Gene expression during the formation of resting spores in the marine diatom Chaetoceros socialis

Angela Pelusi
School of Earth and Environmental Sciences, Cardiff University
Luca Ambrosino
Department of Research Infrastructures for Marine Biological Resources, Stazione Zoologica Anton Dohrn
Marco Miralto
Department of Research Infrastructures for Marine Biological Resources, Stazione Zoologica Anton Dohrn
Maria Luisa Chiusano
Department of Agricultural Sciences, University of Naples “Federico II"
Alessandra Rogato
Institute of Bioscience and BioResources, IBBR- CNR
Maria Immacolata Ferrante
Department of Integrative Marine Ecology, Stazione Zoologica Anton Dohrn
Marina Montresor (marina.montresor@szn.it)
Department of Integrative Marine Ecology, Stazione Zoologica Anton Dohrn

Research Article

Keywords: Transcriptomics, resting stages, nitrogen transporters, nitrogen starvation, diatoms, Chaetoceros socialis, cell signalling, lipoxygenase

Posted Date: August 1st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1882667/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Genomics on March 10th, 2023. See the published version at https://doi.org/10.1186/s12864-023-09175-x.

Abstract

Background

Dormancy is widespread in both multicellular and unicellular organisms. Among diatoms, unicellular microalgae at the base of all aquatic food webs, several species produce dormant cells (spores or resting cells) that can withstand long periods of adverse environmental conditions.

Results

We present the first gene expression study during the process of spore formation in the marine planktonic diatom *Chaetoceros socialis*. Spore formation was induced by nitrogen depletion. Genes related to photosynthesis and nitrate assimilation, including high-affinity nitrate transporters (*NTRs),* were downregulated. While the former result is a common reaction among diatoms under nitrogen stress, the latter seems to be exclusive of the spore-former *C. socialis.* The upregulation of catabolic pathways, such as tricarboxylic acid cycle, glyoxylate cycle and beta-oxidation, suggests that this diatom could use lipids as a source of energy during the process of spore formation. Furthermore, the upregulation of a lipoxygenase and several aldehyde dehydrogenases (*ALDHs*) advocates the presence of oxylipin-mediated signalling, while the upregulation of genes involved in dormancy-related pathways conserved in other organisms (e.g. serine/threonine-protein kinases *TOR* and its inhibitor *GATOR*) provides interesting avenues for future explorations.

Conclusions

Our results demonstrate that the transition from an active growth phase to a resting one is characterized by marked metabolic changes and provide evidence for the presence of signalling pathways related to intercellular communication.

Background

All over the phylogenetic tree of life, many organisms can undergo a dormant phase in which they stop the developmental program or the vegetative division cycle for a variable length of time in response to unfavourable conditions for growth. Unicellular species, both prokaryotes and eukaryotes, produce resting stages such as akinetes, cysts, or spores, characterized by markedly reduced metabolism, increased content of storage material and often considerable morphological modifications of the cell wall ([1], [2]). The reversible switch from an active growth phase to a resting one is generally induced by the perception of specific environmental signals, such as temperature, photoperiod or concentration of nutrients. The transition between life cycle phases implies energetic investments and dramatic changes in several metabolic pathways ([1]); however, the molecular mechanisms at the base of this transition are still far to be elucidated and have been mainly explored for a few model species and cell types of medical interest (e.g., [3]).

Diatoms are a very diverse group of unicellular microalgae, at the base of aquatic food web and key players in different biogeochemical cycles ([4], [5]). Several species produce resting stages that can be either morphologically differentiated spores or resting cells, apparently indistinguishable from the vegetative ones ([6]). Resting stages are characterized by high carbon (C) content and can be responsible for considerable C export to the benthic compartment ([7], [8]) where they can remain viable for decades, ensuring the persistence of genotypes produced in different years through time ([9], [6], [10]).

Chaetoceros is one of the most species-rich and abundant genera of planktonic diatoms with a worldwide distribution ([11]). Many species in this genus produce morphologically differentiated resting spores ([12], [13]), which are considered proxies of palaeo-productivity ([14]). Among *Chaetoceros* species, *C. socialis* often dominates phytoplankton assemblages in open oceanic and coastal waters (e.g., [15]) including the Gulf of Naples (Tyrrhenian Sea, Mediterranean Sea) where blooms are detected in spring and autumn ([16]) and its spores dominate the dormant diatom assemblage in surface sediments ([17], [18]). Recent studies have shown that the production of spores in *C. socialis* is a density-dependent process, most probably mediated by a chemical cue ([19]), and that spores act as a defense strategy against viral attacks ([20]). Similar to what was observed for other diatoms ([6]), a massive transition from vegetative cells to spores in *C. socialis* is induced by inoculating exponentially growing cells into a culture medium with low nitrogen (N) ([21]). Previous RNA-seq studies have investigated the physiological response of diatoms to N starvation ([22], [23]), a condition that negatively impacts microalgal growth in the marine environment ([24]), but none of them focused on species known to form spores.

Here, we explore gene expression patterns during spore formation induced by N starvation in *C. socialis*. After the assembly of its transcriptome, we tested the differential regulation of the major metabolic pathways, the presence of potential oxylipin-mediated signalling and other conserved pathways known to be involved in cell dormancy in other organisms. Our results provide the first account of the molecular responses during this important transition in the diatom life cycle.

Results And Discussion

The efficacy of N starvation in inducing a rapid and massive transition from vegetative cells to resting spores in *C. socialis*, first shown by Pelusi et al. ([19], [21]), is here confirmed. After three days of exponential growth, *C. socialis* cells grown in N depleted medium entered in stationary phase and vegetative cells (**Figure 1a**) progressively turned into spores (**Figure 1b**), reaching ~30% and ~75% of the total cell number on T3 and T4, respectively (**Figure 1c**). In control conditions, spores were present with a much lower percentage (~15%) only at T4 (**Figure 1d**). The experiment was conducted with exactly the same set-up illustrated in Pelusi et al. ([19]), where concentrations of inorganic nutrients in the medium and of organic N and C in cells were monitored in both treatment and control conditions. The cell and spore concentrations over the four days of the experiment here illustrated matched the overall dynamics observed in Pelusi et al ([19]), where spore formation in the treatment coincided with a marked increase of intracellular C, while the intracellular N pool remained almost constant, notwithstanding the fact that

inorganic N decreased in the medium, suggesting a qualitative shift in intracellular nitrogenous compounds in the cells.

Transcriptome assembly

RNA-seq experiments were performed on *C. socialis* samples collected in N deplete medium before the formation of spores (T2), when spore formation started (T3), and when spores reached > 75% of the whole population (T4); control samples were collected from cells in mid-exponential growth phase on day 2 (C2). Since the *C. socialis* genome is not available yet, raw reads have been assembled with a *de novo* approach. The assembled transcriptome accounted for a total of 32,224 transcripts (**Table S1, S2**), with a mean GC content of 43.98%, an average and median contig length of 776 and 536 bp, respectively, and a N50 of 1140 bp (**Table S1**). *In silico* encoded protein sequences were 19,153. Among these, 6693 (34.9%) were full-length proteins, 5401 (28.2%) were 5' partial proteins lacking a stop codon, 2328 (12.2%) were 3' partial proteins lacking a start methionine, and 4731 (24.7%) lacked both a starting methionine and a stop codon (**Table S1**). The function of 8,070 protein sequences (42.1% of the total predicted proteins) was successfully annotated (**Table S2**). Among these, 7,974 proteins were associated with at least to one gene ontology (GO) term.

The comparisons between each time point (T2, T3 and T4) and the control (C2) were used to detect the molecular pathways involved in spore formation. More than 65% of the total transcripts (21,873) resulted to be differentially expressed genes (DEGs), with a fold change $|FC| \geq 1.5$, and the majority of them belonged to T3 and T4 (**Table 1, Table S2**). The hierarchical clustering of the DEGs dataset highlighted a suite of genes specifically characterizing T3 and T4 compared to the control condition (**Figure 2**).

GO enrichment

A summary of the 20 most significantly enriched biological processes detected by the GO enrichment analysis at the different time points of each comparison is presented in **Figure 3**. A total of 3,057 GO terms were significantly enriched in T2, T3 and T4 *versus* control (**Table S3**). The enriched GO terms for both up- and downregulated DEGs in T2 *versus* control were very few and redundant suggesting that the major metabolic rearrangements started at T3. In fact, the presence of spores at T3 and T4 resulted in an increase in the number of the enriched GO terms, both for up- and downregulated DEGs (**Table 1, Table S3**), highlighting a higher level of transcriptional regulation at these sampling points. In addition to terms related to metabolic processes (*cellular aromatic compounds* and *cellular nitrogen compounds*) and to nuclear activity (*nucleobase-containing compound metabolic process,* and *nucleic acids metabolic processes*), an enrichment of terms relative to upregulated DEGs linked to *cell cycle* was present both at T3 and T4 comparisons and to *cell division* at T4 comparison *versus* control. These latter two observations are paralleled by the high expression of cyclin D3-2 (**Table 2**), whose homolog in *Phaeodactylum tricornutum* was suggested to play a role in G2-to-M transition ([25]). The enrichment of terms related to cell division seems at first counterintuitive because N limitation should negatively affect cell duplication. However, this finding is supported by observations in time-lapse microscopy demonstrating that two consecutive acytokinetic nuclear divisions immediately precede the

transformation of vegetative cells into spores ([21]). After the mitotic division, one nucleus degenerates and one new thick theca of the spore is synthesized; the process is repeated when the second theca of the spore is produced.

Photosynthesis was the most enriched term among the downregulated genes in T3 and T4 compared to the control, followed by *protein-chromophore linkage* and *generation of precursor metabolites and energy generation* (**Figure 3, Table S3**). The decline of photosynthesis during N starvation is a common response in photosynthetic organisms due to the tight connection between C and N metabolisms ([26]). N starvation induces an excess of light absorption relative to C fixation, resulting in photo-oxidative damage to a cell that could lead to cell death; thus, cells must control photosynthesis to mitigate the damage. Liefer et al. ([27]) demonstrated a significant decline of light-harvesting complexes and PSII reaction centers during N starvation in *Thalassiosira pseudonana* and *T. weissflogii.* We can thus conclude that the downregulation of genes related to photosynthesis is a general response in diatoms, well-fitting with the lower metabolism that characterizes resting stages ([2]). The enrichment of terms of downregulated genes such as the *translation elongation* and *organonitrogen compound biosynthetic process*, especially at T3 *versus* control, recalls the downregulation of the same processes in *P. tricornutum* during N starvation ([22]). Surprisingly, also the *reactive oxygen metabolic process* term was enriched at all the downregulated sampling points, being among the first 20 GO terms at T2 together with other stress-related terms (**Figure 3**). However, a number of transcripts specifically related to oxidative stress were found upregulated at T3 and T4 (see below).

Primary metabolism during spore formation

In the following, we provide evidence derived from gene expression analysis for several metabolic pathways detected in *C. socialis* during spore formation. The 10 annotated genes with the highest up- and downregulation values in all pairwise comparisons are reported in **Table 2**; upregulated genes had a FC between 13 and 12 both at T3 and T4 and downregulated ones were not differentially expressed at T2, indicating their involvement during spore formation. Some of the genes were present with a high number of isoforms, such as the high-affinity urea active transporters *DUR3* and the serine/threonine-protein kinases *TOR* (see **Table S4, S7**), suggesting a still high level of redundancy among transcripts even though the filtering applied during the bioinformatic analysis.

Two and four out of the six DEGs tested were found significantly differentiated in APC12 and MCA6, respectively (**Figure S1**). The most differentially expressed high-affinity nitrate transporter *NRT 2.6* and the silicon efflux transporter *LSI3* confirmed their expression in both strains.

Nitrogen assimilation

Surprisingly, the high-affinity nitrate transporters (*NRTs*) in *C. socials* had a distinctive expression profile in N starved conditions as compared to the one detected in other diatoms ([28], [29]). The assimilation of nitrate was severely impacted by N stress: three out of the four *NRTs* were in fact highly downregulated, both at T3 and T4, with one of them (high-affinity nitrate transporter 2.6, *NRT2.6*) being the most

downregulated transcript of the whole dataset (**Figure 4**, **Table 2 B, Figure S1**). *NRTs* can be constitutive or inducible in microalgae, as in plants, especially in response to low N concentration ([28]). Examples among diatoms are *T. pseudonana*, *Pseudo-nitzschia multiseries* and *P. tricornutum*, which showed upregulation of these genes in N starvation, with the only exception of one transcript of *P. multiseries* (Pm 261779) ([23]). In the haptophyte *Tisochrysis lutea,* the expression of one of the four *NTRs* decreased in N depletion, showing an expression profile matching the decline of the intracellular N:C ratio ([30]). This ratio markedly decreased also in *C. socialis,* due to C accumulation when cells turned into spores ([19]), suggesting that a similar mechanism could explain the observed expression profile in our dataset. In plants, light and C modulate *NRTs* expression as well and nitrate uptake is not only determined by nitrate availability and demand but also by C produced by photosynthesis ([31]).

We could not detect any low-affinity nitrate transporters (*NPFs*) in *C. socialis*, thus confirming the results recently obtained for various *Chaetoceros* species in which *NPFs* are missing ([32]); this could support the suggestion that the species of this genus can preferentially use ammonia, possibly from bacteria, as a source of N in case of fluctuating nitrate concentrations ([33]). While the expression of ammonium transporters was somehow variable in our dataset, the high-affinity urea active transporters *DUR3* were upregulated at least in one sampling point, with two of them already differentially expressed at T2 (**Figure 4, Table S4**).

Several transcripts coding for glutamine synthetase (*GS*) and glutamate synthase (*GOGAT*), key enzymes in N assimilation, have been detected (**Figure 4, Table S4**). Among the *GS* transcripts, one transcript coding for *GSIII* showed a markedly positive FC at T3 and T4 *versus* control, as reported for *T. pseudonana* ([34]) and for the cyanobacterium *Synechococcus* ([35]) during the early phase of N starvation. Hockin et al. ([34]) suggested that in *T. pseudonana* this gene participates in the assimilation of ammonium obtained by intracellular catabolic processes, in combination with a NAD(P)H-*GOGAT*. In our dataset, only a chloroplastic ferredoxin-*GOGAT* enzyme displayed a similar positive trend at all sampling points, while the other enzymes using NADH as cofactor showed only slight positive regulation at T3 (**Figure 4**). It can thus be hypothesized that *C. socialis*, as other diatoms, recovers most of N by recycling the internal pools, as further suggested by the detection of autophagy-related proteins and the hydrolysis of urea by urease, over-expressed at T3 and T4 (**Figure 4**, **Table S4**). On the other hand, despite the upregulation of urease, the urea cycle transcripts did not appear differentially expressed (**Figure 4**). Another important source of N could be the ammonia obtained from the breakdown of phosphoethanolamines through the ethanolamine-phosphate phospho-lyase, an enzyme showing a FC greater than 12 when spores were detected at T3 and T4 (**Table S4**).

Tricarboxylic acid cycle

The tricarboxylic acid (TCA) cycle is responsible for generating energy through the oxidation of pyruvate to form CO_2, ATP, NADH, and carbon skeletons used for biosynthetic processes. This term was enriched in the GO enrichment analysis, although it was not among the first 20 terms, and presented several enzymes upregulated during the time course of the experiment (**Table S4**). The upregulation of two isocitrate

lyases with FCs of 4 and 8, respectively, only when spores were present (i.e., T3 and T4) could imply that the glyoxylate cycle, a variant of TCA cycle that uses lipids as a source of energy, is preferred by diatom spores (**Table S4**). The upregulation of transcripts such as the mitochondrial short-chain specific acyl-CoA dehydrogenase and the peroxisomal multifunctional enzyme type 2 (*MFE-2*) that increased progressively in T3 and T4, suggest the involvement of the beta-oxidation of fatty acids in feeding the glyoxylate cycle with acetate molecules (**Table S4**). Indeed, a calcium-independent phospholipase A2-gamma responsible for membrane lipid degradation has been found among the most up regulated genes (**Table 2**). These metabolic changes are consistent with those involved in the formation of pellicle cysts in the dinoflagellate *Scripsiella trochoidea* ([36]).

Cell wall

One of the most relevant changes during the formation of spores is the deposition of two thick heteromorphic siliceous thecae that confer mechanical protection to the spores. Metabolites such as proline and long-chain polyamines (LCA) are involved in the synthesis of the organic component of the siliceous cell wall and their synthesis is connected to the urea cycle and the TCA cycle ([37]). Transcripts related to the proline biosynthesis were not particularly overexpressed at any sampling point, but evidence of spermidine production, another polyamine, came from the upregulation of several spermidine synthases, especially when spores were present (**Table S4**). Furthermore, polyamines transporters and polyamine oxidases, involved in the regulation of their intracellular concentration, showed high FC at T3 and T4. The deposition of new silica thecae of the spores was supported also by the simultaneous increase of two silicon efflux transporters (*LSI*), with FC between 4 and 5, and of several other transcripts presenting the InterPRO domain 'Silicon transporter' (IPR004693) with extremely high FC (up to 11) (**Table S4**, **Figure S1**).

Carbon skeletons

Part of carbon skeletons for sustaining the cells were obtained from the degradation of chrysolaminaran, the principal storage compound of diatoms. The exo-1,3,-beta-glucanases, presumed enzymes responsible for its degradation ([38]), were found with ten different transcripts, eight of which were upregulated particularly when spores were present with FC ranging from 1.9 to 6 (**Table S4**). These skeletons most probably feed the glycolysis and the pentose phosphate pathways, both with upregulated enzymes (**Table S4**).

The pentose phosphate pathway produces NADH and pentose sugars in the oxidative and reductive phases, respectively. Among the sugars produced, the ribose 5-phosphate is the precursor of nucleotides and thus essential for DNA replication and transcription. In analogy to what was observed in N starved *P. tricornutum* ([22]), enzymes involved in the oxidative pathway - two glucose-6-phosphate 1-dehydrogenases and two 6-phosphogluconate dehydrogenase decarboxylating 1, one chloroplastic and one cytosolic - were upregulated at T3 and T4 (**Table S4**). Transcripts of the non-oxidative part were instead generally downregulated, except for one chloroplastic transketolase, which is responsible for the

production of D-xylulose 5-phosphate and D-ribose 5-phosphate; this latter molecule is the precursor of nucleotide biosynthesis, essential for the formation of spores.

Chemical signaling

Pathways related to oxylipin production

There are several examples of chemically mediated communication in unicellular organisms ([39]), with some of them having a critical role in dormancy ([40]). In diatoms, the most studied infochemicals related to intra- and inter-specific communication are compounds belonging the oxylipin family, which include polyunsaturated aldehydes (PUAs) and linear oxygenated fatty acids (LOFAs), both derived from the oxygenation of fatty acids ([41]). The production of LOFAs has been reported for *C. socialis* although the enzymatic pathway involved is still unknown ([42]). The first step of their synthesis derives from the oxidation of membrane lipids, through several enzymes among which a crucial role is played by lipoxygenases. A single-copy lipoxygenase with an extremely high FC (>10) has only been detected when spores were present (i.e. T3 and T4), supporting the hypothesis of oxylipin-mediated signaling as the trigger of spore formation (**Table S5**). The formation of spores under stress conditions – i.e., N depletion, high cell density or viral attack ([19],[20]) – in which oxylipins can be produced following cell lysis or breakage, further corroborates the involvement of these infochemicals in determining this life cycle transition in *C. socialis*. In addition to the overexpression of the lipoxygenase, several aldehyde dehydrogenases (*ALDHs*), plausibly linked to the detoxification from oxylipins, were highly overexpressed especially at T3 and T4 (**Table S5**). These enzymes are very conserved all over the phylogenetic tree of life and have a variety of functions, spanning from detoxification against oxidative stress ([43]), to being markers for highly proliferating stem cells and cancer cell phenotypes ([44] and reference therein). We hypothesize that the simultaneous production of LOFAs and *ALDHs* in *C. socialis* may be attributed to the presence of two different cell types in the culture: LOFAs could be produced by N starved cells in poor physiological conditions as a response to stress, while the *ALDHs* by healthy cells that could 'react' to stress starting a series of intracellular cascade signals that lead to spore formation.

Programmed cell death and oxidative stress

ALDHs are also associated with programmed cell death (PCD) in addition to stress oxidative response, as shown for N starved diatoms ([45]). In other microalgae, these two processes are involved in cell cycle arrest and/or differentiation ([46], [47], [48]). The exposure to a sub-lethal concentration of reactive oxygen species modifies a broad range of cellular processes, activating or inhibiting transcription factors, membrane channels, etc., and thus acting as secondary messengers for various physiological processes aimed to increase cell survival ([49]). The high expression of antioxidant enzymes, counteracts the undergoing oxidative stress during spore formation: for example, the peroxiredoxin-6 and the quinone oxidoreductase *PIG3* were detected among the most overexpressed genes in the dataset (**Table 2**). Interestingly, the quinone oxidoreductase *PIG3* has a double role in inducing apoptosis or prolonged cell cycle arrest in cancer mammalian cells, depending on the physiological state of cells ([50]). Thus, it is

possible that this enzyme is produced by both resting and dying cells of *C. socialis* in our experimental set-up.

Diatoms have a surveillance system based on Ca^{2+} and nitric oxide (NO) production, which helps to monitor the stress levels within the population ([51], [52]). Although only one out of the three nitric oxide synthases present in *C. socialis* was weakly up regulated in the early stage of spore formation (T3), several Ca^{2+}-dependent protein kinases have been found with FCs ranging from 2 to 12, with the highest FCs detected when spores were present (**Table S5**). The expression among the antioxidants of two chloroplastic enzymes - a glutathione synthetase and a thioredoxin-like 2-1 with FCs ranging from 5 to 7 especially at T3 and T4, hinted an important role for the chloroplast redoxome in defining which cells were dying or forming spores. However, the contemporary presence of the two types of cells in the culture is most probably at the base of the contrasting expression profile of many antioxidants and PCD related genes within the dataset, which hampers a precise interpretation of the biological meaning of these genes.

Genes involved in quiescence

Quiescence, i.e. a reversible state in which a cell does not divide but retains the ability to re-enter cell proliferation, presents very conserved molecular features among evolutionary distant organisms ([3]). For example, both quiescent mammalian cells and yeasts arrest the cell cycle in G1, condense chromosomes, reduce rRNA synthesis and translation and activate autophagy mechanisms becoming more resistant to different stresses ([3] and references therein). The same features were recorded in the *C. socialis* transcriptome, as shown by the presence of genes related to autophagy and low translation, together with the nuclear rearrangement observed during spore formation. Among conserved pathways that have been related to quiescence in yeasts, there are several serine/threonine-protein kinases *TOR*; the activity of this enzyme decreases in N starved *Saccharomyces cerevisiae*, inducing the formation of spores ([53]). Different transcripts related to the *TOR* showed a marked positive expression (FCs from 3.5 to 6) when spores were present (T3 and T4) (**Table S6**). However, the same trend was observed for its inhibitor *GATOR* supporting once more the fact that different signals are produced by cells undergoing different fates, i.e. dying cells and cells that turn into spores.

Recent studies on resting cells in the dinoflagellate *Scrippsiella trochoidea* and the diatom *P. tricornutum* reported the possible involvement of the phytohormone abscisic acid (ABA) in regulating the transition between active cell division and quiescence ([54], [36], [55]). This molecule is a well know signal initiator in seed dormancy but it is still poorly studied in microalgae ([36]). In *C. socialis*, the expression of 9-cis-epoxycarotenoid dioxygenase, the rate-limiting enzyme of the ABA biosynthetic pathway, increased concomitantly to spore production (FCs from 2.6 to 6 in T3 and T4 *versus* control comparisons); the farnesyltransferase subunit beta is an essential part of the farnesyltransferase complex, an ABA negative regulator and maintained high FCs (from 5.3 to 6.7) at the same sampling points (**Table S6**).

Conclusion

The results of our study provide a first insight into the metabolic pathways activated in the centric diatom *C. socialis* during the transition from vegetative growth to the formation of resting spores. Genes related to photosynthesis and nitrate assimilation were down-regulated when spores were produced and the concomitant upregulation of glutamine synthetase (*GS*), ureases, ethanolamine-phosphate phospho-lyase suggests that N is recovered by recycling internal pools. The upregulation of spermidine synthases and polyamines transporters, together with silicon transporters and the enrichment of terms related to cell division provide functional evidence for the active synthesis of the spores' siliceous walls. Transcriptome data provided also evidence for signalling cues and mechanisms involved in the formation of *Chaetoceros* spores. A highly overexpressed lipoxygenase together with several aldehyde dehydrogenases were detected during spore formation, suggesting the involvement of oxylipins as chemical cues. The detection of different transcripts related to the serine/threonine-protein kinases *TOR* together with its inhibitor (*GATOR*) suggest their possible involvement in the formation of diatom spores, in analogy to their role in the induction of quiescence in yeast.

The complex molecular signals detected in our dataset can be explained by the simultaneous presence at T3 and T4 of two distinct cell types undergoing distinct fates: vegetative cells negatively impacted by N starvation together with cells that could react to stress by turning into resting spores. Detailed investigations at the level of individual cells, taking advantage of recently established protocols ([56]) together with the availability of the genome of *C. socialis*, will hopefully enable to advance our understanding of the molecular mechanisms that regulate this important life cycle transition.

Materials And Methods

Experimental set-up, RNA extraction and sequencing

The experiment was carried out using a newly established clonal strain of *C. socialis,* APC12, genotyped by sequencing the LSU rDNA region ([21]). A non-axenic stock has been maintained in a culture chamber at 18 ± 2 °C, under sinusoidal illumination (12L:12D h photoperiod, ~ 90 µmol photons·m^{-2}· s^{-1} daily average) in control medium made with artificial seawater at a salinity of 36 (Sea salts, Sigma-Aldrich; [21]) and with the following concentration of inorganic nutrients: 580 µM of $NaNO_3$, 300 µM of Na_2SiO_3 and 29 µM of NaH_2PO_4 . An exponentially growing culture was used to inoculate, at an initial cell density of ~ 3×10^3 cells·mL^{-1}, three 5 L glass flasks filled with 3 L of control medium and three flasks filled with low nitrate medium (23 µM of $NaNO_3$, 300 µM Na_2SiO_3, 29 µM NaH_2PO_4). Temperature and light conditions were monitored during the experiment with a HOBO Pendant® Temperature/Light Data Logger. To estimate cell concentration, 4 mL of sample were collected every day, fixed with 1.6% formaldehyde solution, and vegetative cells and spores were enumerated using a Sedgwick-Rafter chamber on a Zeiss Axiophot (ZEISS, Oberkochen, Germany) microscope at 400x magnification.

Total RNA was extracted from each replicate of the control in mid-exponential growth phase at day 2 (C2) and from the replicates growing in N deplete conditions on three consecutive days: before the formation of spores (T2), when spore formation started (T3), and when they reached > 75% of the whole population (T4) **(Figure 5)**. A total of ~1.2×10^7 cells were harvested from each replicate by filtration onto 1.2 µm pore size filters (RAWP04700 Millipore) and extracted with Trizol™ (Invitrogen) following manufacturer's instructions. A DNase I (Qiagen) treatment was applied to remove gDNA contamination, and RNA was further purified using RNeasy Plant Mini Kit (Qiagen). All samples were quantified with Qubit® 2.0 Fluorometer (Invitrogen) and quality checked with an Agilent 2100 bioanalyzer (Agilent Technologies, California, USA) and a NanoDrop ND-1000 Spectrophotometer (Nanodrop Tecnologies Inc., Wilmington, USA). Samples were then pooled in equal concentrations of 100 $ng \cdot \mu l^{-1}$ for sequencing at the Molecular Service of Stazione Zoologica with an Ion Proton™ sequencer (Life Technologies, Carlsbad, USA) using an Ion P1 sequencing Kit v2, generating single-read sequences. Raw reads coming from each replicate were collected in fastqc format files. One of the T3 replicates was removed from downstream analyses due to a sequencing error during library construction. The resulting raw reads were deposited in the Sequence Read Archive (SRA) partition at NCBI with the accession number PRJNA826817.

Reads quality check, transcriptome assembly and functional annotation

The reads quality check was performed using FASTQC ([57]). A trimming step of the low-quality bases at 5' and 3' was performed using Trimmomatic ([58]). Low-quality nucleotides were trimmed from the ends of the reads (first 8 bases), setting the minimum quality per base at a Phread score of 20 and minimum and maximum length of the reads after cleaning at 25 bp and 240 bp, respectively. Cleaned reads were assembled into transcript sequences using Trinity v.2.11.0 ([59]) with *in silico* read normalization, setting the -min_kmer_cov parameter at 2. The clustering of the transcriptome was performed using the CD-hit-est software (v. 4.6.8, [60],) with 90% identity threshold in order to remove transcriptome redundancy. The whole transcriptome was aligned with BLASTx software ([61]) versus the Uniprot SwissProt database (downloaded in July 2020), setting the e-value threshold to $1e^{-3}$. A filtering step was performed at this stage for removing all the matches against bacterial sequences from the transcriptome.

The prediction of the encoded proteins from the assembled transcripts was obtained via TransDecoder v 5.3.0 (https://github.com/TransDecoder/TransDecoder/releases). Coding sequences were identified by the software based on: 1) a minimum length Open Reading Frame (100 by default to minimize the number of false positives); 2) an internal score system; 3) if a candidate ORF is entirely included within the coordinates of another candidate ORF, the longer one is reported. The functional annotation of the predicted proteins was performed by InterProScan (version 5.33) ([62]).

Differentially Expressed Genes (DEGs) calling and gene ontology enrichment

All the cleaned reads were mapped on the assembled *C. socialis* transcriptome using the Bowtie2 aligner (default settings, [63]). Reads count and FPKM calculation per tag for each replicate was performed using the eXpress software ([64]). DEGs calling was performed using two tools implementing two different

statistical approaches: DESeq2 ([65]) and edgeR ([66]). . The mean of the log2 FC values (Log2FC) obtained with the two tools was calculated for each transcript. The thresholds for the DEGs calling were FDR ≤0.05, P-adjusted ≤0.05, and |Log2 FC| ≥ 1.5|.

A Gene Ontology enrichment analysis of the detected DEGs was performed with Ontologizer software ([67]). The threshold used to identify significantly enriched functional terms was P ≤0.05. Genes known to be related to different metabolic pathways were manually searched within the transcriptome considering their SwissProt annotation

qPCR

Six DEGs were validated through a real-time qPCR analysis (**Table S7**). Three DEGs were randomly chosen, in addition to the most downregulated high-affinity nitrate transporter (*NTR2:6*) and one NADH-nitrate reductase, which are related to nitrate uptake and a silicon efflux transporter (*LSI3*) related to the deposition of silicon in spore valves. Two genotyped strains of *C. socialis*, namely APC12 and MCA6 were used for this purpose: the former strain is the one used for the transcriptome experiment, while MCA6 is a strain isolated at station LTER-MC in the Gulf of Naples and for which the D1–D3 region of the nuclear-encoded large subunit ribosomal DNA (partial 28S rDNA) has been sequenced as in [68] to confirm its identity.

Triplicate cultures of both strains were maintained in control and low N media for 2 and 3 days respectively. Cells were harvested when the percentage of spores was zero in the control and ~33 and ~38% for APC12 and MCA6, respectively, in treatment; the latter condition corresponded to that at T3 of the transcriptome experiment. RNA extraction and purification were performed as illustrated above. Total RNA was reverse-transcribed using the QuantiTect® Reverse Transcription Kit (Qiagen, Venlo, Limburgo, Nederlands).

qPCR amplification was performed with cDNA diluted 1:10, in a 10 µl reaction containing each primer at a final concentration of 1 µM and Fast SYBR Green Master mix with ROX (Applied Biosystems) using a ViiA™ 7 Real-Time PCR System (Applied Biosystems by Life Technologies, Carlsbad, CA, USA) and the following cycling parameters: 95 °C for 20 s, 40 cycles at 95 °C for 1 s, 60 °C for 20 s, 95 °C for 15 s, 60 °C 1 min, and a gradient from 60 °C to 95 °C for 15 min. Raw results were processed using the ViiA™ 7 Software and exported into Microsoft Excel for further analyses. The reference gene used was the tubulin gamma chain (*TUB* G) designed using sequence information from the transcriptome and the software Primer3Plus v.2.4.2 ([69]). The sequences for the forward and reverse primers are 5'-TGCAGAGTTTGGTCGATGAG -3'and 5'-GGAAGCCAAAGAGTCTGCTG-3', respectively, yielding a PCR product of 197 bp (**Table S7**). Primers for all other tested DEGs were designed using the same approach. FCs were obtained with the Relative Expression Software Tool-Multiple Condition Solver (REST-MCS) ([70]). A pairwise fixed reallocation randomisation test has been used to identify statistically significant results (P ≤0.05).

Abbreviations

NTR: high-affinity nitrate transporter

ALDHs: aldehyde dehydrogenases

GO: Gene ontology

DEGs: Differentially expressed genes

FC: Fold change

NPFs: low-affinity nitrate transporters

DUR3: urea active transporters

GS: glutamine synthetase

GOGAT: glutamate synthase

TCA: tricarboxylic acid

MFE: multifunctional enzyme type

LCA: long-chain polyamines

LSIs: silicon efflux transporters

PUAs: polyunsaturated aldehydes

LOFAs: linear oxygenated fatty acids

PCD: programmed cell death

NO: nitric oxide

TORC1: target of rapamycin complex I

ABA: abscisic acid

qPCR: Real-time quantitative polymerase chain reaction

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The dataset generated during the current study is available in the Archive Sequence Read (SRA) partition at NCBI repository deposited with the accession number PRJNA826817 [Temporary link:https://dataview.ncbi.nlm.nih.gov/object/PRJNA826817?reviewer=2c2v67pphm5rcnqu7n7snpsbqe]. All data analysed during this study are included in this published article and its supplementary information files.

Competing interests

The authors declare that they have no competing interests.

Funding

AP has been supported by a PhD fellowship from Stazione Zoologica Anton Dohrn (SZN) and by a fellowship of the project ABBaCo (Restauro Ambientale e Balneabilita' del SIN Bagnoli-Coroglio), funded by the Italian Ministry for Education, University and Research (grant no. C62F16000170001)

Author Contributions Statement

AP, MIF and MMo conceived the study, AP performed the experiment. LA, MMi and MLC assembled the transcriptome. AP, LA, AR a MMI analyzed the data. AP and LA prepared figures and tables. AP wrote the paper that was discussed and revised by all the authors. All authors read and approved the final manuscript.

Acknowledgement

We would like to thank Pasquale De Luca and the Research Infrastructures for marine biological resources Department (SZN) for their assistance during the sequencing.

References

1. Lennon JT, Jones SE. Microbial seed banks: the ecological and evolutionary implications of dormancy. Nat Rev Microbiol. 2011;9(2):119-30.
2. Ellegaard M, Ribeiro S. The long-term persistence of phytoplankton resting stages in aquatic 'seed banks'. Biol Rev. 2018;93(1):166-83.
3. Valcourt JR, Lemons JMS, Haley EM, Kojima M, Demuren OO, Coller HA. Staying alive Metabolic adaptations to quiescence. Cell Cycle. 2012;11(9):1680-96.

4. Treguer P, Bowler C, Moriceau B, Dutkiewicz S, Gehlen M, Aumont O, et al. Influence of diatom diversity on the ocean biological carbon pump. Nat Geosci. 2018;11(1):27-37.
5. Smetacek V. Diatoms and the ocean carbon cycle. Protist. 1999;150(1):25-32.
6. McQuoid MR. Diatom resting stages. J. Phycol. 1996;32:889-902.
7. Rynearson TA, Richardson K, Lampitt RS, Sieracki ME, Poulton AJ, Lyngsgaard MM, et al. Major contribution of diatom resting spores to vertical flux in the sub-polar North Atlantic. Deep-Sea Res Pt I. 2013;82:60-71.
8. Rembauville M, Manno C, Tarling GA, Blain S, Salter I. Strong contribution of diatom resting spores to deep-sea carbon transfer in naturally iron-fertilized waters downstream of South Georgia. Deep-Sea Res Pt I. 2016;115:22-35.
9. Sicko-Goad L, Stoermer E, Kociolek J. Diatom resting cell rejuvenation and formation: time course, species records and distribution. J Plankton Res. 1989;11(2):375-89.
10. Härnström K, Ellegaard M, Andersen TJ, Godhe A. Hundred years of genetic structure in a sediment revived diatom population. P Natl Acad Sci USA. 2011;108(10):4252-7.
11. Malviya S, Scalco E, Audic S, Vincenta F, Veluchamy A, Poulain J, et al. Insights into global diatom distribution and diversity in the world's ocean. P Natl Acad Sci USA. 2016;113(11):E1516-E25.
12. Ishii KI, Iwataki M, Matsuoka K, Imai I. Proposal of identification criteria for resting spores of *Chaetoceros* species (Bacillariophyceae) from a temperate coastal sea. Phycologia. 2011;50(4):351-62.
13. Belmonte G, Rubino F. Resting cysts from coastal marine plankton. Oceanogr Mar Biol. 2019:57:1-88.
14. Abelmann A, Gersonde R, Cortese G, Kuhn G, Smetacek V. Extensive phytoplankton blooms in the Atlantic sector of the glacial Southern Ocean. Paleoceanography. 2006;21(1):PA1013.
15. Leblanc K, Aristegui J, Armand L, Assmy P, Beker B, Bode A, et al. A global diatom database - abundance, biovolume and biomass in the world ocean. Earth Syst Sci Data. 2012;4(1):149-65.
16. Zingone A, D'Alelio D, Mazzocchi MG, Montresor M, Sarno D, Balestra C, et al. Time series and beyond: multifaceted plankton research at a marine Mediterranean LTER site. Nat Conserv. 2019;34:273-310.
17. Montresor M, Di Prisco C, Sarno D, Margiotta F, Zingone A. Diversity and germination patterns of diatom resting stages at a coastal Mediterranean site. Mar Ecol-Prog Ser. 2013;484:79-95.
18. Piredda R, Sarno D, Lange CB, Tomasino MP, Zingone A, Montresor M. Diatom resting stages in surface sediments: a pilot study comparing Next Generation Sequencing and Serial Dilution Cultures. Cryptogamie Algol. 2017;38(1):31-46.
19. Pelusi A, Margiotta F, Passarelli A, Ferrante MI, Ribera d'Alcala M, Montresor M. Density-dependent mechanisms regulate spore formation in the diatom *Chaetoceros socialis*. Limnol Oceanogr Letters. 2020;5(5):371-8.
20. Pelusi A, De Luca P, Manfellotto F, Thamatrakoln K, Bidle KD, Montresor M. Virus-induced spore formation as a defense mechanism in marine diatoms. New Phytol. 2021;229(4): 2251–9.

21. Pelusi A, Santelia ME, Benvenuto G, Godhe A, Montresor M. The diatom *Chaetoceros socialis*: spore formation and preservation. Eur J Phycol. 2020;55(1):1-10.
22. Alipanah L, Rohloff J, Winge P, Bones AM, Brembu T. Whole-cell response to nitrogen deprivation in the diatom *Phaeodactylum tricornutum*. J Exp Bot. 2015;66(20):6281-96.
23. Bender SJ, Durkin CA, Berthiaume CT, Morales RL, Armbrust E. Transcriptional responses of three model diatoms to nitrate limitation of growth. Front Mar Sci. 2014; doi: 10.3389/fmars.2014.00003
24. Berges JA, Falkowski PG. Physiological stress and cell death in marine phytoplankton: induction of proteases in response to nitrogen or light limitation. Limnol Oceanogr. 1998;43(1):129-35.
25. Huysman MJ, Martens C, Vandepoele K, Gillard J, Rayko E, Heijde M, et al. Genome-wide analysis of the diatom cell cycle unveils a novel type of cyclins involved in environmental signaling. Genome Biol. 2010;11(2):1-19.
26. Cullen JJ, Yang XL, Macintyre HL Nutrient limitation of marine photosynthesis. In: Falkowski PG, Woodhead AD, Vivirito K, editors. Primary Productivity and Biogeochemical Cycles in the Sea. Springer; 1992: p. 69-88.
27. Liefer JD, Garg A, Campbell DA, Irwin AJ, Finkel ZV. Nitrogen starvation induces distinct photosynthetic responses and recovery dynamics in diatoms and prasinophytes. PLoS One. 2018; doi.org/10.1371/journal.pone.0195705.
28. Rogato A, Amato A, Iudicone D, Chiurazzi M, Ferrante MI, Ribera d'Alcala M. The diatom molecular toolkit to handle nitrogen uptake. Mar Genom. 2015;24(1):95-108.
29. Busseni G, Vieira FRJ, Amato A, Pelletier E, Karlusich JJP, Ferrante MI, et al. Meta-omics reveals genetic flexibility of diatom nitrogen transporters in response to environmental changes. Mol Biol Evol. 2019;36(11):2522-35.
30. Charrier A, Berard JB, Bougaran G, Carrier G, Lukomska E, Schreiber N, et al. High-affinity nitrate/nitrite transporter genes (Nrt2) in *Tisochrysis lutea*: identification and expression analyses reveal some interesting specificities of Haptophyta microalgae. Physiol Plantarum. 2015;154(4):572-90.
31. Vidal EA, Alvarez JM, Araus V, Riveras E, Brooks MD, Krouk G, et al. Nitrate in 2020: thirty years from transport to signaling networks. Plant Cell. 2020;32(7):2094-119.
32. Santin A, Caputi L, Longo A, Chiurazzi M, Ribera d'Alcala M, Russo MT, et al. Integrative omics identification, evolutionary and structural analysis of low affinity nitrate transporters in diatoms, diNPFs. Open Biology. 2021;11:200395.
33. Olofsson M, Robertson EK, Edler L, Arneborg L, Whitehouse MJ, Ploug H. Nitrate and ammonium fluxes to diatoms and dinoflagellates at a single cell level in mixed field communities in the sea. Sci Rep. 2019;9:1424
34. Hockin NL, Mock T, Mulholland F, Kopriva S, Malin G. The response of diatom central carbon metabolism to nitrogen starvation is different from that of green algae and higher plants. Plant Physiol. 2012;158(1):299-312.

35. Sauer J, Dirmeier U, Forchhammer K. The *Synechococcus* strain PCC 7942 glnN product (glutamine synthetase III) helps recovery from prolonged nitrogen chlorosis. J Bacteriol. 2000;182(19):5615-9.
36. Guo X, Wang ZH, Liu L, Li Y. Transcriptome and metabolome analyses of cold and darkness-induced pellicle cysts of *Scrippsiella trochoidea*. Bmc Genomics. 2021;22:526.
37. Prihoda J, Tanaka A, de Paula WBM, Allen JF, Tirichine L, Bowler C. Chloroplast-mitochondria cross-talk in diatoms. J Exp Bot. 2012;63(4):1543-57.
38. Kroth PG, Chiovitti A, Gruber A, Martin-Jezequel V, Mock T, Parker MS, et al. A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis. PLoS One. 2008;3(1):):e1426.
39. Venuleo M, Raven JA, Giordano M. Intraspecific chemical communication in microalgae. New Phytol. 2017;215(2):516-30.
40. Waters CM, Bassler BL. Quorum sensing: cell-to-cell communication in bacteria. Annu Rev Cell Dev Biol. 2005;21(1):319-46.
41. Orefice I, Di Dato V, Sardo A, Lauritano C, Romano G. Lipid mediators in marine diatoms. Aquatic Ecol. 2022;56:377–397.
42. Fontana A, d'Ippolito G, Cutignano A, Miralto A, Ianora A, Romano G, et al. Chemistry of oxylipin pathways in marine diatoms. Pure Appl Chem. 2007;79(4):481-90.
43. Kotchoni SO, Jimenez-Lopez JC, Kayode APP, Gachomo EW, Baba-Moussa L. The soybean aldehyde dehydrogenase (ALDH) protein superfamily. Gene. 2012;495(2):128-33.
44. Hegab AE, Ha VL, Bisht B, Darmawan DO, Ooi AT, Zhang KX, et al. Aldehyde dehydrogenase activity enriches for proximal airway basal stem cells and promotes their proliferation. Stem Cells Dev. 2014;23(6):664-75.
45. Wang HL, Chen F, Mi TZ, Liu Q, Yu ZG, Zhen Y. Responses of marine diatom *Skeletonema marinoi* to nutrient deficiency: programmed cell death. Appl Environ Microbiol. 2020;86(3): e02460-19.
46. Nedelcu AM, Marcu O, Michod RE. Sex as a response to oxidative stress: a twofold increase in cellular reactive oxygen species activates sex genes. Proc. R. Soc. Lond. B. 2004;271:1591-6.
47. Nedelcu AM, Michod RE. Sex as a response to oxidative stress: the effect of antioxidants on sexual induction in a facultatively sexual lineage. Proc. R. Soc. Lond. B. 2003;270:136-9.
48. Vardi A, Berman-Frank I, Rozenberg T, Hadas O, Kaplan A, Levine A. Programmed cell death of the dinoflagellate *Peridinium gatunense* is mediated by CO2 limitation and oxidative stress. Current Biol. 1999;9(18):1061-4.
49. Winterbourn CC, Hampton MB. Thiol chemistry and specificity in redox signaling. Free Radical Biol Med. 2008;45(5):549-61.
50. Polyak K, Xia Y, Zweier JL, Kinzler KW, Vogelstein B. A model for p53-induced apoptosis. Nature. 1997;389(6648):300-5.
51. Vardi A, Formiggini F, Casotti R, De Martino A, Ribalet F, Miralto A, et al. A stress surveillance system based on calcium and nitric oxide in marine diatoms. Plos Biol. 2006;4(3):411-9.

52. Vardi A, Bidie KD, Kwityn C, Hirsh DJ, Thompson SM, Callow JA, et al. A diatom gene regulating nitric-oxide signaling and susceptibility to diatom-derived aldehydes. Current Biol. 2008;18(12):895-9.
53. Sun SY, Gresham D. Cellular quiescence in budding yeast. Yeast. 2021;38(1):12-29.
54. Liu XH, Wang LJ, Wu SC, Zhou L, Gao S, Xie XJ, et al. Formation of resting cells is accompanied with enrichment of ferritin in marine diatom *Phaeodactylum tricornutum*. Algal Research. 2022;61:102567.
55. Deng YY, Hu ZX, Shang LX, Peng QC, Tang YZ. Transcriptomic analyses of *Scrippsiella trochoidea* reveals processes regulating encystment and dormancy in the life cycle of a dinoflagellate, with a particular attention to the role of abscisic acid. Front Microbiol. 2017;8:2450.
56. Annunziata R, Balestra C, Marotta P, Ruggiero A, Manfellotto F, Benvenuto G, et al. An optimised method for intact nuclei isolation from diatoms. Sci Rep. 2021;11:1681.
57. Andrews S. Babraham bioinformatics-FastQC a quality control tool for high throughput sequence data. URL: https://www bioinformatics babraham ac uk/projects/fastqc. 2010.
58. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.
59. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nature Biotechnol. 2011;29(7):644.
60. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics. 2006;22(13):1658-9.
61. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10(1):1-9.
62. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30(9):1236-40.
63. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nature Methods. 2012;9(4):357-9.
64. Roberts A, Trapnell C, Donaghey J, Rinn JL, Pachter L. Improving RNA-Seq expression estimates by correcting for fragment bias. Genome Biol. 2011;12(3):1-14.
65. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):1-21.
66. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-40.
67. Bauer S, Grossmann S, Vingron M, Robinson PN. Ontologizer 2.0—a multifunctional tool for GO term enrichment analysis and data exploration. Bioinformatics. 2008;24(14):1650-1.
68. Gaonkar CC, Kooistra W, Lange CB, Montresor M, Sarno D. Two new species in the *Chaetoceros socialis* complex (Bacillariophyta): *C. sporotruncatus* and *C. dichatoensis*, and characterization of its relatives, *C. radicans* and *C. cinctus*. J Phycol. 2017;53(4):889-907.

69. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3-new capabilities and interfaces. Nucleic Acids Res. 2012;40(15):e115.
70. Pfaffl MW, Horgan GW, Dempfle L. Relative expression software tool (REST (c)) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res. 2002;30(9):e36.

Tables

Table 1: Number of differentially expressed genes (DEGs) and Gene Ontology (GO) enriched terms in the comparisons between treatments (T) and control (C).

Comparison	UP-REG DEGs	DOWN-REG DEGs	OVER- REP GO terms	DOWN-REP GO terms
T2 vs C2	4838	388	469	142
T3 vs C2	11704	3851	704	494
T4 vs C2	13851	4801	699	549

Table 2. The top 10 most over (A) and under (B) expressed genes at the different sampling points (T2, T3, T4) versus control (C2). FCs are reported in the last three columns.

A)

TRANSCRIPT_ID	SwissProt_ID	SwissProt_function	T2vsC2	T3vsC2	T4vsC2
TRINITY_DN4348_c0_g1_i3	Q9FGQ7	Cyclin-D3-2	5.11	13.57	14.31
TRINITY_DN725_c0_g1_i4	Q53FA7	Quinone oxidoreductase PIG3 (EC 1.-.-.-)	4.77	13.03	13.73
TRINITY_DN339_c0_g1_i5	Q7S045	Non-histone chromosomal protein 6	5.44	13.01	13.68
TRINITY_DN2312_c0_g1_i11	Q8BWU8	Ethanolamine-phosphate phospho-lyase (EC 4.2.3.2) (	4.76	12.99	14.16
TRINITY_DN1112_c0_g1_i2	P30041	Peroxiredoxin-6 (EC 1.11.1.27)	-	12.98	-
TRINITY_DN2429_c0_g1_i6	Q9NP80	Calcium-independent phospholipase A2-gamma (EC 3.1.1.5)	4.99	12.57	13.43
TRINITY_DN2494_c0_g1_i2	Q9FK72	Heat stress transcription factor A-4c (AtHsfA4c)	5.54	12.56	14.01
TRINITY_DN195_c0_g1_i4	Q9FM67	Protein TIC 20-v, chloroplastic	5.89	12.46	13.81
TRINITY_DN790_c0_g1_i6	F4K2K3	ARF guanine-nucleotide exchange factor GNL2 (	-	12.40	12.95
TRINITY_DN1629_c0_g1_i2	P40301	Proteasome subunit alpha type-2	5.39	12.09	12.67

B)

TRANSCRIPT_ID	SwissProt_ID	SwissProt_function	T2vsC2	T3vsC2	T4vsC2
TRINITY_DN247_c0_g1_i1	Q9LXH0	High affinity nitrate transporter 2.6 (AtNRT2:6)	-	-11.55	-12.20
TRINITY_DN2228_c0_g1_i3	O14283	Transcription factor prr1	-	-8.28	-4.52
TRINITY_DN22327_c0_g1_i7	B4JTF5	Protein hedgehog [Cleaved into: Protein hedgehog N-product; Protein hedgehog C-product]	-	-8.27	-6.87
TRINITY_DN7356_c0_g1_i1	Q02073	20 kDa chaperonin, chloroplastic (Chaperonin 10)	-	-8.26	-2.17
TRINITY_DN80_c0_g1_i2	Q7Z8P9	Nucleoside diphosphate kinase (NDK) (NDP kinase) (EC 2.7.4.6)	-	-8.03	-3.52
TRINITY_DN32657_c0_g1_i1	P49534	Uncharacterized protein ycf39 (EC 1.-.-.-) (ORF319)	-	-7.71	-3.28
TRINITY_DN20497_c0_g1_i1	Q9LYR5	Peptidyl-prolyl cis-trans isomerase FKBP19, chloroplastic (PPIase FKBP19) (EC 5.2.1.8)	-	-7.17	-3.84
TRINITY_DN4637_c0_g1_i6	O80832	UPF0187 protein At2g45870, chloroplastic	-	-7.17	-3.23
TRINITY_DN22327_c0_g1_i3	Q98862	Indian hedgehog B protein (IHHB)	-	-7.09	-6.44
TRINITY_DN262_c0_g1_i3	Q4ING3	Cytochrome c peroxidase, mitochondrial (CCP) (EC 1.11.1.5)	-	-7.02	-3.90

Figures

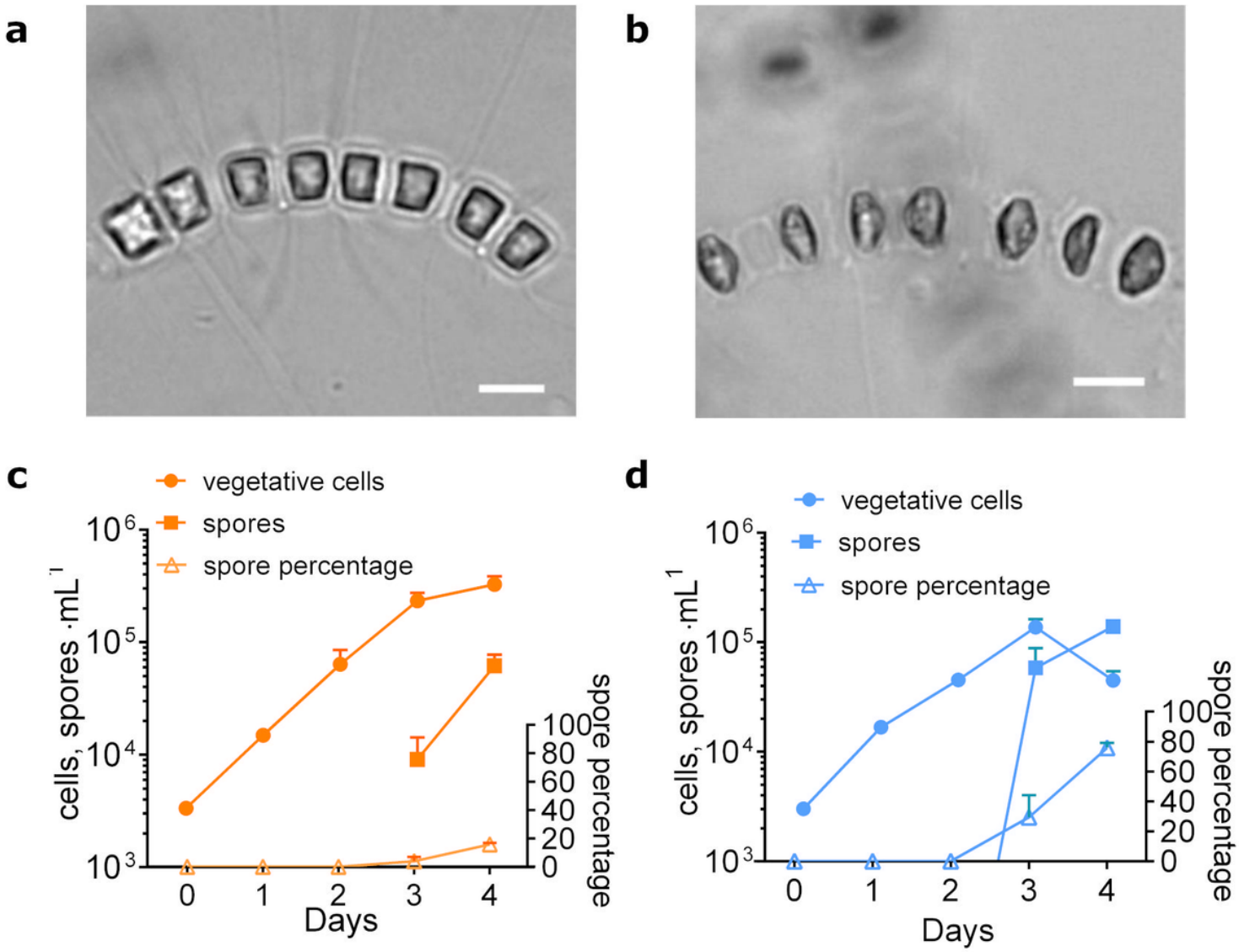

Figure 1

Nitrogen depletion induces the formation of resting spores in the diatom Chaetoceros socialis. Light micrograph of a colony of vegetative cells (**a**) and a colony in which vegetative cells turned into resting spores (**b**); scales bars=10 µm. Time course concentration of vegetative cells and spores (cells, spores mL^{-1}; left axis) and percentage of spores (right axis) in cultures grown in nitrogen deplete treatment (**c**) and in nutrient replete control (**d**); data are shown as average ± S.D. (n=3). Samples for the differential gene expression analysis were collected on day 2 (T2), 3 (T3) and 4 (T4) for the treatment and on day 2 in the control.

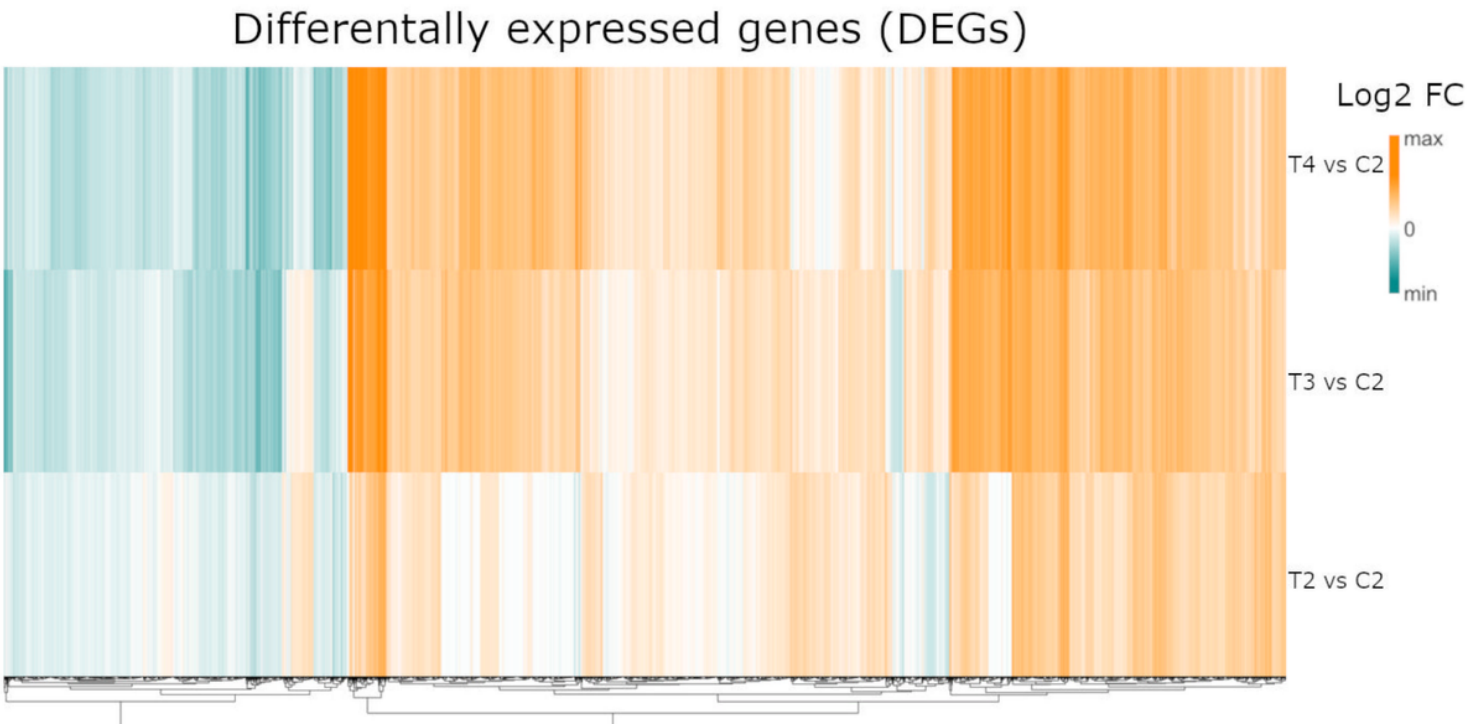

Figure 2

Hierarchical clustering analysis of FCs expression. The heatmap represents genes displaying statistical significance (adjusted P-value ≤ 0.05) for differential expression between sampling points. The up-regulated genes are reported in orange, the down-regulated genes in blue and the genes not differentially expressed genes in white.

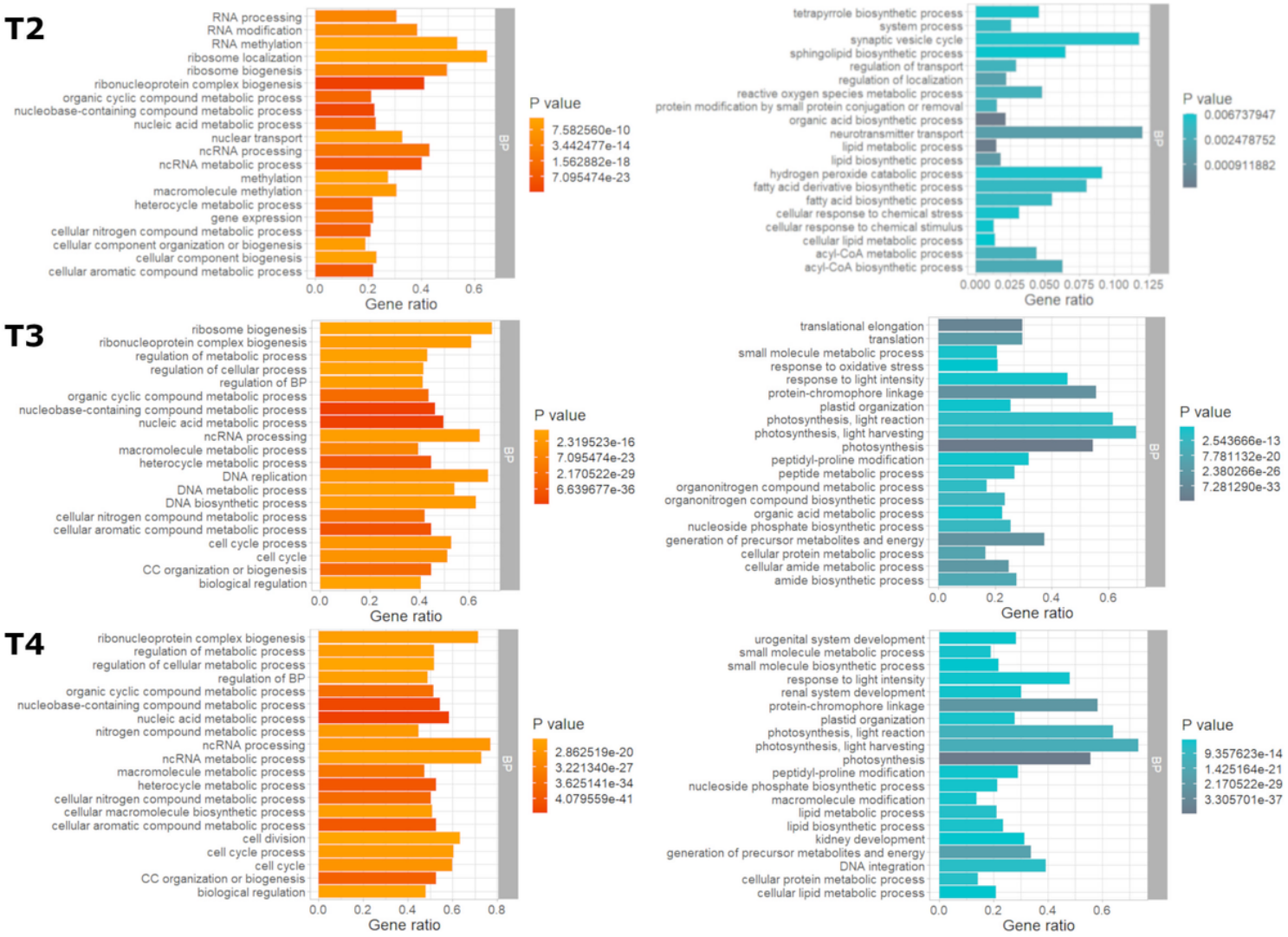

Figure 3

The first 20 biological processes related to GOs terms. Positively (right side) and negatively (left side) enriched terms at T2, T3 and T4 compared to control conditions are reported. The color scale of bars indicates P. values, their length indicates gene ratio i.e., the number of genes found enriched in each comparison over all those annotated under the same term in the transcriptome. Note that the color scale differs among comparisons because of the wide differences in their P. values.

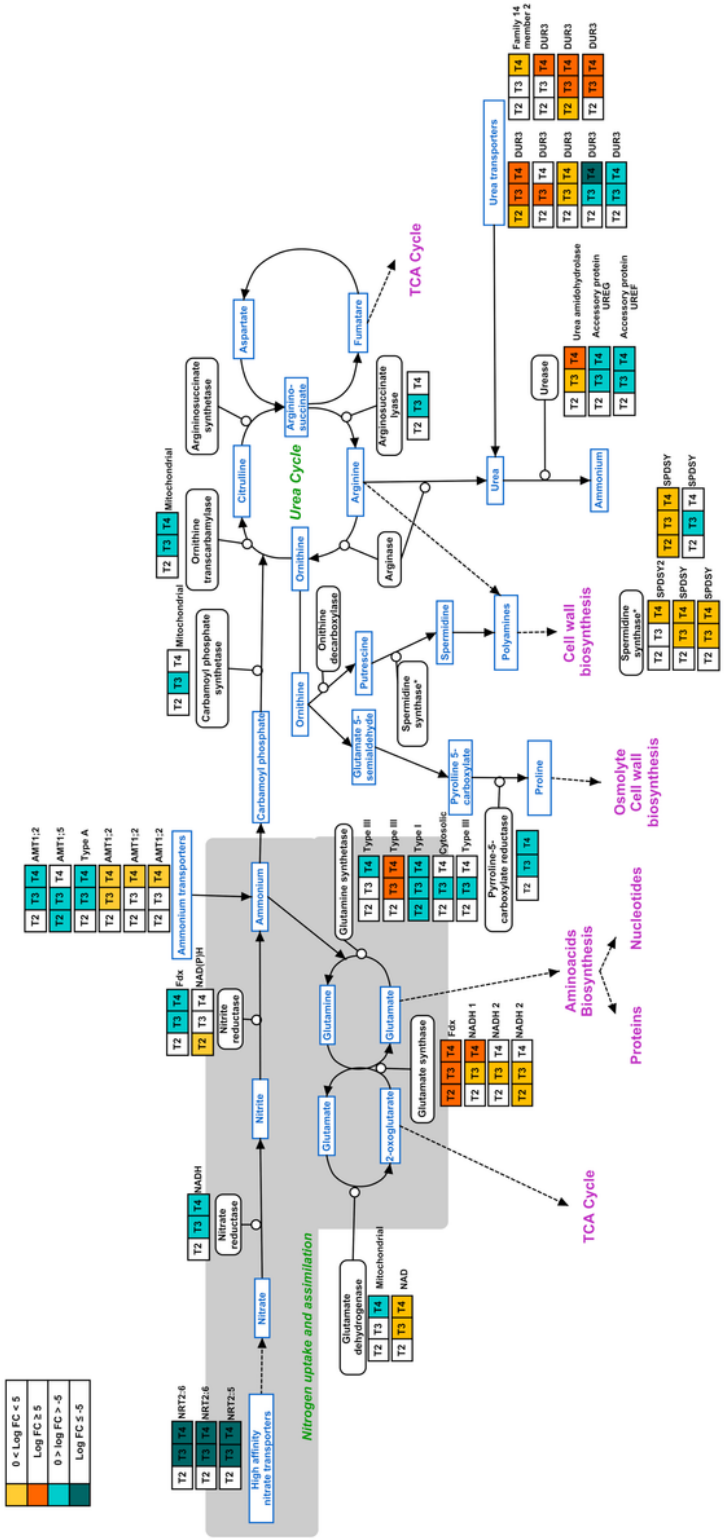

Figure 4

DEGs related to nitrogen metabolism during spore formation in *C. socialis*. Box colors indicate the relative FC detected in the comparisons between T2 vs C2, T3 vs C2, T4 vs C2; orange and blue indicate up- and down- regulation, respectively. The additional information obtained by SwissProt annotation is reported on the right side of each box. .

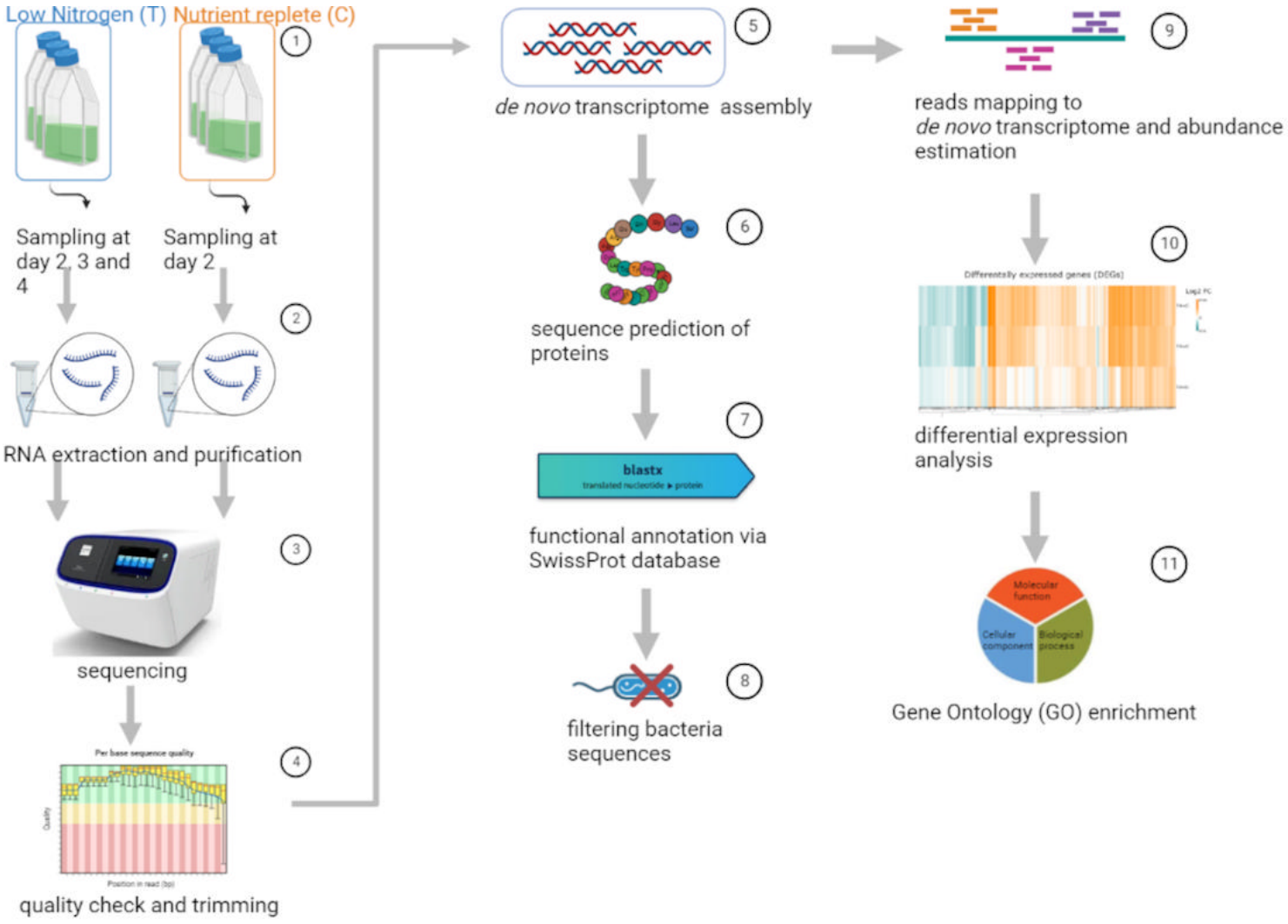

Figure 5

Schematic representation of the bioinformatic pipeline used in this study.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.xls
- Additionalfile2.xls
- Additionalfile3.xls
- Additionalfile4.xls
- Additionalfile5.xls
- Additionalfile6.xls
- Additionalfile7.xls
- Additionalfile8.tif

- Listofadditionalfiles.docx