



# *In Situ* Molecular Ecological Analyses Illuminate Distinct Factors Regulating Formation and Demise of a Harmful Dinoflagellate Bloom

 Liying Yu,<sup>a,b</sup>  Tangcheng Li,<sup>a</sup>  Hongfei Li,<sup>a</sup>  Minglei Ma,<sup>a</sup>  Ling Li,<sup>a</sup>  Senjie Lin<sup>a,c</sup>

<sup>a</sup>State Key Laboratory of Marine Environmental Science and College of Ocean and Earth Sciences, Xiamen University, Xiamen, China

<sup>b</sup>Central Laboratory, Second Affiliated Hospital of Fujian Medical University, Quanzhou, China

<sup>c</sup>Department of Marine Sciences, University of Connecticut, Groton, Connecticut, USA

**ABSTRACT** The development and demise of a harmful algal bloom (HAB) are generally regulated by multiple processes; identifying specific critical drivers for a specific bloom is important yet challenging. Here, we conducted a whole-assemblage molecular ecological study on a dinoflagellate bloom to address the hypothesis that energy and nutrient acquisition, defense against grazing and microbial attacks, and sexual reproduction are critical to the rise and demise of the bloom. Microscopic and molecular analyses identified the bloom-causing species as *Karenia longicanalis* and showed that the ciliate *Strombidinopsis* sp. was dominant in a nonbloom plankton community, whereas the diatom *Chaetoceros* sp. dominated the after-bloom community, along with remarkable shifts in the community structure for both eukaryotes and prokaryotes. Metatranscriptomic analysis indicated that heightened energy and nutrient acquisition in *K. longicanalis* significantly contributed to bloom development. In contrast, active grazing by the ciliate *Strombidinopsis* sp. and attacks by algicidal bacteria (*Rhodobacteraceae*, *Cryomorpaceae*, and *Rhodobacteraceae*) and viruses prevented (at nonbloom stage) or collapsed the bloom (in after-bloom stage). Additionally, nutrition competition by the *Chaetoceros* diatoms plausibly contributed to bloom demise. The findings suggest the importance of energy and nutrients in promoting this *K. longicanalis* bloom and the failure of antimicrobial defense and competition of diatoms as the major bloom suppressor and terminator. This study provides novel insights into bloom-regulating mechanisms and the first transcriptomic data set of *K. longicanalis*, which will be a valuable resource and essential foundation for further elucidation of bloom regulators of this and related species of Kareniaceae in the future.

**IMPORTANCE** HABs have increasingly occurred and impacted human health, aquatic ecosystems, and coastal economies. Despite great efforts, the factors that drive the development and termination of a bloom are poorly understood, largely due to inadequate *in situ* data about the physiology and metabolism of the causal species and the community. Using an integrative molecular ecological approach, we determined that heightened energy and nutrient acquisition promoted the bloom, while resource allocation in defense and failure to defend against grazing and microbial attacks likely prevented or terminated the bloom. Our findings reveal the differential roles of multiple abiotic and biotic environmental factors in driving the formation or demise of a toxic dinoflagellate bloom, suggesting the importance of a balanced biodiverse ecosystem in preventing a dinoflagellate bloom. The study also demonstrates the power of whole-assemblage metatranscriptomics coupled to DNA barcoding in illuminating plankton ecological processes and the underlying species and functional diversities.

**KEYWORDS** *Karenia longicanalis*, metatranscriptome, DNA barcoding, HAB, microeukaryotes, prokaryotic microbes, algae, dinoflagellate

**Editor** Eva C. Sonnenschein, Swansea University

**Copyright** © 2023 Yu et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Senjie Lin, [senjie.lin@uconn.edu](mailto:senjie.lin@uconn.edu).

The authors declare no conflict of interest.

**Received** 15 December 2022

**Accepted** 27 March 2023

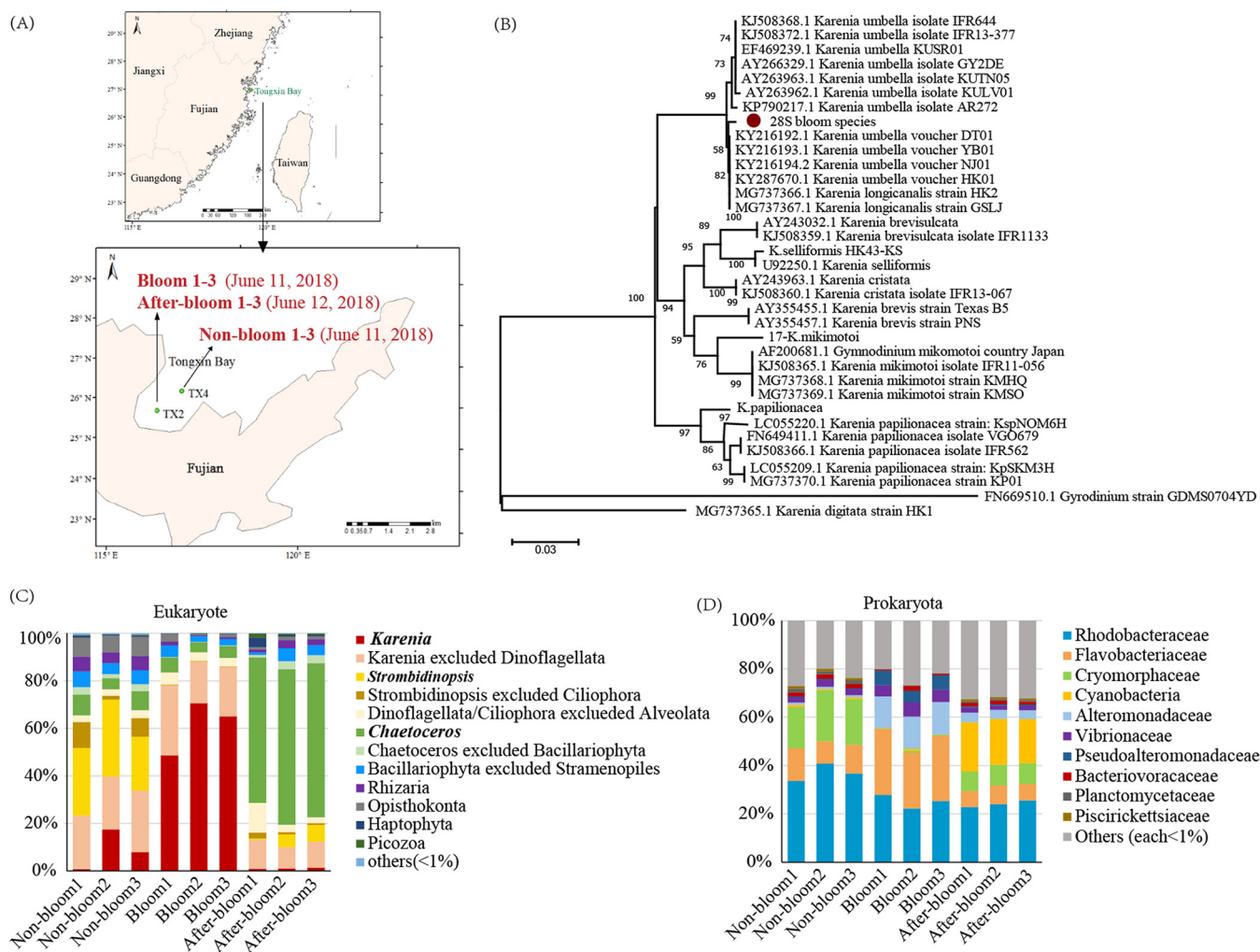
**Published** 19 April 2023

The aquatic ecosystem is threatened by harmful algal blooms (HABs), which have increased in recent decades in frequency, bloom species diversity, geographic extent, and severity of ecological and economic damages (1, 2). HABs result from the rapid growth of one or few microalgae due to natural processes (such as upwelling relaxation) or eutrophication from anthropogenic loadings (3, 4). However, algal bloom dynamics are regulated by cell division rate ( $\mu$ ), current-mediated cell-concentrating (import or I) or -dissipating (export or Ex) factors, grazing by animals (G), and mortality (M) due to environmental stress and microbial attacks (algicidal bacteria and viruses). These factors can be conceptually expressed as follows:  $dN/dt = \mu + I - Ex - G - M$ , where  $N$  and  $t$  are population abundance and time period, respectively. Some blooms are believed to be initiated from seed populations brought onshore by currents (5) or dissipated by rain and currents (6). Currents also can have important indirect effects by bringing in nutrients. However, most HABs result from the dynamics of  $\mu$ , G, and M, whereas I or Ex has minimal direct effects. To date, studies on the regulators of HABs have mainly focused on abiotic factors that regulate  $\mu$ , primarily water temperature and nutrient conditions (7–10). In contrast, biotic interactions reflected in G and M, which are also crucial regulators of bloom development and decline, have been less studied. Among the associated microbes, bacteria are a major constituent, and up to 20% of them are attached to algae or particles (11). The main effects of microbes (bacteria and viruses) on HABs include the lysis of algal cells (12–14) and inhibition of algal growth (15), although some bacteria can promote algal growth (16–18). It is highly challenging to determine which of the multiple growth-promoting variables (e.g., light, nutrients) and bloom-depressing variables (e.g., grazing, microbial attacks) are the drivers of bloom emergence and decline, because this requires simultaneous and *in situ* measurements at the species level. An integrative molecular ecological approach, however, makes such an endeavor possible, since it allows investigators to analyze field samples directly to reconstruct the metabolic profile of the bloom species *in situ*, in comparison to that of other organisms in the same habitat, in the course of bloom evolution.

Metabarcoding and metatranscriptome profiling with next-generation sequencing have proven to be effective ways to profile community composition and metabolic landscape (19, 20) and to detect major shifts in metabolic activities and resource utilization in the course of phytoplankton succession (21, 22). Metatranscriptomic studies have successfully tracked cellular processes associated with life history transitions, nutrient acquisition, defense, and energy production during dinoflagellate blooms by community gene expression analyses (23–30). Previous data also suggested that sexual reproduction (meiosis) could contribute to bloom formation for dinoflagellate blooms (25, 31). How grazing by zooplankton and attacks by microbes influence the bloom dynamics were not addressed in those studies.

Marine HAB causative species are predominantly dinoflagellates (32–36). Species of the genus *Karenia* are frequent bloom-forming dinoflagellates in coastal waters (37–40). Most *Karenia* species, such as *Karenia brevis* and *Karenia mikimotoi*, produce toxins that directly kill marine organisms, causing immense economic losses and environmental damages (19, 41). Extensive research has been conducted to investigate the ecological drivers (mainly nutrients and other environmental factors) of *Karenia* blooms (42–45). For example, a study suggested that N, P, and Si nutrients are important drivers of a *K. mikimotoi* bloom (44). Similarly, inorganic and organic N and P nutrients have been shown to be important to *K. brevis* bloom outbreaks (46). However, due to the complexity of interactions between the species (genetics) and environmental (abiotic and biotic) factors in each bloom, it is hard to generalize the findings, and each bloom's precise driver and terminator are still poorly understood.

In this study, we investigated a *Karenia* bloom that occurred in Tongxin Bay of Fujian, China, in June 2018 using an integrative molecular ecological approach. We collected samples from the bloom, nonbloom (as a surrogate of prebloom), and after-bloom conditions within the bay. Metabarcoding analyses based on small subunit (SSU) ribosomal DNA (rDNA; 16S and 18S) were performed to profile algal and bacterial community structures and their changes across the different bloom conditions. In parallel, whole-assemblage

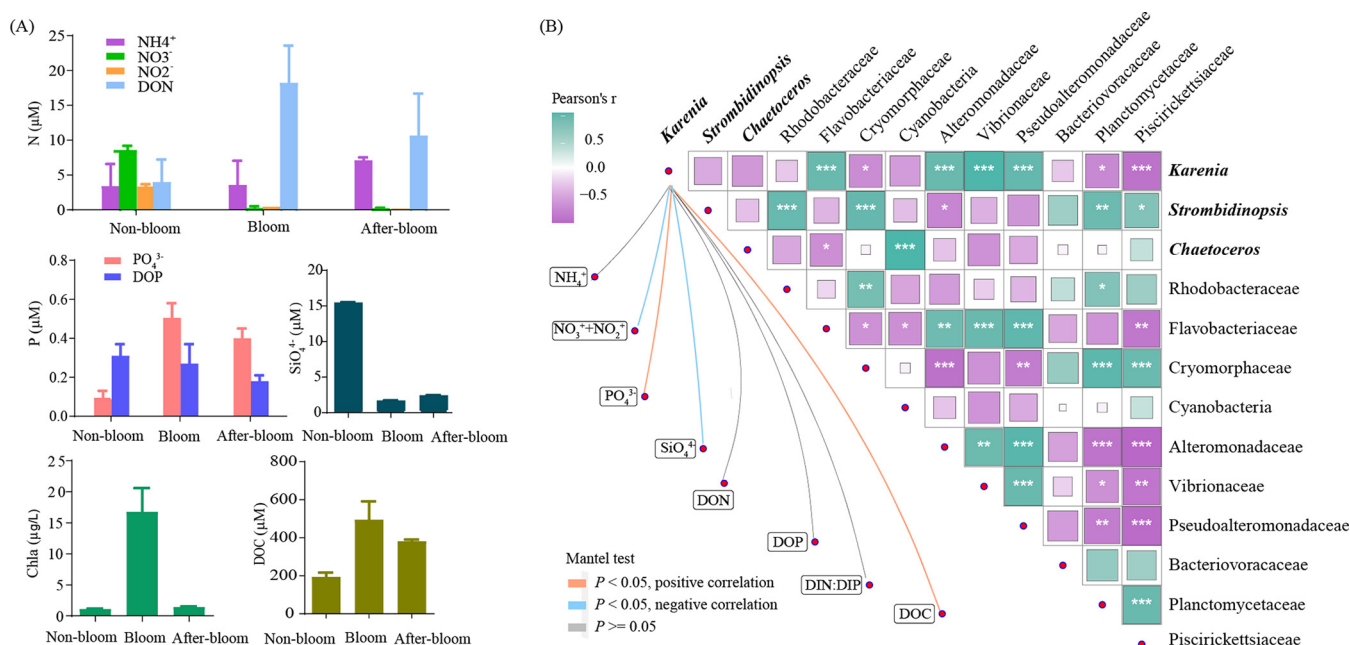


**FIG 1** Study site and bloom community profile. (A) Map of study site and sampling stations and dates. (B) Maximum-likelihood phylogenetic analysis of 28S rDNA sequences showing that the bloom species is *Karenia longicanalis* (= *K. umbella*). Scale bar represents number of substitutions per site. (C) Eukaryotic plankton community composition showing *Karenia* bloom (red) development and demise. (D) Prokaryotic plankton community composition showing shifts of dominant bacterial lineages throughout different stages of the bloom.

metatranscriptomic analyses were carried out to identify critical metabolic processes associated with the different bloom conditions. This research portrays prokaryotic and eukaryotic community shifts among the bloom, nonbloom, and after-bloom stages and sheds light on what main biotic factors and metabolic processes associated with energy and nutrient acquisition, cell proliferation, and defense drive the bloom development and demise in a seriously understudied dinoflagellate.

## RESULTS

**Bloom profile and plankton community shift through bloom stages.** At the community level, the chlorophyll *a* concentration was 16.8  $\mu\text{g/liter}$  in the bloom samples, compared with 1.1  $\mu\text{g/liter}$  in the nonbloom samples and 1.5  $\mu\text{g/liter}$  in the after-bloom samples (Fig. 1A), indicating a  $>10$ -fold-higher phytoplankton community biomass in the bloom than the nonbloom and after-bloom conditions, the latter two of which showed similarly low phytoplankton community biomass. Microscopic observations and molecular analysis (28S rDNA) indicated that the bloom species was *K. longicanalis* (formerly *K. umbella* [47]) (Fig. 1B), and its abundance was  $1.51 \times 10^7$  cells/liter in the bloom samples, compared to  $2.4 \times 10^4$  cells/liter (a 600-fold-lower concentration) in the nonbloom samples, and it declined by 137-fold ( $1.1 \times 10^5$  cells/liter) in the after-bloom samples. Based on 18S rDNA metabarcoding of the community (Fig. 1C),



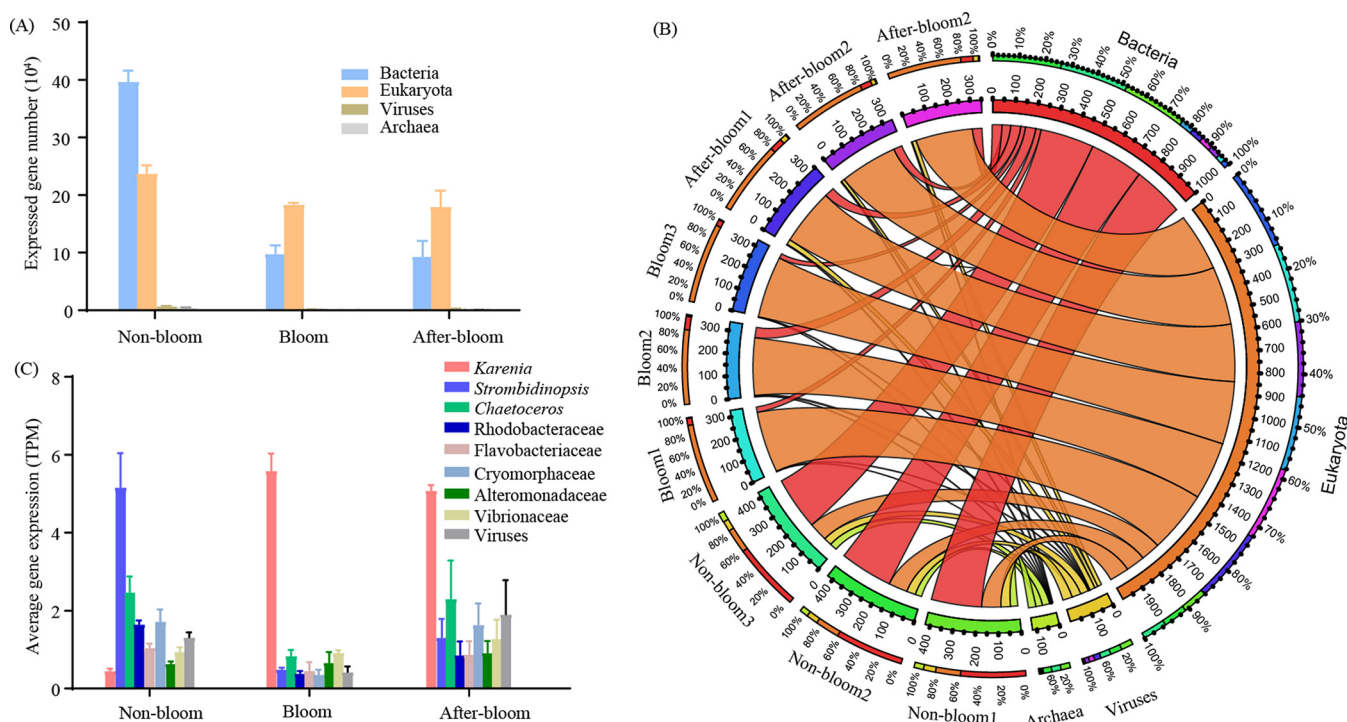
**FIG 2** Quantitative relationships of lineages within the community with nutrient variables. (A) Changes of nutrient concentrations, chlorophyll, and dissolved organic carbon concentrations with bloom conditions. (B) Correlations of dominant lineages with each other and with nutrient conditions. Correlations between lineages (filled squares) were analyzed using the Pearson test. Deeper color and larger squares mark stronger correlations. \*,  $P = 0.01$  to  $0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ . Correlations between *Karenia* and nutrient factors (connecting lines) were analyzed using the Mantel test. The orange and blue connecting lines mark positive and negative correlations between the relative abundance levels of *Karenia* and nutrient concentrations.

*Karenia* sp. accounted for 8.71%, 61.42%, and 1.09% in the nonbloom, bloom, and after-bloom groups of samples, respectively.

The 18S rDNA metabarcoding analysis also showed that the nonbloom plankton community was dominated by the ciliate *Strombidinopsis* sp. (accounting for 27.87%), and the after-bloom community was dominated by the diatom *Chaetoceros* sp. (accounting for 63.86%) (Fig. 1C). Other major eukaryotic phyla included Opisthokonta (0.58% to 8.20%), Rhizaria (0.36% to 6.25%), Haptophyta (0.16% to 3.94%), and *Picozoa* (0.10% to 1.75%) throughout this study period (Fig. 1C).

To explore what algae-associated microbes might be involved in the bloom development and decline, the microbial community was analyzed using 16S rDNA metabarcoding. Ten dominant families of bacteria were identified (Fig. 1D): Rhodobacteraceae (22.29% to 40.81%), Flavobacteriaceae (6.72% to 27.28%), Cryomorpaceae (0.34% to 21.22%), Cyanobacteria (0.04% to 20.40%), Alteromonadaceae (8.71% to 13.54%), Vibrionaceae (1.72% to 5.86%), Pseudoalteromonadaceae (0.31% to 6.10%), Bacteriovoraceae (0.43% to 2.05%), Planctomycetaceae (0.06% to 1.47%), and Rickettsiaceae (0.01% to 1.17%). Rhodobacteraceae were consistently abundant across sample groups. Flavobacteriaceae, Alteromonadaceae, Vibrionaceae, and Pseudoalteromonadaceae had a higher relative abundance in the bloom sample group than in the nonbloom and after-bloom sample groups. Cryomorpaceae showed the highest relative abundance in the nonbloom sample group, whereas Cyanobacteria were most abundant in the after-bloom sample group.

**Environmental and biotic factors showing correlations with the abundance of bloom species.** The relative abundance of *Karenia* (based on 18S metacoding) was negatively correlated with  $\text{NO}_3^-$ ,  $\text{NO}_2^-$ , and  $\text{SiO}_4^{3-}$  but positively correlated with  $\text{PO}_4^{3-}$  and dissolved organic carbon (DOC) concentrations (Fig. 2B). Besides, the *Karenia*-dominated bloom samples displayed the highest dissolved oxygen, pH, and surface irradiance in the environment (see Table S1 in the supplemental material), indicating active photosynthesis. The Pearson test showed that the relative abundance of *Karenia* was positively correlated with the abundances of bacteria from the families Flavobacteriaceae, Alteromonadaceae, Vibrionaceae, and Pseudoalteromonadaceae but negatively correlated

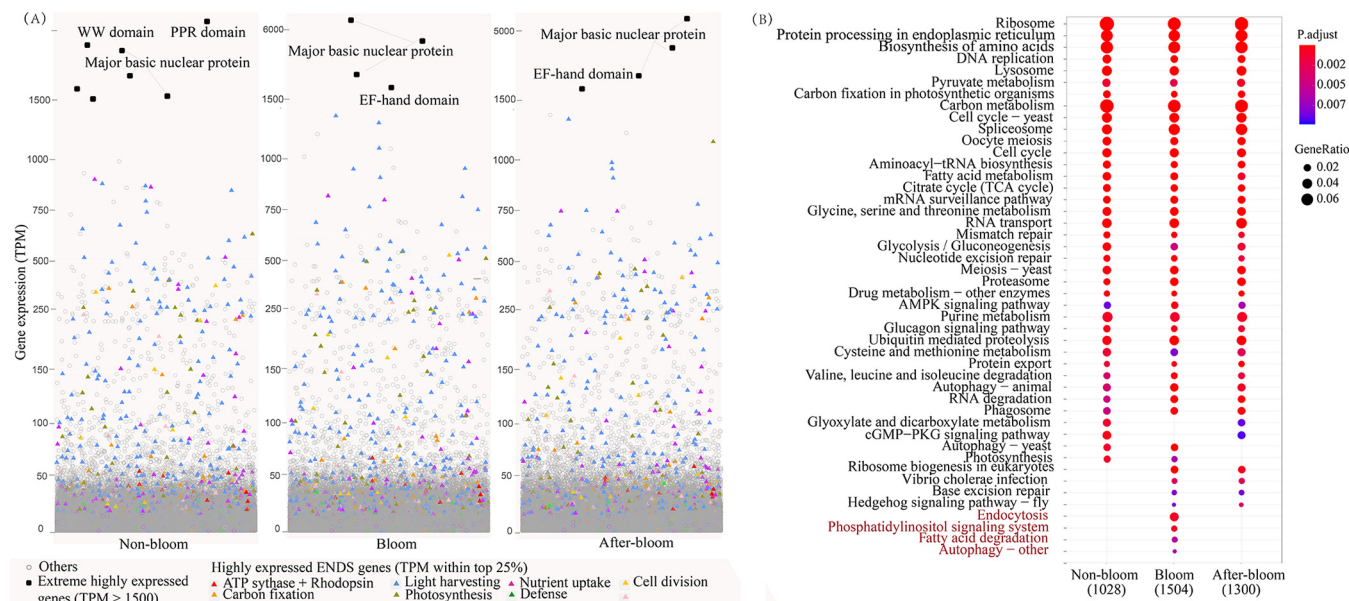


**FIG 3** Community metabolic profiles based on community TPM in the three bloom stages. (A) Expressed gene diversity in the four kingdoms. The average gene diversity of three biological replicates is displayed for each group. (B) Circos plot summarizing gene expression (TPM) contribution of the four kingdoms to each sample. The different colors of the inner ring mark total TPM of different samples or kingdoms. The outer ring shows the gene expression percentages of kingdoms in samples or gene expression percentages of samples in kingdoms, with colors corresponding to the inner ring. Wider connecting ribbons between kingdom and sample indicate a larger gene expression contribution of this kingdom in the sample. (C) Average gene expression level (total TPM divided by number of genes with read count of >0) in the dominant lineages.

with that of Cryomorpaceae, Planctomycetaceae, and Piscirikettsiaceae (Fig. 2B), suggesting potential promoting and depressing effects on the bloom, respectively.

**Community metatranscriptomic profile and its shift through bloom stages.** Our whole-assembly metatranscriptome sequencing yielded a total of over 1.15 billion cDNA reads (Table S2). After quality filtering, the resulting 0.91 billion (79.11%) clean reads were assembled into a reference metatranscriptome composed of 1,389,260 unigenes with an  $N_{50}$  of 868 bp and a maximum length of 131,207 bp (Table S3). The reference metatranscriptome showed a total read mapping rate of 77.63 to 90.88% (Table S4), indicating an overall good assembly quality. Functional annotation based on the eggNOG, Swiss-Prot, KEGG, GO, and nrMegPhylodb databases gave a total annotation rate (genes annotated by at least one of these databases) of 74.87% (Table S5).

Among the 1,389,260 unigenes, 74.32% were assigned taxa, of which 58.20% were from bacteria, 40.04% from eukaryotes, 1.08% from viruses, and 0.69% from archaea. The number of expressed genes and the expression levels of the genes from bacteria, viruses, and archaea decreased markedly from the nonbloom samples to the bloom samples (Fig. 3A). Gene expression (transcript per million mapped reads [TPM]) of the bloom and the after-bloom assemblages were dominated by eukaryotes (about 80%), whereas that of the nonbloom samples were dominated by bacteria (~60%) followed by ~20% eukaryotes (Fig. 3B). At the family or genus level, however, the bloom metatranscriptomes were dominantly contributed by *Karenia*, while the nonbloom metatranscriptomes were contributed mainly by the ciliate *Strombidinopsis*, the diatom *Chaetoceros*, the bacteria *Cryomorpaceae* and *Rhodobacteraceae*, and viruses, in decreasing order. The after-bloom metatranscriptomes were mainly composed of transcripts from *Karenia*, *Chaetoceros*, viruses, *Cryomorpaceae*, and *Strombidinopsis* (Fig. 3C). The gene expression levels of *Strombidinopsis* in the bloom sample group were >20-fold lower than in the nonbloom sample group. Moreover, *Chaetoceros*, *Rhodobacteraceae*, *Cryomorpaceae*, and virus gene



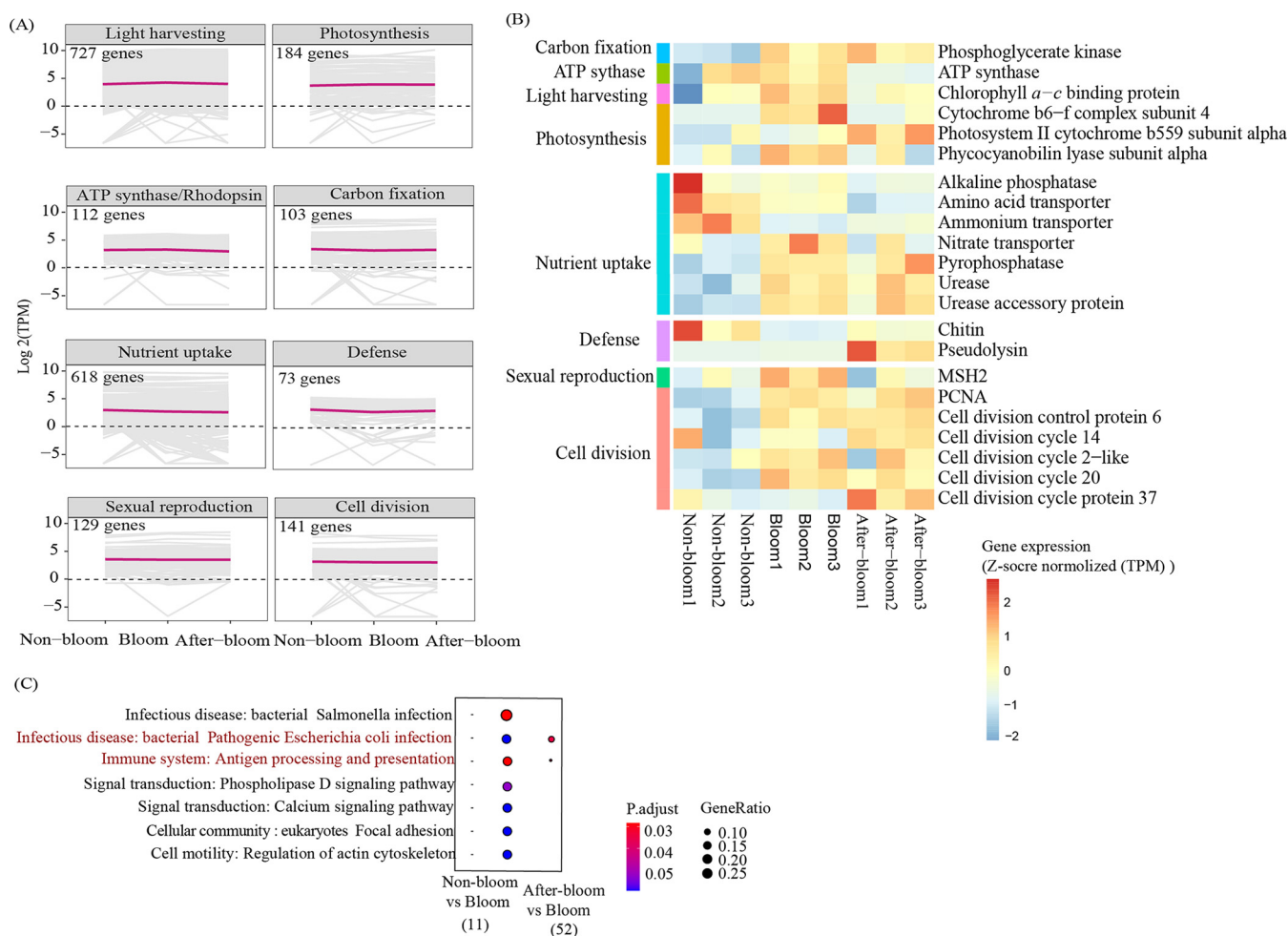
**FIG 4** Abundantly expressed genes in *Karenia* under the three bloom conditions (nonbloom, bloom, and after-bloom). (A) Extreme highly expressed genes (TPM of  $>1,500$ ; black dot) and highly expressed (top 25% expressed genes; colored triangles) ENDS genes in *Karenia*. (B) KEGG enrichment of highly expressed genes (top 25% expressed genes). The pathway in red font represents those with DEGs uniquely enriched under the bloom condition.

expression levels were much higher ( $>3$ -fold) in the nonbloom samples than in the bloom samples.

***K. longicanalis* transcriptome and bloom-associated differential expression of genes related to energy and nutrient acquisition, defense, and cell division cycle (ENDS).** As this is the first transcriptome ever reported for *K. longicanalis*, particularly during a natural bloom, it is of interest to catalog the expressed genes. From the meta-transcriptomes, 94,785 unigenes were identified as *Karenia* genes. Four genes were expressed constantly at  $>1,500$  TPM across all sample groups, including three major basic nuclear protein genes and one EF-hand domain-containing protein gene (Fig. 4A). Using a TPM of the  $>75\%$  quartile as the criterion of highly expressed genes (HEGs), we found 9,021, 17,533, and 13,623 HEGs of *Karenia* in the nonbloom, bloom, and after-bloom samples, respectively. KEGG enrichment of the HEGs revealed that 33 pathways were commonly enriched in all three sample groups, such as the ribosome, protein processing in the endoplasmic reticulum, and biosynthesis of amino acids (Fig. 4B). Besides, we found that HEGs of endocytosis, phosphatidylinositol signaling system, fatty acid degradation, and autophagy, representing pathways of mixotrophic uptake and recycling of nutrients, energy metabolism, and signaling, were uniquely enriched in the bloom sample group (Fig. 4B).

To examine how the capacity of ENDS genes influenced the bloom dynamics, we documented the features of related genes or pathways associated with the bloom. In total, we detected 1,126 energy acquisition genes (E) involved in light harvesting, photosynthesis, proton pump rhodopsin, and carbon fixation and genes of ATP synthase; 618 nutrient uptake genes (N), such as N-, P-, and Si-related transporters; 73 defense genes (D); and 270 cell proliferation genes comprising those regulating sexual reproduction (meiosis) and cell division cycle (S) in *Karenia* (Fig. 5A; Fig. S1). These ENDS genes were relatively stable across the three samples (Fig. 5A). Among the *Karenia* 2,087 ENDS genes identified, 891 (42.69%), 1,001 (47.96%), and 897 (42.98%) were highly expressed in nonbloom, bloom, and after-bloom samples, at higher percentages than the overall percentage of HEGs (25%) (Fig. 4A).

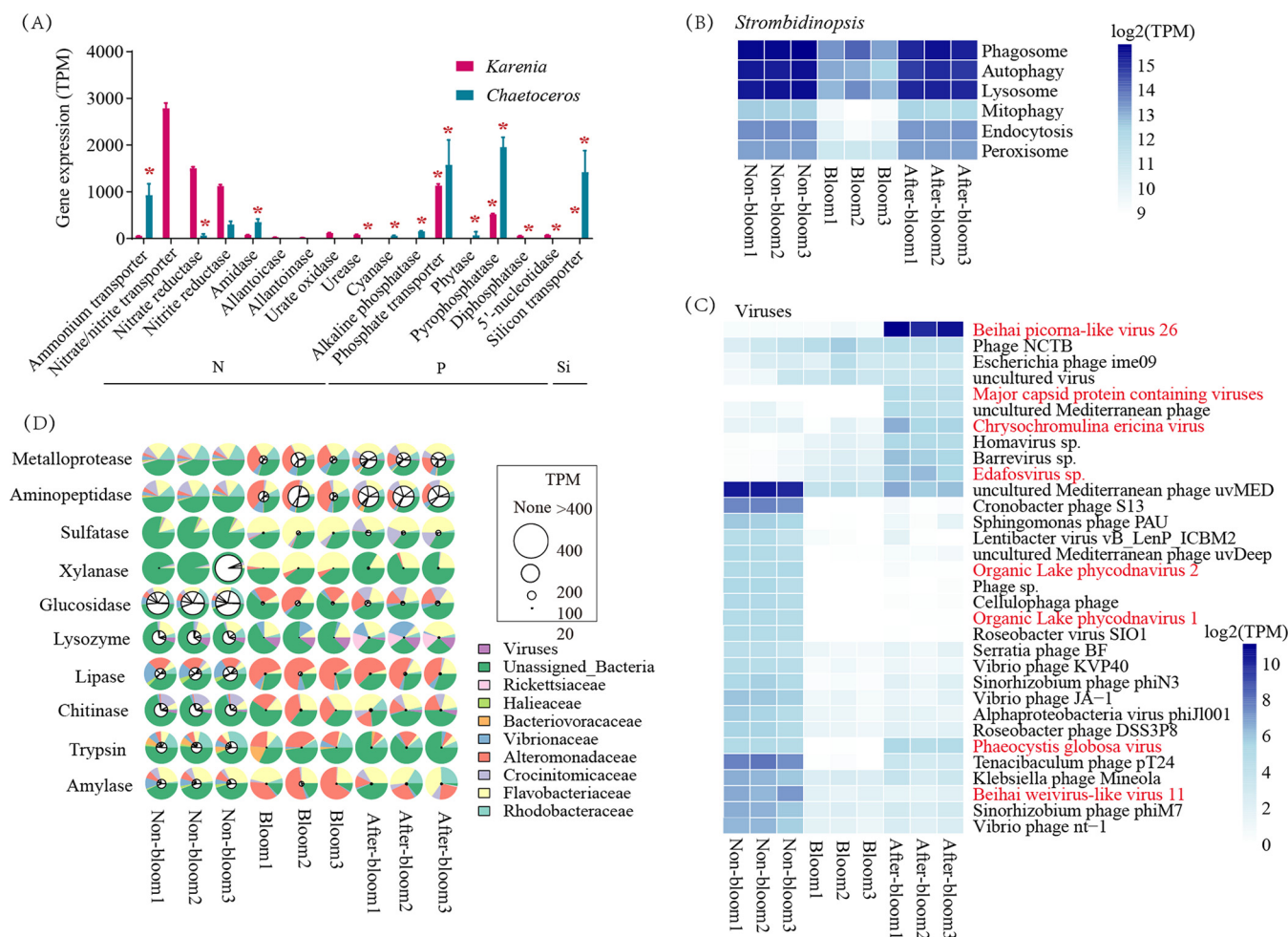
Despite the overall stable high expression of ENDS genes, we observed some slight changes. There were 22 differentially expressed ENDS genes (Fig. 5B). Among them, most of the ENS genes were upregulated, but defense genes were downregulated in the bloom samples compared to the nonbloom or after-bloom samples. Energy metabolism-related



**FIG 5** Energy and nutrient acquisition, defense, and cell reproduction (ENDS)-related genes and differentially expressed genes in *Karenia* under the three bloom conditions (nonbloom, bloom, and after-bloom). (A) Expression patterns of ENDS genes in *Karenia*. (B) Differentially expressed ENDS genes in the bloom stage compared to nonbloom and after-bloom stages. Differentially expressed genes were defined as those with  $P$  values of  $<0.05$  using the  $t$  test. (C) KEGG enrichment of differentially expressed genes (adjusted  $P$  value of  $<0.05$  and fold change of  $\geq 2$ ) in the bloom stage compared to nonbloom and after-bloom stages. The pathway in red font represents common enriched pathways in both bloom versus nonbloom and bloom versus after-bloom comparisons by DEGs upregulated in the bloom condition.

differentially expressed genes (DEGs) that were upregulated during bloom included ATP synthase and genes involved in light harvesting (chlorophyll  $a$  to  $c$  binding protein gene family), photosynthesis, and carbon fixation. Nutrient uptake-related upregulated DEGs included amino acid transporter, nitrate transporter, pyrophosphatase, urease and urease accessory protein. Related to cell proliferation were DEGs of the cell cycle proteins and MSH2, which were upregulated in the bloom (Fig. 5B). Regarding defense (D), the chitin and pseudolysin genes annotated as encoding antimicrobial peptides were significantly downregulated in bloom samples relative to the nonbloom or after-bloom sample groups (Fig. 5B). Additionally, we detected 3,627 total DEGs between the bloom and nonbloom sample groups, but only 257 DEGs between the bloom and after-bloom sample groups, indicating more similar transcriptomic status between the bloom and after-bloom stages. KEGG enrichment results showed that the upregulated DEGs in the nonbloom samples were enriched in defense-related pathways against two bacteria-mediated infectious diseases and a pathway related to the immune system. Two of the three defense-related pathways were also upregulated in the after-bloom samples (Fig. 5C).

**Differential nutrient uptake strategies between *Karenia* bloom and *Chaetoceros* dominance.** To explore what molecular factors drove the transition from the *Karenia* bloom to *Chaetoceros* sp. dominance in the after-bloom stage, we identified HEGs when one of the two species was dominant and DEGs across three sample groups



**FIG 6** Expression of genes in the dominant species of phytoplankton, ciliate *Strombidinopsis*, viruses, and prokaryotes at the three bloom stages. (A) Highly expressed genes (top 25% expressed genes) related to nitrogen (N), phosphate (P), and silicon (Si) uptake in *Karella* (at the bloom) and *Chaetoceros* (at the after-bloom) when one of the two species was dominant. For each gene family, the summed gene expression (TPM) is displayed. Asterisks mark upregulated genes in the bloom stage relative to nonbloom or after-bloom in *Karella*, or upregulated genes in the after-bloom stage relative to the bloom or nonbloom sample groups in *Chaetoceros* ( $n = 3$ ). (B) Phagotrophy and related pathways in *Strombidinopsis*. (C) Top 50 highly expressed viral genes. The viruses in red font represent algae-infecting viruses. (D) Expression of highly expressed microbial genes (TPM, >100) related to cell wall lysis. Larger inner white circles mark higher expression levels of genes. The outer colorful circles represent the gene expression contribution of each species.

related to N, P, and Si uptake in *Karella* and *Chaetoceros* species. *Chaetoceros* ammonium transporters were the most highly expressed N-related HEGs when the species was dominant (the after-bloom), followed by amidase, with both being upregulated in the after-bloom samples compared to the bloom or nonbloom samples, while nitrate-nitrite transporter genes and nitrate-nitrite reductase genes were the top two most highly expressed N-related genes in *Karella* during its bloom (Fig. 6A). Furthermore, a diversity of dissolved organic nitrogen (DON)-related HEGs were detected in *Karella* at the bloom stage but not in *Chaetoceros* during after-bloom, including allantoicase, allantoinase, amidase, urate oxidase, and urease (Fig. 6A). For P-related genes, both *Chaetoceros* and *Karella* phosphate transporter and pyrophosphatase genes were not only among the top two highly expressed but also were upregulated (HEGs and DEGs) under their respective dominance conditions. *Karella* also showed high expression of diphosphatase and 5'-nucleotidase in the bloom sample. In addition, *Chaetoceros* silicon transporters were more highly expressed and upregulated in the after-bloom condition than in the bloom and nonbloom conditions (Fig. 6A). These findings indicated different nutrient uptake strategies between *Karella* and *Chaetoceros*, which were potentially responsible for the transition from the dinoflagellate bloom to the diatom dominance.



**Transcriptomic signals of active grazing and microbial attacks in nonbloom and after-bloom stages.** The upregulated defense genes in the nonbloom and after-bloom samples prompted us to interrogate whether metabolic pathways related to grazing and microbial attack were correspondingly upregulated. We first examined genes controlling transport and grazing activities in the ciliate *Strombidinopsis*, including genes in the lysosome, phagosome, autophagy, endocytosis, and mitophagy pathways. Results showed that these pathways were indeed expressed at much higher levels in the nonbloom and after-bloom than in the bloom stage (Fig. 6B), involving 261 genes related to endocytosis, phagocytosis, or phagosome (Table S6). These included highly expressed genes encoding 1-phosphatidylinositol-4-phosphate 5-kinase (PIP5K), ADP-ribosylation factor 1 (ARF1), cathepsin L, actin beta/gamma 1, cofilin (CFL), and protein transport protein SEC61 subunit alpha (SEC61A).

We also explored whether viral infection was involved in bloom prevention and collapse and indeed detected viral mRNA in our metatranscriptomic data. The top 50 most highly expressed viral genes belonged to 33 functional types. Twenty-three of these, including four algal viruses (annotated as organic lake phycodnavirus 1 and 2, *Phaeocystis globosa* virus, and Beihai weivirus-like virus 11), were most highly expressed in the nonbloom samples. Ten of these showed the highest expression in the after-bloom stage, including Beihai picorna-like virus, major capsid protein-containing viruses, *Chrysochromulina ericina* virus, and *Edafosvirus* sp., four types of algal viruses different from those in the nonbloom stage (Fig. 6C).

Bacteria are another potential bloom killer. Highly expressed genes (TPM, >100 in at least one sample) from prokaryotes that encode enzymes potentially involved in algal cell wall lysis were detected, including genes with expression (listed from high to low expression): aminopeptidase, metalloprotease, sulfatase, xylanase, glucosidase, lysozyme, lipase, chitinase, trypsin, and amylase (Fig. 6D). These 10 bacterial enzymes showed the highest expression in the nonbloom samples, followed by the after-bloom samples. Aminopeptidase, metalloprotease, glucosidase, and trypsin were mainly contributed by Rhodobacteraceae and unassigned bacteria. Sulfatase, xylanase, lysozyme, and amylase were mainly contributed by Flavobacteriaceae. Lipase was dominantly contributed by Alteromonadaceae. Chitinase was dominantly from Crocinitomicaceae. The cytolytic capacities of bacteria likely contributed to the convergent *Karenia* cell-removing activities, which led to the prevention of bloom formation in the nonbloom stage and the destruction of the bloom in the after-bloom stage. However, the current data did not allow us to exclude the possibility that the attack of grazers (ciliates) and viruses might have led to the release of undigested algal remnants in the nonbloom or after-bloom samples, and bacteria might have used the same enzymes (as reported above) to utilize this rich food source and support their growth.

## DISCUSSION

Multiple processes regulate the dynamics of HABs, including  $\mu$ , I, Ex, G, and M. Determining which of these most critically influences a specific bloom is crucial for understanding and managing HABs but is a daunting challenge because it requires simultaneous *in situ* measurements of these terms. However, integrative ecological and molecular analyses provide a possibility to approach the problem. Although transcriptomes are indirect indications of *in situ* physiological and ecological dynamics, they are still powerful evidence of these processes. By taking the molecular ecological approach, clues and insights about the most important bottom-up (E, N, D, S) and the top-down (G, M) factors can be interrogated (29), where a strong signal of D in the bloom species corresponds to high activity of G from grazers and/or M from microbes. In the present study, we conducted a study on a *K. longicanalis* bloom using this approach. Based on a shipboard field campaign, triplicated samples were collected from nonbloom, bloom, and after-bloom assemblages and processed immediately on site to capture the *in situ* ecological status and species-specific metabolic profiles. The resulting metatranscriptomes provided the first functional gene catalog of *K. longicanalis*, enriching the existing

dinoflagellate genomic resource. The data set also revealed the major basic nuclear protein (MBNP) and growth- and development-associated EF-hand domain-containing protein (48, 49) as highly expressed genes in the bloom-causing dinoflagellate. MBNP in dinoflagellates is located at the periphery of the permanently condensed chromosomes and the nucleolar organizing region, and although its role is unclear (50), it has been consistently reported as one of the most highly expressed genes in dinoflagellates (51), signaling its functional importance. In addition, we simultaneously documented eukaryotic and prokaryotic DNA barcodes (SSU) and their mRNAs under different bloom conditions. With both eukaryotic and prokaryotic plankton covered, the metabarcoding and metatranscriptomic data sets represent the most comprehensive data sets for a bloom study to date, and it enabled a deep inquiry into the complex drivers that regulated bloom growth and demise.

**Strong energy and nutrient acquisition as a major driver of the *K. longicanalis* bloom.** Despite the overall stable transcriptional expression of ENDS genes, probably because most dinoflagellate genes are posttranscriptionally regulated (51), a few genes were remarkably regulated in our transcriptomic data set. The acquisition of energy and nutrients plays a fundamental role in algal growth to form and maintain an algal bloom (52, 53). Under bloom conditions, *Prorocentrum shikokuense* (formerly *P. donghaiense*) is advantaged in energy acquisition with high expression levels of genes encoding proton pump rhodopsin, nutrient transporters, and enzymes involved in the utilization of organophosphates (27, 29). The present study found that many genes related to energy acquisition, including photosynthesis and carbon fixation, were significantly upregulated under the bloom condition, generally in congruence with the bloom of *P. shikokuense* (27, 29). The HEGs were uniquely enriched in the fatty acid degradation pathway, suggesting active catabolism of fatty acids to generate energy to support bloom growth. Furthermore, the ATP synthase gene of *Karenia* was also remarkably upregulated under the bloom condition. All these suggest energy acquisition was active during the bloom.

The strongest expression of nitrate-nitrite transporters and reductase genes represented in the N-related HEGs is suggestive of active utilization by the abundant cells of the bloom species. Indeed, the ambient dissolved inorganic nitrogen (DIN;  $\text{NH}_4^+$ ,  $\text{NO}_2^-$ , and  $\text{NO}_3^-$ ) was depleted under the bloom condition ( $<5 \mu\text{M}$ ) relative to the nonbloom and after-bloom conditions (Fig. 2A), evidently as a result of rapid consumption. Despite the DIN depletion, DON was copious ( $>15 \mu\text{M}$ ) during the bloom, probably derived from unconsumed aquaculture feed in the intense aquaculture operation (54). In support of this proposed source of DON, we found that DOC was also most abundant at the bloom stage. Likely, this source of DON played an important role in supporting *Karenia* bloom growth, because our metatranscriptomic data indicated that *Karenia* was actively scavenging DON during the bloom. These bloom-active DON genes included urate oxidase, which catalyzes the oxidative decarboxylation of uric acid to allantoin, which is essential in nitrogen assimilation in many plants (55). Allantoicase, functioning in ureide catabolism to generate a nitrogen source (56), was found to be inducible in the green alga *Chlamydomonas reinhardtii* when taking up ureide (57). Allantoinase, an enzyme that catalyzes the conversion of allantoin into ureidoglycolate and urea (58), was shown to increase in *C. reinhardtii* when it was using allantoin (59). Urease catalyzes urea degradation to  $\text{NH}_4^+$  and  $\text{CO}_2$  to support the growth of many algae on urea as the nitrogen source (60). In addition, amidase, which catalyzes the hydrolysis of amides to ammonia and free carboxylic acids (61), was also highly expressed in *K. longicanalis* during the bloom. Moreover, there was strong and elevated expression of amino acid transporters, suggesting the potential that *K. longicanalis* was actively utilizing amino acids during the bloom. In addition, the HEGs found under the bloom condition were enriched in the endocytosis pathway, suggesting potential involvement of mixotrophy in nutrient acquisition for bloom development.

Similar to N nutrition, the high expression and/or upregulation of phosphate transporter and phosphatases in the bloom stage suggest active inorganic P uptake and organic P scavenging in *Karenia* under bloom conditions. Among the phosphatases highly

expressed in *K. longicanalis* were pyrophosphatase (upregulated), diphosphatase, and 5'-nucleotidase (5NT). Pyrophosphatase, which hydrolyzes the high-energy pyrophosphate into two inorganic phosphate molecules (62), has been reported in the dinoflagellate *Alexandrium tamarense* to hydrolyze phosphodiester for P nutrient (63). Diphosphatase catalyzes diphosphate into phosphates (64). 5NT has been linked to ATP utilization in *K. mikimotoi* (65). All the N and P nutrition results indicated that the bloom species was exploiting all sources of nutrients, including DIN, DON, DIP, DOP, and possibly phagotrophy during the bloom.

Active energy and nutrient acquisition is expected to promote cell proliferation via sexual and asexual reproduction. Previous molecular studies on dinoflagellate blooms have also revealed increases in metabolic activities of energy production, carbon metabolism, transport, and synthesis of cellular membrane components (an indicator of cell division) as potential drivers of blooms (23, 24, 26, 28). As expected, we detected significantly elevated expression of genes coding for cell division proteins. As none of the canonical meiosis core genes was found upregulated in the bloom reported here, contrary to the *K. mikimotoi* and other dinoflagellate blooms (25), the role of sexual reproduction in this *K. longicanalis* bloom was likely weak, although it is not impossible that a significant role might have escaped our detection. Nevertheless, our data provide strong evidence that energy and nutrient acquisition was the major metabolic driver of the bloom of *K. longicanalis*.

**Microbial attack, ciliate grazing, and diatom competition linked to *K. longicanalis* bloom prevention and termination.** Microbial attacks and zooplankton grazing are two major top-down controls of a phytoplankton bloom that can prevent or terminate a bloom, and for a bloom to occur the species needs to have strong defense against these cell-damaging and removing processes. The higher expression of defense mechanisms in *K. longicanalis* in the nonbloom sample group suggests that a more vigorous microbial attack might have diverted intracellular resources to defense at the cost of cell proliferation. Conversely, the considerable decreases in diversity and activity of viral genes under the bloom conditions suggest a lower infection activity during the bloom. Several viruses were transcriptionally very active under nonbloom or after-bloom conditions. Picorna-like viruses are positive-sense single-stranded RNA viruses that are major pathogens of plants, animals, insects, and marine phytoplankton (66, 67). They are widely distributed in marine water (68, 69) and have been reported to infect the bloom species *Heterosigma akashiwo* (66, 70). Mimivirus edafos virus is a type of nucleocytoplasmic large DNA virus (NCDLV) (71) that can infect diverse eukaryotes, such as protists and algae (72, 73). The *TARA Oceans* expedition data showed that NCDLVs outnumbered eukaryotic organisms in the photic layer (74). *Chrysochromulina ericina* viruses and phycodnaviruses are algae-infecting large DNA viruses widely distributed in aquatic environments (75–77). *Phaeocystis globaosa* virus (78), Beihai weivirus-like virus 11 (79), and many major capsid protein-containing viruses have been reported to be microalga-infecting viruses (80, 81). With the highest transcriptional activities under the after-bloom or nonbloom condition, picorna-like viruses, edafos virus, and other detected algal viruses might play an important role in *K. longicanalis*' decline. Further laboratory experiments are needed to demonstrate the infection of these viruses in *K. longicanalis*.

Bacterial genes encoding cytolytic enzymes were also actively expressed in the nonbloom and after-bloom stages (Fig. 6D), suggesting their potential degradation of *K. longicanalis* cell walls. These enzymes have been reported to function in lysing or inhibiting algal cells. Extracellular enzymes, including aminopeptidase and lipase produced by *Cytophaga* sp., have been linked to *Alexandrium catenella* cell lysis (82). Chitinase, sulfatase, and trypsin can inhibit the growth of various *Chlorella* strains by digesting the cell walls (83). Amylase from marine bacteria can be used to saccharify marine microalgae (84). Glucosidase and lysozyme can hydrolyze the cell walls of *C. vulgaris* (85). Xylanase (86, 87) and metalloproteases (88) produced by bacteria have tremendous potential to degrade cell walls. Rhodobacteracea was the most abundant bacteria family with the highest gene expression in the nonbloom samples, and its abundance was negatively correlated with the abundance of *K. longicanalis*. Coincidentally, a marine Rhodobac-

teraceae bacterium strain has been reported to exact an 87% inhibitory rate on *K. mikimotoi* (89), a sister species of *K. longicanalis*. The highest expression of aminopeptidase, metalloprotease, glucosidase, and trypsin in Rhodobacteracea bacteria further supported the bacteria's role in inhibiting *K. longicanalis* growth. In addition, Flavobacteriaceae and Vibrionaceae both had the highest gene expression activities in the nonbloom samples. Many Flavobacteriaceae bacteria have been reported to have algicidal activity against algae, including *K. mikimotoi* (90, 91) and *K. brevis* (92). The highest expression of sulfatase, xylanase, lysozyme, and amylase in Flavobacteriaceae in the nonbloom samples provided strong evidence for the ability of Flavobacteriaceae to disrupt algal cell walls. Vibrionaceae bacteria, especially *Vibrio*, have been reported to have algicidal activity for *K. mikimotoi* (93), *A. tamarense* (94), and *Akashiwo sanguinea* (95). This bacterial lineage probably also played an active role in inhibiting the growth of *K. longicanalis*. All these microbes showed the highest cell wall lysis potential under the nonbloom or after-bloom conditions, suggesting their potential roles in influencing the bloom's initiation and termination. However, in the natural plankton assemblage, the microbial interactive network might be much more complex. For instance, the increase of bacteria and bacterial hydrolysis in the nonbloom and after-bloom stages might have been a coincidence resulting from a microbial trophic network where ciliate slopping grazing and viral cytolytic activities produced organic matter that fueled bacterial growth. Thus, further investigation is needed in the future to directly attribute the bacterial activities to the bloom suppression and termination.

In marine ecosystems, ciliates and other heterotrophic protists are the major phytoplankton consumers (96) and strongly affect phytoplankton population dynamics (97); they regulate harmful algal bloom development as a top-down control (98–100). It was previously reported that *Strombidinopsis* sp. (ca. 150  $\mu\text{m}$  in cell length) could feed on many species of dinoflagellates (101–104). In the present study, endocytosis- and phagosome-related genes were more highly expressed in *Strombidinopsis* in the nonbloom and after-bloom sample groups, indicating that grazing of *K. longicanalis* by *Strombidinopsis* might have prevented bloom formation or accelerated the bloom decline in the nonbloom area or period. For instance, PIP5K modulates the actin cytoskeleton during the attachment and ingestion phases of phagocytosis (105). CFL is involved in the formation and disruption of the phagocytic cup (106). ARF are key proteins of endocytosis in plants (107). The expression levels of these genes and the relative abundance of *Strombidinopsis* were the highest in the nonbloom samples but were also high in the after-bloom samples.

Finally, competition by coexisting phytoplankton can also shape the dynamics of a phytoplankton species. Upon the decline of the *K. longicanalis* bloom under the after-bloom condition, the diatom *Chaetoceros* sp. grew to be dominant in the assemblage. The high and upregulated expression of the diatom silicon transporter in the after-bloom stage compared to the bloom stage was evidence that *Chaetoceros* grew more quickly in the after-bloom stage and hence had a higher demand for nutrients. For N-nutrient uptake, *Chaetoceros* showed the highest expression of ammonium transporter, consistent with the increased  $\text{NH}_4^+$  concentration that likely resulted from bloom decline in the after-bloom stage, indicating that  $\text{NH}_4^+$  might contribute to the growth of this diatom. The high and elevated expression of phosphate transporter and pyrophosphatase in *Chaetoceros* during after-bloom was also evidence of rapid DIP and DOP uptake, imposing competition with *K. longicanalis*. This nutrient competition by the diatom might have contributed to the termination of the dinoflagellate bloom.

While the high activities of ciliates, bacteria, and viruses likely substantially contributed to the prevention and termination of the *K. longicanalis* bloom, it is important to note that other factors might have been at play in the natural marine environment. The data reported here only represent snapshots of the bloom emergence and demise processes and might have missed some changes in the environment, other activities of the coexisting organisms, or possible vertical migration of the bloom-causing dinoflagellate. Future studies need to investigate those other potential factors with simultaneous cell-level physiological measurements with more intense time- and space (vertical)-resolved sampling.

**Conclusions.** This study provides the first transcriptome of *K. longicanalis* and represents the first natural bloom investigation of this species. Acquisition of energy and nutrients, defense, cell proliferation, and cell mortality were explored as potential regulators of a *Karenia longicanalis* bloom using an integrated molecular ecological approach. Our integrative data analysis indicated that the promoted cell proliferation and bloom formation in the bloom stage were supported by high or increased expression of energy and nutrient acquisition and metabolism genes in *Karenia*. Cell mortality due to microbial attack and ciliate grazing appeared to be the major inhibitor of bloom formation in the nonbloom stage and the promoter of bloom demise in the after-bloom stage. The competition of the *Chaetoceros* sp. likely also contributed to the decline of the *K. longicanalis* population and the rise of the *Chaetoceros* sp. population to dominance in the after-bloom stage. Our meta-transcriptomic data suggested that *K. longicanalis* was heavily “invested” in defense during the nonbloom or after-bloom stages, but defense was inadequate to overcome the overwhelming grazing and microbial attacks. Therefore, mechanisms underlying the *K. longicanalis* bloom appear to be different from those regulating the bloom of *P. shikokuense*, which requires the upregulation of genes related to energy and nutrient acquisition, defense, and sexual reproduction (25, 27, 29), or of *K. mikimotoi*, for which sexual reproduction appears to be influential (25). Of the four elements in ENDS, the most notable bloom-promoting factors for the *K. longicanalis* bloom were E and N, with the contribution of D and S being apparently weak, and the biotic bloom-inhibiting or -dissipating factors were evidently the failure of *K. longicanalis* to defend against microbial attack and grazing and to outcompete diatoms for nutrients. This novel insight into HAB-regulating mechanisms, along with the first transcriptomic data for *K. longicanalis* and the integrative molecular ecological approach, will prove to be a valuable resource and essential foundation for further elucidation of bloom regulators of this and related species of Kareniaceae in the future.

## MATERIALS AND METHODS

**Sample collection and environmental factor measurements.** Three sampling events, each in triple biological replicates, were launched to represent a nonbloom condition (nonbloom 1 to 3), a bloom condition (bloom 1 to 3), and an after-bloom condition (after-bloom 1 to 3) (Fig. 1A). Fifty-milliliter seawater samples were collected and fixed in Lugol's solution (2%) for subsequent microscopic examination to identify species and enumerate cell concentrations in a Sedgwick-Rafter counting chamber. The bloom condition was sampled on June 11 when the cell concentration of the bloom species was  $1.51 \times 10^7$  cells/liter. The nonbloom assemblage was sampled on the same day but from outside the bloom patch within the bay, where the cell concentration of the bloom species was 600-fold lower ( $2.4 \times 10^4$  cells/liter). The after-bloom condition was sampled in the same location as the previously sampled bloom area on June 12, when the abundance of the bloom species declined by 137-fold ( $1.1 \times 10^5$  cells/liter). At each sampling event, three separate samples were collected as replicates, giving a total of 9 samples. For each sample, 20 liters of water was taken from the subsurface (~0.5 m) and prescreened through a 200- $\mu\text{m}$  mesh to exclude large zooplankton. Immediately afterward, the sample was filtered onto 3- $\mu\text{m}$  polycarbonate filters (47 mm). Filtration was completed in 15 min to minimize changes in gene expression. Each filter was split into four equal parts using clean scissors, one for DNA analysis and the other three for RNA analysis. Cells retained on the filter membranes included planktonic eukaryotes (3 to 200  $\mu\text{m}$ ) and microbes that were infecting or attached onto the eukaryotic cells. Filters for RNA sequencing were immediately immersed in 1 mL TRIzol in a 2-mL tube and kept frozen in liquid  $\text{N}_2$  during the cruise and, upon return to the laboratory, were stored at  $-80^\circ\text{C}$  until RNA extraction. Samples for DNA sequencing were transferred to a 2-mL tube containing 1 mL DNA lysis buffer (10 mM Tris-HCl [pH 8.0], 100 mM EDTA [pH 8.0], 0.5% SDS, 200  $\mu\text{g mL}^{-1}$  proteinase K) and held at  $-20^\circ\text{C}$  until DNA extraction.

A comprehensive set of environmental factors was also measured. Seawater temperature, dissolved oxygen, conductivity, pH, and salinity were measured *in situ* using YSI Professional Plus, and surface light was measured using a digital lux meter. Two sets of 50-mL seawater samples from each of the three sampling events were filtered onto GF/F membranes and a 0.2- $\mu\text{m}$  membrane and stored at  $-80^\circ\text{C}$  for chlorophyll *a* and nutrient measurements, respectively. Chlorophyll *a* was measured using a Trilogy laboratory fluorometer (Turner Designs, USA). The concentrations of N nutrients ( $\text{NH}_4^+$ ,  $\text{NO}_3^-$ ,  $\text{NO}_2^-$ ),  $\text{PO}_4^{3-}$ , and  $\text{SiO}_4^{4-}$  were measured using the Bran-Lubbe AAIII system (Germany) on water samples that had been filtered through a 0.22- $\mu\text{m}$  membrane and kept at  $-80^\circ\text{C}$ . DOC and total dissolved nitrogen (TDN) were measured using the high-temperature catalytic oxidation method with a Multi N/C 3100 system (Germany). DON concentration was estimated based on TDN minus DIN ( $\text{NH}_4^+ + \text{NO}_3^- + \text{NO}_2^-$ ). For particulate organic phosphorus and particulate organic nitrogen (108–110), samples were first digested with the high-temperature and high-pressure REDOX methods, and the resulting inorganic products were measured using the Bran-Lubbe AAIII system (Germany).

**DNA extraction and barcoding sequencing.** DNA was extracted as previously reported (111). The DNA was used as the template for PCR amplifications of the 16S rRNA genes and 18S rDNA from prokaryotes and eukaryotes, respectively. The 16S rRNA V4 variable region (~290 bp) was amplified with the 515F-806R bacterial/archaeal primer pair 515F (5'-GTGCCAGCMGCCGCGTAA-3') and 806R (5'-GGACTA CHVGGGTWTCTA AT-3') (112). The 18S variable region (V4, 450 bp) was amplified with primers 18S V4-F (5'-GGCAAGTCTGGTGCCAG-3') and 18SV4-R (5'-GACTACGACGGTATCTRATCCTTCG-3') (113). The 28S rRNA gene was amplified with primers LSU rDNA-D1R-F (5'-ACCCGCTGAATTAAGCATA-3') and LSU rDNA-D2C-R (5'-CCTTGGTCCGTGTTCAAGA-3'). The PCR products were then sequenced. The 16S rRNA gene amplicons were sequenced on the HiSeq2500 platform (25,000 2 × 250-bp read pairs). The 18S rDNA amplicons were sequenced on the Miseq platform (25,000 2 × 300-bp read pairs). The 28S rRNA gene was amplified using nested PCR with primers KI\_28sF (5'-TAAGCGGAGGATAAGAACTAAATAGG-3'), KI\_28sR1 (5'-CCGTGTTCAAGACGGGTGCAATAA-3'), and KI\_28sR2 (5'-AACCATTTGTCATCGTACTTATGTC-3').

**RNA extraction and sequencing.** RNA extraction was performed as previously reported (27, 29) with minor modifications. Briefly, cells were rinsed off the filter with TRIzol using a pipette. The cell suspension was mixed with beads (Biospec, USA) and shaken at 6 m/s on a FastPrep-24 bead mill (MP Biomedicals, USA) two or three times to ensure complete cell breakage. The cell lysate was then subjected to RNA extraction following protocols of the TRI reagent and the Direct-zol RNA columns (114). One microgram of RNA with the RNA integrity number (RIN) above seven was used for library preparation according to the manufacturer's protocol (NEBNext Ultra directional RNA library prep kit for Illumina). To yield both prokaryotic and eukaryotic transcriptomes, instead of using oligo(dT)-based mRNA isolation, we removed rRNA from total RNA using the Ribo-Zero rRNA removal kit (bacteria; Illumina). cDNA libraries were prepared from the resulting rRNA-depleted RNA and sequenced following Illumina HiSeq PE150 protocols.

**Bioinformatics analyses. (i) Community classification and taxon interactions.** The raw data from the 16S and 18S amplicon sequencing were quality filtered to eliminate the adapters and low-quality reads to obtain clean reads with TrimmomaticPE. The parameters were set as follows: leading, 5; trailing, 5; sliding window, 4:15; minlen, 50. The clean data were used for feature identification and quantification using the UPRASE with NOISE3 method (115). Taxonomic assignment was carried out using SILVA 138.1 and the NCBI nucleotide database. The taxon-taxon and taxon-environment correlations were analyzed and visualized using ggcor packages in R software (116). Correlations between taxa and environmental factors were analyzed using the Pearson model, and those between species were analyzed using the Mantel test. Those with *P* values of <0.05 were considered significantly correlated.

**(ii) Metatranscriptomic read assembly.** Sequencing reads from each RNA sample were assembled with MEGAHIT v1.0 (117). The assembled unigenes from the triplicate samples of each bloom condition (group) were then clustered using CD-HIT (v4.6.5) (118) with the default parameters of "-d 0 -c 0.95 -G 0 -aL 0.95 -AL 100 -aS0.95 -AS 30" and TGICL (v4.6.5) (119). Finally, the unigenes from the three groups were merged into the final nonredundant reference metatranscriptome using the CD-HIT and TGICL software.

**(iii) Gene taxonomic assignment and functional annotation.** To identify the species origin of the unigenes in the reference metatranscriptome, a comprehensive in-house database (nrMegPhylodb) (120) was created by integrating several existing databases. These included protein sequences of the NCBI non-redundant protein database (NR, June 2019), PhyloDB (version 1.076), and proteomes of four dinoflagellates from our laboratory, including *P. shikokuense*, *K. mikimotoi*, *Effrenium* sp., and *Karlodinium veneficum*. BLASTX was used to match our reference metatranscriptome against the in-house database with an *E* value of 1E-3, and the taxonomic source of the best-hit gene was assigned to the query unigene sequence.

Functional annotation of the unigenes in the reference metatranscriptome was conducted using Diamond (121), based on a Blastx search against such databases nrMegPhylodb, eggNOG (v4.5), KEGG (April 2018), and Swiss-Prot (October 2018). Gene ontology was assigned based on NR annotation.

**(iv) Gene expression quantification and differential expression analysis.** The raw reads of the metatranscriptomes were trimmed and quality-filtered using Cutadapt (v1.9.1) with parameters setting at "-m 75 -max-n 0.1 -discard-trimmed -q 20." Clean reads from each sample were mapped to the above-described reference metatranscriptome using the aligner Bowtie2 and quantified using RSEM (v1.2.4) with default parameters. Gene expression at the community level (community TPM) was calculated as transcripts of the specific gene per million of total mapped reads of the community and used to compare gene expression contributions of different lineages or kingdoms to the whole community. While comparing gene expression differences in the same species between sample groups, gene expression at the species level was renormalized as transcripts of the specific gene from the specific species per million total mapped reads of the species (species TPM) was used. Only genes with an average gene expression TPM of ≥0.1 were used in differential expression analyses within species, including *K. longicanalis* and *Chaetoceros* sp., for which the DESeq package in R software was employed. DEGs were defined as those with an adjusted *P* value of <0.05 and fold change of ≥2. DEGs associated with energy and nutrient acquisition, defense, and cell reproduction were identified using a *t* test based on average gene expression (TPM) of the gene family, as previously reported (25).

**Data availability.** Raw sequence reads of mRNA and the DNA metabarcodes were deposited in the NCBI Sequence Read Archive (BioProject ID [PRJNA689700](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA689700)).

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**SUPPLEMENTAL FILE 1**, PDF file, 0.3 MB.

**SUPPLEMENTAL FILE 2**, XLSX file, 0.1 MB.

## ACKNOWLEDGMENTS

We thank Keqiang Li's group of the Ocean University of China for measuring and providing data of environmental factors. We appreciate Xin Lin's suggestions on data analyses. We thank Huatao Yuan for collecting samples and Chentao Guo for providing logistic support. This work was supported by the National Key Research and Development Program of China (grants 2022YFC3105301 and 2017YFC1404302) and the Natural Science Foundation of China (grant NSFC 42276096).

We declare no conflict of interest.

Liyang Yu contributed Conceptualization, Methodology, Writing – Original Draft. Tangcheng Li contributed Investigation. Hongfei Li contributed Investigation. Minglei Ma contributed Investigation. Ling Li contributed Resources. Senjie Lin contributed Conceptualization, Supervision, Funding Acquisition, Writing – Reviewing & Editing.

## REFERENCES

- Anderson DM, Cembella AD, Hallegraeff GM. 2012. Progress in understanding harmful algal blooms: paradigm shifts and new technologies for research, monitoring, and management. *Annu Rev Mar Sci* 4:143–176. <https://doi.org/10.1146/annurev-marine-120308-081121>.
- Gobler CJ, Doherty OM, Hattenrath-Lehmann TK, Griffith AW, Kang Y, Litaker RW. 2017. Ocean warming since 1982 has expanded the niche of toxic algal blooms in the North Atlantic and North Pacific oceans. *Proc Natl Acad Sci U S A* 114:4975–4980. <https://doi.org/10.1073/pnas.1619575114>.
- Heisler J, Glibert P, Burkholder J, Anderson D, Cochlan W, Dennison W, Gobler C, Dortch Q, Heil C, Humphries E, Lewitus A, Magnien R, Marshall H, Sellner K, Stockwell D, Stoecker D, Suddleson M. 2008. Eutrophication and harmful algal blooms: a scientific consensus. *Harmful Algae* 8:3–13. <https://doi.org/10.1016/j.hal.2008.08.006>.
- Sellner KG, Doucette GJ, Kirkpatrick GJ. 2003. Harmful algal blooms: causes, impacts and detection. *J Ind Microbiol Biotechnol* 30:383–406. <https://doi.org/10.1007/s10295-003-0074-9>.
- Thyng KM, Hetland RD, Ogle MT, Zhang X, Chen F, Campbell LJJ, Fluids O. 2013. Origins of *Karenia brevis* harmful algal blooms along the Texas coast. *Limnol Oceanogr* 3:269–278. <https://doi.org/10.1215/21573689-2417719>.
- Guo Y, Lin S, Huang L, Chen Y, Hu S, Liu S, Tan Y, Huang X, Qiu D. 2022. Dissipation of a *Polykrikos geminatum* Bloom after wind events in Pearl River Estuary. *Water* 14:2313. <https://doi.org/10.3390/w14152313>.
- Behrenfeld MJ, Boss ES. 2014. Resurrecting the ecological underpinnings of ocean plankton blooms. *Annu Rev Mar Sci* 6:167–194. <https://doi.org/10.1146/annurev-marine-052913-021325>.
- Lips I, Lips U. 2008. Abiotic factors influencing cyanobacterial bloom development in the Gulf of Finland (Baltic Sea). *Hydrobiologia* 614:133–140. <https://doi.org/10.1007/s10750-008-9449-2>.
- Huber V, Wagner C, Gerten D, Adrian R. 2012. To bloom or not to bloom: contrasting responses of cyanobacteria to recent heat waves explained by critical thresholds of abiotic drivers. *Oecologia* 169:245–256. <https://doi.org/10.1007/s00442-011-2186-7>.
- Tao M, Xie P, Chen J, Qin B, Zhang D, Niu Y, Zhang M, Wang Q, Wu L. 2012. Use of a generalized additive model to investigate key abiotic factors affecting microcystin cellular quotas in heavy bloom areas of Lake Taihu. *PLoS One* 7:e32020. <https://doi.org/10.1371/journal.pone.0032020>.
- Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil L-A, Thingstad F. 1983. The ecological role of water-column microbes in the sea. *Mar Ecol Prog Ser* 10:257–263. <https://doi.org/10.3354/meps010257>.
- Graneli E, Turner JT (ed). 2006. Ecology of harmful algae, vol 189. Springer, Cham, Switzerland.
- Li Y, Liu L, Xu Y, Guan C, Lei X, Zheng W, Wang H, Zheng T. 2016. First evidence of *Altererythrobracter* sp. LY02 with indirect algicidal activity on the toxic dinoflagellate, *Alexandrium tamarense*. *Curr Microbiol* 73:550–560. <https://doi.org/10.1007/s00284-016-1093-x>.
- Shi X, Liu L, Li Y, Xiao Y, Ding G, Lin S, Chen J. 2018. Isolation of an algicidal bacterium and its effects against the harmful-algal- bloom dinoflagellate *Prorocentrum donghaiense* (Dinophyceae). *Harmful Algae* 80:72–79. <https://doi.org/10.1016/j.hal.2018.09.003>.
- Adachi M, Kanno T, Okamoto R, Itakura S, Yamaguchi M, Nishijima T. 2003. Population structure of *Alexandrium* (Dinophyceae) cyst formation-promoting bacteria in Hiroshima Bay, Japan. *Appl Environ Microbiol* 69:6560–6568. <https://doi.org/10.1128/AEM.69.11.6560-6568.2003>.
- Ferrier M, Martin JL, Rooney-Varga JN. 2002. Stimulation of *Alexandrium fundyense* growth by bacterial assemblages from the Bay of Fundy. *J Appl Microbiol* 92:706–716. <https://doi.org/10.1046/j.1365-2672.2002.01576.x>.
- Zhou J, Richlen ML, Sehein TR, Kulis DM, Anderson DM, Cai Z. 2018. Microbial community structure and associations during a marine dinoflagellate bloom. *Front Microbiol* 9:1201. <https://doi.org/10.3389/fmicb.2018.01201>.
- Uribe P, Espejo RT. 2003. Effect of associated bacteria on the growth and toxicity of *Alexandrium catenella*. *Appl Environ Microbiol* 69:659–662. <https://doi.org/10.1128/AEM.69.1.659-662.2003>.
- Van Dolah FM, Lidie KB, Monroe EA, Bhattacharya D, Campbell L, Doucette GJ, Kamykowski D. 2009. The Florida red tide dinoflagellate *Karenia brevis*: new insights into cellular and molecular processes underlying bloom dynamics. *Harmful Algae* 8:562–572. <https://doi.org/10.1016/j.hal.2008.11.004>.
- Marchetti A, Schrueth DM, Durkin CA, Parker MS, Kodner RB, Berthiaume CT, Morales R, Allen AE, Armbrust EV. 2012. Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proc Natl Acad Sci U S A* 109: E317–25. <https://doi.org/10.1073/pnas.1118408109>.
- Alexander H, Jenkins BD, Rynearson TA, Dyhrman ST. 2015. Metatranscriptome analyses indicate resource partitioning between diatoms in the field. *Proc Natl Acad Sci U S A* 112:E2182–E2190. <https://doi.org/10.1073/pnas.1421993112>.
- Carradec Q, Pelletier E, Da Silva C, Alberti A, Seeleuthner Y, Blanc-Mathieu R, Lima-Mendez G, Rocha F, Tirichine L, Labadie K, Kirilovsky A, Bertrand A, Engelen S, Madoui M-A, Méheust R, Poulain J, Romac S, Richter DJ, Yoshikawa G, Dimier C, Kandels-Lewis S, Picheral M, Searson S, Tara Oceans C, Jaillon O, Aury J-M, Karsenti E, Sullivan MB, Sunagawa S, Bork P, Not F, Hingamp P, Raes J, Guidi L, Ogata H, de Vargas C, Iudicone D, Bowler C, Wincker P, Tara Oceans Coordinators. 2018. A global ocean atlas of eukaryotic genes. *Nat Commun* 9: 373–373. <https://doi.org/10.1038/s41467-017-02342-1>.
- Cooper ED, Bentlage B, Gibbons TR, Bachvaroff TR, Delwiche CF. 2014. Metatranscriptome profiling of a harmful algal bloom. *Harmful Algae* 37: 75–83. <https://doi.org/10.1016/j.hal.2014.04.016>.
- Gong W, Browne J, Hall N, Schrueth D, Paerl H, Marchetti A. 2017. Molecular insights into a dinoflagellate bloom. *ISME J* 11:439–452. <https://doi.org/10.1038/ismej.2016.129>.
- Lin S, Yu L, Wu X, Li M, Zhang Y, Luo H, Li H, Li T, Li L. 2022. Active meiosis during dinoflagellate blooms: a 'sex for proliferation' hypothesis. *Harmful Algae* 118:102307. <https://doi.org/10.1016/j.hal.2022.102307>.
- Xu X, Yu Z, He L, Cao X, Chen N, Song X. 2020. Metabolic analyses by metatranscriptomics highlight plasticity in phosphorus acquisition during monospecific and multispecies algal blooms. *Hydrobiologia* 847: 1071–1085. <https://doi.org/10.1007/s10750-019-04169-x>.
- Yu L, Zhang Y, Li M, Wang C, Lin X, Li L, Shi X, Guo C, Lin S. 2020. Comparative metatranscriptomic profiling and microRNA sequencing to reveal active metabolic pathways associated with a dinoflagellate bloom. *Sci Total Environ* 699:134323. <https://doi.org/10.1016/j.scitotenv.2019.134323>.
- Zhuang Y, Zhang H, Hannick L, Lin S. 2015. Metatranscriptome profiling reveals versatile N-nutrient utilization, CO<sub>2</sub> limitation, oxidative stress, and active toxin production in an *Alexandrium fundyense* bloom. *Harmful Algae* 42:60–70. <https://doi.org/10.1016/j.hal.2014.12.006>.
- Zhang Y, Lin X, Shi X, Lin L, Luo H, Li L, Lin S. 2019. Metatranscriptomic signatures associated with regime shift from diatom dominance to a dinoflagellate bloom. *Front Microbiol* 10:590. <https://doi.org/10.3389/fmicb.2019.00590>.

30. Zhang Y, Lin X, Li T, Li H, Lin L, Luo H, Li L, Ji N, Lin S. 2020. High throughput sequencing of 18S rRNA and its gene to characterize a *Prorocentrum shikokuense* (Dinophyceae) bloom. *Harmful Algae* 94:101809. <https://doi.org/10.1016/j.hal.2020.101809>.
31. Luo H, Wang J, Goes JI, Gomes H, Al-Hashmi K, Tobias C, Koerting C, Lin S. 2022. A grazing-driven positive nutrient feedback loop and active sexual reproduction underpin widespread Noctiluca green tides. *ISME Comm* 2: 103. <https://doi.org/10.1038/s43705-022-00187-4>.
32. Smayda TJ. 2002. Adaptive ecology, growth strategies and the global bloom expansion of dinoflagellates. *J Oceanogr* 58:281–294. <https://doi.org/10.1023/A:1015861725470>.
33. Warns A, Hense I, Kremp A. 2013. Modelling the life cycle of dinoflagellates: a case study with *Biecheleria baltica*. *J Plankton Res* 35:379–392. <https://doi.org/10.1093/plankt/fbs095>.
34. Zhou Z-X, Yu R-C, Zhou M-J. 2017. Seasonal succession of microalgal blooms from diatoms to dinoflagellates in the East China Sea: a numerical simulation study. *Ecol Modell* 360:150–162. <https://doi.org/10.1016/j.ecolmodel.2017.06.027>.
35. Fischer AD, Hayashi K, McGaraghan A, Kudela RM. 2020. Return of the “age of dinoflagellates” in Monterey Bay: drivers of dinoflagellate dominance examined using automated imaging flow cytometry and long-term time series analysis. *Limnol Oceanogr* 65:2125–2141. <https://doi.org/10.1002/lno.11443>.
36. Li X-Y, Yu R-C, Geng H-X, Li Y-F. 2021. Increasing dominance of dinoflagellate red tides in the coastal waters of Yellow Sea, China. *Mar Pollut Bull* 168:112439. <https://doi.org/10.1016/j.marpolbul.2021.112439>.
37. Hetland RD, Campbell L. 2007. Convergent blooms of *Karenia brevis* along the Texas coast. *Geophys Res Lett* 34. <https://doi.org/10.1029/2007GL030474>.
38. Li X, Yan T, Yu R, Zhou M. 2019. A review of *Karenia mikimotoi*: bloom events, physiology, toxicity and toxic mechanism. *Harmful Algae* 90: 101702. <https://doi.org/10.1016/j.hal.2019.101702>.
39. Li X, Yan T, Lin J, Yu R, Zhou M. 2017. Detrimental impacts of the dinoflagellate *Karenia mikimotoi* in Fujian coastal waters on typical marine organisms. *Harmful Algae* 61:1–12. <https://doi.org/10.1016/j.hal.2016.11.011>.
40. Stumpf R, Culver M, Tester P, Tomlinson M, Kirkpatrick G, Pederson B, Truby E, Ransibrahmanakul V, Soracco M. 2003. Monitoring *Karenia brevis* blooms in the Gulf of Mexico using satellite ocean color imagery and other data. *Harmful Algae* 2:147–160. [https://doi.org/10.1016/S1568-9883\(02\)00083-5](https://doi.org/10.1016/S1568-9883(02)00083-5).
41. Rolton A, Rhodes L, Hutson KS, Biessy L, Bui T, MacKenzie L, Symonds JE, Smith KF. 2022. Effects of harmful algal blooms on fish and shellfish species: a case study of New Zealand in a changing environment. *Toxins (Basel)* 14:341. <https://doi.org/10.3390/toxins14050341>.
42. Feki W, Hamza A, Frossard V, Abdennadher M, Hannachi I, Jacquot M, Belhassen M, Aleya L. 2013. What are the potential drivers of blooms of the toxic dinoflagellate *Karenia selliformis*? A 10-year study in the Gulf of Gabes, Tunisia, southwestern Mediterranean Sea. *Harmful Algae* 23:8–18. <https://doi.org/10.1016/j.hal.2012.12.001>.
43. Barnes MK, Tilstone GH, Smyth TJ, Widdicombe CE, Gloël J, Robinson C, Kaiser J, Suggett D. 2015. Drivers and effects of *Karenia mikimotoi* blooms in the western English Channel. *Prog Oceanogr* 137:456–469. <https://doi.org/10.1016/j.pocean.2015.04.018>.
44. Kok JWK, Leong SCY. 2019. Nutrient conditions and the occurrence of a *Karenia mikimotoi* (Kareniaceae) bloom within East Johor Straits, Singapore. *Reg Stud Mar Sci* 27:100514. <https://doi.org/10.1016/j.rsmas.2019.100514>.
45. Medina M, Huffaker R, Jawitz JW, Muñoz-Carpena R. 2020. Seasonal dynamics of terrestrially sourced nitrogen influenced *Karenia brevis* blooms off Florida's southern Gulf Coast. *Harmful Algae* 98:101900. <https://doi.org/10.1016/j.hal.2020.101900>.
46. Heil CA, Dixon LK, Hall E, Garrett M, Lenes JM, O'Neil JM, Walsh BM, Bronk DA, Killberg-Thoreson L, Hitchcock GL, Meyer KA, Mulholland MR, Prociak L, Kirkpatrick GJ, Walsh JJ, Weisberg RW. 2014. Blooms of *Karenia brevis* (Davis) G. Hansen & Ø. Moestrup on the West Florida Shelf: nutrient sources and potential management strategies based on a multi-year regional study. *Harmful Algae* 38:127–140. <https://doi.org/10.1016/j.hal.2014.07.016>.
47. Wang J, Cen J, Li S, Lü S, Moestrup Ø, Chan K-K, Jiang T, Lei X. 2018. A re-investigation of the bloom-forming unarmored dinoflagellate *Karenia longicanalis* (syn. *Karenia umbella*) from Chinese coastal waters. *J Ocean Limnol* 36:2202–2215. <https://doi.org/10.1007/s00343-019-7191-4>.
48. Kaur A, Upadhyay SK. 2022. EF-hand domain-containing proteins: diversity and role in plants, p 185–203. In Upadhyay SK (ed), *Cation transporters in plants*. Elsevier, Amsterdam, Netherlands.
49. Chen Z, Wang X, Li S, Yao J, Shao Z, Duan D. 2019. Verification of the *Saccharina japonica* translocon Tic20 and its localization in the chloroplast membrane in diatoms. *Int J Mol Sci* 20:4000. <https://doi.org/10.3390/ijms20164000>.
50. Sala-Rovira M, Geraud M, Caput D, Jacques F, Soyer-Gobillard M, Vernet G, Herzog MJC. 1991. Molecular cloning and immunolocalization of two variants of the major basic nuclear protein (HCC) from the histone-less eukaryote *Cryptothecodinium cohnii* (Pyrrhophyta). *Chromosoma* 100: 510–518. <https://doi.org/10.1007/BF00352201>.
51. Lin S. 2011. Genomic understanding of dinoflagellates. *Res Microbiol* 162: 551–569. <https://doi.org/10.1016/j.resmic.2011.04.006>.
52. Anderson DM, Burkholder JM, Cochlan WP, Glibert PM, Gobler CJ, Heil CA, Kudela RM, Parsons ML, Rensel JEJ, Townsend DW, Trainer VL, Vargo GA. 2008. Harmful algal blooms and eutrophication: examining linkages from selected coastal regions of the United States. *Harmful Algae* 8: 39–53. <https://doi.org/10.1016/j.hal.2008.08.017>.
53. Glibert PM. 2017. Eutrophication, harmful algae and biodiversity: challenging paradigms in a world of complex nutrient changes. *Mar Pollut Bull* 124:591–606. <https://doi.org/10.1016/j.marpolbul.2017.04.027>.
54. Zhang Y, Song X, Shen H, Cao X, Yuan Y, Wu Z, Yu Z. 2020. The effects of modified clay on abalone (*Haliotis discus hannai*) based on laboratory and field experiments. *Environ Toxicol Chem* 39:2065–2075. <https://doi.org/10.1002/etc.4822>.
55. Theimer RR, Heidinger P. 1974. Control of particulate urate oxidase activity in bean roots by external nitrogen supply. *Zeitschrift Für Pflanzenphysiol* 73:360–370. [https://doi.org/10.1016/S0044-328X\(74\)80136-4](https://doi.org/10.1016/S0044-328X(74)80136-4).
56. Duran VA, Todd CD. 2012. Four allantoinase genes are expressed in nitrogen-fixing soybean. *Plant Physiol Biochem* 54:149–155. <https://doi.org/10.1016/j.plaphy.2012.03.002>.
57. Piedras P, Munoz A, Aguilar M, Pineda M. 2000. Allantoate amidinohydrolase (allantoicase) from *Chlamydomonas reinhardtii*: its purification and catalytic and molecular characterization. *Arch Biochem Biophys* 378: 340–348. <https://doi.org/10.1006/abbi.2000.1833>.
58. Leulliot N, Quevillon-Cheruel S, Sorel I, Graille M, Meyer P, Liger D, Blondeau K, Janin J, Van Tilbeurgh HJ. 2004. Crystal structure of yeast allantoinase reveals a repeated jelly roll motif. *J Biol Chem* 279:23447–23452. <https://doi.org/10.1074/jbc.M401336200>.
59. Piedras P, Cárdenas J, Pineda MJPA. 1995. Solubilization and extraction of allantoinase and allantoinase from the green alga *Chlamydomonas reinhardtii*. *Phytochem Anal* 6:239–243. <https://doi.org/10.1002/pca.2800060503>.
60. Bekheet IA, Syrett PJ. 1977. Urea-degrading enzymes in algae. *Br Phycol J* 12:137–143. <https://doi.org/10.1080/000716177006050151>.
61. Fu L, Li X, Xiao X, Xu J. 2014. Purification and characterization of a thermostable aliphatic amidase from the hyperthermophilic archaeon *Pyrococcus yayanosii* CH1. *Extremophiles* 18:429–440. <https://doi.org/10.1007/s00792-014-0628-y>.
62. Pratt AC, Dewage SW, Pang AH, Biswas T, Barnard-Britson S, Cisneros GA, Tsodikov OV. 2015. Structural and computational dissection of the catalytic mechanism of the inorganic pyrophosphatase from *Mycobacterium tuberculosis*. *J Struct Biol* 192:76–87. <https://doi.org/10.1016/j.jsb.2015.08.010>.
63. Oh SJ, Yamamoto T, Kataoka Y, Matsuda O, Matsuyama Y, Kotani Y. 2002. Utilization of dissolved organic phosphorus by the two toxic dinoflagellates, *Alexandrium tamarense* and *Gymnodinium catenatum* (Dinophyceae). *Fish Sci* 68:416–424. <https://doi.org/10.1046/j.1444-2906.2002.00440.x>.
64. Sano S-I, Matsuda Y, Miyamoto S, Nakagawa H. 1984. Thiamine pyrophosphatase and nucleoside diphosphatase in rat brain. *Biochem Biophys Res Commun* 118:292–298. [https://doi.org/10.1016/0006-291x\(84\)91099-4](https://doi.org/10.1016/0006-291x(84)91099-4).
65. Luo H, Lin X, Li L, Lin L, Zhang C, Lin S. 2017. Transcriptomic and physiological analyses of the dinoflagellate *Karenia mikimotoi* reveal non-alkaline phosphatase-based molecular machinery of ATP utilisation. *Environ Microbiol* 19:4506–4518. <https://doi.org/10.1111/1462-2920.13899>.
66. Culley AI, Lang AS, Suttle CA. 2003. High diversity of unknown picorna-like viruses in the sea. *Nature* 424:1054–1057. <https://doi.org/10.1038/nature01886>.
67. Koonin EV, Wolf YI, Nagasaki K, Dolja VV. 2008. The big bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. *Nat Rev Microbiol* 6:925–939. <https://doi.org/10.1038/nrmicro2030>.
68. Culley AI, Lang AS, Suttle CA. 2006. Metagenomic analysis of coastal RNA virus communities. *Science* 312:1795–1798. <https://doi.org/10.1126/science.1127404>.
69. Kristensen DM, Mushegian AR, Dolja VV, Koonin EV. 2010. New dimensions of the virus world discovered through metagenomics. *Trends Microbiol* 18:11–19. <https://doi.org/10.1016/j.tim.2009.11.003>.



70. Lang AS, Culley AI, Suttle CA. 2004. Genome sequence and characterization of a virus (HaRNAV) related to picorna-like viruses that infects the marine toxic bloom-forming alga *Heterosigma akashiwo*. *Virology* 320: 206–217. <https://doi.org/10.1016/j.virol.2003.10.015>.
71. Schulz F, Alteio L, Goudeau D, Ryan EM, Yu FB, Malmstrom RR, Blanchard J, Woyke T. 2018. Hidden diversity of soil giant viruses. *Nat Commun* 9: 4881. <https://doi.org/10.1038/s41467-018-07335-2>.
72. Fischer MG. 2016. Giant viruses come of age. *Curr Opin Microbiol* 31: 50–57. <https://doi.org/10.1016/j.mib.2016.03.001>.
73. Wilson WH, Van Etten JL, Allen MJ. 2009. The Phycodnaviridae: the story of how tiny giants rule the world. *Curr Top Microbiol Immunol* 328:1–42.
74. Hingamp P, Grimsley N, Acinas SG, Clerissi C, Subirana L, Poulain J, Ferrera I, Sarmiento H, Villar E, Lima-Mendez G, Faust K, Sunagawa S, Claverie J-M, Moreau H, Desdevises Y, Bork P, Raes J, de Vargas C, Karsenti E, Kandels-Lewis S, Jaillon O, Not F, Pesant S, Wincker P, Ogata H. 2013. Exploring nucleocytoplasmic large DNA viruses in Tara Oceans microbial metagenomes. *ISME J* 7:1678–1695. <https://doi.org/10.1038/ismej.2013.59>.
75. Gallot-Lavallée L, Blanc G, Claverie J-M. 2017. Comparative genomics of Chrysochromulina ericina virus and other microalga-infecting large DNA viruses highlights their intricate evolutionary relationship with the established Mimiviridae family. *J Virol* 91:e00230-17. <https://doi.org/10.1128/JVI.00230-17>.
76. Fischer MG, Allen MJ, Wilson WH, Suttle CA. 2010. Giant virus with a remarkable complement of genes infects marine zooplankton. *Proc Natl Acad Sci U S A* 107:19508–19513. <https://doi.org/10.1073/pnas.1007615107>.
77. Zhang W, Zhou J, Liu T, Yu Y, Pan Y, Yan S, Wang Y. 2015. Four novel algal virus genomes discovered from Yellowstone Lake metagenomes. *Sci Rep* 5:15131. <https://doi.org/10.1038/srep15131>.
78. Santini S, Jeudy S, Bartoli J, Poirot O, Lescot M, Abergel C, Barbe V, Wommack KE, Noordeloos AAM, Brussaard CPD, Claverie J-M. 2013. Genome of *Phaeocystis globosa* virus PgV-16T highlights the common ancestry of the largest known DNA viruses infecting eukaryotes. *Proc Natl Acad Sci U S A* 110:10800–10805. <https://doi.org/10.1073/pnas.1303251110>.
79. Benites LF, Stephens TG, Bhattacharya D. 2022. Multiple waves of viral invasions in Symbiodiniaceae algal genomes. *Virus Evol* 8:veac101. <https://doi.org/10.1093/ve/veac101>.
80. Graves MV, Meints RH. 1992. Characterization of the major capsid protein and cloning of its gene from algal virus PBCV-1. *Virology* 188:198–207. [https://doi.org/10.1016/0042-6822\(92\)90750-j](https://doi.org/10.1016/0042-6822(92)90750-j).
81. Park Y, Lee K, Lee YS, Kim SW, Choi T-J. 2011. Detection of diverse marine algal viruses in the South Sea regions of Korea by PCR amplification of the DNA polymerase and major capsid protein genes. *Virus Res* 159: 43–50. <https://doi.org/10.1016/j.virusres.2011.04.024>.
82. Amaro AM, Fuentes MS, Ogalde SR, Venegas JA, Suárez-Isla BA. 2005. Identification and characterization of potentially algal-lytic marine bacteria strongly associated with the toxic dinoflagellate *Alexandrium catenella*. *J Eukaryot Microbiol* 52:191–200. <https://doi.org/10.1111/j.1550-7408.2005.00031.x>.
83. Gerken HG, Donohoe B, Knoshaug EP. 2013. Enzymatic cell wall degradation of *Chlorellavulgaris* and other microalgae for biofuels production. *Planta* 237:239–253. <https://doi.org/10.1007/s00425-012-1765-0>.
84. Matsumoto M, Yokouchi H, Suzuki N, Ohata H, Matsunaga T. 2003. Saccharification of marine microalgae using marine bacteria for ethanol production. *Appl Biochem Biotechnol* 105–108:247–254. <https://doi.org/10.1385/ABAB:105:1-3:247>.
85. Cho H-S, Oh Y-K, Park S-C, Lee J-W, Park J-Y. 2013. Effects of enzymatic hydrolysis on lipid extraction from *Chlorella vulgaris*. *Renew Energy* 54: 156–160. <https://doi.org/10.1016/j.renene.2012.08.031>.
86. Maki M, Leung KT, Qin W. 2009. The prospects of cellulase-producing bacteria for the bioconversion of lignocellulosic biomass. *Int J Biol Sci* 5: 500–516. <https://doi.org/10.7150/ijbs.5.500>.
87. Murashima K, Kosugi A, Doi RH. 2003. Synergistic effects of cellulosomal xylanase and cellulases from *Clostridium cellulovorans* on plant cell wall degradation. *J Bacteriol* 185:1518–1524. <https://doi.org/10.1128/JB.185.5.1518-1524.2003>.
88. Demuez M, González-Fernández C, Ballesteros M. 2015. Algicidal microorganisms and secreted algicides: new tools to induce microalgal cell disruption. *Biotechnol Adv* 33:1615–1625. <https://doi.org/10.1016/j.biotechadv.2015.08.003>.
89. Zheng N, Ding N, Gao P, Han M, Liu X, Wang J, Sun L, Fu B, Wang R, Zhou J. 2018. Diverse algicidal bacteria associated with harmful bloom-forming *Karenia mikimotoi* in estuarine soil and seawater. *Sci Total Environ* 631–632:1415–1420. <https://doi.org/10.1016/j.scitotenv.2018.03.035>.
90. Adachi M, Fukami K, Kondo R, Nishijima T. 2002. Identification of marine hybrid *Flavobacterium* sp. 5 N-3 using multiple probes and wholecell hybridization. *Fish Sci* 68:713–720. <https://doi.org/10.1046/j.1444-2906.2002.00484.x>.
91. Iwata Y, Sugahara I, Kimura T, Kowa H, Matsumoto A, Noritake K. 2004. Properties of an algicidal bacterium (*Flavobacterium* sp.) against *Karenia mikimotoi* isolated from Ise Bay, Japan. *Bull Jpn Soc Fish (Japan)* 61: 780–786. <https://doi.org/10.2331/fishsci.61.780>.
92. Roth PB, Mikulski CM, Doucette GJ. 2008. Influence of microbial interactions on the susceptibility of *Karenia* spp. to algicidal bacteria. *Aquat Microb Ecol* 50:251–259. <https://doi.org/10.3354/ame01167>.
93. Yoshinaga I, Kawai T, Ishida Y. 1997. Analysis of algicidal ranges of the bacteria killing the marine dinoflagellate *Gymnodinium mikimotoi* isolated from Tanabe Bay, Wakayama Pref, Japan. *Fish Sci* 63:94–98. <https://doi.org/10.2331/fishsci.63.94>.
94. Wang BX, Zhou YY, Bai SJ, Su JQ, Tian Y, Zheng TL, Yang XR. 2010. A novel marine bacterium algicidal to the toxic dinoflagellate *Alexandrium tamarense*. *Lett Appl Microbiol* 51:552–557. <https://doi.org/10.1111/j.1472-765X.2010.02936.x>.
95. Wang Y, Li S, Liu G, Li X, Yang Q, Xu Y, Hu Z, Chen C-Y, Chang J-S. 2020. Continuous production of algicidal compounds against *Akashiwo sanguinea* via a *Vibrio* sp. co-culture. *Bioresour Technol* 295:122246. <https://doi.org/10.1016/j.biortech.2019.122246>.
96. Bjorbækmo MFM, Evenstad A, Rosæg LL, Krabberød AK, Logares R. 2020. The planktonic protist interactome: where do we stand after a century of research? *ISME J* 14:544–559. <https://doi.org/10.1038/s41396-019-0542-5>.
97. Calbet A, Landry MR. 2004. Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. *Limnol Oceanogr* 49: 51–57. <https://doi.org/10.4319/lo.2004.49.1.0051>.
98. Calbet A, Vaqué D, Felipe J, Vila M, Sala MM, Alcaraz M, Estrada M. 2003. Relative grazing impact of microzooplankton and mesozooplankton on a bloom of the toxic dinoflagellate *Alexandrium minutum*. *Mar Ecol Prog Ser* 259:303–309. <https://doi.org/10.3354/meps259303>.
99. Jeong HJ, Kim JS, Song JY, Kim JH, Kim TH, Kim SK, Kang NS. 2007. Feeding by protists and copepods on the heterotrophic dinoflagellates *Pfiesteria piscicida*, *Stoeckeria algicida*, and *Luciella masanensis*. *Mar Ecol Prog Ser* 349:199–211. <https://doi.org/10.3354/meps07094>.
100. Jeong HJ, Kim JS, Yoo YD, Kim ST, Kim TH, Park MG, Lee CH, Seong KA, Kang NS, Shim JH. 2003. Feeding by the heterotrophic dinoflagellate *Oxyrrhis marina* on the red-tide raphidophyte *Heterosigma akashiwo*: a potential biological method to control red tides using mass-cultured grazers. *J Eukaryot Microbiol* 50:274–282. <https://doi.org/10.1111/j.1550-7408.2003.tb00134.x>.
101. Jeong HJ, Shim JH, Lee CW, Kim JS, Koh SM. 1999. Growth and grazing rates of the marine planktonic ciliate *Strombidinopsis* sp. on red-tide and toxic dinoflagellates. *J Eukaryot Microbiol* 46:69–76. <https://doi.org/10.1111/j.1550-7408.1999.tb04586.x>.
102. Jeong HJ, Kim JS, Lee KH, Seong KA, Yoo YD, Kang NS, Kim TH, Song JY, Kwon JE. 2017. Differential interactions between the nematocyst-bearing mixotrophic dinoflagellate *Paragymnodinium shiwhaense* and common heterotrophic protists and copepods: killer or prey. *Harmful Algae* 62:37–51. <https://doi.org/10.1016/j.hal.2016.12.005>.
103. Ok JH, Jeong HJ, Lim AS, Lee KH. 2017. Interactions between the mixotrophic dinoflagellate Takayama helix and common heterotrophic protists. *Harmful Algae* 68:178–191. <https://doi.org/10.1016/j.hal.2017.08.006>.
104. Yoo YD, Jeong HJ, Kang NS, Kim JS, Kim TH, Yoon EY. 2010. Ecology of *Gymnodinium aureolum*. II. Predation by common heterotrophic dinoflagellates and a ciliate. *Aquat Microb Ecol* 59:257–272. <https://doi.org/10.3354/ame01401>.
105. Mao YS, Yamaga M, Zhu X, Wei Y, Sun H-Q, Wang J, Yun M, Wang Y, Di Paolo G, Bennett M, Mellman I, Abrams CS, De Camilli P, Lu CY, Yin HL. 2009. Essential and unique roles of PIP5K- $\gamma$  and - $\alpha$  in Fc $\gamma$  receptor-mediated phagocytosis. *J Cell Biol* 184:281–296. <https://doi.org/10.1083/jcb.200806121>.
106. Bierne H, Gouin E, Roux P, Caroni P, Yin HL, Cossart P. 2001. A role for cofilin and LIM kinase in *Listeria*-induced phagocytosis. *J Cell Biol* 155: 111–112. <https://doi.org/10.1083/jcb.200104037>.
107. Naramoto S, Kleine-Vehn J, Robert S, Fujimoto M, Dainobu T, Paciorek T, Ueda T, Nakano A, Van Montagu MCE, Fukuda H, Friml J. 2010. ADP-ribosylation factor machinery mediates endocytosis in plant cells. *Proc Natl Acad Sci U S A* 107:21890–21895. <https://doi.org/10.1073/pnas.1016260107>.
108. Pujo-Pay M, Raimbault P. 1994. Improvement of the wet-oxidation procedure for simultaneous determination of particulate organic nitrogen and phosphorus collected on filters. *Mar Ecol Prog Ser* 105:203–207. <https://doi.org/10.3354/meps105203>.

109. Raimbault P, Diaz F, Pouvesle W, Boudjellal B. 1999. Simultaneous determination of particulate organic carbon, nitrogen and phosphorus collected on filters, using a semi-automatic wet-oxidation method. *Mar Ecol Prog Ser* 180:289–295. <https://doi.org/10.3354/meps180289>.
110. Gardner WD, Richardson MJ, Carlson CA, Hansell D, Mishonov AV. 2003. Determining true particulate organic carbon: bottles, pumps and methodologies. *Deep Sea Res* 50:655–674. [https://doi.org/10.1016/S0967-0645\(02\)00589-1](https://doi.org/10.1016/S0967-0645(02)00589-1).
111. Yuan J, Li M, Lin S. 2015. An improved DNA extraction method for efficient and quantitative recovery of phytoplankton diversity in natural assemblages. *PLoS One* 10:e0133060. <https://doi.org/10.1371/journal.pone.0133060>.
112. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A* 108:4516–4522. <https://doi.org/10.1073/pnas.1000080107>.
113. Bråte J, Logares R, Berney C, Ree DK, Klaveness D, Jakobsen KS, Shalchian-Tabrizi K. 2010. Freshwater Perkin Sea and marine-freshwater colonizations revealed by pyrosequencing and phylogeny of environmental rDNA. *ISME J* 4:1144–1153. <https://doi.org/10.1038/ismej.2010.39>.
114. Lin S, Zhang H, Zhuang Y, Tran B, Gill J. 2010. Spliced leader-based meta-transcriptomic analyses lead to recognition of hidden genomic features in dinoflagellates. *Proc Natl Acad Sci U S A* 107:20033–20038. <https://doi.org/10.1073/pnas.1007246107>.
115. Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 10:996–998. <https://doi.org/10.1038/nmeth.2604>.
116. Huang H. 2020. ggcor: extended tools for correlation analysis and visualization. R package version 0.97. <https://rdrr.io/github/houyunhuang/ggcor/>.
117. Li D, Luo R, Liu C-M, Leung C-M, Ting H-F, Sadakane K, Yamashita H, Lam T-W. 2016. MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* 102:3–11. <https://doi.org/10.1016/j.ymeth.2016.02.020>.
118. Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22:1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>.
119. Pertea G, Huang X, Liang F, Antonescu V, Sultana R, Karamycheva S, Lee Y, White J, Cheung F, Parvizi B, Tsai J, Quackenbush J. 2003. TIGR gene indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics* 19:651–652. <https://doi.org/10.1093/bioinformatics/btg034>.
120. Yu L, Li T, Li L, Lin X, Li H, Liu C, Guo C, Lin SJD. 2020. SAGER: a database of Symbiodiniaceae and algal genomic resource. *Database* 2020. <https://doi.org/10.1093/database/baaa051>.
121. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>.