An attempt to semi-quantify potentially toxic diatoms of the genus *Pseudo-nitzschia* in Tokyo Bay, Japan by using massively parallel sequencing technology

SATOSHI NAGAI¹,*, SHINGO URUSIZAKI², YUKI HONGO¹, HUNGYEN CHEN¹ & NINA DZHEMBEKOVA³

¹Research Center for Bioinformatics and Biosciences, National Research Institute of Fisheries Science, 2–12–4 Fukuura, Kanazawa-ku, Yokohama, Kanagawa 236–8648, Japan
²AXIOHELIX Co. Ltd., 5–11 Hakozaki, Nihonbashi, Chuouku, Tokyo 103–0015, Japan
³Institute of Oceanology, Bulgarian Academy of Sciences, First May Street 40, 152, Varna 9000, Bulgaria

Received 19 June 2017; Accepted 7 September 2017 Responsible Editor: Satoshi Nagai

**Abstract:** Currently, no tool is available that precisely identifies and quantifies all 48 species of *Pseudo-nitzschia* simultaneously. In the present study, we used massively parallel sequencing-based technology with genus-specific *Pseudo-nitzschia* primers to investigate the distribution and population dynamics of the genus in Tokyo Bay, Japan. Furthermore, we attempted to quantify the abundance of each species using both relative abundance data detected in the massively parallel sequencing-based survey, and cell count data, obtained by light microscopy. Several species of *Pseudo-nitzschia* were detected continuously from July 10 to September 9, 2016 in the sampling location. Cell densities ranged from 2–300 cells mL⁻¹ (sample no. = 10). Fourteen operational taxonomic units (OTUs) attributable to *Pseudo-nitzschia* were detected, and 99.9% of sequences detected by massively parallel sequencing belonged to the genus. Each of the most abundant OTUs comprised a single species, and were identified, in decreasing order of abundance, as *P. multiseries* (Takano) Takano, *P. pungens* (Grunow ex Cleve) Hasle, *P. fraudulenta* (Cleve) Hasle, *P. multiseries* (Hasle) Hasle, and *P. galaxiae* Lundholm & Moestrup. Species identification was possible in 9 of the 14 OTUs detected, of which several were detected in Tokyo Bay for the first time. The cell abundance ranges of the five most abundant species were estimated as 1,976–246,395, 0.03–52,999, 5–42,820, 4–24,617, and 0.03–1,007 cells L⁻¹, respectively. The detection limit was 0.03 cells L⁻¹.

**Key words:** massively parallel sequencing, monitoring, operational taxonomic units, *Pseudo-nitzschia*, quantification, Tokyo Bay

**Introduction**

Marine diatoms of the genus *Pseudo-nitzschia* became a cause for concern globally following the first documented case of amnesic shellfish poisoning (ASP) in 1987 in Canada, which resulted in three human fatalities and 105 confirmed cases of acute intoxication (Perl et al. 1990). Interest in the genus increased again following the publication of evidence showing that the potent toxin produced by *Pseudo-nitzschia*, domoic acid (DA), is transferred through the food web, potentially causing mass mortality in wildlife (reviewed by (Bejarano et al. 2008)). *Pseudo-nitzschia* is distributed globally and includes 48 described species; 23 of these species are toxigenic, whereas the others are considered nontoxic (Teng et al. 2016). Furthermore, both toxic and benign *Pseudo-nitzschia* species can cause dense blooms, which may increase in frequency and duration depending on changes in nutrient loading (Anderson et al. 2002). In the United States, Canada, and the European Union, the regulatory limit for DA in bivalves intended for human consumption is set at 20 ppm (20 µg DA g⁻¹ shellfish tissue) to minimize the threat of acute DA exposure and ASP (Fernandes et al. 2014). Given the ecological and economic importance of *Pseudo-nitzschia*, and the health risks posed by DA accumulation, effective monitor-
ing of the genus is crucial. However, observations by light microscopy (LM) are not sufficiently precise to identify *Pseudo-nitzschia* species, particularly cryptic and pseudo-cryptic species; accordingly, data produced using this approach are considered equivocal (Lelong et al. 2012). Recent investigations using more sophisticated molecular analyses have revealed higher *Pseudo-nitzschia* species diversity than previously described using LM observations (McDonald et al. 2007, Ruggiero et al. 2015), and the use of such methods has become essential when dealing with the taxonomic difficulties of the genus (Lundholm et al. 2012).

New sequencing technologies, such as Roche 454 pyrosequencing and the Illumina MiSeq platform, have made it possible to obtain millions of sequence reads in a single experiment, and massively parallel sequencing (MPS) is currently revolutionizing surveys of eukaryotic diversity (Amaral-Zettler et al. 2009, Medinger et al. 2010, Tanabe et al. 2015). Massively parallel sequencing technology is able to detect most eukaryotes present in water samples from various ecosystems, and to detect low-abundance populations within complex eukaryote communities (Cheung et al. 2010, Nolte et al. 2010, Edgcomb et al. 2011, Orsi et al. 2011, Lindeque et al. 2013). However, MPS techniques remain problematic for two key reasons. First, such techniques are prone to overestimating the number of OTUs (operational taxonomic units), due to the formation of PCR-generated artifacts such as chimeric sequences (Qiu et al. 2001, Haas et al. 2011) and sequence errors derived from MPS processes (Kunin et al. 2010, Schloss et al. 2011, Tanabe et al. 2015); second, MPS techniques lack taxonomic identification power, due to sequencing length limitations and the limited quantities of sequences available in international nucleotide sequence databases (INSDs) (Bazinet & Cummings 2012, Tanabe & Toju 2013, Tanabe et al. 2015). Currently, in MPS surveys targeting the 18S-rRNA gene, the identification of eukaryote species is possible for 15–45% of OTUs detected in 500–1,000 mL seawater (Nagai et al. unpublished data).

*Pseudo-nitzschia* species are among the most abundant diatom groups in the coastal waters of Japan (Nishikawa et al. 2010, Nagai et al. 2016a). A survey targeting DA-producing *Pseudo-nitzschia* carried out in Ofunato Bay, Iwate Prefecture, Japan, identified a *P. multiseries* isolate that produced levels of DA comparable to those produced by Canadian strains. However, low levels of DA production (<2.5 pg cell\(^{-1}\)) were found in culture experiments, indicating that the toxic *P. multiseries* does not bloom in dangerously high densities in this bay (Kotaki et al. 1999, Kotaki et al. 2008). In other Japanese coastal waters, however, full studies of the distribution, population dynamics, and population genetics of *P. multiseries* have not yet been carried out. Recently, we conducted time-series monitoring experiments using amplicon sequences of partial 18S-rRNA genes in Japanese coastal waters (Nagai et al. 2016a, Nagai et al. 2016b, Nagai et al. 2016c). Unfortunately, identifying species in the genus *Pseudo-nitzschia* was rendered difficult by the low resolution of the target gene region (Nagai et al. 2016a, Nagai et al. 2016b). Thus, in the present study, we used primers specific to *Pseudo-nitzschia* (developed by McDonald et al. (2007)) in an MPS-based survey to investigate the distribution and population dynamics of *Pseudo-nitzschia* species in Tokyo Bay, Japan. We then evaluated the effectiveness of the genus-specific primer pair, i.e., the taxonomic identification power and of its capacity to quantify each species, using both the relative abundances of *Pseudo-nitzschia* species detected by MPS monitoring, and total cell counts at the genus level observed by LM.

**Materials and Methods**

**Environmental data**

Water temperature, salinity, and chlorophyll \(a\) content are measured every 20 min in Tokyo Bay at an offshore site in Kawasaki, nearby the Yokohama Port and Airport Technology Investigation Office, Kanto Regional Development Bureau, Japanese Ministry of Land, Infrastructure, Transport and Tourism. These data are published online (http://www.tbeic.go.jp CENTER/index.asp). We included surface layer data taken daily at 12:00 during the study period for our analysis.

**Sampling and DNA extraction**

Seawater samples were collected from the surface layer, using a plastic bucket, weekly from July 10 to September 9, 2016 at a sampling station in Tokyo Bay (35°34.60′N, 139°65.72′E), Japan (Fig. 1). The station is situated in water approximately 10 m deep, and seawater samples were collected in the harbor at 11:00 a.m. The abundance of *Pseudo-nitzschia* species in one mL of the seawater was counted by the morphology-based observation method according to Tomas et al. (1996) using an inverted microscope (TE-300, Nikon, Tokyo, Japan). To collect *Pseudo-nitzschia* from the samples, 1,000-mL seawater samples were filtered through polycarbonate filters with pores 8 \(\mu\)m in diameter (Nuclepore membrane; GE Healthcare, Tokyo, Japan), followed by filtering through filters with pores 1 \(\mu\)m in diameter (GE Healthcare). DNA was extracted shortly after filtration. In preparation for DNA extraction, a 5% Chelex\(^{\text{®}}\) suspension (Chelex 100 Molecular Biology Grade Resin; Bio-Rad Laboratories Inc., Richmond, CA, USA) was prepared by dispersing the resin into ultra-pure water (Tanabe et al. 2015, Nagai et al. 2016a, Nagai et al. 2016b). To ensure effective extraction of DNA from planktonic elements trapped on the filters, filters were cut in half and were placed in 1.5-mL tubes (A.150; Assist; Tokyo, Japan) with 150 \(\mu\)L 5% Chelex buffer. Plankton cells were crushed using a motorized pellet pestle (Kontes Glass; Vineland, NJ, USA) for 60 s, and 350 \(\mu\)L buffer was added to achieve a final volume of 500 \(\mu\)L. DNA was
extracted by heating the 1.5-mL tubes at 97°C for 20 min (Tanabe et al. 2015, Nagai et al. 2016a, Nagai et al. 2016b). DNA samples extracted from the 8-µm filter and the 1-µm filter were mixed equally (50 µL + 50 µL) and were used as template DNA. The DNAs were quantified using a Qubit 2.0 128 Fluorometer (Life technology, Carlsbad, CA, USA). Template DNA concentrations were 0.30–2.34 ng L⁻¹ (0.92±0.71, mean±SD).

**Paired-end library preparation and MiSeq sequencing**

In the present study, we used a genus-specific primer to perform time-series monitoring of *Pseudo-nitzschia* species (McDonald et al. 2007). Direct amplification of DNA from natural samples using *Pseudo-nitzschia*-specific primers is generally unsuccessful (McDonald et al. 2007), likely due to low copy numbers of the target gene in the genus. Therefore, we instead applied a nested PCR technique using first the universal and then the genus-specific primers, in a manner similar to that described by McDonald et al. (2007). The universal large subunit (LSU) rRNA primers D3Ca, F: ACG AAC GAT TTG CAC GTC AG and DIR, R: ACC CGC TGA ATT TAA GCA TA were used in the first PCR (ca. 800 bp in length). This PCR was carried out using a thermal cycler (PC-808; ASTEC, Fukuoka, Japan) in a reaction mixture (25 µL) containing 1.0 µL template DNA, 0.2 mM each dNTP, 1×PCR buffer, 1.5 µM Mg²⁺, 1.0 U KOD-Plus-ver.2 (TOYOBO, Osaka, Japan), which has intensive 3′→5′ exonuclease activity, and 1.0 µM of each primer. The following PCR cycling conditions were used: initial denaturation at 94°C for 3 min, followed by 30 cycles at 94°C for 15 s, 56°C for 30 s, and 68°C for 40 s. PCR amplification was verified using 1.5% agarose gel electrophoresis.

To construct DNA libraries in the MiSeq 250 PE platform (Illumina, San Diego, CA, USA), we used a set of genus-specific *Pseudo-nitzschia* LSU primers (D1-186F, F: GTT CCT TGG AAA AGG ACA GCT GA; D1-548R, R: AGA CAT CAA CTC TGA CTG; ca. 360 bp in length). Massively parallel paired-end sequencing using the MiSeq platform requires PCR amplicons to be flanked by: (i) primer-binding sites for sequencing; (ii) dual-index (i.e. barcode) sequences; and (iii) adapter sequences for binding to the flow cells of the MiSeq. The workflow we used was derived from the document ‘16S metagenomic sequencing library preparation: preparing 16S ribosomal gene amplicons for the Illumina MiSeq system,’ distributed by Illumina (part no. 15044223 Rev. B).

We used a two-step PCR approach to construct paired-end libraries. The first PCR step was carried out using the primers 5′-ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT+D1-186F (forward) and 5′-GTG ACT GGA GT T CAG AC G TGT G C T C T C G AT C T (reverse). PCR products from the LSU region (amplified by the D3Ca and DIR primer pair) were diluted 15 times in Milli-Q water and were used as the templates in the first step. The first PCR in the two-step approach was prepared in the same manner as that targeting the LSU region, with the following cycling conditions: initial denaturation at 94°C for 3 min, followed by 30 cycles at 94°C for 15 s, 56°C for 30 s, and 68°C for 40 s. PCR amplification was verified using 1.5% agarose gel electrophoresis. PCR products were purified using an Agencourt AMPure XP kit and eluted in 25 µL TE buffer, according to the protocol specified by the manufacturer (Beckman Coulter, Brea, California, USA).

The second PCR in the two-step approach would require further details to complete the description.
used the PCR products of the first step as a template, and amplified regions using the primers 5′-AAT GAT ACG GGC ACC GAC GAT AC-3′ and 5′-CAA GCA GAA GAC GCC ATA CGA GAT-3′. The eight base segments function as dual-index sequences that identify each sample; the 5′-end sequences are adapters that allow the final product to bind to or hybridize with short oligonucleotides on the surface of the Illumina flow cell; and the 3′-end sequences are priming sites for MiSeq sequencing. The second PCR step was carried out using the same protocols as the first, except that the volume of the reaction mixture was 50 µL due to the addition of 2.0 µL purified DNA. PCR cycling was carried out as follows: initial denaturation at 94°C for 3 min, followed by 7–8 cycles at 94°C for 15 s, 59°C for 30 s, and 68°C for 40 s. PCR amplification was verified using agarose gel electrophoresis, and PCR products were purified using an Agencourt AMPure XP kit (Beckman Coulter, USA). Amplified PCR products were quantified, and the indexed products from the second PCR were pooled in equal concentrations and stored at −30°C until subsequent sequencing analyses.

A PhiX DNA spike-in control (Illumina, USA) was mixed with the pooled DNA library to improve the quality of data generated by low-diversity samples such as single PCR amplicons. DNA concentrations in the pooled library and in the PhiX DNA were adjusted to 4 nM using buffer EB (10 mM Tris-HCl pH8.5), and mixed at a ratio of 7:3. 4-nM library DNA (5 µL) was denatured using 5 µL fresh 0.1 N NaOH. The denatured library DNA (10 µL; 2 nM), including HT1 buffer (sourced from the Illumina MiSeq v.2 Reagent kit for 2×150 bp PE), was diluted to a final concentration of 12 pM for sequencing on the MiSeq platform.

Treatment of MPS data and selection of OTUs

Nucleotide sequences were demultiplexed according to the 5′-multiplex identifier (MID) tag, and primer sequences were demultiplexed according to the default format in MiSeq. Sections containing (1) palindrome clips longer than 30 bp and (2) monomonomers longer than 9 bp were trimmed from the sequences at both ends. 3′ tails with an average quality score lower than 30 at the end of the final 25-bp window were also trimmed from each sequence. 5′ and 3′ tails with an average quality score lower than 20 at the end of the final window were also trimmed. Sequences longer than 250 bp were truncated to 250 bp by trimming the 3′ tails. Trimmed sequences shorter than 200 bp were filtered out. Demultiplexing and trimming were carried out using Trimmomatic version 0.35 (http://www.usadellab.org/cms/?page=trimmomatic). The remaining sequences were merged into paired reads using Usearch version 8.0.1517 (http://www.drive5.com/usearch/). Sequences were aligned with each other using Clustal Omega v 1.2.0.

(http://www.clustal.org/omega/). Multiple sequences were aligned with each other, and sequences were only extracted if they appeared in more than 75% of the read positions. Filtering and some of the multiple alignment processes were carried out using the screen.seqs and filter.seqs commands in Mothur, based on MiSeq SOP (http://www.mothur.org/wiki/MiSeq_SOP) (Schloss et al. 2011). Identical sequences (100% similarity) were collated using the unique.seqs command in Mothur. Erroneous and chimeric sequences were detected and removed in Mothur using the pre.cluster (diffs=4) and chimera.uchime (minh=0.1; http://drive5.com/usearch/manual/uchime_algo.html) (Edgar et al. 2011) commands, respectively. Contiguous sequences identified using count.seqs in Mothur were considered OTUs and were used in the subsequent taxonomic analysis. Demultiplexed and filtered sequence data that were not trimmed were deposited in the DDBJ Sequence Read Archive under accession no. DRA005785.

Taxonomic identification of OTUs

Selected OTUs were identified as follows. A subset of nucleotide databases, consisting of sequences that satisfied the following conditions, was prepared for BLAST search. To populate this subset, one search keyword was selected from among “ribosomal,” “rrna,” and “rdna;” however, “protein” was not included in the title. In the taxonomy search, the keywords “metagenome,” “uncultured,” and “environmental” were not included. Sequences of GenBank IDs retrieved from the nucleotide database that matched these criteria, downloaded from the NCBI FTP server, were extracted on July 2, 2012, and were used to construct a template sequence database. Subsequently, we identified each OTU using BLAST search (Cheung et al. 2010). The search was conducted using the default parameters in NCBI BLAST+ 2.2.26+ (Camacho et al. 2009), using the nucleotide subset described above as the database, and using all OTU sequences identified as above as the query. Taxonomic information was obtained from the BLAST hit with the highest bitscore for each query sequence. Singletons or OTUs with few reads contain artifacts caused by sequence or chimeric errors (Achaz 2008, Knudsen & Miyamoto 2009) that may, in turn, lead to the overestimation of species diversity (Huse et al. 2010, Bazinet & Cummings 2012). To avoid such overestimation, OTUs with few sequences (<15 sequences in 10 samples) were excluded from the analysis.

Checks of primer specificity and universality

The specificity and universality of the genus-specific Pseudo-nitzschia primers provided by MacDonald et al. (2007) were checked using NCBI/Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) against the NT database (nr, non-redundant). In the in silico amplification, up to two mismatches between each target primer and the template sequence were allowed. Sequences of the D1/D2 regions in 41 of the 48 species of Pseudo-nitzschia are
currently available in GenBank. These sequences were selected at random and were aligned with the representative sequences in each OTU obtained in the present study using the Clustal W algorithm in Molecular Evolutionary Genetics Analysis (MEGA) version 7.0 (Kumar et al. 2016), using default settings. The target sequences, all of which were susceptible to amplification using the primer pair, were edited manually. Primer regions were removed during the editing process. An optimal base substitution model was calculated using the default settings of Modeltest in MEGA, based on which K2+G+I was chosen as the best-fit model. All sites, including indels were used for model selection and phylogenetic tree construction. A maximum likelihood (ML) tree was generated using the default settings in MEGA. Bootstrapping with 100 replications was used to estimate the reliability of the phylogenetic tree generated in this way.

Results

Environmental conditions during sampling

Water temperature, salinity, and chlorophyll $a$ content ranged from 23.7–28.6°C, 20.0–31.2, and 3.7–92.4 µg L$^{-1}$, respectively, during the survey period (Fig. 2, Fig. 3). Water temperature exceeded 28°C on several days during the survey. Salinity was relatively stable until mid-August; however, salinity decreased substantially on four occasions following August 21, 2016, during which times it fluctuated between 20–30, depending on rainfall. Chlorophyll $a$ content in the water was high throughout the survey period (20.8±16.9 µg L$^{-1}$, mean±SD), and increased after August 21.

Pseudo-nitzschia abundance

Pseudo-nitzschia abundance was estimated by counting under a light microscope. Pseudo-nitzschia species appeared consistently throughout the survey period. Cell densities ranged from 2–300 cells mL$^{-1}$ (114.2±85.2 cells mL$^{-1}$, mean±SD; Fig. 3).

Numbers of OTUs and sequences detected by MPS belonging to Pseudo-nitzschia

Basic information concerning the total numbers of OTUs and sequences detected by MPS attributable to Pseudo-nitzschia is summarized in Table 1. The number

<table>
<thead>
<tr>
<th>No.</th>
<th>similarity tophit_name</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pseudo-nitzschia multistriata</td>
<td>521,176</td>
</tr>
<tr>
<td>2</td>
<td>Pseudo-nitzschia pungens</td>
<td>133,676</td>
</tr>
<tr>
<td>3</td>
<td>Pseudo-nitzschia fraudulenta</td>
<td>50,213</td>
</tr>
<tr>
<td>4</td>
<td>Pseudo-nitzschia multiseries</td>
<td>29,711</td>
</tr>
<tr>
<td>5</td>
<td>Pseudo-nitzschia galaxiae</td>
<td>10,003</td>
</tr>
<tr>
<td>6</td>
<td>Pseudo-nitzschia fakuyoi</td>
<td>4,618</td>
</tr>
<tr>
<td>7</td>
<td>Pseudo-nitzschia cuspidata</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Pseudo-nitzschia pseudodelicatissima</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Pseudo-nitzschia americana</td>
<td>2,148</td>
</tr>
<tr>
<td>10</td>
<td>Pseudo-nitzschia subfraudulenta</td>
<td>434</td>
</tr>
<tr>
<td>11</td>
<td>Pseudo-nitzschia brasiliiana</td>
<td>315</td>
</tr>
<tr>
<td>12</td>
<td>Pseudo-nitzschia linea</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Pseudo-nitzschia multiseries</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Pseudo-nitzschia decipiens</td>
<td>266</td>
</tr>
<tr>
<td>15</td>
<td>Pseudo-nitzschia brasiliiana</td>
<td>141</td>
</tr>
<tr>
<td>16</td>
<td>Pseudo-nitzschia dolorosa</td>
<td>79</td>
</tr>
<tr>
<td>17</td>
<td>Pseudo-nitzschia delicatissima</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Pseudo-nitzschia micropora</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Pseudo-nitzschia sp.</td>
<td>28</td>
</tr>
<tr>
<td>20</td>
<td>Pseudo-nitzschia subpacifica</td>
<td>25</td>
</tr>
<tr>
<td>21</td>
<td>Pseudo-nitzschia heimii</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Neodenticula seminae</td>
<td>25</td>
</tr>
<tr>
<td>23</td>
<td>Skeletonema marinoi</td>
<td>17</td>
</tr>
<tr>
<td>24</td>
<td>Skeletonema dohrnii</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Skeletonema costatum</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Minidiscus conicus</td>
<td>15</td>
</tr>
<tr>
<td>27</td>
<td>after data treatment</td>
<td>825,764</td>
</tr>
<tr>
<td>28</td>
<td>raw data</td>
<td>1,142,367</td>
</tr>
</tbody>
</table>
of sequences detected by MPS ranged from 94,621–130,179 per sample (114,237±10,211, mean±SD), and the number of sequences detected by MPS after removing sequence errors and chimeras ranged from 75,289–89,856 per sample (82,576±5,698, mean±SD). The maximum number of sequences detected by MPS was 1.19 times larger than the minimum number of sequences detected by MPS. However, as a precaution, any potential effects of this bias on subsequent analyses was minimized by carrying out random re-sampling of the samples with the minimum number of sequences detected by MPS (75,289); following these procedures, 17 OTUs were detected. Fourteen of the 17 OTUs were attributed to *Pseudo-nitzschia*, and 99.9% of the sequences detected by MPS belonged to the genus. The most abundant OTU was identified as *P. multistriata*, followed in decreasing order of abundance by OTUs identified as *P. pungens, P. fraudulenta, P. multiseries, and P. galaxiae*, and the relative abundances were 69.2, 17.8, 6.7, 4.0 and 1.3%, respectively, of all *Pseudo-nitzschia* species detected during monitoring. Species identification was possible in 10 of the 14 *Pseudo-nitzschia* OTUs (Table 1).

**Changes in species composition over time**

The relative abundance of *P. multistriata* among all *Pseudo-nitzschia* species ranged from 17.5–98.8% over time. *Pseudo-nitzschia multistriata* monopolized on July 31, following which it was strongly predominant (Fig. 4). The relative abundance of *P. pungens* ranged from 0.01–66.2%, with a peak in abundance on July 17. Similarly, the abundance of *P. fraudulenta* ranged from 0.07–26.8%, and exceeded 20% on July 10 and July 24. The combined relative abundances of the five most abundant species ranged from 93.4–99.9% (mean±SD, 98.9±0.02%) during the survey period.

The cell densities of each species were estimated based on their relative abundances, supplemented by cell counts at the genus level carried out by light microscopy. The cell abundance ranges of the five most abundant species were estimated as 1,976–246,395, 0.03–52,999, 5–42,820, 4–24,617, and 0.03–1,007 cells L⁻¹, respectively (Fig. 5). The detection limit was 0.03 cells L⁻¹, which is equivalent to 1.2–2.5 copies of the rRNA gene (Penna et al. 2013). The OTU identified as *P. subpacifica* (Hasle) Hasle and/or *P. heimii* Manguin had the lowest abundance, with cell density ranging from 0.03–33 cells L⁻¹.

**Checks of primer universality and specificity**

The universality and the specificity of the genus-specific *Pseudo-nitzschia* primer pair was checked using NCBI/Primer-BLAST in the NT database, resulting in successful amplification of the target sequences in 39 of the 41 *Pseudo-nitzschia* species detected in the present study. However, due to mismatches with the reverse primer, sequences belonging to *P. abrensis* Pérez-Aicua & Orive and *P. batesiana* Lim, Teng, Leaw & Lim were not amplified. In the *in silico* amplification, in which up to 2 bp mismatches were permitted, sequences attributed to *Fragilariopsis curta* (Van Heurck) Hustedt (AF417659), *F. cylindrus* (Grunow) Helmcke & Krieger (AF417657), *F. kerguelensis* (O’Meara) Hustedt (KC833024), *F. rhombica* (O’Meara) Hustedt 1952 (AF417656), *F. vanheurkii* (Peragallo) Hustedt (AF417660), *Neodenticula seminae* (Simonsen & Kanaya) Akiba & Yanagisawa (AF417660), *Nitzschia cf. fonticola* (HF679151), and *Nitzschia soratensis* Morales & Vis (HF679198) were amplified. In the field survey, we detected three OTUs belonging to three other genera (*Neodenticula, Skeletonema*, and *Minidiscus*); however, the proportion of sequences belonging to these species was <0.01% of the total number of sequences detected by MPS.

A maximum likelihood phylogenetic tree was constructed to investigate whether species could be identified by checking the genetic relatedness of the target region between species (Fig. 6). Overall, the proportion of nodes with >70% bootstrap values was not high (11 of 33 nodes), and individual branches often contained several species, indicating that sequence variation between species was not high enough in this region to enable the use of this primer pair for species identification. In fact, species identification of four OTUs (1. *Pseudo-nitzschia cuspidata* (Hasle) Hasle/P. fukuyoi Lim, Teng, Leaw & Lim/P. pseudodelicatissima (Hasle) Hasle; 2. *P. linea* Lundholm, Hasle and Fryxell/P. brasiliiana Lundholm, Hasle & Fryxell; 3. *P. dolorosa* Lundholm & Moestrup/P. delicatissima (Cleve) Heiden/P. micropora Priisholm, Moestrup & Lundholm; 4. *P. subpacifica/heimii*), was not possible at all using this method (Table 1, Fig. 6).

**Discussion**

Yap-Dejeto et al. (2010) identified 11 species of *Pseudo-nitzschia* in water sampled in Tokyo Bay based on morphological characteristics observed by light microscope and transmission electron microscope: *P. americana* (Hasle) Fryxell in Hasle, *P. brasiliiana, P. cayantha* Lundholm, Moestrup & Hasle, *P. calliantha* Lundholm, Moestrup...
S. Nagai et al.

254

& Hasle, P. delicatissima, P. galaxiae, P. fraudulenta, P. multistriata, P. multiseries, P. pseudodelicatissima, and P. pungens. In the present study, we detected 9 species, including seven of those identified by Yap-Dejeto et al. (2010) (all except P. caciantha, P. calliantha, P. delicatissima, and P. pseudodelicatissima; Table 1), and the newly detected species P. subfraudulenta (Hasle) Hasle, P. decipiens, and P. sp., suggesting at least 14 species of Pseudo-nitzschia species are present in Tokyo Bay. However, it is possible that we were unable to detect P. abrensis and P. batesiana, the remaining two species, because of mismatches in the reverse primer we used to identify sequences; if this is the case, it may be necessary to design a new reverse primer for these two species, and to combine this new reverse primer with the existing reverse primer to detect all possible species. Nagai et al. monitored Pseudo-nitzschia at several localities in Japan using the same technique as that used in the present study; their results indicate that species composition is region-specific (Nagai et al. unpublished data), and that P. multistriata is one of the most abundant species in Tokyo Bay (Fig. 4). The detection limit of Pseudo-nitzschia cells in the present study was calculated as 0.03 cells L\(^{-1}\), a value much lower than the detection limit of microscopic observation (the practical limit of which may be 5–10 cells L\(^{-1}\)), clearly demonstrating the advantages of MPS technology.

Fig. 5. Changes in cell abundances of Pseudo-nitzschia species during the survey period, calculated by combining massively parallel sequencing data with relative abundance data calculated by cell counts using a light microscope.
MPS-based monitoring of *Pseudo-nitzschia*

Among the factors that affect calculations of the relative abundances of species detected using MPS-based technology, three are considered critical: the copy number of rRNA genes in a single cell; DNA extraction efficiency; and PCR amplification efficiency, which varies according to the presence of PCR inhibitors (Nagai et al. 2012, Schrader et al. 2012). However, we were unable to take these factors into account when quantifying cell abundance in the present study because it is prohibitively difficult to evaluate these factors separately, particularly where environmental DNA is concerned. Fortunately, an existing study, carried out by Penna et al. (2013), reports copy numbers per cell of the ITS-5.8S-rRNA gene, as quantified with qPCR using 106 clonal strains of three *Pseudo-nitzschia* species: *P. delicatissima*, *P. calliantha*, and *P. pungens*, which have similar copy numbers, at 39±22, 62±30, and 83±67, respectively. Given these data, it may not be necessary to take into account per-cell differences in rDNA copy numbers between species. Furthermore, PCR amplification is known to introduce a significant bias favoring more abundant species, which become even more predominant during amplification. To obtain a precise correction formula, future studies might consider introducing plasmid DNA specific to the target gene in several *Pseudo-nitzschia* species at different concentrations into eDNA samples. Additionally, we have to consider comparing cell abundances using other quantification methods, such as qPCR, to evaluate estimates calculated by other techniques. We also recommend to filter seawater through filters with pore 1 µm in diameter directly, and not to use two filters (8 and 1 µm) to extract the DNA for more precious quantification of the abundance.

Two important factors that influence the power of taxonomic identification are amplicon sequence variability and the number of sequences of the target region deposited in INSDs (Tanabe et al. 2015). We found sequences of the target region in INSDs for 41 of the 48 *Pseudo-nitzschia* species, and data concerning the remaining seven species have not yet to be deposited in any INSDs; furthermore, we confirmed that it is possible to amplify these sequences using the genus-specific primer pair in 39 of these 41 species, indicating that the primer is adequately universal. However, sequence variability was not high enough to differentiate between several species (Fig. 6). In particular, it was impossible to differentiate between *P. cuspidata* (Hasle) Hasle, *P. fukuyoi*, *P. plurisecta* Orive & Pérez-Aicua and *P. pseudodelicatissima* (OTU 6); *P. dolorosa* and *P. micropora* (OTU 12); *P. heimii* and *P. subpacific* (OTU 14); and *P. brasiliana* and *P. linea* (OTU 9). Using the study carried out by Yap-Dejeto et al. (2010) as a basis for speculation, OTUs 6 and 12 may have represented *P. pseudodelicatissima* and *P. delicatissima*, respectively (Table 1). It is likely we would have been able to identify all species if, in addition to the MPS procedure carried out, we had also collected *Pseudo-nitzschia* cells (>100 cells) using a capillary pipette once every three or four samples, and obtained LSU including ITS sequences using sub-cloning techniques carried out on the pipetted cells (Nagai et al. 2008, Teng et al. 2016). Therefore, time-series monitoring of *Pseudo-nitzschia* using MPS-based techniques alongside the sub-cloning of known *Pseudo-nitzschia* cells may prove useful, especially in regions where highly toxic and non-toxic *Pseudo-nitzschia* species co-appear.

Under culture conditions, cellular DA concentrations reach up to 140 pg cell$^{-1}$ in the highly toxic *P. multiseriess*

---

**Fig. 6.** A maximum-likelihood (ML) tree, created using the sequences of a partial 28S-rRNA (LSU) gene, showing the relationships between species of the genus *Pseudo-nitzschia*. ML bootstrap values (≥50%) are placed close to each node. Accession numbers are shown to the right of the species name. Black circles indicate the representative sequence of each OTU identified in the present study. Numbers beside the black bars correspond to the OTUs in which several species were equally likely to represent the candidate.
(Trainer et al. 2012), depending on cell size, species, and growth conditions; in the field, cellular DA concentrations reach up to 117 pg cell⁻¹ in *P. australis* and *P. cuspidata* (Caron et al. 2010). Furthermore, due to bioaccumulation, DA concentrations up to 3,100 µg g⁻¹ tissue have been measured in scallops (*Placopecten magellanicus* Gmelin; reviewed in Landsberg (2002)). Accordingly, *Pseudo-nitzschia australis* is well known for its high toxicity, and was responsible for mortality in sea birds (pelicans and cormorants) and sea lions that predated on DA-contaminated anchovies (Fritz et al. 1992, Scholin et al. 2000).

In Japan, the rate of DA production by *Pseudo-nitzschia* isolates in culture was substantially lower than rates reported by Kotaki (2008) and Trainer et al. (2012). Domoic acid has been detected in bivalves in several localities in Japan; however, the levels of DA detected were substantially lower than those detected in the United States and Canada (Fernandes et al. 2014, Suzuki 2016). Consequently, exhaustive monitoring of *Pseudo-nitzschia* has not been carried out to date in Japanese coastal waters. More recent studies have indicated that chronic exposure to even low levels of DA in humans can cause negative effects at the cellular level, which may in turn cause chronic health problems (Hiolski et al. 2014). Thus, monitoring programs that focus only on the detection of DA concentrations higher than the regulatory limit in shellfish do not guarantee the safety of food or the protection of human health. Furthermore, there is compelling evidence that the geographic ranges of both the toxic dinoflagellate *Alexandrium tamarense* species complex and the bivalve killer dinoflagellate *Heterocapsa circularisquama* have been expanded by human-assisted means, such as translocation of oyster spats and/or cargo vessel ballast water (Anderson 1989, Hallegraeff & Bolch 1992, Honjo et al. 1998, Nagai et al. 2007, Nagai et al. 2009, Matsuyama et al. 2010). These findings suggest that, in the future, populations of highly toxic *Pseudo-nitzschia* may be introduced from elsewhere and may coexist with existing populations in Japanese coastal waters. Therefore, we recommend monitoring the cell abundances and toxicity levels of *Pseudo-nitzschia* populations, or monitoring DA levels in bivalves, to ensure the safety of seafood is maintained.

**Acknowledgements**

We thank Dr. Chui Pin Leaw, Institute of Ocean and Earth Sciences, University of Malaya, for her technical support and encouragement during this study. This work was supported by a Grant-in-Aid ("Marine Metagenomics for Monitoring the Coastal Microbiota") from the Ministry of Agriculture; a Grant-In-Aid for Scientific Research ("Kiban-B") from the Japan Society for the Promotion of Science (No. 25292130); and a Grant-in-Aid from the Japanese Society for the Promotion of Science Fellows (grant number: 02523180).

**References**


Hiolski EM, Kendrick PS, Frame ER, Myers MS, Bammle TK,


Teng ST, Tan SN, Lim HC, Dao VH, Bates SS, Leaw CP (2016) High diversity of *Pseudo-nitzschia* along the northern coast of Sarawak (Malaysian Borneo), with descriptions of *P. biperitita* sp. vov. and *P. limii* sp. nov. (Bacillariophyceae). J Phycol. DOI: 10.1111/jpy.12448