Title

Insights into mechanisms of high productivity and cultivation flexibility resulting from physiological and omic analyses of the microalga Marinichlorella kaistiae KAS603

Author

Sanchez-Alvarez, Eva L.

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Insights into mechanisms of high productivity and cultivation flexibility resulting from physiological and omic analyses of the microalga *Marinichlorella kaistiae* KAS603

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Marine Biology

by

Eva Luz Sánchez Alvarez

Committee in charge:

Mark Hildebrand, Chair
William Gerwick
Brian Palenik
Julian Schroeder
José Luis Stephano Hornedo
Victor Vacquier

2017

The Dissertation of Eva Luz Sánchez Alvarez is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

_____
Chair

University of California, San Diego

2017

DEDICATION

This dissertation is dedicated to my son, Guillermo Salomón Sánchez Alvarez, for your love, for being you, and for putting up with my crazy ideas and mood swings. I truly admire the wonderful person you have become and I am sure that only great things await you. With love and admiration from your mom.

EPIGRAPH

“Sometimes the questions are complicated and the answers are simple.”

Dr. Seuss

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Mark Hildebrand for his belief, encouragement, and mentoring. Mark always kept his office door open, with a welcoming spirit for interaction and healthy discussions. He is a natural teacher, a great friend, and a dedicated mentor. He is incredibly knowledgeable, but always ready to learn. He always knows where to find interesting and fun things, but is prepared for the long, careful observation it sometimes takes to see them. He showed me that passion, good science, and effective work are inextricably linked. Mark, it has been a great honor being under your guidance.

I'd also like to thank my committee: William Gerwick, Brian Palenik, Julian Schroeder, José Luis Stephano, and Victor Vacquier. William provided a new perspective to my work. Brian provided balance and strength. Julian provided strong support. José provided constructive criticism, and Victor provided great comments and crucial guidance as I move into the world of teaching. Finally, a big thank you to my SIO family, for a long string of valuable advise inside and outside of this dissertation. To students everywhere!

Chapter 1, in full, is a reprint of the material as it appears in Algal Research 2017. Sánchez-Alvarez, Eva L.; González-Ledezma, Grisele.; Bolaños-Prats, José A.; Stephano-Hornedo, José L., and Hildebrand, Mark. "Evaluating *Marinichlorella kaistiae* KAS603 cell size variation, growth and TAG accumulation resulting from rapid adaptation to highly diverse trophic and salinity

cultivation regimes", Algal Research, 25:12-24, 2017. The dissertation author was the primary investigator and author of this material.

Chapters 2 and 3 contain unpublished work with co-authors Mark Hildebrand, Matteo Pellegrini, David López, Jing Lu, Megan Inkeles, Weihong Yan and Shawn Cokus. The dissertation author was the primary investigator and author of chapters 2 and 3.

VITA

1998	Bachelor of Science, Universidad Autónoma de Baja California
2009	Master of Science, Universidad Autónoma de Baja California
2015	Master of Science, University of California, San Diego
2017	Doctor of Philosophy, University of California, San Diego

ABSTRACT OF THE DISSERTATION

Insights into mechanisms of high productivity and cultivation flexibility resulting from physiological and omic analyses of the microalga *Marinichlorella kaistiae* KAS603

by

Eva Luz Sánchez Alvarez

Doctor of Philosophy in Marine Biology

University of California, San Diego, 2017

Mark Hildebrand, Chair

Microalgae are essential primary producers and CO_2 sequesters in aquatic environments. They represent a bioenergy source in the global energy system and can be exploited as feedstock for biofuel, food supplements, vaccines, wastewater treatment, fertilizers, bioplastics, and as part of urban infrastructure in architectural designs of eco-communities. Microalgal productivity can be favored by optimizing nutrients in cultivation systems and by using environmentally flexible algal strains. Describing the potential mechanisms of high (TAG and

biomass) productivity and cultivation flexibility followed by microalgal strains can improve basic phycology knowledge as well as algal management in industry. The green unicellular alga, *Mk* KAS603 is a strain that can thrive in freshwater, seawater and hypersalinity. With specific amounts of sodium bicarbonate, sodium acetate and sodium chloride added to cultivation systems, the strain is able to accumulate biomass and TAG. In a pilot plant, hypersalinity culturing of *Mk* KAS603 improves productivity mostly due to diminishing growth of invading organisms. As *Mk* KAS603 grows in different trophic and cultivation conditions, it changes its multiple fission reproductive mode, implicating changes in cell size, cell growth rate and mitotic rounds that is reflected in biomass and TAG accumulation. Through the combination of physiological, genomic and transcriptomic data analysis we were able to predict metabolic changes in different cell compartments in *Mk* KAS603 while being cultivated in different conditions. The most relevant predictions were the possible accumulation of starch in two compartments: cytoplasm and chloroplast; the possible synthesis of osmolytes (glycerol during log phase and glycine and serine during stationary phase) and the shutdown of FA beta oxidation in order to accumulate FA-TAG, in hypersalinity. Our work is aimed at providing insights into algal metabolism and productivity, and we propose that the capacity of *Mk* KAS603 to thrive in different cultivation and trophic environments is due to a coordinated regulation of energy and carbon flux in different cellular compartments while modifying its multiple fission division mode.

Introduction

Microscopic algae are photosynthetic organisms that are efficient at capturing CO_2 and converting it to a variety of carbon compounds. Cellular constituents to enable division are made from this carbon, as well as carbon storage compounds as carbohydrate, such as starch, or fatty acids in the form of triacylclycerols (TAGs) contained in oil droplets. The latter can be extracted from the cells and converted into biofuel through a process called transesterification, in which fatty acids are detached from the glycerol backbone of TAGs and converted to fatty acid methyl esters (FAMEs). The advantage of biofuel is that it is renewable, easily biodegradable and when combusted it emits less greenhouse gases and toxins than fossil fuel (Reda *et al*, 2010). An increase in TAG usually leads to a decrease in algal cell growth, affecting the overall TAG yield in a culture (Sharma *et al*, 2013).

Carbon processing and carbohydrate storage in microalgae are essential to productivity and vary among algal classes. The current knowledge is that the form of carbohydrate and storage location has varied throughout evolutionary history (Hildebrand *et al*, 2013). One possible explanation is that it allows optimization of carbon flux. For example in the model green alga *Chlamydomonas reinhardtii*, starch storage and breakdown occur in the chloroplast. These are competing processes that require tight regulation as well as temporal separation of storage and breakdown processes. This arrangement can lead to inefficiency in processing carbon and restrictions on the time of day when division can occur, which could negatively affect overall productivity. A working strain with carbohydrate storage (and processing) in

more than one compartment might facilitate carbon flow in competing pathways and augment productivity.

Algae are usually environmentally specific, and can be isolated from a variety of natural aqueous habitats ranging from freshwater to brackish water, marine and hypersaline environments and soil (Round, 1984), and can be easily affected if transferred to a new cultivation condition. Environmental flexibility of a strain can be useful in a production setting, because high yield could be maintained under diverse seasonal or cultivation conditions. Examining factors that enable culturing flexibility in a single species could provide clearer insights than when comparing different species because it will reduce interspecies variability. This work is focused on characterizing an environmentally flexible green microalgae, *Marinichlorella kaistiae* KAS603 (*Mk* KAS603) under a diverse variety of cultivation conditions where biomass accumulation and TAG yield are induced (including conditions where TAG accumulation occurs during growth). *Mk* KAS603 divides by multiple fission, where one cell grows and divides multiple times producing various daughter cells that are held captive inside the original mother cell until it ruptures and releases the daughter cells. Such process seems beneficial for the algal strain because smaller cells are more productive: with a large surface area to volume ratio, the daughter cells can absorb more light, uptake more nutrients (even when they are scarce) and enhance biomass.

The overall goal of this thesis was to take advantage of the cultivation flexibility of *Mk* KAS603 to investigate cellular mechanisms which allow production of high amounts of TAG and environmental adaptation. The first chapter describes

phenotypic traits of *Mk KAS603* under highly diverse cultivation conditions (autotrophic, mixotrophic, freshwater, saline and hypersaline) during log and stationary phases of growth. These traits include TAG accumulation and cell growth life cycle stages. We document that this strain can be adapted from freshwater to marine and hypersalinity, that it can use bicarbonate or acetate for growth and TAG accumulation (and hypersalinity adaptation for TAG accumulation without decreasing growth), and that such changes in productivity are linked to changes in size during its multiple fission division mode. The second chapter introduces the genomic sequence determination and characterization of *Mk KAS603*, it analyses mitochondria, chloroplast and nuclear genomes as well as predicted protein targeting into different cellular compartments. A major finding was the prediction of starch accumulation in both the chloroplast and cytoplasm in *Mk* KAS603. We also compared *Mk* KAS603 to other green microalgae (*Chlamydomonas reinhardtii*, *Coccomyxa ellipsoidea*, *Scenedesmus obliquus, Chlorella sorokiniana,* and *Chlorella variabilis* NC64A), and observed that non-oleaginous strains accumulate starch in the chloroplast and oleaginous strains are predicted to accumulate starch in the chloroplast, plus either the ER or cytoplasm. Chapter 3 analyzes transcript changes in *Mk KAS603* during log and stationary phases of growth when cultivated with organic and/or inorganic carbon supplementation and in low or high salinities. From such analyses we were able to predict changes in carbon flow related to productivity and multiple fission division in *Mk* KAS603. The major predicted changes are activation of multiple anabolic and catabolic routes while growing mixotrophically, synthesis of osmolytes in seawater

and hypersalinity and the shutdown of fatty acid catabolism to favor fatty acid and TAG accumulation in hypersalinity.

In summary, this work combined physiological data with genomics and transcriptomics analysis to render explicit knowledge about *Mk KAS603* cellular and metabolic mechanisms that trigger TAG accumulation, and provide the ability to adapt to different trophic and culture conditions. Through protein targeting analysis comparison of *Mk* KAS603 and other green algae, we have predicted that pathways such as glycolysis, gluconeogenesis, fatty acid and TAG synthesis can take place in different cell compartments. Our hypothesis for why this is the case is that compartmentation minimizes the occurrence of competing reactions, which improves overall efficiency and productivity. Our findings also show that specific amounts of bicarbonate and/or acetate supplementation can enhance growth and TAG yield (and that hypersalinity can enhance TAG yield without compromising growth), and modify multiple fission division mode. We propose that the capacity of *Mk* KAS603 to thrive in different cultivation and trophic environments is due to a coordinated regulation of energy and carbon flux in different compartments in the cell while modifying its mode of division related to changes in mitotic rounds reflected in the proportion of daughter and mother cells.

REFERENCES FOR THE INTRODUCTION

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

Reda AI, Shanab A, Jeon BH, Song H, Kim Y, Hwang JH (2010) Algae-biofuel: potential use as sustainable alternative green energy. Online J Power Energy Eng 1(1):4–6

Round FE (1984) The ecology of algae. Cambridge University Press, Cambridge. ISBN-10: 0521269067

Sharma P, Tandon GD, Khetmalas MB (2013) Biofuels from green microalgae. R. K. Salar *et al* (eds.) Biotechnology: Prospects and Applications. Springer India. (9):95–112

Chapter 1

Evaluating *Marinichlorella kaistiae* KAS603 cell size variation, growth and TAG accumulation resulting from rapid adaptation to highly diverse trophic and salinity cultivation regimes

Algal Research 25 (2017) 12–24

Contents lists available at ScienceDirect

Algal Research

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

Evaluating *Marinichlorella kaistiae* KAS603 cell size variation, growth and TAG accumulation resulting from rapid adaptation to highly diverse trophic and salinity cultivation regimes

Eva L. Sánchez-Alvarez[a], Grisele González-Ledezma[a], José A. Bolaños Prats[b], José L. Stephano-Hornedo[b], Mark Hildebrand[a,*]

[a] Marine Biology Research Division, Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0202, USA
[b] Laboratorio Meredith Gould, Facultad de Ciencias, Universidad Autónoma de Baja California, Campus Ensenada, Km 103 Carretera Tijuana-Ensenada, 22860 Ensenada, B.C., Mexico

ARTICLE INFO

Keywords:
Marinichlorella kaistiae KAS603
TAG
Mixotrophy
Hypersalinity
Multiple fission
Mother cell

ABSTRACT

The ability of a microalgal species to adapt to changes in cultivation environment is likely to be beneficial for a successful biofuel/bioproduct production system, because the species could maintain high yields under diverse seasonal or cultivation conditions. Examining factors that enable culturing flexibility in a single species could provide clearer insights than when comparing different species because it will reduce interspecies variability. *Marinichlorella kaistiae* KAS603 is an excellent model to evaluate mechanisms involved in adaptation to different cultivation regimes. We have studied cell growth, size, triacylglycerol (TAG) accumulation and life cycle stages in this multiple fission dividing strain under a wide variety of conditions, ranging from autotrophic growth in freshwater to mixotrophic growth in fresh and seawater, and to autotrophic growth under saline and hypersaline conditions. Such conditions influence the division mode of the strain, which is linked to biomass and TAG yield. Based on lab and pilot plant experiments, we have discovered that the fastest TAG accumulation takes place under mixotrophy in freshwater, highest yield (culture density plus TAG) condition occurs under mixotrophy in sea water, and the best outdoor culture condition to achieve growth with fewer invasive species is hypersaline natural seawater. In addition to characterizing growth and TAG accumulation characteristics under a wide variety of cultivation conditions, this work sets the stage for investigation into the mechanisms that enable diverse cultivation adaptations, and contribute to the development of this environmentally flexible microalga as a production feedstock.

1. Introduction

Most microalgal strains are environmental specialists, and are not necessarily environmentally flexible. Understanding cellular factors enabling environmental adaptation can be challenging when comparing different species, since equivalent cultivation conditions comparing species may not be possible. *Marinichlorella kaistiae* strain KAS603 has a remarkable ability to adapt to different cultivation conditions, and for growth under different trophic modes. It has been a model production development organism due to its productivity and ability to accumulate lipid [1,2]. Examining a single strain such as KAS603 could simplify understanding the mechanisms of cultivation flexibility by allowing a direct comparison of physiological differences relative to control conditions in the same organism. This information could be used to enable optimization of those traits in other strains, leading to an overall improvement of the technology.

Triacylglycerols (TAGs) are important storage products in microalgae, which can be converted into a bio-product, such as biofuel. In such case TAG accumulation and the rate at which they are accumulated is relevant, along with maximum cellular yield acquisition. In a production system, reduced time spent in culturing will translate into increased productivity and reduced cost.

Cell size is also very important for culture productivity and maintenance. Phytoplankton growth requires the acquisition of nutrients and their assimilation, or conversion into biomass, to form new cells [3]. In a similar manner to other members of the Chlorellaceae family, *M. kaistiae* KAS603 undergoes a multiple fission reproduction mode, which involves different cell sizes. A small “free” cell, called autospore or daughter cell, grows in size (2 fold volume increase) and then will divide into two cells that are kept captive inside the original

* Corresponding author.
E-mail address: mhildebrand@ucsd.edu (M. Hildebrand).

http://dx.doi.org/10.1016/j.algal.2017.03.027
Received 27 July 2016; Received in revised form 26 February 2017; Accepted 30 March 2017

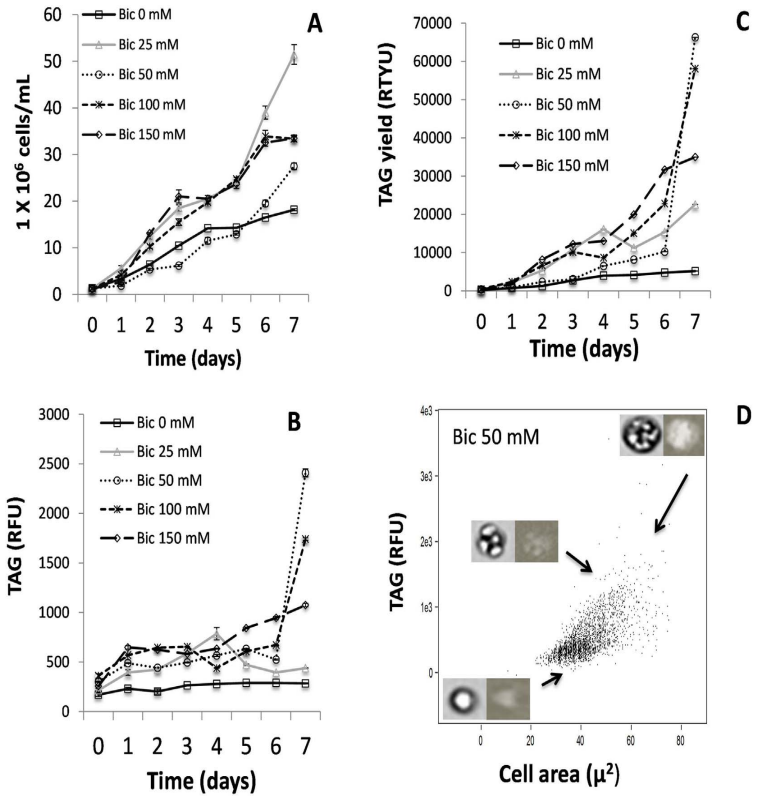

Fig. 1. *Marinichlorella kaistiae* KAS603 response to different molar concentrations of bicarbonate in freshwater medium. Cells were cultured in freshwater modified Shuisheng-4 medium (MS4) with increasing sodium bicarbonate (denoted as Bic) concentrations (mM). A) Culture density response with different concentrations of bicarbonate, B) TAG accumulation response, C) TAG yield (culture density x relative TAG content per cell) response, D) TAG accumulation response relative to cell area of a population with 50 mM bicarbonate on day 6 of cultivation. Cell images inside squares are bright field (left) and BODIPY (TAG) fluorescence (right). Black arrows indicate the place on scatter plot where cells in square images are located. RFU are relative fluorescence units. RTYU are relative TAG yield units, as defined in Materials and methods. The error bars indicate standard error (n = 3).

cell, which is now called a mother cell or autosporangium [4,5]. Division can continue within the mother cell and up to 32 (captive) daughters can be generated. The size of the mother cell is proportional to the number of captive daughters, and the mother cell volume can increase about 4 fold [4,5,6]. The differential cell sizes are another adaptation that enables cultivation flexibility and higher productivity for *M. kaistiae* KAS603 because it can accumulate different amounts of biomass under different nutrient conditions due inherent variations of surface area to cell volume ratio. For example, small daughter cells have higher surface area to cell volume ratio than large daughter cells, which favors nutrient acquisition as well as rapidity of biomass accumulation.

The main goal of our work was to perform a comparative analysis of growth and TAG accumulation in relation to life cycle-based cell size changes in different culture media, in relation to the adaptability of *M. kaistiae* KAS603. Such information should provide insight into underlying physiological and metabolic changes involved in adaptation and productivity. To achieve this goal, we cultured the strain under a wide variety of conditions: autotrophic and mixotrophic growth in freshwater and seawater, and autotrophic growth under saline and hypersaline conditions. We measured cell growth, TAG accumulation, chlorophyll and monitored population cell size tendencies per condition with imaging flow cytometry analysis. We selected the top 4 cultivation conditions and discovered that TAG accumulation is affected by life cycle stage changes during multiple fission division. Our work demonstrates the high adaptability of *M. kaistiae* KAS603 and provides insights into its growth, TAG accumulation, and life cycle responses under diverse cultivation conditions.

2. Materials and methods

2.1. Microalgal strain and cultivation conditions

Wild type *Marinichlorella kaistiae* strain KAS603 provided by Kuehnle AgroSystems, Inc. in Hawaii, was adapted to grow under several culture conditions (biological triplicates were performed for each culture condition tested), from starting liquid stocks cultured in freshwater modified Shuisheng-4 (MS4) medium [7], consisting of (per liter) 0.025 g $CaCl{\cdot}H_20$; 0.1 g $MgSO_4$; 0.033 g $NH_4H_2PO_4$; 0.129 g KCl; 0.250 g $(NH_2)_2CO$; 15 g $NaHCO_3$; 3.15 mg $FeCl_3$; 0.309 mg H_3BO_3; 0.169 mg $MnSO_4{\cdot}H_2O$; 0.0288 mg $ZnSO_4{\cdot}7H_2O$; 0.0618 mg $(NH_4)_6Mo_7O_{24}{\cdot}4H_2O$; 0.0146 mg $Co(NO_3)_2{\cdot}6H_2O$; 0.0016 mg $VOSO_4{\cdot}6H_2O$; 0.0474 mg $AlKO_8S_2{\cdot}12H_2O$; 0.0198 mg $(NH_4)_2(SO_4)_2{\cdot}6H_2O$; 0.0154 mg $Cd(NO_3)_2{\cdot}4H_2O$; 0.0040 mg $CrN_3O_9{\cdot}9H_2O$; 0.0033 mg $Na_2O_4W{\cdot}2H_2O$; 0.0019 mg KBr; 0.0083 mg KI; 0.0125 mg $CuSO_4{\cdot}5H_2O$. The medium was adjusted to pH 10.5 prior to autoclaving and remained constant throughout the duration of the experiments. Algal cells were cultivated for 3 days (log phase) with orbital shaking at 150 rpm under continuous cool fluorescent light intensity of 150 μmol photons $m^{-2}\,s^{-1}$, and a temperature of 25 °C, in

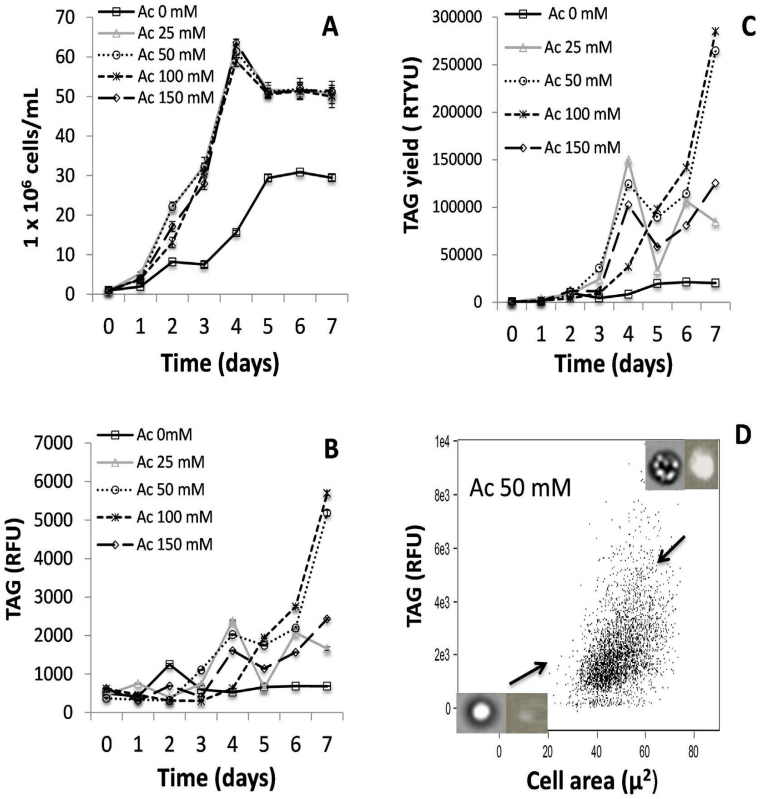

Fig. 2. *M. kaistiae* KAS603 response to different molar concentrations of acetate in freshwater cultures. Cells were cultured in MS4 medium with increasing sodium acetate (denoted as Ac) concentrations (mM). A) Culture density response, B) TAG accumulation response, C) TAG yield response, D) TAG accumulation response relative to cell area of a population with 50 mM acetate on day 6 of cultivation. Cell images inside squares are bright field (left) and BODIPY (TAG) fluorescence (right). Black arrows indicate the place on scatter plot where cells in square images are located, RFU are relative fluorescence units. RTYU are relative TAG yield units. The error bars indicate standard error (n = 3).

125 mL Erlenmeyer flasks containing 50 mL of MS4 medium (containing 178 mM $NaHCO_3$). Algal cells in exponential growth phase were harvested, pelleted and used as inocula (to grow immediately) for freshwater (MS4 with 0 to 150 mM sodium bicarbonate or 0 to 150 mM sodium acetate), and seawater (ASW, 0.4 M NaCl with 0 or 50 mM sodium acetate) batch cultures. From seawater (ASW, 0.4 M NaCl) they were adapted to hypersalinity (ASW, 1.5 M NaCl) during a three-month period. In seawater and hypersaline conditions, a constant pH 8.5 was maintained. All flask cultures were grown with ambient aeration. Culture density was measured by hemocytometer counts, performed every 24 h. Specific growth rate was determined among conditions and time points, using the following equation [8,9]:

$$\mu = \ln\ (X_2) - \ln\ (X_1)/(t_2 - t_1)$$

in which X_1 and X_2 are the mean cell numbers of *M. kaistiae* KAS603 at the times t_1 and t_2 respectively. Since t was expressed in days, growth rate was converted to division or doublings per day by dividing μ by the natural log of 2 = 0.6931 [9]

$$K = \mu/0.6931$$

2.2. Imaging flow cytometry analysis

1×10^6 cells were harvested for each condition every 24 h, pelleted in microfuge tubes, the supernatant was removed and the pellet immediately frozen at − 20 °C. At the end of each trial, all pellets were thawed, resuspended in 1 mL of potassium phosphate buffer (0.1 M, pH 7) for freshwater cultures, or 1 mL of 2.3% NaCl solution for seawater cultures (or 7.0% NaCl solution for hypersaline adapted cultures), and stained with 2.6 μL of a 1 mg·mL^{-1} BODIPY (4,4-difluoro-3a,4a-diaza-*s*-indacene Invitrogen, USA) stock dissolved in DMSO. Stained cells were maintained at room temperature for 15 min, put on ice and processed (10,000 to 20,000 cells per run) through an Amnis ImageStream X imaging flow cytometer. IDEAS™ software was used to analyze the data in order to obtain a final measurement of mean BODIPY fluorescence per cell (indicating the relative TAG content per cell - reported as relative fluorescence units [RFU]), mean chlorophyll autofluorescence per cell (to indicate relative chlorophyll content per cell), and cell area population distribution patterns (to indicate cell size variability and tendency per treatment/population). The number and % of mother and daughter cells was calculated by measuring cell area of thousands (10,000 to 20,000) of individual cells per population and per condition in log and stationary phase using imaging flow cytometry. Frequency histograms of cell area population distribution were created per condition, and microscopic images from such histograms were examined to determine cell division/size tendencies among daughters and mothers.

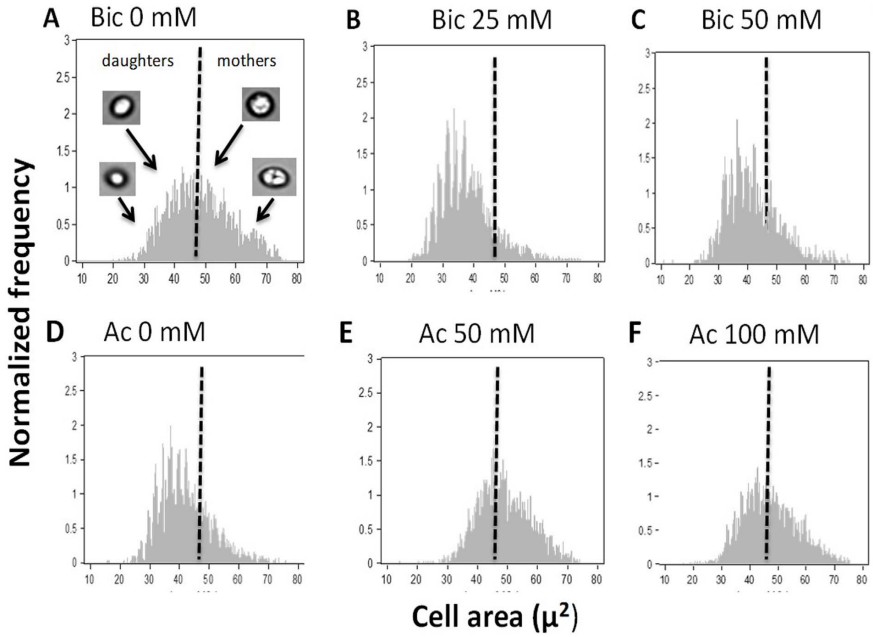

Fig. 3. Cell area population distribution comparison under bicarbonate and acetate cultivation in freshwater. Histograms represent the cell area population frequency distribution during day 6 of culturing among cells cultured in absence (A and D) or presence of sodium bicarbonate (B = 25 mM and C = 50 mM) or sodium acetate (E = 50 mM and F = 100 mM) in MS4 medium. The mean cell area for A is 48.5 μ^2, for B is 37.9 μ^2, for C is 41.9 μ^2, for D is 41.8 μ^2, for E is 48.8 μ^2 and for F is 47.2 μ^2. Arrows in A indicate specific size for daughter or mother cells (images inside squares). Black dotted vertical line separates daughter cells (left side) from mother cells (right side).

2.3. TAG yield

M. kaistiae KAS603 TAG yield under all previously mentioned culture conditions was calculated as:

$$TY = (Cd) * (RTC)$$

In which, TY represents relative TAG yield, Cd represents culture density, and RTC is the relative TAG content per cell. TY was expressed in relative TAG yield units (RTYU).

2.4. Outdoor cultures at a pilot plant

M. kaistiae KAS603 low saline (ASW, 0.4 M NaCl) and hypersaline (ASW, 1.2 M NaCl) adapted lab cultures (50 mL), were up-scaled to 18 L × 2, transported from La Jolla, CA to Ensenada, Baja California; México, and cultivated at pilot plant scale inside a greenhouse at the Universidad Autónoma de Baja California (UABC), during July–August 2013. Cells were adapted to grow in open tanks (100 L volume) during a four-week period; after that, cultures were scaled up from 100 L to 1000 L. Continuous sterile air and agitation were provided by plastic tubing, directly immersed in the cultures. A constant pH of 8.5 was maintained throughout the experiment. Culture density was measured by hemocytometer counts, performed every 24 h. Live ciliates and rotifers were detected and counted from 1 mL subsamples fixed with neutral Lugol's solution using settling chambers and inverted microscopy. Relative percentage was calculated based on total population and size.

2.5. Biomass determination

Daily samples were collected from the pilot plant tank cultures, centrifuged, washed two times with distilled water, dried at 80 °C overnight and weighed (dry cell weight).

3. Results

3.1. Evaluation of growth, TAG accumulation and cell reproduction changes under diverse cultivation conditions

Our goal in this study was to evaluate the ability of *M. kaistiae* KAS603 to grow and accumulate TAG under a diversity of cultivation conditions, and to determine cell reproduction characteristics associated with these conditions. We tested growth and productivity parameters in this multiple fission dividing strain in terms of growth rate, maximum culture density, TAG accumulation, and cell size under different concentrations of carbon added to freshwater (MS4) or seawater (ASW), or under different amounts of NaCl in seawater (ASW). We further calculated the size and percentage of mother and daughter cells by measuring cell area of thousands of individual cells per population, using imaging flow cytometry. We also related TAG accumulation capacity per cell relative to cell population during onset of TAG accumulation.

Our initial cultivation condition was autotrophy in freshwater with a bicarbonate-containing medium [Modified Shuisheng-4 (MS4)]. As seen in Fig. 1A, cells cultured under a wide range of sodium bicarbonate amounts (0 mM to 150 mM), displayed two distinct growth patterns, with 0 and 50 mM having a slower growth rate and maximum biomass accumulation than with 25, 100 and 150 mM. On days 6 and 7, TAG accumulation (Fig. 1B) and TAG yield [described as the product of culture density times the relative TAG content per cell] (Fig. 1C) were maximum at 50 mM and 100 mM sodium bicarbonate, but were less at 150 mM bicarbonate. During onset of TAG accumulation (day 6/late

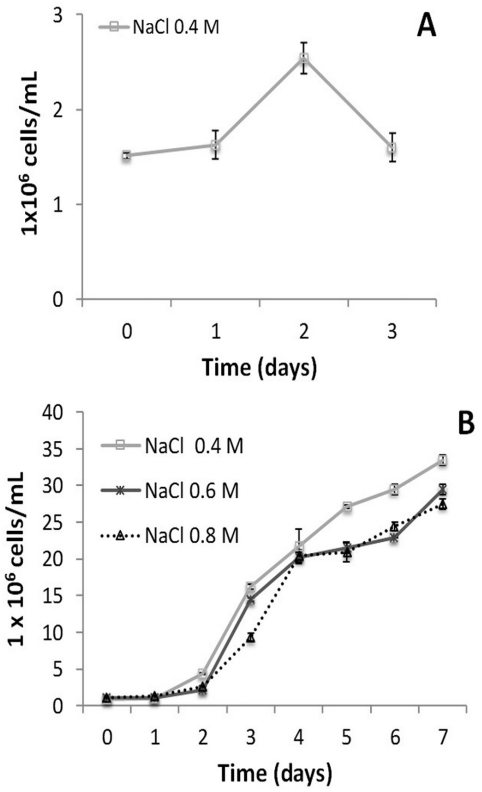

Fig. 4. Growth adaptation process of *M. kaistiae* KAS603 to increasing salinities. A. Initial growth at lowest salinity prior to adaptation. B. Intermediate two-month growth comparison of serial transfers into 3 different salinities. The error bars indicate standard error (n = 3).

log phase) at 50 mM bicarbonate, larger cells (mothers containing captive daughters) tended to accumulate more TAG (Fig. 1D). We measured the size distribution of cells to determine if different life cycle stages were enriched comparing different conditions. Examination of histogram plots enables identification of the percentage of cells in different life cycle stages of *M. kaistiae* KAS603, and changes in their distribution. Fig. 3A shows that different sized daughter cells (whole single cells) which are identified on the basis of lack of internal structure in the image, are prevalent below 47 μ^2, and different sized mother cells with captive daughter cells resulting in visible internal structure, are prevalent above that size. Such cut-off size of 47 μ^2 between mothers and daughters remained the same for all conditions tested. We examined size distributions comparing day 6 grown in 0, 25 and 50 mM bicarbonate. At 0 mM, a normal distribution in size with a trend towards larger sizes (Fig. 3A) was observed, and 55.8% of the total population corresponded to mother cells. At 25 mM, enrichment in daughter cells (86.9%) occurred, suggesting a more rapid release of daughters from mothers. At 50 mM bicarbonate, a similar shift to daughters (61.7%) is seen, but there is an increase in mothers (38.3%) relative to 25 mM (13.1%). The increase in mothers at 50 mM correlates with the higher TAG accumulation ability under these conditions (Fig. 1B–D). In summary, only a low amount (25 mM) of bicarbonate is needed to enhance growth and favors an early release of daughters. A medium amount (50 mM) of bicarbonate is needed to increase TAG accumulation mostly in mother cells with a later release of daughters. This concentration of bicarbonate inhibited growth however, which was restored at higher bicarbonate concentrations, with continued TAG accumulation ability at 100 mM, but a lack of TAG accumulation at 150 mM.

We next evaluated freshwater mixotrophic conditions in the presence of acetate. As shown in Fig. 2, cells were supplemented with various amounts of sodium acetate (the same molar concentrations as for sodium bicarbonate), in order to find the optimum concentration of an organic carbon source (acetate) in MS4 for high growth and TAG yield. Cell growth was enhanced analogously at all concentrations of acetate tested (Fig. 2A). TAG accumulation (Fig. 2B) and TAG yield (Fig. 2C) were maximal at 50 mM and 100 mM acetate (as they were at the same bicarbonate concentrations). During TAG accumulation (day 6/stationary phase) at 50 mM (Fig. 2B), cells accumulating the highest TAG tended to be larger than 50 μ^2, indicating the presence of larger mother cells (Fig. 2D). The size distribution assessment (Fig. 3) confirmed enrichment for mother cells cultured at 50 mM (61.9%) (Fig. 3E) and 100 mM acetate (63.3%) (Fig. 3F) on day 6, relative to 0 mM (28.6% mother cells) (Fig. 3D). In summary, acetate supplementation in a freshwater culture induced cell growth indistinctively of concentration, 50 and 100 mM acetate favored higher TAG yield, larger cells occurred for all concentrations except for 0 mM (during stationary phase), and mother cells tended to predominate during TAG accumulation at 50 mM and 100 mM, implying a late release of daughters.

We adapted *M. kaistiae* KAS603 to grow in salt water (ASW, 0.4 M NaCl), and then to hypersaline medium (ASW, 1.2 M NaCl), to evaluate its salinity tolerance, as well as TAG yield at different salinity conditions. Exponentially grown cells in MS4 (pH 10.5) were transferred to artificial seawater (ASW, pH 8.5) medium with a starting concentration of 0.4 M NaCl; rendering a slower growth than in MS4. Serial transfers into higher salinities (0.6 M NaCl - 0.8 M NaCl–1.2 M NaCl, pH 8.5) were made during 90 days, during which the cells adapted to the new conditions. Adaptability was evaluated through growth measurements until equivalent growth was demonstrated among low, medium and high NaCl concentrations, although growth in all NaCl conditions was lower than in MS4. During initial transfers into 0.4 M NaCl, cells were not able to reach log phase and growth tended to decline after 2 days (Fig. 4A), also cells tended to aggregate. After 60 days, cells grew without aggregation and were transferred (log phase) stepwise during a two-month period into 0.6 M NaCl and into 0.8 M NaCl, displaying slightly slower growth than at 0.4 M (Fig. 4B); until the 3rd month when they reached equivalent growth and were transferred (log phase) to 1.2 M NaCl. After 56 generations through all of the treatments, the cells were fully adapted from low (0.4 M) to high salt (1.2 M NaCl) in ASW. Once adaptation was confirmed, an 11-day trial was conducted in which cells were cultivated concurrently under NaCl 0.4 M, NaCl 0.8 M and NaCl 1.2 M. As shown in Fig. 5A, similar growth rate was displayed among all salinities. The cells achieved a maximum of 1.5 doublings per day. Growth in 1.2 M NaCl resulted in a 2.7 fold increase, compared to the lowest salinity, in TAG accumulation (Fig. 5B), which occurred in stationary phase. Highest TAG yield also occurred under this condition (Fig. 5C). A narrow cell size (area) range was apparent under hypersalinity during TAG accumulation (day 9/stationary phase) (Figs. 5D and 7C), where a large portion of the population tended to accumulate TAG indistinctively of cell size (Fig. 5D). Size distribution analysis on day 9 showed a distinctive shift away from smaller cells under hypersalinity, with a relative absence of small daughters (20–33 μ^2), as well as an enrichment in a specific population of daughters (63%) centered at 40μ^2 (Fig. 7C). The data indicate that large daughters are contributing substantially to TAG accumulation under hypersaline conditions. The mother cell contribution was similar in all three conditions. In summary, wild type *M. kaistiae* KAS603 adapted to hypersalinity displayed a similar growth as under low salinity conditions, and had higher TAG yield, with large daughter cells contributing more substantially to TAG accumulation than in previous conditions tested.

To test trophic mode change and productivity in seawater, we

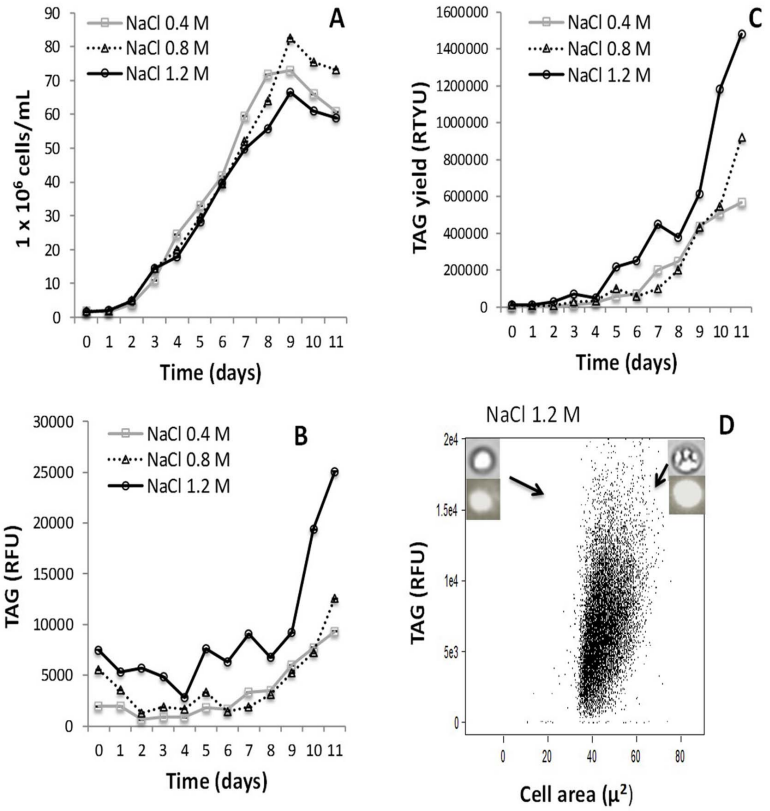

Fig. 5. Response comparison of salinity adapted *M. kaistiae* KAS603 in three NaCl concentrations. The base culture medium was artificial seawater (ASW), which contains 400 mM NaCl (sodium chloride alone was added to the higher salinities). A) Culture density response, B) TAG accumulation response, C) TAG yield response, D) TAG accumulation response relative to cell area of a population with 1.2 M NaCl on day 9 of cultivation. Cell images inside squares are bright field (upper squares) and BODIPY (TAG) fluorescence (lower squares). Black arrows indicate the place on the scatter plot where cells in square images are located. RFU denote relative fluorescence units. RTYU are relative TAG yield units. The error bars indicate standard error (n = 3).

cultured ASW (0.4 M NaCl) adapted cells for 9 days under acetate-supplemented vs non-supplemented conditions. Acetate supplementation resulted in faster growth and higher maximum culture density (Fig. 6A). A maximum of 2 doublings per day was achieved under this condition. Higher TAG accumulation (Fig. 6B) and higher TAG yield (Fig. 6C) occurred with acetate supplementation. TAG accumulation initiated during late log phase and continued throughout stationary phase in the presence of acetate. During TAG accumulation (day 6/late log), small size daughter cells contributed to most of the TAG (Fig. 6D). Examination of the size distribution of cells on day 6 grown under 50 mM acetate, determined that 98% of the population was composed of daughter cells (Fig. 7E) vs 75% when cultivated in 0 mM acetate (Fig. 7D). In summary, sodium acetate (50 mM) supplementation of *M. kaistiae* KAS603 seawater culture, improved cell growth, favored fast TAG accumulation, high TAG yield and enriched for small daughter cells (suggesting an early release from mothers) which contributed the most to TAG accumulation.

3.2. Chlorophyll

To evaluate changes in chlorophyll content under autotrophic vs mixotrophic cultivation (under different substrates: sodium acetate, sodium bicarbonate and sodium chloride) in *M. kaistiae* KAS603; we measured mean chlorophyll auto-fluorescence per cell under excitation with a 488 nm laser, using imaging flow cytometric profiling. Although chlorophyll levels varied during the course of experiments, the only consistent trend was increased chlorophyll with increased culture density. There was no evidence of chlorosis under any conditions.

3.3. Pilot plant cultivation of lab-adapted strains and invasive species control

We monitored the ability of the hypersaline-adapted strain to resist invasive species in a semi-controlled outdoor environment. *M. kaistiae* KAS603 strains (previously adapted to low and high salinities) were cultivated for 15 days at pilot scale level under natural seawater as the control condition, and hypersaline (1.2 M NaCl) seawater as the experimental condition (Fig. 8). The light regime averaged 14 h during day and 10 h during night. The temperature in the tank cultures fluctuated between 22 and 26 °C, the pH remained constant throughout the experiment at 8.5. As depicted in Fig. 8, cells under hypersalinity displayed a healthy growth and proceeded through lag, log and an extended stationary phase; while the control experienced a rapid decrease in culture density (Fig. 8A) from days 1 through 4 due to the presence of invasive organisms (Fig. 8B), and was able to recover by day 5 of the experiment, although experiencing a slower growth rate (days 3 to 5) and a later arrival (day 6) into logarithmic phase than the hypersaline cultured cells (day 2). Ciliates and rotifers grew in the

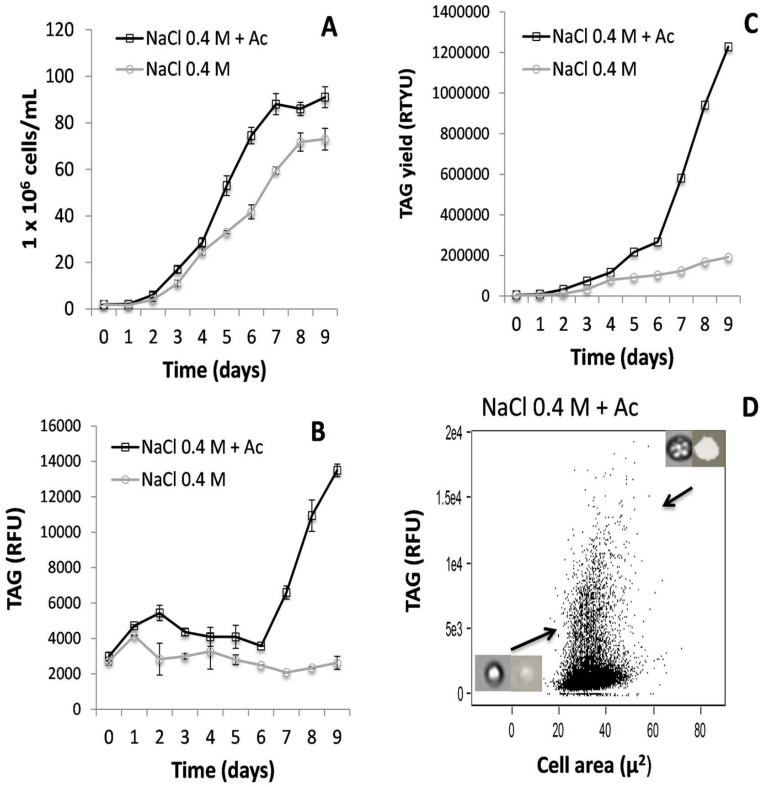

Fig. 6. Seawater mixotrophic vs autotrophic cultures of *M. kaistiae* KAS603. Cells were cultured in ASW (0.4 M NaCl) supplemented with 50 mM sodium acetate (Ac) vs non acetate supplemented ASW (0.4 M NaCl). A) Culture density response, B) TAG accumulation response, C) TAG yield response, D) TAG accumulation response relative to cell area of a population with Ac supplemented ASW (0.4 M NaCl) on day 6 of cultivation. Cell images inside squares are bright field (left) and BODIPY (TAG) fluorescence (right). Black arrows indicate the place on scatter plot where cells in square images are located, RFU denote relative fluorescence units. RTYU are relative TAG yield units. The error bars indicate standard error (n = 3).

control culture condition from days 1 to 5, with more abundant (up to 10 fold) ciliates than rotifers, causing an immediate effect of algal population decline (Fig. 8). The hypersaline cultures had negligible counts/amounts of invading organisms and developed a healthy algal growth (data not shown).

4. Discussion

Our research focused on evaluating and enhancing *M. kaistiae* KAS603 adaptation ability, growth and TAG accumulation under 14 cultivation conditions, the results of which are summarized in Table 1. The experimental system was developed stepwise, cultivating under different amounts of either bicarbonate or acetate in freshwater, then adapting cultured cells from freshwater to seawater and hypersalinity, and finally cultivating with acetate in seawater. We tested the ability of hypersaline artificial seawater adapted lab cultures to resist invasive species in a semi-controlled outdoor system.

This is the first study to have developed such a broad variety of environmental adaptations in one microalgal strain in a short amount of time. This can be taken advantage of in terms of evaluating the effect of different cultivation conditions in a controlled single-species setting, and to eventually understand the underlying basis of adaptability related to high yield. One question to address is what is the basis of the ability to adapt? *M. kaistiae* KAS603 required adaptation over 56 generations to grow at the highest salinity. The adaptation could have involved selection of cells pre-adapted to growth at high salinity, or a genetic or epigenetic change that enabled growth under that condition. Adaptation through mutations is relatively slow, for example, mutations in the green alga *Chlamydomonas reinhardtii* occur in 3 of every 10 billion positions in the genome per generation [10], and since its genome size is 121 Mbp, then on average 0.036 mutations would occur per generation. There are three possibilities for *M. kaistiae* KAS603, 1) a pre-existing mutant was amplified by selection under hypersalinity, 2) the selection generated a mutation which enabled hypersaline growth, or 3) epigenetic changes occurred. If a pre-existing mutant as a single cell was present and then amplified, it would take an average of 20 generations for it to become dominant in the population. That did not occur, which rules out that possibility. The mutation rate in the genome is so low that inducing a mutation would also not explain the data. This leaves epigenetic changes as the most plausible explanation. During adaptation of *M. kaistiae* KAS603 to high salt, methylation and/or acetylation which affect expression levels of genes that regulate osmolyte production might be occurring [11,12,13]. As for the rest of the conditions we evaluated, the adjustment period was immediate, as such, they were considered acclimated vs adapted as occurred under hypersalinity. Interestingly, although it took many generations to adapt *M. kaistiae* KAS603 to hypersaline growth, hypersaline-adapted cells could immediately grow when inoculated into freshwater modified MS4 medium (data not shown).

An unanticipated result was the non-linear response of *M. kaistiae* KAS603 to various treatments, for example, TAG accumulation was optimal at two intermediate concentrations of bicarbonate, but not at

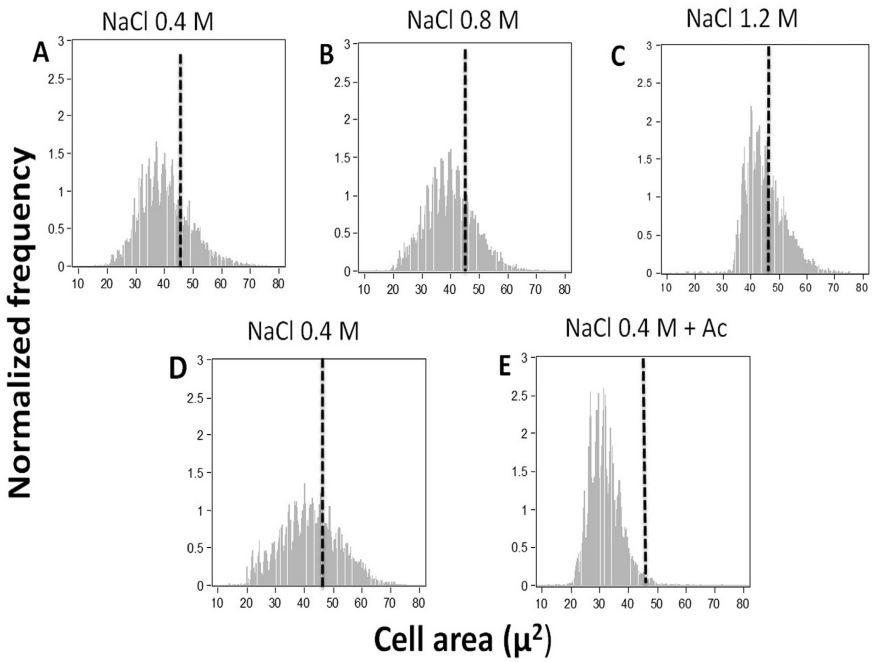

Fig. 7. Cell area population distribution comparison among cells cultured in ASW with low (A), medium (B) and high (C) NaCl amounts during day 9 of cultivation; as well as cells cultured in ASW with low NaCl alone (D) and with low NaCl plus medium sodium acetate (E) during day 6 of cultivation. The mean cell area for A is 40.2 μ^2, for B is 39.9 μ^2, for C is 45.2 μ^2, for D is 42.2 μ^2 and 31.9 μ^2 for E. Black dotted vertical line separates daughter cells (left side) from mother cells (right side).

lower or higher concentrations. These are highly reproducible responses, as three biological replicates were included in the data. Although the specifics may not be identifiable, one can possibly explain these types of results in a general sense, by considering that the uptake, maintenance, and utilization of bicarbonate and acetate, and possible consequences of high salinity growth, will involve different metabolic and energetic requirements for the cell (Figs. 9 and 10). In green algae, bicarbonate is actively transported across the cell membrane by an ABC type transporter that utilizes ATP [14] - thus there is an energetic requirement in terms of ATP hydrolysis for uptake (Fig. 9). Inside the cell, the presence of bicarbonate in the cytoplasm or other cellular compartments could affect pH, which will require adaptations such as synthesis of buffering compounds by the cell [15,16,17], which also will require energy. Transport of bicarbonate into the chloroplast involves an anion transporter located in the plastid envelope [14], which will involve an electrochemical gradient. Utilization of bicarbonate by the cell for growth ultimately involves its conversion by carbonic anhydrase to CO_2, followed by fixation by RubisCO. Thus, there is an intrinsic need for photosynthesis for an algal cell to utilize bicarbonate for growth. Even if CO_2 generation occurs in close proximity to RubisCO, CO_2 losses from the chloroplast can be substantial, up to 70% [18]. Because CO_2 readily diffuses across membranes, it can be easily lost from the cell, and although it can be recaptured prior to loss by intracellular carbonic anhydrases, the repeated interconversion between bicarbonate and CO_2 leads to inefficiency. From this discussion, it is clear that cellular utilization of bicarbonate imposes energetic costs at different levels that should be considered relative to the benefits of providing a relatively abundant source of carbon (Fig. 9). The high pH, which is maintained by the high bicarbonate, increases the solubility of CO_2 (when bicarbonate is utilized, hydroxyl ions are released, resulting in the production of hydroxyl-based alkalinity from carbonate alkalinity effectively increasing the driving force for additional CO_2 dissolution into the medium [17]), which could compensate for the negative energetic effects of the bicarbonate.

In contrast to bicarbonate, the utilization of acetate by a cell has fewer energetic requirements (Fig. 9). Acetate is transported into the cell via a cation/acetate symporter, which relies on a favorable concentration gradient by the cation. Eventually, the cation will have to be pumped out of the cell by an energy requiring process (e.g. sodium/potassium ATPase). Once in the cytoplasm, acetate can be converted to acetyl-CoA. The glyoxylate cycle (in glyoxysomes) incorporates acetyl groups of acetyl-CoA into carbon skeletons; as has been seen in other *Chlorella* strains [19], avoiding a CO_2 fixation step entirely. The glyoxylate cycle generates NADH. Acetyl-CoA is also incorporated into the TCA cycle (in mitochondria) producing 1 ATP, 1 $FADH_2$, 3 NADH and 2 CO_2 molecules per acetyl-CoA. Acetate can also diffuse from the cytoplasm into the chloroplast. Once in the plastid, it can be used to form acetyl-CoA and incorporate directly into fatty acid synthesis [20,21] and TAGs in the endoplasmic reticulum, using mainly NADPH (Fig. 9). In general the energetic requirements for utilizing acetate should be substantially lower than for bicarbonate.

Maintenance of metabolism at high salinity also involves cellular energy (Fig. 10). Sodium can enter the cell by a plasma membrane Na^+/H^+ antiporter, and an increase in Na^+ (hypersalinity), can generate a sharp transient increase in intracellular Ca^{2+} that modulates (in conjunction with other proteins) Na^+ extrusion (through the Na^+/H^+ antiporter) [22]. Enduring calcium levels/signaling in hypersaline adapted cells can induce neutral lipid biosynthesis by increasing calmodulin activity and ATP production through cyclic electron flow (CEF) increase in the chloroplast, as has been reported for *Chlorella sorokiniana* [23,24]. Thus hypersalinity can increase ATP production for TAG synthesis via CEF. The cell would have to adapt to these changes by energy-involved maintenance of ionic gradients. Hypersalinity can

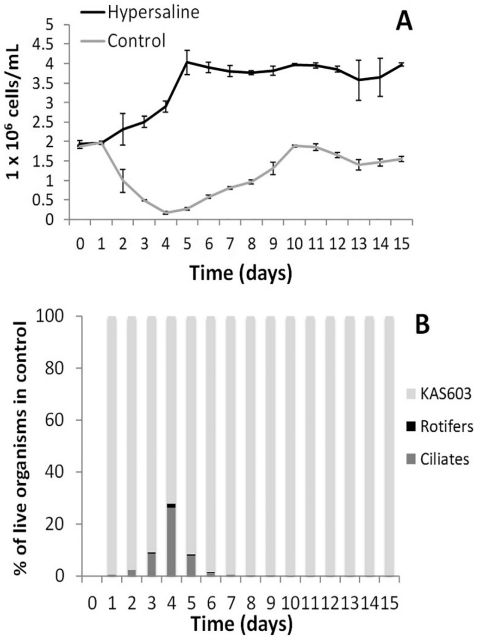

Fig. 8. Outdoor cultures of *M. kaistiae* KAS603 at a pilot plant. Cells were cultured in natural seawater (control) and natural seawater adjusted to 1.2 M (hypersaline) in 1000 L open tank cultures. A) Culture density, B) Relative percentage comparison of invading organisms (rotifers and ciliates) and *M. kaistiae* KAS603 (referred to as KAS603) over time. The error bars represent standard deviation (n = 2).

Table 1
Summary of *M. kaistiae* KAS603 top 4 cultivation conditions. MS4 medium with 50 mM sodium acetate was used for mixotrophy in freshwater. MS4 medium with 100 mM sodium bicarbonate was used for autotrophy in freshwater. ASW medium with 0.4 M NaCl and 50 mM NaAc was used for mixotrophy in seawater. ASW medium with 1.2 M NaCl was used for hypersalinity. Culture density corresponds to maximum TAG yield for each condition. TAG accumulation period displays the growth phase at which accumulation starts, as well as the entire period duration during cultivation. TAG yield (maximum TAG yield is expressed along with the day of occurrence) was calculated as the product of culture density x relative TAG content per cell, and is expressed in relative TAG yield units (RTYU).

Cultivation conditions	Culture density (cells/mL)	Max cell growth (doublings/day)	TAG accumulation period	Maximum TAG yield (RTYU)
Mixotrophy in freshwater	5×10^7 (day 7)	2.7	Early log onset (day 2–7)	264,441 (day 7)
Autotrophy in freshwater	3×10^6 (day 7)	1.7	Late log onset (day 4–7)	58,000 (day 7)
Mixotrophy in seawater	9×10^7 (day 9)	2	Late log onset (day 6–9)	1,227,668 (day 9)
Hypersalinity	6×10^7 (day 11)	1.5	Stationary onset (day 9–11)	1,478,526 (day 11)

also trigger formation of osmolytes (such as glycine and proline), which can be produced in proportion to external osmolarity [11,25,26]. These osmolytes will require energy to produce, but potentially can serve as precursors for fatty acid synthesis [11,27]. In *D. salina* it has been proposed that hypersalinity leads to the conversion of starch into glycerol in the chloroplast [11,28]. It is possible that under hypersalinity, both calcium signaling and osmolytes (such as glycine, proline or glycerol) might act in synergy to increase TAG production. We propose glycerol as the most plausible osmolyte contributor to TAG synthesis in *M. kaistiae* KAS603, because it can readily be incorporated into TAG, it is the least energy expensive to produce and follows a direct route towards TAG anabolism (Fig. 10).

M. kaistiae KAS603 generates autospores (daughter cells) and autosporangia (mother cells), which impose a life cycle stage impact on cell size. In microalgae, a direct relationship among (single) cell size and nutrient uptake has been established: small cells have faster nutrient uptake than large cells due to a higher surface area to volume ratio [3], but large cells have a higher storage capacity and fill up more slowly than smaller cells when acquiring nutrients, and thus can sustain uptake rates for longer periods of time [29,30,31]. As *M. kaistiae* KAS603 daughter cells are increasing in size and accumulating biomass, nutrients are being internalized more slowly. If a nutrient (nitrogen or carbon for example) is added to the culture medium, the small daughters can internalize it more rapidly, favoring their growth and sustainability under low nutrient conditions. Phytoplankton density and biomass generally have size-dependent maxima, so that under optimal conditions, larger algae support lower biomasses and a less dense population than smaller algae [32]. *M. kaistiae* KAS603 multiple fission reproduction mode provides smaller cells either as free or captive daughters inside the mother cell. Such predominance gives *M. kaistiae* KAS603 a continuous nutrient uptake and biomass accumulation advantage. The larger (mother) cell mostly provides a physical form of protection for daughters in division. A general trend has been observed in *Chlorella* species undergoing multiple-fission in that under fast growth conditions with abundant nutrients, mother cells are more prevalent in the population [5]. This could be a strategy to reduce energy and biochemical input into making larger size offspring during the mitotic process – a mother cell containing higher numbers of daughters would seed a culture with more offspring at the burst period, and in the presence of abundant nutrients, the small daughters could more quickly increase their volume to become mother cells. Under nutrient-poor conditions, the maintenance of daughters would be favored because small daughters would be more efficient at nutrient uptake.

The concepts described above serve as framework to understand that multiple energetic, metabolic, and life cycle processes are involved in the utilization of bicarbonate and acetate and adaptation to higher salinity, and that individual aspects of these processes could impose constraints at particular substrate concentrations. We interpret the data below with this in mind.

The growth response of *M. kaistiae* KAS603 to bicarbonate addition was non-linear, with optimal growth at 25 mM, but less growth at lower and higher concentrations (Fig. 1A). TAG accumulation occurred best at two intermediate bicarbonate concentrations (50 and 100 mM) (Fig. 1B). Previous work by Lohman et al. (2015) [17] on *Chlorella vulgaris* also found optimal growth with a low bicarbonate concentration (5 mM $NaHCO_3$) and highest TAG accumulation at 50 mM $NaHCO_3$. Other microalgal species (*Tetraselmis suecica*, *Nannochloropsis salina*, and *Chlorella vulgaris*) have demonstrated an optimum threshold tolerance to bicarbonate addition above which reduced cell densities/biomass have been recorded [33,34]. These data suggest that utilization and processing of bicarbonate in the cells depends on particular thresholds of availability. We propose that at 0 mM bicarbonate, the cell is limited for carbon, resulting in the slowest growth rate and no TAG accumulation. At 25 mM bicarbonate, sufficient extra carbon may be available to stimulate growth, but not enough is present to stimulate TAG accumulation. This suggests a preferential partitioning of carbon towards growth processes and not storage. At 50 mM bicarbonate, growth is suppressed to the 0 mM levels, but TAG accumulation is stimulated. This could result from an energetic drain to process more bicarbonate entering the cell, reducing growth rate, but once the bicarbonate is in the cell, the excess carbon stimulates TAG accumulation. Another consideration is the relation between growth and TAG accumulation in microalgae, in general, they are inversely correlated

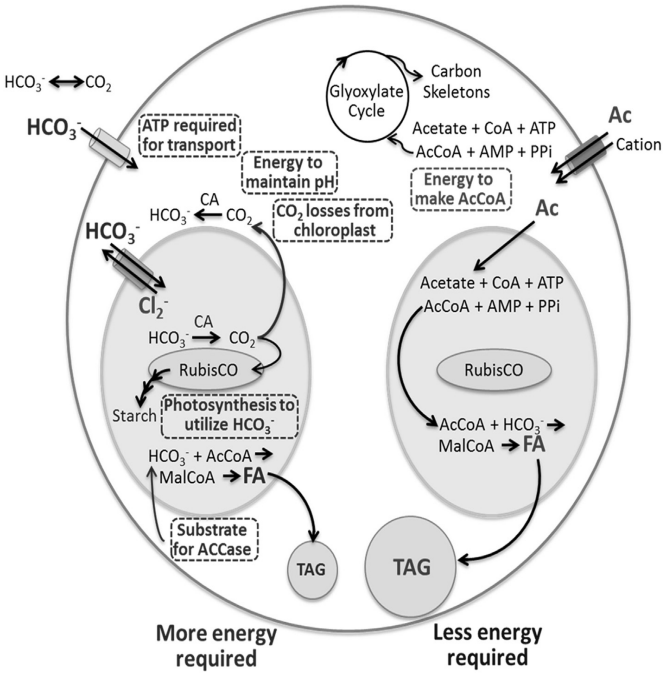

Fig. 9. Proposed steps involved in bicarbonate (left) and acetate (right) uptake and utilization. More energy is required for bicarbonate uptake and usage compared to acetate. Acetate in the cell can be directly incorporated into fatty acids (FA) and triacylglycerids (TAG) synthesis. Ac denotes acetate, CA denotes carbonic anhydrase, CoA denotes coenzyme A, AcCoA is acetyl coenzyme A, ACCase is acetyl coenzyme A carboxylase, MalCoA is malonyl coenzyme A, PPi is pyrophosphate, ATP is adenosine triphosphate, AMP is adenosine monophosphate, RubisCO is ribulose 1,5-bisphosphate carboxylase/oxygenase.

under autotrophic conditions [35,36,37], so the slowing in growth rate may enable more carbon for TAG synthesis. At 100 mM bicarbonate, growth is stimulated again as is TAG accumulation, suggesting that the excess carbon available overcomes an energetic limitation. At 150 mM bicarbonate, growth is reasonable, but minimal TAG accumulation occurs, perhaps due to oversaturation of carbon that can affect/alter pH and cause inefficient carbon flux towards fatty acid and TAG formation.

There is a shift in the population of cells towards daughters in the presence of bicarbonate (Fig. 3), suggesting that either release of daughters from mothers is promoted, or a slower conversion of daughters turning into mothers occurs. Because growth is generally stimulated in the presence of bicarbonate (Fig. 1A), the shift towards daughters suggests that release of daughters from mothers is promoted. Based on the literature, daughters should be favored under lower nutrient conditions [5]. In spite of the improved growth rate under most bicarbonate concentrations, the shift towards daughters suggests that bicarbonate is eliciting a “stress”, and perhaps more specifically, a metabolic or energetic drain on the cells. Another factor to consider is that daughter cells captive in the mother cell should shade each other, reducing overall photosynthetic capability. Since bicarbonate utilization for growth requires photosynthesis, release of daughter cells to reduce shading could be favored under these conditions. TAG accumulation occurred with 50 mM bicarbonate specifically in the mother cells, which also may be indicative of a stress condition where cell division is inhibited.

Higher culture density and higher TAG yield were achieved with acetate addition compared to bicarbonate. All additions of acetate displayed equally improved growth (Fig. 2A). As previously suggested [38] for *Chlorella pyrenoidosa*, acetate could provide a significant increase in cellular respiration rate to produce energy for growth and reduce dependence on light. According to the scheme in Fig. 9, imported acetate would be preferentially used for energy and carbon skeleton generation, which is consistent with an equivalent effect on growth regardless of concentration (Fig. 2A). In contrast, bicarbonate would have to be processed through photosynthetic carbon fixation in order to improve growth (Fig. 9), which may explain the non-linear response of bicarbonate addition on growth (Fig. 1) Only at higher concentrations of acetate, 50 and 100 mM, is the TAG accumulation process stimulated (Figs. 2B and 9), according to the scheme in Fig. 9, carbon in excess of that needed for growth could be channeled into TAG synthesis, but growth is initially favored. With bicarbonate, TAG was also induced at 50 and 100 mM; since bicarbonate is a direct substrate for ACCase, this can bypass the steps required for growth processes (Fig. 9). TAG accumulation is not promoted at 150 mM acetate because acetate could be toxic at high concentrations [19] directly affecting internal pH. Another possibility is that an excessive amount of acetate being processed could inhibit TAG synthesis due to oversaturation and possibly a negative metabolic feedback effect under such excess of substrate, lowering Acetyl CoA levels. Due to equal culture density, TAG yield was driven only by TAG accumulation, being equally high at 50 mM and 100 mM (Fig. 2B and C) during stationary phase. During day 6 of stationary phase, acetate-added conditions produced more mother cells (62–63%) than without acetate (28.6%) (Fig. 3D–F). This is in contrast to the bicarbonate treatments (Fig. 3A–C), in which more daughters were present (62–87%). Since both treatments provide additional carbon to the cells and improve growth, this suggests that the mode of utilization of carbon by the cells is altering the proportion of daughters and mothers. In the case of acetate, the data suggest a favoring of mitotic processes leading to more daughter cells encased within the mother cell, and less frequent release of daughters from

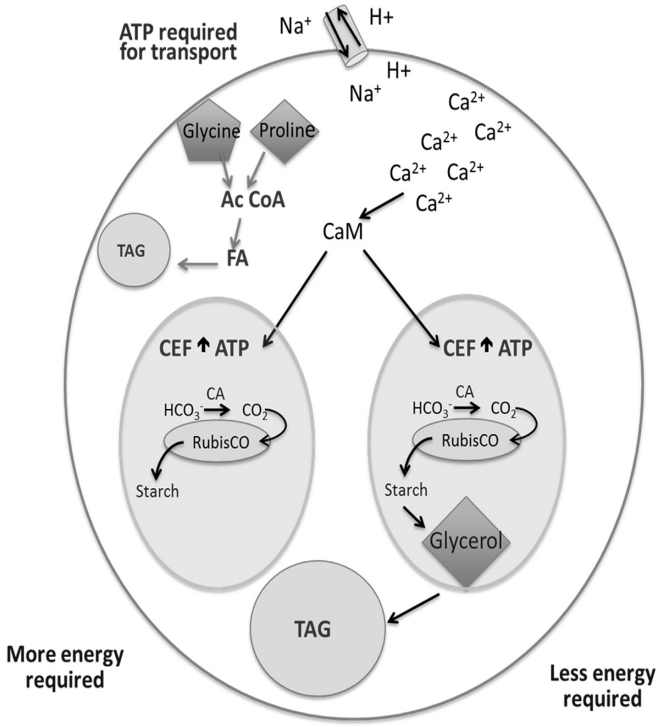

Fig. 10. Proposed processes involved in dealing with hypersaline conditions. If sodium in excess enters the cell, it can trigger calcium (Ca^{2+}) and calmodulin (CaM) release into the cytoplasm, which increases cyclic electron flow (CEF) and ATP production in the chloroplast, favoring energy for TAG synthesis. Osmolytes such as glycine or proline can be produced in the cytoplasm yielding acetyl coenzyme A (AcCoA), fatty acids (FA) and a low amount of triacylglycerol (TAG) with a high energy demand. Glycerol can be produced in the chloroplast from starch, diffuse (lipid soluble) into the cytoplasm and directly be incorporated into a larger amount of TAG in lipid droplet form, which requires less energy than glycine or proline to form (more detailed information in the text Discussion section).

mothers. Because the growth rate with acetate is faster than with bicarbonate, the accumulation of more daughters per mother is more efficient at increasing biomass. When cell area and TAG were correlated (day 6 / stationary phase), there was a trend towards mother cells with higher TAG (Fig. 2D) at 50 mM, which was corroborated by the population distribution differences (Fig. 3E).

Under different concentrations of NaCl, the growth response of adapted cells of *M. kaistiae* KAS603 was similar (Fig. 5A). One might expect that adaptation to hypersalinity might involve an energetic or metabolic drain due to a requirement for osmolyte synthesis, but if that did occur with *M. kaistiae* KAS603, it did not correlate with the degree of salinity. Growth in any of the saline conditions was slower than in freshwater MS4 medium, but because the saline medium lacked bicarbonate, it is not possible to link the difference in growth to salinity per se. It is possible that a particular osmolyte may minimize metabolic drain on the cells, or it could provide a compensating benefit [39]. TAG accumulation was enhanced under the highest (1.2 M) concentration of NaCl (Fig. 5B). Osmolytes such as proline or glycine can be augmented in response to increases in external salinity as previously reported in terrestrial plants, *Dunaliella salina* and *Chlorella autotrophica* [11,25,26]. Glycine can be converted to serine and then to pyruvate yielding Acetyl CoA, fatty acids (FA) and TAGs. Proline can be converted to glutamate and then to alpha ketoglutarate, entering the TCA cycle, yielding Acetyl CoA, which can be used for FA and TAG synthesis. Both amino acids might be able to favor TAG accumulation under hypersalinity in *M. kaistiae* KAS603 if they are overproduced as osmolytes. Another intracellular osmotic metabolite could be glycerol, as previously reported in *Dunaliella* [11]. Glycerol can be easily incorporated into TAG anabolism in the cytoplasm. TAGs were induced during stationary phase (Fig. 5B and C), and we propose that the presence of excess osmolytes in the 1.2 M adapted strains provided substrates for TAG synthesis.

Although there was no trend in overall size change for cells while accumulating TAG under hypersalinity (Fig. 5D), the entire cell population shifted towards larger daughter cells (63%) (Fig. 7C) relative to lower salinities (Fig. 7A and B). This could possibly be due to osmolyte-induced volume changes. Under hypersaline conditions, both mothers and daughters were capable of accumulating TAG (Fig. 5D). We hypothesize that the presence of osmolytes in all cells can be readily converted to fatty acids and TAGs.

Cells in acetate in seawater grew faster and to higher density than without acetate (Fig. 6A), consistent with results in Fig. 2A. TAG accumulation initiated during late log phase and increased through stationary phase in the presence of acetate (Fig. 6A and B). There was a dramatic shift to daughter cells (98.4%) in the presence of acetate (Fig. 7E), which is in contrast to the response in freshwater-based medium (36–38% daughters) (Fig. 3E and F). This suggests that while more energy is required to thrive in seawater, mother cells burst at an earlier stage (when acetate is present). The general trend of larger cells under higher salinity [40] may impose a stress on mother cells – if the daughters are attempting to increase in size within the mother, bursting of the mother cell could be favored.

Factors that control the shift in cell size and life cycle stage (mothers and daughters) are incompletely understood, in part because obtaining good statistics is challenging using standard microscopy. The application of imaging flow cytometry enabled us to evaluate 10,000 + cells in

a population. The results reveal distinct trends, in the presence of bicarbonate, when growth rate was enhanced relative to the control, a shift towards daughter cells was observed (Fig. 3A–C), but in acetate, which also increased growth rate, a shift towards mothers occurs (Fig. 3D–F). Also observed was the ability of either mothers or daughters to accumulate TAG, the underlying mechanism under different culture conditions is unclear. This analysis indicates that factors associated with the multiple fission process have an impact on productivity.

To evaluate whether hypersalinity could serve as a preventive measure for the growth of contaminating species, we tested the performance of *M. kaistiae* KAS603 under natural seawater and hypersaline conditions at pilot plant scale. The pilot plant hypersaline condition was advantageous for *M. kaistiae* KAS603 because it promoted higher culture density, and apparently lessened the impact of invasive organisms (Fig. 8), facilitating *M. kaistiae* KAS603 survival.

5. Conclusions

We have evaluated cell growth, TAG accumulation and reproductive mode in *M. kaistiae* KAS603 by testing its trophic capacity (autotrophy/mixotrophy) and adaptability from freshwater to seawater and hypersalinity. Overall, *M. kaistiae* KAS603 was able to grow fastest (2.7 doublings per day) mixotrophically in freshwater (MS4, acetate 50 mM), it was able to have intermediate growth (2 doublings per day) mixotrophically in seawater (ASW, acetate 50 mM); and it grew the slowest (1.5 doublings per day) in hypersalinity (ASW, 1.5 mM NaCl) but produced the most TAG (Table 1). The non-linear response to bicarbonate or acetate addition in terms of growth and TAG accumulation had a general trend – growth processes were favored first, and TAG accumulation was promoted only with additional added carbon (Figs. 1 and 2). This suggests preferential partitioning of carbon into growth processes, which makes ecological sense. Life-cycle stage analysis revealed that 1) autotrophic and mixotrophic freshwater conditions favored larger size mother cells for TAG accumulation, 2) mixotrophic seawater conditions favored daughter cells for TAG accumulation, and 3) hypersaline conditions favored large daughter cells for TAG accumulation.

The multiple fission cell stages (which are associated to growth and TAG yield) are impacted by the culture conditions tested: mother cells are prevalent high TAG accumulators in freshwater and can experience an early (bicarbonate) or late release of daughters (acetate). In seawater with acetate, early released daughter cells accumulate high TAG, and achieve the highest culture density. Late released daughter cells accumulate high TAG during hypersalinity. The impact of such differences is relevant to productivity because the amount/type of supplemented carbon or NaCl in a culture can be manipulated to increase or decrease the number and size of high TAG accumulating cells.

The fact that algae genotypes are able to adjust to their habitat through phenotypic plasticity has not been extensively documented. Our work indicates that *M. kaistiae* KAS03 is especially adept at adaptation, conferring the advantage of survival in changing environments in a short (acclimation) and long term (adaptation). This feature can be taken advantage of 1) to understand the underlying molecular basis for adaptability in the same strain under different conditions, and 2) for applications in a wide range of production scenarios, potentially in diverse geographic locations with different water sources.

Acknowledgements

This work was funded by AFOSR grant FA9550-08-1-0178 and UC MEXUS-CONACYT grant CN-12-622. Special thanks to the Meredith Gould lab (UABC) and the Glycotechnology Core Resource (UCSD) for their technical support, and to Kuehnle AgroSystems, Inc., for providing the algal strain.

References

[1] General Atomics, Affordable algal-derived JP-8 surrogate. Project Final Report for Defense Advanced Research Projects Agency, (2012) (proprietary).
[2] A. Kuehnle, R. Schurr, M. Perez, Some microalgae from the Hawaiian islands with a focus on industrial applications, Cur. Biotech. 4 (2015) 499–513.
[3] E. Marañón, P. Cermeño, D.C. López, T. Rodríguez, C. Sobrino, M. Huete, J.M. Blanco, J. Rodríguez, Unimodal size scaling of phytoplankton growth and the size dependence of nutrient uptake and use, Ecol. Lett. 16 (2012) 371–379.
[4] M. Yamamoto, M. Fujishita, A. Hirata, K. Shigeyuki, Regeneration and maturation of daughter cell walls in the autospore-forming green alga *Chlorella vulgaris* (Chlorophyta, Trebouxiophyceae), J. Plant Res. 117 (2004) 257–264.
[5] K. Bisova, V. Zachleder, Cell-cycle regulation in green algae dividing by multiple fission, J. Exp. Bot. 65 (2014) 2585–2602.
[6] R.K. Mandalam, B.O. Palsson, Cell cycle of *Chlorella vulgaris* can deviate from the synchronous binary division model, Biotechnol. Lett. 19 (1997) 587–591.
[7] J. Kim, J.-Y. Lee, C. Ahting, R. Johnstone, T. Lu, Growth of *Chlorella vulgaris* using sodium bicarbonate under no mixing condition, Asia Pac. J. Chem. Eng. 9 (2014) 604–609.
[8] Y. Chisti, Biodiesel from microalgae, Biotechnol. Adv. 25 (2007) 294–306.
[9] R.C. Goswami, N. Kalita, M.C. Kalita, A study on growth and carbon dioxide mitigation by microalgae *Selenastrum* sp.: its growth behavior under different nutrient environments and lipid production, Ann. Biol. Res. 3 (2012) 499–510.
[10] R.W. Ness, A.D. Morgan, N. Colegrave, P.C. Keightley, Estimate of the spontaneous mutation rate in *Chlamydomonas reinhardtii*, Genetics 192 (2012) 1447–1454.
[11] A. Ramos, J. Polle, D. Tran, J.C. Cushman, E. Jin, J.C. Varela, The unicellular green alga *Dunaliella salina* Teod. as model for abiotic stress tolerance: genetic advances and future perspectives, Algae 26 (2011) 3–20.
[12] A. Veluchamy, X. Lin, F. Mamus, M. Rivarola, J. Bhavsar, T. Creasy, K. O'Brien, N. Sengamalay, L. Tallon, A. Smith, E. Rayko, I. Ahmed, S. Le Crom, G. Farrant, J.Y. Sgro, S.A. Olson, S.S. Bondurant, A.E. Allen, P.D. Rabinowicz, M.R. Sussman, C. Bowler, L. Tirichine, Insights into the role of DNA methylations in diatoms by genome-wide profiling in *Phaeodactylum tricornutum*, Nat. Commun. 4 (2013) 2091.
[13] B.R. Lyon, T. Mock, Polar microalgae: new approaches towards understanding adaptations to an extreme and changing environment, Biology 3 (2014) 56–80.
[14] S. Subramanian, A.N. Barry, S. Pieris, R.T. Sayre, Comparative energetics and kinetics of autotrophic lipid and starch metabolism in chlorophytic microalgae: implications for biomass and biofuel production, Biotech. Biof. 6 (2013) 150.
[15] J.A. Raven, F.A. Smith, Intracellular pH regulation in the giant-celled marine alga *Chaetomorpha darwinii*, J. Exp. Bot. 31 (1980) 1357–1369.
[16] R.D. Gardner, K.E. Cooksey, F. Mus, R. Macur, K. Moll, E. Eustance, R.P. Carlson, R. Gerlach, M.W. Fields, B.M. Peyton, Use of sodium bicarbonate to stimulate triacylglycerol accumulation in the chlorophyte *Scenedesmus* sp. and the diatom *Phaeodactylum tricornutum*, J. Appl. Phycol. 24 (2012) 1311–1320.
[17] E.J. Lohman, R.D. Gardner, T. Pedersen, B.C. Peyton, K.E. Cooksey, R. Gerlach, Optimized inorganic carbon regime for enhanced growth and lipid accumulation in *Chlorella vulgaris*, Biotechnol. Biofuel. 8 (2015) 82.
[18] B.M. Hopkinson, C.L. Dupont, A.E. Allen, F.M. Morel, Efficiency of the CO_2-concentrating mechanism of diatoms, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) (2011) 3830–3837.
[19] O. Pérez-García, F.M. Escalante, L.E. de-Bashan, Y. Bashan, Heterotrophic cultures of microalgae: metabolism and potential products, Water Res. 45 (2011) 11–36.
[20] V. Kumar, M. Muthuraj, B. Palabhanvi, A.K. Ghoshal, D. Das, High cell density lipid rich cultivation of a novel microalgal isolate *Chlorella sorokiniana* FC6 IITG in a single-stage fed-batch mode under mixotrophic condition, Bioresour. Technol. 170 (2014) 115–124.
[21] P.B. Heifetz, B. Forster, C.B. Osmond, L.J. Giles, J.E. Boynton, Effects of acetate on facultative autotrophy in *Chlamydomonas reinhardtii* assessed by photosynthetic measurements and stable isotope analyses, Plant Physiol. 122 (2000) 1439–1445.
[22] B. Gupta, B. Huang, Mechanism of salinity tolerance in plants: physiological, biochemical, and molecular characterization, Int. J. Genom. (2014) 701596, http://dx.doi.org/10.1155/2014/701596.
[23] H. Chen, Y. Zhang, C. He, Q. Wang, Ca^{2+} Signal transduction related to neutral lipid synthesis in an oil-producing green alga *Chlorella* sp. C2, Plant Cell Physiol. 55 (2014) 634–644.
[24] H. Chen, J. Hu, Y. Qiao, W. Chen, J. Rong, Y. Zhang, C. He, Q. Wang, Ca^{2+}-regulated cyclic electron flow supplies ATP for nitrogen starvation-induced lipid biosynthesis in green algae, Sci. Rep. 5 (2015) 15117.
[25] I. Ahmad, J.A. Hellebust, Osmoregulation in the extreme euryhaline marine microalga *Chlorella autotrophica*, Plant Physiol. 74 (1984) 1010–1015.
[26] M. Ashraf, M.R. Foolad, Roles of glycine betaine and proline in improving plant abiotic stress resistance, Environ. Exp. Bot. 59 (2007) 206–216.
[27] M. Azachi, A. Sadka, M. Fisher, P. Goldshlag, I. Gokhman, A. Zamir, Salt induction of fatty acid elongase and membrane lipid modifications in the extreme halotolerant alga *Dunaliella salina*, Plant Physiol. 129 (2002) 1320–1329.
[28] E. Chitlaru, U. Pick, Regulation of glycerol synthesis in response to osmotic changes in *Dunaliella*, Plant Physiol. 96 (1991) 50–56.
[29] W. Stolte, R. Riegman, Effect of phytoplankton cell size on transient-state nitrate and ammonium uptake kinetics, Microbiology 141 (1995) 1221–1229.
[30] A. Verdy, M. Follows, G. Flierl, Optimal phytoplankton cell size in an allometric model, Mar. Ecol. Prog. Ser. 379 (2009) 1–12.
[31] P.G. Falkowski, M.J. Oliver, Mix and match: how climate selects phytoplankton, Nat. Rev. Microbiol. 5 (2007) 813–819.
[32] S. Agusti, J. Kalff, The influence of growth conditions on the size dependence of maximal algal density and biomass, Limnol. Oceanogr. 34 (1989) 1104–1108.

ACKNOWLEDGEMENTS

Chapter 1, in full, is a reprint of the material as it appears in Algal Research 2017. Sánchez-Alvarez, Eva L.; González-Ledezma, Grisele.; Bolaños-Prats, José A.; Stephano-Hornedo, José L., and Hildebrand, Mark. "Evaluating *Marinichlorella kaistiae* KAS603 cell size variation, growth and TAG accumulation resulting from rapid adaptation to highly diverse trophic and salinity cultivation regimes", Algal Research, 25:12-24, 2017. The dissertation author was the primary investigator and author of this material.

Chapter 2

Genome sequence determination for characterization of metabolic processes in *Marinichlorella kaistiae KAS603*

2.1 ABSTRACT

In this chapter, we determined and analyzed the chloroplast, mitochondrial, and nuclear genomes of *Marinichlorella kaistiae* KAS603. This analysis has enabled the identification of genes and gene products that participate in several metabolic processes that relate to the strain's productivity and adaptability. Predicted intracellular topology metabolic maps have given insights into predicted carbon flow within and between cell compartments. Starch is predicted to be stored in both the plastid (soluble form) and cytoplasm (granular form), which could be a way of enhancing short (plastid) or long (cytoplasm) term availability and avoiding competing metabolic flux in the same compartment. In contrast to the model green alga *Chlamydomonas reinhardtii*, glycolysis and gluconeogenesis are predicted to occur in the cytoplasm and AcCoA synthesis capability has been identified in multiple cellular compartments, including the ER and glyoxysome (which weren't predicted for C. *reinhardtii*). The glyoxylate cycle and photorespiration are suggested to interact in the glyoxysome and enhance energy production and carbon skeleton generation, as well as produce glycine, which is proposed to serve as an osmolyte during hypersalinity adaptation. Beta oxidation is predicted to occur in the glyoxysome as a possible way to rapidly manage and use stored lipids and biosynthetic precursors. Enzyme targeting comparison among various green algal species predicts that starch storage in non- oleaginous strains occurs in the plastid, while all oleaginous strains appear to have the capability of extraplastidial starch storage. A proposed main advantage of moving starch out of the chloroplast is that it

avoids the competing pathways of glycolysis and gluconeogenesis in the same compartment, lessening a metabolic and regulatory burden. Oleaginous strains also appear to have carbon flux in and through the endoplasmic reticulum (ER), which we hypothesize to provide another compartment to separate competing reactions or metabolites in order to facilitate FA and triacylglycerol (TAG) accumulation in *S. obliquus, C. sorokiniana, Mk KAS603 and Coccomyxa C-169.* We propose a multi-compartmentalized coordinate carbon flow among oleaginous algal strains for an efficient carbon supply for storage and partitioning to enable energy production (for growth and cell division), as well as fatty acid and TAG accumulation; which was not predicted to occur in the non-oleaginous strains analyzed.

2.2 INTRODUCTION

Marinichlorella kaistiae KAS603 (which will be abbreviated as *Mk* KAS603) cultivation flexibility and ability to accumulate lipid under different conditions was demonstrated on chapter 1. The goal of the current chapter is to determine and analyze the *Mk* KAS603 genome sequence to identify the genetic and metabolic capability of the species. Understanding the genetic basis for productivity and adaptability is essential for its use in the bioenergy field. This chapter focuses on identifying potential genes that are involved in core metabolic processes that relate to its lipid and biomass accumulation ability and capability to adapt to changing cultivation environments. *Mk* KAS603 nuclear, chloroplast and mitochondrial genomes were sequenced and bioinformatically processed at the Pellegrini lab at

UCLA, to enable our genomic analysis. We have examined gene content, gene clustering and genome size, to deduce metabolic capacity. In particular, we performed an in-depth analysis of the predicted intracellular targeting of enzymes, and compared these features with other sequenced chlorophytes. We have identified a gene repertoire that encodes various enzymatic functions that are likely to represent hallmarks of its adaptation to environmental changes and oleaginousness. This work is complemented by *Mk* KAS603 transcriptomic analysis, described in chapter 3.

2.3. METHODS

2.3.1 DNA isolation

Mk KAS603 was harvested during log phase, algae pellets were frozen overnight at -80°C, DNA for sequencing was isolated from ground frozen pellets by the proteinase K-phenol-chloroform method (Sambrook, 2001), and purified on CsCl/Hoechst 33258 dye gradients (Jacobs *et al*, 1992). Three bands, representing nuclear (nDNA), chloroplast (ctDNA) and mitochondrial (mtDNA) DNA were isolated from the gradient.

2.3.1 Sequencing

After RNaseA treatment, each DNA sample was sonicated using a Covaris S_2 focused ultrasonicator to obtain 100 to 300 bp fragments (Duty Cycle = 10%; Intensity = 5; Cycles/Burst = 200; for 6 minutes). DNA libraries were prepared using Illumina TruSeq DNA Sample Prep Kit v2 according to manufacturer's instruction.

To account for size differences, the three barcoded libraries were pooled in this ratio 1:1/125:1/500 = nDNA:chDNA:mtDNA, before sequencing.

One paired-end HiSeq 150 run of the pooled DNA was performed, using the HiSeq 2500 system from Illumina.

2.3.2 Assembly

Pooled reads were separated into mitochondrial, chloroplast, and genomic data by barcodes. Each data set was assembled into "contigs" comprising overlapping, adjacent reads using the ABySS assembly tool. Default parameters were used and adjusted to program output. Assembled contigs were filtered by length and retained if longer than 3,000 base pairs (mitochondrial and chloroplast) or 10,000 base pairs (nuclear). Contig length distributions were calculated by plotting the length of assembled contigs against the total base pairs in all contigs less than or equal to that length.

2.3.3 Annotation

Mitochondrial genome. A long circular contig was assembled for the mitochondrial genome, and its genes were manually annotated using BLAST (Altschul *et al*, 1990). Genes were annotated based on their similarity with *Coccomyxa sp* C-169 and *Chlorella sp. ArM0029B*.

Plastid genome. A long circular contig was assembled for the chloroplast genome, and its genes were manually annotated using BLAST. As comparative

plastid genomes, *Chlorella sp*. ArM0029B, *Parachlorella kessleri*, *Chlorella variabilis NC64A* and *Chlorella vulgaris* were used.

Nuclear genome. The *Mk* KAS603 nuclear genome was annotated, using 4 approaches:

1. From long contig nuclear assemblies, genes were identified through BLAST.
2. Genes were manually annotated from long contig nuclear assemblies (>550 bp) using the program DNA subway (Hilgerth *et al*, 2014)
3. A first round of gene predictions was generated from long contig assemblies (>550 bp), using the FGENESH (Solovyev *et al*, 2006) gene prediction algorithm and the *Chlamydomonas reinhardtii* training set (this was the closest related training set available).
4. A second round of gene predictions for the *Mk* KAS603 assembly was generated using AUGUSTUS (Stanke *et al*, 2004) under self-training mode (using PASA [Haas *et al*, 2003]) and BLAT (Kent W, 2002), and incorporating the transcriptome assembly (described in chapter 3).

All the contigs generated were concatenated and displayed on the *Mk* KAS603 Genome Browser (http://genomes.mcdb.ucla.edu/cgi-bin/hgGateway?hgsid=66493)

2.3.4 Nuclear genome size estimation using flow cytometry

The relative nuclear DNA content of *Mk* KAS603 was estimated by associating DNA fluorescence to genome size among known genome size microalgal strains (3 positive controls) and *Mk* KAS603 (experimental value), using linear

regression analysis. Diploid *Thalassiosira pseudonana (Tp)* (34.5 Mbp), diploid *Phaeodactylum tricornutum (Pt)* (27.4 Mbp), haploid *Chlorella protothecoides (Cp)* (22.9 Mbp) and haploid *Mk* KAS603, were analyzed. *Mk* KAS603 was cultivated in modified Shuisheng-4 medium (MS4) and *Cp* was cultivated in Bristol medium with a temperature of 25 °C. The culture medium for *Tp* and *Pt* was ASW (Darley and Volcani, 1969) with a temperature 18°C. All strains were cultured in continuous light with 150 μmol photons m^{-2} $s^{-}1$. Cells (1-2 x 10^6 total) for each microalgal strain were harvested in mid-exponential phase and concentrated by centrifugation (15 min at 2000 x g). The pellet was resuspended in 15 mL of 100% methanol at 4°C for 48 h at -20^{o} C for fixation and to extract chlorophyll. Samples were centrifuged and resuspended in 3 mL PBS and treated with RNase A (0.3 mg ml^{-1}) at 37°C for 60 minutes. The DNA was stained with SYBR® Green I (1X final concentration from a 100X SYBR® Green I stock made in DMSO, Life Technologies^{TM}), for 20 min. Stained samples were kept on ice in the dark until run on the flow cytometer. A Becton Dickinson Influx sorting flow cytometer was used. Cells were excited with a 488-nm laser, and emission was evaluated at 530/40 nm. Duplicate samples of 10,000 to 20,000 cells were analyzed for each strain. Flow cytometric data for all strains was analyzed by associating mean fluorescence to genome size and curve fitting in Microsoft Excel using linear regression analysis. The nuclear genome size of *Mk* KAS603 was predicted based on comparative DNA fluorescence with the other known genome size microalgae.

2.3.5 Genome comparison and visualization or orthologous clusters

We performed an analysis of orthologous clusters across green algal species based on the protein sequence of three *Chlorella* strains by using OrthoVenn (Wang *et al*, 2015). A modified OrthoMCL heuristic approach is used by OrthoVenn to identify ortholog groups by BLASTP alignment. A Venn diagram was constructed to identify clusters of genes in three different species: *Mk* KAS603, *A. protothecoides* and *C. variabilis* NC64A.

2.3.6 Topological analysis of metabolic pathways

The AUGUSTUS gene models from the assembled genome and the assembled transcriptome (described on chapter 3) were translated. Two peptide sets (FASTA file containing amino acid sequences/protein models) were created. Using standard protein BLASTP (Altschul *et al*, 1990) reference proteins were aligned, and proteins of interest were identified. Predicted subcellular localization was assigned to each protein identified based on output from 9 different prediction programs: TargetP 1.1 (Emanuelsson *et al*, 2007), Predotar 1.04 (Small *et al*, 2004) SignalP 3.0 (Bendtsen *et al*, 2004), Mitoprot II 1.01 (Claros and Vincens, 1996) ChloroP 1.1 (Emanuelsson *et al*, 1999), PredPlantPTS1 (Reumann *et al*, 2012), PTS1 (Neuberger *et al*, 2003), Plant-mPLoc 2.0 (Chou and Shen, 2010) and PredAlgo 1.0 (Tardif *et al*, 2012). When 2 or more programs predicted the same localization, with other considerations, it was considered strong support for a predicted target. When no organellar targeting prediction occurred (denoted as Elsewhere or Other in Tables 4-6), it defaulted into

cytoplasmic targeting. It has been determined in terrestrial plants and chlorophytes that chloroplast localized proteins can have both a robust chloroplast and mitochondrial predicted targeting (Mitschke *et al*, 2009, Peeters and Small, 2001; Carrie *et al*, 2009). Truly mitochondrially-targeted proteins will have a high mitochondrial prediction but low chloroplast prediction. For proteins with both high chloroplast and mitochondrial predicted targeting in our dataset, after considering the metabolic processes the protein was involved in, we generally assigned the protein to the chloroplast (with two exceptions: ME [g9349g] and accA [g6872g] were targeted to both chloroplast and mitochondrion, see tables 5-6). Dual targeting to more than one organelle is rare (Peeters and Small, 2001), therefore for genes with high predicted targeting to more than one organelle, we mostly defaulted to the one most likely. N-terminal extension for plastid predicted protein models was supported by performing alignments with Clustal Omega (Sievers *et al*, 2011) of *Mk* KAS603 and *C. reinhardtii* starch, glucose metabolism and pyruvate hub enzyme model sequences. Topology maps for metabolic pathways of interest in *Mk* KAS603 were constructed. A specific glycolysis, gluconeogenesis, starch metabolism and pyruvate hub comparison between *Mk* KAS603 and other oleaginous and non-oleaginous green microalgae: *Chlamydomonas reinhardtii*, *Coccomyxa subellipsoidea C-169*, *Scenedesmus obliquus*, *Chlorella variabilis* NC64A, and *Chlorella sorokiniana*, was developed.

2.4 RESULTS

2.4.1 DNA isolation

Three DNA bands were isolated from CsCl/Hoechst 33258 dye gradients (Fig 2-1): lower (nuclear DNA, nDNA), middle (chloroplast DNA, ctDNA) and upper (mitochondrial DNA, mtDNA).

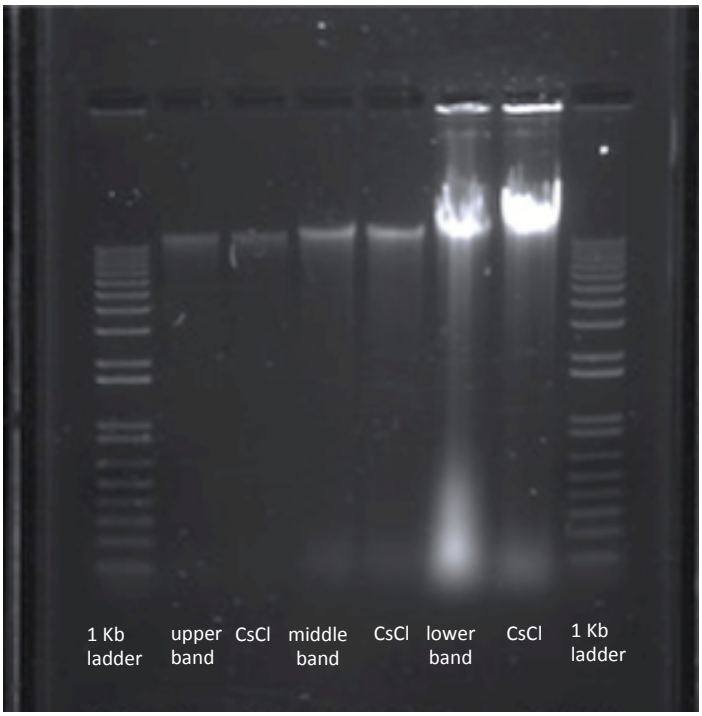

Figure 2-1. Electrophoresis gel run for *Mk* KAS603 DNA purified from CsCl/Hoechst 33258 dye gradients. 3 bands were isolated. The upper band corresponds to mitochondrial DNA, the middle band to chloroplast DNA and the lower band to nuclear DNA.

2.4.2 Mitochondrial Genome (Table 2-1)

The sequenced mitochondrial genome size for *Mk KAS603* was 75,113 bp, which is slightly larger than *Coccomyxa* sp. C-169 (65,497 bp) (Smith, 2011) and *Chlorella sp*. ArM0029B (65,049 bp) (Jeong, 2014). All three species have a circular mitochondrial genome. A total of 38 mt protein coding genes and 3 rRNA genes were identified in *Mk* KAS603 (Table 2-1). The gene number is higher than *Chlorella sp.* (32 mt protein coding genes) and lower than *Coccomyxa sp*. (59 mt protein coding genes). 5 genes were repeated (Cox 1, nad1, Cob, nad4L and nad7 gene) in *Mk* KAS603. There are no reports of repeated mitochondrial genes in the other two algal strains. Both *Chlorella sp*. ArM0029B and *Mk KAS603* have a group I intron with an ORF containing a LAGLIDAG motif in the cox1 gene. LAGLIDAG participates in intron propagation and/or intron splicing (Burger and Nedelcu, 2012).

Genes of interest related to energy generation are (Table 2-1):

- nad 1, 2, 3, 4, 4L, 4M, 5, 6, 7, 9, 10 genes which code for 11 subunits of the electron carrier NADH dehydrogenase, this enzyme is part of complex I of the electron transport chain and catalyzes the transfer of electrons from NADH to ubiquinone, that is the initial step in the respiratory chain.
- Cob gene, that codes for the first electron carrier of complex III called cytochrome b, also part of the respiratory chain
- Cox 1,2,3, genes that code for 3 subunits of cytochrome c oxidase, which catalyzes the transfer of electrons from cytochrome c to

molecular oxygen, reducing it to H_2O in complex IV of the respiratory chain

- atp 1,4,6,8,9 genes that code for subunits of ATP synthase, which catalyzes the synthesis of ATP from ADP by chemiosmosis
- Tat C gene whose product is involved in transmembrane transport and cytochrome C maturation

Table 2-1 *Mk KAS603* list of genes from mitochondrial DNA

Gene Products	Genes
Large subunit ribosomal proteins	rpl 5, 6, 10,
Small subunit ribosomal proteins	rps 2, 3, 4, 7, 8, 10, 11, 12, 13, 14
Cob complex	Cob
Cytochrome c oxidase	Cox 1, 2, 3
NADH dehydrogenase	nad 1, 2, 3, 4, 4L, 4M, 5, 6, 7, 9, 10
DUF genes/unkown function	DUF 4323, 4324
ATP synthase	atp 1,4,6,8,9
Ribosomal RNAs	rns, rnl, rrnS
Translocase	Tat C

2.4.3 Chloroplast genome (Table 2-2)

This genome sequence was most similar to *Chlorella sp*. ArM0029, based on overall sequence conservation and gene arrangement. The size of *Mk KAS603* chloroplast genome was 145,626 bp, which is larger than *Chlorella sp*. ArM0029B (119,989 bp) (Jeong, 2014), *Chlorella variabilis* NC64A (124,597bp) (Accession no. NC_015359), and *Parachlorella kessleri* (123,994 bp) (Turmel *et al*, 2009b); and smaller than *Chlorella vulgaris* (150,613 bp) (Wakasugi *et al*, 1997). *Mk* KAS603 lacks an inverted repeat (IR) region. The biological role of the IR region is likely increased gene dosage for ribosomal components (Lang, 2012). The IR region is present in *P. kessleri* (Turmel et al, 2009b), whereas *C. vulgaris* (Wakasugi *et al*, 1997), *Chlorella sp*. ArM0029B (Jeong, 2014) and *C. variabilis* NC64A, lack an IR region. A total of 78 genes were identified in the *Mk* KAS603 plastid genome**,** of which 75 are protein-encoding genes, 31 are tRNA genes and 3 are rRNA genes (Table 2-2). This is very similar to *Chlorella sp*. ArM0029, whose chloroplast genome contains 79 genes encoding proteins, 32 tRNA genes 3 rRNA genes. As displayed in Fig. 2-2, three distinctive arrangements of plastid genomes between *P. kessleri*, *Chlorella sp*. ArM0029 and *Mk* KAS603 were found: 1) 45 genes have identical arrangement (these include rRNA clusters, atp gene clusters, ribosomal protein clusters and photosystem II clusters) in the 3 strains (Fig. 2-2A), 2) the rpoB-rpoC1-rpoC2-rbcL gene cluster (Fig. 2-2B) has an inverted order in *Chlorella sp*. ArM0029 and *Mk* KAS603. 3) *P. kessleri* is missing the rbcL gene (Fig. 2-2B) from the formerly mentioned cluster. Identified genes of interest related to carbon fixation

and lipid metabolism in *Mk* KAS603, are the rbcL gene, that codes for the large subunit of RubisCO enzyme that is responsible for the catalysis of the first step of carbon fixation, and the accD gene, which codes for a subunit of Acetyl-CoA carboxylase, an enzyme that catalyzes the first committed step of fatty acid synthesis and also plays a role in fatty acid degradation (Guarnieri, 2011). rbcL and accD genes are considered critical genes for lipid metabolism in *Chlorella pyrenoidosa* (Fan, 2014).

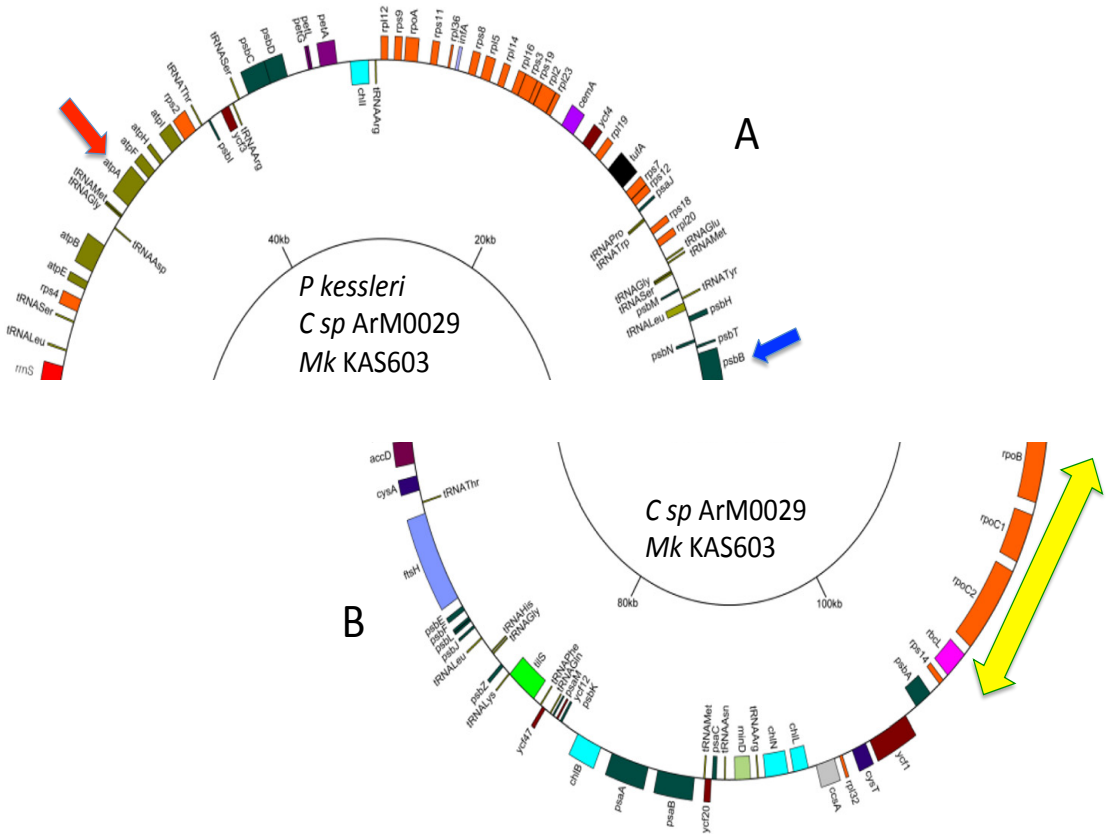

Figure 2-2. Plastid genome arrangement among three Chlorella strains. A. *P kessleri*, *C sp* ArM0029 and *Mk* KAS603 strains have 45 genes in identical order (from red to blue arrows). B. C sp ArM0029 and *Mk* KAS603 strains have a four gene cluster (measured by yellow arrow) in inverted gene order, while *P kessleri* only has 3 genes (orange boxes) forming part of this cluster.

Table 2-2. *Mk KAS603* list of genes from chloroplast DNA

Gene Products	Genes
Large subunit ribosomal proteins	rpl 2, 5, 7, 12, 14, 16, 19, 20, 23, 32, 36
Small subunit ribosomal proteins	rps 2, 3, 4, 7, 8, 9, 11, 12, 14, 18, 19
RNA polymerase	rpo A, B, C1, C2
Cytochrome b6/f	pet A, G
Heme attachment to plastid cytochrome c	ccs A
Photosystem I	psa A, B, C, I, J
Photosystem II	psb A, B, C, D, E, F, H, I, J, K, L, M, N, T,Y,Z
ATP synthase	atp A, B, E, F, H, I
RuBisCO	rbcL
Ribosomal RNAs	rrf, rrl, rrs
Lipid metabolism	accD
ABC transporters	cys A, cys T
Chloroplast division	ftsI, ftsW, minD
Cell cycle control, cell division	Mes J
Chlorophyll biosynthesis	chl B, I, L, N
Translation factors	inf A, EFTU II (tuf gene), EFTU III, tuf A
Transfer RNAs	tRNA T (ggu)

2.4.4 Nuclear genome

Genome sequencing and flow cytometry were used to estimate the genome size and number of genes of *Mk* KAS603. The results for both methods were similar. The estimated *Mk KAS603* genome size by flow cytometric was 45 Mb (Fig. 2-3), and the estimate by sequencing was 50 Mb. The genome size of *Mk* KAS603 should be smaller than 50 Mb (as was estimated by flow cytometry) because, as suggested by Gao *et al*, 2014, sequencing can provide a larger size estimate than flow cytometry due to sequencing errors.

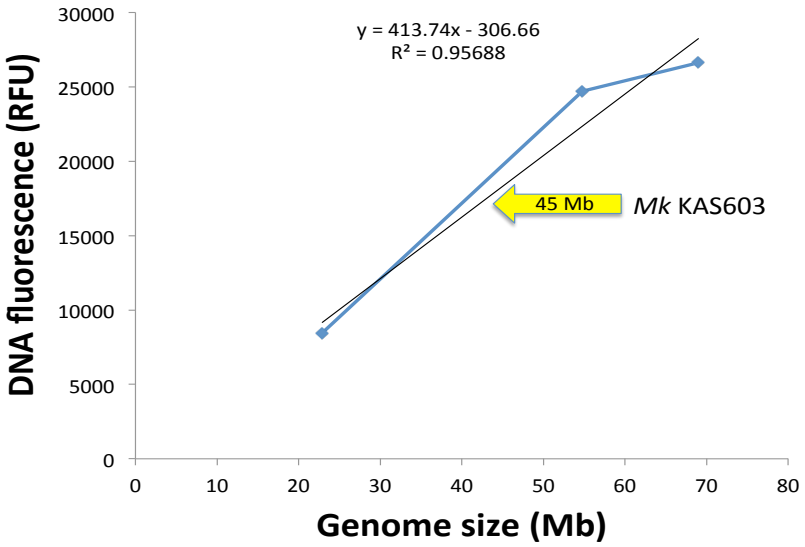

Figure. 2-3 Genome size estimation of *Mk* KAS603 determined by flow cytometry. Yellow arrow depicts the predicted size of the genome of *Mk* KAS603 at 45 Mb based on DNA fluorescence. A linear regression method was used to fit fluorescence/genome size data. Control (genome size) species were *Thalassiosira pseudonana* (34.5 Mb), *Phaeodactylum tricornutum* (27.4 Mb) and *Chlorella protothecoides* (22.9 Mb).

Two genome assembly rounds were performed. FGENESH was used in the first round and it generated fewer gene predictions than other tools, such as AUGUSTUS (Table 2-3) (used in the second round). AUGUSTUS was used under self-training mode with PASA (Hass *et al*, 2003) that stands for Program to Assemble Spliced Alignments, which incorporates assembled RNA data in the training step. BLAST-Like Alignment Tool or BLAT (Kent, 2002), produced a "hints" file (providing information for splice sites and transcript starts/stops) for alignments of the RNA data. Most contigs in *Mk* KAS603 were relatively gene-dense and quite a few were nearly saturated with gene content, suggesting that a number of genes predicted to be fused to other genes using FGENESH were split into distinct genes using AUGUSTUS. Our initial count (using FGENESH) of 4,629 genes, was considered very low compared the other mentioned species: *C. protothecoides* has a gene count of 7,039, *C. variabilis* NC64A has 9,791, *C. subellipsoidea* C-169 has 9,851, and *C. reinhardtii* has 15,143 (Gao *et al*, 2014). A second comparative run with AUGUSTUS (using contigs and transcripts) improved the original count, generating a total of 9,373 genes.

As displayed in Table 2-3 and Fig 2-4 (below), the number of gene predictions has nearly doubled and while the average protein length has hardly changed, the gene size has been nearly cut in half. The discrepancy is mostly due to the combination of much shorter introns and fewer introns per gene, along with slightly longer exons.

The two-fold increase in gene models is mostly due to splitting (presumably improperly fused) FGENESH models rather than the addition of new genes in previously gene-deplete regions of the genome. The plot of contig lengths vs. gene mass (Fig. 2-4) supports this (there is a slight decrease in gene density with the AUGUSTUS models, despite more models [probably due to the regions between fused genes no longer being called a gene]), as does the amount of overlap between the two prediction sets. Taking their genomic locations into account, the typical FGENESH gene is covered by 1.9 AUGUSTUS genes (excluding positionally novel AUGUSTUS genes).

Table 2-3. *Mk* KAS603 gene/protein assembly comparison of AUGUSTUS and FGENESH.

	AUGUSTUS	FGENESH
Total # genes	9373	4629
Total # genes on negative strand	4645 (49.6% of all genes)	2332 (50.4%)
Average gene length (nucleotides)	2534	5598
Average exon length (nucleotides)	178 (of 59312 exons)	132 (40167)
Average # of exons per gene	6.3279	8.6773
Average # of introns per gene	5.3279	7.6773
Average intron length (nucleotides)	263 (of 49939 introns)	487 (35538)
Average protein length (aa)	374	380

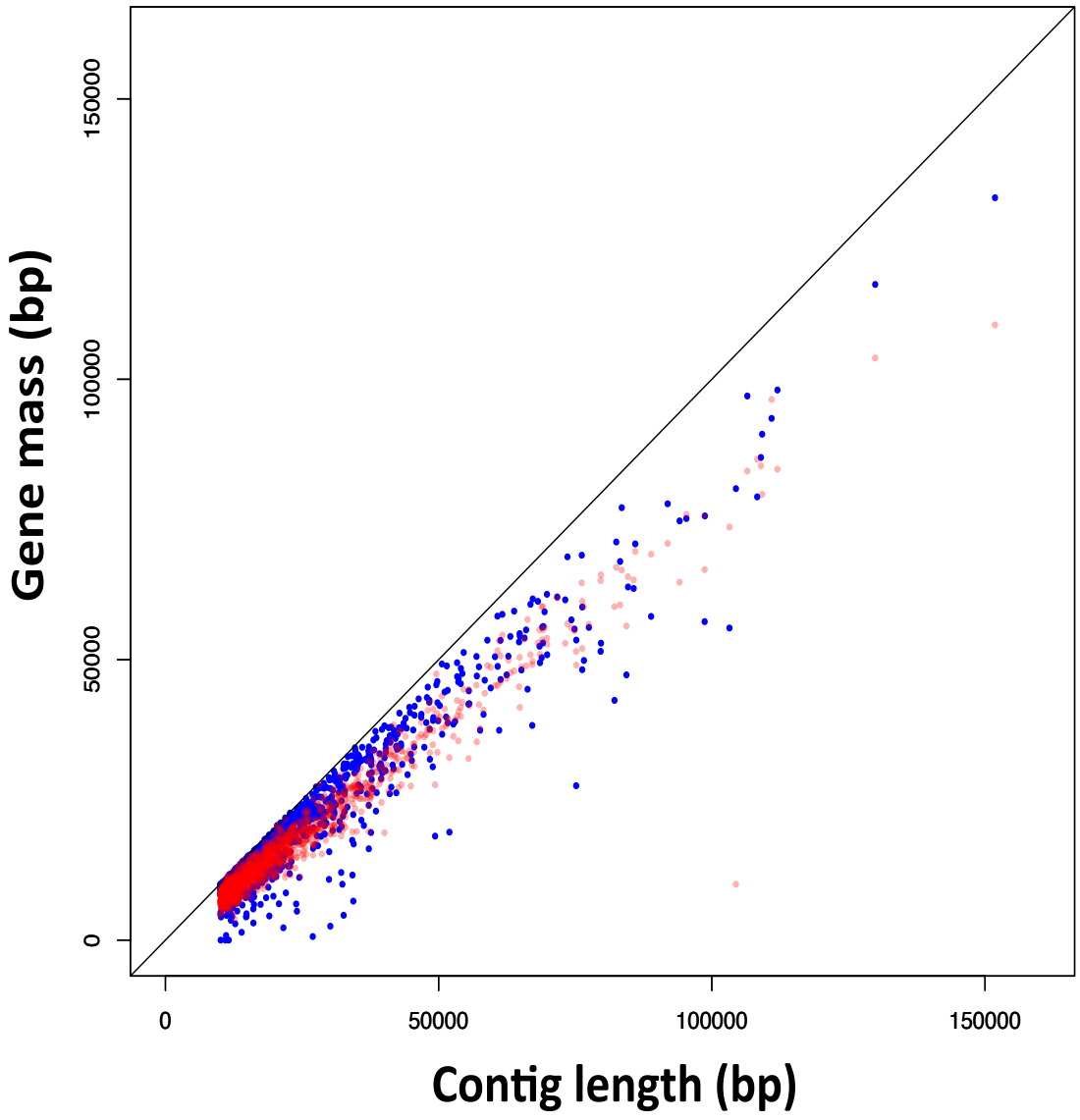

Figure 2-4. *Mk* KAS603 correlation of gene mass and length of contigs. Blue color corresponds to FGENESH and red color to AUGUSTUS.

We compared the estimated genome sizes with the total number of genes determined in four reference genomes and estimated that 9300 genes should be present in *Mk* KAS603 (Fig. 2-5) if the gene density was similar.

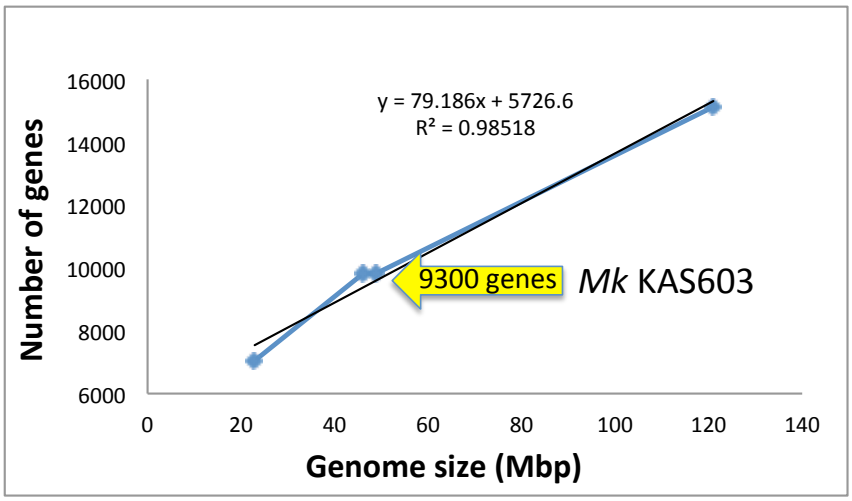

Figure 2-5. Number of genes vs genome size in *Mk* KAS6003. Yellow arrow depicts the predicted number of genes based on the size of the genome of *Mk* KAS603 at 45 Mb. A linear regression method was used to fit fluorescence/genome size relative to gene number data. Control (genome size/gene number) species were *C. protothecoides* sp. 0710 (22.9 Mbp/7039 genes), *C. variabilis NC64A* (46.2Mbp/9791 genes), *C. subellipsoidea* C-169 (48.8 Mbp/9851 genes), and *C. reinhardtii* (121 Mbp/15143 genes).

The genome sequencing estimated genome size for *Mk KAS603* (50 Mbp), is 22.1 Mbp larger than *Chlorella protothecoides* sp. 0710 (22.9 MB), 3.8 Mbp larger than *Chlorella variabilis* NC64A (46.2 Mb), 1.2 Mb larger than *Coccomyxa subellipsoidea* C-169 (48.8 Mbp), and 71 Mbp smaller than *Chlamydomonas reinhardtii* (121 Mbp) *(*Gao *et al,* 2014*). Mk KAS603* nuclear genome sequences have a lower GC content (56.78%) than *C. variabilis* NC64A (67.2%), *C. protothecoides* sp. 0710 (63%), *C. reinhardtii* (64%) and a slightly higher GC content than *C. subellipsoidea* C-169 (53%). We also studied shared gene families among *Mk* KAS603, *C. variabilis* NC64A and *A. protothecoides.* As displayed in Fig. 2-6 (Venn diagram), 3379 orthologs were identified among the 3 strains, suggesting their conservation in the lineage after speciation. Additionally, there were 124 clusters specific to *Mk* KAS603, 50 to *A. protothecoides* and 234 to *C. variabilis* NC64A; such gene clusters are likely to be within multiple genes or in-paralog clusters. The presence of in-paralog clusters suggests a possible lineage specific gene expansion in these gene families for each algal strain, their involvement in biological processes is unknown due to non-specific BLASTP/Swiss-Prot/GO annotation hits. In *Mk* KAS603 the number of proteins per cluster ranged from 2 to 27, in *C. variabilis* NC64A from 2 to 19 and in *A. protothecoides* from 2 to 6 proteins per cluster, which is indicative of a broad similarity range. *Mk* KAS603 and *C. variabilis* NC64A share 4659 gene clusters, *C. variabilis* NC64A and *A. protothecoides* share a similar number (4536 orthologs), while *A. protothecoides* and *Mk* KAS603 share only 3724, which is indicative of less gene conservation with *A. protothecoides*. Genome

comparison and ortholog gene analysis supports the accuracy of our genome size estimations. Using the AUGUSTUS pipeline to incorporate the transcriptome assembly, successfully improved the gene models for *Mk* KAS603.

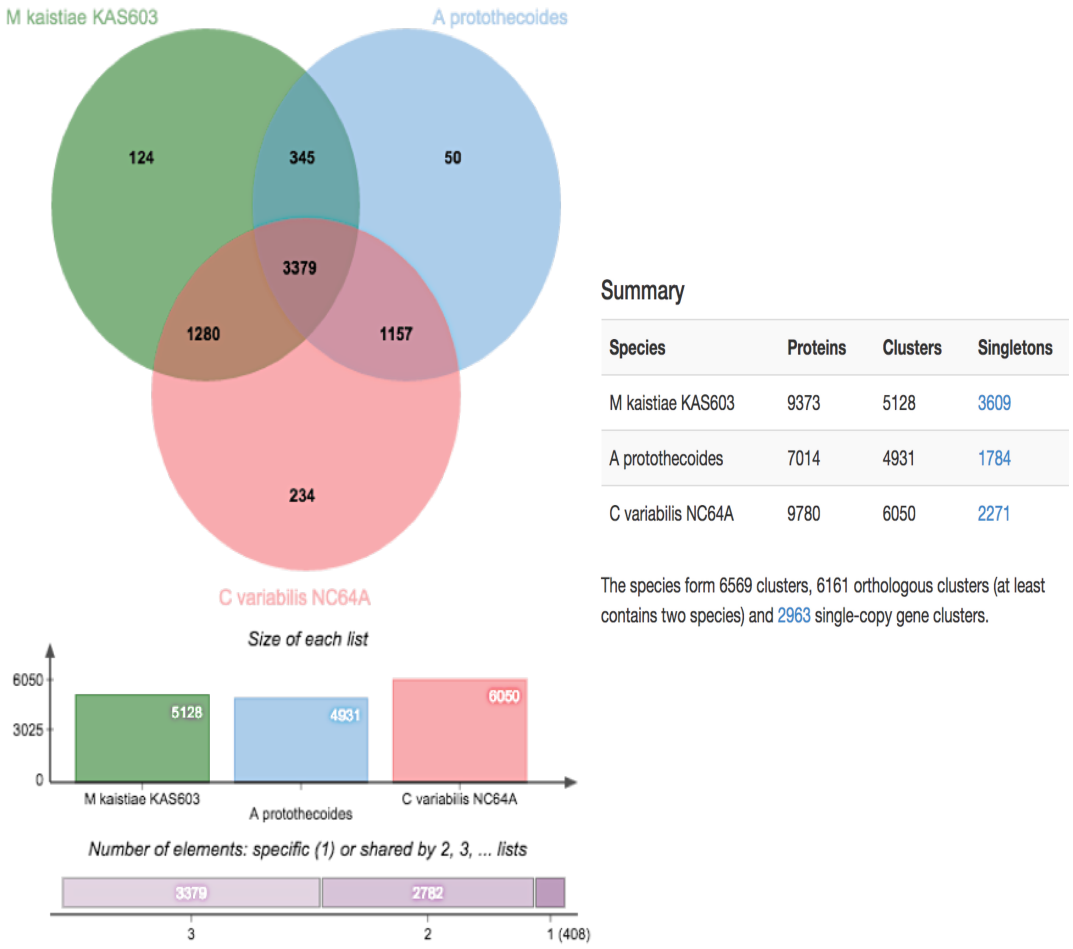

Figure 2-6. Genome comparison among 3 members of the family Chlorellacea*:* *Marinichlorella kaistiae* KAS603, *Auxenochlorella protothecoides* and *Chlorella variabilis* NC64A. The numbers in the Venn diagram represent the number of orthologous clusters that *Mk* KAS603 shares with the other 2 algal species. 3379 orthologous clusters were formed based on the protein sequences from the 3 species. The summary table displays the number of clusters in each genome.

2.4.5. Mk KAS603 cellular localization of metabolic pathways of interest

Targeting analysis of proteins of interest was performed by using multiple subcellular localization prediction programs. If two or more programs agreed, with other considerations, that decided targeting. In cases where gene models were truncated, they were assigned in Figs. 2-7 to 2-16 with question marks, a definitive location could not be predicted, but the possible locations are shown. Full names of enzymes that are abbreviated on Figs. 2-7 to 2-16, are mentioned in supplemental tables.

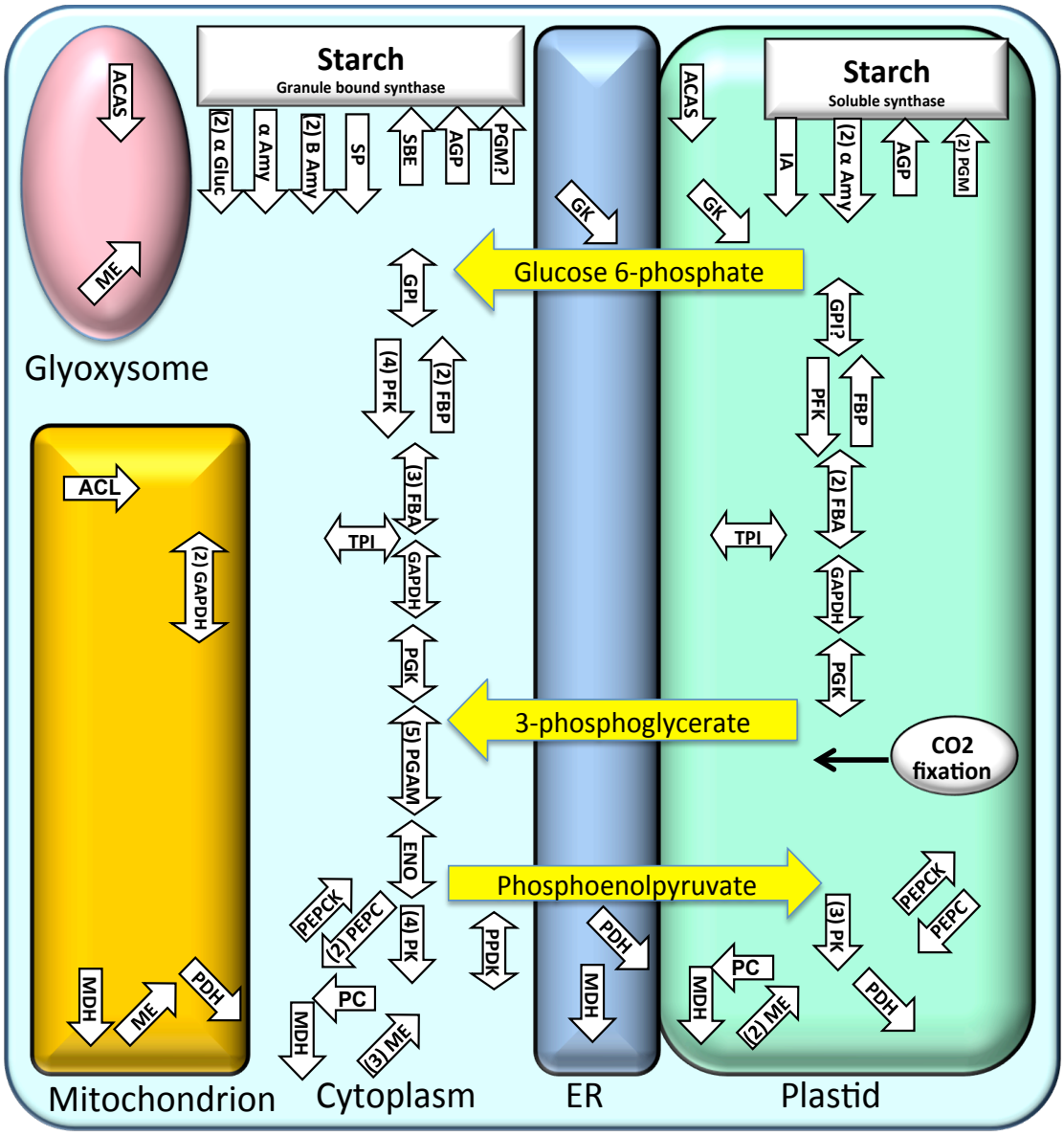

Figure 2-7. *Mk* KAS603 predicted Glycolysis, glucoenogenesis, starch metabolism and pyruvate hub topology map. Double and single white arrows denote bidirectional and unidirectional enzymatic reactions in metabolic pathways. Black letters inside white arrows are enzyme abbreviations (see enzyme names list in supplemental table 1). The number of identified enzyme isoforms is denoted in parenthesis. Question-marked sequences were truncated, and predicted location could not be determined, but possible locations are denoted. Yellow arrows denote proposed movement of carbon compounds among compartments.

Glycolysis and the pyruvate hub.

As displayed in table 2-4 and Fig. 2-7, a partial (upper half) glycolysis pathway was predicted to be localized in the plastid and a complete glycolysis pathway was predicted to be localized in the cytoplasm with the exception of glucose kinase, which was predicted to be targeted to the chloroplast and ER. This enzyme initiates glycolysis by catalyzing the conversion of glucose to glucose-6-phosphate. The lack of its localization in the cytoplasm suggests that glucose-6-phosphate might be exported to the cytoplasm and incorporated into the glycolytic route in this compartment. PFK catalyzes a key unidirectional regulatory step in glycolysis, and from a total of 5 gene copies found in the *Mk* KAS603 genome, 4 had predicted cytoplasmic localization and 1 had predicted plastid localization. Two genes encoding GPI, the preceding step in glycolysis, were identified in the genome; one of which was full length and predicted to be targeted to the cytoplasm, and the other whose sequence was truncated, therefore targeting could not be unambiguously assigned, hence the question mark on its possible plastid localization in Fig. 2-7. Data in support of a plastid copy of GPI is the predicted presence of a PFK, which is an indicator of plastid glycolysis. For glycolysis to occur to process starch breakdown products, both GPI and PFK would need to be present. The presence of 3 isoforms of FBA targeted to the cytoplasm, along with TPI, GAPDH, PGK, 5 PGAM isoforms, ENO and 4 PK isoforms, provides strong evidence that a complete glycolysis pathway is present in the cytoplasm. ENO and PGAM are not predicted to be localized in the plastid, as has been documented in other green algae (Polle *et al*,

2014), indicating that the lower half of glycolysis is not complete in this organelle. Based on the targeting analysis, 3-phosphoglycerate is predicted to be exported from the plastid and enter the cytoplasm for further processing by cytoplasmic PGAM until a final glycolytic product (pyruvate) is synthesized through this route. The cytoplasmic pathway in Mk KAS603 provides a possible means of glycolytic processing of starch breakdown products outside of the plastid, which may be advantageous in avoiding competing reactions in the same organelle. In terms of the pyruvate hub (Table 2-5 and Fig. 2-7), three PK isoforms were predicted to be localized in the plastid and four isoforms of PK were predicted to be localized in the cytoplasm, suggesting that pyruvate might be produced in both compartments. For the plastid, PEP would have to be imported from the cytoplasm. PDH is predicted to be localized in the mitochondria, ER and plastid; indicating a possible pyruvate conversion to AcCoA in those 3 compartments. Malate dehydrogenase was predicted to be localized in the mitochondria, cytoplasm, ER and plastid; suggesting that malate can be reversibly converted to oxaloacetate in four compartments. PEPC was predicted to be localized in the cytoplasm and plastid, suggesting oxaloacetate production in the plastid and cytoplasm. Malic enzyme was predicted to be localized in the glyoxysome, mitochondria, cytoplasm and plastid, indicating a possible reversible malate decarboxylation to pyruvate (with NADPH production) in four compartments. The pyruvate hub targeting results for the mitochondria are consistent with AcCoA (PDH), malate (ME) and oxaloacetate (MDH) involvement in the TCA cycle, and are likely to contribute to energy production (NADH, $FADH_2$ and ATP or

GTP) through this cycle. In the plastid, the enzymes could generate pyruvate and subsequently AcCoA as precursors for fatty acid synthesis. For the cytoplasm, a complete pyruvate hub is suggested to present, enabling the generation of pyruvate and OAA to possibly contribute to gluconeogenesis (and starch accumulation), which is also predicted to occur in the cytoplasm (see below). In the ER, AcCoA (PDH) for protein acetylation and aminoacid synthesis is suggested to be present, as well as oxaloacetate (MDH) for a possible export to the cytoplasm toward gluconeogenesis.

Table 2-4. *Mk* KAS603 protein localization prediction for enzymes involved in starch metabolism, glucose anabolism and catabolism. 9 different prediction programs were used. The consensus predicted compartment is denoted in the predicted targeting column. Highlighted green sections denote a chloroplast predicted enzyme, light blue denotes cytoplasm, orange denotes mitochondrion, dark blue denotes ER. Lower case red "g" denotes protein identification based on the genome assembly, otherwise it is based on the transcriptome assembly. Y = High chloroplast signal prediction, ppPTS1 = PredPlantPTS1 prediction program, E = Elsewhere (cytoplasm), O = other (cytoplasm), P = chloroplast, M = Mitochondrion, C = Cytoplasm, ER = Endoplasmic Reticulum, No = No ER signal predicted, Yes = ER signal predicted, SP = Secretory pathway, Nt = Not targeted to peroxisome (glyoxysome), t = targeted to peroxisome (glyoxysome). Value numbers represent the total prediction score to the corresponding organelle. Full name of proteins mentioned in this table, are listed in supplemental table S2-1.

Pathway	Protein name	Protein ID	Predicted targeting	ChloroP	Mitoprot	TargetP	Predotar	Signal P	PredAlgo	PTS1	ppPTS1	Plant-mPLoc
Starch Anabolism	SSs	g15135	Chloroplast	Y, 0.520	0.225	P, 0.836	E, 0.89	No	Other	Nt	Nt	Chloroplast
	SSg	g16066	Cytoplasm	0.429	0.0048	O, 0.786	E, 0.98	No	Other	Nt	Nt	Chloroplast
	AGP	g16078	Chloroplast	0.475	0.7819	O, 0.427	E, 0.80	No	Other	Nt	Nt	Chloroplast
	AGP	g9209	Truncated/no prediction									
	AGP	g9193	Cytoplasm	0.45	0.842	M, 0.402	E, 0.82	No	Other	Nt	Nt	Chloroplast
	SBE II	g14640	Cytoplasm	0.435	0.0459	O, 0.828	E, 0.99	No	Other	Nt	Nt	Chloroplast
	PGM	g16057	Truncated/no prediction									
	PGM	g8899	Chloroplast	0.499	0.691	P, 0.422	E, 0.98	No	P, 0.9295	Nt	Nt	Chloroplast
	PGM	g3736g	Chloroplast	Y, 0.518	0.9211	P, 0.945	P, 0.71	No	P, 0.561	Nt	Nt	Chloroplast
Starch catabolism	β amy	g2918	Cytoplasm	0.437	0.3429	O, 0.707	E, 0.95	No	Other	Nt	Nt	Chloroplast
	β amy	g4130g	Cytoplasm	0.482	0.8329	M, 0.435	M, 0.52	No	Other	Nt	Nt	Cytoplasm
	β amy	g8637g	Truncated/no prediction									
	IA	g959g	Chloroplast	Y, 0.532	0.9315	P, 0.812	P, 0.32	No	P, 0.5319	Nt	Nt	Chloroplast
	α amy	g6337g	Chloroplast	Y, 0.533	0.7386	P, 0.812	P, 0.35	No	P, 1.059	Nt	Nt	Chloroplast
	α amy	g7903g	Chloroplast	Y, 0.556	0.998	P, 0.790	M, 0.43	No	P, 1.100	Nt	Nt	Chloroplast
	α amy	g15090	Cytoplasm	0.426	0.639	O, 0.779	E, 0.97	No	Other	Nt	Nt	Extracellular
	α gluc	g7514g	Cytoplasm	0.451	0.0046	O, 0.824	E, 0.93	No	Other	Nt	Nt	Cell wall
	α gluc	g15306	Cytoplasm	0.434	0.1223	O, 0.753	E, 0.94	No	Other	Nt	Nt	Cell wall
	SP	g15447	Cytoplasm	0.469	0.0074	O, 0.920	E, 0.99	No	Other	Nt	Nt	Chloroplast
Glycolysis	GK	g3427g	ER	0.436	0.9637	M, 0.855	ER, 0.86	Yes, 0.954	M, 0.426	Nt	Nt	ER
	GK	g6066g	Chloroplast	0.479	0.8906	M, 0.731	M, 0.66	No	P, 1.448	Nt	Nt	Chloroplast
	GPI	g2310g	Cytoplasm	0.431	0.0893	O, 0.845	E, 0.97	No	Other	Nt	Nt	Cytoplasm
	GPI	g12427	Truncated/no prediction									
	PFK	g4546g	Cytoplasm	0.437	0.0809	S, 0.862	E, 0.89	No	Other	Nt	Nt	Cytoplasm
	PFK	g4547g	Cytoplasm	0.437	0.76	O, 0.629	E/M, 0.50/0.5(	No	SP, 0.22	Nt	Nt	Cytoplasm
	PFK	g2763g	Cytoplasm	0.454	0.0089	O, 0.895	E, 0.99	No	Other	Nt	Nt	Cytoplasm
	PFK	g8531	Chloroplast	Y, 0.514	0.7953	M, 0.661	E, 0.81	No	C, 1.222	Nt	Nt	Cytoplasm
	PFK	g14327	Truncated/no prediction									
	PFK	g9975	Truncated/no prediction									
	PFK	g8804	Cytoplasm	0.433	0.0314	O, 0.553	E, 0.98	No	Other	Nt	Nt	Cytoplasm
	FBA	g12970	Cytoplasm	0.431	0.1016	O, 0.437	E, 0.91	No	Other	Nt	Nt	Cytoplasm
	FBA	g13839	Cytoplasm	0.432	0.023	O, 0.838	E, 0.99	No	Other	Nt	Nt	Cytoplasm
	FBA	g8684	Cytoplasm	0.427	0.2565	O, 0.862	E, 0.99	No	Other	Nt	Nt	Mitochondrion
	FBA	g15917	Chloroplast	Y, 0.569	0.989	P, 0.860	P, 0.96	No	P, 2.983	Nt	Nt	Chloroplast
	FBA	g13657	Chloroplast	0.495	0.9815	M, 0.803	E, 0.62	No	P, 3.371	Nt	Nt	Chloroplast
	TPI	g4947g	Cytoplasm	0.454	0.2898	O, 0.636	E, 0.73	No	Other	Nt	Nt	Chloroplast
	TPI	g7559g	Chloroplast	Y, 0.542	0.9036	M, 0.130	E, 0.86	No	P, 0.575	Nt	Nt	Chloroplast
	GAPDH	g11624	Mitochondrion	0.435	0.9784	M, 0.783	M, 0.42	No	Other	Nt	Nt	Cytoplasm
	GAPDH	g11619	Cytoplasm	0.436	0.0125	O, 0.749	E, 0.98	No	Other	Nt	Nt	Cytoplasm
	GAPDH	g880g	Mitochondrion	0.434	0.9049	M, 0.760	E, 0.76	No	Other	Nt	Nt	Mitochondrion
	GAPDH	g9565	Chloroplast	Y, 0.573	0.9169	M, 0.522	P, 0.53	No	C, 3.356	Nt	Nt	Chloroplast
	GAPDH	g13039	Truncated/no prediction									
	PGK	g457g	Cytoplasm	0.446	0.1276	O, 0.824	E, 0.99	No	Other	Nt	Nt	Chloroplast
	PGK	g16467	Chloroplast	Y, 0.573	0.9768	P, 0.959	P, 0.95	No	P, 3.05	Nt	Nt	Chloroplast
	PGAM	g1683g	Cytoplasm	0.442	0.4026	O, 0.505	E, 0.91	No	Other	Nt	Nt	Nucleus
	PGAM	g2873g	Cytoplasm	0.438	0.045	O, 0.669	E, 0.86	No	Other	Nt	Nt	Cytoplasm
	PGAM	g1961g	Cytoplasm	0.442	0.0293	O, 0.864	E, 0.99	No	M, 0.7635	Nt	Nt	Chloroplast
	PGAM	g5368g	Cytoplasm	0.431	0.0104	O, 0.789	E, 0.98	No	Other	Nt	Nt	Cytoplasm
	PGAM	g10700	Cytoplasm	0.429	0.0216	O, 0.863	E, 0.99	No	Other	Nt	Nt	Cytoplasm
	ENO	g13252	Truncated/no prediction									
	ENO	g13732	Cytoplasm	0.436	0.2072	O, 0.852	E, 0.98	No	Other	Nt	Nt	Cytoplasm
	PK	g4286g	Chloroplast	0.455	0.1824	P, 0.691	P, 0.33	No	Other	Nt	Nt	Chloroplast
	PK	g8823g	Chloroplast	Y, 0.512	0.7975	M, 0.631	E, 0.91	No	P, 1.1256	Nt	Nt	Chloroplast
	PK	g12965	Chloroplast	Y, 0.517	0.0129	P, 0.230	E, 0.98	No	Other	Nt	Nt	Cytoplasm
	PK	g15572	Cytoplasm	0.445	0.2514	O, 0.816	E, 0.99	No	Other	Nt	Nt	Chloroplast
	PK	g13471	Cytoplasm	0.433	0.132	M, 0.704	E, 0.73	No	Other	Nt	Nt	Chloroplast
	PK	g13734	Cytoplasm	0.43	0.1567	M, 0.697	E, 0.92	No	Other	Nt	Nt	Chloroplast
	PK	g11454	Cytoplasm	0.434	0.3709	M, 0.735	E, 0.78	No	Other	Nt	Nt	Cytoplasm
Gluconeogenesis	PC	g9967	Cytoplasm	0.444	0.5401	O, 0.689	E, 0.90	No	Other	Nt	Nt	Chloroplast
	PC	g16269	Truncated/no prediction									
	PC	g6872	Chloroplast	Y, 0.518	0.5926	M, 0.745	E, 0.88	No	C, 1.957	Nt	Nt	Mitochondrion
	PEPCK	g13156	Chloroplast	0.471	0.5635	M, 0.537	E, 0.66	No	P, 0.996	Nt	Nt	Cytoplasm
	PEPCK	g984g	Cytoplasm	0.429	0.0851	O, 0.787	E, 0.99	No	Other	Nt	Nt	Nucleus
	PEPCK	g489	Truncated/no prediction									
	FBP	g15165	Cytoplasm	Y, 0.520	0.0361	O, 0.572	E, 0.99	No	Other	Nt	Nt	Cytoplasm
	FBP	g13902	Chloroplast	Y, 0.591	0.9823	P, 0.863	P, 0,72	No	P, 3.058	Nt	Nt	Chloroplast
	FBP	g14899	Cytoplasm	0.431	0.3854	O, 0.647	E, 0.99	No	Other	Nt	Nt	Chloroplast

Table 2-5. *Mk* KAS603 protein localization prediction for enzymes involved in the pyruvate hub. 9 different prediction programs were used. The consensus predicted compartment (s) are denoted in the predicted targeting column. Highlighted green sections denote a chloroplast predicted to be localized enzyme, light blue denotes cytoplasm, orange denotes mitochondrion, dark blue denotes ER and pink denotes glyoxysome. Lower case red "g" denotes protein identification based on the genome assembly, otherwise it is based on the transcriptome assembly. Y = High chloroplast signal prediction, ppPTS1 = PredPlantPTS1 prediction program, E = Elsewhere (cytoplasm), O = Other (cytoplasm), P = chloroplast, M = Mitochondrion, C = Cytoplasm, ER = Endoplasmic Reticulum, No = No ER signal predicted, Yes = ER signal predicted, SP = Secretory pathway, Nt = Not targeted to peroxisome (glyoxysome), t = targeted to peroxisome (glyoxysome). Value numbers represent the total prediction score to corresponding organelle. Full name of proteins mentioned in this table, are listed in supplemental table S2-1.

Pathway	Protein name	Protein ID	Predicted targeting	ChloroP	Mitoprot	TargetP	Predotar	Signal P	PredAlgo	PTS1	ppPTS1	Plant-mPLoc
Pyruvate hub	PEPC	g14708	Cytoplasm	0.434	0.1258	O, 0.795	E, 0.99	No	Other	Nt	Nt	Cytoplasm
	PEPC	g8642g	Chloroplast	0.456	0.7004	O, 0.559	P, 0.27	No	P, 0.500	Nt	Nt	Cytoplasm
	PEPC	g14334	Cytoplasm	0.424	0.0117	O, 0.961	E, 0.99	No	Other	Nt	Nt	Cytoplasm
	MDH_{NADP}	g6527g	Chloroplast	Y, 0.552	0.9871	M, 0.656	M, 0.61	No	P, 1.294	Nt	Nt	Chloroplast
	MDH_{NADP}	g5419g	ER	0.454	0.888	SP, 0.784	ER 0.93	Yes, 0.848	SP, 0.374	Nt	Nt	Chloroplast
	MDH_{NAD}	g12064	Cytoplasm	0.437	0.4704	M, 0.602	E, 0.72	No	Other	Nt	Nt	Chloroplast
	MDH_{NAD}	g12897	Mitochondrion	0.463	0.924	M, 0.896	E, 0.84	No	M, 0.579	Nt	Nt	Mitochondrion
	ME	g3218	Cytoplasm	0.431	0.1064	O, 0.651	E, 0.98	No	Other	Nt	Nt	Chloroplast
	ME	g13572	Cytoplasm	0.435	0.1044	O, 0.729	E, 0.99	No	Other	Nt	Nt	Chloroplast
	ME	g11602	Cytoplasm	0.444	0.0119	O, 0.915	E, 0.99	No	Other	Nt	Nt	Chloroplast
	ME	g9349g	Chloroplast.Mitochondrion	Y, 0.502	0.9727	M, 0.865	M, 0.45	Yes, 0.951	P/M. 1.3/1.4	Nt	Nt	Chloroplast
	ME_{NADP}	g5839g	Glyoxysome	0.445	0.3688	O, 0.440	E, 0.97	No	P, 0.716	t, 2.002	t, 0.591	Chloroplast
	PPDK	g7648g	Cytoplasm	0.455	0.0054	O, 0.866	E, 0.99	No	Other	Nt	Nt	Chloroplast
	PDH E1α	g730g	Mitochondrion	0.492	0.9532	M, 0.318	M, 0.49	No	Other	Nt	Nt	Mitochondrion
	PDH E1 β	g3086g	Mitochondrion	0.469	0.6279	M, 0.936	M, 0.67	No	M, 2.31	Nt	Nt	Mitochondrion
	PDH E2	g5303g	Mitochondrion	0.485	0.8774	M, 0.470	E, 0.57	No	M, 1.59	Nt	Nt	Mitochondrion
	PDH E2	g1987g	Chloroplast	0.445	0.2381	M, 0.915	M, 0.47	No	P, 1.009	Nt	Nt	Chloroplast
	PDH E2	g4680g	ER	0.485	0.3418	M, 0.831	E, 0.89	Yes, 0.087	SP, 0.974	Nt	Nt	Nucleus

Table 2-6. *Mk* KAS603 Intracellular targeting predictions for enzymes that act in fatty acid and TAG anabolism. 9 different prediction programs were used. The consensus predicted compartment is denoted in the predicted targeting column. Highlighted green sections denote a chloroplast predicted to be localized enzyme, light blue denotes cytoplasm, orange denotes mitochondrion, dark blue denotes ER and pink denotes glyoxysome. Lower case red "g" denotes protein identification based on the genome assembly, otherwise it is based on the transcriptome assembly. Y = High chloroplast signal prediction, ppPTS1 = PredPlantPTS1 prediction program, E = Elsewhere (cytoplasm), O = Other (cytoplasm), P = chloroplast, M = Mitochondrion, C = Cytoplasm, ER = Endoplasmic Reticulum, No = No ER signal predicted, Yes = ER signal predicted, SP = Secretory pathway, Nt = Not targeted to peroxisome (glyoxysome), t = targeted to peroxisome (glyoxysome). Value numbers represent the total prediction score to corresponding organelle. Full name of proteins mentioned in this table, are listed in supplemental table S2-2.

Pathway	Protein name	Protein ID	Targeted to	ChloroP	Mitoprot	TargetP	Predotar	Signal P	PredAlgo	PTS1	ppPTS1	Plant-mPLoc
FA-TAG anabolism	ACAS	g5262g	Chloroplast	0.478	0.4759	M, 0.451	P, 0.29	No	P, 2.288	Nt	Nt	Peroxisome
	ACAS	g3121g	Glyoxysome	0.439	0.0692	SP, 0.857	E, 0.87	No	SP, 0.788	t, 13.273	t, 0.811	Peroxisome
	ACL a	g14950	Mitochondrion	0.435	0.669	M, 0.369	M, 0.26	No	Other	Nt	Nt	Mitochondrion
	accA α	g10849	Cytoplasm	0.427	0.0136	O, 0.753	E, 0.99	No	Other	Nt	Nt	Mitochondrion
	accA α	g14318	Cytoplasm	0.47	0.83	O, 0.780	E, 0.91	No	Other	Nt	Nt	Mitochondrion
	accA	g5977g	Chloroplast	Y, 0.540	0.2159	P, 0.902	P, 0.47	No	Other	Nt	Nt	Mitochondrion
	accA	g6872g	Chlorop.Mito	Y, 0.518	0.7511	M, 0.745	E, 0.88	No	P, 1.957	Nt	Nt	Mitochondrion
	accA	g3245g	Cytoplasm	0.426	0.322	O, 0940	E, 0.99	No	Other	Nt	Nt	Chloroplast
	MCAT	g8909g	Cytoplasm	0.431	0.007	O, 0.898	E, 0.99	No	Other	Nt	Nt	Chloroplast
	MCAT	g6288g	Chloroplast	0.453	0.0589	O, 0.942	E, 0.99	No	Other	Nt	Nt	Chloroplast
	KASIII	g103	Chloroplast	Y, 0.511	0.9837	M, 0.752	E, 0.72	No	M, 0.85	Nt	Nt	Chloroplast
	KAS III	g15763	ER	0.446	0.2324	SP, 0.562	ER, 0.99	Yes, 1.00	SP, 1.95	Nt	Nt	Chloroplast
	KAS II	g2394	Chloroplast	Y, 0.533	0.6219	M, 0.648	E, 0.81	No	P, 2.08	Nt	Nt	Chloroplast
	KAS II	g12220	Cytoplasm	0.429	0.2027	O, 0.866	E, 0.99	No	Other	Nt	Nt	Cytoplasm
	KAS I	g5311	Chloroplast	Y, 0.533	0.8598	M, 0.716	P, 0.44	No	P, 1.849	Nt	Nt	Chloroplast
	KAR	g12319	Chloroplast	0.445	0.8913	M, 0.860	E, 0.77	No	Other	Nt	Nt	Chloroplast
	HD	g1393	Chloroplast	0.433	0.2029	O, 0.770	E, 0.96	No	Other	Nt	Nt	Chloroplast
	HD	g11450	Chloroplast	0.431	0.0348	O, 0.851	E, 0.99	No	Other	Nt	Nt	Chloroplast
	ENR	g2718	Chloroplast	0.438	0.0062	O, 0.704	P, 0.21	No	Other	Nt	Nt	Chloroplast
	O3D-III	g3602g	Chloroplast	Y, 0.576	0.8691	P, 0.966	P, 0.95	No	P, 2.864	Nt	Nt	ER
	O3D-III	g13802	ER	0.478	0.907	M, 0.846	ER, 0.98	Yes, 0.212	Other	Nt	Nt	Chloroplast
	O6D	g15681	ER	0.442	0.2078	O, 0.591	E, 0.96	Yes, 0.361	Other	Nt	Nt	ER
	O6D	g10484	Cytoplasm	0.465	0.2209	O, 0.837	E, 0.95	No	Other	Nt	Nt	ER
	O6D	g9590	Chloroplast	Y, 0.575	0.9548	P, 0.671	E, 0.49	No	P, 2.116	Nt	Nt	ER
	18:1D	g5577g	Cytoplasm	0.437	0.0303	O, 0.940	E, 0.95	No	Other	Nt	Nt	ER
	18:1D	g5578g	Cytoplasm	Y, 0.507	0.5169	O, 0.708	E, 0.90	No	Other	Nt	Nt	ER
	16:0D	g14707	ER	0.451	0.3465	O, 0.961	E, 0.98	Yes, 0.973	SP, 0.5127	Nt	Nt	NP
	ACS	g14809	Cytoplasm	0.46	0.2523	O, 0.549	E, 0.99	No	Other	Nt	Nt	Persoxisome
	ACS	g16510	Cytoplasm	0.428	0.0072	O, 0.799	E, 0.99	No	Other	Nt	Nt	Peroxisome
	G3PDH	g15769	ER	0.455	NP	SP, 0.777	NP	Yes, 0.991	SP, 1.585	Nt	Nt	Cytoplasm
	G3PDH	g16176	Chloroplast	Y, 0.527	0.1253	M, 0.458	E, 0.95	No	P, 0.534	Nt	Nt	Nucleus
	G3PDH	g4232g	Cytoplasm	0.433	0.0707	O, 0.847	E, 0.98	No	Other	Nt	Nt	Chloroplast
	GK	g8308	Cytoplasm	0.433	0.0102	O, 0.767	E, 0.99	No	Other	Nt	Nt	Cytoplasm
	GK	g2824g	Cytoplasm	0.433	0.0136	O, 0.876	E, 0.99	No	Other	Nt	Nt	Cytoplasm
	GPAT	g1475	Cytoplasm	0.434	0.0188	O, 0.870	E, 0.98	No	Other	Nt	Nt	Chloroplast
	GPAT	g1495g	Cytoplasm	0.432	0.1675	M, 0.732	E, 0.72	No	Other	Nt	Nt	Mitochondrion
	LPAAT	g13372	ER	0.424	0.2804	O, 0.989	E, 0.97	No	SP, 0.252	Nt	Nt	ER
	PAP	g10757	ER	Y, 0.514	0.0282	SP, 0.850	ER, 0.94	Yes, 0.355	Other	Nt	Nt	Golgi apparatus
	PAP	g4052g	Cytoplasm	0.441	0.1284	M, 0.702	E, 0.81	Yes, 0.086	Other	Nt	Nt	Peroxisome
	DGAT	g16438	ER	0.475	0.9953	M, 0.737	ER, 0.97	Yes, 0.452	Other	Nt	Nt	Chloroplast
	DGAT	g8531	ER	0.449	0.8448	SP, 0.858	ER, 0.99	Yes, 0.910	SP, 0.917	Nt	Nt	Cell membrane
	MLDP	g1295	Cytoplasm	0.443	0.3098	O, 0.756	E, 0.80	No	Other	Nt	Nt	Cytoplasm

Gluconeogenesis

As depicted in Fig. 2-7, a full set of gluconeogenesis enzymes (with the exception of glucose 6-phosphatase) was predicted for location in the cytoplasm. Two steps in gluconeogenesis (ENO, PGAM) were not predicted for the chloroplast, and steps that were (FBP, FBA, TPI, GAPDH PGK) are also involved in ribulose regeneration in the Calvin-Benson cycle. Since 1) gluconeogenesis would not function without GPI, 2) other steps that can also be involved in gluconeogenesis have predicted plastid localization, and 3) analysis of starch metabolism genes indicate that starch storage occurs in the plastid, we assume that the GPI truncated sequence which did not allow for a targeting prediction, is present in the plastid. In the event that GPI was not plastid localized, a series of transport steps of intermediates in and out of the plastid would have to occur to provide the precursor for starch synthesis, which is theoretically possible, but which would be inefficient.

Starch metabolism

As displayed in Fig. 2-7, granule bound starch synthase was predicted to be targeted to the cytoplasm, and plastid starch soluble synthase, to the plastid. This indicates a possible presence of different forms of starch in the chloroplast (readily available soluble starch) and cytoplasm (compact storage starch form). Other appropriate enzymes present to enable starch synthesis to occur in both locations were strongly suggested by our targeting predictions, lending support to starch presence in both locations. AGP, which generates the immediate precursor for starch

biosynthesis, was predicted to the cytoplasm and to the plastid. PGM (which catalyzes the interconversion of G6P to G1P – a precursor for AGP) was targeted to the plastid, suggesting that the plastid would have to export its product, G1P, to the cytoplasm for starch synthesis to occur there. A truncated PGM sequence was identified, and although it couldn't be targeted, a possible location is the cytoplasm (shown with a question mark in Fig 2-7). If this is the case, then there would be no need to export G1P from the plastid. A starch branching enzyme was targeted to the cytoplasm. Alpha amylase (which catalyzes the breakdown of starch to branched glucans and glucose formation) was predicted to be localized to the cytoplasm and to the plastid. A starch debranching enzyme (isoamylase) was targeted to the plastid. Beta amylase (which catalyzes linear glucans and maltose formation from starch), alpha glucosidase (which catalyzes the breakdown of starch and disaccharides to glucose) and starch phosphorylase (which catalyzes the phosphorolytic degradation of starch) were all predicted to be localized to the cytoplasm, suggesting a preference of starch degradation toward maltose and glucose in the cytoplasm, although starch degradation toward glucose could occur in both compartments.

Alignment analysis of starch metabolism, glycolysis, gluconeogenesis and pyruvate hub enzymes

The presence of starch storage and glycolysis in the cytoplasm is distinct from what occurs in the established model green alga *C. reinhardtii*. The presence of multiple steps in these processes predicted in the *Mk* KAS603 analysis, and lack of

corresponding prediction in *C. reinhardtii*, provides strong support that these cytoplasmic processes are authentic. A consequence of plastid targeting would be the presence of an additional targeting sequence at the amino terminus. As a means of further testing the hypothesis that processing of carbon occurs to the predicted extent in the plastid, we therefore analyzed the enzymes involved in starch metabolism, glycolysis, gluconeogenesis and pyruvate hub for extensions only on the predicted plastid localized enzymes and lack of extension on the cytoplasmically localized ones. As can be seen in Fig. 2-8 (where we present 2 examples per pathway), the analysis supports the predicted targeting.

Figure 2-8. Plastid amino terminal extension alignments for starch synthesis, glycolysis, gluconeogenesis and pyruvate hub enzymes. Two enzymes per metabolic route are displayed. Protein sequences for *Mk* KAS603 (denoted as Mk followed by the predicted targeting location) and *C. reinhardtii* (denoted as Chlre followed by the predicted targeting location) are compared. For FPB only predicted plastid and cytoplasmic forms in *Mk* KAS603 are compared. Longer sequences display plastid predicted localization proteins. Sequences that lacked the amino terminal extension were targeted to the cytoplasm. See supplemental table S2-1 for full enzyme names.

Glyoxylate cycle

As displayed in Fig. 2-9, an isozyme of Acetyl CoA synthetase (ACAS), which catalyzes the formation of Acetyl CoA, was predicted to be targeted to the glyoxysome, and is proposed to promote the glyoxylate cycle when *Mk* KAS603 is cultivated with acetate (this will be thoroughly discussed in chapter 3). The glyoxylate cycle enzymes citrate synthase, isocitrate lyase and malate synthase were predicted to be localized in the glyoxysome. Aconitase hydratase and malate dehydrogenase were predicted to be localized to the cytoplasm, suggesting the export of malate to the cytoplasm and import of isocitrate and oxaloacetate to the glyoxysome, to complete the glyoxylate cycle. Key enzymes for this cycle, which catalyze sequential steps (for succinate/glyoxylate and malate synthesis), were identified both as fused (IL-MS) or as individual (IL and MS) proteins (Fig. 2-9). Fused activities might contribute to a more efficient substrate processing in the glyoxylate cycle (with a possible enhancement on succinate, glyoxylate and malate production, as needed). To the best of our knowledge, such occurrence hasn't been reported for other green microalgae. Malic enzyme was predicted to be targeted to the glyoxysome, and even though it is not part of the glyoxylate cycle, it is suggested to provide malate for a possible incorporation to the cycle.

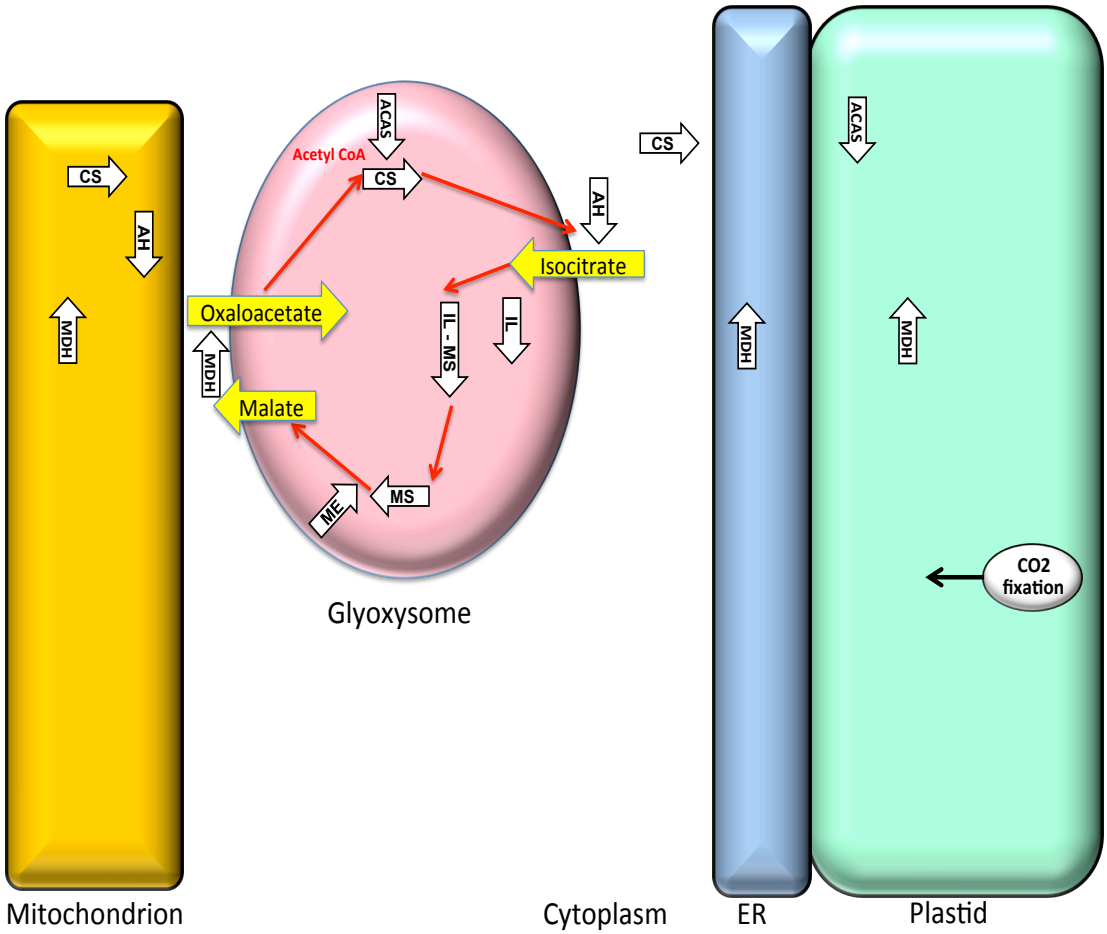

Figure 2-9. *Mk* KAS603 predicted glyoxylate cycle topology map. The glyoxylate cycle is predicted to occur in the glyoxysome and cytoplasm. Flow of carbon is displayed in red arrows. ACAS is Acetyl-coenzyme A synthetase, CS is citrate synthase, AH is aconitase hydratase, IL is isocitrate lyase, MS is malate synthase, MDH is malate dehydrogenase. Single white arrows denote unidirectional enzymatic reactions in the pathway. Black letters inside white arrows are enzyme abbreviations. Identified fused proteins IL and MS are denoted inside the same white arrow.

Fatty acid (FA) and triacylglycerol (TAG) catabolism

As depicted in Fig. 2-10, TAG catabolism via beta oxidation was predicted to be cytoplasmically localized (three TAG lipases were predicted for cytoplasmic location), suggesting lipid droplet degradation in this compartment. Also, fatty acids released upon hydrolysis of membrane structural lipids can provide fatty acid precursors for beta oxidation (Kong *et al*, 2017). The resulting fatty acyl groups might be activated in the cytoplasm (2 acyl CoA synthetases were predicted for location to cytoplasm) and possibly incorporate into the beta oxidation pathway in the glyoxysome (ACOX, ECH, HADH, ACAT, DECR predicted for location to glyoxysome), yielding Acetyl CoA. Predicted glyoxysomal located ECH (which hydrates the double bond between the 2^{nd} and 3^{rd} carbons on acyl-CoA) and HADH (which catalyzes the oxidation of straight-chain 3-hydroxyacyl-CoAs with the production of $NADH_2$) were identified as individual and as fused proteins. To the best of our knowledge fused ECH-HADH hasn't been reported in other chlorophytes. Beta oxidation in other green microalgae can occur in the mitochondria, in peroxisomes or in both organelles (Kong *et al*, 2017; Tan and Lee, 2016). In our work we only found evidence of glyoxysomal (peroxisome) beta oxidation in *Mk* KAS603, which might boost the formerly mentioned glyoxylate cycle with Acetyl CoA (see beta oxidation in discussion section).

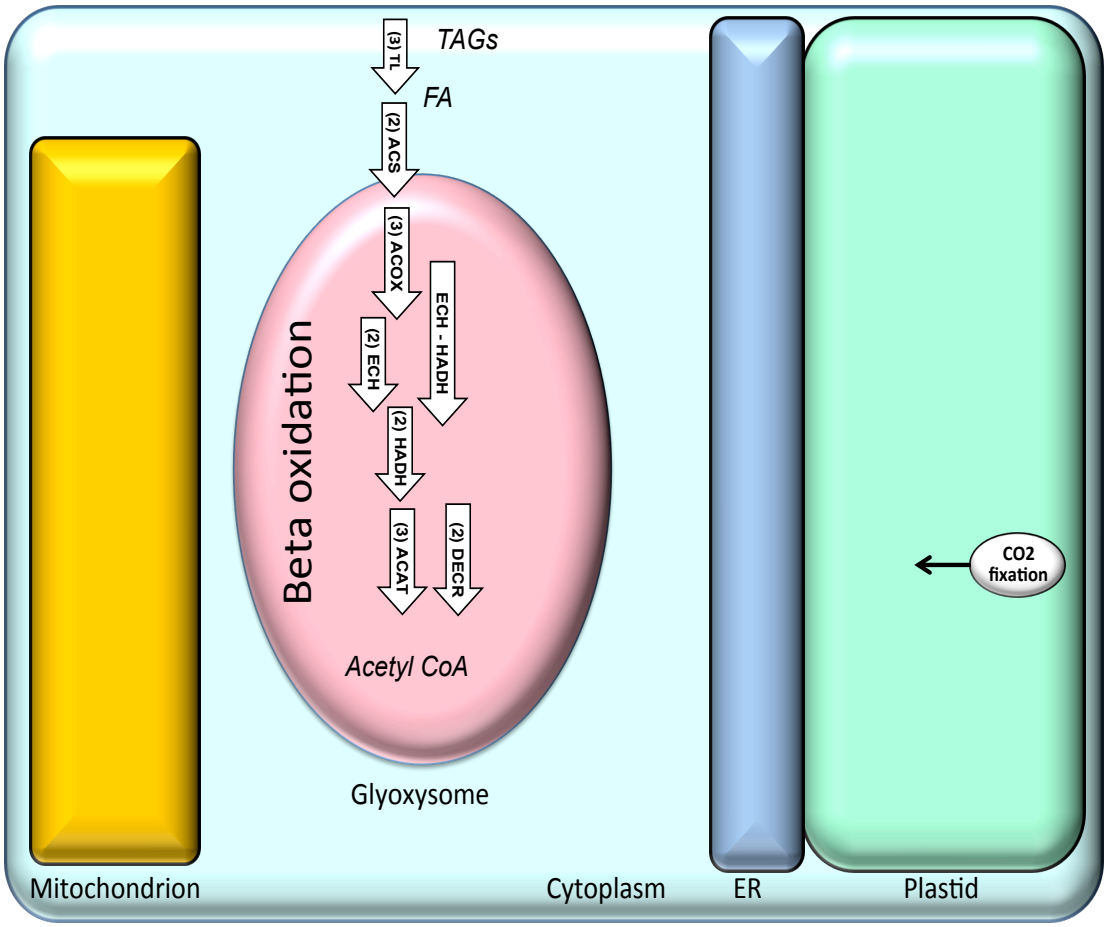

Figure 2-10. *Mk* KAS603 predicted TAG catabolism and FA beta oxidation topology map. TAG lipases (TL) were targeted to the cytoplasm. All beta oxidation enzymes were targeted to the glyoxysome. Available glyoxysomal Acetyl CoA is denoted as a product of beta oxidation. White arrows denote unidirectional enzymatic reactions in the pathways. Black letters inside white arrows are enzyme abbreviations (see enzyme names in supplemental table S2-4). Identified fused proteins ECH and NADH are denoted inside the same arrow. Number of identified enzyme isoforms is denoted in parenthesis.

Fatty acids and TAG synthesis

As displayed in Fig. 2-11, the majority of the fatty acid biosynthesis enzymes, including AccA, MCAT, KAS I, KAS II, KAS III, KAR, HD and ENR, were predicted for plastid localization. Two desaturases (O3D and O6D) were also predicted to be localized to the plastid. This suggests that the synthesis and desaturation of FA may take place in the plastid of *Mk* KAS603. AccA, MCAT, KAS II and 2 long chain ACS, were predicted to be located in the cytoplasm, which may indicate a possible alternate compartment for FA synthesis and FA activation in the cytoplasm, including possible FA desaturation (O6D and 18:1D were predicted to be located to cytoplasm). The ER was also included for predicted targeting, and KAS III, O3D, O6D and 16:0D were predicted to be localized to the ER, suggesting a possible partial processing of FA in the ER, mainly for desaturation. Similarly to other green algae (Fan *et al*, 2011; Huerlimann and Heimann, 2013; Blatti *et al*, 2013), the cytoplasmic set of enzymes in *Mk* KAS603 seem to be involved in long chain fatty acid synthesis, with shorter chain FAs generated in the plastid as precursors. Fatty acids formed in the plastid are proposed for export to the cytoplasm where they are activated and Acyl-CoA groups might be formed (Fig. 2-10, yellow arrows). G3PDH, 2 GPAT isoforms, GK and PAP were predicted to be localized in the cytoplasm, suggesting further processing of fatty acids into TAG assembly. G3PDH, LPAAT, PAP and 2 DGAT isoforms were predicted to be localized to the ER, which might be a compartment where TAG assembly takes place in conjunction with the cytoplasm. G3PDH was predicted to be localized to the plastid, consistent with glycerol 3-

phosphate being exported from plastid to cytoplasm, to be incorporated into TAG synthesis. In *C. reinhardtii*, GPAT and LPAT (which are enzymes that participate in the synthesis of DAG from glycerol-3-phosphate and acyl-CoA) are reported to be localized in the plastid and cytosol (Merchant *et al*, 2012), consistent with the formation of TAG in those compartments. In *Mk* KAS603, these enzymes were predicted to be absent from the plastid; and only G3PDH corresponding to TAG synthesis was localized in this compartment, therefore we discarded the possibility of TAG assembly in the plastid.

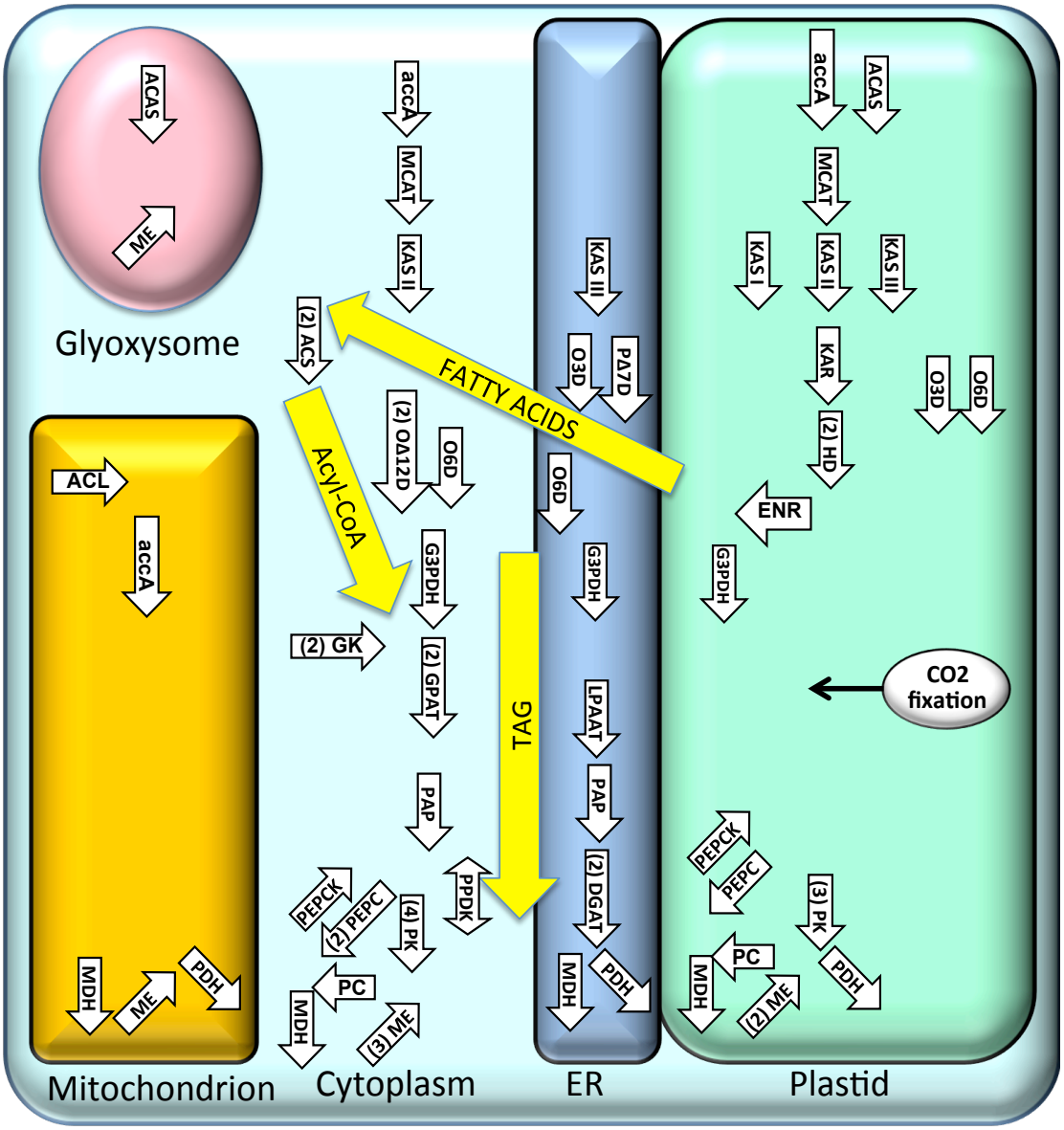

Figure 2-11. *Mk* KAS603 predicted FA-TAG synthesis and Acetyl CoA formation topology map. A predicted main chloroplastic localization of FA synthesis and export to cytoplasm for activation and predicted incorporation to TAG assembly in the cytoplasm in conjunction with ER, are displayed. Yellow arrows denote predicted directional movement of FA, acyl-CoA and TAG assembly in the cell. Acetyl CoA is proposed for availability in the glyoxysome, mitochondria, cytoplasm, ER and chloroplast, consistent with the inability of AcCoA to cross membranes. Double and single white arrows denote bidirectional and unidirectional enzymatic reactions in metabolic pathways. Black letters inside white arrows are enzyme abbreviations (see enzyme names in supplemental table S2-2). Number of

identified enzyme isoforms is denoted in parenthesis.

Photorespiration

As displayed in Fig. 2-12, the photorespiration pathway was predicted to be localized to the mitochondrion (AGAT, GDC and SHMT targeted to mitochondrion), plastid (RubisCO and PGP targeted to plastid) and glyoxysome (GOX, AGAT, SGAT, HPR targeted to the glyoxysome). A (fused) enzyme (GDH/GOX) was predicted to be localized to the cytoplasm. RubisCO in the plastid can fix oxygen instead of CO_2 and catalyze the condensation of O_2 with ribulose 1,5 bisphophate to form 3-phosphglycerate and 2-phosphglycolate. PGP (predicted to be localized to plastid) converts 2-phosphoglycolate to glycolate, which is proposed to be exported to the cytoplasm/glyoxysome for glyoxylate conversion (GOX was targeted to the glyoxysome and fused GDH/GOX was targeted to the cytoplasm). Glyoxylate in (or imported from the cytoplasm) the glyoxysome can be converted to glycine (AGAT targeted to the glyoxysome), which is proposed to be exported to the mitochondrion and be converted to serine (GDC and SHMT targeted to the mitochondrion). Serine is proposed to be exported back to the glyoxysome and converted to glycerate (HPR targeted to the glyoxysome). Glycerate is proposed to be exported into the plastid to start a new cycle. Other proposed fates of glyoxylate are: a) glyoxysomal glyoxylate is suggested to be exported to the mitochondria and converted to glycine (AGAT targeted to the mitochondria) and serine (GDC and SHMT targeted to the mitochondria). b) Glyoxysomal glyoxylate is also suggested to simply incorporate into the glyoxylate cycle (which occurs in the same cellular compartment). c)

Cytoplasmic glyoxylate is also suggested to enter the glyoxysome and incorporate into to the glyoxylate cycle. Similar photorespiratory events were described by Subramanian *et al* (2013) for chlorophytic microalgae and by Xie *et al* (2015) for *Chlorella sorokiniana* under high light in the presence of acetate.

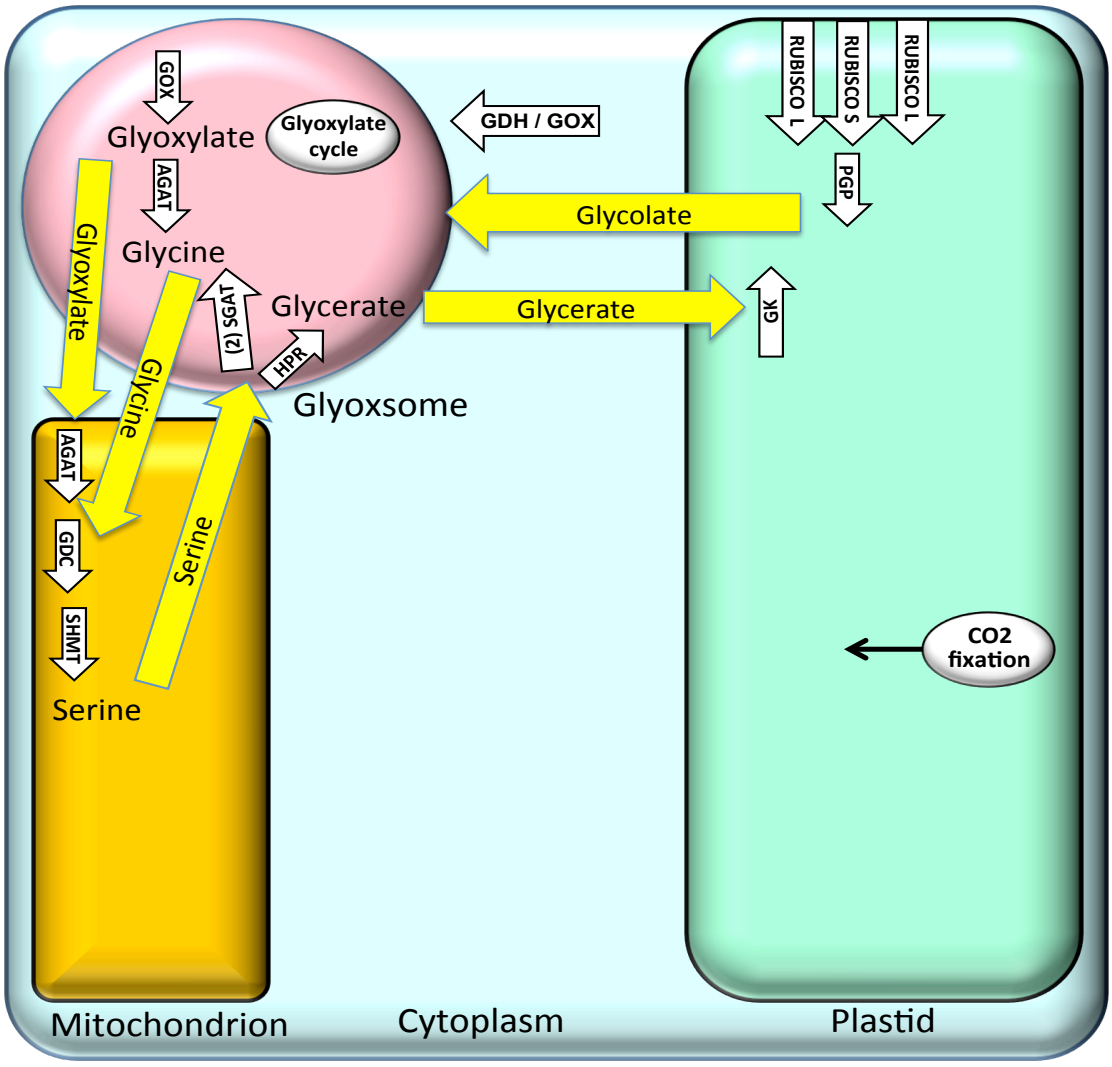

Figure 2-12. *Mk* KAS603 predicted photorespiration topology map.
Photorespiratory activity was predicted to be localized to the chloroplast, glyoxysome and mitochondria. Glycolate is proposed to exit the plastid and enter the glyoxysome to convert into glyoxylate and incorporate to the glyoxylate cycle or to continue photorespiration with glycine production and export (of glyoxylate and glycine) to the mitochondria where serine is suggested to be produced and exported back to the glyoxysome (for more glycine conversion) of for glycerate synthesis and export to the chloroplast to reinitiate the pathway. Single white arrows denote unidirectional enzymatic reactions. Black letters inside white arrows are enzyme abbreviations (see enzyme names in supplemental table S2-5). Number of identified enzyme isoforms is denoted in parenthesis Yellow arrows denote proposed flow of carbon components as they are suggested to enter or exit cell compartments.
Pentose phosphate pathway (PPP)

The pentose phosphate pathway is proposed as source of NADPH for FA anabolism and for producing plastid precursors for starch in *Mk* KAS603. As displayed in Fig. 2-13, the pentose phosphate pathway is suggested to take place in two compartments (plastid and cytoplasm): the oxidation of glucose 6-phosphate and formation of 6-phosphogluconolactone is suggested to occur in the plastid (G6PD targeted to plastid). This reaction releases reducing equivalents in NADPH form. 6-phosphogluconolactone is proposed to be converted to 6-phosphogluconate in the plastid (PGL targeted to plastid) which is proposed to be exported to the cytoplasm (Fig. 2-13, left yellow arrow), and suggested to be converted to ribulose 5-phosphate (6PGD targeted to the cytoplasm) with NADPH production, and to ribose 5-phosphate (R5PI) (which serves as a precursor for nucleotides, coenzymes and nucleic acids). Ribulose 5-phosphate is also proposed to enter the plastid (Fig. 13, right yellow arrow) for conversion to either ribose 5-phosphate (R5PI targeted to plastid) or xylulose 5-phosphate (RP3E targeted to plastid) and processed by TK and TA (plastid targeted) until forming fructose 6-phosphate. In *Chlamydomonas reinhardtii,* G6PD and 6PGD were both reported to be located in the plastid (Klein 1986). In *Mk* KAS603, the shift in location of (2 isoforms) 6PGD suggests that the PP pathway partly occurs in the cytoplasm, similarly to cyanobacteria and red algae (Maruyama *et al*, 2008). The importance of such occurrence lies in the predicted existence of an extraplastidial FA-TAG synthesis that can use NADPH provided by the cytoplasmically predicted section of the PP pathway. In *Chlorella pyrenoidosa* an increased glucose flux through the oxidative phase of the pentose phosphate was

reported to provide NADPH during heterotrophic growth (Tan and Lee, 2016).

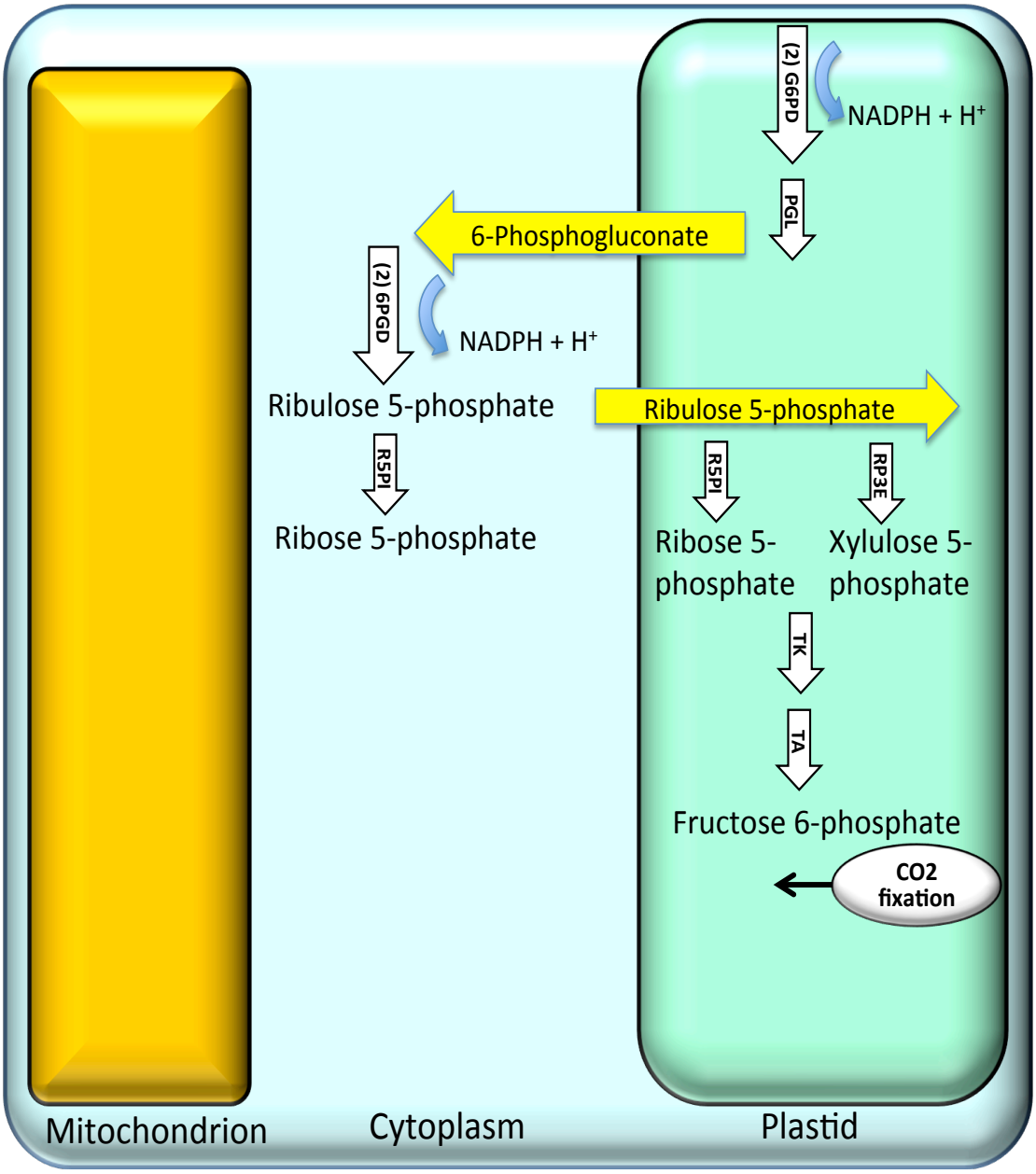

Figure 2-13. *Mk* KAS603 predicted pentose phosphate pathway topology map. This pathway is predicted to take place in the plastid and cytoplasm (for $NADPH_2$ generation) and suggested to generate ribose 5-phosphate for nucleic acid formation or fructose 6-phosphate as a precursor for starch. Single white arrows denote unidirectional enzymatic reactions. Black letters inside white arrows are enzyme abbreviations (see enzyme names in supplemental table S2-6). The number of identified enzyme isoforms is denoted in parenthesis. Yellow arrows denote proposed flow of carbon components as they are suggested to enter or exit cell compartments.

Acetyl CoA

As depicted in Figs. 2-7 and 2-11, targeting predictions suggest that Acetyl CoA can be generated in multiple compartments by distinct means. ACAS predicted to be localized in the glyoxysome suggests a possible incorporation of AcCoA in the glyoxylate cycle. ACL was predicted to be localized in the mitochondrion, consistent with AcCoA being processed in the TCA cycle. PDH was predicted to be localized in the mitochondrion, ER and plastid, which suggests another possible source of AcCoA in those organelles that might be incorporated toward fatty acid synthesis in the plastid, energy generation in the mitochondrion and acyl groups provision for amino acid synthesis in the ER and/or ER export of acetyl groups for histone acetylation (Avidan *et al*, 2015; Goncalves *et al*, 2016). The ER is proposed to act as a storage location for AcCoA and release of acyl groups as needed, and since ER predicted evidence of FA-TAG formation was found, we propose the participation of ER located AcCoA in such formation. ACAS, PDH, ME and MDH were predicted to be plastid localized. This suggests multiple routes of plastidic Acetyl CoA production, which could facilitate FA synthesis in that organelle. ME and MDH can contribute to pyruvate (production) which is converted to AcCoA by PDH, and ACAS is directly involved in AcCoA synthesis.

2.4.6 Carbon flux comparison in green algae

The suggestion that starch synthesis may occur in two different intracellular compartments in *Mk* KAS603 and that glycolysis was located in the cytoplasm rather

than the plastid as in *C. reinhardtii*, prompted an examination of carbon flux topology in other green algae. Using the same enzyme location prediction approach as for *Mk* KAS603, we examined the topology of glycolysis, gluconeogenesis, starch metabolism and the pyruvate hub in other chlorophytes with sequenced genomes.

Glycolysis and gluconeogenesis

As displayed in Figs. 2-14A and 2-15A, *C. reinhardtii* has a predicted plastid-localized upper phase of glycolysis (GK, GPI, PFK, FBA, TPI, GAPDH,) and a cytoplasmic predicted location of 3 enzymes (PGAM, ENO, PK) for the lower phase of glycolysis. The predicted localization of gluconeogenesis was the plastid (GPI, FBP, GAPDH, TPI, and PGK are not predicted in the cytoplasm). The flux of photosynthetically fixed carbon towards starch synthesis and from starch breakdown is confined to the plastid, and further processing to pyruvate occurs in the cytoplasm. As displayed in Figs. 2-14B and 2-15B, *S. obliquus* was predicted to have a robust set of glycolytic/gluconeogenic enzymes in the plastid (except for an extraplastidial ENO and PK, which were targeted to the cytoplasm). A plastidial predicted synthesis and degradation of glucose is proposed for *S. obliquus*, as it was with *C. reinhardtii*. *C. sorokiniana* had a predicted complete glycolytic/gluconeogenic pathway (with the exception of ENO) to be localized in the plastid, but also contained all components split between the cytoplasm and ER (Figs. 2-14C and 2-15C). This offers the possibility that flux for starch synthesis and breakdown could occur either inside or outside the plastid. Cytoplasmic gluconeogenic processes for starch synthesis would

generate G1P which would be imported into the plastid. *C. variabilis* NC64A had predicted gluconeogenic enzymes in the plastid (PGK, GAPDH, TPI, FBA, FBP, GPI), but lacked a plastidial PFK (Figs. 2-13D and 2-15D), suggesting that glycolytic breakdown does not occur in that organelle. It contained a complete glycolytic/gluconeogenic pathway in the cytoplasm, with no components predicted in the ER. The data suggest that glycolysis would occur in the cytoplasm and gluconeogenic processing of fixed carbon could occur both in the plastid and cytoplasm. For the latter, 3PG exported from the plastid is suggested to be re-imported as G6P into the plastid for starch synthesis. Starch breakdown products would be exported from the plastid, for further processing by cytoplasmic glycolysis. *C. subellipsoidea* C-169 (Figs. 2-14F and 2-15F) had a strong predicted glycolytic pathway in the cytoplasm and ER. Due to lack of FBP, gluconeogenesis is predicted to be incomplete in the cytoplasm. The plastid lacked GPI, and has the only copy of FBP, suggesting that gluconeogenic flux could occur up to fructose-6-phosphate (F6P), which could be exported to the cytoplasm for further processing. The plastid may lack PFK for chloroplast glycolytic flux (the gene model lacks a methionine start codon). PFK and GAPDH have predicted ER targeting, suggesting ER import and processing of fructose-6-phosphate and glyceraldehyde 3-phosphate. Comparison of glycolytic and gluconeogenic pathways in these distinct chlorophyte species, provides support for the model developed for *Mk* KAS603, with the fundamental concept that movement of glycolytic, gluconeogenic, or both processes out of the plastid has occurred.

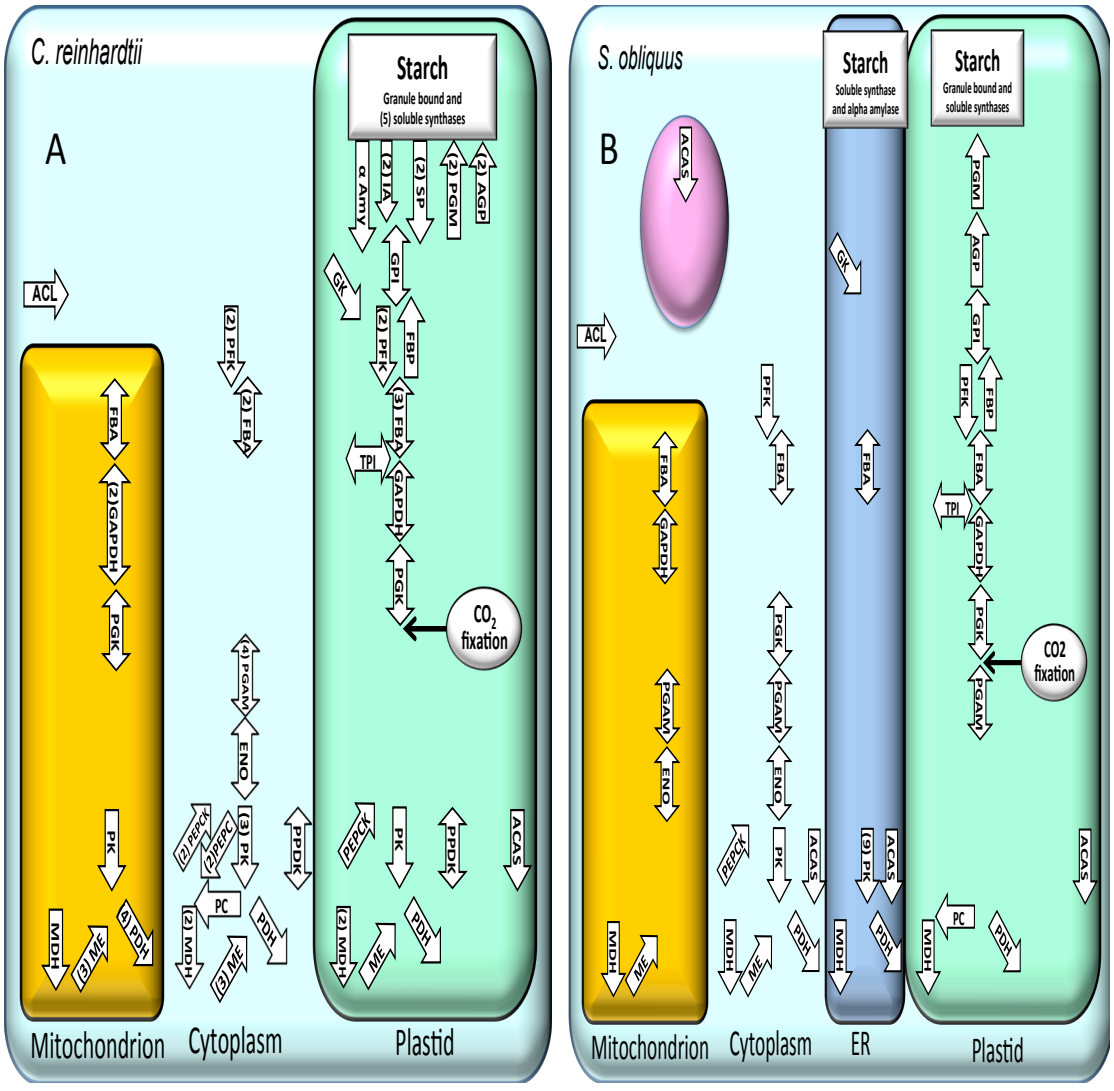

Figure 2-14. Topology map comparison of predicted localization of enzymes involved in starch metabolism, glycolysis, gluconeogenesis and the pyruvate hub in six green microalgal species. Double and single white arrows denote bidirectional and unidirectional enzymatic reactions in a metabolic pathway. Question marks indicate uncertainty about the targeted site. Numbers inside white arrows denote the number of identified enzyme isoforms.

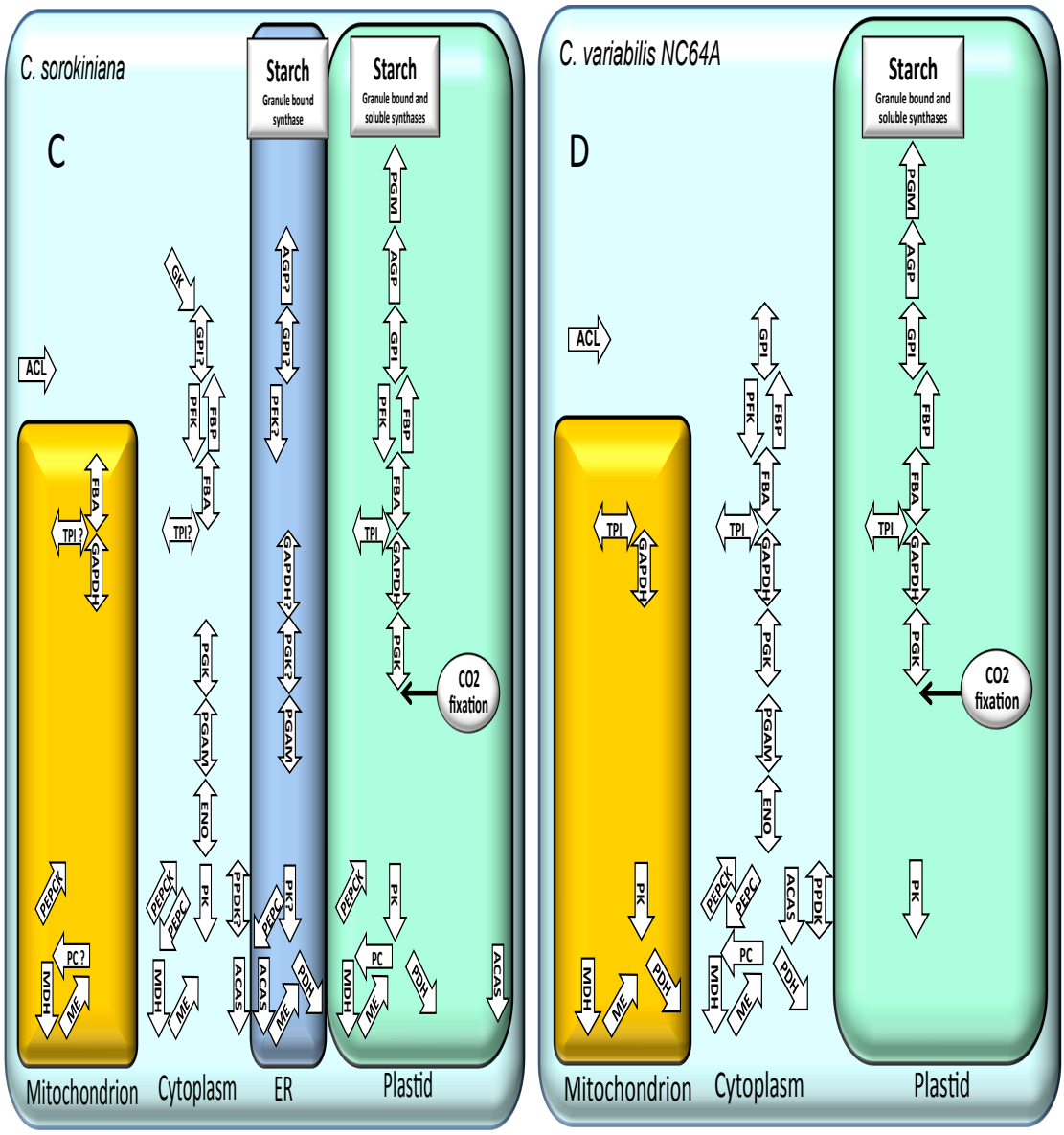

Figure 2-14. Topology map comparison of predicted localization of enzymes involved in starch metabolism, glycolysis, gluconeogenesis and the pyruvate hub in six green microalgal species. Continued.

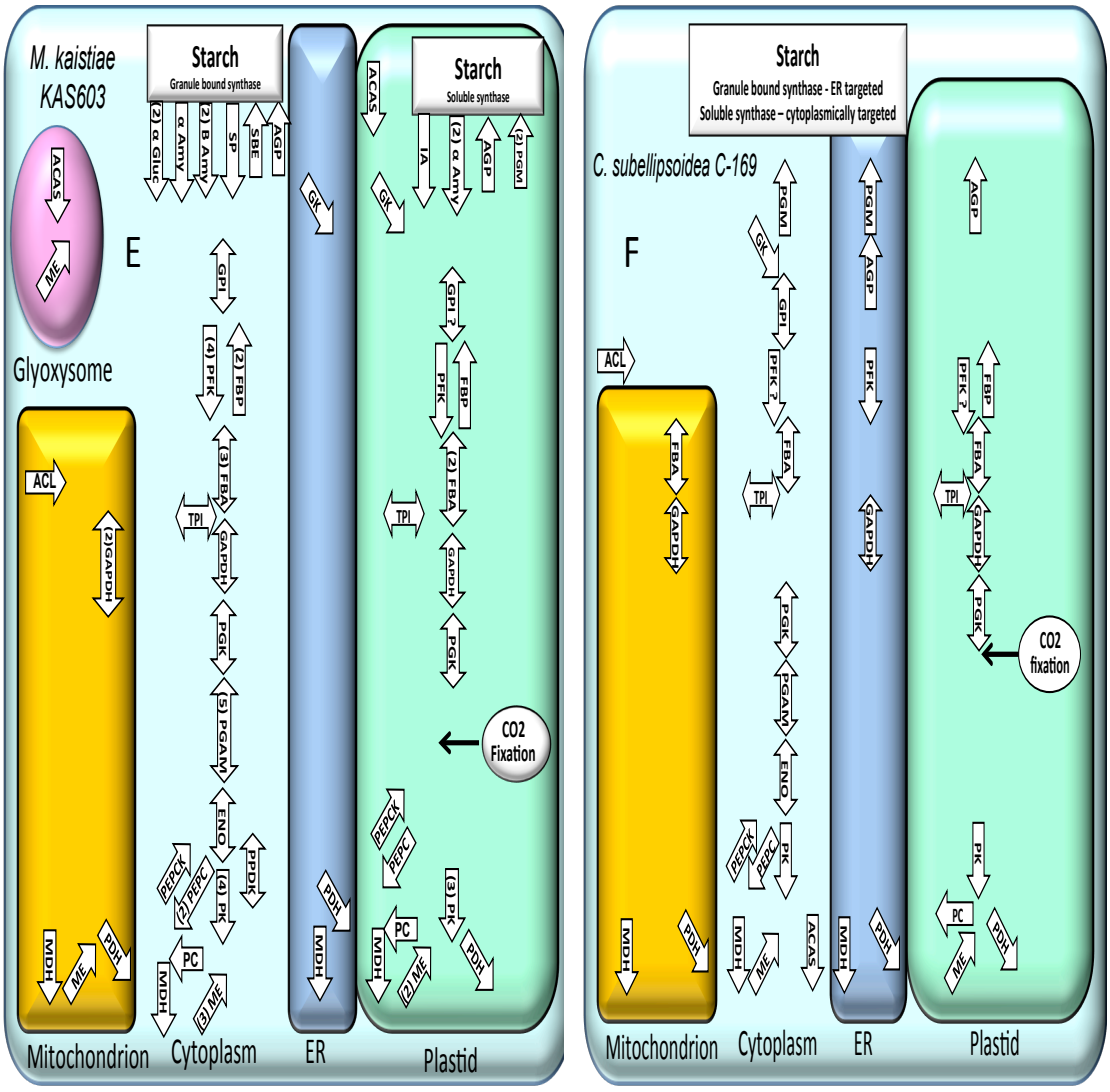

Figure 2-14. Topology map comparison of predicted localization of enzymes involved in starch metabolism, glycolysis, gluconeogenesis and the pyruvate hub in six green microalgal species. Continued.

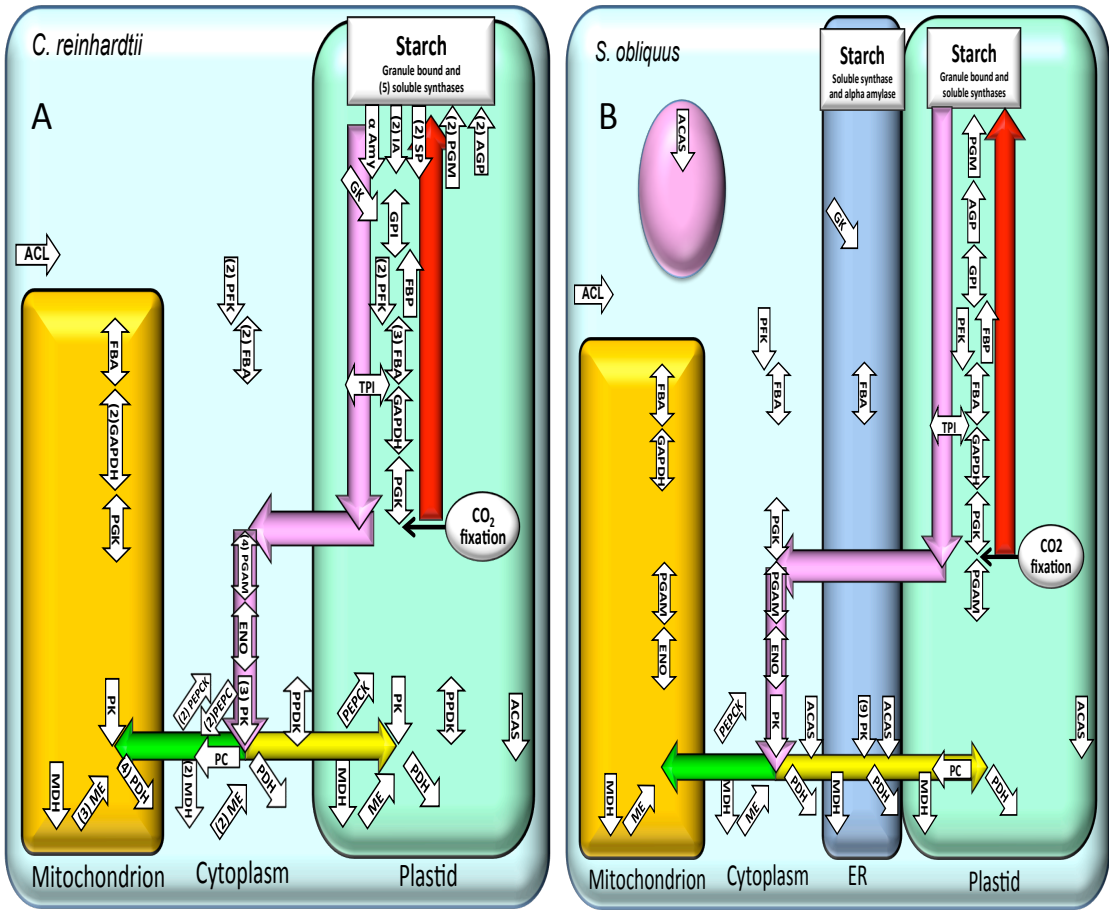

Figure 2-15. Predicted flux of carbon in six green microalgal species. Red arrows denote carbon flux towards carbohydrate, pink denotes breakdown of carbohydrate, bright green denotes carbon flux into the mitochondrion for energy generation for cell division, and yellow denotes carbon flux for fatty acid synthesis. Double and single white arrows denote bidirectional and unidirectional enzymatic reactions in a metabolic pathway. Question marks indicate uncertainty about the targeted site. Numbers inside white arrows denote the number of identified enzyme isoforms.

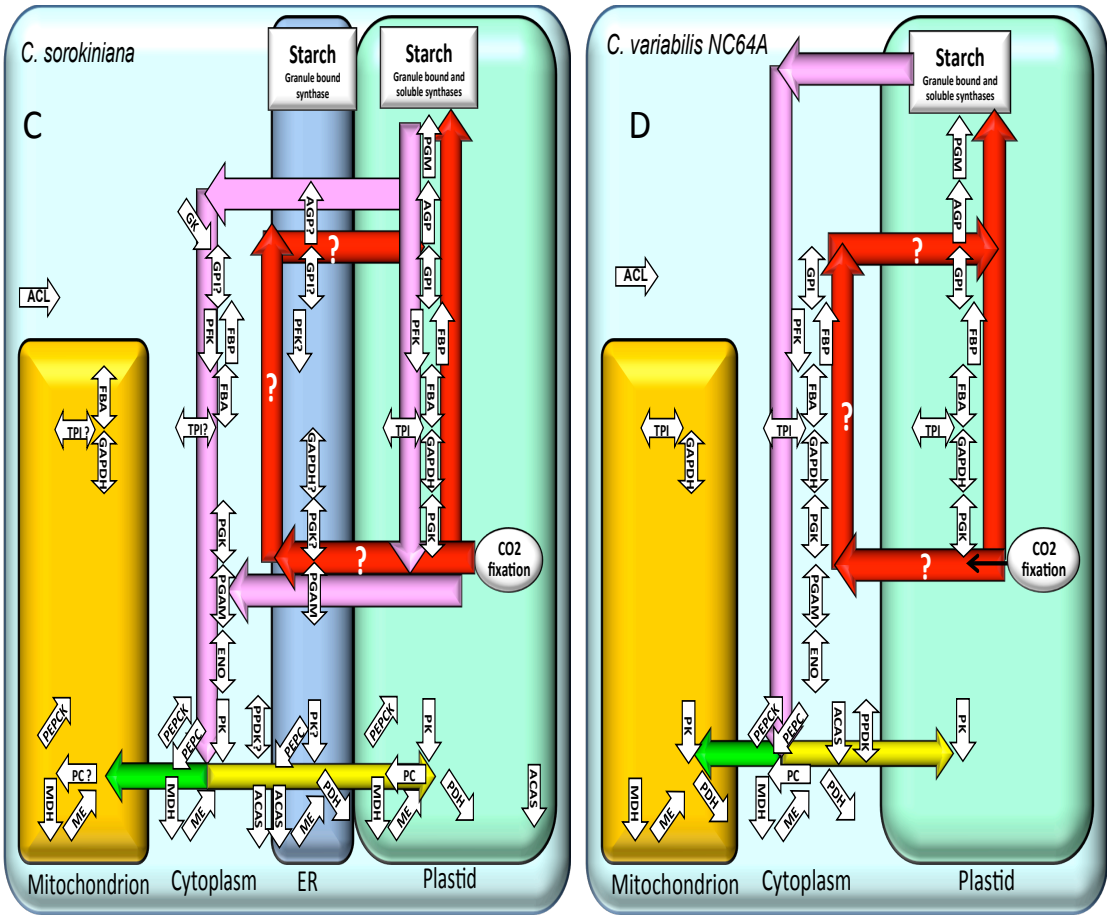

Figure 2-15. Predicted flux of carbon in six green microalgal species. Continued.

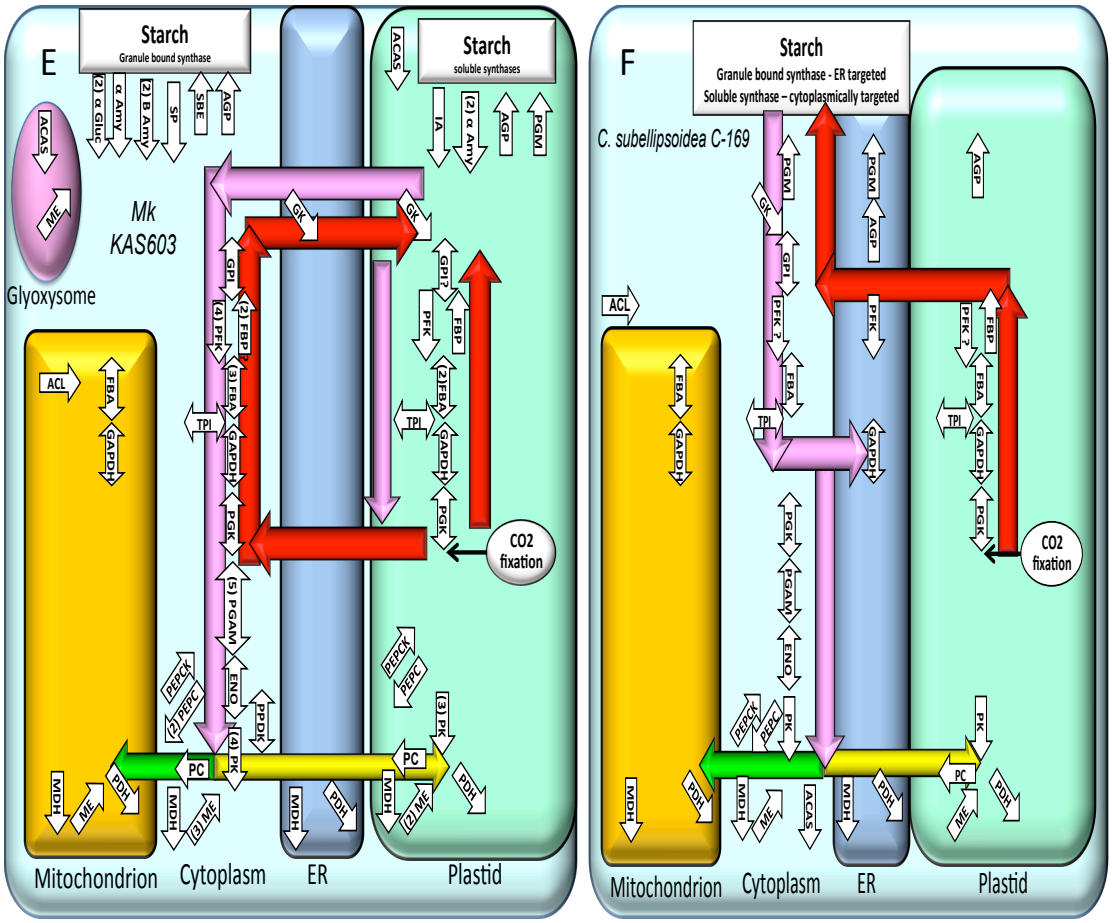

Figure 2-15. Predicted flux of carbon in six green microalgal species. Continued

Starch synthesis

As depicted in Figs. 2-14A and 2-15A, *C. reinhardtii* had predicted plastid targeting for starch catabolism enzymes (alpha amylase, 2 isoforms of IA and 2 isoforms of SP) and of starch formation soluble (5 isoforms) and granule bound synthase, as well as 2 PGM isoforms and 2 AGP isoforms; suggesting starch storage in the plastid, which has been confirmed by visual observation (Busi *et al*, 2014, Buleon *et al*, 1997). *S. obliquus* had soluble and granule bound starch synthases predicted to be located in the plastid, and starch soluble synthase and alpha amylase predicted to be localized in the ER. *S. obliquus* is proposed to store starch both in the plastid and ER. *C. sorokiniana* had predicted starch storage in the plastid (granule and soluble synthases, PGM and AGP were targeted to the plastid) and ER (granule bound synthase and a truncated form of AGP were targeted to the ER). Although the ER prediction is not as strong as the plastid, we present the possibility of extra-plastidial starch accumulation in *C. sorokiniana*. *C. variabilis* NC64A had predicted starch synthesis in the plastid (granule bound and soluble synthases, PGM and AGP are targeted to plastid), with no predicted starch storage in the ER or cytoplasm. As discussed previously, *Mk* KAS603 had starch synthesis and degradation enzymes predicted to be localized in the plastid [soluble starch synthase, PGM (2 isoforms), AGP, alpha amylase (2 isoforms), and isoamylase were targeted to plastid] and in the cytoplasm [granule bound starch synthase, AGP, SBE, SP, beta amylase (2 isoforms), alpha amylase, alpha glucosidase (2 isoforms) were targeted to the cytoplasm],

suggesting starch storage in and outside (cytoplasm) the plastid. As displayed in Figs. 2-14F and 2-15F, *C. subellipsoidea* C-169 is robustly proposed to store starch outside of the plastid, specifically in a compartment with predicted ER targeting. This is supported by predicted ER localization of PGM, AGP, and starch synthases (out of 9 synthases, 8 were predicted to the cytoplasm or ER, only one had predicted chloroplast targeting). Extraplastidial forms of starch are suggested to occur in *S. obliquus, C. sorokiniana, Mk* KAS603 and *C. subellipsoidea* C-169 in contrast to *C. reinhardtii* and *Chlorella sp.* NC64A with only plastidial starch storage predicted. ER targeted starch enzymes enzymes can be acetylated in the ER which can affect their activity and thus ER starch accumulation capacity. The prevalence of extraplastidal starch storage is in contrast to what is currently believed for chlorophytes because models are based on *Chlamydomonas reinhardtii*.

Pyruvate hub

According to our analysis, carbon processing through the pyruvate hub appears to have significant differences between the six green algal strains evaluated (Fig. 2-16). Mitochondrial pyruvate generated by ME is suggested to exist in all strains except in *C. subellipsoidea* C-169, which may use a pyruvate transporter to internalize pyruvate from the cytoplasm. This strain is predicted to have the capacity to produce AcCoA from pyruvate using PDH (as do *Mk* KAS603, *C. variabilis* NC64A and *C. reinhardtii*), which is the classical way of generating AcCoA for entry into the TCA cycle. Two strains (*S. obliquus* and *C. sorokiniana*) may have a non

classical AcCoA generation, which as reported for other green microalgae may possibly produce AcCoA from mitochondrial beta oxidation (Kong *et al*, 2017; Tan and Lee, 2016). Also, due to existence of the mitochondrial citrate-oxalate shuttle (Tan and Lee, 2016) and ACL targeted to the cytoplasm, it is possible that in *S. obliquus* and *C. sorokiniana,* transported citrate derived from the TCA cycle may be converted to AcCoA by ACL in between the membrane or the mitochondrial matrix, and possibly ultimately remaining in the mitochondrial compartment. The plastid of *Mk* KAS603 seems to be enriched in enzymes for production of pyruvate and AcCoA, which may contribute to FA synthesis. *S. obliquu*s is the only strain predicted not to generate plastidial pyruvate (which might be transported/imported from the ER or cytoplasm), but is has strong predicted AcCoA generation (plastid predicted ACAS and PDH). *C. sorokiniana* and *S. obliquu*s possibly process pyruvate into AcCoA in the ER, suggested as a potential provider of acetyl groups for histone acetylation or for amino acid synthesis. *C. reinhardtii* and *S. obliquus* are suggested to convert pyruvate into AcCoA in the cytoplasm for possible FA or amino acid synthesis. Pyruvate generation in the cytoplasm seems to be prevalent in all strains, except in *C. variabilis* NC64A, which may be enriched in PEP and oxaloacetate for predicted cytoplasmic gluconeogenesis. Taken together, these results suggest the possibility of accumulation of pyruvate and OAA in the cytoplasm. In the plastid, PK is predicted in 5/6 species, so pyruvate could be generated in this organelle. PDH is also predicted in 5/6 species, indicating that AcCoA for FA synthesis could be generated by this route. AcCoA synthetase is predicted in the plastid of *C. reinhardtii, S. obliquus, C.*

sorokiniana, and *Mk* KAS603, but not *C. variabilis* NC64A, or *C. subellipsoidea* C-169. ACAS is predicted in the cytoplasm of *S. obliquus, C. sorokiniana and C. subellipsoidea* C-169, and in the ER of *S. obliquus and C. sorokiniana.* Glyoxysomal ACAS is predicted for *S. obliquus and Mk* KAS603. As displayed in Fig. 16, the algae predicted for the broadest compartment ACAS prediction were the oleaginous strains *S. obliquus* (4 compartments) and *C. sorokiniana* (3 compartments). The other option for AcCoA synthesis, ATP citrate lyase, is predicted in the cytoplasm of all species (except for *Mk* KAS603, with predicted mitochondrial ACL).

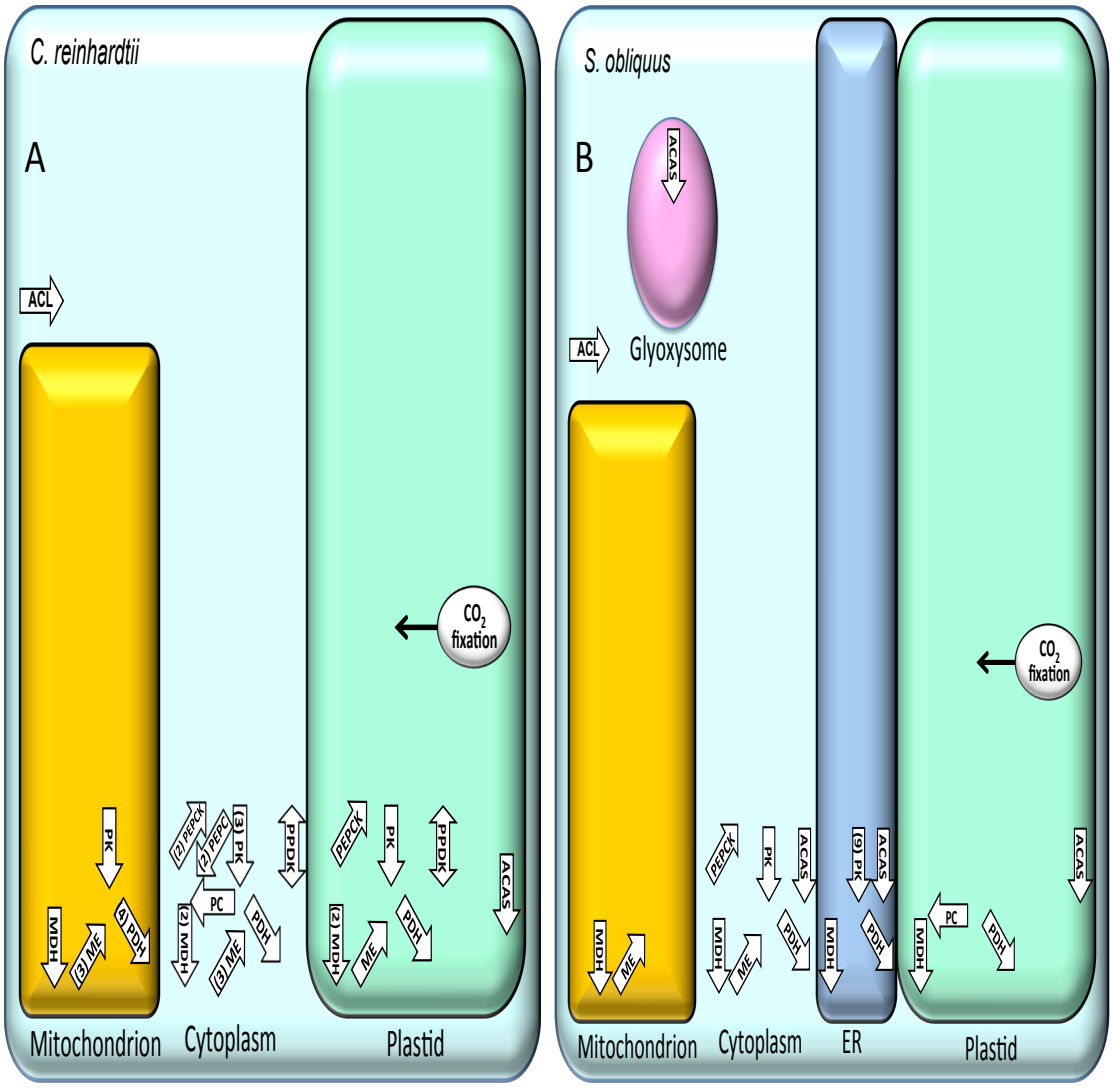

Figure 2-16. Topology map comparison of predicted localization of enzymes involved in AcCoA synthesis in six green microalgal species. Acetyl CoA is proposed not to cross membranes and can be stored in more than 1 compartment in all microalgal species. Double and single white arrows denote bidirectional and unidirectional enzymatic reactions. Black letters inside white arrows are enzyme abbreviations (see enzyme names in supplemental table S2-1). Question marks indicate uncertainty about the targeted site. Numbers inside white arrows denote the number of identified enzyme isoforms.

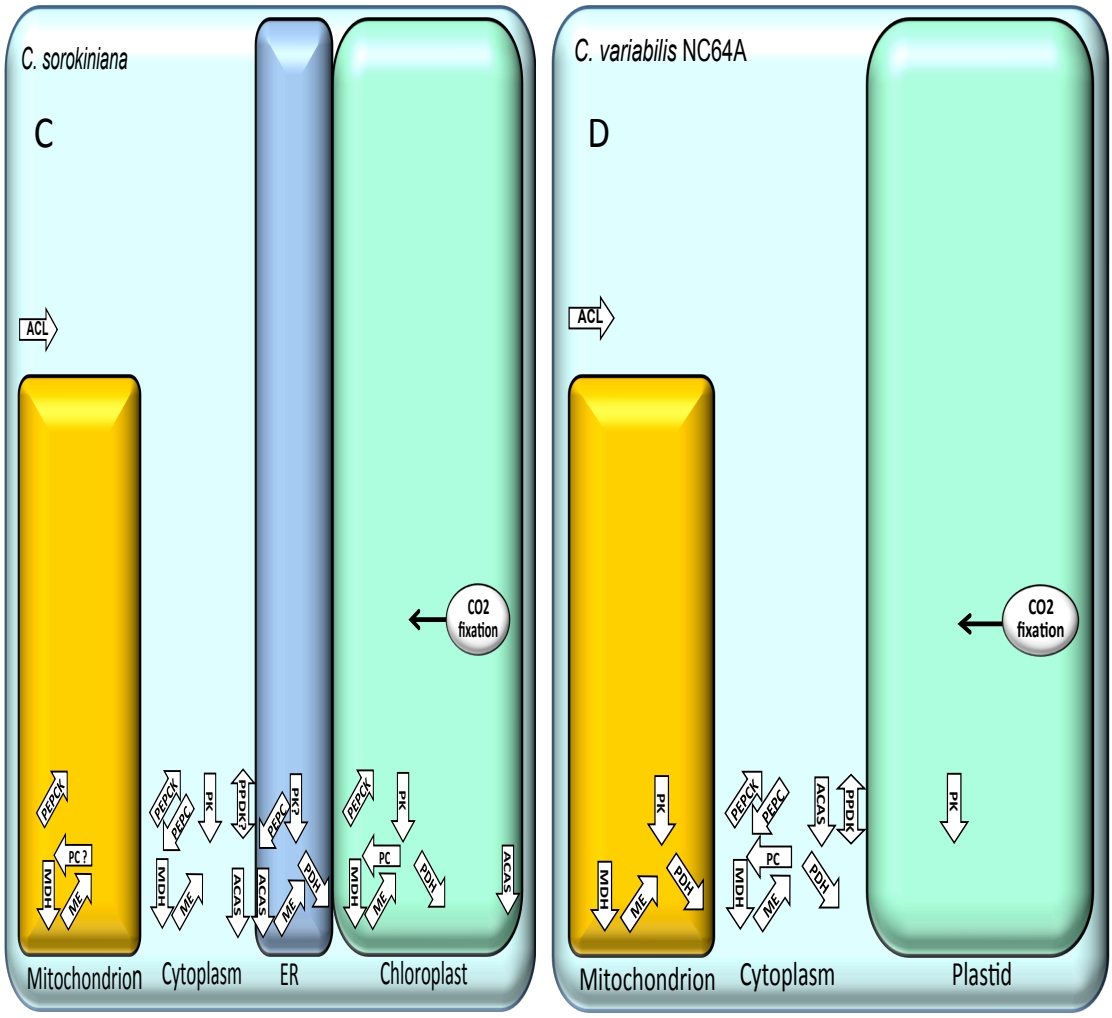

Figure 2-16. Topology map comparison of predicted localization of enzymes involved in AcCoA synthesis in six green microalgal species. Continued.

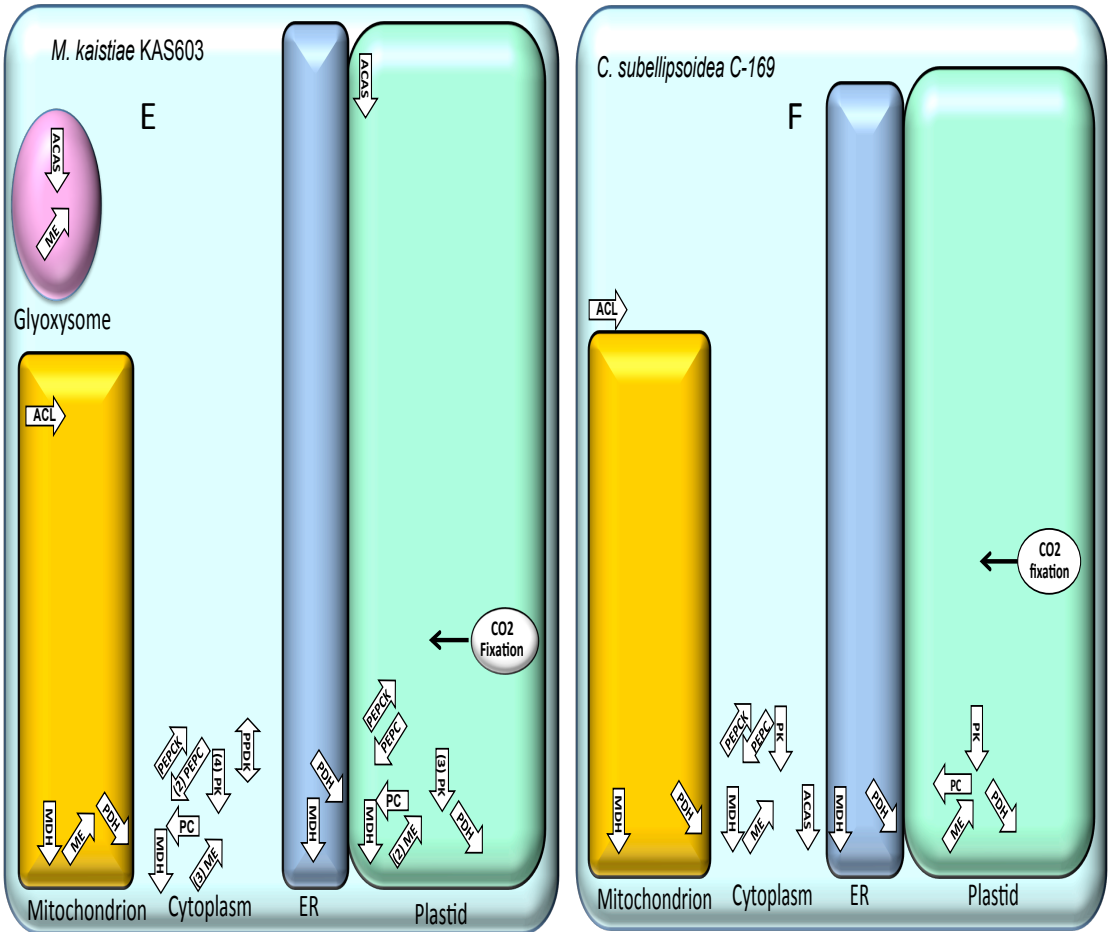

Figure 2-16. Topology map comparison of predicted localization of enzymes involved in AcCoA synthesis in six green microalgal species. Continued.

2.5 DISCUSSION

2.5.1 Sequencing and assembly

The AUGUSTUS pipeline assembly for *Mk* KAS603 rendered 9373 gene models with an average gene length of 2534 nucleotides (Table 2-3), and the genome size and predicted number of genes is very similar to *Chlorella variabilis* NC64A which has 9791 genes models with an average length of 2928 nucleotides (Gao *et al*, 2014). The total length of the genome assembly was considered as 50 Mb (50,214,863 base pairs/non-X nucleotides). The AUGUSTUS genome assembly rendered 31,692,245 base pairs in 1,347 contigs, because it excluded contigs that were less than 10 kb. The nuclear genome size estimation using flow cytometry was 45 Mb (Fig. 3). The level of confidence of the assembly was acceptable mostly due to the support of the flow cytometric analysis and to comparison with another member of the family Chlorellacea, *C. vulgaris* and *C. subellipsoidea* (which had a similar number of genes), and to the full sequence models obtained for the majority of genes/proteins of interest. Most of the missing or incomplete sequences were obtained by combining *Mk* KAS603 protein predictions using genome data plus transcriptome data.

The overlap among orthologous clusters enabled us to compare gene content among 3 members of the family Chlorellacea. As displayed in Fig. 2-6, *Mk* KAS603 formed 3379 ortholog clusters with *C. variabilis* NC64A and *A. protothecoides*, indicating a high proportion of conserved genes among these green microalgal strains.

The 3 species have 2963 single copy genes shared in all 3 genomes, suggesting that they maintained single copy status through evolutionary time after the divergence of these species. *Mk* KAS603 and *C. variabilis* NC64A share a similar number of genes that *C. variabilis* NC64A and *A. protothecoides* do, but *Mk* KAS603 shares fewer with *A. protothecoides,* inferring that they are the least closely related species.

2.5.2 Metabolic capacity of Mk KAS603

Below is a detailed explanation of the proposed metabolic capacity of *Mk* KAS603 based on: 1) physiological responses (under different cultivation conditions described in chapter 1), 2) genomic and transcriptome data that identifies the gene content, and 3) extensive protein targeting analysis that includes a comparative alignment of the cytoplasmic and plastid full sequence protein models using *Chlamydomonas reinhardtii* as a template for longer plastid targeted (signal peptide) sequences (Fig. 2-8). We end our discussion with a proposed carbon flux comparison of glycolysis, gluconeogenesis, starch metabolism and the pyruvate hub among *Mk* KAS603 and five other multiple fission dividing green microalgae.

Glycolytic and gluconeogenic processes related to starch synthesis and breakdown in the cytoplasm and plastid.

The cytoplasmic location of granular starch synthase, ADP-glucose pyrophosphorylase, SBE, SP, α glucosidase, α and β amylases (Fig. 2-7), denote the possible presence of starch storage outside the plastid, similarly to red algae and

glaucophytes, which has been substantiated by visual observation (Viola *et al*, 2001; Busi *et al*, 2014). By definition, granule bound starch synthase is bound exclusively to the starch granule and soluble starch synthase is localized predominantly in the plastid stroma (Patron and Keeling, 2005). Predicted cytoplasmic starch granules in *Mk* KAS603 might be rich in amylose because granule bound synthase is responsible for synthesis of amylose (Busi *et al*, 2014). If true, this distinctive metabolic trait in *Mk* KAS603 is proposed to exist as a possible longer-term storage form. The predicted plastidial starch form might be enriched in amylopectin (soluble starch synthase is responsible for the synthesis of amylopectin [Busi *et al*, 2014]). If true, this suggests a more rapid mobilization of stored carbon to provide substrates for the Calvin-Benson cycle in the plastid. The upper portion of glycolysis and gluconeogenesis is predicted for both the cytoplasm and plastid, but only the cytoplasm would enable glycolytic processing of 3PG to pyruvate through the activity of PGAM and ENO (Fig. 2-7). This data suggests that precursors and breakdown products for starch can be processed in both compartments. There is a plastid targeted GK which would enable phosphorylation of glucose breakdown products of starch, which could be further glycolytically processed in the plastid, but also could be exported as G6P to be glycolytically processed in the cytoplasm (Fig. 2-7). The latter possibility would avoid possible conflicting reactions from occurring in the plastid, which may improve the ability of the cell to process carbon with fewer constraints. For cytoplasmic glycolysis, A HK (which was considered a glucose kinase) with a nucleotide domain (which might be capable of processing UDP-Glucose) is predicted

to be ER targeted. It could process glucose generated by starch degradation, which is transported into the ER (similarly to *S. cerevisiae* and UDP-glucose transport into the ER [Castro *et al*, 1999]), phosphorylated, then exported to the cytoplasm where further glycolytic processing occurs. We identified two enolase isoforms, one with a full-length sequence with predicted cytoplasmic localization, and the other which had a truncated ENO sequence fused to FTS1. FTS1 is a cell division protein in bacteria, and homologs are involved in mitochondrial and plastid replication in photosynthetic eukaryotes. Since ENO is considered a multifunctional or moonlight protein that can bind to microtubule network proteins such as F-actin and tubulin (McAlister and Holland, 1982; Walsh *et a*l, 1989; Díaz- Ramos *et al*, 2012), FTS1 would only be involved in plastid or mitochondrial division. According to Polle *et al* (2014) ENO is proposed to be localized only in the cytoplasm of green algae to facilitate carbon flow and carbon exchange between plastid and cytosol, as it interconverts 2-phosphoglycerate to PEP (and PEP is a precursor to pyruvate). If in fact ENO is only localized to the cytoplasm of *Mk* KAS603, it implies that PEP is imported into the plastid to be converted to pyruvate and AcCoA for FA accumulation. If ENO is localized to the plastid (the truncated ENO-FTS1 fusion suggests that it might be plastid-targeted), PEP might not be needed for import to the plastid, but 2-phosphoglycerate would.

Glyoxysomal activity

As displayed in Fig. 2-9, similarly to plants and other green microalgae, the

glyoxylate cycle in *Mk* KAS603 is predicted to take place mainly in the glyoxysome where energy for growth and malate for fatty acid synthesis are predicted to be produced. The glyoxylate cycle allows plants and some green algae such as *C. reinhardtii* or *Chlorella vulgaris* to grow on acetate because the cycle bypasses the decarboxylation steps of the citric acid cycle. The enzymes that permit the conversion of acetate into succinate, glyoxylate and malate, are isocitrate lyase and malate synthase. (Perez-García *et al*, 2011; Morales-Sánchez *et al*, 2015). Such enzymes were identified as fused (and as single) proteins in the glyoxysome, and are proposed to promote cycle efficiency because of their close proximity and duplicity in *Mk* KAS603 (Fig. 2-9). Glyoxysomal targeted ACAS is proposed to generate AcCoA for incorporation to the glyoxylate cycle (Fig. 2-9). As displayed in Figs. 2-9 and 2-12, glyoxylate is also proposed to be produced from glycolate in the glyoxysome during photorespiration, and could incorporate into the glyoxylate cycle to enhance malate production for plastid/cytoplasm export for FA synthesis. Also, we proposed an enzyme that is not typically part of the glyoxylate cycle (glyoxysomally targeted ME) to provide malate to the cycle (or be exported from the glyoxysome). Malic enzyme isn't targeted to the glyoxysome/peroxisome in *C. reinhardtii* (or in other green microalgae), which might not have a glyoxysomal malate boost as the one predicted for *Mk* KAS603. "Extra" glyoxysomal malate is predicted to follow 3 paths: a) incorporate to the glyoxylate cycle, b) be exported to the plastid for FA synthesis or c) be exported to the mitochondria and TCA cycle to provide carbon skeletons, energy (ATP) and reductive molecules (NADPH). Other differences are that in *C.*

reinhardtii, isocitrate lyase is reported as being located in the cytoplasm, and aconitase hydratase and malate dehydrogenase are reported as being located in peroxisomal-like microbodies (Lauersen et al, 2016). In *Chlorella vulgaris* cultured with acetate in the dark, glyoxysomal MS and IL showed a noticeable increase in their activity (Morales-Sanchez *et a*l, 2015), confirming their essential role in the glyoyxylate cycle. If key enzymes for the glyoxylate cycle are located outside of the glyoxysome/peroxisome, then acetate conversion may not take place as effectively (reduced acetate assimilation), succinate production might be hindered in the glyoxysomes and there might be a need for glyoxylate (which is a key component in the cycle). If key enzymes are all located in the glyoxysome/peroxisome (as single and fused) this might boost the cycle favoring cell growth by generating reducing equivalents and carbon skeletons, and FA synthesis through malate export to the plastid. Having glyoxysomes (and key glyoxylate cycle enzymes located in the glyoxysome) instead of peroxisomal microbodies can be an advantage for biomass and TAG production strains.

Beta oxidation of fatty acids in glyoxysome

Glyoxysomal beta oxidation of fatty acids in green microalgae is not well understood and studies describing such processes in the glyoxysome are very scarce. Kong *et al*, 2017 demonstrated that in *Chlamydomonas reinhardtii* an acyl-CoA oxidase was responsible for the first step of peroxisomal (microbodies) fatty beta oxidation, and highlighted the importance of this compartmentalized route in algal

physiology. The entire set of enzymes for beta oxidation in *Mk* KAS603 were predicted to be targeted to the glyoxysome. Fig. 2-10 predicts TAGs and possibly membrane lipids being catabolized in the cytoplasm and fatty acids being oxidized in the glyoxysome of *Mk* KAS603. This route generates $NADH_2$ and $FADH_2$ and it can contribute to fatty acid accumulation if it is shut down (which might be facilitated due to a predicted localization to only one compartment) or if Acyl CoA groups are recycled into fatty acid and TAG formation (outside the glyoxysome) (Trentacoste *et al* 2013; Tan and Lee, 2016). Two ECH isoforms and a fused ECH-HADH were identified, which might boost beta oxidation during the corresponding hydration and dehydration steps, for $NADH_2$ production/supply. Since the glyoxylate cycle and beta oxidation are proposed to take place (partially or totally) in the glyoxysome and AcCoA is proposed to be present in this same organelle (and apparently not transported across membranes); we propose a glyoxysomal interactive management of AcCoA among these routes to accommodate cellular needs in *Mk* KAS603, similarly to what is proposed to occur in *C. reinhardtii* (Lauersen *et al,* 2016) and plants (Oliver *et al,* 2009) where glyoxysomal AcCoA produced from FA breakdown feeds into the glyoxylate shunt for an efficient acetate assimilation.

Multicompartmentalized fatty acid and TAG synthesis

Current general knowledge of microalgal fatty acid biosynthesis is that it occurs in the plastid whereas TAG biosynthesis occurs in the endoplasmic reticulum (Sharma *et al*, 2015), however in some genuses such as *Chlamydomonas* and

Dunaliella, TAG synthesis can also take place in the plastid (Tan and Lee, 2016). In *Mk* KAS603 it is proposed that fatty acids are preferentially assembled in the plastid to a particular chain length, which can be assembled and elongated into long chain FAs in the cytoplasm. Desaturation of long chain polyunsaturated fatty acids is predicted to take place in the cytoplasm, and since there is one desaturase in the ER that makes C16:1 from C16:0, it is possible that short chain FAs are desaturated in the ER. As displayed in Fig. 2-11, TAG synthesis is proposed to take place in the cytoplasm and ER (where TAG enzymes are targeted) of *Mk* KAS603, similarly to *Chlorella pyrenoidosa* FACHB-9 (Fan *et al*, 2015).

Photorespiration and its proposed benefits for certain conditions in *Mk* KAS603

RubisCO oxygenase activity reduces the efficiency of CO_2 fixation through the requirement to regenerate the substrate RuBP, because fixed carbon is lost as CO_2 during the decarboxylation of serine in the photorespiratory pathway of green microalgae (Subramanian *et al*, 2013). Similar to higher plants and green algae, *Mk* KAS603 photorespiration series of reactions were localized to three subcellular compartments; the plastid, glyoxysome, and mitochondria (Fig. 2-12).

Photorespiration is a rich producer of glyoxylate, serine and glycine, which are suggested compounds to aid *Mk* KAS603 to enhance the glyoxylate cycle activity, or to regulate internal osmotic pressure when needed (glycine is proposed to serve as an osmolyte in *Mk* KAS603 when growing in hypersaline environments [see Chapter 1]). Although photorespiration comes to a high (energetic cell) cost, it might prove

useful for *Mk* KAS603 to produce osmolytes and for reinforcement of the glyoxlate cycle for reducing equivalents and carbon skeletons for cell division.

The pentose phosphate cycle and its potential involvement in FA synthesis

The metabolic steps required for the oxidation of glucose 6-phosphate to pentose phosphates and their recycling into glucose 6-phosphate, constitute the pentose phosphate pathway. Similarly to *C. pyrenoidosa* and to *C. protothecoides* (Tan and Lee, 2016) the pentose phosphate pathway in *Mk* KAS603 is proposed as a major source of reducing power for biosynthetic processes (such as FA-TAG biosynthesis). Previously reported (Tan and Lee, 2016) plastidial glucose 6-phosphate flux through the oxidative phase of the pentose phosphate pathway stimulates the incorporation of carbon from pyruvate into fatty acids. As displayed in Fig. 2-13, in *Mk* KAS603, initial oxidative pentose phosphate synthesis is localized to the plastid and it continues in the cytoplasm in order to produce energy and reducing equivalents as NADPH, potentially contributing for fatty acid synthesis in alternate compartments. Similarly to other Chlorella strains (Heath, 1984; Perez-Garcia, 2011; Morales-Sanchez, 2015) the oxidative phase is also proposed to contribute mainly for cell division and nucleic acid formation (using the pentose ribose 5-phoshate) and the non-oxidative phase with hexose phosphates (such as fructose 6-phosphate) for plastid starch formation. The pentose phosphate pathway of *Mk* KAS603 could generate NADPH in two compartments (plastid and cytoplasm) where FA synthesis is proposed to take place.

2.5.3 Carbon flux differences among green microalgae

Starch, glycolysis and gluconeogenesis

C. reinhardtii and *C. variabilis* NC64A were suggested to synthesize/breakdown and store starch in the plastid only, and G6P generated from starch breakdown may be exported from the plastid for glycolytic processing in the cytoplasm in *C. variabilis* NC64A (Figs. 2-14A, 2-15A, 2-14D and 2-15D). The other 4 algal species seemed to have relocated starch storage to other compartments. *S. obliquus* and *C. sorokiniana* were predicted to conduct starch metabolism in the ER and plastid, while *Mk* KAS603 in the cytoplasm and plastid, and *C. subellipsoidea* C-169 in the ER and cytoplasm – excluding the plastid entirely. Extraplastidial starch metabolism might give an advantage to the aforementioned algal species because it reduces competing reactions and increases specificity for particular reactions, if in the ER they may modified in their performance by acetylation. *C. reinhardtii* and *S. obliquus* were predicted to have only the upper portion of the glycolytic pathway present in the plastid and the lower portion in the cytoplasm, and gluconeogenesis was predicted to be located to the plastid (Figs. 2-14A, 2-15A, 2-14C and 2-15C). The predicted topology of carbon flux was different in the four other species being compared. *C. variabilis* NC64A was predicted to have glycolysis in the cytoplasm and gluconeogenesis in the plastid. *Mk* KAS603 and *C. subellipsoidea* C-169 were predicted to have glycolysis and gluconeogenesis in the cytoplasm and plastid (and a possible partial ER glycolysis for *C. subellipsoidea* C-169). *C. variabilis* NC64A and

C. reinhardtii were the only species that lacked localization of glycolysis/gluconeogenesis enzymes to the ER. On the other hand, *S. obliquus, C. sorokiniana, Mk* KAS603 and *C. subellipsoidea* C-169 had ER predicted targeting for some enzymes, especially *C. sorokiniana* and *C. subellipsoidea* C-169, that seemed to have a higher number of ER targeted proteins, including a key glycolytic enzyme (PFK). Separating carbon flux between different compartments could be a mechanism to increase metabolic pathway efficiency, since it may separate conflicting processes. In addition, different compartments could have particular conditions (e.g. reducing or energetic status, pH, etc.) that favor particular metabolic reactions. A commonality among all species analyzed is ENO being targeted to the cytoplasm only, which has been suggested as a major regulatory point in adjusting carbon partitioning in response to environmental changes in green microalgae (Polle *et al*, 2014) because it can not only facilitate or block PEP or pyruvate (and AcCoA) production, it might also regulate cell division and growth (that are affected by environmental changes). We have strong predicted evidence supporting such specific location and the ENO gene expression will be discussed further in chapter 3 for *Mk* KAS603 supporting some of Polle *et al* findings.

Pyruvate Hub, pyruvate fate and sources of AcCoA

C. variabilis NC64A is the only strain analyzed that seems to lack the capacity of producing AcCoA in the plastid (for FA synthesis) because it lacks PDH, ACAS and ACL. This might be a reason to be considered a non-oleaginous strain. All other

strains seem to have the potential to produce plastid AcCoA for FA anabolism. The cytoplasm of all strains except for *Mk* KAS603 seem to have AcCoA, which can be used for de novo FA synthesis of aminoacid synthesis. The mitochondrion of *S. obliquus* and *C. sorokininana* seem to lack AcCoA production, which is needed for the TCA cycle. It may be possible that another potential source of acetyl-CoA that utilizes acetate (acetate kinase and phosphate acetyltransferase identified in *C. reinhardtii* [Shtadia *et al*, 2015]) may be present in both strains. The ER of *S. obliquus, C. sorokiniana, Mk* KAS603 and *C. subellipsoidea* C-169, seem to be able process pyruvate through PDH to generate AcCoA for acetylation of ER-based protein substrates, for synthesis of aminoacids or for Acyl-CoA groups for TAG synthesis in this compartment (an example for ER TAG synthesis would be Mk KAS603 which has predicted TAG anabolism in the ER).

Oleaginous vs non-oleaginous strains and carbon management.

Compared to other green microalgae, *Mk* KAS603 (which is considered an oleaginous strain) displayed a commonality with the oleaginous strains (*C. sorokiniana*, *S. obliquus*, and Coccomyxa C-169) where predicted enzymes involved in carbohydrate synthesis/degradation or AcCoA synthesis were targeted to ER. This didn't occur in *C. reinhardtii* and *Chlorella sp.* NC64A (which are considered non-oleaginous strains). Enzymes that are modified by acetylation can also modify their activity and carbon flux through a pathway (Xing *et al*, 2012) We propose the ER as a very important organelle for regulating/coordinating global carbon flux and

facilitating FA-TAG formation in oleaginous green algal strains because they seem to have the capacity to synthesize ER AcCoA which can participate in the direct acetylation of proteins already in the ER involved in starch to glucose metabolism and TAG synthesis, and such enzyme/route modification might cause oil overproduction. Another suggested common trait for the oleaginous green microalgae analyzed (described in section 2.5.3) is predicted starch storage outside the plastid, which spatially separates the processes of storage and breakdown. The relevance of our work is that a possible mechanism for oleaginous strains where three main organelles (chloroplast, endoplasmic reticulum and cytoplasm) coordinate carbon flow for an efficient supply and partitioning for storage, energy production (for growth and cell division), as well as fatty acid and TAG accumulation; is a general trend in all TAG accumulating strains analyzed. Such multi-compartmentalized mechanism is not predicted to occur in the non-oleaginous strains analyzed.

2.6 ACKNOWLEDGEMENTS

We would like to acknowledge the Matteo Pellegrini lab at UCLA for *Mk* KAS603 genome sequencing and assembly, and for developing the genome browser.

Special thanks to Mark Hildebrand at Scripps Institution of Oceanography, UCSD for his guidance, patience and encouragement throughout this wonderful voyage of knowledge and learning. Funding for this project was provided by UC-MEXUS-CONACYT grant CN-12-622 , by AFOSR grant FA9550-08-1-0178 and by UCLA Genomics and Proteomics grant DE-FC02-02ER63421. Funding for Eva Sánchez was provided by the UC-MEXUS-CONACYT fellowship, and by Irving Tragen.

Chapter 2 contains unpublished work with co-authors: Mark Hildebrand, Matteo Pellegrini, David López, Jing Lu, Megan Inkeles, Weihong Yan and Shawn Cokus. The dissertation author was the primary researcher and author of this material.

2.7 REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990). Basic local alignment search tool. J. Mol. Biol. 215:403-410.

Avidan O, Brandis A, Rogachev I, Pick U. (2015) Enhanced acetyl-CoA production is associated with increased triglyceride accumulation in the green alga *Chlorella desiccata*. *J Exp Bot* . 66 (13): 3725-3735.

Bendtsen JD, Nielsen H, Heijne GV, Brunak S. (2004) Improved prediction of signal peptides: SignalP 3.0 J. Mol. Biol. 340:783-795.

Blatti JL, Michaud J, Burkart MD (2013) Engineering fatty acid biosynthesis in microalgae for sustainable biodiesel. Curr Opin Chem Biol 17:496–505.

Burger G, Nedelcu AM: Mitochondrial Genomes of Algae. In Genomics of Chloroplasts and Mitocchondria. vol. 35th Edition. Edited by Bock R, Knoop V. Netherlands: Springer; 2012:127-157.

Buleon A, Gallant DJ, Bouchet B, Mouille G, D'Hulst D, Kossmann J, Ball S. (1997) Starches from A to C (Chlamydomonas reinhardtii as a Model Microbial System to Investigate the Biosynthesis of the Plant Amylopectin Crystal. Plant Physiology. 115 (3) 949-957; DOI: 10.1104/pp.115.3.949.

Busi, M. V, Barchiesi, J., Martín, M., Gomez-Casati, D.F. (2014). Starch metabolism in green algae. Starch – Stärke. 66: 28–40.

Carrie C, Giraud E, Whelan J. (2009) Protein transport in organelles: Dual targeting of proteins to mitochondria and chloroplasts. FEBS Journal, 276: 1187–1195. doi:10.1111/j.1742-4658.2009.06876.x

Castro O, Chen LY, Parodi AJ, Abelijón C (1999) Uridine diphosphate-glucose transport into the endoplasmic reticulum of *Saccharomyces cerevisiae*: In vivo and in vitro evidence. Mol. Biol. Cell. 10:1019-1030.

Claros MG, Vincens P. (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur J Biochem. 241 (3):779-786.

Chow K-C, Shen H-B. (2010) Plant-mPLoc: A Top-Down Strategy to Augment the Power for Predicting Plant Protein Subcellular Localization. PLoS ONE 5(6): e11335.

Darley, W.M., and Volcani, B.E. (1969) Role of silicon in diatom metabolism. A silicon requirement for deoxyribonucleic acid synthesis in the diatom Cylindrotheca fusiformis Reimann and Lewin." *Experimental Cell Research*58.2-3 (1969): 334-342.

Díaz-Ramos A, Roig-Borrellas A, García-Melero A, López-Alemany R (2012) α-Enolase, a Multifunctional Protein: Its Role on Pathophysiological Situations. Journal of Biomedicine and Biotechnology. 2012:156795

Emanuelsson O, Brunak S, Von Heijne G, Nielsen H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. Nature protocols 2(4):953.

Emanuelsson O, Nielsen H, Von Heijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci. 8(5):978-984.

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

Fan J, Andre C, Xu Changcheng. (2011) A chloroplast pathway for the de novo biosynthesis of triacylglycerol in *Chlamydomonas reinhardtii.* FEBS Letters 585:1985-1991.

Fan J, Ning K, Zeng X, Luo Y, Wang D, Hu J, Li J, Xu H, Huang J, Wan M, Wang W. (2015) Genomic foundation of starch-to-lipid switch in Oleaginous Chlorella spp. Plant physiol. 1;169(4):2444-61.

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.

Goncalves EC, Wilkie AC, Kirst M, Rathinasabapathi B. (2016). Metabolic regulation of triacylglycerol accumulation in the green algae: identification of potential targets for engineering to improve oil yield. *Plant Biotechnology Journal*, *14*(8), 1649–1660.

Guarnieri MT, Nag A, Smolinski S, Darzins A, Seibert M, Pienkos PT. 2011. "Examination of Triacylglycerol Biosynthetic Pathways via De Novo Transcriptomic and Proteomic Analyses in an Unsequenced Microalga". PLoS ONE. 6(10): e25851.

Hass BJ, Delcher AL, Mount SM, Wortman JR, Smith Jr, RK Jr, Hannick LI, Maiti R, Ronning CM, Rusch DB, Town CD et al. (2003). Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. Nucleic Acids Res. 31:5654-5666 .

Heath R.L. (1984) Decline in energy reserves of *Chlorella sorokiniana* upon exposure to ozone. Planty physiology. 76(3):700-704.

Hilgert U., McKay S, Ghiban C, Khalfan M, Micklos D, Williams J (2014). DNA Subway: Making genome analysis egalitarian. In ACM International Conference Proceeding Series [70]. Association for Computing Machinery. DOI: 10.1145/2616498.2616575. https://dnasubway.cyverse.org/

Huerlimann, R. and Heimann, K. (2013) Comprehensive guide to acetyl-carboxylases in algae. Crit. Rev. Biotechnol. 33:49–65.

Jacobs J. D., Ludwig J. R., Hildebrand M., Kukel A., Feng T., Ord R. W., Volcani B. E. (1992). "Characterization of two circular plasmids from the marine diatom *Cylindrotheca fusiformis*: plasmids hybridize to chloroplast and nuclear DNA". Mol Gen Genet. 233:302-310.

Jeong H, Lim J M, Park J, Sim Y M, Choi H G, Lee J, Jeong W J. (2014) "Plastid and mitochondrion genomic sequences from Arctic *Chlorella* sp. ArM0029B". BMC Genomics.15:286.

Kent W James. (2002). BLAT-the BLAST-like alignment tool. Genome research 12.4:656-664.

Klein U. (1986). Compartmentation of glycolysis and of the oxidative pentose-phosphate pathway in *Chlamydomonas reinhardtii.* Planta.167:81-86.

Kong F, Liang Y., Légeret B, Beyly-Adriano A, Blangy S, Haslam RP, Napier JA, Beisson F, Peltier G, Li-Beisson Y. (2017). *Chlamydomonas* carries out fatty acid β-oxidation in ancestral peroxisomes using a bona fide acyl-CoA oxidase. Plant J, 90: 358–371. doi:10.1111/tpj.13498

Lauersen KJ, Willamme R, Coosemans N, Joris M, Kruse O and Remacle C (2016) Peroxisomal microbodies are at the crossroads of acetate assimilation in the green microalga *Chlamydomonas reinhardtii*. Algal Research. 16:266-276.

McAlister L. and Holland M. J. (1982). Targeted deletion of a yeast enolase structural gene. Identification and isolation of yeast enolase isozymes. *J. of Biol. Chem.* 257(12):7181–7188.

Maruyama S, Misawa K, Iseki M, Watanabe M, Nozaki H. (2008) Origins of a cyanobacterial 6-phosphogluconate dehydrogenase in plastid lacking eukaryotes. BMC Evol Biol. 8:151.

Mitschke J, Fuss J, Blumm T, Hoglund A, Reski R, Kohlbacher, Rensing S. (2009). Prediction of dual protein targeting to plant organelles. New Phytologist. 183:224-226.

Morales-Sanchez D, Martinez-Rodriguez O.A., Kyndt J, Martinez A (2015) Heterotrophic growth of microalgae: metabolic aspects. World J Microbiol Biotechnol. 31:1-9.

Neuberger G, Maurer-Stroh S, Eisenhaber B, Hartig A, Elsenhaber F. (2003) Prediction of peroxisomal targeting signal 1 containing proteins from amino acid sequences. J Mol Biol. 328(3):581-592.

Oliver DJ, Nikolau BJ, Wurtele ES. (2009). Acetyl-CoA-life at the metabolic nexus. Plant Science. 176(5):597-601.

Patron NJ, Keeling PJ. (2005). Common evolutionary origin of starch biosynthetic enzymes in green and red algae. J. Phycol. 41:1131-1141.

Peeters N, Small I.D, (2001) Dual targeting to mitochondria and chloroplasts. Biochimica et Biophysica Acta 1541(1-2):54-63.

Perez-Garcia O, Escalante F.M.E., de-Bashan L.E., Bashan Y. (2011). Heterotrophic cultures of microalgae: Metabolism and potential products. 45:11-36.

Polle J.E.W, Neofotis P, Huang A, Chang W, Sury K, Wiech E.M. (2014). Carbon partitioning in green algae (Chlorophyta) and the enolase enzyme. Metabolites. 4:612-628.

Reumann S, Buchwald D, Lingner T. (2012). PredPlantPTS1: A Web Server for the Prediction of Plant Peroxisomal Proteins. Frontiers in Plant Science. 3:194.

Sambrook, J. and Russel, D. W. 2001. *Molecular Cloning: A Laboratory Manual*. No. Ed. 3. Cold Spring Harbor Laboratory Press, NY.

Sharma T, Gour RS, Kant A, Chauhan RS. (2015). Lipid content in Scenedesmus species correlates with multiple genes of fatty acid and triacylglycerol biosynthetic pathways. Algal Research. 12:341-349.

Shtaida N, Khozin-Goldberg I, Boussiba S. (2015) The role of pyruvate hub enzymes in supplying carbon precursors for fatty acid synthesis in photosynthetic microalgae. Photosynthesis research. 1;125(3):407-22.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. (2011). Fast, scalable

generation of high-quality protein multiple sequences alignments using Clustal Omega. Molecular Systems Biology. 7:539.

Small I, Peeters N, Legeai F, Lurin C (2004) Predorar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics. 4(6):1581-1590.

Smith D R, Burki F, Yamada T, Grimwood J, Grigoriev I V, Van Etten J L, Keeling P J, (2011) "The GC-Rich Mitochondrial and Plastid Genomes of the Green Alga *Coccomyxa* Give Insight into the Evolution of Organelle DNA Nucleotide Landscape". PLoS ONE. 6 (8) : 1-10.

Solovyev V, Kosarev P, Seledsov I, Vorobyev D. (2006) Automatic annotation of eukaryotic genes, pseudogenes and promoters. Genome Biol. 7(1):10.1-10.12.

Subramanian S, Barr A.N., Pieris S, Sayre R.T. (2013). Comparative energetics and kinetics of autotrophic lipid and starch metabolism in chlorophytic microalgae: implications for biomass and biofuel production. Biotechnology for Biofuels. 6:150.

Stanke M, Steinkamp R, Waack S, Morgenstern B (2004). Nucleic Acids Res. 32(2):w309-w312. DOI: https://doi.org/10.1093/nar/gkh379

Tan K.W.M and Lee Y.K. (2016). The dilemma for lipid productivity in green microalgae: importance of substrate provision in improving oil yield without sacrificing growth. Biotechnol Biofuels 9:255.

Tardif M, Atteia A, Specht M, Cogne G, Rolland N, Brugiere S, Hippler M, Ferro M, Peltier G, Vallon O, Cournac L. (2012). PredAlgo: a new subcellular localization predition tool dedicated to green algae. Molecular biology and evolution 29(12):3625-3639.

Trentacoste, E.M., Shrestha R.P., Smith S.R., Glé C., Hartmann A.C., Hildebrand M, Gerwick W. H. (2013), "Metabolic engineering of lipid catabolism increases microalgal lipid accumulation without compromising growth". PNAS. 110 (49): 19748–19753.

Turmel M, Otis C, Lemieux C (2009b) The chloroplast genomes of the green algae *Pedinomonas minor*, *Parachlorella kessleri*, and *Oocystis solitaria* reveal a shared ancestry between the Pedinomonadales and Chlorellales. Mol Biol Evol 26:2317–2331.

Viola R, Nyvall P, Pedersen M (2001) The unique features of starch metabolism in red algae. Proc. Biol. Sci. 268:1417–1422.

Walsh JL, Keith TJ, Knull HR (1989) Glycolytic enzyme interactions with tubulin and microtubules *Biochimica et Biophysica Acta*. 999(1):64–70.

Wang Y, Coelmann-Derr D, Chen G, Gu YQ. (2015). OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. Nucleic Acids Research (web server issue). 43:W78-W84.

Wakasugi T, Nagai T, Kapoor M, Sugita M, Ito M, Ito S, Tsudzuki J, Nakashima K, Tsudzuki T, Suzuki Y, Hamada A, Ohta T, Inamura A, Yoshinaga K, Sugiura M: Complete nucleotide sequence of the chloroplast genome from the green alga Chlorella vulgaris: The existence of genes possibly involved in chloroplast division. Proc Natl Acad Sci U S A 1997, 94(11):5967–5972.

Xie X, Huang A, Gu W, Zang Z, Pan G, Gao S, He L, Zhang B, Niu J, Lin A, Wang G. (2015). Photorespiration participates in the assimilation of acetate in Chlorella sorokiniana under high light. New Phytologist. 209:987-998.

Xing S, Poirier Y. (2012). The protein acetylome and the regulation of metabolism. Trends Plant Sci. 17:423-430.

Supplemental Table S2-1. Abbreviations for enzymes corresponding to starch metabolism, glycolysis, gluconeogenesis, sucrose metabolism and pyruvate hub, in *Mk* KAS603.

Pathway	Protein name	Abbreviation
Starch Anabolism	Soluble starch synthase II	SS
	Starch synthase granule-bound	SS
	ADP glucose pyrophosphorylase	AGP
	Starch branching enzyme II	SBE
	Phosphoglucomutase	PGM
Starch catabolism	Beta–amylase	β amy
	Alpha-amylase	α amy
	Alpha-amylase	α amy
	Starch phosphorylase	SP
	Isoamylase	IA
	Alpha-glucosidase	α gluc
Sucrose metabolism	Bifunctional 6-phosphofructo-2-kinase/fructose 2,6-bisphophate 2-phosphatase	PFK2/FBP2
	Sucrose 6-phosphate synthase	S6PS
	Sucrose 6-phosphate phosphatase	S6PP
Glycolysis	Glucokinase	GK
	Glucose 6-phosphate isomerase	GPI
	6-Phosphofructokinase	PFK
	Fructose 1,6 bisphosphate aldolase	FBA
	Triose phosphate isomerase	TPI
	Glyceraldehyde 3-phosphate dehydrogenase	GAPDH
	Phosphoglycerate kinase	PGK
	Phosphoglycerate mutase	PGAM
	Enolase	ENO
	Pyruvate kinase	PK
Gluconeogenesis	Pyruvate carboxylase	PC
	Phosphoenol pyruvate carboxykinase	PEPCK
	Fructose 1,6 bisphosphatase	FBP
	Phosphoenol pyruvate carboxylase	PEPC
Pyruvate hub	Malate dehydrogenase,	MDH
	Malic enzyme NAD dependent	ME
	Malic enzyme NADP dependent	ME_{NADP}
	Pyruvate phosphate dikinase	PPDK
	Pyruvate dehydrogenase E1 component subunit alpha	PDH E1α
	Pyruvate dehydrogenase E1 component subunit beta	PDH E1 β
	Pyruvate dehydrogenase E2 component	PDH E2

Supplemental Table S2-2. Abbreviations for enzymes corresponding to fatty acid and TAG anabolism, in *Mk* KAS603

Pathway	Protein name	Abbreviation
FA-TAG anabolism	Acetyl-coenzyme A synthetase	ACAS
	ATP citrate lyase	ACL
	Acetyl-coenzyme A carboxylase	accA
	Acetyl-coenzyme A carboxylase carboxyltransferase subunit alpha	accA α
	Malonyl-CoA-acyl carrier protein transacylase	MCAT
	Ketoacyl-acyl carrier protein synthase III	KASIII
	ketoacyl-acyl carrier protein synthase II	KAS II
	Ketoacyl-acyl carrier protein synthase	KAS I
	3-oxoacyl-ACP reductase	KAR
	3-hydroxyacyl-[acyl-carrier-protein] dehydratase	HD
	Enoyl-[acyl-carrier-protein] reductase [NADH]	ENR
	Omega-3 FA desaturase	O3D-III
	Omega-6 FA desaturase	O6D
	Long-chain acyl-CoA synthetase	ACS
	Glycerol-3-phosphate dehydrogenase	G3PDH
	Glycerol kinase	GK
	Glycerol-3-phosphate acyltransferase	GPAT
	Lysophosphatidic acid acyltransferase	LPAAT
	Phosphatidate phosphatase	PAP
	Diacylglycerol O-acyltransferase	DGAT

Supplemental Table S2-3. Abbreviations for enzymes corresponding to the glyoxylate cycle, in *Mk* KAS603.

Pathway	Protein name	Abbreviation
Glyoxylate cycle	Acetyl-coenzyme A synthetase	ACAS
	Citrate synthase	CS
	Aconitase hydratase	AH
	Isocitrate lyase	IL
	Malate synthase	MS
	Malate dehydrogenase	MDH

Supplemental Table S2-4. Abbreviations for enzymes corresponding to lipolysis, in *Mk* KAS603.

Pathway	Protein name	Abbreviation
FA-TAG catabolism	TAG lipase SDP1	TL
	Long chain acyl-CoA synthetase	ACL
	Acetyl-coenzyme A oxidase	ACOS
	Enoyl-CoA hydratase	ECH
	3-hydroxyacyl-CoA dehydrogenase	HADH
	Acyl CoA acetyltransferase (thiolase)	ACAT
	2,4-dienoyl-CoA reductase	DECR

Chapter 3

Transcriptomic analysis of genes involved in multiple fission and carbon metabolism in *Marinichlorella kaistiae* KAS603

3.1 ABSTRACT

Mk KAS603 was evaluated for multiple fission division and carbon metabolism gene expression response under 10 different cultivation and trophic regimes where log and stationary phases were compared. While growing at an accelerated rate with acetate and bicarbonate, the cell is predicted to activate all carbon core metabolic routes (mostly in the cytoplasm), accumulating starch and initiating log phase TAG accumulation. FA synthesis and FA oxidation seem to continuously occur in growing cells to replenish energy and carbon. Routes that can contribute to FA-TAG accumulation are glycolysis, PPP and glyoxylate cycle. While growing rapidly in acetate supplemented ASW, the cell seems to activate gluconeogenesis in the plastid for starch accumulation, FA are possibly being produced for cell membrane formation, and glycerol seems to accumulate as a possible osmolyte in growing cells. TAG only accumulates during stationary phase in acetate supplemented ASW and both FA synthesis and degradation seem to be active to replenish energy and carbon storage, cytoplasmic glycolysis and PPP seem contribute to TAG synthesis. While actively growing in hypersalinity, the cell is predicted to activate gluconeogenesis in the plastid to oxidate glucose through the PPP for ribose-5-phosphate for DNA and RNA synthesis. The TCA cycle is predicted as activated for energy generation, glycerol (osmolyte) synthesis is predicted in growing cells. During stationary phase in hypersalinity, gluconeogenesis seems to favor starch accumulation in the plastid, FA beta oxidation seems to shut down as FA

synthesis is turned on, enriching FA availability for TAG accumulation. Photorespiration is predicted to be activated for glycine and serine (predicted osmolites) synthesis. Gene expression heat maps and predictive metabolic models were developed per condition. This work has provided insights into mechanisms of high vs low productivity and cultivation flexibility in the oleagionous green microalgae *Mk KAS603*.

3.2 INTRODUCTION

Transcriptomic analysis in oleaginous green microalgae has provided insights into metabolic pathway interactions and regulatory mechanisms associated to TAG accumulation (Rismani-Yazdi *et al*, 2012; Gao *et al*, 2014; Fan *et al*, 2015; Sharma *et al*, 2015). Such analysis has been performed under a limited number of experimental conditions, due to the lack of strains with a broad capacity to adapt to various cultivation environments. For this chapter, the environmentally flexible green microalgae (*Marinichlorella kaistiae* KAS603) was examined under a broad variety of conditions for gene expression during log and stationary phases of growth while cultivated with bicarbonate, with acetate, with acetate and bicarbonate, under high salinity, and with low salinity. Through transcriptomic analysis of genes involved in multiple fission and carbon metabolism we have been able to identify key metabolic steps in different cell compartments followed by *Mk* KAS603 during the aforementioned conditions. Coupling physiological responses of the strain described in chapter 1 with the genetic and metabolic capability described in chapter 2, as well

as gene expression analysis described in this chapter, has enabled us to understand possible survival mechanisms of *Mk* KAS603 in different environmental conditions while maintaining productivity. The answer seems to lie in coordinated growth related metabolic shifts in different cell compartments to save energy. During mixotrophic growth, starch accumulation seems to be highly stimulated in the cytoplasm, but in stationary phase, it seems to return to normal. During hypersalinity, different osmolytes seem to be produced during log (glycerol) and during stationary (serine, glycine) phase, and FA seem to accumulate (in cytoplasm and plastid) after accelerated growth mostly due to shutdown of FA beta oxidation. We are including a mixotrophic condition that is not mentioned in chapter 1, which is the supplementation of high bicarbonate combined with medium acetate in MS4 medium, where TAG synthesis is early induced (day 3 of cultivation), mother cells (with captive daughters) are prevalent during TAG accumulation and all cell sizes seemed to accumulate TAG. This was the condition where most metabolic routes where induced in more than one cell compartment.

3.3 METHODS

3.3.1 Cultivation conditions

Wild type *Marinichlorella kaistiae* strain KAS603 provided by Kuehnle AgroSystems, Inc. in Hawaii, was adapted to grow under several culture conditions (biological triplicates were performed for each culture condition tested), from starting liquid stocks cultured in freshwater modified Shuisheng-4 (MS4) medium (Kim et al,

2014), consisting of (per liter) 0.025 g $CaCl \cdot H_20$; 0.1 g $MgSO_4$; 0.033 g $NH_4H_2PO_4$; 0.129 g KCl; 0.250 g $(NH_2)_2CO$; 15 g $NaHCO_3$; 3.15 mg $FeCl_3$; 0.309 mg H_3BO_3; 0.169 mg $MnSO_4 \cdot H_2O$; 0.0288 mg $ZnSO_4 \cdot 7H_2O$; 0.0618 mg $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$; 0.0146 mg $Co(NO_3)_2 \cdot 6H_2O$; 0.0016 mg $VOSO_4 \cdot 6H_2O$; 0.0474 mg $AlKO_8S_2 \cdot 12H_2O$; 0.0198 mg $Ni(NH_4)_2(SO_4)_2 \cdot 6H_2O$; 0.0154 mg $Cd(NO_3)_2.4H_2O$; 0.0040 mg $CrN_3O_9 \cdot 9H_2O$; 0.0033 mg $Na_2O_4W \cdot 2H_2O$; 0.0019 mg KBr; 0.0083 mg KI; 0.0125 mg $CuSO_4.5H_2O$. The medium was adjusted to pH 10.5 prior to autoclaving and pH remained constant throughout the duration of the experiments. Algal cells were cultivated for 3 days (log phase) with orbital shaking at 150 rpm under continuous cool fluorescent light intensity of 150 μmol photons m^{-2} s^{-1}, and a temperature of 25°C, in 125 mL Erlenmeyer flasks containing 50 mL of MS4 medium (containing 178 mM $NaHCO_3$). Algal cells in exponential growth phase were harvested, pelleted and used as inoculant (to grow immediately) for freshwater (MS4 with 178 mM $NaHC0_3$ and MS4 with 178 mM sodium bicarbonate + 50 mM sodium acetate), and seawater (ASW, 0.4 M NaCl with 0 or 50 mM sodium acetate) batch cultures. From seawater (ASW, 0.4 M NaCl) they were adapted to hypersalinity (ASW, 1.5 M NaCl) during a three-month period (see chapter 1). In seawater and hypersaline conditions, a constant pH of 8.5 was maintained. All flask cultures were grown without supplemental CO_2. Culture density was measured by hemocytometer counts, performed every 24 hours. Specific growth rate was determined among conditions and time points, using the following equation:

$$\mu = \ln (X_2) - \ln (X_1)/(t_2-t_1)$$

in which X_1 and X_2 are the mean cell numbers of *M. kaistiae* KAS603 at the times t_1 and t_2 respectively. Since t was expressed in days, growth rate was converted to division or doublings per day by dividing μ by the natural log of 2 = 0.6931

$$K= \mu / 0.6931$$

3.3.2 Imaging flow cytometry analysis

$1x10^6$ cells were harvested for each condition, pelleted in microfuge tubes, the supernatant was removed and the pellet immediately frozen at -20°C. At the end of each trial, all pellets were thawed, resuspended in 1 mL of potassium phosphate buffer (0.1 M, pH 7) for freshwater cultures, or 1 mL of 2.3% NaCl solution for seawater cultures (or 7.0% NaCl solution for hypersaline adapted cultures), and stained with 2.6 μL of a 1 $mg.mL^{-1}$ BODIPY (4,4-difluoro-3a,4a-diaza-s-indacene Invitrogen,USA) stock dissolved in DMSO. Stained cells were maintained at room temperature for 15 minutes, put on ice and processed (10,000 to 20,000 cells per run) through an Amnis ImageStream X imaging flow cytometer. IDEAS™ software was used to analyze the data in order to obtain a final measurement of mean BODIPY fluorescence per cell (indicating the relative TAG content per cell - reported as relative fluorescence units [RFU]), mean chlorophyll autofluorescence per cell (to indicate relative chlorophyll content per cell), and cell area population distribution patterns (to indicate cell size variability and tendency per treatment/population). The

number and % of mother and daughter cells was calculated by measuring cell area of thousands (10,000 to 20,000) of individual cells per population and per condition in log and stationary phase using imaging flow cytometry. Frequency histograms of cell area population distribution were created per condition, and microscopic images from such histograms were examined to determine cell division/size tendencies among daughters and mothers.

3.3.3 TAG yield

M. kaistiae KAS603 TAG yield under all previously mentioned culture conditions was calculated as:

$$TY = (Cd) * (RTC)$$

In which, TY represents relative TAG yield, Cd represents culture density, and RTC is the relative TAG content per cell. TY was expressed in relative TAG yield units (RTYU).

3.3.4 RNA isolation and sequencing

The conditions for this study are as follows (conditions 5-10 were introduced in chapter 1 [Materials and Methods]):

1. Log phase autotrophic freshwater culture. *Mk KAS603* cultured in sodium bicarbonate (178 mM) freshwater modified Shuisheng-4 (MS4) medium. *Control for 3**

2. Stationary phase autotrophic freshwater culture. *Mk KAS603* cultured in sodium bicarbonate (178 mM) freshwater modified Shuisheng-4 (MS4) medium. *Control for 4**

3. Log phase mixotrophic freshwater culture. *Mk KAS603* cultured in sodium bicarbonate (178 mM) freshwater modified Shuisheng-4 (MS4) medium supplemented with sodium acetate (50 mM). Experimental condition.

4. Stationary phase mixotrophic freshwater culture. *Mk KAS603* cultured in sodium bicarbonate (178 mM) freshwater modified Shuisheng-4 (MS4) medium, supplemented with sodium acetate (50 mM). Experimental condition.

5. Log phase autotrophic seawater culture. *Mk KAS603* cultured in ASW medium (0.4 M NaCl). *Control for 7 and 9**

6. Stationary phase autotrophic seawater culture. *Mk KAS603* cultured in ASW medium (0.4 M NaCl) *Control for 8 and 10**

7. Log phase mixotrophic seawater culture. *Mk KAS603* cultured in ASW medium (0.4 M NaCl), supplemented with sodium acetate (50 mM). Experimental condition.

8. Stationary phase mixotrophic seawater culture. *Mk KAS603* cultured in ASW medium (0.4 M NaCl), supplemented with sodium acetate (50 mM). Experimental condition.

9. Log phase autotrophic hypersaline culture. *Mk KAS603* cultured in ASW medium with high NaCl (1.5 M). Experimental condition.

10. Stationary phase autotrophic hypersaline culture. *Mk KAS603* cultured in ASW medium with high NaCl (1.5 M). Experimental condition.

*Autotrophic cultures 1,2,5 and 6, are control conditions and all mixotrophic (3,4,7,8) and hypersaline (9,10) cultures are experimental conditions.

Mk KAS603 was harvested for each condition and algae pellets were frozen overnight at -80°C. The pellets were ground in dry ice and total RNA was extracted and purified from each condition using established methods (Hildebrand and Dahlin, 2000), and sent to the Pellegrini lab at UCLA for library generation and sequencing. The quality of purified RNA was determined on an Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). RNA-seq libraries were generated using the Illumina TruSeq Sample Preparation Kit v2 according to manufacturer's instruction, starting from 1 ug of DNaseI-treated total RNA. Briefly, the mRNA was isolated from total RNA using oligo-dT beads (Illumina). Bound RNA was eluted and fragmented in the presence of divalent cations at 94°C, and subsequently converted into double stranded cDNA using random hexamer primers. After polishing the ends of the cDNA, a single adenine base was added to the 3' ends of cDNA fragments. TruSeq RNAseq (Illumina) adapters were then ligated to cDNA 3' ends. Adapter-ligated cDNA was PCR-amplified for 15 cycles, and the DNA was purified with AMPure XP magnetic beads (Beckman Coulter). The concentration of the final

library was measured with Qubit BR Assay (Life Technologies) and the average size determined on a 2% Agarose Gel. Each library was diluted to 10 nM in EB buffer + 0.1% Tween-20 and up to six samples were combined together for each 100 bp paired end lane run on an Illumina HiSeq2000. Paired-end unstranded 100 bp reads with an average of 30 million paired reads per condition were generated. For quality control, raw sequencing reads with perfect index were analyzed by FastQC tool (v0.10.0) (Andrews, 2011).

3.3.5 Assembly

De novo transcriptomics data was assembled at UCLA using an assembler for short read sequence data (SOAPdenovo-Trans). This pipeline was designed for assembling transcriptomes with alternative splicing and different expression level among transcripts. It provided a complete way to construct full-length transcript sets. The formerly mentioned 10 samples were assembled together, yielding over 50 million pre-assembly reads with a mean size of 442 bp. The transcripts were mostly fragmented, and were subsequently run using AUGUSTUS (Stanke *et al*, 2004) to elongate or truncate *Mk* KAS603 genes from the nuclear genome. Final clustering of transcripts was obtained and non-redundant contigs were generated. *De novo* assembly of transcripts was used to make gene predictions. The reads (paired-end 50+50bp) were aligned with STAR (Dobin *et al*, 2013), quantified with HTSeq (Anders *et al*, 2015), and normalized with DESeq2 (Love *et al*, 2014).

3.3.6 Annotation

From the final set of contigs, a search to identify protein coding genes was conducted using FGenesH (Solovyev *et al*, 2006), AUGUSTUS (Stanke *et al*, 2004), and MAKER (Cantarel *et al*, 2008). Individual (unconcatenated) genome contigs and transcripts were submitted to AUGUSTUS for annotation. The first round of the AUGUSTUS pipeline aligned the transcript data to the genome, and used tools such as PASA (Haas *et al*, 2003) to assemble putative, high-confidence gene models from the regions with the most alignments and highest fidelity, from which training parameters were called. In a second round, those training parameters were tested and calibrated against the remaining transcript evidence, and final gene set was produced from those final parameters. Contigs with significant matches were annotated using the Algal Functional Annotation tool. To determine transcript abundances and differential expression, high quality reads from each experimental condition were individually mapped to the assembled transcriptome using Bowtie software (v0.12.7). Reads mapping to each contig were counted using SAM-tools (v0.1.16) and transcript abundances were calculated as reads per kilobase of exon model per million mapped reads (RPKM). The counts were displayed in a tab-delimited text file (that could be opened in Excel) and the headers described the different experiments. All differential expression analysis (fold changes) and related statistical computations was conducted by feeding non-normalized read counts into the DESeq package (v1.5.1).

3.3.7 Gene expression heatmaps

After sequence assembly and annotation, the data was mined to identify genes of interest with significant up- or down-regulation under the different test conditions relative to controls. Gene expression changes per metabolic route were displayed as heatmaps (which were constructed using RStudio v0.98.1103) to identify pathway-specific changes. Quantitative comparison of transcript levels between each condition allowed the identification of points of up- or down-regulation that were diagnostic of metabolic changes. A proposed global metabolic heatmap (Fig. 3-3) [based on combined genome (described on chapter 2) and transcriptome data, as well as protein targeting] (also described on chapter 2) of cellular adaptation responses to all conditions, was developed for *Mk* KAS603. In order to obtain the most accurate gene expression coverage, two different transcriptome data sets (with a different RPKM scale value) were combined to construct the heatmaps, and since the RPKM values couldn't be directly compared, we compared log based 2 fold changes for our analysis.

3.4 RESULTS AND DISCUSSION

3.4.1 Quality of gene expression data

The sequencing data (paired-end 50bp reads) mapped well to the nuclear genome (over the 10 datasets, only about 2-4% of reads are multi-mappers) and the normalization and quantification steps were precise. The normalized counts for the 10 RNAseq experiments aligned to the transcriptome. The level of multi-mapped reads was very low across all samples (~2-3% of all reads), and most reads were aligned to

the transcriptome. Despite the fact that single samples per condition were used for transcriptomic analysis, we are confident about the numbers because in addition to having good quality gene expression data, we also have experimental data for cell growth, TAG accumulation, TAG yield, TAG accumulation response relative to cell area of a population and cell area population distribution comparison among conditions; which complements and gives support and meaning to the transcriptomic analysis. Since we don't have reproducible data, we used a 1-fold (log based 2) change in expression as a conservative criterion for significance, taking into account that experiment to experiment variation would be expected to be much less than that.

3.4.2 Comparison of bicarbonate and acetate vs bicarbonate alone.

To evaluate the relative contribution of bicarbonate and acetate to growth and productivity, wild type *Mk KAS603* was cultured autotrophically in MS4 medium under very high $NaHCO_3$ (178 mM), and in a combination of medium acetate (50 mM) and very high bicarbonate (178 mM). Very high bicarbonate was used in this trial to detect if inorganic carbon (in excess) alone or in combination with organic carbon might have a different effect on productivity than the previously tested amount (150 mM - Chapter 1). As displayed in Fig. 3-1, the combination of acetate and bicarbonate had a greater effect on maximum cell density than bicarbonate. A maximum of 2.8 doublings per day were achieved under the organic/inorganic carbon source combination, this was a 1.4 fold growth rate increase compared to only bicarbonate in the medium. Highest TAG accumulation (Fig. 3-1B) and TAG yield

(Fig. 3-1C) occurred under the combination of bicarbonate and acetate. Indistinctive of size, cells tended to accumulate TAG (Day 6) in the combination of bicarbonate and acetate (Fig. 3-1D). From our examination of the size distribution of cells during day 6, an enrichment for small cells (ranging from 30 to 40 μ^2) occurred with high bicarbonate alone (Fig. 3-2A) and a distribution toward bigger cells with enrichment from 50 to 75 μ^2 occurred under the carbon combination (Fig. 3-2B); this can be explained as more daughters per mother cell accumulating under carbon addition. In summary, medium acetate (50 mM) combined with very high bicarbonate (178 mM) provided fast growth, high TAG, and an enrichment for bigger (mother) cells. Very high bicarbonate alone doesn't contribute to TAG productivity.

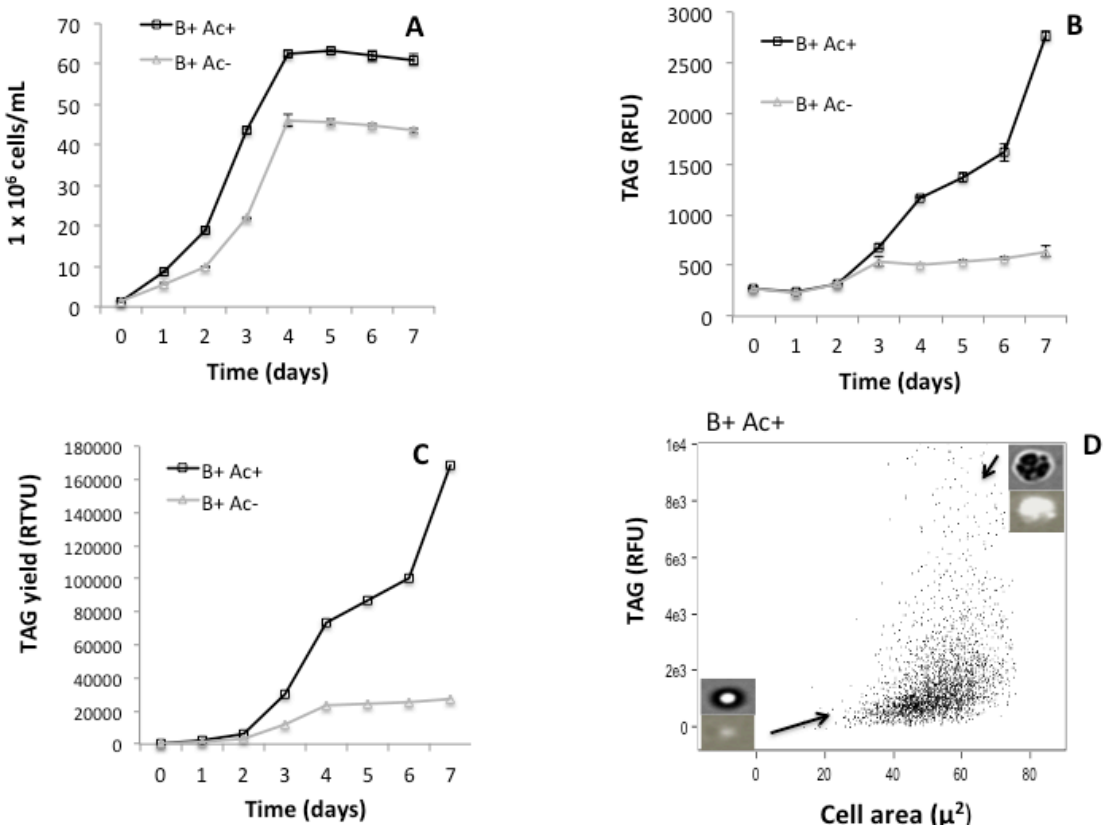

Figure 3-1. *M. kaistiae* KAS603 response to bicarbonate alone and in combination with acetate in freshwater MS4 medium. A) Culture density, B) TAG accumulation, C) TAG yield, D) TAG accumulation response relative to cell area of a population in MS4 supplemented with bicarbonate and acetate on day 6 of cultivation. Cell images inside squares are bright field (upper squares) and BODIPY (TAG) fluorescence (lower squares). Black arrows indicate the place on scatter plot where cells in square images are located. B+ represents sodium bicarbonate (178 mM), Ac+ represents sodium acetate (50 mM), B- represents absence of sodium bicarbonate and Ac- represents absence of sodium acetate. RFU denote relative fluorescence units. RTYU are relative TAG yield units. The error bars indicate standard error (n=3).

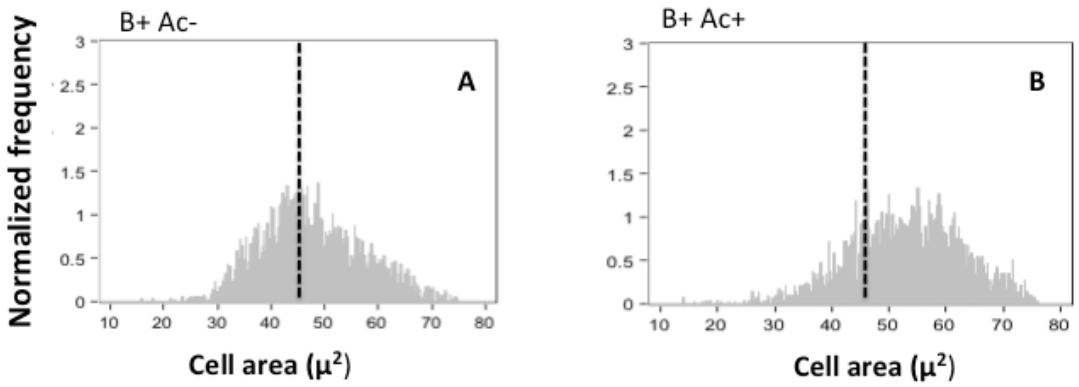

Figure 3-2. Cell area population distribution comparison among cells cultivated under A) presence (B+ Ac-) of high (178 mM) sodium bicarbonate, and B) a combination (B+ Ac+) of sodium acetate (50 mM) and high sodium bicarbonate (178 mM) in MS4 medium. All histograms represent the cell area population frequency distribution during day 6 of the cultures. The mean cell areas are: 47.7 μ^2 for A, and 52.9 μ^2 for B. Black dotted vertical line separates daughter cells (left side) from mother cells (right side)

3.4.3 Main highlights of cultivation conditions

All cultivation conditions analyzed are mentioned in Table 3-1, which includes abbreviated name, description and productivity characteristics for each condition. Cells cultivated in Acb displayed faster growth, larger size (during TAG accumulation), and higher TAG yield, than in b. Cells cultivated in Acsw displayed faster growth, smaller size (during TAG accumulation), and higher TAG yield, than in sw. Cells cultivated in hs displayed similar growth, larger size, and higher TAG yield than in sw. In a cell population, all cell sizes accumulated TAG in the Acb condition (Fig. 3-1A), smaller cells accumulated TAG in the Acsw condition, and medium and large cells accumulated TAG in the hs condition. Growth, TAG content, TAG productivity, and cell area data for cultivation conditions in seawater and hypersalinity (conditions 5-10) are mentioned in the results section of chapter 1.

Table 3-1. Summary of *Mk* KAS603 cultivation conditions used for gene expression comparison. Growth rate is displayed during log phase. Mean cell size (area) and TAG yield are displayed during stationary phase. RTYU represent relative TAG yield units (see methods section 3.3.3).

Condition name	Condition description	Growth rate	Mean cell size	TAG yield (RTYU)
Acb.L	178 mM sodium bicarbonate + 50 mM sodium acetate MS4 during log phase	2.7		
b.L	178 mM sodium bicarbonate MS4 during log phase	1.7		
Acsw.L	50 mM sodium acetate ASW during log phase	2		
sw.L	0.4 M NaCl ASW during log phase	1.5		
hs.L	1.5 M NaCl ASW during log phase	1.5		
Acb.S	178 mM sodium bicarbonate + 50 mM sodium acetate MS4 during stationary phase		53 μ^2	100,000
b.S	178 mM sodium bicarbonate in MS4 during stationary phase		47 μ^2	20,000
Acsw.S	0.4 M NaCl + 50 mM sodium acetate in ASW during stationary phase		32 μ^2	220,000
sw.S	0.4 M NaCl ASW during stationary phase		40 μ^2	50,000
hs.S	1.5 M NaCl ASW during stationary phase		45 μ^2	600,000

3.4.4 Distinctive gene expression patterns per growth stage and culture condition

Specific changes of expression were detected in genes that coded for proteins involved in metabolic pathways and multiple fission cell division. Such changes are proposed to contribute to adaptability and productivity under the various culture conditions tested. We will describe specific metabolic/physiological processes in *Mk* KAS603 starting with an overview, followed by specific comments (to relevant key gene product changes per cultivation condition) and a brief summary of such changes.

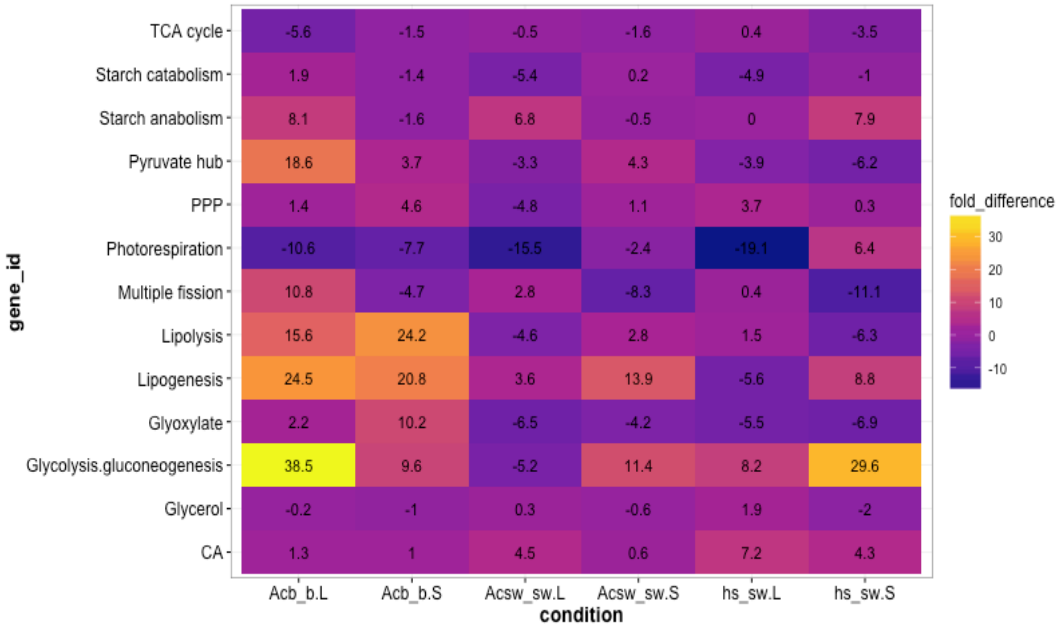

Figure 3-3. *Mk* KAS603 global gene expression heatmap. Acb_b represents acetate (50 mM) + bicarbonate (178 mM) compared to bicarbonate (178 mM) in MS4 freshwater medium. Acsw_sw represents acetate (50 mM) in ASW compared to ASW. hs_sw represents ASW (1.5 M NaCl) compared to ASW (0.4 M NaCl). L denotes log phase and S denotes stationary phase. On the Y axis, genes for CA (carbonic anhydrase); for PPP (the pentose phosphate pathway), for Glyoxylate (glyoxylate cycle); and for Glycerol (glycerol metabolism), are abbreviated. The numbers inside each rectangle as well as the colors, represent total gene expression fold difference per condition. All numbers are log 2 based and represent the sum of the fold change values for each gene being analyzed per pathway.

Multiple fission division

Overview

Multiple fission is a type of cell division that starts with a small unitary cell, called a daughter cell that acquires nutrients and CO_2 from its liquid environment and can grow to acquire an optimum size for the start of cell division. The daughter cell then divides into two siblings, those two cells can divide into 4, 8, 16 and up to 32 cells. This division process takes place inside the cell wall of the original cell, which is called a mother cell. We propose that enzymes that participate in DNA replication, transcription, translation, as well as cell cycle enzymes and cell wall degradation; can participate in multiple fission division in *Mk* KAS603. We are also proposing two transcription factors that might be involved in such a process. As displayed in Fig. 3-4 (depicting summed fold change values for each gene analyzed to derive a total value), the multiple fission gene set (which is composed of selected genes that participate in the aforementioned process), displays the highest global upregulation (10.8 total fold) during log phase in Acb [condition in which *Mk* KAS603 has the fastest growth rate (2.7 doublings per day)], followed by log phase in Acsw with a 2.8 total fold gene upregulation [and the second fastest growth rate (2 doublings per day)], and by log phase in hs, with 0.4 total fold gene expression [with the slowest growth rate (1.5 doublings per day)]. During stationary phase in all conditions, the multiple fission genes were downregulated the most in hs (-11.1 total fold downregulation), followed by Acsw (-8.3 total fold downregulation) and by Acb (-4.7 total fold downregulation). This suggests that genes involved in multiple fission seem

to be upregulated according to growth rate and downregulated as growth decreases.

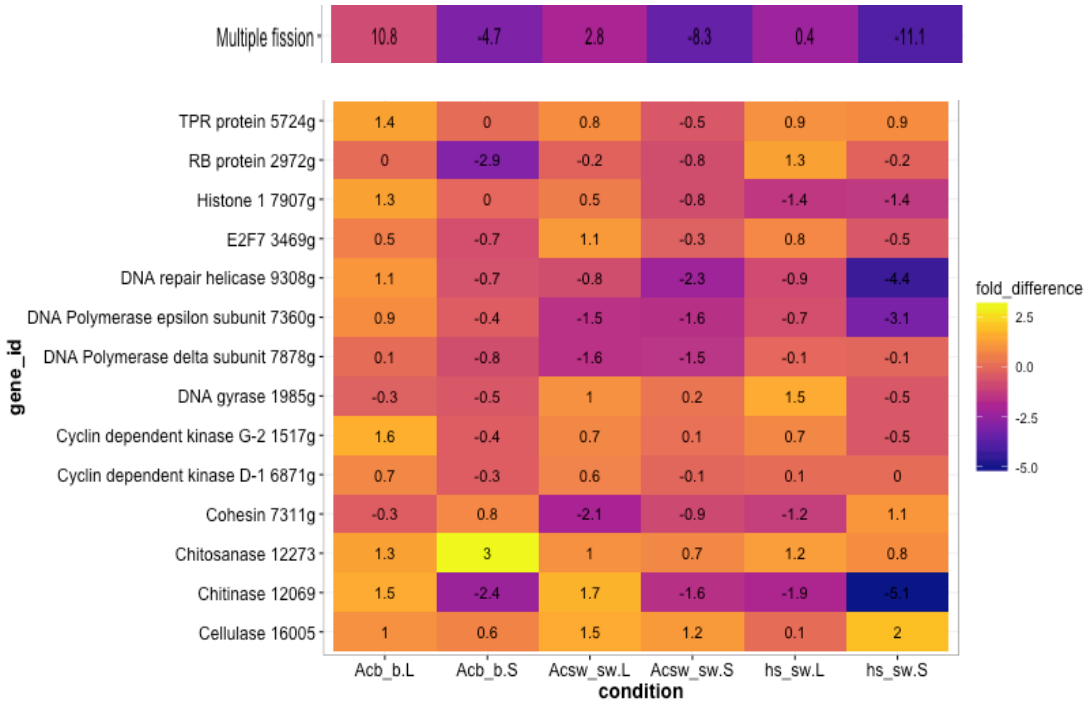

Figure 3-4. *Mk* KAS603 multiple fission gene expression heatmap. Acb_b represents acetate (50 mM) + bicarbonate (178 mM) compared to bicarbonate (178 mM) in MS4 freshwater medium. Acsw_sw represents acetate (50 mM) in ASW compared to ASW. hs_sw represents ASW (1.5 M NaCl) compared to ASW (0.4 M NaCl). L denotes log phase and S denotes stationary phase. The numbers inside each rectangle as well as the colors, represent gene expression fold difference per condition. On the "y" axis, the gene/enzyme names are followed by their gene ID number and by "g" if the sequence was obtained from *Mk* KAS603 genome assembly data, otherwise it corresponds to the transcriptome assembly data. All numbers are log 2 based. The upper part of the figure (denoted as multiple fission) represents the total fold change per condition, as depicted in Fig. 3-3.

Transcription factor/regulator genes possibly involved in multiple fission.

TPR gene.

Proteins that contain tetratricopeptide peptide repeat regions (TPR) are involved in cell cycle regulation and transcriptional control (Paddy, 1998). As displayed in Fig. 3-4, we identified a TPR protein that is upregulated during log phase in all experimental conditions, suggesting an involvement as a transcriptional regulator in multiple fission division. During stationary phase, this gene is only upregulated in hypersalinity, which might be due to an interesting phenomenon that happens only in this condition: in an entire population of cells adapted to hypersalinity, during stationary phase and high TAG accumulation, there is a lack of small daughters, only daughter cells above 30 μ^2 exist, while in lower salinities in the same stage cell size ranges from 20 to 40 μ^2 (Figs. 5 and 7 in chapter 1). This implies that even though growth has slowed down, cells are larger, which could be considered an attempt at growth. The TPR protein might be a regulator of cell size determination in *Mk* KAS603.

E2F7 and RB genes.

E2F can work together with RB protein, regulating the G1-S transition (Shen, 2002). Their upregulation could correspond to the growth and acquisition of biomass of daughter cells and the duration of DNA synthesis, in *Mk* KAS603. As displayed in Fig. 3-4, the E2F isoform 7 gene is upregulated during logarithmic phase and

downregulated in stationary phase during all the experimental conditions tested, which might favor biomass acquisition in preparation for cell division, with the aid of cyclin-dependent kinases D-1 and G2 (whose genes are upregulated under all conditions during log phase) that can phosphorylate RB. The RB gene is only upregulated during log phase in hypersalinity (and cyclin-dependent kinases D-1 and G2 are also upregulated in hypersalinity just not as much as in the other conditions), which displayed the slowest growth rate/biomass acquisition among all experimental conditions. Together E2F7, RB and cyclin-dependent kinases D-1 and G2 gene expression might delay biomass acquisition in *Mk* KAS603 when cultivated in hypersalinity.

Histone 1, DNA repair helicase and DNA polymerase genes' possible role in mitotic rounds.

H1 kinase (which modulates Histone 1 function) activity has been observed to increase at the end of the cell cycle during multiple rounds of DNA replication and mitosis in *C. reinhardtii* (Zachleder *et al*, 1997). As shown in Fig. 3-4, the histone 1 gene displayed the highest upregulation during log phase in Acb, which is the fastest cell growth-cell division condition. Also, cyclin dependent kinase G-2 and D1 genes, displayed the highest upregulation under the same condition, suggesting their involvement in fast growth and several mitotic rounds in multiple fission (in a condition where mother cells (with daughters undergoing mitosis) were predominant.

Genes for two DNA polymerase subunits as well as a DNA repair helicase

gene, are only upregulated during this condition, suggesting more DNA replication events as well as DNA repair need during the fastest cell growth rate condition.

Genes involved in cell wall rupture and daughter cell release.

As evidence for rupture of mother cells and release of daughter cells (correlated to cell size data), we present gene expression data of genes corresponding to enzymes that have been previously reported to be involved in cell wall metabolism in other Chlorella's (Blanc *et al*, 2010; Yamamoto *et al*, 2004).

The chitosanase gene is upregulated in all conditions and growth phases, which suggests a rupture of cells both in log and stationary phase. This gene is highly upregulated during stationary phase in Acb where cell size data indicates abundant (70%) mother cells (Fig. 3-2B) that can be going through a degradation process of their thick cell wall [mother cells have a thicker cell wall than captive daughter cells (Yamamoto et al, 2004), and thus need an intensive enzymatic activity for cell wall rupture and release of daughters]. Acb is the condition where the highest expected amount of rupture and release may be correlated with the highest extent of upregulation of the chitosanase gene. Also, pre-apoptotic daughter cells might be going through chitosan cell wall degradation.

The chitinase gene is only upregulated during log phase in the two mixotrophic condition (Acb and Acsw), which are the 2 fastest cell growth (and possibly the greatest number of cell wall ruptures) conditions. The cellulase gene is upregulated under all conditions (which suggests cell wall rupture of cells during log

and stationary phase) except during log phase in hs (neutral expression) where daughter cells might take longer to develop into mother cells. Cellulose might be a predominant component of the cell wall of *Mk* KAS603.

Cohesin during multiple mitotic rounds.

Cohesin is involved in chromosome condensation, and can mediate sister chromatid cohesion during mitosis (Parisi *et al*, 1999; Lee and Orr-Weaver, 2001) and possibly convey correct chromatid segregation during multiple mitotic rounds in *Mk* KAS603. As depicted in Fig. 3-4, the cohesin gene was upregulated in two conditions only: during stationary phase in hypersalinity and during stationary phase in Acb, in which more mitotic rounds might be needed. During stationary phase in Acb, larger mother cells (with more captive daughters) are predominant compared to the control (only bicarbonate) (Fig. 3-2B), this might correlate to more divisions. During hypersalinity, there is an abundance of mothers in a subpopulation immediately above 47 μ^2 (where daughters transition into mothers) (Fig. 7C in chapter 1), such occurrence might be correlated to an upregulation of the cohesion gene possibly due to more divisions in a small transitioning subpopulation of cells.

Conclusion

The gene expression data correlated well with the physiological data, indicating that the cultivation condition in which *Mk* KAS603 displayed the highest multiple fission gene expression (and possibly the condition with the highest

metabolic activity due to fastest growth), was acetate + bicarbonate, followed by acetate in seawater and hypersalinity. Cells can grow and become mother cells faster during mixotrophy in freshwater than in all other conditions, creating a need for a faster flux of carbon, energy flux and metabolic trade among pathways and cell compartments. Cell size and multiple fission cell division can be important processes for adaptation and survival of *Mk* KAS603 to different cultivation and trophic conditions.

Starch and glucose metabolism in *Mk* KAS603

Overview.

Starch synthesis and degradation were predicted to take place in the cytoplasm and chloroplast of *Mk* KAS603, while glycolysis and gluconeogenesis were strongly (full pathway) predicted in the cytoplasm and partially (upper half of glycolysis) in the plastid (Fig. 2-7 in chapter 2), indicating 2 possible compartments where starch could be accumulated. During log phase in acetate + bicarbonate glycolysis, gluconeogenesis, starch anabolism and catabolism genes were globally upregulated (Fig. 3-5). This might be due to a greater need for carbon skeletons and energy (due to an energy drain caused by transport and processing of bicarbonate) during accelerated growth in the presence (uptake) of acetate, where cells are not only growing and undergoing multiple fission, but also accumulating TAG. During stationary phase in Acb cultivation, glycolysis/gluconeogenesis genes remained globally upregulated and starch metabolism genes were downregulated, as a possible

indication of less energy and carbon skeletons needed during lack of growth.

In ASW supplemented with acetate, growth is enhanced but cells don't accumulate TAG while growing (Fig. 6 in chapter 1). Also, acetate metabolism is predicted to be less energy draining for the cell than bicarbonate. As displayed in Fig. 3-3, a global upregulation for glucose metabolism and starch anabolism genes occurred during Acsw log phase (possibly to maintain cell growth), while starch catabolism genes displayed a global downregulation (possibly to maintain starch accumulation). During Acsw stationary phase, starch metabolism genes were globally downregulated, possibly due to reduced starch accumulation, and glucose metabolism genes were upregulated (Fig. 3-3) possibly for pyruvate (and AcCoA) generation for FA synthesis during this TAG accumulation phase (Fig. 6 in chapter 1).

As displayed in Fig. 3-3, during log phase in hs, glucose metabolism genes were upregulated possibly to provide energy for growth and a precursor (DHAP) for glycerol (predicted osmolyte) synthesis; and starch metabolism genes were downregulated, possibly due to non-enhancement of starch accumulation. During stationary phase, glucose metabolism genes displayed a high upregulation possibly to promote starch accumulation (suggested by upregulation of starch synthesis genes) and to provide energy and pyruvate for AcCoA synthesis for FA-TAG accumulation. Starch might also be accumulated due to a suggested low activity of starch catabolism genes (that display downregulation). In hypersalinity, growth (and possibly starch) was not enhanced and cells didn't accumulate TAG while growing (Fig. 5 in chapter 1).

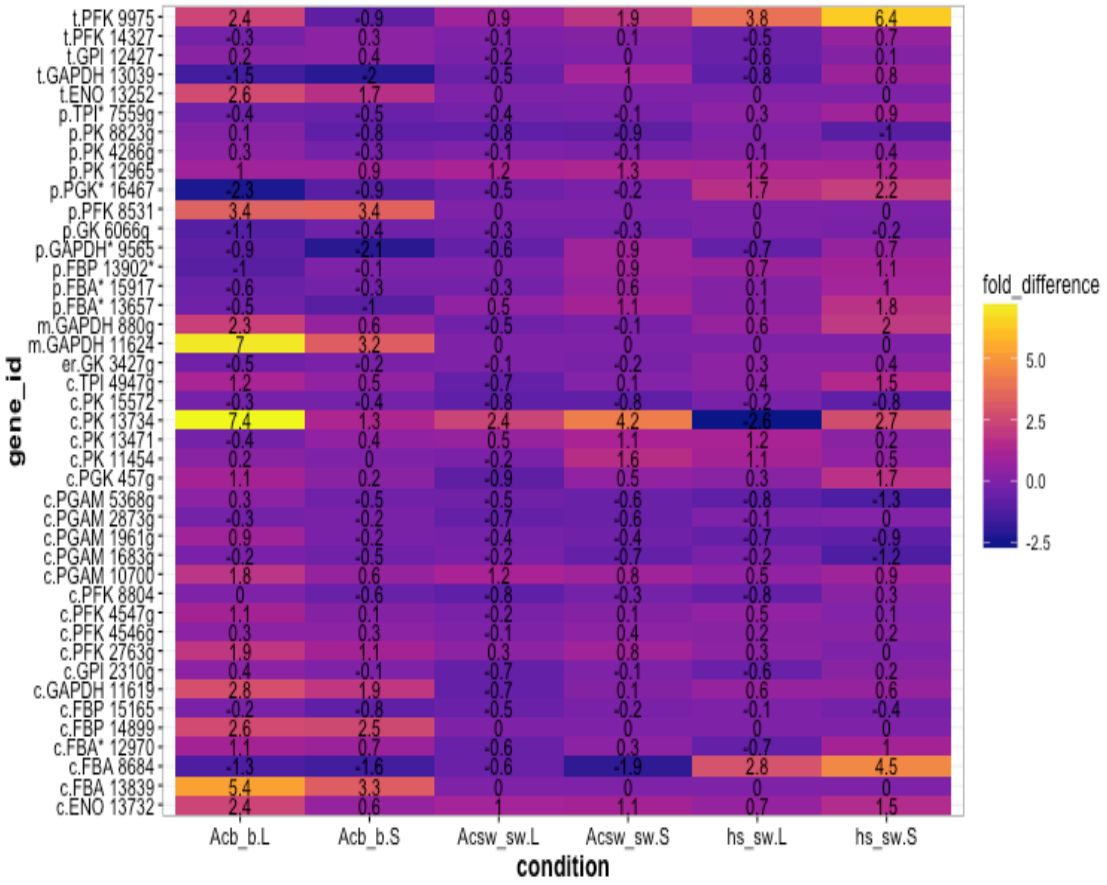

Figure 3-5. *Mk* KAS603 glycolysis and gluconeogenesis gene expression heatmaps. Acb_b represents acetate (50 mM) + bicarbonate (178 mM) compared to bicarbonate (178 mM) in MS4 freshwater medium. Acsw_sw represents acetate (50 mM) in ASW compared to ASW. hs_sw represents ASW (1.5 M NaCl) compared to ASW (0.4 M NaCl). L denotes log phase and S denotes stationary phase. The numbers inside each rectangle as well as the colors, represent gene expression fold difference. The upper half of the figure displays individual gene expression responses per condition. On the Y axis, "t" denotes a truncated non targeted sequence, "p" denotes plastid targeting, "m" denotes mitochondrial targeting, "er" denotes endoplasmic reticulum targeting, "c" denotes cytoplasmic targeting. The abbreviated gene/enzyme names are followed by their gene ID number and by "g" if the sequence was obtained from *Mk* KAS603 genome assembly data, otherwise it corresponds to the transcriptome assembly data. Asterisks denote genes/enzymes that also participate in the Calvin-Benson Cycle. The lower half of the figure indicates total expression per compartment (cytoplasm [c] or plastid [p]) as well as the overall cell gene expression (.Total). All numbers are log 2 based.

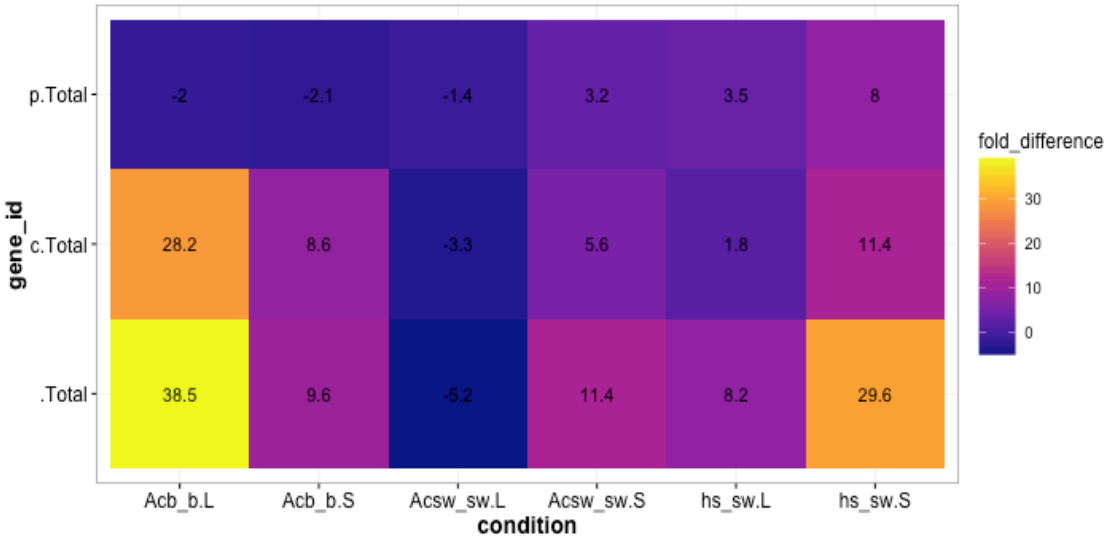

Figure 3-5. *Mk* KAS603 glycolysis and gluconeogenesis gene expression heatmaps. Continued.

PFK and FBP

As displayed in Fig. 3-5, in the acetate + bicarbonate condition, genes corresponding to enzymes (targeted to cytoplasm) involved in glucose metabolism displayed a global upregulation (as well as a specific upregulation of 2 PFK glycolysis genes and 1 FBP gluconeogenesis gene), while genes corresponding to enzymes targeted to plastid were globally downregulated (with the exception of a PFK gene, which is a key glycolytic gene). The former occurrence, suggests an activation of glycolysis in the plastid and activation of both glucose synthesis and degradation in the cytoplasm, for a possible avoidance of conflicting routes in the plastid and for (predicted) induced starch accumulation in the cytoplasm. ***The significant and coordinated changes in expression of the cytoplasmic glycolysis and gluconeogenesis genes supports the targeting analysis that the pathway exists in the cytoplasm, and that it plays an important role in carbon flux in Mk KAS603.*** As shown in Fig. 5, in the acetate supplemented ASW condition, genes encoding for both plastid and cytoplasmic targeted glucose metabolism enzymes, displayed a global downregulation during log phase possibly due to a direct utilization of acetate, not requiring the glycolytic pathway; and a global upregulation during stationary phase (with an upregulation of cytoplasmically targeted PFK and plastid targeted FBP encoding genes), possibly to activate glycolytic activity in the cytoplasm for energy and pyruvate production for AcCoA for FA after active growth. Since gluconeogenesis is only partially predicted in the plastid, a possible activation of this

route might provide carbon molecules for the Calvin Benson cycle and/or glucose-6-phosphate to be exported to the cytoplasm to be processed through glycolysis.

Enolase

The gene that codes for cytoplasmically targeted enolase, was upregulated in all conditions, suggesting an activation of PEP production for pyruvate and ATP production, and also for AcCoA for further energy generation. The highest upregulation was displayed in Acb during log phase to possibly direct carbon flow toward pyruvate generation (PEP/pyruvate/AcCoA) for FA synthesis (genes for FA synthesis are upregulated) for cell membrane formation (multiple fission genes are highly upregulated), as well as for for TAG (onset) synthesis (TAG synthesis genes are upregulated). In hs during stationary phase (which is the second highest upregulation) it might favor PEP/pyruvate/AcCoA for FA-TAG accumulation after growth (FA-TAG synthesis genes are upregulated and multiple fission genes are distinctively downregulated).

Starch metabolism

As displayed in Fig. 3-6, the majority of genes that code for plastid and cytoplasmically targeted starch anabolism enzymes were upregulated during log phase in Acb and Acsw, suggesting starch accumulation in both compartments during growth with an organic carbon source. The opposite occurred during stationary phase, where downregulation of most starch anabolism genes was displayed. During hs in log phase, the starch synthesis genes remained almost neutral, indicating no change

relative to the control, and in hs during stationary phase (where the majority of genes coding for enzymes targeted to cytoplasm and plastid were upregulated) starch is suggested to exist in 2 compartments. Starch synthesis during hypersalinity may be favored during stationary phase (along with glycolysis/gluconeogenesis) for synthesis of metabolic precursors (such as 3-phosphoglycerate) for serine and glycine synthesis (predicted osmolytes). Starch catabolism genes whose products were targeted to plastid and cytoplasm were mostly upregulated during log phase in Acb. In all other conditions they were either neutral or downregulated in majority. This provides a possible distinction for acetate + bicarbonate cultivated cells to possibly be able to accumulate and also utilize starch in the cytoplasm and plastid, during accelerated cell growth. Relative changes in starch catabolism and anabolism could lead to similar outcomes. For example, in hs during log phase catabolism genes are substantially downregulated, with no change for anabolism genes. In hs during stationary phase, anabolism genes were substantially upregulated, and catabolism genes slightly downregulated. The suggested net effect for both would be increased starch accumulation. The cell might have a metabolic advantage to manage starch accumulation by two different mechanisms where saving energy might be a key factor. During log phase cells are investing energy mostly for growth, and having a possible mechanism to shut down starch degradation is the most plausible was to promote starch accumulation. During stationary phase, cells can invest more energy toward activation of starch synthesis, while starch catabolism inhibition contributes to further starch accumulation.

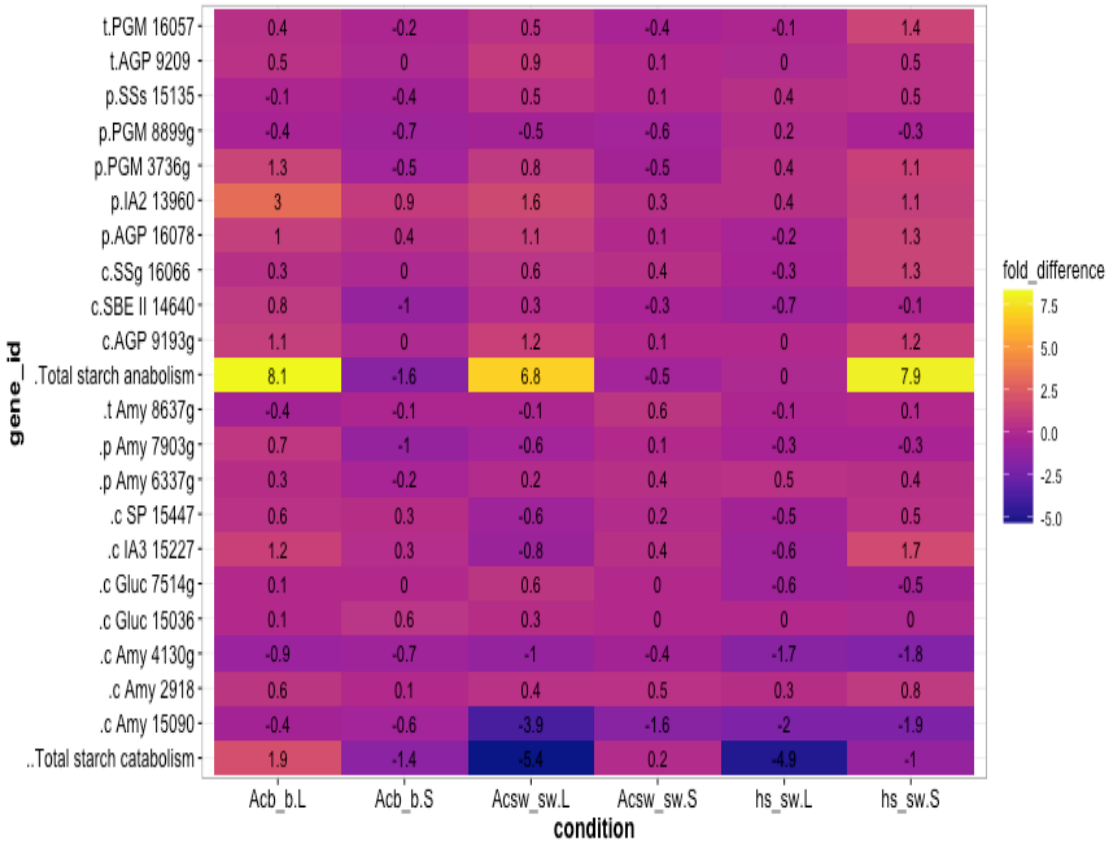

Figure 3-6. *Mk* KAS603 starch metabolism gene expression heatmap. Acb_b represents acetate (50 mM) + bicarbonate (178 mM) compared to bicarbonate (178 mM) in MS4 freshwater medium. Acsw_sw represents acetate (50 mM) in ASW compared to ASW. hs_sw represents ASW (1.5 M NaCl) compared to ASW (0.4 M NaCl). L denotes log phase and S denotes stationary phase. The numbers inside each rectangle as well as the colors, represent gene expression fold difference per condition. Individual as well as total gene expression changes for starch anabolism and catabolism are denoted. On the Y axis, "t" denotes a truncated non-targeted sequence, "p" denotes plastid targeting, "c" denotes cytoplasmic targeting. The abbreviated gene/enzyme names are followed by their gene ID number and by "g" if the sequence was obtained from *Mk* KAS603 genome assembly data, otherwise it corresponds to the transcriptome assembly data. All numbers are log 2 based.

Conclusion

Cytoplasmic predicted glycolysis and gluconeogenesis (as well as cytoplasmic and plastid predicted starch anabolism and catabolism) seem to be activated during acetate + bicarbonate supplementation. Glycolysis, gluconeogenesis and starch catabolism seem to be attenuated in the cytoplasm and plastid during growth with only acetate supplementation and seem to be induced during stationary phase (while starch accumulation seems to be active during growth and decreased when reaching stationary phase). This might be caused by a greater energy need for the cell during uptake of bicarbonate and acetate, compared to only acetate. Plastid and cytoplasmic predicted glycolysis, gluconeogenesis and starch synthesis seem to be enhanced during stationary phase in hypersalinity, when TAG accumulation is highest, suggesting a possible conversion from starch to lipid during stationary phase.

Fatty acid and TAG metabolism in *Mk* KAS6003

Overview.

FA synthesis was predicted to take place in the plastid and cytoplasm. FA desaturation was predicted to occur in the plastid, ER and cytoplasm, and TAG synthesis was predicted to the cytoplasm and ER. (Fig. 2-11 in chapter 2). TAG catabolism is proposed to take place in the cytoplasm and FA catabolism (beta oxidation) in the glyoxysome, in *Mk* KAS603 (Fig. 2-10 in chapter 2). FA-TAG metabolism seems to extend across several compartments in *Mk* KAS603, possibly due to a necessary movement of metabolites from one compartment to another as an

interactive exchange of precursors mechanism among routes, that facilitates global metabolism in the cell. As displayed in Fig. 3-3, genes involved in lipogenesis and lipolysis displayed a global upregulation during log and stationary phases in Acb, possibly due to an early onset of TAG accumulation occurring during active growth that continued throughout stationary phase (Fig. 3-1A-B) as well as an activation of FA-TAG degradation for energy generation needed during and after accelerated growth. In the acetate vs. no acetate condition, a global increase in the expression of lipogenesis genes was displayed, although this increase was 4-fold less during log phase than in stationary phase. Such a pattern might suggest an active synthesis of FA for cell membrane formation (daughter cells) during accelerated growth and an active synthesis of FA and TAG for lipid droplet formation during stationary phase (cells experience TAG accumulation during stationary phase [Fig. 6A-B, in chapter 1] in Acsw). This is a consistent pattern with acetate + bicarbonate, which strongly suggests that acetate boosts FA and TAG synthesis. The lipolysis genes displayed a weak global upregulation in Acsw during stationary phase, suggesting activation of lipid catabolism to fulfill a minimized energy need after accelerated growth. Since starch anabolism genes are upregulated during log phase and downregulated during stationary (and lipolysis genes display upregulation during stationary), the cells seem to shift from using starch during growth to using lipid during stationary phase. In hypersalinity, a global downregulation of lipogenesis genes occurred during log phase (possibly due to a preferential use of starch for energy), and an upregulation during stationary phase; possibly to indicate a non-induced TAG accumulation during log

phase and an active TAG accumulation during stationary phase (Fig. 5 A-B, chapter 1). A significant total downregulation of lipolysis genes was displayed during stationary phase (TAG accumulation phase) in hypersalinity. This might suggest that in hs stationary phase, TAGs and FA are not being catabolized in order to accumulate and serve as energy reserves . The hs stationary condition decreases lipolysis gene expression more and increases lipogenesis gene expression less than occurs under the other conditions in stationary phase. This might be a saving energy mechanism for the cell, where it shuts down catabolism to favor anabolism instead activating an (high energy) anabolic route alone.

Acetyl CoA formation in the plastid

As displayed in Fig. 3-7, plastid targeted PK and PDH encoding genes were upregulated during active growth (log phase) in Acb, which might indicate an active pyruvate and AcCoA synthesis in the plastid directed toward FA production (possibly for both cell membrane formation and TAG accumulation). During stationary phase, PK was upregulated (and PDH was slightly downregulated), possibly due to an activation of pyruvate synthesis to sustain a constant PDH activity for AcCoA production for FA-TAG accumulation, as it seems to also occur in acetate supplemented ASW and hypersalinity, where only the PK gene was upregulated and not the PDH gene.

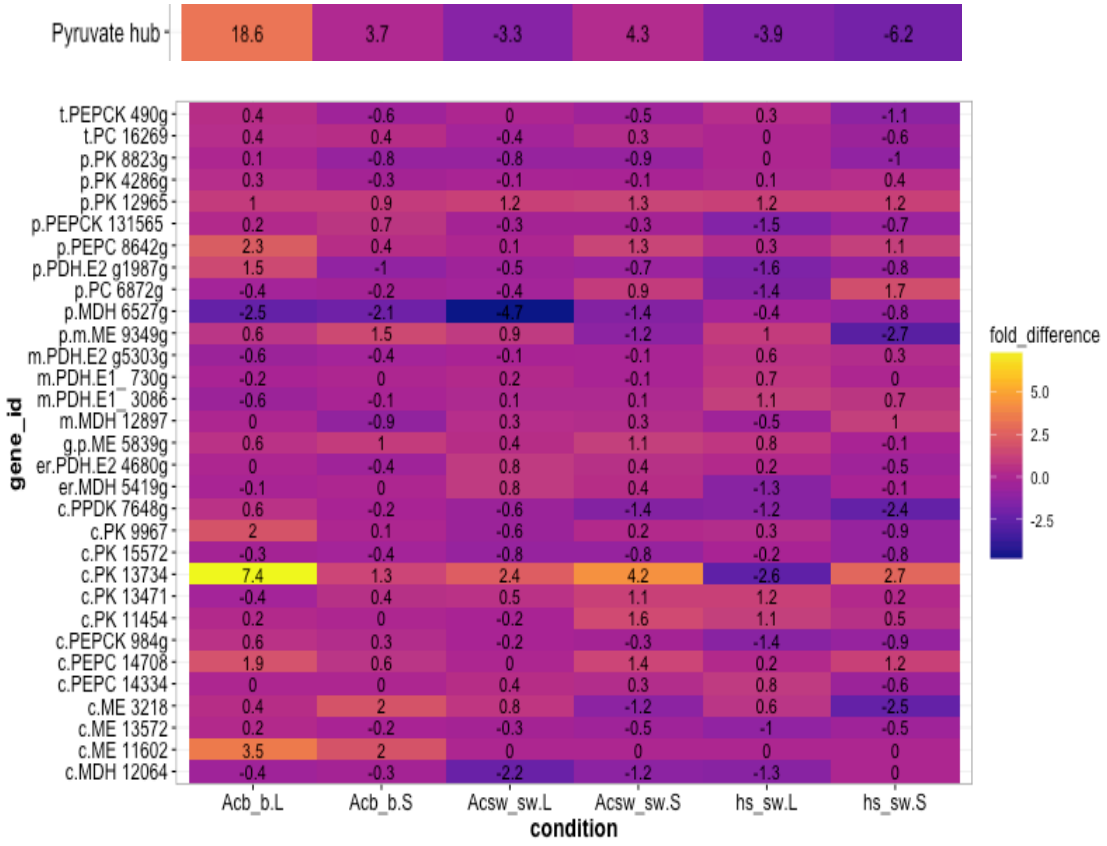

Figure 3-7. *Mk* KAS603 pyruvate hub gene expression heatmap. Acb_b represents acetate (50 mM) + bicarbonate (178 mM) compared to bicarbonate (178 mM) in MS4 freshwater medium. Acsw_sw represents acetate (50 mM) in ASW compared to ASW. hs_sw represents ASW (1.5 M NaCl) compared to ASW (0.4 M NaCl). L denotes log phase and S denotes stationary phase. The numbers inside each rectangle as well as the colors, represent gene expression fold difference per condition. On the Y axis, "t" denotes a truncated non-targeted protein sequence, "p" denotes plastid targeting, "er" denotes endoplasmic reticulum targeting, "c" denotes cytoplasmic targeting, and "m" denotes mitochondrial targeting. Some enzymes were dually targeted. The abbreviated gene/enzyme names are followed by their gene ID number and by "g" if the sequence was obtained from *Mk* KAS603 genome assembly data, otherwise it corresponds to the transcriptome assembly data. All numbers are log 2 based. The upper part of the figure (denoted as Pyruvate hub) represents the total fold change per condition, as depicted in Fig. 3-3.

FA and TAG anabolism genes

Plastid

Acetyl CoA carboxylase (accA) gene

accA catalyzes the formation of malonyl-CoA from acetyl CoA, which is an irreversible process and a key step for FA synthesis. As displayed in Fig. 3-8, a gene that codes for plastid Acetyl CoA carboxylase was upregulated during stationary phase (TAG accumulation) in Acsw and hs; suggesting an activation of FA synthesis for TAG during those conditions. During stationary phase in Acb the accA gene was neutral, possibly due to minimal TAG accumulation in the control (Fig. 3-1B)

KAR and HD genes

Genes that code for plastid targeted KAR and HD were distinctively upregulated in Acb during log and stationary phase. Such upregulation only occurred during this condition, which might suggest an activation of specific consecutive reduction and dehydration steps during FA synthesis.

ER

All ER targeted FA and TAG synthesis corresponding genes were upregulated during log phase in Acb, suggesting an active lipid anabolism activity in this compartment and supporting our targeting prediction analysis.

Cytoplasm

Cytoplasmically targeted MCAT, KASII and PAP corresponding genes were upregulated during log phase in Acb, suggesting a possible activation of predicted FA and TAG synthesis in the cytoplasm.

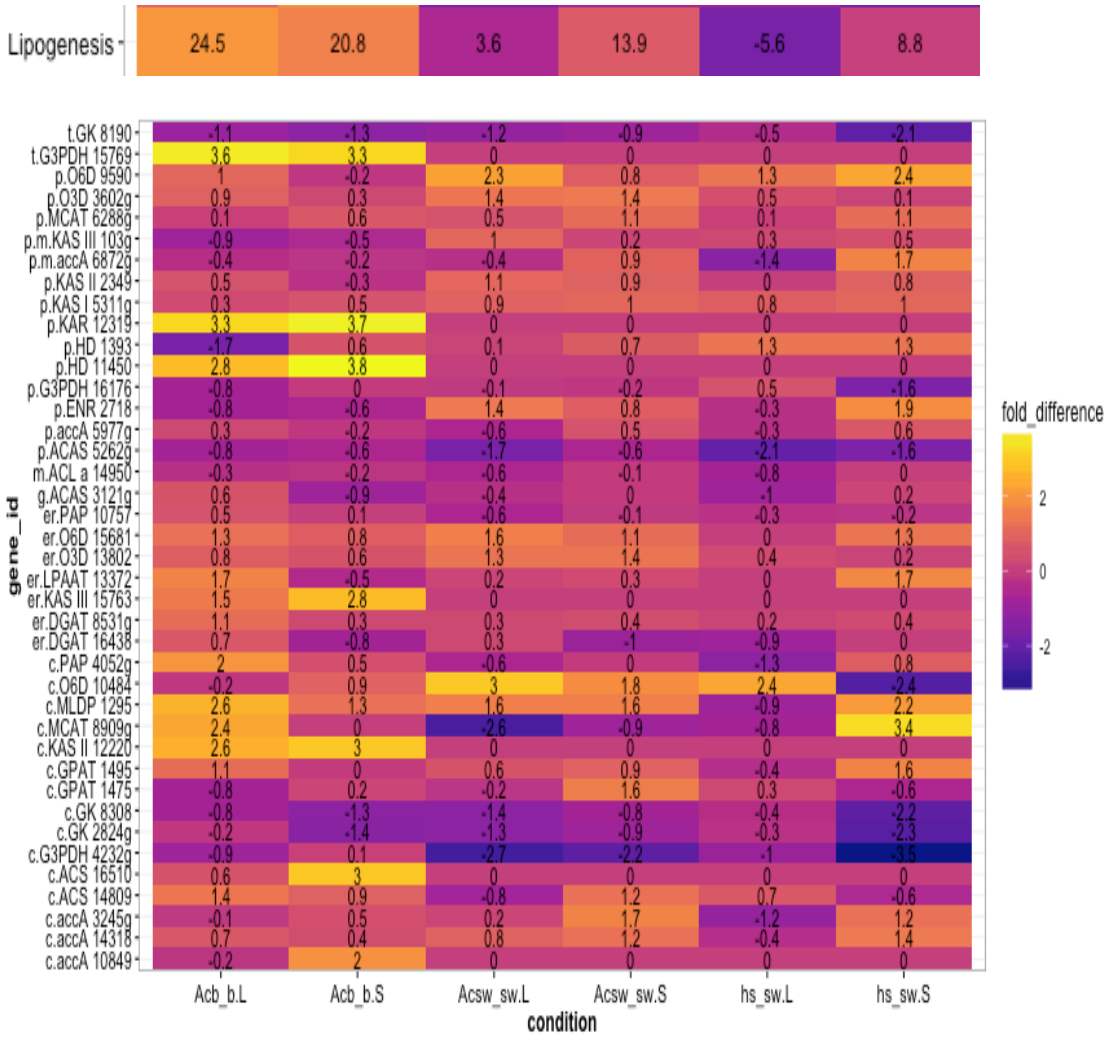

Figure 3-8. *Mk* KAS603 lipogenesis gene expression heatmap. Acb_b represents acetate (50 mM) + bicarbonate (178 mM) compared to bicarbonate (178 mM) in MS4 freshwater medium. Acsw_sw represents acetate (50 mM) in ASW compared to ASW. hs_sw represents ASW (1.5 M NaCl) compared to ASW (0.4 M NaCl). L denotes log phase and S denotes stationary phase. The numbers inside each rectangle as well as the colors, represent gene expression fold difference per condition. On the Y axis, "t" denotes a truncated non-targeted protein sequence, "p" denotes plastid targeting, "er" denotes endoplasmic reticulum targeting, "c" denotes cytoplasmic targeting, and "m" denotes mitochondrial targeting. Some enzymes were dually targeted. The abbreviated gene/enzyme names are followed by their gene ID number and by "g" if the sequence was obtained from *Mk* KAS603 genome assembly data, otherwise it corresponds to the transcriptome assembly data. All numbers are log 2 based. The upper part of the figure (denoted as lipogenesis) represents the total fold change per condition, as depicted in Fig. 3-3.

FA and TAG catabolism genes

As displayed in Fig. 3-9, TAG lipase genes as well as 10 FA beta oxidation genes were upregulated during log and stationary phases in Acb cultured cells. Specifically, genes encoding ACAT and HADH (which catalyze final consecutive steps of beta oxidation, that lead to NADH and AcCoA production), displayed a strong upregulation; which suggests a possible activation of FA oxidation for AcCoA and energy generation. TAG lipase genes and most of the FA genes were downregulated during log phase in Acsw; suggesting a possible low TAG catabolism due to the lack of TAG accumulation during cell growth (Fig. 6 in chapter 1), and a possible beta oxidation repression to maintain a higher concentration of FA, enough to support cell growth (and FA for cell membrane formation). During stationary phase in Acsw, TAG lipase genes were upregulated, possibly due to activation of TAG catabolism for FA release, and 3 FA catabolism genes were upregulated, possibly due to a derepression of beta oxidation. During log phase in hypersalinity most FA-TAG catabolism genes displayed a neutral expression, suggesting a possible low activity and minimal energy generation for the cell through this route. During stationary phase in hypersalinity the majority of FA-TAG catabolism genes were downregulated, specifically genes encoding ACOX and ACAT (that catalyze the first and last steps of FA beta oxidation). Since hypersaline cultivated cells accumulate TAG during stationary phase (Fig. 5 in chapter 1), a possible repression of FA degradation might be favoring FA-TAG accumulation in this condition.

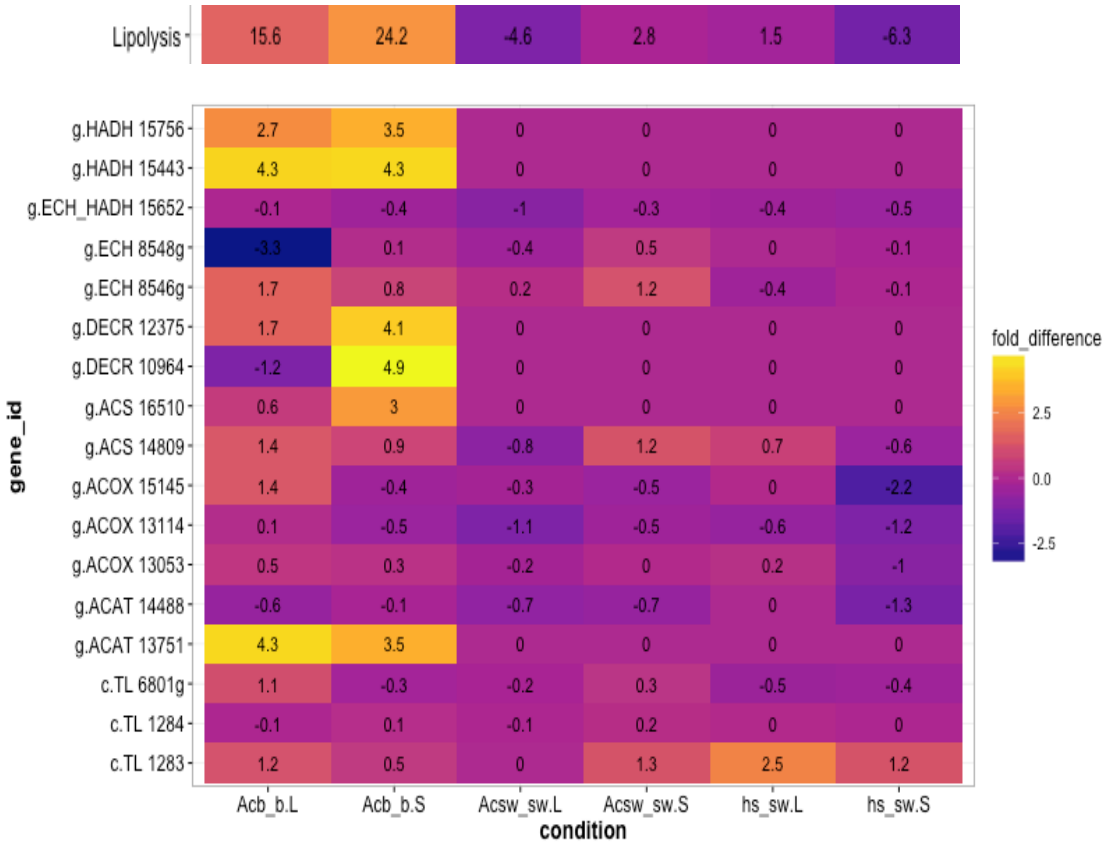

Figure 3-9. Lipolysis gene expression heatmap. Acb_b represents acetate (50 mM) + bicarbonate (178 mM) compared to bicarbonate (178 mM) in MS4 freshwater medium. Acsw_sw represents acetate (50 mM) in ASW compared to ASW. hs_sw represents ASW (1.5 M NaCl) compared to ASW (0.4 M NaCl). L denotes log phase and S denotes stationary phase. The numbers inside each rectangle as well as the colors, represent gene expression fold difference per condition. On the Y axis, "g" denotes glyoxysomal targeting and "c" denotes cytoplasmic targeting. The abbreviated gene/enzyme names are followed by their gene ID number and by "g" if the sequence was obtained from *Mk* KAS603 genome assembly data, otherwise it corresponds to the transcriptome assembly data. All numbers are log 2 based. The upper part of the figure (denoted as Lipolysis) represents the total fold change per condition. As depicted in Fig. 3-3.

Conclusion

Cooperation between intracellular compartments might facilitate lipid metabolism in *Mk* KAS603. During the fastest TAG accumulating condition (acetate + bicarbonate), the lipogenesis and lipolysis gene sets displayed the highest total global expression. Under this condition, acetate can be converted to acetyl CoA, malate and/or pyruvate, favoring FA accumulation in plastid/cytoplasm and TAG accumulation in cytoplasm/ER. The cells may be storing more FA-TAG, but also needing to use it as a carbon/energy source for metabolism or daughter cell formation (new membranes needed), cells are also proposed to maintain FA-TAG catabolism (in cytoplasm and glyoxysome) active to fulfill an energy drain in the cells, due to bicarbonate (transport and processing), accelerated growth and simultaneous accumulation of TAG (during log phase). Cells cultivated in medium acetate supplemented ASW, are suggested to only maintain FA-TAG catabolism active after growth, since only at this stage the cells have available TAG to degrade; the lipogenesis gene set displayed a significant upregulation, in accordance to lipid induction . Also, in Acsw (log phase) and hs (stationary phase), lipogenesis is upregulated but lipolysis is downregulated, which would have the same net effect of TAG accumulation as in Acb. With this approach, the cell might be saving energy by slowing down a (lipid) catabolic route to favor (lipid) anabolism. In hs, a total shutdown of lipid catabolism is suggested due to a highly downregulated beta oxidation to maintain high FA-TAG with a very low energy investment. *Mk* KAS603 seems to maintain homeostasis and TAG storage by energy saving processes.

Glyoxylate cycle

Overview

The glyoxylate cycle is a short and simple (yet essential) cycle that converts two-carbon acetyl units into four-carbon units (succinate) for energy production and biosynthesis. In *Mk* KAS603 it is predicted to take place in the glyoxysome and cytoplasm (Fig. 2-9 in chapter 2). As displayed Figs. 3-3 and 3-10, the glyoxylate cycle is only upregulated in Acb (during log and stationary phases), possibly caused by a high demand of energy due to fast growth while processing bicarbonate and taking up acetate (into the glyosyome for AcCoA conversion by ACAS). Also, FA synthesis takes place during cell growth in Acb and continues throughout stationary phase, which possibly generates an even greater energy need. The glyoxylate cycle can also play a role in generating reducing equivalents. During stationary phase in Acb, genes encoding fused IL and MS and AH, display a distinctive upregulation (Fig. 3-10), suggesting isocitrate, glyoxylate and malate as primary products from the cycle during this condition. These products can be imported to other organelles and incorporated to other routes such as TCA cycle, photorespiration and FA synthesis. In all the other conditions, those same genes display a neutral activity (Fig. 3-10).

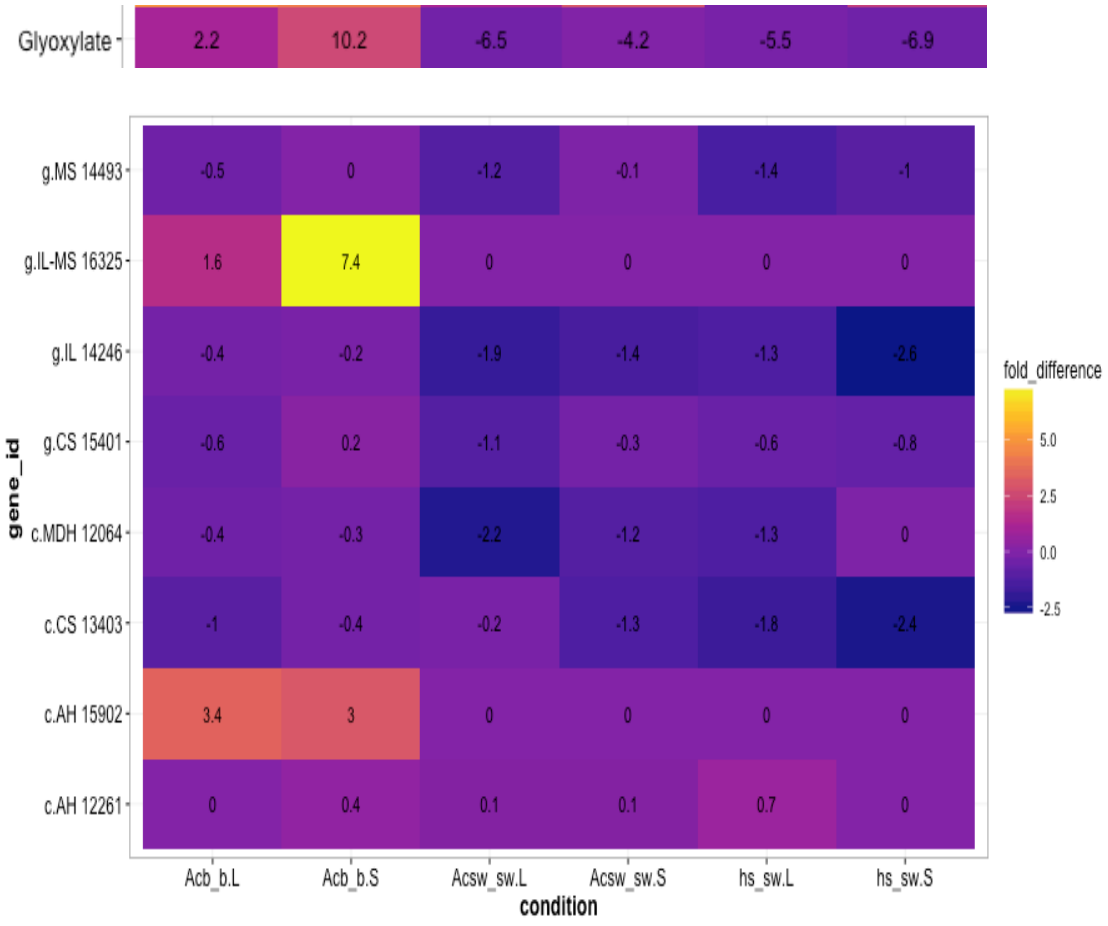

Figure 3-10. Glyoxylate cycle gene expression heatmap. Acb_b represents acetate (50 mM) + bicarbonate (178 mM) compared to bicarbonate (178 mM) in MS4 freshwater medium. Acsw_sw represents acetate (50 mM) in ASW compared to ASW. hs_sw represents ASW (1.5 M NaCl) compared to ASW (0.4 M NaCl). L denotes log phase and S denotes stationary phase. The numbers inside each rectangle as well as the colors, represent gene expression fold difference per condition. On the Y axis, "g" denotes glyoxysomal targeting and "c" denotes cytoplasmic targeting. The abbreviated gene/enzyme names are followed by their gene ID number in the transcriptome assembly data. All numbers are log 2 based. The upper part of the figure represents the total fold change per condition, as depicted in Fig. 3-3.

ACAS

Acetate can be processed through the glyoxylate cycle. It is possible that acetate entering the cell might be imported into the glyoxysome and converted to AcCoA by ACAS. The gene encoding glyoxylate targeted ACAS was upregulated during log phase in Acb (as well as IL-MS and AH genes), suggesting activation of ACAS for AcCoA production in the glyoxysome for NADH, isocitrate, succinate and malate for growth and FA synthesis.

Glyoxylate can contribute to osmolyte production

Since glycine and serine can serve as potential osmolytes for hypersaline cultured *Mk* KAS603 and the photorespiration pathway (which manages conversion of glyoxylate to glycine and serine) is only upregulated during hypersalinity with specific upregulation for GDC (that converts glycine to a serine precursor) and SGAT (that converts serine to glycine) genes (as displayed in Fig. 3-3 and 3-12); we propose that glyoxylate produced during the glyoxylate cycle, can also be incorporated into photorespiration or directly to glycine/serine production.

Conclusion

The glyoxylate cycle can provide carbon skeletons and reducing equivalents for multiple fission division, as well as substrates for other routes such as photorespiration, gluconeogenesis, TCA cycle and FA-TAG synthesis. This cycle

might be favoring biomass and TAG accumulation by providing NADH, isocitrate, glyoxylate, succinate and malate in *Mk* KAS603 cultivated in acetate + bicarbonate in freshwater and possibly osmolyte production (in conjunction with photorespiration) in hypersalinity. The cycle seems to be downregulated under all conditions (that might use the TCA cycle for reducing equivalent production) other than with Acb, which implies that acetate + bicarbonate play a key role in activation of this cycle that processes acetate as AcCoA and may sustain an energy drain created by bicarbonate transport and processing in *Mk* KAS603.

TCA cycle

Overview

It was not be expected that the TCA cycle genes would be regulated primarily at the transcript level, but we did identify significant changes in particular transcripts, that would be significant, such as IDH gene upregulation in Acsw (for NADH generation both in log and stationary phases) and CS and AH (that initiate the cycle) gene upregulation in log phase hs and MDH upregulation during stationary phase in hs (to possibly contribute with NADH or malate for export to the plastid for FA synthesis). As displayed in Figs. 3-3 and 3-11, the TCA cycle genes were globally downregulated in all conditions (except for a slight upregulation in hs during log phase), displaying a possible constitutive expression or an attenuation of the cycle due to ATP generation outside the mitochondria, such as plastidial ATP production (generated due to constant light exposure in all conditions.

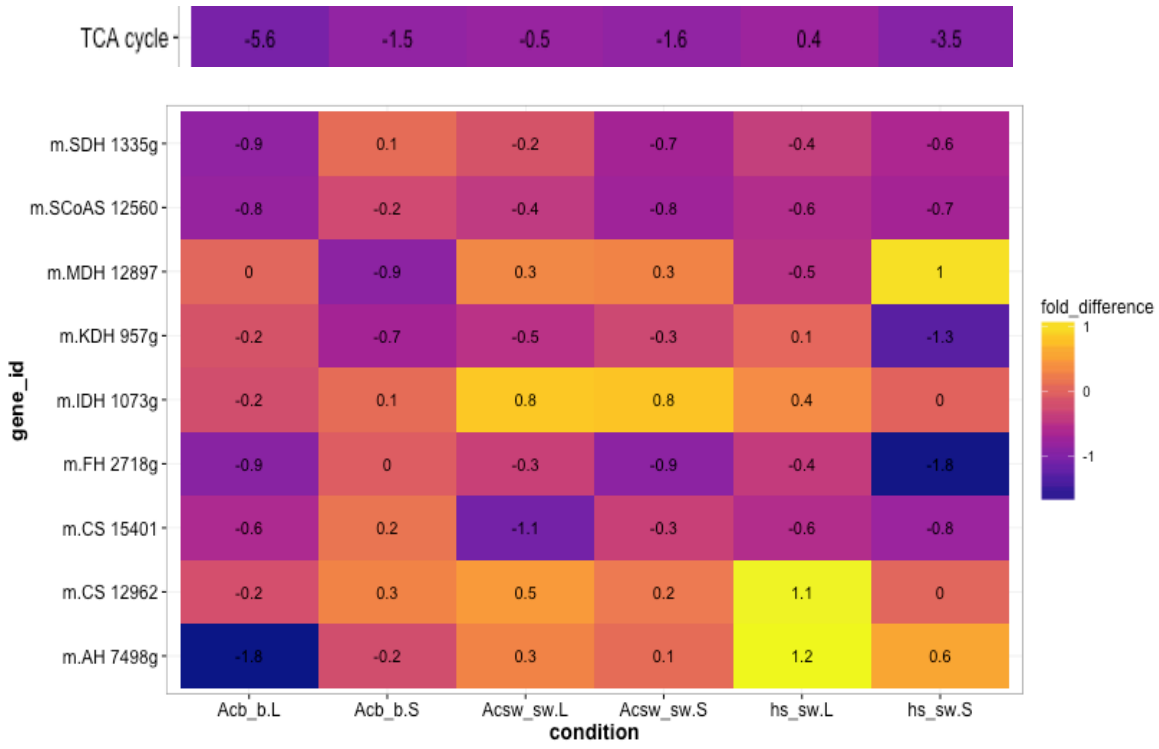

Figure 3-11. TCA cycle gene expression heatmap. Acb_b represents acetate (50 mM) + bicarbonate (178 mM) compared to bicarbonate (178 mM) in MS4 freshwater medium. Acsw_sw represents acetate (50 mM) in ASW compared to ASW. hs_sw represents ASW (1.5 M NaCl) compared to ASW (0.4 M NaCl). L denotes log phase and S denotes stationary phase. The numbers inside each rectangle as well as the colors, represent gene expression fold difference per condition. On the Y axis, "m" denotes mitochondrial targeting. The abbreviated gene/enzyme names are followed by their gene ID number and by "g" if the sequence was obtained from *Mk* KAS603 genome assembly data, otherwise it corresponds to the transcriptome assembly data. All numbers are log 2 based. The upper part of the figure (denoted as TCA cycle) represents the total fold change per condition, as depicted in Fig. 3-3.

CS and AH

CS and AH catalyze subsequent initial reactions in the TCA cycle. To begin a turn of the cycle, AcCoA donates its acetyl group to oxaloacetate to form citrate (catalyzed by CS). Citrate is transformed to isocitrate by AH. In hypersalinity, CS and AH genes were both upregulated during log phase (Fig. 3-11), opening the possibility of those two enzymes initiating and inducing the cycle during this condition.

IDH

Isocitrate dehydrogenase is a TCA irreversible enzyme that catalyzes oxidative decarboxylation of isocitrate to form alpha-ketoglutarate with NADH production. As displayed in Fig. 3-11, in Acsw during log and stationary phases, the IDH gene is upregulated (equally for both phases of growth). This occurrence may imply a need for NADH during and after growth for cells cultivated with acetate, possibly due to glyoxylate cycle downregulation which also provides NADH (specifically a significative cytoplasmically targeted MDH (which oxidizes malate to oxaloacetate with NADH production) gene downregulation that is predicted to be part of the glyoxylate cycle) (Fig. 3-10) in Acsw,

MDH

During stationary phase in hypersalinity, the MDH targeted to mitochondria encoding gene is upregulated (Fig. 3-11), possibly inducing oxaloacetate production

for gluconeogenesis and starch accumulation. This is supported by gene expression upregulation of plastid predicted gluconeogenesis (Fig. 3-5) and starch synthesis (Fig. 3-6) genes during this condition.

Conclusion

The TCA cycle seemed to be substantially downregulated with addition of acetate (Acb), since acetate (as AcCoA) seems to be preferentially processed through the glyoxylate cycle (upregulation) for energy, carbon skeletons, reducing equivalents and malate for growth and FA synthesis. Although in Acsw the glyoxylate cycle genes were downregulated and TCA cycle seemed to provide NADH (specifically through IDH upregulation). In hs during cell growth the TCA cycle seemed to be upregulated (specifically induced by two consecutive enzymes that manage AcCoA and the first condensation and dehydration steps in the cycle). In hs after cell growth, MDH seemed to be significantly upregulated to provide NADH and a precursor for starch synthesis.

Photorespiration

Overview

Photorespiration is a high-energy costly respiratory pathway that consumes oxygen and produces CO_2, and is driven by light. This pathway is proposed to place in 3 compartments in *Mk* KAS603: chloroplast-glyoxysome-mitochondrion (see Fig. 2-12 in chapter 2). In the plastid, RUBISCO can fix O_2 (instead of CO_2) as substrate,

forming 2-phosphoglycolate, which is dephosphorylated and turned into glycolate. In the glyoxysome, glycolate is predicted to be converted to glyoxylate, and glycine. In the mitochondria glycine is converted to serine, hydroxypyruvate and glycerate and finally in the chloroplast, glycerate is converted to 3-phosphoglycerate. As depicted in Figs. 3-3 and 3-12, a global upregulation of photorespiration genes is only displayed during stationary phase in hypersalinity, suggesting that photorespiration might be favoring the production of glycine/serine to be used as osmolytes. Glycine can also serve as a precursor for AcCoA and subsequently favor lipogenesis (which might also be enhanced by a suggested shut down of lipolysis), during stationary phase in hypersalinity. In Acb and Acsw, the downregulation might be due to acetate presence and less dependence of photosynthesis for carbon. In all conditions, photorespiration is downregulated more during growth than in stationary phase. More photo-stress would be expected when cells aren't dividing during stationary.

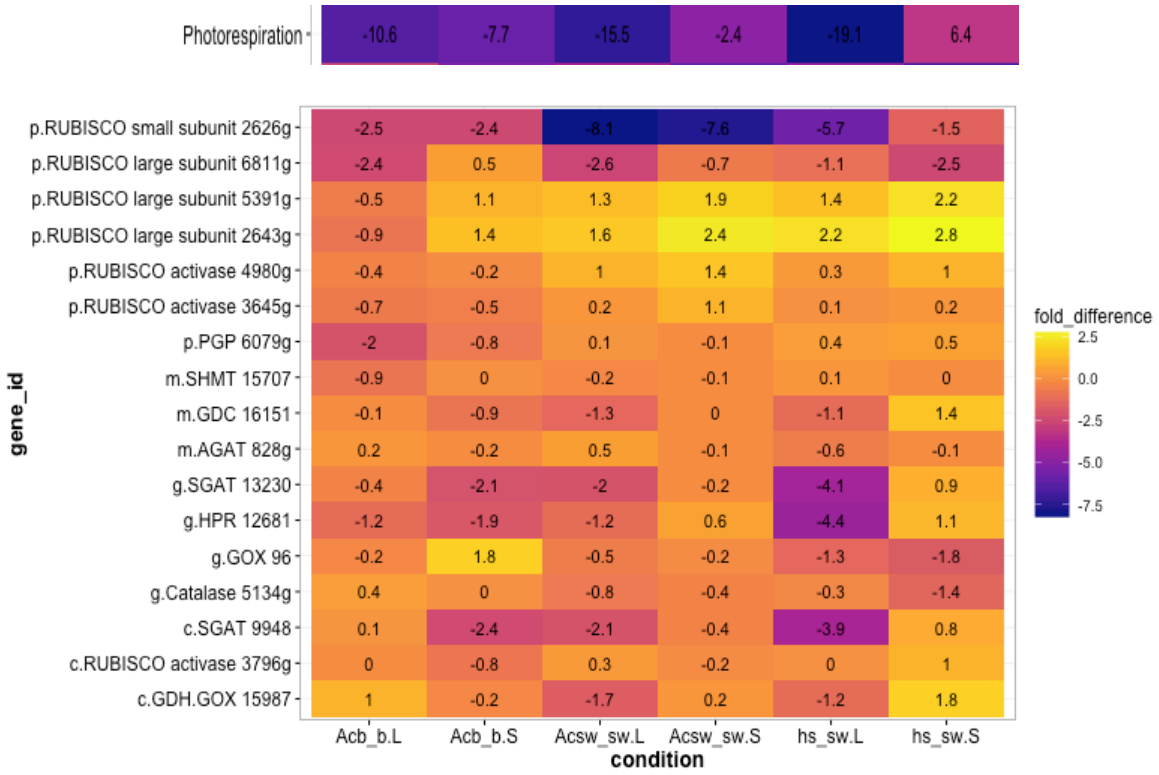

Figure 3-12. Photorespiration gene expression heatmap. Acb_b represents acetate (50 mM) + bicarbonate (178 mM) compared to bicarbonate (178 mM) in MS4 freshwater medium. Acsw_sw represents acetate (50 mM) in ASW compared to ASW. hs_sw represents ASW (1.5 M NaCl) compared to ASW (0.4 M NaCl). L denotes log phase and S denotes stationary phase. The numbers inside each rectangle as well as the colors, represent gene expression fold difference per condition. On the Y axis, "p" denotes plastid targeting, "g" denotes glyoxysomal targeting, "c" denotes cytoplasmic targeting, and "m" denotes mitochondrial targeting. The abbreviated gene/enzyme names are followed by their gene ID number and by "g" if the sequence was obtained from *Mk* KAS603 genome assembly data, otherwise it corresponds to the transcriptome assembly data. All numbers are log 2 based. The upper part of the figure (denoted as photorespiration) represents the total fold change per condition, as depicted in Fig. 3-3.

SGAT

As shown in Fig. 3-12, SGAT genes are upregulated during stationary phase in hypersalinity. SGAT transfers the amino group of serine to glyoxylate and forms glycine, which might serve as an osmolyte during this condition.

GDC

As depicted in Fig. 3-12, the GDC gene is upregulated in hypersalinity during stationary phase. GDC converts glycine to serine. As glycine, serine might also serve as an osmolyte.

Conclusion

Photorespiration seemed to be particularly high in hypersalinity where the cell may use this route to make compounds that are useful to the cell, such as osmolytes (glycine/serine), and FA synthesis in plastid (as glycine might also be exported from glyoxysome to chloroplast to be converted into AcCoA with subsequent FA-TAG formation). Photorespiration may be downregulated with acetate (Acb and Acsw) since photosynthesis can be less active during mixotrophy.

Pentose phosphate pathway

Overview

PPP is a catabolic pathway where glucose-6-phosphate is converted to pentose

phosphates and produces NADPH for biosynthetic reactions. We are proposing that PPP takes place mainly in the plastid (70% enzymes were targeted to the plastid, Fig. 2-13 in chapter 2) in *Mk* KAS603, where the primary function is to provide NADPH that can be used in FA synthesis in this compartment, enhancing lipid production. As shown in Figs. 3-3 and 3-13, the PPP gene set displays the highest upregulation during stationary phase in Acb when TAG displayed high accumulation (Fig. 3-1B), followed by log in hypersalinity when there is very low TAG accumulation (Fig. 6B in chapter 1). The former condition can use NADPH from the oxidative phase of PPP to facilitate and increase lipid synthesis, and the latter condition can utilize products from the non-oxidative phase like pentose ribose 5-phosphate to be converted to 3-phosphoglycerate, and DHAP for glycerol synthesis (proposed osmolyte for *Mk* KAS603 in hypersalinity). During log phase in Acsw a global PPP gene downregulation is displayed, since TAGs are not induced (Figs. 3-3 and 3-13).

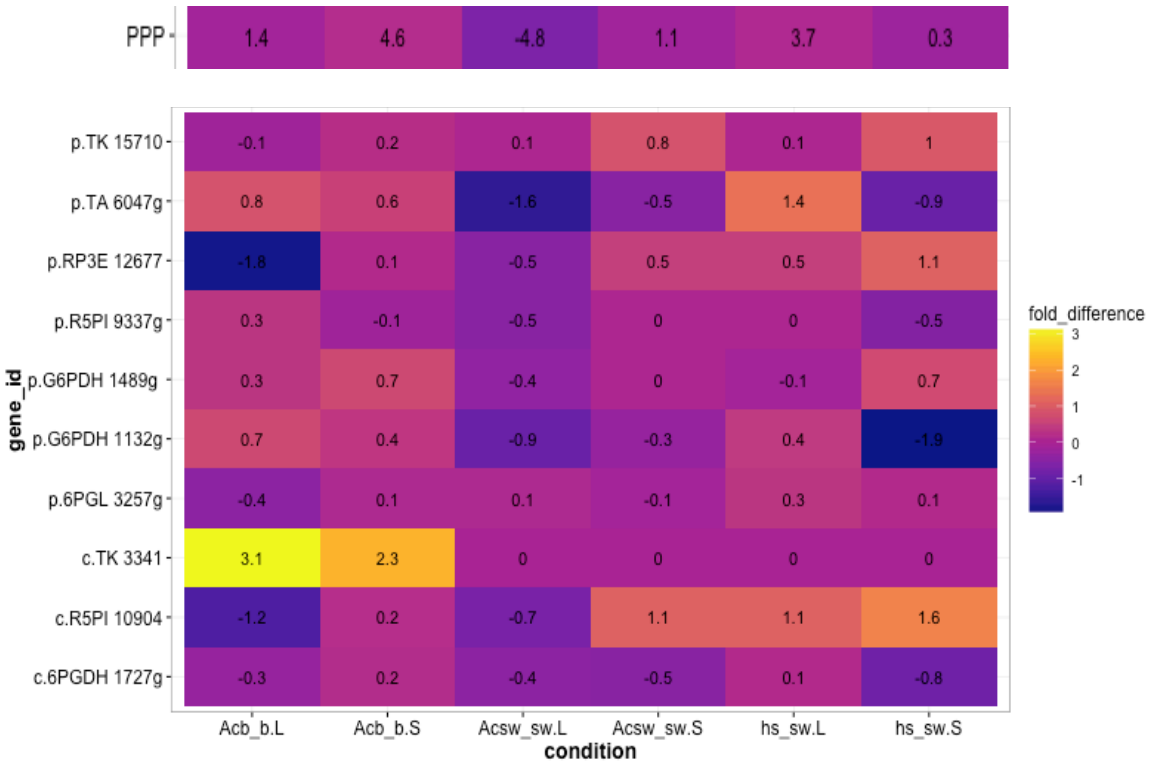

Figure 3-13. Pentose phosphate pathway gene expression heatmap. Acb_b represents acetate (50 mM) + bicarbonate (178 mM) compared to bicarbonate (178 mM) in MS4 freshwater medium. Acsw_sw represents acetate (50 mM) in ASW compared to ASW. hs_sw represents ASW (1.5 M NaCl) compared to ASW (0.4 M NaCl). L denotes log phase and S denotes stationary phase. The numbers inside each rectangle as well as the colors, represent gene expression fold difference per condition. On the Y axis, "p" denotes plastid targeting and "c" denotes cytoplasmic targeting. The abbreviated gene/enzyme names are followed by their gene ID number and by "g" if the sequence was obtained from *Mk* KAS603 genome assembly data, otherwise it corresponds to the transcriptome assembly data. All numbers are log 2 based. The upper part of the figure (denoted as PPP) represents the total fold change per condition, as depicted in Fig. 3-3.

Transketolase (TK) and glucose metabolism

As displayed in Fig. 3-13, in Acb, a gene encoding cytoplasmically targeted TK was distinctively upregulated, suggesting an activation of glyceraldehyde 3-phosphate synthesis possibly for glucose and starch accumulation in the cytoplasm (especially during log phase where TK gene upregulation was highest).

R5PI predicted dual role in growth and osmolyte synthesis

A gene encoding cytoplasmically targeted R5PI was upregulated in Acsw stationary phase and in hs log and stationary phases. Such upregulation may indicate activation of ribose 5-phosphate synthesis for nucleic acids or energy (ATP), but also for fructose 6-phosphate, glucose and starch synthesis; and for glyceraldehyde 3-phosphate also for starch or glycerol synthesis. Since during log phase in hs the glycerol genes displayed upregulation (Fig. 3-15), a proposed PPP contribution with glyceraldehyde 3-phosphate toward proposed osmolyte (glycerol) synthesis may be occurring.

Conclusion

The PPP can contribute with NADPH for FA synthesis in Acb (where cells accumulate TAG during log and stationary phases), with glucose (for starch) and with precursors for nucleotides and nucleic acids in Acsw, and with glycerol synthesis in

hs.

Carbonic anhydrases

Overview

As displayed in Figs. 3-3 and 3-14, carbonic anhydrase (CA) genes were distinctively globally upregulated during log phase in Acsw, and during log and stationary phases in hs. The solubility of CO_2 is decreased in high salt compared to low salt concentrations (Liu et al, 2011). This suggests that CA activation might be needed in hs due to less CO_2 available. In Acsw log there might be a greater CO_2 need due to accelerated growth.

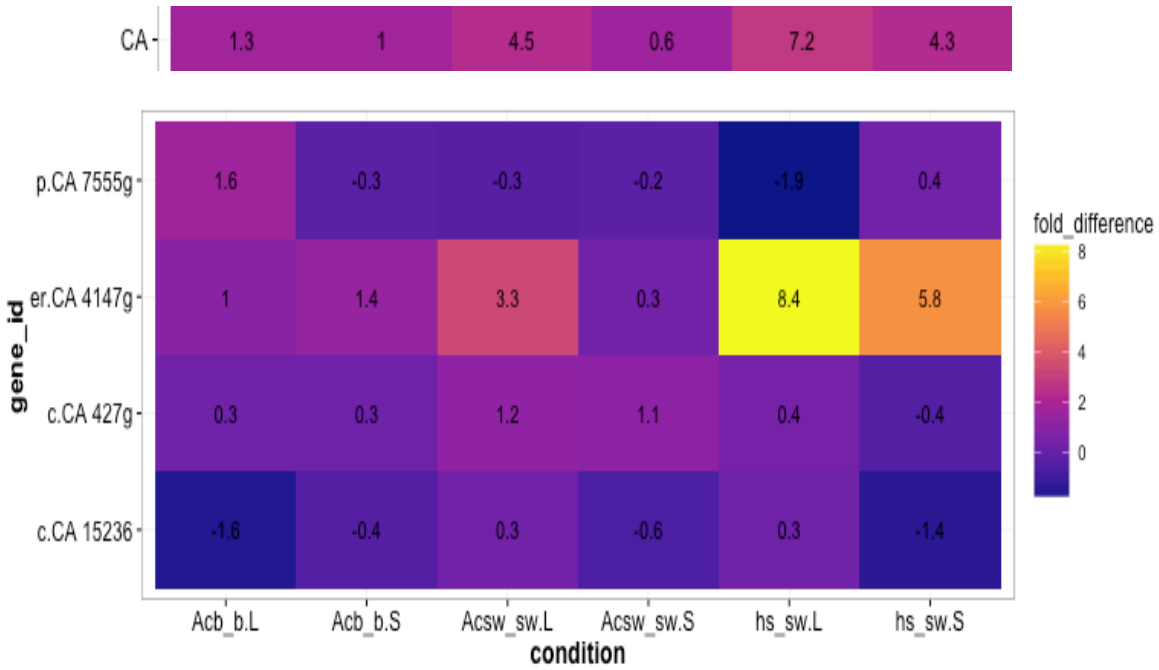

Figure 3-14. Carbonic anhydrase (CA) genes expression heatmap. Acb_b represents acetate (50 mM) + bicarbonate (178 mM) compared to bicarbonate (178 mM) in MS4 freshwater medium. Acsw_sw represents acetate (50 mM) in ASW compared to ASW. hs_sw represents ASW (1.5 M NaCl) compared to ASW (0.4 M NaCl). L denotes log phase and S denotes stationary phase. The numbers inside each rectangle as well as the colors, represent gene expression fold difference per condition. On the Y axis, "p" denotes plastid targeting and "c" denotes cytoplasmic targeting. The abbreviated gene/enzyme names are followed by their gene ID number and by "g" if the sequence was obtained from *Mk* KAS603 genome assembly data, otherwise it corresponds to the transcriptome assembly data. All numbers are log 2 based. The upper part of the figure (denoted as CA) represents the total fold change per condition, as depicted in Fig. 3-3.

ER CA

Two CAs were predicted to be targeted to the cytoplasm, one CA was plastid targeted and another was ER targeted. The ER CA encoding gene displayed the highest upregulation during log and stationary phases in hs. During hypersalinity there is less available CO_2 and CA should be more active. Since the ER is a continuous membrane network of a series of flattened sacs (with a large surface area) in close proximity to the plasma membrane, the upregulation of the ER targeted CA encoding gene might suggest activity to sequester and interconvert CO_2 (entering the cell) and bicarbonate to facilitate CO_2 release into the cell and consequently to the plastid (maximizing CO_2 sequestration). During log phase in Acsw, this gene was also upregulated, possibly due to a higher requirement of CO_2 during cell division.

Plastid CA

As denoted in Fig. 3-14, the plastid targeted CA encoding gene was uniquely upregulated during log phase in Acb, possibly due to a high demand of CO_2 during accelerated growth (induced by acetate) and the combined presence of bicarbonate possibly being transported into the plastid for CO_2 interconversion.

Conclusion

CA were targeted to 3 different compartments, and ER seemed to be the most active compartment possibly with the role of an extensive membrane system that facilitates incoming CO_2 interconversion to bicarbonate during conditions of higher

demand for CO_2 , such as hypersalinity (less solubility of CO_2 in high salt) and accelerated cell division (more cells uptaking CO_2). In the plastid CA seems to also be active when acetate and bicarbonate enter the cell and create an "extra" need for carbon fixation in growing cells.

Glycerol

Overview

Glycerol is proposed to accumulate in *Mk* KAS603 to serve as an osmoregulator. Glycerol metabolism enzymes were targeted to the cytoplasm and plastid. As displayed in Figs. 3-3 and 3-15, a global downregulation of the glycerol synthesis genes was displayed in Acb (during log and stationary phases) as well as Acsw and hs during stationary phase, where cells might not have an active interconversion of starch to glycerol. A significant upregulation was displayed during log phase in hs, where there might be a need for glycerol in growing cells, to be used as an osmolyte.

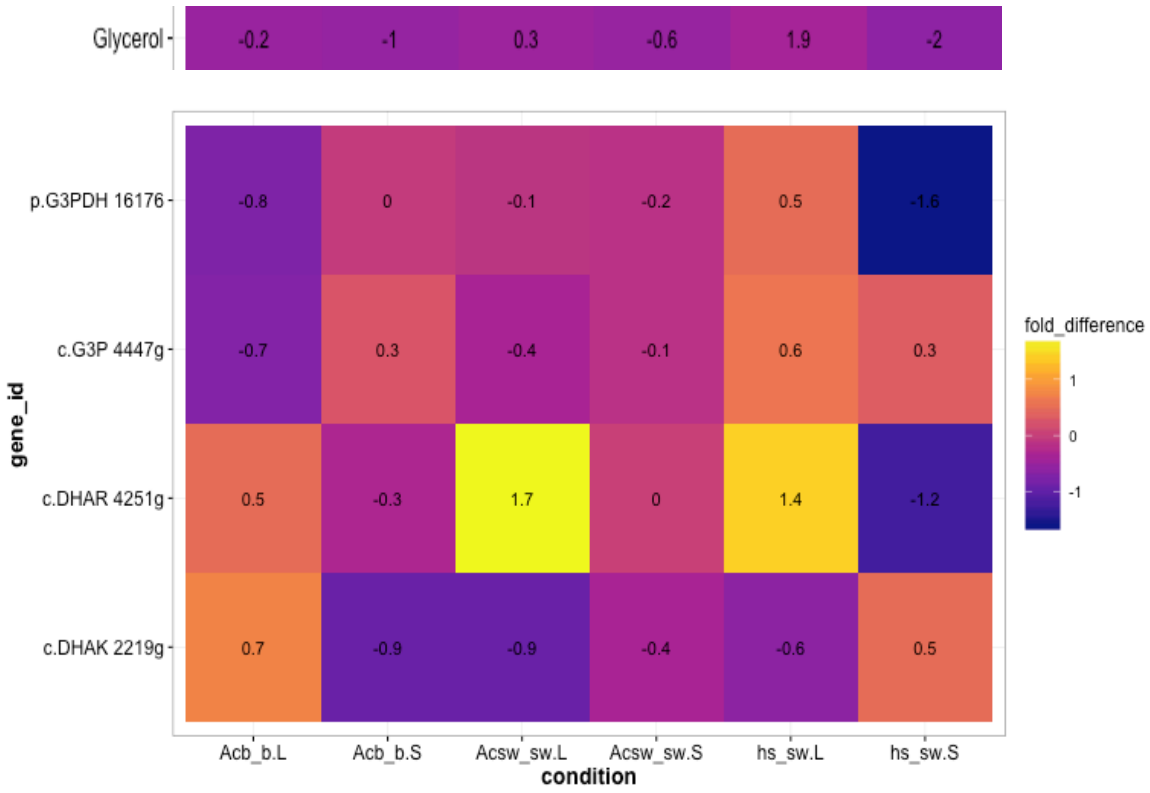

Figure 3-15. Glycerol metabolism gene expression heatmap. Acb_b represents acetate (50 mM) + bicarbonate (178 mM) compared to bicarbonate (178 mM) in MS4 freshwater medium. Acsw_sw represents acetate (50 mM) in ASW compared to ASW. hs_sw represents ASW (1.5 M NaCl) compared to ASW (0.4 M NaCl). L denotes log phase and S denotes stationary phase. The numbers inside each rectangle as well as the colors, represent gene expression fold difference per condition. On the Y axis, "p" denotes plastid targeting and "c" denotes cytoplasmic targeting. The abbreviated gene/enzyme names are followed by their gene ID number and by "g" if the sequence was obtained from *Mk* KAS603 genome assembly data, otherwise it corresponds to the transcriptome assembly data. All numbers are log 2 based. The upper part of the figure (denoted as glycerol) represents the total fold change per condition, as depicted in Fig. 3-3.

DHAR

Starch might be converted to glycerol to meet osmotic requirements (as has been proposed for *Dunaliella* [Shariati and Hadi, 2011]) in Acsw (during log phase) cultivated *Mk* KAS603 cells, and possibly facilitate survival. As displayed in Fig. 3-15, in *Mk* KAS603 the gene corresponding to dihydroxyacetone reductase (DHAR) that catalyzes the interconversion of glycerol to dihydroxyacetone, was the only gene upregulated in the glycerol synthesis pathway gene set during log phase Acsw. In *Dunaliella salina*, DHAR has been reported as an enzyme that could balance the glycerol cycle by its forward and backward reactions (Chen and Jiang, 2012). We suggest a similar occurrence in *Mk* KAS603, where DHAR possibly promotes the glycerol pathway toward glycerol synthesis. In hypersalinity (log phase) three out of four glycerol genes were upregulated, with distinctive upregulation by DHAR, suggesting the synthesis of a proposed osmolyte (glycerol) during growth in a hypersaline environment.

Conclusion

Seawater supplemented with acetate seems to be a condition where fast dividing cells might accumulate more glycerol than slower dividing cells (control without acetate). During stationary phase in Acsw the cell might not have a need to further produce glycerol. In hypersaline adapted *Mk* KAS603 where cells don't experience growth rate increase, glycerol accumulation is proposed to take place for a

possible use as an osmolyte. During stationary phase in hs, *Mk* KAS603 seems to switch to glycine and serine (predicted osmolytes) production through photorespiration (photorespiration genes were only upregulated during this condition).

Overall Summary

Mk KAS603 was cultivated under 2 conditions where growth and TAG accumulation were induced (Acb and Acsw) and 1 condition where TAG accumulation was induced (hs) without decreasing growth. From those, log and stationary phases were analyzed for expression changes in genes encoding enzymes (targeted to different cell compartments) involved in several metabolic routes. The suggested overall metabolic behavior per condition and growth phase, are described as follows. As displayed in Fig. 3-16, cells cultivated in Acb during log phase are proposed to maintain an activated cytoplasmic glucose metabolism and starch accumulation for energy generation during accelerated growth. FA synthesis is proposed to be activated for membrane formation (and for initial TAG formation), PPP for ribose for nucleic acids, glyoxylate cycle for energy and NADH for growth. A constant TCA cycle for energy generation and a low photorespiration was predicted during this mixotrophic condition. Cells cultivated in Acb during stationary phase were predicted to maintain an activated cytoplasmic/plastid glycolysis for energy, pyruvate and AcCoA for FA synthesis and TAG. Starch accumulation was predicted to be downregulated, Beta oxidation was predicted as activated for AcCoA and

energy generation, as were PPP for NADPH for FA-TAG, and glyoxylate cycle for malate and NADH for FA. A constant TCA cycle for energy generation was proposed. A low photorespiration was predicted during mixotrophy. FA synthesis was predicted as activated in the plastid and cytoplasm. TAG synthesis was predicted as activated in the cytoplasm and ER.

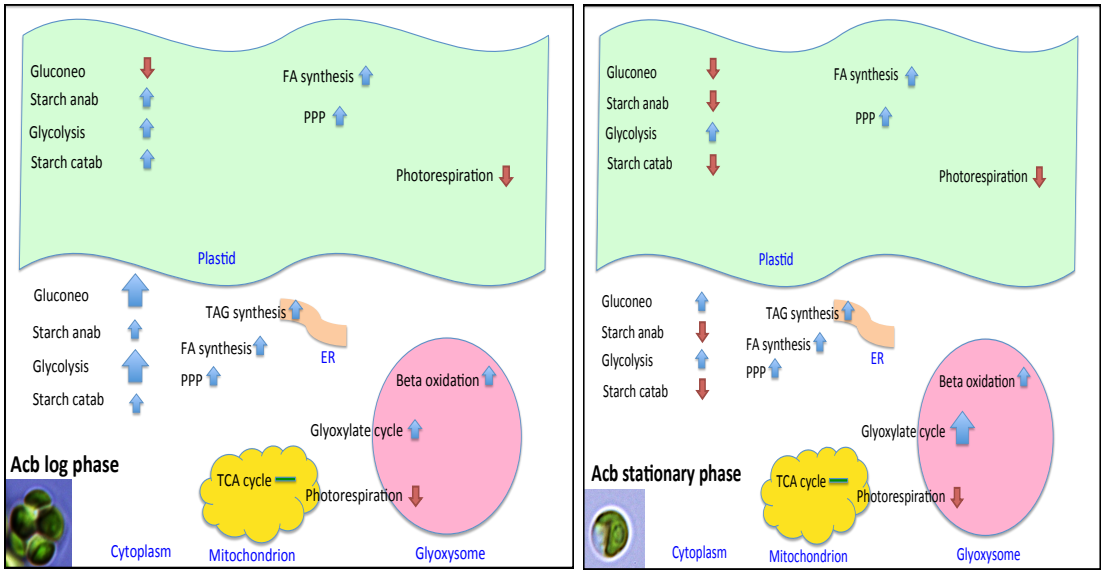

Figure 3-16. *Mk* KAS603 proposed metabolic behavior during log and stationary phases in Acb cultivation. Size of arrows display proposed level of activity. Blue arrows display upregulation and red arrows display downregulation, horizontal green lines display no change. Abbreviations are Gluconeo for Gluconeogenesis, anab for anabolism, catab for catabolism, PPP for pentose phosphate pathway.

As displayed in Fig. 3-17, cells cultivated in Acsw during log phase were predicted to activate starch synthesis (and to downregulate starch catabolism) in the plastid and cytoplasm, favoring starch accumulation in both compartments. FA synthesis was predicted as activated in the plastid and cytoplasm and a predicted downregulation of beta oxidation seemed to favor FA accumulation for new cell membrane formation during accelerated growth. A constant PPP, glyoxylate and TCA cycle for energy and reductive agents generation, was suggested. A low photorespiration was predicted during mixotrophy. TAG synthesis wasn't induced. A predicted activation of glycerol synthesis was suggested to sustain an osmotic balance of cells growing at accelerated rates (due to acetate) in seawater. Cells cultivated in Acsw during stationary phase are predicted to activate plastid gluconeogenesis for glucose oxidation through (activated) PPP for NADPH directed to (activated) FA-TAG synthesis. Starch metabolism seems not to be induced (neutral), as the cell may be using lipids instead of starch for energy because the use of acetate may be reducing the overall energy drain on the cell from fixing carbon, so that the cell can "afford" to use lipid.. Activated FA synthesis for TAG, and upregulated FA degradation for energy and AcCoA generation, are predicted. Glycolysis is predicted to be upregulated in the cytoplasm for energy and pyruvate for AcCoA for FA-TAG. A downregulation of the glyoxylate cycle and a constant activity of the TCA cycle for energy and reductive agents generation, was predicted. A low photorespiration was predicted during this mixotrophic condition. Glycerol synthesis was not predicted during stationary phase.

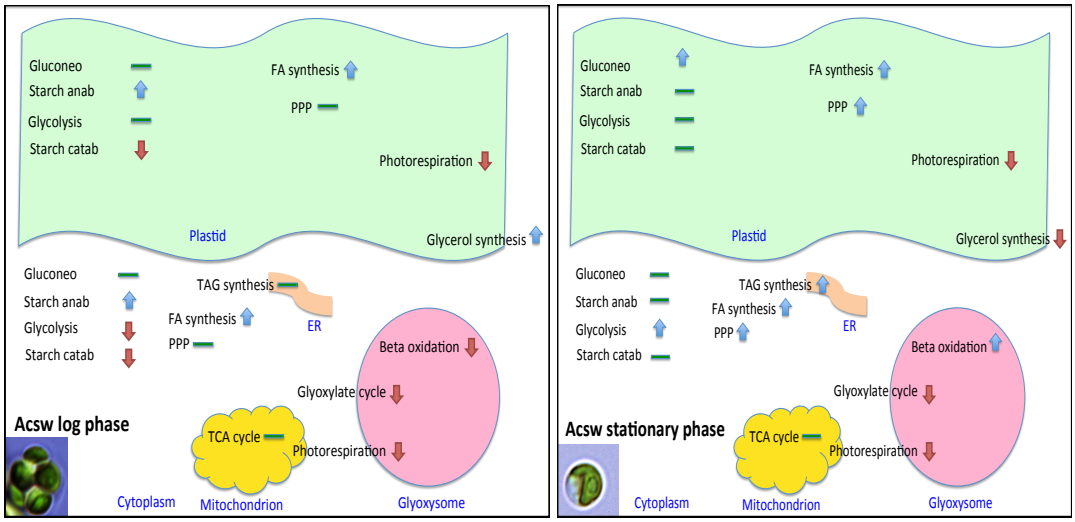

Figure 3-17. *Mk* KAS603 proposed metabolic behavior during log and stationary phases in Acsw cultivation. Size of arrows display proposed level of activity. Blue arrows display upregulation and red arrows display downregulation, horizontal green lines display no change. Abbreviations are Gluconeo for Gluconeogenesis, anab for anabolism, catab for catabolism, PPP for pentose phosphate pathway.

As displayed in Fig. 3-18, cells cultivated in hs during log phase were predicted to have a downregulated starch catabolic route to favor starch accumulation in the cytoplasm and plastid. Gluconeogenesis was predicted to be activated in the plastid for a possible glucose oxidation through the (activated) PPP for ribose for nucleic acids. Glycolysis was predicted as activated in the cytoplasm for energy production during growth. FA beta oxidation and FA-TAG synthesis were not induced. The glyoxylate cycle was predicted as downregulated since TCA seems to be activated for energy and NADH production for growth and glycerol synthesis (which was predicted as activated). Glycerol is predicted to have an osmotic regulation role in *Mk* KAS603 during growth. The cells were predicted to experience a low photorespiration possibly to save energy from a costly side reaction of photosynthesis. Cells cultivated in hs during stationary phase were predicted to have an activated plastid gluconeogenesis for (activated) starch synthesis. A predicted downregulation of starch degradation in the plastid and cytoplasm may favor starch accumulation in both compartments. High TAG accumulation may be favored by predicted activated FA synthesis and by inactivation of FA degradation (beta oxidation). Upregulation of PPP can produce NADPH for TAG accumulation. Glyoxylate is predicted to be produced through (activated) photorespiration due to downregulated glyoxylate cycle, where it can be converted to glycine and serine (proposed osmolytes). Glycerol synthesis is predicted to be downregulated and the TCA cycle to be constantly generating energy.

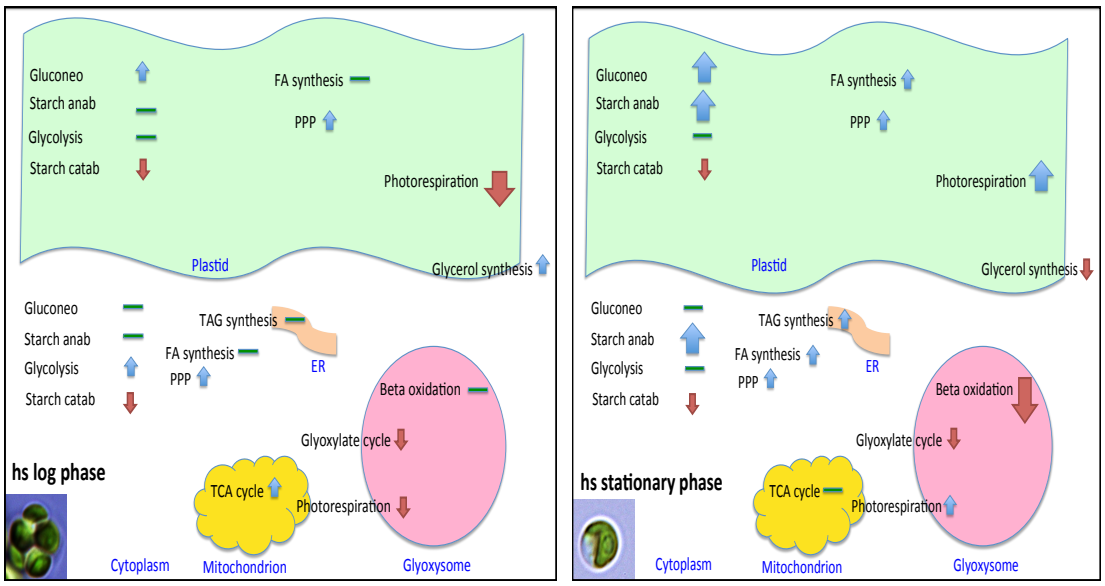

Figure 3-18. *Mk* KAS603 proposed metabolic behavior during log and stationary phases in hs cultivation. Size of arrows display proposed level of activity. Blue arrows display upregulation and red arrows display downregulation, horizontal green lines display no change. Abbreviations are Gluconeo for Gluconeogenesis, anab for anabolism, catab for catabolism, PPP for pentose phosphate pathway.

Most routes analyzed were predicted to be activated during log phase Acb since this is the most active condition where cells can grow, divide and initiate TAG accumulation. During stationary phase in Acb, TAG is accumulated and predicted as favored by PPP that provides NADPH and by a distincltively upregulated glyoxylate cycle that provides malate and NADH. PPP was upregulated in Acb during stationary phase since it is coupled with gluconeogenesis (also upregualted). In Acsw log phase the PPP was not induced possibly due to less glucose availability through a non-induced gluconeogenesis. Cells in Acsw and hs seem to accumulate glycerol during growth for a possible use as an osmoregulator. The TCA cycle seems to be preferentially active during log phase in hs, which may be due to "extra" energy needed to sustain growing cells in a hypersaline environment (osmolytes, more cell transporters, possible enhaced CCM due to less soluble CO_2 in hs) and cells in hs during stationary phase seem to activate photorespiration to contribute to the synthesis of serine and glycine (predicted osmolytes). Distinctive downregulation of beta oxidation in hs stationary phase, was predicted to contribute to FA-TAG accumulation

3.6 ACKNOWLEDGMENTS

We would like to acknowledge the Matteo Pellegrini lab at UCLA for sequencing, assembly and annotation of transcripts as well as for calculating transcript abundances. A special thanks to Mark Hildebrand (Scripps Institution of Oceanography, UCSD), for his guidance and for making this analysis so enjoyable. Funding for this project was provided by UC-MEXUS-CONACYT grant CN-12-622, by AFOSR grant FA9550-08-1-0178 and by UCLA Genomics and Proteomics grant DE-FC02-02ER63421. Funding for Eva Sánchez was provided by the UC-MEXUS-CONACYT Fellowship, by the Shirley Boyd Memorial Fellowship F-1116 and by The Mexican Marine Sciences Scholarship Fund F-1438.

Chapter 3 contains unpublished work with co-authors: Mark Hildebrand, Matteo Pellegrini, David López, Jing Lu, Megan Inkeles, Weihong Yan and Shawn Cokus. The dissertation author was the primary researcher and author of this material.

3.7 REFERENCES

Anders S, Pyl PT, Huber W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 31(2):166-169.

Andrews S. (2011). FastQC: A quality control tool for high throughput sequence data (Babraham Bioinformatics, Cambridge, UK).

Blanc G, Duncan G, Agarkova I, Borodovsky M, Gurnon J, Kuo A, Lindquist E, Lucas S, Pangilinan J, Polle J, Salamov A, Terry A, Yamada T, Dunigan DD, Grigoriev IV, Claverie JM, Van Etten JL. (2010). The Chlorella variabilis NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. Plant Cell. 22(9):2943-2955.

Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, Holt C, Sanchez-Alvarado A, Yandell M. (2008). MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 18(1):188–196.

Coleman JR, Grossman AR (1984). Biosynthesis of carbonic anhydrase in *Chlamydomonas reindardtii* during adaptation to low CO2. Prov. Natl. Acad. Sci. USA. 81(19):6049-6053.

Chen H, Lu Y, Jiang J-G. (2012). Comparative analyisis of the key enzymes of the glycerol cycle metabolic pathway in *Dunaliella salina* under osmotic stresses. PLoS ONE. 7(6):p.e37578

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 29(1):15-21.

Fan J, Ning K, Zeng X, Luo Y, Wang D, Hu J, Li J, Xu H, Huang J, Wan M, Wang W. (2015) Genomic foundation of starch-to-lipid switch in Oleaginous Chlorella spp. Plant physiol. 1;169(4):2444-61.

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.

Gee CW, Niyogi KK (2017). The carbonic anhydrase CAH1 is an essential component of the carbon- concentrating mechanism in *Nannochloropsis oceanica*. Prov. Natl. Acad. Sci. USA. 114(17):4537-4542.

Hass BJ, Delcher AL, Mount SM, Wortman JR, Smith Jr, RK Jr, Hannick LI, Maiti R, Ronning CM, Rusch DB, Town CD et al. (2003). Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. Nucleic Acids Res. 31:5654-5666.

Hildebrand, M. and K. Dahlin (2000). Nitrate transporter genes from the diatom *Cylindrotheca Fusiformis* (Bacillariophyceae): mRNA levels controlled by nitrogen source and by the cell cycle. Journal of Phycology 36: 702-713.

Kim J, Lee JY, Ahting C, Johnstone R, Lu T. (2014). Growth of *Chlorella vulgaris* using sodium bicarbonate under no mixing condition. Asia-Pacific J of Chem Eng. 9(4):604-609.

Lee JY, Orr-Weaver TL. (2001). The molecular basis of sister-chromatid cohesion. Annu. Rev. Cell Dev. Biol. 17:753-777.

Liu Y, Hou M, Yang G, Han B. (2011). Solubility of CO_2 in aqueous solutions of NaCl, KCl, $CaCl_2$ and their mixed salts at different temperatures and pressures. J. Sup Fluids. 56(2):125-129.

Love M, Huber W, Anders S. (2014). Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome Biol. 15(12):550.

Paddy MR (1998). The Tpr Protein: Linking structure and function in the nuclear interior? Am. J. Hum. Genet. 63:305-310.

Parisi S, Mckay MJ, Molnar M, Thompson MA, van der Spek PJ, et al. (1999). Rec8p, a meiotic recombination and sister chromatid cohesion phosphoprotein

of the Rad21p family conserved from fission yeast to humans. Mol. Cell. Biol. 19: 3515–3528.

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

Sharma T, Gour RS, Kant A, Chauhan RS. (2015). Lipid content in Scenedesmus species correlates with multiple genes of fatty acid and triacylglycerol biosynthetic pathways. Algal Research. 12:341-349.

Shariati M, Hadi MR. (2011). Microalgal biotechnology and bionergy in *Dunaliella*. Progress in Molecular and Environmental Bioengineering – From analysis and modeling to technology applications. pp. 483-506.

Shen WH. (2002). The plant E2F-Rb pathway and epigenetic control. TREND in Plant Science. 17(11):505-511.

Solovyev V, Kosarev P, Seledsov I, Vorobyev D. (2006) Automatic annotation of eukaryotic genes, pseudogenes and promoters. Genome Biol. 7(1):10.1-10.12.

Stanke M, Steinkamp R, Waack S, Morgenstern B (2004). Nucleic Acids Res. 32(2):w309-w312. DOI: https://doi.org/10.1093/nar/gkh379

Yamamoto M, Fujishita M, Hirata A, Kawano S. (2004). Regeneration and maturation of daughter cell walls in the autospore-forming green alga Chlorella vulgaris (Chlorophyta, Trebuxiophyceae). J Plant Res. 117:257-264.

Zachleder V, Schlafli O, Boschetti A. (1997). Growth-controlled oscillation in activity of H1 kinase during the cell cycle of *Chlamydomonas reinhardtii* (Chlorophyta). J Phycol. 33:673-681.