Diversity and Spatial Distribution of Chromophytic Phytoplankton in the Bay of Bengal Revealed by RuBisCO Genes (rbcL)

Laxman Pujari1,2, Chao Wu1,3, Jinjun Kan4*, Nan Li5, Xingzhou Wang1,2, Guicheng Zhang1,2, Xiaomei Shang1,2, Min Wang6, Chun Zhou7 and Jun Sun1,2*

1 Research Center for Indian Ocean Ecosystem, Tianjin University of Science and Technology, Tianjin, China, 2 Tianjin Key Laboratory of Marine Resources and Chemistry, Tianjin University of Science and Technology, Tianjin, China, 3 Institute of Marine Science and Technology, Shandong University, Qingdao, China, 4 Stroud Water Research Center, Avondale, PA, United States, 5 Key Laboratory of Environment Change and Resources Use in Beibu Gulf, Nanning Normal University, Nanning, China, 6 College of Marine Life Science, Ocean University of China, Qingdao, China, 7 Key Laboratory of Physical Oceanography/CIMST, Ocean University of China, Qingdao, China

Phytoplankton are the basis of primary production and play important roles in regulating energy export in marine ecosystems. Compared to other regions, chromophytic phytoplankton are considerably understudied in the Bay of Bengal (BOB). Here, we investigated community structure and spatial distribution of chromophytic phytoplankton in the BOB by using RuBisCO genes (Form ID rbcL). High throughput sequencing of rbcL genes revealed that diatoms, cyanobacteria (Cyanophyceae), Pelagophyceae, Haptophyceae, Chrysophyceae, Eustigmatophyceae, Xanthophyceae, Cryptophyceae, Dictyochophyceae, and Pinguiophyceae were the most abundant groups recovered in the BOB. Abundances and distribution of diatoms and Pelagophyceae were further verified using quantitative PCR analyses which showed the dominance of these groups near the Equator region (p < 0.01) where upwelling was likely the source of nutrients. Further, redundancy analysis demonstrated that temperature was an important environmental driver in structuring distributions of Cyanophyceae and dominant chromophytic phytoplankton. Morphological identification and quantification confirmed the dominance of diatoms, and also detected other cyanobacteria and dinoflagellates that were missing in our molecular characterizations. Pearson’s correlations of these morphologically identified phytoplankton with environmental gradients also indicated that nutrients and temperature were key variables shaping community structure. Combination of molecular characterization and morphological identification provided a comprehensive overview of chromophytic phytoplankton. This is the first molecular study of chromophytic phytoplankton accomplished in the BOB, and our results highlight a combination of molecular analysis targeting rbcL genes and microscopic detection in examining phytoplankton composition and diversity.

Keywords: Bay of Bengal, winter monsoon, rbcL gene, chromophytic phytoplankton, high throughput sequencing, morphological identification
INTRODUCTION

Characterization of primary productivity in oceans has been mainly focused in euphotic zones (Sigman and Hain, 2012). Marine phytoplankton are important primary producers in euphotic zones and carry out more than half of the photosynthetic carbon fixation in oceans (Falkowskia et al., 1998). Therefore, it is vital to investigate the detailed composition of phytoplankton and study their spatiotemporal distributions in marine environments. Phytoplankton community structures are traditionally characterized by a suite of approaches including light microscopy (Savin et al., 2004; Vaulot et al., 2008; Joo et al., 2010), flow cytometry (Urbach and Chisholm, 1998; Collier, 2000; Cellamare et al., 2010), and pigment analysis (Bidigare and Ondrusek, 1996; Peeken, 1997). Recently, molecular methods have been applied to decipher detailed phytoplankton communities including using small subunit ribosomal RNA genes (18S rRNA) (Read et al., 2014; Needham et al., 2018). However, because of their evolutionarily conserved nature and multiple copies, rRNA genes may not provide high resolution and sometimes fail to identify organisms at the species or even at the genus level (Pochon et al., 2013; Fletcher et al., 2017). In contrast, functional genes (mostly single copy per cell) are based on selected functional aspects of organisms and likely achieve greater resolution than ribosomal RNA genes (Imhoff, 2016). Functional genes are also involved in major element metabolisms (e.g., carbon and nitrogen) and therefore provide a broad view of both genetic and functional diversity of phytoplankton, which are potentially linked to biogeochemical cycling in oceans.

Currently, functional gene markers such as rbcL (large subunit of RuBisCO) (Samanta and Bhadury, 2014) have been increasingly used and provided improved resolution in characterizing phytoplankton assemblages across a variety of environments. RuBisCO catalyzes the assimilation of carbon dioxide into organic carbon via the Calvin-Benson Cycle and as a fact, it has been found to be the most abundant enzyme on the Earth (Chase et al., 1993). It is principally involved in sequestering carbon dioxide from environments and reductively assimilating them into organic carbon and cellular biomass (Andersson, 1996; Yoon et al., 2000; Tabita et al., 2007). RuBisCO has four different forms (I, II, III, and IV): Form I (L8S8) contains eight large (L8) and eight small subunits (S8), and exists in all plants and some bacteria, Form II has two large units (L2) and is found in Dinoflagellates and some photosynthetic bacteria, Form III has been identified in some archaea, while form IV [also called RuBisCo-like Protein (RLP)] is widely distributed in variety of microbial groups (Kitano et al., 2001; Tabita et al., 2008). Compared to small subunits, large subunits (LSU) encoded by rbcL gene is widely distributed in all four forms of RuBisCO, and it contains more conserved sequences meriting it a suitable gene marker for phylogenetic identifications (Xu and Tabita, 1996). Previous studies based on rbcL LSU have recognized its significance in deciphering community structures of chromophytic phytoplankton and their wide distributions in different ecological niches and geographical locations in global oceans and rivers including Tampa, FL, United States and Southeastern Gulf of Mexico (Pichard et al., 1993), Mid-Atlantic Ridge and seas around Japan (Elsaied and Naganuma, 2001), California upwelling and English Channel (Bhadury and Ward, 2009), Station ALOHA in North Subtropical Gyre (Li et al., 2013), Sundarbans Mangroves reserve forest (Samanta and Bhadury, 2014), Northern South China Sea (Li et al., 2016), and Bering Sea (Endo et al., 2015). Specifically, Form I RuBisCO gene was further divided into two lineages, green and red where most of the non-green phytoplankton contain ID RuBisCO and they were termed as chromophytic phytoplankton (Tabita, 1999). By using the form ID rbcL gene, major groups of chromophytic phytoplankton such as Bacillariophyceae, Haptophyceae, Pelagophyceae, Cryptophyceae, and Bolidiophyceae have been identified (Samanta and Bhadury, 2016). However, up to date, there is no molecular characterization available for rbcL genes and their distributions in the BOB.

Traditionally, the BOB is considered less biologically productive compared to its counterpart, the Arabian Sea. As the BOB is one of the largest tropical basins that is landlocked in the north, it is frequently influenced by semi-annual monsoons which create distinct seasonal changes in physicochemical parameters (Prasanna Kumar et al., 2004). As the Bay is located in a tropical region, growth of phytoplankton is usually impacted by monsoonal cloud cover and the presence of suspended sediments resulting from river discharge that may inhibit light availability for primary production (Gomes et al., 2000; Kumar et al., 2007). Additionally, vertical processes transporting nutrients into euphotic layers are generally weak in the BOB, which is likely due to strong stratification and weak upwelling (Krishna and Sastry, 1985). However, cyclones and mesoscale eddies can disturb stratification of upper layer and bring deep nutrients to the euphotic zone (Vinayachandran and Mathew, 2003; Prasanna Kumar et al., 2004; Muraleedharan et al., 2007; Reddy et al., 2008; Prasannakumar et al., 2010; Maneesha et al., 2011; Chen et al., 2013). Furthermore, BOB is characterized by a large seasonal influx of fresh water, which makes the upper water layer greatly less saline (McCready et al., 1993; Subramanian, 1993; Kumar and Unnikrishnan, 1995; Vinayachandran and Mathew, 2003). This seasonal freshwater input is from the neighboring rivers or from excess precipitation over evaporation. A recent study showed that runoff from medium to heavy rainfall could increase phytoplankton concentration by fourfold (Maneesha et al., 2011). The BOB is also influenced by severe tropical cyclones, which occur annually during inter-monsoon periods of April–May and October–November.

The dynamic processes in the southern BOB including strong currents, intense mixing, and upwelling make BOB one of the active region of the northern Indian Ocean (Vinayachandran et al., 2004; Thushara et al., 2019). Southern BOB experiences weak salinity stratification which results in a deeper mixed layer. Moreover, monsoon circulation leads to prominent chlorophyll blooms in coastal and open ocean zone of the southern BOB. In summer, high Chl-a of southern BOB is linked to coastal upwelling of nutrients. The Southwest Monsoon Current (SMC) intrudes into southern BOB from Indian and Sri Lankan coasts which carries biologically rich water and supports elevated level of Chl-a (Vinayachandran et al., 2013). Phenomena of upwelling continue in winter monsoon at the southwestern open ocean.
zone of BOB which leads to the formation of moderate blooms (Vinayachandran and Mathew, 2003). However, biophysical interaction in the BOB is limited, probably due to the scarcity of observational data. Finally, BOB also experiences large seasonal freshening and semiannual reversal of the upper-layer circulation, as well as global climate events such as El Nino and Indian Ocean Dipole (IOD) (Shankar et al., 1996; Shetye et al., 1996; Kumar et al., 2007). Altogether, the BOB provides a unique environment for investigating the composition and we hypothesize that distribution of chromophytic phytoplankton and morphologically identified phytoplankton are responding to local physicochemical environmental gradients such as upwelling in southern part of the BOB with presence of different currents and freshwater influx from rivers with monsoonal precipitation in the northern BOB.

Up to date, most studies on phytoplankton in BOB have been largely restricted to satellite information derived from ocean color imagery or morphologically identified phytoplankton community composition. Although lacking detailed taxonomic information, satellite-derived phytoplankton composition provides reasonable resolution and explains the forces driving spatial distribution in BOB (Nagamani et al., 2011). Instead, microscopic approaches are able to reveal the community composition of phytoplankton with respect to spatial distribution in BOB at a given time (Boonyapiwat et al., 2008; Paul et al., 2008). Previous studies carried out in the BOB using morphologically identified phytoplankton showed diatoms as the major group which outnumbered other groups such as Chlorophyta, Pyrrophyta, Chrysophyta, and Cyanophyta (Paul et al., 2008). Further, this study also showed that morphological identified phytoplankton composition and abundance were higher in southern than in northern BOB with a higher concentration of phytoplankton present in surface water (upper 20 m depth) (Paul et al., 2008). However, identification of phytoplankton through a microscope is very challenging and some species (Synechococcus or Prochlorococcus) may escape from microscopic detection due to their pico size. Additionally, overestimation of certain phytoplankton species has been observed in previous studies (Palińska and Surosz, 2008), where conspecific individuals were misidentified as different species because of their phenotypic variations. In contrast, DNA-based molecular methods provide more feasible techniques and a large dataset of sequence repositories can be generated from environmental samples. As a result, molecular techniques including high throughput sequencing have become powerful tools for phytoplankton identification (Ebenezer et al., 2012; Xiao et al., 2014). However, very limited information on the molecular composition of phytoplankton in BOB is available, which represents a significant knowledge gap.

The main objectives of this study were: (1) to investigate community composition of chromophytic phytoplankton based on high throughput sequencing of *rbcl* genes (form ID), and (2) to understand their spatial distribution in the BOB and how they respond to environmental gradients. The high throughput sequencing analyses provided comprehensive insights into the community structure of chromophytic phytoplankton. In addition, quantitative PCR analysis of diatoms and Pelagophyceae confirmed abundances of these groups under the influence of physicochemical parameters in BOB. Finally, since *rbcl* genes form ID mainly covers algal groups including Stramenopiles, Rhodophyta and Haptophytes (Tabita et al., 2008), microscopic identification was applied to all water samples. Morphological identifications of phytotplankton provided supplementary information and revealed groups that were not detected in our molecular characterization.

**MATERIALS AND METHODS**

**Study Area and Sample Collection**

Surface water samples were collected in BOB during a cruise aboard research vessel Dongfang II in the winter monsoon season from November 3rd to December 5th, 2016. Surface water samples were collected from 12 selected stations (Figure 1) by a rosette multi-sampler mounted with probes and sensors measuring conductivity, temperature, and depth (Sea-Bird SBE 911Plus, Sea-Bird Electronics, United States). Subsamples (100 ml) were saved in HCL-rinsed bottles and stored at 4°C for nutrient analysis by a Technicon AA3 Auto-Analyzer (Bran+ Luebbe, Norderstedt, Germany) for phosphate (PO$_4^{3-}$), ammonia (NH$_3$), nitrate (NO$_3^-$), nitrite (NO$_2^-$), and silicate (SiO$_4^{2-}$). Further, we vacuum filtered (<10 mm Hg) 500 ml of seawater for chlorophyll a (Chl-a) measurement on 25 mm GF/F filter membranes (Waterman, Florham Park, NJ, United States). After filtration, the filters were placed in aluminum foil and stored at −20°C in the dark until further analysis. Chl-a was extracted with 90% acetone and kept in a refrigerator at 4°C overnight, and was then analyzed using a fluorometer (CHL NA, Model # 046, Turner designs, San Jose, CA, United States). For molecular analysis, 2–3 L of seawater was filtered from all stations with 0.22-µm GFF filter (Millipore, Eschborn, Germany) using a peristaltic pump. The filters were then flash frozen and stored in liquid nitrogen. Upon returning to the laboratory, the filters were transferred to −80°C deep freezer until DNA extraction.

For morphological identification of phytoplankton, we collected 1 L of water from each station and immediately fixed with buffered formalin (final concentration 3%). The samples were stored in a dark and cool place until further analysis. Upon returning to the laboratory, the samples were re-suspended by gentle shaking. Phytoplankton taxa were identified and quantified under a light microscope (Motic AE2000, Xiamen, China) at magnifications of 200 or 400× following the protocol described by Utermöhl (Utermöhl, 1958).

**DNA Extraction and Form ID rbcl Gene Amplification**

Genomic DNA was extracted using DNeasyPowerWater Kits according to manufacturer’s instructions (QIAGEN, Hilden, Germany). DNA quantity and quality were checked on 1% agarose gel electrophoresis and evaluated via A260/A280 ratio using a Nanodrop 2000 Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, United States). Form ID *rbcl* gene fragments (554 bp) were amplified with previously published *rbcl*
primers (Wawrik et al., 2002). PCRs were performed using a Veriti 9902 thermocycler (Applied Biosystems, Foster City, CA, United States) with each reaction containing 2 µL template DNA, 10 µL Premix Taq (Takara, Tokyo, Japan), 1 mM each primer and 6 µL of double-distilled water to make a final volume of 20 µL. PCR conditions were as follows: initial denaturation at 95°C for 5 min, 30 cycles of 95°C for 1 min, 56°C for 1 min, 72°C for 1 min, and a final extension at 72°C for 10 min. PCR reactions for each environmental sample were performed in triplicate, pooled together and subsequently purified using Universal DNA purification kits (Tiangen Biotech, Beijing, China) following the manufacturer’s instruction. All libraries were constructed and sequenced via paired-end approach on an Illumina MiSeq PE300 platform (Illumina, San Diego, CA, United States) at Allwegene Technology, Co., Ltd., Beijing, China. Subsequent image analysis, base calling and error estimation were performed using the Illumina Analysis Pipeline Version 2.6.

qPCR Analysis
The abundance of rbcL genes for diatoms (Bacillariophytes) and pelagophytes were estimated using qPCR on an ABI Step One Plus Real-Time PCR system (Applied Biosystems, Foster City, CA, United States). In order to quantify Diatoms, we used primers (Forward: 5’-GATGATGARAAATTACCT-3’ and Reverse: 5’-TAWGAACCTTTWACTTCCWCC-3’) specified by John et al. (2007) and for Pelagophyceae, primers (Forward: 5’-CRACACWTTATTARAGACTAAG-3’ and Reverse: 5’-ATTTGDCCACAGTGDATACCA-3’) were used Li et al. (2013). Reactions were performed in a 20 µl volume containing 1 µl of DNA, 0.4 µM of each primer and 10 µl of TB Green Premix Ex Taq II (TliRNaseH Plus, Takara, Tokyo, Japan). Thermal cycling conditions for qPCR reactions were: 95°C for 1.2 min; 45 cycles of 15 s at 95°C; 1 min at 56°C; 15 s at 95°C. Melting curves were run between 60 and 95°C. The specificity of the amplification product was verified by melting curve analysis. The R² for standard curves of Bacillariophytes and Pelagophyceae were 0.996 and 0.992, respectively, whereas efficiencies were 105.0 and 167.2%, respectively. Plasmid DNA with inserted rbcL gene was measured with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, United States) and used as standards. Copy numbers of the rbcL gene were then calculated based on the concentration of plasmid DNA and amplicon size as previously described (Okano et al., 2004; He et al., 2007).

Data Processing and Statistical Analysis
The raw sequence data was transformed into sequence reads by base calling. These reads were stored in FASTQ files which contained raw sequence data with respective sequencing quality. The raw sequence data were quality-filtered and analyzed through open source software QIIME (v1.8).
Further, paired reads were merged into full-length sequences by FLASH software (v1.2.7) and filtered with Trimmomatic software (v0.33). Sequences meeting the following three criteria were included in downstream analyses: (1) sequences with precise primers and bar-codes; (2) quality score > 30; and (3) sequences > 200 bp in length. The software package USEARCH (v1.8) was then used to further eliminate erroneous and chimeric sequences. After removing non-\textit{rbcL} sequence reads, sequences were clustered into operational taxonomic units (OTUs) at a 97\% similarity level using UCLUST (v1.2.22). Low-abundance OTUs (fewer than 2 reads, including singletons) were excluded from the subsequent analyses (Dickie, 2010). The remaining high-quality sequences were queried against the GenBank database at NCBI using local BLASTn. The MEGAN program (Huson et al., 2007) was used to assign BLAST hits to taxa in the NCBI database. The sequences obtained from this study have been deposited in NCBI Sequence Read Archive with Accession No. SUB4422106.

Richness estimator (Chao1) and diversity indices (Shannon and Simpson) were calculated using QIIME v1.8. Gene abundances were calculated and plotted with Golden Software Surfer 11 (Golden Software, United States). Redundancy analysis (RDA) (Canoco, v 4.5) was performed to examine correlations between OTU abundance and environmental parameters. The number of sequences in each OTU was log transformed before conducting RDA. Pearson Correlation analysis (PAST, Version-3.23, Natural History Museum, University of Oslo, Norway) was performed for testing significant differences between gene copy numbers of diatoms and pelagophytes with environmental parameters including nutrients, chlorophyll, temperature, and salinity. Pearson correlation analysis was also used for to obtain significant difference between morphologically identified phytoplankton and environmental variables. The differences of various environmental parameters as well as qPCR detected \textit{rbcL} gene abundances of chromophytic phytoplankton between Northern BOB and Southern BOB were evaluated by \textit{t}-test using Excel (Command TTEST, two-tailed).

**RESULTS**

**Environmental Parameters**

Environmental parameters are listed in Table 1 and detailed contour maps for environmental parameter gradients are presented in Figure 2. Phosphate, ammonium, nitrate, nitrite, and silicate concentrations ranged from 0.09 to 0.574 \(\mu\)M (average 0.048 \pm 0.03 \(\mu\)M), 0.457 to 1.571 \(\mu\)M (average 0.86 \pm 0.30 \(\mu\)M), 0.021 to 0.200 \(\mu\)M (average \(\pm 0.78 \mu\)M), 0.021 to 0.2 \(\mu\)M (average 0.076 \pm 0.05 \(\mu\)M) and 0.790 to 1.379 \(\mu\)M (average 1.03 \pm 0.18 \(\mu\)M), respectively (Table 1). Among all measured nutrients, phosphate \((p < 0.01)\) was fairly in low concentrations at all stations and the highest concentration was observed at station BOB 10 (Figure 2). In contrast, ammonium, nitrate, and silicate concentrations were generally higher in BOB than stations near the Equator region \((p > 0.05)\) (Figures 2b–d). Surface water temperature in BOB fairly varied from northern to southern near the Equator with an average of 28.6 \pm 0.34\(^\circ\)C \((p > 0.05)\) (Table 1). Higher salinities were observed in the southern region, which ranged from 32.38 (BOB-1) to 34.68 (BOB-8) with an average of 33.5 \pm 0.69 \((p < 0.01)\). Geographically, the BOB contained lower surface salinity and temperature than the equator region because of freshwater input from adjacent rivers in the northern BOB and massive rainfall during the prevailing monsoon. Chl-\(a\) concentrations ranged from 0.017 to 0.560 \(\mu\)g L\(^{-1}\) with an average of 0.24 \pm 0.16 \(\mu\)g L\(^{-1}\) (Table 1). Chl-\(a\) concentrations showed a similar trend as salinity and temperature: increased from northern to the southern region of the BOB \((p < 0.05)\) (Figures 2a–c).

**Community Composition and Phylogenetic Analysis of Chromophytic Phytoplankton**

A Total of 385,077 raw sequences were generated from our samples and 152,556 clean sequences were included in our downstream analysis (Table 2). The rarefaction curves were shown in Supplementary Figure 1, and most curves were plateaued. A total of 662 OTUs were retrieved after clustering at a 97\% similarity level (Table 2). Shannon-Weiner H index was calculated for all the stations and the highest value occurred at station BOB-12 (6.24) whereas the lowest value was recorded at BOB-1 (3.24) (Table 2). OTU richness was calculated using non-parametric Chao1 estimator. The highest Chao1 value was observed at BOB-2 whereas the lowest Chao1 value was recorded at BOB-10 (Table 2). The species richness and Good’s coverage were also listed in Table 2.

All major chromophytic phytoplankton groups containing form ID \textit{rbcL} genes such as Bacillariophyceae, Haptophyceae, Pelagophyceae, Chrysophyceae, Cryptophyceae, Dictyochophyceae, Cyanophyceae, Euglenoza, Chrysophyceae, and Bolidomonas were detected within class Pelagophyceae (4 OTU), and among them, \textit{Bolidomonas}, a genus containing picoplanktonic flagellated algae, was also recovered from station BOB-12. \textit{Peridinium} was another genus that occurred at station BOB-12, which is symbiotic with diatoms and was discussed further in this study. Four genera were detected within class Pelagophyceae (4 OTU), and among them, \textit{Pelagomonas} was dominant at most of the stations \((e.g., \text{BOB-1})\) whereas lowest abundance was recovered from station BOB-8. Other genera clustered in Pelagophyceae were \textit{Pelagococcus}, \textit{Aureococcus}, and \textit{Sarcinochrysis}. Regarding Haptophyceae, it was represented by 12 OTUs which were dominated by genera \textit{Chrysocromulina} and \textit{Calyptraphera}. Chrysocromulina was
TABLE 1 | Temperature (Temp), salinity (S), chlorophyll a (Chl-a) concentration and dissolved inorganic nutrients (PO\textsubscript{4}\textsuperscript{3−}, NH\textsubscript{4}\textsuperscript{+}, NO\textsubscript{3}−, NO\textsubscript{2}−, and SiO\textsubscript{2}−) measured on the surface layer at the survey stations.

| Stations | Longitude (°E) | Latitude (°N) | Temperature (°C) | Salinity | Chl-a (µg L\textsuperscript{−1}) | PO\textsubscript{4}\textsuperscript{3−} (µM L\textsuperscript{−1}) | NH\textsubscript{4}\textsuperscript{+} (µM L\textsuperscript{−1}) | NO\textsubscript{3}− (µM L\textsuperscript{−1}) | NO\textsubscript{2}− (µM L\textsuperscript{−1}) | SiO\textsubscript{2}− (µM L\textsuperscript{−1}) |
|----------|----------------|---------------|------------------|----------|-------------------------------|------------------|-------------------|------------------|------------------|------------------|
| BOB-1    | 89.499         | 17.014        | 28.03            | 32.38    | 0.167                         | 0.023            | 0.821             | 0.671            | 0.029            | 1.000            |
| BOB-2    | 90.997         | 14.998        | 28.92            | 32.66    | 0.017                         | 0.000            | 0.779             | 1.595            | 0.121            | 1.379            |
| BOB-3    | 88.007         | 15.489        | 28.35            | 33.00    | 0.195                         | 0.019            | 0.721             | 0.467            | 0.050            | 0.768            |
| BOB-4    | 88.497         | 15.001        | 28.72            | 33.08    | 0.176                         | 0.074            | 1.364             | 1.980            | 0.200            | 1.190            |
| BOB-5    | 87.001         | 15.005        | 28.77            | 33.04    | 0.277                         | 0.000            | 0.564             | 0.111            | 0.042            | 1.129            |
| BOB-6    | 87.964         | 12.987        | 28.79            | 33.77    | 0.093                         | 0.045            | 0.943             | 0.459            | 0.064            | 0.939            |
| BOB-7    | 87.998         | 10.509        | 28.93            | 34.03    | 0.391                         | 0.048            | 1.529             | 1.562            | 0.136            | 1.258            |
| BOB-8    | 91.003         | 05.004        | 28.82            | 33.68    | 0.150                         | 0.045            | 0.843             | 0.628            | 0.086            | 1.020            |
| BOB-9    | 87.994         | 05.000        | 28.82            | 33.68    | 0.150                         | 0.045            | 0.843             | 0.628            | 0.086            | 1.020            |
| BOB-10   | 87.497         | 05.001        | 28.72            | 33.08    | 0.176                         | 0.074            | 1.364             | 1.980            | 0.200            | 1.190            |
| BOB-11   | 85.610         | 05.000        | 28.84            | 34.38    | 0.564                         | 0.081            | 0.850             | 0.578            | 0.050            | 0.888            |
| BOB-12   | 83.000         | 05.003        | 28.88            | 33.78    | 0.500                         | 0.081            | 0.850             | 0.578            | 0.050            | 1.107            |

FIGURE 2 | Contour maps demonstrating environmental parameter measurements on the surface water in the BOB. (a) Temperature (°C), (b) salinity, (c) Chl-a (µg/L), (d) phosphate (PO\textsubscript{4}\textsuperscript{3−}), (e) ammonia (NH\textsubscript{4}\textsuperscript{+}), (f) nitrate (NO\textsubscript{3}−), (g) nitrite (NO\textsubscript{2}−), (h) silicate (SiO\textsubscript{2}−).

TABLE 2 | Summary of OTU numbers and diversity indices for rbcL sequences recovered from the BOB.

| Stations       | BOB-1 | BOB-2 | BOB-3 | BOB-4 | BOB-5 | BOB-6 | BOB-7 | BOB-8 | BOB-9 | BOB-10 | BOB-11 | BOB-12 |
|----------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|--------|--------|
| No. of OTUs    | 298   | 350   | 293   | 346   | 262   | 286   | 323   | 323   | 324   | 280    | 313    | 332    |
| Raw tags       | 37992 | 24636 | 20597 | 20673 | 33122 | 32092 | 34223 | 41633 | 26069 | 45731  | 42549  | 26760  |
| Final Tags     | 12713 | 12713 | 12713 | 12713 | 12713 | 12713 | 12713 | 12713 | 12713 | 12713  | 12713  | 12713  |
| Species richness| 290   | 343   | 288   | 339   | 254   | 278   | 315   | 316   | 316   | 275    | 304    | 328    |
| Chao1          | 335.08| 401.51| 324.67| 385.54| 349.98| 361.6 | 400.03| 378.56| 374.69| 309.02 | 369.23 | 359.4  |
| Good's Coverage| 0.99  | 0.99  | 1     | 0.99  | 0.99  | 0.99  | 0.99  | 0.99  | 0.99  | 1      | 0.99   | 1      |
| Shannon-Weiner | 3.25  | 4.94  | 4.32  | 4.53  | 4     | 3.83  | 4.13  | 4.6   | 3.99  | 4.64   | 4.14   | 6.24   |

dominant at station BOB-12 whereas Calyptrosphaera was dominant at BOB-5. Among class Cyanophyceae (5 OTUs), Prochlorococcus was the only dominant genus in recovered sequences. Further, Dictyochophyceae clade contained 6 OTUs and was represented by Apedinella at station BOB-12. Other classes included Cryptophyceae (2 OTUs), Eustigmatophyceae (1 OTU), Xanthophyceae (2 OTUs), and Chrysophyceae (3 OTUs) (Figure 3).
Quantification of rbcL Genes From Diatoms (Bacillariophytes) and Pelagophytes
Abundances of rbcL genes (copies L$^{-1}$) for diatoms (Bacillariophytes) and pelagophytes and their spatial distribution are shown in Figure 5. Both diatoms and pelagophytes showed similar spatial patterns with higher copy numbers occurred near the Equator and lower in the northern BOB ($p < 0.01$) (Figure 5). The copy numbers of rbcL genes for diatoms ranged from $40.4 \pm 5.46 \times 10^4$ copies L$^{-1}$ (BOB-5) to $30.8 \pm 4.61 \times 10^5$ copies L$^{-1}$ (BOB-12). Pelagophyceae rbcL gene copy numbers varied across stations: the highest number ($11.4 \pm 3.33 \times 10^6$) copies L$^{-1}$ was found near the Equator (BOB-10) whereas the lowest was recorded ($32.2 \pm 4.79 \times 10^4$ copies L$^{-1}$) at station BOB-2.

Morphological Identification of Phytoplankton and Quantification
A total of 96 taxa of phytoplankton (morphologically identified phytoplankton) were identified through microscopic observation. Dinoflagellates and diatoms were observed to be the most diverse and abundant groups while cyanobacteria were identified as less diverse but predominant in some stations of the study area (e.g., BOB-9 and BOB-11) (Supplementary Table 1). Contour maps for abundance and spatial distribution of diatoms and dinoflagellates (identified using microscope) at BOB are shown in Figure 6. Diatoms rbcL gene ranged from $1.4 \times 10^4$ cells L$^{-1}$ (BOB-8) to $4.84 \times 10^5$ cells L$^{-1}$ (BOB-11). Whereas dinoflagellates, ranged from $1.8 \times 10^4$ cells L$^{-1}$ (BOB-1) to $1.42 \times 10^5$ cells L$^{-1}$ (BOB-12) (Supplementary Table 1).

Compared to other groups, Chrysophyta was less diverse and abundant, and its abundance ranged from $8 \times 10^3$ cells L$^{-1}$ (BOB-7, BOB-8, and BOB-12) to $1.6 \times 10^4$ cells L$^{-1}$ (BOB-4). Overall through microscopic identification, we identified major groups such as diatoms, dinoflagellates, cyanophytes, and chrysophytes. Among these, dinoflagellates were barely detected from our molecular characterization (Figure 4). In addition, *Trichodesmium thiebaultii* was successfully identified through microscopy as a major cyanobacterial group which was missing from molecular characterization. But due to the small size of picocyanobacterial cells (e.g., *Prochlorococcus* and *Synechococcus*) very likely we missed them in our microscopic identifications.

Correlations Between Phytoplankton Quantified Through Morphological Observation and qPCR Detected Gene Copy Number With Environmental Variables
Correlations between chromophytic phytoplankton communities and the associated environmental factors are shown by RDA plot (Figure 7). The environmental factors in the first two axes explained 55.1% of the total variances. Our RDA analysis indicated that among all environmental variables, temperature ($p < 0.05$) was the major factor determining the distribution of chromophytic phytoplankton (Figure 7). In addition, RDA also showed correlations of other environmental variables with chromophytic phytoplankton groups. For instance, Chrysophyceae positively responded to NO$_3^-$, temperature, and
FIGURE 4 | A neighbor-joining phylogenetic tree constructed based on rbcL amino acid sequences. The topology of the tree was inferred from 1000 bootstrap resampling, and bootstrap values greater than 50% were labeled with black dots at branches.
SiO$_2^-$). Further, Dictyochophyceae and Coscinodiscophyceae responded to PO$_3^{4-}$, while Cyanophyceae showed significant positive relationships with temperature, NO$_2^-$ and NH$_4^+$ (Figure 7). In contrast, Cryptophyceae, Pelagophyceae, and Bacillariophyceae were negatively correlated to NO$_3^-$, SiO$_2^-$, and temperature. Moreover, Pearson’s correlations of rbcL gene abundances and various physicochemical parameters were listed in Table 3. Diatoms and pelagophytes positively correlated to Salinity ($p < 0.01$), Chl-a ($p < 0.01$), PO$_3^{4-}$ ($p < 0.01$), and NO$_2^-$ ($p < 0.05$). Diatoms also showed a significantly positive relationship with temperature but Pelagophytes were negatively correlated with temperature. Negative correlations were also observed between SiO$_2^-$ with diatoms and pelagophytes. Further, Pearson’s correlation matrix of physicochemical parameters and morphologically identified phytoplankton indicated diatoms exhibited positive correlations with salinity and Chl-a, but a negative relationship with NH$_4^+$ (Table 4). As a group present in morphological identification but not recovered in our molecular characterization, dinoflagellates showed positive correlations with all of the physicochemical parameters besides NO$_3^-$. Cyanophyta was positively correlated with temperature and salinity but negatively correlated with PO$_3^{4-}$, NH$_4^+$, NO$_3^-$, and SiO$_2^-$. 

**DISCUSSION**

**Phylogenetic Characterization and Global Distribution of Chromophytic Phytoplankton in the BOB**

It is noticeable that few effort and studies have been conducted to characterize chromophytic phytoplankton communities in the BOB, especially at the molecular level. Recent studies primarily focused on phylogenetic analyses based on 16S rRNA genes in both coastal and open oceans as well as in oxygen minimum zone (OMZ) of the BOB (Bristow et al., 2017; Choi et al., 2017; Menezes et al., 2018; Rajpathak et al., 2018). To our best knowledge, no studies have investigated community structures of chromophytic phytoplankton based on functional genes such as rbcL. This study serves as the first report on the molecular population structure of photosynthetic organisms in BOB and provides baseline data for future studies.
FIGURE 7 | Redundancy analysis (RDA) for rbcL gene-based chromophytic phytoplankton community distribution and environmental factors. The black arrows represented different classes of phytoplankton whereas environmental variables were shown with red arrows. Correlation between environmental variables (Temperature, Salinity, PO$_{4}^{3-}$, Chl-a, NH$_4^{+}$, NO$_3^{-}$, NO$_2^{-}$, and SiO$_2^{3-}$) and RDA axes were shown by both length and angle of arrows. The two RDA axes explained 55.1% of total variation.

TABLE 3 | Pearson correlation matrix of qPCR-detected gene copy numbers and physicochemical parameters.

|          | Temperature | Salinity | chl-a     | PO$_{4}^{3-}$ | NH$_4^{+}$ | SiO$_2^{3-}$ | NO$_3^{-}$ | NO$_2^{-}$ |
|----------|-------------|----------|-----------|--------------|------------|--------------|------------|------------|
| Diatoms  | 0.367**     | 0.635**  | 0.646**   | 0.544**      | 0.066      | -0.212*      | -0.183     | -0.285*    |
| Pelagophytes | -0.288* | 0.544**  | 0.392**   | 0.742**      | -0.304*    | -0.305**     | -0.334**   | -0.249*    |

*p < 0.05; **p < 0.01.

Our high throughput sequencing data revealed major chromophytic groups including Bacillariophyceae, Haptophyceae, Pelagophyceae, Cyanophyceae and less abundant groups such as Cryptophyceae, Chrysophyceae, Eustigmatophyceae, Xanthophyceae, Dictyochophyceae, and Pinguioephyceae. These results were consistent with the results from previous investigations carried out in different environments, suggesting great homogeneity of chromophytic phytoplankton occurring in most marine ecosystems of the world (Pichard et al., 1993; Paul et al., 1999; Bhadury and Ward, 2009; Samanta and Bhadury, 2014; Li et al., 2016). The most diverse and predominant chromophytic phytoplankton group retrieved in our study, diatoms (Bacillariophyceae) are globally distributed and they constitute 20% of the global primary production (Nelson et al., 1995; Round et al., 2007). Undoubtedly, diatoms are the most successful algae present from freshwater to seawater, and from the surface to benthic habitats (Werner, 1977; Mann and Droop, 1996). Since there is no characterization of diatoms at the molecular level in the BOB, a few studies that performed morphological characterizations supported our observation of diversity and abundance of Bacillariophyceae (Paul et al., 2007; Boonyapiwat et al., 2008; Li et al., 2012; Manjumol et al., 2018).

Our phylogenetic analysis also revealed that *Bolidomonas* (Triparma)-like rbcL sequences shared high similarity (94%) with Bacillariophyceae rbcL sequences (Figure 3). With isolated *Bolidomonas* strains and by comparing nuclear, plastidial and
mtDNA gene markers, Ichinomiya et al. (2016) have combined *Bolidomonas* into *Triparma* and included both of them into *Parmales* (Bolidophyceae). The phylogenetic assessment suggested that *Parmales* (Bolidophyceae) was closely related to diatoms (Bacillariophyceae). The study also demonstrated that Bolidophyceae were ubiquitously distributed but only constituted a minor component of the phytoplankton community (Ichinomiya et al., 2016). Nonetheless, this is the first record of *Bolidomonas* (*Triparma*) from the BOB. Similarly, we recovered *Peridinium* (Dinophyceae)-like *rbcL* sequences which were also clustered with Bacillariophyceae in the phylogenetic tree (Figure 3). By a systematic investigation of SSU rDNA sequences, Inagaki et al. (2000) succeeded in recovering the closest sister relationship between two dinoflagellates *Peridinium balticum* and *Peridinium foliaceum*, and it was suggested that the ancestors of these two dinoflagellates engulfed a pennate diatom, and therefore showed a close affinity to diatom *rbcL* gene sequences.

Pelagophyceae was another dominant group observed in our chromophytic phytoplankton analysis. Among this class, *Pelagomonas calceolata* has been referred to as an open ocean species (Saunders et al., 1997). From the discovery of *Pelagomonas calceolata* to recent studies based on 18S rRNA genes, this species has been documented a global distribution in surface or subsurface waters in North Pacific Ocean (Andersen et al., 1993), South Pacific Ocean (Moon-van der Staya et al., 2000), and the Mediterranean Sea (Diez et al., 2001). *Pelagomonas* sequences were also recovered from the surface layer in the central North Pacific gyre of the Global Ocean Survey (Rusch et al., 2007). More recently, a study in the subarctic Pacific Ocean obtained *Pelagomonas* sequences from deep sediments, suggesting a wide vertical distribution of *Pelagomonas* from the euphotic zone to deep oceans (Worden et al., 2012). In this study, our functional gene (*rbcL*) analysis verified the presence and dominance of *Pelagomonas* species at BOB. This is the first record of *Pelagomonas calceolata* in BOB.

Further, considerable Haptophyceae sequences were retrieved from the BOB. Previous observations have demonstrated that these picoplanktonic flagellates are successful cosmopolitan picophytoplankton that inhabits a great variety of environments ranging from freshwater to marine habitats (Thomsen, 1994; Jordan and Chamberlain, 1997; Jordan et al., 2004; Edvardsen et al., 2011). Within Haptophyceae, most sequences were closely related to *Chrysolomonadina*, a ubiquitous genus that has been recovered in European waters (Parke et al., 1955; Wujek, 1986), South Africa, Australian coasts (Hallegraeff, 1983; Pienaar and Bandu, 1984), and in the Arctic region of Canada and Alaska as well (Manton, 1978; Kling and Kristiansen, 1983).

### Influence of Physicochemical Parameters on the Spatial Distribution of Chromophytic Phytoplankton

The BOB is a semi-enclosed tropical sea, which is highly influenced by seasonal monsoons. In addition to the freshwater influx from riverine discharge (Ganges River), monsoons bring a significant amount of fresh water through high precipitation. Moreover, ocean currents [e.g., East India Coastal Current (EICC) and Southwest Monsoon Current (SMC)] and eddies are able to erode stratification as well as bring upwelling nutrients to surface waters (Gomes et al., 2000; Schott and McCreary, 2001; Shankar et al., 2002). Such physical forces change the surface physicochemical parameters and consequently affect the composition and productivity of chromophytic phytoplankton. Our RDA analysis, qPCR results and microscopic cell counts confirmed that the spatial distribution of phytoplankton responded to changes in physicochemical parameters such as temperature and nutrients (Figure 7 and Tables 3, 4). Abundance and community structure of diatoms (Bacillariophyceae) responded to temperature but no consistent results were observed among different analyses. RDA analysis indicated that diatoms were negatively related to temperature (Figure 7), qPCR results indicated a positive relationship (Table 3), while no significant relationship was found between microscopic cell counts vs. temperature (Table 4). In general, the temperature is a key environmental parameter driving phytoplankton
distribution and activity (Patrick, 1948; Montagnes and Franklin, 2001). Inconsistent results from this study suggested that in addition to temperature, other environmental parameters also impact on diatom abundance and distribution at BOB. Smaller phytoplankton groups with large surface area to volume ratio can outcompete diatoms during nutrient limiting conditions and warm water temperatures (Read et al., 2014). Moreover, a laboratory-grown culture experiment has shown that diatoms were not an exception to the rule that cell size reduces with increase in temperature (Montagnes and Franklin, 2001), where size-selective grazing may become a limiting factor for phytoplankton (Peter and Sommer, 2013). Pelagophyceae showed similar distribution patterns and most of the genera recovered in this study have been reported being dominant in warm surface water (Agawin et al., 2000; Mitbavkar et al., 2015).

Another major chromophytic phytoplankton group responding to warm temperature was Cyanobacteria (mainly Prochlorococcus). Prochlorococcus is ubiquitously distributed from 40°N to 40°S, and from the surface to 150 m deep (Flombaum et al., 2013). Recent research based on functional genes (rbcL) verified the global distribution (involving Atlantic, Pacific, and the Indian Ocean) of Prochlorococcus and confirmed their distribution from the surface to DCM (Fernández-Pinos et al., 2017). Prochlorococcus generally proliferate in tropical and subtropical oligotrophic waters which are influenced by seasonality (Campbell et al., 1994; Campbell et al., 1997; Partensky et al., 1999). Our RDA analysis showed a strong positive correlation of Cyanophyceae with temperature. Different strains of Prochlorococcus responded differently to light availability and temperature and their population size declined below 15°C (Johnson et al., 2006). Our study was conducted during prevailing monsoon, which brought dense clouds and limited the availability of sunlight, and therefore influenced growth of Cyanobacteria especially Prochlorococcus. Further, we could not neglect the emerging OMZ in BOB and its impact on the distribution of Cyanobacteria. Recent studies have concluded that BOB, currently referred to as weak OMZ, is in the verge of becoming the next OMZ (Bristow et al., 2017). At South Pacific OMZ, Synechococcus were abundantly present in the surface oxic waters while Prochlorococcus were present in the subsurface low-oxygenated water (Lavin et al., 2010). Contrastingly, we found more prevalence of Prochlorococcus and very low abundances of Synechococcus sequences at surface water. This probably can be linked to a developing OMZ in the BOB. Large riverine network successfully transported organic matter to the northern part of the BOB that enhanced microbial respiration and resulted in less oxygen and nitrogen-based oxidants thereby producing weak OMZ and suppressed denitrification (Al Azhar et al., 2017; Bristow et al., 2017). Investigating interactions between Prochlorococcus and Synechococcus and their distribution at depths (including OMZs) will provide insightful information on primary producers in the BOB and is greatly needed. Although Cyanobacteria mostly favor warm and oligotrophic waters, they are also present in coastal areas where the nutrient is not a limiting factor. For example, they were found in coastal areas in Plume of the Rhone river, Mediterranean Sea (Veldhuis and Kraay, 1990) and Suruga Bay, Japan (Shimada et al., 1995). Lastly, the co-occurrence of Cyanobacteria with other chromophytic phytoplankton sequences suggests the existence of a possible symbiotic relationship. Jyothibabu et al. (2006) explained that during a spring inter-monsoon, nitrate limitation was pronounced at the surface water of the BOB and Cyanobacteria would benefit from symbiotic relationships with other algae or dinoflagellates. In fact, symbiotic relationships between Cyanobacteria and diatoms have been previously explored within studies undertaken in different parts of Indian Ocean such as the BOB, the Arabian Sea and the Andaman Sea (Kulkarni et al., 2010; Jabir et al., 2013; Manjumol et al., 2018).

In the BOB, possible nutrients sources originated from two regions: (1) northern riverine runoffs in northern BOB and (2) Southwest upwelling region near the Equator. Most abundant chromophytic phytoplankton groups including Pelagophyceae, Haptophyceae, and Bacillariophyceae showed higher abundances at both northern BOB and near the Equator region (Figures 3, 4), suggesting that high nutrient concentration likely support growth of these chromophytic phytoplankton. A previous study based on microscopic identification of phytoplankton found a high abundance of phytoplankton in northern BOB and deduced nutrient discharging into the BOB from riverine runoff as the factor driving phytoplankton distribution (Boonyapiwat et al., 2008). Meanwhile, during the northeastern monsoon, the southwestern part of the BOB experiences cyclonic gyre. The cold and dense waters would benefit from symbiotic relationships with other algae or dinoflagellates. In fact, symbiotic relationships between Cyanobacteria and diatoms have been previously explored within studies undertaken in different parts of Indian Ocean such as the BOB, the Arabian Sea and the Andaman Sea (Kulkarni et al., 2010; Jabir et al., 2013; Manjumol et al., 2018).

Phytoplankton Identified by High Throughput Sequencing vs. Microscopic Morphology

Morphological identification of phytoplankton is very challenging and extensive expertise is needed. Through microscopy analysis, we observed diatoms, dinoflagellates, chrysophytes, and cyanophytes were the major groups occurred in the BOB surface waters. Most of the previous studies in characterizing phytoplankton through morphological identification were restricted to coastal zones of the BOB, and very limited investigations were carried out in open oceans at the BOB (Paul et al., 2007). Boonyapiwat et al. (2008) conducted a survey during northwest monsoon and the results showed a high abundance of diatoms occurred in the northern BOB. Most sampling stations in this study were located in the coastal zone and northern BOB, and the abundance of diatoms was probably influenced by riverine nutrient input (Boonyapiwat et al., 2008). In our study, we found a higher abundance of
phytoplankton in the southern open ocean zone of the BOB which was likely impacted by SMC current and upwelling near Sri Lanka (Figures 5–7).

When we compared chromophytic phytoplankton sequences using high throughput sequencing to qPCR assays on morphologically identified phytoplankton, a significant difference was observed between gene copy numbers detected by qPCR vs. direct microscopic counts (Figures 5, 6) in diatoms. Specifically, morphologically identified diatoms ranged between $1.8 \times 10^5$ and $4.8 \times 10^5$ cells L$^{-1}$ whereas qPCR derived diatom (Bacillariophyceae) gene copy numbers showed as low as $4.04 \times 10^2$ to as high as $3.08 \times 10^6$ copies L$^{-1}$. One possible explanation was each picoplanktonic diatom cell contained multiple copies of \textit{rbcL} genes (Pichard et al., 1993). For instance, in addition, two copies of \textit{rbcL} genes spanning on their genomes, diatom cells may contain up to 2000 chloroplast-encoding \textit{rbcL} genes (Douglas, 1988). Therefore, the discrepancy observed between gene copies and morphologically counted cell numbers could be justified as due to the presence of multiple copies of \textit{rbcL} genes in diatom cells.

Further, we observed that some species and groups of chromophytic phytoplankton were missing from our molecular characterization. For instance, \textit{Trichodesmium thiebaultii} was identified as the dominant cyanobacterial species under microscopic counts but it was completely missing in our high throughput sequencing analysis (Figure 3 and Supplementary Table 1). This could most likely be attributed to the bias introduced by the primers used in our study (form ID \textit{rbcL}) for PCR amplification. Evolution of RuBisCO-encoding \textit{rbcL} genes gave rise to different forms used for identifying different groups of phytoplankton. Form IA, IB, IC, and ID encode green and red like RuBisCOs and they exist in different groups of phytoplankton (Watson and Tabita, 1997). Form IB occurred in green plants lineage and some Cyanobacteria, whereas form ID was widely distributed in chromophytic algal groups such as diatoms, prymnesiophytes, pelagophytes, and chrysophytes (Paul et al., 1999). RuBisCO in \textit{Trichodesmium} was likely encoded by \textit{rbcL} form IB (Paul et al., 1999; Chen et al., 2004; Tomitani et al., 2006) and was, therefore, missing from our PCR amplification. Our morphological observation also identified dinoflagellates as another major phytoplankton group (Supplementary Table 1) which agreed with a previous study (Madhupratap et al., 2003). Previous studies carried out in the Chesapeake Bay (MD, United States), the Jiaozhou Bay (Qingdao, China), and the Bay of Biscay (western coast of France) have shown that morphologically identified diatoms, as well as dinoflagellates, were consistently dominant and diverse in different ecosystems (Marshall and Lacouture, 1986; Shen, 2001; Smythe-Wright et al., 2014). However, dinoflagellates were under-represented in our molecular analysis. This could perhaps be explained by Morse's finding (1995) that RuBisCO in dinoflagellates was possibly encoded by RuBisCO form II instead of form I. Also, dinoflagellates RuBisCO were perhaps encoded by genomic DNA and not by chloroplast DNA (Morse et al., 1995). A recent study undertaken to investigate ID \textit{rbcL} gene diversity spanning seven different seas around the world did not recover dinoflagellates sequences (Samanta and Bhadury, 2016). As a single gene marker, form ID \textit{rbcL} gene covers bacteria (Alpha, Beta, and Gamma-proteobacteria) and micro-phytoplankton such as Bacillariophyceae, Haptophyceae, Pelagophyceae, Cryptophyceae, and Bolidiophyceae (Samanta and Bhadury, 2016), but may have missed other groups including dinoflagellates and cyanobacteria. Our results highlight the necessity of a combination of molecular analysis targeting \textit{rbcL} genes and microscopic detection in examining phytoplankton composition and diversity.

CONCLUSION

Our observations on chromophytic phytoplankton demonstrate the diversity of these primary producers and their spatial distribution in the BOB. Environmental variables such as temperature and nutrients were the most significant factor which influenced the community structure of chromophytic phytoplankton. Primarily we observed: (1) greater diversity and community composition of chromophytic phytoplankton in the BOB that has never been studied before, (2) morphological identification in combination with molecular characterization provided insights of overall community compositions, and (3) spatial distribution of chromophytic phytoplankton through \textit{rbcL} genes and morphologically identification showed variation which was likely impacted by coastal freshwater input, winter monsoons, and upwelling. This study serves as the first molecular investigation of phytoplankton communities in the BOB. Inconsistent results were obtained via \textit{rbcL} gene characterization vs. morphological identification suggesting that a careful coupling of molecular analysis and microscopic identification will be needed before the importance of these chromophytic phytoplankton, and their global distribution, are adequately understood.

AUTHOR CONTRIBUTIONS

LP designed the research, wrote the manuscript, and carried out the molecular and statistical analysis. CW carried out the sample collection and experimental analysis. NL, XS, MW, CZ, and GZ helped in the experimental analysis. JK drafted and proofread the manuscript. XW carried out the morphological identification of the phytoplankton. SJ designed the research and drafted the manuscript.

FUNDING

This research was financially supported by the National Key Research and Development Program of China (Grant Nos. 2015CB954002 and 2016YFC1402600) and the National Natural Science Foundation of China (Grant Nos. 41876134, 41676112, and 41276124), the University Innovation Team Training Program for Tianjin (TD12-5003), the Tianjin 131 Innovation Team Program (Grant No. 20180314), the Changjiang Scholar Program of Chinese Ministry of Education (T2014253) to JS, the Endowment Fund from the Stroud Water Research Center. 

Frontiers in Microbiology | www.frontiersin.org 13 July 2019 | Volume 10 | Article 1501
to JK, the Guangxi Natural Science Foundation (Grant Nos. 2017GXNSFAA198081 and 2018JJD130026), and 100 Program of the Introduction of High-level Overseas Talents for Colleges and Universities in Guangxi to NL.

ACKNOWLEDGMENTS

We thank Prof. Wei Zhao from the Ocean University of China for providing the hydrographic (CTD) data. We also gratefully acknowledge the crew of the R/V "Dongfang2" for their assistance and all the participants for their inputs and contributions during the cruise.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.01501/full#supplementary-material

REFERENCES

Agawin, N. S., Duarte, C. M., and Agusti, S. (2000). Nutrient and temperature control of the contribution of picoplankton to phytoplankton biomass and production. Limnol. Oceanogr. 45, 591–600. doi: 10.4319/lo.2000.45.8.1891a

Al Azhar, M., Lachkar, Z., Lévy, M., and Smith, S. (2017). Oxygen minimum zone contrasts between the Arabian Sea and the Bay of Bengal implied by differences in remineralization depth. Geophys. Res. Lett. 44, 11106–11114. doi: 10.1002/2017GL075317

Andersen, R. A., Saunders, G. W., Paskind, M. P., and Sexton, J. P. (1993). Ultrastructure and 18s rRNA gene sequence for pelagomonas calceolata gen. nov. Et Sp. nov. of a new descriptive algal class, the pelagophyceae classis nov. J. Phycol. 29, 701–715. doi: 10.1111/j.0022-3646.1993.00701.x

Anderson, I. (1996). Large structures at high resolution: the 1.6 Å crystal structure of the ribulose-b-5-phosphate carboxylase/oxygenase complex with 2-carboxyraminobitatls phosphate. J. Mol. Biol. 259, 160–174. doi: 10.1006/jmbi.1996.0310

Bhadury, P., and Ward, B. B. (2009). Molecular diversity of marine phytoplankton communities based on key functional genes 1. J. Phycol. 45, 1335–1347. doi: 10.1111/j.1529-8817.2009.00766.x

Bidigare, R. R., and Ondrusek, M. E. (1996). Spatial and temporal variability of phytoplankton pigment distributions in the central equatorial Pacific Ocean. Deep Sea Res. Part II Top. Stud. Oceanogr. 43, 809–833. doi: 10.1016/0967-0645(96)00019-7

Boonyapiwat, S., Sada, M. N., Mandal, J. K., and Sinha, M. K. (2008). Species Composition, Abundance and Distribution of Phytoplankton in the Bay of Bengal. The Ecosystem-Based Fishery Management in the Bay of Bengal. Bangkok: SEAFDEC, 53–64.

Bristow, L. A., Calbeck, C. M., Larsen, M., Altabet, M. A., Dekaezemacker, J., Forth, M., et al. (2017). N 2 production rates limited by nitrite availability in the Bay of Bengal oxygen minimum zone. Nat. Geosci. 10, 24–29. doi: 10.1038/ngeo2847

Campbell, L., Liu, H., Nolla, H. A., and Vaulot, D. (1997). Annual variability of phytoplankton and bacteria in the subtropical North Pacific Ocean at Station ALOHA during the 1991–1994 ENSO event. Deep Sea Res. Part I Oceanogr. Res. Pap. 44, 167–192. doi: 10.1016/S0967-0637(96)00010-2

Campbell, L., Nolla, H., and Vaulot, D. (1994). The importance of Prochlorococcus to community structure in the central North Pacific Ocean. Limnol. Oceanogr. 39, 954–961. doi: 10.4319/lo.1994.39.4.0954

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336. doi: 10.1038/nmeth.f.303

Cellamare, M., Rolland, A., and Jacquet, S. (2010). Flow cytometry sorting of freshwater phytoplankton. J. Appl. Phycol. 22, 87–100. doi: 10.1007/s10811-009-9439-4

Chase, M. W., Solitis, D. E., Olmstead, R. G., Morgan, D., Les, D. H., Misher, B. D., et al. (1993). Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene rbcL. Ann. Mol. Bot. Gard. 80, 528–580. doi: 10.2307/2399846

Chen, F., Wang, K., Kan, J., Bachiou, D. S., Lu, J., Lau, S., et al. (2004). Phylogenetic diversity of Synechococcus in the Chesapeake Bay revealed by Ribulose-1, 5-bisphosphate carboxylase-oxygenase (RuBisCO) large subunit gene (rbcL) sequences. Aquat. Microb. Ecol. 36, 153–164. doi: 10.3354/ame036153
Kumar, S. P., and Unnikrishnan, A. (1995). Seasonal cycle of temperature
Jyothibabu, R., Vinayachandran, P., Madhu, N., Robin, R., Karnan, C., Jagadeesan, J.,
Elkin, M., Jordan, R. W., Cros, L., and Young, J. R. (2004). A revised classification scheme
Joo, S., Lee, S.-R., and Park, S. (2010). Monitoring of phytoplankton community
Jabir, T., Dhanya, V., Jesmi, Y., Prabhakaran, M., Saravanane, N., Gupta, G., et al.
Imhoff, J. (2016). New dimensions in microbial ecology-functional genes in studies
Kulkarni, V. V., Chitari, R. R., Narale, D. D., Patil, J. S., and Anil, A. C.
Hallegraeff, G. M. (1983). Scale-bearing and loricate nanoplankton from the East
1619ñ1640. doi: 10.1016/j.dsr.2007.06.002
Nord. J. Bot. 9, 473ñ481. doi: 10.1016/S0969-2126(01)00608-6
Menezes, L. D., Fernandes, G. L., Mulla, A. B., Meena, R. M., and Damare, M.
Manton, I. (1978). Chrysochromulina hirta sp. nov., a widely distributed species
Mandoji, C., Libini, C. L., Idu, K. A., Mohamed, K., and Kripa, V. (2018).
Maddurpratap, M., Gauns, M., Ramaiah, N., Kumar, S. P., and Muralleedharan, P.
Biogeochemistry of the Bay of Bengal: physical, chemical and primary productivity characteristics of the central and western Bay of Bengal during summer monsoon 2001.
Mar Biol 30, 2075ñ2081. doi: 10.1099/00207713-50-6-2075
John, D. E., Patterson, S. S., and Paul, J. H. (2007). Phytoplankton-group specific
Kabir, T., Dhanya, V., Prabhakaran, M., Saravanane, N., Gupta, G., et al.
Kumar, S. P., and Unnikrishnan, A. (1995). Seasonal cycle of temperature
Kumar, S. P., Nuncio, M., Ramaiah, N., Sardesi, S., Narvekar, J., Fernandes, V., et al. (2007). Eddy-mediated biological productivity in the Bay of Bengal during fall and spring intermonsoons. Deep Sea Res. Part I Oceanogr. Res. Pap. 54, 1619ñ1640. doi: 10.1016/j.dsr.2007.06.002
Kumar, S. P. and Unnikrishnan, A. (1995). Seasonal cycle of temperature and associated wave phenomena in the upper layers of the Bay of Bengal. J. Geophys. Res. Oceans 100, 13585ñ13593. doi: 10.1029/95JC01037
Lavin, P., González, B., Santibañez, J. F., Scanlan, D. J., and Ullhoa, O. (2010). Novel lineages of Prochlorococcus thrive within the oxygen minimum zone of the eastern tropical South Pacific. Env. Microbiol. Rep. 2, 728ñ738. doi: 10.1111/j.1756-1848.2010.00167.x
Li, B., Karl, D. M., Letelier, R. M., Bigdike, R. R., and Church, M. J. (2013). Variability of chromophytic phytoplankton in the North Pacific Subtropical Gyre. Deep Sea Res. Part II Top. Stud. Oceanogr. 93, 84ñ95. doi: 10.1016/j.dsr2.2013.03.007
Li, G., Ke, Z., Lin, Q., Ni, G., Shen, P., and Liu, H. (2012). Longitudinal patterns of spring-intermonsoon phytoplankton biomass, species compositions and size structure in the Bay of Bengal. Acta Oceanol. Sin. 31, 121ñ128. doi: 10.1007/s12524-010-0562-3
Madhupratap, M., Gauns, M., Ramaiah, N., Kumar, S. P., and Muralleedharan, P. (2003). Biogeochemistry of the Bay of Bengal: physical, chemical and primary productivity characteristics of the central and western Bay of Bengal during summer monsoon 2001. Deep Sea Res. Part II Top. Stud. Oceanogr. 50, 881ñ896. doi: 10.1016/S0011-7471(02)00661-2
Mannesha, K., Saras, V., Reddy, N., Sadhuram, Y., Murty, R. T., Saras, V., et al. (2011). Meso-scale atmospheric events promote phytoplankton blooms in the coastal bay of Bengal. J. Earth Syst. Sci. 120, 773ñ782. doi: 10.1007/s12040-011-0089-y
Mishra, K., Libini, C. L., Idu, K. A., Mohamed, K., and Kripa, V. (2018). Occurrence of Diatom-Dioctophytoplankton in the coastal surface waters of south Andaman, India. Symbiosis 76, 293ñ302. doi: 10.1007/s13199-018-0559-y
Mann, D., and Droop, S. (1996). “Biodiversity, biogeography and conservation of Diatoms,” in Biogeography of Freshwater Algae, ed. J. Kristiansen (Dordrecht: Springer), 19ñ32. doi: 10.1007/978-94-017-0908-8_2
Manton, I. (1978). Chrysochromulina hirta sp. nov., a widely distributed species with unusual spines. Br. Phycol. J. 13, 3ñ14. doi: 10.1000/1760-5950(1978)13<3::AID-BPHY400>3.0.CO;2-3
Marshall, H. G., and Lacouture, R. (1986). Seasonal patterns of growth and composition of phytoplankton in the lower Chesapeake Bay and vicinity. Estuar. Coast. Shelf Sci. 23, 115ñ130. doi: 10.1016/0278-4343(99)00135-8
McCready, J. P., Jr., Kundu, P. K., and Molinari, R. L. (1993). A numerical investigation of dynamics, thermodynamics and mixed-layer processes in the Indian Ocean. Prog. Oceanogr. 31, 181ñ244. doi: 10.1016/0079-6613(93)90202-U
Menezes, L. D., Fernandes, G. L., Mullu, A. B., Meena, R. S., and Damare, S. R. (2018). Diversity of culturable Sulphur-oxidising bacteria in the oxygen minimum zones of the northern Indian Ocean. J. Mar. Syst. doi: 10.1016/j.jmarsys.2018.05.007
Mithbabkar, S., Patil, J. S., and Rajaneesh, K. (2015). Picophytoplankton as tracers of environmental forcing in a tropical monsoonal bay. Microb. Ecol. 70, 659ñ676. doi: 10.1007/s00248-015-0599-2
Montagnes, D. J., and Franklin, M. (2001). Effect of temperature on diatom volume, growth rate, and carbon and nitrogen content: reconsidering some paradigms. Limnol. Oceanogr. 46, 2008ñ2018. doi: 10.4319/lo.2002.47.4.1272
Moon- van der Steay, S. Y., Van der staa, G. W., Guillou, L., Vaulot, D., Claustre, H., and Medlin, L. K. (2000). Abundance and diversity of prymnesiophytes in the picoplankton co-munity from the equatorial Pacific Ocean inferred from 18S rDNA sequences. Limnol. Oceanogr. 45, 98ñ109. doi: 10.4319/lo.2000.45.1.0098
Morse, D., Salsol, P., Markovic, P., and Hastings, J. W. (1995). A nuclear-encoded form II RuBisCO in Dinoflagellates. Science 268, 1622ñ1624. doi: 10.1126/science.7777861
Muralleedharan, K., Jasmine, P., Achuthanikutty, C., Revichandran, C., Kumar, P. D., Rejomon, A. G., et al. (2007). Influence of basin-scale and mesoscale physical processes on biological productivity in the Bay of Bengal during the summer monsoon. Prog. Oceanogr. 72, 364ñ383. doi: 10.1016/j.pocean.2006.09.012
Nagaman, P., Shikhakoli, R., and Chauhan, P. (2011). Phytoplankton variability in the Bay of Bengal during winter monsoon using Oceanasat-1 Ocean Colour Monitor data. Indian J. Soc. Remote 39, 117ñ126. doi: 10.1111/j.1758-2229.2010.00167.x
via automated high-frequency sampling. ISME J. 12:2417. doi: 10.1038/si.1396-018-0169-y
Nelson, D. M., Tréguer, P., Brzezinski, M. A., Luyten, A., and Quéguiner, B. (1995). Production and dissolution of biogenic silica in the ocean: revised global estimates, comparison with regional data and relationship to biogenic silification. Glob. Biogeochem. Cycles 9, 359–372. doi: 10.1029/95GB01070
Okano, Y., Hristova, K. R., Leutenegger, C. M., Jackson, L. E., Denison, R. F., Gebreyesus, B., et al. (2004). Application of real-time PCR to study effects of ammonium on population size of ammonia-oxidizing bacteria in soil. Appl. Environ. Microb. 70, 1008–1016. doi: 10.1128/AEM.70.2.1008-1016.2004
Palińska, K. A., and Surosz, W. (2008). Population of Aphanizomenon from the Gulf of Gdańsk (Southern Baltic Sea): differences in phenotypic and genotypic characteristics. Hydrobiologia 607, 163–173. doi: 10.1007/s10789-008-9388-y
Parke, M., Manton, I., and Clarke, B. (1955). Studies on marine flagellates II. three new species of chrysochromulina. J. Mar. Biol. Assoc. UK. 34, 579–609. doi: 10.1017/S0025315400008833
Partensky, F., Hass, W. R., and Vaulot, D. (1999). Prochlorococcus, a marine photosynthetic prokaryote of global significance. Microbiol. Mol. Biol. R. 63, 106–127.
Patrick, R. (1948). Factors effecting the distribution of diatoms. Botan. Rev. 14, 473–524. doi: 10.1007/BF02861575
Paul, J. H., Pichard, S. L., Kang, J. B., Watson, G. M., and Tabita, F. R. (1999). Evidence for a clade-specific temporal and spatial separation in ribulose bisphosphate carboxylase gene expression in phytoplankton populations off cape hatteras and bermuda. Limnol. Oceanogr. 44, 12–23. doi: 10.4319/lo.1999.44.1.0012
Paul, J. T., Ramaiah, N., Gauns, M., and Fernandes, V. (2007). Preponderance of a few diatom species among the highly diverse microphytoplankton assemblages in the Bay of Bengal. Mar. Biol. 152, 63–75. doi: 10.1007/s00227-007-0657-5
Paul, J. T., Ramaiah, N., and Sardessai, S. (2008). Nutrient regimes and their effect on distribution of phytoplankton in the Bay of Bengal. Mar. Environ. Res. 66, 337–344. doi: 10.1016/j.marenvres.2008.05.007
Peeken, I. (1997). Photosynthetic pigment fingerprints as indicators of phytoplankton biomass and development in different water masses of the Southern Ocean during austral spring. Deep Sea Res. Part II Top. Stud. Oceanogr. 44, 261–282. doi: 10.1016/S0967-0645(96)00077-X
Peter, K. H., and Sommer, U. (2013). Phytoplankton cell size reduction in response to warming mediated by nutrient limitation. PLoS One 8:e71528. doi: 10.1371/journal.pone.0071528
Pichard, S. L., Frischer, M. E., and Paul, J. H. (1993). Ribulose-bisphosphate carboxylase gene expression in subtropical marine phytoplankton populations. Mar. Ecol. Prog. Ser. 101:55. doi: 10.3354/meps10105.5
Pienaar, R., and Bandu, V. (1984). A new species of Chrysochromulina (Prymnesiophyceae) from Natal inshore waters. Electr. Microsc. Soc. S. Afr. 14, 65–66.
Pochon, X., Bott, N. J., Smith, K. F., and Wood, S. A. (2013). Evaluating detection limits of next-generation sequencing for the surveillance and monitoring of international marine pests. PLoS One 8:e73935. doi: 10.1371/journal.pone.0073935
Prasanna Kumar, S., Nuncio, M., Narvekar, J., Kumar, A., Sardessai, S., De souza, S., et al. (2004). Are eddies nature's trigger to enhance biological productivity in the Bay of Bengal? Geophys. Res. Lett. 31:L07309.
Prasannakumar, S., Nuncio, M., Narvekar, J., Ramaiah, N., Sardessai, S., Gauns, M., et al. (2010). Seasonal cycle of physical forcing and biological response in the Bay of Bengal. Indian J. Mar. Sci. 39, 388–405.
Rajpathak, S. N., Banerjee, R., Mishra, P. G., Khedkar, A. M., Patil, Y. M., Joshi, S. R., et al. (2018). An exploration of microbial and associated functional diversity in the OMZ and non-OMZ areas in the Bay of Bengal. J. Biosci. 43, 635–648. doi: 10.1007/s12038-018-9781-2
Read, D. S., Bowes, M. J., Newbold, L. K., and Whiteley, A. S. (2014). Weekly flow cytometric analysis of riverine phytoplankton to determine seasonal bloom dynamics. Environ. Science. Process. Impacts 16, 594–603. doi: 10.1039/C3ern00557c
Reddy, P. R. C., Salvekar, P., and Nayak, S. (2008). Super cyclone induces a mesoscale phytoplankton bloom in the Bay of Bengal. IEEE Geosci. Remote Sens. 5, 588–592. doi: 10.1109/LGRS.2008.2006650
Round, F. E., Crawford, R. M., and Mann, D. G. (2007). Diatoms: Biology and Morphology of the Genera. Cambridge: Cambridge university press.
Mitteilungen: Internationale Vereinigung für theoretische und angewandte Limnologie.

Vaulot, D., Eikrem, W., Viprey, M., and Moreau, H. (2008). The diversity of small eukaryotic phytoplankton ($\leq 3 \mu m$) in marine ecosystems. *FEMS Microbiol. Rev.* 32, 795–820. doi: 10.1111/j.1574-6976.2008.00121.x

Veldhuis, M. J., and Kraay, G. W. (1990). Vertical distribution and pigment composition of a picoplanktonic prochlorophyte in the subtropical North Atlantic: a combined study of HPLC-analysis of pigments and flow cytometry. *Mar. Ecol. Prog. Ser.* 68, 121–127. doi: 10.3354/meps068121

Vinayachandran, P., Chauhan, P., Mohan, M., and Nayak, S. (2004). Biological response of the sea around Sri Lanka to summer monsoon. *Geophys. Res. Lett.* 31:L01302. doi: 10.1029/2003GL018533

Vinayachandran, P., and Mathew, S. (2003). Phytoplankton bloom in the Bay of Bengal during the northeast monsoon and its intensification by cyclones. *Geophys. Res. Lett.* 30:1572. doi: 10.1029/2002GL016717

Vinayachandran, P., Shankar, D., Vernekar, S., Sandeep, K., Amol, P., Neema, C., et al. (2013). A summer monsoon pump to keep the Bay of Bengal salty. *Geophys. Res. Lett.* 40, 1777–1782. doi: 10.1002/grl.50274

Watson, G. M., and Tabita, F. R. (1997). Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: a molecule for phylogenetic and enzymological investigation. *FEMS Microbiol. Lett.* 146, 13–22. doi: 10.1111/j.1574-6968.1997.tb10165.x

Wawrik, B., Paul, J., and Tabita, F. (2002). Real-time PCR quantification of rbcL (ribulose-1, 5-bisphosphate carboxylase/oxygenase) mRNA in Diatoms and pelagophytes. *Appl. Environ. Microbiol.* 68, 3771–3779. doi: 10.1128/AEM.68.2.3771-3779.2002

Werner, D. (1977). “Introduction with a note on taxonomy,” in *The Biology of Diatoms. Botanical Monographs*, ed. D. Werner (London: Blackwell), 1–17.