Comparative Shotgun Proteomic Analysis of Wastewater-Cultured Microalgae: Nitrogen Sensing and Carbon Fixation for Growth and Nutrient Removal in *Chlamydomonas reinhardtii*

Anil K. Patel[*,†], Eric L. Huang[†], Etienne Low-Décarie[‡] and Mark G. Lefsrud[†]

[†]Department of Bioresource Engineering, McGill University, Ste. Anne de Bellevue, Quebec H9X 3V9, Canada

[‡]School of Biological Sciences, University of Essex, Colchester CO4 3SQ, United Kingdom

Journal of Proteome Research **(2015), DOI: http://dx.doi.org/10.1021/pr501316h**

SUPPORTING INFORMATION

ABSTRACT:

Chlamydomonas reinhardtii was batch cultured for 12 days under continuous illumination to study nitrogen uptake and metabolic responses to wastewater processing. Our approach compared two conditions: (1) artificial wastewater containing nitrate and ammonia and (2) nutrient-sufficient control containing nitrate as sole form of nitrogen. Treatments did not differ in final biomass; however, comparison of group proteomes revealed significant differences. Label-free shotgun proteomic analysis identified 2358 proteins, of which 92 were significantly differentially abundant. We report additional differentially abundant proteins of interest not discussed previously, including biosynthesis related, stress related, poorly annotated, and unknown proteins. Wastewater cells showed higher relative levels of biosynthesis proteins

related to B complex vitamins and cobalamin, and stress response proteins related to iron deficiency. Control cells showed higher levels of histone superfamily proteins, sugar metabolism and nitrogen regulation proteins, and nutrient limitation and oxidative stress response proteins. Control had a higher abundance of poorly annotated and unknown proteins, but not proteins associated with neutral lipid biosynthesis. Enriched pathways in wastewater, included RNA degradation, thiamine metabolism and one carbon pool by folate, and those in nitrate only control, notably, calcium signaling, systemic lupus erythematosus, protein processing in endoplasmic reticulum and Arf/Sar family. Proteome analysis found that N-limitation did not adversely affect growth or biomass yields in control cells at the flask bench scale, due in part to the role of stress response proteins. Over-enrichment of ammonia-N did not adversely affect cellular physiology or yields in wastewater cells. Enriched pathways related to the stress response in control, most notably, glycolysis/gluconeogenesis, glutathione metabolism, oxidative phosphorylation, peroxisome and proximal tubule bicarbonate reclamation represent potential targets for genetic improvement requiring targeted elucidation.

1. INTRODUCTION

As reported in the article,[1] our experimental design involved propagating *C. reinhardtii* cells in batch cultures for 12 days under continuous illumination to compare two treatment conditions: (1) artificial wastewater containing nitrate and ammonia, and (2) nutrient sufficient control containing nitrate as sole form of nitrogen. Despite the increased availability of nitrogen in the form of ammonia, a preferred source of nitrogen for green algae and higher plants,[2,3] wastewater cells did not grow faster, accumulate higher biomass, or uptake nutrients faster. Nonetheless, *Chlamydomonas* showed a remarkable ability to adapt to local conditions in each treatment.

Surprisingly, even after 12 days the wastewater group was still engaged in photosynthetic carbon fixation, cell division, and biosynthesis of amino acids, fatty acids, secondary metabolites and chlorophyll. Control cells showed continued photosynthesis, synthesis and use of starch, and recycling of amino acids, which suggested adaptations to the onset of N-limitation. The typical cellular response to N-deprivation by C and N responsive pathways is down-regulation of carbon assimilation and chlorophyll biosynthesis, up-regulation of lipid biosynthesis, and higher abundances of oxidative stress response enzymes.[4,5] Different from other differential gene expression studies on nitrogen stress response in *C. reinhardtii*,[6,7] and proteomic and/or metabolomic time-course studies manipulating only ammonia[5,8,9] or nitrate,[10] higher levels of enzymes and proteins related to lipid biosynthesis was not found in the control condition. Furthermore, a pattern search of *A. thaliana* GO annotation terms related to nitrogen starvation found no proteins identified in this experiment. Here we report additional differentially abundant proteins of interest not discussed in the article, including biosynthesis related, stress related, poorly annotated, and unknown proteins. Proteins lacking KEGG pathway association are evaluated by use of functional annotation tools such as Pfam, Panther, GO and KOG ontology databases, JGI Phytozome, and literature related to *Chlamydomonas* and *Arabidopsis*.

2. MATERIALS AND METHODS

Details can be found in the Materials and Methods section of the article for experimental design, organism, and culture conditions (Section 2.1), growth and nutrient removal (Section 2.2), protein extraction and digestion (Section 2.3), 2D LC–MS/MS (Section 2.4), and database searching (Section 2.5).

2.1. Selection of Differentially Abundant Proteins, Functional Enrichment and Pathway Analysis

A minimum of 4 spectra per protein was required for spectral counting and bioinformatics analysis.[11] As described in the article (see additional details in Materials and Methods, Sections 2.6 and 2.7, and Results and Discussion, Section 3.3),[1] we used the QSpec software to identify statistically significant differentially abundant proteins by culture group using spectral count normalization and estimates of false discovery rates (FDR).[12] A threshold of Z-statistic $\geq |3.69|$ and corresponding to a directional FDRup of 5% (0.05) was used, unless the less stringent Z-statistic $\geq |3.14|$ threshold, corresponding to a directional FDRup of 10% (0.10) was specified. Functional attributes of the data set were analyzed using Phytozome annotation (*Chlamydomonas* v9.1 and *Arabidopsis* v9.1). Pattern searches of GO annotation data from *A. thaliana* was used to produce GOSLIM files targeting specific processes and functions. Functional enrichment of biological pathways was performed using the web-based annotation integration tool Algal Functional Annotation Tool (AFAT).[13] Phytozome v5 alphanumerical transcript IDs of significantly differentially abundant proteins identified by QSpec were queried against the entire *C. reinhardtii* genome. Manual curation of enrichment results was performed for all proteins discussed, including those that received multiple hits in AFAT-assigned KEGG pathways. Significantly differentially abundant proteins were represented in both tabular and graphical format (Table 1 and Figure 3), where metabolic mapping was performed using the KEGG pathway mapping software iPath2.0. The open source R statistical package was used for data analysis and generation of plots not performed with the aforementioned software.

3. RESULTS AND DISCUSSION

QSpec software identified 92 statistically significant differentially abundant proteins with a directional FDRup of 5% (0.05), of which 42 and 50 showed higher levels of relative abundance in wastewater and control, respectively (Supporting Information Table S1). Of these 92 differential proteins, 58 (32 in wastewater and 26 in control) were assigned KEGG group association (*ie*., KEGG KO) by the AFAT tool (Table 1). Of the 34 differential proteins lacking KEGG association, 8 in wastewater and 20 in control had some functional annotation or description in the Phytozome databases, AFAT tool, and the literature. The remaining 6 differential proteins, 2 and 4 in wastewater and control, respectively, were unknown proteins lacking any annotation including description or Pfam[14] family membership. Despite this limitation in functional annotation, the AFAT tool matched the 58 responsive differential proteins to 61 KEGG pathway terms (ie., KEGG ko) in both conditions, and manual curation of functionally enriched KEGG pathway terms (ie., KEGG ko) was performed (described in detail in Materials and Methods). Manual curation produced 32 proteins enriched in 17 KEGG pathways in the wastewater group, and 26 proteins enriched in 16 pathways in control (Table 1). In order to visualize functional differences between treatment groups, differentially abundant proteins from Table 1 were mapped to KEGG-derived metabolic pathways using iPath2.0.[15,16] Only those differential proteins with an assigned function from AFAT (KEGG KO and/or Enzyme EC number) could be inputted, omitting several differentially abundant proteins (Supporting Information Table S1), and not all protein functions were matched by the current iPath2.0 metabolic map. Despite the limitations in functional annotation, responsive proteins mapped to 97 metabolic map elements (60 in wastewater and 37 in control) to elucidate a pathway-centric overview of specific functional reactions of potential interest for wastewater processing (Figure 3).[15,16]

In wastewater, RNA degradation, Thiamine metabolism pathway, and One carbon pool by folate make up the enriched pathways by AFAT significance score and manual curation not discussed in the article (Table 1). In the control group Calcium signaling pathway, Systemic lupus erythematosus, Arf/Sar family, and Protein processing in endoplasmic reticulum make up the enriched pathways by AFAT significance score and manual curation. Also, Huntington's disease pathway was enriched in wastewater cells, and suggested iron limitation. In the control group Glycolysis/Gluconeogenesis pathway was enriched due to the onset of N-limitation, and Glutathione metabolism, Oxidative phosphorylation, Proximal tubule bicarbonate reclamation and Peroxisome pathway due to oxidative stress. Interestingly, KEGG pathways not enriched in control, included Fatty acid biosynthesis and Biosynthesis of unsaturated fatty acids.

Significant supplemental proteins with KEGG association (Table 1 and Figure 3) and differential proteins lacking KEGG association (Supporting Information Table S1), including biosynthesis related proteins are discussed in Sections 3.1.1 – 3.1.3. Sections 3.2.1 – 3.2.3 discusses stress related proteins found to be significant in both conditions. Poorly annotated and unknown proteins are discussed in Sections 3.3.1 – 3.3.2. At first mention the putative functional name/description of a significant differential protein is given, followed by the Phytozome gene accession name in brackets and gene alias, if any, in italics. Subsequent mention uses putative enzyme abbreviations, or gene aliases, if any. Upper case is used for nuclear encoded genes and proteins, lower case for organelle encoded proteins, where all genes in Phytozome are nuclear encoded.[17]

3.1. Differential Analysis and Functional Enrichment

3.1.1 Additional Biosynthesis Proteins Enriched in Wastewater by AFAT score.

Proteins related to the B complex vitamins cobalamin (B_{12}), thiamine (B_1), and folate (B_9) were found to be significantly more abundant in wastewater cells. Two cobalamin synthesis proteins (g15619.t1 and g15199.t1) (Table 1 and Figure 3), described as plastid transcriptionally active 17 in *A. thaliana*, were enriched in the RNA degradation KEGG pathway. In *Chlamydomonas* cobalamin is a cofactor for enzymes catalyzing rearrangement-reduction reactions or methyl transfer reactions in the cytosol.[18] The thiazole biosynthetic enzyme (Cre04.g214150.t1.3; *THI4*), with KOG function of thiamine biosynthesis and DNA damage tolerance was enriched in the Thiamine metabolism pathway, but lacked a score in the AFAT tool. It is involved in the synthesis of the thiazole moiety of thiamine pyrophosphate (B_1), which plays a fundamental role in energy metabolism by acting as an essential cofactor for enzymes such as transketolase and α-ketoglutarate dehydrogenase. Methylenetetrahydrofolate reductase (NADPH) (MTHFR) (Cre10.g433600.t1.2) had higher abundance in the One carbon pool by folate KEGG pathway.

3.1.2 Additional Proteins Enriched in Control by AFAT score. In the control group additional proteins enriched in KEGG pathways included the mitochondrial ADP/ATP carrier protein (Cre09.g386650.t1.2; *ANT1*) with higher abundance in the Calcium signaling pathway (Table 1 and Figure 3); and the histone super family proteins Histone H2B (Cre13.g590750.t1.2; *HTB37*) and Histone 3 (Cre06.g265000.t1.2; *HTR13*), both members of the Pfam Core histone H2A/H2B/H3/H4 protein family (PF00125). The H3 and H4 histone DNA binding superfamily proteins are involved in transcriptional activation through hyperacetylation, and chromatin compaction and gene repression through hypoacetylation, as shown in gene expression in *Arabidopsis*.[19] Histone H2A (Cre13.g567700.t1.2; *HAV*) is a core histone for nucleosome formation. All three histone proteins were enriched in the Systemic lupus erythematosus KEGG

pathway. Higher relative abundances of these proteins was not expected. Histone proteins play a critical role in preserving chromatin organization, but are not known to be greatly effected by nutrient deprivation.[7] Proteomic analysis of ammonia-N deprived cells showed levels of histones and other core proteins remained mostly unchanged over time.[8]

The ADP-ribosylation factor family protein (g2791.t1, K0 7977) was enriched in the Arf/Sar family, other KEGG pathway, lacked a score in AFAT, and showed GO biological process of small GTPase mediated signal transduction and molecular function of GTP binding. Protein disulfide-isomerase (Cre02.g088200.t1.2; *RB60*) also lacked an AFAT score and was enriched in the Protein processing in endoplasmic reticulum pathway. This enzyme has Pfam Thioredoxin family membership (PF00085) and GO biological process of cell redox homeostasis. In the chloroplasts of *C. reinhardtii*, PDI RB60 is a redox sensor component of an mRNA-binding protein complex associated with photo-regulation of translation in the dark, and is also implicated in formation of regulatory disulfide bonds in chloroplasts.[20]

3.1.3 Additional Proteins in Control Lacking AFAT Annotation. In the control group 3 additional proteins showed higher relative abundances, but lacked KEGG association (Supporting Information Table S1). First, triose-phosphate transporter (Cre01.g045550.t1.2; *TPT3*) is described as glucose-6-phosphate/phosphate translocator-related in *A. thaliana*, belongs to the Pfam Triose-phosphate transporter family (PF03151) and is associated with sugar metabolism. Located between the inner and outer membrane of the chloroplast, GPTs mediate the transport of carbon between plastids and the cytosol in the form of glucose 6-phosphate (Glc-6-P) for glycolysis, and the oxidation of glucose in the pentose phosphate pathway.[21] This transporter had KOG function of phosphoenolpyruvate/phosphate antiporter, in the export of triose phosphate in exchange for inorganic phosphate used for ATP regeneration in the light

reactions. Second, the *A. thaliana* described carbohydrate-binding-like fold protein (Cre12.g492750.t1.3) has Pfam Starch binding domain family membership (PF00686), GO biological process of carbohydrate metabolic process, and GO molecular function of starch binding. Third, the *A. thaliana* described plastid transcriptionally active 16 (Cre02.g081250.t1.2) showed KOG predicted dehydrogenase function, nitrogen metabolic regulation (Nmr) in Panther (14194), and Nmr-A-like family membership in Pfam (PF05368). It is one of eighteen components, called plastid transcriptionally active chromosome proteins (pTACs). pTACs are associated with plastid gene expression, and may play roles in post-transcriptional processes, including RNA processing and/or mRNA stability.[22] Interestingly, pTAC 17 is associated with cobalamin synthesis and showed higher abundance in wastewater cells (Supporting Information, Section 3.1.1).

3.2. Stress Response Proteins

Similar growth curves suggested cells in neither culture group were in a complete stress response (Figure 1a). The absence of aggressive perturbation by nutrient deprivation is also confirmed by the lack of a dramatic shift at the proteome level between treatment groups, and was not unexpected given the subtle manipulation of macronutrients. Our experimental design did not use a control group held in exponential phase with continuous harvesting of cells and nutrient replenishment. Instead, addition of ammonia-N in the wastewater condition ensured over-enrichment of nitrogen, and cells in both groups were allowed to proceed to steady state, similar to high-density wastewater batch cultures. As discussed in the article, a fundamental question presented itself, namely, whether the biological variability detected and reported here, was caused by differences in nitrogen source and loading associated with wastewater processing, and

not simply indirect compensatory effects i.e., the control group cells employing physiological mechanisms to reverse the onset of nutrient perturbation. A pattern search of *A. thaliana* GO annotation terms related to nitrogen starvation found no proteins identified in this experiment. Furthermore, only 1 differentially expressed protein related to fatty acid biosynthesis or TAG accumulation was identified in the control group (Section 3.2.2 below), and none with KEGG pathway enrichment. Overall, no clear pattern is discernable for proteins related to a stress response, although evidence of oxidative stress was detected in control cells, and iron deficiency in wastewater cells.

3.2.1 Stress Related Proteins in Wastewater. In the wastewater cultures, the superoxide dismutase [Fe] (Cre10.g436050.t1.2; *FSD1*) was enriched in the Huntington's disease KEGG pathway (Table 1 and Figure 3). It belongs to the Iron/manganese superoxide dismutase families in Pfam (PF00081, PF02777), has GO biological process of reactions and pathways involving the superoxide anion O_2- (superoxide free radical) and molecular function of catalyzing 2 superoxide + 2 H+ = O_2 + hydrogen peroxide. Higher levels of this enzyme increases the capacity of *Chlamydomonas* to detoxify superoxide during Fe limitation stress.[23] Also enriched in this KEGG pathway was the Porin/voltage-dependent anion-selective channel protein (Cre05.g241950.t1.2; *ASC2*) with GO cellular component of mitochondrial outer membrane and biological process of regulation of anion transport.

3.2.2 Control Cells Showed Evidence of Nutrient Limitation. In control cells glyceraldehyde 3-phosphate dehydrogenase (GAPDH) [EC 1.2.1.12] (Cre12.g485150.t1.2; GAP1), discussed in the article (Section 3.3.4), was enriched in the Glycolysis/Gluconeogenesis pathway. This enzyme is associated with a stress response to low nutrient conditions.[24] Purple acid phosphatase 15 (PAP) (Cre11.g476700.t1.2; *MPA9*) showed higher abundance (Supporting

Information Table S1) and has GO molecular function of protein serine/threonine phosphatase activity. PAPs are metalloenzymes that hydrolyse phosphate esters and anhydrides under acidic conditions, and are implicated in acclimation to nutritional phosphate deprivation.[25] In phosphate-starved *Arabidopsis* plants, PAPs have been shown to scavenge phosphate from extracellular phosphate esters by root cells, and function in intracellular vacuolar phosphate recycling.[26] Triacylglycerol lipase and Flagellar Associated Protein (g9673.t1; *FAP12*) showed higher abundance, has Lipase (class 3) family membership in Pfam (PF01764) and is discussed in the article (Section 3.3.4). As mentioned, pathways not enriched in control included Fatty acid biosynthesis and Biosynthesis of unsaturated fatty acids, which demonstrated that the N-deprivation survival response had not been induced in control cells.

3.2.3 Oxidative Stress Proteins Were Enriched in Control. The typical cellular response to N-deprivation by carbon and nitrogen responsive pathways is down-regulation of carbon assimilation and chlorophyll biosynthesis, and up-regulation of lipid biosynthesis and oxidative stress response.[4,5] A number of proteins associated with oxidative stress were enriched in control cells. Glutathione peroxidase (GPx) (g8342.t1) was enriched in the Glutathione metabolism KEGG pathway and is associated with the oxidative stress response in GO biological process. The enzyme has GO molecular function of glutathione peroxidase activity and catalyzes the reaction 2 glutathione + hydrogen peroxide = oxidized glutathione + 2 H_2O. The ubiquinol-cytochrome *c* reductase subunit 7 protein (complex 3) (g5901.t1; *QCR7*) was enriched in the Oxidative phosphorylation pathway. In *A. thaliana*, GO biological process and GO molecular function identify this multi-subunit enzyme as complex III, containing ten polypeptide subunits that include four redox centers: cytochrome *b*/*b*6, cytochrome *c*1 and an 2Fe–2S cluster. It is located in the mitochondrial inner membrane and forms part of the mitochondrial respiratory

chain that catalyzes oxidation of ubiquinol by oxidized cytochrome *c*1. The sodium/potassium-transporting ATPase (Cre06.g263950.t1.3) was enriched in Proximal tubule bicarbonate reclamation pathway. Described in *A. thaliana* as an endoplasmic reticulum-type calcium-transporting ATPase 3, it has KOG function Na+/K+ ATPase, alpha subunit (membrane protein, Ca2+-transporting ATPase). Alcohol dehydrogenase (ADH) (g15030.t1; *ADH7*) was more abundant. It is a KOG alcohol dehydrogenase (class V) enzyme belonging to the aldehyde dehydrogenase (ALDH) gene superfamily of enzymes performing NAD^+- or $NADP^+$- dependent conversion of aldehydes to carboxylic acids. In *Arabidopsis* and *Chlamydomonas* they function as 'aldehyde scavengers' that remove reactive aldehydes during the oxidative degradation of lipid membranes (lipid peroxidation).[27] ADH was enriched in the Glycolysis/Gluconeogenesis pathway, and discussed in the article (Section 3.3.4).

Aconitate hydratase (Cre01.g042750.t1.2; *ACH1*) was enriched in the Biosynthesis of alkaloids derived from terpenoid and polyketide pathway. It is part of the KOG functional group RNA-binding translational regulator IRP (aconitase superfamily). In *A. thaliana* it has multiple GO annotations including biological processes of mitochondrial, isocitrate metabolic process, response to oxidative stress, and GO molecular function of aconitate hydratase activity for the interconversion of citrate and isocitrate. 2-cys peroxiredoxin (2-cys Prx) (Cre02.g114600.t1.2; *PRX2*) was listed as K0 3386 in Phytozome and enriched in the Peroxisome KEGG pathway. The enzyme has KOG (0852) Alkyl hydroperoxide reductase, thiol specific antioxidant and related enzymes membership and GO molecular function of antioxidant activity. It is part of the antioxidant defense system of chloroplasts in *Arabidopsis* and mitigates oxidative damage in photosynthetic membranes by reducing toxic peroxides to their corresponding alcohols.[28] Carbonic anhydrase (CA1) (Cre04.g223100.t1.2; *CAH1*), described as carbonate dehydratase 1

in *C. reinhardtii*, was enriched in the Nitrogen metabolism pathway. This enzyme is localized in the periplasmic space, is low-CO_2 inducible gene regulated and has carbonate dehydratase activity.[29] Finally, 4 pherophroin protein family members showed significantly higher relative abundances but lacked further annotation (Section 3.3.2 below).

3.3. Poorly Annotated and Unknown Proteins

A continuing challenge in plant and algae proteomics is the complexity of the respective eukaryotic genomes, gaps in annotation databases and tools, and not an insignificant number of proteins with unknown or poorly described functions.[17] Association with a protein family would at the very least indicate functional units present to hypothesize a biological role.[15] Of those proteins lacking KEGG association, 8 wastewater and 20 control proteins had some functional annotation or description, and 2 and 4 proteins, respectively, were unknown proteins lacking any annotation including Pfam family membership or Phytozome description.

3.3.1 Poorly Annotated and Unknown Proteins with Higher Levels in Wastewater.

The *A. thaliana* described NAD(P)-binding Rossmann-fold superfamily protein (g2755.t1) lacked KEGG pathway association in AFAT, and showed nitrogen metabolic regulation (Nmr) in Panther, and NmrA-like Pfam membership (PF05368). It had GO chloroplast cellular component and biological process of regulation of nitrogen utilization, and molecular function of repressor activity. The cobalamin synthesis protein (Cre02.g118400.t1.3), described in *A. thaliana* as plastid transcriptionally active 17, showed cobalamin synthesis protein function in KOG, and GO cellular component of plastid chromosome and chloroplast stroma (Section 3.1.1 above). It is one of eighteen components called plastid transcriptionally active chromosome proteins (pTACs) that are associated with plastid gene expression and may play roles in posttranscriptional processes,

including RNA processing and/or mRNA stability.[22] Interestingly, pTAC 16 is associated with nitrogen metabolism regulation and showed significantly higher abundance in control cells, as mentioned (Section 3.1.3 above).

The tetratricopeptide repeat (TPR)-like superfamily protein (Cre09.g404000.t1.3) in *A. thaliana* belonged to the Pfam Pentapeptide repeats (8 copies) family (PF00805). The protein (Cre12.g524600.t1.2) had GO biological process of proteolysis and molecular function of aminopeptidase activity, and Pfam X-Pro dipeptidyl-peptidase (S15 family) membership (PF02129). The KOG 5'-AMP-activated protein kinase, gamma subunit (Cre12.g528000.t1.2) had Pfam CBS domain (PF00571). The protein (Cre07.g340450.t1.3) had Pfam memberships of IPT/TIG domain (PF01833), G8 domain (PF10162), and PA14 domain (PF07691) and GO molecular function of protein binding. Two proteins (Cre10.g452800.t1.2 and Cre06.g307500.t1.1) had the Phytozome description of Low-CO_2 inducible protein. *C. reinhardtii* has been shown to modulate photosynthesis to acclimate to CO_2-limiting stress by the induction of the carbon concentrating mechanism (CCM) including carbonic anhydrases and inorganic carbon transporters.[30]

The remaining two unknown differential proteins (Cre07.g352000.t1.2 and Cre11.g467400.t1.2) lacked any functional annotation whatsoever from Pfam, the KEGG, MetaCyc, Panther, and Reactome biological pathway databases, or GO, Mapman and KOG ontology databases. Using the protein sequence we attempted to elucidate a biological function by using BLAST to associate sequence homology with higher plants and organisms. Searches did not return any insight into identity or function.

3.3.2 Poorly Annotated and Unknown Proteins with Higher Levels in Control. Additional proteins enriched in control contained a description and minimal annotation, or no

annotation whatsoever. The KOG Predicted Ca^{2+}-dependent phospholipid-binding protein (Cre02.g110350.t1.2) showed C2 domain protein family membership (PF00168). The flagella membrane glycoprotein (major form) (g9144.t1; *FMG1-B*) had KOG function of proteins containing Ca^{2+}-binding EGF-like domains. The predicted protein with ankrin repeats (Cre03.g173350.t1.2; *ANK22*) had Ankyrin repeat Pfam membership (PF00023), KOG function of myotrophin and similar proteins and GO molecular function of protein binding. The KOG actin regulatory proteins (gelsolin/villin family) (Cre12.g524400.t1.2) had family membership in Villin headpiece domain (PF02209) and Gelsolin repeat (PF00626). Finally the Pfam YbaB/EbfC DNA-binding family protein (Cre10.g458550.t1.2) (PF02575) also showed higher levels.

The proteins containing only descriptions, included, a mitochondrial F1F0 ATP synthase associated 31.2 kDa protein (Cre13.g581600.t1.2*; ASA4*); the low-CO_2 inducible protein (Cre04.g222800.t1.2*; LCID*); the *A. thaliana* described chloroplast thylakoid membrane cellular component proteins (Cre05.g233950.t1.2; *CGL129*) and (Cre10.g433950.t1.3); and the Flagellar Associated Protein (Cre02.g077750.t1.2*; FAP211*) with Pfam NYN domain membership (PF01936), described in *A. thaliana* as a putative endonuclease or glycosyl hydrolase.

A number of Pfam Pherophorin protein family members (PF12499) showed significantly higher abundance in control. These included, the cell wall protein pherophorin-C8 (Cre17.g717850.t1.3; *PHC8*); the cell wall protein pherophorin-C3 (g6305.t1; *PHC3*); and two proteins lacking further description (g17202.t1 and Cre04.g221450.t1.2). In the multicellular green alga *Volvox carteri*, and closely related unicellular *C. reinhardtii*, the presence of reactive oxygen species and pherophorin family proteins suggested an evolutionary connection between

sex and stress at the gene level.[31] As mentioned, N-starvation is known to induce gametogenesis and sexual reproduction in *C. reinhardtii*.[32,33]

The control group contained 4 unknown proteins with significantly higher levels that lacked any functional annotation whatsoever from Pfam, the KEGG, MetaCyc, Panther, and Reactome biological pathway databases, or GO, Mapman and KOG ontology databases. The gene accession names were Cre17.g707900.t1.2, Cre12.g544450.t1.2, Cre09.g405500.t1.3 and g2947.t1. Using the protein sequence we attempted to elucidate a biological function by using BLAST to associate sequence homology with higher plants and organisms. Searches did not return any insight into identity or function.

4. CONCLUDING REMARKS

Differing from studies on harsh nutrient deprivation and/or single nitrogen source manipulation, our label-free shotgun proteomic analysis of wastewater cultured *C. reinhardtii* contributed novel insights into nitrogen metabolism and carbon fixation by using wastewater containing both nitrate and ammonia. Surprisingly, after 12 days, we found no significant difference in final biomass and OD_{680}. This finding demonstrated that even in increasingly N-limited media *C. reinhardtii* continued to grow and attempt to balance its C/N ratio. Wastewater cells showed higher relative abundances of biosynthesis proteins related to B complex vitamins and cobalamin, and stress response proteins related to iron deficiency. Control cells showed higher levels of histone superfamily proteins, proteins associated with carbohydrate metabolism and nitrogen regulation, and stress proteins in response to nutrient limitation and oxidative stress. Control cells demonstrated a higher relative abundance of poorly annotated and unknown proteins, but not lipid biosynthesis. Up-regulated pathways in wastewater, included, RNA

degradation, Thiamine metabolism and One carbon pool by folate, and those in nitrate only control, notably, Calcium signaling, Systemic lupus erythematosus, Protein processing in endoplasmic reticulum and Arf/Sar family. Up-regulated pathways related to the stress response in the control condition, included, Glycolysis/Gluconeogenesis, Glutathione metabolism, Oxidative phosphorylation, Peroxisome and Proximal tubule bicarbonate reclamation.

Proteome analysis found that N-limitation did not adversely affect growth or biomass yields in control cells at the flask bench scale cultured in nitrate-N only medium after 12 days. This finding is due in part to the role of stress response proteins, and further demonstrated the resilience of *C. reinhardtii* to abiotic stressors.[34] Over-enrichment of ammonia-N did not adversely affect cellular physiology or yields in wastewater cells. Both these findings are intriguing and support optimism for improvements in bioprocess optimization and future yields in outdoor ponds using different sources of wastewater of varying quality and macronutrient concentration. Up-regulated pathways related to the stress response in the control condition merit further study. Also interesting are the up-regulated biosynthesis pathways identified in wastewater cells such as One carbon pool by folate that lead to secondary compounds. Differential abundance analysis and functional enrichment identified a number of enzymes and potential gene targets for genetic improvement requiring further elucidation using discovery and targeted proteomic techniques. We will continue the identification of candidate pathways at the bench and pilot scale for development of robust strains suitable for scalable outdoor treatment of real wastewaters.

REFERENCES

(1) Patel, A. K.; Huang, E. L.; Low-Décarie, E.; Lefsrud, M. G. Comparative shotgun proteomic analysis of wastewater-cultured microalgae: nitrogen sensing and carbon fixation for growth and nutrient removal in *Chlamydomonas reinhardtii*. *J. Proteome Res.* **2015**, doi: 10.1021/pr501316h.

(2) Hase, T.; Schürmann, P.; Knaff, D. B. The interaction of ferredoxin with ferredoxin-dependent enzymes. In *Photosystem 1*; Golbeck, J., Ed.; Springer: Dordrecht, The Netherlands, 2006; pp. 477–498.

(3) Hirasawa, M.; Tripathy, J. N.; Sommer, F.; Somasundaram, R.; Chung, J.-S.; Nestander, M.; Kruthiventi, M.; Zabet-Moghaddam, M.; Johnson, M. K.; Merchant, S. S.; Allen, J. P.; Knaff, D. B. Enzymatic properties of the ferredoxin-dependent nitrite reductase from *Chlamydomonas reinhardtii*. Evidence for hydroxylamine as a late intermediate in ammonia production. *Photosynth. Res.* **2010**, *103*, 67–77.

(4) Gutiérrez, R. A.; Lejay, L. V; Dean, A.; Chiaromonte, F.; Shasha, D. E.; Coruzzi, G. M. Qualitative network models and genome-wide expression data define carbon/nitrogen-responsive molecular machines in *Arabidopsis*. *Genome Biol.* **2007**, 8, R7.

(5) Lee, D. Y.; Park, J.-J.; Barupal, D. K.; Fiehn, O. System response of metabolic networks in *Chlamydomonas reinhardtii* to total available ammonium. *Mol. Cell. Proteomics* **2012**, *11*, 973–988.

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

(7) Msanne, J.; Xu, D.; Konda, A. R.; Casas-Mollano, J. A.; Awada, T.; Cahoon, E. B.; Cerutti, H. Metabolic and gene expression changes triggered by nitrogen deprivation in the photoautotrophically grown microalgae *Chlamydomonas reinhardtii* and *Coccomyxa* sp. C-169. *Phytochemistry* **2012**, *75*, 50–59.

(8) Valledor, L.; Furuhashi, T.; Recuenco-Muñoz, L.; Wienkoop, S.; Weckwerth, W. System-level network analysis of nitrogen starvation and recovery in *Chlamydomonas reinhardtii* reveals potential new targets for increased lipid accumulation. *Biotechnol. Biofuels* **2014**, *7*, 171.

(9) Wase, N.; Black, P. N.; Stanley, B. A.; DiRusso, C. C. Integrated quantitative analysis of nitrogen stress response in *Chlamydomonas reinhardtii* using metabolite and protein profiling. *J. Proteome Res.* **2014**, *13*, 1373–1396.

(10) Guarnieri, M. T.; Nag, A.; Yang, S.; Pienkos, P. T. Proteomic analysis of *Chlorella vulgaris*: potential targets for enhanced lipid accumulation. *J. Proteomics* **2013**, *93*, 245–253.

(11) Old, W. M.; Meyer-Arendt, K.; Aveline-Wolf, L.; Pierce, K. G.; Mendoza, A.; Sevinsky, J. R.; Resing, K. A.; Ahn, N. G. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol. Cell. Proteomics* **2005**, *4*, 1487–1502.

(12) Choi, H.; Fermin, D.; Nesvizhskii, A. I. Significance analysis of spectral count data in label-free shotgun proteomics. *Mol. Cell. Proteomics* **2008**, *7*, 2373–2385.

(13) Lopez, D.; Casero, D.; Cokus, S. J.; Merchant, S. S.; Pellegrini, M. Algal Functional Annotation Tool: a web-based analysis suite to functionally interpret large gene lists using integrated annotation and expression data. *BMC Bioinf.* **2011**, *12*, 282.

(14) Finn, R. D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R. Y.; Eddy, S. R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; Sonnhammer, E. L. L.; Tate, J.; Punta, M. Pfam: the protein families database. *Nucleic Acids Res.* **2014**, *42*, 222–230.

(15) Abraham, P.; Giannone, R. J.; Adams, R. M.; Kalluri, U.; Tuskan, G. A.; Hettich, R. L. Putting the pieces together: high-performance LC-MS/MS provides network-, pathway-, and protein-level perspectives in populus. *Mol. Cell. Proteomics* **2013**, *12*, 106–119.

(16) Yamada, T.; Letunic, I.; Okuda, S.; Kanehisa, M.; Bork, P. iPath2.0: interactive pathway explorer. *Nucleic Acids Res.* **2011**, *39*, W412–W415.

(17) Blaby, I. K.; Blaby-Haas, C. E.; Tourasse, N.; Hom, E. F. Y.; Lopez, D.; Aksoy, M.; Grossman, A.; Umen, J.; Dutcher, S.; Porter, M.; King, S.; Witman, G. B.; Stanke, M.; Harris, E. H.; Goodstein, D.; Grimwood, J.; Schmutz, J.; Vallon, O.; Merchant, S. S.; Prochnik, S. The *Chlamydomonas* genome project: a decade on. *Trends Plant Sci.* **2014**, *19*, 672–680.

(18) Croft, M. T.; Warren, M. J.; Smith, A. G. Algae need their vitamins. *Eukaryotic Cell* **2006**, *5*, 1175–1183.

(19) Benhamed, M.; Bertrand, C.; Servet, C.; Zhou, D.-X. *Arabidopsis GCN5*, *HD1*, and *TAF1/HAF2* interact to regulate histone acetylation required for light-responsive gene expression. *Plant Cell* **2006**, *18*, 2893–2903.

(20) Wittenberg, G.; Danon, A. Disulfide bond formation in chloroplasts: formation of disulfide bonds in signaling chloroplast proteins. *Plant Sci.* **2008**, *175*, 459–466.

(21) Niewiadomski, P.; Knappe, S.; Geimer, S.; Fischer, K.; Schulz, B.; Unte, U. S.; Rosso, M. G.; Ache, P.; Flügge, U.-I.; Schneider, A. The *Arabidopsis* plastidic glucose 6-phosphate/phosphate translocator *GPT1* is essential for pollen maturation and embryo sac development. *Plant Cell* **2005**, *17*, 760–775.

(22) Pfalz, J.; Liere, K.; Kandlbinder, A.; Dietz, K.-J.; Oelmüller, R. pTAC2 , -6 , and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. *Plant Cell* **2006**, *18*, 176–197.

(23) Page, M. D.; Allen, M. D.; Kropat, J.; Urzica, E. I.; Karpowicz, S. J.; Hsieh, S. I.; Loo, J. A.; Merchant, S. S. Fe sparing and Fe recycling contribute to increased superoxide dismutase capacity in iron-starved *Chlamydomonas reinhardtii*. *Plant Cell* **2012**, *24*, 2649–2665.

(24) Dastoor, Z.; Dreyer, J.-L. Potential role of nuclear translocation of glyceraldehyde-3-phosphate dehydrogenase in apoptosis and oxidative stress. *J. Cell Sci.* **2001**, *114*, 1643–1653.

(25) Hurley, B. A.; Tran, H. T.; Marty, N. J.; Park, J.; Snedden, W. A.; Mullen, R. T.; Plaxton, W. C. The dual-targeted purple acid Phosphatase isozyme AtPAP26 is essential for efficient acclimation of *Arabidopsis* to nutritional phosphate deprivation. *Plant Physiol.* **2010**, *153*, 1112–1122.

(26) Tran, H. T.; Hurley, B. A.; Plaxton, W. C. Feeding hungry plants: the role of purple acid phosphatases in phosphate nutrition. *Plant Sci.* **2010**, *179*, 14–27.

(27) Brocker, C.; Vasiliou, M.; Carpenter, S.; Carpenter, C.; Zhang, Y.; Wang, X.; Kotchoni, S. O.; Wood, A. J.; Kirch, H.-H.; Kopečný, D.; Nebert, D. W.; Vasiliou, V. Aldehyde dehydrogenase (ALDH) superfamily in plants: gene nomenclature and comparative genomics. *Planta* **2013**, *237*, 189–210.

(28) Dietz, K. J.; Horling, F.; König, J.; Baier, M. The function of the chloroplast 2-cysteine peroxiredoxin in peroxide detoxification and its regulation. *J. Exp. Bot.* **2002**, *53*, 1321–1329.

(29) Eriksson, M.; Karlsson, J.; Ramazanov, Z.; Gardeström, P.; Samuelsson, G. Discovery of an algal mitochondrial carbonic anhydrase: molecular cloning and characterization of a low-CO_2-induced polypeptide in *Chlamydomonas reinhardtii*. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 12031–12034.

(30) Ohnishi, N.; Mukherjee, B.; Tsujikawa, T.; Yanase, M.; Nakano, H.; Moroney, J. V; Fukuzawa, H. Expression of a low-CO_2-inducible protein, LCI1, increases inorganic carbon uptake in the green alga *Chlamydomonas reinhardtii*. *Plant Cell* **2010**, *22*, 3105–3117.

(31) Nedelcu, A. M. Sex as a response to oxidative stress: stress genes co-opted for sex. *Proc. R. Soc., B* **2005**, *272*, 1935–1940.

(32) Moseley, J.; Grossman, A. R. Phosphate metabolism and responses to phosphorus deficiency. In *The Chlamydomonas sourcebook*, 2nd ed.; Stern, DB.; Harris, EH., Eds.; Academic Press: New York, 2009; Vol. 2, pp. 189–216.

(33) Abe, J.; Kubo, T.; Takagi, Y.; Saito, T.; Miura, K.; Fukuzawa, H.; Matsuda, Y. The transcriptional program of synchronous gametogenesis in *Chlamydomonas reinhardtii*. *Curr. Genet.* **2004**, *46*, 304–315.

(34) Behrenfeld, M. J.; Halsey, K. H.; Milligan, A. J. Evolved physiological responses of phytoplankton to their integrated growth environment. *Philos. Trans. R. Soc., B* **2008**, *363*, 2687–2703.