Research Article

Algae 2018, 33(2): 167-180 https://doi.org/10.4490/algae.2018.33.4.3

Open Access



Next-generation sequencing reveals the diversity of benthic diatoms in tidal flats

Sung Min An¹, Dong Han Choi^{1,2}, Howon Lee¹, Jung Ho Lee³ and Jae Hoon Noh^{1,2,*}

¹Marine Ecosystem and Biological Research Center, Korea Institute of Ocean Science & Technology, Busan 49111, Korea

Benthic diatoms are ubiquitous in tidal flats and play major roles in maintaining coastal ecosystems. Spatio-temporal variations in diatom diversity have not been well-studied, mainly because of difficulties in morphological identification and the lack of appropriate genetic tools. To overcome these problems, we used the gene encoding the ribulose bisphosphate carboxylase large-subunit (rbcL) as a molecular marker, and sequenced these genes with the aid of the MiSeq platform. In this manner, we explored the genetic diversity of benthic diatoms in tidal flats of Guenso Bay on the west coast of Korea; differences in the spatial distributions of benthic diatoms were evident. The diatom communities were dominated by Nitzschia, Navicula, and Amphora; their relative distributions were affected by the sand proportion, grain size, and air exposure time. Our results suggest that meta-barcoding of the rbcL gene and next-generation sequencing can be used to explore the diversity of benthic diatoms.

Key Words: benthic diatom; diversity; MiSeq; rbcL gene; tidal flat

INTRODUCTION

The intertidal zone is one of the most productive ecosystems in the world (Hughes and Sherr 1983, Watermann et al. 1999). Benthic microalgae are ubiquitous in tidal flats and play major roles in coastal ecosystems. Microalgae contribute significantly to primary production, biomass generation, nutrient exchange via seawater and sediment interfaces, and sediment stabilization (Admiraal 1984, Yallop et al. 1994, Underwood and Barnett 2006). Many taxa are found in tidal flats, including diatoms, cyanobacteria, and euglenophytes; and sometimes dinoflagellates and chlorophytes (Meadows and Anderson 1968, Rejil 2013). Of these, diatoms are the most abundant, being the principal contributors to primary production and algal biomass in the intertidal zone (Admiraal 1976, Sullivan and Moncreiff 1988). Their ubiquitous distribution renders diatoms one of the most important groups of organisms of ocean ecosystems (Yallop et al. 1994, Mann et al. 2010).

Despite the importance of these taxa, knowledge on the diversity of benthic diatoms is limited. As samples from tidal flats contain a variety of organic matter and sedimentary particles, acid treatment is required prior to traditional morphological analysis (Baldi et al. 2011). However, this removes all organic material; we cannot tell whether frustules are those of living diatoms or dead diatoms deposited from overlying water. Furthermore,

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Com-

mercial License (http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received October 30, 2017, Accepted April 3, 2018

*Corresponding Author

E-mail: jhnoh@kiost.ac.kr Tel: +82-51-664-3260, Fax: +82-51-955-3981

²Department of Marine Biology, University of Science and Technology, Daejeon 34113, Korea

³Department of Biology Education, Daegu University, Gyeongsan 38453, Korea

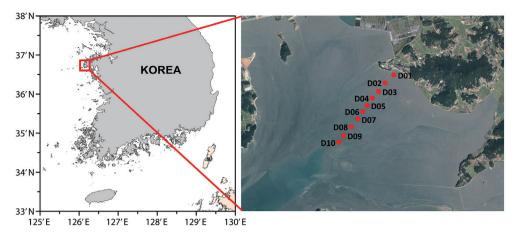


Fig. 1. A map showing the sampling stations on the eastern shore of the Yellow Sea.

diatoms are traditionally identified and classified on the basis of morphological characteristics such as the cell shape, size, and the fine structure of the frustule. Thus, a great deal of time and effort are required for morphological identification and classification. Even experienced taxonomists may differ in their conclusions depending on their skill level and the keys that they follow. For these reasons, morphological data on benthic diatoms of tidal flats are lacking (Zimmermann et al. 2015).

Molecular-based approaches have emerged as useful alternatives to morphology-based approaches and are now widely used to study microorganisms. Molecular approaches based on polymerase chain reaction (PCR) such as denaturing gradient gel electrophoresis, analysis of terminal restriction length polymorphisms, and nextgeneration sequencing (NGS) render it possible to analyze microbial community structures in a time- and costeffective manner (Visco et al. 2015). Of these techniques, NGS is the most powerful. Millions of DNA sequences are obtained, facilitating large-scale biodiversity analyses (Shokralla et al. 2012). In addition, a combination of NGS with DNA barcoding is synergistic when performing high-throughput biodiversity analyses (Hajibabaei et al. 2012). Recently, high-throughput sequencing using the 454 platform of 18S rRNA genes has been used to study diatom diversity (Kermarrec et al. 2013, 2014, Stief et al. 2013, Nanjappa et al. 2014, Visco et al. 2015, Zimmermann et al. 2015). However, very few studies have used NGS to study benthic diatoms (Visco et al. 2015). Further, the phylogenetic resolution afforded by benthic diatom 18S rRNA sequences is inadequate to identify diatoms to the species level (An et al. 2017).

Here, we developed a NGS technique to investigate

the diversity of benthic diatoms. To improve taxonomic resolution, we designed primers specifically targeting the diatom *rbcL* gene and used these for MiSeq sequencing of samples from tidal flats of the Yellow Sea on the western coast of Korea. We studied diatom spatial distribution and identified factors regulating the distributions of various diatom lineages.

MATERIALS AND METHODS

Study site

Sediment samples were collected from 10 tidal flats of the Guenso Bay of Taean on the west coast of Korea (Fig. 1). Guenso Bay is a semi-enclosed bay with an area of 87 km², and the seawater flow is fan-shaped, similar to the shape of the bay. As no river flows into the bay, any influence of fresh water is minimal. The water depth is 2-4 m and the mean tidal difference 6 m. At low tide, about 90% of the total area is exposed; the mean exposure duration is 9.7 h per day. The water depth at high tide ranges from 1.4 to 4.7 m (Kim and Kim 2008).

Sample collection

Surface sediments were collected in October 2011 and January 2012 using a transparent acrylic corer (2 cm in internal diameter and 15 cm in length). As benthic diatoms tend to form heterogeneous mats on the surfaces of sediments, five replicate samples were collected at each station, immediately placed in an ice box, and transported in the cold to the laboratory. Five core samples from

each station were extruded and the top 1-cm amounts of sediment pooled in one conical tube and thoroughly mixed. Samples for DNA extraction were obtained by subsampling the mixed sample and were stored in a deep freezer (-80° C).

Primer design

Diatom-specific primers targeting the *rbcL* gene were designed using the Primer3 (Untergasser et al. 2012) add-on of Geneious Pro v.5.3.6 software (http://www. geneious.com). The primers were designed by reference to an rbcL database containing 1,240 diatom sequences contained in GenBank and our in-house collection (An et al. 2017). After alignment of amino acid sequences, a degenerate primer pair was designed with consideration of amplicon size, specificity, and degeneracy. An appropriate annealing temperature was determined via thermalgradient PCR using cultured strains and natural sediment samples as templates. To assess specificity, the PCR products of a natural sample were purified and cloned into the T-Blunt vector using the T-Blunt PCR Cloning Kit (SolGent, Daejeon, Korea). Cloned products were sequenced by Macrogen (Seoul, Korea) using an M13 primer and an automatic DNA sequencer (ABI Prism 377; Applied Biosystems, Foster City, CA, USA).

DNA extraction and PCR amplification for NGS

Genomic DNAs from 20 sediment samples were extracted using PowerSoil DNA isolation kits (MO BIO Laboratories Inc., Carlsbad, CA, USA) following the manufacturer's protocol. The first PCR was performed using the novel diatom-specific primer pair, which amplified 431-bp rbcL fragments. The PCR conditions were: initial denaturation at 94°C for 5 min; followed by 25 cycles of amplification (95°C for 30 s, 56°C for 30 s, and 72°C for 45 s), and a final extension at 72°C for 10 min. Purification of the first PCR fragments and performance of the second PCR were as suggested in the MiSeq manual (Illumina 2013) and Choi et al. (2016). However, Ex-Tag polymerase (Takara, Ohtsu, Japan) was used for PCR reactions and the products were purified using Agencourt AMPure XP beads (Beckman Coulter, Brea, CA, USA). Each final purified DNA sample was quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). Identical amounts of each PCR product were pooled and the mixed samples sequenced using the Illumina MiSeq 2 × 300 PE platform of Chun Lab (Seoul, Korea).

MiSeq data analysis

We constructed a reference database composed of 2,240 *rbcL* gene sequences from Bacillariophyta Engler & Gilg, 1919 and two *rbcL* gene sequences from Ochrophyta Cavalier-Smith, 1995 using sequences from GenBank and those of strains isolated in a previous study (An et al. 2017). DNA alignment of *rbcL* genes was based on alignment of amino acids translated from the DNA, using the ARB package (Ludwig et al. 2004). Incorrect sequences and those with nonsense codons were excluded.

The sequences were processed using the Mothur package v.1.33.3 (http://www.mothur.org/) (Schloss et al. 2009) following the standard MiSeq operating procedure. Pairs of fastq files were first assembled into contigs using the 'contig.seqs' command and screened to remove poor-quality sequences employing the 'screen. segs (minlength = 400, maxlength = 525, maxambig = 0, maxhomop = 8)' routine. All sequences were translated into amino acid sequences to allow removal of nonsense codons using the Jemboss program v.1.5 (http://emboss. sourceforge.net/Jemboss/) (Carver and Bleasby 2003). Such sequences were removed by the 'remove.segs' command of the Mothur package. Next, single representative sequences selected by the 'unique.seqs' command were aligned to the diatom reference database with the aid of the 'align.seqs' procedure. PCR artifacts and chimeras were removed using the 'pre.cluster (diffs = 4)' and 'chimera.uchime (dereplicate = t)' commands, respectively. The remaining sequences were classified with the aid of the 'classify.seqs (method = knn, numwanted = 1)' command employing the reference database. The number of reads of each sample was confirmed by 'count.group' command. To normalize the number of reads in each sample to sample with the smallest number of reads, all reads were randomly subsampled by the 'sub.sample (size = 9,236)' command of the Mothur package. To build the operational taxonomic unit (OTU) matrix, a distance matrix was generated using the 'dist.seqs (cutoff = 0.20)' command. Next, the sequences were clustered into OTUs based on the distance matrix employing the 'cluster (with cutoffs ranging from 0.01 to 0.06)' command using the opticlust algorithm. Each OTU shared file was generated at each cutoff value using the 'make.shared (labels = from 0.01 to 0.06)' command and used to estimate the number of OTUs and species at each identity level. Finally, representative OTU sequences were generated using the 'get.oturep (method = abundance)' command. The alpha diversity was calculated employing the 'summary. single (calc = nseqs-coverage-sobs-ace-chao-shannonsimpson, subsample = 9,236)' command. The sequences obtained in this study were deposited to the National Center for Biotechnology Information of the US National Library of Medicine (NCBI) sequence read archive (SRA, https://trace.ncbi.nlm.nih.gov/Traces/sra/; accession number SRP126754).

Phylogenetic analysis

OTU taxonomic assignments were checked by performing phylogenetic analyses. Maximum-likelihood (ML) trees of sequences representing major OTUs were constructed using Randomized Axelerated Maximum Likelihood v.8.2.1 software (RAxML) (Stamatakis 2014). Two Ochrophyta species [Triparma pacifica (Guillou & Chrétiennot-Dinet) Ichinomiya & Lopes dos Santos and Triparma mediterranea (Guillou & Chrétiennot-Dinet) Ichinomiya & Lopes dos Santos | served as the outgroup. We used the "-f a" option for rapid bootstrap analysis and the best likelihood tree search employing "-# 100" and the default settings (i.e., "-m GTRGAMMA") of the substitution model with rate heterogeneities "-i" for the automatically optimized subtree pruning and regrafting rearrangement (heuristic search) and "-c" for each of the 25 distinct rate categories. A heat map containing the ML tree was generated and the relative abundance of each OTU displayed using the Interactive Tree Of Life (http:// itol.embl.de) (Letunic and Bork 2007).

Other analyses

Sediment temperature was measured using a thermometer with a stainless steel probe. The exposure time of each station on the tidal flat was calculated using the digital elevation model and tidal data, as described by Choi et al. (2011). Samples for measurement of chlorophyll a (Chl-a) concentration, water content, and grain size were stored in pre-weighed 15-mL tubes and immediately frozen. Sediment samples were dried using a freeze-dryer (FDU-1200; Eyela Co., Tokyo, Japan) and water contents calculated by comparing weights before and after freeze-drying. Chl-a was extracted into 95% (v/v) acetone for 24 h, the samples filtered through 0.2-µmpore-sized hydrophobic syringe filters, and Chl-a levels were measured using the method of Parsons et al. (1984) employing a Turner 10-AU fluorometer (Turner Designs, San Jose, CA, USA) and converted to mass per unit area (mg m⁻²) by reference to the cross-sectional area of the corer. Sediment samples were pretreated prior to particle size analysis. To remove organic matter and carbonate (CaCO₃), 30% (v/v) hydrogen peroxide was added to sediment samples followed by 24 h of incubation on a hot plate (about 80°C); 10% (v/v) hydrochloric acid was then added and the mixture incubated for an additional 24 h on the hot plate. Each sample was washed three times over three days with deionized distilled water. The grain sizes of the sediment samples were then analyzed using a laser diffraction method (Mastersizer 3000 grain size analyzer; Malvern Instruments, Ltd., Worcestershire, UK). Each sample was analyzed in triplicate, and the mean of the three measurements calculated. Ultrasound was applied during analysis to prevent clumping. The size distributions and statistical parameters were calculated using GRADISTAT v.8.0 software (http://www.kpal. co.uk/gradistat.html) (Blott and Pye 2001). Canonical correspondence analysis (CCA) comparing the proportions of major diatom genera and six environmental factors was conducted with the aid of CANOCO (CANOnical Community Ordination v.4.5) software (Ter Braak and Smilauer 2002). In this analysis, the relative proportions of diatom genera were square-root transformed and all variables excluding grain size were log(x + 1)-transformed.

RESULTS AND DISCUSSION

Primer design and specificity

Various DNA regions within nuclear, mitochondrial, and chloroplast genes have been used for the phylogenetic analysis of diatoms (Moniz and Kaczmarska 2010). Previous studies focused principally on ribosomal RNA (i.e., 18S rDNA and 28S rDNA). In particular, 18S rDNA has been extensively used and a database containing a relatively large number of sequences has been constructed (Beszteri et al. 2001, Jones et al. 2005, Sato et al. 2008). However, 18S rDNA affords only low-resolution discrimination of diatom species (Theriot et al. 2009, Mann et al. 2010, Zimmermann et al. 2011, 2015, Kermarrec et al. 2013, An et al. 2017). Although long 18S rRNA gene fragments can be used to identify diatoms to the species level, resolution falls off rapidly with shorter sequences and, moreover, the long-branch attraction artefact of certain diatom species compromises the utility of 18S rDNA sequences in terms of studying benthic diatom diversity (Bruder and Medlin 2008, An et al. 2017). Furthermore, long read sizes are not appropriate when employing NGS. Therefore, alternative genes such as those encoding the ribulose-1,5-bisphosphate carboxylase large subunit (*rbcL*), the cytochrome c oxidase subunit I (*cox*I), the internal transcribed spacer, photosystem II protein D1 (psbA), and ribulose-1,5-bisphosphate carboxylase small subunit (rbcS) have been proposed as alternative genetic markers affording better resolution than 18S rRNA (Ehara et al. 2000, Amato et al. 2007, Evans et al. 2007, Roer 2008, Moniz and Kaczmarska 2010, Trobajo et al. 2010, Delaney et al. 2011, Zimmermann et al. 2011, Kermarrec et al. 2013, An et al. 2017). For NGS analysis, the most important thing is to build a good database. However, the number of sequences for diatoms genes seems to be insufficient. The number of sequences registered in Gen-Bank is 6,624 for 18S rDNA, 4,072 for rbcL, 3,305 for 28S rDNA, and 539 for coxI. Considering the resolution, the number of available sequences, ease of alignment and low risk of contamination by fungi or bacteria (Evans et al. 2007, MacGillivary and Kaczmarska 2011), we selected the *rbcL* gene as a molecular marker for MiSeq analysis.

We designed a primer pair by reference to the *rbcL* sequence database; we sought specificity for benthic diatoms and a fragment size of ca. 450 bp. The two degenerate

primers were as follows: forward: DrbcL265_F, 5'-TAYC-GYGTAGATCCAGTTCCA-3'; and reverse DrbcL695_R, 5'-GCACGRTTRATASCTTCCAT-3'. *In silico* analysis showed that the primers well-matched the sequences of most pennate diatoms, including the most common benthic diatoms (Table 1). When one mismatch was allowed, the primers targeted most diatom sequences of the three predominant benthic orders (Naviculales, Bacillariales, and Thalassiophysales) (Round 1971, Underwood 1994, Cibic et al. 2007) suggesting that the primer pair could be used to study the diversity of benthic diatoms. However, the primers poorly matched the sequences of centric diatoms which are mainly planktonic (Table 1), indicating that the primers should be applied cautiously in studies on planktonic diatom diversity.

PCR-cloning / sequencing supported the idea that the primers were specific for benthic diatoms. All 17 clones from a tidal flat sample were assigned to the diatom *rbcL* sequence by blastn (GenBank), with sequence identities of 92-99% compared with the closest diatom strains (Supplementary Table S1).

Table 1. Taxonomic coverage of the rbcL primer pair designed for this study

| | No. of | No mismatch (%) | | One mismatch (%) | | Two mismatches (%) | |
|--|-----------|-----------------|---------|------------------|---------|--------------------|---------|
| | sequences | Forward | Reverse | Forward | Reverse | Forward | Reverse |
| Class Bacillariophyceae Haeckel, 1878 | | | | | | | |
| Order Naviculales Bessey, 1907 | 660 | 90 | 92 | 100 | 99 | 100 | 100 |
| Order Bacillariales Hendey, 1937 | 366 | 93 | 91 | 99 | 99 | 100 | 100 |
| Order Surirellales D. G. Mann, 1990 | 214 | 96 | 97 | 100 | 100 | 100 | 100 |
| Order Cymbellales D. G. Mann, 1990 | 177 | 82 | 96 | 100 | 99 | 100 | 100 |
| Order Cocconeidales E. J. Cox, 2015 | 72 | 88 | 89 | 99 | 100 | 100 | 100 |
| Order Thalassiophysales D. G. Mann, 1990 | 58 | 74 | 95 | 97 | 100 | 98 | 100 |
| Order Eunotiales P. C. Silva, 1962 | 48 | 0 | 2 | 94 | 90 | 100 | 100 |
| Order Rhopalodiales D. G. Mann, 1990 | 22 | 41 | 82 | 91 | 100 | 100 | 100 |
| Order Achnanthales P. C. Silva, 1962 | 10 | 20 | 50 | 80 | 70 | 100 | 100 |
| Order Fragilariales P. C. Silva, 1962 | 46 | 33 | 100 | 89 | 100 | 98 | 100 |
| Order Licmophorales Round, 1990 | 40 | 63 | 88 | 90 | 98 | 100 | 100 |
| Order Plagiogrammales E. J. Cox, 2015 | 25 | 80 | 96 | 100 | 100 | 100 | 100 |
| Order Rhabdonematales Round & R. M. Crawford, 1990 | 11 | 36 | 55 | 64 | 82 | 73 | 100 |
| Order Rhaphoneidales Round, 1990 | 10 | 40 | 100 | 40 | 100 | 100 | 100 |
| Other Bacillariophyceae | 35 | 49 | 57 | 69 | 86 | 83 | 100 |
| Class Coscinodiscophyceae Round & R. M. Crawford, 1990 | | | | | | | |
| Order Rhizosoleniales P. C. Silva, 1962 | 44 | 34 | 93 | 61 | 95 | 100 | 95 |
| Order Coscinodiscales Round & R. M. Crawford, 1990 | 36 | 17 | 72 | 83 | 97 | 100 | 100 |
| Order Melosirales R. M. Crawford, 1990 | 21 | 57 | 71 | 86 | 95 | 100 | 100 |
| Other Coscinodiscophyceae | 28 | 29 | 86 | 54 | 96 | 64 | 100 |
| Class Mediophyceae (Jousé & Proshkina-Lavrenko) Medlin & Kaczmarska, 2004 | | | | | | | |
| Order Thalassiosirales Glezer & Makarova, 1986 | 155 | 1 | 15 | 13 | 15 | 96 | 98 |
| Order Eupodiscales V. A. Nikolaev & D. M. Harwood, 200 | 0 45 | 2 | 82 | 2 | 82 | 20 | 100 |
| Order Cymatosirales Round & R. M. Crawford, 1990 | 34 | 68 | 91 | 88 | 100 | 100 | 100 |
| Order Biddulphiales Krieger, 1954 | 30 | 30 | 80 | 53 | 90 | 93 | 97 |
| Order Chaetocerotales Round & Crawford, 1990 | 17 | 24 | 82 | 53 | 88 | 82 | 100 |
| Other Mediophyceae | 29 | 21 | 72 | 45 | 83 | 90 | 100 |

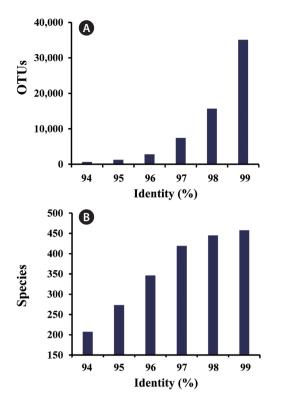


Fig. 2. The number of operational taxonomic units (OTUs) (A) and species (B) estimated over a range of identity thresholds used to assign OTUs of *rbcL* gene sequences.

OTU assignment and classification of reads

The criterion used to separate OTUs is important when it is sought to understand diatom composition and diversity. In general, a 98.7% similarity level of the 16S rRNA gene serves as the threshold defining OTUs at the prokaryote species level (Yarza et al. 2014). However, no consensus criteria in terms of diatom clustering at the species or genus level are available. Nanjappa et al. (2014) analyzed diatom diversity and distribution by clustering OTUs at the 97% identity level employing pyrosequencing of the V4 and V9 regions of small subunit rDNA. Zimmermann et al. (2015) compared pyrosequencing data from 18S V4 diatom regions by clustering OTUs at 98% identity in terms of the traditional morphological classification. With rbcL, Hamsher et al. (2011) reported that inter-species sequence divergences in the genus Sellaphora Mereschowsky, 1902 were very low, ranging from 0.14 to 0.73%. Using a mock community of known freshwater diatoms, Kermarrec et al. (2014) suggested that a 99 and 98% rbcL gene sequence similarity could be used as a clustering threshold at the species and genus level, respectively.

However, recent work showed that the inter-species

distances of *rbcL* genes within the predominant genera of tidal flats ranged from 0.021 to 0.078 (An et al. 2017), suggesting that *rbcL* variation is relatively large among benthic diatom species. In the present study, we estimated the numbers of OTUs and species at different identity proportions (94-99%) of *rbcL* sequence (Fig. 2). The number of OTUs increased exponentially as the threshold increased. However, the number of species tended to become saturated at identity levels over 98%, indicating that 98% was a conservative threshold when used to discriminate species. Therefore, we used 98% *rbcL* sequence similarity for diversity analyses at the species level.

The Mothur program features two methods for classifying reads and assigning taxonomy. One is the k-Nearest Neighbor (k-NN) algorithm used in the present study and the other is the Wang approach. The latter method is implemented by the Ribosomal Database Project classifier (Wang et al. 2007), which analyzes communities in a very reliable manner using a statistical approach termed the Bayesian method. However, when using the Wang approach, more than half of all reads could not be classified even to the genus level, probably because the genetic diversity of benthic diatoms is very poorly understood; the taxonomic information in the database is inadequate (data not shown). Therefore, all reads were classified by reference to the closest database relationship evident using the k-NN algorithm. Next, the classification of the major OTUs was confirmed by phylogenetic analysis to overcome any possible inaccuracies (Jiang et al. 2012). We found some mismatches between the tree-based approach and the k-NN algorithm output; we corrected the classifications of the major OTUs as revealed by the k-NN algorithm by reference to the phylogenetic analyses.

Richness and diversity of benthic diatoms as revealed by the MiSeq approach

To explore the utility of the MiSeq platform when studying the diversity of benthic diatoms, we sequenced 20 samples using the newly designed primers and obtained 972,664 raw reads of an average length of 431 bp (Supplementary Table S2). After removal of low quality / erroneous reads and chimeras, 384,646 reads with 431 bp in length remained. Because the number of reads obtained from each sample ranged from 9,246 to 38,793, the reads were subsampled to normalize the number of sequences in each sample (9,246 reads). In total, 15,655 OTUs were generated after clustering at the 98% similarity level and 103 OTUs exhibited relative frequencies of 1% or more. Diversity indices and number of OTUs for

each sample shown in Table 2. The number of OTUs differed greatly (four-fold) between October and January. Shannon indices in this study ranged from 4.6 to 6.1. There are no previous studies on benthic diatoms, especially using molecular techniques, in tidal flats adjacent to this study area. However, the observed diversity was much higher than benthic algal diversities (1.0-3.3), which were estimated for benthic algae on the basis of microscopic observation, found in other tidal flats (Choi 2002, Yoo 2004, Park et al. 2013). This suggests that traditional morphology-based analysis techniques could significantly underestimate the benthic algal diversity. The richness indices (ACE and Chao1) were also much higher (10- to 20-fold) in January than in October, suggesting large seasonal variations in benthic diatom compositions. The spatial variations in the richness and diversity indices were relatively small compared to the temporal variation. Despite the pronounced seasonal variation, the high-level richness of benthic diatoms suggests that a large proportion of benthic diatom diversity has not yet been explored.

Spatio-temporal variation in benthic diatom levels

The OTUs obtained in the present study clustered into 3 classes, 30 orders, 57 families, 140 genera, and 445 taxa excluding organisms considered *incertae sedis* (e.g., genus *Schizostauron* Grunow, 1867). The OTUs associated

with *Nitzschia*, *Navicula*, *Amphora* Ehrenberg ex Kützing, 1844, and *Planothidium* Round & L. Bukhtiyarova, 1996 predominated in tidal flats, but the particular OTU in differed among the samples, suggesting that the distribution of benthic diatoms was highly heterogeneous in terms of both time and space (Fig. 3).

Interestingly, the 10 most abundant diatom genera constituted over 60% of all diatom communities and exhibited distinct spatio-temporal variations (Fig. 4). Several genera including Nitzschia, Navicula, Amphora, Berkeleya Greville, 1827, Planothidium, Sellaphora, and Surirella Turpin, 1828 were predominant in October and January. However, Nitzschia and Navicula overwhelmingly dominated tidal flats in January. Spatial variation in terms of generic distribution was evident. In October 2011, Navicula and Amphora generally predominated at most stations, but Nitzschia was the most common organism in the low-intertidal areas. In addition, Planothidium and Sellaphora levels tended to decrease toward the sea. In January 2012, the relative Berkeleya proportion tended to fall in the low-intertidal area. Thus, several environmental factors regulated spatial distribution (see below).

The major benthic diatoms of tidal flats of Guenso Bay were *Navicula, Nitzschia,* and *Amphora.* Similarly, *Nitzschia* and *Navicula* were present at relatively high levels during winter in the Geum Estuary of Korea (Kim and Cho 1985) and *Amphora* was a major diatom group in Marennes-Oléeron Bay (Haubois et al. 2005). Benthic

Table 2. Operational taxonomic units (OTUs), abundance-based coverage estimators (ACEs), Chao1 richness values, and Shannon and inverse Simpson diversity indices of sediment samples collected from the Guenso tidal flat

| Date | Station | OTUs | Coverage | ACE | Chao1 | Shannon | Simpson |
|----------|---------|-------|----------|--------|--------|---------|---------|
| Oct 2011 | D01 | 582 | 0.974 | 1,163 | 1,019 | 4.678 | 0.026 |
| | D02 | 546 | 0.978 | 955 | 948 | 4.717 | 0.028 |
| | D03 | 550 | 0.981 | 744 | 823 | 4.774 | 0.026 |
| | D04 | 674 | 0.967 | 1,579 | 1,365 | 4.874 | 0.020 |
| | D05 | 623 | 0.975 | 1,083 | 1,042 | 4.903 | 0.021 |
| | D06 | 594 | 0.975 | 1,170 | 1,265 | 4.998 | 0.017 |
| | D07 | 638 | 0.975 | 1,117 | 969 | 5.097 | 0.014 |
| | D08 | 509 | 0.979 | 1,060 | 1,017 | 4.872 | 0.017 |
| | D09 | 521 | 0.978 | 991 | 984 | 4.747 | 0.022 |
| | D10 | 392 | 0.985 | 827 | 769 | 4.591 | 0.026 |
| Jan 2012 | D01 | 1,857 | 0.849 | 17,841 | 7,769 | 5.600 | 0.015 |
| | D02 | 2,720 | 0.770 | 26,548 | 11,317 | 6.038 | 0.015 |
| | D03 | 2,428 | 0.794 | 22,960 | 10,220 | 5.711 | 0.018 |
| | D04 | 2,602 | 0.776 | 29,298 | 12,119 | 5.970 | 0.013 |
| | D05 | 2,138 | 0.819 | 23,639 | 9,141 | 5.580 | 0.018 |
| | D06 | 2,206 | 0.817 | 21,646 | 10,076 | 5.769 | 0.014 |
| | D07 | 2,237 | 0.813 | 24,813 | 11,139 | 5.841 | 0.013 |
| | D08 | 2,307 | 0.806 | 24,585 | 9,821 | 5.843 | 0.012 |
| | D09 | 2,523 | 0.784 | 27,501 | 11,276 | 6.143 | 0.008 |
| | D10 | 962 | 0.945 | 2,742 | 1,953 | 5.185 | 0.013 |

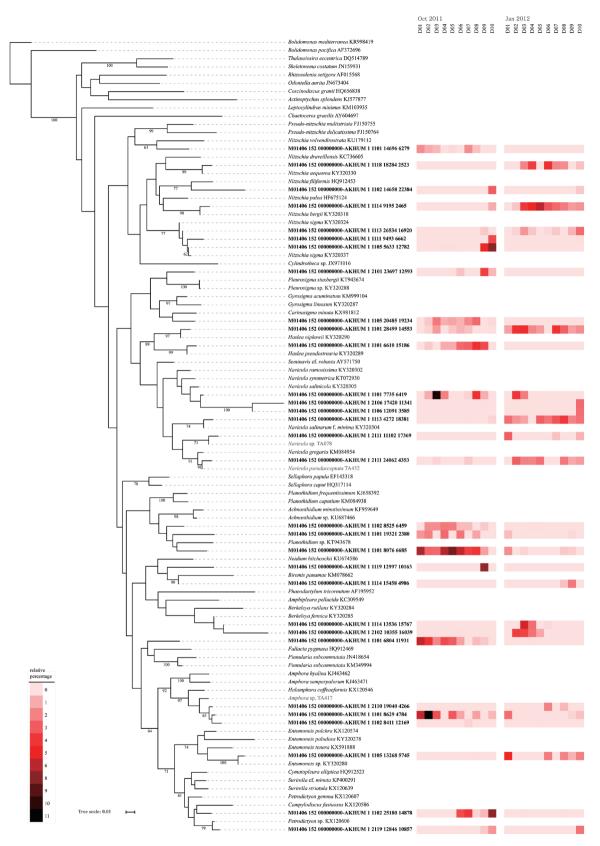


Fig. 3. The maximum likelihood phylogenetic tree of diatoms showing the relationships among genotypes found at 10 stations. The heatmap shows the relative proportions of the major genotypes (at least one sample >3%) in each sample. The color range indicates the proportional abundance of each genotype.

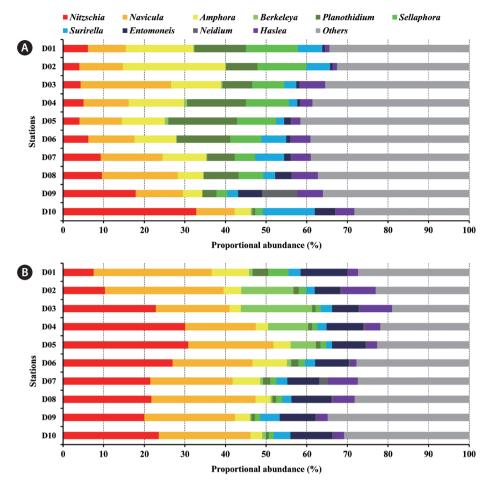


Fig. 4. Temporal variations in the proportional abundance of 10 major genera in the October 2011 (A) and January 2012 (B) samples.

diatoms adhere to a variety of substrates, including mud, sand, stones, plants, algae, and animals (Zimmermann et al. 2015). Diatoms inhabiting tidal flats are mostly epipelic or episammic. *Amphora, Navicula*, and *Nitzschia* are epipelic diatoms with raphes on both sides of the frustule, allowing them to live freely in mud (Round 1971). As the diatoms can move with the aid of an extracellular polymer secreted by the raphe, they grow in coastal environments wherein sediments are continuously deposited and extreme environmental changes occur periodically (Kwon et al. 2012). Therefore, *Amphora, Navicula*, and *Nitzschia* are major taxa of tidal flats worldwide (Round 1971, Dijkema and Wolff 1983, Underwood 1994, Hamels et al. 1998, Cibic et al. 2007, Lee 2013).

The relative frequency might not be consistent with the proportion of absolute abundances of taxa considering the putative PCR efficiency. However, the primer pair used in this study shows similar good coverage for dominant taxa in the tidal flats, indicating that the efficiency problem would not be significant in major benthic diatom groups.

Environmental factors affecting the distributions of benthic diatom genera

To assess environmental factors associated with the spatio-temporal distributions of diatom genera, we subjected benthic diatom assemblages and environmental factors (Table 3) to CCA analysis (Fig. 5). As a result of CCA analysis, we found that grain size, sand percentage, and exposure time all significantly influenced diatom distribution. The variation of benthic diatom assemblages could be accounted for 90.9% by the first two axes (70.8% for first axis and 20.1% for second axis), and the factors that have the largest influence on the community structure were the percentage composition of sand, followed by sediment grain size. De Sève et al. (2010) noted that the relative frequencies of epipelic and episammic

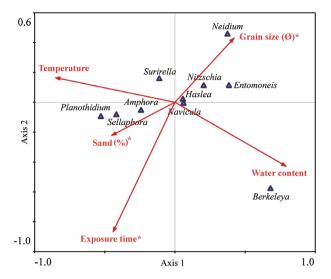


Fig. 5. Canonical Correspondence Analysis ordination diagrams of the 10 major genera with respect to environmental variables. Statistically significant variables revealed by permutation testing are marked with asterisks (p < 0.01).

species were closely associated with sediment grain size. The relative frequencies of the major genera correlated significantly with sediments of particular sizes. Oh et al. (2009) reported that *Nitzschia* grew best when the diameter of the attachment substrate was 90-150 µm (the smallest particle used in the experiments); the smaller the particle size, the higher the growth rate and biomass. In the present study, the proportions of *Nitzschia* and *En-*

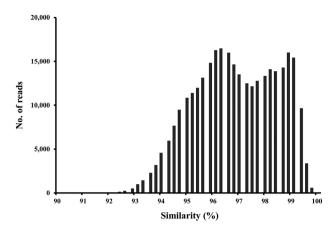


Fig. 6. Histogram showing the similarity distributions of sequence reads and their closest relatives in the reference database.

tomoneis correlated principally (and positively) with the level of muddy sediment. On the other hand, *Amphora* levels correlated positively with the proportion of sand, consistent with the finding of a previous study to the effect that *Amphora* predominated when the particle diameter was about 250 µm (Mitbavkar and Anil 2002).

Applicability and limitations of NGS when studying diatoms

Application of molecular techniques to identify benthic diatoms has obvious advantages, but also several

Table 3. Chlorophyll a (Chl a) levels and environmental parameters

| Sampling time | Station | Chl <i>a</i> (mg m ⁻²) | Sedimentgrain size (Ø) | Sand ratio (%) | Sediment temperature (°C) | Water content (%) | Exposure duration (min) |
|---------------|---------|------------------------------------|---------------------------|----------------|------------------------------|-------------------|----------------------------|
| Oct 2011 | D01 | 26.5 | 5.763 | 14.3 | 19.7 | 33.2 | 420 |
| | D02 | 31.7 | 5.297 | 28.4 | 20.3 | 31.6 | 420 |
| | D03 | 40.5 | 5.275 | 25.8 | 21.2 | 35.3 | 390 |
| | D04 | 37.2 | 5.067 | 35.6 | 21.6 | 31.9 | 360 |
| | D05 | 42.0 | 4.698 | 38.1 | 21.0 | 32.4 | 330 |
| | D06 | 50.5 | 4.875 | 33.7 | 20.7 | 33.6 | 330 |
| | D07 | 56.1 | 5.066 | 38.1 | 21.3 | 31.8 | 310 |
| | D08 | 46.1 | 5.536 | 24.4 | 21.0 | 32.3 | 330 |
| | D09 | 52.7 | 5.839 | 18.2 | 20.4 | 38.8 | 240 |
| | D10 | 77.0 | 5.911 | 17.7 | 19.9 | 31.3 | 210 |
| Jan 2012 | D01 | 74.1 | 5.003 | 28.3 | 2.8 | 30.4 | 390 |
| | D02 | 109.9 | 5.308 | 24.8 | 3.8 | 42.3 | 390 |
| | D03 | 119.4 | 5.396 | 21.4 | 4.3 | 48.9 | 360 |
| | D04 | 111.1 | 5.593 | 19.5 | 4.3 | 48.1 | 330 |
| | D05 | 108.7 | 5.678 | 20.3 | 5.1 | 45.6 | 300 |
| | D06 | 81.2 | 4.838 | 35.2 | 5.2 | 37.2 | 300 |
| | D07 | 101.5 | 5.351 | 23.6 | 5.2 | 34.5 | 290 |
| | D08 | 94.4 | 5.401 | 28.0 | 5.1 | 40.9 | 280 |
| | D09 | 58.5 | 5.714 | 21.4 | 3.5 | 41.1 | 225 |
| | D10 | 76.4 | 5.781 | 20.2 | 2.6 | 31.5 | 190 |

The exposure times represent daytime minutes only.

limitations. The extent of taxonomic coverage by a reference database is important in terms of sequence classification (Kermarrec et al. 2013). Although the number of *rbcL* sequences in the database is relatively greater than those of other protein-encoding genes (except the ribosomal proteins), many of our OTUs lacked close database relatives (Fig. 6). Considering that the estimated species richness was ca. 30,000 in the present study, and the fact that diatoms may feature up to 100,000 species (Mann and Vanormelingen 2013), the database of 2,200 sequences used in the present study is grossly inadequate in terms of thorough phylogenetic analyses. Therefore, continued efforts must be made to expand and improve the database (Visco et al. 2015, Zimmermann et al. 2015).

CONCLUSION

Elucidation of benthic diatom diversity remains challenging. In this study, we developed an ampliconsequencing technique for the *rbcL* gene, using a novel primer set specifically targeting benthic diatoms and the MiSeq high-throughput sequencing platform. This allowed us to estimate diversity indices and explore the spatio-temporal variability of species and environmental factors moderating spatial distribution. The high-level richness of benthic diatoms suggests that a large proportion of such diatoms remains unknown and efforts to reveal the hidden diversity are required. In addition, further studies would increase the quantity and quality of database information, improving the accuracy of future sequence classifications.

ACKNOWLEDGEMENTS

We would like to thank Hwa Young Lee for help with sampling. We also thank Drs. Bon Joo Koo and Jin-Soon Park for help in the analyses of exposure time and grain size, respectively. This study was supported by research programs (PE99512 and PE99611) of the Korea Institute of Ocean Science and Technology.

SUPPLEMENTARY MATERIAL

Supplementary Table S1. Blastn results of clone sequences obtained from a clone library constructed using the primer pair designed in this study (http://e-algae.org).

Supplementary Table S2. Changes in the number of reads according to the pre-processing of MiSeq data (http://e-algae.org).

REFERENCES

- Admiraal, W. 1976. Influence of light and temperature on the growth rate of estuarine benthic diatoms in culture. Mar. Biol. 39:1-9.
- Admiraal, W. 1984. The ecology of estuarine sediment-in-habiting diatoms. Prog. Phycol. Res. 3:269-322.
- Amato, A., Kooistra, W. H. C. F., Ghiron, J. H. L., Mann, D. G., Pröschold, T. & Montresor, M. 2007. Reproductive isolation among sympatric cryptic species in marine diatoms. Protist 158:193-207.
- An, S. M., Choi, D. H., Lee, J. H., Lee, H. & Noh, J. H. 2017. Identification of benthic diatoms isolated from the eastern tidal flats of the Yellow Sea: comparison between morphological and molecular approaches. PLoS ONE 12:e0179422.
- Baldi, F., Facca, C., Marchetto, D., Nguyen, T. N. M. & Spurio, R. 2011. Diatom quantification and their distribution with salinity brines in coastal sediments of Terra Nova Bay (Antarctica). Mar. Environ. Res. 71:304-311.
- Beszteri, B., Ács, É., Makk, J., Kovács, G., Márialigeti, K. & Kiss, K. T. 2001. Phylogeny of six naviculoid diatoms based on 18S rDNA sequences. Int. J. Syst. Evol. Microbiol. 51:1581-1586.
- Blott, S. J. & Pye, K. 2001. GRADISTAT: a grain size distribution and statistics package for the analysis of unconsolidated sediments. Earth Surf. Proc. Land. 26:1237-1248.
- Bruder, K. & Medlin, L. K. 2008. Morphological and molecular investigations of naviculoid diatoms. II. Selected genera and families. Diatom Res. 23:283-329.
- Carver, T. & Bleasby, A. 2003. The design of Jemboss: a graphical user interface to EMBOSS. Bioinformatics 19:1837-1843.
- Choi, D. H., Noh, J. H., An, S. M., Choi, Y. R., Lee, H., Ra, K., Kim, D., Rho, T., Lee, S. H., Kim, K. -T., Chang, K. -I. & Lee, J. H. 2016. Spatial distribution of cold-adapted *Synechococcus* during spring in seas adjacent to Korea. Algae 31:231-241.
- Choi, H. C. 2002. A study on the microphytobenthos on an intertidal mud flat of Ganghwa is. in Korea. M.S. thesis, Inha University, Incheon, Korea, 67 pp.
- Choi, J. -K., Ryu, J. -H., Woo, H. J. & Eom, J. 2011. A study on the flushing characteristics in Geunso bay using hydrohypsographic analysis. J. Wetlands Res. 13:45-52 (in Korean with English abstract).

- Cibic, T., Blasutto, O., Hancke, K., & Johnsen, G. 2007. Microphytobenthic species composition, pigment concentration, and primary production in sublittoral sediments of the Trondheimsfjord (Norway). J. Phycol. 43:1126-1137.
- Delaney, J. A., Ulrich, R. M. & Paul, J. H. 2011. Detection of the toxic marine diatom *Pseudo-nitzschia multiseries* using the RuBisCO small subunit (*rbcS*) gene in two real-time RNA amplification formats. Harmful Algae 11:54-64.
- De Sève, M. A., Poulin, P., Pelletier, É. & Lemarchand, K. 2010.

 Benthic diatom communities from two salt marshes of the St. Lawrence estuary (Canada). J. Water Sci. 23:349-358
- Dijkema, K. S. & Wolff, W. J. 1983. Flora and vegetation of the Wadden Sea islands and coastal areas. August Aimé Balkema, Rotterdam, 413 pp.
- Ehara, M., Inagaki, Y., Watanabe, K. I. & Ohama, T. 2000. Phylogenetic analysis of diatom *cox1* genes and implications of a fluctuating GC content on mitochondrial genetic code evolution. Curr. Genet. 37:29-33.
- Evans, K. M., Wortley, A. H. & Mann, D. G. 2007. An assessment of potential diatom "Barcode" genes (*cox1*, *rbcL*, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). Protist 158:349-364.
- Hajibabaei, M., Spall, J. L., Shokralla, S. & Van Konynenburg, S. 2012. Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol, BMC Ecol. 12:28.
- Hamels, I., Sabbe, K., Muyleart, K., Barranguet, C., Lucas, C., Herman, P. & Vyverman, W. 1998. Organisation of microbenthic communities in intertidal estuarine flats, a case study from the Molenplaat (Westerschelde estuary, The Netherlands). Eur. J. Protistol. 34:308-320.
- Hamsher, S. E., Evans, K. M., Mann, D. G., Poulíčková, A. & Saunders, G. W. 2011. Barcoding diatoms: exploring alternatives to COI-5P. Protist 162:405-422.
- Haubois, A. -G., Sylvestre, F., Guarini, J. -M., Richard, P. & Blanchard, G. F. 2005. Spatio-temporal structure of the epipelic diatom assemblage from an intertidal mudflat in Marennes-Oléeron Bay, France. Estuar. Coast. Shelf Sci. 64:385-394.
- Hughes, E. H. & Sherr, E. B. 1983. Subtidal food webs in a Georgia estuary: δ^{13} C analysis. J. Exp. Mar. Biol. Ecol. 67:227-242.
- Illumina. 2013. 16S metagenomic sequencing library preparation. Available from: http://support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparation.html. Accessed Sep 20, 2017.
- Jiang, S., Pang, G., Wu, M. & Kuang, L. 2012. An improved K-

- nearest-neighbor algorithm for text categorization. Expert Syst. Appl. 39:1503-1509.
- Jones, H. M., Simpson, G. E., Stickle, A. J. & Mann, D. G. 2005. Life history and systematics of *Petroneis* (Bacillariophyta), with special reference to British waters. Eur. J. Phycol. 40:61-87.
- Kermarrec, L., Franc, A., Rimet, F., Chaumeil, P., Frigerio, J. -M., Humbert, J. -F. & Bouchez, A. 2014. A next-generation sequencing approach to river biomonitoring using benthic diatoms. Freshw. Sci. 33:349-363.
- Kermarrec, L., Franc, A., Rimet, F., Chaumeil, P., Humbert, J. F. & Bouchez, A. 2013. Next-generation sequencing to inventory taxonomic diversity in eukaryotic communities: a test for freshwater diatoms. Mol. Ecol. Resour. 13:607-619.
- Kim, D. & Kim, K. H. 2008. Tidal and seasonal variations of nutrients in Keunso Bay, the Yellow Sea. Ocean Polar Res. 30:1-10 (in Korean with English abstract).
- Kim, J.-H. & Cho, K.-J. 1985. The physico-chemical properties of sediment, the species composition and biomass of benthic diatoms in the intertidal zone of Kum River estuary. Korean J. Ecol. 8:21-29 (in Korean with English abstract).
- Kwon, H. K., Yang, H. S., Yu, Y. M. & Oh, S. J. 2012. Effects of substrate size on the growth of 4 microphytobenthos species (*Achnanthes* sp., *Amphora* sp., *Navicula* sp. and *Nitzschia* sp.). J. Environ. Sci. 21:105-111 (in Korean with English abstract).
- Lee, H. Y. 2013. Diversity and biomass of benthic diatoms in Hampyeong Bay tidal flats. Korean J. Environ. Biol. 31:295-301 (in Korean with English abstract).
- Letunic, I. & Bork, P. 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics 23:127-128.
- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T., Steppi, S., Jobb, G., Förster, W., Brettske, I., Gerber, S., Ginhart, A. W., Gross, O., Grumann, S., Hermann, S., Jost, R., König, A., Liss, T., Lüssmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A. & Schleifer, K. H. 2004. ARB: a software environment for sequence data. Nucleic Acids Res. 32:1363-1371.
- MacGillivary, M. L. & Kaczmarska, I. 2011. Survey of the efficacy of a short fragment of the *rbcL* gene as a supplemental DNA barcode for diatoms. J. Eukaryot. Microbiol. 58:529-536.
- Mann, D. G., Sato, S., Trobajo, R., Vanormelingen, P. & Souffreau, C. 2010. DNA barcoding for species identification and discovery in diatoms. Cryptogam. Algol. 31:557-577.

- Mann, D. G. & Vanormelingen, P. 2013. An inordinate fondness? The number, distributions, and origins of diatom species. J. Eukaryot. Microbiol. 60:414-420.
- Meadows, P. S. & Anderson, J. G. 1968. Micro-organisms attached to marine sand grains. J. Mar. Biol. Assoc. U. K. 48:161-175.
- Mitbavkar, S. & Anil, A. 2002. Diatoms of the microphytobenthic community: population structure in a tropical intertidal sand flat. Mar. Biol. 140:41-57.
- Moniz, M. B. J. & Kaczmarska, I. 2010. Barcoding of diatoms: nuclear encoded ITS revisited. Protist 161:7-34.
- Nanjappa, D., Audic, S., Romac, S., Kooistra, W. H. & Zingone, A. 2014. Assessment of species diversity and distribution of an ancient diatom lineage using a DNA metabarcoding approach. PLoS ONE 9:e103810.
- Oh, S. J., Yoon, Y. H., Yamamoto, T. & Yang, H. -S. 2009. Effect of attachment substrate size on the growth of a benthic microalgae *Nitzschia* sp. in culture condition. J. Korean Soc. Mar. Environ. Energy 12:91-95 (in Korean with English abstract).
- Park, S. K., Kim, B. Y., Choi, H. G., Oh, J. -S., Chung, S. -O., An, K. -H. & Park, K. -J. 2013. Seasonal variation in species composition and biomass of microphytobenthos at Jinsanri, Taean, Korea. Korean J. Fish. Aquat. Sci. 46:176-185.
- Parsons, T. R., Maita, Y. & Lalli, C. M. 1984. *A manual of chemical and biological methods for seawater analysis*. Pergamon Press, Oxford, 173 pp.
- Rejil, T. 2013. Microalgal vegetation in the selected Mangrove Ecosystems of Kerala. Ph.D. dissertation, Cochin University of Science and Technology, Kalamassery, 262 pp.
- Roer, R. D. 2008. A genetic marker for coastal diatoms based on *psbA*. M.S. thesis, University of North Carolina, Wilmington, NC, 64 pp.
- Round, F. E. 1971. Benthic marine diatoms. Oceanogr. Mar. Biol. Ann. Rev. 9:83-139.
- Sato, S., Kooistra, W. H. C. F., Watanabe, T., Matsumoto, S. & Medlin, L. K. 2008. A new araphid diatom genus *Psammoneis* gen. nov. (Plagiogrammaceae, Bacillariophyta) with three new species based on SSU and LSU rDNA sequence data and morphology. Phycologia 47:510-528.
- Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J. & Weber, C. F. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microb. 75:7537-7541.
- Shokralla, A., Spall, J. L., Gibson, J. F. & Hajibabaei, M. 2012.

- Next-generation sequencing technologies for environmental DNA research, Mol. Ecol. 21:1794-1805.
- Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312-1313.
- Stief, P., Kamp, A. & de Beer, D. 2013. Role of diatoms in the spatial-temporal distribution of intracellular nitrate in intertidal sediment. PLoS ONE 8:e73257.
- Sullivan, M. J. & Moncreiff, C. A. 1988. Primary production of euphotic algal communities in a Mississippi salt marsh. J. Phycol. 24:49-58.
- Ter Braak, C. J. & Smilauer, P. 2002. CANOCO reference manual and CanoDraw for Windows user's guide: software for canonical community ordination (version 4.5). Available from: http://www.canoco5.com/. Accessed Sep 20, 2017.
- Theriot, E. C., Cannone, J. J., Gutell, R. R. & Alverson, A. J. 2009. The limits of nuclear-encoded SSU rDNA for resolving the diatom phylogeny. Eur. J. Phycol. 44:277-290.
- Trobajo, R., Mann, D. G., Clavero, E., Evans, K. M., Vanormelingen, P. & McGregor, R. C. 2010. The use of partial *cox*1, *rbc*L and LSU rDNA sequences for phylogenetics and species identification within the *Nitzschia palea* species complex (Bacillariophyceae). Eur. J. Phycol. 45:413-425.
- Underwood, G. J. C. 1994. Seasonal and spatial variation in epipelic diatom assemblages in the Severn estuary. Diatom Res. 9:451-472.
- Underwood, G. J. C. & Barnett, M. 2006. What determines species composition inmicrophytobenthic biofilms? *In* Kromkamp, J. C., de Brouwer, J. F. C., Blanchard, G. F., Forster, R. M. & Créach, V. (Eds.) *Functioning of Microphytobenthos in Estuaries*. Royal Netherlands Academy of Arts and Sciences, Amsterdam, pp. 121-138.
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M. & Rozen, S. G. 2012. Primer3: new capabilities and interfaces. Nucleic Acids Res. 40:e115.
- Visco, J. A., Apothéloz-Perret-Gentil, L., Cordonier, A., Esling, P., Pillet, L. & Pawlowski, J. 2015. Environmental monitoring: inferring the diatom index from next-generation sequencing data. Environ. Sci. Technol. 49:7597-7605.
- Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microbiol. 73:5261-5267.
- Watermann, F., Hillebrand, H., Gerdes, G., Krumbein, W. E. & Sommer, U. 1999. Competition between benthic cyanobacteria and diatoms as influenced by different grain sizes and temperatures. Mar. Ecol. Prog. Ser. 187:77-87.
- Yallop, M. L., de Winder, B., Paterson, D. M. & Stal, L. J. 1994. Comparative structure, primary production and bio-

- genic stabilization of cohesive and non-cohesive marine sediments inhabited by microphytobenthos. Estuar. Coast. Shelf Sci. 39:565-582.
- Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F. O., Ludwig, W., Schleifer, K. -H., Whitman, W. B., Euzéby, J., Amann, R. & Rosselló-Móra, R. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat. Rev. Microbiol. 12:635-645.
- Yoo, M. H. 2004. Seasonal distribution of microphytobenthos and primary production on an intertidal mud flat in Janghwa, Ganghwa Island of Korea. M.S. thesis, Inha

- University, Incheon, 96 pp. (in Korean with English abstract).
- Zimmermann, J., Glöckner, G., Jahn, R., Enke, N. & Gemeinholzer, B. 2015. Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. Mol. Ecol. Resour. 15:526-542.
- Zimmermann, J., Jahn, R. & Gemeinholzer, B. 2011. Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. Org. Divers. Evol. 11:173-192.