**Mass Spectrometry-based Metabolomics of Value-added Biochemicals from *Ettlia*
*oleoabundans***

*Eryn K. Matich*¹, *Deena M. Butryn*¹, *Mohsen Ghafari*², *Virginia del Solar*¹, *Ersan Camgoz*³,
*Blaine A. Pfeifer*⁴, *Diana S. Aga*¹, *Berat Z. Haznedaroglu*⁵, *G. Ekin Atilla-Gokcumen*^{1*}

¹ Department of Chemistry, University at Buffalo, The State University of New York (SUNY),
Buffalo, NY USA 14260

² Department of Civil, Structural and Environmental Engineering, University at Buffalo, The
State University of New York (SUNY), Buffalo, NY USA 14260

³ Department of Chemical and Biological Engineering, Koc University, Turkey 34450

⁴ Department of Chemical and Biological Engineering, University at Buffalo, The State
University of New York (SUNY), Buffalo, NY USA 14260

⁵ Institute of Environmental Sciences, Bogazici University, Turkey 34342

ABSTRACT

Microalgae are promising feedstocks for biofuels and other value-added biochemicals. *Ettlia oleoabundans*, a non-model green microalgae, produces increased levels of key precursors for lipid-based biofuels under nutrient stress conditions. Here, we introduce a method for analysis of hydrophobic metabolites in *E. oleoabundans* and apply this method to microalgae grown under different nutrient conditions. We combined targeted analysis using gas chromatography/ tandem mass spectrometry (GC-MS/MS) with targeted and untargeted metabolomics using liquid chromatography-quadrupole-time-of-flight mass spectrometry (LC-Q-ToF MS). Our method allows simultaneous detection of compounds from major lipid families, and other value-added biochemicals using less biomass, hence allowing analysis of larger number of samples to be processed simultaneously. In this study, we showed that 202 metabolites were enhanced or depleted under nutrient stress and annotated 163 of these species. A complex regulation of different classes of metabolites under nutrient stress was observed. Overall, triacylglycerols and sulfoquinovosyldiacylglycerols were accumulated, chlorophylls were depleted, and diacylglycerols were differentially regulated at the molecular level when grown under limited nutrient conditions. The method that we developed will serve as a template for targeted and untargeted metabolomics in *E. oleoabundans* and other non-model microalgae to understand biochemical regulations under different experimental conditions.

Keywords: Microalgae, metabolomics, biofuels, biochemicals, *Ettlia oleoabundans*, GC-MS/MS, LC-Q-ToF MS.

ABBREVIATIONS

In order as they appear: GC-MS/MS, gas chromatography tandem mass spectrometry; LC-Q-ToF MS, liquid chromatography quadrupole-time-of-flight mass spectrometry; DoE, U.S. Department of Energy; ASP, Aquatic Species Program; N⁻, nitrogen deplete; N⁺, nitrogen replete; FAs, fatty acids; PGs, phosphatidylglycerols; PEs, phosphatidylethanolamines; PIs, phosphatidylinositols; MGMGs, monogalactosylmonoacylglycerols; DGDGs, digalactosyldiacylglycerols; SQDGs, sulfoquinovosyldiacylglycerols; DAGs, diacylglycerols, TAGs, triacylglycerols, MeOH, methanol; CHCl₃, chloroform; H₂O, water; IPA, isopropanol; HPLC, high performance liquid chromatography; ESI, electrospray ionization; +ESI, positive mode; -ESI, negative mode; N₂, nitrogen gas; MPP, Mass Profiler Professional; FDR, false discovery rate; MS/MS, tandem mass spectrometry; 13-PCB-52, ¹³C-polychlorinated biphenyl; SRM, selected reaction monitoring; SIM, selected ion monitoring; PBRs, photobioreactors; PAs, phosphatidic acids; PSs, phosphatidylserines; PCs, phosphatidylcholines; MGDGs, monogalactosyldiacylglycerols; LC-Q-ToF MS/MS, liquid chromatography quadrupole-time-of-flight tandem mass spectrometry; MAGs, monoacylglycerols; PSI, photosystem I; PSII, photosystem II; NMR, nuclear magnetic resonance.

1. Introduction

Department of Energy (DoE) Algal Biofuels Technology roadmap fully engages algal research and development, and addresses challenges associated with producing algal biofuels at commercial scale [1]. On this track, it is critical to adjoin all environmental sustainability components algae offers with prospects of economical feasibility to compete with petroleum based fuels [2]. To achieve commercially applicable rates of return on algal biofuels, economic feasibility models suggest that there is a need to produce value-added chemicals within an integrated biorefinery scheme [3-5]. This requires targeting not only lipids as biodiesel feedstock, but also other biomolecules having higher value per dry biomass weight with potential applications such as food additives, health supplements, and pharmaceuticals [6]. Key examples of value added algal products include phytosterols,[7] carotenoids,[8] and vitamins [9]. Phytosterols are steroid alcohols that function in cell proliferation and signal transduction in microalgae [10]. The commercial interests in algal phytosterols are due to their potential for lowering total and low-density lipoprotein cholesterol [7]. Carotenoids, such as β -carotene, lycopene, and zeaxanthin, have wider consumer applications and are utilized in food coloring/nutrition and cosmetics [6]. Vitamins C and E are also among the high-value products pursued from algae [11]. Vitamin E, synthesized only by photosynthetic organisms, is a lipid-soluble metabolite with antioxidant and radical scavenging activities [11].

Expanding on value-added biochemical portfolios often requires exploration and utilization of novel algae species, including unique non-model species. Among more than 3,000 species of algae isolated by the DoE Aquatic Species Program (ASP), fifty-one varieties were characterized as potential high-value strains [12]. However, cellular and molecular characterization of these non-model species still lags behind established model organisms such

as *Chlamydomonas reinhardtii*, *Chlorella variabilis*, *Micromonas*, *Nannochloropsis gaditana*, *Nannochloropsis oceanica*, *Ostreococcus tauri*, *Phaeodactylum tricornutum*, *Thalassiosira pseudonana*, and *Volvox carteri* [13]. In this work, we provide the first comprehensive metabolomics analysis of the non-model green microalgae species *Ettlia oleoabundans* (formerly known as *Neochloris oleoabundans*). *E. oleoabundans* is a Chlorophyceae class green microalgae with unique characteristics. Originally isolated from a sand dune ecosystem, it can be grown both in freshwater[14] and seawater [15, 16] suggesting a strong adaptability to different growth conditions and is a high lipid producer [17-20]. Generating and analyzing high throughput metabolomics data will assist in improving our knowledgebase in novel non-model microalgae domain and gradually enable *E. oleoabundans* to be considered as a candidate model species.

Here, we provide a step-by-step methodology, from sample preparation to data analysis, for untargeted and targeted metabolomics [21-23] of a non-model algae species utilizing liquid chromatography quadrupole-time-of-flight mass spectrometry (LC-Q-ToF MS) and gas chromatography tandem mass spectrometry (GC-MS/MS). Nutrient variations, especially nitrogen, are known to affect primary and secondary metabolisms in algae and were used in this study as a strategy to optimize yields of lipids and other value-added biochemicals [24-27]. For this reason, we focused on comparative analysis of nitrogen deplete (N-) and nitrogen replete (N+) growth conditions to test the metabolomic coverage we can achieve with the method we developed. Our protocols allowed simultaneous detection and comparative analysis of major lipid families, including fatty acids (FAs), phosphatidylglycerols (PGs), phosphatidylethanolamines (PEs), phosphatidylinositols (PIs), monogalactosylmonoacylglycerols (MGMGs), digalactosyldiacylglycerols (DGDGs),

98 sulfoquinovosyldiacylglycerols (SQDGs), diacylglycerols (DAGs) and triacylglycerols (TAGs)
99 in addition to key biofuel feedstock lipids and other hydrophobic metabolites such as
100 chlorophylls, terpenes, phytosterols, and vitamins. Following this methodology, we analyzed the
101 changes in metabolite composition in *E. oleoabundans* under N⁻ conditions and found that
102 TAGs and plant specific lipids such as SQDGs strongly accumulated, whereas chlorophylls and
103 DGDGs were depleted in N⁻. The broad coverage of information obtained at the molecular level
104 paves the way towards experimental designs where molecule-specific interpretation of the
105 changes in metabolite composition will lead to detailed understanding of the metabolic networks.

106 **2. Materials and Methods**

107 **2.1 Materials.** LC-MS grade Methanol (MeOH) and HPLC grade isopropanol (IPA) were
108 purchased from EMD Millipore (Billerica, MA). Formic acid, ammonium formate and
109 ammonium hydroxide were purchased from Sigma Aldrich (St. Louis, MO). HPLC grade
110 chloroform (CHCl₃) was purchased from Honeywell Burdick & Jackson (Muskegon, MI).
111 Distilled water was provided by a Milli-Q purification system. A detailed list of all standards
112 purchased for LC-Q-ToF MS and GC-MS/MS are presented in supporting information. Glass
113 beads (0.5 mm) were purchased from Bertin Technologies (Montigny, France). Stock solutions
114 of the standards were prepared in CHCl₃ and stored at -80°C.

115 **2.2 Microalgae Seeding, Growth, and Harvesting.** *E. oleoabundans* (UTEX 1185) was
116 obtained from Culture Collection of Algae at The University of Texas at Austin. Algal cultures
117 were maintained and grown axenically with Modified Bold 3N growth medium [28]. For each
118 experimental condition, five biological replicates were used. Cultures were grown in 500 mL
119 glass photobioreactors (PBRs) operated at room temperature (25°C ± 2°C), and with a 14:10 h
120 light:dark cycle of exposure to fluorescent light (32 Watt Ecolux, General Electric, Fairfield, CT)

at a photosynthetic photon flux density of $110 \mu\text{mol-photon m}^{-2}\text{s}^{-1}$. PBRs were continuously aerated with sterile, activated carbon filtered air at a flow rate of 200 mL/min using a mass flow controller (Cole-Parmer Instrument Company, IL). No additional CO_2 was provided to air supply.

For experimental conditions, *E. oleoabundans* was grown under nitrogen replete (N+) (7.1 mM, 603.46 $\mu\text{g/mL}$), and nitrogen deplete (N-) (0.71 mM, 60.35 $\mu\text{g/mL}$) as the total $[\text{NO}_3^-]$ was adjusted in growth media. Five biological replicates, harvested from different PBRs, were used for each condition. The nitrate concentration of culture media was determined daily by NECi Microplate Nitrate Kit (Cat # M-NTK-301, Lake Linden, MI). Algal growth was monitored by the optical density measured daily at 730 nm using a microwell-plate reader (SpectraMax i3, Molecular Devices, CA). The microalgae were harvested, centrifuged, decanted, and frozen at -80°C until further analyses. Different aliquots, from each biological replicate, were prepared for biomass calculations, determination of total lipid composition and metabolite extraction. The biomass corresponding to each samples is reported in Table S1.

2.3 Metabolite extraction. Approximately 30-60 mg of microalgae biomass was spiked with 700 μL of MeOH containing squalene-*d*6, estrone-*d*4 and oleic acid-*d*9. The final concentrations of standards were 300-660 ng/mL for squalene-*d*6 and estrone-*d*4 and 8-14 $\mu\text{g/mL}$ for oleic acid-*d*9 depending on the biomass of each sample. These isotope-labeled standards were used to calculate extraction efficiencies. The percent recoveries are reported in Table S2. Glass beads (250 mg) were added to this solution and vortexed five times (1 min each, with 1 min incubation on ice between the cycles). This suspension was transferred into a glass dounce homogenizer and the beads were rinsed with additional 400 μL of methanol by vortexing for an additional cycle; all the extracts were combined. This solution was then mixed with CHCl_3 and H_2O

(MeOH/CHCl₃/H₂O, 1:2:1), and homogenized on ice (×40 strokes). After homogenization, the solution was centrifuged at 4°C, 3000 $\times$ g for 40 min, and the organic layer was separated from the aqueous layer and the protein disk. To improve extraction efficiency, the remaining protein disk and the aqueous layer was extracted two more times using additional CHCl₃. All organic layers were combined and evaporated to complete dryness. The remaining metabolite residues were normalized based on biomass and resuspended in CHCl₃ accordingly. The normalized concentration was 9.20 mg of biomass/mL. The hydrophobic metabolite extracts were stored in -80°C until analysis.

2.4 Sample preparation and analysis. For untargeted profiling and targeted analysis of plant lipids, FAs and DAGs, original solutions were used. For targeted analysis of TAGs, a 20-fold dilution was carried out for each sample. The injection volume was 30 μ L corresponding to metabolites extracted from 0.28 mg or 0.014 mg (for TAG analysis) biomass before extraction.

2.5 LC-Q-ToF MS analysis of extracts. Analyses were carried out using an HPLC-Q-ToF MS system equipped with Infinity 1260 HPLC (Agilent Technologies) consisting of a solvent rack degasser, auto sampler, binary pump with solvent selection valve, a column oven and an isopump coupled to the Q-ToF 6530 dual Jetstream electrospray (ESI) (Agilent Technologies).

There was a continuous flow of a reference solution containing purine and hexakis-(1H,1H,3H-tetrafluoropropoxy)phosphazine used as reference masses for mass correction and correction for other instrumental drifts. The reference m/z 's for positive mode were 121.0509 and 922.0098 and for negative mode were 112.9856 and 1033.9881. The instrument was calibrated before each run according to manufacturer's instructions. The calibration masses (m/z 's) for positive mode were 118.0862, 322.0481, 622.0289, 922.0098, 1221.9906, and 1521.9715 and for negative mode were 112.9856, 301.9981, 601.9789, 1033.9881, 1333.9689, and 1633.9498.

In each run, samples from different growth conditions (n=5, for each condition) were analyzed in an alternating manner. Mobile phase A was composed of 95:5 H₂O/MeOH and mobile phase B was made of 60:35:5 isopropyl alcohol (IPA)/MeOH/H₂O, supplemented with 0.1% ammonium hydroxide (-ESI) or 0.1% formic acid and ammonium formate (+ESI). The starting flow rate was 0.1 mL/min of 100% A, which was then increased to 0.5 mL/min after 5 minutes; the mobile phase composition was changed from 100% A to 100% B from 5 to 45 min, and was maintained at 100% B for 10 min. The mobile phase was then equilibrated back to 100% A and kept there for 8 min at 0.5 mL/min. The total run length was 63 min.

For analysis under positive ESI mode (+ESI) a Phenomenex C5 Luna column (50 x 4.6 mm, 5 μ m) was used for chromatographic separation, while a Phenomenex C18 Gemini column (50 x 4.6 mm, 5 μ m) was used for analysis under negative ESI mode (-ESI). The capillary and fragmentor voltages were 3500 V and 175 V, respectively. The drying gas (N₂) temperature and flow rate were 350°C and 12 L/min, respectively. Finally, the nebulizer pressure was 30 psi. Species with an *m/z* range of 50-1700 were detected in extended dynamic range.

2.6 Untargeted profiling of hydrophobic metabolites. Two independent profiling experiments, each with five biological replicates from the two growth conditions (N⁻ and N⁺), were performed. The raw data was imported into MassHunter Profinder (version B.06.00, Agilent Technologies) for peak alignment and extraction of unique species. The resulting data was analyzed in Agilent Mass Profiler Professional (MPP, version B12.6.1) for significance and fold change determination. The species were grouped, filtered by frequency (species kept if present in 80% or 4 out of 5 replicates for one sample group), and statistically significant species were identified (unpaired *t*-test, *P* < 0.05). Species with < 3 fold change were eliminated. The resulting list of species was then compared between two independent profiling experiments and

reproducible species were identified. The chromatographic peaks of these species were manually interrogated in Profinder and reintegrated when necessary. The species with poor mass accuracy (> 20 ppm) and low abundances ($< 10^4$ ion counts) were eliminated.

2.7 LC-Q-ToF MS/MS identifications. The resulting 202 species were searched in METLIN [29] and LipidMaps [30] databases to identify a candidate molecule based on accurate mass. Known lipid standards (or lipids belonging to the same lipid families) were purchased for the candidate lipids. The standards used are listed in Table S3. MS/MS data is collected similar to described above in “**LC-Q-ToF MS analysis of extracts**”, except data acquisition was in 4 GHz high-resolution mode and fragments were observed at collision energies of 15, 35, 55, and 75 eV. The MS/MS fragmentation patterns of the species of interest and candidate lipids were compared. Searches based on MS/MS fragments provided in METLIN were also used in addition to the fragmentation patterns of known standards.

2.9 Targeted analysis of major lipid families. For targeted analysis of different FAs, DAGs, TAGs, glycerophospholipids, SQDGs, DGDGs, and MGDGs, the most abundant adducts ($[M-H]^-$ for FAs, SQDGs, DGDGs, MGDGs, PGs, PAs, PSs, and PIs; $[M+H]^+$ for DAGs and PCs and $[M+NH_4]^+$ for TAGs) were identified and extracted in Agilent Qualitative Analysis version B.06.00. Each peak was integrated and the abundances were calculated from the areas under the curve. Fold changes, abundances, m/z 's, adducts observed and retention times are provided in Table S4 and S5. Fold changes were calculated accordingly. Structural confirmation of the lipids analyzed by targeted analysis was carried out as described in the “**LC-Q-ToF MS/MS identifications**” section.

2.10 Derivatization of hydroxylated metabolites for GC-MS/MS analysis and their analysis. 25 μ L aliquots of the *E. oleoabundans* extracts were transferred to 2 mL amber vials for selective

derivatization of the hydroxylated metabolites to *tert*-butyldimethylsilyl substituents. These metabolites included α -tocopherol, β -sitosterol, stigmasterol, vitamin A, campesterol, oxidosqualene, which ring opened (see GC-MS/MS section for details), and the surrogate estrone-*d*4. For the derivatization reaction, each extract was evaporated to dryness under N₂ (g), and dissolved in 500 μ L of acetonitrile and 50 μ L of *N*-*tert*-butyldimethylsilyl-*N*-methyltrifluoroacetamide, vortexed, and incubated at 80°C for 30 min. After derivatization, extracts were transferred to vials with glass inserts. Following, they were fully evaporated to dryness and resuspended in 50 μ L isooctane that was spiked with 50 ng/mL of ¹³C-polychlorinated biphenyl (13C-PCB-52) as an internal standard. Sample analysis was performed on a Trace GC Ultra coupled to a TSQ Quantum XLS triple quadrupole MS system (Thermo Fisher Scientific, West Palm Beach, FL.).

Separation was achieved on a 30 m DB-5HT capillary column, with a 0.25-mm inner diameter, and a 0.10- μ m film thickness (Agilent Technologies, Santa Clara, CA). The oven temperature programming started at an initial temperature of 125°C, which was held for 3 min, and then ramped at 30°C/min to 350°C, and held at this temperature for 3 min. A programmable temperature vaporizer injector was used with a 2 μ L injection volume. Helium (99.99% purity) was used as the carrier gas with a constant flow rate of 1 mL/min.

The GC-MS/MS was initially operated under the full scan mode with an ionization energy of 70 eV to obtain the retention times and mass spectra of each derivatized analyte. To optimize MS conditions for selected reaction monitoring (SRM), product ion scans were conducted from 10-30 eV at 5 eV intervals to obtain the collision energies that will give the highest signal-to-noise (S/N) ratios for the selected fragment ions, but retaining a 5-10% relative intensity for the precursor ions (the base peaks from the full MS scan).

The base peak for each derivatized phytosterol (stigmasterol, β -sitosterol, and campesterol) was the loss of the *t*-butyl moiety, $[M-C(CH_3)_3]^+$, for α -tocopherol was $[M]^+$, and for vitamin A was $[M-CH_2OSi(CH_3)_2C(CH_3)_3]^+$ (Table S6). Although oxidosqualene does not have a hydroxyl substituent, the epoxide moiety of the molecule ring opens upon derivatization [31], and leads to the subsequent derivatization of the hydroxyl group, pertaining to either its base peak of $m/z = 341$ or the loss of $[M-C_5H_{10}OSi(CH_3)_2C(CH_3)_3]^+$.

During the analysis of extracts, which contained low concentrations of squalene, selected ion monitoring (SIM) of the base peak for squalene ($m/z=69$, $[C_5H_9]^+$) and its second most abundant fragment ion ($m/z=81$, $[C_6H_9]^+$), were monitored to improve the detection limit for squalene. The fragmentation of squalene-*d*6 also included $m/z=69$ and 81, along with its distinctive $m/z=75$ ($[C_5H_3D_6]^+$) from the deuterated species. As such, for squalene-*d*6, $m/z=75$ was used for quantification. In order to account for the overlapping signals of squalene-*d*6 and squalene due to the presence of the same fragment ions, $m/z=69$ and 81, the calibration curve for squalene was adjusted by subtracting the contribution of $m/z=69$, 81 from squalene-*d*6 on the squalene signal.

Once the m/z values for the SRM transitions were identified, a GC-MS/MS quantification method based on SRM was created for the analysis of all targeted analytes, with the exception of squalene and squalene-*d*6, which were analyzed using SIM. Therefore, a separate segment was created in the MS acquisition method to perform SIM for squalene and squalene-*d*6, while the rest of the target compounds were analyzed under SRM ($n=5$ for N- and N+).

2.11 Percent recoveries of oleic acid-*d*9, squalene-*d*6 and estrone-*d*4. *E. oleoabundans* ($n=3$, for both the N- and N+ conditions) pellets were spiked with 17.49 μ g/mL oleic acid-*d*9, extracted as described above and normalized based on biomass. Matrix-matched calibration curves were prepared in N- and N+ extracts by using concentrations ranging from 1.45-29.1

μg/mL oleic acid-*d*9. Calibration curves of the concentration vs. abundance were created after manually integrating the extracted ion chromatograms (EICs) of the oleic acid-*d*9 ([M-H]⁺, *m/z*=290.3046) in each extract. These extracts spiked after the extraction at different concentrations were later analyzed by LC-Q-ToF MS along with the extracts spiked prior to the extraction to calculate the extraction efficiency. The abundances were then fit into the calibration curve to calculate the concentration after the extraction.

Stock standard solutions of squalene-*d*6 and estrone-*d*4 were combined to make a 6-pt matrix matched calibration curve ranging from 0.03125 to 1 μg/mL to determine recoveries of the squalene-*d*6 and estrone-*d*4 surrogates by GC-MS/MS. In order to account for the different matrix of the two nutrient conditions, a blank extract for each matrix was used for the calibration curves for samples from N⁻ and N⁺ PBRs, respectively, with good linearity ($R^2 > 0.995$) for both surrogates. The surrogate used for the derivatized analytes was estrone-*d*4, while squalene-*d*6 was used for the quantification of squalene. A labeled surrogate for each targeted analyte would be most ideal for quantification, however deuterated or other isotopically-labeled standards were not available for all of the target analytes for mass spectrometric detection.

3. Results and Discussion

3.1 Analysis of biomass and total lipid content under varying N conditions. *E. oleoabundans* cells were grown under N⁺ and N⁻ conditions and harvested at early stationary phase (day 7, Figure 1a). Five biological replicates from different PBRs were used for analyses. At this sampling point, there was a significant decrease in growth rate ($P < 0.01$) of *E. oleoabundans* in the PBRs with N⁻ conditions due to limited nitrogen resources (Figure 1a). The NO₃⁻ concentration in N⁺ was 154 μM, and below the detection level in N⁻ (<36 μM), indicating that the limited condition was reached (Figure 1b). Total lipids were extracted and quantified based

on dry biomass of the microalgae (Table S1). Consistent with our previous studies [32], there was a significant increase ($P = 0.015$) in lipid content from 23% of dry cell weight (DCW) in N⁺ to 30% in N⁻ conditions (Figure 1c).

3.2 Extraction of metabolites and targeted LC-Q-ToF MS analysis. To study the nature of changes in lipid composition under varying NO₃⁻ concentrations and identify the lipid species that contribute to this significant difference under N⁻ using comparative metabolomics, we extracted hydrophobic metabolites from *E. oleoabundans* (Figure 1d). Extraction efficiencies in N⁻ and N⁺ matrices were determined based on the recoveries of spiked oleic acid-*d*9. The average recoveries of oleic acid-*d*9 were 86% in N⁻ and 87% in N⁺ (Table S2).

Figure 1. Characterization of cell growth in the N⁺ and N⁻ conditions and schematics for sample preparation and hydrophobic metabolite profiling. (a) Growth curve of *E. oleoabundans* under N⁻ and N⁺ conditions. Cells at early stationary phase (day 7) were used for metabolic profiling. (b) Remaining [NO₃⁻] in the growth media showing that the N⁻ condition has limited [NO₃⁻] (c) Percent lipid composition of cells grown in the N⁻ (29.9 +/- 2.5) and N⁺ (22.9 +/- 4.5) conditions. (d) *E. oleoabundans* pellets were harvested and spiked with internal standards prior to extraction. A biphasic extraction yielded separation of hydrophobic and hydrophilic metabolites. Hydrophobic metabolites were analyzed further.

Using standards that are representative of the metabolites of interest (Table S3), we utilized reversed-phase liquid chromatography for separation of analytes in complex microalgal extracts. Figure S1 shows the mass (amu) vs. retention time of unique species detected in any one of the two experimental conditions, *i.e.* N⁻ vs. N⁺. We observed that species started eluting at the beginning of the gradient (95:5, H₂O/MeOH) and continued to elute throughout the gradient. This continuous elution profile throughout the gradient indicates that despite the complexity of the hydrophobic extracts, the mobile and stationary phases chosen allowed sufficient separation of the unique species detected in the samples.

We analyzed species from major lipid families in two sets of positive and negative LC-Q-ToF MS runs, which allowed comprehensive analysis of different lipid families (Table 1, Table S4, Table S5). We targeted FAs, DAGs, and TAGs, glycerophospholipids such as phosphatidic acids (PAs), phosphatidylserines (PSs), PGs, PIs, phosphatidylcholines (PCs), and plant lipids such as SQDGs, DGDGs, and monogalactosyldiacylglycerols (MGDGs). We analyzed the levels of these lipids in five biological replicates (*i.e.*, algae that have been grown independently in different PBRs) for both N⁻ and N⁺ in two independent LC-Q-ToF MS runs. This multi-layered technical and biological replication strategy helped us to increase and assess the reproducibility of our analyses. We present data from all 20 data points (five biological replicates, two conditions, and two independent runs) for each lipid species in Table S5. We observed that the standard deviation in abundances is less than 10% within biological replicates in ~84% of the lipid species. We attribute the remaining relatively high standard deviations observed within biological replicates to actual biological variance in the samples for these lipid species. To formally assess the technical robustness of our approach, we compared the fold

changes in two independent technical runs for lipids analyzed (Figure S2). Indeed, we found very high correlation with $R^2=0.9867$ validating the robustness of our approach.

After confirming the identity of these species by LC-QTOF-MS/MS, we investigated the changes in the abundance of different lipid families. We observed moderate depletions of FAs (1.1-3.3 fold) under the N⁻ growth condition independent of the acyl chain length and saturation states. DAGs, on the other hand, showed both accumulation and depletions under the N⁻ condition (Table 1, Table S4). There seems to be an overall accumulation in species with lower unsaturation states compared to highly unsaturated species. For instance, for C32 and C34 DAGs, species with two or more double bonds were depleted; whereas, saturated species were slightly enriched under the N⁻ condition (Table 1). On the other hand, C36 DAGs showed an overall accumulation, except for C36:4 DAG which was significantly depleted in the N⁻ condition (Table 1). TAG accumulation under nutrient limitation has previously been shown in microalgae [33-37]. Our results are consistent with these observations. Except for C48:0 TAG, all the TAG species that were analyzed accumulated in the N⁻ condition. Their accumulation was substantial (2.5-7.9 fold) and significant ($P < 0.001$) (Table 1). Other glycerophospholipid we targeted including PAs, PSs, PIs and PCs did not exhibit any trends in their changes (Table S4).

342 Table 1. Targeted LC-Q-ToF MS analysis of FAs, DAGs, TAGs, PGs and algal lipids

343

Lipid family	lipid	<i>m/z</i>	<i>t_R</i>	FC ^[a] (N-/N+)	<i>P</i> ^[b] value	lipid	<i>m/z</i>	<i>t_R</i>	FC ^[a] (N-/N+)	<i>P</i> ^[b] value
FA ^[c]	C12:0	199.1704	20	1.5	***	C18:0	283.2637	28	0.9	ns
	C14:0	227.2011	24	1.3	**	C20:5	301.2168	21	1.2	ns
	C16:1	253.2168	22	0.5	****	C20:4	303.2324	21	1.1	ns
	C16:0	255.2324	26	0.7	*	C20:1	309.2799	28	0.3	****
	C18:3	277.2168	22	0.5	****	C20:0	311.295	24	0.9	ns
	C18:2	279.2324	25	0.6	**	C24:0	367.3576	33	0.9	ns
	C18:1	281.2481	26	0.7	**					
DAG ^[d]	C32:2	547.4729	43	0.6	****	C34:1	577.5196	44	1.1	*
	C32:1	549.4883	43	1.3	ns	C36:4	599.5039	41	0.2	****
	C32:0	551.5039	43	2.2	***	C36:3	601.5196	44	6.4	****
	C34:4	571.4729	43	0.3	****	C36:2	603.5352	44	4.3	****
	C34:3	573.4883	43	0.2	****	C36:1	605.5509	45	3.4	****
	C34:2	575.5039	43	0.4	****					
TAG ^[e]	C48:0	824.7707	49	0.9	****	C54:4	900.8020	49	3.4	****
	C50:2	848.7707	49	4.5	****	C54:3	902.8177	49	5.0	****
	C50:1	850.7864	49	5.3	****	C54:2	904.8333	50	5.8	****
	C52:2	876.8020	49	6.3	****	C54:1	906.8490	50	4.3	****
	C54:6	896.7707	48	2.5	****	C56:2	932.8646	50	7.9	****
	C54:5	898.7864	49	3.5	****					
SQDG ^[c]	C32:3	787.4672	32	Accum. ^[f]		C34:2	817.5141	34	25.0	****
	C32:2	789.4828	33	Accum. ^[f]		C34:1	819.5298	35	8.6	****
	C32:1	791.4985	34	Accum. ^[f]		C34:0	821.5454	36	Accum. ^[f]	
	C32:0	793.5141	35	1.7	****	C36:6	837.4828	33	Accum. ^[f]	
	C34:5	811.4672	32	Accum. ^[f]		C36:5	839.4985	33	15.0	****
	C34:4	813.4828	33	Accum. ^[f]		C36:4	841.5141	34	Accum. ^[f]	
	C34:3	815.4985	34	57.9	****					

DGDG^[c]	C32:4	883.5424	42	Deplete ^[f]		C34:2	915.6050	45	0.9	****
	C32:3	885.5581	43	Deplete ^[f]		C34:1	917.6207	46	0.9	**
	C32:2	887.5737	44	0.3	****	C36:6	935.5737	43	0.2	****
	C34:5	909.5581	43	0.4	****	C36:5	937.5894	44	0.2	****
	C34:4	911.5737	43	0.5	****	C36:4	939.6050	45	0.1	****
	C34:3	913.5894	44	0.5	****					
MGDG^[c]	C28:0	673.4896	30	4.4	****	C34:3	751.5366	45	1.9	****
	C30:2	697.4896	29	4.0	****	C34:2	753.5522	46	1.3	****
	C32:4	721.4896	43	0.4	****	C36:6	773.5209	44	1.1	ns
	C34:6	745.4896	43	0.9	****	C36:5	775.5366	45	0.5	****
	C34:5	747.5053	44	1.5	****	C36:4	777.5522	45	0.3	****
	C34:4	749.5209	44	2.2	****					
PG^[c]	LPG C16:0	483.2723	26	Deplete ^[f]	****	C34:2	745.5020	35	0.2	****
	LPG C18:1	509.2879	27	Deplete ^[f]		C34:3	743.4863	34	0.3	****
	C32:1	719.4863	35	1.5	****	C36:4	769.5020	34	0.7	****
	C32:0	721.5020	35	0.7	****	C36:2	773.5333	36	2.1	***
	C34:1	747.5176	35	0.4	****					

344 Abbreviations: t_R, retention time; FC, fold change; N⁻, nitrogen deplete; N⁺, nitrogen replete; FA, fatty acid; DAG, diacylglycerol;
345 TAG, triacylglycerol; SQDG, sulfoquinovosyldiacylglycerol; DGDG, digalactosyldiacylglycerol; MGDG,
346 monogalactosyldiacylglycerol; PG, phosphatidylglycerol.

347 ^a FC was determined as [Abundance N⁻] / [Abundance N⁺] for each lipid. Abundance is the total ion count for a given ion. Each ion
348 corresponds to a mass-to-charge ratio (*m/z*), which is used to assign the lipid specie.

349 ^b ns, *P*>0.05; *, *P*<0.05; **, *P*<0.01; ***, *P*<0.001; ****, *P*<0.0001.

350 ^c *m/z*'s corresponding to [M-H]⁻. ^d *m/z*'s corresponding to [M+H]⁺

351 ^e *m/z*'s corresponding to [M+NH₄]⁺. ^f Fold changes could not be determined due to the low abundance in either N⁻ or N⁺ samples.

3.3 GC-MS/MS analysis of value-added biochemicals. Phytosterols, squalene, oxidosqualene, and vitamins were not detected by LC-Q-ToF MS analysis because of their poor ionization efficiencies under electrospray MS and the signal suppression caused by the complexity of the sample matrix. For these compounds, we used GC-MS/MS, which requires chemical derivatization (except for squalene). This approach provides better sensitivity for quantitative analysis of phenol-containing compounds. Specifically, the levels of squalene, oxidosqualene, vitamin A, α -tocopherol, campesterol, stigmasterol, and β -sitosterol were investigated using GC-MS/MS and the fold changes were calculated. Extraction efficiencies were determined based on the recoveries of spiked squalene-*d*6 and estrone-*d*4 in N⁻ and N⁺ matrices using GC-MS/MS. The average recoveries for squalene-*d*6 and estrone-*d*4 were 39% and 67%, respectively, for both N⁻ and N⁺ matrices (Table S2). Once the *m/z* values for the transitions were identified, a GC-MS/MS quantification method based on selective reaction monitoring mode (SRM) was created for the analysis of all targeted analytes, with the exception of squalene and squalene-*d*6, which were analyzed using selection ion monitoring (SIM). The observed retention times, ions monitored, and collision energies for all analytes are presented in Table S6. We did not observe any significant changes in the concentrations of phytosterols and vitamin A (Table 2) between N⁻ and N⁺ matrices. However, squalene and α -tocopherol showed a 2.8- and 3.4-fold accumulation under N⁻ conditions, respectively (Table 2, Table S6).

371 **Table 2. GC-MS/MS analysis of phytosterols, and other value-added biochemicals.**

Metabolite	t_R (min)	m/z	FC^[a] (N-/N+)	Std. Dev.
Vitamin A	6.01	255	1.51	0.38
Oxidosqualene	8.25	341	1.36	0.18
Squalene	9.13	69	2.79	0.59
α-tocopherol	10.46	545	3.44	1.71
Campesterol	10.76	457	1.21	0.27
Stigmasterol	10.83	469	0.88	0.09
β-sitosterol	10.96	471	0.95	0.10

372 Abbreviations: t_R, retention time; FC, fold change

373 ^a FC was determined as [concentration in N-] / [concentration in N+] for each metabolite

374

3.4 Untargeted LC-Q-ToF MS analysis to identify metabolites that show significant differentiation under the N⁻ condition. To test if there are global changes in metabolite composition under nitrogen limitation, we comparatively analyzed the MS-based metabolite profiles of *E. oleoabundans* cells grown under the N⁻ and N⁺ conditions. We produced a heat map by hierarchical clustering. Plotting the ion abundances showed a high correlation among five biological replicates of N⁻ and N⁺ samples (Figure 2). Overall, we observed three distinct clusters based on abundance: metabolites that are differentially regulated, either enhanced or depleted (Figure 2, clusters that are presented by an asterisk), and metabolites that do not exhibit strong changes between the N⁻ and N⁺ conditions (Figure 2, represented by full circles).

385

386

387 **Figure 2. Significant differences in metabolite composition between N- and N+ conditions.**

388 We observed distinct clustering of metabolites into three distinct groups among N- and N+
389 conditions without any *a priori* input, suggesting major compositional differences.

390

These analyses suggested significant differences in the metabolite profiles of algae grown under the N⁻ and N⁺ conditions. To investigate this further, we performed untargeted, comparative LC-Q-ToF MS-based analysis of the metabolome (Figure 3a). After alignment of chromatograms belonging to different samples and extraction of unique species based on different adducts, resulting species were analyzed based on their statistical significance and fold changes. Species that show significant differential ($P < 0.05$) exhibiting a fold change of >3 between the N⁻ and N⁺ conditions, and also are reproducible in two independent profiling runs were investigated (n=5 for each condition and each independent profiling experiment). Afterwards, we manually inspected the resulting species and reintegrated the chromatograms if necessary (Figure 3b). Based on the updated abundances, the fold changes and P values were recalculated. The resulting refined list of species was then subjected to more stringent statistical criteria: a t -test at 99% confidence interval followed a Benjamini-Hochberg FDR (false discovery rate) [38] multiple testing corrections. At the end of this analysis, we obtained a total of 202 unique species (Figure 3c, Table S7). We carried out targeted MS/MS experiments to determine the identity of these species and annotated 163 (81%) of them (Figure 3c, Table S7). A majority of these species (46%) were chlorophylls and their derivatives, 15% were plant lipids such as SQDGs, MGMGs and DGDGs, 11% were simple lipids such as FAs and derivatives, 15% were other simple lipids (MAGs, DAGs and TAGs), 6% were glycerophospholipids, and 7% were amino acids, vitamins, terpene derivatives, sphingolipids, and phytosterols (Figure 4). Fragmentation patterns of representative metabolites can be found in Figure S3. A complete list of fragments observed for annotated species is provided in Table S7.

Figure 3. Data analysis workflow for LC-Q-ToF MS-based untargeted metabolic profiling.

(a) Two independent profiling experiments were performed in positive (+ESI) and negative (-ESI) ionization modes. LC-Q-ToF MS chromatograms were aligned, species detected were grouped based upon different adducts and unique species were identified in Profinder (Agilent Technologies). (b) Statistical analysis and filtering based on significance and absolute fold change were conducted. Species that were differentially present in the N⁻ condition compared to the N⁺ condition ($P < 0.05$) were identified in two independent profiling runs (MPP, Agilent). (c) After manual investigation, we identified 202 species. 163 of these species were annotated by LC-Q-ToF MS/MS (Complete list is provided in Table S7).

3.5 Depletion of chlorophylls under the N⁻ condition. Chlorophyll and chlorophyll derivatives constitute 46% of the changes we observe under the N⁻ condition. Under nutrient limitation, cellular proliferation rate decreases as well as the overall photosynthetic activity [39]. Consistent with this phenotype, most of the chlorophylls (66/75) accumulate in N⁺ condition (Figure 4, Table S7). Our findings provide information on specific chlorophyll species and suggest that they were regulated differently under N⁻ condition. Moreover, based on the accompanying MS/MS fragmentation data, we infer that some of the chlorophyll derivatives are PG modified chlorophylls and porphyrins. In addition to plant lipids such as SQDGs, DGDGs and MGDGs, PGs are the only major phospholipids found in the thylakoid membrane where the photosynthetic reactions and electron transport takes place. X-ray crystallographic data showed the presence of PGs in photosystem I (PSI) and photosystem II (PSII) [40]. It has been suggested that PGs, along with other lipids, are important for the formation and stabilization of membranes and the assembly of the subunit of PSI and PSII.[40] In *C. reinhardtii* mutants where PSII was inactive, some PGs were absent and thylakoid membrane structures were compromised [41]. The observed decrease in concentrations of chlorophylls, their derivatives, and PGs are consistent with the decreased photosynthetic activity under the N⁻ condition, and reinforces the significant role of PGs during photosynthesis.

3.6 Changes in the composition of plant lipids under the N⁻ condition. SQDGs, DGDGs, and MGDGs constitute the majority of the photosynthetic membranes in microalgae. Among these plant-specific lipids, SQDGs show profound accumulations under the N⁻ condition (Figure 4, Table S7). The impact of nutrient limitation on SQDG levels is not widely studied, and the effect on SQDG varies greatly depending on the organism and environmental stress [39, 42]. Unlike SQDGs, DGDGs were strongly depleted under the N⁻ condition (Figure 4, Table S7). Our

447 findings suggest a complex regulation of the metabolic pathway in which the biosynthesis of
448 thylakoid membrane lipids such as SQDGs are activated *via* the use of DAGs in an attempt to
449 restore photosynthetic efficiency.

450

Figure 4. Results of untargeted profiling: Metabolites that are differentially regulated in cells grown under N⁻ compared to N⁺ conditions. 163 species (81% of total hits) were annotated by MS/MS and their distribution reported. A complete list of these species with their *m/z*'s, fold changes, and retention times and MS/MS fragments can be found in Table S7. Representative fragmentation patterns are provided in Figure S3.

4. Conclusions

Uncompetitive prices of algal biofuels compared with its petroleum-based counterparts commenced a biorefinery concept to harvest multiple value-added biochemicals from the biomass [43-45]. Intuitively, such an approach requires better understanding of cellular and molecular mechanisms that leads to biochemical regulations that result in the changes in metabolite composition. Recent advancements in comprehensive and integrative characterization of metabolite composition have opened doors to understanding the metabolic regulation of different value-added bio-chemicals [22, 23]. Therefore, in this study we developed a combined LC- and GC-MS-based method to not only analyze neutral lipids as biodiesel feedstock but also other hydrophobic metabolites with high market value and as food supplements and pharmaceutical ingredients. Investigating percent recoveries of internal standards that are spiked to biological replicates prior to sample preparation and analysis (Table S2) and consistency of the fold changes we measured in two independent runs (Figure S2) validate the reproducibility and robustness of this method. Studies by us[32] and others have already established the link between nutrient limitation and increased lipid levels in different green microalgae. Specifically, studies with model microalgae species such as *C. reinhardtii* [46-48], *Nannochloropsis*[49, 50] and *Chlorella*[50] have shown changes in lipid composition under nitrogen limitation, which vary depending on the organism and the extent of nitrogen deprivation. The study presented here provides a template to conduct high-throughput metabolomics to investigate oleaginous non-model microalgae, *E. oleoabundans*, a key valuable strain and commercially viable candidate for biofuels production as determined by DoE [12]. In addition, *E. oleoabundans* is capable of accumulating high-value chemicals explored herewith to be used in biorefinery platforms towards a more sustainable, bioeconomy-driven future.

We provide a set of complementary methods for a comparative targeted and untargeted characterization of the differences in metabolite composition under two different nutrient conditions in early stationary phase of *E. oleoabundans*. These methods allow simultaneous analysis of lipids along with other hydrophobic metabolites such as chlorophylls, certain vitamins, and terpenes from a microalgal extract. Our results provide information at the molecular level and suggest complex regulation of lipids and other lipid-like metabolites under N⁻ and N⁺ conditions. While the origins and the crosstalk between the different biosynthetic pathways that regulate these changes are currently unknown, we expect that comprehensive metabolomics approaches like ours, in combination with genomics and transcriptomics analyses, will pave the way to the generation of new ideas for expanding the bioproduct portfolio of both model and non-model green microalgae.

Author Contributions

The experiments were designed by G.E.A.-G., B.Z.H, B.A.P., and D.S.A. The experiments were conducted by E.K.M., D.M.B., M.G., and E.C. V.dS., D.M.B, and E.K.M carried out the MS/MS analyses. The manuscript was written by E.K.M., B.Z.H, and G.E.A.-G. The study was directed by G.E.A.-G.

Acknowledgements

This research was supported by the National Science Foundation (Award No: CBET-1438172). B.Z.H is a Newton Advanced Fellow and supported by Royal Society of UK (Award No: NA140481). B.Z.H is also supported by The Scientific and Technological Research Council of

503 Turkey (Award No: 115C019) and Boğaziçi University BAP Research Fund (Award No:
504 14Y00SUP10).

505

506

507

508

509

510

511

512 REFERENCES

- 513 [1] National Algal Biofuels Technology Roadmap. U.S. Department of Energy, Office of Energy
514 Efficiency and Renewable Energy, Biomass Program, U.S. DOE 2010, Energy Efficiency &
515 Renewable Energy.
- 516 [2] S.A. Scott, M.P. Davey, J.S. Dennis, I. Horst, C.J. Howe, D.J. Lea-Smith, A.G. Smith,
517 Biodiesel from algae: challenges and prospects, *Curr Opin Biotechnol*, 21 (2010) 277-286.
- 518 [3] F. Ahmed, Y. Li, P.M. Schenk, Algal biorefinery: sustainable production of biofuels and
519 aquaculture feed?, *The Science of Algal Fuels*, Springer 2012, pp. 21-41.
- 520 [4] E. Stephens, I.L. Ross, Z. King, J.H. Mussgnug, O. Kruse, C. Posten, M.A. Borowitzka, B.
521 Hankamer, An economic and technical evaluation of microalgal biofuels, *Nat Biotechnol*, 28
522 (2010) 126-128.
- 523 [5] T. Dong, E.P. Knoshaug, R. Davis, L.M.L. Laurens, S. Van Wychen, P.T. Pienkos, N. Nagle,
524 Combined algal processing: A novel integrated biorefinery process to produce algal biofuels and
525 bioproducts, *Algal Res*, (2016).
- 526 [6] R. Cameron Coates, E. Trentacoste, W.H. Gerwick, Bioactive and novel chemicals from
527 microalgae, *Handbook of Microalgal Culture: Applied Phycology and Biotechnology*, Second
528 Edition 2013, pp. 504-531.
- 529 [7] M. Francavilla, P. Trotta, R. Luque, Phytosterols from *Dunaliella tertiolecta* and *Dunaliella*
530 *salina*: a potentially novel industrial application, *Bioresource Technol*, 101 (2010) 4144-4150.
- 531 [8] S. Takaichi, Carotenoids in algae: distributions, biosyntheses and functions, *Mar drugs*, 9
532 (2011) 1101-1118.
- 533 [9] M. Plaza, A. Cifuentes, E. Ibáñez, In the search of new functional food ingredients from
534 algae, *Trends Food Sci Tech*, 19 (2008) 31-39.
- 535 [10] J.K. Volkman, A review of sterol markers for marine and terrigenous organic matter, *Org*
536 *Geochem*, 9 (1986) 83-99.
- 537 [11] C.L. Rock, R.A. Jacob, P.E. Bowen, Update on the biological characteristics of the
538 antioxidant micronutrients: vitamin C, vitamin E, and the carotenoids, *J Am Diet Assoc*, 96
539 (1996) 693-702.
- 540 [12] J. Sheehan, T. Dunahay, J. Benemann, P. Roessler, A look back at the US Department of
541 Energy's aquatic species program: biodiesel from algae, *National Renewable Energy Laboratory*
542 *Golden* 1998.
- 543 [13] L. Tirichine, C. Bowler, Decoding algal genomes: tracing back the history of photosynthetic
544 life on Earth, *Plant J*, 66 (2011) 45-57.
- 545 [14] A.J. Klok, D.E. Martens, R.H. Wijffels, P.P. Lamers, Simultaneous growth and neutral lipid
546 accumulation in microalgae, *Bioresource Technol*, 134 (2013) 233-243.
- 547 [15] C.A. Popovich, C. Damiani, D. Constenla, A.M. Martínez, H. Freije, M. Giovanardi, S.
548 Pancaldi, P.I. Leonardi, *Neochloris oleoabundans* grown in enriched natural seawater for
549 biodiesel feedstock: evaluation of its growth and biochemical composition, *Bioresource Technol*,
550 114 (2012) 287-293.
- 551 [16] A. Santos, M. Janssen, P. Lamers, W. Evers, R. Wijffels, Growth of oil accumulating
552 microalga *Neochloris oleoabundans* under alkaline-saline conditions, *Bioresource Technol*, 104
553 (2012) 593-599.
- 554 [17] M.J. Griffiths, S.T. Harrison, Lipid productivity as a key characteristic for choosing algal
555 species for biodiesel production, *J Appl Phycol*, 21 (2009) 493-507.

556 [18] J. Pruvost, G. Van Vooren, G. Cogne, J. Legrand, Investigation of biomass and lipids
557 production with *Neochloris oleoabundans* in photobioreactor, *Bioresource Technol*, 100 (2009)
558 5988-5995.

559 [19] Y. Li, M. Horsman, B. Wang, N. Wu, C.Q. Lan, Effects of nitrogen sources on cell growth
560 and lipid accumulation of green alga *Neochloris oleoabundans*, *Appl Microbiol Biotechnol*, 81
561 (2008) 629-636.

562 [20] C. Baldisserottoa, C. Popovichb, M. Giovanardia, A. Sabiaa, L. Ferronia, D. Constenlad, P.
563 Leonardib, S. Pancaldi, Photosynthetic aspects and lipid profiles in the mixotrophic alga
564 *Neochloris oleoabundans* as useful parameters for biodiesel production, *Algal Res*, 16 (2016)
565 255-265.

566 [21] C. Brunetti, R.M. George, M. Tattini, K. Field, M.P. Davey, Metabolomics in plant
567 environmental physiology, *J Exp Bot*, 64 (2013) 4011-4020.

568 [22] X.H. Yu, X.F. Niu, X.Q. Zhang, G.S. Pei, J. Liu, L. Chen, W.W. Zhang, Identification and
569 mechanism analysis of chemical modulators enhancing astaxanthin accumulation in
570 *Haematococcus pluvialis*, *Algal Res*, 11 (2015) 284-293.

571 [23] Y.X. Su, J.X. Wang, M.L. Shi, X.F. Niu, X.H. Yu, L.J. Gao, X.Q. Zhang, L. Chen, W.W.
572 Zhang, Metabolomic and network analysis of astaxanthin-producing *Haematococcus pluvialis*
573 under various stress conditions, *Bioresource Technol*, 170 (2014) 522-529.

574 [24] G. Dragone, B.D. Fernandes, A.P. Abreu, A.A. Vicente, J.A. Teixeira, Nutrient limitation as
575 a strategy for increasing starch accumulation in microalgae, *Appl Energ*, 88 (2011) 3331-3335.

576 [25] I.A. Guschina, J.L. Harwood, Lipids and lipid metabolism in eukaryotic algae, *Prog Lip*
577 *Res*, 45 (2006) 160-186.

578 [26] L. Rodolfi, G. Chini Zittelli, N. Bassi, G. Padovani, N. Biondi, G. Bonini, M.R. Tredici,
579 Microalgae for oil: Strain selection, induction of lipid synthesis and outdoor mass cultivation in a
580 low-cost photobioreactor, *Biotechnol Bioeng*, 102 (2009) 100-112.

581 [27] A. Shrivastava, S.K. Mishraa, W.I. Suha, W. Farooq, M. Moonb, T. Kimb, K. Kumara, G.
582 Choib, M.S. Parka, J. Yang, Characterization of newly isolated oleaginous microalga
583 *Monoraphidium* sp. for lipid production under different conditions, *Algal Res*, 12 (2015) 289-
584 294.

585 [28] S. Karampudi, K. Chowdhury, Effect of Media on Algae Growth for Bio-Fuel Production,
586 *Notulae Scientia Biologicae*, 3 (2011) 33-41.

587 [29] Z.J. Zhu, A.W. Schultz, J. Wang, C.H. Johnson, S.M. Yannone, G.J. Patti, G. Siuzdak,
588 Liquid chromatography quadrupole time-of-flight mass spectrometry characterization of
589 metabolites guided by the METLIN database, *Nat Protoc*, 8 (2013) 451-460.

590 [30] E. Fahy, M. Sud, D. Cotter, S. Subramaniam, LIPID MAPS online tools for lipid research,
591 *Nucleic Acids Res*, 35 (2007) W606-612.

592 [31] S.J. Khan, H.S. Weinberg, E.C. Bedford, Aqueous-phase aminolysis: approach for the
593 analysis of epoxides in water, *Anal Chem*, 78 (2006) 2608-2616.

594 [32] H. Rismani-Yazdi, B.Z. Haznedaroglu, C. Hsin, J. Peccia, Transcriptomic analysis of the
595 oleaginous microalga *Neochloris oleoabundans* reveals metabolic insights into triacylglyceride
596 accumulation, *Biotechnol Biofuels*, 5 (2012) 74.

597 [33] Y.M. Gong, X.J. Guo, X. Wan, Z. Liang, M.L. Jiang, Triacylglycerol accumulation and
598 change in fatty acid content of four marine oleaginous microalgae under nutrient limitation and
599 at different culture ages, *J Basic Microb*, 53 (2013) 29-36.

600 [34] M.J. Griffiths, R.P. van Hille, S.T.L. Harrison, The effect of degree and timing of nitrogen
601 limitation on lipid productivity in *Chlorella vulgaris*, *Appl Microbiol Biotechnol*, 98 (2014)
602 6147-6159.

603 [35] G. Mujtaba, W. Choi, C.G. Lee, K. Lee, Lipid production by *Chlorella vulgaris* after a shift
604 from nutrient-rich to nitrogen starvation conditions, *Bioresource Technol*, 123 (2012) 279-283.

605 [36] S.N. Zhu, W. Huang, J. Xu, Z.M. Wang, J.L. Xu, Z.H. Yuan, Metabolic changes of starch
606 and lipid triggered by nitrogen starvation in the microalga *Chlorella zofingiensis*, *Bioresource*
607 *Technol*, 152 (2014) 292-298.

608 [37] M.P. Davey, I. Horst, G.H. Duong, E.V. Tomsett, A.C. Litvinenko, C.J. Howe, A.G. Smith,
609 Triacylglyceride production and autophagous responses in *Chlamydomonas reinhardtii* depend
610 on resource allocation and carbon source, *Eukaryotic cell*, 13 (2014) 392-400.

611 [38] Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate - a Practical and Powerful
612 Approach to Multiple Testing, *J Roy Stat Soc B Met*, 57 (1995) 289-300.

613 [39] M. Siaut, S. Cuine, C. Cagnon, B. Fessler, M. Nguyen, P. Carrier, A. Beyly, F. Beisson, C.
614 Triantaphylides, Y.H. Li-Beisson, G. Peltier, Oil accumulation in the model green alga
615 *Chlamydomonas reinhardtii*: characterization, variability between common laboratory strains and
616 relationship with starch reserves, *Bmc Biotechnol*, 11 (2011).

617 [40] H. Wada, N. Mizusawa, The Role of Phosphatidylglycerol in Photosynthesis, *Lipids in*
618 *Photosynthesis: Essential and Regulatory Functions* 2009, pp. 243-263.

619 [41] G. Dubertret, C. Gerard-Hirne, A. Tremolieres, Importance of trans-Delta(3)-hexadecenoic
620 acid containing phosphatidylglycerol in the formation of the trimeric light-harvesting complex in
621 *Chlamydomonas*, *Plant Physiol Bioch*, 40 (2002) 829-836.

622 [42] K. Sugimoto, N. Sato, M. Tsuzuki, Utilization of a chloroplast membrane sulfolipid as a
623 major internal sulfur source for protein synthesis in the early phase of sulfur starvation in
624 *Chlamydomonas reinhardtii*, *Febs Lett*, 581 (2007) 4519-4522.

625 [43] E. Gunerken, E. D'Hondt, M.H.M. Eppink, L. Garcia-Gonzalez, K. Elst, R.H. Wijffels, Cell
626 disruption for microalgae biorefineries, *Biotechnol Adv*, 33 (2015) 243-260.

627 [44] P. Cheali, C. Gargalo, K.V. Gernaey, G. Sin, A Framework for Sustainable Design of Algal
628 Biorefineries: Economic Aspects and Life Cycle Analysis, *Algal Biorefineries*, Springer
629 International Publishing 2015, pp. 511-535.

630 [45] J. Trivedi, M. Aila, D.P. Bangwal, S. Kaul, M.O. Garg, Algae based biorefinery—How to
631 make sense?, *Renew Sust Energ Rev*, 47 (2015) 295-307.

632 [46] W.R. Riekhof, M.E. Ruckle, T.A. Lydic, B.B. Sears, C. Benning, The sulfolipids 2'-O-acyl-
633 sulfoquinovosyldiacylglycerol and sulfoquinovosyldiacylglycerol are absent from a
634 *Chlamydomonas reinhardtii* mutant deleted in SQD1, *Plant Physiol*, 133 (2003) 864-874.

635 [47] D. Yang, D. Song, T. Kind, Y. Ma, J. Hoefkens, O. Fiehn, Lipidomic Analysis of
636 *Chlamydomonas reinhardtii* under Nitrogen and Sulfur Deprivation, *PloS one*, 10 (2015)
637 e0137948.

638 [48] D.Y. Lee, J.J. Park, D.K. Barupal, O. Fiehn, System Response of Metabolic Networks in
639 *Chlamydomonas reinhardtii* to Total Available Ammonium, *Mol Cell Proteomics*, 11 (2012)
640 973-988.

641 [49] J. Li, D.X. Han, D.M. Wang, K. Ning, J. Jia, L. Wei, X.Y. Jing, S. Huang, J. Chen, Y.T. Li,
642 Q. Hu, J. Xu, Choreography of Transcriptomes and Lipidomes of *Nannochloropsis* Reveals the
643 Mechanisms of Oil Synthesis in Microalgae, *Plant Cell*, 26 (2014) 1645-1665.

644 [50] G.J.O. Martin, D.R.A. Hill, I.L.D. Olmstead, A. Bergamin, M.J. Shears, D.A. Dias, S.E.
645 Kentish, P.J. Scales, C.Y. Botte, D.L. Callahan, Lipid Profile Remodeling in Response to

646 Nitrogen Deprivation in the Microalgae *Chlorella* sp (Trebouxiophyceae) and *Nannochloropsis*
647 sp (Eustigmatophyceae), PloS one, 9 (2014).

648