Bioresource Technology 244 (2017) 1281–1293

Contents lists available at ScienceDirect

Bioresource Technology

journal homepage: www.elsevier.com/locate/biortech

RNA-Seq and metabolic flux analysis of *Tetraselmis* sp. M8 during nitrogen starvation reveals a two-stage lipid accumulation mechanism

David K.Y. Lim [a], Holger Schuhmann [a], Skye R. Thomas-Hall [a], Kenneth C.K. Chan [a,d], Taylor J. Wass [a], Felipe Aguilera [b], T. Catalina Adarme-Vega [a], Cristiana G.O. Dal'Molin [c], Glen J. Thorpe [c], Jacqueline Batley [a,d], David Edwards [a,d], Peer M. Schenk [a,*]

[a] School of Agriculture and Food Sciences, The University of Queensland, Brisbane 4072, Australia
[b] School of Biological Sciences, The University of Queensland, Brisbane 4072, Australia
[c] Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, Brisbane 4072, Australia
[d] School of Plant Biology, The University of Western Australia, Perth 6009, Australia

HIGHLIGHTS

- *Tetraselmis* RNA-Seq was performed at early lipid accumulation after N deprivation.
- qRT-PCR analyses revealed distinct phases for lipid biosynthesis and catabolism.
- Gene expression data was combined with metabolic reconstruction modeling.
- *Tetraselmis* shifts from reduced lipid consumption to active lipid production.
- This process appears independent from *DGAT* expression.

ARTICLE INFO

Article history:
Received 30 March 2017
Received in revised form 31 May 2017
Accepted 1 June 2017
Available online 6 June 2017

Keywords:
Beta-oxidation
Metabolic engineering
Microalgae
RNA-Seq
TAG biosynthesis

ABSTRACT

To map out key lipid-related pathways that lead to rapid triacylglyceride accumulation in oleaginous microalgae, RNA-Seq was performed with *Tetraselmis* sp. M8 at 24 h after exhaustion of exogenous nitrogen to reveal molecular changes during early stationary phase. Further gene expression profiling by quantitative real-time PCR at 16–72 h revealed a distinct shift in expression of the fatty acid/triacylglyceride biosynthesis and β-oxidation pathways, when cells transitioned from log-phase into early-stationary and stationary phase. Metabolic reconstruction modeling combined with real-time PCR and RNA-Seq gene expression data indicates that the increased lipid accumulation is a result of a decrease in lipid catabolism during the early-stationary phase combined with increased metabolic fluxes in lipid biosynthesis during the stationary phase. During these two stages, *Tetraselmis* shifts from reduced lipid consumption to active lipid production. This process appears to be independent from *DGAT* expression, a key gene for lipid accumulation in microalgae.

1. Introduction

Currently, microalgae are considered as one of the most promising feedstocks for oil production. Under the appropriate conditions (e.g. nutrient deprivation), oleaginous microalgae can be induced to accumulate neutral lipids or triacylglycerides (Hu et al., 2008), which, among other applications, can be converted into biodiesel via transesterification. Microalgae can produce significantly more lipids than oil palms (Ahmad et al., 2011), and without competing for precious arable land, biodiverse landscapes (e.g. rainforests) or freshwater resources. Despite their potential, the high cost of large-scale production still needs to be reduced in order for microalgal biofuel to achieve full commercialization and wide-scale use. Currently, algal strain development remains one of the most important aspects of microalgae-for-biofuel development. Research efforts are continuously advancing bioprospecting (Nascimento et al., 2013), selective breeding (Zayadan et al., 2014) and genetic engineering (Gimpel et al., 2013) of microalgae in an effort to maximize growth and lipid accumulation of the highest performing strains. Importantly, several lipid induction techniques have been identified in microalgae (Rodolfi et al., 2009; Sharma et al., 2012). Microalgae typically reduce cell divi-

* Corresponding author.
E-mail address: p.schenk@uq.edu.au (P.M. Schenk).

http://dx.doi.org/10.1016/j.biortech.2017.06.003
0960-8524/

sion during adverse conditions, such as nutrient starvation or UV radiation, but are still able to accumulate starch or lipids during photosynthesis as a survival mechanism (Timmins et al., 2009; Wang et al., 2009; Sharma et al., 2014).

Metabolic engineering via genetic modification or modulation of cultivation techniques provides a promising area for increased lipid accumulation. For example, recently the overexpression of diacylglycerol acyltransferase (DGAT)-encoding genes in *Tetraselmis chui* and *Phaeodactylum tricornutum* has resulted in 2-fold increases in triacylglyceride content (Úbeda-Mínguez et al., 2017; Dinamarca et al., 2017). Metabolic pathway engineering can be greatly assisted by comprehensive genomic, transcriptomic, proteomic and metabolomic knowledge. For example, key lipid-related pathways must be mapped out, and important bottleneck enzymes and their genes identified as targets for manipulation. To that effect, global transcriptional profiling of microalgal cells during lipid accumulation enables the identification of the underlying transcriptional networks. Even without pre-existing reference genomes, comparative transcriptomics analyses have been used in microalgae to successfully identify pathways and observe changes during induced lipid accumulation (Rismani-Yazdi et al., 2011; Radakovits et al., 2012; Sun et al., 2013). In most studies, the focus has been on metabolic pathway reconstruction and gene discovery at a single time-point, usually 48–96 h into starvation phase when lipid accumulation is at its peak (Guarnieri et al., 2011; Rismani-Yazdi et al., 2012; Sun et al., 2013). While this approach successfully allowed for the reconstruction of fatty acid (FA), triacylglyceride (TAG), β-oxidation and other metabolic pathways, the limited scope of these studies restricts our understanding how the expression of these pathways change, particularly during early stationary phase as cells transition from growth phase into starvation phase. The few studies that have monitored the transcriptional profile of microalgae at various growth stages have observed more transcriptional changes during early-stationary phase compared to stationary phase. These changes occur particularly in photosynthesis, carbon and lipid synthesis pathways, and can be linked to physiological changes (e.g. reduced cell division and increased lipid synthesis) observed during that phase (Valenzuela et al., 2012; Lv et al., 2013). The appropriate lipid induction conditions and time point of RNA sampling are crucial in obtaining distinct expression profiles between control and treatments cultures. Nitrogen depletion is a commonly used method to induce lipid accumulation in microalgae (Hu et al., 2008; Rodolfi et al., 2009; Miller et al., 2010).

The flagellate green microalga *Tetraselmis* sp. is widely mentioned in the literature, but little functionally annotated sequence information is available on this genus in public databases, apart from a recent study on temperature tolerance that used *de novo* transcript assembly (Shin et al., 2016). *Tetraselmis* sp. (Prasinophyceae) presents a good model organism, based on its reported ability to accumulate high lipid content as well as its robustness to tolerate a range of environmental conditions (Chini Zitelli et al., 2006; Rodolfi et al., 2009). *Tetraselmis* sp. M8 cells accumulate approximately 20–30% lipids and it has been shown that they lose their flagella during stressful conditions, resulting in rapid settling, a feature that can significantly reduce harvesting/dewatering costs and provide an avenue for commercial production (Lim et al., 2012; Sharma et al., 2014). The growth characteristics of *Tetraselmis* sp. M8 strain and its lipid accumulation capability and composition were previously found suitable, in principle, for biodiesel production under both laboratory and outdoor cultivation conditions (Lim et al., 2012; Sharma et al., 2014; Narala et al., 2016).

Even in the absence of a fully sequenced and annotated genome, transcriptomic analysis by microarrays or RNA-Seq can provide a powerful tool to improve our understanding of the underlying physiological networks that allow microalgae to respond to environmental changes (Rismani-Yazdi et al., 2012; Valenzuela et al., 2012; Sun et al., 2013). In the present study, we gain insights into the lipid accumulation mechanism of the genus *Tetraselmis*, particularly the expression of genes in the FA synthesis, TAG synthesis and β-oxidation oxidation pathways, as cells transition from growth phase into stationary phase. Physiological observations such as growth, lipid accumulation and FA profiles were linked to transcriptional data obtained first by global transcriptomic sequencing, followed by quantitative reverse transcriptase PCR (qRT-PCR) time-course analysis of each of the aforementioned pathways, and metabolic reconstruction modeling.

2. Materials and methods

2.1. Culture growth conditions

To detect changes in lipid-related pathways as cells transition from continuous exponential growth in log phase to stationary phase, it was important that RNA sampling was carried out on concurrently-grown control cultures that were maintained in log phase. Therefore, semi-continuous cultures of *Tetraselmis* sp. M8 were first established to maintain cells under constant nutrient-replete conditions and exponential growth phase before the start of each experiment. This way, cells could be maintained in constant growth phase and cell density by feeding and dilution in a constant cycle until the start of experiment. Three 1 L-master cultures were maintained by replacing half the culture (500 mL each) with autoclaved 25 PSU artificial seawater (Aquasonic) supplemented with F/2 medium (Guillard and Ryther, 1962; enriched with an additional 100 μM of phosphate) every 48 h for 2 weeks. The cultures were grown in 1 L-Schott bottles with constant bubbling at 24 °C under 16:8 light/dark photoperiod of fluorescent white lights (80 μmol photons $m^{-2}s^{-1}$). For RNA-Seq, semi-continuous cultures were maintained with a regime as above. At the start of the experiment, master cultures were mixed and distributed to nine cultures (three cultures per treatment). Nitrogen-deprived and phosphate-deprived cultures had media replaced with nitrogen-deficient or phosphate-deficient F/2 medium to induce lipid production, while control cultures received N/P-replete medium. For the time course experiment, the semi-continuous cultures were maintained by diluting to 5×10^5 cells/mL and feeding with F medium (enriched with an additional 100 μM of phosphate) every 48 h. Full strength F medium was used as larger differences in nutrient levels between treatments and were expected to lead to more pronounced lipid induction. At the start of the experiment the nitrogen-starvation treatment was supplied by replacing with nitrogen deficient F-medium. In both experiments, the nitrogen and phosphate concentration of the cultures were measured daily to ensure nutrient-deplete conditions only occurred at 48 h after feeding (Supplementary Fig. 1). Nitrogen use of the medium is part of algae cultivation and therefore also occurred in the control cultures, but not to a point that cells had run out of nitrogen reserves. The dilution and feeding regime was altered in the time course experiment to reduce the duration in which cultures experienced nutrient-deplete conditions at 48 h before feeding.

2.2. Physiological parameter analysis

During the course of the experiments, various physiological parameters such as cell density, Nile red fluorescence, nitrate and phosphate concentration, chlorophyll *a* and *b*, and fatty acid (FA) content were measured. Total nitrate and phosphate contents in the media were measured as described by Adarme-Vega et al. (2014) using API Aquarium pharmaceutical Nitrate NO^{3-} and Phos-

phate PO_4^{3-} test kits with absorbance measurements taken on a spectrophotometer (Hitachi U-2800 UV-VIS) at 545 nm and 690 nm, respectively. For chlorophyll extraction, 90% acetone and glass beads were added to a 5 mL microalgal pellet (extracted via centrifugation, 10000*g*, 7 min) and then vortexed for 3 min before being stored in the dark at 4 °C for 2 h. Cellular debris was then pelleted (500*g*, 20 min) and the optical density (OD) of the acetone supernatant was measured on a spectrophotometer at 664 nm, 647 nm and 630 nm. The calculations for the concentration of chlorophyll *a* and *b* were performed as described by Franson and Eaton (2005).

For lipid accumulation measurements, 1 mL of culture was stained with 6 μL of Nile red in DMSO solution (250 mg/mL). Samples were then gently vortexed and incubated in the dark for 10 min. 200 μL and was loaded into a 96 well-microtiter plate (Sarsted) in triplicates. Yellow-gold fluorescence was then measured on a POLARstar OPTIMA (BMG Labtech) plate reader using fluorescence intensity mode. Gain was set at 3000, with excitation and emission wavelengths of 485 nm and 590 nm selected. Specific fluorescence was obtained by dividing the Nile red fluorescence intensity by the cell number. Cell density was monitored via cell counts using a haemocytometer. FA profiles were generated using GC/MS by Metabolomics Australia as described previously (Lim et al., 2012), with the exception that 5 mL of culture was used, instead of 4 mL culture.

2.3. Microscopic analyses

Cells were stained with Nile red (250 μg/mL) at 24 h after nitrogen and phosphorus deplete media were added to cultures. Photographs were taken using an Olympus BX60 microscope and an Olympus DP50 digital camera. Epifluorescent (excitation: 510–550 nm; emission: 590 nm) images were captured at 400x magnification.

2.4. RNA sampling and extraction

Sampling for RNA was performed on control and nitrogen-starved cultures at the desired time-points of 24 h for the RNA-Seq experiments and at 16, 24, 32, 48 and 72 h for the time-course experiment. At these time points, 10 mL of culture was collected by centrifugation (10000*g*, 7 min) from each replicate; the supernatant was discarded and the collected cell pellets immediately flash-frozen with liquid nitrogen and stored at −80 °C. Just before RNA extraction, cell pellets were resuspended in lysis buffer (SV Total RNA Isolation System, Promega) and then ground using a micro pestle. Total RNA was then extracted following the manufacturer's instructions, with the exception that the incubation at 70 °C was done at room temperature. Total RNA was not pooled but kept as respective replicates and then stored at −80 °C.

2.5. cDNA library construction and sequencing

cDNA libraries were made from replicates of the RNA-seq experiment, following the TruSEQ RNA V2 kit protocol; each replicate with their own adapters to barcode samples. cDNA products were then quantified on a Qubit ® 2.0 Fluorometer (Invitrogen) and checked for quality on a Bioanalyzer 2100 (Agilent). 151 bp paired-end sequencing of the cDNA libraries was then performed on an Illumina Mi-Seq platform using standard manufacturer protocols. Libraries of the same treatment were pooled together, with each treatment being sequenced on a separate run. Sequencing data is deposited and publically available at https://www.ncbi.nlm.nih.gov/bioproject/PRJNA389671.

2.6. Sequence analysis and Differential Kmer Analysis Pipeline (DiffKAP)

Standard RNA-Seq analysis relies on mapping individual short sequence reads to a reference genome or transcriptome and then applying statistical tests to identify differentially expressed genes. We attempted to apply the popular Tophat/Cufflinks (Trapnell et al., 2012) method on the RNA-Seq data using the high quality annotated genome of the chlorophyte *Chlamydomonas reinhardtii* (JGI v5.5), as the reference. However, the analysis was unsuccessful as less than 0.02% reads were mapped. Therefore, we applied the Differential Kmer Analysis Pipeline (DiffKAP) that enabled identification of differentially expressed reads (DERs) in RNA-Seq data between the two treatments without using a reference (Rosic et al., 2014). These DERs can be annotated to public databases like Swiss-Prot, enabling functional classification of differentially expressed sequences. The DiffKAP program is available from http://appliedbioinformatics.com.au/index.php/DiffKAP.

The DiffKAP pipeline consists of six steps, as shown in Supplementary Fig. 2. DiffKAP uses Jellyfish (Marçais and Kingsford, 2011) to perform k-mer counting, and automatically determines the optimal k-mer size by finding the 'knee point' in a k-mer uniqueness graph (Kurtz et al., 2008). For this study, an optimal k-mer size of 17 was found to give the best balance between the specificity and sensitivity of the information content. The abundance of each k-mer was normalized by dataset size, and differentially expressed k-mers (DEKs) were determined using the following formula, where k is the query k-mer, c_1 and c_2 represent the normalized k-mer occurrence in datasets 1 and 2, respectively, X represents the minimum difference of the k-mer occurrence and Y is the minimum fold change of k-mer occurrence between the two datasets required to call a k-mer as differentially expressed. In this study, the minimum difference (X) used was 3, while the minimum fold change (Y) used was 2.

$$(k = DEK) \iff \left(|c_1 - c_2| \geqslant X \bigvee \left(Y \leqslant \frac{c_2}{c_1} \bigwedge Y \leqslant \frac{c_1}{c_2}\right)\right)$$

A set of unique reads was obtained by filtering duplicate reads within the datasets. DERs were determined by a strict criterion to minimize false positives and defined as when all constituent k-mers in the read are differentially expressed k-mers (DEK), where r is the query read and $\mathbb{K}$ denotes all constituent k-mers in r.

$$(r = DER) \iff (\forall k \in \mathbb{K} = DEK)$$

For each DER, the median k-mer abundance was calculated for each of the two datasets, and the ratio of median k-mer abundance (RoM) provided as a prediction of relative expression to the control. These reads were then categorized based on their expression ratios following the rules below:

1. $RoM = 0$: Only present in nitrogen-starved
2. $0 < RoM < 0.5$: Highly induced in nitrogen-starved
3. $0.5 \leq RoM \leq 2$: Not differentially expressed
4. $2 < RoM < \infty$: Highly induced in control
5. $RoM = \infty$: Only present in control

All DERs were annotated by comparison with the Swiss-Prot database (Boutet et al., 2007), with an e-value of 10^{-16}. If DERs were annotated to the same Swiss-Prot entry, the relative expression level was considered as the median RoM of all DERs which mapped to the respective database entry.

2.7. Functional annotation and pathway assignments

Successfully-annotated DERs from both treatments were fed into BLAST2GO software in order to assign associated gene ontology (GO) terms, with an annotation cutoff of 55, and GO weight of 5 (Gotz et al., 2008). Genes in FA synthesis, TAG synthesis and lipid catabolism pathways were identified by BLAST matching the *Tetraselmis* sp. M8 transcriptome that was first uploaded into TAGdb (Marshall et al., 2010) using reference sequences obtained from DiffKAP or NCBI. Similar reads from the BLAST results were also extracted from TAGdb for primer design. Automated gene assembly was not carried out to avoid the generation of potentially false contig sequences that could then be wrongly used by other studies for further assembly of other sequence data. Gene assembly for individual genes was carried out for qRT-PCR primer design (see below).

2.8. Quantitative reverse transcriptase real-time PCR

For optimal primer design, reads extracted from TAGdb were assembled to the initial reference gene using Geneious software to yield a consensus sequence, which was then used as the next reference sequence for TAGdb to BLAST the transcriptome for more similar reads for assembly. This procedure was repeated until a sequence length of more than 500 bp was obtained. This sequence was then used for Primer Express to generate a primer pair. The full primer list in shown in Supplementary Table 1.

Extracted RNA from all replicates and experiments were used for cDNA synthesis using Superscript III reverse transcriptase (Invitrogen) following the manufacturer's instructions. For qRT-PCR, each reaction was performed in a final volume of 10 μL, and contained 1 μl cDNA (10 ng/μL), 1 μL of each primer (1 μM), 5 μL SYBR Green using the 7900 HT Fast Real-time PCR system (Applied Biosystems). Thermal cycling conditions consisted of 10 min at 95 °C and 45 cycles of 15 s at 95 °C and 1 min at 60 °C prior to 2 min at 25 °C. Transcript levels were normalized to the expression of β-*ACTIN* gene.

2.9. Metabolic flux modeling

This study used the genome-scale model of *Chlamydomonas reinhardtii* AlgaGEM (de Oliveira Dal'Molin et al. 2011) coded for use with COBRA Toolbox in MATLab. Using time series data of growth and fatty acid content of *Tetraselmis* sp. M8 (Fig. 1) as constraints for the growth rate and biomass composition, Flux Balance Analysis (FBA) was performed under the assumption of the optimality condition being energy efficient growth (i.e. minimized photon or acetate usage depending on growth phase). This is a well-studied optimality criterion for use in plants and photosynthetic microorganisms, and states that the cell will attempt to fulfil the constraints placed upon the model by using the least amount of photons or carbon source. Further, Flux Variability Analysis (Gudmundsson and Thiele, 2010) was used to determine the minimum and maximum metabolic capacity of the network under the given conditions, which could then be used to analyse how well-constrained the FBA solutions may be. Gene expression data for ACCase was later integrated as a constraint by multiplying the predicted flux through reaction R01280_x in the control condition by

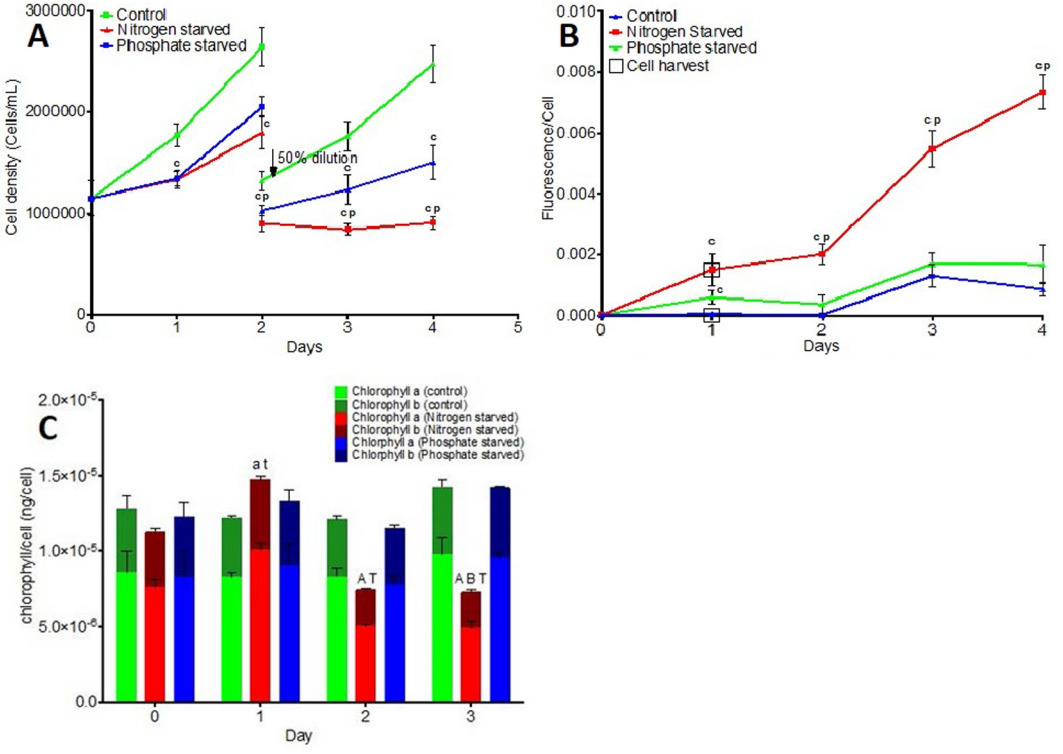


Fig. 1. Growth, lipid accumulation and chlorophyll content of *Tetraselmis* sp. M8 under nitrogen and phosphate-starved conditions. A. Cell density growth curve of *Tetraselmis* sp. M8, cultured semi-continuously with dilutions every 2 days, shows reduced cell growth in nutrient deplete cultures. B. Characterization of lipid accumulation via Nile red fluorescence (shown as intensity per cell number) displays increased lipid accumulation under nitrogen-starved conditions. Boxed data points indicate time and treatments at which cells were sampled for RNA sequencing. "c" and "p" indicate significant differences from control and phosphate-starved treatments, respectively (Student's *T*-test, $P < 0.05$). C. Characterization of chlorophyll content (*a*, *b* and total) per cell number shows reduced chlorophyll content after 48 h of nitrogen starvation. "a" indicates significant differences of chlorophyll *a* content to control. "A" indicates significant differences of chlorophyll *a* content to both control and phosphate-starved treatment. "B" indicates significant differences of chlorophyll *b* content to both control and phosphate-starved treatments. "T" indicates significant differences of total chlorophyll content to both control and phosphate-starved treatments.

the DiffKAP estimated fold-change (−4.1), which was used to investigate a claim regarding the effect of β-oxidation on lipid production during early stationary phase.

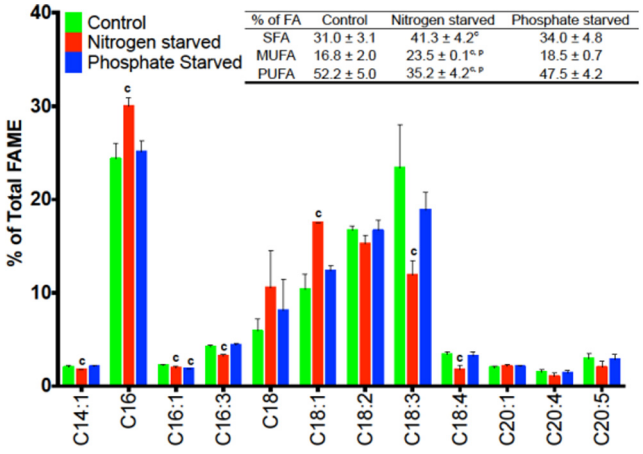
3. Results and discussion

3.1. Nitrogen rather than phosphate deprivation leads to lipid biosynthesis in Tetraselmis sp. M8

Nutrient deprivation is a well-established method for inducing lipid biosynthesis in microalgae and it was used to determine whether nitrogen or phosphorus starvation have the strongest inductive effect in *Tetraselmis* sp. M8. Semi-continuous cultures of *Tetraselmis* sp. M8 were allowed to grow in parallel for three feeding cycles to adapt the cultures to exponential growth with regular nutrient supply. Nitrogen and phosphorus deprivation treatments were then initiated by feeding with nitrogen- and phosphate-deficient F/2 medium, respectively. Physiological parameters such as cell density, lipid accumulation and chlorophyll content were monitored for the duration of the experiment using replicates from three separately-grown cultures for each treatment (Fig. 1). The mock-treated control cultures displayed the highest cell density before and after dilution on day 2. Both nitrogen and phosphate-starved cultures were found to have significantly reduced cell accumulation from day 1 onwards (Fig. 1A, $P < 0.05$). Phosphate-starved cultures reduced growth rates when compared to control cultures but accumulated cells throughout the experiment, while nitrogen-starved cultures only underwent cell divisions until day 2. Lipid accumulation was observed via measurement of Nile red fluorescence per cell number of the cultures (Fig. 1B). Nitrogen-starved cultures showed significantly higher Nile red fluorescence than the other treatments from day 1 onwards ($P < 0.05$), with a marked increase particularly after day 2 of starvation. However, Nile red fluorescence for phosphate-starved cultures was only significantly higher than the controls on day 1, and had similar fluorescence levels on subsequent days, suggesting that N, rather than P starvation have the strongest effect on lipid accumulation in *Tetraselmis* sp. M8. Microscopic analysis of day 1-cultures also confirmed that nitrogen-deprived cells have the largest lipid bodies, followed by phosphate-deprived and control cells, which showed very small lipid bodies.

This is similar to many other microalgae (Rodolfi et al., 2009; Miller et al., 2010; Rismani-Yazdi et al., 2012; Valenzuela et al., 2012) and shows that the transition to starvation phase, detected via a significant increase in lipid accumulation compared to control, was as early as 16 h after exhaustion of exogenous nitrogen. Chlorophyll content was also measured under the different nutrient deprivation treatments (Fig. 1C). Compared to other treatments, N starvation led to significant increases in chlorophyll *a* on day 1, but was followed by significant decreases on day 2 and an overall decrease in chlorophyll content (*a* and *b*) on day 3.

N-deprived cultures were found to undergo just one doubling period within the first 48 h, after which cell growth ceased and the rate of lipid accumulation significantly increased. This sudden halt in growth and increase in lipid content after 1 doubling period also coincides with a significant decrease in chlorophyll *a* content (Fig. 1C), and may indicate the depletion of internal nitrogen stores and the transition from early-starvation to starvation phase. Changes in these physiological parameters during the transition from early-stationary to stationary phase within the first 48 h have also been observed in *Botryosphaerella sudeticus* (Sun et al., 2013), *Phaeodactylum tricornutum* (Valenzuela et al., 2012) and *Neochloris oleoabundans* (Rismani-Yazdi et al., 2012).



% of FA	Control	Nitrogen starved	Phosphate starved
SFA	31.0 ± 3.1	41.3 ± 4.2[c]	34.0 ± 4.8
MUFA	16.8 ± 2.0	23.5 ± 0.1[c,p]	18.5 ± 0.7
PUFA	52.2 ± 5.0	35.2 ± 4.2[c,p]	47.5 ± 4.2

Fig. 2. Fatty acid profile of control and nutrient deficient culture at day 1 of the RNA-Seq experiment. Inserted table shows the percentage of saturated FA (SFA), monounsaturated FA (MUFA) and polyunsaturated FA (PUFA). "c" indicates significant differences between nitrogen-starved and control cultures. (Student's *T*-test; $P < 0.05$).

GC/MS analyses were performed on day 1 after treatments to determine FA composition. Nitrogen-starved cultures exhibited significantly higher saturated (SFA) and monounsaturated FAs (MUFAs) composition compared to controls, while polyunsaturated FAs (PUFAs) were significantly lower ($P < 0.05$, Fig. 2). C16 and C18:1 FAs showed the most significant increases, 5.6 % and 7.1% respectively, while C18:3 decreased the most, 11.5%. It was also found that while there is a general increase in SFAs and MUFAs, there is significant decreases in C14:1 and C16:1 FAs, indicating that *Tetraselmis* sp. M8 could be storing lipids in the form of C16 and C18:1.

3.2. Under nitrogen starvation conditions metabolic flux through lipid biosynthesis pathway is increased

To investigate the endocellular fluxes under conditions of nitrogen starvation/lipid accumulation, the genome-scale model AlgaGEM (de Oliveira Dal'Molin et al. 2011) has been used to simulate lipid accumulation with particular interest in the effect of beta-oxidation on lipid production. To investigate the ability of AlgaGEM to predict fatty acid production in *Tetraselmis*, a fatty acid biomass drain for hexadecanoic acid (R01706_p) was added to the model to allow for fatty acid accumulation in excess of the model's stoichiometric incorporation of hexadecanoic acid into the predicted biomass. The degree of predicted flux through this reaction was used to infer the amount of fatty acid accumulated in the biomass composition, and thus the changes induced by added transcriptomic constraints. Using growth rates calculated from experimental data, FBA was used to simulate flux through the model under both control and nitrogen deprived conditions. The control simulation ran under the assumption that the algal cell would seek to optimize use of absorbed light, such that photon usage is most efficient. The nitrogen deprivation model was then simulated under the assumption that when subjected to stress of this nature, the cell would seek to maximize lipid production, as seen experimentally in Fig. 1B, subject to the previous condition of minimizing photon usage.

The periods from 8–16 h, and 32–40 h after N deprivation were determined to be under heterotrophic growth (night phases), such that the points at 8 and 32 h are modelled using acetate as the carbon/energy source. All other points were modelled using autotrophic growth conditions. Under nitrogen-deprivation

conditions, the rate of lipid production as a percentage of biomass increases compared to control conditions (Fig. 1B). Flux Variability Analysis was then used to determine how tightly-constrained the fatty acid flux was under the nitrogen-deprivation condition. A low variance (±10%) would indicate that the simulated lipid flux is reliable, while a high variance would indicate that the simulation needs to be better constrained.

The flux range was tightly constrained to the predicted metabolic fluxes of the model, which suggests that the minimal constraints used were sufficient for generating a reliable flux prediction. After determining that the model was sufficiently well constrained, further investigation of the metabolic consequences of imposing the growth constraints was undertaken. The control and nitrogen deprivation fluxes were then compared across key reactions, which could indicate possible target reactions for metabolic engineering. Based on model highlights, a flux map was then constructed of the central carbon metabolism of microalgae under the lipid accumulation conditions measured for *Tetraselmis* sp. M8 under nitrogen deprivation at 72 h (Supplementary Fig. 3). The flux map suggests that the plastidic glycolysis pathway and lipid biosynthesis pathway should be both upregulated, while the TCA cycle should be downregulated. To compare whether these predicted flux data match with experimental data, transcriptomics of treated and control cultures were carried out.

3.3. RNA-Seq of Tetraselmis sp. M8 cells reveals distinct sequences compared to other known microalgal sequences

To determine which genes would be required for N deprivation-induced lipid accumulation in *Tetraselmis* sp. M8, cell sampling for RNA sequencing was performed on day 1 control and nitrogen-starved cultures (replicates from three separately-grown cultures each). This time point was selected because it was the earliest time point to exhibit significant lipid accumulation. The concentration of cDNA was normalized and then pooled for each treatment. Sequencing using the Illumina Mi-Seq platform of the cDNA libraries produced approximately 36,000,000 reads per treatment, at an average length of 151 bp per read. Initially, Tophat/Cufflinks was used to assemble the RNA-Seq data using the closest available genome, *Chlamydomonas reinhardtii*. However, the analysis was unsuccessful with less than 0.02% reads mapped and so this approach was abandoned. The Differential Kmer Analysis Pipeline (DiffKAP) approach was then used to identify Differentially Expressed Reads (DERs) between control and N-starved treatments (see Materials and Methods for details). A total of 990,249 DERs were identified as higher expressed in the controls (also considered as down-regulated in N-deprived samples), while 1,046,741 DERs were identified as higher expressed in N-deprived cultures. These DERs were then annotated by BLAST-matching reads to Swiss-Prot. A total of 195,291 DERs that were higher expressed in controls and 24,400 that were higher expressed in N-deprived samples could be annotated. This revealed a total of 593 unique genes that were differentially expressed between treatments (Supplementary Table 2). It should be noted that this only represents 10.68% of total DERs due to stringent BLAST criteria, as well as low similarity to other available sequences. Out of those, the majority of DERs matched to *Arabidopsis* (18%), while the closest-related alga, *Chlamydomonas reinhardtii,* only had a 3.6% match.

3.4. Functional annotation of differentially expressed genes during N-deprivation of Tetraselmis sp. M8

To assign potential functions to differentially expressed genes, annotated DERs in both treatments were assigned with Gene Ontology (GO) terms using the BLAST2GO pipeline. The distribution of DERs from nitrogen-deprived treatments was found to be distinctively different from those of control samples (Supplementary Fig. 4 and Supplementary Table 3). In the N-starved treatment, carbohydrate metabolic processing (45%) and nucleotide binding (42%) accounted for the largest percentage of DERs, followed by DERs related to catabolic processes, generation of precursor metabolites and response to stress (~20–30%). In comparison, the functional categories that accounted for the largest percentage of DERs in control sequences were plastid, thylakoid, generation of precursor metabolites and protein complex (~70–90%), followed by cellular protein modification process, carbohydrate metabolic process, catabolic process and response to abiotic stimulus (~10–33%). DERs linked to lipid metabolic process was found in both control and N-starved treatments at 7% and 11%, respectively, while DERs linked to lipid particle were found at 0.9% and 0.1%, respectively. This was also reflected in the larger lipid particles observed by microscopic analyses.

This finding, when taken into account the high abundance of DERs linked to carbohydrate metabolic process, nucleotide binding, catabolic process and stress response in nitrogen-starved cultures, suggests a shift in carbon flux away from photosynthesis as the cells respond to unfavorable growth conditions and transition into stationary phase. These changes in transcript abundance can be linked to the reduction in chlorophyll content 24 h later, as thylakoids are degraded and not replaced. Other microalgae transcriptome studies have documented down-regulation of photosynthesis-related genes during nitrogen deprivation, and suggest that light harvesting proteins may be in excess, and that this response is linked to the recycling of nitrogen-rich proteins, but with no immediate effect on photosynthesis capacity (Valenzuela et al., 2012; Lv et al., 2013; Sun et al., 2013). This response, along with the observed reduction of transcripts linked to protein complex suggests that *Tetraselmis* sp. M8 cells begin to convert sugar and change nitrogen allocation during early-stationary phase.

While the reduction in protein synthesis support the observations of reduced cell accumulation, it is interesting to highlight that transcripts linked to cell growth, cell cycle and reproduction were more abundant in the transcriptome from nitrogen-deprived cells. The latter group may include negative regulators of cell growth and reproduction, as cell density in starved cultures did not increase (Fig. 1A).

To identify and construct the lipid biosynthesis and degradation (including β-oxidation and lipases) pathways in *Tetraselmis* sp. M8, genes coding for key enzymes in these pathways were identified by BLAST-searching against the TAGDB. In addition, following the Illumina Mi-Seq transcriptomics experiment, a subsequent independent time course experiment was performed to investigate the expression of these molecular pathways in more details at 16, 24, 32, 48 and 72 h after nitrogen starvation (samples from three separately-grown and treated cultures, each) using qRT-PCR. Monitoring of the physiological parameters revealed cell accumulation in the N-starved cultures, similar to the RNA-seq experiment (Supplementary Fig. 5). As previously observed, cell numbers in N-starved cultures increased only until 48 h, and then subsequently stopped dividing, with significant differences between treatments occurring from 32 h onwards (Supplementary Fig. 5A, $P < 0.05$). In comparison, cell numbers in control cultures tripled between 0 and 48 h, and continued to increase after dilution and feeding. Lipid accumulation in the time-course experiment exhibited similar results to that of the RNA-Seq experiment. The nitrogen-starved cultures were found to accumulate significantly more lipids as early as 16 h onwards (Supplementary Fig. 5B, $P < 0.05$), and also exhibited another marked increase after 48 h. RNA was extracted from control and N-starved cultures at every time point, and qRT-PCR analysis was performed to determine the expression of the various gene in the lipid pathways. Primers for qRT-PCR exper-

iments were based on conserved regions, using RNA-Seq data, in an effort to capture most gene family members. However, it should be considered that not all members of the gene families may have been covered in this study.

3.5. Transcripts related to fatty acid biosynthesis and catabolism were down-regulated at 24 h despite an increase in lipid content

Figs. 3–5 show the reconstructed pathways for FA synthesis, TAG synthesis and lipid degradation (β-oxidation) based on the identified enzymes. Transcripts were quantified using combined DiffKAP and qRT-PCR results. Based on the DiffKAP results, the entire FA synthesis pathway, except for 3-ketoacyl-ACP reductase (KAR) and Enoyl-ACP reductase (ENR) was down-regulated, compared to the controls. The expression of genes coding for the entire TAG synthesis pathway remained unchanged according to the DiffKAP data, except for Diacylglycerol O-acyltransferase (DGAT), which decreased 2-fold. In the β-oxidation pathway, the DiffKAP results suggested that genes encoding acyl-CoA synthetase (ACSase) and Enoyl-CoA hydratase (ECH) were down-regulated 2- and 10-fold, respectively, although ACSase was not downregulated in time course analyses.

Similarly, the qRT-PCR results from the timecourse analysis found the entire fatty acid synthesis pathway to be down-regulated under N-starvation, except for Malonyl-CoA:ACP transacylase (MAT) and ENR, whose expression examined unchanged. Transcript abundances of Acetyl-CoA carboxylase (ACCase), 3-ketoacyl-ACP synthase (KAS), 3-ketoacyl-ACP reductase (KAR), 3-hydroxyacyl-ACP dehydratase (HD) significantly decreased at 24 h (4-fold, 3-fold, 8-fold and 75-fold, respectively). No change in transcript abundance was observed in the TAG synthesis pathway in the qRT-PCR data, however the DiffKAP results suggested that the final step of this pathway, DGAT, was down-regulated 2-fold. In the β-oxidation pathway, the qRT-PCR data showed no change in transcript abundance, except for the initial step, TAG lipase, which was 4-fold down-regulated, in contrast with the DiffKAP data, which showed no change.

When taken together, the DiffKAP and qRT-PCR results suggest that the FA synthesis and β-oxidation pathways were down-regulated 24 h after N-starvation. The majority of differential expression data between DiffKAP and qRT-PCR were in agreeance. When making biological conclusions from this data, we used the qRT-PCR data as a reference, as we considered it to be a more reliable measure of transcript abundance.

3.6. Lipid accumulation during early-stationary phase of Tetraselmis sp. M8 could be due to reduced β-oxidation

In *Tetraselmis* sp. M8, although FA synthesis was down-regulated, the down-regulation of genes (TAG lipase, ACSase) at the committing steps, as well as at the ECH-gene of the β-oxidation pathway may indicate that the observed increase in lipid accumulation at this time point is a result of reduced lipid degradation, rather than increased lipid synthesis. This observation has also been observed in *Nannochloropsis gaditana*, where the lack of up-regulation amongst lipid biosynthesis genes despite increased lipid production has been attributed to sufficiently abundant existing lipid production machinery carried over from growth phase, coupled with a shift in carbon flux away from carbohydrate synthesis (Radakovits et al., 2012). There is evidence supporting this in *Tetraselmis* sp. M8, as the maintenance of basal levels lipid production coupled with the decrease of lipid catabolism would result in an overall increase in lipid production. Also similar to *N. gaditana* and *P. tricornutum,* the observed down-regulation of genes encoding fructose-1,6-biphosphate and fructose-1,6-biphosphate aldolase (Supplementary Table 2), key regulatory enzymes of carbon metabolism (Calvin cycle and gluconeogenesis) in *Tetraselmis* sp. M8, suggests a possible shift in the carbon flux away from carbohydrate synthesis to lipid synthesis. This could contribute to the increase of lipid production as carbon is being "pushed" into FA synthesis and not being "pulled" by increased FA synthesis genes (Radakovits et al., 2012; Valenzuela et al., 2012; Yang et al., 2013). This diversion of carbon towards lipid metabolism is further supported by the observed reduction in starch synthase genes (Supplementary Table 2). Overall, results in this study suggest that the observed lipid accumulation of *Tetraselmis* sp. M8 at 24 h is the result of a reduction in lipid catabolism, coupled with a possible shift in carbon flux towards lipid synthesis.

3.7. Despite earlier signs of cellular lipid accumulation, FA and TAG synthesis genes were only up-regulated at 48 h after N deprivation

Irrespective of experimental approach taken, gene expression levels in the FA synthesis pathway of nitrogen-starved cells were mostly down-regulated or unchanged in the first 24 h, but were then up-regulated significantly from 48 h onwards (Fig. 3). In the first 24 h, the expression levels of genes encoding ACCase, KAR and 3-hydroxyacyl-ACP dehydratase (HD) (FA synthesis) in nitrogen-starved cells were 3-, 5.9- and 113-fold lower, respectively, ($P < 0.05$) than in control cells, but ACCase, KAR, ENR and HD were then significantly up-regulated 3- to 11-fold higher ($P < 0.05$) than control cells at 48 h and 72 h. FA synthesis pathway upregulation at later time points is consistent with the flux model at 72 h after nitrogen deprivation (Supplementary Fig. 3). For the MAT-encoding gene(s), gene expression between treatments was similar until 72 h, when it was down-regulated by 2.5-fold ($P < 0.05$). The expression level of ENR-encoding gene was similar between the treatments until 48 h onwards when this gene was found to be significantly down-regulated in control cells (8-fold, $P < 0.05$). In the nitrogen deprivation treatment, the KAS-encoding gene(s) was consistently down-regulated throughout the experiment. In the TAG synthesis pathway (Fig. 4), gene expression remained largely unchanged between control and treated cells, particularly in the first 24 h. After 48 h onwards, up-regulation of the PP-encoding gene(s) in nitrogen-starved cells was observed ($P < 0.05$), followed by the up-regulation of the gene(s) encoding GK at 72 h (4-fold, $P < 0.05$). No differences in gene expression were observed between control and treatment for genes encoding glycerol-3-phosphate O-acyltransferase (GPAT) and 1-acyl-sn-glycerol-3-phosphate acyltransferase (AGPAT) throughout the entire experiment. DGAT-encoding gene(s) was found to be significantly down-regulated ($P < 0.05$) at 48 h and 72 h after nitrogen starvation. In the lipid degradation pathway (Fig. 5), only gene(s) encoding TAG lipase was observed to be down-regulated (4.6-fold, $P < 0.05$) in N-starved cells, while other genes remained unchanged between treatments. Changes in gene expression can be observed from 48 h onwards, with genes encoding ACSase being up-regulated (5-fold, $P < 0.05$) in nitrogen-starved cells and ECH being down-regulated (83-fold, $P < 0.05$) at 72 h.

In summary, as FA and TAG synthesis genes were mostly down-regulated at 24 h after N deprivation (Figs. 3 and 4), the observed increased lipid accumulation of cells harvested at this time (Fig. 1, Supplementary Fig. 5B) is likely to be attributed to a reduced rate of FA degradation by β-oxidation (Fig. 5). Cells at a later stage then clearly show up-regulation of FA and TAG synthesis that coincides with further increases in lipid fluorescence.

3.8. High lipid accumulation during stationary phase of Tetraselmis sp. M8 is due to active FA synthesis

The expression of lipid-related genes of *Tetraselmis* sp. M8 at 48 h and 72 h after nitrogen deprivation was analyzed in the time

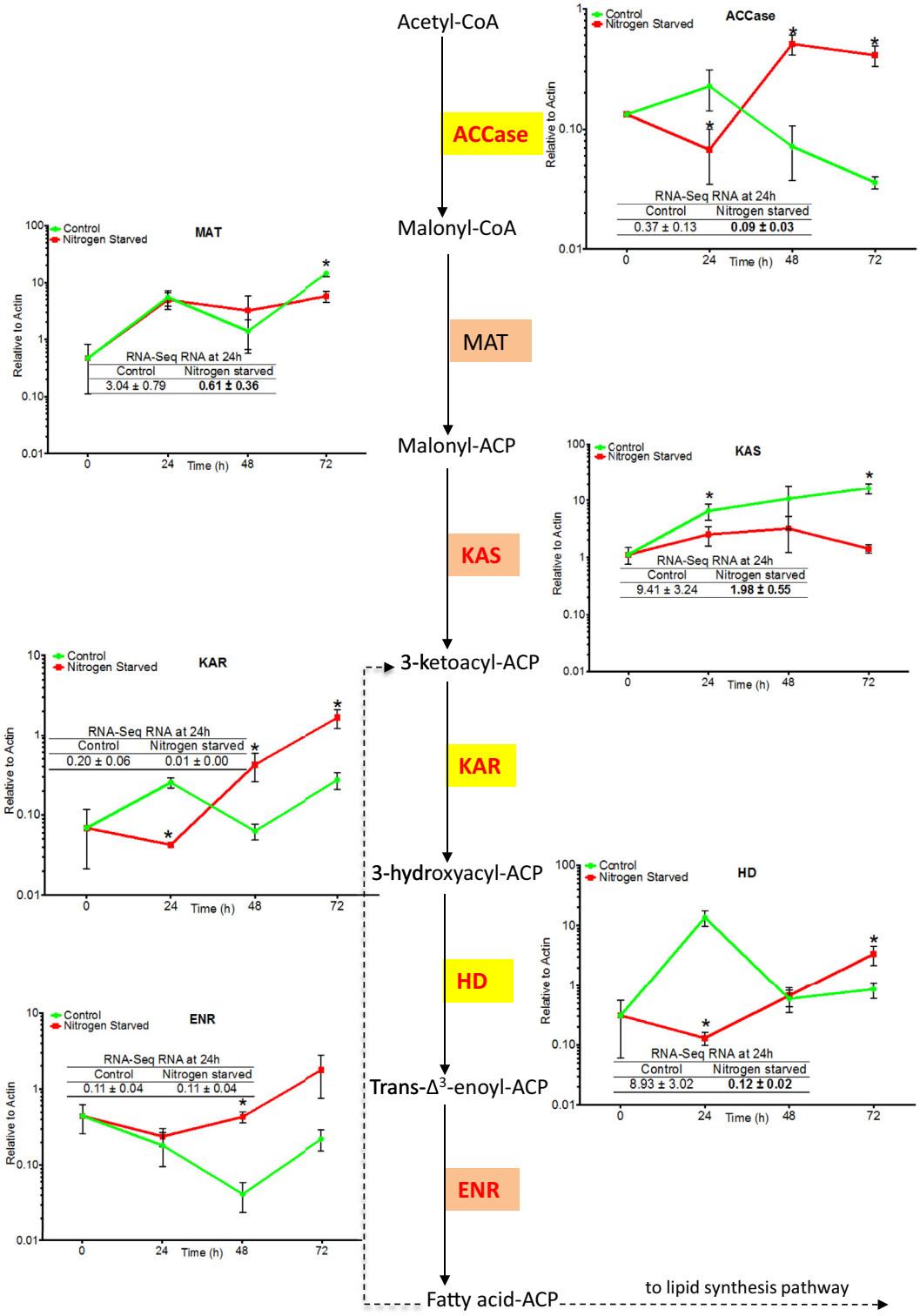


Fig. 3. Fatty acid synthesis pathway and gene expression profiles under nitrogen deprivation. Genes in red font were down-regulated in nitrogen-starved treatment according to DiffKAP analysis. Inserted tables show qRT-PCR analysis of the RNA-seq RNA, with bold numbers indicating significant differences (Student's *T*-test; $P < 0.05$). Graphs show qRT-PCR expression analysis of genes at 0, 24, 48 and 72 h of the time course experiment, with asterisks (*) indicating significant differences (Student's *T*-test; $P < 0.05$). Genes affected by the circadian rhythm were highlighted in yellow (both control and nitrogen-starved affected) and orange (only control affected). Acetyl-CoA carboxylase (ACCase); Malonyl-CoA:ACP transacylase (MAT); 3-ketoacyl-ACP synthase (KAS); 3-ketoacyl-ACP reductase (KAR); 3-hydroxyacyl-ACP dehydratase (HD); Enoyl-ACP reductase (ENR); FAT (Acyl-ACP thioesterase).

course experiment. The aim was to investigate the increase in rate of lipid accumulation, as well as determine if the lack of expression in FA synthesis and TAG synthesis pathways were consistent throughout the entire stationary phase. The expression of the entire FA pathway in nitrogen-deprived cultures was significantly higher than the control cultures after 24 h (Fig. 3). The committing

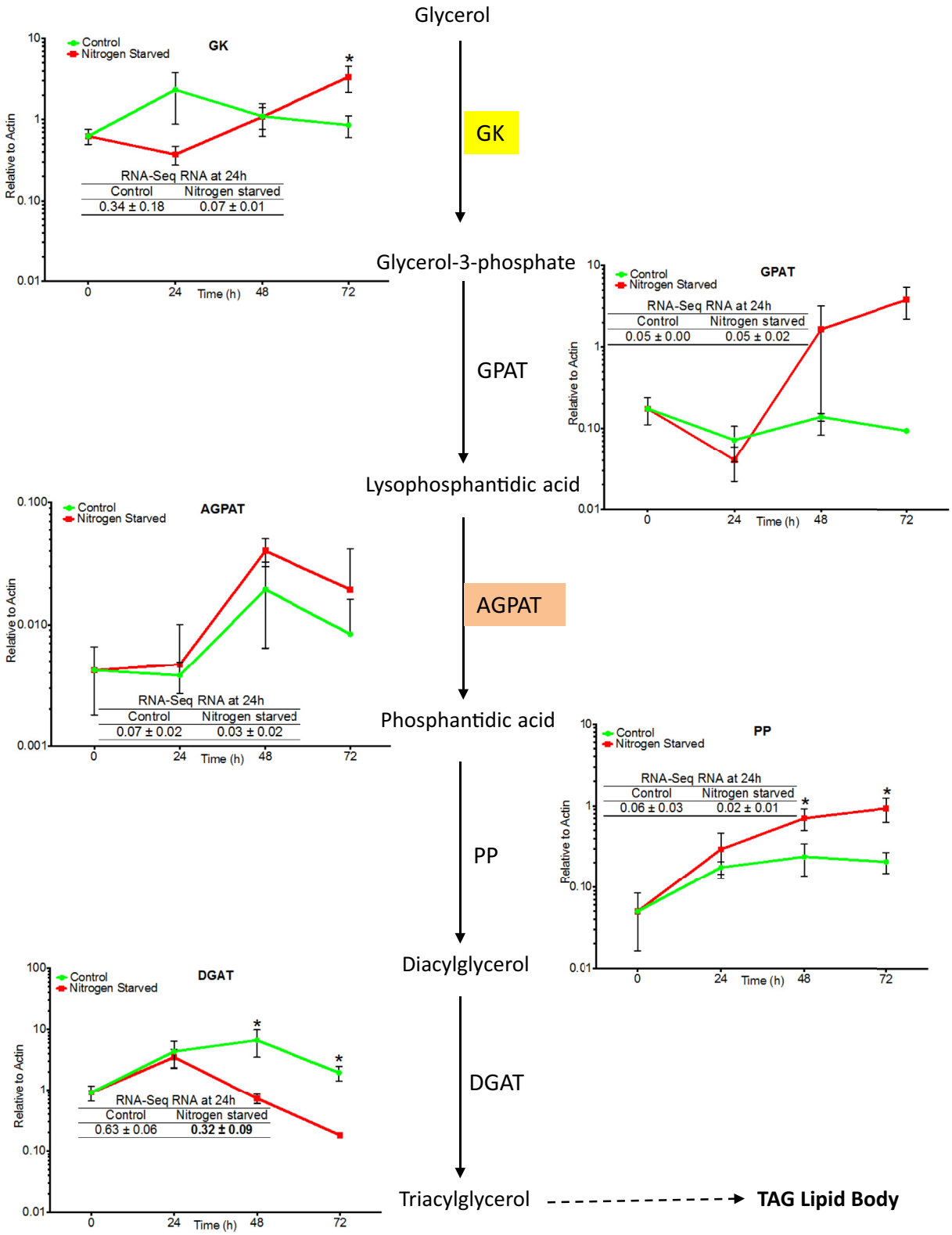


Fig. 4. Triacylglyceride synthesis pathway and gene expression profiles under nitrogen deprivation. Genes in red font were found down-regulated in DiffKAP analysis of RNA-seq RNA. Inserted tables show qRT-PCR analysis of the RNA-seq RNA, with bold numbers indicating significant differences (Student's T-test; $P < 0.05$). Inserted graphs show qRT-PCR expression analysis of genes at 0, 24, 48 and 72 h of the time course experiment, with asterisks (*) indicating significant differences (Student's *T*-test; $P < 0.05$). Genes affected by the circadian rhythm were highlighted in yellow (both control and nitrogen-starved affected) and in orange (only control affected). Glycerol kinase (GK); Glycerol-3-phosphate O-acyltransferase (GPAT); 1-acyl-sn-glycerol-3-phosphate acyltransferase (AGPAT); Phosphatidate phosphatase (PP); Diacylglycerol O-acyltransferase (DGAT); Triacylglyceride (TAG) lipase.

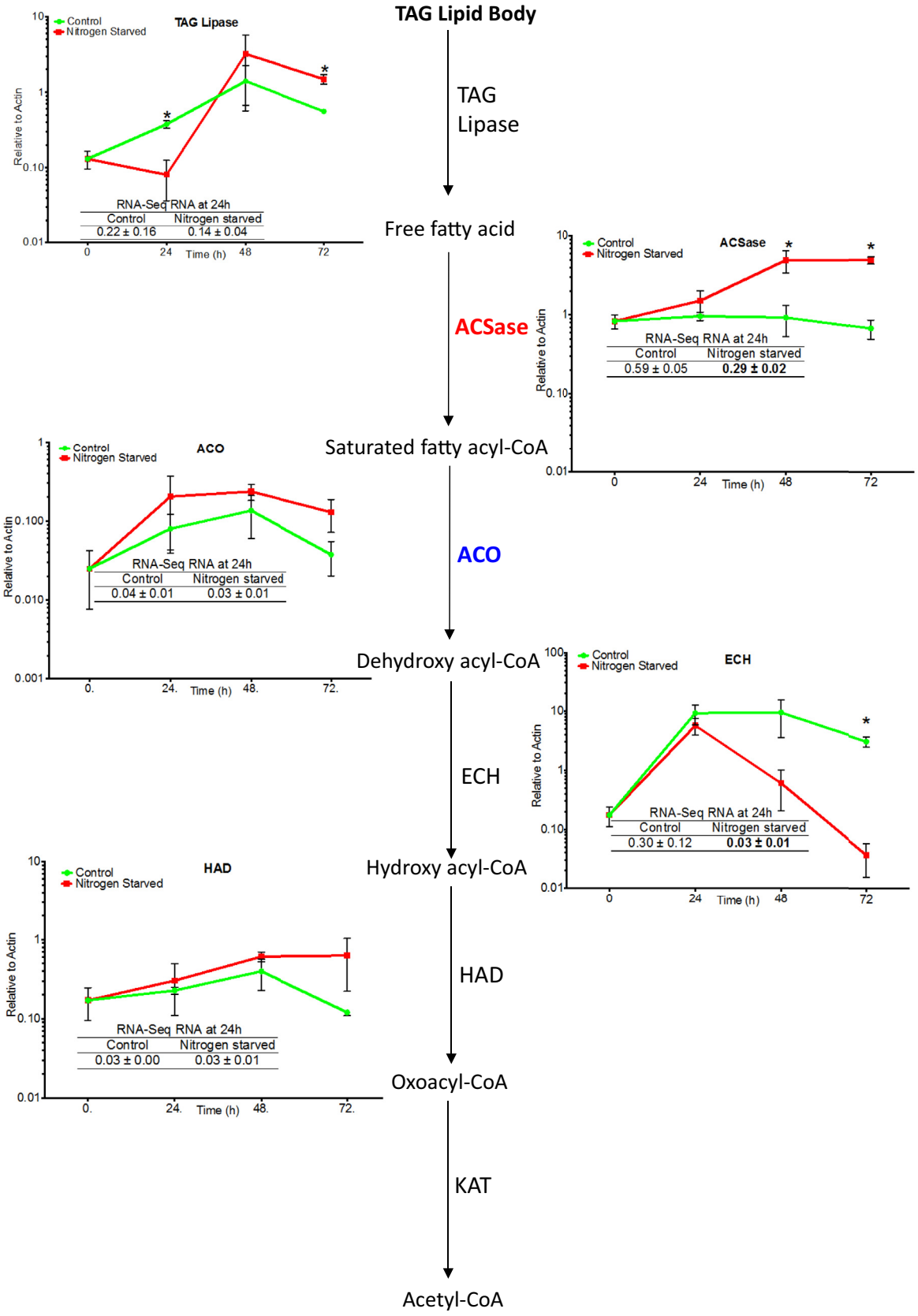


Fig. 5. Lipid degradation pathway and gene expression profiles under nitrogen deprivation. Genes in red font were found down-regulated in DiffKAP analysis of RNA-seq RNA. Inserted tables show qRT-PCR analysis of the RNA-Seq RNA, with bold numbers indicating significant differences (Student's *T*-test; $P < 0.05$). Inserted graphs show qRT-PCR expression analysis of genes at 0, 24, 48 and 72 h of the time course experiment, with asterisks (*) indicating significant differences (Student's *T*-test; $P < 0.05$). Acyl-CoA synthetase (ACSase); Acyl-CoA oxidase (ACO); Enoyl-CoA hydratase (ECH); Hydroxyacyl- CoA dehydrogenase (HAD); Ketoacyl- CoA thiolase (KAT).

steps of TAG synthesis (GK- and GPAT-encoding genes) and the phosphatidate phosphatase (PP) gene were also found to be up-regulated in nitrogen-deprived cultures (Fig. 4). Interestingly, both TAG lipase and ACSase-encoding genes were similarly up-regulated in nitrogen-starved treatments, with only the ECH-gene being down-regulated at 72 h. These observations suggest that lipid accumulation after 24 h has switched from a result of decrease in catabolism to an increase in FA and TAG synthesis activity. *Tetraselmis* sp. M8 cells would be actively producing more lipids as opposed to just consuming less, which would explain the increase in the rate of lipid accumulation. The lack of a clear significant increase in many TAG synthesis genes may suggest the TAG assembly pathway on a whole, may be more post-transcriptional controlled in *Tetraselmis* sp. M8, particularly for DGAT that exhibited reduced expression during nitrogen-starvation. Post-transcriptional control of DGAT has been reported in proteomics studies of *Chlorella vulgaris* (Guarnieri et al., 2011) and *Brassica napus* (Nykiforuk et al., 2002), and has also been suggested for *Neochloris oleoabundans* (Rismani-Yazdi et al., 2012). The evidence for post-transcriptional control of DGAT is further supported by the lack of the phospholipid:diacylglycerol acyltransferase (PDAT) gene within the *Tetraselmis* sp. M8 transcriptome comprising more than 73,000,000 reads. This indicates that *Tetraselmis* sp. M8 may lack the acyl-CoA-independent mechanism for TAG biosynthesis as has been found in certain microalgae (e.g. *Dunaliella tertiolecta*; Rismani-Yazdi et al., 2011), and thus relies solely on the TAG synthesis pathway for lipid production. Further proteomic and metabolic studies have to be performed to confirm this, but it is interesting to note that heterologous expression of *DGAT* genes from a plant and yeast in *Tetraselmis chui* resulted in increased TAG accumulation (Úbeda-Mínguez et al., 2017). The up-regulation of genes (e.g. for TAG-lipase and ACSase) within the β-oxidation pathway during starvation phase is not completely unexpected, and has previously been observed in *P. tricornutum* in association with changing membrane dynamics to cope with nutrient depletion (Valenzuela et al., 2012). The ECH-encoding gene however, was found to be progressively down-regulated after 24 h, and could present a potential bottleneck in the pathway. Nevertheless, the increased rate in lipid accumulation observed after 24 h, coupled with the up-regulation of the FA synthesis pathway is indicative that *Tetraselmis* sp. M8 has transitioned from a state of reduced lipid consumption during early-stationary phase to a state of active lipid production during stationary phase. It appears that under nutrient stress conditions, one of the first responses of *Tetraselmis* cells is to down-regulate gene expression which is energetically preferable to upregulating genes where the newly synthesized proteins could benefit from the additional nitrogen made available from early protein degradation.

To examine whether the early lipid accumulation at 24 h after N deprivation can be explained by a reduced rate of FA degradation by β-oxidation, TAG lipase was constrained (down-regulated) for metabolic flux analyses, assuming the transcriptome has a regulatory effect on pathway flux. Supplementary Fig. 6 shows that under this assumption of regulatory constraint, the lipid biosynthesis and catabolism pathways are both significantly downregulated, despite an overall increase in lipid accumulation. These results seem to support the notion that during early stationary phase growth, the downregulation of lipid catabolism, rather than an upregulation of biosynthesis, drives the increase in lipid accumulation in microalgae.

3.9. Expression profiling of genes related to lipid pathways and their effect on circadian rhythms

The expression levels of genes involved in lipid pathways were also investigated at 16 h (at the end of the dark cycle) and at 32 h (at the end of the light cycle) to determine the effects of the circadian rhythm on gene expression (Supplementary Fig. 7). The complete set of graphs for all genes including the 16 h and 32 h time points is available in Supplementary Figs. 7–9. In genes encoding ACCase, KAR, HD and glycerol kinase (GK) (Fig. 3, highlighted yellow), a circadian effect was observed in both control and nitrogen-starved treatments, whereby expression was significantly up-regulated after and before the dark cycle. However in genes encoding MAT, 3-ketoacyl-ACP synthase (KAS), ENR and 1-acyl-sn-glycerol-3-phosphate acyltransferase (AGPAT) (Figs. 3 and 4, highlighted orange), the circadian effect was only observed in the control treatment, while gene expression in N-starved cells remained consistent throughout the 16 h and 32 h time points. Furthermore, while all the genes affected by the circadian cycle presented a diurnal expression, ENR- and GK-encoding genes showed increased expression only towards the end of the light cycle at 32 h. This was different from those in *P. tricornutum* (Chauton et al., 2013), where most of the FA synthesis genes showed increased expression only at the onset of the light cycle, while β-oxidation genes showed increased expression towards the end. This could suggest that FA synthesis in *Tetraselmis* sp. M8 could be linked to photosynthesis and the changing influx of carbon, while TAG and β-oxidation are not. Furthermore, nitrogen-starvation appears to have an overriding effect on certain genes such as those encoding MAT, KAS, ENR and AGPAT, where a spike in expression at the start or end of the light cycle was no longer observed in comparison with control cultures. Lastly, the circadian rhythm had measurable effects on genes encoding TAG lipase, ACSase and ECH, involved in lipid catabolism. The regular medium replacement after 48 h may have also had an additional effect on gene expression, similar to a circadian rhythm. Both overarching influences were accounted for, as controls were equally exposed to these external factors.

3.10. Nitrogen-starvation improves FA profile of Tetraselmis sp. M8 for potential oil production

Oil produced by *Tetraselmis* sp. presents a suitable feedstock for variety of applications, including biodiesel, based on its suitable FA composition (Lim et al., 2012). Under nitrogen-deplete conditions, the FA profile of *Tetraselmis* sp. M8 was found to increase in its proportion of saturated (C16) and monounsaturated FA (C18:1), and to decrease in polyunsaturated FA (C16:3, C18:3, C18:3; Fig. 2). Several genes encoding enzymes involved in FA desaturation were identified by DiffKAP and have decreased expression in nitrogen-deprived cultures (Supplementary Table 2), such as genes encoding omega-6 FA desaturase, palmitoyl-monogalactosyldiacylglycerol delta-7 desaturase, lipid desaturase ADS3.2 and delta-9 acyl-lipid desaturase. Desaturases such as delta-9 acyl- lipid desaturase have been found in other microalgae that have a similar reduction in poly-unsaturated FA and an increase in saturated FA (C16), improving their FA content's cetane number and resulting biodiesel (Miller et al., 2010; Rismani-Yazdi et al., 2012).

3.11. DiffKAP vs qRT-PCR and experimental limitations

There are several limitations associated with this study, most of which are present due to the fact that the genome of *Tetraselmis* sp. has not been sequenced. The initial attempt at transcriptome mapping to *C. reinhardtii* genome had a <0.02% match, and we therefore used DiffKAP to identify differentially-expressed reads between the two treatments. Reads that were revealed as either highly expressed in control or nitrogen-starved treatments were then annotated to Swiss-Prot by BLAST analyses. Limitations arose due to the high stringency of the BLAST parameters (e-value 10^{-16}), coupled with the fact that *Tetraselmis* sp. M8 sequences were not

closely related to any available genomes. This resulted in only 10.68% of DERs being annotated, leaving nearly 90% of the DiffKAP results unannotated. This could have caused certain genes in our DiffKAP analysis of the identified lipid metabolism pathways to not be identified as differentially expressed. For example, genes encoding MAT, DGAT and ECH, which were classified as "not differentially expressed" by DiffKAP, although qRT-PCR analysis revealed these to be significantly down-regulated at 24 h in both RNA-Seq and time-course experiments. Primers for this experiment were designed based on the consensus sequence of all reads (extracted via TAGDB) that were related to a reference gene, and thus the qRT-PCR results would more likely reflect the overall expression of whole gene families, and not be limited to what was annotated. The nature of qRT-PCR analysis and the primers would also explain the expression of genes that encode ACO and ENR, whose differential expression was identified by DiffKAP, but not by qRT-PCR analysis. This is because DiffKAP would have identified individual genes as being differentially expressed, while qRT-PCR analysis revealed the gene family as being unchanged. Therefore, due to the limitations caused by the low degree of annotation in DiffKAP, results from this analysis were used primarily as an overview of gene expression in *Tetraselmis* sp. M8. This was then followed up with qRT-PCR analysis, providing a more accurate representation of gene expression in the lipid pathways.

Another limitation in this study is the discrepancies in the feeding regime of the RNA-Seq and time-course experiments. Cultures in the time-course experiment were fed with full strength F-media and diluted to 0.5×10^6 cells/mL every 48 h, compared to F/2 media and diluted by half in the RNA-Seq experiment, effectively giving the cells in the time-course experiment more nutrients per cell. This was done to induce a more extreme starvation difference, and by extension a more distinct transcription profile between the two treatments as more nutrients would be available for control cultures towards the end of every 48 h feeding cycle. The extra exogenous nutrients would allow N-deprived cultures to have additional internal nutrient stores, and could exhibit a more delayed expression profile compared to the RNA-Seq experiment. This could explain the discrepancies in the DGAT-encoding gene expression, where the significant difference observed in the RNA-seq experiment was only observed 24 h later in the time-course experiment. Future studies may involve knock-down or overexpression experiments in *Tetraselmis* coupled with phenotypic analysis. The identification of key regulatory genes in this process could help to bioengineer cells that favor rapid TAG biosynthesis rather than starch accumulation.

4. Conclusion

Nitrogen-induced lipid accumulation in *Tetraselmis* sp. M8 is not a simple case of increased FA and TAG synthesis. A distinct early-stationary phase, characterized by reduced cell division and increased lipid accumulation, was followed by a stationary phase characterized by a cessation of cell division and a significant lipid accumulation. A shift from reduced lipid degradation to active FA production explains the increase in the rate of lipid accumulation at 48 h and 72 h, while DGAT gene expression was not a bottleneck in *Tetraselmis* sp. M8 lipid biosynthesis.

Acknowledgements

We wish to thank the Australian Research Council (LP0990558) and Meat and Livestock Australia for research funding.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.biortech.2017.06.003.

References

Adarme-Vega, T.C., Thomas-Hall, S.R., Lim, D.K.Y., Schenk, P.M., 2014. Effects of long chain fatty acid synthesis and associated gene expression in microalgae *Tetraselmis* sp. Mar. Drugs 12, 3381–3398.

Ahmad, A.L., Mat Yasin, N.H., Derek, C.J.C., Lim, J.K., 2011. Microalgae as a sustainable energy source for biodiesel production: a review. Renew. Sustain. Energy Rev. 15, 584–593.

Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M., Bairoch, A., 2007. UniProtKB/Swiss-Prot. Methods Mol. Biol. 406, 89–112.

Chauton, M.S., Winge, P., Brembu, T., Vadstein, O., Bones, A.M., 2013. Gene regulation of carbon fixation, sotrage and utlisation in the diatom *Phaeodactylum tricornutum* acclumated to light/dark cycles. Plant Physiol. 161, 1034–1048.

Chini Zitelli, G., Rodolfi, L., Biondi, N., Tredici, M.R., 2006. Productivity and photosynthetic efficiency of outdoor cultures of *Tetraselmis suecica* in annular columns. Aquaculture 261, 932–943.

C.G., de Oliveira Dal'Molin, Quek, L.E., Palfreyman, R.W., Nielsen, L.K., 2011. AlgaGEM–a genome-scale metabolic reconstruction of algae based on the Chlamydomonas reinhardtii genome. BMC Genomics 12, S5.

Dinamarca, J., Levitan, O., Kumaraswamy, G.K., Lun, D.S., Falkowski, P., 2017. Overexpression of a diacylglycerol acyltransferase gene in *Phaeodactylum tricornutum* directs carbon towards lipid biosynthesis. J. Phycol. http://dx.doi.org/10.1111/jpy.12513.

Franson, M.A.H., Eaton, A.D., 2005. Standard Methods for the Examination of Water & Wastewater. American Public Health Association, Washington, DC.

Gimpel, J.A., Specht, E.A., Geogianna, D.R., Mayfield, S.P., 2013. Advances in microalgae engineering and synthetic biology applications for biofuel production. Curr. Opin. Chem. Biol. 17, 489–495.

Gotz, S., Garcia-Gomez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talon, M., Dopazo, J., Conesa, A., 2008. High-throughput functional annotation and data mining with the Blast2GO suit. Nucleic Acids 36, 3420–3435.

Guarnieri, M.T., Nag, A., Smolinski, S.L., Darzins, A., Seibert, M., Pienkos, P.T., 2011. Examination of triacylglycerol biosynthetic pathways via *de novo* transcriptomic and proteomic analyses in an unsequenced microalga. PLoS One 6, e25851.

Gudmundsson, S., Thiele, I., 2010. Computationally efficient flux variability analysis. BMC Bioinf. 11, 489.

Guillard, R.R., Ryther, J.H., 1962. Studies of marine planktonic diatoms.1. Cyclotella nana Hustedt, and Detonula confervacea (Cleve) Gran. Can. J. Microbiol. 8, 229–239.

Hu, Q., Sommerfeld, M., Jarvis, E., Ghirardi, M., Posewitz, M., Seibert, M., Darzins, A., 2008. Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. Plant J. 54, 621–639.

Kurtz, S., Narechania, A., Stein, J., Ware, D., 2008. A new method to compute k-mer frequencies and its application to annotate large repetitive plant genomes. BMC Genomics 9, 517.

Lim, D.K.Y., Garg, S., Timmins, M., Zhang, E.S.B., Thomas-Hall, S.R., Schumann, H., Li, Y., Schenk, P.M., 2012. Isolation and evaluation of oil-producing microalgae from subtropical coastal and brackish waters. PLoS One 7, e40751.

Lv, H., Qu, G., Qi, X., Lu, L., Tian, C., Ma, Y., 2013. Transcriptome analysis of *Chlamydomonas reinhardtii* during the process of lipid accumulation. Genomics 101, 229–237.

Marçais, G., Kingsford, C., 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics 27, 764–770.

Marshall, D.J., Hayward, A., Eales, D., Imelfort, M., Stiller, J., Berkman, P.J., Clark, T., McKenzie, M., Lai, K., Duran, C., Batley, J., Edwards, D., 2010. Targeted identification of genomic regions using TAGdb. Plant Methods 6, 19.

Miller, R., Wu, G., Deshpande, R.R., Vieler, A., Gartner, K., Li, X., Moellering, E.R., Zauner, S., Cornish, A.J., Liu, B., Bullard, B., Sears, B.B., Kuo, M., Hegg, E.L., Shachar-Hill, Y., Shiun, S., Benning, C., 2010. Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen deprivation predict diversion of metabolism. Plant Physiol. 154, 1737–1752.

Narala, R.R., Garg, S., Sharma, K.K., Thomas-Hall, S.R., Deme, M., Li, Y., Schenk, P.M., 2016. Comparison of microalgae cultivation in photobioreactor, open raceway pond, and a two-stage hybrid system. Front. Energy Res. 4, 29.

Nascimento, I.A., Marques, S.S.I., Cabanelas, I.T.D., Pereira, S.A., Druzian, J.I., de Souza, C.O., Vich, D.V., de Carvalho, G.C., Nascimento, M.A., 2013. Screening microalgae strains for biodiesel production: lipid productivity and estimation of fuel quality based on fatty acids profiles as selective criteria. Bioenergy Res. 6, 1–13.

Nykiforuk, C.L., Furukawa-Stoffer, T.L., Puff, P.W., Sarna, M., Laroche, A., Moloney, M.M., Weselake, R.J., 2002. Characterization of cDNAs encoding diacylglycerol acyltransferase from cultures of *Brassica napus* and sucrose-mediated induction of enzyme biosynthesis. Biochim. Biophys. Acta Mol. Cell Biol. Lipids 1580, 95–109.

Radakovits, R., Jinkerson, R.E., Fuerstenberg, S.I., Tae, H., Settlage, R.E., Boore, J.L., Posewitz, M., 2012. Draft genome sequence and genetic transformation of the oleaginous alga *Nannochloropsis gaditana*. Nat. Commun. 3, 686.

Rismani-Yazdi, H., Haznedaroglu, B.Z., Bibby, K., Peccia, J., 2011. Transcriptome sequencing and annotation of the microalgae *Dunaliella tertiolecta*: Pathway description and gene discovery for production of next-generation biofuels. BMC Genomics 12, 148.

Rismani-Yazdi, H., Haznedaroglu, B.Z., Hsin, C., Peccia, J., 2012. Transcriptomic analysis of the oleaginous microalga *Neochloris oleoabundans* reveals metabolic insights into triacylglyceride accumulation. Biotechnol. Biofuels 5, 74.

Rodolfi, L., Zittelli, G.C., Bassi, N., Padovani, G., Biondi, N., Bonini, G., Tredici, M.R., 2009. Microalgae for oil: strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. Biotechnol. Bioeng. 102, 100–112.

Rosic, N., Kaniewska, P., Chan, C.K.K., Ling, E.Y.S., Edwards, D., Dove, S., Hoegh-Guldberg, O., 2014. Early transcriptional changes in the reef-building coral *Acropora aspera* in response to thermal and nutrient stress. BMC Genomics 15, 1052.

Sharma, K., Li, Y., Schenk, P.M., 2014. UV-C-mediated lipid induction and settling: a step change towards economical microalgal biodiesel production. Green Chem. 16, 3539–3548.

Sharma, K.S., Schuhmann, H., Schenk, P.M., 2012. High lipid induction in microalgae for biodiesel production. Energies 5, 1532–1553.

Shin, H., Hong, S.J., Yoo, C., Han, M.A., Lee, H., Choi, H.K., Cho, S., Lee, C.G., Cho, B.K., 2016. Genome-wide transcriptome analysis revealed organelle specific responses to temperature variations in algae. Sci. Rep. 6, 37770.

Sun, D., Zhu, J., Fang, L., Zhang, X., Chow, Y., Liu, J., 2013. *De novo* transcriptome profiling uncovers a drastic downregulation of photosynthesis upon nitrogen deprivation in nonmodel green alga *Botryosphaerella sudeticus*. BMC Genomics 14, 715.

Timmins, M., Zhou, W.R., Rupprecht, J., Lim, L., Thomas-Hall, S.R., Doebbe, A., Kruse, O., Hankamer, B., Schenk, P.M., 2009. The metabolome of *Chlamydomonas reinhardtii* following induction of anaerobic H_2 production by sulfur depletion. J. Biol. Chem. 284, 23415–23425.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., Pachter, L., 2012. Differential gene and transcript expression analysis of RNA-Seq experiments with TopHat and Cufflinks. Nat. Protoc. 7, 562–578.

Úbeda-Mínguez, P., García-Maroto, F., Alonso, D.L., 2017. Heterologous expression of *DGAT* genes in the marine microalga *Tetraselmis chui* leads to an increase in TAG content. J. Appl. Phycol. http://dx.doi.org/10.1007/s10811-017-1103-9.

Valenzuela, J., Mazurie, A., Carlson, R.P., Gerlack, R., Cooksey, K.E., Peyton, B.M., Fields, M.W., 2012. Potential role of multiple carbon fixation pathways during lipid accumulation in *Phaeodactylum tricornutum*. Biotechnol. Biofuels 5, 40.

Wang, Z.T., Ullrich, N., Joo, S., Waffenschmidt, S., Goodenough, U., 2009. Algal lipid bodies: stress induction, purification, and biochemical characterization inwild-type and starchless *Chlamydomonas reinhardtii*. Eukaryot. Cell 8, 1856–1868.

Yang, Z., Niu, Y., Ma, Y., Xue, J., Zhang, M., Yang, W., Liu, J., Lu, S., Guan, Y., Li, H., 2013. Molecular and cellular mechanisms of neutral lipid accumulation in diatom following nitrogen deprivation. Biotechnol. Biofuels 6, 1–67.

Zayadan, B.K., Purton, S., Sadvakasova, A.K., Userbaeva, A.A., Bolatkhan, K., 2014. Isolation, mutagenesis, and optimization of cultivation conditions of microalgal strains for biodiesel production. Russ. J. Plant Physiol. 61, 124–130.