Genomic Foundation of Starch-to-Lipid Switch in Oleaginous *Chlorella* spp.[1]

Jianhua Fan, Kang Ning, Xiaowei Zeng, Yuanchan Luo, Dongmei Wang, Jianqiang Hu, Jing Li, Hui Xu, Jianke Huang, Minxi Wan, Weiliang Wang, Daojing Zhang, Guomin Shen, Conglin Run, Junjie Liao, Lei Fang, Shi Huang, Xiaoyan Jing, Xiaoquan Su, Anhui Wang, Lili Bai, Zanmin Hu, Jian Xu*, and Yuanguang Li*

State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, Shanghai 200237, China (J.F., Y.Lu., H.X., J.Hua., M.W., W.W., D.Z., G.S., C.R., J.Lia., L.F., Y.Li.); Single-Cell Center, Chinese Academy of Sciences Key Laboratory of Biofuels and Shandong Key Laboratory of Energy Genetics, Qingdao Institute of BioEnergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao, Shandong 266101, China (K.N., X.Z., D.W., J.Hu., J.Li., S.H., X.J., X.S., A.W., J.X.); and Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China (L.B., Z.H.)

The ability to rapidly switch the intracellular energy storage form from starch to lipids is an advantageous trait for microalgae feedstock. To probe this mechanism, we sequenced the 56.8-Mbp genome of *Chlorella pyrenoidosa* FACHB-9, an industrial production strain for protein, starch, and lipids. The genome exhibits positive selection and gene family expansion in lipid and carbohydrate metabolism and genes related to cell cycle and stress response. Moreover, 10 lipid metabolism genes might be originated from bacteria via horizontal gene transfer. Transcriptomic dynamics tracked via messenger RNA sequencing over six time points during metabolic switch from starch-rich heterotrophy to lipid-rich photoautotrophy revealed that under heterotrophy, genes most strongly expressed were from the tricarboxylic acid cycle, respiratory chain, oxidative phosphorylation, gluconeogenesis, glyoxylate cycle, and amino acid metabolisms, whereas those most down-regulated were from fatty acid and oxidative pentose phosphate metabolism. The shift from heterotrophy into photoautotrophy highlights up-regulation of genes from carbon fixation, photosynthesis, fatty acid biosynthesis, the oxidative pentose phosphate pathway, and starch catabolism, which resulted in a marked redirection of metabolism, where the primary carbon source of glycine is no longer supplied to cell building blocks by the tricarboxylic acid cycle and gluconeogenesis, whereas carbon skeletons from photosynthesis and starch degradation may be directly channeled into fatty acid and protein biosynthesis. By establishing the first genetic transformation in industrial oleaginous *C. pyrenoidosa*, we further showed that overexpression of an NAD(H) kinase from Arabidopsis (*Arabidopsis thaliana*) increased cellular lipid content by 110.4%, yet without reducing growth rate. These findings provide a foundation for exploiting the metabolic switch in microalgae for improved photosynthetic production of food and fuels.

[1] This work was supported by the National Basic Research Program of China (grant nos. 2011CB200904 and 2012CB721101), the National Natural Science Foundation of China (grant nos. 31300295, 31010103907, and 31271410), the Chenguang Program by Shanghai Education Development Foundation and Shanghai Municipal Education Commission, the National Special Fund for State Key Laboratory of Bioreactor Engineering (grant no. 2060204), the National Key Technologies Research and Development Program of China (grant nos. 2011BAD14B02 and 2011BAD23B04), the Fundamental Research Funds for the Central Universities, and the Open Funding Project of the State Key Laboratory of Bioreactor Engineering.

* Address correspondence to ygli@ecust.edu.cn and xujian@qibebt.ac.cn.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Yuanguang Li (ygli@ecust.edu.cn).

J.F., J.Hua., Y.C., Y.Lu., W.W., M.W., and G.S. cultured the strain; J.Hua. isolated the microalgal DNA and mRNA; J.F., J.Hu., L.B., C.R., W.W., X.Z., X.J., D.W., H.X., J.Lia., L.F., D.Z., and Z.H. contributed to the experimental data collection and processing; J.F., K.N., J.Li., X.S., J.Hu., S.H., and A.W. analyzed the genomic and transcriptomic data; J.F., K.N., Y.Li., and J.X. designed the study and wrote the article.

www.plantphysiol.org/cgi/doi/10.1104/pp.15.01174

Microalgae are considered promising feedstock for sustainable, large-scale production of commodities such as food, feed, chemicals, materials, and biofuels (Chisti, 2007; Hu et al., 2008; Scott et al., 2010; Wijffels and Barbosa, 2010; Ho et al., 2011; Savage, 2011; Hildebrand et al., 2013). Compared to traditional land crops such as corn (*Zea mays*), rice (*Oryza sativa*), and wheat (*Triticum aestivum*), one key feature of microalgae-based production is their ability to rapidly switch their energy storage forms from polysaccharides (e.g., starch) to lipids or to high-value compounds such as proteins under certain environmental stresses (Sansawa and Endo, 2004; Fan et al., 2012b; Li et al., 2014b). This metabolic feature is of significance in commercialization of large-scale microalgal production systems for biofuels, as the switch between food and fuel production allows nimbleness and flexibility of operation in the volatile market. However, the molecular mechanism underlying such a metabolic switch in microalgae remains poorly understood (Merchant et al., 2012).

Chlorella spp. are a group of unicellular freshwater green microalgae that evolved over one billion years

ago (Heckman et al., 2001). They are among the most widely cultivated microalgae by mankind (Klein-Marcuschamer et al., 2013), with an annual net production of about 4,000 tons biomass worldwide. *Chlorella* spp. have served as a popular food supplement and animal feed in Asia, the United States, and Europe for centuries because of their high protein, lipid, and chlorophyll content (the highest amount among known plants; Lee, 1997; Spolaore et al., 2006). However, strain improvement of industrial *Chlorella* spp. has been hindered by the lack of genomics research models. Moreover, although genetic modification of *Chlorella vulgaris* (Chow and Tung, 1997) and *Chlorella ellipsoidea* (Chen et al., 2001) were achieved, none of the *Chlorella* spp. strains shown to be amenable to transformation are of commercial interest (Tran et al., 2013). Thus transformation of industrial production *Chlorella* spp. strains remains a key goal.

One important feature of *Chlorella* spp. is their ability to support high biomass productivity under both autotrophic and heterotrophic culture modes (Gladue and Maxey, 1994; Radmer and Parker, 1994; Lee, 1997; Doucha and Livansky, 2012). Under photoautotrophic conditions, most *Chlorella* spp. produce abundant proteins and lipids, but 1 to 2 weeks of cultivation is usually required from inoculation to harvest. Such a relatively long production cycle often results in low productivity and vulnerability to changes in the weather (Perez-Garcia et al., 2011). However, when cultivated heterotrophically, they reach much higher biomass productivity (>100 g L^{-1}; Doucha and Livansky, 2012) while accumulating significant levels of starch and much lower levels of lipids and proteins (Sansawa and Endo, 2004; Marsalkova et al., 2010; Choix et al., 2012). Remarkably, when cells are shifted from heterotrophic to photoautotrophic conditions, a short duration (1 to 2 d depending on the weather) of illumination induces a prominent change in the intracellular profile, with marked decreases in starch yet 70% to 120% increases in lipids and proteins (Fan et al., 2012a, 2012b; Han et al., 2012; Li et al., 2014b), resulting in both high target product contents (lipids and proteins) and high biomass productivity (Fig. 1; Supplemental Fig. S1). Such a controlled switch that exploits the advantages of both autotrophy and heterotrophy shows potential in large-scale food and biofuel production. Similar switches have been found in many microalgae and have been exploited in cultivation (e.g., via high-light induction or change of nutritional mode) to improve the productivity of valuable cellular components such as astaxanthin, ketocarotenoid, and α-tocopherol (Ogbonna et al., 1999; Hata et al., 2001; Zhang and Lee, 2001). Nevertheless, the molecular mechanism underlying the switch remains unclear.

In this study, we established an oleaginous industrial *Chlorella pyrenoidosa* strain FACHB-9 as a model to investigate the genomic foundation of heterotrophy-to-photoautotrophy metabolism and the associated metabolic switch. We sequenced and annotated the 56.8-Mbp genome sequence of *C. pyrenoidosa* FACHB-9. Furthermore, we employed replicated deep messenger RNA sequencing (mRNA-seq) to track transcriptome dynamics during the transition from heterotrophy to photoautotrophy at six time points along the complete process of the

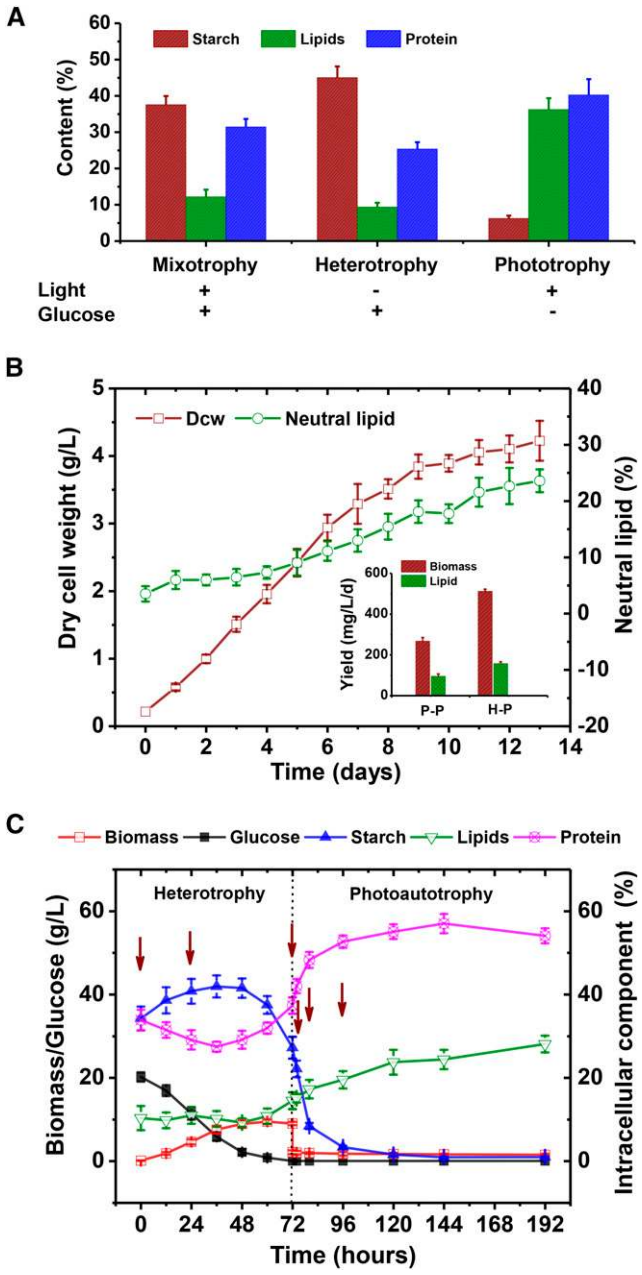

Figure 1. *C. pyrenoidosa* FACHB-9 as a research model for the starch-to-lipid switch and scalable production of lipids. A, Macromolecule profile under different culture models with heterotrophic cells as seed. B, Production of biomass and lipids by photoautotrophic growth in a 2-L column photobioreactor with 2% CO_2. Comparisons of lipid production for heterotrophic and photoautotrophic seed for subsequent photoautotrophic culture are also presented. P-P indicates photoautotrophic culture with photoautotrophic cells as seed, and H-P indicates photoautotrophic culture with heterotrophic cells as seed. C, Dynamics of algal growth and cellular components during SHDP. Arrows represent the six sampling time points for mRNA-seq (0, 24, 72, 74, 80, and 96 h). Dcw, Dry cell weight.

metabolic switch from starch-centric to lipid-centric production, revealing the underpinning molecular machineries and mechanism. Finally, by demonstrating the first genetic transformation of industrial oleaginous *Chlorella* spp., we showed that introducing a NAD(H) kinase gene into FACHB-9 increased the lipid content by 45.3% to 110.4% but did not affect the growth rate of the host cells under either heterotrophic or photoautotrophic conditions. These findings provide a foundation for exploiting the metabolic switch in microalgae for improved photosynthetic production of food and fuels.

RESULTS AND DISCUSSION

C. pyrenoidosa as a Research Model for Large-Scale Switchable Production of Starch and Lipids

C. pyrenoidosa is one of the most heavily studied species in the *Chlorella* genus. Due to its rich protein content, relatively fast growth rate, and tolerance of a wide range of environmental conditions, *C. pyrenoidosa* has been employed as commercial microalgal feedstock of proteins, lipids, and chlorophyll for both heterotrophic fermentation and outdoor mass cultivation (Running et al., 1994; Li et al., 1998; Mason, 2001). Under a standalone heterotrophic culture mode, *C. pyrenoidosa* strain FACHB-9 exhibited a competitive growth rate (average 2 g L^{-1} h^{-1}) and can reach greater than 150 g L^{-1} dry cell weight (Supplemental Fig. S2), which is comparable to a commercial yeast (*Saccharomyces cerevisiae*) fermentation system (approximately 130 g L^{-1}; Chen, 1996). In addition, the maximal biomass of the strain reached 4.2 g L^{-1} after 13-d photoautotrophic growth using heterotrophic cells as seed, with the intracellular neutral lipid content increasing from 3.5% to 23.6% (1.91 and 1.66 times higher in biomass and lipid productivity, respectively, than with photoautotrophic seed; Fig. 1B). These observations have allowed us to develop a process for *C. pyrenoidosa* cultivation called sequential heterotrophy-dilution-photoautotrophy (SHDP; Fan et al., 2012b).

SHDP consists of two phases: first, the heterotrophic culture mode for biomass accumulation, and subsequently, the photoautotrophic culture mode for production of target metabolites (Fig. 1C). In the heterotrophic mode, which lasts several days, Glc is gradually depleted while cell density steadily increased, resulting in accumulation of cellular starch to approximately 40% of dry cell weight (Fig. 1, A and C). This is followed by dilution, which is critical to improve light penetration of high-density cultures (e.g., >5 g L^{-1}) before moving into the photoautotrophic mode. After the dilution, treatments of heterotrophically growing cells with a sustained high light irradiation of 250 μmol m^{-2} s^{-1} for 24 h resulted in a marked increases in lipid, protein, and chlorophyll content but marked decreases in starch (Fig. 1C; Supplemental Fig. S1; Fan et al., 2012b). These SHDP-cultivated cells reach a state in which the metabolite profile is as rich in target products (lipid, protein, and chlorophyll) as that of autotrophically grown cells, yet the cultivation cycle is cut from the typical 1 to 2 weeks under the autotrophic mode to 3 to 5 d under SHDP (Fig. 1A). The high productivity demonstrated capability of commercial-scale cultivation and flexibility to switch cellular metabolites suggests *C. pyrenoidosa* FACHB-9 as an advantageous research model for the switchable production of starch, protein, and lipids.

Genomic Foundation of Starch and Lipid Production

General Feature of the Genome

The 56.6-Mbp nuclear genome of *C. pyrenoidosa* FACHB-9 was assembled from a combination of 15× shotgun and 10× paired-end sequence coverage of Roche 454 Titanium (average read length of 400 bp; Table I; Supplemental Table S1). Half of the assembled genome sequences are contained in nine scaffolds, each longer than 1.39 Mbp. The *C. pyrenoidosa* nuclear genome is slightly larger than *Chlorella variabilis* NC64A (46.2 Mbp; Blanc et al., 2010) and considerably larger than *Chlorella protothecoides* (22.9 Mbp; Gao et al., 2014), with *C. pyrenoidosa* predicted to contain more genes (10,284 as opposed to 9,791 and 7,039 genes, respectively; Supplemental Table S2), which suggests that great changes have taken place in the size of the genome after the diversification in *Chlorella* genus. This may be due to the different number of unique genes, multicopy genes, and genome duplication /rearrangements among close relatives (Gao et al., 2014).

To produce a high-quality map of gene structures, total RNA samples from both heterotrophic and autotrophic culture conditions were pooled and sequenced on 454 Titanium (533,300 raw reads produced; Supplemental Table S1) and assembled into 13,018 unique complementary DNA (cDNA) isotigs. Furthermore, to detect rare transcripts, 187,503,403 raw reads of mRNA collected from SHDP culture conditions and six time-points were collected (Supplemental Table S3). All of the cDNA reads that passed quality control were used for transcript-based gene prediction. The vast majority (97%) of the predicted genes were supported by these mRNA-seq reads (fragments per kilobase of exon model per million mapped fragments [FPKM] value > 1; Supplemental Data Set S1), and 89% of the predicted proteins were complete models with both start and stop codons (Table I).

The predicted *C. pyrenoidosa* protein-coding genes are rich in introns, averaging nearly nine introns per gene (the largest in known chlorophytes; Supplemental Table S2) and only 1.1% intron-free genes (Table I). Repeated sequences (such as retroelements, DNA transposons, simple repeats, and low-complexity sequences) made up merely 3.1% of the whole genome (Supplemental Table S4), suggesting a much lower repeat content than *Chlamydomonas reinhardtii* (Merchant et al., 2007).

Table I. *General features of C. pyrenoidosa FACHB-9 genome*

Assembly Features for Nuclear Genome	Results
Genome size	56.8 Mbp
Sequencing depth	26×
No. of contigs	8,193
No. of scaffolds (>2 kb)	1,336
Contig N/L50	1,265/11.14 kb
Scaffold N/L50	9/1.39 Mbp
Guanine and cytosine content	66.6%
Gene model quality	
Gene count	10,284
Multiexon genes	10,171 (98.9%)
Complete models with both start and stop codons	9,228 (89.7%)
Complete models extend both to 5′ and 3′ untranslated regions	9,159 (89.1%)
Gene model support	
Models with cDNA support (454 RNA sequencing [RNA-seq])	8,317 (80.9%)
Models with homology support (NCBI nr database)	8,992 (87.4%)
Models with Pfam domain alignments	6,249 (60.8%)
Models with CDD alignments	5,378 (52.3%)
Models with Kyoto Encyclopedia of Genes and Genomes (KEGG)/ European Commission assignments	7,799 (75.8%)
Models with GO assignments	5,639 (54.8%)

Genomic Features of C. pyrenoidosa That Drive Metabolic Versatility

Green algae have a complex phylogenetic history and constitute a large resource of genetic and metabolic diversity (Falkowski et al., 2004). A whole-genome phylogeny of Chlorophyta species (consistent with 18S rDNA-based phylogeny; Supplemental Fig. S3) was constructed from the 1,685 single-copy orthologous groups identified from seven sequenced green algal genomes (Fig. 2; Supplemental Results S1). Mapping these sets of single-copy orthologs allowed the selective pressure to be traced at every branch of chlorophyte evolution. At the Chlorophyta level, all 52 functional categories of genes are strongly constrained, with median estimates of ω (the K_a/K_s ratio of nonsynonymous to synonymous nucleotide divergence) much less than 1 (Fig. 3). The most constrained biological processes include photosynthesis, organelle organization, and protein phosphorylation, whereas the least constrained ones are lipid metabolism, phospholipid biosynthesis, glycolysis, and carbohydrate metabolism. When compared at the genus level (*C. pyrenoidosa* FACHB-9 and *C. variabilis* NC64A), the most constrained processes are protein folding, the ubiquitin-dependent protein catabolic process, and cellular component organization, and the least constrained are tRNA aminoacylation for protein translation, the biosynthesis process, and photosynthesis (Supplemental Fig. S4). Most of the functional categories, including lipid and carbohydrate metabolisms, are moderately constrained. Between the two *Chlorella* species, most genes are under strong negative selection ($\omega < 0.1$); however, 51 genes were found under positive selection with an average ω greater than 1 (Supplemental Table S5). The largest subclasses are oxidation-reduction process, protein binding, and transport-related proteins. The functional and evolutionary implications of such quickly evolving enzymes appear crucial to driving the diverse metabolic capacity of *Chlorella* spp.

To probe the evolutionary forces behind the expanded starch and lipid metabolic gene repertoire, phylogenetic analysis was employed to investigate the horizontal gene transfer (HGT) events in *C. pyrenoidosa* genome. Among the 231 genes related to starch metabolism (49 genes), fatty acid biosynthesis (45 genes), glycerolipid biosynthesis (48 genes), lipid activation (47 genes), and lipase (42 genes), 10 genes (all were related to the lipid metabolism, accounting for 5.5%) might have originated from bacteria via HGT (Supplemental Figs. S5 and S6).

Three genes encoding fatty acyl-CoA synthase (FACS; g11, g937, and g2922) identified were phylogenetically close to bacterial homologs, suggesting their possible evolutionary origin from bacteria via HGT. From the starting point of phototrophic process, transcriptional levels of all of the three genes exhibited significant up-regulation, while transcriptional levels of the other six long-chain (LC)-FACS genes slightly increased or decreased. The synchronous expression pattern of the HGT candidates FACS, along with the increased production of glycerolipids, indicates their important contribution in converting carbon flux from fatty acids into lipid.

Glycerol-3-phosphate acyltransferase (GPAT) genes usually have stringent one to two copy numbers in microalgae, even in heterokonts *Nannochloropsis* spp., which have highly expanded gene dose in both fatty acid biosynthesis and triacylglycerol (TAG) assembly pathways (Vieler et al., 2012; Wang et al., 2014). However, in *Chlorella* spp., in addition to the homologous gene inherited in microalgae, another GPAT gene (g2197) was identified and inferred as being originated from bacteria via HGT. This gene was transcriptionally active and up-regulated in the late heterotrophic process, suggesting its potential roles in lipid metabolism.

In addition to the above four HGT candidates, one β-ketoacyl-acyl-carrier protein (ACP) reductase gene (g7443) in fatty acid biosynthesis pathway, two enoyl-CoA hydratase genes (g4146, g4040) in fatty acid β-oxidation pathway in mitochondrion, and three lipase genes (g3730, g7283, and g10256) were inferred as HGT candidates. All of the three β-oxidation-related HGT candidates were dramatically up-regulated in late heterotrophic process and in the first several hours of phototrophic process. Thus, the diverse evolutionary origin of the *Chlorella* spp. lipid biosynthesis genes has underlain their massive genetic pools and likely contributed to the extraordinary capacity for lipid accumulation.

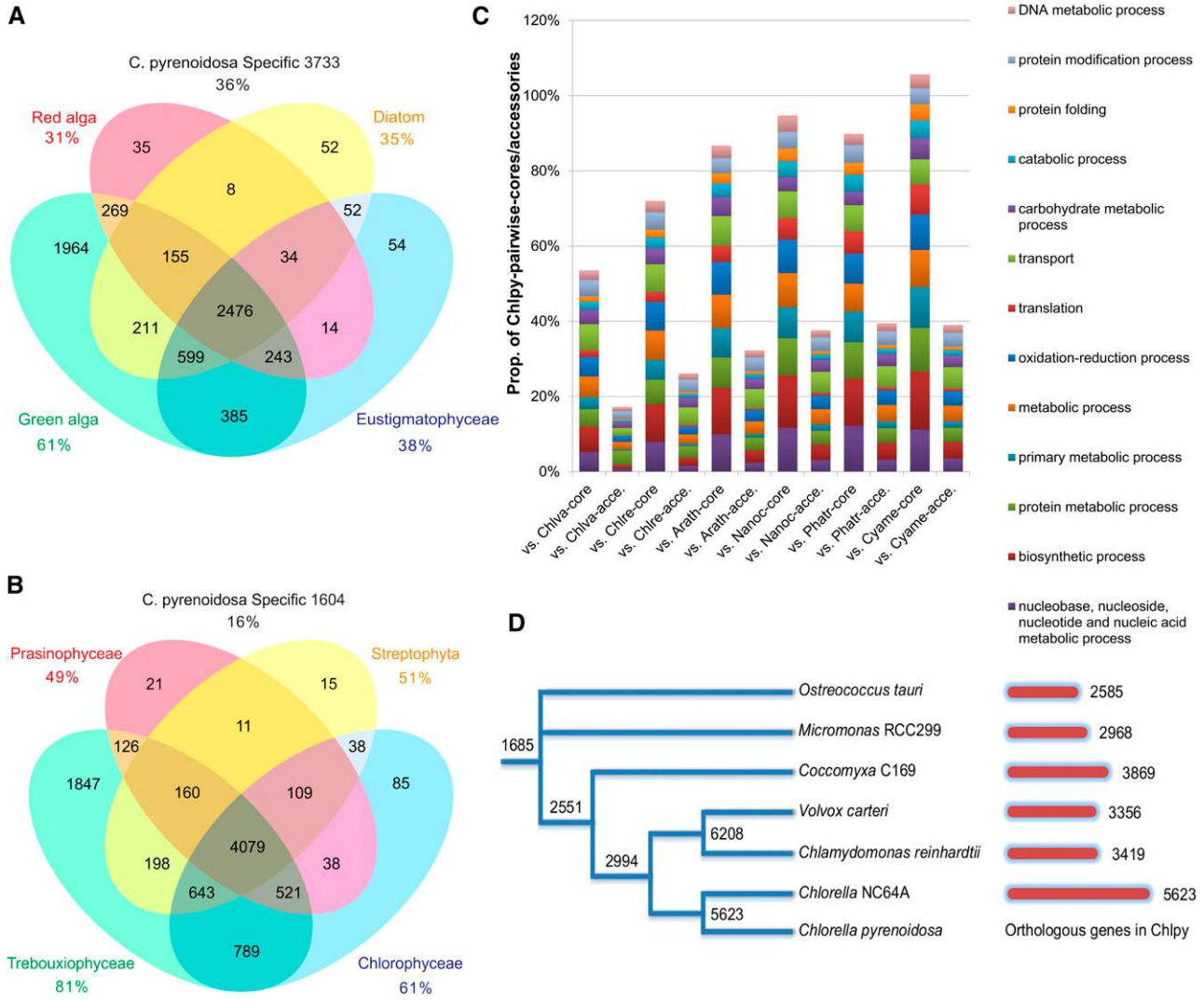

Figure 2. Phylogenetic analysis of the *C. pyrenoidosa* genome. A, Venn diagram representation of shared/unique genes of *C. pyrenoidosa* compared with red algae, diatoms, Eustigmatophyceae, and green algae. B, Venn diagram representation of shared/unique genes of *C. pyrenoidosa* compared with with other sequenced Chlorophyceae algae and land plants. C, Functional profiles of the six Chlpy-pairwise-cores and the six Chlpy-pairwise-accessories. D, Whole-genome-based phylogeny of Chlorophyta alga. A maximum-likelihood (consensus) tree generated based on the 1,685 single-copy orthologous gene sets identified from the seven selected green algal phylogenomes. At each branching point on the tree, the number of orthologous genes found within the corresponding branch is shown. The number of single-copy orthologous genes pairwise to *C. pyrenoidosa* is indicated along with strain names. Red alga, *Cyanidioschyzon merolae*; diatom, *P. tricornutum*; green alga, *C. reinhardtii*; Eustigmatophyceae, *N. oceanica*; Prasinophyceae, *Micromonas* RCC299; Streptophyta, Arabidopsis; Trebouxiophyceae, *C. variabilis*; and Chlorophyceae, *C. reinhardtii*.

Genomic Basis of Carbohydrate and Starch Metabolism in C. pyrenoidosa

Polysaccharides, including starch, are one major sink for photosynthetic carbon in many microalgae. The *C. pyrenoidosa* genome encodes a distinct functional profile in carbohydrate metabolism (as represented by Gene Ontology [GO] terms) from those of *C. reinhardtii*, *Phaeodactylum tricornutum*, and *Nannochloropsis oceanica* (Supplemental Table S6). In *C. pyrenoidosa*, GO terms in the carbohydrate metabolic process (GO:0005975) are prominently enriched, including the subcategories chitin metabolic process (GO:0006030), polysaccharide metabolic process (GO:0005976), lipopolysaccharide biosynthetic process (GO:0009103), and amino sugar metabolic process (GO:0006040). These findings suggest that carbohydrate metabolic machinery is a prominent feature of the *C. pyrenoidosa* genome.

Starch is a common form of polysaccharides and the major storage carbohydrate in plants and green algae (Zeeman et al., 2010). In the *C. pyrenoidosa* genome, for 12 of the 17 nodes along the starch metabolic pathway, multiple genes were found in each step (Supplemental Table S7). Comparison of the gene doses in each step of the starch metabolic pathways among six sequenced green

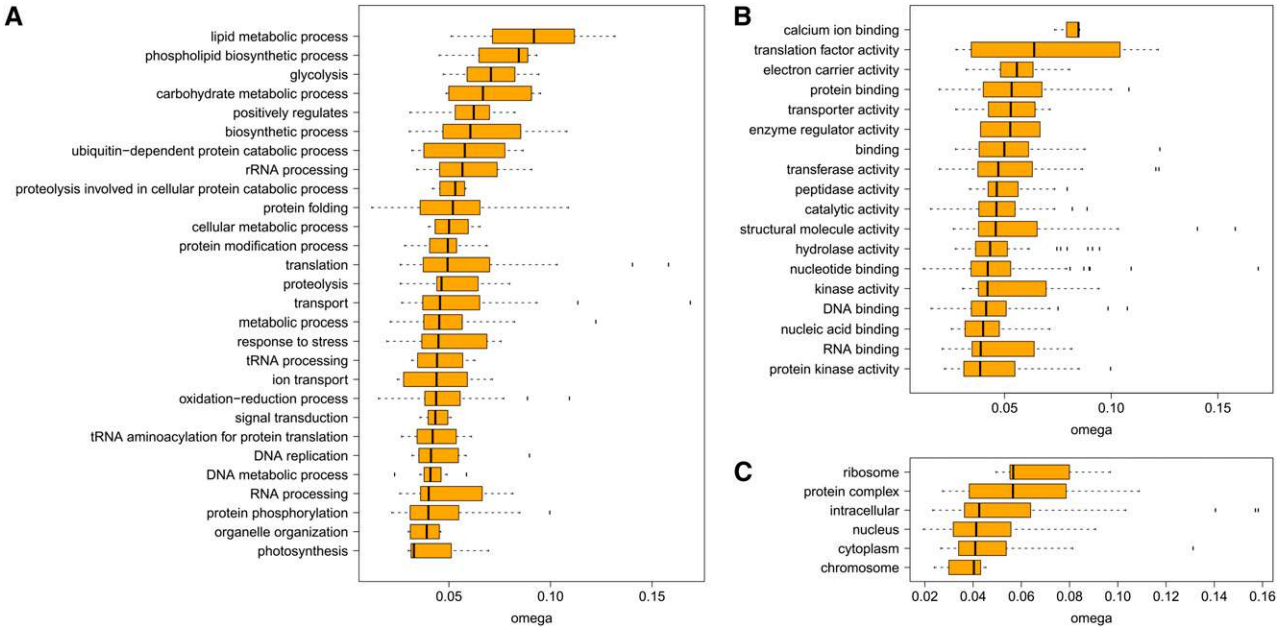

Figure 3. Selection pressure on protein-coding sequences in green algae. For each of the 1,685 seven-way single-copy orthologous gene sets identified among seven selected Chlorophyta species, Phylogenetic Analysis by Maximum Likelihood (PAML) model M0 was used to estimate a single ω (K_a/K_s, ratio of nonsynonymous to synonymous nucleotide divergence) that is fixed across the reconstructed whole-genome phylogeny. The associated GO slim terms that have at least three genes are shown for biological process (A), molecular function (B), and cellular component (C). rRNA, Ribosomal RNA.

algal species and strains revealed that *C. pyrenoidosa* harbors an expanded repertoire of genes involved in starch biosynthesis and degradation (Supplemental Table S8), in which selective steps such as ADP-Glc pyrophosphorylase, glucan-water dikinase, and 1,4-α-glucan branching enzyme were highly elevated in gene dose, while isoamylase and starch phosphorylase were less variable and identical in gene dose to other green algae. This suggests that, in green algae, the evolution of starch synthesis pathways is characterized by conservation of pathway backbones and divergence of gene doses in selected pathway nodes.

Genomic Foundation of Lipid Metabolism

To probe the mechanism behind the robust accumulation of neutral lipids and the rapid starch to lipid switch in *C. pyrenoidosa*, we reconstructed the fatty acid biosynthesis, TAG assembly, and lipid activation/degradation pathways (Supplemental Table S9). In *C. pyrenoidosa*, glycerolipids biosynthesis is presumably carried out by two distinct pathways: the prokaryotic pathway (located in the chloroplast), which is widely distributed in prokaryotes, microalgae, and higher plants, and the eukaryotic pathway, which occurs in the endoplasmic reticulum (Shen et al., 2010). Genes encoding both heteromeric (multisubunit; g2736, g8911, and g9519) and homomeric (multifunctional; g6472 and g6473) acetyl-CoA carboxylase (ACCase) were identified. Six putative LC-FACS genes were identified in the genome. Six copies of diacylglycerol acyltransferase (DGAT) genes were discovered in the *C. pyrenoidosa* genome (Supplemental Table S9). Five of the six *DGAT* genes encode putative type-2 DGATs, whereas the remaining one encodes type-1 DGAT (g7494). Previous research has indicated that there is an alternate pathway for TAG synthesis in yeast, plants (Dahlqvist et al., 2000), and *Chlamydomonas* spp. (Riekhof and Benning, 2009) that involves phospholipid:diacylglycerol acyltransferase (PDAT) to generate TAG using phospholipids as donors. One PDAT gene was identified in *C. pyrenoidosa* (g2349; Supplemental Table S9).

Comparison of GO term profile in *C. pyrenoidosa* with those in *C. reinhardtii*, *P. tricornutum*, and *N. oceanica* revealed the overrepresentation in *C. pyrenoidosa* of the GO terms for the phospholipid metabolic process (Supplemental Table S6). Moreover, the expansion of several other gene families was apparent, which may contribute to the robust biomass and lipid production phenotypes of *C. pyrenoidosa*. These include genes involved in photosynthesis, responses to stress and other organisms (including bacteria and fungi), phosphorus utilization, and the cell cycle (Supplemental Table S6).

TFs *Involved in Starch and Lipid-Related Metabolic Pathways*

Genome-wide identifications of transcriptional factors (TFs) by computational methods were reported for green algae such as *C. reinhardtii* and *Volvox carteri* (Pérez-Rodríguez et al., 2010; Zhang et al., 2011), red

algae such as *Galdieria sulphuraria* (Pérez-Rodríguez et al., 2010; Zhang et al., 2011), and Eustigmatophyceae strains such as *Nannochloropsis* spp. (Vieler et al., 2012; Hu et al., 2014). Here, we performed such analysis in *Chlorella* spp. via the characteristic domains of plant TFs using the methods of PlantTFDB (Zhang et al., 2011; Jin et al., 2014). The results revealed 195 TFs in *C. pyrenoidosa* (1.9% of the proteome; Supplemental Data Set S2). Each predicted TF was then assigned into a specific TF family based on its DNA-binding domain. MYB, Cys3His zinc finger domain (C3H), and basic Leucine Zipper (bZIP) are the three largest TFs families, together accounting for 36% of all TFs in *C. pyrenoidosa*. Orthologs in *C. pyrenoidosa* genome of higher plant TFs experimentally shown related to sugar signaling (e.g., WRKY, g3291; Sun et al., 2003) and lipid accumulation (e.g. Dof, g1875; Wang et al., 2007) were identified.

Fourteen and eight TFs were found putatively involved in the regulation of starch and lipid metabolism, respectively (Supplemental Data Set S2; Supplemental Methods S1). In starch-related TF genes, eight TF families that include bZIP, C3H, SQUAMOSA promoter binding proteins, MYB, STERILE APETALA, GATA, B3 (the third basic domain in the VIVIPAROUS1), and ARR-B (a Myb-like DNA binding domain called ARRM [type B]) were identified, whereas in lipid-related TFs, seven families that include Calmodulin binding transcription factors, DNA binding with one finger, Heat stress transcription factors, GATA, MYB-related, CONSTANS-like, and Nodule inception-like were found. Notably, the Dof family of TFs, which contains a single C2C2-type zinc-finger-like motif, was shown to regulate in plants processes that include lipid accumulation and photosynthesis (Papi et al., 2002; Wang et al., 2007; Shaw et al., 2009). One of the family members, *Gmdof4*, was found involved in lipid synthesis in *Glycine max* by activating the ACCase gene (Wang et al., 2007). Moreover, total fatty acid and lipid content were significantly increased in *Gmdof4*-overexpressing Arabidopsis (*Arabidopsis thaliana*; Wang et al., 2007) and *C. ellipsoidea* (Zhang et al., 2014). These TFs thus represent potential opportunities for engineering the regulatory network for enhanced starch or oil productivity.

Transcriptome Dynamics during the Switch Process

To model the temporal process featuring the switch of the intracellular metabolite profile from one dominated by starch to one dominated by proteins and lipids, we changed the *C. pyrenoidosa* cultivation mode from heterotrophy to photoautotrophy at 72 h (Fan et al., 2012a, 2012b; Han et al., 2012), followed by a 48-h light illumination. At each of the six time points during this SHDP process, three biological replicates of separate culture experiments were each sampled for transcriptome analysis via deep mRNA sequencing and metabolite analysis (Fig. 1C).

A total of 187,503,403 trimmed paired-end reads (50 nucleotides in length after trimming, representing 22× average coverage of the genome at each sample; Supplemental Table S3). Our results suggested that 1,578 and 1,491 nuclear genome-encoded genes were respectively up- and down-regulated greater than 3-fold at least one of the time points (Supplemental Fig. S7; Supplemental Data Sets S3 and S4). A set of 75 transcripts was found whose expression levels together were able to distinguish the cellular state among the six time points in a highly reliable manner, suggesting a transcript-based strategy to quantitatively model the SHDP process (Supplemental Fig. S8).

Overview of the Switch Process

To determine the time-dependent differential gene expression when cells were transferred from heterotrophic growth to photoautotrophy, the temporal patterns of the relative transcript abundance over the six time points were grouped by *k*-means clustering, which revealed 16 distinct groups of protein-coding genes (Fig. 4). During the course of the heterotrophic growth phase, the number of differentially expressed genes steadily increased, suggesting distinct cellular transcriptomic states during growth (Supplemental Fig. S7). The number of genes differentially expressed in the phototrophic process at time point 80 h was significantly higher than the other sampling points (Supplemental Fig. S7). This suggests a vigorous gene regulatory response in photoautotrophy, which can partly explain the rapid and massive changes in intracellular carbon storage forms within the first several hours (Fig. 1C; Fan et al., 2012b).

In the heterotrophic growth process, the genes most strongly up-regulated were those from the citrate cycle and amino acid metabolism, whereas those most down-regulated were from fatty acid biosynthesis and metabolism, pentose phosphate metabolism, and nucleotide metabolism (Supplemental Fig. S9). The shift from heterotrophy into photoautotrophy highlights the increased expression of genes involved in carbon fixation and photosynthesis that would be expected owing to light illumination (Fig. 5). Fatty acid biosynthesis, pentose phosphate pathway, and starch catabolism genes were also transcriptionally up-regulated.

How Was Starch Synthesized?

In *C. reinhardtii*, starch biosynthesis and degradation are under circadian clock control that appears uncorrelated with illumination (Ral et al., 2006). However, there is evidence for starch turnover in *Chlorella* spp. during growth in response to light intensity or nutritional status (Rosen et al., 1986; Fan et al., 2012b). In our experiments, the starch content of *C. pyrenoidosa* cells showed a comparable decrease between the exponential and stationary phases of heterotrophic growth (48–72 h; Fig. 1C; Sansawa and Endo, 2004; Fan et al., 2012b), indicating a highly active starch catabolism during the SHDP process (Fig. 6). Three transcripts encoding phosphoglucomutase (g7581) and ADP-Glc pyrophosphorylase (g5866 and

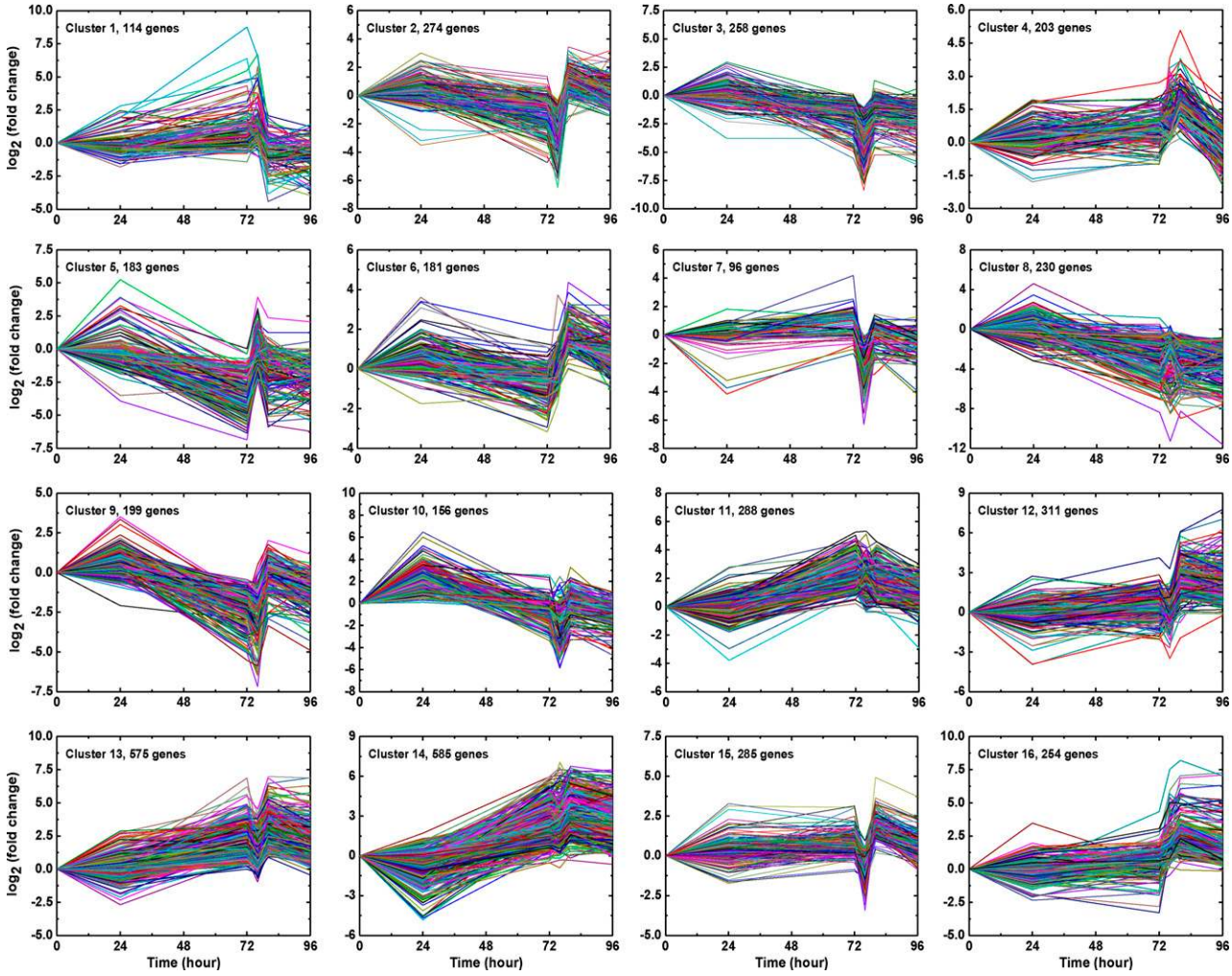

Figure 4. Clustering results of gene expression patterns in *C. pyrenoidosa* when shifting heterotrophic cells to photoautotrophy. The differentially expressed genes were clustered into nine groups using the *k*-means clustering method and visualized with TM4 software. The horizontal axis indicates the time points of the culture process, and the vertical axis is the $\log_2$ expression ratio. Fold expression changes between different time points (heterotrophy 24 h/heterotrophy 0 h, heterotrophy 72 h/heterotrophy 0 h, phototrophy 2 h/heterotrophy 72 h, phototrophy 8 h/heterotrophy 72 h, and phototrophy 24 h/heterotrophy 72 h) were calculated using the $\log_2$ ratios.

g9503), which define the beginning committed steps of starch biosynthesis, were down-regulated by the end of heterotrophy (72 h). Two granule-bound starch synthase I genes (g2429 and g7634) showed continuous up-regulation during the process of light induction (72–96 h). Transcripts responsible for starch breakdown, including glucan-water dikinase (g5653, g5659, and g7110), β-amylase (g3412 and g9517), phosphoglucan water dikinase (g4469 and g8607), and starch phosphorylase (g4743 and g4758), showed remarkable increases both in the heterotrophic (24–72 h) and phototrophic (74–96 h) phases. The enzyme activities of both ADP-glucose pyrophosphorylase and β-amylase showed the consistent trend of starch metabolism (Supplemental Table S10). Thus, the redirection of carbon flux away from storage carbohydrates in dark-grown cells at 72 h suggests that starch serves as a carbon and energy supply when external supplies of organic carbon are depleted after the end of heterotrophy.

How Was Lipid Synthesized?

The heterotrophy-to-photoautotrophy change of cultivation mode increased cellular lipid content but reduced starch content in *C. pyrenoidosa* (Fig. 1C) by an unknown mechanism. Glc was exhausted at the end of the heterotrophic process, and sudden illumination was supplied subsequently. Accordingly, genes encoding enzymes associated with fatty acid and neutral lipid metabolic pathways were differentially expressed (Fig. 7). Of the four genes encoding subunits of heteromeric ACCase, two (biotin carboxylase [g8911] and biotin carboxyl carrier protein [g2736]) decreased

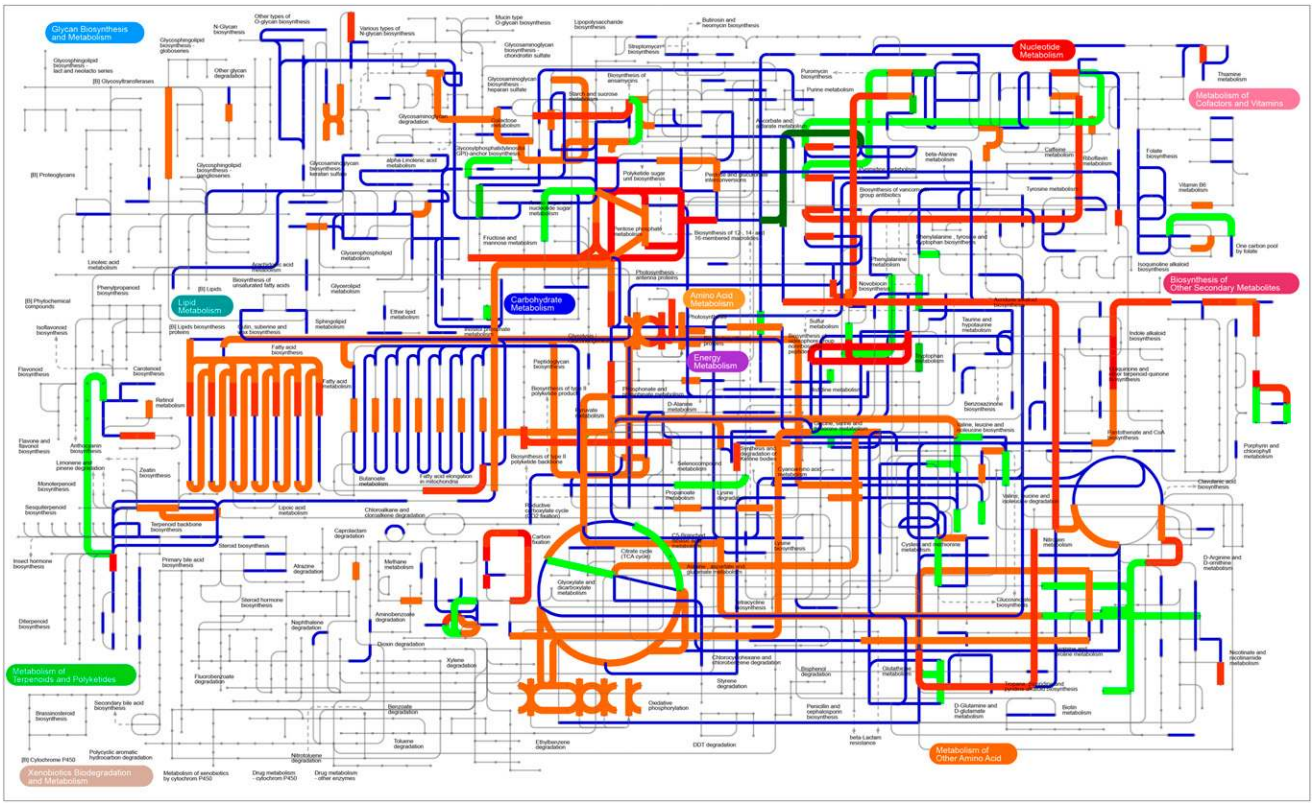

Figure 5. Overview of metabolic pathways and regulation during the heterotrophy to photoautotrophy transition in *C. pyrenoidosa*. KEGG pathways in blue, red, or green are present in the *C. pyrenoidosa* genome. Light-gray background lines indicate KEGG pathways not encoded by the *C. pyrenoidosa* genome. Genes that are up- or down-regulated (compared with heterotrophy 72 h) are labeled in red and green, respectively. Deeper red and green indicates greater fold changes of differential expression.

transcriptionally at the end of heterotrophic phase and then recovered quickly following 2 h of light induction. Enzyme activity of ACCase confirmed its up-regulation during the process (Supplemental Table S10). Similar patterns were also observed in fatty acid synthetase genes, including *MALONYL-COA:ACP TRANSACYLASE, BETA KETOACYL-ACP SYNTHASE, BETA KETOACYL-ACP REDUCTASE, BETA-HYDROXYACYL-ACP DEHYDRATASE, and ENOYL-ACP REDUCTASE* (Fig. 7; Supplemental Table S9). Thus, de novo fatty acid synthesis is modulated by light/dark conditions, with synthesis increasing in the light and decreasing in the dark (Sasaki et al., 1997).

Transferring the dark-grown cells (72 h) to light conditions (80 h) resulted in a slight increase in transcript abundance of acyl-ACP thioesterase (g1189), suggesting increased fatty acid export from the plastid to the endoplasmic reticulum, where TAG assembly occurs. Among the six putative *LC-FACS* genes identified in *C. pyrenoidosa* genome data set, two were almost completely suppressed under either or both growth conditions (FPKM < 5; g2654 and g6978; Supplemental Table S9). Another three genes showed a 4.4- to 17.5-fold decrease in mRNA abundance in the dark. Two of the three, g3983 and g9628, were up-regulated under light illumination; the remaining gene (g6510) showed an increase in transcript abundance at the log phase upon heterotrophic growth. These observations suggest distinct roles for individual *LC-FACS* genes and the presence of an intricate regulatory mechanism.

Transcriptional levels of TAG assembly pathways, including genes encoding glycerol-3-phosphate dehydrogenase (g2897 and g5566), lysophospholipid acyltransferase (LPAAT; g3070 and g9996), phosphatidic acid phosphatase (g6947), and diacylglycerol acyltransferase (type-2 DGAT; g3280 and g7566) showed a synchronized increase by the end of heterotrophic growth (72 h). In *Chlamydomonas* and *Nannochloropsis* spp., transcriptome and genetic analysis revealed contribution of acyltransferases to TAG accumulation and their role as potential targets for manipulating TAG hyperaccumulation (Boyle et al., 2012; Sanjaya et al., 2013; Li et al., 2014a). Several carotenoid biosynthesis genes exhibited similar expression patterns (Supplemental Data Set S3); one exception is that a LPAAT gene (g8553), predicted to function in the chloroplast prokaryotic pathway, was down-regulated at 72 h in dark-grown cells and then recovered by light induction. Thus, TAG accumulation, which usually accompanies secondary carotenoid synthesis (Rabbani

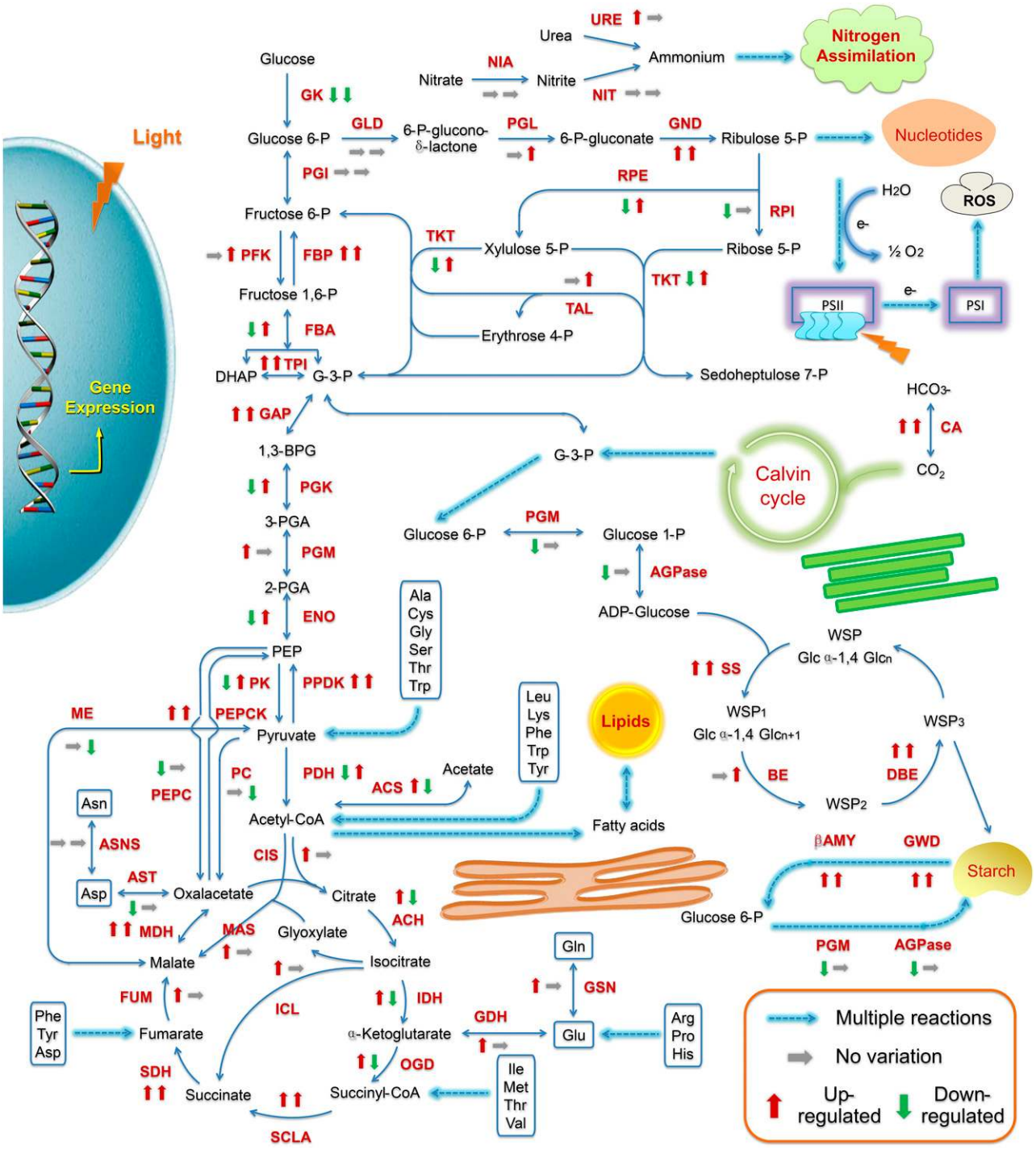

Figure 6. Genomic and transcriptomic features of central carbon metabolism in *C. pyrenoidosa*. Metabolic steps are represented by arrows. Dashed lines represent multiple metabolic steps. Boxes indicate those nodes where carbon skeletons from amino acid degradation feed into the pathway. Genes encoding the enzymes of these pathways are labeled in red. Up- or down-regulation of mRNA expression under heterotrophic growth (compared with heterotrophy 0 h, leftward arrows) and the heterotrophy to photoautotrophy transition (compared with heterotrophy 72 h, rightward arrows) based on mRNA-seq data are indicated with red upward arrows and green downward arrows, respectively. The full names of the corresponding genes are given in Supplemental Table S9.

et al., 1998; Zhekisheva et al., 2002), might be regulated by the heterotrophy-to-autotrophy switch, thus protecting the chloroplast from potential photooxidative damage. Therefore, the dynamic modulation of eukaryotic and prokaryotic pathways for glycerolipid synthesis, which accompanies the heterotrophy-to-photoautotrophy switch, may be crucial for forming the various intracellular energy storage compounds.

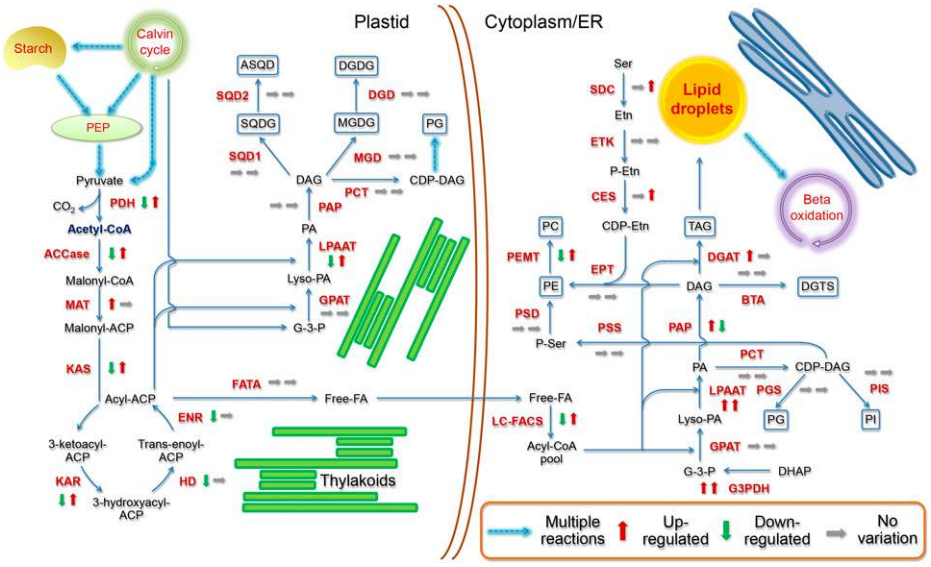

Figure 7. Fatty acid and glycerolipid biosynthesis and regulation in *C. pyrenoidosa* FACHB-9 as reconstructed from genomic and transcriptomic evidence. Metabolic steps are represented by arrows. Dashed lines represent the presence of multiple metabolic steps. End products are shown in boxes. Genes encoding the enzymes in these pathways are labeled in red. Up- or down-regulation of mRNA expression under heterotrophic growth (compared with heterotrophy 0 h, leftward arrows) and the heterotrophy to photoautotrophy transition (compared with heterotrophy 72 h, rightward arrows) based on mRNA-seq data are indicated with red upward arrows and green downward arrows, respectively. The full names of the corresponding genes are given in Supplemental Table S9. Because the glycolytic enolase is coded by only one gene (g2571) in *C. pyrenoidosa*, it is proposed that in *Chlorella* spp., the glycolytic enolase is only active in the cytosol, resulting in a specific pattern of carbon flow following fixation that involves export of carbon from the plastid and then reimport for pyruvate generation.

Dynamics of the Central Metabolism That Supports the Switch

Central metabolism pathways played a prominent role in the metabolic switch of *C. pyrenoidosa* (Fig. 6). Key enzymes from the tricarboxylic acid (TCA) cycle), respiratory chain, and oxidative phosphorylation were transcriptionally up-regulated at the 24-h time point in the dark (Fig. 5; Supplemental Data Set S3). The energy economy of cell growth is dominated by ATP production and the reducing agent NADH via the TCA cycle, which supports various metabolic processes. Consistent with the enrichment of starch granules at the beginning of the heterotrophic phase (Fan et al., 2012b), up-regulation of gluconeogenesis genes was evident within the first 24 h of heterotrophy, which directed carbon from central carbon metabolism to intracellular carbohydrate stores. One Fru-bisphosphate aldolase (g6760) and two glyceraldehyde-3-P dehydrogenase genes (g4391 and g3826) were up-regulated 3- to 16-fold (Supplemental Data Set S3). Although these enzymes can also catalyze the reversible reactions of glycolysis, transcripts of a phosphoenolpyruvate carboxykinase (g2374), one pyruvate phosphate dikinase (g7054), and three Fru 1, 6-bisphosphatase genes (g5024, g4411, and g3655) also increased markedly.

The glyoxylate cycle serves as a shunt of the TCA cycle and allows cells to utilize simple carbon compounds as carbon source when complex sources such as Glc are unavailable (Lorenz and Fink, 2002). This permits the synthesis of Glc from lipids via acetate generated in fatty acid β-oxidation. Transcripts of isocitrate lyases (g1361 and g4926) and malate synthases (g10080), which are specific to the glyoxylate cycle, showed a 3- to 25-fold increase at the end of the heterotrophic culture conditions (72 h). Accordingly, transcripts of acetyl-CoA synthetase (ACS; g7446, g7414, and g4467) increased 7- to 12-fold at the end of the heterotrophic process. ACS catalyzes acetate to produce acetyl-CoA, a precursor in the glyoxylate cycle, TCA cycle, and fatty acid biosynthesis. In contrast to the initial heterotrophic phase, majority of the transcripts for key enzymes of glycolysis, gluconeogenesis, Calvin and TCA cycles, and the pentose phosphate pathway declined at the end of heterotrophic process (i.e., after 72 h, when Glc became limiting), exhibiting a global down-regulation of central carbon metabolism (Supplemental Data Sets S3 and S4).

In addition, oxidative pentose phosphate (OPP) pathway genes, including 6-phosphogluconate dehydrogenase (decarboxylating; g4585) and 6-phosphogluconolactonase (g3624), were transcriptionally up-regulated approximately 3-fold under light conditions (80 h). The elevated OPP activity resulted in the generation of reducing equivalents (in the form of NADPH) that drive reductive biosynthesis reactions (e.g., fatty acid synthesis); thus, up-regulation of the OPP pathway may eventually contribute to lipid accumulation.

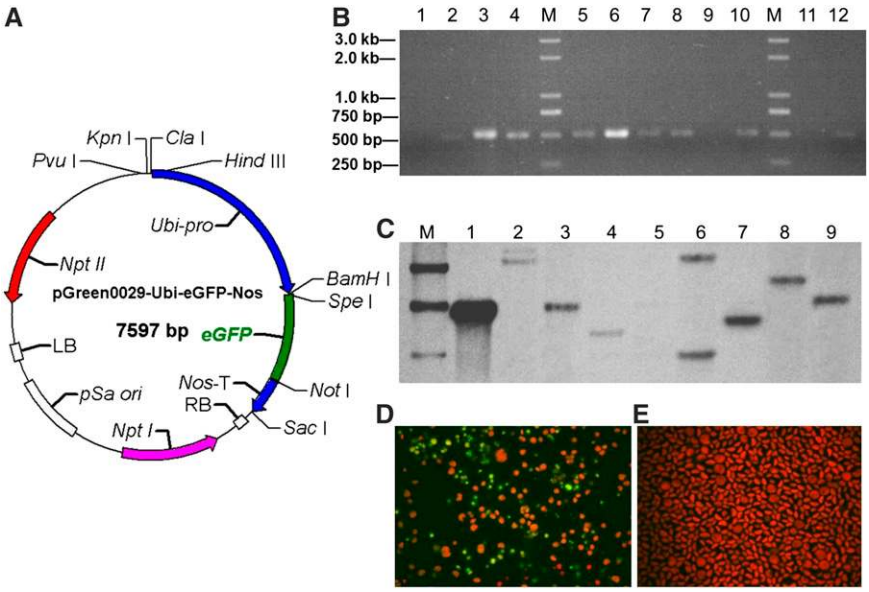

Figure 8. Transformation of *C. pyrenoidosa* FACHB-9 and validation of transgenic strains. A, A schematic map of the pGreen0029-Ubi-eGFP-Nos plasmid. The pGreen0029-Ubi-eGFP-Nos vector contained an expression box of the *eGFP* gene, controlled by the *Ubiquitin gene1* promoter (*Ubi-pro*) and terminated by the *Nos* terminator (*Nos-T*) and an expression box of the neomycin phosphotransferase enzyme (*NptII*), which conferred resistance to aminoglycoside antibiotics. B, PCR detection of the *NptII* gene using primers F and R. M, DNA molecular weight marker; 1, wild-type *C. pyrenoidosa* cells; 3, positive control, the pGreen0029-Ubi-eGFP-Nos plasmid; 2 and 4 to 12, strains of transgenic cells. C, Southern-blot analysis of *C. pyrenoidosa* transformants. The DNA from the wild type and transformants were digested with *Pst*I and *Eco*RI and then hybridized with a 464-bp fragment of the partial *eGFP* gene. M, DNA molecular weight marker; 1, positive control; the pGreen0029-Ubi-eGFP-Nos plasmid; 5, wild-type cells; 2 to 4 and 6 to 9, strains of transgenic cells. D, Identification of the transgenic strains using fluorescence microscopy observation. E, Identification of the negative control (which underwent the entire electroporation protocol without any plasmid DNA) using fluorescence microscopy.

The pyruvate dehydrogenase complex (PDC) subunits (g2537, g9493, g1999, and g1804) were also up-regulated following 2-h light illumination (Supplemental Data Set S4). The PDC transforms pyruvate into acetyl-CoA, which may then be used in fatty acid biosynthesis or as an input to the TCA cycle to carry out cellular respiration. Five fatty acid biosynthesis genes (g4578, g2736, g8879, g8526, and g8911) were rapidly up-regulated upon transferring cells from dark conditions to photoautotrophy (from 72–80 h; Supplemental Data Set S4), suggesting that carbon from photosynthesis or glycolysis might be channeled through acetyl-CoA and then into this pathway. Furthermore, changes of enzymatic activities of the key metabolic notes, including those from glycolysis, TCA cycle, gluconeogenesis, and oxidative pentose phosphate pathway, were consistent with the transcriptome analysis (Supplemental Table S10). Collectively, these findings reveal the complex interactions of the central metabolic pathways that jointly support the switch.

Reboot of Photosynthetic Pathways Promoted the Switch

Down-regulation of photosynthesis is one response to external shocks in microalgae such as nutrient deficiency, stress, or nutritional alteration (González-Ballester et al., 2010; Miller et al., 2010; Xiong et al., 2010; Msanne et al., 2012; Fan et al., 2014a, 2014b; Schmollinger et al., 2014). Most of the photosystem proteins, light-harvesting complexes, and the cytochrome b_6f complex were not down-regulated until 72 h in darkness (Supplemental Data Set S3). In contrast to their inert and decreased expression in heterotrophy (24 and 72 h), most of the photosynthesis-related genes showed a fast response, reaching their maximal expression levels at 2 h after photoautotrophy began (74 h; Supplemental Data Set S4), indicating that *C. pyrenoidosa* is able to rapidly reboot its photosynthetic apparatus to protect from photo damage after the onset of light.

Notably, carbon fixation genes showed a moderate response during heterotrophy (Supplemental Fig. S9; Supplemental Data Set S3). By contrast, a rapid increase in the transcript abundance of predicted chloroplast Calvin cycle genes was observed at the onset of light induction (74 h; Fig. 5; Supplemental Data Set S4). Moreover, up-regulation was observed for C4 dicarboxylic acid cycle enzymes, including NADP-malate dehydrogenase (NADP-MDH; g4904), carbonic anhydrase (g5902), pyruvate phosphate dikinase (PPDK; g7054), and phosphoenolpyruvate carboxykinase (g2374), following the nutritional conversion from heterotrophy to photoautotrophy (74–96 h; Supplemental Data Set S4). NADP-MDH and PPDK are both reported

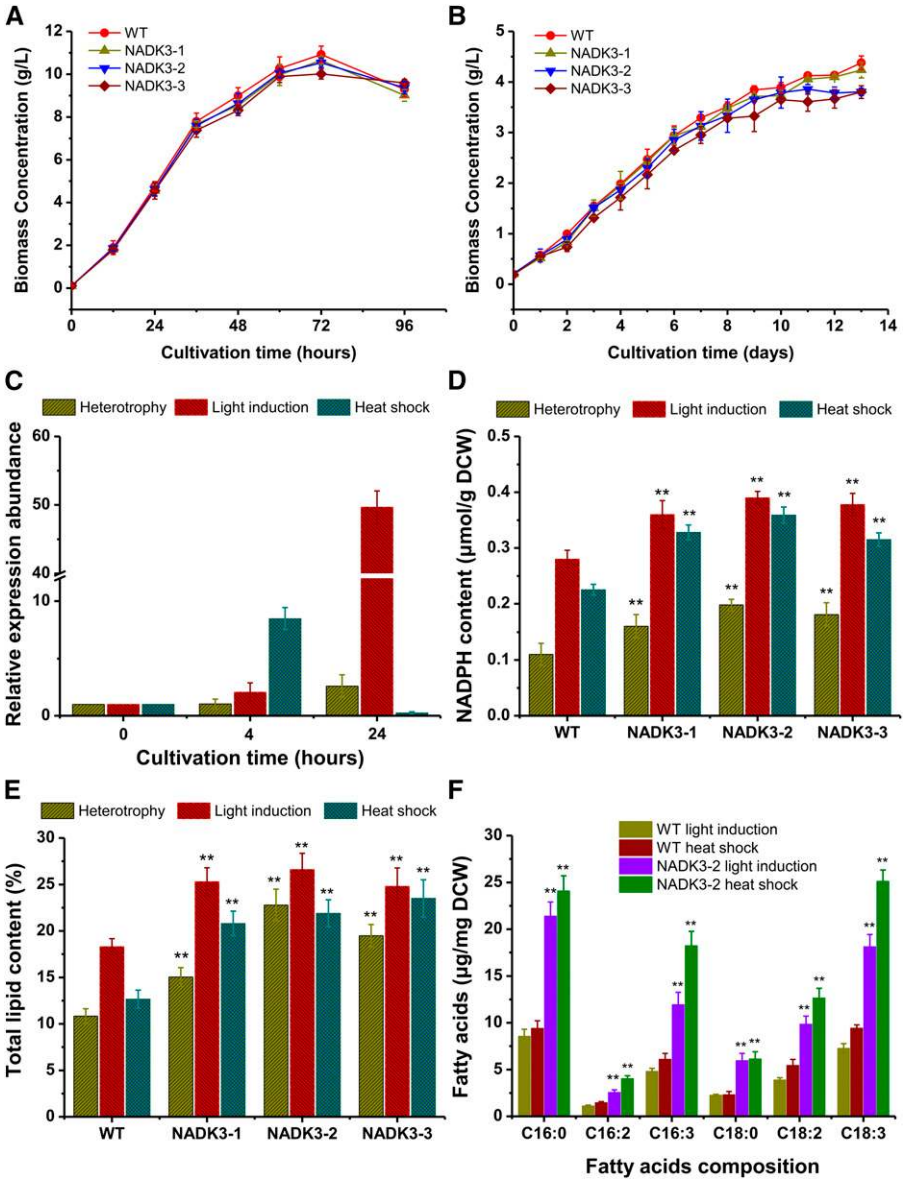

Figure 9. Characterization of *AtNADK3*-overexpressing *C. pyrenoidosa* FACHB-9 strains under different culture conditions. A, Growth curves of three *AtNADK3*-overexpressing strains under heterotrophic culture conditions. B, Growth curves of the three *AtNADK3*-overexpressing strains under photoautotrophic culture conditions. C, Relative expression levels of *AtNADK3* in the NADK3-2 strain under different culture conditions. D, NADPH content of the three *AtNADK3*-overexpressing strains under different culture conditions. E, Total lipid contents of the three *AtNADK3*-overexpressing strains under different culture conditions. F, The fatty acids content and composition in the NADK3-2 strain under different culture conditions. Data represent the means ± SD of three biological replicate experiments and were analyzed by Student's *t* test ($n = 3$). Asterisks indicate a significant difference from wild-type strains (**, $P < 0.01$). For heterotrophy, 72-h cultivation was carried out. For light induction (100 μmol m^{-2} s^{-1}) and heat shock (42°C), 24-h cultivation was carried out using heterotrophic cells as seed. DCW, Dry cell weight; WT, wild-type *C. pyrenoidosa*.

to be light-modulated enzymes that function in the C4 pathway of photosynthesis (Hatch and Slack, 1969; Carr et al., 1999; Ocheretina et al., 2000). Thus, nimble response of the carbon fixation pathway genes might underpin rapid adaptation of the microalga to the phototrophic culture.

Genetic Transformation of *C. pyrenoidosa* FACHB-9 Using Electroporation

To demonstrate amenability of industrial production *Chlorella* spp. strains to genetic engineering, electroporation of FACHB-9 was performed without prior enzymatic treatments and selection on 25 mg L^{-1} G418. A transformation rate of about 1.07×10^{-5} per microgram plasmid DNA was achieved (Supplemental Methods S1; Supplemental Table S11). The efficiency of the transformations was strongly affected by the promoter and field strength (Supplemental Table S11), as the highest number of colonies was generated using 3,300 V cm^{-1} field strength, and the promoter producing the highest number of transformants was *Heat shock protein70* (*Hsp70*; followed by *Tubulin*; Supplemental Table S11). Both genomic PCR and Southern blot confirmed presence and integration of the transgenes in selected colonies (Fig. 8, B and C). Fluorescence microscopy observation analysis confirmed successful stable expression of the exogenous GFP gene in resistant *C. pyrenoidosa* lines (Fig. 8, D and E). Stable integration and expression of the exogenous gene were confirmed after five generations of growth via antibiotic selection in both agar plates and liquid media.

Overexpression of the *AtNADK3* Significantly Enhances the Fatty Acids and Lipid Content in *C. pyrenoidosa* FACHB-9

Our transcriptome analysis indicated the switch from heterotrophy to light environment might lead to generation of reducing energy equivalents (e.g., in the form of NADPH) that drive reductive biosynthesis reactions (e.g., fatty acid synthesis). Thus, increasing the level of the intracellular reducing equivalents may eventually contribute to lipid accumulation. In Arabidopsis, NAD(H) kinase3, a key source of the cellular reductant NADPH, is required for various abiotic stress responses (Chai et al., 2006). The corresponding gene *AtNADK3* encodes a NADH kinase that can synthesize NADPH with NADH as the preferred substrate and also utilizes NAD+ (while the other two *AtNADK* genes can utilize only singular substrate; Chai et al., 2006). Intriguingly, in *C. pyrenoidosa* FACHB-9 genome, only one endogenous copy of NADK (g6804) was found. Transcriptome analysis showed that the *NADK3* native to *C. pyrenoidosa* was up-regulated by 360% and 41.6% in the heterotrophic and light induction process (Supplemental Data Set S1), suggesting an active role of this transcript in lipid production. To test whether higher dose of this enzyme could lead to NADPH accumulation and thus promote fatty acid synthesis, we introduced into FACHB-9 one more copy of *AtNADK3* using the genetic transformation system we developed (Supplemental Methods S1).

Growth curves of the transgenic *AtNADK3* strains and wild-type *C. pyrenoidosa* were similar at either heterotrophy or photoautotrophy (Fig. 9, A and B), suggesting absence of obvious deleterious effects on growth as introduced by the transformation. Because *Hsp70* promoter, as a transcriptional enhancer, promoted transgene expression in *C. reinhardtii* under inducing conditions (Rosales-Mendoza et al., 2012), we compared three culture conditions (heterotrophy, photoinduction, and heat shock) to probe whether the native *Hsp70* promoter works as constitutive or inducible type in *C. pyrenoidosa*. All the three conditions induced *AtNADK3* expression, yet with distinct effects (Fig. 9C; Supplemental Fig. S10). Compared with the wild type, all three transgenic strains exhibited significantly increased NADPH and lipid content, by 39.3% to 79.9% (Fig. 9D) and 45.3% to 110.4% (Fig. 9E), respectively. Main fatty acid components of the transgenic strain NADK3-2 and the wild type were similar, as revealed by gas chromatography-mass spectrometry analysis (Fig. 9F). Both light induction and heat shock increased the cellular content of the fatty acids (Fig. 9F) compared with the wild type, yet the relative compositions of total fatty acids were not significantly different from that in the wild type (Fig. 9F). Among the three investigated induction strategies, light was the most effective way (followed by heat shock and then heterotrophy) to induce both gene expression and NADPH accumulation as well as lipid biosynthesis (Supplemental Figs. S10 and S11).

Overexpression of *NADK* causes perturbation of NADP(H) pool and has positive effects on growth or stress tolerance (Pollak et al., 2007; Panagiotou et al., 2009), e.g., stimulation of photosynthesis metabolism and tolerance of oxidative damages in higher plants (Takahashi et al., 2009; Takahara et al., 2010). Our study is, to our knowledge, the first to show that the NADP(H) pool can be engineered to increase lipid content, by introducing a single NAD(H) kinase gene in microalgae. Thus, these findings suggest a new strategy for strain development for higher lipid productivity in not just industrial microalgae but also higher plants.

CONCLUSION

The green algae *Chlorella* spp. are among the most widely cultivated microalgae in the world (Klein-Marcuschamer et al., 2013). Genome of the industrial strain *C. pyrenoidosa* FACHB-9 revealed positive selection and gene family expansion in lipid and carbohydrate metabolism genes, as well as those related to the cell cycle and response to stimuli. HGTs were inferred in several lipid metabolism nodes. Temporal analysis of transcriptomes during the metabolic switch from a starch-rich heterotrophy mode to a protein- and lipid-rich photoautotrophy mode revealed a marked redirection of metabolism in which the primary carbon source, Glc, is no longer supplied to cell building blocks by the TCA cycle and gluconeogenesis, while carbon skeletons from photosynthesis and starch degradation may be channeled directly into fatty acid and protein biosynthesis. By establishing, to our knowledge, the first genetic transformation for industrial *Chlorella* spp., we show that overexpression of the *AtNADK3* significantly enhances the fatty acids and lipid content without affecting its growth. This is, to our knowledge, the first report of increasing microalgal fatty acid and lipid content by genetically engineering cellular-reducing equivalents. These findings thus serve as a foundation for engineering the metabolic switch apparatus in this and related algae for optimized and controllable production of food and fuels.

MATERIALS AND METHODS

For a detailed description of all methods, see Supplemental Methods S1.

Strain and Culture Conditions

Chlorella pyrenoidosa FACHB-9 (Freshwater Algae Culture Collection of the Institute of Hydrobiology, Chinese Academy of Sciences) was cultured heterotrophically in Endo medium (Endo et al., 1974) in flasks or in 50- and 500-L fermenters with Glc and urea as the organic carbon and nitrogen source. F-Si medium (Sato et al., 2006) was chosen for photoautotrophic culture lipid production with 2% (v/v) CO_2/air supplementation.

For SHDP cultivation, the heterotrophically grown cells (in which Glc was already depleted) were resuspended in water at a final cell concentration of about 2 g L^{-1} and then transferred to 5-d phototrophic cultivation under indoor continuous light at approximately 300 μmol m^{-2} s^{-1} at 30°C. For each of

cultivation processes above, samples were taken at specific intervals to determine biomass and cellular component production.

Genome Sequencing and Assembly

For *C. pyrenoidosa* FACHB-9 genome sequencing, we employed Roche 454 Titanium to collect single-end and paired-end reads. We first generated a total of 3,730,590 454-Titanium sequence reads (average read length, approximately 400 bp, with pair-end distances of 3, 8, and 20 kb, respectively). Assembly of the single- and pair-ended 454 sequences (1.46 Gbp in total and approximately 26-fold coverage) with Newbler version 2.0 (Roche) resulted in 8,193 contigs that assembled into 1,345 scaffolds longer than 2 kb. The assembled scaffolds were further screened and filtered by searching against all bacterial sequences from the SILVA (Pruesse et al., 2007) and National Center for Biotechnology Information (NCBI) nr database to filter out possible contaminants. The resultant assembly of the *C. pyrenoidosa* nuclear genome contains 1,336 scaffolds totaling 56.8 Mbp in size, with a contig N/L50 of 1,265/11 kb and a scaffold N/L50 of 9/1.39 Mbp.

Algal cells grown under different heterotrophic and phototrophic conditions (±Glc, low-/high-light illumination, and logarithmic/stationary phase) were collected and harvested by centrifugation. Each sample was then pooled to prepare libraries of cDNA for mRNA sequencing on 454 Titanium (Roche). A total of 533,300 raw reads were produced using GS-Titanium. All raw reads were trimmed based on quality value before further analysis. All cDNA reads that passed quality control were used for transcript-based gene prediction. About 81% (8,317 genes) of the total number of predicted protein-coding genes were covered by the 454 cDNA sequencing reads.

Genome Annotation and Analyses

The assembled draft nuclear genome of *C. pyrenoidosa* was annotated using the QIBEBT-FUNGEA annotation pipeline (accessibility), which combines three different ab initio gene predictors (Augustus version 2.5 [Stanke and Morgenstern, 2005], GeneID v1.2 [Parra et al., 2000], and SNAP [Korf, 2004]) and one evidence-based gene predictor (Cufflinks version 1.1.0 [Trapnell et al., 2010]). The cDNA reads that passed quality control were also used for transcript-based gene prediction. Functional annotation of the predicted protein-coding genes was carried out based on the identification of protein domains using InterProScan (Quevillon et al., 2005) database and BLASTp matches against the NCBI nr, Kyoto Encyclopedia of Genes and Genomes (KEGG) and Conserved Domain Database (CDD) databases, with an E-value $< 10^{-5}$. Protein sequences from *C. pyrenoidosa* were used as BLASTp queries against the sequenced alga species, and hits with an E-value $< 10^{-10}$ were considered homologous proteins. For selection pressure analysis of protein-coding genes in green algae, orthologs among seven selected Chlorophyta strains were first identified by a Markov Clustering algorithm (Li et al., 2003). PAML (Yang, 2007) codon substitution models and likelihood ratio tests (codeml) were then used to estimate the rate of evolution and to test the selection pressure. To connect gene function with sequence evolution, GO term assignments for each of the genes were retrieved from InterproScan results.

RNA-Seq and Differential Expression Analysis

Total RNA was extracted from *C. pyrenoidosa* cells during the SHDP process. Briefly, the SHDP cultivations were performed in 500-mL shaker flasks. Algal cells were first cultivated heterotrophically to achieve high cell density. After Glc was completely consumed, the broth was diluted to approximately 2 g L^{-1} dry biomass and transferred to a light environment for photoautotrophic culture. Three separate flasks (i.e., three biological replicates) were analyzed in parallel (Supplemental Methods S1). As a result, 18 samples in total, with the three biological replicates collected from each of the six time points (0, 24, and 72 h from heterotrophic cells and 2, 8, and 24 h from phototrophic cells; Fig. 1C), were used for mRNA-seq library preparation and then submitted to Solexa GAIIx (Illumina) for sequencing (short insert [approximately 400 bp] paired-end, 2×100 and 2×90 bp). Default parameters were used to screen reads using the Illumina quality control pipeline. The processed sequence files (a total of 187,503,403 reads) were then mapped to the assembled genome by TopHat version 2.0.4 (Trapnell et al., 2009) with a tolerance of up to two mismatches and three indels. For each of the RNA-seq data sets, the transcript relative abundance (TRA) was expressed as the number of aligned reads to annotated gene models using Cufflinks (version 1.1.0; Trapnell et al., 2010) and then normalized to FPKM values. Fold expression changes between different time points were calculated using the TRA $\log_2$ ratios. Genes were regarded as differentially expressed if they had at least a 3-fold change and 5% or less false discovery rate. Furthermore, the gene sets with positive TRA $\log_2$ ratios that had numerators greater than 50 FPKM were regarded as up-regulated. Down-regulated genes were those with negative TRA $\log_2$ ratios that had denominators greater than 50 FPKM.

Nuclear Transformation by Electroporation

FACHB-9 cells were cultured up to logarithmic phase, harvested following centrifugation at 3,000 rpm for 10 min, and then processed with a high-exudate buffer containing 0.2 M sorbitol and mannitol on ice. The processed cells were recentrifuged at 3,000 rpm and resuspended with a buffer at a concentration of approximately 5×10^6 cells mL^{-1} and then immediately mixed with a final concentration of 30 μg mL^{-1} plasmid pGreen0029-Ubi-eGFP-Nos (freely available by contacting Jianhua Fan), a final concentration of 10 μg mL^{-1} plasmid pSoup, and 25 μg mL^{-1} salmon sperm DNA. The mixture was placed on ice for 10 min before a 100-μL mixture was supplied into the 2-mm electroporation cuvette for transformation with a Bio-Rad Gene Pulser II Electroporation System. The cells were transformed at a field strength of 1.8 to 4.3 kV cm^{-1} leading to time constants of 1 to 5 ms. Following electroporation, cells were screened via selective medium (in agar plates with 25 mg L^{-1} G418; Nichols, 1973) and validated.

Construction, Validation, and Phenotyping of *AtNADK3*-Overexpression Strains of *C. pyrenoidosa*

To test whether higher dose of this enzyme could lead to NADPH accumulation and thus promote fatty acid synthesis, we introduced into FACHB-9 one more copy of *AtNADK3* using the genetic transformation system we developed. The cDNA of *AtNADK3* was amplified from the leaves of Arabidopsis (*Arabidopsis thaliana*) according to the published *AtNADK3* cDNA sequence (NCBI accession no. NM_106506), and the fragments were then inserted into the site between *Spe*I and *Not*I in the plastid pGreen0029-Ubi-eGFP-Nos to replace *eGFP*. Similarly, the promoter *Ubiquitin1* gene was also replaced with endogenous promoter with *Hin*dIII and *Bam*HI to create the pGreen0029-Hsp70-AtNADK3-Nos.

Fatty acid composition of mutant cells was qualitatively and quantitatively determined using a gas chromatography mass spectrometer. For gene expression, the first-strand cDNA synthesis and real-time quantitative PCR were performed using the ReverTra Ace qPCR RT Master Mix with gDNA Remover and SYBR Green Realtime PCR Master Mix (Toyobo), respectively. A pair of specific primers (forward, ATCGCATCCCATTCGTGAT, and reverse, TGCCAAACAAAATGTCATCCA) were designed and used for expression-level quantitative detection. The comparative threshold cycle method was used to analyze gene expression relative to control based on the average fold change. The content of NADPH was determined using the coenzyme II NADP (H) content kit (SuZhou Comin).

The *C. pyrenoidosa* whole-genome sequencing project (PRJNA171991) has been deposited at DDBJ/EMBL/GenBank under the accession number ANZC00000000. The version described in this paper is the first version, ANZC01000000. The mRNA-seq raw data are available at the Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE40028. The genome and transcriptome data can also be freely accessed through the project Web site (http://www.bioenergychina.org:8989/indexs/Chlorella_pyrenoidosa_index.html).

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Confocal microscopy images of Nile red-stained *C. pyrenoidosa* cells during SHDP process and nitrogen deprivation under photoautotrophic conditions.

Supplemental Figure S2. Scale-up the *C. pyrenoidosa* heterotrophic growth in 50 to 500-L fermentors by repeated fed-batch cultivation for seeds supply.

Supplemental Figure S3. Phylogenetic position of *C. pyrenoidosa* among algae, plants, and stramenopiles.

Supplemental Figure S4. Selection pressure on protein-coding sequences in two *Chlorella* spp.

Supplemental Figure S5. Phylogenetic analysis of lipid synthesis genes that were inferred to have originated from HGT using the neighbor-joining method.

Supplemental Figure S6. Phylogenetic analysis of lipid synthesis genes that were inferred to have originated from HGT, using the maximal-likelihood method.

Supplemental Figure S7. Number of differentially expressed genes identified at each time point for both heterotrophic growth and phototrophic process.

Supplemental Figure S8. Transcriptomic signature analysis of the SHDP process.

Supplemental Figure S9. Overview of metabolic pathways and regulation under heterotrophic growth in *C. pyrenoidosa*.

Supplemental Figure S10. Expression of *AtNADK3* into *C. pyrenoidosa* cells under different culture conditions.

Supplemental Figure S11. Analyses of ultrastructure for *C. pyrenoidosa* transformant NADK3-2 expressing *AtNADK3* under different culture conditions.

Supplemental Figure S12. Guanine-cytosine content distribution of raw reads from the two 454 sequenced libraries for genome assembly.

Supplemental Figure S13. Saturation curves of transcriptome sequencing reads.

Supplemental Figure S14. Spearman correlation upon biological replicate samples.

Supplemental Figure S15. Comparison of fold difference in RNA abundance during SHDP process estimated from mRNA-seq versus real-time PCR experiments.

Supplemental Figure S16. Quality of reads after filtering via FastQC.

Supplemental Table S1. The genome sequencing data for *C. pyrenoidosa* using Roche 454.

Supplemental Table S2. Comparison of *C. pyrenoidosa* genome statistics with those of sequenced chlorophytes genomes.

Supplemental Table S3. Details of the transcriptome sequencing data for *C. pyrenoidosa* using an Illumina solexa platform.

Supplemental Table S4. Repeat sequences in the *C. pyrenoidosa* genome.

Supplemental Table S5. Annotation of single-copy orthologs between two *Chlorella* spp. with an average K_a/K_s value greater than one.

Supplemental Table S6. Selected GO term enrichment and depletion of *C. pyrenoidosa* proteins compared with other alga.

Supplemental Table S7. Starch metabolic pathway genes in *C. pyrenoidosa*.

Supplemental Table S8. Starch metabolic pathway gene comparison among green algae.

Supplemental Table S9. Fatty acid and lipid metabolic pathway genes in *C. pyrenoidosa*.

Supplemental Table S10. Enzyme activities in *C. pyrenoidosa* grown under the sequential heterotrophy-dilution-photoautotrophy model.

Supplemental Table S11. Transformation rates achieved by electroporation of *C. pyrenoidosa* FACHB-9 cells in the presence of pGreen0029 plasmid with different voltage field strengths and promoters.

Supplemental Data Set S1. BLASTp annotation of *C. pyrenoidosa* nuclear genome-encoded gene in NCBI nr database.

Supplemental Data Set S2. TFs involved in starch- and lipid-related metabolic pathways.

Supplemental Data Set S3. Expression levels of differentially expressed genes grouped by putative functions under heterotrophic process.

Supplemental Data Set S4. Expression levels of differentially expressed genes grouped by putative functions under photo-induced process.

Supplemental Methods S1. Details for the methods used in this article.

Supplemental Results S1. A more detailed description of the results aside from the main text.

ACKNOWLEDGMENTS

We thank Dr. Qiang Hu for critical discussions and for reading the article.

Received July 29, 2015; accepted October 20, 2015; published October 20, 2015.

LITERATURE CITED

Blanc G, Duncan G, Agarkova I, Borodovsky M, Gurnon J, Kuo A, Lindquist E, Lucas S, Pangilinan J, Polle J, et al (2010) The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. Plant Cell **22:** 2943–2955

Boyle NR, Page MD, Liu B, Blaby IK, Casero D, Kropat J, Cokus SJ, Hong-Hermesdorf A, Shaw J, Karpowicz SJ, et al (2012) Three acyltransferases and nitrogen-responsive regulator are implicated in nitrogen starvation-induced triacylglycerol accumulation in *Chlamydomonas*. J Biol Chem **287:** 15811–15825

Carr PD, Verger D, Ashton AR, Ollis DL (1999) Chloroplast NADP-malate dehydrogenase: structural basis of light-dependent regulation of activity by thiol oxidation and reduction. Structure **7:** 461–475

Chai MF, Wei PC, Chen QJ, An R, Chen J, Yang S, Wang XC (2006) NADK3, a novel cytoplasmic source of NADPH, is required under conditions of oxidative stress and modulates abscisic acid responses in Arabidopsis. Plant J **47:** 665–674

Chen F (1996) High cell density culture of microalgae in heterotrophic growth. Trends Biotechnol **14:** 421–426

Chen Y, Wang Y, Sun Y, Zhang L, Li W (2001) Highly efficient expression of rabbit neutrophil peptide-1 gene in *Chlorella ellipsoidea* cells. Curr Genet **39:** 365–370

Chisti Y (2007) Biodiesel from microalgae. Biotechnol Adv **25:** 294–306

Choix FJ, de-Bashan LE, Bashan Y (2012) Enhanced accumulation of starch and total carbohydrates in alginate-immobilized *Chlorella* spp. induced by *Azospirillum brasilense*. II. Heterotrophic conditions. Enzyme Microb Technol **51:** 300–309

Chow KC, Tung WL (1997) Electrotransformation of *Chlorella vulgaris*. Plant Physiol **114:** 1610

Dahlqvist A, Stahl U, Lenman M, Banas A, Lee M, Sandager L, Ronne H, Stymne S (2000) Phospholipid:diacylglycerol acyltransferase: an enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. Proc Natl Acad Sci USA **97:** 6487–6492

Doucha J, Livansky K (2012) Production of high-density *Chlorella* culture grown in fermenters. J Appl Phycol **24:** 35–43

Endo H, Nakajima K, Chino R, Shirota M (1974) Growth characteristics and cellular components of *Chlorella regularis*, heterotrophic fast growing strain. Agric Biol Chem **38:** 9–18

Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJR (2004) The evolution of modern eukaryotic phytoplankton. Science **305:** 354–360

Fan J, Cui Y, Huang J, Wang W, Yin W, Hu Z, Li Y (2012a) Suppression subtractive hybridization reveals transcript profiling of *Chlorella* under heterotrophy to photoautotrophy transition. PLoS One **7:** e50414

Fan J, Cui Y, Wan M, Wang W, Li Y (2014a) Lipid accumulation and biosynthesis genes response of the oleaginous *Chlorella pyrenoidosa* under three nutrition stressors. Biotechnol Biofuels **7:** 17

Fan J, Cui Y, Zhou Y, Wan M, Wang W, Xie J, Li Y (2014b) The effect of nutrition pattern alteration on *Chlorella pyrenoidosa* growth, lipid biosynthesis-related gene transcription. Bioresour Technol **164:** 214–220

Fan J, Huang J, Li Y, Han F, Wang J, Li X, Wang W, Li S (2012b) Sequential heterotrophy-dilution-photoinduction cultivation for efficient microalgal biomass and lipid production. Bioresour Technol **112:** 206–211

Gao C, Wang Y, Shen Y, Yan D, He X, Dai J, Wu Q (2014) Oil accumulation mechanisms of the oleaginous microalga *Chlorella protothecoides* revealed through its genome, transcriptomes, and proteomes. BMC Genomics **15:** 582

Gladue RM, Maxey JE (1994) Microalgal feeds for aquaculture. J Appl Phycol **6:** 131–141

González-Ballester D, Casero D, Cokus S, Pellegrini M, Merchant SS, Grossman AR (2010) RNA-seq analysis of sulfur-deprived *Chlamydomonas* cells reveals aspects of acclimation critical for cell survival. Plant Cell **22:** 2058–2084

Han F, Huang J, Li Y, Wang W, Wang J, Fan J, Shen G (2012) Enhancement of microalgal biomass and lipid productivities by a model of photoautotrophic culture with heterotrophic cells as seed. Bioresour Technol **118:** 431–437

Hata N, Ogbonna JC, Hasegawa Y, Taroda H, Tanaka H (2001) Production of astaxanthin by *Haematococcus pluvialis* in a sequential heterotrophic-photoautotrophic culture. J Appl Phycol **13:** 395–402

Hatch MD, Slack CR (1969) Studies on the mechanism of activation and inactivation of pyruvate, phosphate dikinase. A possible regulatory role for the enzyme in the C4 dicarboxylic acid pathway of photosynthesis. Biochem J **112:** 549–558

Heckman DS, Geiser DM, Eidell BR, Stauffer RL, Kardos NL, Hedges SB (2001) Molecular evidence for the early colonization of land by fungi and plants. Science **293:** 1129–1133

Hildebrand M, Abbriano RM, Polle JE, Traller JC, Trentacoste EM, Smith SR, Davis AK (2013) Metabolic and cellular organization in evolutionarily diverse microalgae as related to biofuels production. Curr Opin Chem Biol **17:** 506–514

Ho SH, Chen CY, Lee DJ, Chang JS (2011) Perspectives on microalgal CO_2-emission mitigation systems: a review. Biotechnol Adv **29:** 189–198

Hu J, Wang D, Li J, Jing G, Ning K, Xu J (2014) Genome-wide identification of transcription factors and transcription-factor binding sites in oleaginous microalgae *Nannochloropsis*. Sci Rep **4:** 5454

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

Jin J, Zhang H, Kong L, Gao G, Luo J (2014) PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. Nucleic Acids Res **42:** D1182–D1187

Klein-Marcuschamer D, Chisti Y, Benemann JR, Lewis D (2013) A matter of detail: assessing the true potential of microalgal biofuels. Biotechnol Bioeng **110:** 2317–2322

Korf I (2004) Gene finding in novel genomes. BMC Bioinformatics **5:** 59

Lee YK (1997) Commercial production of microalgae in the Asia-Pacific rim. J Appl Phycol **9:** 403–411

Li J, Han D, Wang D, Ning K, Jia J, Wei L, Jing X, Huang S, Chen J, Li Y, et al (2014a) Choreography of transcriptomes and lipidomes of *Nannochloropsis* reveals the mechanisms of oil synthesis in microalgae. Plant Cell **26:** 1645–1665

Li L, Stoeckert CJ Jr, Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res **13:** 2178–2189

Li SW, Li HQ, Zhang JJ (1998) Progress in studies on large-scale culture of *Chlorella*. Chin Bull Bot **15:** 45–50

Li YQ, Xu H, Han FX, Mu JX, Chen D, Feng B, Zeng HY (2014b) Regulation of lipid metabolism in the green microalga *Chlorella protothecoides* by heterotrophy–photoinduction cultivation regime. Bioresour Technol dx.doi.org/10.1016/j.biortech.2014.1007.1028

Lorenz MC, Fink GR (2002) Life and death in a macrophage: role of the glyoxylate cycle in virulence. Eukaryot Cell **1:** 657–662

Marsalkova B, Sirmerova M, Kurec M, Branyik T, Branyikova I, Melzoch K, Zachleder V (2010) Microalgae Chlorella sp. as an alternative source of fermentable sugars. Pres 2010: 13th International Conference on Process Integration, Modelling and Optimisation for Energy Saving and Pollution Reduction 21**:** 1279-1284

Mason R (2001) *Chlorella* and *Spirulina*: Green supplements for balancing the body. Altern Complement Ther **7:** 161–165

Merchant SS, Kropat J, Liu B, Shaw J, Warakanont J (2012) TAG, you're it! *Chlamydomonas* as a reference organism for understanding algal triacylglycerol accumulation. Curr Opin Biotechnol **23:** 352–363

Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, Witman GB, Terry A, Salamov A, Fritz-Laylin LK, Maréchal-Drouard L, et al (2007) The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. Science **318:** 245–250

Miller R, Wu G, Deshpande RR, Vieler A, Gärtner K, Li X, Moellering ER, Zäuner S, Cornish AJ, Liu B, et al (2010) Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen deprivation predict diversion of metabolism. Plant Physiol **154:** 1737–1752

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

Nichols HW (1973) Growth media-freshwater. Handbook of Phycological Methods, Culture Methods and Growth Measurements. University Press, Cambridge, United Kingdom

Ocheretina O, Haferkamp I, Tellioglu H, Scheibe R (2000) Light-modulated NADP-malate dehydrogenases from mossfern and green algae: insights into evolution of the enzyme's regulation. Gene **258:** 147–154

Ogbonna JC, Tomiyama S, Tanaka H (1999) Production of α-tocopherol by sequential heterotrophic-photoautotrophic cultivation of *Euglena gracilis*. J Biotechnol **70:** 213–221

Panagiotou G, Grotkjaer T, Hofmann G, Bapat PM, Olsson L (2009) Overexpression of a novel endogenous NADH kinase in *Aspergillus nidulans* enhances growth. Metab Eng **11:** 31–39

Papi M, Sabatini S, Altamura MM, Hennig L, Schäfer E, Costantino P, Vittorioso P (2002) Inactivation of the phloem-specific Dof zinc finger gene *DAG1* affects response to light and integrity of the testa of Arabidopsis seeds. Plant Physiol **128:** 411–417

Parra G, Blanco E, Guigó R (2000) GeneID in *Drosophila*. Genome Res **10:** 511–515

Perez-Garcia O, Escalante FM, de-Bashan LE, Bashan Y (2011) Heterotrophic cultures of microalgae: metabolism and potential products. Water Res **45:** 11–36

Pérez-Rodríguez P, Riaño-Pachón DM, Corrêa LG, Rensing SA, Kersten B, Mueller-Roeber B (2010) PlnTFDB: updated content and new features of the plant transcription factor database. Nucleic Acids Res **38:** D822–D827

Pollak N, Niere M, Ziegler M (2007) NAD kinase levels control the NADPH concentration in human cells. J Biol Chem **282:** 33562–33571

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res **35:** 7188–7196

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R (2005) InterProScan: protein domains identifier. Nucleic Acids Res **33:** W116–W120

Rabbani S, Beyer P, Lintig J, Hugueney P, Kleinig H (1998) Induced β-carotene synthesis driven by triacylglycerol deposition in the unicellular alga *Dunaliella bardawil*. Plant Physiol **116:** 1239–1248

Radmer RJ, Parker BC (1994) Commercial applications of algae: opportunities and constraints. J Appl Phycol **6:** 93–98

Ral JP, Colleoni C, Wattebled F, Dauvillée D, Nempont C, Deschamps P, Li Z, Morell MK, Chibbar R, Purton S, et al (2006) Circadian clock regulation of starch metabolism establishes GBSSI as a major contributor to amylopectin synthesis in *Chlamydomonas reinhardtii*. Plant Physiol **142:** 305–317

Riekhof WR, Benning C (2009) Glycerolipid biosynthesis. *In* DB Stern, EE Harris, eds, The *Chlamydomonas* Sourcebook: Organellar and Metabolic Processes, Ed 2 **Vol 2**. Academic Press pp 41–68

Rosales-Mendoza S, Paz-Maldonado LMT, Soria-Guerra RE (2012) *Chlamydomonas reinhardtii* as a viable platform for the production of recombinant proteins: current status and perspectives. Plant Cell Rep **31:** 479–494

Rosen BH, Berliner MD, Petro MJ (1986) Analysis of starch content in *Chlorella pyrenoidosa* (Chlorophyta) by morphological and quantitative techniques. Am J Bot **73:** 1372–1375

Running JA, Huss RJ, Olson PT (1994) Heterotrophic production of ascorbic-acid by microalgae. J Appl Phycol **6:** 99–104

Sanjaya MR, Miller R, Durrett TP, Kosma DK, Lydic TA, Muthan B, Koo AJ, Bukhman YV, Reid GE, Howe GA, et al (2013) Altered lipid composition and enhanced nutritional value of *Arabidopsis* leaves following introduction of an algal diacylglycerol acyltransferase 2. Plant Cell **25:** 677–693

Sansawa H, Endo H (2004) Production of intracellular phytochemicals in *Chlorella* under heterotrophic conditions. J Biosci Bioeng **98:** 437–444

Sasaki Y, Kozaki A, Hatano M (1997) Link between light and fatty acid synthesis: thioredoxin-linked reductive activation of plastidic acetyl-CoA carboxylase. Proc Natl Acad Sci USA **94:** 11096–11101

Sato T, Usui S, Tsuchiya Y, Kondo Y (2006) Invention of outdoor closed type photobioreactor for microalgae. Energy Convers Manage **47:** 791–799

Savage N (2011) Algae: the scum solution. Nature **474:** S15–S16

Schmollinger S, Mühlhaus T, Boyle NR, Blaby IK, Casero D, Mettler T, Moseley JL, Kropat J, Sommer F, Strenkert D, et al (2014) Nitrogen-sparing mechanisms in *Chlamydomonas* affect the transcriptome, the proteome, and photosynthetic metabolism. Plant Cell **26:** 1410–1435

Scott SA, Davey MP, Dennis JS, Horst I, Howe CJ, Lea-Smith DJ, Smith AG (2010) Biodiesel from algae: challenges and prospects. Curr Opin Biotechnol **21:** 277–286

Shaw LM, McIntyre CL, Gresshoff PM, Xue GP (2009) Members of the Dof transcription factor family in *Triticum aestivum* are associated with light-mediated gene regulation. Funct Integr Genomics **9:** 485–498

Shen W, Li JQ, Dauk M, Huang Y, Periappuram C, Wei Y, Zou J (2010) Metabolic and transcriptional responses of glycerolipid pathways to a perturbation of glycerol 3-phosphate metabolism in Arabidopsis. J Biol Chem **285:** 22957–22965

Spolaore P, Joannis-Cassan C, Duran E, Isambert A (2006) Commercial applications of microalgae. J Biosci Bioeng **101:** 87–96

Stanke M, Morgenstern B (2005) AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res **33:** W465–W467

Sun C, Palmqvist S, Olsson H, Borén M, Ahlandsberg S, Jansson C (2003) A novel WRKY transcription factor, SUSIBA2, participates in sugar signaling in barley by binding to the sugar-responsive elements of the iso1 promoter. Plant Cell **15:** 2076–2092

Takahara K, Kasajima I, Takahashi H, Hashida SN, Itami T, Onodera H, Toki S, Yanagisawa S, Kawai-Yamada M, Uchimiya H (2010) Metabolome and photochemical analysis of rice plants overexpressing Arabidopsis NAD kinase gene. Plant Physiol **152:** 1863–1873

Takahashi H, Takahara K, Hashida SN, Hirabayashi T, Fujimori T, Kawai-Yamada M, Yamaya T, Yanagisawa S, Uchimiya H (2009) Pleiotropic modulation of carbon and nitrogen metabolism in Arabidopsis plants overexpressing the NAD kinase2 gene. Plant Physiol **151:** 100–113

Tran M, Van C, Barrera DJ, Pettersson PL, Peinado CD, Bui J, Mayfield SP (2013) Production of unique immunotoxin cancer therapeutics in algal chloroplasts. Proc Natl Acad Sci USA **110:** E15–E22

Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics **25:** 1105–1111

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol **28:** 511–515

Vieler A, Wu G, Tsai CH, Bullard B, Cornish AJ, Harvey C, Reca IB, Thornburg C, Achawanantakun R, Buehl CJ, et al (2012) Genome, functional gene annotation, and nuclear transformation of the heterokont oleaginous alga *Nannochloropsis oceanica* CCMP1779. PLoS Genet **8:** e1003064

Wang D, Ning K, Li J, Hu J, Han D, Wang H, Zeng X, Jing X, Zhou Q, Su X, et al (2014) *Nannochloropsis* genomes reveal evolution of microalgal oleaginous traits. PLoS Genet **10:** e1004094

Wang HW, Zhang B, Hao YJ, Huang J, Tian AG, Liao Y, Zhang JS, Chen SY (2007) The soybean Dof-type transcription factor genes, GmDof4 and GmDof11, enhance lipid content in the seeds of transgenic Arabidopsis plants. Plant J **52:** 716–729

Wijffels RH, Barbosa MJ (2010) An outlook on microalgal biofuels. Science **329:** 796–799

Xiong W, Gao C, Yan D, Wu C, Wu Q (2010) Double CO_2 fixation in photosynthesis-fermentation model enhances algal lipid synthesis for biodiesel production. Bioresour Technol **101:** 2287–2293

Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol **24:** 1586–1591

Zeeman SC, Kossmann J, Smith AM (2010) Starch: its metabolism, evolution, and biotechnological modification in plants. Annu Rev Plant Biol **61:** 209–234

Zhang DH, Lee YK (2001) Two-step process for ketocarotenoid production by a green alga, *Chlorococcum* sp. strain MA-1. Appl Microbiol Biotechnol **55:** 537–540

Zhang H, Jin J, Tang L, Zhao Y, Gu X, Gao G, Luo J (2011) PlantTFDB 2.0: update and improvement of the comprehensive plant transcription factor database. Nucleic Acids Res **39:** D1114–D1117

Zhang J, Hao Q, Bai L, Xu J, Yin W, Song L, Xu L, Guo X, Fan C, Chen Y, et al (2014) Overexpression of the soybean transcription factor GmDof4 significantly enhances the lipid content of *Chlorella ellipsoidea*. Biotechnol Biofuels **7:** 128

Zhekisheva M, Boussiba S, Khozin-Goldberg I, Zarka A, Cohen Z (2002) Accumulation of oleic acid in *Haematococcus pluvialis* (Chlorophyceae) under nitrogen starvation or high light is correlated with that of astaxanthin esters. J Phycol **38:** 325–331