AN ALGAL-BACTERIAL SYMBIOTIC SYSTEM OF FORMATE UTILIZATION FOR CO_2 CAPTURE AND UTILIZATION

By

Yurui Zheng

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Biosystems Engineering – Doctor of Philosophy

2023

ABSTRACT

Microalgae is one type of photosynthetic microorganism which can utilize carbon dioxide (CO_2) through photosynthesis and convert exogenous carbon into microalgal biomass. Microalgal cultivation is considered as one of promising ways for carbon capture to reduce greenhouse gas effects. Formate is known as a single-carbon chemical, which possesses good solubility and stability under a wide range of pH. This work represents a comprehensive examination of microalgal culture of the green microalgae *Chlorella sorokiniana* by utilizing formate and effects of formate on this symbiotic algal-bacterial system.

In Chapter 2, an algal-bacterial symbiotic system was explored to investigate utilization of formate as a carbon source. The algal-bacterial assemblage, after a 400-day adaptive evolution using the formate medium, has demonstrated a new capability to assimilate both formate and CO_2 to promote biomass production. ^{13}C isotope tracing and microbial community analysis were conducted to indicate a uniquely evolved formate utilizing culture. This study demonstrates a new route of using electrochemical-derived formate to support mutualistic algae-bacteria biorefinery while the formate as an alternative carbon source could repel pests for outdoor algal cultivations.

In Chapter 3, cultivation parameters including light intensity, formate feeding rate and a novel approach of alternating carbon feeding were tested for enhancing microalgal biomass and productivity during the culture. Effects of formate on microbial communities and algal assemblage were reported. The results showed that formate was a good carbon source for microalgal cultivation, and the highest biomass concentration of 1.4 g/L was achieved during the culture. Microbial community analysis revealed a stable microalgal-bacteria consortia under the formate feeding system. Microalgal biomass was further increased to 1.6 g/L compared to the

formate feeding method with the alternating carbon feeding method.

In Chapter 4, effects of formate as an additional carbon source for microalgal culture with pumping flue gas within a 100 L photobioreactor (PBR) were investigated. Results showed that formate addition group exhibited better carbon capture efficiency than group without formate addition. Mass and energy balance analysis showed that formate group required 20% less energy consumption and showed nearly 33% higher biomass yield on average compared to control group.

In summary, this work presents a symbiotic algal-bacterial system of utilizing formate. The work establishes a stable microalgal cultivation method with formate feeding in both bench-scale and pilot-scale PBRs. Notably, this work advances carbon capture efficiency in microalgal cultivation field, as well as innovative methods and techniques that elucidate the viability of formate utilization in microalgal cultivation for carbon capture.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my PhD advisor, Dr. Wei Liao, for his unwavering support and guidance during my entire PhD study. His expertise and patience have always been invaluable for me and led me to find the right path in my academic journey. I have learned so much from him not only about my research, but also ways of thinking, attitudes towards difficulties in my life. I feel truly grateful for the opportunity to have worked under his supervision.

I greatly appreciate my dissertation committee members Drs. Yan Liu, Yinjie Tang, and R. Jan Stevenson for their support and feedbacks to my dissertation. I also would like to thank Liao lab members for their kind support and help during my PhD study. I have learned a lot from working with them.

Finally, I would like to express my appreciation to the financial support from the funding agency U. S Department of Energy (DOE).

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

CO_2 : Carbon dioxide

ANOVA: Analysis of variance

PBR : Photobioreactor

APB: Algae photobioreactor

BECCS: Bioenergy with carbon capture and storage

TAP: Tris-Acetate-Phosphate

TN: Total nitrogen

TP: Total phosphorous

HPLC: High-performance liquid chromatography

LED: Light-emitting diode

PCR: Polymerase chain reaction

DNA: Deoxyribonucleic acid

RNA: Ribonucleic Acid

CHAPTER 1: INTRODUCTION

1. Problem statement

Global warming is one of major issues that human activities created in recent years. The direct reason of global warming is caused by the large emission of greenhouse gas [1]. The deterioration of the natural environment caused by greenhouse gas emissions is approaching a critical point and presents a very serious challenge to humankind. Greenhouse gas emissions, particularly CO_2, must be controlled and reduced in a fast and efficient manner. Reduction of CO_2 emission and capture of emitted CO_2 are the ways to address this challenge.

Gaseous mass transfer largely restricts the growth of microalgae [2, 3], which are photosynthetic organisms capable of capturing CO_2 [4] and converting it into macromolecules of protein, carbohydrate, and lipid [5].These macromolecules have the potential to be used to produce a variety of value-added commodities, such as animal feedstock, pharmaceutical compounds, cosmetic areas, and health care products [6]. To address the issue of CO_2 transfer enhancement in culture media, bicarbonate has been used, but its applications are limited due to requirements of high alkalinity and fine pH control [7]. Formate, a liquid product from CO_2 reduction, is a better option for microalgal cultivation as it is stable under a wide range of pH and can support the growth of microalgal assemblage as a carbon source. Moreover, formate can address the challenge of algal predators for large-scale cultivations [8]. Formic acid, which is toxic to many insects [9], can decrease protozoa population [10]. A high level of formate (greater than 1,000 mg/L) during microalgal-bacterial cultivation can serve as a contamination control strategy to repel insects and protozoa in non-sterilized environments and achieve long-term culture stability. Formate requires relatively less energy [11, 12] compared to other reduction products of CO_2 and is a soluble compound that can bypass the issues of gas-liquid mass transfer limitation and pH requirement. In addition, formate also plays an important metabolic role [13,

14], such as a compound to promote NADH generation through formate dehydrogenase ($HCOO^- \rightarrow CO_2 + H^+ + 2e^-$) [15, 16].

The use of microalgal-bacterial communities for phototrophic applications, including wastewater treatment [17], biomass production [18], and lipid accumulation [19], has been extensively investigated. In these systems, microalgae provide dissolved oxygen and carbohydrates to support bacterial growth, while the bacteria generate metabolites to promote microalgae growth [20]and maintain the stability of the cultivation by preventing invasions from other organisms [21]. This symbiotic strategy is especially important for large-scale operations where culture media cannot be easily sterilized, such as in open pond cultivation, and where the culture conditions are largely influenced by environmental factors [22]. Microalgal-bacteria assemblage offers stability for long-term microalgal cultivation in carbon capture. By having such symbiotic relationship between algae and bacteria, the system is resistant to external invading species and exhibits resilience to outer pressure, which is beneficial for large-scale application of microalgal culture for carbon capture. Therefore, coculturing microalgal-bacteria assemblage could be an excellent way for carbon capture.

Therefore, studying an algal-bacterial symbiotic system to utilize formate as a carbon source would address the identified knowledge and research gaps in microalgal cultivation for carbon capture. However, microalgal-bacterial communities that use formate as a carbon source have not yet been reported.

2. Literature review

2.1 Carbon capture technologies

Global climate change which was caused by greenhouse gas emission has been intensively reported since 21st century [23]. Carbon capture is considered as a necessary way to effectively control the greenhouse gas emissions in the atmosphere. CO_2 is usually the major

target gas since it has the largest ratio in composition of greenhouse gas [24]. Technologies for carbon capture and storage have been investigated by many researchers. Among these researches and findings, carbon capture technologies can be broadly into three types: chemical, physical and biological.

Chemical carbon capture technologies involve chemicals in different ways. Absorption is one of the major chemical ways to capture CO_2. Absorption normally involves using a liquid solvent to absorb carbon dioxide from a flue gas stream. The solvent is then separated from the CO_2, allowing for the capture of the CO_2 for storage or use. Chemicals such as sodium hydroxide, monoethanolamine (MEA), piperazine, potassium carbonate, are used as solvents to react with CO_2 directly to achieve the carbon capture [25]. Another example of chemical carbon capture is calcium looping, which involves the use of calcium-based sorbents to capture carbon dioxide. In this process, calcium oxide reacts with carbon dioxide to form calcium carbonate, which can be further processed to release the captured carbon dioxide. Chemical carbon capture technologies have been developed and tested at a commercial scale, but they can be energy-intensive and costly to operate [26].

Physical carbon capture technologies: Compared to chemical ways, physical carbon capture technologies focus on separating or detaining CO_2 without chemical reactions due to CO_2 physical attributes. Common physical carbon capture technologies include adsorption and membrane separation [27]. Adsorption refers to using a solid material, such as activated carbon, zeolites, or metal-organic frameworks, to capture CO_2 from a gas stream [28]. The CO_2 is then desorbed from the solid material, allowing for the capture and storage or utilization of the CO_2. Membrane separation means using a membrane to selectively allow carbon dioxide to pass through while blocking other gases [29]. This process is known as gas permeation. Due to

various size and solubility of gases, the corresponding diffusion rate can be different. This allows the separation of multiple gas components. Different types of membranes can be used, such as polymeric, ceramic, or metallic membranes [30].

Biological carbon capture technologies: Biological carbon capture technologies use natural or engineered biological systems to capture and store carbon dioxide. There are two main approaches to biological carbon capture: biological sequestration and bioenergy with carbon capture and storage (BECCS) [31]. Biological sequestration involves using natural or engineered ecosystems to capture and store carbon dioxide from the atmosphere . This can be done through various methods such as afforestation [32], reforestation [33], and ocean fertilization [34]. Afforestation involves planting trees on land that was previously not forested, while reforestation involves restoring degraded forests. Both methods increase the amount of carbon dioxide absorbed by the plants through photosynthesis. Ocean fertilization involves adding nutrients to the ocean to stimulate the growth of phytoplankton, which absorb carbon dioxide through photosynthesis [34]. BECCS is a technology that combines bioenergy production with carbon capture and storage [35]. BECCS involves using biomass (e.g., plants, crop residues, or animal waste) to generate energy through combustion or other processes. The carbon dioxide emitted during the combustion process is then captured and stored in geological formations or other suitable storage sites. BECCS has the potential to reduce greenhouse gas emissions by providing a source of renewable energy and removing carbon dioxide from the atmosphere. Microalgae cultivation is one of the promising BECCS strategies for carbon capture. Compared to terrestrial plants, microalgal cells have higher photosynthetic efficiency [36].

2.2 Microalgal cultivation

Microalgal cultivation is mainly influenced by reactor types, nutrient availability, carbon

source and environmental factors. As for reactor, there are two main types of culture system: open system and closed system [37]. An open pond is a typical open system that involves growing microalgae in an outdoor pond or shallow pool. Open ponds are typically made of concrete or lined with synthetic material to prevent leakage and are exposed to sunlight to promote photosynthesis in the microalgae. Nutrient-rich water is continuously pumped into the pond and circulated to maintain optimal conditions for microalgal growth [38]. Open ponds are relatively low-cost and easy to operate, but they can be susceptible to contamination from environmental factors such as wind, rain, and insects [38]. Additionally, open ponds may have lower productivity compared to other reactor types due to limitations in their ability to regulate temperature and light exposure [39]. Raceway ponds are an improvement over open ponds, offering higher productivity and reduced contamination risk [40]. The major advantages of open pond culture system are minimal capital and operating cost, and a low energy requirement [22]. However, open pond systems need a large land area to build and are susceptible to contamination and weather conditions.

Photobioreactors (PBRs) are closed systems that use artificial light sources to grow microalgae. Common PBR types include tubular PBRs, flat-panel PBRs, and stirred-tank PBRs [41]. Tubular PBRs are cylindrical tubes made of transparent materials that are used to grow microalgae for carbon capture. Tubular PBRs consist of a long, transparent tube coiled around a frame and a pumping system that circulates the growth medium and the microalgae. The tubes are usually made of glass or plastic and can range in diameter from a few centimeters to several meters. Flat-panel photobioreactors (FPPBR) consist of flat panels made of transparent materials, such as glass or plastic, which are placed parallel to each other to form a channel. The panels are typically thin and have a large surface area-to-volume ratio, allowing for efficient light

penetration and gas exchange [42]. The microalgae culture is circulated through the channel using a pump or other means of mixing, while aeration is provided through a sparging system. The panels are often arranged in a modular fashion, allowing for easy scaling up or down of the system. Another example is stirred-tank PBR. These reactors consist of a cylindrical vessel with an impeller or stirrer that provides agitation to ensure proper mixing of nutrients, light, and microalgae. The impeller is driven by a motor and is designed to maintain a homogenous suspension of microalgae cells in the culture medium. The reactor is equipped with a light source, which provides optimal light conditions for photosynthesis. The temperature, pH, and dissolved oxygen levels are also closely controlled to ensure the growth of microalgae. Stirred-tank PBRs are widely used for large-scale microalgae cultivation because of their high productivity and scalability [43]. Compared to open ponds, PBRs are more efficient and stable due to highly controlled operational conditions. PBRs require less space than open systems and have a more stable environment for microalgal growth regardless of weather conditions [44]. Nevertheless, the high capital cost is the major obstacle preventing PBRs from scaling up and commercialization.

Nutrient components in the culture medium are also important for microalgal cultivation. Carbon, nitrogen, and phosphorous are major elements and other nutrients such as magnesium (Mg), calcium (Ca), Sulfur (S), Iron (Fe) are minor elements in the medium. Carbon is the main microalgal biomass element, inorganic carbon is fixed inside the microalgal cells by the Calvin cycle. In this study, carbon is introduced into the system as formate. Nitrogen is the second most abundant element in microalgal biomass. Nitrogen is an essential chemical compound present in DNA, RNA, proteins, and pigments in microalgal cells. The glutamine synthetase enzyme system is the main pathway for metabolizing nitrogen in microalgal cells [45]. Phosphorous is

another important nutrient for microalgal growth, which is a fundamental material for DNA, RNA, and ATP in cells. Other micronutrients (Mg, Ca, S, and Fe) are also imperative for microalgal growth, which are involved in photosynthesis, respiration, and cell division.

Carbon source is an essential factor for microalgal photosynthesis and the production of organic matter. Carbon sources in microalgal cultivation can be divided into two categories: organic and inorganic. In microalgal culture, organic carbon is added such as glucose, fructose, sucrose, acetate, and so forth [46-48]. It has been reported that the addition of organic carbon source can promote the growth rate and lipid production of microalgae [49]. The utilization of organic carbon source help microalgal culture become one practical way for wastewater treatment [50]. Compared to traditional fertilizer, organic carbon sources are often less expensive than traditional chemical fertilizers, making them a cost-effective option for microalgal culture [51]. Inorganic carbon sources in microalgal cultivation include CO_2, bicarbonate, and carbonate. Photosynthesis in microalgal cells can directly uptake CO_2 from atmospheric environment to produce organic compounds. Studies showed that bicarbonate addition can boost microalgal growth and carbon capture efficiency from soybean wastewater [52]. Both production cost and energy inputs can be reduced through utilizing bicarbonate for microalgal cultivation [53]. Compared to carbon sources discussed above, utility of formate as a carbon source still remain unclear.

Environment is another decisive factor for microalgal cultivation, which include temperature, light intensity, pH stability, harvesting rate/amount, contamination control, and so forth. Previous research indicated that temperature can have significant influence on microalgal growth [54]. Methods reported to controlling pH in microalgal cultivation such as adding acid/base, making stoichiometrically-balanced medium [55] are commonly used in microalgal

culture. Harvesting rate/amount is another important parameters for successful microalgal cultivation because they play an important role in balancing biomass productivity and biomass concentration during a continuous run. A suitable harvesting amount for cultivation is essential for maintaining cultural stability. Biological contamination in microalgae cultivation is hard to avoid, which makes contamination control very crucial. Common methods include adding chemical pesticides and physical filtration [56]. However, these methods have either high cost or low effect. A resilient microalgal-bacterial system is a more promising way for future microalgal cultivation applications.

2.3 Microalgal-bacterial symbiotic system

The microalgal-bacterial symbiotic system refers to a type of biological system in which microalgae and bacteria form a symbiotic relationship. In this relationship, the microalgae provide the bacteria with organic compounds produced through photosynthesis, and the bacteria provide the microalgae with essential nutrients and other substances that the microalgae need to grow and survive. Many studies show applications of microalgal-bacterial in wastewater treatment [57], biomass production [58], and lipid accumulation [59]. Heterotrophic metabolism of aerobic bacteria and algal capabilities of nutrient assimilation and photosynthetic oxygen generation can be mutually symbiotic if the growth of two different microbial communities are compatible. Photoautotrophic microalgae need nitrogen and phosphorous to consume and transform CO_2 or dissolved carbon (such as bicarbonate) into their biomass. Through photosynthesis, O_2 is released which can be utilized by aerobic bacteria as an electron acceptor [26]. Meanwhile, aerobic bacteria also produce more CO_2 in this system. Compared to growing microalgae alone, the growth rate of microalgae can be increased by 10%-70% and subsequently higher productivity can be achieved when co-cultivating microalgae and growth-promoting

bacteria [60]. An engineered bacterial consortium showed a significant enhancement of the microalgal biomass and lipid productivity through carbon exchange upon co-cultivation of *Chlorella vulgaris* with four different growth-enhancing bacteria [61]. Aerobic bacteria co-cultured with microalgae can modify the microalgal environment by consumption of excessive dissolved O2 to lower the net photosynthetic carbon fixation by favoring Rubisco activity, thereby creating a more favorable condition for microalgal growth [62]. These interactions between microalgae and bacteria enable the co-culture system to share the metabolites and endure nutrient limitation, and resist the invasion of other species.

However, some challenges need to be overcome to scale up microalgal-bacterial symbiotic systems for commercial use. The first one is culture stability. Maintaining stability in the microalgal-bacterial symbiotic system can be challenging, as changes in environmental conditions can alter the balance between the microalgae and bacteria. For example, changes in temperature, light, or pH can have a negative impact on the growth and survival of either the microalgae or bacteria. Another obstacle is to find the right combination of microalgae and bacteria that will form a symbiotic relationship. In some cases, bacteria can compete with the microalgae for nutrients, or they can produce toxic byproducts that harm the microalgae. Careful selection of the microalgae and bacteria is therefore essential for the success of a microalgal-bacterial symbiotic system. Overall, microalgal-bacterial symbiotic systems show great promise as a sustainable and renewable source of energy and bioproducts. By combining the strengths of both microalgae and bacteria, it is possible to achieve high levels of productivity and efficiency. However, for further scaling up, challenges such as stability in long-term culture, compatibility in strain selection, need to be solved.

2.4 Biological utilization of formate

Formate is a simple organic acid that can be utilized by various microorganisms for energy and carbon. Biological utilization of formate occurs through a process called formate oxidation, which involves the transfer of electrons from formate to an electron acceptor, such as oxygen or nitrate, to generate energy [63]. Many bacteria and archaea are capable of utilizing formate as an energy source [64]. One well-known example is the bacterium *Escherichia coli*, which can use formate as an alternative carbon and energy source when glucose is limited [65]. *Escherichia coli* and other formate-utilizing bacteria contain formate dehydrogenase, an enzyme that catalyzes the conversion of formate to carbon dioxide and generates electrons that can be used for energy production [66]. In addition to bacteria and archaea, some methanogenic archaea are capable of utilizing formate as a substrate for methanogenesis. During this process, formate is converted to methane and carbon dioxide through a series of biochemical reactions catalyzed by enzymes such as formate dehydrogenase and formylmethanofuran dehydrogenase [67]. Formate has also been studied as a potential substrate for microbial electrochemical systems, which use microorganisms to catalyze the transfer of electrons from an organic substrate to an electrode [68]. Formate can be used as a substrate for both anode-respiring bacteria [69], which transfer electrons to the anode, and cathode-respiring microorganisms [70], which receive electrons from the cathode.

Biological utilization of formate is a diverse and important process that plays a role in many microbial metabolic pathways and has potential applications in various fields, including biotechnology and environmental science.

2.5 Formate production

The commercial processes of formic acid production mainly are methyl formate

hydrolysis and oxidation of alkanes [71]. Other methods like electrochemical production [72] and biosynthesis were also reported [73]. Because of the low-cost and large-scale availability of formic acid by carbonylation of methanol and hydrolysis of the resulting methyl formate, formate is usually prepared by neutralizing formic acid with a base like sodium hydroxide.

Formate can be produced through a variety of methods, including electrochemical synthesis, biological processes, and the conversion from carbon dioxide [74-76]. In general, chemical synthesis is the most cost-effective method for producing formate on a large scale, as it leverages the economies of scale and the availability of large quantities of cheap raw materials, such as methanol [77]. However, the production of formate through chemical synthesis can also result in the generation of waste products and greenhouse gas emissions, which can increase the overall cost of production if the proper measures are not taken to manage these environmental impacts. In contrast, newer methods for producing formate, such as electrochemical conversion and photosynthetic production, offer the potential to produce formate in a more sustainable and environmentally friendly manner [78], but they are typically more expensive than chemical synthesis due to the higher cost of the technology and the need for specialized equipment and expertise. In order to reduce the cost of formate production and make it more economically viable, ongoing research and development efforts are needed to optimize the efficiency and scalability of the production methods, and to develop new technologies that can lower the cost and environmental impact of formate production. To encourage research of formate production, an important large scale use of formate is needed.

3. Research goals and objectives

The primary aim of my dissertation was to collect cultivation data and perform analyses that would validate the viability of a symbiotic microalgal-bacteria system of formate utilization

and advance carbon capture research. Three sub-tasks were designed for accomplishing the primary aim: (1) Developing an algal-bacterial symbiotic system of carbon fixation using formate as a carbon source; (2) Optimizing the operation of algae photobioreactors for biomass productivity and biomass concentration; (3) Using the established microalgal-bacterial system to study the effects of formate on continuous algal cultivation on a pilot scale. A secondary goal of this study was to investigate current limitations in microalgal cultivation and opportunities for future applications and research.

CHAPTER 2: AN ALGAL-BACTERIAL SYMBIOTIC SYSTEM FOR CARBON FIXATION USING FORMATE AS A CARBON SOURCE

1. Introduction

Photosynthetic organisms such as microalgae can capture CO_2 [4] and convert it into macromolecules of protein, carbohydrate, and lipid. These macromolecules can be used directly or indirectly to produce value-added commodities, such as animal feedstock, pharmaceutical compounds, cosmetics, and health care products [6]. While microalgae growth is largely restricted by gaseous mass transfer [2, 3], bicarbonate has been used as a vehicle for enhancing CO_2 transfer in culture media to address this issue, though, requirements of high alkalinity and fine pH control limit its applications [7]. Therefore, this study aimed to develop an algal-bacterial symbiotic system to utilize formate, a liquid product from CO_2 reduction that is stable under a wide range of pH, as a carbon source to support the growth of the algal-bacterial assemblage.

Formate can be produced by electrocatalysis of CO_2 [79, 80]. Compared to other catalytic processes, electrocatalysis is more technically sound and economically feasible to reduce CO_2 to formate for algae cultivation since the reaction takes place at ambient conditions [81]. In addition, the use of formate could address another large-scale cultivation challenge of algal predators[8]. Formic acid is a well-known toxicant against many microbiota, protozoa, and insects[9, 10]. For example, a study showed that protozoa population in rumen fluid decreased with formic acid supplementation [10]. It has also been reported that freshwater organisms and marine crustaceans were adversely affected by formic acid at concentrations ranging between 111 – 400 mg/L [82]. The presence of formate at a relatively high level (e.g., greater than 1,000 mg/L) during the algal-bacterial cultivation may serve as a control strategy for biological contamination to repel protozoa, insects, and other species in non-sterilized environments (e.g., open-pond cultivation) and enable long-term culture stability.

Many investigations have explored algal-bacterial communities for stable and effective phototrophic applications, including wastewater treatment [83], biomass production [18], and lipid accumulation [19]. In these algal-bacterial systems, algae provide photosynthetic products like dissolved oxygen and carbohydrates for bacteria growth. Meanwhile, the bacteria typically promote algae growth through the provision of some metabolites [20], and also keep the whole environment stable by preventing invading organisms [21]. The algal-bacterial symbiotic strategy is quite useful especially for large-scale operations where the culture media cannot be easily sterilized and the culture conditions are largely influenced by other environmental factors (such as open pond cultivation)[22]. Meanwhile, alga-bacterial communities as mixotrophic systems can utilize a wide variety of inorganic and organic carbon and nitrogen sources, such as bicarbonate, carbonate, glucose, glycerol, acetate, nitrate, and ammonia [84, 85] [86, 87]. However, algal-bacterial communities using formate as a carbon source have not been reported to date. Compared to other reduction products of CO_2 (e.g., acetic acid, CO, methane, and methanol), formate requires relatively less energy [11, 12] and is a dense, stable, and soluble compound that can bypass the issues of gas-liquid mass transfer limitation and pH requirement. In addition, formate also plays an important metabolic role [13, 14], such as a compound to promote nicotinamide adenine dinucleotide hydrogen (NADH) generation through formate dehydrogenase ($HCOO^- \rightarrow CO_2 + H^+ + 2e^-$) [15, 16].

The objective of this study is to explore an algal-bacterial symbiotic assemblage to utilize formate as a carbon source to accumulate microbial biomass. After a long-term culture, a symbiotic microalgal-bacteria assemblage was obtained which can take formate as carbon source. Compared to initial algal seed, this assemblage had higher relative abundances of bacteria, and showed higher absolute abundances of both microalgae and bacteria. We expect

that formate utilization bacteria in the assemblage oxidize formate into localized CO_2 to improve carbon assimilation by algae while the algal growth releases oxygen and other metabolites to support bacterial growth. Batch and continuous cultivations along with isotopic tracing, proteomics, and amplicon sequencing were carried out to analyze consortia metabolism and population interactions during the formate utilization.

2. Materials and methods

2.1 Algal assemblage and cultivation system

The algal assemblage containing a selected microalga *Chlorella sorokiniana* MSU from the Great Lakes region and several bacteria (mainly *Bacteroidetes* and *Proteobacteria*) [88] was continuously cultured in flasks on tris-acetate-phosphate (TAP) medium [89] at room temperature under constant fluorescent light to use for seeding the algae photobioreactors (APBs). Modified liquid TAP medium (without acetic acid and tris base) was used for microalgal cultures, which contains 7.5 mmol L^{-1} of NH_4Cl, 0.34 mmol L^{-1} of $CaCl_2 \cdot 2H_2O$, 0.4 mmol L^{-1} of $MgSO_4 \cdot 7H_2O$, 0.68 mmol L^{-1} of K_2HPO_4 (anhydrous), 0.45 mmol L^{-1} of KH_2PO_4 (anhydrous), 0.09 mmol L^{-1} $FeCl_3 \cdot 6H_2O$, and 1ml TAP trace elements solution. The modified TAP medium was unsterilized. The microbial community was analyzed before seeding the photobioreactors.

2.2 Photobioreactors

Both lab-scale and pilot-scale APBs were used. Lab-scale APBs were modified based on 10 L Eppendorf BioFlo®/CelliGen® 115 Benchtop fermenters with a working volume of 7.5 L (Fig. S2.1 a). Metal shells with adjustable light-emitting diode (LED) light strips installed inside were placed around the fermenters. The lab-scale APBs were used for kinetic study and semi-continuous microalgal cultivation with formate utilization.

The pilot-scale APB was located at the T.B. Simon Power Plant at Michigan State University. The effective volume of the pilot-scale APB is 100 L (Fig. S2.1b). The pilot-scale APB configuration and operating mechanism were described in a previous study [88]. The control of continuous culture on saturated CO_2 was carried out using the pilot-scale APB.

2.3 Batch cultivation using ^{13}C labeled formate and ^{13}C labeled bicarbonate

^{13}C labeling of culture metabolites was performed to trace carbon fates using batch cultivation. Cultivation with an inoculum of 0.1g/L of the seed was applied in two lab-scale APBs containing 4 L modified liquid TAP medium. The temperature was controlled at 22°C under a continuous light of either 50 or 500 μmol/m^2/s. The APB was mixed by a mechanical agitation at 250 rpm. pH was maintained between 6.5-7.2 by the automatic addition of sulfuric acid (5% vol/vol). Cultures were pulsed with concentrated ^{13}C-formate or ^{13}C-bicarbonate as the carbon source during early growth stage. At specified time points after the pulse (1 min, 20 min, 4 h, 8 h, and 24 h), 20 mL of culture medium was sampled and quenched with an equal volume of medium salts (containing no carbon or nitrogen sources) in a liquid nitrogen bath. Samples were then pelleted at 4°C, the supernatant discarded, and stored at -80°C until further analysis. Additional 50 mL algal samples at each timepoint were collected for analysis of nutrient concentration, biomass concentration, and carbon utilization.

2.4 Semi-continuous algal cultivation on formate and CO_2

Semi-continuous cultivations were carried out to compare algal growth and biomass accumulation on formate and CO_2. Lab-scale APBs were used to run semi-continuous algal cultivation on formate. The APB contained a 4 L modified liquid TAP medium with an initial algal biomass concentration of 0.35 g/L. The light intensity was maintained at 180 μmol/m^2/s for the entire cultivation. The pH was maintained between 6.5-7.2 by automatic addition of sulfuric

acid (5% vol/vol). The temperature was kept at 22±2°C. The APB was mixed by a mechanical agitation at 250 rpm. The culture was initiated as a batch culture for 48 hours. 1 g/L of formate was fed to the APB every 24 hours in the first 48 hours. After the initial biomass concentration reached 0.7 g/L at the end of the 48-hour batch culture, the semi-continuous cultivation started with a daily formate feeding rate of 1 g/L/day (the formate was added to the reactors once per day) and a daily harvesting ratio of 30% (v/v). The experiment was continuously run for 14 days.

The semi-continuous cultivation on CO_2 was run in the pilot-scale APB. The light intensity for the cultivation was 407 μmol/m^2/s. The natural gas-fired flue gas, containing 7.2 v/v of CO_2 was directly pumped from the stack into the APB at a flow rate of 120 L/m^3/min to provide CO_2 (2647.5 g/day) to the culture medium. The modified liquid TAP medium was used as the nutrients. The pilot-scale APB has been continuously running for 33 months on the flue gas as the CO_2 source. The data for the comparison were from 20-day continuous cultivation with the same nutrient condition and same harvesting ratio of 30% (v/v).

2.5 Chemical analysis

Samples were analyzed for dry biomass weight, pH, and nutrient (total nitrogen (TN), total phosphorus (TP), nitrate (NO_3-N), and ammonia (NH_3-N)) concentrations. Algal biomass was pelleted for dry weight measurement using a Thermo Electron Corporation IEC Centra CL2 Centrifuge at 3800 rpm for 5 minutes. Biomass was washed once and resuspended using deionized water, and then dried at 105°C for 24 hours. Sample pH was measured using a pH meter (Fisherbrand™ accumet™ AB15 + Basic, Fisher Scientific Co., Pittsburgh, PA). Nutrient concentrations were tested in the liquid supernatant using nutrient test kits (HACH Company, Loveland, Colorado) equivalent to the Environmental Protection Agency (EPA) methods (hach.com/epa). Algal biomass composition was analyzed using the standard forage analysis

method [90]. Standard forage analysis is a common method used to determine the composition and nutritive value of plant materials, including biomass. The analysis typically involves several different tests that measure the concentration of various components in the plant material, such as protein, fiber, fat, and ash.

Formate concentration of algal samples in kinetic study was determined by high-performance liquid chromatography (HPLC) (Shimadzu Corp., Kyoto, Japan) equipped with an analytical column (Aminex HPX-87H, Bio-Rad Laboratories, Inc., Hercules, CA) and a refractive index detector (Shimadzu Corp., Kyoto, Japan). The mobile phase was 0.005 mol/L sulfuric acid at a flow rate of 0.6 mL/min. The oven temperature was set at 65 °C. The bicarbonate concentration of algal samples in the kinetic study was determined by the alkalinity test kit (HACH Company, Loveland, CO).

2.6 Isotopomer analysis

^{13}C carbon incorporation into biomass proteins was quantified via gas chromatography-mass spectrometry (GC-MS) analysis of proteinogenic amino acids [91]. Briefly, the pelleted biomass was hydrolyzed with 6 N HCL at 100°C for approximately 20 h. The supernatant was transferred to a new vial and dried with air for 12 h. The amino acids were then derivatized with N-(tert-butyldimethylsilyl)-N-methyltrifluoroacetamide) at 70 °C for 1 hour and analyzed on an Agilent GC (7820A)-MS (MS5970E) equipped with HP-5ms column and a temperature gradient previously described [91].

The labeling in fast turnover-free metabolites was also measured as previously described [92]. In brief, biomass was centrifuged at ~0°C and the metabolites were extracted from the biomass pellet using a cold methanol-chloroform solution. The aqueous phase was collected and diluted with liquid chromatography-mass spectrometry (LC-MS) grade water. Then the samples

were frozen, lyophilized, and reconstituted in 200 μL of 60:30:10 acetonitrile:methanol:water. Labeling was analyzed using a hydrophobic interaction liquid chromatography (HILIC) method on a Shimadzu Prominence-xR ultra-fast liquid chromatography (UFLC) system and a SCIEX hybrid triple quadrupole-linear ion trap mass spectrometry equipped with Turbo VTM electrospray ionization (ESI) source.

2.7 Microbial community analysis

Samples (1 mL) collected for DNA analysis were kept frozen at -20°C until analysis. To remove nutrient media, the algae sample was centrifuged using an Eppendorf 5416R centrifuge at 10,000 rpm for 5 min and the supernatant was discarded. The remaining pellet was used for DNA extraction using the DNeasy® PowerSoil® Kit (Qiagen, Germany). DNA was eluted with 100 μL of 10 mM Tris-HCl (pH 8.5) and the concentration and purity determined using a NanoDrop Lite spectrophotometer (Thermo Fisher Scientific, USA). The DNA extracts were stored at -80°C for several weeks and then used for polymerase chain reaction (PCR) and Illumina DNA sequencing.

Illumina sequencing was performed for the 16S rRNA gene to assess the bacterial community. Prior to PCR, extracted DNA samples were diluted 10x due to high DNA concentrations. The PCR conditions were as follows: 1.0 μL DNA template (10x diluted), 0.5 μL of 100 μM forward primer (IDT, Pro341F), 0.5 μL of 100 μM reverse primer (IDT, Pro805R), 12.5 μL 2x Supermix (Invitrogen, USA), and 10.5 μL PCR grade water. The PCR program used for all assays is as follows: 96°C for 2 min, followed by 30 cycles of 95°C for 20 s, 52°C for 30 s, and 72°C for 1 min, and a final elongation period of 72°C for 10 min. After PCR, samples were diluted to normalize DNA concentrations within a range of 5-10 ng/μL. DNA concentration was determined using the PicoGreen® dsDNA quantitation assay (Invitrogen, USA) and Fluostar

Optima microplate reader (BMG Labtech, Germany). The PicoGreen® conditions were as follows: 95 μL 1x TE buffer solution, 100 μL 1:200 diluted PicoGreen® reagent, 5 μL DNA template. Samples with known DNA concentrations were also prepared for standard curve generation. Illumina library preparation and sequencing were performed at the Michigan State University Genomics Laboratory, East Lansing, USA.

The 16S rRNA gene sequencing was also used to determine *C. sorokiniana* in the assemblage [88]. It has been reported that *Cyanobacteria* have 85-93% of 16S rRNA gene sequences similar to *C. sorokiniana* [93, 94], while, color, shape, and size of both species are very different [95]. Therefore, after microscopic imaging verification of each sample, the *Cyanobacteria* sequence was interpreted as microalga *C. sorokiniana* for all samples.

2.8 Proteomic analysis

Samples from cultures using formate and bicarbonate under low and high light intensities at the time point of 24 hours were used for proteomic analysis. Biological duplicates at each condition were analyzed and each sample was measured by a reversed-phase liquid chromatography-electrospray ionization-mass spectrometry and tandem mass spectrometry (RPLC-ESI-MS/MS) in technical duplicate.

The samples were quickly spun down to remove the supernatant. 100 μL lysis buffer (200 mM NaCl, 4% SDS, a phosphatase inhibitor, and a protease inhibitor) was then added to the sample. For protein extraction, samples were first sonicated for 5 min, and heated at 95°C for 10 min. After cell lysis, 600 μL acetone was added to the sample for protein precipitation. The samples were incubated at room temperature for 10 min, and then centrifuged at 14,000 rpm for 10 min. The supernatant was taken out and the precipitate was washed once with 400 μL acetone. 4% sodium dodecyl sulfate (SDS) buffer was used to dissolve the protein precipitate. To

facilitate protein dissolving, the samples were heated at 95 °C for 5 min. The heat-treated samples were then centrifuged at 14,000 rpm for 5 min and the supernatant was saved for protein concentration measurement using bicinchoninic acid (BCA) assay.

100 µg of the protein sample was used for protein analysis. For reduction, 2 µL 100 mM dithiothreitol (DTT) was added to each sample, and samples were incubated at 37 °C for 30 min. For alkylation, 4 µL 100 mM iodoacetamide (IAA) was added to each sample, and samples were incubated at room temperature for 20 min in dark. After reduction and alkylation, those samples were processed with the SP3 sample preparation method [96]. The proteins captured on the beads were resuspended in 100 µL NH_4HCO_3 (pH 8.0). 3 µg trypsin was added to the sample for digestion at 37 °C overnight. After protein digestion, 1 µg of peptides for each sample was analyzed by RPLC-ESI-MS/MS platform in technical duplicate.

A C18 RPLC column (100 µm i.d. x 50 cm, C18, 1.9 µm, 100 Å) connected to an EASY nanoLC-1200 system (Thermo Fisher Scientific) was used for the nanoRPLC separation. Buffer A containing 0.1% (v/v) formic acid (FA) and buffer B containing 80% (v/v) acetonitrile (ACN) and 0.1% (v/v) FA were used to generate gradient separation. The sample was loaded onto the RPLC column with buffer A at 800-bar pressure. Peptides retained on the column were separated by a linear gradient. The flow rate was 400 nL/min. The gradient for RPLC separation was as follows: 60 minutes from 5% to 30% (v/v) B, 28 minutes from 30% to 50% (v/v) B, 2 minutes from 50% to 80% (v/v) B, and 15 minutes maintained at 80% (v/v) B.

A Q-Exactive HF mass spectrometer (Thermo Fisher Scientific) was used for MS analysis. The resolution for full MS was 60 000, the AGC was 3E6, the maximum injection time was 50 ms, and the scan range was 400−1800 m/z. A Top10 data-dependent acquisition (DDA) method was applied. Only ions with a charge of 2 or higher were isolated in the quadrupole and

fragmented in the high-energy collision dissociation (HCD) cell with normalized collision energy (NCE) as 28%. The resolution for MS/MS was 60 000, the AGC target for MS/MS was 2E5, and the maximum injection time for MS/MS was 50 ms. The isolation window was 2 m/z, and the intensity threshold for fragmentation was 5E4. Dynamic exclusion was 30 s.

All MS raw files were processed with MaxQuant 1.5.5.1 [97]. The proteome databases of proteobacteria bacterium, *C. sorokiniana*, and cyanobacterium were downloaded from UniProt and combined for database search. All the parameters were set to default. The match between runs (MBR) function and label-free quantification was turned on [98]. The false discovery rates (FDRs) were controlled to be lower than 1% at the peptide and protein group levels.

2.9 Statistical analysis

Data was collected from sample analysis of two biological replicates. All data collected were analyzed using the statistical tools of R (version 3.6.3). To determine whether a parametric or non-parametric test was necessary, the data were first tested for normality and equal variance using a Shapiro-Wilk's test and an F-test, respectively. Data that were normal with equal variance were tested using an analysis of variance (ANOVA) and a Tukey test was used when applicable to compare individual factors. Data with non-normal distribution and unequal variance were tested using the Kruskal-Wallis test. All tests were performed with a significance value of $\alpha = 0.05$. QIIME 2™ was used on 16S rRNA gene sequences to obtain taxonomic/phylogenetic data of amplicon sequence variant (ASV) [99]. Microbial community analysis was then completed on the ASV data using Vegan, ggplot2, phyloseq, and MASS R libraries. The diversity index (Shannon's index, H), community evenness (Pielou's index, J), and rarefaction curve for algal assemblage were calculated. ASV data was also applied to graph the relative abundances of individual samples.

3. Results and discussion

3.1. The algae assemblage utilizing formate as a carbon source

Formate and bicarbonate feeding cultures under different light intensities were used to study the growth of algae assemblage (Fig. 2.1). Under the high light intensity of 500 µmol/m2/s, biomass concentrations reached 0.12±0.00 and 0.16±0.03 g/L at 24 hours of the culture for the formate and bicarbonate carbon sources, respectively (Fig. 1a). Under the low light intensity of 50 µmol/m2/s, the corresponding biomass concentrations were 0.15±0.02 and 0.13±0.02 g/L for formate and bicarbonate (Fig. 2.1b). The growth patterns under individual light intensities and carbon sources show that the algae assemblage grew similarly on bicarbonate and formate under both light intensities. The statistical analysis concluded that biomass concentrations of all four cultures (with different light intensities and carbon substrates) were not significantly ($P>0.05$) different at early culture phase (within 24 hours). However, the consumptions of formate and bicarbonate by the assemblage show different patterns between different carbon sources and light intensities (Fig. 2.1c and d). Under the high light intensity, 79% of bicarbonate was consumed in 24 hours of the culture. While 49.5% of bicarbonate was consumed by the culture under the low light intensity. Corresponding sodium bicarbonate concentrations were dropped from 1g/L to 0.32±0.10 and 0.46±0.02 g/L for 500 and 50 µmol/m2/s, respectively. As for formate consumption, the data show that the assemblage consumed formate slower than bicarbonate under corresponding light intensities. This observation indicates algal consortia has a much higher photosynthesis activity than formate oxidations at early growth phase (when total biomass concentration is low and shade effect is minimal). After 24 hours of cultivation, 25 and 26% of the provided formate were consumed under 500 and 50 µmol/m2/s, respectively. The identical sodium formate consumptions indicated that bacterial formate utilization were not affected by

light intensity at the early growth stage (Fig. 2.1). Despite similar final biomass concentrations, different observed rates of formate and bicarbonate consumption between different conditions led to significant differences in biomass yield. After 24 hours of the cultivation, biomass yields under different light and carbon source conditions are shown in Table 2.1. The biomass yields of the cultures grown on formate were significantly ($P<0.05$) higher than the cultures grown on bicarbonate. The data are consistent with the fact that formate has higher degree of reduction than bicarbonate, leading to higher biomass yields.

TN and TP concentrations were also monitored during the culture (Fig.2.1e, f, g, and h). The cultures from the different conditions show similar trends. The TN reductions were 25.2, 26.6, 22.2, and 26.2 mg/L, for bicarbonate under 500 μmol/m^2/s, formate under 500 μmol/m^2/s, bicarbonate under 50 μmol/m^2/s, and formate under 50 μmol/m^2/s, respectively, and the corresponding TP reductions were 4.7, 8.0, 8.4, and 8.9 mg/L. There were no significant ($P>0.05$) differences in TN reduction between the four cultures under different carbon sources and light intensities. As for TP reduction, the culture on bicarbonate under high light intensity had slightly less TP reduction than the other three cultures that had similar TP usage (Fig.1e, f). Since a large quantity of TP and TN remained at the end of 24 hours, they were not the limiting nutrients during growth. Therefore, carbon and light intensity are the main factors that influence carbon assimilation of the algal assemblage under the studied conditions.

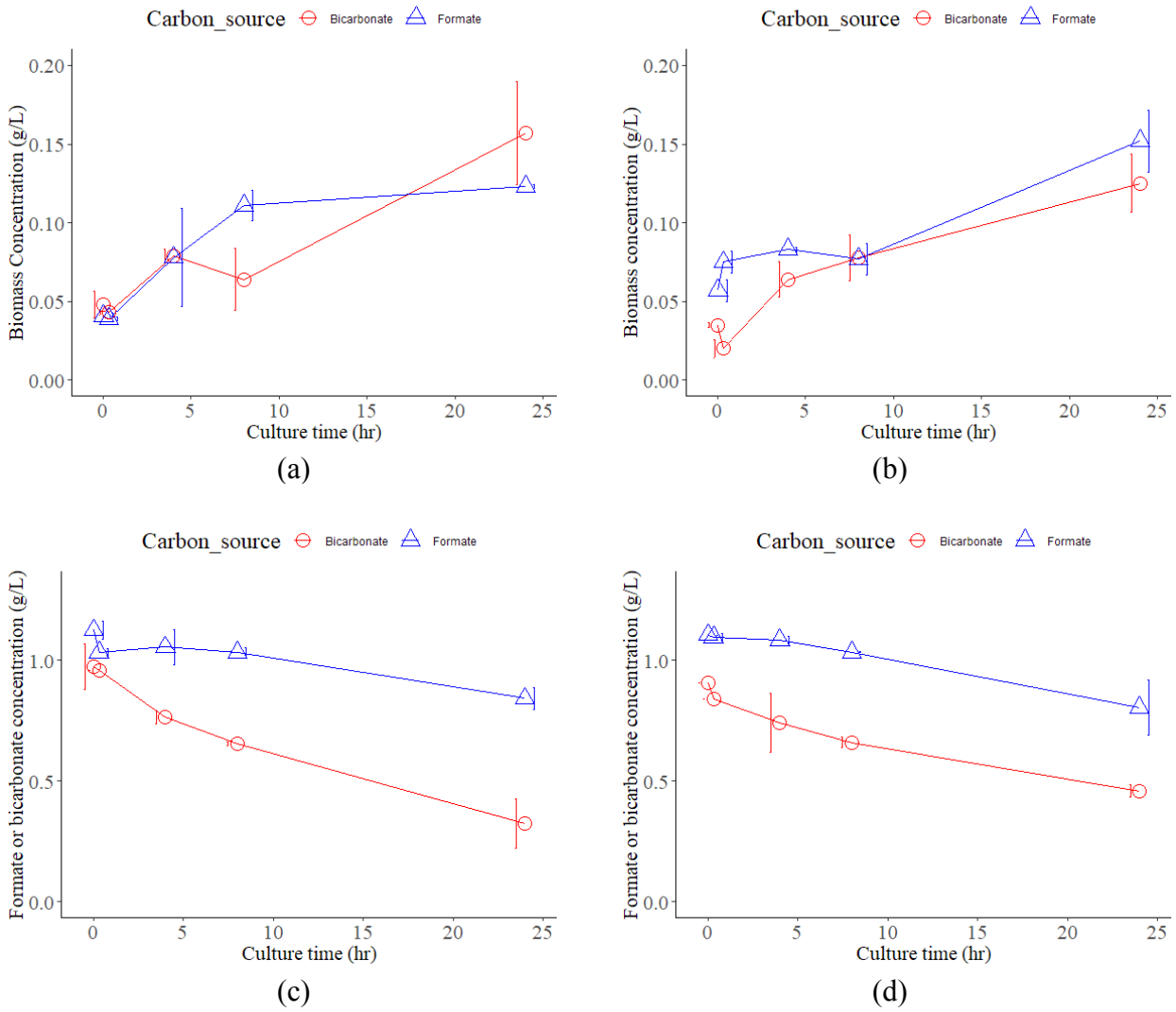


Figure 2.1 Time course of batch algae cultivation on formate and bicarbonate*

(a). Biomass concentration under 500 µmol/m²/s
(b). Biomass concentration under 50 µmol/m²/s
(c). Formate/bicarbonate concentration under 500 µmol/m²/s
(d). Formate/bicarbonate concentration under 50 µmol/m²/s
(e). TP concentration under 500 µmol/m²/s
(f). TP concentration under 50 µmol/m²/s
(g). TN concentration under 500 µmol/m²/s
(h). TN concentration under 50 µmol/m²/s
**: Data for biomass concentration, formate/bicarbonate concentration, TP concentration, TN concentration are the mean of 2 replicates*

Figure 2.1(cont'd)

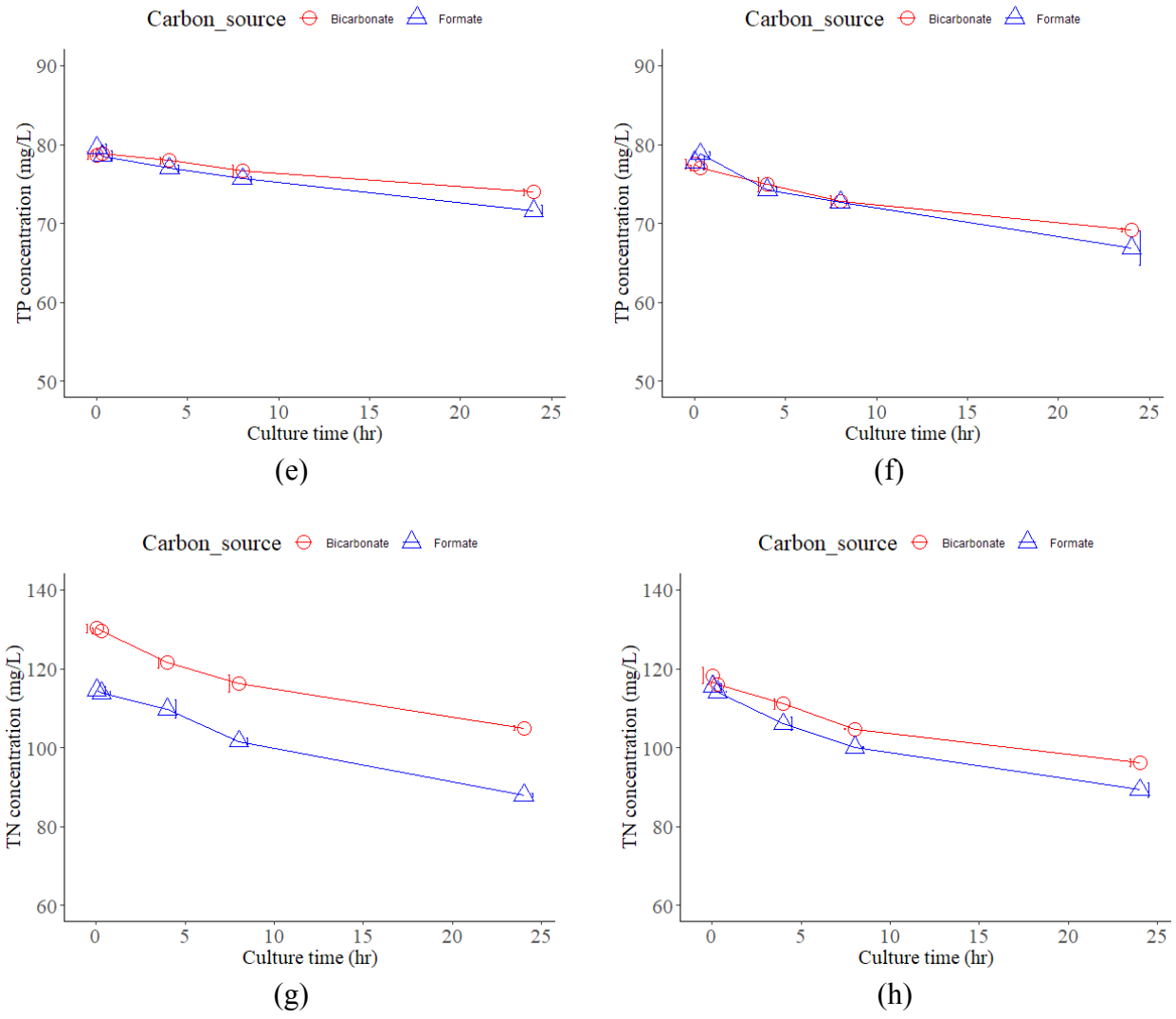


Table 2.1 Biomass concentration and yield on formate and bicarbonate from batch culture

Culture conditions		Biomass yield with respect to carbon source (g biomass/g substrate)*	Final biomass concentration (g biomass/L)
500 µmol/m^2/s	Formate	0.44 ± 0.01	0.12± 0.00
	Bicarbonate	0.23 ± 0.06	0.16± 0.03
50 µmol/m^2/s	Formate	0.49 ± 0.04	0.15± 0.02
	Bicarbonate	0.28 ± 0.05	0.13± 0.02

**: Biomass yield is calculated using biomass dry matter at the end of a culture divided by total formate and bicarbonate ions (without counting sodium ion) consumed.*

3.2. Metabolic analysis via ^{13}C tracing

The metabolic activity of the community was assessed using ^{13}C pulse-trace experiments. Labeled formate or bicarbonate was pulsed to the cultures at the beginning of the culture, and the resulting label incorporation into the metabolites or proteins was determined (Fig. 2.2). The turnover rates of free metabolites (sugar phosphates and Calvin cycle intermediates) show that at the beginning of the culture (within the first 4 hours), ^{13}C carbon from formate incorporation into the central metabolism of the assemblage was slower than bicarbonate under both light intensities (Fig. 2.2a and b; Fig. S2.2). This observation confirms the algal photosynthesis is dominant for biosynthesis, while formate has additive growth contributions. In addition, the pure culture of *C. sorokiniana* indicates that the alga is not able to efficiently utilize formate (unpublished data). Therefore, the formate utilization by algal assemblage requires the additional steps of bacterial conversion of formate to CO_2, transportation of CO_2 to the algae, and incorporation of CO_2 into the Calvin cycle. Interestingly, low light cultivation did not significantly decrease labeling rates of free metabolites from glycolysis and the Calvin cycle (i.e., the dark reactions) comparing to high light conditions, this observation suggests our community has robust photomixotrophic metabolism under light limited conditions. Moreover, we observe that glycerol-3-phosphate (Glycerol-3P) labeling rates were different between two light conditions. Glycerol-3P is a key intermediate related to glycerol production and lipid accumulation. Its turnover rate is dependent on NADPH and ATP from photosynthesis (i.e., light-dependent reactions). Low light conditions could prevent Glycerol-3P from consumption for lipid synthesis.

The labeled data from protein-based amino acids confirm that ^{13}C-formate was used as a carbon source to synthesize biomass (i.e., protein) of the algal assemblage (Fig. 2.2c and d; Fig.

S2.3). Under high light intensity, the culture on formate shows moderately faster rates of labeling in proteinogenic amino acids than the cultures grown with bicarbonate (Fig. 2.2c). Labeled serine and methionine from the culture on formate were 41.5±0.9 and 41.7±3.1%, respectively, at 24 hours, which were significantly ($P<0.05$) higher than the culture on bicarbonate (35.7±3.1 and 31.1±0.0% respectively). Under low light intensity, amino acid labeling in ^{13}C formate and ^{13}C-bicarbonate cultures was higher (Fig. 2.2c and d) than the culture under the high light intensity condition. There were no significant ($P>0.05$) differences between the labeling contents for the measured amino acids (Alanine, glycine, serine, phenylalanine, aspartic acid, glutamic acid, and methionine) between the two carbon sources at 24 hours of the culture (Fig. 2.2d). The results were consistent with the biomass yield of different cultures, with the cultures on formate resulting in higher biomass yields than the cultures grown on bicarbonate, and the beneficial impact of low light intensity on biomass yield (Table 2.1). In addition, it is important to note that protein labeling experiments were conducted over a longer duration (24 hours) than free metabolites labeling experiments (4 hours). The results indicate that formate continuously supported biomass growth over an extended period (24 hours), which could guide the design of semi-continuous cultures (presented in the next section). In summary, the kinetic data and ^{13}C pulse-trace data indicate that syntrophic interactions between alga and bacteria were established by the algal assemblage using formate as the carbon source (Figure 2.3) [100]. Bacteria in the assemblage first utilize formate as an energy source for their metabolism and release CO_2 that satisfies the need for algal photosynthesis; photosynthesis then generates oxygen and other micro-nutrients to support bacterial growth.

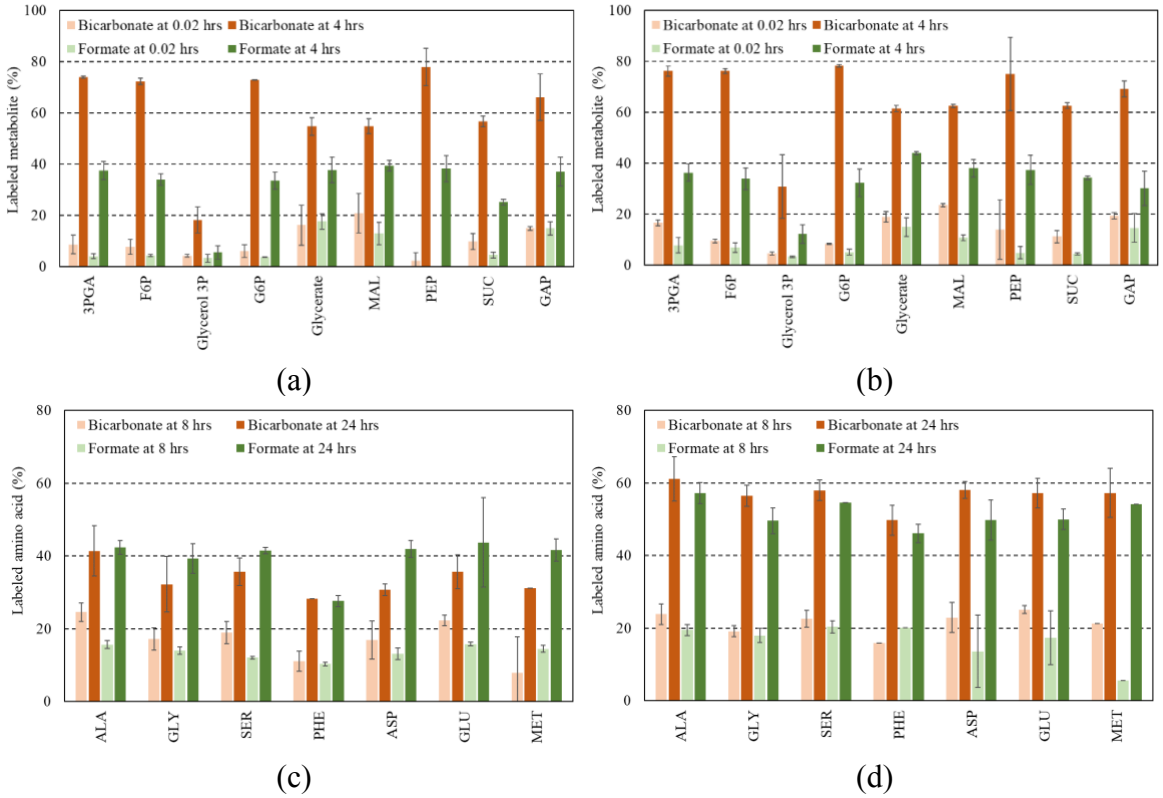


Figure 2.2 Labeled free metabolites and proteinogenic amino acids from the cultures on formate and bicarbonate

a. Free metabolites under 500 µmol/m²/s; b. Free metabolites under 50 µmol/m²/s; c. Amino acids under 500 µmol/m²/s; d. Amino acids under 50 µmol/m²/s
Free metabolites: 3PGA is 3-phosphoglyceric acid; F6P is D-fructose-6-phosphate; Glycerol 3P is glycerol 3-phosphate; G6P is glucose 6-phosphate; MAL is malate; PEP is phosphoenolpyruvic acid; SUC is succinate; and GAP is D-glyceraldehyde-3-phosphate
Amino acids: ALA is alanine; GLY is glycine; SER is serine; PHE is phenylalanine; ASP is aspartic acid; GLU is glutamic acid; and MET is methionine

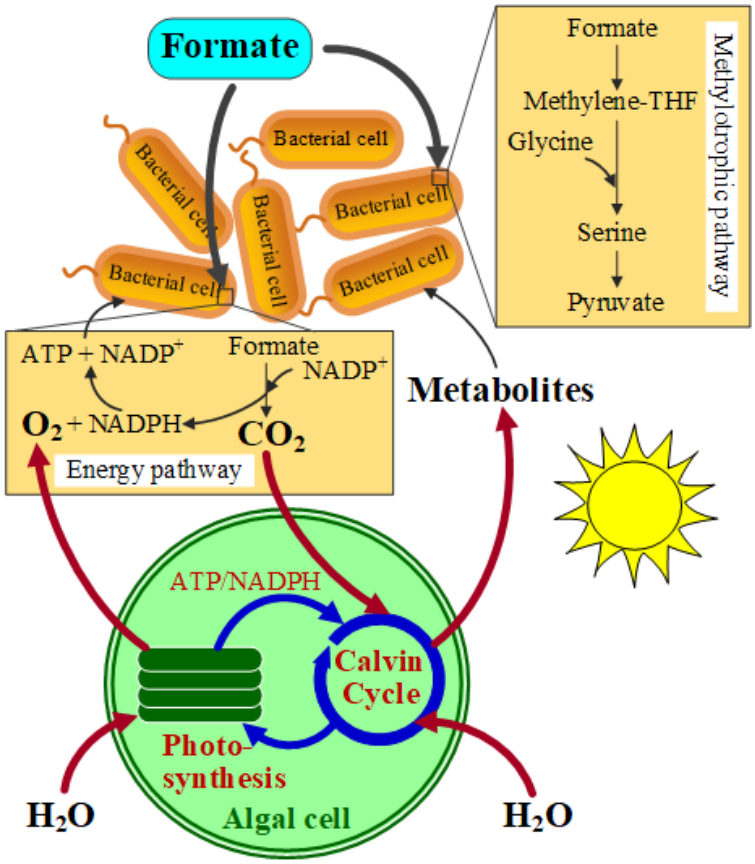


Figure 2.3 Symbiosis of alga and bacteria on formate

3.3. Amplicon sequencing and proteomics

Amplicon sequencing was applied to the samples taken at the beginning and end of the cultivation. The data demonstrate that 16S rRNA gene sequences in samples ranged from 18,326 to 38,996 reads (Figure S2.4a). The number of species in the algal assemblages was stabilized after sampling 10,000 sequences. The sequences were rarified at 18,000 reads. A steep gradient of the rank abundance at the rank less than 30 (lower rank means higher abundance) presents low evenness as the high-ranking species (algae) have much higher abundances than the low-ranking species (bacteria) (Figure S2.4b). A three-way ANOVA further concludes that light intensity had significant ($P<0.05$) influences on alpha diversity and evenness (Figure S2.5 and Table S2.1). The algal assemblages under 500 µmol/m^2/s were more diverse than the algal assemblages under

50 μmol/m^2/s (Figure S2.5). Carbon source and culture time had no significant (P>0.05) influences on both diversity and evenness. Permutation one-way ANOVA of light intensity, carbon source, and culture time on the batch cultures shows that these factors had no significant (P>0.05) influences on the number of microbial species change in the algal assemblages between different culture conditions (Table S2.2). The diversity analysis indicates that the studied algal-bacterial symbiotic system is relatively resilient to changes in culture conditions during the batch culture.

A total of 32 microbial genera from both were identified in both cultures (Table S2.3). The results show that the dominating phylum was Chlorphyta for all four cultures at 24 hours (Fig. 2.4a), which corresponds to *C. sorokiniana* in the assemblage. The relative abundances of *C. sorokiniana* were 92±1.35, 80.8±1.9, 86.9±3.6, and 83.6±0.4% for the cultures with formate under 500 μmol/m^2/s, formate under 50 μmol/m^2/s, bicarbonate under 500 μmol/m^2/s, and bicarbonate under 50 μmol/m^2/s, respectively. Compared to its abundance at the beginning of the culture (86.2%), the cultures under high light intensity maintained similar abundances of *C. sorokiniana*, while its abundances in the cultures under low light intensity were significantly (P<0.05) reduced. Correspondingly, the bacterial community percentage in the cultures under low light intensity at 24 hours was significantly (P<0.05) increased compared to the cultures grown under high light intensity (Fig. 2.4b). *Rhizobiales* were the dominant bacterial order in the cultures. An unclassified *Rhizobiales* family and *Methylobacteriaceae* are two dominant *Rhizobiales* families (Fig. 2.4c and d). Carbon sources and light intensity had significant (P<0.05) influences on their abundances in the culture. The data clearly show that *Methylobacteriaceae* were more abundant from the cultures on formate (3.35 and 10.4% for 500 and 50 μmol/m^2/s, respectively) than those on bicarbonate (0.7 and 0.6% for 500 and 50

μmol/m^2/s, respectively). Unclassified *Rhizobiales* family was more abundant in the cultures with bicarbonate than formate. As it is well known, *Methylobacteriaceae* is one of the methylotrophs that are capable of growth on single carbon compounds [101]. It has also been reported that *Methylobacteriaceae* are accumulated under abiotic stress, which is consistent with the observation of this study [102, 103]. The abundance of *Methylobacteriaceae* in the culture with formate under 50 μmol/m^2/s was much higher than the other three cultures. The results from microbial community analysis, along with data of free metabolites, amino acids, and biomass yields, confirmed the high metabolic resilience of the algal-bacterial symbiotic system under a wide range of light and carbon conditions. Particularly, such resilience enables the algal-bacterial system to efficiently utilize formate as a carbon source.

Furthermore, proteomic analysis was performed to gain insights into the population interactions under different light and substrate conditions. The statistical analysis (t-test) and volcano plots were performed using Perseus software [104]. The false discovery rate (FDR) and s0 value for the statistical analysis were 0.05 and 0.1, respectively. The label-free quantification data revealed that high light intensity statistically ($P<0.05$) affects the abundance of photosynthesis enzymes (PSI and PSII) and enzymes in Calvin cycle/amino acid synthesis pathways (90 differentially expressed proteins were identified between formate and $NaHCO_3$ at the high light intensity) (Figure 2.5, Table S2.4 and S2.5). This is not surprising that light controls photosynthesis. Under low light, the addition of bicarbonate or formate did not impact algal proteomics. Because of low light, algal has low CO_2 uptakes, and mass transfer is not rate-limiting. As shown in our recent paper [100], high light condition is beneficial to enhance mutualistic interaction between algae and bacteria (enhanced exchange of O_2 and CO_2). This benefit increases photosynthesis enzyme expressions in *C. sorokiniana*. On the other hand,

proteomics mainly captured the *C. sorokiniana* enzymes because *C. sorokiniana* is the dominant single species. Bacterial protein identification was not successful because of the highly diverse bacterial species and relatively low protein level of each species.

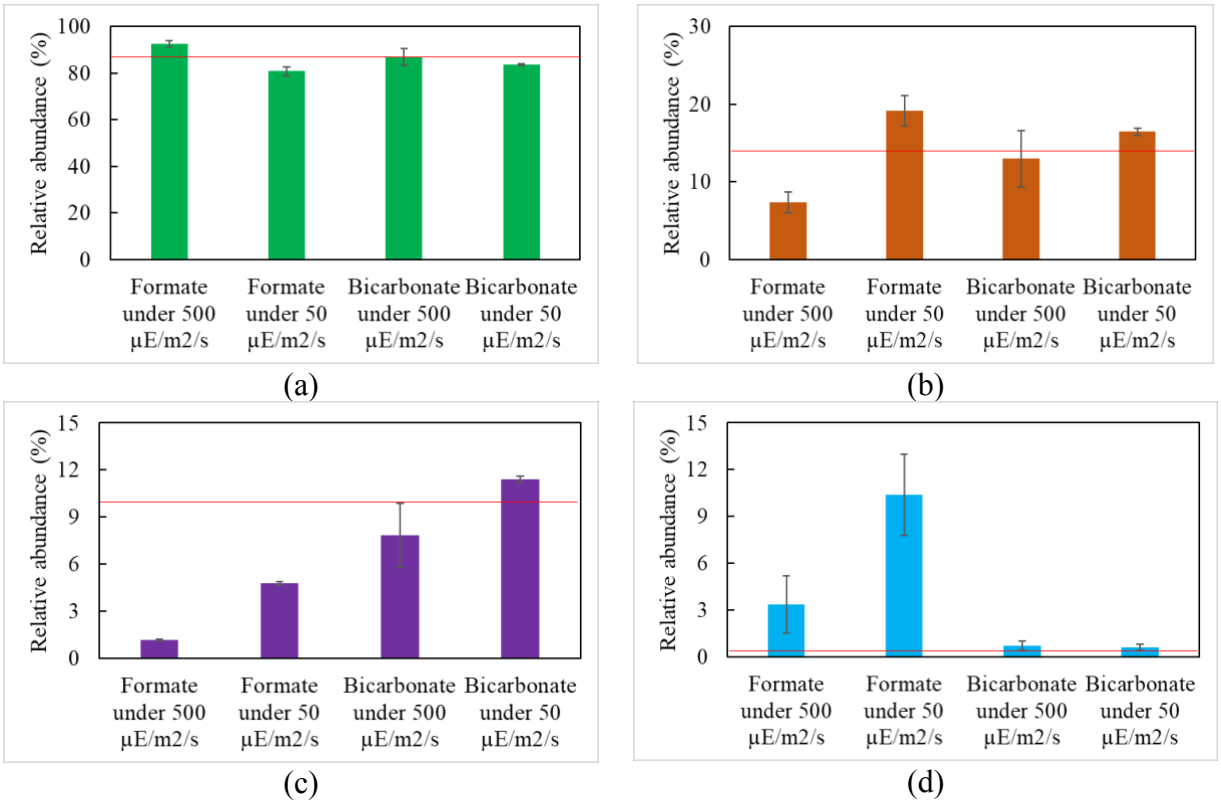


Figure 2.4 Changes of microbial communities from the batch cultures on formate and bicarbonate*

A. Eukarya; b. Bacteria; c. Unclassified Rhizobiales family; d. Methylobacteriaceae
**: Relative abundances of Eukarya, Bacteria, Unclassified Rhizobiales family, and Methylobacteriaceae were 86.2, 13.7, 10.0 and 0.3% at the beginning of the cultures. They were presented as the red lines in the figures. The detailed relative abundances of all communities were presented in Figure S4*

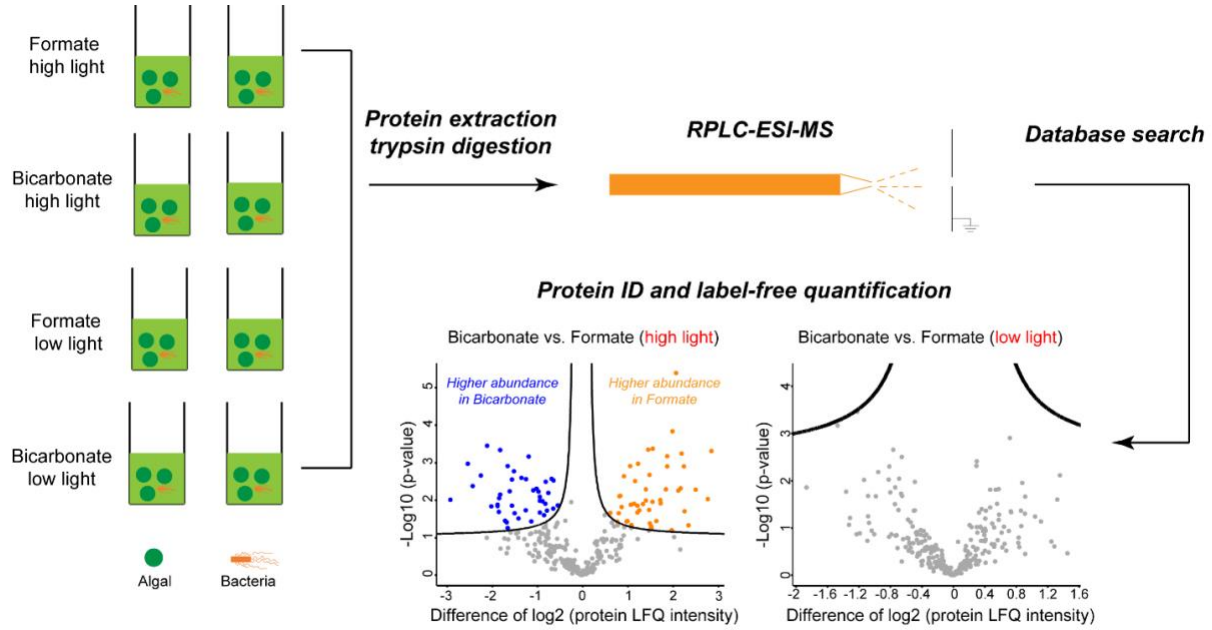


Figure 2.5 Quantitative proteomics of the algal assemblage cultured under bicarbonate and formate conditions using reversed-phase liquid chromatography (RPLC)-electrospray ionization (ESI)-mass spectrometry (MS) and tandem MS*

**: Samples are taken from 24 hours of the culture.*

3.4. Semi-continuous culture of the algae assemblage using formate as a carbon source

The algal assemblage was fed at 1 g sodium formate/L/day to study the performance of carbon capture and biomass accumulation during stable continuous cultivation. 30% of the culture volume was harvested daily. A comparison experiment was run on saturated CO_2 using the same harvesting amount and light intensity. An intermediate light intensity of 180 µmol/m2/s was used for the continuous culture. Figure 2.6 shows the effects of carbon sources on the continuous culture of the assemblage. Under stable culture conditions, there were no significant ($P>0.05$) differences in phosphorous and nitrogen consumption between the two cultures (Fig. 2.6c and d). As for biomass concentration and productivity, biomass concentrations of the cultures on formate and CO_2 were 0.92±0.12 and 0.97±0.19 g/L, respectively, with no significant difference ($P>0.05$) from each other (Fig. 2.6a). Biomass productivity of the culture on formate (0.31±0.04 g/L/day) was significantly ($P<0.05$) higher than that of the culture on CO_2 (0.24±0.06

g/L/day) (Fig. 2.6b). 1 g/L formate was completely consumed by the algal assemblage every day. The elemental composition of the biomass from the formate culture was measured as carbon (47.2±2.7%), hydrogen (6.9±0.3%), and nitrogen (8.0±0.9%), which was not significantly ($P>0.05$) different from the biomass composition from the CO_2 cultures (48.9±0.5, 7.4±0.0, and 8.8±0.6% for C, H, and N, respectively) (Table 2.2). Based on carbon content in the biomass, the carbon balance calculation shows that carbon capture efficiency (calculated by the carbon mass (g) in the harvested biomass divided by the carbon mass (g) in the substrates of formate and flue gas being added to the APBs) of the culture on formate was 89.8±10.2%, which is much higher than the culture on CO_2 (13.4±2.2%) (Fig. 2.6e). The results indicate that the algal assemblage utilizing formate can be a new route to fix carbon and accumulate algal biomass, which is more efficient than the algal cultivation with direct CO_2 aeration.

Microbial community analysis was further conducted on continuous cultures to elucidate the effects of carbon sources on alga and bacteria in the assemblage. 16S rRNA gene sequences were rarified at 5,000 reads, which indicated sufficient sample coverage (Figure S2.7a). A rank abundance curve with a gentle slope between 10 and 60 species shows a more even distribution of gene sequences than the batch culture (Figure S2.7b and Figure S2.4b). Statistical analysis on alpha-diversity and evenness shows that carbon sources had a significant ($P<0.05$) influence on alpha diversity under steady-state semi-continuous culture conditions (Table 2.3 and Table S2.6). The community of the culture on formate had significantly ($P<0.05$) higher H and J (1.5±0.2 and 0.5±0.1, respectively) than the community on CO_2 (0.4±0.1 and 0.1±0.0, respectively), which means that more microbial species were in the formate culture than the bicarbonate culture. Permutation one-way ANOVA of carbon source on beta diversity of the semi-continuous cultures shows similar results from the batch culture. There were no significant ($P>0.05$)

influences of carbon sources on the number of microbial species in the algal assemblages (Table S2.7).

Under the stable culture condition, there were 33 microbial genera identified in both semi-continuous cultures (Table S2.8), which is similar to genera numbers in the batch cultures (Table S2.3). The relative abundances of algae at the domain level for the cultures on formate and CO_2 were 57.9 and 92.3%, respectively, and corresponding abundances of bacteria at the domain level were 42.3 and 7.6% (Fig. 2.7a). The microbial community of the semi-continuous culture on CO_2 was similar to the community of the culture on bicarbonate in the batch culture and a previous study [88], while, community data of the culture on formate demonstrate that compared to the batch culture, formate significantly ($P<0.05$) shifted the community and increased bacterial distribution to facilitate formate utilization. The dominant bacterial phyla of the culture on formate were Bacteroidetes (19.3%) and Proteobacteria (22%) (Fig. 2.7b). Flavobacteriaceae (3.3%) and unclassified Bacteroidetes family (12.8%) are two major families in the phylum Bacteroidetes (Fig. 7c). Unclassified Alphaproteobacteria family (3.2%), Unclassified Rhizobiales family (3.4%), Methylobacteriaceae (7.3%), Unclassified Betaproteobacteria family (5.3%) are four key families in the phylum Proteobacteria (Fig. 2.7d). Besides the two dominant Proteobacteria, Methylobacteriaceae and the unclassified Rhizobiales, an unclassified Bacteroidetes became another dominant bacterial family. It is well known that Bacteoridetes are mainly responsible for degrading carbohydrates in the medium. Enrichment of them in the assemblage of the culture on formate could be interpreted as some carbohydrates from algae needing to be degraded to provide nutrients to other bacteria to grow and convert formate into CO_2 which enhances the symbiosis of alga and bacteria to utilize formate. The

results from microbial community analysis of the continuous cultivation further demonstrated the symbiosis of alga and bacteria in the assemblage of formate utilization (Fig. 2.3).

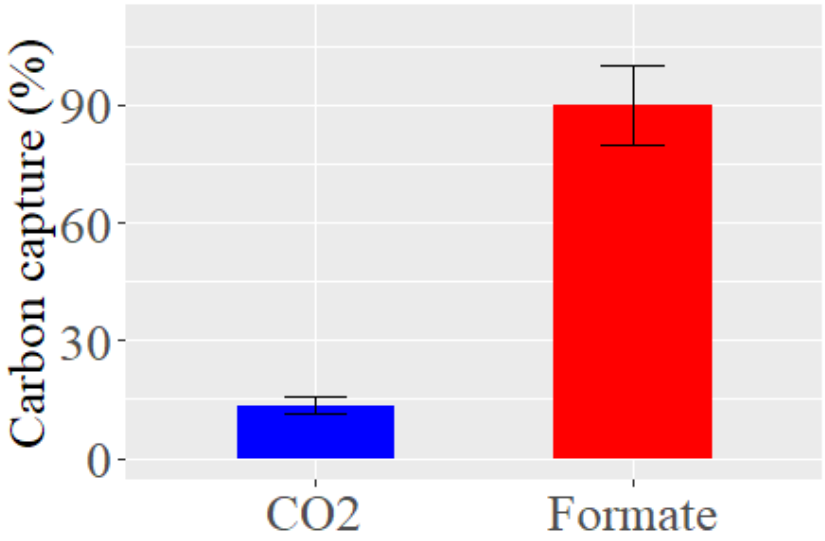
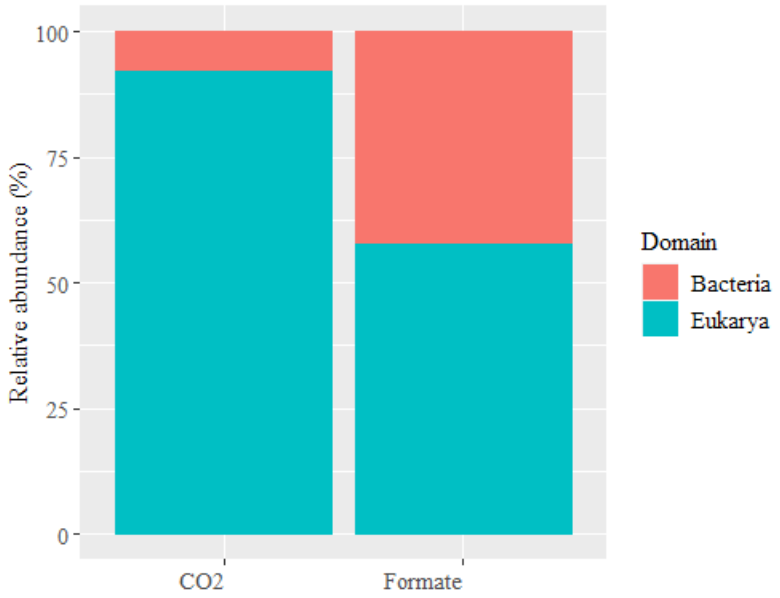


Figure 2.6 Performance of the algae assemblage on formate and (carbonic acid) dissolved CO_2*

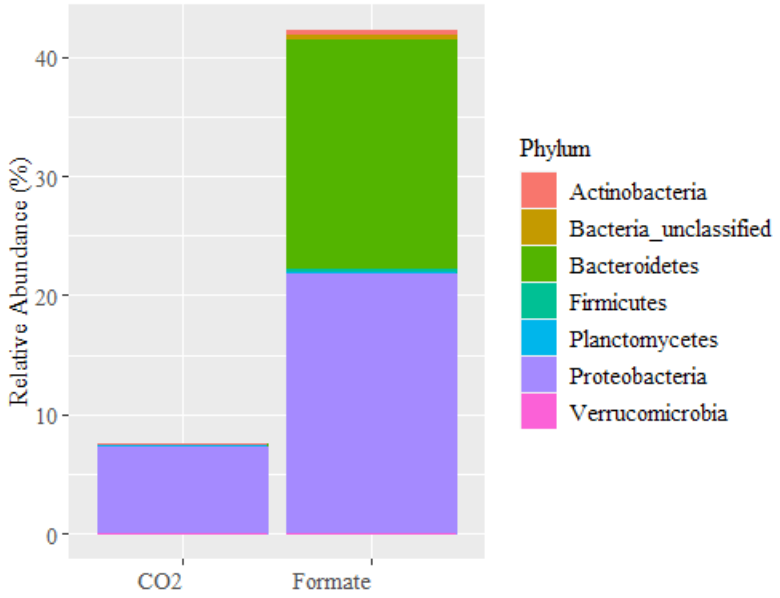
*a. Biomass concentration; b. Biomass productivity; c. TN consumption; d. TP consumption; e. carbon capture***

**: Data for biomass concentration, biomass productivity, TN consumption, TP consumption are the average of two replicates*

***: The carbon capture is calculated by the carbon content in the harvested biomass divided by the carbon in the substrates (formate and flue gas) being added into the reactors*



(a)

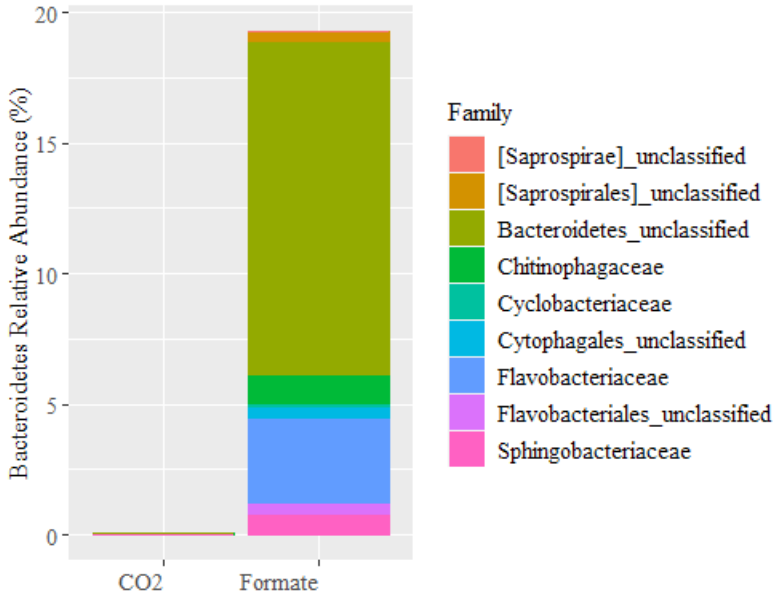


(b)

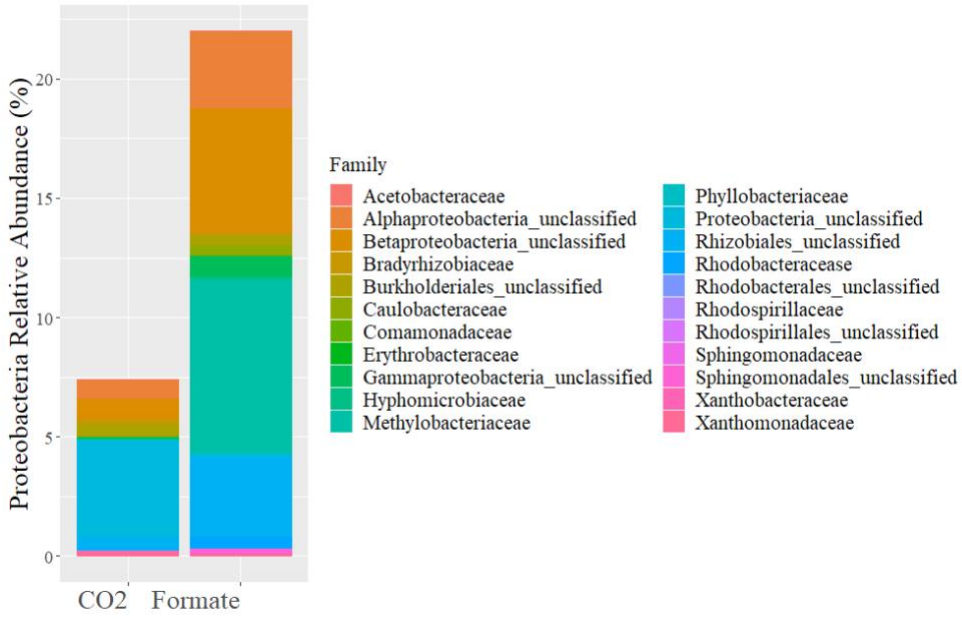
Figure 2.7 Microbial community of the algae assemblage on formate and CO_2

(a) Abundance at the domain level; (b) Bacteria abundance at the phylum level; (c) Bacteroidetes abundance at the family level; (d) Proteobacteria abundance at the family level

Figure 2.7 (cont'd)



(c)



(d)

Table 2.2 Element contents of algal biomass from semi-continuous cultures on different carbon sources

Carbon source	Carbon (% dry matter)	Hydrogen (% dry matter)	Nitrogen (% dry matter)
Formate	47.24±2.67	6.92±0.27	7.99±0.85
CO_2	48.85±0.46	7.39±0.04	8.78±0.56

Table 2.3 Diversity and evenness of microbial communities of the algae assemblage on formate and CO_2 during semi-continuous cultures

Carbon source	Frequency [a]	H[b]	J[c]
Formate	26	1.52±0.20	0.47±0.06
CO_2	18	0.39±0.08	0.13±0.03

[a] *Frequency: numbers of observed frequency*
[b] *H: Shannon's index indicates the diversity of the microbial community*
[c] *J: Pielou's index indicates the evenness of the microbial community*

4. Conclusions

A new and robust algal-bacterial assemblage containing *C. sorokiniana* and bacteria (*Proteobacteria* and *Bacteroidetes*) has been adaptively evolved and selected to utilize formate as a carbon source. Formate significantly enhances the methylotrophic population in the assemblage. Isotope tracing results conclude a significant contribution of formate as a carbon source for photomixotrophic growth. Formate can be used as an alternative carbon form to bicarbonate or CO_2. Particularly, formate as a carbon source allows algal growth under a wide range of pH, from weakly acidic to alkaline conditions. In addition, the assemblage of formate utilization has strong resilience under different light intensities. A high carbon capture of 90% was achieved from semi-continuous cultivation of the assemblage on formate. With the advancement of current research on the electrochemical conversion of CO_2 to formate, this study provides a novel, flexible, and efficient route to fix CO_2 into algal biomass for value-added uses. Finally, the capability of the assemblage to handle relatively high formate concentration is highly advantageous to repel protozoa, insects, and other contaminated species during long-term, continuous cultivation.

CHAPTER 3: EFFECTS OF FORMATE ON AN ALGAL ASSEMBLAGE AND CORRESPONDING CULTIVATION STRATEGY

1. Introduction

Human activities such as industrialization, deforestation, and energy generation have led to massive greenhouse gas emissions into the atmosphere. This global warming effects have become increasingly severe since the beginning of 21st century [105], triggering many side effects such as climate change [106], crop failure [107], and species extinction [107]. CO2 as one of the major greenhouse gases has seen a dramatic increase in the atmosphere over the past several decades [108]. Urgent action is needed to address this problem, and carbon capture is a critical solution to reducing greenhouse gas emissions and mitigating the global warming effects.

Microalgae, as one of the oldest organisms on the earth, can generate oxygen through photosynthesis and help reduce greenhouse gas emissions. Due to their good CO2 assimilation capacity [109], microalgal cultivation has been recognized as a promising carbon capture method [110]. Compared to terrestrial plants, microalgae exhibited better photosynthetic efficiency of the CO2 capture and microalgal biomass accumulation [111]. The algal biomass is a valuable feedstock for a range of applications including biofuels [112], animal feeds [113], and cosmetics [114]. In the biofuel production area, microalgae offer several advantages over other biomass feedstocks such as corn and soybeans. Microalgae have higher lipid content [115] and can produce more biomass per unit of land area [116]. Microalgae can be cultivated on non-arable land, alleviating its competition with human food resources. As a source of animal feed, microalgae are rich in protein [117] and vitamins [118], providing high nutritional value. In cosmetics area, microalgae are frequently used to extract valuable cellular products (i.e., astaxanthin) [119].

While bicarbonate [120], glucose [121], and carbon dioxide [122] are extensively studied as carbon sources in microalgal culture, research on formate as carbon source is limited. Formate

is a soluble chemical that has much better mass transfer efficiency than CO_2 during the microalgal cultivation. It is also relatively stable and can withstand a wide range of pH conditions compared to other carbon sources (i.e., bicarbonate and CO_2). These advantages make formate an promising carbon source for microalgal cultivation. Formate can also serve as an inhibitor for protozoa and insects [10], which may help prevent their contaminations during microalgal cultivation. In this study, we aimed to elucidate the effects of formate on microbial community in algal-bacterial system, and develop a strategy to promote long-term and stable algal cultivation.

2. Materials and methods

2.1 Algal assemblage and cultivation system

The microbial assemblage containing a selected microalga Chlorella sorokiniana MSU from the Great Lakes region and several bacteria (mainly Bacteroidetes and Proteobacteria etc.) was continuously cultured in flasks on Tris-Acetate-Phosphate (TAP) medium at room temperature under constant fluorescent light. The assemblage was used as the seed to inoculate the algae photobioreactors (APBs). Modified liquid TAP medium (without acetic acid and tris base) was used for microalgal cultures, which contains 7.5 mmol L− 1 of NH_4Cl, 0.34 mmol L− 1 of $CaCl_2 \cdot 2H_2O$, 0.4 mmol L− 1 of $MgSO_4 \cdot 7H_2O$, 0.68 mmol L− 1 of K_2HPO_4 (anhydrous), 0.45 mmol L− 1 of KH_2PO4 (anhydrous), 0.09 mmol L -1 $FeCl_3 \cdot 6H_2O$, and 1ml TAP trace elements solution. The modified TAP medium was unsterilized. The microbial community of the assemblage was analyzed before the inoculation.

Lab-scale APBs were modified based on 10 L Eppendorf BioFlo®/CelliGen® 115 Benchtop fermenters with a working volume of 7.5 L (Fig. S1a). Metal shells with adjustable light-emitting diode (LED) light strips installed inside were placed around the fermenters.

Two different light intensities (180 μmol/m^2/s and 500 μmol/m^2/s) and two formate feeding rate (1 g/L/day and 2 g/L/day) were tested to study the effects of formate addition on growth of the microalgal assemblage. The culture time was 336 hours for each condition. During the cultivation, 30 samples were collected to monitor biomass concentration and nutrient consumption for each reactor.

An alternative carbon source (sodium bicarbonate) was added into the formate-feeding microalgal culture system to study the effects of alternating carbon sources on biomass accumulation and microbial communities. The culture time was 42 days. The light intensity was 180 μmol/m^2/s with 30% v/v daily harvesting amount. During the first week of cultivation, formate was fed at 1g/L/day. At the second week of cultivation, bicarbonate was added into the photobioreactor to replace formate as the alternative carbon source. This culture mode was repeated twice until culture time reached 42 days.

2.2 Chemical analysis

Samples were analyzed for dry biomass weight, pH, and nutrient (total nitrogen (TN), total phosphorus (TP), nitrate (NO_3-N), and ammonia (NH_3-N)) concentrations. Algal biomass was collected using a Thermo Electron Corporation IEC Centra CL2 Centrifuge at 3800 rpm for 5 minutes. Biomass was washed once and resuspended using deionized water, and then dried at 105°C for 24 hours. Sample pH was measured using a pH meter (Fisherbrand™ accumet™ AB15 + Basic, Fisher Scientific Co., Pittsburgh, PA). Nutrient concentrations were tested in the liquid supernatant using nutrient test kits (HACH Company, Loveland, Colorado) equivalent to the Environmental Protection Agency (EPA) methods (hach.com/epa). Algal biomass composition was analyzed using the standard forage analysis method [90].

Formate concentration of samples for the kinetic study was determined by high-

performance liquid chromatography (HPLC) (Shimadzu Corp., Kyoto, Japan) equipped with an analytical column (Aminex HPX-87H, Bio-Rad Laboratories, Inc., Hercules, CA) and a refractive index detector (Shimadzu Corp., Kyoto, Japan). The mobile phase was 0.005 mol/L sulfuric acid at a flow rate of 0.6 mL/min. The oven temperature was set at 65 °C. The bicarbonate concentration of algal samples in the kinetic study was determined by the alkalinity test kit (HACH Company, Loveland, CO).

2.3 Microbial community analysis

Samples (1 mL) collected for DNA analysis were kept frozen at -20°C until analysis. To remove nutrient media, algae sample was centrifuged using an Eppendorf 5416R centrifuge at 10,000 rpm for 5 min and the supernatant was discarded. The remaining pellet was used for DNA extraction using the DNeasy® PowerSoil® Kit (Qiagen, Germany). DNA was eluted with 100 μL of 10 mM Tris-HCl (pH 8.5) and the concentration and purity determined using a NanoDrop Lite spectrophotometer (Thermo Fisher Scientific, USA). The DNA extracts were stored at -80°C for several weeks and then used for PCR and Illumina DNA sequencing.

Illumina sequencing was performed for both 16S rRNA genes and Internal Transcribed Spacers (ITS) to assess the bacterial and fungal communities, respectively. The primers used for bacterial community analysis are: the forward primer Pro341F (5'CCTACGGGNBGCASCAG-3') and the reverse primer Pro 805R (3'GACTACNVGGGTATCTAATCC-5') that target 16S rRNA genes in the V3−V4 region. The primers used for fungal community analysis are: the forward primer ITS1F and the reverse primer ITS2R, targeting ITS rRNA genes in the V3-V4 region. Prior to PCR, extracted DNA samples were diluted 10x due to high DNA concentrations. The PCR conditions were as follows: 1.0 μL DNA template (10x diluted), 0.5 μL of 100 μM forward primer, 0.5 μL of 100 μM reverse primer, 12.5 μL 2x Supermix (Invitrogen, USA), and

10.5 μL PCR grade water. The PCR program used for all assays is as follows: 96°C for 2 min, followed by 30 cycles of 95°C for 20 s, 52°C for 30 s, and 72°C for 1 min, and a final elongation period of 72°C for 10 min. After PCR, samples were diluted to normalize DNA concentrations within a range of 5-10 ng/μL. DNA concentration was determined using the PicoGreen® dsDNA quantitation assay (Invitrogen, USA) and Fluostar Optima microplate reader (BMG Labtech, Germany). The PicoGreen® conditions were as follows: 95 μL 1x TE buffer solution, 100 μL 1:200 diluted PicoGreen® reagent, 5 μL DNA template. Samples with known DNA concentrations were also prepared for standard curve generation. Illumina library preparation and sequencing were performed at the Michigan State University Genomics Laboratory, East Lansing, USA. QIIME 2™ was used for all sequence analyses.

In addition, the 16S rRNA gene sequencing was also used to determine *C. sorokiniana* in the assemblage [88]. It has been reported that *Cyanobacteria* have 85-93% of 16 rRNA gene sequences similar with *C. sorokiniana*, while, color, shape, and size of both species are very different [95]. Therefore, after microscopic imaging verification of each sample, *Cyanobacteria* sequence was interpreted as microalga *C. sorokiniana* for all samples.

2.4 Statistical analysis

All data collected was analyzed using the statistical tools of R (version 3.6.3). To determine whether a parametric or non-parametric test was necessary, the data were first tested for normality and equal variance using a Shapiro-Wilk's test and an F-test, respectively. Data that were normal with equal variance were tested using an analysis of variance (ANOVA) and a Tukey test was used when applicable to compare individual factors. Data with non-normal distribution and unequal variance were tested using the Kruskal-Wallis test. All tests were performed with a significance value of $\alpha = 0.05$. Microbial community analysis was completed

using Vegan, ggplot2, phyloseq, and MASS R libraries. Taxonomic/phylogenetic data was analyzed in order to graph relative abundances of samples.

3. Results and discussion

3.1 Effects of formate on biomass accumulation of the algal assemblage

Figure 3.1 illustrates biomass concentration and formate utilization under different light intensities and formate feeding rates. Notably, both algal assemblages under 180 μmol/m^2/s exhibited higher biomass concentrations compared to those under 500 μmol/m^2/s (Figures 3.1a and 3.1b). Under the formate feeding rate of 1 g/L/day, biomass concentration slightly decreased during the first 96 hours after inoculation, but gradually increased thereafter, reaching a stable concentration of 1.12±0.15 g/L and 0.89±0.04 g/L for light intensities of 180 μmol/m^2/s and 500 μmol/m^2/s, respectively, at the end of the culture period (Figure 1a). Similarly, under the formate feeding rate of 2 g/L/day, biomass concentration dropped slightly in the first 96 hours, and then increased thereafter, maintaining a stable concentration of 1.01±0.02 g/L and 0.80±0.09 g/L under light intensities of 180 μmol/m^2/s and 500 μmol/m^2/s, respectively.

Formate utilization under different light intensities and formate feeding rate was monitored during the cultivation (Figures 3.1c and 3.1d). With the formate feeding rate of 1 g/L/day, formate started to accumulate in the culture of microalgae under 500 μmol/m^2/s (Figure 3.1c), and reached to the highest concentration of 0.65±0.20 g/L at the 96th hour. While, under 180 μmol/m^2/s, no significant formate accumulation was observed (Figure 3.1c). With a formate feeding rate of 2 g/L/day, formate utilization under two different light intensities displayed a similar trend. The average formate concentration under 500 μmol/m^2/s reached 0.30±0.06 g/L, and the formate concentration remained nearly zero under 180 μmol/m^2/s.

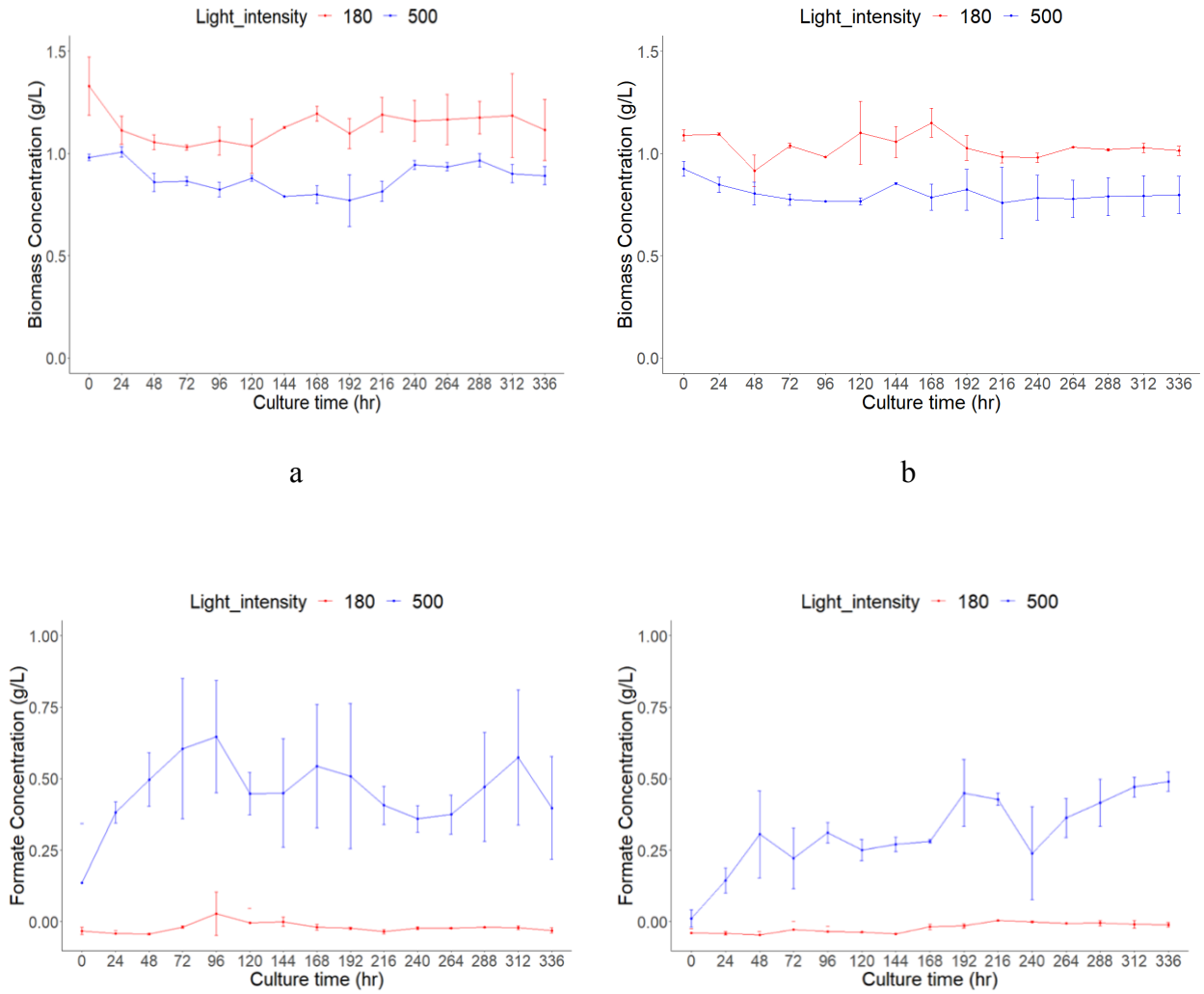


Figure 3.1. Time course of formate utilization and biomass accumulation under different light intensities

a. Biomass concentration with 1g/L/day formate feeding rate; b. Biomass concentration with 2g/L/day formate feeding rate; c. Formate concentration with 1g/L/day formate feeding rate; d. Formate concentration with 2g/L/day formate feeding rate

The effects of varying light intensity and formate feeding rate on the biomass concentration, total nitrogen, and total phosphorus of the microalgal culture were statistically compared (Figure 3.2, Table 3.1, Figures S3.1 and S3.2). The difference between biomass concentration on various light intensities was significant ($P<0.05$) (Table 3.1). Biomass concentration decreased when light intensity increased (Figure 3.2a). While, low formate feeding rate (1 g/L/day) significantly ($P<0.05$) benefits biomass accumulation of the cultivation than high formate feeding rate (2 g/L/day). As for nutrients (TN and TP), TN data show no significant differences ($P>0.05$) between the cultivations under 180 µmol/m2/s and 500 µmol/m2/s with 2 g/L/day formate feeding rate. While under 1g/L/day formate feeding rate, the cultivation with low light intensity had significantly ($P<0.05$) higher TN than the cultivation with high light intensity. As for TP, the data show no significant differences ($P>0.05$) between groups under 180 µmol/m2/s and 500 µmol/m2/s for both formate feeing rates.

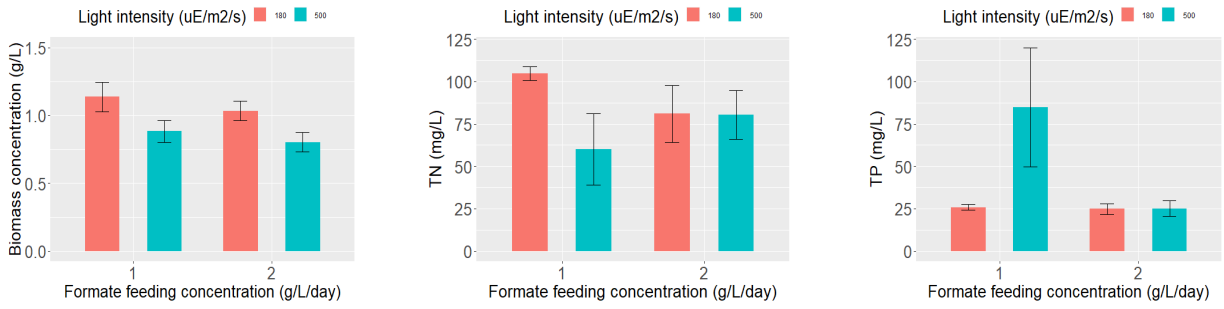


Figure 3.2 Comparison of biomass accumulation under different formate concentrations and light intensities

a. Comparison of biomass concentration under different conditions; b. Comparison of total nitrogen under different conditions; c. Comparison of total phosphorous under different conditions

Table 3.1 Effects of light intensity and formate feeding rate on cultivation parameters

Factors and levels	Light intensity (μmol/m²/s)		Formate feeding rate (g/L/day)	
	180	500	1	2
Biomass concentration	P<0.05		P<0.05	
Total nitrogen	P<0.05		P>0.05	
Total phosphorus	P<0.05		P<0.05	

3.2 Effects of formate on microbial dynamics of the algal assemblage

Real-time polymerase chain reaction (qPCR) was conducted to quantitively compare bacterial and fungal communities in the assemblage (Figure 3.3). The average C_q values were 26 ±2 and 19±2 for eukaryotes and prokaryotes, respectively. The difference in C_q value suggests that eukaryotes have much lower expression levels of target genome compared to prokaryotes, as a higher Cq value indicates a lower amount of amplified target gene. The result indicates that fungal community is minuscule in contrast to bacteria community.

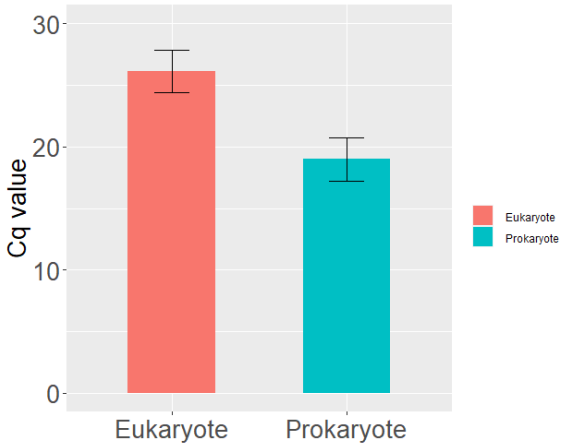


Figure 3.3 Comparison of Cq values of two sequencing methods

Microbial analysis was then conducted on both bacterial and fungal communities to reveal the interaction between the algal assemblage and cultivation conditions (Table 3.2). The gene sequences was rarified at 150,000 reads. The rarefaction analysis indicates sufficient sample coverage (Figure S3.3a). A rank abundance curve between 10 and 689 species shows that

gene sequences were evenly distributed (Figure S3.3b). Shannon's (H) and Pielou's (J) indices were determined to assess alpha-diversity and evenness, respectively, within the microbial community of the case scenarios. As shown in Table 3.3, microbial communities exhibited stability with different light intensities and formate feeding rates. Both alpha-diversity and evenness remained no significant differences ($P>0.05$) between individual conditions. Similarly, there are no significant differences ($P>0.05$) on beta-diversity between individual conditions (Table 3.4).

The average relative abundance of microalgae and bacteria at domain level were 73±4% and 27±4% respectively. As seen in Figure 3.4a, there were no significant differences ($P>0.05$) at domain level between individual conditions. At phylum level (Figure 3.4b), *Proteobacteria* and *Bacteroidetes* were two major bacterial groups in the cultivations. Evidences were reported that *Proteobacteria* were able to decompose organic carbon in the ecosystem [123]. The result suggests formate feeding is possibly favorable for *Proteobacteria* group to grow and generate soluble carbon for microalgal growth.

Microbial analysis of fungal community shows that *Ascomycota* and *Eukarya* were two main fungal communities in the assemblage with average relative abundances of 58±4% and 42 ±4%, respectively, at phylum level (Figure 3.5a). Th statistical analysis concludes that different light intensities and formate feeding rate did not significantly affect ($P>0.05$) the relative abundance of Ascomycota and Eukarya. It has been reported that Ascomycota was capable of decomposing organic carbon in the environment [124]. In Figure 3.5b, *Cladosporiaceae* is a family of filamentous fungi belonging to the Ascomycota phylum. Some species are plant pathogens, causing diseases in crops and ornamental plants [125]. The presence of *Sordariomycetes_unclassified* also was reported by some research that it can lead to diseases in

crops and plants [126]. This may prevent other invading species from growing in the system, which helps the consortia stay stable.

Table 3.2 Four different cultivation conditions

Conditions	Light intensity (μmol/m^2/s)	Formate feeding rate (g/L/day)	Harvesting amount (v/v)	Culture time (day)
Condition 1	180	1	30%	14
Condition 2	500	1	30%	14
Condition 3	180	2	30%	14
Condition 4	500	2	30%	14

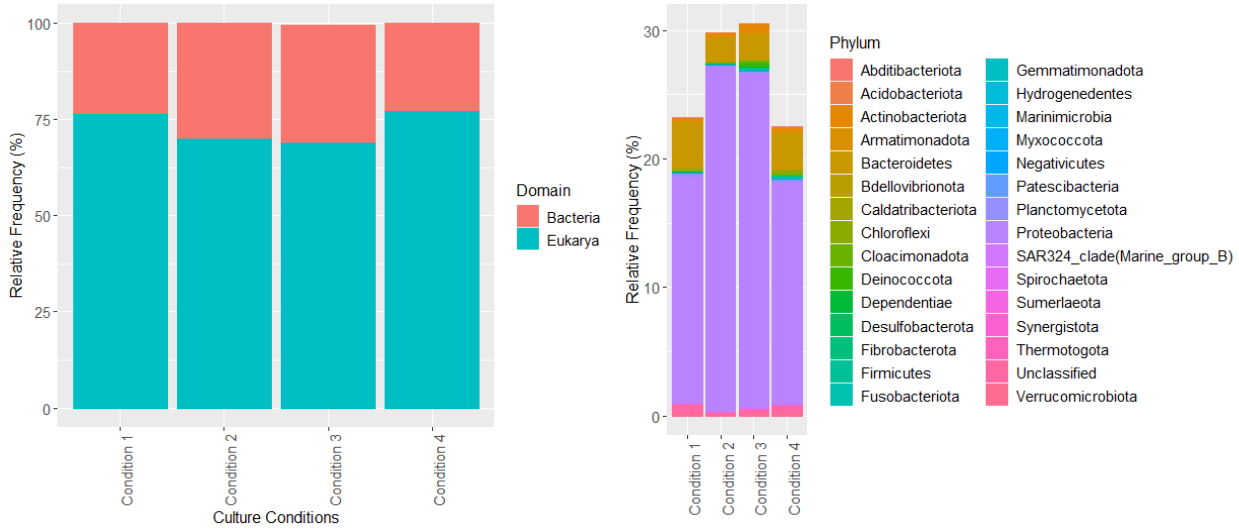
Table 3.3 One-way analysis of variance of light intensity and formate feeding rate on alpha-diversity and evenness of microbial communities

[a] H: Shannon's index which indicates the diversity of the microbial community. [b] J: Pielou's index which indicates the evenness of the microbial community

Parameter		Light intensity	Formate feeding rate
H[a]	Degree of freedom	1	1
	Sum squared	0.011	0.010
	P	0.616	0.627
J[b]	Degree of freedom	1	1
	Sum squared	0.0000384	0.0000019
	P	0.852	0.967

Table 3.4 Permutational analysis of variance of light intensity and formate feeding rate on beta-diversity and evenness of microbial communities

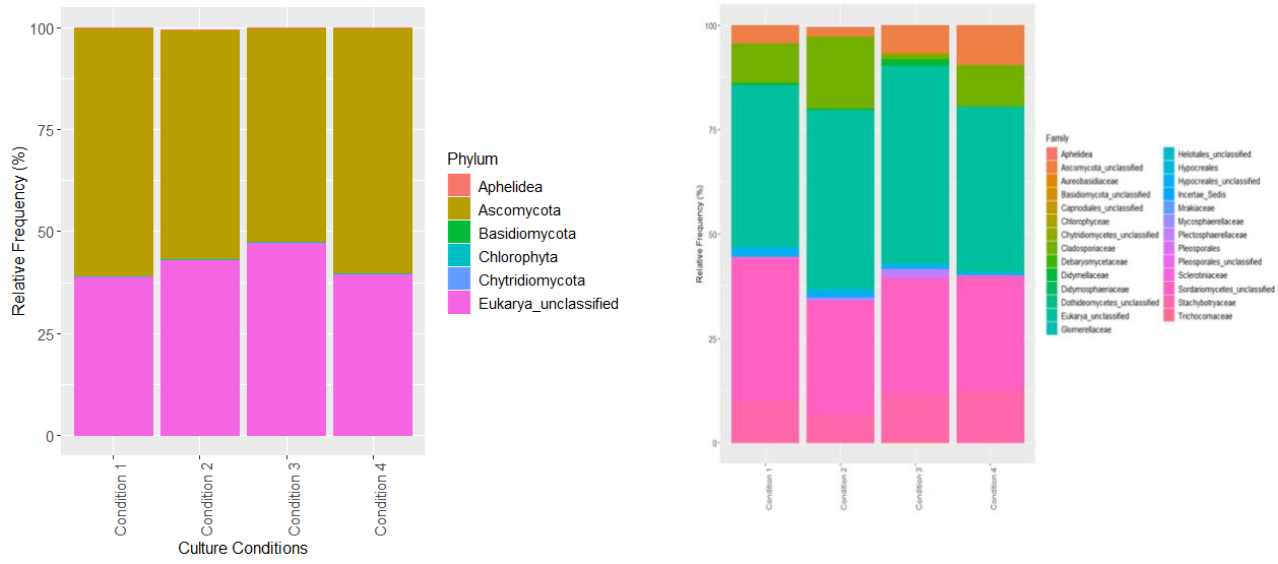
Parameter		Light intensity	Formate feeding rate
Beta-diversity	Degree of freedom	1	1
	Sum squared	0.105	0.127
	P	1	0.333



a b

Figure 3.4 Relative abundance of eukarya and bacteria communities in the algal assemblage

a. Relative abundance at domain level; b. Bacteria relative abundance at phylum level.



a b

Figure 3.5 Relative abundance of fungal communities in the algal assemblage

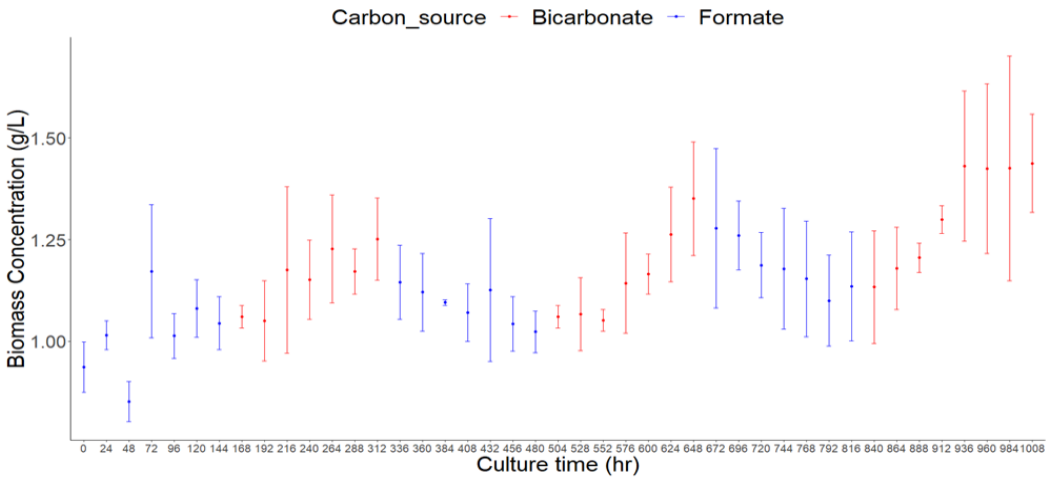
a. Fungi relative abundance at phylum level; b. Fungi relative abundance at family level

3.3 A new cultivation method of altering formate and bicarbonate to enhance biomass accumulation.

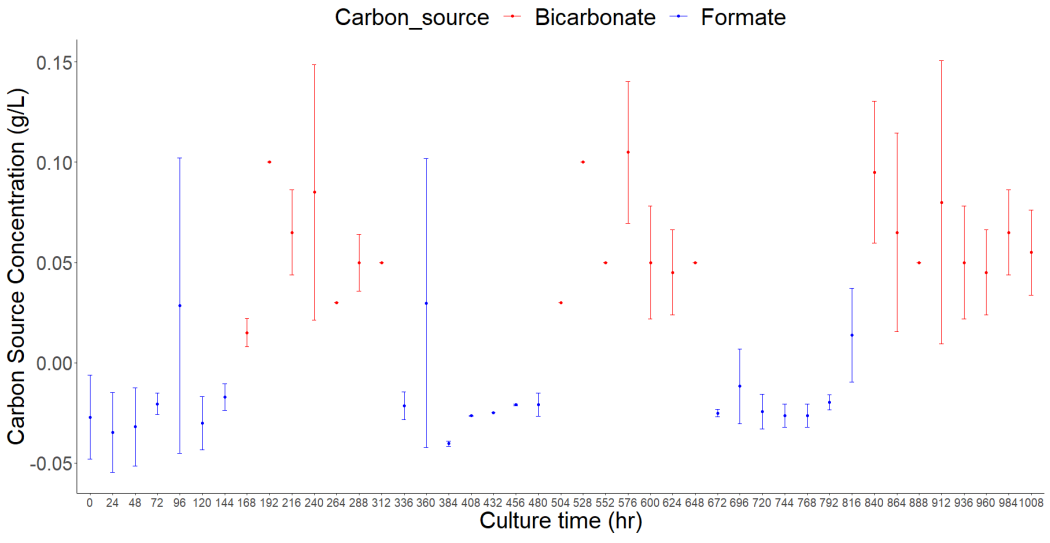
The effects of alternating carbon sources on biomass accumulation and microbial communities are shown in Figure 3.6. The average biomass concentration during formate feeding phases was 1.10±0.12 g/L, and the average biomass concentration during bicarbonate feeding phases was 1.21±0.16 g/L (Figure 3.6a). Correspondingly, formate and bicarbonate concentrations were monitored during the cultivation (Figure 3.6b). Both carbon sources were effectively consumed by the assemblage during the cultivation. Compared to bicarbonate phases, formate feeding phases exhibited slightly lower biomass concentration during the cultivation. Since formate has slower incorporation rate than bicarbonate in microalgal biomass [127], which could possibly lead to lower biomass accumulation. Table 5 indicated effects of carbon source on biomass, TN, and TP. Alternating carbon source had significant influence ($P<0.05$) on both biomass concentrations and TP concentrations.

Microbial community information was collected during the culture to study the effects of two different carbon sources on microbial communities (Figure 3.7). The average relative abundances of microalgae under formate feeding and bicarbonate feeding were 81±2% and 83 ±4% respectively. The average relative abundances of bacteria under formate feeding and bicarbonate feeding were 19±2% and 17±4% respectively. There was no significant difference between these two groups regarding the domain level ($P>0.05$). When looking at bacteria at phylum level. The dominant bacteria at phylum level were *Proteobacteria*. The relative abundance of *Proteobacteria* in the formate feeding and bicarbonate groups were 12±1% and 9 ±1% respectively. The formate group has significantly higher relative abundance of *Proteobacteria* than the bicarbonate group ($P<0.05$). Microbial analysis was also applied for

fungal community in the culture system (Figure 3.8). There was no significant difference (P>0.05) in fungal community at phylum level between two different carbon source groups. qPCR was applied to investigate the relative number of prokaryotes and eukaryotes. As shown in Figure 3.9, fungal community is a small portion in the culture system compared to prokaryotic communities for the alternative carbon cultivation as well.



a

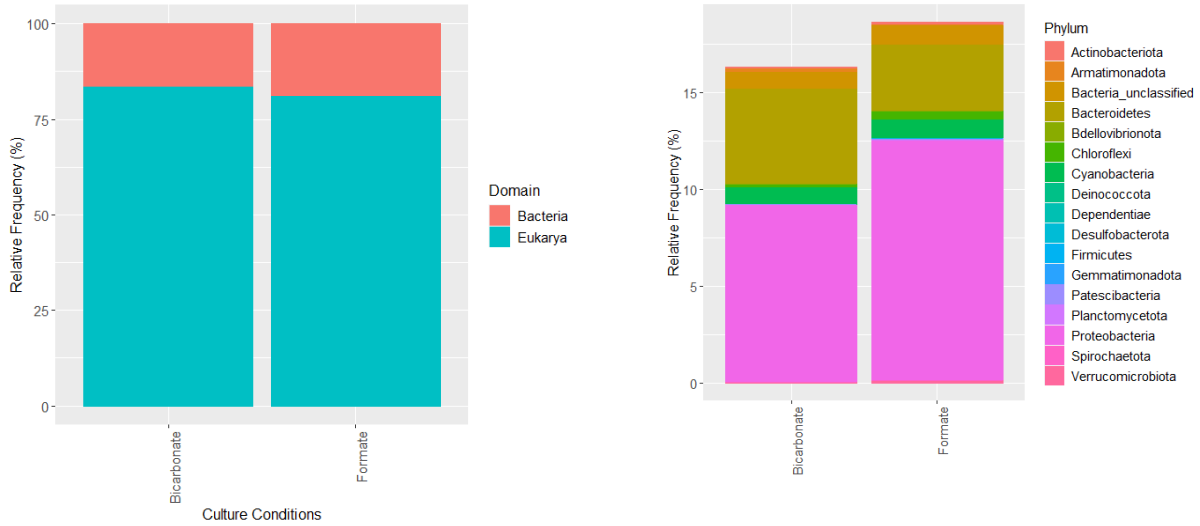


b

Figure 3.6 Time course of formate and bicarbonate utilization and biomass accumulation
a. Biomass concentration during the cultivation. b. Formate and bicarbonate concentration during the cultivation.

Table 3.5 Effects of carbon source on cultivation parameters

Factors and levels	Carbon source	
	Formate	Bicarbonate
Biomass concentration (g/L)	P<0.05	
Total nitrogen (mg/L)	P>0.05	
Total phosphorus (mg/L)	P<0.05	



a b

Figure 3.7 Relative abundance of eukarya and bacteria communities during the cultivation

a. Relative abundance at domain level; b. Bacteria relative abundance at phylum level

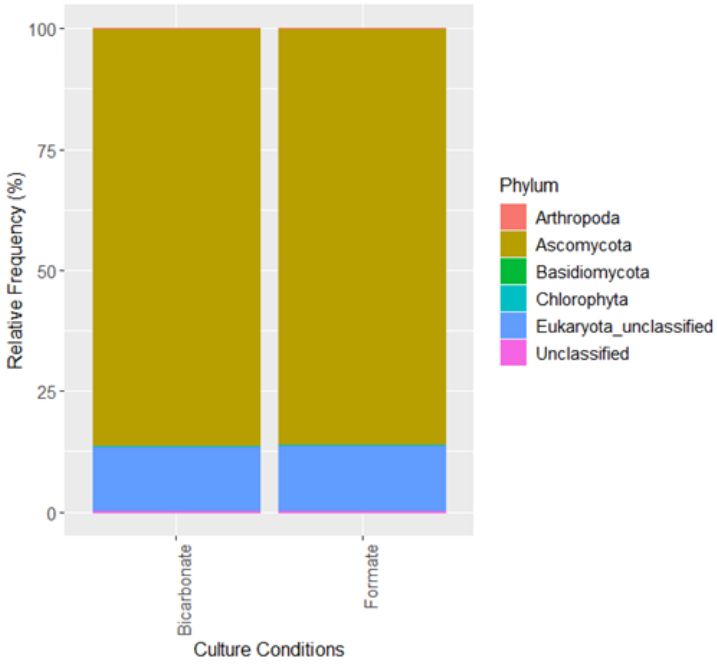


Figure 3.8 Relative abundance of fungal communities during the cultivation at phylum level

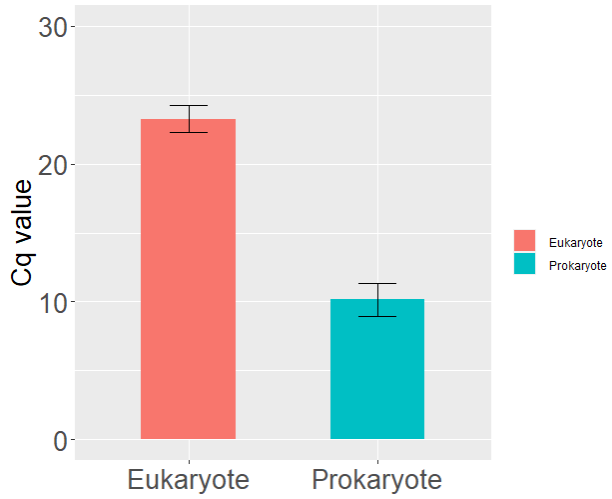


Figure 3.9 Comparison of Cq values between Eukaryote and Prokaryote during the alternative carbon cutlivation

3.4 Comparison of two feeding techniques

To elucidate the effects of two feeding methods of formate-only and alternative carbon sources (formate and bicarbonate) on the cultivation, statistical analysis was conducted (Figure 3.10). Data of Condition 1 from the formate-only cultivation was selected to compare with the data from the alternative carbon cultivation (the last two weeks, from 672 to 1008 hours). The

average biomass concentration under the formate-only cultivation was 1.13±0.1 g/L. The average biomass concentration under the alternative carbon cultivation was 1.26±0.2 g/L. As shown in Table 6, biomass concentration of the alternative carbon cultivation was significantly higher (P<0.05) than that from the formate-only cultivation. The result indicated that alternating formate and bicarbonate as carbon sources enhances biomass accumulation of the microalgal cultivation. Correspondingly, the alternative carbon cultivation had significantly lower (P<0.05) TN concentration than formate culture group.

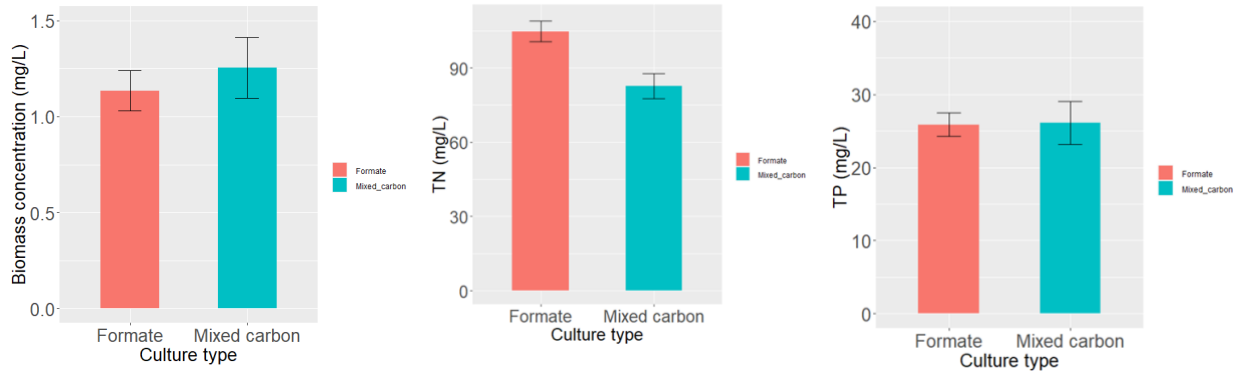


Figure 3.10 Comparison of two culture methods on cultivation parameters

a. Comparison of biomass concentration across two culture methods; b. Comparison of total nitrogen across two culture methods; c. Comparison of total phosphorous across two culture methods.

Table 3.6 Effects of culture techniques on cultivation parameters

Factors and levels	Culture type	
	Formate	Mixed carbon
Biomass concentration (g/L)	P<0.05	
Total nitrogen (mg/L)	P<0.05	
Total phosphorus (mg/L)	P>0.05	

4. Conclusion

Formate as a carbon source for microalgal cultivation has been investigated through culture performance and microbial community analysis. Results showed that a formate feeding rate of 1 g/L/day, under a light intensity of 180 μmol/m^2/s, outperforms other conditions in the study. A new microalgal cultivation technique of alternative carbon feeding showed promising results with the highest biomass concentration of 1.6 g/L during the cultivation. Statistical analysis indicated that this new technique significantly enhanced the cultivation performance of biomass accumulation. This study highlights the efficacy of formate as a carbon source for microalgal culture and provides new insights for microalgal cultivation through the alternative carbon feeding strategy.

CHAPTER 4: EFFECTS OF FORMATE ADDITION ON CONTINUOUS ALGAL CULTIVATION OF CO_2 CAPTURE FROM POWER PLANT FLUE GAS

1. Introduction

Global warming brings the long-term increase in Earth's average surface temperature, primarily caused by human activities such as the burning of fossil fuels, deforestation, and land-use changes. The primary cause of global warming is the increase in greenhouse gases in the atmosphere, mainly carbon dioxide (CO_2). Carbon capture technologies are crucial to control greenhouse gas amount in the atmosphere and mitigate the global warming impacts. Regarding gas emission at global scale, CO_2 accounts for 76% of greenhouse gas emissions [128]. In CO_2 emission, around 85% of carbon emission comes from fossil fuel and industrial process [128]. According to the data, countries with the highest CO_2 emissions in 2014 were China, the United States, the European Union and India, accounting for 30%, 15%, 9% and 7% of global CO_2 emissions, respectively [129]. As a result, carbon capture for fossil fuels and power industry is necessary.

Microalgal cultivation for carbon capture is considered as an innovative solution to global warming. Through photosynthesis, microalgae can absorb solar energy and utilize CO_2 to grow and release O_2 during the process [130]. Compared to terrestrial plants, microalgae have higher photosynthetic efficiency with lower arable land requirements and water consumptions [131]. Beyond carbon capture, microalgae biomass is a great source of value-added products, such as biofuel [132] and bio-fertilizers [133]. These value-added products can create economic benefits for developers and lower the cost of culture system, which is advantageous to carry out large-scale commercialization.

Microalgal cultivation is commonly conducted in either algae photobioreactor or open pond. Differences between these two systems including: (1) Scale: Photobioreactors are typically smaller and more controlled systems, while open ponds can cover larger areas and can be more

difficult to control [134]. (2) Productivity: Photobioreactors can have higher microalgal densities and faster growth rates compared to open ponds, which can result in higher biomass and product yields per unit of area and time [135]. (3) Nutrient and light availability: Photobioreactors can provide more controlled and consistent nutrient and light conditions for microalgal growth, which can result in higher quality and more consistent biomass and product yields [136]. In contrast, open ponds may have variable nutrient and light availability, depending on factors such as weather conditions and water quality [137]. (4) Capital and operational costs: Photobioreactors can be more expensive to build and operate compared to open ponds, due to their more complex design, materials, and monitoring systems [138]. However, they can also be more efficient in terms of resource use and product yields [139]. (5) Environmental impacts: Open ponds can have a larger environmental footprint and potential impacts on local water quality and biodiversity, due to the use of large areas of land and water [140]. In contrast, photobioreactors can be designed to minimize environmental impacts and can be used in urban or indoor settings.

Microalgal-bacterial systems has been reported by many researches in applications such as CO_2 capture [140], wastewater treatment [141], biofuel production [142]. In a microalgal-bacterial system, the microalgae provide oxygen and organic matter through photosynthesis, which can support the growth and metabolism of bacteria. In return, the bacteria can provide nutrients and growth factors to the microalgae, as well as help to stabilize the microbial community and reduce harmful contaminants [143]. Microalgal-bacterial systems can support a high diversity of microorganisms, which can enhance their resilience and adaptability to changing conditions [144]. This diversity can also lead to a lifted production of various valuable metabolites, such as lipids [145], pigments [146], and enzymes [147]. Evidence showed that

microalgal-bacterial system exhibited excellent efficiency in carbon removal within flue gas sparged system for sewage treatment [148].

Carbon sources for microalgal culture can be classified into organic carbon source and inorganic carbon source. Organic carbon includes glucose, acetate, and lactose. Studies showed that organic carbon can largely improve biomass production in the coculture of microalgae and bacteria[149]. Other research indicated that organic carbon could enhance carbon sequestration and nutrient removal in wastewater treatment [17]. Inorganic carbon source such as CO_2 [150], carbonate [151], bicarbonate [152] were also reported in microalgal cultivation. Evidences were found that bicarbonate could boost microalgal growth rate and leaded to higher lipid production [153]. Bicarbonate was also considered as a helpful carbon source to enhance lutein biosynthesis in the culture of green microalgae *Chlorella pyrenoidosa* [154]. However, few research was reported about application of formate as carbon source or addition in microalgal culture for carbon capture. Studies on effects of formate on microalgal-bacterial system still remain vague. The objectives of this study are to elucidate effects of formate addition on algal cultivation of carbon capture from flue gas, to understand effects of formate on microbial community of the algal-bacterial assemblage, and to demonstrate performance of the symbiotic system on carbon capture.

2. Materials and methods

2.1 Algal assemblage and cultivation system

The algal assemblage containing a selected microalga *Chlorella sorokiniana* MSU from the Great Lakes region and several bacteria (mainly *Bacteroidetes* and *Proteobacteria* etc.) was continuously cultured in flasks on Tris-Acetate-Phosphate (TAP) medium at room temperature under constant fluorescent light to use as the seed for the algae photobioreactors (APBs). Modified liquid TAP medium (without acetic acid and tris base) was used for microalgal

cultures, which contains 7.5 mmol L^{-1} of NH_4Cl, 0.34 mmol L^{-1} of $CaCl_2 \cdot 2H_2O$, 0.4 mmol L^{-1} of $MgSO_4 \cdot 7H_2O$, 0.68 mmol L^{-1} of K_2HPO_4 (anhydrous), 0.45 mmol L^{-1} of KH_2PO_4 (anhydrous), 0.09 mmol L^{-1} $FeCl_3 \cdot 6H_2O$, and 1ml TAP trace elements solution. The modified TAP medium was unsterilized. The microbial community was analyzed before seeding the photobioreactors.

2.2 Pilot algal cultivation system

The pilot-scale APB was located at the T.B. Simon Power Plant at Michigan State University. The effective volume of the pilot-scale APB is 100 L. The pilot-scale APB configuration and operating mechanism were described in a previous study [155]. Various cultivation conditions were carried out to test the effects of formate addition: formate feeding rate (0.25 g/L/day, 0.5 g/L/day, 0.75 g/L/day, 1 g/L/day), harvesting amount (30% and 50%). The culture temperature was maintain at 20 ± 2℃. Fresh water was used to refill the reactor after harvesting. Nutrients and trace elements were supplied based on modified TAP medium to sustain N/P molar ratio at 6.65. The control of continuous culture on saturated CO_2 was carried out using the pilot-scale APB under identical conditions. In total seven scenarios of study were tested to collect data which was used to study the effects of formate on continuous microalgal culture.

2.3 Chemical analysis

Samples were analyzed for dry biomass weight, pH and nutrient (total nitrogen (TN), total phosphorus (TP), nitrate (NO_3-N) and ammonia (NH_3-N)) concentrations. Algal biomass was pelleted for dry weight measurement using a Thermo Electron Corporation IEC Centra CL2 Centrifuge at 3800 rpm for 5 minutes. Biomass was washed once and resuspended using deionized water, and then dried at 105°C for 24 hours. Sample pH was measured using a pH

meter (Fisherbrand™ accumet™ AB15 + Basic, Fisher Scientific Co., Pittsburgh, PA). Nutrient concentrations were tested in the liquid supernatant using nutrient test kits (HACH Company, Loveland, Colorado) equivalent to EPA methods (hach.com/epa). Algal biomass composition was analyzed using the standard forage analysis method.

Formate concentration in culture media was determined by high performance liquid chromatography (HPLC) (Shimadzu Corp., Kyoto, Japan) equipped with an analytical column (Aminex HPX-87H, Bio-Rad Laboratories, Inc., Hercules, CA) and a refractive index detector (Shimadzu Corp., Kyoto, Japan). The mobile phase was 0.005 mol/L sulfuric acid at a flow rate of 0.6 mL/min. The oven temperature was set at 65 °C. Bicarbonate concentration of algal samples in kinetic study was determined by the alkalinity test kit (HACH Company, Loveland, CO).

2.4 Microbial community analysis

Samples (1 mL) collected for DNA analysis were kept frozen at -20°C until analysis. To remove nutrient media, the algae sample was centrifuged using an Eppendorf 5416R centrifuge at 10,000 rpm for 5 min and the supernatant was discarded. The remaining pellet was used for DNA extraction using the DNeasy® PowerSoil® Kit (Qiagen, Germany). DNA was eluted with 100 μL of 10 mM Tris-HCl (pH 8.5) and the concentration and purity determined using a NanoDrop Lite spectrophotometer (Thermo Fisher Scientific, USA). The DNA extracts were stored at -80°C for several weeks and then used for polymerase chain reaction (PCR) and Illumina DNA sequencing.

Illumina sequencing was performed for the 16S rRNA gene to assess the bacterial community. Prior to PCR, extracted DNA samples were diluted 10x due to high DNA concentrations. The PCR conditions were as follows: 1.0 μL DNA template (10x diluted), 0.5 μL

of 100 μM forward primer (IDT, Pro341F), 0.5 μL of 100 μM reverse primer (IDT, Pro805R), 12.5 μL 2x Supermix (Invitrogen, USA), and 10.5 μL PCR grade water. The PCR program used for all assays is as follows: 96°C for 2 min, followed by 30 cycles of 95°C for 20 s, 52°C for 30 s, and 72°C for 1 min, and a final elongation period of 72°C for 10 min. After PCR, samples were diluted to normalize DNA concentrations within a range of 5-10 ng/μL. DNA concentration was determined using the PicoGreen® dsDNA quantitation assay (Invitrogen, USA) and Fluostar Optima microplate reader (BMG Labtech, Germany). The PicoGreen® conditions were as follows: 95 μL 1x TE buffer solution, 100 μL 1:200 diluted PicoGreen® reagent, 5 μL DNA template. Samples with known DNA concentrations were also prepared for standard curve generation. Illumina library preparation and sequencing were performed at the Michigan State University Genomics Laboratory, East Lansing, USA.

The 16S rRNA gene sequencing was also used to determine *C. sorokiniana* in the assemblage [88]. It has been reported that *Cyanobacteria* have 85-93% of 16S rRNA gene sequences similar to *C. sorokiniana* [93, 94], while, color, shape, and size of both species are very different [95]. Therefore, after microscopic imaging verification of each sample, the *Cyanobacteria* sequence was interpreted as microalga *C. sorokiniana* for all samples.

2.5 Mass and energy balance

A mass balance analysis was conducted on a 1 m3 APB unit to compare formate/CO_2 and CO_2 cultivations. The APB unit was described in a previous study [156]. The envisioned APB unit has 10 tubes that each of which is the same size as the pilot APB unit. The tubes share the same upflow tube, so that the 1 m3 APB could have a similar performance as the pilot unit. The gas transfer in the APB is operated through airlift. An excessive amount of CO_2 was pumped into the APB unit to ensure CO_2 saturation in the culture medium. The harvesting amount of the

medium is determined by the selected ratio from the previous task. The exact amount of the feeding medium with the same nutrient composition as the pilot unit is fed to the APB reactor. Carbon capture efficiency is calculated by the carbon mass (g) in the harvested biomass divided by the carbon mass (g) dissolved in the medium using the data of carbon content in absorbed CO_2 and formate, algal biomass productivity, carbon content of algal biomass.

An energy balance analysis was conducted for the 1 m^3 APB unit based on the data from the 0.1 m3 pilot APB as well. A 5 kW centrifuge runs 3 min/day to collect algal biomass. The water pump transferring the medium to the centrifuge is a 0.5 kW pump with an average running time of 3 min/day. The feeding pump to feed the APB is a 0.5 kW pump with an average running time of 8 min/day. The flue gas pump is a 0.1 kW unit with an average running time of 24 hours/day.

2.6 Statistical analysis

All data collected was analyzed using the statistical tools of R (version 3.6.3). In order to determine whether a parametric or non-parametric test was necessary, the data were first tested for normality and equal variance using a Shapiro-Wilk's test and an F-test, respectively. Data that were normal with equal variance were tested using an analysis of variance (ANOVA) and a Tukey test was used when applicable to compare individual factors. Data with non-normal distribution and unequal variance were tested using the Kruskal-Wallis test. All tests were performed with a significance value of $\alpha = 0.05$. Microbial community analysis was completed using Vegan, ggplot2, phyloseq, and MASS R libraries. Taxonomic/phylogenetic data was analyzed in order to graph relative abundances of samples.

3. Results and discussion

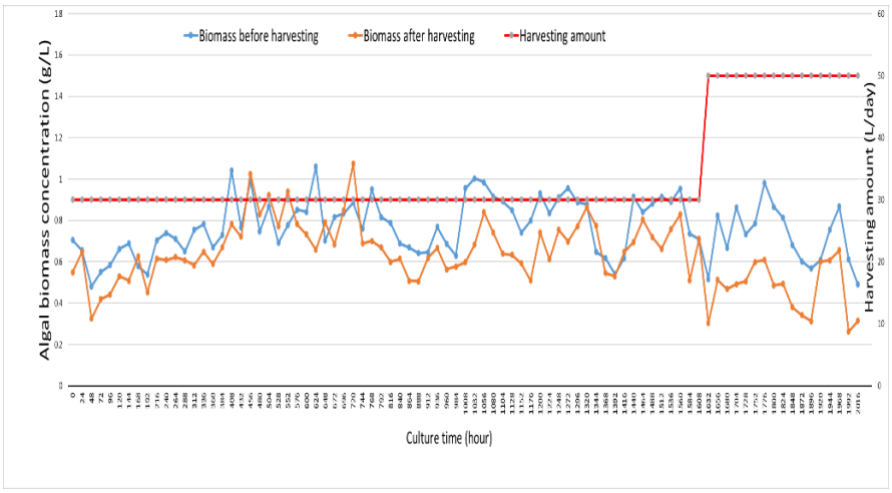
3.1 Effects of formate on algal growth and CO_2 capture

The algal assemblage was fed by five different sodium formate concentrations (0g/L, 0.25g/L, 0.5 g/L, 0.75 g/L, 1 g/L) to study the performance of carbon capture and biomass accumulation during steady state continuous cultivation (Figure 4.1a). Light intensity of 180 μmol/m^2/s was used for the continuous culture. Control group study was conducted on flue gas from power plant using the same light condition. Figure 4.1 showed that the biomass concentration under various formate concentration conditions. Data showed that biomass concentration of the control culture was 0.62±0.10 g/L. Under 0.5g/L formate concentration, average biomass concentration maintained at 0.81±0.13g/L. Under 1 g/L formate concentration, biomass concentration reached 0.85±0.10 g/L. The effect of formate addition on algal biomass was studied. Figure 4.1b compared the biomass concentration differences between control and various sets of formate group ranging from 0.25 to 1 g/L. Results showed significant increases ($P<0.05$) in biomass concentration were observed under 0.5 g/L, 0.75 g/L, and 1g/L formate concentration, respectively, compared to the control and the culture with the formate addition of 0.25g/L (Table 4.1). Figures 1c and d also showed the TN (total nitrogen) and TP (total phosphorus) consumption of five different groups. There are no significant differences ($P>0.5$) on phosphorus consumption among five cultures (Table 4.1). As for total nitrogen consumption, the culture with the formate addition of 0.75 g/L showed significantly less consumption compared to other four conditions ($P<0.5$) (Table 4.1). There are no significant ($P>0.05$) differences among these five groups in terms of biomass productivity (Table 4.1).

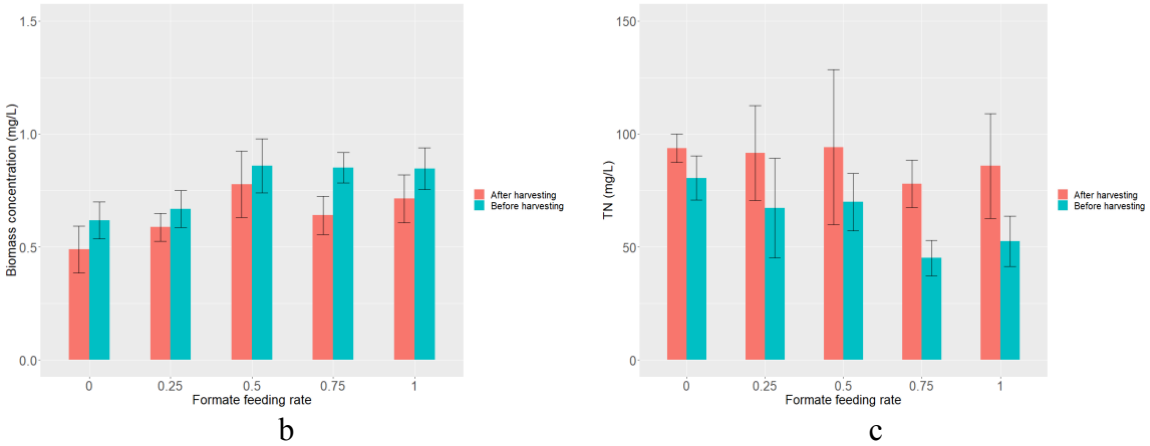
The effect of harvesting rate on algal biomass was also compared under 1g/L formate addition condition. Biomass concentration under 30% and 50% harvesting rate reached 0.85±

0.09 and 0.82±0.10 respectively. Total nitrogen consumptions of two different harvesting rates were 33.1±22.5 and 35.4±17.6. As shown in Table 4.1, different formate concentration had significant differences among biomass concentrations, TN concentrations and TP concentrations in samples. Similarly, two varying harvesting rate also showed significant differences in TN concentrations. However, there were no significant differences in biomass concentrations and TP concentrations under two harvesting rate.

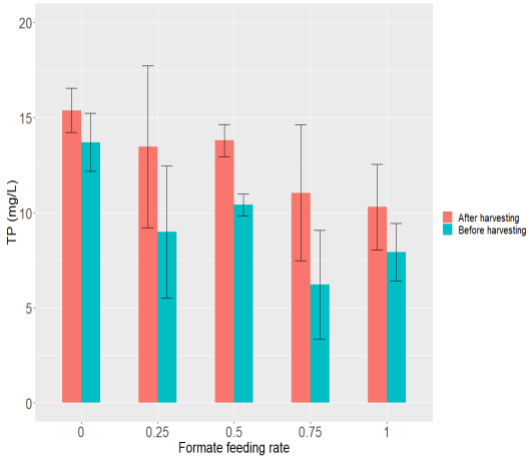
According to the results of statistical analysis of the microalgal growth with formate addition, the cultivation conditions with 0.5 g/L of formate addition were selected for the mass and energy balance analysis of the cultivation system in Section 3.3



a



b c



d

Figure 4.1 Algal cultivation with formate addition under different formate concentrations

a. Continuous algal cultivation on the flue gas from a power plant; b. Biomass concentration; c. TN concentration; d. TP concentration.

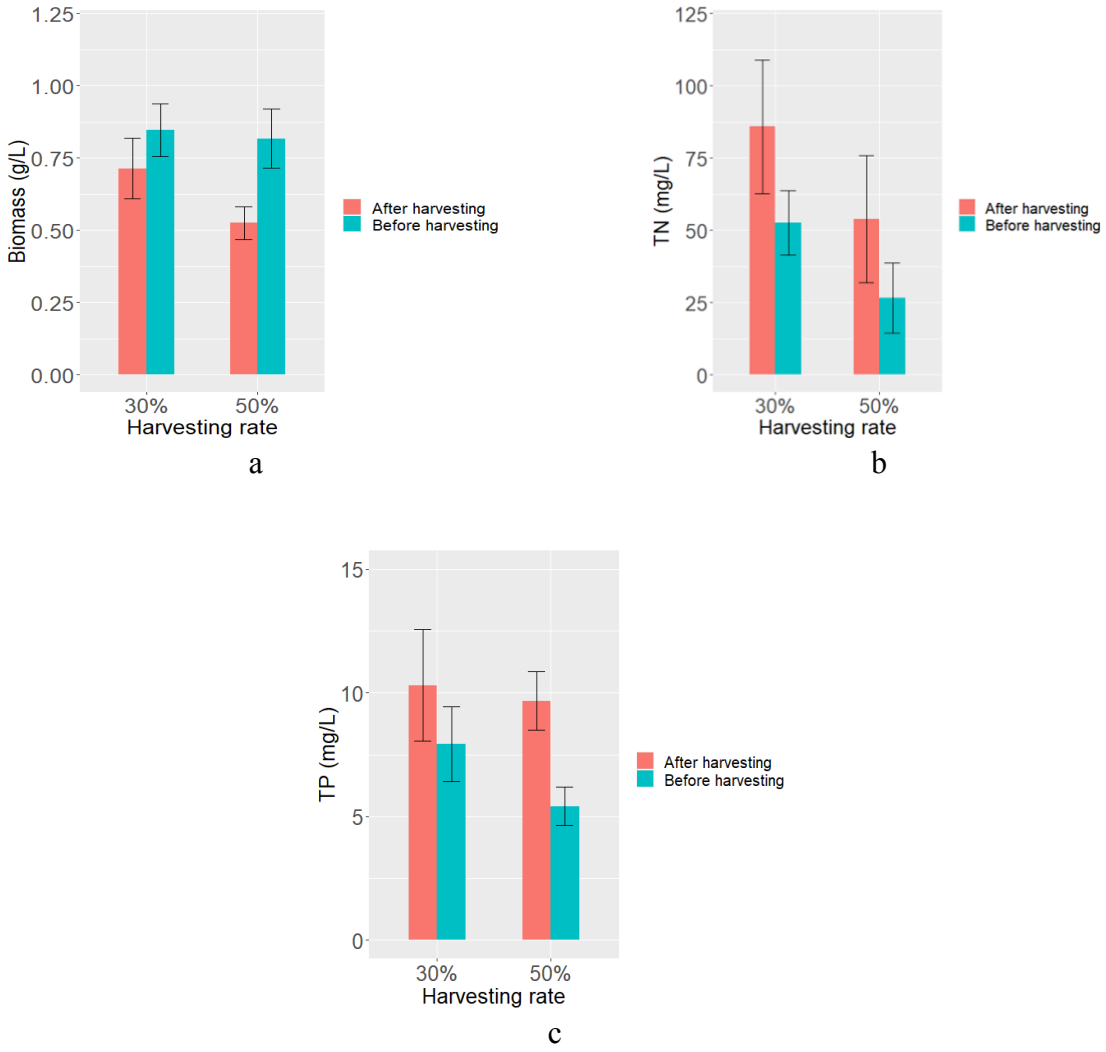


Figure 4.2 Culture parameters under 1g/L formate feeding rate with different harvesting amount

a. Biomass concentration; b. TN concentration; c. TP concentration

Table 4.1 ANOVA of formate concentration and harvesting rate on algal cultivation

Parameter	Formate concentration	Harvesting rate
Biomass	P<0.05	P>0.05
TN	P<0.05	P<0.05
TP	P<0.05	P>0.05

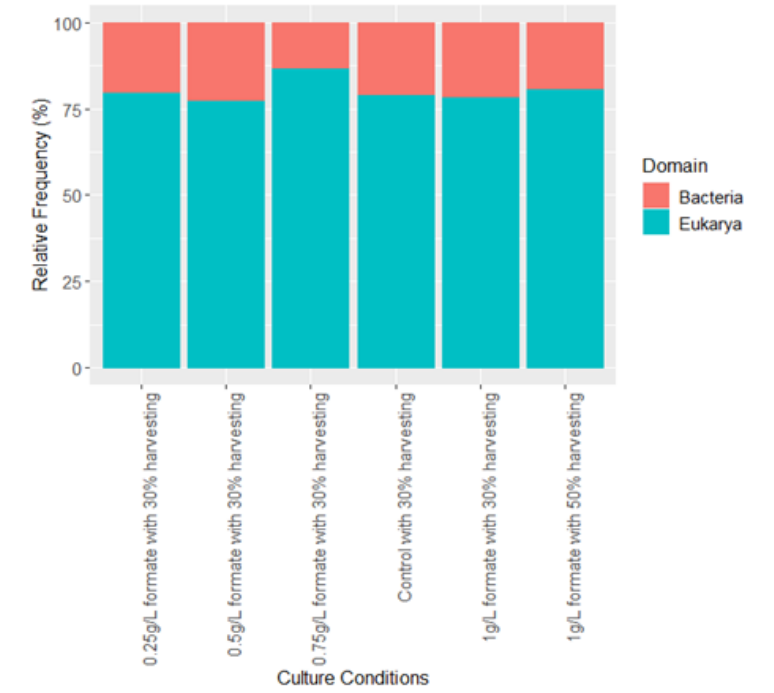
3.2 Effects of formate on algae and bacteria in the assemblage during the culture

Microbial community analysis was conducted based on DNA extracted from 16 samples. This analysis tended to illustrate the relationship between microalgae and bacteria under various operational conditions. The dataset of 16S rRNA gene sequences was rarified at 25,000 reads. Shannon's (H) and Pielou's (J) indices were determined to assess alpha-diversity and evenness, respectively. The relative average abundance of algae and bacteria were 80.43±2.91% and 19.57 ± 2.91% respectively at domain level without significant differences between each condition (Figure 4.2). This result indicated that C. sorokiniana is a robust algal strain over different formate feeding conditions.

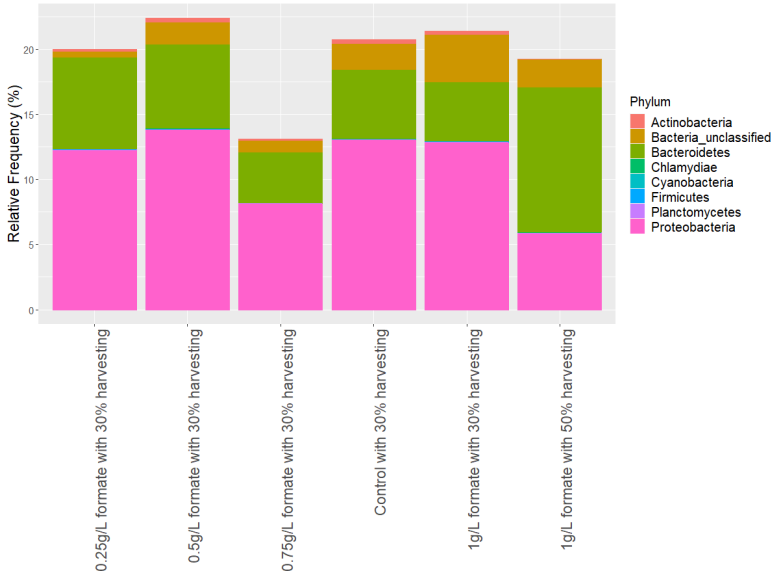
As seen in Figure 4.2b, Proteobacteria and Bacteroidetes are two major bacteria phyla in the microalgal-bacterial system. This information is consistent with reports on microbial community of pilot-scale microalgal cultivation system [157]. At the class level, Bacteroidetes communities remained relatively stable under different formate feeding rate and harvesting rate (Figure 4.2c). Moreover, in Proteobacteria communities, Alphaproteobacteria also remained stable throughout each condition. Research found that some species of Alphaproteobacteria such as those in Nitrobacter genus, can play the role of oxidizing nitrite to nitrate in the broth [158]. This result may indicate that Alphaproteobacteria can help microalgal cells utilize nitrogen nutrients in the solution and facilitate the formate utilization. Evidences were found that the presence of Alphaproteobacteria as the dominant bacteria with microalgae in the co-cultured system, which has strong ability to utilize ammonia and phosphate in the solution [159]. Notably, in Figure 4.2d, compared to control group, all formate addition groups have higher relative abundances of Gammaproteobacteria. It was reported that Gammaproteobacteria was major contributor of dark carbon fixation in the coastal sediments [160]. About 70-86% of dark carbon

fixation was accomplished by Gammaproteobacteria. This microbial information indicated that formate feeding may serve as carbon source for Gammaproteobacteria to grow, which resulted in higher abundances in formate-added groups than control group. It may show that an established way to use formate in the system as the carbon source for both microalgae and Gammaproteobacteria to grow, which needs further studies on its mechanisms.

One-way ANOVA was conducted to further investigate the effects of various formate feeding rate and harvesting rate on alpha-diversity and evenness of microbial communities during the culture. In Table 4.2, it shows that different formate feeding rates don't have significant differences in alpha-diversity and evenness among groups ($P>0.05$). Similarly, both 30% and 50% harvesting rate showed no significant differences in alpha-diversity and evenness of samples. This information demonstrates that this co-cultured system is stable under different culture condition. Overall, microbial community analysis suggests that Proteobacteria with its different subdivisions play an important role in nitrification and fixation for microalgae growth and formate utilization, thus making the algal-bacterial system stable across different culture conditions.



a

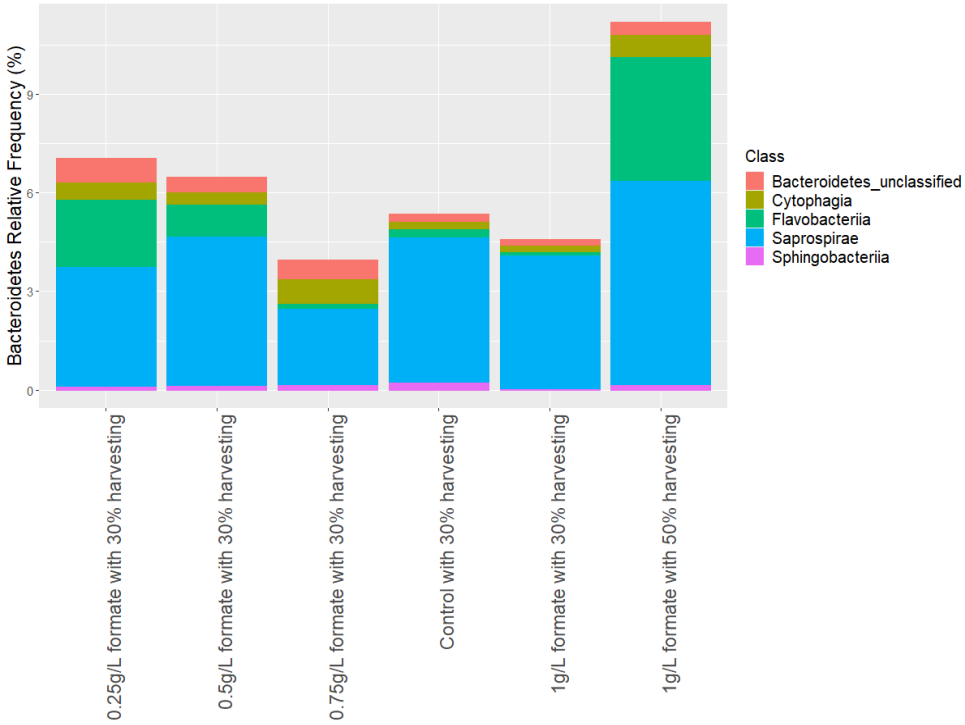


b

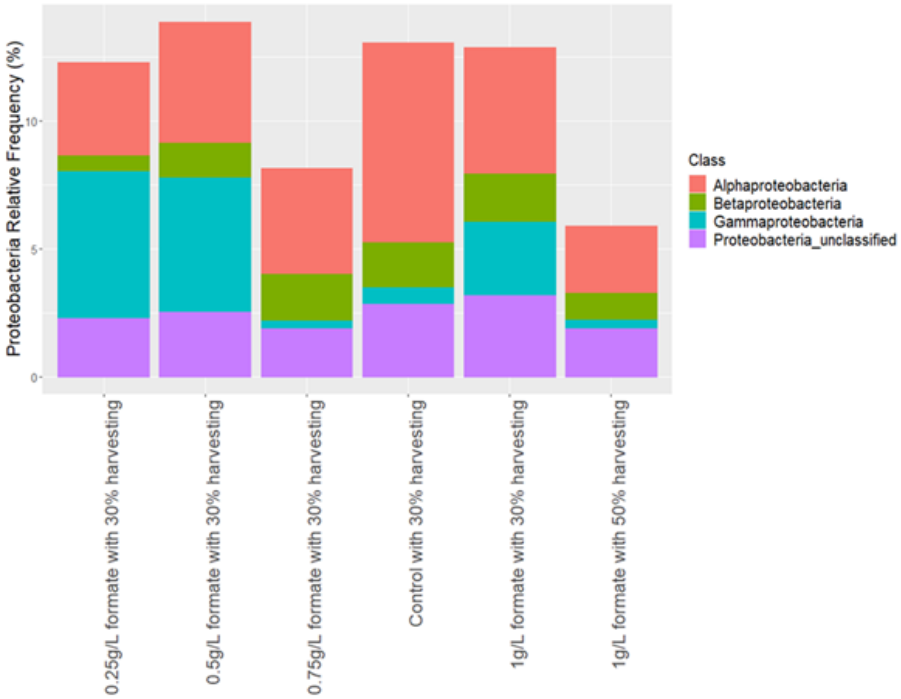
Figure 4.3 Relative abundance of microbial communities in the algal assemblage

a. Relative abundance at the domain level; b. Relative abundance of bacteria at the phylum level; c. Bacteroidetes relative abundance at the class level; d. Proteobacteria relative abundance at the class level

Figure 4.3 (cont'd)



c



d

Table 4.2 One-way analysis of variance of formate feeding rate and harvesting rate on alpha-diversity and evenness of microbial communities

[a] H: Shannon's index which indicates the diversity of the microbial community. [b] J: Pielou's index which indicates the evenness of the microbial community

Parameter		Formate feeding rate	Harvesting rate
H[a]	Degree of freedom	4	1
	Sum squared	0.08265	0.00057
	P	0.33	0.879
J[b]	Degree of freedom	4	1
	Sum squared	0.006098	0.000114
	P	0.392	0.804

3.3 Mass and energy balance of formate facilitated algal cultivation of CO_2 capture and utilization

Based on the experimental data, the algal cultivation with the addition of 0.5 g/L formate was selected to conduct the mass and energy balance analysis on an APB unit with 1 m^3 effective volume (Fig. 4.3). The control was algal cultivation without formate addition. Other than formate, both cultivations were under the same conditions. The control cultivation produces 0.93 kg wet biomass with 20% dry matter. The dry algal biomass contains 49.6% of carbon. The control cultivation then captures 0.34 kg of CO_2 per day. 300 kg/day of water with P and N nutrients is needed to replace 299 kg of effluent discharged from the centrifuge and replace water contained in the algal biomass. The cultivation with formate addition has a biomass production of 1.22 kg wet biomass per day with 20% dry matter. The dry algal biomass from the culture with formate addition has a carbon content of 45.4% (w/w). The cultivation captures 0.41 kg of CO_2 per day. The same amount of water with P and N nutrients is added to the culture.

As for carbon balance, the total carbons being dissolved in the medium are 0.12 and 0.17 kg/day for the control cultivation and cultivation with formate addition, respectively, based on the saturated CO_2 solubility and solidum formate. The carbons captured in the algal biomass are 0.09 and 0.11 kg/day for the control cultivation and cultivation with formate addition,

respectively. The corresponding carbon capture efficiencies (calculated by the carbon mass (g) in the harvested biomass divided by the carbon mass (g) dissolved in the medium) are 43.76 and 48.29%. The mass balance clearly indicated that formate addition improves carbon capture efficiency.

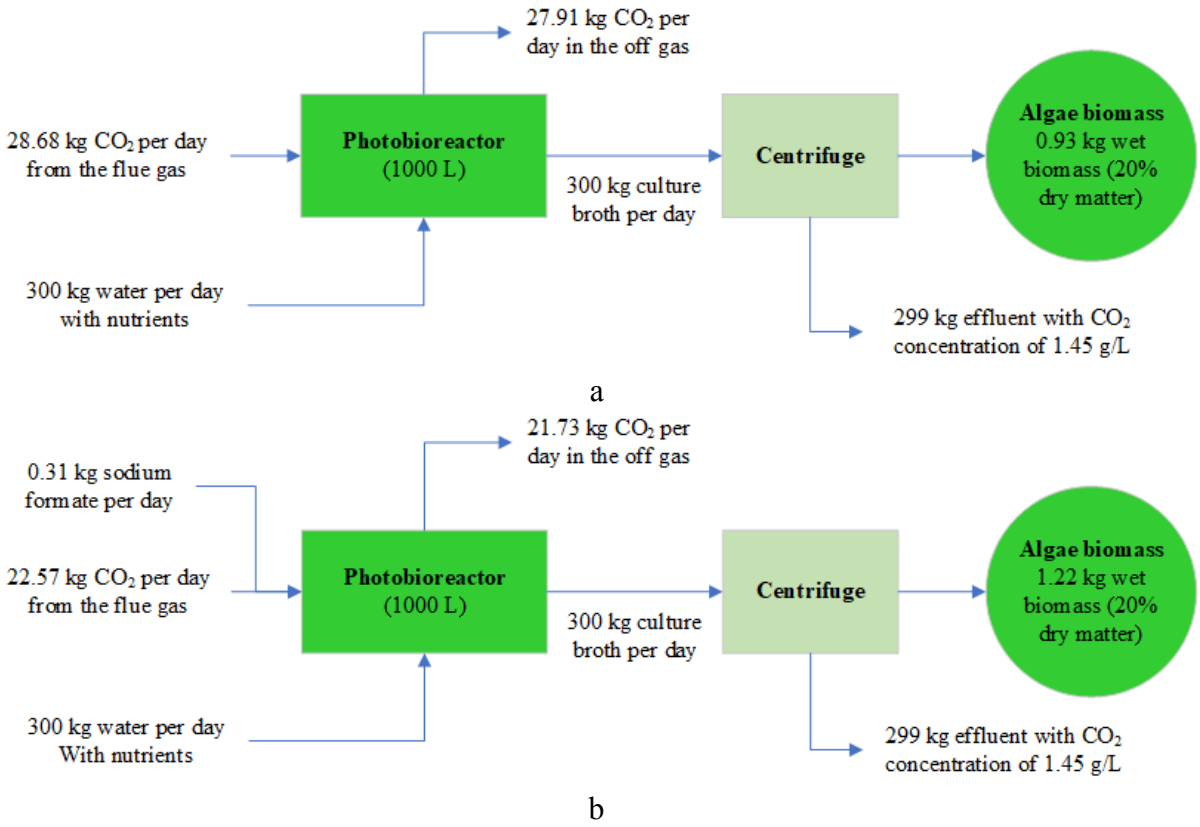


Figure 4.4 Mass balance of two cultivations in a 1000 L APB

a. Mass balance under CO_2 as the carbon source . b. Mass balance under CO_2 + sodium formate as the carbon sources

The energy balance demonstrated that the cultivation with formate addition demands much less energy (25 kJ/kg CO_2 captured) that the control cultivation (30 kJ/kg CO_2 captured) to capture CO_2 (Table 2). Among all of the unit operations, the most significant energy demand comes from the flue gas delivery due to the excessive amount of flue gas delivered to the cultivation system to ensure CO_2 saturation. The cultivation with formate addition and the control cultivation consumed 21 and 25 kJ/kg CO_2 captured, respectively. The centrifugation of

biomass collection was the unit operation with the second highest energy input. The cultivation with formate addition and the control cultivation demanded 2.2 and 2.7 kJ/kg CO_2 captured, respectively for the centrifugation. The pumps that fed the APB and transferred the medium to the centrifuge consumed the least energy because of their short daily operational time. Since formate addition improved CO_2 capture of the algal cultivation, the energy efficiency of the cultivation with formate addition is significantly better than the control cultivation.

Table 4.3 Energy balance of two cultivations in a 1000 L APB [a]

Energy demand for unit operations	The control cultivation (kJ/kg CO_2 captured)	The cultivation with formate addition (kJ/kg CO_2 captured)
Flue gas delivery [b]	25.4	21.1
Feeding water and nutrients [c]	0.7	0.6
Transferring the medium to the centrifuge [d]	0.3	0.3
Biomass collection by centrifugation [e]	2.7	2.2
Biomass drying [f]	0.5	0.7
Total energy demand	**29.6**	**24.9**

a. Data are based on the pilot operation
b. The flue gas pump with the capacity of 120 L/min requires 2.4 kWh/day (8.64 MJ/day) to deliver flue gas to the APB
c. The feeding pump requires 0.07 kWh/day (0.25 MJ/day) to feed the APB
d. The transferring pump requires 0.03 kWh/day (0.11 MJ/day) to transfer culture medium to the centrifuge
e. The centrifuge demands 0.25 kWh/day (0.9 MJ/day) to collect the wet biomass
f. The drying energy required is equal to heat energy to raise temperature to 100°C plus latent heat to remove water. The specific heat of the wet algal biomass is 3.8 kJ/kg °C. The initial temperature of the wet biomass is 20°C. The latent heat of vaporization of water at 100°C under standard atmospheric pressure is 2,257 kJ/kg. The thermal efficiency for the drum dryer is 70%. The energy needed to dry 1 kg wet algal biomass with 80% moisture for the drum dryer is calculated as follow: the dry energy = [1 kg x 3.8 kg/kg°C x (100-20) + 2,257 kJ/kg x (1 kg – 0.2 kg)]/70% = 2.88 MJ/kg dry algal biomass

4. Conclusion

This study concluded a new strategy to use formate to facilitate algal cultivation of carbon capture. The cultivation performance and microbial community analyses indicated that a continuous, stable culture on flue gas was achieved with the algal assemblage. The biomass yield was significantly improved to 0.82 g/L with 30% (v/v) daily harvesting once formate concentration was 0.5 g/L/day. The microalgal-bacteria assemblage was stable during the entire pilot operation with average relative abundances of 80.4% and 19.6% at domain level for the microalga and bacteria, respectively. Under the formate addition of 0.5 g/L/day and 30% (v/v) harvesting, mass and energy balance analysis revealed that the formate addition increased algal biomass yield by nearly 33% and reduce energy demand per carbon capture by 20% compared to the control cultivation on flue gas only.

CHAPTER 5: CONCLUSIONS AND FUTURE WORK

A new and robust algal-bacterial assemblage has been developed by this research to utilize formate as a carbon source. The assemblage contains C. sorokiniana and bacteria (Proteobacteria and Bacteroidetes), and it is highly adaptable to utilize formate. The use of formate has significantly enhanced the methylotrophic population in the assemblage, and isotope tracing results confirmed its significant contribution as a carbon source for photomixotrophic growth. Formate can be used as an alternative carbon source to bicarbonate or CO_2, which allows algal growth under a wide range of pH conditions, from weakly acidic to alkaline. The assemblage has shown strong resilience under different light intensities, and a high carbon capture rate of 90% was achieved from semi-continuous cultivation on formate. This assemblage provides a novel, flexible, and efficient route to fix CO_2 into algal biomass for value-added uses, particularly with the advancement of current research on the electrochemical conversion of CO_2 to formate. One of its advantages is the ability to handle relatively high formate concentrations, making it highly advantageous to repel protozoa, insects, and other contaminated species during long-term continuous cultivation. Hence, it holds great promise as an option for large-scale cultivation in non-sterile environments, such as open-pond cultivation.

The results also showed that formate is a highly effective carbon source for microalgal cultivation and has potential to be used for a mixed carbon feeding strategy in this field. Through comprehensive analysis of culture performance and information gathered from the microbial community, it has been demonstrated that formate is an excellent carbon source for microalgal cultivation. Among various experimental conditions, the cultivation setup with a formate feeding rate of 1 g/L/day under a light intensity of 180 μmol/m2/s proved to be the most outstanding. This innovative microalgal cultivation technique, which employs an alternative carbon feeding method, has yielded promising results. During the study, the highest biomass concentration

achieved was 1.6 g/L, a notable accomplishment made possible by this new cultivation technique. Statistical analysis conducted on the data clearly indicated a significant increase in biomass concentration, attributable to the utilization of this novel approach. By employing formate as the carbon source, the researchers have not only confirmed its effectiveness in microalgal cultivation but also provided new insights into the potential of mixed carbon feeding strategies.

Finally, this study applied the algal assemblage to utilize formate to enhance a pilot-scale algal cultivation of carbon capture. The results show that a continuous and stable culture can be achieved with the algal assemblage on flue gas. The biomass yield was significantly increased to 0.82 g/L with daily harvesting at 30% (v/v) once the formate concentration reached 0.5 g/L/day. The microbial community analysis revealed that the microalgal-bacterial assemblage remained stable during the pilot operation. Mass and energy balance analysis showed that the addition of formate at a rate of 0.5 g/L/day and 30% (v/v) harvesting increased the algal biomass yield by nearly 33% and reduced the energy demand per carbon capture by 20% compared to the control cultivation on flue gas only. These findings provide a promising strategy for improving the efficiency of carbon capture using microalgae and have potential applications in various industries. The microbial community analysis showed that the microalgal-bacterial assemblage remained stable during the pilot operation, indicating that the use of formate as a carbon source did not have a negative impact on the microbial community. This is an important consideration for the long-term sustainability of the cultivation process and suggests that formate can be used as a safe and effective carbon source for microalgal cultivation.

The promising results obtained from using formate as a carbon source in microalgal cultivation open up several avenues for future research and development. Some potential

directions for future work include:

- Efficient formate production methods: Investigating and develop more efficient and cost-effective methods for producing formate from various sources, such as biomass, natural gas, or carbon dioxide captured from the atmosphere or industrial processes.
- Investigation of downstream processing: Evaluating the impact of using formate as a carbon source on the downstream processing of microalgal biomass, such as extraction of lipids, proteins, and other valuable compounds. This would provide a better understanding of how formate utilization affects the overall production process and product quality.
- Metabolic pathway analysis: Conducting comprehensive metabolic pathway analysis to better understand the integration of formate metabolism with other cellular processes in microalgae. This may reveal potential bottlenecks, regulatory points, or opportunities for metabolic engineering to improve formate utilization and overall productivity.

BIBLIOGRAPHY

1. Al-Ghussain, L., Global warming: review on driving forces and mitigation. 2019. 38(1): p. 13-21.

2. Fu, J., et al., Photo-bioreactor design for microalgae: A review from the aspect of CO2 transfer and conversion. Bioresour Technol, 2019. 292: p. 121947.

3. Ho, S.-H., et al., Perspectives on microalgal CO2-emission mitigation systems — A review. Biotechnology Advances, 2011. 29(2): p. 189-198.

4. Wang, B., et al., CO(2) bio-mitigation using microalgae. Appl Microbiol Biotechnol, 2008. 79(5): p. 707-18.

5. Li-Beisson, Y., et al., The lipid biochemistry of eukaryotic algae. Progress in Lipid Research, 2019. 74: p. 31-68.

6. de Morais, M.G. and J.A. Costa, Biofixation of carbon dioxide by Spirulina sp. and Scenedesmus obliquus cultivated in a three-stage serial tubular photobioreactor. J Biotechnol, 2007. 129(3): p. 439-45.

7. Kim, G.-Y., K. Roh, and J.-I. Han, The use of bicarbonate for microalgae cultivation and its carbon footprint analysis. Green Chemistry, 2019. 21(18): p. 5053-5062.

8. Flynn, K.J., P. Kenny, and A. Mitra, Minimising losses to predation during microalgae cultivation. Journal of Applied Phycology, 2017. 29(4): p. 1829-1840.

9. Chen, J., T. Rashid, and G. Feng, Toxicity of formic acid to red imported fire ants, Solenopsis invicta Buren. Pest Management Science, 2012. 68(10): p. 1393-1399.

10. Kara, K., et al., Effect of dietary formic acid on the in vitro ruminal fermentation parameters of barley-based concentrated mix feed of beef cattle. Journal of Applied Animal Research, 2018. 46(1): p. 178-183.

11. Albo, J., et al., Towards the electrochemical conversion of carbon dioxide into methanol. Green Chemistry, 2015. 17(4): p. 2304-2324.

12. Kim, S., et al., Conversion of CO2 to formate in an electroenzymatic cell using Candida boidinii formate dehydrogenase. Journal of Molecular Catalysis B: Enzymatic, 2014. 102: p. 9-15.

13. Hwang, H.W., et al., Two-stage bioconversion of carbon monoxide to biopolymers via formate as an intermediate. Chemical Engineering Journal, 2020. 389: p. 124394.

14. Liu, Z., et al., Exploring eukaryotic formate metabolisms to enhance microbial growth and lipid accumulation. Biotechnol Biofuels, 2017. 10: p. 22.

15. Ferry, J.G., Formate dehydrogenase. FEMS Microbiol Rev, 1990. 7(3-4): p. 377-82.

16. Hou, C.T., et al., NAD-linked formate dehydrogenase from methanol-grown Pichia pastoris NRRL-Y-7556. Arch Biochem Biophys, 1982. 216(1): p. 296-305.

17. Mohamad, S., et al., Advanced wastewater treatment using microalgae: effect of temperature on removal of nutrients and organic carbon. IOP Conference Series: Earth and Environmental Science, 2017. 67(1): p. 012032.

18. Wang, Y., et al., Optimization of Chlorella vulgaris and bioflocculant-producing bacteria co-culture: enhancing microalgae harvesting and lipid content. Lett Appl Microbiol, 2015. 60(5): p. 497-503.

19. Magdouli, S., K. Brar, and J.-F. Blais, Co-culture for lipid production: Advances and challenges. Biomass and Bioenergy, 2016. 92: p. 20-30.

20. Riquelme, C. and R. Avendaño-Herrera, Interacción bacteria-microalga en el ambiente marino y uso potencial en acuicultura. Revista Chilena De Historia Natural - REV CHIL HIST NAT, 2003. 76.

21. Kazamia, E., et al., Mutualistic interactions between vitamin B12 -dependent algae and heterotrophic bacteria exhibit regulation. Environ Microbiol, 2012. 14(6): p. 1466-76.

22. Narala, R.R., et al., Comparison of Microalgae Cultivation in Photobioreactor, Open Raceway Pond, and a Two-Stage Hybrid System. 2016. 4(29).

23. Manabe, S.J.T.A.D.M. and Oceanography, Role of greenhouse gas in climate change. 2019. 71(1): p. 1620078.

24. Yoro, K.O. and M.O. Daramola, CO2 emission sources, greenhouse gases, and the global warming effect, in Advances in carbon capture. 2020, Elsevier. p. 3-28.

25. Yu, C.-H., et al., A review of CO2 capture by absorption and adsorption. 2012. 12(5): p. 745-769.

26. Mantripragada, H.C. and E.S.J.E.P. Rubin, Calcium looping cycle for CO2 capture: Performance, cost and feasibility analysis. 2014. 63: p. 2199-2206.

27. Pires, J., et al., Recent developments on carbon capture and storage: An overview. 2011. 89(9): p. 1446-1460.

28. Ben-Mansour, R., et al., Carbon capture by physical adsorption: materials, experimental investigations and numerical modeling and simulations–a review. 2016. 161: p. 225-255.

29. Khalilpour, R., et al., Membrane-based carbon capture from flue gas: a review. 2015. 103: p. 286-300.

30. Chen, G., et al., Membrane materials targeting carbon capture and utilization. 2022. 2: p. 100025.

31. Kraxner, F., et al., Negative emissions from BioEnergy use, carbon capture and sequestration (BECS)—the case of biomass production by sustainable forest management from semi-natural temperate forests. 2003. 24(4-5): p. 285-296.

32. Dou, X., et al., Greenhouse gas (CO2, CH4, N2O) emissions from soils following afforestation in central China. 2016. 126: p. 98-106.

33. Woodwell, G.J.S., CO2 Reduction and reforestation. 1988. 242(4885): p. 1493-1493.

34. Bertram, C.J.E.P., Ocean iron fertilization in the context of the Kyoto protocol and the post-Kyoto process. 2010. 38(2): p. 1130-1139.

35. Hanssen, S., et al., The climate change mitigation potential of bioenergy with carbon capture and storage. 2020. 10(11): p. 1023-1029.

36. Perin, G., et al., The potential of quantitative models to improve microalgae photosynthetic efficiency. 2019. 166(1): p. 380-391.

37. Chew, K.W., et al., Effects of water culture medium, cultivation systems and growth modes for microalgae cultivation: A review. Journal of the Taiwan Institute of Chemical Engineers, 2018. 91: p. 332-344.

38. Costa, J.A.V., et al., Open pond systems for microalgal culture, in Biofuels from algae. 2019, Elsevier. p. 199-223.

39. Costa, J.A.V. and M.G. de Morais, An open pond system for microalgal cultivation, in Biofuels from algae. 2014, Elsevier. p. 1-22.

40. Sompech, K., Y. Chisti, and T. Srinophakun, Design of raceway ponds for producing microalgae. Biofuels, 2012. 3(4): p. 387-397.

41. Chen, C.-Y., et al., Cultivation, photobioreactor design and harvesting of microalgae for biodiesel production: a critical review. Bioresource technology, 2011. 102(1): p. 71-81.

42. Molina, E., et al., Tubular photobioreactor design for algal cultures. Journal of biotechnology, 2001. 92(2): p. 113-131.

43. Camacho, F.G., et al., Photobioreactor scale-up for a shear-sensitive dinoflagellate microalga. Process Biochemistry, 2011. 46(4): p. 936-944.

44. Wang, S.K., et al., Microalgae cultivation in photobioreactors: An overview of light characteristics. Engineering in Life Sciences, 2014. 14(6): p. 550-559.

45. Zuccaro, G., et al. Microalgae Cultivation Systems. 2020.

46. Ma, X., et al., A comprehensive review on carbon source effect of microalgae lipid accumulation for biofuel production. Science of The Total Environment, 2022. 806: p. 151387.

47. Plöhn, M., et al., Utilization of Different Carbon Sources by Nordic Microalgae Grown Under Mixotrophic Conditions. 2022. 9.

48. Bashir, K.M.I., et al., Effect of organic carbon sources and environmental factors on cell growth and lipid content of Pavlova lutheri. Annals of Microbiology, 2019. 69(4): p. 353-368.

49. Liu, N., et al., Effects of organic carbon sources on the biomass and lipid production by the novel microalga Micractinium reisseri FM1 under batch and fed-batch cultivation. South African Journal of Botany, 2021. 139: p. 329-337.

50. Gao, F., et al., Effect of organic carbon to nitrogen ratio in wastewater on growth, nutrient uptake and lipid accumulation of a mixotrophic microalgae Chlorella sp. Bioresource Technology, 2019. 282: p. 118-124.

51. Pooja, K., et al., Cost-effective treatment of sewage wastewater using microalgae Chlorella vulgaris and its application as bio-fertilizer. Energy Nexus, 2022. 7: p. 100122.

52. Song, C., et al., Microalgae carbon fixation integrated with organic matters recycling from soybean wastewater: Effect of pH on the performance of hybrid system. Chemosphere, 2020. 248: p. 126094.

53. Zhu, C., et al., Progress toward a bicarbonate-based microalgae production system. Trends in Biotechnology, 2022. 40(2): p. 180-193.

54. Bernard, O. and B. Rémond, Validation of a simple model accounting for light and temperature effect on microalgal growth. Bioresource Technology, 2012. 123: p. 520-527.

55. Scherholz, M.L. and W.R. Curtis, Achieving pH control in microalgal cultures through fed-batch addition of stoichiometrically-balanced growth media. BMC Biotechnology, 2013. 13(1): p. 39.

56. Zhu, Z., J. Jiang, and Y. Fa, Overcoming the Biological Contamination in Microalgae and Cyanobacteria Mass Cultivations for Photosynthetic Biofuel Production. 2020. 25(22): p. 5220.

57. Saravanan, A., et al., A review on algal-bacterial symbiotic system for effective treatment of wastewater. Chemosphere, 2021. 271: p. 129540.

58. Cho, D.-H., et al., Enhancing microalgal biomass productivity by engineering a microalgal–bacterial community. Bioresource Technology, 2015. 175: p. 578-585.

59. Xue, L., et al., Analysis of growth and lipid production characteristics of Chlorella vulgaris in artificially constructed consortia with symbiotic bacteria. 2018. 58(4): p. 358-367.

60. Ramanan, R., et al., Algae–bacteria interactions: Evolution, ecology and emerging applications. Biotechnology Advances, 2016. 34(1): p. 14-29.

61. Cho, D.H., et al., Enhancing microalgal biomass productivity by engineering a microalgal-bacterial community. Bioresour Technol, 2015. 175: p. 578-85.

62. Wang, H., et al., Effects of bacterial communities on biofuel-producing microalgae: stimulation, inhibition and harvesting. Crit Rev Biotechnol, 2016. 36(2): p. 341-52.

63. Mao, W., et al., Recent progress in metabolic engineering of microbial formate assimilation. Applied Microbiology and Biotechnology, 2020. 104: p. 6905-6917.

64. Müller, V., A synthetic bacterial microcompartment as production platform for pyruvate from formate and acetate. Proceedings of the National Academy of Sciences, 2022. 119(9): p. e2201330119.

65. Hu, G., et al., Synergistic metabolism of glucose and formate increases the yield of short-chain organic acids in Escherichia coli. ACS Synthetic Biology, 2022. 11(1): p. 135-143.

66. Hartmann, T., N. Schwanhold, and S. Leimkühler, Assembly and catalysis of molybdenum or tungsten-containing formate dehydrogenases from bacteria. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 2015. 1854(9): p. 1090-1100.

67. Stadtman, T.C. and H. Barker, Studies on the Methane Fermentation X: A New Formate-Decomposing Bacterium, Methanococcus vannielii. Journal of bacteriology, 1951. 62(3): p. 269-280.

68. Han, N., et al., Promises of main group metal–based nanostructured materials for electrochemical CO2 reduction to formate. Advanced Energy Materials, 2020. 10(11): p. 1902338.

69. Katuri, K.P., et al., Electroactive biofilms on surface functionalized anodes: The anode respiring behavior of a novel electroactive bacterium, Desulfuromonas acetexigens. Water Research, 2020. 185: p. 116284.

70. Li, L., Recombineering for Precise Genome Editing in Shewanella oneidensis. 2022.

71. Hietala, J., et al., Formic Acid, in Ullmann's Encyclopedia of Industrial Chemistry. p. 1-22.

72. Innocent, B., et al., Electro-reduction of carbon dioxide to formate on lead electrode in aqueous medium. Journal of Applied Electrochemistry, 2009. 39(2): p. 227-232.

73. Hefetz, A. and M.S. Blum, Biosynthesis of formic acid by the poison glands of formicine ants. Biochim Biophys Acta, 1978. 543(4): p. 484-96.

74. Lu, X., et al., Electrochemical Reduction of Carbon Dioxide to Formic Acid. 2014. 1(5): p. 836-849.

75. Pan, H. and M.D. Heagy, Photons to Formate: A Review on Photocatalytic Reduction of CO2 to Formic Acid. 2020. 10(12): p. 2422.

76. Pan, H. and M.D.J.N. Heagy, Photons to Formate: A Review on Photocatalytic Reduction of CO2 to Formic Acid. 2020. 10.

77. Bulushev, D.A. and J.R.H. Ross, Towards Sustainable Production of Formic Acid. 2018. 11(5): p. 821-836.

78. Xiang, H., et al., Production of formate by CO2 electrochemical reduction and its application in energy storage. 2020. 4(1): p. 277-284.

79. Li, H. and C. Oloman, Development of a continuous reactor for the electro-reduction of carbon dioxide to formate - Part 1: Process variables. Journal of Applied Electrochemistry, 2006. 36(10): p. 1105-1115.

80. Li, H. and C. Oloman, Development of a continuous reactor for the electro-reduction of carbon dioxide to formate - Part 2: Scale-up. Journal of Applied Electrochemistry, 2007. 37(10): p. 1107-1117.

81. Philips, M.F., et al., Optimizing the Electrochemical Reduction of CO2 to Formate: A State-of-the-Art Analysis. Acs Sustainable Chemistry & Engineering, 2020. 8(41): p. 15430-15444.

82. EPA, U.S., Research and Development: Health and Environmental Effect Document for Formic Acid - Final Report, U.S.E.P. Agency, Editor. 1990: Cincinnati, OH.

83. Mujtaba, G. and K. Lee. Advanced Treatment of Wastewater Using Symbiotic Co-culture of Microalgae and Bacteria. 2016.

84. Wang, J., H. Yang, and F. Wang, Mixotrophic Cultivation of Microalgae for Biodiesel Production: Status and Prospects. Applied Biochemistry and Biotechnology, 2014. 172(7): p. 3307-3329.

85. Fawzy, M.A., et al., Growth behavior, phenol removal and lipid productivity of microalgae in mixotrophic and heterotrophic conditions under synergistic effect of phenol and bicarbonate for biodiesel production. Journal of Applied Phycology, 2022. 34(6): p. 2981-2994.

86. Gao, P., et al., Regulation of carbon source metabolism in mixotrophic microalgae cultivation in response to light intensity variation. Journal of Environmental Management, 2022. 302: p. 114095.

87. Poddar, N., R. Sen, and G.J.O. Martin, Bacterial abundance and diversity in Microchloropsis salina (formerly Nannochloropsis salina) cultures in response to the presence of ammonium, nitrate and glycerol. Journal of Applied Phycology, 2020. 32(2): p. 839-850.

88. Cutshaw, A., et al., A long-term pilot-scale algal cultivation on power plant flue gas – Cultivation stability and biomass accumulation. Algal Research, 2020. 52: p. 102115.

89. Wang, X., et al., Two-stage photoautotrophic cultivation to improve carbohydrate production in Chlamydomonas reinhardtii. Biomass and Bioenergy, 2015. 74: p. 280-287.

90. Undersander, D.J., D.R. Mertens, and N.J. Thiex, - Forage analyses procedures. 1993.

91. Hollinshead, W., L. He, and Y.J. Tang, (13)C-Fingerprinting and Metabolic Flux Analysis of Bacterial Metabolisms. Methods Mol Biol, 2019. 1927: p. 215-230.

92. Czajka, J.J., et al., Application of Stable Isotope Tracing to Elucidate Metabolic Dynamics During Yarrowia lipolytica α-Ionone Fermentation. iScience, 2020. 23(2): p. 100854.

93. Burja, A.M., et al., Identification of the green alga, Chlorella vulgaris (SDC1) using cyanobacteria derived 16S rDNA primers: targeting the chloroplast. FEMS Microbiol Lett, 2001. 202(2): p. 195-203.

94. Burja, A.M., et al., Identification of the green alga, Chlorella vulgaris (SDC1) using cyanobacteria derived 16S rDNA primers: targeting the chloroplast. Fems Microbiology Letters, 2001. 202(2): p. 195-203.

95. Sukenik, A., T. Zohary, and J. Padisák, Cyanoprokaryota and Other Prokaryotic Algae, in Encyclopedia of Inland Waters, G.E. Likens, Editor. 2009, Academic Press: Oxford. p. 138-148.

96. Hughes, C.S., et al., Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. Nature Protocols, 2019. 14(1): p. 68-+.

97. Cox, J. and M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnology, 2008. 26(12): p. 1367-1372.

98. Cox, J., et al., Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. Molecular & Cellular Proteomics, 2014. 13(9): p. 2513-2526.

99. Bolyen, E., et al., Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology, 2019. 37(8): p. 852-857.

100. Xiao, Z., et al., Development of a kinetic model to describe six types of symbiotic interactions in a formate utilizing microalgae-bacteria cultivation system. Algal Research-Biomass Biofuels and Bioproducts, 2021. 58.

101. Lidstrom, M.E., Aerobic Methylotrophic Prokaryotes. Prokaryotes: a Handbook on the Biology of Bacteria, Vol 2, Third Edition: Ecophysiology and Biochemistry, 2006: p. 618-634.

102. Kelly, D.P., I.R. McDonald, and A.P. Wood, The Family Methylobacteriaceae, in The Prokaryotes: Alphaproteobacteria and Betaproteobacteria, E. Rosenberg, et al., Editors. 2014, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 313-340.

103. Kumar, M., et al., Biodiversity of methylotrophic microbial communities and their potential role in mitigation of abiotic stresses in plants. Biologia, 2019. 74(3): p. 287-308.

104. Tyanova, S., et al., The Perseus computational platform for comprehensive analysis of (prote)omics data. Nature Methods, 2016. 13(9): p. 731-740.

105. Herzog, H., B. Eliasson, and O. Kaarstad, Capturing greenhouse gases. Scientific American, 2000. 282(2): p. 72-79.

106. Crowley, T.J. and R.A. Berner, CO2 and climate change. Science, 2001. 292(5518): p. 870-872.

107. Bach, W. co2 LEVELS IN THE ATMOSPHERE. in Carbon Dioxide, Climate and Society: Proceedings of a IIASA Workshop cosponsored by WMO, UNEP, and SCOPE, February 21-24, 1978. 2013. Elsevier.

108. Doney, S.C., et al., Ocean acidification: the other CO2 problem. Ann Rev Mar Sci, 2009. 1: p. 169-92.

109. Ghosh, A. and B. Kiran, Carbon concentration in algae: reducing CO2 from exhaust gas. Trends in biotechnology, 2017. 35(9): p. 806-808.

110. Chelf, P. and L.M. Brown, Microalgal mass culture and the greenhouse effect: resources update. Proceedings of the Aquatic Species Program. Solar Energy Research Institute, 1989.

111. Randrianarison, G. and M.A. Ashraf, Microalgae: a potential plant for energy production. Geology, Ecology, and Landscapes, 2017. 1(2): p. 104-120.

112. Ghasemi, Y., et al., Microalgae biofuel potentials. Applied Biochemistry and Microbiology, 2012. 48: p. 126-144.

113. Lum, K.K., J. Kim, and X.G. Lei, Dual potential of microalgae as a sustainable biofuel feedstock and animal feed. Journal of animal science and biotechnology, 2013. 4: p. 1-7.

114. Mourelle, M.L., C.P. Gómez, and J.L. Legido, The potential use of marine microalgae and cyanobacteria in cosmetics and thalassotherapy. Cosmetics, 2017. 4(4): p. 46.

115. Gatrell, S., et al., Nonruminant Nutrition Symposium: Potential of defatted microalgae from the biofuel industry as an ingredient to replace corn and soybean meal in swine and poultry diets. Journal of animal science, 2014. 92(4): p. 1306-1314.

116. Singh, N.K. and D.W. Dhar, Microalgae as second generation biofuel. A review. Agronomy for Sustainable Development, 2011. 31: p. 605-629.

117. Becker, E.W., Micro-algae as a source of protein. Biotechnology advances, 2007. 25(2): p. 207-210.

118. Brown, M., et al., The vitamin content of microalgae used in aquaculture. Journal of Applied Phycology, 1999. 11: p. 247-255.

119. Tominaga, K., et al., Cosmetic benefits of astaxanthin on humans subjects. Acta Biochimica Polonica, 2012. 59(1).

120. White, D., et al., The effect of sodium bicarbonate supplementation on growth and biochemical composition of marine microalgae cultures. Journal of applied phycology, 2013. 25: p. 153-165.

121. Cheirsilp, B. and S. Torpee, Enhanced growth and lipid production of microalgae under mixotrophic culture condition: effect of light intensity, glucose concentration and fed-batch cultivation. Bioresource technology, 2012. 110: p. 510-516.

122. Zheng, Q., et al., Critical review of strategies for CO2 delivery to large-scale microalgae cultures. Chinese journal of chemical engineering, 2018. 26(11): p. 2219-2228.

123. Tao, X., et al., Winter warming in Alaska accelerates lignin decomposition contributed by Proteobacteria. Microbiome, 2020. 8(1): p. 1-12.

124. Ma, A., et al., Ascomycota members dominate fungal communities during straw residue decomposition in arable soil. PloS one, 2013. 8(6): p. e66146.

125. Wang, Y., et al., Gut fungal microbiome responses to natural cryptosporidium infection in horses. Frontiers in Microbiology, 2022. 13.

126. Zhang, N., et al., An overview of the systematics of the Sordariomycetes based on a four-gene phylogeny. Mycologia, 2006. 98(6): p. 1076-1087.

127. Zheng, Y., et al., An algal-bacterial symbiotic system of carbon fixation using formate as a carbon source. 2021.

128. IPCC, Climate Change 2014: Mitigation of Climate Change 2014.

129. Boden, T.A., Marland, G., and Andres, R.J. (2017). National CO2 Emissions from Fossil-Fuel Burning, Cement Manufacture, and Gas Flaring: 1751-2014, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, doi.

130. Daneshvar, E., et al., Biologically-mediated carbon capture and utilization by microalgae towards sustainable CO2 biofixation and biomass valorization – A review. Chemical Engineering Journal, 2022. 427: p. 130884.

131. Singh, U.B. and A.S. Ahluwalia, Microalgae: a promising tool for carbon sequestration. Mitigation and Adaptation Strategies for Global Change, 2013. 18(1): p. 73-95.

132. Lam, M.K. and K.T. Lee, Microalgae biofuels: A critical review of issues, problems and the way forward. Biotechnology Advances, 2012. 30(3): p. 673-690.

133. Dineshkumar, R., et al., Exploring the Microalgae Biofertilizer Effect on Onion Cultivation by Field Experiment. Waste and Biomass Valorization, 2020. 11(1): p. 77-87.

134. Narala, R.R., et al., Comparison of Microalgae Cultivation in Photobioreactor, Open Raceway Pond, and a Two-Stage Hybrid System. 2016. 4.

135. Torzillo, G., et al., Production of Spirulina biomass in closed photobioreactors. Biomass, 1986. 11(1): p. 61-74.

136. Pruvost, J., et al., Microalgae culture in building-integrated photobioreactors: Biomass production modelling and energetic analysis. Chemical Engineering Journal, 2016. 284: p. 850-861.

137. Borowitzka, M.A. and N.R. Moheimani, Open Pond Culture Systems, in Algae for Biofuels and Energy, M.A. Borowitzka and N.R. Moheimani, Editors. 2013, Springer Netherlands: Dordrecht. p. 133-152.

138. Singh, R.N. and S. Sharma, Development of suitable photobioreactor for algae production – A review. Renewable and Sustainable Energy Reviews, 2012. 16(4): p. 2347-2353.

139. Raes, E.J., et al., Comparison of growth of Tetraselmis in a tubular photobioreactor (Biocoil) and a raceway pond. Journal of Applied Phycology, 2014. 26(1): p. 247-255.

140. Deruyck, B., et al., Modeling the impact of rotifer contamination on microalgal production in open pond, photobioreactor and thin layer cultivation systems. Algal Research, 2019. 38: p. 101398.

141. Ogbonna, J.C., H. Yoshizawa, and H. Tanaka, Treatment of high strength organic wastewater by a mixed culture of photosynthetic microorganisms. Journal of Applied Phycology, 2000. 12(3): p. 277-284.

142. Zhu, L.J.B., Bioproducts and Biorefining, Microalgal culture strategies for biofuel production: a review. 2015. 9(6): p. 801-814.

143. Novoveská, L., et al., Stabilizing continuous mixed cultures of microalgae. 2016. 13: p. 126-133.

144. Benemann, J.J.E., Microalgae for biofuels and animal feeds. 2013. 6(11): p. 5869-5886.

145. Magdouli, S., S.K. Brar, and J.F. Blais, Co-culture for lipid production: Advances and challenges. Biomass and Bioenergy, 2016. 92: p. 20-30.

146. Perera, I.A., et al., Microalgal–bacterial consortia unveil distinct physiological changes to facilitate growth of microalgae. 2021. 97(3): p. fiab012.

147. Tesfaw, A., F.J.B. Assefa, and m.B. reviews, Co-culture: A great promising method in single cell protein production. 2014. 9(2): p. 12-20.

148. Van Den Hende, S., et al., Bioflocculation of microalgae and bacteria combined with flue gas to improve sewage treatment. New Biotechnology, 2011. 29(1): p. 23-31.

149. Han, J., et al., Co-culturing bacteria and microalgae in organic carbon containing medium. Journal of Biological Research-Thessaloniki, 2016. 23(1): p. 8.

150. Molazadeh, M., et al., The Use of Microalgae for Coupling Wastewater Treatment With CO2 Biofixation. 2019. 7.

151. Geng, Y., et al., Resourceful treatment of harsh high-nitrogen rare earth element tailings (REEs) wastewater by carbonate activated Chlorococcum sp. microalgae. Journal of Hazardous Materials, 2022. 423: p. 127000.

152. Burns, B.D. and J. Beardall, Utilization of inorganic carbon by marine microalgae. Journal of Experimental Marine Biology and Ecology, 1987. 107(1): p. 75-86.

153. Zhang, R.-L., et al., Selection of microalgae strains for bicarbonate-based integrated carbon capture and algal production system to produce lipid. International Journal of Green Energy, 2019. 16(11): p. 825-833.

154. Sampathkumar, S.J. and K.M. Gothandam, Sodium bicarbonate augmentation enhances lutein biosynthesis in green microalgae Chlorella pyrenoidosa. Biocatalysis and Agricultural Biotechnology, 2019. 22: p. 101406.

155. Cutshaw, A., et al., A long-term pilot-scale algal cultivation on power plant flue gas - Cultivation stability and biomass accumulation. Algal Research-Biomass Biofuels and Bioproducts, 2020. 52.

156. Daiek, C. and t.l.w.o.i.a.n.w. Link to external site, Effects of Water Recirculation on Pilot-Scale Microalgae Cultivation Using Flue Gas CO2. 2020, Michigan State University: United States -- Michigan. p. 100.

157. Fulbright, S.P., et al., Bacterial community changes in an industrial algae production system. Algal research, 2018. 31: p. 147-156.

158. Teske, A., et al., Evolutionary relationships among ammonia-and nitrite-oxidizing bacteria. Journal of bacteriology, 1994. 176(21): p. 6623-6630.

159. Ji, B., et al., Microalgal-bacterial granular sludge for municipal wastewater treatment under simulated natural diel cycles: Performances-metabolic pathways-microbial community nexus. Algal Research, 2021. 54: p. 102198.

160. Dyksma, S., et al., Ubiquitous Gammaproteobacteria dominate dark carbon fixation in coastal sediments. The ISME Journal, 2016. 10(8): p. 1939-1953.

APPENDIX A: ORIGINAL R CODES

CHAPTER 2

```
#Load libraries
   library(vegan)
   library(phyloseq)
   library (MASS)
   library(ggplot2)
   library(grid)
   library(gridExtra)
   library(ggpubr)
   library(extrafont)
   loadfonts(device="win")
# Plot bar chart with standard deviation
#data : a data frame
#varname : the name of a column containing the variable to be summarized
#groupnames : vector of column names to be used as
#grouping variables
data_summary <- function(data, varname, groupnames){
   require(plyr)
   summary_func <- function(x, col){
      c(mean = mean(x[[col]], na.rm=TRUE),
        sd = sd(x[[col]], na.rm=TRUE))
   }
   data_sum<-ddply(data, groupnames, .fun=summary_func,
                           varname)
   data_sum <- rename(data_sum, c("mean" = varname))
   return(data_sum)
}
#Choose data file
con <-file.choose(new = FALSE)
metadata <- read.table(con, header = T, row.names = 1, fill = TRUE)
# Define factors for metadata
metadata$Carbon_source <- factor(metadata$Carbon_source)

## Statistical analysis ##
# Each data set is tested for normality
# Normal data is tested for equal variance with an f test using var.test
# Non-normal data is tested for equal variance with levenes test using infer_levene_test
# Normal & equal variance--> ANOVA, Tukey pairwise comparison
# Normal & non-equal variance-->t test, var.equal=FALSE
# Non-normal & equal variance--> Kruskal Willis test
# Non-normal & non-equal varianve-->Kruskal Willis test
# At 0.02 hours ----
```

```
# Select the data at 0.02 hours
data4 <- metadata[which(metadata$Culture_time=="0.02"),]
data4$Carbon_source<-factor(data4$Carbon_source)
data4$Light_intensity<-factor(data4$Light_intensity)
data4

# Biomass concentration
# Normality
shapiro.test(metadata$Biomass_concentration)
# Variance
var.test(Biomass_productivity ~ Formate_conc, data = data1)
var.test(Biomass_conc ~ Light_intensity, data = data3)

# Two-way ANOVA and pair-wise comparison
fit1 <- aov(Biomass_productivity ~ Formate_conc, data=data1)
summary(fit1)
Tukey1 <- TukeyHSD(fit1, conf.level=0.95)
Tukey1

aov=aov(TN~Light_intensity*Formate_conc,data=metadata)
summary(aov)
Tukey1 <- TukeyHSD(aov, conf.level=0.95)
Tukey1
# At 0.31 hours ------

# Select the data at 0.31 hours
data4 <- metadata[which(metadata$Culture_time=="0.31"),]
data4$Carbon_source<-factor(data4$Carbon_source)
data4$Light_intensity<-factor(data4$Light_intensity)
data4

# Biomass concentration

# Normality
shapiro.test(data4$Biomass_conc)
# Variance
var.test(Biomass_conc ~ Carbon_source, data = data4)
var.test(Biomass_conc ~ Light_intensity, data = data4)

# Two-way ANOVA and pair-wise comparison
fit2 <- aov(Biomass_concentration~ Formate_conc*Light_intensity, data=metadata)
summary(fit2)
Tukey2 <- TukeyHSD(fit2, conf.level=0.95)
Tukey2

# Carbon source consumption
```

```
# Normality
shapiro.test(data4$Carbon_source_consumption)
# Variance
var.test(Carbon_source_consumption ~ Carbon_source, data = data4)
var.test(Carbon_source_consumption ~ Light_intensity, data = data4)

# Two-way ANOVA and pair-wise comparison
fit3 <- aov(Carbon_source_consumption ~ Carbon_source*Light_intensity, data=data4)
summary(fit3)
Tukey3 <- TukeyHSD(fit3, conf.level=0.95)
Tukey3

# NH4-N consumption

# Normality
shapiro.test(data4$NH4_N_consumption)
# Variance
var.test(NH4_N_consumption ~ Carbon_source, data = data4)
var.test(NH4_N_consumption ~ Light_intensity, data = data4)

# Two-way ANOVA and pair-wise comparison
fit4 <- aov(NH4_N_consumption ~ Carbon_source*Light_intensity, data=data4)
summary(fit4)
Tukey4 <- TukeyHSD(fit4, conf.level=0.95)
Tukey4

# TN consumption

# Normality
shapiro.test(data4$TN_consumption)
# Variance
var.test(TN_consumption ~ Carbon_source, data = data4)
var.test(TN_consumption ~ Light_intensity, data = data4)

# Significance
kruskal.test(TN_consumption ~ Carbon_source, data = data4)
kruskal.test(TN_consumption ~ Light_intensity, data = data4)


# TP consumption

# Normality
shapiro.test(data4$TP_consumption)
# Variance
var.test(TP_consumption ~ Carbon_source, data = data4)
```

```
var.test(TP_consumption ~ Light_intensity, data = data4)

# Two-way ANOVA and pair-wise comparison
fit5 <- aov(TP_consumption ~ Carbon_source*Light_intensity, data=data4)
summary(fit5)
Tukey5 <- TukeyHSD(fit4, conf.level=0.95)
Tukey5


# At 4 hours ------

# Select the data at 4 hours
data5 <- metadata[which(metadata$Culture_time=="4"),]
data5$Carbon_source<-factor(data5$Carbon_source)
data5$Light_intensity<-factor(data5$Light_intensity)
data5

# Biomass concentration

# Normality
shapiro.test(data5$Biomass_conc)
# Variance
var.test(Biomass_conc ~ Carbon_source, data = data5)
var.test(Biomass_conc ~ Light_intensity, data = data5)

# Two-way ANOVA and pair-wise comparison
fit6 <- aov(Biomass_conc ~ Carbon_source*Light_intensity, data=data5)
summary(fit6)
Tukey6 <- TukeyHSD(fit6, conf.level=0.95)
Tukey6

# Carbon source consumption

# Normality
shapiro.test(data5$Carbon_source_consumption)
# Variance
var.test(Carbon_source_consumption ~ Carbon_source, data = data5)
var.test(Carbon_source_consumption ~ Light_intensity, data = data5)

# Two-way ANOVA and pair-wise comparison
fit7 <- aov(Carbon_source_consumption ~ Carbon_source*Light_intensity, data=data5)
summary(fit7)
Tukey7 <- TukeyHSD(fit7, conf.level=0.95)
Tukey7

# NH4-N consumption
```

```
# Normality
shapiro.test(data5$NH4_N_consumption)
# Variance
var.test(NH4_N_consumption ~ Carbon_source, data = data5)
var.test(NH4_N_consumption ~ Light_intensity, data = data5)

# Significance
kruskal.test(NH4_N_consumption ~ Carbon_source, data = data5)
kruskal.test(NH4_N_consumption ~ Light_intensity, data = data5)


# TN consumption

# Normality
shapiro.test(data5$TN_consumption)
# Variance
var.test(TN_consumption ~ Carbon_source, data = data5)
var.test(TN_consumption ~ Light_intensity, data = data5)

# Two-way ANOVA and pair-wise comparison
fit8 <- aov(TN_consumption ~ Carbon_source*Light_intensity, data=data5)
summary(fit8)
Tukey8 <- TukeyHSD(fit8, conf.level=0.95)
Tukey8

# TP consumption

# Normality
shapiro.test(data5$TP_consumption)
# Variance
var.test(TP_consumption ~ Carbon_source, data = data5)
var.test(TP_consumption ~ Light_intensity, data = data5)

# Two-way ANOVA and pair-wise comparison
fit9 <- aov(TP_consumption ~ Carbon_source*Light_intensity, data=data5)
summary(fit9)
Tukey9 <- TukeyHSD(fit9, conf.level=0.95)
Tukey9


# At 8 hours ------

# Select the data at 8 hours
data6 <- metadata[which(metadata$Culture_time=="8"),]
data6$Carbon_source<-factor(data5$Carbon_source)
```

```
data6$Light_intensity<-factor(data6$Light_intensity)
data6

# Biomass concentration

# Normality
shapiro.test(data6$Biomass_conc)
# Variance
var.test(Biomass_conc ~ Carbon_source, data = data6)
var.test(Biomass_conc ~ Light_intensity, data = data6)

# Two-way ANOVA and pair-wise comparison
fit10 <- aov(Biomass_conc ~ Carbon_source*Light_intensity, data=data6)
summary(fit10)
Tukey10 <- TukeyHSD(fit10, conf.level=0.95)
Tukey10

# Carbon source consumption

# Normality
shapiro.test(data6$Carbon_source_consumption)
# Variance
var.test(Carbon_source_consumption ~ Carbon_source, data = data6)
var.test(Carbon_source_consumption ~ Light_intensity, data = data6)

# Two-way ANOVA and pair-wise comparison
fit11 <- aov(Carbon_source_consumption ~ Carbon_source*Light_intensity, data=data6)
summary(fit11)
Tukey11 <- TukeyHSD(fit11, conf.level=0.95)
Tukey11

# NH4-N consumption

# Normality
shapiro.test(data6$NH4_N_consumption)
# Variance
var.test(NH4_N_consumption ~ Carbon_source, data = data6)
var.test(NH4_N_consumption ~ Light_intensity, data = data6)

# Two-way ANOVA and pair-wise comparison
fit12 <- aov(NH4_N_consumption ~ Carbon_source*Light_intensity, data=data6)
summary(fit12)
Tukey12 <- TukeyHSD(fit12, conf.level=0.95)
Tukey12

# TN consumption
```

```
# Normality
shapiro.test(data6$TN_consumption)
# Variance
var.test(TN_consumption ~ Carbon_source, data = data6)
var.test(TN_consumption ~ Light_intensity, data = data6)

# Two-way ANOVA and pair-wise comparison
fit13 <- aov(TN_consumption ~ Carbon_source*Light_intensity, data=data6)
summary(fit13)
Tukey13 <- TukeyHSD(fit13, conf.level=0.95)
Tukey13

# TP consumption

# Normality
shapiro.test(data6$TP_consumption)
# Variance
var.test(TP_consumption ~ Carbon_source, data = data6)
var.test(TP_consumption ~ Light_intensity, data = data6)

# Two-way ANOVA and pair-wise comparison
fit14 <- aov(TP_consumption ~ Carbon_source*Light_intensity, data=data6)
summary(fit14)
Tukey14 <- TukeyHSD(fit14, conf.level=0.95)
Tukey14



# At 24 hours ------

# Select the data at 24 hours
data7 <- metadata[which(metadata$Culture_time=="24"),]
data7$Carbon_source<-factor(data5$Carbon_source)
data7$Light_intensity<-factor(data6$Light_intensity)
data7

# Biomass concentration

# Normality
shapiro.test(data7$Biomass_conc)
# Variance
var.test(Biomass_conc ~ Carbon_source, data = data7)
var.test(Biomass_conc ~ Light_intensity, data = data7)

# Two-way ANOVA and pair-wise comparison
```

```
fit15 <- aov(Biomass_conc ~ Carbon_source*Light_intensity, data=data7)
summary(fit15)
Tukey15 <- TukeyHSD(fit15, conf.level=0.95)
Tukey15

# Carbon source consumption

# Normality
shapiro.test(data7$Carbon_source_consumption)
# Variance
var.test(Carbon_source_consumption ~ Carbon_source, data = data7)
var.test(Carbon_source_consumption ~ Light_intensity, data = data7)

# Two-way ANOVA and pair-wise comparison
fit16 <- aov(Carbon_source_consumption ~ Carbon_source*Light_intensity, data=data7)
summary(fit16)
Tukey16 <- TukeyHSD(fit16, conf.level=0.95)
Tukey16

# NH4-N consumption

# Normality
shapiro.test(data7$NH4_N_consumption)
# Variance
var.test(NH4_N_consumption ~ Carbon_source, data = data7)
var.test(NH4_N_consumption ~ Light_intensity, data = data7)

# Two-way ANOVA and pair-wise comparison
fit17 <- aov(NH4_N_consumption ~ Carbon_source*Light_intensity, data=data7)
summary(fit17)
Tukey17 <- TukeyHSD(fit17, conf.level=0.95)
Tukey17

# TN consumption

# Normality
shapiro.test(data7$TN_consumption)
# Variance
var.test(TN_consumption ~ Carbon_source, data = data7)
var.test(TN_consumption ~ Light_intensity, data = data7)

# Two-way ANOVA and pair-wise comparison
fit18 <- aov(TN_consumption ~ Carbon_source*Light_intensity, data=data7)
summary(fit13)
Tukey18 <- TukeyHSD(fit18, conf.level=0.95)
Tukey18
```

```
# TP consumption

# Normality
shapiro.test(data6$TP_consumption)
# Variance
var.test(TP_consumption ~ Carbon_source, data = data7)
var.test(TP_consumption ~ Light_intensity, data = data7)

# Two-way ANOVA and pair-wise comparison
fit19 <- aov(TP_consumption ~ Carbon_source*Light_intensity, data=data7)
summary(fit19)
Tukey19 <- TukeyHSD(fit19, conf.level=0.95)
Tukey19

##Plots##

# Biomass concentration under 500 µmol/m2/s
data <- metadata[which(metadata$Light_intensity=="500"),]
Biomass_concentration <- data_summary(data, varname="Biomass_concentration",
                                                groupnames=("Carbon_source"))
Biomass_concentration$ Carbon_source =as.factor(Biomass_concentration$ Carbon_source)
plot_1 <- ggplot(Biomass_concentration, aes(x=Culture_time, y=Biomass_concentration, group=
Carbon_source)) +
   geom_errorbar(aes(ymin=Biomass_concentration-sd,          ymax=Biomass_concentration+sd,
color= Carbon_source), width=0.2, position=position_dodge(0))+
   geom_line(aes(color= Carbon_source))+
   geom_point(aes(color= Carbon_source))+
   xlab("Culture time (hr)")+
   ylab("Biomass Concentration (g/L)") + labs(title = "", subtitle=NULL) +
   theme_classic() +
   theme(title=element_text(size=30, family="Times New Roman"),
            axis.text.x = element_text(size=25, family="Times New Roman"),
            axis.text.y=element_text(size=25, family="Times New Roman"),
            axis.title.y = element_text(size = 30, family="Times New Roman"),
            axis.title.x=element_text(size=30, family="Times New Roman"),
            legend.position ="top",
            legend.title = element_text(size=30),
            legend.text = element_text(size=30),
            legend.key.size=unit(1,'cm'))+
   scale_shape_manual(values=c(1,2,3,4,5))+
   scale_color_manual(values=c("red","blue","green","black","Purple"))+
   scale_size_manual(values=c(5,5,5,5,5))
plot_1
```

```
# Biomass concentration under 50 µmol/m2/s
data <- metadata[which(metadata$Light_intensity=="50"),]
Biomass_concentration <- data_summary(data, varname="Biomass_concentration",
                                                groupnames=("Carbon_source"))
Biomass_concentration$ Carbon_source =as.factor(Biomass_concentration$ Carbon_source)
plot_2 <- ggplot(Biomass_concentration, aes(x=Culture_time, y=Biomass_concentration, group=
Carbon_source)) +
   geom_errorbar(aes(ymin=Biomass_concentration-sd,          ymax=Biomass_concentration+sd,
color= Carbon_source), width=0.2, position=position_dodge(0))+
   geom_line(aes(color= Carbon_source))+
   geom_point(aes(color= Carbon_source))+
   xlab("Culture time (hr)")+
   ylab("Biomass Concentration (g/L)") + labs(title = "", subtitle=NULL) +
   theme_classic() +
   theme(title=element_text(size=30, family="Times New Roman"),
            axis.text.x = element_text(size=25, family="Times New Roman"),
            axis.text.y=element_text(size=25, family="Times New Roman"),
            axis.title.y = element_text(size = 30, family="Times New Roman"),
            axis.title.x=element_text(size=30, family="Times New Roman"),
            legend.position ="top",
            legend.title = element_text(size=30),
            legend.text = element_text(size=30),
            legend.key.size=unit(1,'cm'))+
   scale_shape_manual(values=c(1,2,3,4,5))+
   scale_color_manual(values=c("red","blue","green","black","Purple"))+
   scale_size_manual(values=c(5,5,5,5,5))
plot_2

# Carbon concentration under 500 µmol/m2/s
data <- metadata[which(metadata$Light_intensity=="500"),]
Carbon_concentration <- data_summary(data, varname="Carbon_concentration",
                                                groupnames=("Carbon_source"))
Carbon _concentration$ Carbon_source =as.factor(Carbon_concentration$Carbon_source)
plot_3 <- ggplot(Carbon_concentration, aes(x=Culture_time, y=Carbon_concentration, group=
Carbon_source)) +
   geom_errorbar(aes(ymin=Carbon_concentration-sd,  ymax=Carbon_concentration+sd,  color=
Carbon_source), width=0.2, position=position_dodge(0))+
   geom_line(aes(color= Carbon_source))+
   geom_point(aes(color= Carbon_source))+
   xlab("Culture time (hr)")+
   ylab("Carbon Concentration (g/L)") + labs(title = "", subtitle=NULL) +
   theme_classic() +
   theme(title=element_text(size=30, family="Times New Roman"),
            axis.text.x = element_text(size=25, family="Times New Roman"),
            axis.text.y=element_text(size=25, family="Times New Roman"),
            axis.title.y = element_text(size = 30, family="Times New Roman"),
```

```
            axis.title.x=element_text(size=30, family="Times New Roman"),
            legend.position ="top",
            legend.title = element_text(size=30),
            legend.text = element_text(size=30),
            legend.key.size=unit(1,'cm'))+
   scale_shape_manual(values=c(1,2,3,4,5))+
   scale_color_manual(values=c("red","blue","green","black","Purple"))+
   scale_size_manual(values=c(5,5,5,5,5))
plot_3

# Carbon concentration under 50 µmol/m2/s
data <- metadata[which(metadata$Light_intensity=="50"),]
Carbon_concentration <- data_summary(data, varname="Carbon_concentration",
                                                groupnames=("Carbon_source"))
Carbon _concentration$ Carbon_source =as.factor(Carbon_concentration$Carbon_source)
plot_4 <- ggplot(Carbon_concentration, aes(x=Culture_time, y=Carbon_concentration, group=
Carbon_source)) +
   geom_errorbar(aes(ymin=Carbon_concentration-sd, ymax=Carbon_concentration+sd, color=
Carbon_source), width=0.2, position=position_dodge(0))+
   geom_line(aes(color= Carbon_source))+
   geom_point(aes(color= Carbon_source))+
   xlab("Culture time (hr)")+
   ylab("Carbon Concentration (g/L)") + labs(title = "", subtitle=NULL) +
   theme_classic() +
   theme(title=element_text(size=30, family="Times New Roman"),
            axis.text.x = element_text(size=25, family="Times New Roman"),
            axis.text.y=element_text(size=25, family="Times New Roman"),
            axis.title.y = element_text(size = 30, family="Times New Roman"),
            axis.title.x=element_text(size=30, family="Times New Roman"),
            legend.position ="top",
            legend.title = element_text(size=30),
            legend.text = element_text(size=30),
            legend.key.size=unit(1,'cm'))+
   scale_shape_manual(values=c(1,2,3,4,5))+
   scale_color_manual(values=c("red","blue","green","black","Purple"))+
   scale_size_manual(values=c(5,5,5,5,5))
plot_4

# TP concentration under 500 µmol/m2/s
data <- metadata[which(metadata$Light_intensity=="500"),]
TP_concentration <- data_summary(data, varname="TP_concentration",
                                                groupnames=("Carbon_source"))
TP _concentration$ Carbon_source =as.factor(TP_concentration$Carbon_source)
plot_5 <- ggplot(TP_concentration, aes(x=Culture_time, y=TP_concentration, group=
Carbon_source)) +
   geom_errorbar(aes(ymin=TP_concentration-sd, ymax=TP_concentration+sd, color=
```

```
Carbon_source), width=0.2, position=position_dodge(0))+
   geom_line(aes(color= Carbon_source))+
   geom_point(aes(color= Carbon_source))+
   xlab("Culture time (hr)")+
   ylab("TP Concentration (mg/L)") + labs(title = "", subtitle=NULL) +
   theme_classic() +
   theme(title=element_text(size=30, family="Times New Roman"),
            axis.text.x = element_text(size=25, family="Times New Roman"),
            axis.text.y=element_text(size=25, family="Times New Roman"),
            axis.title.y = element_text(size = 30, family="Times New Roman"),
            axis.title.x=element_text(size=30, family="Times New Roman"),
            legend.position ="top",
            legend.title = element_text(size=30),
            legend.text = element_text(size=30),
            legend.key.size=unit(1,'cm'))+
   scale_shape_manual(values=c(1,2,3,4,5))+
   scale_color_manual(values=c("red","blue","green","black","Purple"))+
   scale_size_manual(values=c(5,5,5,5,5))
plot_5

# TP concentration under 50 µmol/m2/s
data <- metadata[which(metadata$Light_intensity=="50"),]
TP_concentration <- data_summary(data, varname="TP_concentration",
                                                  groupnames=("Carbon_source"))
TP _concentration$ Carbon_source =as.factor(TP_concentration$Carbon_source)
plot_6 <- ggplot(TP_concentration, aes(x=Culture_time, y=TP_concentration, group=
Carbon_source)) +
   geom_errorbar(aes(ymin=TP_concentration-sd, ymax=TP_concentration+sd, color=
Carbon_source), width=0.2, position=position_dodge(0))+
   geom_line(aes(color= Carbon_source))+
   geom_point(aes(color= Carbon_source))+
   xlab("Culture time (hr)")+
   ylab("TP Concentration (mg/L)") + labs(title = "", subtitle=NULL) +
   theme_classic() +
   theme(title=element_text(size=30, family="Times New Roman"),
            axis.text.x = element_text(size=25, family="Times New Roman"),
            axis.text.y=element_text(size=25, family="Times New Roman"),
            axis.title.y = element_text(size = 30, family="Times New Roman"),
            axis.title.x=element_text(size=30, family="Times New Roman"),
            legend.position ="top",
            legend.title = element_text(size=30),
            legend.text = element_text(size=30),
            legend.key.size=unit(1,'cm'))+
   scale_shape_manual(values=c(1,2,3,4,5))+
   scale_color_manual(values=c("red","blue","green","black","Purple"))+
   scale_size_manual(values=c(5,5,5,5,5))
```

```
plot_6

# TN concentration under 500 µmol/m2/s
data <- metadata[which(metadata$Light_intensity=="500"),]
TN_concentration <- data_summary(data, varname=" TN_concentration",
                                                groupnames=("Carbon_source"))
TP _concentration$ Carbon_source =as.factor(TP_concentration$Carbon_source)
plot_7 <- ggplot(TN_concentration, aes(x=Culture_time, y=TN_concentration, group=
Carbon_source)) +
   geom_errorbar(aes(ymin=TN_concentration-sd, ymax=TN_concentration+sd, color=
Carbon_source), width=0.2, position=position_dodge(0))+
   geom_line(aes(color= Carbon_source))+
   geom_point(aes(color= Carbon_source))+
   xlab("Culture time (hr)")+
   ylab("TN Concentration (mg/L)") + labs(title = "", subtitle=NULL) +
   theme_classic() +
   theme(title=element_text(size=30, family="Times New Roman"),
            axis.text.x = element_text(size=25, family="Times New Roman"),
            axis.text.y=element_text(size=25, family="Times New Roman"),
            axis.title.y = element_text(size = 30, family="Times New Roman"),
            axis.title.x=element_text(size=30, family="Times New Roman"),
            legend.position ="top",
            legend.title = element_text(size=30),
            legend.text = element_text(size=30),
            legend.key.size=unit(1,'cm'))+
   scale_shape_manual(values=c(1,2,3,4,5))+
   scale_color_manual(values=c("red","blue","green","black","Purple"))+
   scale_size_manual(values=c(5,5,5,5,5))
plot_7

# TN concentration under 50 µmol/m2/s
data <- metadata[which(metadata$Light_intensity=="50"),]
TN_concentration <- data_summary(data, varname=" TN_concentration",
                                                groupnames=("Carbon_source"))
TP _concentration$ Carbon_source =as.factor(TP_concentration$Carbon_source)
plot_8<- ggplot(TN_concentration, aes(x=Culture_time, y=TN_concentration, group=
Carbon_source)) +
   geom_errorbar(aes(ymin=TN_concentration-sd, ymax=TN_concentration+sd, color=
Carbon_source), width=0.2, position=position_dodge(0))+
   geom_line(aes(color= Carbon_source))+
   geom_point(aes(color= Carbon_source))+
   xlab("Culture time (hr)")+
   ylab("TN Concentration (mg/L)") + labs(title = "", subtitle=NULL) +
   theme_classic() +
   theme(title=element_text(size=30, family="Times New Roman"),
            axis.text.x = element_text(size=25, family="Times New Roman"),
```

```
            axis.text.y=element_text(size=25, family="Times New Roman"),
            axis.title.y = element_text(size = 30, family="Times New Roman"),
            axis.title.x=element_text(size=30, family="Times New Roman"),
            legend.position ="top",
            legend.title = element_text(size=30),
            legend.text = element_text(size=30),
            legend.key.size=unit(1,'cm'))+
   scale_shape_manual(values=c(1,2,3,4,5))+
   scale_color_manual(values=c("red","blue","green","black","Purple"))+
   scale_size_manual(values=c(5,5,5,5,5))
plot_8


#Box plot
# Select data file
con <-file.choose(new = FALSE)
metadata <- read.table(con, header = T, row.names = 1)

#Biomass concentration
Biomass_concentration <- data_summary(metadata, varname="Biomass_concentration",
                                                groupnames=c("Carbon_source "))
Biomass_concentration$ Carbon_source =as.factor(Biomass_concentration$ Carbon_source)

box_1 <- ggplot(Biomass_concentration, aes(x=Carbon_source, y=Biomass_concentration, fill=
Carbon_source)) +
   geom_bar(stat="identity", position=position_dodge(0.9), width=0.5)+
   geom_errorbar(aes(ymin=Biomass_concentration-sd,        ymax=Biomass_concentration+sd),
width=0.2, position=position_dodge(0.9))+
   xlab("")+
   ylab("Biomass concentration (g/L)")    + labs(title = "", subtitle=NULL) +
   theme(title=element_text(size=20, family="Times New Roman"),
            axis.text.x = element_text(size=20, family="Times New Roman"),
            axis.text.y=element_text(size=20, family="Times New Roman"),
            axis.title.y = element_text(size = 20, family="Times New Roman"),
            axis.title.x=element_text(size=20, family="Times New Roman"),
            legend.position="top",)
box_1

#Biomass productivity
Biomass_productivity <- data_summary(metadata, varname="Biomass_productivity ",
                                                groupnames=c("Carbon_source "))
Biomass_productivity $ Carbon_source =as.factor(Biomass_productivity $ Carbon_source)

box_2<-  ggplot(Biomass_productivity,  aes(x=Carbon_source,  y=Biomass_productivity,  fill=
Carbon_source)) +
   geom_bar(stat="identity", position=position_dodge(0.9), width=0.5)+
```

```
   geom_errorbar(aes(ymin=Biomass_productivity -sd, ymax=Biomass_productivity +sd),
width=0.2, position=position_dodge(0.9))+
   xlab("")+
   ylab("Biomass productivity (g/L/day)") + labs(title = "", subtitle=NULL) +
   theme(title=element_text(size=20, family="Times New Roman"),
            axis.text.x = element_text(size=20, family="Times New Roman"),
            axis.text.y=element_text(size=20, family="Times New Roman"),
            axis.title.y = element_text(size = 20, family="Times New Roman"),
            axis.title.x=element_text(size=20, family="Times New Roman"),
            legend.position="top",)
box_2


#TN consumption
TN <- data_summary(metadata, varname="TN_consumption ",
                                                groupnames=c("Carbon_source "))
TN $ Carbon_source =as.factor(TN $ Carbon_source)

box_3<- ggplot(TN, aes(x=Carbon_source, y=TN_consumption, fill= Carbon_source)) +
   geom_bar(stat="identity", position=position_dodge(0.9), width=0.5)+
   geom_errorbar(aes(ymin= TN_consumption -sd, ymax= TN_consumption+sd), width=0.2,
position=position_dodge(0.9))+
   xlab("")+
   ylab("TN consumption (mg/L/day)")    + labs(title = "", subtitle=NULL) +
   theme(title=element_text(size=20, family="Times New Roman"),
            axis.text.x = element_text(size=20, family="Times New Roman"),
            axis.text.y=element_text(size=20, family="Times New Roman"),
            axis.title.y = element_text(size = 20, family="Times New Roman"),
            axis.title.x=element_text(size=20, family="Times New Roman"),
            legend.position="top",)
box_3

#TP consumption
TP<- data_summary(metadata, varname="TP_consumption ",
                                                groupnames=c("Carbon_source "))
TP $ Carbon_source =as.factor(TP $ Carbon_source)

box_4<- ggplot(TP, aes(x=Carbon_source, y=TP_consumption, fill= Carbon_source)) +
   geom_bar(stat="identity", position=position_dodge(0.9), width=0.5)+
   geom_errorbar(aes(ymin= TP_consumption -sd, ymax= TP_consumption+sd), width=0.2,
position=position_dodge(0.9))+
   xlab("")+
   ylab("TP consumption (mg/L/day)") + labs(title = "", subtitle=NULL) +
   theme(title=element_text(size=20, family="Times New Roman"),
            axis.text.x = element_text(size=20, family="Times New Roman"),
            axis.text.y=element_text(size=20, family="Times New Roman"),
```

```
        axis.title.y = element_text(size = 20, family="Times New Roman"),
        axis.title.x=element_text(size=20, family="Times New Roman"),
        legend.position="top",)
box_4

#Carbon capture
CC<- data_summary(metadata, varname="Carbon_capture ",
                                                groupnames=c("Carbon_source "))
CC $ Carbon_source =as.factor(CC $ Carbon_source)

box_5<- ggplot(CC, aes(x=Carbon_source, y= Carbon_capture, fill= Carbon_source)) +
  geom_bar(stat="identity", position=position_dodge(0.9), width=0.5)+
  geom_errorbar(aes(ymin=Carbon_capture-sd,    ymax=Carbon_capture    +sd),    width=0.2,
position=position_dodge(0.9))+
  xlab("")+
  ylab("Carbon capture (%)")    + labs(title = "", subtitle=NULL) +
  theme(title=element_text(size=20, family="Times New Roman"),
        axis.text.x = element_text(size=20, family="Times New Roman"),
        axis.text.y=element_text(size=20, family="Times New Roman"),
        axis.title.y = element_text(size = 20, family="Times New Roman"),
        axis.title.x=element_text(size=20, family="Times New Roman"),
        legend.position="top",)
box_5


##Microbial community analysis##
## Load libraries -----
  library(vegan)
  library(phyloseq)
  library (MASS)
  library(ggplot2)
  library(grid)
  library(gridExtra)
  library(ggpubr)
  library(extrafont)
  loadfonts(device="win")

## Import data files -----
  #Choose Frequency_Table.txt (change gene frequency to relative frequency (%))
  con <- file.choose(new = FALSE)
  Frequency_Table <- read.table(con, header = T, row.names = 1)
  #Choose Frequency_Table_taxonomy.txt
  con1 <-file.choose(new = FALSE)
  Frequency_Table_taxonomy <- read.delim(con1, header = T, row.names = 1)

## Phyloseq -----
```

```
   Full_Frequency <- cbind.data.frame(Frequency_Table, Frequency_Table_taxonomy)
   Frequency <- otu_table(Frequency_Table,taxa_are_rows = TRUE) #Frequency table production for phyloseq
   TAX <- tax_table(as.matrix(Frequency_Table_taxonomy)) #Taxanomy production for phyloseq
   physeq <- phyloseq(Frequency, TAX) #physeq document production
   physeq0 <- tax_glom(physeq, taxrank=rank_names(physeq)[7], NArm=TRUE, bad_empty=c(NA, "", " ", "\t"))
   tax_table(physeq0)
   TAX

#Relative abundance at Domain level
physeqa <-tax_glom(physeq, taxrank=rank_names(physeq)[1], NArm=TRUE, bad_empty=c(NA, "", " ", "\t"))
   tablea <- otu_table(physeqa)
   tablea
   a = plot_bar(physeqa, fill = "Domain") +
      geom_bar(aes(color=Domain, fill=Domain), stat = "identity", position = "stack") +
      xlab("Culture Conditions") + ylab("Relative Frequency (%)") +
      theme(legend.position="right",
            axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust = 1),
            axis.text.y = element_text(size = 11, family="Times New Roman"),
            axis.title.x = element_text(size = 12, family="Times New Roman"),
            axis.title.y = element_text(size = 12, family="Times New Roman"),
            legend.text = element_text(size = 11, family="Times New Roman"),
            legend.title= element_text(size = 12, family="Times New Roman"))
a

#Relative abundance at Phylum level
physeqa1 <-tax_glom(physeq, taxrank=rank_names(physeq)[2], NArm=TRUE, bad_empty=c(NA, "", " ", "\t"))
   tablea1 <- otu_table(physeqa1)

   a1 = plot_bar(physeqa1, fill = "Phylum") +
      geom_bar(aes(color=Phylum, fill=Phylum), stat = "identity", position = "stack") +
      xlab("Culture Conditions") + ylab("Relative Frequency (%)")    +
      theme(legend.position="right",
               axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust = 1),
               axis.text.y = element_text(size = 11, family="Times New Roman"),
               axis.title.x = element_text(size = 12, family="Times New Roman"),
               axis.title.y = element_text(size = 12, family="Times New Roman"),
               legend.text = element_text(size = 11, family="Times New Roman"),
               legend.title= element_text(size = 12, family="Times New Roman"))
a1
```

```
# Bacteroidetes abundance at the family level
physeq3 <-subset_taxa(physeq, Phylum == "Bacteroidetes")
  physeq3_1 <-tax_glom(physeq3, taxrank=rank_names(physeq3)[5], NArm=TRUE,
bad_empty=c(NA, "", " ", "\t"))
  table3_1 <- otu_table(physeq3_1)

  d = plot_bar(physeq3_1, fill = " Family")+ geom_bar(aes(color=Family, fill= Family), stat =
"identity",position = "stack") +
    xlab("") + ylab("Bacteroidetes Relative Frequency (%)") +
    theme(legend.position="right",
          axis.text.x = element_text(size = 17, family="Times New Roman", angle = 90, hjust
= 1),
          axis.text.y = element_text(size = 11, family="Times New Roman"),
          axis.title.x = element_text(size = 12, family="Times New Roman"),
          axis.title.y = element_text(size = 17, family="Times New Roman"),
          legend.text = element_text(size = 15, family="Times New Roman"),
          legend.title= element_text(size = 15, family="Times New Roman"))
d
# Proteobacteria abundance at the family level
physeq3 <-subset_taxa(physeq, Phylum == "Proteobacteria")
  physeq3_1 <-tax_glom(physeq3, taxrank=rank_names(physeq3)[5], NArm=TRUE,
bad_empty=c(NA, "", " ", "\t"))
  table3_1 <- otu_table(physeq3_1)

  d = plot_bar(physeq3_1, fill = " Family")+ geom_bar(aes(color=Family, fill= Family), stat =
"identity",position = "stack") +
    xlab("") + ylab("Proteobacteria Relative Frequency (%)") +
    theme(legend.position="right",
          axis.text.x = element_text(size = 17, family="Times New Roman", angle = 90, hjust
= 1),
          axis.text.y = element_text(size = 11, family="Times New Roman"),
          axis.title.x = element_text(size = 12, family="Times New Roman"),
          axis.title.y = element_text(size = 17, family="Times New Roman"),
          legend.text = element_text(size = 15, family="Times New Roman"),
          legend.title= element_text(size = 15, family="Times New Roman"))
d
```

```
# Alpha-diversity

library(vegan)
library(phyloseq)
library (MASS)
library(ggplot2)
library(grid)
library(gridExtra)
library(ggpubr)

library(extrafont)
font_import() #It may take a few minutes to import.
loadfonts(device="win")

## the .txt file needs to be saved as the type of "Tab delimited".

#Gene frequency data from QIIME2

## Choose data files -----

#Choose the Frequency_Table.txt
con <- file.choose(new = FALSE)
Frequency_Table <- read.table(con, header = T, row.names = 1)
#Choose the Frequency_Table_taxanomy.txt
con1 <-file.choose(new = FALSE)
Frequency_Table_taxonomy <- read.delim(con1, header = T, row.names = 1)

## Alpha Diversity -----
#Create a matrix object with the data frame
t.Frequency.table <- t(Frequency_Table)    # Transpose the data
class(t.Frequency.table) # Check the class of the table

#Alpha diversity analysis indexes
#Shannon
H <- diversity(t.Frequency.table, index = "shannon", MARGIN = 1, base = exp(1))
#Simpson
D <- diversity(t.Frequency.table, "simpson", MARGIN = 1, base = exp(1))
#Inverse Simpson
iD <- diversity(t.Frequency.table, "inv")
#Pielou's evenness
J<-H/log(specnumber(t.Frequency.table))
#List all indexes
IN <- cbind(H,D,iD,J)
IN
write.csv(IN, "diversity.csv")
```

```
#Plot H, D, iD, and J
plot(H)
plot(D)
plot(iD)
plot(J)

#Estimate Chao1 and ACE
estimateR(t.Frequency.table)

## ANOVA for Alpha Diversity -----

#Using the H, D, iD, and J data to generate "diversity.txt" to run ANOVA
#Choose diversity_30L_10_2019.txt

con2 <-file.choose(new = FALSE)
alphadiversity <- read.table(con2, header = T, row.names = 1)

#Define factor for alpha diversity
alphadiversity$Carbon_source <- factor(alphadiversity$Carbon_source)
# Normality
shapiro.test(alphadiversity$H)
shapiro.test(alphadiversity$D)
shapiro.test(alphadiversity$iD)
shapiro.test(alphadiversity$J)
#ANOVA of H index
Hfit <- aov(H ~ Carbon_source, data = alphadiversity)
summary(Hfit)

#ANOVA of J index
Jfit <- aov(J ~ Carbon_source, data = alphadiversity)
summary(Jfit)


## Rarefaction -----

col <- c("black", "darkred", "forestgreen", "orange", "blue", "yellow", "hotpink")
lty <- c("solid", "dashed", "longdash", "dotdash")
pars <- expand.grid(col = col, lty = lty, stringsAsFactors = FALSE)
head(pars)
ra <- rarecurve(t.Frequency.table, step = 20, col =col,lty = lty, cex = 0.6) # curve of rarefication
rad <- rad.lognormal(t.Frequency.table) # Rank of Abundance
rad1 <- plot(rad, xlab = "Rank", ylab = "Abundance") # Plotting the rank

##Beta Diversity##
con3 <-file.choose(new = FALSE)
metadata <- read.table(con3, header = T, row.names = 1, fill = TRUE)
```

```
# Define factors for metadata -----
metadata$Light_intensity <- factor(metadata$Light_intensity)

#Permutational analysis of variance
t.Frequency.table <- t(Frequency_Table)    #transpose the data
class(t.Frequency.table) #check the class of the table
View(Frequency_Table)
View(t.Frequency.table)

betad <-betadiver(t.Frequency.table, 'z')
betad

#adonis
adonis(betad~Formate_conc, metadata, perm=200)
```

CHAPTER 3

```
library (MASS)
library(ggplot2)
library(grid)
library(gridExtra)
library(ggpubr)
library(plyr)
library(inferr)
library(extrafont)
font_import() #It may take a few minutes to import.
loadfonts(device="win")

data_summary <- function(data, varname, groupnames){
   require(plyr)
   summary_func <- function(x, col){
      c(mean = mean(x[[col]], na.rm=TRUE),
         sd = sd(x[[col]], na.rm=TRUE))
   }
   data_sum<-ddply(data, groupnames, .fun=summary_func,
                           varname)
   data_sum <- rename(data_sum, c("mean" = varname))
   return(data_sum)
}

con <-file.choose(new = FALSE)
#metadata<-read_xlsx(con)
```

```
metadata <- read.table(con, header = T, row.names = 1, fill = TRUE)
head(metadata)

metadata$Formate_conc <- factor(metadata$Formate_conc)
metadata$Light_intensity <- factor(metadata$Light_intensity)


## Plots ##
#1. Biomass concentration ---1g/L formate concentration Harvesting rate of 30% under Light intensity 180 umol/m2/s and 500 umol/m2/s

data1 <- metadata[which(metadata$Formate_conc=="1"),]
data1$Light_intensity<-factor(data1$Light_intensity)
data1$Culture_time<-factor(data1$Culture_time)
Biomass_concentration <- data_summary(data1, varname="Biomass_concentration",

groupnames=c("Formate_conc","Light_intensity","Culture_time"))
Biomass_concentration$Formate_conc=as.factor(Biomass_concentration$Formate_conc)
Biomass_concentration$Culture_time=as.factor(Biomass_concentration$Culture_time)
Biomass_concentration$Light_intensity=as.factor(Biomass_concentration$Light_intensity)
head(Biomass_concentration)
Biomass_concentration

plot_1 <- ggplot(Biomass_concentration, aes(x=Culture_time, y=Biomass_concentration, group=Light_intensity)) +
  geom_errorbar(aes(ymin=Biomass_concentration-sd, ymax=Biomass_concentration+sd, color=Light_intensity), width=0.2, position=position_dodge(0))+
  geom_line(aes(color=Light_intensity))+
  geom_point(aes(color=Light_intensity))+
  xlab("Culture time (hr)")+
ylim(0,1.5)+
  ylab("Biomass Concentration (g/L)") + labs(title = "", subtitle=NULL) +
  theme_classic() +
  theme(title=element_text(size=30, family="Times New Roman"),
        axis.text.x = element_text(size=25, family="Times New Roman"),
        axis.text.y=element_text(size=25, family="Times New Roman"),
        axis.title.y = element_text(size = 30, family="Times New Roman"),
        axis.title.x=element_text(size=30, family="Times New Roman"),
        legend.position ="top",
        legend.title = element_text(size=30),
        legend.text = element_text(size=30),
        legend.key.size=unit(1,'cm'))+
  scale_shape_manual(values=c(1,2,3,4,5))+
  scale_color_manual(values=c("red","blue","green","black","Purple"))+
  scale_size_manual(values=c(5,5,5,5,5))
plot_1
```

#2. Biomass concentration --- 2g/L formate concentration Harvesting rate of 30% under Light intensity 180 umol/m2/s and 500 umol/m2/s

```
data2 <- metadata[which(metadata$Formate_conc=="2"),]
data2$Light_intensity<-factor(data2$Light_intensity)
data2$Culture_time <-factor(data2$Culture_time)

Biomass_concentration <- data_summary(data2, varname="Biomass_concentration",

groupnames=c("Formate_conc","Light_intensity","Culture_time"))
Biomass_concentration$Formate_conc=as.factor(Biomass_concentration$Formate_conc)
Biomass_concentration$Culture_time=as.factor(Biomass_concentration$Culture_time)
Biomass_concentration$Light_intensity=as.factor(Biomass_concentration$Light_intensity)
head(Biomass_concentration)
Biomass_concentration

plot_2 <- ggplot(Biomass_concentration, aes(x=Culture_time, y=Biomass_concentration, group=Light_intensity)) +
  geom_errorbar(aes(ymin=Biomass_concentration-sd, ymax=Biomass_concentration+sd, color=Light_intensity), width=0.2, position=position_dodge(0))+
  geom_line(aes(color=Light_intensity))+
  geom_point(aes(color=Light_intensity))+
  xlab("Culture time (hr)")+
ylim(0,1.5)+
  ylab("Biomass Concentration (g/L)") + labs(title = "", subtitle=NULL) +
  theme_classic() +
  theme(title=element_text(size=30, family="Times New Roman"),
        axis.text.x = element_text(size=25, family="Times New Roman"),
        axis.text.y=element_text(size=25, family="Times New Roman"),
        axis.title.y = element_text(size = 30, family="Times New Roman"),
        axis.title.x=element_text(size=30, family="Times New Roman"),
        legend.position ="top",
        legend.title = element_text(size=30),
        legend.text = element_text(size=30),
        legend.key.size=unit(1,'cm'))+
  scale_shape_manual(values=c(1,2,3,4,5))+
  scale_color_manual(values=c("red","blue","green","black","Purple"))+
  scale_size_manual(values=c(5,5,5,5,5))
plot_2

data3 <- metadata[which(metadata$Formate_conc=="1"),]
data3$Light_intensity<-factor(data3$Light_intensity)
data3$Culture_time <-factor(data3$Culture_time)
```

```
Formate_conc <- data_summary(data3, varname="Formate_concentration", groupnames=c("Formate_conc","Light_intensity","Culture_time"))
Formate_conc$Formate_conc=as.factor(Formate_conc$Formate_conc)
Formate_conc$Culture_time=as.factor(Formate_conc$Culture_time)
Formate_conc$Light_intensity=as.factor(Formate_conc$Light_intensity)
head(Formate_conc)
Formate_conc

plot_3 <- ggplot(Formate_conc, aes(x=Culture_time, y=Formate_concentration, group=Light_intensity)) +
  geom_errorbar(aes(ymin=Formate_concentration-sd, ymax=Formate_concentration+sd, color=Light_intensity), width=0.2, position=position_dodge(0))+
  geom_line(aes(color=Light_intensity))+
  geom_point(aes(color=Light_intensity))+
  xlab("Culture time (hr)")+
ylim(-0.05,1)+
  ylab("Formate Concentration (g/L)") + labs(title = "", subtitle=NULL) +
  theme_classic() +
  theme(title=element_text(size=30, family="Times New Roman"),
        axis.text.x = element_text(size=25, family="Times New Roman"),
        axis.text.y=element_text(size=25, family="Times New Roman"),
        axis.title.y = element_text(size = 30, family="Times New Roman"),
        axis.title.x=element_text(size=30, family="Times New Roman"),
        legend.position ="top",
        legend.title = element_text(size=30),
        legend.text = element_text(size=30),
        legend.key.size=unit(1,'cm'))+
  scale_shape_manual(values=c(1,2,3,4,5))+
  scale_color_manual(values=c("red","blue","green","black","Purple"))+
  scale_size_manual(values=c(5,5,5,5,5))
plot_3

#4. Formate utilization---2g/L formate concentration Harvesting rate of 30% under Light intensity 180 umol/m2/s and 500 umol/m2/s
data4 <- metadata[which(metadata$Formate_conc=="2"),]
data4$Light_intensity<-factor(data4$Light_intensity)
data4$Culture_time <-factor(data4$Culture_time)


Formate_conc <- data_summary(data4, varname="Formate_concentration", groupnames=c("Formate_conc","Light_intensity","Culture_time"))
Formate_conc$Formate_conc=as.factor(Formate_conc$Formate_conc)
Formate_conc$Culture_time=as.factor(Formate_conc$Culture_time)
Formate_conc$Light_intensity=as.factor(Formate_conc$Light_intensity)
head(Formate_conc)
Formate_conc
```

```
plot_4 <- ggplot(Formate_conc, aes(x=Culture_time, y=Formate_concentration, group=Light_intensity)) +
  geom_errorbar(aes(ymin=Formate_concentration-sd, ymax=Formate_concentration+sd, color=Light_intensity), width=0.2, position=position_dodge(0))+
  geom_line(aes(color=Light_intensity))+
  geom_point(aes(color=Light_intensity))+
  xlab("Culture time (hr)")+
ylim(-0.05,1)+
 ylab("Formate Concentration (g/L)") + labs(title = "", subtitle=NULL) +
  theme_classic() +
  theme(title=element_text(size=30, family="Times New Roman"),
        axis.text.x = element_text(size=25, family="Times New Roman"),
        axis.text.y=element_text(size=25, family="Times New Roman"),
        axis.title.y = element_text(size = 30, family="Times New Roman"),
        axis.title.x=element_text(size=30, family="Times New Roman"),
        legend.position ="top",
        legend.title = element_text(size=30),
        legend.text = element_text(size=30),
        legend.key.size=unit(1,'cm'))+
  scale_shape_manual(values=c(1,2,3,4,5))+
  scale_color_manual(values=c("red","blue","green","black","Purple"))+
  scale_size_manual(values=c(5,5,5,5,5))
 plot_4
#5. Total Nitrogen---1g/L formate concentration Harvesting rate of 30% under Light intensity 180 umol/m2/s and 500 umol/m2/s
data5 <- metadata[which(metadata$Formate_conc=="1"),]
data5$Light_intensity<-factor(data5$Light_intensity)
data5$Culture_time <-factor(data5$Culture_time)


TN <- data_summary(data5, varname="TN",

groupnames=c("Formate_conc","Light_intensity","Culture_time"))
TN$Formate_conc=as.factor(TN$Formate_conc)
TN$Culture_time=as.factor(TN$Culture_time)
TN$Light_intensity=as.factor(TN$Light_intensity)
head(TN)
TN

plot_5 <- ggplot(TN, aes(x=Culture_time, y=TN, group=Light_intensity)) +
  geom_errorbar(aes(ymin=TN-sd, ymax=TN+sd, color=Light_intensity), width=0.2, position=position_dodge(0))+
  geom_line(aes(color=Light_intensity))+
  geom_point(aes(color=Light_intensity))+
  xlab("Culture time (hr)")+
  ylab("Total Nitrogen (mg/L)") + labs(title = "", subtitle=NULL) +
```

```
  theme_classic() +
   theme(title=element_text(size=30, family="Times New Roman"),
         axis.text.x = element_text(size=25, family="Times New Roman"),
         axis.text.y=element_text(size=25, family="Times New Roman"),
         axis.title.y = element_text(size = 30, family="Times New Roman"),
         axis.title.x=element_text(size=30, family="Times New Roman"),
         legend.position ="top",
         legend.title = element_text(size=30),
         legend.text = element_text(size=30),
         legend.key.size=unit(1,'cm'))+
  scale_shape_manual(values=c(1,2,3,4,5))+
  scale_color_manual(values=c("red","blue","green","black","Purple"))+
  scale_size_manual(values=c(5,5,5,5,5))plot_5

#6. Total Phosphorus---1g/L formate concentration Harvesting rate of 30% under Light intensity
180 umol/m2/s and 500 umol/m2/s
data6 <- metadata[which(metadata$Formate_conc=="1"),]
data6$Light_intensity<-factor(data6$Light_intensity)
data6$Culture_time <-factor(data6$Culture_time)


TP <- data_summary(data6, varname="TP",
groupnames=c("Formate_conc","Light_intensity","Culture_time"))
TP$Formate_conc=as.factor(TP$Formate_conc)
TP$Culture_time=as.factor(TP$Culture_time)
TP$Light_intensity=as.factor(TP$Light_intensity)
head(TP)
TP

plot_6 <- ggplot(TP, aes(x=Culture_time, y=TP, group=Light_intensity)) +
  geom_errorbar(aes(ymin=TP-sd, ymax=TP+sd, color=Light_intensity), width=0.2,
position=position_dodge(0))+
  geom_line(aes(color=Light_intensity))+
  geom_point(aes(color=Light_intensity))+
  xlab("Culture time (hr)")+
  ylab("Total Phosphorus (mg/L)") + labs(title = "", subtitle=NULL) +
  theme_classic() +
  theme(title=element_text(size=30, family="Times New Roman"),
         axis.text.x = element_text(size=25, family="Times New Roman"),
         axis.text.y=element_text(size=25, family="Times New Roman"),
         axis.title.y = element_text(size = 30, family="Times New Roman"),
         axis.title.x=element_text(size=30, family="Times New Roman"),
         legend.position ="top",
         legend.title = element_text(size=30),
         legend.text = element_text(size=30),
         legend.key.size=unit(1,'cm'))+
```

```
  scale_shape_manual(values=c(1,2,3,4,5))+
  scale_color_manual(values=c("red","blue","green","black","Purple"))+
  scale_size_manual(values=c(5,5,5,5,5))plot_6

#7. Total Nitrogen---2g/L formate concentration Harvesting rate of 30% under Light intensity 180 umol/m2/s and 500 umol/m2/s

data7 <- metadata[which(metadata$Formate_conc=="2"),]
data7$Light_intensity<-factor(data4$Light_intensity)
data7$Culture_time <-factor(data4$Culture_time)


TN <- data_summary(data7, varname="TN", groupnames=c("Formate_conc","Light_intensity","Culture_time"))
TN$Formate_conc=as.factor(TN$Formate_conc)
TN$Culture_time=as.factor(TN$Culture_time)
TN$Light_intensity=as.factor(TN$Light_intensity)
head(TN)
TN

plot_7 <- ggplot(TN, aes(x=Culture_time, y=TN, group=Light_intensity)) +
  geom_errorbar(aes(ymin=TN-sd, ymax=TN+sd, color=Light_intensity), width=0.2, position=position_dodge(0))+
  geom_line(aes(color=Light_intensity))+
  geom_point(aes(color=Light_intensity))+
  xlab("Culture time (hr)")+
  ylab(" Total Nitrogen (mg/L)") + labs(title = "", subtitle=NULL) +
  theme_classic() +
  theme(title=element_text(size=30, family="Times New Roman"),
        axis.text.x = element_text(size=25, family="Times New Roman"),
        axis.text.y=element_text(size=25, family="Times New Roman"),
        axis.title.y = element_text(size = 30, family="Times New Roman"),
        axis.title.x=element_text(size=30, family="Times New Roman"),
        legend.position ="top",
        legend.title = element_text(size=30),
        legend.text = element_text(size=30),
        legend.key.size=unit(1,'cm'))+
  scale_shape_manual(values=c(1,2,3,4,5))+
  scale_color_manual(values=c("red","blue","green","black","Purple"))+
  scale_size_manual(values=c(5,5,5,5,5))plot_7

#8. Total Phosphorus---2g/L formate concentration Harvesting rate of 30% under Light intensity 180 umol/m2/s and 500 umol/m2/s
data8 <- metadata[which(metadata$Formate_conc=="2"),]
data8$Light_intensity<-factor(data8$Light_intensity)
data8$Culture_time <-factor(data8$Culture_time)
```

```
TP <- data_summary(data7, varname="TP", groupnames=c("Formate_conc","Light_intensity","Culture_time"))
TP$Formate_conc=as.factor(TP$Formate_conc)
TP$Culture_time=as.factor(TP$Culture_time)
TP$Light_intensity=as.factor(TP$Light_intensity)
head(TP)
TP

plot_8 <- ggplot(TP, aes(x=Culture_time, y=TP, group=Light_intensity)) +
  geom_errorbar(aes(ymin=TP-sd, ymax=TP+sd, color=Light_intensity), width=0.2, position=position_dodge(0))+
  geom_line(aes(color=Light_intensity))+
  geom_point(aes(color=Light_intensity))+
  xlab("Culture time (hr)")+
  ylab("Total Phosphorus (mg/L)") + labs(title = "", subtitle=NULL) +
  theme_classic() +
  theme(title=element_text(size=30, family="Times New Roman"),
        axis.text.x = element_text(size=25, family="Times New Roman"),
        axis.text.y=element_text(size=25, family="Times New Roman"),
        axis.title.y = element_text(size = 30, family="Times New Roman"),
        axis.title.x=element_text(size=30, family="Times New Roman"),
        legend.position ="top",
        legend.title = element_text(size=30),
        legend.text = element_text(size=30),
        legend.key.size=unit(1,'cm'))+
  scale_shape_manual(values=c(1,2,3,4,5))+
  scale_color_manual(values=c("red","blue","green","black","Purple"))+
  scale_size_manual(values=c(5,5,5,5,5))plot_8

#9. Biomass comparison

Biomass_concentration <- data_summary(metadata, varname="Biomass_concentration",

groupnames=c("Light_intensity","Formate_conc"))

Biomass_concentration$Light_intensity=as.factor(Biomass_concentration$Light_intensity)
Biomass_concentration$Formate_conc=as.factor(Biomass_concentration$Formate_conc)
head(Biomass_concentration)

box_1 <- ggplot(Biomass_concentration, aes(x=Formate_conc, y=Biomass_concentration, fill=Light_intensity)) +
  geom_bar(stat="identity", position=position_dodge(0.9), width=0.5)+
  geom_errorbar(aes(ymin=Biomass_concentration-sd, ymax=Biomass_concentration+sd), width=0.2, position=position_dodge(0.9))+
  xlab("Formate feeding concentration (g/L/day)")+
  ylab("Biomass concentration (g/L)") + ylim(0, 1.5)    + labs(title = "", subtitle=NULL) +
```

```
  theme(title=element_text(size=20, family="Times New Roman"),
        axis.text.x = element_text(size=20, family="Times New Roman"),
        axis.text.y=element_text(size=20, family="Times New Roman"),
        axis.title.y = element_text(size = 20, family="Times New Roman"),
        axis.title.x=element_text(size=20, family="Times New Roman"),
        legend.position="top",)+
  labs(fill="Light intensity (uE/m2/s)")

box_1

#10. TN comparison

TN <- data_summary(metadata, varname="TN",
                   groupnames=c("Light_intensity","Formate_conc"))

TN$Light_intensity=as.factor(TN$Light_intensity)
TN$Formate_conc=as.factor(TN$Formate_conc)
head(TN)

box_2 <- ggplot(TN, aes(x=Formate_conc, y=TN, fill=Light_intensity)) +
  geom_bar(stat="identity", position=position_dodge(0.9), width=0.5)+
  geom_errorbar(aes(ymin=TN-sd, ymax=TN+sd), width=0.2, position=position_dodge(0.9))+
  xlab("Formate feeding concentration (g/L/day)")+
  ylab("TN (mg/L)") + ylim(0, 120)    + labs(title = "", subtitle=NULL) +
  theme(title=element_text(size=20, family="Times New Roman"),
        axis.text.x = element_text(size=20, family="Times New Roman"),
        axis.text.y=element_text(size=20, family="Times New Roman"),
        axis.title.y = element_text(size = 20, family="Times New Roman"),
        axis.title.x=element_text(size=20, family="Times New Roman"),
        legend.position="top",)+
  labs(fill="Light intensity (uE/m2/s)")

#11. TP comparison
TP <- data_summary(metadata, varname="TP",
                   groupnames=c("Light_intensity","Formate_conc"))

TP$Light_intensity=as.factor(TP$Light_intensity)
TP$Formate_conc=as.factor(TP$Formate_conc)
head(TN)

box_3 <- ggplot(TP, aes(x=Formate_conc, y=TP, fill=Light_intensity)) +
  geom_bar(stat="identity", position=position_dodge(0.9), width=0.5)+
  geom_errorbar(aes(ymin=TP-sd, ymax=TP+sd), width=0.2, position=position_dodge(0.9))+
  xlab("Formate feeding concentration (g/L/day)")+
  ylab("TP (mg/L)") + ylim(0, 120)    + labs(title = "", subtitle=NULL) +
  theme(title=element_text(size=20, family="Times New Roman"),
```

```
        axis.text.x = element_text(size=20, family="Times New Roman"),
        axis.text.y=element_text(size=20, family="Times New Roman"),
        axis.title.y = element_text(size = 20, family="Times New Roman"),
        axis.title.x=element_text(size=20, family="Times New Roman"),
        legend.position="top",)+
  labs(fill="Light intensity (uE/m2/s)")

box_3


##Statistical analysis##
metadata$Formate_conc <- factor(metadata$Formate_conc)
metadata$Light_intensity <- factor(metadata$Light_intensity)
metadata$Culture_time <- factor(metadata$Culture_time)

# Normality
shapiro.test(metadata$Biomass_concentration)
shapiro.test(metadata$TN)
shapiro.test(metadata$TP)

# Variance
var.test(Biomass_concentration ~ Formate_conc, data = metadata)
var.test(Biomass_concentration ~ Light_intensity, data = metadata)
var.test(TP ~ Formate_conc, data = metadata)
var.test(TP ~ Light_intensity, data = metadata)
var.test(TN ~ Formate_conc, data = metadata)
var.test(TN ~ Light_intensity, data = metadata)

# Two-way ANOVA and pair-wise comparison
aov=aov(Biomass_concentration~ Formate_conc*Light_intensity, data=metadata)
summary(fit1)
Tukey1 <- TukeyHSD(fit1, conf.level=0.95)
Tukey1
aov=aov(TN~Light_intensity*Formate_conc,data=metadata)
summary(aov)
Tukey2 <- TukeyHSD(aov, conf.level=0.95)
Tukey2

aov=aov(TP~Light_intensity*Formate_conc,data=metadata)
summary(aov)
Tukey3 <- TukeyHSD(aov, conf.level=0.95)
Tukey3

#qPCR
#Select data file
con <- file.choose(new = FALSE)
```

```
   metadata <- read.table(con, header = T, row.names = 1)
metadata$Cell <- factor(metadata$Cell )

   data_summary <- function(data, varname, groupnames){
      require(plyr)
      summary_func <- function(x, col){
         c(mean = mean(x[[col]], na.rm=TRUE),
            sd = sd(x[[col]], na.rm=TRUE))
      }
      data_sum<-ddply(data, groupnames, .fun=summary_func,
                              varname)
      data_sum <- rename(data_sum, c("mean" = varname))
      return(data_sum)
   }

data <- data_summary(metadata, varname="Cq_value", groupnames=("Cell"))

box_1 <- ggplot(data, aes(x=Cell, y=Cq_value, fill=Cell)) +
   geom_bar(stat="identity", position=position_dodge(0.9), width=0.5)+
   geom_errorbar(aes(ymin=Cq_value-sd,                ymax=Cq_value+sd),                width=0.2,
position=position_dodge(0.9))+
   xlab("")+
   ylab("Cq value") + ylim(0, 30)    + labs(title = "", subtitle=NULL) +
   theme(title=element_text(size=20, family="Times New Roman"),
            axis.text.x = element_text(size=20, family="Times New Roman"),
            axis.text.y=element_text(size=20, family="Times New Roman"),
            axis.title.y = element_text(size = 20, family="Times New Roman"),
            axis.title.x=element_text(size=20, family="Times New Roman"),
            legend.position="right",)+
   labs(fill="")
box_1


##Microbial community analysis##
library(vegan)
   library(phyloseq)
   library (MASS)
   library(ggplot2)
   library(grid)
   library(gridExtra)
   library(ggpubr)
   library(extrafont)
   loadfonts(device="win")

## Import data files -----
   #Choose Frequency_Table_YZ_PartB_Plots.txt (change gene frequency to relative frequency
```

```
(%))
   con <- file.choose(new = FALSE)
   Frequency_Table <- read.table(con, header = T, row.names = 1)
   #Choose Frequency_Table_taxonomy_YZ.txt
   con1 <-file.choose(new = FALSE)
   Frequency_Table_taxonomy <- read.delim(con1, header = T, row.names = 1)
   Full_Frequency <- cbind.data.frame(Frequency_Table, Frequency_Table_taxonomy)
   Frequency <- otu_table(Frequency_Table,taxa_are_rows = TRUE) #Frequency table
production for phyloseq
   TAX <- tax_table(as.matrix(Frequency_Table_taxonomy)) #Taxanomy production for
phyloseq
   physeq <- phyloseq(Frequency, TAX) #physeq document production
   physeq0 <- tax_glom(physeq, taxrank=rank_names(physeq)[7], NArm=TRUE,
bad_empty=c(NA, "", " ", "\t"))
   tax_table(physeq0)

   # Plot 1
   p = plot_bar(physeq0, fill = "Class", facet_grid=Domain~Phylum) +
      xlab("") + ylab("Relative Frequency (%)") +
      geom_bar(color = "black", size = .1, stat = "identity", position = "stack")+
      theme(legend.position="right",
            axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust =
1),
            axis.text.y = element_text(size = 11, family="Times New Roman"),
            axis.title.x = element_text(size = 12, family="Times New Roman"),
            axis.title.y = element_text(size = 12, family="Times New Roman"),
            legend.text = element_text(size = 11, family="Times New Roman"),
            legend.title= element_text(size = 12, family="Times New Roman"))
   p

# Plot 2
physeqa <-tax_glom(physeq, taxrank=rank_names(physeq)[1], NArm=TRUE, bad_empty=c(NA,
"", " ", "\t"))
   tablea <- otu_table(physeqa)

   a = plot_bar(physeqa, fill = "Domain") +
      geom_bar(aes(color=Domain, fill=Domain), stat = "identity", position = "stack") +
      xlab("") + ylab("Relative Frequency (%)") +
      theme(legend.position="right",
            axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust =
1),
            axis.text.y = element_text(size = 11, family="Times New Roman"),
            axis.title.x = element_text(size = 12, family="Times New Roman"),
            axis.title.y = element_text(size = 12, family="Times New Roman"),
            legend.text = element_text(size = 11, family="Times New Roman"),
            legend.title= element_text(size = 12, family="Times New Roman"))
```

```
   a

# Plot 3
physeqa1 <-tax_glom(physeq, taxrank=rank_names(physeq)[2], NArm=TRUE,
bad_empty=c(NA, "", " ", "\t"))
   tablea1 <- otu_table(physeqa1)

   a1 = plot_bar(physeqa1, fill = "Phylum") +
      geom_bar(aes(color=Phylum, fill=Phylum), stat = "identity", position = "stack") +
      xlab("") + ylab("Relative Frequency (%)")    +
      theme(legend.position="right",
               axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust
= 1),
               axis.text.y = element_text(size = 11, family="Times New Roman"),
               axis.title.x = element_text(size = 12, family="Times New Roman"),
               axis.title.y = element_text(size = 12, family="Times New Roman"),
               legend.text = element_text(size = 11, family="Times New Roman"),
               legend.title= element_text(size = 12, family="Times New Roman"))
   a1

# Plot 4

physeqa2 <-tax_glom(physeq, taxrank=rank_names(physeq)[3], NArm=TRUE,
bad_empty=c(NA, "", " ", "\t"))
   tablea2 <- otu_table(physeqa2)

   a2 = plot_bar(physeqa2, fill = "Class") +
      geom_bar(aes(color=Class, fill=Class), stat = "identity", position = "stack") +
      xlab("") + ylab("Relative Frequency (%)") +
      theme(legend.position="right",
               axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust
= 1),
               axis.text.y = element_text(size = 11, family="Times New Roman"),
               axis.title.x = element_text(size = 12, family="Times New Roman"),
               axis.title.y = element_text(size = 12, family="Times New Roman"),
               legend.text = element_text(size = 11, family="Times New Roman"),
               legend.title= element_text(size = 12, family="Times New Roman"))
   a2

# Plot 5
physeq2 <-subset_taxa(physeq, Domain== "Bacteria")
   physeq2_1 <-tax_glom(physeq2, taxrank=rank_names(physeq2)[2], NArm=TRUE,
bad_empty=c(NA, "", " ", "\t"))
   table2_1 <- otu_table(physeq2_1)

   c = plot_bar(physeq2_1, fill = "Phylum") +
```

```
      geom_bar(aes(color=Phylum, fill=Phylum), stat = "identity",position = "stack") +
      xlab("") + ylab("Relative Frequency (%)") +
      theme(legend.position="right",
            axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust
= 1),
            axis.text.y = element_text(size = 11, family="Times New Roman"),
            axis.title.x = element_text(size = 12, family="Times New Roman"),
            axis.title.y = element_text(size = 12, family="Times New Roman"),
            legend.text = element_text(size = 11, family="Times New Roman"),
            legend.title= element_text(size = 12, family="Times New Roman"))
  c

# Plot 6
  physeq3 <-subset_taxa(physeq, Phylum == "Bacteroidetes")
  physeq3_1     <-tax_glom(physeq3,     taxrank=rank_names(physeq3)[5],     NArm=TRUE,
bad_empty=c(NA, "", " ", "\t"))
  table3_1 <- otu_table(physeq3_1)

  d = plot_bar(physeq3_1, fill = "Family")+ geom_bar(aes(color=Family, fill=Family), stat =
"identity",position = "stack") +
      xlab("") + ylab("Bacteroidetes Relative Frequency (%)") +
      theme(legend.position="right",
            axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust
= 1),
            axis.text.y = element_text(size = 11, family="Times New Roman"),
            axis.title.x = element_text(size = 12, family="Times New Roman"),
            axis.title.y = element_text(size = 12, family="Times New Roman"),
            legend.text = element_text(size = 11, family="Times New Roman"),
            legend.title= element_text(size = 12, family="Times New Roman"))
 d

# Plot 7
physeq5 <-subset_taxa(physeq, Phylum == "Proteobacteria")
  physeq5_1     <-tax_glom(physeq5,     taxrank=rank_names(physeq5)[5],     NArm=TRUE,
bad_empty=c(NA, "", " ", "\t"))
  table5_1 <- otu_table(physeq5_1)
  # write.csv(table5_1, "ProteobacteriaFamily.csv")

  f = plot_bar(physeq5_1, fill = "Family")+ geom_bar(aes(color=Family, fill=Family), stat =
"identity",position = "stack") +
      xlab("") + ylab("Proteobacteria Relative Frequency (%)") +
      theme(legend.position="right",
            axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust
= 1),
            axis.text.y = element_text(size = 11, family="Times New Roman"),
            axis.title.x = element_text(size = 12, family="Times New Roman"),
```

```
            axis.title.y = element_text(size = 12, family="Times New Roman"),
            legend.text = element_text(size = 11, family="Times New Roman"),
            legend.title= element_text(size = 12, family="Times New Roman"))
  f

## ANOVA for Alpha Diversity##

#Using the H, D, iD, and J data to generate "diversity.txt" to run ANOVA
#Choose diversity_30L_10_2019.txt

con2 <-file.choose(new = FALSE)
alphadiversity <- read.table(con2, header = T, row.names = 1)

#Define factor for alpha diversity
alphadiversity$Light_intensity <- factor(alphadiversity$Light_intensity)
alphadiversity$Formate_conc <- factor(alphadiversity$Formate_conc)
# Normality
shapiro.test(alphadiversity$H)
shapiro.test(alphadiversity$D)
shapiro.test(alphadiversity$iD)
shapiro.test(alphadiversity$J)

## Rarefaction -----

col <- c("black", "darkred", "forestgreen", "orange", "blue", "yellow", "hotpink")
lty <- c("solid", "dashed", "longdash", "dotdash")
pars <- expand.grid(col = col, lty = lty, stringsAsFactors = FALSE)
head(pars)
ra <- rarecurve(t.Frequency.table, step = 20, col =col,lty = lty, cex = 0.6) # curve of rarefication
rad <- rad.lognormal(t.Frequency.table) # Rank of Abundance
rad1 <- plot(rad, xlab = "Rank", ylab = "Abundance") # Plotting the rank


## ANOVA for Beta Diversity ##
con3 <-file.choose(new = FALSE)
metadata <- read.table(con3, header = T, row.names = 1, fill = TRUE)

# Define factors for metadata -----
metadata$Light_intensity <- factor(metadata$Light_intensity)

#Permutational analysis of variance
t.Frequency.table <- t(Frequency_Table)    #transpose the data
class(t.Frequency.table) #check the class of the table
View(Frequency_Table)
View(t.Frequency.table)
```

```
betad <-betadiver(t.Frequency.table, 'z')
betad

#adonis
adonis(betad~Formate_conc, metadata, perm=200)


##ITS sequencing results
   library(vegan)
   library(phyloseq)
   library (MASS)
   library(ggplot2)
   library(grid)
   library(gridExtra)
   library(ggpubr)
   library(extrafont)
   loadfonts(device="win")

# Fungi relative abundance at phylum level
  physeqa1        <-tax_glom(physeq,        taxrank=rank_names(physeq)[2],        NArm=TRUE,
bad_empty=c(NA, "", " ", "\t"))
   tablea1 <- otu_table(physeqa1)

   a1 = plot_bar(physeqa1, fill = "Phylum") +
      geom_bar(aes(color=Phylum, fill=Phylum), stat = "identity", position = "stack") +
      xlab("") + ylab("Relative Frequency (%)")    +
      theme(legend.position="right",
               axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust
= 1),
               axis.text.y = element_text(size = 11, family="Times New Roman"),
               axis.title.x = element_text(size = 12, family="Times New Roman"),
               axis.title.y = element_text(size = 12, family="Times New Roman"),
               legend.text = element_text(size = 11, family="Times New Roman"),
               legend.title= element_text(size = 12, family="Times New Roman"))+
scale_x_discrete(labels=c("Condition 1","Condition 2","Condition 3","Condition 4"))

   a1
# Fungi relative abundance at family level
physeqa2        <-tax_glom(physeq,        taxrank=rank_names(physeq)[5],        NArm=TRUE,
bad_empty=c(NA, "", " ", "\t"))
   tablea2 <- otu_table(physeqa2)

   a2 = plot_bar(physeqa2, fill = "Family") +
      geom_bar(aes(color=Family, fill=Family), stat = "identity", position = "stack") +
      xlab("") + ylab("Relative Frequency (%)") +
      theme(legend.position="right",
```

```
            axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust
= 1),
            axis.text.y = element_text(size = 11, family="Times New Roman"),
            axis.title.x = element_text(size = 12, family="Times New Roman"),
            axis.title.y = element_text(size = 12, family="Times New Roman"),
            legend.text = element_text(size = 11, family="Times New Roman"),
            legend.title= element_text(size = 12, family="Times New Roman"))+
scale_x_discrete(labels=c("Condition 1","Condition 2","Condition 3","Condition 4"))
  a2



library (MASS)
library(ggplot2)
library(grid)
library(gridExtra)
library(ggpubr)

# Installing the font package ------------------------------------------
Sys.setenv(R_GSCMD="C:/Program Files/gs/gs9.05/bin/gswin32c.exe")
install.packages("extrafontdb")
library(extrafont)
font_import() #It may take a few minutes to import.
loadfonts(device="win")

# PROGRAM TO PLOT BAR CHART WITH STANDARD DEVIATION ----------------------

#+++++++++++++++++++++++++
# Function to calculate the mean and the standard deviation
# for each group
#+++++++++++++++++++++++++
# data : a data frame
# varname : the name of a column containing the variable
#to be summariezed
# groupnames : vector of column names to be used as
# grouping variables
data_summary <- function(data, varname, groupnames){
  require(plyr)
  summary_func <- function(x, col){
    c(mean = mean(x[[col]], na.rm=TRUE),
      sd = sd(x[[col]], na.rm=TRUE))
  }
  data_sum<-ddply(data, groupnames, .fun=summary_func,
                  varname)
  data_sum <- rename(data_sum, c("mean" = varname))
  return(data_sum)
```

```
}


# ANALYSIS---------------------------------------------------------------

## the .txt file needs to be saved as the type of "Tab delimited".

##load meta_data_RegularStat(30L).txt
con <-file.choose(new = FALSE)
metadata <- read.table(con, header = T, row.names = 1)

## DEFINING FACTORS

metadata$Formate_conc <- factor(metadata$carbon_concentration)


data <- data_summary(metadata, varname="Biomass_concentration",
groupnames=c("Carbon_source","Culture_time"))
data$Carbon_source=as.factor(data$Carbon_source)
data$Culture_time=as.factor(data$Culture_time)

plot_1 <- ggplot(data, aes(x=Culture_time, y=Biomass_concentration)) +
  geom_errorbar(aes(ymin=Biomass_concentration-sd, ymax=Biomass_concentration+sd,
color=Carbon_source), width=0.2, position=position_dodge(0))+
  geom_line(aes(color=Carbon_source))+
  geom_point(aes(color=Carbon_source))+
  xlab("Culture time (hr)")+
  ylab("Biomass Concentration (g/L)") + labs(title = "", subtitle=NULL) +
  theme_classic() +
  theme(title=element_text(size=30, family="Times New Roman"),
        axis.text.x = element_text(size=15, family="Times New Roman"),
        axis.text.y=element_text(size=25, family="Times New Roman"),
        axis.title.y = element_text(size = 30, family="Times New Roman"),
        axis.title.x=element_text(size=30, family="Times New Roman"),
        legend.position ="top",
        legend.title = element_text(size=30),
        legend.text = element_text(size=30),
        legend.key.size=unit(1,'cm'))+
  scale_shape_manual(values=c(1,2,3,4,5))+
  scale_color_manual(values=c("red","blue","green","black","Purple"))+
  scale_size_manual(values=c(5,5,5,5,5))

plot_1

data <- data_summary(metadata, varname="carbon_concentration",
groupnames=c("Carbon_source","Culture_time"))
```

```
data$Carbon_source=as.factor(data$Carbon_source)
data$Culture_time=as.factor(data$Culture_time)


plot_3 <- ggplot(data, aes(x=Culture_time, y=carbon_concentration)) +
   geom_errorbar(aes(ymin=carbon_concentration-sd,          ymax=carbon_concentration+sd,
color=Carbon_source), width=0.2, position=position_dodge(0))+
   geom_line(aes(color=Carbon_source))+
   geom_point(aes(color=Carbon_source))+
   xlab("Culture time (hr)")+
   ylab("Carbon Source Concentration (g/L)") + labs(title = "", subtitle=NULL) +
   theme_classic() +
   theme(title=element_text(size=30, family="Times New Roman"),
            axis.text.x = element_text(size=15, family="Times New Roman"),
            axis.text.y=element_text(size=25, family="Times New Roman"),
            axis.title.y = element_text(size = 30, family="Times New Roman"),
            axis.title.x=element_text(size=30, family="Times New Roman"),
            legend.position ="top",
            legend.title = element_text(size=30),
            legend.text = element_text(size=30),
            legend.key.size=unit(1,'cm'))+
   scale_shape_manual(values=c(1,2,3,4,5))+
   scale_color_manual(values=c("red","blue","green","black","Purple"))+
   scale_size_manual(values=c(5,5,5,5,5))
plot_3

library(vegan)
   library(phyloseq)
   library (MASS)
   library(ggplot2)
   library(grid)
   library(gridExtra)
   library(ggpubr)
   library(extrafont)
   loadfonts(device="win")

## Import data files -----
   #Choose Frequency_Table_YZ_PartB_Plots.txt (change gene frequency to relative frequency
(%))
   con <- file.choose(new = FALSE)
   Frequency_Table <- read.table(con, header = T, row.names = 1)
   #Choose Frequency_Table_taxonomy_YZ.txt
   con1 <-file.choose(new = FALSE)
   Frequency_Table_taxonomy <- read.delim(con1, header = T, row.names = 1)

## Phyloseq -----
```

```
  Full_Frequency <- cbind.data.frame(Frequency_Table, Frequency_Table_taxonomy)
  Frequency <- otu_table(Frequency_Table,taxa_are_rows = TRUE) #Frequency table production for phyloseq
  TAX <- tax_table(as.matrix(Frequency_Table_taxonomy)) #Taxanomy production for phyloseq
  physeq <- phyloseq(Frequency, TAX) #physeq document production
  physeq0 <- tax_glom(physeq, taxrank=rank_names(physeq)[7], NArm=TRUE, bad_empty=c(NA, "", " ", "\t"))
  tax_table(physeq0)
  TAX

#Plot 1
  physeqa <-tax_glom(physeq, taxrank=rank_names(physeq)[1], NArm=TRUE, bad_empty=c(NA, "", " ", "\t"))
  tablea <- otu_table(physeqa)

  a = plot_bar(physeqa, fill = "Domain") +
    geom_bar(aes(color=Domain, fill=Domain), stat = "identity", position = "stack") +
    xlab("") + ylab("Relative Frequency (%)") +
    theme(legend.position="right",
          axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust = 1),
          axis.text.y = element_text(size = 11, family="Times New Roman"),
          axis.title.x = element_text(size = 12, family="Times New Roman"),
          axis.title.y = element_text(size = 12, family="Times New Roman"),
          legend.text = element_text(size = 11, family="Times New Roman"),
          legend.title= element_text(size = 12, family="Times New Roman"))
  a
#Plot 2
physeqa1 <-tax_glom(physeq, taxrank=rank_names(physeq)[2], NArm=TRUE, bad_empty=c(NA, "", " ", "\t"))
  tablea1 <- otu_table(physeqa1)

  a1 = plot_bar(physeqa1, fill = "Phylum") +
    geom_bar(aes(color=Phylum, fill=Phylum), stat = "identity", position = "stack") +
    xlab("") + ylab("Relative Frequency (%)")  +
    theme(legend.position="right",
          axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust = 1),
          axis.text.y = element_text(size = 11, family="Times New Roman"),
          axis.title.x = element_text(size = 12, family="Times New Roman"),
          axis.title.y = element_text(size = 12, family="Times New Roman"),
          legend.text = element_text(size = 11, family="Times New Roman"),
          legend.title= element_text(size = 12, family="Times New Roman"))
  a1
```

```
## Load libraries -----
   library(vegan)
   library(phyloseq)
   library (MASS)
   library(ggplot2)
   library(grid)
   library(gridExtra)
   library(ggpubr)
   library(extrafont)
   loadfonts(device="win")

## Import data files -----
   #Choose Frequency_Table_YZ_PartB_Plots.txt (change gene frequency to relative frequency
(%))
   con <- file.choose(new = FALSE)
   Frequency_Table <- read.table(con, header = T, row.names = 1)
   #Choose Frequency_Table_taxonomy_YZ.txt
   con1 <-file.choose(new = FALSE)
   Frequency_Table_taxonomy <- read.delim(con1, header = T, row.names = 1)


# Relative abundance of fungal communities during the cultivation at phylum level.
physeqa1 <-tax_glom(physeq, taxrank=rank_names(physeq)[2], NArm=TRUE,
bad_empty=c(NA, "", " ", "\t"))
   tablea1 <- otu_table(physeqa1)

   a1 = plot_bar(physeqa1, fill = "Phylum") +
      geom_bar(aes(color=Phylum, fill=Phylum), stat = "identity", position = "stack") +
      xlab("") + ylab("Relative Frequency (%)")    +
      theme(legend.position="right",
               axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust
= 1),
               axis.text.y = element_text(size = 11, family="Times New Roman"),
               axis.title.x = element_text(size = 12, family="Times New Roman"),
               axis.title.y = element_text(size = 12, family="Times New Roman"),
               legend.text = element_text(size = 11, family="Times New Roman"),
               legend.title= element_text(size = 12, family="Times New Roman"))
   a1

#qPCR figure

con <- file.choose(new = FALSE)
   metadata <- read.table(con, header = T, row.names = 1)
metadata$Cell <- factor(metadata$Cell )

   data_summary <- function(data, varname, groupnames){
```

```
    require(plyr)
    summary_func <- function(x, col){
      c(mean = mean(x[[col]], na.rm=TRUE),
        sd = sd(x[[col]], na.rm=TRUE))
    }
    data_sum<-ddply(data, groupnames, .fun=summary_func,
                    varname)
    data_sum <- rename(data_sum, c("mean" = varname))
    return(data_sum)
  }

data <- data_summary(metadata, varname="Cq_value", groupnames=("Cell"))

box_1 <- ggplot(data, aes(x=Cell, y=Cq_value, fill=Cell)) +
  geom_bar(stat="identity", position=position_dodge(0.9), width=0.5)+
  geom_errorbar(aes(ymin=Cq_value-sd,          ymax=Cq_value+sd),          width=0.2,
position=position_dodge(0.9))+
  xlab("")+
  ylab("Cq value") + ylim(0, 30)    + labs(title = "", subtitle=NULL) +
  theme(title=element_text(size=20, family="Times New Roman"),
        axis.text.x = element_text(size=20, family="Times New Roman"),
        axis.text.y=element_text(size=20, family="Times New Roman"),
        axis.title.y = element_text(size = 20, family="Times New Roman"),
        axis.title.x=element_text(size=20, family="Times New Roman"),
        legend.position="right",)+
  labs(fill="")
box_1

#comparison of two culture
library (MASS)
library(ggplot2)
library(grid)
library(gridExtra)
library(ggpubr)

# Installing the font package ---------------------------------------------
Sys.setenv(R_GSCMD="C:/Program Files/gs/gs9.05/bin/gswin32c.exe")
install.packages("extrafontdb")
library(extrafont)
font_import() #It may take a few minutes to import.
loadfonts(device="win")

# PROGRAM TO PLOT BAR CHART WITH STANDARD DEVIATION ----------------------

#++++++++++++++++++++++++++
# Function to calculate the mean and the standard deviation
```

```
# for each group
#++++++++++++++++++++++++++
# data : a data frame
# varname : the name of a column containing the variable
#to be summariezed
# groupnames : vector of column names to be used as
# grouping variables
data_summary <- function(data, varname, groupnames){
   require(plyr)
   summary_func <- function(x, col){
      c(mean = mean(x[[col]], na.rm=TRUE),
        sd = sd(x[[col]], na.rm=TRUE))
   }
   data_sum<-ddply(data, groupnames, .fun=summary_func,
                           varname)
   data_sum <- rename(data_sum, c("mean" = varname))
   return(data_sum)
}


# ANALYSIS----------------------------------------------------------------

## the .txt file needs to be saved as the type of "Tab delimited".

##load meta_data_RegularStat(30L).txt
con <-file.choose(new = FALSE)
metadata <- read.table(con, header = T, row.names = 1)

Biomass_concentration <- data_summary(metadata, varname="Biomass_concentration",
                                                  groupnames=c("Culture_type"))

Biomass_concentration$Culture_type=as.factor(Biomass_concentration$Culture_type)
Biomass_concentration
box_1 <- ggplot(Biomass_concentration, aes(x=Culture_type, y= Biomass_concentration,
fill=Culture_type)) +
   geom_bar(stat="identity", position=position_dodge(0.9), width=0.5)+
   geom_errorbar(aes(ymin= Biomass_concentration -sd, ymax= Biomass_concentration +sd),
width=0.2, position=position_dodge(0.9))+
   xlab("Culture type")+
   ylab("Biomass concentration (mg/L)") + ylim(0, 1.5)    + labs(title = "", subtitle=NULL) +
   theme(title=element_text(size=20, family="Times New Roman"),
            axis.text.x = element_text(size=20, family="Times New Roman"),
            axis.text.y=element_text(size=20, family="Times New Roman"),
            axis.title.y = element_text(size = 20, family="Times New Roman"),
            axis.title.x=element_text(size=20, family="Times New Roman"),
            legend.position="right",)+
```

```
   scale_x_discrete(labels = c('Formate','Mixed carbon'))+
   labs(fill="")
box_1

#TN comparison
TN <- data_summary(metadata, varname="TN",
                                                    groupnames=c("Culture_type"))

TN$Culture_type=as.factor(TN$Culture_type)
TN$Culture_time=as.factor(TN$Culture_time)
TN
box_2 <- ggplot(TN, aes(x=Culture_type, y=TN, fill=Culture_type)) +
   geom_bar(stat="identity", position=position_dodge(0.9), width=0.5)+
   geom_errorbar(aes(ymin=TN-sd,ymax=TN+sd),
width=0.2, position=position_dodge(0.9))+
   xlab("Culture type")+
   ylab("TN (mg/L)") + ylim(0, 110)    + labs(title = "", subtitle=NULL) +
   theme(title=element_text(size=20, family="Times New Roman"),
            axis.text.x = element_text(size=20, family="Times New Roman"),
            axis.text.y=element_text(size=20, family="Times New Roman"),
            axis.title.y = element_text(size = 20, family="Times New Roman"),
            axis.title.x=element_text(size=20, family="Times New Roman"),
            legend.position="right",)+
   scale_x_discrete(labels = c('Formate','Mixed carbon'))+
   labs(fill="")
box_2

#TP comparison
TP <- data_summary(metadata, varname="TP",
                                                    groupnames=c("Culture_type"))

TP$Culture_type=as.factor(Biomass_concentration$Culture_type)
TP$Culture_time=as.factor(Biomass_concentration$Culture_time)
TP
box_3 <- ggplot(TP, aes(x=Culture_type, y=TP, fill=Culture_type)) +
   geom_bar(stat="identity", position=position_dodge(0.9), width=0.5)+
   geom_errorbar(aes(ymin=TP-sd, ymax=TP+sd), width=0.2, position=position_dodge(0.9))+
   xlab("Culture type")+
   ylab("TP (mg/L)") + ylim(0, 40)    + labs(title = "", subtitle=NULL) +
   theme(title=element_text(size=20, family="Times New Roman"),
            axis.text.x = element_text(size=20, family="Times New Roman"),
            axis.text.y=element_text(size=20, family="Times New Roman"),
            axis.title.y = element_text(size = 20, family="Times New Roman"),
            axis.title.x=element_text(size=20, family="Times New Roman"),
            legend.position="right",)+
scale_x_discrete(labels = c('Formate','Mixed carbon'))+
```

```
   labs(fill="")
box_3
```

CHAPTER 4

```
library (MASS)
library(ggplot2)
library(grid)
library(gridExtra)
library(ggpubr)

# Installing the font package ------------------------------------------
Sys.setenv(R_GSCMD="C:/Program Files/gs/gs9.05/bin/gswin32c.exe")
install.packages("extrafontdb")
library(extrafont)
font_import() #It may take a few minutes to import.
loadfonts(device="win")

# PROGRAM TO PLOT BAR CHART WITH STANDARD DEVIATION ----------------------

#+++++++++++++++++++++++++
# Function to calculate the mean and the standard deviation
# for each group
#+++++++++++++++++++++++++
# data : a data frame
# varname : the name of a column containing the variable
#to be summariezed
# groupnames : vector of column names to be used as
# grouping variables
data_summary <- function(data, varname, groupnames){
   require(plyr)
   summary_func <- function(x, col){
      c(mean = mean(x[[col]], na.rm=TRUE),
         sd = sd(x[[col]], na.rm=TRUE))
   }
   data_sum<-ddply(data, groupnames, .fun=summary_func,
                           varname)
   data_sum <- rename(data_sum, c("mean" = varname))
   return(data_sum)
}


# ANALYSIS---------------------------------------------------------------

## the .txt file needs to be saved as the type of "Tab delimited".
```

```
##load meta_data_RegularStat(30L).txt
con <-file.choose(new = FALSE)
metadata <- read.table(con, header = T, row.names = 1)

#Biomass_concentration
Biomass_concentration <- data_summary(metadata, varname=" Biomass_concentration ",
                                                groupnames=c("Type","Formate"))
Biomass_concentration
box_1 <- ggplot(Biomass_concentration, aes(x=Formate, y=Biomass_concentration, fill=Type))
+
  geom_bar(stat="identity", position='dodge', width=0.5)+
  geom_errorbar(aes(ymin= Biomass_concentration -sd, ymax= Biomass_concentration +sd),
width=0.3, position=position_dodge(0.5))+
  xlab("Formate feeding rate")+
  ylab("Biomass_concentration (g/L)") + labs(title = "", subtitle=NULL) +
  theme(title=element_text(size=20, family="Times New Roman"),
        axis.text.x = element_text(size=20, family="Times New Roman"),
        axis.text.y=element_text(size=20, family="Times New Roman"),
        axis.title.y = element_text(size = 20, family="Times New Roman"),
        axis.title.x=element_text(size=20, family="Times New Roman"),
        legend.position="right",
        legend.text = element_text(size = 15))+
  scale_fill_discrete(labels = c('After harvesting','Before harvesting'))+
  labs(fill="")
box_1

#TN concentration
TN<- data_summary(metadata, varname="TN",
                                                groupnames=c("Type","Formate"))
TN
box_2<- ggplot(TN, aes(x=Formate, y=TN, fill=Type)) +
  geom_bar(stat="identity", position='dodge', width=0.5)+
  geom_errorbar(aes(ymin=TN-sd,ymax=TN+sd), width=0.3, position=position_dodge(0.5))+
  xlab("Formate feeding rate")+
  ylab("TN (mg/L)") + labs(title = "", subtitle=NULL) +
  theme(title=element_text(size=20, family="Times New Roman"),
        axis.text.x = element_text(size=20, family="Times New Roman"),
        axis.text.y=element_text(size=20, family="Times New Roman"),
        axis.title.y = element_text(size = 20, family="Times New Roman"),
        axis.title.x=element_text(size=20, family="Times New Roman"),
        legend.position="right",
        legend.text = element_text(size = 15))+
  scale_fill_discrete(labels = c('After harvesting','Before harvesting'))+
  labs(fill="")
box_2
```

```
#TP concentration
TP<- data_summary(metadata, varname="TP",
                                                groupnames=c("Type","Formate"))
TP
box_3 <- ggplot(TP, aes(x=Formate, y=TP, fill=Type)) +
   geom_bar(stat="identity", position='dodge', width=0.5)+
   geom_errorbar(aes(ymin=TP-sd,ymax=TP+sd), width=0.3, position=position_dodge(0.5))+
   xlab("Formate feeding rate")+
   ylab("TP (mg/L)") + labs(title = "", subtitle=NULL) +
   theme(title=element_text(size=20, family="Times New Roman"),
         axis.text.x = element_text(size=20, family="Times New Roman"),
         axis.text.y=element_text(size=20, family="Times New Roman"),
         axis.title.y = element_text(size = 20, family="Times New Roman"),
         axis.title.x=element_text(size=20, family="Times New Roman"),
         legend.position="right",
         legend.text = element_text(size = 15))+
   scale_fill_discrete(labels = c('After harvesting','Before harvesting'))+
   labs(fill="")
box_3


##Microbial community analysis##

library(vegan)
   library(phyloseq)
   library (MASS)
   library(ggplot2)
   library(grid)
   library(gridExtra)
   library(ggpubr)
   library(extrafont)
   loadfonts(device="win")

## Import data files -----
   #Choose Frequency_Table_YZ_PartB_Plots.txt (change gene frequency to relative frequency
(%))
   con <- file.choose(new = FALSE)
   Frequency_Table <- read.table(con, header = T, row.names = 1)
   #Choose Frequency_Table_taxonomy_YZ.txt
   con1 <-file.choose(new = FALSE)
   Frequency_Table_taxonomy <- read.delim(con1, header = T, row.names = 1)

## Phyloseq -----
   Full_Frequency <- cbind.data.frame(Frequency_Table, Frequency_Table_taxonomy)
   Frequency <- otu_table(Frequency_Table,taxa_are_rows = TRUE) #Frequency table
production for phyloseq
```

```
  TAX <- tax_table(as.matrix(Frequency_Table_taxonomy)) #Taxanomy production for phyloseq
  physeq <- phyloseq(Frequency, TAX) #physeq document production
  physeq0 <- tax_glom(physeq, taxrank=rank_names(physeq)[7], NArm=TRUE, bad_empty=c(NA, "", " ", "\t"))
  tax_table(physeq0)
  TAX

#Abundance Plotbar Domain
  physeqa <-tax_glom(physeq, taxrank=rank_names(physeq)[1], NArm=TRUE, bad_empty=c(NA, "", " ", "\t"))
  tablea <- otu_table(physeqa)
  tablea
  a = plot_bar(physeqa, fill = "Domain") +
    geom_bar(aes(color=Domain, fill=Domain), stat = "identity", position = "stack") +
    xlab("Culture Conditions") + ylab("Relative Frequency (%)") +
    theme(legend.position="right",
          axis.text.x = element_text(size = 11, family="Times New Roman", angle = 90, hjust = 1),
          axis.text.y = element_text(size = 11, family="Times New Roman"),
          axis.title.x = element_text(size = 12, family="Times New Roman"),
          axis.title.y = element_text(size = 12, family="Times New Roman"),
          legend.text = element_text(size = 11, family="Times New Roman"),
          legend.title= element_text(size = 12, family="Times New Roman"))+
    scale_x_discrete(labels=c("0.25g/L formate with 30% harvesting","0.5g/L formate with 30% harvesting","0.75g/L formate with 30% harvesting","Control with 30% harvesting","1g/L formate with 30% harvesting","1g/L formate with 50% harvesting" ))
  a


physeqa1 <-tax_glom(physeq, taxrank=rank_names(physeq)[2], NArm=TRUE, bad_empty=c(NA, "", " ", "\t"))
  tablea1 <- otu_table(physeqa1)
 physeq2 <-subset_taxa(physeq, Domain== "Bacteria")
  physeq2_1 <-tax_glom(physeq2, taxrank=rank_names(physeq2)[2], NArm=TRUE, bad_empty=c(NA, "", " ", "\t"))
  table2_1 <- otu_table(physeq2_1)
  table2_1
  c = plot_bar(physeq2_1, fill = "Phylum") +
    geom_bar(aes(color=Phylum, fill=Phylum), stat = "identity",position = "stack") +
    xlab("") + ylab("Relative Frequency (%)") +
    theme(legend.position="right",
          axis.text.x = element_text(size = 17, family="Times New Roman", angle = 90, hjust = 1),
          axis.text.y = element_text(size = 11, family="Times New Roman"),
          axis.title.x = element_text(size = 12, family="Times New Roman"),
```

```
            axis.title.y = element_text(size = 17, family="Times New Roman"),
            legend.text = element_text(size = 15, family="Times New Roman"),
            legend.title= element_text(size = 15, family="Times New Roman"))+
scale_x_discrete(labels=c("0.25g/L formate with 30% harvesting","0.5g/L formate with 30%
harvesting","0.75g/L formate with 30% harvesting","Control with 30% harvesting","1g/L formate
with 30% harvesting","1g/L formate with 50% harvesting" ))
c


physeq3 <-subset_taxa(physeq, Phylum == "Bacteroidetes")
  physeq3_1 <-tax_glom(physeq3, taxrank=rank_names(physeq3)[3], NArm=TRUE,
bad_empty=c(NA, "", " ", "\t"))
  table3_1 <- otu_table(physeq3_1)

  d = plot_bar(physeq3_1, fill = "Class")+ geom_bar(aes(color=Class, fill=Class), stat =
"identity",position = "stack") +
    xlab("") + ylab("Bacteroidetes Relative Frequency (%)") +
      theme(legend.position="right",
            axis.text.x = element_text(size = 17, family="Times New Roman", angle = 90, hjust
= 1),
            axis.text.y = element_text(size = 11, family="Times New Roman"),
            axis.title.x = element_text(size = 12, family="Times New Roman"),
            axis.title.y = element_text(size = 17, family="Times New Roman"),
            legend.text = element_text(size = 15, family="Times New Roman"),
            legend.title= element_text(size = 15, family="Times New Roman"))+
scale_x_discrete(labels=c("0.25g/L formate with 30% harvesting","0.5g/L formate with 30%
harvesting","0.75g/L formate with 30% harvesting","Control with 30% harvesting","1g/L formate
with 30% harvesting","1g/L formate with 50% harvesting" ))
  d


physeq5 <-subset_taxa(physeq, Phylum == "Proteobacteria")
  physeq5_1 <-tax_glom(physeq5, taxrank=rank_names(physeq5)[3], NArm=TRUE,
bad_empty=c(NA, "", " ", "\t"))
  table5_1 <- otu_table(physeq5_1)
  f = plot_bar(physeq5_1, fill = "Class")+ geom_bar(aes(color=Class, fill=Class), stat =
"identity",position = "stack") +
    xlab("") + ylab("Proteobacteria Relative Frequency (%)") +
    theme(legend.position="right",
            axis.text.x = element_text(size = 17, family="Times New Roman", angle = 90, hjust
= 1),
            axis.text.y = element_text(size = 11, family="Times New Roman"),
            axis.title.x = element_text(size = 12, family="Times New Roman"),
            axis.title.y = element_text(size = 17, family="Times New Roman"),
            legend.text = element_text(size = 15, family="Times New Roman"),
            legend.title= element_text(size = 15, family="Times New Roman"))+
```

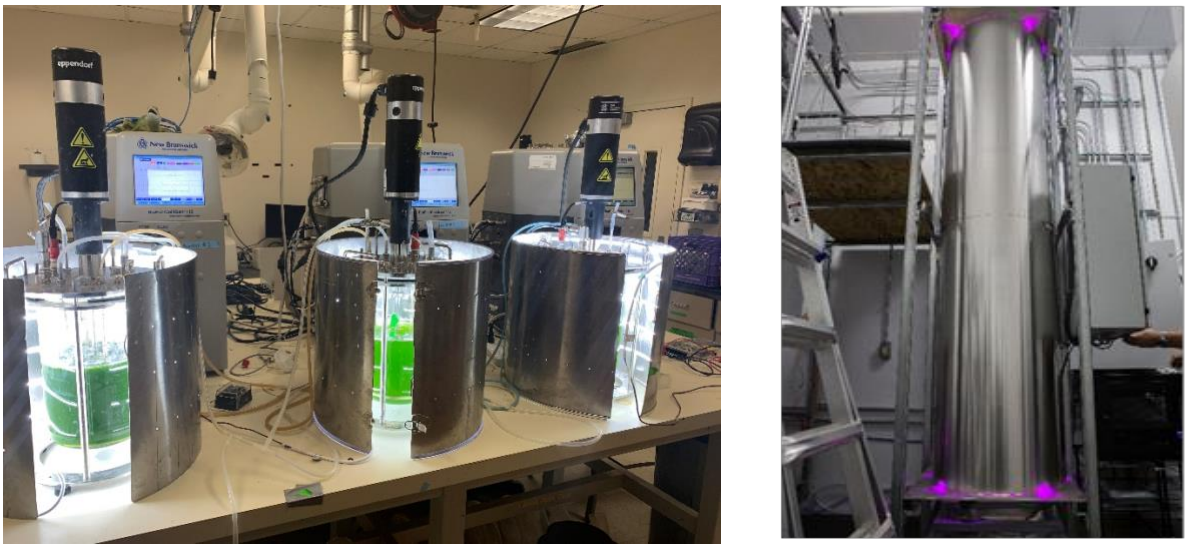
```
scale_x_discrete(labels=c("0.25g/L formate with 30% harvesting","0.5g/L formate with 30% harvesting","0.75g/L formate with 30% harvesting","Control with 30% harvesting","1g/L formate with 30% harvesting","1g/L formate with 50% harvesting" ))
f


##Statitical analysis##
library(vegan)
library(phyloseq)
library (MASS)
library(ggplot2)
library(grid)
library(gridExtra)
library(ggpubr)

library(extrafont)
font_import() #It may take a few minutes to import.
loadfonts(device="win")

## the .txt file needs to be saved as the type of "Tab delimited".

#Gene frequency data from QIIME2

## Choose data files -----

#Choose the Frequency_Table.txt
con <- file.choose(new = FALSE)
Frequency_Table <- read.table(con, header = T, row.names = 1)
#Choose the Frequency_Table_taxanomy.txt
con1 <-file.choose(new = FALSE)
Frequency_Table_taxonomy <- read.delim(con1, header = T, row.names = 1)

## Alpha Diversity -----
#Create a matrix object with the data frame
t.Frequency.table <- t(Frequency_Table)    # Transpose the data
class(t.Frequency.table) # Check the class of the table

#Alpha diversity analysis indexes
#Shannon
H <- diversity(t.Frequency.table, index = "shannon", MARGIN = 1, base = exp(1))
#Simpson
D <- diversity(t.Frequency.table, "simpson", MARGIN = 1, base = exp(1))
#Inverse Simpson
iD <- diversity(t.Frequency.table, "inv")
#Pielou's evenness
J<-H/log(specnumber(t.Frequency.table))
```

```
#List all indexes
IN <- cbind(H,D,iD,J)
IN
write.csv(IN, "diversity.csv")

#Plot H, D, iD, and J
plot(H)
plot(D)
plot(iD)
plot(J)

#Estimate Chao1 and ACE
estimateR(t.Frequency.table)

## ANOVA for Alpha Diversity -----

#Using the H, D, iD, and J data to generate "diversity.txt" to run ANOVA
#Choose diversity_30L_10_2019.txt

con2 <-file.choose(new = FALSE)
alphadiversity <- read.table(con2, header = T, row.names = 1)

#Define factor for alpha diversity
alphadiversity$Harvesting <- factor(alphadiversity$Harvesting)
alphadiversity$Formate_conc <- factor(alphadiversity$Formate_conc)
# Normality
shapiro.test(alphadiversity$H)
shapiro.test(alphadiversity$D)
shapiro.test(alphadiversity$iD)
shapiro.test(alphadiversity$J)

Hfit <- aov(H ~ Harvesting, data = alphadiversity)
summary(Hfit)

#ANOVA of J index
Jfit <- aov(J ~ Harvesting, data = alphadiversity)
summary(Jfit)

Hfit <- aov(H ~ Formate_conc, data = alphadiversity)
summary(Hfit)

#ANOVA of J index
Jfit <- aov(J ~ Formate_conc, data = alphadiversity)
summary(Jfit)

## Rarefaction -----
```

```
col <- c("black", "darkred", "forestgreen", "orange", "blue", "yellow", "hotpink")
lty <- c("solid", "dashed", "longdash", "dotdash")
pars <- expand.grid(col = col, lty = lty, stringsAsFactors = FALSE)
head(pars)
ra <- rarecurve(t.Frequency.table, step = 20, col =col,lty = lty, cex = 0.6) # curve of rarefication
rad <- rad.lognormal(t.Frequency.table) # Rank of Abundance
rad1 <- plot(rad, xlab = "Rank", ylab = "Abundance") # Plotting the rank


## ANOVA for Beta Diversity ##
con3 <-file.choose(new = FALSE)
metadata <- read.table(con3, header = T, row.names = 1, fill = TRUE)

# Define factors for metadata -----
metadata$Light_intensity <- factor(metadata$Light_intensity)

#Permutational analysis of variance
t.Frequency.table <- t(Frequency_Table)    #transpose the data
class(t.Frequency.table) #check the class of the table
View(Frequency_Table)
View(t.Frequency.table)

betad <-betadiver(t.Frequency.table, 'z')
betad

#adonis
adonis(betad~Formate_conc, metadata, perm=200)
```

APPENDIX B: SUPPLEMENTAL TABLES AND FIGURES

CHAPTER 2



a b

Figure S2.1 Lab-scale and pilot-scale photobioreactors

a. Lab-scale algae photobioreactors (two of them were used for this study); b. Pilot-scale algae photobioreactor

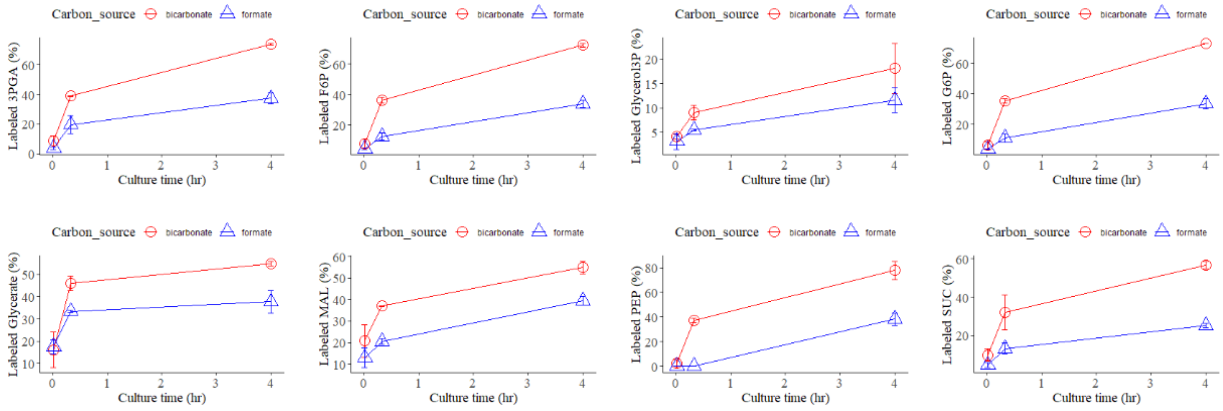
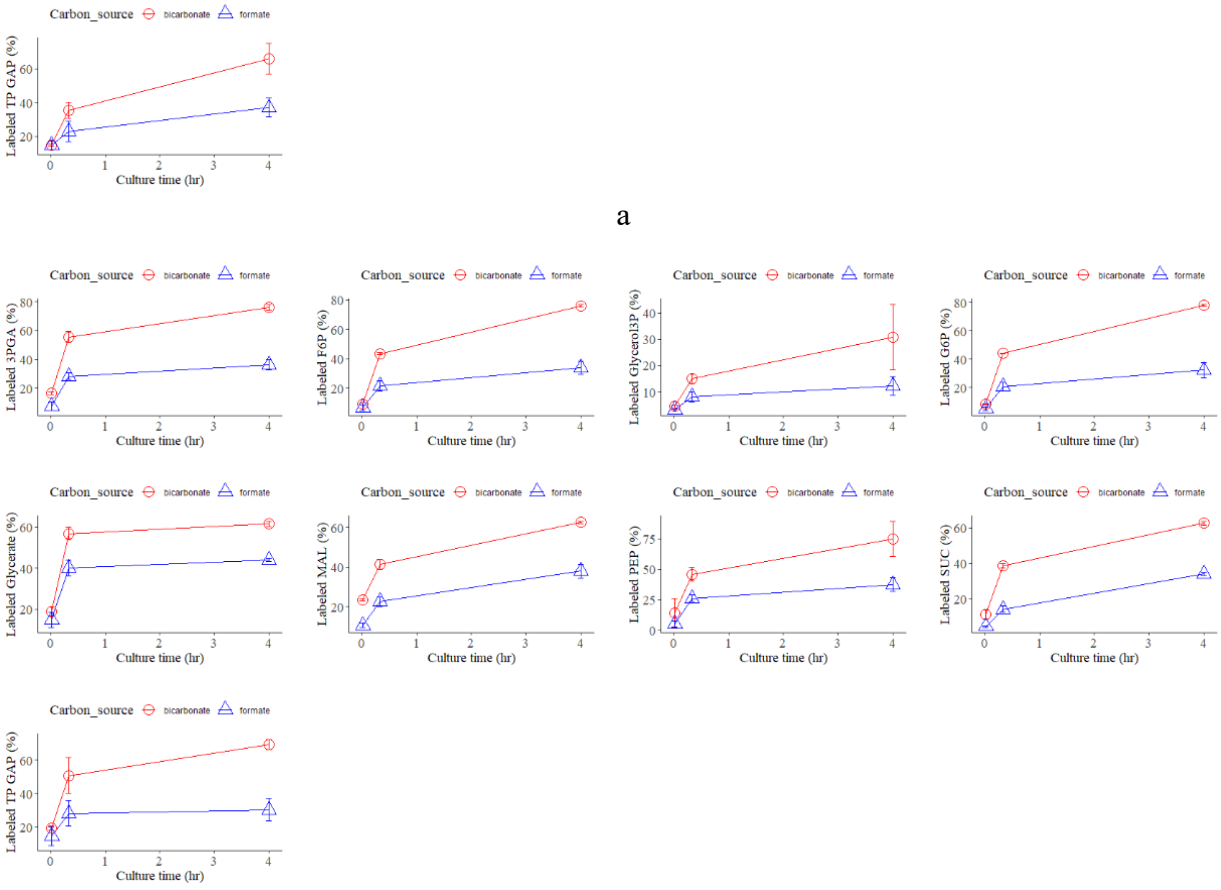


Figure S2.2 Labeled free metabolites from the cultures on formate and bicarbonate

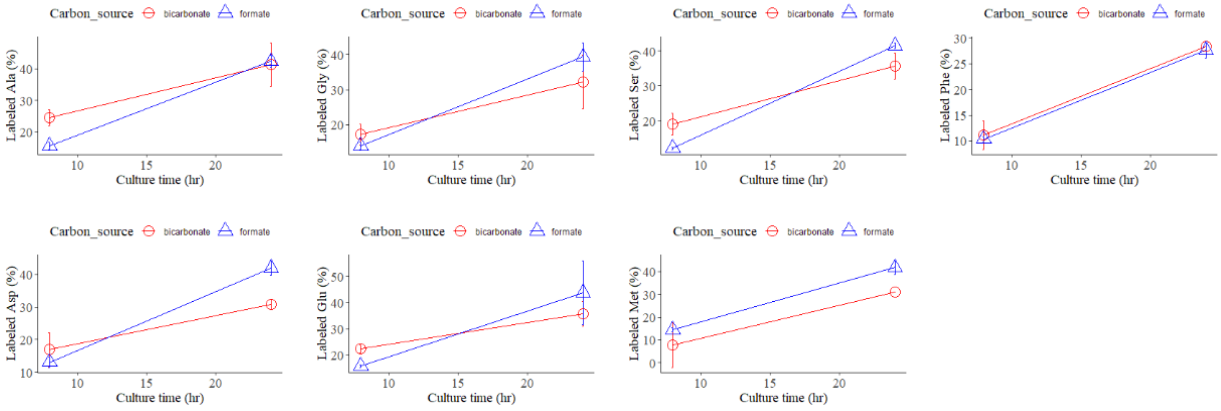
a. Labeled free metabolites under 500 µmol/m²/s; b. Free metabolites under 50 µmol/m²/s.

Figure S2.2 (cont'd)



a

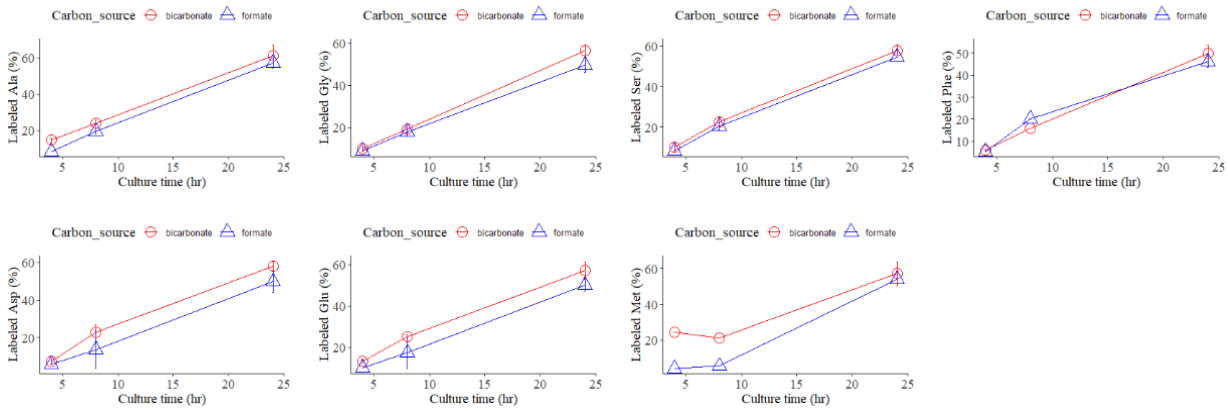
b



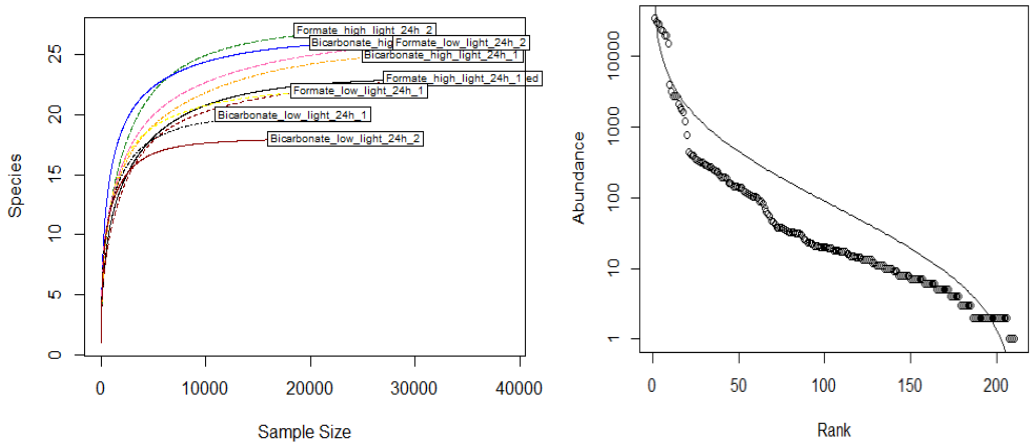
a

Figure S2.3 Labeled proteinogenic amino acids from the cultures on formate and bicarbonate
a. Labeled proteinogenic amino acids under 500 µmol/m^2/s; b. Labeled proteinogenic amino acids under 50 µmol/m^2/s

Figure S2.3 (cont'd)



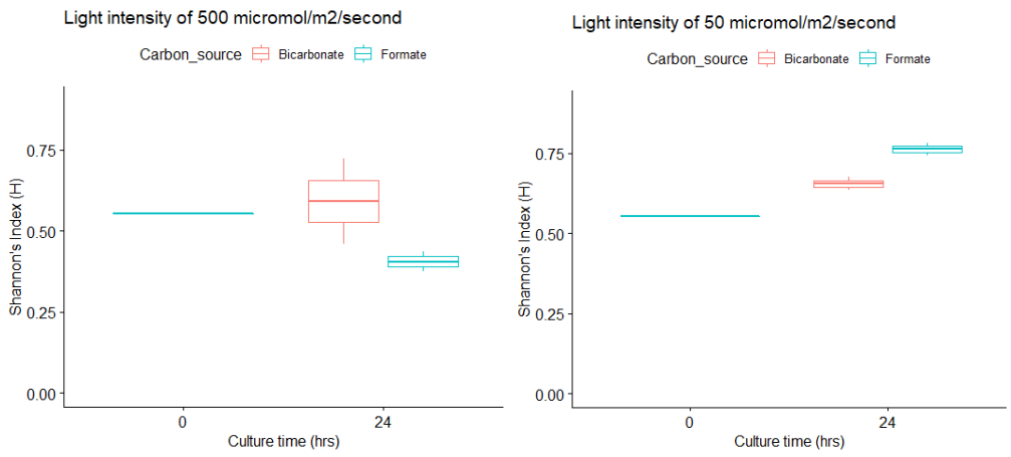
b



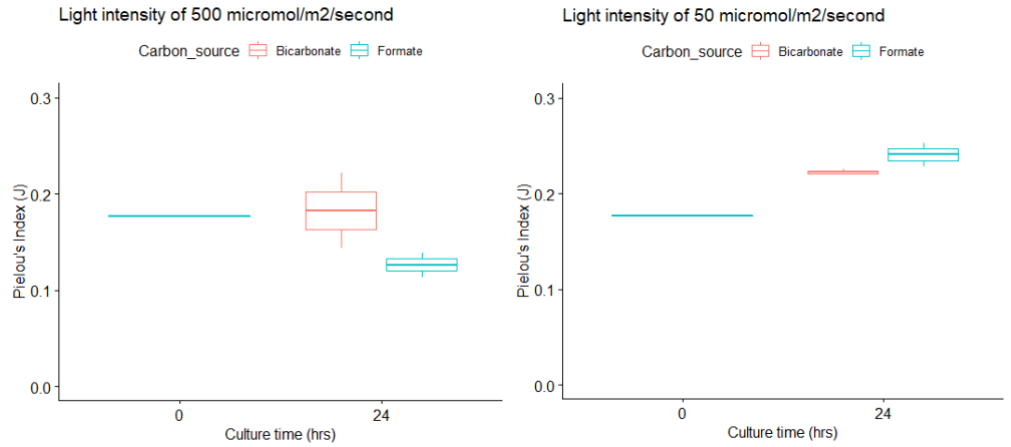
a b

Figure S2.4 Rarefaction and rank abundance of the batch culture

a. Rarefaction curves for gene sequences from all samples; b. Rank abundance



a



b

Figure S2.5 Ecological diversity indexes of the batch cultures

a. Shannon (H) diversity on two light intensities; b. Pielou's (J) evenness on two light intensities

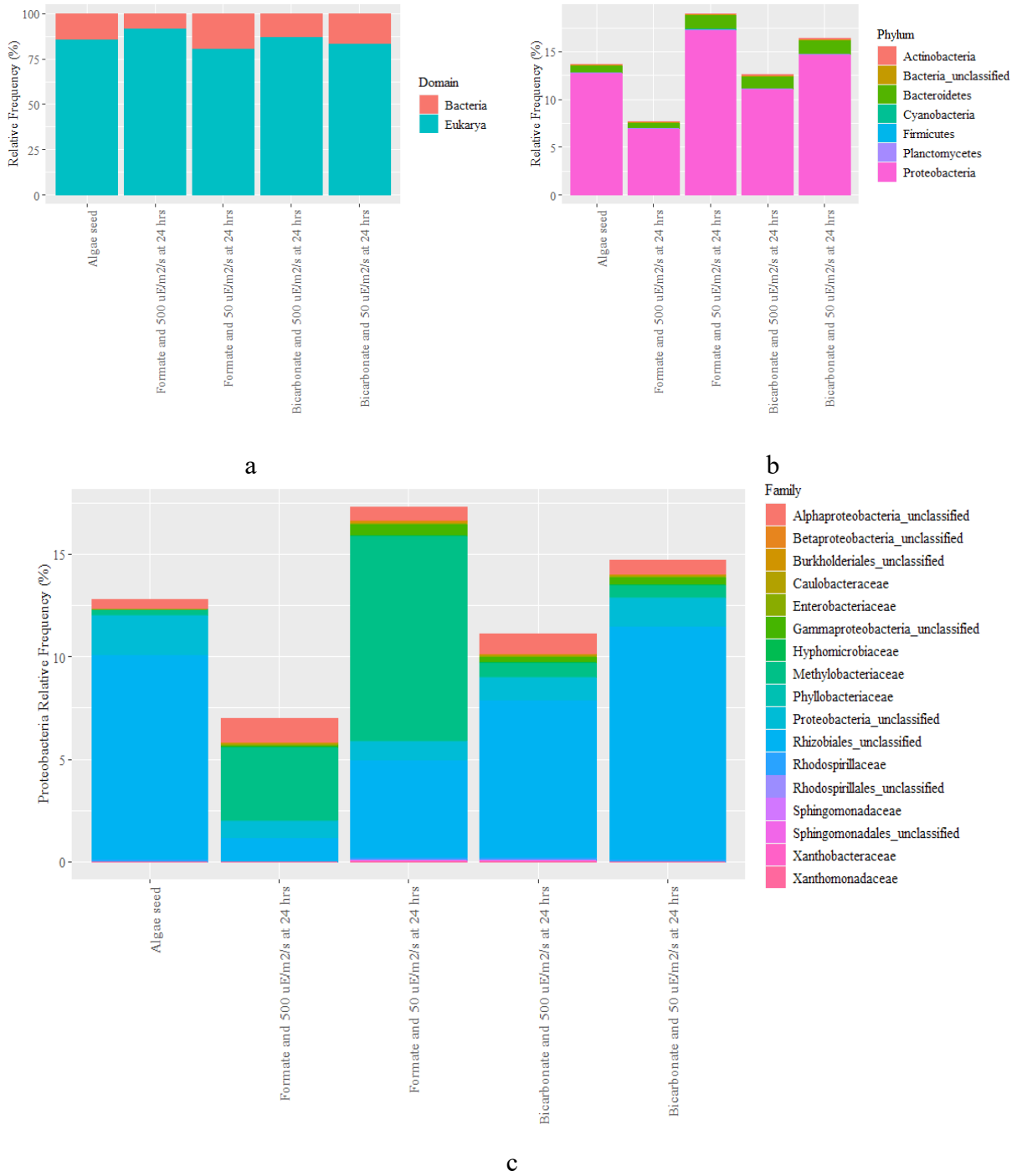
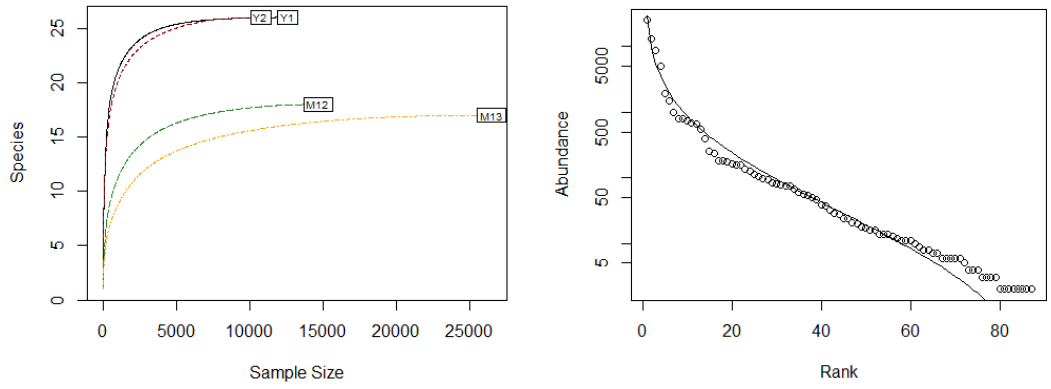


Figure S2.6 Relative abundance of microbial communities from the batch cultures on formate and bicarbonate

a. Microbial domain; b. Bacterial phylum; c. Proteobacteria family



a b

Figure S2.7 Rarefaction and rank abundance of the semi-continuous culture*

a. Rarefaction curves for gene sequences from all samples; b. Rank abundance

**: Y1 and Y2 are two samples from the semi-continuous culture on formate. M12 and M13 are two samples from the semi-continuous culture on CO_2*

Table S2.1 Two-way ANOVA of light intensity, carbon source, and culture time on alpha diversity and evenness of the batch culture

Parameter		Light intensity	Carbon source	Culture time	Light intensity: Carbon source	Light intensity: culture time	Residuals
Pielou's index (J)	Degree of freedom	1	1	1	1	1	4
	Sum of squares	0.009566	0.001079	0.00060	0.00794	0.004379	0.003728
	F value	10.265	1.157	0.065	0.852	4.699	
	P (>F)	0.03*	0.34	0.81	0.41	0.09	
Shannon's index (H)	Degree of freedom	1	1	1	1	1	4
	Sum of squares	0.07121	0.00579	0.00119	0.01855	0.04288	0.03758
	F value	7.579	0.616	0.126	1.974	4.563	
	P (>F)	0.05*	0.48	0.74	0.23	0.10	

** means a significant factor*

Table S2.2 Permutation one-way ANOVA of light intensity, carbon source, and culture time on beta diversity of the batch culture

	Degree of freedom	Sum of squares	F value	P (>F)
Light intensity	2	0.04149	1.1798	0.32
Carbon source	1	0.006346	0.3159	0.86
Culture time	1	0.02393	1.3616	0.33

**: permutation is free. The number of permutations is 200*

Table S2.3 Microbial genus identified in the batch cultures

	Domain	Phylum	Class	Order	Family	Genus
1	Bacteria	Bacteria_unclassified	Bacteria_unclassified	Bacteria_unclassified	Bacteria unclassified	Bacteria unclassified
2	Bacteria	Actinobacteria	Actinobacteria	Actinomycetales	Actinomycetales unclassified	Actinomycetales unclassified
3	Bacteria	Bacteroidetes	Bacteroidetes unclassified	Bacteroidetes unclassified	Bacteroidetes unclassified	Bacteroidetes unclassified
4	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidales unclassified	Bacteroidales unclassified
5	Bacteria	Bacteroidetes	Cytophagia	Cytophagales	Cytophagales unclassified	Cytophagales unclassified
6	Bacteria	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Flavobacteriaceae unclassified
7	Bacteria	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Sphingobacteriaceae	Sphingobacteriaceae unclassified
8	Bacteria	Bacteroidetes	[Saprospirae]	[Saprospirales]	Chitinophagaceae	Chitinophagaceae unclassified
9	Eukarya	Chlorophyta	Trebouxiophyceae	Chlorellales	Chlorellaceae	Chlorella
10	Bacteria	Cyanobacteria	Chloroplast	Chlorophyta	Chlorophyta unclassified	Chlorophyta unclassified
11	Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Clostridium
12	Bacteria	Planctomycetes	Planctomycetia	Planctomycetia unclassified	Planctomycetia unclassified	Planctomycetia unclassified

Table S2.3 (cont'd)

13	Bacteria	Proteobacteria	Proteobacteria unclassified	Proteobacteria unclassified	Proteobacteria unclassified	Proteobacteria unclassified
14	Bacteria	Proteobacteria	Alphaproteobacteria	Alphaproteobacteria unclassified	Alphaproteobacteria unclassified	Alphaproteobacteria unclassified
15	Bacteria	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Caulobacteraceae unclassified
16	Bacteria	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Brevundimonas
17	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiales unclassified	Rhizobiales unclassified
18	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Hyphomicrobiaceae unclassified
19	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methylobacteriaceae	Methylobacteriaceae unclassified
20	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	Phyllobacteriaceae unclassified
21	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	Xanthobacter
22	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Rhodobacteraceae unclassified
23	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillales unclassified	Rhodospirillales unclassified
24	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	Acetobacteraceae unclassified
25	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Rhodospirillaceae unclassified
26	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadales unclassified	Sphingomonadales unclassified
27	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonadaceae unclassified
28	Bacteria	Proteobacteria	Betaproteobacteria	Betaproteobacteria unclassified	Betaproteobacteria unclassified	Betaproteobacteria unclassified
29	Bacteria	Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiales unclassified	Burkholderiales unclassified
30	Bacteria	Proteobacteria	Gammaproteobacteria	Gammaproteobacteria unclassified	Gammaproteobacteria unclassified	Gammaproteobacteria unclassified

Table S2.3 (cont'd)

31	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Enterobacteriaceae unclassified
32	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Xanthomonadaceae unclassified

Table S2.4 Comparison of proteins highly expressed in the formate culture with them from the bicarbonate culture under 500 µmol/m^2/s at 24 hours of the culture

Protein names	Gene names	Organism	Mean LFQ in the formate culture (LOG2)	Mean LFQ in the bicarbonate (LOG2)	LFQ difference between two cultures
60S ribosomal L14	C2E21_1106	Chlorella sorokiniana (Freshwater green alga)	30.27957667	27.4020725	-2.8775
Peptidylprolyl isomerase (EC 5.2.1.8)	C2E21_7932	Chlorella sorokiniana (Freshwater green alga)	27.70004333	24.90500667	-2.79504
F-box SKIP16	C2E21_1762	Chlorella sorokiniana (Freshwater green alga)	28.10517333	25.57441	-2.53076
Histone H2A	C2E21_4012 C2E21_4031 C2E21_4058 C2E21_4068 C2E21_4106 C2E21_4168 C2E21_8032 C2E21_8050 C2E21_8098 C2E21_9088 C2E21_9091	Chlorella sorokiniana (Freshwater green alga)	32.52807333	30.1565425	-2.37153

Table S2.4(cont'd)

PsbP chloroplastic	C2E21_6264	Chlorella sorokiniana (Freshwater green alga)	27.45091	25.16561667	-2.28529
40S ribosomal S10	C2E21_8389	Chlorella sorokiniana (Freshwater green alga)	29.70546667	27.43614	-2.26933
Regulator of chromosome condensation RCC1	C2E21_1268	Chlorella sorokiniana (Freshwater green alga)	29.98240333	27.7643825	-2.21802
Ferredoxin component	C2E21_8942	Chlorella sorokiniana (Freshwater green alga)	28.42739	26.232795	-2.1946
Photosystem I reaction center subunit VI-chloroplastic-like	C2E21_5226	Chlorella sorokiniana (Freshwater green alga)	27.38434667	25.2378975	-2.14645
40S ribosomal S28	C2E21_0253	Chlorella sorokiniana (Freshwater green alga)	30.16748333	28.07339	-2.09409
Histone H4	C2E21_9090	Chlorella sorokiniana (Freshwater green alga)	33.15795	31.1459875	-2.01196
Ribosomal S1	C2E21_5682	Chlorella sorokiniana (Freshwater green alga)	29.18863667	27.1908175	-1.99782
50S ribosomal chloroplastic-like	C2E21_7040	Chlorella sorokiniana (Freshwater green alga)	26.54203	24.6448175	-1.89721

Table S2.4(cont'd)

Calmodulin	C2E21_8040	Chlorella sorokiniana (Freshwater green alga)	27.88988667	25.99539	-1.8945
Rhodanese-like domain-containing chloroplastic	C2E21_7293	Chlorella sorokiniana (Freshwater green alga)	28.0447425	26.1752425	-1.8695
Glycine-rich 2	C2E21_8624	Chlorella sorokiniana (Freshwater green alga)	29.19301333	27.4376525	-1.75536
Chlorophyll a-b binding protein, chloroplastic	C2E21_5761	Chlorella sorokiniana (Freshwater green alga)	29.8619	28.1143475	-1.74755
Chloroplast ATP synthase subunit delta	C2E21_5501	Chlorella sorokiniana (Freshwater green alga)	27.61971667	25.9584125	-1.6613
40S ribosomal S13	C2E21_4948	Chlorella sorokiniana (Freshwater green alga)	27.26081667	25.6809	-1.57992
Glutaredoxin-dependent peroxiredoxin (EC 1.11.1.25)	C2E21_7738	Chlorella sorokiniana (Freshwater green alga)	27.70572	26.1317375	-1.57398
Ribose-5-phosphate isomerase (EC 5.3.1.6)	C2E21_4453	Chlorella sorokiniana (Freshwater green alga)	28.04598333	26.49711667	-1.54887
Protein disulfide-isomerase (EC 5.3.4.1)	C2E21_0663	Chlorella sorokiniana (Freshwater green alga)	27.80271667	26.25984667	-1.54287

Table S2.4(cont'd)

Peptidylprolyl isomerase (EC 5.2.1.8)	C2E21_7308	Chlorella sorokiniana (Freshwater green alga)	26.956345	25.4440925	-1.51225
40S ribosomal S19-1	C2E21_2453	Chlorella sorokiniana (Freshwater green alga)	28.67901	27.1727425	-1.50627
R3H domain-containing protein	EOP07_17760	Proteobacteria bacterium	30.120355	28.62316	-1.4972
Mitochondrial outer membrane porin 4	C2E21_3515	Chlorella sorokiniana (Freshwater green alga)	29.03696	27.5605125	-1.47645
ATP synthase subunit mitochondrial	C2E21_4028	Chlorella sorokiniana (Freshwater green alga)	28.65461	27.2249425	-1.42967
Acyl carrier protein	C2E21_5060	Chlorella sorokiniana (Freshwater green alga)	31.53778	30.11347	-1.42431
Serine arginine-rich splicing factor RSZ22	C2E21_1354	Chlorella sorokiniana (Freshwater green alga)	26.83751	25.47503	-1.36248
Programmed cell death 4-like isoform B	C2E21_3642	Chlorella sorokiniana (Freshwater green alga)	25.78204667	24.4982375	-1.28381
Histone acetyltransferase (EC 2.3.1.48)	C2E21_1396	Chlorella sorokiniana (Freshwater green alga)	31.77522	30.5259325	-1.24929

Table S2.4(cont'd)

Clathrin heavy chain	C2E21_0246	Chlorella sorokiniana (Freshwater green alga)	27.37363667	26.1334	-1.24024
Plastid-lipid-associated chloroplastic-like	C2E21_1628	Chlorella sorokiniana (Freshwater green alga)	27.47864	26.241065	-1.23758
Low molecular mass early light-induced	C2E21_8756	Chlorella sorokiniana (Freshwater green alga)	27.67826333	26.4938775	-1.18439
RETICULATA-RELATED chloroplastic-like isoform B	C2E21_0197	Chlorella sorokiniana (Freshwater green alga)	26.85447	25.6925975	-1.16187
1,4-alpha-glucan branching enzyme (EC 2.4.1.18)	C2E21_5287	Chlorella sorokiniana (Freshwater green alga)	25.947215	24.8417125	-1.1055
Thylakoid lumenal chloroplastic	C2E21_1644	Chlorella sorokiniana (Freshwater green alga)	28.19691	27.1039	-1.09301
Photosystem I reaction center subunit IV	C2E21_0167	Chlorella sorokiniana (Freshwater green alga)	29.84331	28.7665025	-1.07681
V-type proton ATPase subunit G	C2E21_6393	Chlorella sorokiniana (Freshwater green alga)	26.67865333	25.66186	-1.01679
Chlorophyll a-b binding protein, chloroplastic	C2E21_5185	Chlorella sorokiniana (Freshwater green alga)	30.27658	29.2809725	-0.99561

Table S2.4(cont'd)

Cytochrome c peroxidase	C2E21_9152	Chlorella sorokiniana (Freshwater green alga)	27.20883333	26.26526	-0.94357
ATP synthase subunit beta (EC 7.1.2.2) (ATP synthase F1 sector subunit beta) (F-ATPase subunit beta)	atpD atpB C7271_03120	filamentous cyanobacterium CCP5	27.42362667	26.53659667	-0.88703
Putative membrane-associated 30 kDa chloroplastic	C2E21_2641	Chlorella sorokiniana (Freshwater green alga)	29.17042667	28.30892	-0.86151
Translational inhibitor	C2E21_6778	Chlorella sorokiniana (Freshwater green alga)	28.34465667	27.5046275	-0.84003
Chlorophyll a-b binding protein, chloroplastic	C2E21_5760	Chlorella sorokiniana (Freshwater green alga)	28.77956333	27.9522025	-0.82736
40S ribosomal S15	C2E21_7240	Chlorella sorokiniana (Freshwater green alga)	28.57213667	27.77318	-0.79896
40S ribosomal S18	C2E21_4711	Chlorella sorokiniana (Freshwater green alga)	30.02448667	29.38918	-0.63531

Table S2.5 Comparison of proteins highly expressed in the bicarbonate culture with them from the formate culture under 500 µmol/m²/s at 24 hours of the culture

Protein names	Gene names	Organism	Mean LFQ in the formate culture (LOG2)	Mean LFQ in the bicarbonate (LOG2)	LFQ difference between two cultures
Glutamate 5-kinase	C2E21_4687	Chlorella sorokiniana (Freshwater green alga)	25.19615	25.71774333	0.521593
Nitrate reductase	C2E21_6288	Chlorella sorokiniana (Freshwater green alga)	27.96938	28.590695	0.621315
Phosphoserine aminotransferase (EC 2.6.1.52)	C2E21_2225	Chlorella sorokiniana (Freshwater green alga)	25.630055	26.25395667	0.623902
Glutathione peroxidase	C2E21_7473	Chlorella sorokiniana (Freshwater green alga)	28.62905333	29.2925675	0.663514
Chlorophyll a-b binding protein, chloroplastic	C2E21_2201	Chlorella sorokiniana (Freshwater green alga)	28.51437333	29.2495675	0.735194
Calcium sensing chloroplastic	C2E21_4514	Chlorella sorokiniana (Freshwater green alga)	28.46617	29.2272475	0.761077
Pre-mRNA-processing factor 19 (EC 2.3.2.27)	C2E21_9058	Chlorella sorokiniana (Freshwater green alga)	24.82261333	25.61738667	0.794773

Table S2.5 (cont'd)

Nadh:ubiquinone oxidoreductase complex i intermediate-associated 30	C2E21_5050	Chlorella sorokiniana (Freshwater green alga)	26.47398	27.2748525	0.800873
Nucleolar 56	C2E21_8588	Chlorella sorokiniana (Freshwater green alga)	25.47142	26.3001175	0.828697
Protein disulfide-isomerase (EC 5.3.4.1)	C2E21_2729	Chlorella sorokiniana (Freshwater green alga)	27.034535	27.8983675	0.863833
Plastoquinol--plastocyanin reductase (EC 7.1.1.6)	C2E21_5551	Chlorella sorokiniana (Freshwater green alga)	27.9237	28.84526	0.92156
N-acetyl-glutamate semialdehyde dehydrogenase (EC 1.2.1.38)	C2E21_1983	Chlorella sorokiniana (Freshwater green alga)	26.97181	27.8945325	0.922722
Endopeptidase Clp (EC 3.4.21.92)	C2E21_2049	Chlorella sorokiniana (Freshwater green alga)	26.63490333	27.560135	0.925232
Malate dehydrogenase (NADP(+)) (EC 1.1.1.82)	C2E21_1149	Chlorella sorokiniana (Freshwater green alga)	26.66042667	27.610645	0.950218
Serine hydroxymethyltransferase (EC 2.1.2.1)	C2E21_3819	Chlorella sorokiniana (Freshwater green alga)	26.117595	27.109615	0.99202
Leucyl-tRNA synthetase (EC 6.1.1.4)	C2E21_2873	Chlorella sorokiniana (Freshwater green alga)	27.38565	28.45414	1.06849

Table S2.5 (cont'd)

Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.-)	C2E21_5374	Chlorella sorokiniana (Freshwater green alga)	27.10652	28.2155425	1.109023
Rab family GTPase	C2E21_2008	Chlorella sorokiniana (Freshwater green alga)	26.37358667	27.545615	1.172028
Pentose-5-phosphate 3-epimerase (EC 5.1.3.1)	C2E21_4587	Chlorella sorokiniana (Freshwater green alga)	28.53828667	29.7658625	1.227576
Aminopeptidase family	C2E21_3120	Chlorella sorokiniana (Freshwater green alga)	26.631345	27.863495	1.23215
D-fructose-1,6-bisphosphate 1-phosphohydrolase (EC 3.1.3.11)	C2E21_6233	Chlorella sorokiniana (Freshwater green alga)	27.18380667	28.5151975	1.331391
Aminomethyltransferase, mitochondrial (EC 2.1.2.10) (Glycine cleavage system T protein)	C2E21_7296	Chlorella sorokiniana (Freshwater green alga)	25.56968667	26.961485	1.391798
40S ribosomal protein S6	C2E21_6316	Chlorella sorokiniana (Freshwater green alga)	26.65690333	28.1390725	1.482169
TIC chloroplastic	C2E21_1189	Chlorella sorokiniana (Freshwater green alga)	27.661165	29.1569	1.495735

Table S2.5 (cont'd)

60S ribosomal L10-3	C2E21_8800	Chlorella sorokiniana (Freshwater green alga)	26.61655	28.1533175	1.536768
Adenosylhomocysteinase (EC 3.3.1.1)	C2E21_8126	Chlorella sorokiniana (Freshwater green alga)	27.4533	28.990405	1.537105
Puromycin-sensitive aminopeptidase isoform X1 isoform A	C2E21_5257	Chlorella sorokiniana (Freshwater green alga)	24.43628	26.023	1.58672
40S ribosomal protein S8	C2E21_0182	Chlorella sorokiniana (Freshwater green alga)	26.26368667	27.895575	1.631888
Signal peptide peptidase-like 3	C2E21_6391	Chlorella sorokiniana (Freshwater green alga)	27.14985333	28.78248	1.632627
Glyceraldehyde-3-phosphate dehydrogenase	C2E21_9115	Chlorella sorokiniana (Freshwater green alga)	30.479695	32.1380525	1.658358
Chlorophyll a-b binding protein, chloroplastic	C2E21_4122	Chlorella sorokiniana (Freshwater green alga)	27.76483	29.4531675	1.688338
Ferredoxin-NADP+ reductase	C2E21_5690	Chlorella sorokiniana (Freshwater green alga)	29.281515	31.0780925	1.796578
Phosphoenolpyruvate carboxylase (EC 4.1.1.31)	C2E21_0194	Chlorella sorokiniana (Freshwater green alga)	26.38904	28.1905775	1.801537

Table S2.5 (cont'd)

Chlorophyll a-b binding protein, chloroplastic	C2E21_5089	Chlorella sorokiniana (Freshwater green alga)	29.63762333	31.45713	1.819507
ADP,ATP carrier	C2E21_7012	Chlorella sorokiniana (Freshwater green alga)	27.9943425	29.820525	1.826183
Chlorophyll a-b binding protein, chloroplastic	C2E21_0580	Chlorella sorokiniana (Freshwater green alga)	27.88650667	29.7427375	1.856231
Sedoheptulose-1,7-chloroplastic	C2E21_5298	Chlorella sorokiniana (Freshwater green alga)	28.332205	30.1903675	1.858163
Glyceraldehyde-3-phosphate cytosolic	C2E21_7692	Chlorella sorokiniana (Freshwater green alga)	28.4321975	30.42999	1.997793
Pyruvate, phosphate dikinase (EC 2.7.9.1)	C2E21_8120	Chlorella sorokiniana (Freshwater green alga)	28.1339975	30.2270825	2.093085
Serine hydroxymethyltransferase (EC 2.1.2.1)	C2E21_0021	Chlorella sorokiniana (Freshwater green alga)	26.4116775	28.63954	2.227863
Beta-Ig-H3 fasciclin	C2E21_4613	Chlorella sorokiniana (Freshwater green alga)	24.122535	26.5344525	2.411918
Chlorophyll a-b binding protein, chloroplastic	C2E21_0372	Chlorella sorokiniana (Freshwater green alga)	27.365735	29.8838675	2.518133

Table S2.5 (cont'd)

PSI subunit V	C2E21_9213	Chlorella sorokiniana (Freshwater green alga)	25.344805	28.247905	2.9031

Table S2.6 Three-way ANOVA of carbon source on alpha diversity and evenness of the semi-continuous cultures

Parameter		Carbon source	Residuals
Pielou's index	Degree of freedom	1	2
	Sum of squares	0.1102	0.0094
	F value	23.46	0.005
	P (>F)	0.04*	
Shannon's index	Degree of freedom	1	2
	Sum of squares	1.2859	0.0971
	F value	26.47	
	P (>F)	0.36*	

** means a significant factor*

Table S2.7 Permutation one-way ANOVA of carbon source on beta diversity of the semi-continuous cultures*

	Degree of freedom	Sum of squares	F value	P (>F)
Carbon source	1	0.1572	10.15	0.33
Residuals	2	0.0310		

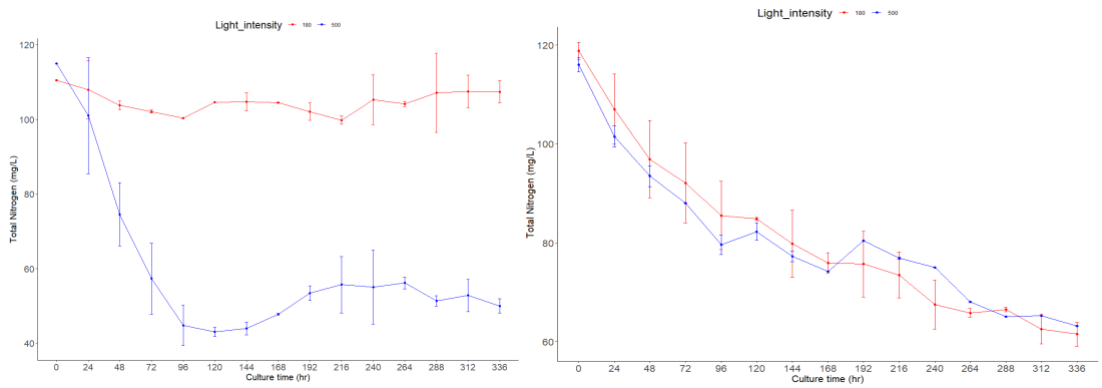
**: permutation is free. The number of permutations is 23*

Table S2.8 Microbial genus identified in the semi-continuous cultures

	Domain	Phylum	Class	Order	Family	Genus
1	Bacteria	Bacteria unclassified	Bacteria unclassified	Bacteria unclassified	Bacteria unclassified	Bacteria unclassified
2	Bacteria	Actinobacteria	Actinobacteria	Actinomycetales	Actinomycetales unclassified	Actinomycetales unclassified
3	Bacteria	Bacteroidetes	Bacteroidetes unclassified	Bacteroidetes unclassified	Bacteroidetes unclassified	Bacteroidetes unclassified
4	Bacteria	Bacteroidetes	Cytophagia	Cytophagales	Cytophagales unclassified	Cytophagales unclassified
5	Bacteria	Bacteroidetes	Cytophagia	Cytophagales	Cyclobacteriaceae	Cyclobacteriaceae unclassified
6	Bacteria	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriales unclassified	Flavobacteriales unclassified
7	Bacteria	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Flavobacteriaceae unclassified
8	Bacteria	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Sphingobacteriaceae	Sphingobacteriaceae unclassified
9	Bacteria	Bacteroidetes	[Saprospirae]	[Saprospirales]	[Saprospirales] unclassified	[Saprospirales] unclassified
10	Bacteria	Bacteroidetes	[Saprospirae]	[Saprospirales]	Chitinophagaceae	Chitinophagaceae unclassified
11	Eukarya	Chlorophyta	Trebouxiophyceae	Chlorellales	Chlorellaceae	Chlorella
12	Bacteria	Firmicutes	Bacilli	Bacilli unclassified	Bacilli unclassified	Bacilli unclassified
13	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillales unclassified	Bacillales unclassified
14	Bacteria	Proteobacteria	Proteobacteria unclassified	Proteobacteria unclassified	Proteobacteria unclassified	Proteobacteria unclassified
15	Bacteria	Proteobacteria	Alphaproteobacteria	Alphaproteobacteria unclassified	Alphaproteobacteria unclassified	Alphaproteobacteria unclassified
16	Bacteria	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Brevundimonas

Table S2.8 (cont'd)

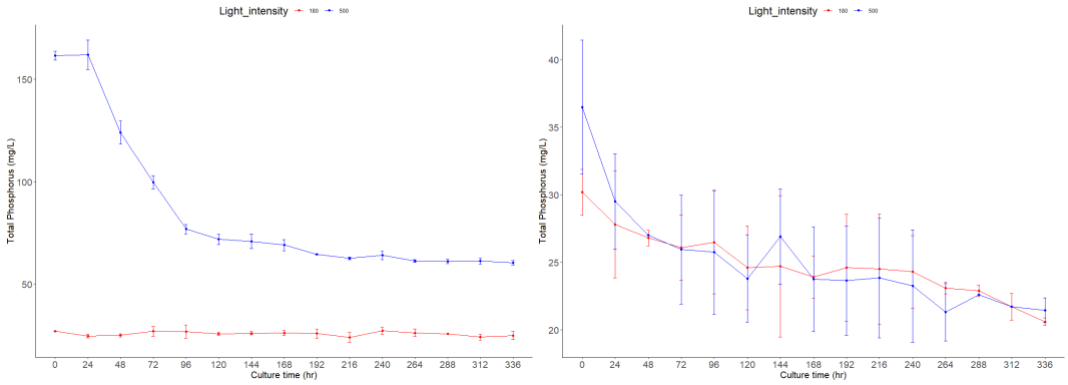
17	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiales unclassified	Rhizobiales unclassified
18	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobiaceae unclassified
19	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Hyphomicrobiaceae unclassified
20	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methylobacteriaceae	Methylobacteriaceae unclassified
21	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	Phyllobacteriaceae unclassified
22	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	Xanthobacter
23	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Rhodobacteraceae unclassified
24	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillales unclassified	Rhodospirillales unclassified
25	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	Roseomonas
26	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Rhodospirillaceae unclassified
27	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadales unclassified	Sphingomonadales unclassified
28	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Erythrobacteraceae	Erythrobacteraceae unclassified
29	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonadaceae unclassified
30	Bacteria	Proteobacteria	Betaproteobacteria	Betaproteobacteria unclassified	Betaproteobacteria unclassified	Betaproteobacteria unclassified
31	Bacteria	Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiales unclassified	Burkholderiales unclassified
32	Bacteria	Proteobacteria	Gammaproteobacteria	Gammaproteobacteria unclassified	Gammaproteobacteria unclassified	Gammaproteobacteria unclassified
33	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Xanthomonadaceae unclassified



a b

Figure S3.1 Time course of TN concentration under different light intensities

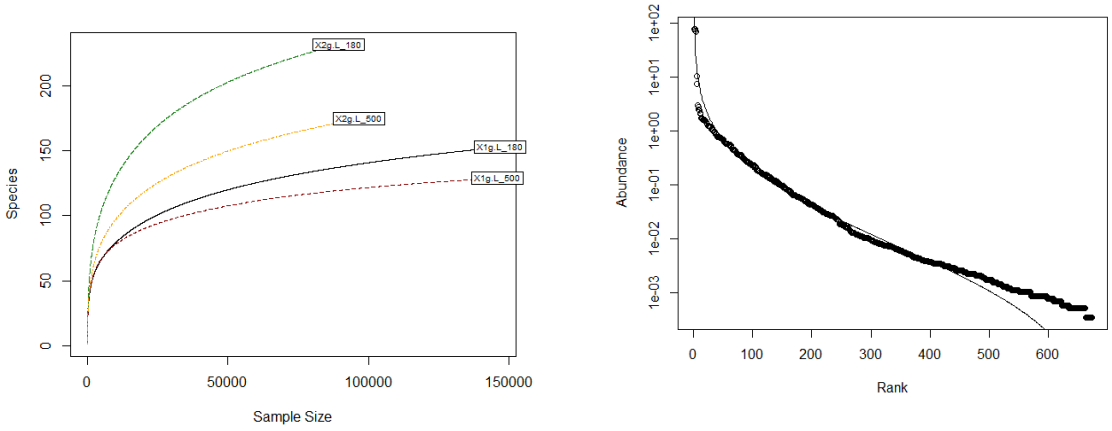
a. TN concentration with 1g/L/day formate feeding rate; b. TN concentration with 2g/L/day formate feeding rate



a b

Figure S3.2 Time course of TP concentration under different light intensities

a. TP concentration with 1g/L/day formate feeding rate; b. TP concentration with 2g/L/day formate feeding rate.

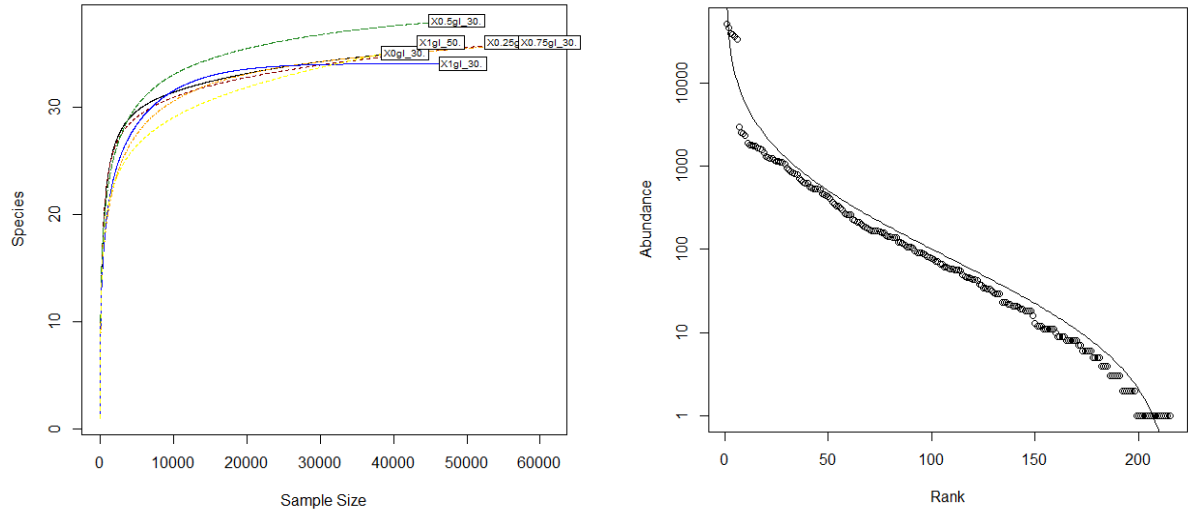


a b

Figure S3.3 Rarefaction and rank abundance of the long-term culture*

a. Rarefaction curves for gene sequences from all samples; b. Rank abundance

CHAPTER 4



a b

Figure S4.1 Rarefaction and rank abundance of the long-term culture*

a. Rarefaction curves for gene sequences from all samples; b. Rank abundance