SCIENTIFIC REPORTS

natureresearch

OPEN

Depth-dependent transcriptomic response of diatoms during spring bloom in the western subarctic Pacific Ocean

Shigekatsu Suzuki[1], Takafumi Kataoka[2], Tsuyoshi Watanabe[3], Haruyo Yamaguchi [1], Akira Kuwata[3] & Masanobu Kawachi[1]

Diatoms play important roles in primary production and carbon transportation in various environments. Large-scale diatom bloom occurs worldwide; however, metabolic responses of diatoms to environmental conditions have been little studied. Here, we targeted the Oyashio region of the western subarctic Pacific where diatoms bloom every spring and investigated metabolic response of major diatoms to bloom formation by comparing metatranscriptomes between two depths corresponding to different bloom phases. *Thalassiosira nordenskioeldii* and *Chaetoceros debilis* are two commonly occurring species at the study site. The gene expression profile was drastically different between the surface (late decline phase of the bloom; 10 m depth) and the subsurface chlorophyll maximum (SCM, initial decline phase of the bloom; 30 m depth); in particular, both species had high expression of genes for nitrate uptake at the surface, but for ammonia uptake at the SCM. Our culture experiments using *T. nordenskioeldii* imitating the environmental conditions showed that gene expression for nitrate and ammonia transporters was induced by nitrate addition and active cell division, respectively. These results indicate that the requirement for different nitrogen compounds is a major determinant of diatom species responses during bloom maturing.

Diatoms contribute up to 20–40% of oceanic primary productivity forming large-scale blooms mainly in coastal and upwelling regions[1–3]. They are key organisms for carbon flux in the ocean, acting as large carbon sinks by vertical carbon transportation by sinking to deeper layers[4]. Several tens of species occur in diatom blooms; however, a few species are usually dominant (60–90% abundance). For example, *Chaetoceros* spp. and *Thalassiosira* spp. are dominant in the Oyashio region of the western North Pacific[5,6] and the world ocean[7].

Diatom blooms are induced by the supply of resources that promote growth, such as light, macronutrients, and iron. Water column stability and local mixing are indicated as factors promoting diatom blooms[8]. Increase of light intensity causes diatom blooms in coastal regions, such as the Narragansett Bay[9], the fjord of western Norway[10], and the Southern Ocean[11]. Diatoms are known to response to upwelled nutrients and grow rapidly[12,13]. Iron enrichment experiments, such as the western North Pacific[14], suggest that supply of iron is an important factor of blooming of diatoms[15]. However, details of physiological processes in blooming diatoms, e.g., nutritional utilizations and cell responses of bloom-forming species, remain unclear.

Metatranscriptome analyses are useful for studying algal cell responses in natural environments[16–18]. Alexander *et al.*[19] compared gene expression patterns of two bloom-forming diatoms in Narragansett Bay in the western North Atlantic and found differential gene expression in nitrogen and phosphorus metabolism pathways among the diatoms. However, there are methodological difficulties to applying metatranscriptome analysis to natural microbial communities. For example, sampling without disturbing transcripts is difficult because transcripts are susceptible to environmental change during sampling. Lack of reference genome information in public databases also prevents the selection of appropriate control organisms.

[1]Center for Environmental Biology and Ecosystem Studies, National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba, Ibaraki, Japan. [2]Faculty of Marine Science and Technology, Fukui Prefectural University, 1-1 Gakuen-cho, Obama, Fukui, Japan. [3]Tohoku National Fisheries Research Institute, Japan Fisheries Research and Education Agency, 3-27-5 Shinhama-cho, Shiogama, Miyagi, Japan. Correspondence and requests for materials should be addressed to S.S. (email: suzuki.shigekatsu@nies.go.jp)
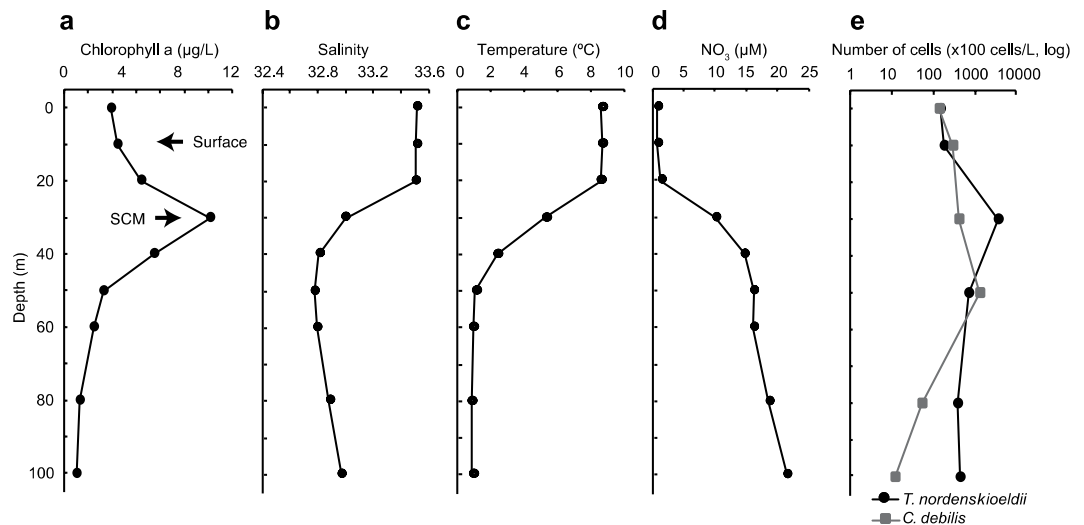
Figure 1. Features of the environment and diatom blooms. Concentration of chlorophyll *a* (**a**), salinity (**b**), temperature (**c**), nitrate (**d**) along depth are shown. The chlorophyll *a* concentration peaks at around 30 m, the subsurface chlorophyll maximum (SCM). The sample at 10 m was used as the surface sample. (**e**) Number of cells of *Tharassiosira nordenskioeldii* (black) and *C. debilis* (gray) along depth. The number of *T. nordenskioeldii* peaks at the SCM, but that of *C. debilis* gradually increases from the surface to 50 m depth.

The western subarctic Pacific is a preferable field to study cell responses of bloom-forming species because large-scale diatom blooms occur annually in spring in this region[5,20,21]. The hydrography of this region in spring is complex, with dominant water columns: Oyashio Water (OW), Coastal Oyashio Water (COW), and modified Kuroshio Water (MKW)[22]. Warming by increasing solar radiation and inflow of the low-salinity COW lead to formation of a shallow surface layer and make phytoplankton use nutrients supplied by winter mixing to form massive diatom blooms[21–24]. The alleviation of light limitation due to the mixed layer shoaling and the accumulation of diatom cells due to the weakening of turbulence are also related to onset of diatom blooms[25].

Here, we performed metatranscriptome analyses of diatom communities during spring blooms in the western North Pacific to reveal bloom-specific cell responses, focusing on differences in depth. We compared two diatom communities obtained from the sea surface (late decline phase of the bloom; 10 m) and subsurface chlorophyll maximum (SCM, initial decline phase of the bloom; 30 m), particularly targeting *T. nordenskioeldii* and *C. debilis*, which commonly occur during spring bloom in this region[21]. Moreover, gene expression analysis of nutrient-controlled cultures was conducted to confirm the metatranscriptomic analysis.

Results and Discussion
Characterization of the environments and diatom blooms.
We sampled an algal bloom at Station A4 located in the Oyashio region of the western subarctic Pacific. The SCM was observed at a depth of around 30 m. Here, the chlorophyll *a* concentration (12.1 μg L^{-1}) was about three times higher than that at the sea surface (10 m depth; 4.4 μg L^{-1}) (Fig. 1a; Supplementary Table S1). These chlorophyll levels are similar to the values observed during spring diatom blooms in this region[23,26]. Total cell numbers of diatoms were 275,830 and 1,600,465 cells L^{-1} in the surface and SCM, respectively. The environmental conditions between the surface and SCM were different in terms of temperature, salinity, and concentration of nitrate, nitrite, silicic acid, and phosphate (Fig. 1b–d; Supplementary Table S1). Temperature and salinity at the SCM (5.2 °C and 33.0) were lower than at the surface (8.7 °C and 33.5), indicating that the SCM and surface were affected by the water columns COW and MKW, respectively[27]. Upper layers including the layer of the surface sample (10 m) had lower nutrient concentrations, whereas lower layers, where the SCM sample was collected, had higher nutrient concentrations. Thus, the phytoplankton bloom we analysed would be in nutrient-rich conditions at the SCM (30 m) due to the influence of cold nutrient-rich COW, whereas low-nutrient conditions existed at the surface (10 m). To elucidate actively transcriptionally responding taxa, we performed taxonomic analyses of the RNA-seq reads (Supplementary Fig. S1). The RNA-seq reads were mainly composed of diatoms, Opisthokonta, Alveolata, and Viridiplantae both in the surface and SCM samples. Diatoms were the predominant group in both samples. The proportion of diatoms in the SCM layer (76.7%) was higher than that in the surface layer (51.4%), suggesting that diatoms have more active transcriptional responses in the SCM layer than at the surface. To elucidate development stages of the diatom blooms at the surface and SCM, we used an index, I_{devel}[20]. I_{devel} was 0.95 and 0.77 at the surface and SCM, respectively, and thus diatoms of the surface and SCM are under late and initial decline phase of the diatom bloom. In the surface, nitrate concentration was lower than the SCM (Fig. 1d; Supplementary Table S1), however, the nitrogen concentration is not limiting factors for the diatom growth rates in the surface sample (Supporting Text; Supplementary Fig. S2). For light intensities, it has been reported that growth of *T. nordenskioeldii* is limited in cold water at high light intensity[28]. Thus, the cells in the surface can be affected by high light intensity.
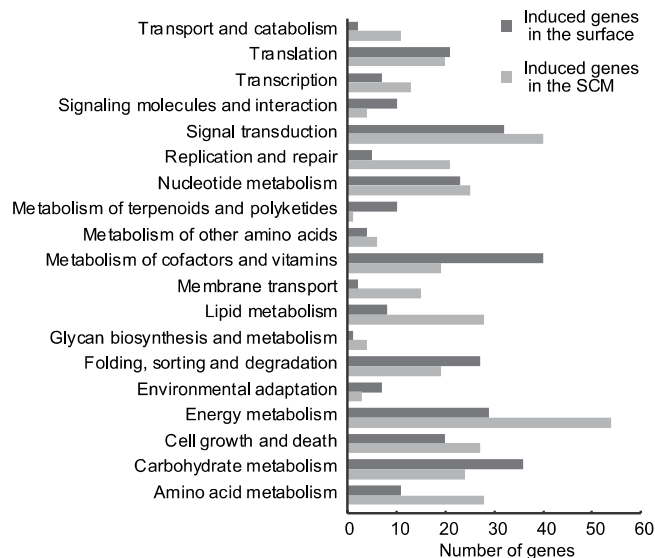
Figure 2. Functional classification of differentially expressed genes in *T. nordenskioeldii*. Functional analyses were performed using Kyoto Encyclopedia of Genes and Genomes (KEGG). Black and gray bars represent numbers of significantly induced genes (FDR < 0.01) at the surface and SCM, respectively.

Our direct microscopy counting showed that the main diatom species in the blooms were *T. nordenskioeldii*, *C. compressus*, *C. socialis*, *C. diadema*, *C. debilis*, and *Fragilariopsis cylindrus* (Supplementary Fig. S3a–e). RNA-seq or genome data of *F. cylindrus*, *C. socialis*, and *C. debilis* are available and fortunately, we could acquire the strain of *T. nordenskioeldii* from the same experiment region. Therefore, we selected those four diatoms as the representative transcripts in this study. Large cell numbers for *T. nordenskioeldii*, *C. socialis*, and *C. debilis* were also observed in the Oyashio region during spring diatom bloom[6,21], and their relatives were observed in our past sampling in 2015 as described in Supporting Text and Supplementary Fig. S4.
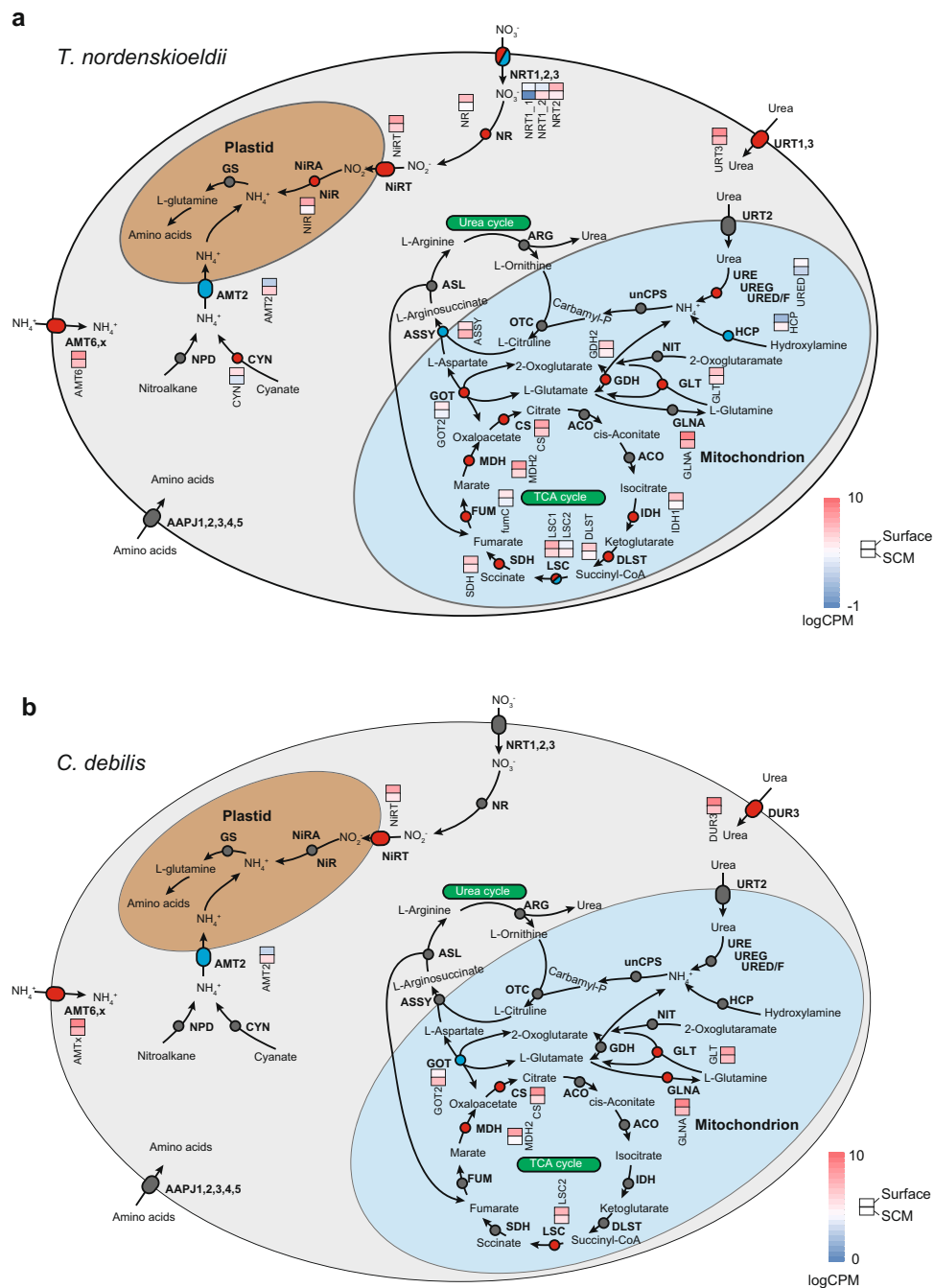
Overall gene expression patterns in *Thalassiosira nordenskioeldii* and *Chaetoceros debilis*. RNA-seq reads of both the surface and SCM samples were mapped to 91.6%, 59.0%, 6.0%, and 24.1% of all of the protein-coding genes originating from *T. nordenskioeldii*, *C. debilis*, *C. socialis*, and *F. cylindrus*, respectively. For more detail on the gene expression, the top two diatoms, *T. nordenskioeldii* and *C. debilis*, were subjected to DEG analysis because mapping rates of all replicates of *T. nordenskioeldii* and *C. debilis* were higher than *C. socialis*, and *F. cylindrus* (Supplementary Fig. S5). Few reads were mapped to *C. socialis*, which contradicted our microscopic cell counting showing that *C. socialis* was one of the major species (Supplementary Fig. S3a). This might be due to intraspecific diversity of the gene sequences[29] and identification by microscopic observation is difficult to separate the cryptic genetic diversity.

Among the protein-coding genes, 2,882 and 542 genes of *T. nordenskioeldii* and *C. debilis*, respectively, were significantly differently expressed between the two depths (FDR < 0.01) (Supplementary Fig. S6a,b, Tables S2, S3. Higher gene expression was detected in the SCM sample (63.0% of total DEGs) for *T. nordenskioeldii* and in the surface sample (55.2% of total DEGs) for *C. debilis*. In addition, KEGG category-based functional analysis showed that genes for cofactors and vitamin metabolism dominated only in surface samples of *T. nordenskioeldii*, while genes for energy metabolism, lipid metabolism, and amino acid metabolism dominated only in the SCM (Fig. 2).

In the SCM, *T. nordenskioeldii* and *C. debilis* induced expression of some genes for S-G_2/M phase-specific cyclin A and DNA replication system such as replication factor C and DNA helicase. In addition, expression of genes for silicon transporters (*SIT1*) were induced in the SCM samples (Supplementary Fig. S7). Silicon limitation alone induces transcription of *SIT1*[30] and thus diatom cells at the SCM are under high silicon requirement. These results suggest that the diatoms in the SCM had more frequent cell division than the surface, which is consistent with our microscopic cell counting showing that the cell numbers of *T. nordenskioeldii* and *C. debilis* in the SCM samples were greater than those in the surface sample (Fig. 1e).

Depth-dependent nitrogen metabolism in *Thalassiosira nordenskioeldii* and *Chaetoceros debilis*. For *T. nordenskioeldii*, expression profiles of nitrogen metabolism-related genes were significantly different between the surface and SCM (Fig. 3a): high-affinity nitrate transporter (*NRT2*), nitrate reductase (*NR*), nitrite transporter (*NiRT*), nitrite reductase (*NiR*), urea transporter (*URT3*), ammonium transporter (*AMT6*), cyanate hydratase (*CYN*), urease accessory protein (*URED/F*), glutamate dehydrogenase (*GDH*), and aspartate aminotransferase (*GOT2*) were induced in surface samples; whereas ammonium transporter (*AMT2*), hydroxylamine reductase (*HCP*), and argininosuccinate synthase (*ASSY*) were induced in SCM samples. Two genes for nitrate transporters (*NRT1_2* and *NRT1_1*) were induced in surface and SCM samples, respectively. Among the induced genes in surface samples, *NRT1_1*, *NRT2*, *NR*, *NiRT*, and *NiR* contribute to an intake of nitrate from outside and inside of the cells and a reduction of nitrate to ammonia in the plastids[31]. Therefore, *T. nordenskioeldii* preferably uses extracellular and intracellular nitrate/nitrite as a nitrogen source for amino acids at the surface. *CYN* is

Figure 3. Metabolism map of nitrogen metabolism and the different gene expression patterns in bloom-forming diatoms. Metabolites, enzymes, and transporters of *T. nordenskioeldii* (**a**) and *C. debilis* (**b**) are shown in this figure. The names of enzymes (circles) and transporters (ovals) are shown in bold type. Genes induced at the surface and SCM are depicted as red- and blue-colored circles/ovals, respectively. For differently expressed genes (FDR < 0.01), heatmaps of relative expression level (logarithmic read count per million; RPM) are shown. AAPJ: General L-amino acid transport system, ACO: acyl-CoA oxidase, AMT: ammonium transporter, ARG: arginase, ASL: argininosuccinate lyase, ASSY: argininosuccinate synthase, CYN: cyanate hydratase, CS: citrate synthase, DLST: dihydrolipoamide S-succinyltransferase, DUR: urea transporter, FUM: fumarate hydratase, GS: glutamine synthase, GDH: glutamate dehydrogenase, GLNA: glutamine synthetase, GLT: glutamate synthase, GOT: aspartate aminotransferase, HCP: hydroxylamine reductase, IDH: isocitrate dehydrogenase, LSC: succinate–CoA ligase, MDH: malate dehydrogenase, NPD: nitronate monooxygenase, NiR: nitrite reductase, NiRA: nitrite reductase, NiRT: nitrite transporter, NIT: 2-oxoglutaramate amidase, NR: nitrate reductase, NRT: nitrate transporter, OTC: ornithine carbamoyltransferase, SDH: succinate dehydrogenase, unCPS: carbamoyl-phosphate synthase, URE: urease, UREG: urease accessory protein, URED/F: urease accessory protein, URT: urea transporter.

involved in conversion of cyanate to NH_3 in the cytoplasm[32]. Cyanate is a major nitrogen source for marine phytoplankton in oligotrophic regions of the mid Atlantic (up to 10% of total N uptake[33]). Widner *et al.*[33] also showed heterogeneous distribution of cyanate throughout the water column and suggested release of cyanate to the outside of cells in SCM. In our analysis, *CYN* of *T. nordenskioeldii* was induced in the surface samples where nitrate concentration was low (Fig. 1d; Supplementary Table S1), suggesting that insufficient nitrogen was compensated by using cyanate at the surface. The induction of *CYN* was accompanied by high expression of *NR* in surface samples (Fig. 3a). This can be explained by consumption of intercellular cyanate, because cyanate is known to inhibit NR activity post-transcriptionally in *Chlorella vulgaris*[34]. *URT3* and *AMT6* were also induced in surface sample, contributing uptake of nitrogen sources, i.e. urea and ammonia, from the outside of cells. Urea hydrolysis for *T. nordenskioeldii* was uncertain because *URED* was induced but induction of other genes related to the urease complex was not observed as reported in Bender *et al.*[35]. On the other hand, among the induced genes in the SCM samples, an ammonia transporter, *AMT2*, was induced, while *AMT6* was induced in surface samples. AMT2 is located in plastid membranes and transports NH_4^+ from cytoplasm into plastids[36]. Thus, *T. nordenskioeldii* could activate the ammonium metabolism of plastids for use of intracellular ammonia as a nitrogen source for amino acids at the SCM during the matured spring bloom in this region. AMT6 is located in cytoplasmic membranes and transports NH_4^+ from outside of cells. The induction of *AMT6* is consistent with nitrogen limitation at the surface. Together with nitrogen metabolism, HCP and NR serve as electron sinks for redox balancing for hydroxylamine/NO and NO_2^- reduction, respectively[37,38]. In our analyses, *HCP* and *NR* induction in SCM and surface samples, respectively, would explain diatoms' control of intracellular redox state. Diatoms are unique among algal species with respect to possessing a complete urea cycle[31,39]. We detected gene expression for all enzymes of the urea cycle in surface and SCM samples; however, only *ASSY* was significantly induced in the SCM samples. ASSY catalyses the biosynthesis of L-argininosuccinate from aspartate and L-citrulline by the urea cycle[40]. These results suggest that the activity of the urea cycle might be regulated by the expression level of *ASSY*, closely related to TCA cycle activity. The aspartate is shared between the urea cycle and the TCA cycle because aspartate is reversibly synthesized from oxaloacetate by GOT2. The TCA cycle was active in surface samples because genes for TCA cycle enzymes were significantly induced: citrate synthase, isocitrate dehydrogenase, dihydrolipoamide succinyltransferase, succinyl-CoA synthetase, and fumarate hydrase. These results indicate that *T. nordenskioeldii* had increased mitochondrial respiration at the surface. The inactivation of urea cycle in the surface might prevent competition of oxaloacetate with TCA cycle. In the reaction of oxaloacetate biosynthesis by GOT2, the genes encoding GOT2 and GDH were both induced in surface samples. These suggest induction of 2-oxoglutarate production as a substrate for GOT2. In the SCM, nutrient concentration was high (Fig. 1d, Supplementary Table S1), and thus urea cycle might be activated to detoxify ammonium that was an intermediate metabolite derived from the nitrogen sources (e.g., nitrate). The activation of urea cycle can inactivate TCA cycle in the SCM.

The gene expression patterns of *C. debilis* were similar to those of *T. nordenskioeldii* in terms of nitrogen metabolism (Fig. 3b). For *C. debilis*, *AMTx*, *NiRT*, urea transporter (*DUR3*), and glutamine synthases (*GLNA* and *GLT*) were induced in surface samples. Interestingly, expression of *GOT2* had an opposite pattern compared to that for *T. nordenskioeldii*; *GOT2* of *C. debilis* and *T. nordenskioeldii* was induced in surface and SCM samples, respectively. Although reasons for this difference are unclear, *C. debilis* may regulate activity of urea cycle by transcriptional level of *GOT2* unlike *T. nordenskioeldii*.

Factors for induction of *NRT2*, *NR*, and *AMT2*. To elucidate factors causing the differential gene expression in nitrogen metabolism, we conducted qPCR analysis of *T. nordenskioeldii*, cultured in a combination of nitrogen forms (i.e., NH_4^+ and NO_3^-) under various conditions described in Supporting Text and Supplementary Fig. S8. In particular, we focused on *NRT2*, *NR*, and *AMT2*, which were highly expressed in the surface and SCM samples, respectively, in *T. nordenskioeldii* (Fig. 3a). After adding nitrogen in the form of both NH_4^+ and NO_3^- to nitrogen-depleted *T. nordenskioeldii*, the cells began to grow and kept growing for 4 days (Fig. 4a). The relative expression values of *NRT2* increased to 90-fold within 30 min after nitrogen addition ($p < 0.01$, t-test), and then decreased in the next 1 and 2 days (Fig. 4b). The significant increase in the *NRT2* transcript was also seen when only NO_3^- was added (Supplementary Fig. S8c). *NR* expression was also induced by the NO_3^- addition (Supplementary Fig. S8c). On the other hand, *AMT2* expression was not significantly different ($p > 0.05$) (Supplementary Fig. S8c–e). These results indicate that *NRT2* in *T. nordenskioeldii* is induced by nitrate, as shown in an alga and plants[41–44]. Next, we investigated relationships between population growing state and gene expression by comparing *AMT2* gene expression in the culture strains (Fig. 5a,b). The *AMT2* expression values in nutrient-replete culture increased to 22-fold in 6 days' culturing ($p < 0.01$) and the level was maintained toward the stationary phase (8 and 11 days' culturing) (Fig. 5b). Meanwhile, NH_4^+ addition did not affect *AMT2* expression (Supplementary Fig. S8d,e). Therefore, *T. nordenskioeldii* AMT2 was induced during the actively growing phase, particularly late log phase, in the presence of NH_4^+ outside of the cell. Our results indicate that diatoms preferably use nitrate, but temporally use ammonium in the actively growing phase during the bloom-maturing period in their natural environment. Under active cell growth, the accessible but toxic ammonium in the cells would be immediately converted into detoxified forms, thus ammonium is less available.

Low light intensity also induced expression of *NRT2* (Supplementary Fig. S8b). The relative expression values of *NRT2* under low light (~80 μ photons $m^{-2} s^{-1}$) increased to three-fold compared to high light (~260 μ photons $m^{-2} s^{-1}$). However, the induction of *NRT* by nitrate addition is more efficient than that by low light intensity. These results indicate that the supply of nitrate can be a main factor to regulate *NRT2* expression.

The results of the culturing experiments are consistent with the cellular response of major diatoms in nature. Our field study showed that the NO_3^- concentration at the surface was low but sharply increased below ~20 m depth (Fig. 1d; Supplementary Table S1), implying that the surface is nitrogen limited. Additionally, small amount of nitrate might be supplied by turbulence. This situation can explain high expression of genes for nitrate uptake in both *T. nordenskioeldii* and *C. debilis* in surface samples as described above for *NRT2* of *T. nordenskioeldii*.
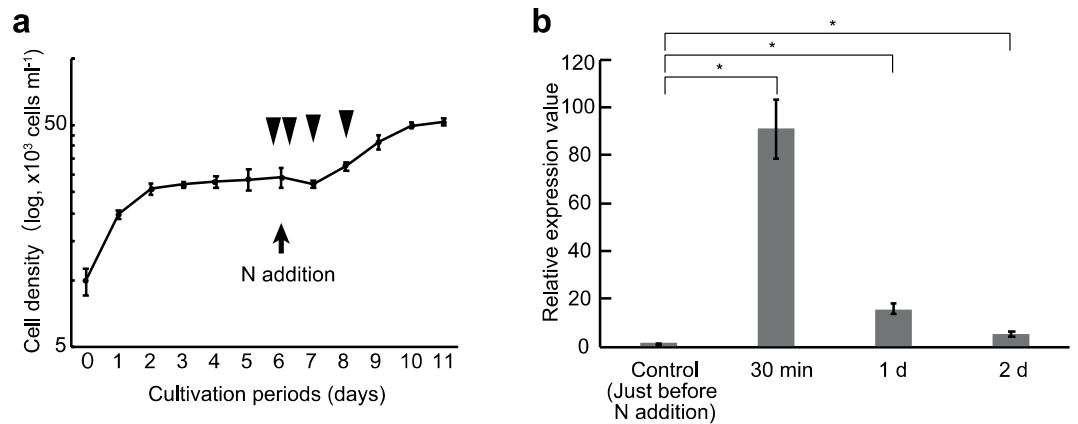
Figure 4. Culture experiments for induction of *NRT2* in *T. nordenskioeldii*. Transcriptional differences of *NRT2* in *T. nordenskioeldii* before/after nitrogen addition under nitrogen-depleted conditions. (**a**) Growth curve of the *T. nordenskioeldii* strain. The strain was cultivated for 6 days, then both $NaNO_3$ and NH_4Cl were added (arrow). The cells were sampled four times, just before the nitrogen addition, and after 30 min, 1 day, and 2 days (arrowheads). Error bars represent standard deviation (SD). (**b**) Relative expression values of *NRT2* in *T. nordenskioeldii*, normalized to the expression value of 18S rRNA. Error bars represent SD. P-values are shown on bars (t-test). Asterisks represent significantly different expression values ($p < 0.01$).
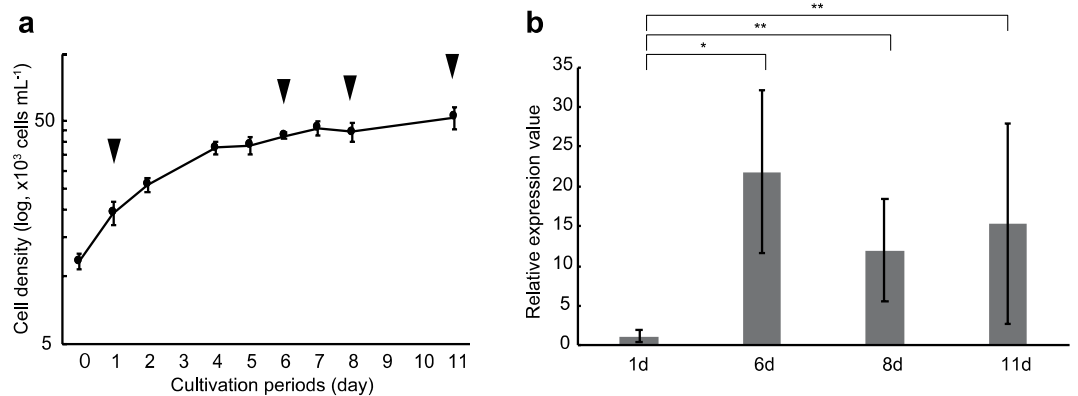
Figure 5. Culture experiments for induction of *AMT2* in *T. nordenskioeldii*. Transcriptional differences of *AMT2* in *T. nordenskioeldii* under nitrogen-repleted conditions. (**a**) Growth curve of the *T. nordenskioeldii* strain. The strain was cultivated for 11 days, and sampled four times, 1, 6, 8, and 11 days after starting cultivation (arrowheads). The sampled cells were washed, transferred into new medium (i.e., nitrogen-repleted 1/10 N K medium), and incubated for 2 h in the dark, and the RNA was extracted. Error bars represent SD. (**b**) Relative expression values of *AMT2* in *T. nordenskioeldii*, normalized to the expression value of 18S rRNA. Error bars represent SD. P-values are shown on bars (t-test). Asterisks represent significantly different expression values ($p < 0.01$). Double asterisks represent non-significant expression values ($p \geq 0.01$).

Similar situation that diatoms had quick response to upwelled nitrate was reported[13]. Diatom cells at the SCM is under initial decline phase of the bloom, probably corresponding to late log phase in the culturing experiments, which can explain the induction of *AMT2* in SCM sample. The transcriptional induction of *NRT2*, *NR*, and *AMT2* can be used as a direct index of development stages of diatom blooms.

Unique transcriptional response of genes for iron transporters in *Chaetoceros debilis*. A mesoscale iron-enrichment experiment in the western subarctic North Pacific (neighbouring region for our experiment) showed that the dominant diatom species changed from *Pseudo-nitzschia turgidula* to *C. debilis* after iron addition[14]. During the experiment, *C. debilis* was more sensitive to iron concentration change than the other diatoms in this region, and had a high growth rate, 2.6 doublings per day, after iron addition. Concentration of dissolved iron initially reached 2.9 nM in a patch and subsequently rapidly decreased. In our analyses, both *C. debilis* and *T. nordenskioeldii* had high expression of *ISIP3*, encoding an iron starvation-induced protein, at the surface (Fig. 6a). *C. debilis* also induced another gene for ISIP (*ISIP1*). ISIP1 and ISIP3 are putative receptors and co-receptors, respectively, which are localized on the cell membrane, and induced in low iron environments[45]. Additionally, *T. nordenskioeldii* and *C. debilis* had high expression of *ISIP2a*, which concentrates iron at the cell surface[46], in the surface layer (Fig. 6a). Therefore, *T. nordenskioeldii* and *C. debilis* are iron-depleted in iron-limited
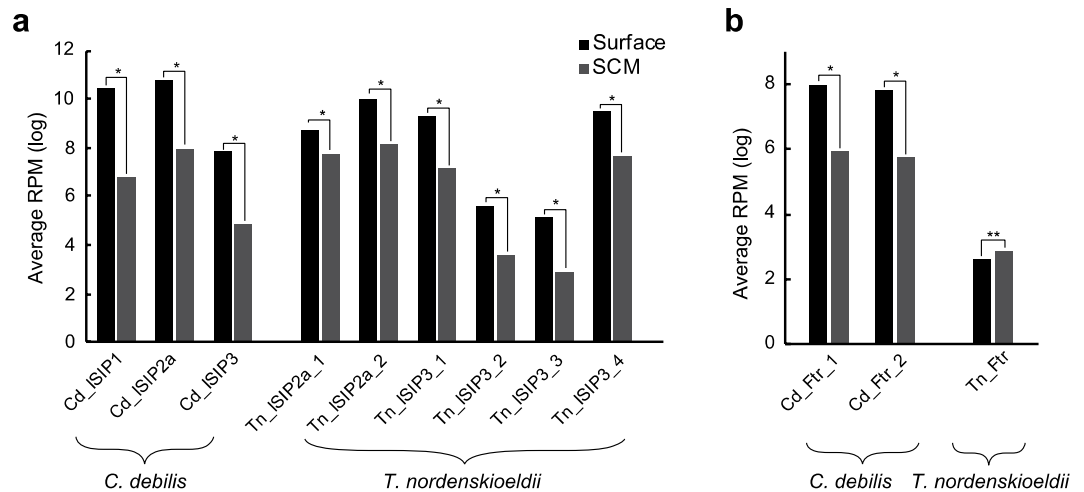
Figure 6. Changes in expression values of genes for sensing and intake of iron. (**a**) Average reads per million (RPM) of genes for iron sensing (*ISIP1* and *ISIP3*) and iron accumulation (*ISIP2a*) in *C. debilis* and *T. nordenskioeldii*. (**b**) Average RPM of genes for iron intake (*Ftr*) in *C. debilis* and *T. nordenskioeldii*. FDRs are shown above bars. Asterisks represent significantly different RPM (FDR < 0.01). Double asterisks represent non-significant RPM (FDR ≥ 0.01).

environments such as the surface of the subarctic Pacific[47]. Interestingly, another iron-related gene, iron permease (*Ftr*), showed a different expression pattern between *T. nordenskioeldii* and *C. debilis* (Fig. 6b): *C. debilis* had significantly higher expression of *Ftr* at the surface than SCM, whereas *T. nordenskioeldii* did not have significantly different expression. These results suggest that *C. debilis*, but not *T. nordenskioeldii*, transcriptionally regulates iron intake systems during spring diatom bloom in this region. Moreover, those results could also explain the high sensitivity of *C. debilis* to the temporal increase in iron concentration in the mesoscale iron-enrichment experiments. *C. debilis* can rapidly grow corresponding to the iron addition because *C. debilis* produces more iron intake systems than the other diatoms (e.g., *T. nordenskioeldii*) in the surface layer.

Suspending strategy at the SCM. Changing the buoyancy is one of the main strategies of diatoms to adapt to marine environments[48]. In the SCM, genes for phospholipid:diacylglycerol acyltransferase (PDAT), which is involved in triacylglycerol synthesis from phospholipid[49], were induced in *T. nordenskioeldii* and *C. debilis* (Supplementary Fig. S9). Lipid accumulation is responsible for the floating strategy of marine and freshwater algae[50–52]. However, the floating strategies are not simple for diatoms; indeed lipid content of a freshwater diatom *Asterionella* is not enough to make it float[53]. Body shape, altered, for example, by chitin fibres, can also affect buoyancy[51,54]. Our data showing high expression of *T. nordenskioeldii* chitin synthase genes in surface samples indicate production of chitin fibres (Supplementary Fig. S10a), which would form long fibres on the cell membrane[55,56]. Thus, *T. nordenskioeldii* might differentially express genes to regulate their buoyancy to adapt to local environments (i.e., surface or SCM) within a spring diatom bloom in the Oyashio region. Interestingly, the gene for plastid ftsZ, which is a plastid division protein[57,58], was induced at the surface (Supplementary Fig. S10b), suggesting frequent plastid division at the surface, which may correspond to light intensity. In contrast, the gene of *C. debilis* was not induced because *C. debilis* has a single plastid per cell[59].

Conclusions

Here, we performed metatranscriptome analyses of the bloom-forming diatoms, focusing differences between two depths, and cultivation experiments to reproduce the expression patterns of genes for nitrogen intake. We reveal the transcriptional complexity of diatom bloom formation. It is thought that suspension and sinking are important for diatoms to grow in a suitable environment, and we indicate that the vertical relocation should affect cell responses of diatoms, and trigger bloom formation. Expression of *AMT2* is likely to be closely related to the bloom-maturing, and thus it can be used for quantitative monitoring of diatom bloom formation. Combining metatranscriptome and strain-based analysis using isolated cultures enables elucidation of population dynamics and cellular responses in various types of algal blooms, such as raphidophyte bloom, cyanobacterial bloom, and *Aureococcus* bloom.

Materials and Methods

Seawater sampling and prediction of development stages of the diatom bloom. Seawater sampling was conducted in the Oyashio region of the western North Pacific (Station A4: 42.25°N, 145.13°E) ~2 h after sunset on May 16, 2016, onboard the R/V *Wakataka Maru*. Seawater samples were collected using the Niskin bottles. 110 L of seawater samples for metatranscriptome were collected at 10 m depth within the surface mixing layer (10 m depth) and at 30 m depth of SCM layer after sunset. Each sample was pre-filtered using a plankton net (100 μm mesh size) and the cells >20 μm were concentrated to 45 mL using a plankton net (20 μm mesh size), then filtered onto a 0.2 μm pore size polycarbonate Nuclepore filter (Whatman, Dassel, Germany) with gentle vacuum pressure (<150 mmHg). The sampling, concentrating, and filtering the cells for metatranscriptome

analysis was completed within 15 min. The filters were immersed in RNAlater (Ambion, Austin, TX, USA) for 15 min at room temperature, and flash frozen in liquid nitrogen, then stored at −80 °C until analysis. Samples for nutrients and chlorophyll *a* concentration analysis were collected from nine layers (0, 10, 20, 30, 40, 50, 60, 80, and 100 m depth). Subsamples of 1 L were immediately fixed with acid Lugol's solution (4% final concentration) and preserved at 4 °C. Vertical profiles of temperature and salinity were measured using a conductivity-temperature-depth profiler. To elucidate development stages of the diatom bloom, an index for the development of spring diatom blooms (I_{devel})[20] was used. For this index, diatom blooms are divided into three stages: pre-bloom ($I_{devel} < 0.15$), growth phase of the bloom ($I_{devel} = 0.15$ to 0.50), and decline phase of the bloom ($I_{devel} > 0.50$).

Identification of major diatom group.   For enumeration of phytoplankton cells, fixed samples were concentrated by reverse filtration through 2 μm nucleopore filters. In the concentrated samples, diatom species were identified following Hasle and Syvertsen[60], and their abundance (cells L^{-1}) was estimated using inverted light microscopy.

RNA extraction, quality control, and sequencing.   RNA was extracted from three independent samples per depth using RNeasy Mini Kit (Qiagen Ltd., Venlo, The Netherlands), with slight modifications. The diatoms on filters with RNAlater were incubated with 700 μL of buffer RLT and 7 μL of 2-mercaptoethanol on ice. Cells were collapsed using a Mini Bead Beater (Cole-Parmer, Vernon Hills, IL, USA) with glass beads (0.05 g of 0.1 mm beads and 0.05 g of 0.5 mm beads, Bertin, Rockville, MD, USA) at 5,000 rpm for 50 s, and the filters were removed after centrifugation at 15,000 rpm for 3 min. Subsequently, the manufacturer's protocol was followed. Genomic DNA was removed using the TURBO DNA-free Kit (Thermo Fisher Scientific, Waltham, MA, USA), and the quality of the RNA was checked using TapeStation (Agilent Technologies, Santa Clara, USA). Genomic DNA was checked by using PCR and details were described in Supporting Information and Supplementary Fig. S11. RNA libraries were constructed using NEBNext Ultra Directional RNA Library Prep Kit for Illumina (New England BioLabs, Ipswich, MA, USA) according to the manufacturer's protocol. Sequencing was performed using the MiSeq System (Illumina Inc., San Diego, CA, USA) with MiSeq Reagent Kits v2 (250 bp × 2) or v3 (300 bp × 2) (Illumina) (Supplementary Table S4). In total, 42,073,468 paired-reads (8,390,580,790 bp) were acquired, with more than 4.9 million paired-reads (1.0 Gbp) per sample.

For reference, RNA-seq analysis of *T. nordenskioeldii* NIES-4227, which was isolated from the Oyashio region of the western North Pacific, was performed. The cells were cultivated in 100 mL of L1 medium[61] for 2 days under white LED light (~20 μmol photons $m^{-2} s^{-1}$) with 14 h:10 h light:dark cycles. Cells were collected by gentle centrifugation during the light and dark phases, and equal amounts of the light and dark RNA samples were mixed. RNA was extracted using the RNeasy Mini Kit (Qiagen) with slight modifications. The cells were incubated with 350 μL of buffer RLT and 3.5 μL of 2-mercaptoethanol on ice, and vortexed to mix. The cells were collapsed using a Mini Bead Beater (Cole-Parmer) with glass beads (0.05 g of 0.5 mm beads, Bertin) at 5,000 rpm for 30 s. Subsequently, the manufacturer's protocol was followed. DNase treatment and RNA library construction were performed as described above. Sequencing was performed using the MiSeq System (Illumina) with MiSeq Reagent Kits v3 (300 bp × 2) (Illumina). In total, 13,440,632 paired-reads (2,939,690,202 bp) were obtained. All of the reads were trimmed using Trimmomatic 0.36 with default options[62]. The RNA-seq reads of the metatranscriptome were deposited in GenBank/ENA/DDBJ with accession numbers DRA006756, DRA008069, and DRA008900. The RNA-seq reads of *T. nordenskioeldii* were deposited in GenBank/ENA/DDBJ with accession number DRA006754.

Assembly of RNA-seq reads for reference of genes.   Four major species of the diatom bloom were selected: *T. nordenskioeldii*, *C. socialis*, *C. debilis*, and *F. cylindrus*. For references, transcripts data of *F. cylindrus* were downloaded from JGI website (https://genome.jgi.doe.gov/Fracy1/Fracy1.home.html)[63], and RNA-seq data of *C. socialis* and *C. debilis* were downloaded from the NCBI SRA database (accession numbers SRR1205835 and SRR1296919, respectively). RNA-seq of *T. nordenskioeldii* was obtained in this study as described above. All of the RNA-seq reads were assembled using Trinity v 2.2.0[64]. Coverages were calculated by RSEM[65] and bowtie2[66]. Isoforms with the highest coverages were extracted using a script in Trinity package as reference genes. Finally, 34,797, 69,789, and 28,069 transcripts for *T. nordenskioeldii*, *C. socialis*, and *C. debilis*, respectively, were acquired.

Mapping and DEG analyses.   RNA-seq reads from the environment samples were mapped to the reference genes using segemehl version 0.3.2[67] with similarity >85%. The hit strategy was best-hit only, which aligns only the most similar read to the reference under similarity criteria. Read count data were split into respective species and used for differentially expressed gene (DEG) analyses. Genes without any mapped reads were removed from the dataset. Read count normalization and DEG analyses were performed with likelihood ratio tests using edgeR 3.18.1[68], and DEGs were defined with false discovery rate (FDR) < 0.01.

Functional analyses.   Functional analyses of DEGs were performed using Kyoto Encyclopedia of Genes and Genomes (KEGG)[69] and eggNOG mapper 1.0.3[70,71]. Genes for nitrogen metabolism were also identified by performing blastp with proteins of *T. pseudonana*[31].

Cultivation experiments for variable gene expression of *AMT2*, *NR* and *NRT2*.   To find factors influencing gene expression of *AMT2*, *NR*, and *NRT2*, *T. nordenskioeldii* NIES-4227 cells were cultivated under various conditions and subjected to qPCR analyses. All cultivations were performed using filter-sterilized 1/10 N K medium or 1/100 N K medium with artificial sea water (Wako, Osaka, Japan or Red Sea, Huston, TX, USA). These are modified K media[72,73] with 1/10 or 1/100 concentrations of nitrogen source (original concentrations in K medium: 882 μM $NaNO_3$ and 50 μM NH_4Cl, respectively). Cell numbers were counted manually using a 4 Grid Cell Counter Plate (Fuchs Rosental model, Watoson Co., Ltd., Tokyo, Japan). Detailed methods for the cultivation and qPCR analyses are described in Supporting Text. All of the cultivation and qPCR analyses had three independent technical and biological replicates.

References

1. Nelson, D. M., Tréguer, P., Brzezinski, M. A., Leynaert, A. & Quéguiner, B. 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 (1995).
2. Mann, D. G. The species concept in diatoms. *Phycologia* **38**, 437–495 (1999).
3. Armbrust, E. V. The life of diatoms in the world's oceans. *Nature* **459**, 185–192 (2009).
4. Tréguer, P. *et al*. Influence of diatom diversity on the ocean biological carbon pump. *Nat. Geosci.* **11**, 27–37 (2018).
5. Chiba, S., Ono, T., Tadokoro, K., Midorikawa, T. & Saino, T. Increased stratification and decreased lower trophic level productivity in the Oyashio Region of the North Pacific: A 30-year retrospective study. *J. Oceanogr.* **60**, 149–162 (2004).
6. Suzuki, K. *et al*. Population dynamics of phytoplankton, heterotrophic bacteria, and viruses during the spring bloom in the western subarctic Pacific. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **58**, 575–589 (2011).
7. Leblanc, K. *et al*. A global diatom database – abundance, biovolume and biomass in the world ocean. *Earth Syst. Sci. Data* **4**, 149–165 (2012).
8. Sverdrup, H. U. On conditions for the vernal blooming of phytoplankton. *ICES J. Mar. Sci.* **18**, 287–295 (1953).
9. Hitchcock, G. L. & Smayda, T. J. The importance of light in the initiation of the 1972–1973 winter-spring diatom bloom in Narragansett Bay. *Limnol. Oceanogr.* **22**, 126–131 (1977).
10. Erga, S. R. & Heimdal, B. R. Ecological studies on the phytoplankton of Korsfjorden, western Norway. The dynamics of a spring bloom seen in relation to hydrographical conditions and light regime. *J. Plankton Res.* **6**, 67–90 (1984).
11. Lancelot, C., Hannon, E., Becquevort, S., Veth, C. & De Baar, H. J. W. Modeling phytoplankton blooms and carbon export production in the Southern Ocean: dominant controls by light and iron in the Atlantic sector in Austral spring 1992. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **47**, 1621–1662 (2000).
12. Margalef, R. Life-forms of phytoplankton as survival alternatives in an unstable environment. *Oceanol. Acta* **1**, 493–509 (1978).
13. Fawcett, S. & Ward, B. Phytoplankton succession and nitrogen utilization during the development of an upwelling bloom. *Mar. Ecol. Prog. Ser.* **428**, 13–31 (2011).
14. Tsuda, A. *et al*. A mesoscale iron enrichment in the western subarctic Pacific induces a large centric diatom bloom. *Science* **300**, 958–61 (2003).
15. Martin, J. H. & Fitzwater, S. E. Iron deficiency limits phytoplankton growth in the north-east Pacific subarctic. *Nature* **331**, 341–343 (1988).
16. Frias-Lopez, J. *et al*. Microbial community gene expression in ocean surface waters. *Proc. Natl. Acad. Sci.* **105**, 3805–3810 (2008).
17. Poretsky, R. S. *et al*. Analysis of microbial gene transcripts in environmental samples. *Appl. Environ. Microbiol.* **71**, 4121–4126 (2005).
18. Marchetti, A. *et al*. Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proc. Natl. Acad. Sci.* **109**, E317–E325 (2012).
19. Alexander, H., Jenkins, B. D., Rynearson, T. A. & Dyhrman, S. T. Metatranscriptome analyses indicate resource partitioning between diatoms in the field. *Proc. Natl. Acad. Sci.* **112**, E2182–E2190 (2015).
20. Yoshie, N., Suzuki, K., Kuwata, A., Nishioka, J. & Saito, H. Temporal and spatial variations in photosynthetic physiology of diatoms during the spring bloom in the western subarctic Pacific. *Mar. Ecol. Prog. Ser.* **399**, 39–52 (2010).
21. Kuroda, H. *et al*. Influence of Coastal Oyashio water on massive spring diatom blooms in the Oyashio area of the North Pacific Ocean. *Prog. Oceanogr.* **175**, 328–344 (2019).
22. Kono, T. & Sato, M. A mixing analysis of surface water in the Oyashio region: Its implications and application to variations of the spring bloom. *Deep. Res. Part II Top. Stud. Oceanogr.* **57**, 1595–1607 (2010).
23. Kasai, H., Saito, H., Yoshimori, A. & Taguchi, S. Variability in timing and magnitude of spring bloom in the Oyashio region, the western subarctic Pacific off Hokkaido, Japan. *Fish. Oceanogr.* **6**, 118–129 (1997).
24. Yoshimori, A. *et al*. Modeling of spring bloom in the western subarctic Pacific (off Japan) with observed vertical density structure. *J. Oceanogr.* **51**, 471–488 (1995).
25. Shiozaki, T. *et al*. Regional variability of factors controlling the onset timing and magnitude of spring algal blooms in the northwestern North Pacific. *J. Geophys. Res. Ocean.* **119**, 253–265 (2014).
26. Saito, H., Tsuda, A. & Kasai, H. Nutrient and plankton dynamics in the Oyashio region of the western subarctic Pacific Ocean. *Deep. Res. Part II Top. Stud. Oceanogr.* **49**, 5463–5486 (2002).
27. Hanawa, K. & Mitsudera, H. Variation of water system distribution in the Sanriku Coastal Area. *J. Oceanogr.* **42**, 435–446 (1986).
28. Suzuki, Y. & Takahashi, M. Growth responses of several diatom species isolated from various environments to temperature. *J. Phycol.* **31**, 880–888 (1995).
29. Degerlund, M., Huseby, S., Zingone, A., Sarno, D. & Landfald, B. Functional diversity in cryptic species of *Chaetoceros socialis* Lauder (Bacillariophyceae). *J. Plankton Res.* **34**, 416–431 (2012).
30. Mock, T. *et al*. Whole-genome expression profiling of the marine diatom *Thalassiosira pseudonana* identifies genes involved in silicon bioprocesses. *Proc. Natl. Acad. Sci.* **105**, 1579–1584 (2008).
31. Armbrust, E. V. *et al*. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* **306**, 79–86 (2004).
32. Anderson, P. M. & Little, R. M. Kinetic properties of cyanase. *Biochemistry* **25**, 1621–1626 (1986).
33. Widner, B., Mulholland, M. R. & Mopper, K. Distribution, sources, and sinks of cyanate in the coastal North Atlantic Ocean. *Environ. Sci. Technol. Lett.* **3**, 297–302 (2016).
34. Pistorius, E. K., Gewitz, H.-S., Voss, H. & Vennesland, B. Reversible inactivation of nitrate reductase in *Chlorella vulgaris in vivo*. *Planta* **128**, 73–80 (1976).
35. Bender, S. J., Durkin, C. A., Berthiaume, C. T., Morales, R. L. & Armbrust, E. V. Transcriptional responses of three model diatoms to nitrate limitation of growth. *Front. Mar. Sci.* **1**, 1–15 (2014).
36. Bender, S. J., Parker, M. S. & Armbrust, E. V. Coupled effects of light and nitrogen source on the urea cycle and nitrogen metabolism over a diel cycle in the marine diatom *Thalassiosira pseudonana*. *Protist* **163**, 232–51 (2012).
37. Allen, A. E., Vardi, A. & Bowler, C. An ecological and evolutionary context for integrated nitrogen metabolism and related signaling pathways in marine diatoms. *Curr. Opin. Plant Biol.* **9**, 264–273 (2006).
38. Lomas, M. W. & Glibert, P. M. Temperature regulation of nitrate uptake: A novel hypothesis about nitrate uptake and reduction in cool-water diatoms. *Limnol. Oceanogr.* **44**, 556–572 (1999).
39. Allen, A. E. *et al*. Evolution and metabolic significance of the urea cycle in photosynthetic diatoms. *Nature* **473**, 203–207 (2011).
40. Slocum, R. D. Genes, enzymes and regulation of arginine biosynthesis in plants. *Plant Physiol. Biochem.* **43**, 729–745 (2005).
41. Clarkson, D. T. & Lüttge, U. Mineral Nutrition: Inducible and Repressible Nutrient Transport Systems. In *Progress in Botany* 61–83, https://doi.org/10.1007/978-3-642-76293-2_5 (Springer Berlin Heidelberg, 1991).
42. Navarro, M. T., Prieto, R., Fernandez, E. & Galvan, A. Constitutive expression of nitrate reductase changes the regulation of nitrate and nitrite transporters in *Chlamydomonas reinhardtii*. *Plant J.* **9**, 819–827 (1996).
43. Crawford, N. M. & Glass, A. D. Molecular and physiological aspects of nitrate uptake in plants. *Trends Plant Sci.* **3**, 389–395 (1998).
44. Daniel-Vedele, F., Filleur, S. & Caboche, M. Nitrate transport: a key step in nitrate assimilation. *Curr. Opin. Plant Biol.* **1**, 235–9 (1998).

45. Lommer, M. *et al.* Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation. *Genome Biol.* **13**, R66 (2012).
46. Morrissey, J. *et al.* A novel protein, ubiquitous in marine phytoplankton, concentrates iron at the cell surface and facilitates uptake. *Curr. Biol.* **25**, 364–371 (2015).
47. Nishioka, J., Ono, T., Saito, H., Sakaoka, K. & Yoshimura, T. Oceanic iron supply mechanisms which support the spring diatom bloom in the Oyashio region, western subarctic Pacific. *J. Geophys. Res.* **116**, C02021 (2011).
48. Kuwata, A. & Jewson, D. H. Ecology and Evolution of Marine Diatoms and Parmales. in *Marine Protists: Diversity and Dynamics* (eds Ohtsuka, S., Suzaki, T., Horiguchi, T., Suzuki, N. & Not, F.) 251–275, https://doi.org/10.1007/978-4-431-55130-0_10 (Springer Japan, 2015).
49. Yoon, K., Han, D., Li, Y., Sommerfeld, M. & Hu, Q. Phospholipid:diacylglycerol acyltransferase is a multifunctional enzyme involved in membrane lipid turnover and degradation while synthesizing triacylglycerol in the unicellular green microalga *Chlamydomonas reinhardtii. Plant Cell* **24**, 3708–3724 (2012).
50. Belcher, J. H. Notes on the physiology of *Botryococcus braunii* Kützing. *Arch. Mikrobiol.* **61**, 335–346 (1968).
51. Reynolds, C. S. *The Ecology of Phytoplankton*, https://doi.org/10.1017/CBO9780511542145 (Cambridge University Press, 2006).
52. Muto, M. *et al.* Potential of water surface-floating microalgae for biodiesel production: Floating-biomass and lipid productivities. *J. Biosci. Bioeng.* **123**, 314–318 (2017).
53. Reynolds, C. S. *The ecology of freshwater phytoplankton* (Cambridge University Press, 1984).
54. Walsby, A. E. & Xypolyta, A. The form resistance of chitan fibres attached to the cells of *Thalassiosira fluviatilis* Hustedt. *Br. Phycol. J.* **12**, 215–223 (1977).
55. McLachlan, J. & Craigie, J. S. Chitan fibres in *Cyclotella cryptica* and growth of *C. cryptica* and *Thalassiosira fluviatilis*. in *Some contemporary studies in marine science* (ed. Barnes, H.) 511–517 (George Allen and Unwin Ltd., 1966).
56. Durkin, C. A., Mock, T. & Armbrust, E. V. Chitin in diatoms and its association with the cell wall. *Eukaryot. Cell* **8**, 1038–1050 (2009).
57. Miyagishima, S. *et al.* Two types of FtsZ proteins in mitochondria and red-lineage chloroplasts: the duplication of FtsZ is implicated in endosymbiosis. *J. Mol. Evol.* **58**, 291–303 (2004).
58. TerBush, A. D., Yoshida, Y. & Osteryoung, K. W. FtsZ in chloroplast division: structure, function and evolution. *Curr. Opin. Cell Biol.* **25**, 461–470 (2013).
59. Kooistra, W. H. C. F. *et al.* Comparative molecular and morphological phylogenetic analyses of taxa in the Chaetocerotaceae (Bacillariophyta). *Phycologia* **49**, 471–500 (2010).
60. Hasle, G. R. & Syvertsen, E. E. Marine Diatoms. in *Identifying Marine Phytoplankton* (ed. Tomas, C. R.) 5–385, https://doi.org/10.1016/B978-012693018-4/50004-5 (Elsevier, 1997).
61. Guillard, R. R. L. & Hargraves, P. E. *Stichochrysis immobilis* is a diatom, not a chrysophyte. *Phycologia* **32**, 234–236 (1993).
62. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
63. Mock, T. *et al.* Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus. Nature* **541**, 536–540 (2017).
64. Haas, B. J. *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).
65. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
66. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
67. Hoffmann, S. *et al.* A multi-split mapping algorithm for circular RNA, splicing, trans-splicing and fusion detection. *Genome Biol.* **15**, R34 (2014).
68. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
69. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40**, D109–114 (2012).
70. Huerta-Cepas, J. *et al.* EGGNOG 4.5: A hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* **44**, D286–D293 (2016).
71. Huerta-Cepas, J. *et al.* Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Mol. Biol. Evol.* **34**, 2115–2122 (2017).
72. Keller, M. D. & Guillard, R. R. L. Factors significant to marine diatom culture. in *Toxic Dinoflagellates* (eds Anderson, D. M., White, A. W. & Baden, D. G.) 113–116 (Elsevier, 1985).
73. Keller, M. D., Seluin, R. C., Claus, W. & Guillard, R. R. L. Media for the culture of oceanic ultraphytoplankton. *J. Phycol.* **23**, 633–638 (1987).

Acknowledgements

Author Contributions

A.K., H.Y. and M.K. conceived the study. T.K., T.W. and A.K. sampled seawater. A.K. isolated the *T. nordenskioeldii* strain. T.K. performed 18S rRNA cloning analysis. T.W. calculated I_{devel}. S.S. performed RNA-seq, assembly, annotation, gene expression analysis, and culture experiments. S.S., T.K. and A.K. wrote the manuscript. All authors contributed in discussing ideas, and read and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-51150-8.

Competing Interests: The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.