HOKKAIDO UNIVERSITY

Title	Photosynthetic carbon fixation capability of diatom communities in the southern Sea of Okhotsk from winter to spring as estimated using their rbcL genes
Author(s)	Sun, Yanxu; Nishioka, Jun; Toyota, Takenobu et al.
Citation	Journal of oceanography, 82, 135-154 https://doi.org/10.1007/s10872-025-00778-4
Issue Date	2025-10-07
Doc URL	https://hdl.handle.net/2115/98133
Rights	This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://dx.doi.org/10.1007/s10872-025-00778-4
Type	journal article
File Information	Sun et al (2025) Photosynthetic carbon fixation capability of diatom communities in the southern Sea of Okhotsk from winter to spring as estimated using their rbcL genes .pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Photosynthetic carbon fixation capability of diatom communities in the southern Sea of Okhotsk from winter to spring as estimated using their *rbcL* genes

Yanxu Sun[1*], Jun Nishioka[2], Takenobu Toyota[2], Koji Suzuki[3*]

[1] Graduate School of Environmental Science, Hokkaido University, Sapporo, Japan

[2] Pan-Okhotsk Research Center, Institute of Low Temperature Science, Hokkaido University, Sapporo, Japan

[3] Faculty of Environmental Earth Science, Hokkaido University, Sapporo, Japan

*** Corresponding authors:**

Yanxu Sun syx0620@ees.hokudai.ac.jp

Koji Suzuki kojis@ees.hokudai.ac.jp

1. Abstract

The transition from ice-covered to open waters is a recurring environmental feature in seasonal sea-ice regions. Yet the differences in the photosynthetic activity of phytoplankton communities during this period are poorly known. The southern Sea of Okhotsk (SSO), a sizeable seasonal ice zone in the northern hemisphere, experiences significant phytoplankton bloom dynamics, mainly composed of diatoms, driven by sea ice retreat in spring. Here, we compared the differences in the expression level of the *rbcL* gene, which encodes the large subunit of the photosynthetic carbon fixation enzyme RuBisCO, among dominant diatom families in the SSO from winter to spring in 2022 and compared the results with environmental factors. Higher diatoms' *rbcL* transcription efficiency in terms of mRNA (cDNA) /DNA copy number ratio was observed in winter-coastal and spring waters than in winter under-ice waters, suggesting light availability as a key driver of their photosynthetic activity. Additionally, centric diatoms exhibited significantly higher *rbcL* transcription efficiency than pennate diatoms throughout the study period. The transcription efficiency of three dominant diatom families-Chaetocerotaceae, Thalassiosiraceae, and Bacillariaceae showed strong positive correlations with both the maximum quantum yield of photochemistry in algal photosystem II (F_v/F_m) and chlorophyll *a* concentrations. This suggests that light-harvesting capacity and carbon fixation activity (mediated by *rbcL*-encoded RuBisCO) jointly contribute to spring bloom in the SSO. The *rbcL* cDNA copy number correlated strongly with fucoxanthin, a proxy for diatom biomass. These findings highlight the identification of photosynthetically active diatom groups contributing to high primary productivity in spring blooms and their sensitivity to environmental changes in seasonally ice-covered regions.

Keyword: Phytoplankton, Sea of Okhotsk, Photosynthetic activity, *rbcL* gene

2. Introduction

In recent decades, the retreat of sea ice in the Arctic Ocean has significantly expanded open-water habitats, creating favorable conditions for phytoplankton growth (Arrigo et al., 2008; Kahru et al., 2016). Between 1998 and 2012, the annual net primary production (NPP) across the Arctic Ocean open waters increased by 30% (Arrigo and van Dijken, 2015). Similarly, in sub-Arctic regions such as the Sea of Okhotsk, the extent of seasonal ice cover has significantly decreased over the past decades, with notable changes observed in the seasonal ice zone (SIZ) (Comiso and Nishio, 2008; Toyota et al., 2022). These changes are expected to profoundly impact marine ecosystems and biogeochemical processes.

Diatoms are among the most successful phytoplankton groups in modern oceans, contributing approximately 20% of the Earth's annual primary production (Field et al., 1998). In the Arctic and sub-Arctic regions, diatoms dominate the phytoplankton community in open waters during the ice-free period (Ardyna and Arrigo, 2020; Degerlund and Eilertsen, 2010; Mundo et al., 2023; Selz et al., 2018). Their dominance is attributed to their exceptional adaptability to polar environmental conditions, including their ability to survive within sea ice (Von Quillfeldt, 2001). Some species inhabit brine channels in sea ice and produce extracellular polymeric substances (EPS) to tolerate high salinity, low temperatures, and limited light (Arrigo, 2014). They can endure prolonged darkness during polar winters and rapidly resume growth under returning light (McMinn and Martin, 2013), supported by metabolic downregulation and preservation of key photosynthetic components (e.g., photosystem II, pigments, RuBisCO) (Kvernvik et al., 2018; Lacour et al., 2019).

Like other phytoplankton, photosynthetic carbon fixation in diatoms is mediated by the enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) via the Calvin-Benson-Bassham (CBB) reductive pentose phosphate pathway (Raven, 1998). RuBisCO consists of eight large subunits (LSU) and eight small subunits (SSU) and is the rate-limiting enzyme in the first step of photosynthetic carbon dioxide fixation in phytoplankton (Corrector et al., 2004; John et al., 2007b). Among them, the *rbcL* gene encoding the RuBisCO large subunit has been widely used as a functional gene to estimate the abundance and community composition of diatoms (e.g., John et al., 2007b; Johnson and Martiny, 2015; Kong et al., 2012).

The mRNA (cDNA) transcripts of the *rbcL* gene can indicate phytoplankton photophysiological responses to environmental changes (Endo et al., 2015). Changes in gene expression and protein synthesis levels of RuBisCO occur within a few hours in response to environmental changes such as light and nutrient availability (Granum et al., 2009; John et al., 2012). These dynamics provide direct insights into phytoplankton activity and their responses to physical and chemical environmental changes. For instance, in the Bering Sea and the Northwest Pacific, diatom-specific *rbcL* transcripts have been shown to decrease following a 2- to 3-day incubation under elevated CO_2 levels (Endo et al., 2015, 2016). Additionally, several studies have indicated that an increase in maximum photosynthetic rate (P_{max}) was accompanied by an increase in diatom-specific *rbcL* gene transcript abundance, demonstrating a significant correlation between these photosynthetic parameters (Corrector et al., 2004; Endo et al., 2017; John et al., 2007b). Furthermore, differences in RuBisCO expression among diatom species under varying environmental conditions have been observed. For example, during incubation experiments, Valegård et al. (2018) demonstrated variations in RuBisCO expression among five Arctic diatom species under different temperature and light conditions. However, in natural environments, the *rbcL* gene expression level in diatoms is likely to be simultaneously influenced by multiple environmental factors. Despite this, the differences in *rbcL* gene expression among diatom groups in natural Arctic and subarctic marine environments remain unclear.

The southern Sea of Okhotsk (SSO) is the southernmost seasonally ice-covered zone (SIZ) in the Northern Hemisphere, characterized by high primary productivity and abundant fisheries resources (Kasai and Hirakawa, 2016; Shimada et al., 2012; Mundo et al., 2023). During winter, the SSO is strongly influenced by sea ice transported from the northern continental shelf and the eastern Sakhalin Island coast via the East Sakhalin Current (ESC), as well as by prevailing northerly winds (Kuga et al., 2023). In spring, rising temperatures lead to sea ice melt, which triggers substantial phytoplankton blooms in surface waters, contributing significantly to the annual primary productivity of the SSO (Kasai et al., 2010; Ohwada, 1957). Previous studies have emphasized the dominance of diatoms in this region from winter to spring (Matsumoto et al., 2021; Yan et al., 2022), with large centric diatoms *Thalassiosira*

contributing most significantly to the spring bloom, followed by the small-sized *Chaetoceros* and pennate sea ice diatom *Fragilariopsis* (Sun et al., 2025). However, the molecular mechanisms underlying photosynthetic carbon assimilation by different diatoms remain unexplored.

In this study, using quantitative real-time PCR (qPCR) and next-generation sequencing (NGS) techniques, we quantified and sequenced the *rbcL* gene fragments in diatoms and their transcripts to investigate the differences in *rbcL* gene expression level under natural environmental conditions in the SSO. In addition, we validated the quantification of the *rbcL* gene for diatoms by integrating HPLC pigment signatures and 18S rRNA gene analysis. The expression levels of the *rbcL* gene in dominant diatom groups help us identify major primary producers (carbon fixers) in the SSO.

3. Materials and methods

3.1 Sample collection and environmental data analysis

The field sampling was conducted at 14 stations in the southern Sea of Okhotsk (SSO) during the transition from winter to spring in 2022. Among these, five stations were sampled aboard the icebreaker P/V Soya (Japan Coast Guard) from February 11 to 15, while nine stations were sampled aboard R/V *Shinsei Maru* (JAMSTEC) from May 1 to 12 (Table 1, Figure 1). Station 18 was sampled three times on May 1, 9, and 10 and is referred to as Stations 18, 18R, and 18TM, respectively (Figure 1).

To examine changes in chlorophyll *a* (Chl *a*) during the sampling period in the entire SSO, we utilized 8-day composite Chl *a* data from the GCOM-C/SGLI satellite sensor (Japan Aerospace Exploration Agency, JAXA), with a spatial resolution of 250 m. Image standardization and processing were conducted using SeaDAS (version 9.0.1, NASA; Figure 1).

Given that icebreaking operations might have influenced sea ice conditions at the sampling sites, we further analyzed sea ice concentrations during the sampling period using high-resolution sea ice concentration (HSI) data from GCOM-W/AMSR-2 (JAXA), which has a spatial resolution of 5 km. The HSI data accurately described the ice conditions on the sampling days at winter stations and were consistent with visual observations made during sampling (Table 1).

Seawater samples were collected in February using acid-cleaned, 5 L Teflon-coated Niskin-X bottles. Seawater temperature and salinity were measured using a CTD profiler (ASTD-102, JFE Advantec). In May, hydrographic data and seawater samples were obtained using a CTD-carousel system (SBE-9 plus, Sea-Bird Electronics) equipped with 12 L acid-cleaned Teflon-coated Niskin-X bottles for 5 m depth sampling or a clean plastic bucket for surface (0 m depth) sampling. The mixed layer depth is determined by $(\rho - \rho_m) \geq 0.15$ (kg m^{-3}), whichever is shallower, where ρ is density and ρ_m is the mixed layer density, taken at a depth of 5 m, following Ohshima et al. (2005). To assess the stability of the water column in the mixed layer, we used MLD stratification indices, "the Brunt-Väisälä frequency (N^2)", to estimate the vertical stability (i.e., stratification) of the mixed layer depth (MLD) (Table 1). This frequency was

calculated based on a 3-meter moving average of the density profile at the MLD, derived from CTD data (Schaeffer et al., 2012; Wahl and Teague, 1983).

$$N^2 = \frac{g}{\rho} \cdot \frac{\partial \rho}{\partial z}$$

Here, g represents the gravitational acceleration, ρ is the seawater density at each depth (kg m^{-3}), and z is the depth (m).

Macronutrients, including nitrate (NO_3^-), nitrite (NO_2^-), ammonium (NH_4^+), phosphate (PO_4^{3-}), and silicate ($Si(OH)_4$) were analyzed. Duplicate nutrient samples were collected in 10 mL clean acrylic tubes and immediately frozen at −20 °C for later analysis. Nutrient concentrations were determined in a laboratory using a continuous segmented flow auto-analyzer (QuAAtro-2, BLTEC), adhering to the methodology described by Nishioka et al. (2020).

3.2 Pigment analysis using ultra-high performance liquid chromatography (UHPLC)

Water samples (1.2 L) were filtered onto 25 mm Whatman GF/F glass-fiber filters (nominal 0.7 μm pore size) under low vacuum pressure (<0.013 MPa). The obtained filters were blotted with qualitative filter papers and stored in a deep freezer (−80°C) or liquid nitrogen until analysis on land. Phytoplankton pigments were extracted with *N*,*N*-dimethylformamide (DMF) using the bead-beating method and analyzed by ultra-high-performance liquid chromatography (UHPLC) (Suzuki et al., 2015). The concentrations of Chl *a* and a marker pigment for diatoms, that is, fucoxanthin (Fuco), were quantified and used in this study, following Suzuki et al. (2015).

3.3 PAM fluorometry

The maximum photochemical quantum efficiency of photosystem II (PSII), expressed as F_v/F_m, was assessed during each sampling period using a pulse-amplitude-modulated fluorometer (Water-PAM, Walz, Germany) after the dark acclimation for 30 min, following (Suzuki et al., 2014). The F_v/F_m ratios were determined using the formula $F_v/F_m = (F_m - F_o)/F_m$, where F_o represents the minimum fluorescence yield recorded after 30 min of the dark acclimation, and F_m denotes the maximum fluorescence yield obtained under the saturating pulse by the fluorometer.

3.4 qPCR assay for diatom-specific 18S rRNA gene abundance

The seawater of 200–300 mL was filtered onto Supor polyethersulfone (PES) membrane filters (pore size 0.2 μm, 25 mm, Pall Corporation) under low vacuum pressure (< 0.013 MPa) and then stored in a −80 °C freezer until analysis. The abundance of diatom-specific 18S rRNA gene fragments was measured by qPCR. The process and general primer design were based on the methods described in Endo et al. (2018).

3.5 qPCR and qRT-PCR assay for rbcL gene abundance

For DNA analysis, seawater samples (200–300 mL) were filtered onto Supor polyethersulfone (PES) membrane filters (pore size 0.2 μm, 25 mm, Pall Corporation) under low vacuum pressure (< 0.013 MPa). The filters were then stored at −80 °C until further analysis. DNA extraction and purification followed the protocol outlined by Endo et al. (2013). Extracted DNA pellets were resuspended in 100 μL of 10 mM Tris-HCl buffer (pH 8.5).

For RNA analysis, seawater samples of 200–300 mL were filtered onto Supor polyethersulfone (PES) membrane filters (0.2 μm pore size, 25 mm diameter; Pall Corporation) under low vacuum conditions (< 0.013 MPa). The filters obtained were soaked in 600 μL RNAlater solution (Thermo Fisher Scientific Inc.) in each cryovial and stored either in liquid nitrogen or a −80 °C freezer until RNA extraction. The detailed RNA extraction procedures were described in Endo et al. (2015). The extracted RNA was reverse transcribed into Complementary DNA (cDNA) using the PrimeScript RT Reagent Kit with gDNA Eraser (Takara).

A qPCR assay was performed with triplicate 10μL reactions containing 1 μL of genomic DNA (or cDNA) 5 μL TB Green Premix Ex Taq™ II (Takara, Japan), 0.4 μL of each forward and reverse primer, and 3.2 μL DNase-free Milli-Q water. The *rbcL* gene (DNA) and its transcripts (cDNA) were quantified using quantitative polymerase chain reaction (qPCR) and quantitative reverse-transcription PCR (qRT-PCR) on a CFX Connect Real-Time PCR Detection System (Bio-Rad, USA) with the following thermal cycling conditions: 95 °C for 1 min, followed by 39 cycles of 95 °C for 5 s and 52 °C for 60 s. After amplification, a melting curve analysis was performed by heating the samples to 65 °C for 5 s, followed by a stepwise increase to 95 °C in

0.5 °C increments, with fluorescence data collection at each step. The forward primer sequence was 5′-GATGATGARAAYATTAACTC-3′, and the reverse primer sequence was 5′-TAWGAACCTTTWACTTCWCC-3′, as reported by John et al. (2007a). Standard curves were generated using plasmid DNA (pUC-GW-Kan; Azenta Life Sciences, Germany) containing a synthetic 113-bp fragment of the *rbcL* gene derived from the diatom *Thalassiosira weissflogii* (CCMP1336). The plasmid was serially diluted in 10-fold steps to construct the standard curves, which were used to quantify the copy numbers of both *rbcL* genomic DNA and cDNA.

3.6 DNA and cDNA sequencing

Gene fragments of *rbcL* sequences were amplified from the extracted DNA or cDNA samples, using the same primers as in the qPCR analysis (John et al., 2007a) and the DNA polymerase KOD FX Neo (Toyobo) according to the manufacturer's protocol. Library construction and sequencing were carried out by Bioengineering Lab Co., Ltd. The libraries were sequenced on a MiSeq platform (Illumina) to generate paired-end reads.

Raw reads were processed with the fastx_barcode_splitter tool in the FASTX Toolkit (ver. 0.0.14, http://hannonlab.cshl.edu/fastx_toolkit/), retaining only those reads that completely matched the primer sequence at the beginning. The primer sequences were then removed from these reads using fastx_trimmer in the FASTX Toolkit. Next, quality filtering was performed using sickle (ver. 1.33, https://github.com/najoshi/sickle) to remove reads with a quality score below 20. Reads of 40 bases or shorter, along with their paired reads, were discarded. Subsequently, paired-end reads were merged using FLASH (ver. 1.2.11), with the minimum overlap set to 10 bases.

Following these steps, the resulting sequences were analyzed using QIIME2 (ver. 2022.8; Bolyen et al., 2019) in conjunction with the DADA2 plugin (Callahan et al., 2016). The command "dada2 denoise-pyro" was used to denoise, dereplicate, and remove chimeric reads from the quality-controlled sequences, resulting in 196 amplicon sequence variants (ASVs).

For taxonomic assignment, each of the 196 ASVs that accounted for ≥ 0.2% of the total reads was compared to *rbcL* sequences in the GenBank database (http://www.ncbi.nlm.nih.gov) using the BLASTN search engine. All other parameters were set to their default values. Each ASV

that has > 95% sequence similarity with known *rbcL* sequences was classified into the groups:Chaetocerotaceae,Thalassiosiraceae,Odontellaceae,Bacillariaceae,Attheyaceae,Naviculaceae,Amphipleuraceae,Rhizosoleniaceae,Coscinodiscaceae,Plagiogrammaceae,Stephanodiscaceae,Pleurosigmataceae,Rhizosoleniaceae,Ulnariaceae,Fragilariaceae,unidentified diatoms, and other eukaryotes. The ASVs related to two or more diatom families with the same sequence similarity were assigned as unidentified diatoms. In total, we identified 60 identified diatom ASVs and 32 unidentified diatom ASVs (Figure S1). Since the *rbcL* gene fragment used for sequencing was relatively short, we followed previous studies (e.g. Endo et al., 2016, 2015) and grouped the identified diatom ASVs at the family level for subsequent analyses.

In this study, the same set of primers was used for both qPCR quantification and NGS sequencing from John et al., (2007a). To quantify the *rbcL* gene abundance of diatoms from all phytoplankton groups, the absolute abundance of the diatom-specific *rbcL* gene was calculated. This was achieved by multiplying the *rbcL* gene copy number (from both DNA and cDNA) by the relative contribution of the 92 diatom-derived ASVs to the total of 196 ASVs.

3.7 Statistical analysis

To investigate differences in *rbcL* gene transcription efficiency at the ASV level in diatoms, we multiplied the copy numbers of diatom DNA and cDNA by the relative abundance of each ASV in the respective DNA and cDNA sequences. This approach allowed us to estimate an 'absolute abundance' for each ASV at both DNA and cDNA levels. We then used the ratio of cDNA to DNA absolute abundance as an indicator of *rbcL* gene expression level (hereafter, transcription efficiency) for each diatom ASV.

The absolute abundance of top 25 ASV (which means the highest read counts) in cDNA and DNA at each station were visualized as heatmaps using the R packages 'pheatmap' (Kolde, 2022). In this heatmap, the absolute abundance of each ASV in both cDNA and DNA was separately normalized by z-scores method and visualized as circles, with the size and color of the circles indicating the standardized absolute abundance.

To investigate differences in *rbcL* gene expression level among various diatom taxa (e.g., centric vs. pennate diatoms) or among different families, the identified ASVs were grouped

according to their taxonomic classification. Then, the results were visualized as box plots using the R packages ‘ggplot2’. Statistical significance was assessed using the non-parametric Mann–Whitney U test, with $p < 0.05$ indicating a significant difference. To investigate the relationships between *rbcL* gene expression level at the ASV level for the major spring-blooming diatom family (Chaetocerotaceae, Thalassiosiraceae, and Bacillariaceae) and various environmental factors, as well as the interrelationships among those environmental factors, we performed Mantel tests and constructed a correlation heatmap using the R package ‘ggcor’. In addition, to investigate how these three phytoplankton genera respond to environmental factors, we employed a linear regression model and used the R package ‘ggplot2’ for visualization. Redundancy analysis (RDA) was performed using the R package 'vegan' to analyze relationships between transcriptional efficiency of *rbcL* and environmental factors. The RDA results were tested by a non-parametric permutation test using the package 'nptest'.

4. Results

4.1 hydrography

We classified the sampling stations into two groups based on the sampling time: winter stations (February) and spring stations (May). As summarized in Table 1, the winter stations were characterized by lower water temperatures, low Chl *a* concentrations, and high concentrations of macronutrients. Among these, stations 10 and 14, located on or near the slope of the Kuril Basin (Figure 1), exhibited distinct characteristics of being almost entirely ice-covered (sea ice concentration > 80%). These stations, referred to as Under-ice stations, displayed lower average surface temperatures (-1.74±0.01°C), higher average salinity (32.91±0.07), and high concentrations of macronutrients such as NOx and $Si(OH)_4$. In contrast, stations 17w, 21, and 4 were situated closer to the Hokkaido coastline, where sea ice was either absent (station 21 and 17w) or minimal (station 4). These stations, referred to as Coastal stations, showed lower average salinity (31.92±0.19), higher NH_4 concentrations, and slightly higher average Chl *a* concentrations compared to under-ice stations. For all winter stations, F_v/F_m values were below 0.3, with under-ice stations even lower, below 0.2.

For the spring stations, seawater samples collected between May 1–8 at stations 18, 17s, 204, Bio1, and Bio3 exhibited Chl *a* concentrations exceeding 1 mg m^{-3}. Among these, stations 18 and 17s had particularly high Chl *a* concentrations, exceeding 3 mg m^{-3}, defined as bloom stations. The bloom stations were characterized by significantly reduced macronutrient concentrations and markedly high fucoxanthin and F_v/F_m levels. However, seawater samples collected after May 9 from stations 18 (Station 18R and TM) showed rapid replenishment of macronutrients; however, Chl *a* concentrations had already declined to below 1 mg m^{-3}. Station 11, the furthest from the coast, retained winter-like characteristics with high macronutrient concentrations (Table 1).

Apart from stations 18 and 17s, all other spring stations were classified as non-bloom stations, exhibited Chl *a* concentrations < 1.5 mg m^{-3}.

4.2 Copy numbers of rbcL DNA and cDNA of diatoms

Across all 14 samples, Fucoxanthin concentration had a significant positive correlation with

Chl *a* concentration (R^2=0.96, $p < 0.001$, n =14) (Figure 2a). We also observed a highly significant positive correlation between fucoxanthin concentration and the cDNA copy number of the diatom *rbcL* gene ($R^2 = 0.82$, $p < 0.001$, n =14) (Figure 2c). While the fucoxanthin concentration also showed a significant correlation with the DNA copy number of the *rbcL* gene ($R^2 = 0.38$, $p < 0.05$, n =14), but this correlation was weaker compared to that with cDNA (Figure 2c). Additionally, the DNA copy number of the diatom *rbcL* gene exhibited a stronger and highly significant positive correlation with the copy number of the 18S rRNA gene ($R^2 = 0.93$, $p < 0.001$, n =14) (Figure 2b).

4.3 Phytoplankton composition and rbcL gene transcription efficiency analyses

For both DNA and cDNA sequencing, diatom-derived ASVs accounted for 83 ± 20% of the total ASVs across all stations (Figure 3). At the winter stations, the relative contribution of diatom ASVs were 80 ± 6%, whereas for the spring stations excluding Station 11, diatom ASVs contributed 95 ± 5% of the total ASVs (Figure 3).

In this study, the dominant diatom families based on total DNA reads were Thalassiosiraceae (31% ± 19%), Bacillariaceae (16% ± 9%), and Chaetocerotaceae (15% ± 15%). Thalassiosiraceae showed the highest contribution in winter stations (39 ±5%) and spring bloom stations (62%± 0%). In contrast, Chaetocerotaceae showed higher contributions at non-bloom stations during spring (24% ± 17%) compared to Thalassiosiraceae (17% ± 13%). The pennate diatom family Bacillariaceae contributed notably at under-ice stations in winter (28% ±14%) and at non-bloom stations in spring (16% ±16%) but accounted for less than 8% at bloom stations.

The cDNA sequence result revealed that within the Diatom (Bacillariophyta), the dominant groups shifted to Chaetocerotaceae (36% ± 32%), Thalassiosiraceae (23% ± 19%), and Bacillariaceae (9% ± 15%) in that order. Among these, Chaetocerotaceae exhibited a particularly high contribution across all spring stations (54% ± 26%), whereas the contribution of Thalassiosiraceae decreased substantially compared to its DNA-based contribution, retaining only 14% ± 17% at spring stations. Bacillariaceae made a substantial contribution in winter (26% ± 18%), while at other stations, its contribution remained close to or below 10%.

For *rbcL* gene transcription efficiency (cDNA/DNA), which includes diatoms and other heterokont, the average *rbcL* gene transcription efficiency at winter coastal stations was 3.88 ± 0.53, while at under-ice stations, it was only 0.29 ± 0.37. During the spring bloom, *rbcL* gene transcription efficiency exceeded 10 at all bloom stations, with an average of 25.43 ± 18.79. In contrast, non-bloom stations exhibited values between 4 and 10, with an average of 6.79 ± 1.90.

4.4 Difference in ASV-based diatom groups rbcL gene transcription efficiency

To investigate the differences in *rbcL* gene transcription efficiency among diatom groups and families, we used absolute abundance to represent the *rbcL* copy number of each ASV. The absolute abundance of an ASV was calculated by multiplying the diatom DNA and cDNA copy numbers by their relative contributions in DNA and cDNA, respectively. Figure 4 presents a heatmap of the absolute abundance of the top 25 diatom ASVs. Among these, Chaetocerotaceae-derived ASVs comprised five ASVs, with ASV_001 being the most abundant. Their absolute DNA abundance was noticeably lower than their cDNA abundance at spring stations. Thalassiosiraceae-derived ASVs included nine ASVs, with ASV_002 as the most dominant. For this family, the DNA absolute abundance consistently exceeded the cDNA abundance across all stations. Representative pennate diatoms from Bacillariaceae included four ASVs, which were primarily found at winter and non-bloom spring stations in both DNA and cDNA datasets, with higher absolute abundance observed in DNA.

To further investigate the *rbcL* gene transcription efficiency of different diatom types, we categorized 60 diatom ASVs (excluding unidentified ASVs) into centric diatoms and pennate diatoms and calculated their *rbcL* gene transcription efficiency across all stations. The results showed that the normalized *rbcL* gene transcription efficiency ($\log_{10}$ cDNA/DNA) of centric diatom ASVs was significantly higher than that of pennate diatoms (Mann–Whitney U test, $p < 0.001$), with a median difference exceeding 220% (Figure 5a). From the perspective of station differences, no significant variation in normalized *rbcL* gene transcription efficiency was observed between centric and pennate diatoms at winter stations or spring bloom stations. However, a significant difference was detected at spring non-bloom stations (Mann–Whitney U test, $p < 0.001$) (Figure 5b).

For the three representative diatom families, the normalized *rbcL* gene transcription efficiency of ASVs from the centric diatoms Chaetocerotaceae and Thalassiosiraceae was significantly higher than that of the pennate diatom Bacillariaceae (Mann–Whitney U test, $p < 0.05$) (Figure 6a). As the most prominent centric diatoms during spring, Chaetocerotaceae and Thalassiosiraceae exhibited no significant differences in normalized *rbcL* gene transcription efficiency in winter. However, during spring, at both bloom and non-bloom stations, the normalized *rbcL* gene transcription efficiency of Chaetocerotaceae was significantly higher than that of Thalassiosiraceae (Mann–Whitney U test, $p < 0.01$) (Figure 6b).

4.5 Differences in rbcL gene transcription efficiency among major diatom families in relation to environmental factors

A correlation heatmap was used to analyze the relationships among environmental variables, and a Mantel test was conducted to examine the relationships between the average *rbcL* gene transcription efficiency of ASVs from three major diatom families and environmental variables (Figure 7). The results showed a strong positive correlation between Chl *a*, Fuco concentrations, and F_v/F_m, as well as among macronutrients (NO_3, PO_4, and SiO_2). The Mantel test indicated that the *rbcL* gene transcription efficiency of three major diatom families was all correlated with NO_3, PO_4 and F_v/F_m. Additionally, the *rbcL* gene transcription efficiency of Chaetocerotaceae correlated with water temperature, Fuco, and Chl *a*, while two centric diatom families showed correlations with SiO_2. All Mantel's correlation coefficients were below 0.5.

We used a linear regression model to further investigate potential linear relationships between the *rbcL* gene transcription efficiency of the three diatom families and environmental variables. The results showed a significant positive linear relationship between the normalized *rbcL* gene transcription efficiency of Chaetocerotaceae and water temperature (Figure 8a), as well as significant negative linear relationships with the three macronutrients (Figure 8d, e and f). For Bacillariaceae, a significant negative linear relationship was observed between its normalized *rbcL* gene transcription efficiency and salinity (Figure 8b). The normalized *rbcL* gene transcription efficiency of the three dominant diatom families all showed a strong positive correlation with F_v/F_m (Figure 8c). Additionally, significant positive correlations were observed between the *rbcL* gene transcription efficiency of Chaetocerotaceae, Thalassiosiraceae, and

Bacillariaceae and Chl *a* concentration, with R^2 values of 0.70, 0.53, and 0.51, respectively (Figure 8g).

5. Discussion

5.1 Classification of Sampling Stations Based on the hydrographic conditions and their rbcL gene expression level

In this study, sampling stations were categorized into four groups based on sampling time and hydrographic characteristics. Winter stations were divided into non-sea ice stations and under-ice stations, determined primarily by whether the station was covered by high concentrations of sea ice (ref. Table 1). Coastal stations were either fully dominated by the East Sakhalin Current (ESC; e.g., stations 21 and 4) or strongly influenced by the ESC (e.g., station 17w). The mixed layer depth (MLD) at these stations did not exceed 15 m (Table 1). In contrast, basin stations were dominated by the Okhotsk Sea Intermediate Water (OSIW), formed through density-driven mixing processes involving nutrient-rich Dense Shelf Water (DSW), the East Kamchatka Current (or Western Subarctic Water), and d-SWC (Itoh et al., 2003; Mundo et al., 2023). The MLD at these basin stations was approximately 30 m (Table 1). All winter station Brunt-Väisälä frequency (N^2) are extremely low (almost all around 10^{-5} s^{-2}), indicating that the vertical water column has undergone thorough mixing and is in a relatively uniform state.

All winter stations in this study were characterized by high concentrations of macronutrients critical for diatom growth. Dissolved inorganic nitrogen (DIN, including NO_2, NO_3 and NH_4), which is essential for diatoms (Yan et al., 2020), was about 20 μmol L^{-1} at under-ice stations and 10 μmol L^{-1} at coastal stations. Another essential macronutrient, SiO_2, necessary for the construction of diatom frustules, was consistently above 20 μmol L^{-1} at all winter stations, which may support diatom blooms in the surface layer after sea ice melt. Interestingly, at coastal stations with little or no ice cover, *rbcL* transcription efficiency (cDNA/DNA) was around 5 (Figure 3c), which is comparable to the ratio (~3.5) observed in a previous sea ice melt simulation experiment in the SSO (Yan et al., 2020). In contrast, under-ice stations showed almost zero transcription efficiency of *rbcL* gene, and the F_v/F_m values dropped below 0.2. This indicates that the light reactions of photosynthesis were severely limited under the ice, which may affect the functioning of the dark reactions, light availability is the most critical environmental factor limiting carbon fixation in winter.

In spring, we defined bloom stations as those with Chl *a* concentrations exceeding 3 mg m^{-3}

(e.g., stations 18 and 17s) and non-bloom stations as those with Chl *a* concentrations below 1.5 mg m^{-3}. Based on satellite ocean color remote sensing study, the spring bloom in the SSO during 2022 began in early April in the ice-melting coastal water. The bloom then expanded across the continental shelf, reaching its maximum extent by mid-April. In contrast, at winter under-ice stations located on the Kuril basin slope (e.g., stations 14 and 10) and some spring stations (e.g., stations 18, 17, Bio3, and NS5), the bloom occurred later, starting in mid-April, peaking in early May, and declining rapidly thereafter. This trend is consistent with our three consecutive samplings at station 18, where Chl *a* concentrations rapidly decreased from >3 mg m^{-3} on May 1 (St. 18) to <0.7 mg m^{-3} on May 9 (St. 18R) and May 10 (St. 18TM). Due to the variability of the spring bloom, we opted to define bloom and non-bloom stations based on Chl *a* concentration. The non-bloom stations (Chl *a* concentrations < 1mg m^{-3}) represented the pre-bloom and post-bloom states in spring. All spring stations were characterized by Sea of Okhotsk Surface Water (OSW) in the surface layer and Okhotsk Sea Intermediate Water (OSIW) at a depth of 100 m, with a mixed layer depth (MLD) of less than 50 m (Table 1). The Brunt-Väisälä frequency (N^2) at spring stations is higher than in winter, but still close to 0, indicating that vertical stratification is not yet significant. Among them, the spring bloom stations showed higher N^2 levels (Table 1), suggesting stronger vertical stratification compared to stations. It may provide a stable environment for the rapid proliferation of phytoplankton in the surface waters during this period.

At spring bloom stations, 18 and 17s, the concentration of the diatom marker pigment fucoxanthin was approximately 2 mg m^{-3}, more than four times the average value observed at non-bloom stations (Table 1). The transcription efficiency of *rbcL* gene at spring bloom stations exceeded 10, with station 17s reaching a value of 37 (Figure 3c). This ratio was much higher than that observed during the initial ice melting period in the simulation experiment (Yan et al., 2020). Such high transcription efficiency indicates exceptionally high photosynthetic activity during this stage, reflecting rapid proliferation and growth of phytoplankton. Meanwhile, we observed that the F_v/F_m values never exceeded 0.6, which is below the 0.65-0.8 range typically seen in phytoplankton under optimal conditions (Samson et al., 1999; Suggett et al., 2004). This could suggest that even in the spring bloom stations, the

light reactions of photosynthesis remain limited by other environmental stresses.

When comparing stations 18 and 17s, both spring bloom stations were sampled around the same time (Table 1), and station 17s exhibited higher transcription efficiency than station 18 (Figure 3c). At station 18, the NO_2+NO_3 concentration had already dropped to 0.21 μmol L^{-1}, with an N:P ratio of only 1.45, lower than the N:P ratio of 2.6 at station 17s (Table 1). Inorganic nitrogen is essential for the synthesis of proteins and nucleic acids, as well as for the formation of critical photosynthetic components—including the major light-harvesting pigment Chl *a*, protein-rich electron transport complexes, and key enzymes such as RuBisCO—that collectively drive photosynthetic electron transport and carbon fixation (Hughes et al., 2018; Rhiel et al., 1986). Nitrogen deficiency reduces the efficiency of light-harvesting in photosynthesis, meaning that nitrogen availability directly limits photosynthetic activity at the cellular physiological level (Kolber et al., 1998). Therefore, the bloom at station 18 is nearing its end due to low nitrogen availability, as evidenced by the sharp decline in Chl *a* concentration to below 1 mg m^{-3} a week later.

At non-bloom stations, *rbcL* transcription efficiency ranged between 4 and 10, slightly higher than that observed at winter coastal stations. Meanwhile, macronutrients consumed during the bloom phase were beginning to replenish at these stations (e.g. station 18, Table 1).

Station 11, located in the Kuril Basin, was the only station during the sampling period where diatoms were not the dominant group (diatoms contributed only ~20% of DNA sequences and ~10% of cDNA sequences) (Figure 3a and b). According to previous studies, diatoms are dominant most of the time, but dinoflagellates, cryptophytes and haptophytes can occasionally make considerable contributions (Hamao et al., 2022; Sildever et al., 2019; Sun et al., 2025; Yan et al., 2020). Since many dinoflagellates and cryptophytes are mixotrophic or heterotrophic, they either lack the key photosynthetic enzyme RuBisCO (in heterotrophic groups) or possess only the Form II RuBisCO, which is more favorable for photorespiration (in mixotrophic groups) (Glibert et al., 2016; Morse et al., 1995). Therefore, at this station, the *rbcL* gene exclusively represents photosynthetic phytoplankton groups. The *rbcL* transcription efficiency of these phytoplankton was relatively high (>9, Figure 3c), but the Chl *a* concentration was only 0.11 mg m^{-3} (Table 1). Only at diatom-dominated stations did high carbon fixation activity

translate into substantial Chl *a* accumulation, underscoring the pivotal role of diatoms in sustaining seasonal productivity through efficient carbon-to-biomass conversion from winter to spring.

5.2 Transcription Efficiency of the rbcL Gene in Different Diatom Groups

Diatoms are one of the dominant phytoplankton and ice algae groups in the Arctic and sub-Arctic seas, as summarized in Table 2, which highlights the main ice algae and bloom phytoplankton groups. Pennate diatom Bacillariaceae, represented by *Fragilariopsis* and *Nitzschia*, typically dominate ice algal communities (Poulin et al., 2011; Szymanski and Gradinger, 2016). In contrast, abundant species of *Thalassiosira* found in the Okhotsk Sea ice have also been reported in Arctic sea ice, although their dominance is generally much lower than that of pennate diatoms (Hegseth, 1992; Okolodkov, 1992). Notably, the Sea of Okhotsk appears to be unique in that centric diatoms Thalassiosiraceae, such as *Thalassiosira*, frequently occur in high abundance in the sea ice (Ohwada, 1956; McMinn et al., 2008; Yan et al., 2022). In the open waters of the Arctic and sub-Arctic regions following ice melt, sporadic blooms of haptophyte *Phaeocystis pouchetii*, chlorophyte *Micromonas*, or dinoflagellates are occasionally reported. However, diatoms consistently play a central role, with families such as Thalassiosiraceae, Chaetocerotaceae, and Bacillariaceae dominating in most areas (ref. Table 2).

In this study, we focused on the transcription efficiency of the *rbcL* gene in different diatom genera before and after the winter and spring bloom periods in the SSO, with a particular emphasis on dominant diatom families. Overall, during the transition from winter to spring in the SSO, the transcription efficiency of the *rbcL* gene was higher in centric diatoms than in pennate diatoms (Figure 5a). This difference was not significant in winter but became evident at non-bloom stations in spring (Figure 5b). When representative families of centric and pennate diatoms were compared, the transcription efficiency in small centric diatoms such as Chaetocerotaceae and large centric diatoms such as Thalassiosiraceae was significantly higher than in pennate diatoms such as Bacillariaceae (Figure 6a).

An incubation experiment with sea ice from the SSO reported that pennate diatoms, although

initially abundant, gradually became negligible during incubation due to their lower photosynthetic efficiency under osmotic stress compared to centric diatoms, which led to reduced energy availability, cell death, and the rapid sedimentation of some pennate ice diatoms (Yan et al., 2020). These findings are consistent with our result, where only the pennate diatom Bacillariaceae showed a negative correlation between its *rbcL* transcription efficiency and salinity ($R^2 = 0.34$, $p < 0.05$, $n = 14$). As the most abundant diatoms in sea ice, many pennate ice diatoms are characterized by their rapid sedimentation during the ice melting period (Selz et al., 2018; Sourisseau et al., 2024), resulting in shorter residence times in the surface layer and hindering their contribution to phytoplankton bloom. Some pennate diatoms show a preference for benthic habitats (Gradinger and Ikävalko, 1998; Nakov et al., 2015), and there is evidence suggesting that diatom sinking is directly influenced by metabolic processes, which are closely linked to cell size and energy availability, and thus to photosynthesis and respiration (Gemmell et al., 2016; Waite et al., 1992). Yan et al. (2020) demonstrated that in the SSO, pennate diatoms exhibit high mortality and sedimentation rates during ice melt due to their limited adaptability to salinity changes. In this study, we found that pennate diatoms suspended in surface seawater exhibited significantly lower transcription efficiency of *rbcL* gene compared to centric diatoms (Figure 5). However, the dominant pennate diatom family Bacillariaceae showed a significant positive correlation with Chl *a* concentration like centric diatoms (Figure 8g, $R^2 = 0.51$, $p < 0.01$). This indicates that despite the general pattern of lower carbon fixation activity (mediated by *rbcL*-encoded RuBisCO) among pennate diatoms, Bacillariaceae represent an important exception that actively contributes to phytoplankton blooms in open waters.

In this study, Chaetocerotaceae and Thalassiosiraceae, as representative centric diatoms, exhibited a strong positive correlation between their *rbcL* gene transcription efficiency and Chl *a* concentrations (Figure 8g), highlighting the importance of their high photosynthetic activity in driving higher Chl *a* concentrations during spring blooms. Chaetocerotaceae has been reported as a dominant diatom group throughout the Arctic during blooms (Booth et al., 2002; Freyria et al., 2021; Sergeeva et al., 2010) and is known to persist even after nutrient depletion during bloom periods (Schiffrine et al., 2020). In contrast, Thalassiosiraceae, which represented

by larger diatoms- *Thalassiosira*, typically appear at the onset of spring blooms in Arctic and sub-Arctic waters but decline in abundance as nutrient concentrations begin to decrease (Freyria et al., 2021; Lovejoy et al., 2002; Tremblay et al., 2006). Comparing the transcription efficiency of Chaetocerotaceae and Thalassiosiraceae, no significant differences were observed in winter; however, during spring, Chaetocerotaceae consistently showed significantly higher transcription efficiency than Thalassiosiraceae, particularly at non-bloom stations (Figure 6b). The transcription efficiency of the *rbcL* gene may reflect differences in their photosynthetic activity, a trend consistent with their abundance changes. In this study, we found a significant positive correlation between the cDNA copy number of Chaetocerotaceae and the diatom pigment fucoxanthin (Figure 2d), which confirms its importance to spring bloom. Mantel test results revealed that the *rbcL* gene transcription efficiency of both Chaetocerotaceae and Thalassiosiraceae correlates with macronutrients and F_v/F_m (Figure 7). The linear model further indicates a significant positive correlation between their transcription efficiency and F_v/F_m (Figure 8c), suggesting a coupling between the photosynthetic capacity and carbon fixation process in these two diatom families. In optimal environmental conditions, they may upregulate the expression of the *rbcL* gene to enhance their photosynthetic and carbon fixation abilities. Meanwhile, Chaetocerotaceae' transcription efficiency has a significant positive correlation with water temperature but a significant negative correlation with macronutrients (Figure 8) suggesting a stronger response of *Chaetoceros* to rising temperatures. Notably, differences in diatom RuBisCO catalytic properties significantly influence photosynthetic resource allocation, balancing between RuBisCO content and carbon concentrating mechanisms (CCMs) (Young et al., 2012). Specifically, Thalassiosiraceae maintains lower RuBisCO content but requires more resource investment in CCMs to saturate their low-CO_2-affinity (high Kc) RuBisCO. Conversely, diatoms like Chaetocerotaceae have higher RuBisCO content and lower Kc values, enabling carboxylation saturation at lower CO_2 concentrations (Young et al., 2016). Therefore, the observed differences in *rbcL* gene expression between these two families during open-water blooms in this study require further analysis of their physiological mechanisms in the future.

In this study, 32 diatom ASVs could not be identified to the species level and were classified as 'Unidentified diatoms', with ASV_003 serving as a representative (ref. Figure 3a and 4).

These unidentified diatom ASVs contributed an average of 15% ± 10% to the diatom ASV reads across all sites. This limitation may arise from both the relatively short *rbcL* primers designed by John et al. (2007a), which may reduce sequencing resolution, and the limited coverage of reference sequences in existing databases, hindering sequence annotation. To enhance identification accuracy, future research could consider employing longer *rbcL* primers or incorporating additional genetic markers to improve the recognition of diatom ASVs.

5.3 Estimating the Absolute Abundance of Diatoms Based on Multiple Methods

In diatom research, the absolute abundance is typically quantified using three main approaches. 1) morphological methods involve counting diatom cells under a microscope (or using flow cytometry) and evaluated diatom carbon biomass based on biovolume (e.g.Hillebrand et al., 1999; Sun and Liu, 2003). 2) HPLC pigment analysis uses marker pigments, such as fucoxanthin, as proxies for diatom abundance (e.g.Jeffrey et al., 2012; Kramer and Siegel, 2019). 3) molecular biology methods amplify specific genetic markers using diatom-specific primers, such as the commonly targeted V4 and V9 regions of the 18S rRNA gene (e.g.Vasselon et al., 2017; Zimmermann et al., 2011) or the *rbcL* gene (e.g.John et al., 2007). The amplified copy number is then used as an indicator of diatom abundance. In our previous research conducted at the same stations, we observed a stronger correlation between fucoxanthin concentration and diatom carbon biomass, while both metrics showed weaker correlations with the copy number of the diatom-specific 18S rRNA gene (rDNA) (Sun et al., 2025).

In this study, we found a very strong correlation between the DNA copy numbers of the *rbcL* gene and 18S rDNA ($R^2 = 0.93$, $p < 0.001$, $n = 14$) (Figure 2b), but a weaker correlation between *rbcL* gene DNA copy numbers and fucoxanthin concentration ($R^2 = 0.38$, $p < 0.05$, $n = 14$) (Figure 2c). Decoupling between morphological and molecular biology methods in diatom quantification is commonly observed and is generally attributed to the limitations of individual methods (e.g. Abad et al., 2016; Xiao et al., 2014). For instance, microscopy may overlook small or broken diatom cells (Naik et al., 2011), while fucoxanthin, although a marker pigment for diatoms, is also found in some dinoflagellates and haptophytes, potentially leading to overestimation (Higgins et al., 2012; Mackey et al., 1996). On the other hand, for molecular biology methods, the reliability of outputs depends on factors such as primer pairs for PCR

amplification (e.g. Alberdi et al., 2018; Piñol et al., 2019) and differences in cellular copy number of target genes among phytoplankton taxa, which can affect their estimates significantly (Kembel et al., 2012). Interestingly, in this study, fucoxanthin concentration showed a much stronger correlation with the cDNA copy numbers of the diatom *rbcL* gene ($R^2 = 0.82$, $p < 0.001$, $n = 14$) compared to DNA copy numbers (Figure 2c). This finding challenges the conventional approach of using *rbcL* DNA copy number as a proxy for quantifying diatom abundance, which often overlooks the importance of mRNA levels. As unicellular organisms, the *rbcL* gene DNA copy number in diatoms is primarily determined by the number of chloroplasts per cell and the number of genomes per chloroplast, rather than cell size (Ersland et al., 1981; Treusch et al., 2012). Vasselon et al. (2018) noted a positive correlation between *rbcL* gene DNA copy number and cell biovolume; however, this relationship exhibited considerable variability, necessitating the application of correction factors (CF) for accurate quantification. Generally, larger diatom cells require increased carbon fixation to support their growth and reproduction (Wirtz, 2011). Consequently, cDNA (mRNA) copy number, reflecting transcriptional activity, may more accurately represent the real-time abundance of diatoms in the environment. Our findings suggest that cDNA copy number, indicative of transcriptional activity, may more effectively represent diatom functional dynamics and align more closely with other photosynthesis-related biomass indicators. However, due to the relatively limited sample size in this study, this conclusion should currently be regarded as a preliminary hypothesis. Therefore, future research should consider adopting cDNA copy numbers as a quantitative indicator of diatom abundance and conduct larger-scale validation studies. Such efforts would enable closer integration of molecular methods with other analytical approaches, ultimately leading to more accurate biomass assessments.

6. Conclusion

This study investigated the transcriptional dynamics of the *rbcL* gene in diatoms during the transition from winter to spring in the southern Sea of Okhotsk (SSO), focusing on differences among diatom taxa and their relationship to environmental factors. Our findings highlight the critical role of diatoms in driving primary production in this sub-Arctic region, particularly during spring blooms. The *rbcL* transcription efficiency of diatoms varied significantly across seasons and diatom families. In winter, under-ice seawater exhibited negligible *rbcL* transcription efficiency and low F_v/F_m, suggesting that light availability was the primary limiting factor for photosynthetic activity. In contrast, spring bloom stations with high Chl *a* concentrations showed extremely high transcription efficiency and F_v/F_m.

Among diatom taxa, centric diatoms, particularly Chaetocerotaceae and Thalassiosiraceae, displayed higher *rbcL* transcription efficiency than pennate diatoms, such as Bacillariaceae. Chaetocerotaceae demonstrated the highest transcription efficiency during spring, particularly at bloom stations, and showed strong positive correlations with water temperature, F_v/F_m and Chl *a* concentrations. These results underscore the adaptability of Chaetocerotaceae to changing environmental conditions and its pivotal role in sustaining high primary production during spring blooms. Conversely, pennate diatoms exhibited relatively lower transcription efficiency and were more negatively impacted by environmental factors such as salinity changes, limiting their contribution to phytoplankton blooms in open waters.

Furthermore, our analysis revealed that cDNA copy number of the *rbcL* gene correlated more strongly with diatom biomass-related proxies, such as fucoxanthin compared to DNA-based measurements. This suggests that transcriptional activity provides a more accurate representation of the functional dynamics of diatoms in response to environmental changes.

This study emphasizes the importance of diatom composition and *rbcL* transcription efficiency in shaping phytoplankton community dynamics and primary productivity in seasonally ice-covered marine ecosystems. Future research should further explore the physiological mechanisms driving differences in *rbcL* gene expression level among diatom genera and their implications for carbon cycling under ongoing climate change.

Reference

Abad, D., Albaina, A., Aguirre, M., Laza-Martínez, A., Uriarte, I., Iriarte, A., Villate, F., Estonba, A., 2016. Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. Mar. Biol. 163. https://doi.org/10.1007/s00227-016-2920-0

Alberdi, A., Aizpurua, O., Gilbert, M.T.P., Bohmann, K., 2018. Scrutinizing key steps for reliable metabarcoding of environmental samples. Methods Ecol. Evol. 9. https://doi.org/10.1111/2041-210X.12849

Ardyna, M., Arrigo, K.R., 2020. Phytoplankton dynamics in a changing Arctic Ocean. Nat. Clim. Chang. https://doi.org/10.1038/s41558-020-0905-y

Arrigo, K.R., 2014. Sea Ice Ecosystems. Ann. Rev. Mar. Sci. 6. https://doi.org/10.1146/annurev-marine-010213-135103

Arrigo, K.R., van Dijken, G., Pabi, S., 2008. Impact of a shrinking Arctic ice cover on marine primary production. Geophys. Res. Lett. 35. https://doi.org/10.1029/2008GL035028

Arrigo, K.R., van Dijken, G.L., 2015. Continued increases in Arctic Ocean primary production. Prog. Oceanogr. 136. https://doi.org/10.1016/j.pocean.2015.05.002

Booth, B.C., Larouche, P., Bélanger, S., Klein, B., Amiel, D., Mei, Z.P., 2002. Dynamics of Chaetoceros socialis blooms in the North Water. Deep. Res. Part II Top. Stud. Oceanogr. 49. https://doi.org/10.1016/S0967-0645(02)00175-3

Comiso, J.C., Nishio, F., 2008. Trends in the sea ice cover using enhanced and compatible AMSR-E, SSM/I, and SMMR data. J. Geophys. Res. Ocean. 113. https://doi.org/10.1029/2007JC004257

Corrector, J.E., Wawrik, B., Paul, J.H., Tran, H., Kerkhof, L., López, J.M., Dieppa, A., Cárdenas, O., 2004. Geochemical rate-RNA integration study: Ribulose-1,5-bisphosphate carboxylase/oxygenase gene transcription and photosynthetic capacity of planktonic photoautotrophs. Appl. Environ. Microbiol. 70.

https://doi.org/10.1128/AEM.70.9.5459-5468.2004

Degerlund, M., Eilertsen, H.C., 2010. Main Species Characteristics of Phytoplankton Spring Blooms in NE Atlantic and Arctic Waters (68-80° N). Estuaries and Coasts 33. https://doi.org/10.1007/s12237-009-9167-7

Endo, H., Hattori, H., Mishima, T., Hashida, G., Sasaki, H., Nishioka, J., Suzuki, K., 2017. Phytoplankton community responses to iron and CO2 enrichment in different biogeochemical regions of the Southern Ocean. Polar Biol. 40. https://doi.org/10.1007/s00300-017-2130-3

Endo, H., Sugie, K., Yoshimura, T., Suzuki, K., 2016. Response of spring diatoms to CO2 availability in the western North Pacific as determined by next-generation sequencing. PLoS One 11. https://doi.org/10.1371/journal.pone.0154291

Endo, H., Sugie, K., Yoshimura, T., Suzuki, K., 2015. Effects of CO2 and iron availability on rbcL gene expression in Bering Sea diatoms. Biogeosciences 12. https://doi.org/10.5194/bg-12-2247-2015

Ersland, D.R., Aldrich, J., Cattolico, R.A., 1981. Kinetic Complexity, Homogeneity, and Copy Number of Chloroplast DNA from the Marine Alga Olisthodiscus luteus . Plant Physiol. 68. https://doi.org/10.1104/pp.68.6.1468

Freyria, N.J., Joli, N., Lovejoy, C., 2021. A decadal perspective on north water microbial eukaryotes as Arctic Ocean sentinels. Sci. Rep. 11. https://doi.org/10.1038/s41598-021-87906-4

Gemmell, B.J., Oh, G., Buskey, E.J., Villareal, T.A., 2016. Dynamic sinking behaviour in marine phytoplankton: Rapid changes in buoyancy may aid in nutrient uptake. Proc. R. Soc. B Biol. Sci. 283. https://doi.org/10.1098/rspb.2016.1126

Glibert, P.M., Wilkerson, F.P., Dugdale, R.C., Raven, J.A., Dupont, C.L., Leavitt, P.R., Parker, A.E., Burkholder, J.M., Kana, T.M., 2016. Pluses and minuses of ammonium and nitrate uptake and assimilation by phytoplankton and implications for productivity and community composition, with emphasis on nitrogen-enriched conditions. Limnol.

Oceanogr. 61. https://doi.org/10.1002/lno.10203

Gradinger, R., Ikävalko, J., 1998. Organism incorporation into newly forming arctic sea ice in the greenland sea. J. Plankton Res. 20. https://doi.org/10.1093/plankt/20.5.871

Granum, E., Roberts, K., Raven, J.A., Leegood, R.C., 2009. Primary carbon and nitrogen metabolic gene expression in the diatom thalassiosira pseudonana (bacillariophyceae): Diel eriodicity and effects of inorganic carbon and nitrogen. J. Phycol. 45. https://doi.org/10.1111/j.1529-8817.2009.00728.x

Hamao, Y., Matsuno, K., Mitani, Y., Yamaguchi, A., 2022. Spatial distribution of the protist community in the southern part of the Okhotsk Sea off Hokkaido during summer. J. Oceanogr. 78. https://doi.org/10.1007/s10872-021-00630-5

Higgins, H.W., Wright, S.W., Schlüter, L., 2012. Quantitative interpretation of chemotaxonomic pigment data, in: Phytoplankton Pigments. https://doi.org/10.1017/cbo9780511732263.010

Hillebrand, H., Dürselen, C.D., Kirschtel, D., Pollingher, U., Zohary, T., 1999. Biovolume calculation for pelagic and benthic microalgae. J. Phycol. 35. https://doi.org/10.1046/j.1529-8817.1999.3520403.x

Hughes, D.J., Varkey, D., Doblin, M.A., Ingleton, T., Mcinnes, A., Ralph, P.J., van Dongen-Vogels, V., Suggett, D.J., 2018. Impact of nitrogen availability upon the electron requirement for carbon fixation in Australian coastal phytoplankton communities. Limnol. Oceanogr. 63. https://doi.org/10.1002/lno.10814

Jeffrey, S.W., Wright, S.W., Zapata, M., 2012. Microalgal classes and their signature pigments, in: Phytoplankton Pigments. https://doi.org/10.1017/cbo9780511732263.004

John, D.E., López-Díaz, J.M., Cabrera, A., Santiago, N.A., Corredor, J.E., Bronk, D.A., Paul, J.H., 2012. A day in the life in the dynamic marine environment: How nutrients shape diel patterns of phytoplankton photosynthesis and carbon fixation gene expression in the Mississippi and Orinoco River plumes. Hydrobiologia 679. https://doi.org/10.1007/s10750-011-0862-6

John, D.E., Patterson, S.S., Paul, J.H., 2007a. Phytoplankton-group specific quantitative polymerase chain reaction assays for RuBisCO mRNA transcripts in seawater. Mar. Biotechnol. 9. https://doi.org/10.1007/s10126-007-9027-z

John, D.E., Wang, Z.A., Liu, X., Byrne, R.H., Corredor, J.E., López, J.M., Cabrera, A., Bronk, D.A., Tabita, F.R., Paul, J.H., 2007b. Phytoplankton carbon fixation gene (RuBisCO) transcripts and air-sea CO"2 flux in the Mississippi River plume. ISME J. 1. https://doi.org/10.1038/ismej.2007.70

Johnson, Z.I., Martiny, A.C., 2015. Techniques for quantifying phytoplankton biodiversity. Ann. Rev. Mar. Sci. 7. https://doi.org/10.1146/annurev-marine-010814-015902

Kahru, M., Lee, Z., Mitchell, B.G., Nevison, C.D., 2016. Effects of sea ice cover on satellite-detected primary production in the Arctic Ocean. Biol. Lett. 12. https://doi.org/10.1098/rsbl.2016.0223

Kasai, H., Hirakawa, K., 2016. Seasonal changes of primary production in the southwestern okhotsk sea off hokkaido, Japan during the ice-free period. Plankt. Benthos Res. 10. https://doi.org/10.3800/pbr.10.178

Kasai, H., Nakano, Y., Ono, T., Tsuda, A., 2010. Seasonal change of oceanographic conditions and chlorophyll a vertical distribution in the southwestern Okhotsk Sea during the non-iced season. J. Oceanogr. 66. https://doi.org/10.1007/s10872-010-0002-3

Kembel, S.W., Wu, M., Eisen, J.A., Green, J.L., 2012. Incorporating 16S Gene Copy Number Information Improves Estimates of Microbial Diversity and Abundance. PLoS Comput. Biol. 8. https://doi.org/10.1371/journal.pcbi.1002743

Kolber, Z.S., Prášil, O., Falkowski, P.G., 1998. Measurements of variable chlorophyll fluorescence using fast repetition rate techniques: Defining methodology and experimental protocols. Biochim. Biophys. Acta - Bioenerg. 1367. https://doi.org/10.1016/S0005-2728(98)00135-2

Kong, W., Ream, D.C., Priscu, J.C., Morgan-Kiss, R.M., 2012. Diversity and expression of RubisCO genes in a perennially ice-covered antarctic lake during the polar night

transition. Appl. Environ. Microbiol. 78. https://doi.org/10.1128/AEM.00029-12

Kramer, S.J., Siegel, D.A., 2019. How Can Phytoplankton Pigments Be Best Used to Characterize Surface Ocean Phytoplankton Groups for Ocean Color Remote Sensing Algorithms? J. Geophys. Res. Ocean. 124. https://doi.org/10.1029/2019JC015604

Kuga, M., Ohshima, K.I., Kishi, S., Kimura, N., Toyota, T., Nishioka, J., 2023. Backward-tracking simulations of sea ice in the Sea of Okhotsk toward understanding of material transport through sea ice. J. Oceanogr. 1–12.

Kvernvik, A.C., Hoppe, C.J.M., Lawrenz, E., Prášil, O., Greenacre, M., Wiktor, J.M., Leu, E., 2018. Fast reactivation of photosynthesis in arctic phytoplankton during the polar night1. J. Phycol. 54. https://doi.org/10.1111/jpy.12750

Lacour, T., Morin, P.I., Sciandra, T., Donaher, N., Campbell, D.A., Ferland, J., Babin, M., 2019. Decoupling light harvesting, electron transport and carbon fixation during prolonged darkness supports rapid recovery upon re-illumination in the Arctic diatom Chaetoceros neogracilis. Polar Biol. 42. https://doi.org/10.1007/s00300-019-02507-2

Lovejoy, C., Legendre, L., Martineau, M.J., Bâcle, J., Von Quillfeldt, C.H., 2002. Distribution of phytoplankton and other protists in the North Water. Deep. Res. Part II Top. Stud. Oceanogr. 49. https://doi.org/10.1016/S0967-0645(02)00176-5

Mackey, M.D., Mackey, D.J., Higgins, H.W., Wright, S.W., 1996. CHEMTAX - A program for estimating class abundances from chemical markers: Application to HPLC measurements of phytoplankton. Mar. Ecol. Prog. Ser. 144. https://doi.org/10.3354/meps144265

Matsumoto, T., Matsuno, K., Katakura, S., Kasai, H., Yamaguchi, A., 2021. Seasonal variability of the protist community and production in the southern Okhotsk Sea revealed by weekly monitoring. Reg. Stud. Mar. Sci. 43. https://doi.org/10.1016/j.rsma.2021.101683

McMinn, A., Martin, A., 2013. Dark survival in a warming world. Proc. R. Soc. B Biol. Sci. https://doi.org/10.1098/rspb.2012.2909

Morse, D., Salois, P., Markovic, P., Hastings, J.W., 1995. A nuclear-encoded form II RuBisCO in dinoflagellates. Science (80-.). 268. https://doi.org/10.1126/science.7777861

Mundo, R., Matsunaka, T., Nakanowatari, T., Taniuchi, Y., Inoue, M., Kasai, H., Mashita, K., Mitsunushi, H., Seiya, N., 2023. A review of the oceanographic structure and biological productivity in the southern Okhotsk Sea. Prog. Oceanogr. 103194.

Naik, R.K., Anil, A.C., Narale, D.D., Chitari, R.R., Kulkarni, V. V., 2011. Primary description of surface water phytoplankton pigment patterns in the Bay of Bengal. J. Sea Res. 65. https://doi.org/10.1016/j.seares.2011.03.007

Nakov, T., Ashworth, M., Theriot, E.C., 2015. Comparative analysis of the interaction between habitat and growth form in diatoms. ISME J. 9. https://doi.org/10.1038/ismej.2014.108

Nishioka, J., Obata, H., Ogawa, H., Ono, K., Yamashita, Y., Lee, K., Takeda, S., Yasuda, I., 2020. Subpolar marginal seas fuel the North Pacific through the intermediate water at the termination of the global ocean circulation. Proc. Natl. Acad. Sci. U. S. A. 117. https://doi.org/10.1073/pnas.2000658117

Ohwada, M., 1957. Diatom Communities in the Okhotsk sea, Principally on the West Coast of Kamchatka, Spring to Summer 1955. J. Oceanogr. Soc. Japan 13. https://doi.org/10.5928/kaiyou1942.13.29

Piñol, J., Senar, M.A., Symondson, W.O.C., 2019. The choice of universal primers and the characteristics of the species mixture determine when DNA metabarcoding can be quantitative. Mol. Ecol. 28. https://doi.org/10.1111/mec.14776

Raven, J.A., 1998. The twelfth Tansley Lecture. Small is beautiful: The picophytoplankton. Funct. Ecol. https://doi.org/10.1046/j.1365-2435.1998.00233.x

Rhiel, E., Krupinska, K., Wehrmeyer, W., 1986. Effects of nitrogen starvation on the function and organization of the photosynthetic membranes in Cryptomonas maculata (Cryptophyceae). Planta 169. https://doi.org/10.1007/BF00392132

Samson, G., Prášil, O., Yaakoubd, B., 1999. Photochemical and thermal phases of chlorophyll a fluorescence, in: Photosynthetica. https://doi.org/10.1023/A:1007095619317

Schaeffer, B.A., Kurtz, J.C., Hein, M.K., 2012. Phytoplankton community composition in nearshore coastal waters of Louisiana. Mar. Pollut. Bull. 64. https://doi.org/10.1016/j.marpolbul.2012.03.017

Schiffrine, N., Tremblay, J.É., Babin, M., 2020. Growth and Elemental Stoichiometry of the Ecologically-Relevant Arctic Diatom Chaetoceros gelidus: A Mix of Polar and Temperate. Front. Mar. Sci. 6. https://doi.org/10.3389/fmars.2019.00790

Selz, V., Saenz, B.T., van Dijken, G.L., Arrigo, K.R., 2018. Drivers of Ice Algal Bloom Variability Between 1980 and 2015 in the Chukchi Sea. J. Geophys. Res. Ocean. 123. https://doi.org/10.1029/2018JC014123

Sergeeva, V.M., Sukhanova, I.N., Flint, M. V., Pautova, L.A., Grebmeier, J.M., Cooper, L.W., 2010. Phytoplankton community in the Western Arctic in july-august 2003. Oceanology 50. https://doi.org/10.1134/S0001437010020049

Shimada, H., Sawada, M., Tanaka, I., Asami, H., Fukamachi, Y., 2012. A method for predicting the occurrence of paralytic shellfish poisoning along the coast of Hokkaido in the Okhotsk Sea in summer. Fish. Sci. 78. https://doi.org/10.1007/s12562-012-0513-5

Sildever, S., Kawakami, Y., Kanno, N., Kasai, H., Shiomoto, A., Katakura, S., Nagai, S., 2019. Toxic HAB species from the Sea of Okhotsk detected by a metagenetic approach, seasonality and environmental drivers. Harmful Algae 87. https://doi.org/10.1016/j.hal.2019.101631

Sourisseau, M., Font-Muñoz, J., Bellouche, S., Fauvarque, O., Rouxel, J., Tardivel, M., Sauvey, A., 2024. Sinking rates, orientation, and behavior of pennate diatoms. J. Phycol.

Suggett, D.J., Macintyre, H.L., Geider, R.J., 2004. Evaluation of biophysical and optical determinations of light absorption by photosystem II in phytoplankton. Limnol. Oceanogr. Methods 2. https://doi.org/10.4319/lom.2004.2.316

Sun, J., Liu, D., 2003. Geometric models for calculating cell biovolume and surface area for phytoplankton. J. Plankton Res. 25. https://doi.org/10.1093/plankt/fbg096

Sun, Y., Nishioka, J., Toyota, T., Suzuki, K., 2025. Determining dominant phytoplankton assemblages and their controlling factors in the winter sea-ice-covered Southern Sea of Okhotsk and the spring open water through a multiple analytical approach. Mar. Environ. Res. 107297.

Suzuki, K., Hattori-Saito, A., Sekiguchi, Y., Nishioka, J., Shigemitsu, M., Isada, T., Liu, H., McKay, R.M.L., 2014. Spatial variability in iron nutritional status of large diatoms in the Sea of Okhotsk with special reference to the Amur River discharge. Biogeosciences 11. https://doi.org/10.5194/bg-11-2503-2014

Suzuki, K., Kamimura, A., Hooker, S.B., 2015. Rapid and highly sensitive analysis of chlorophylls and carotenoids from marine phytoplankton using ultra-high performance liquid chromatography (UHPLC) with the first derivative spectrum chromatogram (FDSC) technique. Mar. Chem. 176. https://doi.org/10.1016/j.marchem.2015.07.010

Toyota, T., Kimura, N., Nishioka, J., Ito, M., Nomura, D., Mitsudera, H., 2022. The Interannual Variability of Sea Ice Area, Thickness, and Volume in the Southern Sea of Okhotsk and Its Likely Factors. J. Geophys. Res. Ocean. 127. https://doi.org/10.1029/2022JC019069

Tremblay, J.É., Michel, C., Hobson, K.A., Gosselin, M., Price, N.M., 2006. Bloom dynamics in early opening waters of the Arctic Ocean. Limnol. Oceanogr. 51. https://doi.org/10.4319/lo.2006.51.2.0900

Treusch, A.H., Demir-Hilton, E., Vergin, K.L., Worden, A.Z., Carlson, C.A., Donatz, M.G., Burton, R.M., Giovannoni, S.J., 2012. Phytoplankton distribution patterns in the northwestern Sargasso Sea revealed by small subunit rRNA genes from plastids. ISME J. 6. https://doi.org/10.1038/ismej.2011.117

Valegård, K., Andralojc, P.J., Haslam, R.P., Pearce, F.G., Eriksen, G.K., Madgwick, P.J., Kristoffersen, A.K., van Lun, M., Klein, U., Eilertsen, H.C., Parry, M.A.J., Andersson,

I., 2018. Structural and functional analyses of Rubisco from arctic diatom species reveal unusual posttranslational modifications. J. Biol. Chem. 293. https://doi.org/10.1074/jbc.RA118.003518

Vasselon, V., Bouchez, A., Rimet, F., Jacquet, S., Trobajo, R., Corniquel, M., Tapolczai, K., Domaizon, I., 2018. Avoiding quantification bias in metabarcoding: Application of a cell biovolume correction factor in diatom molecular biomonitoring. Methods Ecol. Evol. 9. https://doi.org/10.1111/2041-210X.12960

Vasselon, V., Rimet, F., Tapolczai, K., Bouchez, A., 2017. Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte island, France). Ecol. Indic. 82. https://doi.org/10.1016/j.ecolind.2017.06.024

Von Quillfeldt, C.H., 2001. Identification of some easily confused common diatom species in arctic spring blooms. Bot. Mar. 44. https://doi.org/10.1515/BOT.2001.048

Wahl, R.J., Teague, W.J., 1983. Estimation of Brunt-Väisälä Frequency from Temperature Profiles. J. Phys. Oceanogr. 13. https://doi.org/10.1175/1520-0485(1983)013<2236:eobvff>2.0.co;2

Waite, A.M., Thompson, P.A., Harrison, P.J., 1992. Does energy control the sinking rates of marine diatoms? Limnol. Oceanogr. 37. https://doi.org/10.4319/lo.1992.37.3.0468

Wirtz, K.W., 2011. Non-uniform scaling in phytoplankton growth rate due to intracellular light and CO2 decline. J. Plankton Res. 33, 1325–1341.

Xiao, X., Sogge, H., Lagesen, K., Tooming-Klunderud, A., Jakobsen, K.S., Rohrlack, T., 2014. Use of high throughput sequencing and light microscopy show contrasting results in a study of phytoplankton occurrence in a freshwater environment. PLoS One 9. https://doi.org/10.1371/journal.pone.0106510

Yan, D., Nishioka, J., Toyota, T., Suzuki, K., 2022. Winter microalgal communities of the southern Sea of Okhotsk: A comparison of sea ice, coastal, and basinal seawater. Prog. Oceanogr. 204, 102806.

Yan, D., Yoshida, K., Nishioka, J., Ito, M., Toyota, T., Suzuki, K., 2020. Response to Sea Ice Melt Indicates High Seeding Potential of the Ice Diatom Thalassiosira to Spring Phytoplankton Blooms: A Laboratory Study on an Ice Algal Community From the Sea of Okhotsk. Front. Mar. Sci. 7. https://doi.org/10.3389/fmars.2020.00613

Young, J.N., Heureux, A.M.C., Sharwood, R.E., Rickaby, R.E.M., Morel, F.M.M., Whitney, S.M., 2016. Large variation in the Rubisco kinetics of diatoms reveals diversity among their carbon-concentrating mechanisms. J. Exp. Bot. 67. https://doi.org/10.1093/jxb/erw163

Young, J.N., Rickaby, R.E.M., Kapralov, M. V., Filatov, D.A., 2012. Adaptive signals in algal Rubisco reveal a history of ancient atmospheric carbon dioxide. Philos. Trans. R. Soc. B Biol. Sci. 367. https://doi.org/10.1098/rstb.2011.0145

Zimmermann, J., Jahn, R., Gemeinholzer, B., 2011. Barcoding diatoms: Evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. Org. Divers. Evol. 11. https://doi.org/10.1007/s13127-011-0050-6

Tables, Figures and Supplementary

Table 1 Summary of environmental variables from the samples collected during the P/V Soya (winter) and R/V Shinsei Maru (spring) cruises, 2022.

Season	Stations	Sampling Date	Sampling Time (JST)	Sampling Depths (m)	Mixed layer Depth (m)	Brunt-Väisälä Frequency (s^{-2})	Temperature (°C)	Salinity	Sea ice concentration (%)	Fv/Fm	NO_2+NO_3 (μmol L^{-1})	NH_4 (μmol L^{-1})	PO_4 (μmol L^{-1})	SiO_4 (μmol L^{-1})	Chl *a* (mg m^{-3})	Fucoxanthin (mg m^{-3})
Winter	21	2022/2/11	7:15	5	12	6.09×10^{-5}	-1.43	31.71	0	0.28±0.09	9.85±0.07	1.70±0.18	1.17±0.00	20.00±0.03	0.34±0.03	0.14±0.02
	17w	2022/2/12	7:15	5	18	2.94×10^{-5}	-1.59	32.07	0	0.23±0.05	12.06±0.03	0.94±0.09	1.33±0.00	23.29±0.08	0.27±0.00	0.09±0.00
	14	2022/2/13	7:15	5	34	4.77×10^{-5}	-1.73	32.86	95	0.20±0.05	18.57±0.09	1.14±0.74	1.65±0.00	26.65±0.09	0.27±0.00	0.09±0.00
	10	2022/2/14	7:15	5	31	6.52×10^{-5}	-1.75	32.96	88	0.19±0.07	20.33±0.01	0.67±0.33	1.68±0.11	30.02±0.04	0.26±0.01	0.08±0.00
	4	2022/2/15	7:15	5	11	5.92×10^{-4}	-1.46	31.99	24	0.20±0.02	10.23±0.02	2.26±0.17	1.18±0.00	20.28±0.24	0.37±0.00	0.11±0.00
Spring	18	2022/5/1	11:20	0	33	1.52×10^{-4}	2.19	32.53	0	0.55±0.03	0.21±0.05	0.34±0.17	0.38±0.02	2.50±0.33	3.10±0.73	1.96±0.24
	17s	2022/5/2	14:22	5	25	2.40×10^{-4}	1.94	32.51	0	0.58±0.02	1.53±0.07	1.00±0.30	0.97±0.62	4.42±0.03	4.13±0.47	2.70±0.31
	204	2022/5/3	12:53	5	43	7.93×10^{-5}	1.91	32.61	0	0.52±0.01	5.45±0.21	0.45±0.16	0.81±0.06	9.78±0.37	1.24±0.11	0.85±0.09
	Bio1	2022/5/5	13:48	5	17	-6.69×10^{-4}	2.98	32.53	0	0.42±0.02	1.02±0.03	0.61±0.03	0.46±0.00	3.66±0.01	1.23±0.26	0.79±0.07
	Bio3	2022/5/8	8:33	0	31	1.42×10^{-4}	2.89	32.59	0	0.53±0.02	5.39±0.17	0.90±0.32	0.83±0.02	12.87±0.34	1.33±0.03	0.65±0.02
	11	2022/5/8	17:31	0	NA	NA	NA	NA	0	0.32±0.03	11.15±0.32	0.96±0.11	1.22±0.06	23.98±0.76	0.75±0.00	0.11±0.00
	18R	2022/5/9	14:12	5	28	2.82×10^{-4}	2.52	32.54	0	0.52±0.02	3.98±0.11	0.63±0.23	0.69±0.04	9.36±0.32	0.58±0.04	0.43±0.00
	18TM	2022/5/10	11:23	0	29	1.58×10^{-5}	2.43	32.6	0	0.47±0.02	3.97±0.34	0.91±0.11	0.71±0.03	10.22±0.39	0.62±0.06	0.71±0.10
	NS5	2022/5/12	12:05	0	15	2.05×10^{-5}	4.33	32.54	0	0.11±0.04	0.15±0.02	1.03±0.21	0.29±0.01	1.41±0.01	0.16±0.02	0.12±0.00

Table 2 Comparison of This Study with Previous Research on RuBisCO gene expression in Phytoplankton

Study/Reference	Ocean/Sea	Dominant groups	Type of study	RuBisCO Form	Notes
This study	Sea of Okhotsk (Seasonal Ice Zone)	Chaetocerotaceae, Thalassiosiraceae, and Bacillariaceae (diatoms)	*rbcL* gene expression level (mRNA/DNA) in field; *Fv*/*Fm* measurements	Form ID	*rbcL* gene expression level (cDNA/DNA) strongly correlated with *Fv*/*Fm* and Chl *a*
(John et al., 2007a)	Gulf of Mexico (Oligotrophic site)	*Phaeodactylum tricornutum* (diatom)	Real-time qPCR of *rbcL* mRNA in culture and field	Form ID	Field: Maximum *rbcL* mRNA at chlorophyll max depth Lab: Diurnal expression under light/dark cycle
(Corrector et al., 2004)	Northwest Atlantic Coastal Ocean	diatoms, prymnesiophytes, chrysophytes, green flagellates	Real-time qPCR of *rbcL* mRNA; photosynthetic capacity (Pmax) measurements	Form ID, IB, IA	Significant correlation between *rbcL* mRNA and P_{max} *rbcL* mRNA is a significant control point for carbon fixation during diatom bloom events
(Endo et al., 2015)	Bering Sea	diatoms, haptophytes	rbcL gene expression level (mRNA/DNA) On-deck incubation (manipulating CO_2 and Fe levels)	Form ID	rbcL expression level decreased under high CO_2; Diatom biomass declined with rising CO_2 under Fe limitation
(Kong et al., 2012)	Lake Bonney, Antarctica	stramenopiles, cryptophytes, haptophytes	Seasonal transition monitoring (summer → polar night), qPCR & sequencing of RubisCO genes	Form ID, IB, IA	Strong vertical and seasonal patterns in RubisCO gene expression; Nutrient stress affects shallow phytoplankton RubisCO activity
(Fernández-Pinos et al., 2017)	Atlantic, Indian, and Pacific Oceans	*Prochlorococcus* (cyanobacteria)	Field incubation with complex mixtures of organic pollutants	Form IA	Reduced *rbcL* expression at ~2× environmental organic pollutant levels
(Valegård et al., 2018)	Arctic Ocean	5 Arctic diatom species	Cultivation; kinetic assays; Rubisco crystal structure determination	Form I C/D	High CO_2/O_2 specificity and low O_2 sensitivity, with unique small subunit structures; Multiple unusual posttranslational modifications indicate adaptation to low-light, cold environments

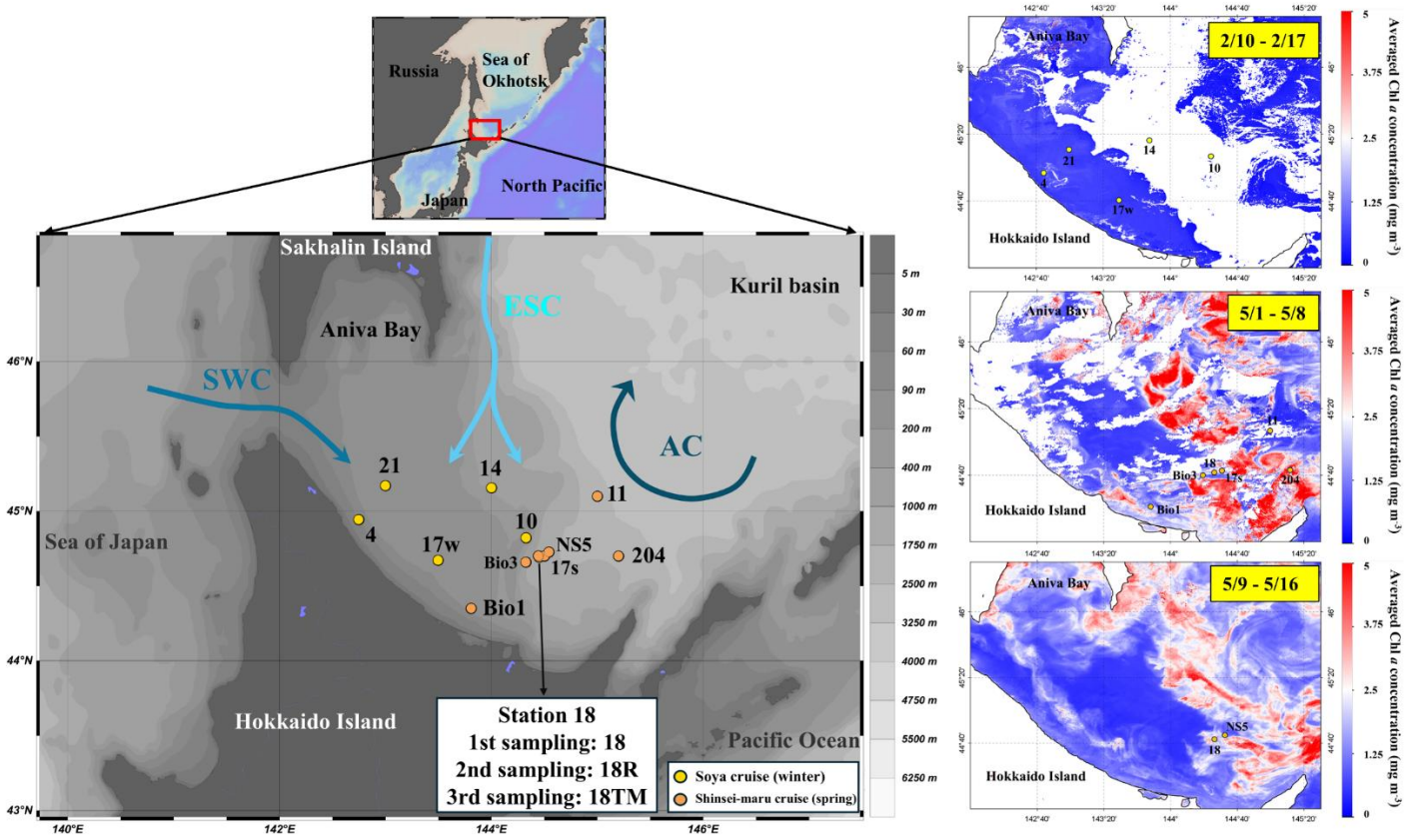


Figure 1 Sampling site map (left) and chlorophyll *a* distribution map (right) based on GCOM-C/SGLI data at the time of sampling for each station. In the left map, yellow and orange circles indicate sampling locations during the P/V Soya (winter) and R/V Shinsei Maru (spring) cruises, respectively. Arrows represent the directions of the major currents in the region: ESC (East Sakhalin Current), SWC (Soya Warm Current), and AC (Anticyclonic Circulation). In the right map, white areas indicate no data due to sea ice or cloud cover.

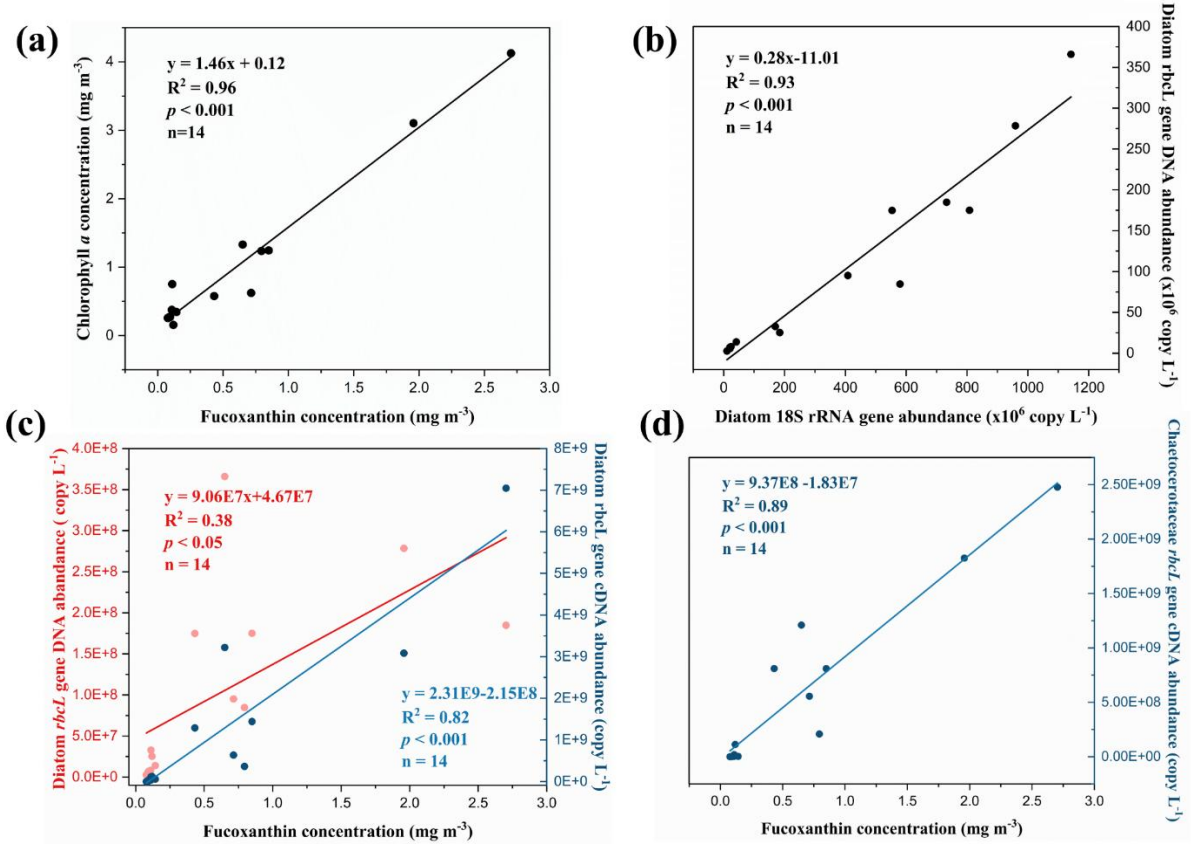


Figure 2 (a) Relationship between fucoxanthin concentration and chlorophyll *a* concentration analyzed using UHPLC; (b) correlation between diatom-specific *rbcL* gene DNA abundance and 18S rRNA gene abundance determined by qPCR; (c) comparison of fucoxanthin concentration with diatom-specific *rbcL* gene DNA abundance (red line, left) and cDNA abundance (blue line, right); (d) correlation between fucoxanthin concentration and *rbcL* gene cDNA abundance derived from Chaetocerotaceae ASVs.

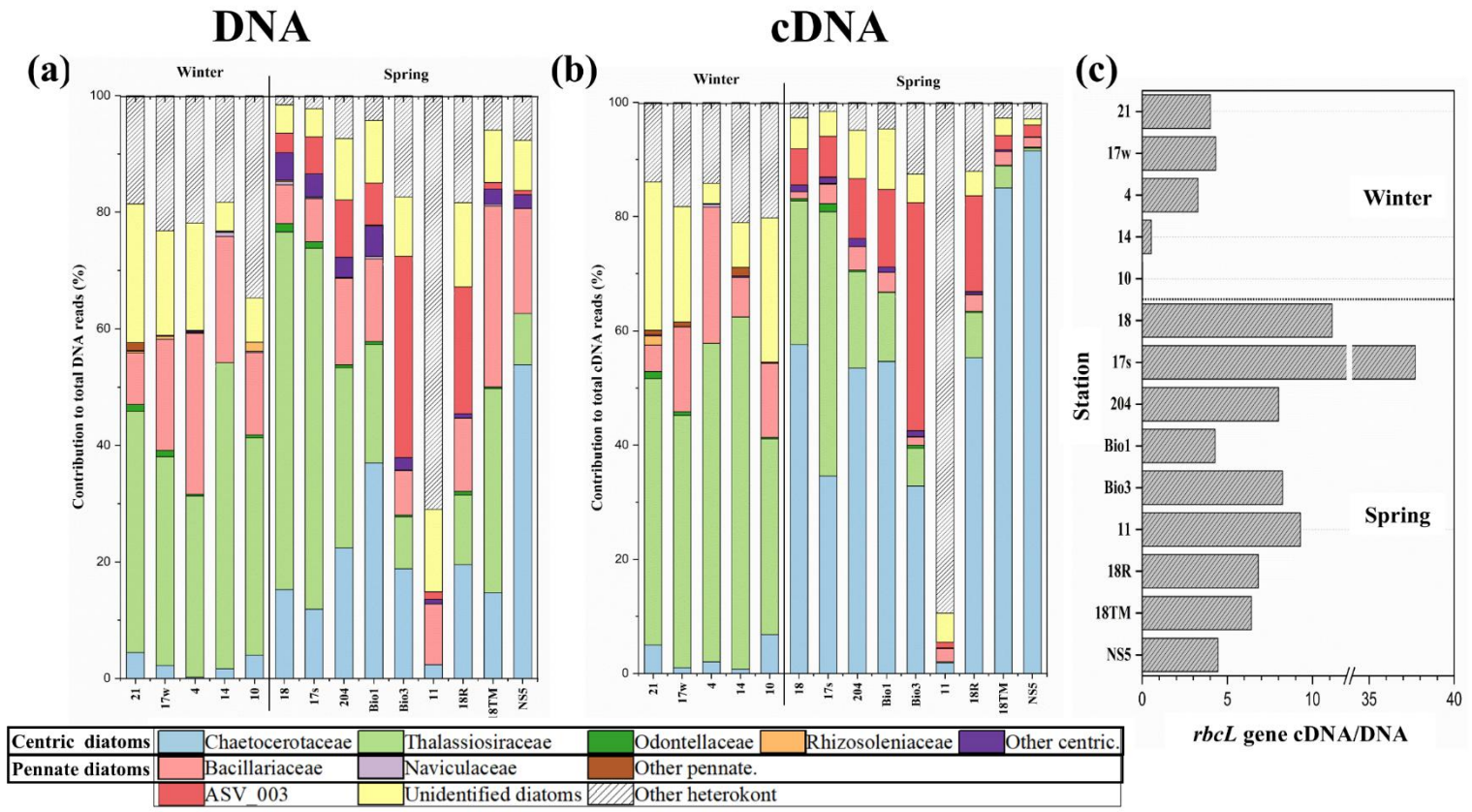


Figure 3 (a) Relative taxonomic contributions (%) of different phytoplankton groups at each station in the *rbcL* DNA libraries, and (b) in the *rbcL* cDNA libraries. (c) The *rbcL* gene transcription efficiency (cDNA/DNA) at each station. In panels (a) and (b), colored bars represent different diatom groups, while black and white bars represent other heterokonts excluding diatoms.

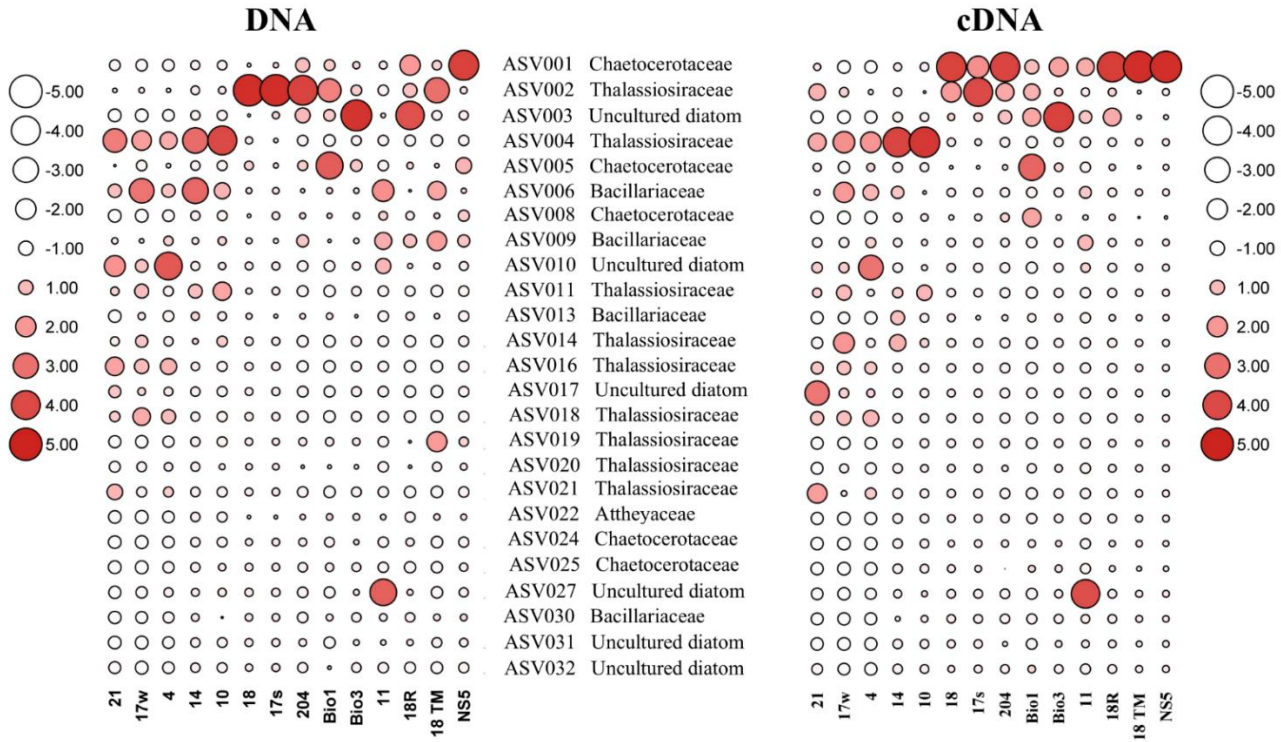


Figure 4 Heatmap of the top 25 diatom-derived ASVs with the highest read counts during winter and spring in the SSO. The left panel shows the absolute abundance of ASVs based on *rbcL* DNA, while the right panel shows their absolute abundance based on *rbcL* cDNA. The color and size of each circle represent normalized values of one ASV absolute abundance within different stations based on Z-score method, absolute abundance of each ASV is calculated by multiplying *rbcL* cDNA and DNA gene abundance by their relative contributions. The top ASVs included members of centric diatom families (Chaetocerotaceae, Thalassiosiraceae, Attheyaceae) and the pennate diatom family Bacillariaceae.

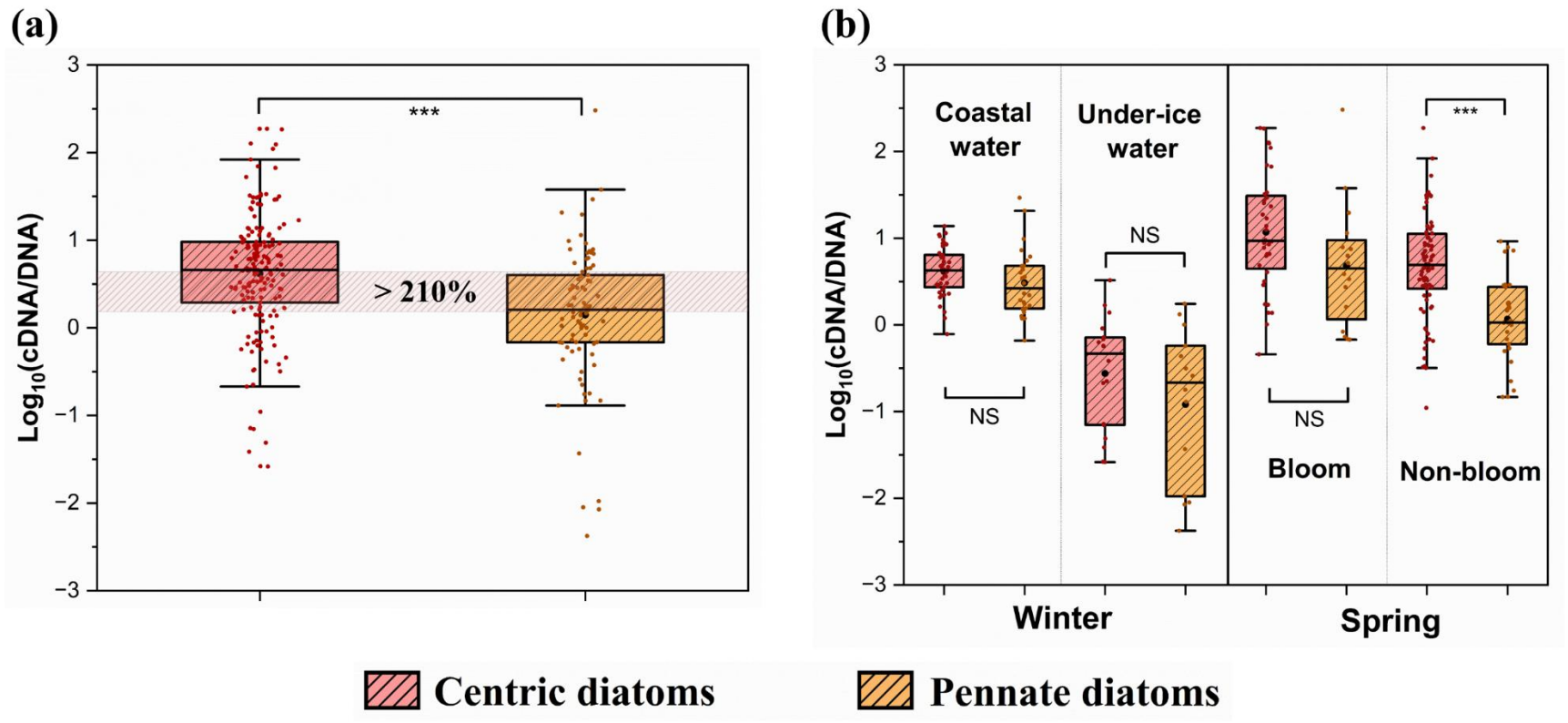


Figure 5 Variations in *rbcL* gene transcription efficiency (cDNA/DNA) at the ASV level for centric and pennate diatoms in the SSO during the sampling period. (a) Differences in *rbcL* transcription efficiency between centric and pennate diatoms across the sampling period. The shaded region indicates a >70% difference in the medians between centric and pennate diatoms. (b) Spatial differences (winter) and temporal differences (spring). "Bloom" and "non-bloom" period in spring are defined as chlorophyll a concentrations greater than 3 mg m^{-3} and less than 1.5 mg m^{-3}, respectively. Asterisks indicate statistical significance levels from Mann-Whitney U tests with FDR correction: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

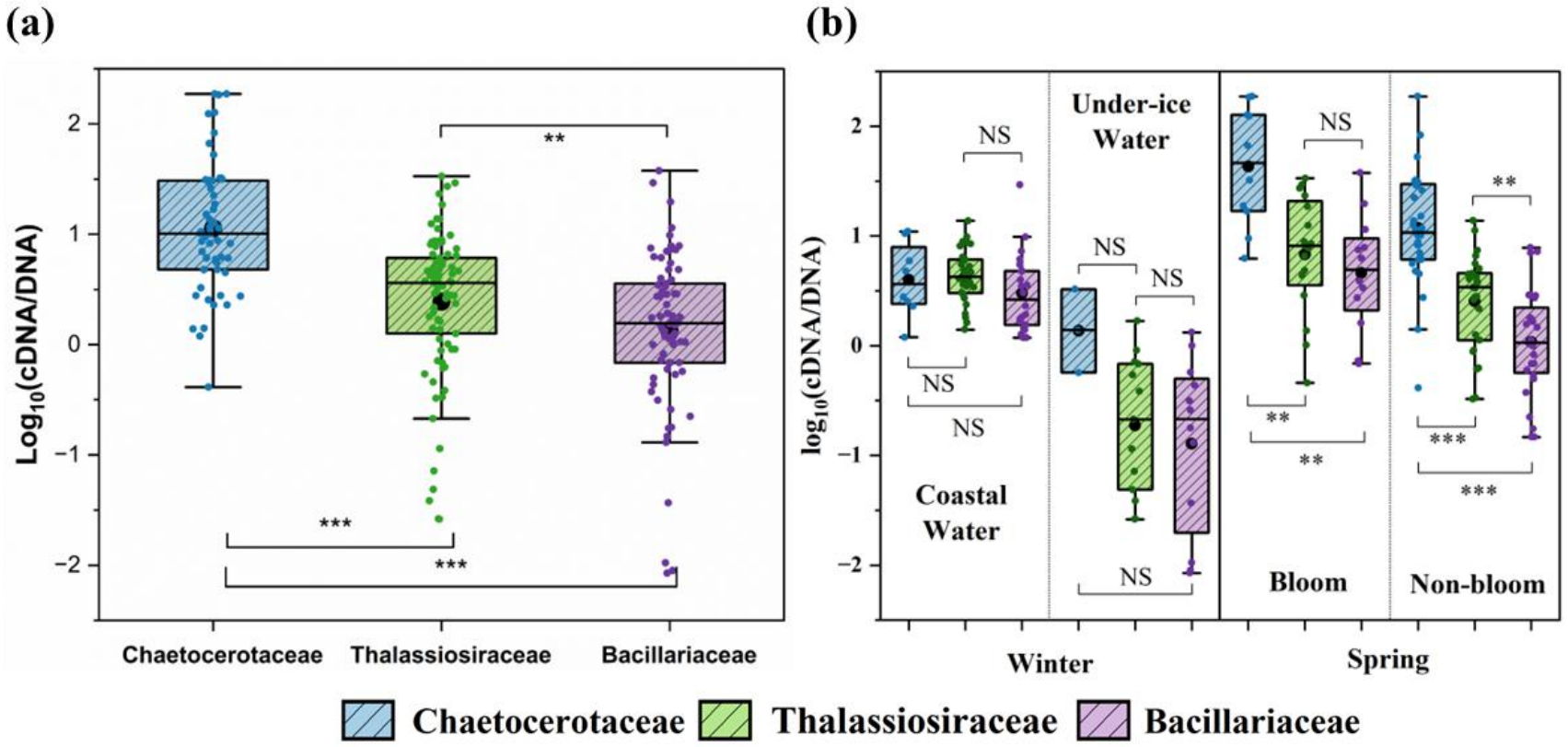


Figure 6 Variations in *rbcL* transcription efficiency (cDNA/DNA) at the ASV level for represented diatoms in the SSO during the sampling period. (a) Differences in *rbcL* transcription efficiency between three dominant diatoms families across the sampling period. (b) Spatial differences (winter) and temporal differences (spring) in *rbcL* gene transcription efficiency for centric diatom family Chaetocerotaceae and Thalassiosiraceae. "Bloom" and "Non-bloom" periods in spring are defined as chlorophyll a concentrations greater than 3 mg m^{-3} and less than 1.5 mg m^{-3}, respectively. Asterisks indicate statistical significance levels from Mann-Whitney U tests with FDR correction: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

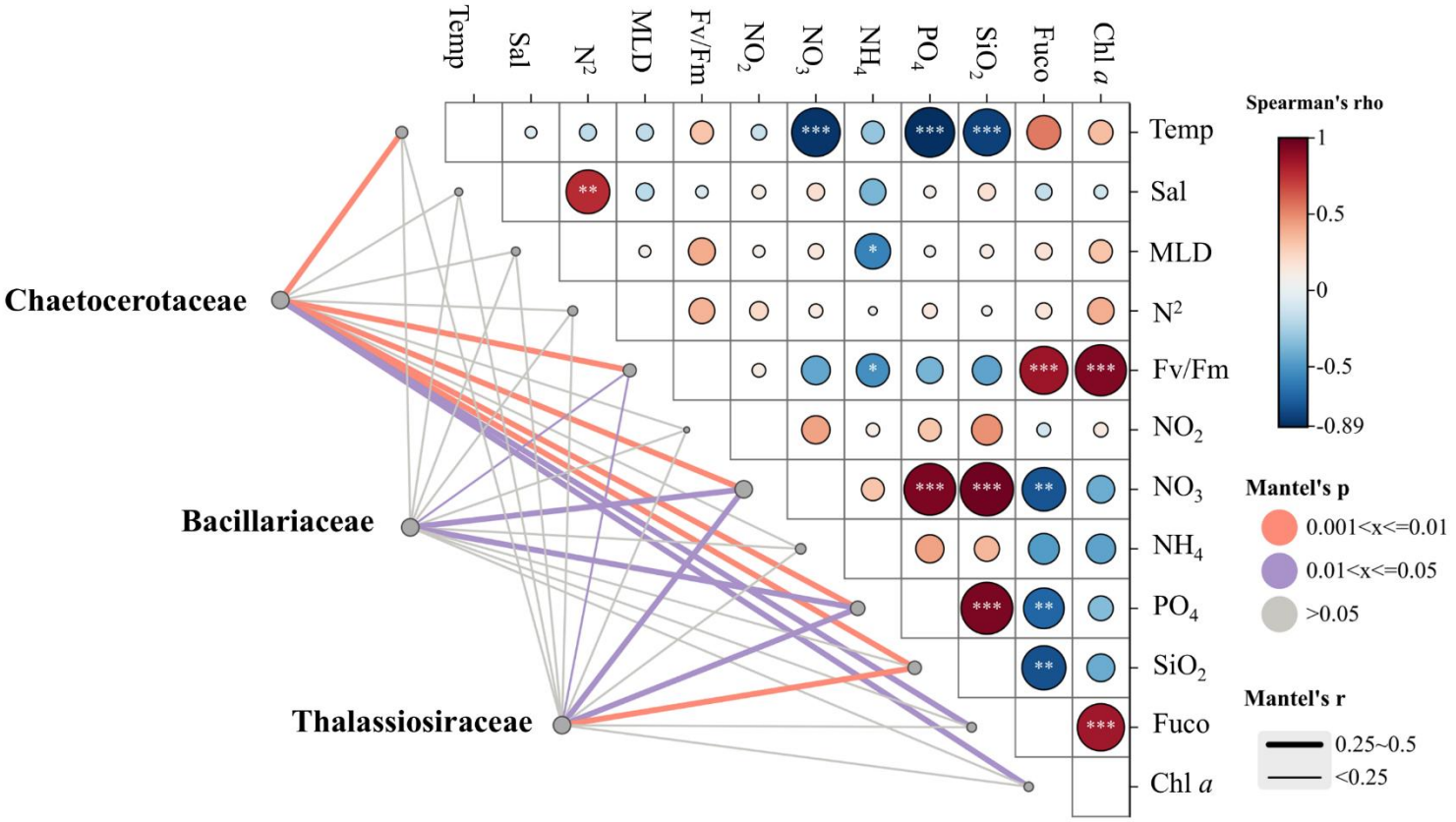


Figure 7. Partial mantel tests between environmental variables and the averaged *rbcL* gene transcription efficiency (cDNA/DNA) of three dominant diatom families. The heatmap shows the pairwise correlations between environmental variables. The lines denote mantel test results, with the line width represents Mantel's r statistic, and the color represents spearman correlation coefficient. Asterisks denote levels of statistical significance based on Spearman correlation tests: *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

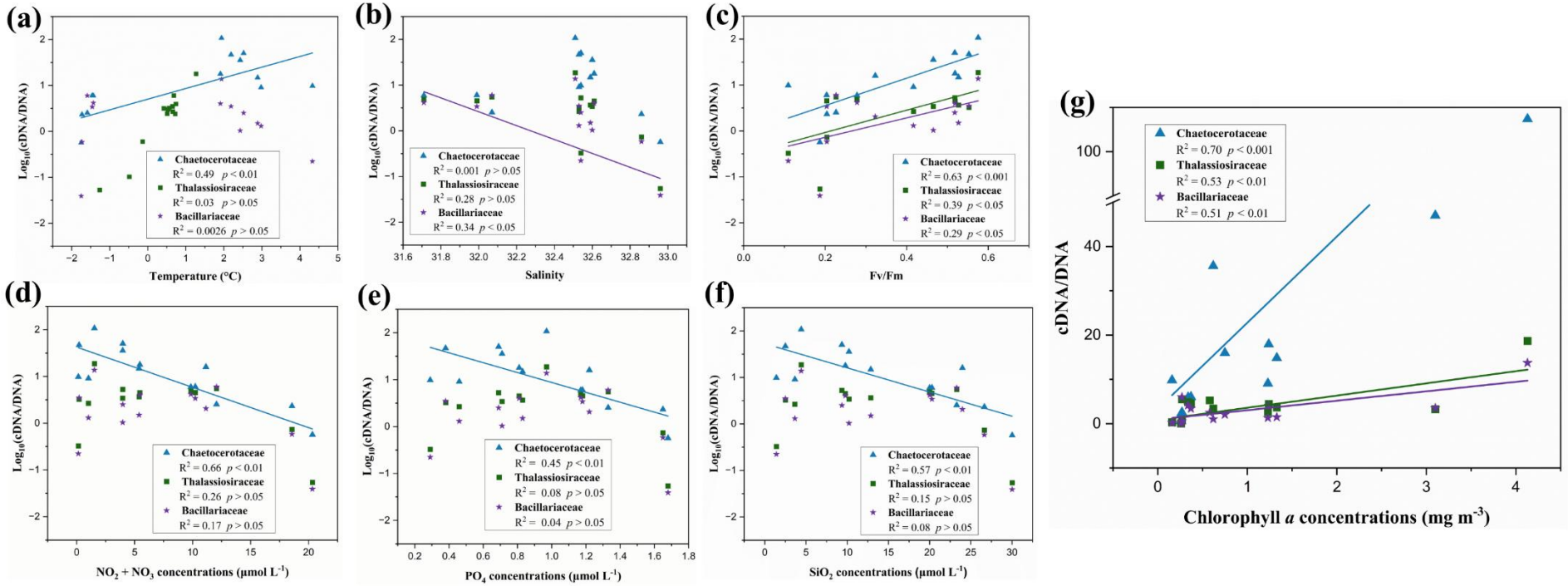


Figure 8 Linear relationship model illustrates the responses of the *rbcL* gene transcription efficiency of three dominant diatom families to environmental variables (a-e) and their relationship with chlorophyll a concentration (f) at each station. Blue triangles, green squares, and purple stars represent Chaetocerotaceae, Thalassiosiraceae, and Bacillariaceae, respectively. Solid lines are shown only where significant linear relationships exist ($p < 0.05$).

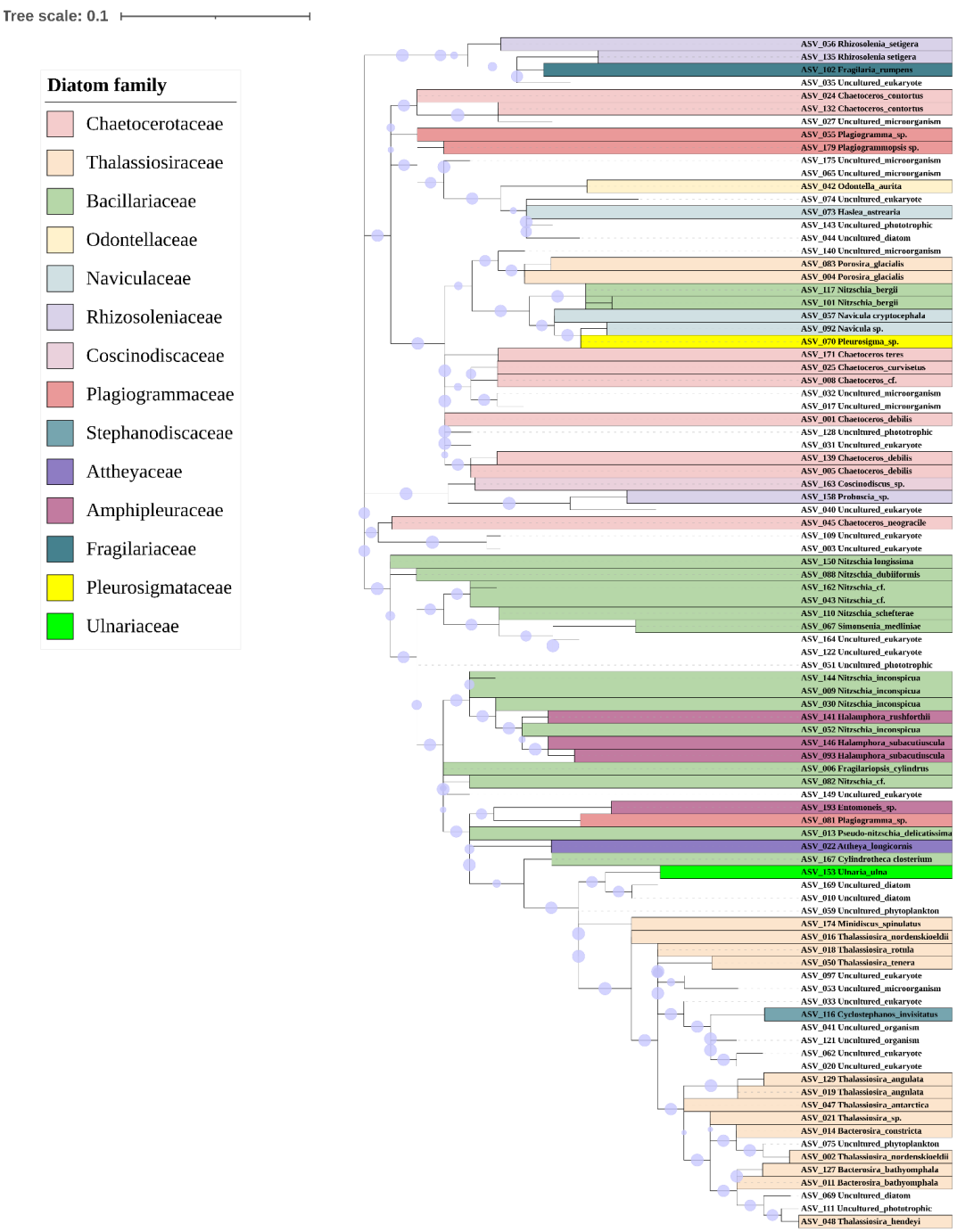


Figure S1 Neighbor-joining tree based on diatom *rbcL* gene reference sequences used for the taxonomic classification. Taxonomic groups are distinguished by different colors in the tree. Reference sequences which could not be identified as a specific diatom family are shown without color.

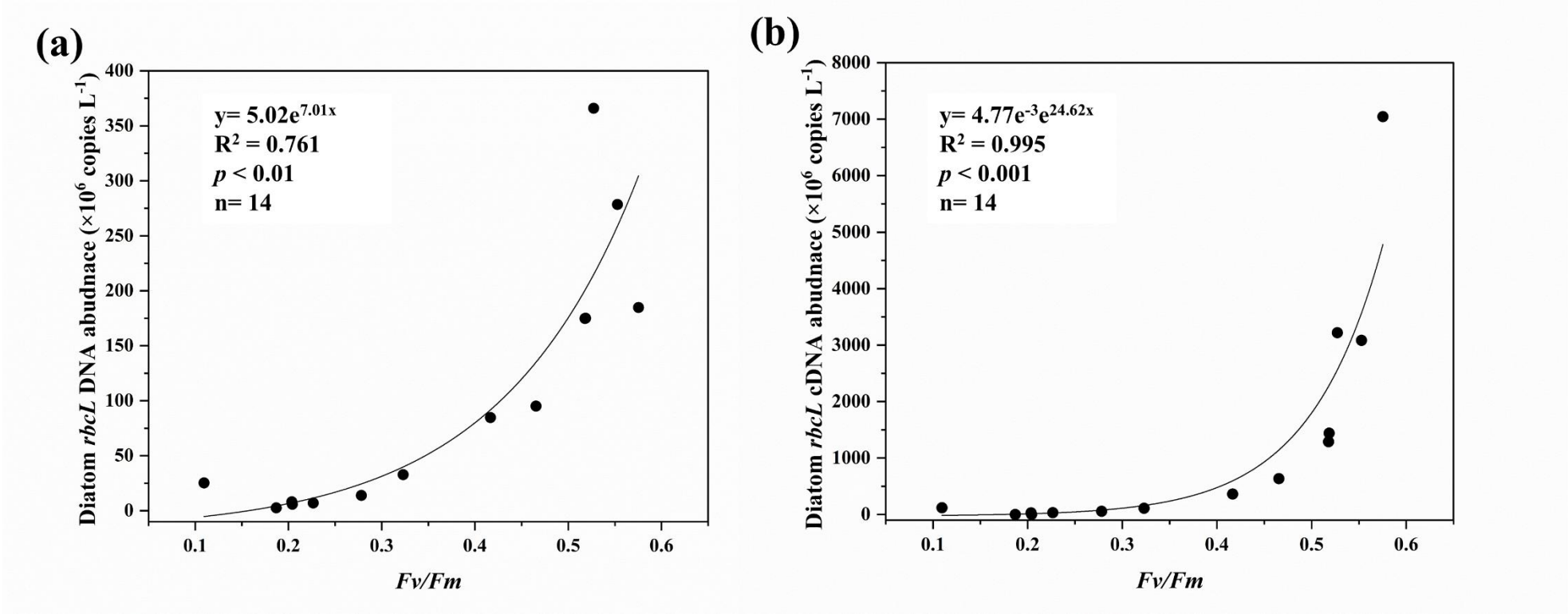


Figure S2. Relationships between the maximum quantum yield of photosystem II (*Fv/Fm*) and diatom-specific *rbcL* gene abundance at the DNA (a) and cDNA (b) levels.

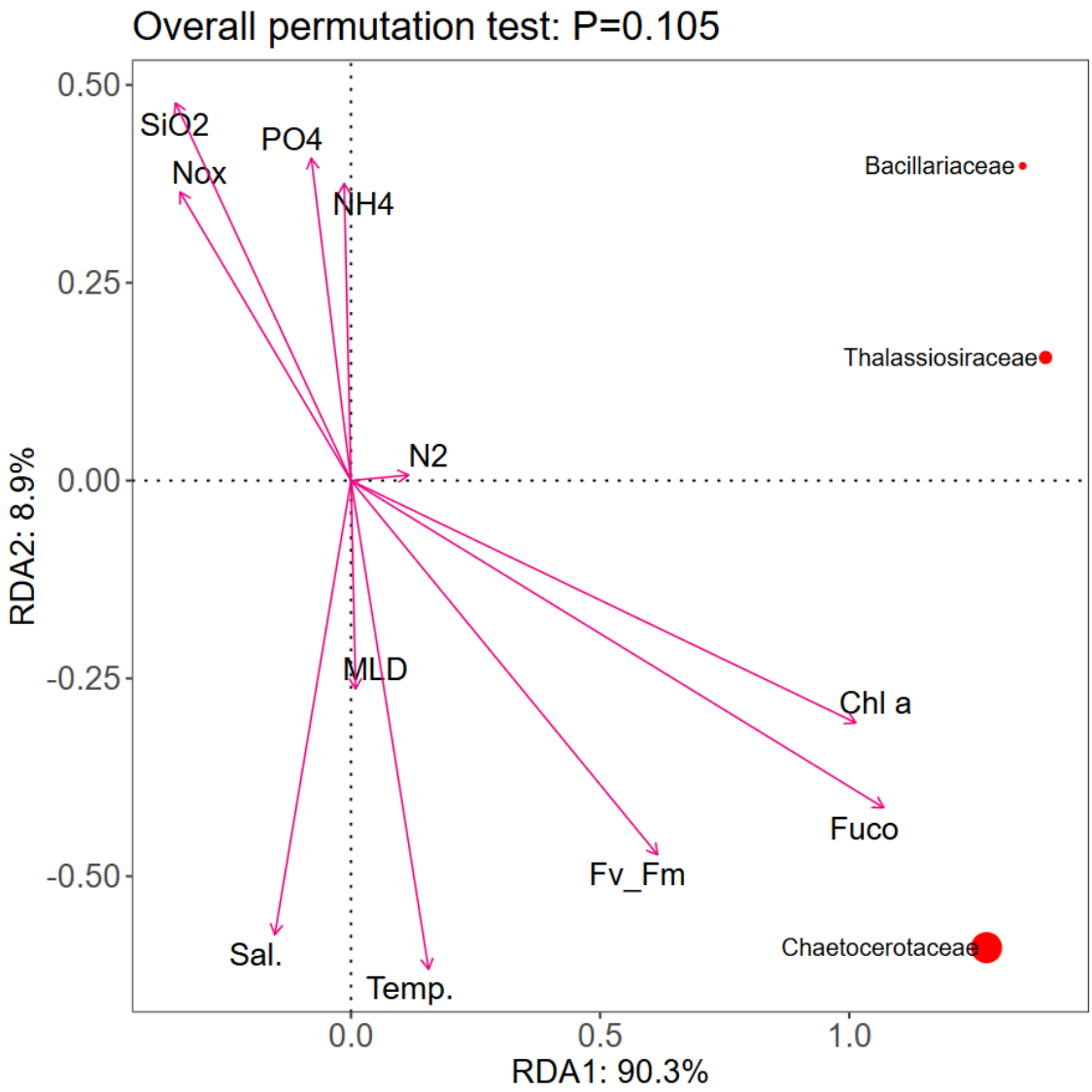


Figure S3. Redundancy analysis (RDA) of average *rbcL* gene transcription efficiency (cDNA/DNA) of three dominant diatom families with environmental parameters. Environmental variables are represented by red arrows. The dot size of diatom groups means their average transcription efficiency of *rbcL* gene.