nature communications

Article

https://doi.org/10.1038/s41467-026-77753-0

Abrupt pH changes overestimate the responses of *Chaetoceros muelleri* to ocean acidification

Received: 1 October 2025

Accepted: 1 September 2026

Published online: 24 September 2026

Check for updates

Xin Zhao (赵欣)[1,2], Hanxi Tan (谭涵茜)[1], Xu Zhang (张旭)[1], Dong Xu (徐东)[3], Zhaowen Hu (胡朝文)[4], Jiazhen Sun (孙嘉振)[5], Shanlin Wang (王杉霖)[1], Kunshan Gao (高坤山)[1], Tianze Zheng (郑天泽)[1], Xin Lin (林昕)[1], Thomas Mock[6,7] ✉, Naihao Ye (叶乃好)[3] ✉ & Guang Gao (高光)[1,2] ✉

Ocean acidification, characterized by a gradual seawater pH decline, influences marine primary producers such as diatoms, which contribute substantially to oceanic carbon fixation. Experimental studies aiming to project climate change impacts often simulate acidification via abrupt pH shifts, potentially misrepresenting ecological and evolutionary responses. Here we conduct a long-term evolution experiment with the marine diatom *Chaetoceros muelleri*, comparing abrupt and gradual CO_2 increases from 440 to 1000 ppm over 1100 generations. Integrating evolutionary genomics, physiological assays, elemental stoichiometry, and Earth System Modeling, we find that abrupt acidification exaggerates diatom growth and carbon fixation responses by up to 15%, driven by enhanced divergent genomic signals and persistent gene expression changes. Gradual acidification elicits weaker phenotypic and genomic shifts, aligning more closely with natural ocean conditions. These findings suggest that models based on abrupt CO_2 increases may overestimate the contribution of *Chaetoceros muelleri* and maybe also other diatoms to productivity and biogeochemical cycling under ocean acidification, highlighting the importance of gradual pH decline in experimental designs to accurately predict marine ecosystem responses.

The global ocean has absorbed 20–30% of total CO_2 emissions since the 1980s, which led to a significant pH decline[1], termed ocean acidification (OA)[2]. By the end of this century, the atmospheric CO_2 level is likely to increase to 800–1000 ppm[1]. Ocean acidification impacts a variety of biological processes including calcification, silicification and carbon fixation—to name only a few—and therefore has an impact on oceanic biodiversity and associated food webs and biogeochemical cycles.

Diatoms are responsible for 40% of total oceanic primary production and play important roles in the marine biological pump through marine food webs and biogeochemical cycling[3,4]. Amorphous silica is the main component of the cell wall (frustule) of diatoms, facilitating carbon transport from surface to deep oceans[5]. Since the early Cenozoic, diatoms have become the most abundant and diverse siliceous marine microorganisms[6], and have been associated with the

[1]State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen, China. [2]Fujian Ocean Innovation Center, Xiamen, China. [3]State Key Laboratory of Mariculture Biobreeding and Sustainable Goods, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao, China. [4]Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, China. [5]National Biopesticide Engineering Research Centre, Hubei Biopesticide Engineering Research Centre, Hubei Academy of Agricultural Sciences, Wuhan, China. [6]School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich, UK. [7]College of Environmental Science and Engineering, Ocean University of China, Qingdao, China. ✉e-mail: T.Mock@uea.ac.uk; yenh@ysfri.ac.cn; guang.gao@xmu.edu.cn

biogeochemical cycling of carbon and silicon in the ocean[7,8]. Since the concentrations of dissolved CO_2 in seawater are much lower than the half-saturation constant (the substrate concentration at which the reaction rate is half of the maximum) of RUBISCO (ribulose-1,5-bisphosphate carboxylase/oxygenase), diatoms have evolved highly efficient CO_2 concentration mechanisms[9,10]. Many ocean acidification studies on diatoms have examined responses by abruptly shifting cells to target CO_2 levels and tracking performance over short-term exposures (e.g., 8–20 generations) to long-term (1000–2000 generations) exposure for investigating their response[11–13]. However, the global average rate of pH decline is about 0.002 units per year and therefore ocean acidification is a slow and gradual process[14]. Thus, compared to the long-term gradual acidification characteristic of climate change, directly exposing organisms to a target concentration of high CO_2 without any transitional acclimation represents a significantly different methodological approach. Such an abrupt change of environmental conditions therefore may not accurately reflect natural biological processes and therefore may lead to biased projections of how marine ecosystems will respond to ongoing ocean acidification[15]. Because diatoms have short generation times, gradual environmental change may allow natural selection to act over successive generations, thereby facilitating evolutionary adaptation. To address this issue, we have set up different acidification regimes: ambient CO_2 (AmbC, 440 ppm) as a control, abrupt acidification (AbruA, 1000 ppm), and gradual acidification (GradA, with CO_2 increasing from 440 to 1000 ppm in 40 ppm increments every 70 generations). Our model diatom was *Chaetoceros muelleri* because it represents the most abundant diatom genus in the oceans[16] and it can be cultivated reasonably well under laboratory conditions. A combination of targeted evolutionary genomics, physiology, elemental stoichiometry, and Earth System Modelling was used with a focus on the evolution of phenotypic traits because they underpin primary productivity and biogeochemical cycling most relevant for assessing the ecosystem response to ocean acidification and because they bridge adaptive molecular evolution with Earth System Modelling.

Here, we show that the results from this integrative work indicate that applying a gradual increase in the levels of CO_2 provides a more realistic estimate as to how much diatoms will mitigate the effects of ocean acidification and an overall increase in atmospheric CO_2 concentrations due to their enhanced primary productivity, with associated effects on the cycling of key elements such as carbon and silicon.

Results

The response of phenotypic traits and stoichiometry

Throughout the entire cultivation period, the carbonate system remained stable (Fig. 1). In particular, for the GradA group, both pH (ranging from 8.281 to 7.796) and pCO_2 (ranging from 347 to 910 µatm) exhibited a consistent trend of gradual decrease and increase, respectively, which largely aligned with the original experimental design (Fig. 1a). The specific growth rates of *C. muelleri* under three CO_2 conditions and over the entire length of the experiment ranged from 0.843 to 1.391 (Fig. 2a and Supplementary Fig. 1). In the early phase (Day 34, 55–61 generations) of the evolution experiment, AbruA caused an increase in the specific growth rate of up to 9.0% (GAMMs predicted data, ranging from 3.5% to 14.5%; measured data was 6.5%; Fig. 2a–c) compared to AmbC. However, even after ca. 550 days (over 900 generations), there was still a slight stimulatory effect (Fig. 2a, b). Compared to AbruA, GradA had a much-reduced impact on the growth rates (Fig. 2a, b). Generalized additive mixed models (GAMMs) indicated that average growth significantly benefited from AbruA (Mean: 3.2%, ranging from 2.9% to 3.5%; $t(1330) = 3.673$, Cohen's $f^2 = 0.485$, 95% CI = [0.015, 0.051], $p = 0.0002$) but not GradA (Mean: −0.2%, ranging from −0.4% to 0.1%; $t(1330) = 1.955$, Cohen's $f^2 = 0.485$, 95% CI = [8×10^{-6}, 0.020], $p = 0.051$), compared to AmbC (Fig. 2d). The results were similar for average carbon fixation because it underpins growth although there was a significant effect between GradA and AmbC based on the statistical results of the GAMMs analysis. Thus, AbruA (Mean: 12.4%, ranging from 11.2% to 13.5%; $t(120) = 3.953$, Cohen's $f^2 = 1.077$, 95% CI = [5.409, 16.046], $p = 0.0001$) was highest compared to AmbC and GradA (Mean: 3.7%, ranging from 3.2% to 4.1%; $t(120) = 2.945$, Cohen's $f^2 = 1.077$, 95% CI = [1.544, 7.685], $p = 0.004$) with strong fluctuations especially between days 200 to 500 (about 320–850 generations) (Fig. 2e, f). Specifically, the positive effects of AbruA on the growth rate and the carbon fixation rate were higher by 3.4% (ranging from 3.1% to 3.6%; $t(1330) = 4.407$, Cohen's $f^2 = 0.485$, 95% CI = [0.013, 0.033], $p < 0.0001$) and 8.2% (ranging from 7.3% to 9.0%; $t(120) = 3.902$, Cohen's $f^2 = 1.077$, 95% CI = [3.042, 9.184], $p = 0.0002$), respectively, than those of GradA. Concentrations of the main photosynthetic pigments (Chlorophyll *a* and carotenoids) did not appear to have been impacted differently by either AbruA or GradA (Fig. 3 and Supplementary Tables 1–2), but overall, AbruA had a greater negative effect on pigment concentrations on average (Fig. 3). Particulate organic carbon (POC) content on average did not significantly change between the treatments over time but there were significant oscillations identified for GradA, which suggests a different adaptive strategy under GradA compared to AbruA. The persistent oscillations in GradA might indicate ongoing adaptation to gradually changing conditions, where selection never fully stabilizes on a single phenotype because the optimum is shifting over time[17]. Although the synthesis of biogenic silica (BSi) did not seem to significantly change on average between treatments, our results suggest that the BSi cellular content decreased on average from AmbC to GradA (Mean: 11.1%, ranging from −11.9% to −10.2%; $t(352) = -5.714$, Cohen's $f^2 = 1.440$, 95% CI = [−0.015, −0.007], $p < 0.0001$) and AbruA (ranging from −17.6% to −15.7%; $t(352) = -5.852$, Cohen's $f^2 = 1.440$, 95% CI = [−0.026, −0.013], $p < 0.0001$) with the latter showing a 16.6% decrease compared to AmbC. Therefore, stoichiometric ratios between BSi and POC (BSi/POC) and PON (BSi/PON) also decreased on average from AmbC to GradA and AbruA (Fig. 4).

The response of the genome and transcriptome underpinning the phenotypes

To investigate how genes underpinning the phenotypic response have evolved, we sequenced the genome of *C. muelleri* at three time points (Days 113 (161–174 generations), 376 (612–639 generations), and 727 (1133–1202 generations)) (Supplementary Table 3) for all experimental conditions (AmbC, GradA, AbruA) and compared the level of genetic differentiation by focusing on the rate of single nucleotide polymorphism (SNP) accumulations. AbruA overall caused the fastest and strongest rate and accumulation of SNPs compared to both AmbC and GradA. The increased number of SNPs under AbruA likely contributes to the slightly elevated levels in the ratio of nonsynonymous to synonymous nucleotide substitutions (pN/pS) (Fig. 5a–c). For instance, mutations accumulating in certain genes under AbruA belonged to those involved in carbon fixation and the concentration of inorganic carbon such as CO_2. This increase in SNP numbers under AbruA therefore leads to greater genetic variation. However, pN/pS remained largely invariant across most experimental treatments. The fact that pN/pS values stayed within a narrow range of 0.08–0.18 implies that the efficacy of purifying selection was consistent. At the final time point in the GradA vs AmbC comparison, the value of pN/pS on average was 12% lower than that of AbruA vs AmbC (LSD test, $t(18) = 1.761$, Cohen's $d = 1.438$, 95% CI = [−0.004, 0.044], $p = 0.095$), which suggests slightly relaxed purifying selection in AbruA. Furthermore, differences in the SNP levels of key genes were observed throughout the cultivation period, particularly at the end of the experiment (e.g., *FBA*, *FBP*, and HCO_3^- transporter) (Fig. 5d–f). This therefore suggests that despite a relatively constrained evolution of genes (purifying selection) long-term cultivation still may have led to genetic changes in

https://doi.org/10.1038/s41467-026-77753-0

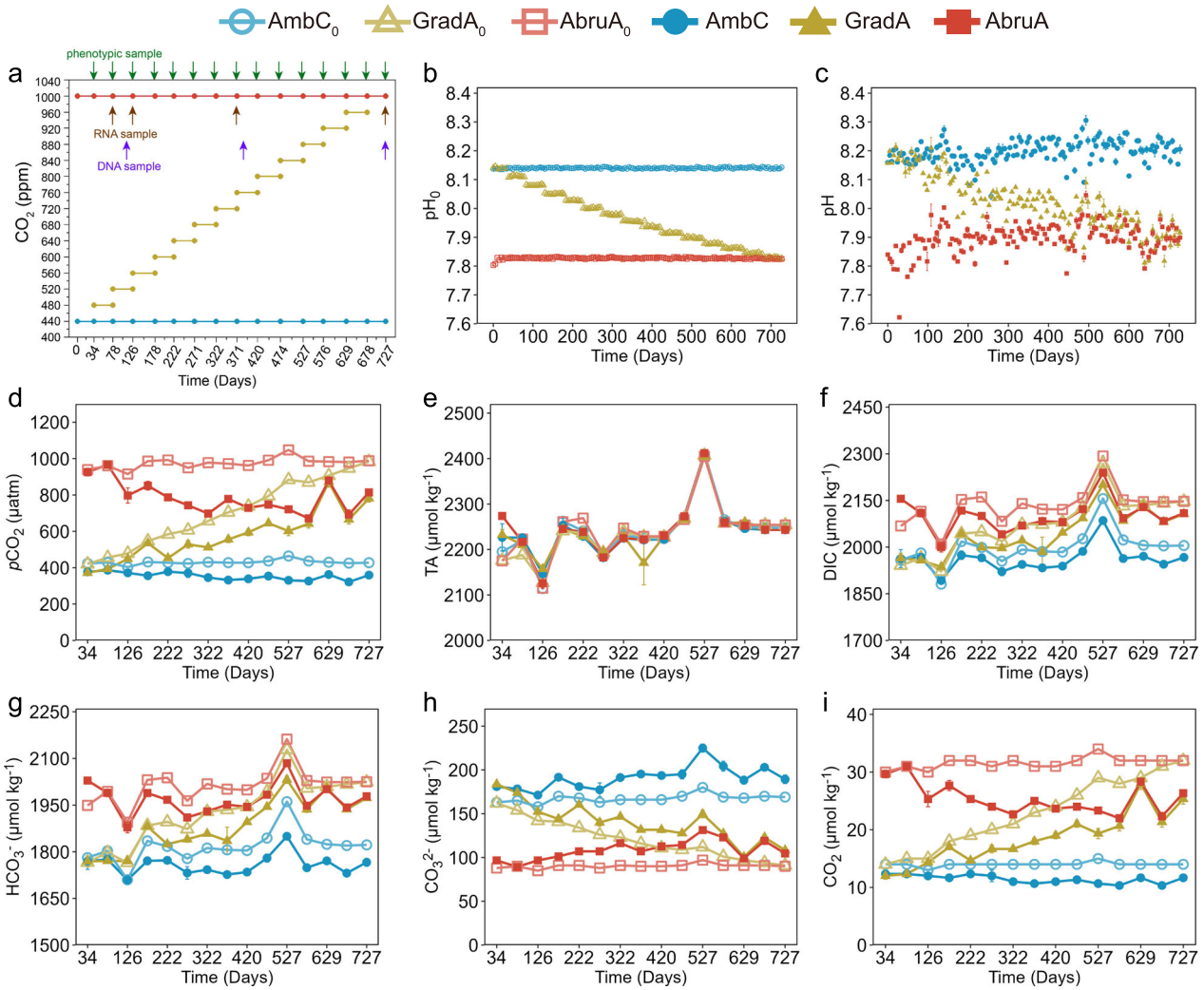

Fig. 1 | The experimental design and carbonate system of seawater before and after dilution. AmbC (ambient CO_2: 440 ppm), GradA (gradual increase of CO_2 from 440 to 1000 ppm with an interval of 40 ppm for each 70 generations), and AbruA (abrupt increase of CO_2 from 440 to 1000 ppm) represent different CO_2 treatment groups before dilution during the long-term culture. $AmbC_0$, $GradA_0$, and $AbruA_0$ denote the corresponding groups after dilution. Culture, sampling time, and the corresponding CO_2 concentration **a**; the initial pH of seawater in the diluted medium **b**; the pH of seawater before dilution **c**; the concentrations of pCO_2 **d**, TA **e**, DIC **f**, HCO_3^- **g**, CO_3^{2-} **h**, and CO_2 **i** in seawater before and after dilution. The pre-dilution data are presented as mean ± SEM ($n = 3$ biological replicates). Post-dilution parameters are reported as single measurements, as seawater for each treatment group was aliquoted from a common source into replicate culture polycarbonate bottles, precluding parallel biological replicates for chemical analysis. TA and pH are directly measured by the total alkalinity analyzer and pH meter, while other parameters are calculated using CO2SYS based on TA and pH. For more detailed information, please refer to the Supplementary Fig. 17 and Supplementary Data 2. Source data are provided as a Source data file.

certain genes. But it is unclear if those changes provide a fitness advantage.

To shed light on the differences in gene regulation between the different experimental conditions, transcriptome sequencing was conducted, and data were analyzed according to temporal changes of each individual transcript over the entire period of experimental evolution starting at day 78 (T1, the point of first CO_2 step completion in GradA, about 121–133 generations). The latter was chosen for T1 to account for the burn-in period, which usually represents the lag phase of evolution where more stochastic processes and non-adaptive responses might take place (e.g., founder effect). Four sampling points for RNA-seq were chosen, including T1 to represent the period of adaptive evolution under AbruA, GradA, and AmbC (Supplementary Figs. 2–6). Of all identified clusters, AbruA was the only experimental condition that caused significant divergence of gene expression changes over time (Figs. 6–7 and Supplementary Figs. 7–8) compared to GradA and AmbC. Notably, some of these divergences in gene expression persisted until the end of the evolution experiment (Day 727) with either consistently increased or decreased gene expression (Fig. 7d).

A GO-term analysis has revealed that for AbruA in cluster 1, for instance, the processes of photosynthesis, including light harvesting, inorganic carbon concentration, electron transport, and pigment synthesis, were most significantly enriched, with average gene expression being significantly higher compared to AmbC and GradA. In contrast, in cluster 3, genes under AbruA were consistently down-regulated from day 126 (205 to 218 generations) to the end of the experiment. Those genes were enriched in GO terms representing proton-transporting ATP synthase complexes and weaker enrichment in nucleoside triphosphate biosynthetic processes involved in DNA and RNA synthesis and energy transfer (Supplementary Fig. 9). Gene expression changes over time between GradA and AmbC were much more similar, with only minor differences in a small number of genes. In addition to regulatory changes over time, we also identified sets of genes, which were uniquely expressed under either AbruA or GradA in comparison to AmbC for all time points sampled throughout the

https://doi.org/10.1038/s41467-026-77753-0

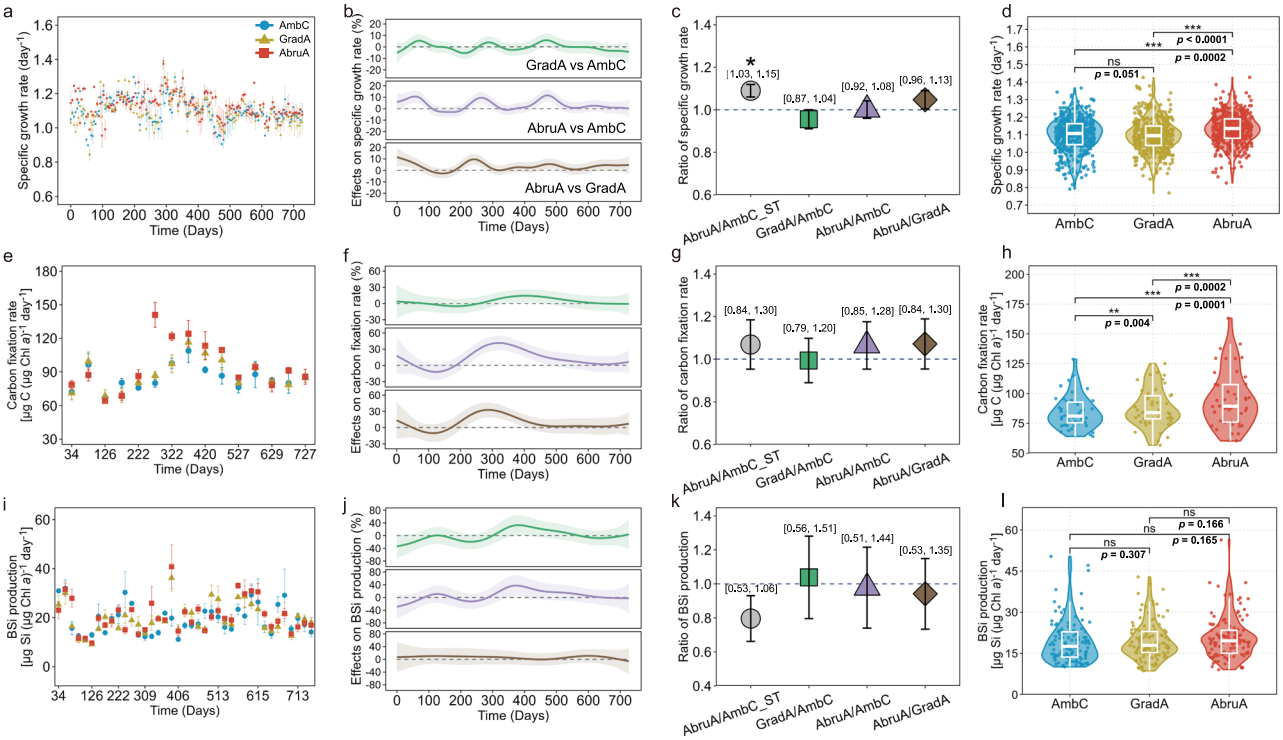

Fig. 2 | Acidification effects on specific growth rate, carbon fixation rate, and BSi production. The specific growth rate **a**, carbon fixation rate **e**, and BSi production **i** of *C. muelleri* under AmbC (~1160 generations), GradA (~1156 generations) and AbruA (~1189 generations) over the long-term acidification experiment ($n = 3$ biological replicates); effects of GradA and AbruA on specific growth rate **b**, carbon fixation rate **f**, and BSi production **j** based on GAMMs analysis, and the shaded areas represent the 95% confidence interval (CI). Intervals that exclude zero indicate statistically significant differences; GAMMs-predicted mean ratios of the specific growth rate **c**, carbon fixation rate **g,** and BSi production **k** between two groups. AbruA/AmbC_ST represents short-term acidification results on day 34, selected to ensure direct comparability with previous short-term ocean acidification studies[32,70]. GradA/AmbC, AbruA/AmbC, and AbruA/GradA all represent long-term acidification results on day 727. Values above the error bars indicate the 95% CI, and " * " denotes significance when the CI does not include 1 based on the statistical results of GAMMs analysis; violin and box plots showed the specific growth rate **d**, carbon fixation rate **h,** and BSi production **l** based on direct measurements during long-term culture. Data points represent pooled measurements from three biological replicates ($n = 3$ per time point) across all sampling time points, yielding 453, 45, and 117 total data points per treatment for **d**, **h**, and **l**, respectively (151, 15, and 39 time points × 3 biological replicates). Pairwise comparisons are performed using Wald-type *t*-tests based on GAMMs parameter coefficients (two-sided, no adjustment for multiple comparisons). In the box plots, the center line represents the median, box limits represent the 25th and 75th percentiles, and whiskers extend to the minimum and maximum values. AmbC, GradA, and AbruA represent ambient CO_2, gradual acidification mode, and abrupt acidification mode, respectively. Values in panels **a**, **c**, **e**, **g**, **i**, and **k** are presented as means ± SEM. Source data are provided as a Source data file.

evolution experiment. AbruA had 176 genes that were unique and therefore not expressed under either GradA or AmbC whereas GradA only had 85 of those uniquely expressed genes (Supplementary Fig. 10). For instance, genes uniquely regulated under AbruA encoded the following proteins: acetyl-CoA carboxylase/biotin carboxylase, electron transfer flavoprotein FAD-binding domain, radical SAM ThiC family, protein tyrosine kinase and a serine/threonine kinase. Those proteins are involved in the biosynthesis and metabolism of malonyl-CoA, acyl-CoA, purine nucleoside bisphosphate, nucleoside bisphosphate, diverse fatty acids, and sulfur compounds. In contrast, GO enrichment for GradA revealed proteins involved in DNA integration, proteolysis, removal of superoxide radicals, superoxide metabolic process, cellular oxidant detoxification, and diverse DNA metabolic processes. However, their enrichment was not statistically different (p-adj > 0.05), as with AmbC, but not for AbruA in relation to AmbC (Supplementary Figs. 9 and 11).

Projected *Chaetoceros*-based carbon and silicon synthesis in the global ocean

The projected global marine carbon fixation rate of *Chaetoceros* in 2091–2100 using the Community Earth System Model Version 2 (CESM2) is estimated to be 3.27 Pg C yr^{-1} under the SSP5-8.5 scenario. However, when the experimental results of the two acidification modes are superimposed, the carbon fixation rates were 3.48 Pg C yr^{-1} (GradA) and 3.99 Pg C yr^{-1} (AbruA), which is an increase of 6.4% [0.2%, 15.8%] and 22.1% [11.6%, 28.4%], respectively (Fig. 8a, b; Supplementary Table 4; Supplementary Fig. 12). The increase is particularly pronounced in the North Pacific, North Atlantic, eastern equatorial Pacific, and generally between 30–60°S and coastal waters (Fig. 8a, b). For example, abrupt acidification is predicted to increase carbon fixation rates by an average of 22.8% and 28.7% in the North Pacific and North Atlantic, respectively, whereas gradual acidification increased rates are only 6.6% and 8.3% according to our predictions. Like the carbon fixation rate, the BSi production is predicted to increase by 1.3% (GradA) [−11.4%, 13.1%] and 6.5% (AbruA) [−11.3%, 22.9%] respectively under the two modes. In the North Pacific and North Atlantic, the BSi production is predicted to increase by 6.3% and 12.2% respectively under the AbruA scenario, and by 1.2% and 2.4% respectively under the GradA scenario (Fig. 8c, d; Supplementary Table 4; Supplementary Fig. 13). Under the SSP5-8.5 scenario, the global average marine BSi/POC ratio of *Chaetoceros* is projected to be 0.14. Two acidification patterns are predicted to cause a decrease of 43.8% (under the GradA scenario) [−50.3%, −38.5%] and 45.0% (under the AbruA scenario) [−53.5%, −43.3%], respectively. However, coastal and open ocean BSi/POC effects differ significantly between the two acidification scenarios. For instance, the nearshore and pelagic decreases were 59.1% and 33.2%, respectively, at AbruA, and 56.7% and 33.0%, respectively, at GradA. The predicted reduction is most pronounced in the Antarctic

https://doi.org/10.1038/s41467-026-77753-0

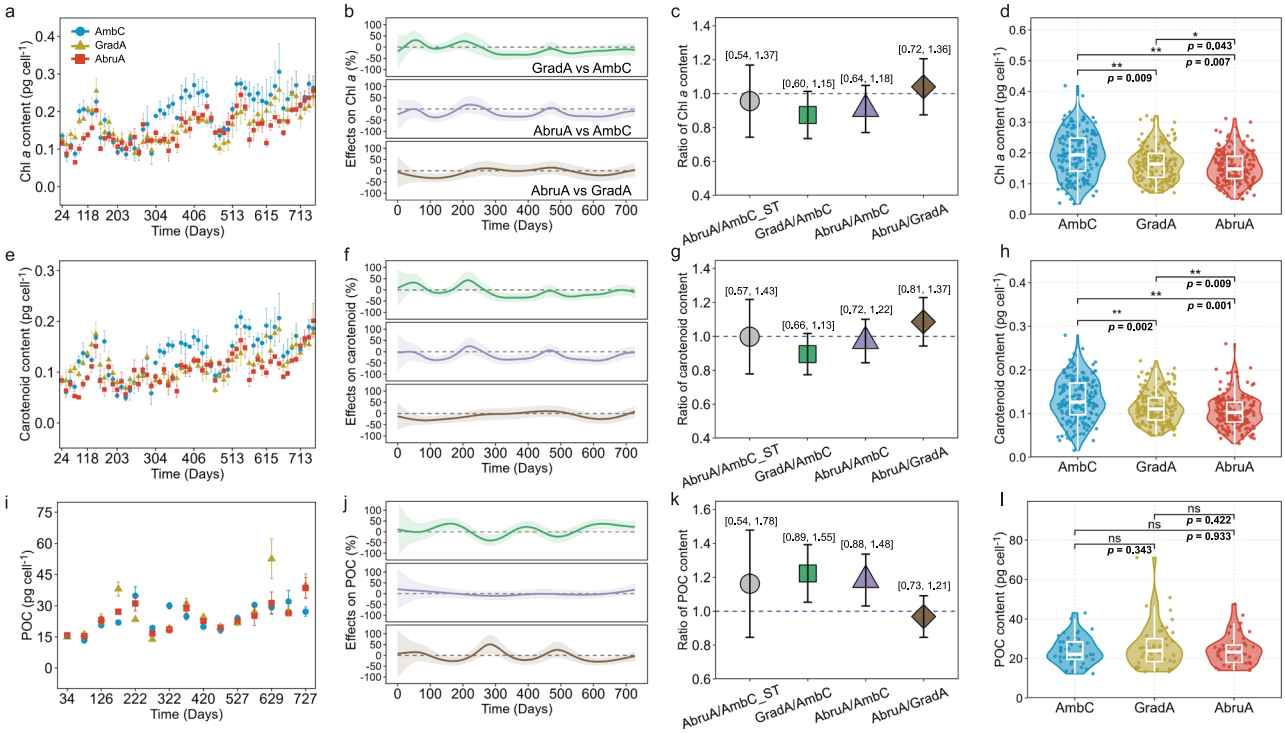

Fig. 3 | Acidification effects on cellular chlorophyll *a* (Chl *a*), carotenoid, and particulate organic carbon (POC) contents. The Chl *a* **a**, carotenoid **e,** and POC contents **i** of *C. muelleri* under AmbC (~1160 generations), GradA (~1156 generations) and AbruA (~1189 generations) over the long-term acidification experiment ($n = 3$ biological replicates); effects of GradA and AbruA on Chl *a* **b**, carotenoid **f,** and POC contents **j** based on GAMMs analysis, and the shaded areas represent the 95% confidence interval (CI). Intervals that exclude zero indicate statistically significant differences; GAMMs-predicted mean ratios of Chl *a* **c**, carotenoid **g,** and POC contents **k** between two groups. AbruA/AmbC_ST represents short-term acidification results on day 34, selected to ensure direct comparability with previous short-term ocean acidification studies[32,70]. GradA/AmbC, AbruA/AmbC, and AbruA/GradA all represent long-term acidification results on day 727. Values above the error bars indicate the 95% CI, and " * " denotes significance when the CI does not include 1 based on the statistical results of GAMMs analysis; violin and box plots showed the Chl *a* **d**, carotenoid **h**, and POC **l** contents based on direct measurements during long-term culture. Data points represent pooled measurements from three biological replicates ($n = 3$ per time point) across all sampling time points, yielding 180, 180, and 45 total data points per treatment for panels **d**, **h**, and **l**, respectively (60, 60, and 15 time points × 3 biological replicates). Pairwise comparisons are performed using Wald-type *t*-tests based on GAMMs parameter coefficients (two-sided, no adjustment for multiple comparisons). In the box plots, the center line represents the median, box limits represent the 25th and 75th percentiles, and whiskers extend to the minimum and maximum values. AmbC, GradA, and AbruA represent ambient CO_2, gradual acidification mode, and abrupt acidification mode, respectively. Values in panels **a**, **c**, **e**, **g**, **i**, and **k** are presented as means ± SEM. Source data are provided as a Source data file.

regions (as shown in Fig. 8e, f; Supplementary Table 4; Supplementary Fig. 14) where the predicted BSi/POC ratio will decline by 25.1% and 26.0% respectively under the AbruA and GradA scenarios. It is important to note, however, that this modeling framework does neither account for potential nonlinear responses to ocean acidification nor incorporate pigment changes, a limitation we acknowledge and plan to address in future work using either an online CESM framework or mechanistic models.

Discussion

While our work confirms previous studies that ocean acidification enhances diatom carbon fixation, we provide novel evidence that the design of experiments to simulate OA directly drives phenotypic shifts and adaptive/evolutionary responses in the model diatom *Chaetoceros muelleri* and therefore potentially other diatom species. Not only long-term experimental evolution but also a gradual increase of CO_2 in laboratory experiments is likely required to best match the natural process of ocean acidification. Our integrative approach has not only revealed that an abrupt change in CO_2 overestimates the response of *C. muelleri*, but it also provides insights as to why, including potential implications for future research and Earth System Models which usually take these laboratory results into account[18–20].

Previous studies that have employed abrupt acidification showed that elevated CO_2 stimulates growth of diatoms and marine primary production[21]. Consistent with these findings, our work also found that abrupt acidification increased specific growth rate (+6.5%), especially early in our evolution experiments (e.g., day 34). This early response to abrupt ocean acidification can be attributed mainly to two aspects. The first is the direct effect on photosynthesis. Seawater is CO_2 limited for algal photosynthesis, and thus a sudden increase in the CO_2 concentration can accelerate photosynthetic carbon fixation, confirmed by increased carbon fixation rates of *C. muelleri* in our study[9,22]. This fast response is likely underpinned by the upregulation of genes (e.g., *PetH*) related to photosynthetic electron transport, indicating that light reactions of photosynthesis were enhanced by acidification, which may have been caused by enhanced CO_2 fixation of RUBISCO to match the demand of electrons and protons for the synthesis of NADPH and ATP, respectively[23]. The second aspect likely is related to a reduced demand for converting bicarbonate into CO_2, commonly known as the inorganic carbon concentrating mechanisms (CCMs) in algae[24]. Consequently, genes encoding proteins involved in the CCMs were downregulated subsequent to abrupt acidification (Fig. 6). The saved energy from downregulated CCMs genes and related processes (e.g., pigment synthesis and light harvesting) possibly was reallocated to other metabolic processes such as nitrogen assimilation (Supplementary Fig. 15) and therefore supporting cell division (e.g., cell-cycle genes were upregulated) and growth[25]. In addition, increased glycolysis and TCA cycle activity might have mitigated acid-based disturbance of cellular homeostasis caused by abrupt acidification[26]. While we acknowledge that the peak in carbon fixation around day 250

https://doi.org/10.1038/s41467-026-77753-0

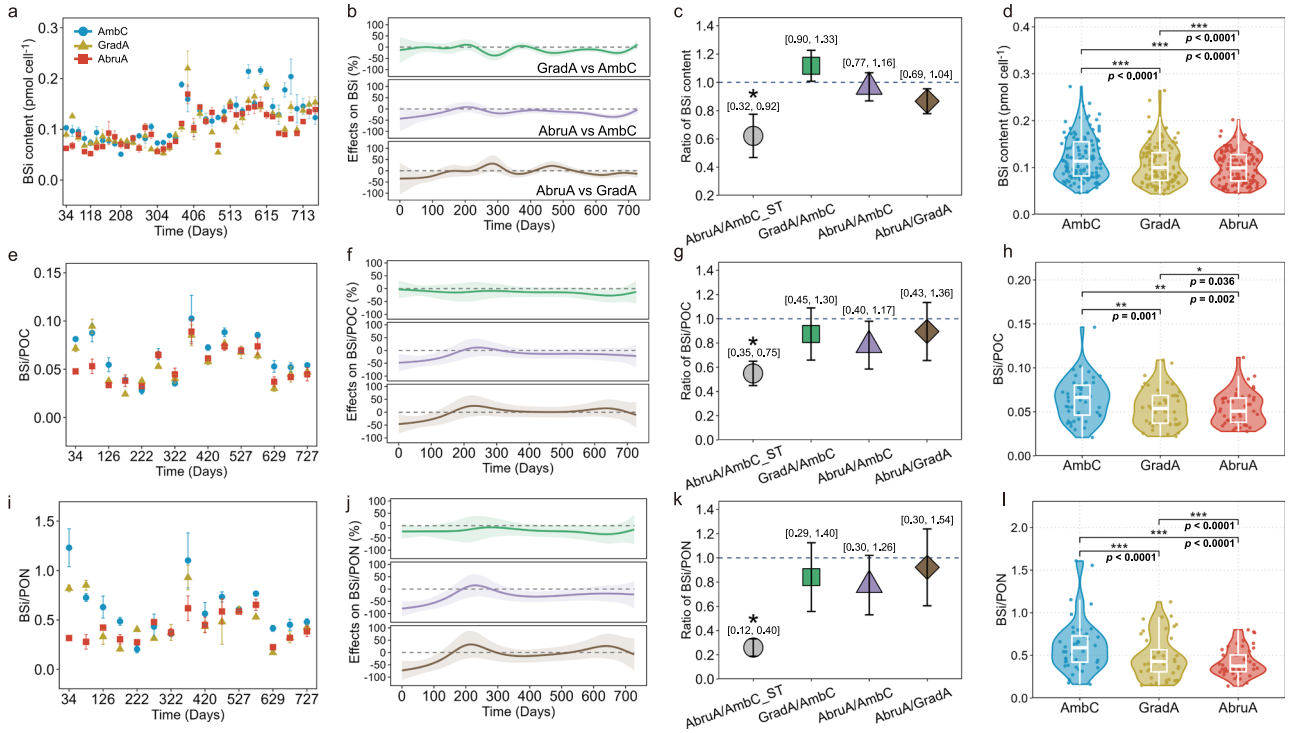

Fig. 4 | Acidification effects on biogenic silica (BSi) content, BSi/POC, and BSi/PON. The BSi content **a**, BSi/POC **e,** and BSi/PON **i** of *C. muelleri* under AmbC (~1160 generations), GradA (~1156 generations) and AbruA (~1189 generations) over the long-term acidification experiment ($n = 3$ biological replicates); effects of GradA and AbruA on BSi content **b**, BSi/POC **f,** and BSi/PON **j** based on GAMMs analysis, and the shaded areas represent the 95% confidence interval (CI). Intervals that exclude zero indicate statistically significant differences; GAMMs-predicted mean ratios of the BSi content **c**, BSi/POC **g,** and BSi/PON **k** between two groups. AbruA/AmbC_ST represents short-term acidification results on day 34, selected to ensure direct comparability with previous short-term ocean acidification studies[32,70]. GradA/AmbC, AbruA/AmbC, and AbruA/GradA all represent long-term acidification results on day 727. Values above the error bars indicate the 95% CI, and " * " denotes significance when the CI does not include 1 based on the statistical results of GAMMs analysis; violin and box plots showed the BSi content **d**, BSi/POC **h**, and BSi/PON **l** based on direct measurements during long-term culture. Data points represent pooled measurements from three biological replicates ($n = 3$ per time point) across all sampling time points, yielding 126, 45, and 45 total data points per treatment for panels **d**, **h**, and **l**, respectively (42, 15, and 15 time points × 3 biological replicates). Pairwise comparisons are performed using Wald-type *t*-tests based on GAMMs parameter coefficients (two-sided, no adjustment for multiple comparisons). In the box plots, the center line represents the median, box limits represent the 25th and 75th percentiles, and whiskers extend to the minimum and maximum values. AmbC, GradA, and AbruA represent ambient CO_2, gradual acidification mode, and abrupt acidification mode, respectively. Values in panels **a**, **c**, **e**, **g**, **i**, and **k** are presented as means ± SEM. Source data are provided as a Source data file.

significantly impacts the average, the overall decoupling between carbon fixation and growth across the experiment likely reflects a metabolic shift. Fixed carbon appears to be preferentially allocated to respiratory energy production and lipid turnover rather than biomass accumulation, likely driven by oxidative stress and nitrogen assimilation. Gradual acidification, however, did not affect the growth rate of *C. muelleri*, but significantly enhanced carbon fixation compared to AmbC. Interestingly, this increase in carbon fixation occurred without any significant changes in the expression of genes involved in CCMs, pigment synthesis, photosynthetic electron transport, and the Calvin cycle, unlike the rapid transcriptional adjustments commonly observed in previous studies and under AbruA[27,28], suggesting a potentially distinct acclimation strategy under gradual, multigenerational exposure. This decoupling between physiological and transcriptional responses indicates that the observed enhancement of carbon fixation was likely driven by post-translational regulation, such as adjusting the activity of enzymes involved in photosynthesis, rather than by de novo gene expression, as the abundance of many proteins in diatoms is transcriptionally regulated[29–31]. Thus, it appears that the gradually increased availability of CO_2 enhanced the catalytic efficiency of key enzymes such as RUBISCO, or altered the allocation of fixed carbon towards storage compounds without requiring transcriptional reprogramming under GradA, whereas AbruA caused transcriptional adjustments. Thus, these results highlight distinct physiological responses of *C. muelleri* to abrupt versus gradual acidification, suggesting that laboratory experiments based on abrupt acidification may not accurately reflect how diatoms respond to the gradual long-term process of natural ocean acidification. In addition, the sinusoidal pattern observed in growth rates possibly reflects intrinsic biological rhythms[32] or transient adaptive dynamics[33], especially during the early phase of gradual and multigenerational exposure to elevated CO_2, a phenomenon that was not observed in short-term experiments[27,34]. Thus, the synchronized fluctuations of growth, chlorophyll *a*, and carotenoids possibly reflect a coordinated biological response rather than a methodological artifact, a conclusion further supported by our GAMMs analysis, which accounts for non-linear temporal patterns.

In addition to CO_2, silicate [Si $(OH)_4$] is another essential nutrient, especially for diatoms. Diatoms acquire silicate via silicate transporters (SIT) to produce silicified cell walls, which are considered to have multiple functions including defense, to provide ballast for sinking and carbon export, and for pH buffering essential to optimize extracellular carbonic anhydrase activity[35,36]. However, AbruA reduced the silica content of *C. muelleri* likely leading to a decreased ballast effect[6,37]. Although our work cannot explain why, we found downregulation of SITs, suggesting a lower demand of silicate post-AbruA. Although overall BSi production has not been affected by AbruA, due to an increased growth rate, reduced cellular silica quotas decreased the ratio of BSi/POC. These results therefore mislead if used to interpret the response of diatoms to the natural process of ocean acidification

https://doi.org/10.1038/s41467-026-77753-0

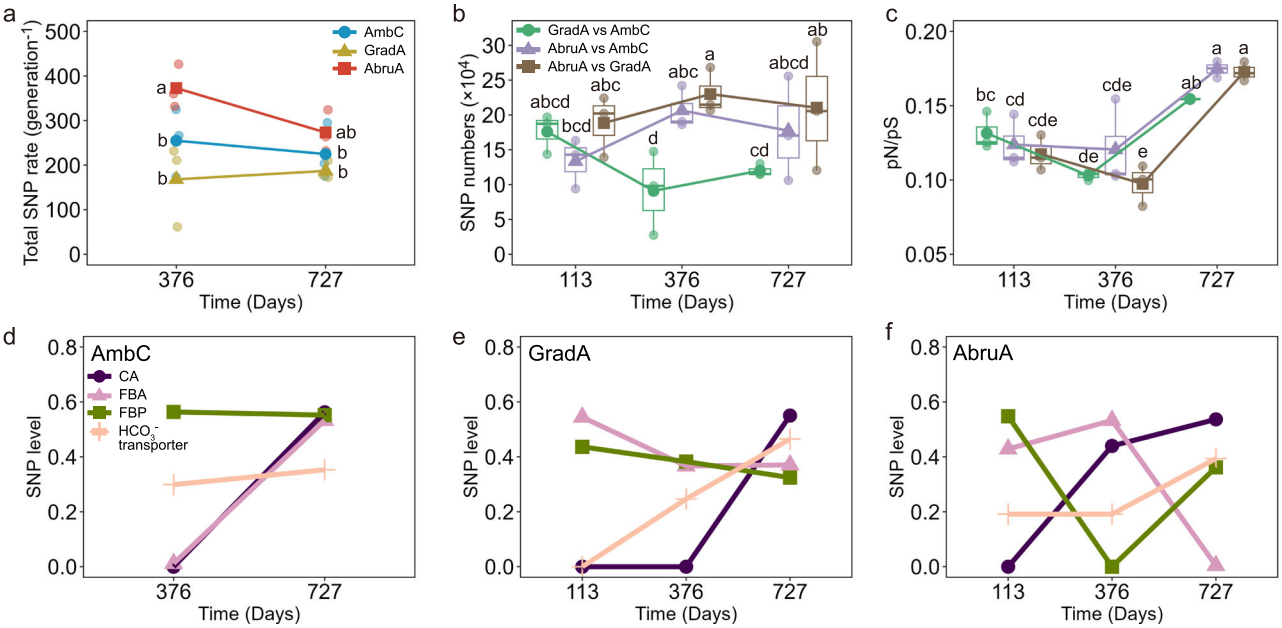

Fig. 5 | Genetic mutation of *C. muelleri* after long-term adaptation to CO_2 conditions. The rates of total SNPs on days 376 and 727 **a**, SNP numbers **b**, pN/pS **c**, SNP levels of critical genes in AmbC **d**, GradA **e**, and AbruA **f** on days 113, 376, and 727. Total SNP rate = Total SNP numbers/generation. SNP level = SNP counts/gene length. In **a**–**c**, different lowercase letters indicate significant differences between treatments ($p < 0.05$, one-way ANOVA with LSD post-hoc tests, two-sided, no adjustment for multiple comparisons; ANOVA F and p values: **a**, $F(5, 12) = 4.152$, $p = 0.020$; **b**, $F(8, 18) = 2.397$, $p = 0.059$; **c**, $F(8, 18) = 12.312$, $p = 7.05 \times 10^{-6}$). In the box plots, the center line represents the median, box limits represent the 25th and 75th percentiles, and whiskers extend to the minimum and maximum values. Individual data points are overlaid as dots ($n = 3$ biological replicates). AmbC, GradA, and AbruA represent ambient CO_2, gradual acidification mode, and abrupt acidification mode, respectively. The full names of genes are shown in Supplementary Table 7. Source data are provided as a Source data file.

because one could conclude that a lower ratio of BSi/POC would lead to an increase in POC remineralization in the upper oceans and therefore reduced diatom-based carbon sequestration under ocean acidification[27,38]. Additionally, the decreased BSi/PON ratio can be interpreted as being the consequence of enhanced conversion of PON (Supplementary Fig. 15) to inorganic nitrogen, potentially accelerating nitrogen cycle in upper oceans and thus alleviating nutrient limitation[39,40]. Compared to AbruA, GradA had a weaker effect on the BSi/POC and BSi/PON ratios. These results, especially in light of some previous studies claiming that ocean acidification could reduce silicification of diatom communities by more than 60% when fCO_2 increased from 343 to 1641 μatm[41], underscore a word of caution in terms of how much the natural process of ocean acidification affects diatom-driven biogeochemical cycles. Likely, the effects are rather minor, without a drastic impact on diatom-driven biogeochemical cycles under ocean acidification.

The underlying mechanism for why there is a different response between AbruA and GradA appears to be related to how a disruption of the environment is being sensed and processed by *C. muelleri*. Our results show that abrupt changes of environmental conditions lead to divergent genomic signals based on a higher number of SNPs, increased pN/pS, and differential expression of genes underpinning adaptive processes as a response to elevated CO_2 (e.g., photosynthesis, protein synthesis). Hence, these genomic and transcriptomic changes enable *C. muelleri* to respond quickly, which corroborates previous work to explain their "bloom and bust" behavior in the oceans[42,43]. For instance, genes encoding Chl *a* biosynthesis, enzymes of the Calvin cycle, TCA cycle, and linear electron transport were upregulated subsequent to abrupt CO_2 increase (Fig. 6). And some of these genes remained upregulated (e.g., pyruvate metabolism, ribosome biogenesis in eukaryotes, porphyrin metabolism, purine metabolism, ABC transporters, arginine and proline metabolism) even after long-term adaptive evolution (Figs. 6–7). However, GradA did only show a very weak response compared to ambient CO_2 in terms of SNPs, pN/pS, and changes in the transcriptome. These results suggest that a gradual increase in CO_2 over a longer period of time similar to the natural process of ocean acidification, can likely be accommodated by adjusting metabolic fluxes through posttranscriptional and even translational processes, such as the modification of enzyme activity[28,44]. At the very least, our data suggest that additional CO_2 added in small doses to an evolving *C. muelleri* population does not induce adaptive divergences and significant transcriptional remodeling, which usually has an impact on the phenotypes. Hence, the natural process of ocean acidification does not appear to be a strong selecting agent like temperature or the growth limitation by essential nutrients such as iron and nitrate.

Interestingly, the relatively fast response following AbruA appears to leave a signal in the genomes and transcriptomes of *C. muelleri,* as evidenced by persistent SNPs and differential gene expression until the end of the evolution experiment (Day 727) (Fig. 5). This strongly suggests that, compared to GradA, AbruA not only differentially impacts the phenotypes of diatoms through quantitative changes in SNPs located in coding parts and their related transcripts, but also indicates that regulatory components of gene expression might have been changed (e.g., promoters, transcription factors)[45,46]. As some of these changes appear to persist throughout the entire evolution experiment (Figs. 6–7; Supplementary Figs. 2–6), it suggests that abrupt environmental changes can induce persistent phenotypic changes that are maintained over successive cell divisions. However, we recognize that the observed differences between treatments cannot be attributed solely to evolutionary divergence because genomic evidence does not yet allow inference of parallel evolutionary change across independent replicate lines. Instead, they likely reflect a complex interplay between plastic phenotypical responses and genetic changes over hundreds of generations. Future studies therefore will be necessary to fully disentangle phenotypic plasticity and genetic change to reveal their individual contributions to a gain in fitness for coping with environmental change such as gradual ocean acidification.

https://doi.org/10.1038/s41467-026-77753-0

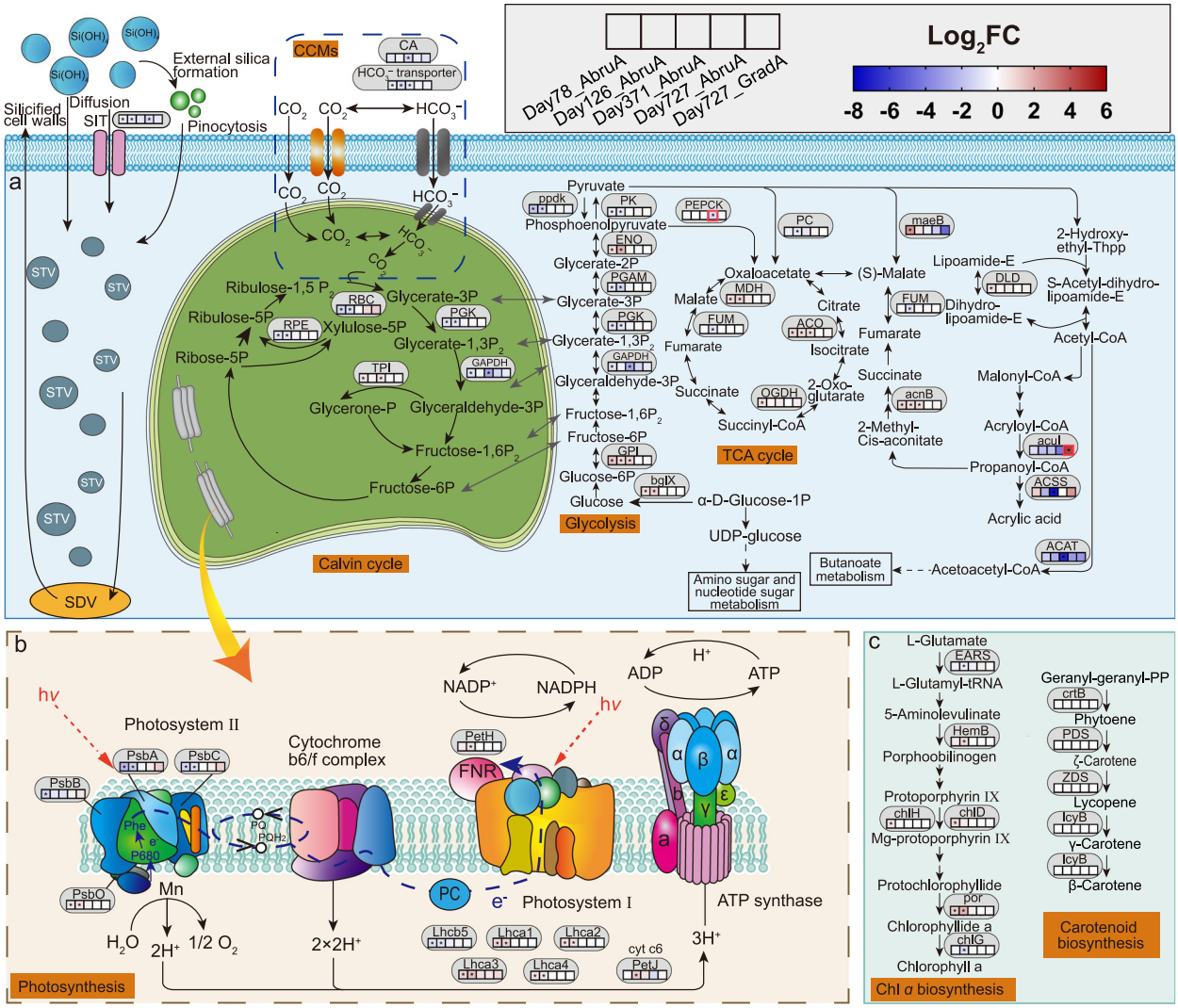

Fig. 6 | Transcriptomic responses of carbon metabolism and photosynthesis related pathways to different acidification modes. Responses of carbon metabolism **a**, photosynthetic electron transport **b**, and pigment synthesis pathways **c** to different acidification modes over the long-term culture using transcriptome data at four time points. The five squares from left to right below the gene name indicate AbruA vs AmbC at day 78, AbruA vs AmbC at day 126, AbruA vs AmbC at day 371, AbruA vs AmbC at day 727, GradA vs AmbC at day 727. " * " indicates a significant difference (|Fold Change| ≥ 1.5 & *p*-adj < 0.05, based on two-sided Wald tests with Benjamini–Hochberg correction for multiple comparisons). AmbC, GradA, and AbruA represent ambient CO_2, gradual acidification mode, and abrupt acidification mode, respectively. The full names of genes are shown in Supplementary Table 7.

However, the differential molecular responses between AbruA and GradA undoubtedly match the phenotypic response of *C. muelleri* with consequences for our understanding of how they drive biogeochemical cycles of carbon and silicon. For instance, results based on the AbruA experiments overestimate projected carbon fixation of *C. muelleri* by more than three-fold across much of the global ocean, with consequences for BSi/POC ratios and therefore diatom-driven biogeochemical cycles. Whether the observed responses to elevated CO_2 are accompanied by evolutionary trade-offs—where adaptation to high CO_2 reduces performance under low CO_2 conditions—remains an open question. Testing this hypothesis using the retained cultures from our experiment represents a valuable direction for future research.

In summary, although our work confirms previous work in general, an abrupt decline of pH and/or increase of CO_2 in laboratory experiments overestimates the contribution of *C. muelleri* and likely other *Chaetoceros* species in mitigating the effect of ocean acidification. The reason for this overestimation is that the evolutionary response between a sudden change in environmental conditions (AbruA) is accommodated differently compared to a gradual change (GradA) that takes place over decades or even centuries. Hence, matching experimental conditions in the laboratory to natural processes within given constraints may provide more accurate estimates of how key primary producers such as diatoms respond to natural processes. This, in turn, will provide more accurate predictions of how natural populations of diatoms will mitigate the effect of ocean acidification through regulating biogeochemical cycles and carbon sequestration. It is worth noting this study did not consider the interactive effects between ocean acidification and other drivers, including species competition, food web dynamics, and climate-induced changes in resource availability. Exploring these interactions will be an important direction for future work. Moreover, future work using multiple diatom species will be essential to determine whether the observed patterns in *C. muelleri* hold across the entire class of diatoms. If so, it would have consequences for how the response of diatoms to OA needs to be represented in Earth System Models.

Methods

Cell culture

Chaetoceros muelleri (MMDL50116) was obtained from the diatom laboratory in the College of Life Sciences of Xiamen University, which

https://doi.org/10.1038/s41467-026-77753-0

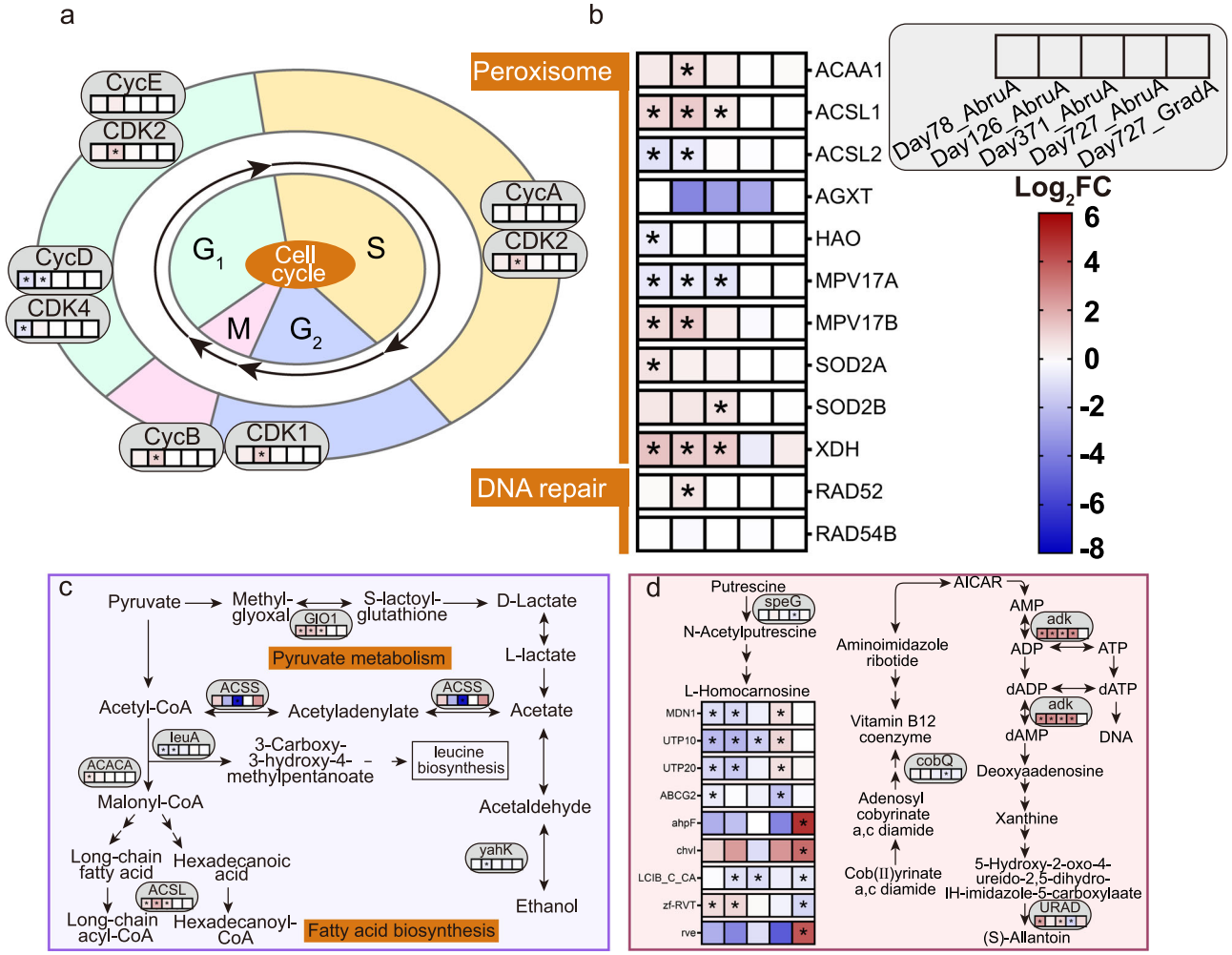

Fig. 7 | Transcriptomic responses of cell cycle, DNA repair, and key metabolic pathways to different acidification modes. Responses of cell cycle **a**, peroxisome, DNA repair **b**, pyruvate metabolism, fatty acid biosynthesis **c**, and nitrogen metabolism **d** to different acidification modes over the long-term culture using transcriptome data at four time points. The five squares from left to right below the gene name indicate AbruA vs AmbC at day 78, AbruA vs AmbC at day 126, AbruA vs AmbC at day 371, AbruA vs AmbC at day 727, GradA vs AmbC at day 727. " * " indicates a significant difference (|Fold Change| ≥ 1.5 & p-adj < 0.05, based on two-sided Wald tests with Benjamini–Hochberg correction for multiple comparisons). AmbC, GradA, and AbruA represent ambient CO_2, gradual acidification mode, and abrupt acidification mode, respectively. The full names of genes are shown in Supplementary Table 7.

was originally isolated from coastal waters of the East China Sea near Xiamen. Three different CO_2 regimes were set, including AmbC (ambient CO_2: 440 ppm), AbruA (abrupt increase of CO_2 from 440 to 1000 ppm), and GradA (gradual increase of CO_2 from 440 to 1000 ppm in steps of 40 ppm every 70 generations). Although we have made every effort to simulate the natural ocean acidification process, the CO_2 ramping rate in the GradA group was 37 times higher than the IPCC-projected rate for ocean acidification toward 1000 ppm by 2100, due to the limitations in instrument precision and culture duration[47,48]. Specifically, we achieved the three required CO_2 concentrations by separately bubbling three bottles of autoclaved seawater with three customized CO_2 enrichers (0A100, Ruihua, Wuhan, China) that can output set CO_2 levels. After each treatment's seawater was saturated with the respective gas, it was dispensed into three polycarbonate bottles ($n = 3$ biological replicates), with algal inoculum and seawater added according to the dilution concentration (sealed culture). This approach ensured that all biological replicates within each treatment had the same CO_2 concentration. To avoid confounding treatment effects with physical differences between incubators (e.g., in temperature or light intensity), all cultures (nine bottles) were then transferred to a single illuminating incubator (GXZ-380E, Jiangnan, Ningbo, China) with a cool white fluorescent light of 200 μmol photons m^{-2} s^{-1} (US-SQS/WB, Walz, Germany), randomly swapped positions and shaken every 4 h during the light period of the light-dark cycle. A simple schematic has also been provided in Supplementary Fig. 16 to further clarify the experimental setup. Cells were semi-continuously cultured in 500 mL polycarbonate bottles using sterilized seawater supplemented with f/20 medium[49], under a 12 h light/12-h dark cycle at 20 °C, with dilution every five days. To prevent the influence of headspace air on the carbonate system within the culture setup, we actually filled the 500 mL polycarbonate bottles with 600 mL of algal culture solution. Moreover, the semi-continuous culture regime was implemented to maintain stable pH and pCO_2 throughout each culture cycle. The dilution factor was determined based on the algal specific growth rate, ensuring that after five days of cultivation, the cell density did not exceed 3×10^4 cells mL^{-1} when starting from an initial concentration of 150–300 cells mL^{-1}. Under this regime, approximately 0.5–1% of the algal population was transferred to fresh, pre-equilibrated medium at each cycle, serving as the inoculum for the subsequent growth phase. The variation of CO_2 concentration with time is shown in Supplementary Data 1.

https://doi.org/10.1038/s41467-026-77753-0

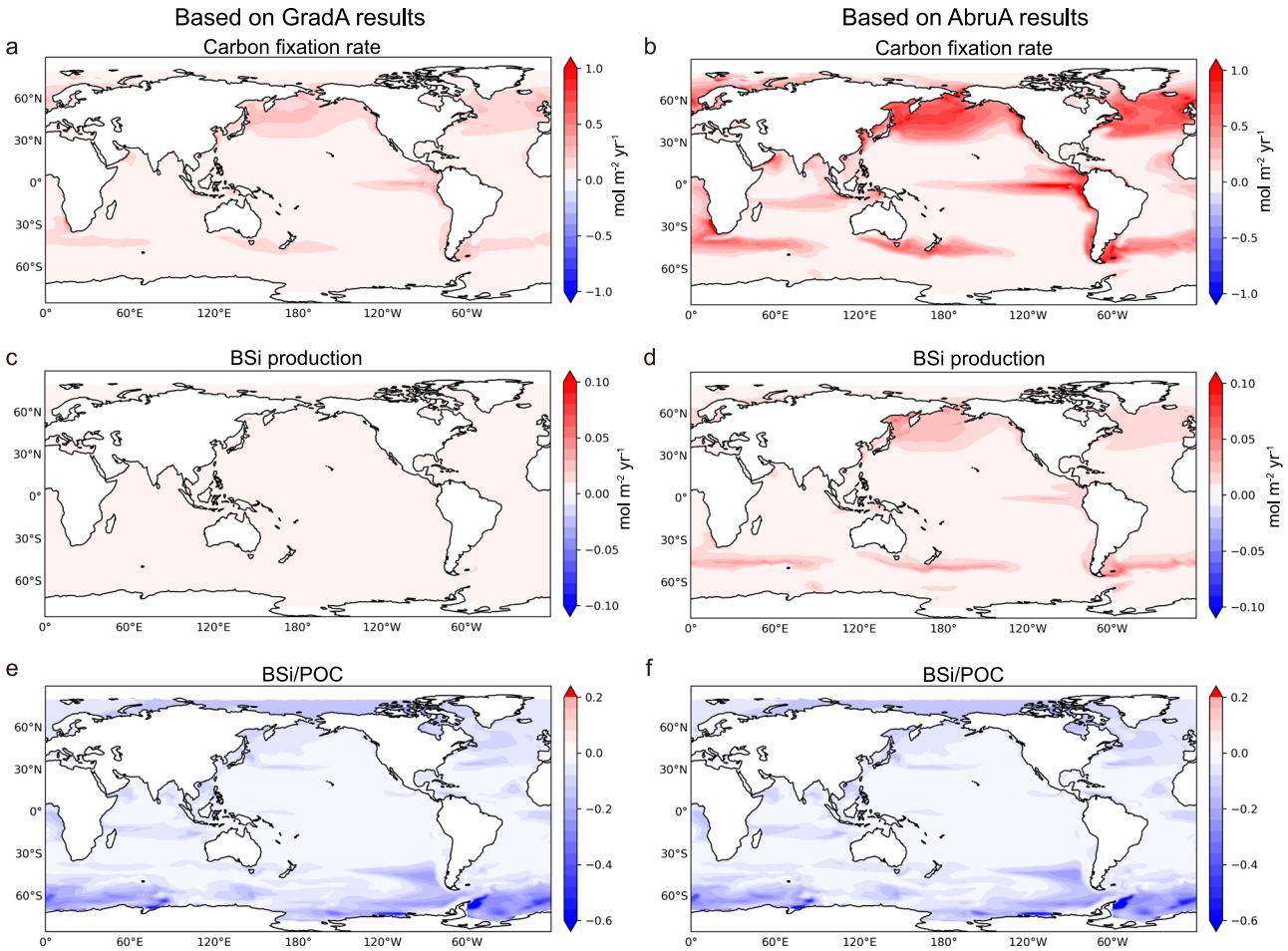

Fig. 8 | Modeled contributions of *Chaetoceros* to global marine carbon fixation and biogenic silica (BSi) production in different acidification modes. Projected changes in carbon fixation rate (**a**, **b**), BSi production (**c**, **d**), and BSi/POC (**e**, **f**) of global marine *Chaetoceros* caused by ocean acidification in 2091–2100. Panels **a**, **c**, **e** are based on gradual acidification (GradA) results, while **b**, **d**, **f** are based on abrupt acidification (AbruA) results.

Carbonate chemistry in cultures

The pH_{NBS} was measured before and after dilution of the culture medium using a pH meter (Orion StarA211, Thermo, Indonesia) that was three-point calibrated with NBS buffer, to ensure that the difference in pH before and after dilution was within 0.2. In practice, the difference in pH was less than 0.1 units in more than 80% of all data points. Total alkalinity was also measured using a Total Alkalinity Titrator (AS-ALK2, Apollo SciTech, USA). Other parameters (pCO_2, DIC, HCO_3^-, CO_3^{2-}, CO_2, ΩCa, ΩAr) of the carbonate system were calculated using the CO2SYS software[50] from pH_{NBS} and TA (Supplementary Data 2). ΩCa and ΩAr are presented in Supplementary Fig. 17. Because seawater for each treatment was aliquoted from a common source, post-dilution carbonate chemistry parameters are reported as single measurements per treatment, and bottle-level variability was not directly measured but inferred from the pre-dilution distributions.

Nutrients

In order to ensure that there was no nutrient restriction in the semi-continuous culture process, we randomly measured the nutrient concentrations in the culture system. The filtrate filtered through a 0.45 μm cellulose acetate lipid membrane was collected in a 50 mL centrifuge tube, and then stored directly in the refrigerator at −20 °C. The sample was determined by an AA3 Auto-Analyzer (SEAL, Germany) (Supplementary Table 5). The concentrations of NO_X, NO_2^- and PO_4^{3-} were obtained directly by AA3 Auto-Analyzer, while the concentration of SiO_3^{2-} was corrected by the following formula:

$$SiO_3^{2-} = \frac{SiO_3^{2-}\ \text{measured}}{-0.00023 \times \text{Salinity} + 0.9898} \quad (1)$$

Where SiO_3^{2-} measured is the direct result of AA3 Auto-Analyzer and Salinity is the corresponding salinity (‰).

Specific growth rate

Cell concentration was measured using a Z-2™ Coulter particle Counter (Z2, Beckman Coulter, USA) before the dilutions, and the particle size range was set as 3–9 μm. The specific growth rate (μ, day^{-1}) was calculated using the following equation:

$$\mu = \frac{\ln N_1 - \ln N_0}{t_1 - t_0} \quad (2)$$

Where N_1 and N_0 represent cell concentrations at t_1 (before the dilutions) and t_0 (initial), respectively.

Carbon fixation rate

The carbon fixation rate was determined by the ^{13}C isotope tracer method. Carbon stable isotope compositions were measured with an Isotope Ratio Mass Spectrometer (Thermo Scientific Delta V

Advantage, Germany) coupled with an Elemental Analyzer (Flash IRMS, Thermo, Germany). L-glutamic acid USGS 40 and USGS41a were used as the certified reference materials. The carbon fixation rates were calculated by using similar equations proposed by Hama et al.[51].

Chlorophyll *a* (Chl *a*) and carotenoids

Fresh cells were filtered onto GF/F membranes (Whatman, 25 mm) by suction filtration within a pressure of 0.02 MPa to prevent cell damage from excessive pressure. And then, 4 mL of absolute methanol was added to extract pigment at 4 °C for 24 h in darkness. The extracts were centrifuged at 8000 × *g* for 10 min at 4 °C (320 R, Hettich, Germany). The absorption values of the supernatant were measured with a UV-VIS spectrophotometer (TU-1810, PERSEE, China) across the wavelength range 400–800 nm. The contents of Chl *a* and carotenoid were then estimated according to the formulas of Porra et al.[52] and Strickland and Parsons[53], respectively.

Particulate organic carbon (POC) and particulate organic nitrogen (PON)

Cell samples for POC and PON were collected onto pre-combusted Whatman GF/F filters (450 °C, 4 h) and frozen at −20 °C. Before measuring, the filter membranes were exposed to HCl for 24 hours to completely remove any inorganic carbon and then dried at 60 °C for 24 h. The dried samples were analyzed by an elemental analyzer (Vario EL cube, Elementar, Germany).

Biogenic Silica (BSi) content and production

Cell samples for biogenic silica measurement were collected onto 0.8 μm polycarbonate membranes (25 mm, Isopore, Ireland). They were then dried at 60 °C for 24 h and stored at dry room temperature until measured. The biogenic silica contents were measured by molybdenum-blue spectrophotometry[54]. BSi production was calculated by BSi content multiplied by the specific growth rate [μ (day^{-1})] and Chl *a* content [Chl *a* ($pg\ cell^{-1}$)], which was expressed in [μg Si (μg Chl *a*) $^{-1}$ day^{-1}], and the calculation formula is as follows:

$$\text{BSi production}\left[\mu g\ Si\ (\mu g\ Chl\ a)^{-1}\ day^{-1}\right] = \frac{\text{BSi content}\ (pg\ cell^{-1})}{\text{Chl}\ a\ (pg\ cell^{-1}) \times \mu\ (day^{-1})} \quad (3)$$

The ratios of BSi/POC and BSi/PON were obtained by dividing BSi content by POC and PON content, respectively, with the calculation formulas provided below:

$$\text{BSi/POC} = \frac{\text{BSi content}\ (pg\ cell^{-1})}{\text{POC content}\ (pg\ cell^{-1})} \quad (4)$$

$$\text{BSi/PON} = \frac{\text{BSi content}\ (pg\ cell^{-1})}{\text{PON content}\ (pg\ cell^{-1})} \quad (5)$$

Transcriptome

To explore the mechanisms by which *C. muelleri* responded to different acidification modes during the long-term experiment, transcriptome sequencing was performed at four time points: Days 78, 126, 371, and 727. Day 78 was selected as the initial time point because it marked the completion of the first CO_2 elevation step in the gradual acidification (GradA) treatment, representing the point at which all three treatments were actively under their respective CO_2 regimes. For each treatment group at each time point, all biological replicate samples were subjected to RNA-seq. To prevent RNA degradation, the samples from the four time points were shipped for sequencing in three separate batches: samples from Day 78 and Day 126 were shipped together, while samples from Day 371 and Day 727 were each shipped immediately after collection.

Total RNA was isolated and purified using the modified CTAB (Bio Basic, Toronto, Canada) method. Tissue samples were ground to a fine powder in liquid nitrogen and then transferred to 1.5 mL of preheated CTAB lysis reagent (add 2% β-mercaptoethanol) at 65 °C, followed by incubation at 65 °C for 15 min. After incubation, the tube was cooled to room temperature and centrifuged at 4 °C with 12,000 × *g* for 5 min. The supernatant was transferred into a new 2.0 mL EP tube. Chloroform/isoamyl alcohol (24:1) was added at a volume of 200 μL per mL of CTAB buffer, and the mixture was vortexed and centrifuged at 12,000 × *g* for 10 min at 4 °C. Next, the clear upper aqueous layer was transferred to a 2.0 mL EP tube, followed by an equal-volume extraction with phenol/chloroform/isoamyl alcohol (25:24:1). After vortexing and centrifugation under the same conditions, the aqueous layer was recovered and subjected to a final extraction with an equal volume of chloroform/isoamyl alcohol (24:1). The resulting aqueous phase was transferred to a new 1.5 mL tube, and RNA was precipitated by adding two-thirds volumes of isopropanol. The mixture was gently inverted and incubated at −20 °C for 2 h. The RNA pellet was collected by centrifugation at 12,000 × *g* for 20 min at 4 °C, washed with 1 mL of 75% ethanol, and centrifuged again at 17,500 × *g* for 3 min at 4 °C. The supernatant was discarded, and the pellet was air-dried in a biosafety cabinet for 3–5 min. The dried RNA was dissolved in 20–200 μL of DEPC-treated RNase-free water. RNA quality and quantity were assessed using a Fragment Analyzer or Agilent 2100 Bioanalyzer (Agilent, CA, USA). All samples were sent to BGI (Wuhan, China) for transcriptome sequencing (DNBSEQ platform).

To correct for potential batch effects introduced by this shipping schedule, the "sva" package in R was used to remove batch effects prior to downstream analyses. The assembled Unigene was annotated using three major functional databases (KEGG, GO, and Pfam), and transcription factors were predicted. For pathway-level enrichment analysis, KEGG and GO pathways with a *p*-adjust value < 0.05 were considered significantly enriched. Subsequently, the expression patterns of genes within these significantly enriched pathways were examined in detail. Differential expression analysis was performed within each treatment group using DESeq[55]. Genes satisfying both |Fold Change| ≥ 1.5 and adjusted *p*-value (*p*-adj) <0.05 were considered statistically significant. The raw sequence data reported in this paper have been deposited in the NCBI Sequence Read Archive database under accession code PRJNA1512098 (https://www.ncbi.nlm.nih.gov/sra/PRJNA1512098).

Genomic variation

Genomic samples were collected on days 113, 376, and 727, and each replicate from every treatment group at each time point was sequenced. Genomic DNA was extracted from the samples using the universal genomic DNA extraction kit (magnetic beads) (Magorbio, Shanghai, China) according to the manufacturer's instructions. The genome sequence analysis of *C. muelleri* strain was carried out using the Illumina sequencing platform (MajorBio Co., Shanghai, China). Genomic DNA was fragmented to an average size of ~400 bp using a Covaris M220 Focused Acoustic Shearer following the manufacturer's standard protocol. Illumina sequencing libraries were prepared using the NEXTFLEX Rapid DNA-Seq Kit following manufacturer's recommendations. Fragmented DNA underwent end-repair and 5′ phosphorylation, followed by 3′ adenylation and adapter ligation. The adapter-ligated products were then enriched by PCR. The final libraries were sequenced on the Illumina NovaSeq 6000 platform (Illumina, Inc., San Diego, CA, USA) in paired-end mode (2 × 150 bp). We assembled de novo genomes for *C. muelleri* samples AmbC, GradA, and AbruA using SPAdes[56] (-k 21, 33, 55, 77, 127), obtaining scaffold-level assemblies. Repetitive sequences were masked via RepeatModeler[57]/RepeatMasker[58], and gene annotation was performed by training

BRAKER with in situ transcriptomes (Trinity-assembled) followed by MAKER[59] structural annotation. For variant detection, experimental-group sequencing reads were aligned (Bowtie2[60]) to the control-group genome; SAM outputs were converted to sorted BAM (SAMtools[61]), followed by variant calling/filtering (bcftools[61]: -g 3 -G 10 -e "% QUAL < 10") and functional annotation (ANNOVAR[62] using MAKER-derived GFFs). Key-gene SNPs were identified through homologous sequence alignment (MAFFT[63]) with SNP quantification via a custom Python script. Finally, selective pressure (pN/pS) was assessed using ParaAT and KaKs_Calculator: homologous gene pairs with corresponding CDS/PEP sequences underwent ParaAT analysis; after pN/pS integration, whole-genome CDS/PEP files were automatically formatted for ParaAT based on homologous pairs prior to KaKs_Calculator processing. The raw sequence data reported in this paper have been deposited in the NCBI Sequence Read Archive database under accession code PRJNA1512098 (https://www.ncbi.nlm.nih.gov/sra/PRJNA1512098) .

Statistical analyses

Data analyses were performed with R language (version 4.3.1). To analyze physiological changes with time in the long-term experiments, generalized additive mixed models (GAMMs) were constructed using the "gamm" function from "mgcv" package[64]. Individual bottles ($n = 3$) were treated as a random effect to capture replicate-level variability. Model diagnostics were performed by examining the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the residuals from initial models; based on the presence of temporal autocorrelation indicated by these plots, an appropriate correlation structure (corCAR1) was incorporated into the models. The optimal model for each response variable was selected based on the lowest Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values. Following model fitting, three primary contrasts of biological interest were pre-defined to test specific hypotheses: (1) GradA versus AmbC, (2) AbruA versus AmbC, and (3) AbruA versus GradA. For each contrast, statistical significance was assessed, and effect sizes with their corresponding 95% confidence intervals were reported to facilitate interpretation of the magnitude and biological relevance of the observed differences. Model building, selection, and validation followed the principles and guidelines outlined in Pinheiro and Bates[65], Zuur et al.[66] and Wood[64]. The selected models and corresponding results are provided in Supplementary Tables 1–2. Data for days 34 and 727 were derived from GAMMs predicted results, with statistical significance determined by whether the 95% confidence interval included zero. A significant difference was indicated if the confidence interval did not include zero. Significant differences between treatment groups are indicated by asterisks (*). All parameters are presented as mean ± standard error (SEM). The mean, SEM values, and effect sizes were calculated using R language.

Due to the limited sample size for the gene-level data (fewer time points compared to the main physiological dataset), GAMMs would be inappropriate and prone to overfitting. Therefore, for the gene data, one-way analysis of variance (ANOVA) followed by Fisher's Least Significant Difference (LSD) post-hoc test was conducted using the "stats" and "agricolae" packages in R. To minimize the risk of false-positive results, all treatment groups from both time points were included together in the post-hoc comparisons. Different letters indicate the significant differences between two groups ($p < 0.05$). The source data are provided with this paper, and the code has been deposited in the Zenodo database[67] (https://doi.org/10.5281/zenodo.21910957).

Model estimations

Our study used an offline model of the Marine Biogeochemical Library (MARBL)[68] embedded in Community Earth System Model Version 2 (CESM2)[69] to estimate the impacts of OA on marine phytoplankton, *Chaetoceros*. The growth of phytoplankton function types in MARBL (expressed as Carbon/day) is limited by 3 factors: the availability of nutrients, light and temperature. The formula is as follows:

$$\mu = \mu_{ref} \times f_{nut} \times T_{func} \times light_{lim} \quad (6)$$

where μ_{ref} is the resource-unlimited growth rate, f_{nut} represents the limitation of nutrient, T_{func} is a Q10 curve[68], and the light limitation ($light_{lim}$) is mainly computed with photosynthetically active radiation and the initial slope of the P-E curve, α. Other elements in living biomass are calculated by multiplying their cell quotas normalized to carbon. In simulations of future scenarios, CESM2 can represent changes in physical circulation, temperature, nutrient availability, and light conditions, as well as the impacts of these changes on phytoplankton and marine ecosystem. Our offline approach is just to superimpose the experimentally derived ocean acidification effects onto the future mean climate state. The detailed calculation method is as follows:

Since there is only one diatom functional group in MARBL, we first compiled the dataset on the ratio of *Chaetoceros* to diatoms from Tara Oceans, which provides relatively comprehensive observations of *Chaetoceros* abundance. And then we binned the dataset within 10° of latitude and fitted fourth-degree polynomial function with quartiles, mean, median ratio of each box-and-whisker plot. We chose the set of parameters with the minimum RMSE and established a relationship between this ratio and latitude. In each model grid, we calculated the specific ratio and multiplied it by model projection of diatoms to obtain the global pattern of *Chaetoceros*. In subsequent offline simulation, we linearly calculate the OA effect on each model grid based on laboratory results. In detail, we used the carbon fixation rate of *Chaetoceros* as an example to illustrate the offline estimates of the OA effect on primary production under Shared Socioeconomic Pathways 5–8.5 (SSP5–8.5) emission scenario.

$$intppChae_{CESM2015\sim2024} = intppdiat_{CESM2015\sim2024} \times r \quad (7)$$

$$intppChae_{CESM2091\sim2100} = intppdiat_{CESM2091\sim2100} \times r \quad (8)$$

$$intppChae_{oa\ effect} = \left(\frac{\delta[H^+]_{CESM}}{\delta[H^+]_{lab}} \times oa_{cfix}\right) \times intppChae_{2015\sim2024} \quad (9)$$

$$intppChae_{SSP+oa} = intppChae_{oa\ effect} + intppChae_{CESM2091\sim2100} \quad (10)$$

where $intppdiat_{CESM}$ represents the 10-year average of the present global net primary production of diatom. $\delta[H^+]_{CESM}$ is the average $[H^+]$ from 2091 to 2100 minus the average $[H^+]$ from 2015 to 2024 in CESM, and $\delta[H^+]_{lab}$ is the difference between measured $[H^+]$ in the AmbC and GradA (and/or AbruA) experiments, the oa_{cfix} specific ocean acidification coefficient of carbon fixation rate [μgC (μg Chl a)$^{-1}$ day^{-1}], $\frac{Cfix_{1000}ppm - Cfix_{440}ppm}{Cfix_{440}ppm} \approx 3.69\%$(GradA) and $\frac{Cfix_{1000}ppm - Cfix_{440}ppm}{Cfix_{440}ppm} \approx 12.69\%$(AbruA), r is the ratio of *Chaetoceros*, and the $intppChae_{SSP+oa}$ represents the global *Chaetoceros* primary production with the consideration of the direct OA effect under SSP5–8.5 scenario. $intppChae_{CESM2015\sim2024}$ and $intppChae_{CESM2091\sim2100}$ represent the average global net primary production of *Chaetoceros* over 2015–2024 and 2091–2100, respectively; $intppChae_{oa\ effect}$ represents the experimentally derived ocean acidification effects of *Chaetoceros* under SSP5-8.5 scenario. In the same way, when estimating the change of BSi production and BSi/POC, we used relevant biogeochemical outputs of CESM, calculated the respective OA coefficients, and then acquired their changes

considering the OA effect:

$$\mathrm{pbsiChae}_{\mathrm{CESM2015\sim2024}} = \mathrm{pbsidiat}_{\mathrm{CESM2015\sim2024}} \times r \quad (11)$$

$$\mathrm{pbsiChae}_{\mathrm{CESM2091\sim2100}} = \mathrm{pbsidiat}_{\mathrm{CESM2091\sim2100}} \times r \quad (12)$$

$$\mathrm{pbsiChae}_{\mathrm{oa\ effect}} = \left(\frac{\delta[\mathrm{H}^+]_{\mathrm{CESM}}}{\delta[\mathrm{H}^+]_{\mathrm{lab}}} \times \mathrm{oa}_{\mathrm{BSiprod}}\right) \times \mathrm{pbsiChae}_{\mathrm{2015\sim2024}} \quad (13)$$

$$\mathrm{pbsiChae}_{\mathrm{SSP+oa}} = \mathrm{pbsiChae}_{\mathrm{oa\ effect}} + \mathrm{pbsiChae}_{\mathrm{CESM2091\sim2100}} \quad (14)$$

$\mathrm{pbsidiat}_{\mathrm{CESM}}$ represents the 10-year average of the global biogenic silica production, the $\mathrm{oa}_{\mathrm{BSiprod}}$ is specific ocean acidification coefficient of biogenic silica production [μg Si (μg Chl a)$^{-1}$ day^{-1}], $\frac{\mathrm{BSiprod}_{1000\mathrm{ppm}} - \mathrm{BSiprod}_{440\mathrm{ppm}}}{\mathrm{BSiprod}_{440\mathrm{ppm}}} \approx 1.44\%\,(\mathrm{GradA})$ and $\frac{\mathrm{BSiprod}_{1000\mathrm{ppm}} - \mathrm{BSiprod}_{440\mathrm{ppm}}}{\mathrm{BSiprod}_{440\mathrm{ppm}}} \approx 7.42\%\,(\mathrm{AbruA})$.

$$\mathrm{physiChae}_{\mathrm{CESM2015\sim2024}} = \mathrm{physidiat}_{\mathrm{CESM2015\sim2024}} \times r \quad (15)$$

$$\mathrm{physiChae}_{\mathrm{CESM2091\sim2100}} = \mathrm{physidiat}_{\mathrm{CESM2091\sim2100}} \times r \quad (16)$$

$$\mathrm{physiChae}_{\mathrm{oa\ effect}} = \left(\frac{\delta[\mathrm{H}^+]_{\mathrm{CESM}}}{\delta[\mathrm{H}^+]_{\mathrm{lab}}} \times \mathrm{oa}_{\mathrm{Si}}\right) \times \mathrm{physiChae}_{\mathrm{2015\sim2024}} \quad (17)$$

$$\mathrm{physiChae}_{\mathrm{SSP+oa}} = \mathrm{physiChae}_{\mathrm{oa\ effect}} + \mathrm{physiChae}_{\mathrm{CESM2091\sim2100}} \quad (18)$$

$\mathrm{physidiat}_{\mathrm{CESM}}$ represents the 10-year average of Si content in total diatoms, the $\mathrm{oa}_{\mathrm{Si}}$ is specific ocean acidification coefficient of BSi (pmol cell^{-1}), $\frac{\mathrm{BSi}_{1000\mathrm{ppm}} - \mathrm{BSi}_{440\mathrm{ppm}}}{\mathrm{BSi}_{440\mathrm{ppm}}} \approx -11.25\%(\mathrm{GradA})$ and $\frac{\mathrm{BSi}_{1000\mathrm{ppm}} - \mathrm{BSi}_{440\mathrm{ppm}}}{\mathrm{BSi}_{440\mathrm{ppm}}} \approx -16.28\%(\mathrm{AbruA})$.

$$\mathrm{phyChae}_{\mathrm{CESM2015\sim2024}} = \mathrm{phydiat}_{\mathrm{CESM2015\sim2024}} \times r \quad (19)$$

$$\mathrm{phyChae}_{\mathrm{CESM2091\sim2100}} = \mathrm{phydiat}_{\mathrm{CESM2091\sim2100}} \times r \quad (20)$$

$$\mathrm{phyChae}_{\mathrm{oa\ effect}} = \left(\frac{\delta[\mathrm{H}^+]_{\mathrm{CESM}}}{\delta[\mathrm{H}^+]_{\mathrm{lab}}} \times \mathrm{oa}_{\mathrm{POC}}\right) \times \mathrm{phyChae}_{\mathrm{2015\sim2024}} \quad (21)$$

$$\mathrm{phyChae}_{\mathrm{SSP+oa}} = \mathrm{phyChae}_{\mathrm{oa\ effect}} + \mathrm{phyChae}_{\mathrm{CESM2091\sim2100}} \quad (22)$$

$\mathrm{phydiat}_{\mathrm{CESM}}$ represents the 10-year average of carbon content in total diatoms, the $\mathrm{oa}_{\mathrm{POC}}$ is specific ocean acidification coefficient of POC (pmol cell^{-1}), $\frac{\mathrm{POC}_{1000\mathrm{ppm}} - \mathrm{POC}_{440\mathrm{ppm}}}{\mathrm{POC}_{440\mathrm{ppm}}} \approx 8.69\%(\mathrm{GradA})$ and $\frac{\mathrm{POC}_{1000\mathrm{ppm}} - \mathrm{POC}_{440\mathrm{ppm}}}{\mathrm{POC}_{440\mathrm{ppm}}} \approx 0.47\%(\mathrm{AbruA})$.

$$\frac{\mathrm{Si}}{\mathrm{C}_{\mathrm{Chae\,SSP+oa}}} = \frac{\mathrm{physiChae}_{\mathrm{SSP+oa}}}{\mathrm{phyChae}_{\mathrm{SSP+oa}}} \quad (23)$$

All ocean acidification coefficients were calculated from the mean values of the entire dataset from the long-term experiment for both the GradA and AbruA groups. It is important to note that, in this offline framework, the acidification response is superimposed onto the late-century climate state based on the mean $[H^+]$ change, effectively representing a direct jump to the target CO_2 level rather than a transient forcing with gradual rate changes, which inevitably neglects potential variability observed in laboratory culture experiments. The upper and lower effect estimates further complement this approach by providing a reasonable range of uncertainty (Supplementary Table 4). We chose $[H^+]$ as the metric for ocean acidification because pH was directly measured in the laboratory experiments, making $[H^+]$ more readily comparable to the model-derived values. In contrast, using pCO_2 would introduce additional uncertainties associated with converting between carbonate chemistry parameters. Therefore, $[H^+]$ provides a more direct and conservative basis for scaling the acidification response in our offline modelling framework. Additionally, we compared the predicted values derived from variations in pCO_2, which yielded results that are largely consistent with those calculated from $[H^+]$. Further details are provided in Supplementary Table 4 and 6.

All calculations were conducted using code written with Python 3.8. The model data are provided with this paper, and the code has been deposited in the Zenodo database[67] (https://doi.org/10.5281/zenodo.21910957).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Source data are provided as a Source data file. The datasets generated during and/or analysed during the current study are available in the Figshare repository (https://doi.org/10.6084/m9.figshare.30412849). The raw sequence data reported in this paper have been deposited in the NCBI Sequence Read Archive (SRA) database under accession code PRJNA1512098 (https://www.ncbi.nlm.nih.gov/sra/PRJNA1512098). Source data are provided with this paper.

Code availability

The R and Python codes used to perform the analyses are publicly available the Zenodo database[67] (https://doi.org/10.5281/zenodo.21910957).

References

1. Gulev, S. K. et al. Changing state of the climate system. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I To The Sixth Assessment Report of The Intergovernmental Panel on Climate Change* 287–422 (Cambridge University Press, 2021).
2. Orr, J. C. et al. Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature* **437**, 681–686 (2005).
3. Caldeira, K. & Wickett, M. E. Anthropogenic carbon and ocean pH. *Nature* **425**, 365 (2003).
4. Boyd, P. W. & Hutchins, D. A. Understanding the responses of ocean biota to a complex matrix of cumulative anthropogenic change. *Mar. Ecol. Prog. Ser.* **470**, 125–135 (2012).
5. Martin-Jézéquel, V., Hildebrand, M. & Brzezinski, M. A. Silicon metabolism in diatoms: implications for growth. *J. Phycol.* **36**, 821–840 (2000).
6. Tréguer, P. et al. Influence of diatom diversity on the ocean biological carbon pump. *Nat. Geosci.* **11**, 27–37 (2018).
7. Sugie, K., Yoshimura, T. & Wakita, M. Impact of CO_2 on the elemental composition of the particulate and dissolved organic matter of marine diatoms emerged after nitrate depletion. *Limnol. Oceanogr.* **63**, 1924–1943 (2018).
8. Street-Perrott, F. A. & Barker, P. A. Biogenic silica: a neglected component of the coupled global continental biogeochemical

cycles of carbon and silicon. *Earth Surf. Proc. Land.* **33**, 1436–1457 (2008).
9. Badger, M. R. et al. The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO_2-concentrating mechanisms in algae. *Can. J. Bot.* **76**, 1052–1071 (1998).
10. Reinfelder, J. R. Carbon concentrating mechanisms in eukaryotic marine phytoplankton. *Annu. Rev. Mar. Sci.* **3**, 291–315 (2011).
11. Taucher, J. et al. Combined effects of CO_2 and temperature on carbon uptake and partitioning by the marine diatoms *Thalassiosira weissflogii* and *Dactyliosolen fragilissimus*. *Limnol. Oceanogr.* **60**, 901–919 (2015).
12. Xu, D. et al. Acclimation and adaptation to elevated pCO_2 increase arsenic resilience in marine diatoms. *ISME J.* **15**, 1599–1613 (2021).
13. Li, F. T., Beardall, J., Collins, S. & Gao, K. S. Decreased photosynthesis and growth with reduced respiration in the model diatom *Phaeodactylum tricornutum* grown under elevated CO_2 over 1800 generations. *Global Change Biol.* **23**, 127–137 (2017).
14. Lauvset, S. K., Gruber, N., Landschützer, P., Olsen, A. & Olsen, A. Trends and drivers in global surface ocean pH over the past 3 decades. *Biogeosciences* **12**, 1285–1298 (2015).
15. Li, W. et al. Functional responses of smaller and larger diatoms to gradual CO_2 rise. *Sci. Total. Environ.* **680**, 79–90 (2019).
16. Malviya, S. et al. Insights into global diatom distribution and diversity in the world's ocean. *Proc. Natl. Acad. Sci. USA* **113**, 1516–1525 (2016).
17. Kaneko, K. *Universal Biology: The Physics of Life through the Macro-Micro Consistency Principle* (Cambridge University Press, 2025).
18. Luo, Y. W. et al. Reduced nitrogenase efficiency dominates response of the globally important nitrogen fixer *Trichodesmium* to ocean acidification. *Nat. Commun.* **10**, 1521 (2019).
19. Blackford, J. C. Predicting the impacts of ocean acidification: challenges from an ecosystem perspective. *J. Marine Syst.* **81**, 12–18 (2010).
20. Zhang, F. T. et al. Phosphate limitation intensifies negative effects of ocean acidification on globally important nitrogen fixing cyanobacterium. *Nat. Commun.* **13**, 6730 (2022).
21. Hein, M. & Sand-Jensen, K. CO_2 increases oceanic primary production. *Nature* **388**, 526–527 (1997).
22. Hoppe, C. J. M., Holtz, L. M., Trimborn, S. & Rost, B. Ocean acidification decreases the light-use efficiency in an Antarctic diatom under dynamic but not constant light. *New Phytol.* **207**, 159–171 (2015).
23. Roberts, K., Granum, E., Leegood, R. C. & Raven, J. A. Carbon acquisition by diatoms. *Photosynth. Res.* **93**, 79–88 (2007).
24. Burlacot, A. et al. Alternative photosynthesis pathways drive the algal CO_2-concentrating mechanism. *Nature* **605**, 366–371 (2022).
25. Hopkinson, B. M., Dupont, C. L., Allen, A. E. & Morel, F. M. M. Efficiency of the CO_2-concentrating mechanism of diatoms. *Proc. Natl. Acad. Sci. USA* **108**, 3830–3837 (2011).
26. Flynn, K. J. et al. Changes in pH at the exterior surface of plankton with ocean acidification. *Nat. Clim. Change.* **2**, 510–513 (2012).
27. Zhang, M., Zhen, Y., Mi, T. Z. & Lin, S. J. Integrated transcriptome sequencing (RNA-seq) and proteomic studies reveal resource reallocation towards energy metabolism and defense in *Skeletonema marinoi* in response to CO_2 increase. *Appl. Environ. Microb.* **87**, e02614–e02620 (2021).
28. Jin, P. et al. A reduction in metabolism explains the tradeoffs associated with the long-term adaptation of phytoplankton to high CO_2 concentrations. *New Phytol.* **5**, 2155–2167 (2021).
29. Balparda, M. et al. Regulation of plant carbon assimilation metabolism by post-translational modifications. *Plant J.* **114**, 1059–1079 (2023).
30. Matiolli, C. C., Soares, R. C., Alves, H. L. S. & Abreu, I. A. Turning the knobs: the impact of post-translational modifications on carbon metabolism. *Front. Plant Sci.* **12**, 781508 (2022).
31. Zheng, K. et al. Regulation of plant glycolysis and the tricarboxylic acid cycle by posttranslational modifications. *Plant J.* **122**, e70142 (2025).
32. Annunziata, R. et al. bHLH-PAS protein RITMO1 regulates diel biological rhythms in the marine diatom *Phaeodactylum tricornutum*. *Proc. Natl. Acad. Sci. USA* **116**, 13137–13142 (2019).
33. Peng, B. Y. et al. Adaptation to warming and elevated CO_2 influences diatom response norms and multi-trait variations across CO_2 gradients. *J. Plant Ecol.* **18**, rtae116 (2025).
34. Cai, T. et al. The differential responses of coastal diatoms to ocean acidification and warming: a comparison between *Thalassiosira* sp. and *Nitzschia closterium f.minutissima*. *Front. Microbiol.* **13**, 851149 (2022).
35. Falciatore, A. & Mock, T. *The Molecular Life of Diatoms* (Springer, 2022).
36. Milligan, A. J. & Morel, F. M. M. A proton buffering role for silica in diatoms. *Science* **297**, 1848–1850 (2002).
37. Agusti, S. et al. Ubiquitous healthy diatoms in the deep sea confirm deep carbon injection by the biological pump. *Nat. Commun.* **6**, 7608 (2015).
38. Le Moigne, F. A. C., Pabortsava, K., Marcinko, C. L. J., Martin, P. & Sanders, R. J. Where is mineral ballast important for surface export of particulate organic carbon in the ocean? *Geophys. Res. Lett.* **41**, 8460–8468 (2014).
39. Finkel, Z. V. et al. Phytoplankton in a changing world: cell size and elemental stoichiometry. *J. Plankton Res.* **32**, 119–137 (2010).
40. Taucher, J. et al. Enhanced silica export in a future ocean triggers global diatom decline. *Nature* **605**, 696–700 (2022).
41. Petrou, K. et al. Acidification diminishes diatom silica production in the Southern Ocean. *Nat. Clim. Change.* **9**, 781–786 (2019).
42. Ogura, A. et al. Comparative genome and transcriptome analysis of diatom, *Skeletonema costatum*, reveals evolution of genes for harmful algal bloom. *BMC Genomics* **19**, 765 (2018).
43. Sett, S., Schulz, K. G., Bach, L. T. & Riebesell, U. Shift towards larger diatoms in a natural phytoplankton assemblage under combined high-CO_2 and warming conditions. *J. Plankton Res.* **4**, 391–406 (2018).
44. Wu, F. H. et al. The dynamics of adaptive evolution in microalgae in a high-CO_2 ocean. *New Phytol.* **4**, 1608–1624 (2025).
45. Necsulea, A. & Kaessmann, H. Evolutionary dynamics of coding and non-coding transcriptomes. *Nat. Rev. Genet.* **15**, 734–748 (2014).
46. Carrier, G. et al. Comparative transcriptome of wild type and selected strains of the microalgae *Tisochrysis lutea* provides insights into the genetic basis, lipid metabolism and the life cycle. *PLoS ONE* **9**, e86889 (2014).
47. Low-Décarie, E., Fussmann, G. F. & Bell, G. The effect of elevated CO_2 on growth and competition in experimental phytoplankton communities. *Global Change Biol.* **17**, 2525–2535 (2011).
48. IPCC. Climate change 2023: synthesis report. *Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Lee, H. & Romero, J.), 35–115 (IPCC, 2023).
49. Guillard, R. R. L. & Ryther, J. H. Studies of marine planktonic diatoms. I. *Cyclotella nana Hustedt* and *Detonula confervaceae* (Cleve). *Gran. Can. J. Microbiol.* **8**, 229–239 (1962).
50. Lewis, E., Wallace, D. W. R. & Allison, L. J. Program developed for CO_2 system calculations. In: oceanCO$_2$, Available at: https://www.osti.gov/servlets/purl/639712 (1998).
51. Hama, T. et al. Measurement of photosynthetic production of a marine phytoplankton population using a stable ^{13}C isotope. *Mar. Biol.* **73**, 31–36 (1983).

52. Porra, R. J. The chequered history of the development and use of simultaneous equations for the accurate determination of chlorophylls *a* and *b*. *Photosynth. Res.* **73**, 149–156 (2002).
53. Strickland, J. D. H. & Parsons, T. R. A practical handbook of seawater analysis. *Bull. Fish. Res. Boar. Can.* **167**, 49–80 (1968).
54. Brzezinski, M. A. & Nelson, D. M. The annual silica cycle in the Sargasso Sea near Bermuda. *Deep Sea Res. I Oceanogr. Res. Papers* **42**, 1215–1237 (1995).
55. Wang, L. K., Feng, Z. X., Wang, X., Wang, X. W. & Zhang, X. G. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **26**, 136–138 (2010).
56. Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A. & Korobeynikov, A. Using SPAdes *de novo* assembler. *Curr. Protoc. Bioinformatics.* **70**, e102 (2020).
57. Flynn, F. M. et al. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl. Acad. Sci. USA* **117**, 9451–9457 (2020).
58. Tarailo-Graovac, M. & Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinformatics*. **4**, 4.10.1–4.10.14 (2009).
59. Cantarel, B. L. et al. MAKER: an easy -to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18**, 188–196 (2008).
60. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* **9**, 357–359 (2012).
61. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
62. Wang, K., Li, M. Y. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
63. Katoh, K., Misawa, K. & Kuma, K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
64. Wood, S. N. *Generalized Additive Models: An Introduction With R* (CRC Press, 2017).
65. Pinheiro, J. C. & Bates, D. M. *Mixed-Effects Models in S and S-PLUS* (Springer, 2006).
66. Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A. & Smith, G. M. *Mixed Effects Models and Extensions in Ecology With R* (Springer, 2009).
67. Zhao, X. et al. Data from: Abrupt pH changes overestimate the responses of *Chaetoceros muelleri* to ocean acidification. *Zenodo*, https://doi.org/10.5281/zenodo.21910957 (2026).
68. Long, M. C. et al. Simulations with the marine biogeochemistry library (MARBL). *J. Adv. Model. Earth Sy.* **13**, e2021MS002647 (2021).
69. Danabasoglu, G. et al. The community earth system model version 2 (CESM2). *J. Adv. Model. Earth Syst.* **12**, e2019MS001916 (2020).
70. Bach, L. T. et al. Effects of elevated CO_2 on a natural diatom community in the subtropical NE Atlantic. *Front. Mar. Sci.* **6**, 75 (2019).

Acknowledgements

T.M. acknowledges partial support by the School of Environmental Sciences, University of East Anglia, United Kingdom. We were grateful for the help from the technicians Wenyan Zhao and Xianglan Zeng.

Author contributions

X.Z. (Xin Zhao) performed the experiments, collected and visualized the data, conducted data analysis and interpretation, and wrote and revised the drafts. H.X.T. and S.L.W. constructed and visualized the CESM model. X.Z. (Xu Zhang) performed the downstream analysis of genomic data. D.X. contributed to data analysis and interpretation. Z.W.H. and T.Z.Z. performed the downstream analysis of transcriptomic data. J.Z.S. contributed to data analysis and visualization. K.S.G. contributed to data analysis and manuscript revision. X.L. performed the transcriptomic and genomic analyses. T.M. and N.H.Y. contributed to data analysis, manuscript writing, and revision. G.G. conceptualized the study, offered guidance in the research and manuscript development, contributed to manuscript writing and revision, and acquired resources and funding.

Funding

This work was supported by the National Natural Science Foundation of China (42576209), "Sustainable Aquaculture under Climate Change" as a project forming part of the UN Decade of Ocean Science for Sustainable Development 2021-2030 (No. 119.9), the International Cooperation Seed Funding Project for China's Ocean Decade Actions (GHZZ3702840002024020000020), the Open Project of Key Laboratory of Marine Environmental Survey Technology and Application, Ministry of Natural Resources, P.R. China (MESTA-2024-A001), and the Ocean Negative Carbon Emissions (ONCE) program.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-026-77753-0.

Correspondence and requests for materials should be addressed to Thomas Mock, Naihao Ye or Guang Gao.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2026