Marine Life Science & Technology (2026) 8:221–236
https://doi.org/10.1007/s42995-025-00343-4

RESEARCH PAPER

Environmental drivers and key taxa shaping diatom–dinoflagellate ratios in eutrophic coastal waters

Lingshuai Zhang[1,2] · Zhiqin Wang[1,2] · Jin-Yu Terence Yang[3] · Shuqun Song[4] · Zhaozhang Chen[3] · Kuanbo Zhou[3] · Hao Zhang[1] · Yehui Tan[1] · Hongbin Liu[5] · Dazhi Wang[6] · Xiaomin Xia[7]

Received: 17 May 2025 / Accepted: 22 November 2025 / Published online: 9 February 2026

Abstract

Diatoms and dinoflagellates are two pivotal phytoplankton groups present in coastal ecosystems that play key roles in marine food webs and biogeochemical cycles. The diatom-to-dinoflagellate ratio (diat/dino ratio) serves as an indicator of ecosystem status and phytoplankton community dynamics; however, the specific taxa that contribute to its variability remain poorly understood. This study investigated the phytoplankton community composition and diat/dino ratios in the coastal regions of the East China Sea and northern South China Sea during summer using quantitative PCR (qPCR) and 18S rRNA gene pyrosequencing. The qPCR results revealed that diatoms dominated in the estuarine and nearshore waters, whereas dinoflagellates prevailed in the offshore regions. Random Forest analysis identified dissolved oxygen (DO) and the nitrogen-to-phosphorus (N:P) and silicon-to-nitrogen (Si:N) ratios as the primary drivers of variation in the diat/dino ratio. The influence of N:P ratios was further modulated by the absolute nitrogen and phosphorus concentrations. Taxonomic profiling revealed that Thalassiosiraceae and Chaetocerotaceae were enriched in nutrient-rich estuarine waters, while Leptocylindraceae, Bacillariaceae, and Skeletonemaceae dominated in regions with low N:P ratios. In contrast, dinoflagellate families, such as Thoracosphaeraceae, Pyrocystaceae, Peridiniaceae, and Heterocapsaceae, were more abundant in environments with high DO and elevated N:P ratios. Notably, the northward expansion of *Scrippsiella* (Thoracosphaeraceae) drove changes in the bloom dynamics that threaten the coastal ecosystem balance. These findings demonstrate that nutrient stoichiometry and oxygen availability influence the diat/dino ratio by favoring distinct phytoplankton taxa, thus offering insights into how anthropogenic nutrient inputs shape community structure and guide coastal ecosystem management.

Keywords Coastal waters · Diatom · Dinoflagellate · Diat/dino ratio · Eutrophication · Nutrients

Edited by Chengchao Chen.

✉ Xiaomin Xia
xxm1984@xmu.edu.cn

[1] State Key Laboratory of Tropical Oceanography, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China

[2] University of Chinese Academy of Sciences, Beijing 100049, China

[3] State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen 361102, China

[4] CAS Key Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China

[5] Department of Ocean Science, The Hong Kong University of Science and Technology, Hong Kong SAR 999077, China

[6] State Key Laboratory of Marine Environmental Science/College of the Environment and Ecology, Xiamen University, Xiamen 361102, China

[7] Innovation Research Center for Carbon Neutralization, Fujian Key Laboratory of Marine Carbon Sequestration, Xiamen University, Xiamen 361102, China

Introduction

In recent years, shifts in phytoplankton community composition have become an increasing concern (Michael Beman et al. 2005; Rajapaksha et al. 2024), particularly regarding changes in the diatom-to-dinoflagellate ratio (diat/dino ratio), which is a widely used indicator of phytoplankton community structure. Diatoms and dinoflagellates are two of the most ecologically important phytoplankton groups, although they fulfill markedly different roles within marine ecosystems. Diatoms are key primary producers that contribute substantially to biogeochemical cycles and marine food webs (B-Béres et al. 2023). In contrast, dinoflagellates include numerous toxic and mixotrophic species, and their rising abundance is often associated with harmful algal blooms (HABs), thereby posing potential threats to marine biodiversity and ecosystem health (Gu et al. 2013; Jeong et al. 2016).

Changes in the diat/dino ratio have been documented across various marine environments, including the coastal waters of Europe (Hinder et al. 2012; Nohe et al. 2020; Spilling et al. 2018; Wasmund et al. 2017), the Bohai Sea (Wei et al. 2024), and the East China Sea (ECS) (Xiao et al. 2018). However, despite extensive research on the diat/dino ratios, the drivers of shifts therein remain to be fully elucidated. For instance, in the ECS, a decline in diatom dominance over the past 50 years has been attributed to nutrient enrichment (Jiang et al. 2014), whereas alterations in the silicon-to-nitrogen (Si:N) ratio following the Three Gorges Dam construction have favored dinoflagellate proliferation (Zhou et al. 2016). Similarly, in the Bohai Sea, eutrophication characterized by elevated nitrogen-to-phosphorus (N:P) ratios has been linked to a transition from diatom dominance to a codominance with dinoflagellates (Wei et al. 2024). In contrast, studies in the Pearl River Estuary (PRE) have revealed that increased nitrate levels were associated with higher diatom-to-dinoflagellate ratios (Cheung et al. 2021). These contrasting findings may reflect (1) regional differences in species composition, as various diatom and dinoflagellate taxa have distinct physiological traits and environmental tolerances, and (2) variability in the intensity and type of eutrophication across locations.

Despite the complexity of these communities, many studies treat diatoms and dinoflagellates as homogeneous functional groups, often overlooking the ecological roles and environmental sensitivities of the specific taxa (Agirbas et al. 2017; Liang et al. 2019; Nohe et al. 2020; Spilling et al. 2018; Wang et al. 2024). However, different species within these groups possess unique morphological, behavioral, and nutritional traits that influence their ecological niches (Litchman et al. 2010). For example, *Chaetoceros* and *Thalassiosira* are both common diatoms, but they differ considerably in form and function. *Chaetoceros* typically forms long chains with silica-based setae, which reduces sinking and deters grazers, thus making them well suited to the turbulent, nutrient-rich coastal waters (Hoffmann et al. 2007). In contrast, *Thalassiosira* species are often solitary, more grazing-sensitive, and broadly distributed across coastal and open ocean regions (Liu et al. 2024).

To better evaluate the ecological influence of diat/dino ratio shifts, it is essential to focus on the key indicator taxa rather than relying solely on broad functional groupings. However, traditional approaches such as light microscopy and pigment analysis offer limited taxonomic resolution, and they often fail to detect small, cryptic, or heterotrophic taxa (Cheung et al. 2021; Hinder et al. 2012; Wei et al. 2024; Xiao et al. 2018). These limitations highlight the urgent need for molecular tools, such as high-throughput sequencing, to more accurately characterize the phytoplankton community structure and reveal the nutrient-driven mechanisms behind community shifts.

To address the knowledge gaps concerning (1) key environmental drivers of the diat/dino ratio variations and (2) specific diatom and dinoflagellate taxa involved in these shifts, we examined the phytoplankton abundance and community composition across diverse environmental niches. By integrating random forest (RF) models with correlation analyses, we identified the key taxa that contribute to shifts in the diat/dino ratio and the environmental variables that are most strongly associated with these changes. In summer, coastal waters of China experience relatively stable temperatures but high variability in nutrient concentrations and stoichiometric ratios, thereby providing a natural experimental setting to separate the influence of nutrient dynamics from that of temperature on the phytoplankton community structure.

Materials and methods

Study sites and sample collections

Sampling was performed in the coastal water of the ECS and northern South China Sea (SCS) in August 2023, with sampling locations including the Changjiang River Estuary (CJE), PRE, Daya Bay (DYB), and coastal waters of Fujian and Zhejiang (MZC) (Fig. S1). A total of 92 sampling stations were established and parameters were measured, including temperature (Tem), salinity (Sal), dissolved oxygen (DO), ammonium (NH_4^+), nitrate (NO_3^-), nitrite (NO_2^-), phosphate (PO_4^{3-}), and silicic acid ($Si(OH)_4$). At each station, one surface seawater sample (at approximately 3 m depth, within the surface mixed layer) was collected

using a conductivity–temperature–depth rosette system (CTD, Sea-Bird Electronics) fitted with 12-L Niskin bottles. For DNA sample collection, 500 mL of seawater collected from each station was prefiltered with a 200 µm mesh to exclude mesozooplankton and then filtered through a 3-µm polycarbonate membrane (PC, Millipore Corporation). The polycarbonate membrane samples were flash-frozen in liquid nitrogen and stored at −80 °C. The temperature, salinity, and DO of the samples were measured with CTD sensors. To assess the chlorophyll content, 250 mL of prefiltered seawater was passed through GF/F filters. Chlorophyll was extracted from these filters using a 90% acetone solution, and the extract was subsequently measured using a Trilogy fluorometer (Turner Designs) following the previously described method by Strickland and Parsons (1972). Nutrient samples were collected by filtering water through a 0.45 µm acetate fiber membrane. Silicate samples were stored at 4 °C, while NO_3^-, NO_2^-, NH_4^+, and PO_4^{3-} samples were stored at −20 °C. The concentrations of NO_3^-, NO_2^-, PO_4^{3-}, and $Si(OH)_4$ were determined using a Bran+Luebbe Auto-Analyzer 3 (AA3, Germany) via gas-segmented continuous flow analysis, which is an automated colorimetric method based on spectrophotometry. The quantification limits for NO_3^-, NO_2^-, PO_4^{3-}, and $Si(OH)_4$ were 0.10, 0.04, 0.08, and 0.08 umol/L, respectively. The analytical precision was ±1% for NO_3^- and NO_2^-, ±2% for PO_4^{3-}, and ±2.8% for $Si(OH)_4$ (Murphy and Riley 1962). The NH_4^+ concentrations were measured using a Bran+Luebbe AA500 AutoAnalyzer (Germany), with a quantification limit of 0.1 umol/L (Holmes et al. 1999). Moreover, in this study, NO_x represents the combined concentration of NO_3^- and NO_2^-. The photosynthetically active radiation (PAR) data were downloaded from the NOAA database at https://coastwatch.pfeg.noaa.gov/data.html.

Genomic DNA extraction, qPCR of diatoms and dinoflagellates, and bioinformatics analysis

Genomic DNA was extracted from the 3 µm polycarbonate membranes collected from the 92 stations using a modified enzyme/phenol-chloroform extraction protocol (Xia et al. 2020). The diatom and dinoflagellate cell abundances were quantified by qPCR using the primer pairs 5′-CAGGTC TGTGATGCCCTT-3′/5′-CAATGCAGWTTGATGAWC TG-3′ and 5′-CCGCGGTAATTCCAGCTC-3′/5′-GAGCCA GATRCDCACCCA-3′, respectively (Godhe et al. 2008). To determine the optimal annealing temperature for qPCR amplification, a gradient PCR was performed. The reaction was set up in a total volume of 50 µL, containing 25 µL of 2× Taq Mix (Takara, China), 2 µL of forward primer, 2 µL of reverse primer, 2 µL of DNA template, and 19 µL of double-distilled water. The thermal cycling conditions included an initial denaturation at 95 °C for 3 min, followed by 30 cycles of denaturation at 95 °C for 20 s, annealing across a temperature gradient ranging from 52 °C to 65 °C for 20 s, and extension at 72 °C for 30 s. The diatom and dinoflagellate PCR products were cloned into the pMD18-T vector (Takara), and plasmids containing the diatom and dinoflagellate rDNA fragments were extracted using the Mini Plasmid Kit (Takara). Based on the formula reported by Zhu et al. (2005), the initial concentrations of diatom and dinoflagellate plasmids were determined to be 2.23×10^{10} and 1.92×10^{10} copies per µL, respectively. Standard curves were generated by performing 10 serial dilutions of the diatom and dinoflagellate plasmids. The qPCR reaction was performed in a 20 µL volume consisting of 10 µL of qPCR master mix, 0.3 µL of each primer, 2 µL of template DNA, and 7.4 µL of double-distilled water. The amplification reactions were performed in triplicate with the following PCR protocol: initial denaturation at 95 °C for 3 min, followed by 30 cycles of denaturation at 95 °C for 20 s and annealing/extension at 55 °C for diatoms or 60 °C for dinoflagellates for 25 s. Standard curves and negative controls without template DNA were included in each qPCR run. No PCR amplification was detected in the negative controls. Cell abundance was calculated by relating the qPCR-derived gene copy number to the standard curves, which provided a linear relationship between the Ct values and known gene copy number (GCN). It is important to note that the copy number of the 18S rRNA gene varies across the different diatom and dinoflagellate species. To improve the reliability of our quantitative data, a correction factor (GCN per cell) of 166 for diatoms and 4919 for dinoflagellates was applied, as previously reported by Martin et al. (2022). The distribution of diatom and dinoflagellate abundances across the studied region was visualized using Surfer Version 23.

Diat/dino ratio

The diat/dino ratio was calculated using the formula: $\text{diat/dino} = \frac{\text{Abundance\_diatom}}{\text{Abundance\_dinoflagellate}}$, where abundance represents the number of cells per liter. To normalize the distribution, the ratio was $\log_{10}$-transformed. A positive log-transformed value indicates diatom dominance, whereas a negative value reflects a higher dinoflagellate abundance. To examine the relationship between the diat/dino ratio and environmental variables, three statistical approaches were employed: generalized additive models (GAMs), RF, and a cubic polynomial regression model. GAMs were implemented using the R package *mgcv* (version 1.9-1), which provides flexibility in modeling nonlinear relationships (Pedersen et al. 2019). For the GAM selection, we first calculated the relative importance of each predictor and sequentially removed variables with the lowest importance. This stepwise procedure continued until the model with the lowest Akaike

information criterion value was obtained. The optimal GAM had the following structure, and the statistical model is given by

$$\log_{10}(\text{diat/dino}) \sim \text{s(Tem)} + \text{s(Sal)} + \text{s(DO)} + \text{s(Chl}a\text{)} + \text{s}\left(\text{Si(OH)}_4\right) + \text{s}\left(\text{PO}_4^{3-}\right) + \text{s}\left(\text{NO}_x\right) + \text{s}\left(\text{NO}_2^{-}\right) + \text{s}\left(\text{NH}_4^{+}\right) + \text{s(N/P)} + \text{s(Si/N)} \quad (1)$$

To further investigate the potential nutrient interactions, we constructed alternative models that incorporated tensor (te) product smooths between the N:P ratio and PO_4^{3-} or NO_x, which are expressed as

$$\log_{10}(\text{diat/dino}) \sim \text{te}\left(\text{N/P},\ \text{PO}_4^{3-}\right) + \text{s(DO)} + \text{s(PAR)} \quad (2)$$

$$\log_{10}(\text{diat/dino}) \sim \text{te}\left(\text{N/P},\ \text{NO}_x\right) + \text{s(DO)} + \text{s(PAR)} \quad (3)$$

Partial dependence plots and variable importance curves were generated using the *randomForest* package (version 4.7-1.2) in R (Breiman 2001; Tyralis et al. 2019). In addition, a cubic polynomial regression model was applied using the geom_smooth() function from the ggplot2 package (version 3.5.1) in R, with the formula *y ~ poly(x, 3)*, to explore the nonlinear relationships among diatom abundances, dinoflagellate abundances, the diat/dino ratio, and environmental factors.

Pyrosequencing of the 18S rRNA gene and bioinformatics analysis

The V9 hypervariable region of the 18S ribosomal RNA (rRNA) gene was amplified using standard primers 1389F (5′-TTGTACACACCGCCC-3′) and 1510R (5′-CCTTCY GCAGGTTCACCTAC-3′) (Amaral-Zettler et al. 2009). The PCR reaction was performed in a 50 μL reaction volume containing 25 μL of 2× Premix Taq, 2 μL of each primer, 2 μL of reverse primer, 2 μL of DNA template, and 19 μL of nuclease-free water. Thermal cycling consisted of initial denaturation at 94 °C for 30 s, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 52 °C for 30 s, extension at 72 °C for 30 s, and a final extension at 72 °C for 10 min. The amplicons were then gel-purified using the EZNA® Gel Extraction Kit (Omega, USA) and sequenced on a NovaSeq 6000 platform (Magigene, China).

High-quality 18S rRNA gene sequences were filtered using fastp (version 0.23.4) and analyzed with QIIME2, following the standard operating procedure previously described by Bolyen et al. (2019). The DADA2 pipeline was employed to generate denoised, chimera-free, and nonsingleton amplicon sequence variants (ASVs), along with their representative sequences (Callahan et al. 2016). The taxonomic assignment of the representative sequences was performed using DIAMOND (blastn, E-value ≤ 1e-5) against the PR2 database (version 5.0.0) (Buchfink et al. 2015). Sequences classified as Bacteria, Archaea, or Metazoa, as well as ASVs with fewer than 10 reads, were removed. The taxonomic ranks were assigned at the species, genus, or phylum level based on the sequence similarity thresholds of ≥97%, ≥95%, and ≥83%, respectively. Furthermore, species richness was calculated using the *qiime diversity* command in QIIME2 after rarefying all samples to a total of 13,500 reads.

The geographic distribution of microeukaryotic plankton, diatom, and dinoflagellate communities was visualized using the R package *scatterpie* (version 0.2.4). Cluster analysis of the microeukaryotic plankton communities was performed using the R package vegan (version 2.6-10) based on the Bray–Curtis dissimilarity at the family level. The relationship between the diat/dino ratio and relative abundances of the diatom and dinoflagellate families was analyzed using a combination of RF models and Spearman's correlation. In addition, GAMs were applied to explore associations between the key families and environmental factors. All RF, Spearman's correlation, and GAMs analyses were performed with $\log_{10}$-transformed data to ensure normalization.

Results

Geographic distribution patterns of diatom and dinoflagellate abundances

During the investigation, the abundance of diatoms in China's coastal waters ranged from 3.9×10^2 cells L^{-1} to 4.8×10^6 cells L^{-1}, while dinoflagellate abundances ranged from 48 cells L^{-1} to 5.8×10^6 cells L^{-1} (Fig. 1). Spatially, diatoms and dinoflagellates exhibited distinct distribution patterns. Diatoms were predominantly found in the nearshore stations and estuarine brackish waters. Hotspots of increased diatom abundances (≥ 10^6 cells L^{-1}) were observed at stations Z10_3 (4.8×10^6 cells L^{-1}) and Z12_1 (2.2×10^6 cells L^{-1}), both of which were located in the diluted waters of the PRE. Additional hotspots were found at station ZH_8 (4.5×10^6 cells L^{-1}), which was influenced by river run-off from the Oujiang River, and at station D45 (1.0×10^6 cells L^{-1}) in DYB (Fig. 1A). In contrast, dinoflagellates generally exhibited a higher abundance than diatoms at offshore stations. The identified dinoflagellate hotspots were at stations C1_1 (4.9×10^6 cells L^{-1}) and C2_3 (5.8×10^6 cells L^{-1}), which were located in the nearshore region north of the CJE (Fig. 1B).

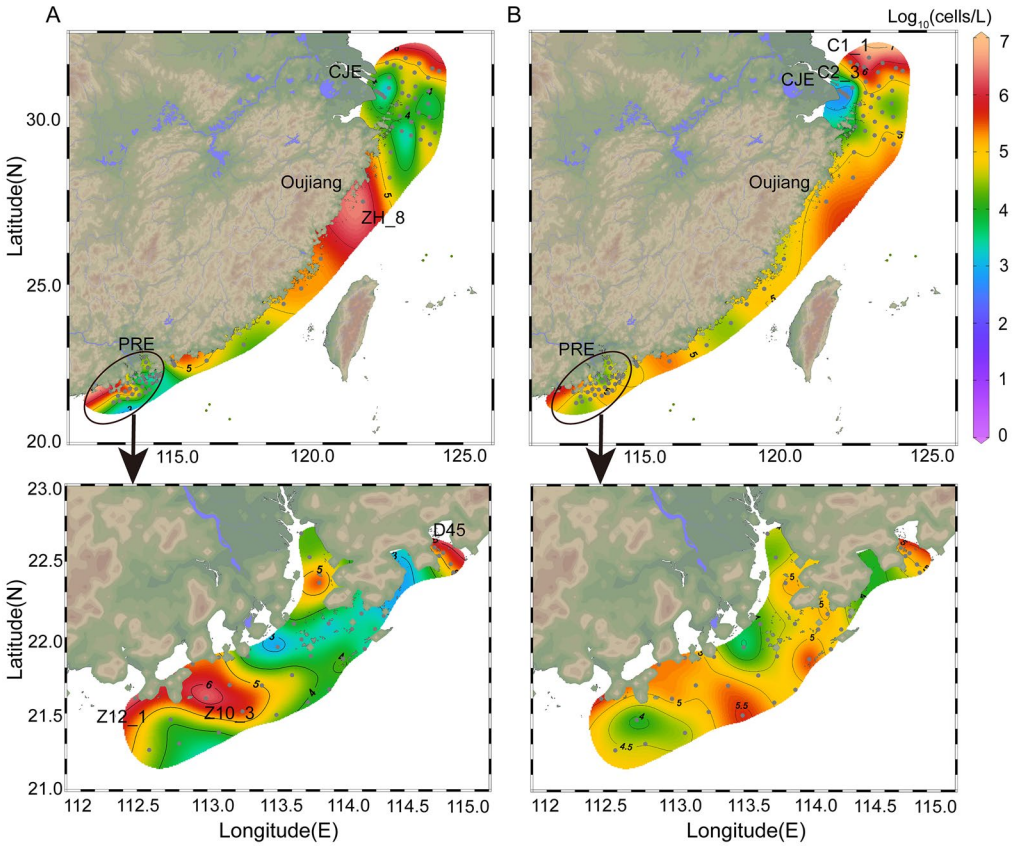

Fig. 1 Abundance of diatoms and dinoflagellates in the study region. **A** Diatom abundance. **B** Dinoflagellate abundance. Cell concentrations (log_{10}-transformed, cells L^{-1}) are visualized using different color gradients on contour maps, which were generated using Ocean Data View

Key environmental drivers of diat/dino ratio variations

The diat/dino ratio exhibited substantial spatial variability across the study area, ranging from 0.01 to 34.10. The elevated diat/dino ratios were primarily observed in the CJE and western nearshore zones of the PRE, whereas the low ratios were primarily detected in the northern nearshore regions of the CJE and southeastern regions of the PRE (Fig. 2A). The RF analysis results revealed that the DO, N:P ratio, and Si:N ratio were the most important factors that influenced the diat/dino ratio, with increase in mean squared error (IncMSE) values of 23.07%, 19.13%, and 21.49%, respectively (Fig. 2B). Results from further analysis using cubic polynomial regression showed that diatom abundance was significantly negatively correlated with the N:P ratio ($p < 0.01$; Fig. 2C) and positively associated with the Si:N ratio when values were below 8 ($p = 0.05$; Fig. 2C). In contrast, the abundance of dinoflagellates exhibited a significant positive correlation with DO concentration ($p < 0.01$; Fig. 2D), which suggests that these two phytoplankton groups possess competitive advantages under different biogeochemical conditions.

Diversity and community composition of diatoms and dinoflagellates across regions with different diat/dino ratios

The relative abundances of diatoms and dinoflagellates in the eukaryotic communities are presented in Fig. S2. Consistent with the spatial diat/dino ratio pattern, diatoms displayed a higher relative abundance than dinoflagellates in the estuarine regions and western nearshore zones of the PRE. In contrast, dinoflagellates were more abundant in the adjacent coastal waters of both the CJE and PRE (Fig. S2A, B).

Based on the cluster analysis, we identified two major groups within the eukaryotic communities, where one was dominated by dinoflagellates (Cluster I) and the other by diatoms (Cluster II) (Fig. 3). Cluster I was further divided into four subclusters, each reflecting distinct variations in the dominant dinoflagellate families. Subcluster 1.1 was predominant at stations from the ECS and exhibited a relatively balanced distribution of dominant families, which included Dinophyceae_XX (17.49%), Dino-Group-II-Clade-22 (8.52%), Gymnodiniaceae (8.15%), and Pyrocystaceae (7.17%). Subcluster 1.2, representing the DYB stations, was largely dominated by Dinophyceae_XX (33.67%) and

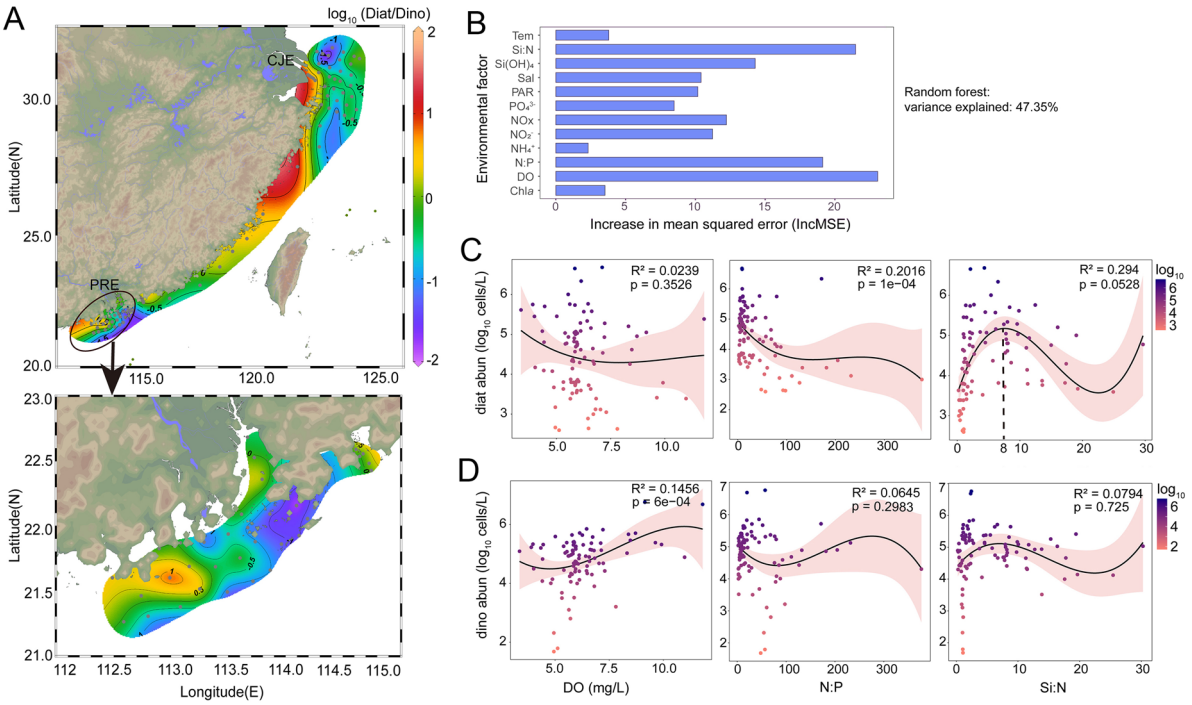

Fig. 2 Spatial patterns of the diatom-to-dinoflagellate ratio (diat/dino ratio) and key environmental drivers in the study region. **A** Spatial distribution of the diat/dino ratio (log_{10}-transformed). **B** Variable importance ranking identified using the random forest (RF) model showing the environmental factors that influence the diat/dino ratio (log_{10}-transformed). A higher increase in the mean squared error (IncMSE) values indicates greater importance. NO_x refers to the combined concentrations of nitrate (NO_3^-) and nitrite (NO_2^-). Cubic polynomial regressions for diatom abundance **C** and dinoflagellate abundance **D** (both log_{10}-transformed) with the DO, N:P ratio, and Si:N ratio

Dino-Group-II-Clade-22 (8.45%). Subcluster 1.3, which primarily consisted of offshore stations in the ECS and SCS, was overwhelmingly dominated by Dinophyceae_XX (42.31%). Furthermore, subcluster 1.4, covering stations from the northern CJE and southern/eastern PRE, was dominated by Thoracosphaeraceae (23.34%), followed by Dinophyceae_XX (16.10%), Pyrocystaceae (10.14%), Peridiniaceae (9.95%), and Heterocapsaceae (8.62%) (Fig. S3).

In addition, Cluster II was further divided into two subclusters: 2.1 and 2.2. Subcluster 2.1, primarily consisting of the estuarine stations, was dominated by the diatom families Thalassiosiraceae (18.99%) and Chaetocerotaceae (12.97%). In contrast, subcluster 2.2, which encompassed stations from DYB and coastal regions of MZC, was characterized by a higher relative abundance of Leptocylindraceae (10.56%), Bacillariaceae (9.53%), Skeletonemaceae (8.74%), and Thalassiosiraceae (8.01%) (Fig. S3).

To investigate how the nutrient conditions at the stations influenced the diat/dino ratio under elevated N:P scenarios, we focused on two representative subclusters, namely subcluster 1.4, a dinoflagellate-dominated nearshore region, and subcluster 2.1, a diatom-dominated estuarine region. Both subclusters exhibited elevated N:P ratios (Fig. S4), but the PO_4^{3-} and NO_x concentrations were markedly lower in subcluster 1.4 than in subcluster 2.1. GAMs analysis identified opposite patterns in the diat/dino ratio responses to nutrient conditions. Specifically, when the N:P ratio ranged approximately from 30 to 100, the diat/dino ratio increased as the concentrations of PO_4^{3-} and NO_x rose. In contrast, under relatively low PO_4^{3-} and NO_x conditions, the diat/dino ratio declined as the N:P ratio increased (Fig. 4).

We further examined the diatom and dinoflagellate families that were associated with algal blooms in the coastal waters of China during summer. Based on previous studies, the bloom thresholds for diatoms and dinoflagellates are 2×10^6 and 1×10^6 cells L^{-1}, respectively (Costa et al. 2021; Gong et al. 2017). In this study, diatom blooms were detected at stations Z10_3 (4.8×10^6 cells L^{-1}), Z12_1 (2.2×10^6 cells L^{-1}), and ZH_8 (4.5×10^6 cells L^{-1}), while dinoflagellate blooms occurred at stations C1_1 (4.9×10^6 cells L^{-1}) and C2_3 (5.8×10^6 cells L^{-1}). The dominant families responsible for the diatom blooms varied across stations. At station Z10_3, Bacillariaceae (20.21%), Skeletonemaceae (14.40%), and Hemiaulaceae (12.32%) were the most abundant bloom-associated families. In contrast, at station Z12_1, the bloom-associated families were dominated by Leptocylindraceae (24.43%) and

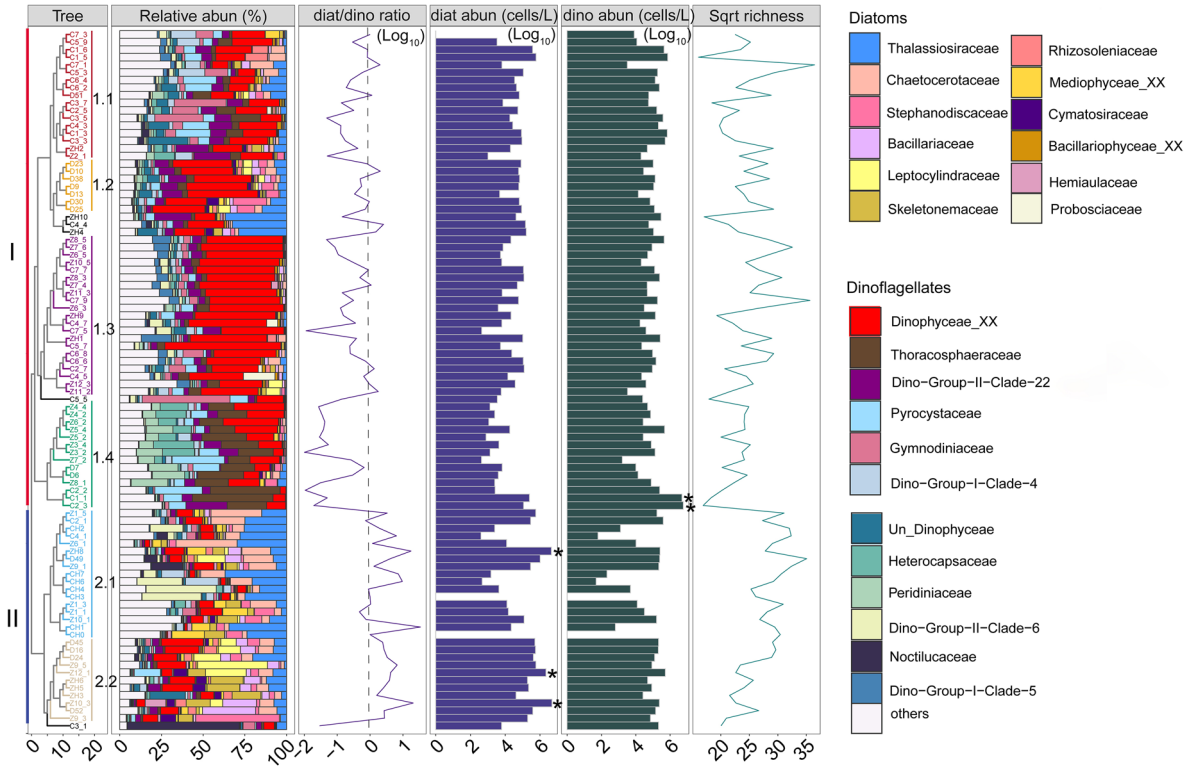

Fig. 3 Community composition, diatom-to-dinoflagellate ratio (diat/dino ratio) and diversity across clusters. The dendrogram, which was constructed based on Bray–Curtis similarity of diatom and dinoflagellate community composition, was accompanied by stacked bar plots illustrating their relative abundances at the family level. The line plot in the middle represents the spatial variation of the diat/dino ratio (log_{10}-transformed) across the study area. The two bar plots depict the distribution of diatom and dinoflagellate abundance, with asterisks (*) indicating stations where blooms were observed. The line plot on the right indicates richness (square root-transformed)

Skeletonemaceae (10.48%). At station ZH_8, Chaetocerotaceae (17.16%) and Mediophyceae_XX (11.15%). In the case of dinoflagellate blooms, stations C1_1 and C2_3 were both predominantly occupied by Thoracosphaeraceae, with relative abundances of 48.34% and 23.64%, respectively. Further analysis revealed that within Thoracosphaeraceae, *Scrippsiella* was the primary bloom-forming genus, with an average relative abundance of 31.68% at the bloom sites (Table S1). In addition, a similar distribution pattern was observed between the diat/dino ratio and overall phytoplankton species richness across the surveyed stations. Furthermore, a regression model revealed a significant positive correlation between the overall phytoplankton species richness and diat/dino ratio ($p < 0.01$; Fig. S5), which suggests that a higher dominance of dinoflagellates may negatively affect the overall phytoplankton diversity.

Key environmental drivers of diatom and dinoflagellate community composition

An RF model combined with Spearman correlation analysis identified the key taxa that contributed to the diat/dino ratio shifts. Among diatoms, the families Thalassiosiraceae, Chaetocerotaceae, Bacillariaceae, and Hemiaulaceae played significant roles, with high IncMSE values of 13.30%, 19.46%, 16.80%, and 15.57%, respectively, and strong positive correlations with the diat/dino ratio ($r > 0.6$; $p < 0.001$; Fig. 5). Spearman correlation analysis further revealed that these four key families were significantly negatively correlated with DO and significantly positively correlated with Chl *a*, $Si(OH)_4$, and PO_4^{3-} ($p < 0.05$; Fig. 5). In addition, Bacillariaceae exhibited a significant negative correlation with the N:P ratio ($p < 0.05$; Fig. 5). Among dinoflagellates, Thoracosphaeraceae, Dino_Group_II_Clade_22, and Peridiniaceae were identified as the key contributors, revealing IncMSE values of 14.02%, 11.45%, and 8.03%, respectively, and a significant negative correlation with the diat/dino ratio ($r < -0.4$; $p < 0.001$). Moreover, Spearman correlation analysis revealed a significant positive correlation between the families Thoracosphaeraceae and Peridiniaceae with DO, while both families exhibited a significant negative correlation with $Si(OH)_4$ and PO_4^{3-} ($p < 0.05$; Fig. 5). Dino_Group_II_Clade_22 only revealed a significant negative correlation with temperature ($p < 0.05$; Fig. 5).

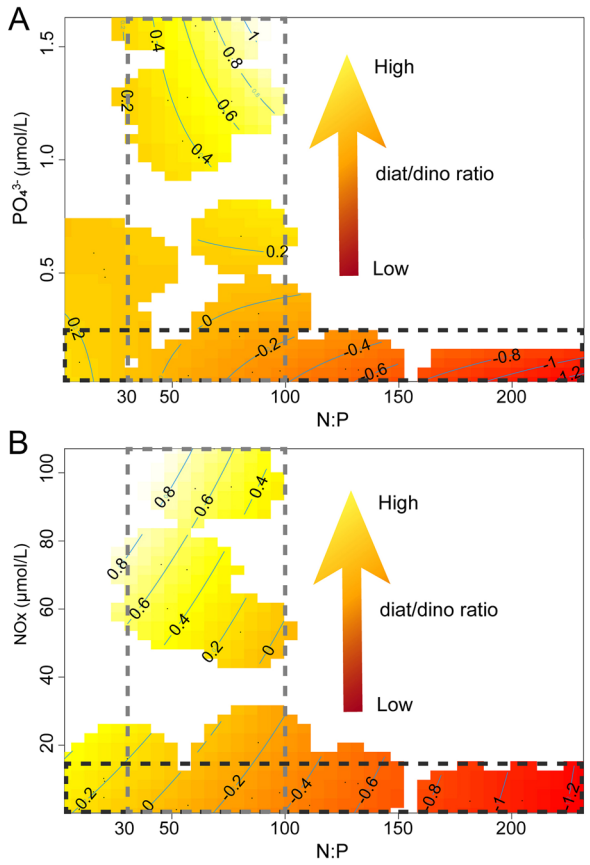

Fig. 4 Generalized additive models (GAMs) reveal the dynamics of the diatom-to-dinoflagellate ratio (diat/dino ratio) in response to pairwise interactions between the N:P ratio and PO_4^{3-} or NO_x. **A** N:P ratio and PO_4^{3-}. **B** N:P ratio and NO_x. The gray dashed boxes indicate the response patterns of the diat/dino ratio to changes in PO_4^{3-} or NO_x concentrations within an N:P range of approximately 30–100. The black dashed boxes highlight the response of the diat/dino ratio to variations in the N:P ratio under relatively low PO_4^{3-} and NO_x concentrations. Red areas indicate low values of the diat/dino ratio, and yellow areas indicate high values. NO_x refers to the combined concentrations of NO_3^- and NO_2^-

The environmental factor relationships were inversely correlated between dinoflagellates and diatoms.

To further validate these findings, GAMs were applied to the identified key phytoplankton families and taxa characterized by high relative abundance or ecological significance, which included the diatom families Leptocylindraceae and Skeletonemaceae and the dinoflagellate families Dinophyceae_XX and Pyrocystaceae (Figs. 6, S6). The results confirmed significant differences between diatoms and dinoflagellates for the environmental responses. Within each group, different families exhibited similar response patterns but varied in sensitivity. Notably, after controlling for other environmental factors, Thalassiosiraceae, Chaetocerotaceae, Bacillariaceae, and Skeletonemaceae revealed significant negative partial effects of temperature at 27.5 °C–30 °C, whereas Thoracosphaeraceae, Dinophyceae_XX, and Pyrocystaceae displayed positive effects of temperature, thus indicating a preference for higher temperatures among dinoflagellates ($p < 0.05$; Fig. 6). Regarding the $Si(OH)_4$ concentrations, Thalassiosiraceae, Chaetocerotaceae, and Bacillariaceae were positively correlated with $Si(OH)_4$ ($p > 0.05$), while Thoracosphaeraceae, Dino_Group_II_Clade_22, Dinophyceae_XX, and Pyrocystaceae were significantly negatively correlated ($p < 0.05$; Figs. 6, S6). This reflects the greater $Si(OH)_4$ dependency in diatoms and inhibitory effects on dinoflagellates. The N:P ratio was negatively correlated with Thalassiosiraceae, Chaetocerotaceae, and Bacillariaceae, with significant negative effects on Thalassiosiraceae and Bacillariaceae ($p < 0.05$). In contrast, Thoracosphaeraceae and Dino_Group_II_Clade_22 exhibited positive correlations with the N:P ratio ($p > 0.05$). In addition, Thalassiosiraceae, Chaetocerotaceae, and Bacillariaceae were negatively correlated with DO ($p > 0.05$), while Thoracosphaeraceae and Pyrocystaceae exhibited significant positive correlations ($p < 0.05$), thereby suggesting a preference among dinoflagellates for high DO conditions. Consistent with these findings, Thoracosphaeraceae and Pyrocystaceae were dominant in stations within Cluster 1.4, which demonstrated significantly higher DO levels than other stations ($p < 0.05$; Table S2), along with elevated N:P ratios and lower $Si(OH)_4$ concentrations (Fig. S4).

Discussion

In this study, we applied an integrative analytical framework that combined RF and GAMs to identify the key environmental drivers that influenced the spatial variability in the diat/dino ratio across China's coastal waters during summer. This multi-model approach outperformed traditional single-model methods in robustness and interpretability and highlighted the dominant roles of $Si(OH)_4$ availability, N:P ratios, and DO in shaping the phytoplankton community structure. By incorporating high-throughput sequencing data, we further resolved the spatial niches of the key taxa at a high resolution, thereby bridging crucial gaps in understanding phytoplankton responses to varying environmental conditions (Fig. 7). Together, these insights advance our mechanistic understanding of nutrient stoichiometry and oxygen dynamics in regulating the dominance of diatoms and dinoflagellates, which offers valuable implications for the prediction and mitigation of harmful algal blooms within eutrophic coastal ecosystems.

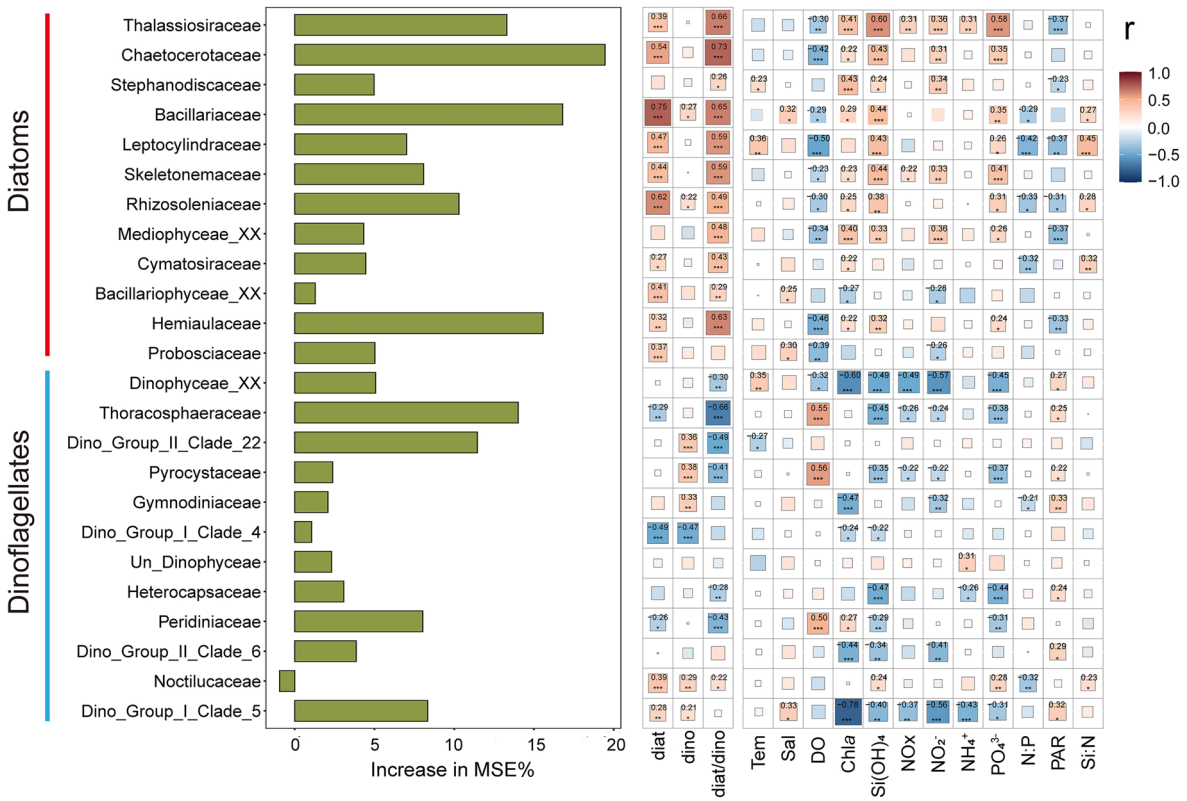

Fig. 5 Influence of key families and environmental factors on the diatom-to-dinoflagellate ratio (diat/dino ratio). An RF model ranks the importance of diatom and dinoflagellate families in influencing the diat/dino ratio (log_{10}-transformed), with an IncMSE% representing their contribution to the ratio's variation. The correlation heatmap depicts the relationships among specific families, diatom and dinoflagellate abundances, the diat/dino ratio, and environmental factors. Red and blue squares represent positive and negative correlations, respectively. Significant correlations are depicted by the size of the squares and marked with asterisks (***, $p < 0.001$; **, $p < 0.01$; and *, $p < 0.05$)

Key environmental drivers of the diat/dino ratio–Si:N ratio

RF analyses revealed that the diat/dino ratio in the coastal waters of China during summer was primarily influenced by the Si:N ratio, N:P ratio, and levels of DO availability (Fig. 2). Diatoms dominated in regions characterized by abundant $Si(OH)_4$, Si:N ratios ranging from 1 to 8, and N:P ratios of approximately 16:1, whereas dinoflagellates were more prevalent in areas characterized by higher N:P ratios and DO levels (Fig. 7). These findings indicate that diatom abundances increase when the Si:N ratio is below 8, thus emphasizing the essential role of Si:N stoichiometric balance in supporting the growth of diatoms (Fig. 2C). This pattern aligns with the physiological demand of diatoms for silicate, which is essential for the synthesis of their siliceous frustules and completion of their cell cycle (Martin-Jézéquel et al. 2000). Silicate availability directly regulates the diatom cell cycle and morphology. The diatom cell cycle is classically divided into four phases: G1, S, G2, and M (Inzé and De Veylder 2006). Under Si-limited conditions (i.e., an Si:N ratio lower than 0.02), diatom cells are arrested in the G2 and M phases, thus preventing division and leading to cell elongation (Brzezinski et al. 2024; Hoffmann et al. 2007; Xu et al. 2014). Furthermore, low Si:N ratios alter the diatom morphology, increasing their vulnerability to grazing (Assmy et al. 2013; Durbin 1977; Liu et al. 2016). For instance, some species, such as *Chaetoceros*, rely on sufficient silicate to form longer chains, and when Si:N ratios are low, these species produce shorter chains, which are more susceptible to grazing, thus further reducing their abundance (Hoffmann et al. 2007). In the ECS, reduced silicate input because of the regulation of riverine discharge by the Three Gorges Dam may lower the Si:N ratios, thereby limiting diatom growth and altering their morphology. Over time, this could contribute to the long-term decline in diatom abundance (Zhou et al. 2016).

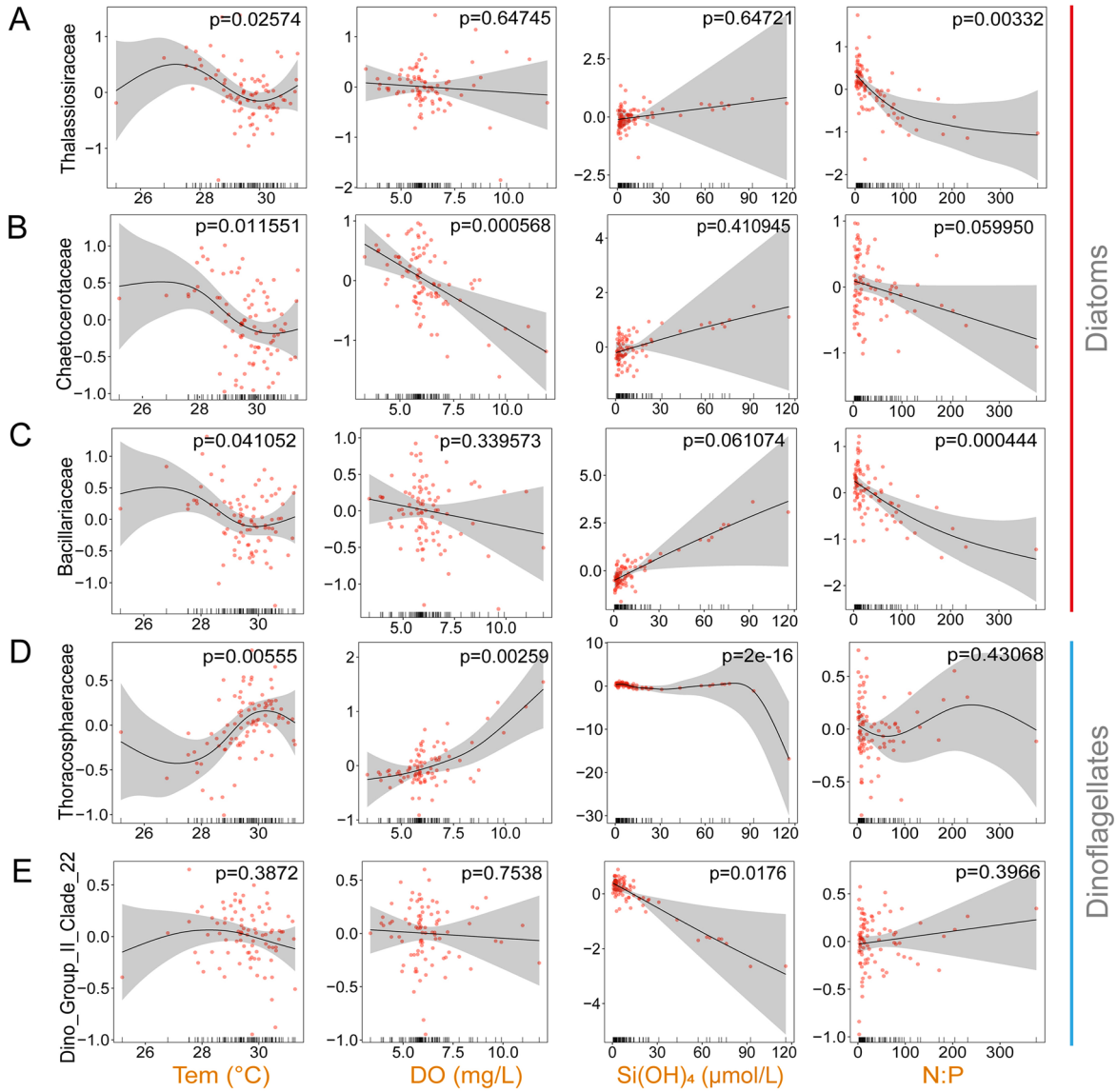

Fig. 6 GAMs reveal the partial effects of temperature, DO, $Si(OH)_4$ and N:P ratios on the relative abundance of key families (log_{10}-transformed). **A** Thalassiosiraceae, **B** Chaetocerotaceae, **C** Bacillariaceae, **D** Thoracosphaeraceae, and **E** Dino_Group_II_Clade_22. The shaded areas represent the 95% confidence intervals, and tick marks on the x-axis represent the distribution of the observed data

Key environmental drivers of the diat/dino ratio–N:P ratio

Disparities in nitrogen and phosphorus supply also significantly affected the diat/dino ratio. In this study, high N:P ratios were observed in the estuarine stations for subcluster 2.1 and nearshore stations for subcluster 1.4 (Figs. 7, S4), which is consistent with the stoichiometric shifts reported in previous studies (Jiang et al. 2014; Zhou et al. 2008). For example, substantial nitrogen inputs in the ECS have resulted in the dramatic increase in N:P ratios, where the ratio increased from 40 in the 1960s to 200 in the 2010s (Moon et al. 2021; Ou et al. 2020). Subcluster 1.4 was dinoflagellate-dominated, likely because of the limited availability of phosphorus (Figs. 3, S4). This observation is consistent with the resource allocation theory proposed by Klausmeier et al. (2004), which posits that under high N:P ratios and phosphorus-limited conditions, phytoplankton allocate more resources toward phosphorus acquisition and thus favor taxa with inherently higher cellular N:P stoichiometry. Dinoflagellates, which generally exhibit higher internal N:P ratios and lower phosphorus requirements than diatoms (John and Flynn 2000; Rhee and Gotham 1980), are therefore better adapted to thrive in high N:P, phosphorus-depleted

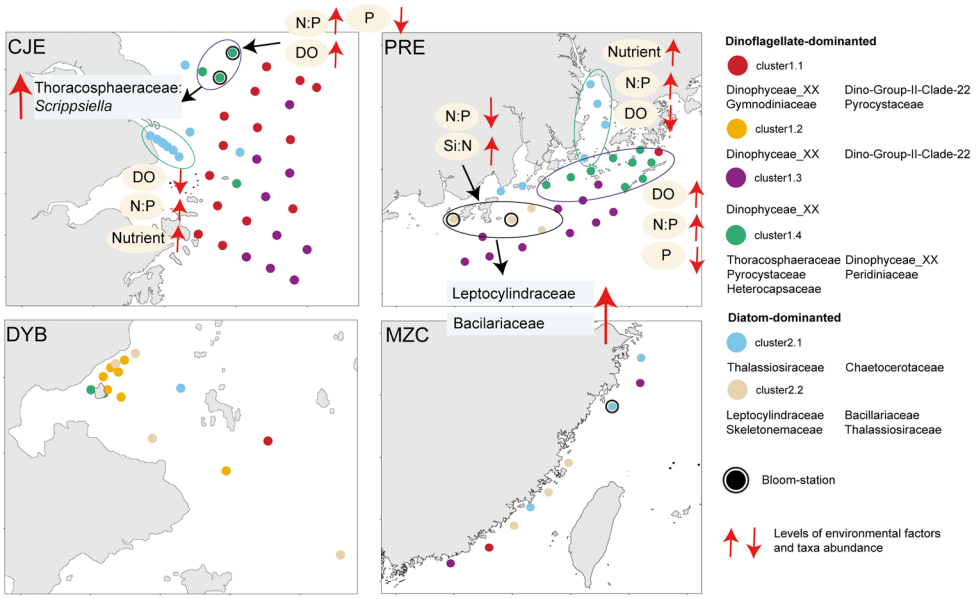

Fig. 7 Conceptual diagram of the diatom-to-dinoflagellate ratio (diat/dino ratio) across different clusters in China's coastal waters. Colored circles represent different clusters. Red arrows indicate high (↑) or low (↓) levels of environmental factors or taxa relative abundance. The map shows a simplified coastline for visual clarity and does not represent national boundaries. CJE: Changjiang River estuary; PRE: Pearl River estuary; DYB: Daya Bay; MZC: coastal waters off Fujian and Zhejiang

environments (Xiao et al. 2018; Zhang et al. 2024). Consistent with this mechanism, previous studies in the ECS have demonstrated that rising N:P ratios and warming, together with PO_4^{3-} concentrations dropping as low as 0.01 umol/L, contributed to the long-term decline in diat/dino ratios (Xiao et al. 2018). Similarly, Zhang et al. (2024) reported that limited phosphorus concentrations under high N:P ratio conditions significantly constrained diatom growth, and they were outcompeted by dinoflagellates in phosphorus-depleted environments. However, subcluster 2.1 was diatom-dominated (Figs. 3, S4), likely because sufficient phosphorus and nitrogen availability support diatom growth under high N:P conditions. Our GAM analysis supports this notion, revealing that within an N:P ratio range of approximately 30–100, increased PO_4^{3-} and NO_x concentrations were positively associated with higher diat/dino ratios (Fig. 4). This pattern aligns with findings from Hong Kong coastal waters, where the relatively high PO_4^{3-} (0.65–1.29 umol/L) and rising NO_3^- levels were associated with the elevated diat/dino ratios (Cheung et al. 2021). Furthermore, experimental studies have shown that the diatom *Phaeodactylum tricornutum* maintains a competitive edge over the dinoflagellate *Prorocentrum minimum* under nutrient-replete conditions (P > 70 umol/L; N > 4400 umol/L), regardless of the temperature conditions and N:P ratio (Bi et al. 2021). Overall, our findings suggest that while high N:P ratios generally favor dinoflagellates, this effect is modulated by nutrient availability. Under phosphorus-limited conditions, increasing N:P ratios exacerbate the growth suppression of diatoms, thereby enhancing dinoflagellates' dominance. Conversely, in nutrient-rich environments, additional nitrogen inputs may stimulate diatom growth and help them maintain their competitive advantage. These results emphasize the crucial role of phosphorus availability in shaping the phytoplankton community composition, indicating that the ecological outcome of nitrogen enrichment is highly dependent on the concurrent phosphorus supply.

Key environmental drivers of the diat/dino ratio–DO

This study also found a significant negative correlation between DO and the diat/dino ratio. However, the underlying mechanisms behind this relationship are likely to be complex. On the one hand, phytoplankton can directly affect oxygen levels within aquatic ecosystems through photosynthesis. Moreover, intense phytoplankton blooms can lead to oxygen depletion due to the increased activity of heterotrophic bacteria that accompany the bloom (Chen et al. 2024a). On the other hand, changes in oxygen levels can indirectly influence phytoplankton communities. Oxygen availability can affect grazers and microbial communities that interact with phytoplankton, thereby indirectly shaping the community structure (Li et al. 2022). In addition, some phytoplankton, particularly mixotrophs, are known to prefer oxic conditions, thus further linking oxygen levels to phytoplankton dynamics (Cohen et al. 2021; Jeong et al.

2010). In this study, estuarine stations dominated by diatoms (subcluster 2.1) exhibited lower DO levels (mean: 4.9 mg L^{-1}; Fig. S4), likely because of organic matter degradation driven by nutrient-rich inputs (Gray et al. 2002; Wang et al. 2017). The prevalence of diatoms in low-oxygen waters is consistent with previous findings (Broman et al. 2017) and may be explained by the experimental evidence that demonstrated that reduced oxygen enhanced diatom growth by limiting photorespiration and mitochondrial activity (Sun et al. 2022). Rapid diatom blooms, such as at station ZH_8, may further deplete the oxygen levels and reinforce hypoxia through positive feedback (Mohd-Din et al. 2020; Wallace and Gobler 2021). In contrast, the dinoflagellate-dominated stations (subcluster 1.4) exhibited higher DO levels (mean: 7.9 mg L^{-1}), which suggests that oxygen-rich waters may favor dinoflagellates because of their mixotrophic characteristics and dependence on oxidative metabolism (Cohen et al. 2021; Jeong et al. 2010). For instance, *Lepidodinium* sp. enhances energy production via oxidative phosphorylation when prey availability supports mixotrophy (Chen et al. 2024b). These findings highlight the complex role of oxygen in the structure of phytoplankton communities. While diatoms and dinoflagellates produce oxygen through photosynthesis, their contrasting responses to DO levels suggest that oxygen availability is a key, yet often overlooked, driver of community composition. Future models of phytoplankton succession should incorporate these DO dynamics to better capture these interactions.

Overall, the spatial heterogeneity of Si:N ratios, N:P ratios, and DO levels collectively shapes the diat/dino ratio. Offshore regions were primarily limited by silicate availability, whereas nearshore and estuarine zones were influenced by nutrient imbalances. Moreover, oxygen availability further modulates the phytoplankton community structure. These spatial patterns emphasize the need for region-specific strategies to manage eutrophication and anticipate the climate-driven shifts in phytoplankton dynamics.

Taxon-specific contributions to shifts in the diat/dino ratio across environmental conditions

The present study further identified key phytoplankton groups that may drive a shift in the diat/dino ratio in the coastal waters of China during summer. These findings reveal that different diatom families dominated in the two regions with high diat/dino ratios (subclusters 2.1 and 2.2 in Fig. 3). In the estuarine environments (subcluster 2.1), which were characterized by low-salinity and high-nutrient loads, diatom communities were dominated by the Thalassiosiraceae and Chaetocerotaceae families. The dominance of these families suggests their adaptation to high-nutrient, low-salinity conditions (Alverson et al. 2007; Bharathi et al. 2022; De Luca et al. 2019). In contrast, the Leptocylindraceae, Bacillariaceae, Skeletonemaceae, and Thalassiosiraceae families prevailed in the nearshore regions with low N:P ratios and sufficient $Si(OH)_4$ availability (subcluster 2.2, Fig. S4), thus indicating the role of nutrient stoichiometry in shaping the taxonomic composition (Eissler et al. 2025; Pang et al. 2024; Takabayashi et al. 2006). These diatom families are known for their rapid proliferation under favorable conditions, which often results in bloom formation (Kužat et al. 2022; Ma et al. 2022; Marić Pfannkuchen et al. 2018). Such conditions likely contributed to the bloom development in subcluster 2.2, as seen at stations with low N:P ratios and high $Si(OH)_4$ concentration (e.g., Z10_3 and Z12_1). These cases highlight how nutrient conditions characterized by low N:P ratios and sufficient silicate availability in nearshore regions can promote diatom dominance and trigger bloom formation. Furthermore, our results highlight niche differentiation among the diatom taxa, where species-specific ecological preferences drive spatial heterogeneity within the community structure. This differentiation has considerable ecological implications, as shifts in diatom assemblages directly influence nutrient cycling, primary production, and carbon export in coastal ecosystems (Fu et al. 2022; Henson et al. 2019; Tréguer et al. 2018). Given that the different diatom families exhibited distinct ecological niches and responded variably to nutrient stoichiometry, understanding these patterns is essential for predicting ecosystem responses to nutrient enrichment.

Our study revealed a significant positive correlation between the diat/dino ratio and phytoplankton species richness (Fig. S5), which suggests that shifts toward dinoflagellate dominance may contribute to biodiversity loss. This aligns with previous studies revealing that elevated dinoflagellate abundances can suppress other phytoplankton groups through resource competition, mixotrophy, water condition modifications, and the release of allelopathic compounds, which ultimately reduces species richness and ecosystem stability (Cheung et al. 2021; Gu et al. 2013; Jeong et al. 2010).

Extreme diat/dino ratios that were observed in the dinoflagellate-dominated regions were closely associated with the proliferation of specific dominant taxa, thereby highlighting species-specific ecological preferences and the niche differentiation across environmental conditions. An unclassified lineage within Dinophyceae, denoted as Dinophyceae_XX in the PR2 database (Guillou et al. 2012), was consistently abundant and exhibited peak relative abundance in the offshore zones that were characterized by low-nutrient levels (subcluster 1.3 in Fig. 3). The prevalence of this unclassified lineage in oligotrophic waters suggests that it may comprise taxa with adaptations to nutrient-poor conditions. Notably, numerous dinoflagellates possess mixotrophic capabilities, which combine phototrophy and heterotrophy as a flexible nutritional strategy (Lambert

et al. 2022; Mitra et al. 2023). Although the trophic mode of Dinophyceae_XX remains unresolved, its dominance within oligotrophic environments suggests that mixotrophy may underlie its ecological success. Such traits support the findings of previous studies revealing that mixotrophic dinoflagellates are often well suited to low-nutrient, stratified offshore environments, where mixotrophy provides a competitive advantage (Cohen et al. 2021; Stamieszkin et al. 2024). In contrast, nearshore waters with high N:P ratios and elevated DO levels were predominantly characterized by the prevalence of the Thoracosphaeraceae, Pyrocystaceae, Peridiniaceae, and Heterocapsaceae families (subcluster 1.4 in Fig. 3). Among these, *Scrippsiella* (Thoracosphaeraceae) emerged as a key bloom-forming genus. Its dominance in high-nutrient, oxygen-rich environments is consistent with previous studies demonstrating that *Scrippsiella* thrives under eutrophic conditions and elevated N:P ratios (Ge et al. 2012). Similar to numerous dinoflagellates, *Scrippsiella* is mixotrophic (Cooper et al. 2016; Jeong et al. 2005), capable of photosynthesis and phagotrophy via endocytosis, an oxygen-consuming process that may explain the prevalence of blooms in well-oxygenated environments (Beisner et al. 2019; Cohen et al. 2021). In addition to nutrient stoichiometry and DO levels, elevated surface water temperatures during summer may further favor the dominance of *Scrippsiella*. Laboratory and field observations have shown that *Scrippsiella trochoidea* grows rapidly at temperatures of approximately 30 °C to 32.1 °C (Tian et al. 2021). Moreover, *Scrippsiella* readily forms resting cysts under adverse conditions, thereby enhancing its persistence, dispersal capacity, and potential for bloom recurrence (Li et al. 2025). Historically, *Scrippsiella*-dominated blooms were largely confined to the SCS (Tian et al. 2021; Wang et al. 2011). However, recent observations, including those from the present study, indicate a northward expansion (Wang and Wu 2009), with blooms at stations C1_1 and C2_3 in the ECS showing clear *Scrippsiella* dominance. With the increase in its dispersal and recurrence potential, *Scrippsiella* is increasingly contributing to biodiversity loss and ecosystem instability. Its blooms have been shown to exert rapid and lethal effects on the larvae of economically important shellfish species, such as *Crassostrea virginica* and *Mercenaria mercenaria* (Tang and Gobler, 2012), thus emphasizing the sustained ecological and economic risks associated with its expansion.

Conclusions

This study reveals the crucial roles of Si:N ratios, N:P ratios, and DO levels in regulating the distribution and dominance of diatoms and dinoflagellates within coastal ecosystems. Low Si:N ratios limit diatoms' growth, thereby favoring dinoflagellates, especially under nutrient-imbalanced conditions. Elevated N:P ratios further reinforce this shift by suppressing diatom abundances and enhancing dinoflagellate proliferation, thus emphasizing the urgency of nutrient management to prevent HABs and their ecological consequences. Moreover, the positive correlation between DO and dinoflagellate abundances (Figs. 2D, 7) suggests that well-oxygenated environments may promote dinoflagellate dominance, thereby adding another layer of complexity to phytoplankton dynamics. These findings highlight the need for integrated management approaches that consider nutrient stoichiometry and oxygen dynamics to maintain balanced phytoplankton communities and mitigate the risk of HABs.

Future research should focus on the long-term monitoring of nutrient stoichiometry and oxygen fluctuations to predict shifts in phytoplankton community structures. Integrating high-throughput sequencing with metatranscriptomics could provide valuable insights into the metabolic responses of diatoms and dinoflagellates to environmental changes (Xia et al. 2024). In parallel, laboratory experiments simulating nutrient limitations may help clarify the physiological mechanisms that underlie diatom–dinoflagellate competition. Collectively, these efforts will inform more effective strategies for managing coastal eutrophication and reducing the ecological and socioeconomic impacts of HABs.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42995-025-00343-4.

Acknowledgments This work was supported by the National Key Research and Development Program of China (2022YFC3105301), the National Natural Science Foundation of China (32470101, 42188102), the development fund of South China Sea Institute of Oceanology of the Chinese Academy of Sciences (SCSIO202205), the Ocean Negative Carbon Emissions (ONCE) Program, the Science and Technology Planning Project of Guangdong Province (2023B1212060047), and the Guangdong Basic and Applied Basic Research Foundation (2023A1515012095). Shaorong Fang from the Information and Network Center of Xiamen University is acknowledged for the help with the high-performance computing.

Author contributions LZ, HL, DW and XX conceived of the presented idea. LZ, ZW, JY, SS, ZC and HZ conducted the field sampling and laboratory work. LZ and XX performed statistical analysis and visualization. LZ wrote the manuscript with support from XX. KZ and YT provided academic guidance and critical revisions. All the authors reviewed and approved the final manuscript.

Data availability The sequences obtained in this study have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject numbers: PRJNA1255418.

Declarations

Conflict of interest The authors declare that they have no conflict of interest. The author Dazhi Wang is one of the editorial board members, but he was not involved in the journal's review of, or decision related to, this manuscript.

Human and animal rights This article does not contain any studies with human participants or animals performed by the authors.

References

Agirbas E, Koca L, Aytan U (2017) Spatio-temporal pattern of phytoplankton and pigment composition in surface waters of southeastern Black Sea. Oceanologia 59:283–299

Alverson AJ, Jansen RK, Theriot EC (2007) Bridging the Rubicon: phylogenetic analysis reveals repeated colonizations of marine and fresh waters by thalassiosiroid diatoms. Mol Phylogenet Evol 45:193–210

Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM (2009) A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. PLoS One 4:e6372

Assmy P, Smetacek V, Montresor M, Klaas C, Henjes J, Strass VH, Arrieta JM, Bathmann U, Berg GM, Breitbarth E, Cisewski B, Friedrichs L, Fuchs N, Herndl GJ, Jansen S, Krägefsky S, Latasa M, Peeken I, Röttgers R, Scharek R et al (2013) Thick-shelled, grazer-protected diatoms decouple ocean carbon and silicon cycles in the iron-limited Antarctic Circumpolar Current. Proc Natl Acad Sci U S A 110:20633–20638

B-Béres V, Stenger-Kovács C, Buczkó K, Padisák J, Selmeczy GB, Lengyel E, Tapolczai K (2023) Ecosystem services provided by freshwater and marine diatoms. Hydrobiologia 850:2707–2733

Beisner BE, Grossart H-P, Gasol JM (2019) A guide to methods for estimating phago-mixotrophy in nanophytoplankton. J Plankton Res 41:77–89

Bharathi MD, Venkataramana V, Sarma VVSS (2022) Phytoplankton community structure is governed by salinity gradient and nutrient composition in the tropical estuarine system. Cont Shelf Res 234:104643

Bi R, Cao Z, Ismar-Rebitz SMH, Sommer U, Zhang H, Ding Y, Zhao M (2021) Responses of marine diatom-dinoflagellate competition to multiple environmental drivers: abundance, elemental, and biochemical aspects. Front Microbiol 12:731786

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK et al (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37:852–857

Breiman L (2001) Random Forests. Mach Learn 45:5–32

Broman E, Sachpazidou V, Dopson M, Hylander S (2017) Diatoms dominate the eukaryotic metatranscriptome during spring in coastal "dead zone" sediments. Proc R Soc Lond B Biol Sci 284:20171617

Brzezinski MA, Johnson L, Estapa M, Clevenger S, Roca-Martí M, Romanelli E, Buck KN, Jenkins BD, Jones JL (2024) Physical mechanisms sustaining silica production following the demise of the diatom phase of the north Atlantic spring phytoplankton bloom during exports. Glob Biogeochem Cycles 38:e2023GB008048

Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods 13:581–583

Chen C-C, Chou W-C, Hung C-C, Gong G-C (2024) Nutrient sources, phytoplankton blooms, and hypoxia along the Chinese coast in the East China Sea: insight from summer 2014. Mar Pollut Bull 205:116692

Chen J, Deng L, Pang M, Li Y, Xu Z, Zhang X, Liu H (2024) Transcriptomic insights into the shift of trophic strategies in mixotrophic dinoflagellate Lepidodinium in the warming ocean. ISME Commun 4:ycae087

Cheung YY, Cheung S, Mak J, Liu K, Xia X, Zhang X, Yung Y, Liu H (2021) Distinct interaction effects of warming and anthropogenic input on diatoms and dinoflagellates in an urbanized estuarine ecosystem. Glob Change Biol 27:3463–3473

Cohen NR, McIlvin MR, Moran DM, Held NA, Saunders JK, Hawco NJ, Brosnahan M, DiTullio GR, Lamborg C, McCrow JP, Dupont CL, Allen AE, Saito MA (2021) Dinoflagellates alter their carbon and nutrient metabolic strategies across environmental gradients in the central Pacific Ocean. Nat Microbiol 6:173–186

Cooper JT, Sinclair GA, Wawrik B (2016) Transcriptome analysis of *Scrippsiella trochoidea* CCMP 3099 reveals physiological changes related to nitrate depletion. Front Microbiol 7:639

Costa RR, Mendes CRB, Ferreira A, Tavano VM, Dotto TS, Secchi ER (2021) Large diatom bloom off the Antarctic Peninsula during cool conditions associated with the 2015/2016 El Niño. Commun Earth Environ 2:252

De Luca D, Kooistra W, Sarno D, Gaonkar CC, Piredda R (2019) Global distribution and diversity of *Chaetoceros* (Bacillariophyta, Mediophyceae): integration of classical and novel strategies. PeerJ 7:e7410

Durbin EG (1977) Studies on the autecology of the marine diatom *Thalassiosira nordenskioeldii*. 2. The influence of cell size on growth rate, and carbon, nitrogen, chlorophyll a and silica content. J Phycol 13:150–155

Eissler Y, Anabalón J, Celis-Plá PSM, Kieft B, Junier P, Rivera P, Cruces F, Ascencio E, Dorador C, Molina V (2025) High nutrient requirement on the photosynthetic performance of a diatom species of the genus Nitzschia (Bacillariophyta) isolated from the high-altitude wetland, Salar de Huasco, Chile. Phycol Res. https://doi.org/10.1111/pre.12581. (**Advance online publication**)

Fu W, Shu Y, Yi Z, Su Y, Pan Y, Zhang F, Brynjolfsson S (2022) Diatom morphology and adaptation: current progress and potentials for sustainable development. Sustain Horiz 2:100015

Ge W, Wang F, Chai C (2012) Effect of nitrogen and phosphorus concentration on interspecific competition between *Skeletonema costatum* and *Scrippsiella trochoidea*. Chin J Plant Ecol 36:697–704

Godhe A, Asplund ME, Härnström K, Saravanan V, Tyagi A, Karunasagar I (2008) Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. Appl Environ Microbiol 74:7174–7182

Gong W, Browne J, Hall N, Schruth D, Paerl H, Marchetti A (2017) Molecular insights into a dinoflagellate bloom. ISME J 11:439–452

Gray J, Wu R, Or YY (2002) Effects of hypoxia and organic enrichment on the coastal marine environment. Mar Ecol Prog Ser 238:249–279

Gu H, Zeng N, Liu T, Yang W, Müller A, Krock B (2013) Morphology, toxicity, and phylogeny of Alexandrium (Dinophyceae) species along the coast of China. Harmful Algae 27:68–81

Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, Boutte C, Burgaud G, de Vargas C, Decelle J, del Campo J, Dolan JR, Dunthorn M, Edvardsen B, Holzmann M, Kooistra WHCF, Lara E, Le Bescot N, Logares R, Mahé F et al (2012) The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Res 41:D597–D604

Henson S, Le Moigne F, Giering S (2019) Drivers of carbon export efficiency in the global ocean. Glob Biogeochem Cycles 33:891–903

Hinder SL, Hays GC, Edwards M, Roberts EC, Walne AW, Gravenor MB (2012) Changes in marine dinoflagellate and diatom abundance under climate change. Nat Clim Change 2:271–275

Hoffmann LJ, Peeken I, Lochte K (2007) Co-limitation by iron, silicate, and light of three Southern Ocean diatom species. Biogeosci Discuss 2007:209–247

Holmes RM, Aminot A, Kérouel R, Hooker BA, Peterson BJ (1999) A simple and precise method for measuring ammonium in marine and freshwater ecosystems. Can J Fish Aquat Sci 56:1801–1808

Inzé D, De Veylder L (2006) Cell cycle regulation in plant development. Annu Rev Genet 40:77–105

Jeong HJ, Ok JH, Lim AS, Kwon JE, Kim SJ, Lee SY (2016) Mixotrophy in the phototrophic dinoflagellate *Takayama helix* (family Kareniaceae): predator of diverse toxic and harmful dinoflagellates. Harmful Algae 60:92–106

Jeong HJ, Yeong DY, Park JY, Jae YS, Seong TK, Lee S-H, Kim K, Yih W (2005) Feeding by phototrophic red-tide dinoflagellates: five species newly revealed and six species previously known to be mixotrophic. Aquat Microb Ecol 40:133–150

Jeong HJ, Yoo YD, Kim JS, Seong KA, Kang NS, Kim TH (2010) Growth, feeding and ecological roles of the mixotrophic and heterotrophic dinoflagellates in marine planktonic food webs. Ocean Sci J 45:65–91

Jiang Z, Liu J, Chen J, Chen Q, Yan X, Xuan J, Zeng J (2014) Responses of summer phytoplankton community to drastic environmental changes in the Changjiang (Yangtze River) estuary during the past 50 years. Water Res 54:1–11

John EH, Flynn KJ (2000) Growth dynamics and toxicity of *Alexandrium fundyense* (Dinophyceae): the effect of changing N[ratio] P supply ratios on internal toxin and nutrient levels. Eur J Phycol 35:11–23

Kužat N, Marić Pfannkuchen D, Smodlaka Tanković M, Baričević A, Ivančić I, Vrana I, Gašparović B, Pfannkuchen M (2022) Morpho-physiological adaptations of Leptocylindrus aporus and L. hargravesii to phosphate limitation in the northern Adriatic. Sci Rep 12:2687

Lambert BS, Groussman RD, Schatz MJ, Coesel SN, Durham BP, Alverson AJ, White AE, Armbrust EV (2022) The dynamic trophic architecture of open-ocean protist communities revealed through machine-guided metatranscriptomics. Proc Natl Acad Sci USA 119:e2100916119

Li F, Yue C, Deng Y, Tang YZ (2025) Full-length transcriptome analysis of a bloom-forming dinoflagellate *Scrippsiella acuminata* (Dinophyceae). Sci Data 12:352

Li Y, Chen Z, Chen J, Castellano MJ, Ye C, Zhang N, Miao Y, Zheng H, Li J, Ding W (2022) Oxygen availability regulates the quality of soil dissolved organic matter by mediating microbial metabolism and iron oxidation. Glob Change Biol 28:7410–7427

Liang Y, Koester JA, Liefer JD, Irwin AJ, Finkel ZV (2019) Molecular mechanisms of temperature acclimation and adaptation in marine diatoms. ISME J 13:2415–2425

Litchman E, de Tezanos Pinto P, Klausmeier CA, Thomas MK, Yoshiyama K (2010) Linking traits to species diversity and community structure in phytoplankton. Hydrobiologia 653:15–28

Liu H, Chen M, Zhu F, Harrison PJ (2016) Effect of diatom silica content on copepod grazing, growth and reproduction. Front Mar Sci 3:89

Liu K, Liu S, Cui Z, Zhao Y, Chen N (2024) Rich diversity and active spatial-temporal dynamics of Thalassiosira species revealed by time-series metabarcoding analysis. ISME Commun 4:ycad009

Ma X, Johnson KB, Gu B, Zhang H, Li G, Huang X, Xia X (2022) The in-situ release of algal bloom populations and the role of prokaryotic communities in their establishment and growth. Water Res 219:118565

Marić Pfannkuchen D, Godrijan J, Smodlaka Tanković M, Baričević A, Kužat N, Djakovac T, Pustijanac E, Jahn R, Pfannkuchen M (2018) The ecology of one cosmopolitan, one newly introduced and one occasionally advected species from the genus *Skeletonema* in a highly structured ecosystem, the northern adriatic. Microb Ecol 75:674–687

Martin-Jézéquel V, Hildebrand M, Brzezinski MA (2000) Silicon metabolism in diatoms: implications for growth. J Phycol 36:821–840

Martin JL, Santi I, Pitta P, John U, Gypens N (2022) Towards quantitative metabarcoding of eukaryotic plankton: an approach to improve 18S rRNA gene copy number bias. Metabarc Metagenom. https://doi.org/10.3897/mbmg.6.85794

Michael Beman J, Arrigo KR, Matson PA (2005) Agricultural runoff fuels large phytoplankton blooms in vulnerable areas of the ocean. Nature 434:211–214

Mitra A, Caron DA, Faure E, Flynn KJ, Leles SG, Hansen PJ, McManus GB, Not F, do Rosario Gomes H, Santoferrara LF, Stoecker DK, Tillmann U (2023) The Mixoplankton Database (MDB): diversity of photo-phago-trophic plankton in form, function, and distribution across the global ocean. J Eukaryot Microbiol 70:e12972

Mohd-Din M, Abdul-Wahab MF, Mohamad SE, Jamaluddin H, Shahir S, Ibrahim Z, Hii KS, Tan SN, Leaw CP, Gu H, Lim PT (2020) Prolonged high biomass diatom blooms induced formation of hypoxic-anoxic zones in the inner part of Johor Strait. Environ Sci Pollut Res 27:42948–42959

Moon J-Y, Lee K, Lim W-A, Lee E, Dai M, Choi Y-H, Han I-S, Shin K, Kim J-M, Chae J (2021) Anthropogenic nitrogen is changing the East China and Yellow Seas from being N deficient to being P deficient. Limnol Oceanogr 66:914–924

Murphy J, Riley JP (1962) A modified single solution method for the determination of phosphate in natural waters. Anal Chim Acta 27:31–36

Nohe A, Goffin A, Tyberghein L, Lagring R, De Cauwer K, Vyverman W, Sabbe K (2020) Marked changes in diatom and dinoflagellate biomass, composition and seasonality in the Belgian Part of the North Sea between the 1970s and 2000s. Sci Total Environ 716:136316

Ou L, Qin X, Shi X, Feng Q, Zhang S, Lu S, Qi Y (2020) Alkaline phosphatase activities and regulation in three harmful *Prorocentrum* species from the coastal waters of the East China Sea. Microb Ecol 79:459–471

Pang M, Liu K, Chen B, Zhang X, Gao Z, Xu Z, Tan Y, Yang J, Liu H (2024) Nutrient availability influences the thermal response of marine diatoms. Limnol Oceanogr 69:2318–2331

Pedersen EJ, Miller DL, Simpson GL, Ross N (2019) Hierarchical generalized additive models in ecology: an introduction with mgcv. PeerJ 7:e6876

Rajapaksha RP, Wu M, Wang Y, Bandara G, Atapaththu KSS, Wang Y (2024) Long-term alterations of nutrient dynamics and phytoplankton communities in Daya Bay, South China Sea. Mar Pollut Bull 208:116955

Rhee G-Y, Gotham IJ (1980) Optimum N:P ratios and coexistence of planktonic algae. J Phycol 16:486–489

Spilling K, Olli K, Lehtoranta J, Kremp A, Tedesco L, Tamelander T, Klais R, Peltonen H, Tamminen T (2018) Shifting diatom—dinoflagellate dominance during spring bloom in the Baltic Sea and its potential effects on biogeochemical cycling. Front Mar Sci 5:327

Stamieszkin K, Millette NC, Luo JY, Follett E, Record NR, Johns DG (2024) Large protistan mixotrophs in the North Atlantic Continuous Plankton Recorder time series: associated environmental conditions and trends. Front Mar Sci 11:1320046

Strickland JDH, Parsons TR (1972) A practical handbook of seawater analysis. Ottawa: fisheries research board of Canada. Int Rev Hydrobiol 55:167–167

Sun J-Z, Wang T, Huang R, Yi X, Zhang D, Beardall J, Hutchins DA, Liu X, Wang X, Deng Z, Li G, Gao G, Gao K (2022) Enhancement of diatom growth and phytoplankton productivity with reduced O2 availability is moderated by rising CO2. Commun Biol 5:54

Takabayashi M, Lew K, Johnson A, Marchi A, Dugdale R, Wilkerson FP (2006) The effect of nutrient availability and temperature on chain length of the diatom, *Skeletonema costatum*. J Plankton Res 28:831–840

Tian Y, Hu S, Lin X, Huang H, Song X, Yan Y, Xie X, Li T, Liu S (2021) Mechanisms of high-frequency dinoflagellate blooms of *Scrippsiella trochoidea* in Daya Bay, South China Sea. J Oceanol Limnol 39:1293–1304

Tréguer P, Bowler C, Moriceau B, Dutkiewicz S, Gehlen M, Aumont O, Bittner L, Dugdale R, Finkel Z, Iudicone D, Jahn O, Guidi L, Lasbleiz M, Leblanc K, Levy M, Pondaven P (2018) Influence of diatom diversity on the ocean biological carbon pump. Nat Geosci 11:27–37

Tyralis H, Papacharalampous G, Langousis A (2019) A brief review of random forests for water scientists and practitioners and their recent history in water resources. Water 11:910

Wallace RB, Gobler CJ (2021) The role of algal blooms and community respiration in controlling the temporal and spatial dynamics of hypoxia and acidification in eutrophic estuaries. Mar Pollut Bull 172:112908

Wang B, Chen J, Jin H, Li H, Huang D, Cai W-J (2017) Diatom bloom-derived bottom water hypoxia off the Changjiang estuary, with and without typhoon influence. Limnol Oceanogr 62:1552–1569

Wang J, Wu J (2009) Occurrence and potential risks of harmful algal blooms in the East China Sea. Sci Total Environ 407:4012–4021

Wang Q, Yang Q, Zhu L, Cui Z, Qu K, Wei Y (2024) Environmental controls on the seasonal variations of diatoms and dinoflagellates in the Qingdao coastal region, the Yellow Sea. Mar Environ Res 198:106524

Wang Z, Mu D, Li Y, Cao Y, Zhang Y (2011) Recent eutrophication and human disturbance in Daya Bay, the South China Sea: dinoflagellate cyst and geochemical evidence. Estuar Coast Shelf Sci 92:403–414

Wasmund N, Kownacka J, Göbel J, Jaanus A, Johansen M, Jurgensone I, Lehtinen S, Powilleit M (2017) The Diatom/Dinoflagellate index as an indicator of ecosystem changes in the Baltic Sea 1 principle and handling instruction. Front Mar Sci 4:22

Wei Y, Luan Q, Shan X, Cui H, Qu K, Cui Z, Sun J (2024) Temperature and nutrients drive distinct successions between diatoms and dinoflagellates over the past 40 years: implications for climate warming and eutrophication. Sci Total Environ 931:172997

Xia X, Cheung S, Zhang S, Lu Y, Leung SK, Shi Z, Xu H, Gu B, Tan Y, Zeng H, Li Y, Liu H (2024) *Noctiluca scintillans* bloom alters the composition and carbohydrate utilization of associated bacterial community and enriches potential pathogenic bacterium *Vibrio anguillarum*. Water Res 249:120974

Xia X, Ki Leung S, Cheung S, Zhang S, Liu H (2020) Rare bacteria in seawater are dominant in the bacterial assemblage associated with the bloom-forming dinoflagellate *Noctiluca scintillans*. Sci Total Environ 711:135107

Xiao W, Liu X, Irwin AJ, Laws EA, Wang L, Chen B, Zeng Y, Huang B (2018) Warming and eutrophication combine to restructure diatoms and dinoflagellates. Water Res 128:206–216

Xu B, Luo C-S, Liang J-R, Chen D-D, Zhuo W-H, Gao Y-H, Chen C-P, Song S-S (2014) Cellular metabolic responses of the marine diatom *Pseudo-nitzschia multiseries* associated with cell wall formation. Mar Genomics 16:29–38

Zhang X, Yu K, Li M, Jiang H, Gao W, Zhao J, Li K (2024) Diatom-dinoflagellate succession in the Bohai Sea: the role of N/P ratios and dissolved organic nitrogen components. Water Res 251:121150

Zhou M, Shen Z, Yu R (2008) Responses of a coastal phytoplankton community to increased nutrient input from the Changjiang (Yangtze) River. Cont Shelf Res 28:1483–1489

Zhou Y, Zhang Y, Li F, Tan L, Wang J (2016) Nutrients structure changes impact the competition and succession between diatom and dinoflagellate in the East China Sea. Sci Total Environ 574:499–508

Zhu F, Massana R, Not F, Marie D, Vaulot D (2005) Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. FEMS Microbiol Ecol 52:79–92