Patterns of diatom diversity correlate with dissolved trace metal concentrations and longitudinal position in the northeast Pacific coastal–offshore transition zone

P. Dreux Chappell[1,2,*], E. Virginia Armbrust[3], Katherine A. Barbeau[4],
Randelle M. Bundy[3,4], James W. Moffett[5], Jagruti Vedamati[5], Bethany D. Jenkins[1]

[1]Department of Cell and Molecular Biology, University of Rhode Island, Kingston, RI 02881, USA
[2]Department of Ocean, Earth and Atmospheric Sciences, Old Dominion University, Norfolk, VA 23529, USA
[3]School of Oceanography, University of Washington, Seattle, WA 98195, USA
[4]Scripps Institution of Oceanography, University of California at San Diego, La Jolla, CA 92093, USA
[5]Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089, USA

ABSTRACT: Diatoms are important primary producers in the northeast Pacific Ocean, with their productivity closely linked to pulses of trace elements in the western high nitrate, low chlorophyll (HNLC) region of the oceanographic time series transect 'Line P.' Recently, the coastal–HNLC transition zone of the Line P transect was identified as a hotspot of phytoplankton productivity, potentially controlled by a combination of trace element and macronutrient concentrations. Here we describe diatom community composition in the eastern Line P transect, including the coastal–HNLC transition zone, with a method using high-throughput sequencing of diatom 18S gene amplicons. We identified significant correlations between shifting diatom community composition and longitude combined with concentrations of dissolved copper and 2 other dissolved trace metals (dissolved Fe [dFe] and/or dissolved zinc) and/or a physical factor (salinity or density). None of these variables on its own was significantly correlated with shifts in community composition, and 3 of the factors (dFe, salinity, and density) correlated with one another. Longitude could incorporate multiple factors that may influence diatom communities, including distance from shore, proximity of sampling stations, and an integration of previous pulses of macro- and micro-nutrients. We also evaluated *in situ* Fe limitation of the diatom *Thalassiosira oceanica* using a quantitative reverse-transcription polymerase chain reaction method, and found biological evidence of Fe stress in samples from the coastal–HNLC transition zone. Combined, our results support a prior hypothesis that dissolved trace metals as well as longitudinal distance may be important to diatom diversity in the coastal–HNLC transition zone of the Line P transect.

KEY WORDS: Diatom diversity · Iron · Line P · Transition zone

1. INTRODUCTION

Diatoms are a diverse group of eukaryotic photosynthetic organisms, with an estimated 200 000 species (Mann & Droop 1996) found throughout marine and freshwater environments (Round et al. 1990). Diatoms are responsible for up to 40% of oceanic primary production and are especially important in nutrient-rich coastal zones, where they generate up to 90% of the organic carbon fueling the ecosystem (Nelson et al. 1995). In addition to nutrient-rich coastal zones, where they dominate phytoplankton communities, diatoms can account for a large amount of carbon and biogenic silica content in open ocean environments (Leblanc et al. 2012). Diatoms are also particularly responsive to iron (Fe) inputs in Fe-limited systems, including the Subarctic North Pacific Ocean (Boyd et al. 2005). These results imply that

*Corresponding author: pdchappe@odu.edu

diatoms have the ability to persist when Fe supplies are low, and can respond quickly to Fe inputs.

The Line P program is an important oceanographic time series transect that has been running since the late 1950s and has provided valuable insights into what drives phytoplankton productivity in the northeast Pacific Ocean (Freeland 2007). The transect traverses a range of oceanographic environments, starting from the nutrient-replete coastal region driven by periodic upwelling, through a transition zone to the high nitrate, low chlorophyll (HNLC) region in the farthest west portion of the transect (Whitney et al. 1998). Phytoplankton productivity at the westernmost station of the transect, Stn P26 (50.00° N, 145.00° W), has been well established as a site of chronic Fe limitation (Harrison 2002) and is the site where Martin developed his 'Iron Hypothesis' that Fe insufficiency limits phytoplankton growth and productivity (Martin & Fitzwater 1988). Since Martin's groundbreaking work (Martin 1990), a number of other studies have shown that Fe is an important driver of phytoplankton productivity at P26, and that chronic Fe-limiting conditions stretch along much of the Line P transect, at least as far east as Stn P20 located at 49.57° N, 138.67° W (La Roche et al. 1996, Boyd et al. 1998, 2004, Harrison 2002, Marchetti et al. 2006, 2012). More recently, the transition zone between the coastal and HNLC regions has become an area of increasing interest, as it has been associated with elevated productivity (Peña & Varela 2007, Ribalet et al. 2010). The exact boundaries of the transition zone vary as far west as P12 and as far east as P2, depending on whether it is defined by nutrient concentrations or salinity gradients, but consistently includes P4–P6 (Whitney et al. 1998, Ribalet et al. 2010; see Fig. 1 for P4–P6). Questions persist as to the potential drivers of productivity in this coastal–HNLC transition zone, with one hypothesis being that trace element variability is influencing phytoplankton productivity in this region (Ribalet et al. 2010).

In addition to evaluating phytoplankton productivity as a whole, it is important to examine how the phytoplankton community composition shifts along the transect in response to these changing environmental gradients. Large diatoms dominate the phytoplankton community in the western Vancouver coastal zone, where the Line P transect begins (Taylor & Haigh 1996). As the Line P transect moves west into the coastal–HNLC transition zone, the larger diatoms (>5 µm) become rare, except during occasional blooms (Boyd & Harrison 1999, Boyd et al. 2004, Peña & Varela 2007). On the westernmost portion of the transect, the phytoplankton community is usually dominated by smaller phytoplankton including cyanobacteria like *Synechoccocus* spp. (Miller et al. 1991, Booth et al. 1993). However, even in the western portion of the transect, approaching and including P26, diatoms remain present at low concentrations, as they have been shown to bloom rapidly in response to Fe inputs in incubation experiments (Boyd et al. 2004, Marchetti et al. 2006). A recent study using optical estimates of phytoplankton size and taxonomic groups along Line P that included historical high-performance liquid chromatography (HPLC) pigment data since 2007, confirmed that diatoms dominate phytoplankton communities when chlorophyll *a* (chl *a*) is above 1 mg m^{-3}, and are usually at least 20% of phytoplankton biomass, even when chl *a* is 0.1 mg m^{-2} (Zeng et al. 2018). Because of the persistence and the blooming potential of the diatoms, gaining an understanding of how the diatom community shifts along the transect is important to our broader understanding of phytoplankton productivity in this region.

Many diatom taxa cannot be identified and enumerated at the species level with light microscopy and cell counts, a fact that complicates our understanding of diatom community composition in mixed-assemblage field samples. While there have been recent advances developing molecular methods to identify species assemblages for individual genera of diatoms (Hubbard et al. 2008, 2014, Chappell et al. 2013), these methods have thus far remained limited in their ability to evaluate diatom species assemblages across multiple genera. Identification of diatoms to the species level permits the application of methods targeting biogeochemically relevant functional genes within prevalent taxa to explore the relationship between species and gene activity, in addition to allowing for evaluation of environmental drivers of community composition. Previous indicators of diatom Fe stress have largely been obtained from field incubation studies that observe diatom growth after adding Fe (Boyd et al. 2007), thus assuming *in situ* diatom Fe stress after observing a chl *a* response following the relief of Fe limitation. However, the metabolic status of the diatom community *in situ* cannot be understood from these experiments, as incubation experiments themselves can drive diatom communities into Fe starvation (Marchetti et al. 2017). Two diatom-specific genes have been shown to be upregulated in times of Fe stress in cultures and in natural populations (Whitney et al. 2011, Chappell et al. 2015) and can thus be used as a quantitative real-time indicator of metabolic Fe status. This type of molecular approach is especially

useful for field expeditions that do not allow time for growth-dependent incubation experiments to observe diatom Fe limitation (such as large-scale GEOTRACES cruises), and instead provides a diatom-specific genetic proxy for Fe limitation.

In this work, we describe diatom community composition using a method that involves high-throughput sequencing of diatom 18S amplicons in samples from 2 distinct regions along Line P: (1) the nutrient-replete coastal region and (2) the coastal–HNLC transition zone often characterized by sharp nutrient gradients. The short sampling expedition did not allow us to reach the third region, the historically severely Fe-limited HNLC region. We also present evidence for longitudinal distance and dissolved trace metals severly (Fe, zinc [Zn], and copper [Cu]) as potential drivers of the shifts in the diatom community, by examining diatom Fe status *in situ* using previous quantitative molecular methods developed by Chappell et al. (2015) in combination with hydrographic trace element data.

2. MATERIALS AND METHODS

2.1. Sample collection

Samples were collected from the northeast Pacific Ocean aboard the RV 'Thomas G. Thompson' (cruise TN280) from 17–22 May 2012 (Fig. 1). Historical Line P stations were sampled starting from P8 and moving shoreward to P1. A rosette mounted with a conductivity/temperature/depth profiler (CTD; Sea-Bird SBE-911+) was used to collect seawater for molecular and nutrient analyses at 5 m depth at each station. Fluorescence was measured by a WET Labs ECO-AFL/FL incorporated into the SeaBird SBE-911+ system. Samples for ammonium and nitrate concentration measurements were 0.2 μm filtered and measured aboard ship. Ammonium was measured using the o-phthaldialdehyde fluorescence method (Holmes et al. 1999), and nitrate was measured using the standard spectrophotometric method (Grasshoff et al. 1999). Samples for macronutrient analysis (nitrate, phosphate, and silicate) were 0.2 μm filtered and frozen at sea and later measured, according to the recommendations of Gordon et al. (1993), with a Technicon autoanalyzer II system by the University of Washington marine chemistry laboratory.

For molecular sampling, approximately 1 l of surface water was filtered through 25 mm, 3.0 μm polyester filters using a peristaltic pump on a low setting. Filters for DNA extraction were placed in 2 ml cryotubes filled with Qiagen® AP1 Buffer, frozen in liquid nitrogen, and stored at −80°C until processing. Filters for RNA extraction were placed in 2 ml cryotubes filled with Qiagen® RLT Buffer, frozen in liquid nitrogen, and stored at −80°C until processing. Samples for trace element analysis were collected using the GEOTRACES sampling rosette system (Cutter & Bruland 2012). This system uses 12 l GO-FLO sampling bottles (General Oceanics) that feature close-open-close operation such that the bottles remain
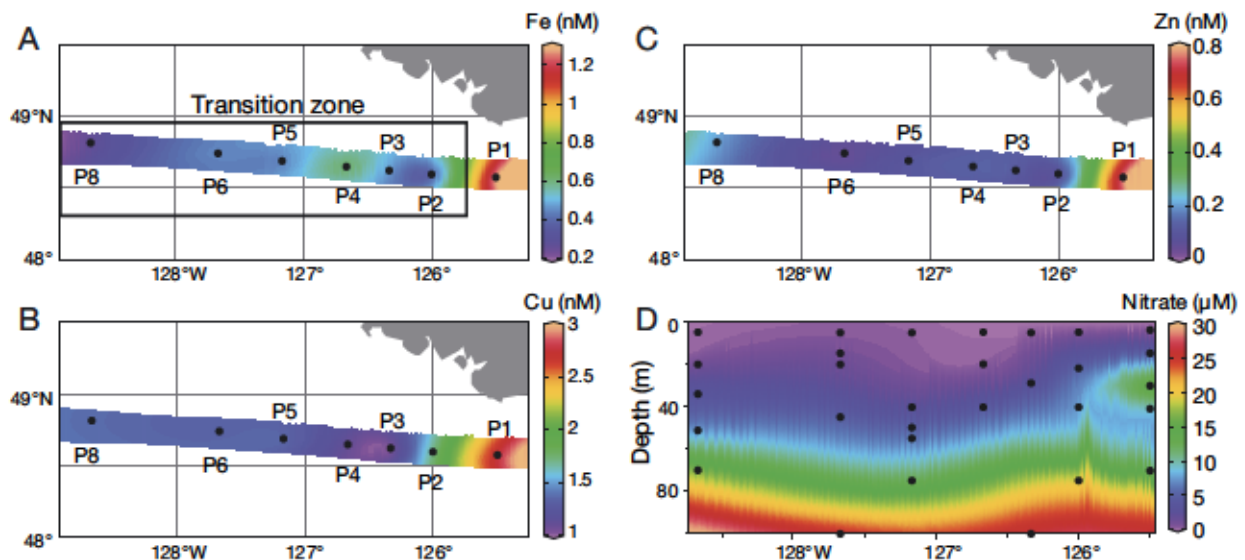

Fig. 1. Stations sampled on cruise TN280 from 17–22 May 2012 shown over a background of surface concentrations of (A) dissolved iron (Fe), (B) dissolved copper (Cu), and (C) dissolved zinc (Zn). (D) Depth section plot of nitrate (NO_3^-) along the transect. Concentrations are shown as data interpolating variational analysis gridded fields generated in Ocean Data View (Schlitzer 2017). Station numbers are the same as those used by the Line P program. The box in (A) shows the stations that were in the transition zone during our sampling period

closed on deck, are opened by hydrostatic pressure at approximately 10 m depth, and are closed by remote command. Samples were collected within the mixed layer between 10 and 30 m depth. After the rosette was brought on deck, the GO-FLO bottles were detached and sampled inside a Class-100 clean van kept under positive pressure with HEPA-filtered air. Total dissolved metal samples were collected in acid-cleaned, low-density polyethylene bottles after in-line filtration through a 0.2 μm pore size Acropak 200 filter using gas pressurization of the Niskin bottles with ultrapure N_2 gas. Samples for dissolved organic Fe-binding ligands were similarly filtered and collected in acid-cleaned, fluorinated low-density polyethylene bottles and frozen at −20°C until analysis.

2.2. Dissolved total metal and organic speciation analyses

Total dissolved Fe, Zn, and Cu concentrations were determined at the University of Southern California (USC) using a single batch nitrilotriacetatic acid (NTA) resin extraction and isotope dilution inductively coupled plasma mass spectrometry (ICP-MS) method adapted from Lee et al. (2011). They were analyzed in triplicate using the Finnegan Element 2 (Thermo Scientific) ICP-MS in medium resolution mode equipped with an Apex desolvation system.

Dissolved metals were pre-concentrated in the samples by adding a chelating resin (NTA Superflow resin; Qiagen) in the preparatory stage. The NTA resin was cleaned using the following procedure (Lee et al. 2011): 25 ml of the NTA resin solution were poured into a clean 50 ml polypropylene centrifuge tube (Corning) and washed 5 times with 18.2 M (Milli-Q) water. Between washes, the tube was spun down in a 5810-R centrifuge (Eppendorf) maintained at 8°C for 10 min at $3320 \times g$. After decanting the supernatant, Milli-Q water was added for the next wash. The resin was then washed 5 times with 1.5 M trace metal grade hydrochloric acid (HCl; Optima, Fisher) and several more times with Milli-Q water after that to bring the pH of the solution above 4, indicating that all of the HCl had been removed from the solution. For the final cleaning step, the resin solution was washed 5 times with 0.5 M trace metal grade nitric acid (HNO_3; Optima, Fisher). The resin solution was placed on an analog shaker for several hours for the first wash, and then left overnight on the shaker for the last wash. After the final wash, the resin solution was again washed at least 5 times with Milli-Q water

until the pH had risen above 4 in order to remove all of the HNO_3. The resin solution was diluted 2-fold with 25 ml Milli-Q water and stored in the refrigerator for future use. A 25 μl working volume of the resin suspension contained ~100–400 beads, representing a 1:50 dilution of the primary resin solution.

Samples were pre-concentrated for analysis in 15 ml polypropylene centrifuge tubes, which were first cleaned in a 2-step process by soaking them in 10% HCl at 60°C for 48 h and then rinsing each tube at least 5 times with Milli-Q water. After the rinses, the tubes were filled to a positive meniscus with 0.5% trace metal grade HCl, capped and baked at 60°C overnight. After retrieving them from the oven, the tubes were left capped and stored until used. Before sample addition, the tubes were emptied and rinsed 3 times with Milli-Q water and at least once with the sample.

The centrifuge tubes were filled with ~7.5 ml of sample (with the exact volume determined gravimetrically) and spiked with enough [57]Fe-enriched spike and [67]Zn enriched spike (BDH Aristar Plus) to bring the final concentration to ~2 nM. No spike was added for Cu. Subsequently, 0.1 ml of 1.5 M trace metal grade hydrogen peroxide (Optima, Fisher) was added to each sample and left to equilibrate for at least 1 h at room temperature, to completely oxidize any Fe(II) to Fe(III) (Lee et al. 2011). Next, 200 μl (~800 beads) of the working resin suspension were added to each sample, and the tubes were placed on a shaker for 2 to 3 d. The samples were centrifuged for 10 min at $3320 \times g$, and the seawater was siphoned off to leave the resin beads. The beads were washed twice with 3 ml Milli-Q water to remove salts, and the tubes were once again centrifuged using the same settings. After the final wash, 1 ml of 5% trace metal grade HNO_3 was added to each tube and, after leaving them on the shaker again for 1 d, the samples were ready for analysis.

Procedural seawater blanks were prepared in triplicate in the same way as samples using ~0.2 ml low trace metal surface seawater from the 2004 Sampling and Analysis of Iron (SAFe) cruise (Fe = 0.09 ± 0.007 nM l^{-1} SD). The average detection limit and internal blank value for this method (n = 3, 1σ) for Fe were 0.01 and 0.06 nM, respectively. The accuracy of the method was evaluated by measuring SAFe reference standards S1 and D1 (Johnson et al. 2007). The Fe values we obtained by this method for S1 and D1 were 0.09 ± 0.01 and 0.65 ± 0.02 nM, respectively. The latest consensus values are 0.09 ± 0.01 nM Fe (S1) and 0.67 ± 0.04 nM Fe (D1) (www.geotraces.org/science/intercalibration).

Dissolved organic Fe-binding ligands were determined using competitive ligand exchange adsorptive cathodic stripping voltammetry (CLE-AdCSV) as described in detail elsewhere (Bundy et al. 2014). In brief, organic speciation samples were slowly thawed and shaken vigorously before analysis. Samples were then aliquoted into 12 separate 10 ml acid-cleaned and conditioned Teflon vials (Savillex), and dissolved Fe was added in a range of 0–15 nM along with 50 µl of a 1.5 M ammonium borate buffer (pH 8.2, NBS scale) to each aliquot. The dissolved Fe additions and buffer were left to equilibrate with the natural ligands for 2 h, after which 25 µM of the well-characterized competing ligand salicylaldoxime was added to each sample aliquot. This was equilibrated for 15 min, and then each aliquot was analyzed using AdCSV on a controlled growth mercury electrode (BASi). Peak heights generated during the cathodic stripping step were determined using ECDSoft using a curved baseline (Omanovic et al. 2015), and titration data were analyzed in ProMCC using an optimized sensitivity (Hogle et al. 2016). Multiple ligand classes were determined in these samples using several titrations at multiple analytical windows (Bundy et al. 2014), but only the data for the strongest ligand class, L_1, are presented here. This ligand class has conditional stability constants, or Fe-binding strengths, akin to siderophores, and likely represents microbially produced metabolites with high binding affinities for dissolved Fe, or other strong Fe-binding ligands (Gledhill & Buck 2012).

2.3. DNA extraction and 18S amplicon sequencing

DNA filters were extracted using the Qiagen® DNeasy Plant DNA extraction kit with an additional bead-beating step and homogenation using the QIA-Shredder column (Qiagen). The V4 region of the 18S rDNA was amplified in triplicate using primers designed and evaluated as diatom barcoding primers (Zimmermann et al. 2011) that were modified to include the Illumina overhang adapter sequences for 2 step amplicon sequencing (www.illumina. com/content/dam/illumina-support/documents/ documentation/chemistry_documentation/16s/16s-

metagenomic-library-prep-guide-15044223-b.pdf; Table 1). Triplicate 25 µl polymerase chain reaction (PCR) reactions were run separately for each sample and then pooled prior to purification. Each 25 µl PCR reaction consisted of 7.5 ng DNA and a final concentration of 0.05 µM of each primer and 1× BIO-X-ACT™ Short Mix (Bioline). PCR reactions were performed using a MasterCycler© thermal cycler (Eppendorf). PCR conditions consisted of 95°C for 10 min, then 35 cycles of 95°C for 60 s, 61.8°C for 60 s, 72°C for 30 s; then a final 72°C for 10 min. We used 5 µl of each pooled PCR reaction to confirm that the PCR was successful via agarose gel electrophoresis. The remaining 70 µl were PCR purified using the QIAquick PCR Purification Kit (Qiagen). Purified PCR products were diluted to 10 ng µl^{-1} and sent to the Rhode Island Genomics & Sequencing Center at the University of Rhode Island for the Index PCR of the amplicon sequencing protocol (www.illumina. com/content/dam/illumina-support/documents/ documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf), which includes a short PCR to add Nextera XT Indices, and sequencing on the Illumina MiSeq sequencer using the 2× 300 bp sequencing kit. DNA extracted from 2 stations (P5 and P8) were sequenced in duplicate.

2.4. Sequence analysis

Diatom 18S rDNA amplicon sequences were uploaded to the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under SRA accession number SRP144011. Sequences were de-multiplexed and imported into the CLC Genomics Workbench (Qiagen). Reads were imported in pairs, trimmed, and merged in CLC before being exported in FASTA format. Parameters for trimming were as follows: no ambiguous nucleotides allowed, quality limit 0.02, and discard reads <150 nucleotides in length post trimming. Read merging parameters were as follows: mismatch cost = 1, minimum score = 8, gap cost = 2, and maximum unaligned end matches = 5. Following the merge, reads were searched for the primer sequences, and any

Table 1. Diatom 18S primers with MiSeq adapters

Primer	Sequence
MiSeqD512	TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG ATT CCA GCT CCA ATA GCG
MiSeqD978	GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GGA CTA CGA TGG TAT CTA ATC

merged reads without the primers were discarded. The remaining merged reads had the primers trimmed off. A final trimming step had the following parameters: no ambiguous nucleotides, quality limit = 0.05, reads were discarded if they were not within the range of 390–415 nucleotides in length. Reads were exported out of CLC, concatenated into 1 FASTA file, and community composition was determined using the minimum entropy decomposition (MED) pipeline (Eren et al. 2015). The MED parameters that were altered from the default parameters were as follows: minimum-substantive-abundance = 10 and outliers were relocated. MED grouped the sequences by similarity into MED nodes (operational taxonomic units) using an entropy minimization principle, generating a table with sequence counts per node, per sample, as well as a FASTA file with representative sequences for each node. The representative MED node FASTA file was used as an input for a nucleotide BLAST (Altschul et al. 1990) against an in-house database that combined stramenopile 18S sequences from NCBI (downloaded as of 20 August 2018) and the SILVA eukaryote 18S database. Parameters for BLAST were as follows, program = blastn, no. hits per input read = 1, e-value = 1E − 100.

While the primers used were designed and evaluated as diatom barcoding primers (Zimmermann et al. 2011), because the 18S gene is so well conserved, some non-diatom sequences were amplified. As the ability of these primers to quantitatively amplify non-diatom species has not been evaluated, non-diatom MED nodes and their count data were filtered out after BLAST was performed. This is similar to the practice of removing 18S rRNA sequences or 16S sequences from *Archaea*, chloroplasts, and mitochondria, which is common practice in analysis of 16S bacterial datasets (Schloss et al. 2009). Stn P2 had the most non-diatom sequences recovered. Almost all non-diatom sequences that were removed were also from the stramenopiles/heterokont division of the eukaryotes. The most abundant non-diatom sequence retrieved was 90% similar to the diatom parasite *Pirsonia verrucosa*, which accounted for 32% of the reads at P2, 15% of the reads at P6, and 5% of the reads at P5. Sequences similar to *Bolidomonas*, *Hyphochytrium*, and *Paraphysomonas* were also occasionally recovered. *Hyphochytrium* accounted for 7% of sequences at Stn P2, otherwise none of the non-diatom sequences accounted for >5% of sequences at any station. Following non-diatom sequence removal, sample reads were normalized down to 40 000 using the rrarify program of 'vegan' (Oksanen et al. 2017). Because of trimming parameters, some outputs re-

sulted in sequences from the same species being a base or 2 shorter in length. Even though the sequences were identical over the lengths that they co-existed, they ended up getting pooled into different MED nodes. For graphing purposes, the count data for nodes were pooled if the nodes had <1% identity difference and nodes were classified to the species level if they had >99% identity to a classified organism. Nodes were classified as being species-like (e.g. *Thalassiosira pacifica*-like) if they had 98–99% identity. Sequences with percent ID <98% were only classified to the genus level and given a number in cases where there were multiple different unidentified species in a given genus (e.g. *Thalassiosira* sp. 1).

2.5. Statistical analyses

All statistical analyses were done on original non-grouped MED node count data using PRIMER v.6 (PRIMER-E) (Clarke & Gorley 2006) or in GraphPad Prism version 7 for Macintosh. The Shannon-Wiener Index of diversity (Shannon 1948) was computed on the relative abundance data for each station. A similarity matrix of Bray-Curtis (Bray & Curtis 1957) coefficients was used to compare the relative abundance data and to establish a cluster dendrogram using the group average mode of clustering. To determine the single environmental variable or group of environmental variables (temperature, salinity, density, bottom depth, mixed layer depth, depth of the chlorophyll maximum, photosynthetically active radiation [PAR], degree of longitude, fluorescence, nitrate, nitrite, phosphate, silicate, dissolved metals [dCu, dFe, dZn], concentration of L_1 Fe-binding ligands, and nitrate:dFe ratio) that best explained the similarity distribution of these data, the environmental data and Bray-Curtis similarity matrices were analyzed using the BEST analysis with the BIOENV algorithm (Clarke & Ainsworth 1993, Clarke et al. 2008) using the Spearman rank correlation method and D1 Euclidean distance as the resemblance measure and 999 random permutations to assess the significance of the sample statistics. Additionally, we ran a non-parametric (Spearman) correlation on all of the environmental variables against each other, and report the Spearman correlations (Spearman's ρ) and the 2-tailed p-values of correlated variables.

2.6. RNA extraction and gene expression analysis

RNA extraction and expression of the Fe-responsive *T. oceanica* flavodoxin gene (FLDA1), the gene

encoding Fe stress-induced protein 3 (ISIP3), and the stably expressed β-actin gene (ACT) was performed using the methods described by Chappell et al. (2015). Briefly, RNA extraction was done using the Qiagen® RNeasy Plant RNA extraction kit using an additional bead beating step. RNA was converted to cDNA using the Applied Biosystems High-Capacity cDNA Reverse Transcription kit (Life Technologies) using an oligo dT primer. Quantitative reverse transcription PCR (qRT-PCR) for FLDA1, ISIP3, and the housekeeping gene ACT was done with triplicate technical replication on a Light Cycler 480 Real-Time PCR System (Roche Applied Science) using the primers, plasmids, kit, and conditions described by Chappell et al. (2015).

3. RESULTS

3.1. Hydrographic and trace element conditions along Line P

Near-surface environmental data used in the BIOENV analysis are shown in Table 2 and Fig. 1. While only near surface values were used in analysis, a section plot of nitrate in the upper 100 m is also shown to provide context on upper water column structure (Fig. 1D). In order to avoid contamination from the ship by collecting near surface waters, trace metal samples were not collected above 15 m, except

at the final station, P1, where a sample was collected at 10 m because of a shallow mixed layer. Thus, trace element samples were always collected within the mixed layer but were not collected from the same depths as the other variables that were collected at 5 m. Macronutrient data (nitrate, silicate, and phosphate) tracked density structure and remained constant throughout the mixed layer at all stations. Along the transect, temperature in surface waters ranged from 10.0–10.6°C and salinity ranged from 30.8–31.9. Density (sigma-theta) ranged from 23.61–24.45 kg m^{-3} – 1000. Bottom depth of the sampling sites ranged from 120–2450 m. The 2 highest surface fluorescence measurements were at P1 (6.70 mg m^{-3}) and P8 (4.34 mg m^{-3}) and the lowest fluorescence was measured at P3 (1.76 mg m^{-3}). Silicate ranged from 1.14–6.59 μM, and both phosphate and nitrate did not vary substantially across the transect, with ranges of 0.10–0.36 μM for phosphate and 0.04–0.55 μM for nitrate. Nitrite and ammonium were also very low in surface samples, as would be expected. Nitrite was only above detection at 3 stations, and the maximum surface concentration was 0.02 μM. Ammonium was below detection at P2 and otherwise ranged from 0.05–0.09 μM. On the whole, dissolved trace metal concentrations were highest inshore at P1, decreased rapidly between P1 and P2, and stayed low all the way to P8. Dissolved Cu ranged from 1.02–2.76 nM with its lowest concentration measured at Stn P3. Dissolved Zn ranged from 0.04–0.76 nM

Table 2. Environmental data used in BIO-ENV analysis. PAR: photosynthetically active radiation, dissolved iron (dFE), zinc (dZn) and copper (dCu), FeL$_1$: Fe-binding L$_1$ ligand, Fe': inorganically bound Fe

| | Station | | | | | | |
	P1	P2	P3	P4	P5	P6	P8
Longitude (°W)	125.500	125.000	126.333	126.667	127.168	127.667	128.667
Bottom depth (m)	120	114	750	1300	2100	2500	2450
Temperature (°C)	10.6	10.4	10.0	10.6	10.0	10.0	10.5
PAR (μE m^{-2} s^{-1})	12.5	48.4	35.5	0.0	70.4	34.2	0.00
Salinity	30.8	31.3	31.8	31.6	31.7	31.5	31.9
Density (kg m^{-3} – 1000)	23.6	24.0	24.5	24.2	24.3	24.2	24.4
Fluorescence (mg m^{-3})	6.7	2.8	1.8	3.2	2.1	3	4.3
Silicate (μM)	2.80	3.49	1.14	4.64	6.59	5.33	3.85
Phosphate (μM)	0.20	0.21	0.10	0.31	0.36	0.36	0.32
Nitrate (μM)	0.33	0.04	0.06	0.13	0.45	0.55	0.18
Nitrite (μM)	0.01	0.00	0.00	0.00	0.01	0.02	0.00
Ammonium (μM)	0.09	0.00	0.08	0.07	0.08	0.05	0.05
Depth metal sample (m)	10	22	28	20	30	20	20
dFe (nM)	1.28	0.55	0.36	0.64	0.44	0.46	0.27
FeL$_1$ (nM)	3.09	2.31	1.89	1.28	1.66	2.66	0.88
Fe' (pM)	0.27	0.30	0.07	0.28	0.28	0.12	0.03
dZn (nM)	0.76	0.06	0.15	0.14	0.12	0.04	0.22
dCu (nM)	2.76	1.74	1.02	1.13	1.39	1.35	1.39

with its lowest measured concentration at Stn P2, while dFe continued to decrease throughout the transect ranging from 1.28 nM in the coastal stations to 0.27 nM offshore. Dissolved Fe-binding L_1 ligands (FeL$_1$) ranged from 0.88–3.09 nM, and were always in excess of dFe. Inorganically bound Fe (Fe'), calculated from the L_1 concentrations, conditional stability constants, and the total dFe concentrations, ranged from 0.03–0.28 pM.

3.2. Diatom diversity

The number of different sequences found at a given station (S), the total number of sequences analyzed per station (N), the evenness (J'), Shannon diversity index (H'), and Gini-Simpson index ($1 - \lambda$) of sequence data at each station are given in Table 3. Stn P5 was the most even (highest J') and was also the highest in both diversity measures (H' and $1 - \lambda$). Stns P2 and P6 were both similarly even and diverse with values for J', H', and $1 - \lambda$ slightly less than those for Stn P2. Stns P8 and P1 were the least even and diverse with the lowest values for J', H', and $1 - \lambda$.

Bray-Curtis similarity of the relative abundance of diatom 18S sequences groups the stations into 4 or 5 'clusters' depending on the % similarity used as a cutoff (Fig. 2A). The most coastal station, P1, and the most oceanic station, P8, were the least similar to one another and to all the other stations. Stns P3 and P4 grouped together and were >70% similar. Stns P2, P5, and P6 grouped together into 1 larger cluster where the diatom community composition was >60% similar between all 3 stations. The diatom community composition at Stns P5 and P6 were the most similar

Table 3. Diversity measures for diatom 18S sequence analysis. Stations P5 and P8 were sampled in duplicate. For each station, the total number of sequences analyzed (N) was 40 000. S: number of different sequences found at a given station, J': evenness, H': Shannon diversity index, $1 - \lambda$: Gini-Simpson index

Station	S	J'	H'	$1 - \lambda$
P1	1179	0.63	4.43	0.91
P2	1530	0.68	5.01	0.96
P3	1347	0.66	4.78	0.95
P4	1696	0.65	4.82	0.95
P5A	1776	0.70	5.20	0.98
P5B	1759	0.70	5.24	0.98
P6	1715	0.70	5.13	0.97
P8A	1172	0.60	4.24	0.90
P8B	1163	0.60	4.26	0.90

at >75% similarity. The duplicate samples, P5 (A/B) and P8 (A/B), were >90% similar to one another.

The relative abundance of sequences associated with different diatom genera (Fig. 2B) and different diatom species (Fig. 2C) at each station are shown as columns, which are displayed in the order of the Bray-Curtis similarity cluster dendrogram. The most abundant diatom genus based on sequencing corresponded to *Guinardia*. The second and third most abundant diatom species based on sequencing, *Guinardia delicatula* and *Rhizosolenia delicatula*, respectively, were counted as part of this genus. *R. delicatula* has previously been classified as a synonymous species to *G. delicatula* (Bisby et al. 2011), and was proposed to be included in the genus *Guinardia* by Round et al. (1990). We should note that the *G. delicatula* and *R. delicatula* 18S sequences are not identical. In fact, they are only 97% similar, thus we discuss them as separate species. However, as the *R. delicatula* 18S sequence is only 92% similar to the most similar other *Rhizosolenia* genus sequences that were recovered, we support grouping *R. delicatula* within the *Guinardia* genus. Sequences corresponding to *G. delicatula* were the most abundant sequences at Stn P4, and were among the top 3 sequences recovered at all stations, except P6 and P8. Sequences corresponding to *R. delicatula* accounted for 48% of the sequences recovered at P8. The second most abundant diatom genus based on sequencing was *Thalassiosira*, which dominated the sequencing reads at Stns P4 and P5, and ranked in the top 3 most abundant genera recovered based on sequencing at all stations except P8. This was driven mainly by *T. pacifica*, which was among the top 5 most abundant species based on sequencing at Stns P1, P3, and P4. Additionally, at all stations except P8, at least 2 *Thalassiosira* species were in the top 10 most abundant species based on sequencing. The most abundant diatom species based on sequencing overall can only be classified to the genus level, as it was only 93% similar to *Leptocylindrus danicus*. This sequence was the most abundant sequence recovered at Stns P3 and P6, and ranked in the top 3 most abundant sequences recovered at Stns P1, P2, P4, and P5, which resulted in *Leptocylindrus* being the third most abundant diatom genus based on sequencing. The fourth most abundant diatom genus based on sequencing was *Pseudonitzschia*. Stn P1 was dominated by sequences that corresponded to *P. pungens*, which accounted for 48% of the diatom 18S sequences recovered at that station. At Stn P8, the third most abundant diatom species based on sequencing was *P. granii*. Sequences from the top 4 genera accounted for at least 55% of the
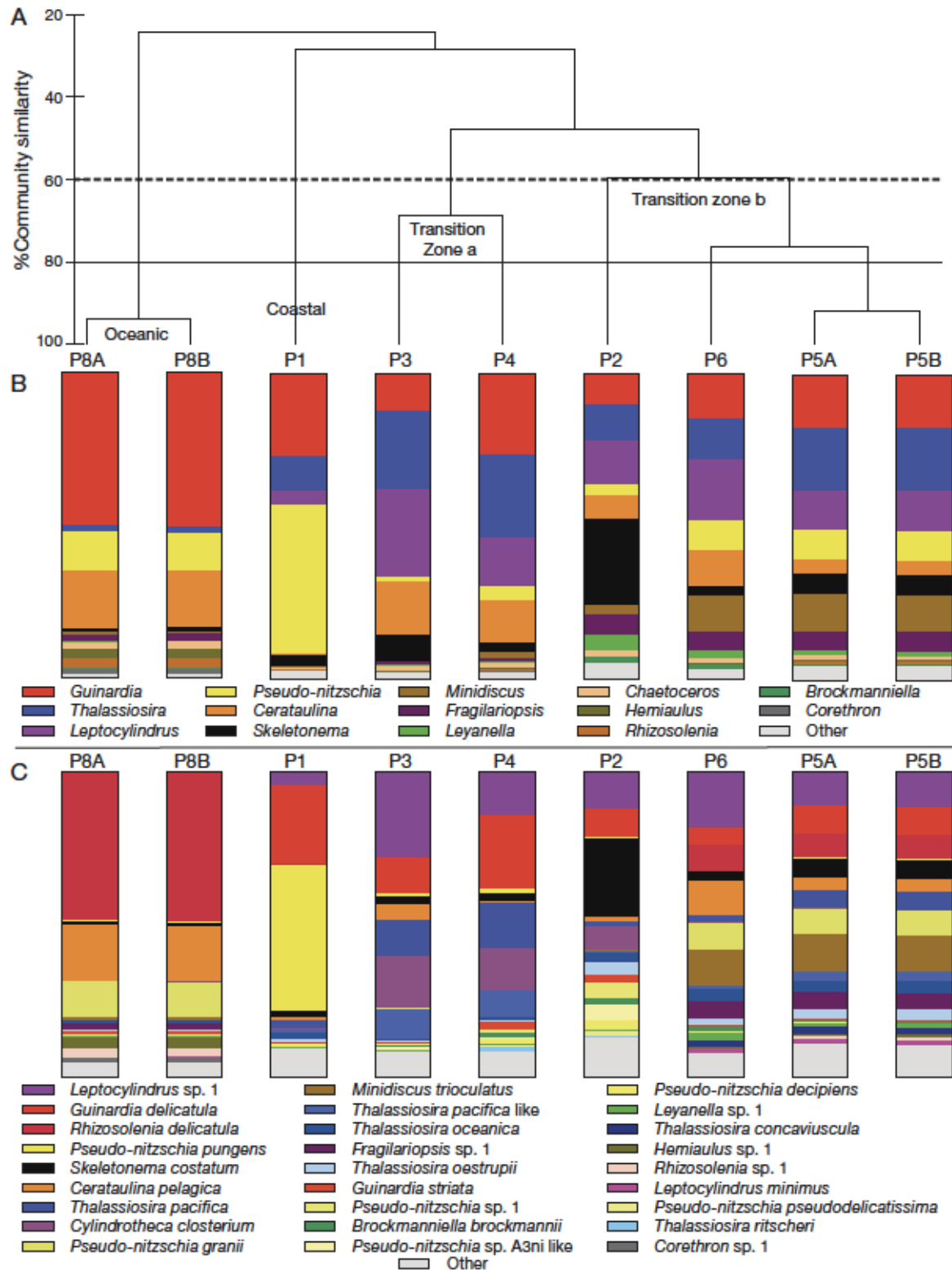
Fig. 2. (A) Bray-Curtis similarity of the diatom 18S amplicon sequencing results from the samples collected from 5 m depth at 7 stations along Line P. Samples from 2 stations (P5 and P8) were sequenced in duplicate. The solid line is at 80% similarity and the dashed line is at 60% similarity. Additionally, labels associated with the cluster designations referred to in the text have been added. (B,C) Relative abundances, showing the fraction of sequences at each station attributable to a given genus (B) or species (C) of diatom. Only genera and species that accounted for >1% of sequences at one of the stations are identified. The remaining genera and species are grouped together in the 'other' category. Genera and species are ordered from largest to smallest in total read abundance among all the stations

diatom 18S sequences at any given station, except P2 (where they only accounted for 40% of the sequencing reads). The next 3 most abundant diatom genera based on sequencing were *Cerataulina*, *Skeletonema*, and *Minidiscus*. Stn P2 was dominated by sequences that corresponded to *S. costatum*. Sequences that corresponded to *M. tricolatus* were the most abundant sequences recovered from Stn P5, and were the second most abundant sequences recovered from Stn P6. While not the most abundant diatom species recovered at any single station based on sequencing, sequences that corresponded to *C. pelagica* were in the top 3 most abundant sequences recovered at Stns P6 and P8.

The shifts along the transect between the dominant species of the genus *Guinardia* and the dominant species of the genus *Pseudo-nitzschia* are highlighted in Fig. 3. It is important to note that these graphs are plotted on a log scale, as the read counts for individual sequences represented a range of over 5 orders of magnitude. The dominant *Guinardia* species shifted from *G. delicatula* to *R. delicatula* at Stn P5. The dominant *Pseudo-nitzschia* species shifted from *P. pungens* at P1 to *P. granii* at P5.
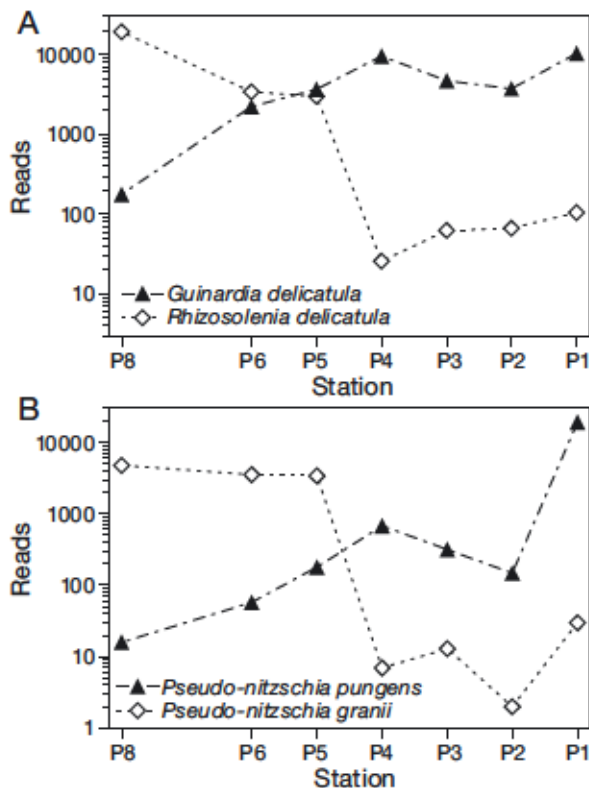

Fig. 3. Number of sequences per station attributed to (A) *Guinardia delicatula* and *Rhizosolenia delicatula* and (B) *Pseudo-nitzschia pungens* and *P. granii*. Note that both (A) and (B) are plotted on a log scale

3.3. Statistical analyses of correlations within biological and chemical datasets

A few of the environmental variables significantly (p < 0.05) correlated with one another, including phosphate with silicate ($\rho = 0.95$, $p \leq 0.004$), nitrate with nitrite ($\rho = 0.91$, $p \leq 0.01$), fluorescence with temperature ($\rho = 0.86$, $p \leq 0.02$), fluorescence with PAR ($\rho = -0.77$, $p \leq 0.05$), salinity with dFe ($\rho = -0.89$, $p \leq 0.01$), salinity with density ($\rho = 0.93$, $p \leq 0.01$), density with dFe ($\rho = -0.93$, $p \leq 0.01$), and salinity with FeL_1 ligands ($\rho = -0.81$, $p \leq 0.03$). The BIOENV analysis found no significant correlations between diatom community composition and a single environmental factor; however, combining multiple factors did result in significant (p < 0.05) correlations based on permutation tests. Due to the small number of stations involved in the analysis and the large number of environmental variables considered, only correlations with $\rho > 0.825$ could be considered significant (p < 0.05) based on permutation tests. Multiple combinations of 4 environmental variables yielded a correlation of between $\rho = 0.829$ and 0.834 (Table 4). This is almost certainly the result of correlation between the environmental variables that were part of these groupings. Longitude and dCu were always factors included in the significant correlations. The other 2 factors involved in the correlation included a combination of dFe, dZn, salinity, and density. As stated above, salinity and density are positively correlated with one another and negatively correlated with dFe.

3.4. Indicators of diatom Fe stress

To determine whether members of the diatoms were experiencing Fe limitation, the relative transcript abundance of FLDA1 relative to ACT, the stably expressed housekeeping gene, for *T. oceanica* was compared and plotted on a log scale at stations where *T. oceanica* ACT transcripts were detectable

Table 4. BIO-ENV biota and environmental Spearman correlations (ρ). dCu (dFe, dZn): dissolved copper (iron, zinc)

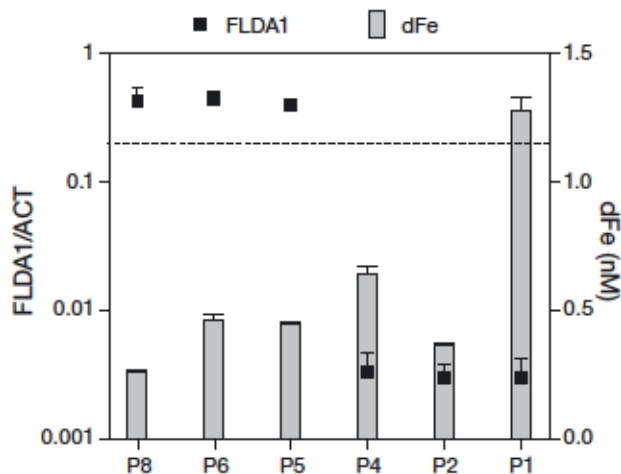| Spearman ρ	Environmental variables included in correlation			
0.834	Longitude	dCu	dFe	Salinity
0.834	Longitude	dCu	dFe	Density
0.832	Longitude	dCu	dZn	Salinity
0.830	Longitude	dCu	dZn	Density
0.829	Longitude	dCu	dFe	dZn

Fig. 4. *Thalassiosira oceanica* flavodoxin (FLDA1) transcript abundance relative to the β-actin gene (ACT) transcript abundance for the 6 stations where *T. oceanica* was present in measurable quantities (note the log scale). The dashed line is the expression level above which *T. oceanica* is growth limited by low dissolved Fe (dFe) according to Chappell et al. (2015). Also plotted are the dFe concentrations associated with these samples. Error bars represent the standard error for technical triplicates

(Fig. 4). FLDA1 relative transcript abundance ranged from 0.003 at the more coastal stations (P1, P2, and P4) to 0.39–0.45 at the more oceanic stations (P5, P6, and P8). Dissolved Fe at the corresponding stations is also plotted (Fig. 4). ISIP3 transcripts were only above detection limit at Stns P6 and P8. The values for ISIP3 expression relative to ACT at those 2 stations were 0.07 ± 0.005 and 0.06 ± 0.004, respectively. Note that ISIP3 expression could not be assessed at P5 because of a sample processing issue.

4. DISCUSSION

Our study area focused on the first 8 stations (P1–P8) of the Line P transect, as recent work in this region suggests that the transition zone between the coastal and HNLC regions is an area of elevated productivity (Peña & Varela 2007, Ribalet et al. 2010). Previous researchers have identified different boundaries for the 3 zones (coastal, transition, and oceanic–HNLC) along the Line P, transect depending on whether the study focused on macro- or micro-nutrient concentrations or on physical parameters as the main distinctions between the 3 zones. Whitney et al. (1998) defined the coastal–HNLC transition zone as between Stns P4 and P12, where nitrate levels remain low throughout the summer. In the late summer and early fall, coastal upwelling results in high levels of nitrate at Stns P1 and P2 (Peña & Varela 2007), commonly dropping to their lowest along the transect by Stns P3 and P4, and then begin to rise as the transect moves westward. In early fall, nitrate values close to detection limits have been observed at the surface as far west as P12 (Varela & Harrison 1999, Peña & Varela 2007). During our cruise, surface macronutrients (nitrate, phosphate, and silicate) were relatively low throughout the transect, thus it is difficult to assess the boundaries of the transition zone using the parameters of Whitney et al. (1998). A smaller coastal–HNLC transition zone area from P2 to P6 was defined based on the area of steep surface salinity gradients moving offshore (Ribalet et al. 2010). Based on the definition of the coastal–HNLC transition zone according to Ribalet et al. (2010), the transition zone for our cruise would only include Stns P2 and P3, which does not seem appropriate based on the rest of our results. While we cannot use the extent of elevated coastal nitrate to define the eastern boundary of the coastal–HNLC transition zone, we measured the lowest surface salinity (30.8) at Stn P1, which supports previous work classifying P1 as the only truly coastal station of the transect (Ribalet et al. 2010). Additionally, while nitrate concentrations were higher at Stns P5 and P6 than at the other sampled stations, they still remained below 1 µM and dropped down to 0.18 µM at Stn P8. The May–June 2012 Line P transect cruise (www.waterproperties. ca/linep/cruises.php), which occurred just after our cruise, measured elevated nitrate (1.3 µM) at Stn P8, with even higher concentrations (8.5 µM) beginning at Stn P10. While we refer to Stn P8 as 'oceanic' because it is our most oceanic station, P8 nitrate concentrations are an order of magnitude lower than values typically observed in HNLC waters in this region (Harrison 2002 and references therein), thus technically it is a transition zone station based on the criteria of Whitney et al. (1998).

Ribalet et al. (2010) highlighted the transition zone as a phytoplankton 'hotspot' with high productivity and increased diversity of species in the diatom genus *Pseudo-nitzschia*, hypothesizing that this 'hotspot' occurs as a result of coastal water containing higher Fe concentrations mixing periodically with water masses containing elevated nitrate in the transition zone. Other work has found that high productivity is occasionally observed at Stn P4, with low and invariable chl a concentrations west of Stn P4 (Peña & Varela 2007). We observed elevated fluorescence throughout our transect. Based on the Pacific Fisheries Environmental Laboratory upwelling indices

from the 3 monitoring sites closest to our sampling transect (48° N, 125° W; 51° N, 131° W; 45° N, 125° W; https://www.pfeg.noaa.gov/products/PFEL/modeled/indices/PFELindices.html), our sampling did not occur during a period of upwelling-favorable winds; however, a 10 d period of upwelling-favorable winds did precede our cruise. The previous period of upwelling likely resulted in the elevated surface chlorophyll fluorescence we measured along the transect out to Stn P8. The May–June 2012 Line P transect cruise, which sampled the whole Line P transect starting at Stn P1 3 d after we sampled there, recorded even higher surface chlorophyll (22 mg m^{-3}) at Stn P1, although near-surface chlorophyll remained low (<1 mg m^{-3}) throughout the remainder of the transect, except at Stn P7 where it was also slightly elevated (1.8 mg m^{-3}). While there was limited satellite chlorophyll data from the region because of cloud cover during our cruise, the available data agree with both sets of shipboard results. The 3 d composite centered on 15 May 2012 shows elevated chlorophyll offshore, while the 3 d composite centered on 23 May 2012 only shows an elevated chlorophyll signal around the location of Stn P1 (Fig. A1 in the Appendix). Our results support the hypothesis that the transition zone is an area with periodic increases in phytoplankton productivity. Despite seasonal variations in the boundary of the coastal–HNLC transition zone, we will discuss the diatom diversity data in the context of both previous studies and our results that have found elevated productivity associated with increased diatom diversity in the coastal–HNLC transition zone of the Line P transect.

4.1. Patterns in diatom diversity

Bray-Curtis similarity analysis of the diatom community identifies 4 different station groupings. For simplicity, we have classified these 4 groups of stations as: coastal (P1), transition zone A (P3 and P4), transition zone B (P2, P5, and P6), and oceanic (P8), while recognizing that our most 'oceanic' station is not representative of the oceanic HNLC typically found in the western portion of the Line P transect. Sequences attributed to the top 4 genera, *Guinardia*, *Thalassiosira*, *Leptocylindrus*, and *Pseudo-nitzschia*, accounted for 65% of the sequences retrieved.

As the MED analysis allows us to differentiate between 18S sequences with <1% differences, and most species of diatoms have more than a 1% difference in sequence in the region that our primers amplify, we were able to examine the relative abun-

dance of sequences from individual diatom species, and observe distinct shifts in the dominant species of diatom genera along the transect. Two diatom species, *P. pungens* and *G. delicatula*, which (combined) accounted for 75% of the reads at the most coastal station, were not well represented in the sequences returned from the most oceanic station sampled. Instead, sequences associated with 2 different *Pseudo-nitzschia* and *Guinardia* species were abundant at our most oceanic station. The abundance of *P. pungens* sequences dropped rapidly as we moved away from the coast, which is in agreement with previous studies indicating that *P. pungens* is commonly found in coastal embayments on the west coast of North America, and does not extend significantly beyond the coastal zone (Stehr et al. 2002), although it is cosmopolitan in its distribution in coastal regions globally (Hasle 1995, Casteleyn et al. 2008). West of Stn P4, the most abundant *Pseudo-nitzschia* species was *P. granii*, which is an oceanic species that is common in the coastal–HNLC transition zone (Ribalet et al. 2010) and is an early responder to pulses of Fe (El-Sabaawi & Harrison 2006). *G. delicatula* dropped in relative abundance as we moved offshore, though not as rapidly as *P. pungens*. West of Stn P5, *Rhizosolenia delicatula* sequences were more abundant than *G. delicatula* sequences. While these 2 species have 18S sequences that are 3% different over the region amplified, supporting the fact they are distinct species, *R. delicatula* has been classified as a synonymous species to *G. delicatula* (Bisby et al. 2011). Because of this classification issue, it is difficult to determine from the literature what is known about the distribution of these 2 species, but *R. delicatula* has been associated with coastal (Huang et al. 2004) and upwelling (Blasco et al. 1980) environments.

Based on the distribution of individual diatom species, our diatom community composition data support previous hypotheses that the coastal–HNLC transition zone is an environment where both coastal and oceanic species of phytoplankton can be found. Sequences 93% similar to *L. danicus*, a diatom species commonly associated with coastal, estuarine, and upwelling environments (Davis et al. 1980, Marshall & Cohn 1983, French & Hargraves 1986, Trigueros & Orive 2000, Huang et al. 2004, Aizawa et al. 2005), were the most abundant sequences recovered overall, and abundant at all stations except P8. With such a low sequence identity, even classifying this sequence to the genus level is putative. However, a high recovery of *Leptocylindrus* sequences is consistent with prior research suggesting they are an abundant coastal species. *Skeletonema costatum* se-

quences, another common estuarine and shelf diatom species (Marshall & Cohn 1983, Huang et al. 2004), and sequences matching *Cerataulina pelagica*, a common bloom-forming diatom species in coastal and estuarine environments (Carstensen et al. 2015), were both abundant across the transect. A sequence matching *Thalassiosira pacifica*, a species known to be associated with near-coastal and upwelling environments in the North Pacific Ocean (Aizawa et al. 2005), was abundant at all stations except P8. Sequences attributable to *Minidiscus tricolatus* were abundant in samples from transition zone B stations (P5 and P6). Previous work has shown that *Minidiscus* spp. are very abundant in this region (Aizawa et al. 2005, Chappell et al. 2013) and, like *T. pacifica*, appear to be most closely associated with coastal and upwelling environments (Aizawa et al. 2005, Buck et al. 2008). Sequences associated with 2 additional *Thalassiosira* species, *T. oestrupii*, known to be abundant in the non-coastal regions of the North Pacific Ocean (Aizawa et al. 2005) and *T. oceanica*, which has previously been shown to be an abundant diatom species both coastally and in open ocean samples from the region (Aizawa et al. 2005, Chappell et al. 2013), were also abundant along the transect.

West of Stn P4, an abundant sequence recovered was 99% similar to *Fragilariopsis kerguelensis, F. curta* or *F. sublineata*. It is important to note that in the region amplified by our primers, *F. kerguelensis, F. curta*, and *F. sublineata* are all identical in sequence. All 3 of these species of *Fragilariopsis* are found in the Southern Ocean and have not previously been reported in the North Pacific or Arctic Oceans. There are 2 *Fragilariopsis* species, *F. oceanica* and *F. pseudonana*, that are abundant in the non-coastal regions of the North Pacific and Arctic Oceans (Aizawa et al. 2005). 18S sequence data for the region covered by our primers is not available for either of these 2 species, and thus we cannot differentiate what *Fragilariopsis* species we recovered, only that there was a *Fragilariopsis* sp. sequence that was abundant along the transect. Our results fit with what has previously been found in the North Pacific Ocean for *Fragilariopsis* distribution (Aizawa et al. 2005), in that the sequences were not recovered at the most coastal station.

The high relative abundance of many species linked to upwelling environments in transition zone samples supports the previous hypothesis that the increased transition zone productivity is a result of physical dynamics similar to upwelling zones, where higher nutrient waters from the oceanic HNLC mix

with waters with higher trace metal concentrations from the coast (Ribalet et al. 2010). We also find that our sequencing results support the conclusion that the coastal–HNLC transition zone is an area of high phytoplankton diversity. We observed the highest diatom diversity and most even community in the transition zone B stations (P2, P5, and P6), while the transition zone A stations (P3 and P4) had slightly lower diatom diversity and evenness, but these values were still higher than those at the most coastal (P1) and oceanic (P8) stations. This then leads to the question: What environmental factors may be driving this increased diversity?

4.2. Distance and trace metals as drivers of diatom community composition

Understanding the limitations of correlation analyses, specifically that correlation does not mean causation, we examined correlations between the environmental variables we measured and shifts in community composition. A challenge with doing this on a small dataset and evaluating a wide array of environmental variables, is that only combinations of variables that yielded a high correlation coefficient could be considered significant based on permutation tests. While no single environmental factor correlated significantly with changing diatom community composition, multiple sets of 4 environmental factors resulted in significant correlations. Some environmental variables included in these groupings were correlated with one another, which both explains why multiple groups of variables yielded similar correlation coefficients and complicates our interpretation of the BIO-ENV results.

Two environmental factors that did not correlate with other environmental variables, longitude and dCu, were part of every combination of 4 environmental factors that resulted in significant correlations with shifts in diatom community composition. Longitudinal position incorporates a number of factors that could influence diatom communities. Changes in longitude mark the distance along the transect, so it could be a manifestation of the transition from a coastal high nutrient regime to an oligotrophic and stratified regime. However, the other factors associated with that transition, like distance from shore, macronutrient concentrations, chlorophyll fluorescence, bottom depth, mixed layer depth, and depth of the chlorophyll maximum, did not result in strong correlations. Longitude also incorporates proximity of sampling stations. Research on bacterioplankton

has found that while immediate proximity (<50 km) may be associated with community similarity, community similarity has not been found to correlate with proximity in distances beyond 50 km, except in samples from a narrow shallow straight (Hewson et al. 2006). The total distance covered by our sampling transect was 234 km; the closest pair of stations (P2–P3) were 24.66 km apart and were <45% similar to one another. The next closest pair (P3–P4) were 24.71 km apart and were 70% similar. Additionally, Stns P2 and P6 were 2 of our furthest stations at 123.5 km apart from each other and were 60% similar.

While diatom community composition strongly correlated with longitude, it was not the whole story. As previously mentioned, dCu was a factor that was included in all groups of environmental variables that significantly correlated with diatom community shifts. Two other trace metals, dZn and dFe, were also included in groups of environmental factors that correlated significantly with diatom community composition. Fe limitation has been a major focus of research along Line P for some time. Although concentrations and potential biological importance of dZn to phytoplankton have previously been explored in the region (Lohan et al. 2002, Janssen & Cullen 2015), there has not been as much focus on this potentially important micronutrient. Our results suggest that the impact of dZn and dCu on phytoplankton should be explored further in this region. Either density or salinity, which were positively correlated with one another and negatively correlated with dFe, was included in 4 out of 5 sets of environmental factors that correlated with shifts in diatom community composition. Because of the correlation between dFe, salinity, and density, it is impossible to deconvolute which factor, if any, could be an important driver of diatom community composition, though all 3 have the potential to impact the productivity of individual phytoplankton species. Previous work has shown that low dFe concentrations are typical in the eastern portion of the Line P transect, although chronic Fe limitation is not usually associated with stations east of P12 (Boyd et al. 1998). Interestingly, while dFe correlated with shifting diatom community composition, neither of the Fe binding ligand associated parameters, FeL concentration or calculated Fe', or the nitrate:dFe ratio, which has often been found to be a better indicator of Fe limitation than dFe concentrations, were found to significantly increase the correlation.

In light of the correlation analyses and the previous work indicating that dFe is an important driver of phytoplankton productivity along Line P (Harrison 2002 and references therein), we examined whether any portion of the diatom community was exhibiting signs of Fe limitation along our transect. We examined the gene expression of *T. oceanica*, which accounted for ≥1% of the reads at all stations except P3, but was not the most abundant diatom at any of the stations sampled. A benefit of the recovery of *T. oceanica* sequences is that it is a diatom relatively tolerant to low Fe (Sunda & Huntsman 1995), and one of the few diatom species for which we have a molecular method by which to test for Fe limitation (Chappell et al. 2015). Using the FLDA1 gene expression assay, we found evidence of Fe limitation of *T. oceanica* at Stns P5, P6, and P8. While the gene expression we measured was not as high as previously seen in other areas of the northeast Pacific Ocean (Chappell et al. 2015), the dFe was also not as low as was measured in that previous work (0.2 nM dFe), which sampled both further north and further offshore than the current work. Nevertheless, the expression at Stns P5, P6, and P8 was over 100× higher than what was found at Stns P1, P2, and P4. N-limitation can lead to a slight (3-fold) increase in FLDA1 expression (Chappell et al. 2015), but that is well below the 100-fold increase observed here. An alternative measure of Fe limitation is ISIP3 gene expression. However, ISIP3 transcripts were above detection limit only at Stns P6 and P8 (a sample processing issue precluded any measurement for Stn P5). The ISIP3 expression values at P6 and P8 were 2.6× lower than the established ISIP3 threshold for Fe limitation, but 33× greater than the replete expression levels from *T. oceanica* cultures (Chappell et al. 2015). Thus the *T. oceanica* ISIP3 induction is indicative of Fe stress, but perhaps not Fe starvation.

Non-detection of ISIP3 transcripts at Stns P1, P2, and P4 is meaningful, as ACT expression levels suggested that *T. oceanica* was replete for Fe at these locations (agreeing with the FLDA1 results). Given the potential for a slight bump in FLDA1 expression associated with N limitation and the elevated but sub-limitation-cutoff ISIP3 expression, we classify the *T. oceanica* at Stns P5, P6, and P8 as Fe stressed, rather than as Fe growth-limited. We note that while diatom communities were >60% similar at Stns P2, P5, and P6, there was no evidence of Fe stress at Stn P2. Some of the changes in diatom community composition that were apparent between P2 and P5/P6 are consistent with what might be expected as a phytoplankton community begins to experience Fe stress. When compared to P2, the diatom communities at P5/P6 show a reduction in shelf species, such as *S. costatum*, and an increase in abundance of low-Fe tolerant oceanic diatoms like *P. granii*, which is

common in the open ocean portion of the transect (Ribalet et al. 2010) and has numerous strategies for surviving in low-Fe environments (Marchetti et al. 2009, 2012, 2017). We also note that the dFe levels at P2, where there was no sign of Fe stress, were similar to those at P5, P6, and P8, where *T. oceanica* was Fe stressed.

A challenge of discrete sampling of a transect is that we have no knowledge of prior dFe concentrations at any of the sites, which may have been only recently drawn down at P2. We also do not know the source, composition, or bioavailability of the dFe at any of the sites. These complications are precisely why we need methods like the *T. oceanica* Fe stress assay, which allows us to use the diatoms themselves as *in situ* monitors of Fe status. As *T. oceanica* represents a diatom relatively tolerant to low Fe (Sunda & Huntsman 1995), this indicates that other diatoms at Stns P5, P6, and P8 may also be experiencing Fe stress, which could be resulting in the diatom community composition shifts we observed. In order to evaluate this, we would need to explore the Fe status of additional diatom species. These results showing Fe limitation of at least 1 member of the diatom community, combined with the correlation between shifting diatom community composition and trace metal concentrations, further support the proposed hypothesis that changing Fe concentrations may be important in controlling phytoplankton community diversity and productivity in the coastal–HNLC transition zone (Ribalet et al. 2010). It is a possibility, however, that dFe is an important factor for only some diatoms, such as *T. oceanica*, and we cannot discount the fact that non-concentration associated factors — longitude, salinity, and density — were correlated as well.

5. CONCLUSIONS

In conclusion, our results support prior work indicating that the coastal–HNLC transition zone is an area of high phytoplankton productivity and diatom diversity, where coastal and oceanic diatom species are both present. The current study expanded on this previous work by examining the distribution of a wider range of diatom genera, reaching similar conclusions regarding diversity and species composition. In this study, the shifts in diatom community composition were correlated with a combination of physical (longitude, salinity, density) and chemical (trace metal concentrations of dCu, dFe, and dZn) factors. To further explore the importance of trace metals, which have previously been shown to be important

in controlling productivity further west along the Line P transect, we used molecular assays for Fe limitation of the diatom *T. oceanica* to show evidence of Fe stress on the oceanic end of the transect. While the transcript abundance of FLDA1 indicated Fe limitation of *T. oceanica*, the level of expression was not as high as has been observed further offshore in waters with lower dFe, which suggests that the populations were Fe stressed, rather than Fe limited for growth. Overall, these data support previous work suggesting one of the drivers of coastal–HNLC transition zone diatom communities and productivity likely includes the mixing of high nutrient waters from the HNLC with high trace element waters from the coast. Our results, in light of data from the Line P cruise that immediately followed ours, further highlight that the transition zone region is an area that experiences dynamic shifts in productivity, especially early in the season when upwelling is not consistent.

LITERATURE CITED

Aizawa C, Tanimoto M, Jordan R (2005) Living diatom assemblages from North Pacific and Bering Sea surface waters during summer 1999. Deep Sea Res II 52: 2186–2205

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Bisby FA, Roskov YR, Orrell TM, Nicolson D and others (2011) Species 2000 & ITIS Catalogue of Life: 2011 annual checklist. www.catalogueoflife.org/annual-checklist/2011/

Blasco D, Estrada M, Jones B (1980) Relationship between the phytoplankton distribution and composition and the hydrography in the northwest African upwelling region near Cabo-Corbeiro. Deep Sea Res 27:799–821

Booth BC, Lewin J, Postel JR (1993) Temporal variation in the structure of autotrophic and heterotrophic communities in the Sub-Arctic Pacific. Prog Oceanogr 32:57–99

Boyd P, Harrison PJ (1999) Phytoplankton dynamics in the NE subarctic Pacific. Deep Sea Res II 46:2405–2432

Boyd P, Berges JA, Harrison PJ (1998) In vitro iron enrich-
ment experiments at iron-rich and -poor sites in the NE
subarctic Pacific. J Exp Mar Biol Ecol 227:133–151

Boyd PW, Law CS, Wong CS, Nojiri Y and others (2004) The
decline and fate of an iron-induced subarctic phyto-
plankton bloom. Nature 428:549–553

Boyd PW, Strzepek R, Takeda S, Jackson G and others
(2005) The evolution and termination of an iron-induced
mesoscale bloom in the northeast subarctic Pacific. Lim-
nol Oceanogr 50:1872–1886

Bray JR, Curtis JT (1957) An ordination of the upland forest
communities of southern Wisconsin. Ecol Monogr 27:
325–349

Buck KR, Chavez FR, Davis AS (2008) *Minidiscus triocula-
tus*, a small diatom with a large presence in the up-
welling systems of central California. Nova Hedwigia
Beih 133:1–6

Bundy RM, Biller DV, Buck KN, Bruland KW, Barbeau KA
(2014) Distinct pools of dissolved iron-binding ligands in
the surface and benthic boundary layer of the California
Current. Limnol Oceanogr 59:769–787

Carstensen J, Klais R, Cloern JE (2015) Phytoplankton
blooms in estuarine and coastal waters: seasonal patterns
and key species. Estuar Coast Shelf Sci 162:98–109

Casteleyn G, Chepurnov VA, Leliaert F, Mann DG and oth-
ers (2008) *Pseudo-nitzschia pungens* (Bacillariophyceae):
a cosmopolitan diatom species? Harmful Algae 7:
241–257

Chappell PD, Whitney LP, Haddock TL, Menden-Deuer S,
Roy EG, Wells ML, Jenkins BD (2013) *Thalassiosira* spp.
community composition shifts in response to chemical
and physical forcing in the northeast Pacific Ocean.
Front Microbiol 4:273

Chappell PD, Whitney LP, Darer AI, Jean-Charles S, Jenkins
BD (2015) Genetic indicators of iron limitation in wild
populations of *Thalassiosira oceanica* from the northeast
Pacific Ocean. ISME J 9:592–602

Clarke KR, Ainsworth M (1993) A method of linking multi-
variate community structure to environmental variables.
Mar Ecol Prog Ser 92:205–219

Clarke KR, Gorley RN (2006) Primer v6: user manual/tutorial.
PRIMER-E, Plymouth

Clarke KR, Somerfield PJ, Gorley RN (2008) Testing of null
hypotheses in exploratory community analyses: similar-
ity profiles and biota-environment linkage. J Exp Mar
Biol Ecol 366:56–69

Cutter GA, Bruland KW (2012) Rapid and noncontaminating
sampling system for trace elements in global ocean sur-
veys. Limnol Oceanogr Methods 10:425–436

Davis CO, Hollibaugh JT, Seibert DLR, Thomas WH, Harri-
son PJ (1980) Formation of resting spores by *Leptocylin-
drus danicus* (Bacillariophyceae) in a controlled experi-
mental ecosystem. J Phycol 16:296–302

El-Sabaawi R, Harrison PJ (2006) Interactive effects of irra-
diance and temperature on the photosynthetic physiol-
ogy of the pennate diatom *Pseudo-nitzschia granii* (Bacil-
lariophyceae) from the northeast subarctic Pacific.
J Phycol 42:778–785

Eren AM, Morrison HG, Lescault PJ, Reveillaud J, Vineis
JH, Sogin ML (2015) Minimum entropy decomposition:
unsupervised oligotyping for sensitive partitioning of
high-throughput marker gene sequences. ISME J 9:
968–979

Freeland H (2007) A short history of Ocean Station Papa and
Line P. Prog Oceanogr 75:120–125

French FW, Hargraves PE (1986) Population dynamics of the
spore-forming diatom *Leptocylindrus danicus* in Narra-
gansett Bay, Rhode Island. J Phycol 22:411–420

Gledhill M, Buck KN (2012) The organic complexation of
iron in the marine environment: a review. Frontiers
Microbiol: 3

Gordon LI, Jennings JC, Ross AA, Krest JM (1993) A sug-
gested protocol for continuous flow automated analysis
of seawater nutrients (phosphate, nitrate, nitrite, and sili-
cic acid) in the WOCE Hydrographic Program and the
Joint Global Ocean Fluxes Study. WOCE Hydrographic
Program Office Methods Manual WHPO 91-1. https://
www.nodc.noaa.gov/woce/woce_v3/wocedata_1/whp/
manuals.htm

Grasshoff K, Kremling K, Erhard M (1999) Methods of sea-
water analysis, Vol 3. Wiley-VCH, Weinheim

Harrison PJ (2002) Station Papa time series: insights into
ecosystem dynamics. J Oceanogr 58:259–264

Hasle GR (1995) *Pseudo-nitzschia pungens* and *P. multi-
series* (Bacillariophyceae)—nomenclatural history, mor-
phology, and distribution. J Phycol 31:428–435

Hewson I, Steele JA, Capone DG, Fuhrman JA (2006) Tem-
poral and spatial scales of variation in bacterioplankton
assemblages of oligotrophic surface waters. Mar Ecol
Prog Ser 311:67–77

Hogle S, Bundy RM, Blanton JM, Allen EE, Barbeau KA
(2016) Copiotrophic marine bacteria are associated with
strong iron-binding ligand production during phyto-
plankton blooms. Limnol Oceanogr Lett 1:36–43

Holmes RM, Aminot A, Kerouel R, Hooker BA, Peterson BJ
(1999) A simple and precise method for measuring am-
monium in marine and freshwater ecosystems. Can J
Fish Aquat Sci 56:1801–1808

Huang L, Jian W, Song X, Huang X and others (2004)
Species diversity and distribution for phytoplankton of
the Pearl River estuary during rainy and dry seasons.
Mar Pollut Bull 49:588–596

Hubbard KA, Rocap G, Armbrust EV (2008) Inter- and intra-
specific community structure within the diatom genus
Pseudo-nitzschia (Bacillariophyceae). J Phycol 44:
637–649

Hubbard KA, Olson CE, Armbrust EV (2014) Molecular
characterization of *Pseudo-nitzschia* community struc-
ture and species ecology in a hydrographically complex
estuarine system (Puget Sound, Washington, USA). Mar
Ecol Prog Ser 507:39–55

Janssen DJ, Cullen JT (2015) Decoupling of zinc and silicic
acid in the subarctic northeast Pacific interior. Mar Chem
177:124–133

Johnson KS, Elrod V, Fitzwater S, Plant J and others (2007)
Developing standards for dissolved iron in seawater.
EOS Trans Am Geophys Union 88:131–132

La Roche J, Boyd PW, McKay RML, Geider RJ (1996) Flavo-
doxin as an in situ marker for iron stress in phytoplank-
ton. Nature 382:802–805

Leblanc K, Arístegui J, Armand L, Assmy P and others
(2012) A global diatom database—abundance, biovol-
ume and biomass in the world ocean. Earth Syst Sci Data
4:149–165

Lee JM, Boyle EA, Echegoyen-Sanz Y, Fitzsimmons JN,
Zhang R, Kayser RA (2011) Analysis of trace metals (Cu,
Cd, Pb, and Fe) in seawater using single batch nitrilotri-
acetate resin extraction and isotope dilution inductively
coupled plasma mass spectrometry. Anal Chim Acta 686:
93–101

Lohan MC, Statham PJ, Crawford DW (2002) Total dissolved zinc in the upper water column of the subarctic North East Pacific. Deep Sea Res II 49:5793–5808

Mann DG, Droop SJM (1996) Biodiversity, biogeography and conservation of diatoms. Hydrobiologia 336:19–32

Marchetti A, Sherry ND, Kiyosawa H, Tsuda A, Harrison PJ (2006) Phytoplankton processes during a mesoscale iron enrichment in the NE subarctic Pacific. I. Biomass and assemblage. Deep Sea Res II 53:2095–2113

Marchetti A, Parker MS, Moccia LP, Lin EO and others (2009) Ferritin is used for iron storage in bloom-forming marine pennate diatoms. Nature 457:467–470

Marchetti A, Schruth DM, Durkin CA, Parker MS and others (2012) Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. Proc Natl Acad Sci USA 109:E317–E325

Marchetti A, Moreno CM, Cohen NR, Oleinikov I and others (2017) Development of a molecular-based index for assessing iron status in bloom-forming pennate diatoms. J Phycol 53:820–832

Marshall HG, Cohn MS (1983) Distribution and composition of phytoplankton in northeastern coastal waters of the United States. Estuar Coast Shelf Sci 17:119–131

Martin JH (1990) Glacial-interglacial CO_2 change: the iron hypothesis. Paleoceanography 5:1–13

Martin JH, Fitzwater SE (1988) Iron-deficiency limits phytoplankton growth in the northeast Pacific subarctic. Nature 331:341–343

Miller CB, Frost BW, Wheeler PA, Landry MR, Welschmeyer N, Powell TM (1991) Ecological dynamics in the Sub-Arctic Pacific, a possibly iron-limited ecosystem. Limnol Oceanogr 36:1600–1615

Nelson DM, Treguer P, Brzezinski MA, Leynaert A, Queguiner B (1995) Production and dissolution of biogenic silica in the ocean — revised global estimates, comparison with regional data and relationship to biogenic sedimentation. Global Biogeochem Cycles 9:359–372

Oksanen J, Blanchet FG, Friendly M, Kindt R and others (2017) vegan: community ecology package. https://cran.r-project.org/web/packages/vegan/index.html

Omanovic D, Gamier C, Pizeta I (2015) ProMCC: an all-in-one tool for trace metal complexation studies. Mar Chem 173:25–39

Peña MA, Varela DE (2007) Seasonal and interannual variability in phytoplankton and nutrient dynamics along Line P in the NE subarctic Pacific. Prog Oceanogr 75:200–222

Ribalet F, Marchetti A, Hubbard KA, Brown K and others (2010) Unveiling a phytoplankton hotspot at a narrow boundary between coastal and offshore waters. Proc Natl Acad Sci USA 107:16571–16576

Round FE, Crawford RM, Mann DG (1990) The diatoms. Cambridge University Press, Cambridge

Schlitzer R (2017) Ocean Data View. http://odv.awi.de/

Schloss PD, Westcott SL, Ryabin T, Hall JR and others (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–7541

Shannon CE (1948) A mathematical theory of communication. The Bell Syst Tech J 27:379–423

Stehr CM, Connell L, Baugh KA, Bill BD, Adams NG, Trainer VL (2002) Morphological, toxicological, and genetic differences among Pseudo-nitzschia (Bacillariophyceae) species in inland embayments and outer coastal waters of Washington State, USA. J Phycol 38:55–65

Sunda WG, Huntsman SA (1995) Iron uptake and growth limitation in oceanic and coastal phytoplankton. Mar Chem 50:189–206

Taylor FJR, Haigh R (1996) Spatial and temporal distributions of microplankton during the summers of 1992–1993 in Barkley Sound, British Columbia, with emphasis on harmful species. Can J Fish Aquat Sci 53:2310–2322

Trigueros JM, Orive E (2000) Tidally driven distribution of phytoplankton blooms in a shallow, macrotidal estuary. J Plankton Res 22:969–986

Varela DE, Harrison PJ (1999) Seasonal variability in nitrogenous nutrition of phytoplankton assemblages in the northeastern subarctic Pacific Ocean. Deep Sea Res II 46:2505–2538

Whitney FA, Wong CS, Boyd PW (1998) Interannual variability in nitrate supply to surface waters of the Northeast Pacific Ocean. Mar Ecol Prog Ser 170:15–23

Whitney LP, Lins JJ, Hughes MP, Wells ML, Chappell PD, Jenkins BD (2011) Characterization of putative iron responsive genes as species-specific indicators of iron stress in Thalassiosiroid diatoms. Front Microbiol 2:234

Zeng C, Rosengard SZ, Burt W, Peña MA and others (2018) Optically-derived estimates of phytoplankton size class and taxonomic group biomass in the Eastern Subarctic Pacific Ocean. Deep Sea Res I 136:107–118

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

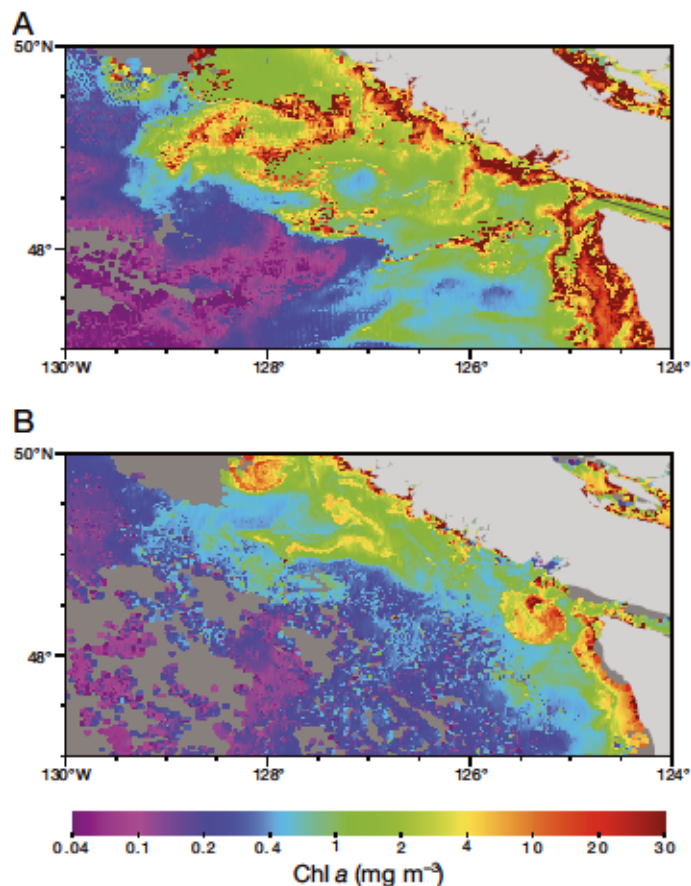**APPENDIX.** Additional data

A

B

Fig. A1. Chlorophyll *a* 3 d composite satellite image from Aqua-MODIS NPP 0.0125 degree resolution for the region sampled during (A) 14–16 May 2012 and (B) 22–24 May 2012. NOAA Coast Watch; data courtesy of NASA GSFC (OBPG)