Supplementary information

The genome-scale metabolic model for *Chlorella vulgaris* UTEX 395 reveals intracellular flux distributions and compartmental pathway activity during autotrophic, heterotrophic, and mixotrophic growth

Cristal Zuñiga, Chien-Ting Li, Tyler Huelsman, Jennifer Levering, Daniel C. Zielinski, Brian O. McConnell, Christopher P. Long, Eric P. Knoshaug, Michael T. Guarnieri, Maciek R. Antoniewicz, Michael J. Betenbaugh, Karsten Zengler*.

Index

Supplementary text

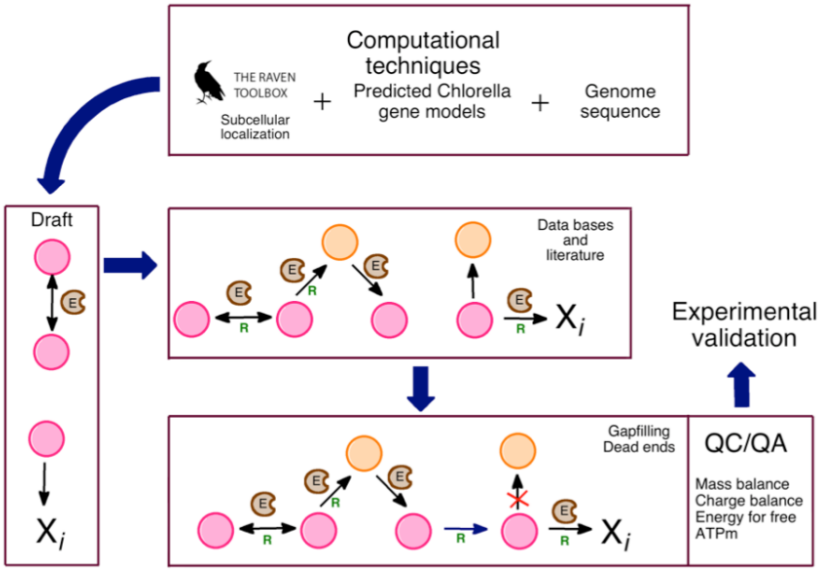

Figure S1. Manual curation workflow used for the metabolic network reconstruction.

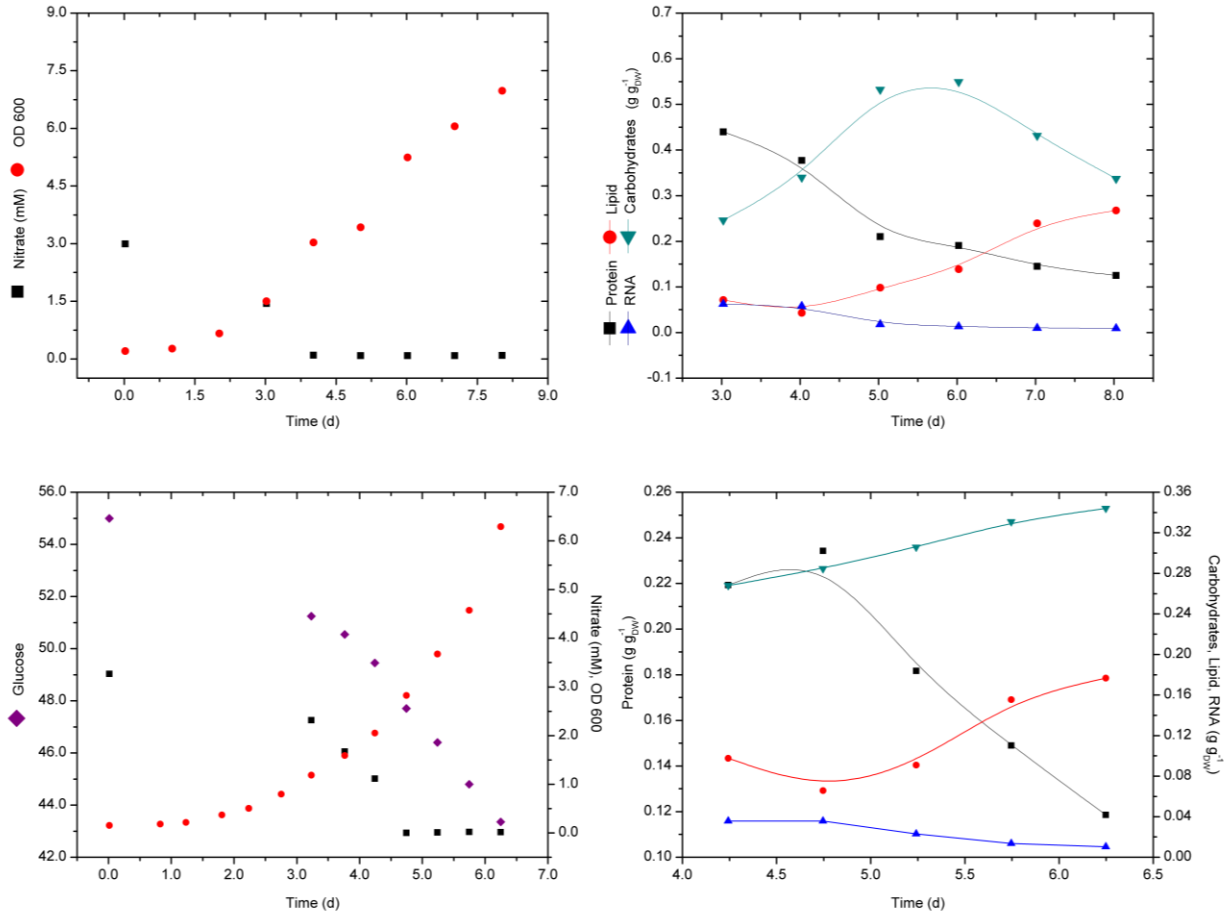

Figure S2. Experimental data for *C. vulgaris* UTEX 395. A) Time-course of nitrate uptake (mM) and growth (OD600) under photoautotrophic conditions. B) Time-course of protein and RNA decrease and accumulation of carbohydrates and lipids under photoautotrophic conditions. C) Time-course of nitrate uptake (mM), growth (OD600) and glucose consumption (mM) under heterotrophic conditions. D) Time-course of protein and RNA decrease and accumulation of carbohydrates and lipids under heterotrophic conditions.

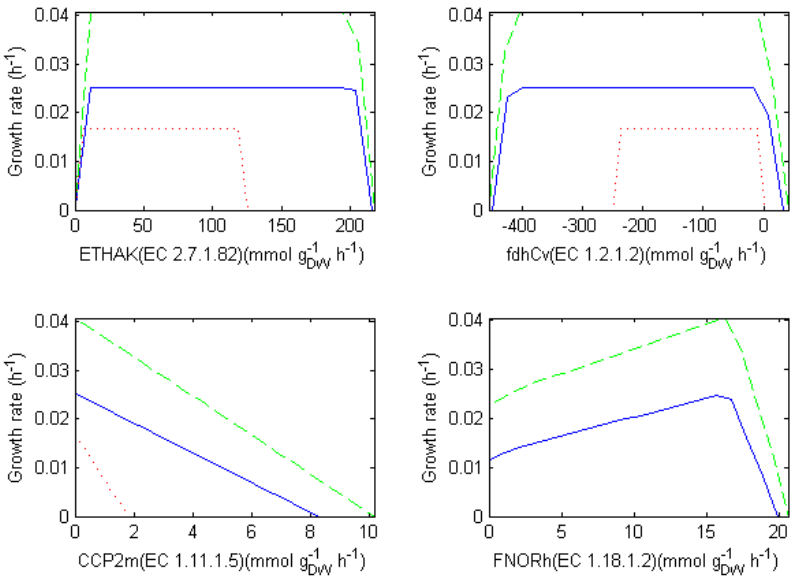

Figure S3. Robustness analysis. The predicted optimal growth rate of *C. vulgaris* under heterotrophic (red line), photoautotrophic (blue line), and mixotrophy (green line) conditions as a function of ethanolamine kinase (EC 2.7.1.82), formate dehydrogenase (EC 1.2.1.2), cytochrome C peroxidase (EC 1.11.1.5), and ferredoxin-NADP+ reductase (EC 1.18.1.2).

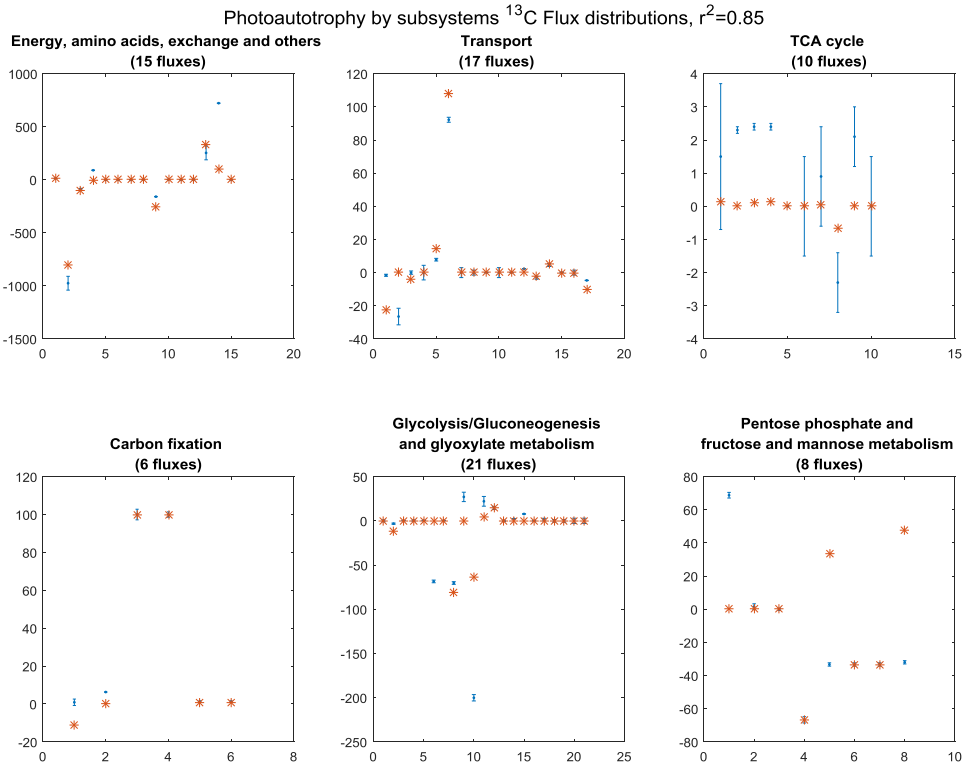

Figure S4. Comparison of the experimental flux distribution of *Chlorella protothecoides* and predicted data by *i*CZ843. The Pearson correlation (R^2) was determined 0.85 under photoautotrophy.

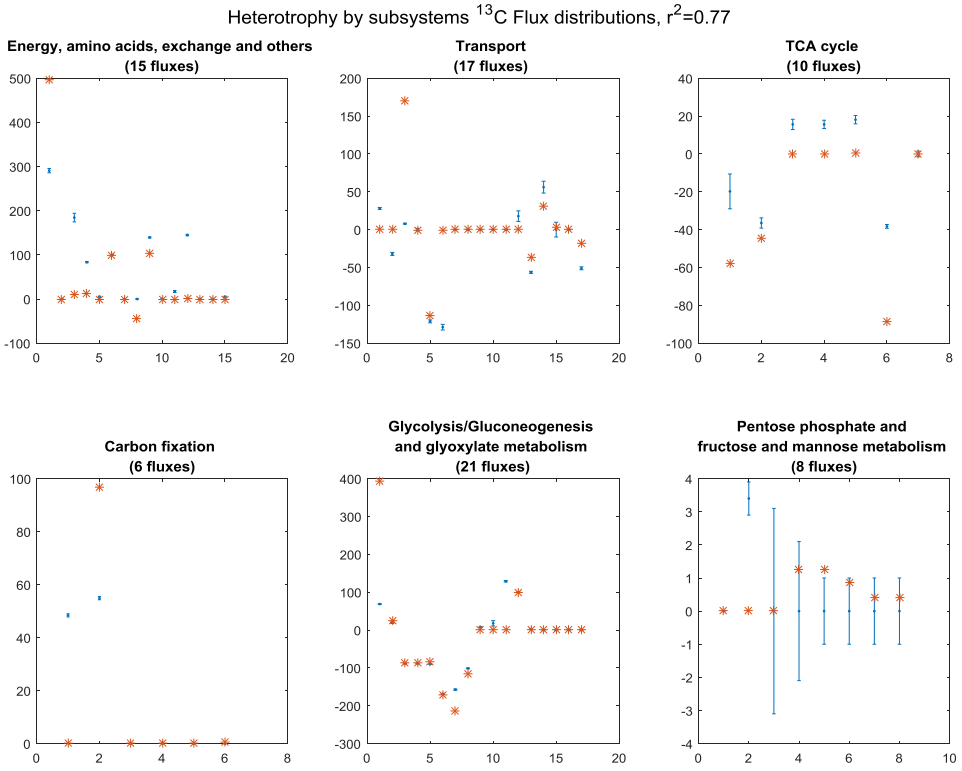

Figure S5. Comparison of the experimental flux distribution of *Chlorella protothecoides* and predicted data by *i*CZ843. The Pearson correlation (R^2) was determined 0.77 under heterotrophy.

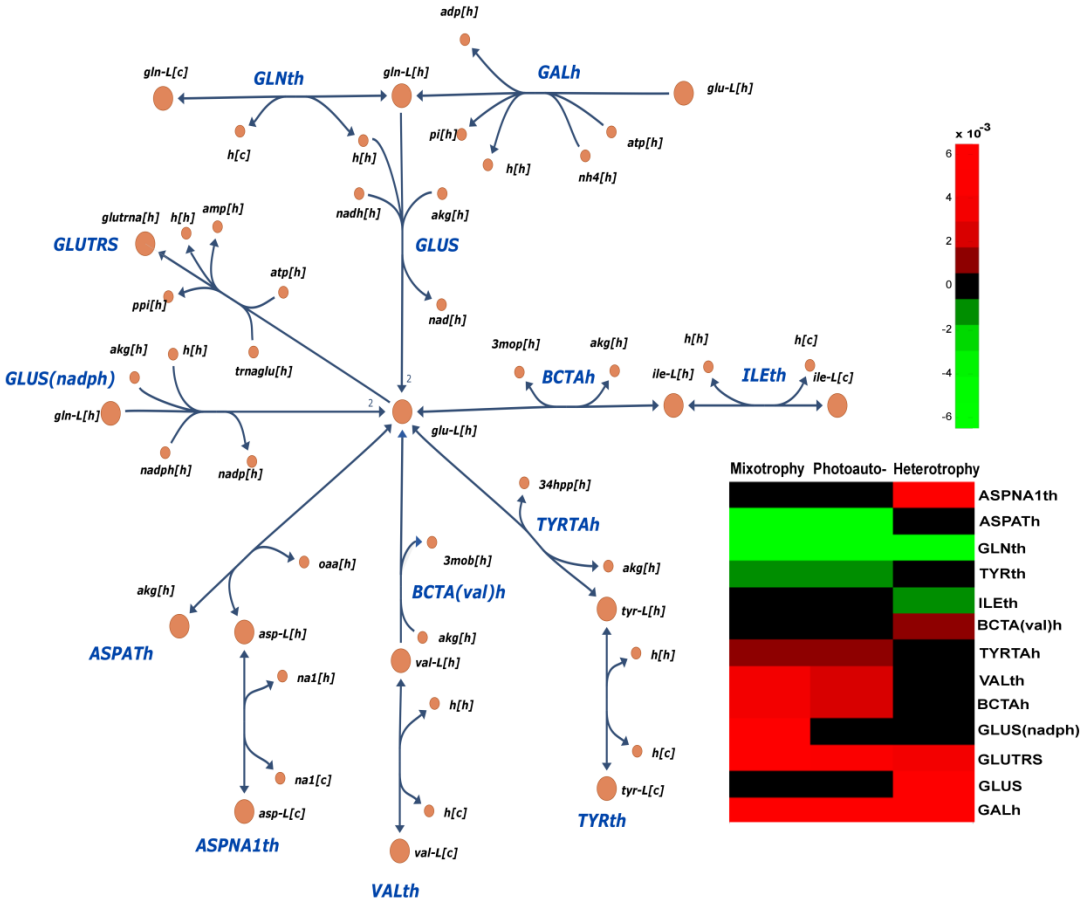

Figure S6. Flux distribution associated to glutamate metabolism in the chloroplast and cytoplasm. For abbreviation of reactions and metabolites see Table S1 and S2.

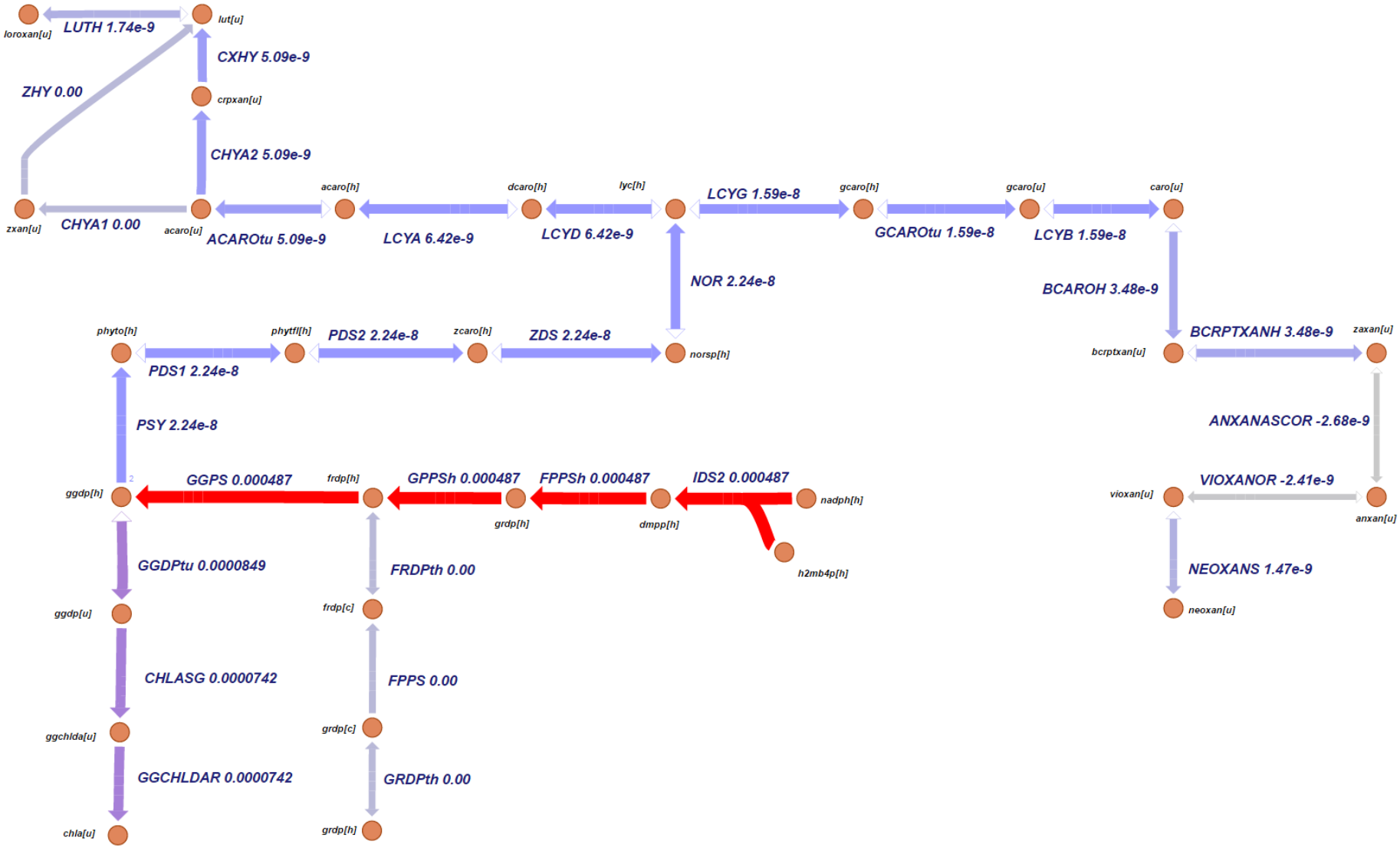

Figure S7. Flux distribution of pigment biosynthesis pathways under heterotrophic growth.

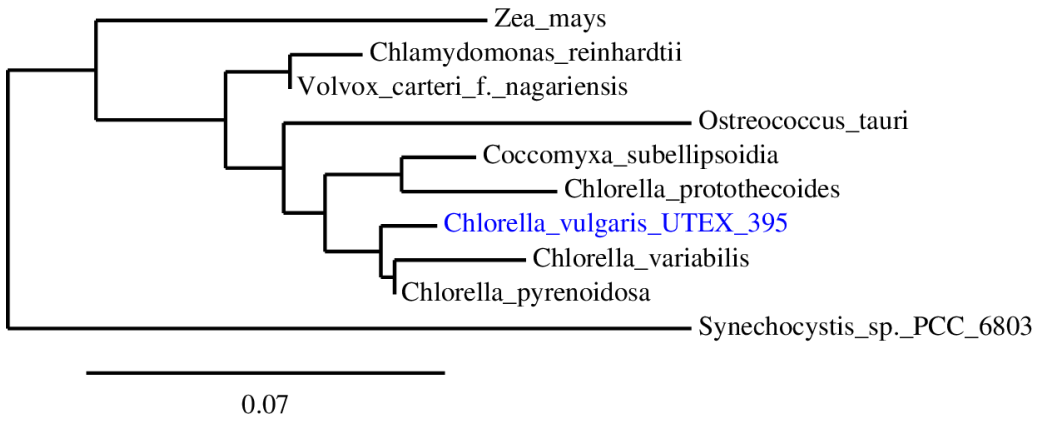

Figure S8. Arbitrarily rooted phylogenetic tree based on multiple sequence alignments of the big subunit of RuBisCO (EC 4.1.1.39) inferred using Phylogeny.fr.

Supplementary text

Refinement and gap filling

Enzymes such as diphosphate phospho-alpha-D-ribosyl-transferase (ADPART EC 2.4.2.17), chorismate synthase (CHRS EC 4.2.3.5), prephenate dehydratase (PPDH EC4.2.1.51), and threonine aldolase (THRA EC 4.1.2.5) were included in the model without gene association. In each case gene-associated neighbor enzymes were already in the model and the non-GPR reaction was indispensable in order to get a functional reconstruction. Additional non-GPR reactions related to nucleotides and mainly inositol metabolisms were included in *i*CZ843. Most of these enzymes are neither mentioned in the literature nor retrieved any hits in the Table 1 databases for *C. vulgaris*, as well as for other green algae. Although the inositol metabolism is not well known in green algae, several publications indicate the presence and functionality of this pathway based on enzyme activity and intermediate production (Nichols et al., 1967). The inositol metabolism is mainly involved in the synthesis of glycerophospholipids, cell wall and galactose.

The model contains GPR associations for the enzymes Δ-5-desaturase and ω-3/6-desaturase involved in the polyunsaturated fatty acids (PUFAs) biosynthesis; the Δ-desaturase transforms oleate to linoleate and ω-desaturase converts linoleate to α-linolenate. Following bioinformatic localization prediction algorithms and information obtained from other organisms/models, such as *C. rheinhardtii*/*i*RC1080 and *Saccharomyces cerevisiae/i*ND750 (Duarte and Herrg, 2004; Mo et al., 2009), these reactions were located in the chloroplast and cytoplasm, respectively. According to literature, the *de novo* fatty acid biosynthesis in green algae is predominantly localized in the chloroplast as well as in the non-green plastid. We split synthesis of 56 lipids in the cytoplasm, and the synthesis of 43 fatty acids such as MGDG, DGDG and SQDG in the chloroplast (Stitt and Heldt, 1981).

We were able to find gene evidence for enzymes, such as serine O-acetyltransferase (EC 2.3.1.30), histidiol-phosphatase (EC 3.1.3.15), 1-deoxy-D-xylulose-5-phosphate reductoisomerase (EC 1.1.1.267), and uroporphyrinogen III synthase (EC 4.2.1.75), opposed to what was reported for other Chlorella reconstructions of core metabolism (Wu, 2014). Predicted gene models of *C. vulgaris* often aligned with the green algae *Micromonas* sp. and *Volvox carteri* (Worden et al., 2009; Prochnik et al., 2010).

Studies about the transport of acetyl-CoA showed that it cannot cross the chloroplast membrane (Roughan et al., 1979), however, this essential fatty acid precursor needs to be generated in the chloroplast and then transported. We found a strong gene homology with the transporter of acetyl-CoA:CoA antiporter 1 (TC 2.A.1.25.1) and therefore incorporated this transporter into the model. The enzyme PYRShi carries out the conversion of pyruvate and coenzyme A to acetyl-coA and CO2 using the ferredoxin electron transporter. This reaction was identified as essential for model consistency after the gap filling analysis under heterotrophy. PYRShi was assumed present in Chlorella vulgaris metabolism

because a stacked alignment analysis for pyruvate synthase showed homologous sequences in green algae such as Volvox carteri, Chlamydomonas reinhardtii and Chlorella variabilis (Noth et al., 2013).

Comparison with Chlorella core models

The existing core models for Chlorella represent mainly central carbon metabolism (see Table 2). We compared *in silico* growth rates using the same experimentally determined uptake rates from these models with *i*CZ843. Despite all Chlorella species belong to the same genus some differences in the metabolism can be found using a full genome-scale reconstruction. *C. protothecoides* can consume glucose at a rate of 0.3028 mmol $(g_{DW}h)^{-1}$ and it can grow at 0.0257 h^{-1} under heterotrophic conditions, with a specific growth rate of 0.0251 h^{-1} (Wu et al., 2015). Applying the same constraints, *i*CZ843 predicts a comparable growth rate of 0.0264 h^{-1}, which can be expected because *C. vulgaris* UTEX 395 and *C. protothecoides* are closely related species (see Figure S8). Figure 3B shows the growth rate under photoautotrophic growth, the blue line represents the simulation by *i*CZ843, the red circles and red diamond are experimental data from (Nascimento et al., 2015; Wu et al., 2015), respectively. The simulated growth rates have over 85% of fit. Mixotrophic growth of *C. pyrenoidosa* resulted in uptake rates of 5.28 $mmol_{CO2}$ $(g_{DW}h)^{-1}$ and 1.15 $mmol_{Glc}$ $(g_{DW}h)^{-1}$, that allow a growth rate of 0.066 h^{-1} (Yang et al., 2000). Under the same constraints the simulated growth rates by that model and *i*CZ843 were 0.653 h^{-1} and 0.524 h^{-1}. It is important to note that a metabolic genome-scale model allows more realistic and detailed analysis of the stoichiometry in the complete metabolic network. Aditionally, these models can be used to link the metabolic activity of reactions with omics data such as expression data through gene associations.

Experimental data and uptake rates calculations

C. vulgaris was not impeded to grow after nitrate's depletion, this fact occurred after 3.5 days for photoautotrophy and after 5 days under heterotrophic conditions (Figure S2 A,C). Instead, the culture continued to grow and ended the experiment with twice the total biomass as at the onset of nitrate depletion.

Before nitrogen starvation, the final biomass mass concentration was 0.84 and 1.31 $g_{DW}\,L^{-1}$ for photoautotrophy and heterotrophy, respectively, and the protein content increased up to 0.44 and 0.24 g g_{DW}^{-1}, respectively. Once in nitrogen starvation, the protein and RNA content decreased to similar values as during photoautotrophic and heterotrophic growth (0.009 and 0.010 g g_{DW}^{-1}, respectively), revealing that nitrogen availability affects both amino acid and nucleotide metabolism independent of the carbon and energy source. Nitrogen starvation resulted first in significant accumulation of carbohydrates to the two growth conditions (Figure S2 B,D). Once the carbohydrate content reached about 50% of dry weight under photoautotrophic growth, the production of lipids increased (days 5 to 6) and reached 33% at the end of the experiment (day 8). Concurrent with the lipid production, the carbohydrate content declined to 35% of dry weight.

During heterotrophic growth, the carbohydrate and lipid accumulation reached 35% (carbohydrates) and 17% (lipids) at approximately the same time. Both carbohydrates and lipids began to accumulate after protein degradation at day 4.

The specific rates for synthesis and degradation of carbohydrates were determined experimentally in order to constrain the model. Data were fitted to a Gompertz model (Wang and Zuidhof, 2004) with a R^2 value of 0.98 (see Supplementary Table S6).

The biomass composition under heterotrophy and photoautotrophy was measured using samples taken after day 6 of experiment, assuring that the highest lipids accumulation was reached.

Light uptake modeling

*i*CZ843 can simulate growth for 11 different light sources. The theoretical effective incident photon flux was used to constrain the model, depending of the selected light. The incidence of photons is a parameter that correlates the spectral photon irradiance with the wavelength range; it shows peaks for wavelengths at the highest photosynthetically active radiation. The constraint values were specified according to (Chang et al., 2013; http://www.mv.helsinki.fi/aphalo/photobio/lamps.html).

The parameter photosynthetic oxygen evolution is the main player determining the predicted growth rates under light conditions; this constraint limits the biomass production to 0.025 h^{-1} for CO_2 uptake rates over 8 mmol g_{DW}^{-1} h^{-1}. This asymptotic behavior was modeled using an oxygen demand reaction located in the thylakoid (DM_o2D(u)). The maximum oxygen evolution rate for *Chlorella vulgaris* is 21 $nmol_{O2}$ μg_{Chl}^{-1} h^{-1} (Kozłowska-Szerenos et al., 2004), corresponding to 8.31 mmol g_{DW}^{-1} h^{-1}, and experimentally we measured chlorophyll content of 396 mg g_{DW}^{-1}. Using these data we constrained the lower bound of the DM_o2D(u) reaction to 0 and the upper bound to the experimentally reported value of 8.31 mmol g_{DW}^{-1} h^{-1}. This value is similar to the one used for Chlamydomonas model (*i*RC1080) of 8.28 mmol g_{DW}^{-1} h^{-1}.

Literature cited in the Supplementary text

Duarte NC, Herrg MJ (2004) Reconstruction and Validation of. Genome Res 1–12

Kozłowska-Szerenos B, Bialuk I, Maleszewski S (2004) Enhancement of photosynthetic O2 evolution in Chlorella vulgaris under high light and increased CO2 concentration as a sign of acclimation to phosphate deficiency. Plant Physiol Biochem **42**: 403–409

Nascimento IA, Cabanelas-Dominguez IT, Nunes JD, Nascimento MA, Sousa L, Sansone G (2015) Biodiesel yields and fuel quality as criteria for algal-feedstock selection: Effects of CO2-supplementation and nutrient levels in cultures. Algal Res **8**: 53–60

Nichols BW, James a T, Breuer J (1967) Interrelationships between fatty acid biosynthesis and acyl-lipid synthesis in Chlorella vulgaris. Biochem J **104**: 486–496

Noth J, Krawietz D, Hemschemeier A, Happe T (2013) Pyruvate:Ferredoxin Oxidoreductase Is Coupled to Light-independent Hydrogen Production in Chlamydomonas reinhardtii. J Biol Chem **288**: 4368–4377

Prochnik SE, Umen J, Nedelcu AM, Hallmann A, Miller SM, Nishii I, Ferris P, Kuo A, Mitros T, Fritz-laylin LK, et al (2010) Genomic Analysis of Organismal. Science (80-) **329**: 223–226

Roughan PG, Mudd JB, McManus TT, Slack CR (1979) Linoleate and alpha-linolenate synthesis by isolated spinach (*Spinacia oleracea*) chloroplasts. Biochem J **184**: 571–4

Stitt M, Heldt HW (1981) Physiological rates of starch breakdown in isolated intact spinach chloroplasts. Plant Physiol **68**: 755–61

Wang Z, Zuidhof MJ (2004) Estimation of growth parameters using a nonlinear mixed Gompertz model. Poult Sci **83**: 847–852

Worden AZ, Lee J-H, Mock T, Rouze P, Simmons MP, Aerts AL, Allen AE, Cuvelier ML, Derelle E, Everett M V., et al (2009) Green Evolution and Dynamic Adaptations Revealed by Genomes of the Marine Picoeukaryotes Micromonas. Science (80-) **324**: 268–272

Wu C, Xiong W, Dai J, Wu Q (2015) Genome-based metabolic mapping and 13 C flux analysis reveal systematic properties of an oleaginous microalga Chlorella protothecoides. Plant Physiol **167**: 586–599

Yang C, Hua Q, Shimizu K (2000) Energetics and carbon metabolism during growth of microalgalcells under photoautotrophic, mixotrophic and cycliclight-autotrophic/dark-heterotrophic conditions. Biochem Eng J **6**: 87–102