1996) model in relation to the data, but it is not critical for our overall analysis. Because the permeability is a constant in the denominator of our τ calculation, ignoring it would not change the relative trends in our results. Better understanding the CO_2 membrane permeability in coccolithophores and whether it is a fixed or dynamic parameter will be useful for determining the sensitivity of these organisms to CO_2 supply in the future.

2.5 Results

2.5.1 New *G. oceanica* cultures

These batch culture experiments targeted 200, 400, 600, 800, and 1000 μatm pCO_2. All treatments were conducted in triplicate (e.g. 200A, 200B, 200C). We treated each individual replicate as a unique experiment and conducted all measurements on seawater and particulate matter collected from each vessel. Data are reported in Appendix A, Table S2.

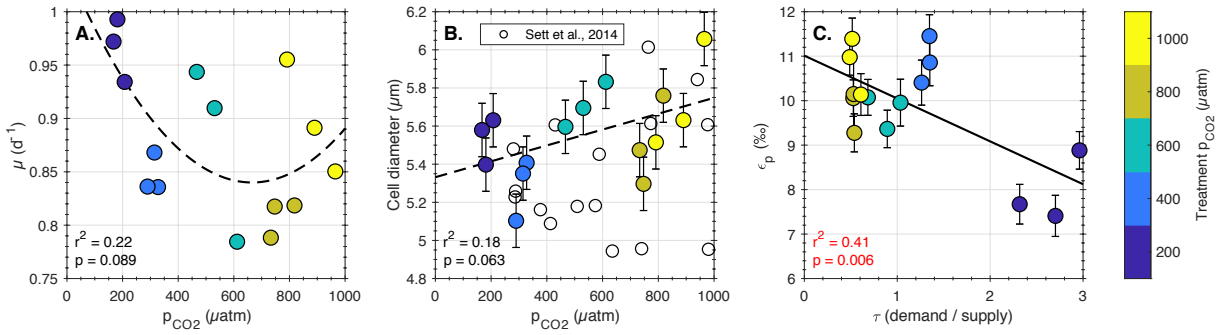

Figure 1. Physiological and geochemical results from *Gephyrocapsa oceanica* (RCC1303) batch cultures. (A) Growth rates with respect to media pCO_2 in new batch cultures; dashed black line is a quadratic fit of μ to pCO_2 from all discrete batch cultures. (B) Estimated cell radius from flow cytometry as a function of pCO_2; dashed line shows linear fit to new culture data. Shown for reference are the results of Sett et al. (2014) in batch cultures of this same strain at 20°C. Cell diameters for those cultures are estimated using measured POC cell^{-1} and assuming a constant POC density of 17.5 fmol C cell^{-1}. (C) Carbon isotope fractionation (ε_p) in new *G. oceanica* cultures determined from alkenone $\delta^{13}C$ as a function of carbon demand relative to carbon supply. ε_p is calculated as the isotope discrimination between algal biomass and $\delta^{13}C_{CO2aq}$ and using a 4.44‰ offset between $\delta^{13}C_{Mk37:2}$ and algal biomass ($\delta^{13}C_{POC}$). Error bars on ε_p are one standard deviation (‰).

2.5.1.1 Cell growth

Measured cellular growth rates range from 0.78 to 0.99 d^{-1}. Instantaneous (photoperiod-normalized via Eq. 3) growth rates range from 1.51 d^{-1} to 1.91 d^{-1}, demonstrating a low-amplitude growth rate response to our experimental range of pCO_2 under the nutrient-replete, constant light, and constant temperature growth conditions. Maximum growth rates were found at minimum pCO_2 (201±3 µatm, 0.97±0.03 d^{-1}; mean ± one standard deviation of three replicates). Our observed response of growth rate to pCO_2 variation is moderately well fit by a second-order polynomial (Fig. 1A; $r^2 = 0.29$, p = 0.051, n = 15), with average minimum growth rates in the 800-µatm treatment and a slight rebound in growth rate at the 1000-µatm treatment. In batch cultures of this same strain (RCC1303) at 20°C, Sett et al. (2014) found roughly similar growth rates over a similar pCO_2 range (0.79 to 1.06 d^{-1} from 280 to 1000 µatm). The observed growth rates are also broadly consistent with a recent compilation by Krumhardt et al. (2017).

Their compilation generally shows higher growth rates at ~200 µatm compared to CO_2 levels above preindustrial, though there is considerable scatter. *G. oceanica* growth rates appear to be most sensitive to pCO_2 at very low partial pressures (< 100 µatm), where growth rates drop below 0.5 d^{-1} (Krumhardt et al., 2017). We visualized cells at the time of harvest to confirm they calcified during these experiments. Final cell concentrations ranged between 1.89 x 10^4 and 1.24 x 10^5 cells mL^{-1}, with an average of 6.39 x 10^4 cells mL^{-1}.

2.5.1.2 Carbonate chemistry

The realized pCO_2 was generally slightly lower than the target pCO_2 of each treatment due to biological uptake, resulting in a pCO_2 range of 168 to 965 µatm. At 18°C and salinity of 32.53, this results in a range of $[CO_{2(aq)}]$ from 5.8 to 33.4 µmol kg^{-1} (we assume ±10% uncertainty). Total alkalinity was measured at t_0 and t_f, while pH measurements were made daily. The mean t_f alkalinity was 2131 µmol kg^{-1} with a standard deviation across all treatments of 47 µmol kg^{-1}, and a mean percent change from t_0 to t_f of 1.75%. The mean pH drift from t_0 to t_f across all treatments was 0.05 units. The carbonate system was therefore relatively chemically stable throughout the majority of the experiments. The largest drift in pH occurred in the 1000C treatment (0.1 units), and the largest drift in alkalinity was found in the 400C treatment (85.5 µmol kg^{-1}). The pH and total alkalinity drift is shown in Fig. S8, and the initial and final values are reported in Table S3.

Because photosynthesis preferentially removes ^{12}C, in a closed system, $\delta^{13}C_{DIC}$ would become more positive as cell growth and the culture experiment progresses. However, because our system was open to the atmosphere and continuously purged by isotopically light CO_2, the $\delta^{13}C_{DIC}$ became more negative from t_0 to t_f. In some cases (e.g. the 400-µatm experiments), the

drift is quite large, with a maximum of -9.1‰. However, the vast majority (81-89%) of the total algal biomass in each treatment is synthesized during the last two cell divisions, with a mean of 62% generated during the final division. The biomass evolution of each treatment is shown in Fig. S7. Because we do not have a full time series of carbonate chemistry or $\delta^{13}C_{DIC}$, we use the t_f measurements as the benchmark against which to calculate carbon isotope fractionation (ε_p) in our cultures.

2.5.1.3 Cell size

Estimated cell diameters at the time of sampling ranged from 5.0 to 5.7 μm. When normalized to a common time (here the middle of the photoperiod, Eq. 5) to account for cell growth and expansion throughout the day, cell diameters range from 5.1 μm to 6.1 μm (n = 15), a relatively narrow span. There was no significant relationship between cell size and pCO_2 across treatments ($r^2 = 0.14$, $p > 0.1$), although the largest cells were found in the highest pCO_2 treatment (~1000 μatm). Our mean cell-size estimates from flow cytometry agree well with POC-estimated cell size from the batch cultures of Sett et al., (2014), who grew this same strain (Fig. 1B). A few micrographs of fixed *G. oceanica* cells from our culture stock, as well as the batch culture data of Sett et al. (2014) and Faucher et al. (2017)—all the same strain of *G. oceanica* (RCC1303)—suggest our cell size estimates are representative for nutrient-replete batch cultures of this strain.

2.5.1.4 Carbon isotope fractionation in organic matter

Alkenone carbon isotope ratios varied from -30±0.4 ‰ to -41±0.4 ‰. Alkenone $\delta^{13}C$ reflects the source $\delta^{13}C_{DIC}$ as well as the effect of CO_2 concentration and physiology on isotope

discrimination by RubisCO. In our cultures, $\delta^{13}C_{Mk37:2}$ is highly correlated with $[CO_{2(aq)}]$ (r^2 = 0.78) because the pCO_2 of the culture vessels was established by continuously aerating the headspace with isotopically light CO_2; this is further evident from the relationship between $\delta^{13}C_{DIC}$ and $[CO_{2(aq)}]$ (r^2 = 0.83) (Fig. S1). The slope of a linear regression between $[CO_{2(aq)}]$ and $\delta^{13}C_{Mk37:2}$ is ~20% steeper than that of $[CO_{2(aq)}]$ vs. $\delta^{13}C_{DIC}$, demonstrating the presence of the CO_2 effect on $\delta^{13}C_{Mk37:2}$ in addition to the carbonate system manipulation (Fig. S1). This CO_2 effect is clear when isotope fractionation is examined as ε_p, the isotope discrimination with respect to $\delta^{13}C_{CO2aq}$ values (Fig. S2). Measured ε_p falls between 7.4‰ and 11.5‰, a range comparable to that observed in many batch cultures, and sediments spanning the past ~23 Ma (Pagani, 2014). As observed in many other datasets and as expected from theory, we find an inverse relationship between ε_p with both $1/[CO_{2(aq)}]$ ($R = -0.62$, $p < 0.05$) and growth rate ($R = -0.52$, $p < 0.05$) (Fig. S2). Under these growth conditions, higher growth rates therefore lead to smaller ε_p, and higher $[CO_{2(aq)}]$ is associated with larger ε_p.

2.5.1.5 Carbon demand and diffusive CO_2 supply

In the conventional framework, carbon demand is set by the instantaneous cellular growth rate and the carbon content per cell; diffusive supply is calculated as the product of the cellular surface area, the permeability of the cell membrane to CO_2, and the $[CO_{2(aq)}]$. Alkenone carbon isotope fractionation and τ are significantly correlated in our dataset, demonstrating a strong effect of physiology and CO_2 supply on alkenone ε_p. In our culture experiments, τ is primarily driven by changing $[CO_{2(aq)}]$ and growth rate, and ranged from approximately 0.5 to 2.5. Estimated cell radius (and by inference, POC $cell^{-1}$) varied by less than 20% across treatments, and instantaneous growth rate changed by a maximum of ~25%. The largest range in

carbon supply or demand was driven by $[CO_{2(aq)}]$, (5.8 to 33.4) μmol kg^{-1}, a ~5.8-fold change). Indeed, the strongest correlation between ε_p and the components of τ was found with $1/[CO_{2(aq)}]$ (Fig. S2). Our results support the foundational observations that carbon isotope fractionation carries a signature of the ambient $[CO_{2(aq)}]$. However, our dataset covers a small ε_p range and does not reach low τ values where carbon demand is very low relative to supply, so we are unable to determine the continuity of the ε_p response from our experiments alone.

2.5.2 Culture synthesis

We synthesized existing culture data of alkenone-producing coccolithophores that report ε_p and sufficient physiological information and metadata to estimate τ in each sample. These additional data include *E. huxleyi* strains BT6 (Bidigare et al., 1997; Popp et al., 1998b; Popp et al., 1998a), PML B92/11 (Bidigare et al., 1997; Popp et al., 1998b; Popp et al., 1998a; Riebesell et al., 2000; Rost et al., 2002), RCC1216, and RCC1256 (McClelland et al., 2017), and *G. oceanica* strains RCC1211, and RCC1314 (McClelland et al., 2017). While there may be differences at the strain level, including data from all strains in our analysis is most applicable to sediment studies, as one cannot distinguish alkenones at the strain level in the sedimentary record. Our compilation enables data from all included studies to be directly compared because we account for differences in sampling times and τ. Our synthesis permits a robust evaluation of the carbon isotope response of alkenone-producing algae to changes in CO_2 supply and cellular carbon demand.

2.5.2.1 Relationships between carbon supply, cellular carbon demand, and carbon isotope fractionation

Combining our new *G. oceanica* data with published Noelaerhabdaceae ε_p data, we find a significant inverse linear correlation between ε_p and $1/[CO_{2(aq)}]$ ($R = -0.32$; $p < 0.01$; Fig. S2A). Diffusive CO_2 supply—or CCM activity that responds to the ambient CO_2 concentration like diffusive supply—can therefore explain a small portion of the variance in ε_p. Adding instantaneous growth rate (μ_i) and relating ε_p to $\mu_i/[CO_{2(aq)}]$ increases the correlation ($R = -0.50$; $p < 0.001$; Fig S2B). Finally, when accounting for the cellular organic carbon demand relative to diffusive CO_2 supply (τ, Fig. 2A), there is a clear, though still scattered, linear relationship between ε_p and τ ($R = -0.53$, $p < 0.001$). However, visual inspection of this relationship suggests a non-linear fit is more appropriate; the Akaike's information criterion (AIC) for a two-term exponential function is substantially lower than the AIC for a linear model (485 vs. 511). Fitting ε_p as a two-term exponential function of τ, we are able to explain ~45% of the variance in ε_p (Fig. 2A). This fit suggests that the parameters used to construct τ are indeed primary factors determining carbon isotope fractionation, and the algae record information about the ambient CO_2 environment at the time of growth and carbon fixation. However, the non-linear relationship suggests that CCMs or other processes are present, as a linear relationship is expected for diffusive supply alone (Rau et al., 1996; Laws et al., 2002).

Finally, particularly at low values of τ (< 1), we observe a wide range of possible ε_p responses. The variable that is most strongly correlated with the range of ε_p responses at constant τ is the light conditions in the cultures (photoperiod and light intensity). For the same τ range, high integrated daily irradiance (product of photoperiod length and instantaneous photon flux density) yields more positive ε_p values compared to low-irradiance cultures. The effect of light is

most evident at low τ values as those cultures have the largest differences in light conditions. The effect is further evident when plotting ε_p against integrated daily irradiance (Fig. 2B). Below we discuss the non-linear relationship of ε_p and τ, the effect of light on ε_p, as well as steps forward for alkenone paleobarometry.

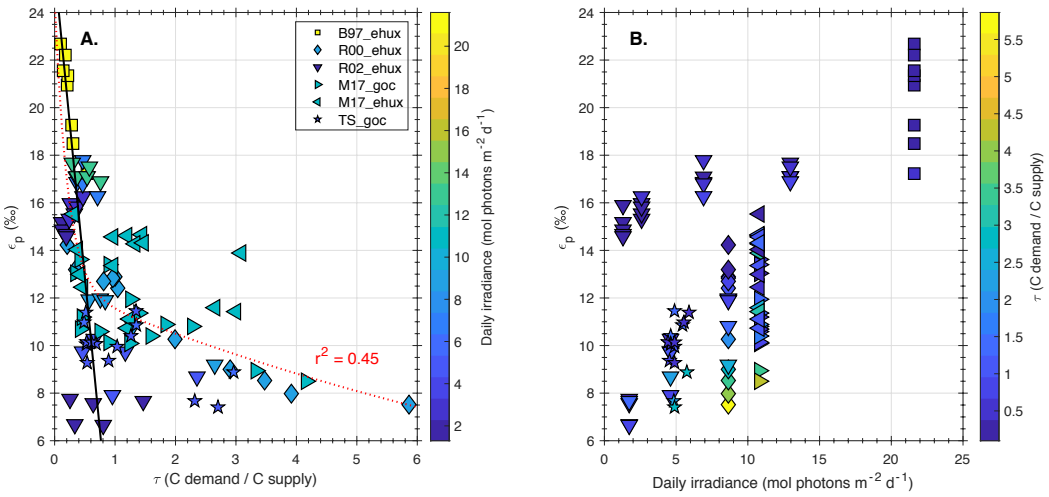

Figure 2. Carbon isotope fractionation in organic matter (ε_p) from new *G. oceanica* cultures (this study) and published data from other alkenone-producing algae. (A) ε_p as a function of carbon fixation relative to diffusive CO_2 supply to the cell (τ), colored by total daily irradiance. Red line is a two-term exponential fit to all data plotted, black line is the solution to the Rau et al. (1996) model, which assumes only diffusive CO_2 supply. (B) ε_p as a function of total daily irradiance, colored by τ. Total daily irradiance is calculated as the product of light intensity and the number of seconds of illumination, converted to units of mol m^{-2} d^{-1}. Symbols are the same in both panels and indicate the publication and species, where the abbreviation is the first initial of the first author's last name, the year, and the species studied (ehux = *Emiliania huxleyi*; goc = *Gephyrocapsa oceanica*). Reference abbreviations are as follows: B97: Bidigare et al., 1997 (including data from Popp et al., 1998a; 1998b). R00: Riebesell et al., 2000. R02: Rost et al., 2002. M17: McClelland et al., 2017. TS: this study.

2.6 Discussion

Our analysis of existing and new ε_p and τ determinations from cultures of alkenone-producing algae demonstrates a number of relationships that will be important to address in any understanding and application of alkenone ε_p values:

i) ε_p as a function of τ appears to follow a non-linear relationship rather than the linear relationship predicted by the Rau et al. (1996) model and the alkenone- CO_2 proxy.

ii) The scatter—and much of the relationship itself—in ε_p vs. τ is correlated to the light conditions of the cultures.

iii) The values of ε_p in many cultures exceed the value of ε_f measured on Form ID RubisCO in the model coccolithophore *E. huxleyi* (Boller et al., 2011).

Here we present an empirical analysis of the relationship between ε_p and τ, where we consider the Rau et al. (1996) diffusive model as well as the presence and influence of active carbon uptake by CCMs. This analysis uses the previously assumed value of 25‰ for ε_f rather than the experimentally determined 11-18‰ for RubisCO Form ID (Boller et al., 2011; Boller et al., 2015). While the 25‰ value for ε_f is probably not correct, we feel it is important to address published interpretations of carbon isotope fractionation in algae based upon this value and the diffusive model. In the remainder of the discussion we examine which experimental and environmental factors may be responsible for elevating ε_p above ε_f. Finally, we conclude with a discussion of future directions for alkenone carbon isotope ratios as a proxy of past environmental conditions.

2.6.1 The shape of the ε_p-τ response: non-diffusive CO_2 supply

The foundational works on which the alkenone-CO_2 proxy is based assume a linear relationship between carbon isotope fractionation (ε_p) and carbon demand relative to carbon supply (τ) (Rau et al., 1996; Bidigare et al., 1997; Popp et al., 1998b; Pagani et al., 2000).

Contrary to this theory and framework, but consistent with other recent studies (Laws et al., 2002; Cassar et al., 2006; McClelland et al., 2017; Stoll et al., 2019), we find a highly non-linear relationship between ε_p and τ. This conclusion is not original, and firmly supports those of previous studies. However, the carbon isotope response of *G. oceanica* to variable [$CO_{2(aq)}$] through pH modification (our experiments) has not previously been studied; existing cultures examining the response of ε_p to variable [$CO_{2(aq)}$] in *G. oceanica* maintained constant pH but modified alkalinity and dissolved inorganic carbon concentrations dramatically (McClelland et al., 2017). Combining our new *G. oceanica* batch cultures with the published data we compile here increases the size of the database by ~15%. Furthermore, the strict standardization techniques used in our compilation account for changes in cell size and carbon fixation rate in ways that have been overlooked in the past, allowing us to better compare data across different studies.

Empirically, the data strongly demonstrate that ε_p decreases with increasing τ (Fig. 2A). This relationship is one of the most significant features of the dataset: it suggests that the organic carbon isotope fractionation by alkenone-producing algae is indeed recording information about the cellular carbon economics, including ambient [$CO_{2(aq)}$]. There is, therefore, some utility in algal paleobarometry, if other physiological and environmental factors can be well constrained. However, the diffusive model does not accurately predict isotope fractionation, as evidenced by the increasing separation between the Rau et al. (1996) diffusive model and the experimental data as τ increases (Fig. 2A). Furthermore, applications of the alkenone paleobarometer that attempt to account for changes in cell size or growth rate (e.g. Pagani et al., 2000; Henderiks and Pagani, 2007) following the scaling factors of Popp et al. (1998b) would only be valid if the ε_p response is linear, as the slope of the Rau et al. (1996) line is a linear function of cell size and

growth rate. The data are much better fit with an exponential relationship (red line, Fig. 2A), which suggests the scaling factor between cell size and ε_p changes as a function of τ.

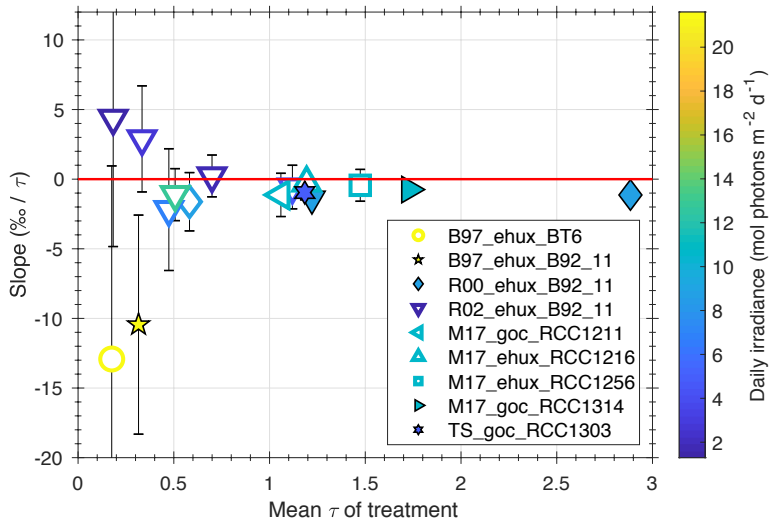

Figure 3. Meta-analysis of individual experimental treatments in culture data of alkenone-producing algae where τ is well constrained. Slope of the ε_p vs. τ relationship in each treatment plotted as a function of the mean τ in each treatment. Error bars are 90% confidence intervals on the estimate of the slope; symbols are colored by total daily irradiance. Solid-filled symbols have ε_p vs. τ slopes that are different from 0 at the 90% confidence level. The abbreviations in the legend are similar to those in Fig. 2, with the first initial of the first author's last name, the year of publication, the species studied, and the algal strain. There are more symbols plotted than entries in the legend because one study (R02; Rost et al., 2002) has multiple light conditions.

We conduct an additional meta-analysis on the data presented in Fig. 2. Here we examine how the slope between ε_p and τ, and the sensitivity of ε_p as a paleobarometer, changes with τ across experimental treatments. We define a "treatment" as a unique combination of lab, algal strain, and light conditions. From five studies with sufficient data to calculate the statistics on this relationship, there are 15 unique treatments, which subject a total of four strains of *E. huxleyi* and three strains of *G. oceanica* to various CO_2 conditions. We find that all slopes are either negative (decreasing ε_p with increasing τ) or indistinguishable from 0 at the 90% confidence level (Fig. 3). Although discrete factors unique to each culture experiment—such as cell density, $\delta^{13}C_{DIC}$ measurement time, or seawater composition—may affect the absolute ε_p in each

treatment, this analysis shows that ε_p is most sensitive to τ at low τ values, and the sensitivity declines as τ increases. Therefore, the τ range of the treatment is important in determining both the amplitude of the ε_p response and the magnitude of the slope.

Some authors have argued that the nonlinearity of the ε_p vs. τ relationship is indicative of non-diffusive CO_2 supply (Laws et al., 1997; Laws et al., 2002; Stoll et al., 2019). The apparent asymptote in carbon isotope discrimination suggests that as τ increases, the photosynthetic apparatus reaches a CO_2 threshold below which carbon fixation does not proceed because O_2 outcompetes CO_2 at the active site of RubisCO (Raven, 1997). Approaching this threshold where CO_2 becomes less available, algae rely more heavily on import and/or transport of HCO_3^- across intracellular membranes to meet their carbon fixation requirement (Anning et al., 1996; Bach et al., 2013; Holtz et al., 2015b; Holtz et al., 2015a; Holtz et al., 2017; McClelland et al., 2017). Import of HCO_3^- and movement across internal cellular membranes follows facilitated diffusion or active transport pathways because it is a charged species. It is then converted to CO_2—likely by the enzyme carbonic anhydrase (CA) (Quiroga and González, 1993; Mackinder et al., 2011) or through acid-base equilibrium set by internal pH gradients (Falkowski and Raven, 2007; Holtz et al., 2015a)—and used for carbon fixation. This type of CCM has been termed the "chloroplast pump model" (Hopkinson, 2014). Greater HCO_3^- uptake at higher τ values would raise the chloroplast CO_2 concentration above what is expected or possible from CO_2 diffusion alone (Hopkinson et al., 2011; Bach et al., 2013; Bolton and Stoll, 2013; Isensee et al., 2014). At thermodynamic equilibrium in seawater, $\delta^{13}C$ of CO_2 is approximately 10‰ lower than HCO_3^-. Dehydration of HCO_3^- by CA imparts a negative 10.1‰ isotope effect, generating CO_2 that is isotopically similar to ambient equilibrium CO_2 (Riebesell and Wolf-Gladrow, 1995). This type of CCM is therefore isotopically undetectable and it appears as an artificial enhancement of CO_2

concentrations (Zeebe and Wolf-Gladrow, 2001). Other CCMs may indeed be active and responsible for some of the ε_p variance we see, but we do not have measurements to directly test for them.

Another possible explanation for the nonlinear ε_p response as τ increases invokes changes in the effective permeability of cellular membranes to CO_2. Using the Rau et al. (1996) diffusive model, Stoll et al. (2019) solved for the permeability needed to accurately reconstruct the measured coccolithophore ε_p in a compiled culture dataset similar to the one presented here. They found that the calculated permeability spanned over three orders of magnitude, a much broader range than that measured in several diatoms (Hopkinson et al., 2011). The authors argued based on the documented diatom CCM activity (Tortell et al., 2000; Hopkinson, 2014; Young et al., 2016), and the limited range of measured permeability but a similar range of ε_p data in diatoms, that coccolithophores are unlikely to modify membrane permeability by three orders of magnitude in order to match the observed ε_p trends. However, as we discuss next, the larger challenge to the conventional framework is the kinetic fractionation by RubisCO, regardless of the specific CCM activity.

2.6.2 Kinetic fractionation by RubisCO and implications for the classical alkenone paleobarometer

The premise of the canonical diffusive model is that the kinetic fraction by RubisCO sets the maximum possible ε_p, fractionation associated with CO_2 diffusion sets the minimum possible ε_p, and carbon demand relative to CO_2 supply defines the slope of the line between these two end members. However, *in vitro* measurements of ε_f on RubisCO isolated from *E. huxleyi* yielded ~11‰ (95% CI: 9.8-12.6‰) (Boller et al., 2011). *In vitro* measurements made in the same lab on

ε_f from the marine diatom *S. costatum*, which also expresses form ID, gave a value of 18.5 ‰ (95% CI: 17.0-19.9‰) (Boller et al., 2015). These measurements suggest the ε_f of 25‰ traditionally used for alkenone paleobarometry requires reconsideration.

The implications for the diffusive model of a coccolithophore ε_f value less than 25‰ are profound. Coccolithophore ε_p values larger than ε_f ($\varepsilon_p > 11$‰) are abundant in culture, common in many geographic settings in the modern ocean, and present in Cenozoic marine sediments. These observations cannot be explained by the traditional diffusive model, or any model that has ε_f of 11‰ as the only significant fractionation step. The mechanism for large ε_p—abundant CO_2 supply, very low cellular carbon demand, and full expression of RubisCO's isotope preferences—can only produce a *maximum* ε_p approximately equal to ε_f. Explaining coccolithophore ε_p greater than 11‰ requires additional fractionation steps that can supply ^{13}C depleted CO_2 to the chloroplast prior to fixation by RubisCO.

The observed variation in form ID RubisCO $\delta^{13}C$ fractionation is a fundamental hurdle to a mechanistic understanding of carbon isotope fractionation in algae. Given that model alkenone-producer *E. huxleyi* and marine diatom *S. costatum* both possess form ID RubisCO, why are their measured ε_f values so different? Further, if there is such diversity in form ID ε_f, could other alkenone-producing coccolithophores (e.g., *G. oceanica* and its ancestors) have dramatically different ε_f values than measured in *E. huxleyi* by Boller et al. (2011)?

Limited data exist to evaluate the systematics and diversity of RubisCO carbon isotope fractionation. Some authors have argued for a close relationship between the catalytic properties of RubisCO and the magnitude of isotope discrimination (Tcherkez et al., 2006); the differences observed in ε_f between *E. huxleyi* and *S. costatum* could therefore be revealed by the kinetics of their RubisCO isolates. Unfortunately, detailed catalytic measurements have not been made for

these two organisms. Measured Michelis-Menten half-saturation constants (K_C) for *E. huxleyi* RubisCO range from ~40μM to 72μM, depending on the strain (Boller et al., 2011; Young, 2011). This range is much higher than that of other haptophytes (~15 to 24 μM; (Heureux et al., 2017)). Boller et al. (2015) measured a K_C of ~50 μM for *S. costatum*, which is consistent with other diatom species (~25 to 65 μM; (Young et al., 2016)). Tcherkez et al. (2006) documented relationships between the enzymatic turnover rate (k_{cat}), the half-saturation constant, and the specificity for CO_2 over O_2 ($S_{C/O}$) across RubisCO forms isolated from various organisms including bacteria, higher plants, diatoms, and green algae. The authors found higher K_C values (lower CO_2 affinities) are associated with lower values of $S_{C/O}$, as are smaller ε_f values. Although no coccolithophores have been examined in this framework, we would predict based on these relationships and the high K_C values for *E. huxleyi* that ε_f would be low relative to other RubisCO forms analyzed by Tcherkez et al. (2006), consistent with the findings of Boller et al. (2011) and Boller et al. (2015). However, more recent work with diatoms suggests while these characteristic relationships may generally describe RubisCO catalytic behavior across photosynthetic organisms, within a group of phytoplankton the relationship is more scattered (Young et al., 2016). Understanding the range of variability in ε_f is important for testing the mechanistic underpinnings of $\delta^{13}C$ fractionation in algae, and should be a priority for future research.

In lieu of robust catalytic constraints, we can examine RubisCO phylogenetics. It is plausible that carbon isotope fractionation is related to the genetic diversity of the RubisCO amino acid sequences that encode the catalytic activity of this enzyme. The active site of carboxylation is found at the union of two large subunits of RubisCO (Andersson, 2008). Phylogenies for the *rbcL* genes encoding the large subunit of RubisCO place haptophytes and

diatoms as quite distinct groups (Young et al., 2012), perhaps consistent with the ~7‰ difference in ε_f measurements from these groups. However, *G. oceanica* and *E. huxleyi* are statistically indistinguishable in their *rbcL* and *rbcS* (small subunit gene) phylogenies (Young, 2011), suggesting features of RubisCO activity, including carbon isotope fractionation, may be conserved within this lineage. Our observation of a relatively consistent ε_p signal across strains with respect to changing τ and irradiance suggests that non-RubisCO fractionation processes are conserved as well. Either ε_f varies with τ systematically across the strains that we have examined here, or ε_p is indeed recording changes in τ and light for mechanistic reasons that are still being uncovered.

2.6.3 What explains variation in ε_p?

Using an exponential model, cellular carbon demand and diffusive carbon supply (τ) can explain only ~45% of the variance in ε_p that we observe (Fig. 2A). With 56% of the variance unexplained by an exponential relationship with τ, other factors must also play a role in determining carbon isotope fractionation. Fundamental parameters for algal growth, including nutrient concentrations, light intensity, and temperature, are likely candidates. The algal cultures assessed in this compilation were grown under nutrient-limited chemostat conditions or in nutrient-replete batch conditions (published data and our new experiments). In chemostat cultures, the cells are maintained at a constant growth rate by adjusting the rate that fresh medium is added to the growth chamber, and are often grown under high and continuous light intensity. In batch culture, medium with excess macronutrients (5- to 100-fold natural open-ocean concentrations) is inoculated with cells, which are allowed to proliferate exponentially for a handful of generations until harvest, or one nutrient becomes limiting. Because nutrients are in

abundance under these conditions, temperature and light (and possibly micronutrients) are likely limiting growth. It therefore seems reasonable to surmise that these physiologically influential parameters, including light and nutrients, could impact ε_p independent of τ. Indeed, a fundamental determinant of the observed ε_p response is the nutrient availability and light environment. The highest ε_p (>20‰) is achieved only under nutrient-limited, continuous and high-intensity light experiments (Bidigare et al., 1997; Popp et al., 1998b; Wilkes et al., 2018). The next highest ε_p (~18‰) is achieved in the (nutrient-replete) batch incubations of (Rost et al., 2002), which were also grown under continuous light at relatively high light intensity (150 μmol m^{-2} s^{-1}). Most batch cultures with a light-dark cycle yield ε_p in the 8 to 14‰ range (Riebesell et al., 2000; Rost et al., 2002; McClelland et al., 2017; this study). Riebesell et al. (2000) also noted a distinct difference in the behavior of the nutrient-limited chemostat incubations of Bidigare et al. (1997) and their light-limited batch cultures, which had comparable $\mu/[CO_{2(aq)}]$ but very different ε_p responses. Following these observations, we suggest that the interplay between light and nutrient availability in modifying physiological responses is paramount to understanding ε_p and improving the use of $\delta^{13}C$ as a paleobarometer.

Across the compiled culture dataset, we see a significant relationship between ε_p and the integrated daily irradiance (μmol photons m^{-2} d^{-1}), where higher ε_p is achieved at higher total daily irradiance (Fig. 2B). As discussed above, these large ε_p values likely require a highly ^{13}C-depleted internal DIC pool in the vicinity of RubisCO, because the maximum fractionation RubisCO can impart is 11‰ (Boller et al., 2011). The data, particularly the low-light experiments of Rost et al. (2002) (dark blue inverted triangles, Fig. 2A), also demonstrate that high CO_2 supply relative to demand (low τ) is necessary for explaining high ε_p, but it is not sufficient on its own (data populating lower left of Fig. 2A). Whatever process produces more

negative $\delta^{13}C_{CO2}$ inside the cell is enhanced at high irradiance, longer photoperiods, and perhaps low nutrient availability.

Why might light be responsible for this phenomenon? The photosynthetic machinery undergoes significant physiochemical change when exposed to light; proton and carbon fluxes across internal membranes establish intracellular pH gradients, and in some organisms, RubisCO assembles into densely packed pyrenoids at the center of cells (Nassoury et al., 2005). Under high photon flux density, photosystem activity is stimulated, dissipating some incident light energy by splitting water molecules and generating protons in the thylakoid lumen that drive ATP generation. Proton movement may be coupled with bicarbonate transport to help maintain charge balance. At high, continuous irradiance, the thylakoid lumen remains acidic with continuous water splitting and proton generation. HCO_3^- transported into the lumen would be converted to CO_2 rapidly because of the low lumenal pH, which could then diffuse into the pyrenoid and be fixed by RubisCO. If the HCO_3^- that crosses the thylakoid membrane was derived from carbonic anhydrase, it would carry a light isotopic signature, yielding isotopically light organic matter.

Wilkes and Pearson (2019) synthesized these ideas into the concept of a photocatalytic CCM, which is more active at high light intensity and low nutrient conditions, is responsible for the large isotope fractionation in chemostat experiments. High irradiance and low nutrient conditions must therefore result in more ^{13}C-depleted DIC in the vicinity of RubisCO. Their model, which includes this biochemical fractionation step during hydroxylation of CO_2 to HCO_3^- during transit across the thylakoid membrane under high photon flux, was able to explain a large proportion of the observed culture ε_p data across a range of eukaryotic species.

We explore the role of light and nutrients by examining the relationship between ε_p at the mean τ of each treatment and the ratio of daily irradiance to the $[NO_3^-]$ in the growth medium (Fig. 4). We find that both τ and light/nutrients affect ε_p. When viewed through the relationship of ε_p and τ (Fig. 4A), smaller ε_p corresponds to lower light/high nutrients within a constant τ range. Alternatively, when viewed through the relationship of light/nutrients and ε_p (Fig. 4B), smaller τ values correspond to larger ε_p values. This analysis reinforces the observations that high light/low nutrient experiments produce the highest ε_p, and increasing τ at a given light to nutrient ratio reduces ε_p.

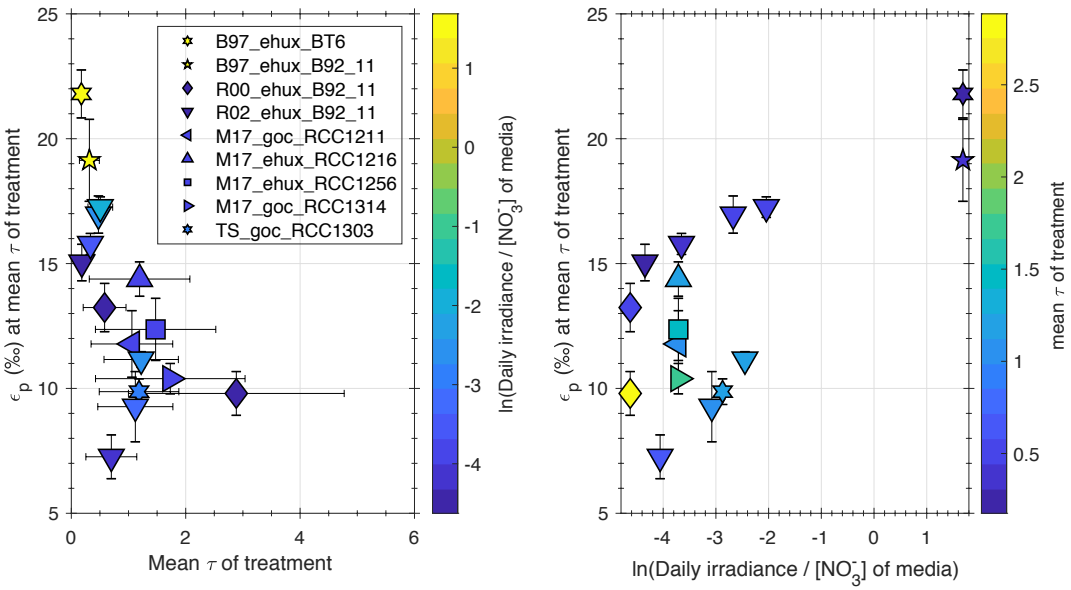

Figure 4. Calculated carbon isotope fractionation (ε_p) in alkenone-producing coccolithophores at the mean τ of each treatment with respect to (A) the mean τ of each treatment, (B) ratio of daily irradiance to the $[NO_3^-]$ in the culture media. Error bars on ε_p are 95% confidence intervals on the predicted ε_p at the mean τ based on the linear fit from the metanalysis. Horizontal error bars in (A) show the maximum and minimum τ values in the treatment. Legend abbreviations are the same as in Fig. 3. Symbols in (A) are colored by the natural log of the ratio of lights to $[NO_3^-]$; symbols in (B) are colored by the mean τ of the treatment.

Although we do not understand in detail, or have experimental evidence for, the mechanism of an irradiance effect on overall carbon isotope fractionation, we can probe the influence of light on top of carbon demand and supply through multivariate linear regressions (Stoll et al., 2019). First, we develop a linear regression model to predict ε_p as a function of the components of τ: instantaneous growth rate (μ_{inst}), cell radius (μm), and [$CO_{2(aq)}$]. We also allow for a quadratic term in [$CO_{2(aq)}$] because it improves the quality of the fit and the sensitivity of ε_p to [$CO_{2(aq)}$]. With these variables, we are able to model 47% of the variance in ε_p (Fig. 5A), slightly better than the exponential fit of ε_p to τ (section 4.2.1.). This model is markedly improved if we include the light parameters. We can explain 88% of the variance in ε_p if we instead model ε_p as a function of growth rate (μ), cell radius, [$CO_{2(aq)}$] (and a quadratic term), the light intensity, and light hours (Fig. 5B). A third model that uses the integrated daily irradiance instead of light hours and light intensity independently also performs better than the τ model (r^2 = 0.66; Fig. S6). We elect to use light hours and cellular growth rate (μ) rather than photoperiod-adjusted instantaneous growth rate (μ_{inst}) because μ_{inst} is a function of photoperiod and requires an assumption about the magnitude of respiratory carbon loss in the dark (see Eq. 3). These regression models indicate that light intensity and duration are essential parameters for understanding carbon isotope fractionation in algae (independent of their effect on growth rate) and must be considered in natural samples.

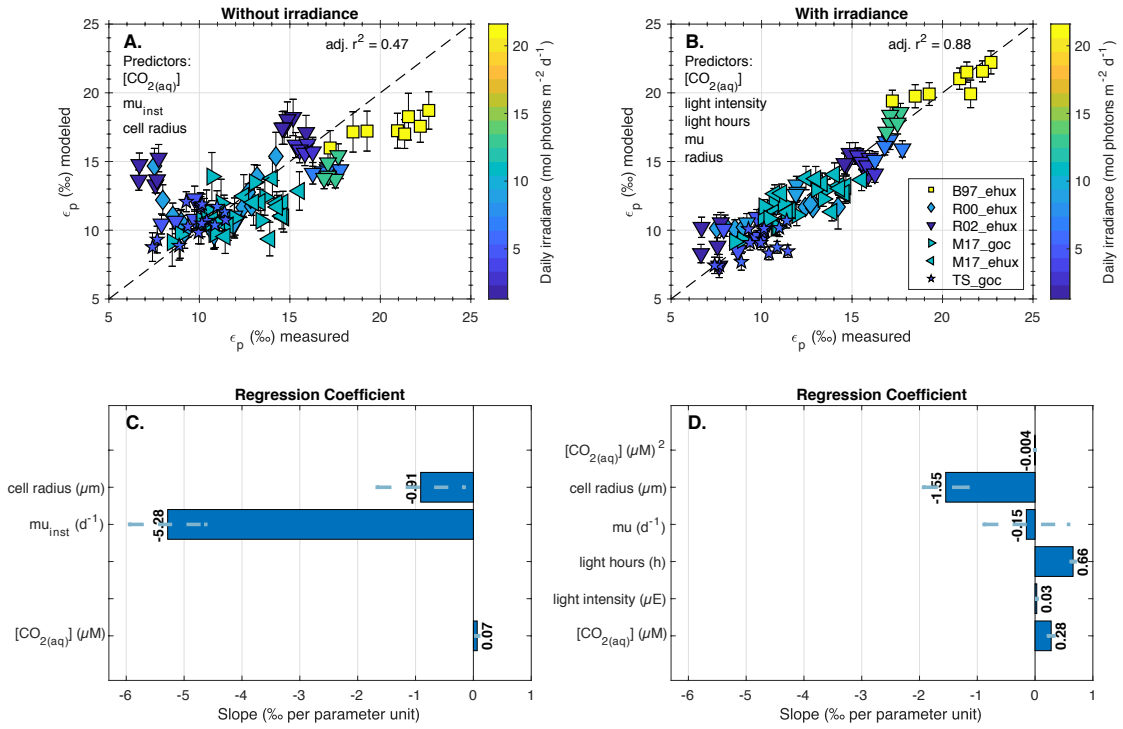

Figure 5. Multiple linear regression models to predict ε_p in cultures of alkenone-producing algae. Error bars show the 95% confidence intervals on the modeled ε_p, and samples are colored by total daily irradiance. (A) Predictor variables include only the components of τ, the ratio of instantaneous carbon fixation rate to diffusive CO_2 supply. (B) Predictor variables add light intensity (μmol photons m^{-2} s^{-1}) and light duration (hours), and use growth rate (μ) instead of instantaneous growth rate (μ_i) because light hours are an independent parameter in the "with irradiance" model. (C) and (D) show the magnitude and value of the coefficients for each predictor variable in the linear regressions for (A) and (B) respectively. Dashed lines in (C) and (D) indicate 66% confidence intervals on the estimate of the coefficient value.

2.6.4 Next steps for proxy and cultures

Our ultimate goal, through the development and improved understanding of this proxy, is to use alkenone carbon isotope fractionation in ancient sediments to estimate past atmospheric CO_2 concentrations. The theoretical limit of maximum ε_p at 25‰ was previously explained by cellular CO_2 supply in such excess of CO_2 demand that carbon fixation was the rate-limiting step in CO_2 consumption, which leads to a full expression of the kinetic fractionation by RubisCO. However, because the ε_f of form ID in alkenone-producing algae may be ~11‰ (Boller et al.,

2011), more complex mechanisms must be at play. Measurements to identify these mechanisms should be a target of future culture studies and field campaigns.

Empirical calibrations, like the one we present in Figs. 2 and 5, are useful, but paleoreconstructions using these calibrations are only as strong as their (as yet limited) calibration matrix. There is a clear distinction in the ε_p and τ space that is occupied by different culture conditions: chemostats tend to produce the highest ε_p and lowest τ, while batch cultures generally yield lower ε_p and higher τ. Conditions like those in the chemostat cultures included in our analysis (e.g. high-intensity, 24-hour light) are rarely achieved in the ocean—especially not in the subtropical regions that are oceanographically suited for paleoreconstructions of air-sea equilibrium CO_2. In order to apply these empirical relationships to sediments with confidence, future work should focus on identifying whether the factors investigated here (carbon economics and light energy) are indeed the primary determinants of carbon isotope fractionation in algae in the modern ocean. There is an abundance of existing alkenone carbon isotope data from marine particulates and core-top sediments, but our analysis shows that previous validation studies (e.g. Pagani et al., 2002) have not yet rigorously tested the fundamental principles of carbon isotope fractionation in alkenones. That analysis is beyond the scope of this study, but is the focus of future work.

2.7 Conclusion

The alkenone paleobarometer is one of only a handful of methods for estimating past atmospheric CO_2 concentrations. The proxy is built on the assumption that the coccolithophorid algae producing alkenones acquire carbon solely through diffusion, and the extent of carbon isotope fractionation—ranging between the kinetic fractionation factor of the carboxylating

enzyme RubisCO and the isotope fractionation of CO_2 diffusion—is inversely proportional to CO_2 availability. We use new cultures of alkenone-producing *G. oceanica* in conjunction with a rigorous compilation and re-calculation of existing culture data from other alkenone-producing coccolithophores to test these fundamental underpinnings of the alkenone paleobarometer. We found a strong non-linear dependence of ε_p on carbon demand relative to carbon supply, indicating that CO_2 information is indeed recorded in the carbon isotope fractionation of these algae. However, as existing studies have noted, the observed relationship is quite different from the linear form of conventional proxy framework, and suggests additional processes such as carbon concentrating mechanisms must also be operating. We also show that the amount of light energy available to these organisms can explain an additional ~40% of the variance in the ε_p data. The strong covariance between irradiance and ε_p suggests that applications of the alkenone paleobarometer must take light into account. A more serious challenge to the alkenone proxy as currently applied is that the measured maximum kinetic fractionation by the RubisCO enzyme responsible for the first step in carbon fixation is ~11‰, rather than the commonly-employed and assumed 25‰. Many culture, modern ocean, and geologic samples have ε_p values larger than 11‰, which implies that additional isotope fractionation steps likely occur before carboxylation to produce a pool of ^{13}C-depleted CO_2 in the vicinity of RubisCO. There is a strong correlation between maximum ε_p and irradiance, suggesting that light and the photosynthetic machinery may be responsible for generating this isotopically light DIC pool. Although the mechanism for this phenomenon is still being uncovered, empirical relationships present a means for interpreting ε_p in terms of the physiological and environmental parameters controlling alkenone $\delta^{13}C$.

2.8 Acknowledgements

This work was supported by the Center for Climate and Life at Columbia University, as well as the Lamont Climate Center and the G. Unger Vetlesen Foundation. SRP was supported by a U.S. National Science Foundation graduate research fellowship (grant no. DGE16-44869). Additional support was provided by the National Science Foundation Biological Oceanography Program (grant no. OCE1314336, STD), with partial support also provided by WSL Pure in partnership with Columbia University's Center for Climate and Life, the Paul M. Angell Family Foundation, Columbia Earth Institute Research Assistantship (OMW), and The Columbia University Bridge to PhD Program in the Natural Sciences Research Assistantship (MDHL). We thank Hugh Ducklow and Naomi Shelton for the use of and training on the autotitrator and Andrew Juhl for contributions to the original experimental design and construction of a CO_2 flowmeter manifold. We thank Heather Stoll, Peter deMenocal, and Bärbel Hönisch for discussions and constructive comments on the manuscript, and Wei Huang for laboratory assistance.

2.9 References

Aloisi G. (2015) Covariation of metabolic rates and cell size in coccolithophores. *Biogeosciences* **12**, 4665–4692.

Andersson I. (2008) Catalysis and regulation in Rubisco. *J. Exp. Bot.* **59**, 1555–1568.

Anning T., Nimer N., Merrett M. J. and Brownlee C. (1996) Costs and benefits of calcification in coccolithophorids. *J. Mar. Syst.* **9**, 45–56.

Bach L. T., MacKinder L. C. M., Schulz K. G., Wheeler G., Schroeder D. C., Brownlee C. and Riebesell U. (2013) Dissecting the impact of CO2 and pH on the mechanisms of photosynthesis and calcification in the coccolithophore Emiliania huxleyi. *New Phytol.* **199**, 121–134.

Badger M. P. S., Foster G. L., Chalk T. B., Gibbs S. J., Badger M. P. S., Pancost R. D., Schmidt D. N., Sexton P. F., Mackensen A., Bown P. R. and Pälike H. (2019) Insensitivity of alkenone carbon isotopes to atmospheric CO2 at low to moderate CO2 levels. *Clim. Past* **15**, 539–554.

Badger M. R., Andrews T. J., Whitney S. M., Ludwig M., Yellowlees D. C., Leggat W. and Price G. D. (1998) The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO_2-concentrating mechanisms in algae. *Can. J. Bot.* **76**, 1052–1071.

Balch W. M., Holligan P. M. and Kilpatrick K. A. (1992) Calcification, photosynthesis and growth of the bloom-forming coccolithophore,Emiliania huxleyi. *Cont. Shelf Res.* **12**, 1353–1374.

Bidigare R. R., Fluegge A., Freeman K. H., Hanson K. L., Hayes J. M., Hollander D., Jasper J. P., King L. L., Laws E. A., Milder J., Millero F. J., Pancost R., Popp B. N., Steinberg P. A. and Wakeham S. G. (1997) Consistent fractionation of 13C in nature and in the laboratory: Growth-rate effects in some haptophyte algae. *Global Biogeochem. Cycles* **11**, 279–292.

Boller A. J., Thomas P. J., Cavanaugh C. M. and Scott K. M. (2015) Isotopic discrimination and kinetic parameters of RubisCO from the marine bloom-forming diatom, Skeletonema costatum. *Geobiology* **13**, 33–43.

Boller A. J., Thomas P. J., Cavanaugh C. M. and Scott K. M. (2011) Low stable carbon isotope fractionation by coccolithophore RubisCO. *Geochim. Cosmochim. Acta* **75**, 7200–7207.

Bolton C. T., Hernández-Sánchez M. T., Fuertes M.-Á., González-Lemos S., Abrevaya L., Mendez-Vicente A., Flores J.-A., Probert I., Giosan L., Johnson J. and Stoll H. M. (2016) Decrease in coccolithophore calcification and CO2 since the middle Miocene. *Nat. Commun.* **7**, 10284.

Bolton C. T. and Stoll H. M. (2013) Late Miocene threshold response of marine algae to carbon dioxide limitation. *Nature* **500**, 558–562.

Brassell S. C., Dumitrescu M., Bralower T. J., Premoli-Silva I., Malone M. J., Arthur M. A., Averyt K., Bown P., Channell J. E. T., Clarke L. J., Dutton A. L., Eleson J. W., Frank T. D., Gylesjö S., Hancock H., Kano H., Leckie R. M., Marsaglia K. M., McGuire J., Moe K. T., Petrizzo M. R., Robinson S., Röhl U., Sager W. W., Takeda K., Thomas D., Williams T. and Zachos J. C. (2004) Recognition of alkenones in a lower Aptian porcellanite from the west-central Pacific. *Org. Geochem.* **35**, 181–188.

Burkhardt S., Riebesell U. and Zondervan I. (1999) Effects of growth rate , CO2 concentration , and cell size on the stable carbon isotope fractionation in marine phytoplankton. *Geochim. Cosmochim. Acta* **63**, 3729–3741.

Burky D. (1971) Cenozoic calcareous nannofossils from the Pacific Ocean. *Trans. San Diego Soc. Nat. Hist.* **16**, 303–328.

Cassar N., Laws E. a. and Popp B. N. (2006) Carbon isotopic fractionation by the marine diatom Phaeodactylum tricornutum under nutrient- and light-limited growth conditions. *Geochim. Cosmochim. Acta* **70**, 5323–5335.

Dickson Andrew G., Sabine Christopher L. and Christian James R. (2007) *Guide to Best Practices for Ocean CO2 Measurements*. eds. A.G. Dickson, C.L. Sabine, and J.R. Christian,

Falkowski P. G. and Raven J. A. (2007) *Aquatic Photosynthesis (Second edition).*,

Faucher G., Hoffmann L., Bach L. T., Bottini C., Erba E. and Riebesell U. (2017) Impact of trace metal concentrations on coccolithophore growth and morphology: laboratory simulations of Cretaceous stress. *Biogeosciences* **14**, 3603–3613.

Freeman K. H. and Hayes J. M. (1992) Fractionation of carbon isotopes by phytoplankton and estimates of ancient CO2 levels. *Global Biogeochem. Cycles* **6**, 185–198.

Freeman K. H. and Pagani M. (2005) Alkenone-Based Estimates of Past CO2 Levels: A Consideration of Their Utility Based on an Analysis of Uncertainties. *A Hist. Atmos. CO2 Its Eff. Plants, Anim. Ecosyst.*, 35–61.

Gartner S. (1969) Correlation of Neogene planktonic foraminifer and calcareous nannofossil zones.; Geology of the American Mediterranean. *Trans. Gulf Coast Assoc. Geol. Soc.* **19**, 585–599.

Geitzenauer K. R., Roche M. B. and McIntyre A. (1976) Modern pacific coccolith assemblages: Derivation and application to late Pleistocene paleotemperature analysis. *Mem. Geol. Soc. Am.* **145**, 423–448.

Giordano M., Beardall J. and Raven J. A. (2005) CO2 Concentrating Mechanisms in Algae: Mechanisms, Environmental Modulation, and Evolution. *Annu. Rev. Plant Biol.* **56**, 99–131.

Goericke R., Montoya J. P. and Fry B. (1994) Physiology of isotopic fractionation in algae and cyanobacteria. In *Stable isotopes in ecology and environmental science* (eds. K. Lajtha and R. H. Michener). Methods in ecology. Blackwell Scientific Publications, Oxford; Boston. pp. 187–221.

Guillard R. R. L. and Hargraves P. E. (1993) Stichochrysis immobilis is a diatom, not a chrysophyte. *Phycologia* **32**, 234–236.

Hagino K. and Young J. R. (2015) Biology and Paleontology of Coccolithophores (Haptophytes). In *Marine Protists: Diversity and Dynamics* (eds. S. Ohtsuka, T. Suzaki, T. Horiguchi, N. Suzuki, and F. Not). Springer Japan, Tokyo. pp. 311–330.

Henderiks J. (2008) Coccolithophore size rules - Reconstructing ancient cell geometry and cellular calcite quota from fossil coccoliths. *Mar. Micropaleontol.* **67**, 143–154.

Henderiks J. and Pagani M. (2007) Refining ancient carbon dioxide estimates: Significance of coccolithophore cell size for alkenone-based pCO2 records. *Paleoceanography* **22**, 1–12.

Hennon G. M. M., Williamson O. M., Hernández Limón M. D., Haley S. T. and Dyhrman S. T. (2019) Non-linear Physiology and Gene Expression Responses of Harmful Alga Heterosigma akashiwo to Rising CO2. *Protist* **170**, 38–51.

Heureux A. M. C., Young J. N., Whitney S. M., Eason-Hubbard M. R., Lee R. B. Y., Sharwood R. E. and Rickaby R. E. M. (2017) The role of Rubisco kinetics and pyrenoid morphology in shaping the CCM of haptophyte microalgae. *J. Exp. Bot.* **68**, 3959–3969.

Holtz L.-M., Wolf-Gladrow D. and Thoms S. (2015a) Numerical cell model investigating cellular carbon fluxes in Emiliania huxleyi. *J. Theor. Biol.* **364**, 305–315.

Holtz L.-M., Wolf-Gladrow D. and Thoms S. (2015b) Simulating the effects of light intensity and carbonate system composition on particulate organic and inorganic carbon production in Emiliania huxleyi. *J. Theor. Biol.* **372**, 192–204.

Holtz L. M., Wolf-Gladrow D. and Thoms S. (2017) Stable carbon isotope signals in particulate organic and inorganic carbon of coccolithophores – A numerical model study for Emiliania huxleyi. *J. Theor. Biol.* **420**, 117–127.

Hopkinson B. M. (2014) A chloroplast pump model for the CO2 concentrating mechanism in the diatom Phaeodactylum tricornutum. *Photosynth. Res.* **121**, 223–233.

Hopkinson B. M., Dupont C. L., Allen A. E. and Morel F. M. M. (2011) Efficiency of the CO2-concentrating mechanism of diatoms. *Proc. Natl. Acad. Sci.*, 1–8.

Isensee K., Erez J. and Stoll H. M. (2014) Detection of a variable intracellular acid-labile carbon pool in Thalassiosira weissflogii (Heterokontophyta) and Emiliania huxleyi (Haptophyta) in

response to changes in the seawater carbon system. *Physiol. Plant.* **150**, 321–338.

Jasper J. P. and Hayes J. M. (1990) A carbon isotope record of CO2 levels during the late Quaternary. *Nature* **347**, 462–464.

Jasper J. P., Hayes J. M., Mix A. C. and Prahl F. G. (1994) Photosynthetic fractionation of 13C and concentrations of dissolved CO2 in the central equatorial Pacific during the last 255,000 years. *Paleoceanography* **9**, 781–798.

Juhl A. R. and Latz M. I. (2002) Mechanisms of fluid shear-induced inhibition of population growth in a red-tide dinoflagellate. *J. Phycol.* **38**, 683–694.

Keller K. and Morel F. M. M. (1999) A model of carbon isotopic fractionation and active carbon uptake in phytoplankton. *Mar. Ecol. Prog. Ser.* **182**, 295–298.

Kottmeier D. M., Rokitta S. D., Tortell P. D. and Rost B. (2014) Strong shift from HCO3 − to CO2 uptake in Emiliania huxleyi with acidification: new approach unravels acclimation versus short-term pH effects. *Photosynth. Res.* **121**, 265–275.

Krumhardt K. M., Lovenduski N. S., Iglesias-Rodriguez M. D. and Kleypas J. A. (2017) Coccolithophore growth and calcification in a changing ocean. *Prog. Oceanogr.* **159**, 276–295.

Laws E. A. (1997) *Mathematical Methods for Oceanographers: An Introduction*., Wiley.

Laws E. A. and Bannister T. T. (1980) Nutrient- and light-limited growth of Thalassiosira fluviatilis in continuous culture with implications for phytoplankton growth in the ocean. *Limnol. Oceanogr.* **25**, 457–473.

Laws E. A., Bidigare R. R. and Popp B. N. (1997) Effect of growth rate and CO2 concentration on carbon isotopic fractionation by the marine diatom Phaeodactylum tricornutum. *Limnol. Oceanogr.* **42**, 1552–1560.

Laws E. A., Popp B. N., Bidigare R. R., Kennicutt M. C. and Macko S. A. (1995) Dependence of phytoplankton carbon isotopic composition on growth rate and [CO2]aq: Theoretical considerations and experimental results. *Geochim. Cosmochim. Acta* **59**, 1131–1138.

Laws E. A., Popp B. N., Cassar N., Tanimoto J., Cassar N. and Tanimoto J. (2002) 13C discrimination patterns in oceanic phytoplankton: likely influence of CO2 concentrating mechanisms, and implications for paleoreconstructions. *Funct. plant Biol.* **29**, 323–333.

Lewis E. and Wallace D. (1998) CO2SYS: Program developed for CO2 system calculations. *Oak Ridge Natl. Lab. Dioxide Inf. Anal. Cent. 105*.

Lorimer G. H. and Miziorko H. M. (1980) Carbamate Formation on theε-Amino Group of a Lysyl Residue as the Basis for the Activation of Ribulosebisphosphate Carboxylase by

Carbon Dioxide and Magnesium(2+). *Biochemistry* **19**, 5321–5328.

Lueker T. J., Dickson A. G. and Keeling C. D. (2000) Ocean pCO2 calculated from dissolved inorganic carbon, alkalinity, and equations for K1 and K2: Validation based on laboratory measurements of CO2 in gas and seawater at equilibrium. *Mar. Chem.* **70**, 105–119.

Mackinder L., Wheeler G., Schroeder D., von Dassow P., Riebesell U. and Brownlee C. (2011) Expression of biomineralization-related ion transport genes in Emiliania huxleyi. *Environ. Microbiol.* **13**, 3250–3265.

Marlowe I. T. T., Green J. C. C., Neal A. C. C., Brassell S. C. C., Eglinton G. and Course P. A. (1984) Long chain (n-C37–C39) alkenones in the Prymnesiophyceae. Distribution of alkenones and other lipids and their taxonomic significance. *Br. Phycol. J.* **19**, 203–216.

McClelland H. L. O., Barbarin N., Beaufort L., Hermoso M., Ferretti P., Greaves M. and Rickaby R. E. M. (2016) Calcification response of a key phytoplankton family to millennial-scale environmental change. *Sci. Rep.* **6**, 1–11.

McClelland H. L. O., Bruggeman J., Hermoso M. and Rickaby R. E. M. (2017) The origin of carbon isotope vital effects in coccolith calcite. *Nat. Commun.* **8**, 14511.

McIntyre A. and Bé A. W. H. (2003) Modern coccolithophoridae of the atlantic ocean—I. Placoliths and cyrtoliths. *Deep Sea Res. Oceanogr. Abstr.* **14**, 561–597.

Montagnes D. J. S., Berges J. A., Harrison P. J. and Taylor F. J. R. (1993) Estimating carbon, nitrogen, protein, and chlorophyll a from volume in marine phytoplankton. *Limnol. Oceanogr.* **39**, 1044–1060.

Moolna A. and Rickaby R. E. M. (2012) Interaction of the coccolithophore Gephyrocapsa oceanica with its carbon environment: response to a recreated high-CO 2 geological past. *Geobiology* **10**, 72–81.

Müller M. N., Antia A. N. and LaRoche J. (2008) Influence of cell cycle phase on calcification in the coccolithophore Emiliania huxleyi. *Limnol. Oceanogr.* **53**, 506–512.

Nassoury N., Wang Y. and Morse D. (2005) Brefeldin A inhibits circadian remodeling of chloroplast structure in the dinoflagellate Gonyaulax. *Traffic* **6**, 548–561.

O'Leary M. H. (1984) Measurement of the isotope fractionation associated with diffusion of carbon dioxide in aqueous solution. *J. Phys. Chem.* **88**, 823–825.

Olsen A., Key R. M., Van Heuven S., Lauvset S. K., Velo A., Lin X., Schirnick C., Kozyr A., Tanhua T., Hoppema M., Jutterström S., Steinfeldt R., Jeansson E., Ishii M., Pérez F. F. and Suzuki T. (2016) The global ocean data analysis project version 2 (GLODAPv2) - An internally consistent data product for the world ocean. *Earth Syst. Sci. Data* **8**, 297–323.

Pagani M. (2014) Biomarker-Based Inferences of Past Climate: The Alkenone pCO2 Proxy. In *Treatise on Geochemistry* Elsevier. pp. 361–378.

Pagani M., Arthur M. A. and Freeman K. H. (1999) Miocene evolution of atmospheric carbon dioxide. *Paleoceanography* **14**, 273–292.

Pagani M., Arthur M. A. and Freeman K. H. (2000) Variations in Miocene phytoplankton growth rates in the southwest Atlantic: Evidence for changes in ocean circulation. *Paleoceanography* **15**, 486–496.

Pagani M., Freeman K. H., Ohkouchi N. and Caldeira K. (2002) Comparison of water column [CO2aq] with sedimentary alkenone-based estimates: A test of the alkenone-CO2 proxy. *Paleoceanography* **17**, 21-1-21–12.

Pagani M., Huber M., Liu Z., Bohaty S. M., Henderiks J., Sijp W., Krishnan S. and DeConto R. M. (2011) The Role of Carbon Dioxide During the Onset of Antarctic Glaciation. *Science (80-.).* **334**, 1261–1264.

Polissar P. J. and D'Andrea W. J. (2014) Uncertainty in paleohydrologic reconstructions from molecular δD values. *Geochim. Cosmochim. Acta* **129**, 146–156.

Popp B. N., Kenig F., Wakeham S. G., Laws E. A. and Bidigare R. R. (1998a) Does growth rate affect ketone unsaturation and intracellular carbon isotopic variability in Emiliania huxleyi ? *Paleoceanography* **13**, 35–41.

Popp B. N., Laws E. A., Bidigare R. R., Dore J. E., Hanson K. L. and Wakeham S. G. (1998b) Effect of phytoplankton cell geometry on carbon isotopic fractionation. *Geochim. Cosmochim. Acta* **62**, 69–77.

Quiroga O. and González E. L. (1993) Carbonic anhydrase in the chloroplast of a coccolithophorid (Prynesiophyceae). *J. Phycol.* **29**, 321–324.

Rau G. H., Riebesell U. and Wolf-Gladrow D. (1996) A model of photosynthetic 13C fractionation by marine phytoplankton based on diffusive molecular CO2 uptake. *Mar. Ecol. Prog. Ser.* **133**, 275–285.

Rau G. H., Takahashi T., Desmarais D. J., Repeta D. J. and Martin J. H. (1992) The Relationship between d13C of Organic-Matter and [CO2(aq)] in Ocean Surface-Water: Data from a JGOFS Site in the Northeast Atlantic-Ocean and a Model. *Geochim. Cosmochim. Acta* **56**, 1413–1419.

Rau G. H., Takahashi T. and Des Marais D. J. (1989) Latitudinal variations in plankton δ13C: Implications for CO2 and productivity in past oceans. *Nature* **341**, 516–518.

Raven J. A. (1997) Inorganic Carbon Acquisition by Marine Autotrophs. *Adv. Bot. Res.* **27**, 85–209.

Raven J. A. and Beardall J. (2003) *Carbon acquisition mechanisms of algae: carbon dioxide diffusion and carbon dioxide concentrating mechanisms*. eds. A. W. Larkum, S. E. Douglas, and J. A. Raven, Kluwer Academic Publishers.

Reinfelder J. R. (2011) Carbon Concentrating Mechanisms in Eukaryotic Marine Phytoplankton. *Ann. Rev. Mar. Sci.* **3**, 291–315.

Rickaby R. E. M., Henderiks J. and Young J. N. (2010) Perturbing phytoplankton: Response and isotopic fractionation with changing carbonate chemistry in two coccolithophore species. *Clim. Past* **6**, 771–785.

Riebesell U., Revill A. T., Holdsworth D. G. and Volkman J. K. (2000) The effects of varying CO2 concentration on lipid composition and carbon isotope fractionation in Emiliania huxleyi. *Geochim. Cosmochim. Acta* **64**, 4179–4192.

Riebesell U. and Wolf-Gladrow D. A. (1995) Growth limits on Phytoplankton. *Nature* **373**, 28.

Roeske C. A. and O'Leary M. H. (1985) Carbon isotope effect on carboxylation of ribulose bisphosphate catalyzed by ribulose bisphosphate carboxylase from Rhodospirillum rubrum. *Biochemistry* **24**, 1603–1607.

Rost B., Riebesell U. and Sültemeyer D. (2006) Carbon acquisition of marine phytoplankton: Effect of photoperiod length. *Limnol. Oceanogr.* **51**, 12–20.

Rost B., Zondervan I. and Riebesell U. (2002) Light-dependent carbon isotope fractionation in the coccolithophorid Emiliania huxleyi. *Limnol. Oceanogr.* **47**, 120–128.

Sett S., Bach L. T., Schulz K. G., Koch-Klavsen S., Lebrato M. and Riebesell U. (2014) Temperature Modulates Coccolithophorid Sensitivity of Growth, Photosynthesis and Calcification to Increasing Seawater pCO2. *PLoS One* **9**, e88308.

Stoll H. M., Guitian J., Hernandez-Almeida I., Mejia L. M., Phelps S. R., Polissar P. J., Rosenthal Y., Zhang H. and Ziveri P. (2019) Upregulation of phytoplankton carbon concentrating mechanisms during low CO2 glacial periods and implications for the phytoplankton pCO2 proxy. *Quat. Sci. Rev.* **208**, 1–20.

Tabita F. R., Satagopan S., Hanson T. E., Li H., Chan S. and Singh J. (2007) Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. *Microbiol. Mol. Biol. Rev.* **71**, 576–599.

Tcherkez G. G. B., Farquhar G. D. and Andrews T. J. (2006) Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc. Natl. Acad. Sci.* **103**, 7246–7251.

Thompson P. a. and Calvert S. E. (1995) Carbon isotope fractionation by Emiliania huxleyi.

Limnol. Oceanogr. **40**, 673–679.

Tortell P. D., Rau G. H. and Morel F. M. M. (2000) Inorganic carbon aquisition in coastal Pacific phytoplankton communities. *Limnol. Ocean.* **45**, 1485–1500.

Uppström L. R. (1974) The boron/chlorinity ratio of deep-sea water from the Pacific Ocean. *Deep Sea Res. Oceanogr. Abstr.* **21**, 161–162.

Verity P. G., Robertson C. Y., Tronzo C. R., Andrews M. G., Nelson J. R. and Sieracki M. E. (1992) Relationships between cell volume and the carbon and nitrogen content of marine photosynthetic nanoplankton. *Limnol. Oceanogr.* **37**, 1434–1446.

Wilkes E. B., Carter S. J. and Pearson A. (2017) CO2-dependent carbon isotope fractionation in the dinoflagellate Alexandrium tamarense. *Geochim. Cosmochim. Acta* **212**, 48–61.

Wilkes E. B., Lee R. B. Y., McClelland H. L. O., Rickaby R. E. M. and Pearson A. (2018) Carbon isotope ratios of coccolith–associated polysaccharides of Emiliania huxleyi as a function of growth rate and CO2 concentration. *Org. Geochem.* **119**, 1–10.

Wilkes E. B. and Pearson A. (2019) A general model for carbon isotopes in red-lineage phytoplankton: Interplay between unidirectional processes and fractionation by RubisCO. *Geochim. Cosmochim. Acta* **265**, 163–181.

Young J. (2011) Past and future adaptations of phytoplankton to carbon dioxide. Oxford University.

Young J. N., Heureux A. M. C., Sharwood R. E., Rickaby R. E. M., Morel F. M. M. and Whitney S. M. (2016) Large variation in the Rubisco kinetics of diatoms reveals diversity among their carbon-concentrating mechanisms. *J. Exp. Bot.* **67**, 3445–3456.

Young J. N., Rickaby R. E. M., Kapralov M. V. and Filatov D. A. (2012) Adaptive signals in algal Rubisco reveal a history of ancient atmospheric carbon dioxide. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 483–492.

Young J. R., Geisen M., Cros L., Kleijne A., Sprengel C., Probert I. and Østergaard J. (2003) A guide to extant coccolithophore taxonomy. *J. Nannoplankt. Res.*, 125.

Zeebe R. E., Ridgwell A. J. and Zachos J. C. (2016) Anthropogenic carbon release rate unprecedented during the past 66 million years. *Nat. Geosci.* **9**, 325–329.

Zeebe R. E. and Wolf-Gladrow D. A. (2001) *CO2 in Seawater: Equilibrium, Kinetics, Isotopes*. 1st ed., Elsevier.

Zhang Y. G., Pagani M., Henderiks J. and Ren H. (2017) A long history of equatorial deep-water upwelling in the Pacific Ocean. *Earth Planet. Sci. Lett.* **467**, 1–9.

Zhang Y. G., Pagani M., Liu Z., Bohaty S. M. and DeConto R. (2013) A 40-million-year history of atmospheric CO2. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **371**, 20130096.

Zhang Y. G., Pearson A., Benthien A., Dong L., Huybers P., Liu X. and Pagani M. (2019) Refining the alkenone-pCO2 method I: Lessons from the Quaternary glacial cycles. *Geochim. Cosmochim. Acta* **260**, 177–191.

Chapter 3: Controls on alkenone carbon isotope fractionation in the natural environment and revisions to the alkenone-pCO_2 proxy

3.1 Abstract

Carbon isotope discrimination in alkenones (ε_p) is one of the primary proxies for past atmospheric CO_2 levels, but it has yielded inconsistent results during periods where CO_2 variations are known. Recent syntheses of algal cultures have quantitatively demonstrated that ε_p indeed records CO_2 information: ε_p increases as aqueous CO_2 concentrations increase relative to cellular carbon demand. However, this relationship is complicated by irradiance, which also yields larger ε_p values. Here we examine the roles of physiology and environment in setting alkenone ε_p in the modern ocean. We compile surface water particulate and core-top alkenone ε_p data and add new core-top measurements. To constrain cellular carbon demand, we generate new estimates of cell sizes and growth rates of the alkenone-producing population, though cell size is a much more important control on ε_p than growth rate. Supporting culture studies, we find irradiance to be a key control on alkenone ε_p in the modern ocean. We test recent culture-derived multiple linear regression models on particulate and sediment samples and find that relationships from cultures can be used to predict alkenone ε_p in sediments. In particulate samples, ε_p is substantially overestimated, largely resulting from higher irradiance in the surface where samples were collected than the integrated light conditions under which the sampled biomass was produced. Our results demonstrate that quantitative relationships derived from laboratory experiments have predictive skill in the natural environment. We further argue that, with an expanded dataset, the theory underpinning the conventional diffusive alkenone carbon isotope

fractionation model, including the '*b*' parameter, is not supported by the field data and should not be used to reconstruct past variations in CO_2. Future estimates of CO_2 from alkenone ε_p should use empirical or mechanistic models to quantitatively account for irradiance and cell size variations.

3.2 Introduction

Understanding the role of CO_2 as both a feedback on, and forcing of, global climate is essential for better predicting how present and future anthropogenic CO_2 emissions will shape the planet. Carbon isotope fractionation ($\delta^{13}C$) in photosynthetic organic matter has long been used to estimate ancient atmospheric CO_2 concentrations (Popp et al., 1989). The vast majority of carbon isotope-based CO_2 records from marine sediments derive from long-chain alkenones, unsaturated ketones produced by a specific family of coccolithophorid algae, the Noelaerhabdaceae, whose modern forms include the widespread genera *Gephyrocapsa* and *Emiliania*, as well as the less abundant *Reticulofenestra* (Bendif et al., 2016). Because of the constant isotopic offset between alkenone $\delta^{13}C$ and bulk cellular organic matter (Popp et al., 1998a; Riebesell et al., 2000b; Wilkes et al., 2018; Chapter 2), these molecules are faithful recorders of the carbon isotope discrimination by photosynthesis (ε_p) in these algae.

Past field observations have shown that ε_p is related to the growth environment, including ambient CO_2 concentrations, temperature, and irradiance, as well as physiological factors such as cellular growth rate and cell size (Fry and Wainwright, 1991; Rau et al., 1992; Francois et al., 1993; Laws et al., 1995; Popp et al., 1998b). Decades of culture work identified the role of these processes in modifying ε_p values in alkenone-producing algae. Recent syntheses of culture data have elucidated the quantitative relationships between various physiological and environmental

influences, and demonstrated that ε_p does indeed decline as [$CO_{2(aq)}$] decreases (Stoll et al., 2019; Chapter 2). However, the relationship between ε_p and carbon demand relative to CO_2 supply is nonlinear: the slope declines as demand relative to supply increases, making ε_p less sensitive to changes in CO_2 at lower CO_2 values, higher growth rates, and larger cell sizes. Furthermore, culture experiments suggest light and nutrient conditions are nearly as important controls on ε_p as physiology and CO_2, with higher ε_p occurring at higher irradiance (Rost et al., 2002; Wilkes and Pearson, 2019; Chapter 2). Though alkenone ε_p has been extensively used to generate paleo-CO_2 records, only one study has quantitatively considered cell size, growth rate, and irradiance when interpreting alkenone ε_p in sediment sequences (Stoll et al., 2019). Whether the same key factors observed in culture experiments control alkenone ε_p in the modern ocean remains an open question.

Here we focus on the modern and late Holocene ocean and conduct an empirical analysis to test whether the same relationships identified in culture also control alkenone ε_p in the natural environment. We explore this question with a careful compilation of published alkenone $\delta^{13}C$ measurements from modern suspended particulate organic matter samples and from late Holocene core-tops, and add new alkenone $\delta^{13}C$ measurements and coccolith size measurements from 22 core-top sediments. We estimate growth rates of the alkenone-producing Noelaerhabdaceae using Michaelis-Menten parameterizations, and compile published coccolith assemblage and size data to estimate cellular carbon demand at each sample location. We first examine the relationship between alkenone ε_p and carbon demand and carbon supply, then the influence of irradiance and nutrients on ε_p. We test the theory of the conventional diffusive alkenone ε_p model using growth rate data, and finally evaluate whether multiple linear regression models of ε_p developed from cultures are applicable to the ocean.

3.3 Methods and data sources

The goal of this work is to determine the controls on alkenone carbon isotope fractionation in the natural environment. We compile a rich body of existing literature and update and standardize all calculations to interpret our new data and all published data in a common framework.

3.3.1 Sample origins and locations

Our sample set can most broadly be divided into "particulate" or "sediment." Particulate samples, taken from water column suspended particulate organic matter, have all been previously reported (Fig. 1) (Bidigare et al., 1997; Bidigare et al., 1999; Eek et al., 1999; Laws et al., 2001; Harada et al., 2003; Schulte et al., 2003; Prahl et al., 2005; Benthien et al., 2007; Tolosa et al., 2008; Wallsgrove, 2008; Wolhowe et al., 2014; Wolhowe et al., 2015). New core-top sediment samples (n = 22) come from the Lamont-Doherty Earth Observatory Core Repository (LDEO Repository). Several of these samples were obtained during the R/V Oceanus Cruise 437-7 along the NW African margin (McGee et al., 2013). The remainder come from various sediment cores in the LDEO Repository. We exclude the particulate data from three datasets (Harada et al., 2003; Schulte et al., 2003; Tolosa et al., 2008) as well as the sediment trap results of Yamamoto et al. (2007) because they did not measure $\delta^{13}C_{DIC}$ during alkenone sampling, prohibiting a proper calculation of ε_p.

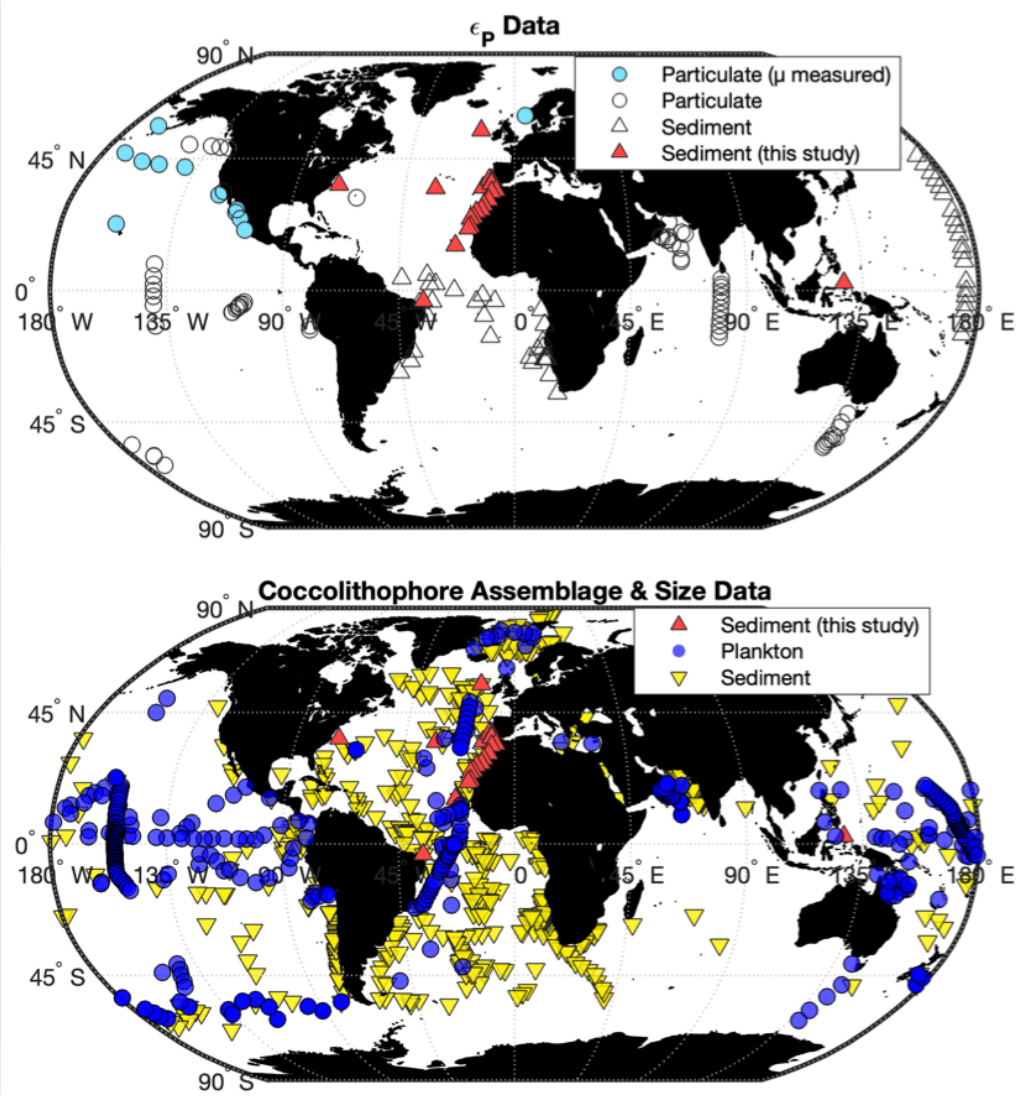

Figure 1. Locations of samples used in this study. *Top panel*: particulate and sediment sample locations sourcing ε_p data. Filled aqua circles are "training set" particulate samples where growth rate (μ) was measured in-situ; filled red triangles are new surface sediments; open triangles and circles are previously reported surface sediment and particulate data, respectively. *Bottom panel*: coccolithophore assemblage and size data mined from the literature, as well as new sediment samples (red, same as above).

3.3.2 Oceanographic data

Relevant oceanographic data for this analysis include temperature, salinity, nutrients, aqueous CO_2 ($[CO_{2(aq)}]$), mixed layer depth, surface irradiance (photosynthetically available radiation, PAR), and the diffuse attenuation coefficient of PAR (Kd_{490}). In cases where this information is not reported in the original publication or primary cruise documents, we estimate these parameters from gridded climatologies. In MATLAB R2019a we use the interp3 function for three-dimensional interpolation or interp2 function for two-dimensional interpolation to estimate parameter values at each sample collection latitude, longitude, and depth by natural neighbor interpolation. For samples with an exact collection date reported, we linearly interpolate between monthly climatological averages to the collection day, with day of year as the time axis and assuming the monthly climatology is reported for the middle of the month. For sediment samples, we take the mean annual average of the monthly climatologies of each oceanographic parameter. While there can be a strong seasonality to algal production in the mixed layer, this signal is greatly attenuated by enhanced remineralization in the subsurface such that deep ocean particulate flux and surface sediments more closely track mean annual conditions (Rosell-Melé and Prahl, 2013).

3.3.2.1 Irradiance and photoperiod

We calculated irradiance at the collection depth by:

$$PAR_z = PAR_{surface} \times e^{-(Kd_{490} \times z)} \quad \text{[Eq. 1]}$$

where z = depth (meters, m), $PAR_{surface}$ is the incident photosynthetically available radiation at the ocean surface (mol photons m^{-2} d^{-1}), and Kd_{490} is the diffuse attenuation coefficient of light at 490nm (m^{-1}) (Falkowski and Raven, 2007). Surface irradiance and Kd_{490} at each location were obtained averaging NASA SeaWiFS and MODIS Aqua data (NASA Goddard Space Flight Center, Ocean Biology Processing Group, 2018a-d).

We calculated the hours of daylight at each collection location and date following the equations outlined in the NOAA Sunrise/Sunset and Solar Position Calculators (https://www.esrl.noaa.gov/gmd/grad/solcalc/calcdetails.html). For sediment samples, the alkenone unsaturation index ($U^{K'}_{37}$) has been shown to best correlate with mean annual sea surface temperatures, though $U^{K'}_{37}$ is explained nearly as well with 0-20 m mean temperatures (Müller et al., 1998). Studies of alkenone fluxes out of the photic zone have shown that the $U^{K'}_{37}$ signal reaching the sediments is best described globally by mean annual conditions (Rosell-Melé and Prahl, 2013). We therefore assume all sediment alkenone $\delta^{13}C$ values reflect mean annual production, and employ a photoperiod of 12.05 hours for all sediment samples.

3.3.2.2 Collection depth

The collection depth of all particulate samples was reported in their original publications. For sediment samples, we set the collection depth as 75% of the mixed layer depth (MLD). We estimate the depth of the mixed layer at each sample location from the gridded dataset of de Boyer Montégut et al. (2004). We use their fixed-threshold temperature criterion for MLD calculation, which determines the base of the mixed layer as the depth of a 0.2°C change from the temperature at 10 m depth. This method therefore assumes the minimum mixed layer depth is

10 m, because the upper few meters of the surface ocean are generally subject to significant daily density variations (Schneider and Müller, 1990).

3.3.2.3 Carbonate system and [$CO_{2(aq)}$]

In the particulate samples, all studies (except Schulte et al. (2003)) measured [$CO_{2(aq)}$] while underway, typically through simultaneous measurement of DIC and total alkalinity. For sediment samples, we must estimate the preindustrial surface ocean [$CO_{2(aq)}$]. We do this by assuming surface ocean mean annual alkalinity has remained constant over the last several thousand years, and speciate the carbonate system with estimated preindustrial total CO_2 (PI_TCO2) from the GLODAPv2.2016b mapped climatology (Key et al., 2015; Olsen et al., 2016). When speciating the carbonate system, we use the temperature, salinity, and nutrient data of the gridded GLODAPv2.2016b, rather than WOA18, because the GLODAP dataset is internally consistent. We use the CO2SYS MATLAB program to speciate the carbonate system, using the K_1 and K_2 constants of (Lueker et al., 2000), K_{SO4} of Dickson (Dickson, 1990), and total boron of Uppström (Uppström, 1974).

3.3.3 Alkenone ε_p and $\delta^{13}C_{DIC}$

Determining alkenone ε_p requires carbon isotope measurements of the $C_{37:2}$ methyl ketone ($\delta^{13}C_{Mk37:2}$), the carbon isotope ratio of dissolved inorganic carbon in seawater at the collection depth ($\delta^{13}C_{DIC}$), and temperature, to calculate the carbon isotope ratio of dissolved CO_2 ($\delta^{13}C_{CO2aq}$). For particulate samples, we use the $\delta^{13}C_{DIC}$ measured at the time of sample collection. For sediment samples, we must estimate the pre-industrial surface ocean $\delta^{13}C_{DIC}$. Because anthropogenic CO_2 has been preferentially adding $^{12}CO_2$ to the ocean since the

industrial revolution (Keeling, 1979), we cannot use modern values. Instead, we use the GLODAPv2.2019 bottle dataset to first estimate modern $\delta^{13}C_{DIC}$ values at each sample location (Olsen et al., 2019). We then estimate the anthropogenic contribution to $\delta^{13}C_{DIC}$ using the gridded output of the isotope enabled biogeochemical model of (Schmittner et al., 2013), which modeled the anthropogenic contribution to $\delta^{13}C_{DIC}$ throughout the global ocean. We subtract the $\delta^{13}C_{DIC\_ANT}$ from modern $\delta^{13}C_{DIC}$ at each sample location, yielding preindustrial $\delta^{13}C_{DIC}$ values. We use this method to estimate $\delta^{13}C_{DIC}$ and recalculate ε_p for the core-top datasets of Andersen et al. (1999), Benthien et al. (2002) and Pagani et al. (2002), putting all sediment ε_p data in a common framework.

3.3.3.1 Calculation of ε_p

We calculate ε_p following the definition of Freeman and Hayes (1992):

$$\varepsilon_p = 1000 \times [(1000 + \delta^{13}C_{CO2aq}) / (1000 + \delta^{13}C_{POC}) - 1] \quad \text{[Eq. 1]}$$

where $\delta^{13}C_{CO2aq}$ is the carbon stable isotope ratio of ambient aqueous CO_2 (in ‰, VPDB), and $\delta^{13}C_{POC}$ is the carbon stable isotope ratio of algal biomass.

$\delta^{13}C_{CO2aq}$ can be determined using the temperature-dependent fractionation between DIC and dissolved CO_2 of Rau et al. (1996), which is derived from Mook et al. (1974) and Freeman and Hayes (1992):

$$\delta^{13}C_{CO2aq} = \delta^{13}C_{DIC} + 23.644 - (9701.5 / T_K) \quad \text{[Eq. 2]}$$

where T_K is temperature in Kelvin. For all samples where $\delta^{13}C_{DIC}$ is reported, we recalculate $\delta^{13}C_{CO2aq}$ using this equation, reducing any systematic bias in the $\delta^{13}C_{CO2aq}$ data due to preexisting calculation differences. In samples where $\delta^{13}C_{DIC}$ is not reported but $\delta^{13}C_{CO2aq}$ is, we use the reported $\delta^{13}C_{CO2aq}$ value. The vast majority (97%) of shipboard particulate sampling studies reported SSTs from simultaneous CTD casts. For sediment samples, we use alkenone unsaturation ratios and the calibration of (Müller et al., 1998) to translate $U^{K'}_{37}$ to SST.

The $C_{37:2}$ methyl ketone has a constant isotopic offset from bulk biomass ($\varepsilon_{bio/alk}$) of 4.44 ± 0.05 ‰ (1σ) (Popp et al., 1998a; Riebesell et al., 2000b; Wilkes et al., 2018; Chapter 2). The carbon isotope ratio of biomass is calculated after Hayes (1993) using the equation:

$$\varepsilon_{bio/alk} = 4.44 = [(1000 + \delta^{13}C_{POC} / (1000 + \delta^{13}C_{Mk37:2}) - 1] \times 1000 \qquad \text{[Eq. 3]}$$

To standardize all ε_p data, we recalculate ε_p in published samples using $\varepsilon_{bio/alk}$ of 4.44‰ (Chapter 2). In a few samples, only ε_p values and no $\delta^{13}C_{Mk37:2}$ or $\delta^{13}C_{DIC}$ or $\delta^{13}C_{CO2aq}$ values are reported, preventing a proper recalculation of ε_p. In these cases, we subtract 0.44‰ from these ε_p measurements if $\varepsilon_{bio/alk}$ was reported as 4.0‰, and subtract 0.64‰ if $\varepsilon_{bio/alk}$ was reported as 3.8‰. Because ε_p is calculated as a difference in isotope ratios relative to a standard instead of a simple arithmetic difference, our adjustment may be imprecise by up to 0.1‰, which does not impact our findings.

3.3.4 Coccolithophore distributions and coccolith size compilation

Here we generate new estimates of Noelaerhabdaceae coccolith size in the global ocean. Because coccolith size is well correlated to cell radius in the alkenone-producing

Noelaerhabdaceae (Henderiks and Pagani, 2007; McClelland et al., 2016), these data allow us to evaluate the influence of cell size on the spatial variation of ε_p in the modern ocean. Our algorithm is described in detail in Appendix B. Briefly, we compile coccolith size and abundance data from published core-top and water column surveys (Fig. 1B), and using the mean coccolith size defining each species, we estimate a biomass-weighted ("BMW") coccolith length for the alkenone-producing population at every sample location in Fig. 1B. The details of this biomass-weighting approach are also described in Appendix B, Section 1. We then estimate the coccolith length at each alkenone ε_p sample location (Fig. 1A) by taking the inverse-distance weighted mean of all coccolith length estimates within 5° of the alkenone ε_p sample. Finally, we convert coccolith length to cell radius using well-defined linear functions relating Noelaerhabdaceae coccolith morphometry to cell dimensions (McClelland et al., 2016).

3.3.5 Noelaerhabdaceae growth rate modeling

We estimate Noelaerhabdaceae growth rates in all samples using a modified approach of Krumhardt et al. (2017), referred to henceforth as "K17". This method models coccolithophore growth rates using the temperature-dependent maximum growth rates determined from an analysis of all available *Emiliania huxleyi* culture data (Fielding, 2013). The realized growth rate is modified by growth-limiting parameters following Michaelis-Menten uptake kinetics and using published substrate half-saturation (K_M) constants for *Emiliania huxleyi*, which are calculated from fitting curves of the growth rate response to substrate concentrations. We include the influence of nitrate, phosphate, CO_2, and irradiance as growth substrates in our model. The equations used to calculate growth rate can be found in Appendix B, Section 2. We evaluate the skill of this model using a subset of particulate samples where growth rates of alkenone-

producing algae were measured either by cell counts Benthien et al. (2007), or during in-situ incubations by monitoring the uptake of isotopically-labeled DIC into alkenones (Prahl et al., 2005; Wallsgrove, 2008; Wolhowe et al., 2014).

3.3.6 Laboratory methods

All biomarker analysis was conducted at the Lamont-Doherty Earth Observatory of Columbia University (LDEO; New York, USA). All nannofossil measurement work was conducted at the Centre Européen de Recherche et d'Enseignement des Géosciences et de l'Environnement (CEREGE; Aix-en-Provence, France). The details of lipid extraction, isotope measurement, and nannofossil slide preparation and size analyses are described in detail in Section 3 of Appendix B and are summarized briefly here.

Bulk sediment samples were freeze dried and free lipids were extracted using a DIONEX ASE 350. Alkenones were purified from the total lipid extract using silica gel flash column chromatography and DCM as the eluent. The DCM fraction containing the alkenones was base-hydrolyzed to remove interfering long-chain fatty-acid methyl esters. Alkenone unsaturation ratios were measured by gas chromatography with flame ionization detection, which were converted to temperature using the calibration of Müller et al. (1998). Alkenone carbon isotope ratios were measured by gas-chromatograph isotope ratio mass spectrometry (GC-irMS), with an average error of 0.31‰, which includes realizing sample delta values on the VPDB scale (Polissar and D'Andrea, 2014).

Coccolith size measurements were made by automated light microscopy (SYRACO). Bulk sediment was sieved at 25 μm in ~pH 9 deionized water with dilute ammonium hydroxide. At least two coverslips for each sample were prepared by random settling of the <25μm fraction.

On a transmission light microscope with circularly-polarized and inverse circularly-polarized light, between 165 and 297 images of each coverslip were automatically acquired. Carbonate particles were segmented from the black background, and nannofossils were identified and grouped roughly at the genus level using a neural network that was trained on tens of thousands of Cenozoic coccoliths. Morphometric parameters, such as coccolith length, were determined by the segmentation and measurement program. We further manually vetted all coccolith images that were programmatically identified to ensure that coccolith particles were properly classified, and removed those that were incorrectly identified.

3.3.7 Carbon demand relative to diffusive CO_2 supply (τ)

Laboratory cultures have demonstrated that the expression of RubisCO's preference for ^{12}C over ^{13}C decreases as the ratio of cellular carbon demand to $[CO_{2(aq)}]$ increases (Laws et al., 1995; Popp et al., 1998b). A convenient way to describe the ratio of carbon demand to CO_2 supply is the dimensionless unit τ (McClelland et al., 2017):

$$\tau = \frac{r \rho \mu_i}{3 P_C C_e} \qquad \text{[Eq. 4]}$$

where r is the cell radius (m), μ_i is the instantaneous (photoperiod-adjusted) growth rate (Rost et al., 2002) (Rost et al., 2002), ρ is the cellular carbon density (pmol C m^{-3}), C_e is the ambient aqueous CO_2 concentration (mol C m^{-3}), P_C is the cellular permeability to the diffusion of aqueous CO_2 (m s^{-1}). This term is equivalent to other demand/supply ratios in the literature. Cell permeability to CO_2 diffusion is not well known and is an ongoing area of research, but we apply a value of $2.0e^{-5}$ m s^{-1} so τ values used here are comparable to those in recent culture work

(Chapter 2). We treat it as a constant in all of our ensuing analysis, because measurements of membrane permeability to CO_2 in diatoms have shown to be relatively constant in four species across a range of growth environments (Hopkinson et al., 2011).

3.4 Results

3.4.1 Growth rates of alkenone-producing algae

To test and refine the Krumhardt et al. (2017) growth rate model, we take advantage of one mesocosm experiment where daily cell counts were made (Benthien et al., 2007), as well as four particulate alkenone $\delta^{13}C$ studies that measured growth rates through the incorporation of isotopically-labeled DIC in controlled in-situ incubations (Prahl et al., 2005; Popp et al., 2006; Wallsgrove, 2008; Wolhowe et al., 2014). These data constitute the best-controlled sample set to investigate the relationship between ε_p and μ, and serve as the primary training set estimating growth rates of natural Noelaerhabdaceae populations.

Using the exact equations and approach of Krumhardt et al. (2017), we substantially overestimate growth rates of alkenone producers. This systematic error likely results from the low irradiance in the training set and the lack of irradiance in the original K17 model. Many of the samples in our training set come from depths below the surface where irradiance drops significantly and growth is light-limited. Roughly one-third of the training set is below the 5% surface irradiance level (average ~2.4 mol photons m^{-2} d^{-1}, or ~15 μmol m^{-2} s^{-1} at these photoperiods). For comparison, these daily integrated and instantaneous irradiance values are at the very low end of the range in culture experiments (Rost et al., 2002).

We find the optimum growth rate model requires nitrate, phosphate, and light limitation on the temperature-dependent maximum growth rate, and that the best fit is achieved with

maximum half-saturation constants. Importantly, pCO_2 is not the growth-limiting parameter in any of the samples in the training dataset or full compilation (Fig. 2A). Krumhardt et al. (2017) found pCO_2 to be growth-limiting only at very low ambient CO_2 levels (< 100ppm), which we do not encounter in the modern ocean. The lack of pCO_2 limitation may be due to ubiquity of carbon-concentrating mechanisms to acquire DIC from the natural environment (Raven and Johnston, 1991).

The growth rate model performs reasonably well, with a correlation coefficient 0.55 and a root-mean-squared error of 0.16 d^{-1}. Figure 2 shows the relationship between measured and modeled growth rate, as well as the dependence of growth rate on individual parameters in the model. In this framework, the vast majority of samples are light-limited (yellow symbols, Fig. 2A). It is instead possible that this results from overestimating the K_M value for light, or the measured K_M values for nitrate and phosphate are too low. The model tends to generally overestimate growth rate, particularly at low measured growth rates. This overestimation could be explained in several ways. It is possible that the half-saturation constants determined in culture experiments are underestimates of realized nutrient limitation in the ocean, possibly due to competition with other algae for nutrients, the episodic and non-uniform nature of nutrient delivery to the surface ocean (Arístegui et al., 2009; Calil et al., 2011; Macias et al., 2012), pressures from predation, or lower effective nutrient concentrations due to scavenging by particles in the water column (Clegg and Sarmiento, 1989; Bruland et al., 2014). None of these factors are incorporated in this model. Our model also does not incorporate key vitamins or trace metals (e.g. iron) that are essential for phytoplankton growth, but for which we do not have robust constraints on the catalytics or concentrations at our sample locations in the field. Regardless, as discussed in Section 4.4, growth rate uncertainty does not dramatically affect our

findings because ε_p is rather insensitive to growth rate in the culture-derived models (Chapter 2), with coefficients of -0.18‰ to -0.33‰ d^{-1}. This suggests growth rate need not be particularly well constrained to understand ε_p variations in space and time.

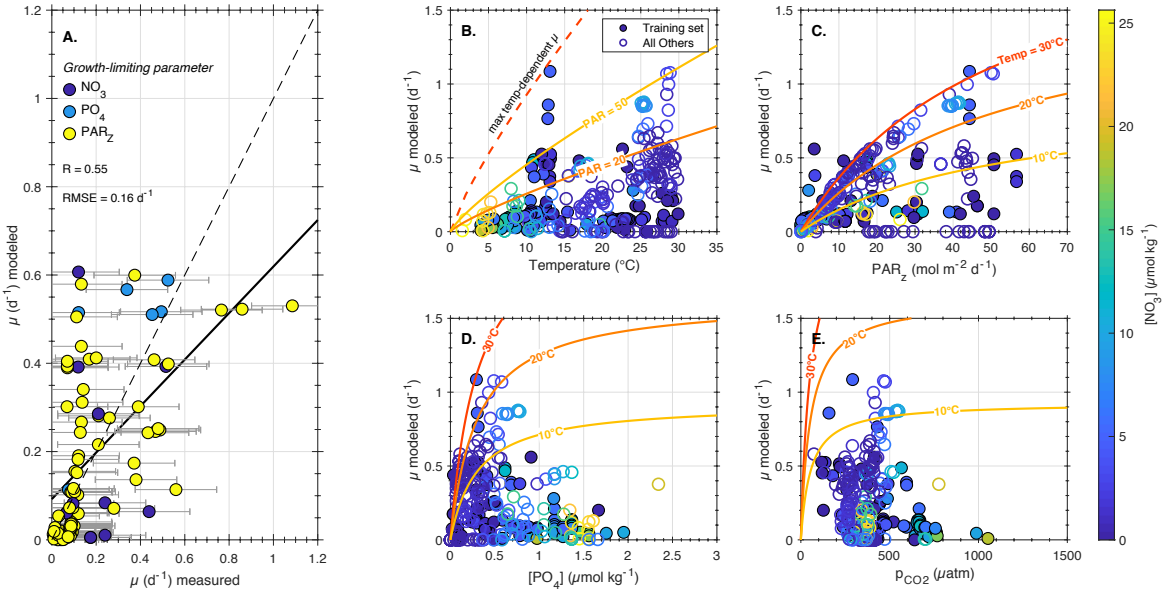

Figure 2. Measured and modeled growth rates of alkenone-producing algae. (A) Measured and modeled growth rates from "training set" samples where growth rates were determined in-situ by cell counts or isotope-labeled incubations. Samples are colored by the growth-limiting parameter (see Supplemental Material for details). In (A), solid black line is the least-squares fit, black dashed line is the 1:1 line. (B–E) Growth rates of alkenone-producing algae in the global particulate and sediment dataset with respect to the growth-influencing parameters in the model. Filled symbols are the training set shown in (A), and measured growth rates are plotted in (B–E). Empty symbols are the remainder of the ε_p particulate and sediment dataset where growth rates are estimated using the growth rate model described in Appendix B.

3.4.2 Noelaerhabdaceae coccolith size and cell radius

In our core-top samples, biomass-weighted (BMW) coccolith length ranged from 3.46 to 5.05 µm (mean = 4.12 µm ± 0.58, 1σ). In this subset, we do not find a statistically significant

linear correlation (defined as $p < 0.05$) between biomass-weighted mean coccolith size and ε_p or any of the components of the growth rate model (temperature, irradiance, macronutrients, pCO_2). When we combine our new data with size estimates in the full alkenone compilation (n = 253), coccolith length ranges from 2.58 to 5.05 μm (mean = 3.56 μm ± 0.42, 1σ). Cell radius is a linear translation of coccolith length, and ranges from 1.92 μm to 3.33 μm (mean = 2.48 μm ± 0.24, 1σ). Mean cell radius in sediment samples is 2.66 μm ± 0.26 (1σ) and 2.42 μm ± 0.20 (1σ) in the particulates, which are statistically different by a Kolmogorov-Smirnov test.

There is no statistically significant correlation between BMW coccolith length and estimated growth rate. We observe weak but significant inverse correlations between BMW coccolith length and phosphate, nitrate, pCO_2, irradiance, and ε_p, and a weak positive correlation with temperature (Fig. 3). Though quite scattered, the sign of the relationship between ε_p and cell radius is the same in marine samples as in culture experiments: larger cells have, on average, smaller ε_p (Fig. 3F; slope = -0.83 ‰ μm^{-1}, $R = -0.16$, $p < 0.05$).

In our combined alkenone dataset, we observe both large and small coccolith sizes at low irradiance, but a trend toward smaller coccoliths as irradiance increases ($R = -0.25$, $p < 0.001$). This relationship contrasts with that in planktonic foraminifera, where irradiance is an important driver of larger test sizes (Spero and Lea, 1993; Lombard et al., 2010). In coccolithophore populations, cell size is most closely associated with species composition, thus the observed trends may reflect top down (ecosystem) rather than bottom up (resource availability) controls. It is possible therefore that Noelaerhabdaceae size is controlled by more complex competitive interactions, particularly in locations with strong seasonal cycles. For example, although *E. huxleyi* is the dominant species around the world, different regions are occupied by genetically distinct strains (Iglesias-Rodriguez et al., 2006), thus there may be niche partitioning related to

the presence and absence of other phytoplankton groups, or micronutrient availability governing which strain or species dominates, both of which we do not have the data to test.

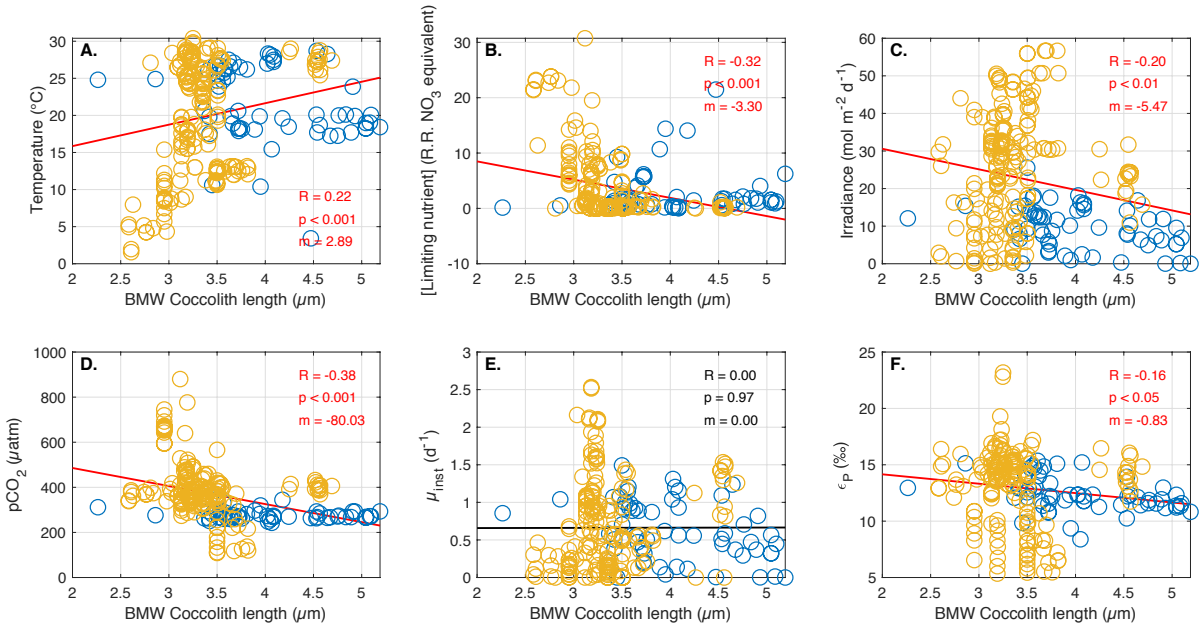

Figure 3. Biomass-weighted coccolith lengths in the alkenone ε_p sample set. Blue symbols in all panels are sediment samples, and orange symbols are particulate samples. Coccolith length regressed against: (A) temperature; (B) the concentration of the limiting nutrient, adjusted to nitrate equivalence by the Redfield Ratio; (C) irradiance at the collection or production depth; (D) in-situ pCO_2; (E) instantaneous growth rate; (F) ε_p. Regressions are shown as red lines if the slope is significant at the $p < 0.05$ confidence interval; otherwise they are shown in black.

3.4.3 Alkenone ε_p and carbon demand relative to CO_2 supply

In our new sediment samples, alkenone ε_p ranges from 9.8‰ to 13‰, which is about half the modern sediment range (~8‰ to 15‰). The ε_p range in particulate samples is about twice as large that in sediments, with a minimum of 5.2‰ and a maximum of 19.5‰ (Fig. 4). The larger variability in particulate ε_p likely results from the sampling of depths and environmental conditions that do not contribute substantially to the sedimentary record.

In the sediment sample set, [$CO_{2(aq)}$] ranges from to 6.5 to 13.5 μmol kg^{-1}, in particulates the range is 4.3 μmol kg^{-1} to 40 μmol kg^{-1}. The higher maximum [$CO_{2(aq)}$] results from sampling higher latitude sites with lower temperatures and higher $CO_{2(aq)}$ solubilities, as well as sampling at depths lower in the photic zone, where respiration elevates $CO_{2(aq)}$ concentrations relative to the surface. When the larger dataset is considered, there is no universal relationship between ε_p and 1/[$CO_{2(aq)}$] (Fig. 4). Culture observations demonstrate an overall decrease in ε_p as 1/[$CO_{2(aq)}$] increases; sediment samples, however, have a significant *positive* relationship ($r^2 = 0.11$, $p < 0.01$) with 1/[$CO_{2(aq)}$] (Fig. 4C). The theory underpinning the conventional diffusive alkenone CO_2 proxy (Bidigare et al., 1997) implies large ε_p values should exist at low 1/[$CO_{2(aq)}$] values, which is not the case in the sediment data (Fig. 4C). Underlying this correlation in sediment samples is a clear relationship between 1/[$CO_{2(aq)}$] and the ratio of irradiance to nutrients (see Section 4.1 for details on this metric), which increases as 1/[$CO_{2(aq)}$] increases (Fig. S1). The large positive effect of higher irradiance on ε_p (discussed below) appears to largely offset and mask any effect of 1/[$CO_{2(aq)}$] on ε_p in the data.

We also examine trends in ε_p with respect to τ because it is the cellular carbon demand relative to ambient CO_2 supply, rather than CO_2 supply alone, that has been shown to influence ε_p in cultures (Laws et al., 1995; Popp et al., 1998b; Burkhardt et al., 1999). The τ range in sediments (~0.1 to ~2.15) is larger than that in particulate samples (~0.01 to ~1.6), resulting from the combination of lower pre-industrial CO_2 concentrations and relatively high modeled growth rates due to high temperatures and high irradiance in tropical locations. The τ range in the modern ocean is about 1/3 of the range achieved in Noelaerhabdaceae culture experiments (Chapter 2). Between τ values of 0 and 2 in cultures, ε_p spans nearly a 15‰ range, depending on the nutrient and light conditions. This phenomenon is also apparent in particulate samples, with

higher irradiance yielding higher ε_p at low τ values (Fig. 4B). In the particulates, two populations are apparent: one with generally high irradiance and a mean ε_p of ~15‰, and a second population with lower irradiance, in which ε_p is below ~11‰ and declines with increasing τ. However, in sediment samples, ε_p increases with increasing τ (Fig. 4D; $p < 0.01$), in contrast to the relationship observed in cultures. Largely reinforcing the trend seen in Fig. 4C, higher irradiance promotes higher growth rates and larger τ values in the sediment samples studied here.

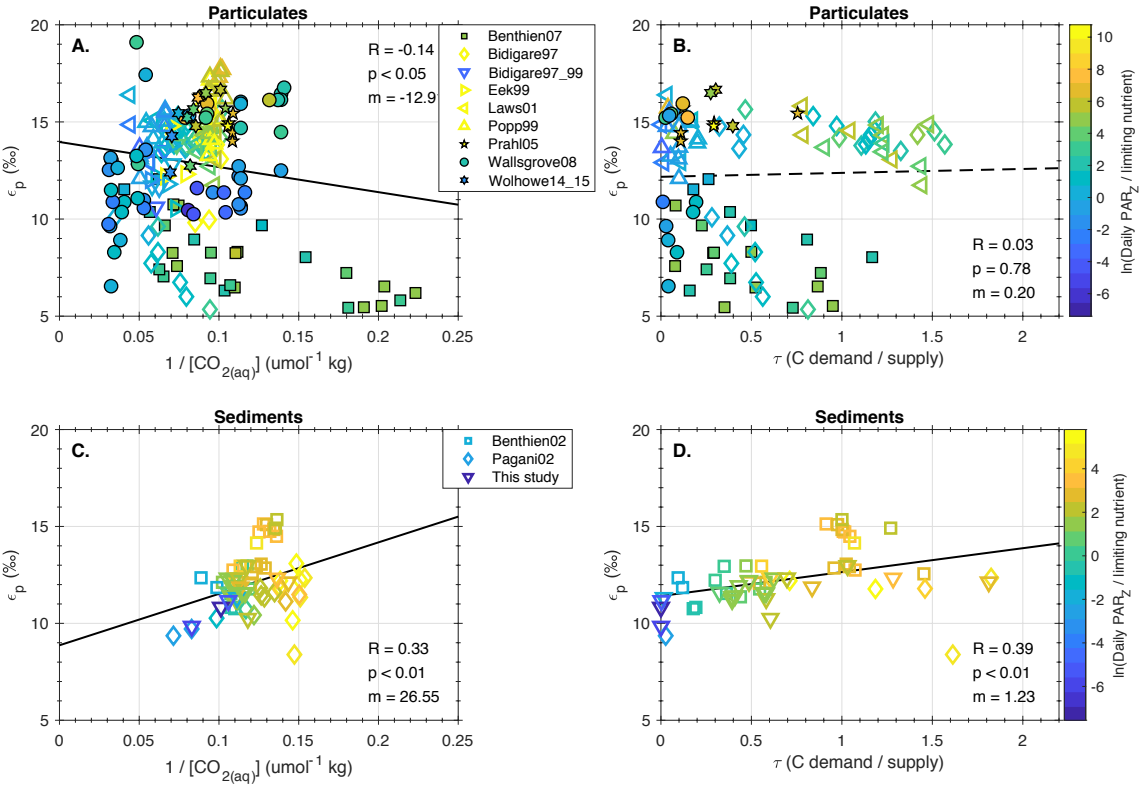

Figure 4. Alkenone carbon isotope discrimination in the natural environment. Carbon isotope fractionation in alkenones (ε_p) as a function of $1/[CO_{2(aq)}]$ (A, C) and τ (B, D), cellular carbon demand divided by CO_2 supply. Filled symbols in A and B are samples with measured growth rates, all other growth rates are modeled using the model described in Section 2.5. There are fewer data in panels B and D than A and C because cell size could not be estimated in all samples. Black lines are linear regressions to all data in each panel. Solid lines have statistically significant slopes at $p < 0.05$.

3.5 Discussion

3.5.1 Controls on alkenone ε_p in the ocean

Culture studies have shown that irradiance, environmental conditions (e.g. batch or chemostat cultures), and algal carbon demand relative to diffusive CO_2 supply (τ) are important determinants of coccolithophore carbon isotope fractionation, ε_p (Popp et al., 1998b; Riebesell et al., 2000b; Rost et al., 2002; McClelland et al., 2017). Here we examine whether these same factors are primary controls on ε_p in the ocean.

We expect from culture experiments to see an inverse proportionality between ε_p and $1/[CO_{2(aq)}]$ (Laws et al., 1995; Bidigare et al., 1997). Instead, there is a statistically significant positive correlation between ε_p and $1/[CO_{2(aq)}]$ in sediment samples, where higher ε_p is found at lower $CO_{2(aq)}$ concentrations (Fig. 4D). The absence of the expected relationship likely arises because $[CO_{2(aq)}]$ and integrated irradiance—which has a coefficient roughly equal to $[CO_{2(aq)}]$ in the empirical culture-based multilinear regression models (Chapter 2)— are significantly inversely proportional ($R = -0.73$, $p < 0.001$). The combination of high irradiance and high $[CO_{2(aq)}]$ does not exist (Fig. S1). Because of the temperature-dependent solubility of CO_2, higher $[CO_{2(aq)}]$ generally occurs at higher latitudes, where surface irradiance is generally lower, resulting from greater seasonality and a more oblique angle of incidence for incoming solar rays. The low irradiance at these locations would reduce ε_p, compensating for the effect of higher CO_2.

Chapter 2 showed from culture experiments of Noelaerhabdaceae coccolithophores that the ratio of irradiance to nutrients, in addition to τ, is an important factor controlling ε_p. Previous investigators have noted differences between ε_p from chemostat cultures and batch cultures of both diatoms and coccolithophores (Riebesell et al., 2000a; Pagani, 2014), suggesting the interplay of light energy and nutrient availability modulate carbon acquisition and isotope

fractionation (Wilkes and Pearson, 2019). Here we calculate an "energy quotient" (Eq. 5) as the natural log of the ratio of integrated daily irradiance (mol photons m^{-2} d^{-1}) to the Redfield ratio-adjusted macronutrient concentration. We use the natural log to transform the exponential decay in irradiance from surface to depth. For each sample, we multiply the phosphate concentration by 16 to make it numerically equivalent to the nitrate concentration according to the Redfield ratio, and then take the minimum of $\left[PO_4^{3-}\right]\times 16$ or $[NO_3^-]$ (both in μmol kg^{-1}) to use as the divisor in our energy quotient.

$$\text{energy quotient} = \frac{\ln(\text{daily PAR}_Z)}{\min(PO_4^{3-} \times 16,\ NO_3^-)} \qquad \text{[Eq. 5]}$$

Samples with limiting nutrient concentrations below the detection limit (i.e. values of zero) produce infinite energy quotients; we set these values to the maximum of the calculable energy quotient in the dataset. When we regress ε_p vs. this energy quotient (Fig. 5), we find a significant relationship in both the particulate and sediment samples, where higher ε_p is found at higher irradiance relative to nutrient concentrations ($R = 0.34$, $p < 0.001$). Although this parameter does not explain a majority of the variance in ε_p, the finding is in direct agreement with the results from culture data, in which higher integrated irradiance yields higher ε_p values (Chapter 2). A comparison of Figures 4 and 5 suggests that in the natural environment, oceanographic conditions are an even more important predictor of ε_p than the physiological components encapsulated in τ.

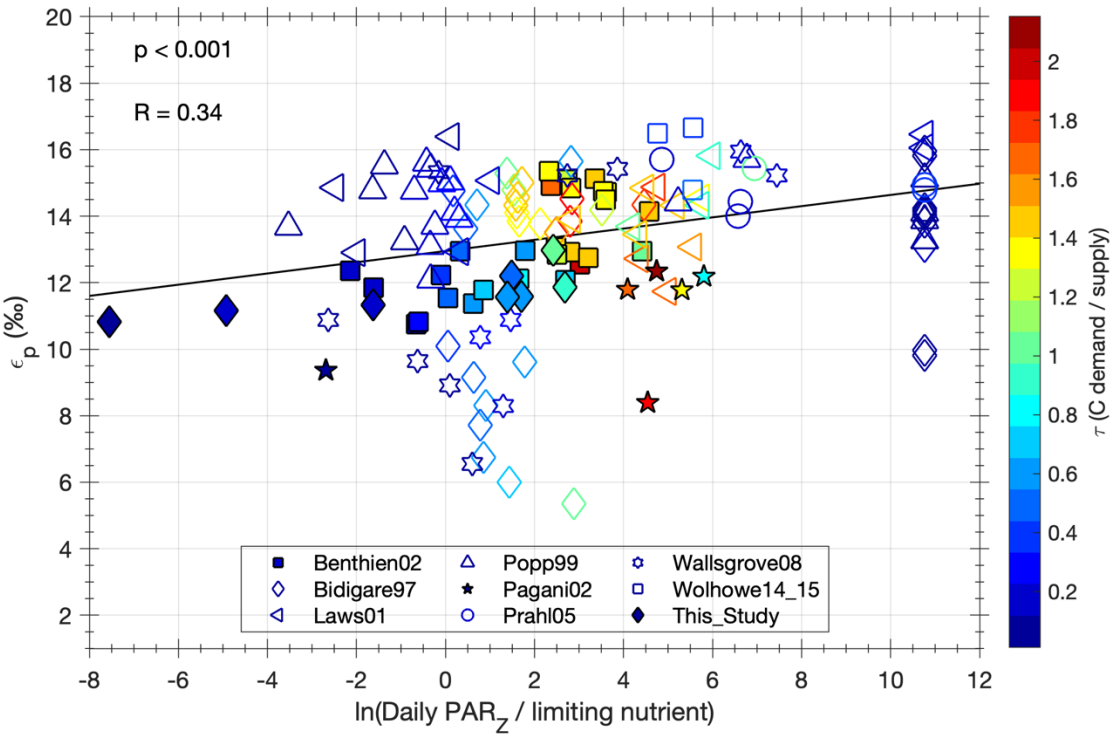

Figure 5. Alkenone ε_p as a function of the energy quotient. The energy quotient is calculated by taking the natural log of the ratio of daily irradiance at the collection depth to the Redfield ratio-adjusted limiting nutrient concentration (see Eq. 5, Section 4.1). Filled symbols are sediment samples, open symbols are particulate samples, both are colored by τ (C demand / diffusive CO_2 supply). While the ratio of irradiance to nutrients does not explain a majority of the variance in ε_p, it is significantly positively correlated with ε_p spatially in the modern ocean, as in culture experiments. The correlation coefficient and *p*-value of the regression are calculated using the *fitlm* function in MATLAB R2019a with "robust" fitting to reduce outlier effects.

3.5.2 Application of a culture-based ε_p model to marine samples

A recent empirical study by Stoll et al. (2019) demonstrated that Noelaerhabdaceae ε_p could be modeled with high accuracy and precision using instantaneous cellular growth rate, cell size, $[CO_{2(aq)}]$, and light intensity; in Chapter 2, we update these multiple linear regressions to include new pH manipulation experiments of *G. oceanica*. The latter presented three multiple linear models for predicting ε_p: one including only the variables in τ, one adding light intensity and photoperiod as separate predictors, and one adding daily integrated irradiance (the product of

light intensity and photoperiod). Here we examine whether the integrated irradiance model (Chapter 2, Appendix A), which is derived from well-controlled culture experiments, is applicable to samples from the natural environment (Fig. 6).

In the culture calibrations, ε_p is best modeled using $[CO_{2(aq)}]$, μ, cell radius, light intensity, and photoperiod as independent predictors (Chapter 2). In marine sediments, we find ε_p is best modeled using the integrated irradiance model, which also performs well in the culture calibrations ($r^2 \cong 0.7$; Chapter 2, Appendix A). We calculate the prediction error using this integrated irradiance model as the difference between ε_p measured and ε_p modeled. The distribution of measured-modeled values in the sediment sample set is remarkably similar to that in the cultures (Fig. 6D). This further suggests the same factors controlling Noelaerhabdaceae carbon isotope fractionation in cultures are at play in the ocean. This is a significant finding, as the canonical diffusive model has had mixed results explaining ε_p variations in sediments where CO_2 concentrations are relatively well-constrained (Benthien et al., 2002; Pagani et al., 2002; Badger et al., 2019; Zhang et al., 2019). There is a large spread to the measured-modeled values (Figs. 6B,C), and indeed some modeled values fall far from measured values, but given the breadth of geographies and environmental conditions included in the sediment dataset, and the generalized and consistent treatment of samples (i.e. setting the alkenone production depth to 75% of the base of the mixed layer depth for all samples)—the result is robust.

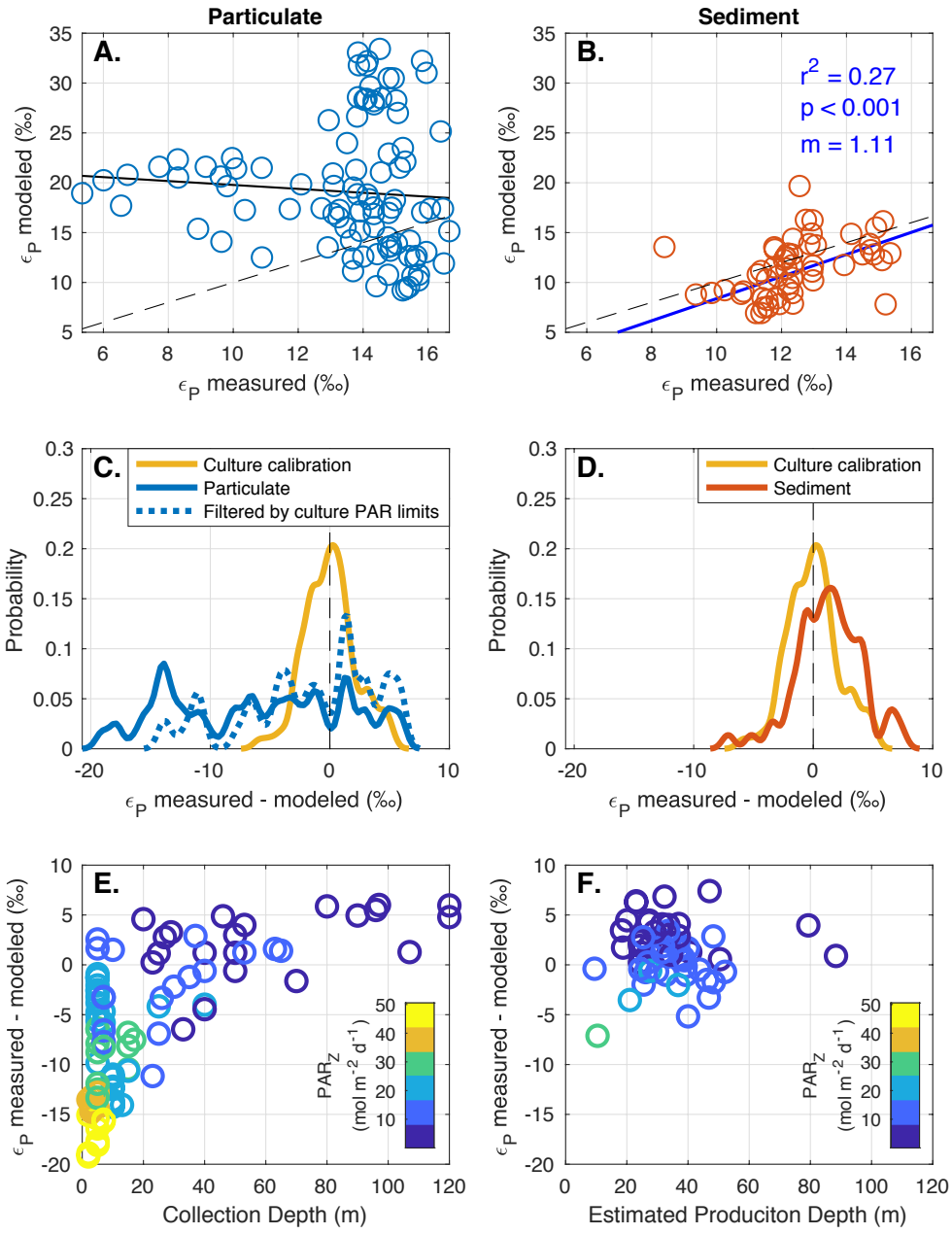

Figure 6. Application of culture-based linear regression model to marine samples. We test the integrated irradiance multiple linear regression model on particulate (A, C, E) and sediment (B, D, F) samples. ε_p is modeled as a linear function of daily irradiance, cell size, CO_2, and growth rate. (A) and (B): modeled vs. measured ε_p values in particulates and sediments, showing no relationship in the particulates, but a reasonably good agreement in the sediments. Dashed line in both panels is the 1:1 line, solid line is a least-squares linear regression in MATLAB R2019a with the "robust" option in the *fitlm* function. (C) and (D): probability density functions (pdfs, calculated using the *ksdensity* function in MATLAB R2019a with a bandwidth of 0.5) of measured minus modeled ε_p values from the culture calibration (solid yellow line), particulates (blue), and sediments (red). Dashed line in (C) shows pdf of measured minus modeled values filtered by the irradiance range in the culture calibration. The close correspondence between the error pdfs of the sediments and cultures demonstrates that the relationships

and coefficients derived from culture experiments can be used to predict ε_p in surface sediments. (E) and (F): error analysis plotting measured minus modeled ε_p values against the collection depth and production depth, respectively; the error in ε_p prediction is highly correlated to irradiance in the sample, as shown by the coloring of samples in panels E and F.

Error analysis of this model (Fig. 6E-F) shows the prediction error in ε_p is highly correlated with integrated irradiance in the particulate sample set, but less so in the sediment sample set. In the particulate dataset, we are unable to model ε_p with much precision or accuracy. The positive correlation between irradiance and model error suggests the coefficient on integrated light developed from the cultures may be too large for the ocean, leading to higher-than-measured ε_p at high light. Filtering the samples plotted to include only those with irradiance values less than the maximum irradiance in the calibration dataset (21.6 mol m^{-2} d^{-1}) improves the agreement between the culture and particulate error distributions (Fig. 6C, dashed line). It is also possible that the too-high calculated ε_p results from using irradiance values that are too high. The majority of the particulate samples come from $\leq$ 10 m depth, while the median mixed-layer depth is ~20 m. Roughly two-thirds of the particulate samples were collected within the mixed-layer. Most of these samples were collected by filtering at a discrete depth for less than one day. However, because the upper water column is mixing during sample collection, it is unlikely that the entirety of the biomass collected during sampling was generated under the environmental conditions at that collection depth. Alkenone ε_p in the particulate samples is modeled using the irradiance at the collection depth, yet much of the biomass was likely produced under lower light (and thus lower ε_p) conditions at greater depths within the mixed layer, and mixed up to the sampling depth. Furthermore, in many of the particulate samples, we had to estimate irradiance from satellite-based mean monthly climatological datasets. It is entirely possible, however, that the days preceding alkenone collection were more or less cloudy than mean monthly conditions,

leading to higher or lower light intensity when the sampled biomass and alkenones were produced. Finally, it is possible that interpolation of the monthly-average climatology to the sampling day may be systematically biased if the location has high interannual variability in irradiance. In an additional error analysis, we solve for the irradiance needed to match ε_p (Fig. S2). We find that for the ~75% of the particulate samples, the solved irradiance is much lower than the prescribed irradiance at the sampling depth. This suggests that a large portion of the alkenones collected during water column sampling may have been produced under lower light conditions, and at deeper depths in the mixed layer than the mode ($\leq$ 10 m) of the particulate dataset.

The model performs better at predicting ε_p in the sediment samples. That we are able to grossly model alkenone ε_p in sediment samples assuming a common depth of production (75% of the mixed-layer depth) and using mean annual irradiance is a highly significant finding. In contrast to the water column particulate samples, this is likely because the depth-integrated irradiance signal that is incorporated into the sediment reflects production at lower mean irradiance than the surface particulate sampling. Figure 6F supports this argument, as there is a much weaker correlation between prediction error and irradiance. Because the sediment samples integrate hundreds to thousands of years of alkenone production (10^3 to 10^7 generations of Noelaerhabdaceae), and the estimation of model parameters is directly analogous to the methods for reconstructing ε_p in paleo samples, the ability to reasonably model ε_p with a consistent irradiance treatment across all samples supports the use of this culture-based model to recover paleoenvironmental information from alkenone ε_p in ancient sediments.

3.5.3 Conventional application of the alkenone-pCO_2 proxy is not supported by existing culture or marine data

Past applications of the alkenone-CO_2 proxy have assumed the Noelaerhabdaceae acquire CO_2 for fixation exclusively via diffusion (Pagani et al., 1999; Zhang et al., 2019). The theory underpinning alkenone carbon isotope fractionation is derived from studies of terrestrial plants (Farquhar et al., 1982), in which ε_p can be described by as the balance of two fractionating steps: diffusion of CO_2 into the cell (~0.7‰ for algae in aqueous solution), and carboxylation by RubisCO (~25‰). The ratio of the cellular CO_2 consumption rate (a function of the carbon fixation rate) to the diffusive CO_2 supply rate (a function of the ambient CO_2 concentration and the cell membrane surface area and permeability) defines the slope of the line between these two fractionating steps. Surveys of marine $\delta^{13}C_{POC}$ indeed found an inverse relationship between ε_p and CO_2, which led to the empirical relationship of

$$\varepsilon_p = \varepsilon_f - b/CO_2 \qquad \text{[Eq. 6]}$$

where ε_f is the fractionation by RubisCO (~25‰), and the *b* term accounts for all non-CO_2 effects on ε_p (Jasper et al., 1994; Bidigare et al., 1997). This equation was introduced by Jasper et al. (1994) and has been the basis for alkenone-based CO_2 reconstructions for over two decades (Pagani et al., 1999; Zhang et al., 2019). A theoretical mechanistic model by Rau et al. (1996) formalized various physiological influences and demonstrated that *b* is directly proportional to growth rate and cell size, a finding that seemed confirmed by a linear relationship between ε_p and μ/CO_2 in nitrate-limited chemostat cultures with high and continuous irradiance (Bidigare et al., 1997). Because growth rates have been difficult to estimate by proxy, empirical calibrations from modern particulate samples (a subset of the "particulate" dataset presented here) related *b* to in-

situ phosphate concentrations, which, in the original calibration set, were found to be very highly correlated (Bidigare et al., 1997). Past [$CO_{2(aq)}$] is then calculated using modern or modeled phosphate concentrations at the core location to estimate the *b* value, and Eq. 6 is solved for CO_2 using the measured ε_p. However, this approach has yielded inconsistent results in sediment studies (Badger et al., 2019; Zhang et al., 2019), and fails to describe alkenone ε_p in culture experiments (Stoll et al., 2019; Chapter 2). Here we revisit the statistical and theoretical basis for the conventional alkenone-p_{CO2} proxy and argue that it should no longer be used to generate CO_2 reconstructions.

The original rationale for predicting *b* from phosphate concentrations was that it had additional explanatory power over CO_2 alone. Because *b* is solved from Eq. 4 using particulate samples and [$CO_{2(aq)}$], a correlation between [PO_4^{3-}] and *b* could arise from an underlying correlation between [$CO_{2(aq)}$] and [PO_4^{3-}]. Bidigare et al. (1997) considered this possibility, and concluded that [PO_4^{3-}] contained information about *b* in addition to the minor influence of CO_2 because the correlation between *b* and PO_4 ($r^2 = 0.95$) in their calibration dataset was much greater than the correlation coefficient between CO_2 and [PO_4^{3-}] ($r^2 = 0.65$). They ascribed this greater correlation to a likely relationship between [PO_4^{3-}] and growth rate effects (that should be present in the *b* term).

We find the relationship between PO_4 and *b* is less robust than previously demonstrated (Fig. 7A), and [PO_4^{3-}] and [$CO_{2(aq)}$] are almost exactly as well correlated as [PO_4^{3-}] and *b* ($r^2 = 0.36$ and $r^2 = 0.37$, respectively, Fig. 7A-B). Our analysis, using additional calibration data published since the original study in 1997, suggests there is no extra information about ε_p contained in [PO_4^{3-}].

This result demonstrates that any apparent success reconstructing past CO_2 with the conventional alkenone-pCO_2 method results from the construction of the *b* parameter from $[CO_{2(aq)}]$ and the strong correlation between near-surface $[PO_4^{3-}]$ and $[CO_{2(aq)}]$ in the modern ocean. If the modern PO_4^{3-} and $CO_{2(aq)}$ concentrations at the core site fall on the global relationship, then the mean $[CO_{2(aq)}]$ reconstructed using ε_p and Eq. 6 will as well, because *b* is solved from $[CO_{2(aq)}]$ and ε_p, and $[PO_4^{3-}]$ is directly translatable to $[CO_{2(aq)}]$. For example, the core-top study of Pagani et al. (2002) aimed to test the alkenone-pCO_2 proxy with modern samples. However, the modern $[PO_4^{3-}]$ and $[CO_{2(aq)}]$ values at the core sites of Pagani et al. (2002) fall very close to the calibration regression line. Therefore the measured ε_p values are scaled by *b* values that were statistically determined to yield the correct modern $[CO_{2(aq)}]$ by the algebraic formulation of Eq. 6. Any study of core-top/modern samples that uses a match between modern and reconstructed CO_2 values will yield a false-positive result because the statistical relationship between ε_p, $[PO_4^{3-}]$ and $[CO_{2(aq)}]$ is algebraically included in the determination of *b*.

Bidigare et al. (1997) argued that *b*, and by correlation, $[PO_4^{3-}]$, are directly proportional to Noelaerhabdaceae growth rates. This reasoning was supported with available evidence at the time from 24-hr light, nitrate-limited chemostat cultures with *E. huxleyi*, which yielded a linear relationship between ε_p and μ/CO_2 (Bidigare et al., 1997). Additional batch culture experiments with light:dark cycles have since shown that the slope between ε_p and τ (proportional to μ/CO_2 x V/SA) is dependent on the τ range of the treatment, and there is in fact a non-linear relationship between ε_p and τ. Using published field data, we can test the hypothesis that *b* is proportional to growth rate. Fig. 7C shows relationships between *b* and μ for samples in our growth rate training set (where growth rates were measured using in-situ isotope-labelling incubations), as well as the entire particulate dataset (where growth rates are modeled as described in Section 2.5). If the Rau

et al. (1996) diffusion-CO_2 model and the conventional alkenone-pCO_2 proxy accurately described carbon isotope fractionation, we would see a positive correlation between *b* and μ. Instead, we observe the opposite: a wide range of *b* values are found at low growth rates, but small *b* values are found at high growth rates (Fig. 7C). This is because the highest growth rates occur at high light, high temperature, near-surface locations, where $[CO_{2(aq)}]$ is low and ε_p is high as a result of high irradiance. Arithmetically, these combine in Eq. 6 to produce a low *b* value. The theoretical underpinnings of the conventional construction of alkenone-pCO_2 proxy (Eq. 6) are therefore not supported when the entirety of existing calibration data is considered.

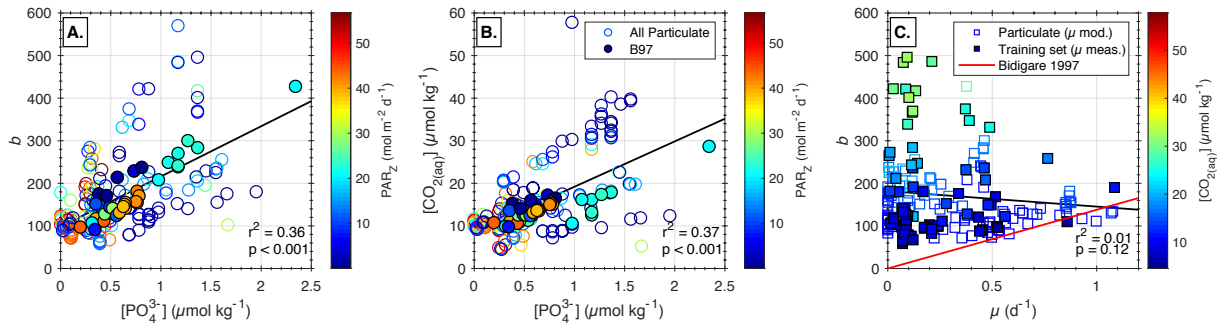

Figure 7. Calibration of the conventional alkenone-pCO_2 proxy. (A) *b* value calculated using Eq. 4 as a function of in-situ phosphate. In (A) and (B), filled symbols are the data used by Bidigare et al. (1997) for the *b*-value calibration from particulate samples; open symbols are all data generated from particulate samples since 2000. (B) Aqueous CO_2 concentration regressed against phosphate concentration. (C) *b* plotted against cellular growth rate; filled symbols are samples with measured growth rates in the field, open symbols are samples where growth rate is modeled. The conventional alkenone-pCO_2 proxy argues that *b* and μ are directly proportional (red line in panel C), but existing data show the opposite relationship.

One could argue that the relationship between ε_p and $[CO_{2(aq)}]$ through *b* is simply empirical, and all non-CO_2 effects—including but not limited to cell size, growth rate, CCM activity, and irradiance—are encapsulated in *b*. While that is mathematically true, there is no immediately apparent way to quantitatively scale *b* for changes in these non-CO_2 effects. Because the theory behind the CO_2-diffusion model does not accurately describe photosynthetic carbon isotope fractionation in these algae, attempts to modify *b* for changes in cell size

(Henderiks and Pagani, 2007; Henderiks and Pagani, 2008) have limited quantitative basis when the full culture and modern datasets are considered. Therefore, *b,* as originally designed and as subsequently modified, is not supported by the breadth of culture or modern ocean observations, and should not be used for alkenone paleobarometry.

3.5.4 Recommendations for paleobarometry and the sensitivity of alkenone ε_p to known CO_2 change

Our ultimate goal is to extract paleoenvironmental information from the carbon isotope signature recorded in long-chain alkenones. Culture experiments demonstrate that physiology (cell size and growth rate) and environment (light and $[CO_{2(aq)}]$) are key determinants of alkenone ε_p (Bidigare et al., 1997; Popp et al., 1998b; Riebesell et al., 2000b; Rost et al., 2002; McClelland et al., 2017; Stoll et al., 2019; Chapter 2). Our analysis of sediment samples (Fig. 6D) suggests these same factors are also controlling alkenone carbon isotope fractionation in the natural environment, providing a framework for interpreting alkenone ε_p through time from sediment samples.

To illustrate the approach we model alkenone ε_p at a hypothetical tropical deep-sea sediment core site in the western Pacific warm pool—an area largely in CO_2 equilibrium with the atmosphere—for three times in the past, and calculate how changes in CO_2 and other model parameters would manifest in ε_p (Fig. 8). We prescribe atmospheric CO_2 levels of 180 ppm, 280 ppm, and 400 ppm for the cold Last Glacial Maximum (LGM), Holocene, and warm mid-Pliocene, with assumed sea surface temperatures of 26°C, 28°C, and 30°C, respectively. We assume a constant salinity of 35 psu and estimate in-situ $[CO_{2(aq)}]$ using the temperature-dependent solubility of CO_2 (Weiss, 1974). We treat our down-core ε_p modeling the same way as

a core-top sediment sample, setting the production depth as 75% of the mixed layer depth, and calculating nutrients and light at this location from the gridded global datasets used in the analysis above, and modeling μ using our growth rate model described in Section 2.5. From $\Delta[CO_{2(aq)}]$ alone, we calculate an ε_p change of -0.91‰ from Holocene to LGM, and of +0.95‰ from Holocene to Pliocene.

We model changes in coccolith length of ±0.75 μm, irradiance changes of ±25%, and growth rate changes of ±50% and examine the resulting ε_p. The modeled ε_p is very insensitive to growth rate, which, given the uncertainties in modeling growth rate (Fig. 2) using our temperature and substrate-limitation model, bodes well for extracting $[CO_{2(aq)}]$ from ε_p. Because the coefficient on growth rate is small, it does not need to be well constrained to interpret past ε_p changes (Chapter 2). Cell radius (inferred from coccolith length) and irradiance both have strong effects on the modeled ε_p values, and will need to be considered carefully in any paleobarometry applications. Changes in coccolith length have been reconstructed in many locations through the Miocene and Pliocene; some locations show a drop of ~1.5 μm from Pliocene to Holocene (Badger et al., 2019), and a global reduction in Noelaerhabdaceae size in the late Miocene is well documented (Young, 1990; Beaufort, 1992; Takayama, 1993; Kameo and Bralower, 2000), which may explain the apparent rise in ε_p from the mid-Miocene to Pliocene (Pagani et al., 1999; Pagani et al., 2000). Changes in irradiance also have a large influence on the modeled ε_p, with a 25% increase in irradiance (+4 mol m^{-2} d^{-1}) roughly equivalent to a 0.75 μm reduction in coccolith length or a 3 μmol kg^{-1} increase in $[CO_{2(aq)}]$. Explanations for ε_p variations in sedimentary sequences will therefore be non-unique, but environmental information can still be extracted by constraining coccolith size and carefully considering irradiance changes. While understanding the temporal variations in the parameters influencing ε_p records is beyond the

scope of this manuscript, this study provides a quantitative framework for considering their effects on CO_2 reconstructions.

Our ability to invert this model and solve for the predictors of ε_p will also depend on the uncertainty in realizing ε_p, which is a function of alkenone and foraminifera $\delta^{13}C$ measurement precision, as well as the uncertainty in translating foraminiferal calcite $\delta^{13}C$ to $\delta^{13}C_{DIC}$. Typical precision in modern and paleo ε_p samples is between 0.3 and 0.7‰, substantial when compared to the hypothetical changes during the LGM and Pliocene. Increasing the number of sites and samples for each reconstruction should help overcome this issue through cancellation of site-specific or random uncertainties, increasing the confidence in estimated paleo-pCO_2 values. As discussed in Section 3.4.1, because irradiance is the only important parameter that cannot be constrained in sediment samples, target locations should be those most likely to have a stable irradiance regime. Such locations are pelagic sites with deep and stable mixed layers, with low seasonality or interannual variability in mixed layer depth, surface irradiance, and productivity, as all three of these parameters can modulate the penetration of irradiance in the water column. Warm tropical locations that are hydrographically stable will therefore be best-suited for alkenone paleobarometry while high-latitude or dynamic oceanographic environments should be avoided.

The use of other paleoecological and paleoenvironmental proxies in tandem with alkenone ε_p may help constrain whether light conditions or trophic status of the core location have changed dramatically through time. Nonetheless, that the regression coefficients derived from culture experiments produce reasonable hypothetical changes in ε_p over the Plio/Pleistocene suggests it may be possible to tease apart past variations in ε_p in terms of three fundamental influences: irradiance, cell size, and CO_2. Investigations of alkenone ε_p and cell size

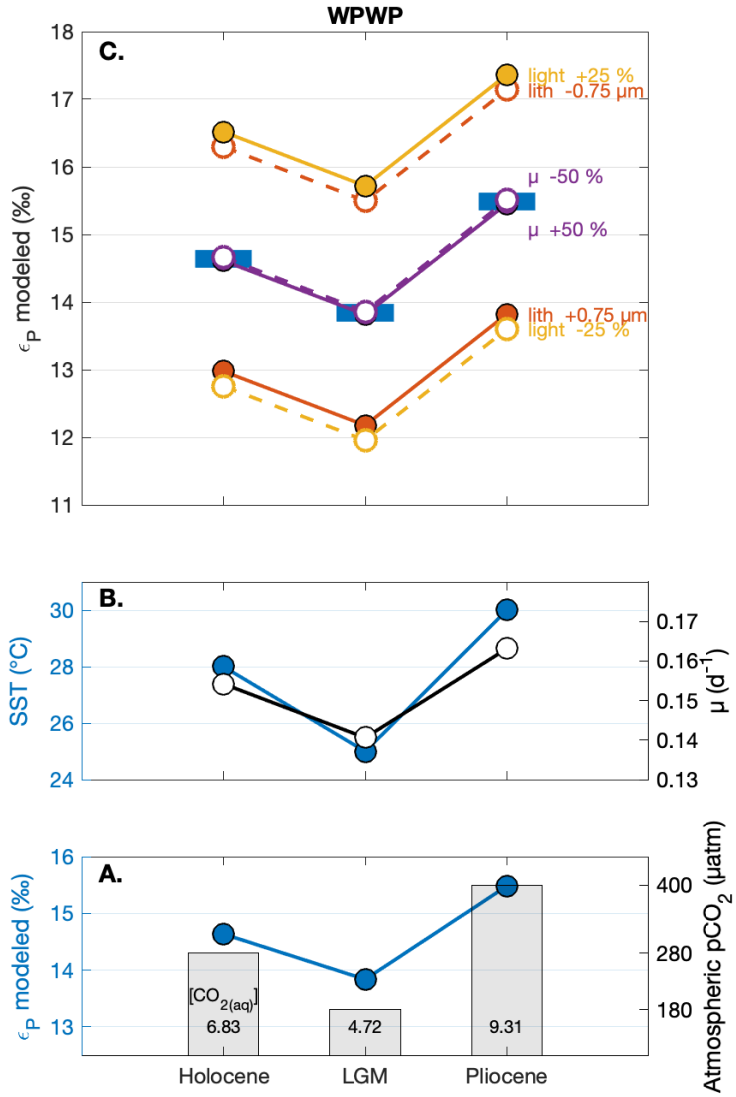

Figure 8. Modeled alkenone ε_p in the Pliocene/Pleistocene at a hypothetical location in the Western Pacific Warm Pool (WPWP). (A) In-situ [$CO_{2(aq)}$] for the Holocene, LGM, and Pliocene (3.2-3.8 Ma), from prescribed atmospheric CO_2 concentrations of 280, 180, and 400 ppm, respectively (gray bars; right axis); and modeled ε_p resulting from [$CO_{2(aq)}$] variations alone. Irradiance and mixed-layer depths are determined using the same algorithm as for the surface sediment samples, and alkenone ε_p is modeled using the integrated irradiance multilinear model. We use the mean cell radius of the global ocean (2.5 µm) for our baseline case. (B) Sea surface temperatures (blue) and estimated growth rates (black), modeled using the growth rate model described in the text. [$CO_{2(aq)}$] in (A) is calculated using the SSTs shown in (B), and assuming constant salinity of 35.5. (C) Sensitivity of ε_p to changes in model parameters on top of ΔCO_2: irradiance (yellow), growth rate (purple), coccolith length (red). Dashed lines and open symbols show a reduction in the model parameter value, while solid lines filled symbols show the effect of an increase in the parameter value. Blue bars show the baseline ε_p case with only CO_2 sensitivity. For example, the solid red lines show the effect on ε_p of increasing coccolith length by 0.75 µm from the baseline scenario with only CO_2 changes (blue symbols).

through time periods of known atmospheric CO_2 change can provide further validation of our model and may lead to an enhanced understanding of atmospheric CO_2 variations in the Cenozoic.

3.6 Conclusion

Although alkenone carbon isotope fractionation (ε_p) has a long history of being used as an atmospheric pCO_2 paleobarometer, recent culture and sediment studies have called into question the fidelity of the proxy. Culture observations have demonstrated that, in addition to cell size, $[CO_{2(aq)}]$, and growth rate, irradiance is a key parameter driving ε_p variability. We observe a similar dependence of ε_p on irradiance in the modern ocean, in both sediment and particulate samples. With a multiple linear regression model derived from culture experiments, we are not able to model ε_p well in particulate samples, likely because the alkenones collected during particulate sampling were synthesized throughout the water column, and the irradiance at the depth of sampling is too high compared to the conditions under which the algae were growing. However, we demonstrate that ε_p can be modeled in marine sediments with reasonable accuracy, and with a measured versus modeled slope near 1. This finding is the first successful application of culture-based models of alkenone ε_p to sediments where $[CO_{2(aq)}]$ is known. Using measurements mined from the literature of the growth rates of alkenone producers in nature, we show that the theory underpinning the diffusive model is not supported by the field data. This study presents a framework for beginning to quantitatively interpret alkenone ε_p variations in ancient sediments, and the extraction of CO_2 information may be possible if irradiance and cell size can be constrained.

3.7 Acknowledgements

We thank the Lamont-Doherty Earth Observatory Core Repository for CHEETA core-top samples used in this analysis. We thank Angela Dial and Yves Gally for sample and laboratory assistance. We thank the Lamont Climate Center and the Geological Society of America for financial support. SRP acknowledges the NSF Graduate Research Fellowship Program (Grant # DGE-16-44869). This material is based upon research supported by the Chateaubriand Fellowship of the Office for Science & Technology of the Embassy of France in the United States.

3.8 References

Andersen N., Müller P. J., Kirst G. and Schneider R. R. (1999) Alkenone δ13C as a Proxy for Past PCO2 in Surface Waters: Results from the Late Quaternary Angola Current. In *Use of Proxies in Paleoceanography* Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 469–488.

Arístegui J., Barton E. D., Álvarez-Salgado X. A., Santos A. M. P., Figueiras F. G., Kifani S., Hernández-León S., Mason E., Machú E. and Demarcq H. (2009) Sub-regional ecosystem variability in the Canary Current upwelling. *Prog. Oceanogr.* **83**, 33–48.

Badger M. P. S., Foster G. L., Chalk T. B., Gibbs S. J., Badger M. P. S., Pancost R. D., Schmidt D. N., Sexton P. F., Mackensen A., Bown P. R. and Pälike H. (2019) Insensitivity of alkenone carbon isotopes to atmospheric CO2 at low to moderate CO2 levels. *Clim. Past* **15**, 539–554.

Beaufort L. (1992) Size variations in late Miocene Reticulofenestra and implication for paleoclimatic interpretation. *Mem. die Sci. Geol.* **43**, 339–350.

Bendif E. M., Probert I., Díaz-Rosas F., Thomas D., van den Engh G., Young J. R. and von Dassow P. (2016) Recent reticulate evolution in the ecologically dominant lineage of coccolithophores. *Front. Microbiol.* **7**.

Benthien A., Andersen N. and Schulte S. (2002) Carbon isotopic composition of the C37:2 alkenone in core top sediments of the South Atlantic Ocean: Effects of CO2 and nutrient concentrations. *Global Biogeochem. Cycles* **16**, 12-1-12–12.

Benthien A., Zondervan I., Engel A., Hefter J., Terbrüggen A. and Riebesell U. (2007) Carbon isotopic fractionation during a mesocosm bloom experiment dominated by Emiliania huxleyi: Effects of CO2 concentration and primary production. *Geochim. Cosmochim. Acta* **71**, 1528–1541.

Bidigare R. R., Fluegge A., Freeman K. H., Hanson K. L., Hayes J. M., Hollander D., Jasper J. P., King L. L., Laws E. A., Milder J., Millero F. J., Pancost R. D., Popp B. N., Steinberg P. A. and Wakeham S. G. (1999) Correction to "Consistent fractionation of 13C in nature and in the laboratory: Growth-rate effects in some haptophyte algae" by R.R. Bidigare et al. **13**, 251–252.

Bidigare R. R., Fluegge A., Freeman K. H., Hanson K. L., Hayes J. M., Hollander D., Jasper J. P., King L. L., Laws E. A., Milder J., Millero F. J., Pancost R., Popp B. N., Steinberg P. A. and Wakeham S. G. (1997) Consistent fractionation of 13C in nature and in the laboratory: Growth-rate effects in some haptophyte algae. *Global Biogeochem. Cycles* **11**, 279–292.

de Boyer Montégut C., Madec G., Fischer A. S., Lazar A. and Iudicone D. (2004) Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *J. Geophys. Res. C Ocean.* **109**, 1–20.

Bruland K. W., Middag R. and Lohan M. C. (2014) Controls of Trace Metals in Seawater. In *Treatise on Geochemistry* Elsevier. pp. 19–51.

Burkhardt S., Riebesell U. and Zondervan I. (1999) Effects of growth rate , CO2 concentration , and cell size on the stable carbon isotope fractionation in marine phytoplankton. *Geochim. Cosmochim. Acta* **63**, 3729–3741.

Calil P. H. R., Doney S. C., Yumimoto K., Eguchi K. and Takemura T. (2011) Episodic upwelling and dust deposition as bloom triggers in low-nutrient, low-chlorophyll regions. *J. Geophys. Res.* **116**, C06030.

Clegg S. L. and Sarmiento J. L. (1989) The hydrolytic scavenging of metal ions by marine particulate matter. *Prog. Oceanogr.* **23**, 1–21.

Dickson A. G. (1990) Thermodynamics of the dissociation of boric acid in potassium chloride solutions from 273.15 to 318.15 K. *J. Chem. Eng. Data* **35**, 253–257.

Eek M. K., Whiticar M. J., Bishop J. K. B. and Wong C. S. (1999) Influence of nutrients on carbon isotope fractionation by natural populations of Prymnesiophyte algae in NE Pacific. *Deep. Res. Part II Top. Stud. Oceanogr.* **46**, 2863–2876.

Falkowski P. G. and Raven J. A. (2007) *Aquatic Photosynthesis (Second edition).*,

Farquhar G. D., O'Leary M. H. and Berry J. A. (1982) On the Relationship Between Carbon Isotope Discrimination and the Intercellular Carbon Dioxide Concentration in Leaves. *Funct. Plant Biol.* **9**, 121.

Fielding S. R. (2013) Emiliania huxleyi specific growth rate dependence on temperature. *Limnol. Oceanogr.* **58**, 663–666.

Francois R., Altabet M. a., Goericke R., McCorkle D. C., Brunet C. and Poisson A. (1993) Changes in the δ13C of surface water particulate organic matter across the subtropical convergence in the SW Indian Ocean. *Global Biogeochem. Cycles* **7**, 627.

Freeman K. H. and Hayes J. M. (1992) Fractionation of carbon isotopes by phytoplankton and estimates of ancient CO2 levels. *Global Biogeochem. Cycles* **6**, 185–198.

Fry B. and Wainwright S. C. (1991) Diatom sources of 13C-rich carbon in marine food webs. *Mar. Ecol. Prog. Ser.* **76**, 149–157.

Harada N., Shin K. H., Murata A., Uchida M. and Nakatani T. (2003) Characteristics of alkenones synthesized by a bloom of emiliania huxleyi in the Bering Sea. *Geochim. Cosmochim. Acta* **67**, 1507–1519.

Hayes J. (1993) Factors controlling 13C contents of sedimentary organic compounds: Principles and evidence. *Mar. Geol.* **113**, 111–125.

Henderiks J. and Pagani M. (2008) Coccolithophore cell size and the Paleogene decline in atmospheric CO2. *Earth Planet. Sci. Lett.* **269**, 575–583.

Henderiks J. and Pagani M. (2007) Refining ancient carbon dioxide estimates: Significance of coccolithophore cell size for alkenone-based pCO2 records. *Paleoceanography* **22**, 1–12.

Hopkinson B. M., Dupont C. L., Allen A. E. and Morel F. M. M. (2011) Efficiency of the CO2-concentrating mechanism of diatoms. *Proc. Natl. Acad. Sci.*, 1–8.

Iglesias-Rodriguez M. D., Schofield O. M., Batley J., Medlin L. K. and Hayes P. K. (2006) INTRASPECIFIC GENETIC DIVERSITY IN THE MARINE COCCOLITHOPHORE EMILIANIA HUXLEYI (PRYMNESIOPHYCEAE): THE USE OF MICROSATELLITE ANALYSIS IN MARINE PHYTOPLANKTON POPULATION STUDIES1. *J. Phycol.* **42**, 526–536.

Jasper J. P., Hayes J. M., Mix A. C. and Prahl F. G. (1994) Photosynthetic fractionation of 13C and concentrations of dissolved CO2 in the central equatorial Pacific during the last 255,000 years. *Paleoceanography* **9**, 781–798.

Kameo K. and Bralower T. (2000) Neogene Calcareous Nannofossil Biostratigraphy of Sites 998, 999, and 1000, Caribbean Sea. *Proc. Ocean Drill. Progr.* **165**, 3–17.

Keeling C. D. (1979) The Suess effect: 13Carbon-14Carbon interrelations. *Environ. Int.* **2**, 229–300.

Key R. M., Olsen A., van Heuven S., Lauvset S. K., Velo A., Lin X., Schirnick C., Kozyr A., Tanhua T. and Hoppema M. (2015) *Global Ocean Data Analysis Project, Version 2 (GLODAPv2)*.,

Krumhardt K. M., Lovenduski N. S., Iglesias-Rodriguez M. D. and Kleypas J. A. (2017) Coccolithophore growth and calcification in a changing ocean. *Prog. Oceanogr.* **159**, 276–295.

Laws E. A., Popp B. N., Bidigare R. R., Kennicutt M. C. and Macko S. A. (1995) Dependence of phytoplankton carbon isotopic composition on growth rate and [CO2]aq: Theoretical considerations and experimental results. *Geochim. Cosmochim. Acta* **59**, 1131–1138.

Laws E. A., Popp B. N., Bidigare R. R., Riebesell U., Burkhardt S. and Wakeham S. G. (2001) Controls on the molecular distribution and carbon isotopic composition of alkenones in certain haptophyte algae. *Geochemistry, Geophys. Geosystems* **2**.

Lombard F., da Rocha R. E., Bijma J. and Gattuso J.-P. (2010) Effect of carbonate ion concentration and irradiance on calcification in planktonic foraminifera. *Biogeosciences* **7**, 247–255.

Lueker T. J., Dickson A. G. and Keeling C. D. (2000) Ocean pCO2 calculated from dissolved inorganic carbon, alkalinity, and equations for K1 and K2: Validation based on laboratory measurements of CO2 in gas and seawater at equilibrium. *Mar. Chem.* **70**, 105–119.

Macias D., Landry M. R., Gershunov A., Miller A. J. and Franks P. J. S. (2012) Climatic control of upwelling variability along the Western North-American coast. *PLoS One* **7**, 20–23.

McClelland H. L. O., Barbarin N., Beaufort L., Hermoso M., Ferretti P., Greaves M. and Rickaby R. E. M. (2016) Calcification response of a key phytoplankton family to millennial-scale environmental change. *Sci. Rep.* **6**, 1–11.

McClelland H. L. O., Bruggeman J., Hermoso M. and Rickaby R. E. M. (2017) The origin of carbon isotope vital effects in coccolith calcite. *Nat. Commun.* **8**, 14511.

McGee D., DeMenocal P. B., Winckler G., Stuut J. B. W. and Bradtmiller L. I. (2013) The magnitude, timing and abruptness of changes in North African dust deposition over the last 20,000 yr. *Earth Planet. Sci. Lett.* **371**–**372**, 163–176.

Mook W. G., Bommerson J. C. and Staverman W. H. (1974) Carbon isotope fractionation between dissolved bicarbonate and gaseous carbon dioxide. *Earth Planet. Sci. Lett.* **22**, 169–176.

Müller P. J., Kirst G., Ruhland G., von Storch I. and Rosell-Melé A. (1998) Calibration of the alkenone paleotemperature index U37K′ based on core-tops from the eastern South Atlantic and the global ocean (60°N-60°S). *Geochim. Cosmochim. Acta* **62**, 1757–1772.

NASA Goddard Space Flight Center, Ocean Ecology Laboratory and Ocean Biology Processing Group (2018a) *Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Downwelling Diffuse Attenuation Coefficient Data; 2018 Reprocessing*., Greenbelt, MD, USA.

NASA Goddard Space Flight Center, Ocean Ecology Laboratory and Ocean Biology Processing Group (2018b) *Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Photosynthetically Available Radiation Data; 2018 Reprocessing*., NASA OB.DAAC, Greenbelt, MD, USA.

NASA Goddard Space Flight Center, Ocean Ecology Laboratory and Ocean Biology Processing Group (2018c) *Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Downwelling Diffuse Attenuation Coefficient Data; 2018 Reprocessing*., Greenbelt, MD, USA.

NASA Goddard Space Flight Center, Ocean Ecology Laboratory and Ocean Biology Processing Group (2018d) *Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Photosynthetically Available Radiation Data; 2018 Reprocessing*., Greenbelt, MD, USA.

Olsen A., Key R. M., Van Heuven S., Lauvset S. K., Velo A., Lin X., Schirnick C., Kozyr A., Tanhua T., Hoppema M., Jutterström S., Steinfeldt R., Jeansson E., Ishii M., Pérez F. F. and

Suzuki T. (2016) The global ocean data analysis project version 2 (GLODAPv2) - An internally consistent data product for the world ocean. *Earth Syst. Sci. Data* **8**, 297–323.

Olsen A., Lange N., Key R. M., Tanhua T., Álvarez M., Becker S., Bittig H. C., Carter B. R., Cotrim da Cunha L., Feely R. A., van Heuven S., Hoppema M., Ishii M., Jeansson E., Jones S. D., Jutterström S., Karlsen M. K., Kozyr A., Lauvset S. K., Lo Monaco C., Murata A., Pérez F. F., Pfeil B., Schirnick C., Steinfeldt R., Suzuki T., Telszewski M., Tilbrook B., Velo A. and Wanninkhof R. (2019) GLODAPv2.2019 - an update of GLODAPv2. *Earth Syst. Sci. Data Discuss.*, 1–39.

Pagani M. (2014) Biomarker-Based Inferences of Past Climate: The Alkenone pCO2 Proxy. In *Treatise on Geochemistry* Elsevier. pp. 361–378.

Pagani M., Arthur M. A. and Freeman K. H. (1999) Miocene evolution of atmospheric carbon dioxide. *Paleoceanography* **14**, 273–292.

Pagani M., Arthur M. A. and Freeman K. H. (2000) Variations in Miocene phytoplankton growth rates in the southwest Atlantic: Evidence for changes in ocean circulation. *Paleoceanography* **15**, 486–496.

Pagani M., Freeman K. H., Ohkouchi N. and Caldeira K. (2002) Comparison of water column [CO2aq] with sedimentary alkenone-based estimates: A test of the alkenone-CO2 proxy. *Paleoceanography* **17**, 21-1-21–12.

Polissar P. J. and D'Andrea W. J. (2014) Uncertainty in paleohydrologic reconstructions from molecular δD values. *Geochim. Cosmochim. Acta* **129**, 146–156.

Popp B. N., Bidigare R. R., Deschenes B., Laws E. A., Prahl F. G., Tanimoto J. K. and Wallsgrove R. J. (2006) A new method for estimating growth rates of alkenone-producing haptophytes. *Limnol. Oceanogr. Methods* **4**, 114–129.

Popp B. N., Kenig F., Wakeham S. G., Laws E. A. and Bidigare R. R. (1998a) Does growth rate affect ketone unsaturation and intracellular carbon isotopic variability in Emiliania huxleyi ? *Paleoceanography* **13**, 35–41.

Popp B. N., Laws E. A., Bidigare R. R., Dore J. E., Hanson K. L. and Wakeham S. G. (1998b) Effect of phytoplankton cell geometry on carbon isotopic fractionation. *Geochim. Cosmochim. Acta* **62**, 69–77.

Popp B. N., Takigiku R., Hayes J., Louda J. and Baker E. W. (1989) The post-paleozoic chronology and mechanism of 13C depletion in primary marine organic matter.pdf. *Am. J. Sci.* **289**, 436–454.

Prahl F. G., Popp B. N., Karl D. M. and Sparrow M. A. (2005) Ecology and biogeochemistry of alkenone production at Station ALOHA. *Deep. Res. Part I Oceanogr. Res. Pap.* **52**, 699–719.

Rau G. H., Riebesell U. and Wolf-Gladrow D. (1996) A model of photosynthetic 13C fractionation by marine phytoplankton based on diffusive molecular CO2 uptake. *Mar. Ecol. Prog. Ser.* **133**, 275–285.

Rau G. H., Takahashi T., Desmarais D. J., Repeta D. J. and Martin J. H. (1992) The Relationship between d13C of Organic-Matter and [CO2(aq)] in Ocean Surface-Water: Data from a JGOFS Site in the Northeast Atlantic-Ocean and a Model. *Geochim. Cosmochim. Acta* **56**, 1413–1419.

Raven J. A. and Johnston A. M. (1991) Mechanisms of inorganic-carbon acquisition in marine phytoplankton and their implications for the use of other resources. *Limnol. Oceanogr.* **36**, 1701–1714.

Riebesell U., Burkhardt S., Dauelsberg A. and Kroon B. (2000a) Carbon isotope fractionation by a marine diatom: Dependence on the growth-rate-limiting resource. *Mar. Ecol. Prog. Ser.* **193**, 295–303.

Riebesell U., Revill A. T., Holdsworth D. G. and Volkman J. K. (2000b) The effects of varying CO2 concentration on lipid composition and carbon isotope fractionation in Emiliania huxleyi. *Geochim. Cosmochim. Acta* **64**, 4179–4192.

Rosell-Melé A. and Prahl F. G. (2013) Seasonality of UK'37 temperature estimates as inferred from sediment trap data. *Quat. Sci. Rev.* **72**, 128–136.

Rost B., Zondervan I. and Riebesell U. (2002) Light-dependent carbon isotope fractionation in the coccolithophorid Emiliania huxleyi. *Limnol. Oceanogr.* **47**, 120–128.

Schmittner A., Gruber N., Mix A. C., Key R. M., Tagliabue A. and Westberry T. K. (2013) Biology and air-sea gas exchange controls on the distribution of carbon isotope ratios (δ13C) in the ocean. *Biogeosciences* **10**, 5793–5816.

Schneider N. and Müller P. (1990) The Meridional and Seasonal Structures of the Mixed-Layer Depth and its Diurnal Amplitude Observed during the Hawaii-to-Tahiti Shuttle Experiment. *J. Phys. Oceanogr.* **20**, 1395–1404.

Schulte S., Benthien A., Andersen N., Müller P. J., Rühlemann C. and Schneider R. R. (2003) Stable Carbon Isotopic Composition of the C37:2 Alkenone: A Proxy for CO2(aq) Concentration in Oceanic Surface Waters? *South Atl. Late Quat.* **2**, 195–211.

Spero H. J. and Lea D. W. (1993) Intraspecific stable isotope variability in the planktic foraminifera Globigerinoides sacculifer: Results from laboratory experiments. *Mar. Micropaleontol.* **22**, 221–234.

Stoll H. M., Guitian J., Hernandez-Almeida I., Mejia L. M., Phelps S. R., Polissar P. J., Rosenthal Y., Zhang H. and Ziveri P. (2019) Upregulation of phytoplankton carbon

concentrating mechanisms during low CO2 glacial periods and implications for the phytoplankton pCO2 proxy. *Quat. Sci. Rev.* **208**, 1–20.

Takayama T. (1993) Notes on Neogene calcareous nanofossil biostratigraphy of the Ontong Java plateau and size variations of Reticulofenestra coccoliths. *Proc. Ocean Drill. Program, Sci. Results* **130**, 179–229.

Tolosa I., Miquel J.-C., Gasser B., Raimbault P., Azouzi L. and Claustre H. (2008) Ecology and biogeochemistry of contrasting trophic environments in the South East Pacific by carbon isotope ratios on lipid biomarkers. *Biogeosciences Discuss.* **4**, 4653–4696.

Uppström L. R. (1974) The boron/chlorinity ratio of deep-sea water from the Pacific Ocean. *Deep Sea Res. Oceanogr. Abstr.* **21**, 161–162.

Wallsgrove R. J. (2008) Alkenone-producing algae as a function of specific growth. .

Weiss R. F. (1974) Carbon dioxide in water and seawater: the solubility of a non-ideal gas. *Mar. Chem.* **2**, 203–215.

Wilkes E. B., Lee R. B. Y., McClelland H. L. O., Rickaby R. E. M. and Pearson A. (2018) Carbon isotope ratios of coccolith–associated polysaccharides of Emiliania huxleyi as a function of growth rate and CO2 concentration. *Org. Geochem.* **119**, 1–10.

Wilkes E. B. and Pearson A. (2019) A general model for carbon isotopes in red-lineage phytoplankton: Interplay between unidirectional processes and fractionation by RubisCO. *Geochim. Cosmochim. Acta* **265**, 163–181.

Wolhowe M. D., Prahl F. G., Langer G., Oviedo A. M. and Ziveri P. (2015) Alkenone δD as an ecological indicator: A culture and field study of physiologically-controlled chemical and hydrogen-isotopic variation in C37 alkenones. *Geochim. Cosmochim. Acta* **162**, 166–182.

Wolhowe M. D., Prahl F. G., White A. E., Popp B. N. and Rosas-Navarro A. (2014) A biomarker perspective on coccolithophorid growth and export in a stratified sea. *Prog. Oceanogr.* **122**, 65–76.

Yamamoto M., Shimamoto A., Fukuhara T., Naraoka H., Tanaka Y. and Nishimura A. (2007) Seasonal and depth variations in molecular and isotopic alkenone composition of sinking particles from the western North Pacific. *Deep. Res. Part I Oceanogr. Res. Pap.* **54**, 1571–1592.

Young J. (1990) Size variation of Neogene Reticulofenestra coccoliths from Indian Ocean DSDP Cores. *J. Micropalaeontology* **9**, 71–85.

Zhang Y. G., Pearson A., Benthien A., Dong L., Huybers P., Liu X. and Pagani M. (2019) Refining the alkenone-pCO2 method I: Lessons from the Quaternary glacial cycles. *Geochim. Cosmochim. Acta* **260**, 177–191.

Chapter 4: Disentangling the CO_2 signal in alkenone carbon isotope fractionation in the Pleistocene

4.1 Abstract

Carbon isotope fractionation of coccolithophore algae reconstructed from alkenone biomarkers ($\varepsilon_{p37:2}$) have a long history as a proxy for reconstructing past atmospheric CO_2. Recent studies have found disagreement between the estimates using the conventional alkenone pCO2 model and known atmospheric CO_2 fluctuations in the Pleistocene, raising questions about the quantitative scaling between CO_2 and $\varepsilon_{p37:2}$ under the classical diffusive model. Recent empirical models based on culture observations suggest the sensitivity of $\varepsilon_{p37:2}$ to CO_2 is relatively small (~1‰/2μM), and that $\varepsilon_{p37:2}$ is similarly sensitive to irradiance, a parameter with no true paleoproxy. Here we generate two new Pleistocene $\varepsilon_{p37:2}$ and coccolith size records from ODP 807 and ODP 1207 in the western Pacific. We use these new records along with a compilation of all existing Pleistocene $\varepsilon_{p37:2}$ data to identify locations where measured $\varepsilon_{p37:2}$ corresponds most closely to $\varepsilon_{p37:2}$ predicted from ice-core CO_2 variations. These locations are the best candidates for reconstructing CO_2 during older time periods as our results suggest that the coccolithophore growth environment was relatively stable over the Pleistocene with respect to irradiance and other factors that also affect $\varepsilon_{p37:2}$. We find that the best locations for paleobarometry are oligotrophic and have relatively stable modern irradiance conditions (surface irradiance, attenuation coefficients, and mixed layer depths) on a mean annual basis. We also evaluate uncertainties associated with estimating CO_2 from the model, determining that Pleistocene- and Pliocene-scale CO_2 changes continue to be difficult to resolve, but larger changes, such as those over the Neogene or Cenozoic, will be quantifiable with this method.

4.2 Introduction

Atmospheric carbon dioxide is a primary driver of global climate and plays a fundamental role as both a direct control and as a feedback (Foote, 1856; Shakun et al., 2012). As anthropogenic carbon emissions continue unabated, it is imperative to improve our understanding of how increased CO_2 concentrations will affect the biosphere and climate system. Critical to this goal is reconstructing atmospheric CO_2 variations during past warm periods (Rohling et al., 2012). The alkenone-CO_2 proxy is one of the most widely applied tools for determining ancient atmospheric CO_2 levels in the Cenozoic (Pagani et al., 1999a; Pagani, 2005; Beerling and Royer, 2011). This method estimates CO_2 from the stable carbon isotope ratio ($\delta^{13}C$) of algal biomass, which is preserved in the $\delta^{13}C$ values of long-chain alkenones (Popp et al., 1998a). These unique biomarkers are produced by Noelaerhabdaceae coccolithophores, a geographically wide-spread and abundant family of plankton (Young et al., 2003). Theory and empirical evidence from cultures and field studies demonstrate that larger photosynthetic isotope fractionation (termed ε_p or $\varepsilon_{p37:2}$ for alkenones specifically) is expressed at higher aqueous CO_2 ($CO_{2(aq)}$) concentrations, and smaller ε_p values are found under lower [$CO_{2(aq)}$]. Published alkenone ε_p records document a dramatic ~15‰ decline in ε_p from the Eocene to Pleistocene, reflecting the global shift from the early Cenozoic greenhouse to modern icehouse (Zachos et al., 2001; Pagani, 2005). However, recent work has demonstrated that CO_2 estimates from alkenone ε_p in the Pleistocene are inconsistent, and, generally disagree with our knowledge of CO_2 variations as recorded in polar ice cores (Badger et al., 2019; Zhang et al., 2019; Stoll et al., 2019).

The vast majority of $\varepsilon_{p37:2}$ records to date have been interpreted using a diffusive model, which assumes that all CO_2 used for photosynthesis arrives at the site of carboxylation via

passive diffusion (Jasper and Hayes, 1990; Bidigare et al., 1997; Zhang et al., 2013). The breadth and discordance of Pleistocene $\varepsilon_{p37:2}$ results have led some to argue that $\varepsilon_{p37:2}$ is less sensitive to $[CO_{2(aq)}]$ than the diffusive model presumes, and that we do not understand the quantitative scaling between $[CO_{2(aq)}]$ and $\varepsilon_{p37:2}$, possibly resulting from non-diffusive CO_2 acquisition by these algae (Badger et al., 2019; Stoll et al., 2019). Others have argued that $\varepsilon_{p37:2}$ changes must be large to extract the low-amplitude $[CO_{2(aq)}]$ variations using the diffusive model (Zhang et al., 2019). Recent culture experiments and core-top sediment reconstructions have found that alkenone $\varepsilon_{p37:2}$ is indeed less sensitive to CO_2 than prescribed by the diffusive model (Stoll et al., 2019; Chapter 2). These studies also find that cell size and irradiance, which are not included in the diffusive model, are more important influences on $\varepsilon_{p37:2}$ than $[CO_{2(aq)}]$. However, there is still little consensus on why $\varepsilon_{p37:2}$ records deviate from the diffusive model, and how to account for changes in irradiance through time.

Here we generate new records of alkenone $\varepsilon_{p37:2}$ and Noelaerhabdaceae coccolith length from two sites in the western Pacific to explore the spatiotemporal controls on alkenone carbon isotope fractionation as well as the CO_2 signature in these molecules. We also compile all published late Pleistocene alkenone $\delta^{13}C$ data, update age models, standardize sea surface temperature estimates, and recalculate $\varepsilon_{p37:2}$ using a common framework. We use the culture-derived multiple linear regression model developed in Chapter 2 to model $\varepsilon_{p37:2}$ in these Pleistocene sequences as a function of $[CO_{2(aq)}]$, cell radius, growth rate, and irradiance. This model provides a framework for quantitatively evaluating the variability in Pleistocene $\varepsilon_{p37:2}$ records, allowing us to identify which locations contain extractable CO_2 information and why different locations perform better than others. First, we review the status of the conventional

diffusive model, as the quantitative relationship between $\varepsilon_{p37:2}$ and $[CO_{2(aq)}]$ has been critiqued in recent publications (Badger et al., 2019; Zhang et al., 2019; Stoll et al., 2019; Chapter 3).

4.3 Background

The conventional alkenone-CO_2 proxy is modified from descriptions of carbon isotope discrimination in plants (Farquhar et al., 1982), and assumes that alkenone-producing coccolithophores acquire CO_2 for photosynthesis via passive diffusion from the ambient environment. This model of carbon isotope fractionation has thus been referred to as the "diffusive model" (Rau et al., 1996; Pagani, 2002). Algal ε_p has been shown to increase with increasing $[CO_{2(aq)}]$ and irradiance, and decrease with increasing growth rate, cell size, and carbon content (Laws et al., 1995; Rau et al., 1996; Laws et al., 1997; Popp et al., 1998b; Burkhardt et al., 1999; Rost et al., 2002; Stoll et al., 2019; Wilkes and Pearson, 2019; Chapter 2). Building on an abundance of field and culture studies examining $\delta^{13}C$ values of algal biomass in relation to $[CO_{2(aq)}]$ and physiology (Rau et al., 1992; Francois et al., 1993; Goericke et al., 1994; Jasper et al., 1994), all influences other than $[CO_{2(aq)}]$ have been consolidated into a single variable ("*b*"), and carbon isotope discrimination is related to CO_2 through the hyperbolic form:

$$\varepsilon_p = \varepsilon_f - b/C_e \qquad \text{[Eq. 1]}$$

where ε_p is the carbon isotope fractionation between biomass and $\delta^{13}C_{CO2aq}$, ε_f is the maximum fractionation by RubisCO, the carboxylating enzyme (typically taken to be 25‰), C_e is the *in-situ* aqueous CO_2 concentration (μmol kg^{-1}), and *b* is the combination of all biological, non-CO_2 effects, including cell size, growth rate, carbon acquisition pathway, membrane permeability, and

others. Under the original diffusive theory, *b* is directly proportional to growth rate and cell radius (Rau et al., 1992; Francois et al., 1993; Laws et al., 1995; Popp et al., 1998b), and the value of *b* defines the slope of the line of ε_p with respect to $1/[CO_{2(aq)}]$. In alkenone ε_p samples from modern suspended particulate matter, Bidigare et al. (1997) found a strong relationship between in-situ phosphate concentrations and *b* values calculated from measured $\varepsilon_{p37:2}$ and in-situ $[CO_{2(aq)}]$. Because cell size was thought to vary little among modern alkenone producers, the large range in *b* and its relationship with phosphate was therefore interpreted to most likely represent growth rate (Bidigare et al., 1997). However, recent reanalysis of the calibration data with an expanded sample set that includes measurements of in-situ growth rates and estimates of cell size from coccolith lengths demonstrates the relationship between *b* and $[PO_4^{3-}]$ arises from an underlying correlation between $[PO_4^{3-}]$ and $[CO_{2(aq)}]$, and that *b* is not correlated with cell size or growth rate (Chapter 3).

After initial apparent success reconstructing $[CO_{2(aq)}]$ over a full glacial cycle (Jasper and Hayes, 1990), the diffusive method was adopted, and has led to an extensive body of Cenozoic alkenone ε_p-based CO_2 estimates (Pagani et al., 1999b; Pagani et al., 2000; Pagani, 2005; Pagani et al., 2011; Badger et al., 2013a; Badger et al., 2013b; Zhang et al., 2013; Bolton et al., 2016; Zhang et al., 2017). Recent works, however, have called into question the fundamental controls on carbon isotope fractionation in these algae (Boller et al., 2011; Stoll et al., 2019; Wilkes and Pearson, 2019; Chapter 2; Chapter 3), and the quantitative relationships embedded in the diffusive model (Badger et al., 2019; Zhang et al., 2019). Stoll et al. (2019) were the first to present multiple linear regression models trained on culture experiments that could predict $\varepsilon_{p37:2}$ as a function of cell radius and the natural logs of $[CO_{2(aq)}]$, light intensity, and DIC. These authors argued that the sensitivity of $\varepsilon_{p37:2}$ to $\Delta[CO_{2(aq)}]$ is much lower than that prescribed by the

modern calibration of the diffusive model, a conclusion supported by the analysis of Zhang et al. (2019), who inverted Pleistocene $\varepsilon_{p37:2}$ data to solve for *b* and found that lower *b* values (lower sensitivity) are required in Pleistocene sediments than those predicted from the modern particulate calibration. Badger et al. (2019) argued that the conventional diffusive model, even when accounting for size as proposed by Popp et al. (1998b) and Henderiks and Pagani (2007) does not accurately capture known CO_2 change in the Pleistocene. Wilkes and Pearson (2019) identified a possible mechanism for the variation in algal (and alkenone) $\delta^{13}C$ by which higher light intensity and nutrient limitation leads to larger ε_p values, requiring that irradiance be considered in algal ε_p interpretations. Stoll et al. (2019) were the first to demonstrate the influence of irradiance on Quaternary and Neogene CO_2 reconstructions. While large growth rate uncertainties were found to contribute to large $[CO_{2(aq)}]$ uncertainties, the approach provides a framework for quantitatively scaling the multiple factors that influence algal carbon isotope fractionation. In Chapter 2, we updated the multilinear regression models to include additional *Gephyrocapsa oceanica* culture data and account for variations in photoperiod duration, yielding a multiple linear model with lower unexplained variance. In Chapter 3, we tested this model on modern $\varepsilon_{p37:2}$ samples from suspended particulate matter and core-tops, demonstrating that alkenone ε_p can be modeled in core-top sediments using these laboratory-derived relationships. Importantly, the empirical model identified a strong dependence of $\varepsilon_{p37:2}$ on irradiance, which is supported by the analysis of ocean particulate and core-top data. This summarizes the development and current state of the alkenone paleo-CO_2 proxy. Here we use the integrated irradiance model developed in Chapter 2 to determine what factors are responsible for the mean and variability of Pleistocene $\varepsilon_{p37:2}$ records. We use the ice-core CO_2 record along with site-specific temperature changes and coccolith size changes to model the $\varepsilon_{p37:2}$ record expected at

each site assuming no changes in irradiance. We then identify sites where measured $\varepsilon_{p37:2}$ and modeled $\varepsilon_{p37:2}$ are linearly related with a slope close to one. This indicates that size and $[CO_{2(aq)}]$ are the dominant parameters driving $\varepsilon_{p37:2}$ variations. By inverting the model and solving for irradiance, we present a framework for identifying locations best suited for paleobarometry and for calibrating the proxy for application beyond the Pleistocene.

4.4 Materials and methods

4.4.1 Site descriptions and age models

We use sediment cores from ODP Site 807 in the western equatorial Pacific and ODP Site 1207 in the northwest Pacific. At present, both sites are close to CO_2 equilibrium with the atmosphere on a mean annual basis (0.5% and ~11% disequilibrium, respectively). Sediments at both sites are nannofossils oozes with good calcium carbonate preservation. Nearby ODP Sites 806 and 1208 have been used for alkenone paleobarometry (Pagani et al., 2010), suggesting these sites might yield reliable alkenone $\delta^{13}C$ records as well.

Sites considered in this compilation come from a variety of oceanographic locations, and the details of each age model and core location can be found in the Supplementary Material. A well-defined chronology is necessary to estimate atmospheric CO_2 during the Pleistocene and determine in-situ $CO_{2(aq)}$ concentrations at each location. For consistency, we reevaluate all published age models, generating new tie points where necessary. We update stratigraphic datums to the GTS2012 timescale, and isotope chronologies to the LR04 stack (Lisiecki and Raymo, 2005; Gradstein et al., 2012). After vetting the data from eleven studies and adding our new data, there are 14 unique sites with $\varepsilon_{p37:2}$ measurements overlapping the ice core CO_2 record. Some sites cannot be included in our analysis: 05PC-21 (Bae et al., 2015) does not have

sufficient information to recalculate $\varepsilon_{p37:2}$ in our consistent framework (see Appendix C for detail), and ODP Sites 1241 (Seki et al., 2010), 882, 982, 1012, and 1208 (Pagani et al., 2010) do not have precise age control in the Pleistocene to identify samples as glacial or interglacial. Below we briefly describe the age models for ODP Sites 807 and 1207, the new records in this study.

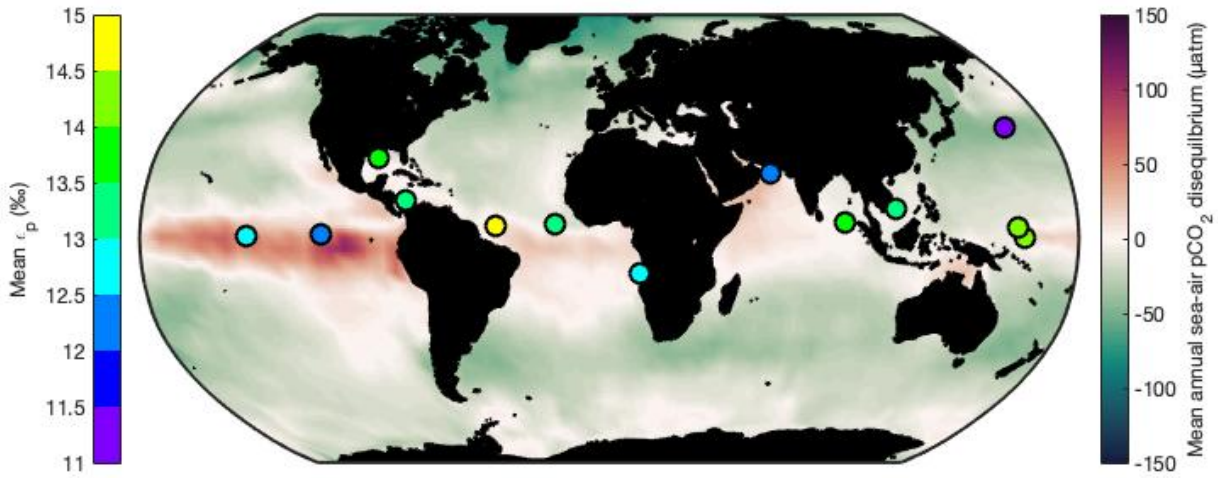

Figure 1. Map of core locations used in this study. The mean $\varepsilon_{p37:2}$ of each time series is shown in colored circles. Core locations overlay a gridded map of the mean annual pCO_2 difference between the surface ocean and atmosphere (ΔpCO_2, µatm) for the reference year of 2005 (Takahashi et al., 2014). Positive values indicate the surface ocean is a source of CO_2 to the atmosphere, and negative values indicate the region is a net CO_2 sink. Landmask was created using the "landmask" function (Greene, 2019). The colormap is from Thyng et al. (2016). Because some sites do not have late Holocene or core-top samples, we use the mean to generalize the data.

4.4.1.1 ODP Site 807

ODP Site 807 is located on the Ontong Java Plateau in the Western Equatorial Pacific (WEP, 3.607°N, 156.625°E, water depth 2805 m), above the modern lysocline of ~3300 m (Prentice et al., 1993). This sediment core has been sampled and studied thoroughly, and there is a reasonably high resolution Pleistocene planktic oxygen isotope stratigraphy using the surface

dwelling planktic foraminifer Trilobatus trilobus (Prentice et al., 1993), and a more detailed benthic foraminifera isotope stratigraphy covering the last ~550 kyr (Zhang et al., 2007). We use the LR04-tuned age model of Zhang et al. (2007) for the last 550 kyr, and update the SPECMAP-tuned age model of Prentice et al. (1993) for 550–800 ka to LR04.

4.4.1.2 ODP Site 1207

ODP Site 1207 is located in the northwest Pacific, situated on the Northern High of Shatsky Rise (37.79°N, 162.75°E). At a water depth of 3103 m, this site is well above the CCD in this region (Berger et al., 1976). No detailed chronostratigraphy has been published for the late Pleistocene section of ODP 1207A. We generate a new age model, initially based on the published biostratigraphic and magnetostratigraphic markers (Shipboard Scientific Party, 2002), which we update from the Cande & Kent 1995 timescale to GTS2012 (Cande and Kent, 1995; Gradstein et al., 2012). We then generate a high-resolution chronology by band-pass filtering and tuning the shipboard core-scanning reflectance data to obliquity variations (Appendix C).

4.4.2 Laboratory Methods

All analytical methods, calculations, and error propagation are described in great detail in Appendix C. All organic geochemistry and stable isotope work were conducted at the Lamont-Doherty Earth Observatory (NY, USA), while morphometric nannofossil work was carried out at the Centre Européen de Recherche et d'Enseignement des Géosciences et de l'Environnement (Aix-en-Provence, France). Briefly, we extracted free lipids from lyophilized sediments via accelerated solvent extraction. The total lipid extract was spiked with an internal standard, and the alkenone fraction was isolated by silica gel flash column chromatography. Alkenone

fractions from ODP 1207 were saponified to remove interfering wax esters. Unsaturation ratios were measured by gas chromatography with a flame ionization detector, and carbon isotope ratios were determined by continuous flow gas chromatography-isotope ratio mass spectrometry. At ODP 807, we use the *T. trilobus* record of Prentice et al. (1993) linearly interpolated to our sample depths. We note that the record shows very little low- or high-frequency variability. At ODP 1207, we picked and analyzed *Orbulina universa* from the same sediment samples as our alkenone measurements. We used *O. universa* because it was the most abundant surface-dwelling foraminifera available in all samples. Coccolith size measurements were made by automated light microscopy (SYRACO) on our ODP 807 and ODP 1207 samples. For several published records, population-mean coccolith size was estimated from nannofossil assemblages (see Appendix C)**.**

4.5 Results

4.5.1 In-situ [$CO_{2(aq)}$] and temperature ranges

Atmospheric CO_2 values from the ice core records at the same ages as our samples cover a range of ~100 ppm, from 188 to ~285 ppm. Alkenone-derived sea surface temperatures (SST) are much warmer at ODP 807 than those at ODP 1207 (Fig. 2), reflecting the general oceanographic setting of each location. The glacial-interglacial SST range, estimated from alkenone unsaturation ratios, is ~1°C at ODP 807, while it is 5.5°C at ODP 1207. SSTs range between 12.29°C and 17.81°C at ODP 1207, and 27.96°C and 28.85°C at ODP 807, which is approaching the upper limit of the alkenone-SST proxy of 28.97°C (Müller et al., 1998). The modern mean annual pCO_2 disequilibrium at ODP 807 is +2.7% (+10 µatm), whereas it is -11% at ODP 1207 (-41 µatm). The estimated in-situ [$CO_{2(aq)}$], applying the same fractional

disequilibrium factor downcore (cf. Supplementary Material section 2.1), ranges from 5.3 to 7.6 μmol kg^{-1} at ODP 807 and from 6.1 to 10.0 μmol kg^{-1} at ODP 1207.

The distribution of atmospheric CO_2 conditions sampled by the alkenone ε_p sample ages in the entire compilation are similar to the mean and distribution of CO_2 values in the ice core record (Fig. S1). This correspondence demonstrates that the dataset is accurately sampling the range of atmospheric CO_2 conditions possible in the Pleistocene. However, alkenone ε_p is sensitive to changes in $[CO_{2(aq)}]$, and not directly to atmospheric CO_2. Temperature and atmospheric CO_2 changes on glacial-interglacial timescales are antagonistic to that goal. Lower atmospheric CO_2 will lead to lower surface ocean $[CO_{2(aq)}]$, but the reduction in temperature increases K_0, attenuating the CO_2 signal. This effect is smaller in warm locations where the relative change in K_0 is smaller, leading to a ~15% attenuation of the atmospheric CO_2 change in the tropics (e.g., ODP 807) as opposed to ~33% attenuation in the mid-latitudes (e.g., ODP 1207). From equilibrium calculations using temperature and coeval atmospheric CO_2, the dataset samples a $[CO_{2(aq)}]$ range of 4.8 to 10.2 μmol kg^{-1} (mean = 6.6 μmol kg^{-1}, 1σ = 1 μmol kg^{-1}). The largest range is found at ODP 1207 (this study) and the smallest range at ODP 668 (Zhang et al., 2019). Across the entire compilation, SST and coeval atmospheric CO_2 are generally well correlated, consistent with a robust body of literature examining the surface ocean temperature response to atmospheric CO_2 change (Hostetler and Mix, 1999; Shakun et al., 2012).

4.5.2 Alkenone ε_p records

We generated two new alkenone ε_p records in the western Pacific. At ODP 807, the stable carbon isotope ratios of the di-unsaturated alkenone ($\delta^{13}C_{37:2}$) varies between -25.4 and -23.6‰.

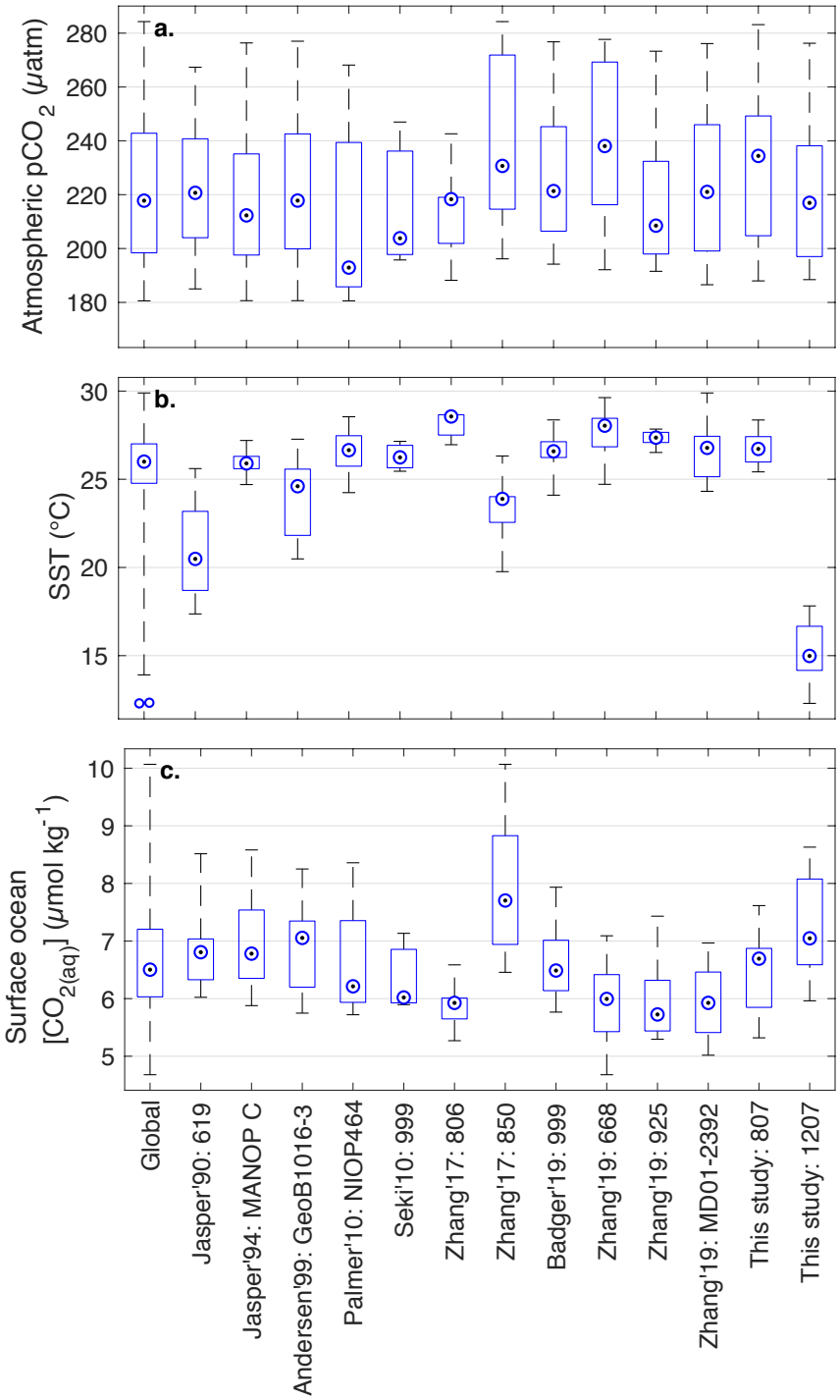

Figure 2. Box-and-whisker plots of environmental data at each core location. (a) ice-core atmospheric CO_2 values sampled by each Pleistocene record, (b) the range of proxy-derived SST values from either the alkenone unsaturation or TEX_{86} proxy, and (c) in-situ $CO_{2(aq)}$ concentrations (product of coeval atmospheric pCO_2 and solubility coefficient), accounting for the modern disequilibrium at each site. The centroid of each box shows the median value, the box edges mark the 25th and 75th percentiles, and the whiskers mark the range of each variable at each site. Small circles beyond the reach of the whiskers show individual data points that are ±2.7 standard deviations from the mean.

Foraminiferal $\delta^{13}C$ ranges from 1.42 to 2.0‰, giving rise to a $\delta^{13}C_{CO2aq}$ range of -6.6 to -6.0‰. Together, these data result in an alkenone ε_p range of 13.4 to 15.0‰ (mean 14.3‰, 1σ = 0.6‰). At ODP 1207, $\delta^{13}C_{37:2}$ ranges from -24.2 to -22.9‰, while $\delta^{13}C_{foram}$ varies from 0.9 to 2.2‰, and estimated $\delta^{13}C_{CO2aq}$ varies between -8.6 and -7.0‰, leading to an alkenone ε_p range of 10.5 to 12.6‰ (mean = 11.4‰, 1σ = 0.6‰). The two locations have similar $\delta^{13}C_{foram}$ and estimated $\delta^{13}C_{DIC}$ values, but the lower temperatures at ODP 1207 result in larger carbon isotope fractionation between DIC and $CO_{2(aq)}$. At ODP 807, the range in estimated $\delta^{13}C_{CO2aq}$ is about one-third of the amplitude of the signal in $\delta^{13}C_{Mk37}$ (0.6‰ versus 1.6‰, respectively). In contrast, the amplitude of the $\delta^{13}C_{CO2aq}$ signal at ODP 1207 is 1.6‰, while alkenone $\delta^{13}C$ varies by only 1.34‰. Because the glacial/interglacial temperature changes are larger at ODP 1207, the equilibrium temperature-dependent fractionation between $\delta^{13}C_{DIC}$ and $\delta^{13}C_{CO2aq}$ also has larger amplitude variability. The more negative $\delta^{13}C_{CO2aq}$ and the relatively more positive $\delta^{13}C_{Mk37:2}$ values at ODP 1207 combine for average lower ε_p values compared to ODP 807. Neither site has a significant correlation between $\delta^{13}C_{Mk37:2}$ and $\delta^{13}C_{CO2aq}$ (ODP 1207: R = 0.02, p = 0.92; ODP 807: R = 0.2, p = 0.26). The only previously published record with a statistically significant linear correlation between $\delta^{13}C_{Mk37}$ and $\delta^{13}C_{CO2aq}$ is MANOP C in the central Equatorial Pacific (Jasper et al., 1994) (R = 0.29, $p < 0.05$). The $\varepsilon_{p37:2}$ signal therefore most likely reflects changes in algal physiology, rather than shifts in the ambient $\delta^{13}C_{DIC}$ pool alone. When all sites from the compilation are considered, $\delta^{13}C_{foram}$ and $\delta^{13}C_{CO2aq}$ range from -0.56 to 2.24‰ and from -8.6‰ to -5.1‰, respectively, while alkenone $\delta^{13}C$ varies from -26.8‰ to -21.3‰. Across the entire late Pleistocene compilation, once $\varepsilon_{p37:2}$ is recalculated in our consistent framework, values range from 9.9 to 16.2‰ (Fig. 4d), with the maximum and minimum both found at DSDP 619 (Jasper

and Hayes, 1990). The smallest mean ε_p values (mean = 11.4‰) are found at ODP Site 1207 (this study), and the largest values (mean = 14.7‰) at ODP Site 925 (Zhang et al., 2013).

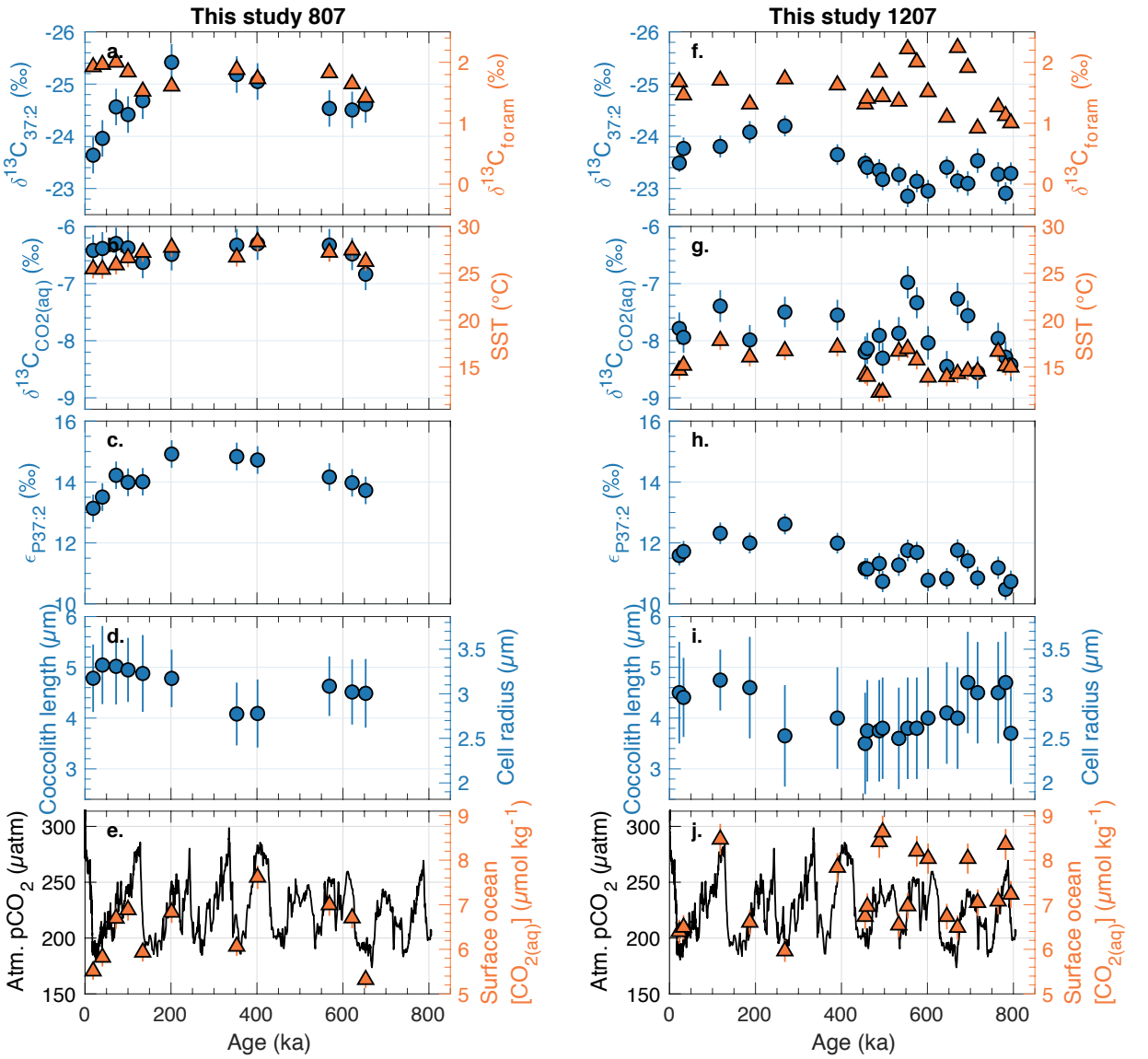

Figure 3. New alkenone carbon isotope records. ODP 806/807 in the Western Equatorial Pacific (a-e) and ODP 1207 in the Northwest Pacific (f-j). (a, f) Measured alkenone carbon isotope ratios (left axis) and measured foraminifera carbon isotope ratios (right axis). (b,g) Estimated dissolved CO_2 carbon isotope ratios (left axis) and measured sea surface temperatures (right axis). (c, h) Calculated $\varepsilon_{p37:2}$ values. (d, i) Measured biomass-weighted mean coccolith length values (left axis) and estimated biomass-weighted mean cell radius (right axis), which is a linear transformation of coccolith length. The standard error of the mean is smaller than the symbol size. (e, j) Global atmospheric CO_2 (left axis) and estimated in-situ $CO_{2(aq)}$ from atmospheric CO_2 and measured SST. Error bars on all data are propagated 1σ uncertainties.

4.5.3 Coccolith size

We measured coccolith length by automated circular-polarized light microscopy (SYRACO) in our new ODP 807 and ODP 1207 samples, from the same material used for alkenone ε_p measurements. We also estimated coccolith length at DSDP 619, and ODP Sites 668B, 925, and 807 from coccolith assemblage data at those core locations or at a core nearby. We compare the assemblage-based results at ODP 807, interpreted from data reported by Chiyonobu et al. (2006), to our actual measurements to evaluate the assemblage approach. Table S2 shows all coccolith data sources, related alkenone ε_p records, and the core locations from which these data derive. All values we report are the biomass-weighted-mean estimate (BMW), which accounts for the greater contribution of alkenones to sediment by larger cells (Chapter 3; see Appendices B and C for details).

At ODP 807, measured BMW-mean coccolith length ranges from 4.1 to 5.1 μm across the time series, while at ODP 1207 values range between 3.5 and 4.8 μm. Across all sites and samples, BMW-mean coccolith length ranges from 2.4 to 5.4 μm. We use the transfer function from Chapter 2, which is based on culture data of McClelland et al. (2016), to relate coccolith size to cell radius. Estimated Noelaerhabdaceae cell radius spans 1.8 μm to 3.3 μm across all sites and samples. Some locations show strong negative or positive correlations between coccolith length and $\varepsilon_{p37:2}$ or $[CO_{2(aq)}]$, though we find no universal relationship across all locations examined (Figs. S3, S4).

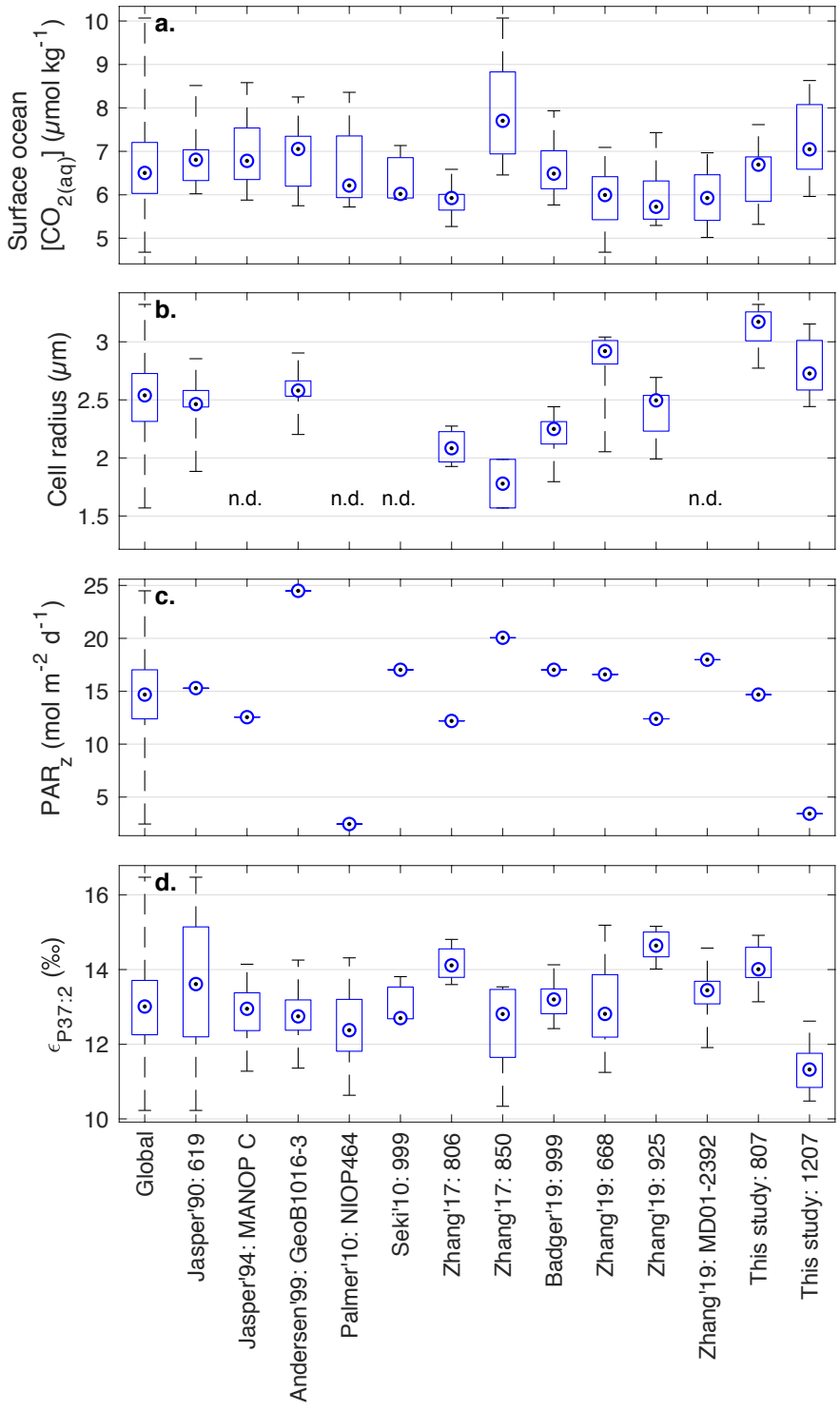

Figure 4. Box-and-whisker plots of the time series at each core site showing the key variables driving alkenone carbon isotope fractionation. (a) Calculated in-situ surface ocean [$CO_{2(aq)}$] as a function of coeval atmospheric CO_2 and proxy-derived temperature, (b) estimated cell radius inferred from a linear transformation of coccolith length, (c) modern estimated mean-annual irradiance at depth, and (d) measured $\varepsilon_{p37:2}$ at each location. As in Fig. 2, the box is centered on the median value, edges mark the 25th and 75th percentiles, and the whiskers mark the full range of measured data.

4.6 Discussion

We generated new records of $\varepsilon_{p37:2}$ and coccolith length from two locations in the western Pacific and compiled all existing Pleistocene $\varepsilon_{p37:2}$ and available related coccolith length data. We found that in many locations, $\varepsilon_{p37:2}$ and in-situ [$CO_{2(aq)}$] are correlated (Fig. S2) as well as $\varepsilon_{p37:2}$ and coccolith length (Fig. S3), suggesting alkenone carbon isotope fractionation is indeed recording past changes in environment and physiology. However, the signs and magnitudes of the observed correlations are only consistent with culture relationships in some cases, as we will examine in detail in the following sections. Culture-based models provide a quantitative framework for evaluating how $\varepsilon_{p37:2}$ should vary given the known in-situ CO_2 change (from atmospheric CO_2 and proxy-derived SST estimates), and cell size change from measured coccolith length. Growth rate is a negligible contributor to $\varepsilon_{p37:2}$ variability in the model, and we ignore it here for that reason (Chapter 2; Appendix A). We use our analyses to identify locations where CO_2 information can be faithfully extracted from $\varepsilon_{p37:2}$ variations (locations where the slope of $\varepsilon_{p37:2}$ measured vs. modeled is close to one), then examine the environmental conditions that are associated with measured vs. modeled slopes of 1, and discuss the prospects for resolving past surface ocean CO_2 changes with this method.

4.6.1 Controls on alkenone carbon isotope fractionation in space and time

In the following subsections we test whether measured $\varepsilon_{p37:2}$ variations agree with $\varepsilon_{p37:2}$ values predicted using the integrated irradiance model from Chapter 2, extending the core-top analysis of Chapter 3. We probe the $\varepsilon_{p37:2}$ response to CO_2 variability first independently (Section 4.6.1.1.), then include size variations as well (Section 4.6.1.2.). We determine irradiance at the depth of alkenone production as in Chapter 3. We calculate the mixed-layer depth at each

location from the gridded output of de Boyer Montégut et al. (2004), and determine mean annual surface irradiance (PAR_{SURF}) and diffuse attenuation coefficients (Kd_{490}) from SeaWiFS and MODIS Aqua gridded data (NASA Goddard Space Flight Center, Ocean Biology Processing Group, 2018a-d). We apply a mean annual photoperiod of 12.05 hours and assume that the integrated alkenone signal in organic matter exported from the photic zone reflects mean annual conditions, as appears to be true in most of the global ocean, especially in the tropics and subtropics (Müller et al., 1998; Rosell-Melé and Prahl, 2013). We set the alkenone production depth at each location as 75% of the mixed layer depth, which was found to be the global average alkenone $\varepsilon_{p37:2}$ signal depth in core-top sediments based on the agreement between modeled and measured $\varepsilon_{p37:2}$ (Chapter 3). Irradiance at depth is calculated as:

$$PAR_z = PAR_{surface} \times e^{-(Kd_{490} \times z)} \quad \text{[Eq. 2]}$$

We note that the prescribed constant irradiance value is simply a baseline initial condition that we hold constant in these analyses, as it does not affect the *slope* of the measured vs. modeled $\varepsilon_{p37:2}$ plots (Figs. 5, 6). Because $\varepsilon_{p37:2}$ is linear with respect to irradiance (Fig. S5), modifying the parameterization of irradiance and production depth would simply vertically shift all data. Variably changing irradiance in each sample, however, would affect the regression slopes. If cell radius can be constrained using measurements of coccolith length, the key question becomes: Where are $\varepsilon_{p37:2}$ variations *not* influenced by irradiance so that the embedded CO_2 signal can be extracted?

4.6.1.1 The influence of $[CO_{2(aq)}]$ on $\varepsilon_{p37:2}$

To examine the relationship between $\varepsilon_{p37:2}$ and $[CO_{2(aq)}]$, we first hold cell radius constant. If all variation in $\varepsilon_{p37:2}$ was simply the result of known CO_2 changes, the slope of the modeled vs. measured $\varepsilon_{p37:2}$ would be 1. Figure 5 shows modeled and measured $\varepsilon_{p37:2}$ at the fourteen sites analyzed in our compilation. Eight have significant slopes ($p<0.1$), all of which are positive, while the rest are indistinguishable from zero. Because $\varepsilon_{p37:2}$ and CO_2 are generally positively correlated (Fig. S2), this model result (Fig. 5) is expected as the model prescribes a positive coefficient to $[CO_{2(aq)}]$, and confirms the model is capturing real $\varepsilon_{p37:2}$ variations resulting from CO_2 alone. No site has a slope close to unity; the closest is ODP 925 (slope = 0.43), indicating that other parameters are driving much of the variability in $\varepsilon_{p37:2}$ when considered in the framework of this culture-based model.

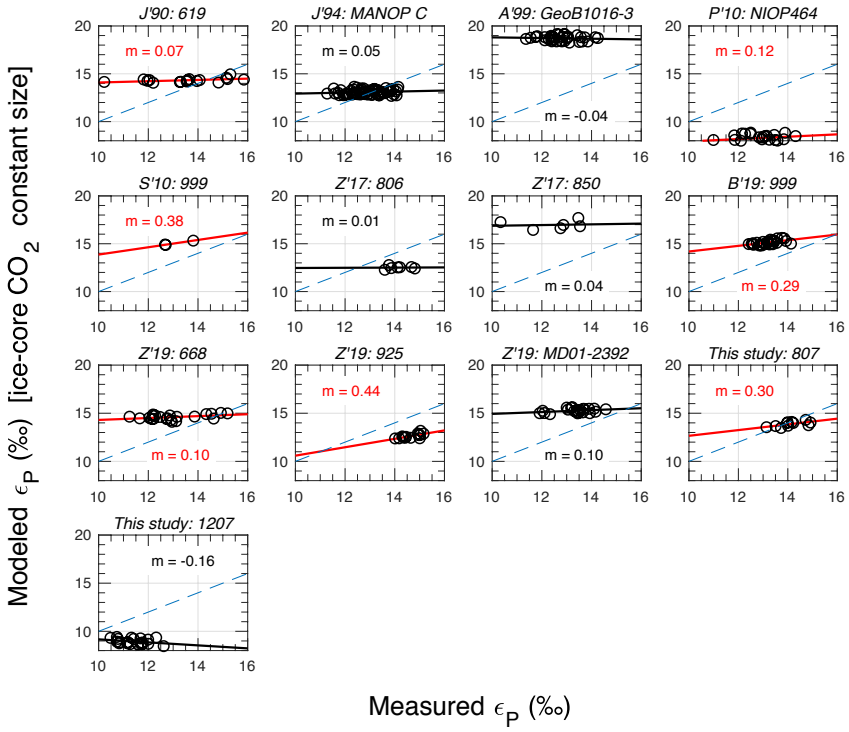

Figure 5. Modeled and measured $\varepsilon_{p37:2}$ holding cell radius and irradiance constant. Axes limits are identical across all panels, and panels are in publication order from top left to bottom right. Solid red lines are linear regressions where the slope is statistically significant ($p<0.1$). Insignificant linear regressions

are shown as black lines. The 1:1 line is shown in all panels as a dashed blue line. Abbreviations above each panel show the first initial of the first author's last name, the last two digits of the publication year, and the site. Author abbreviations are as follows: J'90: Jasper and Hayes (1990). J'94: Jasper et al. (1994). A'99: Andersen et al. (1999). P'10: Palmer et al. (2010). S'10: Seki et al. (2010). Z'17: Zhang et al. (2017). B'19: Badger et al. (2019). Z'19: Zhang et al. (2019).

4.6.1.2 Combined influence of $[CO_{2(aq)}]$ and cell size on $\varepsilon_{p37:2}$

To identify which sites yield the expected sensitivity (modeled vs. measured slope = 1) when all parameters can be constrained to the best degree possible, we model $\varepsilon_{p37:2}$ using the known in-situ CO_2 and reconstructed cell size changes, again holding irradiance constant. We find that three sites have statistically significant ($p<0.1$) modeled vs. measured $\varepsilon_{p37:2}$ relationships: ODP 668 and ODP 925 (Zhang et al., 2019), and ODP 807 (this study), with ODP 668 and ODP 807 within 20% of a slope of one. ODP 806 also shows good agreement, but is only significant at $p<0.25$ as a result of the small sample size. Here we do not combine data from ODP Sites 806 and 807 because coccolith size was measured by two different microscopy techniques. We are also unable to calculate biomass-weighted coccolith size at ODP 806 because coccolith distributions are not reported. At ODP 925, the slope is far too steep (m = 2.46), while all other sites have slopes indistinguishable from zero. We caution that coccolith lengths at ODP 925 were estimated from nearby assemblage data, and suggest that reducing uncertainty by measuring coccolith length may change the slope and bring the modeled and measured data to closer agreement. Our new record at ODP 1207, on the other hand, performs very poorly, with a weak negative relationship between modeled and measured $\varepsilon_{p37:2}$. These findings suggest that measuring coccolith size in conjunction with $\varepsilon_{p37:2}$ will not necessarily allow one to model $\varepsilon_{p37:2}$ with known CO_2 variations. This is the case at DSDP 619 and ODP Sites 925, 999, and 1207, and GeoB1016-3. The implication is that irradiance is varying to such an extent at these locations on glacial/interglacial timescales that it overwhelms the effects of size and CO_2 on $\varepsilon_{p37:2}$.

The offset of modeled $\varepsilon_{p37:2}$ values relative to the 1:1 line suggests the satellite-estimated modern irradiance conditions are either slightly too high or too low. Irradiance at each location is prescribed using surface irradiance measurements and attenuation coefficients, and the production depth (at which to calculate in-situ irradiance) is set to 75% of the modern mixed layer depth. In Chapter 2, we found this to be representative of the global $\varepsilon_{p37:2}$ signal in core-top sediments, albeit with a broad residual distribution that is very similar to the $\varepsilon_{p37:2}$ residuals from the culture calibration. While these uncertainties in the irradiance values can lead to absolute differences in the modeled vs. measured $\varepsilon_{p37:2}$ values, sites with slopes close to one indicate that temperature, coccolith length, and aqueous CO_2 are the only factors changing through time and that changes in $\varepsilon_{p37:2}$ can be modeled with constant irradiance.

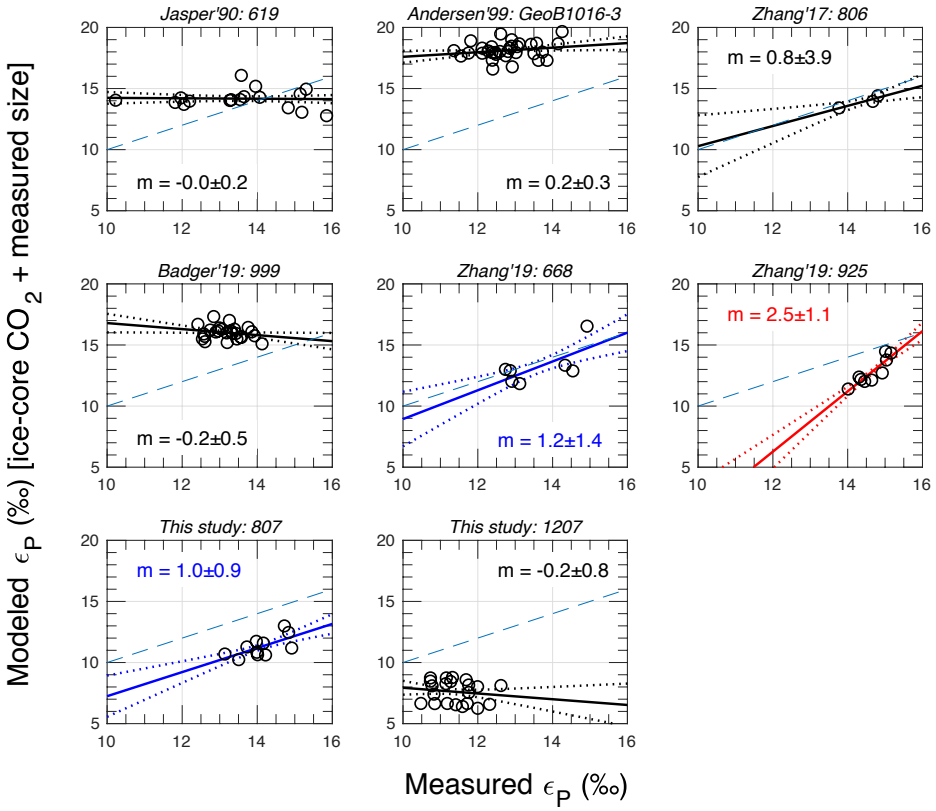

Figure 6. Modeled and measured $\varepsilon_{p37:2}$ using the in-situ [$CO_{2(aq)}$], modern irradiance, and measured cell radius to drive ε_p variations. Black lines have slopes with p>0.1; red lines are significant at p<0.1, and blue lines are significant and the slope is within 1 SE of one. Stippled lines show 68% confidence intervals on slopes.

4.6.2 Site selection for paleobarometry: what determines a well-modeled $\varepsilon_{p37:2}$ record?

The analysis in the previous section (Fig. 6) presents a fundamental criterion for site selection for alkenone paleobarometry using this model: using constant irradiance, the modeled $\varepsilon_{p37:2}$ and measured $\varepsilon_{p37:2}$ relationship should have a slope close to 1. A slope near unity suggests irradiance is stable enough that variations in $\varepsilon_{p37:2}$ reflect size and $[CO_{2(aq)}]$. Variance in the measured vs. modeled relationship comes from measurement and model uncertainty, but primarily from variations in irradiance, which we held constant in all analyses. Because irradiance is the only unconstrained parameter in the model, and given that there is no bona fide proxy for irradiance in the past, it is essential to minimize the influence of this parameter in applications of this model to ancient sediments. To examine the range of irradiance values, we can invert the $\varepsilon_{p37:2}$ model and solve for the irradiance needed to match the measured $\varepsilon_{p37:2}$. If the solved irradiance values in the Pleistocene vary dramatically, assuming stationarity in irradiance conditions deeper in time is highly uncertain and may lead to large systematic errors in reconstructed CO_2. Next, we assess the modern and paleoenvironmental conditions that may lead to stable irradiance conditions through time.

4.6.2.1 Modern oceanography

To evaluate which modern oceanographic parameters may describe a suitable core site with stable irradiance conditions in the Pleistocene, we consider the annual variability in irradiance parameters at each core location. Equation 2 outlines the parameters that could affect irradiance at depth and at a given location through time; it is logical to surmise that high

variability in these parameters during the instrumental record could be indicative of high variability in the past.

We examine the site-by-site "performance" (measured vs. modeled slope) with respect to the monthly-averaged variability in irradiance parameters: mixed-layer depth, the attenuation coefficient (Kd_{490}), and surface irradiance. We find that sites ODP 807, 806, and 668, which have a modeled/measured slope near 1, cluster at the low end of irradiance variability (Fig. 7). Locations that perform poorly, such as ODP 1207 (this study), DSDP 619 (Jasper and Hayes, 1990), and GeoB1016-3 (Andersen et al., 1999) all have high intra-annual variability in at least one of the three parameters that define the irradiance at the depth of alkenone production. In contrast, ODP 999 has low modern variability in irradiance parameters. Although these two sites are stable today, it is possible they experienced greater hydrographic variability in the past, as we will consider in the following section.

We also examine the parameters explaining the mean $\varepsilon_{p37:2}$ value in each timeseries, as changes in the absolute values produced at each location could modulate $\varepsilon_{p37:2}$ values through time. If $[CO_{2(aq)}]$ were the dominant influence on $\varepsilon_{p37:2}$, fractionation should be larger outside the tropics, where in-situ $CO_{2(aq)}$ concentrations are higher resulting from the lower mean sea surface temperatures and higher gas solubility. Instead, we see the highest mean $\varepsilon_{p37:2}$ at Site 925 in the Western Tropical Atlantic, which has very low mean in-situ $[CO_{2(aq)}]$ (Fig. 4). The mean in-situ $[CO_{2(aq)}]$, a function of temperature and the balance of gas exchange and biological productivity (Weiss, 1974; Maier-Reimer and Hasselmann, 1987; Wanninkhof et al., 2009), is significantly inversely correlated with the average $\varepsilon_{p37:2}$ value at each core site (Fig. 8a). This correlation is likely due to the relationship between $[CO_{2(aq)}]$ and nutrient concentrations, which are highly correlated in this dataset and in the global surface ocean, and assuming that the relationship

between CO_2 and nutrient concentrations at one location is relatively stable through the Pleistocene. In direct agreement with observations from the modern ocean (Chapter 3), the ratio of mean annual irradiance to the concentration of the limiting nutrient (adjusted by the Redfield ratio) is a key predictor of the mean $\varepsilon_{p37:2}$ value in a given location (Fig. 8b). The high modern irradiance to nutrient ratio at Site 925 results in a large mean $\varepsilon_{p37:2}$ value for the Pleistocene, suggesting this ratio is generally characterizing the trophic status of a given location and the resulting mean $\varepsilon_{p37:2}$. If the irradiance to nutrient ratio is a direct control on carbon isotope fractionation as suggested by culture experiments (Wilkes and Pearson, 2019), orbital- and millennial-scale variability in irradiance and nutrients could therefore shift $\varepsilon_{p37:2}$ values dramatically. Locations with minimal variability in hydrography, nutrient supply, and upper water column structure are prime targets for stable irradiance regimes and alkenone paleobarometry.

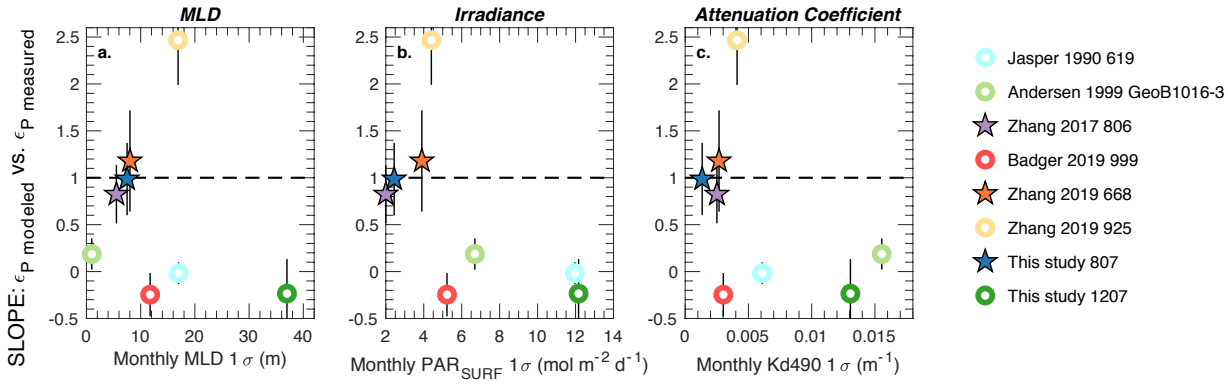

Figure 7. Slopes from Fig. 6 plotted against monthly variability in the parameters that contribute to the calculation of irradiance at depth. (a) mixed layer depth (MLD), (b) surface irradiance, and (c) the diffuse attenuation coefficient (Kd490). Variability is calculated as the standard deviation of the modern monthly climatological mean values from the MODIS Aqua and SeaWiFS gridded datasets, extracted at the latitude and longitude of each core location. Error bars on the slopes are the 1 standard error values shown in Fig. 6. Sites with slopes within 20% of 1 are shown as filled stars; all others are shown as open circles.

4.6.2.2 Paleoceanographic variability

In the previous section we identified modern climatological variables that are associated with better model performance, and found that stability in irradiance parameters is key. Here we examine site-specific paleoceanography to identify what features may indicate whether a site is suitable for CO_2 reconstructions using the $\varepsilon_{p37:2}$ proxy through time. We observed that the range in measured $\varepsilon_{p37:2}$ at each location is strongly correlated with the magnitude of glacial-interglacial SST change at each location (Fig. 8c). While our model does not include any direct temperature effects on $\varepsilon_{p37:2}$, the temperature change is likely indicative of other local dynamical changes, such as changes in upwelling or turbulent mixing, which could affect the weighted-mean depth of alkenone production, the delivery of nutrients to the mixed layer and the turbidity of the water column, and ultimately the irradiance seen by the alkenone-producing population.

We do not suggest that $\varepsilon_{p37:2}$ can be corrected for temperature variations to yield CO_2, nor should the regression shown in Fig. 8c be used quantitatively. Because $\varepsilon_{p37:2}$ partially depends on temperature, we conduct a sensitivity test (Fig. S10) where we hold $\delta^{13}C_{37:2}$ constant in all samples and reevaluate the relationships in Fig. 8. In brief, there is no statistically significant linear correlation between $\Delta\varepsilon_{p37:2}$ and ΔSST (Fig. S10c) or the solved irradiance variability and ΔSST (Fig. S10d). We therefore conclude that the relationships in Fig. 8 are not an artifact of variable SST and insensitive $\delta^{13}C_{37:2}$. Sea surface temperature is one of the most readily accessible types of data in paleoclimate studies, and we argue that it is a blunt but useful tool for generalizing past surface hydrography. Large SST changes likely suggest that the structure of the mixed layer is changing dramatically, which can lead to changes in nutrient availability, the average depth of alkenone production, and the turbidity of the water column, all of which would affect the irradiance conditions seen by the alkenone-producing population.

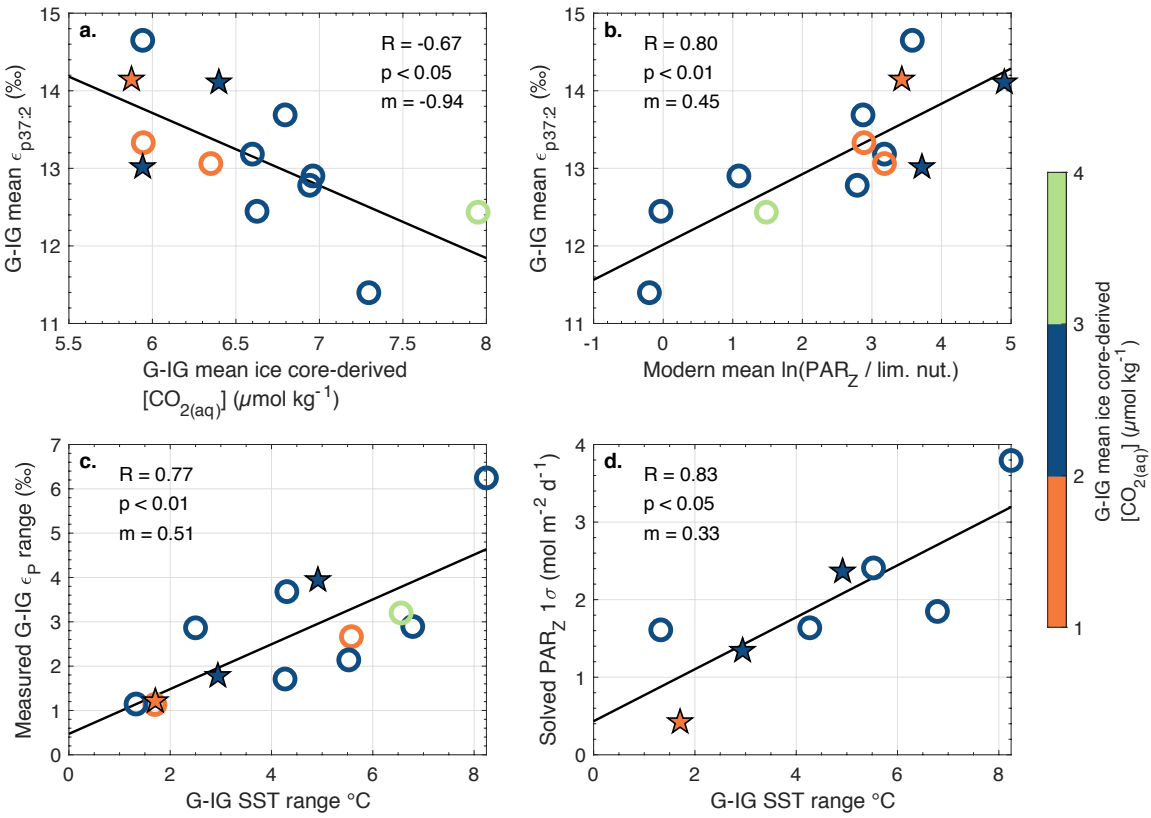

Figure 8. Drivers of glacial-interglacial mean $\varepsilon_{p37:2}$ values and variability. (a) Mean $\varepsilon_{p37:2}$ in each time series regressed against the mean ice core-derived in-situ $CO_{2(aq)}$, showing greater fractionation at lower average CO_2. Symbols are colored by the G-IG CO_2 change. Stars are sites with slopes between 0.8 and 1.2 (ODP 668, 806, 807). (b) Mean $\varepsilon_{p37:2}$ is associated with higher irradiance relative to nutrients, as seen in cultures, surface sediments, and modern particulates (Chapters 2 and 3). (c) Measured $\varepsilon_{p37:2}$ range in each time series is associated with the G-IG SST range, (d) as is the standard deviation of the solved irradiance values.

Alkenone paleobarometry efforts should therefore target locations where glacial-interglacial temperature changes are minimal. Warm, oligotrophic, tropical sites with low-amplitude variability in irradiance parameters are those that appear best suited for paleo-CO_2 reconstructions. However, why is the low SST variability site ODP 999 so different from ODP 807 in terms of observed vs. predicted $\varepsilon_{p37:2}$? ODP Site 999 in the Caribbean Sea is near the mouth of the Magdalena River system, a drainage of the Andes mountains. Fluvial sediment discharge is sensitive to changes in precipitation and the position of the ITCZ, which is known to

migrate on seasonal and orbital timescales (Haug et al., 2001; Mora and Martínez, 2005; McGee et al., 2014). Drying of the Colombian Andes during El Niño years (Poveda et al., 2011) and during glacials reduces river discharge compared to interglacial periods, likely reducing riverine sediment flux as well (Mora and Martínez, 2005). In the Pleistocene, sediment at ODP 999 is generally < 50 % calcium carbonate by weight, and the remainder is largely terrigenous or ash (Peters et al., 2000). The high fraction of aluminosilicates in the sediment suggests the water column may have experienced substantial variations in turbidity, and similar to DSDP 619, led to large changes in integrated mixed-layer irradiance. Martinez et al. (2007) argued for significant changes in surface ocean structure and trophic status at ODP 999 over the last 0.5 Ma associated with changes in the Magdalena River system. At ODP 807, the core material is characterized by ~90% $CaCO_3$ for its entirety, indicating more consistent pelagic sediment sourcing and less allochthonous variability (Shipboard Scientific Party, 1991). It therefore appears that applications of the alkenone paleobarometer should seek out sites with low variability in sediment composition, as changes in sediment composition cam be related to variable surface conditions.

ODP 999 (Badger et al., 2019) is also distinct in that $\varepsilon_{p37:2}$ and coccolith length are positively correlated (Fig. S4). In contrast, cultures show that ε_p and cell size are strongly inversely related (Popp et al., 1998b; Burkhardt et al., 1999). Although this site has the expected directionality between $\varepsilon_{p37:2}$ and CO_2, higher CO_2 is associated with *larger* cell sizes (Henderiks and Pagani, 2007; McClelland et al., 2016). Together, the larger cells compensate for the higher CO_2 concentrations, leading to relatively flat modeled $\varepsilon_{p37:2}$ values, which therefore requires large changes in irradiance to bring the measured and modeled $\varepsilon_{p37:2}$ into agreement.

Our new record from ODP 807 is an example of a location where variability in temperature and irradiance is small. It is a warm, tropical site, just north of the equator in the Western Pacific Warm Pool. Several lines of evidence point to this location being relatively stable in temperature and irradiance during the late Pleistocene. The *T. sacculifer* $\delta^{13}C$ record exhibits low variability, especially compared to ODP 1207 in the Northwest Pacific. Single foraminifera $\delta^{18}O$ analyses at nearby Site 806 show no change in variability of subsurface temperature from the LGM to the Holocene (Ford et al., 2015). More analyses are needed to confirm this lack of trend over other glacial-interglacial transitions, but the existing data support stability on orbital timescales. Furthermore, benthic foraminiferal accumulation rates suggest low and stable productivity from 0.4 to 0.8 Ma (Diester-Haass et al., 2018) at nearby Site 806. As a pelagic site far from the influence of rivers, dust sources, and major oceanographic fronts, this region seems to have the most stable surface hydrography and irradiance parameters over the Late Pleistocene, and is therefore a prime target for applying paleobarometry. ODP 1207, on the other hand, is located near the Kuroshio Extension current, which is the boundary between the subtropical and subpolar gyres in the North Pacific (Zhang and Hanawa, 1993). Diatom assemblages at DSDP 579, less than 1° north of ODP 1207, suggest the ecology and surface ocean hydrography were not stable through the Pleistocene, particularly surrounding the Mid-Brunhes event between ~0.6 and 0.4 Ma (Koizumi, 1985). Southward shifts of the Kuroshio extension during glacial periods could bring subpolar waters to ODP Site 1207, modifying the trophic status of the surface ocean, along with the Noelaerhabdaceae assemblage. Ideal target locations for alkenone paleobarometry are those with an abundance of additional proxy data to evaluate the stability of the surface ocean hydrography through time.

4.6.3 Reconciliation with the diffusive model

Zhang et al. (2019) argued that the best sites for alkenone paleobarometry are those where the glacial to interglacial change in $\varepsilon_{p37:2}$ is large, which they found are sites with higher modern phosphate concentrations. The analysis and findings we present here suggest the opposite is true: the best locations for paleobarometry are warm and oligotrophic ($[PO_4^{3-}] \leq 0.2$), where changes in CO_2 and coccolith size are the primary drivers of $\varepsilon_{p37:2}$ and variations in irradiance are small. This supports the early work and recommendations of Pagani et al. (1999), who focused on oligotrophic sites to minimize the influence of growth rate in their interpretive framework, albeit for a different reason than our model. The contrasting findings of this study and those of Zhang et al. (2019) can be explained by the difference in sensitivity of $\varepsilon_{p37:2}$ to CO_2. A similar argument was laid out in Stoll et al. (2019), which we reinforce here with our updated version of the culture-based multilinear model. In the diffusive model, the relationship between ε_p and $[CO_{2(aq)}]$ is hyperbolic and the shape, or "steepness," equivalent to the sensitivity of ε_p to $[CO_{2(aq)}]$, is *b*-dependent. For example, a typical *b* value applied to an oligotrophic site is ~100; this is approximately the value used by at Zhang et al. (2013) at ODP 925. In the range of Pleistocene $[CO_{2(aq)}]$ variations (~5-10 μmol kg^{-1}), the sensitivity to $CO_{2(aq)}$ is between 1 and 2‰ per μM $CO_{2(aq)}$. This is three- to six-fold higher than the sensitivity derived from the synthesis of available culture observations (Stoll et al., 2019; Chapter 2). This realized lower sensitivity than what is prescribed by the diffusive model calibration was apparent from the back-calculation of the *b* value at each location by Zhang et al. (2019), who demonstrated *b* values in particulate samples are on average greater than those from core-top sediments. Furthermore, *b* values of core-top sediments are also greater than those in Pleistocene sediments. The $\varepsilon_{p37:2}$ change required by the diffusive model for the known CO_2 variations is too large when compared to the

expected sensitivity from culture experiments, requiring that cell size and/or irradiance are varying instead (Stoll et al., 2019).

To illustrate the difference in sensitivity between the two models, we consider the record of Jasper and Hayes (1990) from DSDP 619, where the CO_2 change from MIS 5 to Holocene was reconstructed successfully using the diffusive model. This site, the only record that has been able to match ice-core CO_2 variations with the diffusive model, did so because the glacial-interglacial variation in $\varepsilon_{p37:2}$ was ~4‰ for the ~2 μmol kg^{-1} [$CO_{2(aq)}$] change (Fig. 4). However, our $\varepsilon_{p37:2}$ modeling results show that such large $\varepsilon_{p37:2}$ variability must arise from changes in cell size or irradiance. Estimated cell sizes at DSDP 619 from nannofossil assemblage data are relatively invariant (Fig. 4). Consequently, we must ask: Is there evidence to support large fluctuations in irradiance over the last 100 ka? This core was taken on the continental slope in the Gulf of Mexico. Though it is ~225 km from the coast today, it is only ~100 km from the 120m isobath, a rough estimate of where the LGM shoreline was (Donoghue, 2011). Meckler et al. (2008) found a higher flux of terrigenous material in the nearby Orca Basin in the Gulf of Mexico during glacial periods compared to the Holocene, likely resulting from the shorter glacial distance from river mouth to core site and enhanced glacial erosion. A shorter flowpath and greater erosion could increase turbidity and the attenuation of light in the upper water column, reducing irradiance in the mixed layer. In the analyses shown in Figures 5 and 6, we model $\varepsilon_{p37:2}$ using constant modern irradiance. High measured $\varepsilon_{p37:2}$ values are modeled reasonably well, while low measured $\varepsilon_{p37:2}$ are over-predicted using modern irradiance, in agreement with the proposition that irradiance was lower during glacial periods. As we noted above, target sites for alkenone paleobarometry should be far from modern and ancient shorelines, and as oceanographically and ecologically stable on glacial-interglacial timescales as possible.

4.6.4 Prospects for alkenone paleobarometry

Our analysis in the preceding sections provides a framework for identifying locations best-suited for alkenone paleobarometry. Regressing measured $\varepsilon_{p37:2}$ against modeled $\varepsilon_{p37:2}$ identifies locations where irradiance through time is stable enough that $\varepsilon_{p37:2}$ variations can be interpreted in terms of changes in $[CO_{2aq)}]$ and cell size. Of the sites examined, those on the Ontong Java Plateau in the Western Pacific as well as Site 668 in the eastern tropical Atlantic are well-behaved: irradiance conditions are found to be relatively stable in the Pleistocene, and the relationship between measured and modeled $\varepsilon_{p37:2}$ has a slope very close to one. Because we must assume constant irradiance with some uncertainty in paleobarometry applications, it is essential to demonstrate with Pleistocene samples that $\varepsilon_{p37:2}$ variations can be modeled reasonably well with constant irradiance. As we have shown above, such locations have low intra-annual variability in modern irradiance parameters (Fig. 7), are oligotrophic (Fig. 8b), and have relatively low glacial-interglacial SST change (Fig. 8c). We solve for the irradiance value (PAR_Z, mol m^{-2} d^{-1}) in each sample needed to match the measured $\varepsilon_{p37:2}$ with $\varepsilon_{p37:2}$ modeled using the known $[CO_{2aq)}]$ and cell radius, and find that the standard deviation in solved irradiance is also strongly correlated with the glacial-interglacial SST range, again suggesting temporal SST variability is related to irradiance variability, either through changes in the mean depth of production, or the turbidity of the water column.

To apply the paleobarometer at a core location, we recommend first demonstrating that changes in $\varepsilon_{p37:2}$ can be modeled reasonably well with constant irradiance (measured vs. modeled slope of 1). Because irradiance is a free parameter, we will have to tune for the correct irradiance value using the Pleistocene data. We recommend calibrating irradiance using Pleistocene

samples, and taking the mean and standard error as the envelope of uncertainty on irradiance through time.

An additional challenge for alkenone-based reconstructions is the magnitude of the $[CO_{2(aq)}]$ change we are attempting to resolve using this proxy relative to sources of uncertainty in the proxy parameters. As we describe in detail in Section 3.10 of Appendix C, typical errors in the input parameters produce 1σ uncertainties of ~2.5 μmol $CO_{2(aq)}$ kg^{-1}. The fully propagated $[CO_{2aq)}]$ uncertainty depends on the mean $[CO_{2aq)}]$ at each location, but using the average values of ODP 807 as an example, uncertainty in reconstructed CO_2 would be on the order of ± 90 μatm (1σ). Because standard deviations add in quadrature, distinguishing differences in the mean reconstructed $[CO_{2aq)}]$ between two time periods will require a ~3.5 μmol kg^{-1} separation for 66% confidence that the means are indeed different. Locations or time periods with higher $[CO_{2aq)}]$ will have lower relative CO_2 uncertainties. However, the warm, stable, oligotrophic locations that are the best candidates for CO_2 reconstructions because they minimize irradiance variability will also have low aqueous CO_2 relative to the rest of the surface ocean for a given atmospheric CO_2 level. These competing influences will make it difficult to estimate absolute atmospheric CO_2 values in the Pleistocene or on orbital timescales. Estimating absolute mid-Pliocene pCO_2, for which the most recent foraminiferal $\delta^{11}B$ estimates yield ~400 ± 100 μatm, (Sosdian et al., 2018), is pushing the limits of the proxy. However, changes over the Cenozoic larger than ~200 μatm will be quantifiable, and relative changes at a given location will be easier to reconstruct. Foraminifera-derived $\delta^{11}B$ pH records require a second carbonate system parameter to convert pH to pCO_2. Most existing Neogene records use an assumed evolution of carbonate ion based on reconstructions of the global calcium carbonate compensation depth, or an estimate of whole ocean DIC. By pairing pH from $\delta^{11}B$ with $[CO_{2aq)}]$ change quantified using

alkenone ε_p, there exists an opportunity to use the alkenone-derived [$CO_{2aq)}$] estimates as the second parameter to constrain the temporal evolution of aqueous carbonate parameters such as total dissolved inorganic carbon or alkalinity.

Reducing uncertainty and improving paleo-CO_2 estimates with this method can be achieved with (1) better calibration of the culture-based model through more culture experiments at variable irradiance and cell size, (2) identification of Pleistocene sites where $\varepsilon_{p37:2}$ can be successfully modeled with constant irradiance to improve the calibration and uncertainty envelopes on irradiance values, (3) development of independent irradiance proxies to constrain temporal changes in this unknown parameter, and (4) reconstructing contemporaneous atmospheric CO_2 from multiple sites to reduce the standard error of the mean estimate through time.

4.7 Conclusions

Carbon isotope fractionation recorded in long-chain alkenones has been one of the primary tools for reconstructing ancient CO_2 variations. Recent work using the conventional application of this method have questioned the fundamental quantitative framework for extracting CO_2 information from $\varepsilon_{p37:2}$, while other recent advances have quantified the significance of irradiance on $\varepsilon_{p37:2}$. We have demonstrated with our new data that we can identify locations where $\varepsilon_{p37:2}$ variations are controlled primarily by [$CO_{2(aq)}$] and size, and where the influence of variable irradiance through time can be minimized. We identified modern and paleoceanographic conditions that help explain the mean and amplitude of glacial-interglacial $\varepsilon_{p37:2}$ variations. We revisit the sensitivity of $\varepsilon_{p37:2}$ in the diffusive model compared to our updated multiple regression model, reaffirming that the diffusive model prescribes a too large

slope sensitivity of $\varepsilon_{p37:2}$ with respect to CO_2. We provide a calibration test that can be used to screen locations for deeper-time applications of the alkenone CO_2 proxy: inversion of Pleistocene $\varepsilon_{p37:2}$ records that also constrain cell size should yield relatively invariant irradiance values, and the slope of modeled versus measured $\varepsilon_{p37:2}$ should be close to one. Locations where the solved irradiance values vary wildly in the Pleistocene—when CO_2 and cell radius are constrained—cannot confidently be assumed stable enough in deeper time to extract variations in CO_2. Considering the uncertainties in the measurement and model inversion, we present several recommendations for the best practices when applying the alkenone-CO_2 proxy, as well as avenues for reducing uncertainties and further improving the proxy. Uncertainties associated with inverting the model and solving for CO_2 are large, but the proxy should be able to capture large, long-term changes. Increasing the number of sites and samples over a time interval of interest will reduce the uncertainty in estimated global atmospheric CO_2 and can improve confidence in CO_2 reconstructions.

4.8 Acknowledgements

We thank the International Ocean Drilling Program for providing the sediment samples used in this study. This work was supported by the Lamont Climate Center and the G. Unger Vetlesen Foundation. SRP acknowledges the NSF Graduate Research Fellowship Program (Grant # DGE-16-44869) and the IODP Schlanger Fellowship. This material is based upon research supported by the Chateaubriand Fellowship of the Office for Science & Technology of the Embassy of France in the United States. We thank Wei Huang for assistance with foraminifera stable isotope measurements, and Yves Gally for assistance with microscopy work. We are grateful for Dr. Bärbel Hönisch and Dr. Peter deMenocal for their constructive comments on the manuscript.

4.9 References

Andersen N., Müller P. J., Kirst G. and Schneider R. R. (1999) Alkenone δ13C as a Proxy for Past PCO2 in Surface Waters: Results from the Late Quaternary Angola Current. In *Use of Proxies in Paleoceanography* Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 469–488.

Badger M. P. S., Foster G. L., Chalk T. B., Gibbs S. J., Badger M. P. S., Pancost R. D., Schmidt D. N., Sexton P. F., Mackensen A., Bown P. R. and Pälike H. (2019) Insensitivity of alkenone carbon isotopes to atmospheric CO2 at low to moderate CO2 levels. *Clim. Past* **15**, 539–554.

Badger M. P. S., Lear C. H., Pancost R. D., Foster G. L., Bailey T. R., Leng M. J. and Abels H. A. (2013a) CO2 drawdown following the middle Miocene expansion of the Antarctic Ice Sheet. *Paleoceanography* **28**, 42–53.

Badger M. P. S., Schmidt D. N., Mackensen A. and Pancost R. D. (2013b) High-resolution alkenone palaeobarometry indicates relatively stable pCO2 during the Pliocene (3.3–2.8 Ma). *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **371**, 20130094.

Beerling D. J. and Royer D. L. (2011) Convergent Cenozoic CO2 history. *Nat. Geosci.* **4**, 418–420.

Berger W. H., Adelseck C. G. and Mayer L. A. (1976) Distribution of carbonate in surface sediments of the Pacific Ocean. *J. Geophys. Res.* **81**, 2617–2627.

Bidigare R. R., Fluegge A., Freeman K. H., Hanson K. L., Hayes J. M., Hollander D., Jasper J. P., King L. L., Laws E. A., Milder J., Millero F. J., Pancost R., Popp B. N., Steinberg P. A. and Wakeham S. G. (1997) Consistent fractionation of 13C in nature and in the laboratory: Growth-rate effects in some haptophyte algae. *Global Biogeochem. Cycles* **11**, 279–292.

Boller A. J., Thomas P. J., Cavanaugh C. M. and Scott K. M. (2011) Low stable carbon isotope fractionation by coccolithophore RubisCO. *Geochim. Cosmochim. Acta* **75**, 7200–7207.

Bolton C. T., Hernández-Sánchez M. T., Fuertes M.-Á., González-Lemos S., Abrevaya L., Mendez-Vicente A., Flores J.-A., Probert I., Giosan L., Johnson J. and Stoll H. M. (2016) Decrease in coccolithophore calcification and CO2 since the middle Miocene. *Nat. Commun.* **7**, 10284.

de Boyer Montégut C., Madec G., Fischer A. S., Lazar A. and Iudicone D. (2004) Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *J. Geophys. Res. C Ocean.* **109**, 1–20.

Burkhardt S., Riebesell U. and Zondervan I. (1999) Effects of growth rate , CO2 concentration , and cell size on the stable carbon isotope fractionation in marine phytoplankton. *Geochim. Cosmochim. Acta* **63**, 3729–3741.

Cande S. C. and Kent D. V. (1995) Revised calibration of the geomagnetic polarity timescale for the Late Cretaceous and Cenozoic. *J. Geophys. Res. Solid Earth* **100**, 6093–6095.

Chiyonobu S., Sato T., Narikiyo R. and Yamasaki M. (2006) Floral changes in calcareous nannofossils and their paleoceanographic significance in the equatorial Pacific Ocean during the last 500 000 years. *Isl. Arc* **15**, 476–482.

Diester-Haass L., Billups K. and Lear C. H. (2018) Productivity changes across the mid-Pleistocene climate transition. *Earth-Science Rev.* **179**, 372–391.

Donoghue J. F. (2011) Sea level history of the northern Gulf of Mexico coast and sea level rise scenarios for the near future. *Clim. Change* **107**, 17–33.

Farquhar G. D., O'Leary M. H. and Berry J. A. (1982) On the Relationship Between Carbon Isotope Discrimination and the Intercellular Carbon Dioxide Concentration in Leaves. *Funct. Plant Biol.* **9**, 121.

Foote E. N. (1856) ART. XXXI.--Circumstances affecting the Heat of the Sun's Rays;: (Read before the American Association, August 23d, 1856.). *Am. J. Sci. Arts* **22**, 382.

Ford H. L., Ravelo A. C. and Polissar P. J. (2015) Reduced El Nino-Southern Oscillation during the Last Glacial Maximum. *Science (80-.).* **347**, 255–258.

Francois R., Altabet M. a., Goericke R., McCorkle D. C., Brunet C. and Poisson A. (1993) Changes in the δ13C of surface water particulate organic matter across the subtropical convergence in the SW Indian Ocean. *Global Biogeochem. Cycles* **7**, 627.

Goericke R., Montoya J. P. and Fry B. (1994) Physiology of isotopic fractionation in algae and cyanobacteria. In *Stable isotopes in ecology and environmental science* (eds. K. Lajtha and R. H. Michener). Methods in ecology. Blackwell Scientific Publications, Oxford; Boston. pp. 187–221.

Gradstein F. M., Ogg J. G., Schmitz M. D. and Ogg G. M. (2012) *The Geologic Time Scale 2012*.,

Greene C. A. (2019) landmask. Available at: https://www.mathworks.com/matlabcentral/fileexchange/48661-landmask [Accessed December 17, 2019].

Haug G. H., Hughen K. a, Sigman D. M., Peterson L. C. and Röhl U. (2001) Southward migration of the intertropical convergence zone through the Holocene. *Science* **293**, 1304–1308.

Henderiks J. and Pagani M. (2007) Refining ancient carbon dioxide estimates: Significance of coccolithophore cell size for alkenone-based pCO2 records. *Paleoceanography* **22**, 1–12.

Hostetler S. W. and Mix A. C. (1999) Reassessment of ice-age cooling of the tropical ocean and atmosphere. *Nature* **399**, 673–676.

Jasper J. P. and Hayes J. M. (1990) A carbon isotope record of CO2 levels during the late Quaternary. *Nature* **347**, 462–464.

Jasper J. P., Hayes J. M., Mix A. C. and Prahl F. G. (1994) Photosynthetic fractionation of 13C and concentrations of dissolved CO2 in the central equatorial Pacific during the last 255,000 years. *Paleoceanography* **9**, 781–798.

Koizumi I. (1985) Late Neogene Paleoceanography in the Western North Pacific. In *Initial Reports of the Deep Sea Drilling Project, 86* U.S. Government Printing Office. pp. 0–9.

Laws E. A., Bidigare R. R. and Popp B. N. (1997) Effect of growth rate and CO2 concentration on carbon isotopic fractionation by the marine diatom Phaeodactylum tricornutum. *Limnol. Oceanogr.* **42**, 1552–1560.

Laws E. A., Popp B. N., Bidigare R. R., Kennicutt M. C. and Macko S. A. (1995) Dependence of phytoplankton carbon isotopic composition on growth rate and [CO2]aq: Theoretical considerations and experimental results. *Geochim. Cosmochim. Acta* **59**, 1131–1138.

Lisiecki L. E. and Raymo M. E. (2005) A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. *Paleoceanography* **20**.

Maier-Reimer E. and Hasselmann K. (1987) Transport and storage of CO2 in the ocean-an inorganic ocean-circulation carbon cycle model. *Clim. Dyn.* **2**, 63–90.

Martinez J. I., Mora G. and Barrows T. T. (2007) Paleoceanographic conditions in the western Caribbean Sea for the last 560 kyr as inferred from planktonic foraminifera. *Mar. Micropaleontol.* **64**, 177–188.

McClelland H. L. O., Barbarin N., Beaufort L., Hermoso M., Ferretti P., Greaves M. and Rickaby R. E. M. (2016) Calcification response of a key phytoplankton family to millennial-scale environmental change. *Sci. Rep.* **6**, 1–11.

McGee D., Donohoe A., Marshall J. and Ferreira D. (2014) Changes in ITCZ location and cross-equatorial heat transport at the Last Glacial Maximum, Heinrich Stadial 1, and the mid-Holocene. *Earth Planet. Sci. Lett.* **390**, 69–79.

Meckler A. N., Schubert C. J., Hochuli P. A., Plessen B., Birgel D., Flower B. P., Hinrichs K. U. and Haug G. H. (2008) Glacial to Holocene terrigenous organic matter input to sediments from Orca Basin, Gulf of Mexico - A combined optical and biomarker approach. *Earth Planet. Sci. Lett.* **272**, 251–263.

Mora G. and Martínez J. I. (2005) Sedimentary metal ratios in the Colombia Basin as indicators for water balance change in northern South America during the past 400,000 years.

Paleoceanography **20**.

Müller P. J., Kirst G., Ruhland G., von Storch I. and Rosell-Melé A. (1998) Calibration of the alkenone paleotemperature index U37K′ based on core-tops from the eastern South Atlantic and the global ocean (60°N-60°S). *Geochim. Cosmochim. Acta* **62**, 1757–1772.

NASA Goddard Space Flight Center, Ocean Ecology Laboratory and Ocean Biology Processing Group (2018a) *Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Downwelling Diffuse Attenuation Coefficient Data; 2018 Reprocessing*., Greenbelt, MD, USA.

NASA Goddard Space Flight Center, Ocean Ecology Laboratory and Ocean Biology Processing Group (2018b) *Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Photosynthetically Available Radiation Data; 2018 Reprocessing*., NASA OB.DAAC, Greenbelt, MD, USA.

NASA Goddard Space Flight Center, Ocean Ecology Laboratory and Ocean Biology Processing Group (2018c) *Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Downwelling Diffuse Attenuation Coefficient Data; 2018 Reprocessing*., Greenbelt, MD, USA.

NASA Goddard Space Flight Center, Ocean Ecology Laboratory and Ocean Biology Processing Group (2018d) *Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Photosynthetically Available Radiation Data; 2018 Reprocessing*., Greenbelt, MD, USA.

Pagani M. (2002) The alkenone-CO2 proxy and ancient atmospheric carbon dioxide. *Philos. Trans. A. Math. Phys. Eng. Sci.* **360**, 609–32.

Pagani M., Arthur M. A. and Freeman K. H. (1999a) Miocene evolution of atmospheric carbon dioxide. *Paleoceanography* **14**, 273–292.

Pagani M., Arthur M. A. and Freeman K. H. (2000) Variations in Miocene phytoplankton growth rates in the southwest Atlantic: Evidence for changes in ocean circulation. *Paleoceanography* **15**, 486–496.

Pagani M., Freeman K. H. and Arthur M. A. (1999b) Late Miocene Atmospheric CO2 Concentrations and the Expansion of C4 Grasses. *Science (80-.).* **285**, 876–879.

Pagani M., Huber M., Liu Z., Bohaty S. M., Henderiks J., Sijp W., Krishnan S. and DeConto R. M. (2011) The Role of Carbon Dioxide During the Onset of Antarctic Glaciation. *Science (80-.).* **334**, 1261–1264.

Pagani M., Liu Z., LaRiviere J. and Ravelo A. C. (2010) High Earth-system climate sensitivity determined from Pliocene carbon dioxide concentrations - Supplement. *Nat. Geosci.* **3**, 27–30.

Pagani M., Zachos J. C., Freeman K. H., Tipple B. and Bohaty S. (2005) Marked Decline in

Atmospheric Carbon Dioxide Concentrations During the Paleogene. *Science (80-.).* **309**, 600–603.

Palmer M. R., Brummer G. J. A., Cooper M. J., Elderfield H., Greaves M. J., Reichart G. J., Schouten S. and Yu J. (2010) Multi-proxy reconstruction of surface water pCO2 in the northern Arabian Sea since 29ka. *Earth Planet. Sci. Lett.* **295**, 49–57.

Peters J. L., Murray R. W., Sparks J. W. and Coleman D. S. (2000) Terrigenous matter and dispersed ash in sediment from the Caribbean Sea: results from Leg 165. In *Proceedings of the Ocean Drilling Program, 165 Scientific Results* Ocean Drilling Program. pp. 115–124.

Popp B. N., Kenig F., Wakeham S. G., Laws E. A. and Bidigare R. R. (1998a) Does growth rate affect ketone unsaturation and intracellular carbon isotopic variability in Emiliania huxleyi ? *Paleoceanography* **13**, 35–41.

Popp B. N., Laws E. A., Bidigare R. R., Dore J. E., Hanson K. L. and Wakeham S. G. (1998b) Effect of phytoplankton cell geometry on carbon isotopic fractionation. *Geochim. Cosmochim. Acta* **62**, 69–77.

Poveda G., Álvarez D. M. and Rueda Ó. A. (2011) Hydro-climatic variability over the Andes of Colombia associated with ENSO: a review of climatic processes and their impact on one of the Earth's most important biodiversity hotspots. *Clim. Dyn.* **36**, 2233–2249.

Prentice M. L., Freiz J. K., Simonds G. G. and Matthews R. K. (1993) Neogene Trends in Planktonic Foraminifer 18O from Site 807: Implications for Global Ice Volume and Western Equatorial Pacific Sea-Surface Temperatures. In *Proceedings of the Ocean Drilling Program, 130 Scientific Results* Ocean Drilling Program. pp. 281–305.

Rau G. H., Riebesell U. and Wolf-Gladrow D. (1996) A model of photosynthetic 13C fractionation by marine phytoplankton based on diffusive molecular CO2 uptake. *Mar. Ecol. Prog. Ser.* **133**, 275–285.

Rau G. H., Takahashi T., Desmarais D. J., Repeta D. J. and Martin J. H. (1992) The Relationship between d13C of Organic-Matter and [CO2(aq)] in Ocean Surface-Water: Data from a JGOFS Site in the Northeast Atlantic-Ocean and a Model. *Geochim. Cosmochim. Acta* **56**, 1413–1419.

Rohling E. J., Sluijs A., Dijkstra H. A., Köhler P., van de Wal R. S. W., von der Heydt A. S., Beerling D. J., Berger A., Bijl P. K., Crucifix M., DeConto R. M., Drijfhout S. S., Fedorov A. V., Foster G. L., Ganopolski A., Hansen J., Hönisch B., Hooghiemstra H., Huber M., Huybers P., Knutti R., Lea D. W., Lourens L. J., Lunt D., Masson-Demotte V., Medina-Elizalde M., Otto-Bliesner B. L., Pagani M., Pälike H., Renssen H., Royer D. L., Siddall M., Valdes P., Zachos J. C. and Zeebe R. E. (2012) Making sense of palaeoclimate sensitivity. *Nature* **491**, 683–691.

Rosell-Melé A. and Prahl F. G. (2013) Seasonality of UK'37 temperature estimates as inferred

from sediment trap data. *Quat. Sci. Rev.* **72**, 128–136.

Rost B., Zondervan I. and Riebesell U. (2002) Light-dependent carbon isotope fractionation in the coccolithophorid Emiliania huxleyi. *Limnol. Oceanogr.* **47**, 120–128.

Seki O., Foster G. L., Schmidt D. N., Mackensen A., Kawamura K. and Pancost R. D. (2010) Alkenone and boron-based Pliocene pCO2 records. *Earth Planet. Sci. Lett.* **292**, 201–211.

Shakun J. D., Clark P. U., He F., Marcott S. A., Mix A. C., Liu Z., Otto-Bliesner B., Schmittner A. and Bard E. (2012) Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation. *Nature* **484**, 49–54.

Shipboard Scientific Party (2002) Site 1207. In *Proceedings of the Ocean Drilling Program, 198 Initial Reports* (eds. T. J. Bralower, I. Premoli Silva, and M. J. Malone). Ocean Drilling Program. p. Ch. 3, 1-140.

Shipboard Scientific Party (1991) Site 807. In *Proceedings of the Ocean Drilling Program, 130 Initial Reports* (eds. L. W. Kroenke, W. H. Berger, and T. R. Janecek). Ocean Drilling Program, College Station, TX. pp. 369–493.

Sosdian S. M., Greenop R., Hain M. P., Foster G. L., Pearson P. N. and Lear C. H. (2018) Constraining the evolution of Neogene ocean carbonate chemistry using the boron isotope pH proxy. *Earth Planet. Sci. Lett.* **498**, 362–376.

Stoll H. M., Guitian J., Hernandez-Almeida I., Mejia L. M., Phelps S. R., Polissar P. J., Rosenthal Y., Zhang H. and Ziveri P. (2019) Upregulation of phytoplankton carbon concentrating mechanisms during low CO2 glacial periods and implications for the phytoplankton pCO2 proxy. *Quat. Sci. Rev.* **208**, 1–20.

Takahashi T., Sutherland S. C., Chipman D. W., Goddard J. G., Ho C., Newberger T., Sweeney C. and Munro D. R. (2014) Climatological distributions of pH, pCO2, total CO2, alkalinity, and CaCO3 saturation in the global surface ocean, and temporal changes at selected locations. *Mar. Chem.* **164**, 95–125.

Thyng K. M., Greene C. A., Hetland R. D., Zimmerle H. M. and DiMarco S. F. (2016) True colors of oceanography. *Oceanography* **29**, 9–13.

Wanninkhof R., Asher W. E., Ho D. T., Sweeney C. and McGillis W. R. (2009) Advances in Quantifying Air-Sea Gas Exchange and Environmental Forcing. *Ann. Rev. Mar. Sci.* **1**, 213–244.

Weiss R. F. (1974) Carbon dioxide in water and seawater: the solubility of a non-ideal gas. *Mar. Chem.* **2**, 203–215.

Wilkes E. B. and Pearson A. (2019) A general model for carbon isotopes in red-lineage phytoplankton: Interplay between unidirectional processes and fractionation by RubisCO.

Geochim. Cosmochim. Acta **265**, 163–181.

Young J. R., Geisen M., Cros L., Kleijne A., Sprengel C., Probert I. and Østergaard J. (2003) A guide to extant coccolithophore taxonomy. *J. Nannoplankt. Res.*, 125.

Zachos J., Pagani M., Sloan L., Thomas E. and Billups K. (2001) Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present. *Science (80-.).* **292**, 686–693.

Zhang J., Wang P., Li Q., Cheng X., Jin H. and Zhang S. (2007) Western equatorial Pacific productivity and carbonate dissolution over the last 550 kyr: Foraminiferal and nannofossil evidence from ODP Hole 807A. *Mar. Micropaleontol.* **64**, 121–140.

Zhang R.-C. and Hanawa K. (1993) Features of the water-mass front in the northwestern North Pacific. *J. Geophys. Res. Ocean.* **98**, 967–975.

Zhang Y. G., Pagani M., Henderiks J. and Ren H. (2017) A long history of equatorial deep-water upwelling in the Pacific Ocean. *Earth Planet. Sci. Lett.* **467**, 1–9.

Zhang Y. G., Pagani M., Liu Z., Bohaty S. M. and DeConto R. (2013) A 40-million-year history of atmospheric CO2. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **371**, 20130096.

Zhang Y. G., Pearson A., Benthien A., Dong L., Huybers P., Liu X. and Pagani M. (2019) Refining the alkenone-pCO2 method I: Lessons from the Quaternary glacial cycles. *Geochim. Cosmochim. Acta* **260**, 177–191.

Chapter 5: CO_2 decline in the Neogene accompanied global cooling, C_4 grassland expansion, and decreased carbonate production

5.1 Abstract

With atmospheric CO_2 concentrations above 400 ppm and rising, studying warm periods in the paleoclimate record can help society understand and prepare for the effects of CO_2 levels similar to those predicted for the next several decades. One target is the Miocene (~23–5 million years ago, Ma), which has estimated atmospheric pCO_2 levels between ~200 and 1000 µatm. This epoch bears many signatures of greenhouse forcing: intense global warmth with polar amplification, reduced ice sheet extent, and changes in terrestrial ecosystems that are sensitive to CO_2 variations. Existing CO_2 data for this critical time period in Earth history are notably variable, in part because of the different approaches used to estimate past CO_2. Here we present a record of atmospheric CO_2 for the past ~20 Ma that uses a new culture-based model to calculate CO_2 from geochemical and physiological signatures of algal remains preserved in deep-sea sediments. Our data show a dramatic decrease in surface ocean pCO_2 of ~600 µatm from the Early Miocene to Pleistocene, with a large step-wise drop between ~9 and 8 Ma. Our CO_2 data suggest the global warmth of the Miocene was the result of greenhouse gas forcing. By pairing our CO_2 record with independent records of surface ocean pH, we estimate trends in surface ocean carbonate chemistry, finding higher dissolved inorganic carbon and alkalinity inventories in the Miocene compared to today. We provide first-order estimates of Earth system climate sensitivity, which suggest a coupling between temperature and CO_2 that is similar to what has been observed for other warm periods in the Cenozoic. The timing of CO_2 decline broadly agrees with the timing of C_4 grassland expansion, suggesting CO_2 played a primary role in the radiation of these CO_2-sensitive ecosystems.

5.2 Introduction

Atmospheric CO_2 is currently higher than at any time in the past several million years (Martínez-Botí et al., 2015). The "business-as-usual" scenario of the IPCC (RCP 8.5) estimates atmospheric CO_2 levels of ~930 ppm by 2100 (IPCC, 2013). Understanding the sensitivity of the climate system, cryosphere, and biosphere to atmospheric CO_2 variability during past warm periods is critical for better predicting the magnitude and impact of ensuing anthropogenic warming. For example, the Pliocene epoch (~5.3-2.6 Ma) is characterized by CO_2 levels of around 400 ppm (Martínez-Botí et al., 2015), which resulted in mean global temperatures ~3°C warmer than today (Lunt et al., 2010). As atmospheric CO_2 is likely to exceed maximum Pliocene levels before the end of the century, understanding climate variability under the impending CO_2 conditions requires paleoclimate analysis of warm periods with even higher CO_2 than the Pliocene.

The Miocene (~23 to 5 Ma) was a time of intense global warmth and dramatic global environmental change, during which many present-day biomes took shape. The arctic had periods of ice-free summers (Stein et al., 2016); SSTs were higher by ~10°C in the high latitudes (Herbert et al., 2016); Antarctic ice sheet extent was greatly reduced and Antarctica was home to tundra vegetation; temperate forests existed into regions presently occupied by boreal forests (Pound et al., 2012); grasses employing the C_4 photosystem—an advantage at lower CO_2 levels—emerged and expanded globally between 10 and 6 Ma (Polissar et al., 2019). The marine carbon cycle also underwent significant change. Sedimentological records suggest carbonate burial was higher in the Miocene, possibly resulting from higher weathering fluxes (Suchéras-Marx and Henderiks, 2014; Si and Rosenthal, 2019). Geochemical proxies indicate a more acidic surface ocean in the Miocene compared to preindustrial (Sosdian et al., 2018). However,

atmospheric CO_2 records for this time period are equivocal: recent iterations of the algal carbon isotope proxy suggest substantial CO_2 decline through the late Miocene (Bolton et al., 2016; Mejía et al., 2017; Stoll et al., 2019), while foraminiferal boron isotope records show elevated CO_2 in the Middle Miocene but suggest a decoupling between temperature and CO_2 during the global Late Miocene cooling (Fig. 1) (Foster et al., 2012; Greenop et al., 2014; Sosdian et al., 2018). Better understanding the timing and magnitude of CO_2 decline can shed light on the earth system sensitivity to CO_2 change, the role of CO_2 in the global environmental transitions of the Miocene, and the evolution of the marine carbon cycle.

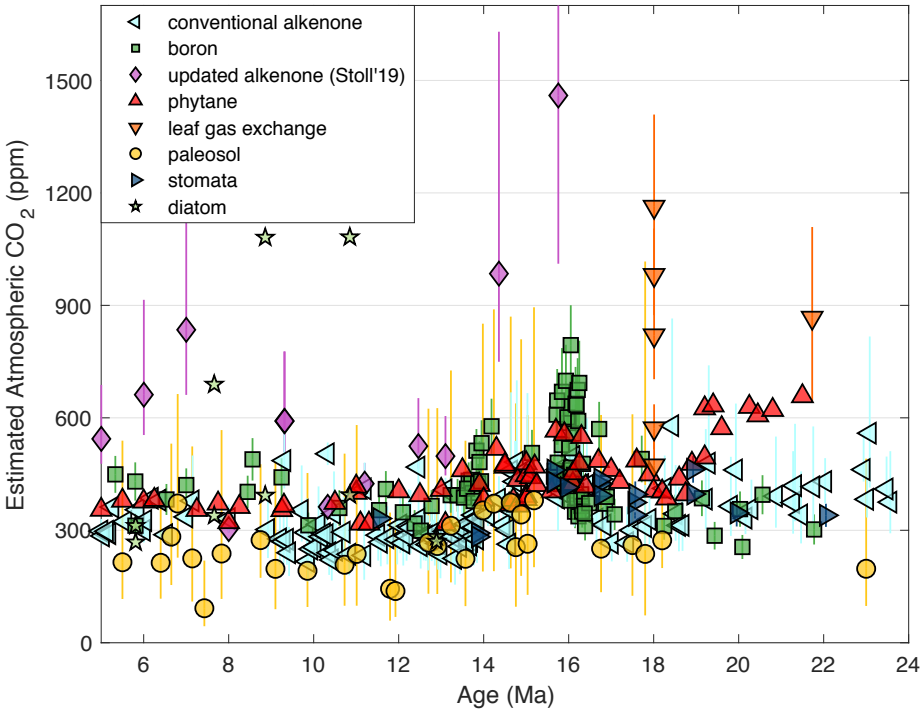

Figure 1. Existing atmospheric CO_2 estimates for the Miocene (~23 – 5 Ma). Proxy data shown use the conventional alkenone diffusive model (Seki et al., 2010; Pagani et al., 2011; Badger et al., 2013; Zhang et al., 2013; Bolton et al., 2016; Zhang et al., 2017; Super et al., 2018); foraminiferal boron isotopes (Sosdian et al., 2018); updated alkenone methods, similar to the approach in this study (Stoll et al., 2019); marine phytane biomarkers (Witkowski et al., 2018); the leaf gas exchange method (Tesfamichael et al., 2017; Londoño et al., 2018); paleosol carbonates (Breecker and Retallack, 2014; Ji et al., 2018); stomatal indices (Kürschner and Kvaček, 2009); and diatom $\delta^{13}C$ (Mejía et al., 2017). Error bars show 84th and 16th percentiles.

Here we present new estimates of atmospheric CO_2 in the Miocene from the organic ($C_{37:2}$ alkenone biomarkers) and inorganic (coccolith) remains of a specific family of coccolithophorid algae, the Noelaerhabdaceae. We use an empirical model of carbon isotope discrimination ($\varepsilon_{p37:2}$) in Noelaerhabdaceae algae to interpret measurements of alkenone $\delta^{13}C$ and coccolith length in terms of past variations in CO_2. We document a dramatic decline in surface ocean pCO_2 of ~600 μatm over the last ~19 million years. Given the large uncertainties associated with estimating CO_2 using this method (roughly ± 100 μatm, 1σ; Chapter 4), our aim here is to examine the secular trends across the Neogene (~23 – 2.6 Ma). We additionally use this new atmospheric CO_2 record in conjunction with published foraminiferal boron isotope ($\delta^{11}B_{foram}$) pH records to determine the temporal evolution of surface ocean carbonate chemistry. We examine the relationship between atmospheric CO_2 and global temperatures, and the role of CO_2 in the Late Miocene expansion of C_4 grasslands. We find that global warmth in the Miocene is associated with greenhouse gas forcing, and that CO_2 decline accompanied the spread of C_4 ecosystems in the Late Miocene.

5.3 Materials and methods

We use data from deep-sea sediment cores from the Western Equatorial Pacific (WEP) obtained by the Ocean Drilling Program. Data from ODP Site 807 and all sample preparation and analytical methods are reported in Chapter 4 and Appendix C. All analytical methods for nearby ODP Site 806 samples were identical; the two sample sets were analyzed in tandem. The laboratory methods are identical to those of Chapter 4, and calculations are described in detail in Appendix C and D. We briefly describe our approach for calculating CO_2 from alkenone ε_p here.

5.3.1 Proxy approach

We estimate [$CO_{2(aq)}$] using the "integrated irradiance" multiple linear regression model developed in Chapter 2. Briefly, alkenone ε_p can be described as a linear combination of cell radius, [$CO_{2(aq)}$], irradiance, and growth rate, though the influence of growth rate is insignificant over the range of physiologically possible values (cf. Chapter 4; Appendix C). Irradiance is a free parameter that varies as a function of the depth of alkenone production, surface irradiance, and the attenuation of photosynthetically available radiation, itself a function of the turbidity of the water column. Because there is no robust paleo-proxy for subsurface irradiance, we solve for the irradiance values needed to match the Pleistocene $\varepsilon_{p37:2}$ variations. At ODP 807, we find irradiance to be stable enough on glacial-interglacial timescales that ε_p can be modeled in the Pleistocene with a constant irradiance value (Chapter 4). We assume that ODP 806 behaves similarly; reanalysis of three samples reported by Zhang et al. (2017) suggests this is likely (Chapter 4). We then take the mean and standard deviation of the solved irradiance values in the Pleistocene to apply downcore as the mean value and uncertainty. Our fundamental argument is that stability on longer timescales is possible if $\varepsilon p_{37:2}$ can be modeled with a constant irradiance value over the large environmental changes associated with high amplitude CO_2 and climate oscillations. In the converse, if irradiance changes dramatically on orbital timescales over the late Pleistocene, there is no basis for assuming stability in older sediments

This site-specific calibration of irradiance is necessary to anchor past changes in [$CO_{2(aq)}$] relative to the Pleistocene. As a sensitivity test, we examine the irradiance variations required for the proxy to produce no apparent atmospheric CO_2 change in the Neogene (Fig. S4; see discussion in Appendix D). We estimate that irradiance would have to be roughly twice as high in the Middle Miocene compared to today to produce the $\varepsilon_{p37:2}$ and size variations that we

observe and have pCO_2 remain constant at 278 μatm. With a modern irradiance attenuation coefficient, this would require the integrated alkenone production depth to shoal to within 5 meters of the ocean surface. While not physically impossible, we argue that large increases in irradiance are unlikely at our core site for the following reasons. The western Pacific warm pool has a relatively deep and stable mixed layer (Andreasen and Ravelo, 1997), has been within ~4° latitude of its current position for the last 20 Ma (Brückmann et al., 1999), and likely had a deeper thermocline in the past compared to today as a result of warmer subsurface temperatures (Ford et al., 2015). Together, shallower production (and greater irradiance at the depth of production) is unlikely from a dynamical standpoint.

5.4 Results

5.4.1 Atmospheric CO_2 decline after the Early Miocene

The $\varepsilon_{p37:2}$ record is relatively stable over the last 8 million years, but shows an overall amplitude of ~3.5‰. There is a local minimum at ~10 Ma, after which $\varepsilon_{p37:2}$ increases steadily to its maximum value at 16.4 Ma during the Miocene Climatic Optimum (MCO), then declines to the minimum value of the record at 18.8 Ma. Biomass-weighted mean coccolith size (see Supplementary Methods) declines from an average of ~7.7 μm in the early Miocene to ~4.7 μm in the Pleistocene. This change is largely driven by the loss of large morphotypes *Cyclicargolithus floridanus* and *Reticulofenestra pseudoumbilicus* in the Middle and Late Miocene (Takayama, 1993). Our measurements capture the Late Miocene *Reticulofenestra* size-reduction event at ~8 Ma—a global phenomenon where large forms of this genus disappeared from the fossil record in the low-latitude oceans (Young, 1990; Takayama, 1993; Kameo and Takayama, 1999; Kameo and Bralower, 2000).

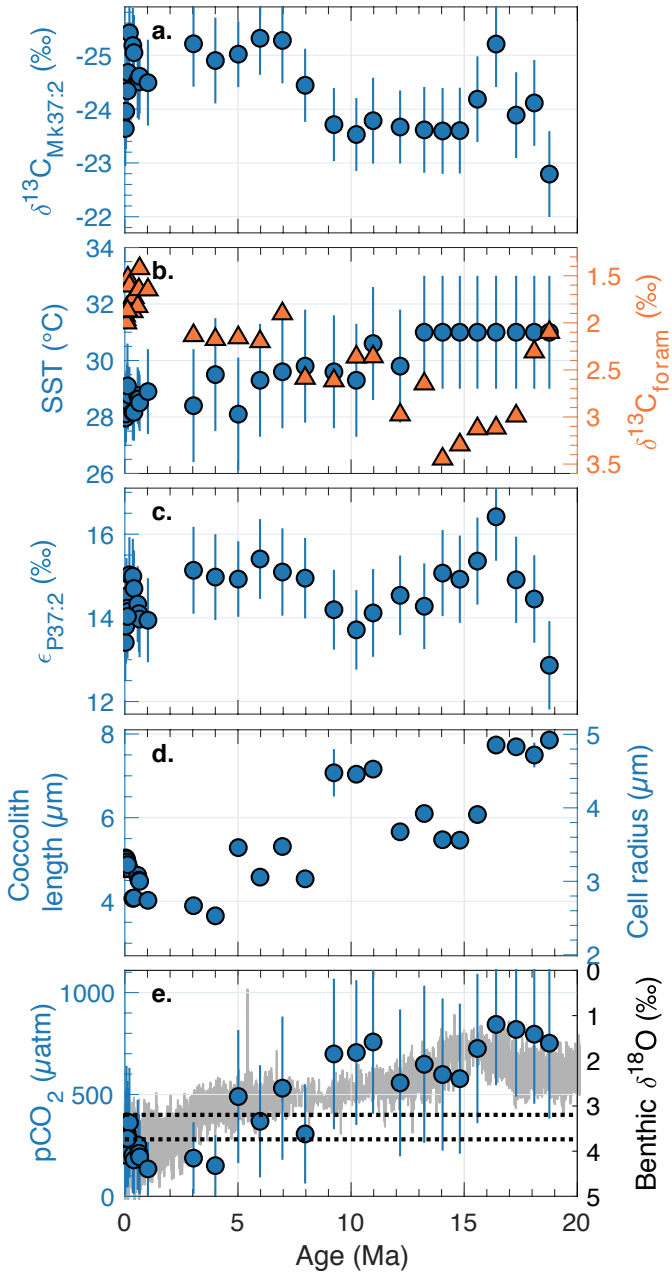

Figure 2. New record of atmospheric pCO_2 through the Neogene. (a) Measured alkenone ($C_{37:2}$ di-unsaturated methyl ketone) carbon isotope ratios; note inverted axis because more negative $\delta^{13}C$ values are associated with higher CO_2. (b) Left, sea surface temperature (SST) estimates from a combination of foraminiferal Mg/Ca and TEX_{86} . Right, planktonic foraminifera $\delta^{13}C$; also note reversed axis because carbon isotope discrimination ($\varepsilon_{p37:2}$) is calculated relative to ambient $\delta^{13}C_{CO2}$ (inferred from $\delta^{13}C_{foram}$), and higher $\varepsilon_{p37:2}$ is associated with higher CO_2. (c) Calculated $\varepsilon_{p37:2}$; larger values indicate greater difference between alkenone $\delta^{13}C$ and seawater $\delta^{13}C$. (d) Measured size of coccoliths from alkenone-producing algae (left); cell radius (right) is a linear function of coccolith length. (e) Calculated pCO_2 required to match the measured $\varepsilon_{p37:2}$ and coccolith length using the integrated irradiance model from Chapter 2. All error bars are 2σ, except for SST (1σ). 2σ errors are used for propagating uncertainty.

The overall trend in $\varepsilon_{p37:2}$ appears to be driven primarily by changes in cell size in our record. Larger cell radii are associated with smaller $\varepsilon_{p37:2}$ values, while higher [$CO_{2(aq)}$] is associated with larger $\varepsilon_{p37:2}$ values. In an alternative scenario, if there were no [$CO_{2(aq)}$] change through our study interval, the larger cell sizes in the Miocene would make the average $\varepsilon_{p37:2}$ from 8 to 20 Ma ~10.2‰ instead of the measured 14.6‰ (Fig. S5). To compensate for the larger cell size and produce the measured $\varepsilon_{p37:2}$, the model requires higher [$CO_{2(aq)}$]. We invert the integrated irradiance model (Chapter 2) to solve for the aqueous CO_2 concentration through time that explains the measured $\varepsilon_{p37:2}$ and coccolith length, assuming constant irradiance with an uncertainty envelope of ~8% (1σ). Our resulting pCO_2 estimates demonstrate a dramatic decrease from the Middle Miocene (average 11–19 Ma = 702 ± 177 µatm, 1σ is mean of 1σ values in bin) to the Plio-Pleistocene (average 0 to 5 Ma = 241 ± 104 µatm, 1σ). We observe a large pCO_2 drop between samples at 9.24 and 7.97 Ma, where pCO_2 declines from ~699 ± 185 µatm to 306 ± 122 (both 1σ).

5.4.2 Consistency with other CO_2 records

Our new CO_2 record for the Neogene is much higher than published estimates using the conventional alkenone CO_2 proxy (Fig. 1). Two records using the conventional model have reliable SST estimates during this interval permitting proper $\varepsilon_{p37:2}$ determinations (Zhang et al., 2013; Super et al., 2018). Together they yield an average pCO_2 in the Middle Miocene (16.5 to 14.5 Ma) of ~410 µatm, and a mean from ~14 to 8 Ma equivalent to preindustrial levels. Neither of these records accounted for coccolith size, which exerts a stronger influence on $\varepsilon_{p37:2}$ than CO_2 (Chapter 2). In the broadest sense, our record is more similar to recent iterations of the algal $\delta^{13}C$ proxy, most of which have tighter constraints on algal physiology, and all of which show higher

pCO_2 in the Middle and Late Miocene compared to the Pleistocene (Bolton et al., 2016; Mejía et al., 2017; Witkowski et al., 2018; Stoll et al., 2019). Our new record and that of Mejía et al. (2017), which uses diatom-bound $\delta^{13}C$ from sediments in the Eastern Equatorial Pacific but a different model to estimate CO_2, both show a pCO_2 increase at ~11 Ma and a pCO_2 decrease after 9 Ma (Fig. 1). The magnitude of the decline is similar (~300 to 400 µatm), but the diatom record suggests a much more dramatic rise (~200 to 800 µatm, depending on which growth rate scenario is used). The diatom $\delta^{13}C$ measurements are from a single diatom size fraction, so cell size variation is not a contributor to the estimated CO_2 variability.

Terrestrial proxies also provide insight to past atmospheric pCO_2 variations (Fig. 1). Londoño et al. (2018) estimated atmospheric pCO_2 at ~18 Ma from fossil leaves using the gas exchange model (Franks et al., 2014), resulting in a bimodal distribution with modes at 528 µatm and 912 µatm. Our pCO_2 estimate at 18.1 Ma overlaps with their reconstruction within uncertainty (804 ± 346 µatm, 2σ). Early Miocene (~22 Ma) fossil leaf deposits from tropical Africa yielded similarly high estimates of ~810 ppm (Tesfamichael et al., 2017), while stomatal index records (Kürschner and Kvaček, 2009) suggest lower CO_2 values (~300-500 ppm) for the Middle and Early Miocene. Estimates from paleosols are similarly diverse, spanning a range of ~50 to 900 ppm in the Middle and Late Miocene (16^{th}-84^{th} percentiles), though recent estimates identify a relative peak during the MCO (Breecker and Retallack, 2014; Ji et al., 2018).

The boron isotope ratios of foraminifera provide high precision and highly-resolved records of marine carbonate chemistry. pCO_2 estimates based on foraminiferal $\delta^{11}B$ pH records depend on the chosen scenario for the boron isotope ratio of seawater ($\delta^{11}B_{sw}$) and the constraint on a second carbonate system parameter, but average ~400 µatm in the Late Miocene (5–11 Ma) and ~460 µatm in the Middle/Early Miocene (11–19 Ma) (Sosdian et al., 2018). The data show

orbital-scale pCO_2 fluctuations on the order of 200 µatm in the Middle Miocene, with several samples yielding values of ~600-800 µatm between 15.5 and 16.5 Ma (Sosdian et al., 2018). In summary, while some existing pCO_2 estimates from a variety of proxies agree with our new record within uncertainty, our alkenone-based pCO_2 estimates are higher than those from foraminiferal boron isotopes for the majority of the Miocene. Below we consider the implications of this mismatch on marine carbonate chemistry.

5.5 Discussion

With our new estimates of the evolution of atmospheric CO_2 through the Neogene, we can attempt to put quantitative constraints on surface ocean carbonate chemistry, the relationship between CO_2 and global climate, and the evolution of terrestrial ecosystems over the last 20 Ma.

5.5.1 Higher Miocene CO_2 requires higher surface ocean alkalinity and dissolved inorganic carbon concentrations

Our record of atmospheric CO_2 provides a single constraint on the global carbon cycle. The marine carbonate system has two degrees of freedom: with knowledge of a second carbonate system parameter, seawater temperature, salinity, pressure, and major-ion composition, we can speciate the dissolved carbon system (Zeebe and Wolf-Gladrow, 2001). To this end, we utilize published records of planktic foraminifera $\delta^{11}B$ (e.g. Sosdian et al., 2018), which is a high-fidelity recorder of seawater $\delta^{11}B_{borate}$ and therefore pH (Hemming and Hanson, 1992; Hönisch and Hemming, 2005). Taking both datasets at face value and accounting for uncertainty in all measured and inferred values, we calculate the probabilistic evolution of the carbonate system using a Monte Carlo routine (see Supplementary Material for full approach in detail). Given the

low resolution of our record and our interest in long-term trends, we pair alkenone-derived pCO_2 values with each $\delta^{11}B_{foram}$ value by taking the average of all pCO_2 values within 0.25 Ma of each $\delta^{11}B_{foram}$ measurement. Then, for a DIC range of 0.2 to 8 mmol kg^{-1}, we speciate the carbonate system and solve for $\delta^{11}B_{SW}$ using the paired $\delta^{11}B_{foram}$ and pCO_2 values. We use such a large range of DIC to account for all possible solutions that occur at the tails of Monte Carlo distributions of the input variables. This exercise yields a system of solutions in DIC and $\delta^{11}B_{SW}$ space for each sample, such as those shown in Fig. 3A-B. We then multiply the independent $\delta^{11}B_{SW}$ probability distribution for each sample (determined from Greenop et al. (2017), Fig. 3A-B, blue distributions) by the bivariate $\delta^{11}B_{SW}$ and DIC distribution to produce the probability distribution of DIC in each sample that agrees with $\delta^{11}B_{foram}$ measurement and our alkenone-based pCO_2 (Fig. 3A-B; red distributions). Though pH is the second carbonate system constraint, we plot DIC against $\delta^{11}B_{SW}$ instead of pH because the absolute pH at any point in time is a function of both $\delta^{11}B_{foram}$ and $\delta^{11}B_{SW}$. The $\delta^{11}B_{SW}$ value prescribes a pH to the $\delta^{11}B_{foram}$ value, which determines the resulting probability distribution of DIC. The crossplots in Fig. 3A-B give two examples of how changing the $\delta 11B_{SW}$ constraint at a given time would affect the calculated DIC. These calculations allow us to evaluate, within uncertainty, the carbonate system conditions that are required to explain both sets of observations.

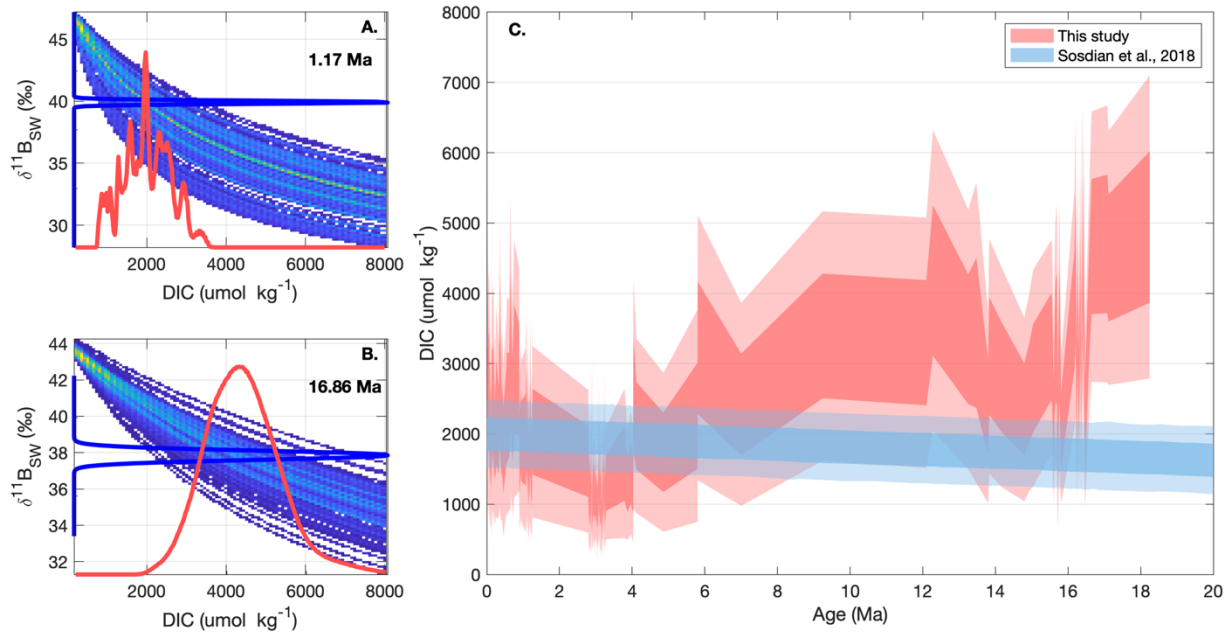

Figure 3. Calculated history of surface ocean dissolved inorganic carbon from new alkenone CO_2 data and existing foraminiferal boron isotope pH estimates. (A, B) Examples of the system of DIC and $\delta^{11}B_{SW}$ values that are compatible with the measured $\delta^{11}B_{foram}$ and estimated atmospheric CO_2, with the spread in values reflecting uncertainties in parameters used in the calculations. Brighter colors in the crossplots relate to the density of solutions within a given bin. The probability distribution of DIC (red) at each time point is calculated using independent estimates of $\delta^{11}B_{SW}$ (blue probability density function in A and B). (C) calculated evolution of surface ocean DIC (pink) and the value applied by Sosdian et al. (2018) (their scenario HO-PA-GR) to calculate pCO_2 from $\delta^{11}B_{foram}$ (blue). Dark shading is 1σ range; light shading is 2σ.

We find that surface ocean DIC (and alkalinity, Fig. S2) was likely higher in the Miocene than today (Fig. 3C). Average solved DIC values in the Late Miocene (5–11 Ma) are around 3.0 mmol kg^{-1}, while those for the Middle/Early Miocene (11–19 Ma) are even higher at roughly 3.4 mmol kg^{-1}. The average for the Plio-Pleistocene (1.8 ± 0.6 mmol kg^{-1}) agrees with pre-industrial DIC estimates (Olsen et al., 2019), suggesting we may reasonably be reconstructing relative changes. Our finding of elevated DIC is at odds with the surface ocean DIC history used for Neogene $\delta^{11}B_{foram}$ pCO_2 reconstructions (Fig. 3C) (Caves et al., 2016; Sosdian et al., 2018; Boudreau et al., 2019; Zeebe and Tyrrell, 2019). The disagreement is intuitive considering basic carbonate chemistry: higher DIC and alkalinity are required to raise the $\delta^{11}B$-based pCO_2

estimates and bring them in line with the higher alkenone-derived pCO_2 values. The implication from our atmospheric CO_2 record is that the carbonate system evolution may be different than what many have modeled for the Neogene.

Is our calculated evolution of the carbonate system supported by independent data? Recent sediment records suggest that pelagic carbonate burial fluxes were 2-3x greater in the Middle and Late Miocene compared to the Pleistocene (Suchéras-Marx and Henderiks, 2014; Si and Rosenthal, 2019). An analysis of the stable strontium isotope ratio of seawater ($\delta^{88/86}Sr_{SW}$), which is sensitive to variations in marine carbonate production, weathering, and burial—among other factors—also suggests carbonate burial fluxes were higher in the early Neogene (Vollstaedt et al., 2014). These higher burial fluxes were achieved despite enhanced dissolution and worse carbonate preservation in the deep sea, as evidenced by improved preservation of foraminifera in the Pleistocene compared to the Miocene (Si and Rosenthal, 2019). Based on the observed ~2.5-fold decline in $CaCO_3$ accumulation rates from ~15 Ma to present at a suite of globally-distributed deep-sea core sites, Si and Rosenthal (2019) argued that alkalinity delivery to the ocean was higher in the Miocene to support such high carbonate fluxes. Because carbonate saturation in the deep ocean has increased since the Middle Miocene (Pälike et al., 2012; Keating-Bitonti and Peters, 2019), and because we estimate ~50% higher alkalinity inventories in the surface ocean during the Middle Miocene, the data imply that the gradient between surface ocean and deep ocean carbonate chemistry was enhanced during the Miocene, a different result than modeling studies have generated (Ridgwell, 2005; Caves et al., 2016; Zeebe and Tyrrell, 2019). The differing conclusions could be explained by the range of input data used for modeling studies, and a shallower calcite saturation horizon in the Miocene (see Supplementary Discussion). While uncertainties in the data allow for a wide range of solutions, we argue that the

general trend of higher surface ocean DIC and alkalinity in the Miocene is a robust conclusion from the existing data.

5.5.2 Miocene global warmth occurs under elevated CO_2 concentrations

The Middle and Late Miocene have hallmarks of global warmth: the Northern Hemisphere was largely unglaciated until ~7 Ma (Larsen et al., 1994) and the Antarctic ice sheet shrank and experienced substantial waxing and waning (Gasson et al., 2016). Sea surface temperature proxies document dramatic cooling over the last ~ 15 Ma, with high latitudes as much as ~10°C warmer in the Middle Miocene compared to today. Our pCO_2 data suggest global warmth was associated with higher than modern pCO_2 levels (Fig. 4), in agreement with recent reconstructions from algal $\delta^{13}C$, foraminifera $\delta^{11}B$, and some terrestrial proxies (Bolton et al., 2016; Tesfamichael et al., 2017; Mejía et al., 2017; Sosdian et al., 2018; Londoño et al., 2018; Stoll et al., 2019). The largest pCO_2 decline in our record (~9–8 Ma) corresponds to an acceleration of global cooling, particularly in the higher latitudes (Fig. 4a), where many existing pCO_2 records show no change (Fig. 1).

What do our new pCO_2 estimates imply for the sensitivity of the Earth system to greenhouse gas forcing? Because our record is low resolution, we limit our inferences here to long-term trends. We take the average of our pCO_2 estimates for the Late Miocene (5–11 Ma) and the Middle/Early Miocene (11–19 Ma), for which global mean annual temperature estimates from pollen and floral assemblage proxy data are 4.5±1.0°C and 7.6±1.2°C (1σ) warmer than today (Fig. 4; Pound et al., 2012; Goldner et al., 2014). For each of these time periods, we calculate the climate sensitivity to a doubling of CO_2 (S_{2xCO2}). This approach treats all non-CO_2 factors—such as ice volume, tectonics, and vegetation—as feedbacks to a CO_2 forcing (Lunt et

al., 2010). We calculate S_{2xCO2} to be 4.6°C ± 1.7°C (1σ) in the Late Miocene and 5.7°C ± 1.2°C (1σ) in the Middle/Early Miocene (Fig. 4C). These findings generally agree with other long-term Earth system sensitivity estimates, including those for the Early Eocene, but are slightly lower than estimates for the Pliocene (Lunt et al., 2010; Martínez-Botí et al., 2015). While relevant for future anthropogenic warming, our long-term Earth system sensitivity estimates are not directly comparable to the Charney (fast-feedback) climate sensitivity estimates imminent for the next several decades (Lunt et al., 2010; Rohling et al., 2012; Royer, 2016). Parsing out the radiative forcing of the slow feedbacks, for which there are reasonable temporal reconstructions, is beyond the scope of this study and is a difficult problem because of synergistic interactions between these slow processes, the magnitude of which can change as a function of the background CO_2 level (Knorr et al., 2011; Bradshaw et al., 2015). However, our new proxy data provide target atmospheric CO_2 levels for the next generation of climate models aimed at determining the magnitude of fast-feedback processes under different boundary conditions.

5.5.3 High atmospheric CO_2 in the Miocene likely supported by volcanic outgassing

The warm climate of the Miocene has been recognized for decades (Savin et al., 1975; Shackleton and Kennett, 1975). Past reconstructions of low (< ~400 ppm) atmospheric CO_2 during the Miocene resulted in arguments of paleogeography, ocean circulation, or vegetation as explanations for the decoupling between climate and CO_2 (Pagani et al., 1999a; Knorr et al., 2011; Foster et al., 2012; LaRiviere et al., 2012). Model studies suggest that Miocene paleogeography, smaller ice sheets, and lower albedo due to more expansive forested areas are responsible for some of the warming (Knorr et al., 2011; Bradshaw et al., 2012; Goldner et al., 2014; Bradshaw et al., 2015), but elevated CO_2 (800 ppm) provides the best agreement between

modeled and proxy temperatures. Our new data, in tandem with a handful of recent CO_2 estimates for the Miocene, suggest greenhouse forcing by CO_2 is a primary explanation for the higher global temperatures (Bolton et al., 2016; Tesfamichael et al., 2017; Mejía et al., 2017; Sosdian et al., 2018; Witkowski et al., 2018; Londoño et al., 2018; Stoll et al., 2019).

If atmospheric CO_2 in the Middle Miocene was indeed 2-3x preindustrial values, as our data suggest, what sustained such high steady-state CO_2? The primary fluxes balancing the CO_2 content of the ocean-atmosphere system are volcanism, and the burial and weathering of organic carbon, carbonate, and pyrite (Derry and France-Lanord, 1996; Zeebe and Caldeira, 2008; Torres et al., 2014). Tectonics, outgassing, and CO_2 are linked on various timescales. Periods of higher volcanic activity are associated with warmer temperatures and higher CO_2, for example in the Cretaceous (Lee et al., 2013), Paleozoic, and Mesozoic (McKenzie et al., 2016). A compilation of crustal rift length shows global rift length was ~2-3 times higher in the early Neogene compared to modern, which is likely to result in higher volcanic CO_2 fluxes to the ocean-atmosphere system (Brune et al., 2017). Furthermore, the Columbia River Flood basalts were emplaced during the Middle Miocene, and about 200,000 km^3 of lava erupted between ~16.7 Ma and 15.9 Ma (Kasbohm and Schoene, 2018). In sum, we find that the global thermal evolution of the last ~20 Ma is consistent with our atmospheric CO_2 record and may have been supported by enhanced volcanic outgassing during the middle Miocene compared to today.

5.5.4 CO_2 decline accompanied the expansion of C_4 ecosystems in the Late Miocene

Terrestrial ecosystems saw dramatic floral and faunal turnover throughout the Late Miocene, as global cooling and environmental change transformed biomes around the world. The ancient forests and wooded grasslands of the tropics and subtropics were replaced with savanna

ecosystems that dominate these regions today. The ecosystem changes cascaded through fauna, leading to the diversification of mammals, their anatomy, and their diets on multiple continents (Cerling et al., 2005; MacFadden, 2005; Uno et al., 2011; Gordon and Prins, 2019). In particular, abundant C_4 grasslands emerged at ~10 Ma across northern Africa and dominated the landscape by ~5 Ma (Feakins et al., 2013; Uno et al., 2016; Polissar et al., 2019). The expansion of C_4 grasslands occurred on the Indian subcontinent around 8 Ma (Quade and Cerling, 1995; Freeman and Colarusso, 2001), and in Southern Africa, China, and South and North America by ~6.5 Ma (Passey et al., 2009; Hoetzel et al., 2013; Dupont et al., 2013) (Fig. 4B). C_4 vegetation has a competitive advantage over C_3 vegetation at higher growing season temperatures, higher irradiance, more arid or seasonably-variable precipitation conditions, and at lower CO_2 concentrations (Ehleringer, 1978). Original explanations for the expansion of C_4 ecosystems favored a decline in atmospheric CO_2 (Ehleringer et al., 1997; Cerling et al., 1997). However, early CO_2 estimates argued against a role for CO_2 in this transition (Pagani et al., 1999b), which seemed to be supported by additional proxy reconstructions (Henderiks and Pagani, 2007; Zhang et al., 2013).

We document a decline in CO_2 from ~11.0 to ~9.2 Ma of ~45 μatm Myr^{-1}, and drop from ~700 μatm at 9.2 Ma to roughly preindustrial levels at ~8.0 Ma, coincident with the global expansion of C_4 ecosystems (Fig. 4B-C). C_3 plants show a steep reduction in photosynthetic quantum yield (~mol CO_2 fixed per photon) as growing season temperature increases and ambient CO_2 decreases, while C_4 photosynthesis is insensitive to variations in these parameters (Ehleringer and Björkman, 1977). From a simple model of photosynthetic quantum yield accounting for this temperature- and CO_2-dependence in C_3 grasses, Ehleringer et al. (1997) predicted that C_4 plants should be favored over C_3 plants below ~600-700 ppm CO_2 at a growing

season temperature of ~35°C (cf. their Fig. 2: monocot crossover curves). Northern East and West Africa, which today have daytime temperatures of ~30°C (Guichard et al., 2009), could have reasonably had temperatures ~35°C at 10 Ma, given the global sea surface temperature anomalies for this time period (Herbert et al., 2016). Although this is a very simplified treatment of vegetation dynamics, the emergence of C_4 grasslands at ~10 Ma when our pCO_2 estimates are 700±180 μatm (1σ) is therefore broadly in agreement with biophysical predictions.

Our new data yield a pCO_2 estimate at ~12.2 Ma of 550±180 μatm (1σ), yet C_4 vegetation did not appear in measurable abundance in the subtropics until ~10 Ma, when our reconstructed pCO_2 value is slightly higher. The vegetation that exists on a landscape—particularly in places that can support savanna ecosystems—is a complex function of seasonality, precipitation, fire, herbivory, soil composition, competitive dynamics, and phylogeny (Lehmann et al., 2011). For example, much of South America has the temperature and moisture conditions to support larger savanna ecosystems than currently exist, possibly due to poor soils (Lehmann et al., 2014). Lower CO_2 is likely a necessary precondition for the expansion of C_4 grasses, but other local and regional factors may have shaped the trajectories on each continent (Griffith et al., 2015). The simplest explanation for the delayed emergence of C_4 ecosystems is that atmospheric CO_2 ~12-14 Ma was in fact towards the upper distribution of our uncertainty envelope. Increasing the number of sites and the temporal resolution of estimates could provide tighter constraints on the evolution of CO_2 through this interval and test this hypothesis. However, the timing of the global C_4 expansion agrees extremely well with our documented CO_2 decline in the Late Miocene, supporting the original hypothesis of Cerling et al. (1997).

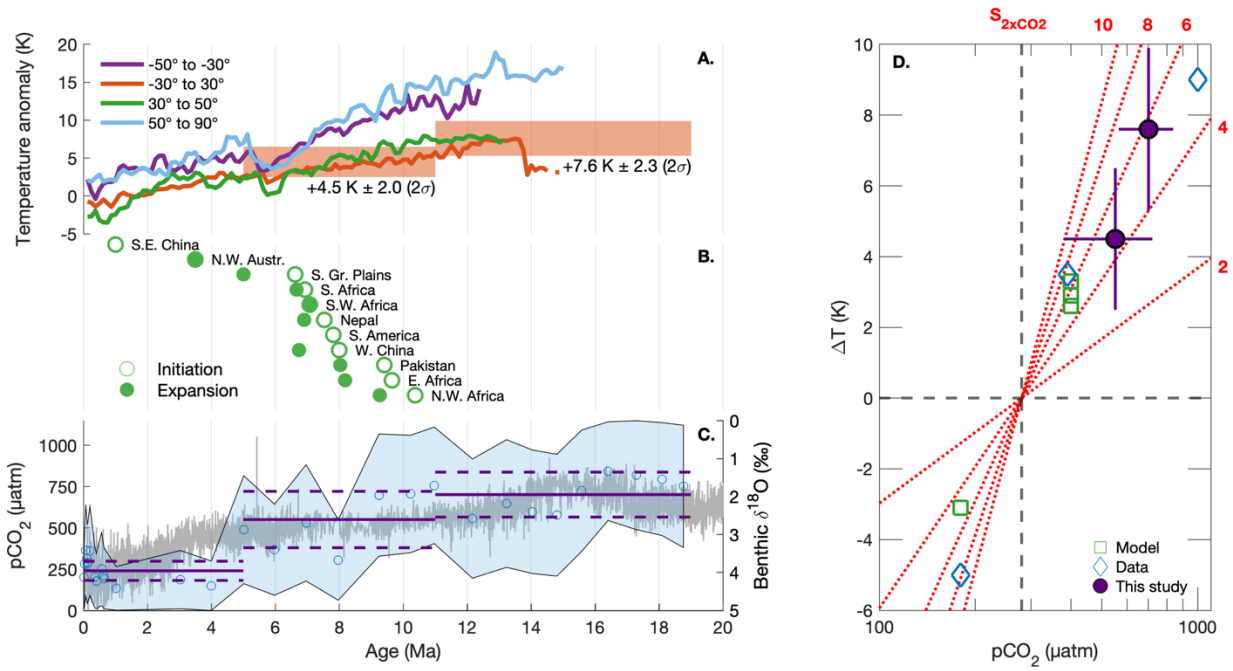

Figure 4. Atmospheric CO_2, global temperatures, C_4 grassland expansion, and earth system sensitivity in the Neogene. (A) Estimates of mean annual sea surface temperature, averaged by latitudinal bin (Herbert et al., 2016). Red boxes are Late Miocene (5–11 Ma) and Middle/Early Miocene (11–19 Ma) global mean annual temperature estimates from fossil pollen and floral assemblage data (Pound et al., 2012; Goldner et al., 2014). (B) Timing of the initiation (open circles) and expansion (filled circles) of C_4 ecosystems at different global localities. Data are from Polissar et al. (2019). (C) pCO_2 (this study) derived from alkenone $\delta^{13}C$ and coccolith records from Western Equatorial Pacific sediment cores (same as Fig. 2e). To compare our estimates to global temperatures, we take the mean CO_2 (± 2 SEM) of our new data during the same time periods for which the global temperature estimates are derived. Shown also is the global benthic foraminifera $\delta^{18}O$ record of Zachos et al. (2008). (D) Global average temperature change with respect to average CO_2 (log-scale) for the Late Miocene and Middle/Early Miocene (filled purple circles). Dashed gray lines show ΔT and ΔCO_2 zero values (CO_2 of equal to 278 ppm, the preindustrial average used here). Red dashed lines show slopes for a given climate sensitivity (temperature change for a doubling of CO_2). Open green squares show model-based estimates of temperature change and CO_2 change for LGM (Bradshaw et al., 2015), Pliocene (Lunt et al., 2010), and Late Miocene time slices (Bradshaw et al., 2015). Open blue diamonds show data-based sensitivity estimates for the Pliocene (Pagani et al., 2010), and LGM and Eocene (Hansen et al., 2008).

5.6 Conclusion

Records of algal carbon isotope fractionation and cell size can provide key insights to past variation in atmospheric CO_2. We document a large pCO_2 decline from a maximum of ~850 µatm in the Middle Miocene to an average of ~250 µatm in the Pleistocene, a change of ~600 µatm. The elevated CO_2 levels of the Miocene supported global temperatures ~4 to 8 K above

preindustrial. Though uncertain, our estimates of Miocene Earth system climate sensitivity are ~3 to 7 K per CO_2 doubling, which is broadly consistent with other warm intervals in the Cenozoic. Our new pCO_2 estimates for much of the Late and Early Miocene contrast with those derived from the boron isotope pH proxy, while they are more consistent during the Middle Miocene. Taking our pCO_2 record and the surface ocean pH record together implies higher dissolved inorganic carbon concentrations in the Miocene compared to the Pleistocene. We find that the timing of major pCO_2 decline corresponds with the expansion of C_4 grasslands across the globe, lending support to original hypotheses for the emergence of these iconic ecosystems. Ultimately, we find CO_2 was a key driver of environmental change through the Neogene.

5.7 Acknowledgements

This work was supported by the Lamont Climate Center and the G. Unger Vetlesen Foundation. SRP acknowledges the NSF Graduate Research Fellowship Program (Grant # DGE-16-44869) and the IODP Schlanger Fellowship. This material is based upon research supported by the Chateaubriand Fellowship of the Office for Science & Technology of the Embassy of France in the United States. We thank the International Ocean Drilling Program for providing sediment samples. We thank Yves Gally for SYRACO microscopy assistance, Dr. Wei Huang for stable isotope assistance, and Jason Swann and Nicole deRoberts for laboratory assistance. We are very appreciative of constructive comments on the manuscript from Dr. Bärbel Hönisch, and for thoughtful conversations with Dr. Mathis Hain.

5.8 References

Andreasen D. J. and Ravelo A. C. (1997) Tropical Pacific Ocean thermocline depth reconstructions for the Last Glacial Maximum. *Paleoceanography* **12**, 395–413.

Badger M. P. S., Lear C. H., Pancost R. D., Foster G. L., Bailey T. R., Leng M. J. and Abels H. A. (2013) CO2 drawdown following the middle Miocene expansion of the Antarctic Ice Sheet. *Paleoceanography* **28**, 42–53.

Bolton C. T., Hernández-Sánchez M. T., Fuertes M.-Á., González-Lemos S., Abrevaya L., Mendez-Vicente A., Flores J.-A., Probert I., Giosan L., Johnson J. and Stoll H. M. (2016) Decrease in coccolithophore calcification and CO2 since the middle Miocene. *Nat. Commun.* **7**, 10284.

Boudreau B. P., Middelburg J. J., Sluijs A. and van der Ploeg R. (2019) Secular variations in the carbonate chemistry of the oceans over the Cenozoic. *Earth Planet. Sci. Lett.* **512**, 194–206.

Bradshaw C. D., Lunt D. J., Flecker R. and Davies-Barnard T. (2015) Disentangling the roles of late Miocene palaeogeography and vegetation - Implications for climate sensitivity. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **417**, 17–34.

Bradshaw C. D., Lunt D. J., Flecker R., Salzmann U., Pound M. J., Haywood A. M. and Eronen J. T. (2012) The relative roles of CO2 and palaeogeography in determining late Miocene climate: Results from a terrestrial model-data comparison. *Clim. Past* **8**, 1257–1285.

Breecker D. O. and Retallack G. J. (2014) Refining the pedogenic carbonate atmospheric CO2 proxy and application to Miocene CO2. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **406**, 1–8.

Brückmann W., Cepek M., Hay W. W., Matschkowski T., Söding E., Spieß V., Thiede J., Tiedemann R. and Wefer G. (1999) The Ocean Drilling Stratigraphic Network (www.odsn.de). *Joides J.* **25**, 25–27.

Brune S., Williams S. E. and Müller R. D. (2017) Potential links between continental rifting, CO2 degassing and climate change through time. *Nat. Geosci.* **10**, 941–946.

Caves J. K., Jost A. B., Lau K. V. and Maher K. (2016) Cenozoic carbon cycle imbalances and a variable weathering feedback. *Earth Planet. Sci. Lett.* **450**, 152–163.

Cerling T. E., Harris J. M. and Leakey M. G. (2005) Environmentally Driven Dietary Adaptations in African Mammals. In *A History of Atmospheric CO2 and Its Effects on Plants, Animals, and Ecosystems* (eds. J. R. Ehleringer, T. E. Cerling, and M. D. Dearing). Springer Verlag, New York. pp. 258–272.

Cerling T. E., Harris J. M., Macfadden B. J., Leakey M. G., Quadek J., Eisenmann V., Ehleringer J. R., Quade J., Eisenmann V. and Ehleringer J. R. (1997) Global vegetation change through

the Miocene/Pliocene boundary. *Nature* **389**, 153–158.

Derry L. A. and France-Lanord C. (1996) Neogene growth of the sedimentary organic carbon reservoir. *Paleoceanography* **11**, 267.

Dupont L. M., Rommerskirchen F., Mollenhauer G. and Schefuß E. (2013) Miocene to Pliocene changes in South African hydrology and vegetation in relation to the expansion of C4 plants. *Earth Planet. Sci. Lett.* **375**, 408–417.

Ehleringer J. and Björkman O. (1977) Quantum Yields for CO(2) Uptake in C(3) and C(4) Plants: Dependence on Temperature, CO(2), and O(2) Concentration. *Plant Physiol.* **59**, 86–90.

Ehleringer J. R. (1978) Implications of Quantum Yield Differences on the Distributions of C_3 and C_4 Grasses. *Oecologia* **31**, 255–267.

Ehleringer J. R., Cerling T. E. and Helliker B. R. (1997) C4 photosynthesis, atmospheric CO2, and climate. *Oecologia* **112**, 285–299.

Feakins S. J., Levin N. E., Liddy H. M., Sieracki A., Eglinton T. I. and Bonnefille R. (2013) Northeast african vegetation change over 12 m.y. *Geology* **41**, 295–298.

Ford H. L., Ravelo A. C., Dekens P. S., Lariviere J. P. and Wara M. W. (2015) The evolution of the equatorial thermocline and the early Pliocene El Padre mean state. *Geophys. Res. Lett.* **42**, 4878–4887.

Foster G. L., Lear C. H. and Rae J. W. B. (2012) The evolution of pCO2, ice volume and climate during the middle Miocene. *Earth Planet. Sci. Lett.* **341**–**344**, 243–254.

Franks P. J., Royer D. L., Beerling D. J., Van de Water P. K., Cantrill D. J., Barbour M. M. and Berry J. A. (2014) New constraints on atmospheric CO2 concentration for the Phanerozoic. *Geophys. Res. Lett.* **41**, 4685–4694.

Freeman K. H. and Colarusso L. a. (2001) Molecular and isotopic records of C4 grassland expansion in the late miocene. *Geochim. Cosmochim. Acta* **65**, 1439–1454.

Gasson E., DeConto R. M., Pollard D. and Levy R. H. (2016) Dynamic Antarctic ice sheet during the early to mid-Miocene. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 201516130.

Goldner A., Herold N. and Huber M. (2014) The challenge of simulating the warmth of the mid-Miocene climatic optimum in CESM1. *Clim. Past* **10**, 523–536.

Gordon I. J. and Prins H. H. T. (2019) The Ecology of Browsing and Grazing II. In pp. 1–4.

Greenop R., Foster G. L., Wilson P. A. and Lear C. H. (2014) Middle Miocene climate instability associated with high-amplitude CO2 variability. *Paleoceanography* **29**, 845–853.

Greenop R., Hain M. P., Sosdian S. M., Oliver K. I. C., Goodwin P., Chalk T. B., Lear C. H., Wilson P. A. and Foster G. L. (2017) A record of Neogene seawater δ11B reconstructed from paired δ11B analyses on benthic and planktic foraminifera. *Clim. Past* **13**, 149–170.

Griffith D. M., Anderson T. M., Osborne C. P., Strömberg C. A. E., Forrestel E. J. and Still C. J. (2015) Biogeographically distinct controls on C3 and C4 grass distributions: merging community and physiological ecology. *Glob. Ecol. Biogeogr.* **24**, 304–313.

Guichard F., Kergoat L., Mougin E., Timouk F., Baup F., Hiernaux P. and Lavenu F. (2009) Surface thermodynamics and radiative budget in the Sahelian Gourma: Seasonal and diurnal cycles. *J. Hydrol.* **375**, 161–177.

Hansen J., Sato M., Kharecha P., Beerling D., Berner R., Masson-Delmotte V., Pagani M., Raymo M. E., Royer D. L. and Zachos J. C. (2008) Target Atmospheric CO2: Where Should Humanity Aim? *Open Atmos. Sci. J.* **2**, 217–231.

Hemming N. G. and Hanson G. N. (1992) Boron isotopic composition and concentration in modern marine carbonates. *Geochim. Cosmochim. Acta* **56**, 537–543.

Henderiks J. and Pagani M. (2007) Refining ancient carbon dioxide estimates: Significance of coccolithophore cell size for alkenone-based pCO2 records. *Paleoceanography* **22**, 1–12.

Herbert T. D., Lawrence K. T., Tzanova A., Peterson L. C., Caballero-Gill R. and Kelly C. S. (2016) Late Miocene global cooling and the rise of modern ecosystems. *Nat. Geosci.* **9**, 843–847.

Hoetzel S., Dupont L., Schefuß E., Rommerskirchen F. and Wefer G. (2013) The role of fire in Miocene to Pliocene C4 grassland and ecosystem evolution. *Nat. Geosci.* **6**, 1027–1030.

Hönisch B. and Hemming N. G. (2005) Surface ocean pH response to variations in pCO2 through two full glacial cycles. *Earth Planet. Sci. Lett.* **236**, 305–314.

IPCC (2013) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. eds. T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Ji S., Nie J., Lechler A., Huntington K. W., Heitmann E. O. and Breecker D. O. (2018) A symmetrical CO2 peak and asymmetrical climate change during the middle Miocene. *Earth Planet. Sci. Lett.* **499**, 134–144.

Kameo K. and Bralower T. (2000) Neogene Calcareous Nannofossil Biostratigraphy of Sites 998, 999, and 1000, Caribbean Sea. *Proc. Ocean Drill. Progr.* **165**, 3–17.

Kameo K. and Takayama T. (1999) Biostratigraphic significance of sequential size variations of the calcareous nannofossil genus Reticulofenestra in the Upper Pliocene of the North Atlantic. *Mar. Micropaleontol.* **37**, 41–52.

Kasbohm J. and Schoene B. (2018) Rapid eruption of the Columbia River flood basalt and correlation with the mid-Miocene climate optimum. *Sci. Adv.* **4**, eaat8223.

Keating-Bitonti C. R. and Peters S. E. (2019) Influence of increasing carbonate saturation in Atlantic bottom water during the late Miocene. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **518**, 134–142.

Knorr G., Butzin M., Micheels A. and Lohmann G. (2011) A warm Miocene climate at low atmospheric CO2 levels. *Geophys. Res. Lett.* **38**, 1–5.

Kürschner W. M. and Kvaček Z. (2009) Oligocene-Miocene CO2 fluctuations, climatic and palaeofloristic trends inferred from fossil plant assemblages in central Europe. *Bull. Geosci.*, 189–202.

LaRiviere J. P., Ravelo A. C., Crimmins A., Dekens P. S., Ford H. L., Lyle M. W. and Wara M. W. (2012) Late Miocene decoupling of oceanic warmth and atmospheric carbon dioxide forcing. *Nature* **486**, 97–100.

Larsen H. C., Saunders A. D., Clift P. D., Beget J., Wei W., Spezzaferri S. and ODP Leg 152 Scientific Party (1994) Seven Million Years of Glaciation in Greenland. *Science (80-.).* **264**, 952–955.

Lee C.-T. A., Shen B., Slotnick B. S., Liao K., Dickens G. R., Yokoyama Y., Lenardic A., Dasgupta R., Jellinek M., Lackey J. S., Schneider T. and Tice M. M. (2013) Continental arc-island arc fluctuations, growth of crustal carbonates, and long-term climate change. *Geosphere* **9**, 21–36.

Lehmann C. E. R., Anderson T. M., Sankaran M., Higgins S. I., Archibald S., Hoffmann W. A., Hanan N. P., Williams R. J., Fensham R. J., Felfili J., Hutley L. B., Ratnam J., San Jose J., Montes R., Franklin D., Russell-Smith J., Ryan C. M., Durigan G., Hiernaux P., Haidar R., Bowman D. M. J. S. and Bond W. J. (2014) Savanna vegetation-fire-climate relationships differ among continents. *Science (80-.).* **343**, 548–552.

Lehmann C. E. R., Archibald S. A., Hoffmann W. A. and Bond W. J. (2011) Deciphering the distribution of the savanna biome. *New Phytol.* **191**, 197–209.

Londoño L., Royer D. L., Jaramillo C., Escobar J., Foster D. A., Cárdenas-Rozo A. L. and Wood A. (2018) Early Miocene CO 2 estimates from a Neotropical fossil leaf assemblage exceed 400 ppm. *Am. J. Bot.* **105**, 1929–1937.

Lunt D. J., Haywood A. M., Schmidt G. A., Salzmann U., Valdes P. J. and Dowsett H. J. (2010) Earth system sensitivity inferred from Pliocene modelling and data. *Nat. Geosci.* **3**, 60–64.

MacFadden B. J. (2005) Terrestrial Mammalian Herbivore Response to Declining Levels of Atmospheric CO2 During the Cenozoic: Evidence from North American Fossil Horses (Family Equidae). In *A History of Atmospheric CO2 and Its Effects on Plants, Animals, and Ecosystems* (eds. J. R. Ehleringer, T. E. Cerling, and M. D. Dearing). Springer-Verlag, New York. pp. 273–292.

Martínez-Botí M. a., Foster G. L., Chalk T. B., Rohling E. J., Sexton P. F., Lunt D. J., Pancost R. D., Badger M. P. S. and Schmidt D. N. (2015) Plio-Pleistocene climate sensitivity evaluated using high-resolution CO2 records. *Nature* **518**, 49–54.

McKenzie N. R., Horton B. K., Loomis S. E., Stockli D. F., Planavsky N. J. and Lee C. T. A. (2016) Continental arc volcanism as the principal driver of icehouse-greenhouse variability. *Science (80-.).* **352**, 444–447.

Mejía L. M., Méndez-Vicente A., Abrevaya L., Lawrence K. T., Ladlow C., Bolton C. T., Cacho I. and Stoll H. M. (2017) A diatom record of CO2 decline since the late Miocene. *Earth Planet. Sci. Lett.* **479**, 18–33.

Olsen A., Lange N., Key R. M., Tanhua T., Álvarez M., Becker S., Bittig H. C., Carter B. R., Cotrim da Cunha L., Feely R. A., van Heuven S., Hoppema M., Ishii M., Jeansson E., Jones S. D., Jutterström S., Karlsen M. K., Kozyr A., Lauvset S. K., Lo Monaco C., Murata A., Pérez F. F., Pfeil B., Schirnick C., Steinfeldt R., Suzuki T., Telszewski M., Tilbrook B., Velo A. and Wanninkhof R. (2019) GLODAPv2.2019 - an update of GLODAPv2. *Earth Syst. Sci. Data Discuss.*, 1–39.

Pagani M., Arthur M. A. and Freeman K. H. (1999a) Miocene evolution of atmospheric carbon dioxide. *Paleoceanography* **14**, 273–292.

Pagani M., Freeman K. H. and Arthur M. A. (1999b) Late Miocene Atmospheric CO2 Concentrations and the Expansion of C4 Grasses. *Science (80-.).* **285**, 876–879.

Pagani M., Huber M., Liu Z., Bohaty S. M., Henderiks J., Sijp W., Krishnan S. and DeConto R. M. (2011) The Role of Carbon Dioxide During the Onset of Antarctic Glaciation. *Science (80-.).* **334**, 1261–1264.

Pagani M., Liu Z., LaRiviere J. and Ravelo A. C. (2010) High Earth-system climate sensitivity determined from Pliocene carbon dioxide concentrations - Supplement. *Nat. Geosci.* **3**, 27–30.

Pälike H., Lyle M. W., Nishi H., Raffi I., Ridgwell A. J., Gamage K., Klaus A., Acton G., Anderson L., Backman J., Baldauf J., Beltran C., Bohaty S. M., Bown P., Busch W., Channell J. E. T., Chun C. O. J., Delaney M., Dewangan P., Dunkley Jones T., Edgar K. M., Evans H., Fitch P., Foster G. L., Gussone N., Hasegawa H., Hathorne E. C., Hayashi H., Herrle J. O., Holbourn A., Hovan S., Hyeong K., Iijima K., Ito T., Kamikuri S., Kimoto K., Kuroda J., Leon-Rodriguez L., Malinverno A., Moore Jr T. C., Murphy B. H., Murphy D.

P., Nakamura H., Ogane K., Ohneiser C., Richter C., Robinson R., Rohling E. J., Romero O., Sawada K., Scher H., Schneider L., Sluijs A., Takata H., Tian J., Tsujimoto A., Wade B. S., Westerhold T., Wilkens R., Williams T., Wilson P. a., Yamamoto Y., Yamamoto S., Yamazaki T. and Zeebe R. E. (2012) A Cenozoic record of the equatorial Pacific carbonate compensation depth. *Nature* **488**, 609–614.

Passey B. H., Ayliffe L. K., Kaakinen A., Zhang Z., Eronen J. T., Zhu Y., Zhou L., Cerling T. E. and Fortelius M. (2009) Strengthened East Asian summer monsoons during a period of high-latitude warmth? Isotopic evidence from Mio-Pliocene fossil mammals and soil carbonates from northern China. *Earth Planet. Sci. Lett.* **277**, 443–452.

Polissar P. J., Rose C., Uno K. T., Phelps S. R. and DeMenocal P. B. (2019) Synchronous rise of African C4 ecosystems 10 million years ago in the absence of aridification. *Nat. Geosci.* **12**, 657–660.

Pound M. J., Haywood A. M., Salzmann U. and Riding J. B. (2012) Global vegetation dynamics and latitudinal temperature gradients during the Mid to Late Miocene (15.97–5.33Ma). *Earth-Science Rev.* **112**, 1–22.

Quade J. and Cerling T. E. (1995) Expansion of C4 grasses in the Late Miocene of Northern Pakistan: evidence from stable isotopes in paleosols. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **115**, 91–116.

Ridgwell A. J. (2005) A Mid Mesozoic Revolution in the regulation of ocean chemistry. *Mar. Geol.* **217**, 339–357.

Rohling E. J., Sluijs A., Dijkstra H. A., Köhler P., van de Wal R. S. W., von der Heydt A. S., Beerling D. J., Berger A., Bijl P. K., Crucifix M., DeConto R. M., Drijfhout S. S., Fedorov A. V., Foster G. L., Ganopolski A., Hansen J., Hönisch B., Hooghiemstra H., Huber M., Huybers P., Knutti R., Lea D. W., Lourens L. J., Lunt D., Masson-Demotte V., Medina-Elizalde M., Otto-Bliesner B. L., Pagani M., Pälike H., Renssen H., Royer D. L., Siddall M., Valdes P., Zachos J. C. and Zeebe R. E. (2012) Making sense of palaeoclimate sensitivity. *Nature* **491**, 683–691.

Royer D. L. (2016) Climate Sensitivity in the Geologic Past. *Annu. Rev. Earth Planet. Sci.* **44**, 277–293.

Savin S. M., Douglas R. G. and Stehli F. G. (1975) Tertiary marine paleotemperatures. *Bull. Geol. Soc. Am.* **86**, 1499–1510.

Seki O., Foster G. L., Schmidt D. N., Mackensen A., Kawamura K. and Pancost R. D. (2010) Alkenone and boron-based Pliocene pCO2 records. *Earth Planet. Sci. Lett.* **292**, 201–211.

Shackleton N. J. and Kennett J. P. (1975) Paleotemperature History of the Cenozoic and the Initiation of Antarctic Glaciation: Oxygen and Carbon Isotope Analyses in DSDP Sites 277, 279 and 281. In *Initial Reports of the Deep Sea Drilling Project, 29* U.S. Government

Printing Office. pp. 743–755.

Si W. and Rosenthal Y. (2019) Reduced continental weathering and marine calcification linked to late Neogene decline in atmospheric CO2. *Nat. Geosci.* **12**, 833–838.

Sosdian S. M., Greenop R., Hain M. P., Foster G. L., Pearson P. N. and Lear C. H. (2018) Constraining the evolution of Neogene ocean carbonate chemistry using the boron isotope pH proxy. *Earth Planet. Sci. Lett.* **498**, 362–376.

Stein R., Fahl K., Schreck M., Knorr G., Niessen F., Forwick M., Gebhardt C., Jensen L., Kaminski M., Kopf A., Matthiessen J., Jokat W. and Lohmann G. (2016) Evidence for ice-free summers in the late Miocene central Arctic Ocean. *Nat. Commun.* **7**, 1–13.

Stoll H. M., Guitian J., Hernandez-Almeida I., Mejia L. M., Phelps S. R., Polissar P. J., Rosenthal Y., Zhang H. and Ziveri P. (2019) Upregulation of phytoplankton carbon concentrating mechanisms during low CO2 glacial periods and implications for the phytoplankton pCO2 proxy. *Quat. Sci. Rev.* **208**, 1–20.

Suchéras-Marx B. and Henderiks J. (2014) Downsizing the pelagic carbonate factory: Impacts of calcareous nannoplankton evolution on carbonate burial over the past 17 million years. *Glob. Planet. Change* **123**, 97–109.

Super J. R., Thomas E., Pagani M., Huber M., O'Brien C. and Hull P. M. (2018) North Atlantic temperature and pCO2 coupling in the early-middle Miocene. *Geology* **46**, 519–522.

Takayama T. (1993) Notes on Neogene calcareous nanofossil biostratigraphy of the Ontong Java plateau and size variations of Reticulofenestra coccoliths. *Proc. Ocean Drill. Program, Sci. Results* **130**, 179–229.

Tesfamichael T., Jacobs B., Tabor N., Michel L., Currano E., Feseha M., Barclay R., Kappelman J. and Schmitz M. (2017) Settling the issue of "decoupling" between atmospheric carbon dioxide and global temperature: [CO2]atm reconstructions across the warming Paleogene-Neogene divide. *Geology* **45**, 999–1002.

Torres M. A., West A. J. and Li G. (2014) Sulphide oxidation and carbonate dissolution as a source of CO2 over geological timescales. *Nature* **507**, 346–349.

Uno K. T., Cerling T. E., Harris J. M., Kunimatsu Y., Leakey M. G., Nakatsukasa M. and Nakaya H. (2011) Late Miocene to Pliocene carbon isotope record of differential diet change among East African herbivores. *Proc. Natl. Acad. Sci.* **108**, 6509–6514.

Uno K. T., Polissar P. J., Jackson K. E. and DeMenocal P. B. (2016) Neogene biomarker record of vegetation change in eastern Africa. *Proc. Natl. Acad. Sci.* **113**, 6355–6363.

Vollstaedt H., Eisenhauer A., Wallmann K., Böhm F., Fietzke J., Liebetrau V., Krabbenhöft A., Farkaš J., Tomašových A., Raddatz J. and Veizer J. (2014) The Phanerozoic δ88/86Sr

record of seawater: New constraints on past changes in oceanic carbonate fluxes. *Geochim. Cosmochim. Acta* **128**, 249–265.

Witkowski C. R., Weijers J. W. H., Blais B., Schouten S. and Sinninghe Damsté J. S. (2018) Molecular fossils from phytoplankton reveal secular PCO2 trend over the phanerozoic. *Sci. Adv.* **4**.

Young J. (1990) Size variation of Neogene Reticulofenestra coccoliths from Indian Ocean DSDP Cores. *J. Micropalaeontology* **9**, 71–85.

Zachos J. C., Dickens G. R. and Zeebe R. E. (2008) An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* **451**, 279–283.

Zeebe R. E. and Caldeira K. (2008) Close mass balance of long-term carbon fluxes from ice-core CO2 and ocean chemistry records. *Nat. Geosci.* **1**, 312–315.

Zeebe R. E. and Tyrrell T. (2019) History of carbonate ion concentration over the last 100 million years II: Revised calculations and new data. *Geochim. Cosmochim. Acta* **257**, 373–392.

Zeebe R. E. and Wolf-Gladrow D. A. (2001) *CO2 in Seawater: Equilibrium, Kinetics, Isotopes*. 1st ed., Elsevier.

Zhang Y. G., Pagani M., Henderiks J. and Ren H. (2017) A long history of equatorial deep-water upwelling in the Pacific Ocean. *Earth Planet. Sci. Lett.* **467**, 1–9.

Zhang Y. G., Pagani M., Liu Z., Bohaty S. M. and DeConto R. (2013) A 40-million-year history of atmospheric CO2. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **371**, 20130096.

Appendix A: Supplementary Material for Chapter 2

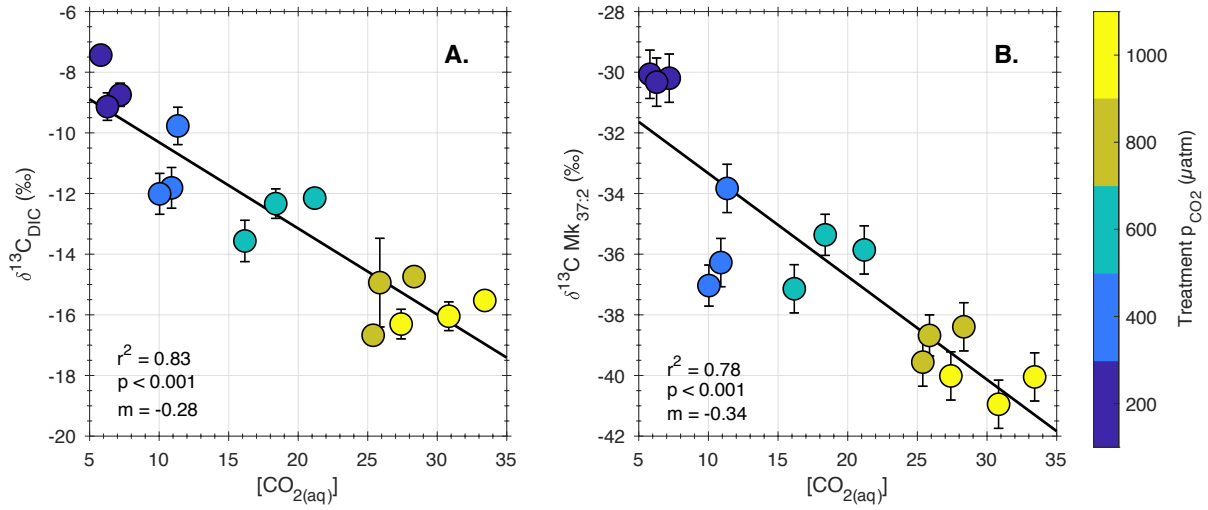

Figure S1. (A) Carbon isotope ratio ($\delta^{13}C$, ‰) of dissolved inorganic carbon (DIC) in each culture vessel with respect to the aqueous CO_2 concentration ($[CO_{2(aq)}]$, µmol kg^{-1}). (B) Carbon isotope ratio of the $C_{37:2}$ methyl ketone as a function of $[CO_{2(aq)}]$. Symbols are colored by the target treatment p_{CO2}. Linear fits on both panels are ordinary least squares regressions. Higher CO_2 in the culture treatments was achieved through headspace aeration and diffusion through the boundary layer; the inverse relationship between $\delta^{13}C_{DIC}$ and $[CO_{2(aq)}]$ reflects the light isotope ratio of the CO_2 tank used for these experiments. The slope of the relationship in (B) is ~20% steeper than that in (A), indicating that the carbon isotope fractionation in alkenones of *G. oceanica* records additional information on top of the $\delta^{13}C_{DIC}$.

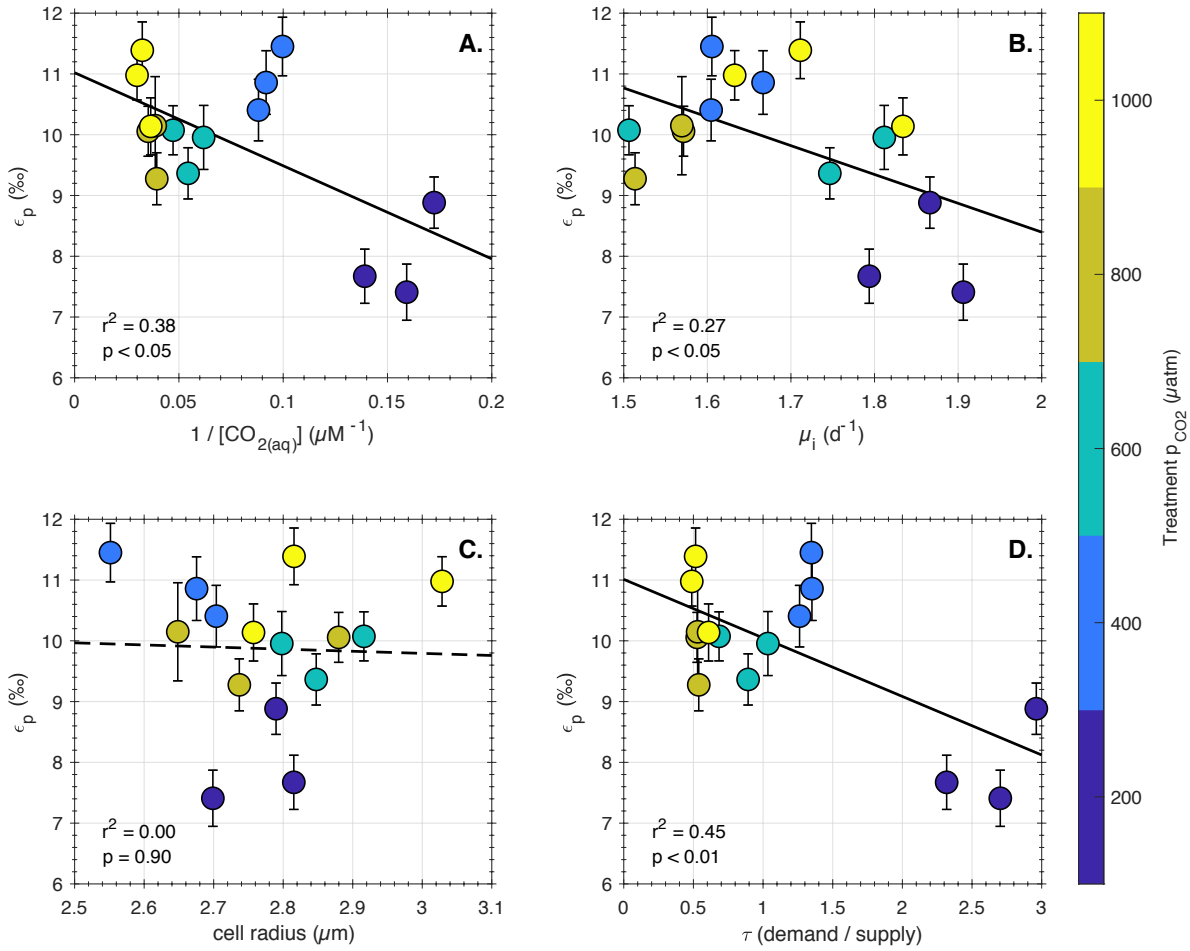

Figure S2. Carbon isotope fractionation relative to in-situ $\delta^{13}C_{CO2aq}$ (ε_p) in new batch culture experiments of *Gephyrocapsa oceanica*, as a function of (A) 1/[$CO_{2(aq)}$], (B) instantaneous growth rate (μ_i), (C) cell radius, and (D) the ratio of carbon demand relative to carbon supply (τ), which includes an estimate of POC $cell^{-1}$ from measured cell size. Symbols are colored by the target pCO_2 of the treatment, which are within ~10% of the actual measured pCO_2. Linear regressions between ε_p and the independent variables are plotted in each panel, and are shown as solid lines if the relationship is significant at the 95% confidence level.

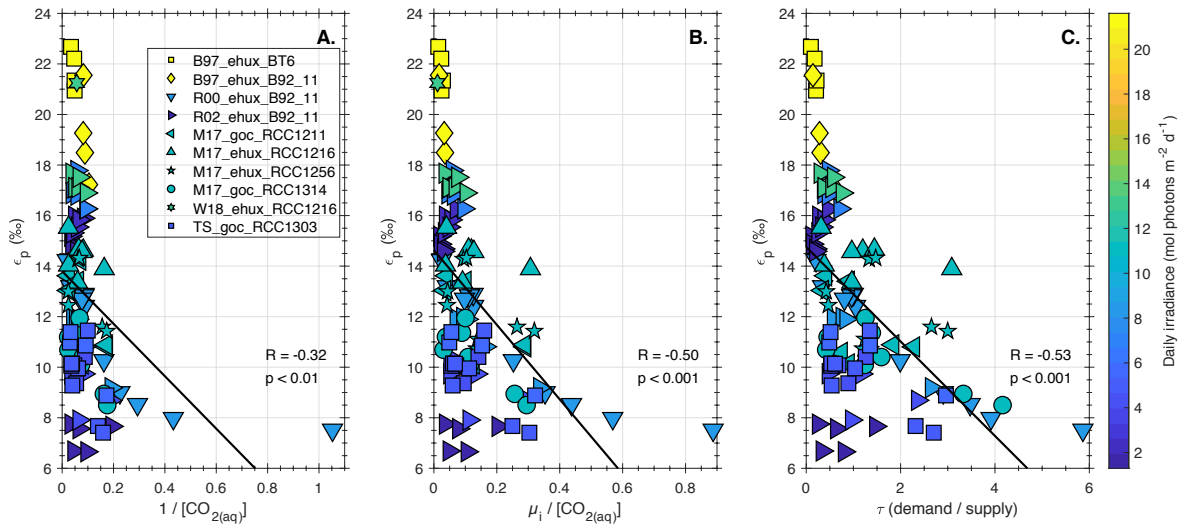

Figure S3. Carbon isotope fractionation relative to in-situ $\delta^{13}C_{CO2aq}$ (ε_p) in a new compilation of culture experiments of alkenone-producing coccolithophores where the components of τ are constrained. ε_p is plotted against (A) $1/[CO_{2(aq)}]$, (B) instantaneous growth rate (μ_i) divided by $[CO_{2(aq)}]$, and (C) the ratio of carbon demand relative to carbon supply (τ). Symbols are colored by the integrated daily irradiance in the culture, and different symbols represent different publications, species, and strains. Shorthand in the legend reflects the publication (e.g. B97 = Bidigare et al., 1997), species cultured (ehux = *Emiliania huxleyi*; goc = *Gephyrocapsa oceanica*), and strain identifier. Lines in each plot are simple linear regressions between ε_p and the independent variable. Linear correlation coefficients and significance values of the slope estimate are shown in each panel.

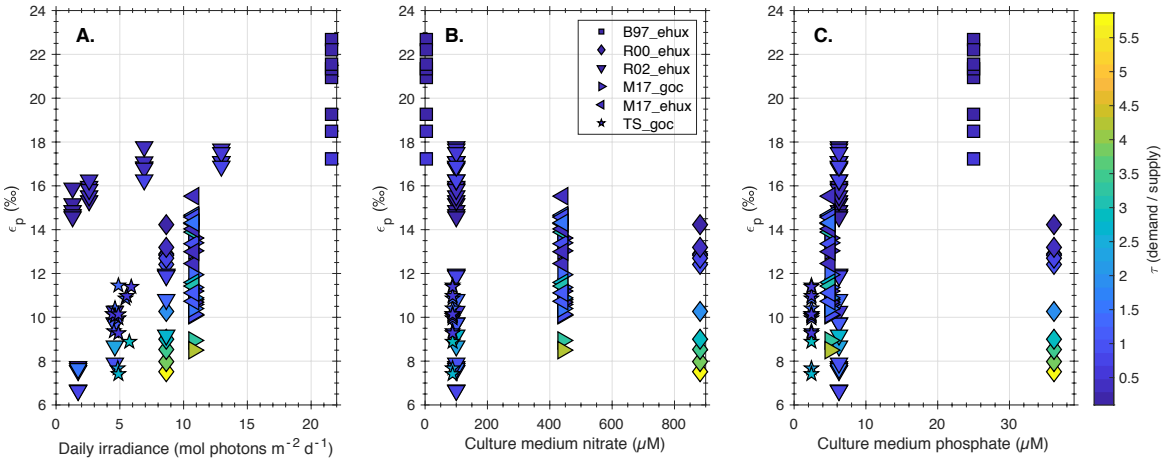

Figure S4. Carbon isotope fractionation relative to in-situ $\delta^{13}C_{CO2aq}$ (ε_p) in compiled dataset with respect to (A) integrated daily irradiance, (B) medium nitrate, and (C) medium phosphate. Symbols are assigned based on publication and the species cultured, and are colored by the calculated τ value (see Chapter 2, Equation 2).

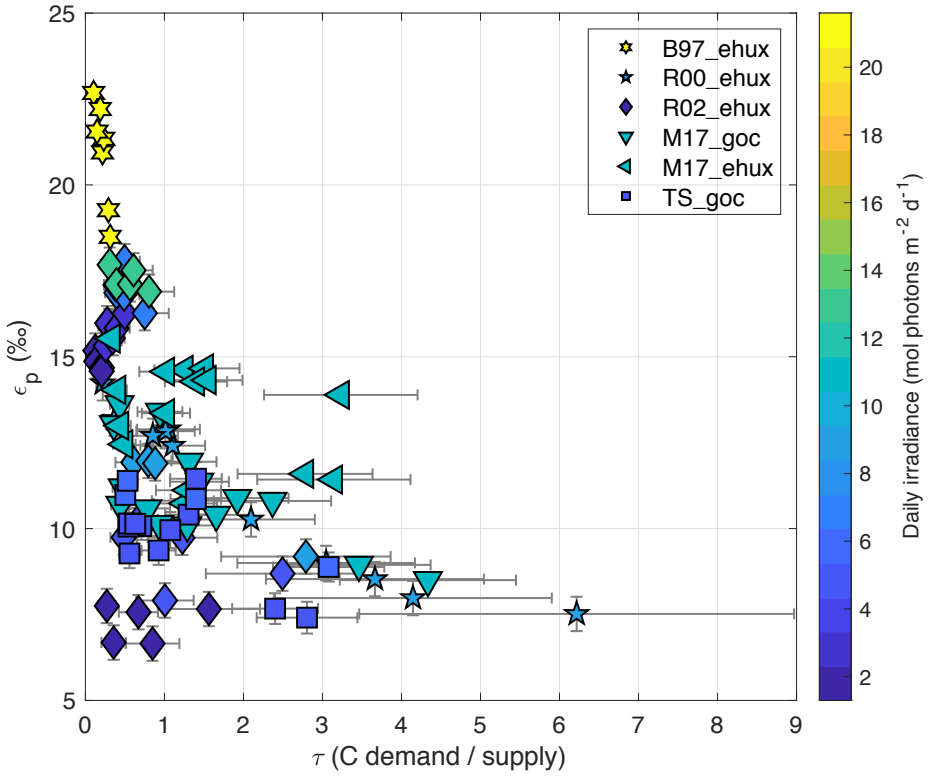

Figure S5. Carbon isotope fractionation (ε_p) with respect to the ratio of cellular carbon demand to diffusive CO_2 supply (τ). This plot is identical to Chapter 2, Figure 2A, but here the we include ± 1 standard deviations on the estimated τ values. Errors in τ are calculated using a Monte Carlo routine with 10,000 simulated observations from the components of τ (POC cell^{-1}, instantaneous growth rate, surface area, and $[CO_{2(aq)}]$). Permeability is held constant. Errors are assumed to be normally distributed and are either measured and reported in the original publications, or assumed to be 10-20% relative error (one standard deviation). Errors on ε_p are one standard deviation from the mean.

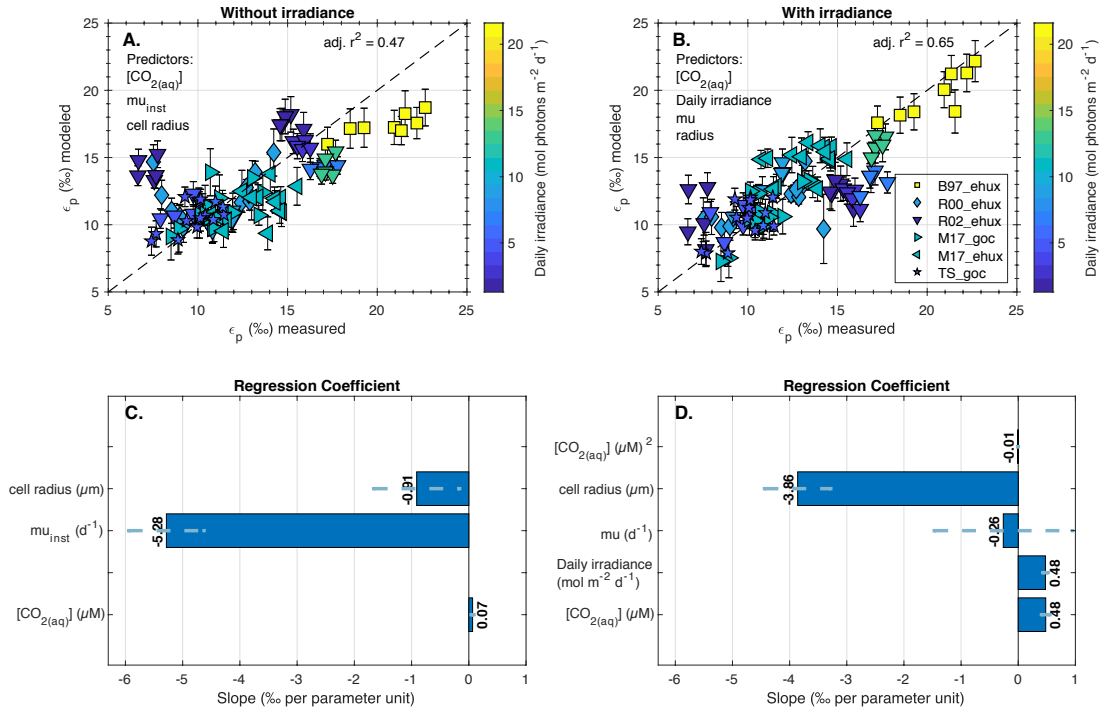

Figure S6. Additional multiple linear regression model using total daily irradiance ("integrated irradiance") as a predictor instead of light hours and light intensity independently. (A) and (C) are identical to Figure 5 in Chapter 2. (B) and (D) are modified to use the product of light hours and light intensity. This "integrated irradiance" model performs better than the "Without irradiance" model, though explains ~20% less variance than the irradiance model shown in the main text (Chapter 2, Fig. 5) where light intensity and light hours are independent predictors.

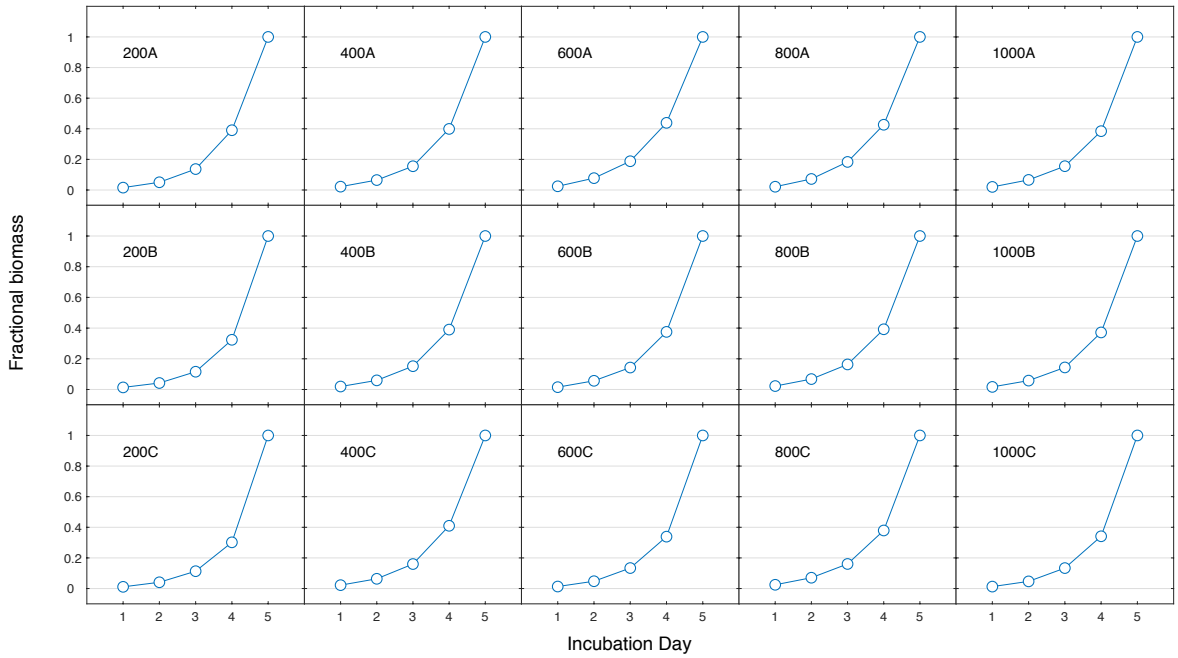

Figure S7. Temporal evolution in biomass in each treatment and incubation. Cumulative biomass in each replicate as inferred from relative fluorescence. The total fractional biomass on the final incubation day is 1, and the initial fractional biomass is slightly higher than zero. The majority of the biomass in all treatments (~60%) is generated during the final cell division. Labels show the target CO_2 of the treatment as well as the replicate (A, B, or C).

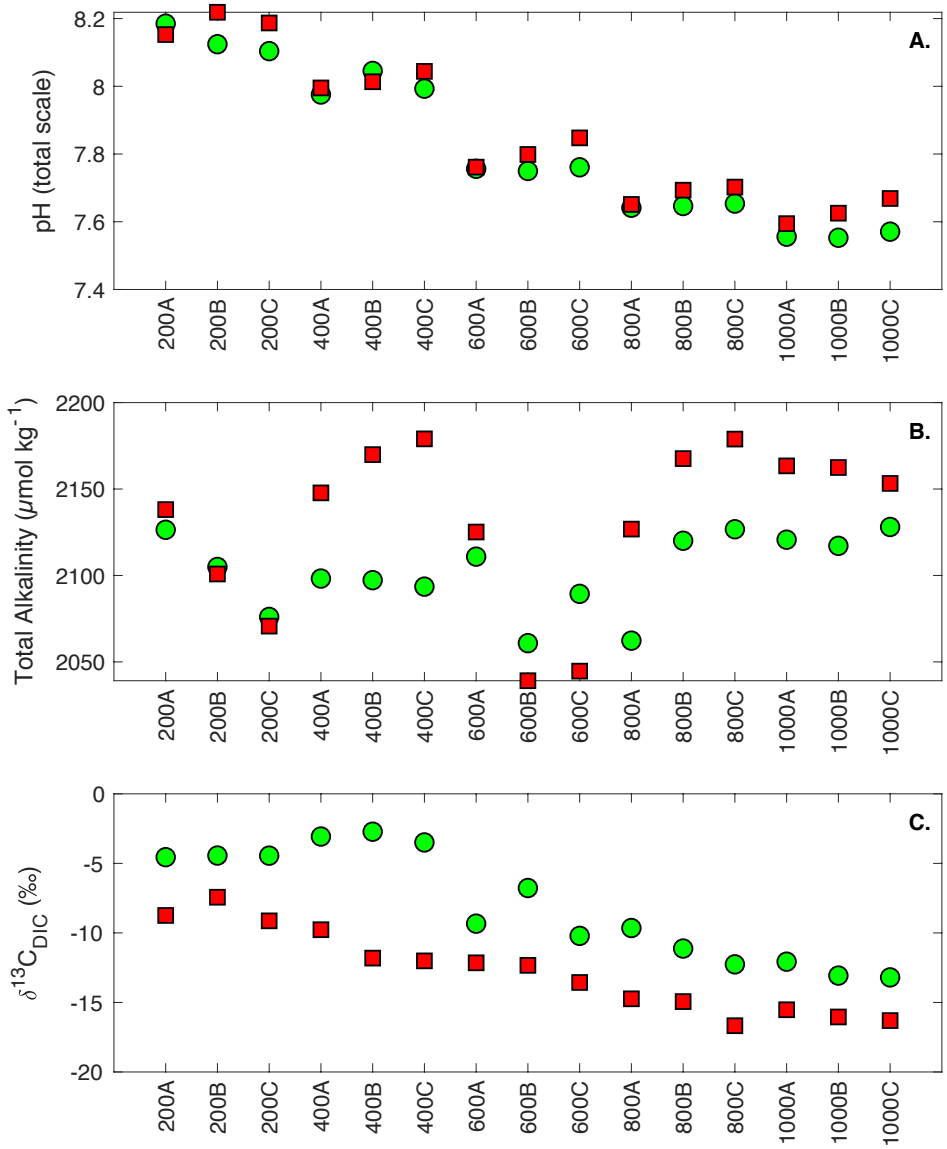

Figure S8. Temporal evolution of measured carbonate system parameters in each treatment. Green circles are initial (t_0) measurements and red squares are measurements at the time of harvest (t_f). (A) pH; (B) total alkalinity; (C) $\delta^{13}C_{DIC}$. Labels on the abscissa show the treatment and replicate. The evolution of the carbonate system is different than most closed-system batch incubations because the headspace of the culture vessel was continuously purged with air containing the target CO_2 but the vessel was partially open to the atmosphere.

Table S1. Size measurements and normalization from new *G. oceanica* cultures.

Treatment & Replicate	Sampling Time (hour of day)	Sampling Time (hours after light on)	Photoperiod length (hours)	Normalization Time (hours after light on)	Cells mL^{-1}	log(FSC_cells)[a]	log(FSC_beads)[a]	Response factor, (log(FSC_cells) / log(FSC_beads))	Prescribed baseline cell diameter (µm)	Prescribed baseline cell volume ($µm^3$)	Baseline response factor[b]	Volume relative to baseline ($µm^3$)	Volume relative to baseline, time-normalized ($µm^3$)	Cell radius at sampling time (µm)	Cell radius, time-normalized (µm)
200A	10.30	3.30	14	7	9.4E+04	118.10	151.27	0.78	5.2	73.62	0.75	77.00	93.47	2.64	2.82
200B	11.78	4.78	14	7	1.3E+05	120.66	146.29	0.82	5.2	73.62	0.75	81.34	90.94	2.69	2.79
200C	13.33	6.33	14	7	1.1E+05	124.60	154.13	0.81	5.2	73.62	0.75	79.73	82.34	2.67	2.70
400A	12.88	5.88	14	7	1.9E+04	111.22	139.92	0.79	5.2	73.62	0.75	78.39	82.80	2.65	2.70
400B	12.95	5.95	14	7	2.5E+04	110.95	143.59	0.77	5.2	73.62	0.75	76.20	80.22	2.63	2.68
400C	13.00	6.00	14	7	2.3E+04	98.27	146.25	0.67	5.2	73.62	0.75	66.27	69.58	2.51	2.55
600A	10.67	3.67	14	7	7.3E+04	131.46	148.37	0.89	5.2	73.62	0.75	87.38	103.87	2.75	2.92
600B	11.87	4.87	14	7	8.9E+04	132.97	150.93	0.88	5.2	73.62	0.75	86.89	96.71	2.75	2.85
600C	13.43	6.43	14	7	9.3E+04	131.60	145.38	0.91	5.2	73.62	0.75	89.28	91.75	2.77	2.80
800A	10.47	3.47	14	7	4.3E+04	153.81	182.31	0.84	5.2	73.62	0.75	83.21	100.04	2.71	2.88
800B	11.90	4.90	14	7	4.9E+04	104.82	147.62	0.71	5.2	73.62	0.75	70.03	77.81	2.56	2.65
800C	13.63	6.63	14	7	2.2E+04	131.88	154.15	0.86	5.2	73.62	0.75	84.38	85.87	2.72	2.74
1000A	10.52	3.52	14	7	6.0E+04	137.70	139.93	0.98	5.2	73.62	0.75	97.05	116.35	2.85	3.03
1000B	11.93	4.93	14	7	6.7E+04	131.11	153.40	0.85	5.2	73.62	0.75	84.29	93.49	2.72	2.82
1000C	13.52	6.52	14	7	7.6E+04	130.48	150.00	0.87	5.2	73.62	0.75	85.79	87.81	2.74	2.76

(a) measured by flow cytometry

(b) average of the 400-ppm treatments

Table S2. Carbonate chemistry measurements

Treatment ID	Initial conditions (t_0)			Final condtions (t_f)		
	pH (total)	Total alkalinity ($\mu mol\ kg^{-1}$)	$\delta^{13}C_{DIC}$ (‰ VPDB)	pH (total)	Total alkalinity ($\mu mol\ kg^{-1}$)	$\delta^{13}C_{DIC}$ (‰ VPDB)
200A	8.293	2126.4	-12.07	8.260	2138.1	-15.53
200B	8.232	2104.9	-13.07	8.327	2100.9	-16.05
200C	8.211	2076.0	-13.20	8.295	2070.7	-16.30
400A	8.081	2098.2	-4.56	8.101	2147.8	-8.75
400B	8.152	2097.3	-4.44	8.119	2169.9	-7.44
400C	8.099	2093.5	-4.44	8.150	2179.0	-9.13
600A	7.858	2110.9	-3.07	7.863	2125.1	-9.77
600B	7.851	2060.8	-2.72	7.901	2039.1	-11.81
600C	7.862	2089.4	-3.49	7.951	2044.7	-12.01
800A	7.739	2062.3	-9.34	7.749	2126.8	-12.15
800B	7.745	2120.1	-6.77	7.793	2167.6	-12.33
800C	7.752	2126.7	-10.21	7.802	2178.9	-13.56
1000A	7.651	2120.7	-9.65	7.691	2163.4	-14.74
1000B	7.648	2117.1	-11.13	7.723	2162.5	-14.94
1000C	7.666	2128.0	-12.26	7.768	2153.2	-16.67

Table S3. Results from new *G. oceanica* culture data.

Treatment ID	Photoperiod	Light intensity (µmol photons / m^2 / s)	Nutrient media recipe	Nutrient media recipe notes	Nitrate (µmol/L)	Phosphate (µmol/L)	Alkalinity (µmol/kg)	pH (total scale)	pCO2 (µatm)	CO2aq (µmol/kg)	Temperature (°C)	Salinity
200A	14:10	96	L1, reduced	L1/10 N, L1/	88.2	2.413	2138.11	8.260396	207.567944	7.18901509	18	32.53
200B	14:10	114	L1, reduced	L1/10 N, L1/	88.2	2.413	2100.85	8.21857686	167.507789	5.80155107	18	32.53
200C	14:10	97	L1, reduced	L1/10 N, L1/	88.2	2.413	2070.69	8.18698139	181.357472	6.28122812	18	32.53
400A	14:10	92	L1, reduced	L1/10 N, L1/	88.2	2.413	2147.76694	7.99515542	327.516942	11.3433905	18	32.53
400B	14:10	108	L1, reduced	L1/10 N, L1/	88.2	2.413	2169.90358	8.01342438	314.485579	10.8920555	18	32.53
400C	14:10	97	L1, reduced	L1/10 N, L1/	88.2	2.413	2179.0065	8.04400953	289.732141	10.0347322	18	32.53
600A	14:10	92	L1, reduced	L1/10 N, L1/	88.2	2.413	2125.13	7.7612468	611.755922	21.1878697	18	32.53
600B	14:10	90	L1, reduced	L1/10 N, L1/	88.2	2.413	2039.11	7.79870615	530.792852	18.3837531	18	32.53
600C	14:10	97	L1, reduced	L1/10 N, L1/	88.2	2.413	2044.66	7.84765025	466.719966	16.1646197	18	32.53
800A	14:10	92	L1, reduced	L1/10 N, L1/	88.2	2.413	2126.82	7.65114157	818.069793	28.3334506	18	32.53
800B	14:10	88	L1, reduced	L1/10 N, L1/	88.2	2.413	2167.61	7.69323639	746.957769	25.8705202	18	32.53
800C	14:10	97	L1, reduced	L1/10 N, L1/	88.2	2.413	2178.89	7.70239159	733.064032	25.3893173	18	32.53
1000A	14:10	110	L1, reduced	L1/10 N, L1/	88.2	2.413	2163.39	7.59442595	964.885982	33.4183581	18	32.53
1000B	14:10	117	L1, reduced	L1/10 N, L1/	88.2	2.413	2162.45	7.62527096	890.106114	30.8283937	18	32.53
1000C	14:10	97	L1, reduced	L1/10 N, L1/	88.2	2.413	2153.24	7.66872075	791.180133	27.4021403	18	32.53

Treatment ID	δ^{13}C DIC (‰ VPDB)	δ^{13}C DIC, 1std (‰ VPDB)	δ^{13}C biomass (‰ VPDB)	δ^{13}C biomass, 1std (‰ VPDB)	δ^{13}C Mk37:2 (‰ VPDB)	δ^{13}C Mk37:2, 1std (‰ VPDB)	ε_p (‰ VPDB)	ε_p 1std (‰ VPDB)	Membrane permeability (m/s)	Cell radius (µm)[a]	POC (pgC/cell)[a]	τ[c]
200A	-8.75	0.19	-25.90	0.40	-30.20	0.40	7.67	0.45	2.00E-05	2.82	19.65	2.32
200B	-7.44	0.13	-25.77	0.40	-30.07	0.40	8.88	0.42	2.00E-05	2.79	19.12	2.96
200C	-9.13	0.23	-26.03	0.40	-30.33	0.40	7.41	0.46	2.00E-05	2.70	17.31	2.70
400A	-9.77	0.31	-29.55	0.40	-33.83	0.40	10.41	0.51	2.00E-05	2.70	17.40	1.26
400B	-11.81	0.34	-32.00	0.40	-36.28	0.40	10.86	0.52	2.00E-05	2.68	16.86	1.35
400C	-12.01	0.34	-32.76	0.34	-37.04	0.34	11.45	0.48	2.00E-05	2.55	14.63	1.35
600A	-12.15	0.04	-31.58	0.40	-35.86	0.40	10.07	0.40	2.00E-05	2.92	21.83	0.68
600B	-12.33	0.24	-31.08	0.34	-35.36	0.34	9.36	0.42	2.00E-05	2.85	20.33	0.89
600C	-13.56	0.34	-32.87	0.40	-37.14	0.40	9.96	0.53	2.00E-05	2.80	19.28	1.03
800A	-14.74	0.09	-34.13	0.40	-38.39	0.40	10.06	0.41	2.00E-05	2.88	21.03	0.53
800B	-14.94	0.73	-34.41	0.34	-38.68	0.34	10.15	0.81	2.00E-05	2.65	16.35	0.53
800C	-16.67	0.15	-35.30	0.40	-39.56	0.40	9.28	0.43	2.00E-05	2.74	18.05	0.54
1000A	-15.53	0.06	-35.79	0.40	-40.05	0.40	10.98	0.41	2.00E-05	3.03	24.46	0.49
1000B	-16.05	0.24	-36.69	0.40	-40.95	0.40	11.39	0.47	2.00E-05	2.82	19.65	0.52
1000C	-16.30	0.24	-35.76	0.40	-40.02	0.40	10.14	0.47	2.00E-05	2.76	18.46	0.61

(a) sampling-time normalized using Eq. 5 in the main text (Ch. 2)
(b) calculated from Eq. 4 in the main text (Ch. 2)
(c) calculated from Eq. 2 in the main text (Ch. 2)

Appendix B: Supplementary Material for Chapter 3

1. Coccolithophore distributions and coccolith size compilation

Cell radius is an important determinant of algal ε_p (Popp et al., 1998; Burkhardt et al., 1999; McClelland et al., 2017); it is therefore necessary to constrain cell size in order to understand alkenone ε_p variations in the ocean. Coccolithophore species definitions in the alkenone-producing family Noelaerhabdaceae are largely based on morphometry (Young et al., 2003) (Young et al., 2003). Species in this family have well-defined coccolith size ranges. We can therefore use any assemblage-level data and the characteristic sizes of the species observed to generate a reasonable estimate of the average cell radius.

Where species percentages are reported, we convert them to synthetic observations of each species with a sum of 100. Where coccolith lengths (mean, 1 standard deviation, and N measured) are reported, we generate N measured observations from these descriptive statistics assuming a gaussian underlying distribution. We calculate a mean coccolith length from the maximum and minimum coccolith dimensions in the Nannotax database and prescribe these mean values to each species in the compilation. Coccolith length and cell radius in the Noelaerhabdaceae are highly correlated, which allows us to estimate cell dimensions from coccolith dimensions. We convert coccolith length to cell radius using the culture data of McClelland et al. (2016) and the equation:

$$\text{Cell radius } (\mu m) = 0.57 \times \text{coccolith length } (\mu m) + 0.45\ (\mu m) \qquad \text{[Eq. S1]}$$

An important distinction in our average size calculations is our use of biomass-weighting. Because biomass is proportional to the cube of the radius and alkenone concentration per cell scales with biomass (Riebesell et al., 2000), using the arithmetic mean coccolith size in a sample would underestimate the average radius of the cells that produced the ε_p signal. We therefore weight our average coccolith length and cell radius estimates by the contribution of each species to the estimated total biomass in the sample, yielding a biomass-weighted (BMW) mean. We calculate the biomass that each species or coccolith size observation contributes to the total population biomass in a sample using the relationships from culture experiments that found Noelaerhabdaceae organic carbon density is approximately 17.5 fmol C μm^{-3} (Aloisi, 2015; McClelland et al., 2016; Chapter 2). The relationship between cell volume and organic carbon is therefore:

$$\text{POC (fmol C)} = 17.5 \times \text{volume } (\mu m^3) \qquad \text{[Eq. S2]}$$

We merge our size estimates with published alkenone ε_p data using an inverse-distance weighting algorithm. In attempt to minimize any bias resulting from preferential preservation or changes in assemblage over the industrial era, we only combine like archives: cell size estimates for sediment ε_p data is merged only with coccolith sediment data, and particulate ε_p data with particulate assemblage data. For every sample in our coccolithophore assemblage database (n = 1657), we calculate the biomass-weighted coccolith length as described above. Then, for each ε_p sample, we calculate the arc length distance to every coccolithophore size sample using the *distance* function in MATLAB R2019a, and keep all that are within 5° of the ε_p sample. We take a weighted mean using the inverse of the arc-length distance as the weighting function, so

samples in closer proximity are given greater weight in the final estimate. Alkenone ε_p samples with no close coccolith assemblage data are excluded from any analysis that requires a cell radius estimate.

2. Growth rate modeling

We model growth rates following the method developed by Krumhardt et al. (2017), but modify it to include irradiance. In this model, maximum growth rates are set by power-law function of temperature. The realized growth rate is determined by nutrient limitation, which is governed by Michaelis-Menten kinetics.

Maximum possible growth rates (in d^{-1}) are a power function of temperature:

$$\mu_{max} = 0.1419 \times T(°C)^{0.8151} \quad \text{[Eq. S3]}$$

We note that the slope and exponential coefficients in this equation are identical to those in (Fielding, 2013), but the slope coefficient was misreported in (Krumhardt et al., 2017).

Actual growth rates are modified by substrate limitation according to:

$$\mu = \mu_{max} \times \left(\frac{N}{N + K_M}\right) \quad \text{[Eq. S4]}$$

where μ_{max} is the temperature-dependent maximum growth rate (in d^{-1}), N is the in-situ substrate concentration (typically μmol kg^{-1}), and K_M is the half-saturation constant for a given nutrient (units vary depending on substrate, but are the same as units for N). The "limited" growth rates are calculated for each nutrient using:

$$N_{lim} = \left(\frac{N}{N + K_N}\right) \quad \text{[Eq. S5]}$$

The value of N_{lim} is calculated for each growth-dependent substrate in the model. In the initial K17 approach, only p_{CO2} and phosphate were considered; here we include nitrate and daily PAR as well. The realized growth rate is ultimately determined by the abundance of each growth-dependent nutrient relative to the half-saturation constant and is calculated by:

$$\mu = \mu_{max} \times \min(pCO2_{lim}, PO4_{lim}, NO3_{lim}, light_{lim}) \quad \text{[Eq. S6]}$$

A list of the half-saturation constants used is reported in Table S1. Because nutrient half-saturation constants have not been determined for *Gephyrocapsa* to the same degree as *E. huxleyi*, we apply the *E. huxleyi* values to all alkenone producers and assume that the Noelaerhabdaceae have similar nutrient catalytics. *E. huxleyi* and *G. oceanica* cannot be distinguished by their nuclear DNA or 18s rDNA sequences, and have a single substitution differentiating their 16s rDNA sequences (Bendif et al., 2014; Bendif et al., 2016), suggesting many key proteins and traits are conserved. However, there are locations around the world where *E. huxleyi* and *G. oceanica* do not coexist. This may imply that there are some environmental conditions – possibly liked to one of the model parameters – that lead to the dominance of one of these groups due to differences in nutrient acquisition ability. However, we do not currently have a way of testing that line of inquiry.

3. Laboratory Methods

3.1. Lipid extraction and separation

Where possible, the exterior of each sample was scraped with methanol- and dichloromethane (DCM)-cleaned metal spatula to remove potential organic or drilling-mud contamination emplaced during the sampling process. Some trigger-weight samples were too unconsolidated to allow this cleaning step. Bulk sediment samples were lightly disaggregated by hand in sample bags and lyophilized for ~24 hours. For the R/V Oceanus sediments, the bulk sediment was freeze-dried and first washed for foraminifera analysis; for this sample set only, the <63 μm fine-fraction split was used for the ensuing nannofossil and lipid biomarker analysis. Free lipids were extracted with a dichloromethane and methanol mixture (9:1 v/v) via accelerated solvent extraction using a DIONEX ASE 350 at a maximum temperature of 100°C and 1000 psi. A general recovery standard was added to the total lipid extract (TLE) of each sample, which was then evaporated to dryness under a stream of N2 gas. The general recovery standard contained ~2000 ng of each 5α-androstane, 1-1' Binaphthyl, 5α-androstan-3-one, stearyl stearate, 11-eicosenol, 5α-androstan-17b-ol, 11-eicosenoic acid, and 19-methyleicosanoic acid. The TLE was split into aliphatic, ketone/ester (containing the alkenones), and polar fractions using silica gel flash column chromatography by elution with 4-ml hexane, 4-ml DCM, and 4-ml methanol, respectively. The ketone fraction was further treated to eliminate fatty acid methyl esters that coelute with the target alkenone molecules. Roughly 1.5-ml of a saponification mixture (1 N KOH in 95:5 MeOH:ultra-pure H_2O) was added to each vial. Headspace was purged with N_2, and the vial was tightly capped and heated for ~2.5 hours at 75°C. Approximately 0.5 mL of a 5% NaCl in ultrapure H2O solution was added to each vial to improve phase separation. Toluene (~1 mL) was then added to each vial, which was shaken vigorously and centrifuged for 5 minutes at 4000 rpm to facilitate phase separation, and the toluene phase containing the alkenones was recovered. This process was repeated three times,

evaporated to dryness under a pure N_2 stream, and transferred in DCM to a 2 mL borosilicate vial for analysis.

3.2. Alkenone carbon isotope measurements

Carbon isotope ratios of alkenone molecules were measured by continuous flow in a helium carrier gas using a Thermo TraceGC–Isolink system coupled to a Thermo DeltaV isotope ratio mass spectrometer. All target molecules were converted to CO_2 gas using a home-built combustion reactor. A ceramic alumina tube (1.5-mm o.d. x 0.5-mm i.d. x 320-mm length) was packed with a 24-inch long strand of each nickel, copper, and platinum wire (all 0.1-mm diameter, twisted into a braid). The combustion furnace was heated to 1000°C and oxidized overnight with 100% O_2. To maintain stable reactor chemistry, samples were run with a small additional trickle of 1% O_2 in helium, which was introduced to the analyte stream prior to entering the oxidation reactor.

Samples used in this study were run at different times and were analyzed using slightly different GC columns and water removal systems. The LDEO core-tops were analyzed on an Agilent J&W DB-1 (60m x 0.25 mm i.d. x 250 μm phase thickness) with a 10 m guard column. Water removal was achieved with an in-line Nafion membrane with a dry helium counter-flow. The R/V Oceanus core-tops were analyzed using an Agilent J&W VF-200 column (60m x 0.25 mm i.d. x 250 μm phase thickness) coupled to a 10 m guard column. Water was removed using an ethanol cryotrap, which was cooled to -80°C and thawed daily. We ran an internal alkenone laboratory standard across both measurement sessions and found no systematic change in the measured $\delta^{13}C$. All values are reported on the VPDB scale using delta notation, and the standard error of the mean was better than 0.3‰ for all but one sample (0.55‰). Error analysis and

realization of sample delta values on the VPDB scale were conducted using the MATLAB routine of Polissar and D'Andrea (2014).

3.3. Coccolith size measurements

In the Noelaerhabdaceae, coccolith morphometry is a robust proxy for cell diameter (Henderiks and Pagani, 2007; McClelland et al., 2016). New sediment samples analyzed in this study include estimates of coccolith length using an automatic particle identification and measurement program. Measurements were made at the Centre Européen de Recherche et d'Enseignement des Géosciences de l'Environnement (CEREGE) using the Système de reconnaissance automatique de coccoliths (SYRACO).

Samples were prepared by random settling on 12mm x 12mm coverslips, which were weighed before use. The upward-facing side of the coverslip was licked and placed in the center of a glass petri dish, and covered until sediment was added. Sediments were suspended using a pH ~9 dilute ammonium hydroxide solution in deionized water, which was prepared daily. Approximately 10 mL of this solution was added to a 50-mL centrifuge tube. A small amount of sediment (~one-half of a micro spoon) was added to the centrifuge tube, and the sample was lightly sonicated for three seconds, then vortexed for thirty seconds. The suspension was filtered through a 25-μm sieve into another centrifuge tube, then poured into a petri dish with the coverslip already placed in the center of the petri dish. Air bubbles were tamped out of the coverslip using plastic forceps, and all materials were rinsed liberally with tap water and shaken dry prior to use with to the subsequent sample. Petri dishes were covered for approximately four hours, after which the buffered water was pipetted off using a vacuum filtration line. Coverslips

in petri dishes were oven-dried at 60°C to remove any remaining water. Eight coverslips were mounted on each microscope slide using EPOTENCY NOA81 optical adhesive.

At least two coverslips were prepared for each sample, and between 165 and 297 images of each coverslip were taken. Images were acquired with circularly-polarized light and inverse circularly-polarized light, and these two images were overlain. The brightness of the resulting image is directly proportional to the mass of calcite in the image (Beaufort, 2005). Individual carbonate particles are segmented from the background using a proprietary routine in LabView that identifies bright pixels as distinct from the black background. Approximately 90 morphometric parameters of each particle are calculated and used to classify the particles as coccoliths or non-coccoliths using a random forest. A subsequent series of neural networks that has been trained on tens of thousands of Cenozoic nannofossils are used to classify the particles further into morphometric groups, which roughly cluster at the family or genus level.

To verify that all coccoliths were properly classified, we manually examined every image that was identified as a coccolith, and particles that were deemed "potential coccoliths," but were not assigned a morphometric group during the neural network classifier. Segmented images were either coccoliths that were properly classified and manually kept in their SYRACO-assigned group, coccoliths that were improperly classified and manually reassigned to another group, non-coccoliths that were improperly classified and manually removed entirely, or coccoliths that were not classified and manually added to the proper group. Here we take all members of the Noelaerhabdaceae lineage as alkenone producers, and use the coccolith length measurements of this family for our cell size estimation. In modern and Pleistocene pelagic ecosystems, species in this family include *Emiliania huxleyi*, *Gephyrocapsa spp*., *Pseudoemiliania lacunosa*, and *Reticulofenestra spp.* (Young et al., 2003).

3.4. References

Aloisi G. (2015) Covariation of metabolic rates and cell size in coccolithophores. *Biogeosciences* **12**, 4665–4692.

Balch W. M., Holligan P. M. and Kilpatrick K. A. (1992) Calcification, photosynthesis and growth of the bloom-forming coccolithophore,Emiliania huxleyi. *Cont. Shelf Res.* **12**, 1353–1374.

Beaufort L. (2005) Weight estimates of coccoliths using the optical properties (birrefringence) of calcite. *Micropaleontology* **51**, 289–298.

Bendif E. M., Probert I., Carmichael M., Romac S., Hagino K. and de Vargas C. (2014) Genetic delineation between and within the widespread coccolithophore morpho-species Emiliania huxleyi and Gephyrocapsa oceanica (Haptophyta). *J. Phycol.* **50**, 140–148.

Bendif E. M., Probert I., Díaz-Rosas F., Thomas D., van den Engh G., Young J. R. and von Dassow P. (2016) Recent reticulate evolution in the ecologically dominant lineage of coccolithophores. *Front. Microbiol.* **7**.

Burkhardt S., Riebesell U. and Zondervan I. (1999) Effects of growth rate , CO2 concentration , and cell size on the stable carbon isotope fractionation in marine phytoplankton. *Geochim. Cosmochim. Acta* **63**, 3729–3741.

Eppley R. W., Rogers J. N. and McCarthy J. J. (1969) HALF-SATURATION CONSTANTS FOR UPTAKE OF NITRATE AND AMMONIUM BY MARINE PHYTOPLANKTON1. *Limnol. Oceanogr.* **14**, 912–920.

Feng Y., Roleda M. Y., Armstrong E., Boyd P. W. and Hurd C. L. (2017) Environmental controls on the growth, photosynthetic and calcification rates of a Southern Hemisphere strain of the coccolithophore Emiliania huxleyi. *Limnol. Oceanogr.* **62**, 519–540.

Fielding S. R. (2013) Emiliania huxleyi specific growth rate dependence on temperature. *Limnol. Oceanogr.* **58**, 663–666.

Henderiks J. and Pagani M. (2007) Refining ancient carbon dioxide estimates: Significance of coccolithophore cell size for alkenone-based pCO2 records. *Paleoceanography* **22**, 1–12.

Krumhardt K. M., Lovenduski N. S., Iglesias-Rodriguez M. D. and Kleypas J. A. (2017) Coccolithophore growth and calcification in a changing ocean. *Prog. Oceanogr.* **159**, 276–295.

McClelland H. L. O., Barbarin N., Beaufort L., Hermoso M., Ferretti P., Greaves M. and Rickaby R. E. M. (2016) Calcification response of a key phytoplankton family to millennial-scale environmental change. *Sci. Rep.* **6**, 1–11.

McClelland H. L. O., Bruggeman J., Hermoso M. and Rickaby R. E. M. (2017) The origin of

carbon isotope vital effects in coccolith calcite. *Nat. Commun.* **8**, 14511.

Nanninga H. J. and Tyrrell T. (1996) Importance of light for the formation of algal blooms by Emiliania huxleyi. *Mar. Ecol. Prog. Ser.* **136**, 195–203.

Perrin L., Probert I., Langer G. and Aloisi G. (2016) Growth of the coccolithophore Emiliania huxleyi in light- and nutrient-limited batch reactors: relevance for the BIOSOPE deep ecological niche of coccolithophores. *Biogeosciences* **13**, 5983–6001.

Polissar P. J. and D'Andrea W. J. (2014) Uncertainty in paleohydrologic reconstructions from molecular δD values. *Geochim. Cosmochim. Acta* **129**, 146–156.

Popp B. N., Laws E. A., Bidigare R. R., Dore J. E., Hanson K. L. and Wakeham S. G. (1998) Effect of phytoplankton cell geometry on carbon isotopic fractionation. *Geochim. Cosmochim. Acta* **62**, 69–77.

Riebesell U., Revill A. T., Holdsworth D. G. and Volkman J. K. (2000) The effects of varying CO2 concentration on lipid composition and carbon isotope fractionation in Emiliania huxleyi. *Geochim. Cosmochim. Acta* **64**, 4179–4192.

Young J. R., Geisen M., Cros L., Kleijne A., Sprengel C., Probert I. and Østergaard J. (2003) A guide to extant coccolithophore taxonomy. *J. Nannoplankt. Res.*, 125.

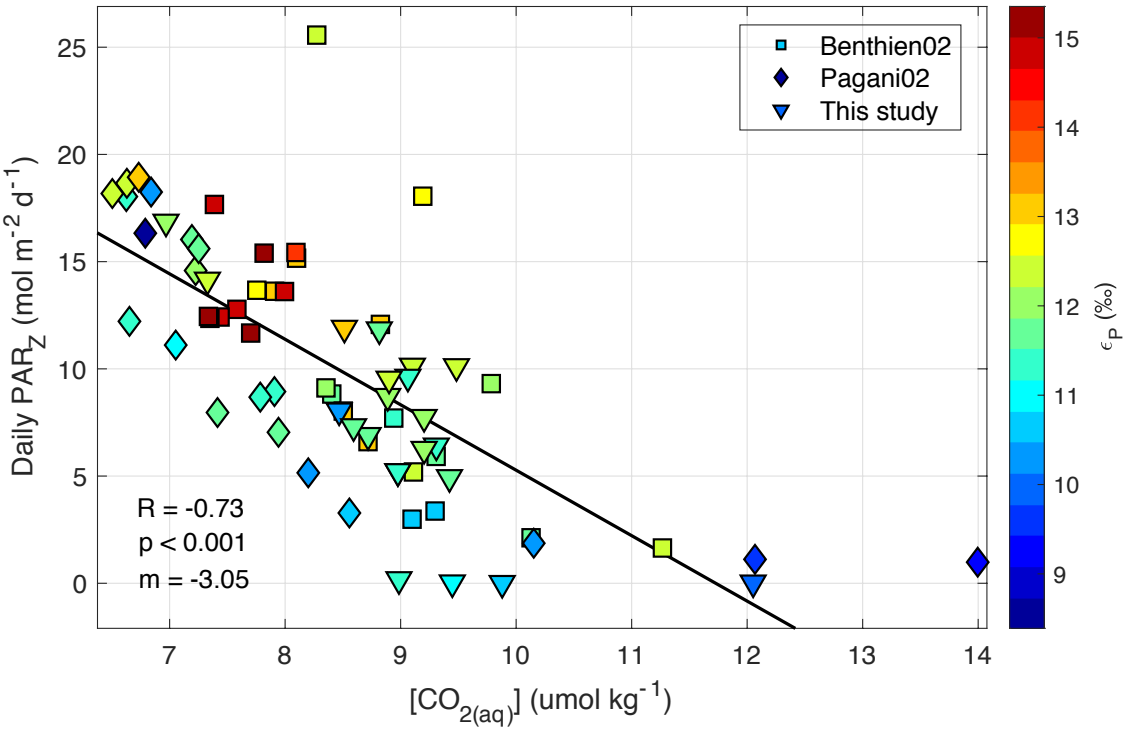

Figure S1. CO_2 and irradiance at the production depth for sediment samples. The production depth at each location is set to 75% of the mixed-layer depth, and irradiance at depth (PAR_Z) is calculated from surface irradiance and the diffuse attenuation coefficient. We find a strong relationship between irradiance, $[CO_{2(aq)}]$, and ε_p. Warm colors (high ε_p) cluster at high irradiance and low $[CO_{2(aq)}]$, while low ε_p occurs at high $[CO_{2(aq)}]$ and low irradiance.

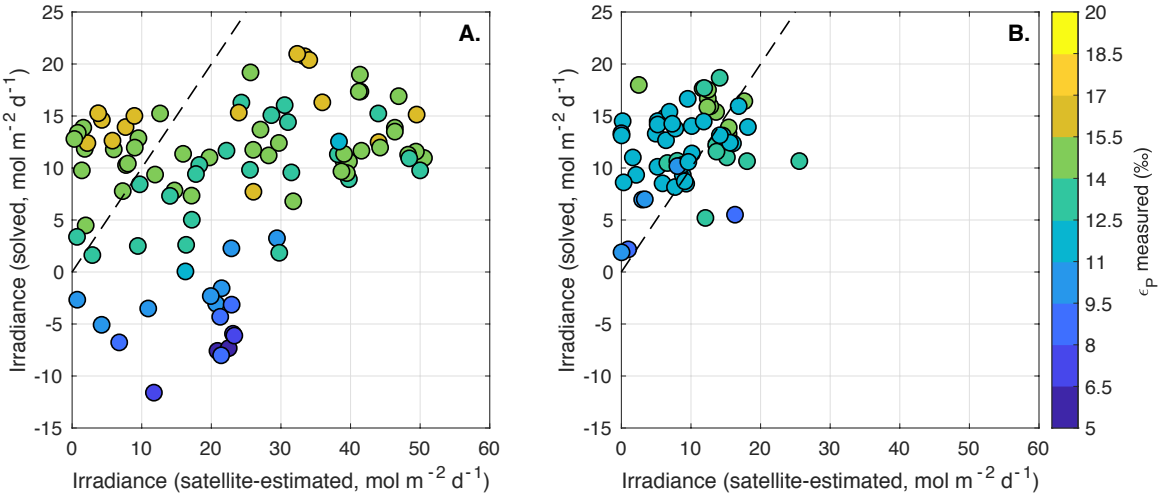

Figure S2. Solved irradiance compared to satellite-estimated irradiance. Using the measured ε_p, cell size, and known [$CO_{2(aq)}$], we solve for the irradiance value needed to perfectly model ε_p using the integrated irradiance model of Chapter 2 in (A) particulates and (B) sediments. Symbols are colored by the measured ε_p value. Dashed lines in each panel are 1:1 lines. In the particulate samples, the satellite-estimated irradiance at the collection depth overestimates the solved irradiance by up to a factor of 5. We suggest this is because the alkenones collected at a given particulate sampling depth are towards the top of the mixed layer. The sample does not solely reflect the environmental conditions of that sampling depth, and the sample entrains a history of surrounding environmental conditions. In contrast, the solved irradiance is on average higher than the satellite-estimated irradiance from the sediment samples. This offset explains why the regression line in Figure 6B (Chapter 3) is slightly below the 1:1 line.

Table S1. Michaelis-Menten half-saturation constants for *Emiliania huxleyi*

Substrate	Min	Reference	Max	Reference	Value used	Reference
NO_3^- ($\mu mol\ kg^{-1}$) [a]	0.1	Eppley et al. (1969)	13.71	Feng et al. (2017)	0.35	Perrin et al. (2016)
PO_4^{3-} ($\mu mol\ kg^{-1}$)	0.051	Perrin et al. (2016)	0.31	Perrin et al. (2016)	0.31	Perrin et al. (2016)
pCO_2 (μatm)	40.2	Krumhardt et al. (2017)	54.3	Krumhardt et al. (2017)	54.3	Krumhardt et al. (2017)
Light (mol photons $m^{-2}\ d^{-1}$) [b]	1.8	Nanninga and Tyrrell (1996)	52.1	Balch et al. (1992)	52.1	Balch et al. (1992)

[a] Max value is an outlier compared to other measurements and is ~40-fold higher than others. We take max of similarly-ranged estimates.
[b] Listed as the measured ranges in the caption of Krumhardt et al. (2017), Figure 6.

Appendix C: Supplementary Material for Chapter 4

1. Site descriptions and samples

1.1. ODP Site 807 (this study)

We selected 11 samples from ODP Site 807A (3.607°N, 156.625°E, water depth 2805 m) spanning the last ~650 ka. The Pleistocene section of this core has been studied extensively, and at the time of our sampling, the Late Pleistocene section was almost entirely depleted. Late Pleistocene sediments consist primarily of foraminifera nannofossil ooze, with between 86% and 92% calcium carbonate and good nannofossil preservation (Shipboard Scientific Party, 1991). The age model is described in the main text. Briefly, we used published age models from Zhang et al. (2007) and Prentice et al. (1993) for the 0-550 ka and 550-800 ka, respectively, and updated that of Prentice et al. (1993) to the LR04 stack (Lisiecki and Raymo, 2005). This location is ~0.5% out of equilibrium today on a mean annual basis, making it a good recorder of global atmospheric CO_2 (see Equation S1 for calculation of disequilibrium) (Takahashi et al., 2014).

1.2. ODP Site 1207 (this study)

Sediments in the Pleistocene section of ODP Site 1207 (37.79°N, 162.75°E, 3103 m water depth) are mainly nannofossil ooze, interspersed with some diatom and clay contributions (Shipboard Scientific Party, 2002). Because no detailed isotope stratigraphy exists for ODP 1207A, we first generate a coarse age model ("1207 initial age model") from the available Pleistocene biostratigraphic and magnetostratigraphic markers. In Analyseries 2.0.8 (Paillard et al., 1996), we evenly resample the reflectance data between 1.5 ka and 998.1 ka at the mean

sample spacing (1.55 kyr) using piecewise simple linear interpolation. We then bandpass filter the resampled reflectance data at 41 kyr (frequency = 0.024 kyr^{-1}; bandwidth = 0.005; Gaussian shape) and tune this bandpass-filtered timeseries to the obliquity solution of (Laskar et al., 2004). Peaks in the bandpass-filtered reflectance data correlate with obliquity peaks, and we identify 41 tie points in this 1 Ma section. We then export the tie points from Analyseries, convert the ages on the "1207 initial age model" back to depth, and assign these depths "true" ages using the obliquity tie points. The tuned age model agrees very well with the biostratigraphic and magnetostratigraphic constraints, passing through the upper and lower bounds of these markers. This site is approximately -11% out of equilibrium today on a mean annual basis (Takahashi et al., 2014).

1.3. ODP Site 806

ODP Site 806 has been studied extensively and used to generate some canonical paleoceanographic records (Lea et al., 2000; Wara et al., 2005). Site 806 is located on a topographic high on the Ontong-Java Plateau in the western equatorial Pacific (0.3185°N, 159.3615°E, water depth 2305 m). This site is ~3‰ out of equilibrium today (Takahashi et al., 2014). Sediments examined in this study come from holes 806A and 806B. The vast majority of the alkenone ε_p data in the Plio-Pleistocene were originally published by Pagani et al. (2010), and two Late Pleistocene samples were added by Zhang et al. (2017). High-resolution benthic and planktonic foraminifera isotope records exist for holes 806B and 806A, which are used as age constraints. Because neither Pagani et al. (2010) or Zhang et al. (2017) report original sample intervals with core, section, and depth information, we obtained these details from the sample request repository available through the Janus database maintained by IODP

(http://iodp.tamu.edu/janusweb/links/links_all.shtml). The foraminifera carbon isotope data used for this record were originally published by Wara et al. (2005), who did report hole, section, and sample intervals. We confirmed that these materials were from the same samples by comparing the joint $\delta^{18}O_{foram}$ and $\delta^{13}C_{foram}$ data in Wara et al. (2005) with the values reported by Zhang et al. (2017), finding them to be identical, allowing us to assign sample information to the samples of Zhang et al. (2017). However, we identified a slight mismatch between the δ18O record of Wara et al. (2005) and those of Berger et al. (1993) and Medina-Elizalde and Lea (2005). With the primary sample depths, we converted all mbsf values to updated meters composite depth (mcd) by linearly interpolating between tiepoints correlating hole A to hole B, as opposed to applying a standard offset. This improves the $\delta^{18}O$ agreement in the upper 2 meters mcd between the record of Wara et al. (2005) and the others. We then converted depth (mcd) to age using the age model of Medina-Elizalde and Lea (2005).

1.4. ODP Site 850

ODP Site 850 is located in the eastern equatorial Pacific, on the edge of the eastern equatorial Pacific cold tongue (1°17.83'N, 110°31.29'W, 4197 m water depth), where cold, high p_{CO2} thermocline waters upwell to the surface due to wind shear along the equator. Because of this, it is ~12% supersaturated with respect to atmospheric pCO_2. Zhang et al. (2017) measured alkenone ε_p in 42 samples from ODP 850; of these, six are younger than 800 ka. That study examined broad, long-term changes in alkenone ε_p over the last ~7 Ma; accordingly, ages were assigned using low-resolution biostratigraphy and magnetostratigraphy. As with ODP 806, we reconstructed primary sample information using the Janus database. To extract contemporaneous atmospheric CO_2 values from the ice core CO_2 record, we update ages using the age model of

Lyle et al. (2019), which ties Site 850 to nearby ODP Site 849 using physical properties, the isotope stratigraphy of which is correlated to the Ceara Rise $\delta^{18}O$ stack (Wilkens et al., 2017). Ages for the six samples < 800 ka change by an average of ~33 kyr in absolute terms with the new chronology.

1.5. ODP Site 999

Seki et al. (2010) and Badger et al. (2019) generated alkenone ε_p data for much of the Pleistocene and Pliocene at ODP Site 999. This core was taken in the Caribbean Sea (12.74°N, 78.74°W, 2838 m water depth) and has an orbitally-constrained age model in the late Pleistocene, which was tuned to the LR04 stack (Schmidt et al., 2006). Three samples from Seki et al. (2010) overlap the ice core CO_2 record. We update the ages of the two Seki et al. (2010) samples that are within the Pleistocene time series of Badger et al. (2019) by linear interpolation, shifting the ages by ~3 kyr. Seki et al. (2010) report one additional sample overlapping the ice core record at 450 ka, but it is too far from the Badger et al. (2019) to update in the same manner. Given the minor age modification in the overlapping section, we keep the Seki et al. (2010) age assignment of 450 ka, and do not rebuild the chronology from the isotope records. This site is +0.4% out of equilibrium on an annually-averaged basis (Takahashi et al., 2014).

1.6. ODP Site 1241

ODP Site 1241 is located in the eastern equatorial Pacific (5.85°N, 86.45°W, 2027 m water depth), in the Eastern Pacific Warm Pool. Seki et al. (2010) generated a low-resolution alkenone $\delta^{13}C$ record over the last ~5 Ma at ODP Site 1241, with four samples overlapping the ice core CO_2 record. However, the chronology for this site from 0 to 2.5 Ma is based exclusively

on biostratigraphy, and the location of samples with respect to glacial-interglacial variability is not known. For this reason, we exclude these data from our analysis.

1.7. GeoB 1016-3

The core site GeoB 1016-3 is located in the Angola Basin (11.77°S, 11.68°E, 3410 m water depth) and is about 185 km from the Angolan coast. This location has an mean annual disequilibrium of +1.3% (Takahashi et al., 2014). An alkenone ε_p record spanning the last ~200 ka from this core was reported by Andersen et al. (1999). The chronology is based on *G. ruber* (pink) $\delta^{18}O$ stratigraphy correlated to the SPECMAP chronology. We update these ages to the LR04 timescale by linear interpolation, resulting in an average shift of approximately +2 kyr.

1.8. DSDP 619

The alkenone ε_p record from DSDP 619 was published by Jasper and Hayes (1990) and was the first to suggest that carbon isotope discrimination recorded in alkenones could be used to reconstruct known variations in past CO_2 levels. The core site is on the continental slope, about 225 kilometers off the coast of Louisiana, in the Pigmy Basin in the Gulf of Mexico (27.1935°N, 91.409°W, 2259 m water depth). The age model for this study is based on the *G. ruber* $\delta^{18}O$ record from this core, which is correlated to the SPECMAP chronology; we update the ages to the LR04 timescale. This core site is a net sink of CO_2 on an annual basis, with an average disequilibrium of ~ -2.9% (Takahashi et al., 2014).

1.9. MANOP C

The gravity core W8402A-14GC was taken from site MANOP C in the central equatorial Pacific (0.953°N, 138.955°W, 4287 m water depth). Originally published by Jasper et al. (1994), the age model was derived by correlating the oxygen isotope records of deep-dwelling planktonic foraminifera *Globorotalia tumida* and benthic foraminifera *Cibcidoides wuellerstorfi* to the SPECMAP chronology. We update the published ages to the LR04 timescale. The authors do not publish the exact foraminifera $\delta^{13}C$ values used for their calculations, as the depths and ages of the alkenone samples are slightly different from the foraminifera samples. We linearly interpolate the reported *G. tumida* $\delta^{13}C$ record to the sample levels used for alkenone $\delta^{13}C$ analysis. Being in the central equatorial Pacific, this location is characterized by persistent upwelling, and is quite supersaturated in CO_2 with respect to the atmosphere, with a +13.7% disequilibrium on a mean annual basis (Takahashi et al., 2014).

1.10. ODP Site 925

ODP Site 925 is located on the Ceara Rise in the western tropical Atlantic (4°12.25′N, 43°29.33′W, 3042 m water depth). The core site is about 625 km from the nearest coastline, and about 800 km from the Amazon delta; it is approximately 450 km from the 120m isobath, which we estimate to be the LGM coastline. The majority of the Pleistocene record was generated by Pagani et al. (2010) (13 samples overlapping the ice core era), and Zhang et al. (2013) added three samples from the last ~30 ka. All Pleistocene samples come from hole C, and Zhang et al. (2013) report sample depths on the mcd scale. A Pleistocene benthic foraminifera $\delta^{18}O$ record was originally published by , which was included in the LR04 stack. The original benthic $\delta^{18}O$ record did not report primary sample intervals or which hole the samples came from, only the sample depths on mcd scale. Fortunately, because the age model is part of the LR04 stack and

the depths of Zhang et al. (2013) are on the mcd scale, we do not need to update this chronology. However, we note that the published alkenone ε_p values in the electronic supplement of Zhang et al. (2019) cannot be directly compared to our recalculations, because there is a clerical error in their supplemental table, giving rise to $\varepsilon_{p37:2}$ values of ~25‰. This site is a slight source of CO_2 to the atmosphere on a mean annual basis, with an average disequilibrium of +1.4% (Takahashi et al., 2014).

1.11. MD01-2392

Core site MD01-2392 is located in the southern South China Sea (09°51.13'N, 110°12.64'E, 1966 m water depth), and is approximately 215 m from the coast and ~175 m from the LGM shoreline. Zhang et al. (2019) report data from 23 samples over the last ~400 ka. The chronology for this core was developed by tuning a high-resolution TEX_{86} temperature record (Dong, 2016) to the EPICA Dome C CO_2 record, which do not require an update to be consistent with our study. This site is -1.8% out of equilibrium today (Takahashi et al., 2014).

1.12. ODP Site 668

ODP Site 668 is located in the eastern tropical Atlantic on the Sierra Leone Rise (4°46.23'N, 20°55.62'W, water depth 2693 m). Zhang et al. (2019) generated an alkenone ε_p record covering the last ~800 ka with 22 samples from hole 668B. Because of its position above the lysocline and near-equilibrium of surface waters with atmospheric pCO_2, this site has been used to generate CO_2 reconstructions through the Pleistocene from the boron isotope ratios in planktonic foraminifera (Hönisch and Hemming, 2005; Hönisch et al., 2009; Dyez et al., 2018). The core site is about 920 km from the coast and over 800 km from the estimated LGM shoreline

(120 m isobath). This site has a mean annual relative disequilibrium of ~ -2.7% (Takahashi et al., 2014).

1.13. NIOP464

Palmer et al. (2010) published an alkenone ε_p record covering the last ~30 ka from sediment core NIOP464 on the Murray Ridge in the Arabian Sea (22.15°N, 63.35°E; 1470 m water depth). The site is ~340 m from the nearest coastline and ~315 m from the nearest 120 m isobath. The age model for this site is based on a $\delta^{18}O$ record of *Neogloboquadrina dutertrei*, which was originally correlated to the SPECMAP chronology. We update the ages to the LR04 timescale. As the Arabian Sea is a site of strong seasonal upwelling, this site has a mean annual disequilibrium of +5.7% (Takahashi et al., 2014).

1.14. 05PC-21

Sediment core 05PC-21 was taken from the East Sea/Japan Sea (38.40°N, 131.55°E, 1721 m water depth). Bae et al. (2015) reported alkenone ε_p data over the last ~200 kyr. However, we are unable to resolve whether their data includes the fractionation between $CO_{2(g)}$ and $CO_{2(aq)}$, and it is unclear whether their ε_p values account for the lipid to biomass isotope fractionation, because it is not mentioned in the text. If $\varepsilon_{bio/alk}$ has not been included, it would make the measured ε_p values some of the lowest in the global ocean. We were unable to obtain the original data from the study's authors.

2. Supplementary methods

2.1. In-situ [$CO_{2(aq)}$] estimation

In order to reconstruct past variations in atmospheric CO_2, the core location should be close to CO_2 equilibrium with the atmosphere. However, here we are interested in understanding the alkenone ε_p response in different oceanographic regimes to known variations in CO_2, and therefore consider upwelling areas important ecosystems where the ε_p response—due to elevated CO_2 levels—may be different from oligotrophic regions. We handle disequilibrium by calculating the ratio of in-situ pCO_2 at each location to the reference year atmospheric CO_2, and then apply this constant fractional disequilibrium down-core. The in-situ aqueous CO_2 at a given location at a given time is a function of the atmospheric CO_2 concentration:

$$[CO_2]_{aq}\{in\text{-}situ\}|t,s = K_0 * pCO_2\{atm\} * k_{md} \quad \text{[Eq. S1]}$$

where K_0 is the first dissociation constant (Henry's constant), and k_{md} is the modern fractional disequilibrium (defined as in-situ pCO_2 / reference year atmospheric pCO_2). With the GLODAP Mapped Climatology, the reference year is 2002 (Key et al., 2015; Lauvset et al., 2016; Olsen et al., 2016) with an atmospheric CO2 of 373.15 ppm. For the Takahashi et al. (2014) dataset, the reference year is 2005 (Takahashi et al., 2014), with an atmospheric CO_2 of 379.46 ppm. Reference year atmospheric pCO_2 values were retrieved from (ftp://aftp.cmdl.noaa.gov/products/trends/co2/co2_annmean_mlo.txt).

2.2. Lipid extraction and handling

We obtained 11 samples from ODP 807A and 21 samples from ODP 1207A from the International Ocean Drilling Program core repositories. Where possible, we scraped the exterior surfaces of samples to remove any organic contaminants from sampling. Samples were then lightly disaggregated, frozen, and freeze-dried for ~48 hours. All glassware was combusted at 450°C for 12 hours before use. Free lipids from sediment samples (8.5 to 25.7 g dry weight) were extracted using a Dionex ASE 350 at 100°C and maximum 1000 psi using a 9:1 v/v dichloromethane and methanol mixture (VWR Omnisolv) using four ten-minute static cycles and a rinse volume of 150%. An internal standard of 20.03 µg stearyl sterate (stearic acid stearyl ester) was added to the total lipid extract (TLE) of each sample, and the TLE was then evaporated to dryness under a pure N_2 stream. Lipid extracts were split into aliphatic, ketone/ester, and polar fractions by silica gel flash column chromatography and using hexane, dichloromethane, and methanol, respectively, as eluents. The ketone/ester (F2) fraction containing the alkenones was evaporated to dryness, transferred in dichloromethane to a 2 mL borosilicate vial for analysis. Initial screening of the F2 fraction from ODP 1207 identified long-chain unsaturated fatty acid methyl esters that would interfere with alkenone carbon isotope measurements. An aliquot (~1/2 volumetrically) of each ODP 1207 F2 fraction was transferred in dichloromethane to a 4 mL borosilicate vial and evaporated to dryness. Approximately 2 ml of 0.5 M KOH in MeOH/H_2O (95:5 v/v) was added to this vial, which was capped tightly and heated to ~75°C for approximately 2 hours. About 0.5 ml of 5% NaCl in ultrapure H_2O was added to each vial, followed by 1.5 ml of toluene. Samples were shaken vigorously for ~30 seconds, centrifuged for four minutes at 4000 rpm to facilitate phase separation, and the toluene fraction containing the ketones was removed and transferred to a new vial. This toluene extraction was carried out two additional times, and the ketone fraction was evaporated,

transferred to a 2 ml borosilicate vial, and dissolved in 100 µl of toluene for gas chromatographic analysis.

2.3. Alkenone SST estimates

We measured alkenone abundance and unsaturation ratios by gas chromatograph flame ionization detection (GC-FID). ODP 807 samples were measured on an Agilent J&W DB-1 column (60 m x 0.25 mm i.d. x 0.1 µm phase thickness) with a 10 m guard column. Samples were dissolved in toluene and injected using a split/splitless (SSL) injector operating in splitless mode at 300°C, with a splitless time of 1.5 minutes and helium as the carrier gas at a constant flow of 1.0 ml/min. The oven temperature program initiated at 90°C, which was held for 3 minutes, then ramped at 25°C/min to 250°C, ramped 1.0°C/min to 305°C, then heated at 10°C/min to 320°C and held for 20 minutes. An internal laboratory standard (alkenones from an *Emiliania huxleyi* culture) was run every six samples to assess instrument drift. ODP 1207 samples were also analyzed on a 60 m DB-1 with a 10 m guard column, but using a programmable temperature vaporizing (PTV) injector in splitless mode, also using toluene as the carrier solvent and helium as the GC carrier gas. The inlet was held at 90°C for 0.1 minutes, then heated at 10°C/second to 325°C, with ramped pressure and a splitless time of 5 minutes. The GC oven was held at 90°C for 1.5 minutes, then ramped at 25°C/min to 250°C, then ramped at 1°C/min to 313°C, and finally ramped at 10°C/min to 320°C and held for 20 minutes. We use the calibration of Müller et al. (1998) to convert $U_{37}^{K'}$ values to sea surface temperatures. Where unsaturation ratios are reported in our literature compilation, we recalculate SST using the Müller et al. (1998) calibration.

2.4. Alkenone carbon isotope measurements

Alkenone carbon isotope ratios were measured by gas chromatograph isotope ratio mass spectrometry (GC-irMS). We used a Thermo GC Ultra coupled to an Isolink, Thermo ConFlo IV interface, and a Thermo Delta V isotope ratio mass spectrometer. Samples were injected in toluene using a PTV inlet in operating in splitless mode with a 2 mm borosilicate liner with deactivated glass wool.

ODP 1207 samples were analyzed in March, 2018 using an Agilent J&W VF-200 column (60 m x 0.25 mm i.d. x 0.1 μm phase thickness) with a 10 m guard column. The PTV inlet was held at 90°C for 0.1 minutes during injection, then ramped to 275°C and held in splitless mode for three minutes, then heated to 350°C for a cleaning phase. The GC program started at 90°C, ramped 25°C/min to 255°C, then 2°C/min to 300°C, then 10°C/min to 320°C, and held isothermal for 15 minutes. The helium carrier gas flow rate was 1.0 ml/min. For some sample injections, water was removed using in-line Nafion® water removal system, though the majority were analyzed using an ethanol cold trap at -85°C. We observed a systematic offset of ~0.1‰ between the water removal systems, which was accounted for during the reduction using the Polissar and D'Andrea (2014) uncertainty determination.

ODP 807 samples were analyzed July, 2015 using an Agilent J&W DB-1 (60 m x 0.25 mm i.d. x 0.1 μm phase thickness) column, also fitted with a 10 m guard column. The injector initiated at 90°C, which was held for 0.1 minutes, then ramped at 14.5°C/sec to 275°C and held in splitless mode for 4.9 minutes, and then heated to 350°C and the split vent opened. The GC program started at 90°C, ramped 25°C/min to 250°C, then 1°C/min to 313°C, then 10°C/min to 320°C. Water from combustion was removed using an in-line Nafion® membrane with a dry helium counter flow. In both analytical sessions, the guard column was trimmed daily and the liner was changed every few days.

Sample $\delta^{13}C$ values were realized on the Vienna Pee Dee Belemnite (VPDB) scale using Mix A and B (acquired from A. Schimmelman, Indiana University) and the MATLAB routine of Polissar and D'Andrea (2014). We also monitored instrument stability and reproducibility using an internal laboratory alkenone standard derived from *Emiliania huxleyi* cultures. The *E. huxleyi* culture standard was identical within error of the reference gas value during the two analytical sessions, showing no difference in $\delta^{13}C$ with respect to GC method, GC column, or time. Rabinowitz et al. (2017) demonstrated alkenone unsaturation ratios are stable once compounds are extracted and stored in solvent; our analyses show that the $\delta^{13}C$ values of dissolved alkenones are stable as well. Uncertainty in our $\delta^{13}C_{37:2}$ our measurements includes the transfer from the internal laboratory reference gas scale to the VPDB scale, as well as sample measurement uncertainty. Samples were analyzed two to four times, and the average precision was 0.26‰ (1 SE).

2.5. Foraminifera isotope measurements

Between ~6 and 15 g of dry sediment was washed to isolate the coarse fraction. At LDEO, the bulk sediment recovered from ASE lipid extraction was weighed, placed in a 125 ml Nalgene bottle, suspended in tap water, and disaggregated on a Cambridge rotating wheel for ~1.5 hours. Sediments were then sieved at 63 μm using tap water, and the fine and coarse fractions were collected. The coarse fraction was oven-dried at 60°C, then weighed. The fine fraction was allowed to settle, and the supernatant was decanted, the sediment concentrated, then oven-dried at 60°C. *Orbulina universa* were picked from the 250-300, 300-355, 425-500, and 500-600 μm size fractions. *C. wuellerstorfi* and *Uvigerina spp*. were picked from the 250-300 μm and 300-355 μm size fraction.

2.6. SYRACO coccolith size measurements

We measured coccolith size by automated light microscopy using the SYRACOv2 system. The specifics of this approach are described in detail in Beaufort and Dollfus (2004), and we followed the same specific preparation technique as in Chapter 2 (Appendix B). Briefly, sediment was suspended in ~10 ml dilute ammonium hydroxide solution and randomly settled for ~4 hours in petri dishes on to 12 mm x 12 mm coverslips. At least two coverslips were prepared for each sample, which were mounted on microscope slides with EPOTENCY NOA81. Between 165 and 297 images per coverslip were acquired automatically using a LEICA microscope with circularly polarized light. Particles in each image were programmatically isolated from the black background, and classified into groups, roughly at the genus level, using an artificial neural network. Morphometric parameters are generated automatically by the program.

2.7. Biomass-weighted coccolith size estimates

In Chapter 3, we argued that, because larger cells have more biomass, and alkenone concentration is proportional to cellular carbon content (Popp et al., 1998a; Riebesell et al., 2000; Prahl et al., 2003), the alkenone ε_p signal preserved in sediments is likely biased towards larger cells. We therefore calculate the biomass-weighted mean cell radius estimate the biomass following the approach in Chapter 3. This method treats each coccolith as an observation of one individual coccolithophore. In each sample, we take the distribution of coccolith sizes, and calculate the cell radius equivalent to that coccolith size after McClelland et al. (2016). Noelaerhabdaceae coccolith length is related to cell radius by:

$$R_C = 0.57 \times \mathrm{CL} + 0.45 \qquad \text{[Eq. S2]}$$

We assume a spherical geometry for the Noelaerhabdaceae and calculate the cellular volume for each individual. We convert cell volume to biomass (particulate organic carbon, POC) using the relationship derived in Chapter 1 (Eq. S3), and we note that using a different transfer function, such as that of Popp et al. (1998b) does not have a marked impact on the resulting biomass-weighting.

$$\mathrm{POC} = 17.5 \text{ fmol C } \mu\mathrm{m}^{-3} \qquad \text{[Eq. S3]}$$

We then sum the individual observations of POC, and generate a total population biomass in each sample. We divide each individual POC value by this total population biomass to get the fractional contribution of each individual to the total biomass in the sample, and these fractional values ("weights") are used as the weighting function in the biomass-weighted mean. We then take the vector product of the weights times the individual coccolith size observations to give a biomass-weighted (BMW) coccolith size estimate. We then calculate the biomass-weighted radius from this BMW coccolith length value using Equation S1, and use this radius value for all ensuing analysis.

2.8. Coccolith size estimates from assemblage data

ODP 807, 1207, 806, 850, and 999 have coccolith size measurements accompanying the alkenone ε_p data to varying degrees of completeness. The best-case scenario (ODP 807, 1207, and 999) is the complete list of individual coccolith size observations in each sample so that the biomass-weighted cell radius can be calculated (see Section 3.2.6). At ODP 806 and 850, only an arithmetic mean coccolith length is reported for a subset alkenone ε_p measurements. However, we aim to take advantage of as much data as possible to evaluate the influence of cell size changes on alkenone ε_p variations. At DSDP 619, ODP 668, GeoB 1016-3, ODP 925, and ODP 807, we derive estimates of coccolith size using existing assemblage-level coccolith relative abundance timeseries (Constans and Parker, 1986; Müller et al., 1997; Kinkel et al., 2000; Henriksson, 2000; Chiyonobu et al., 2006). We evaluate the assemblage-estimation method using these data and our own actual coccolith length measurements at ODP 807.

2.8.1. DSDP 619

We compile coccolith assemblage data from Constans and Parker (1986), who published semiquantitative nannofossil abundances from DSDP 619, which are digitally available at the NGDC data repository. Different species were identified in terms of relative abundance, with levels between "trace" and "highly abundant," which relate to the number of individuals of a given species per field of view. The authors do not report the number of fields of view examined, but we can compare the relative abundances of the Noelaerhabdaceae morphotypes reported. We convert the relative abundances to quantitative values using the following number per field conversion factors: trace = 0.001, rare = 0.00545, few = 0.05, common = 0.03, abundant = 5.5, very abundant = 55, highly abundant = 100. We determine these values as the average of the inverse of the ranges of fields of view per specimen given in the caption of their Table 11. We

calculate the relative abundances of different Noelaerhabdaceae species using these the assemblage data and the aforementioned conversion factors. *Emiliania huxleyi* dominates the assemblage through the record, but because there is no morphotype distinction there is considerable uncertainty in the size estimate.

2.8.2. ODP Site 668

Manivit (1989) reports semiquantitative nannofossil abundances in ODP 668A. However, these data are low resolution and have very loose morphotype information. Importantly, no *E. huxleyi* abundances are reported, though this species is found in abundance in this region today and in the Pleistocene (Henriksson, 2000; Bendif et al., 2019). Instead, we take advantage of a record of coccolith abundances over the last ~200 ka at the nearby core 16772 (1°21'S, 11°58'W, 3912 m water depth) investigated by Henriksson (2000). That study published relative and absolute abundance data of alkenone producing coccoliths. We digitize their figure 5 and estimate the abundance of each nannofossil group through time. We update the age model of Henriksson (2000) from (Martinson et al., 1987) to the LR04 stack, to which Zhang et al. (2019) is tuned. The largest uncertainty in the digitization arises during large and rapid fluctuations in the relative abundances, which is only an issue during the 105 to 120 ka interval, which overlaps with 1 sample from Zhang et al. (2019). Based on the morphotype data in the CLIMAP surface dataset, we assume the *E. huxleyi* morphotype at this location is Type A (CLIMAP Project Members, 1976).

2.8.3. ODP Site 925

Kinkel et al. (2002) report percentages of *E. huxleyi*, *G. oceanica*, and *G. ericsonii* from sediment core GeoB 1523-1 (03°49.9'N, 41°37.3'W, water depth 3292 m) over the last ~120 ka. Nine samples overlap with the record of Zhang et al. (2019), and we extrapolate ~10 kyr (~4 m) from the oldest sample to extend the overlap to 10 samples.

2.8.4. GeoB 1016-3

Müller et al. (1997) measured abundances of *E. huxleyi*, *G. oceanica*, *G. muellerae*, *G. aperta*, *G. ericsonii*, and *Gephyrocapsa spp.* >3 μm and <3 μm, as well as two other non-alkenone producers at sediment core GeoB1028-5 on the Walvis Ridge over the last ~400 kyr (20.1°S, 9.18°E, 2209 m water depth) . The data were available digitally through the PANGAEA database, so we did not digitize any figures. We calculate the relative abundances of each of these species, combining G. aperta and Gephyrocapsa spp. <3μm into a single group denoted "G. small," following the morphological definitions of the Nannotax web database. The data show an increase in the abundance of *E. huxleyi* from ~10% to 90% over the last ~80 kyr. Prior to this floral shift, the assemblage was dominated by small *Gephyrocapsa*, with fluctuating minor percentages of the other taxa. We update the age model to LR04.

2.8.5. ODP Site 807

Chiyonobu et al. (2006) published a record of the relative abundances of various coccolithophore species at ODP Site 807 over the last 500 ka. Of the Noelaerhabdaceae, the authors counted *E. huxleyi*, *G. oceanica*, *G. parallela*, *Gephyrocapsa* spp. (small), *P. lacunosa*, and *Reticulofenestra* spp. (small). We digitize their Figure 3, and with our algorithm for biomass-weighted coccolith length estimates from assemblages, estimate biomass-weighted

mean coccolith length at equivalent depths in our samples. We compare assemblage-based inferences to our actual measurements at ODP 807 as a test of the assemblage method.

2.9. Determination of ε_p

The carbon isotope discrimination by photosynthesis (ε_p) is defined by Freeman and Hayes (1992) as:

$$\varepsilon_p \equiv [(\delta^{13}C_{CO2aq} + 1000) / (\delta^{13}C_{POC} + 1000) - 1] \times 1000 \quad \text{[Eq. S4]}$$

where the subscript CO_{2aq} indicates aqueous carbon dioxide, and POC indicates particulate organic carbon; in this case, the primary photosynthate of Noelaerhabdaceae algae.

In alkenone ε_p sediment studies, neither of these parameters is measured directly. $\delta^{13}C_{POC}$ is estimated from $\delta^{13}C_{Mk37:2}$, and $\delta^{13}C_{CO2aq}$ is estimated from the $\delta^{13}C$ of planktonic foraminifera. Calculation of $\delta^{13}C_{POC}$ is straightforward. Lipids are generally depleted by ~3-5‰ relative to bulk organic carbon (Hayes, 1993), and alkenones are no exception. The $C_{37:2}$ methyl ketone is approximately 4.44‰ more negative than bulk particulate organic matter (Chapter 2). The alkenone to biomass fractionation is described mathematically as:

$$\varepsilon_{bio/alk} = 4.44 = [(1000 + \delta^{13}C_{POC} / (1000 + \delta^{13}C_{Mk37:2}) - 1] \times 1000 \quad \text{[Eq. S5]}$$

The carbon isotope ratio of the photosynthetic organic matter ($\delta^{13}C_{POC}$) is calculated by rearranging Equation S5:

$$\delta^{13}C_{POC} = (\varepsilon_{bio/alk} / 1000 + 1) \times (\delta^{13}C_{Mk37:2} + 1000) - 1000 \quad \text{[Eq. S6]}$$

Calculating $\delta^{13}C_{CO2aq}$ is less straightforward. All estimates use the carbon isotope ratio of calcifying microorganisms as tracers of mixed-layer $\delta^{13}C$. The classical method (e.g. Pagani et al., 1999) uses planktonic foraminifera and relates $\delta^{13}C_{foram}$ to $\delta^{13}C_{CO2(g)}$ using the laboratory-derived temperature-dependent fractionation of Romanek et al. (1992):

$$\varepsilon_{cal/CO2(g)} = 11.98(\pm0.13) - 0.12(\pm0.01) \times T_C \quad \text{[Eq. S7]}$$

where the subscript "cal" refers to $\delta^{13}C_{calcite}$ and "CO2(g)" refers to $\delta^{13}C_{CO2(g)}$, and T_C is temperature in Celsius. $\delta^{13}C_{CO2(g)}$ is then calculated by equation 8:

$$\delta^{13}C_{CO2(g)} = (\delta^{13}C_{calcite} + 1000) / (\varepsilon_{cal/CO2(g)} / 1000 + 1) - 1000 \quad \text{[Eq. S8]}$$

The fractionation between $CO_{2(g)}$ and $CO_{2(aq)}$ is determined using the temperature dependence of Mook et al. (1974):

$$\varepsilon_{CO2(g)/CO2(aq)} = -373 / T_K + 0.19 \quad \text{[Eq. S9]}$$

where T_K is temperature in Kelvin. Finally, $\delta^{13}C_{CO2aq}$ is calculated by:

$$\delta^{13}C_{CO2(aq)} = (\varepsilon_{CO2(g)/CO2(aq)} / 1000 + 1) \times (\delta13C_{CO2(g)} + 1000) - 1000 \quad \text{[Eq. S10]}$$

In most paleo applications, $\delta^{13}C_{CO2aq}$ is calculated using Equations S7 – S10. This method assumes that foraminiferal calcite is precipitated in isotopic equilibrium with the ambient seawater carbonate chemistry and can be treated the same as synthetic equilibrium calcite. However, many planktic foraminifera have obligate symbiotic relationships with photosynthetic dinoflagellates (Bé et al., 1982). Their photosynthetic activity preferentially removes ^{12}C from the foraminiferal microenvironment, generally increasing local $\delta^{13}C_{DIC}$ relative to the ambient mixed-layer chemistry. Although Romanek et al. (1992) demonstrated a constant, temperature-independent 1‰ enrichment in the $\delta^{13}C$ of synthetic calcite relative to HCO_3^-, the DIC used for foraminiferal calcification is not in isotopic equilibrium with its surroundings because of the symbiotic activity described above, and can be isotopically depleted due to metabolic fractionation (Wefer and Berger, 1991; Spero et al., 1997). The magnitude of this isotopic "vital effect" is a function of several environmental parameters, including shell size, light intensity, carbonate ion concentration, species, and habitat depth (Ezard et al., 2015). Intriguingly, Spero et al. (2003) have shown that the same species and size fraction can have a different $\delta^{13}C$ offset relative to DIC in different geographies.

Because of this complexity, we take a simple empirical approach to dealing with vital effects. At each location, we calculate the arithmetic isotope ratio difference between the core-top (or Late Holocene) $\delta^{13}C_{foram}$ value and the estimated preindustrial mixed-layer $\delta^{13}C_{DIC}$:

$$\Delta_{VE} = \delta^{13}C_{DIC} - \delta^{13}C_{foram} \qquad \text{[Eq. S11]}$$

We then add the vital effect offset to the measured $\delta^{13}C_{foram}$ value to estimate $\delta^{13}C_{DIC}$, and calculate $\delta^{13}C_{CO2aq}$ after Rau et al. (1996):

$$\delta^{13}C_{CO2(aq)} = \delta^{13}C_{DIC} + 23.644 - 9701.5 / T_K \quad \text{[Eq. S12]}$$

We do therefore assume that the direction and magnitude of the vital effect remains constant through time. Spero et al. (1999) suggested that $\delta^{13}C_{DIC}$ could be reconstructed from paired measurements of $\delta^{13}C_{foram}$ of different species with different $\delta^{13}C_{foram}$ sensitivities to carbonate ion concentration changes. While that approach may prove fruitful, we do not have those data at our disposal in this compilation. Because our global analysis aims to take advantage of all published Pleistocene alkenone $\delta^{13}C$ data, a consistent treatment of the data is necessary to compare the changes in alkenone ε_p between locations.

Our reported ε_p values will be different from published values for the following reasons. First, we recalculate $\delta^{13}C_{DIC}$ using the approach described above. Second, we use a standard $\varepsilon_{bio/alk}$ value of 4.44‰ to calculate $\delta^{13}C_{POC}$, which appears in Equation S4 and S5. Finally, we elect to use the Müller et al. (1998) alkenone temperature calibration, which is in some cases different from the one used in the original publication.

3. Error calculation and propagation

Errors considered in our data treatment are both measurement and calibration error. Full propagation of errors is essential for comparing the trends and absolute value of different variables in our study. Unless otherwise stated, we assume that errors are normally distributed and uncorrelated. We assign and calculate propagated errors as follows.

3.1. Sea surface temperature and salinity

Proxy-derived sea surface temperature estimates in this study come from alkenone unsaturation ratios ($U_{37}^{K'}$), which has a standard error of prediction of ~1°C (Müller et al., 1998). We apply modern salinity values from GLODAPv2 (Lauvset et al., 2016) and an uncertainty of ±1 psu (normal distribution), though we note that salinity has an essentially negligible influence on our calculations.

3.2. Foraminifera $\delta^{13}C$

Errors in carbon isotope measurements of foraminifera are among the smallest relative errors in the entire dataset. Analytical uncertainties are on the order of ~0.05‰ (1σ). We use the reported 1σ measurement uncertainty and assume a normal distribution. For published datasets that did not report $\delta^{13}C_{foram}$ uncertainties, we use the median of the entire dataset (0.04‰).

3.3. Alkenone $\delta^{13}C$

Many studies report uncertainty $\delta^{13}C_{Mk37:2}$ as the standard deviation of replicate measurements, without the number of measurements, and often it is not stated whether the reported uncertainty includes uncertainty in realizing analytes on the VPDB scale as opposed to the analysis-specific reference gas scale. Reported uncertainties in the literature range between 0‰ and ~0.7‰ (1σ). We calculate uncertainties (1 SEM) using the MATLAB routine of Polissar and d'Andrea (2014). Where uncertainties are not reported, we apply a value of 0.3‰, the median value of all reported 1σ uncertainties in the datasets compiled here. Because existing publications do not report the standard error of the mean and only the standard deviation of replicate measurements, we take the reported 1σ errors as 1 SEM and treat them in an equivalent

way as our calculated 1 SEM errors. We could attempt to convert reported 1σ errors to 1 SEM errors by assuming all samples were measured in duplicate, but because it is unclear whether existing studies accounted for reference gas uncertainty and the actual number of measurements is unknown, we choose to take the reported errors as 1 SEM.

3.4. Foraminifera $\delta^{13}C$ vital effect

There are few data available to determine the true value of this uncertainty, which is likely location- and species-dependent. We estimate pre-industrial surface ocean $\delta^{13}C_{DIC}$ values from visual interpretation of Figure 3 in Schmittner et al. (2013). The figure has contours of 0.2‰, which gives an uncertainty of ±0.1‰. We estimate the mean "preindustrial" $\delta^{13}C_{foram}$ value from as many late Holocene carbon isotope measurements as possible. From our compilation of foraminifera vital effect offsets, the average standard deviation of multiple core-top samples surrounding the core location, or multiple samples at the same location in the Late Holocene, is 0.17‰, though the standard error of the mean is much smaller (0.06‰) where we have the correct statistics. Adding the contouring uncertainty from Schmittner et al. (2013) and the observed $\delta^{13}C_{foram}$ variability in quadrature gives a typical uncertainty of 0.19‰ (1σ). To be conservative, we use an uncertainty of 0.25‰ (1 SEM) on the calculated vital effect.

3.5. DIC and $CO2_{(aq)}$ $\delta^{13}C$

We calculate $\delta^{13}C_{DIC}$ using Equation S11, which simply adds the vital effect offset to the $\delta^{13}C_{foram}$ value. The uncertainty in $\delta^{13}C_{DIC}$ is given by:

$$\sigma_{DIC}^2 = \sigma_{VE}^2 + \sigma_{foram}^2 \qquad \text{[Eq. S13]}$$

where the subscript "DIC" is the dissolved inorganic carbon $\delta^{13}C$ value, "VE" is the $\delta^{13}C$ vital effect, and "foram" is the measured foraminifera value. Typical 1σ errors are ~0.3‰.

The carbon isotope ratio of dissolved CO_2 depends on the carbon isotope ratio of dissolved inorganic carbon as well as the in-situ temperature (Mook et al., 1974). We calculate $\delta^{13}C_{CO2aq}$ using Equation 2 of Rau et al. (1996) which follows the treatment of Freeman and Hayes (1992). We determine uncertainty in $\delta 13C_{CO2aq}$ through Monte Carlo simulations (n = 10,000) using the proxy SST and foraminifera-estimated $\delta^{13}C_{DIC}$ and normal distribution of the respective uncertainties. We take the standard deviation of the resulting distribution as the uncertainty in $\delta^{13}C_{CO2aq}$, with a typical value of ~0.33‰.

3.6. Henry's coefficient and ice core-derived in-situ [$CO2_{(aq)}$]

We calculate Henry's law constant for CO_2 (K_0) using equation 12 of Weiss (1974), using gravimetric units. Using the sea surface temperature and salinity uncertainties as described above, we use a Monte Carlo approach (n = 10,000) to calculate K_0 and take the standard deviation of the Monte Carlo distribution as the 1σ uncertainty in K_0. Relative uncertainty in K_0 (σ_{K0} / K_0) is between 2 and 5%, depending on the precision of the SST estimate, with larger K_0 uncertainties at less precise SST estimates. Uncertainty in salinity has a minute and effectively negligible impact on the K_0 uncertainty.

The average precision on ice-core atmospheric CO_2 measurements is <2 ppm (Bereiter et al., 2015), while the annual cycle at one location has an amplitude of ca. 5-10 ppm (Keeling, 1960). We assume a relative 1σ error of 2.5% on our estimation of atmospheric CO_2 at our sample ages, which is an average of ~3 ppm across the Pleistocene compilation. This is in-line

with the applied CO_2 uncertainty of Badger et al. (2019). Uncertainty in the estimated in-situ $[CO_2]_{aq}$ is calculated by propagating this atmospheric pCO_2 uncertainty and the SST- and salinity-dependent uncertainty in K_0 using the formula:

$$\sigma_{CO2}^2 = \overline{CO_2} \times \sqrt{\left(\frac{\sigma_{pCO2}}{pCO_2}\right)^2 + \left(\frac{\sigma_{K_0}}{K_0}\right)^2} \quad \text{[Eq. S14]}$$

where the horizontal bar indicates the mean $[CO_{2(aq)}]$ value.

3.7. Algal biomass and alkenone ε_p

Propagated uncertainty in alkenone ε_p has been extensively discussed by previous studies (Freeman and Pagani, 2005; Badger et al., 2019). We estimate the uncertainty in algal biomass by adding the uncertainties in $\delta^{13}C_{Mk37:2}$ and $\varepsilon_{bio/alk}$ in quadrature, using a 1 SEM uncertainty on $\varepsilon_{bio/alk}$ of 0.058‰ (Chapter 2). Uncertainty in ε_p is calculated by:

$$\sigma_{\varepsilon p}^2 = \sigma_{CO2aq}^2 + \sigma_{biomass}^2 \quad \text{[Eq. S15]}$$

where the subscripts "CO2aq" and "biomass" indicate the carbon isotope ratios of dissolved CO_2 and algal biomass, respectively. On average, the error in ε_p is around 0.45‰.

3.8. Irradiance

We estimate the uncertainty on the irradiance value to apply to a sediment sequence from the distribution of solved values from the Pleistocene calibration exercise. Using the known CO_2,

measured cell radius, and measured $\varepsilon_{p37:2}$, we solve each sample for the PAR_Z value that brings the measured and modeled $\varepsilon_{p37:2}$ to agreement. At a given location, we recommend taking the standard error of the mean of these Pleistocene values, and using this as the uncertainty in irradiance to propagate downcore. At ODP 807, for example, solved irradiance has a standard error of ~0.43 mol m^{-2} d^{-1}. This equates to an uncertainty in CO_2 of ~ 0.42 μmol kg^{-1}. The resulting pCO_2 uncertainty from a 0.42 μmol kg^{-1} $CO_{2(aq)}$ uncertainty is a function of Henry's Law. Uncertainty in the dissociation constant (K_0) is described above (Section 3.6.). At an average temperature of 28°C, and salinity of 35, the pCO_2 uncertainty is ~16 μatm. This uncertainty is SST-dependent because K_0 and temperature are inversely proportional. Irradiance error will propagate to larger pCO_2 uncertainty at warmer SSTs, though the effect of K_0 is much smaller than that of $CO_{2(aq)}$: if the temperature used in this exercise were 15°C instead of 28°C, the resulting pCO_2 uncertainty would be ~11 μatm instead of ~16 μatm, a small difference relative to the total error propagation.

3.9. Cell radius

We calculate cell radius using a linear transfer function relating coccolith length to radius, which is developed from culture experiments of different species of alkenone-producing coccolithophores under varying CO_2 levels. Several thousand coccoliths from each sample were imaged, providing a robust estimate of the population for the calibration of the transfer function (McClelland et al., 2016). In sediment samples, we treat each measured coccolith as an individual, and uncertainty in the population mean coccolith length is best described using the standard error of the mean. Uncertainty in the estimated cell radius includes this standard error of the population mean coccolith length, as well as uncertainty in the slope and intercept of the

transfer function. We estimate uncertainty in cell radius resulting from the transfer function (termed $\sigma_{transfer}$) using the "fitlm" and "predict" functions in MATLAB R2019a, which use to calculate 1σ uncertainties on the cell radius estimated from a given coccolith length. We then estimate the combined uncertainty by adding the standard error of the mean population coccolith length (σ_{lith}) and the transfer function uncertainty ($\sigma_{transfer}$) in quadrature.

$$\sigma_{radius} = \sqrt{\sigma_{transfer}^2 + \sigma_{lith}^2} \quad \text{[Eq. S16]}$$

Typical uncertainties for σ_{lith} are around 0.05 µm, and uncertainty from the transfer function is approximately 0.03 µm, summing to ~0.06 µm. Following the same approach for estimating the uncertainty resulting from an irradiance perturbation, a 0.06 µm uncertainty in cell radius equates to a CO_2 uncertainty of 0.48 µmol kg^{-1}, which at 28°C and 35 psu equals a pCO_2 uncertainty of ~17.5 µatm (1σ).

3.10. Uncertainty in model coefficients

Because errors in least-squares linear model parameters are correlated, we estimate uncertainty in the model coefficients through Monte Carlo re-fitting simulations (nRF = 1,000). First, we calculate the error (modeled minus measured ε_p for n = 99 observations, nObs) from the "best" model fit of the culture data, and create an empirical cumulative distribution function (eCDF) of these residuals. We resample this eCDF nObs x nRF times, providing an estimate of the error in modeled ε_p for each Monte Carlo simulation. We add these errors to the original culture ε_p data, and then re-fit the multiple linear regression with the original input predictor

values. This approach synthetically loads all the error in the Y value (ε_p), but because the residuals in Y are a function of the uncertainty in the predictor terms, resampling the residuals of the best fit is a reasonable way to estimate the uncertainty in the model terms. We fit nRF synthetic models that have been perturbed with the resampled residuals as described above.

When solving the model equation for CO_2, we use each of these synthetic error-propagated models and calculate CO_2 using the perturbed coefficients, but the measured down-core ε_p, cell size, and the prescribed irradiance. This returns nRF estimates of CO_2, and we take the mean and standard deviation of this distribution as the mean and 1σ uncertainty on the calculated CO_2. We term this uncertainty "σ_{MDL}" because it is solely from the uncertainty in the multiple linear model coefficients. We calculate σ^2_{MDL} for each sample in the Pleistocene compilation dataset using nRF of 10,000, estimating an average uncertainty of ± 1.9 μmol kg^{-1}, which at 28°C and 35 psu equates to ~70 μatm.

3.11. Model inversion for $CO_{2(aq)}$ and conversion to pCO_2

We can solve the ε_p linear regression equation for the CO_2 needed to match the measured ε_p, cell radius, and prescribed irradiance and growth rate. As mentioned above, growth rate is an insignificant contributor in the model and could be assumed constant, but we model growth rate as a function of irradiance, temperature, and modern phosphate and nitrate concentrations (see Chapter 3, Appendix B for details). Because the term is such a small contributor to ε_p, we ignore it in the error consideration here. The multilinear regression equation takes the form of:

$$\varepsilon_p = \beta_0 + \beta_1 \times CO_2 + \beta_2 \times CO_2^2 + \beta_3 \times PAR_Z + \beta_4 \times \mu + \beta_5 \times radius \quad \text{[Eq. S17]}$$

The analytical solution for CO_2 is:

$$CO_2 = -\frac{b \pm \sqrt{b^2 - 4f(a - \varepsilon_p + cPAR_Z + rR + d\,\mu)}}{2f} \quad \text{[Eq. S18]}$$

This equation has two roots. We ignore the larger roots in all solutions because the model is empirical and not mechanistic, and the larger CO_2 values would result in unnaturally high (~2500 µatm) pCO_2. Any solutions that yield complex numbers are discarded and only real solutions are kept. We estimate theoretical uncertainty in reconstructed CO_2 using Monte Carlo simulations (nMC = 1,000). For each sample, we generate nMC synthetic observations of each input variable (ε_p, cell radius, and PAR_Z) assuming a normal distribution around the mean and using the errors as described above. Growth rate is held constant because any reasonable µ variation has a minute effect on ε_p in the model (Fig. S5). We solve each of these Monte Carlo iterations using each of the re-fit models generated as described in Section 3.10 (above). As a result of the spread in the normal distributions, the lower tail can yield negative values. We therefore remove all solutions less than 2.5 µmol kg (~100 µatm at 28°C, even less at colder temperatures), then take the mean and standard deviation of the resulting values. Using the typical uncertainties as described above ($\sigma_{\varepsilon p}$ = 0.45 ‰, σ_{radius} = 0.06 µm, σ_{PAR} = 0.5 mol m^{-2} d^{-1}) and the model coefficient uncertainty combine for an average σ_{CO2} of ~2.5 µmol kg^{-1}. This error is large relative to the CO_2 signal in the Pleistocene and Pliocene, the implications which we discuss in detail in the main text (Section 5.4).

We can approximate the individual uncertainties by setting different variables constant and solving for the change in CO_2 associated with the uncertainties in each variable. This exercise with uncertainties of $\sigma_{\varepsilon p}$ = 0.45 ‰, σ_{radius} = 0.06 µm, and σ_{PAR} = 0.43 mol m^{-2} d^{-1} yield

CO_2 uncertainties of 0.94 µmol kg^{-1}, 0.48 µmol kg^{-1}, and 0.42 µmol kg^{-1}, respectively. The total uncertainty in CO_2 can be approximated by:

$$\sigma_{CO2,total} = \sqrt{\sigma^2_{CO2,MDL} + \sigma^2_{CO2,PAR} + \sigma^2_{CO2,radius} + \sigma^2_{CO2,\varepsilon p}} \quad \text{[Eq. S19]}$$

Solving this relationship for $\sigma_{CO2,total}$ of 2.5, we estimate the $\sigma_{CO2,MDL}$ to be 2.2µmol kg^{-1}. The majority of the error in estimating past CO_2 therefore comes from the model uncertainty. At 28°C and 35 psu, 2.5 µmol kg^{-1} CO_2 equates to ~95 µatm pCO_2.

Finally, we account for uncertainty in the conversion of $CO_{2(aq)}$ to pCO_2 using the uncertainty in $CO_{2(aq)}$ and uncertainty in K_0, Henry's Law coefficient:

$$\sigma^2_{pCO2} = \overline{pCO_2} \times \sqrt{\left(\frac{\sigma_{CO2}}{CO_2}\right)^2 + \left(\frac{\sigma_{K_0}}{K_0}\right)^2} \quad \text{[Eq. S20]}$$

Where the horizontal bar indicates the mean value. We estimate σ_{K_0} as described in Section 3.6. Equation S20 demonstrates that the fully propagated pCO_2 uncertainty is sample specific, but an average error of 2.5 µmol kg^{-1} and an average $CO_{2(aq)}$ of 6.5 µmol kg^{-1} at a warm tropical site with a relative K_0 error of ~5% (Section 3.6.) would result in pCO_2 errors of ~39%.

4. Supplementary discussion

4.1. Assemblage-estimated versus measured coccolith size

There is a generally good agreement between our measured and assemblage-estimated approach at ODP 807; the two show the same general trends (Fig. S7a). The measured values are

offset from the estimated values by ~0.5μm on average. This offset arises because the assemblage estimate applies the mean coccolith size of each species present, though most holotypes have a ~2μm range in length according to the species definition. However, because we are interested in the relative sensitivity, and the $\varepsilon_{p37:2}$ model is linear with respect to coccolith length, our assemblage estimates are reasonable enough to evaluate general trends.

4.2. Controls on coccolith size

We examine the relationship between coccolith size and SST and in-situ $CO_{2(aq)}$ concentrations through time at each location by regressing coccolith length against these variables at each location. Five sites (GeoB1016-3, ODP 999, ODP 806, and ODP 925), have a significant relationship ($p \leq 0.1$) between in-situ $CO_{2(aq)}$ and coccolith length, two with negative slopes and three with positive slopes. Four sites (GeoB1016-3, ODP 999, and ODP 807) have a significant relationship between SST and coccolith length, in this case half have positive slopes. We therefore find no universal directional relationship between coccolith length and either CO_2 or SST—the two environmental parameters on which we have the best constraints—through time.

One additional question we consider is whether modern oceanographic conditions explain the mean coccolith size of a timeseries. In other words, are there structural ecological explanations for the spatial variability in the mean coccolith size at the locations we examine. We take the BMW mean coccolith length of each site-specific time series (n = 11) and regress these mean values against the collective modern oceanographic conditions, including latitude, nutrient concentrations, temperature, irradiance, and mixed layer depth, and find no statistically significant relationship between mean site-specific coccolith size and modern environmental

variables. However, we do find a statistically significant ($p < 0.05$) correlation between the timeseries-mean coccolith length and the ratio of irradiance to Redfield-ratio adjusted nutrient concentrations. Many culture experiments demonstrate that higher irradiance leads to larger cell sizes (Aloisi, 2015; Perrin et al., 2016). However, changes in nutrient concentrations from replete to limiting can have opposite effects on cell size, depending on the nutrient that becomes limiting (Aloisi, 2015). However, by comparing in-situ nutrient concentrations and irradiance levels to the half-saturation constants needed to best model growth rate in modern alkenone-producing populations (Chapter 3), we find that growth in >90% of the locations examined here is light-limited. This suggests that the sites we are considering are, at least in modern conditions, nutrient replete, and variance in size reflect the modern irradiance range we observe, in agreement with culture data. This analysis, however, assumes that the modern irradiance and nutrient conditions generally reflect the past environmental conditions at each location. There are many reasons to think—including our own analysis examining variations in irradiance through the ε_p model—that this assumption is tenuous. We do not put too much emphasis on our finding of mean coccolith size, and instead leave it as an observation that could be examined in greater detail in the future.

4.3. Temporal variability in irradiance

Three independent parameters can induce changes in the irradiance at depth seen by the alkenone producing population: surface irradiance (PAR_{surf}), the attenuation coefficient in the upper water column (Kd_{490}), and the average depth of alkenone production within the mixed layer (Z_{prod}). Together, these three values combine to determine the irradiance at depth (PAR_Z):

$$PAR_z = PAR_{surface} \times e^{-(Kd_{490} \times Z)} \quad \text{[Eq. S21]}$$

The integrated daily irradiance at depth (in mol m^{-2} d^{-1}) is an input in the culture-based ε_p model (Chapter 2). In theory, variations in PAR_Z through time could arise from variations in any of the three parameters on the right side of this equation.

Incoming solar radiation at a given core location varies on orbital timescales (Hays et al., 1976). Using the solutions of Laskar et al. (2004) stored in Analyseries 2.0.8 (Paillard et al., 1996), we determined mean annual insolation over the last 800 kyr at several latitudes. Calculated mean annual insolation at the equator has a maximum amplitude of variability of about 1% over the last one million years; mean annual insolation varies the most—by up to roughly 10%—at the poles, but even less at mid-latitudes than at the equator. Given these modest variations, changes in total irradiance due to Earth's orbital configuration have a negligibly small effect on ε_p.

The seasonality of insolation also changes as a function of Earth's orbital configuration, and it is entirely possible that there are concomitant changes in alkenone production with irradiance. However, we do not have a way of resolving variations in the seasonality of alkenone production in our down-core sediment samples, which integrate hundreds to thousands of years. Other ways to change irradiance include increasing or decreasing cloud cover, changing the water column turbidity, or changing the depth of alkenone production. Because PAR_Z is exponential with Kd_{490} and depth, both of these have a larger influence than the surface irradiance. We do not, as of yet, have an independent method for evaluating changes in depth of production or the attenuation coefficient.

4.4. Growth rate

In the integrated irradiance ε_p model presented in Chapter 2, the coefficient on growth rate is very small compared to the range of the data (0.3‰ d^{-1}), meaning growth rate variations between 0 and 1.5 d^{-1} are essentially within the noise of the measurable alkenone ε_p signal. Given that the largest measured growth rate in the field is ~1 d^{-1} (c.f. Chapter 3), growth rate variations have a negligible impact when interpreting ε_p in our framework. In the diffusive model, growth rate has a much larger influence on ε_p. Using the relationships of Bidigare et al. (1997), which underpin the quantitative scaling of growth rate and ε_p in the diffusive model, we can calculate the sensitivity of $\varepsilon_{p37:2}$ to changes in μ. The slope of ε_p as a function of μ is always negative, but is CO_2-dependent, and gets steeper as CO_2 declines. In the diffusive model, growth rate therefore has a larger effect on ε_p at low CO_2 than high CO_2. At constant CO_2 levels of 10 µM and 5 µM—roughly the range captured by the samples in this compilation—the slope of the relationship between ε_p and μ steepens from -14‰/d^{-1} to -28‰/d^{-1}, respectively. This sensitivity to μ is nearly two orders of magnitude greater than that in the integrated irradiance multilinear regression model of Chapter 2, which was calibrated using culture experiments.

Explaining our $\varepsilon_{p37:2}$ and $[CO_{2(aq)}]$ observations with the diffusive model would require higher latitude sites to have higher growth algal growth rates, which would lower $\varepsilon_{p37:2}$ despite higher $[CO_{2(aq)}]$. Mean cell size at ODP 1207 and ODP 807 are roughly similar, with about a 0.2µm discrepancy at different time periods, which permits a ~0.5‰ difference in $\varepsilon_{p37:2}$. The mean $\varepsilon_{p37:2}$ at ODP 807 is 14.1‰ whereas the mean at ODP 1207 is 11.4‰. With similar cell sizes, but higher CO_2 concentrations at ODP 1207, the diffusive model would require higher growth rates at ODP 1207 to reduce $\varepsilon_{p37:2}$ by ~2.5‰. The compilation of field measurements of Noelaerhabdaceae growth rates in Chapter 3 found growth rates were generally higher at higher temperatures and irradiances (low-latitude locations), as irradiance and temperature are

fundamental controls on growth rate (Fielding, 2014). For example, Popp et al. (2006) found that increasing irradiance in natural populations by shifting incubations vertically in the water column led to much greater production rates of alkenones at station ALOHA in the subtropical Pacific.

In field incubations from the North Pacific, Popp et al. (2006a) estimated Noelaerhabdaceae growth rates declined from ~0.6 d^{-1} near the surface to ~0.1 d^{-1} at 70 m depth, over an irradiance change of ~250 μmol m^{-2} s^{-1}. At a CO_2 concentration of ~15 μM, a reasonable estimate for surface waters in the North Pacific (Reinfelder, 2011), using the culture-based relationship of Bidigare et al. (1997), this decline in growth rate from surface to deep would lead to an *increase* in $\varepsilon_{p37:2}$ of about 5‰. Even over a 30 m depth change, as irradiance drops by ~100 μmol m^{-2} s^{-1}, growth rate declines by 0.3 d^{-1}. Under the diffusive model, this should cause a ~3‰ increase in $\varepsilon_{p37:2}$ at depth. However, in the vast majority of field samples taken at multiple depths at one location, carbon isotope values at the surface are as depleted or more depleted than at depth, giving rise to larger $\varepsilon_{p37:2}$ near the surface, in contrast to the growth-rate dependent predictions from the diffusive model (Bidigare et al., 1997; Eek et al., 1999; Laws et al., 2001; Tolosa et al., 2008; Wolhowe et al., 2014). If we account for the fact that CO_2 typically increases with depth as respiration starts to overwhelm the CO_2 drawdown by photosynthesis, the elevated CO_2 at depth would lead to even higher $\varepsilon_{p37:2}$ values than our estimates above. While a few locations show increases of ~1-2‰ in $\varepsilon_{p37:2}$ with small depth changes of ~25m, we do not see the magnitude or direction of the response predicted by the diffusive model. It is more likely that, as the culture regression models predict, the higher irradiance at the surface is directly engendering greater isotope fractionation, and the decline in $\varepsilon_{p37:2}$ with depth is associated with reductions in irradiance.

4.5. Alternative explanations for a measured vs. modeled slope of 1

As discussed in detail in the main text, measured versus modeled $\varepsilon_{p37:2}$ slopes close to values of 1 suggest that irradiance is relatively stable through time. While this may not hold beyond the late Pleistocene, this helps identify whether CO_2 information can be extracted from alkenone ε_p, and provides a way of estimating a reasonable envelope of uncertainty on irradiance through time. We consider this a necessary precondition for attempting to reconstruct CO_2 from $\varepsilon_{p37:2}$, as a location that yields wildly variable irradiance in the Pleistocene cannot be assumed to remain constant with much confidence in the past.

One concern with our findings is that the two sites with slopes close to 1 (ODP 807 and ODP 668) have small sample sizes (n = 12 and n = 6, respectively). It could therefore be possible that the relationship arises from a small random sampling of a diverse $\varepsilon_{p37:2}$ distribution, such as that at ODP 999, and increasing the number of samples at ODP 807 or ODP 668 could flatten out the slope. The sample selection at ODP 807 also targeted glacial and interglacial maxima so as to minimize the likelihood of sampling high amplitude CO_2 variability. The record at ODP 999 used a more continuous sampling strategy, but covered an entire glacial cycle. While we cannot increase the number of samples, we can subsample records with more data to see if the positive slopes are an artifact of a small sample size.

We use a Monte Carlo resampling routine (n = 10,000) to randomly draw two populations from the ODP 999 record with sample sizes roughly equal to those of ODP 807 and ODP 668 (n = 12 and n = 6). We draw half of each Monte Carlo resampling from glacial periods, and half from interglacials, setting the pCO_2 cutoff defining the boundary as 240 μatm. In each synthetic population, we calculate the slope between measured ε_p and modeled ε_p, and find that a slope of 1 is effectively impossible with a sample size of 12, and a slope of 1 occurs in the 98th

percentile of the Monte-Carlo generated slopes with a sample size of 6, suggesting the findings at ODP 807 and ODP 668 are robust.

Furthermore, the ODP 806 samples (n = 3) of Zhang et al. (2017) also yield a positive slope close to 1, albeit with less statistical confidence ($p < 0.25$). We do not combine these data with our own at nearby ODP 807 because they measured coccolith length by a different method and report only the average coccolith size, limiting the estimated size to the arithmetic mean alone instead of the biomass-weighted mean, which requires knowledge of the distribution of coccolith size in a sample. The arithmetic mean will always be lower than the biomass-weighted mean, which explains why the modeled $\varepsilon_{p37:2}$ for *Zhang'17: 806* is higher than the modeled $\varepsilon_{p37:2}$ at ODP 807: the larger BMW-mean coccolith size in our record reduces $\varepsilon_{p37:2}$ more in the model result than the arithmetic mean used by Zhang et al. (2017).

Another possible explanation for the behavior at ODP 807 and ODP 668 compared to other oligotrophic sites is the age distribution of the sample sets. For example, if irradiance varies more on precessional timescales than eccentricity timescales, a shorter, high-resolution record may be subject to greater irradiance variability than a longer low-resolution record. ODP 807 represents one of these longer low-resolution records, as it covers the last 650 ka at an average sample spacing of ~60 kyr. ODP 668, on the other hand, only has coccolith size data for the last ~140 ka, limiting application of the $\varepsilon_{p37:2}$ model to this time period. Neither differences in sampling resolution nor the age range of each record explain differences between site behavior.

4.6. Influence of cell radius on alkenone ε_p

To emphasize the influence of cell radius on alkenone carbon isotope fractionation, we hold the in-situ CO_2 constant at 6.5 μmol kg^{-1}, the mean of the entire dataset, and model $\varepsilon_{p37:2}$ as

a function of the reconstructed cell radius alone (Fig. S9). Nine sites have sufficient coccolith length timeseries data to estimate statistics on the relationship between modeled and measured $\varepsilon_{p37:2}$; of these, three are statistically significant ($p<0.1$). In contrast to the modeled versus measured relationships account for CO_2 alone (Fig. 5), we find both positive and negative slopes, suggesting in some locations that CO_2 and coccolith length are positively correlated and in others they are anti-correlated; which is indeed the case (Fig. S4). Importantly, the slopes of these relationships vary between -0.3 and +2, demonstrating that temporal variance in cell size is a more significant contributor to the measured $\varepsilon_{p37:2}$ variation than CO_2 in the Late Pleistocene.

The culture-based model identifies the sensitivity of $\varepsilon_{p37:2}$ to cell radius as -3.9‰ μm^{-1}, a high slope that suggests changes in cell radius can have cause large variations in $\varepsilon_{p37:2}$ through time (Chapter 3). In our Pleistocene compilation, the average coccolith length amplitude (maximum length minus minimum length in each record) is 1.2 μm, which equates to a cell radius value of 0.7 μm. In the absence of compensatory changes in CO_2 and/or irradiance, a 0.7 μm cell radius fluctuation would produce a 2.7‰ deviation in $\varepsilon_{p37:2}$, large enough to explain most of the variation in $\varepsilon_{p37:2}$ in the dataset, and mask changes in $\varepsilon_{p37:2}$ due to CO_2.

The magnitude of the impact of cell radius is evident from the range of environmental and physiological parameters in each site-specific timeseries (Fig. 4, main text) and the known slope sensitivities of $\varepsilon_{p37:2}$ to each parameter (Fig. S5), but the site-by-site regression allows for a more pointed analysis of the roles of cell size and CO_2 in explaining the measured $\varepsilon_{p37:2}$ data. Accounting for cell radius has often been considered a "correction" to the alkenone ε_p records, but our model and the results presented here suggest it is a more important control on $\varepsilon_{p37:2}$ than CO_2, and $\varepsilon_{p37:2}$ records cannot be accurately interpreted without considering cell size.

5. References

Aloisi G. (2015) Covariation of metabolic rates and cell size in coccolithophores. *Biogeosciences* **12**, 4665–4692.

Badger M. P. S., Foster G. L., Chalk T. B., Gibbs S. J., Badger M. P. S., Pancost R. D., Schmidt D. N., Sexton P. F., Mackensen A., Bown P. R. and Pälike H. (2019) Insensitivity of alkenone carbon isotopes to atmospheric CO2 at low to moderate CO2 levels. *Clim. Past* **15**, 539–554.

Bé A. W. H., Spero H. J. and Anderson O. R. (1982) Effects of symbiont elimination and reinfection on the life processes of the planktonic foraminifer Globigerinoides sacculifer. *Mar. Biol.* **70**, 73–86.

Beaufort L. and Dollfus D. (2004) Automatic recognition of coccoliths by dynamical neural networks. *Mar. Micropaleontol.* **51**, 57–73.

Bendif E. M., Nevado B., Wong E. L. Y., Hagino K., Probert I., Young J. R., Rickaby R. E. M. and Filatov D. A. (2019) Repeated species radiations in the recent evolution of the key marine phytoplankton lineage Gephyrocapsa. *Nat. Commun.* **10**, 4234.

Bereiter B., Eggleston S., Schmitt J., Nehrbass-Ahles C., Stocker T. F., Fischer H., Kipfstuhl S. and Chappellaz J. (2015) Revision of the EPICA Dome C CO2 record from 800 to 600-kyr before present. *Geophys. Res. Lett.* **42**, 542–549.

Berger W. H., Bickert T., Schmidt H. and Wefer G. (1993) Quaternary Oxygen Isotope Record of Pelagic Foraminifers: Site 806, Ontong Java Plateau. In *Proceedings of the Ocean Drilling Program, 130 Scientific Results* (eds. W. H. Berger, L. W. Kroenke, and L. A. Mayer). Ocean Drilling Program. pp. 381–395.

Bidigare R. R., Fluegge A., Freeman K. H., Hanson K. L., Hayes J. M., Hollander D., Jasper J. P., King L. L., Laws E. A., Milder J., Millero F. J., Pancost R., Popp B. N., Steinberg P. A. and Wakeham S. G. (1997) Consistent fractionation of 13C in nature and in the laboratory: Growth-rate effects in some haptophyte algae. *Global Biogeochem. Cycles* **11**, 279–292.

Chiyonobu S., Sato T., Narikiyo R. and Yamasaki M. (2006) Floral changes in calcareous nannofossils and their paleoceanographic significance in the equatorial Pacific Ocean during the last 500 000 years. *Isl. Arc* **15**, 476–482.

CLIMAP Project Members (1976) The Surface of the Ice-Age Earth. *Science (80-.).* **191**, 1131–1137.

Constans R. E. and Parker M. E. (1986) Calcareous Nannofossil Biostratigraphy and Paleoclimatic Indices for the Late Quaternary, Deep Sea Drilling Project Leg 96, Gulf of Mexico. In *Initial Reports of the Deep Sea Drilling Project, 96* U.S. Government Printing Office. pp. 601–630. Available at: http://deepseadrilling.org/96/volume/dsdp96_32.pdf.

Dyez K. A., Hönisch B. and Schmidt G. A. (2018) Early Pleistocene Obliquity-Scale pCO2 Variability at ~1.5 Million Years Ago. *Paleoceanogr. Paleoclimatology* **33**, 1270–1291.

Eek M. K., Whiticar M. J., Bishop J. K. B. and Wong C. S. (1999) Influence of nutrients on carbon isotope fractionation by natural populations of Prymnesiophyte algae in NE Pacific. *Deep. Res. Part II Top. Stud. Oceanogr.* **46**, 2863–2876.

Ezard T. H. G., Edgar K. M. and Hull P. M. (2015) Environmental and biological controls on size-specific 13C and 18O in recent planktonic foraminifera. *Paleoceanography*, 1–23.

Fielding S. (2014) Emiliania huxleyi population growth rate response to light and temperature: a synthesis. *Aquat. Microb. Ecol.* **73**, 163–170.

Freeman K. H. and Hayes J. M. (1992) Fractionation of carbon isotopes by phytoplankton and estimates of ancient CO2 levels. *Global Biogeochem. Cycles* **6**, 185–198.

Freeman K. H. and Pagani M. (2005) Alkenone-Based Estimates of Past CO2 Levels: A Consideration of Their Utility Based on an Analysis of Uncertainties. *A Hist. Atmos. CO2 Its Eff. Plants, Anim. Ecosyst.*, 35–61.

Hayes J. . (1993) Factors controlling 13C contents of sedimentary organic compounds: Principles and evidence. *Mar. Geol.* **113**, 111–125.

Hays J. D., Imbrie J. and Shackleton N. J. (1976) Variations in the Earth's Orbit: Pacemaker of the Ice Ages. *Science (80-.).* **194**, 1121–1132.

Henriksson A. S. (2000) Coccolithophore response to oceanographic changes in the equatorial Atlantic during the last 200,000 years. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **156**, 161–173.

Hönisch B. and Hemming N. G. (2005) Surface ocean pH response to variations in pCO2 through two full glacial cycles. *Earth Planet. Sci. Lett.* **236**, 305–314.

Hönisch B., Hemming N. G., Archer D., Siddall M., McManus J. F., Honisch B., Hemming N. G., Archer D., Siddall M. and McManus J. F. (2009) Atmospheric Carbon Dioxide Concentration Across the Mid-Pleistocene Transition. *Science (80-.).* **324**, 1551–1554.

Keeling C. D. (1960) The Concentration and Isotopic Abundances of Carbon Dioxide in the Atmosphere. *Tellus* **12**, 200–203.

Key R. M., Olsen A., van Heuven S., Lauvset S. K., Velo A., Lin X., Schirnick C., Kozyr A., Tanhua T. and Hoppema M. (2015) *Global Ocean Data Analysis Project, Version 2 (GLODAPv2).*,

Kinkel H., Baumann K. H. and Cepek M. (2000) Coccolithophores in the equatorial Atlantic

Ocean: Response to seasonal and Late Quaternary surface water variability. *Mar. Micropaleontol.* **39**, 87–112.

Laskar J., Robutel P., Joutel F., Gastineau M., Correia A. C. M. and Levrard B. (2004) A long-term numerical solution for the insolation quantities of the Earth. *Astron. Astrophys.* **428**, 261–285.

Lauvset S. K., Key R. M., Olsen A., Van Heuven S., Velo A., Lin X., Schirnick C., Kozyr A., Tanhua T., Hoppema M., Jutterström S., Steinfeldt R., Jeansson E., Ishii M., Perez F. F., Suzuki T. and Watelet S. (2016) A new global interior ocean mapped climatology: The 1° × 1° GLODAP version 2. *Earth Syst. Sci. Data* **8**, 325–340.

Laws E. A., Popp B. N., Bidigare R. R., Riebesell U., Burkhardt S. and Wakeham S. G. (2001) Controls on the molecular distribution and carbon isotopic composition of alkenones in certain haptophyte algae. *Geochemistry, Geophys. Geosystems* **2**.

Lea D. W., Pak D. K. and Spero H. J. (2000) Climate Impact of Late Quaternary Equatorial Pacific Sea Surface Temperature Variations. *Science (80-.).* **289**, 1719–1724.

Lisiecki L. E. and Raymo M. E. (2005) A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. *Paleoceanography* **20**.

Manivit H. (1989) Calcareous Nannofossil Biostratigraphy of Leg 108 Sediments. In *Proceedings of the Ocean Drilling Program, 108 Scientific Results* Ocean Drilling Program. pp. 35–69.

Martinson D. G., Pisias N. G., Hays J. D., Imbrie J., Moore T. C. and Shackleton N. J. (1987) Age Dating and the Orbital Theory of the Ice Ages: Development of a High-Resolution 0 to 300,000-Year Chronostratigraphy. *Quat. Res.* **27**, 1–29.

McClelland H. L. O., Barbarin N., Beaufort L., Hermoso M., Ferretti P., Greaves M. and Rickaby R. E. M. (2016) Calcification response of a key phytoplankton family to millennial-scale environmental change. *Sci. Rep.* **6**, 1–11.

Medina-Elizalde M. and Lea D. W. (2005) The Mid-Pleistocene Transition in the Tropical Pacific. *Science (80-.).* **310**, 1009–1012.

Mook W. G., Bommerson J. C. and Staverman W. H. (1974) Carbon isotope fractionation between dissolved bicarbonate and gaseous carbon dioxide. *Earth Planet. Sci. Lett.* **22**, 169–176.

Müller P. J., Čepek M., Ruhland G. and Schneider R. R. (1997) Alkenone and coccolithophorid species changes in late Quaternary sediments from the Walvis Ridge: Implications for the alkenone paleotemperature method. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **135**, 71–96.

Müller P. J., Kirst G., Ruhland G., von Storch I. and Rosell-Melé A. (1998) Calibration of the

alkenone paleotemperature index U37K′ based on core-tops from the eastern South Atlantic and the global ocean (60°N-60°S). *Geochim. Cosmochim. Acta* **62**, 1757–1772.

Olsen A., Key R. M., Van Heuven S., Lauvset S. K., Velo A., Lin X., Schirnick C., Kozyr A., Tanhua T., Hoppema M., Jutterström S., Steinfeldt R., Jeansson E., Ishii M., Pérez F. F. and Suzuki T. (2016) The global ocean data analysis project version 2 (GLODAPv2) - An internally consistent data product for the world ocean. *Earth Syst. Sci. Data* **8**, 297–323.

Pagani M., Arthur M. A. and Freeman K. H. (1999) Miocene evolution of atmospheric carbon dioxide. *Paleoceanography* **14**, 273–292.

Pagani M., Liu Z., LaRiviere J. and Ravelo A. C. (2010) High Earth-system climate sensitivity determined from Pliocene carbon dioxide concentrations - Supplement. *Nat. Geosci.* **3**, 27–30.

Paillard D., Labeyrie L. and Yiou P. (1996) Macintosh Program performs time-series analysis. *Eos, Trans. Am. Geophys. Union*.

Palmer M. R., Brummer G. J. A., Cooper M. J., Elderfield H., Greaves M. J., Reichart G. J., Schouten S. and Yu J. (2010) Multi-proxy reconstruction of surface water pCO2 in the northern Arabian Sea since 29ka. *Earth Planet. Sci. Lett.* **295**, 49–57.

Perrin L., Probert I., Langer G. and Aloisi G. (2016) Growth of the coccolithophore Emiliania huxleyi in light- and nutrient-limited batch reactors: relevance for the BIOSOPE deep ecological niche of coccolithophores. *Biogeosciences* **13**, 5983–6001.

Polissar P. J. and D'Andrea W. J. (2014) Uncertainty in paleohydrologic reconstructions from molecular δD values. *Geochim. Cosmochim. Acta* **129**, 146–156.

Popp B. N., Bidigare R. R., Deschenes B., Laws E. A., Prahl F. G., Tanimoto J. K. and Wallsgrove R. J. (2006a) A new method for estimating growth rates of alkenone-producing haptophytes. *Limnol. Oceanogr. Methods* **4**, 114–129.

Popp B. N., Kenig F., Wakeham S. G., Laws E. A. and Bidigare R. R. (1998a) Does growth rate affect ketone unsaturation and intracellular carbon isotopic variability in Emiliania huxleyi ? *Paleoceanography* **13**, 35–41.

Popp B. N., Laws E. A., Bidigare R. R., Dore J. E., Hanson K. L. and Wakeham S. G. (1998b) Effect of phytoplankton cell geometry on carbon isotopic fractionation. *Geochim. Cosmochim. Acta* **62**, 69–77.

Popp B. N., Prahl F. G., Wallsgrove R. J. and Tanimoto J. (2006b) Seasonal patterns of alkenone production in the subtropical oligotrophic North Pacific. *Paleoceanography* **21**.

Prahl F. G., Wolfe G. V. and Sparrow M. A. (2003) Physiological impacts on alkenone paleothermometry. *Paleoceanography* **18**, 1–7.

Prentice M. L., Freiz J. K., Simonds G. G. and Matthews R. K. (1993) Neogene Trends in Planktonic Foraminifer 18O from Site 807: Implications for Global Ice Volume and Western Equatorial Pacific Sea-Surface Temperatures. In *Proceedings of the Ocean Drilling Program, 130 Scientific Results* Ocean Drilling Program. pp. 281–305.

Rabinowitz H. S., Polissar P. J. and Savage H. M. (2017) Reaction kinetics of alkenone and n-alkane thermal alteration at seismic timescales. *Geochemistry, Geophys. Geosystems* **18**, 204–219.

Rau G. H., Riebesell U. and Wolf-Gladrow D. (1996) A model of photosynthetic 13C fractionation by marine phytoplankton based on diffusive molecular CO2 uptake. *Mar. Ecol. Prog. Ser.* **133**, 275–285.

Reinfelder J. R. (2011) Carbon Concentrating Mechanisms in Eukaryotic Marine Phytoplankton. *Ann. Rev. Mar. Sci.* **3**, 291–315.

Riebesell U., Revill A. T., Holdsworth D. G. and Volkman J. K. (2000) The effects of varying CO2 concentration on lipid composition and carbon isotope fractionation in Emiliania huxleyi. *Geochim. Cosmochim. Acta* **64**, 4179–4192.

Romanek C. S., Grossman E. L. and Morse J. W. (1992) Carbon isotopic fractionation in synthetic aragonite and calcite: Effects of temperature and precipitation rate. *Geochim. Cosmochim. Acta* **56**, 419–430.

Seki O., Foster G. L., Schmidt D. N., Mackensen A., Kawamura K. and Pancost R. D. (2010) Alkenone and boron-based Pliocene pCO2 records. *Earth Planet. Sci. Lett.* **292**, 201–211.

Shipboard Scientific Party (2002) Site 1207. In *Proceedings of the Ocean Drilling Program, 198 Initial Reports* (eds. T. J. Bralower, I. Premoli Silva, and M. J. Malone). Ocean Drilling Program. p. Ch. 3, 1-140.

Shipboard Scientific Party (1991) Site 807. In *Proceedings of the Ocean Drilling Program, 130 Initial Reports* (eds. L. W. Kroenke, W. H. Berger, and T. R. Janecek). Ocean Drilling Program, College Station, TX. pp. 369–493.

Spero H. J., Bijma J., Lea D. W. and Bemis B. E. (1997) Effect of seawater carbonate concentration on foraminiferal carbon and oxygen isotopes. *Nature* **390**, 497–500.

Spero H. J., Bijma J., Lea D. W. and Russell A. D. (1999) Deconvolving Glacial Ocean Carbonate Chemistry from the Planktonic Foraminifera Carbon Isotope Record. *Reconstr. Ocean Hist.*, 329–342.

Spero H. J., Mielke K. M., Kalve E. M., Lea D. W. and Pak D. K. (2003) Multispecies approach to reconstructing eastern equatorial Pacific thermocline hydrography during the past 360 kyr. *Paleoceanography* **18**, n/a-n/a.

Takahashi T., Sutherland S. C., Chipman D. W., Goddard J. G., Ho C., Newberger T., Sweeney C. and Munro D. R. (2014) Climatological distributions of pH, pCO2, total CO2, alkalinity, and CaCO3 saturation in the global surface ocean, and temporal changes at selected locations. *Mar. Chem.* **164**, 95–125.

Tolosa I., Miquel J.-C., Gasser B., Raimbault P., Azouzi L. and Claustre H. (2008) Ecology and biogeochemistry of contrasting trophic environments in the South East Pacific by carbon isotope ratios on lipid biomarkers. *Biogeosciences Discuss.* **4**, 4653–4696.

Wara M. W., Ravelo A. C. and Delaney M. L. (2005) Permanent El Nino-Like Conditions During the Pliocene Warm Period. *Science (80-.).* **309**, 758–761.

Wefer G. and Berger W. H. (1991) Isotope paleontology: growth and composition of extant calcareous species. *Mar. Geol.* **100**, 207–248.

Weiss R. F. (1974) Carbon dioxide in water and seawater: the solubility of a non-ideal gas. *Mar. Chem.* **2**, 203–215.

Wilkens R. H., Westerhold T., Drury A. J., Lyle M. W., Gorgas T. and Tian J. (2017) Revisiting the Ceara Rise, equatorial Atlantic Ocean: isotope stratigraphy of ODP Leg 154 from 0 to 5 Ma. *Clim. Past* **13**, 779–793.

Wolhowe M. D., Prahl F. G., White A. E., Popp B. N. and Rosas-Navarro A. (2014) A biomarker perspective on coccolithophorid growth and export in a stratified sea. *Prog. Oceanogr.* **122**, 65–76.

Zhang J., Wang P., Li Q., Cheng X., Jin H. and Zhang S. (2007) Western equatorial Pacific productivity and carbonate dissolution over the last 550 kyr: Foraminiferal and nannofossil evidence from ODP Hole 807A. *Mar. Micropaleontol.* **64**, 121–140.

Zhang Y. G., Pagani M., Henderiks J. and Ren H. (2017) A long history of equatorial deep-water upwelling in the Pacific Ocean. *Earth Planet. Sci. Lett.* **467**, 1–9.

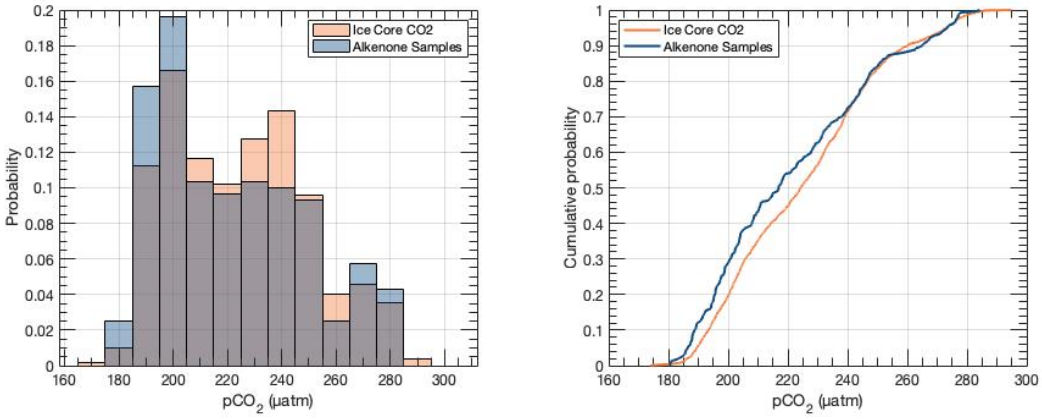

Figure S1. (Left) Histogram of ice core CO_2 data resampled to the median sampling density and histogram of atmospheric CO_2 at the ages of alkenone ε_p samples, both normalized the total number of samples in each group. (Right) Empirical cumulative distribution functions of the data plotted in the left panel, showing the median of the two are within ~10ppm. The alkenone $\varepsilon_{p37:2}$ dataset closely captures the range and distribution of atmospheric CO_2 over the last 800 ka, which has a slight bias towards glacial CO_2 levels.

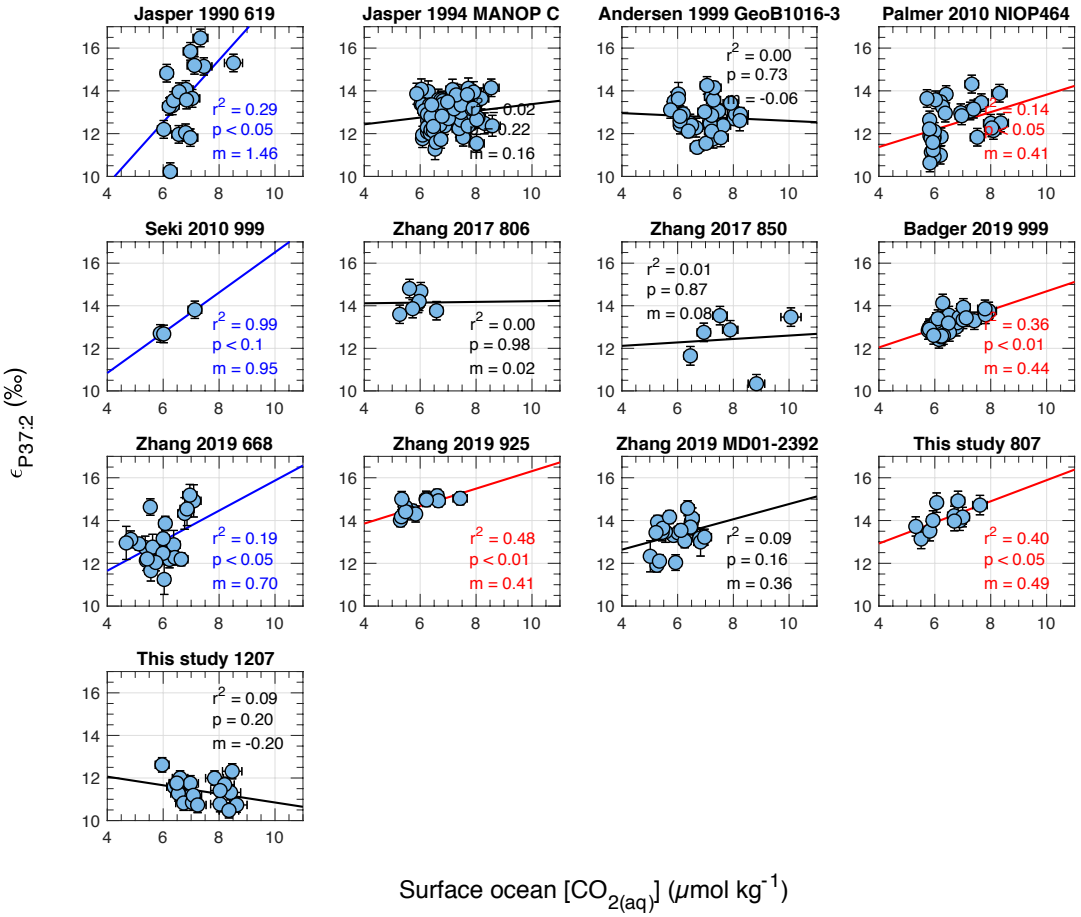

Figure S2. Correlation between measured $\varepsilon_{p37:2}$ and in-situ aqueous CO_2 concentrations for all sites considered in this compilation. Our new data (ODP 807 and ODP 1207) are shown the lower two panels. Text and regression lines are red where linear relationships are statistically significant at a threshold of $p<0.1$, and in blue where regressions are significant and the regression slope ± 1 standard error overlaps with a slope of 1. All statistically significant correlations are positive; no site shows a statistically significant decrease in $\varepsilon_{p37:2}$ with increasing CO_2.

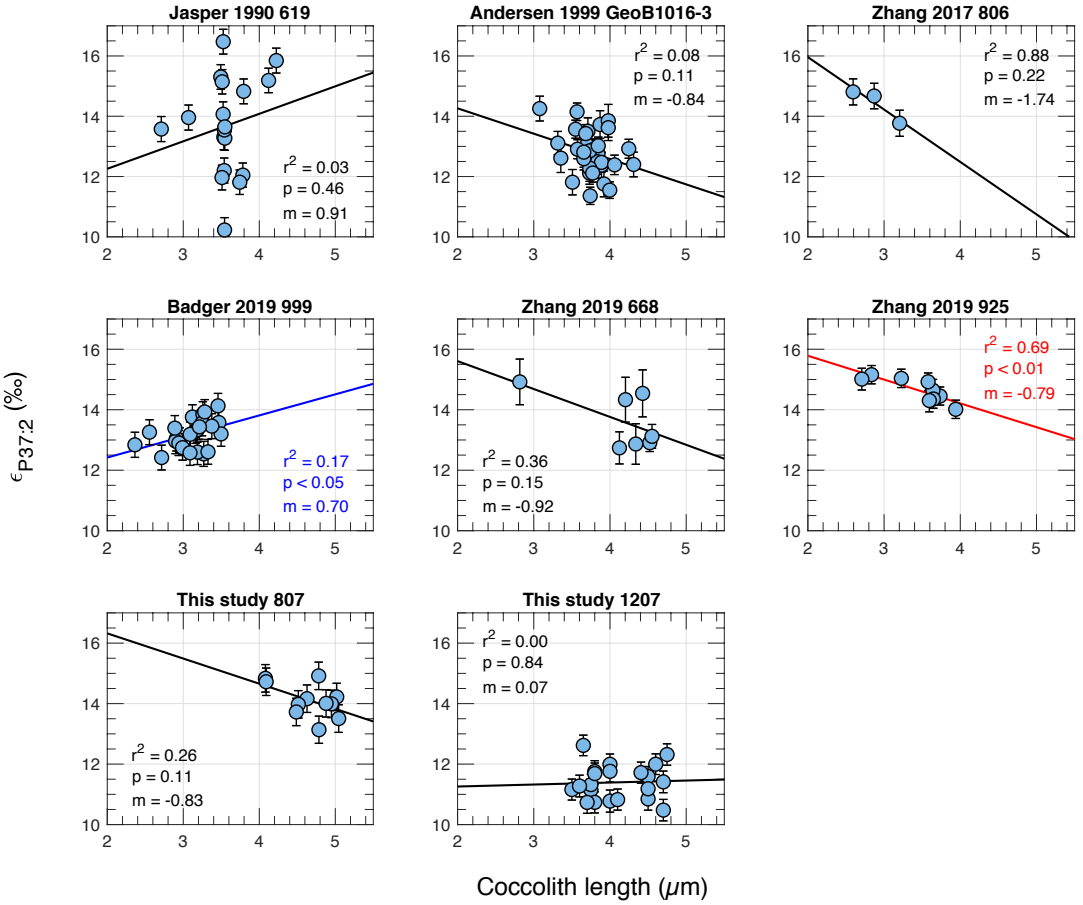

Figure S3. Correlation between measured $\varepsilon_{p37:2}$ and measured or estimated coccolith length. Coccolith length was measured by microscopy at ODP 807, ODP 1207, and ODP 999. Coccolith length was estimated from assemblage data at ODP 668, GeoB1016-3, and DSDP 619. As in Fig. S2, red text and regression lines indicate statistical significance at $p<0.1$; blue lines and text indicate the slope is within 1 SE of one. As opposed to the relationship between $\varepsilon_{p37:2}$ and CO_2, several sites show increases in $\varepsilon_{p37:2}$ with coccolith length while others show decreases in $\varepsilon_{p37:2}$ with coccolith length.

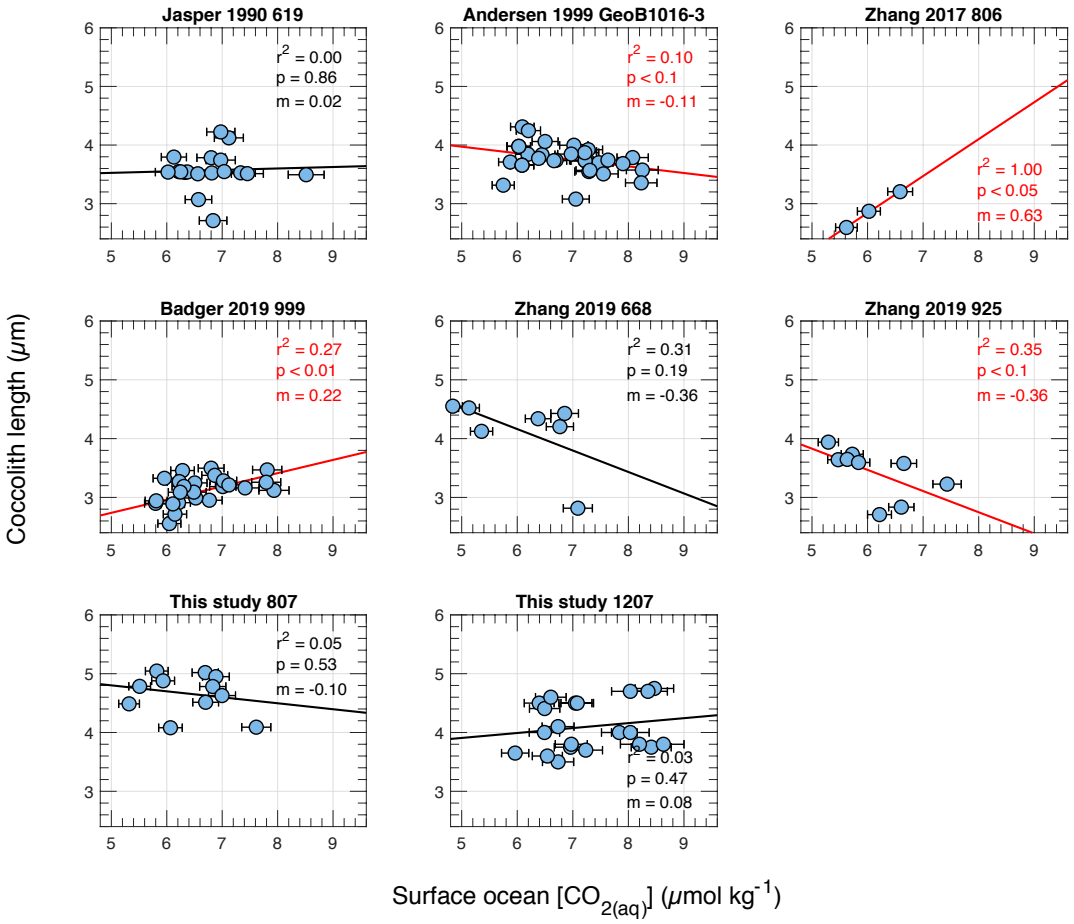

Figure S4. Site-by-site relationships between coccolith length and in-situ CO_2. As in Figures S2 and S3, red text and regression lines are those where the correlation is significant at the p<0.1 level. There are mixed correlations between coccolith length and CO_2, though the sites with the worst modeled versus measured relationships have positive or no correlation between coccolith length and CO_2 (compare to Fig. 7 in the main text).

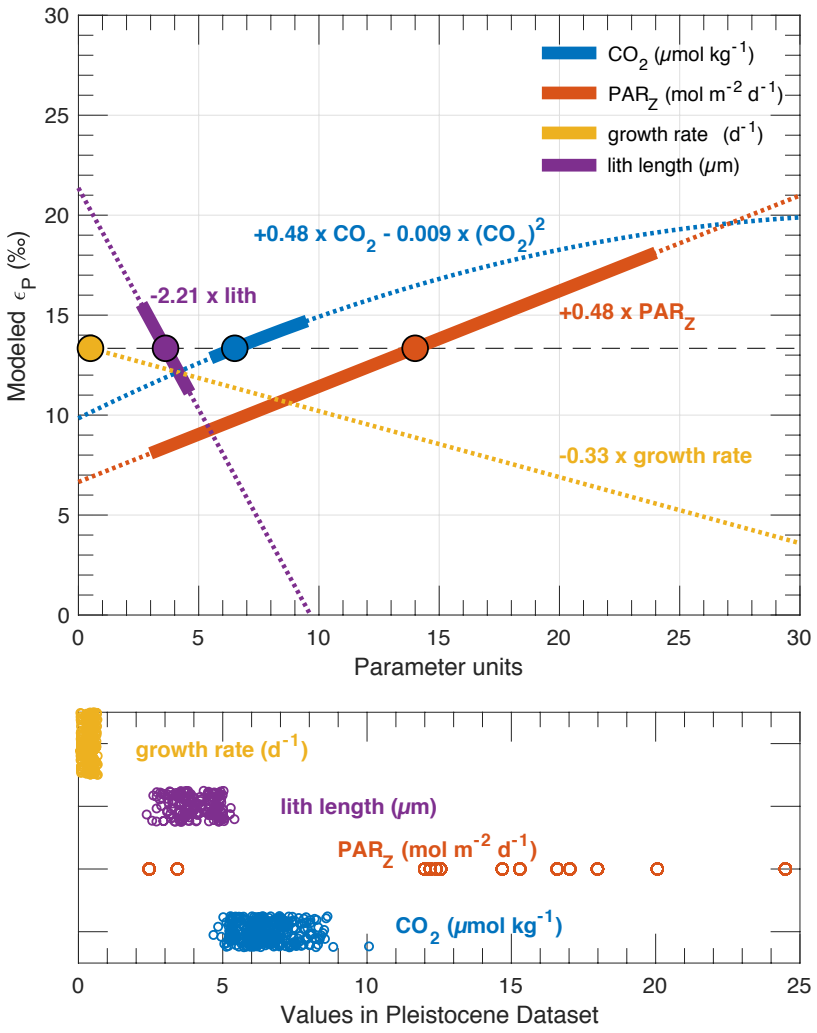

Figure S5. Effect plot for linear model used to calculate ε_p. Each line in the top panel displays the effect of changing the parameter (CO_2, PAR_Z, growth rate, or coccolith length) by one unit, with the units shown in the legend. The largest influence on ε_p is cell radius, here shown as coccolith size because they are linearly related. Irradiance and CO_2 have similar slopes, but the sensitivity to CO_2 declines as CO_2 increases above ~15 µmol kg^{-1}. Initial points are set at the mean of each value in the dataset and the mean ε_p (~13‰) to show the influence of the range in each parameter on ε_p. Of course, the effect plot does not include interactions; it is possible to reach ε_p values above ~20‰ through the combination of smaller cells, higher irradiance, and higher CO_2. The bottom panel shows the range of the measured values in the dataset (growth rate is modeled), with data slightly offset vertically for visual clarity. Growth rate has a negligibly small effect on ε_p and would have to increase to ~2 d^{-1} to have a measurable influence on ε_p, a value larger than any measured in the modern ocean (cf. Chapter 3).

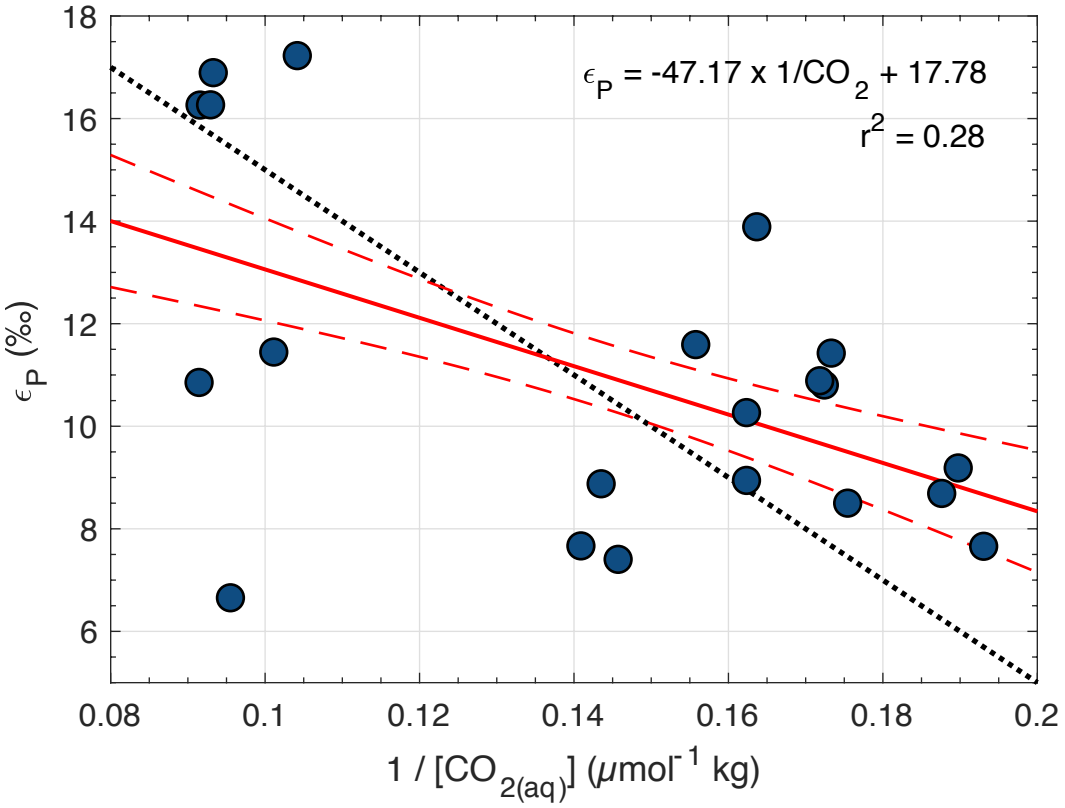

Figure S6. Relationship between ε_p and $1/CO_2$ from Noelaerhabdaceae cultures with CO_2 concentrations similar to those in the Pleistocene (approximately 5 – 11 µmol kg^{-1}). Data were taken from Chapter 2. The culture data show a lower slope and intercept than the diffusive model as it has been applied to sediments, which has an intercept of 25 and an average slope (the *b* value) of 80 to 120. The dotted black line is $\varepsilon_p = 25 - 100/[CO_{2(aq)}]$, showing the relationship between ε_p and $1/CO_2$ for a *b* value of 100.

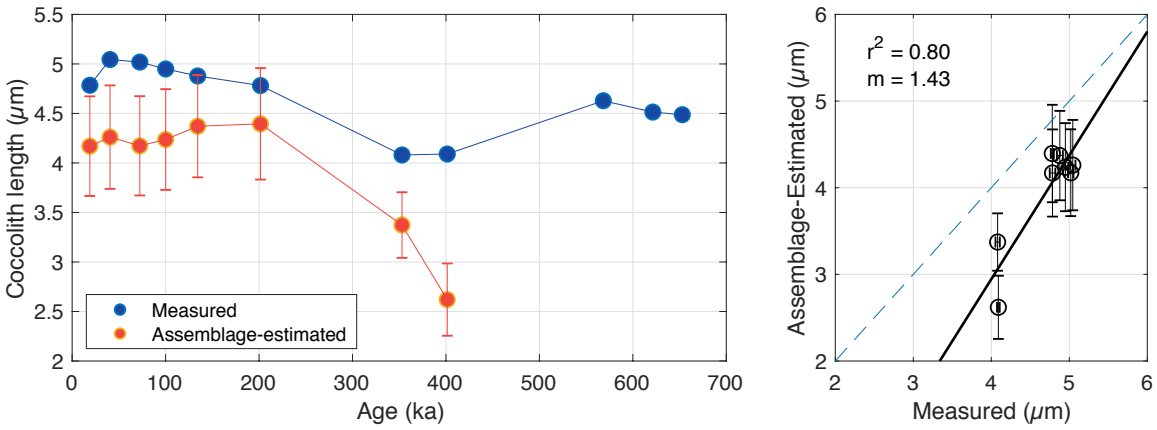

Figure S7. Measured vs. assemblage-estimated coccolith length at ODP 807. Coccolith length was estimated from the relative abundance of different alkenone-producing coccoliths reported by Chiyonobu et al. (2006). The two show generally good agreement, overlapping within uncertainty for the majority of the timeseries.

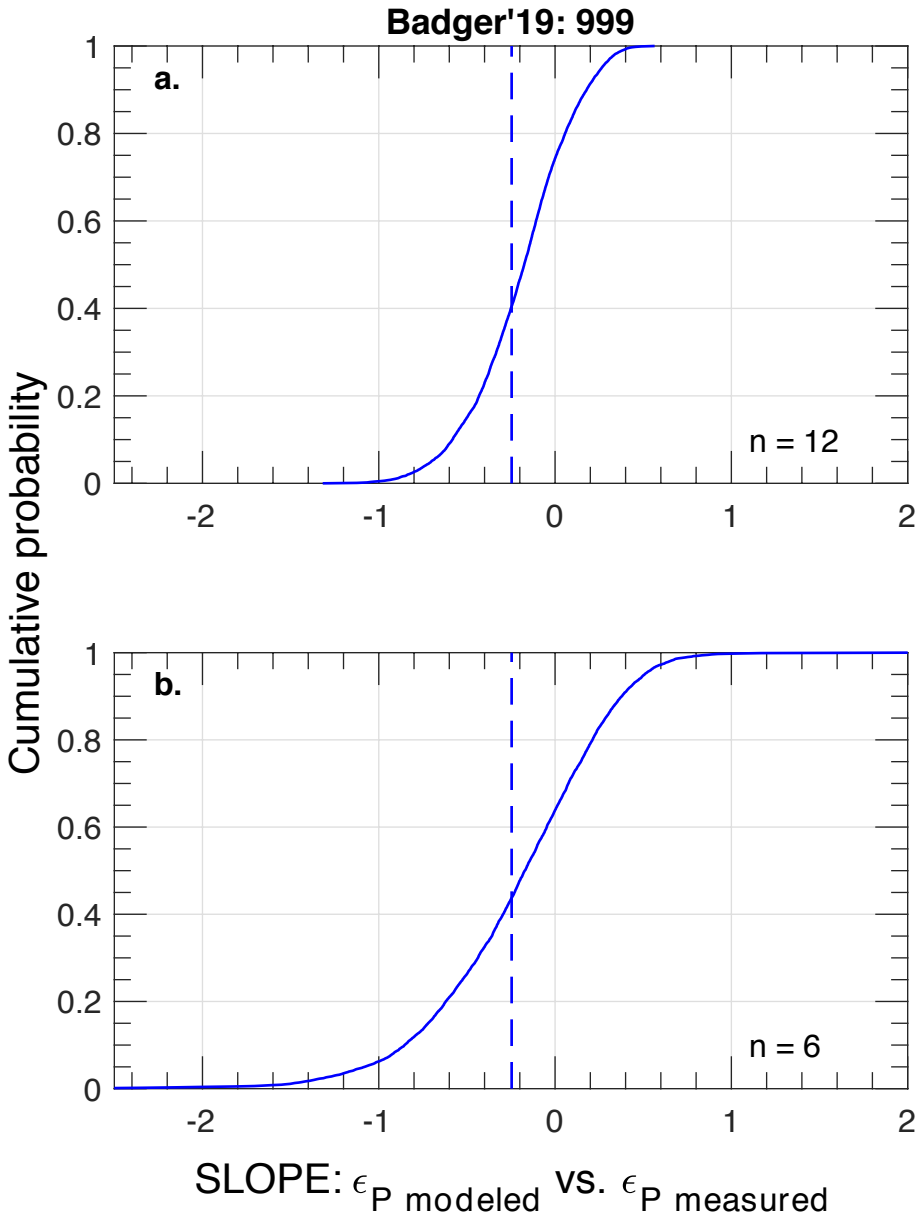

Figure S8. Results from Monte Carlo resampling of larger datasets to smaller sample sizes to determine whether the ε_p modeled versus ε_p measured slopes at ODP 807 and ODP 668 arise from their small sample sizes. The Badger et al. (2019) record resampled to (a) 12 samples and (b) 6 samples. In the resampling, samples are chosen so half of the samples come from glacial periods (< 240 µatm atmospheric pCO_2) and half from non-glacial periods (> 240 µatm atmospheric pCO_2). The mean slopes, the same as those in Fig. 6 in the main text, are shown as dashed lines on each plot. With a sample size of 12, it is virtually impossible to resample the ODP 999 record to achieve a slope greater than 0.8. With a sample size of 6, the probability of generating a positive slope is much higher, but a slope of 1 occurs in the >95th percentiles of each record, suggesting the observed slope at ODP 668 is unlikely an artifact of the small sample size.

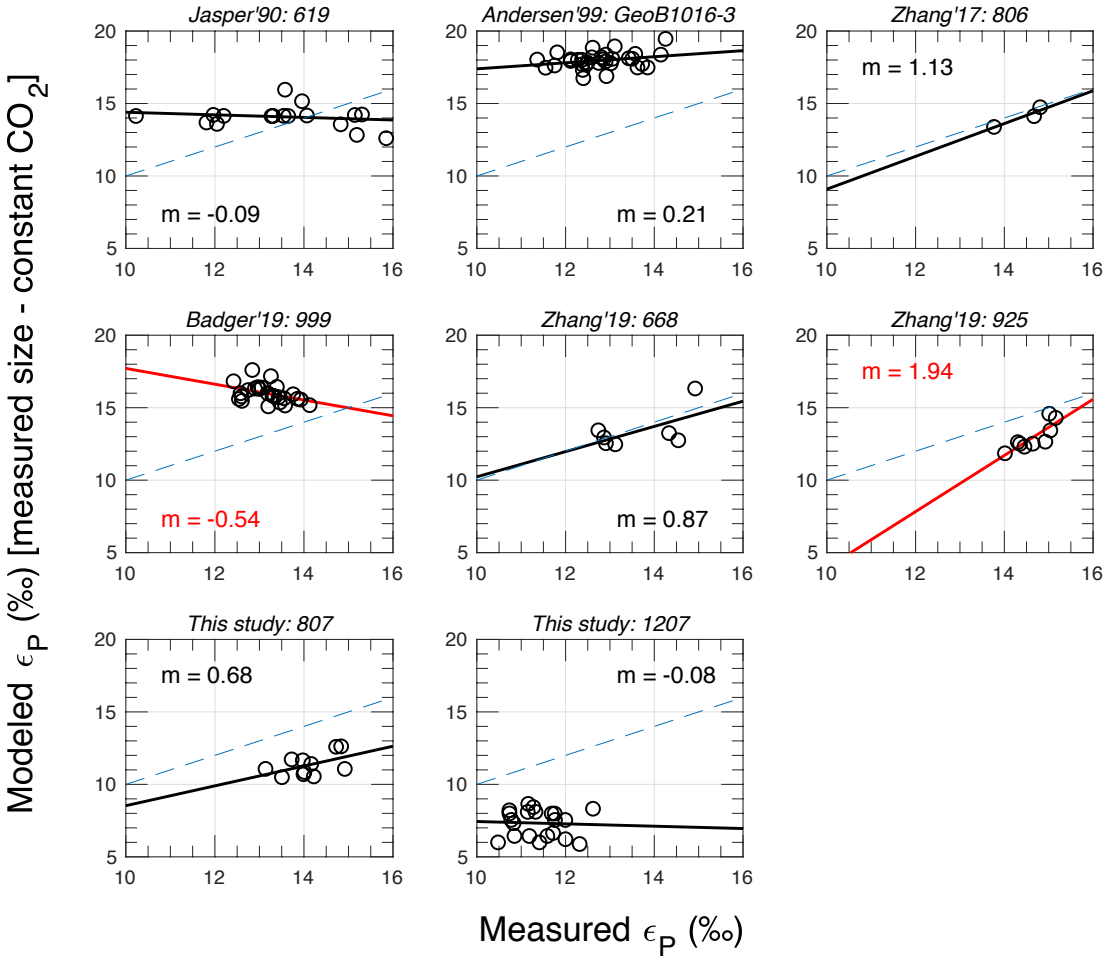

Figure S9. Effect of modeling $\varepsilon_{p37:2}$ while holding CO_2 values constant and varying only cell radius using the measured values. Cell size changes explain much of the variance in measured (and modeled) ε_p.

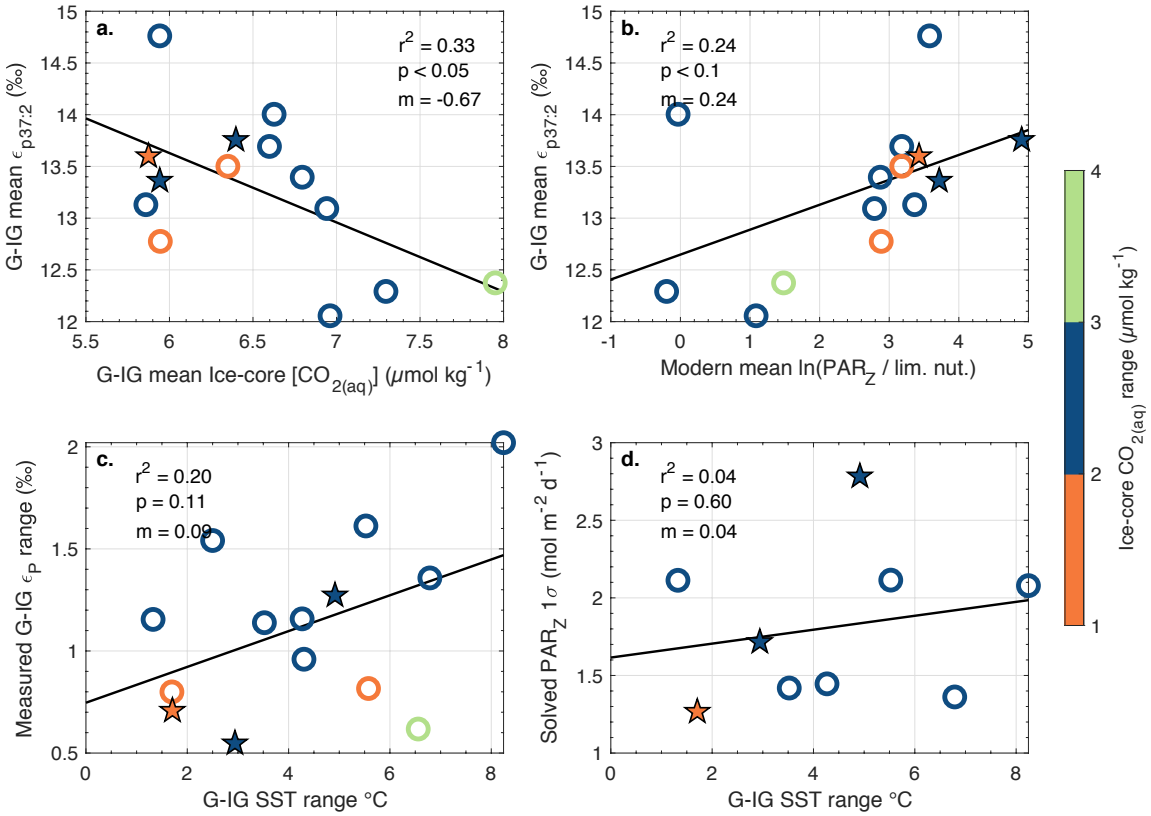

Figure S10. Same as Figure 8 in the main text, except alkenone $\delta^{13}C$ is held constant at the mean of the entire dataset in all samples (-24.2‰). The objective of this analysis is to identify where temperature has a strong influence on $\varepsilon_{p37:2}$ and our inferences. $\varepsilon_{p37:2}$ and temperature are intimately linked because an increase in temperature lessens the equilibrium carbon isotope fractionation between DIC and $CO_{2(aq)}$ (see Equation S12). There is no way to examine the influence of SST changes and normalize for this effect on $\delta^{13}C_{CO2aq}$, as it is a true component of the signal, but the results shown in panels C and D suggest that the actual measured alkenone $\delta^{13}C$ variability is responsible for the two most important findings. The glacial/interglacial $\varepsilon_{p37:2}$ variability is linked to the SST variability (C), likely through effects on irradiance (D).

Appendix D: Supplementary Material for Chapter 5

1. Introduction

This appendix includes a detailed explanation of the materials and methods used in this study, including our approach for solving the carbonate system using paired alkenone-derived CO_2 estimates and foraminifera-derived pH estimates. We also include an expanded discussion of the implications of our work as highlighted in the main text.

2. Methodology

2.1. Core locations and samples

Twenty samples from ODP Site 806B (0°19.11'N, 159°21.68'E, 2520 m water depth) and eleven samples from nearby ODP Site 807A (3°36.42'N, 156°37.49'E, 2804 m water depth) were selected for this study. Both core sites are on the Ontong-Java Plateau in the Western Equatorial Pacific (WEP) and have been studied extensively for paleoclimate investigations on Neogene timescales. We refer to the samples collectively as "WEP" samples. Both sites are above the modern-day calcite saturation horizon and likely were in the past (Pälike et al., 2012; Si and Rosenthal, 2019). Sediments at both sites are characterized by nannofossil and foraminiferal ooze and chalks, with persistently high calcium carbonate content (~90 wt.%) over the last 20 Ma (Shipboard Scientific Party, 1991b; Shipboard Scientific Party, 1991a).

2.2. Chronology

The age model for this core is based on a combination of isotope stratigraphy and nannofossil and foraminifera biostratigraphy. The upper ~100 meters of Site 806B represent the

last ~4 million years, and they are included in the LR04 benthic isotope stack (Berger et al., 1993; Lisiecki and Raymo, 2005). The age model for sediments beyond Pliocene age is based on those presented by Nathan and Leckie (2009) and Suchéras-Marx and Henderiks (2014). We update the foraminiferal datums to ages reported by Wade et al. (2011). We then fit a third-order polynomial to the age-depth data (Fig. S1). The age model for ODP 807A is described in Chapter 4. Briefly, we use a combination of the published age models of Zhang et al. (2007) and Prentice et al. (1993), and update both to the LR04 age scale. We note that because our interest is in secular, long-term changes, age uncertainties in our Neogene samples of even 0.5 Ma are insignificant to the trends we observe.

2.3. Laboratory methods

2.3.1. Biomarker analysis

Samples obtained from the IODP Gulf Coast Repository were frozen and lyophilized for ~24 hours. All glassware used in sample handling and preparation was ashed for at least 8 hours at 450°C to remove all organic contaminants. Free lipids were extracted from sediments by accelerated solvent extraction (DIONEX ASE 350) at 1000 psi and a maximum temperature of 100°C, using a dichloromethane and methanol mixture (9:1 v/v, DCM:MeOH). Total lipid extracts (TLEs) were spiked with an internal standard then evaporated to dryness under a stream of N_2 gas. The TLE was passed over a sodium sulfate column in DCM to remove sediment clay-bound water that was present after extraction. TLEs were evaporated to dryness then split into aliphatic, ketone/ester (containing the alkenones), and polar fractions. We achieved this by flash silica gel column chromatography (~0.5 g organics-free deactivated Si gel in a 5-inch pipette) and eluting compounds with hexane, DCM, and methanol for the three respective fractions. The

DCM fraction (containing the alkenones) was evaporated, transferred to a 2-mL borosilicate vial for analysis, and reconstituted in toluene for analysis. The analytical methods were identical to those in Chapter 4.

2.3.2. Foraminifera analysis

We wet-sieved approximately 8 grams of the homogenized ASE-extracted sediment using a 63-µm sieve and a 2% ammonia solution in MilliQ ultrapure (≥17.9 Ω) water to prevent dissolution. For samples younger than 10 Ma (n = 10), we picked *Trilobatus trilobus* from the 355–425 µm size fraction. Specimens were sonicated briefly to remove clays, rinsed twice with MilliQ water, once with methanol, and left to dry at ~55°C until isotopic analysis. Stable carbon and oxygen isotope ratios were determined by dual-inlet isotope ratio mass spectrometry using a Thermo DeltaV with a Kiel IV Carbonate Device. Analytical precision was approximately 0.05‰ for $\delta^{13}C$ and 0.07‰ for $\delta^{18}O$.

We compiled published foraminifera stable isotope data from ODP 806/807 over the Neogene to supplement our analyses. We interpolate records of *T. trilobus* and *Dentoglobigerina altispira* to our sample depths (Corfield and Cartlidge, 1993; Nathan and Leckie, 2009). We generate a "mixed-layer stack" by determining the average offset between *T. trilobus* and *D. altispira* where they overlap, and adjust the *D. altispira* values to equivalent *T. trilobus* values. The approach is based on comparing nearby core-top *T. sacculifer* data on the Ontong Java Plateau to an estimate of the surface ocean preindustrial $\delta^{13}C_{DIC}$ at ODP 806/807. We use a 1σ error of 0.25‰ on our conversion of $\delta^{13}C_{foram}$ to $\delta^{13}C_{DIC}$, which we assume remains constant throughout the entire study interval.

2.3.3. Proxy-derived sea surface temperature estimates analysis

Sea surface temperatures for ODP 807 samples were reported in Chapter 4. At ODP 806, because temperatures are above 28°C for the entirety of the record, the alkenone SST proxy cannot provide much information about temperature, particularly in the warm Miocene (Müller et al., 1998). We estimate sea surface temperature by linearly interpolating published SST records from ODP 806 to our sample ages. For 0-2 Ma, we use the planktonic foraminiferal Mg/Ca record at ODP 806 of Wara et al. (2005). Beyond 2 Ma, we use the Tex86 record of Zhang et al. (2014) which extends to 12 Ma. For samples older than 12 Ma, we assume a constant temperature of 31°C. The temperature variations in the record (~3°C) are small relative to the calibration uncertainty of the proxies (~2-3°C), making uncertainties in our mean temperature estimates a small source of error in the trends we observe on Neogene timescales.

2.3.4. Coccolith length and cell size

We measured coccolith length using the SYRACO system at the Centre Européen de Recherche de l'Environnement et des Géosciences. The details of the image acquisition and processing used for these samples is identical to that used for the ODP 807 samples and are detailed in Appendix B. Briefly, this technique employs automated microscopy and circularly polarized transmitted light to image carbonate particles prepared on microscope slides. Particles are identified from the black background, classified into groups—roughly at the taxonomic Family level—by a neural network that has been trained on Cenozoic and recent nannofossils. Morphometric parameters (such as length and width) of the identified coccoliths are automatically produced using an algorithm in LabView. We vetted every identified coccolith to ensure that it was classified correctly. Because alkenone content is proportional to biomass and

because biomass scales with the volume of the cell, larger cells likely produce more alkenones than smaller cells (Chapter 3). Our estimate of the population-mean coccolith size weights contributions based on the estimated total biomass in the sample and the fractional contribution of each coccolith (which we take to represent one individual) to the total biomass. This approach yields slightly higher mean cell size estimates than the arithmetic mean. We then convert coccolith length to cell radius using the same approach as in Chapter 3, which is based on coccolithophore culture data (McClelland et al., 2016). Cell radius is linearly related to coccolith length by:

$$\text{Cell radius } (\mu m) = 0.57 \times \text{coccolith length } (\mu m) + 0.45 \qquad \text{[Eq. S1]}$$

3. Solving the marine carbonate system

We use existing data describing the marine boron isotope system and our new atmospheric pCO_2 estimates to speciate the carbonate system though time. We also require estimates of temperature, salinity, and the $[Mg^{2+}]$ and $[Ca^{2+}]$ through time to constrain the thermodynamic dissociation constants dictating the acid-base chemistry of seawater.

3.1. The boron isotope-pH proxy and $\delta^{11}B_{borate}$ from $\delta^{11}B_{foram}$

Boron isotope fractionation in seawater is tightly related to seawater pH, providing a key constraint on the temporal evolution of the marine carbonate system (Dickson, 1990; Hain et al., 2018). Dissolved boron exists primarily as borate ion ($B(OH)_4^-$) and boric acid ($B(OH)_3$). The partitioning between these two species is a function of seawater pH, similar to the speciation of carbonic acid in seawater. While the $\delta^{11}B$ of each species varies with pH, there is a constant

isotopic offset between borate ion and boric acid of ~27‰, and the isotope mass balance of these two species equals the boron isotope ratio of seawater ($\delta^{11}B_{SW}$) (Hönisch et al., 2019).

Borate ions are incorporated into marine carbonates, and culture and core-top studies with foraminifera demonstrate a linear relationship between $\delta^{11}B_{borate}$ and $\delta^{11}B_{foram}$. Given the pH-dependence of $\delta^{11}B_{borate}$, it is possible to estimate seawater pH from $\delta^{11}B_{foram}$ (Hemming and Hanson, 1992; Hönisch et al., 2019) using the equation:

$$pH = pK^*_B - \log\left(-\frac{\delta^{11}B_{SW} - \delta^{11}B_{borate}}{(\delta^{11}B_{SW} - (\alpha_B \times \delta^{11}B_{borate}) - 1000 \times (\alpha_B - 1)}\right) \quad \text{[Eq. S2]}$$

This formulation demonstrates the dependence of pH estimates on the boron isotope ratio of seawater, borate ion, the fractionation factor (α_B), and the dissociation constant (pK^*_B), which is a function temperature, salinity, pressure and seawater magnesium and calcium ion concentration. Flux studies suggest the residence time of boron in seawater is long, ranging from ~14 to ~20 Myr (Spivack and Edmond, 1987; Lemarchand et al., 2000; Lemarchand et al., 2002). However, estimates of the boron isotope ratio of seawater show ~2‰ variations over the last 20 (Raitzsch and Hönisch, 2013; Greenop et al., 2017). Our approach here is to pair the published $\delta^{11}B_{foram}$ ($\delta^{11}B_{borate}$, by inference) measurements with our alkenone-derived atmospheric pCO_2 estimates to solve for the probabilistic temporal evolution of the carbonate system. We explain the calculations in detail below.

We compile measurements of $\delta^{11}B_{foram}$ measurements from the literature covering the last ~20 Ma (Hönisch et al., 2009; Seki et al., 2010; Bartoli et al., 2011; Foster et al., 2012; Badger et al., 2013; Greenop et al., 2014; Martínez-Botí et al., 2015; Henehan et al., 2017; Sosdian et al., 2018; Dyez et al., 2018). Our primary focus here, however, is the correspondence between our

Miocene CO_2 estimates and the Miocene pH estimates. The only Late Miocene $\delta^{11}B$-pH data are from Sosdian et al. (2018), and Middle and Early Miocene samples are from three other studies (Foster et al., 2012; Badger et al., 2013; Greenop et al., 2014). As such, we are mainly interested in comparing the secular trends during the Neogene from these studies to our results.

All measurements were made on either *Globigerinoides ruber* (300-355 μm size fraction), or *Trilobatus trilobus* (>500 μm, 515-865 μm, 425-500 μm, and 300-355 μm size fractions). Measurements of $\delta^{11}B_{foram}$ are made using either a multi collector-inductively coupled plasma-mass spectrometer (MC-ICP-MS; NEPTUNE) or a thermal ionization mass spectrometer (TIMS; NBS or TRITON). The translation of $\delta^{11}B_{foram}$ to $\delta^{11}B_{borate}$ depends on the species, size fraction, and measurement technique (e.g. Farmer et al., 2016; Dyez et al., 2018).

For *G. ruber* samples measured by MC-ICP-MS, we use the calibration of Henehan et al. (2013):

$$\delta^{11}B_{borate} = (\delta^{11}B_{G.ruber} - 8.87) / 0.60 \quad \text{[Eq. S3]}$$

For *T. trilobus* data measured by MC-ICP-MS, we use the calibration reported by Sosdian et al. (2018) to convert measured $\delta^{11}B_{foram}$ to $\delta^{11}B_{borate}$:

$$\delta^{11}B_{borate} = (\delta^{11}B_{T.trilobus} - 2.69) / 0.833 \quad \text{[Eq. S4]}$$

This relationship uses the slope determined by Sanyal et al. (2001), but adjusts the intercept so the core-top value at ODP 999 matches measured in-situ $^{\delta 11}B_{borate}$ in the Caribbean Sea (Henehan et al., 2016). See Sosdian et al. (2018) for more details.

For *T. trilobus* measured by N-TIMS, we use the calibration reported by Dyez et al. (2018):

$$\delta^{11}B_{borate} = (\delta^{11}B_{T.trilobus} - 6.42) / 0.73 \quad \text{[Eq. S5]}$$

We use the constant fractionation factor (α_B) of 1.0272 ‰ in our calculations using Equation S2 (Klochko et al., 2006).

3.2. Thermodynamic dissociation constants and the major-ion chemistry of seawater

We determine pK^*_B values using the MyAMI ion-interaction model, which accounts for the effect of changes in seawater $[Mg^{2+}]$ and $[Ca^{2+}]$ on dissociation constants (Hain et al., 2015). As an approximation for the oceanic $[Mg^{2+}]$ and $[Ca^{2+}]$ history, we use the trendlines derived by Sosdian et al. (2018) which are based on the fluid-inclusion data of (Brennan et al., 2013) and (Horita et al., 2002). We assume $[Ca^{2+}]$ evolves linearly from 10 mmol kg^{-1} at present to 14.37 mmol kg^{-1} at 25 Ma, and $[Mg^{2+}]$ evolves linearly from 55.24 mmol kg^{-1} to 41.43 mmol kg^{-1} at 25 Ma. We assume a uniform uncertainty in $[Mg^{2+}]$ and $[Ca^{2+}]$ of 0 mmol kg^{-1} at modern that increases linearly to ±3 mmol kg^{-1} at 25 Ma. We recalculate sea surface temperatures from the Mg/Ca ratios of planktonic foraminifera, which are in most cases reported with the $\delta^{11}B_{foram}$ data. We use our estimated Mg/Ca ratio of seawater (Mg/Ca_{SW}) from the method described above and correct for the influence of variable Mg/Ca_{SW} on SST following the approach of Sosdian et al. (2018):

$$Mg/Ca_{SW,corr} = \left(\frac{Mg/Ca_{SW,T}}{Mg/Ca_{SW,modern}}\right)^{0.41} \quad \text{[Eq. S6]}$$

Temperature is calculated using the equation of Anand et al. (2003):

$$T(°C) = \ln\left(\frac{Mg/Ca_{foram}}{0.38*Mg/Ca_{SW,corr}}\right) / 0.09 \quad \text{[Eq. S7]}$$

Where the foraminiferal Mg/Ca ratio is not reported and we cannot recalculate SST, we use the reported temperature values.

We assume a constant salinity of 35 psu with a uniform uncertainty of 1 psu. Salinity has a small effect on carbonate system calculations and pH reconstructed using $\delta^{11}B_{foram}$ (Hain et al., 2015; Dyez et al., 2018).

To determine the dissociation constants (K_0^*, K_1^*, K_2^*, K_B^*, K_W^*) to apply to each sample, we first estimate the Pitzer interaction coefficients defining their sensitivity to $[Mg^{2+}]$ and $[Ca^{2+}]$ variations using the output of the MyAMI model from Hain et al. (2015). For the estimated seawater $[Mg^{2+}]$ and $[Ca^{2+}]$ values in each sample, we use nearest-neighbor interpolation on their Table S1, which lists Pitzer coefficients for seawater $[Mg^{2+}]$ and $[Ca^{2+}]$ from 0 to 60 mmol kg^{-1} at a spacing of 1 mmol kg^{-1}. We use their equations 1a-1h to calculate dissociation constants as a function of temperature, salinity, and ionic strength, which we calculate after Dickson et al. (2007):

$$I = \frac{19.924 \times \text{salinity}}{(1000\text{-}1.005 \times \text{salinity})} \quad \text{[Eq. S8]}$$

Because we only consider planktonic foraminifera in these calculations, we ignore the very small pressure effect on dissociation constants. Sensitivity calculations demonstrate that changing depth (pressure) from 0m to 100m—the typical range of calcification depths of mixed-layer planktonic foraminifera—changes the calculated pH by ~0.004 units (~0.05%). We do not consider variations in pressure through time. Ignoring the pressure effect does not affect the relative trends in our calculations, and has a miniscule effect on the absolute values.

3.3. Speciating the carbonate system

Because an estimate of pH is from $\delta^{11}B_{foram}$ is dependent on $\delta^{11}B_{SW}$, we solve for the system of carbonate chemistry solutions that is possible given our atmospheric pCO_2 and the measured $\delta^{11}B_{foram}$. We then filter this range of solutions using independent estimates of $\delta^{11}B_{SW}$.

We propagate uncertainties using a Monte Carlo routine (nMC = 150). First, we generate Monte Carlo distributions of salinity, $[Mg^{2+}]$ and $[Ca^{2+}]$. We calculate SST in each of these Monte Carlo simulations using the estimated Mg/Ca_{SW}. We then add the SST error of ± 1.5°C (1σ) from the calibration of Mg/Ca_{foram} to SST (°C). For each Monte Carlo simulation, we calculate dissociation constants as a function of temperature, salinity, $[Mg^{2+}]$ and $[Ca^{2+}]$, as described above in Appendix D, Section 3.2. For each $\delta^{11}B$ sample, we take the average of all alkenone-CO_2 samples within 0.25 Ma of the $\delta^{11}B$ sample, then generate Monte Carlo distributions of atmospheric CO_2 for each sample, assuming that surface ocean pCO_2 is in equilibrium with atmospheric pCO_2. We set a lower threshold on the allowable pCO_2 of 100 µatm, and remove all Monte Carlo simulations that produce pCO_2 values less than this cutoff. We finally propagate error in $\delta^{11}B_{foram}$ measurement uncertainty, but do not include uncertainty in the conversion of $\delta^{11}B_{foram}$ to $\delta^{11}B_{borate}$, as that is only ~20% of the $\delta^{11}B_{borate}$ uncertainty, and the majority (~80%) is from $\delta^{11}B_{foram}$ analytical uncertainty. This provides pairs of $\delta^{11}B_{borate}$ and CO_2 that can be used to constrain the carbonate system.

For each $\delta^{11}B_{foram}$ sample and each Monte Carlo iteration, we then solve the carbonate system using the paired CO_2 value and a DIC range of 200 to 8000 µmol kg^{-1}. We first calculate dissolved CO_2 from pCO_2 and K_0^* using the equation:

$$pCO_2 = K_0^* \times [CO_2] \qquad \text{[Eq. S9]}$$

We then calculate pH after Zeebe and Wolf-Gladrow (2001) equation 1.5.79:

$$DIC \times [H^+]^2 = CO_2 \times ([H^+]^2 + K_1^*[H^+] + K_1^*K_2^*) \qquad \text{[Eq. S10]}$$

We calculate the hydroxyl ion concentration from pH and K_w^*:

$$[OH^-] = K_w^* / [H^+] \quad \text{[Eq. S11]}$$

We then calculate the concentrations of bicarbonate ion and carbonate ion from pH and DIC after Zeebe and Wolf-Gladrow (2001):

$$[HCO_3^-] = \frac{DIC}{\left(1+\frac{[H^+]}{K_1^*}\right)+\left(\frac{K_2^*}{[H^+]}\right)} \quad \text{[Eq. S12]}$$

$$[CO_3^{2-}] = \frac{DIC \times K_1^* K_2^*}{[H^+]^2 + K_1^*[H^+] + K_1^* K_2^*} \quad \text{[Eq. S13]}$$

We assume total boron [B_T] remained constant at 432.6 μmol kg^{-1}, and determine the concentrations of borate ion and boric acid using K_B^*. We calculate practical alkalinity as:

$$PA = [HCO_3^-] + 2 \times [CO_3^{2-}] + [BOH_4^-] + [OH^-] - [H^+] \quad \text{[Eq. S14]}$$

Finally, we solve for $\delta^{11}B_{SW}$ using:

$$\delta^{11}B_{SW} = \frac{\left(\delta^{11}B_{borate} + 10^{\left(pK_B^* - pH\right)} \times \left(1000 \times \alpha_B + \alpha_B \times \delta^{11}B_{borate} - 1000\right)\right)}{10^{\left(pK_B^* - pH\right)} + 1}$$

[Eq. S15]

This Monte Carlo exercise provides a set of all possible solutions of $\delta^{11}B_{SW}$ for the given pCO_2 and $\delta^{11}B_{foram}$, for the domain of DIC = 200 to 8000 μmol kg^{-1} at a spacing of 30 μmol kg^{-1} (nDIC = 251). We refer to this set of $\delta^{11}B_{SW}$ values as "solved" ($\delta^{11}B_{SW\_SOLVED}$), as they are derived from our measured pCO_2 and the DIC search grid. We can then calculate the

probabilities of allowable DIC values using the independently known $\delta^{11}B_{SW}$ and its probability distribution function.

3.4. Calculating the DIC probability distribution from $\delta^{11}B$ constraints

Ultimately, our two constraints on the carbonate system are CO_2 and pH, but we do not impose a pH value initially. We speciate the carbonate system with a single CO_2 value for the entire DIC range above, solve for $\delta^{11}B_{SW}$ using the measured $\delta 11B_{borate}$, and then determine which DIC values yield convergent solutions by comparison with the independent $\delta^{11}B_{SW}$ value.

To constrain the set of DIC values that agree with our alkenone-based pCO_2 and existing $\delta^{11}B_{foram}$ records, we use the estimate of $\delta^{11}B_{SW}$ from Greenop et al. (2017). We refer to this as $\delta^{11}B_{SW\_G17}$ or "G17." We choose this reconstruction instead of Lemarchand et al. (2002) or Raitzsch and Hönisch (2013) because Greenop et al. (2017) provide an easily accessible timeseries of estimated $\delta^{11}B_{SW}$ and the associated uncertainty. However, we note that the three are not dramatically different for the last 20 Ma. The pH estimates shown in Sosdian et al. (2018) are quite similar for the last 16 Ma regardless of the $\delta^{11}B_{SW}$ history used (see their Fig. 2E).

We calculate the probability distribution of DIC that agrees with the measured CO_2, the measured $\delta^{11}B_{foram}$, and $\delta^{11}B_{SW\_G17}$ through two-dimensional histogram counting. For each $\delta^{11}B_{foram}$ sample, we first bin the $\delta^{11}B_{SW\_SOLVED}$ values into nDIC equally spaced bins ("BIN_{SW_SOLVED}"), by rounding down (up) the minimum (maximum) of the set of $\delta^{11}B_{SW\_SOLVED}$ solutions to the nearest integer, and shifting the bin edges down by half of the distance between each bin.

Next, we estimate the mean $\delta^{11}B_{SW\_G17}$ at the sample age by linear interpolation from the G17 dataset. We assume the $\delta^{11}B_{SW\_G17}$ uncertainty is normally distributed and generate a

synthetic distribution of 1,000,000 $\delta^{11}B_{SW\_G17}$ values using the mean and standard deviation. We then calculate the probability of the $\delta^{11}B_{SW\_G17}$ from the synthetic distribution in each BIN_{SW_SOLVED}. These probabilities will be used as the multiplicand on the joint probability distribution of DIC and $\delta^{11}B_{SW\_SOLVED}$. We calculate that joint probability using the *histcounts2* function in MATLAB R2019a. We bin the DIC solutions into nDIC equally-spaced bins ("BIN_{DIC}"), shifting the bin edges down by half of the bin width (as described above), and using DIC, $\delta^{11}B_{SW\_S}$, and BIN_{DIC}, BIN_{SW_S} as the data and bin definitions, respectively, in *histcounts2*. This returns a probability of a DIC–$\delta^{11}B_{SW\_S}$ solution pair in that bivariate space.

Because the $\delta^{11}B_{SW\_G17}$ distribution has been binned to the same "BIN_{SW_S}" resolution, we can multiply that probability mass function by the two-dimensional probability space of our $\delta^{11}B_{SW\_S}$ and DIC solutions. We take this product, which effectively forces $\delta^{11}B_{SW\_S}$ and DIC values outside the probability distribution of $\delta^{11}B_{SW\_G17}$ to zero, and increases the probability of those at the mean $\delta^{11}B_{SW\_G17}$ value. **This product defines the probability of allowable DIC values that solves the system of $\delta^{11}B_{foram}$ and alkenone-derived pCO_2 with fully propagated uncertainties, and constrained by the coeval $\delta^{11}B_{SW\_G17}$ and its associated uncertainty.**

To determine the mean DIC, we take the vector product of the probabilities multiplied by the DIC value. We calculate the variance by subtracting the mean from each unique DIC value, squaring this difference, multiplying the squared difference by the relative abundance (probability) of each DIC value, and taking the sum of these products. Finally, we estimate the standard deviation as the square root of the variance.

Because we determine ancillary components of the carbonate system in the process of calculating $\delta^{11}B_{SW\_S}$ from pCO_2 and DIC, we also have a system of solutions for alkalinity, its components, and the surface carbonate saturation state (Ω_{Ca}), adjusted for secular changes in

[Ca^{2+}] and the effects on the solubility constant. We repeat the joint probability calculation as described above for each carbonate system component, always using nDIC bins that are uniformly spaced from the minimum to the maximum value, with edges shifted by half the linear distance.

4. Supplementary Discussion

4.1. Argument for constant irradiance

Our interpretations of $\varepsilon_{p37:2}$ are based on the "integrated irradiance" linear model presented in Chapter 2, which treats ε_p as a linear combination of CO_2 (μmol kg^{-1}), daily integrated irradiance (mol m^{-2} d^{-1}), growth rate (d^{-1}), and cell radius (μm). As described in section 2.1 in the main text, we determine the irradiance value to use by solving in the late Pleistocene samples where CO_2 is known from the composite ice-core record. We then apply the mean late Pleistocene irradiance value downcore, using the standard deviation of the solved values as the uncertainty envelope. This centers the average pCO_2 estimates of the past 800 ka roughly on the ice-core average. In the Late Pleistocene at ODP 806/807, changes in $\varepsilon_{p37:2}$ are reliably modeled using the known CO_2 change, measured coccolith size, and constant irradiance (Chapter 4), suggesting irradiance may be stable on longer timescales. Nevertheless, we consider how variations in irradiance would affect our interpretation of $\varepsilon_{p37:2}$.

Figure S4 shows the result of the sensitivity analysis that solves for the irradiance value required to explain the observed measured $\varepsilon_{p37:2}$ and coccolith lengths at a constant surface ocean pCO_2 of 280 μatm. The necessary irradiance in the peak Middle Miocene $\varepsilon_{p37:2}$ and cell size is around 40 mol m^{-2} d^{-1}, which is ~2 times the Late Pleistocene average at ODP 806/807 (Fig. S4). While not physically impossible, we argue that such an increase was unlikely. First, we calculate

the depth at which an irradiance value of 40 mol m^{-2} d^{-1} exists using modern mean annual incident surface irradiance and the modern mean annual attenuation coefficient of photosynthetically available radiation, as estimated from spatial interpolation of satellite-derived climatological datasets. We then examine records of temperature change at ODP 806 over the late Neogene and argue that dynamical changes are unlikely to have made the depth of production shallower in the Miocene than today.

The average modern incident surface irradiance at ODP 806/807 is ~45 mol m^{-2} d^{-1}. Using the modern mean annual Kd_{490} value of ~0.03 m^{-1}, this would require an integrated production depth of ~4 meters to yield 40 mol m^{-2} d^{-1}. An analysis of late Holocene samples with the irradiance model we use here found that the average depth of alkenone production varies by location, but globally it is towards the base of the mixed layer (Chapter 3). The modern average mixed layer depth is ~55 meters (de Boyer Montégut et al., 2004), which suggests the average production depth well below the 4 m calculated here.

Paleoceanographic evidence suggests the mixed layer was likely deeper in the Neogene compared to the Pleistocene. Ford et al. (2015) measured magnesium calcium ratios of surface-dwelling planktonic foraminifera *G. sacculifer* and subsurface-dwelling *G. tumida* over the last 5 Ma at ODP 806. Their analyses showed that the difference between subsurface and surface temperatures of ~8°C in the Late Pleistocene was reduced to ~2°C in the early Pliocene. These data suggest that the two species were either calcifying at the same depth, or that the depth of the subsurface waters was much closer to surface temperatures. Given the modern habitats of these two species, the latter explanation of a warmer subsurface is more likely. Subsurface waters in the equatorial Pacific are sourced from shallow overturning cells in the subtropical Pacific (Zeller et al., 2019). These regions were also several degrees warmer in the early Pliocene and

Late to Middle Miocene (LaRiviere et al., 2012), which would support a warmer subsurface in the Western Equatorial Pacific. A deeper thermocline is consistent with GDGT-based temperature estimates over the last 12 Ma at this same location (Zhang et al., 2014). Because GDGTs are produced throughout the upper several hundred meters and not in the uppermost surface waters (Hurley et al., 2019), the temperature estimates likely reflect the integrated temperature change of the upper water column (Ho and Laepple, 2016). Taken together, these data suggest that the thermal gradient between the surface and the subsurface was reduced during the Pliocene and Miocene. Assuming constant salinity, this would reduce the density contrast between the surface and subsurface, and equatorial winds would more effectively mix the upper water column. We assume here that salinity variations in the surface and subsurface are small enough such that temperature is the dominant control on density in the western equatorial Pacific on Neogene timescales. As discussed in the preceding paragraph, the irradiance levels required for the null hypothesis (atmospheric pCO_2 = constant 280 ppm) necessitate a mean alkenone production depth of ~4 meters, which we argue is unlikely given the temperature considerations here.

Although the density structure of the upper water column is an important component governing the mixing of phytoplankton, the mean depth of alkenone production will be a function of several factors that control algal growth and photosynthesis, including irradiance, nutrient availability, temperature, and predation. Our analysis and argument here assume that the depth of optimum growth at ODP 806 was not shallower in the past compared to the Pleistocene at ODP 806. A decoupling of the nutricline from the local hydrography and a shallowing of the depth of nutrient remineralization could shift the average depth of production to shallower depths. However, nutrient availability is a function of local remineralization and vertical and

lateral advection of nutrient-rich waters, and it is unclear how these processes operated during the Neogene. If new proxy data can provide additional constraints on the irradiance conditions experienced by alkenone producing algae during the Neogene, we will be able to refine our CO_2 estimates.

4.2. Comparison to existing calculations of the marine carbonate system

Our carbonate system solution is notably different from existing calculations of carbonate chemistry through time (Ridgwell, 2005; Caves et al., 2016; Boudreau et al., 2019; Zeebe and Tyrrell, 2019). One contribution to this difference is the input parameters used by past studies. Because most published Neogene atmospheric CO_2 records are lower than our data (e.g. Beerling and Royer, 2011), inferences of DIC and alkalinity based on lower pCO_2 return lower values, as is seen in Zeebe and Tyrrell (2019) (c.f. their Fig. 5). A second explanation is the approach used to constrain the second carbonate parameter. Previous studies have inferred changes in surface carbonate ion concentration from reconstructions of the carbonate compensation depth (CCD), an assumed constant whole-ocean calcite saturation state (Ω_C), and assumed proportionality between deep ocean and surface ocean carbonate chemistry (Tyrrell and Zeebe, 2004; Caves et al., 2016). Recent sediment and modeling studies suggest that the depth of the CCD is not necessarily tightly linked to surface ocean chemistry, because large changes in dissolution and burial occur over several kilometers between the surface and CCD (Greene et al., 2019; Si and Rosenthal, 2019).

Greene et al. (2019) analyzed the global abundance of calcium carbonate in deep-sea sediments from the early Cenozoic and found that the CCD can be an insensitive gauge of the supply of alkalinity from weathering. Through analyses with a global carbon cycle model, the

authors found that the increased $CaCO_3$ burial required under higher weathering fluxes was achieved by burying more carbonate at depths above than the CCD, without changing the depth of the CCD. While the CCD is undoubtedly linked to the saturation state of the deep ocean, the depth of the calcite saturation horizon (CSH, where $\Omega_C = 1$) appears to be much more sensitive to CO_2 perturbations (Greene et al., 2019). Comparing the Middle Miocene to the Pleistocene, Si and Rosenthal (2019) argued for an expanded lysocline (the depth between the saturation horizon and the CCD; (Boudreau et al., 2018)) based on large changes in carbonate preservation at ~3km depth, even though the CCD remained constant. Boudreau and Luo (2017) also estimate deep ocean carbonate ion from records of the CCD and carbonate burial fluxes, though with different assumptions than Tyrrell and Zeebe (2004) or Caves et al. (2016). Their model does not account for dissolution in the water column. Given the estimated variability in the CSH (Si and Rosenthal, 2019), it is possible that water-column dissolution is an important component, and that the CSH was higher than they estimated for the Neogene. If the Miocene carbonate burial flux data of Suchéras-Marx and Henderiks (2014) and Si and Rosenthal (2019) are correct and indeed represent higher global carbonate burial, the inferred history of surface ocean alkalinity by Boudreau et al. (2019), which is based on lower $CaCO_3$ burial during the Miocene, may not be compatible.

One implication of our calculated surface ocean carbonate chemistry and the above discussion of carbonate preservation is that there was likely a strengthened gradient between surface and deep carbonate chemistry during the Miocene. We consider this premise through the relationship between organic matter oxidation, temperature, and pH. The oxidation of marine organic matter is a complex system of reactions, but degradability, temperature, and electron acceptor availability are some of the fundamental controls (Arndt et al., 2013). In marine

sediments, the temperature below the seafloor is a key determinant of the sedimentary organic carbon content (Malinverno and Martinez, 2015), and oxidation of the oceanic dissolved organic carbon pool is quite sensitive to an increase in temperature (Lønborg et al., 2018). Given the overall higher temperatures in the Miocene (Billups and Schrag, 2002; Shevenell et al., 2008; Lear et al., 2015), organic matter remineralization in the ocean interior may have been enhanced. This could act to increase the CO_2 content of deep water, making it more acidic, which could allow greater carbonate production in the surface ocean and greater dissolution at depth. A reanalysis of existing paired planktic and benthic $\delta^{11}B_{foram}$ data from the same or adjacent samples suggests this was possible: though the data are sparse, the estimated pH gradient between surface and deep ocean was on average higher from 20 to 10 Ma compared to 10 to 0 Ma (Fig. S3). The reconstruction of $\delta^{11}B_{SW}$ we use here is based on $\delta^{11}B$ measurements of planktic and benthic foraminifera in the same samples, but also on indirect estimates of deep ocean pH (Greenop et al., 2017). We therefore consider ΔpH under different assumptions of $\delta^{11}B_{SW}$, finding that even with the estimated change in $\delta^{11}B_{SW}$ from 15 Ma to present, the mean ΔpH from 20 to 10 Ma is still higher than 10 to 0 Ma by about 0.07 pH units. The uncertainty bounds on the means of the two time periods do have some overlap (Fig. S3C), so the data are not conclusive at a high confidence interval. However, there is a clear spatiotemporal bias in the data. The estimates come from 4 different sites, and only one site has paired $\delta^{11}B$ planktic and benthic in both the 0-10 Ma and 10-20 Ma windows. This site is ODP 926, which shows a large change in surface-to-deep pH gradient between ~10 and 8 Ma (Fig. S3B, circles), when the $\delta^{11}B_{SW}$ probably did not change by more than 0.2‰, assuming a maximum rate of change of 0.1‰ / Myr (Dyez et al., 2018). Given the limited geographic coverage of this analysis and changes in the data availability by site through time, we are hesitant to ascribe too much

significance to this finding. Additional measurements of paired benthic and planktonic $\delta^{11}B$ can robustly assess how the surface-to-deep pH gradient has changed through time.

5. References

Anand P., Elderfield H. and Conte M. H. (2003) Calibration of Mg/Ca thermometry in planktonic foraminifera from a sediment trap time series. *Paleoceanography* **18**, 1050 2002PA000846.

Arndt S., Jørgensen B. B., LaRowe D. E., Middelburg J. J., Pancost R. D. and Regnier P. (2013) Quantifying the degradation of organic matter in marine sediments: A review and synthesis. *Earth-Science Rev.* **123**, 53–86.

Badger M. P. S., Lear C. H., Pancost R. D., Foster G. L., Bailey T. R., Leng M. J. and Abels H. A. (2013) CO2 drawdown following the middle Miocene expansion of the Antarctic Ice Sheet. *Paleoceanography* **28**, 42–53.

Bartoli G., Hönisch B. and Zeebe R. E. (2011) Atmospheric CO2 decline during the Pliocene intensification of Northern Hemisphere glaciations. *Paleoceanography* **26**, 1–14.

Beerling D. J. and Royer D. L. (2011) Convergent Cenozoic CO2 history. *Nat. Geosci.* **4**, 418–420.

Berger W. H., Bickert T., Schmidt H. and Wefer G. (1993) Quaternary Oxygen Isotope Record of Pelagic Foraminifers: Site 806, Ontong Java Plateau. In *Proceedings of the Ocean Drilling Program, 130 Scientific Results* (eds. W. H. Berger, L. W. Kroenke, and L. A. Mayer). Ocean Drilling Program. pp. 381–395.

Billups K. and Schrag D. P. (2002) Paleotemperatures and ice volume of the past 27 Myr revisited with paired Mg/Ca and 18 O/ 16 O measurements on benthic foraminifera. *Paleoceanography* **17**, 3-1-3–11.

Boudreau B. P. and Luo Y. (2017) Retrodiction of secular variations in deep-sea CaCO3 burial during the Cenozoic. *Earth Planet. Sci. Lett.* **474**, 1–12.

Boudreau B. P., Middelburg J. J. and Luo Y. (2018) The role of calcification in carbonate compensation. *Nat. Geosci.* **11**, 894–900.

Boudreau B. P., Middelburg J. J., Sluijs A. and van der Ploeg R. (2019) Secular variations in the carbonate chemistry of the oceans over the Cenozoic. *Earth Planet. Sci. Lett.* **512**, 194–206.

de Boyer Montégut C., Madec G., Fischer A. S., Lazar A. and Iudicone D. (2004) Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *J. Geophys. Res. C Ocean.* **109**, 1–20.

Brennan S. T., Lowenstein T. K. and Cendon D. I. (2013) The major-ion composition of cenozoic seawater: the past 36 million years from fluid inclusions inmarine halite. *Am. J. Sci.* **313**, 713–775.

Caves J. K., Jost A. B., Lau K. V. and Maher K. (2016) Cenozoic carbon cycle imbalances and a variable weathering feedback. *Earth Planet. Sci. Lett.* **450**, 152–163.

Corfield R. M. and Cartlidge J. E. (1993) Oxygen and Carbon Isotope Stratigraphy of the Middle Miocene, Holes 805B and 806B. In *Proceedings of the Ocean Drilling Program, 130 Scientific Results* Ocean Drilling Program. pp. 307–322.

Dickson A. G. (1990) Thermodynamics of the dissociation of boric acid in potassium chloride solutions from 273.15 to 318.15 K. *J. Chem. Eng. Data* **35**, 253–257.

Dickson Andrew G., Sabine Christopher L. and Christian James R. (2007) *Guide to Best Practices for Ocean CO2 Measurements*. eds. A.G. Dickson, C.L. Sabine, and J.R. Christian,

Dyez K. A., Hönisch B. and Schmidt G. A. (2018) Early Pleistocene Obliquity-Scale pCO2 Variability at ~1.5 Million Years Ago. *Paleoceanogr. Paleoclimatology* **33**, 1270–1291.

Farmer J. R., Hönisch B. and Uchikawa J. (2016) Single laboratory comparison of MC-ICP-MS and N-TIMS boron isotope analyses in marine carbonates. *Chem. Geol.* **447**, 173–182.

Ford H. L., Ravelo A. C., Dekens P. S., Lariviere J. P. and Wara M. W. (2015) The evolution of the equatorial thermocline and the early Pliocene El Padre mean state. *Geophys. Res. Lett.* **42**, 4878–4887.

Foster G. L., Lear C. H. and Rae J. W. B. (2012) The evolution of pCO2, ice volume and climate during the middle Miocene. *Earth Planet. Sci. Lett.* **341**–**344**, 243–254.

Greene S. E., Ridgwell A. J., Kirtland Turner S., Schmidt D. N., Pälike H., Thomas E., Greene L. K. and Hoogakker B. A. A. (2019) Early Cenozoic Decoupling of Climate and Carbonate Compensation Depth Trends. *Paleoceanogr. Paleoclimatology* **34**, 930–945.

Greenop R., Foster G. L., Wilson P. A. and Lear C. H. (2014) Middle Miocene climate instability associated with high-amplitude CO2 variability. *Paleoceanography* **29**, 845–853.

Greenop R., Hain M. P., Sosdian S. M., Oliver K. I. C., Goodwin P., Chalk T. B., Lear C. H., Wilson P. A. and Foster G. L. (2017) A record of Neogene seawater δ11B reconstructed from paired δ11B analyses on benthic and planktic foraminifera. *Clim. Past* **13**, 149–170.

Hain M. P., Foster G. L. and Chalk T. (2018) Robust Constraints on Past CO2 Climate Forcing From the Boron Isotope Proxy. *Paleoceanogr. Paleoclimatology* **33**, 1099–1115.

Hain M. P., Sigman D. M., Higgins J. A. and Haug G. H. (2015) The effects of secular calcium and magnesium concentration changes on the thermodynamics of seawater acid/base chemistry: Implications for Eocene and Cretaceous ocean carbon chemistry and buffering. *Global Biogeochem. Cycles* **29**, 517–533.

Hemming N. G. and Hanson G. N. (1992) Boron isotopic composition and concentration in modern marine carbonates. *Geochim. Cosmochim. Acta* **56**, 537–543.

Henehan M. J., Evans D., Shankle M., Burke J. E., Foster G. L., Anagnostou E., Chalk T. B., Stewart J. A., Alt C. H. S., Durrant J. and Hull P. M. (2017) Size-dependent response of foraminiferal calcification to seawater carbonate chemistry. *Biogeosciences* **14**, 3287–3308.

Henehan M. J., Foster G. L., Bostock H. C., Greenop R., Marshall B. J. and Wilson P. A. (2016) A new boron isotope-pH calibration for Orbulina universa , with implications for understanding and accounting for 'vital effects.' *Earth Planet. Sci. Lett.* **454**, 282–292.

Henehan M. J., Rae J. W. B., Foster G. L., Erez J., Prentice K. C., Kucera M., Bostock H. C., Martínez-Botí M. A., Milton J. A., Wilson P. A., Marshall B. J. and Elliott T. (2013) Calibration of the boron isotope proxy in the planktonic foraminifera Globigerinoides ruber for use in palaeo-CO2 reconstruction. *Earth Planet. Sci. Lett.* **364**, 111–122.

Ho S. L. and Laepple T. (2016) Flat meridional temperature gradient in the early Eocene in the subsurface rather than surface ocean. *Nat. Geosci.* **9**, 606–610.

Hönisch B., Eggins S. M., Haynes L. L., Allen K. A., Holland K. D. and Lorbacher K. (2019) *Boron Proxies in Paleoceanography and Paleoclimatology*., John Wiley & Sons, Ltd, Chichester, UK.

Hönisch B., Hemming N. G., Archer D., Siddall M., McManus J. F., Honisch B., Hemming N. G., Archer D., Siddall M. and McManus J. F. (2009) Atmospheric Carbon Dioxide Concentration Across the Mid-Pleistocene Transition. *Science (80-.).* **324**, 1551–1554.

Horita J., Zimmermann H. and Holland H. D. (2002) Chemical evolution of seawater during the Phanerozoic. *Geochim. Cosmochim. Acta* **66**, 3733–3756.

Hurley S. J., Close H. G., Elling F. J., Jasper C. E., Gospodinova K., McNichol A. P. and Pearson A. (2019) CO2-dependent carbon isotope fractionation in Archaea, Part II: The marine water column. *Geochim. Cosmochim. Acta* **261**, 383–395.

Klochko K., Kaufman A. J., Yao W., Byrne R. H. and Tossell J. A. (2006) Experimental measurement of boron isotope fractionation in seawater. *Earth Planet. Sci. Lett.* **248**, 261–270.

LaRiviere J. P., Ravelo A. C., Crimmins A., Dekens P. S., Ford H. L., Lyle M. W. and Wara M. W. (2012) Late Miocene decoupling of oceanic warmth and atmospheric carbon dioxide forcing. *Nature* **486**, 97–100.

Lear C. H., Coxall H. K., Foster G. L., Lunt D. J., Mawbey E. M., Rosenthal Y., Sosdian S. M., Thomas E. and Wilson P. A. (2015) Neogene ice volume and ocean temperatures: Insights from infaunal foraminiferal Mg/Ca paleothermometry. *Paleoceanography* **30**, 1437–1454.

Lemarchand D., Gaillardet J., Lewin É. and Allègre C. . (2002) Boron isotope systematics in large rivers: implications for the marine boron budget and paleo-pH reconstruction over the Cenozoic. *Chem. Geol.* **190**, 123–140.

Lemarchand D., Gaillardet J., Lewin É. and Allègre C. J. (2000) The influence of rivers on marine boron isotopes and implications for reconstructing past ocean pH. *Nature* **408**, 951–954.

Lisiecki L. E. and Raymo M. E. (2005) A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. *Paleoceanography* **20**.

Lønborg C., álvarez-Salgado X. A., Letscher R. T. and Hansell D. A. (2018) Large stimulation of recalcitrant dissolved organic carbon degradation by increasing ocean temperatures. *Front. Mar. Sci.* **4**, 1–11.

Malinverno A. and Martinez E. A. (2015) The effect of temperature on organic carbon degradation in marine sediments. *Sci. Rep.* **5**, 17861.

Martínez-Botí M. a., Foster G. L., Chalk T. B., Rohling E. J., Sexton P. F., Lunt D. J., Pancost R. D., Badger M. P. S. and Schmidt D. N. (2015) Plio-Pleistocene climate sensitivity evaluated using high-resolution CO2 records. *Nature* **518**, 49–54.

McClelland H. L. O., Barbarin N., Beaufort L., Hermoso M., Ferretti P., Greaves M. and Rickaby R. E. M. (2016) Calcification response of a key phytoplankton family to millennial-scale environmental change. *Sci. Rep.* **6**, 1–11.

Müller P. J., Kirst G., Ruhland G., von Storch I. and Rosell-Melé A. (1998) Calibration of the alkenone paleotemperature index U37K′ based on core-tops from the eastern South Atlantic and the global ocean (60°N-60°S). *Geochim. Cosmochim. Acta* **62**, 1757–1772.

Nathan S. a. and Leckie R. M. (2009) Early history of the Western Pacific Warm Pool during the middle to late Miocene (~13.2–5.8 Ma): Role of sea-level change and implications for equatorial circulation. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **274**, 140–159.

Pälike H., Lyle M. W., Nishi H., Raffi I., Ridgwell A. J., Gamage K., Klaus A., Acton G., Anderson L., Backman J., Baldauf J., Beltran C., Bohaty S. M., Bown P., Busch W., Channell J. E. T., Chun C. O. J., Delaney M., Dewangan P., Dunkley Jones T., Edgar K. M., Evans H., Fitch P., Foster G. L., Gussone N., Hasegawa H., Hathorne E. C., Hayashi H., Herrle J. O., Holbourn A., Hovan S., Hyeong K., Iijima K., Ito T., Kamikuri S., Kimoto K., Kuroda J., Leon-Rodriguez L., Malinverno A., Moore Jr T. C., Murphy B. H., Murphy D. P., Nakamura H., Ogane K., Ohneiser C., Richter C., Robinson R., Rohling E. J., Romero O., Sawada K., Scher H., Schneider L., Sluijs A., Takata H., Tian J., Tsujimoto A., Wade B. S., Westerhold T., Wilkens R., Williams T., Wilson P. a., Yamamoto Y., Yamamoto S., Yamazaki T. and Zeebe R. E. (2012) A Cenozoic record of the equatorial Pacific carbonate compensation depth. *Nature* **488**, 609–614.

Prentice M. L., Freiz J. K., Simonds G. G. and Matthews R. K. (1993) Neogene Trends in Planktonic Foraminifer 18O from Site 807: Implications for Global Ice Volume and Western Equatorial Pacific Sea-Surface Temperatures. In *Proceedings of the Ocean Drilling Program, 130 Scientific Results* Ocean Drilling Program. pp. 281–305.

Raitzsch M. and Hönisch B. (2013) Cenozoic boron isotope variations in benthic foraminifers. *Geology* **41**, 591–594.

Ridgwell A. J. (2005) A Mid Mesozoic Revolution in the regulation of ocean chemistry. *Mar. Geol.* **217**, 339–357.

Sanyal A., Bijma J., Spero H. and Lea D. W. (2001) Empirical relationship between pH and the boron isotopic composition of Globigerinoides sacculifer: Implications for the boron isotope paleo-pH proxy. *Paleoceanography* **16**, 515–519.

Seki O., Foster G. L., Schmidt D. N., Mackensen A., Kawamura K. and Pancost R. D. (2010) Alkenone and boron-based Pliocene pCO2 records. *Earth Planet. Sci. Lett.* **292**, 201–211.

Shevenell A. E., Kennett J. P. and Lea D. W. (2008) Middle Miocene ice sheet dynamics, deep-sea temperatures, and carbon cycling: A Southern Ocean perspective. *Geochemistry, Geophys. Geosystems* **9**.

Shipboard Scientific Party (1991a) Site 806. In *Proceedings of the Ocean Drilling Program, 130 Initial Reports* Ocean Drilling Program. pp. 301–343.

Shipboard Scientific Party (1991b) Site 807. In *Proceedings of the Ocean Drilling Program, 130 Initial Reports* (eds. L. W. Kroenke, W. H. Berger, and T. R. Janecek). Ocean Drilling Program, College Station, TX. pp. 369–493.

Si W. and Rosenthal Y. (2019) Reduced continental weathering and marine calcification linked to late Neogene decline in atmospheric CO2. *Nat. Geosci.* **12**, 833–838.

Sosdian S. M., Greenop R., Hain M. P., Foster G. L., Pearson P. N. and Lear C. H. (2018) Constraining the evolution of Neogene ocean carbonate chemistry using the boron isotope pH proxy. *Earth Planet. Sci. Lett.* **498**, 362–376.

Spivack A. . and Edmond J. . (1987) Boron isotope exchange between seawater and the oceanic crust. *Geochim. Cosmochim. Acta* **51**, 1033–1043.

Suchéras-Marx B. and Henderiks J. (2014) Downsizing the pelagic carbonate factory: Impacts of calcareous nannoplankton evolution on carbonate burial over the past 17 million years. *Glob. Planet. Change* **123**, 97–109.

Tyrrell T. and Zeebe R. E. (2004) History of carbonate ion concentration over the last 100 million years. *Geochim. Cosmochim. Acta* **68**, 3521–3530.

Wade B. S., Pearson P. N., Berggren W. A. and Pälike H. (2011) Review and revision of Cenozoic tropical planktonic foraminiferal biostratigraphy and calibration to the geomagnetic polarity and astronomical time scale. *Earth-Science Rev.* **104**, 111–142.

Wara M. W., Ravelo A. C. and Delaney M. L. (2005) Permanent El Nino-Like Conditions During the Pliocene Warm Period. *Science (80-.).* **309**, 758–761.

Zeebe R. E. and Tyrrell T. (2019) History of carbonate ion concentration over the last 100 million years II: Revised calculations and new data. *Geochim. Cosmochim. Acta* **257**, 373–392.

Zeller M., McGregor S. and Spence P. (2019) Hemispheric Asymmetry of the Pacific Shallow Meridional Overturning Circulation. *J. Geophys. Res. Ocean.* **124**, 5765–5786.

Zhang J., Wang P., Li Q., Cheng X., Jin H. and Zhang S. (2007) Western equatorial Pacific productivity and carbonate dissolution over the last 550 kyr: Foraminiferal and nannofossil evidence from ODP Hole 807A. *Mar. Micropaleontol.* **64**, 121–140.

Zhang Y. G., Pagani M. and Liu Z. (2014) A 12-million-year temperature history of the tropical Pacific Ocean. *Science* **344**, 84–7.

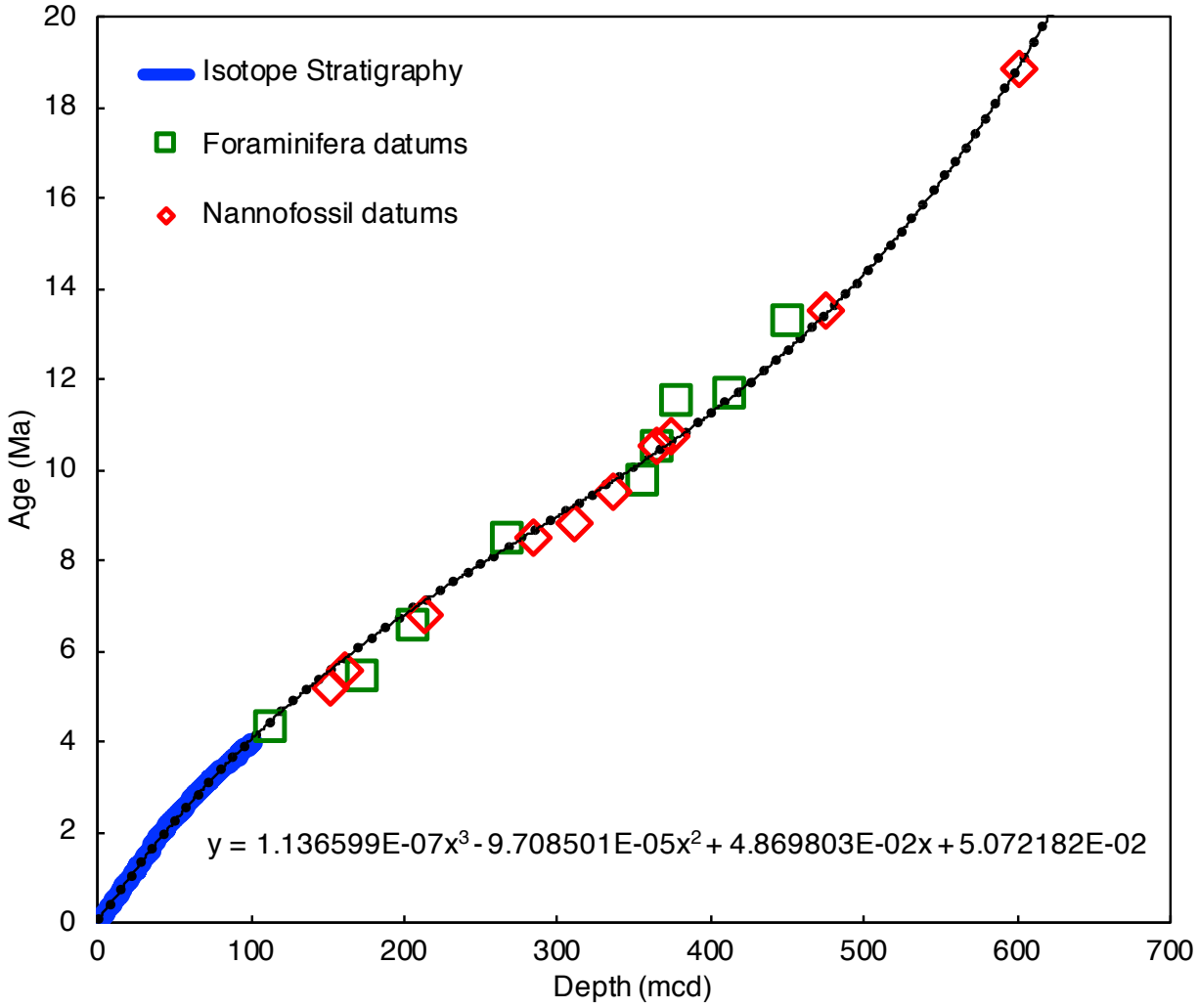

Figure S1. Age model for ODP Site 806. We use the isotope stratigraphy of Berger et al. (1993) which is included in the Lisiecki and Raymo (2005) stack for the past 4 Ma. We use a combination of foraminifera and nannofossils biostratigraphy for 6 to 20 Ma (see Table S1). We fit a third order polynomial to the data to estimate age from depth.

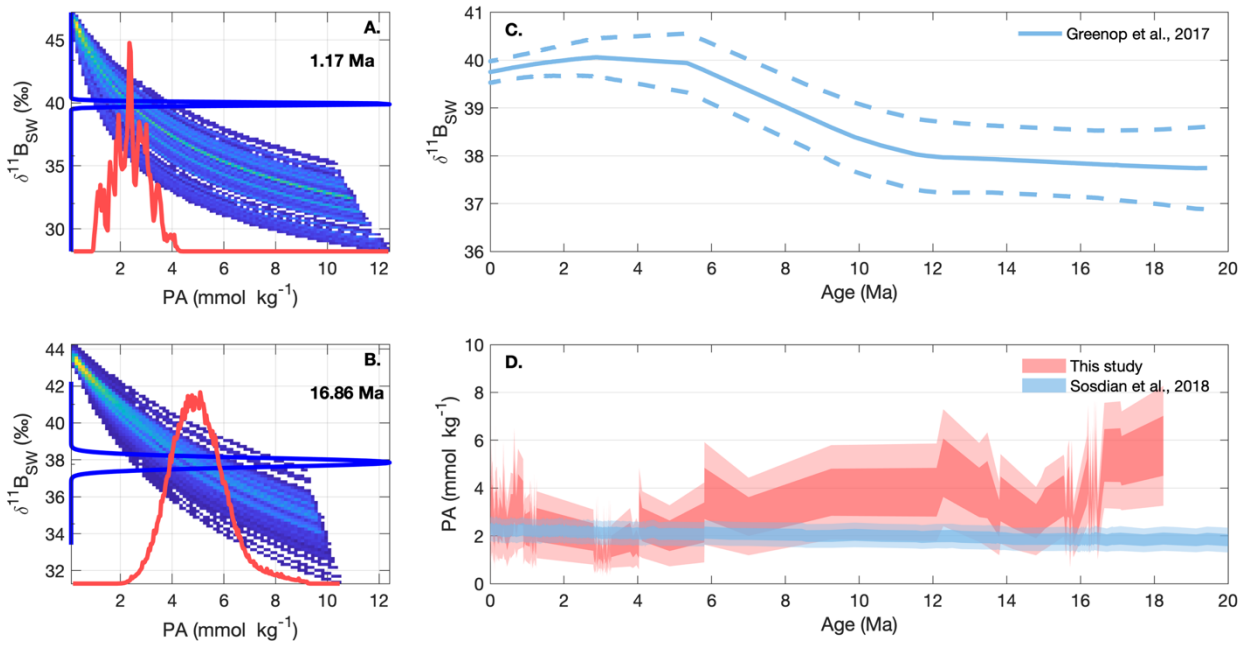

Figure S2. Calculated history of surface ocean alkalinity. (A, B) Example time slices of the system of solutions from paired $\delta^{11}B_{foram}$ and CO_2 values. (C) is the $\delta 11B_{SW}$ history that we use to determine the allowable range of carbonate chemistry. Solid line is the mean and the dashed lines are $\pm 2\sigma$. (D) Calculated timeseries of surface ocean alkalinity through the Neogene. The approach for calculating alkalinity is identical to that for DIC, which is described in detail in Section 3 of this Appendix.

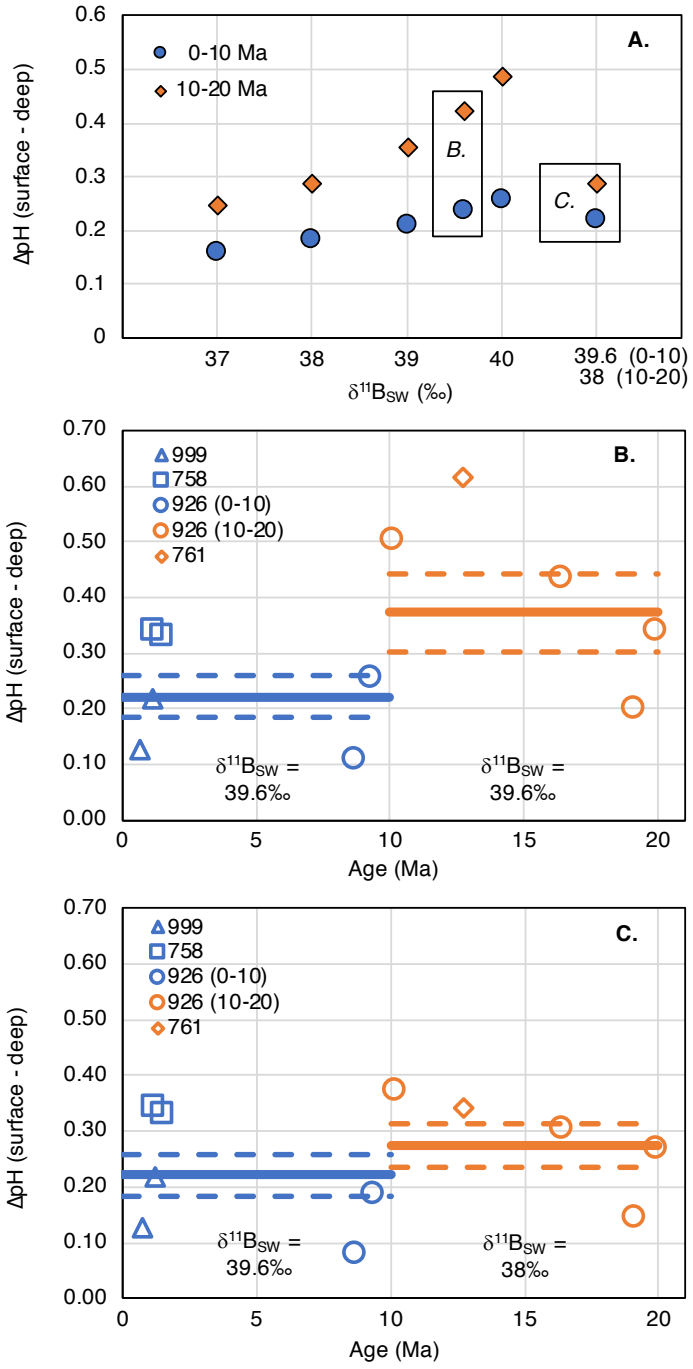

Figure S3. Estimated vertical pH gradient (ΔpH, surface-deep) through the Neogene. (A) average ΔpH with different $\delta^{11}B_{SW}$ assumptions (abscissa). We treat it as constant through the 0-20 Ma interval with values between 37‰ and 40‰, as well as one time-varying scenario (far right). (B) Example timeseries of data summarized in (A) assuming $\delta 11B_{SW}$ is constant at 39.6‰. (C) As in B, but assuming $\delta^{11}B_{SW}$ was 38‰ from 10-20Ma and 39.6‰ from 0-10 Ma. All scenarios we consider show a greater pH gradient in the Miocene compared to Pleistocene. Orange indicates 10-20 Ma samples, blue is 0-10 Ma, and different symbols mark different ODP sites.

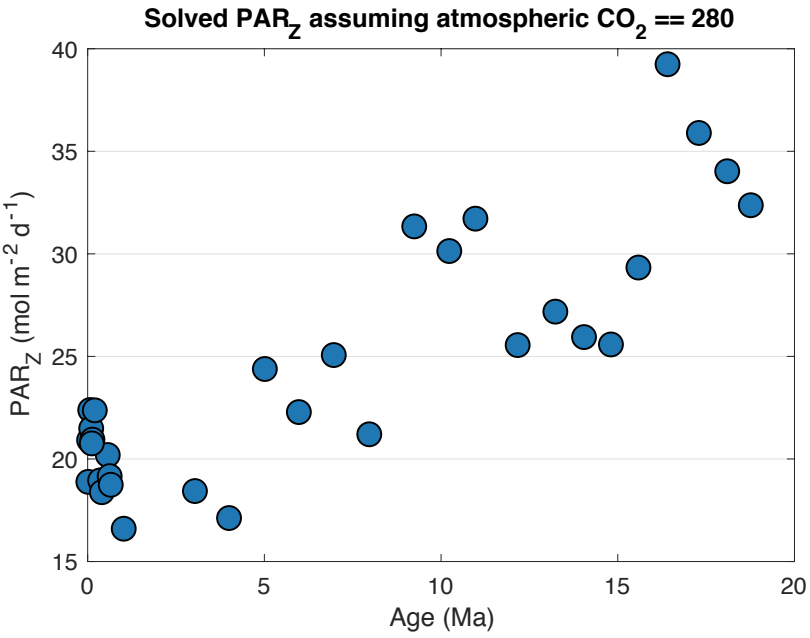

Figure S4. Sensitivity analysis of irradiance. Solved irradiance value assuming atmospheric pCO_2 is constant at 280 ppm for the past 20 Ma. Because the model is a multiple linear regression, the inferred change in CO_2 under constant irradiance is proportional to the inferred change in irradiance under constant CO_2, which makes this figure resemble our calculated CO_2 history.

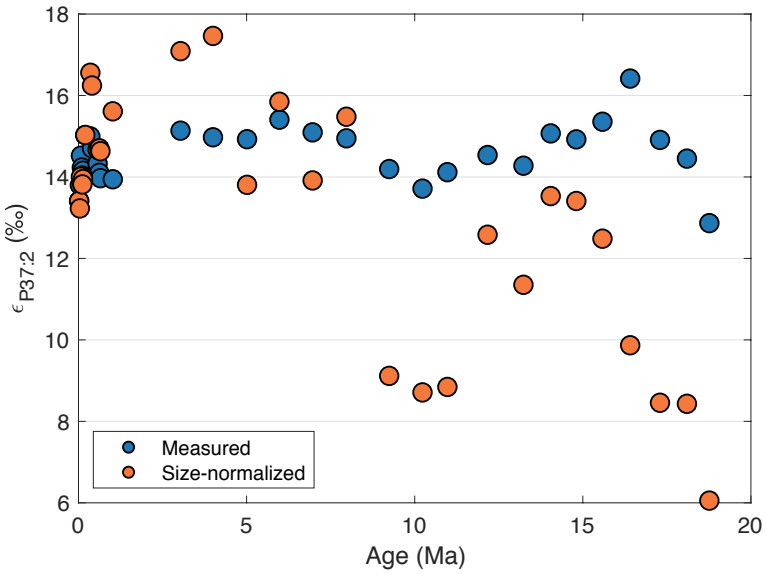

Figure S5. Effect of coccolith size on the measured alkenone ε_p at ODP 806. We calculate the change in cell radius in each sample relative to the youngest sample. Using this Δradius, we then calculate the effect of the cell radius change on $\varepsilon_{p37:2}$ using the "integrated irradiance" multiple linear regression model (Chapter 2). The size-normalized $\varepsilon_{p37:2}$ shows what we would expect to measure if there were no change in $[CO_{2(aq)}]$. The measured $\varepsilon_{p37:2}$ results from higher $[CO_{2(aq)}]$ in the Miocene relative to the Pleistocene.