Increased interglacial atmospheric CO_2 levels followed the mid-Pleistocene Transition

Masanobu Yamamoto[1,2*], Steven C. Clemens[3], Osamu Seki[2,4], Yuko Tsuchiya[2], Yongsong Huang[3], Ryouta O'ishi[5], Ayako Abe-Ouchi[5]

[1]Faculty of Environmental Earth Science, Hokkaido University, Sapporo 060-0810, Japan

[2]Graduate School of Environmental Science, Hokkaido University, Sapporo 060-0810, Japan

[3]Department of Earth, Environmental and Planetary Sciences, Brown University, Box 1864, Providence, RI 02912, U.S.A.

[4]Insititute of Low Temperature Science, Hokkaido University, Sapporo 060-0819, Japan

[5]Atmosphere and Ocean Research Institute, The University of Tokyo, Kashiwa 277-8568, Japan

*Corresponding author. E-mail: myama@ees.hokudai.ac.jp

Atmospheric CO_2 and polar ice volume have been strongly coupled over the past 805,000 years. However, the prior extent of coupling, during times of lower-amplitude ice-volume variability, is unknown because continuous high-resolution CO_2 records are lacking. We reconstructed the past 1,460,000 years of atmospheric CO_2 (~1,700-year sample resolution) by taking advantage of the unique relationship between CO_2 concentration and leaf wax $\delta^{13}C$ resulting from changes in the extent

of C_3 and C_4 vegetation in East India. Notably, reconstructed interglacial CO_2 concentrations were lower before the transition to large volume variability during the mid-Pleistocene transition (MPT; 900,000 years ago). Prior to the MPT, CO_2 had a secular trend similar to that of deep-ocean carbon isotopes. At orbital time scales phase analysis indicates that the CO_2 lead relative to ice volume changed to a lag during the MPT. Combined, these findings suggest that deep-ocean circulation controlled the long-term CO_2 trend, and that interaction between CO_2, continental ice, and deep-ocean circulation was reorganized during the MPT, involving a decrease in the carbon storage in the deep Pacific.

Earth's atmospheric carbon dioxide (CO_2) concentration is now > 415 ppm, far exceeding the highest interglacial values (299 ppm) of the past 805,000 years (805 kyrs) as documented in Antarctic ice cores[1–4] (Fig. 1a). Before the past 805 kyrs, marine sediment foraminifera $\delta^{11}B$, alkenone $\delta^{13}C$, terrestrial pedogenic carbonate $\delta^{13}C$, and C_3 plant $\delta^{13}C$ proxies have been used to reconstruct CO_2[5–12]. Still, they do not yield long continuous records sufficient to resolve orbital-scale cycles, except for some specific periods[6,7,12]. Blue ice yielded spot data at ~1 million years ago (Ma)[13], 1.5 Ma, and 2 Ma[14]: reported values range from 214 to 279 ppm. Compared to blue ice measurements, other proxies indicate far larger ranges over the same interval of time[5–11]. There are significant disagreements among existing proxies, highlighting the need for more efforts to CO_2 reconstruction.

From 1,200 to 700 thousand years ago (ka), the dominant 41 kyr cycles of benthic marine carbonate $\delta^{18}O$ transitioned to dominant quasi-100-kyr cycles, accompanied by increased amplitude and a saw-tooth pattern (rapid deglaciations followed by slower

glacial build-up). This transition is called the Mid-Pleistocene Transition (MPT)[15]. Numerous hypotheses have been proposed to explain the cause of the MPT, such as a gradual lowering of CO_2 (ref.[16]), increase in ice sheet stability[17–19], their combination[7,20], an abrupt decrease in CO_2 around 900 ka[6,21], phase locking of bipolar ice sheets[22], changes in self-sustained oscillation of global carbon cycles[23], and extremely low southern hemisphere summer insolation at MIS 23 as the trigger[24]. The extended CO_2 reconstruction presented here is highly relevant to this discussion.

In this study, we analyze the $\delta^{13}C$ of leaf wax long-chain *n*-fatty acids ($\delta^{13}C_{FA}$) in sediments from International Ocean Discovery Program (IODP) Site U1446 (19.08°N, 85.73°E) on the Indian margin, Bay of Bengal (Extended Data Fig. 1); we generate a record of CO_2 spanning the last 1.46 million years (Myr), with ~1.7 kyr sample resolution. We discuss the use of $\delta^{13}C_{FA}$ as a CO_2 proxy (CO_2^{FA}), compare our results with those of other independent reconstructions, and discuss changes in the relationship among CO_2, continental ice, and deep-ocean circulation over the MPT.

Carbon isotopic ratio of long-chain n-fatty acids

The average $\delta^{13}C$ value of *n*-C_{26}, C_{28}, C_{30}, and C_{32} fatty acids ($\delta^{13}C_{FA}$), leaf wax derived from terrestrial higher plants, ranges from −30.3 to −20.3 ‰ VPDB (Fig. 1a; Data are partly presented[25]). The $\delta^{13}C_{FA}$ is higher during glacials and lower during interglacials (Fig. 1b). The variation is positively correlated with that of benthic foraminifer $\delta^{18}O$, a proxy of global ice volume and bottom-water temperature (r = 0.76 for 0–1460 ka). Of particular interest, variation of $\delta^{13}C_{FA}$ mirrors that of the Antarctic ice core CO_2 concentration over the past 805 kyrs (Fig. 1a). The $\delta^{13}C_{FA}$ in Site U1446 sediments ranges between the end-members of C_3 and C_4 plants (Fig. 2a and Methods), indicating that it

predominantly reflects the relative contributions of C_3 and C_4 plants from the sediment source area, *i.e.,* the Mahanadi River and adjacent coastal regions (Methods and Extended Data Fig. 1).

Effects of CO_2, temperature, and precipitation on $\delta^{13}C_{FA}$

The proportion of C_3 and C_4 vegetation is affected by three factors: CO_2 concentration, water availability, and air temperature[26]. C_3 plants fix carbon from CO_2 by photosynthesis using RuBisCO in mesophyll cells. At the same time, the RuBisCO fixes oxygen and loses carbon through photorespiration. In low CO_2 conditions, photosynthesis of C_3 plants is suppressed relative to photorespiration. C_4 plants convert CO_2 to C_4 diacids in mesophyll cells, transfer it to bundle-sheath cells, and fix reconverted CO_2 using RuBisCO. Because oxygen content is low in the bundle-sheath cells, RuBisCO can fix carbon efficiently. Thus, C_4 plants have an advantage over C_3 plants in low CO_2 environments[27]. Meta-analysis studies of many open-air CO_2 enrichment experiments for native plants indicate that CO_2 enrichment to 550–600 ppm enhances the production rate of C_3 plants relative to C_4 plants by ~25%[28]. To further test these inferences, a series of experiments were run using a vegetation model to assess the sensitivity of regional vegetation functional types to changes in CO_2 (Fig. 3a), showing that changes in CO_2 from 285 to 185 ppm increase the $C_4/(C_3 + C_4)$ vegetation coverage ratio by 0.15 in the sediment source area (Fig. 3b).

The $\delta^{13}C_{FA}$ variation during glacial−interglacial cycles reflected the physiological effects of CO_2 on the carbon isotope fractionation of C_3 plants (green curve in Fig. 2a)[29], in addition to the relative abundance of C_3 and C_4 vegetations. A CO_2 reduction from 280 to 190 ppm induced a change in isotopic fractionation in C_3 plants that resulted in an

increase of 2.2 ‰ in $\delta^{13}C_{FA}$. This implies that about one fourth of the $\delta^{13}C_{FA}$ variation was attributable to the CO_2 effect on C_3 plant $\delta^{13}C_{FA}$. The $C_4/(C_3 + C_4)$ plant ratio and the ice core CO_2 concentration showed good correlation ($r^2 = 0.71$; Fig. 2b), but this correlation was slightly weaker than the correlation between $\delta^{13}C_{FA}$ and the ice core CO_2 concentration ($r^2 = 0.78$; Fig. 2a). This suggests that the CO_2 effect on C_3 plant $\delta^{13}C_{FA}$ enhances the correlation between $\delta^{13}C_{FA}$ and ice core CO_2, although by only a small amount.

Precipitation is the second factor determining the C_3/C_4 vegetation, with C_4 plants having an advantage in drier conditions[26]. $\delta^{13}C_{FA}$ has a different structure compared to precipitation reconstructions based on regional river runoff into the Bay of Bengal (seawater $\delta^{18}O_{sw}$)[25] from the study site, and other Indian summer monsoon records (Extended Data Fig. 2). The $\delta^{18}O_{sw}$ variation shares 12% of the total $\delta^{13}C_{FA}$ variation deduced from $r^2 = 0.12$ in Extended Data Fig. 3. This relevance is minor compared to that of CO_2 concentration (78% of total $\delta^{13}C_{FA}$ variation; Extended Data Fig. 3). Results of nine different general circulation models (GCMs) (Extended Data Fig. 4) show that precipitation decreased by 9 ± 23% from interglacial to glacial conditions (Fig. 3d)[30–32]. The dynamic vegetation model shows that the 9% precipitation decrease at the 285 ppm CO_2 level increases the $C_4/(C_3 + C_4)$ vegetation coverage ratio by 0.02 in the sediment source area (Fig. 3b). This result demonstrates that the impact of precipitation on the C_3/C_4 vegetation structure (0.02) is much smaller than that of atmospheric CO_2 concentration (0.15) in the sediment source region over glacial–interglacial time scales.

Temperature is the third factor determining the C_3/C_4 vegetation, with C_4 plants having an advantage in warmer conditions[26]. However, Site U1446 data indicates that C_4 plants were more abundant ($\delta^{13}C_{FA}$ was higher) when sea surface temperature decreased by

~3°C during glacials (Extended Data Figs. 2 and 4); this is inconsistent with the physiology of C_4 plants as advantaged in warmer conditions. A higher CO_2 increases C_3 plants relative to C_4 plants according to the difference in physiological responses, but it also decreases C_3 plants relative to C_4 plants under warmer conditions induced by the greenhouse effect. Because these two factors shift $\delta^{13}C_{FA}$ in the opposite direction, the temperature effect works against the shift in $\delta^{13}C_{FA}$ caused by CO_2 change. The negative correlation between $\delta^{13}C_{FA}$ and the EPICA CO_2 concentration (Fig. 2a) indicates that the effect of CO_2 concentration on C_3/C_4 vegetation is much stronger than the effect of temperature in the sediment source region.

These considerations, and the strong correspondence of $\delta^{13}C_{FA}$ with ice core CO_2 concentration, indicate that CO_2 is the predominant factor determining $\delta^{13}C_{FA}$ in the sediment source area and, hence, in U1446 sediments. Our dynamic vegetation model demonstrates that the vegetation in the sediment source area (within the range of mean annual temperatures >18°C and annual precipitation of 1000 to 1500 mm) changed from predominantly C_4 during the last glacial maximum (LGM) to C_3 in the preindustrial (PI) period (Fig. 3c). A few other regions exhibit this same response (Extended Data Fig. 5), characterized by hot and seasonally dry (savanna) climates, and could serve as targets for replicating our CO_2 reconstruction.

$\delta^{13}C_{FA}$ as a proxy of CO_2

The $\delta^{13}C_{FA}$ in Site U1446 sediments show nearly identical variation relative to Antarctic ice core CO_2 (Fig. 1a). However, age model differences between the marine and ice core records are somewhat relevant to this comparison. In this study, the linear regression equation after the tuning is used to convert $\delta^{13}C_{FA}$ to CO_2 concentration

(Methods and Extended Data Figs. 6 and 7). The calibration error (root-mean-square error) of $\delta^{13}C_{FA}$-based CO_2 reconstruction (CO_2^{FA}) is 12 ppm, including an analytical error of 1 ppm. The upper limit of the CO_2^{FA} estimation is 325 ppm as C_3 plants attain 100%, while the lower limit is 151 ppm as C_3 plants attain 0% (Fig. 2a).

The calibration stability during the last 800 kyr showed that it remained essentially unchanged and within the range of the above calibration error (Extended Data Fig. 8). Thus, the calibration is likely to be site specific and may not be directly applicable to other sites with significantly different climate and environmental settings (Extended Data Fig. 5).

We should keep in mind that precipitation and temperature potentially affect the $\delta^{13}C_{FA}$ if the background condition changes significantly. Proxy records of temperature and precipitation from the study site show that the ranges of variations are not different before and after the MPT (Extended Data Fig. 2). The mineral, major, trace, and rare earth element compositions in Sites NGHP 18A and 19A near the study site were almost constant in Miocene to Pleistocene sediments, suggesting no significant change in sediment provenance[33,34]. These do not ideally guarantee but encourage us to apply the $\delta^{13}C_{FA}$ to CO_2 reconstruction prior to the MPT.

A comparison with blue ice values show that the $\delta^{13}C_{FA}$-based CO_2^{FA} around 1 Ma (~180 to ~262 ppm) overlap 18 of 23 values for ~1 Ma blue ice (222 to 277 ppm)[13], and the average value is 31 ppm lower than that of blue ice (Extended Data Fig. 9). The CO_2^{FA} values around 1.5 Ma (~190 to ~250 ppm) overlap 27 of the 33 values for ~1.5 Ma blue ice (214 to 279 ppm)[14], and the average value is 10 ppm lower than that of blue ice. Foraminifera $\delta^{11}B$-based CO_2 values show a wider range than that of ice core measurements[8]; the lower values overlap blue ice values and CO_2^{FA} (Extended Data Figs.

9 and 10). Despite the difference in their absolute concentrations, the *G. ruber* $\delta^{11}B$-based CO_2 variation is consistent with CO_2^{FA} variation in the pre- and post-MPT periods (Fig. 1d). This correspondence suggests that both the $\delta^{11}B$ and $\delta^{13}C_{FA}$ proxies reflect the atmospheric CO_2 concentration, but the absolute concentrations remain a calibration issue for converting $\delta^{11}B$ into the CO_2 concentration[35]. Alkenone $\delta^{13}C$ analysis gave higher CO_2 concentrations than blue ice[9], which could be explained by a calibration issue in the physiological parameter[36]. The pedogenic carbonate $\delta^{13}C$ and C_3 plant $\delta^{13}C$ reconstructions showed wider CO_2 values than that of blue ice (Extended Data Figs. 9 and 10)[10,11].

The CO_2^{FA} varied between 180 and 283 (± 12) ppm during the last 900 ka. The highest CO_2^{FA} concentrations appear in marine isotope stages (MIS) 5, 9, and 11, as is also the case for the preindustrial ice core CO_2 data[3,4]. The lowest CO_2^{FA} concentration appears in MIS 22 (Fig. 1b). Prior to the MPT, CO_2^{FA} varied between 188 and 257 ppm, averaging 5 ppm lower than post-MPT values (Fig 1b). The interglacial values in the pre-MPT period were 20 ppm lower than those of the lukewarm interglacial periods (~900 to 450 ka) and 32 ppm lower than those of the last five strong interglacials (~450 ka to present) (Fig. 1d).

The CO_2^{FA} trend is consistent with the SST variations in the western Pacific warm pool (WPWP) region across the MPT (Extended Data Fig. 2d). The SSTs in the WPWP are assumed to directly reflect the greenhouse effect because of the remote location from the ice sheets[23]. On the other hand, the glacial CO_2^{FA} trend differs from the glacial cooling SST trends in the regions other than the WPWP[37]. In those regions, the glacial SSTs are assumed to be influenced by additional factors, such as ice-driven cooling, albedo

feedback, and upwelling. The CO_2^{FA} record seems to be beneficial to understand the greenhouse effect on the climate before the MPT.

It is thus important to test the robustness of our reconstruction; this can be done using future blue ice drilling, vegetation modeling studies, intercomparison with other CO_2 proxies, and $\delta^{13}C_{FA}$-based reconstructions at other sites (Extended Data Fig. 5).

CO_2 variability through the MPT

Variation in CO_2^{FA} shows 96-kyr (eccentricity), 41-kyr (obliquity), 23-, and 19-kyr (precession) cycles during 800–0 ka, while it shows 76-, and 43-kyr cycles during 1460–1000 ka (Fig. 4a, b). The variation in CO_2^{FA} is highly coherent at all significant periodicities with variations in benthic foraminifera *Uvigerina* $\delta^{18}O$ at the study site (U1446 $\delta^{18}O$ reversed)[25] as well as *Uvigerina* $\delta^{13}C$ in the ODP Site 1123 of the deep (3300 m) South Pacific (South Pacific $\delta^{13}C$)[24] (Fig. 4a, b). They show similar amplitude variations with CO_2 in each cycle (Fig 4c, d). These striking similarities indicate that CO_2, ice volume, and deep-ocean circulation are tightly coupled on orbital timescales.

Time-series analysis shows a small but robust change in the relative phase of CO_2^{FA} and benthic $\delta^{18}O$ variations over the MPT. On a 41-kyr cycle, the CO_2^{FA} preceded benthic $\delta^{18}O$ by 2.7 ± 0.9 kyr during 1460–1000 ka, whereas CO_2^{FA} was delayed by 1.7 ± 1.1 kyr during 800–0 ka (Fig. 4a). On the other hand, the CO_2^{FA} was synchronous with South Pacific $\delta^{13}C$ during 1460–1000 ka, whereas it was delayed by 2.3±0.8 kyr during 800–0 ka on a 41-kyr cycle (Fig. 4b). Thus, the shifts of the relative phase of CO_2^{FA} to benthic $\delta^{18}O$ and South Pacific $\delta^{13}C$ (4.4 and 3.3 kyr, respectively) between 1000–800 ka are statistically significant. These results imply that the interaction between CO_2, continental ice, and deep-ocean circulation changed across the MPT. Because the long-chain n-fatty

acids were stored in continental reservoirs (*e.g.*, for ~1 kyr[38] and ~2.7 kyr[39] in the Ganges-Brahmaputra river basin), CO_2^{FA} variation is expected to be delayed by the contribution of older components (Methods). However, even taking this effect into account, the phase shift during the MPT described above is significant.

The long-term mean trends of the CO_2^{FA} and $\delta^{18}O$ records differ (Fig. 1b). The mean $\delta^{18}O$ shows a long-term drift in a positive direction until 700 ka[40]. This long-term shift is not seen in CO_2^{FA} (Fig. 1b). The offset suggests that reduced continental ice volume co-existed with lower CO_2 levels before the MPT compared to afterward.

In contrast to foraminifera $\delta^{18}O$, South Pacific $\delta^{13}C$ (ref. [24]) shows a high correlation with the CO_2^{FA} on their 200-kyr running means ($r = 0.80$; Fig. 1c). The South Pacific $\delta^{13}C$ is interpreted to reflect changes in deep-ocean circulation coupled with changes in the Southern Ocean deep-water ventilation and marine carbon reservoir[24]. Thus, the long-term co-variation of CO_2^{FA} and South Pacific $\delta^{13}C$ suggests that the deep-ocean circulation changes were a major factor determining the long-term trend of the atmospheric CO_2 concentration *via* changes in ocean carbon storage through the MPT[21,41–43].

Mean values of CO_2^{FA} and South Pacific $\delta^{13}C$ (ref.[24]) decreased by 15 ppm and 0.3‰ on average, respectively, from 1,150 to 950 ka during the MPT (Fig. 1c). The decrease in CO_2 may have driven the expansion of continental ice sheets by decreasing the greenhouse effect. However, from 950 ka to 700 ka, the ice sheets continued to expand, despite the increase in mean CO_2 (Fig. 1b). This peculiar change reflected a changing response of ice sheets to CO_2 forcing throughout 950–700 ka, as shown by the two distinct regression trends between CO_2^{FA} and the deep-water (3,300 m) $\delta^{18}O$ in the South Pacific (a proxy of continental ice volume) [24] before and after 900 ka (Fig. 2c). As described

earlier, at a 41,000-year cycle, CO_2 change precedes benthic $\delta^{18}O$ change before the MPT, but benthic $\delta^{18}O$ change slightly precedes CO_2 after 800 ka (Fig. 4a). These results suggest that the CO_2–ice interaction was reorganized during the MPT. This reorganization cooccurred with the beginning of large ice sheet development under relatively high CO_2 concentrations. The expansion of the North American ice sheet could have triggered the onset of the glacial 100-kyr cycle[44].

This study raises the new key question of what mechanism caused this reorganization. The close relationship between CO_2^{FA} and South Pacific $\delta^{13}C$ variability in both the long-term trend (Fig. 1c) and glacial–interglacial cycles (Fig. 4b, d) indicates that the process involved deep-ocean circulation. The Atlantic Meridional Overturning Circulation was consistently active during both glacials and interglacials before 900 ka[21,43], while it damped during glacials after 900 ka. In response to this change, carbonate preservation, with maxima at deglaciations and minima at glaciations, increased in the deep Pacific across the MPT[45]. If the more severe carbonate dissolution before the MPT simply means that the abyssal Pacific or the Pacific sector of the Southern Ocean sequestered more carbon, the carbon storage in those areas would constitute key information for understanding the low CO_2 concentrations in the pre-MPT period. Elucidating this mechanism is critical for understanding past and future changes in the interactions between CO_2, continental ice, and deep-ocean circulation.

Acknowledgments This research used samples provided by the International Ocean Discovery Program Expedition 353. Yuka Sazuka, Akiko Muto, and Kaori Ono are thanked for their analytical assistance. Yoshito Chikaraishi kindly provided the end-member value of wild C_4 plants. M.Y. was funded by JSPS grants JPMXS05R2900001

and 19H05595 and JAMSTEC Exp 353 post-cruise study. S.C. was supported by US NSF OCE1634774. A.A-O. was supported by JSPS grant 17H06104, MEXT grant 17H06323, and JAMSTEC to use the Earth Simulator supercomputer. R.O. was supported by ArCS Project JPMXD1300000000, MEXT Japan.

Author contributions M.Y. designed the study. M.Y., O.S., Y.T., and Y.H. generated $\delta^{13}C_{FA}$ data. S.C. generated foraminifera $\delta^{18}O$ data and age-depth model. M.Y. analyzed data. R.O. and A.A. performed model experiments. M.Y. wrote the manuscript with input from others.

Competing interests All authors declare that they have no competing interests.

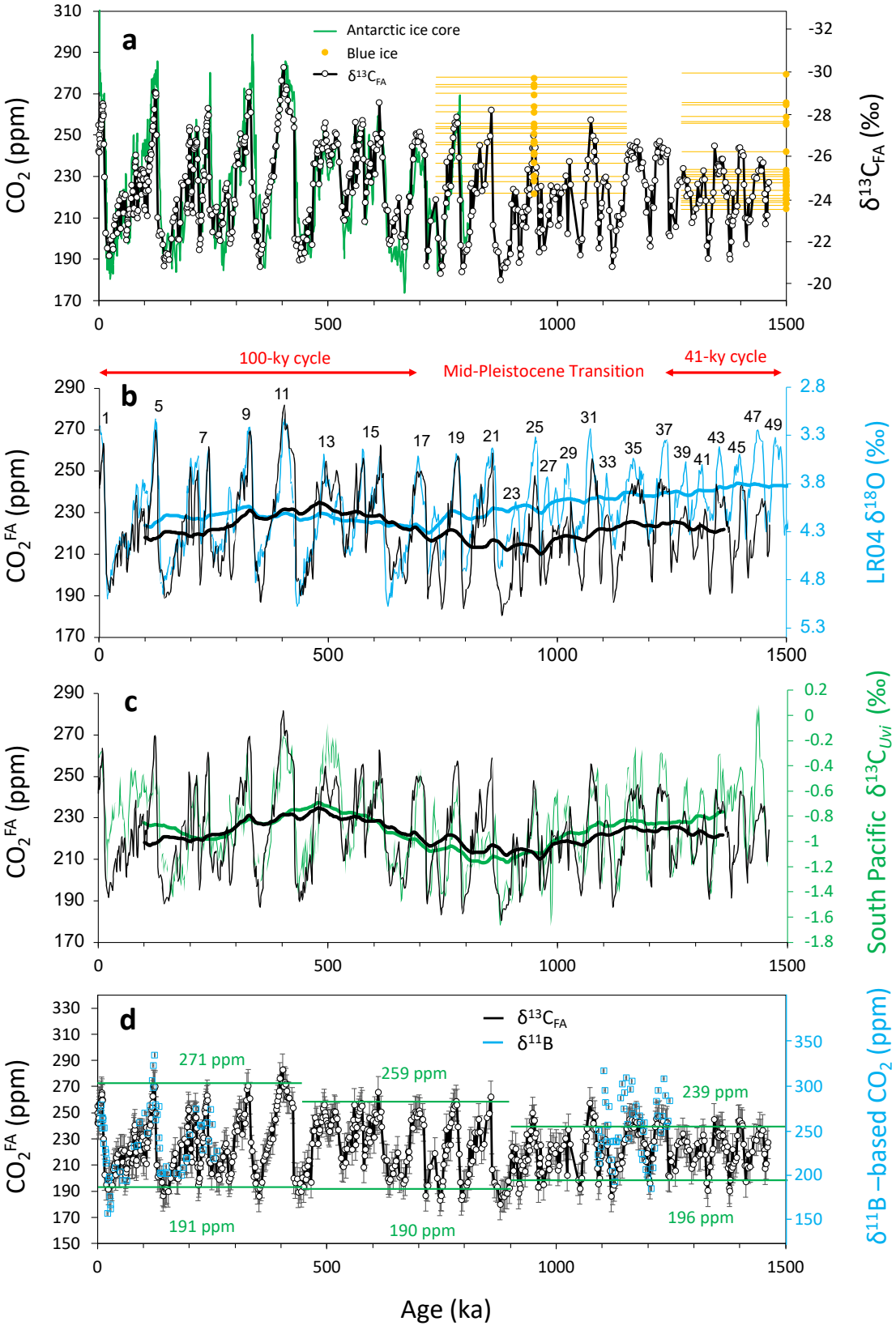


Fig. 1. Reconstructed CO_2 concentration during the last 1.46 Myr. Reconstructed

atmospheric CO_2 concentration (CO_2^{FA}) based on fatty acid $\delta^{13}C$ ($\delta^{13}C_{FA}$) with a calibration error of 12 ppm at Site U1446 during the last 1.46 Myr (black) with **a** the Antarctic ice core CO_2 concentration (green)[4] and CO_2 concentration in Allan Hills blue ice with age uncertainty ±213 kyr (orange)[13,14]. **b,** the benthic foraminifera LR04 $\delta^{18}O$ stack[40] with marine isotope stages. **c,** benthic foraminifera *Uvigerina* $\delta^{13}C$ in the ODP Site 1123 of the deep South Pacific[24]. **d,** the average CO_2 concentrations of interglacials and glacials (green lines) and the $\delta^{11}B$-based CO_2 concentration (blue; with 2σ interval)[7]. Thick lines in panels **b** and **c** indicate the 200-kyr running means. The scale of $\delta^{11}B$-based CO_2 concentration in panel **d** is adjusted to the range of Antarctic ice core CO_2 during the last two glacial cycles following the regression equation: [Ice core CO_2] = 0.678 × [$\delta^{11}B$-based CO_2] + 67 (r = 0.87, n = 57).

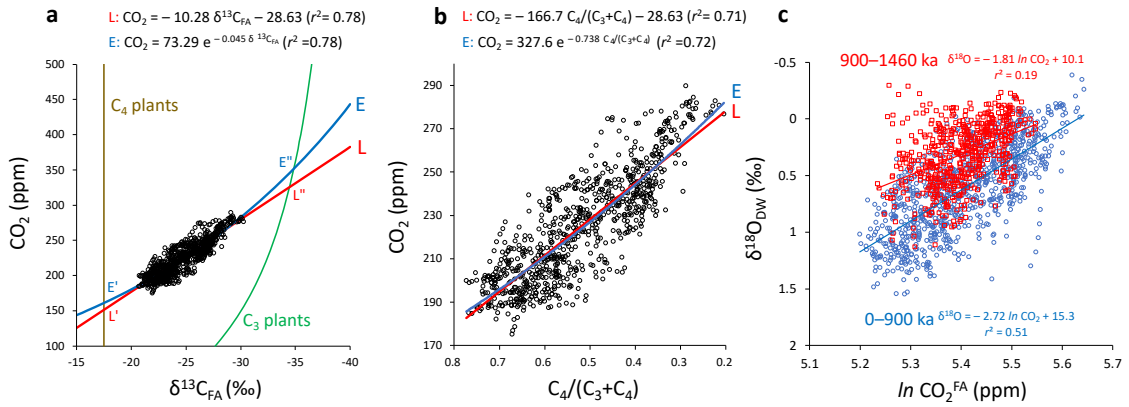


Fig. 2. $\delta^{13}C_{FA}$, $C_4/(C_3+C_4)$ plant ratio, CO_2 concentration, and ice volume. **a,** Plot of $\delta^{13}C_{FA}$ (axis reversed) and the Antarctic ice core CO_2 concentration[4], and end-member values of $\delta^{13}C_{FA}$ of C_3 and C_4 plants predicted based on modern $\delta^{13}C$ values[46,47], the empirical relationship between the isotope fractionation of $\delta^{13}C$ of C_3 plants and CO_2 concentration and the Suess effect (Methods). The Antarctic CO_2 record[4] was tuned to the $\delta^{13}C_{FA}$ record, as shown in Extended Data Fig. 6 (Methods). Letters L and E indicate the linear and exponential regression lines. Letters L' and L" indicate the lower and upper ends (151 and 325 ppm), respectively, of the $\delta^{13}C_{FA}$–CO_2 calibration in the linear regression model. Letters E' and E" indicate the lower and upper ends (161 and 357 ppm), respectively, in the exponential regression model. **b,** The plot of $C_4/(C_3+C_4)$ plant ratio and the ice core CO_2 concentration, **c,** The plot of *ln* CO_2^{FA} and the $\delta^{18}O$ of deep-water in the South Pacific ($\delta^{18}O_{DW}$)[24] before and after 900 ka (red and blue dots, respectively). $\delta^{18}O_{DW}$ reflects the volume of the continental ice sheets[24].

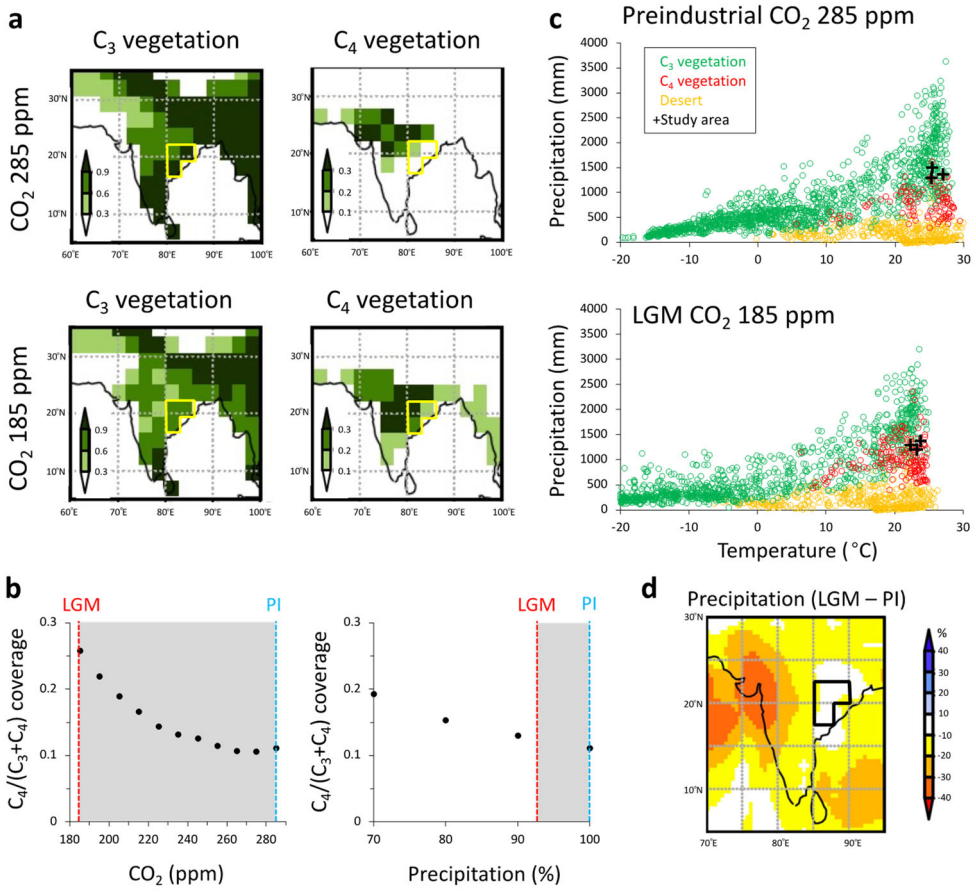


Fig. 3. Dynamic vegetation model results. a, The fractional coverages of C_3 and C_4 vegetation at preindustrial (PI) temperature and precipitation conditions for a CO_2 concentration of 185 and 285 ppm. **b,** Changes in the coverages of C_4 and C_3 plants in the sediment source area (grids with the yellow border in panel **a**) as the CO_2 concentration decreases from 285 to 185 ppm under PI precipitation and temperature[30–32] and as precipitation decreases from 100 (PI condition) to 70% under PI CO_2 and temperature[30–32]. **c,** World C_3/C_4 vegetation under PI and Last Glacial Maximum (LGM) conditions[30–32] estimated with the Lund–Potsdam–Jena Dynamical Global Vegetation Model (LPJ-DGVM)[48]. **d**, The difference in annual mean precipitation between the LGM and PI conditions (LGM – PI), estimated with nine different general circulation models (Extended Data Fig. 4)[30–32].

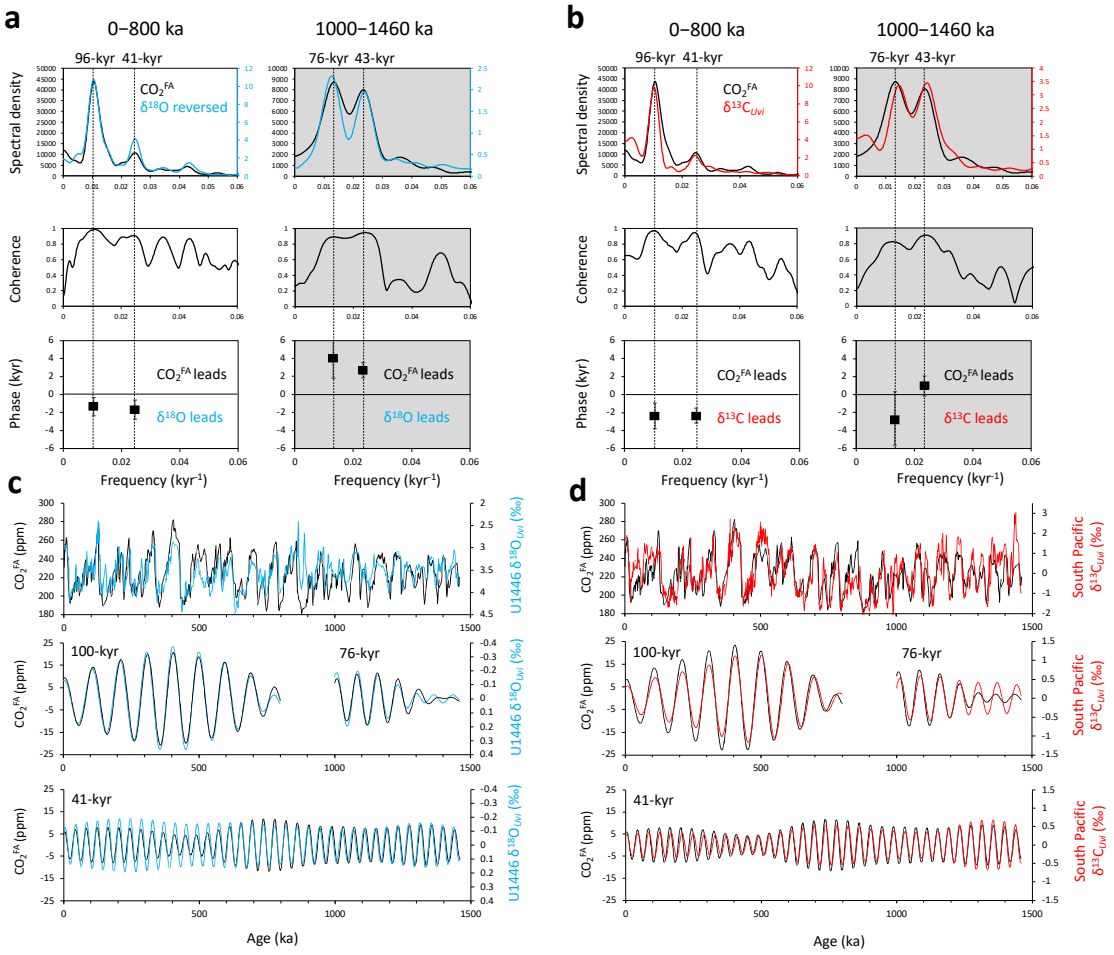


Fig. 4. Phase lead and lag of CO_2^{FA}, ice volume, and deep-ocean circulation. a and **b,** Power spectra, coherence, and phase differences with the 80% confidence level (vertical bars) of the variation in CO_2^{FA}, benthic foraminifera *Uvigerina* $\delta^{18}O$ at the study site (U1446 $\delta^{18}O$ reversed, reflecting global ice volume and bottom-water temperature)[25], and *Uvigerina* $\delta^{13}C$ at ODP Site 1123 in the deep South Pacific (South Pacific $\delta^{13}C_{Uvi}$, reflecting deep-ocean circulation)[24]. **c** and **d,** The variation in CO_2^{FA}, U1446 $\delta^{18}O$ (axis reversed), South Pacific $\delta^{13}C_{Uvi}$, and their 100-, 76-, and 41-kyr filtered components before and after the MPT (1460–1000 and 800–0 ka, respectively). Cross-spectral analysis was performed with the AnalySeries package[49].

References:

1. Barnola, J. M., Raynaud, D., Korotkevich, Y. S. & Lorius, C. Vostok ice core provides 160,000-year record of atmospheric CO_2. *Nature* **329**, 408–414 (1987).
2. Petit, J. R. et al. Climate and atmospheric history of the past 420,000 years from the Vostok ice core. *Nature* **399**, 429–436 (1999).
3. Lüthi, D. et al. High-resolution carbon dioxide concentration record 650,000–800,000 years before present. *Nature* **453**, 379–382 (2008).
4. Bereiter, B. et al. Revision of the EPICA Dome C CO_2 record from 800 to 600 kyr before present. *Geophys. Res. Let.* **42**, 542–549 (2015).
5. Foster, G. L. Seawater pH, pCO_2 and $[CO^{2-}]$ variations in the Caribbean Sea over the last 130 kyr: A boron isotope and B/Ca study of planktic foraminifera. *Earth Planet. Sci. Lett.* **271**, 254–266 (2008).
6. Hönisch, B., Hemming, N. G., Archer, D., Siddall, M. & McManus, J. F. Atmospheric carbon dioxide concentration across the Mid-Pleistocene Transition. *Science* **324**, 1551–1554 (2009).
7. Chalk, T. B. et al. Causes of ice age intensification across the mid-Pleistocene transition. *Proc. Nat. Acad. Sci.* **114**, 13114–13119 (2017).
8. Dyez, K. A., Hönisch, B. & Schmidt, G. A. Early Pleistocene obliquity-scale pCO_2 variability at ~1.5 million years ago. *Paleoceanogr. Paleoclimatol.* **33**, 1270–1291 (2018).
9. Pagani, M., Liu, Z., LaRiviere, J., Ravelo, A. C. High Earth-system climate sensitivity determined from Pliocene carbon dioxide concentrations. *Nature Geoscience* **3**, 27–30 (2010).
10. Da, J., Zhang, Y. G., Li, G. Meng, X. & Ji, J. Low CO_2 levels of the entire Pleistocene

epoch *Nature Comm.* **10**, 4342 (2019)

11. Cui, Y., Schubert, B. A. & Jahren, A.H., A 23 m.y. record of low atmospheric CO_2. *Geology* 48, 888–892 (2020).
12. Martínez-Botí, M. A. et al. Boron isotope evidence for oceanic carbon dioxide leakage during the last deglaciation. *Nature*, **518**, 219–222 (2015).
13. Higgins, J. A. et al. Atmospheric composition 1 million years ago from blue ice in the Allan Hills, Antarctica. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6887–6891 (2015).
14. Yan, Y. et al. Two-million-year-old snapshots of atmospheric gases from Antarctic ice. *Nature* **574**, 663–666 (2019).
15. Pisias, N. G. & Moore, Jr., T. C. The evolution of the Pleistocene climate: a time series approach. *Earth Planet. Sci. Let*. **52**, 450–458 (1981).
16. Clark, P. U. et al. The middle Pleistocene transition: characteristics, mechanisms, and implications for long-term changes in atmospheric pCO_2. *Quat. Sci. Rev.* **25**, 3150–3184 (2006).
17. Clark, P. U. & Pollard, D. Origin of the Middle Pleistocene Transition by ice sheet erosion of regolith. *Paleoceanogr.* **13**, 1–9 (1998).
18. Ganopolski, A. & Calov, R. The role of orbital forcing, carbon dioxide and regolith in 100 kyr glacial cycles. *Clim. Past* **7**, 1415–1425 (2011).
19. Tabor, C. R. & Poulsen, C. J. Simulating the mid-Pleistocene transition through regolith removal. *Earth Planet. Sci. Lett.* **434**, 231–240 (2016).
20. Willeit, M., Ganopolski, A., Calov, R. & Brovkin, V. Mid-Pleistocene transition in glacial cycles explained by declining CO_2 and regolith removal. *Sci. Adv.* **5**, eaav7337 (2019).
21. Pena, L. D. & Goldstein, S. L. Thermohaline circulation crisis and impacts during

the mid-Pleistocene transition. *Science* **345**, 318–322 (2014).

22. Raymo, M. E. & Huybers, P. Unlocking the mysteries of the ice ages. *Nature* **451**, 284–285 (2008).
23. Medina-Elizalde, M. & Lea, D. W. The mid-Pleistocene transition in the tropical Pacific. *Science* **310**, 1009–1012 (2005).
24. Elderfield, H. et al. Piotrowski, Evolution of ocean temperature and ice volume through the mid-Pleistocene climate transition. *Science* **337**, 704–708 (2012).
25. Clemens, S. C. et al. Remote and local drivers of Pleistocene South Asian summer monsoon precipitation: A test for future predictions. *Science Advances* **7**, eabg3848 (2021).
26. Edwards, E. J., Osborne, C. P., Strömberg, C. A. E. & Smith, S. A. C_4 Grasses Consortium, The Origins of C_4 grasslands: Integrating evolutionary and ecosystem science. *Science* **328**, 587–591 (2010).
27. Cerling, T. E., Ehlenringer, J. R. & Harris, J. M. Carbon dioxide starvation, the development of C_4 ecosystem, and mammalian evolution. *Phil. Trans. R. Soc. Lond.* B, **353**, 159–171 (1998).
28. Ainsworth, E. & Long, S. P. What have we learned from 15 years of free-air CO_2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2. *New Phytol.* **165**, 351–372 (2005).
29. Schubert, A. & Jahren, A. H. Global increase in plant carbon isotope fractionation following the Last Glacial Maximum caused by increase in atmospheric pCO_2. *Geology* **43**, 435–438 (2015).
30. Braconnot, P. et al. Evaluation of climate models using palaeoclimatic data. *Nature*

Climate Change **2**, 417-424 (2012).

31. Abe-Ouchi, A., Saito, F., Kageyama, M., Braconnot, P., Harrison, S. P., Lambeck, K., Otto-Bliesner, B. L., Peltier, W. R., Tarasov, L., Peterschmitt, J.-Y. & Takahashi, K. Ice-sheet configuration in the CMIP5/PMIP3 Last Glacial Maximum experiments, *Geosci. Model Dev*. 8, 3621–3637 (2015).
32. O'ishi, R. & Abe-Ouchi, A., Influence of dynamic vegetation on climate change and terrestrial carbon storage in the Last Glacial Maximum. *Clim. Past* **9**, 1571–1587 (2013).
33. Rao, V. P., Reddy, N. P. & Rao, C. M. Clay mineral distribution in the shelf sediments off the northern part of the east coast of India. *Cont. Shelf Res.* **8**, 145–151 (1988).
34. Phillips, S. C. et al. Long-timescale variation in bulk and clay mineral composition of Indian continental margin sediments in the Bay of Bengal, Arabian Sea, and Andaman Sea. *Mar. Petrol. Geol.* **58**, A, 117–138 (2014).
35. Hain, M. P., Foster, G. L. & Chalk, T. Robust constraints on past CO_2 climate forcing from boron isotope proxy. *Paleoceanogr. Paleoclimatol.* **33**, 1099–1115 (2018).
36. Zhang, Y. G. et al. Refining the alkenone-pCO_2 method I: Lessons from the Quaternary glacial cycles. *Geochim. Cosmochim. Acta* **260**, 177–191 (2019).
37. McClymont, E. L., Sosdian, S. M., Rosell-Melé, A. & Rosenthal, Y. Pleistocene sea surface temperature evolution: Early cooling, delayed glacial intensification, and implications for the mid-Pleistocene climate transition. *Earth Sci. Rev*. **123**, 173–193 (2013).
38. French, K. L. et al. Millennial soil retention of terrestrial organic matter deposited in the Bengal Fan. *Sci. Rep.* **8**, 11997 (2018).

39. Hein, C. J., Usman, M., Eglinton, T. I., Haghipour, N. & Galy, V. V. Millennial-scale hydroclimate control of tropical soil carbon storage. *Nature* **581**, 63–66 (2020).

40. Lisiecki, L. E. & Raymo, M. E. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. *Paleoceanogr.* **20**, PA1003 (2005).

41. Hodell, D. A., Venz, K. A., Charles, C. D. & Ninnemann, U. S. Pleistocene vertical carbon isotope and carbonate gradients in the South Atlantic sector of the Southern Ocean. *Geochem. Geophys. Geosyst.* **4**, 1004, doi:10.1029/2002GC000367 (2003).

42. Lisiecki, L. E. A benthic $\delta^{13}C$-based proxy for atmospheric pCO_2 over the last 1.5 Myr. *Geophys. Res. Let.* **37**, L21708 (2010).

43. Farmer, J. R. et al. Deep Atlantic Ocean carbon storage and the rise of 100,000-year glacial cycles. *Nature Geoscience* **12**, 355–360 (2019).

44. Abe-Ouchi, A. et al. Insolation-driven 100,000-year glacial cycles and hysteresis of ice-sheet volume. *Nature* **500**, 190–193 (2013).

45. Sexton, P. F. & Barker, S. Onset of 'Pacific-style' deep-sea sedimentary carbonate cycles at the mid-Pleistocene transition. *Earth Planet. Sci. Let.* **321–322**, 81–94 (2012).

46. Chikaraishi, Y., Naraoka, H. & Poulson, S. R. Hydrogen and carbon isotopic fractionations of lipid biosynthesis among terrestrial (C3, C4 and C.A.M.) and aquatic plants. *Phytochem.* **65**, 1369–1381 (2004).

47. Chikaraishi, Y., Takano, Y. & Ohkouchi, N. Heterotrophic uptake of soil amino acids by C4 plants: unusual stable isotopic composition of C4 plant lipids may be derived from amino acid uptake. *Paper presented at the 27th International Meeting of Organic Geochemistry*, Prague, Czech Republic, 14 September 2015.

48. Sitch, S. et al. Evaluation of ecosystem dynamics, plant geography and terrestrial

carbon cycling in the L.P.J. dynamic global vegetation model. *Global Change Biology* **9**, 161–185 (2003).

49. Paillard, D., Labeyrie, L. & Yiou, P. Macintosh program performs time-series analysis. *Eos, Trans. Amer. Geophys. Union* **77**, 379 (1996).
50. Mazumdar, A. et al. Geochemical characterization of the Krishna–Godavari and Mahanadi offshore basin (Bay of Bengal) sediments: A comparative study of provenance. *Mar. Petrol. Geol*. **60**, 18–33 (2015).
51. Gebregiorgis, D. et al. Southern hemisphere forcing of South Asian monsoon precipitation over the past ~1 million years. *Nature Comm*. **9**, DOI: 10.1038/s41467-081-0706-2 (2018).
52. Bolton, C.T. et al. A 500,000 year record of Indian summer monsoon dynamics recorded by eastern equatorial Indian Ocean upper water-column structure. *Quat. Sci. Rev.* **77**, 167–180 (2013).
53. Clemens, S. C. & Prell, W. L. A 350,000 year summer-monsoon multi-proxy stack from the Own Ridge, Northern Arabian Sea. *Mar. Geol*. **201**, 35–51 (2003).
54. Caley, T. et al. New Arabian Sea records help decipher orbital timing of Indo-Asian monsoon. *Earth Planet. Sci. Lett*. **308**, 433–444 (2011).
55. Wara, M. W., Ravelo, A. C. & Delaney, M. L. Permanent El Niño-like conditions during the Pliocene warm period. *Science* **309**, 758–761 (2005).
56. de Garidel-Thoron, T., Rosenthal, Y., Bassinot, F. & Beaufort, L. Stable sea surface temperatures in the western Pacific warm pool over the past 1.75 million years. *Nature* **433**, 293–297 (2005).
57. McGrath, S. M., Clemens, S. C., Huang, Y. & Yamamoto, M. Greenhouse gas and ice volume drive Pleistocene Indian summer monsoon precipitation isotope

variability. *Geophys. Res. Lett.* **48**, e2020GL092249 (2021).

58. Ansari, A., Noble, J., Deodhar, A. & Kumar, U. S. Atmospheric factors controlling the stable isotopes ($\delta^{18}O$ and $\delta^{2}H$) of the Indian summer monsoon precipitation in a drying region of eastern India. *J. Hydrol.* **584**, 124636 (2020).

METHODS

Samples

IODP Site U1446 is located near the mouth of the Mahanadi River at 19.08°N, 85.73°E and 1425 meters water depth on the slope of the Indian margin in the Bay of Bengal (Extended Data Fig. 1). Sediments were drilled to 180 m below the seafloor and consist of calcareous microfossil-bearing clay and silt[59]. Samples were taken every 30 cm (1500−2000 years) throughout and cover the last 1.46 Myr. The sediments at Site U1446 were derived mainly from the Mahanadi River and the adjacent coastal rivers (Extended Data Fig. 1)[33,34,50], although the potential contribution of the Ganges–Brahmaputra River cannot be ignored[35]. The mineral, major, trace, and rare earth element compositions in Sites NGHP 18A and 19A were almost constant in Miocene to Pleistocene sediments, suggesting no significant change in sediment provenance[34,50]. An age model in calendar years was created by oxygen isotope stratigraphy[40] of benthic foraminifera *Uvigerina spp.* (Supplementary Table 1)[25].

The modern vegetation of the Mahanadi and Godavari River basins consists mainly of tropical thorny vegetation, with tropical dry deciduous forest and tropical moist deciduous forest (Map by Gaussen *et al.* shown in refs.[60,61]). A pollen assemblage in Holocene sediments at Site NGHP-01-16, located ~400 km southwest of the study site, near the Godavari River mouth, demonstrates that Poaceae and Cyperaceae pollen comprise 50% of the total pollen in the Holocene sediments. The content of Poaceae and Cyperaceae pollen reached 71% in Heinrich Stadial 2 (~26 ka) when the CO_2 concentration was ~200 ppm, which was ~80 ppm lower than the preindustrial level. Because 62% of Poaceae and 16% of Cyperaceae species are C_4 plants[62], the increases in Poaceae and Cyperaceae

in the glacial period are consistent with the expansion of C_4 plants in the sediment source area shown by the positive shift of $\delta^{13}C_{FA}$ at Site U1446 during glacial periods.

Fatty acids and their $\delta^{13}C$

Lipids were extracted (×3) from ca. 3 g of dried sediment using a DIONEX Accelerated Solvent Extractor ASE-200 with dichloromethane–methanol (6:4). The extract was separated into neutral and acid fractions by aminopropyl silica gel column chromatography[63]. The acid fraction was methylated with the methanol–acetyl chloride (95:5) and purified with SiO_2 column chromatography. The $\delta^{13}C$ (‰ VPDB) of the methylated n-fatty acid was analyzed using an Agilent 6890 series gas chromatograph combined with a Finnigan MAT DELTA^{plus} mass spectrometer. The $\delta^{13}C$ values of fatty acids were obtained from the measured values of fatty acid–methyl esters by correcting methyl carbon (–34.1 ‰). The reproducibility of the measurement based on repeated analyses was better than ± 0.1 ‰. The $\delta^{13}C_{FA}$ was calculated by averaging the $\delta^{13}C$ of C_{26}, C_{28}, C_{30}, and C_{32} n-fatty acids. The $\delta^{13}C$ values of *n*-C_{26}, C_{28}, C_{30}, and C_{32} fatty acids are within the range of 2‰ in most samples (Supplementary Table 2). The data of samples showing a standard deviation larger than 2 ‰ were removed from the dataset.

All samples from Site U1446 show a bimodal homologous distribution of *n*-fatty acids with maxima at C_{16}/C_{18} and C_{26}. The carbon number preference (4.4±0.5) and averaged chain length (28.2±0.3 in C_{26}–C_{32}) are nearly constant, indicating that the long-chain *n*-fatty acids are leaf wax derived from terrestrial higher plants.

Compound-specific radiocarbon analysis of surface sediments from the Santa Monica Basin indicated that the $\Delta^{14}C$ values of long-chain n-fatty acids are the same as those of phytoplankton biomarkers[64]. By contrast, the $\Delta^{14}C$ values of long-chain n-alkanes are

more depleted than those of phytoplankton biomarkers. This indicates that n-fatty acids are better for monitoring paleovegetation changes than are n-alkanes because of their shorter continental retention time.

In the Bay of Bengal, ref.[38] estimated the continental reservoir age of long-chain n-fatty acids based on the depth profile of the radiocarbon concentration in the Bengal shelf canyon. A two-end-member model indicated that 79–83% of the long-chain n-fatty acids were stored in continental reservoirs for an average of 1.0–1.2 kyr (slow-cycling component), and the remainder was stored for an average of 15 years (fast-cycling component). In the deep Bay of Bengal sediments, the age offset of long-chain n-fatty acids was estimated at 2.7 kyr on average[39]. However, the millennial-scale retention time of the slow-cycling component does not markedly delay the timing of multi-millennial scale $\delta^{13}C_{FA}$ events because it represents a broad distribution of n-fatty acids with different ages that tend to smear each other out[38,39].

End-member $\delta^{13}C_{FA}$ values of C_3 and C_4 plants

The $\delta^{13}C_{FA}$ of the C_3 and C_4 plants sampled in the early 2000s (at CO_2 of ~370 ppm) average −37.1 ± 2.0 ‰ (n = 13) and −19.5 ± 1.8 ‰ (n = 9), respectively[46,47]. However, these end-member values have varied in the past. Contemporary $\delta^{13}C_{FA}$ values for both C_3 and C_4 plants are affected by the ^{13}C Suess effect (decreasing by about 2 ‰ over the past 250 years)[65]. The $\delta^{13}C_{FA}$ values of C_3 plants have been diminished further by increasing the isotopic fractionation governed by the atmospheric CO_2 concentration[29]. The isotopic fractionation is defined as follows:

$$\Delta^{13}C = (\delta^{13}C_{CO_2} - \delta^{13}C) / (1 + \delta^{13}C /1000)$$

where $\delta^{13}C_{CO_2}$ is the $\delta^{13}C$ of the atmospheric CO_2 and $\delta^{13}C$ is the $\delta^{13}C$ of the bulk plant. The empirical relationship between the isotope fractionation of $\delta^{13}C$ of C_3 plants ($\Delta^{13}C$) and CO_2 concentration [CO_2] was obtained from field and chamber experiments[29]:

$$\Delta^{13}C = (A \times B \times ([CO_2] + C))/(A + B \times ([CO_2] + C))$$
$$= (6.22 \times ([CO_2] + 23.9))/(28.26 + 0.22 \times ([CO_2] + 23.9))$$

where A = 28.26, B = 0.22, and C =23.9[29]. Thus, the $\Delta^{13}C$ values at CO_2 concentrations of 370 and 280 ppm are 21.3 and 19.9 ‰, respectively (Supplementary Table 2). The difference is 1.4 ‰. As the $\delta^{13}C_{FA}$ of C_3 plants at a CO_2 level of 370 ppm is −37.1 ‰[46], and the Suess effect is ~2 ‰[65], the $\delta^{13}C_{FA}$ of C_3 plants ($\delta^{13}C_{FA\ (C3)}$) at CO_2 of 280 ppm before 1750 CE can be calculated as follows:

$$\delta^{13}C_{FA\ (C3)280\ ppm} = \delta^{13}C_{370\ ppm} + \text{Suess effect} + \Delta^{13}C_{370\ ppm} - \Delta^{13}C_{280\ ppm}$$
$$= -37.1 + 2.0 + 1.4 = -33.7\ ‰$$

Similarly, the $\delta^{13}C_{FA\ (C3)}$ at a given CO_2 concentration (ppm) is calculated by the following formula:

$$\delta^{13}C_{FA\ (C3)} = \delta^{13}C_{370\ ppm} + \text{Suess effect} + \Delta^{13}C_{370\ ppm} - \Delta^{13}C$$
$$= -37.1 + 2.0 + 21.3 - (6.22 \times ([CO_2] + 23.9))/(28.26 + 0.22 \times ([CO_2] + 23.9))$$

On the other hand, the $\delta^{13}C$ value for C_4 plants is independent of the atmospheric CO_2 concentration. Because the Suess effect is ~2 ‰, the $\delta^{13}C_{FA}$ of C_4 plants ($\delta^{13}C_{FA\ (C4)}$) before 1750 CE can be obtained from the $\delta^{13}C_{FA}$ of −19.5 ‰ of modern C_4 plants[47] as follows:

$$\delta^{13}C_{FA\ (C4)} = -19.5 + 2.0 = -17.5\ ‰$$

The calculated $\delta^{13}C_{FA}$ values for C_3 and C_4 plants are shown in Fig. 2a and included in Supplementary Table 3.

Over the last 30 kyrs, the $\delta^{13}C$ value of trapped atmospheric CO_2 in Antarctic ice cores has varied by 0.7 ‰[66,67]. However, this variation before 30 ka is unknown and thus not included in further calculations.

Precipitation also affects the isotopic fractionation of C_3 plants[68], but its effect is negligible in the range of glacial and interglacial mean annual precipitations (9% difference) estimated in nine GCM simulations (Fig. 3d).

Tuning the Antarctic CO_2 record to the $\delta^{13}C_{FA}$ record

The ages of gas in ice cores have uncertainty of more than ± 5 kyr[69], comparable to uncertainties in the marine record[40]. The U1446 $\delta^{13}C_{FA}$ record plotted on the LR04 marine chronology lags behind the EPICA CO_2 record on the AICC2012 chronology[69] by a maximum ~10 kyr (Extended Data Fig. 6a). Tuning the Antarctic EPICA CO_2 record to the $\delta^{13}C_{FA}$ record improves the correlation between CO_2 concentration and $\delta^{13}C_{FA}$ (linear regression; $r^2 = 0.63$ before and 0.78 after tuning) (Extended Data Fig. 6). Cross-spectral analysis confirmed that tuning increased coherences and reduced the lag between CO_2

and $\delta^{13}C_{FA}$ variations at all orbital periods (Extended Data Fig. 7). The regression lines are similar and would not change the resulting CO_2 reconstruction to any practical degree (within 5 ppm) regardless of which regression is used, with the exception that the calibration error (root-mean-square error; RMSE) is improved with tuning. The exponential regression after tuning gives a correlation of determination ($r^2 = 0.78$) as high as that of the linear regression (Fig. 2a), and the estimated CO_2 concentrations are nearly identical (within 4 ppm).

Model description

We performed 16 sensitivity experiments, examining vegetation in CO_2 levels from 185 to 285 ppm in 10-ppm steps (the control experiment = 285 ppm) and precipitation from 70 to 100% in 10% steps (the control experiment = 100%). We used the Lund–Potsdam–Jena Dynamic Global Vegetation Model (LPJ-DGVM)[48], which predicts the distributions of ten vegetation types (plant functional types, PFTs) based on bioclimatic limits, photosynthesis, respiration, carbon allocation, plant establishment, growth, turnover, mortality, and competition among the PFTs with a spatial resolution of ~2.8°. Temperature, precipitation, and cloud coverage in the control experiment are taken from the ECMWF EAR40[70] and CMAP[71] datasets for 1979–1999, and were performed 50 times to integrate 1000 years. The results from the last 20 years are analyzed. The mean temperature and precipitation in the LGM and PI conditions were calculated from those predicted with nine different GCMs[30–32] to reconstruct C_3/C_4 vegetation in the LGM and PI periods. The combination of PFTs was translated into the MATSIRO vegetation classification according to reported methods[32,72].

Spectral analysis

Spectral analysis quantifies, as a function of frequency, the lead/lag (phase) and linear coherence (calculated with the phase set to zero) relationships between two time series and has been applied routinely to Pleistocene reconstructions over orbital time scales since its early application[73,74]. Here, cross-spectral analysis was performed using the Blackman–Tukey method in the AnalySeries package[49] employing a 30% lag. The bandwidth was 0.0034, the non-zero coherence was > 0.3842, and the error estimation on the power spectrum was 0.6256. All data were interpolated to 1 kyr prior to analysis. Bandpass filtering was calculated using a Gaussian filter with the AnalySeries package and a bandwidth of 0.002.

Data availability

The data are available in supplementary tables and available at https://www.ncei.noaa.gov/pub/data/paleo/paleocean/indian_ocean/yamamoto2021/yamamoto2021-u1446.txt. The CMIP5 and PMIP3 datasets are publicly available at https://cmip.llnl.gov/. CMAP Precipitation data are available at http://www.cdc.noaa.gov/. ECMWF ERA-40 data are available at http://data.ecmwf.int/products/data/archive/.

References

59. Clemens, S. C., Kuhnt, W., LeVay, L. J. & the Expedition 353 Scientists. *Indian Monsoon Rainfall*. *Proc. IODP*, 353, College Station, TX, IODP (2016).

60. Zorzi, C. et al. Indian monsoon variations during three contrasting climate periods: The Holocene, Heinrich Stadial 2 and the last interglacial–glacial transition. *Quat. Sci. Rev.* **125**, 50–60 (2015).

61. Singh, J. S. & Chaturvedi, R. K. Diversity of ecosystem types in India: A review. *Proc. Indian Natn. Sci. Acad.* **83,** 569–594 (2017).

62. Sage, R. F. A portrait of the C_4 photosynthetic family on the 50^{th} anniversary of its discovery: species number, evolutionary lineages, and Hall of Fame. *Jour. Exp. Botany* **68**, e11–e28 (2017).

63. Gao, L., Guimond, J., Thomas, E. & Huang, Y. Major trends in leaf wax abundance, δ^2H and $\delta^{13}C$ values along leaf venation in five species of C_3 plants: Physiological and geochemical implications. *Org. Geochem.* **78**, 144–152 (2015).

64. Pearson, A., McNichol, A. P., Benitez-Nelson, B. C., Hays, J. M. & Eglinton, T. Origins of lipid biomarkers in Santa Monica Basin surface sediments: A case study using compound-specific $\Delta^{14}C$ analysis. *Geochim. Cosmochim. Acta* **65**, 3123–3137 (2001).

65. Keeling, R. F. et al. Atmospheric evidence for a global secular increase in carbon isotopic discrimination of land photosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 10361–10366 (2017).

66. Elsig, J. et al. Stable isotope constraints on Holocene carbon cycle changes from an Antarctic ice core. *Nature* **461**, 507–510 (2009).

67. Lourantou, A. et al. Constraint of the CO_2 rise by new atmospheric carbon isotopic measurements during the last deglaciation. *Global Biogeochem. Cycles* **24**, GB2015 (2010).

68. Diefendorf, A. F., Mueller, K. E., Wing, S. L., Koch, P. L. & Freeman, K. H. Global patterns in leaf ^{13}C discrimination and implications for studies of past and future climate. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 5738–5743 (2010).

69. Bazin, L. et al. An optimized multi-proxy, multi-site Antarctic ice and gas orbital chronology (AICC2012): 120–800 ka. *Clim. Past* **9**, 1715–1731 (2013).

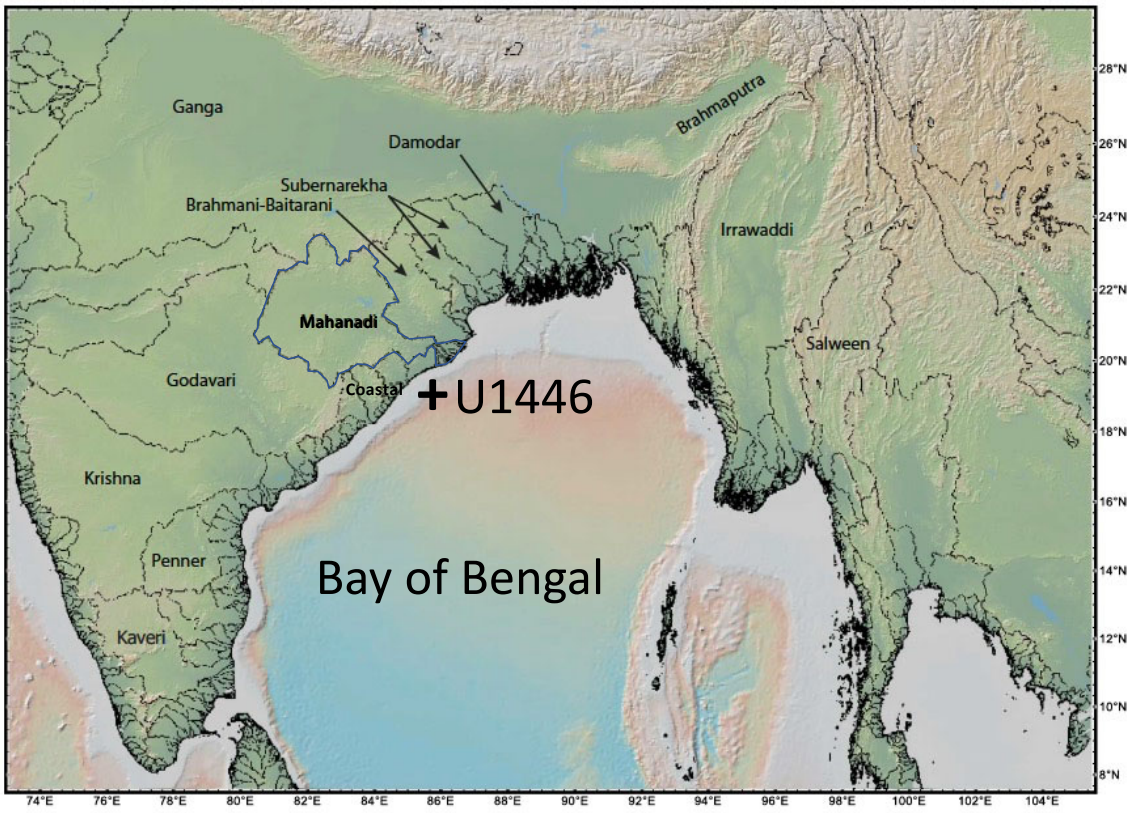
70. Uppala, S. M. et al. The ERA ‐ 40 re-analysis. *Q. J. R. Meteorol. Soc.* **131**, 2961–3012 (2005).

71. Xie, P. & Arkin, P. A. (1997), A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, *Bull. Am. Meteorol. Soc.* **78**, 2539–2558 (1997).

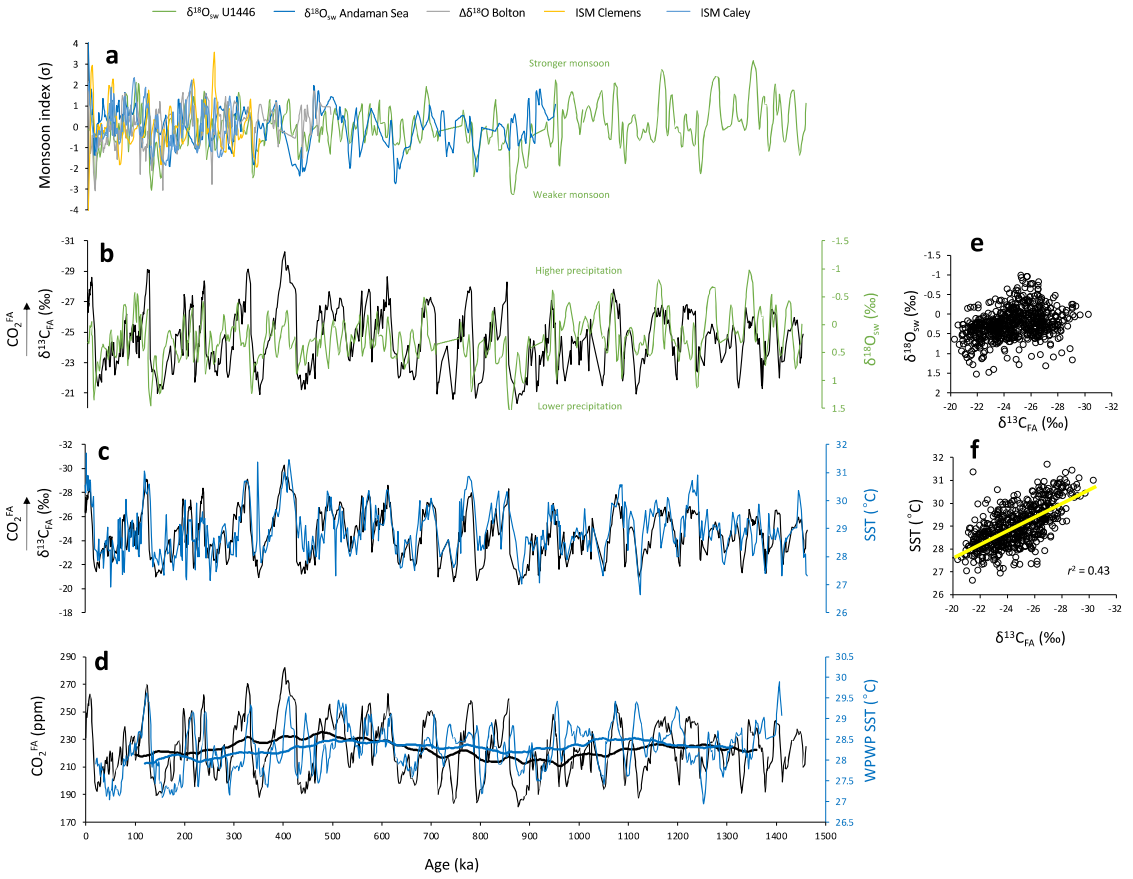
72. O'ishi, R. & Abe-Ouchi, A. Influence of dynamic vegetation on climate change arising from increasing CO_2. *Clim. Dynam.* **33**, 645–663 (2009).

73. Hays, J. D., Imbrie, J. I. & Shackleton, N. J. Variations in the earth's orbit: pacemaker of the ice ages. *Science* **194**, 1121–1132 (1976).

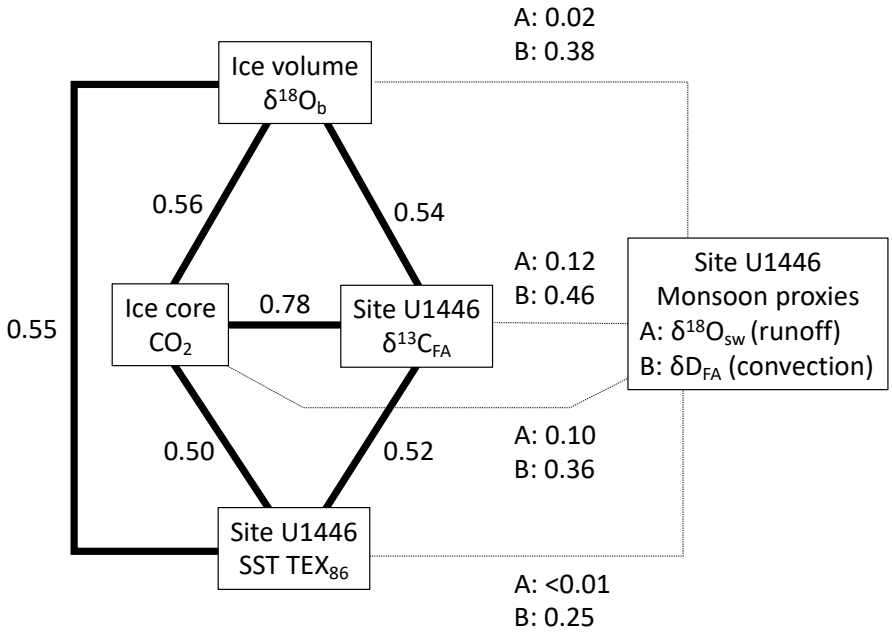
74. Imbrie, J. & Imbrie, J. Z. Modeling the climatic response to orbital variations. *Science* **207**, 943–953 (1980).



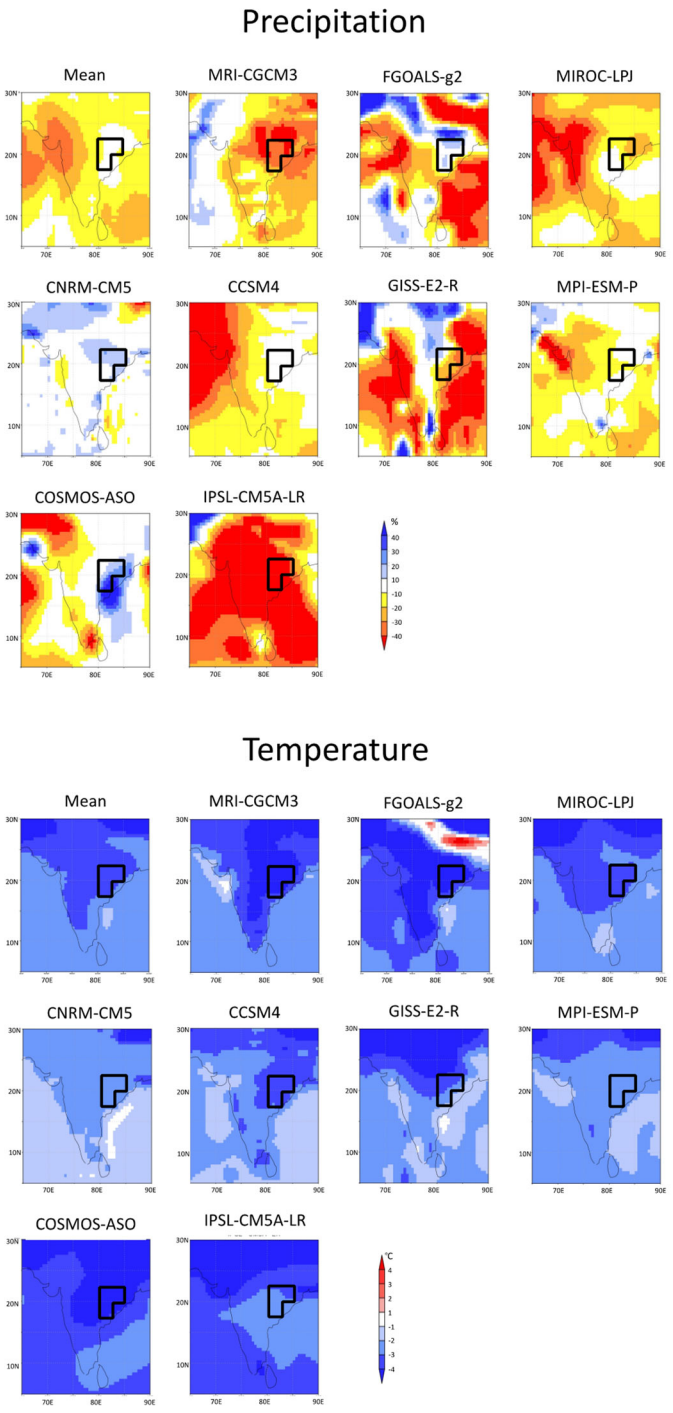
Extended Data Fig. 1. Location of IODP Site U1446. The sediments at Site U1446 were derived mainly from the Mahanadi River (blue) and the adjacent coastal rivers (yellow)[33,34,50].



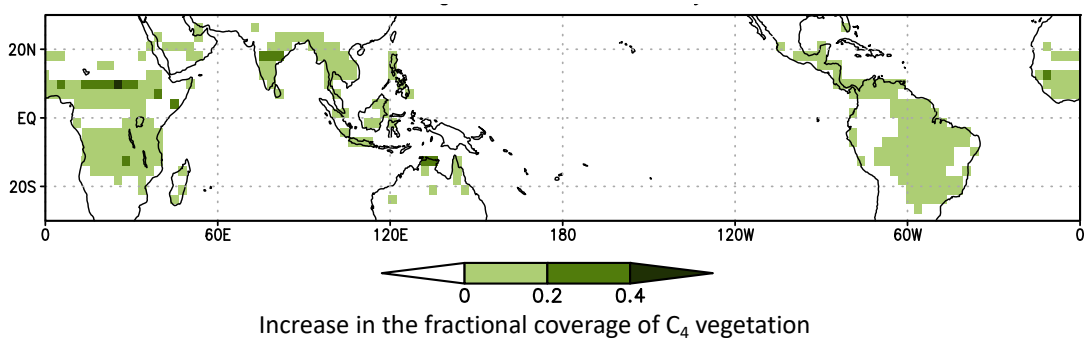
Extended Data Fig. 2. $\delta^{13}C_{FA}$, monsoon indices, and sea surface temperature (SST) at Site U1446 during the last 1.46 Myr. a, Monsoon indices: normalized values of Indian summer monsoon (ISM) records based on the seawater $\delta^{18}O$ (reversed $\delta^{18}O_{sw}$ U1446)[25] at Site U1446, seawater $\delta^{18}O$ of the Andaman Sea (reversed $\delta^{18}O_{sw}$ Andaman Sea)[51], difference between surface and thermocline foraminifera $\delta^{18}O$ ($\Delta\delta^{18}O$) in the equatorial Indian Ocean[52], Indian summer monsoon (I.S.M.) stack of the Arabian Sea[53,54]. **b** and **e,** $\delta^{13}C_{FA}$ and U1446 $\delta^{18}O_{sw}$. **c** and **f,** $\delta^{13}C_{FA}$ and the TEX_{86}^{H}-based sea surface temperature[25]. **d**, CO_2^{FA} and the western Pacific warm pool (WPWP) SST. The SSTs in the WPWP were averaged (Sites ODP Site 806[23,55] and MD97-2140[56]). Thick lines indicate the 200-kyr running mean.



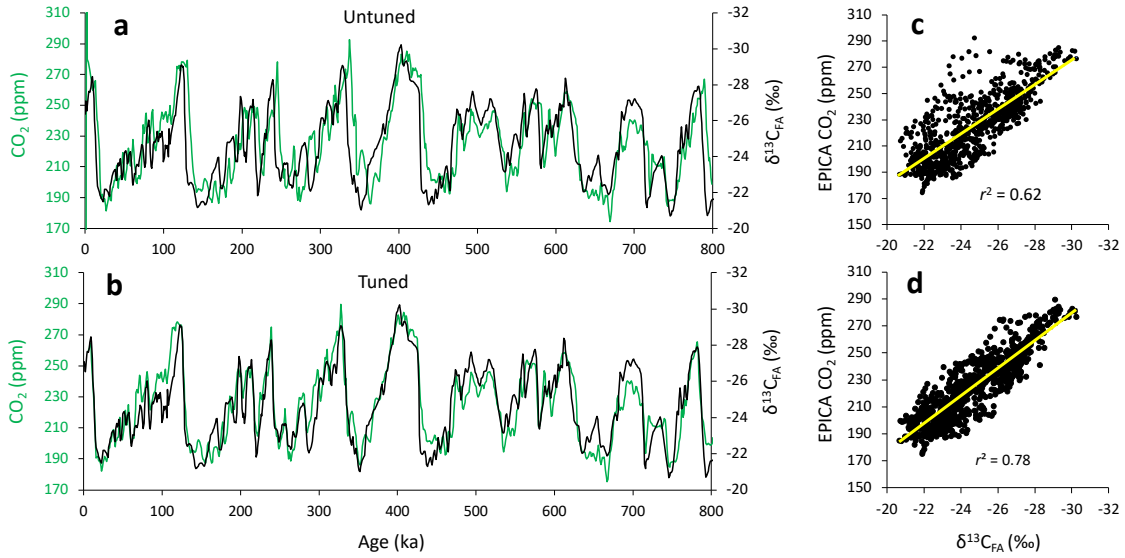
Extended Data Fig. 3. Coefficients of determination (r^2) between $\delta^{13}C_{FA}$ and climate parameters for Site U1446 during the last 805 kyr. Monsoon proxies A and B respectively correspond to the seawater $\delta^{18}O$ ($\delta^{18}O_{sw}$)[25], a proxy of salinity in the Bay of Bengal, and long-chain n-fatty acid δD (for the last 640 kyr: δD_{FA})[57], a proxy of tropical convection activity[58], at Site U1446, respectively. The r^2 value between $\delta^{13}C_{FA}$ and the Antarctic ice core CO_2 concentration with tuned age[4] is highest, indicating that atmospheric CO_2 concentration is a major factor determining $\delta^{13}C_{FA}$. The low r^2 value between $\delta^{13}C_{FA}$ and $\delta^{18}O_{sw}$ suggests that the influence of precipitation on $\delta^{13}C_{FA}$ is limited. The δD_{FA} has higher coefficients with $\delta^{13}C_{FA}$, ice core CO_2 concentration, benthic foraminifera $\delta^{18}O_b$ ($\delta^{18}O_b$)[25] and SST[25]. The higher correlation between $\delta^{13}C_{FA}$ and δD_{FA} is attributable to the response of the tropical convection activity to CO_2-induced global climate.



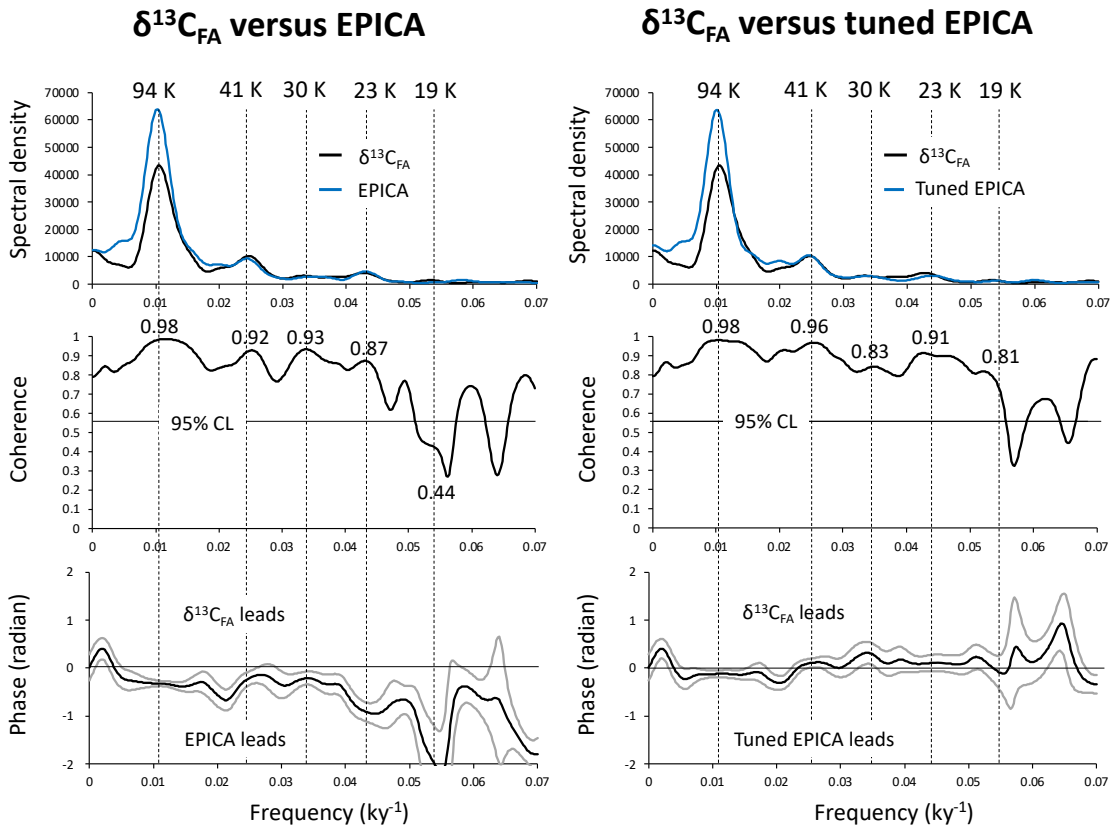
Extended Data Fig. 4. The differences in annual mean precipitation and surface air temperature in the sediment source area between the LGM and preindustrial periods (LGM – PI). They were estimated with nine different GCMs[30–32]. The decreases in annual mean precipitation and temperature in the sediment source area (grids of the black border) were 9 ± 23% and 3.4 ± 0.8°C, respectively.



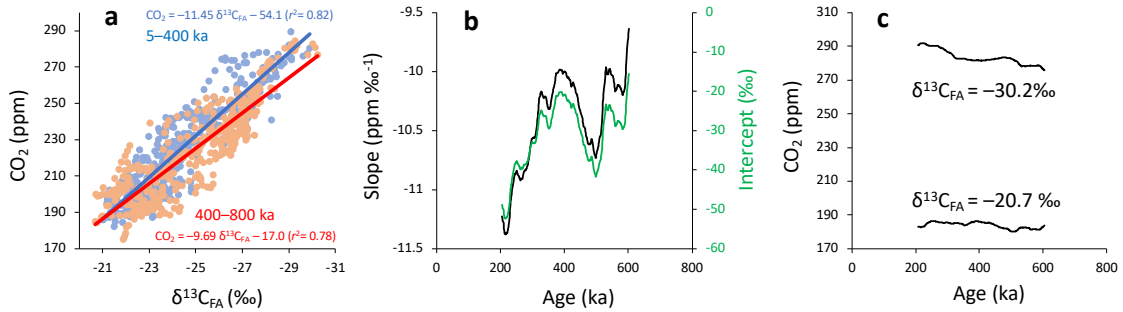
Extended Data Fig. 5. The areas where C_3/C_4 vegetation is expected to respond primarily to CO_2 changes. The green grid shows the site where the increase in C_4 vegetation due to the decrease in CO_2 from 285 to 185 ppm is greater than the increase due to the decrease in precipitation and temperature from the PI level to the LGM level. Vegetation was predicted using LPJ-DGVM under the PI and LGM conditions[30–32]. A very few regions where the increase in C_4 vegetation is significant show an empirical response of C_3/C_4 vegetation to CO_2 variation, characterized by hot and seasonally dry (savanna) climates, and could serve as targets for replicating our CO_2 reconstruction.



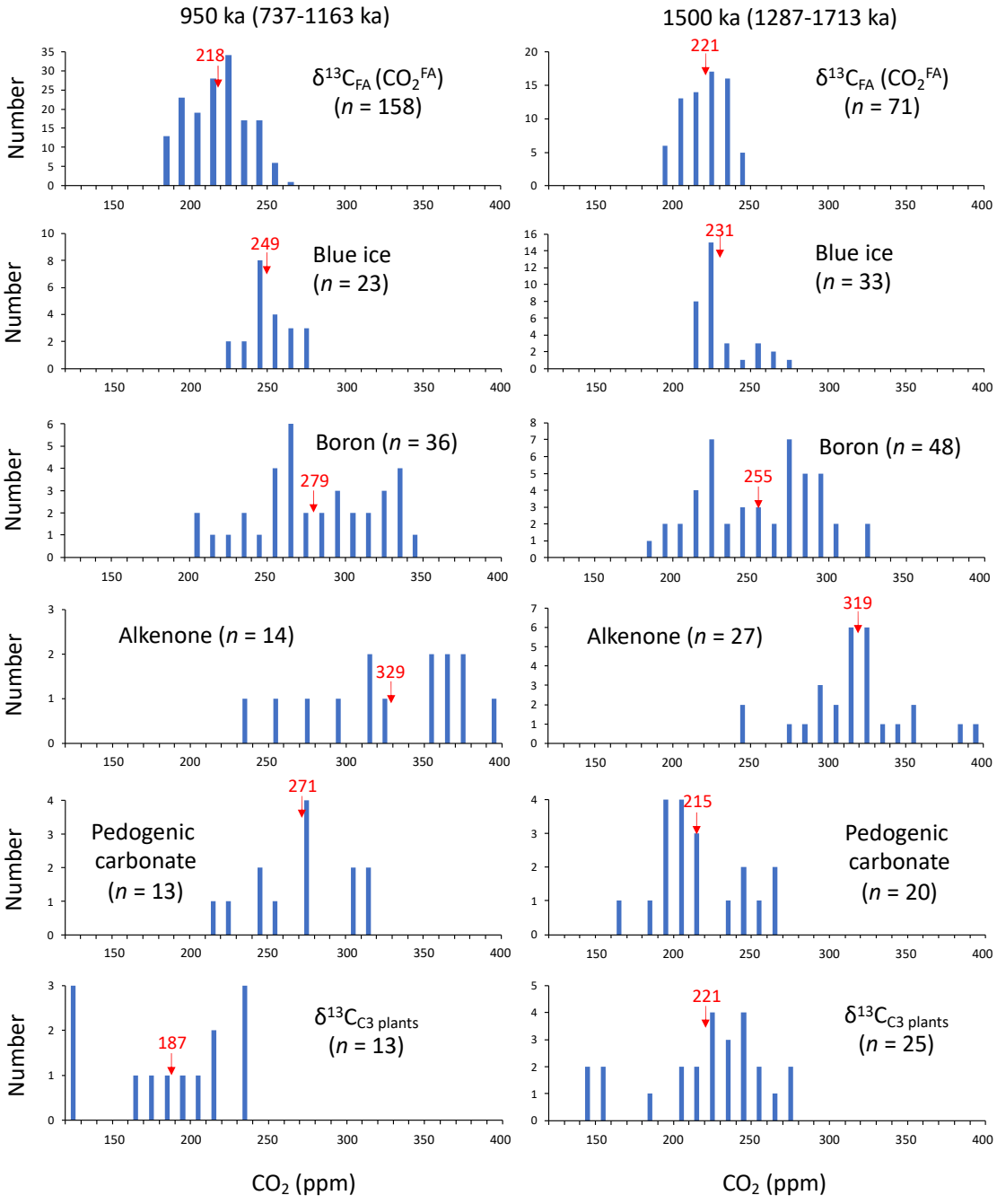
Extended Data Fig. 6. Tuning of records. a and **b,** The $\delta^{13}C_{FA}$ at Site U1446 and Antarctic ice core CO_2 concentrations[4] during the last 805 kyr before and after tuning, and **c** and **d,** the plots of $\delta^{13}C_{FA}$ and Antarctic ice core CO_2 concentration[4] between 5 and 800 ka before and after tuning. The Antarctic CO_2 record[4] was tuned to the $\delta^{13}C_{FA}$ record.



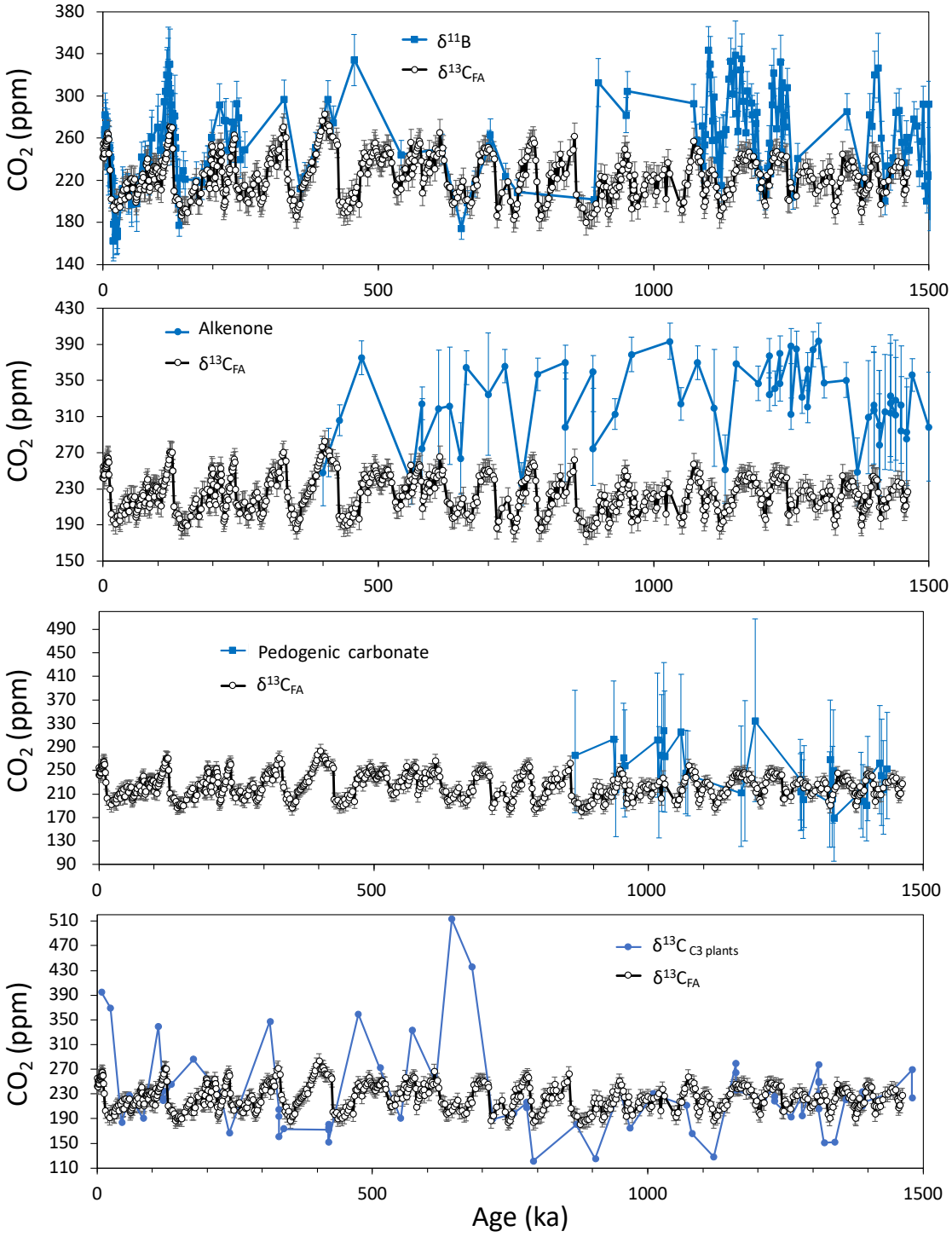
Extended Data Fig. 7. The effect of tuning. Power spectra, coherence, and phase difference of the variation in $\delta^{13}C_{FA}$-based CO_2 (CO_2^{FA}) and Antarctic EPICA CO_2 (ref.[4]) before and after tuning the EPICA record to the $\delta^{13}C_{FA}$-record in the period between 5 and 800 ka. The horizontal line in the coherence panel indicates the 95% confidence level. Faint lines in the phase difference panel indicate the upper and lower limits of the 80% confidence level. The $\delta^{13}C_{FA}$ and EPICA CO_2 records showed similar power spectra and high coherences at orbital cycles, implying that the $\delta^{13}C_{FA}$ reflects the atmospheric CO_2 concentration. The tuning increased coherence in the 41-, 23- and 19-kyr cycles, and reduced the phase lags.



Extended Data Fig. 8. The calibration of $\delta^{13}C_{FA}$ to CO_2 concentration. a, Plot of $\delta^{13}C_{FA}$ (axis reversed) and the Antarctic ice core CO_2 concentration[4] in the 5–400 and 400–800 ka periods, **b**, The slope and intercept of the linear regression equation between $\delta^{13}C_{FA}$ and ice core CO_2 variations in moving 400-kyr windows, **c**, The CO_2 concentrations estimated from constant $\delta^{13}C_{FA}$ values using the calibration equations of the panel b.



Extended Data Fig. 9. Histograms showing the number of samples by CO_2 value during two distinct periods of 950 and 1,500 ka. The values were estimated by the $\delta^{13}C_{FA}$ at Site U1446 (CO_2^{FA}), blue ice[13,14], foraminiferal $\delta^{11}B$ (ref.[8]), alkenone $\delta^{13}C$ (ref.[9]), pedogenic carbonate $\delta^{13}C$ (ref.[10]) and C_3 plant $\delta^{13}C$ (ref.[11]).



Extended Data Fig. 10. Estimates of CO_2 in various methods. CO_2 values estimated by the $\delta^{13}C_{FA}$ at Site U1446 (CO_2^{FA}) with a calibration error of 12 ppm, foraminiferal $\delta^{11}B$ (with 1σ interval; ref.[8] data after refs.[5–7]), alkenone $\delta^{13}C$ (with the upper and lower estimates)[9], pedogenic carbonate $\delta^{13}C$ (ref.[10]) and $\delta^{13}C$ of C_3 plants[11] during the past 1.5 Myr.