Journal of Species Research 10(1):46-56, 2021

Molecular profiling of 18S rRNA reveals seasonal variation and diversity of diatoms community in the Han River, South Korea

Buhari Lawan Muhammad, Yeon-Su Lee and Jang-Seu Ki[1,*]

Department of Biotechnology, Sangmyung University, Seoul 03016, Republic of Korea

**Correspondent: kijs@smu.ac.kr*

Diatoms have been used in examining water quality and environmental change in freshwater systems. Here, we analyzed molecular profiling of seasonal diatoms in the Han River, Korea, using the hypervariable region of 18S V1-V3 rRNA and pyrosequencing. Physicochemical data, such as temperature, DO, pH, and nutrients showed the typical seasonal pattern in a temperate region. In addition, cell counts and chlorophyll-*a*, were recorded at high levels in spring compared to other seasons, due to the diatom bloom. Metagenomic analysis showed a seasonal variation in the phytoplankton community composition, with diatoms as the most frequently detected in spring (83.8%) and winter (69.7%). Overall, diatom genera such as *Stephanodiscus*, *Navicula*, *Cyclotella*, and *Discostella* were the most frequent in the samples. However, a large number of unknown Thalassiosirales diatoms were found in spring (35.5%) and winter (36.3%). Our molecular profiling revealed a high number of diatom taxa compared to morphological observation. This is the first study of diatoms in the Han River using molecular approaches, providing a valuable reference for future study on diatoms-basis environmental molecular monitoring and ecology.

Keywords: phytoplankton, *Cyclotella*, water quality, freshwater diatoms, metagenomics

DOI:10.12651/JSR.2021.10.1.046

Introduction

Diatoms are groups of eukaryotic unicellular microalgae and are some of the most common organic matter in an aquatic ecosystem. They play an important role as the primary producers in freshwater and marine environments, and they contribute significantly to the photosynthesis activity on Earth (Lee *et al*., 2019). In addition, diatoms are used widely across the world for environmental monitoring because different species of diatom typically have specific habitat preferences, and they react quickly to environmental changes (Dixit *et al*., 1992; Round *et al*., 2007). For example, species like *Cyclotella meneghiniana* and *Thalassiosira pseudonana* (Baek *et al*., 2011; Adenan *et al*., 2013) have distinct ranges of pH and salinity where they will grow (Spaulding *et al*., 2019). In addition, they have ranges and tolerance for other environmental factors, such as nutrient concentration, suspended sediment, flow regime, elevation, and for different types of human disturbances (Spaulding *et al*., 2019). Hence, they are considered susceptible indicators of water quality and environmental change. For these reasons, diatoms have received considerable attention from past and ongoing research (Dixit *et al*., 1992; Adenan *et al*., 2013; Hilaluddin *et al*., 2020).

In general, diatoms have been discriminated morphologically by using a light microscope (LM) and scanning electron microscope (SEM). However, some diatoms may have few discernible structural differences when viewed with a microscope. In addition, different lineages have often adapted in similar ways to analogous habitats. As a result, there are many examples of diatoms species that look very similar at the microscopic level, but are not closely related (Graham *et al*., 2009). Since the early 1980s, molecular techniques by using carbohydrates, toxins, proteins, and nucleic acids as markers have been developed to detect and discriminate phytoplankton (Ebenezer *et al*., 2012; Sun *et al*., 2012). Among DNA based methods such as terminal restriction fragment length polymorphism, microarray, real-time PCR, the high-throughput sequencing (HTS) of small-subunit rRNA genes from environmental DNA techniques has already been widely applied for the assessment of microbial diversity and micro-planktonic community structure (Eiler *et al*., 2012; Ghiglione and Murray, 2012). These methods allowed high resolution and rapid analysis of microbial and phytoplankton com-

munities. In addition, recent next generation sequencing NGS-based approaches facilitate the precise identification of nano, pico, rare, and fragile phytoplankton (Faria *et al*., 2014; Boopathi *et al*., 2015).

The Han River is the largest river system in South Korea, which serves as an important resource for the people living in Seoul and nearby cities. People living in cities depend on the river as the primary source of water. Over the decades, phytoplankton seasonal distribution has been well studied in the Han River (Jung *et al*., 2003). In addition, the diatom *Stephanodiscus hantzschii* has been found to be frequently exposed to increased winter blooms, which is considered an indicator of eutrophication (Lee and Yoon, 1996; Kim *et al*., 2001; Hwang *et al*., 2003; Jung *et al*., 2009). However, the molecular data of diatoms in the Han River is lacking. To date, the results from most of the studies were based on microscopic observations and have been inconsistent (Jung *et al*., 2003; Hwang *et al*., 2011; Yun *et al*., 2014). Therefore, considering the current status of the river and its importance to the public, a comprehensive metagenomics study on phytoplankton dynamics is imperative.

In view of this, we aim to institute a detailed study on the diatom community to explore their molecular diversity and seasonal variations in the temperate freshwater Han River. The results from this study are the first diatom studies in the study site using molecular approaches, and can be used as a valuable reference for future study on diatoms for environmental monitoring.

Materials and Methods

Environmental factors and water sample collections

The surface water samples were collected from the Seongsan Bridge of the Han River in March (spring), June (summer), September (autumn), and December (winter) 2012. Water temperature, pH, dissolved oxygen (DO), and conductivity from a monitoring site were measured using a YSI 566 Multi Probe System (Xylem Inc., Yellow Springs, OH) when environmental sampling was carried out. In addition, water samples from the sampling site were collected at the surface by using a 20 L bucket from March to December 2012. In a brief, three hundred milliliters of water samples were fixed with 1% Lugol's solution (Sigma-Aldrich) and subsequently used for the identification and quantification of phytoplankton using a light microscope (Axioskop, Zeiss, Oberkochen, Germany). Additionally, samples for environmental DNA extraction were prepared as follows: Firstly, large size organisms such as zooplankton were removed using a 200 μm-pore size mesh. A total of 500 mL of this pre-filtered freshwater was size-fractionated through a 10 μm (Cat. No. TCTP04700, 47 mm diameter, Millipore, Billerica, MA), followed by a 2.0 μm (TTTP04700, 47 mm diameter, Millipore) and 0.22 μm membrane filters (GVWP04700, 47 mm diameter, Millipore) to prevent clogging. The membrane filters were immersed into 0.8 mL extraction buffer (100 mM Tris- HCl, 100 mM Na2-EDTA, 100 mM sodium phosphate, 1.5 M NaCl, 1% CTAB) and were stored at −80°C until DNA extraction.

Nutrient data and chlorophyll-*a* measurement

As for nutrient data, total nitrogen (TN) and total phosphorus (TP) were obtained from the Han River Basin Environmental Office (http://www.me.go.kr/hg/web/main.do). The Chlorophyll-*a* (Chl-*a*) concentrations were estimated according to Parsons *et al*. (1984). Briefly, a total of 200 mL of water samples were filtered with GF/F filter (Cat. No. 1825047, 47 mm diameter, Whatman, UK), and those filters were placed in 90% acetone for overnight under the dark in order to extract pigments. The supernatants were used to measure the concentration of Chl-*a* using a DU730 Life Science UV/Vis spectrophotometer (Beckman Coulter, Inc., Fullerton, CA).

Environmental DNA extraction

Environmental DNA from the filtered samples was extracted using a modified protocol by Harder *et al*. (2003). A 2 mL microcentrifuge tube containing each membrane filter (10 μm, 2.0 μm, and 0.22 μm) was subjected to freeze-thaw cycles in liquid N_2 and 65°C maintained water bath. Subsequently, 8 μL proteinase K (10 mg/mL in TE buffer) was added and the tube was incubated at 37°C for 30 min. Following incubation, 80 μL 20% sodium dodecyl sulfate (SDS) prepared in double-distilled water (ddH_2O) was added and the sample was incubated at 65°C for 2h. After incubation, the tube was shaken with equal volumes of chloroform-isoamyl alcohol (24 : 1) and centrifuged at 10,000 × g for 5 min. The aqueous phase of the mixture was transferred to a new microcentrifuge tube, to which 0.1 volumes of 3 M sodium acetate (pH 5.1, prepared in ddH_2O) and 0.6 volumes of isopropanol (≥99%) were added. The microcentrifuge tube was centrifuged at 14,000 × g for 20 min, the supernatant was discarded, 1 mL cold 70% ethanol was added to the pellet, and the sample was centrifuged again at 14,000 × g for 15 min. The pellet was air-dried and reconstituted with 100 μL TE buffer (10 mM Tris-HCl, 1 mM EDTA; pH 8).

PCR and pyrosequencing

Based on the study by Ki (2012), the V1−V3 hypervariable region of the 18S rRNA was selected for metagenomic data analyses. Target rDNA retrieved from the environmental samples was amplified using polymerase

chain reaction (PCR). PCR reactions were performed using two universal primers: 18F23 (5′-ACC TGG TTG ATC CTG CCA GTA G-3′) and 3NDf-R (5′-CTG GCA CCA GAC TTG CC-3′); the latter primer was designed as the complementary sequence to the universal primer 3NDf (Cavalier-Smith *et al*. 2009). Each primer was tagged using multiplex identifier (MID) adaptors according to the manufacturer's instructions (Roche, Mannheim, Germany), which allowed for the automatic sorting of the pyrosequencing derived sequencing reads based on MID adaptors. In addition, MID-linked 18F23 and 3NDf-R were linked to pyrosequencing primers, according to manufacturer's instructions (Roche, Mannheim, Germany).

Metagenomic PCR reactions were performed with 20 μL reaction mixtures containing 2 μL of 10× Ex Taq buffer (TaKaRa, Kyoto, Japan), 2 μL of a dNTP mixture (4 mM), 1 μL of each primer (10 pM), 0.2 μL of the Ex Taq polymerase (2.5 U), and 0.1 μg of the environmental DNA template. PCR cycling was performed in an iCycler (Bio-Rad, Hercules, CA) at 94°C for 5 min, followed by 35 cycles at 94°C for 20 s, 52°C for 40 s, and 72°C for 1 min, and a final extension at 72°C for 10 min. The resulting PCR products were electrophoresed in 1.0% agarose gel, stained with ethidium bromide, and viewed under ultraviolet transillumination.

PCR products were individually purified using a Dual PCR Purification Kit (Bionics, Seoul, Korea) and subsequently, equal volumes of each purified PCR product were mixed together. Pyrosequencing of eight MID-tagged PCR amplicons was performed with a GS FLX Titanium system (Roche, Mannheim, Germany) using a commercial service at Macrogen Inc. (Seoul, Korea).

Data cleaning and BLAST (Basic Local Alignment Search Tool) searches

Pyrosequencing data were processed using NGS analysis pipeline of the SILVA rRNA gene database project (SILVAngs 1.0; Quast *et al*., 2013) to perform systematic checks to remove sequencing artifacts and low-quality reads (i.e. reads <300 bp and reads with $>1\%$ ambiguity or 2% homopolymers). In the initial quality check, low-quality reads and sequencing artifacts were discarded, and the remaining sequence data were manually checked, trimmed and edited using the Sequencher v5.1 (Gene Codes Corporation, Ann Arbor, MI). Subsequently, identical reads were identified (dereplication) and clustered at the operational taxonomic units (OTUs) on 97% similarity thresholds, using the program Cd-hit-est (http://www.bioinformatics.org/cd-hit; Li and Godzik, 2006). In order to reduce data size, only single reads with the longest DNA sequences were selected as a reference read for OTU classification. Finally, we constructed a dataset comprised of different genotypic sequence reads.

For OTU classification, the individual reads were subjected to BLASTn searches (Basic Local Alignment Search Tool) against the GenBank database. They were identified and assigned to their respective taxonomic groups. Sequence similarities of $\geq 98\%$ with known species were considered to represent the identical species. Those ranging from 95.0% to 97.9% similarity were considered to represent the identical genus, and sequences with less than 95% similarity were regarded as unknown, while less than 93% were assigned to the meta-group 'No Relative' (Boopathi and Ki, 2016; Penna *et al*., 2017; Banerji *et al*., 2018). All the reads that were assigned to the respective OTU were mapped onto classified OTU reference read, which provides quantitative information about the classified OTU. The thresholds stated herein were set based on 18S rRNA sequence comparisons with strains from different species and genera. In the present study, we determined a total of 145 sequences, which were deposited in GenBank with accession numbers MT361356−MT361500.

Diversity and statistical analyses

For the diatom diversity indices, Good's coverage was calculated in the SILVAngs pipeline. The Shannon-Weaver diversity index (H), Evenness (E), Diversity estimator Chao-1, Rarefaction curve, and Principal Component Analyses (PCA) were calculated using Paleontological Statistics Software Package (PAST) version 3.25 (Hammer *et al*., 2001).

Results

Environmental variables

Environmental variables (Fig. 1) such as water temperature, dissolved oxygen (DO), pH, conductivity, TN, and TP were measured to record their influence on the diatom blooms. Water temperature varied considerably within these months, ranging from zero to 28°C. The water temperature and DO concentration were found to be inversely related to each other; increasing the water temperature decreased the DO concentration (Fig. 1A). TN and TP levels were found vary over the months (Fig. 1B). In addition, the conductivity of the water was found to be high in March (250 μS cm^{-1}), and low in September (160 μS cm^{-1}). The pH varied over the months ranging from 6.6 in June to 7.3 in December (Fig. 1C). As for biological factors, Chl-*a* concentrations (Fig. 1D) were found to be congruent with pattern changes of both cell numbers except in September when there was little variation.

Pyrosequencing characteristics and diversity indices

Pyrosequencing generated a total of 50,594 sequences reads in March (spring), June (summer), September (aut-

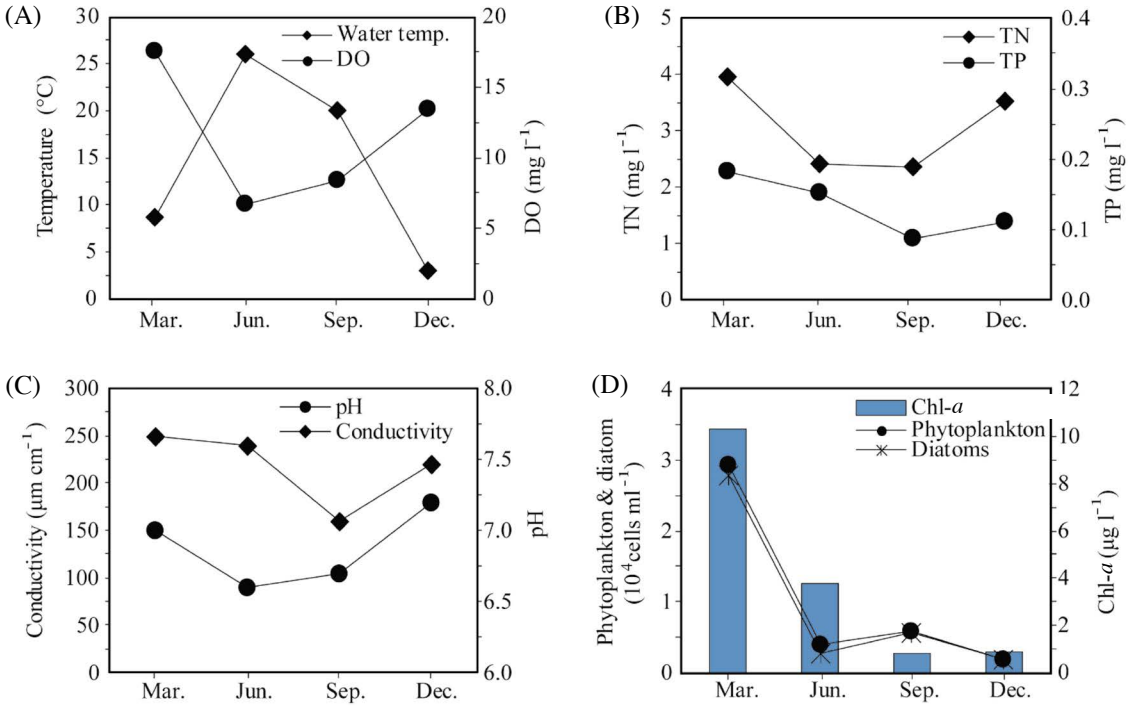

Fig. 1. Seasonal variation in water temperature and DO (A), pH and conductivity (B) and TN and TP (C), and cell counts and Chl-*a* (D) at the Seongsan Bridge of Han River, Korea.

Table 1. Pyrosequencing data (50,594 reads) for micro-eukaryotic community with emphasis on diatoms.

Month	Total reads	Classified reads	Rejected reads	Phytoplankton reads	Diatom reads	No. of Diatom OTU	% of Diatom
Mar.	14468	13761	707	10677	8943	37	64.9
Jun.	12221	11516	705	4964	804	30	6.9
Sep.	17900	17152	748	6332	1670	48	9.7
Dec.	6005	5701	304	2477	1726	34	30.2

umn), and December (winter), of which raw data were further analyzed by the NGS analysis pipeline at the SILVA rRNA gene database project (SILVAngs 1.0). At the initial quality filtering step, 2,464 low-quality reads were rejected, and the remaining 48,130 reads were classified (Table 1). The number of diatom OTU varied among the samples, and most neared the plateau while the March sample reached the plateau (Fig. 2). In March (spring), 64.9% of the total micro-eukaryotic sequence reads belonged to the diatom.

For assessing the sampling completeness, we used Good's coverage estimator and was calculated by randomly selecting sequence reads from a given sample. Good's coverage values ranged between 96.8% and 98.9%. The Shannon-Weaver diversity index (*H'*) was measured to characterize community diversity which ranged from 1.57 (autumn) to 2.65 (winter), indicating a high level of overall biodiversity. In addition, the Chao-1 estimator demonstrated community richness. Chao-1 values ranged between 44 in December and 53 in September samples, indicating these samples were rich in diversity (Table 2).

Diatoms community structures

Our pyrosequencing results uncovered a considerably high eukaryotic diversity which was represented by all known eukaryotic supergroups. Overall, among all the micro-eukaryotic species, diatoms had the highest number of OTU reads in March (spring) (64.9%), while in other months there was a high number of Eukaryota (Fig. 3A). From the overall phytoplankton reads, upon comparisons with diatoms, we found out that diatoms were frequently present in March (83.7%) and December (69.7%), however, the lowest number of diatom reads were recorded in June (16%) where Chlorophyceae (66%) were signifi-

cantly present among the phytoplankton community (Fig. 3B). Among the diatoms, Coscinodiscophycidae were the most frequent in all four seasons (Fig. 3C).

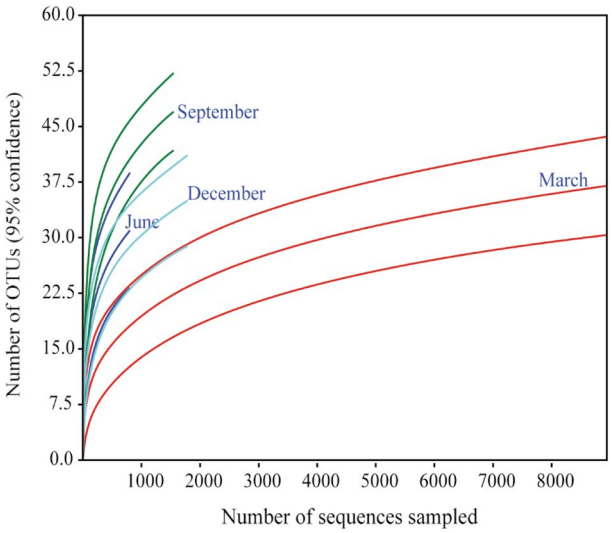

Fig. 2. Rarefaction curves representing the numbers of Operational Taxonomic Units (OTUs) of diatoms vs. the number of tags sampled from pyrosequencing data.

A large number of OTU reads that belong to unknown diatom species (unknown Thalassiosirales 1) were detected in spring (35.5%) and winter (36%) (Table 3). However, reads that belong to *Stephanodiscus* sp. were also detected frequently in spring (31.8%). In summer there was a change in the species compositions where unknown reads of Aulacoseiracea were the most frequently detected (37.8%). Interestingly in autumn a different diatom species was detected (*Navicula* sp., 21.6%) that was not detected in other seasons. The top ten most frequent species within the diatom community exhibited seasonal variations and are listed in Table 3.

Table 2. Diversity indices (Shannon-Weaver diversity index (*H'*), Evennnes (*E*), Chao-1 estimator, and Good's coverage) obtained from phylotypes recovered in the present study.

Diversity indices	Month			
	Mar.	Jun.	Sep	Dec.
Shannon-Weaver (*H'*)	1.57	2.11	2.65	2.20
Evenness (*E*)	0.13	0.27	0.30	0.26
Chao-1	48.3	45	53	44
Good's coverage	97.5	96.8	98.9	96.6

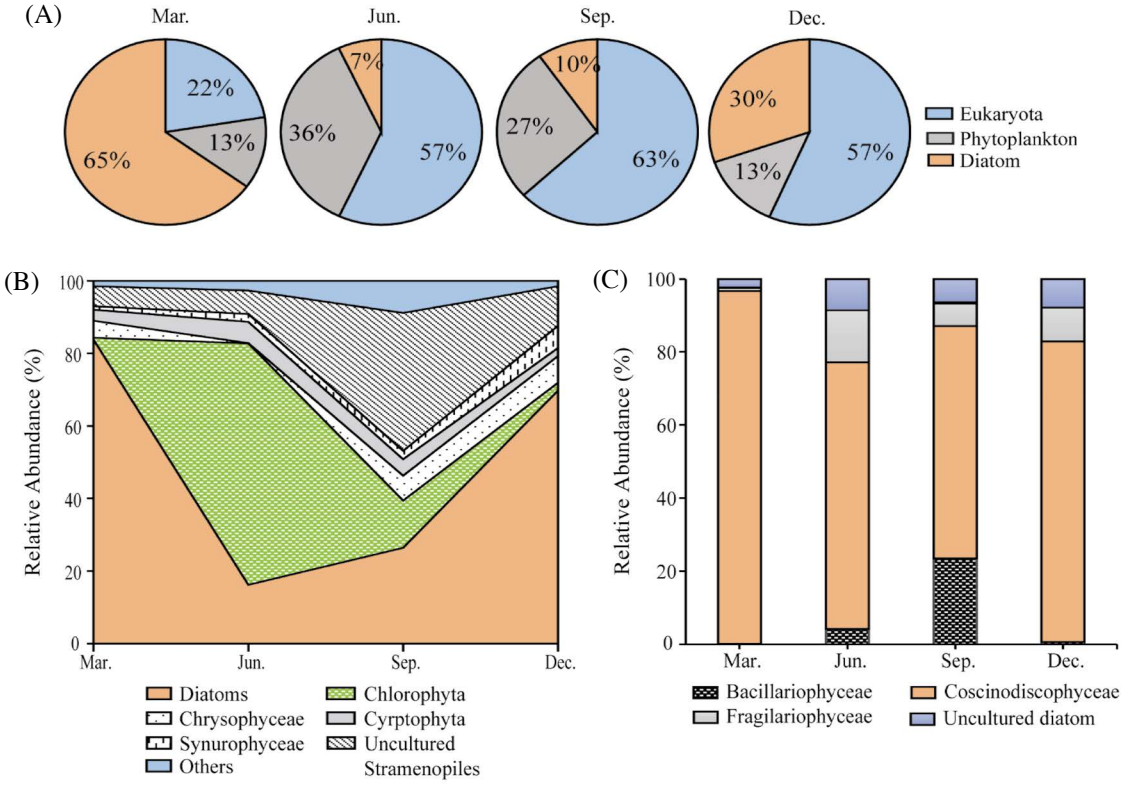

Fig. 3. (A) Proportion of each eukaryotic taxon (eukaryote, phytoplankton, and diatom), (B) relative abundance of phytoplankton, and (C) relative abundance of diatom taxa. These data were calculated by using 18S rRNA pyrosequencing reads. Taxonomic identity of "others" represents taxa with less than 1% composition of total reads.

Table 3. Relative abundance of the frequent taxa detected from the diatom community of the Han River.

Species	Month (%)				GenBank accession number
	Mar.	Jun.	Sep	Dec.	
Unknown Thalassiosirales 1	35.5	0.0	0.0	36	MT361386
Stephanodiscus sp.	31.8	1.2	0.1	0.0	MT361392
Unknown Thalassiosirales 2	21.5	1.6	0.0	0.0	MT361389
Unknown Aulacoseiraceae	0.0	37.8	0.0	0.0	MT361400
Cyclotella choctawhatcheeana	0.0	21.8	6	0.5	MT361399
Unknown Fragilariales	0.0	12.2	0.0	0.0	MT361414
Navicula sp.	0.0	0.0	21.6	0.1	MT361431
Cyclostephanos sp.	0.0	0.0	17	0.0	MT361441
Aulacoseira ambigua	0.0	0.0	15.9	0.1	MT361457
Discostella woltereckii	0.0	0.0	0.0	19.5	MT361479
Unknown Coscinodiscophycea	2.1	0.0	1	11	MT361486

Seasonal variations

We constructed a PCA plot with the diatom relative abundance to determine the variation between the seasonal samples. PC1 and PC2 represented 54.06% and 31.73%, respectively (Fig. 4A). The biplot analyses revealed that most of the OTUs were clustered together; however, some OTUs highly varied especially between March and June. This showed the seasonal variation among some diatom species. In addition, PCA analysis based on the most frequently detected OTU reads revealed the seasonal variations between the diatom species (Fig. 4B).

Microscopic analysis of diatom community

A total of 60 taxa of phytoplankton were identified by morphological analysis in this study, in which 27 taxa were identified as diatoms (Table 4). In all the samples, the diatom community accounted for >70% of the phytoplankton. The highest cell counts of the diatoms were in December (>90%). Among the diatoms, *Stephanodiscus hantzschii* f. *tenuis* was the most abundant taxon in March (49.2%) and December (88.2%) (Tables 4 and 5). In June and September *Aulacoseira* sp. was found to be the most abundant taxon, however, *Fragilaria* sp. was dominant in June (29.9%) (Table 5).

Discussion

The occurrence of diatoms is influenced by many environmental factors, such as temperature, light, pH, and nutrients (Scott and Marcarelli, 2012). However, the conditions necessary to trigger the blooms remain uncertain even after more than 60 years of study (Sathyendranath *et al*., 2015). In the present study, diatoms were the most abundant in spring and winter compared to all other microeukaryotic groups. Similar results were reported by several other phytoplankton diversity studies (Jung *et al*., 2003; Faria *et al*., 2014; Boopathi and Ki, 2016). Typically, the diatom community remains dominant in winter and spring blooms in large lowland river systems around the world (Rumyantseva *et al*., 1993; Wehr and Descy, 1998; Reynolds, 2006). This may be because diatoms tend to have significantly high maximum uptake rates of nutrients than any other group (Malviyaa *et al*., 2016) and have relatively high maximum growth rates. These characteristics make diatoms good nutrient competitors (Wolanski and Elliott, 2016). Due to this, mass occurrence of diatoms in spring affects the surrounding biodiversity and disrupts ecosystem functions worldwide (Wehr and Descy, 1998; Jung *et al*., 2013), therefore, monitoring of the diatom community is imperative to control spring blooms.

Stephanodiscus sp. is an important component of phytoplankton as it plays significant roles in the freshwater phytoplankton. During winter, *Stephanodiscus* species are well-known, bloom-forming diatoms in Korean rivers (Lee and Yoon, 1996; Jung *et al*., 2009), and the bloom density of this species in Korean rivers is much higher compared to other rivers worldwide (Ha *et al*., 1998). In the present study, we found OTU reads of *Stephanodiscus* sp. to be the most frequently detected during spring. Similar cases have been recorded by previous work of this species being abundant during spring and winter (Jung *et al*., 2003; Boopathi and Ki, 2016). This is probably due to the lower temperatures in winter and spring, which is favorable for *Stephanodiscus* spp. growth. Generally, cold-water diatoms dominated phytoplankton spring blooms in temperate regions (Grigorszky *et al*., 2017). In addition, *Stephanodiscus* spp. are normally in high occurrence in eutrophicated water and were reported also in eutrophic freshwater

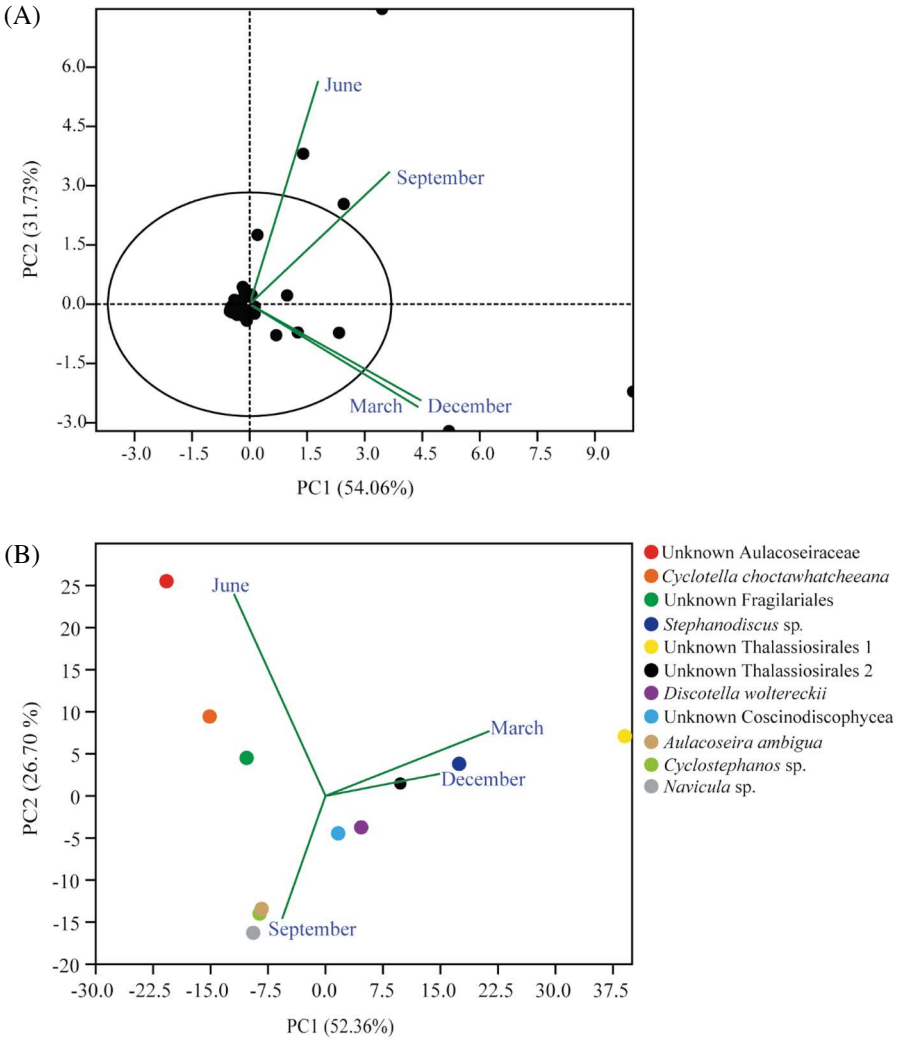

Fig. 4. Principal component analysis (PCA) biplot showing the seasonal variation of (A) all the diatom OTU reads, (B) most frequent diatom OTU detected, sampled in March (spring), June (summer), September (autumn), and December (winter). Calculated based on the number of OTU reads in each sample. Each dot represents diatom OTU recovered in this study.

of North America and Japan (Kiss *et al*., 2013).

Cyclotella choctawhatcheeana was the most frequent taxon detected in summer. The considerable morphological variation among species of *Cyclotella* makes their taxonomy hard to unravel (Håkansson and Kling, 1994; Meyer and Håkansson, 1996). *Cytocella choctawhatcheeana* can grow in different seasons with high and low nutrient availability, which makes them an important contributor to the depth chlorophyll maximum (DCM) present in the stratification period (Mitrovic *et al*., 2010). Despite their small size (5−54 μm), *Cytocella* species are significant components of phytoplankton in the saline water (saline system). However, *C. meneghiniana* growth increased linearly with temperature to a growth maximum at 25°C, then at 28°C the decreased in growth was observed (Mitrovic *et al*., 2010). Therefore, in this study, the high occurrence of *Cyclotella* sp. seems to be related to increased water temperature and a relative decrease in nutrient level in summer and autumn. In addition, this seems to be because *Cyclotella* sp. is an oligotroph (Saros and Anderson, 2015), however, the observed frequent taxa, *Stephanodiscus* sp. is a copiotroph (Jung *et al*., 2011). On the other hand, *Aulacoseira* sp. is one of the most successful, in terms of distribution in time and space, of all freshwater centric diatoms (Spaulding and Edlund, 2008), and in the present study, *Aulacoseira* sp. and *Cyclostephanos* sp. were detected in autumn, but not detected in other seasons. These two species may serve as good bioindicators of water eutrophication level (Wang *et al*., 2017). *Navicula* sp. was found to occur in autumn. Based on the literature review, we could not find any record of this species in high occurrence in the Han River. Therefore, this could potentially have a

Table 4. List and cell counts of diatom taxa present in the Han River from March 2012 to Dec. 2012, identified under microscopy.

Class	Species	Cell counts (cells/mL)			
		Mar.	Jun.	Sep.	Dec.
Bacillariophyceae	*Cymbella* sp. (L)	49	13	0	0
	Cymbella sp. (S)	148	0	51	7
	Gomphonema sp.	0	0	0	0
	Gyrosigma sp.	49	0	0	0
	Navicula sp.	542	91	123	19
	Nitzschia holsatica	0	0	0	0
	Nitzschia sp.	49	0	21	0
	Pinnularia sp.	49	0	0	0
	Cocconeis placentula	99	13	0	0
	Surirella tenera	443	0	0	0
Coscinodiscophyceae	*Actinoptychus* sp.	148	0	93	0
	Aulacoseira sp. (L)	0	0	895	0
	Aulacoseira sp. (M)	549	784	1512	0
	Aulacoseira sp. (S)	0	457	1810	0
	Aulacoseira sp. (XL)	0	0	627	19
	Cyclotella sp.	2808	26	144	5
	Stephanodiscus hantzchii	7637	353	72	77
	Stephanodiscus hantzchii f. *tenuis*	13795	78	0	1560
Fragilariophyceae	*Asterionella formosa*	0	0	0	0
	Asterionella gracillima	99	0	82	24
	Diatoma sp.	99	0	21	0
	Fragilaria sp. (L)	394	823	195	41
	Fragilaria sp. (S)	0	0	0	0
	Synedra sp. (L)	99	0	41	7
	Synedra sp. (S)	739	0	10	2

Table 5. Top 3 diatom phylotypes based on frequency, detected by microscopy and 454 pyrosequencing.

Month	Species detected by microscopy	(%)	Phylotypes detected by pyrosequencing	(%)
March	*Stephanodiscus hantzchii* f. *tenuis*	49.2	Unknown Thalassiosirales 1	35.5
	Stephanodiscus hantzchii	27.2	*Stephanodiscus* sp.	31.8
	Cyclotella sp.	10.0	Unknown Thalassiosirales 2	21.5
June	*Fragilaria* sp. (L)	29.9	Unknown Aulacoseiraceae	37.8
	Aulacoseira sp. (M)	28.6	*Cyclotella choctawhatcheeana*	21.8
	Aulacoseira sp. (S)	16.7	Unknown Fragilariales	12.2
September	*Aulacoseira* sp. (S)	31.7	*Navicula* sp.	21.6
	Aulacoseira sp. (M)	26.5	*Cyclostephanos* sp.	17.0
	Aulacoseira sp. (L)	15.7	*Aulacoseira ambigua*	15.9
December	*Stephanodiscus hantzchii* f. *tenuis*	88.2	Unknown Thalassiosirales 1	36.0
	Fragilaria sp. (L)	2.3	*Discostella woltereckii*	19.5
	Navicula sp.	1.0	Unknown Coscinodiscophycea	11.0

significant ecological impact.

As our pyrosequencing results highlight, a large number of diatom sequences that greatly affected the phytoplankton community in spring and summer belonged to the unidentified diatoms (unknown Thalassiosirales 1). This was due to the lack of comparable sequences from the freshwater environments, since most of the currently available sequences in the databases contain phytoplankton sequences of marine origin (Eiler *et al*., 2013). Expanding the sequence databases with cultured and characterized freshwater phytoplankton is imperative for comparing environmental sequences. Furthermore, improving taxonomic frameworks would address the discrepancies between phytoplankton phylogeny and various morphological classification systems (Gugger *et al*., 2002; Zapomêlová *et al*., 2009).

In the present study, both microscopic observations and molecular approaches revealed species diversity and seasonal variations in diatom community composition (Table 4). However, results from the two methods were incongruent. Previous studies have shown the incongruous diversity pattern associated with diatoms and other phytoplankton identification based on microscopic and molecular methods (Lopes *et al*., 2012; Eiler *et al*., 2013; Xiao *et al*., 2014). This may be due to microscopic ambiguities, taxonomic experience, and identification skill. In addition, different lineages of algae may be similar at the microscopic level, but actually not closely related while some are closely related (Graham *et al*., 2009). For example, in the present microscopic observation, we detected *Stephanodiscus hantzchii* f. *tenuis* to be the predominant species in March and December; however, the OTU reads data from pyrosequencing analysis revealed a large number sequences of an unknown species of Thalassiosirales, which is closely related to *Stephanodiscus* sp., *Thalassiosira*-like sp., and *Cyclostephanos* sp. A similar case was also reported by Boophati and Ki (2016). However, these centric diatoms were found to be closely related in phylogenetic analyses (Kaczmarsk *et al*., 2006; Alverson *et al*., 2007). In consequence, many phycologists use molecular characters as well as structural features and rely on the phylogenetic species concept for taxonomic descriptions of phytoplankton (Graham *et al*., 2009). However, both microscopy and molecular methods can be used to aid in resolving the issues related with phytoplankton diversity. For example, in the present study, our molecular profiling of the 18S rRNA revealed that Unknown Thalassiosirales occurred in high frequency in March and December samples. However, in our microscopic analysis and in previous studies (Lee and Yoon, 1996; Kim *et al*., 2001; Hwang *et al*., 2003; Jung *et al*., 2009), *Stephanodiscus hantzchii* was the dominant taxon in winter and spring. In this case, morphological identification can be used to confirm the result from molecular analysis (Krienitz and Bock, 2012) and therefore the unknown sequence from the molecular methods is likely to be *Stepanodiscus hantzchii*. Thus, both morphological and molecular techniques are crucial in monitoring programs and the study of diatom taxonomy. For an objective assessment of true diversity, future research should optimize and standardize all steps in microscopic and molecular methods, and minimize the discrepancies between the two methods.

The present study provides detailed insight into the diversity and seasonal community structure of freshwater diatoms in the Han River, Korea. We extensively analyzed the molecular taxonomy of these diatoms and compared the results to morphological data. The results from microscopy and molecular data were incongruent. The molecular results revealed that uncultured diatom species belonging to the order Thalassiosirales contributed significantly to the overall diatoms community composition in spring and winter. *Stephanodiscus* sp. were also detected frequently in spring, and as the temperature became warmer in summer, *Cyclotella choctawhatcheeana* occurrence increased. Interestingly, we found a high number of *Navicula* sp. OTU reads in autumn with few or no records in other seasons. The findings of the present study can be used as a valuable reference for comparing diatom diversity in future studies. In addition, we suggest the use of NGS approaches in continuous monitoring of diatoms, which may lead to better management of algal blooms and maintaining good water quality.

Acknowledgements

We thank Mr. MD Lee for field sampling and species identification. We would like to thank two anonymous reviewers for their valuable suggestions and comments. This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2016M3A9E1915578 and 2020R1A2C2013373) and by a research Grant from Sangmyung University.

References

Adenan, N.S., F.M. Yusoff and M. Shariff. 2013. Effect of salinity and temperature on the growth of diatoms and green algae. Journal of Fisheries and Aquatic Science 8(2):397.

Alverson, A.J., R.K. Jansen and E.C. Theriot. 2007. Bridging the Rubicon: phylogenetic analysis reveals repeated colonizations of marine and fresh waters by thalassiosiroid diatoms. Molecular Phylogenetics and Evolution 45(1): 193-210.

Baek, S.H., S.W. Jung and K. Shin. 2011. Effects of temperature and salinity on growth of *Thalassiosira pseudonana* (Bacillariophyceae) isolated from ballast water. Journal of Freshwater Ecology 26(4):547-552.

Banerji, A., M. Bagley, M. Elk, E. Pilgrim, J. Martinson and J. Santo Domingo. 2018. Spatial and temporal dynamics of a freshwater eukaryotic plankton community revealed via 18S rRNA gene metabarcoding. Hydrobiologia 818(1):71-86.

Boopathi, T. and J.S. Ki. 2016. Unresolved diversity and monthly dynamics of eukaryotic phytoplankton in a temperate freshwater reservoir explored by pyrosequencing. Marine and Freshwater Research 67(11):1680-1691.

Boopathi, T., D.G. Faria, M.D. Lee, J. Lee, M. Chang and J.S. Ki. 2015. A molecular survey of freshwater microeukaryotes in an Arctic reservoir (Svalbard, 79 N) in summer by using next-generation sequencing. Polar Biology 38(2): 179-187.

Chou, H.H. and M.H. Holmes. 2001. DNA sequence quality trimming and vector removal. Bioinformatics 17(12):1093-1104.

Dixit, S.S., J.P. Smol, J.C. Kingston and D.F. Charles. 1992. Diatoms: powerful indicators of environmental change. Environmental Science and Technology 26(1):22-33.

Ebenezer, V., L.K. Medlin and J.S. Ki. 2012. Molecular detection, quantification, and diversity evaluation of microalgae. Marine Biotechnology 14(2):129-142.

Eiler, A., F. Heinrich and S. Bertilsson. 2012. Coherent dynamics and association networks among lake bacterioplankton taxa. The ISME Journal 6(2):330-342.

Eiler, A., S. Drakare, S. Bertilsson, J. Pernthaler, S. Peura, C. Rofner and E.S. Lindström. 2013. Unveiling distribution patterns of freshwater phytoplankton by a next generation sequencing based approach. PLoS One 8(1):e53516.

Faria, D.G., M.D. Lee, J.B. Lee, J. Lee, M. Chang, S.H. Youn and J.S. Ki. 2014. Molecular diversity of phytoplankton in the East China Sea around Jeju Island (Korea), unraveled by pyrosequencing. Journal of Oceanography 70(1):11-23.

Ghiglione, J.F. and A.E. Murray. 2012. Pronounced summer to winter differences and higher wintertime richness in coastal Antarctic marine bacterioplankton. Environmental Microbiology 14(3):617-629.

Graham, L.E., J.M. Graham and W.W. Lee. 2009. Algae.- Second edition, Benjamin Cummings, New York.

Grigorszky, I., K. Tihamér Kiss, G. Pór, G. Dévai, A.S. Nagy, I. Somlyai and É. Ács. 2017. Temperature and growth strategies as the essential factors influencing the occurrence of *Stephanodiscus minutulus* (Kützing) Cleve and Möller and *Palatinus apiculatus* (Ehrenberg) Craveiro, Calado, Daugbjerg and Moestrup. Fundamental and Applied Limnology/ Archiv für Hydrobiologie 189(2):167-175.

Gugger, M., C. Lyra, P. Henriksen, A. Coute, J.F. Humbert and K. Sivonen. 2002. Phylogenetic comparison of the cyanobacterial gen era Anabaena and Aphanizomenon. International Journal of Systematic and Evolutionary Microbiology 52(5):1867-1880.

Ha, K., H.W. Kim and G.J. Joo. 1998. The phytoplankton succession in the lower part of hypertrophic Nakdong River (Mulgum), South Korea.- In Pytoplankton and Trophic Gradients Springer, Dordrecht (pp. 217-227).

Håkansson, H. and H. Kling. 1994. *Cyclotella agassizensis* nov. sp. and its relationship to *C. quillensis* Bailey and other prairie *Cyclotella* species. Diatom Research 9(2):289-301.

Hammer, Ø., D.A. Harper and P.D. Ryan. 2001. PAST: Paleontological statistics software package for education and data analysis. Palaeontologia Electronica 4(1):9.

Harder, T., S.C. Lau, S. Dobretsov, T.K. Fang and P.Y. Qian. 2003. A distinctive epibiotic bacterial community on the soft coral *Dendronephthya* sp. and antibacterial activity of coral tissue extracts suggest a chemical mechanism against bacterial epibiosis. FEMS Microbiology Ecology 43(3): 337-347.

Hilaluddin, F., F.M. Yusoff and T. Toda. 2020. Shifts in Diatom Dominance Associated with Seasonal Changes in an Estuarine-Mangrove Phytoplankton Community. Journal of Marine Science and Engineering 8(7):528.

Hwang, S.J., N.Y. Kim, S.A. Yoon, B.H. Kim, M.H. Park, K.A. You, and O.M. Lee. 2011. Distribution of benthic diatoms in Korean rivers and streams in relation to environmental variables. In Annales de Limnologie-International Journal of Limnology (Vol. 47, No. S1, pp. S15-S33). EDP Sciences.

Hwang, S.J., S.K. Kwun and C.G. Yoon. 2003. Water quality and limnology of Korean reservoirs. Paddy and Water Environment 1(1):43-52.

Jung, S.W., J.H. Lee and J.S. Yu. 2003. Environmental studies of the lower part of the Han River V. Blooming characteristics of phytoplankton communities. Algae 18(4):255-262.

Jung, S.W., O.Y. Kwon, J.H. Lee and M.S. Han. 2009. Effects of water temperature and silicate on the winter blooming diatom *Stephanodiscus hantzschii* (Bacillariophyceae) growing in eutrophic conditions in the lower Han River, South Korea. Journal of Freshwater Ecology 24(2):219-226.

Jung, S.W., H. Min Joo, Y.O. Kim, J.H. Lee and M.S. Han. 2011. Effects of temperature and nutrient depletion and reintroduction on growth of *Stephanodiscus hantzschii* (Bacillariophyceae): implications for the blooming mechanism. Journal of Freshwater Ecology 26(1):115-121.

Jung, S.W., Y.H. Kang, S.H. Baek, D. Lim and M.S. Han. 2013. Biological control of *Stephanodiscus hantzschii* (Bacillariophyceae) blooms in a field mesocosm by the immobilized algicidal bacterium Pseudomonas fluorescens HYK0210-SK09. Journal of Applied Phycology 25(1):41-50.

Ki, J.S. 2012. Hypervariable regions (V1-V9) of the dinoflagellate 18S rRNA using a large dataset for marker considerations. Journal of Applied Phycology 24(5):1035-1043.

Kim, B., J.H. Park, G. Hwang, M.S. Jun and K. Choi. 2001. Eutrophication of reservoirs in South Korea. Limnology 2(3):223-229.

Kiss, K.T., S.I. Genkal, L. Ector, L. Molnár, M. Duleba, P. Bíró and É. Ács. 2013. Morphology, taxonomy and distribution of *Stephanodiscus triporus* (Bacillariophyceae) and

related taxa. European Journal of Phycology 48(4):363-379.

Krienitz, L. and C. Bock. 2012. Present state of the systematics of planktonic coccoid green algae of inland waters. Hydrobiologia 698(1):295-326.

Lee, K. and S.K. Yoon. 1996. A study on the phytoplankton in the Paldang Dam Reservoir. III. The changes of diatom community structure. Algae - Korean Phycological Society 11:277-283.

Lee, S.D., S.M. Yun, P.Y. Cho, H.W. Yang and O.J. Kim. 2019. Newly Recorded Species of Diatoms in the Source of Han and Nakdong Rivers, South Korea. Phytotaxa 403(3):143-170.

Li, W. and A. Godzik 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22(13):1658-1659.

Lopes, V.R., V. Ramos, A. Martins, M. Sousa, M. Welker, A. Antunes and V.M. Vasconcelos. 2012. Phylogenetic, chemical and morphological diversity of cyanobacteria from Portuguese temperate estuaries. Marine Environmental Research 73:7-16.

Malviya, S., E. Scalco, S. Audic, F. Vincent, A. Veluchamy, J. Poulain and A. Zingone. 2016. Insights into global diatom distribution and diversity in the world's ocean. Proceedings of the National Academy of Sciences 113(11):E1516-E1525.

Meyer, B. and H. Håkansson. 1996. Morphological variation of *Cyclotella polymorpha* sp. nov. (Bacillariophyceae). Phycologia 35(1):64-69.

Mitrovic, S.M., J.N. Hitchcock, A.W. Davie and D.A. Ryan. 2010. Growth responses of *Cyclotella meneghiniana* (Bacillariophyceae) to various temperatures. Journal of Plankton Research 32(8):1217-1221.

Penna, A., S. Casabianca, A.F. Guerra, C. Vernesi and M. Scardi. 2017. Analysis of phytoplankton assemblage structure in the Mediterranean Sea based on high-throughput sequencing of partial 18S rRNA sequences. Marine Genomics 36:49-55.

Parsons, T.R., Y. Maita and C.M. Lalli. 1984. Amanual of chemical and biological methods for seawater analysis. Pergamon Press.

Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza and F.O. Glöckner. 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research 41(D1):D590-D596.

Reynolds, C.S. 2006. The ecology of phytoplankton. - Cambridge University Press.

Round, F.E., R.M. Crawford and D.G. Mann. 2007. Diatoms: biology and morphology of the genera. Cambridge university press.

Rumyantseva, A., S. Henson, A. Martin, A.F. Thompson, G.M. Damerell, J. Kaiser and K.J. Heywood. 2019. Phytoplankton spring bloom initiation: The impact of atmospheric forcing and light in the temperate North Atlantic Ocean. Progress in Oceanography 178:102202.

Saros, J.E. and N.J. Anderson. 2015. The ecology of the planktonic diatom Cyclotella and its implications for global environmental change studies. Biological Reviews 90(2): 522-541.

Sathyendranath, S., R. Ji and H.I. Browman. 2015. Revisiting Sverdrup's critical depth hypothesis. ICES Journal of Marine Science 72(6):1892-1896.

Spaulding, S. and M. Edlund. 2008. Aulacoseira. In Diatoms of North America. https://diatoms.org/genera/aulacoseira.

Spaulding, S.A., I.W. Bishop, M.B. Edlund, S. Lee, P. Furey, E. Jovanovska and M. Potapova. 2019. What are diatoms? In diatoms of North America. https://diatoms.org/what-are-diatoms.

Sun, F., H.Y. Pei, W.R. Hu and M.M. Song. 2012. A multi-technique approach for the quantification of *Microcystis aeruginosa* FACHB-905 biomass during high algae-laden periods. Environmental Technology 33(15):1773-1779.

Wang, C., S. Lek, Z. Lai and L. Tudesque. 2017. Morphology of Aulacoseira filaments as indicator of the aquatic environment in a large subtropical river: The Pearl River, China. Ecological Indicators 81:325-332.

Wehr, J.D. and J.P. Descy. 1998. Use of phytoplankton in large river management. Journal of Phycology 34(5):741-749.

Wolanski, E. and M. Elliott. 2016. Estuarine ecological structure and functioning. Second Edition. (Elsevier) pp. 157-193.

Xiao, X., H. Sogge, K. Lagesen, A. Tooming-Klunderud, K.S. Jakobsen and T. Rohrlack. 2014. Use of high throughput sequencing and light microscopy show contrasting results in a study of phytoplankton occurrence in a freshwater environment. PLoS One 9:e106510.

Yun, S.M., H.M. Joo, S.W. Jung, C.H. Choi, J.S. Ki and J.H. Lee. 2014. The relationship between epilithic diatom communities and changes in water quality along the lower Han River, South Korea. Journal of Freshwater Ecology 29(3): 363-375.

Zapomělová, E., J.P. Jezberová, D. Hrouzek, Hisem, K. Řeháková and J. Komárková. 2009. Polyphasic characterization of three strains of *Anabaena reniformis* and *Aphanizomenon aphanizomenoides* (cyanobacteria) and their reclassification to *Sphaerospermum* gen. nov. (incl. Anabaena kisseleviana) 1. Journal of Phycology 45(6):1363-1373.

Submitted: September 16, 2020
Revised: November 24, 2020
Accepted: November 27, 2020