THE MICROBIAL COMMUNITY ECOLOGY OF VARIOUS SYSTEMS FOR THE
CULTIVATION OF ALGAL BIODIESEL

by

Tisza Ann Szeremy Bell

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Microbiology

MONTANA STATE UNIVERSITY
Bozeman, Montana

January 2017
DEDICATION

This body of work is dedicated to my beloved best friend, the epitome of compassion and strength, the wild thing I never saw sorry for itself, who lived in love, and never let me quit - even in her absence.

Ryan Marie Patterson

January 6, 1983 – October 9, 2011

i carry your heart with me(i carry it in
my heart)i am never without it(anywhere
i go you go,my dear;and whatever is done
by only me is your doing,my darling)

here is the deepest secret nobody knows
(here is the root of the root and the bud of the bud
and the sky of the sky of a tree called life;which grows
higher than soul can hope or mind can hide)
and this is the wonder that’s keeping the stars apart

i carry your heart(i carry it in my heart)

-EE Cummings
ACKNOWLEDGEMENTS

I would like to thank family and friends for their love, support, and encouragement. I would be nowhere in this world without my amazing parents, Susi and Richard, my brother Devon, and my partner Brian Guyer, and my faithful Puli, Jack, who have loved me, commiserated with me, and tolerated me on my worst days. I am indebted to my late great-uncle, Tom Bell, the first black sheep, for his encouragement, faith, and inspirational life’s work through *High Country News*. I have had endless love and support from my aunt and uncle, Tunde and Dick Dwyer, my cousins Chris & Andy Szeremy, Ajna Akantisz, everyone in Hungary, and the whole of the Patterson clan especially, Uncle Billy, Mike, Marie, and Danielle. My dear friends Tracy Thomas, Flynn Murray, Karen Moll, Emily Wellington, Leila Sterman, Josiah Pisel, Julia Mead, Katie & Steve Breston, Andrea Hobbs, Kevin Clark, Mason Griffith, and Maria Blevens have given me endless strength and support. I am in debt to Julie and Jim Wellington for providing me sanctuary and solitude to write at one of my favorite places in the world, Longhorn Ranch. I am forever indebted to my lab-mates and my advisers, Brent Peyton and Matthew Fields, for their guidance and mentorship, and my committee members Anne Camper and Christine Formeman. I am especially thankful to Christine Foreman for convincing me not to quit my PhD on more than one occasion, to Sarah Codd for her mentorship, and Ann Willis for support that knew no bounds (ex: dumpster diving for samples at 11pm on a Sunday). I am grateful for the formatting prowess of Robin Kelley and to my lab mates past and present especially Kara DeLeon, Karen Moll, Muneeb Rathore, Todd Pedersen, Matthew Jackson, Jackson Herron, Katie Davis, Anna Zalaya, Chiachi Hwang, Heidi Schoen, Lakotah Doig, Emel Sen, Sarah Altenburg, and Jacob Valenzuela.
TABLE OF CONTENTS

1. INTRODUCTION ............................................................................................................. 1
   Renewable Energy and Role of Algal Biodiesel ............................................................... 1
   Current Algal Cultivation and Lipid Augmentation Approaches and Challenges ....... 4
      Strain Selection and Biomass ..................................................................................... 4
      Stress and Nutrient Deprivation .............................................................................. 4
   Photobioreactors ............................................................................................................ 5
   Genetic Engineering ....................................................................................................... 6
   Prospective Methodologies to Further Maximize Lipid Potential .............................. 8
      Identification of Promising Regional Systems and Strains ................................... 8
      Employing Ecologic Principals to Microbial Systems ........................................... 8
         Diversity and Productivity .................................................................................... 8
         Niche Occupancy and Stability ........................................................................... 9
         The Phycosphere .................................................................................................... 10
      Improving Stability by Utilizing Alkalinity .............................................................. 12
   Mass Production ........................................................................................................... 13
      Water Limitations for Mass Production ................................................................ 14
      Viability of Wastewater ............................................................................................ 15
      Wastewater Nutrients .............................................................................................. 16
   Conclusion ..................................................................................................................... 19

2. A LIPID-ACCUMULATING ALGA MAINTAINS GROWTH
   IN OUTDOOR, ALKALIPHILIC RACEWAY POND WITH MIXED
   MICROBIAL COMMUNITIES .......................................................................................... 22
   Contribution of Authors and Co-Authors .................................................................... 22
   Manuscript Information Page ......................................................................................... 23
   Abstract .......................................................................................................................... 24
   Introduction ..................................................................................................................... 25
   Methods ........................................................................................................................... 27
      Site Description and Raceway Pond Conditions ....................................................... 27
      Sample Collection ...................................................................................................... 28
      DNA Extraction and Sequencing ............................................................................. 29
         DNA Extraction ..................................................................................................... 29
         Bar-Coded Pyrosequencing .................................................................................... 29
      Bioinformatic Sequence and Community Analysis .............................................. 30
      Statistical Analysis .................................................................................................... 32
   Results and Discussion ................................................................................................. 33
      Environmental Variables .......................................................................................... 33
      Community Composition and Interaction .............................................................. 34
      Correlation of Environmental Variables and Community Structure ..................... 40
TABLE OF CONTENTS CONTINUED

Correlations between Community Members ...................................................... 43
Conclusions ........................................................................................................... 47
Funding ................................................................................................................ 48
Acknowledgements ............................................................................................. 48

3. BACTERIA AND EUKARYA COMMUNITY DURING EUTROPHICATION AND TOXIC CYANOBACTERIAL BLOOMS IN THE ALKALINE LAKE VELENCE, HUNGARY ....................... 49

Contribution of Authors and Co-Authors .......................................................... 49
Manuscript Information Page ............................................................................ 50
Abstract ............................................................................................................. 52
Introduction ........................................................................................................ 53
Materials and Methods ..................................................................................... 54
  Site Description ................................................................................................. 54
  Sample Collection ............................................................................................. 55
  MALDI-TOF MS Analysis .................................................................................. 57
  Chemical Analyses ............................................................................................ 58
  DNA Extraction and Sequencing ...................................................................... 59
    DNA Extraction ............................................................................................... 59
    Bar-Coded Pyrosequencing ............................................................................ 59
  Processing of Pyrosequencing Data and Community Analysis ................. 61
  Statistical Analysis .......................................................................................... 62
Results and Discussion ..................................................................................... 63
  Chemical Analyses ............................................................................................ 64
Sequencing and Community Analysis .............................................................. 68
  Eukaryotic Diversity ......................................................................................... 70
  Bacterial Diversity ............................................................................................ 75
  Algal and Bacterial Dynamics ......................................................................... 79
Conclusions ......................................................................................................... 84
Acknowledgements ............................................................................................ 84
Funding ............................................................................................................... 85
References .......................................................................................................... 86

4. MONITORING COMMUNITY ECOLOGY IN WASTEWATER TREATMENT LAGOONS FOR THE PRODUCTION OF ALGAL BIODIESEL ................................................................. 111

Contribution of Authors and Co-Authors ........................................................ 111
Manuscript Information Page ........................................................................... 112
### TABLE OF CONTENTS CONTINUED

Abstract ................................................................................................................................. 114  
Introduction ............................................................................................................................ 115  
Material and Methods ........................................................................................................ 118  
  Site Description .................................................................................................................. 118  
  Sample Collection .............................................................................................................. 120  
  Water Chemistry ............................................................................................................... 120  
  Algal Cell Counts .............................................................................................................. 121  
  FAME Analysis .................................................................................................................. 121  
  DNA Extraction and Sequencing ..................................................................................... 122  
  Bioinformatic Sequence and Community Analysis ........................................................ 123  
  Statistical Analysis .......................................................................................................... 125  
Results ................................................................................................................................. 127  
  Water Chemistry .............................................................................................................. 127  
  Cell Number ...................................................................................................................... 128  
  FAME .................................................................................................................................. 129  
  Retention Time .................................................................................................................. 131  
Community Structure and Diversity .................................................................................. 132  
  Diversity Estimates .......................................................................................................... 132  
  Archaeal Community Structure ..................................................................................... 134  
  Bacterial Community Structure ..................................................................................... 134  
  Eukaryotic Community Structure .................................................................................. 138  
  dsDNA Viruses Community Structure ......................................................................... 141  
  Community Structure and FAME ................................................................................... 145  
Discussion ............................................................................................................................ 147  
  Influential Environmental Factors on FAME ................................................................. 149  
  Algae Taxa Potentially Responsible for Elevated FAME ............................................... 150  
  Ecologic Principles and the Functional Role of Interspecific Diversity on FAME .......... 152  
Conclusion ............................................................................................................................ 156  
Funding ................................................................................................................................. 157  
Acknowledgements ............................................................................................................ 157  
Conflict of Interest ............................................................................................................. 158  
References ............................................................................................................................ 159  

5. CONCLUSION ...................................................................................................................... 184  

Knowledge Gaps and Areas for Future Research ............................................................. 184  
  Understanding of Large Scale In Situ Community Ecology ............................................. 184  
  Ecologic Principles Specific to Algae .............................................................................. 185  
  Spectral Niche Theory ....................................................................................................... 185  
  Role of Viruses as Community Members ......................................................................... 186
TABLE OF CONTENTS CONTINUED

Closing ..........................................................................................................................................................187

REFERENCES CITED .....................................................................................................................................192

APPENDICES ...............................................................................................................................................218

APPENDIX A: Supporting Tables and Figures for
Chapter 1 - Introduction ..........................................................................................................................219

APPENDIX B: Supporting Tables and Figures for
Chapter 3 - Bacteria and Eukarya Community During Eutrophication and Toxic Cyanobacterial Blooms in the Alkaline Lake Velence, Hungary ........................................................................................................222

APPENDIX C: Supporting Tables and Figures for
Chapter 4 - Monitoring Community Ecology in Wastewater Treatment Lagoons for the Production of Algal Biodiesel .............................................................................................................................................225

APPENDIX D: Sources and Re-Sources: Importance of Nutrients, Resource Allocation, and Ecology in Microalgal Cultivation for Lipid Accumulation .................................................................................................................264
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1: Water chemistry over the course of the pond run.</td>
<td>34</td>
</tr>
<tr>
<td>2.2: 454 Pyrosequencing statistics.</td>
<td>35</td>
</tr>
<tr>
<td>3.1: Summary of water chemistry analysis at six time points taken approximately every week for 6 weeks.</td>
<td>64</td>
</tr>
<tr>
<td>3.2: 454 Pyrosequencing statistics for SSU rRNA gene sequences</td>
<td>68</td>
</tr>
<tr>
<td>3.3: Hill numbers showing changes in the diversity of <em>Bacteria</em> and <em>Eukarya</em> over the sampling period.</td>
<td>70</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1. Global liquid fossil fuel production and consumption in millions of barrels per day (b/d) from 2011 to 2016 with projections for 2017</td>
<td>10</td>
</tr>
<tr>
<td>2.1. Total relative abundances of A) archaeal and B) bacterial taxa by order (as order, genus, or unknown) with &gt;2% relative abundance over the 16 day sample period</td>
<td>35</td>
</tr>
<tr>
<td>2.2. Cell concentration (log cells/mL) and the relative abundance of C. vulgaris over 16-day sample period. The transfer from the 200 L to 2,000 L raceway is indicated by the vertical line</td>
<td>37</td>
</tr>
<tr>
<td>2.3. Chao diversity for each domain plotted over time</td>
<td>39</td>
</tr>
<tr>
<td>2.4. Canonical Correspondence Analysis (CCA) plotting chemical variables that correlated with variation in Archaeal taxa (with &gt;2% relative abundance) and sample points</td>
<td>42</td>
</tr>
<tr>
<td>2.5. Canonical Correspondence Analysis (CCA) plotting chemical variables that correlated with variation in Bacterial taxa (with &gt;2% relative abundance) and sample points</td>
<td>44</td>
</tr>
<tr>
<td>2.6. Heat map showing taxa with greater than 2% relative abundance at each sampling time point</td>
<td>45</td>
</tr>
<tr>
<td>2.7: SparCC network map showing significant (p&lt;0.05) of 122 Interactions with a 0.85 correlation between different OTUs incorporating all time points</td>
<td>47</td>
</tr>
<tr>
<td>3.1. Aerial view of Lake Velence and its location within Hungary indicated by the star on the inset map</td>
<td>56</td>
</tr>
<tr>
<td>3.2. PCA of the six sampling time points demarcated by the arrows</td>
<td>67</td>
</tr>
<tr>
<td>3.3: Relative abundance of A) Eukaryotic taxa by phylum and B) Bacterial taxa by phylum with at least a 2% relative abundance over the 6 week period</td>
<td>69</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES – CONTINUED

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4. Heat map showing the Pearson correlation between Eukaryotic taxa (listed as genera when possible) with &gt;2% relative abundance at sampling time points</td>
<td>71</td>
</tr>
<tr>
<td>3.5. Canonical Correspondence Analysis (CCA) plotting chemical variables that drive variation in Eukaryotic genera (with &gt;2% relative abundance) and sample points</td>
<td>73</td>
</tr>
<tr>
<td>3.6. 1% TAE gel of viral amplicons for MCP</td>
<td>75</td>
</tr>
<tr>
<td>3.7. Heat map showing the Pearson correlation between Bacterial phyla with &gt;2% relative abundance at sampling time points</td>
<td>76</td>
</tr>
<tr>
<td>3.8. Canonical Correspondence Analysis (CCA) plotting chemical variables that drive variation in Bacterial phyla (with &gt;2% relative abundance) and sample points</td>
<td>80</td>
</tr>
<tr>
<td>3.9. SparCC network map showing significant (p&lt;0.05) of interactions with a 0.99 correlation between different OTUs incorporating each time point</td>
<td>82</td>
</tr>
<tr>
<td>4.1. Aerial schematic of the lagoons at the City of Logan Wastewater Treatment Plant</td>
<td>119</td>
</tr>
<tr>
<td>4.2. A PCA of the chemical variables responsible for causing the most variation between the lagoons</td>
<td>128</td>
</tr>
<tr>
<td>4.3. The log of algal cells per mL and fatty acid methyl esters quantification for each time point as A) weight (%w/w) and B) volume (g/L). In A) September 2014</td>
<td>130</td>
</tr>
<tr>
<td>4.4. A bar graph of the eukaryotic, archaeal, and bacterial Hill numbers for September 2013 (S13), November 2013 (N13), May 2014 (M14), and September 2014 (S14) in Lagoons, A-E. S13A = September 2013/Lagoon A sample, etc.</td>
<td>133</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES – CONTINUED

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5.</td>
<td>Heatmap showing the Pearson correlation between all bacterial phyla with greater than 2% relative abundance for each lagoon culture.</td>
<td>135</td>
</tr>
<tr>
<td>4.6.</td>
<td>A Canonical Correspondence Analysis (CCA) of bacterial phyla with greater than 2% relative abundance.</td>
<td>137</td>
</tr>
<tr>
<td>4.7.</td>
<td>Heatmap showing the Pearson correlation between all eukaryotic genera with greater than 2% relative abundance for each lagoon culture.</td>
<td>139</td>
</tr>
<tr>
<td>4.8.</td>
<td>A Canonical Correspondence Analysis (CCA) of eukaryotic genera with greater than 2% relative abundance.</td>
<td>140</td>
</tr>
<tr>
<td>4.9.</td>
<td>A bar graph of the viral Hill numbers for both 0.01% and 1% relative abundances for September 2013 (S13), November 2013 (N13), May 2014 (M14), and September 2014 (S14) in Lagoons A-E. S13A = September 2013/Lagoon A sample, etc.</td>
<td>141</td>
</tr>
<tr>
<td>4.10.</td>
<td>Heatmap showing the Pearson correlation between viral OTUs with greater than 1% relative abundance for each lagoon culture.</td>
<td>142</td>
</tr>
<tr>
<td>4.11.</td>
<td>A Canonical Correspondence Analysis (CCA) of viral OTUs with greater than 1% relative abundance.</td>
<td>144</td>
</tr>
<tr>
<td>4.12.</td>
<td>Output from LEFSe analysis using the log10 of LDA scores to list the differential taxa/OTUs associated with periods of high and low FAME by volume (g/L) with A) viral OTUs &gt;0.01% and B) viral OTUs &gt;1%.</td>
<td>146</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES – CONTINUED

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.13. A SparCC network map showing significant (p&lt;0.05) of interactions with a 0.99 correlation between different OTUs incorporating all ponds for the November time point</td>
<td>148</td>
</tr>
</tbody>
</table>
Algal based biofuel has the potential to aid in offsetting future fossil fuel consumption and demand, and lowering CO$_2$ emissions. Cultivation strategies are a pivotal component of achieving high biomass yield. Open outdoor pond systems are currently the most economically viable method for large-scale algae cultivation due to less energy for maintenance than closed systems. However, open pond cultivation is subject to microbial colonization, sometimes negatively impacting the algal crop. Thus, large-scale production is hindered by gaps in our fundamental understanding of microbial interactions and ecology. The following research aims to explore the interplay between cultivation methods, nutrient availability, community composition, lipid metabolism, and system ecology and identify cost effective concepts for algal lipid production. Using alkalinity to limit microbial colonization of an open system is investigated in Chapter 2 in which a monoculture of *Chlorella vulgaris* was successfully cultivated. A putative relationship with a *Pseudomonas* sp. was identified in which the exchange of key metabolites could have enhanced algal growth and limited contamination. Such interactions may minimize the need for pesticides and fertilizer subsequently reducing pollution and operating costs. Findings suggested that potentially beneficial algal-bacterial relationships occurring in alkaline conditions supported a productive and stable monoculture. Alkalinity, in addition to nutrient abundance, is further explored in a natural freshwater terminal lake system, presented in Chapter 3. Lake eutrophication coupled with temperature increases led to a toxic cyanobacterial bloom that reduced overall eukaryotic diversity. Insight gained on the interplay between alkalinity, nutrients, and community dynamics from this natural system was then applied to a series of artificial wastewater lagoons Chapter 4. Elevated lipid (g/L) was observed in this system partially facilitated by increased water residence time in the lagoons and elevated nitrogen availability. Differing alga community composition were observed during periods of elevated lipid in addition to higher biomass (cells/mL) suggesting that higher lipid volumes were the result of high biomass concentration and not necessarily the lipid productivity of specific alga taxa. The research presented utilizes traditional ecologic concepts like diversity and contributes to a more comprehensive understanding of community interactions helping to minimize cost, reduce pollution, and ultimately contribute to the realization of viable biodiesel.
CHAPTER 1 - INTRODUCTION

Renewable Energy and Role of Algal Biodiesel

World energy demands are predicted to nearly triple between 2000 and 2050, rising from $350 \times 10^{18}$ Joules/year to $950 \times 10^{18}$ Joules/year with fossil fuels fulfilling the majority of the demand (Fridleifsson 2003). Despite a recent drop in fuel prices, the previous decade was characterized by escalating prices that detrimentally affected the global economy but did little to stunt demand and consumption (Chisti, 2008; Shurin et al., 2013; Smith et al., 2010). In turn, the volatility of fuel prices undermines the value of food and other products (Dillon and Barrett, 2015). Currently, due in part to lower fuel prices, consumption is predicted to escalate from 88 million barrels per day (b/d) in 2011 to 97 million b/d by 2017 (Pankaj, 2016). Inevitably, these prices will inflate, jeopardizing the global economy once more and serving as a tangible reminder of our dependency on a finite resource.

In addition to economic consequences, the extraction and transport of fossil fuels has resulted in severe environmental degradation. A prime example is the Deepwater Horizon Oil Spill that occurred in 2010 releasing an estimated 4.1 million barrels of crude oil from an offshore drilling rig into the Gulf of Mexico (Lamendella, 2014). Even after extraction and transport, the subsequent combustion of liquid fossil fuels (diesel, gasoline, and propane) compounds environmental impacts through the emission of green-house gases, particularly the staggering 457.5 lbs. of CO$_2$ per million British thermal units (Btu) of energy (EIA, 2015). This value is approximately 400 times the amount of carbon global primary productivity can fix annually (Dukes, 2003; Ringsmuth et al., 2016). Increased
atmospheric CO₂ accounted for 78% of global warming potential (GWP) between 1990 and 2010 and made a significant contribution to the effects of climate change, evident in sea level rise and extreme weather anomalies (Trenberth et al., 2015). Despite the threat of economic and environmental peril, the reliance on fossil fuels goes unabated and necessitates the development of renewable alternatives derived from a combination of sources like wind, solar, thermal, and biomass.

The earth supports approximately 713 billion tons of plant and marine single-celled biomass (Kallmeyer et al., 2012). Organisms like plants and algae capture the sunlight energy through photosynthesis. Captured energy in biomass can in turn be converted into other forms like liquid fuels via biochemical/thermochemical refining, fermentation, or transesterification (NREL, 2012). Focusing on liquid fuels derived from biomass, cellulosic ethanol and biodiesel have been successfully synthesized from terrestrial crops such as corn, soy, and palm. However, these have been demonstrated to be environmentally and/or economically unviable due, in large part, to enormous land requirements (Chisti, 2008; Smith et al., 2010). In contrast, biomass derived from algae requires significantly less land area for cultivation and the lipid bodies within algal cells can be harvested and converted into biodiesel. It is estimated that algal biodiesel could replace liquid fossil fuels in the U.S. on 9.5 million acres, a fraction of the approximately 450 million acres of crops in the U.S (Sheehan et al., 1998). Additionally, algae are more productive than terrestrial crops sequestering CO₂ and converting sunlight into biochemical energy at nearly twice the efficiency (Chisti, 2007; 2008; 2013). For example, biodiesel sourced from corn yields 172 L/ha while algae (depending upon strain and conditions) yields ~58,700-137,000 L/ha.
Lipids, the biodiesel precursor, are the product of the biochemical pathway fatty acid synthesis that occurs in algal cells (Dismukes et al., 2008; Sheehan et al., 1998). Algae synthesize two types of lipid, storage lipids and structural lipids. The more energy rich storage lipids typically appear as triacylglycerols (TAG) or are composed of fatty acids. The transesterification of TAG with an alcohol converts the lipid into the final diesel product, methylesters (Sheehan et al., 1998). The subsequent algal biodiesel can be combusted in diesel engines and, unlike ethanol, requires no infrastructure change (Scott et al., 2010).

Despite the great potential algal biodiesel demonstrates as a renewable liquid fuel, the scale up to mass production is hindered by knowledge gaps whose solutions require dynamic and interdisciplinary approaches. The production of algal biodiesel is multifaceted with many contributing variables like nutrients, species selection, growth conditions, lipid extraction/transesterification, and waste disposal. There are opportunities for improvement at every stage of production, ranging from algal cultivation and nutrient supply, to lipid extraction from the algal cells. Thus, mass production of algal biodiesel will demand expertise from a wide spectrum of disciplines ranging from engineering to ecology. An area of research with potential to drastically improve efficiency is the optimization of low-cost algal growth for lipid synthesis. However, limited research has been conducted on algal ecology in this context, and ecology can directly impact lipid metabolism, which turn would impact production scenarios. Therefore, this body of work will focus on microbial community dynamics in natural and man-made systems with the purpose of larger-scale cultivation of biomass and lipid accumulation.
**Strain Selection and Biomass**

There is an estimated one million species of algae, 35,000 of which have been described, offering a wide breadth of physical and metabolic diversity (Borowitzka and Moheimani, 2012; Guiry, 2012). Aiming to identify taxa with the best traits, strain selection for biodiesel is a common approach that considers several factors including lipid composition and content, growth rate, cell size, and environmental tolerances. These traits, like lipid content and composition, vary among species. Some forms of perturbation have been shown to further capitalize on these traits by increasing lipid content up to 60-70% of dry cell weight (Kent and Andrews, 2007). Lipid accumulation in response to perturbation varies across species, and thus requires the assessment of individual strains for biodiesel suitability.

**Stress and Nutrient Deprivation**

Lipids occur in various energy dense forms like fats, glycerides, and waxes. Algae synthesize lipids for the concentrated storage of metabolic energy to be used in unfavorable conditions induced by environmental stressors like temperature extremes and nutrient starvation (Guschina and Harwood, 2006). Thus, lipid production can be elevated by precise applications of stress altering gene expression and causing the cell to abate biomass synthesis diverting carbon supply to lipid synthesis (Fields et al., 2014; Valenzuela et al., 2012). However, application of stress requires metered precision to maintain culture viability as it can stunt biomass production and lead to cell death (Fields et al., 2014; Odum...
et al., 1979). Studies have achieved higher lipid concentrations, specifically triacylglycerol (TAG), by applying stress in controlled systems in the form of nutrient starvation, temperature extremes, radiation, pH change, or osmotic stress at specific point in the growth cycle (Sharma et al., 2012).

These forms of stress have been shown to induce similar metabolic responses among various taxa with the greatest success achieved with nitrogen starvation or a combination of nitrogen and phosphorus depletion (Fields et al., 2014; Guschina and Harwood, 2006). Research on TAG concentrations of various monocultures in closed reactors demonstrated that algae only produced 6-15% of their dry weight as lipid under optimal conditions, but stress could induce lipid composition to be up to 80% of cell weight (Richmond, 2004). As mentioned, the success of stress application relies on the careful control of conditions necessitating the use of enclosed controlled systems referred to as photobioreactors (PBRs).

Photobioreactors

The literature is replete with examples of increased lipid production by algal species cultivated as monocultures in closed PBRs, advantageous because variables like light and nutrients can be tightly regulated (Hu et al., 2008; Rodolfi et al., 2009; Sheehan et al., 1998). As greater interest in algal-based biofuel evolves, various approaches for cultivation are being tested. However, recent reviews have concluded that PBRs may not be commercially viable for large-scale production due to high energy requirements (Chisti, 2013; Smith and Crews, 2013). In contrast, open outdoor pond systems have been shown to be more economically viable while also having applications for wastewater treatment,
discussed in a later section and in Chapter 4. Although PBRs will continue to play an important role in the development of algal biodiesel, they are not commercially viable for mass production (Smith et al., 2010). The logical progression from PBRs that facilitate the external manipulation of gene expression for increased lipid synthesis is to investigate internal means of gene manipulation. By genetically engineering desired traits, algal strains could be made more efficient and productive, but this approach has inherent risks and challenges, which are discussed further in the following section.

**Genetic Engineering**

Algae occupy nearly every photic niche on earth and proffer an enormous range of taxa naturally suited for a gamut of conditions. Given this range of diversity, the necessity of genetic modification to enhance lipid production could be called into question (Rodolfi et al., 2009; Snow and Smith, 2012). Several strains capable of high lipid production have been identified but typically under conditions of stress. To bypass the necessity of applying stress, genetic engineering is believed to be imperative for economic viability (Blatti et al., 2013; Borowitzka and Moheimani, 2012). Manipulating genetic content for enhanced lipid production is contingent upon genome sequencing but as of 2010 there were only ten completed genomes publicly available (Schuhmann and Schenk, 2013) with only the addition of 22 new genomes for one of the most diverse groups on the planet at the conclusion of 2016 (www.ncbi.com). Of these, *Chlamydomonas reinhardtii* is likely the most studied for biofuel application facilitated by the sequencing and genetic transformation of nuclear, chloroplast, and mitochondrial genomes (Borowitzka and Moheimani, 2012). Each of these genomes contain an enormous amount of information
that requires a substantial investment aimed at assigning biological function and location of genes within the genome (Stephens et al., 2015). Efforts to insert fatty acid genes from terrestrial plants into algae has had little success, likely due to gaps in our understanding of algal fatty acid metabolisms (Blatti et al., 2013). Extensive research on carbon fixation and lipid accumulation conducted by Valenzuela et al. (Valenzuela et al., 2012) on the diatom *Phaeodactylum tricornutum* demonstrated the complexity of carbon concentrating mechanisms (CCMs) with several functionally equivalent enzymes that operated under different environmental conditions. Such findings are indicative of the vast amount of research still required for the realization of genetically modified organisms’ (GMO) potential.

In addition to conceptual hurdles associated with GMOs, risk assessments have identified several hazards to this approach. Unlike genetically modified (GM) food crops, algae are microscopic organisms that can unknowingly be transferred and released into the larger environment from closed PBRs with potentially devastating consequences (Henley et al., 2011). In New Zealand the 2004 introduction of the diatom, *Didymosphenia geminata*, generally common to the northern hemisphere, has cost the country an estimated $3 million per year to combat, dominating river ways and fiscally jeopardizing New Zealand’s fishing and tourism industries (Zealand, 2010). Cultivation in open ponds further introduces the risk of horizontal gene transfer allowing GM genes to enter the global algae population unhindered with no *a priori* knowledge of the outcome (Snow and Smith, 2012; Stephens et al., 2015). Although the released gene is unlikely to offer any competitive
advantages in nature and eventually disappear, history is replete with disastrous examples of anthropogenic blunders that permanently altered the system at large.

**Prospective Methodologies to Further Maximize Lipid Potential**

**Identification of Promising Regional Systems and Strains**

The release of the invasive diatom *D. geminata* into New Zealand waterways had long-term detrimental consequences both economically and environmentally. It is vital that the effort to bring algal biodiesel to fruition does not come at the expense of biodiversity. We have identified promising taxa in parts of the U.S., like the alkaliphilic *C. vulgaris* isolated from Soap Lake, WA discussed in Chapter 2, and should take similar approaches to identifying promising taxa in other locales. This prevents the introduction of invasive species into other parts of the world. The research discussed in Chapter 3 investigated the microbial community of the alkaline lake Velence in Hungary. Hungary contains the most numerous and alkaline lakes in Europe that may contain native alga suitable for biodiesel applications that could help the European Union achieve the commitment of 10% renewable transportation fuel by the year 2020 as mandated in the Renewable Energy Directive (European Parliament Council of the European Union, 2009).

**Employing Ecologic Principals to Microbial Systems**

**Diversity and Productivity.** Traditional ecology has demonstrated a direct correlation between diversity and productivity. Although first developed in terrestrial grassland systems, this tenet is thought to apply to microbial systems including algal
cultivation. Demonstrated by mathematical models and field experiments, phytoplankton biodiversity causes increased productivity (Downing and Leibold, 2002; Striebel et al., 2009a; Tilman et al., 1981). More recently in a pair of novel studies, Stockenreiter et al. (Stockenreiter et al., 2011; 2013) demonstrated increased lipid production in a naturally occurring algal community cultivated in open ponds over that of single species cultivated in PBRs. This finding suggests that diversity could play a role in lipid productivity, but the mechanisms are unknown.

Niche Occupancy and Stability. Ecology emphasizes the importance of the relationship between species diversity and ecosystem function and can be applied to algae biomass production by using diversity to promote stability and productivity (Ptacnik et al., 2008). Diversity can be utilized to achieve competitive exclusion in which all functional niches are occupied making the system more resistant to invasion (McGrady-Steed et al., 1997; Tilman, 1977; Tilman et al., 1981). To mitigate invasion, water is treated with antibiotics perpetuating wastewater treatment issues discussed further in the next section (Chisti, 2013; Smith and Crews, 2013; Wang et al., 2013). However, niche occupancy can be used as a tactic for the successful management of algal systems providing stability that would allow for the cultivation of a desired algal species (Kazamia et al., 2012a; Martin et al., 2014; Natrah et al., 2013; Santos and Reis, 2014; Smith and Crews, 2013). This was demonstrated in Chapter 2, in which increasing bacterial diversity may have contributed to the successful cultivation of *C. vulgaris* in an open system (Bell et al., 2016). Numerous studies have documented a positive symbiotic relationship between algal taxa and bacteria (Croft et al., 2005; Sapp et al., 2007; Watanabe et al., 2005; Xie et al., 2013). These
relationships frequently occur as physical associations, in which the bacteria live in the exopolymeric substances (EPS) or “phycosphere” of the alga.

**The Phycosphere** The phycosphere is considered the area immediately surrounding an algal cell (Bell and Mitchell, 1972; Bell, 1983; Sapp et al., 2007; 2008; Yang et al., 2013). Bacteria colonize this niche feeding on extracellular products secreted by the algal cell resulting in a tight physical association, sometimes covering the entire algal cell, as observed on the diatom *Chaetoceros* sp. utilizing confocal microscopy (Figure 1.1).

![Confocal image of the centric diatom Chaetoceros sp. and associated bacteria. The diatom autofluoresces red and bacteria are stained green with SYTO-9.](image)

Figure 1.1: Confocal image of the centric diatom *Chaetoceros* sp. and associated bacteria. The diatom autofluoresces red and bacteria are stained green with SYTO-9.

Bacteria occupying the phycosphere have been shown to increase growth in some algae (Guo and Tong, 2013). The alga not only benefits from exchange of growth promoting and
antibacterial metabolites, but occupation of the phycosphere niche minimizes colonization by parasitic bacteria (Kazamia et al., 2012a; 2012b; Smith and Crews, 2013). The symbiotic nature of this relationship is substantiated by several key metabolite exchanges, such as cobalamin. A large fraction of algae are unable to synthesize cobalamin de novo, a key cofactor in methionine synthesis, and rely on bacterial synthesis and provision (Croft et al., 2006). In turn, algae provide an immediate source of dissolved organic matter (DOM) to the bacteria (Geng and Belas, 2010). These are often highly evolved relationships evident in their complexity. For example, the bacterium Phaeobacter gallaeciensis inhabits the phycosphere of the marine alga Emiliania huxleyi. The alga provides its bacterial symbiont with a carbon source, dimethylsulfoniopropionate (DMSP), and a secure niche space while P. gallaeciensis provides antibacterial metabolites and growth promoting hormones. However, as E. huxleyi cells age, they begin to senesce releasing a metabolite that triggers P. gallaeciensis to become pathogenic by releasing the algicidal toxin Roseobacticides, suggesting a bacterial contribution to algal apoptosis (Seyedsayamdoust et al., 2011a; 2011b).

Not only do these relationships increase productivity but they also promote stability over extended periods of time (Imase et al., 2008). These mutual relationships have applicability for biotechnological applications like biodiesel by offering the potential to artificially select microbial consortia that promote the growth of desired species (Kazamia et al., 2012a; 2014; Natrah et al., 2013; Ortiz-Marquez et al., 2012; Santos and Reis, 2014). Furthermore, such relationships have been shown to have unique capabilities that neither organism could perform individually. For example, the consortia of Chlorella
sorokiniana and a bacteria, related to Microbacterium trichotecenolyticum, were able to degrade propionate in wastewater, but were unable to do so individually (Imase et al., 2008). Previous research has shown higher algal growth rates when grown with associated bacteria (Grossart and Simon, 2007). Relationships between bacteria and algal growth have been studied in lab settings, but very little is known about these relationships in open systems and their effects on rates of lipid production has not been investigated to date. Drawing from the same ecologic principles of diversity, the presence of bacteria may further bolster the stability and productivity of an algal system as demonstrated in Chapter 2.

Improving Stability by Utilizing Alkalinity. In effort to address climate change, there is a growing interest in alkaline systems’ capacity to sequester atmospheric CO₂ through increased dissolved inorganic carbon (DIC). Alkaline water (pH>7.9) absorbs more CO₂ into solution to maintain carbonate equilibrium. In cultures of autotrophic microorganisms, alkaline systems can be a sink for atmospheric CO₂ that results in a more accessible DIC pool for growth (Wetzel, 2001). Thus, alkaline lakes are some of the most productive and diverse in the world (Grant, 2006). As discussed above, diversity can increase algal lipid productivity in systems in which all available niches are occupied contributing to stability and predictability. In addition to niche occupancy, employing extreme environmental conditions can further reduce potential colonization by external microbes that could prove disastrous to targeted algal species. Utilizing high pH systems selects for specialists thus reducing the likelihood of colonization by a broader range of organisms unequipped for life at high pH and providing additional stability (Mendes and
Vermelho, 2013). This concept was investigated in Chapter 2 and 3 in which we monitored community dynamics in two contrasting alkaline systems: an open outdoor raceway pond cultivating the alkaliophilic *Chlorella vulgaris* and the alkaline Lake Velence in Hungary.

**Mass Production**

Systematic stress or starvation techniques to increase lipid content have been successful on monocultures in closed PBRs. PBR research utilizing nutrient replete wastewater suggests that sheer biomass volume could be more important, identifying high initial cell concentration (10^6 cells/mL) to be the most significant factor in productivity measured over 21 days (Gani et al., 2016). Yet, conditions typically required for high lipid production, such as nutrient deprivation, may also result in low biomass concentrations (Choudhary et al., 2016; Fields et al., 2014; Griffiths and Harrison, 2009). Applying what we have learned from PBR experiments to natural systems, one would expect to see substantially higher lipid content in oligotrophic algal communities than in eutrophic. However, work conducted by Stockenreiter and others found that the most eutrophic lake had the highest lipid content over the phosphorus deplete community while Interlandi and Kilham (2001) observed replete nutrients in natural freshwater systems lead to the highest biomass concentrations (Interlandi and Kilham, 2001; Stockenreiter et al., 2011). These findings suggest that for overall lipid productivity, high biomass volume, thus requiring high nutrient availability, could offset the lower lipid production expected in nutrient replete systems. We observed additional support for biomass concentration being of greater importance than lipid composition in the Logan wastewater treatment plant.
(WWTP) lagoons, discussed in Chapter 4. Fatty acid methyl esters (FAME) can be calculated as g/L, factoring in cell number, or as %w/w. We found a correlation between g/L of FAME and cell number and that time points did not coincide with high %w/w of FAME. High %w/w FAME values were due to a limited number of high lipid producing community members while high g/L of FAME was the result of high biomass. Achieving high biomass in an open system will require large volumes of water and ample nutrients, which could hamper scale up efforts.

**Water Limitations for Mass Production**

Open pond systems present the most economic algae cultivation methods known, but required water inputs for commercial-scale outdoor ponds can reach as high as 3,267 million liters per year in large scale ponds (>650 hectares), with other estimates predicting 216L of water for every liter of biodiesel (Borowitzka and Moheimani, 2010; Chisti, 2007; 2013; Shurin et al., 2013; Smith et al., 2010). Regions of the world with an optimal climate for open pond systems are also often severely water limited. For example, drought is common in the western U.S. and the Midwest, which is compounded by the depletion of the region’s major groundwater aquifer, the Ogallala, due to agricultural withdrawals (Subhadra, 2011). In addition, U.S. potable surface water supply and use are almost entirely allocated to various agricultural and municipal sectors (Pate et al., 2011). The use of wastewater for cultivation could help alleviate demands and diminish competition for critical drinking and agricultural water supplies (Borowitzka and Moheimani, 2010; Smith
et al., 2010). However, much work is needed to identify appropriate algal populations and associated communities that can tolerate and thrive in low-quality water.

Viability of Wastewater

Algae have long demonstrated effective removal of excess nutrients and pollutants from wastewater aiding in the control and remediation of eutrophication in natural systems. As of 2012, the Environmental Protection Agency (EPA) estimates that the nation produces 28.3 billion gallons per day (107.1 billion liters per day) of municipal wastewater requiring $271 billion annually for treatment (U.S. Environmental Protection Agency, 2012). Treatment methods vary depending upon the source and content of the wastewater. Large open stabilization ponds or lagoon systems offer an inexpensive method for treating municipal wastewater by allowing natural physical, chemical, and biological processes to breakdown or absorb harmful compounds and excess nutrients in a relatively controlled system (Li et al., 2013). Lagoon wastewater typically contains high concentrations of dissolved organic matter (DOM), an ideal substrate for microbial growth. As bacteria metabolize DOM, CO$_2$ is excreted as waste and subsequently utilized by algae for photosynthesis. Essential to cellular components, nitrogen and phosphorus species in solution are absorbed effectively reducing nutrient loads. Algae further augment the process by adsorbing heavy metals, and degrading some toxic phenolic, polycyclic, and heterocyclic compounds that might otherwise persist in treated water (Christenson and Sims, 2011; Coogan et al., 2007; El-Sheekh et al., 2012). Due to the presence of lipid bodies, algae have been shown to bioaccumulate hazardous lipophilic organic contaminants such as triclosan and hormones (Coogan et al., 2007). Although this research
focused on the bioaccumulation and subsequent trophic transfer of these toxic compounds, utilizing algae’s capacity to absorb pharmaceuticals and toxins for effective wastewater treatment is highly understudied.

The production of algal biomass while simultaneously treating wastewater could be advantageous for both the algal biodiesel and wastewater treatment industries (Christenson and Sims, 2011). Under optimized conditions, an average of 6.5 million L/day of biodiesel using medium strength domestic wastewater operating 9 months of the year could be produced (Christenson and Sims, 2011). This represents a small fraction of the 2.1 billion L/day (12.9 million barrels) of petroleum fuel used for transportation in the U.S. but these wastewater treatment lagoons offer the best opportunity to develop large scale cultivation systems required for the commercial viability of algal biofuel (Christenson and Sims, 2011; U.S. Energy Information Administration, 2016b). Logan, Utah has the largest operating municipal lagoon WWTP in North America. The lagoons cover 460 acres receiving an average of 14 million gallons of wastewater per day providing a unique opportunity to investigate algal lipid production on wastewater in a large operational system (Davies, 2015). Chapter 4 presents the research conducted on these lagoons offering valuable insight into biomass/lipid production in a high nutrient mixed culture open system and while effectively treating municipal wastewater.

**Wastewater Nutrients**

Broadly speaking, municipal wastewater lacks sufficient carbon to generate adequate algal biomass for the total assimilation of nitrogen, whereas phosphorus assimilation can be achieved with far less carbon (Borowitzka and Moheimani, 2012). The
carbon:nitrogen and carbon:phosphorus ratios in domestic sewage C:N 3.5:1; C:P 20:1 and
dairy lagoon water C:N 3:1; C:P 10:1 are low compared to typical ratios in rapidly growing
algae biomass C:N 6:1; C:P 48:1 (Oswald and Gotaas, 1957; Sheehan et al., 1998).
Nitrogen assimilation can be resolved by sparging CO₂ and starting with a more alkaline
system that will also drive atmospheric CO₂ into solution increasing algal productivity by
30-100% (Park and Craggs, 2010). Utilizing alkalinity, natural or induced with CO₂
sparging, the total assimilation of nitrogen is feasible in some wastewater systems
(Benemann, 2003; Borowitzka and Moheimani, 2012).

Nitrogen, changes in nitrogen species’ concentration, and differing bioavailability,
has profound effects on algal metabolism because it is a critical constituent in cellular
components. Especially in terms of algal lipid metabolism, nitrogen has been identified as
the most influential and significant nutrient (Sharma et al., 2012). As a result of the fossil
fuel powered Haber-Bosch process that catalyzes the conversion of atmospheric nitrogen
into nitrate for fertilizer, nitrogen has become prolific even in pristine systems that were
historically nitrogen limited, drastically altering microbial community dynamics (Baron et
al., 2000; Wolfe et al., 2001). Of the nitrogen species, ammonium is the most soluble, is
typically ubiquitous in most types of wastewater, and requires the least amount of energy
to be incorporated into biomass (Markou et al., 2014). Thus, wastewater could provide a
large and reliable source of nitrogen needed for the mass production of algae.

The dominance of ammonium versus ammonia (deprotonated ammonium) is pH
dependent, with pH higher than 9.25 (9.25=pK at 25°C) favoring ammonia (Markou et al.,
2014). Ammonia is the cheapest and most viable source of nitrogen for biomass production,
but at high concentrations it can be toxic to cells (toxicity concentration is species specific) caused by large discrepancies between intra- and extracellular pH facilitating passive diffusion through cell membranes ultimately damaging photosystem II (Markou et al., 2014). For this reason, previous studies utilizing ammonium have emphasized the importance of maintaining a lower pH for algal cultivation (Eustance et al., 2015). However, alkaliphilic algae inherently have a greater resistance to ammonia toxicity due to minimized differences between intra- and extracellular pH (Markou et al., 2014). For example, at pH 9 the alkaliphilic alga *Spirulina platensis* preferentially utilizes ammonia over nitrate and at pH 11 it retained 70% of its maximal photosynthetic capacity whereas a non-alkaliphile only retained 10% (Belkin and Boussiba, 1991). Thus, utilizing wastewater in alkaline conditions near or below a pH of 9.25 offers further advantages by providing high ammonia availability and selecting for taxa with greater resistance to ammonia toxicity.

Phosphorus is another critical constituent to life, found in nearly every metabolic pathway, and is a major component of the cellular currency adenosine triphosphate (ATP) (Elser, 2012). Occurring in relatively low concentrations within the earth’s crust, the natural mobilization of P relies heavily on chemical weathering and thus becomes a limiting nutrient for primary production. Anthropogenic activities have altered P availability, similar to increased N availability due to the Haber-Bosch process, P has been extensively mined from deposits concentrated in just a few countries, the largest found in Morocco (Elser, 2012). Although the exact timing differs vastly between predictions (50 to several hundred years), extractable P is a finite resource that has become integral to
global agriculture including algae cultivation (Edixhoven et al., 2014). Approximately 40 million tons of phosphate is mined annually from the estimated 2.8 billion tons in the U.S., a number that would have to nearly double if algal biodiesel were to replace U.S. petroleum consumption (Hannon et al., 2010). Despite the impending shortage of minable P, there is a very large supply of N- and P-rich wastewater due to industrial processes and municipal waste (Markou et al., 2014; Smith et al., 1999).

**Conclusion**

Examine each question in terms of what is ethically and esthetically right, as well as what is economically expedient. A thing is right when it tends to preserve the integrity, stability, and beauty of the biotic community. It is wrong when it tends otherwise.

-Aldo Leopold (Leopold, 1989)

The interplay between cultivation methods, nutrient availability, community composition, and lipid metabolism are staggering in their complexity. Employing tactics that limit colonization in open systems, such as diversity and alkalinity, could help to reduce some of this complexity. Utilizing the lens of established ecologic tenets to further examine these dynamics provides a relevant and verified framework for experimentation as ultimately, “all of the concepts that ecologists use to describe the continental-scale ecosystems that we see through satellites also apply to ecosystems that we peer at with microscopes” (Yong, 2016). The purpose of the following chapters is to explore cost effective concepts for stable and productive algal lipid production in an ecologic context.
Productivity and stability are commodities that can be sourced from system diversity, echoing that ecology is inherently economy, as advocated by Aldo Leopold. More recently, beneficial goods and services, like food production and waste treatment, resulting from biologic processes has been termed “ecosystem services” and facilitated the fiscal valuation of previously arbitrary concepts (Chapin et al., 1997; Costanza et al., 1997; Edwards and Abivardi, 1998). Typically, diversity increases a system’s value by expanding its functional capacity resulting in increased productivity and delivery of vital services (Beyter et al., 2016). Diverse functional capacity means an array of niche spaces are occupied, thereby reducing successful colonization and contributing to overall system stability. However, artificial algal systems may challenge this concept’s veracity, depending on the scale at which diversity is quantified. Case in point, should diversity be quantified across all three domains of life or be limited to the domain, or kingdom, or phylum, etc. of study? Excluded from the tree of life, how should the viral component of a community be considered? If viral diversity is high at the expense of other community members would the system still be considered diverse and therefore productive? Single alga cells can be home to a high level of bacterial diversity but if the alga is cultivated as a monoculture the system’s overall diversity would be considered low and should result in lower productivity.

The research in Chapter 2 challenges the paradigm that less diverse systems are more vulnerable to invasion by monitoring a stable monoculture of *C. vulgaris* cultivated in an open system for 16 days. This observation suggested that alkalinity could be used as a tool to limit colonization of open niche space and maintain a viable monoculture. The
alga may have also benefitted from exchanges in the phycosphere with *Pseudomonas* sp. that may have further contributed to culture stability. This poses the question, does the success of a monoculture rely upon the diversity of the phycosphere? In algal systems, this answer may depend upon the scale in which diversity is quantified.

Transitioning from a small-scale, man-made system, the stability and community dynamics of bacteria and eukaryotes in a natural alkaline system were investigated. Despite having much higher levels of diversity than the raceway investigated in Chapter 2, the stability of a Hungarian alkaline lake was perturbed by an anthropogenic nutrient source and elevated temperature fostering a substantial and toxic cyanobacterial bloom. This observation underscores how influential nutrient abundance is in the diversity and stability dynamic then prompted inquiry into the role of nutrients on the diversity and productivity dynamic.

Investigation on the role of nutrients in diversity and productivity required a nutrient replete system achieved in a series of connected lagoons at a municipal wastewater treatment plant (Chapter 4). Findings suggested that ample nutrient availability contributed to productivity quantified as lipid concentration and cell number. Eukaryotic, bacterial, and even phycoviral diversity positively correlated with lipid productivity, suggesting that sufficient lipid for biodiesel could be achieved while simultaneously treating wastewater. These findings contribute new approaches by resurrecting established ecologic tenets and applying them to large-scale algal cultivation, challenging the perception that PBRs, nutrient deprivation, and monocultures are the sole means of lucrative lipid production and substantiating the importance of ecology in outdoor cultivation.
CHAPTER 2

A LIPID-ACCUMULATING ALGA MAINTAINS GROWTH IN OUTDOOR, ALKALIPHILIC RACEWAY POND WITH MIXED MICROBIAL COMMUNITIES

Contribution of Authors and Co-Authors

Manuscript in Chapter 2

Author: Tisza Ann Szeremy Bell

Contributions: Collected field data, conducted most experiments, analyzed data and wrote manuscript

Co-author: Bharath Prithiviraj

Contributions: Offered advice on experimental design and conducted some of the downstream sequencing analysis

Co-author: Brad D. Wahlen

Contributions: Oversaw the experimental design and supervised sample collection

Co-author: Matthew W. Fields

Contributions: Principal Investigator

Corresponding Author: Brent Peyton

Contributions: Principal Investigator
Author and Co-authors: Tisza Ann Szeremy Bell, Bharath Prithiviraj, Brad D. Wahlen, Matthew W. Fields, and Brent M. Peyton
Journal: Frontiers in Microbiology
Status of Manuscript:
  ____ Prepared for submission to a peer-reviewed journal
  ____ Officially submitted to a peer-review journal
  ____ Accepted by a peer-reviewed journal
  ___ Published in a peer-reviewed journal
Publisher: Frontiers Media
Date of Submission: August 2, 2015
Date of Publication: January 7, 2016
Issue: 6:1480
Abstract

Algal biofuels and valuable co-products are being produced in both open and closed cultivation systems. Growing algae in open pond systems may be a more economical alternative, but this approach allows environmental microorganisms to colonize the pond and potentially infect or outcompete the algal “crop”. In this study, we monitored the microbial community of an outdoor, open raceway pond inoculated with a high lipid-producing alkaliphilic alga, *Chlorella vulgaris* BA050. The strain *C. vulgaris* BA050 was previously isolated from Soap Lake, Washington, a system characterized by a high pH (approximately 9.8). An outdoor raceway pond (200 L) was inoculated with *C. vulgaris* and monitored for ten days and then the culture was transferred to a 2,000 L raceway pond and cultivated for an additional six days. Community DNA samples were collected over the 16-day period in conjunction with water chemistry analyses and cell counts. Universal primers for the SSU rRNA gene sequences for *Eukarya*, *Bacteria*, and *Archaea* were used for barcoded pyrosequence determination. The environmental parameters that most closely correlated with *C. vulgaris* abundance were pH and phosphate. Community analyses indicated that the pond system remained dominated by the *Chlorella* population (93% of eukaryotic sequences), but was also colonized by other microorganisms. Bacterial sequence diversity increased over time while archaeal sequence diversity declined over the same time period. Using SparCC co-occurrence network analysis, a positive correlation was observed between *C. vulgaris* and *Pseudomonas sp.* throughout the experiment, which may suggest a symbiotic relationship between the two organisms. The putative relationship coupled with high pH may have contributed to the success of *C. vulgaris*. The
characterization of the microbial community dynamics of an alkaliphilic open pond system provides significant insight into open pond systems that could be used to control photoautotrophic biomass productivity in an open, non-sterile environment.

Introduction

The cultivation of algal biomass has many industrial applications ranging from health-care products to biodiesel (Fields et al., 2014). As the number of applications and subsequent demand for biomass increases, a challenge will be to exponentially increase production cost-effectively. The use of alkaliphilic photoautotrophs may help overcome some of the constraints associated with large-scale biomass production in open systems due to limited niche accessibility caused by higher pH values. Algal production of triacylglycerol (TAG) and other lipids are of substantial interest because of being biodiesel precursors that can be transesterified into fatty acid methyl esters (FAMEs) (Dismukes et al., 2008; Scott et al., 2010; Sheehan et al., 1998). Many studies have screened algal species for biodiesel applications based on high lipid content (Dismukes et al., 2008; Griffiths and Harrison, 2009; Griffiths et al., 2011; Scott et al., 2010; Sheehan et al., 1998), and Chlorella vulgaris has been identified as one such species (Duong et al., 2012; Griffiths et al., 2011; Li et al., 2008).

Many studies on C. vulgaris and other species have been conducted in closed photobioreactors (PBRs) whose overall costs are considered a major constraint in the scale-up of algal biodiesel production (Chisti, 2013; Kazamia et al., 2012a; Shurin et al., 2013; Smith et al., 2010). However, in an effort to scale up production at lower capital
investments, open pond systems have been shown to be a viable alternative (Smith et al., 2010). These open systems are, however, prone to colonization by environmental microbes spanning all three domains, and may contain hundreds of distinct taxa whose relative abundances vary by orders of magnitude (Fulbright et al., 2014). By employing parameters from “extreme” environments observed in natural systems, such as high pH, unwanted colonization (e.g., invasion) may be limited (Georgianna and Mayfield, 2012; Selvaratnam et al., 2014; Wang et al., 2013). In addition, alkaline systems favor higher dissolved inorganic carbon (DIC) from atmospheric CO$_2$ thereby providing increased carbon delivery for primary producers. Moreover, some of the highest primary production rates have been reported for microbial, alkaline systems (Hem, 1985; Melack and Kilham, 1974).

Bacterial colonization has benefits and drawbacks for biomass production that are determined by the system’s community structure and composition. Numerous studies have documented positive, symbiotic relationships between algal taxa and bacteria (Croft et al., 2005; Sapp et al., 2007; Watanabe et al., 2005; Xie et al., 2013). Specifically, different species of *Pseudomonas* have been observed living in association with algae, including *C. vulgaris* (Sapp et al., 2007). The physical association, in which the bacteria live in the exopolymeric substances (EPS) or “phycosphere” of the alga, has been shown to increase *C. vulgaris* growth (Guo and Tong, 2013). The alga not only benefits from exchange of growth promoting and antibacterial metabolites in the niche space of the phycosphere, but also via the exclusion of potential opportunistic pathogens (Kazamia et al., 2012a; 2012b; Smith and Crews, 2013). Additional work has shown symbiosis to be critical for adaptation to thermal stress resulting in higher algal biomass (Xie et al., 2013). These mutual positive
relationships may have benefits for biotechnological applications offering the potential to artificially select microbial consortia that promote the growth of desired species (Kazamia et al., 2012a; 2014; Natrah et al., 2013; Ortiz-Marquez et al., 2012; Santos and Reis, 2014).

In the described study, we utilized pyrosequencing to monitor fluctuations in the community structure of an outdoor raceway pond inoculated with C. vulgaris during scale-up from 200 L to 2,000 L. The results indicated that the inoculated algal population could maintain predominance under alkaline conditions, and that bacterial diversity increased while archaeal diversity decreased over time. In addition, particular populations could be correlated with C. vulgaris.

**Methods**

**Site Description and Raceway Pond Conditions**

Outdoor ponds were located in Logan, Utah (July 2011) approximately 40 km west of the northern arm of the Great Salt Lake (GSL). The 200 L oblong pond manufactured by Separations Engineering Inc. was lined with fiberglass and equipped with a paddle wheel promoting gas exchange with ambient air (Separations Engineering Inc., San Diego, CA). Initially, a 200 L raceway was inoculated with C. vulgaris (10% v/v), maintained at a culture depth of 13 cm, and monitored for 10 days. On day 10, the entire culture was transferred into an adjacent 2,000 L raceway and maintained at a culture depth of 20 cm for the remaining 6 days of the experiment. The 2,000 L raceways were constructed of cinder blocks stacked two high with 46 mil EPDM rubber pond liner creating the pond. Marine board was used to divide the pond into a circulating raceway with a paddlewheel
providing circulation. The ponds were inoculated with *Chlorella vulgaris* BA050 that was previously isolated from Soap Lake, Washington, which is characterized by growth at high pH (approximately 9.8) (Dimitriu et al., 2008). The isolate was maintained on agar plates and was streak isolated during each subsequent plating every 2 months. The 18S rRNA gene sequences obtained from isolated DNA confirmed the presence of a single eukaryotic microorganism. A single 200 L raceway was inoculated with a 20 L culture (10% volume) that had been previously cultivated in shaker flasks bubbled with 1% CO₂. A more saline version of Bold’s Basal Medium, consistent with the salinity of seawater (35 ppt), at pH 8.7 was prepared under non-sterile conditions with the addition of dry salts and concentrated solutions (Nichols and Bold, 1965). Inoculation resulted in a cell density of 3.6E+6 cells/mL with the addition of sufficient medium to bring the total volume to 200 L. The pH was not controlled in the pond. Unfiltered tap water was added each day to replace measured evaporative loss. After 10 days, repeating methods from the initial inoculation, the 2,000 L pond was inoculated by transferring 200 L (10%) of culture from the first pond (200 L).

**Sample Collection**

Samples were collected twice daily for cell density ascertained by OD<sub>750</sub> and direct cell counts via optical microscopy (cells/mL). Additionally, 500 mL samples were collected and frozen at -80°C for DNA extraction and 454 sequencing.
DNA Extraction and Sequencing

DNA Extraction. Samples were slowly thawed at 4°C and microbial biomass was collected via filtration through 0.22 μm polyethersulfone membrane filters. The solids were then suspended in the MO BIO PowerMax™ Soil DNA Isolation Kit PowerBead Solution, and the cells were lysed via three cycles of liquid nitrogen freeze–thaw and ground with a mortar and pestle aided by sterile sand (Zhou et al., 1996) (MO BIO Laboratories Inc., Carlsbad, CA, USA). The DNA was cleaned and concentrated with the Wizard® SV Gel and PCR Clean-Up System (Promega Corporation, Madison, WI, USA) according to the manufacturer's protocol.

Bar-Coded. Pyrosequencing. Pyrosequencing was utilized to characterize the microbial population of the ponds. PCR was used to increase the DNA concentration needed for pyrosequencing analysis. Each sample was labeled with a unique 10 nucleotide-barcode for multiplexing. The SSU rRNA gene sequences for Eukarya and Bacteria were amplified via 25 cycles of PCR with the following barcoded primers; 7F (5’-ACCTGGTTGATCCTGCCAG-3’) and 591R (5’-GGAGCTGGGAATTACCG-3’) for Eukarya and FD1 (5’AGAGTTTGATCCTGGCTCAG-3’) and 529R (5’-CGCGGCTGCTGGCAC-3’), which targeted the V1-V3 region of Bacteria (Bowen De León et al., 2012). Archaeal sequences were amplified separately from Bacteria using a nested approach with non-barcoded 21F (5’-TTCYGGTTGATCCYGCCRGA-3’) and 1492R (5’-CGGTTCCTTGTTACGACTT-3’) for 20 cycles followed by an additional 20 cycles with barcoded 751F (5’CCGACGTTGAGRYGAA-3’) and 1204R (5’-
TTMGGGGCATRCNKACCT-3') (Baker et al., 2003; Barnhart et al., 2013). PCR products of the correct size were confirmed using a 1% agarose gel. Products were cut from the gel and pooled using an Ultrafree®-DNA gel extraction column (Millipore Corporation, Bedford, MA, USA). The gel extract was cleaned and concentrated using the Wizard® SV Gel and PCR Clean-Up System, and dsDNA was quantified with a Qubit fluorometer (Invitrogen, Carlsbad, CA, USA). Adaptors for 454 sequencing were ligated to the amplicons and were pyrosequenced on a 454 GS-Junior (454 Life Sciences, Branford, CT, USA). Roche's image analysis separated sequences by barcode. Sequences were trimmed to one standard deviation below the mean length or removed if shorter. Employing the Phred score filter, 15% of the nucleotides were allowed to be below Q27, and removed if primer errors or Ns were observed.

Bioinformatic Sequence and Community Analysis

Data analysis was performed using the Quantitative Insights into Microbial Ecology (QIIME) software package, version 1.4.0 (Caporaso et al., 2010b). Parameter settings for demultiplexing were at a default level of 200 bp and 1000 bp in length. Metadata files were prepared according to a QIIME compatible template taking into account environmental sampling data on pH, temperature, and ionic concentrations. Libraries were split according to barcode for each of the respective domains (Archaea, Bacteria, and Eukarya). Sequences were then concatenated for data normalization needed in downstream analysis. Operational taxonomic units (OTUs) were assigned using the closed reference OTU picking protocol. Clusters were referenced against the Silva 108 database and pre-clustered at 97% identity using UCLUST (Edgar, 2010).
Sequence reads that matched a Silva reference sequence at 97% identity were clustered within an OTU defined by a reference sequence. OTU assignment (and all subsequent steps) was performed for the combined Archaea, Bacteria and Eukarya reads. The singleton OTUs were discarded. The centroid sequence in every cluster was selected to represent the cluster and aligned with the Silva core set using PyNAST (Caporaso et al., 2010a). Chimeric sequences, identified with Chimera Slayer (Haas et al., 2011) and reads that failed to align with PyNAST were excluded from subsequent analyses. PyNAST (v1.1) was used for sequence alignment and filtering through QIIME using default parameters.

Taxonomic assignments were additionally assigned using the retrained RDP Classifier (Wang et al., 2007) on the Silva 108 database for phylogenetic resolution at the genus level. Taxonomic summary for Archaea, Bacteria, and Eukarya was plotted to the genus level with a given relative abundance based on diversity and distribution pattern per domain. Taxa distribution was also summarized by time (sample day). The β- diversity analysis downstream between samples was derived using UniFrac (Lozupone and Knight, 2005) that took into account the phylogenetic structure of the algal pond microbial communities. Taxonomic richness was calculated by rarefaction analysis based upon OTU tables that were rarefied at an even sampling threshold value. Richness was measured on the basis of the Chao index (Chao et al., 2010).

The co-occurrence of community members was illustrated in a heat-map using the R vegan package version 2.0-10 (Oksanen, 2011). Due to the fact that some taxa had a 0% relative abundance at certain time points, 0.1 was added to all values in order to be log transformed. In an effort to enhance the visual distribution of taxa, log transformed values
were cubed and resulting values plotted \((\log(\text{relative abundance } + 0.1)^3)\). A SparCC analysis was subsequently used to construct community correlation networks by estimating linear correlation values between log transformed abundances based on the absolute number of sequences for an OTU rather than a relative abundance (Berry and Widder, 2014; Friedman and Alm, 2012). The key advantage of this analysis was that, for instance, the ratio of the fractions of two OTUs was independent of the fluctuations in other OTUs included in the analysis \((i.e., \text{subcompositional coherence})\) (Friedman and Alm, 2012). Archaeal taxa were not included in this analysis due to the sharp decline in relative abundance between days 1 and 3 and near absence by the end of the pond experiment. We observed that the drastic decline in Archaeal relative abundance would have produced deceptive relationships in the network model.

**Statistical Analysis**

A principal coordinate analysis (PCA) was used to reduce dimensionality and give structure to the water chemistry variables obtained from each time point (Legendre and Gallagher, 2001). In order to incorporate taxonomic data, we used the direct-gradient ordination technique, Canonical Correspondence Analysis (CCA), which concurrently showed pond taxa, time points, and water chemistry (Hall and Smol, 1992). This kind of ordination is appropriate when assessing community dynamics because it does not use Euclidean based metrics that assume linear trends in community change. Not only does the ordination show environmental factors influencing community change, but results suggest potential interactions between taxa (Amaral-Zettler et al., 2010). The first two axes, CCA1 and CCA2, typically account for the majority of observed variation. All axes
are constrained to present a linear combination of the water chemistry that maximizes the dispersion of taxa (Hall and Smol, 1992).

Results and Discussion

Environmental Variables

Water chemistry was monitored daily over the course of the experiment and fluctuations were used to draw correlations to community structure (Table 2.1). The nitrate concentration was lowest on day 8 as *C. vulgaris* achieved stationary phase in the 200 L raceway. Changes in fluoride and chloride anion concentrations at this time may have been due to the use of tap water to compensate for evaporative loss. Nitrogen concentrations recover on day 10 when the culture is transferred to the 2,000 L raceway and combined with new media. The pH values remained high over time and may have benefited *C. vulgaris* ($R^2=0.2$). The decline in pH from 10.7 to 8.39 on day 11 corresponded with the transfer of the pond from the 200 L to 2,000 L raceway.
Table 2.1: Water chemistry over the course of the pond run. Values are in mg/L. B.d. stands for below detection. Volume of pond indicated by 200 L or 2,000 L.

<table>
<thead>
<tr>
<th>(mg/L)</th>
<th>Day 1 (200 L)</th>
<th>Day 3 (200 L)</th>
<th>Day 7 (200 L)</th>
<th>Day 8 (200 L)</th>
<th>Day 11 (2,000 L)</th>
<th>Day 16 (2,000 L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-</td>
<td>0.05</td>
<td>42.9</td>
<td>71.6</td>
<td>0.6</td>
<td>71.6</td>
<td>225.7</td>
</tr>
<tr>
<td>Cl-</td>
<td>8092</td>
<td>1405</td>
<td>1294</td>
<td>2460</td>
<td>1114</td>
<td>537</td>
</tr>
<tr>
<td>NO₃⁻</td>
<td>13.02</td>
<td>60.14</td>
<td>40.98</td>
<td>8.68</td>
<td>210.24</td>
<td>191.50</td>
</tr>
<tr>
<td>SO₄²⁻</td>
<td>427.20</td>
<td>277.44</td>
<td>275.52</td>
<td>345.60</td>
<td>300.48</td>
<td>335.04</td>
</tr>
<tr>
<td>PO₄³⁻</td>
<td>10.56</td>
<td>0.96</td>
<td>b.d.</td>
<td>0.29</td>
<td>9.60</td>
<td>6.72</td>
</tr>
<tr>
<td>P</td>
<td>5.92</td>
<td>2.85</td>
<td>1.46</td>
<td>1.14</td>
<td>5.26</td>
<td>2.90</td>
</tr>
<tr>
<td>Ca</td>
<td>52.4</td>
<td>22.1</td>
<td>14.8</td>
<td>12.4</td>
<td>41.6</td>
<td>20.3</td>
</tr>
<tr>
<td>K</td>
<td>198</td>
<td>226</td>
<td>262</td>
<td>213</td>
<td>302</td>
<td>382</td>
</tr>
<tr>
<td>Mg</td>
<td>93.1</td>
<td>116.0</td>
<td>110.0</td>
<td>97.9</td>
<td>144.0</td>
<td>178.0</td>
</tr>
<tr>
<td>Na</td>
<td>4734</td>
<td>5316</td>
<td>5527</td>
<td>4740</td>
<td>6211</td>
<td>7483</td>
</tr>
<tr>
<td>S</td>
<td>157</td>
<td>150</td>
<td>148</td>
<td>135</td>
<td>175</td>
<td>219</td>
</tr>
<tr>
<td>pH</td>
<td>8.24</td>
<td>9.77</td>
<td>10.41</td>
<td>10.69</td>
<td>8.39</td>
<td>10.02</td>
</tr>
</tbody>
</table>

Community Composition and Interaction

The SSU rRNA for Bacteria, Archaea, and Eukarya were amplified and sequenced for each of the six sample days. After screening sequences for errors (see Methods), 161,731 quality gene sequences with a valid barcode were retrieved. Sequences with a 97% identity were clustered within an OTU totaling 1,349 observed OTUs composed of 748, 249, and 352 OTUs of Bacteria, Archaea, and Eukarya respectively for all sampled days (Table 2.2).
Table 2.2: 454 Pyrosequencing statistics.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of samples</td>
<td>18</td>
</tr>
<tr>
<td>Number of OTUs</td>
<td>2,074</td>
</tr>
<tr>
<td>Number of total sequences</td>
<td>161,731</td>
</tr>
<tr>
<td>Minimum sequences/sample</td>
<td>303</td>
</tr>
<tr>
<td>Maximum sequences/sample</td>
<td>22,256</td>
</tr>
<tr>
<td>Sequences/sample</td>
<td>8,985</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>5,626</td>
</tr>
<tr>
<td>Even sampling depth</td>
<td>3,068</td>
</tr>
</tbody>
</table>

Figure 2.1 shows the relative abundances for the observed archaeal and bacterial taxa.

![Figure 2.1](image)

Figure 2.1: Total relative abundances of A) archaeal and B) bacterial taxa by order (as order, genus, or unknown) with >2% relative abundance over the 16 day sample period.
Figure 2.1 shows phosphate limitation in the plateau of *C. vulgaris* observed from day 3-8, emulating nutrient conditions frequently observed in wastewater and some natural systems. Nitrogen availability has more than doubled due to anthropogenic inputs facilitated by the Haber-Bosch process entering aquatic systems through precipitation, runoff, and dust deposition (Baron et al., 2000; Brahney et al., 2015; Fenn et al., 2003; Gardner et al., 2008; Miller and McKnight, 2012; Wolfe et al., 2001). Thus, phytoplankton is typically phosphorus limited giving a competitive advantage to taxa that can quickly scavenge phosphorus (Baron et al., 2000; Elser et al., 2009; Elser et al., 2012; Wolfe et al., 2001). In addition, some *C. vulgaris* strains have been shown to accumulate polyphosphate (Aitchison and Butt, 1973). Considering that nitrogen inputs are nearly unavoidable in the majority of open systems, most communities will be phosphorus limited, the effects of which are pertinent to endeavors such as commercial algae production in open systems. Moreover, photoautotrophs that have the ability to accumulate and/or scavenge phosphorus will be more competitive in open, mixed systems in which low phosphate levels can be used. Combined, these attributes reduce the need for higher levels of phosphorus and/or the use of low-quality phosphorus.
Effects of phosphorus limitation on *C. vulgaris* did not appear to detrimentally affect ability to compete and grow under the tested conditions as it was observed to be the single dominant Eukaryote during the 16-day experiment. Figure 2.2 shows that *C. vulgaris* recovered from phosphorus limitation upon transfer to the 2,000 L pond, and throughout the pond experiment it was the dominant eukaryotic taxon and did not decline below 93% of eukaryotic relative abundance. The other 7% was composed of pine pollen, insects, and fungi. While it is difficult to ascertain physiology from phylogeny, sequences indicative of *Psuedomonas* were inversely correlated to high phosphate levels (Istvánovics, 2008). Organisms from this genus can be phosphate-accumulating and/or grow under low-phosphate conditions (Sidat et al., 1999). It is likely that phosphate-accumulating bacterial populations would be selected as overall P levels are depleted. As expectations for
inexpensive biomass and feedstock become greater, we will need improved insight into biological responses to low-level and low-quality phosphorus.

Other organisms propagated in the pond but did not appear to have a detrimental effect on the alga population. Several halophilic archaeal taxa were observed in the first 3 time points, especially day 1 leading to the greatest observed diversity that declined over time (Figure 2.3). In addition to the Archaea, several bacteria were present in the first sample. The high archaeal and bacterial diversity did not appear to have a negative effect on *C. vulgaris*, which was the predominant eukaryote, and minimal fluctuation was observed in the eukaryotic diversity. As the Archaeal taxa diversity declined, the bacterial diversity increased (Figure 2.3). It is not known if the decline in archaeal populations was related to the increase in bacterial populations or an independent process, such as lower salinity.

The transitory presence of halophilic taxa could have been due to suboptimal conditions in the ponds. These halophilic microorganisms have an optimal pH range from 6.8 to 9, NaCl concentration between 3.4 and 5.1 M, and generally require at least 0.85 to 3.4 M in order to maintain osmotic pressure for cell integrity and prevent lysis (Bowers and Wiegel, 2011; Cui et al., 2012; DasSarma and DasSarma, 2006; Oren, 2006). Their presence in the ponds could potentially be due to the proximity of the pond to the Great Salt Lake (GSL) that is 40 km to the west and characterized by high salinity and a neutral pH (Baxter et al., 2005; Oren, 1994; 2006; Post, 1977; Tazi et al., 2014). Wind dispersal and precipitation events may have been responsible for the presence of these populations in the pond.
Figure 2.3: Chao diversity for each domain plotted over time. Archaea started with high diversity but quickly declined. Eukarya maintained a steady diversity level composed almost entirely of C. vulgaris. Bacterial diversity steadily increased with time.

We observed numerous haloarchaeal genera including *Halorubellus*, *Haloquadrata*, *Halalkalicoccus*, *Candidatus*, *Halomonas*, *Halobacterium*, and *Haloarcula* and the halobacterial genera *Devosia*, *Aliihoelea*, *Halomonas sp.*, *Seohicola*, *Erthromicrobium*, *Aquiflexum*, and *Rhodobacterales* which have also been found in GSL samples (Baxter et al., 2005; Oren, 1994; 2006; Post, 1977; Tazi et al., 2014). However, another possibility is that these taxa were already present in the salts used to make the medium. The vast majority of these taxa subsisted for only the first two time points (Figure 2.3). It is unknown if the detected sequences were indicative of populations that survived for a given time in the test pond or simply were static and/or dead cells that were transported to the ponds.
In contrast to the halophilic microorganisms, *C. vulgaris* has been shown to be inhibited by concentrations greater than 1 M NaCl and showed substantial declines in cell concentration at 0.5 M (Alyabyev et al., 2007). The highest recorded salinity in the ponds was 0.7 M with an average 0.2 M. Thus, the success of *C. vulgaris* in the ponds further supports that salinity was below the presumptive optima of the halophiles. The changes in archaeal and bacterial community structure did not appear detrimental to *C. vulgaris* and a high relative abundance (>93%) was observed throughout the course of the pond experiment. A decrease in cell number was observed immediately after pond transfer and was likely due to dilution of cells during the transfer of the 200 L inoculum to the 2,000 L pond; however the relative abundance of *C. vulgaris* remained consistent (94-95% relative abundance).

Correlation of Environmental Variables and Community Structure

Environmental factors could have also played a role in the success of *C. vulgaris*. We observed high pH values which ranged between 8.2 and 10.7 (Table 2.1), which may have prevented other algal taxa from successfully colonizing the pond. This finding demonstrates that the cultivation of a single algal strain in an open alkaline pond without the addition of antibiotics or herbicides can be successful (Lundquist et al., 2010; McBride et al., 2014; Smith and Crews, 2013; Smith et al., 2010).

Using canonical correspondence analysis (CCA), the most significant water chemistry variables correlated with fluctuations in archaeal taxa were plotted in Figure 2.4. Taxa and time points were correlated to pH, nitrate, and phosphate (vectors in Figure 2.4),
and the temporal variation in archaeal taxa was observed as the pH increased on day 8 and 16 most likely as a consequence of photosynthesis. Not only does CCA provide insight into the environmental factors influencing the community structure, but it also suggests potential interactions occurring between taxa. The appearance and subsequent decline of plotted taxa was the most influential variable on discrepancies between the time points. For instance, day 8 is more similar to day 16 than other time points due to the loss of three archaeal taxa from day 7 to 8. As discussed, the low levels of salinity were an influential factor associated to the rapid decline of halophilic taxa.

Figure 2.4 also shows that \textit{Methanococcus maripaludis} was most associated with the variation in CCA1 as it was the only detectable archaeal taxa remaining at the last time point on day 16. \textit{M. maripaludis} was detected on days 1, 3, 8, and 16, but not detected on days 7 and 11. While known Methanococci are strict anaerobes, there could be micro-anaerobic niches in the raceway pond related to biomass turnover. It is also possible that the sequences are detected at later time points due to PCR biases. Recent research has shown \textit{M. maripaludis} can survive in anaerobic biofilms (Brileya et al., 2013); and therefore, it is possible that a small population was able to survive within a biofilm matrix on the walls of the raceway or paddle wheel. Its sporadic appearance may also be the result of collection methods that could have disturbed the biofilm or biofilm detachment. CCA2 was most influenced by nitrate and phosphate, which explained 38.4% of the variation in archaeal taxa.
Figure 2.4: Canonical Correspondence Analysis (CCA) plotting chemical variables that correlated with variation in Archaeal taxa (with >2% relative abundance) and sample points. Archaea generally lacked long term survivability in the pond. CCA1=81.8% CCA2=57.3%

We also applied the same CCA metrics for the bacterial taxa to visualize variance in taxa and time points as correlating with changes in chemical variables (Figure 2.5). CCA1 was predominately influenced by increased nitrate concentrations accounting for about 57.2% of the variance in bacterial taxa distribution. The first two time points were
correlated with the initial halophilic bacterial taxa that were unable to maintain a population due to low salinity, aerobic conditions, and/or increasing pH (Baxter et al., 2005; Oren, 1994; 2006; Post, 1977; Tazi et al., 2014). The decrease in phosphate contributed approximately 30.5% of the observed variation in CCA2 and correlated with the increase in some of the most abundant taxa at day 16, (for example *Flavobacteriales* \(R^2=0.3\)) and *Proteobacteria* \(R^2=0.35\)).

**Correlations between Community Members**

The distribution of community member occurrence is shown in Figure 2.6, illustrating the persistent *C. vulgaris* population. The upper dendogram clusters taxa by percent relative abundance and frequency of co-occurrence. As observed via CCA, the time points are grouped by declining archaeal sequences with the exception of *Methanococcus* at day 8 and 16. Post-transfer to 2,000 L, several bacterial OTUs clustered with day 11 and 16. *Flavobacterium* and *Erythromicrobium* are common groundwater/tap water organisms that were likely introduced during volume scale-up, but it is not known if the co-occurrence is direct or indirect. In addition, sequences indicative of *Loktanella* and *Roseicyclus* correlated with Chlorella on day 8, 11, and 16 during cultivation scale-up. Sequences indicative of *Algoriphagus* correlated with Chlorella during the 200 L cultivation but declined during the 2,000 L cultivation.
Figure 2.5: Canonical Correspondence Analysis (CCA) plotting chemical variables that correlated with variation in Bacterial taxa (with >2% relative abundance) and sample points. Bacteria generally trended with an increase nitrate toward end of the experiment, possibly due to an increase in nitrifying taxa. *Pseudomonas sp.* was associated with the early time points (day 1 and 3). CCA1 = 60% and CCA2=8%
Figure 2.6: Heat map showing taxa with greater than 2% relative abundance at each sampling time point. Taxa co-occurrence and each time points were correlated and clustered according to relatedness as shown in the dendogram.

However, co-occurrence does not infer a statistical correlation between taxa. In order to investigate possible correlations we used SparCC to construct community correlation networks. Figure 7 shows a community network map of correlations between
community members (excluding archaea due to their general absence after day 3) (Friedman and Alm, 2012). The most salient of these relationships is the positive 0.85 correlation between *C. vulgaris* and *Pseudomonas spp.* (p<0.05). No other bacteria correlated with *C. vulgaris*, which was the predominant eukaryote. Previous studies have observed different species of *Pseudomonas* living in association with algae including *C. vulgaris* (Sapp et al., 2007). A symbiotic relationship between these organisms was described by Guo and Tong finding that *Pseudomonas sp.* fostered the growth of *C. vulgaris* (Guo and Tong, 2013). When in co-culture with *Pseudomonas sp.*, the cell concentration of *C. vulgaris* was 1.4 times greater than that of axenic cultures under the same conditions. Scanning electron microscope (SEM) images revealed that the bacteria were living in the exopolymeric substances (EPS) or “phycosphere” of *C. vulgaris* (Guo and Tong, 2013).

Figure 2.7: SparCC network map showing significant (p<0.05) of 122 Interactions with a 0.85 correlation between different OTUs incorporating all time points. Green lines are indicative of positive interactions while red lines are negative.
The phycosphere, coined by Bell and Mitchell in 1972, is often colonized by bacteria (Bell and Mitchell, 1972; Goecke et al., 2013; Sapp et al., 2007). This specific niche facilitates a tight exchange of oxygen, carbon, and metabolites minimizing dilution (Bruckner et al., 2011; Gärdes et al., 2012; Martin et al., 2014; Paul et al., 2012; Sapp et al., 2008). Bacteria can provide the alga with sources of growth promoters (e.g., indole-3-acetic acid), and essential vitamins (e.g., cobalamin), while discouraging colonization by other potentially harmful microorganisms with antimicrobial metabolites (Croft et al., 2005; 2006; Gonzalez and Bashan, 2000). In return, bacteria have immediate access to algal exudates that can be a key source of fixed carbon (Bell and Mitchell, 1972; Goecke et al., 2013; Sapp et al., 2007). The correlation of *Pseudomonas* populations with *C. vulgaris* throughout the course of the pond experiment, even following the transfer to the larger raceway, may have contributed to the predominance of the algal culture under open conditions (Figure 2.7). The results suggest that symbiotic-associations could have relevant industrial applications that could result in increased biomass yields (Imase et al., 2008; Natrah et al., 2013). Further work is needed to discern the mechanism(s) of interactions that impact algal biomass and/or lipid accumulation in addition to confirmation of a direct and/or indirect relationship between these two organisms under the tested growth conditions.

**Conclusions**

Our work demonstrated that the cultivation of a single algal strain in an open pond without the addition of antibiotics or herbicides can be successful. The use of high pH
systems and alkaline adapted algal taxa could be a successful strategy for overcoming some of the constraints associated with large-scale biomass production in open systems. Furthermore, certain phycosphere associations could enhance biomass yields and deter colonization by detrimental populations. Further work is needed to determine the longevity and stability of open, outdoor cultivation systems for the production of algal biomass and/or biomolecules.

**Funding**

This work was financially supported by the NSF IGERT Program in Geobiological Systems (DGE 0654336) in addition to the US Department of Energy-Advancements in Sustainable Algal Production (ASAP) program under contract DE-EE0005993, the National Science Foundation Sustainable Energy Pathways (SEP) Program under NSF CHE-1230632, and the US Department of Energy Bioenergy Technologies Office grant DE-FG36-08GO18161.

**Acknowledgements**

The authors would like to thank the members of the Center for Biofilm Engineering, especially Ann Willis, for their support. We would also like to acknowledge Dr. Kara Bowen DeLeón for her guidance and all members of the MSU Algal Biofuels Group for their helpful discussions.
CHAPTER 3

BACTERIA AND EUKARYA COMMUNITY DURING EUTROPHICATION AND TOXIC CYANOBACTERIAL BLOOMS IN THE ALKALINE LAKE VELENCE, HUNGARY

Contribution of Authors and Co-Author:

Author: Tisza Ann Szeremy Bell
Contributions: Collected field data, conducted most experiments, analyzed data and wrote manuscript

Co-Author: Emel Şen
Contributions: Aided in DNA extraction and data analysis

Co-author: Tamás Felföldi
Contributions: Contributed to the experimental design and sample collection

Co-author: Matthew W. Fields
Contributions: Principal Investigator

Corresponding Author: Brent Peyton
Contributions: Principal Investigator
Manuscript Information Page

Author and Co-authors: Tisza Ann Szeremy Bell, Emel Sen-Kilic, Tamas Felföldi, Gábor Vasas, Matthew W. Fields, and Brent M. Peyton

Journal:

Status of Manuscript:

X Officially submitted to a peer-review journal

Prepared for submission to a peer-reviewed journal

Accepted by a peer-reviewed journal

Published in a peer-reviewed journal
Title: Microbial community changes during a toxic cyanobacterial bloom in an alkaline Hungarian lake

Running Page Head: Community dynamics during toxic cyanobacterial bloom

Authors: Bell, T.A.S.1,2, Sen-Kilic, E.2,5, Felföldi, T.3, Vasas, G.6, Fields, M.W.1,2,4, and Peyton, B. M. *2,4,5

Affiliations: 1Department of Microbiology & Immunology, Montana State University, Bozeman, MT USA; 2Center for Biofilm Engineering, Montana State University, Bozeman, MT USA; 3Department of Microbiology, Eötvös Loránd University, Budapest, Hungary; 4Energy Research Institute, Montana State University; 5Department of Chemical and Biological Engineering, Montana State University, Bozeman, MT USA; 6Department of Botany, University of Debrecen, Debrecen, Hungary

Keywords: lake ecology, alkaline lake, phytoplankton, algae, eutrophication, cyanobacteria bloom, microcystin, Microcystis

Corresponding author: B.M. Peyton Department of Chemical and Biological Engineering and Center for Biofilm Engineering, Montana State University, Bozeman, MT 59717, USA
Abstract

The Carpathian Basin is a lowland plain located mainly in Hungary. Due to impermeable alluvial deposits and a bowl shape, many lakes and ponds of the area are characterized by high alkalinity. In this study, we characterized temporal changes in eukaryal and bacterial community dynamics with high throughput sequencing and relate the changes to environmental conditions in Lake Velence located in Fejér county, Hungary. The sampled Lake Velence microbial populations (algal and bacterial) were analyzed to identify potential correlations with other community members and environmental parameters at six time points over approximately 6 weeks in the Spring of 2012. Correlations between community members suggest a positive relationship between certain algal and bacterial populations (e.g., Chlamydomonadaeae with Actinobacteria and Acidobacteria), while other correlations allude to changes in these relationships over time. During the study, high nitrogen availability may have favored non-nitrogen fixing cyanobacteria (e.g., Microcystis), and the eutrophic effect may have been exacerbated by the high calcium and magnesium content of the Carpathian Basin bedrock, which may have fostered exopolymer production and cell aggregation. Cyanobacterial bloom formation could have a negative environmental impact on other community members and potentially affect overall water quality as well as recreational activities. To our knowledge, this is the first prediction for relationships between photoautotrophic eukaryotes and bacteria from an alkaline, Hungarian lake.
Phytoplankton are the primary producers in lake ecosystems forming the foundation of aquatic food chains (Reynolds, 1984). Within these webs exist dynamic relationships between community members, some of which are dualistic in nature, and fluctuate between positive and negative interactions across domains (Amin et al., 2012). Diversity in community composition of the phytoplankton and bacteria provide insight into the biogeochemical cycles occurring within the lake as well as the surrounding watershed due to allochthonous inputs (Wetzel, 2001). Further, overall species richness can be an indicator of overall community function and stability (Ptacnik et al., 2008). Fluctuations in community members in response to changing environmental conditions (e.g., temperature, eutrophication) can reduce biodiversity and overall community function and underlie larger global trends such as climate change and other environmental perturbations (Dodds et al., 2012; Strayer and Dudgeon, 2010).

The shallow lakes of the Carpathian Basin in Hungary are some of the most alkaline in the world (Boros et al., 2014). The basin is surrounded by slowly eroding mountains composed of limestone and carboniferous granite with high magnesium and calcium carbonate content deposited into the lakes and ponds that have formed on top of impermeable alluvial deposits from the Danube and Tisza Rivers (Gyalog and Horváth, 2004). Despite the unique conditions of high pH and salinity, lakes and ponds of the region contain productive, phototrophic communities (Ács et al., 1994; 2003; Felföldi et al., 2009; Pálffy et al., 2014; Vörös et al., 2006). In fact, some of the most productive,
photoautotrophic systems are alkaline (e.g., soda lakes) due to the high availability of carbon (Grant, 2006).

Lake Velence, located in Fejér county, Hungary, is one of the largest lakes in the region (24.17 km²) (Boromisza et al., 2014). Submerged macrophytes differentiate the western portion of the lake from the open water of the eastern portion with close proximity to Budapest (50 km southwest), making the lake a popular holiday destination and recreational resource (Boromisza et al., 2014). Lake Velence is a terminal lake supplied by two streams, surrounding runoff, and adjacent basement springs (Gyalog and Horváth, 2004). These inputs supply the lake with a high concentration of calcium and magnesium cations (Boromisza et al., 2014; Gyalog and Horváth, 2004). Therefore, the limnogeology of Lake Velence is unique due to concentrated calcium carbonate input with limited output further facilitating high pH (8.5-9.9), sodium (~300 mg/L), bicarbonate (~680 mg/L), magnesium (~420 mg/L) and sulfate (~670 mg/L) ions (Borsodi et al., 2005; Gyalog and Horváth, 2004; Haas, 2012).

Due to agricultural runoff, the lake has become eutrophic over the last century (Water Management Directorate of VITUKI Environmental Protection and Water Management Research Institute on behalf of the Ministry for Environment and Water, 2006), and eutrophic systems like Lake Velence are especially susceptible to large blooms of cyanobacteria. Certain non-nitrogen fixing cyanobacteria, like *Microcystis aeruginosa* isolated from Lake Velence, capitalize on the high availability of nitrogen in the system forming large visible blooms capable of producing different forms of the toxin microcystin (MC), of which MC-LR is the most toxic to lake organisms, wildlife, livestock, and humans.
Blooms may be exacerbated by high availability of soluble calcium and magnesium, enhancing the activity of lectins, a key constituent of extracellular polymeric substances (EPS) important in bloom formation (Zhao et al., 2011). Large cyanobacterial blooms threaten biodiversity by limiting light, oxygen, and nutrient availability, and although cyanobacteria release oxygen as a byproduct of photosynthesis, oxygen is consumed when these large blooms begin to decompose (Paerl and Otten, 2013). Climate change induced warming and increased eutrophication may additively increase toxic bloom frequency and extent (Davis et al., 2009; De Senerpont Domis et al., 2012; Paerl and Otten, 2013), and thus decrease microbial community stability. By investigating the microbial diversity and environmental conditions over the summer period, temporal community dynamics were elucidated and potential inter-domain population networks were predicted to provide insight into the microbial community structure of a unique, alkaline system responding to anthropogenic disturbance.

**Materials and Methods**

**Site Description**

Lake Velence is located in Fejér county, Hungary, approximately 50 km southwest of the capital, Budapest. Samples were collected off a dock located in the northwestern corner of the lake (Figure 3.1). The northwest area of the lake is one of the most developed providing easy access for recreation and represents a highly impacted area.
Sample Collection

To analyze biotic and abiotic microbial community interactions in Lake Velence, limnological chemistry and biological samples were collected at 6 time points (1 - 5/18, 2 - 5/25, 3 - 6/1, 4 - 6/8, 5 - 6/13, and 6 - 6/21) approximately one week apart (every five to eight days) during the spring of 2012 (May 18th – June 21st). Surface water samples were collected mid-day from the same location each time into sterile bottles for further chemical
and biologic analysis. Samples for community analyses (1 L) were taken with sterilized bottles, kept at 4°C for transport back to the laboratory, and frozen at -80°C for later filtration and DNA extraction as described previously (Bell et al., 2016). Biomass was also collected on 6/8 from a large bloom that persisted for approximately 10 days. Samples were kept frozen until being lyophilized for mass spectrometer analysis described below.

MALDI-TOF MS Analysis

Lyophilized biomass was screened for microcystins using matrix assisted desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) following methods described previously (Farkas et al., 2014). Lyophilized sample (5 mg) was mixed with 200μl of 50% aqueous methanol, sonicated for 5 minutes, and incubated for 60 minutes. The samples were examined in positive-ion mode using a Bruker Biflex MALDI-TOF mass spectrometer (Bruker Daltonics Inc., Billerica, MA, USA) equipped with delayed-ion extraction. A 337nm nitrogen laser was used for desorption/ionization of the sample molecules. Spectra from multiple (at least 100) laser shots were summarized using 19-kV accelerating and 20-kV reflectron voltage. External calibration was applied using the [M- Na+] 1 peaks of malto-oligosaccharides dp 3–7, m/z values 527.15, 689.21, 851.26, 1013.31, and 1175.36, respectively. The measurement was performed in 2,5-dihydroxybenzoic acid (DHB) matrix, by mixing 0.5 ml of matrix solution with 0.5 ml of sample on the sample target and allowing it to dry at room temperature. DHB matrix solution was prepared by dissolving DHB (10 mg) in a mixture (0.5 ml) of ethanol and water (1:1, v:v). The compounds were identified on the basis of the mass of [M-1 H]⁺ peak. After determination of mass values, post-source decay (PSD) measurements were
performed directly from the same sample on the template and microcystins and other peptides identified by PSD fragment structure analysis.

Chemical Analyses

For chemical analyses, the guidelines given in Standard Methods for the Examination of Water and Wastewater were followed (Eaton and Franson, 2005). The pH was measured by electrochemical method (ASTM 4500-H-B) with a Radelkis OP-264 pH meter (Budapest, Hungary). Specific electric conductivity was determined according to ASTM-2130-B. Aliquots from water samples were filtered through a 0.45µm membrane filter prior to the determination of other parameters. Concentration of dissolved organic carbon (DOC), dissolved inorganic carbon (DIC) and total nitrogen (TN) was measured with a Multi N/C 2100S analyzer (Analytik Jena, Jena, Germany). Ammonium (ASTM 4500-NH3-D), nitrite (ASTM 4500-NO2-B), nitrate (ASTM 4500-NO3-B) and sulfate (ASTM 4500-SO4 2-E) ion concentrations were measured by spectrometry. For ammonium ion quantification, 100 ml of each sample was acidified with 1 mL of 1M sulfuric acid for conservation, and samples were neutralized just before the analysis. Reactive phosphate and total phosphorus (TP) were both examined with the ascorbic acid method. Before the TP measurement, persulfate oxidation was carried out (ASTM 4500-P-E). Sulfide was measured by an iodometric method (ASTM 4500-S-E).

Sodium and potassium ion content was determined with a PFP-7 flame photometer (BUCK Scientific, East Norwalk, USA). Prior to analysis, samples were acidified with nitric acid to a 0.15% final concentration. Hardness was determined with ethylenediamine-tetraacetic acid (EDTA) titrimetric method: calcium concentration was calculated from the titration
result using murexide indicator, and magnesium content determined as the difference of
titrations carried out with murexide and eriochrome black T with appropriate pH buffers.

DNA Extraction and Sequencing

**DNA Extraction.** Samples were slowly thawed at 4°C and microbial biomass was
collected via filtration of 1L through 0.22μm polyethersulfone membrane filters and stored
at -80°C until transport to the USA (Corning Inc., Corning, NY, USA). Filters were
transported from Hungary to the USA on dry ice and immediately stored at -80°C upon
arrival until DNA extraction. For extraction, solids collected on the filters were suspended
in the MO BIO PowerMax™ Soil DNA Isolation Kit PowerBead Solution, and the cells
were lysed via three cycles of liquid nitrogen freeze–thaw and ground with a mortar and
pestle aided by sterile sand (Zhou et al., 1996) (MO BIO Laboratories Inc., Carlsbad, CA,
USA). The DNA was cleaned and concentrated with the Wizard® SV Gel and PCR Clean-
Up System (Promega Corporation, Madison, WI, USA) according to the manufacturer's
protocol.

**Bar-Coded Pyrosequencing.** Pyrosequencing was used to characterize the microbial
population of the lake samples via PCR amplification of the V1-V3 region of the SSU
rRNA gene sequences. Each sample was labeled with a unique 10 nucleotide-barcode for
multiplexing. The SSU rRNA gene sequences for Eukarya and Bacteria were amplified
via 25 cycles of PCR with the following barcoded primers; 7F (5’-ACCTGGTTGATCTGCCAG-3’) and 591R (5’-GGAGCTGGGAATTACCG-3’) for
Eukarya and FD1 (5’AGAGTTTGATCTGGCTGAGC-3’) and 529R (5’-
CGCGGCTGCTGGCAC-3’), which targeted the V1-V3 region of Bacteria (Bowen De León et al., 2012). Archaeal sequences were amplified separately from Bacteria using a nested approach with non-barcoded 21F (5’-TTCYGGTTGATCCYGCRGA-3’) and 1492R (5’-CGGTT ACCTTGTTACGACTT-3’) for 20 cycles followed by an additional 20 cycles with barcoded 751F (5’-CCGACGGTGAGRGRYGAA-3’) and 1204R (5’-TTMGGGCGATRCNKACCT-3’) primers (Baker et al., 2003; Barnhart et al., 2013). Viral dsDNA was amplified by targeting the major capsid protein (MCP) via touchdown PCR using a pair of degenerate primers mcpF (5’-GGYGGYCARCGYATTGA-3’) and mcpR (5’-TGIARYTGYTCRAYIAGGTA-3’) as described previously reported (Larsen et al., 2008). PCR products of the correct size were confirmed using a 1% agarose gel for all amplimers. Eukaryotic and bacterial products were cut from the gel and pooled using an Ultrafree®-DNA gel extraction column (Millipore Corporation, Bedford, MA, USA). The gel extract was cleaned and concentrated using the Wizard® SV Gel and PCR Clean-Up System, and dsDNA was quantified with a Qubit fluorometer (Invitrogen, Carlsbad, CA, USA). Adaptors for 454 sequencing were ligated to the eukaryotic and bacterial amplicons and were sequenced on a 454 GS-Junior (454 Life Sciences, Branford, CT, USA). Roche's image analysis identified sequences by barcode. Sequences were trimmed to one standard deviation below the mean length or removed if shorter. Employing the Phred score filter, 15% of the nucleotides were allowed to be below Q27, and removed if primer errors or Ns were observed (Bowen De León et al., 2012).
Processing of Pyrosequencing Data and Community Analysis

Data analysis was performed using the Quantitative Insights into Microbial Ecology (QIIME) software package, version 1.4.0 (Caporaso et al., 2010b). Parameter settings for demultiplexing were at a default level of 200 bp and 1000 bp in length. Metadata files were prepared according to a QIIME compatible template taking into account environmental data on pH, temperature, and ion concentrations. The sequencing libraries were split according to barcode for each of the respective domains (Archaea, Bacteria, and Eukarya). Sequences were then concatenated for data normalization needed in downstream analysis. Operational taxonomic units (OTUs) were assigned using the closed reference OTU picking protocol. Clusters were referenced against the Silva 108 database and pre-clustered at 97% identity using UCLUST (Edgar, 2010).

Sequence reads that matched a Silva reference sequence at 97% identity were clustered within an OTU defined by a reference sequence. OTU assignment (and all subsequent steps) was performed for the combined Archaea, Bacteria and Eukarya reads. The singleton OTUs were discarded. The centroid sequence in every cluster was selected to represent the cluster and aligned with the Silva core set using PyNAST (Caporaso et al., 2010a). Chimeric sequences, identified with Chimera Slayer (Haas et al., 2011) and reads that failed to align with PyNAST were excluded from subsequent analyses. PyNAST (v1.1) was used for sequence alignment and filtering through QIIME using default parameters.

Taxonomic assignments were additionally made using the retrained RDP Classifier (Wang et al., 2007) on the Silva 108 database for phylogenetic resolution at the genus level. Taxonomic summary for Archaea, Bacteria, and Eukarya was plotted to the genus level.
with a given relative abundance based on diversity and distribution pattern per domain. Taxa distribution was also summarized by time (sample day). The biodiversity analysis downstream between samples was derived using UniFrac that took into account the phylogenetic structure of the microbial communities (Lozupone and Knight, 2005). Diversity was quantified by calculating the effective number of OTUs for *Bacteria* and *Eukarya*. The effective number of OTUs, or Hill Number, was calculated using the R package “rioja” and used to track fluctuations in diversity (Hill, 1973; Jost, 2006; Juggins, 2012).

The co-occurrence of community members was illustrated in a heat-map using the R vegan package version 2.0-10 (Oksanen, 2011). Some taxa had a 0% relative abundance at certain time points; therefore, 0.1 was added to all values to allow a log transformation. To enhance the visual distribution of taxa, log transformed values were cubed and resulting values plotted \((\log(\text{relative abundance} + 0.1)^3)\). A SparCC analysis was subsequently used to construct community correlation networks by estimating linear correlation values between log transformed abundances based on the absolute number of sequences for an OTU rather than a relative abundance (Berry and Widder, 2014; Friedman and Alm, 2012). The key advantage of this analysis was that the ratio of the fractions of two OTUs was independent of the fluctuations in other OTUs included in the analysis (i.e., subcompositional coherence) (Friedman and Alm, 2012).

**Statistical Analysis**

Regression analysis and p values were calculated using Matlab Statistics Toolbox release 2012a (The MathWorks, Inc., Natick, Massachusetts, United States). Only
regressions with $p \leq 0.01$ were considered significant. A principal coordinate analysis (PCA) was used to reduce dimensionality and give structure to the water chemistry variables obtained from each time point (Legendre and Gallagher, 2001). To incorporate taxonomic data, the direct-gradient ordination technique, Canonical Correspondence Analysis (CCA) was used, which concurrently showed lake taxa, time points, and water chemistry (Hall and Smol, 1992). The CCA ordination is appropriate when assessing community dynamics because it does not use Euclidean based metrics that assume linear trends in community change. Not only does the ordination show environmental factors correlating to community changes, but the results suggest potential interactions between taxa (Amaral-Zettler et al., 2010). The first two axes, CCA1 and CCA2, typically account for the majority of observed variation. All axes are constrained to present a linear combination of the water chemistry that maximizes the dispersion of taxa (Hall and Smol, 1992).

**Results and Discussion**

Fluctuations in Lake Velence algal and bacterial populations over the six-week period showed correlations between community members and environmental parameters suggestive of intricate relationships. In addition, some of the observed changes occurred over time. Archaeal sequences were not detected as a significant portion of the total reads, and may be a result of the universal primers used. Therefore, the analyses focus on *Eukaryal* and *Bacterial* sequences.
Chemical Analyses

Chemical parameters were sampled once per week over the course of 6 weeks from May 2012 to June 2012 at the same location of the lake (47° 14’ 3.481” N, 18° 37’ 51.456” E). Over the tested time points, conductivity, nitrate, nitrite, DOC, DIC, and TP remained relatively constant (Table 3.1). The sulfate concentrations increased in the last two weeks, and the ammonium-N levels fluctuated between 0.01 and 0.3 mg/l (Table 3.1). The phosphate (mg/l) increased in weeks 2 and 3. The temperature increased from 14° to 20°C in week 2, fluctuated between 20° and 23°C up to week 5, and increased to almost 30°C in the 6th week (Table 3.1). The overall increase in water temperature coincided with an increase in air temperature.

Table 3.1: Summary of water chemistry analysis at six time points taken approximately every week for 6 weeks. (DOC – dissolved organic carbon; DIC – dissolved inorganic carbon; TP – total phosphorus; TDS – total dissolved solids)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>8.76</td>
<td>8.29</td>
<td>9.60</td>
<td>9.14</td>
<td>8.36</td>
<td>8.26</td>
</tr>
<tr>
<td>Conductivity (mS/cm)</td>
<td>2.93</td>
<td>2.71</td>
<td>2.60</td>
<td>2.36</td>
<td>2.76</td>
<td>2.79</td>
</tr>
<tr>
<td>Sulfate (mg/L)</td>
<td>630</td>
<td>679</td>
<td>639</td>
<td>647</td>
<td>721</td>
<td>705</td>
</tr>
<tr>
<td>Ammonium-N (mg/L)</td>
<td>0.30</td>
<td>&lt;0.01</td>
<td>0.06</td>
<td>0.09</td>
<td>0.10</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Nitrate-N (mg/L)</td>
<td>1.8</td>
<td>2.4</td>
<td>2.1</td>
<td>1.9</td>
<td>1.6</td>
<td>1.8</td>
</tr>
<tr>
<td>Nitrite-N (mg/L)</td>
<td>&lt;0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>DOC (mg/L)</td>
<td>27.7</td>
<td>37.9</td>
<td>35.2</td>
<td>28.6</td>
<td>43.1</td>
<td>35.2</td>
</tr>
<tr>
<td>DIC (mg/L)</td>
<td>179</td>
<td>175</td>
<td>116</td>
<td>137</td>
<td>164</td>
<td>166</td>
</tr>
</tbody>
</table>
Table 3.1 Continued

<table>
<thead>
<tr>
<th>Phosphate (mg/L)</th>
<th>&lt;0.01</th>
<th>0.11</th>
<th>0.15</th>
<th>0.02</th>
<th>0.02</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP (mg/L)</td>
<td>&lt;0.01</td>
<td>1.15</td>
<td>1.74</td>
<td>1.25</td>
<td>1.35</td>
<td>1.03</td>
</tr>
<tr>
<td>TDS (g/L)</td>
<td>1.63</td>
<td>1.62</td>
<td>1.52</td>
<td>1.64</td>
<td>1.63</td>
<td>1.64</td>
</tr>
<tr>
<td>Secchi Depth (cm)</td>
<td>40.6</td>
<td>43.0</td>
<td>48.3</td>
<td>58.4</td>
<td>48.3</td>
<td>47.0</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>14.1</td>
<td>20.0</td>
<td>21.1</td>
<td>23.1</td>
<td>21.9</td>
<td>29.7</td>
</tr>
</tbody>
</table>

During the sampling period, lake water temperature increased, which likely promoted the increase of general microbial activity. Furthermore, weather events may have significant effects on the measured values of chemical variables; for instance, release of nutrients from the sediment due to an intensive mixing of the whole water column by wind (e.g., increases in sulfate, nitrate, DOC, and phosphate after the first week) or cooling the lake water due to a heavy rainfall or cold front (as after the fourth week).

Biotic activity likely impacted some fluctuations in nutrient concentrations. An increase in pH during week 3 and 4 may suggest higher photoautotrophic primary productivity appearing to peak on the third time point, and similar trends have been previously observed (Wetzel, 2001). This was further supported by decreases in the concentration of dissolved inorganic carbon (DIC) (Hem, 1985). Subsequent increases in DIC at weeks 5 and 6 that coincided with declines in pH were suggestive of an overall decrease in photoautotrophic activity during weeks 5 and 6.
Variations in additional chemical variables were likely influenced by metabolic activities associated with dominant taxa at respective sampling time points. Phosphate concentrations in sampling time points 4, 5, and 6 were lower than that of 2 and 3, and these results suggested an interplay between biological storage and different rates of biotic utilization/replenishment (Hudson et al., 2000). Concentrations of calcium, magnesium, sodium, and potassium were measured at the first time point (30mg/L, 420mg/L, 297mg/L and 46 mg/L, respectively) and only slight changes were observed during the subsequent samplings. An ordination of these chemical variables was used to visualize potential relationships and variation between the samples (Figure 3.2). A principle component analysis (PCA) indicated that high ammonium, low temperature, DIC, DOC, and increases in phosphate and TP most influenced the discrepancies between the chemical profile of sample time points. The increase in ammonium accounted for the spatial distance in the PCA between the first and second time points. An increase in sulfate accounted for most of the variation between sample time points 4, 5 and 6.
Due to generally high nutrient concentrations in Lake Velence, especially nitrate, the lake can be characterized as eutrophic-hypertrophic suggesting that phytoplankton were not limited by nutrients during a majority of the sampling period (Wetzel, 2001). High nitrogen and phosphorus concentrations likely originated from an allochthonous source (i.e., the nutrient load provided by surrounding agriculture in addition to aquatic bird feces),
as demonstrated in the case of other shallow alkaline lakes of the Carpathian Basin (Baldi and Kisbenedek, 2000; Boros et al., 2008).

Sequencing and Community Analysis

The SSU rRNA gene sequences for Bacteria and Eukarya were amplified and sequenced for each of the samples (5 samples in the case of Eukarya). After screening of the sequence raw data (see Methods), 118,593 quality gene sequences with a valid barcode were retrieved. Sequences with a 97% identity were clustered within an OTU totaling 7,533 observed OTUs in Bacteria and Eukarya combined (Table 3.2 and Figure 3.3).

Table 3.2: 454 Pyrosequencing statistics for SSU rRNA gene sequences

<table>
<thead>
<tr>
<th></th>
<th>16S</th>
<th>18S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of samples</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Number of Unique OTUs</td>
<td>6,817</td>
<td>716</td>
</tr>
<tr>
<td>Number of total sequences</td>
<td>60,367</td>
<td>58,626</td>
</tr>
<tr>
<td>Minimum sequences/sample</td>
<td>4,487</td>
<td>812</td>
</tr>
<tr>
<td>Maximum sequences/sample</td>
<td>22,159</td>
<td>25,897</td>
</tr>
<tr>
<td>Average number of sequences/sample</td>
<td>10,061</td>
<td>11,725</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>5,807</td>
<td>10,569</td>
</tr>
</tbody>
</table>
Figure 3.3: Relative abundance of A) Eukaryotic taxa by phylum and B) Bacterial taxa by phylum with at least a 2% relative abundance over the 6 week period. Note that no Eukaryotic sequences were amplified for week 4

Amplification of eukaryotic DNA resulted in low concentration and subsequently poor quality 18S reads for the fourth time point and were discarded (data for this 18S time point were removed from the statistical analysis). By conducting tandem 16S and 18S pyrosequencing, the first to be conducted on a Hungarian lake, we were able to identify major
taxa and population distributions during the spring of 2012 at six time points with each approximately one week apart in an anthropogenically impacted lake.

Diversity was quantified by calculating Hill numbers, the effective number of taxa, for *Bacteria* and *Eukarya* for each time point, and fluctuations in bacterial and eukaryotic diversity did not appear to be related to one another ($R^2=0.13$). Bacterial diversity was generally higher than eukaryotic diversity with the exception of the 5th time point (Table 3.3).

Table 3.3: Hill numbers showing changes in the diversity of *Bacteria* and *Eukarya* over the sampling period

<table>
<thead>
<tr>
<th>Time Point</th>
<th><em>Bacteria</em> Hill Number</th>
<th><em>Eukarya</em> Hill Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.0</td>
<td>3.20</td>
</tr>
<tr>
<td>2</td>
<td>11.0</td>
<td>7.20</td>
</tr>
<tr>
<td>3</td>
<td>9.10</td>
<td>2.40</td>
</tr>
<tr>
<td>4</td>
<td>12.2</td>
<td>NA</td>
</tr>
<tr>
<td>5</td>
<td>4.90</td>
<td>6.80</td>
</tr>
<tr>
<td>6</td>
<td>8.80</td>
<td>4.15</td>
</tr>
</tbody>
</table>

**Eukaryotic Diversity**

Eukaryotic diversity fluctuated over the course of the sampling period peaking during the 5th time point (Table 3.3, Figure 3.5). Diversity had a small negative correlation with the relative abundance of the cyanobacteria *Microcystis* ($R^2=0.17$) discussed in greater detail in below. *Arthropoda* abundance, mainly composed of two *Crustacea*, was highest
during the first three weeks and suggestive of low suspended solids benefitting filter feeders (Vörös et al., 2006) (Figure 3.5). In contrast, Cryptophyta, Pavlova, and Ciliophora generally had increased abundances in the final 2 sampling time points (sampling time points 5 and 6) when Microcystis abundance was the highest. Cryptophyta was identified via light microscopy in prior surveys of Lake Velence and demonstrated strong correlations with sulfate and DOC ($R^2=0.84$, $R^2=0.53$), confirming previous observations of sulfate tolerance in Cryptophyta while capitalizing on available carbon for incorporation into biomass (Ács et al., 1994; 2003; Camacho et al., 2001). Pavlova appeared to respond to carbon availability, positively correlating with DOC ($R^2=0.56$). Previous research associated ciliate communities to the presence of reeds, ubiquitous on the western portion of the lake (Boromisza et al., 2014; Mieczan, 2008).

The community was also predominated by sequences closely related to unidentified Chlorophytes, an undescribed Chlorella taxa, and a Thalassiosira like diatom, which was also observed via microscopy. The first sampling time point had low algal relative abundance except for Pseudoschroedena sp., a small Chlorophyte. Overall, algal relative abundance (e.g., Chlorophyta) within the eukaryotes increased with time, particularly the last two weeks. After the third week, sequences indicative of Chlorophyta were the most predominant for detected 18S sequences. During this time, pH remained between 9.2 and 8.2 with relative increases in both DOC and DIC that could promote the growth of green algae (but bacterial phototrophs as well).
Figure 3.4: Heat map showing the Pearson correlation between Eukaryotic taxa (listed as genera when possible) with >2% relative abundance at sampling time points. Sampling time points were also correlated and clustered according to relatedness as shown in the dendogram on top. Note that no Eukaryotic sequences were amplified for week 4 (W4).

Differences in temperature, pH, DIC, and DOC were the major drivers of observed variation between taxa and the chemical composition of sampling time points (Figure 3.5).
Eukaryotic diversity was positively correlated with DIC and DOC ($R^2 = 0.3$ and $R^2 = 0.5$, respectively), particularly sequences indicative of green algae and diatoms.

Figure 3.5: Canonical Correspondence Analysis (CCA) plotting chemical variables that drive variation in Eukaryotic genera (with >2% relative abundance) and sample points. Algae generally responded to increases in temperature and DOC ($CCA1 = 62\%$, $CCA2 = 36\%$).
However, a different green algae sequence group (*Chlorophyceae 2*) and *Botryococcus* sequence group appeared to not respond as strongly to DIC, DOC, and temperature. Lying almost parallel to axis 1, changes in pH accounted for the most variation between taxa and time points and negatively correlated with eukaryotic diversity ($R^2=0.56$). The flagellate *Euglenida* relative abundance increased with decreases in pH suggesting the organism may have a lower pH optimum. Increases in *Chlorophyta* correlated best with increases in temperature, DIC, and DOC, and the combination of increased carbon availability and temperature likely contributed to increased occurrence of certain OTUs and an overall decrease in diversity. Previous studies have shown that increased perturbation (including nutrients) can decrease microbial diversity (and/or evenness) (Fields et al., 2005; 2006).

In addition, we amplified the region for the major capsid protein (MCP) that is unique to double-stranded DNA (dsDNA) viruses that infect algae (Larsen et al., 2008). Many previous studies emphasized the pivotal role these viruses play in aquatic systems (Brussaard, 2004; Hill et al., 1998; Martínez et al., 2007; Rhodes and Martin, 2010; Short, 2012; Van Etten et al., 1991; Weitz and Wilhelm, 2012; Wilhelm and Suttle, 1999; Wommack and Colwell, 2000). Due to the very limited availability of these viral genomes, we were only able to qualitatively confirm presence of the viruses via gel electrophoresis of amplified products from time points 1, 4, 5, and 6 (Figure 3.6). The results indicate the presence of dsDNA algal viruses in Lake Velence, and further work is needed to understand the potential role in algal population dynamics as well as nutrient cycling.
Figure 3.6. 1% TAE gel of viral amplicons for MCP. Time points 1, 4, 5, and 6 had product while 2 and 3 did not show amplification. The MCP region is between ~300-500 base pairs in size. N is representative of the negative control.

**Bacterial Diversity**

The abundance of bacterial taxa varied during the 5-week time period (Figure 3.3, 3.7 and 3.8) but was more consistent over the sampled time period compared to 18S sequences. However, notable changes were observed for the *Cyanobacteria* and *Actinobacteria* groups (Figure 3.3). Diversity negatively correlated with DOC ($R^2=0.75$). *Anaeroplasmatales*, an order of *Mollicutes* often found in the human gut, increased with time (Stearns et al., 2011). This could be the result of increased recreational use, especially swimming, as spring progressed into summer. In the first 3 weeks *Actinobacteria* was the most abundant phylum of bacteria with a positive correlation to elevated pH.
A large visible cyanobacterial bloom, possibly containing *Microcystis aeruginosa* that was previously isolated from Lake Velence, was observed on the 4th and 5th time points (6/8...
and 6/13) during the lowest observed bacterial diversity and causing all water samples to be opaque and green in color (Kós et al., 1995). *Microcystis* abundance negatively correlated with bacterial diversity ($R^2=0.6$). Specifically, a large decline in *Actinobacteria* was observed during the bloom, possibly outcompeted by *Microcystis* sp. for nitrogen as observed in the opposing quadrant of the CCA ($R^2 =0.78$). A converse relationship was observed between *Actinobacteria* and *Microcystis* observed in the eutrophic Lake Taihu (Paerl et al., 2011), in eastern China, in which *Actinobacteria* increased as the *Microcystis* population declined further supporting that the organisms compete for resources.

Due to the inability of *Microcystis* to fix nitrogen, it is often associated with eutrophication, benefitting from high nitrogen availability (Paerl et al., 2011). The bloom may have also benefitted from high calcium availability sourced from the bedrock further promoting cell aggregation leading to bloom formation that can cause hypoxia (Zhao et al., 2011). Increased temperature could have also been an influential variable, as *Microcystis aeruginosa* isolates have been shown to be limited by temperatures lower than 15°C (Jiang et al., 2008; Robarts and Zohary, 1987). Robarts and Zohary observed that growth rate increased with temperature and had an optima of 25°C or higher (Robarts and Zohary, 1987). Such blooms can pose a threat to public safety because they could release the compound, microcystin (MC), which can be lethal to humans and other biota (Paerl and Otten, 2013; Paerl et al., 2001). We detected three chemical forms of microcystin in collected biomass, MC-LR, MC-YR, and MC-WR totaling 3.342mg g$^{-1}$ dry weight (DW) of MC-LR equivalent (Farkas et al., 2014). MC-LR and MC-YR was not detected in Lake Velence until 1997, but were likely responsible for widespread skin, stomach, and eye
ailments related to a very large *Microcystis* bloom in 1992, which resulted in restricted access to Lake Velence and other infected waters (Habermehl et al., 1997). Research has demonstrated that the growth of the toxic strain, unique in containing microcystin synthesis genes (*mcyA-mcyJ*), was exacerbated by warmer water temperatures (25-30°C) (Davis et al., 2009). The peak of the large bloom in Lake Velence (6/13) occurred at 21.9°C, and by the following week the water temperature had risen an additional 7.8°C providing ideal conditions for not only large *Microcystis* blooms, but the synthesis of microcystin. The highest growth rate of the toxic strain was achieved when coupled with temperature increases and high available phosphorus (4-5μM). The combination of temperature and nutrient conditions correlated well with conditions we observed in Lake Velence (Davis et al., 2009). The isolation of a toxic strain of *Microcystis aeruginosa* and identification of MC-LR and MC-YR from Lake Velence in 1995, further supports that large toxic blooms promoted by ongoing eutrophication and temperature increases are probable (Davis et al., 2009; Kós et al., 1995).

In the case of Lake Velence, the increasing frequency and toxicity of blooming *Microcystis* could cause disease and possibly mortality in lake biota, potentially altering the ecosystem, its biodiversity, and severely limiting recreational activities. Microcystin has also been detected in several other Hungarian lakes and ponds suggesting the increasing probability of more toxic blooms throughout the region (Davis et al., 2009; Vasas et al., 2006; 2010; 2013).
Algal and Bacterial Dynamics

The relative algal abundance declined over the course of the five-week sampling period (May 18th – June 21st, 2012) most likely due in part to competition. The simultaneous decline in phosphorus concentration with an increase in the relative abundance of cyanobacteria, particularly *Microcystis*, supports the efficiency of phosphorus uptake by the cyanobacterium, especially following periods of limitation (Ritchie et al., 2001; Shen and Song, 2007) (Figure 3.8). Phosphorus uptake could be exacerbated by a large bloom, further limiting phosphorus availability to other community members. It is possible that the aforementioned *Microcystis* bloom could have outcompeted algae for nitrogen as well, but there was no significant negative correlation observed between alga taxa and *Microcystis*. Rather, we observed positive correlations between *Microcystis* and an alga, *Komma* sp. (R²=0.44). Some members of the algal community may have benefitted from cyanobacteria potentially supplying important growth promoting metabolites (Amin et al., 2012; Foster et al., 2011; Tarakhovskaya et al., 2007). In addition to the possible exchange of growth promoting metabolites, the *Komma* sp. could have been more efficient at sequestering and storing phosphorus at low concentrations, possibly giving the alga a competitive advantage over other taxa.
Figure 3.8: Canonical Correspondence Analysis (CCA) plotting chemical variables that drive variation in Bacterial phyla (with >2% relative abundance) and sample points. Taxa were most influenced by increases in available nitrogen in the first weeks of sampling (CCA1=43% CCA2=14%).

The algae population could have been negatively impacted by viral infection. After amplifying DNA, gel electrophoresis confirmed the presence of dsDNA viruses during some of the sample points (Figure 3.4). The targeted amplicon, MCP, ranges between
~350-500bp differing between viruses (Larsen et al., 2008), and future work includes the characterization of viral loads on population dynamics and resource allocations in anthropogenically-impacted, alkaline lakes.

Early increases in the observed *Crustacea* populations showed correlations to \(\alpha\)-Proteobacteria. Specifically, a significant positive correlation was observed between \(\alpha\)-Proteobacteria and Unk. *Crustacea* 1 \((R^2=0.87)\) and Unk. *Crustacea* 2 \((R^2=0.63)\). Known members of the \(\alpha\)-Proteobacteria can be endosymbionts and/or intracellular parasites (Newton et al., 2011). Half of the \(\alpha\)-Proteobacteria OTUs present in the Lake Velence samples belonged to the *Rickettsiales* order \((50.7\%)\), a taxa that frequently has symbiotic relationships with crustaceans (Fritsche et al., 1999). These results suggest that the corresponding increase between \(\alpha\)-Proteobacteria, Unk. *Crustacea* 1, and Unk. *Crustacea* 2 could represent linked population dynamics through cooperation or competition. The algae that could be available as potential food sources, coupled with increases in pH and sulfate, may have directly or indirectly benefitted the two observed *Crustacea* taxa \((\text{pH}: R^2=0.84\) and sulfate: \(R^2=0.86)\).

SparCC correlation network analysis was employed to identify possible correlations between community members (Friedman and Alm, 2012). As observed in Figure 3.9, the only significant alga taxa in the SparCC correlation network was an unknown *Chlamydomondaceae*. 
Figure 3.9: SparCC network map showing significant (p<0.05) of interactions with a 0.99 correlation between different OTUs incorporating each time point. All lines are indicative of positive interactions. Spotted circles denote bacterial taxa while grey are eukaryotic members of *Chlamydomonadaeae* have also been observed in neighboring systems, and in our case had a positive correlation (R²= 0.99, p<0.05) with the bacterial taxa, ACK-M1 an Actinomycetales and C111 an Acidimicrobiales (Borics et al., 2012). Previous studies have observed members of both Actinobacteria and Acidobacteria living in close, physical association with algae (e.g., the phycosphere or EPS of an alga cell) (Goecke et al., 2013; Green et al., 2015). This association is believed to supply key metabolites to the alga while providing carbon to the bacteria (Bell, 1983; Ramanan et al., 2015; Sapp et al., 2007). Establishment of the phycosphere in this case could have been further encouraged by aqueous geochemistry due to the carbonaceous bedrock characteristic of the Carpathian Basin (Haas, 2012). As previously discussed, the high availability of calcium and magnesium can promote EPS production, fostering cell aggregation and physical
association (Kalakoutskii et al., 1990; Zhao et al., 2011). However, both of these bacteria phyla can occur in soils and the gut of domestic livestock (Lauber et al., 2009; Reti et al., 2013; Shepherd et al., 2011). Therefore, we cannot determine whether the positive correlation is due to an actual physical association with the alga, or is the byproduct of agricultural runoff and/or the aquatic bird population living in the reed islands of Lake Velence (Baldi and Kisbenedek, 2000; De Jesús-Laboy et al., 2012).

In contrast, a negative correlation between *Sphingobacteria*, common in freshwater lakes, and two algal taxa was observed (R² = 0.72 for Unk. *Chlorophyceae* 1 and R² = 0.79 for *Thalassiosira* sp.). Previous research demonstrated the succession of *Sphingobacteria* in late stationary phase of *Thalassiosira rotula* growth. The previous work demonstrated that the bacteria flourished on the organic matter of deceased algae, but did not directly contribute to the decline (Grossart and Simon, 2007). The findings suggest that a negative correlation between an alga and bacteria does not necessarily indicate pathogenesis, but could also be indicative of opportunism. However, the nature of such a relationship is subject to change with algal growth phases; *i.e.* in certain instances bacteria can promote algal growth in early exponential phase, but can potentially transition into competitors or parasites (Grossart and Simon, 2007). It is therefore critical to consider the duality of these relationships, both positive and negative, occurring in the phycosphere (Seyedsayamdost et al., 2011b).
Conclusions

Characterization of the bacterial and eukaryotic microbial community of Lake Velence provided insights into the community composition and potentially influential relationships between taxa and measured abiotic parameters during a cyanobacterial bloom. Identifying positive and negative influences on the algal community is an important component of initial understanding of lake ecology and the potential impacts of human activity ultimately contributing to a broader understanding of ecosystem function. The findings reported here showed the rise of optimal conditions for large cyanobacterial blooms, producing microcystin, over the course of 5 weeks due to pervasive warm water temperatures and nutrient availability. When put in the context of global climate change and increasing eutrophication, these trends could be suggestive of larger and more frequent toxic blooms that will negatively impact biodiversity. The possible toxicity of these blooms should be further investigated especially in the context of human recreation and health, as well as on the altered dynamics of indigenous populations crucial to nutrient cycling and overall water quality in these unique freshwater systems of the Carpathian Basin.

Acknowledgements

The authors wish to thank Katalin Barkács and Laura Jurecska (Cooperative Research Centre for Environmental Sciences, Budapest, Hungary) for the chemical analysis of water samples. We would also like to thank Frank W. Schneider Agriculture Specialist, for U.S. Customs & Border Protection for his assistance and Dr. Bharath Prithiviraj with the Plant Biology Division at the Samuel Roberts Noble Foundation Inc.
for his bioinformatics expertise. Thanks are also due to Devon Andor Bell, Susi Szeremy, Ajna Akantisz, Hugh MacMurray, and Shane Mortensen for sampling assistance.

**Funding**

Support was provided by the US Department of Energy-Advancements in Sustainable Algal Production (ASAP) program under contract DE-EE0005993. T.S.A. Bell was also partially supported by the NSF IGERT Program in Geobiological Systems (DGE 0654336). T. Felföldi was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. Additional support was provided by the Hungarian Scientific Research Fund (grant numbers PD 105407 and K 116275). G. Vasas was supported by the National Research, Development, and Innovation Office of Hungary (NKFIH-119647).


CHAPTER 4

CHARACTERIZING THE MICROBIAL COMMUNITY AND ITS INTRINSIC ABILITY TO PRODUCE ALGAL BIODIESEL IN WASTEWATER TREATMENT LAGOONS

Contribution of Authors and Co-Authors

Manuscript in Chapter 4

Author: Tisza Ann Szeremy Bell
Contributions: Collected field data, conducted most experiments, analyzed data and wrote manuscript

Co-Author: Lakotah Doig
Contributions: Aided in portions of field work, and DNA extraction and amplification

Co-Author: Brent M. Peyton
Contributions: Aided in data analysis and manuscript preparation

Co-Author: Robin Gerlach
Contributions: Principal Investigator

Corresponding Author: Matthew W. Fields
Contributions: Principal Investigator
Manuscript Information Page

Type Tisza Ann Szeremy Bell, Lakotah Doig, Brent M. Peyton, Robin Gerlach, and Matthew W. Fields

Journal Name:

Status of Manuscript: (Put an x in one of the options below)

- X Prepared for submission to a peer-reviewed journal
- Officially submitted to a peer-review journal
- Accepted by a peer-reviewed journal
- Published in a peer-reviewed journal
Title: CHARACTERIZING THE MICROBIAL COMMUNITY AND ITS INTRINSIC ABILITY TO PRODUCE ALGAL BIODIESEL IN WASTEWATER TREATMENT LAGOONS

Authors: Tisza Ann Szeremy Bell 1,2, Lakotah Doig 1,2, Brent M. Peyton 2,3,4, Robin Gerlach 2,3,4, and Matthew W. Fields 1,2,3

Affiliations: 1Department of Microbiology & Immunology, Montana State University, Bozeman, MT USA; 2Center for Biofilm Engineering, Montana State University, Bozeman, MT USA; 3Energy Research Institute, Montana State University, Bozeman, MT USA; 4Department of Chemical and Biological Engineering, Montana State University, Bozeman MT USA

Keywords: algal biomass, algal biofuel, bio-oil, bio-diesel, phycosphere, wastewater

Corresponding author: Matthew W. Fields Department of Microbiology and Immunology and Center for Biofilm Engineering, Montana State University, Bozeman, MT 59717, USA
Abstract

The large-scale production of algae for biofuel synthesis has the potential to help alleviate current energy problems. The vast majority of research on algal biofuel production has been conducted on single species isolates in closed systems that are costly to maintain. Growth of algae in wastewater has potential to help overcome shortages of water, nitrogen, and phosphorus. However, cultivation in open wastewater systems presents several unique challenges that include colonization by a variety of bacteria, archaea, viruses, and other eukaryotic species. Before we can simultaneously control positive and negative microbial interactions for algal biodiesel production in open wastewater systems we must first characterize and understand possible interactions in these systems. Here, we monitored the microbial community, water chemistry, and FAME content of 5 connected wastewater treatment lagoons over the course of a year. DNA was sequenced from all three domains of life in addition to viral DNA specific to eukaryotic algae. Each lagoon demonstrated a unique community profile even when taxa were classified at the coarsest scale. Statistically significant correlations were observed between community members, chemical variables, and FAME concentrations. In contrast to previous findings that correlated elevated FAME with nutrient deprivation, we observed high levels of FAME (up to 48% w/w) in some nutrient replete lagoons. High FAME concentrations were significantly correlated to ammonia, nitrite, TKN, pH, phosphate, and phosphorus ($R^2= 0.74$). However, we found that results were substantially influenced by how FAME concentrations were quantified, either as %w/w in the biomass or by mass per volume (g/L) of water in the system. FAME quantified by volume (g/L) correlated with higher biomass concentrations ($R^2= 0.78$).
whereas high FAME as %w/w coincided with lower cell densities. Under these uncontrolled conditions, high biomass concentrations provided for a higher overall biofuel content (g/L) than high lipid content with low biomass. High FAME volumes also correlated to the abundance of particular bacteria and showed a weak correlation with a dsDNA virus specific to eukaryotic algae (phycovirus) \((R^2=0.2)\). Our findings suggest the feasibility of algal biofuel production using wastewater lagoons and shows both positive and negative interactions within the diverse microbial community. The resulting data can provide significant insight into biomass and/or lipid accumulation in an open system.

**Introduction**

Algal derived biodiesel has demonstrated advantages over other crop-based fuels. Algae require less land area for the same fuel production and photosynthesize at almost double the efficiency of terrestrial plants (Chisti, 2013). Algal biodiesel content, derived from cellular lipid bodies, can be maximized through strain selection, application of environmental stress, and/or nutrient starvation applied at certain points in the growth curve (Chisti, 2013; Shurin et al., 2013; Smith et al., 2010). Utilizing stress mechanisms mirrors Odum and others’ concept in which stress can be quantified as inputs and outputs, the input being the stress and the output being lipid accumulation (Fields et al., 2014; Odum et al., 1979) Successful maximization of lipid content has almost exclusively been limited to monocultures in closed and controlled systems. Currently, the mass cultivation of algae in closed systems that require large energy inputs severely limit the economic viability of algal biodiesel (Smith and Crews, 2013; Smith et al., 2010). Logically open systems that
rely on sunlight for energy proffer a cheaper option but introduce a whole suite of additional variables, specifically, contamination and colonization by other algae, fungi, bacteria, and viruses for which no practical control methods exist (Craggs et al., 2011).

However, rather than combating the environment at large, basic ecologic principles on diversity may instruct us on how to cultivate a stable, predictable, and productive system. For example, algal growth rates have been shown to increase when in tight physical association with certain beneficial bacteria (Guo and Tong, 2013). Termed the “phycosphere” by Bell and Mitchell (1972), this is a volume near the algal cells where algal exopolymeric substances (EPS) provide niche space for bacteria to interact (Bell and Mitchell, 1972; Bell, 1983). The alga capitalizes on a steady supply of essential vitamins and growth promoting and antibacterial metabolites produced by these bacteria while the bacteria receive oxygen and organic compounds. Furthermore, colonization of the phycosphere acts to inhibit parasitic bacteria lending stability to the overall system (Kazamia et al., 2012a; 2012b; Smith and Crews, 2013). In addition, the symbiosis of bacteria and algae possess industrially relevant capabilities. Imase et al. (2008) demonstrated the degradation of propionate by an alga and bacterium consortia not possible by either organism individually (Imase et al., 2008).

Although we have identified a valuable role of bacteria in promoting stable systems for mass algal growth, our appreciation has yet to extend toward viruses. dsDNA viruses specific to eukaryotic algae have often been blamed for mass culture crashes and are typically not considered crucial to system predictability, especially in industrial settings. Recent literature suggests that algae have evolved coping mechanisms for infection citing
that viral infection has yet to adversely affect commercial ponds and may actually enhance production (Borowitzka and Moheimani, 2012; Jacquet et al., 2012; Sanmukh et al.; Thyrhaug et al., 2003). Models including diversity metrics predicted lipid productivity better than biomass alone (Graham et al., 2016). Therefore, improving our understanding of the microbial ecology in diverse systems is vital to the scalable mass production of algal biodiesel.

Despite not being microbiologically based, many basic ecologic principles are applicable to microbial community dynamics. Mathematical models and field experiments showed phytoplankton biodiversity resulted in higher productivity mirroring some of the earlier ecologic studies on grasslands (Downing and Leibold, 2002; McGrady-Steed et al., 1997; Striebel et al., 2009a; Tilman et al., 1981). Applying diversity principles to algal systems, Stockenreiter et al. (2011) demonstrated the legitimacy of this tenet, even when productivity was gauged by lipid content rather than biomass (Stockenreiter et al., 2011). Additionally, wastewater was used as media successfully negating the costly need for nutrient additions and showing the applicability of ecological principles to industrial production settings and challenges (Stockenreiter et al., 2011; 2013; 2016). Algal wastewater remediation has been used for many years to significantly reduce potentially hazardous nutrients (Li et al., 2013; Oswald and Gotaas, 1957). More recently, the reduction of heavy metals, pharmaceuticals, and steroid hormone concentrations has been observed (Li et al., 2013). Municipal and agricultural wastewater are generally high in macronutrients, especially nitrogen and phosphorus, which are imperative for algal cellular and metabolic processes. However, anthropogenic nitrogen and phosphorus released into
the environment have impacted nearly every ecosystem, drastically altering even the most remote alpine systems (Baron et al., 2000; Brahney et al., 2015; Elser et al., 2009; Gardner et al., 2008; Miller and McKnight, 2012; Wolfe et al., 2001). Acting as a sink, algae reduce harmful nitrogen and phosphorus concentrations in large open wastewater systems that perpetuate high growth rates and biomass accumulation.

In this study, we assessed the microbial community of large wastewater lagoons. We utilized high-throughput sequencing to monitor fluctuations in the community structure and lipid productivity of five interconnected municipal wastewater treatment lagoons in Logan City, Utah, U.S.A. Chemical and microbial analysis was conducted to determine factors correlating with community changes both within and between lagoons over time. Each lagoon possessed a unique community structure and different water chemistry profiles. Samples collected from the Logan City lagoons were sequenced revealing correlations in community changes with physicochemical parameters both within and between lagoons, specifically pertaining to biomass and fatty acid methyl ester (FAME) content.

Materials and Methods

Site Description

Logan City, Utah, U.S.A. is located approximately 40 km west of the northern arm of the Great Salt Lake (GSL). The City owns and operates a system of seven interconnected lagoons (A1, A2, B1, B2, C, D, and E) equipped with pontoon-mounted surface aerators that facilitate the treatment of municipal and surrounding community wastewater (Figure
Algae have naturally colonized the 460-acre system of lagoons and aid in the wastewater treatment process by absorbing soluble phosphorus and nitrogen. During the summer months (late May-September), effluent discharged from the final lagoon (lagoon E) supplied water for neighboring field irrigation. The remainder of the year effluent entered a series of wetland polishing cells before being discharged into the Cutler Reservoir.

![Figure 4.1](image)

Figure 4.1: Aerial schematic of the lagoons at the City of Logan Wastewater Treatment Plant. Arrows indicate the flow of water when all lagoons are in use. X’s signify the sampling location.

The Wastewater Treatment Plant (WWTP) receives an average of 53 million liters of wastewater per day (14 million gallons per day) making it one of the largest operating non-mechanized plants in the country. Due to increasing population growth, effluent is projected to nearly double by 2040 (Davies, 2015). Plant managers have anticipated that
the lagoon system will not be capable of handling the increase while continuing to meet nutrient limits, particularly a total phosphorus concentration of less than 0.69 mg/L. Therefore the construction of a mechanical treatment facility has been planned for the end of 2017 (Davies, 2015). Prior to the 2016 re-design, these lagoons offered a unique opportunity to observe the microbial community changes over time in a very large open pond system. The study presented here is specifically focused on correlations between community members, chemical parameters, and lipid concentration aiming to gain insight on nutrient recycling and algal biomass production.

**Sample Collection**

Sampling of five of the operating lagoons A1, B1, C, D, and E (all seven are rarely in operation) for water chemistry and community analysis was conducted in September 2013, November 2013, May 2014, and September 2014 (Figure 4.1). Extractable lipid was also analyzed for lagoons B, C, and D for November 2013 and A1, B1, C, D, and E for May 2014 and September 2014.

**Water Chemistry**

Samples for the five lagoons collected in September 2013, November 2013, May 2014, and September 2014 were analyzed for alkalinity, ammonia, bromide, chloride, conductance, dissolved oxygen (DO), dissolved organic carbon (DOC), fluoride, nitrate, nitrite, ortho-phosphate, phosphorus, sulfate, and total nitrogen (TN) following analytical methods approved by the Environmental Protection Agency (EPA) (Chemtech-Ford Laboratory, Sandy, Utah). In addition, pH, temperature, and dissolved oxygen (DO)
concentrations were recorded in the field during sampling (Hach HQ40d, Loveland, CO). Retention time for each lagoon was calculated by dividing lagoon volume by the influent flow rate.

**Algal Cell Counts**

To quantify algal cell numbers, 500mL from each lagoon at each time point was collected in an amber bottle and preserved with Lugol’s solution (Carolina Biological Supply Company, Burlington, NC) until cell counts could be conducted in the lab (Gardner et al., 2008; Moheimani et al., 2013; Stockenreiter et al., 2013; Striebel et al., 2009a; Tilman et al., 1981). We performed direct cell counts (cells/mL) using a hemocytometer (Sigma-Aldrich, St. Louis, MO) and transmitted/epifluorescence light microscope (Nikon Eclipse E8000), counting a minimum of 400 cells for statistical accuracy (Eustance et al., 2013; Gardner et al., 2013; Post, 1977).

**FAME Analysis**

Concentrated biomass was required for fatty acid methyl ester (FAME) analysis. A continuous centrifuge was used to collect biomass in November for lagoons B1, C, and D. Conditions prevented collecting biomass from lagoons A1 and E. Beginning in May a 5μm pore size filter bag was used for all subsequent collections. Whole cell or in situ transesterification of extractable lipids was performed on lyophalized biomass was performed. (Lohman et al., 2013). Gas chromatography coupled to a mass spectrometer (GC-MS) analysis was used to identify FAMEs based on retention time and quantified via surrogate standards. These data were used to calculate the biodiesel potential (%w/w) and
biodiesel content (g/L) of the wastewater lagoon samples (Eustance et al., 2013; Gardner et al., 2013).

**DNA Extraction and Sequencing**

Water samples were filtered for DNA analysis on site with a 0.22μm polyethersulfone membrane filter, placed on dry ice for transport, and stored at -80°C until DNA extraction in the lab according to methods described by Bell et al. (2016) (Bell et al., 2016). To amplify DNA with PCR and prepare it for sequencing, primers were designed for multiplexed amplicon sequencing and included a linker primer sequence that was compatible with Illumina flow cells. The SSU rRNA gene sequences for *Eukarya*, *Archaea*, and *Bacteria* were amplified via 25 cycles of PCR with the following primers; 7F (5’-ACCTGGTTGATCCTGCCAG-3’) and 591R (5’-GGAGCTGGAATTACCG-3’) for *Eukarya*, FD1 (5´AGAGTTTGATCCTGGCTCAG-3´) and 529R (5´-CGCGGCTGCTGGCAC-3´), which targeted the V1-V3 region of *Bacteria* and 751F (5’CCGACGGTGAGRGRYGAA-3’) and 1204R (5’-TTMGGGGCATRCNKACCT-3’) for *Archaea* (Baker et al., 2003; Barnhart et al., 2013; Bell et al., 2016; Bowen De León et al., 2012). PCR products of the correct size were confirmed using a 1% agarose gel.

To investigate the phycovirus community, we targeted an additional biomarker, the SSU rRNA gene sequence for the major capsid protein (MCP), specific to viruses infecting eukaryotic algae (Larsen et al., 2008). Due to the large size of these viruses (40-400 nm in diameter), the 0.22μm polyethersulfone membrane was an effective filter and the same DNA extract for the previously described amplicons was used (Van Etten et al., 1991). The MCP region was amplified utilizing touchdown PCR and the degenerate primers mcp
Forward (5’-GGYGGYCARCGYATTGA-3’) and mcp Reverse (5’
TGIARYTGYTCRAYIAGGTA-3’) as described by Larsen et al. (2008) (Larsen et al., 2008). Successful amplification was again confirmed using a 1% agarose gel with products ranging from 437-518 base pairs (Larsen et al., 2008). In addition, we used transmission electron microscopy (TEM) (LEO912AB TEM Zeiss, Oberkochen, Germany) to examine some samples for the presence of virus like particles following methods by Børsheim and others (Børsheim et al., 1990).

All amplicons were indexed with the Nextera XT kit and sequenced in house using the Illumina MiSeq platform (Illumina, San Diego, CA) following the manufacturer’s protocol for a paired-end run and V3 Reagent Kit (600 cycles).

**Bioinformatic Sequence and Community Analysis**

Data analysis was performed using the Quantitative Insights into Microbial Ecology (QIIME) software package, version 1.4.0 (Caporaso et al., 2010b). Parameter settings for demultiplexing were at a default level of 200 bp and 1000 bp in length. Metadata files were prepared according to a QIIME compatible template taking into account environmental sampling data on pH, temperature, and ionic concentrations. Libraries were split according to barcode for each of the respective groups (Archaea, Bacteria, Eukarya, and Phycovirus). Sequences were then concatenated for data normalization needed in downstream analysis. Operational taxonomic units (OTUs) were assigned using the closed reference OTU picking protocol. UCLUST was used to pre-cluster Archaea, Bacteria, Eukarya, at 97% identity and Phycoviruses at 99% identity, due to the degree of similarity of hypervariable gene regions within a population (Deng et al.,
Due to the lack of molecular tools to specifically identify viral genera, viral sequences were not carried downstream for taxonomic assignment (Deng et al., 2014; Roux et al., 2015; Wilhelm and Matteson, 2008; Zhong and Jacquet, 2013). Resulting Archaea, Bacteria, and Eukarya clusters were referenced against the Silva 108 database (Edgar, 2010).

Archaea, Bacteria, and Eukarya sequence reads that matched a Silva reference sequence at 97% identity were clustered within an OTU defined by a reference sequence. OTU assignment (and all subsequent steps) was performed for the combined Archaea, Bacteria, and Eukarya reads. The singleton OTUs were discarded. The centroid sequence in every cluster was selected to represent the cluster and aligned with the Silva core set using PyNAST (Caporaso et al., 2010a). Chimeric sequences, identified with Chimera Slayer (Haas et al., 2011) and reads that failed to align with PyNAST were excluded from subsequent analyses. PyNAST (v1.1) was used for sequence alignment and filtering through QIIME using default parameters.

Taxonomic assignments were additionally assigned using the retrained RDP Classifier (Wang et al., 2007) on the Silva 108 database for phylogenetic resolution at the genus level. Taxonomic summary for Archaea, Bacteria, and Eukarya was plotted to the genus level with a given relative abundance based on diversity and distribution pattern per domain. Taxa distribution was also summarized by time (sample day). The biodiversity analysis downstream between samples was derived using UniFrac and took into account the phylogenetic structure of the microbial communities (Lozupone and Knight, 2005).
The co-occurrence of OTUs was illustrated in a heat-map using the R vegan package version 2.0-10 (Oksanen, 2011). Since some taxa had a 0% relative abundance at certain time points, 0.1 was added to all values in order to be log transformed. In an effort to enhance the visual distribution of taxa, log transformed values were cubed and resulting values plotted (log(relative abundance + 0.1)^3). A SparCC analysis was subsequently used to construct community correlation networks by estimating linear correlation values between log transformed abundances based on the absolute number of sequences for an OTU rather than a relative abundance (Berry and Widder, 2014; Friedman and Alm, 2012). The key advantage of this analysis is that the ratio of two OTU fractions is independent of other OTU fluctuations included in the analysis (i.e., subcompositional coherence) (Friedman and Alm, 2012). Archaeal taxa were not included in this analysis due to a very low relative abundance that would have produced misleading relationships in a network model.

Diversity was quantified by calculating the effective number of OTUs for Archaea, Bacteria, Eukarya, and dsDNA viruses. The effective number of OTUs, or Hill Number, was calculated using the R package “rioja” and used to track fluctuations in diversity (Hill, 1973; Jost, 2006; Juggins, 2012).

Statistical Analysis

A principal coordinate analysis (PCA) was used to reduce dimensionality and give structure to the water chemistry variables obtained from each time point (Legendre and Gallagher, 2001). To incorporate taxonomic data, we used the direct-gradient ordination technique, Canonical Correspondence Analysis (CCA), which concurrently showed pond
taxa, time points, and water chemistry (Hall and Smol, 1992). This kind of ordination is appropriate when assessing community dynamics because it does not use Euclidean based metrics that assume linear trends in community change. Not only does the ordination show environmental factors influencing community change, but results suggest potential interactions between taxa (Amaral-Zettler et al., 2010). The first two axes, CCA1 and CCA2, typically account for the majority of observed variation. All axes are constrained to present a linear combination of the water chemistry that maximizes the dispersion of taxa (Hall and Smol, 1992).

Correlation and multiple regression analyses were conducted (November lagoons C and D, all lagoons May and September 2014) to investigate possible relationships between total %w/w FAME and chemical variables significant (p<0.05) correlations. All values were normalized using log_{10} transformation.

Linear discriminate analysis effect size (LEFSe) analysis was employed to determine the significance of effect relevance and biologic consistency (Blankenberg et al., 2010; Giardine et al., 2005; Goecks et al., 2010). We were able to split FAME values as %w/w and concentration into two categories, high (> 0.08mg/L or 19%w/w) and low (<0.08mg/L or 19%w/w), and obtain a list of taxa ordered by effect size on periods of high and low FAME.
Results

Water Chemistry

Water chemistry was monitored at each sampling time point and lagoon to observe fluctuations and correlations to community change and lipid content (Table 1). Using a PCA, we visualized the unique chemical composition and relatedness of each lagoon during the 4 sampling events (Figure 4.2). Overall, lagoons clustered together by sampling time and discrepancies were due to varying concentrations of phosphate, phosphorus, and nitrogen species. Lagoons A and B for each time point grouped near each other with the exception of lagoons A and B in September 2013, which is widely separated due to discrepancies in nitrite, ammonia, and TN. Lagoons C, D, and E did not consistently assemble together or with lagoons A and B. In November, Lagoons A, B, and E cluster together trending with higher concentrations of nitrate, DO, DOC, and pH. Higher concentrations of phosphate, phosphorus, ammonia, and TN were associated with lagoons A, B, and E in May as well as A, B, and D in September 2014. Elevated concentrations of DOC and DO caused the unusual congregation of lagoon B in November and lagoon C in September 2013 while elevated nitrate concentrations separated lagoon A in November, D in May, and C in September 2014.
Figure 4.2: A PCA of the chemical variables responsible for causing the most variation between the lagoons. Lagoons generally clustered together by sampling month. PCA1=34% PCA2=25%

Cell Number

Cell numbers were quantified for each lagoon and time point as log(cells/mL). The highest cell numbers were observed in lagoons A and B in November and May, but declined and remained low upon reaching lagoon C. Little fluctuation was observed in cell
number during September 2013 and 2014 among all lagoons with the exception of a decline in lagoon B in September 2013.

**FAME**

Biofuel content and potential was assessed by characterizing and quantifying biomass lipids and fatty acid methyl esters (FAMEs). The major FAME compounds identified were long and even carbon-number chains (C16:0, C16:1(z-9), C18:0, C18:1(e-9), C18:1(z-9), C18:2(e-9,12), C18:2(z-9,12), C18:3(z-9,12,15), C20:5(z-5,8,11,14,17), C22:6(e-2,4,6,8,10,12)), but odd-numbered fatty acid chains were also observed, a result of the methylmalonyl pathway (Guschina and Harwood, 2006). Figure 4.3 plots the log(cells/mL) with A) the major FAME compounds by %w/w or biodiesel potential and B) the major FAME compounds by concentration (g/L) or biodiesel content (%w/w). The discrepancies between Figure 4.3A and B illustrate the importance of biomass concentrations on overall lipid productivity, because despite observing high weight percent FAME content in Lagoon C September 2014, i.e. high lipid producing algae, there were relatively few cells in the system. This outcome translates to an overall low biofuel content. However, even though Lagoon B November 2013 had lower biodiesel potential (%w/w), the higher biomass concentration (cells/mL) at this time point provided for a higher overall biofuel content (g/L).
Figure 4.3: The log of algal cells per mL and fatty acid methyl esters quantification for each time point as A) weight (%w/w) and B) volume (g/L). In A) September 2014 Lagoon C had the highest FAME value by %w/w, but when cell number was included in the calculation B) November 2013 Lagoon B had the highest FAME value by volume.
Correlation and multiple regression analyses were conducted (November lagoons C and D, for all lagoons May and September 2014) to investigate possible relationships between total \%w/w FAME and chemical variables. Ammonia, Nitrite, TKN, pH, phosphate, and phosphorus had a significant (p<0.05) correlations with total FAME accounting for 74% of the variance in FAME \(F(6,6) = 6.54, \ p<0.05, \text{adjusted } R^2 = 0.74\). We observed fluctuations in ammonia (NH\(_3\)) and total nitrogen (TN) that likely impacted the biologic community. Ammonia was highest during longer retention times, discussed below. However, looking specifically at ammonia and TN in relation to our highest biofuel content in lagoon B for November (0.16 g/L), we observed a decrease in both ammonia and TN in Lagoon B.

Retention Time

To evaluate a potential relationship between biomass concentration and the length of time water resided in each lagoon, we calculated the retention time (days) for each lagoon. Longer retention times were observed in November and September 2014 in which it took approximately 80 and 70 days, respectively, for water to travel through the entire lagoon system. In contrast, September 2013 and May had much shorter residence times, of 54 and 53 days, respectively. Retention time correlated with cell number \(R^2=0.23\) as well as FAME \%w/w \(R^2=0.4\) but not with FAME concentration (g/L). However, cell number and FAME concentration (g/L) strongly correlated with each other indicating the necessity of longer residence times for higher biomass accumulation \(R^2=0.99\).
Community Structure and Diversity

The SSU rRNA for Bacteria, Archaea, Eukarya, and the MCP region for dsDNA viruses were amplified and sequenced for each time point and lagoon totaling 80 unique libraries. After screening sequences for errors (see Methods), 45,634,066 quality gene sequences were retrieved. Sequences with a 97% identity were clustered within an OTU, resulting in a total of 2,075 OTUs composed of 1,572, 58, and 445 OTUs of Bacteria, Archaea, and Eukarya respectively for all sample points. Additionally, we observed a total of 1,048,576 viral OTUs clustered at 99% similarity. Figures 5, 7, and 10 show heat maps of OTUs with greater than 2% relative abundance for Bacteria and Eukarya, and viral OTUs with greater than 1% relative abundance across all sampling time points and lagoons (Heat maps for Archaea (Supplementary Figure 4.3) and viral OTUs with >0.01% relative abundance (Supplementary Figure 4.7) can be found in the supplementary section. This information allowed us to observe fluctuations and trends in the community structure over time.

Diversity Estimates. Figure 4.4 shows the Hill Numbers for Bacteria, Archaea, and Eukarya used to characterize diversity in the lagoons over time (Chao et al., 2010; Hill, 1973; Hill et al., 2006). Archaeal diversity remained low and showed the least amount of variation over the course of the year but showed correlations to pH ($R^2=0.31$). Eukarya diversity also demonstrated little fluctuation but tended to have higher diversity in lagoon C and lower diversity in lagoons B or E. As the middle lagoon, lagoon C could have acted as the chemical intercessor between the lagoons close to the influent and those near the effluent meeting the needs of the general eukaryotic population.
Figure 4.4: A bar graph of the eukaryotic, archaeal, and bacterial Hill numbers for September 2013 (S13), November 2013 (N13), May 2014 (M14), and September 2014 (S14) in Lagoons, A-E. S13A = September 2013/Lagoon A sample, etc.

In contrast, bacterial diversity fluctuated between lagoons and was highest in B, D, and E, especially in September 2014, and strongly correlated with ammonia concentration ($R^2=0.39$) and pH to a lesser degree ($R^2=0.26$). Diversity measures and variation were highest amongst the dsDNA viruses. We did not observe a pattern between diversity and certain lagoons, but did see a general correlation with warmer temperatures and higher diversity, particularly for September 2014 ($R^2=0.4$) also correlating with increased pH ($R^2=0.36$). Viral diversity estimates are further discussed below with viral community structure.
Archaeal Community Structure. We successfully amplified and sequenced six orders of Archaea, mainly methanotrophs ubiquitous to wastewater treatment lagoon sludge (Mara and Horan, 2003). So few were detected due to the aerobic nature of the upper portion of the water column in the lagoons where samples were collected. As a result, these taxa behaved like rare species in many of our downstream analysis biasing the results. Therefore, Archaea were excluded from most of our downstream analysis.

Bacterial Community Structure. Depicted in the heatmap, we detected high abundances of cyanobacteria in lagoons C and E in September 2014 as well as proteobacteria in lagoon D (Figure 4.5). Actinobacteria was elevated in lagoons C, D, and E for September 2013 and November. Noticeable declines were observed in September 2013 and May among most taxa between lagoons B and C but had nearly recovered to their previous abundances in D. An unknown bacteria was pervasive in all of the lagoons until September 2014 when it was no longer detected. Candidatus aquirestis, Rhodocyclales, and an uncultured Sphingobacteriales were mutually observed in all lagoons during September 2013, May, and November.
Figure 4.5: Heatmap showing the Pearson correlation between all bacterial phyla with greater than 2% relative abundance for each lagoon culture. Relative abundances of taxa were log transformed. Taxonomic relatedness is depicted by the dendogram on the top of the heatmap. Similarity between lagoon culture communities was correlated and clustered accordingly depicted by the dendogram on the left.

*Microcystis* was nearly absent in November and May but had elevated abundances in September 2013 (3.13%) and September 2014 (9.9%). Previously undetected, high abundances of an unknown cyanobacterium were observed in the final 3 lagoons for
September 2014 (C=21.29%, D=19.29%, E=11.83%) contributing to the similar bacterial community composition of these 3 lagoons reflected in the dendogram of the heatmap and clustering in the first quadrant of the CCA (Figure 4.5 and 4.6). We did observe *Methylophilales* in previous months below 4% relative abundance, but again saw substantially elevated abundances up to 14.09% in September 2014.

We observed the greatest relative abundance of nitrogen fixing bacteria, especially Cyanobacteria, in September 2014 (53% of bacterial relative abundance). Abundance largely increased in each subsequent lagoon from A to E. In contrast, nitrogen fixer abundance decreased in each subsequent lagoon in November, which had the second highest diversity. Ammonia concentrations trended with the relative abundance of the nitrogen fixers in September 2013 and 2014 suggesting the wastewater provided an ample source of ammonia in November and May.

Potential algal symbionts, possibly inhabiting the phycosphere, were also identified. *Pseudomonas* has been shown to colonize the phycosphere of some algae providing critical nutrients (Bell, 1983; Guo and Tong, 2013; Xie et al., 2013). We also detected *Rhizobiales*, an order of bacteria that includes several phycosphere colonizing genera. In particular, the genus *Rhizobium* has been found to have a near ubiquitous relationship with green algae and can increase growth rates several orders of magnitude (Kim et al., 2014).
Figure 4.6: A Canonical Correspondence Analysis (CCA) of bacterial phyla with greater than 2% relative abundance. The axes are constrained by the three most influential chemical variables. The length of the arrow represents the relative importance of the variable while arrow direction is indicative of the rate of maximum change. The angles between environmental variables and taxa by domain give their approximate correlation. The location of the lagoon with respect to environmental variables gives the approximate value of the environmental variable in the respective lagoon (CCA1=44% and CCA2=29%).
Eukaryotic Community Structure. In the heatmap of eukaryotic relative abundances, the green alga *Chlorophyceae* had the greatest relative abundances when FAME by volume was at its highest. For example, in lagoon B for November the relative abundance of *Chlorophyceae* was 32.38% (Figure 4.7). However, despite *Chlorophyceae* having the greatest relative abundance in lagoon B for November, it did not have the highest correlation of phototrophic taxa with FAME by volume but still maintained one of the higher relative abundances averaging 19.9% of eukaryotes in all lagoons and time points. Instead, *Cercazoa* had the most significant statistical relationship of possible eukaryotic phototrophs with FAME by volume (R²=0.54) potentially influenced by temperature as shown in the CCA (R²=0.5) (Figure 4.8). An *Intramononucleata* also exhibited high relative abundances throughout the year. A protist, it is a subphylum of *Ciliophora*, which has demonstrated the greatest species diversity in wastewater systems responsible for the slime layer that covers piping and filtering systems (Mara and Horan, 2003). We also observed a high abundance of an unknown eukaryote whose role and impact on the system is currently undetermined.
Figure 4.7: Heatmap showing the Pearson correlation between all eukaryotic genera with greater than 2% relative abundance for each lagoon culture. Relative abundances of taxa were log transformed. Taxonomic relatedness is depicted by the dendogram on the top of the heatmap. Similarity between lagoon culture communities was correlated and clustered accordingly depicted by the dendogram on the left.
Figure 8: A Canonical Correspondence Analysis (CCA) of eukaryotic genera with greater than 2% relative abundance. The axes are constrained by the three most influential chemical variables. The length of the arrow represents the relative importance of the variable while arrow direction is indicative of the rate of maximum change. The angles between environmental variables and taxa by domain give their approximate correlation. The location of the lagoon with respect to environmental variables gives the approximate value of the environmental variable in the respective lagoon (CCA1=44% and CCA2=23%).
dsDNA Viruses Community Structure. We identified 81 unique dsDNA eukaryotic algal viruses with at least 0.01% relative abundance with a subset of 31 OTUs having a 1% or higher relative abundance (Figure 4.9).

![Graph showing Hill numbers for dsDNA Viruses with 0.01% and 1% relative abundances.](image)

Figure 4.9: A bar graph of the viral Hill numbers for both 0.01% and 1% relative abundances for September 2013 (S13), November 2013 (N13), May 2014 (M14), and September 2014 (S14) in Lagoons, A-E. S13A = September 2013/Lagoon A sample, etc.

We observed an inverse relationship between the Hill numbers of the total viral population with greater than 0.01% relative abundance and the diversity of a subset of this population with greater than 1% relative abundance ($R^2=0.2$) (Figure 4.10). For instance, the lowest diversity of the total viral population was observed in November, yet the smaller subset of this group with relative abundances greater than 1% showed its highest diversity at this time. Correlations could be an artifact of viral diversity being highest when retention time, and thus algal cell numbers and FAME, were low.
Figure 4.10: Heatmap showing the Pearson correlation between viral OTUs with greater than 1% relative abundance for each lagoon culture. Relative abundances of taxa were log transformed. Taxonomic relatedness is depicted by the dendogram on the top of the heatmap. Similarity between lagoon culture communities was correlated and clustered accordingly depicted by the dendogram on the left.

Prior to culling OTUs that had a less than 0.01% relative abundance, there were over 700,000 OTUs, a function of the strict 99% similarity constraint and the inherently
enormous diversity of this group. Though recently contested, no algal virus has been reported to infect more than one species of algae yet each algae species can be infected by several viral strains (Short, 2012). This number could be even larger due to cases of extreme host specificity that transcended the species level, in which a virus differentiated between strains *Emanlia huxleyi* and life cycle stages (haploid vs. diploid) (Frada et al., 2008). In light of a conservative estimate gauging the number of algae species to be between 100,000 and several million, the are implications of an astounding number of unique and undiscovered algal viruses (Wilson et al., 2009). To account for this complexity, we incorporated viral relative abundances of >0.01%, and two orders of magnitude higher, >1% relative abundance, providing the opportunity to observe trends that would not have otherwise been captured. This also provided the opportunity to compare levels of diversity dictated by relative abundance cutoffs, specifically, contributions of rare OTUs that may have collectively contributed to temporal variability tempering interpretations based on more stringent parameters (Shade and Gilbert, 2015). Although it was beyond the scope of this study to identify each viral OTU and its specific host, we obtained transmission electron microscope (TEM) images of selected lagoon samples that identified the presence of virus like particles with a similar morphology to phycoviruses described by Van Etten and Zhang (See Supplementary Figure 8) (Van Etten et al., 1991; Zhang et al., 2011).

An ordination of the viral population (>1% relative abundance) and environmental parameters identified temperature, nitrate, retention time, and ammonia as the most influential factors on the viral community (Figure 4.11).
Figure 4.11: A Canonical Correspondence Analysis (CCA) of viral OTUs with greater than 1% relative abundance. The axes are constrained by the three most influential chemical variables. The length of the arrow represents the relative importance of the variable while arrow direction is indicative of the rate of maximum change. The angles between environmental variables and taxa by domain give their approximate correlation. The location of the lagoon with respect to environmental variables gives the approximate value of the environmental variable in the respective lagoon (CCA1=56% and CCA2=31%).

The fraction of viruses in lagoon B in November appeared to correspond with increased retention time and higher nitrate concentrations while the majority of viral OTUs were
associated with higher temperatures and lower nitrate concentrations. When the ordination was repeated with viral OTUs with >0.01% relative abundance, temperature, nitrate, retention time, and ammonia were again the most influential with a majority of the OTUs clustering around ammonia and the September time points (Supplementary Figure 9).

**Community Structure and FAME.** For the sake of this study, we operated under the assumption that the majority of FAME quantified was algal derived. Equally proficient fungal lipid producers exist but the absence of solid substrate in the water column likely prevented growth (Meng et al., 2009). Some bacteria taxa can also accrue more than 80% of dry cell weight as TAG, but the large pore size of filters for biomass collection should have allowed the passage of free-living bacteria with the filtrate (Alvarez and Steinbüchel, 2002). However, this method could not account for bacteria inhabiting the phycosphere. Though minimal, a bacterial and fungal contribution to our FAME values cannot be ruled out as methods to differentiate what fractions of FAME identified with GC-MS were synthesized by algae, bacteria, or other eukaryotes.

Two LefSe analyses were conducted incorporating different relative abundance cut-offs for viral OTUs in order to observe the effect relevance of higher and lower viral diversity on FAME. The first LEFSe analysis identified several viral OTUs at the >0.01% that best accounted for periods of low FAME by volume (Figure 4.12A). A subsequent repeat of the analysis including only >1% viral OTUs did not identify any viruses associated with low FAME but rather 3 OTUs with a high effect relevance for periods of high FAME (Figure 4.12B). We also identified a positive correlation between the viral OTU 309525 and FAME by volume ($R^2=0.2$).
Figure 4.12: Output from LEFSe analysis using the log_{10} of LDA scores to list the differential taxa/OTUs associated with periods of high and low FAME by volume (g/L) with A) viral OTUs >0.01% and B) viral OTUs >1%

To further visualize community networks, SparCC was used to create correlation network identifying potential relationships between community members across lagoons and time points. Networks were estimated by linear correlation values between log transformed abundances based on the absolute number of sequences for an OTU rather than a relative abundance percentage (Berry and Widder, 2014; Friedman and Alm, 2012). The key advantage of this analysis is, for instance, that the ratio of the fractions of two
OTUs was independent of the fluctuations in other OTUs included in the analysis (i.e., subcompositional coherence) (Friedman and Alm, 2012). Figure 4.13 show significant correlations ($R^2=0.99$, $p<0.05$) between eukaryotes, bacteria, and viruses across all lagoons in November (highest recorded FAME) and September 2014 (lowest recorded FAME). Networks for the other two time points, May and September 2014, as well as for each individual lagoon (A, B, C, D, and E) incorporating the community composition for all time points (September 2013, November, May, and September 2014) (See supplementary Figures 10-16).

**Discussion**

From our results we found evidence of interplay between ammonium assimilation, direct/indirect nitrification, algal biomass/lipids, water residence time, and fluctuations in community composition and diversity. Diversity was characterized by overall low archaeal diversity regardless of lagoon or season, high bacterial diversity with seasonal fluctuation, and moderate eukaryotic diversity, in the shape of a bell-curve, as water cascaded into the subsequent lagoon (i.e. from lagoon A to B to C etc.). Insight into these dynamic and complex relationships build on our broader understanding of microbial community ecology ultimately allowing us to select for and predict productivity rates for products like lipid (Graham et al., 2016). Specifically, in an open system such as this, the synergy between nitrogen availability and algal lipid synthesis can be improved by mixed community dynamics and diversity that has the capacity to accumulate more light energy as lipid than a monoculture in a closed system (Smith et al., 2010).
Figure 4.13: A SparCC network map showing significant (p<0.05) of interactions with a 0.99 correlation between different OTUs incorporating all ponds for the November time point. All lines are indicative of positive interactions. Green circles denote algal taxa, pink denote bacterial and fungal taxa while purple are viral.
Influential Environmental Factors on FAME

Several observed nutrients likely had an important impact on the biotic community and ultimately FAME concentrations. Fluctuations in nitrogen, carbon, and phosphorus provided the most insight into community dynamics. Several previous studies have observed alterations in lipid metabolism in response to nutrient fluctuations, specifically increased lipid concentrations following periods of nitrogen starvation (Breuer et al., 2012; Chen et al., 2011; Guschina and Harwood, 2006; Hu et al., 2008). However, we observed both the highest FAME volume and nitrate concentration in lagoon B for November, similar to the findings of Stockenreiter and others in which eutrophic systems demonstrated higher lipid content (Stockenreiter et al., 2011). Other factors such as community composition and water retention time, which may have influenced the amount of time algae spent in the photic zone, were also factors and thus elevated FAME did not always ensue eutrophication.

Therefore, in this wastewater system nutrient availability contributed to high biomass and elevated lipid in lagoon B for November. Interlandi and Kilham (2001) also observed a correlation between high biomass and nutrient availability in a freshwater system (Interlandi and Kilham, 2001). The literature does not currently offer any examples of nitrogen starvation triggering lipid accumulation in a mixed community, suggesting that lipid triggers effective on monocultures may not be effective on mixed communities. Through realized niche occupation, a mixed community may have improved overall nitrogen allocation, utilization, and efficiency indirectly contributing to elevated FAME. In this setting, we see potential for the integration of total nitrogen assimilation and
adequate algal lipid production for biodiesel applications, with previous research achieving a maximum lipid productivity of 2.8 g/m²/day utilizing dairy wastewater, with the potential to produce 11,000 L/ha/year of biodiesel (Woertz et al., 2009).

Corresponding with lower FAME volumes, we observed low ammonia concentrations in September 2013 and 2014 that also trended with declines in the relative abundance of nitrogen-fixing bacteria (Supplementary Figure 4). Fluctuations in nitrate in September 2014 followed a similar trend. This may have been indicative of a possible interchange between nitrogen that was biologically fixed in the lagoons and seasonal variation of the wastewater influent chemical profile. Although no correlation between water residence time, a function of season, and FAME volume was identified, water residence time correlated with cell number ($R^2=0.23$), with a strong correlation between cell number and FAME volume (g/L) ($R^2=0.99$), suggesting longer residence time supported higher biomass accumulation.

**Algae Taxa Potentially Responsible for Elevated FAME**

The taxonomic composition of the biomass was investigated to further substantiate whether the amount of biomass was more significant to biofuel potential than the taxonomic composition of the biomass. We observed high relative abundance of *Chlorophyceae* during our highest recorded FAME volume (0.16 g/L) in lagoon B for November but found no significant correlation between *Chlorophyceae* and FAME. Instead, *Cercazoa* had the highest correlation of possible phototrophic eukaryotes with FAME by volume ($R^2=0.54$). *Cercazoa*, a eukaryotic phylum of protists include the chlorarachniophytes, a class of algae, as well as several plant and invertebrate parasites,
including *Vampyrella*, a genus of amoeba that predate on algae like “vampires” by extracting the cytoplasm out of a cell ultimately causing rapid reductions in algal numbers (Gong et al., 2015). However, the *Cercazoa* phyla is largely understudied, especially the chlorarachniophytes, with little sequencing information available despite their vast diversity in environmental systems (Burki and Keeling, 2014). As a result, without additional focused analysis, we were unable to conclude definitively whether this *Cercazoa* was an alga, an amoeba predating on algae (ex: *Vampyrella* sp.), or an invertebrate/plant parasite. If this taxa was not phototrophic, but an algal parasite and/or predator, it could have contributed to elevated FAME values by liberating nutrients trapped in the cells of its prey and opening previously occupied niche space, all of which could have benefitted the higher lipid producing algae at this time.

The second highest FAME volume (0.08 g/L) was observed in Lagoon A May 2014. The abundance of *Chlorophyceae* decreased between November and May and an increase of a *Chlorophyte* and Unknown Eukaryote was observed, suggesting that changes in the taxonomic composition of high biomass concentrations may not have drastically altered the concentration of FAME. Therefore, in an open, outdoor system, the biomass concentration could be more important than community composition (e.g., lipid-accumulating populations). In other words, it may not necessarily be imperative to have the superior lipid-accumulating taxa, but rather taxa capable of high biomass accretion and moderate lipid accumulation in this system. Given the open nature of the system, we wanted to examine what ecologic factors could have promoted periods of higher biomass rather than focusing on the specific taxonomic composition.
Ecologic Principles and the Functional Role of Interspecific Diversity on FAME

Early ecologic research established the tenant that diversity increased productivity (McGrady-Steed et al., 1997; P White et al., 2006; Prosser, 2012; Ptacnik et al., 2008; Stockenreiter et al., 2011; Tilman, 1977; Tilman et al., 1981). Subsequently, numerous studies have demonstrated the applicability of this principle to algal cultivation, observing higher biomass yields from diverse consortia over that of monocultures (Liu, 2016; Smith and Crews, 2013; Smith et al., 2010; Stockenreiter et al., 2011; Wacker et al., 2015). In microbial algal systems, biodiversity promotes carbon assimilation and more comprehensive nutrient uptake for productivity yielding more biomass than the most productive species alone (Liu, 2016; Striebel et al., 2009a). Previous research found that increased productivity was not only observed as biomass, but also as elevated lipid concentrations compared to monocultures, correlating a linear increase in algal lipid content to species richness (Liu, 2016; Stockenreiter et al., 2011; Wacker et al., 2015). The positive correlation of lipid content to species richness is likely a function of interspecific biodiversity resulting in resource use complementarity, in which the joint utilization of a resource results in greater productivity than that of a single taxa (Liu, 2016).

Periods of elevated FAME concentrations in this system, like lagoon B in November, could have been partially due to algal diversity perpetuating high biomass through interspecific rather than intraspecific competition. Specifically, different alga taxa possess different photo pigments excited by different wavelengths (Stomp et al., 2007; Striebel et al., 2009b). Termed “spectral niche theory”, rather than only utilizing one
portion of the spectra, the mixed community most likely exploited a larger fraction maximizing productivity potential, biomass, and FAME concentrations (Stomp et al., 2007; Striebel et al., 2009b). The number of eukaryotic phototroph OTUs ranged from 3 to 7 over the course of a year, with 5 OTUs appearing in lagoon B for November, the highest observed between November and May, supporting the role of resource complementarity in lipid productivity (See Supplementary Figure 17). Alternatively, we observed up to 6 phototrophic OTUs in September 2014, when temperature was elevated, but FAME concentrations by volume were the lowest. Elevated water temperatures and reduced water retention time was unique to the September time points. Elevated temperature may have benefitted taxa characteristic of the region with higher temperature requirements subsequently increasing the number of OTUs. Lower nutrient availability could have stunted resource use complementarity, in addition to reduced water retention, time that may have limited algal reproduction and vertical positioning within the photic zone, key to capturing particular light wavelengths. These factors may have stunted elevated FAME values that could have been anticipated due high diversity.

Increased bacterial diversity coincided with elevated FAME volume, potentially a result of higher algal cell numbers that consequently augmented the availability of phycosphere niche space. Previous research indicated that the phycosphere can support high bacterial diversity as corroborated by the isolation of 11 distinct cultivatable bacteria from a stock monoculture (Schwenk et al., 2014). Several positive correlations, depicted in SparCC network maps, between algal and bacterial taxa were observed in the lagoons over the course of the study period (Supplementary Figures 10-16). The literature offers
several instances in which algae grown in consortia with bacteria produce higher yields of biomass and lipid and have increased tolerance to stress (de-Bashan et al., 2002; 2008; Marshall et al., 2006; Natrah et al., 2013; Paul et al., 2012; Tate et al., 2012). This could be a function of bacterial provisions of critical vitamins like cobalamin and a range of growth promoting hormones, like indole, which was present in the lagoons (Supplementary Figure 18) (de-Bashan et al., 2008; Gonzalez and Bashan, 2000; Guo and Tong, 2013; Mendes and Vermelho, 2013; Tarakhovskaya et al., 2007). The lagoons contained several bacterial groups that have been previously identified as phycosphere dwellers including Betaproteobacteria, Flavobacterium, Rhizobiales, Rhodobacterales, and the genera Algoriphagus, Pseudomonas, Limnohabitans several of which were found to be directly benefitting the algal cells (Abby et al., 2014; Hume and Hann, 1984; Kim et al., 2014; Nedashkovskaya et al., 2004; Ramanan et al., 2016; Schäfer et al., 2002). The LEFSe analysis in Figure 4.12A supports the beneficial role bacteria play in algal productivity. Rhodobacterales, an order of bacteria frequently described in the phycosphere, correlated to periods of high FAME volume (lagoon B in November, lagoon A in May, and lagoon B in May) demonstrated in the LDA analysis (Figure 4.12 A) (Brinkhoff et al., 2008; Geng and Belas, 2010; Goecke et al., 2013; Lachnit et al., 2010). While we found evidence that the importance of biomass concentration may be of greater value than the selection of specific high lipid producing algae, the opposite could be true of bacteria in this system relying on key taxonomic groups that provide specific metabolites. Bacterial composition could provide additional information acting as indicators to system phenomena. Figure 4.12 B, a LefSe analysis including only viral OTUs
with >1% relative abundance, equivocated *Sedimenbacterium* with low FAME volumes. A genus of Sphingobacteria, *Sedimenbacterium* growth has been shown to be promoted by products excreted from algal cells rather than by mutualistic resource allocation (Lee, 2013). Therefore, elevated *Sedimenbacterium* growth could have been indicative of a lytic event, possibly the result of viral infection.

Viruses are generally considered a threat to algal cultures capable of causing mass lytic events or “pond crash” (Gerla et al., 2012; Golinski, 2015; Smith and Crews, 2013). Community dynamics between algae and phycoviruses ultimately influence biomass and FAME concentrations. However, from an ecologic perspective, phycoviruses are critical cornerstones of balanced biologic systems by facilitating the release previously unavailable carbon and nutrients (Wilhelm and Suttle, 1999). The results of our LEFSe analysis, which incorporated viral OTUs with at least 0.01% relative abundance supported viral culpability in low FAME concentrations (Figure 4.12 A). However, when the analysis was repeated including only viral OTUs with a minimum 1% relative abundance, many of which were almost exclusively observed in lagoon B in November, several viral OTUs had a positive correlation with high FAME (Figure 19 B). Specifically, we observed elevated relative abundances of the viral OTUs 709139 and 309525. OTU 309525 positively correlated with the major FAME compounds by volume ($R^2=0.2$). Our SparCC network analysis for the month of November also identified a positive correlation between Chlorophyta and the viral OTU 739342 ($R^2=0.99$, $p < 0.05$). Such correlations could be suggestive of viral predation on algal species possibly competing with dominant taxa resulting in a positive correlation with FAME (Figure 4.13).
A possible alternative, in which phycoviruses directly contributed to increased FAME, has been observed in marine cultures (Rosenwasser et al., 2014). Rosenwasser and others quantified metabolic changes that significantly increased fatty acid synthesis, especially longer carbon chains that compose transacylglycerol (TAG) (C16:0, C18:0, C:20) in infected algal cells within an hour of viral infection and later resulted in a ten-fold increase in TAG after 48 hours (Malitsky et al., 2016; Rosenwasser et al., 2014). Due to their large physical and genomic size (60-600nm and 160 to 560 kb including up to 600 protein encoding genes), phycoviruses siphon large concentrations of cellular scaffolding from their host, like amino acids and FAME, for their own replication and assembly (Rosenwasser et al., 2014; Van Etten et al., 1991). As observed in viruses infecting a strain of *E. huxleyi*, the viral hijacking and subsequent surge in the algal host’s lipid metabolism was strategic and imperative for viral replication (Rosenwasser et al., 2014). In the future we hope to focus on the viral population in this system and investigate whether a similar metabolic rewiring could have contributed to higher FAME volumes in the lagoons. We believe that the viral component of algal systems is highly relevant and applicable that could have major implications for algal biodiesel.

**Conclusion**

The sampling at Logan WWTP allowed us to establish molecular techniques to track community change and lipid potential. These techniques can be applied to other microbial communities including those in industrial settings with the potential to offer valuable information on complex microbial communities. In investigating the Logan
WWWTP lagoons, we observed that a mixed community may have a higher capacity for resource use and allocation contributing to increased biomass and, in turn, higher FAME. When considering open and mixed systems, ecology can elucidate community interactions and mechanisms that promote biomass generation and maintenance, such as those occurring in the phycosphere. It is imperative that we do not fix our gaze on the affairs of a few select lipid producing taxa, but comprehensively approach the system as a dynamic whole. By utilizing the full functional capacity of a system we can achieve the greatest productivity and yield.

**Funding**

This work was financially supported by the US Department of Energy-Advancements in Sustainable Algal Production (ASAP) program under contract DE-EE0005993, the National Science Foundation Sustainable Energy Pathways (SEP) Program under NSF CHE-1230632, and the US Department of Energy Bioenergy Technologies Office grant DE-FG36-08GO18161.

**Acknowledgements**

The authors would like to thank Dr. Ronald Sims and his laboratory at Utah State University for their helpful discussion and use of their facilities and equipment. We are indebted to Dr. Becky Hochstein and Sue Brumfield at Montana State University’s Transmission Electron Microscopy Core Laboratory for their time and expertise. We also offer many thanks to the City of Logan Wastewater Treatment Plant, specifically plant
operator Tim Lindsay and plant engineer Alexandra Rasband for all of their time and willingness to help make this research possible. We would also like to acknowledge Ron Fuller at Chem-Tech Ford laboratories for chemical analysis and interpretation. We are very grateful to Ann Willis and Karen Moll from Montana State University’s Microbiology Department and Center for Biofilm Engineering for their unfailing help in all aspects of this study. We thank Dr. Al Parker at Montana State University’s Center for Biofilm Engineering for bioinformatic guidance and Dr. Jane Turner at the University of Colorado for providing space to store samples over the holidays, and all members of the MSU Algal Biofuels Group for their helpful discussions.

Conflict of Interest

The authors declare no conflict of interest.
References


Knowledge Gaps and Areas for Future Research

World energy consumption is predicted to increase 48% between 2012 and 2040, and the development of renewable alternative energies is critical for the growing population (U.S. Energy Information Administration, 2016a). Algal biodiesel has enormous potential to contribute to renewable solutions, but will require mass cultivation to meet demands. Mass cultivation is most efficient in large open pond systems that are subject to invasion/colonization by other microorganisms. Subsequently, an understanding of applied microbial ecology can help maximize and maintain these systems.

Understanding of Large Scale In Situ Community Ecology

Despite the enormous potential of algal biodiesel, our limited understanding of the microbial community dynamics hinders scalability especially in large, open, mixed systems. As high throughput sequencing technologies improve, becoming more efficient at a lower cost, the technique is becoming more available giving insight into various systems. Even over the course of a year major strides are made that make this technology more accessible. Data processing has become more streamlined and community composition can be ascertained within hours of sequence completion. Ultimately, the aim for mass production of algal biodiesel is to create a probe or primer based surveillance systems for routine monitoring of large-scale microalgae cultures (Carney et al., 2014). With information on community structure, both qualitative and quantitative, the challenge
becomes improving our understanding of ecology necessary for interpretation of complex community interactions.

**Ecologic Principles Specific to Algae**

Substantial headway has been made utilizing ecologic tenets for the study of microbial systems (Fields et al., 2014; Kazamia et al., 2014; McBride et al., 2014; Nalley et al., 2014; Shurin et al., 2014; Smith and Crews, 2013; Smith et al., 2010; Stockenreiter et al., 2011). However, these tenets do not explain these systems in their entirety, such as the role of cross domain relationships on productivity and stability, the influence of viruses, nor do they provide a map of the system’s metagenome, metaproteome, or metabolome, and we must continue to advance our understanding. For instance, Hutchinson’s “Paradox of the Plankton” explored how phytoplankton communities seemingly contradicted the ecologic principle of competitive exclusion as several alga species can exist in a system in relative stability yet be performing similar functions, requiring the same resources, and occupying the same niche space (Hutchinson, 1961). Rather, one would expect to observe instances of competitive exclusion with the dominance of certain species at a given time responding to fluctuations in resource availability.

**Spectral Niche Theory**

The “Paradox of the Plankton” was explained by “spectral niche” theory which demonstrated that light resources are being partitioned by various photosynthetic pigments on a very fine scale (Striebel et al., 2009b). Chlorophylls a and b absorb less than half of the available solar spectrum. Yet, algae can also contain chlorophylls c1, c2, d, and f thus
utilize a broader range of the spectrum (Blankenship and Chen, 2013). Chlorophyll d and f are remarkable because of red shifts, in other words, they may be able to obtain energy from light in the infrared range that goes unaccounted in most productivity calculations (Chen et al., 2010; Manning and Strain, 1943; Striebel et al., 2009b). Thus, different combinations and amounts of various photosynthetic pigments likely allows algae to occupy very specific niche spaces resolving the paradox (Striebel et al., 2009b). Microbial systems may expand ecologic theory by offering unexplored metabolic pathways and relationships not previously considered.

**Role of Viruses as Community Members**

No algal virus OTU has been reported to infect more than one species of algae yet each algae species can be infected by several different viral OTUs (Short, 2012). In light of a conservative estimate, gauging the number of algae species to be between 100,000 and several million, the diversity of these viruses is immense (Wilson et al., 2009). Termed the microbial shunt, the viral-induced lytic process releases crucial organic matter to the environment, contributing to the global carbon cycle, and releasing nutrients to the immediate microbial community (Wilhelm and Suttle, 1999). In addition, to their global relevance, Rosenwasser and others have identified implications for biodiesel by quantifying metabolic changes that significantly increased fatty acid synthesis in infected algae cells within an hour of viral infection, increasing TAG by ten-fold (20%) when compared to the control (2%) (Malitsky et al., 2016; Rosenwasser et al., 2014). Due to their large physical and genomic size (60-600nm and 160 to 560 kb including up to 600 protein encoding genes), these viruses steal large concentrations of cellular scaffolding from their
host, like amino acids and FAME, for replication and assembly (Rosenwasser et al., 2014; Van Etten et al., 1991). As observed in viruses infecting a strain of *Emiliania huxleyi*, the hijacking and subsequent surge in the algal host’s lipid metabolism was imperative for viral replication (Rosenwasser et al., 2014).

These findings suggest viral infection could benefit lipid production for biofuels and represent the inverse of current approaches that aim to minimize infection. Without additional focused research, we cannot conclude whether a viral induced metabolic rewiring is scalable or if all virus/host infections behave similarly. We feel this is an area worthy of intensive and focused research.

**Closing**

The presented research demonstrates the applicability of ecologic principles and contributes to the larger body of knowledge utilizing ecology as a tool for economically viable and environmentally responsible mass cultivation strategies. Each system (raceway, alkaline lake, and wastewater lagoons) highlights the value of community interactions including potential symbiotic relationships that may contribute to higher concentrations of desired products, like algal lipids. However, the mechanistic nature of these relationships largely remains unexplored. Even the most studied of these relationships, like that between *E. huxleyi* and its bacterial symbiont *Phaeobacter* sp., offer new surprises as *Phaeobacter* sp. revealed its parasitic nature as its host ages (Seyedsayamdost et al., 2011b). Failing to acknowledge the dichotomy inherent to many of these relationships, let alone their sheer existence, stymies the fundamental understanding required to manage open algal systems.
Although a putative relationship was observed here between *C. vulgaris* and *Pseudomonas* sp. in Chapter 2, nothing is known about this relationship through time and its effect on lipid accumulation in the respective alga. Subsequent research on this and similar systems should include lipid analysis and techniques like fluorescent in situ hybridization (FISH), that utilize targeted probes for the visualization of the phycosphere and its occupants. Confirmation of bacteria living in the phycosphere coupled with lipid quantification would provide more insight into effects of bacteria on metabolic interactions and lipid production.

The raceway research described in Chapter 2 contested some basic ecologic assumptions about diversity and stability. The use of alkaline pH to limit microbial colonization and maintain a monoculture suggests that the absence of diversity could be mitigated with extreme conditions, like pH. However, is the benefit of employing alkalinity to minimize invasion negated by the subsequent limit in diversity? Could the application of something like spectral niche theory in an alkaline system facilitate the artificial construction of a productive and stable community? If so, is there an optimal number for diversity and how much overlap in niche occupation is tolerated before competition ensues? Additional research is required to explore these concepts which may or may not facilitate greater control and selectivity in cultivation systems by not only characterizing specific relationships between an alga and bacteria in closed controlled systems over time but by increasing the level of diversity (eukaryotic, bacterial, archaeal, and viral), quantifying the effects, and increasing system scale and exposure. Transitioning from a small-scale, man-made system, the community dynamics of bacteria and eukaryotes in a natural alkaline system were investigated and provided insight on the effects of high
nutrient availability (Chapter 3). Quantifying the potential relationships between nutrient availability, algal lipid production, and the larger microbial community composed of bacteria, archaea, fungi, and viruses is critical to further understanding of these systems.

The eukaryotic and bacterial community of a large natural alkaline system (Lake Velence, Hungary) was monitored for 6 weeks. Chapter 3 presents the composition of these communities and the effects of a toxic cyanobacterial bloom in the face of higher alkalinity supported by eutrophication and increasing temperature. The bloom did not appear to negatively impact the eukaryotic algae abundance but may have been responsible for the overall decline in eukaryotic diversity. Larger and more frequent toxic blooms are probable in Lake Velence, and similar eutrophic systems, as climate change continues to increase temperatures. In the context of artificial systems for algae cultivation, these findings suggest that high nutrient availability in alkaline systems could overtly favor a specific taxon and challenges the hypotheses that alkalinity alone can be used as a tool to control community composition and diversity when nutrients are replete. Implementing additional control strategies, like high salinity, in tandem with alkalinity may be required to mitigate similar bloom formations. This prompts future hypothesis to combine various environmental control tactics and ecologic principles to achieve stability and productivity under nutrient replete conditions.

High nutrient availability was further explored in Chapter 4 on a man-made system of WWTP lagoons. Research on the lagoons revealed the diversity and complexity of the microbial community spanning all 3 domains of bacteria, eukarya, and archaea, including phycoviruses. Findings demonstrated the importance of longer water retention time and
nutrient availability on lipid concentrations, suggesting that sufficient lipid for biodiesel can be achieved while simultaneously treating wastewater. The correlation between phycoviral diversity and a specific phycovirus on periods of elevated FAME may support findings of fatty acid synthesis upregulation in infected algae and suggest that viral diversity could enhance productivity. However, these findings must be examined in the context of the system’s large size and complexity. The use of smaller raceways (as used in Chapter 2) to assess community dynamics could reduce the complexity and give clearer insight into the interplay between nutrient availability and lipid production. Future work on the whole lagoon system would benefit from increased sampling frequency and the isolation and lipid screening of algal taxa.

In summation, the research presented in this dissertation contributes to the growing body of knowledge on algal biodiesel and provides a unique ecologic perspective concluding that the selective cultivation of alga taxa is possible in some open systems by utilizing extreme conditions like elevated pH, high nutrient availability may not be detrimental to algal lipid production if algal diversity is maintained, and phycoviruses may not necessarily be detrimental, but may even promote, algal lipid production. Each chapter offers different vantage points and possible approaches for algal growth optimization in open and mixed systems with varying nutrient availability. Continued research in these areas utilizing ecology as a tool is imperative for the realization of algal biodiesel as an economically viable and more sustainable energy source. Identifying the critical nutrients, like nitrogen and phosphorus, and microbial relationships, especially those facilitated by
the phycosphere, that support and/or enhance algal lipid production serve to improve productivity and stability.


APPENDICES
APPENDIX A

SUPPORTING TABLES AND FIGURES FOR CHAPTER 1 - INTRODUCTION
Figure A.2: DAPI stain. Diatoms are red and bacteria are blue.
Figure A.3: SYTO 9 stain. Diatoms are red, bacteria are green. Streaks in the lower and upper left portion of the picture are created by the movement of living bacterial cells during imaging.
APPENDIX B

SUPPORTING TABLES AND FIGURES FOR CHAPTER 3 - BACTERIA AND EUKARYA COMMUNITY DURING EUTROPHICATION AND TOXIC CYANOBACTERIAL BLOOMS IN THE ALKALINE LAKE VELENCE, HUNGARY
Figure A.4: Images from sample time point 5 of A) the Lake Velence shore and B) one of the collected samples depicting the *Microcystis aeruginosa* bloom.
Figure A.5: Images of diatoms collected from Lake Velence and stained with Nile Red for 5 minutes.
APPENDIX C

SUPPORTING TABLES AND FIGURES FOR CHAPTER 4 - MONITORING COMMUNITY ECOLOGY IN WASTEWATER TREATMENT LAGOONS FOR THE PRODUCTION OF ALGAL BIODIESEL
Figure A.6: Aerial view of Logan WWTP.
Table A.1: Illumina MiSeq sequencing statistics for all samples collected from Logan WWTP.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of samples</td>
<td>80</td>
</tr>
<tr>
<td>Number of OTUs</td>
<td>1,050,651</td>
</tr>
<tr>
<td>Number of total sequences</td>
<td>45,634,066</td>
</tr>
<tr>
<td>Minimum sequences/sample</td>
<td>185</td>
</tr>
<tr>
<td>Maximum sequences/sample</td>
<td>2,938,498</td>
</tr>
<tr>
<td>(mg/L)</td>
<td>S13A</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>Ammonia</td>
<td>1.9</td>
</tr>
<tr>
<td>Nitrate</td>
<td>b.d.</td>
</tr>
<tr>
<td>TN</td>
<td>9.2</td>
</tr>
<tr>
<td>DOC</td>
<td>6.6</td>
</tr>
<tr>
<td>Phosphate</td>
<td>1.9</td>
</tr>
<tr>
<td>Phosphorus</td>
<td>2.9</td>
</tr>
<tr>
<td>pH</td>
<td>7.99</td>
</tr>
<tr>
<td>DO</td>
<td>2</td>
</tr>
<tr>
<td>Retention Time (days)</td>
<td>19</td>
</tr>
</tbody>
</table>
Figure A.7: Inorganic water chemistry of the lagoons in September 2013, November 2013, and May 2014. Units are in mg/L with the exception of temperature (in °C) and pH. Abbreviations are as follows: TN (total nitrogen), DO (dissolved oxygen), and DOC (dissolved organic carbon).
Figure A.8: The log of algal cell counts (cells/mL) for each lagoon at all sampling points.
Figure A.9: Relative abundance of archaeal classes showing greater than 2% relative abundance for September 2013, November 2013, and May 2014.
Figure A.11: Relative abundance of bacteria phyla showing greater than 2% relative abundance for September 2013, November 2013, and May 2014.
Figure A.12: Samples testing positive for indole after a 48-hour incubation are indicated by the red ring at the surface of the liquid. Cultures on the left were grown on Bold’s medium and samples on the right were grown in sterilized Logan wastewater.
Figure A.13: The relative abundance of nitrogen fixing bacteria overlaid by line graphs showing the ammonia and nitrate concentrations (mg/L) for September 2013, November 2013, May 2014, and September 2014 in each lagoon.
Figure A.14: Relative abundance of eukaryotic classes showing greater than 2% relative abundance for September 2013, November 2013, May 2014, and September 2014.
Figure A.15: Dominant FAME compounds for Logan WWTP lagoons in November 2013 and May 2014.
Figure A.16: Total FAME concentrations for all sampling points.
Figure A.17: Eukaryotic phototroph $R^2$ values with FAME quantified as %w/w (blue) and mg/L (red).
Figure A.18: SparCC correlation network map of Cercozoa and other community members ($R^2=0.5$). Cercozoa is a phylum of protists that includes green alga, plant parasites, and “vampire” amoeba.
Figure A.19: A SparCC correlation network map of bacteria and eukaryotic phototrophs \((R^2=0.99)\). The lines between taxa indicate whether the correlation is positive (blue) or negative (red).
Figure A.20: Relative abundance of viral OTUs showing greater than 1% relative abundance for September 2013, November 2013, May 2014, and September 2014.
Figure A.21: TEM image of virus like particle with similar morphology to previously dsDNA viruses.
Figure A.22: Heatmap showing the Pearson correlation between viral OTUs with greater than 0.01% relative abundance for each lagoon culture. Relative abundances of taxa were log transformed. Taxonomic relatedness is depicted by the dendogram on the top of the heatmap. Similarity between lagoon culture communities was correlated and clustered accordingly depicted by the dendogram on the left.
Figure A.23: A bar graph of the viral Hill numbers for both 0.01% and 1% relative abundances for September 2013 (S13), November 2013 (N13), May 2014 (M14), and September 2014 (S14) in Lagoons, A-E. S13A = September 2013/Lagoon A sample, etc.
Figure A.24: A Canonical Correspondence Analysis (CCA) of viral OTUs with greater than 0.01% relative abundance. The axes are constrained by the three most influential chemical variables. The length of the arrow represents the relative importance of the variable while arrow direction is indicative of the rate of maximum change. The angles between environmental variables and taxa by domain give their approximate correlation. The location of the lagoon with respect to environmental variables gives the approximate value of the environmental variable in the respective lagoon (CCA1=53% and CCA2=28%).
Figure A.25: The number of viral and eukaryotic phototroph OTUs for September 2013 (S13), November 2013 (N13), May 2014 (M14), and September 2014 (S14) in Lagoons, A-E. S13A = September 2013/Lagoon A sample, etc. Red and purple arrows point out trends in viral diversity.
Figure A.26: SparCC network map of May 2014 community members showing a 0.99 correlation. Negative correlations are shown in red while positive are shown in black. Denovo # refers to an algal virus OTU.
Figure A.27: SparCC network map of September 2013 community members showing a 0.99 correlation. Negative correlations are shown in red while positive are shown in black. Denovo # refers to an algal virus OTU.
Figure A.28: SparCC network map of September 2014 community members showing a 0.99 correlation. Negative correlations are shown in red while positive are shown in black. Denovo # refers to an algal virus OTU.
Figure A.29: SparCC network map of the Lagoon A, including samples from all time points, community members showing a 0.99 correlation. Negative correlations are shown in red while positive are shown in black. Denovo # refers to an algal virus OTU.
Figure A.30: SparCC network map of the Lagoon B, including samples from all time points, community members showing a 0.99 correlation. Negative correlations are shown in red while positive are shown in black. Denovo # refers to an algal virus OTU.
Figure A.31: SparCC network map of the Lagoon C, including samples from all time points, community members showing a 0.99 correlation. Negative correlations are shown in red while positive are shown in black. Denovo # refers to an algal virus OTU.
Figure A.32: SparCC network map of the Lagoon D, including samples from all time points, community members showing a 0.99 correlation. Negative correlations are shown in red while positive are shown in black. Denovo # refers to an algal virus OTU.
Figure A.33: SparCC network map of the Lagoon E, including samples from all time points, community members showing a 0.99 correlation. Negative correlations are shown in red while positive are shown in black. Denovo # refers to an algal virus OTU.
APPENDIX D

SOURCES AND RE-SOURCES: IMPORTANCE OF NUTRIENTS, RESOURCE ALLOCATION, AND ECOLOGY IN MICROALGAL CULTIVATION FOR LIPID ACCUMULATION
Invited Mini-Review

Title: Sources and Re-sources: Importance of nutrients, resource allocation, and ecology in microalgal cultivation for lipid accumulation

Authors: Fields, M.W. 1,3*, Adam Hise2, Egan Lohman3,4, Tisza Bell1,3, Rob Gardner3, Luisa F.A. Corredor1,3, Karen Moll1,3, Brent Peyton3,4, Greg Characklis2, and Robin Gerlach3,4

Affiliations: 1Department of Microbiology and Immunology, Montana State University, Bozeman, MT USA
2Department of Environmental Sciences and Engineering, University of North Carolina, Chapel Hill, NC USA
3Center for Biofilm Engineering, Montana State University, Bozeman, MT USA
4Department of Chemical and Biological Engineering; Montana State University, Bozeman, MT, USA

Corresponding Author: Matthew W. Fields, Associate Professor

109 Lewis Hall
Department of Microbiology & Immunology
Montana State University
Bozeman, MT 59717
406-994-7340
Introduction

In modern societies, petroleum-based products and fuels have strongly influenced human society and infrastructure. For example, energy, food, and chemicals make up approximately 70% of commerce on the planet (www.eia.gov), and petroleum/hydrocarbons directly and indirectly impact these commodities. Petroleum/hydrocarbon markets have become increasingly unpredictable and cause destabilized commodity prices (e.g., fuel, food). In addition, the environmental impacts from increased carbon dioxide (CO$_2$) without balanced CO$_2$ sequestration has contributed to the increases in atmospheric CO$_2$ levels. The amount of carbon released in 1 year from the consumption of fossil fuels is now more than 400-fold the amount of carbon that can be fixed via net global primary productivity (Dukes 2003). In order to offset the massive influx of CO$_2$ into the atmosphere, the utilization of renewable biofuels (e.g., ethanol, butanol, H$_2$, CH$_4$, and biodiesel) is needed.

Diatoms and green algae (microalgea) are eukaryotic, photoautotrophs that can utilize inorganic carbon (e.g., CO$_2$) as a carbon source and sunlight as an energy source, and many microalgae can store carbon and energy in the form of neutral lipids [e.g., triacylglycerides (TAGs)]. In addition to accumulating useful precursors for biofuels and chemical feed-stocks, the use of autotrophic microorganisms can further contribute to reduced CO$_2$ emissions through utilization of atmospheric CO$_2$. For these reasons, eukaryotic photoautotrophs have been studied in the context of lipid accumulation for over 50 years and were a focus of the U.S. Department of Energy’s Aquatic Species Program in the 1980s and 1990s (Sheehan, 1998). However, cheap petroleum prices eventually eroded
monetary support for alternative (and renewable) energy sources until increasing petroleum prices over the last two decades has reinvigorated interest in alternatives. The advent and increased use of fracking technologies has opened up new markets for petroleum and hydrocarbon reservoirs, and almost $190 \times 10^9$ was spent in the United States in 2012 to drill and “frac” for conventional petroleum hydrocarbons (www.eia.gov). However, the process of fracking increases the rate and not the supply of hydrocarbons, and peak hydrocarbon production is predicted to occur around 2030 (www.eia.gov). Regardless of current market conditions and availability of conventional sources, alternatives are needed to circumvent future economic and environmental impacts from continued exploration and harvesting of conventional petroleum hydrocarbons.

Chisti (2007) estimated conservatively (assuming a lipid content of 25–30% (w/w) in microalgae) that an area equivalent to 3% of the arable cropland in the United States would be required to grow sufficient microalgae to replace 50% of the transportation fuel needs in the United States. Although the interest in algal biofuels has been reinvigorated (Courchesne et al. 2009; Greenwell et al. 2010; Razghefard, 2013), significant fundamental and applied research is still needed to fully maximize algal biomass and biochemical production for biofuels and other products.

The accumulation of lipids is of substantial interest because these compounds are energy-rich biodiesel precursors (Dismukes et al., 2008; Hu et al., 2008). Much of the reported research has focused on the increase of algal lipid accumulation upon exposing cultures to a range of environmental stresses prior to harvest (Hu et al 2008; Valenzuela et al, 2012; Valenzuela et al., 2013; Mus et al., 2013; Lohman et al., 2013; references therein).
While the stress event can increase lipid accumulation, it can also limit biomass production, but the stress scenario provides a tractable method to study and understand lipid accumulation at the laboratory-scale. Because of the inherent connection between carbon (C), nitrogen (N), and phosphorus (P) uptake in biological systems, macronutrient deprivation has been proven to significantly enhance lipid accumulation in different diatom and algae species. Nitrogen limitation is the most commonly studied stress in green algae and diatoms; the effect of silica limitation is regularly studied in diatoms (Valenzuela et al., 2012; Lohman et al., 2013; Chu et al., 2013, Schnurr et al., 2013). Keeping in mind that a vast majority of living pools of C, N, and P resides in the microbial realm (Whitman et al., 1998), much work is needed to understand the link between C, N, and P in controlling resource allocation both with respect to natural and man-made systems. In this context, a 50% replacement of transportation fuel will impose a vast nutrient demand (Pate et al., 2012). However, microalgal biomass/product production can be coupled to wastewater resources (e.g., water, N, and P), and wastewater from agricultural, industrial, and municipal activity may provide a cost-effective source of nutrients. Agricultural and municipal wastewater can be high in N and P (Aslan and Kapdan 2006; Hoffmann 2002; Mallick 2002; Pittman et al. 2011), and thus, there is great potential for the integration of wastewater treatment and algal biofuel production (Figure 1).
Figure 1. The biological recycling of carbon, nitrogen, and phosphorus to harvest fuel and food linked to sunlight to reduce net consumption of N and P and net production of C.

However, an improved understanding of the relationship between the effects of N, P, and micronutrient availability on cellular resource allocation (cell growth versus lipid storage) in microalgae is needed. This Mini-Review will briefly discuss the current literature on the use of nutrient-deprivation and other growth conditions to control and optimize microalgal culture growth in the context of cell and lipid accumulation.

Nutrient Dependent Lipid Accumulation

Under optimal growth conditions, (i.e., adequate supply of nutrients including C, N, P and sunlight), algal biomass productivity can exceed 30 g dry weight m$^{-2}$ day$^{-1}$ (Gordon and Polle, 2007); however, the lipid content of the biomass is typically very low (<5% w/w) (Gordon and Polle, 2007). The low lipid content is due to lipid biosynthesis being a metabolic process that is stimulated by stress inducement. Essentially, biomass synthesis and lipid biosynthesis compete for photosynthetic assimilation of inorganic
carbon, and a fundamental metabolic switch is required to shift from biomass production to energy storage metabolism (Schuhmann et al. 2012; Valenzuela et al., 2012). As denoted by Odum (1985), stress is a syndrome that consists of inputs and outputs, and the input is the stressor that is contrasted to the stress, or the output. Lipids (the output) is typically believed to provide a storage function within the cell that enables the organism to endure adverse environmental conditions, i.e., the stressor. The output can be viewed as the cessation of cell production and the accumulation of lipids in response to the input of unbalanced resources (e.g., C, N, P, and/or sunlight). It is likely that there are trade-offs in terms of biomass versus lipid accumulation depending on the different levels of perturbation (Figure 2).

**Figure 2.** Hypothetical performance curve for an increasingly perturbed (i.e., stressed) microalgal system being used to produce photoautotrophic biomass and/or lipids. Adapted from Odum et al. (1979).
Recent research has provided evidence that lipids may also act as a reservoir for specific fatty acids such as poly-unsaturated fatty acids (PUFAs). PUFAs play a key role in the structural components of cell membranes, and as antioxidants (PUFAs can counteract free radical formation during photosynthesis). As such, PUFA-rich TAGs might donate specific compounds necessary to rapidly reorganize membranes through adaptive metabolic responses to sudden changes in environmental conditions (Khozin-Goldberg & Cohen, 2006). However, a recent study showed that PUFA content in lipids can negatively impact biodiesel standards based upon lipids from *Chlorella pyrenoidosa* (Shekh et al., 2013), and this result suggests that lipid composition and just not amounts should be considered. In either case, lipid is an energy-rich storage compound that can be chemically transesterified to produce FAME (fatty acid methyl esters), the biological equivalent to diesel fuel (a.k.a., biodiesel). However, in order to maximize lipid biosynthesis, the organism is typically induced through environmental stress conditions (Hu et al., 2008). In addition, most studies have been based upon axenic cultures with limited understanding of potential bacterial “contamination”, and thus, lipid accumulation may be different at different scales of biological resolution (discussed below).

Significant work has been done to identify and optimize stress inducement strategies that enhance lipid accumulation in microalgal species. Nutrient deprivation, specifically nitrogen depletion, is the most prevalent technique employed (Hu et al., 2008). Temperature variations, pH, salinity, light, osmotic, and chemical stress inducements have also been investigated with varying success (Sharma et al., 2012). This may be due to two factors: 1) Lack of requisite nutrients such as nitrogen limits an organisms’ capacity to
synthesize proteins necessary for biomass production (e.g., cellular division). In order to compensate, the organism must take advantage of alternative metabolic pathways for inorganic carbon fixation, such as fatty acid synthesis and hence store those de novo fatty acids as TAG (Msanne et al., 2012). 2) Photosynthesis and the electron transport chain in eukaryotic microalgae produce ATP and NADPH as energy “storage” and electron carrier metabolites, respectively. These metabolites are consumed during biomass production resulting in ADP and NADP⁺, which in turn are regenerated via photosystems. Under normal growth conditions, this cycle maintains a balanced ratio of the reduced and oxidized forms of these metabolites; however, when biomass production is impaired due to a lack of requisite nutrients, the pool of NADP⁺ and ADP can become depleted. This can lead to a potentially dangerous situation for the cell because photosynthesis is mainly controlled by light availability, and cannot be shut off completely. Fatty acid synthesis consumes NADPH and ATP; therefore, increased fatty acid synthesis replenishes the pool of required electron acceptors in the form of NADP⁺, and de novo fatty acids are most frequently stored as lipid (Brown et al, 2009). Here we will review the most successful strategies involving nutrient stress to induce lipid accumulation in commonly studied microalgal species.

Nitrogen and Phosphorus

Nutrient availability is critical for cell division and intracellular metabolite cycling, and once nutrients such as N or P become depleted or limited in the medium, invariably a steady decline in cellular reproduction rates ensues. Once this occurs, the activated metabolic pathways responsible for biomass production are down-regulated and cells instead divert and deposit much of the available C into lipid (Wang et al, 2009; Valenzuela
et al, 2013). There have been numerous studies to compare different N sources in the context of maximal biomass or lipid accumulation, and the results are different dependent upon the organism. Breuer et al. (2012) accumulated previous literature on 56 eukaryotic, photoautotrophic genera studied in the context of lipid accumulation that included: Amphora, Ankitrodesmus, Biddulphia, Botryococcus, Bracteacoccus, Chaetoceros, Chlamydomonas, Chlorella, Chlorococcum, Chroomonas, Crypteocodium, Cryptomonas, Cyclotella, Cylindrotheca, Dictyospaerium, Dunaliella, Ellipsoidion, Emuliania, Enteromorpha, Euglena, Fragilaria, Glossomastrix, Gymnodinium, Haematococcus, Hantzchia, Hemiselmis, Isochrysis, Monallantus, Monodus, Nannochloris, Nannochloropsis, Navicula, Neochloris, Nephroselmis, Nitschia, Ochromonas, Parietochloris, Pavlova, Phaeodactylum, Pheomonas, Polytoma, Porphyridium, Protosyphon, Prototheca, Rhodomonas, Rhodosorus, Scenedesmus, Scrippsiella, Selenastrum, Skeletonema, Stichococcus, Tetraselmis, Thalassiosira, Ulothrix, and Volvox. The authors chose Chlorella vulgaris, Chlorella zofingiensis, Nannochloris UTEX 1999, Neochloris oleoabundans, Scenedesmus obliquus, Dunaliella tertiolecta, Isochrysis galbana, Phaeodactylum tricornutum, and Prophyridium cruentum to conduct normalized growth and lipid accumulation studies with nitrate as the N-source (Breuer et al, 2012). Under N-deprivation, C. vulgaris, C. zofingiensis, N. oleoabundans, and S. obliquus accumulated over 35% dry weight as TAG, and S. obliquus and C. zofingiensis had the highest TAG productivity (mg l$^{-1}$ day$^{-1}$) among the nine compared strains.
When the model Chlorophyte *Chlamydomonas reinhardtii* was cultivated under N limitation, an increase in lipid was also observed. Interestingly, fully saturated C\textsubscript{16} fatty acids were the most abundantly synthesized compounds, whereas polyunsaturated C\textsubscript{18} fatty acids remained relatively unchanged in this organism under the tested conditions (Lohman et al., 2013). While nitrate supported increased biomass compared to ammonium in *Monoraphidium* sp. SB2 (Wu et al., 2013), *Chlorococcum ellipsoideum* exhibited elevated lipid levels with urea compared to nitrate (Li et al., 2013). A different *Scenedesmus* strain (sp. R-16) was shown to have the highest lipid accumulation with nitrate compared to urea, peptone, or yeast extract (Ren et al., 2013). To date, nitrate is a commonly studied N source used to understand nutrient deprivation to induce lipid accumulation; however, different N sources have different effects dependent upon the organism. This is most likely a consequence of typical habitat for the organism as well as long-term life history that is common for the respective species. As nutrient recycling becomes more evident, different types and mixtures of nutrients (*e.g.*, human, agriculture, industrial) will need to be evaluated. For example, two recent studies investigated the ability of *Chlamydomonas polypyre)noideum* and *Chlorella pyrenoidosa* to grow and accumulate lipids during cultivation on diary wastewater (Kothari et al., 2012; Kothari et al., 2013). We recently grew a green alga isolated from storage ponds of coal-bed water that produced lipids under nutrient deprivation (Fields, Nagy, and Barnhart, unpublished results). Recently, N deprivation was shown to induce lipid accumulation in the wastewater isolates, *Scenedesmus* sp. 131 and *Monoraphidium* sp. 92 with ammonium, nitrate, or urea (Eustance et al., 2013) or nitrate depletion in *Skeletonema marinoi* (Bertozzini et al., 2013).
Interestingly, *Ettlia oleoabundans* initiated lipid accumulation in response to increased temperature before nitrate was completely depleted (Yang et al., 2013). This result suggests that different combinations of potential stressors could impact lipid accumulation in different ways.

In addition to N, P starvation to induce lipid accumulation in microalgae has been studied. In general, greater lipid accumulation due to N deprivation has been observed compared to P deprivation as reported for various *Chlorella* species (Feng et al., 2012; Liang et al., 2012). When the marine diatom *Phaeodactylum tricornutum* was grown under N and P limitation, an increase in lipid accumulation was noticed in all limiting conditions (Valenzuela et al., 2012; Burrows et al., 2012). However, cultures of *P. tricornutum* that were limited exclusively in N showed a far more significant increase in TAG than cultures that were limited solely in P. The combined limitation of both N and P resulted in the highest lipid concentrations in *P. tricornutum* (Valenzuela et al., 2012; Valenzuela et al., 2013). Given the commonly accepted N:P ratio of 16:1 in microalgal biomass (Redfield, 1958), the *P. tricornutum* work demonstrated that the external N:P ratio was 27 and the cellular N:P ratio was 8 to 9:1 when lipid accumulation was detectable (Valenzuela et al., 2012).

Both N- and P-deprivation result in cell cycle cessation, but the relative lipid accumulation response is different, and this observation is most likely a consequence of cellular resource allocation (*e.g.*, protein/chlorophyll vs. nucleotides). Based upon results in *P. tricornutum*, we observed a five-fold greater increase in specific fluorescence of Nile Red, a commonly used indicator of lipid accumulation, when cells were depleted of nitrate
compared to cells depleted of phosphate. In addition, re-supplementation of N or P promoted cellular growth, cessation of lipid accumulation, and increased lipid consumption in *P. tricornutum*.

**Carbon**

It is important to keep in mind that when comparing different nutrient-deprived states, carbon above all else is absolutely required for lipid biosynthesis (Palmqvist et al, 1988; Raven, 2010; Spalding, 2008). Without carbon, independent of nutrient deprivation, biomass or lipid biosynthesis is impossible. Therefore, the most successful reports of lipid induction techniques in microalgal lipid production typically involve elevated concentrations of inorganic carbon in tandem with N and/or P limitation (Gardner et al, 2011; Sharma et al., 2012; Gardner et al, 2012). These strategies often employ a CO₂ sparge to increase dissolved CO₂ above atmospheric concentrations, or chemical addition of soluble inorganic carbon during inoculation or just prior to nutrient depletion (Gardner et al, 2011; 2012). Providing large amounts of dissolved inorganic carbon via a CO₂ gas sparge can contribute significantly to the production cost in an algal biorefinery (e.g., Liu et al. 2013), and alternative methods to a gaseous CO₂-based inorganic carbon supply should be considered. Gardner et al. (2011, 2012) demonstrated that the dosage of fairly small amounts of bicarbonate, solely or in combination with a CO₂ sparge, can achieve similar algal growth and lipid production yields compared to continuous CO₂ sparging at a potentially substantially lower cost. In either case, elevated concentrations of C, combined with N or other nutrient deprivation has been shown to induce lipid accumulation in virtually every microalgal species tested. However, an improved understanding of cellular
and population responses to not only the respective levels but the ratios of macronutrients 
(e.g., C, N, and P) will improve resource utilization and promote efficient, cost-effective 
processes.

**Silicon Limitation**

Reports on silicon limitation have revealed that both marine and freshwater diatoms will accumulate lipid under Si-limiting conditions (Sharma et al, 2012). Diatoms possess immense potential as contributors to biodiesel production. When faced with Si-limitations, most diatoms appear to direct carbon storage towards lipid (Roessler, 1988), albeit the response is dependent on the degree of Si content in the cell wall. Diatoms incorporate biologically available Si as monomeric or dimeric silicic acid into silicious cell walls (frustules), and require approxiamtely 7% of the energy expenditure required for polysaccharide cell wall formation characteristic of green algae (Hildebrand et al 2012; Kroger et al 2008; Raven 1983). Diatoms produce comparatively less cellular starch, such that fixed carbon has increased potential to be allocated to lipid accumulation (Burrows et al 2012; Gardner et al 2011; Roessler 1988; Smith et al 2012). In fact, diatom cells can accumulate enough TAG to cause the frustules to break under silica deplete conditions (Hildebrand et al 2012), potentially reducing the need for energy intensive procedures associated with lipid extraction in green algae. Numerous studies have shown increased lipid accumulation when diatoms are cultured in silica deplete media (Lombardi and Wangersky 1991; McGinnis et al 1997; Obata et al 2013; Valenzuela et al 2012; Valenzuela et al 2013; Yu et al 2009). However, the majority of these studies were performed on marine diatoms (Taguchi et al 1987) (e.g., *Cylindrotheca spp.* (Roessler et
al 1988), Thalassiosira pseudonana, and Phaeodactylum tricornutum) (Yu et al 2009) grown in media containing comparatively lower silica concentrations. The results of Moll et al. (2014) indicate that increasing the silica concentration will increase cell numbers, which is vital for improving algal biodiesel productivity in terms of increased biomass. Therefore, while research on marine diatoms for use in biofuel applications may be advantageous for use in large-scale raceway ponds due to the ability to tolerate saline environments, the actual use may be limited until conditions are optimized for diatom cell growth and lipid accumulation.

While silica limitation is known to increase lipid accumulation, combined with other physiological stresses, lipid accumulation may be enhanced. A recent study investigated the effect of coincident silica and nitrate limitation and HCO$_3^-$ addition to promote lipid accumulation in a freshwater diatom. Moll et al. (2014) observed that combined silica and nitrate limitation, as well as sodium bicarbonate addition increased lipid accumulation compared to individual stressors with or without HCO$_3^-$. One hypothesis for this effect is the effect on the cell cycle. Olsen et al. (13) and Vaulot et al. (20) revealed that for Thalassiosira weisflogii and Hymenomonas carterae, nitrate and silica limitation resulted in halting the cell cycle at G1 and the G1/S and G2/M boundaries, respectively (Darley and Volcani 1969). It is possible that the two combined nutrient limitations at different periods within the cell cycle may contribute to cellular stress and ultimately lead to enhanced lipid accumulation in diatoms.
Iron Limitation

Approximately 30 to 40% of the world’s oceans are iron limited, and studies have investigated “iron fertilization” experiments whereby iron is added to High Nutrient Low Chlorophyll (HNLC) areas to induce phytoplankton growth and CO₂ fixation (Buesseler et al 2004). Iron-limited conditions are thought to alter cell physiology in the following ways: reduce cell volume, chlorophyll content, and photosynthetic activity. Additionally, the following enzymes were down regulated: β-carbonic anhydrase, phosphoribulokinase (PRK), two RuBisCO enzymes and decrease in expression of a HCO₃⁻ transporter, thus potentially resulting in decreased carbon fixation (Allen et al 2008). Iron limitation is known to induce chlorosis, as well as reduced carbon fixation rates, photosynthetic efficiency, and growth rates (Allen et al. 2008). Iron limitation has also been linked to increased rates of silicification, thus increasing cell density and cell sinking. According to Allen et al. (2008), cells grown under limited nitrate conditions fixed carbon at rates 14 times lower compared to cells grown in iron replete conditions (Allen et al 2008). Since iron limitation can result in detrimental physiological effects, it is pertinent to determine the potential for these processes to be useful for commercial scale lipid accumulation.

Biofilm Growth

One of the most significant limitations to the economical use of algae is the high cost of harvesting and concentrating the biomass (Johnson and Wen 2010, Christenson and Sims 2012; Ozkan et al 2012; Schnurr et al 2013). To date, research has been focused on microalgae in suspended phase for lipid production, and few studies have focused on the biofilm growth state. However, the biofilm growth state provides some advantages over
suspended growth systems in terms of biomass accumulation and maintenance that would be beneficial for harvesting and concentration of biomass prior to processing. Algal suspensions are often between 0.02% and 0.06% total suspended solids (TSS), and significant energy is spent to harvest and concentrate the cells to 5 to 25% TSS. Biofilms can range from 6 to 16% TSS (Schnurr et al. 2013), and could potentially minimize biomass-processing costs (Johnson and Wen 2010, Christenson and Sims 2012; Ozkan et al. 2012). In general, the available algal biofilm studies are based upon wastewater treatment, biofilm structure and development, and aquaculture applications (Johnson and Wen 2010; Christenson and Sims 2012; Patil and Anil 2005; Irving and Allen 2011; Avendaño-Herrera and Riquelme 2007). There is a small amount of research on biofilm systems for the production of biomass and lipids in eukaryotic, photoautotrophs, but very little in relation to the influence of environmental stresses (Schnurr et al., 2013; Bernstein et al., 2014).

Recently, Schnurr et al. (2013) reported biofilm growth under nutrient starvation to stimulate lipid accumulation. A semi-continuous flat plate parallel horizontal photobioreactor system (PBR) was designed to control the bulk medium nitrogen and silicon concentrations until nutrient depletion and biofilm onset. Wastewater was used to seed biofilm growth and later replaced by synthetic medium and pure cultures of Nitzchia palea and Scenedesmus obliquus. Well-attached, thick algal biofilms were observed in all experiments, until N and Si levels decreased to below detection limits, resulting in detachment from the substratum. In contrast to suspended algae, the algal biofilms did not accumulate more neutral lipids when exposed to nutrient deficient conditions in these
Similar results were reported by Bernstein et al. (2014) who observed little lipid accumulation in mixed culture wastewater biofilms on the field scale or in laboratory-scale algal biofilm reactors seeded with a *Botryococcus* sp. (strain WC2B). Based upon these results, there appears to be fundamental differences in the way suspended cultures and biofilm cultures respond to nutrient deprivation. It is possible that benthic microorganisms would prove more useful for biofilm growth modes, and we have recently grown a benthic diatom in biofilm reactors that could accumulate lipids (Fields, Whitney, and Valenzuela, unpublished results). These results suggest that the two growth modes can elicit different behaviors, but numerous research approaches and questions need to be explored to better understand the feasibility and cellular responses of microalgal biofilms for biomass and lipid accumulation.

**Ecological Effects**

The literature offers many examples of increased lipid production in numerous algal species cultivated as monocultures in closed photobioreactor (PBR) or open raceway systems under varying nutrient limitations. However, as demonstrated by mathematical models and field experiments, phytoplankton biodiversity can cause increased productivity (Tilman et al., 1981; Downing and Leibold, 2002; Striebel et al., 2009). Furthermore, in natural freshwater systems, productivity, measured as biomass, was highest when there was abundant nutrient availability (Interlandi and Kilham, 2001). These observations underlie the challenge of needing high biomass loads to maximize overall lipid production. Obviously, productivity can fall under the guise of several metrics ranging from biomass, cell number, chlorophyll/pigments, and more recently lipid accumulation. Despite the
success of increased lipid content in nutrient deprived monocultures, recent studies indicate that comparable lipid production can also be achieved in nutrient rich systems with a diverse community.

A diverse community is also more resistant to invasion from other species that could outcompete the desired algal species (Tilman 1977; Tilman et al., 1981). Higher nutrient availability may also aid in algal cultivation by making algae less susceptible to viral infection. Coupled with community diversity, the relative health of each species is an important component. “Healthy” algae (i.e., cells not under nutrient stress) can be more resistant to viral infection that leads to cell lysis (Murray, 1995). Rhodes and Martin (2010) developed a theoretical model that implicated high nutrient availability in significantly reduced viral infection, and such scenarios will be important to consider in microalgal cultivation processes.

A study by Stockenreiter et al. (2012) demonstrated increased lipid production in a naturally occurring algal community cultivated in open ponds compared to that of single species cultivated in PBRs (Stockenreiter et al., 2012). The total phosphorus (TP) of the natural system, ranging from 3 to 302 µg/L of total P, was maintained in batch culture. For freshwater systems, P is the key nutrient responsible for eutrophication and can greatly alter productivity when limited (Schindler 1977; Smith and Bennett 1999; Schindler et al., 2008; Stockenreiter et al., 2012). Based upon observations with PBRs with deplete nutrient availability, one would expect to see substantially higher lipid content in the oligotrophic communities than the eutrophic. However, the most oligotrophic lake in the study had the highest lipid content by an order of magnitude over the P replete community, 5.2 x10^6
pg/mL and $1.1 \times 10^5$ pg/mL, respectively (Stockenreiter et al., 2012). Although in no way conclusive, results such as these suggest comparable lipid values in nutrient replete and deplete systems and indicate the need to further investigate the relationship between nutrient type and abundance in the context of lipid production in mixed communities (Stockenreiter et al., 2013).

Thus, a diverse community of selected algae could be maintained in an open pond system (or closed system) and promote culture stability and resistance to community invasion (cells and/or viruses). Furthermore, open pond systems have been show to be more economically viable and have applications for wastewater treatment, although closed systems have advantages of higher productivity. Due to industrial processes and municipal waste, there is a large supply of N- and P-rich wastewater (Smith et al., 2009), and the recycling of these nutrients requires a diverse and stable community. Algae have long been shown to be very effective at removing excess nutrients and pollutants from wastewater aiding in the control of eutrophication in natural systems that can be linked to biomass/biomaterial production.

In contrast, despite the success of increased lipid accumulation in PBR monocultures under nutrient limiting conditions, economic assessments indicate that PBRs operating on a large scale may not be commercially viable (Smith et al., 2009). However, some have argued for hybrid systems that utilize a combination of both closed and open systems (Singh and Dhar, 2011), or modified PBRs such as solid-state reactors (Naumann et al., 2013). In addition, if ponds are not well mixed, biomass loss due to dark respiration may impact performance for some microalgae (Huesemann et al., 2013).
Therefore, life-cycle analyses should help direct research to identify complementarity between water footprint, regional light, process design, and the organism.

**Integrating Life-Cycle Analysis**

Algal biofuels have the potential to provide a substantial fraction of United States transportation fuel while imposing a relatively small (arable) land footprint (Sander and Murthy, 2010), and providing opportunities for reducing water and nutrient consumption relative to first generation biofuels (Clarens et al., 2010). The degree to which biology and engineering can contribute to these goals will; however, be a function of the entire lifecycle (Zaimes and Khanna, 2013). A circumscribed version of that lifecycle, one involving only the production cycle (distinct from the usage cycle), includes microbial growth, dewatering/drying, extraction/conversion, and energy/input recovery stages, and each stage involves a number of choices (Figure 3).

**Figure 3.** Primary stages and (alternative processes) in the microalgae to fuel production process.
With respect to benefits, LCA (life cycle analysis) has promoted system optimization by highlighting processing alternatives that produce a net increase in system performance, avoiding “burden shifting” that can be obscured when viewing the production system less holistically (Klöpffer, 1997). In addition, choices made in the other three production stages will have implications for the growth stage, with the technology selected for extraction/conversion having particular importance (Brennan and Owende, 2010). While each has respective strengths and weaknesses, two of the most critical distinctions from a life-cycle perspective are (a) the degree of pre-conversion drying required (Lardon et al., 2009) and (b) whether the conversion process involves all of the algal biomass or only the lipid fraction (Kirrolia et al., 2013).

The dependence of transesterification processes on algal lipid content can impose extra costs in the growth stage (Brennan and Owende, 2010), and lipid accumulation procedures typically come at the cost of algal productivity (Davis et al., 2011). As noted by Quinn et al. (2013) and Chowdhury et al. (2012), increasing lipid content can result in an increase in process GHG (greenhouse gas) emissions, because less residual biomass is used in a potential energy recovery stage. Thus, the grid energy requirement increases proportionally with the lipid fraction. Wet extraction transesterification processes, while significantly reducing the drying energy input, typically involve solvent-based extraction, leading to concerns over solvent disposal (Torres et al., 2013), and the recycling of these solvents is made challenging by the high volumes associated with a wet slurry (Rios et al., 2013). Simultaneous extraction and transesterification processes (i.e., “reactive extraction”) offer the potential for increased oil yields and lower process costs (Rawat et
al., 2012), but the effectiveness of this process at an industrial scale is still untested (Nagarajan et al., 2012).

Hydrothermal liquefaction (HTL) processes, despite greater capital expense, also reduce drying/dewatering requirements through the utilization of a wet feedstock (López Barreiro et al., 2013), while converting up to 60% of the total biomass into a useable fuel product (Liu et al., 2013). HTL can result in greater fuel yields than those achieved via transesterification (Frank et al., 2013), and this technology may reduce the importance of advanced culturing methods to enhance algal lipid accumulation for biofuel production (Elliot et al., 2013). However, thermochemical conversion methods such as HTL make nutrient recycling more challenging, as the aqueous byproduct is poorly suited for anaerobic digestion (López Barreiro et al., 2013). In addition, high nitrogen incorporation into the oil results in a substantially increased nutrient requirement in the production process (Liu et al., 2013).

Life cycle analysis has been, and continues to be, successfully utilized to identify optimal algal biofuel production pathways. Ongoing refinement and application of this analytical technique can lead to advances that will guide future research toward a better understanding of the implications of a number of important choices, and thereby promote the development of more cost-effective and environmentally benign biofuel production processes. LCA has successfully identified synergies and tradeoffs between the growth stage and other parts of the production process, and results suggest that parallel research efforts involving both experimental research and life-cycle modeling should be used (Collet et al., 2014).
Conclusion

With the re-invigorated interest in alternative fuels, microalgae provide one option that will likely contribute to an overall plan for biomass, biochemical, and biofuel production in a more sustainable and efficient manner. Given the typical ratio of C:N:P in microalgal biomass (C\textsubscript{106}: N\textsubscript{16}:P\textsubscript{1}), much of the research has focused on N and P (P to a lesser extent) and these two elements are linked in different ways to C through resource allocation at the cellular, population, and community levels. Micronutrients also play a role in cellular responses and activity, and Si and Fe need to be further studied with respect to C:N:P ratios and the allocation of C into desired compounds (e.g., lipids). Much of the nutrient-deprived states have been studied with monocultures (or nearly axenic) as suspended cultures, and regardless of the systems used (e.g., closed reactors vs. ponds), communities will assemble with different characteristics of stability, resiliency, and productivity. In addition, biofilms will likely develop, and may even be desired for the traits of accumulated biomass that can provide advantages for harvesting.

Moreover, while not directly covered in this mini-review, other resources/conditions will affect the cultivation of microalgae and include water, climate (e.g., light and temperature), land, and location (i.e., geography). Water will be essential for any biological process, and the re-cycle of water will be crucial as many parts of the globe become increasingly stressed for potable water. Light is obviously an important parameter for phototrophs, and is inherently related to temperature as the need for light energy and heat-regulation scale at different proportions. Land is an essential commodity whether bioreactors or ponds are used and should not compete with agricultural needs. The
location of growth and processing facilities are crucial aspects to be considered via LCA both for economic implications as well as the biology/ecology (e.g., biogeography) that can differ from region to region. Therefore, targeted science and engineering research is needed to better inform life-cycle analyses and process design to maximize productivity, efficiency, and cost-ratios.

Acknowledgments

The authors would like to thank all members of the MSU Algal Biofuels Group for helpful discussions. Support was provided by the United States Department of Energy-Advancements in Sustainable Algal Production (ASAP) program under contract # DE-EE0005993 and the National Science Foundation Sustainable Energy Pathways (SEP) Program under NSF CHE-1230632. We also would like to acknowledge student support from the Molecular Bioscience Program at Montana State University.
References


Redfield A (1958) The biological control of chemical factors in the environment. American Scientist 46:205–211


