9-2015

Explorations In Advancing Microalgae Productivity

Peter Neofotis
Graduate Center, City University of New York

Recommended Citation
Neofotis, Peter, "Explorations In Advancing Microalgae Productivity" (2015). CUNY Academic Works.
https://academicworks.cuny.edu(gc_etds/1072

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
EXPLORATIONS IN ADVANCING MICROALGAE PRODUCTIVITY

By Peter Neofotis

A dissertation submitted to the Graduate Faculty of Biology (Plant Subprogram) in partial fulfillment of the requirements of a degree of Doctor of Philosophy, the City University of New York
2015
This manuscript has been read and accepted for the Graduate Faculty in Biology to satisfy the dissertation requirements for the degree of Doctor of Philosophy

Jürgen Polle, Chair of Examining Committee

Date ____________  ________________________________

Date ____________  Laurel Eckhardt, Executive Officer  ________________________________

Defense Committee

Amy Litt

Mateo Pellegrini

Louis Bradbury

Theodore Muth

Mara Schvarztein

THE CITY UNIVERSITY OF NEW YORK
Abstract

EXPLORATIONS IN ADVANCING MICROALGAE PRODUCTIVITY

by

Peter Neofotis

Advisor: Jürgen Polle

This thesis is part of an effort to advance microalgae cultivation for biofuel and high value products. The first chapters describe a large scale screening effort in which novel strains that hold great promise in large outdoor cultivation systems were isolated. Among these are strains of *Scenedesmus obliquus*, *Borodinellopsis texensis*, *Chlorella sorokiniana*, *Ankistrodesmus*, *Coelastrella*, and a previously uncharacterized species in the Chlamydomonadales. These species are characterized in terms of their phylogenetic relationship, biomass productivity, and, for some, lipid and carotenoid profiles under growth and stress conditions. Also discussed in regards to growth metabolism and oil productivity is a unique enzymatic characteristic found in *S. obliquus*, the presences of a chloroplast localized enolase. Lastly, we demonstrate a novel bacteria-algae mutualistic relationship that may be of particular interest in cultivating microalgae in wastewater. Overall, the work advances algae productivity with algal species and bacteria that may be extended further upon to push forward the field of algae biofuel production.
Acknowledgments

Part of this research was supported by the Air Force Office of Scientific Research (grant #FA9550-08-1-0170 and #FA9550-08-1-0403) and by the US Department of Energy (grant #DE-EE0003129 and #DE-EE0003046-28302B). The authors greatly appreciate support by their collaborators within the National Alliance for Advanced Biofuels and Bioproducts, specifically Dr. J. Olivares and Dr. R. Sayre. Dr. D. Tran, Dr. Mahendra-Perumal, and Ms. Sophia Registe’s support in the early work of strain isolation is recognized. The authors also thank the members of the boyscout troop 1949 in Katy, TX and the high school students working with Dr. C. McEntee at Brooklyn College for sampling and isolation of strains. Dr. S. Slocombe and Dr. J. Benemann also provided comments and suggestions for improvement of chapter 1. Funding for the work in Chapter 3 was provided by grant #DE-EE0003129 of the US Department of Energy.
Table of Contents

Thesis Introduction .......................................................... 1

Chapter 1: Microalgae Strain Isolation, Screening, and Identification for Biofuels and High Value Products ................. 8

Chapter 2: Characterization and Classification of Highly Productive Microalgae Strains Discovered for Biofuel and Bioproduct Generation .................................................. 47

Chapter 3: The Chloroplast Enolase and its Presence in an Oleaginous Microalgae ...................................................... 96

Chapter 4: Syntrophic Metabolism of Ethanol by Agrobacterium tumefaciens Enhances Micro-algal Growth ....................... 112

Overall Conclusions .......................................................... 129

Bibliography .................................................................. 138
List of Tables

Chapter 2
Table 1: Listed are the top performing strains that were classified in the work reported here. Included in the table are the strain numbers, taxon, and location along with the habitat of collection. Also listed is whether or not the strain is carotenogenic and whether or not it was grown successfully outdoors. Strains with the ID DOE#### are from the NAABB phycoprospecting project. All these strains have been deposited at the University of Texas Culture Collection of Algae and are available to researchers. Strains with the ID EN#### resulted from an Airforce Office of Science funded phycoprospecting effort. The asterix next to Strain ID numbers indicates that the strain was tested successfully in outdoor ponds at NAABB testbed sites (NAABB Final Report, 2014). ................................................................. 60

Chapter 3
Table 1: Information on the genes coding for the enolase in green algae as extracted from the database of the US Joint Genome Institute (as of 05/23/2014). ................................................................. 104

Appendix
Table 1: Advantages and disadvantages in domesticating the Microalgae when compared to the Angiosperms, showing in red, greater challenges, and in green, greater advantages........ 130
List of Figures

Introduction

Figure 0.1: General tree of life for photosynthetic organisms, with solid arrows depicting primary or secondary endosymbiotic events, and dashed arrows representing multiple events, some of which are not fully resolved. .................................3

Chapter 1

Figure 1: An overview for the bioprospecting approach from sampling to screening and testing of the strains is shown. Samples were collected from various habitats, aliquots plated, colonies picked, transferred into liquid medium and then cell-sorted onto agar-media culture plates. Strains were screened by growth in liquid batch cultures. Strains showing good growth characteristics and/or high Nile Red fluorescence were deemed candidates and grown in bubble-columns for biomass productivity measurements. .................................14

Figure 2: Photos showing representative habitats of the wide range of environments from which algal strains could be isolated. Left panel from top to bottom - Freshwater algal bloom in a road side ditch in Texas; Microalgal bloom in the Brooklyn College freshwater pond in Brooklyn, NY; Pacific Ocean in San Diego, CA; Diverse lichens on a rock in Arizona. Right panel from top to bottom - Freshwater birdbath in San Diego, CA; Diatom bloom in the Salton Sea, CA; Green algae of the genus Dunaliella blooming in a brine pond adjacent to the Great Salt Lake in Utah; Biofilm of microalgae at a roadside location in Louisiana. 17

Figure 3: Light microscopic image of an environmental freshwater sample originating from the Brooklyn College pond in Brooklyn, NY. Based on morphological characters the sample contained green algae such as Desmodesmus, Ankistrodesmus, Kirchneriella, Chlamydomonas, and Chlorella type cells as well as a diatom species. In addition, there were some unknown algal species visible. The black bar represents 10 µm. .................17

Figure 4: Algal colonies grown on agar-solid media. Left panel - Aliquots of the same environmental freshwater sample were placed on agar-solid media in regular standard 90 mm Petri dishes. Three different growth media as indicated on the top left of each plate were used supplemented with silicate to allow for growth of diatoms. Right panel - An example of algal colonies grown on agar-solidified medium shown with a magnification of 60x. .........................................................................................................................18

Figure 5: The frequency of the number of microalgae strains isolated per environmental sample from a total of 350 environmental samples. The x axis indicates the number of isolates per sample, and the y axis indicates the frequency of such isolates. .............................................22

Figure 6: An example scatter plot of a mixed sample, with the dots representing single events as visualized by the FACS Aria System II cell sorter. The different populations were selected and single events sorted onto agar-solidified media in Petri dishes. The parameter chlorophyll fluorescence at 695 nm was used to select populations of algal cells via cell sorting. Bacterial cells may display chlorophyll fluorescence, but were usually at values below 10^1 relative fluorescence. .........................................................................................24

Figure 7: FACS cell sorting of events from liquid cultures into 96-well plates. A) Sorted here was the green alga Dunaliella salina strain CCAP19/18. Cells of D. salina are fragile, as they do not have cell walls. Furthermore, the sorted culture was grown in 1.0 M NaCl medium. Nevertheless, it was possible to sort cells and recover viable cultures in liquid medium. In contrast, it was not possible to sort cells of D. salina onto regular plates with solidified medium. B) Sorted here were populations of a new, unknown, and carotenogenic isolate. All wells contained cultures, indicating that the sorting process had a high efficiency and that the cells survived sorting. On this plate, it is evident that the conditions in the wells were not uniform. Cultures in the outer wells had already turned orange indicating that they
were under stress, whereas the cultures in the inner wells were still green. Also, a number of the inner wells show heavy condensation on the cover plate. 

**Figure 8:** Examples of microalgal colonies growing on agar-solidified media in Petri dishes after FACS cell sorting. A) Plate from a non-unialgal liquid culture. Although one strain seemed to dominate, several different colony colors and types indicate the presence of at least 4 different species. B) One unialgal strain only, with most colonies being axenic. C-K) 60x magnification images of algal colonies sorted from a single event that are: C-E) unialgal and axenic; F-H) non-unialgal; I-K) containing bacteria as indicated by the white, beige, or orange rings around the colonies. 

**Figure 9:** Screening of strains of microalgae for growth and lipid accumulation was performed in batch cultures using Erlenmeyer flasks on a shaker. A) Shown is an example for the first-level screen of growth of microalgal cultures in batch cultures on a shaker. B) The strain shown was isolated from a C-Medium plate, but grew best in liquid BG11 medium. C) The strain was isolated from a BG11 medium plate, but grew well on all three media tested. D) The strain was isolated from a BG11 medium plate, but best growth occurred in liquid BBM medium. Note the clumpy growth of this strain in C-Medium and BG11 medium. 

**Figure 10:** Normalized Nile Red fluorescence representing the lipid contents of cells. The Nile Red signal is shown for three strains during the screening period. Error bars represent standard deviation among triplicates. Although strain EN1234 is carotenogenic and over-accumulates triacylglycerols in cytosolic oil bodies (see Figure 11E and 11F), it did not display high relative Nile Red values in our first-level screen. Low Nile Red fluorescence levels for carotenogenic strains were often seen in the first-level screen. This may be due, in part, to thick cell walls prohibiting efficient penetration of Nile Red into the cells. In addition, the carotenoids within the oil bodies might quench Nile Red fluorescence. Although such carotenogenic strains did not pass the Nile Red screen, their orange and reddish culture colors indicated triacylglycerol accumulation and so were advanced to the second level of screening. 

**Figure 11:** Cells from two of the strains deemed to be potential high-producers, showing bright-field (A, C, E) and Nile Red fluorescence images from the same pairwise field of view at 530 nm (B, D, F). A-B) *Scenedesmus obliquus* strain EN0004 with oil bodies visible with fluorescence microscopy. C-D) Strain EN1234 (a putative *Coelastrum*-related Desmid) from a young culture with few significant oil bodies visible. E-F) Stressed cells of strain EN1234 taken from an older culture. Note the orange coloration under bright field indicative of carotenoid accumulation (E) and apparent Nile Red fluorescence (F). The bar indicates 10 μm. 

**Figure 12:** Identification of strains that over-accumulate carotenoids. A) Sometimes colonies of microalgae grown on plates would display from a yellowish, orange, or reddish color indicating that cells over-accumulated carotenoids. B) More often strains of liquid cultures from the first-level screen would display yellowish, orange, or reddish color upon aging. C) Shown is a culture displaying an apricot color indicating accumulation of carotenoids. D) Microscopy of cells from culture shown in C). The cells have apricot color originating from very large oil bodies that sequester the carotenoids (data not shown). The bar represents 10 μm. 

**Figure 13:** Displayed are UHPLC traces that show the absorbance of pigments from cell extracts plotted as the 450nm/550nm difference. Different pigments such as chlorophylls and carotenoids elute at known times and the relative level of absorbance represents the amounts of these pigments present in the strains. Upper panel – Profiles of extracts from strain EN1234. Lower Panel – Profiles of extracts from carotenoid over-accumulating cells of strain EN1234 and *Haematococcus* sp. strain Haema001. 

**Figure 14:** Shown are bubble-columns (numbered 1 to 7 from left to right) with each containing a unique, newly isolated strain. *Scenedesmus obliquus* strain EN0004 was grown in the
leftmost column (column 1). The sequence of photos visualizes the different growth rates of the various strains. Photos were taken 24 hours, 48 hours, and 120 hours after culture inoculation. ................................................................. 39

Figure 15: Plotted are growth curves based on biomass for the strains shown in Figure 14. Some strains such as #5 did not show any growth inhibition during our experiment and could reach biomass concentrations of more than 5 g/L. ................................................................. 40

Chapter 2

Figure 1: Schematic summarizing the overall phycoprospecting process performed in our laboratory as part of the NAABB Consortium to identify the best performing microalgae strains for biofuel production. ................................................................. 60

Figure 2: Shown are growth curves based on the AFDW (grams/liter) of some of the most productive strains. The Chlamydomonadales species was strain DOE0101; the Chlorella species was Chlorella sorokiniana strain DOE1412; The Coelastrella strains were DOE0088, DOE0202, and DOE0369; and the Scenedesmus obliquus strains used were DOE0152 and DOE0111. At least two biological replicates were included for each strain. Shading represents the 95% confidence region for the regression fit of the biological replicates. ........................................................................... 64

Figure 3: Compared are cells and metabolites of the Chlamydomonadales sp. strain DOE0101 (upper panel) and the Coelastrella sp. strain DOE0202 (lower panel). In the center of the figure, PDA profiles (absorbance traces at 450 nm) are displayed of extracts from cells of actively growing cultures and of carotenogenic (orange) cultures. Photos for representative cells from these cultures are shown for both strains. The black bars underneath the cells of strain DOE0101 indicate 10µm. The same scale was used for photos of cells from strain DOE0202. Extractions were conducted from the same volume of cells for cultures of the same age. Comparison with retention times of concurrent MS-TOF measurements suggest the carotenoids are - following the number or profiles - 1) Adonirubin or Astaxanthin; 2) Lutein and/or Zeaxanthin; 3) Canthaxanthin or Adonixanthin; and 4) Echinenone or Cryptoxanthin. Further identification of the isomers was beyond our analytical capabilities. However, based on the late retention times, it is assumed for extracts from carotenogenic cultures that Astaxanthin Diesters (labeled as AstaDE) are represented by the late peaks. The outer panels of this figure show corresponding TAG profiles of the cells of both strains. Peaks visible at longer retention times are indicative of TAG molecules containing fatty acids with greater degrees of saturation. Peaks at shorter retention times represent presence of TAG molecules containing fatty acids with more double bonds. The overall class of TAG molecules shown in each profile is indicated by the labels with the example of 16_16_16 indicating TAG molecules containing three fatty acid chains each having 16 carbon atoms. Additional lipid profiles can be seen in SM-2. ................................................................. 68

Figure 4: ITS2 sequence-structure tree of novel strains for biofuel production and also known strains of the Chlorophyta – rooted with two prasinophyceae. Significant bootstrap values are based on 500 replicates and mapped to the appropriate internode. Branch lengths are drawn proportional to inferred changes. Arrows point to strains discussed in this paper in more detail. ............................................................................. 71

Figure 5: Photos showing a variety of cells of S. obliquus strain DOE0152. A) Side and top views of several four-celled coenobia and one single cell in the center. B) A larger four-cell coenobium where the individual cells began division, as indicated by broadening pyrenoids. C) A four-celled coenobium where one cell already completed cell divisions and released a new four-celled coenobium. ............................................................................. 73

Figure 6: Non-rooted ITS2 sequence tree including our novel strains DOE0088, DOE0155, DOE0202, DOE0369, EN1234, EN1235, and also known strains of the Scenedesmaceae.. 76
Figure 7: Light microscopic photos of cells of *Coelastrella* sp. strain DOE0202 (A-H). A) Several autosporangia containing daughter cells. B) Side view of vegetative, oval, coccoid cells with one cup-shaped chloroplast, which contains one pyrenoid. C) Cells stained with iodine show a ring of starch indicated by black color on the outside of the pyrenoid. Note the Scenedesmus-like coenobia of smaller cells that were released from the autosporangia. D) Larger sized vegetative cells with the hyaline cell wall. E) Daughter cells just released from an autosporangium. The hyaline cell wall of the mother cell is visible in the background. F) A variety of vegetative cells in different growth stages. G) An autosporangium is visible on the lower left side with cells that are in the carotenogenic process. H) Center view of two vegetative carotenogenic cells where protruberances are visible on the cell wall. I) Two SEM images where striated cells ridges are visible among the oval cells. The black bar represents 10 µm for the light microscopic and the electron microscopic images. The SEM images were taken by Dr. Cooke at New Mexico State University.

Figure 8: Shown are images of cells of the novel *Ankistrodesmus* strain DOE0259. The black bar represents 10 µm. A) Two cells accumulating numerous oil bodies in the cytosol. B) Images of several cells illustrating the plasticity in cell morphology. C) Single cell. D) Two cells in the early division process. E) Two cells released from an autosporangium. F) Four daughter cells still in the autosporangium. G) Four cells released from the autosporangium.

Figure 9: Shown are photos of cells of *Borodinellopsis texensis* strain EN1423 from cultures grown in 0.5 M NaCl artificial *Dunaliella* medium. The black bar represents 10 µm. A) Shown is one autosporangium. B) A single vegetative cell with the asteroidal chloroplast originating out of the pyrenoid. C) A variety of cells in different growth stages. D) Cells that are in the process of carotenogenesis by accumulation of numerous brown/orange oil bodies in the cytosol oriented towards the outside of the cells. E) Mature cells, each with large numbers of orange/red oil bodies in the cytosol making the chloroplast invisible.

Figure 10: Shown are images of cells of the strain DOE0101 indicating the variety of cell types. The bar indicates 10 µm. A) Vegetative, coccoid cell beginning division of the pyrenoid. B) Several cells each containing one cup-shaped chloroplast and one pyrenoid per chloroplast. C) One large cell in the beginning phase of carotenogenesis and one autosporangium. D) Single cells with orange oil bodies. E) Orange cells with large oil bodies and unknown white appearing structures. F) Two autosporangia originating from orange cells.

Figure 11: Phylogenetic tree based on the rDNA region including the 3’ end of the 18s gene, the ITS1 region, the 5.8s gene, the ITS2 region, and 5’ end of the 28s gene. Sequences were aligned with ClustalW and reconstructed using Maximum Likelihood with a Tamura-Nei Model and Nearest-Neighbor-Interchange (NNJ) heuristic method, with 1000 bootstrap replications performed. The sequences for the strains *Chlorella sorokiniana* UTEX 1230 and *Chlorella sorokiniana* LANL 1228 were provided by Dr. Starkenburg.

Figure 12: Shown are photos of cells of *Chlorella sorokiniana* strain DOE1412. The black bar represents 10 µm. A) Autosporangia B) Some autosporangia and several cells in the early stages of cell division. C) Several larger cells with the pyrenoid clearly visible. D) Several autosporangia and one larger cell. One autosporangium and an opened autosporangium with released daughter cells. F) Larger cells with the pyrenoids clearly visible within the chloroplast.

Chapter 3

Figure 1: General overview of carbon metabolism in plants, with particular attention paid to glycolysis.

Figure 2: (top) Alignment of enolase protein sequences, showing the DOE0152 enolase EST 201 translated open reading frame has a n-terminal transit peptide followed by an unknown
domain. (Bottom) Alignment of the two Scenedesmus obliquus DOE0152 enolase sequences. .......................................................................................................................... 106

Figure 3: Protein Maximum Likelihood phylogeny reconstruction based on 500 bootstraps of the glycolytic enolase genes including several green algae and the three different glycolytic enolase forms from A. thaliana. Tree was rooted with the human alpha enolase. The multifunctional enolases were not included in this phylogeny, because they are too different from the glycolytic enolases and resulted in very low support values. ...................................................... 107

Figure 4: Simplified schematic showing our proposed path for carbon partitioning within the plastid of a green algal cell focused on starch, lipids, and isoprenoids. ENO indicates the enolase enzyme outside the plastid. C3, C5, and C6 represent monosaccharides with three, five, or six carbon atoms. PEP, abbreviation for phosphoenolpyruvate. The blue circle represents the Calvin-Benson cycle with 1 = Carbon fixation, 2 = reduction phase, 3 = regeneration phase. A shows the location of the enolase in Chlamydomonas. B shows the possibilities with the chloroplast enolase. .............................................................................. 108

Chapter 4

Figure 1: The cycle of carbon dioxide and oxygen in ideal synthetic algae bacteria assemblages, adapted from (Oswald et al., 1955). .......................................................................................................................... 115

Figure 2: A known pathway of ethanol metabolism (Arndt et al., 2008). Enzymes found with high confidence in the genomes of A. tumefaciens are boxed in brown while those found in Chlamydomonas reinhardtii are boxed in green. The acs gene, present in both Chlamydomonas and A. tumefaciens, is shaded both brown and green. ...................... 117

Figure 3: Combined effect of ethanol and different bacteria on growth of Scenedesmus obliquus. When ethanol was not also present in the media, no bacteria had any effect on algae growth. .......................................................................................................................... 118

Figure 4: Photo (top) and OD600 graph (bottom) showing the effect of ethanol and minimal (HS) media on bacteria growth. Although the Pseudomonas aeruginosa also grew, it did so in clumps and to a lesser degree than the Agrobacterium. The experiment was also conducted in BG11 media with similar results. .......................................................................................... 119

Figure 5: Effect of different alcohols on Scenedesmus obliquus growth. None of the carbon sources had an effect when Agrobacterium was not present. Only the combination of Agrobacterium tumefaciens and ethanol had a statistically significant (P>95%) effect on algal growth. ................................................................. 120

Figure 6: Photo (top) and graph (bottom) of the effect of ethanol and A. tumefaciens C58 on Scenedesmus obliquus (DOE 1052) growth. Only the combination of ethanol and Agrobacterium tumefaciens had a significant effect on algal growth (P<001). Photo was taken after cultivation for five days. .......................................................................................... 122

Figure 7: Growth of Chlamydomonas reinhardtii after five days with in HS media with additions of various A. tumefaciens strains and ethanol. A. tumefaciens C58 has the Ti plasmid, C58c1 does not, and GV3101 has a Ti plasmid that has been disarmed. .......... 123

Figure 8: Growth of S. obliquus in autotrophic media (BG11) with A. tumefaciens and ethanol in it for 48 hours, after which the bacteria was removed (red) – compared to S. obliquus in BG11 alone (blue). ................................................................................... 124

Figure 9: Cells of S. obliquus growing in BG11 media treated with A. tumefaciens and ethanol. ........................................................................................................................................................................................................................................... 124

Figure 10: C. reinhardtii exposed to A. tumefaciens and 1 ul ethanol/ml of media, but also bubbling with 1% CO2. When additional CO2 was not supplied, no difference was observed after three days of culturing. Photo was taken after 6 days of culturing. .................................................. 125

Figure 11: Cultures of C-530 ac17 mt-, which requires acetate (Levine & Goodenou, 1970, Ebersold et al., 1962) after 72 hours. Only the cultures grown in Tris-Acetate-Phosphate (TAP) media grew. .......................................................................................................................... 126
Figure 12: Cultures of *C. reinhardtii* treated with *A. tumefaciens* and ethanol with different strength HS media. Photo taken after 7 days, shows where (brown arrow) the culture exposed to ethanol and *A. tumefaciens*, which had bloomed, crashed first. A week later, the culture exposed to ethanol and *A. tumefaciens* (green arrow) was the first of the 1x cultures to crash.
Thesis Introduction

Peter Neofotis\textsuperscript{a,b}

\textsuperscript{a}Department of Biology, Brooklyn College, 2900 Bedford Avenue, Brooklyn, New York 11210, USA
\textsuperscript{b}The Graduate Center of the City University of New York, New York, NY 10016, USA

INTRODUCTION

The harnessing of fire is perhaps that most distinguishing feature of \emph{Homo sapiens}. No other feature sets us so clearly apart from the wild. The use of tools, for instance, is not unique to the Homo genus, as Chimpanzees (Glikson, 2014), beavers (Glikson, 2014), otters (Fujii \textit{et al.}, 2015), and even crows (Martinho \textit{et al.}, 2014) make and/or utilize them. And sophisticated language is exhibited by bird songs, bee dances, as well as whale and dolphin echoes (Glikson, 2014). Yet the \emph{Homo} genus is unique in its ability to ignite and transfer fire (Glikson, 2014). This mastery of fire has allowed man to elevate planetary entropy on a scale of an asteroid impact or cataclysmic volcanic events, at a rate unprecedented in Earth’s timeline since the dawn of the Cenezoic era (65 Ma), after the Cretaceous-Paleogene (k-Pg) extinction, which ended the age of the dinosaurs, occurred (Glikson, 2014). The trait becomes more evident as time’s arrow moves forward.

But this very characteristic of contemporary humans now poses a threat to society. Humans’ burning of fossil fuels releases CO\textsubscript{2} into the air at unprecedented rates. Fossil fuels are, by definition, stores of carbon captured by plants and algae over millions of years and sequestered in the Earth (Mann \textit{et al.}, 2014). The rapid release of this carbon, which took millennia to accumulate, into our atmosphere over a period that can
be measured in a century or two is expected to have profound impacts on Earth’s atmosphere (IPCC, 2014). Between 1750 and 2011, humans released $2040 \pm 320$ Gt of CO$_2$ into the atmosphere. As a result, CO$_2$, along with other greenhouse gases, have all risen exponentially since the start of the Industrial Revolution, with CO$_2$ levels rising from 280 ppm to about over 400 ppm (http://co2now.org/). The emissions, now higher than ever, have led to atmospheric concentrations of carbon dioxide, methane, and nitrous oxide which are unprecedented in the last 800,000 years (IPCC, 2014). This has lead to a warming of the climate system, with the period from 1982 to 2012 likely the warmest 30-year period of the last 1400 years in the Northern Hemisphere. The changes in temperature have had impacts on natural and managed systems in all continents and oceans (IPCC, 2014, Rosenzweig et al., 2008). Projections show that over the course of the next century, global average temperatures may rise by 4°C and sea level may rise by as much as one meter (IPCC, 2014). The oceans will also continue to acidify (IPCC, 2014). Furthermore, since the rate of consumption via combustion exceeds the rate of deposition by several orders of magnitude, the fossil fuel supply will eventually run out, with British Petroleum estimating that oil supplies will be exhausted within 53 years (Tully, 2014).

The current exploitation of fossil fuels, then, is troublesome for human society for two reasons. Firstly, the huge release of carbon dioxide into Earth’s atmosphere has caused levels to rise to the highest amounts in the last 800,000 years (IPCC, 2014). Secondly, the eventual exhaustion of transportable fuel undermines a key component of modern society that archetypically distinguishes it, from the wild, as intelligent life.
Looking back into prehistory, the world’s oil came from fossilized remains of tiny sea organisms, which include the algae (Broad, 2010, Eglinton et al., 2008, Stow, 2012). The evolutionary phylogeny of these algae (oxygenic photoautotrophic organism that do not include the streptophytes) is ultimately confounding, involving multiple endosymbiotic events (Keeling, 2013) (Figure 1).

Figure 1: General tree of life for photosynthetic organisms, with solid arrows depicting primary or secondary endosymbiotic events, and dashed arrows representing multiple events, some of which are not fully resolved.

Microalgae biomass is also seen to have potential as a biofuel feedstock (National Research Council of the National Academy, 2013, Medeiros et al., 2015). Compared to terrestrial biofuel feedstocks, algal biomass production requires less land area (National Research Council of the National Academy, 2013), with this land requirements being 1-2
orders of magnitude lower than that of land crops (Mata et al., 2010, Singh et al., 2011, Pienkos & Darzins, 2009). However, in order to scale up algal biofuel production to sufficient amounts to satisfy 5% of US demand for transportation fuel, unsustainable demands could be placed on energy, water, and nutrients (National Research Council of the National Academy, 2013). It is recognized, though, that a potential exists to shift this dynamic through improvements in biological and engineering variables (Zhu et al., 2013, Singh & Olson, 2011, Um & Kim, 2009, National Research Council of the National Academy, 2013).

In order to make biological advancements in algal biomass productivity, several aims need to be realized:

1. **Identify and Characterize new platform strains.** Within this aim, there are several related questions: Are there previously undiscovered species of microalgae that have higher growth rates then the ones already cultivated at large scales? Are there phylogenetic groups that typically have top biomass producers to serve as potential platform strains?

2. Looking toward improving these potential platform strains, there is an aim to **discern if there are genomic characteristics of these organisms** that might be related to why they are suited to biofuel production (i.e. genotype’s relationship to phenotype).

3. In line with improving production, there is also an aim to **determine if there are “pond probiotics” (algae-bacteria interactions) that could boost algal growth,** particularly relevant to growing algae in wastewater, which is seen as important to alleviating demands on water use.
Here, we focus on the development of the freshwater green algae. With regard to their biology, several wild-types of algae are currently being investigated for biofuel production, but it is doubtful that these are indeed the best species; selection has been limited, as has their modification (Rawat et al., 2011, Benemann, 2010).

Microalgae cultivation can be traced back to at least the 1950s, when investigators under the auspices of the Carnegie Institute of Washington published a report on the growth of algae in mass culture and its possible uses (Burlow, 1953). There has also been an interest in using algae as a protein source for humans, although problems with digestibility exist (Powell et al., 1961, Becker, 1984). In 1978, in response to the oil crisis, the Aquatic Species Program was launched to work on oil from algae (Sheehan et al., 1998). Over two decades, the program isolated more than 3,000 species and screened many of these for lipid production, but offered no final recommendations as to which clades of species are typically the top performers. Nevertheless, a number of cyanobacteria, diatoms, and a few green algae came out as best producers and had been tested outdoors (Sheehan et al., 1998). Since then, there are individual reports describing local scale screening projects (Rodolfi et al., 2009, Nascimento et al., 2013, Valdez-Ojeda et al., 2015), yet there seems to be little consensus or understanding as to which algae groups together are the most exploitable in the biofuels context (Griffiths & Harrison, 2009). Nevertheless, the selection of an appropriate algal strain is of utmost importance (Mutanda et al., 2011), with indigenous species of microalgae with high lipid yields especially valuable (Mutanda et al., 2011). The isolation and characterization of
algae from unique habitats could provide novel insights into the unique mechanisms that algae possess for more efficient lipid/oil production (Hu et al., 2008).

A review of the literature examined 55 different species of microalgae in terms of growth rates, lipid content, and lipid productivities (Griffiths & Harrison, 2009). The survey consisted of a wide swath of phylogenetically diverse organisms from the Chlorophyta, Bacillariophyta, and Cyanobacteria. Although the review showed that certain species like *Chlorella pyrenoidosa* have fast doubling times, the authors concluded that problems with interpreting the research lie in that experiments were conducted under different conditions, with different equipment, protocols, and media. Thus, it is difficult to interpret which clade bears strains suited well to biofuel production.

This large scale screening described in this thesis seeks to overcome this problem. Also, as part of the platform strain characterization, we investigated if there were particular genomic characteristics, related to growth, of one of the best growing strains, a *Scenedesmus obliquus*. In particular, we analyzed the number and cellular localization of the enolase enzyme, which plays an important role in carbon core metabolism.

Once the best strains are selected, genetic improvement techniques can be employed to improve these strain’s productivity. But the culture productivity may also be improved with bacteria pro-biotics. Co-immobilizing of microalgae with growth promoting bacteria, such as *Azospirillum brasilense*, has been shown to increase ammonium and phosphorus ion removal from wastewater (de-Bashan et al., 2002, de-Bashan et al., 2004). Similarly, to alleviate CO$_2$ limitation in large ponds, bacteria may also be used. Because bacteria respire CO$_2$, they may release it with certain organic inputs. Finding a bacteria that quickly respires certain organic inputs, releasing CO$_2$, may
then maximize the growth rate and biomass production of a phototrophe (Schieffer & Caldwell, 1982). However, it is important that these bacteria do not inhibit the growth of algae (He et al., 2013). In the last chapter of this thesis, I detail the possibility of using a plant bacterium – *Agrobacterium* – with algae to bring about mutualistic gas exchange as they syntrophically metabolize human inputs.

The overall goal of this thesis is to push forward the field of algae biomass and biofuel production. The objectives are to discover microalgae strains that are more productive, gain insights into their phylogenetic relationships, discern clues into genetic pathways of carbon utilization and growth, and discover bacteria that can increase algae’s resource utilization and growth.

Though advancing algae productivity will be challenging, this work identifies species of wild algae that demonstrate characteristics already amiable to biofuel production. In comparison to the ‘gold-standard’, or benchmark, for biomass and lipid productivity comparison of newly isolated strains, *Nannochloropsis salina* strain CCMP1776, the newly discovered strains described in this work grow well and, under stress, accumulate lipids and high value products. This growth is exhibited under a wide variety of conditions in the laboratory as well as outdoors. They appear, then, prime candidates for further study of stress biology, with the hopes of harvesting their bioproducts. Finding ways to increase growth rates will also be essential. A group of bacteria that has positive effects on algae growth, with an organic input, is also identified. It is hoped that this research may contribute positively to algal biomass utilization and a more carbon neutral economy.
Chapter 1: Microalgae Strain Isolation, Screening, and Identification for Biofuels and High Value Products

Neofotis, P.\textsuperscript{a,b}, Huang, A.\textsuperscript{a}, Chang, W.\textsuperscript{a}, Joseph, F.\textsuperscript{c}, and Polle, J.E.W.\textsuperscript{a,b}

\textsuperscript{a} Department of Biology, Brooklyn College, 2900 Bedford Avenue, Brooklyn, New York 11210, USA
\textsuperscript{b} The Graduate Center of the City University of New York, New York, NY 10016, USA
\textsuperscript{c} Department of Cell Biology, State University of New York Downtown Medical Center, Brooklyn, NY 11203, USA

Keywords: algae, biofuel, screening, bioprospecting, isolation, carotenoids

Corresponding Author: Dr. Jürgen E. W. Polle, jpolle@brooklyn.cuny.edu

ABSTRACT

Microalgae have been investigated for production of biofuels and higher value bioproducts since the 1950s, but only a few species of microalgae are currently produced commercially. The great diversity of so far untapped microalgal species found in nature holds promise for future practical applications of microalgae biotechnology. Bioprospecting is required to identify novel strains with desired product profiles that could be cultivated in large-scale systems. Here, we describe such a bioprospecting effort specifically geared to biofuels but also applicable to other bioproducts.

1. INTRODUCTION

Microalgae, as yet a mostly untapped biological resource, could be cultivated on non-arable land and utilize industrial flue gas as a carbon source for biofuel production (Gouveia & Oliveira, 2009, U.S. Department of Energy, 2010, Borowitzka & Moheimani, 2013, Passell et al., 2013, Wijffels et al., 2013, Benemann, 2013). With more than 100,000 species described (Guiry et al., 2014), eukaryotic microalgae are classified into several major lineages, which are more widely separated genetically than humans are from fungi (Lemieux et al., 2014, Chan et al., 2011, Cavalier-Smith, 2013, Keeling, 2013). Further, with the classification of even many known species tenuous, it is unknown how many microalgae species may exist (Krienitz & Bock, 2012, Leliaert et al., 2014). Although this chapter focuses on eukaryotic microalgae, the prokaryotic microalgae belonging to the cyanobacteria, formerly called blue-green algae, also display great diversity (Palinska & Surosz, 2014).

Only about 15 of the currently known microalgae species are cultivated at any significant scale and these are used almost exclusively for human food supplements (“nutraceuticals”), cosmetics, and aquaculture feeds (Raja et al., 2008, Borowitzka, 2013). Examples of green microalgae already cultivated commercially belong to the freshwater genera *Chlorella* (Doucha & Livansky, 2006) and *Haematococcus* (for astaxanthin production) (Guerin et al., 2003, Laurens et al., 2012). In addition, the green algal species *Dunaliella salina* is grown in large hyper-saline lagoons and ponds for provitamin A production (Borowitzka, 1991, Ben-Amotz et al., 2009). Furthermore, the cyanobacterium *Arthrospira sp.*, commonly known as *Spirulina*, is also cultivated – essentially exclusively – in raceway type ponds, which are mixed by paddle wheels (Elliott et al., 2015). Worldwide autotrophic (on sunlight and CO₂) microalgae
production is currently no more than 20,000 tons globally, with about the same amount produced heterotrophically (Benemann, 2013). The diversity of microalgae has attracted considerable interest in the large-scale production of biofuels (Elliott et al., 2012a, Ratha & Prasanna, 2012, Ratha et al., 2012), animal feeds (Benemann, 2013), and higher value products such as nutraceuticals (Borowitzka, 2013, Draaisma et al., 2013, Wijffels et al., 2013).

Bioprospecting – the identification of superior, “platform” strains, which exhibit high growth rates and culture stability, and create desired product – is required to realize the potential of algal biotechnology (Davis et al., 2011, Mutanda et al., 2011, Barclay & Apt, 2013). These strains could then be developed further through genetic improvements to achieve cost and production goals (Georgianna & Mayfield, 2012, Gimpel et al., 2013, Krienitz et al., 2015, Kaufnerova & Elias, 2013, Hegwald et al., 2010). The study of these strains may also provide knowledge that can be used to improve other strains.

Microalgal bioprospecting for novel strains for production of biofuels was first carried out systematically during the 1980s U.S Department of Energy’s ‘Aquatic Species Program’ (Sheehan et al., 1998). It was, and remains, a very labor-intensive undertaking, even with the newer high-throughput automated screening technologies now being applied in this field.

This chapter describes our process of isolation, screening, identification, characterization, and maintenance of microalgal strains. Although the focus was on obtaining strains suited for producing biofuels, many strains discovered also have potential applications in the future development of foods, feeds, and high-value products (for example: carotenoids and long chain polyunsaturated omega-3 fatty acids).
2. METHODS

2.1 Strain Isolation

As outlined in Figure 1, microalgal strains can either be obtained from existing culture collections or be isolated from environmental samples. Our approach employed environmental sampling from different habitats followed by strain isolation. Strains from culture collections with known characteristics were used only in our screens as benchmarks to assess the performance of the newly isolated strains. In the following text, we describe the different steps involved in our approach to isolate and screen novel microalgal strains.
Unialgal Strain Storage

Lab Lipid Screening

Small Scale Liquid Culture

Nile Red Lipid Screening

Lipid Screening Data

Proxys: OD750 for Growth
Nile Red Signal for Lipid Accumulation

Potential Producer

Carotenogenic Pigmentation

Further Analysis in the Lab

2nd Level Screen (Multi Media)

3rd Level Screen Bubbling Column Culture

Carboy Culture

DNA for Molecular Work

Testbed
Figure 2: An overview for the bioprospecting approach from sampling to screening and testing of the strains is shown. Samples were collected from various habitats, aliquots plated, colonies picked, transferred into liquid medium and then cell-sorted onto agar-media culture plates. Strains were screened by growth in liquid batch cultures. Strains showing good growth characteristics and/or high Nile Red fluorescence were deemed candidates and grown in bubble-columns for biomass productivity measurements.

**Sampling**

More than 1,000 environmental samples were collected over a period of approximately five years (2007-2011) from a variety of habitats across the United States, including both dry (e.g. soil, lichen) and wet (e.g. fresh, brackish and saltwater) samples. To isolate both summer and winter strains, sampling was performed in all seasons. When possible, multiple samples were taken from the same habitat at different times to also cover annual succession of algae in a location. Examples of sampling habitats are shown in Figure 2. The sampling efforts focused on California, New Mexico, and Texas, which are considered prime potential future sites for large-scale algal operations (Maxwell *et al.*, 1985, Sheehan *et al.*, 1998, Pate *et al.*, 2011). Environments sampled included shallow and temporary aqueous habitats, often with extreme temperature variations, and exposure to full sunlight. Freshwater habitats included birdbaths, roadside ditches, fountains, shallow ponds, creeks, rivers, and lakes. Saline water habitats included estuaries, inland salt lakes, and general coastal waters in Southern Texas, New York, and Connecticut.

Whole water (rather than concentrated) samples were collected and stored in clean containers until plating, as they have been found to provide viable cells when concentrated samples fail (Andersen & Kawachi, 2005). As natural samples often contain zooplankton that feed upon algae, water samples were passed through 70 µm filters.
Often, even some predators – such as amoeba or small ciliates – remained. The filters also, unfortunately, removed algae that form large coenobia.
Figure 3: Photos showing representative habitats of the wide range of environments from which algae strains could be isolated. Left panel from top to bottom - Freshwater algal bloom in a roadside ditch in Texas; Microalgal bloom in the Brooklyn College freshwater pond in Brooklyn, NY; Pacific Ocean in San Diego, CA; Diverse lichens on a rock in Arizona. Right Panel from top to bottom - Freshwater birdbath in San Diego, CA; Diatom bloom in the Salton Sea, CA; Green algae of the genus *Dunaliella* blooming in a brine pond adjacent to the Great Salt Lake in Utah; Biofilm of microalgae at a roadside location in Louisiana.

When possible, water samples were first inspected via microscopy to determine what species might be isolated from the environmental samples. Figure 3 shows an example of the diverse microalgae seen in a freshwater sample. Though it was very time consuming and could only be performed on a fraction of the samples, microscopic analysis allowed us to understand what species were lost during the isolation process. It was noted that colonial green algae of the genera *Volvox* and *Pleodorina* were observed in many samples by microscopy, but were not isolated with our approach. In contrast, many filamentous green algae present in the samples could be isolated.

Figure 4: Light microscopic image of an environmental freshwater sample originating from the Brooklyn College pond in Brooklyn, NY. Based on morphological characters the sample contained green algae such as *Desmodesmus, Ankistrodesmus, Kirchneriella,*
Chlamydomonas, and Chlorella type cells as well as a diatom species. In addition, there were some unknown algal species visible. The black bar represents 10 µm.

**Media and Plating**

We observed that viable algal cells were still present in many original water samples that had been kept for several years in the laboratory. Nevertheless, since many kinds of algae, unless maintained and properly handled, may die from anywhere between a few hours to a few days after collection, we plated aliquots of samples as soon as possible on solidified growth media (1.5 % Agar) (Figure 4). On longer sampling trips, prepared Petri-dishes allowed plating of aliquots of the environmental samples within minutes or hours. Once colonies developed, the plates were then shipped back to the laboratory for further processing.

![Algal colonies grown on agar-solid media](image)

Figure 5: Algal colonies grown on agar-solid media. Left Panel - Aliquots of the same environmental freshwater sample were placed on agar-solid media in regular standard 90 mm Petri dishes. Three different growth media as indicated on the top left of each plate were used supplemented with silicate to allow for growth of diatoms. Right Panel - An example of algal colonies grown on agar-solidified medium shown with a magnification of 60x.

With current technology, complete recovery of all microalgal species present within a given environmental sample is not possible, as many freshwater and marine algal strains do not produce colonies on agar plates, despite this being a well-established method (Andersen & Kawachi, 2005). Yet, plating on agar-solidified media still allows
isolation of large numbers of clonal strains (deriving from a single original cell) in situations where newer high-throughput techniques are not available. Another advantage of this technique is that during plating algal grazers and most infectious agents are removed.

Inoculated plates were sealed with paraffin film to slow drying and were then incubated at about 50 µE m\(^{-2}\) s\(^{-1}\) continuous light provided by daylight fluorescent lamps. As some algae require day/night cycles and will not grow, or grow poorly, in continuous light (Andersen & Kawachi, 2005), this introduced another selection. For some diatoms under continuous lighting on the plates, a ring-growth-pattern, due to alternating growth and resting stages, was observed. Although continuous illumination reduced the diversity of algae isolated, many microalgae, including some of those currently mass-cultivated species such as species of the green algae *Chlorella* and *Dunaliella*, as well as the cyanobacterium *Arthrospira*, do not require day/night cycles. Continuous illumination also achieved rapid algal growth and saved time in the strain isolation process. Ultimately, our isolation and screening procedure resulted in new strains of microalgae that could be cultivated successfully in the greenhouse of Brooklyn College and in outdoor ponds (NAABB Final Report, 2014), thus validating this isolation procedure.

In addition to plating aliquots of liquid samples, initially a transfer through a liquid culture for enrichment of fast growing cells, followed by plating, was also tested. But this resulted in the isolation of many essentially identical strains and thus was discontinued with only the traditional plating method used as the primary isolation tool.

Subaerial and aerial algae were isolated via enrichment cultures as recommended (Andersen & Kawachi, 2005). In brief, soil and lichen samples were first placed into
liquid growth medium. Generally, after two weeks some algal growth was observed, as indicated by coloration of the liquid medium. An aliquot was then plated and colonies of algae were then observed on the plates after about another two weeks. Samples derived from lichen often contained cyanobacteria.

The culture medium was a critical factor in the isolation of microalgae from natural sources. Because this screen was designed primarily for eukaryotic algae, which grow photoautotrophically, freshwater samples were plated on several suitable known minimal media: C Medium, Bold Basal Medium, and BG11 Medium (Andersen & Kawachi, 2005) (see Figure 4). These media contain nitrogen in the form of nitrate. As it contains ammonium as the fixed nitrogen source for growth, HS medium (Harris, 2009) was also used initially, as was Tris-Acetate-Phosphate (TAP) medium (Harris, 2009), to allow for the isolation of strains that would be selected for preferential photoheterotrophic growth on acetate. However, the latter was abandoned early on due to overgrowth of bacteria and fungi, which often displaced the microalgal colonies. Marine strains were placed in F/2 media (Guillard & Ryther, 1962, Guillard, 1975) and artificial seawater media (Keller et al., 1987). Vitamins were not used for any of the freshwater media, and only used initially for the F/2 media, as we were screening for strains suitable for large-scale cultivation, and vitamins can be prohibitively expensive at such scales. To allow for growth of diatoms, silica was added to all the different media to a final concentration of 0.4 mM. To isolate strains from inland highly saline water samples, a minimum medium known to allow the growth of green algae of the genus of Dunaliella, as well as of diatoms, and cyanobacteria, was used (Pick et al., 1986), as we found that many such algae (e.g. from hyper-saline soils) did not grow in F/2 (Sheehan et al., 1998).
Although some cyanobacteria can form lipid droplets (Peramuna & Summers, 2014), in general, cyanobacteria are not known to accumulate significant amounts of lipids as triacylglycerides (Hu et al., 2008). Due to their expected low triacylglyceride content, in our initial screen cyanobacteria were down-selected in this biofuels focused project based on the blue-green color of colonies on the plates during the visual inspection. Still, a number of cyanobacteria were recovered later, due to not all colonies of cyanobacteria having the more typical vibrant blue-green color.

Figure 4 shows examples of colonies on plates. This comparison of growth on plates demonstrates that using different media increased overall diversity of strains recovered from one water sample. Colonies for transfer were chosen based on differences in colony color and overall appearance, to increase the number of unique types of algal strains thus isolated. Analysis of the number of isolated strains from over 300 samples showed that most environmental samples yielded on average five unique strains, typically ranging from 1 to 10 (Figure 5).
Figure 6: The frequency of the number of microalgae strains isolated per environmental sample from a total of 350 environmental samples. The x axis indicates the number of isolates per sample, and the y axis indicates the frequency of such isolates.

Flow Cytometry Cell Sorting

In recent years, flow cytometry – with its single-cell sorting capability – has become a method of choice for high-throughput isolation of algal strains (Sieracki et al., 2005, Pereira et al., 2011, Hyka et al., 2013). In Fluorescence Aided Cell Sorting (FACS), flow cytometry is coupled with detection of fluorescence of cells (Shapiro, 2004). As algal cells have pigments such as chlorophylls and carotenoids, fluorescence of these pigments can be easily detected and thus be used as sorting criteria for algal cell isolation. FACS can be used at different stages in algal strain isolation. For example, sorting of cells can be performed directly from water samples (Figure 1) either based on pigment fluorescence or based on staining with marker dyes, such as BODIPY or Nile Red for oil bodies (Cooper et al., 2010, Montero et al., 2011, Cirulis et al., 2012, Elliott et al., 2012b).

In comparing the number of strains obtained using the plating technique to the number that successfully passed the cell sorting process, we found that about 1/3 of all strains isolated by plating did not survive the flow cytometry process. Therefore, we decided to continue with a strain isolation approach where first traditional isolation methods were utilized, as described above, followed by high-throughput FACS cell sorting.

In our combined approach, following the development of colonies on sampling plates (Figure 1; Figure 4), individual colonies were picked and transferred to index plates (Figure 1). On these index plates, the colonies were allowed to grow further,
accumulating biomass for 2-4 weeks, before being transferred into liquid culture. In the liquid media transfer stage, the strains were grown in culture tubes for about two to four weeks before cell sorting. Growing cells in liquid media allowed determination of whether or not the algal strain grew well in liquid media, an essential characteristic for large scale cultivation in outdoor raceway ponds, shallow ponds containing one channel in the shape of an oval, like an automotive raceway circuit. The 2-4 week growth period was required, as cells from shorter incubation periods did not grow well after cell sorting indicating that the growth phase of cells impacts cell survival. Prior to sorting, 2 mL aliquots of the liquid cultures were filtered through 70 µm or 100 µm filters, to remove cell clumps or large colonies from the liquid samples.

The filtered algal suspension was then cell sorted using a FACS Aria System II cell sorter from BD Biosciences. The nozzle size was generally set to 100 µm or to 120 µm. If larger cells or colonies (“coenobia”) of cells were present, a 120 µm nozzle was used. Pigments in algal cells were excited using a 488 nm Blue Argon laser and, based on the chlorophyll fluorescence signal (695 nm) and the carotenoid fluorescence (530 nm), the algae were sorted onto agar-solidified growth media. In addition, we concomitantly monitored the forward and side scatter. Figure 6 shows an example of clusters of events in plots of chlorophyll fluorescence at 695/20 nm against the carotenoid fluorescence at 530/20 nm obtained from the FACS. Clusters of events indicate populations of cells often consisting of cells from a single original cell (or clone, e.g. strain). Events can be single cells, clusters of cells from one original cell, clusters of cells from multiple clones, and/or algal cells or clusters with attached bacteria.
Figure 7: An example scatter plot of a mixed sample, with the dots representing single events as visualized by the FACS Aria System II cell sorter. The different populations were selected and single events sorted onto agar-solidified media in Petri dishes. The parameter chlorophyll fluorescence at 695 nm was used to select populations of algal cells via cell sorting. Bacterial cells may display chlorophyll fluorescence, but were usually at values below $10^1$ relative fluorescence.

In the FACS experiments, populations were initially sorted into liquid media in 96-well plates, and later onto agar-solidified media, with generally higher survival rates in the liquid media than on the agar plates. We successfully recovered sorted cells of the fragile cell wall-less green alga *Dunaliella salina*, grown in medium containing 1.0 M NaCl (about three-fold the salinity of seawater). Figure 7 shows a 96-well plate with cultures of *D. salina* growing in a number of wells. Only a minority of sorted cells grew up to form visible cultures, indicating that cell survival was low. To achieve even this low survival rate, the cell sorting process had to be modified, including use of a large 120 µm nozzle, low flow rate, and increasing the salinity of the sheath fluid to 1.0 M NaCl.
Higher salinities led to clogging of the fluid lines and prevented sorting. Nevertheless, this example of *D. salina* demonstrates that once the FACS sorting conditions are found that allow for survival, even very fragile cells can be FACS cell sorted. That is however, not always a simple matter, and FACS may lose its advantage for fast isolation where isolation conditions are not standard.

Figure 7 shows that 96-well plates also had some major limitations. Uneven gas exchange and evaporation rates, as well as condensation, led to reduced growth of cultures in the wells at the plate center (Figure 7B). In the future, use of membrane covers rather than regular plastic lids might alleviate such issues. Cells also settled in the wells and growth could not be readily followed. More importantly, only by sorting onto plates could colonies be easily determined as unialgal.
Figure 8: FACS cell sorting of events from liquid cultures into 96-well plates. A) Sorted here was the green alga *Dunaliella salina* strain CCAP19/18. Cells of *D. salina* are fragile, as they do not have cell walls. Furthermore, the sorted culture was grown in 1.0 M NaCl medium. Nevertheless, it was possible to sort cells and recover viable cultures in liquid medium. In contrast, it was not possible to sort cells of *D. salina* onto regular plates with solidified medium. B) Sorted here were populations of a new, unknown, and carotenogenic isolate. All wells contained cultures, indicating that the sorting process had a high efficiency and that the cells survived sorting. On this plate, it is evident that the conditions in the wells were not uniform. Cultures in the outer wells had already turned orange indicating that they were under stress, whereas the cultures in the inner wells were still green. Also, a number of the inner wells show heavy condensation on the cover plate.
As shown in Figure 8, sorting cells onto agar-solidified media allowed for easy examination by microscopy of colonies to identify unialgal colonies. Overall, about 5 - 10 % of all colonies did not appear to have originated from a single algal cell or coenobium. Figure 8 shows examples of colonies obtained which appeared axenic. Lack of visible bacterial contamination on agar plates, however, is no assurance of axenic cultures. Figure 8 also shows algal colonies that were not unialgal and/or had rings of white or slightly pinkish color around them indicating heavy bacterial contamination. Different unialgal colonies were picked from each plate and transferred for storage into tubes with agar-solidified medium overlaid with liquid growth medium. Strains were not made axenic on a routine basis, because the process is too time and labor consuming.

Figure 9: Examples of microalgal colonies growing on agar-solidified media in Petri dishes after FACS cell sorting. A) Plate from a non-unialgal liquid culture. Although one strain seemed to dominate, several different colony colors and types indicate the presence of at least 4 different species. B) One unialgal strain only, with most colonies being
axenic. C-K) 60x magnification images of algal colonies sorted from a single event that are: C-E) unialgal and axenic; F-H) non-unialgal; I-K) containing bacteria as indicated by the white, beige, or orange rings around the colonies.

In summary, our combined approach of traditional cell culture followed by FACS cell sorting allowed us to isolate approximately 1,000 strains per year, a several-fold higher yield than if just the traditional isolation approach involving plating and streaking had been used.

2.2 Strain Maintenance

Following isolation as described above, strains were assembled into a catalogued culture collection. The strains were stored in glass test tubes, which contained growth media as a solidified 1.0 - 1.5% agar slant, overlaid with liquid growth medium. However, sometimes the growth medium was changed based on further growth results, such as the “multi-media” screen described below.

The storage tubes were capped and sealed with plastic paraffin film. They were then maintained with a 12 hour day-night cycle, under 50 µmol photons m$^{-2}$ s$^{-1}$ light, though in some cases less light was more optimal. Strains were stored at room temperature for at least 12 months. After one year most strains had to be transferred to new tubes with fresh medium, although a number of strains could be stored up to two years without transfer. Nevertheless, for longer-term storage without such maintenance, cryopreservation should be considered (Benson, 2008). In that regard, 30 of the best performing strains isolated were transferred to the UTEX culture collection for cryopreservation (NAABB Final Report, 2014).

2.3 Classification of Microalgal Strains
Working with thousands of isolates, any comprehensive molecular classification effort including DNA isolation, PCR, sequencing, and sequence analysis would have been time consuming and costly. Consequently, newly isolated algal strains were classified only after they passed at least the first screening step.

The most promising strains obtained from the second-level screen were subjected to molecular classification. Sequences of nuclear rDNA Internal Transcribed Spacer 2 Regions (ITS2), which usually are about 300 base pairs long, in algal strains have been used frequently for classification and barcoding, because the ITS2 has been shown to be an excellent phylogenetic marker (Koetschan et al., 2010). It has been used for species identification in environmental samples (Landis & Gargas, 2007, Engelmann et al., 2009) and as a target molecule for barcoding diatoms (Moniz & Kaczmarska, 2009). With regard to the green algae, the ITS2 has been used for phylogenies of orders and families within the green algal class Chlorophyceae (Hegewald et al., 2010, Buchheim et al., 2012) as well as in automated reconstruction of the green algae tree of life (Buchheim et al., 2011). Recently, a large downloadable database set of sequences became available that allows for alignment with and comparison to a large number of ITS2 sequences for green algae (http://its2.bioapps.biozentrum.uni-wuerzburg.de/).

For strain classification and bar coding, the DNA of the candidate microalgae was extracted using a MO-BIO DNA Isolation Kit (USA). The PCR products were sent out for purification and sequencing to Eurofins (MWG Operon now Eurofins Genomics, AL, USA). Amplification of the ITS2 regions can be achieved by PCR using universal ITS2 Primers (forward primer ITS2-F GCATCGATGAAGAACGCAGC and reverse primer ITS2-R TCCTCCGCTTATTGATATGC) (White et al., 1990). We also conducted
microscopic analysis of cell morphology. Phylogenetic analysis of the best performing strains from our bioprospecting research is underway (Neofotis et al., In revision).

2.4 Screening for Biomass Productivity and Lipid Content

Following isolation, strains were subjected to a multi-level laboratory screening protocol. The screen was designed to identify strains with the highest lipid productivities. Lipid productivity is defined as the product of biomass productivity times the lipid content of the biomass. The two main current proxies used for lipid content in algae, other than actual chemical analysis, are fluorescent dyes (Cooksey et al., 1987, Cooper et al., 2010) and, at least for green algae, their capacity to over-accumulate carotenoids (Goodwin, 1952). Figure 1 shows the general pipeline of the screen, described in more detail below.

First-Level Screening

First, a high-throughput screening process using the lipophilic fluorescent dye Nile Red (9-diethylamino-5H-benzo[a]phenoxazine-5-one) was used to qualitatively estimate the relative cellular content of lipids (Cooksey et al., 1987). Another lipid dye, BODIPY (boron-dipyrromethane), could not be used for our assay because its background fluorescence was too high. In a study looking at the optimization of staining conditions for microalgae with three lipophilic dyes (Nile Red, BODIPY, and DiO) to reduce precipitation and fluorescence variability, Nile Red was found to have a stable fluorescence intensity that was unaffected by the broadest range of conditions and could be correlated to cellular lipid content (Cirulis et al., 2012).
Because of the above noted limitations of 96 well plates and even 12-well or six-well plates, ultimately 120 batch cultures of unialgal strains were grown in their respective liquid media types in Erlenmeyer flasks in batch culture on a shaker (Innova Platform) at room temperature under continuous illumination of about 50 µmol photons m$^{-2}$ s$^{-1}$ provided by daylight fluorescent lighting (Figure 9). The cultures were then screened for their Nile Red fluorescence and optical density (OD$_{750}$) at weekly intervals. Depending on the strain, these batch cultures grew for about three to five weeks.

![Figure 10: Screening of strains of microalgae for growth and lipid accumulation was performed in batch cultures using Erlenmeyer flasks on a shaker. A) Shown is an example for the first-level screen of growth of microalgal cultures in batch cultures on a shaker. B) The strain shown was isolated from a C-Medium plate, but grew best in liquid BG11 medium. C) The strain was isolated from a BG11 medium plate, but grew well on all three media tested. D) The strain was isolated from a BG11 medium plate, but best growth occurred in liquid BBM medium. Note the clumpy growth of this strain in C-Medium and BG11 medium.](image)

For the weekly measurements of growth and relative lipid content, aliquots of 200 µl of algal culture were taken from the Erlenmeyer flasks and placed into wells of black, clear-bottom, surface non-binding 96-well polystyrene plates (Corning). A set of twelve
strains was tested at a time. For each strain, six wells were used. If necessary, the 200 µl aliquots of the samples were diluted to an OD$_{750}$ of 0.1 to about 0.3 to ensure consistency in measurements. Into three of the six wells for each strain, Nile Red was added to a 3.0 µM final concentration. The other three wells, without Nile Red, were the negative controls. Six wells of the plate were used as negative controls with only medium in them. As standard controls for comparison, six wells were also loaded with aliquots of a culture of the unicellular green alga *Chlamydomonas reinhardtii* strain CW15 in TAP Medium, a low lipid producer with known characteristics in this protocol.

After allowing for 15 minutes for the dye to penetrate intracellular bodies and bind to the cellular lipids, the assay plate was screened with the BioTek Synergy 2 microplate reader with an excitation at 485/20 nm and emission at 590/35 nm for Nile Red fluorescence and absorbance at 750 nm to assess optical density. The OD$_{750}$ was used as a proxy for biomass accumulation with its values representing the growth of the culture. The level of fluorescence of Nile Red at 590 nm indicated the level of lipids in a culture (Cooksey et al., 1987). Nile Red is only suitable for quantification of lipids with saturated fatty acids and thus strains accumulating larger amounts of lipids with polyunsaturated fatty acids may not have been detected in our screen. Then the Nile Red fluorescence values were normalized to the biomass as measured by the OD$_{750}$. After 4-5 weeks, progressive optical density values and normalized Nile Red readings per unit optical density were plotted against time. As the bar for the low lipid producing strain (Figure 10), *C. reinhardtii*’s fluorescence was compared to that of newly isolated strains. The cut-off for selecting novel strains for further study was either a Nile Red fluorescence above 60,000 relative units (compared to 5,000 for *C. reinhardtii*), and/or an
exceptional growth based on the OD$_{750}$. Overall, about 20% of all isolated and screened strains passed the first-level screen. One of the first strains identified with excellent growth and lipid accumulation properties was a strain of *Scenedesmus (Acutodesmus) obliquus* (EN0004) (Neofotis et al., In revision) (Figure 10 and Figure 11).

Figure 11: Normalized Nile Red fluorescence representing the lipid contents of cells. The Nile Red signal is shown for three strains during the screening period. Error bars represent standard deviation among triplicates. Although strain EN1234 is carotenogenic and over-accumulates triacylglycerols in cytosolic oil bodies (see Figure 11E and 11F), it did not display high relative Nile Red values in our first-level screen. Low Nile Red fluorescence levels for carotenogenic strains were often seen in the first-level screen. This may be due, in part, to thick cell walls prohibiting efficient penetration of Nile Red into the cells. In addition, the carotenoids within the oil bodies might quench Nile Red fluorescence. Although such carotenogenic strains did not pass the Nile Red screen, their orange and reddish culture colors indicated triacylglycerol accumulation and so were advanced to the second level of screening.
Figure 12: Cells from two of the strains deemed to be potential high-producers, showing bright-field (A, C, E) and Nile Red fluorescence images from the same pairwise field of view at 530 nm (B, D, F). A-B) *Scenedesmus obliquus* strain EN0004 with oil bodies visible with fluorescence microscopy. C-D) Strain EN1234 (a putative *Coelastrum*-related Desmid) from a young culture with few significant oil bodies visible. E-F) Stressed cells of strain EN1234 taken from an older culture. Note the orange coloration under bright field indicative of carotenoid accumulation (E) and apparent Nile Red fluorescence (F). The bar indicates 10 µm.

In our experience, though many strains grow well in HS medium, containing ammonium as the only nitrogen source, none of the hundreds of strains examined in HS
medium accumulated lipids. Indeed, subsequent tests with several strains, each grown in multiple media, showed that algae grown in HS medium did not accumulate lipids. When grown on nitrate containing C-medium, BGII, and BBM media, these same strains accumulated lipids. We also noticed that many newly isolated *Haematococcus* strains grew well in liquid HS media, but never accumulated the characteristic red astaxanthin carotenoid. Based on these observations, use of HS media was discontinued.

**Screening for Carotenoid Over-accumulation**

As a complementary method of screening (Figure 1), all cultures of strains were also examined for changes in color. In particular, yellowish, orange, and reddish cultures, indicating over-accumulation of carotenoids (Figure 11, Figure 12, and Figure 13), were further inspected through bright field and Nile Red fluorescence microscopy (Olympus BX51). Algae strains observed to accumulate carotenoids were presumed to also have the capability to produce significant levels of triacylglycerols, due to the known sequestration of carotenoids in oil bodies. Cells of one carotenogenic strain (EN1234), a *Coelastrum*-related species of the family of the Scenedesmaceae within the class Chlorophyceae of the green algae (Neofotis et al., In revision), are shown in Figure 11 as an example. Staining of cells with Nile Red or BODIPY (Cooper et al., 2010) confirmed the location of carotenoids within triacylglycerol-containing cytosolic oil bodies.
Figure 13: Identification of strains that over-accumulate carotenoids. A) Sometimes colonies of microalgae grown on plates would display from a yellowish, orange, or reddish color indicating that cells over-accumulated carotenoids. B) More often strains of liquid cultures from the first-level screen would display yellowish, orange, or reddish color upon aging. C) Shown is a culture displaying an apricot color indicating accumulation of carotenoids. D) Microscopy of cells from culture shown in C). The cells have apricot color originating from very large oil bodies that sequester the carotenoids (data not shown). The bar represents 10 µm.

Sometimes, based on the colony color on the initial plates from environmental samples, strains could already be identified as carotenogenic (Figure 12A). However, in batch cultures, the majority of strains identified as carotenoid over-accumulating only
exhibited such characteristics in the first-level screen when they reached the stationary phase (Figure 12B and 11C).

Currently, *D. salina* and *H. pluvialis* are the only species cultivated on commercial scale for carotenoid production. However, we isolated various strains of other green algae genera able to accumulate carotenoids (Neofotis et al., In revision). The pigment profiles of two of these strains, EN1234 and *H. pluvialis* strain Haema001 are compared in Figure 13. To obtain the profiles, dichloromethane/methanol 2:1 (v/v) was used to extract the pigments. The extracts were then separated using a C18 reverse phase column during Ultra High-Performance Liquid Chromatography (UHPLC). The extracts were then run through a Photo Diode Array detector (Flexar FX-15 UHPLC system from Perkin Elmer). Different carotenoids within the spectrum were identified based on absorbance spectrum and based on concomitant molecular mass determination by application of mass spectrometry using Atmospheric-Pressure Chemical Ionization (APCI) in connection with a Time-of-Flight Mass Spectrometer (Perkin Elmer, MA, USA). Similar to the green alga *H. pluvialis*, strain EN1234 contained astaxanthin, although it was only a minor component of its total carotenoids. Further analysis of the carotenoid profiles of strain EN1234 and some other carotenoid-over-accumulating strains showed that they accumulate pre-cursors of astaxanthin in different ratios as well as possibly novel carotenoids that may be of commercial interest for feed formulations or nutraceuticals (Neofotis et al., In revision).
Figure 14: Displayed are UHPLC traces that show the absorbance of pigments from cell extracts plotted as the 450nm/550nm difference. Different pigments such as chlorophylls and carotenoids elute at known times and the relative level of absorbance represents the amounts of these pigments present in the strains. Upper panel – Profiles of extracts from strain EN1234. Lower Panel – Profiles of extracts from carotenoid over-accumulating cells of strain EN1234 and Haematococcus sp. strain Haema001.

Second-level Strain Screening

Strains that had passed the first-level screen were then subjected to a second-level screen to determine their growth in media other than the initial isolation medium. For freshwater strains, BG11, BBM, and C-Medium were all used as these media have very different nutrient concentrations (i.e. nitrate, phosphate, and iron). Color and density of algal cultures were inspected visually (Figure 9) to estimate the growth of the cultures. In many cases, strains had a broad range of media that they would grow in (example shown in Figure 9C). But often one medium was preferred by a strain (examples in Figure 9B and 9D). Based on the results obtained from this ‘multi-media’ screen, the storage
medium for strains was adjusted to the liquid medium that the strains grew in best. The best growth medium as determined in the second-level “multi-media” screen was then used in the third-level screening.

**Third-level Screening for Biomass and Lipid Productivities of Strains of Microalgae**

Potential future platform strains that had undergone the multi-media screen were then tested for their biomass and lipid productivities in bubble-columns (500 mL, 3 cm inner diameter) at 28 °C under continuous illumination with fluorescent daylight lamps of about 200 μmol photons m$^{-2}$ s$^{-1}$ (Figure 14). So that the cultures would not have CO$_2$ limitations or major pH changes, CO$_2$ was provided to cultures at a concentration of 1% in air. This allowed, in a short period of time, the production of enough biomass for gravimetric and lipid content analysis. This third-level screen evaluated the growth potential of the strains under conditions of sufficient CO$_2$, constant pH, and no O$_2$ accumulation, in contrast to the CO$_2$ limiting and uncontrolled pH, and potentially O$_2$ inhibitory conditions present in the second level screening. Strains that grow well under both conditions are of potential interest in process scale-up.

![Figure 15: Shown are bubble-columns (numbered 1 to 7 from left to right) with each containing a unique, newly isolated strain. *Scenedesmus obliquus* strain EN0004 was grown in the leftmost column (column 1). The sequence of photos visualizes the different growth rates of the various strains. Photos were taken 24 hours, 48 hours, and 120 hours after culture inoculation.](image-url)
Figure 16: Plotted are growth curves based on biomass for the strains shown in Figure 14. Some strains such as #5 did not show any growth inhibition during our experiment and could reach biomass concentrations of more than 5 g/L.

When pre-cultures were in their light-limited growth phase, they were used to start new batch cultures (Figure 14) from which growth curves based on ash-free-dry-weight (AFDW) were obtained (Figure 15). Figure 14 shows the succession of cultures of different newly isolated strains in each one of the bubble-columns. As visible from the photos of the cultures, some, but far from all strains, grew well under these laboratory conditions. Overall, more than 100 freshwater and brackish/marine strains were tested in bubble-columns. More than 50 of these newly isolated strains grew well. When the
cultures reached their stationary phase, the total lipid content of the biomass was determined by a gravimetric method (Bigogno et al., 2002). Under the above-described growth conditions in bubble-columns in batch culture mode, the average biomass productivity of the best performing strains during the light-limited phase was in the range of 0.7 to 1.0 g AFDW L⁻¹ day⁻¹. We compared our newly isolated strains to the eustigmatophyte Nannochloropsis salina strain CCMP1776, a strain used in outdoor cultivation trials in raceway ponds (Unkefer et al., 2011, NAABB Final Report, 2014) and also in outdoor photobioreactors (Quinn et al., 2012). In our laboratory bubble column tests, N. salina reached only about 0.5 to 0.7 g AFDW L⁻¹ day⁻¹. Most of the strains in the third-level screen reached the stationary phase after about 5 to 6 days. At that time, the new strain’s total lipid content was about 15% to 25%, with lipid productivities of about 0.2 g Total Lipid L⁻¹ day⁻¹. Although growing slower, the lipid content of N. salina could reach over 30% lipid content within the five days in batch culture, thus ultimately reaching a lipid productivity similar to the strains we isolated. However, these newly isolated strains did exhibit shorter initial lag-phases and higher daily biomass productivities.

Recently, for biofuels applications, hydrothermal liquefaction (HTL) technologies, which convert the entire biomass into crude oils, rather than depending only on extractable lipids, have come to the forefront in this field (Barreiro et al., 2013, Stephens et al., 2013, Zhu et al., 2013). Consequently, higher biomass productivity is now considered a viable alternative to lipid productivity or content in algal liquid fuel production.
In our third-level screen, the strains with the highest biomass productivities belonged to the green algal families Scenedesmaceae and the Chlorellaceae. Visual examples for the rapid growth of some of the Scenedesmaceae are shown in Figure 14. Culture #1 contained the *S. obliquus* strain EN0004, which was one of the first and best biomass and lipid producers. Culture #5 contained a *Desmodesmus* species and culture #6 contained a second *Scenedesmus* species. In addition to strain EN0004, later during our screening several other strains of the species *S. obliquus* – isolated from a variety of different habitats – were identified in our multi-level screen as having high biomass and lipid productivities. *S. obliquus* EN0004 and the *Desmodesmus* sp. shown in Figure 14 as #5, performed successfully in small-scale raceway ponds at the NAABB testbed facility of the Texas A&M University at Pecos, Texas (NAABB Final Report, 2014).

**RESULTS AND DISCUSSION**

More than 5,000 strains were isolated during this project, and dozens of newly isolated strains were identified as potential new platform strains, either for biofuels applications and/or carotenoid production. However, only relatively few strains went through the entire three stage screening process, and we focus here only on two strains of the larger group of freshwater Scenedesmaceae belonging to the green algae: *S. obliquus* EN0004, and strain EN1234, the later being one example of a *Coelastrum*-related, but as of yet unknown species.

EN0004 was classified as *S. obliquus* based on its molecular marker rDNA ITS2 sequence – confirmed via its morphology (see Figure 11A). This strain was chosen for further characterization, because it exhibited one of the highest growth rates and lipid
productivities in the multi-level screen. Others have also considered strains of *S. obliquus* for use in potential biofuels applications (Mandal & Mallick, 2009, Breuer *et al.*, 2012, Vigeolas *et al.*, 2012), and strains of this species were previously found suitable for production in outdoor mass cultures (Becker, 1984, Grobbelaar *et al.*, 1990). The isolation of *S. obliquus* by our bioprospecting method described above validated that the approach taken can select for species, such as *S. obliquus*, which are potentially well suited for outdoor cultivation for biofuels.

The second strain, EN1234, demonstrates that novel strains could be identified as potential production strains for biofuels as well as carotenoid applications. Although this strain had low Nile Red fluorescence readings in the first-level screen (Figure 10), cultures of this strain turned orange, a characteristic of a concomitant lipid and carotenoid producing green alga. Light microscopy confirmed that cells had high amounts of carotenoids in cytosolic oil bodies during the lipid accumulation phase (Figure 11). Since carotenoids are known fluorescence quenchers, there was the possibility that the presence of carotenoids in Nile Red labeled oil bodies could have reduced the Nile Red signal during the screening process. Overall, about 5 to 10% of all strains screened did not exhibit high Nile Red fluorescence levels. But many of these strains could be identified as high lipid producers based on color changes due to their carotenoid co-accumulation with TAGs.

Based on molecular sequences and microscopic investigation, strain EN1234 is a green alga belonging to the family of the Scenedesmaceae (Figure 11). In general, cells of strain EN1234 were coccoid with cells from older batch cultures containing large orange oil bodies. However, cells of EN1234 were clearly different from those of *H. pluvialis*
and further molecular phylogenetic analysis will help clarify the taxonomic position of this novel strain.

Regarding biomass and total lipid productivities, in bubble-columns strain EN1234 performed as well as S. obliquus strain EN0004, but unlike strains from S. obliquus, strain EN1234 has not yet been tested in outdoor culture. In addition to the poorly understood green algal strain EN1234, a number of other carotenogenic strains with similar morphology and physiology were identified. Some of these high-producing strains were also successfully tested outdoors in small-scale high-rate ponds (NAABB Final Report, 2014).

The above described bioprospecting approach, with a combination of traditional cell cultivation and FACS cell sorting, allowed us to isolate about 1,000 strains per year. FACS cell sorting greatly enhanced the speed of strain isolation to a level that many more strains were isolated than could be screened. Because of manpower and time constraints, only about 2/3 of all isolated strains could ultimately be put through the first-level screen. Furthermore, only about 20% of the strains that passed the first-level and second-level screen could then be screened at the third level. Additionally, of the candidate platform strains coming out of our laboratory screening process, only a handful of strains could be taken to the final stage of outdoor ponds or large photobioreactors. Outdoor testing should be carried out as soon as possible after initial isolation. This is due to potential deterioration of strain characteristics including adaptations to low light during growth in the laboratory (Polle et al., 2004). In our case, the turnaround time from isolation to larger scale testing in ponds or bioreactors was about 12 months. However, in a test to
determine the fastest turnaround time, it took less than three months from strain isolation to completion of indoor screening.

Another limitation faced was the classification of isolated strains. As mentioned earlier, obtaining taxonomic data for thousands of strains was not possible. Therefore, we focused on classification of only the best performing strains. Based on microscopic analysis and using molecular markers of potential future platform strains, it was found that in our collection some algal species are represented by multiple strains – isolated from different environmental samples (Neofotis et al., In revision). For example, several strains of the green alga *S. obliquus* emerged as top performers, which compared favorably to commercially grown strains (Neofotis et al., In revision). Having multiple strains of the same species could be valuable as it increases the diversity in germplasm of species, which might be exploited in future strain development.

**CONCLUSION**

There is little reason to believe that the natural organisms most suitable for further algae biofuel development work have already been isolated (Mutanda et al., 2011, Wilkie *et al.*, 2011). The majority of algae species have yet to be identified or characterized. Therefore, bioprospecting for novel natural platform strains suitable for algal mass cultivation and the elucidation of their potential for generation of products is an essential first step for advancing microalgal biotechnology. Of course, multiple strategies can be applied to obtain novel promising strains including univariate or multivariate approaches (Barclay & Apt, 2013). We employed an approach that consisted of isolation of a large number of new strains that were then screened for use in biofuels applications. To that
end, our bioprospecting project resulted in 30 potential platform strains, some of which had been tested successfully in outdoor ponds (NAABB Final Report, 2014). By no means should bioprospecting of microalgae be viewed as a completed process. Any new microalgal-based technology would be expected to go through a phase of identification of the ‘best’ platform strains for that application, which might include new bioprospecting projects.

Following identification of proper platform strains, the next step should be strain development to improve existing qualities of the strain or to introduce new metabolic characteristics. As with higher plants, a number of techniques such as classical breeding, mutagenesis, directional selection, and genetic engineering could be applied. Combining these techniques on a strain, which has already evolved in the natural environment to be well suited to algal mass culturing, offers the most promise. Nevertheless, strain bioprospecting is the first step in making microalgal technologies commercially successful.
Chapter 2: Characterization and Classification of Highly Productive Microalgae Strains Discovered for Biofuel and Bioproduct Generation

Peter Neofotis\textsuperscript{a,b}, Andy Huang\textsuperscript{b}, Kiran Sury\textsuperscript{b}, William Chang\textsuperscript{b}, Florenal Joseph\textsuperscript{b,3}, Arwa Gabr\textsuperscript{b}, Scott Twary\textsuperscript{c}, Weigang Qiu\textsuperscript{d}, Omar Holguin\textsuperscript{c}, Jürgen E.W. Polle\textsuperscript{a,b,4}

\textsuperscript{a}Department of Biology, Graduate Center of the City University of New York, New York, New York 10016, USA
\textsuperscript{b}Department of Biology, Brooklyn College, Brooklyn, New York 11210-2889, USA
\textsuperscript{c}Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA
\textsuperscript{d}Department of Biology, Hunter College, New York, New York 10065, USA
\textsuperscript{e}Department of Plant and Environmental Sciences, New Mexico State University, Las Cruces, New Mexico 88003, USA

ABSTRACT

This paper describes the characteristics of microalgal strains that originated out of a multi-year strain isolation and screening project included within the algal biology division of the National Alliance for Advanced Biofuels and Bioproducts. The project’s goal was to identify new potential platform strains with high growth rates and/or lipid productivities. To classify the best performing strains from our screening, we conducted a combined microscopic and sequence/secondary structure phylogenetic analyses of the nuclear-encoded internal transcribed spacer 2 (ITS2) ribosomal RNA. Among the isolated strains that showed the most promise for future development were coccoid green algae. Several strains belong to the species \textit{Scenedesmus} (\textit{Acutodesmus}) \textit{obliquus}, one

\textsuperscript{2} A version of this paper is currently in revision to resubmit to \textit{Algae Research}
\textsuperscript{3} Present Address: Molecular and Cellular Biology, School of Graduate Studies at State University of New York Downstate Medical Center, Brooklyn, NY 11203
\textsuperscript{4} Author for correspondence: email jpolle@brooklyn.cuny.edu, Telephone 917-951-2025, Fax 718-951-4659

47
strain to the species *Borodilopsis texensis*, and three new strains to the species *Chlorella sorokiniana*. A number of additional strains could not be resolved to the species level and appear to be novel. Identified to the genus level were one *Desmodesmus* strain and one *Ankistrodesmus* strain. Another strain belonged to the order of the Chlamydomonadales and several strains were from the family of the Chlorellaceae. Although *Scenedesmus* and *Chlorella* species were already known to be commercially valuable, our strains expand on the existing germplasm. We also report multiple strains falling into the genus of *Coelastrella*, which are here reported for the first time for the continent of North America. Isolating and reporting on the above strains, some of which have been tested successfully in outdoor ponds and all of which are deposited at the University of Texas Culture Collection of Algae and publically available to researchers, is a step forward in expanding the biological resources available for algae biofuel production.

*Key index words*: biofuel, bioprospecting, ITS2, green algae, *Scenedesmus*, *Coelastrella*, *Desmodesmus*, *Ankistrodesmus*, *Chlorella*, triacylglycerols

*Abbreviations*: UHPLC, Ultra high performance liquid chromatography; PDA, Photo diode array; MS, Mass Spectrometry; APCI, Atmospheric Pressure Chemical Ionization; ITS, internal transcribed spacer; SSU, small-subunit; LSU, large-subunit; TAG, triacylglycerols

1. INTRODUCTION

Algae are seen as a more appealing biofuel feedstock than land plants because of their faster biomass doubling cycles, more accessible forms of stored carbon, and their
ability to thrive on waste or salt water sources (Markou & Nerantziz, 2013, National Research Council of the National Academy, 2013, Wijffels & Barbosa, 2010). But, because the vast majority of algae species have not been characterized in a biofuel context, the best strains for such production are unlikely to be known or utilized at large scales (Wilkie et al., 2011). Therefore, a need exists for “phycoprospecting” (Wilkie et al., 2011) – the identification of novel algal “platform strains” – which sustain high growth rates and accumulate lipids (Barclay & Apt, 2013, Georgianna & Mayfield, 2012, Mutanda et al., 2011, Sheehan et al., 1998). In addition, strains that produce significant amounts of higher-value products, such as carotenoids, which are sold as nutraceuticals, are also desired (Borowitzka, 1986, Markou & Nerantziz, 2013, Mulders et al., 2014, Olaizola, 2003, National Research Council of the National Academy, 2013, Wijffels & Barbosa, 2010). These novel “platform strains” could, in turn, be developed further via artificial selection, genetic engineering, or other “crop improvement” methods and used to generate biofuel feedstocks at lower costs (Davis et al., 2011, Gimpel et al., 2013, Wijffels & Barbosa, 2010).

Commercial scale algae ponds have been operated for more than a decade (Del Campo et al., 2007, Pulz & Gross, 2004, Wijffels & Barbosa, 2010, Vanthoor-Koopmans et al., 2013), mainly to harvest pigments and metabolites that are desired as nutritional supplements (Borowitzka, 2013). The algae cultivated in these ponds include Spirulina (Arthrospira), for high protein powder (Gershwin & Belay, 2008, Guedes et al., 2011), Haematococcus, for the antioxidant astaxanthin (Guerin et al., 2003, Laurens et al., 2012); and Dunaliella salina, for pro-vitamin A production (Borowitzka, 1991, Ben-Amotz et al., 2009)
Due to the success demonstrated with these species (Pulz & Gross, 2004), the as of yet untapped diversity of microalgae appears to offer possibilities in the field of biofuels (Elliott et al., 2012a, Ratha & Prasanna, 2012, Ratha et al., 2012), as well as the production of high-value products (Markou & Nerantziz, 2013), such as pharmaceuticals (Borowitzka, 2013, Wijffels et al., 2013, Draaisma et al., 2013).

It is not clear how many microalgae species there are, with estimates running from 70,000 to one million (Clerck et al., 2013, John, 1994, Norton et al., 1996). Only 44,000 have been described (Clerck et al., 2013, Guiry, 2012). New species and genera are consistently being discovered; this consistent rate of discovery indicates that a large proportion of undocumented species exists (Clerck et al., 2013).

Based on the known utility of already characterized species and the known absence of knowledge regarding the large amount of undocumented strains, it is hypothesized that previously uncharacterized strains of microalgae exist which exhibit high growth rates and lipid productivities. These novel “strains” may be new species or varieties of species with more favorable growth characteristics for biofuel production. Some of these species or strains may also produce high value products. To test this hypothesis, a large-scale effort was conducted to isolate, screen, identify, and characterize microalgae strains that could be used as platform strains for biofuel and/or high-value product generation. For a number of the most promising strains to come out of the screening, a classification was conducted. In deciding on an appropriate barcode, the nuclear rDNA Internal Transcribed Spacer 2 Region (ITS2) was chosen because, unlike the 18S – which often does not vary enough to distinguish algal species (Hegewald et al., 2010) – the ITS2 has proved to be a helpful tool for discrimination at the genus
(Hegewald et al., 2010, Hegewald et al., 2013) and species level (An et al., 1999, van Hannen et al., 2002, Coleman, 2003, Coleman, 2009, Hegewald et al., 2005, Jeon & Hegewald, 2006, Schultz et al., 2006, Schultz & Wolf, 2009). At the same time, its two dimensional structure is highly conserved throughout the eukaryotes (Coleman, 2007, Mai & Coleman, 1997, Schultz et al., 2005). Combining the fast evolving sequences with its slowly evolving structure has allowed for it to be used in high level classifications while still discriminating at the low species level, within the same phylogenetic tree (Coleman, 2003).

The overall goal of our research was to discover new microalgal strains that could enhance the genomic/biological resources available for algae biomass production to be used as feedstock for biofuels and/or for bioproducts generation. The focus of this report is on freshwater algae originating mainly from the NAABB phycoprospecting project (NAABB Final Report, 2014). The objective of this report is to document some of these strains and also begin to have a cladistic understanding of where they fall in the algal tree of life. It is hoped that this work may provide some clues for targeted phycoprospecting approaches in the future.

2. MATERIALS AND METHODS

2.1. Sampling, isolation, and screening

Cells were sampled and screened as described previously in detail (Neofotis et al., 2015). Briefly, environmental samples were collected over the three years of the NAABB
prospect (2010-2012) from a variety of habitats (including freshwater lakes as well as soils) across the continental U.S. and during all seasons (Table 1). As described in detail (Neofotis et al., 2015), for strain isolation an approach was taken that combined the traditional plating method (Andersen & Kawachi, 2005) with high-throughput flow cytometry using fluorescence aided cell sorting (Sieracki et al., 2005, Dolezel et al., 2007).

For the first-level screen, in summary, batch cultures of strains were grown in Erlenmeyer flasks in defined minimal media suitable for this selection: C Medium, Bold Basal Medium, and BG11 Medium (Andersen & Kawachi, 2005), with the only CO₂ enrichment being the sodium bicarbonate which is added to BG11. Flasks were grown under 50 µE m⁻² s⁻¹ continuous light provided by daylight fluorescent lamps (See Supplementary Note 1 for more details on screening methods).

Strains that passed the first-level screen were deemed to be potential producers and were selected for further characterization (Figure 1). In a second-level screen, the strains were cultivated in multiple different media to determine if another growth medium was better suited for use in continued strain characterization.

In a third-level screen for analysis of actual biomass and lipid production potential, strains were grown at 28 Degree Celsius in glass columns (3 cm width with 500 mL volume) into which CO₂ enriched air was bubbled through the cultures from the bottom. When the algal cultures reached their stationary phase, biomass also was taken and analyzed via a gravimetric method to determine the total lipid content (Bigogno et al., 2002).
The daily AFDW was then graphed using the ggplot package, within R (http://cran.r-project.org/), to create weighted polynomial curves, which represented algal growth curves. The doubling time was also calculated in R for the exponential growth phase with the formula log(2)/k, with k being the slope of the growth curve.

2.2 Metabolite analysis

Pigments and lipids were extracted using a modified version of an established protocol (Bligh & Dyer, 1959) (See SM Note 2). Pigments and lipids were analyzed using a Ultra High-Performance Liquid Chromatography (UHPLC) C18 reverse-phase column, connected to a Photo Diode Array detector (Flexar FX-15 UHPLC system from Perkin Elmer), and then passed through an Atmospheric-Pressure Chemical Ionization (APCI) probe connected to a Time-of-Flight Mass Spectrometer (Perkin Elmer, Massachusetts) (UHPLC-PDA-APCI-MS). The samples were run through the LC with solvent A consisting of 39% water, 59% Acetonitrile, and 2% ammonium acetate and solvent B consisting of 88.2% 2 propanol, 9.8% acetonitrile, and 1.9% ammonium acetate. The column was first equilibrated for 15 minutes with 60% A and 40% B and a flow of .4 ml/min. A 10 minute gradient transition then brought solvent B to 100% — and it was held for 6 minutes. For one minute, then, solvent B was brought down to 40% and solvent A was brought to 60%. Following, the column was washed for 5 minutes. For the PDA, the differential absorbance was set at 450nm/550nm to detect the carotenoids. For the MS, the following settings were used – Corona 5uA, Endplate -2000 Volts, Cap

2.3 Benchmark for comparison

Our ‘gold-standard’, or benchmark, for biomass and lipid productivity comparison of newly isolated strains was the species *Nannochloropsis salina* strain CCMP1776. This strain was used by the National Alliance for Advanced Biofuels and Bioproducts (NAABB Final Report, 2014, Unkefer et al., 2011) in outdoor cultivation trials in raceway ponds (NAABB Final Report, 2014). It had also previously been evaluated in outdoor photobioreactors (Quinn et al., 2012).

*DNA Extraction, polymerase chain reaction (PCR), and sequencing.* A 50 ml culture in the light-limited growth phase was concentrated by centrifugation, and then the DNA of 1 ml of the concentrated liquid sample was extracted using a MO-BIO DNA Isolation Kit (USA). The ITS2 rDNA region was amplified using the universal primers ITS3 (GCATCGATGAAGAACGCAGC) and ITS4 (TCCTCCGCTTATTGATATGC) (White et al., 1990). GoTaq® (Promega) master mix was used as the PCR reaction mixture. We performed 30 cycles (1 minute 95°C, 1 minute 48°C, and 1 minute at 72°C – with an initial denaturing step of 5 minutes at 95°C). In cases where, even with the ITS2 barcode, the identity of the species was still not clear, the 18S, ITS1, 5.8S, and ITS2 regions were also amplified using universal primers AL1500af (GCGCGCTACACTGATGC) and LR 1850 (CCTCACGGTACTTGTTC). In these instances, 30 cycles (30 seconds 95°C, 30 seconds 48°C, and 1.5 minutes at 72°C – with
an initial denaturing step of 2 minutes at 95°C) we performed. PCR products were then purified and sequenced by Genewiz (www.genewiz.com). For one of these organisms, DOE0101, the DNA was also extracted using a DNeasy Plant Maxi Kit (Qiagen) and the genome was then sequenced using Illumina at Los Alamos National Laboratory, NM.

2.4. Alignment and phylogenetic analysis

All 28 ITS2 sequences obtained from our lab samples were delimited and cropped with the hidden Markov model (HMM)-based annotation tool present at the ITS2 database (Keller et al., 2009)(E-Value <0.001, Viridiplantae HMMs). In addition to the 28 ITS2 sequences from our top performing strains, we also downloaded ~50 algae ITS2 sequences from the ITS2 database (http://its2.bioapps.biozentrum.uni-wuerzburg.de/). As the sequences were to be used to help determine the taxonomic position of our strains, sequences from algae branches were chosen that NCBI blasts suggested were close to our specimens – within the Chlorophyceae and Trebouxiophyceae clades. From the Trebouxiophyceae, the Chlorellales order seemed particularly well represented. Within the Chlorophyceae, the Sphaeropleales and Chlamydomonadales orders were well represented. Since one of our strains appeared to be an Ankistrodesmus, yet no sequences of that genus were available in the ITS2 database, an Ankistrodesmus sequence was also downloaded from NCBI and cropped using the aforementioned model. As outgroups, two species from the prasinophytes (Mantoniella squamata and Micromonas pusilla) were chosen. Mantoniella squamata has been used previously to root phylogenetic trees showing the evolution relationship of chlorophytes based on SSU rRNA sequence data.
(Kessler et al., 1997). *Micromonas pusilla* is a well characterized prasinophyte whose complete genome had been sequenced (Worden et al., 2009a).

The cropped sequences and structures were then automatically aligned with 4SALE 1.7 using an ITS2-specific scoring matrix for sequences and structures (Seibel et al., 2006, Seibel et al., 2008). To find simultaneously the evolutionary distance between organisms according to sequence and secondary structures, profile neighbor joining (PNJ) was applied using ProfDistS 0.9.9 (Qt-Version) (Friedrich et al., 2005, Wolf et al., 2008). An ITS2-specific general time-reversible substitution model was applied (Seibel et al., 2006). The resulting tree was displayed with FigTree (http://tree.bio.ed.ac.uk/) and further processed in R (http://cran.r-project.org/) using Ape (Paradis, 2006). For a few of the strains, more detailed sub-trees were reconstructed by downloading additional ITS2 or 18S, ITS1, 5.8S, and ITS2 sequences GenBank. The sequences for this subtree were aligned with ClustalW and reconstructed using Maximum Likelihood with a Tamura-Nei Model and Nearest-Neighbor-Interchange (NNJ) heuristic method, with 500 bootstrap replications performed. Significant bootstrap values are based on 1000 replicates and mapped to the appropriate internode. Branch lengths were drawn proportional to inferred changes.

Twenty eight of the strains described in this study were deposited at the UTEX Culture Collection of Algae (http://web.biosci.utexas.edu/), and all sequences generated were also deposited at the National Center for Biotechnology Information (NCBI) (Benson et al., 2008).

2.5. *Secondary Structure Prediction*
The structures of most of the sequences downloaded from the ITS2 database were determined either by direct-fold or homology modeling within the ITS2 database (Wolf et al., 2005, Schultz et al., 2006, Selig et al., 2008). The secondary structures of most of the sequences from our strains were also determined by the same homology model. For four of our sequences, that of DOE0101, EN1423, DOE0259, and the Ankistrodesmus sequences from NCBI – because no related organisms were found in the ITS2 database – RNAstructure 5.6 was used to predict the ITS2 structure (Mathews et al., 2004).

3. RESULTS AND DISCUSSION

The focus of our manuscript is on the phylogeny and characterization of some of the most promising strains that were isolated in the course of our phycoprospecting work for the NAABB consortium (NAABB Final Report, 2014). Nevertheless, the fist part of this section provides an overview of the phycoprospecting project.

3.1. Summary of the NAABB phycoprospecting project

The multi-year strain isolation and screening process, which yielded a variety of novel strains, consisted of multiple stages (Neofotis et al., 2015). For a better understanding of the results, the overall approach that resulted in potential future platform strains is discussed below:

1. A dual spatial and temporal sampling strategy was employed. Diverse sampling sites included many aqueous and several terrestrial environments, with an emphasis on the
southwestern United States, because the southwestern States were at the time viewed as potential sites for algal mass cultivation (U.S. Department of Energy, 2010). Many shallow and temporal water bodies were sampled because it was anticipated that future use of strains would include raceway-type ponds. Whenever possible, repetitive collection at sites accounted for temporal succession of microalgae communities in the selected habitats.

2. Aliquots of environmental samples were plated onto different agar-solidified media in Petri dishes. The rationale for this first step was that different types of algae could be easily distinguished by colony color and size. This first step allowed us to determine if new isolates could grow on agar-solidified media. This would be important for future strain development, which might later be done in high-throughput applications (Zhang et al., 2014).

3. Transfer of colonies onto index plates was done with the rationale that more biomass could be used for inoculation of liquid medium, which reduced the lag-phase of the liquid cultures, thus speeding up the isolation process. Additionally, cyanobacteria, which were apparent because of their blue-green color, were avoided as it was assumed that they do not contain significant amounts of lipids.

4. Re-suspension of cells directly from the original index plates into 2 mL of liquid medium was done initially. But in that case most cells transferred in clumps, thus preventing single cell or coenobial sorting. So, the cells were transferred to liquid medium. This had the additional benefit of demonstrating the cells do indeed grow in liquid, which most later applications would involve with open ponds or photobioreactors.
5. FACS cell sorting onto Petri dishes was applied with the rationale to provide a high-throughput method to obtain a large number of unialgal strains rapidly. Sometimes new strains were also already axenic following cell sorting onto plates. Sorting into liquid wells did not prove as efficient because determination of potentially unialgal strains was more difficult.

6. Unialgal strains were then transferred into storage tubes consisting of agar-solidified medium overlaid with liquid medium. This storage method was chosen because all strains could be maintained in this manner and strains survived longer in agar/liquid-overlay cultures under a 12 h light/dark regime without transfer for up to three years.

Figure 1 summarizes our phycoprospecting efforts for the NAABB program. Within the isolation and screening project, several bottlenecks were encountered at different levels due to the labor-intensive steps involved. Out of our total of 2,465 isolated NAABB strains, only 1,575 strains could be subjected to a first-level screen, of which 334 passed as candidates for potential use as future biofuel/nutraceutical platform species. Subsequently, from the 334 candidate strains only 200 strains could be tested in a second-level screen to determine the optimal growth medium. In the following third-level screen, only 110 strains could be examined for actual biomass and lipid productivities in bubbling columns. Of the 110 strains tested in bubbling columns, more than 30 freshwater strains outperformed our ‘gold’ standard, *Nannochloropsis salina*, in terms of growth rate and final biomass yield. In addition to having high growth rates, several of our best performing strains also accumulated carotenoids in cytosolic oil bodies.
Overall, 30 strains were deposited with the UTEX culture collection. Currently, these strains are not in the main UTEX collection, but are in a special collection. Out of these 30 deposited strains, general information on 25 strains is listed in Table 1. Additionally, Table 1 includes information on three strains from a previous phycoprospecting effort.

Table 1: Listed are the top performing strains that were classified in the work reported here. Included in the table are the strain numbers, taxon, and location along with the
habitat of collection. Also listed is whether or not the strain is carotenogenic and whether or not it was grown successfully outdoors. Strains with the ID DOE#### are from the NAABBB phycoprospecting project. All these strains have been deposited at the University of Texas Culture Collection of Algae and are available to researchers. Strains with the ID EN#### resulted from an Airforce Office of Science funded phycoprospecting effort. The asterix next to Strain ID numbers indicates that the strain was tested successfully in outdoor ponds at NAABBB testbed sites (NAABBB Final Report, 2014).

<table>
<thead>
<tr>
<th>Strain Number</th>
<th>Taxon</th>
<th>State</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Habitat</th>
<th>NCBI Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOE0013</td>
<td><em>Scenedesmus</em></td>
<td>CA</td>
<td>37.97</td>
<td>-122.56</td>
<td>Fountain</td>
<td>KJ434964</td>
</tr>
<tr>
<td>DOE0043</td>
<td><em>Desmodesmus</em></td>
<td>CA</td>
<td>37.9</td>
<td>-122.25</td>
<td>Lake</td>
<td>KJ434974</td>
</tr>
<tr>
<td>DOE0088</td>
<td><em>Coelastrella</em></td>
<td>CA</td>
<td>37.96</td>
<td>-122.34</td>
<td>Birdbath</td>
<td>KJ43972</td>
</tr>
<tr>
<td>DOE0101*</td>
<td><em>Chlamydomonadales</em></td>
<td>TX</td>
<td>29.92</td>
<td>-95.82</td>
<td>Roadside Mud</td>
<td>KJ434990</td>
</tr>
<tr>
<td>DOE0111</td>
<td><em>Scenedesmus obliquus</em></td>
<td>CA</td>
<td>37.96</td>
<td>-122.34</td>
<td>Garden Pot</td>
<td>KJ439665</td>
</tr>
<tr>
<td>DOE0152*</td>
<td><em>Scenedesmus obliquus</em></td>
<td>NY</td>
<td>40.65</td>
<td>-74.01</td>
<td>Turtle Tank</td>
<td>KJ434966</td>
</tr>
<tr>
<td>DOE0155</td>
<td><em>Coelastrella</em></td>
<td>TX</td>
<td>29.84</td>
<td>-95.87</td>
<td>Bayou</td>
<td>KJ434970</td>
</tr>
<tr>
<td>DOE0202*</td>
<td><em>Coelastrella</em></td>
<td>CA</td>
<td>37.96</td>
<td>-122.34</td>
<td>Water in Garden Pot</td>
<td>KJ434971</td>
</tr>
<tr>
<td>DOE0259</td>
<td><em>Ankistrodesmus</em></td>
<td>TX</td>
<td>29.87</td>
<td>-95.81</td>
<td>Roadside Ditch</td>
<td>KJ434991</td>
</tr>
<tr>
<td>DOE0314</td>
<td><em>Dichloster-related</em></td>
<td>TX</td>
<td>29.87</td>
<td>-95.81</td>
<td>Roadside Ditch</td>
<td>KJ434977</td>
</tr>
<tr>
<td>DOE0369</td>
<td><em>Coelastrella</em></td>
<td>CA</td>
<td>38.89</td>
<td>-121.97</td>
<td>Fountain</td>
<td>KJ43967</td>
</tr>
<tr>
<td>DOE0623</td>
<td><em>Chlorellaceae</em></td>
<td>TX</td>
<td>29.77</td>
<td>-95.62</td>
<td>Bayou</td>
<td>KJ434983</td>
</tr>
<tr>
<td>DOE0686</td>
<td><em>Chlorella sorokiniana</em></td>
<td>CT</td>
<td>41.6</td>
<td>-72.7</td>
<td>Greenhouse Hydroponics</td>
<td>KJ434981</td>
</tr>
<tr>
<td>DOE0700</td>
<td><em>Chlorellaceae</em></td>
<td>NM</td>
<td>32.31</td>
<td>-106.78</td>
<td>Wastewater Pond</td>
<td>KJ434979</td>
</tr>
<tr>
<td>DOE0717</td>
<td><em>Chlorellaceae</em></td>
<td>CT</td>
<td>41.6</td>
<td>-72.7</td>
<td>Greenhouse Hydroponics</td>
<td>KJ434986</td>
</tr>
<tr>
<td>DOE1041</td>
<td><em>Chlorellaceae</em></td>
<td>TX</td>
<td>30.29</td>
<td>-97.84</td>
<td>Roadside Ditch</td>
<td>KJ434978</td>
</tr>
<tr>
<td>DOE1044</td>
<td><em>Chlorellaceae</em></td>
<td>TX</td>
<td>30.78</td>
<td>-95.95</td>
<td>Unknown</td>
<td>KJ434987</td>
</tr>
</tbody>
</table>
3.2. Growth evaluation of strains

Following sampling and strain isolation a multi-level screen was applied to identify potential lipid producers. The first-level screen evaluated growth by optical density changes and lipid productivity using Nile Red fluorescence (Neofotis et al., 2015). The second-level screen evaluated performance of strains in different growth media (Neofotis et al., 2015). In the third-level screen strains were cultivated in bubbling columns (Neofotis et al., 2015) and biomass productivities and lipid levels were determined gravimetrically. During the light-limited growth phase the biomass productivity of the best strains identified was between 0.7 to 1.0 g AFDW l\(^{-1}\) d\(^{-1}\) (Figure 2). In comparison, *N. salina*, our ‘standard’, only produced about 0.6 g AFDW l\(^{-1}\) d\(^{-1}\).
For most strains the third-level screen only consisted of one culture having been tested. Only for a few selected groups of strains more than one culture was grown. Figure 2 shows growth curves of selected groups of the best performing microalgal strains identified: what our phylogenetic analysis would classify as Chlamydomonadales strain DOE0101, Chlorella sorokiniana strain DOE1412, three novel Coleastrella sp., and two strains of the species Scenedesmus obliquus - also known as Acutodesmus obliquus (Hegewald et al., 2013). Because all bubbling column cultures were inoculated from pre-cultures that were in the light-limited growth phase, the lag-phase for all the bubbling column cultures were usually very short. The following exponential growth phase was also short. The doubling time during the exponential growth phase was fastest for C. sorokiniana strain DOE1412 (24.3 hours), followed by the S. obliquus strains (26.8 hours). In comparison, the strains of the Coelastrella species and Chlamydomonadales sp. strain DOE0101 were the slower growers, both with doubling times at about 38 hours. Although rigorous comparison across experiments carried out in the screens different conditions is not possible, the doubling times of these strains were also within the range of the reviewed most promising species of algae for lipid productivity (Griffiths & Harrison, 2009).

For all strains shown in Figure 2 the exponential growth phase was followed, between about 50 to 100 hours, by the light-limited growth phase. Then growth gradually slowed and after about 120-150 hours in the bubbling columns, most of our tested strains reached the stationary growth phase with total biomass yields of sometimes more than 4.0 g AFDW L⁻¹.
Figure 2: Shown are growth curves based on the AFDW (grams/liter) of some of the most productive strains. The *Chlamydomonadales* species was strain DOE0101; the *Chlorellales* species was *Chlorella sorokiniana* strain DOE1412; The *Coelastrella* strains were DOE0088, DOE0202, and DOE0369; and the *Scenedesmus obliquus* strains used were DOE0152 and DOE0111. At least two biological replicates were included for each strain. Shading represents the 95% confidence region for the regression fit of the biological replicates.

For the data shown in Figure 2, for *S. obliquus* the growth curve shown represents two strains DOE0111 and DOE0152. The growth curve of *Coelastrella sp.* represents the summary results of three independently isolated strains DOE0088, DOE0202, and DOE0369. Noteworthy is that at about 200 hours, when the batch cultures were terminated, the strains of the *Coelastrella* had still not reached the stationary growth phase. At that time, the biomass yield was about 4.5 g AFDW l\(^{-1}\). When grown for a longer time period, the biomass yields were sometimes upwards of 6.0 g AFDW l\(^{-1}\) and the cultures had still not reached the stationary phase.

When in the stationary growth phase at 180 hours the strains had accumulated the following amounts of total lipids as part of their biomass: *Chlamydomonadales sp.* strain DOE0101 (25.3%); *Chlorella sorokiniana* DOE1412 (25.4%); *Coelastrella sp.* strain DOE0088 (16.5%) and strain DOE0202 (22.5%); *S. obliquus* strain DOE0111 (28.5%) and strain DOE0152 (27.6%). It should be noted that due to their thick cell walls not all lipids could be extracted from cells of the *Coelastrella* strains, thus leading to
underestimating their total lipid contents. Overall, the total lipid productivities of all strains shown in Figure 2 were similar at about 0.2 g lipids l\(^{-1}\) d\(^{-1}\).

Although growing slower, the lipid content of *N. salina* could reach over 30% lipids within the five days in batch culture, thus ultimately also reaching about 0.2 g Total Lipid l\(^{-1}\) day\(^{-1}\). However, compared to *N. salina*, our top strains had shorter lag-phases and higher daily biomass productivities. This is important, as the latest biofuels applications aim at use of Hydrothermal Liquefaction Technologies (HLT), which convert biomass into crude oils rather than extracting lipids only (Zhu et al., 2013, Barreiro et al., 2013, Stephens et al., 2013, Elliott et al., 2015). Consequently, rather than technologies that are dependent on the lipid content of algal cells, Hydrothermal Liquefaction Technologies can rely on fast growing, biomass-accumulating strains to increase higher overall crude oil productivities.

### 3.3. Characterization of carotenogenic capabilities of selected strains

When cultivated in shaker flasks during the first-level screening process, it was noticed that a variety of strains over-accumulated secondary carotenoids in cytosolic lipid bodies (Neofotis et al., 2015) similar to *Haematococcus pluvialis* (Boussiba, 2000, Jin et al., 2006). All *Coelastrella* strains listed in Table 1, the strain DOE0101 belonging to the Chlamydomonadales, and *Borodinellopsis texensis* strain EN1423 displayed such carotenogenic capability. This over-accumulation of secondary carotenoids, including Astaxanthin, may be of value for future biotechnological application in carotenoid production. As shown above under the section of growth evaluation, *Coelastrella sp.*
strain DOE0202 and the *Chlamydomonadales* sp. strain DOE0101 performed well in the third-level screen in the laboratory. In addition, both strains were successfully cultivated in outdoor ponds at the NAABB testbed site in Pecos, TX (NAABB Final Report, 2014). Therefore, these two strains were further characterized. As the secondary carotenoids accumulate in cytosolic oil bodies, the carotenoid and triacylglycerol (TAG) profiles were compared for both strains for actively growing (green) cultures and carotenogenic (orange) cultures (Figure 3). In the center, Figure 3 displays exemplary UHPLC chromatogram traces of extracted pigments. Peaks shown in the UHPLC chromatogram traces are either carotenoids or chlorophylls. Identification of carotenoids was done based on their absorbance spectra and based on their molecular mass (http://metlin.scripps.edu/index.php). Nevertheless, for some of the peaks, identifying the molecules as one carotenoid rather than another could not be done, as they were either stereoisomers or different ions with the same mass. However, carotenoids of pharmaceutical relevance are present in these algae. As indicated by the orange color of cells and cultures, the three strains investigated contained different precursors of Astaxanthin (peak 1 with Adonirubin or Astaxanthin; Peak 3 with Canthaxanthin or Adonixanthin; Peak 4 with Echinenone or Cryptoxanthin), but possibly also minor amounts of Astaxanthin (Peak 3) and Astaxanthin diesters (Peaks at later retention times). In brief, strain DOE0101 seemed to have a distinctly different profile with a particularly high accumulation of either canthaxanthin or adinoxanthin, when compared to the *Coelastrella* species strain DOE0202 (Figure 3).
Figure 3: Compared are cells and metabolites of the *Chlamydomonadales sp.* strain DOE0101 (upper panel) and the *Coelastrella sp.* strain DOE0202 (lower panel). In the center of the figure, PDA profiles (absorbance traces at 450 nm) are displayed of extracts from cells of actively growing cultures and of carotenogenic (orange) cultures. Photos for representative cells from these cultures are shown for both strains. The black bars underneath the cells of strain DOE0101 indicate 10µm. The same scale was used for photos of cells from strain DOE0202. Extractions were conducted from the same volume of cells for cultures of the same age. Comparison with retention times of concurrent MS-TOF measurements suggest the carotenoids are - following the number or profiles - 1) Adonirubin or Astaxanthin; 2) Lutein and/or Zeaxanthin; 3) Canthaxanthin or Adonixanthin; and 4) Echinonene or Crytoxanthin. Further identification of the isomers was beyond our analytical capabilities. However, based on the late retention times, it is assumed for extracts from carotenogenic cultures that Astaxanthin Diesters (labeled as AstaDE) are represented by the late peaks. The outer panels of this figure show corresponding TAG profiles of the cells of both strains. Peaks visible at longer retention times are indicative of TAG molecules containing fatty acids with greater degrees of saturation. Peaks at shorter retention times represent presence of TAG molecules containing fatty acids with more double bonds. The overall class of TAG molecules shown in each profile is indicated by the labels with the example of 16_16_16 indicating TAG molecules containing three fatty acid chains each having 16 carbon atoms. Additional lipid profiles can be seen in SM-2.

Figure 3 also displays profiles of overlaid Extracted Ion Counts (EIC) traces for a variety of triacylglycerol molecules identified for both strains. The individual EIC traces were obtained by UHPLC separation followed by APCI-TOF-MS detection of TAG molecules from lipids extracted from algal strains. Each EIC peak represents a different specific TAG molecule. In general, TAG molecules with longer fatty acid chains eluted at later retention times than TAG molecules with shorter chains (Brydwell, 2005, Buchgraber *et al.*, 2004). More saturated TAG molecules (fewer double bonds) eluted at later retention times than TAG molecules with higher unsaturation grades (Liu *et al.*, 2013, Buchgraber *et al.*, 2004). All strains investigated contained dozens of different molecular classes of TAGs. The main molecular classes of TAG molecules that could be distinguished based on retention time and molecular mass contained:

1) Three fatty acids with 16 carbon atoms (C\textsubscript{16}-C\textsubscript{16}-C\textsubscript{16}),
2) Two fatty acids with 16 carbon atoms, and one with 18 carbons (C\textsubscript{16}-C\textsubscript{16}-C\textsubscript{18}).

3) One fatty acid with 16 carbon atoms and two with 18 carbons (C\textsubscript{16}-C\textsubscript{18}-C\textsubscript{18}).

4) TAG molecules that contained three fatty acids with 18 carbon atoms (C\textsubscript{18}-C\textsubscript{18}-C\textsubscript{18}).

In general, not all TAG molecules could be expected to ionize at the same rate (Brydwell, 2005). Therefore, peak heights or peak areas cannot be used for absolute quantification of TAG molecules in the presented UHPLC-TOF-MS analysis.

Nevertheless, the profiles of EIC traces representing the same TAG molecules can be compared between species. The EIC traces representing the TAG profiles show that the strains DOE0202 and DOE0101 each had distinct TAG profiles. For example, growing cultures of the *Coelastrella* species strain DOE0202 had more unsaturated C\textsubscript{16}-C\textsubscript{16}-C\textsubscript{18}s and C\textsubscript{16}-C\textsubscript{18}-C\textsubscript{18}s. As visualized in Figure 3, when compared to the *Coelastrella* strain DOE0202, the difference of strain DOE0101’s TAG profile under stress conditions remained apparent. Superficially, under the light microscope the cells look somewhat similar (Figure 3). Nevertheless, stressed cells of DOE0101 have comparatively very little C\textsubscript{16}-C\textsubscript{16}-C\textsubscript{18} TAGs, and have mostly saturated C\textsubscript{16}-C\textsubscript{18}-C\textsubscript{18} molecules. In contrast, cells of the *Coelastrella* species have more C\textsubscript{16}-C\textsubscript{16}-C\textsubscript{18} TAGs. Also, its C\textsubscript{16}-C\textsubscript{18}-C\textsubscript{18}s are weighted, if anything, toward the unsaturated molecules (Figure 3). These differences among the cells of the two species may be due to the different amounts of the various secondary carotenoids in the two strains that are sequestered in the oil bodies.

Investigating this further is of pharmaceutical/cosmetic relevance, but is beyond the scope of this work.

3.4. *Strain classification*
For the phylogenetic analysis of microalgal strains coming out of our isolation and screening work the rDNA ITS2 was used as the molecular marker for classification as in other algal studies (Buchheim et al., 2012, Hegewald et al., 2010, Vanormelingen et al., 2007). The phylogenetic analysis was done for two reasons. First, since morphological characteristics are unable reliably to distinguish many microalgae, molecular phylogeny was necessary to properly identify them. Secondly, the phylogeny allowed us to begin to gain a clade-based understanding of what clades typically have top performers.

For a first comparison to our newly isolated strains, sequences of 49 known strains were used. The specific known strains chosen for inclusion represented genuine green algal strains as determined by previous publications. This selection of sequences from genuine strains was important for an accurate phylogenetic classification of the newly isolated strains; often rDNA sequences deposited in databases originate from incorrectly named green algal strains. In the classification, three strains were also included that had come out of a previous phycoprosppecting project of this laboratory. Two of these strains (EN1234 and EN1235) had very similar cell morphology and carotenogenic capabilities to our strains DOE0088, DOE0155, DOE0202, and DOE0369. A third strain EN1423 was also carotenogenic, but very different in cell morphology. Figure 4 shows our recovered phylogenetic tree.
Figure 4: ITS2 sequence-structure tree of novel strains for biofuel production and also known strains of the Chlorophyta – rooted with two prasinophyceae. Significant bootstrap values are based on 500 replicates and mapped to the appropriate internode. Branch lengths are drawn proportional to inferred changes. Arrows point to strains discussed in this paper in more detail.

Within the ITS2 phylogenetic tree - including 25 newly isolated strains from the NAABB project, 3 previously isolated novel strains, and the 49 known strains – it is visible that some clades of the tree were not resolved, because of low support in some internal branches (Figure 4). Nevertheless, the major branches were well supported in...
their monophyly, and for the known sequences from the ITS2 database, our phylogeny mirrors the established relationship among the Chlorophyta (Leliaert et al., 2012). Similarly, in the systematics of the major clade within our study, the Scenedesmaceae, there is low bootstrap support for the bases of the trees and only good support for the groups of species within genera (Eliáš et al., 2010, Hegewald et al., 2010, Hegewald et al., 2013). Therefore, despite the low bootstrap values within our tree, that the known taxa fall in accordance with other published ITS1, ITS2, and 18S trees yields some confidence for the phylogenetic information revealed in the branches. Our investigation corroborates other studies (Buchheim et al., 2011), which showed that sequence-structure analysis of ITS2 provides a taxon-rich means of testing phylogenetic hypothesis at high taxonomic levels.

In general, our best performing strains were coccoid green algae that either fell within the green algal order of the Chlamydomonadales (Chlorophyceae), the Sphaeropleales (Chlorophyceae), or the Chlorellales (Trebouxiophyceae) (Figure 4). Among the thousands of strains that were isolated, preliminary microscopic investigation of many strains during the isolation and screening process demonstrated presence of flagellate, coccoid, and even some filamentous strains of the green algae. In addition, many cyanobacteria, diatoms, and other algae for example Eustigmatophytes were identified based on microscopic examination of cells. Therefore, our overall isolation protocol did not, per se, discriminate against non green algae. This report focuses on the green algae because, among the best performing strains emerging from our multi-level screen, the majority were coccoid green algal species. This may be because non-motile coccoid cells do not ‘waste’ resources on active movement, a process that consumes large
amounts of energy. Moreover, coccoid cells cannot determine their location and may have a need to store high-energy containing molecules as reserve substances. In consequence, many coccoid species could have evolved to store energy in oil bodies, specifically under environmentally unfavorable conditions.

3.4.1. Sphaeropleales (Class of the Chlorophyceae)

With regard to the newly isolated strains that are predicted to be within the class of the Chlorophyceae, many strains belong to the order of the Sphaeropleales and the family of the Scenedesmaceae. Several of our top performing strains (DOE0152, DOE0111, DOE0013) are similar to strains identified in the ITS2 database as \textit{S. obliquus}, the cells of these strains also have the characteristic oblong shape and coenobia. As an example, we show cells of \textit{S. obliquus} strain DOE0152 in Figure 5.

![Figure 5: Photos showing a variety of cells of \textit{S. obliquus} strain DOE0152. A) Side and top views of several four-celled coenobia and one single cell in the center. B) A larger four-cell coenobium where the individual cells began division, as indicated by broadening pyrenoids. C) A four-celled coenobium where one cell already completed cell divisions and released a new four-celled coenobium.](image)

Although the species \textit{S. obliquus} is synonymous to \textit{A. obliquus} (Hegewald et al., 2013), in this paper, we refer to it as \textit{S. obliquus} to keep consistent with the papers we are
comparing our results to.

That several top performers independently isolated were strains of *S. obliquus* was not surprising. Strains of *S. obliquus* have been known to grow well in mass cultures – even demonstrating potential in outdoor raceway ponds in the developing world (Becker, 1984). Strains of the genus *Scenedesmus* have been found to be able to accumulate lipids even under colder temperatures (10°F), which is important as winter algal strains are needed for biofuel production (Li *et al.*, 2011). Other studies have found that *S. obliquus* has a high tolerance and growth rates under elevated levels of CO₂, suggesting that the species has potential to reduce CO₂ in the flue gases emitted by thermoelectric power plants (Tang *et al.*, 2011).

In bubbling columns with 3% CO₂ at 25°C, biomass productivity for *S. obliquus* strains was ~1 g L⁻¹ d⁻¹ (Rodolfi *et al.*, 2009), which is in our range of productivity results (Figure 2). Also an interesting comparison can be found in a study looking at eight microalgae from a total of 33 isolated cultures of water samples from freshwater rivers and livestock wastewater treatment plants at Wunju, South Korea (Abou-Shanab *et al.*, 2011). One of the highest performers was a strain of *S. obliquus*. In that study, among isolated *S. obliquus* strains, which were the same species according to LSU rDNA D1-D2 regions, significant differences in lipid productivity existed. The *S. obliquus* YSR01 strain was able to produce lipids at 0.91 ± 0.36 grams L⁻¹, while the other strains were significantly lower – with *S. obliquus* YSR04 at 0.40 ± 0.03, *S. obliquus* YSR05 at 0.47 ± 0.05, and *S. obliquus* at 0.49 ± 0.06 (Abou-Shanab *et al.*, 2011). Such different productivities within a species from the same geographical region demonstrate that surveying culture collections may not be sufficient, as different strains of the same
species may have different biomass or lipid productivities. Similarly, showing that different strains of the same species can have widely varied physiological parameters, some *S. obliquus* isolates have been found to have 3-4 times higher carbon fixation efficiencies than other *S. obliquus* isolates (Ho et al., 2010). The strain of *S. obliquus* that we isolated not only grew well under laboratory conditions, but also outdoors even when compared to *N. salina* (NAABB Final Report, 2014).

Another group of strains (DOE0088, DOE0155, DOE0202, DOE0369, EN1234, EN1235) nested closely together in a clade with *Coelastrella sp.* strain SAG217.5 (Hegewald et al., 2010, Kaufnerova & Elias, 2013) being the sister group (Figure 6). To gain a more comprehensive taxonomic overview, a second and specific phylogenetic analysis was performed including all publically available ITS2 sequences of *Coelastrella* strains from the NCBI database. Our results showed that our six novel strains fell into one clade together with other *Coelastrella* strains SAG2123 and SAG217-5 (Figure 6).
Figure 6: Non-rooted ITS2 sequence tree including our novel strains DOE0088, DOE0155, DOE0202, DOE0369, EN1234, EN1235, and also known strains of the Scenedesmaceae.
Based on microscopic investigation, we found that the new *Coelastrella* strains had a broad range of coccoid cell morphologies. As one example, images of the strain
DOE0202 are presented in Figure 7. Flagellate cells were never observed under a range of conditions. Reproduction occurs through division of autospores, which release two to eight daughter cells by rupture of the parental cell wall. Depending on the growth conditions, cells were either solitary or found in coenobia (Figure 7C). Solitary cells were round to ellipsoid and 3-10 µm long. When ellipsoid, cells often had polar thickenings. Cells contained one chloroplast with one pyrenoid. In growing cultures, vegetative cells were green. Resting cells from cultures in the stationary phase became first brown and then orange due to accumulation of orange colored oil bodies in the cytosol (Figure 7). Regardless of growth conditions cell walls of all cells were hyaline. As observed by light microscopy, cell walls were mostly smooth. Only sometimes, during the exponential growth phase in batch cultures, coenobia were found that looked similar to coenobia of Scenedesmus species. Individual cells of these coenobia had ridges appearing as protruberances. Such variability in the cell wall appearance and ridges is in agreement with the description of Coelastrella sp. strain SAG 217-5 (Hegewald et al., 2010), which is a close relative according to our recovered ITS2 phylogenetic tree (Figure 7).

All our newly isolated Coelastrella strains originated from different freshwater habitats from a variety of locations within the United States (Table 1), indicating that this Coelastrella species has a broad distribution within the United States (Table 1). In their habitats such as birdbaths, fountains, and temporary waterbodies (roadside ditches), the Coelastrella strains were often found co-existing with strains of the green alga Haematococcus pluvialis. Similar to H. pluvialis, our new isolates were also carotenogenic and formed spores that survive desiccation. To the best of our knowledge, this is the first description of Coelastrella strains isolated from the United States, thus all
these *Coelastrella* isolates appear to be novel strains. These novel *Coelastrella* strains not only performed well in the laboratory, but strain DOE0202 also tested positively in small raceway-type ponds (NAABB Final Report, 2014), indicating that – similar to *H. pluvialis*— the new *Coelastrella* strains might be useful as new platform strains for biofuels and/or bioproducts generation.

Several of the well performing strains (DOE1418, DOE1357, DOE1051, DOE0043) appear to be of the *Desmodesmus* genus. This was confirmed by the strains cell morphology as oval cells also exhibit the characteristic spines at the ends of their cenoeiba (Vanormelingen et al., 2007). For the genus of *Desmodesmus*, detailed species level studies comparing morphology and ITS2 rDNA phylogenies are lacking to date (Vanormelingen et al., 2007), except for an account on four closely related species bearing lateral spines (Hegewald et al., 2005). Although studies analyzing the *Desmodesmus* genus with regards to biofuels are sparse, some species within this genus, such as the ones we isolated, may be excellent platforms for biofuel production.

The strain DOE0259 falls into one clade with *Ankistrodesmus* sp. SP2-15 (Figure 4), although quite some evolutionary distance exists between the two. Currently, only three other *Ankistrodemus* strains besides DOE0259 have their ITS2 region sequences in the NCBI database. All three strains showed similar far distances from our newly isolated strain, so only one was included in the tree shown in Figure 4. Extending a search in the NCBI database with 18S, 5.8S, and ITS1 sequences from DOE0259 (Genebank Accession Number KT274017) also identified this strain as an *Ankistrodesmus*, with it sharing the highest percent identity (99%) with *Ankistrodesmus* RS-2012, although with low query coverage (59%). Lack of exact matches for the 18S gene, the ITS1 spacer, the
5.8S gene, or the ITS1 spacer for strain DOE0259 suggests that it may represent a new species. Unfortunately, the genus of *Ankistrodesmus* is polyphyletic (Krienitz et al., 2001, Krienitz et al., 2011b). It is also difficult to discern the different crescent-shaped morphological features of species, which are highly variable and challenging to identify by light microscopy (Heynig & Krienitz, 1982, Nygaard et al., 1986). Figure 8 shows images of cells of the strain DOE0259. The crescent-shaped cells were always solitary. The basic cell morphology is coccoid with cells being elongated sickle-shaped having rounded ends (Figure 8). Due to the known uncertainty regarding the taxonomic position of species within the genus *Ankistrodesmus* and the family of the Selenastraceae in general (Krienitz et al., 2001, Krienitz et al., 2011a), any further classification of strain DOE0259 was beyond the scope of this manuscript.
Figure 8: Shown are images of cells of the novel *Ankistrodesmus* strain DOE0259. The black bar represents 10 µm. A) Two cells accumulating numerous oil bodies in the cytosol. B) Images of several cells illustrating the plasticity in cell morphology. C) Single cell. D) Two cells in the early division process. E) Two cells released from an autosporangium. F) Four daughter cells still in the autosporangium. G) Four cells released from the autosporangium.

3.4.2. *Chlamydomonadales (Class of the Chlorophyceae)*

Two taxa, strain DOE0101 and strain EN1423, fell into the Chlamydomonadales order (Figure 4). Strain EN1423 was isolated from hyper-saline soil from a roadside ditch close to the Laguna Del Perro in New Mexico, USA (Table 1). Originally the strain was maintained in 0.5 M NaCl containing artificial *Dunaliella* medium and also tested for biomass productivity in bubbling columns. However, in the saline medium large, sticky cell clumps developed that clung to the glass vessel, thus preventing accurate productivity determination. Later the strain was transferred to BG11 freshwater medium and found to grow well in that medium. The rDNA sequence containing the 3’ end of the 18s gene, ITS1, 5.8s gene, ITS2, and the 5’ end of the 28s gene (Accession Number KT274016) was used for a BLASTn search into the NCBI database. This search resulted in a hit with 100% identity for our partial 18s gene sequence to the species *Borodinellopsis texensis* (NCBI Accession number KM020129). In addition to the molecular marker sequence, the cell morphology and physiology of strain EN1423 was investigated. Figure 9 shows images of cells of strain EN1423. The previous description of *Borodinellopsis texensis* (Dykstra, 1971) asserts that zoospores are present, but only motile for very short periods or that autosporangia divide to produce aplanospores, making observation of motile zoospores difficult. This could be why no biflagellate zoospores were observed in our EN1423 cultures. Depending on the culture conditions,
cells of EN1423 were 2-20 µm in diameter. Larger, mature cells had thick hyaline walls and cells often came in diad, triad, or tetrahedral stages (Figure 9). Chloroplasts were of asteroidal nature. Overall, cells of EN1423 had the morphological characteristics previously described (Dykstra, 1971). In addition, in older cultures cells turned orange (Dykstra, 1971) in BG11 medium (=freshwater), but brick-red in 0.5M NaCl artificial Dunaliella medium. Based on the molecular marker, cell morphology, and the carotenogenic characteristic strain EN1423 was classified as *Borodinellopsis texensis*. To our knowledge, this is the first report of this species having potential as a platform strain for biofuel or bioproduct generation.

Figure 9: Shown are photos of cells of *Borodinellopsis texensis* strain EN1423 from cultures grown in 0.5 M NaCl artificial Dunaliella medium. The black bar represents 10 µm. A) Shown is one autosporangium. B) A single vegetative cell with the asteroidal chloroplast originating out of the pyrenoid. C) A variety of cells in different growth
stages. D) Cells that are in the process of carotenogenesis by accumulation of numerous brown/orange oil bodies in the cytosol oriented towards the outside of the cells. E) Mature cells, each with large numbers of orange/red oil bodies in the cytosol making the chloroplast invisible.

While strain DOE0101 may not have been the very best biomass producer in the laboratory (see section 3.2. above), this strain was highly carotenogenic (see section 3.3. above) and it was successfully tested in small race-way type ponds (NAABB Final Report, 2014). Classification of strain DOE0101 to the species or genus level remained challenging. From our genome sequences (unpublished) of the strain DOE0101, we located the 18S gene (GenBank Accession Number AJ249515), the 28S gene (GenBank Accession Number KC145458), the psaB gene (GenBank Accession Number JN63055), and the rbcL sequence (GenBank Accession Number KC145509) and BLASTed them within NCBI. For these four molecular markers, the best hits were to *Dysmorphococcus globosus* (SM 1), which is in the Chlamydomonadales order, but it has no ITS2 sequence available. Moreover, for DOE0101’s psaB and rbcL genes, the query coverage with *Dysmorphococcus globosus* was only 85 and 86 %, respectively. *Dysmorphococcus* is distinguished by having a porous, ornamental envelope and two flagella apertures (Porcella & Walne, 1980, Takeda, 1916). However, regardless of our cultivation conditions, in none of the cells of strain DOE0101 these *Dysmorphococcus* cellular characteristics were ever observed by light microscopy (Figure 10). Therefore, though the ITS2 sequence-structure phylogeny places DOE0101 within the Chlamydomonadales, its identity at a finer taxonomic level remains unclear.
Figure 10: Shown are images of cells of the strain DOE0101 indicating the variety of cell types. The bar indicates 10 µm. A) Vegetative, coccoid cell beginning division of the pyrenoid. B) Several cells each containing one cup-shaped chloroplast and one pyrenoid per chloroplast. C) One large cells in the beginning phase of carotenogenesis and one autosporangium. D) Single cells with orange oil bodies. E) Orange cells with large oil bodies and unknown white appearing structures. F) Two autosporangia originating from orange cells.

In summary, while *Borodinellopsis texensis* strain EN1423 and the Chlamydomonadales strain DOE0101 are in the same order as the species *H. pluvialis*, in contrast to flagellate green cells of *H. pluvialis*, non-stressed green cells of the novel strains were always coccoid under our growth conditions. Also in contrast to *H. pluvialis*, which is well-known for application in mass production of the brick-red carotenoid astaxanthin, cells of DOE0101 appear orange under stress (Figure 3 & 10) with the orange pigments in DOE0101 having been identified as precursors of astaxanthin (Figure
3). As the strains EN1423 and DOE0101 accumulate secondary carotenoids, both strains may find further use in future applications for carotenoid production.

3.4.3. Chlorellales (Class of the Trebouxiophyceae)

Green algae with the spherical morphotype named Chlorella belonging to the class of the Trebouxiophyceae have been used as model organisms for studies of photosynthesis and biotechnological applications for decades (Krienitz et al., 2015). As shown in Figure 4, many of our top performing strains clustered within the class of the Trebouxiophyceae. However, the positions of many of the new strains within the order of the Chlorellales was not too well resolved when the rDNA ITS2 marker was used (Figure 4). The best support for identification was obtained for strain DOE1135, which appears to be closely related to Chlorella vulgaris. DOE0314 falls with Dicloster acutus, though a significant distance exists between the two. The most internal branches of the other new strains were also not well supported, but in the terminal parts of the phylogenetic tree they clearly nested into highly supported clades. As the phylogeny based only on the rDNA ITS 2 spacer region was not providing sufficient information regarding the taxonomic position at the genus or even species level, an additional molecular comparison was performed using the rDNA region including the 3’ end of the 18s gene, the ITS1 region, the 5.8s gene, the ITS2 region, and 5’ end of the 28s gene (Bock et al., 2011). The resulting phylogenetic analysis allowed classification of the strain DOE1412 as C. sorokiniana.
Figure 11: Phylogenetic tree based on the rDNA region including the 3’ end of the 18s gene, the ITS1 region, the 5.8s gene, the ITS2 region, and 5’ end of the 28s gene. Sequences were aligned with ClustalW and reconstructed using Maximum Likelihood with a Tamura-Nei Model and Nearest-Neighbor-Interchange (NNJ) heuristic method, with 1000 bootstrap replications performed. The sequences for the strains *Chlorella sorokiniana* UTEX 1230 and *Chlorella sorokiniana* LANL 1228 were provided by Dr. Starkenburg.

In the past few years the classification of coccoid taxa within the Chlorellaceae including the genus of *Chlorella* has receive a lot of attention and many species were re-classified (Bock et al., 2011, Krienitz et al., 2015, Krienitz & Bock, 2012, Lemieux et al., 2014, Luo *et al.*, 2010). Overall, it is now accepted that molecular and morphological characters are used for delineation of the coccoid species (Bock et al., 2011, Krienitz & Bock, 2012). Consequently, we included images of cells of *C. sorokiniana* strain DOE1412 for reference in Figure 12 below.
Figure 12: Shown are photos of cells of *Chlorella sorokiniana* strain DOE1412. The black bar represents 10 µm. A) Autosporangia B) Some autosporangia and several cells in the early stages of cell division. C) Several larger cells with the pyrenoid clearly visible. D) Several autosporangia and one larger cell. One autosporangium and an opened autosporangium with released daughter cells. F) Larger cells with the pyrenoids clearly visible within the chloroplast.

Although strain DOE1412 is not a representative of a new species and concern about redundancies regarding already existing strains in culture collections may exist, the first analysis of a draft genome of strain DOE1412 revealed that its genome contains significant differences to two other *C. sorokiniana* strains – UTEX1230 and LANL1228 (Starkenburg, 2015). Therefore, it can be concluded that – at the very least – the new isolate strain DOE1412 greatly expands the germplasm of the species *C. sorokiniana*. Comparative genomic work is ongoing, but an analysis of the results is beyond the scope of this manuscript.
Within the Chlorellales clade, *C. sorokiniana* DOE01412 proved to be one of the best growers of the first strains isolated during the NAABB project and was brought forward as a top producer. The fatty acid profile for *C. sorokiniana* DOE1412 was obtained and analyzed for a culture grown in a photobioreactor (Figure SM-1) showing that it had C16:0 and C18:1 fatty acids as the main constituents, which are common in green algae and present in biodiesel as methyl esters.

It is of note that *C. sorokiniana* strain DOE1412 was found to grow well in LED-lighted 800L indoor raceway ponds (Huesemann *et al.*, 2013). Also, when later tested outdoors, it grew very well, significantly better then even gold standard *N. salina*. In 23,000 L NAABB outdoor raceway ponds with paddle wheels, it had a maximum productivity of 30g/m$^2$/day, tolerating temperatures from 40$^\circ$-110$^\circ$F, withstanding a range of salinities, and accumulating about 25% lipids. The strain also faired well in economical models, and oil from it was successfully converted to jet fuel. Consequently, the discovery of DOE1412 was cited as a major deliverable for the NAABB project (NAABB Final Report, 2014).

Like *Scenedesmus*, *Chlorella* strains have also been of interest as biofuel feedstocks. *C. sorokiniana* is regarded as one of the most promising algae feedstocks, with some strains having long-straight chain alkanes and fatty alcohols being major compounds of extracts (Yang *et al.*, 2013). The algae can grow in autotrophic, heterotrophic, and mixotrophic conditions (Kim *et al.*, 2013). Similarly, the related species *Chlorella vulgaris* has shown potential to produce biofuels from high concentration bioethanol wastewater, and simultaneously performs wastewater treatment (Xie *et al.*, 2013). Our new strains may offer an expansion of the genetic resources
available to create even more productive algae from this clade.

4. CONCLUSIONS

This work identified and classified the top algal strains to come out of a multi-year screening effort to find algae suitable for further development as biofuel feedstocks. Thirty of the best performing strains were deposited with the UTEX algal culture collection. The project successfully isolated several strains, which showed potential as future platform strains: One new strain designated as DOE0101 was from the Chlamydomonadales. In addition, several novel strains – including DOE0202 – belonging to the Coelastrella genus were found, which are reported on for the first time for the United States. Both, the Chlamydomonadales strain DOE0101 and all the Coelastrella strains accumulate carotenoids. That these novel strains appear previously uncharacterized in a biofuel context despite their high growth rates and carotenoid-rich profiles demonstrates the untapped diversity of the green microalgae. In addition, several new strains of species already known to be fast growers were isolated, and they now offer an expansion of the genomic resources available for these clades. Lastly, the discovery of C. sorokiniana strain DOE1412 is particularly noteworthy as it is currently one of the most promising algal species being developed for biofuels applications. Understanding the evolution of the highly productive strains isolated in this study and characterizing their metabolic and growth traits is a key step in bringing forth cultivars of microalgae for biofuel, and high-value compound, production.

Contributions
J.P. designed and oversaw the entire isolation and screening project. P.N., J.P., and A.H. conceived and wrote the paper. A.H., W.C. and F.J. conducted the screening, P.N., J.P., S.T., and F.J., conducted the lipid analysis, and P.N., J.P., K.S., and Q.W. conducted the phylogenetic analysis. Light microscopy for many strains was performed by A.G. and the Transmission Electron Micrograph for *Coelastrella* strain DOE0202 was provided by O.H. The fatty acid profile for *C. sorokiniana* strain DOE1412 was provided by S.T. Together, J.P. (jpolle@brooklyn.cuny.edu) and P.N. (pneofotis@gc.cuny.edu) declare the integrity of this work as a whole.

SM-Table 1: Best nucleic acid matches of several barcode genes from the Chlamydomonadales species strain DOE0101. Genes were located within the draft genome of strain DOE0101 and then blasted within the NCBI database.

<table>
<thead>
<tr>
<th>Gene</th>
<th>DOE0101 IsoTig</th>
<th>Best Hit</th>
<th>Query Cover</th>
<th>E Value</th>
<th>Identity</th>
<th>Hit Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>18S</td>
<td>NODE1118, Length 418</td>
<td><strong>Dysmorphococcus globosus strain SAG 20-1 18S ribosomal RNA gene, partial sequence</strong></td>
<td>100%</td>
<td>0</td>
<td>96%</td>
<td>KM020136.1</td>
</tr>
<tr>
<td>28S</td>
<td>NODE 1514, Length 1003</td>
<td><strong>Dysmorphococcus globosus strain UTEX LB 65 26S ribosomal RNA gene, partial sequence</strong></td>
<td>100%</td>
<td>0</td>
<td>97%</td>
<td>DQ015755.1</td>
</tr>
<tr>
<td>psaB</td>
<td>Node 16537, Length 1132</td>
<td><strong>Dysmorphococcus globosus chloroplast psaB gene for P700 chlorophyll a apoprotein A2, partial cds, strain:</strong></td>
<td>85</td>
<td>0</td>
<td>95%</td>
<td>AB451208.1</td>
</tr>
<tr>
<td>rbcL</td>
<td>NODE 3207, Length 734 \ cov 3022.291504</td>
<td><strong>Dysmorphococcus globosus gene encoding ribulose bisphosphate carboxylase large subunit (strain SAG 20-1)</strong></td>
<td>86</td>
<td>0</td>
<td>98%</td>
<td>AJ001885.1</td>
</tr>
</tbody>
</table>
SM-1: Fatty Acid Profiles of DOE1412 grown in a bioreactor for 16 days.
SM-2: TAG Profiles of some of the top performing strains isolated in this study.

Supplemental Material Notes

SM – Note 1: For the first-level screen, in summary, batch cultures of strains were grown in Erlenmeyer flasks in defined minimal media suitable for this selection: C Medium, Bold Basal Medium, and BG11 Medium (Andersen & Kawachi, 2005), with the only CO₂ enrichment being the sodium bicarbonate in the BG11. Flasks were grown under 50 µE m⁻² s⁻¹ continuous light provided by daylight fluorescent lamps. On a weekly basis, samples were taken to determine growth by using the absorption at 750 nm as a proxy for biomass. Concurrently, the relative lipid contents of cells in culture was determined by measuring the fluorescence emission of the lipophilic fluorescent dye Nile Red (9-diethylamino-5H-benzo[a]phenoxazine-5-one) (Cooksey et al., 1987, Cooper et
al., 2010) at 590/35 nm. The unicellular green alga *Chlamydomonas reinhardtii*, which did not accumulate significant amounts of lipids, was used as an *in vivo* control.

Regarding the potential to over-accumulate carotenoids, all batch cultures of the algal strains were also examined for changes in color. In particular, cultures which turned yellow, orange or red – which indicate the accumulation of carotenoids – were further inspected through bright field and Nile Red fluorescence microscopy to gauge a visual understanding of lipid and carotenoid accumulation. Based on visual examination, the most promising of the carotenoid accumulating strains were advanced to the second level screening regardless of their OD$_{750}$ readings.

Strains that passed the first-level screen were deemed to be potential producers and were selected for further characterization (Figure 1). In a second-level screen the strains were cultivated in multiple different media to determine if another growth medium was better suited for use in continued strain characterization.

In a third-level screen for analysis of actual biomass and lipid production potential, strains were grown at 28 Degree Celsius in glass columns (3 cm width with 500 mL volume) into which CO$_2$ enriched air was bubbled through the cultures from the bottom. A concentration of 1% CO$_2$ was chosen because experience and published experiments (Azov, 1982, Fang *et al.*, 2012) have shown the concentration to be high enough to suppress any of the CO$_2$ concentration mechanisms (CCM) that many microalgae possess (Reinfelder, 2011, Raven *et al.*, 2011, Fang *et al.*, 2012). The columns received 24-hour irradiance using fluorescent daylight lamps at about 200 μE m$^{-2}$ s$^{-1}$. 
From pre-cultures that were in their light-limited growth phase, 30 ml of the pre-culture from a flask was used to inoculate a glass column of 300 ml total liquid media. Biomass was determined daily by measuring the ash-free-dry-weight (AFDW). To measure the AFDW, aliquots were taken and biomass was filtered onto 47 mm glass fiber microfilters (Whatman-GE Healthcare; Pittsburg, PA), which had previously been ashed for one hour at 550°C in a Isotemp muffle furnace (Fisher Scientific; Waltham, Massachusetts) and then weighed. The pre-ashed filter with the sample on it was placed overnight into a drying oven at 110°C and then weighed to obtain the dry biomass weight. Following the filter with dry biomass was placed for one hour into the furnace at 550°C to incinerate all organic biomass. Following incineration of the organic biomass the filters with minerals were weighed again. The AFDW was then calculated by subtracting the weight of the minerals from the dry biomass weight. When the algal cultures reached their stationary phase, biomass also was taken and analyzed via a gravimetric method to determine the total lipid content (Bigogno et al., 2002).

SM Note 2: Briefly, 10 ml of algae culture were concentrated by centrifugation and then re-suspended in 5 ml of chloroform, 2.5 ml of methanol, and 2.5 ml of water (2:1:1), and then secured onto a shaker (200 rpm) in the dark for one hour. The glass tube containing the mixture was centrifuged, and the top liquid layer was aspirated off. Following, the bottom liquid layer was transferred to a new glass tube under nitrogen protection and dim light in a 65°C water bath to speed the evaporation of the chloroform. After the evaporation of the chloroform, the residue was then re-suspended in acetonitrile, and filtered through a syringe (3mm PTFE, 0.5 μm pore size).
Chapter 3: The Plastidic Enolase and its Presence in an Oleaginous

Microalgae

Jürgen Polle\textsuperscript{a,b} and Peter Neofotis\textsuperscript{a,b}

\textsuperscript{a}Department of Biology, Brooklyn College, 2900 Bedford Avenue, Brooklyn, New York 11210, USA

\textsuperscript{b}The Graduate Center of the City University of New York, New York, NY 10016, USA

ABSTRACT

Central carbon core metabolism in higher plants involves multiple isoforms of the glycolytic enolase enzyme, coded for by different genes, which could be differentially regulated. The exact mechanisms underlying the distribution of fixed carbon within photosynthetic eukaryotic cells (i.e. carbon partitioning) are still unknown. Part of what is unknown includes the number of genes coding for isozymes and their subcellular localization. Here we report on the number of gene copies coding for the enolase in several genomes of species spanning the major classes of green algae. Our genomic analysis of most green algae revealed the presence of only one gene coding for a glycolytic enolase [EC 4.2.1.11]. In that case, our predicted cytosolic localization would require export of organic carbon from the plastid to provide substrate for the enolase and

\textsuperscript{5} Large portions of this manuscript were updated and revised from an earlier published paper cited as

subsequent re-import of organic carbon back into the plastids. The role a plastid localized enolase, which is present in higher plants and, in contrast to the majority of the investigated green algae, also in *Scenedesmus obliquus* strain DOE0152, is discussed.

**INTRODUCTION**

Understanding of the molecular mechanisms underlying regulation of carbon partitioning is fundamental for genetic engineering and development of novel strains of green algae that, for example, will have improved growth and lipid productivities. However, currently a major gap in knowledge exists regarding the molecular mechanism(s) of regulation of carbon partitioning in algal cells, including the green algae. This lack of knowledge may still be a relict from the historical vision of green algae as simply single celled plants. Consequently, we focus here on the contrast between higher plants and green algae.

Figure 1: General overview of carbon metabolism in plants, with particular attention paid to glycolysis.
Figure 1 shows a general overview of the path of carbon leading from fixation of CO$_2$ to the different major cellular compartments, referred to as ‘carbon partitioning’, in a phototrophically grown green algal cell. Carbon partitioning in cells involves multiple pathways, together representing a carbon metabolic network. In algae, the core carbon metabolism includes the plastid localized Calvin-Benson cycle (=Reductive Pentose Phosphate Pathway). In the light, operation of the Calvin-Benson cycle ‘begins’ with fixation of CO$_2$ through the enzyme Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO, EC 4.1.1.39) into glycerate-3-phosphate. Then glycerate-3-phosphate can be used for conversion into multiple different cellular macromolecules (Figure 1). In general, following reduction of glycerate-3-phosphate into triose phosphates, condensation of two triose phosphate molecules can create hexose phosphates that may either be used for regeneration of ribulose-1,5-bisphosphate or for the synthesis of starch. Also, either triose phosphates or hexose phosphates could leave the Calvin-Benson cycle to be exported from the plastid for use in other cellular compartments. In the context of the core carbon metabolism, specifically important for cells is the connection of carbon fixed in the Calvin-Benson cycle to creation of pyruvate, which is needed in cells as a precursor to form a variety of macromolecules including glycerolipids and isoprenoids. This connection between glycerate-3-phosphate to pyruvate synthesis is determined by the operation of the reactions of glycolysis/gluconeogenesis, and the cellular localization of the essential enzymes. Depending on the organism, various alternative options for conversion of glycerate-3-phosphate to pyruvate may exist for organisms, but metabolic constraints drive certain possible paths (Milo & Last, 2012)
The enolase is a multi-functional protein (Diaz-Ramos et al., 2012, Gerlt et al., 2011, Huberts & van der Klei, 2010). It operates not only as an essential phosphopyruvate hydratase (EC 4.2.1.11) enzyme in glycolytic catabolism or in anabolic gluconeogenesis converting reversibly 2-phospho-D-glycerate (2PGA) into PEP, but the enolase protein was also reported to have non-glycolytic functions making it a moonlighting protein (Henderson & Martin, 2013, Gerlt et al., 2011, Gomez-Arreaza et al., 2014, Huberts & van der Klei, 2010). For example, the enolase protein was implicated as a cell surface receptor in a wide range of organisms from bacteria (Carneiro et al., 2004, Ehinger et al., 2004, Pancholi & Fischetti, 1998) to eukaryotic parasites (Swenerton et al., 2011, Vanegas et al., 2007, Castillo-Romero et al., 2012). It can be a structural component of animal eyelenses (Mathur et al., 1992), or – as an alternate translation product – a partial enolase protein can operate as a transcriptional repressor (Subramanian & Miller, 2000, Ghosh et al., 1999, Kang et al., 2013). In addition to the enolase protein having different functions within one cell, an enolase superfamily exists with not all predicted gene products working as enolases. One example is the recent discovery of family members which function as D-mannonate and d-gluconate dehydratases (Wichelecki et al., 2014).

In our comparative approach to study carbon core metabolism in green algae, we use the known metabolism of higher plants as a basis to expand our understanding. Higher plants such as the model A. thaliana have a gene family coding for the enolase enzyme (www.arabidopsis.org), with the gene products being localized to different cellular compartments (Andriotis et al., 2010). Further, it is known that metabolism in higher plants is differentially regulated at the organelle level based upon the presence or
absence of a family of enolase enzymes. In the model plant *A. thaliana*, overall four nuclear genes code for enolases:

2. An enolase (ENO2 also called LOS2), which is very similar in sequence and structure to those of animals and yeast (Vanderstraeten *et al.*, 1991) and most likely localizes to the cytosol and as an alternative translation product (MBP-1) to the nucleus where it acts as transcriptional suppressor (Ghosh *et al.*, 1999, Kang *et al.*, 2013).
3. An enolase localized to the cytosol (ENOc (Voll *et al.*, 2009)),
4. A multi-functional enolase (DEP1) (Rutschow *et al.*, 2008), which was detected in the chloroplast stroma in *A. thaliana* (Rutschow *et al.*, 2008). DEP1 in Arabidopsis functions in the methionine biosynthesis (Pommerrenig *et al.*, 2011) performing the reaction of the 5'-methylthioribulose-1-phosphate dehydratase (EC 4.2.1.109) and of the Enolase-phosphatase E1 (EC 3.1.3.77). The DEP1 enolase activity converts 2,3-diketo-5-methylthiopentyl-1-phosphate into the intermediate 2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate, which is then dephosphorylated. DEP1 is an essential enzyme in cellular sulfur metabolism in plants (Sauter *et al.*, 2013).

As pyruvate is a central precursor metabolite for many different cellular products such as glycerolipids, isoprenoids, and amino acids, the cellular localization and activity of the enolase enzymes may be at the crossroad for carbon partitioning in plant cells. Furthermore, the cellular localization of enzymes can play a strong role in the regulation of their activity. For the enolase, it is usually perceived as present and active in the cytosol of animal, yeasts, and plant cells. For animals and yeasts, this localization
coordinates with the only source of triose phosphate, which is formed when glucose or fructose is broken down.

Plant cells, on the other hand, also have triose phosphate available at the end of the reduction phase of the Calvin-Benson cycle, which happens in the chloroplast. In order for the triose phosphate to be utilized by the cytosolic enolase, it must be exported out of the chloroplast. In the cytosol, it is then converted by the triose phosphate isomerases into glyceraldehyde-3-phosphate. Though it is known that plants also have a plastid form of the enolase, the functional advantage of this has not been elucidated. Nor was it clear at the onset of this research whether the algae have the plastidial form. Even in the best-studied green model alga *C. reinhardtii* the exact path of carbon remains unresolved. For example, some metabolic models contain the enolase activity present in the plastid (Winck *et al.*, 2013, Dal'Molin *et al.*, 2011), other models have the enolase absent from the plastids (Johnson & Alric, 2013, Kliphuis *et al.*, 2012), or compartmentalization was not taken into account (Christian *et al.*, 2009). It is, therefore, still unknown if one or two complete glycolytic pathways exist in the green algae. In this work, we analyzed the genomes of several green algae to detect the presence/absence of enolase copies to help shed light on the distribution and function of the enolase enzyme in photosynthetic organisms. In addition, the transcriptome of the oleaginous green microalgae *Scenedesmus obliquus* DOE0152 was obtained and analyzed.

**METHODS**

We investigated the genomes of several species belonging to core chlorophytes such as *Chlamydomonas reinhardtii* (Merchant *et al.*, 2007), *Volvox carteri* (Prochnik *et
al., 2010), *Astrochloris* sp., *Chlorella variabilis* NC64A (Blanc et al., 2010), and *Coccomyxa subelipsoidea* C-169 (Blanc et al., 2012), as well as prasinophytes represented by *Micromonas pusilla* strains RCC299 and CCMP1545 (Worden et al., 2009b), *Ostreococcus lucimarinus* and *Ostreococcus tauri* (Palenik et al., 2007). In addition to sequences obtained from the US Department of Energy Joint Genome Institute (JGI) database, for our analysis we included the enolase protein sequence deduced from genomic and transcriptomic data from the unicellular green alga *Dunaliella salina* (www.jgi.doe.gov, project #16719 and project #1014865). We also included the unpublished transcriptome of *Scenedesmus obliquus* strain DOE0152_Z in our analysis. This is a sub clone of strain DOE0152, which had come out of the NAABB project and has been shown to perform well in the laboratory (Neofotis et al., In revision) as well as outdoors (NAABB Final Report, 2014). The transcriptome was obtained via a collaboration with Pacific Northwest National Laboratory and assembled using Oases (Version 0.2.08) in the Polle laboratory.

To learn more about the possible function and cellular location of the green algal enolase(s), we compared the algae enolase protein sequences with the known enolase proteins of the higher plant *Arabidopsis thaliana* – as well as the one found in the human genome. First, we aligned the green algal sequences together with the three sequences obtained from the higher plant *A. thaliana* (www.arabidopsis.org - ENO1 (AT1G74030), ENO2 (AT2G36530), ENOc (AT2G29560)) and with the human alpha-enolase using ClustalW (Thompson et al., 1994). As now more full-length green algal enolase sequences were available than for previous studies, we performed a new phylogenetic analysis. The phylogeny tree was reconstructed in Mega 5.2.2 using a Maximum
Likelihood statistical method with 500 bootstraps. A Jones-Taylor-Thorton (JTT) model with uniform rates of substitution was used with a Nearest-Neighbor-Interchange (NNJ) heuristic method. The sub-cellular localization of the proteins were also predicted by using the program PredAlgo (https://giavap-genomes.ibpc.fr).

RESULTS AND DISCUSSION

Number of genes

Table 1 summarizes our results of the comparative genomic analysis of the search for gene copies present in green algae. Similar to its close relatives C. reinhardtii and V. carteri, the genome of D. salina only contained one enolase. The Trebouxiophytes Asterochloris sp. and Chlorella variabilis NC64A had only one gene coding for enolase. In contrast, the Trebouxiophyte C. subelipsoidea C-169 contained two genes coding for identical enolase proteins and further examination of the genomic sequences showed that they were identical as well, indicating a more recent gene duplication event. The prasinophytes M. pusilla strains RCC299 and CCMP1445 as well as O. lucimarinus and O. tauri had each one copy coding for the enolase. In addition to our results with species of Micromonas and Ostreococcus, the presence of two enolase gene copies had previously been reported for the prasinophyte Pycnococcus provasolii (Harper & Keeling, 2004). Although the sequences were not complete due to missing N-terminus and C-terminus, the alignment of the two enolase protein sequences of P. provasolii (Enolase 1 - GenBank #AAL05456.1; Enolase 2 - GenBank #AAL05457.1;) showed that they were not identical (Data not presented) and three insertions found in the enolase 1 of P. provasolii indicated plasticity within the species. In brief, one or multiple enolase genes could be found in the genomes of green algae.
Table 1: Information on the genes coding for the enolase in green algae as extracted from the database of the US Joint Genome Institute (as of 05/23/2014).

<table>
<thead>
<tr>
<th>Species</th>
<th>JGI Protein ID</th>
<th>JGI Locus Name &amp; Location</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Asterochloris sp.</em> v1.0</td>
<td>35232</td>
<td>scaffold_00088:18492-25512 [JGI Genome Portal]</td>
<td>Glycolytic enolase</td>
</tr>
<tr>
<td><em>Chlamydomonas reinhardtii</em></td>
<td>136652</td>
<td>Cre12.g513200.t1.2; chromosome_12: 3521433 - 3525524 [Phytozome v9.1]</td>
<td>Glycolytic enolase</td>
</tr>
<tr>
<td><em>Chlamydomonas reinhardtii</em></td>
<td>NA</td>
<td>g393; chromosome_1: 2781663 – 2787514 [Phytozome v9.1]</td>
<td>Multi-functional enolase</td>
</tr>
<tr>
<td><em>Chlorella variabilis</em> NC64A</td>
<td>136652</td>
<td>scaffold_17:245647-250804 [JGI Genome Portal]</td>
<td>Glycolytic enolase; Gaps in genomic sequence</td>
</tr>
<tr>
<td><em>Coccomyxa subelipsoidea</em></td>
<td>38308</td>
<td>fgenesh1_pm.19._#_130; scaffold_19: 1328456 – 1334257 [Phytozome v9.1]</td>
<td>Glycolytic enolase</td>
</tr>
<tr>
<td><em>Coccomyxa subelipsoidea</em> C-169</td>
<td>35576</td>
<td>fgenesh1_pm.3._#_251; scaffold_3: 2788580 – 2794381 [Phytozome v9.1]</td>
<td>Glycolytic enolase</td>
</tr>
<tr>
<td><em>Micromonas pusilla RCC299</em></td>
<td>107587</td>
<td>Chr_01: 1470509 – 1472948 [Phytozome v9.1]</td>
<td>Glycolytic enolase</td>
</tr>
<tr>
<td><em>Ostreococcus lucimarinus</em></td>
<td>28765</td>
<td>Chr_1: 271278 – 273263 [Phytozome v9.1]</td>
<td>Glycolytic enolase</td>
</tr>
<tr>
<td><em>Ostreococcus tauri</em> v2.0</td>
<td>27349</td>
<td>Chr_01.0001:228132-230031 [JGI Genome Portal]</td>
<td>Glycolytic enolase</td>
</tr>
<tr>
<td><em>Volvox carteri</em></td>
<td>79991</td>
<td>Vocar20013958m.g; scaffold_6: 3504767 – 3509879 [Phytozome v9.1]</td>
<td>Glycolytic enolase</td>
</tr>
<tr>
<td><em>Volvox carteri</em></td>
<td>42159</td>
<td>Vocar20006493m.g; scaffold_7: 3486914 – 3491916 [Phytozome v9.1]</td>
<td>Multi-functional enolase</td>
</tr>
<tr>
<td><em>Dunaliella salina</em> CCAP19/18</td>
<td>-</td>
<td>GenBank submission ID 1734905</td>
<td>Glycolytic enolase</td>
</tr>
<tr>
<td><em>Dunaliella salina</em> CCAP19/18</td>
<td>-</td>
<td>GenBank submission ID 1734908</td>
<td>Multi-functional enolase</td>
</tr>
<tr>
<td><em>Scenedesmus obliquus</em> DOE0152 (locus 201)</td>
<td>NA</td>
<td>See Figure 2</td>
<td>Glycolytic enolase</td>
</tr>
<tr>
<td><em>Scenedesmus obliquus</em> DOE0152(locus 1181)</td>
<td>NA</td>
<td>See Figure 2</td>
<td>Glycolytic enolase</td>
</tr>
</tbody>
</table>
Gene phylogeny

As expected from the results of the previous study by (Harper & Keeling, 2004), the cytosolic enolase (ENOc) of *A. thaliana* showed closest homology to the green algal glycolytic enolases. Except for *Scenedesmus obliquus*, the enolase alignment (Figure 2 top) and phylogeny (figure 3 bottom) followed the basic taxonomy of the green algae included in our study. *Scenedesmus obliquus* DOE0152 has a protein sequence (Locus201) that falls closely with the algae cytosolic enolases, but has a transit pepide that suggests it is in the chloroplast. Its other copy, Locus1181, also falls within the cytosolic enolases and is predicted to be there (Figure 3).
Figure 2: (top) Alignment of enolase protein sequences, showing the DOE0152 enolase EST 201 translated open reading frame has a n-terminal transit peptide followed by an unknown domain. (Bottom) Alignment of the two Scenedesmus obliquus DOE0152 enolase sequences.
Figure 3: Protein Maximum Likelihood phylogeny reconstruction based on 500 bootstraps of the glycolytic enolase genes including several green algae and the three different glycolytic enolase forms from *A. thaliana*. Tree was rooted with the human alpha enolase. The multi-functional enolases were not included in this phylogeny, because they are too different from the glycolytic enolases and resulted in very low support values.

In photosynthetic eukaryotes, many of the biosynthetic pathways using pyruvate are localized only in the plastid, and it was reported previously that plants without proper supply of PEP as a precursor for pyruvate to plastids are impaired in growth and development (Prabhakar et al., 2010, Voll et al., 2009). An absence of the lower part of glycolysis from the chloroplast and localization only to the cytosol requires that fixed carbon must be exported by the chloroplast into the cytosol fraction of the cell where the glycolytic enolase enzyme could generate PEP, which could then be converted into pyruvate. The cytosolic PEP and/or pyruvate could then be used as a source metabolite for import of carbon back into the plastid (O'Grady et al., 2012). Based on information from higher plants, it would be expected that presence of enzymes in the plastid allowing...
for carbon flow from the Calvin-Benson cycle directly to pyruvate should make dramatic
differences in distribution of carbon within green algal cells.

Figure 4: Simplified schematic showing our proposed path for carbon partitioning
within the plastid of a green algal cell focused on starch, lipids, and isoprenoids. ENO
indicates the enolase enzyme outside the plastid. C3, C5, and C6 represent
monosaccharides with three, five, or six carbon atoms. PEP, abbreviation for
phosphoenolpyruvate. The blue circle represents the Calvin-Benson cycle with 1 =
Carbon fixation, 2 = reduction phase, 3 = regeneration phase. A shows the location
of the enolase in Chlamydomonas. B shows the possibilities with the chloroplast enolase.

Our analysis of available data (Terashima et al., 2011) strongly suggested that the
crucial enzyme enolase is only present in the cytosol of C. reinhardtii and most green
algae. Consequently, Figure 4 shows a simplified summary of our metabolic network
reconstruction results for green algae with proposed major routes of fixed carbon into
starch, glycerolipids, and isoprenoids. The carbon flow as proposed in Figure 4 requires
carbon exchange between plastid and cytosol due to the localization of the enolase only
to the cytosol. Consequently, the enolase activity might represent a major control point
and possible regulator for carbon partitioning in green algae, because the distribution of
carbon into macromolecules is not constitutive. As allocation of fixed carbon into
different cellular components (e.g. glycerolipids and isoprenoids) is altered by green
algae to satisfy the cellular needs during acclimation to changing environments,
regulation of carbon partitioning has to occur. One example is *D. salina* where cells react to reduced availability to nutrients – for example nitrogen – by switching from primary metabolism geared towards growth towards secondary stress metabolism, which includes over-accumulation of β-carotene in plastidic carotene globules (Ben-Amotz *et al.*, 1982). In non-stressed cells the amount of β-carotene is less than 0.1% of the total cellular biomass, but reduced availability of nitrogen results in accumulation of β-carotene to a level of more than 8% of the total cellular biomass (Ben-Amotz *et al.*, 1982). Such massive accumulation of β-carotene is due to de-novo biosynthesis of isoprenoids, which is evidence for massive re-routing of carbon under cellular stress into isoprenoid biosynthesis.

In green algae, isoprenoids are only made through the plastid localized MEP pathway using pyruvate as a precursor molecule. As PEP is a precursor to pyruvate, we hypothesize that the enolase represents a possible bottleneck in carbon partitioning in green algae. Specifically under stress conditions (e.g. low nutrient availability) when photosynthesis in algae is uncoupled from growth, the enolase might become a major regulatory point in adjusting carbon partitioning in response to environmental changes.

Although the plastid enolase seemed to be absent in most algae, it is present in *Scenedesmus obliquus*. At this point, the advantages of having this copy are speculative. Under growth conditions, the export of chloroplast triose phosphate to the cytosol is less likely to occur in favor of reactions related the synthesis of starch, which also takes place in the chloroplast, unless starch synthesis is bottlenecked. The chloroplast enolase in higher plants may give the plant more flexibility as to when it can complete glycolysis, as the triose phosphates do not need to build up (under specific conditions) and be exported
from the chloroplast; instead, glycolysis can occur in the chloroplast. In higher plants, glycolysis is up-regulated in the period of storage lipid synthesis (Troncoso-Ponce et al., 2009). And enolase proteins have been found to be up-regulated in high oil producing sunflower lines (Hajduch et al., 2007). It is perhaps for similar reasons that the chloroplast enolase is often active in the production of sink tissue such as seeds or fruits (Prabhakar et al., 2010, Andriotis et al., 2010).

Some of the triose phosphate incorporated in such sink tissue is not even directly from the Calvin-Benson Cycle, but from hexose phosphate imported into the plastid (Sriram et al., 2004, Alonso et al., 2007). The plastidial enolase is even more necessary in root tissue, which are also sinks (Prabhakar et al., 2009). Having the plastidial enolase, then, may allow for the creating of phosphoenopyruvate there, where it can be fed, directly, into the shikamate pathway for the production of amino acids and a huge variety of secondary plant metabolites (Prabhakar et al., 2009, Schulze-Siebert et al., 1984). Alternatively, phosphoenylpyruvate can be converted to pyruvate which is the precursor for fatty acid biosynthesis (Dennis, 1989, Ohlrogge & Jaworski, 1997), or isoprenoids, whose biosynthesis also occurs in the chloroplast. Most microalgae seem to lack this enolase (Polle et al., 2014), perhaps because microalgae do not produce sink tissue, which accumulate such metabolites, under favorable growth conditions.

The discovery of the chloroplast enolase in a Scenedesmus obliquus (DOE0152), which is a fast grower, both in the lab and also in outdoor ponds (Neofotis et al., In revision, NAABB Final Report, 2014) is of interest. Chlamydomonas reinhardtii, which only has one enolase, uses starch as a principle carbon sink (Krishnan et al., 2015) while Scenedesmus obliquus DOE1052 accumulates more lipids (Neofotis et al., In revision).
The presences of the chloroplast enolase in *Scenedesmus obliquus* DOE0152 may play a role in this difference in carbon allocation and warrants further investigation.

CONCLUSIONS

The enolase enzyme plays a major role in carbon partitioning. It is hypothesized that the number of copies and locations within the plant cell play a role in the regulation of carbon allocation. Though the enolase is primarily located in the cytoplasm, the presence of another enolase in the chloroplast seems to be important in the sink tissues of higher plants, which accumulate oil. Here, we show that although most green algae, such as *Chlamydomonas reinhardtii*, appear to lack this chloroplast enolase, it is present in an oleaginous microalga, *Scenedesmus obliquus*. *Scenedesmus obliquus* grows well outdoors and, unlike *Chlamydomonas reinhardtii*, is known to produce large amounts of oil. The role of the chloroplast enolase in oleaginous microalgae should be further investigated, as it may have much relevance to biofuel applications.
Chapter 4: Syntrophic Metabolism of Ethanol by *Agrobacterium tumefaciens* Enhances Micro-algal Growth


a Department of Biology, Brooklyn College, 2900 Bedford Avenue, Brooklyn, New York 11210, USA
b The Graduate Center of the City University of New York, New York, NY 10016, USA

ABSTRACT

Microalgae are seen as a potential feedstock for biofuels and also other high value products, but improvements in biomass yield in outdoor ponds are needed. One of the limitations of growth is CO$_2$, which is not present in optimal amounts in growth media without an external CO$_2$ supply (i.e. gas bubbling). A potential source of CO$_2$ could come from the respiration of bacteria, which are often used to uptake excess nutrients in wastewater treatment systems with the additional of an organic food source. Here, we describe a synthetic assemblage consisting of *Agrobacterium tumefaciens* and green algae. An organic molecule (ethanol) is first utilized by the bacteria, metabolizing it via respiration. In the process, the bacteria supply the algae with CO$_2$. This work has implications for improving algae cultivation, particularly in wastewater treatment systems, as well as the understanding of bacteria-algae interactions.

INTRODUCTION

Microalgae may be grown on wastewater as biofuel and high value product feedstocks (NAABB Final Report, 2014). In such outdoor pond systems, because CO$_2$
has a 10,000-fold slower diffusion rate in water as compared to air, algae are often severely limited in CO₂ (Moroney & Ynalvez, 2007, Fang et al., 2012) that is needed for the Calvin Cycle of photosynthesis.

Bacteria and other heterotrophs, in the process of respiring organic molecules, release CO₂. Although these organic molecules are often thought of as sugars and carbohydrates, some bacteria also have the capacity to metabolize ethanol. This metabolism is of interest because ethanol is seen as one of the most attractive organic carbon sources to allow bacteria to grow and utilize nitrogen and phosphorus at sufficient rates for the nutrients’ removal in wastewater treatment ponds (Ginige et al., 2009, Mokhayeri et al., 2006, Kazasi et al., 2013). Compared to alternatives, such as methanol, it is less toxic, has higher specific denitrification rates (SDNRs) (Trela 1998; Ginge et al., 2009) and is more readily metabolized at even colder temperatures (Mokhayeri et al., 2006, Cherchi et al., 2009). Ethanol is also less corrosive than methanol and its distinctive odor makes it easy to detect (Kasasi 2013). When bacteria metabolize the ethanol, they release CO₂. This CO₂ can then, in principle, be utilized by microalgae in the Calvin Cycle for autotrophic growth. What is unknown is whether and which bacteria can metabolize ethanol at high enough rates and release CO₂ in sufficient amounts to stimulate algal growth, and whether such bacterial growth will have its own negative effects on algal biomass yield. It is also unclear if ethanol alone, without bacteria, can stimulate algal productivity, although at least one general review has suggested it (Wiessner, 1979).

In the process of taking up the CO₂ in photosynthesis, the algae release oxygen in the water, which in turn allows the bacteria to gain sustained access to the energy stored
in organic matter (Oswald et al., 1955). Thereby, by pairing a bacteria with green algae, the decomposition of a simple organic molecule may occur as more valuable carbon molecules (i.e. lipids, carotenoids) are being synthesized by the algae, provided light is also available (Oswald et al., 1955). This cycle of bacteria respiration and photosynthetic oxygen production is shown in Figure 1.

Ideal pond treatment would involve an overlapping of the bacteria and algae phases, forcing direct and immediate gas exchange and symbiosis. However, such interactions are seldom encountered in nature (Oswald et al., 1953). This is due to several incompatibilities, including resource competition and antibiotic activity (Oswald et al., 1953). The antagonistic relationship between bacteria and algae is one of the reasons that single algae with mixotrophic metabolism (Henkanatte-Gedera et al., 2015) are very attractive options for urban wastewater treatment plants. However, organic matter is most rapidly oxidized by bacteria (Oswald et al., 1955).

The work’s hypothesis is that there are bacteria that can co-exist with algae and metabolize organic carbon (i.e. ethanol); in the process, the bacteria release CO$_2$ into the liquid media, thereby stimulating algal growth. The objective of this work is to gain insight into bacteria-algae relationships that may be utilized in future culture systems, including wastewater, to increase algal productivity. Although the organic utilized here is ethanol, because the bacteria discovered to grow well with ethanol is one of the most readily transformable, the bacteria’s metabolism could also be extended to other organics.
METHODS

Experiments were conducted principally with two species of micro-algae. The first was *Scenedesmus obliquus* (DOE0152) (Neofotis et al., In revision). The second species was *Chlamydomonas reinhardtii* (CC119 SRU-) or the acetate requiring mutant (CC-530 ac17) (Levine & Goodenou, 1970, Ebersold et al., 1962). *Scenedesmus obliquus* was used because it has been grown outdoors in mass culture since the 1980s (Becker, 1984) and even recently has been shown to be one of the most promising strains for further development as a biofuel feedstock (Neofotis et al., In revision). *Chlamydomonas reinhardtii* was used because it is the most studied, and widely-regarded model, of the green micro-algae (Harris, 2009). Strains were grown up in defined autotrophic media, with *Chlamydomonas reinhardtii* grown from single colonies in HS medium (Harris 2009) and *Scenedesmus obliquus* grown in BG11 medium (Anderson 2005) in 50 ml Erlenmeyer flasks on a rotary shaker under about 50 μE m⁻² s⁻¹ continuous light provided by daylight fluorescent lamps for roughly five days (A₆₈₀=0.15-.2). One ml was then
inoculated into 20 ml of media or 2.5 ml into 50 ml, depending on volume desired for the experiment. *Agrobacterium tumefaciens*, having been grown overnight in YEP media ($A_{600} > 1$), were then centrifuged (4500 rpm for 3 minutes) down and re-suspended in appropriate autotrophic media (BG11 or HS) and, based on OD approximation, about $1 \times 10^6$ cells of bacteria were used to inoculate the specific treatment (usually 9 ul of suspension). *Agrobacterium tumefactions* strain C58 was primarily used, although GV3101, EHA105, and C58c1 were also tested. *Saccharomyces cerevisiae*, *Pseudomonas fluorescens*, *Pseudomonas aeruginosa*, *Acetobacter aceti*, and *Sinorhizobium meliloti* were also all tested. For samples treated with ethanol, 1 ul of 100 percent ethanol (Sigma) was used for every ml of the total growth media. All treatments were grown in triplicates. Treatments were either algae alone, algae & bacteria, algae & ethanol, or algae & bacteria & ethanol. Flasks of media with bacteria, and media with bacteria & ethanol were also grown. Algae cells were counted on subsequent days using a Neubauer hemocytometer.

Genes related to ethanol metabolism were searched in National Center for Biotechnology (NCBI) by running protein queries of translated nucleotides (tblastn) against the *Agrobacterium* taxa and *Chlamydomonas reinhardtii* genomes. The *A. tumefaciens* genome (Wood *et al.*, 2001) was also searched using protein queries of translated nucleotides.

**RESULTS AND DISCUSSION**

It is widely known that *Pseudomonads, Saccharomyces*, and *Acetobacter aceti* metabolize ethanol (Arndt *et al.*, 2008). *A. tumefaciens* has been reported to grow on
ethanol (Tsfira & Citovsky, 2008). Enzymes necessary to metabolize ethanol include the ethanol dehydrogenase and the acetate kinase (Ingram-Smith et al., 2006). Some of these enzymes are also found in green algae, such as Chlamydomonas reinhardtii (Figure 2). It does not seem that Chlamydomonas reinhardtii has all the enzymes necessary to metabolize ethanol to acetate. It lacks an ethanol dehydrogenase and acetaldehyde dehydrogenase (AcetalDh), although it does have an alcohol dehydrogenase.

Figure 2: A known pathway of ethanol metabolism (Arndt et al., 2008). Enzymes found with high confidence in the genomes of A. tumefaciens are boxed in brown while those found in Chlamydomonas reinhardtii are boxed in green. The acs gene, present in both Chlamydomonas and A. tumefaciens, is shaded both brown and green.

To determine what microbe could metabolize ethanol and, in the process, stimulate algal growth, we tested the growth stimulation effect of Saccharomyces...
cerevisae, Pseudomonas fluorescens, Pseudomonas aeruginosa, and Acetobacter aceti, along with Agrobacterium tumeficiens C58 (with Ti-Plasmid), C58-c1 (without Ti-Plasmid), and GV3101 (with a disarmed Ti-Plasmid). After 6 days of culturing, the A. tumeficiens had an effect on algal growth – with only the A. tumefaciens strains C58 and GV3101’s effect being significant (P<0.001).

Figure 3: Combined effect of ethanol and different bacteria on growth of Scenedesmus obliquus. When ethanol was not also present in the media, no bacteria had any effect on algae growth.

To determine why Agrobacterium, of all the bacteria, induced the growth response – we also tested what effect ethanol in the minimal media had on bacteria growth alone. After 48 hours, the Agrobacterium’s growth was the most apparent, followed by the Pseudomonas aeruginosa. Sinorhizobium meliloti, which is closely related to Agrobacterium, was also tested and exhibited only a small amount of growth (Figure 4).
Figure 4: Photo (top) and OD$_{600}$ graph (bottom) showing the effect of ethanol and minimal (HS) media on bacteria growth. Although the *Pseudomonas aeruginosa* also grew, it did so in clumps and to a lesser degree than the *Agrobacterium*. The experiment was also conducted in BG11 media with similar results.

From an evolutionary point of view, that *Agrobacterium* – a known plant pathogen that induces excessive, tumor-like plant growth – grew well on ethanol and in turn syntrophically metabolized it to CO$_2$ is noteworthy for the bacteria. Ethanol is produced by plants under biotic and abiotic stress conditions (Hann *et al.*, 2014, Kimmerer & Kozlowski, 1982, Macdonald *et al.*, 1989). Ethanol has also been found to
induce the *Agrobacterium* promoter (P2) which in turn regulates VirG, a protein that induces pathogenesis (Mantis & Winans, 1992).

An experiment was also conducted to see if other carbon sources, such as methanol, could also induce the growth response. The effect of methanol, ethanol, propanol, and butanol with *Agrobacterium* were all tested on the growth of *Scenedesmus obliquus*, and only ethanol induced a significant response (P>95%) (Figure 5).

![Figure 5: Effect of different alcohols on *Scenedesmus obliquus* growth. None of the carbon sources had an effect when *Agrobacterium* was not present. Only the combination of *Agrobacterium tumefactions* and ethanol had a statistically significant (P>95%) effect on algal growth.](image)
To address in more detail the growth response, *S. obliquus* was grown and measured daily in 50ml Erlenmeyer flasks. As before for *S. obliquus*, no statistically significant response was observed when the algae was exposed to ethanol or *Agrobacterium* alone. However, when the algae was treated with ethanol and *Agrobacterium*, an over a 300% increase in growth was observed (Figure 6).
Figure 6: Photo (top) and graph (bottom) of the effect of ethanol and *A. tumefaciens* C58 on *Scenedesmus obliquus* (DOE 1052) growth. Only the combination of ethanol and *Agrobacterium tumefaciens* had a significant effect on algal growth (P<0.001). Photo was taken after cultivation for five days.

We also repeated the experiment for *Chlamydomonas reinhardtii*, and found that, similarly, ethanol and *Agrobacterium* had a stimulation effect.
As media becomes CO₂ enriched, it typically acidifies. We grew the *A. tumefaciens* alone in our growth media and found that indeed, the pH dropped after 48 hours, which corresponds to when the stimulation response typically occurs. The media alone, media with ethanol, and media with ethanol and bacteria all had a pH of 7, while the media with bacteria and ethanol had a 6.5 pH. This drop in pH, though, could not be solely due to the addition of carbon, but also the assimilation of ammonium.

When this media was then filter-sterilized and used to grow algae, a significant stimulation response was observed (P<0.01) (Figure 8). However, the algal growth response was not sustainable and the culture collapsed after the 3rd day. It seems that the *Agrobacterium* continues to respire in response to the stimulated algae growth, as the algae supply the bacteria with oxygen. Also, cultures of *Agrobacterium* and algae with ethanol do show the bacteria attaching to the algal cell walls (Figure 9).
Figure 8: Growth of *S. obliquus* in autotrophic media (BG11) with *A. tumefaciens* and ethanol in it for 48 hours, after which the bacteria was removed (red) – compared to *S. obliquus* in BGII alone (blue).

Figure 9: Cells of *S. obliquus* growing in BG11 media treated with *A. tumefaciens* and ethanol.

As another test to determine if CO$_2$ is contributing to the stimulation effect, when the experiment was conducted under non-limiting CO$_2$ conditions (i.e. bubbling with 1%
CO₂), the effect of the ethanol and *A. tumefaciens* together was no longer apparent when compared to the other treatments (Figure 10).

Figure 10: *C. reinhardtii* exposed to *A. tumefaciens* and 1 ul ethanol/ml of media, but also bubbling with 1% CO₂. When additional CO₂ was not supplied, no difference was observed after three days of culturing. Photo was taken after 6 days of culturing.

Besides the stimulation effect being because of additional CO₂, the other most likely mechanism is that the bacteria, in converting the ethanol to acetate, could be releasing some of this acetate into the media and stimulates the algae (heterotrophically) via the glyoxylate cycle of respiration. However, we found that adding acetate alone, without a buffer, to media does not stimulate *Scenedesmus* growth. Furthermore, when an acetate-requiring mutant strain of *Chlamydomonas reinhardtii* (AC17) was used in the experiment, it was found that in the culture with *Agrobacterium* and ethanol the *Chlamydomonas reinhardtii* AC17 did not survive. This result suggests that even though *A. tumefaciens* may be converting ethanol to acetate, the bacteria is not making the acetate available to the microalgae in significant quantities.
Figure 11: Cultures of C-530 ac17 mt-, which requires acetate (Levine & Goodenou, 1970, Ebersold et al., 1962) after 72 hours. Only the cultures grown in Tris-Acetate-Phosphate (TAP) media grew.

Besides CO$_2$, *Agrobacterium* cells may be stimulating the algae’s growth via other mechanisms. Indeed, *Agrobacteria* can supplement algae with several limiting metabolites, such as biotin, thiamine, cobalamine, and siderophores that help facilitate growth and bioconversion (Wang *et al.*, 2013). We were able to rule out cobalamine, which, although *Agrobacterium* is reported to produce (Tsfira & Citovsky, 2008), we did not find stimulated the growth of *C. reinhardtii*. The other possibilities, particularly the creation of siderophores, which *A. tumefaciens* is known to also produce (Rondon *et al.*, 2004), are other possible additional factors. However, doubling or quadrupling the iron content of the media did not have observable impacts on culture growth (SM-1).
What is clear is that *A. tumefaciens* and small amounts of ethanol allow the culture, as a whole, to accumulate biomass faster – as well as uptake nutrients. When the experiment was repeated with *Chlamydomonas reinhardtii* in 0.5x, 1x, and 2x strength HS media, the cultures of *A. tumefaciens* and ethanol in 0.5x collapsed first, followed by the *A. tumefaciens* and ethanol in 1x (Figure 12). The sequence of culture collapse demonstrated that the synthetic assemblage of algae and ethanol and bacteria are the first to run out of nutrients.

![Figure 12: Cultures of *C. reinhardtii* treated with *A. tumefaciens* and ethanol with different strength HS media. Photo taken after 7 days, shows where (brown arrow) the culture exposed to ethanol and *A. tumefaciens*, which had bloomed, crashed first. A week later, the culture exposed to ethanol and *A. tumefaciens* (green arrow) was the first of the 1x cultures to crash.](image)
CONCLUSIONS

The stability and function of microbial communities depends on metabolite interactions among community members (Seth & Taga, 2014). And while the ability to compete for nutrients can determine whether or not a microbe will be able to persist in an particular niche (Hibbing et al., 2010, Johnson et al., 2012), metabolic cooperation between microbes engaged in syntrophic partnerships might allow access to substrates that neither microbe could metabolize alone. Here, we find that A. tumefaciens syntrophically metabolizes ethanol and, in the process, stimulates green algal growth. A primary mechanism for this stimulation effect is CO$_2$ being respired by the bacteria as it metabolized ethanol. While further work may detail additional mechanisms, this work highlights that synthetic assemblages of A. tumefaciens and green algae, with the additional of small amounts of ethanol, may effectively form a culture that quickly uptakes nutrients from water and also produces useful algae-derived end-products.
Overall Conclusions

This work contributes to improving and understanding the biomass productivity and utility of the green algae. It identified and characterized the best strains to come out of a large scale effort to discover platform strains for biofuel, as well as other high value products, production. Carbon dioxide is often limiting in outdoor ponds, but this thesis finds that there is potential to alleviate this limitation via symbiotic association with Agrobacterium, which can, via respiration, metabolize an alternative carbon source and enter a symbiotic gas exchange relationship with green algae.
Appendix: Directions on Current Knowledge

Peter Neofotis\textsuperscript{a,b}

\textsuperscript{a} Department of Biology, Brooklyn College, 2900 Bedford Avenue, Brooklyn, New York 11210, USA
\textsuperscript{b} The Graduate Center of the City University of New York, New York, NY 10016, USA

Selecting algae that are naturally suited for biofuel production is important because “domesticating” the algae is likely to involve different challenges than those posed by the angiosperms (Table 1). These challenges are important to consider. When thinking about the unrealized possibilities of algal productivity and oil yield, an example often used to highlight the vast improvements possible with microalgae is given with maize, which has large kernels, which arose from the wild and small-grained teosinte (Beadle, 1980, Doebley \textit{et al.}, 1990). Though the development of maize took thousands of years of human selection, its domestication demonstrates, one might argue, the untapped potential that lies with any microalgae once we turn our attention to it; especially now that we have genetic tools available for some of the microalgae (Purton, 2007, Forjan \textit{et al.}, 2015).

Table 1: Advantages and disadvantages in domesticating the Microalgae when compared to the Angiosperms, showing in red, greater challenges, and in green, greater advantages.

<table>
<thead>
<tr>
<th>Fundamental Difference</th>
<th>Microalgae</th>
<th>Angiosperms</th>
<th>Engineering Opportunity for Algae Culturing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differences in conceptual mortality &amp; mutualism</td>
<td>Accumulates lipids, starches, and carotenoids for itself</td>
<td>Accumulates lipids, starches, and carotenoids for progeny (seeds) or (fruits) dissemination</td>
<td>Selecting strains that accumulate these in high quantities, along with high value products</td>
</tr>
</tbody>
</table>
However, there are vast conceptual and biological design differences that set domestication algae apart from angiosperms. Angiosperms are uniquely suited for domestication. Indeed, outside of the angiosperm, one is hard pressed to find a photosynthetic species that is dramatically different from its wild progenitor. Yet, in the angiosperms, such examples readily abound. The first reason that angiosperms are well-adapted for domestication is that they are hard wired to provide a separate nutritional encasement (the seed) for its progeny or an attractant (fruit) for a disperser. The
nutritional tissue inside the seed is called endosperm, and created as a result of a second fertilization event. As seeds became larger to account for the shade stress in forests, yet also where wind was not available for dispersal, fruits then emerged to utilize animals as the disseminators (Willson et al., 1990, Levey, 1988).

The sessile angiosperm then, seemingly aware of its own life on land’s immovability and impermanence, is wired to send resources into what can be ultimately considered a separate structure meant to be gathered or moved – as doing so contributes to their next generation.

The green micro-algae, on the other hand, are able to survive perpetually on asexual reproduction alone. For some, like those in the Trebouxiophyceae, it is not clear if mating occurs, although the genes for sexual reproduction appear to be present (Fucikova et al., 2015). Even if sex does occur, the zygote does not germinate incased with other additional cells to provide nutritional support, like the endosperm of angiosperms. The concept of storing large reserves of starches, lipids, or sugars to be used for another individual (be it progeny or disseminator) is not present in the algal life-plan.

Instead, the energy storage molecules they make for themselves is usually to enable them to make it through periods, such as the night, when light energy is not available. Or, they make pigments to protect them from solar radiation that is beyond the limits of their photosynthetic capacity. So, while a focus of much research has been to try to find a way to get micro-algae that accumulate lipids without, at the same time, suffering a degradation of the photosynthetic apparatus – these goals are contradictory, at least for the most studies algae – *Chlamydomonas reinhardtii*. Unlike the angiosperms,
placing an extreme amount of excess energy into a sink has no clear advantage for the algae when there is an option to divide. Beyond a certain level, many algae only accumulate lipids because they do not have the nutrients available to divide their cells. It is for these reasons that the quest of finding an algae that simultaneously grows fast while accumulating large amounts of lipids is challenging and trade-offs have been demonstrated between growth and lipid accumulation (Miller et al., 2010, Juergens et al., 2015).

In fact, the reason that algae accumulate excess lipids is because they are experiencing periods of nutrient limitation. At this time, it has been suggested that photosynthesis induces significant stress via energetic overflows, bringing about the potential oxidative damage and driving starch and oil accumulation as energy sinks (Roessler, 1990, Hu et al., 2008, Li et al., 2010). Starch and oil accumulate in algae usually because they are deprived of nitrogen and other organic nutrients, growth slows, and then photosynthesis is downregulated. This can be seen in changes in the levels of transcripts and proteins associated with photosynthesis (Juergens et al., 2015), as the expression of most photosynthesis genes is downregulated during N deprivations (Juergens et al., 2015, Merchant et al., 2007, Schmollinger et al., 2014), particularly those in the light reactions, along with chlorophyll and the theoretical maximum quantum efficiency of Photosystem II (PSII). At the same time that cells are downregulating photosynthesis, decreasing carbon fixation, and reducing ATP and NADPH production; cells then start accumulating starch and lipids (Siaut et al., 2011, Fan et al., 2011, Juergens et al., 2015). The accumulation of triglyceride molecules in specialized organelles occurs along with structural changes and a breakdown of intracellular
membrane systems such as the thylakoids and endoplasmic reticulum (Martin & Goodenough, 1975, Moellering & Benning, 2010). And though there is a strong increase in transcript levels for genes encoding the committing step of Triglycerides (TAG) synthesis (Miller et al., 2010), this may not actually reflect a further accumulation of carbon. Rather, it may be a result of a recycling of fatty acids in membrane lipids into TAGs (Miller et al., 2010). The lipids in micro-algae, then, are accumulated because growth conditions are not optimal. The goal then of finding the micro-algae equivalent to something like an peanut or avocado, which creates large vessels of oil best under ideal growth conditions, almost boarders on fantasy because it is like putting a central tenant from one world based on multi-cellularity and mortality – and placing it in another of single cellularity and perpetual renewal. The same visions do not apply.

Because the oil and high-value products we want from algae accumulate in algae under non-optimal growth conditions, the stress biology of these organisms is of great importance. It maybe “too luxurious” to grow algae only for biofuel applications, but growing them for fine chemicals, cosmetics, and medicines which are accumulated under stress offer an avenue for commercialization (Zhu, 2015). A first step is identifying what organisms, when under stress, synthesize the lipid and carotenoid molecules that are desired. The quality of TAG molecules is also important. The longer the unbranched chain of the fatty acid methyl ester (FAME) components, the higher the cetane number (CN) which is indicative of the time delay in the ignition of fuel (Nascimento et al., 2013). Neutral fatty acids with a high degree of saturation are also desired (Talebi et al., 2013) although unsaturated ones are more desired for human health benefits. Once the organisms which produce the desired metabolites are known, and if they have high
growth rates, we may be able to determine a way to switch on the genes related to stress once the organism has reached its optimal growth density. To ease harvesting, algae with large cell sizes (larger then 20 µm (Borowitzka, 1997, Borowitzka & Moheimani, 2013) are particularly desirable.

A second reason that microalgae domestication is likely to be different is that microalgae, being single cellular, must all contain all essential functional characteristics of a photosynthetic cell. Angiosperms, on the other hand, have greater opportunities for spatial heterogeneity. At the cellular level, this allows for an angiosperm’s seed to have markedly different metabolism than a photosynthetic cell (Schwender et al., 2015). Because the seed is collecting nutrients from the mother plant, the cells do not have to save room in the cell for photosynthetic membranes, and so can use the space to store a combination of starches, proteins, and lipids (Schwender et al., 2015). As with the seed, the regulatory mechanisms controlling fruit ripening have nothing to do with photosynthetic stress (Tranbarger et al., 2011).

From a whole plant view, compared to the algae the plant can create large fruits or seeds – apart from the photosynthetic tissue — which can be easily harvested. For the angiosperm, the heterogeneity of their body structure also gives much room for reallocation of resources. For example, the whole plant structure of teosinte was drastically altered to create maize, with the number and architecture of female inflorescences being concentrated, and the number of male tassels reduced and shortened, allowing more resources to be taken to the kernels (Doebley et al., 1990). For the algae, the resource reallocation is limited to the cellular level. But this cellular reallocation has still had
positive results. Mutants that do not produce starch, for instance, have shown increased lipid production (Goodenough et al., 2014).

Though the lack of fruits, seeds, and heterogeneity have disadvantages in terms of the prospects of domestication, compared to land plants, algae have more rapid biomass production (Rittmann, 2008, Groom et al., 2008). This may, in part, be because the cells can reproduce without a need or concern for structural differentiation. Algae’s lack of structural heterogeneity is also part of why they can also be grown in cell culture systems, which may offer more cultivation flexibility as well as contribute to their high yield per acre. As single cellular organisms, they can also be subjected to rapid selections such as those in pH-auxostats, as has been done with the already fast growing yeast Kluyveromyces marxianus (Groeneveld et al., 2009), to select faster even growing strains. Finding organisms with already fast growth rates, then, becomes important.

Biomass productivity remains a top priority. Furthermore, it is essential that these strains exhibit this growth not only in the laboratory, but also outdoors, which involves a variable environment (Borowitzka & Moheimani, 2013, Sheehan et al., 1998). Going along with this, it is important to determine which organisms can be cultivated in saltwater or wastewater, which is more available than freshwater.

Angiosperms were the enablers of human civilization. Intrinsic to many of their natures is a since of gift giving – to an animal to spread to carry its seeds, to the next generation to support its growth on inhospitable land. By enhancing this tendency found in nature, we have created the great grain crops of corn, wheat, soy, and rice – which largely feed modern civilization. Many angiosperms have also developed fleshy fruits, filled with nutrients such as Vitamin C, which, in exchange for, an animal disperses in an
encased seed. Selecting and enhancing the desirable protein, carbohydrate, lipid, and vitamin characteristics of seeds and fruits has been a major cultural achievement in human history. By doing so, complex human societies were able to emerge. It remains to be seen whether or not the micro-algae, who intrinsically lack this since of gift-giving, of yielding to another generation for renewal or to an animal for the sake of mutualistic interaction, may be willing participants in a partnership with the fire species, humans. Domesticating the algae will likely require a level of rewiring of metabolism that never had to be done with the angiosperms. It is for this reason that it is very important to find species with natural tenancies that at least partially align with our needs. From these, work toward improving their growth can then be continued.
Bibliography


Science 324:268-72.


