Towards a Cenozoic History of Atmospheric CO_2

The Cenozoic CO_2 Proxy Integration Project (CenCO_2PIP) Consortium

Abstract: The geological record encodes the relationship between climate and atmospheric carbon dioxide (CO_2) over long and short timescales, as well as potential drivers of evolutionary transitions. However, reconstructing CO_2 beyond direct measurements requires the use of paleo-proxies and herein lies the challenge, as proxies differ in their assumptions, degree of understanding, and even reconstructed values. Here we critically evaluate, categorize, and integrate available proxies to create a high-fidelity and transparently constructed atmospheric CO_2 record spanning the past 66 million years. This provides clearer evidence for higher Earth System Sensitivity in the past and for the role of CO_2 thresholds in biological and cryosphere evolution.

The contribution of atmospheric CO_2 to Earth's greenhouse effect and the potential for variations in the global carbon cycle to cause climate change has been known for more than a century (*1*), but it was only in 1958 that direct measurements of the concentration of CO_2 in the atmosphere (or molar mixing ratio - the mole fraction of a gas in one mole of air) were systematically collected. Alongside reconstructions of the historical rise in Earth's surface temperature (*2*), this record has become one of the most influential and scientifically valuable environmental time-series, documenting the continuous rise in annual mean CO_2 from 315 parts per million (ppm) in 1958 to 419 ppm in 2022 (*3*). Projecting beyond these records to estimate how Earth's climate will respond to further increases in CO_2 requires global climate models (*4*). However, while successful in explaining observed historical climate change (*2*), models leave doubt as to whether global mean temperature will rise linearly as a function of future doubling of CO_2 (i.e., an invariant 'climate sensitivity') or whether climate feedbacks will lead to an increasing (or 'state-dependent') sensitivity of climate to CO_2 in the future (*5, 6*).

We can turn to the geological record to help constrain models and improve our understanding of non-linearities in the climate system (e.g., *7*), as it documents a variety of global climate changes and critically, climate states warmer than today. Leveraging this record, however, requires the paired quantification of both past atmospheric CO_2 and temperature. In parallel with recent efforts to compile and vet paleo-temperature estimates (*8*), here we focus on paleo-CO_2 estimates. Samples of ancient air can be extracted and analyzed from bubbles preserved in ancient polar ice (*9, 10*), but continuous ice core records currently only extend our knowledge of CO_2 back about 800 thousand years (kyr) (for a compilation, see *11*), with isolated time slices extending to ~2 Ma (million years ago) (*12, 13*). Importantly, at no point during the Pleistocene (2.58 Ma to 11,700 years ago) did CO_2 come close to present-day values (419 ppm, year 2022), with 300 ppm being the highest value measured to date (*14*). In contrast, depending on the extent of future human emissions, atmospheric CO_2 could reach 600–1000 ppm by the year 2100 (*2*). Feedbacks between changing climate and the carbon cycle may also amplify or diminish emissions from surficial carbon reservoirs (e.g., thawing permafrost, adjustments in size and composition of the terrestrial biosphere and marine carbon pool), creating additional uncertainty in future CO_2 projections (*15, 16*). Past changes in CO_2 inherently include the role of these feedbacks, and their study could help reduce uncertainty in Earth system models (*17*).

A solid understanding of atmospheric CO_2 variation through geological time is also essential to deciphering and learning from other features of Earth's history. Changes in atmospheric CO_2 and climate are suspected to have caused mass extinctions (e.g., *18, 19*) as well as evolutionary innovations (*20, 21*). During the Cenozoic, long-term declines in CO_2 and associated climate cooling have been proposed as the drivers of changing plant physiology (e.g., carbon-concentrating mechanisms), species competition and dominance, and associated with this, mammalian evolution. A more refined understanding of past trends in CO_2 is therefore central to understanding how modern species and ecosystems arose and may fare in the future.

Extending the CO_2 record beyond the temporally restricted availability of polar ice requires the use of 'proxies'. In essence, a CO_2 proxy could be any biological and/or geochemical property of a fossil or mineral that responds to the concentration of ambient CO_2 when it is formed. Unfortunately, unlike in the case of bubbles of ancient air trapped in polar ice, this response is invariably indirect. The connection between a proxy signal and atmospheric CO_2 is often strongly mediated via biological 'vital effects' (e.g., concentration of or discrimination against certain molecules, elements or isotopes due to physiological processes such as biomineralization, photosynthesis, respiration), may be indirectly connected to the atmosphere via dissolution of carbon in seawater or lakes, may involve isotopic or other chemical fractionation steps, or a combination of these. When preserved in terrestrial or marine sediments, proxy substrates can also be impacted by post-depositional ('diagenetic') processes that must be accounted for. Relationships between proxies and CO_2 are typically calibrated using observations or laboratory experiments; in biological systems, these calibrations are often limited to modern systems (e.g., modern organisms or soils), and applications to the distant past focus on physiologically or physically similar systems preserved in the sediment and rock record (e.g., similar fossil organisms or fossil soils). Most CO_2 proxies also require estimation of one or more additional environmental parameters and hence depend on additional proxy records. The complexity of proxy-enabled paleoclimate reconstructions thus presents a major challenge for creating a self-consistent estimate of atmospheric CO_2 through geological time and requires careful validation.

One of the first paleo-CO_2 proxies to be devised was based on the observation that vascular plants typically optimize the density, size, and opening/closing behavior of stomatal pores on their leaf surfaces to ensure sufficient CO_2 uptake while minimizing water loss (e.g., *22*). A count of stomatal frequencies then provides a simple proxy for the CO_2 concentration experienced by the plant (*23*). Changes in ambient CO_2 can also drive a cascade of interrelated effects on photosynthesis, the flux of CO_2 into the leaf (largely determined by stomatal size and density), and the carbon isotopic fractionation during photosynthesis ($\Delta^{13}C$, *22, 23, 24*). While lacking functional stomata, non-vascular plants like liverworts also exhibit isotopic fractionation during photosynthesis, and their $\delta^{13}C$ values are thus similarly controlled by ambient CO_2. The list of terrestrial paleo-CO_2 proxies also includes inorganic carbonate nodules precipitated in ancient soils (i.e., paleosols) as well as sodium carbonate minerals precipitated in continental lacustrine evaporites. While the paleosol proxy uses the carbon isotope composition of carbonate nodules and deconvolves the mixture of atmospheric and soil-respired CO_2 in soil porewaters using models of soil CO_2 (*25, 26*), the nahcolite proxy is based on the CO_2 dependence of sodium carbonate mineral equilibria (*27, 28*). Analogous to non-vascular plants on land, phytoplankton fractionate carbon isotopes during photosynthesis in response to the concentration of dissolved CO_2 in seawater, creating an isotopic signal stored in organic biomolecules that can be retrieved from ocean sediments (*29*). Boron proxies recorded in fossil shells of marine calcifying organisms are related to seawater pH, which in turn can be related back to atmospheric CO_2 (*30, 31*). A detailed discussion of the analytical details, entrained assumptions, and inherent uncertainties of currently available CO_2 proxies, plus summaries of recent advances and opportunities for further validation, is presented in the Supplemental Material and in Table S1.

Although each of these proxies has been validated extensively, comparing reconstructions from different proxies often reveals discrepancies. Prior compilations of paleo-CO_2 and explorations of the CO_2-climate linkage already exist (*32-34*), however, those studies applied limited proxy vetting, include CO_2 estimates that predate major innovations in some methods, and use rather basic data interpolation to assess broad CO_2 trends. Earlier CO_2 reconstructions are also often insufficiently constrained by ancillary data (e.g., concomitant temperature, isotopic composition of seawater or atmosphere) to be consistent with modern proxy theory, have incomplete or missing uncertainty estimates for CO_2 and/or sample age, and may exhibit fundamental disagreement with other proxies, leaving our current understanding of past CO_2 incomplete.

In this study we present the results of a 7-year endeavor by an international consortium of researchers whose collective expertise spans the reconstruction of paleo-CO_2 from all available terrestrial and marine archives. We have jointly created a detailed, open-source database of published paleo-CO_2 estimates including all raw and ancillary data together with associated analytical and computational methods. Each record was vetted and categorized in view of the most recent proxy understanding, with calculations adopting a common methodology including full propagation of uncertainties. We focus our efforts here on the Cenozoic, when the spatial distribution of continents and ocean basins, as well as the structure of marine and terrestrial ecosystems, was similar to the modern, yet profound changes in CO_2 and climate occurred. Identifying the most reliable Cenozoic CO_2 estimates published to date allows us to quantify important physical (e.g., temperature, ice volume) and biological (i.e., physiological, ecosystem) thresholds and tipping points.

We structure this investigation as follows: First we summarize the methodology by which we assessed the CO_2 proxies and associated estimates. We then apply these methods to derive a series of paleo CO_2 compilations comprised of data with different levels of quality or confidence, and statistically integrate the 'top-tier' data to create a realization of the Cenozoic variability in atmospheric CO_2. This is followed by a discussion of the climatic implications (including climate sensitivity) of the paleo-CO_2 curve, and a presentation of an evolutionary perspective. We finish with a roadmap for further advances in understanding past changes in atmospheric CO_2.

Critical assessment of atmospheric CO_2 proxies

The basis of our synthesis is a set of comprehensive data templates documenting all types of proxy data and their corresponding CO_2 estimates (a total of 6,247 data points). The completed data sheets for each study can be accessed as the paleo-CO_2 'Archive' in NOAA's National Climatic Data Center (NCDC). These 'Archive' sheets report all underlying data at face value from the original publications, but their unprecedented level of detail is designed to facilitate critical evaluation and recalculation of each CO_2 estimate.

From the 'Archive', published CO_2 estimates were evaluated by teams of experts who are active in validating and applying these proxies, and often included the original authors of the respective data. No new proxy data were collected as part of this effort, but estimates were recalculated where needed and possible, and age models were revised where new evidence was readily accessible. Additionally, CO_2 and age uncertainties were updated, as necessary, to consistently reflect propagated 95% confidence intervals. The vetting criteria are summarized in Supplementary Table S1 and detailed in paleo-CO_2 'Product' sheets. These CO_2 estimates are categorized as follows: 'Category 1' estimates (Fig. 1a, 1,673 data points or ~27% of the original total) are based on data whose uncertainty is fully documented and quantifiable in view of current proxy understanding. 'Category 2' estimates (Fig. 1b, 1,813 data points) contain sources of uncertainty that are not yet fully constrained. These uncertainties vary between proxies and datasets, and include, e.g., insufficient replication, poorly constrained proxy sensitivity to parameters other than CO_2, or extrapolation of calibration curves. 'Category 3' estimates (the residual 2,761 data points or ~44% of the Cenozoic paleo-CO_2 estimates published to date) are either superseded by newer, independently published evaluations from the same raw data, are considered unreliable due to factors such as incomplete supporting datasets that prevent full quantification of uncertainties, or outdated sample preparation methods.

Although objective criteria are applied throughout, the vetting process was particularly challenging for the paleosol- and phytoplankton-based proxies because multiple approaches are currently in use for interpreting these proxy data (*35-41*). Given the lack of a universally agreed-upon method, we compare multiple approaches for treating the data of these two proxies whenever possible. For the paleosol proxy, the greatest source of uncertainty is in the estimation of paleo-soil CO_2 concentration derived from respiration. Two different approaches are commonly used to do this. The first method is based on proxy-estimated mean annual rainfall, while the second is based on soil order (i.e., the most general hierarchical level in soil taxonomy, comparable to kingdom in the classification of biological organisms). However, few records in the database allow for a direct comparison between the two approaches. An opportunity for comparison exists with two Eocene records (*37, 42*), where re-calculation using each of the two different methods leads to CO_2 estimates that do not overlap within 95% confidence intervals for most stratigraphic levels (Fig. S6). This implies that the uncertainty in estimating paleo-soil CO_2 concentration derived from respiration cannot be fully quantified with either of these approaches. Thus, most paleosol-based CO_2 estimates were designated as Category 2. For the phytoplankton proxy, routinely applied methods differ in how algal cell size and growth rate are accounted for, as well as the assumed sensitivity of algal $\delta^{13}C$ values to aqueous CO_2 concentration (see Supplementary Materials for details). Where data are available, we compare both newer and traditional methods, finding that although there are deviations between the resulting CO_2 estimates, they do agree within 95% confidence intervals. We hence assign many phytoplankton CO_2 estimates to Category 1 and present mean CO_2 and uncertainty values that reflect the range of results from the different methods.

Towards a Cenozoic history of atmospheric CO_2

Our composite Category 1 and 2 realizations of Cenozoic CO_2 (Figs. 1a and b) display much better agreement among proxies than does the ‘raw’, un-curated collection ('Archive', Fig. 1c). Encouragingly, objective criteria applied to the original data products automatically placed the earlier-reported estimates of 'negative' CO_2, as well as some unusually high values, into Category 3, and without subjective intervention to otherwise filter them. We note that the Category 1 composite is now largely dominated by marine proxy estimates, with some intervals (e.g., the middle Paleocene, ~63-57 Ma) very sparsely sampled. Furthermore, some intervals (e.g., Oligocene, Miocene) still exhibit significant differences between proxies; for instance, marine-based CO_2 estimates start high and decline during the Oligocene (~34-23 Ma), whereas plant-based estimates suggest overall lower and constant CO_2 (Fig. 1a). Estimates of global temperature (Fig. 2b) during this time interval are largely invariant, which leaves us with the questions of whether CO_2 and climate were decoupled during this interval, or whether there is a systematic bias in the marine or plant-based CO_2 proxies and/or in the temperature proxies. All proxies become more uncertain further back in time as our knowledge of vital effects in biological proxy carriers, secular changes in the elemental and isotopic composition of ocean and atmosphere, as well as proxy sensitivity to environmental parameters that change along with CO_2 (e.g., temperature, rainfall, see Supplementary Materials for details) becomes less certain. In some cases, ancillary constraints and uncertainties are shared across multiple proxies (e.g., assumed atmospheric $\delta^{13}C$ is common to proxies based on land plant $\delta^{13}C$, leaf gas exchange, and paleosols), creating interdependence of estimates from seemingly independent proxies. More robust paleo-CO_2 reconstruction thus requires not only continued application of all proxies but also replication from different locations.

Although some uncertainties and proxy disagreements remain, the much-improved agreement within the vetted paleo-CO_2 compilation gives us confidence that a quantitative reconstruction of Cenozoic CO_2 based on the combined Category 1 data is possible. To do so, we statistically model mean CO_2 values at half-million-year intervals, together with uncertainties in age and proxy CO_2 estimates (Fig. 2a, see Supplementary Materials for details). Our choice of a 500-kyr resolution interval reflects a compromise driven by the proxy data compilation. Although parts of the Cenozoic, particularly the Plio-Pleistocene, are sampled at higher temporal resolution, the density of records remains relatively sparse throughout much of the Paleogene (1 datum per 190 kyr on average). As a result, the data (and in some cases the underlying age models) are not suited to interpreting higher-frequency (e.g., Milankovitch-scale) variations in atmospheric composition, and we focus here on low-frequency (e.g., multi-million year) trends and transitions. Proxy sampling within some intervals may be biased toward conditions that deviate from the 500-kyr mean (most notably here, the Paleocene-Eocene Thermal Maximum, PETM). We do not attempt to remove this bias but recommend caution in interpreting any features expressed at sub-million-year timescales.

This curve (Fig. 2a) allows us to constrain Cenozoic paleo-CO_2 and its uncertainty with greater confidence than earlier efforts. The highest CO_2 values of the past 66 Myr appear during the Early Eocene Climatic Optimum (EECO, ~53-51 Ma), while the lowest values occur during the Pleistocene. In contrast to earlier compilations, which suggested early Cenozoic CO_2 concentrations <400 ppm (e.g., *33*), rigorous data vetting and newly published records place early Paleocene mean CO_2 in our reconstruction between 650 and 850 ppm. However, the Paleocene remains data poor, and uncertainty in the curve remains large. Although the Paleocene record is predominantly based on the boron isotope proxy (Fig. 1a), inclusion of other (non-marine) proxy data does influence and refine the reconstruction through this epoch, supporting the value of the multi-proxy approach (Fig. S10). Following the rapid CO_2 rise and fall associated with the PETM at 56 Ma, mean CO_2 steadily rose to peak values of ~1600 ppm around 51 Ma during the EECO. The middle and late Eocene recorded slightly lower values (800-1100 ppm). Mean CO_2 dropped to <600 ppm across the Eocene-Oligocene transition (EOT, 33.9 Ma) and reached values that generally fall between ~400 and 200 ppm during the Miocene through Pleistocene, except for a notable increase during the Middle Miocene (~17-15 Ma) to a mean of ~500 ppm. Uncertainty in the mean CO_2 values drops substantially in the Plio-Pleistocene (see also Fig. S11), as expected given a dramatic increase in data density. Our analysis suggests that ~14.5-14 Ma was the last time 500-kyr-mean CO_2 value was as high as the present (Fig. S11), and that all Plio-Pleistocene peak interglacial CO_2 concentrations were exceptionally likely less than those of the modern atmosphere (Fig. S12). In contrast, prior to the Miocene, there is very little support (<2.5% probability) for Cenozoic 500-kyr-mean CO_2 values reaching or falling below pre-industrial levels.

Climatic implications of the revised CO_2 curve

Relationship with global temperature change and climate sensitivity

Our reconstructed Cenozoic CO_2 trends are broadly coherent with those for global temperature as inferred, e.g., from the oxygen isotopic composition ($\delta^{18}O$) of fossil benthic foraminifera shells (*43, 44*) and compilations of global surface temperature (*45*) (Fig. 2b). The Paleocene and Eocene epochs display overall higher temperatures and atmospheric CO_2 concentrations as compared to the later Oligocene, Miocene, and Pliocene - consistent with a predominantly greenhouse-gas regulated global energy budget. More specifically, the slow rise and subsequent fall of CO_2 over the course of the Paleocene and Eocene are mirrored by global temperatures, just as a transient Miocene CO_2 rise coincides with a period of warming at the Miocene Climatic Optimum (MCO). The EOT is identifiable in both the CO_2 and temperature records, despite the smoothing introduced by the curve fitting and 500-kyr binning interval.

Despite this overall agreement, rates and timing of CO_2 vs. temperature changes in the two records are not always synchronized (Fig. 2a,b). For example, CO_2 appears broadly static or even rising during the late Eocene (37-34 Ma) and late Miocene (11-5 Ma) despite global cooling (see also *46*) at these times. Conversely, decreasing CO_2 during the early Oligocene corresponds with relatively stable global temperatures (Fig. 2b, but see also *47, 48*) and ice volume (Fig. 2c) at that time. We note that the reconstructed Oligocene CO_2 decrease is driven by the contribution of marine proxies to the composite curve, whereas estimates from leaf gas exchange proxies are low and broadly static (Fig. 1c), a discrepancy that cannot be resolved without further experimentation and data collection. We caution that, even at the 500-kyr resolution of our study, the relative timing of CO_2 and temperature change might be unresolved in poorly sampled intervals (i.e., middle Paleocene), but should be well resolved during more recent, well sampled intervals (i.e., late Miocene through present, Fig. S8). Is the occasional divergence of temperature and CO_2 change evidence for occasional disconnects between CO_2 forcing and climate response? Although one might posit bias in the CO_2 reconstruction, the strength of our multiproxy approach is the reduced likelihood that multiple proxies exhibit common bias during particular periods of the Cenozoic. We suggest that some cases of divergence between temperature and CO_2 could reflect non-CO_2 effects on climate (e.g., changes in paleogeography affecting ocean circulation, albedo and heat transport, *49*), or the temperature reconstructions used herein could be biased by non-thermal influences (e.g., uncertain elemental and isotopic composition of paleo-seawater, physiological or pH effects on proxies, *48, 50*).

Our updated CO_2 curve, in conjunction with existing global temperature reconstructions, gives us the opportunity to reassess how climate sensitivity might have evolved through the

Cenozoic. The most commonly reported form of climate sensitivity is equilibrium climate sensitivity (ECS), which focuses on fast feedback processes (e.g., clouds, lapse rate, snow, sea ice) and is therefore best suited for predicting present-day warming (~3°C for a doubling of CO_2 above the pre-industrial condition, *2*). Because the average temporal resolution of our CO_2 database is coarser than 1000 years, we cannot estimate ECS directly. Instead, our data are most appropriate for interpreting an Earth System Sensitivity ($ESS_{[CO2]}$, following the taxonomy of *51*) – the combination of short-term climate responses to doubling CO_2 plus the effects of slower, geological feedback loops such as the growth and decay of continental ice sheets. We compare our reconstructed 500-kyr-mean CO_2 values with two different estimates of global surface temperature. We apply the same Bayesian inversion model used in the CO_2 reconstruction to derive 500-kyr-mean surface temperatures from the benthic foraminiferal $\delta^{18}O$ compilation of Ref. (*43*), which we convert to temperatures using the methodology of Ref. (*44*) (Fig. 2b). In addition, we pair a set of multiproxy global surface temperature estimates for eight Cenozoic time intervals (Fig. 2b, *45*) with posterior CO_2 estimates from time bins corresponding to each interval. The two temperature reconstructions are broadly similar, although the benthic record suggests relatively higher temperatures during the hothouse climate of the Paleocene and Eocene, whereas the multiproxy reconstruction is elevated relative to the benthic record during the Oligocene and Neogene.

The co-evolution of atmospheric CO_2 and global mean surface temperature (GMST) over the Cenozoic is shown in Fig. 3. Because CO_2 is on a log scale, the slopes of lines connecting two adjacent points in time reflect the average intervening $ESS_{[CO2]}$. Benthic $\delta^{18}O$-derived temperatures suggest early Paleocene warming occurs with a very high $ESS_{[CO2]}$ (>8°C per CO_2 doubling), although CO_2 uncertainties are large during this time interval. $ESS_{[CO2]}$ steadily declines towards the peak of Cenozoic warmth ~50 Ma, then steepens again to ~8°C per CO_2 doubling for much of the cooling through to the EOT at ~34 Ma. In contrast, the multiproxy global temperature record suggests a lower $ESS_{[CO2]}$ of ~5°C between the early Eocene and earliest Oligocene. During the Oligocene and early part of the Miocene, both temperature records imply a near-zero $ESS_{[CO2]}$, i.e., CO_2 values appear to decline with no appreciable global cooling. $ESS_{[CO2]}$ implied by both temperature reconstructions steepens again from the middle Miocene (~16 Ma) to present, averaging 8°C per CO_2 doubling over the past 10 Myr.

An alternative perspective on early Cenozoic climate forcing was introduced by Ref. (*44*), who hypothesized that all pre-Oligocene climate change was the response of direct and indirect CO_2 radiative forcing plus long-term change in solar output (i.e., constant albedo). Given this, they converted Paleocene and Eocene benthic $\delta^{18}O$-derived GMST to estimates of CO_2 change required to explain the temperature record. Our reconstruction offers a direct test of this hypothesis, and although it compares well with the $\delta^{18}O$ approach of Ref. (*44*) throughout much of the early Cenozoic, our curve suggests that the late Eocene decline in CO_2 was less severe than expected under the constant albedo assumption (Fig. S13). This result is consistent with a growing contribution of glacier and sea ice albedo effects (e.g., *52, 53*) and the opening of Southern Ocean gateways (e.g., *54*) to climate cooling preceding the Eocene-Oligocene boundary.

In summary, the Cenozoic compilation confirms a strong link between CO_2 and GMST across timescales from 500 kyr to tens of Myr, with $ESS_{[CO2]}$ generally within the range of 5-8°C – patterns consistent with most prior work (*32-34, 45, 51, 55-60*), and considerably higher than the present-day ECS of ~3°C. Both temperature reconstructions imply relatively high $ESS_{[CO2]}$ values during the last 10 Myr of the Cenozoic, when global ice volumes were highest. This agrees with expectations of an amplified $ESS_{[CO2]}$ due to the ice-albedo feedback (*61*). However, even during times with little-to-no ice (Paleocene to early Eocene), we find elevated values of $ESS_{[CO2]}$ (approaching or exceeding 5°C per CO_2 doubling). This implies that fast, non-ice feedbacks, such as clouds or non-CO_2 greenhouse gases (*60, 62-65*) were probably stronger in the early Paleogene than they are in the present-day climate system (see also *5*). The Oligocene to early Miocene is the most enigmatic interval, with an apparent decrease in CO_2 despite relatively stable temperature, implying near zero $ESS_{[CO2]}$. It should be noted that this is one interval where different CO_2 proxies disagree on CO_2 change (Fig. 1a), with relatively stable values from plants but a decline in values from alkenones. More work is needed to confirm these CO_2 and temperature findings, but if these estimates are correct, this could partly reflect transition from a climate state too cold to support the strong fast feedbacks (e.g., clouds) of the early Eocene (*5*), but not cold enough to generate strong ice-albedo feedback. Tectonic changes in the arrangement of continents and the opening of critical ocean gateways may also be confounding derivation of $ESS_{[CO2]}$ at that time (e.g., *49, 54*).

Relationship with the evolution of the cryosphere

Our composite CO_2 record also enables reexamination of the evolution of Earth's cryosphere (Fig. 2c) in relation to CO_2 radiative forcing. We use the sea level estimation of Ref. (*66*) for this comparison because it covers the entire Cenozoic and is somewhat independent of the benthic $\delta^{18}O$ stack (*43*) used for the GMST derivation in Fig. 2b and also of the more recent sea level reconstruction of Ref. (*67*). Although there are significant differences between the two sea level estimates, the main features discussed herein are broadly consistent between them. The establishment of a permanent, continent-wide Antarctic ice shield at the EOT (~34 Ma) comes at the end of a ~10-Myr period of generally slowly decreasing CO_2. There is evidence for isolated, unstable Antarctic glaciers at various points over the 10-Myr interval prior to the EOT (*50, 53, 66, 68*), which is consistent with the increasing paleogeographic isolation of Antarctica and Southern Ocean cooling (*54*), and CO_2 may have been sufficiently low to enable the repeated crossing of a glaciation threshold by periodic orbital forcing. Tectonic cooling of Antarctica would have progressively raised the CO_2 glaciation threshold, which has been modeled to be within 560-920 ppm (*69, 70*). Our composite CO_2 record allows us to further assess this glaciation threshold but requires determining the point during glacial inception when strong positive feedbacks (e.g., ice-albedo and ice sheet elevation) commenced and ice sheet growth accelerated (*71*). Using the sea level curve of Ref. (*66*), we determine this point as 33.75 ±0.25 Ma, where our composite CO_2 record suggests 719 $\frac{+180}{-152}$ ppm (95% CIs). Once established, the land-based Antarctic ice sheet likely persisted for the remainder of the Cenozoic, although significant retreat of land-based ice has been modeled (30-36 m sea level equivalent, *72*) and estimated from proxies (Fig. 2c) for the Miocene Climatic Optimum (MCO). 500-kyr-mean CO_2 values increased to ~500 ppm during the MCO (Figs. 2a, S10), and benthic foraminiferal $\delta^{18}O$ (Fig. 2b, *43*) and clumped isotopes (*50*) indicate warming. While the stability of the land-based Antarctic ice sheet depends on many factors in addition to CO_2-induced global warming (e.g., hysteresis (*73*), bed topography (*74*)), our composite record indicates that significant retreat of land-based ice did not occur below 441-480 ppm (2.5-50 percentiles), and some land-based ice may have persisted up to 563 ppm (97.5 percentile) during the MCO. Excepting the MCO, atmospheric CO_2 has remained below our current value of 419 ppm since the late Oligocene (Figs. 2a, S10), with relatively small sea-level variations (up to ~20m, Fig. 2c and *67*) being driven by orbitally-forced melting of the marine-based ice sheet (e.g., *72, 75*). Finally, at ~2.7 Ma, the transition to intensified northern hemisphere glaciation and orbitally-driven glacial cycles coincided with CO_2 values that began decreasing after a relative high during the Pliocene (Fig. 2a).

Evolutionary implications of the revised CO_2 curve

While geologic trends in terrestrial floral and faunal habitat ranges (e.g., *76, 77*) and diversity (e.g., *78, 79, 80*) are largely thought to be controlled by temperature and associated climate patterns, atmospheric CO_2 has been hypothesized to drive the evolution of biological carbon concentrating mechanisms and their subsequent diversification in terrestrial plants (CCMs, Fig. 2d, *81, 82*). Our realization of how atmospheric CO_2 has varied through the Cenozoic allows us to re-examine this hypothesis. The two primary CCMs in terrestrial plants are the crassulacean acid metabolism (CAM) and C_4 photosynthetic syndromes. CCMs in terrestrial C_4 and CAM plants confer competitive advantages over the ancestral C_3 pathway under higher growing season temperatures, low rainfall, and lower atmospheric CO_2. As a result, C_4 photosynthesis contributes about 23% of today's global terrestrial gross primary production (GPP, *83*).

Plant clades with the C_4 pathway first emerged in the early Oligocene (*84, 85*), yet did not expand to ecological significance until the late Miocene (i.e., <5% GPP before ~10 Ma, Fig. 2d, *86, 87, 88*). CAM plants (e.g., cacti, ice plants, agaves, and some orchids) underwent significant diversification events around the late Oligocene and late Miocene (*89-91*). Taken together, two general biological thresholds emerge based on our CO_2 record: (1) All known origins of C_4 plants occurred when atmospheric CO_2 was lower than ~550 ppm (i.e., after 32 Ma, Fig. 2a,d, *84*), which is in agreement with theoretical predictions (*92, 93*). (2) All major Cenozoic CAM diversification events coincided with intervals when CO_2 was lower than ~430 ppm (i.e., after 27 Ma, *89, 90*). Our record is thus consistent with decreasing atmospheric CO_2 (< 550 ppm) being a critical threshold for the Cenozoic origin, diversification, and expansion of C_4 and CAM plants within grasslands, arid habitats (such as deserts), and habits (such as epiphytes), and provides strong data support for previous hypotheses (*20, 84, 86, 88, 89, 92, 94, 95*). Importantly, following their origin in the early Oligocene, C_4 plants did not immediately proliferate. By ~24 to ~18 Ma, open habitat grasslands are evident on most continents (*96*), yet widespread dispersal of C_4 plants was delayed until the late Miocene, and without any apparent decline in CO_2 (Fig. 2d). Therefore, the rise of C_4 plants to their dominance in many tropical and subtropical ecosystems was likely driven (and maintained today) by other factors such as fire, seasonality of rainfall, and herbivory (i.e., grazing that keeps landscapes open) (*97, 98*). The temporal evolution of these factors warrants further study as we move towards a future where CO_2 may rise above the 550-ppm threshold that was key to the origin, taxonomic diversification, and spread of C_4 plants.

Terrestrial mammals evolved and adapted to the changing and more open floral ecosystems of the late Cenozoic (*99-101*), and are thus indirectly linked to the 550-ppm atmospheric CO_2 threshold discovered herein. In particular, dental wear patterns (such as the shape of the chewing surface of a tooth, i.e., mesowear) and tooth morphology, such as crown height, reflect an increasingly abrasive and tough diet (*102, 103*), and can be traced across many herbivore lineages during this period. For instance, mesowear in North American Equidae (horses and their ancestors, Fig. 2d) began to increase in the late Eocene, and steadily continued into the Quaternary. Similarly, equids evolved high-crowned (hypsodont) teeth in the Miocene (*103-105*), and their body size increased to accommodate higher intake of more abrasive, grassy vegetation (Fig. 2d).

Evolutionary trends are a little less clear in the ocean, because marine algal CCMs are ubiquitous and diverse in form (*106*) and are believed to have an ancient origin. Moreover, the large spatial and seasonal variance of dissolved CO_2 in the surface ocean (as compared to the relatively uniform seasonal and spatial concentration of CO_2 in the air) may somewhat decouple their evolution from geologic trends in atmospheric CO_2. Evidence exists that marine algae, and in particular the coccolithophores (i.e., the source of the alkenone biomarkers), express CCMs to greater extent when CO_2 is lower (e.g., *107, 108, 109*), with estimates of cellular carbon fluxes suggesting enhanced CCM activity in coccolithophores began ~7-5 Ma (*110*). However, our revised CO_2 curve displays mean atmospheric CO_2 broadly constant at 300-350 ppm since at least ~14 Ma (Figs. 2a, S10), suggesting that increased CCM activity may reflect other proximal triggers, perhaps involving changes in ocean circulation and nutrient supply.

Perspectives and opportunities for further advances

Our community-assessed composite CO_2 record and statistically modelled time-averaged CO_2 curve exhibit greater clarity in the Cenozoic evolution of CO_2 and its relationship with climate than was possible in previous compilations, and furthermore highlight the value of cross-disciplinary collaboration and community building. Generating a paleo-CO_2 record with even greater confidence requires targeted efforts using multiple proxies to fill in data gaps, higher resolution and replication from multiple locations, and novel approaches to resolve remaining differences between CO_2 proxy estimates. Specifically: although the number and diversity of paleo-CO_2 proxy records continues to grow, data remain relatively sparse during several key parts of the Cenozoic record (e.g., middle Paleocene, Oligocene). Moreover, records from the Paleocene and Eocene are dominated by estimates from the boron isotope proxy, increasing potential for bias. Targeted efforts are hence needed to expand the number and diversity of data through these intervals and to refine multi-proxy reconstructions. Secondly, despite substantial progress, there remains a lack of consensus regarding the identity and/or quantification of some of the factors underlying each of the proxy systems analyzed here. New experimental and calibration studies, particularly those that isolate and quantify specific mechanistic responses and/or their interactions, need to be undertaken in order to reduce potential biases and uncertainty for each method. For instance, the emerging fields of genomics, evolutionary and developmental biology, and proteomics provide exciting new opportunities for improving and understanding paleo-proxy systematics. Thirdly, and associated with improved experimental quantification, refining our theoretical and mechanistic understanding of how proxies are encoded will allow us to create explicit and self-consistent representations of the processes involved. The development of proxy system forward models provides a promising leap in this direction (e.g., *111*). Bayesian statistical methods can then enable the full suite of models and data to be integrated and constrain the range of environmental conditions, including atmospheric CO_2 and other variables that are consistent with the multiproxy data (*112, 113*). Finally, development of new proxies is also a realistic and desirable aim. For instance, while this study focuses on more established proxies, new proxies such as coccolith calcite stable isotopes (*114*) and mammalian bone and teeth oxygen-17 anomalies (*115*) show promising results for reconstructing paleo-CO_2, but perhaps require further validation before they can be assessed with confidence.

Proxies and proxy-based reconstructions of how atmospheric CO_2 has varied through deep time have improved immeasurably over the past few decades. While they will never allow us to reconstruct past CO_2 with the same fidelity as direct air measurement, our study shows how community-based consensus assessment, together with a critical reanalysis of proxy models and assumptions, can progressively move us towards a quantitative history of atmospheric CO_2 for geological time.

References and Notes

1. S. Arrhenius, *Philosophical Magazine* **41**, 237 (1896).
2. IPCC, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger *et al.*, Eds., (Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021).
3. P. Tans, R. Keeling, NOAA/ESRL, Ed. (NOAA/ESRL, www.esrl.noaa.gov/gmd/ccgg/trends/, 2023).
4. K. E. Taylor, R. J. Stouffer, G. A. Meehl, *Bulletin of the American Meteorological Society* **93**, 485 (2012/04/01, 2011).
5. R. Caballero, M. Huber, *Proceedings of the National Academy of Sciences USA* **110**, 14162 (August 5, 2013, 2013).
6. J. Zhu, C. J. Poulsen, *Geophysical Research Letters* **47**, e2020GL089143 (2020/09/28, 2020).
7. J. Zhu, C. J. Poulsen, B. L. Otto-Bliesner, *Nature Climate Change* **10**, 378 (2020/05/01, 2020).
8. E. J. Judd, J. E. Tierney, B. T. Huber, S. L. Wing, D. J. Lunt, H. L. Ford *et al.*, *Scientific Data* **9**, 753 (2022/12/06, 2022).
9. R. J. Delmas, J.-M. Ascencio, M. Legrand, *Nature* **284**, 155 (1980).
10. A. Neftel, H. Oeschger, J. Schwander, B. Stauffer, R. Zumbrunn, *Nature* **295**, 220 (1982).
11. B. Bereiter, S. Eggleston, J. Schmitt, C. Nehrbass-Ahles, T. F. Stocker, H. Fischer *et al.*, *Geophysical Research Letters* **42**, 542 (2015).
12. Y. Yan, M. L. Bender, E. J. Brook, H. M. Clifford, P. C. Kemeny, A. V. Kurbatov *et al.*, *Nature* **574**, 663 (2019/10/01, 2019).
13. J. A. Higgins, A. V. Kurbatov, N. E. Spaulding, E. Brook, D. S. Introne, L. M. Chimiak *et al.*, *Proceedings of the National Academy of Sciences* **112**, 6887 (June 2, 2015, 2015).
14. C. Nehrbass-Ahles, J. Shin, J. Schmitt, B. Bereiter, F. Joos, A. Schilt *et al.*, *Science* **369**, 1000 (2020).
15. C. Le Quéré, M. R. Raupach, J. G. Canadell, G. Marland, et al., *Nature Geosci* **2**, 831 (2009).
16. P. Friedlingstein, M. O'Sullivan, M. W. Jones, R. M. Andrew, L. Gregor, J. Hauck *et al.*, *Earth Syst. Sci. Data* **14**, 4811 (2022).
17. P. Valdes, *Nat. Geosci.* **4**, 414 (2011/07/01, 2011).
18. W. Kiessling, M. Aberhan, L. Villier, *Nature Geosci* **1**, 527 (2008).
19. J. L. Payne, A. M. Bush, N. A. Heim, M. L. Knope, D. J. McCauley, *Science* **353**, 1284 (2016).
20. E. J. Edwards, C. P. Osborne, C. A. Strömberg, S. A. Smith, W. J. Bond, P. A. Christin *et al.*, *Science* **328**, 587 (Apr 30, 2010).
21. F. A. McInerney, S. L. Wing, *Annual Review of Earth and Planetary Sciences* **39**, 489 (2011/05/30, 2011).
22. W. Konrad, D. L. Royer, P. J. Franks, A. Roth-Nebelsick, *Geological Journal*, (2020/03/18, 2020).

23. J. C. McElwain, M. Steinthorsdottir, *Plant Physiology* **174**, 650 (2017).
24. B. A. Schubert, A. H. Jahren, *Geochim. Cosmochim. Acta* **96**, 29 (2012).
25. T. E. Cerling, *Am. J. Sci.* **291**, 377 (1991).
26. D. O. Breecker, *Geochemistry, Geophysics, Geosystems* **14**, 3210 (2013/08/01, 2013).
27. E. A. Jagniecki, T. K. Lowenstein, D. M. Jenkins, R. V. Demicco, *Geology* **43**, 1075 (December 1, 2015, 2015).
28. T. K. Lowenstein, R. V. Demicco, *Science* **313**, 1928 (2006).
29. M. Pagani, in *Treatise on Geochemistry (Second Edition),* H. D. Holland, K. K. Turekian, Eds. (Elsevier, Oxford, 2014), pp. 361-378.
30. B. Hönisch, S. M. Eggins, L. L. Haynes, K. A. Allen, K. Holland, K. Lorbacher, *Boron proxies in Paleoceanography and Paleoclimatology*. Analytical Methods in Earth and Environmental Science (John Wiley & Sons, Ltd., 2019).
31. J. W. B. Rae, in *Boron Isotopes: The Fifth Element,* H. Marschall, G. Foster, Eds. (Springer International Publishing, Cham, 2018), pp. 107-143.
32. D. J. Beerling, D. L. Royer, *Nat. Geosci.* **4**, 418 (2011).
33. G. L. Foster, D. L. Royer, D. J. Lunt, *Nature Communications* **8**, 14845 (04/04/online, 2017).
34. J. W. B. Rae, Y. G. Zhang, X. Liu, G. L. Foster, H. M. Stoll, R. D. M. Whiteford, *Annual Review of Earth and Planetary Sciences* **49**, 609 (2021).
35. S. Ji, J. Nie, A. Lechler, K. W. Huntington, E. O. Heitmann, D. O. Breecker, *Earth and Planetary Science Letters* **499**, 134 (2018/10/01/, 2018).
36. J. Da, Y. G. Zhang, G. Li, X. Meng, J. Ji, *Nature Communications* **10**, 4342 (2019/09/25, 2019).
37. E. Hyland, N. D. Sheldon, M. Fan, *GSA Bulletin* **125**, 1338 (2013).
38. J. Henderiks, M. Pagani, *Paleoceanography* **22**, (2007).
39. S. R. Phelps, G. M. M. Hennon, S. T. Dyhrman, M. D. Hernández Limón, O. M. Williamson, P. J. Polissar, *Geochemistry, Geophysics, Geosystems* **22**, e2021GC009657 (2021/07/01, 2021).
40. H. M. Stoll, J. Guitian, I. Hernandez-Almeida, L. M. Mejia, S. Phelps, P. Polissar *et al.*, *Quaternary Science Reviews* **208**, 1 (2019/03/15/, 2019).
41. Y. G. Zhang, J. Henderiks, X. Liu, *Geochimica et Cosmochimica Acta* **281**, 118 (2020/07/15/, 2020).
42. E. G. Hyland, N. D. Sheldon, *Palaeogeography, Palaeoclimatology, Palaeoecology* **369**, 125 (2013).
43. T. Westerhold, N. Marwan, A. J. Drury, D. Liebrand, C. Agnini, E. Anagnostou *et al.*, *Science* **369**, 1383 (2020).
44. J. Hansen, M. Sato, G. Russell, P. Kharecha, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **371**, 20120294 (2013/10/28, 2013).
45. S. J. Ring, S. G. Mutz, T. A. Ehlers, *Paleoceanography and Paleoclimatology* **37**, e2021PA004364 (2022/12/01, 2022).
46. T. D. Herbert, K. T. Lawrence, A. Tzanova, L. C. Peterson, R. Caballero-Gill, C. S. Kelly, *Nature Geosci* **9**, 843 (11//print, 2016).

47. D. E. Gaskell, M. Huber, L. O’Brien Charlotte, N. Inglis Gordon, R. P. Acosta, J. Poulsen Christopher *et al.*, *Proceedings of the National Academy of Sciences* **119**, e2111332119 (2022/03/15, 2022).
48. C. L. O'Brien, M. Huber, E. Thomas, M. Pagani, J. R. Super, L. E. Elder *et al.*, *Proceedings of the National Academy of Sciences* **117**, 25302 (2020).
49. D. J. Lunt, A. Farnsworth, C. Loptson, G. L. Foster, P. Markwick, C. L. O'Brien *et al.*, *Clim. Past* **12**, 1181 (2016).
50. A. N. Meckler, P. F. Sexton, A. M. Piasecki, T. J. Leutert, J. Marquardt, M. Ziegler *et al.*, *Science* **377**, 86 (2022/07/01, 2022).
51. PALEOSENS-project-members, *Nature* **419**, 683 (2012).
52. A. Tripati, D. Darby, *Nature Communications* **9**, 1038 (2018/03/12, 2018).
53. H. D. Scher, S. M. Bohaty, B. W. Smith, G. H. Munn, *Paleoceanography* **29**, 628 (2014/06/01, 2014).
54. I. Sauermilch, J. M. Whittaker, A. Klocker, D. R. Munday, K. Hochmuth, P. K. Bijl *et al.*, *Nature Communications* **12**, 6465 (2021/11/09, 2021).
55. D. L. Royer, *Annual Review of Earth and Planetary Sciences* **44**, 277 (2016).
56. R. M. Brown, T. B. Chalk, A. J. Crocker, P. A. Wilson, G. L. Foster, *Nat. Geosci.* **15**, 664 (2022/08/01, 2022).
57. T. E. Wong, Y. Cui, D. L. Royer, K. Keller, *Nature Communications* **12**, 3173 (2021/05/26, 2021).
58. D. L. Royer, M. Pagani, D. J. Beerling, *Geobiology* **10**, 298 (2012/07/01, 2012).
59. J. E. Tierney, J. Zhu, M. Li, A. Ridgwell, G. J. Hakim, C. J. Poulsen *et al.*, *Proceedings of the National Academy of Sciences* **119**, e2205326119 (2022/10/18, 2022).
60. E. Anagnostou, E. H. John, T. L. Babila, P. F. Sexton, A. Ridgwell, D. J. Lunt *et al.*, *Nature Communications* **11**, 4436 (2020/09/07, 2020).
61. J. Hansen, M. Sato, P. Kharecha, D. Beerling, R. Berner, V. Masson-Delmotte *et al.*, *Open Atm. Sci. J.* **2**, 217 (2008).
62. J. T. Kiehl, C. A. Shields, *Philosophical Transactions of the Royal Society A* **371**, 20130093 (October 28, 2013, 2013).
63. J. Zhu, C. J. Poulsen, J. E. Tierney, *Science Advances* **5**, eaax1874 (2019).
64. D. J. Beerling, A. Fox, D. S. Stevenson, P. J. Valdes, *Proceedings of the National Academy of Sciences* **108**, 9770 (2011/06/14, 2011).
65. T. Schneider, C. M. Kaul, K. G. Pressel, *Nat. Geosci.* **12**, 163 (2019/03/01, 2019).
66. K. G. Miller, J. V. Browning, W. J. Schmelz, R. E. Kopp, G. S. Mountain, J. D. Wright, *Science Advances* **6**, eaaz1346 (2020).
67. E. J. Rohling, J. Yu, D. Heslop, G. L. Foster, B. Opdyke, A. P. Roberts, *Science Advances* **7**, eabf5326 (2021).
68. M. J. Henehan, K. M. Edgar, G. L. Foster, D. E. Penman, P. M. Hull, R. Greenop *et al.*, *Paleoceanography and Paleoclimatology* **35**, e2019PA003713 (2020/06/01, 2020).
69. E. Gasson, D. J. Lunt, R. DeConto, A. Goldner, M. Heinemann, M. Huber *et al.*, *Clim. Past* **10**, 451 (2014).
70. R. M. DeConto, D. Pollard, P. A. Wilson, H. Palike, C. H. Lear, M. Pagani, *Nature* **455**, 652 (2008).
71. R. M. DeConto, D. Pollard, *Nature* **421**, 245 (2003).

72. E. Gasson, R. M. DeConto, D. Pollard, R. H. Levy, *Proceedings of the National Academy of Sciences* **113**, 3459 (2016/03/29, 2016).
73. D. Pollard, R. M. DeConto, *Glob. Planet. Change* **45**, 9 (2005/02/01/, 2005).
74. G. J. G. Paxman, E. G. W. Gasson, S. S. R. Jamieson, M. J. Bentley, F. Ferraccioli, *Geophysical Research Letters* **47**, e2020GL090003 (2020/10/28, 2020).
75. J. J. Fürst, G. Durand, F. Gillet-Chaulet, L. Tavard, M. Rankl, M. Braun *et al.*, *Nature Climate Change* **6**, 479 (2016/05/01, 2016).
76. J. C. McElwain, *Annual Review of Plant Biology* **69**, 761 (2018/04/29, 2018).
77. S.-M. Popescu, J.-P. Suc, S. Fauquette, M. Bessedik, G. Jiménez-Moreno, C. Robin *et al.*, *Journal of Biogeography* **48**, 2771 (2021/11/01, 2021).
78. C. Jaramillo, M. J. Rueda, G. Mora, *Science* **311**, 1893 (Mar 31, 2006).
79. J. Y. Lim, H. Huang, A. Farnsworth, D. J. Lunt, W. J. Baker, R. J. Morley *et al.*, *Global Ecology and Biogeography* **31**, 425 (2022/03/01, 2022).
80. A. Graham, *Am J Bot* **98**, 336 (Mar, 2011).
81. J. R. Ehleringer, R. K. Monson, *Annual Review of Ecology and Systematics* **24**, 411 (1993).
82. H. Griffiths, in *Vascular plants as epiphytes,* U. Lüttge, Ed. (Springer, Berlin, 1989), pp. 42–86.
83. C. J. Still, J. A. Berry, G. J. Collatz, R. S. DeFries, *Glob. Biogeochem. Cycle* **17**, 6 (2003/03/01, 2003).
84. P.-A. Christin, G. Besnard, E. Samaritani, M. R. Duvall, T. R. Hodkinson, V. Savolainen *et al.*, *Current Biology* **18**, 37 (2008/01/08/, 2008).
85. A. Vicentini, J. C. Barber, S. S. Aliscioni, L. M. Giussani, E. A. Kellogg, *Glob. Change Biol.* **14**, 2963 (2008/12/01, 2008).
86. P. J. Polissar, C. Rose, K. T. Uno, S. R. Phelps, P. deMenocal, *Nat. Geosci.* **12**, 657 (2019/08/01, 2019).
87. L. Tauxe, S. J. Feakins, *Paleoceanography and Paleoclimatology* **35**, e2020PA003857 (2020/07/01, 2020).
88. T. E. Cerling, J. M. Harris, B. J. MacFadden, M. G. Leakey, J. Quade, V. Eisenmann *et al.*, *Nature* **389**, 153 (1997/09/01, 1997).
89. M. Arakaki, P. A. Christin, R. Nyffeler, A. Lendel, U. Eggli, R. M. Ogburn *et al.*, *Proc Natl Acad Sci U S A* **108**, 8379 (May 17, 2011).
90. T. J. Givnish, D. Spalink, M. Ames, S. P. Lyon, S. J. Hunter, A. Zuluaga *et al.*, *Proc. R. Soc. B.* **282**, (2015).
91. E. J. Edwards, *New Phytologist* **223**, 1742 (2019/09/01, 2019).
92. J. R. Ehleringer, T. E. Cerling, B. R. Helliker, *Oecologia* **112**, 285 (1997/10/01, 1997).
93. H. Zhou, B. R. Helliker, M. Huber, A. Dicks, E. Akçay, *Proceedings of the National Academy of Sciences* **115**, 12057 (2018/11/20, 2018).
94. C. E. R. Lehmann, S. A. Archibald, W. A. Hoffmann, W. J. Bond, *New Phytologist* **191**, 197 (2011/07/01, 2011).
95. W. M. Kürschner, Z. Kvacek, D. L. Dilcher, *Proceedings of the National Academy of Sciences USA* **105**, 449 (2008).
96. C. A. E. Strömberg, *Annual Review of Earth and Planetary Sciences* **39**, 517 (2011/05/30, 2011).

97. A. T. Karp, K. T. Uno, P. J. Polissar, K. H. Freeman, *Paleoceanography and Paleoclimatology* **36**, e2020PA004106 (2021/04/01, 2021).
98. T. Kukla, J. K. C. Rugenstein, D. E. Ibarra, M. J. Winnick, C. A. E. Strömberg, C. P. Chamberlain, *AGU Advances* **3**, e2021AV000566 (2022/04/01, 2022).
99. M. Fortelius, J. T. Eronen, F. Kaya, H. Tang, P. Raia, K. Puolamäki, *Annual Review of Earth and Planetary Sciences* **42**, 579 (2014/05/30, 2014).
100. F. Kaya, F. Bibi, I. Žliobaitė, J. T. Eronen, T. Hui, M. Fortelius, *Nature Ecology & Evolution* **2**, 241 (2018/02/01, 2018).
101. C. M. Janis, J. Damuth, J. M. Theodor, *Proceedings of the National Academy of Sciences* **97**, 7899 (2000).
102. M. Fortelius, N. Solounias, *American Museum Novitates* **3301**, 1 (2000).
103. M. C. Mihlbachler, F. Rivals, N. Solounias, G. M. Semprebon, *Science* **331**, 1178 (2011/03/04, 2011).
104. C. A. E. Strömberg, *Paleobiology* **32**, 236 (2006).
105. C. M. Janis, M. Fortelius, *Biol Rev Camb Philos Soc* **63**, 197 (May, 1988).
106. M. R. Badger, D. Hanson, G. D. Price, *Funct Plant Biol* **29**, 161 (Apr, 2002).
107. M. P. S. Badger, *Biogeosciences* **18**, 1149 (2021).
108. Y. G. Zhang, M. Pagani, Z. Liu, S. M. Bohaty, R. DeConto, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **371**, (October 28, 2013, 2013).
109. P. D. Tortell, *Limnology and Oceanography* **45**, 744 (2000/05/01, 2000).
110. C. T. Bolton, H. M. Stoll, *Nature* **500**, 558 (2013).
111. M. N. Evans, S. E. Tolwinski-Ward, D. M. Thompson, K. J. Anchukaitis, *Quaternary Science Reviews* **76**, 16 (2013/09/15/, 2013).
112. G. J. Bowen, B. Fischer-Femal, G. J. Reichart, A. Sluijs, C. H. Lear, *Clim. Past* **16**, 65 (2020).
113. M. B. Osman, J. E. Tierney, J. Zhu, R. Tardif, G. J. Hakim, J. King *et al.*, *Nature* **599**, 239 (2021/11/01, 2021).
114. C. Godbillot, F. Minoletti, F. Bassinot, M. Hermoso, *Clim. Past* **18**, 449 (2022).
115. A. Pack, A. Gehler, A. Süssenberger, *Geochimica et Cosmochimica Acta* **102**, 306 (2013/02/01/, 2013).
116. C. R. Scotese, *Annual Review of Earth and Planetary Sciences* **49**, 679 (2021/05/30, 2021).
117. P. A. Christin, C. P. Osborne, R. F. Sage, M. Arakaki, E. J. Edwards, *Journal of experimental Botany* **62**, 3171 (2011).

Acknowledgments: We thank Candace Major at NSF and the many researchers who contributed advice and enthusiasm for this project over the years. We are grateful to the staff at NCEI for facilitating curation of the paleo-CO_2 database, and to LDEO, the UCLA Lake Arrowhead Lodge and the Lady Bird Johnson Wildflower Center for hosting the three workshops that framed this work. Editor Jesse Smith and two anonymous reviewers provided valuable comments that improved this manuscript. This study is dedicated to Wally Broecker and Taro Takahashi, who pioneered the study of CO_2 dynamics in a changing climate.

Funding:

National Science Foundation grant OCE 16-36005 (BH, PJP)
Heising-Simons Foundation 2018-0996 (BH, VF)
National Science Foundation grant EAR 21-21649 (BH, VF, JM, CG)
National Science Foundation grant EAR 21-21170 (GJB)
National Science Foundation grant EAR 20-02370 (YC)
National Science Foundation grant 18-43285 (AP)
Columbia University's Center for Climate and Life (KTU)
G. Unger Vetlesen Foundation (KTU)
National Science Foundation grant 21-00537 (AP, PJP)
National Science Foundation grant 21-00509 (PJP)
National Science Foundation grant EAR 21-21165 (AR)
National Science Foundation grant EAR 18-06015 (YGZ)
National Science Foundation grant DGE 16-44869 (SRP)
National Science Foundation grant 18-13703 (EGH)
National Science Foundation grant OCE 16-58023 (JCZ)
National Science Foundation grant 16-02905 (MH)
Swedish Research Council grant NT7-2016 04905 (MS)
European Research Council grant 101020824 (JCM)
SFI/RC/2092 (JCM)
UK Research and Innovation grant 101045371 (MJH)
Natural Environment Research Council grant NE/X000567/1 (MPSB)
Royal Society grant DHF\R1\221014 (CRW)
Australian Research Council grant DP150104007 (PJF)
Deutsche Forschungsgemeinschaft grant RA 2068/4-1 (MR)
European Research Council grant 805246 (JWBR)
ETH Fellowship (JKCR)
National Science Foundation of China grant 42030503 (JD)
The Sandal Society Museum (GR)
Royal Society Tata Fellowship (BDAN)
Natural Environment Research Council grant NE/P019048/1 (GLF)

Author contributions:

Conceptualization: BH, GJB, DOB, MJH, MJK, AP, PJP, SRP, AR, DLR, MS, RSB, PJF, MH, MFS, JKCR, NDS, YGZ

Data Curation: BH, GJB, DOB, YC, MJH, MJK, JCM, PJP, DLR, MS, SRP, EA, MPSB, RSB, TBC, EdlV, KAD, CFG, MG, DTH, LLH, TKL, MG, JNM, MHS, RW, AR-N, NDS, SS, CRW, YGZ, JMC, DJ, DDE, GLF, EGH, BDAN, JBWR, MR, GJR, OS, LZ, CTB

Formal Analysis: BH, GJB, DOB, CY, MJH, PJP, DLR, AM, CEL

Investigation: BH, DLR, DOB, PJP, SRP, GJB, MJH, YC, MS, JCM, MJK, AP, EA, MPSB, RSB, PJF, WK, TKL, MFS, YGZ,

Software: GJB, VF, JJM, RW

Visualization: BH, GJB, MJH, MS, JCM, KTU, PJP, DOB, JA, BAK, CRS, JJM, RW

Funding acquisition: BH, PP, DLR, DOB, GJB, MJK, AP, RMdC, MH, KES, VF

Project administration: BH, DLR, DOB, PJP, GJB

Writing – original draft: BH, DLR, DOB, PJP, GJB, MJH, YC, MS, JCM, MJK, AP, KTU, AR, SRP

Writing – review & editing: all

Competing interests: None.

Data and materials availability: The completed data sheets for each study can be accessed as the paleo-CO_2 'Archive' at NOAA's National Center for Environmental Information (NCEI). The specific choice of category, as well as the updated CO_2 and age estimates, are documented in 'Product' sheets for each data set and proxy. In contrast to the 'Archive', which will grow with new publications but will otherwise remain passive, the paleo-CO_2 'Product' is a living database that will be updated when newly published data or ancillary data constraints become available, and/or methodological improvements are developed that enable modernization of previously underconstrained datasets. The 'Product' sheets created for this study can be accessed in NCEI, and this is also the place where future data updates will be made available in consecutive versions of the data 'Product'.

Authors composing the Consortium:

Bärbel Hönisch[1*], Dana L. Royer[2*], Daniel O. Breecker[3*], Pratigya J. Polissar[4*], Gabriel J. Bowen[5*], Michael J. Henehan[6], Ying Cui[7], Margret Steinthorsdottir[8], Jennifer C. McElwain[9], Matthew J. Kohn[10], Ann Pearson[11], Samuel R. Phelps[12], Kevin T. Uno[1], Andy Ridgwell[13*], Eleni Anagnostou[14], Jacqueline Austermann[1], Marcus P. S. Badger[15], Richard S. Barclay[16], Peter K. Bijl[17], Thomas B. Chalk[18], Christopher R. Scotese[19], Elwyn de la Vega[20], Robert M. DeConto[21], Kelsey A. Dyez[22], Vicki Ferrini[1], Peter J. Franks[23], Claudia F. Giulivi[1], Marcus Gutjahr[14], Dustin T. Harper[5], Laura L. Haynes[24], Matthew Huber[25], Kathryn E. Snell[26], Benjamin A. Keisling[27], Wilfried Konrad[28], Tim K. Lowenstein[29], Alberto Malinverno[1], Maxence Guillermic[30], Luz María Mejía[31], Joseph N. Milligan[16], John J. Morton[1], Lee Nordt[32], Ross Whiteford[33], Anita Roth-Nebelsick[34], Jeremy K. C. Rugenstein[35], Morgan F. Schaller[36], Nathan D. Sheldon[22], Sindia Sosdian[37], Elise B. Wilkes[38], Caitlyn R. Witkowski[6], Yi Ge Zhang[39], Lloyd Anderson[1], David J. Beerling[40], Clara Bolton[18], Thure E. Cerling[5], Jennifer M. Cotton[41], Jiawei Da[3], Douglas D. Ekart[42], Gavin L. Foster[43], David R. Greenwood[44], Ethan G. Hyland[45], Elliot A. Jagniecki[46], John P. Jasper[47], Jennifer B. Kowalczyk[48], Lutz Kunzmann[49], Wolfram M. Kürschner[50], Charles E. Lawrence[48], Caroline H. Lear[37], Miguel A. Martínez-Botí[51], Daniel P. Maxbauer[52], Paolo Montagna[53], B. David A. Naafs[6], James W. B. Rae[33], Markus Raitzsch[54], Gregory J. Retallack[55], Simon J. Ring[56], Osamu Seki[57], Julio Sepúlveda[26], Ashish Sinha[58], Tekie F. Tesfamichael[59], Aradhna Tripati[30], Johan van der Burgh[60], Jimin Yu[61], James C. Zachos[62], Laiming Zhang[63]

Affiliations:

*Corresponding authors. Email: hoenisch@ldeo.columbia.edu, droyer@wesleyan.edu, breecker@jsg.utexas.edu, gabe.bowen@utah.edu, polissar@ucsc.edu, andy@seao2.org

[1] Lamont-Doherty Earth Observatory of Columbia University, Palisades, United States
[2] Wesleyan University, Middletown, United States
[3] The University of Texas at Austin, Austin, United States
[4] University of California Santa Cruz, Santa Cruz, United States
[5] University of Utah, Salt Lake City, United States
[6] University of Bristol, Bristol, United Kingdom
[7] Montclair State University, Montclair, United States
[8] Swedish Museum of Natural History, Stockholm, Sweden
[9] Trinity College Dublin, Dublin, Ireland
[10] Boise State University, Boise, United States
[11] Harvard University, Cambridge, United States
[12] CIM Group, New York, United States
[13]University of California Riverside, Riverside, United States
[14]GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel, Germany
[15] The Open University, Milton Keynes, United Kingdom
[16] Smithsonian Institution, Washington, DC, United States
[17] Utrecht University, Utrecht, Netherlands
[18] Aix Marseille University, CNRS, IRD, INRAE, CEREGE, Aix-en-Provence, France
[19] Northwestern University, Evanston, United States
[20] University of Galway, Galway, Ireland
[21] University of Massachusetts, Amherst, United States
[22] University of Michigan, Ann Harbor, United States
[23] The University of Sydney, Sydney, Australia
[24] Vassar College, Poughkeepsie, United States
[56] Purdue University, West Lafayette, United States
[26] University of Colorado Boulder, Boulder, United States
[27] University of Texas Institute for Geophysics, Austin, United States
[28] University of Tübingen, Tübingen, Germany
[29] Binghamton University, Binghamton, United States
[30] University of California Los Angeles, Los Angeles, United States
[31] MARUM, University of Bremen, Bremen, Germany
[32] Baylor University, Waco, United States
[33] University of St Andrews, St Andrews, United Kingdom
[34] State Museum of Natural History, Stuttgart, Germany
[35] Colorado State University, Fort Collins, United States
[36] Rensselaer Polytechnic Institute, Troy, United States
[37] Cardiff University, Cardiff, United Kingdom
[38] Ginkgo Bioworks, Boston, United States
[39] Texas A&M University, College Station, United States
[40] University of Sheffield, Sheffield, United Kingdom
[41] California State University Northridge, Northridge, United States
[42] n/a, Salt Lake City, United States

[43] University of Southampton, Southampton, United Kingdom
[45] Department of Biology, Brandon University, Brandon, Canada
[45] North Carolina State University, Raleigh, United States
[46] Utah Geological Survey, Salt Lake City, United States
[47] Molecular Isotope Technologies, LLC, Niantic, United States
[48] Brown University, Providence, United States
[49] Senckenberg Natural History Collections, Dresden, Germany
[50] University of Oslo, Oslo, Norway
[51] EIT Urban Mobility, Barcelona, Spain
[52] Carleton College, Northfield, United States
[53] Institute of Polar Sciences - National Research Council, Bologna, Italy
[54] Dettmer Group KG, Bremen, Germany
[55] University of Oregon, Eugene, United States
[56] Deutsches GeoForschungsZentrum GFZ, Potsdsam, Germany
[57] Hokkaido University, Sapporo, Japan
[58] California State University Dominguez Hills, Carson, United States
[59] Addis Ababa University, Addis Ababa, Ethiopia
[60] n/a, Rossum, The Netherlands
[61] Laoshan Laboratory, Qingdao, China
[62] University of California Santa Cruz, Santa Cruz, United States
[63] China University of Geosciences, Beijing, China

List of Supplementary Materials

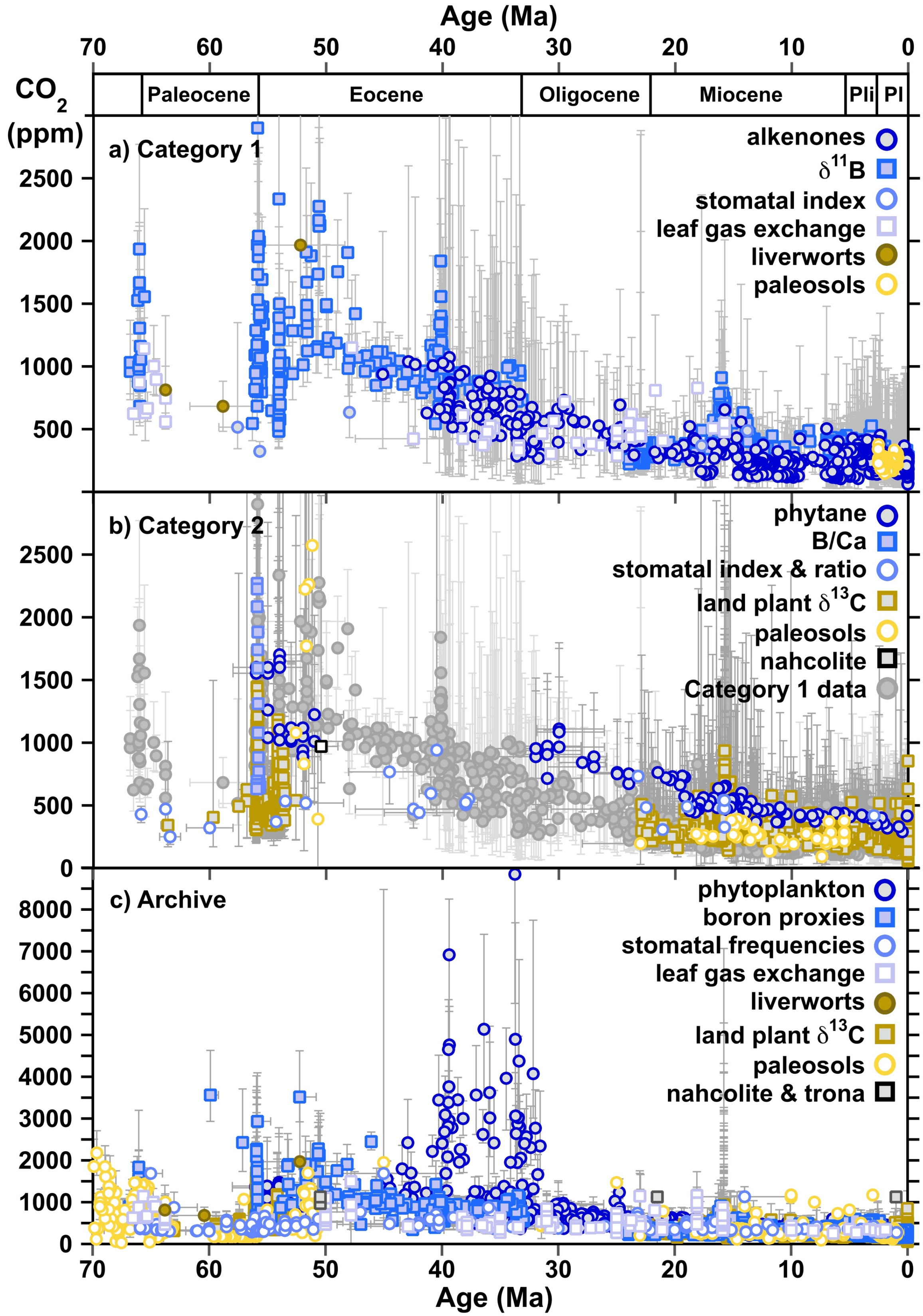
Supplementary Text, Sections 1-10, Figs. S1 to S13

Tables S1 to S3

References (*118-439*)

Figures :

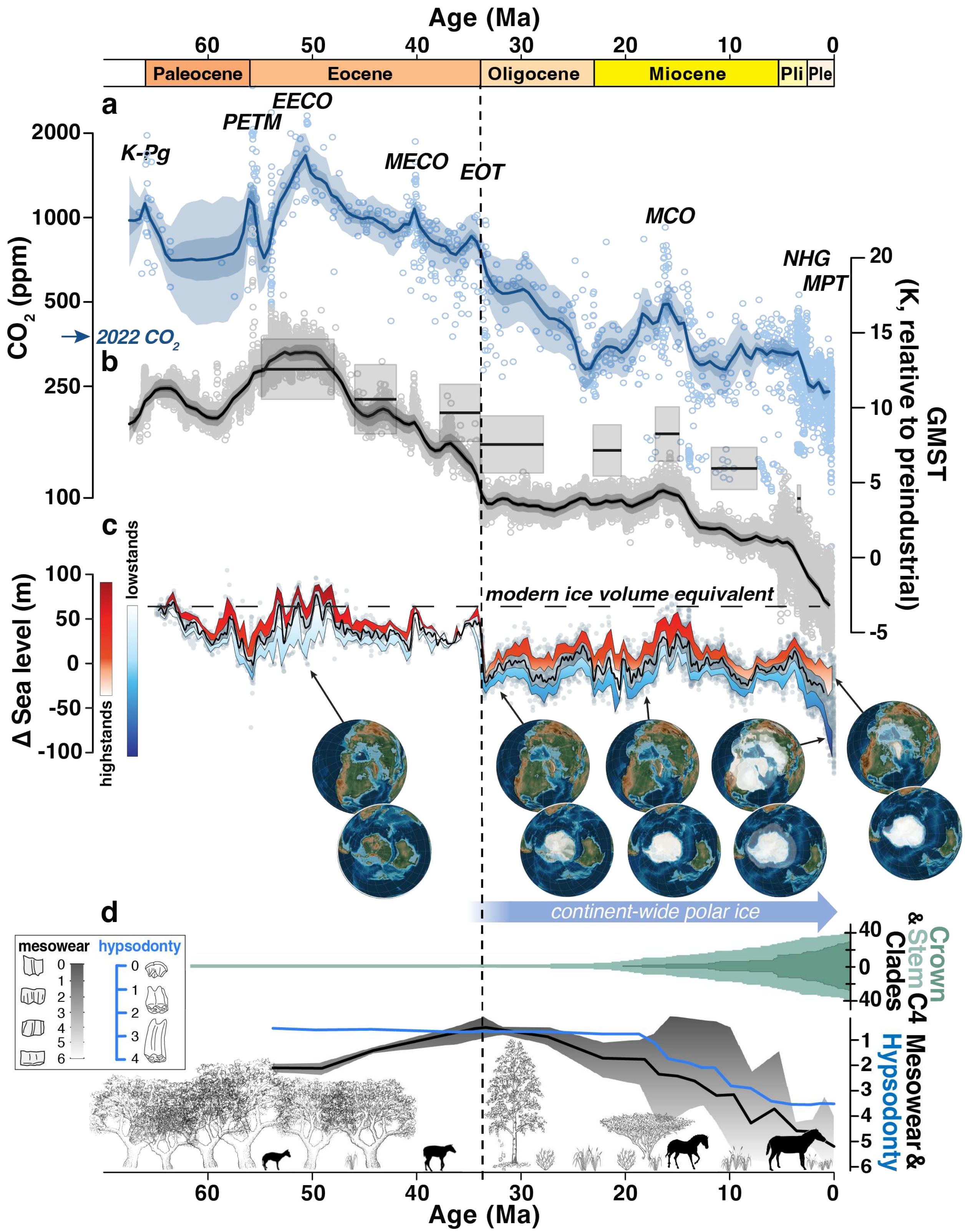
Fig. 1. Documentation and assessment of all Cenozoic paleo-CO_2 estimates published to date. Individual proxy estimates are defined by colored symbols in legends. **(a)** Vetted Category 1 estimates with their fully developed uncertainty estimates (95% CIs); age uncertainties have been updated or established to the best of current understanding. **(b)** Vetted Category 2 estimates whose uncertainty is not yet fully constrained. Category 1 data are shown in grey for reference. **(c)** Archive compilation of all CO_2 estimates in their originally published quantification. To toggle view of individual proxy records in panels (a) and (c), please go to paleo-co2.org (Note: panel (a) visualization will be published on the website after acceptance of the manuscript for publication).



Age (Ma)
70
60
50
40
30
20
10
0
CO2 (ppm)
Paleocene
Eocene
Oligocene
Miocene
Pli
Pl
a) Category 1
alkenones
δ11B
stomatal index
leaf gas exchange
liverworts
paleosols
2500
2000
1500
1000
500
b) Category 2
phytane
B/Ca
stomatal index & ratio
land plant δ13C
paleosols
nahcolite
Category 1 data
0
c) Archive
phytoplankton
boron proxies
stomatal frequencies
leaf gas exchange
liverworts
land plant δ13C
paleosols
nahcolite & trona
8000
7000
6000
5000
4000
3000

839

Fig. 2. Category 1 paleo-CO_2 record compared to global climate signals. The vertical dashed line indicates the onset of continent-wide glaciation in Antarctica. **(a)** Atmospheric CO_2 estimates (symbols) and 500-kyr mean statistical reconstructions (median and 50 and 95% credible intervals - dark and light-blue shading, respectively). Major climate events are highlighted (K-PG - Cretaceous/Paleogene boundary, PETM - Paleocene Eocene Thermal Maximum, EECO - Early Eocene Climatic Optimum, MECO - Middle Eocene Climatic Optimum, EOT - Eocene/Oligocene Transition, MCO - Miocene Climatic Optimum, NHG - onset of Northern Hemisphere Glaciation, MPT - Mid Pleistocene Transition). The 2022 annual average atmospheric CO_2 of 419 ppm is indicated for reference. **(b)** Global mean surface temperatures estimated from benthic $\delta^{18}O$ data after Westerhold et al. (*43*) (solid line, individual proxy estimates as symbols, and statistically reconstructed 500-kyr mean values shown as the continuous curve, with 50 and 95% credible intervals) and from surface temperature proxies (*45*) (grey boxes). **(c)** Sea level after Ref. (*66*) with gray dots displaying raw data; the solid black line reflects median sea level in a 1-Myr running window. High- and lowstands are defined within a running 400-kyr window, with lower and upper bounds of highstands defined by the 75th and 95th percentiles, and lower and upper bounds of lowstands defined by the 5th and 25th percentiles in each window. Globes depict select paleogeographic reconstructions and the growing presence of ice sheets in polar latitudes from Ref. (*116*). **(d)** Crown ages show C_4 clades, with CCMs adapted to low CO_2, initially diversified in the early Miocene and then rapidly radiated in the late Miocene (*117*). Flora transition from dominantly forested and woodland to open grassland habitats based on fossil phytolith abundance data (*96*). North American equids typify hoofed animal adaptations to new diet and environment (*103*), including increasing tooth mesowear (black line, note inverted scale), hypsodonty (blue line), and body size.

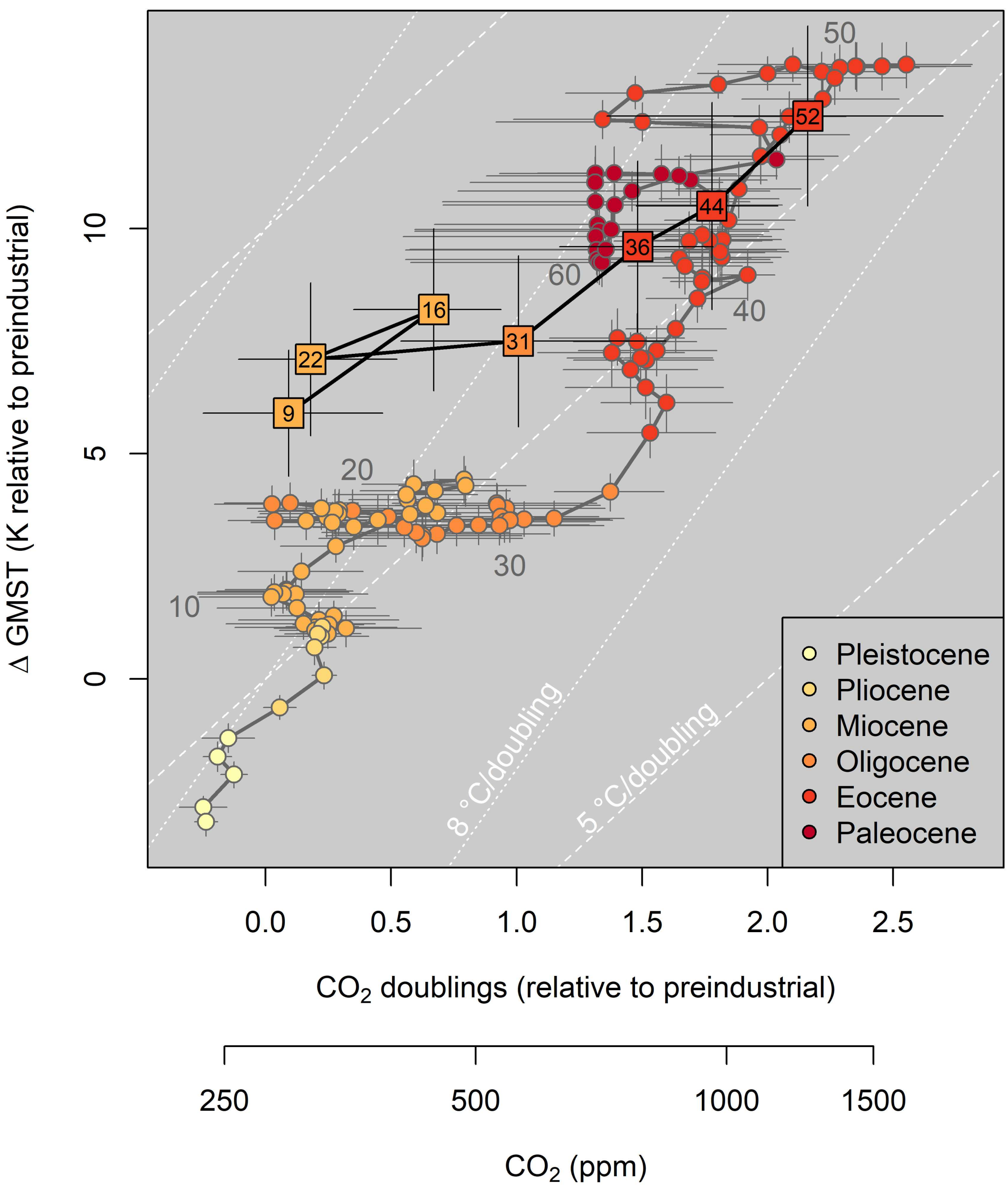


Age (Ma)
60
50
40
30
20
10
0
Paleocene
Eocene
Oligocene
Miocene
Pli
Ple
a
K-Pg
PETM
EECO
MECO
EOT
MCO
NHG
MPT
2000
1000
500
250
100
CO2 (ppm)
2022 CO2
b
GMST (K, relative to preindustrial)
20
15
10
5
0
-5
c
Δ Sea level (m)
100
50
0
-50
-100
lowstands
highstands
modern ice volume equivalent
continent-wide polar ice
d
mesowear
hypsodonty
40
0
40
Crown & Stem C4 Clades
Mesowear & Hypsodonty
1
2
3
4
5
6

865

866

Fig. 3. Application of the Category 1 CO_2 record to determine $ESS_{[CO2]}$. GMST deviation (K) from preindustrial global average surface temperature of 14.15°C is displayed versus paleo-CO_2 doublings relative to the preindustrial baseline of 280 ppm (upper x-axis) and paleo-CO_2 estimates on a log scale (lower x-axis). The slopes between two points in time reflect the average $ESS_{[CO2]}$. Circles reflect 500-kyr binned 'Category 1' CO_2 estimates paired with corresponding GMST-means from Ref. (*43*), squares pair CO_2 and GMST means from compilations of sea surface temperature (*45*) in seven coarsely resolved time intervals. Note that this figure omits the Pliocene temperature estimate of (*45*) because it samples too short a time interval (cf. Fig. 2) to be comparable with mean CO_2. Data from Cenozoic epochs are color coded and shift from red (Paleocene) to yellow (Pleistocene); labels indicate specific age bins (Ma). Dashed lines indicate reference $ESS_{[CO2]}$ lines of 8 and 5°C warming per doubling of CO_2.


Δ GMST (K relative to preindustrial)
CO₂ doublings (relative to preindustrial)
CO₂ (ppm)
0.0
0.5
1.0
1.5
2.0
2.5
250
500
1000
1500
-5
0
5
10
8 °C/doubling
5 °C/doubling
Pleistocene
Pliocene
Miocene
Oligocene
Eocene
Paleocene
10
20
30
40
50
60
9
16
22
31
36
44
52