
Sexual reproduction during diatom bloom

Prigent Lea ¹, Quéré Julien ¹, Plus Martin ¹, Le Gac Mickael ^{1,*}

¹ Ifremer, DYNECO, F-29280 Plouzané, France

* Corresponding author : Mickael Le Gac, email address : mickael.le.gac@ifremer.fr

Abstract :

Phytoplankton supports food webs in all aquatic ecosystems. Ecological studies highlighted the links between environmental variables and species successions in situ. However, the role of life cycle characteristics on phytoplankton community dynamics remains poorly characterized. In diatoms, sexual reproduction creates new genetic combinations and prevents excessive cell size miniaturisation. It has been extensively studied in vitro but seldom in the natural environment. Here, analyzing metatranscriptomic data in the light of the expression patterns previously characterized in vitro, we identified a synchronized and transient sexual reproduction event during a bloom of the toxic diatom species *Pseudo-nitzschia australis*. Despite the complexity of environmental conditions encountered in situ, sexual reproduction appeared to be the strongest differential gene expression signal that occurred during the bloom. The potential link between environmental conditions and the initiation of sexual reproduction remain to be determined, but sexual reproduction probably had a major impact on the bloom dynamic.

1. Introduction

Phytoplankton plays a major ecological role. It has remarkable morphological and functional diversity and contributes to about half of Earth primary production [1]. Among this group, diatoms represent one of the most diverse and dominant protists [2-3]. When environmental conditions favor high growth rate and low mortality, diatoms can temporarily proliferate to high densities (up to millions of cells per liter) and form a natural phenomenon called bloom [4]. Improving our understanding of algal blooms requires the ability to detect entire bloom events (from the start until the end), to identify the parameters that lead the causative species to transiently dominate the community and eventually be replaced by other species. Numerous studies have been carried out to identify the environmental factors that trigger the initiation, influence the magnitude and lead to bloom termination. Abiotic

factors influencing species occurrence and dynamics include hydrodynamic processes, temperature, salinity and nutrient availability [5], whereas biotic factors include grazing, pathogenicity, parasitism, and competitive interactions within the microbial communities [5-6]. Classically, temporal successions between phytoplankton functional groups is conceptualized using Margalef's mandala that originally took into account turbulence and nutrients [7] and was later expanded by integrating twelve effect or response traits ranging from cell motility, pigment composition, to the ability to produce allelopathic compounds [8]. In dinoflagellates, life cycle characteristics, such as the ability to form resting cysts that may constitute "seed banks" have been considered as an important factor that may explain bloom dynamics across years [9]. In diatoms, life cycle is rarely considered as a factor influencing bloom dynamics.

Diatom life cycles consist in a diploid vegetative phase lasting months to years, during which cells divide mitotically, and in a sexual phase lasting a few days [10-11]. The numerous mitotic divisions are associated with a cell size decrease imposed by the rigid silica frustule composed of two theca with the smaller fitting into the larger. During divisions, each of the two daughter cells retains one maternal theca and a smaller one is synthesized, leading to an average decrease in cell size throughout divisions. This cell size decrease could lead to death if cells reach a critical size incompatible with survival. Sexual reproduction generates bigger cells and is the only way to escape the miniaturization process. It may be initiated when cells reach the sexualisation size threshold (SST) [12-13] and is also density dependent, with sexualisation occurring at high densities [14]. In *P. multistriata*, sexual reproduction has been initiated for cells between 40 and 55 μm , compared to a maximum cell size of ~80 μm [15]. Pennate diatoms are predominantly heterothallic and gamete formation occurs when two strains of opposite mating types (MT+ and MT-) come into contact under stable conditions [16]. Pheromones are involved in mating partner recognition in some species [17-19]. The initiation of sexual reproduction has been associated with a population wide growth arrest, even for cells not directly involved in sexual reproduction [14, 17, 20]. Following the initial contact, meiosis produces gametes that conjugate and produce a zygote. Zygotes expand into non-silicated cells, the auxospores, within which large initial vegetative cells are synthesized [13, 21]. Several studies have investigated

diatom sexual reproduction *in vitro* using molecular approaches [22-27], *P. multistriata* having been used as a model to understand the molecular bases of pennate diatoms sexual reproduction [28]. Five mating related genes have been identified [26]. The MRP3 gene is expressed in MT+ strains and regulates the expression of MRP1 and MRP2 in MT+ strains and MRM1 and MRM2 in MT- strains. During mating type perception, at the initial stage of cell sexualisation, the most upregulated genes in MT- were the endopeptidase Cathepsin D, MRM2, the meiotic gene Rad51A1 and an uncharacterized gene 7488. In MT+, MRP1 was the most overexpressed gene in MT+ during sexualisation [25]. Following the initiation of sexual reproduction, major transcriptional changes occurred in numerous metabolic functions such as photosynthesis, photoprotection, carbon assimilation and fatty acid metabolism [20].

Very few studies have reported observations of diatom sexual events in the natural environment. Sexual stages of *Pseudo-nitzschia* species in their natural environment have been observed and reported twice [29-30]. In Europe, sexual stages of two species of *Pseudo-nitzschia* were detected in the Bay of Naples (Mediterranean Sea), where they accounted for 9.2 and 14.3% of the total number of cells of *P. cf. delicatissima* and *P. cf. calliantha*, respectively [29]. Another massive sex event involving *P. australis* and *P. pungens* had been reported along the Washington coast (Pacific Ocean). Auxospores were detected for about 3 weeks and accounted for up to 59% of the *P. australis* population at the end of the bloom [30]. Sexual reproduction stages have also been observed for *Fragilariopsis kerguelensis* and *Corethron criophilum* in the Southern Ocean [31-32].

Here, we showed that synchronous and transient sexual reproduction is the strongest differential gene expression signal that occurred during a bloom of the toxic diatom *P. australis*, producer of domoic acid, and that it was probably a major determinant of the bloom dynamic. These results were obtained using a metatranscriptomic approach and built upon markers of sexual reproduction as well as global expression patterns identified during *P. multistriata in vitro* sexual reproduction [20].

2. Materials and Methods

2.1 Field sampling and metatranscriptomic samples

From April 4th to April 24th 2017, duplicate samples were collected over 14 days during a *P. australis* bloom (Supplementary table 1). Subsurface water samples were collected from shore at a single site (48.3601°N; - 4.5531°E) in the Bay of Brest (Brittany, North-West of France; Supplementary Figure 1) around high tide (+/- 2h). Samples were filtered (6-15 litres, Supplementary table 1) onto polycarbonate filters (10 µm, 47 mm) using a peristaltic pump before being frozen in liquid nitrogen in RNA later (Fisher Scientific, Illkirch, France) and stored at -80 °C before extraction.

2.2 Environmental data

High tide nitrate, phosphate and silicate surface concentrations (µmol/L, weekly data frequency) were obtained from the SOMLIT monitoring network (<https://somalit.fr>) at the Brest-Portzic station located 150 meters from our sampling station. High frequency (20 min) surface temperature (°C), turbidity (NTU) and chlorophyll fluorescence (Fluorescein Fluorescence Units, FFU) were obtained from the nearby COAST-HF-MAREL-Iroise buoy (48.3579°N, -4.5517°E) [33].

Average photosynthetically active radiations (PAR expressed in µmol photon/m²/s) reaching the sea surface at the sampling station and daily rainfall values (mm) were extracted from the Météo-France AROME numerical model (1.3 km resolution) [34]. Finally, hourly tide levels were provided by the National Hydrographic Service (SHOM, REFMAR dataset, <http://dx.doi.org/10.17183/REFMAR>, Brest tide gauge).

2.3 RNA extraction, library preparation and sequencing

Total RNA was extracted by sonicating filters on ice (Vibra-cell 75115, Bioblock Scientific, Illkirch, France) for 30 seconds at 35% intensity in LBP buffer (Macherey-Nagel, Duren, Germany). Extraction was performed using NucleoSpin® RNA Plus kit (Macherey-Nagel) following the manufacturer's protocol. Library preparation was performed using the Illumina mRNA TruSeq stranded kit starting from 0.5 µg of total RNA. Library preparation failed for two samples. A total of 26 samples were paired-end sequenced using 2 × 150 bp cycles on Illumina Novaseq6000 at the GeT-PlaGe France Genomics sequencing platform (Toulouse, France). To avoid batch effects, samples

were randomized for RNA extraction, library preparation and sequencing. Generated fastQ files have been deposited to ENA (Supplementary table 1).

2.4 Bioinformatic analysis

Prior to read mapping, raw reads quality were assessed using FastQC (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/>), and Trimmomatic (V. 0.33) [35], was used to trim ambiguous, low quality reads and sequencing adapters with parameters ILLUMINACLIP: Adapt.fasta: 2 :30 :10 :8 LEADING: 3 TRAILING: 3 MAXINFO:135:0.8 MINLEN: 80.

2.4.1 Community composition

Community composition was estimated using blastn to identify similarity between the forward trimmed reads and two reference databases. The first database, PR2 (18S rRNA) [36] was used to determine eukaryote relative abundance at the class level. The second one, Diat_barcode (*rbcL*) [37] was used to determine diatom relative abundance at the genus or species level. For each database, reads were considered if the hits displaying the minimum e-value all belonged to the same class (PR2, e-value $\leq 10^{-70}$), genus or species (Diat_barcode, e-value $\leq 10^{-30}$). The relative abundances were obtained by dividing the number of reads associated with each species, genera or class, by the total number of taxonomically assigned reads in each sample.

2.4.2. Gene expressions and functions analysis

- Reference transcriptomes and alignment

Using the BWA-MEM aligner [38], trimmed reads were aligned to a metareference corresponding to the combination of 315 species specific reference transcriptomes, representing 213 unique genera [39]. It mostly corresponded to the resources developed during the Marine Microbial Eukaryotic Transcriptome Sequencing Project (MMETSP) [40] with the addition of reference transcriptomes obtained for three *Pseudo-nitzschia* species (*P. australis*, *P. fraudulenta* and *P. pungens*) based on local strains [41]. Samtools was used to discard reads displaying low quality alignments (MapQ<10),

to remove read pairs that did not align to the same transcript and to generate the raw read count expression matrix [42]. Raw read counts corresponding to *P. australis* contigs were then extracted.

-Gene expression dynamics

After preliminary analyses, the April 24th samples were excluded due to low relative abundances of *P. australis*. Transcripts covered by less than 5 reads on average per day were discarded from the analyses. Overall gene expression variability among samples was explored following rlog transformation using a principal component analysis (PCA) as implemented in Deseq2 [43]. After checking their homogeneity, replicates were pooled by summing the expression matrices.

Pairwise differential expression (DE) was tested between sample groups (see results) using Wald tests (DESeq2) considering a negative binomial generalized linear model with a false discovery rate (FDR) set at $qvalue = 0.05$. DE transcripts were clustered based on expression profiles across samples using negative binomial models as implemented in MBCluster.Seq [44]. Different cluster sizes were visually inspected and clusters with similar profiles were merged. Gene expression is reported as the log₂ fold change (log₂FC) of the expression of a given transcript over the median expression considering all samples. It is calculated as: $\log_2(2^{X_i}/2^{\bar{X}})$ where X_i is the rlog transformation of the number of reads mapping to a given transcript for each sample i and \bar{X} is the rlog transformation of the median number of reads mapping to a given transcript for all samples.

The relationship between gene expression (rlog) and environmental factors (PAR, salinity, temperature, fluorescence, turbidity, tidal amplitude) was analysed using a canonical correlation analysis (CCA) as implemented in the vegan R package [45]. Rainfall and nutrients were excluded from this analysis due to very low rainfall and low frequency nutrient analyses, respectively.

-Functional annotation

The reference transcriptome of *P. australis* was previously annotated [41]. It was based on sequence similarity with the manually curated Uniprot database as inferred using blastx and considering e-value $<10^{-3}$. Contigs were classified in Gene Ontology (GO) categories (<http://geneontology.org/>). Overrepresentation of GO categories in the clusters was tested using Fisher Exact tests followed by

FDR correction with a q-value threshold set at 0.05. Fisher Exact tests odds ratios (ORs) were reported. Only GO categories containing more than five DE transcripts in a given cluster were considered. To take account GO redundancy (i.e. a similar set of genes can be found in different GO categories), GO categories displaying an overlap coefficient $\frac{GO_i \cap GO_j}{\min(GO_i, GO_j)} > 0.8$ (where GO_j is the size of the GO category j) were clustered [41].

Genes belonging to cluster 6 (see results) were further annotated based on homology with proteins of identified function in the NCBI nr database using blastx (e-value $<10^{-3}$). Genes involved in domoic acid production and related to mating type locus were annotated based on homology (tblastn) with *P. multiseriata* *dabA*, *dabB*, *dabC* and *dabD* genes [46] and *P. multistriata* MRM1, MRM2, MRM3, MRP1 and MRP2 genes [26, 41], respectively. Homolog of the silicon transporter *sit1* was also manually annotated [47]. Annotations are reported in Supplementary table 2.

Genes differentially expressed during *P. multistriata* crossing experiments were obtained from [20] (datafile_S1). These expression levels were obtained 1 hour after parental cells of opposite MT had been mixed (T1), 24 hours after initial mixing (T2) and five days after initial mixing (T3). At T2, about 4% of parental cells had formed gametes. At T3, gametes, zygotes, auxospores, and F1 cells represented 7% of the total population. Homology with *P. australis* genes was identified using tblastx (e-value $<10^{-10}$). The association between genes belonging to the *P. australis* gene clusters displaying differential gene expression *in situ* (see Results) and *P. multistriata* genes displaying differential expression at T1, T2 and T3 during crossing experiments was tested using Fisher Exact tests.

3 Results

3.1 Community composition

At the beginning of April 2017, a strong increase in chlorophyll fluorescence indicated the first massive phytoplankton development at the sampled station (Figure 1A; Supplementary figure 2). This bloom lasted until the end of the month. The eukaryote community relative abundances at the Class level indicated that diatoms (*Bacillariophyta*) strongly dominated the community throughout the bloom, with relative abundances ranging from 38% to 81% (Figure 1B). *Dinophyceae*,

Prymnesiophyceae and *Spirotrichea* were present but below 14%. *Arthropoda*, mainly copepods, appeared at the end of the bloom, reaching a maximum relative abundance of 20% on April 19th (Figure 1B; Supplementary figure 3). The species dominating the bloom was *P. australis*. It represented ~20% of the diatoms at the beginning of the survey, increased to more than 60% between April 7th and 14th and decreased afterward (Figure 1C). At the beginning of the bloom, the accompanying species mostly belonged to the genus *Cerataulina*, *Chaetoceros* and *Thalassiosira*, while at the end, relative abundances of *Chaetoceros*, *Guinardia* and *Rhizosolenia*, but also of *P. hasleana* increased.

3.2 Environmental conditions

During the sampled period, the environment reflected typical spring transitions with a regular decrease in nutrient concentrations (from 7.2 $\mu\text{mol/L}$ to 0.7 $\mu\text{mol/L}$ for nitrates; from 0.2 $\mu\text{mol/L}$ to 0.03 $\mu\text{mol/L}$ for phosphates and from 2.84 $\mu\text{mol/L}$ to 1.23 $\mu\text{mol/L}$ for silicates) associated with a slight increase in salinity (Figure 2). Such a trend resulted from a decrease in riverine flow due to low rain throughout the period. The weather was mostly sunny, and resulted in a regular increase in sea water temperature, from an average of 11.5°C on April 4th to 12.5°C on April 24th. In parallel, the average PAR values showed variability, decreasing for some days over the second week of monitoring with an average of 293 $\mu\text{mol photons/m}^2/\text{s}$. Sea water temperature was also affected by the night-day cycle, with an increase in temperature during the day. Salinity was affected by the semi-diurnal tidal variation between low and high tide, with low salinity at low tide. Finally, temperature and salinity were affected by changes in tidal amplitude, with higher diurnal temperature and semi-diurnal salinity fluctuations when tidal amplitudes were minimal (April 4th to 6th and 18th to 22^d; Figure 2).

3.3 Gene expression dynamics change over the bloom period

To determine gene expression dynamics of *P. australis*, meta-transcriptomic reads aligning specifically on the *P. australis* reference transcriptome were extracted. To visualize gene expression

variability between samples, a PCA was performed, using the 25,140 *P. australis* expressed genes (Supplementary table 3). The two first PCA components explained 54

% of the gene expression variability (Figure 3A). Three groups of samples displaying similar gene expression were identified. They correspond to samples from April 4th to 7th, April 10th to 14th and April 18th to 21st. These three groups of samples, hereafter named early, middle and late groups, respectively, were obtained in slightly different environmental conditions (Figure 3B). The early group was associated with slightly lower temperature, salinity and chlorophyll fluorescence. The middle group of samples was associated with higher tidal amplitudes and chlorophyll fluorescence. The late group was associated with higher temperature, PAR and turbidity.

Differential gene expression was tested between these three groups. A total of 8,252 genes were identified as differentially expressed between the three groups of samples. These genes were grouped into 6 clusters based on gene expression changes across samples (Figure 4). The clusters contained between 75 (cluster 6) and 2,591 (cluster 2) genes. Overall, these gene clusters displayed pronounced gene expression differences between the early, middle and late groups of samples, with major shifts in gene expression occurring on April 10th and 18th.

The over-representation of gene functions was analyzed in the 6 clusters. At the beginning of the bloom (from April 4th to April 10th), numerous genes involved in photosynthesis were over-expressed (Figure 4, cluster 3, Supplementary table 4). In addition, although decreasing during this first phase, there was also an over-expression of genes involved in energy production and storage (Figure 4, cluster 4, Supplementary table 4), as well as in pigment and fatty acid biosynthesis (cluster 4, Supplementary table 4). During the second phase (from April 10th to 14th), genes involved in ribosomal biogenesis and protein synthesis as well as in chromosomal organization were over-expressed (Figure 4, cluster 1, Supplementary table 4). During the late phase (from April 14th to 21st), expression slightly increased for genes related to photosynthesis (Figure 4, cluster 3, Supplementary table 4) and a more pronounced increase was detected for genes related to energy production and storage, as well as involved in the biosynthesis of pigments and fatty acids (Figure 4, cluster 4, Supplementary table 4). There was a very strong overrepresentation of genes involved in "phosphate-containing compound metabolic process", which may indicate tight regulation of phosphate utilization

during the late phase. It was also characterized by an under-representation of various functions related to the cellular redox balance (Figure 4, cluster 5, Supplementary table 4), suggesting a perturbation of the cellular redox balance towards the end of the bloom.

Throughout the entire bloom, there was a moderate and rather constant gene expression increase for several processes related to excretion and more specifically to protein excretion (Figure 4, cluster 2; Supplementary table 4). Associated with this, there was a strong overrepresentation of genes involved in the biosynthesis of dimethylallyl diphosphate, an important precursor of isoprenoids synthesis, including carotenoids. It is also important to note that the dimethylallyl diphosphate is a domoic acid precursor.

3.4 Sexualisation of *P. australis* cells

Between the early and middle phases, genes belonging to cluster 6 displayed a transient high over-expression. In this cluster, the only over-represented GO category is “peptidase activity” which may indicate a metabolic shift. In order to understand the major changes that occurred on April 10th, the annotation of individual genes belonging to cluster 6 were analyzed.

Out of the 75 genes belonging to cluster 6, 28 could not be annotated (Supplementary table 2). The two annotated genes displaying the highest over-expression levels on April 10th were homologous to the “mating type related plus 1” gene (MRP1, log₂FC of 8.1 and 7.9; Figure 5) identified in *P. multistriata*. This gene had been shown to be the most over-expressed gene in MT+ cells during sexualization [25]. Moreover, homologs of three out of the four most DE genes during MT-sexualization were also in the cluster 6. It was the case of two Rad51 homologs (log₂FC of 2.7 and 2.3; Figure 5), two cathepsin D (log₂FC of 4.8 and 3.7; Figure 5) homologs and one homolog of *P. multistriata* protein 7488 (log₂FC of 3.6; Figure 5). The homolog of the fourth gene displaying high DE during MT- sexualisation, MRM2, did not belong to cluster 6. However, it also displayed a transient increase in expression on April 10th (log₂FC=2). Another mating type related gene, MRP2, also belonged to cluster 6 (Log₂FC=2.1 on April 10th). Of the two remaining mating type related genes MRP3 displayed a rather stable expression level throughout the bloom and MRM1 was not expressed.

3.5 Expression of genes related to sexual reproduction during the bloom

The pattern of expression presented above suggested the initiation of a sexual reproduction event on April 10th. In order to test this hypothesis, we compared the expression of *P. australis in situ*, to the gene expression of the closely related species *P. multistriata* characterised *in vitro* during sexual reproduction. During *in vitro* experiments in *P. multistriata*, from initial contact to F1 generation, sexual reproduction lasted several days. For genes belonging to cluster 6 the odds of having homologs up-regulated one hour after initial parental cells contact during *P. multistriata in vitro* experiment were 24 times higher than randomly expected (two-sided Fisher exact test, Odd-Ratio=24.4, P-value < 2.2e⁻¹⁶; 24 up-regulated homologs out of 38 belonging to cluster 6; Figure 6; Supplementary table 5). This clearly added an argument in favor of *P. australis* sexual reproduction initiation *in situ* around April 10th. To determine whether the abrupt changes in gene expression observed following April 10th could be the result of sexual reproduction, genes belonging to cluster 3 and 4 (i.e. genes down-regulated during the second phase of the bloom) and to cluster 1 (i.e. genes up-regulated during the second phase of the bloom) were compared to *P. multistriata* genes displaying down and up regulation during *in vitro* sexual reproduction, respectively. For genes belonging to cluster 3 or 4, odds of having homologs down-regulated 1 hour, 5 hours and 5 days after initial parental cells contact during *P. multistriata in vitro* experiment were 7, 8 and 7 times higher than randomly expected, respectively (two-sided Fisher exact test, Odd-Ratio=6.96, 8.23, 7.37, P-values < 2.2e⁻¹⁶; 280/791, 400/791 and 304/791; Figure 6; Supplementary table 5). Moreover, the genes most strongly down-regulated in clusters 3 and 4 also appear to be strongly down-regulated during *in vitro* experiments (Figure 6; Supplementary table 5). Nevertheless, there were 208 genes belonging to clusters 3 and 4 that tended to be up-regulated during *in vitro* experiments (average log₂FC>0). Functions related to ribosome, proteolysis, one-carbon metabolic process and phagocytosis were overrepresented in these 208 genes compared to all the genes belonging to clusters 3 and 4 (Supplementary table 6).

To a lesser extent, genes up-regulated during the middle bloom period also tend to be up-regulated from a few hours to a few days after the initiation of *P. multistriata* sexual reproduction *in vitro*. For genes belonging to cluster 1, odds of having homologs up-regulated 1 hour, 5 hours and 5 days after

initial parental cells contact during *P. multistriata* *in vitro* experiment were 1.3, 2.5 and 2.5 times higher than randomly expected, respectively (two-sided Fisher exact test, Odd-Ratio=1.34, 2.48, 2.49; P-value = 0.002, < 2.2e-16, < 2.2e-16; 153/1804, 520/1804 and 401/1804; Figure 6; Supplementary table 5). Despite this strong correspondence, 604 genes tend to be up-regulated during the middle phase *in situ* (cluster 1) but down-regulated during *in vitro* experiments (average $\log_2FC < 0$). Highlighting differences between sexual reproduction *in vitro* and *in situ*, functions related to cell division, mitosis and chromosome condensations (Supplementary table 7) were strongly overrepresented in these 604 genes compared to all the genes belonging to cluster 1.

3.6 Expression of toxins biosynthesis genes

The four genes (*dabABCD*) identified as involved in the biosynthesis of DA in *P. multiseriis* [46] displayed in our dataset a drop in expression during the early phase, a stable low expression during the middle phase followed by a further increase during the late phase (Figure 7). Among these genes, two (*dabA* and *dabB*) showed a significant difference in expression during the survey and belonged to cluster 4.

4 Discussion

Results presented above show strong evidence of synchronous and transient sexual reproduction during a bloom of the toxic diatom *P. australis*.

First, a transient over-expression of a few tens of genes highlighted the sexualisation of the *P. australis* vegetative cells. Indeed, homologs of the five most upregulated genes during *in vitro* sexualisation of *P. multistriata* displayed high expression levels at a single sampling date during our survey. This was true for the mating type plus related gene MRP1, the most overexpressed gene in MT+ during the sexualisation as well as for the mating type minus related gene MRM2, the endopeptidase Cathepsin D, the meiotic gene Rad51 and the uncharacterized gene 7488, which are the most overexpressed genes in MT- during the *P. multistriata* sexualisation [25]. More generally,

besides these marker genes, the transiently up-regulated genes in *P. australis* corresponded to the genes upregulated one hour after opposite mating type contact during *P. multistriata* experiments [20]. This transient signal, peaking on a single day and quickly decreasing over the next couple of days, was a clear indication that the initiation of sexual reproduction was synchronized at the population scale. It doesn't mean that all cells entered into sexual reproduction. Even during sexual reproduction experiments, only a portion of the parental cells, ranging from 10 to 30%, reproduced sexually simultaneously [14, 20]. However, the extremely strong up-regulation (from 0 reads up to several thousands of reads) of the genes mentioned above, indicated that the proportion of cells simultaneously entering sexual reproduction was sufficient to induce a population wide signal. *In situ* reports of sexual reproduction in *Pseudo-nitzschia* species are scarce, but are in line with *in vitro* observations. Sexual stages detected in the Bay of Naples (Mediterranean Sea) accounted for 9 and 14% of the total number of cells of *P. cf. delicatissima* and *P. cf. calliantha*, respectively [29]. Another massive sex event, involving *P. australis* and *P. pungens* had been reported along the pacific coast of North America with auxospores accounting up to 59% of the *P. australis* population at the end of the bloom [28]. These varying time scales mean that studying this process in its natural environment is very difficult [48]. During experiments, sexual reproduction was induced by mixing MT+ and MT- strains under non-limiting conditions [16, 20, 21, 25,26], with higher rates of sexual reproduction occurring when cells were in exponential phase [14]. *In situ*, cells of opposite mating types were co-occurring before sexual reproduction initiation, so one may wonder what was the signal triggering synchronous sexualisation. *In vitro* studies have demonstrated that sexual initiation depends upon the cell sizes. In diatoms, cell size tends to decrease through the successive mitotic divisions spanning several months to several years until it reaches the sexualisation size threshold, a cell size below which sexual reproduction may be induced [12-13]. Based on cells size patterns along a multi annual *P. multistriata* time series, it has been proposed that sexual reproduction in the field could occur every three years [49]. Given a sexual event lasting for a few days, this may explain why reports of sexual reproduction in the field remain extremely rare. Sexualisation also depends upon cell densities and high cell density is required to favor cells physical proximity and sex initiation [14-15]. This is in agreement with the results of the present study, where sexualisation occurred when *P.*

australis cell densities were maximum. During the study period, changes in abiotic environmental conditions were recorded, but not specific change could be identified as triggering sexualization.

Second, following this transient expression pattern, a massive shift in *P. australis* gene expression was recorded *in situ*. This shift lasted for a few days and the observed expression pattern corresponded to the one observed from a few hours to a few days after opposite mating type initial contact during *P. multistriata* experiments, when gametes, zygotes and auxospores were observed [20]. Gene expression similarity was especially strong for downregulated genes, with low expression of genes involved in photosynthesis, glycolysis, phosphate metabolism, fatty acid and carotenoid biosynthesis. This could indicate that down regulation of these genes may be the hallmark of sexual reproduction in *Pseudo-nitzschia*, independently of environmental conditions. For instance, prior to sexual reproduction, *P. multistriata* cells tend to accumulate lipid droplets to store energy and potentially transfer it to F1 cells [14, 20, 22]. Moreover, down-regulation of photosynthesis related genes was described as a method to decrease the input of photochemical energy in a way to protect cell health during the sexual phase [20]. However, during the *P. australis* bloom, PAR was higher during the late bloom phase but also to a lesser extent during the early phase, coinciding with higher expression of photosynthesis related genes. In the same vein, light intensity was high enough during these two phases to be considered as potentially stressful and could have triggered the biosynthesis of photoprotection pigments such as carotenoids [50]. As a result, it is difficult to determine whether differential expression of photosynthesis and carotenoid biosynthesis genes was caused by changes in light conditions, sexual reproduction or both. Of specific interest was the down-regulation of *dab* genes, suggesting lower domoic acid production during sexual reproduction [46]. A significant overlap was also detected between genes up-regulated *in vitro* and *in situ*. However the overlap was less strong than for down-regulated genes, which might suggest that, overall, up-regulated genes are less specific to sexual reproduction, and perhaps more dependent upon environmental conditions, than down-regulated ones. Several *in vitro* studies showed that a population wide growth arrest, lasting a few days, followed the initiation of sexual reproduction in *P. multistriata*, even for cells not directly involved in sexual reproduction [14, 20, 21]. This growth arrest has also been identified *in vitro* in

other pennate diatoms [51]. *In situ*, observed cell densities are the result of a balance between cell division, mortality and hydrodynamic dilution. It is therefore difficult to determine whether a growth arrest took place during the *P. australis* bloom. However, during the sexual reproduction event, both chlorophyll fluorescence (a proxy of primary producers biomass) and *P. australis* relative abundances remained high. In the study zone, tide is by far the strongest hydrodynamic force [52]. As a result, tidal amplitude may be considered as a good proxy of bloom dilution. During the sexual reproduction event, tidal amplitude, and as a result hydrodynamic dilution, were high. The maintenance of high levels of chlorophyll fluorescence and *P. australis* relative abundances was in favour of a maintenance of *P. australis* active growth. Of special interest was the fact that, among the genes displaying high expression during the sexual reproduction event *in situ*, but not *in vitro*, there was a strong over-representation of genes associated with cell division, mitosis and chromosome condensations. This is a further indication that, contrary to what has been observed *in vitro*, active growth may be maintained during sexual reproduction *in situ* [53-54].

After the end of the sexual reproduction event, *P. australis* abundance quickly dropped before virtually disappearing in a matter of a few days. This occurred despite a strong decrease in tidal amplitude, and thus of hydrodynamic dilution, and could indicate a *P. australis* growth arrest that resulted in bloom termination. *In vitro*, *P. multistriata* growth rate tends to increase again after the sexual reproduction growth arrest [20], nonetheless in the few events characterized during environmental blooms, *Pseudo-nitzschia* cell densities quickly dropped following sexual reproduction events [29-30]. In the present case, a transient increase in copepod relative abundance occurred the day after *P. australis* abundances started to drop. This may have accelerated bloom termination through grazing, but, because of this time lag, was unlikely to be the main origin of bloom termination. Beside slight changes in environmental conditions, no specific factor could be identified as explaining the bloom demise. However, during this last bloom phase, decreased expression of genes involved in the cellular redox balance and more specifically related to glutathione was identified. Glutathione being one of the most abundant antioxidants in cells, this expression pattern may be indicative of an oxidative stress, resulting from an imbalance between reactive oxygen species

generation and antioxidant capacity [55]. Such a stress could lead to cell senescence [56-57] that could explain the termination of the bloom. In the absence of any straightforward biotic or abiotic environmental change, it is difficult to determine whether the triggering mechanism could be environmental, or the end of the sexual reproduction event, with cell senescence of the remaining small cells that did not reproduce sexually.

Conclusion

This study identified a transient and synchronous sexual reproduction event in the toxic diatom *P. australis in situ*. Gene expression patterns associated with sexual reproduction of a closely related species had been extensively studied *in vitro*. Thanks to these studies, it was possible to use gene expression patterns as markers of sexual reproduction. Based on the expression of a handful of marker genes as well as on global expression patterns, it appeared that vegetative cells sexualisation occurred on a given day and that sexual reproduction lasted for less than a week. This corresponded the timing observed *in vitro*. The relationship between environmental condition and the initiation of sexual reproduction remained to be determined but it was probably initiated following high cell densities with cell sizes below the sexualisation size threshold. Despite the complexity of biotic and abiotic environmental conditions encountered *in situ* and the genetic heterogeneity of *P. australis* populations, sexual reproduction appeared to be the strongest differential gene expression signal that occurred during the bloom. As a result, it was probably a major determinant of bloom dynamics, with potential impact on growth rate and bloom's demise. The dynamic of a given species in microbial communities is mostly thought to be determined by the abiotic conditions acting on cell divisions as well as by biotic interactions that may modulate resources accesses and or mortality rates. The present study highlighted that life cycle events, such as sexual reproduction, may be a key parameter determining microbial community dynamics *in situ*.

Acknowledgment

This work is part of Lea Prigent PhD project and received financial support from the research project “MaSCoET” (Maintien du Stock de Coquillage en lien avec la problématique des Efflorescences Toxiques) financed by France Filière Pêche and the Brittany Region. The authors would like to thank Mariella Ferrante and Gust Bilcke for discussion. Sequencing has been done at the GeT-PlaGe GenoToul sequencing platform (Toulouse, France). We thank the Ifremer RIC informatic and sebimer bio-informatic teams for support.

Data Availability Statement

Raw reads have been deposited to SRA with accession numbers indicated in Supplementary table 1. Read counts and gene annotations are available in Supplementary tables 2 and 3.

Conflict of interest statement

The authors declare no conflict of interest

Study funding

This work received financial support from the research project “MaSCoET” (Maintien du Stock de Coquillage en lien avec la problématique des Efflorescences Toxiques) financed by France Filière Pêche and the French Brittany Region.

References

1. Field, C.B., Behrenfeld, M.J., Randerson, J.T., Falkowski, P. 1998. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science* 281, 237–240. <https://doi.org/10.1126/science.281.5374.237>

2. Armbrust, E.V. 2009. The life of diatoms in the world's oceans. *Nature* 459, 185–192. <https://doi.org/10.1038/nature08057>
3. Kooistra, W.H.C.F., Gersonde, R., Medlin, L.K., Mann, D.G. 2007. The Origin and Evolution of the Diatoms: Their Adaptation to a Planktonic Existence, in: *Evolution of Primary Producers in the Sea*. Elsevier, pp. 207–249. <https://doi.org/10.1016/B978-012370518-1/50012-6>
4. Ji, N., Lin, L., Li, L., Yu, L., Zhang, Y., Luo, H. et al., 2018. Metatranscriptome analysis reveals environmental and diel regulation of a *Heterosigma akashiwo* (*raphidophyceae*) bloom: Metatranscriptome profiling of *Heterosigma akashiwo*. *Environ. Microbiol.* 20, 1078–1094. <https://doi.org/10.1111/1462-2920.14045>
5. Zhou, J., Richlen, M.L., Sehein, T.R., Kulis, D.M., Anderson, D.M., Cai, Z. 2018. Microbial Community Structure and Associations During a Marine Dinoflagellate Bloom. *Front. Microbiol.* 9, 1201. <https://doi.org/10.3389/fmicb.2018.01201>
6. Carnicer, O., Guallar, C., Andree, K.B., Diogène, J., Fernández-Tejedor, M., 2015. *Ostreopsis cf. ovata* dynamics in the NW Mediterranean Sea in relation to biotic and abiotic factors. *Environ. Res.* 143, 89–99. <https://doi.org/10.1016/j.envres.2015.08.023>
7. Glibert, P.M. 2016. Margalef revisited: A new phytoplankton mandala incorporating twelve dimensions, including nutritional physiology. *Harmful Algae* 55, 25–30. <https://doi.org/10.1016/j.hal.2016.01.008>
8. Margalef, R., Estrada M., Blasco D. 1979. Functional morphology of organisms involved in red tides, as adapted to decaying turbulence. D. Taylor, H. Seliger (Eds.), *Toxic Dinoflagellate Blooms*, Elsevier, New York, pp. 89-94
9. Azanza, R.V., Brosnahan, M.L., Anderson, D.M., Hense, I., Montresor, M., 2018. The Role of Life Cycle Characteristics in Harmful Algal Bloom Dynamics, in: Glibert, P.M., Berdalet, E., Burford,

M.A., Pitcher, G.C., Zhou, M. (Eds.), Global Ecology and Oceanography of Harmful Algal Blooms. Springer International Publishing, Cham, pp. 133–161. https://doi.org/10.1007/978-3-319-70069-4_8

10. Chepurnov, V. A., Mann, D. G., von Dassow, P., Vanormelingen, P., Gillard, J., Inzé, D., et al. 2008. In search of new tractable diatoms for experimental biology. *BioEssays* 30, 692-702

11. Bilcke, G., Ferrante, M. I., Montresor, M., De Decker, S., De Veylder, L. & Vyverman, W. 2022. Life Cycle Regulation. In: Falciatore, A. & Mock, T. (Eds.) *The Molecular Life of Diatoms*. Springer, pp. 205-28

12. von Dassow, P., Montresor, M., 2011. Unveiling the mysteries of phytoplankton life cycles: patterns and opportunities behind complexity. *J. Plankton Res.* 33, 3–12. <https://doi.org/10.1093/plankt/fbq137>

13. Montresor, M., Vitale, L., D’Alelio, D., Ferrante, M.I., 2016. Sex in marine planktonic diatoms: insights and challenges. *Perspect. Phycol.* 3, 61–75. <https://doi.org/10.1127/pip/2016/0045>

14. Scalco, E., Stec, K., Iudicone, D., Ferrante, M.I., Montresor, M., 2014. The dynamics of sexual phase in the marine diatom *Pseudo-nitzschia multistriata* (*Bacillariophyceae*). *J. Phycol.* 50, 817–828. <https://doi.org/10.1111/jpy.12225>

15. D’Alelio, D., Amato, A., Luedeking, A., Montresor, M., 2009. Sexual and vegetative phases in the planktonic diatom *Pseudo-nitzschia multistriata*. *Harmful Algae* 8, 225–232. <https://doi.org/10.1016/j.hal.2008.05.004>

16. Kim, J.H., Ajani, P., Murray, S.A., Kim, J.-H., Lim, H.C., Teng, S.T. et al. 2020. Sexual reproduction and genetic polymorphism within the cosmopolitan marine diatom *Pseudo-nitzschia pungens*. *Sci. Rep.* 10, 10653. <https://doi.org/10.1038/s41598-020-67547-9>

17. Moeys, S., Frenkel, J., Lembke, C., Gillard, J.T.F., Devos, V., Van Den Berge, K. et al. 2016. A sex-inducing pheromone triggers cell cycle arrest and mate attraction in the diatom *Seminavis robusta*. *Sci. Rep.* 6, 19252. <https://doi.org/10.1038/srep19252>

18. Fiorini, F., Borgonuovo, C., Ferrante, M.I., Brönstrup, M., 2020. A Metabolomics Exploration of the Sexual Phase in the Marine Diatom *Pseudo-nitzschia multistriata*. *Mar. Drugs* 18, 313. <https://doi.org/10.3390/md18060313>
19. Gillard, J., Frenkel, J., Devos, V., Sabbe, K., Paul, C., Rempt, M. et al. 2013. Metabolomics Enables the Structure Elucidation of a Diatom Sex Pheromone. *Angew. Chem. Int. Ed.* 52, 854–857. <https://doi.org/10.1002/anie.201208175>
20. Annunziata, R., Mele, B.H., Marotta, P., Volpe, M., Entrambasaguas, L., Mager, S., et al. 2022. Trade-off between sex and growth in diatoms: Molecular mechanisms and demographic implications. *Sci. Adv.* 8, eabj9466. <https://doi.org/10.1126/sciadv.abj9466>
21. Chepurnov, V.A., Mann, D.G., Sabbe, K., Vyverman, W., 2004. Experimental Studies on Sexual Reproduction in Diatoms, in: *International Review of Cytology*. Elsevier, pp. 91–154. [https://doi.org/10.1016/S0074-7696\(04\)37003-8](https://doi.org/10.1016/S0074-7696(04)37003-8)
22. Basu, S., Patil, S., Mapleson, D., Russo, M.T., Vitale, L., Fevola, C. et al. 2017. Finding a partner in the ocean: molecular and evolutionary bases of the response to sexual cues in a planktonic diatom. *New Phytol.* 215, 140–156. <https://doi.org/10.1111/nph.14557>
23. Bilcke, G., Van den Berge, K., De Decker, S., Bonneure, E., Poulsen, N., Bulankova, P. et al. 2021. Mating type specific transcriptomic response to sex inducing pheromone in the pennate diatom *Seminavis robusta*. *ISME J.* 15, 562–576. <https://doi.org/10.1038/s41396-020-00797-7>
24. Ferrante, M.I., Entrambasaguas, L., Johansson, M., Töpel, M., Kremp, A., Montresor, M. et al. 2019. Exploring Molecular Signs of Sex in the Marine Diatom *Skeletonema marinoi*. *Genes* 10, 494. <https://doi.org/10.3390/genes10070494>
25. Marotta, P., Borgonuovo, C., Santin, A., Russo, M.T., Manfellotto, F., Montresor, M. et al. 2022. Mate Perception and Gene Networks Regulating the Early Phase of Sex in *Pseudo-nitzschia multistriata*. *J. Mar. Sci. Eng.* 10, 1941. <https://doi.org/10.3390/jmse10121941>

26. Russo, M.T., Vitale, L., Entrambasaguas, L., Anestis, K., Fattorini, N., Romano, F. et al. 2018. MRP3 is a sex determining gene in the diatom *Pseudo-nitzschia multistriata*. Nat. Commun. 9, 5050. <https://doi.org/10.1038/s41467-018-07496-0>
27. Russo, M.T., Ruggiero, M.V., Manfellotto, F., Scriven, V., Campbell, L., Montresor, M. et al. 2021. New alleles in the mating type determination region of West Atlantic strains of *Pseudo-nitzschia multistriata*. Harmful Algae 103, 101995. <https://doi.org/10.1016/j.hal.2021.101995>
28. Ferrante, M.I., Broccoli, A., Montresor, M., 2023. The pennate diatom *Pseudo-nitzschia multistriata* as a model for diatom life cycles, from the laboratory to the sea. J. Phycol. 59, 637–643. <https://doi.org/10.1111/jpy.13342>
29. Sarno, D., Zingone, A., Montresor, M., 2010. A massive and simultaneous sex event of two *Pseudo-nitzschia* species. Deep Sea Res. Part II Top. Stud. Oceanogr. 57, 248–255. <https://doi.org/10.1016/j.dsr2.2009.09.012>
30. Holtermann, K.E., Bates, S.S., Trainer, V.L., Odell, A., Armbrust, E.V., 2010. Mass Sexual Reproduction in the Toxigenic Diatoms *Pseudo-Nitzschia australis* and *P. pungens* (*Bacillariophyceae*) on the Washington Coast, Usa1. J. Phycol. 46, 41–52. <https://doi.org/10.1111/j.1529-8817.2009.00792.x>
31. Assmy, P., Henjes, J., Smetacek, V., Montresor, M., 2006. AUXOSPORE FORMATION BY THE SILICA-SINKING, OCEANIC DIATOM *FRAGILARIOPSIS KERGUELENSIS* (*BACILLARIOPHYCEAE*). J. Phycol. 42, 1002–1006. <https://doi.org/10.1111/j.1529-8817.2006.00260.x>
32. Crawford, R.M., 1995. The role of sex in the sedimentation of a marine diatom bloom. Limnol. Oceanogr. 40, 200–204. <https://doi.org/10.4319/lo.1995.40.1.0200>
33. Rimmelmaury, P., Charria, G., Repecaud, M., Quemener, L., Beaumont, L., Guillot, A. et al. 2023. COAST-HF-Marel-Iroise buoy's time series (French Research Infrastructure ILICO): long-

term high-frequency monitoring of the Bay of Brest and Iroise sea hydrology.
<https://doi.org/10.17882/74004>

34. Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F. et al. 2011. The AROME-France Convective-Scale Operational Model. *Monthly Weather Review* 139, 976–991.
<https://doi.org/10.1175/2010MWR3425.1>

35. Bolger A. M., Lohse M., Usadel B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>

36. Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L. et al. 2012. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 41, D597–D604.
<https://doi.org/10.1093/nar/gks1160>

37. Rimet, F., Gusev, E., Kahlert, M., Kelly, M.G., Kulikovskiy, M., Maltsev, Y. et al. 2019. Diat.barcode, an open-access curated barcode library for diatoms. *Sci. Rep.* 9, 15116.
<https://doi.org/10.1038/s41598-019-51500-6>

38. Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997*. <https://doi.org/10.48550/arXiv.1303.3997>

39. Metegnier, G., Paulino, S., Ramond, P., Siano, R., Sourisseau, M., Destombe, C. et al. 2020. Species specific gene expression dynamics during harmful algal blooms. *Sci. Rep.* 10, 6182.
<https://doi.org/10.1038/s41598-020-63326-8>

40. Keeling, P.J., Burki, F., Wilcox, H.M., Allam, B., Allen, E.E., Amaral-Zettler, L.A. et al. 2014. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing. *PLoS Biol.* 12, e1001889. <https://doi.org/10.1371/journal.pbio.1001889>

41. Lema, K.A., Metegnier, G., Quéré, J., Latimier, M., Youenou, A., Lambert, C. et al. 2019. Inter- and Intra-Specific Transcriptional and Phenotypic Responses of *Pseudo-nitzschia* under Different Nutrient Conditions. *Genome Biol. Evol.* 11, 731–747. <https://doi.org/10.1093/gbe/evz030>
42. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N. et al. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
43. Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2 (preprint). *Bioinformatics*. <https://doi.org/10.1101/002832>
44. Si Y., Liu P., Li P., Brutnell T. P., 2014. Model-based clustering for RNA-seq data. *Bioinformatics*, 30, 197–205. <https://doi.org/10.1093/bioinformatics/btt632>
45. Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P et al. 2024. *vegan*: Community Ecology Package. R package version 2.7-0, <https://github.com/vegandevs/vegan>, <https://vegandevs.github.io/vegan/>.
46. Brunson, J.K., McKinnie, S.M.K., Chekan, J.R., McCrow, J.P., Miles, Z.D., Bertrand, E.M. et al. 2018. Biosynthesis of the neurotoxin domoic acid in a bloom-forming diatom. *Science* 361, 1356–1358. <https://doi.org/10.1126/science.aau0382>
47. Brunson, J.K., Thukral, M., Ryan, J.P., Anderson, C.R., Kolody, B.C., James, C.C. et al. 2024. Molecular Forecasting of Domoic Acid during a Pervasive Toxic Diatom Bloom ». *Proceedings of the National Academy of Sciences* 121, e2319177121. <https://doi.org/10.1073/pnas.2319177121>.
48. Kim, J.H., Ajani, P.A., Murray, S.A., Kang, S.-M., Kim, S.-H., Lim, H.C. et al. 2023. Abiotic and biotic factors controlling sexual reproduction in populations of *Pseudo-nitzschia pungens* (*Bacillariophyceae*). *Harmful Algae* 123, 102392. <https://doi.org/10.1016/j.hal.2023.102392>

49. D'Alelio, D., Ribera d'Alcalà, M., Dubroca, L., Sarno, D., Zingone, A. & Montresor, M. 2010. The time for sex: a biennial life cycle in a marine planktonic diatom. *Limnol. Oceanogr.* 55, 106-14. <https://doi.org/10.4319/lo.2010.55.1.0106>
50. Ambrosino, L., Miralto, M., Barra, L., Sardo, A., Chiusano, M.L., Cozza, R. et al. 2023. Stress Response in Toxic Diatoms: The Effects of Abiotic Factors on Growth and RNA Content in *Pseudo-nitzschia calliantha* and in *Pseudo-nitzschia multistriata*. *J. Mar. Sci. Eng.* 11, 1743. <https://doi.org/10.3390/jmse11091743>
51. Smayda, T.J., 1997. What is a bloom? A commentary. *Limnol. Oceanogr.* 42, 1132–1136. https://doi.org/10.4319/lo.1997.42.5_part_2.1132
52. Petton, S., Pouvreau, S., Dumas, F., 2020. Intensive use of Lagrangian trajectories to quantify coastal area dispersion. *Ocean Dynamics* 70, 541–559. <https://doi.org/10.1007/s10236-019-01343-6>
53. Gifford, S.M., Sharma, S., Booth, M., Moran, M.A., 2013. Expression patterns reveal niche diversification in a marine microbial assemblage. *ISME J.* 7, 281–298. <https://doi.org/10.1038/ismej.2012.96>
54. Gong, W., Paerl, H., Marchetti, A., 2018. Eukaryotic phytoplankton community spatiotemporal dynamics as identified through gene expression within a eutrophic estuary. *Environ. Microbiol.* 20, 1095–1111. <https://doi.org/10.1111/1462-2920.14049>
55. van Creveld, S.G., Rosenwasser, S., Schatz, D., Koren, I., Vardi, A., 2015. Early perturbation in mitochondria redox homeostasis in response to environmental stress predicts cell fate in diatoms. *ISME J.* 9, 385–395. <https://doi.org/10.1038/ismej.2014.136>
56. Bidle, K.D., 2016. Programmed Cell Death in Unicellular Phytoplankton. *Curr. Biol.* 26, R594–R607. <https://doi.org/10.1016/j.cub.2016.05.056>

57. Kranner, I., Birtić, S., Anderson, K.M., Pritchard, H.W., 2006. Glutathione half-cell reduction potential: A universal stress marker and modulator of programmed cell death? *Free Radic. Biol. Med.* 40, 2155–2165. <https://doi.org/10.1016/j.freeradbiomed.2006.02.013>

Figure legends:

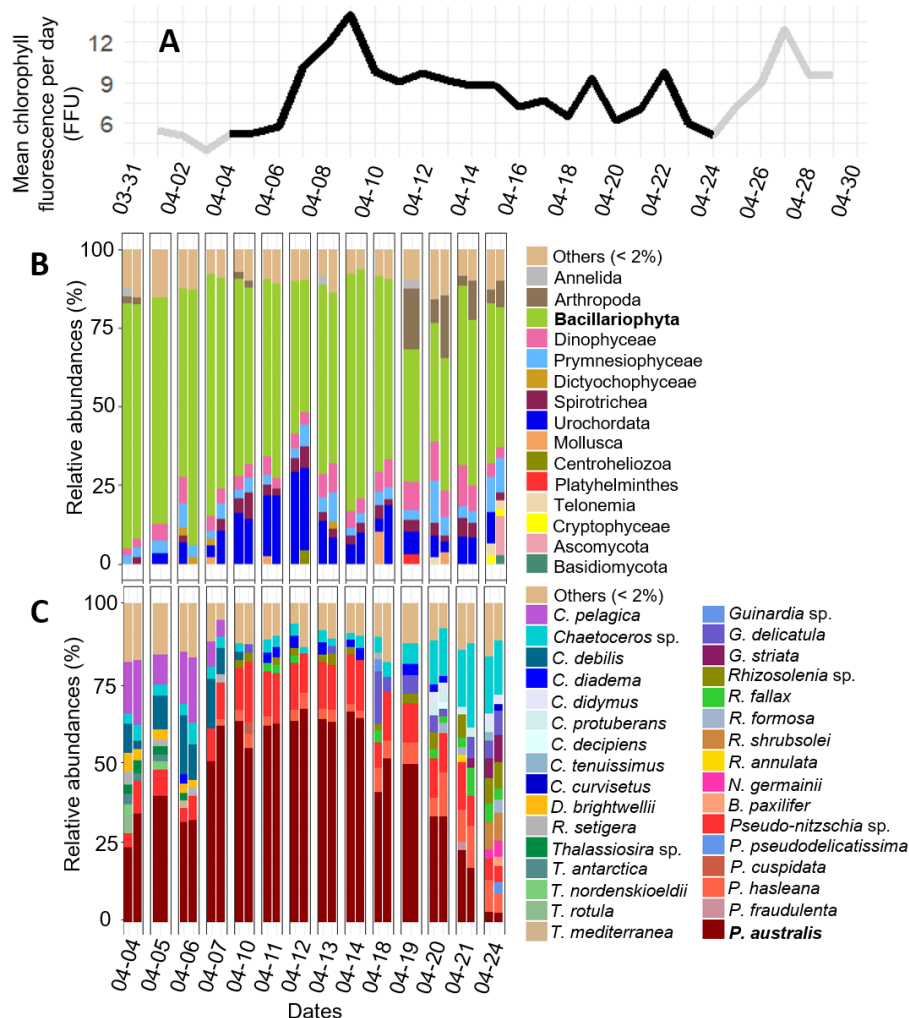


Figure 1: A. Temporal evolution of the mean chlorophyll fluorescence per day between 2017-04-01 and 2017-04-29. The survey period is shown in black. B. Relative abundances of the class community composition during the monitoring. C. Relative abundances of the diatom community composition during the monitoring (C: genus *Cerataulina* then *Chaetoceros*; D: *Ditylum*; T: *Thalassiosira*; G: *Guinardia*, R: *Rhizosolenia*; N: *Navicula*; B: *Bacillaria*; P: *Pseudo-nitzschia*). The category others regroup class, genus, or species lower than 2%.

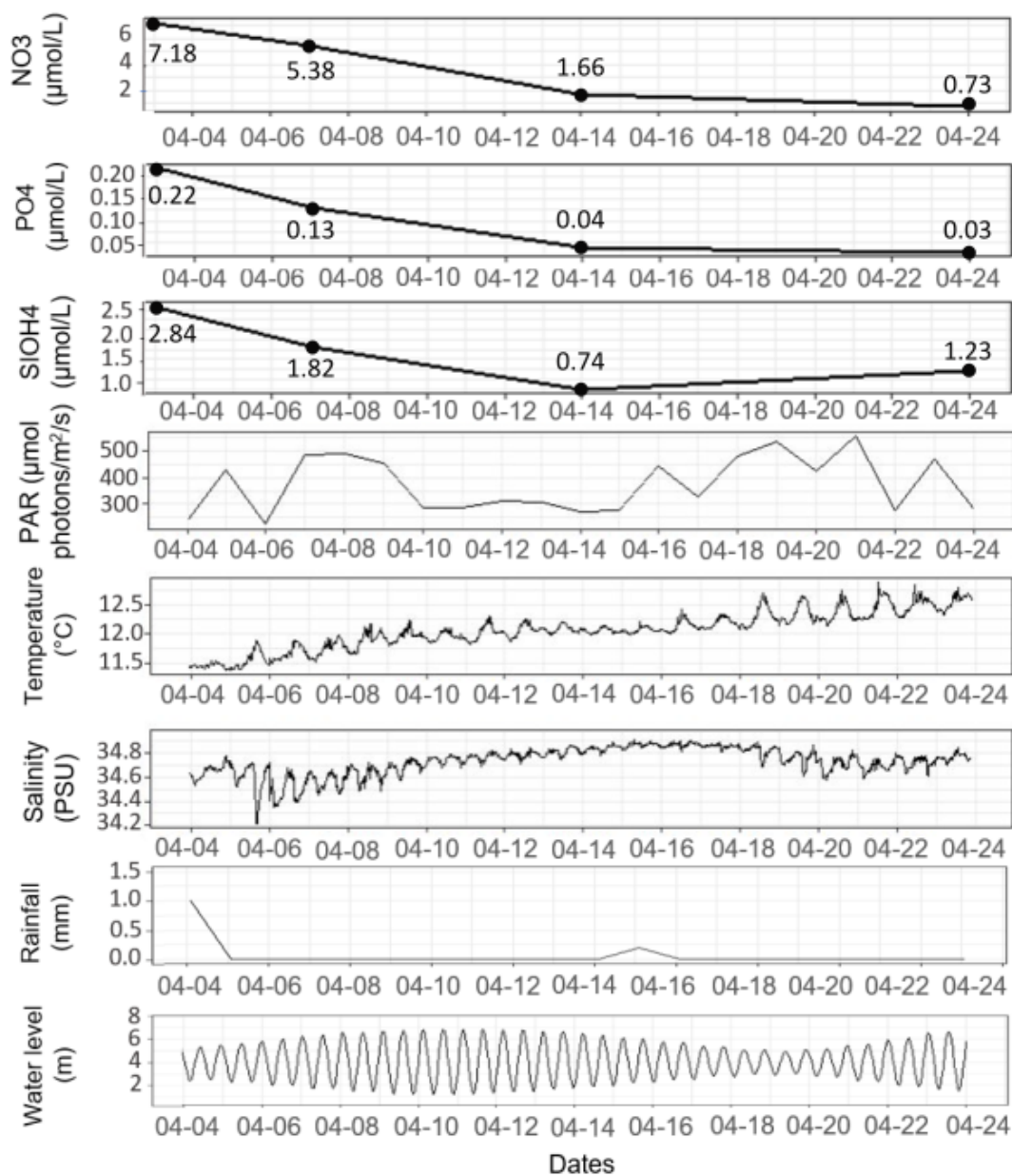


Figure 2: Temporal evolution of abiotic parameters during the sampling period for nutrients: Nitrate (NO₃), Phosphate (PO₄) and Silicate (SiOH₄), average Photosynthetically Active Radiation, temperature, salinity, daily rainfall and water level.

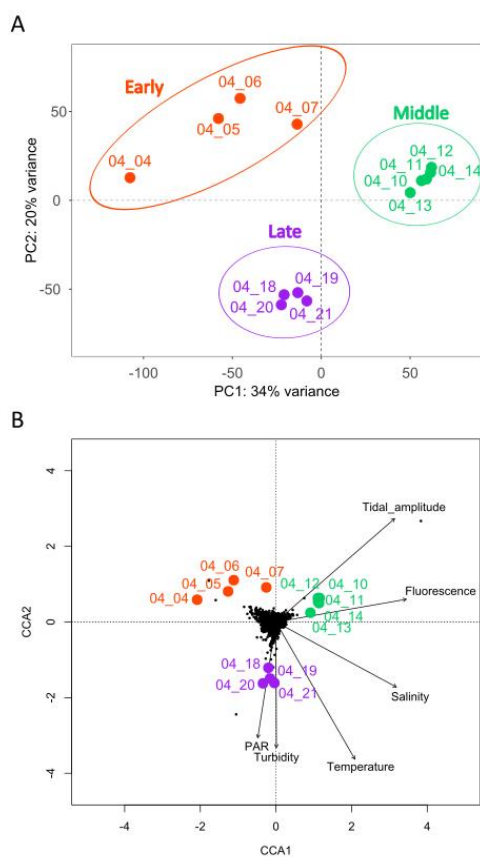


Figure 3: Gene expression variability and relation to environmental conditions. A. Principal component analysis of normalized *P. australis* gene expression (25,140 genes) profiles for each sample. The three sample groups are identified by circles. B. Canonical correspondence analysis of *P. australis* gene expression in relation to environmental conditions. Samples are indicated by coloured dots and text. Black dots correspond to individual genes and black arrow to environmental parameters.

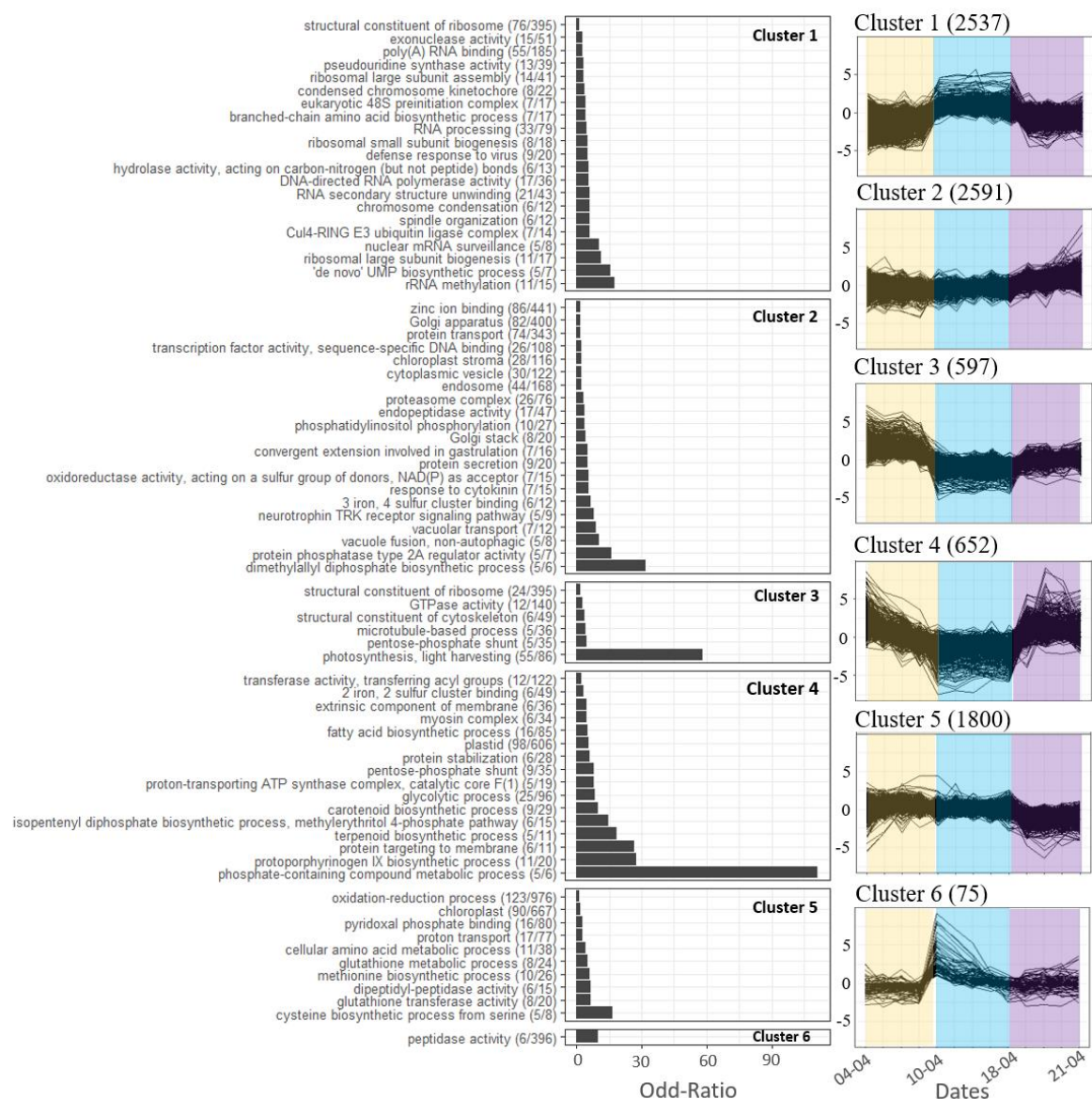


Figure 4: Clusters of *P. australis* genes displaying dynamic expression during the bloom. Right side: clusters of DE genes displaying similar gene expression dynamics (expressed as log₂FC) across samples. The number of transcripts belonging to each cluster is indicated in parenthesis, colors indicate early, middle and late sample groups. Left side: overrepresented biological processes in each cluster (GO terms). GO terms (y axis) were presented based on their odd-ratio (x axis). For each GO term, the number of transcripts belonging to the cluster over the total number of transcripts present in *P. australis* transcriptome is indicated in parenthesis.

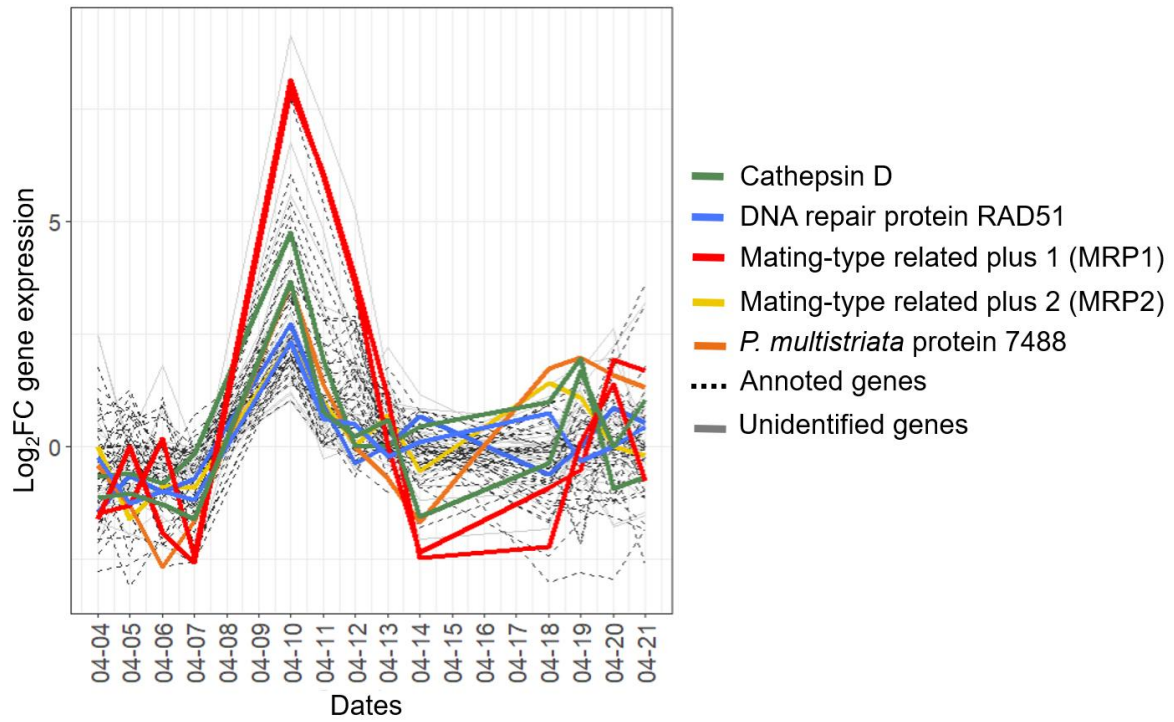


Figure 5: Temporal trends of expression for genes belonging to cluster 6. Genes related to the initiation of sexual reproduction are in colors, genes with other annotated functions are indicated with a dotted black line, and unannotated genes are in grey.

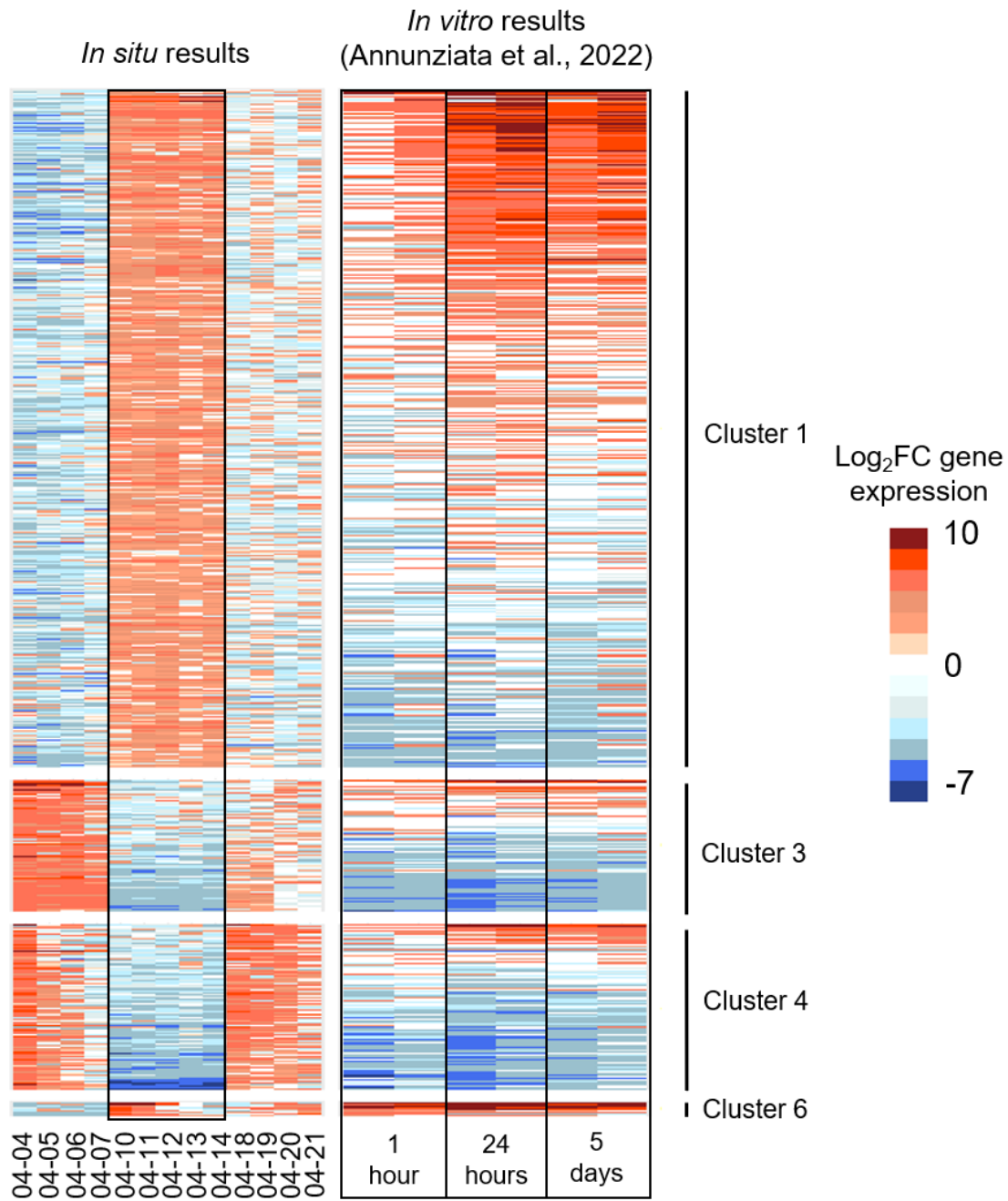


Figure 6: Comparison of the expression levels (under- in blue and over-expressed in red) of *P. multistriata* genes identified as differentially expressed at one hour, 24 hours, and 5 days post-crossing compared to mating type + and - cells *in vitro* (from [20]), with their homologs identified in the *P. australis* reference transcriptome and belonging to clusters 1, 3, 4 and 6 *in situ*. For each *in vitro* time point, the two columns indicate the expression level of the cross against MT+ and MT- strains, respectively. Gene names and log2FC indicated in Supplementary table 5.

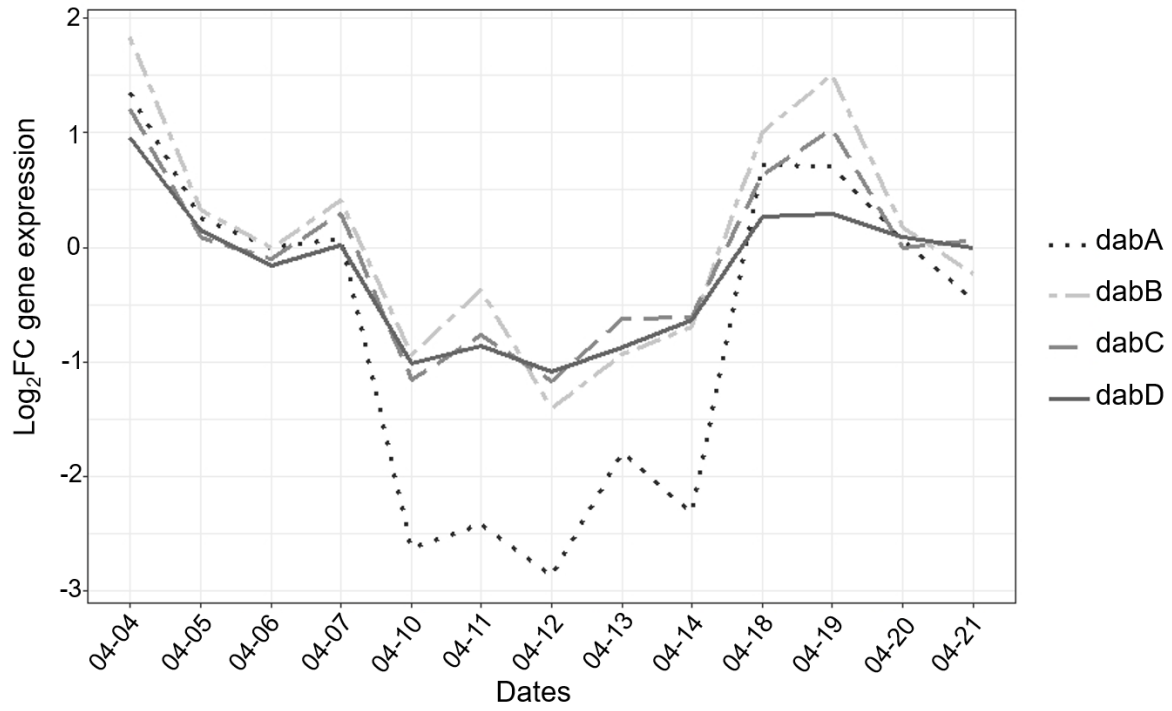


Figure 7: DA biosynthesis gene expression dynamics in *P. australis*. The expression of the *P. australis* genes *dabA*, *dabB*, *dabC* and *dabD* during the time course of the bloom is indicated.

UNCORRECTED MANUSCRIPT