
Chemical &
Biological
Engineering.

# Improving microalgal biotechnology by applying principles from ecological theory

David A. Russo

Supervisors:
Dr. Jagroop Pandhal
Dr. Andrew Beckerman

A thesis presented to The University of Sheffield in fulfilment
of the thesis requirement for the degree of
Doctor of Philosophy in Chemical and Biological Engineering

May 2016

## Table of contents

## I. Declaration

I, David Russo, declare that I am the sole author of this thesis and that the research presented within is the result of my own efforts and achievements, unless acknowledged otherwise in the text. I confirm that this work has not been submitted for any other degrees.

## II. List of figures

## III. List of tables

## IV. List of abbreviations

| | |
|---|---|
| ABC | ATP-binding cassette |
| ACN | Acetonitrile |
| ANOVA | Analysis of variance |
| C | Carbon |
| CCM | Carbon concentrating mechanism |
| CFU | Colony forming units |
| CSI | Captive-spray ionisation |
| DHA | Docosahexaenoic acid |
| DO | Dissolved oxygen |
| DOM | Dissolved organic matter |
| EDTA | Ethylenediaminetetraacetic acid |
| EmPAI | Exponentially modified protein abundance index |
| EPA | Eicosapentaenoic acid |
| FDR | False discovery rate |
| GH | Glycoside hydrolase |
| GMO | Genetically modified organism |
| GRAS | Generally regarded as safe |
| HMW | High molecular weight |
| HPLC | High performance liquid chromatography |
| HRP | High rate pond |
| K | Carrying capacity |
| LC | Liquid chromatography |
| LMVC | Low- to medium-value compounds |
| MS | Mass spectrometer |
| N | Nitrogen |
| P | Phosphorus |
| PAGE | Polyacrylamide gel electrophoresis |
| PBR | Photobioreactor |
| PS | Photosystem |
| PUFA | Polyunsaturated fatty acid |
| RSM | Response surface methodology |
| RuBisCO | Ribulose 1,5-biphosphate carboxylase oxygenase |

| | |
|---|---|
| SDS | Sodium dodecyl sulphate |
| SEM | Structural equation modelling |
| TAG | Triglyceride |
| TAP | Tris-acetate-phosphate |
| TBDR | TonB-dependent receptors |
| TEAB | Tetraethylammonium bromide |
| TFA | Trifluoroacetic acid |
| ToF | Time of flight |
| UniProt | Universal Protein Resource |
| UHPLC | Ultra-high performance liquid chromatography |
| WT | Wild type |

## V. Acknowledgements

This thesis is the outcome of years of my hard work and the direct and indirect influences of people to whom I am grateful to.

To my PhD supervisor Jagroop Pandhal for giving me the opportunity to undertake this project, his endless support, hours of discussion and for being present the whole way through. I extend this gratitude to my second supervisor Andrew Beckerman for many of the same reasons but also for stubbornly pushing me in the right direction and forcing me to leave my comfort zone and improve as a researcher.

To Narciso Couto for being an ever-present mentor. You taught me so much about proteomics and our long conversations made me feel at home.

To everyone in the D72 office, D73/76 labs and the Microalgal Biotechnology Sheffield Network. You are the people that were always present both inside and outside of the department and will always be remembered as the best part of my years in Sheffield. The friendships I made will never be forgotten.

To the Pandhal research group. Although you arrived later in my journey it was great to have shared ideas with you. A special thank you to Umar Jibril. Your joy and enthusiasm was contagious.

To all the CBE technical and administrative staff, especially Dave, Louise, James, Kasia and Mark. Thank you for always being available and helping me in my research.

To my family for being a constant presence and an endless source of support. And to Julie Zedler, for the countless hours of discussion, patience, support and for making me a better researcher.

## VI. Summary

Microalgae are unicellular organisms that can be grown photoautotrophically and their abundance in natural valuable compounds makes their industrial cultivation attractive. Current technology only allows for cost-effective production of high-value compounds. Therefore, this thesis proposes the use of ecological theory and practice to improve the large-scale cultivation of low- to medium-value compounds in microalgae.

In the first study a multivariate modelling approach determined the individual importance of several abiotic factors on the dynamics of a microcosm microbial community under oligotrophic and eutrophic conditions. The application of a simple model illustrated key causal relationships and demonstrated that nutrient enrichment significantly changed the relative importance of the tested abiotic variables to the dynamics of the microbial system. The second study utilised a metaproteomic approach to detail the mechanisms of co-existence and acclimation in the same microbial community. A decrease in microalgal exudation, in eutrophic conditions, affected bacterial acquisition of energy and nutrients. Furthermore, two microalgal-bacterial relationships, of potential use to synthetic ecology, were highlighted. Finally, in the third study, the competitive dynamics between two *C. reinhardtii* strains, a wild type and a high-lipid mutant, were studied utilising response surface methodology. In the co-culture with 25% wild type, intraspecific competition significantly increased triglyceride concentrations. The competition data also suggested there was little risk of the mutant displacing the wild type under any of the experimental treatments. Finally, the highest triglyceride productivity was found in the pure mutant culture, after just 24 hours, demonstrating potential to scale out a small batch biomanufacturing system.

This thesis successfully coupled traditional ecology experiments with modern ‘omics techniques. Several existing hypotheses, regarding microalgal ecophysiology, were assessed based on their potential application in commercial microalgal cultivation. In sum, microalgal biotechnology can benefit from the integration of core principles of microalgal ecophysiology in the transition from laboratory to commercial-scale cultivation.

# VII. Introduction

The concept of microorganisms as biomanufacturing platforms has rapidly expanded in the last decade. The industry of synthetic biology is expected to reach $5.6 million by 2018 (Marketsandmarkets.com, 2014). Currently, yeast, bacteria and mammalian cells are the most popular host organisms, however, microalgae are rapidly becoming an alternative target due to their potential sustainability advantages. The ability to grow photoautotrophically and their abundance in valuable natural compounds makes their industrial cultivation attractive. Another important aspect is that several microalgae have a generally regarded as safe (GRAS) status which is essential for products targeted at human and animal consumption. Present technology allows for the cost-effective production of high value commodities, such as cosmetics and fine chemicals. However technologies for the production of low- to medium-value commodities (LMVCs), such as food and feed additives and biofuels, have not yet reached maturity (Benemann, 2013; Draaisma et al., 2013). Nonetheless, the objective of producing LMVCs in microalgae has sparked an enormous amount of interest. Chisti et al. (2007) noted that if certain microalgae can have up to 70% oil, double every day and have high productivities then microalgae can produce 136 900 L $ha^{-1}$. This study also concluded that "Microalgae appear to be the only source of biodiesel that has the potential to completely displace fossil diesel. Unlike other oil crops, microalgae grow extremely rapidly and many are exceedingly rich in oil" (Chisti, 2007). Although the potential yields stated in this reference are controversial (i.e. it breaks the principle of mass conservation) it served the purpose of bringing microalgae into the biotechnology spotlight. To date, this paper has 2520 citations in the Web of Science Core Collection (Fig. VII.I) with each year surpassing the previous in cumulative number of citations. This is indicative of the increasing interest in LMVC production in microalgae.

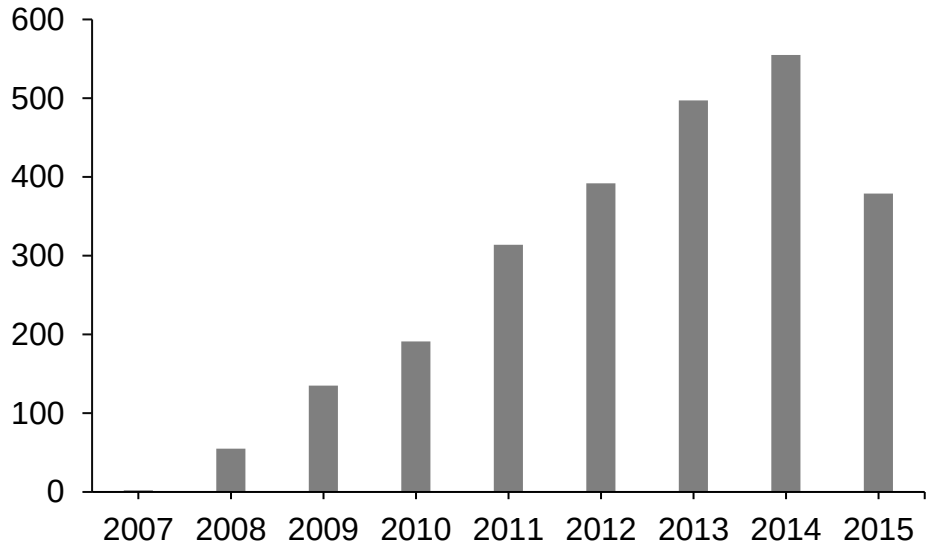


Figure VII.I. – Number of citations, per year, of Chisti (2007) "Biodiesel from microalgae".

There are many issues that hamper the cost effective production of LMVCs in microalgae. Mass cultures are currently grown outdoors, with a uni-algal population (monoculture) and are prone to contamination and sudden crashes. Furthermore, downstream processing (harvesting, cell lysis and product purification) costs are high. Due to the tight profit margins of LMVCs the traditional approach is not economically viable therefore cultivation and harvesting practices have to be improved. Several approaches to achieve economically feasible production of microalgae have been investigated over the last years. One approach that has been suggested is to harness the extensive ecological knowledge of microbial systems and apply this to microalgal biotechnology (Smith et al., 2015). Therefore, utilising ecological principles to study natural communities, which thrive under similar characteristics as those used in industrial cultivation, can provide valuable insight. One such microbial community can be found in aquatic ecosystems. Aquatic microbial communities have a unique characteristic which, traditionally, is considered a problem but has enormous potential for mass microalgal cultivation. This characteristic is the capacity to bloom under nutrient enriched conditions. Microalgal bloom communities are normally uni-microalgal, highly competitive and resistant to grazing and disease. These are precisely the community characteristics that can be harnessed for microalgal cultivation.

There is a variety of ecological tools that have been used to study microbial systems amongst which modelling is one of the most widespread. The ability to describe and translate complex biological systems through the use of mathematical

models has for long been one of the greatest strengths of ecology. Models are useful tools to describe the present function of complex systems and create hypotheses regarding functionality in the future. Amongst the many standard modelling techniques one stands out for its simplicity in describing a complex system while, simultaneously, considering all multivariate relationships in said system. Structural equation modelling (SEM) is a technique where theoretical understanding is mapped into a causal graph, a series of paths that creates a network of direct and indirect effects. The creation of this causal graph and the fitting of experimental data generates quantitative predictions for the effects of the experimental treatments. However, one limitation of this technique is that a statistical description of a relationship fails to reveal information on the underlying biological mechanisms. To achieve this, modern molecular techniques have been developed allowing an unprecedented view of microbial diversity and dynamics. The emergence of metagenomics allowed for the identification of microbial genetic diversity and, with metatranscriptomics, to identify the regulation and expression of said diversity. However, it was the application of proteomics to microbial communities, termed metaproteomics, which started to provide a more accurate insight into community function. Proteins are the final product of transcription and translation thus providing insight into the outcome of the regulation and expression of genetic diversity. Therefore, metaproteomics allows us to identify the functional expression of the metagenome and elucidate the activity of a community at the moment of sampling (Wilmes et al., 2006).

The information gathered from both modelling and molecular tools can provide an unparalleled view of natural microbial interactions and contribute to the long standing objective of community ecology to understand the assembly, structure and function of communities (Tilman, 1982; Chesson, 2000). Furthermore, this overlaps with the objective of creating a stable community, such as the ones found in natural aquatic systems, and applying such a community in microalgal biotechnology. The field that crosses community ecology and biotechnology has been named synthetic ecology and one of its recent focuses has been the design of microbial communities for the production of LMVCs. The objective is to design a microbial community, around a microalgal strain of interest, in order to establish mutualistic relationships which increase productivity beyond that obtained from the sum of the individual microbes. In addition to naturally occurring communities, this can also be done by introducing genetically modified microbial strains which are able to resist disease, outcompete contaminants or establish symbiotic relationships with the microalgal strain of interest.

This network of interactions should, ultimately, stabilise the community and provide resistance from disease and external contamination (Fig. VII.2) (Kazamia et al., 2012).

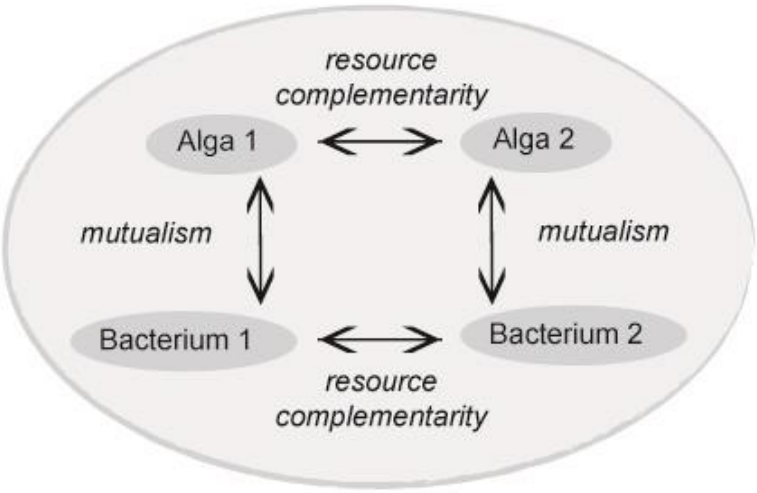


Figure VII.II. Hypothetical synthetic microbial community where the microalgal and bacterial species co-exist through resource complementarity and mutualistic interactions.

The widespread application of synthetic communities, both in open and closed microalgal production systems, also has disadvantages. It is inevitable that these cultivation systems will spread the cultivated organisms through spills, wind dispersion or improper waste disposal. This has led to worries regarding the unpredictable effects of the release of non-native, including genetically modified (GM), organisms into the environment (Gressel et al., 2013). There is beginning to be a general awareness of the potential risk of microalgal mass cultivation (Snow et al., 2012; Flynn et al., 2013), however, more needs to be done to develop a scientific risk analysis. Therefore, it is necessary to undertake studies that, generally, compare both native and non-native communities and, specifically, transgenic organisms with their respective wild type. This will be fundamental to protect both natural systems from non-native microorganisms and artificial systems from their natural counterparts.

In conclusion, the application of ecological theory is fundamental to the development of low-cost technologies for mass microalgal cultivation. Ultimately, culture productivity will depend on the ecophysiology of specific producer strains therefore, understanding the ecological principles that govern microalgae in microbial systems will improve their development, implementation and management in industrial scenarios.

## VIII. Aims and chapter outline

This dissertation will utilise multivariate modelling, metaproteomics and competition experiments to investigate the effect of biotic and abiotic factors on microbial community dynamics and microalgal production of LMVCs.

Chapter 1 reviews the existing literature on microalgal industrial applications, illustrates how ecological theory has been applied to microalgal biotechnology and, finally, how metaproteomics has provided insight into microbial communities.

Chapter 2 combines a traditional growth experiment with multivariate analysis (SEM) to uncover the role of nutrients (nitrogen and phosphorus) and physicochemical variables (pH, temperature and dissolved oxygen) on bacterial dynamics and associated microalgal communities in oligo- and eutrophic conditions.

Chapter 3 applies a metaproteomic approach to explore the identity and function of the microbial community previously explored through multivariate analysis (Chapter 2). Samples were representative of an early stage in the time series, during microalgal and cyanobacterial dominance, and a later stage during bacterial dominance. Both timepoints were analysed in oligo- and eutrophic conditions in order to uncover changes over time and nutrient levels.

Chapter 4 investigates the effects of intraspecific competition, between a *Chlamydomonas reinhardtii* mutant and its wild type, on growth and lipid productivity across different nitrogen concentrations and initial wild type to mutant cell ratios.

Finally, Chapter 5 discusses the results of the experimental chapters and contextualises them in regard to potential applications for LMVC production in microalgae.

# Chapter 1: Literature review

This literature review will give a broad overview of the current and potential future of utilising the aquatic ecology principles in the development of LMVC production in microalgae. Within this objective potential applications and the current situation of cultivation systems and markets for microalgal products will be discussed. The focus is then set on describing the development of synthetic ecology and how it can be applied to microalgal cultivation. This section will summarise the importance of aquatic microbial processes for microalgal biotechnology and emphasise metaproteomics as a discovery tool in aquatic microbial systems. This will demonstrate how findings in natural environments are relevant to microalgal biotechnology and tie into current mass cultivation practices.

## 1.1. Definition of microalgae

Generally, algae are considered to be aquatic, oxygen-evolving photosynthetic autotrophs that are unicellular, colonial or are constructed of filaments or composed of simple tissues. They are a polyphyletic group of organisms (i.e. classified into the same group but derived from different ancestors) with representatives in four kingdoms: Bacteria, Plantae, Chromista and Protozoa (Guiry, 2012). In biotechnology, the term microalga typically refers *sensu strictu* to microscopic algae and cyanobacteria thus, for the extent of this thesis, unless stated otherwise, the term microalgae will be used in its biotechnological sense.

## 1.2. Microalgal metabolism and biochemical characteristics

The possible metabolic modes for microalgae are photoautotrophy and organoheterotrophy. Photoautotrophic microalgae obtain energy from light for the reduction of $CO_2$, utilising water as an electron donor, with consequent release of $O_2$. They only require inorganic nutrients and obligate photoautotrophs cannot grow in the dark. Heterotrophic microalgae organisms generate their energy by degradation and uptake of organic compounds. Photoheterotrophic organism require light as an energy source to use organic compounds as nutrients, while organoheterotrophs utilise organic compounds to satisfy both their energetic and mineral needs (Grobbelaar, 2013). The traditional classification into phototrophs and heterotrophs cannot accurately describe

microalgal diversity therefore the term mixotrophy has been proposed. Here, mixotrophy will be used to describe a metabolic strategy in which at least three modes (photo-, organo-, auto-, and heterotrophy) are combined. An example of mixotrophy can be found in the model organism *Chlamydomonas* sp. This microalgal genus can successfully grow in a photoautotrophic regime, however, the presence of an organic carbon source (e.g. acetate) will increase growth rates (Lalibertè et al., 1993).

### 1.2.1. Nutrient requirements

Microalgal growth depends on the interaction between biotic (e.g. trophic regime, presence of predators) and abiotic factors, such as nutrient concentrations (Andersen, 2005). In natural systems there is a large debate regarding the importance of nitrogen (N) versus phosphorus (P) for microalgal growth (Elser et al., 2007; Schindler et al., 2008). N is mainly supplied in the form of nitrate ($NO_3^-$) and ammonium ($NH_4^+$). The N form selected will depend on the desired species. Certain cyanobacterial species are capable of fixing atmospheric nitrogen ($N_2$) and reducing it to $NH_4^+$ utilising the enzyme nitrogenase (O'Neil et al., 2012). For example, Ortiz-Marquez et al. (2012), co-cultured the microalga *Chlorella sorokiniana* strain RP with an *Azotobacter vinelandii* mutant strain that accumulates several times more ammonium in culture medium than wild-type (WT) cells. This was done in a growth medium with no N addition. The results showed that the microalgal growth in co-culture matched that of the microalgal growth in a medium with addition of 0.5 mM ammonium chloride. P is a macronutrient essential for all living cells and exists as phosphate ($PO_4^{3-}$) is its bioavailable form. It is a widely abundant element incorporated in important bioorganic molecules such as nucleic acids, phospholipids and proteins. Absolute values of N and P are important parameters for microalgal growth, however, our knowledge of nutrient physiology and ecological stoichiometry indicate that the C:N:P ratio can heavily influence productivity and species dominance. The typical ideal nutrient ratio is $C_{106}$:$N_{16}$:$P_1$ and is commonly referred to as the Redfield Ratio (Redfield, 1958). This ratio should be the starting point for any growth medium design. It is important to note that this ratio is an average cellular stoichiometry. Large fluctuations in these numbers have been found when comparing different species, habitats and geographical location. Sterner et al. (2008) measured the chemistry of particulate matter in over 2000 water bodies and found a constant proportionality of $C_{166}$:$N_{20}$:$P_1$ across the entire dataset which implies higher

carbon (C) sequestration per unit of N and P. In small freshwater lakes, a natural system similar to outdoor cultivation systems, natural microalgal communities exhibited a ratio of $C_{224}$:$N_{22}$:$P_1$ indicating an even higher C sequestration for a similar input of N and P (Sterner et al., 2008).

### 1.2.1. Biochemical composition

Microalgae have a diverse chemical composition that changes in response to environmental conditions. Under favourable conditions, a microalga can have 30 – 50% protein, 20 – 40% carbohydrate and 8 – 15% lipids on a dry weight basis. However, under environmental stress, they can accumulate up to 65% of lipids or carbohydrates. This characteristic is an example of what makes microalgae an appealing feedstock for LMVC production such as biofuels (Benemann, 2013). One of the main focuses of biofuel production in microalgae is on techniques to improve lipid accumulation. Many microalgal species have a naturally high lipid content (30 – 50% dry weight), however, it is possible to increase it further by manipulating growth conditions. Environmental stress activates the neutral lipid biosynthetic pathway and the accumulation of neutral lipids, especially triacylglycerols (TAGs). However, an increase in lipid accumulation is not synonymous with an increase in lipid yield. There is a metabolic trade-off between biomass productivity and lipid accumulation that limits the total yield of lipid per biomass unit in microalgae (Hu et al., 2008). Currently, N stress is the most common procedure in lipid accumulation research (Adams et al., 2013). However, other ways of lipid accumulation by P stress in *Nannochloropsis* sp. (Rodolfi et al., 2009), light stress in *Scenedesmus obliquus* (Ho et al., 2012), temperature stress in *Nannochloropsis oculata* (Converti et al., 2009) and salt stress in *Dunaliella tertiolecta* (Takagi et al., 2006) have also been successful.

## 1.3. Historical perspective

The use of microalgae dates back 2000 years to the Chinese, who used *Nostoc* to survive famine. They were also used by native tribes of Africa and South America to supplement their diet (Spolaore et al., 2006). The first description of a microalga occurred under the form of a communication to the Royal Society, in 1703, with a

representation of the freshwater diatom *Tabellaria*. However, it wouldn't be until 1850 that the first attempts of lab culture were reported. There was a slow progress in early times but this saw some growth during the 20$^{th}$ century. Due to meat shortages the idea of industrial applications of microalgae first appeared in Germany as an inexpensive source of protein. In the 50s and 60s various researchers started to research microalgal culture optimisation. These decades also saw the birth of commercial scale production with *Chlorella* culture in Japan (Nihon Chlorella) and *Arthrospira* culture in Mexico (Sosa Texcoco S.A.) (Spolaore et al., 2006). It wouldn't be until the 70s that great progress was achieved with microalgal exploitation extending to pigments, food supplements and vitamins for the pharmaceutical industry. In that decade, marked by two oil crises in 1973 and 1979, the Department of Energy (USA), in collaboration with the Solar Energy Research Institute (now the National Renewable Energy Laboratory), started a program dedicated to optimising open pond cultivation of microalgae for biofuels. The programme became known as the Aquatic Species Program. One of its main objectives was to study the biochemistry of lipid production in oleaginous microalgae. It was concluded that the production of biofuels in microalgae was technically feasible, however, considerable research and development was needed to bring this to fruition. This program would come to be abandoned in 1996, due to the lack of funding, without achieving all its proposed goals (Sheehan et al., 1998). In the last ten years, there has been an exponential increase in microalgal research at universities and companies all around the world. In 2003, there were less than 100 companies worldwide involved in the continuous production of microalgae, but in 2015 there are more than a 1000. However, only 20% of the companies existing in 2003 still exist. These numbers are simultaneously a testament of the current success rate and inherent difficulties faced in the field of microalgal biotechnology (Vieira, 2014).

## 1.4. Current commercial applications of microalgal products

The production of biofuels from microalgae is to date not economically feasible. Therefore, the current microalgal market is mainly for applications in nutraceuticals, food and feed (Table 1). The value of the global microalgae market was estimated at €2.4bn in 2014 with an expected yearly growth of 10%. The health food market represents over 75% of this value (Enzing et al., 2014). The potential of microalgae as food and feed supplements was originally due to their high protein content (~60%) and

the fact that their amino acid composition compares favourably to other food proteins such as whey. They also have high concentrations of polyunsaturated fatty acids (PUFAs) of more than 18 carbons (omega-3 and omega-6). Of these groups of PUFAs the largest markets are for eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) (Adarme-Vega et al., 2012). Microalgae are also a potential source of almost every vitamin: A, B1, B2, B3, B5, B6, B9, B12, C, E and H. Finally, certain species are cultivated for astaxanthin and beta-carotene, both of which are carotenoids with anti-oxidant properties. These pigment rich microalgae play a big role in aquaculture. Their applications are mainly for nutrition and colouring the flesh of certain animals (thus improving their market value). In artificial diets the natural sources of pigments are lacking therefore microalgae provide a source of carotenoid pigments like astaxanthin and beta-carotene. The microalgal groups currently cultivated for the food and feed market are *Arthrospira*, *Chlorella*, *Cryptocodinium*, *Dunaliella salina*, *Haematococcus pluvialis*, *Isochrysis galbana*, *Nitzchia*, *Nannochloropsis*, *Phaeodactylum*, *Porphyridium*, *Schizochrytium* and *Spirulina* (Enzing et al., 2014).

Table 1.1. Market figures of microalgae based food and feed products (adapted from Enzing et al., 2014)

| Microalgal species (product) | Production volume (tons year$^{-1}$ dry weight) | Number of producers (key players) | Value of production volume (yearly turnover) |
|---|---|---|---|
| **Food and feed products: Whole dried microalgae biomass** | | | |
| *Spirulina* | 5000 (2012) (Norsker et al., 2011) | >15 companies (Cyanotech/Earthrise) (Milledge, 2012) | £26m (2005) (Milledge, 2012) |
| *Chlorella* | 2000 (2003) (Norsker et al., 2011) | >70 companies (Milledge, 2012) | £25m (2006) (Milledge, 2012) |
| **Food and feed products: Microalgae components** | | | |
| *Haematococcus pluvialis* (astaxanthin) | 300 (2004) (Norsker et al., 2011) | >8 companies (Fuji Chemicals, Cyanotech) | £6.5m (2004) (Milledge, 2012) |
| Cyanobacteria (pigment colourants) | (NA) | >2 companies (Milledge, 2012) | £32.5m (2004) (Milledge, 2012) |
| *Crypthecodinium* (EPA/DHA) | 240 (2003) (Milledge, 2012) | >4 companies (Martek/DSM) (Eckelberry, 2011) | £9.5bn (2009) (Ismail, 2010) |
| *Dunaliella salina, Schizochytrium, Nannochloropsis* (beta-carotene) | 1200 (2010) (Spolaore et al., 2006) | >10 companies (Cognis/BASF) (Milledge, 2012) | £188m (2012) (Milledge, 2012) |

### 1.4.1. Microalgal industry in the UK

The case for large scale cultivation in the UK, specifically, is constrained by climatic conditions adverse to outdoor microalgal cultivation. Notwithstanding, the interest in microalgae in the UK has increased over the past decade in response to the global interest in biofuels and, more recently, due to their potential as a biomanufacturing platform. Microalgal cultivation in the UK is predominantly still at R&D scale and, despite the increased interest, the overall photoautotrophic production of microalgae in the UK is estimated to be only between one and five tonnes of dry weight per year (Schlarb-Ridley et al., 2013). Several early stage companies are in the process of expansion although do not yet have any significant production capacity (e.g.

AlgaeCytes, Xanthella). The main bottleneck to the development of microalgal technologies in the UK are the lack of access to test, pilot and demonstration facilities in order to prove economic and technological feasibility, environmental sustainability and public acceptability of commercialisation opportunities (Schlarb-Ridley et al., 2013).

## 1.5. Current state of low- to medium-value compound production

LMVCs have been defined as products with a life cycle cost of \$1000 $t^{-1}$ (metric ton of ash-free dry weight of microalgal biomass) or less (Benemann, 2013). These are typically biofuels, bulk chemicals, animal feeds or fertilisers. To become economically viable microalgae will generally have to be grown outdoors, in open ponds or raceways and photoautotrophically, utilising sunlight and $CO_2$. Currently, commercial microalgal production of LMVCs does not exist. Papers published in 2006 and 2007 (e.g. Chisti, 2007) proposed the idea that if certain microalgae can have up to 70% oil, double every day and have high productivities then microalgae can produce 136900 L $ha^{-1}$. These papers led to an explosion in microalgal investment, however, ignored simple biological, physical and technological limitations that made these targets impossible to achieve. In recent years, R&D groups have come to the realisation that existing production costs will need to be reduced by well over an order of magnitude and productivity increased by a hundred-fold to achieve economic feasibility (Benemann, 2013).

Most of the research in this field is focused on liquid biofuels, mainly on the production of microalgal lipids, in the form of triacylglycerides (TAGs) for conversion to biodiesel. The direct conversion of hydrocarbons and the production of ethanol and biological hydrogen ($H_2$) have also received attention (Greenwell et al., 2010). To overcome the limitation of a single use extraction process the concept of the "microalgal biorefinery" has been introduced. This concept proposes to couple product extraction with the utilisation of the residual biomass to offset the overall cost. Furthermore, the utilisation of pre-developed technologies (e.g. oil industry) and the use of brackish, saline or wastewater resources offers the possibility of integration with existing industrial infrastructure (Hariskos et al., 2014).

### 1.5.1. Microalgal growth systems

No matter what product or technology is utilised, both capital and operating costs, must be low. Currently, microalgal cultivation is carried out in high rate ponds (HRPs) or in closed photobioreactors (PBRs). PBRs are closed systems, with many different configurations (e.g. horizontal, vertical), where the abiotic factors are controlled in order to maximise biomass production (Carvalho et al., 2006). They allow high growth rates, parameter control and reduced risk of contaminants. However, due to thermodynamic limitations in gas exchange (maximum length) and light (maximum width), PBRs cannot be scaled much above 100 $m^2$. Acquiring and maintaining tens of thousands of PBRs would lead to prohibitive capital and operating costs, which leaves HRPs as the only feasible solution to produce LMVCs (Benemann, 2013).

HRPs normally consist of a series of outside reservoirs that harness the power of the sun and maintain the microalgal culture in movement through a paddlewheel system until it achieves its ideal harvesting density. The first growth systems were simple and consisted of artificial ponds, under natural conditions, with little to no control over the growth parameters (Oswald et al., 1957). Several decades of engineering have made the raceway pond the most popular design for outdoor cultivation. The raceway format is designed to maximize areal usage while being able to cultivate a large volume of microalgal culture. Engineering and economics limit pond depth to ~30 cm, mixing velocity to 20 – 25 cm $s^{-1}$, pH from 7.5 to 8.5 and harvest rate from 20% to 40% of total culture volume per day (Weissman et al., 1988). As mentioned above HRPs can be built to link up with wastewater treatment plants or agricultural facilities to exploit nutrient-rich process waters. Additionally, $CO_2$ rich industrial waste streams (e.g. flue gas) can be used to supplement microalgal growth (Hariskos et al., 2014).

### 1.5.2. Downstream processing

Downstream processing (i.e. extraction and purification) of microalgal biomass is a key area of research because it can contribute to up to 60% of the total production cost (Molina Grima et al., 2003). The first step in this chain of processes is a solid – liquid separation (i.e. harvesting) and is, typically, a two-step operation. First, bulk harvesting concentrates the biomass by flotation, flocculation or gravity sedimentation to reach up to 7% total solid matter. Second, thickening steps such as centrifugation or filtration are used to concentrate the biomass slurry (Brennan et al., 2010). These steps

are followed either by a dry or wet route. In a typical dry route process chain the biomass slurry is treated to a mechanical dehydration followed by a thermal drying step (Xu et al., 2011). Drying processes for LMVCs should be developed with waste heat sources or solar energy by taking advantage of local resources and geography. A recent study has shown that the dry route has a higher net energy balance and, in the short term, more potential for the production of LMVCs (Xu et al., 2011). The alternative to the dry route is to perform a wet extraction process. Recent studies have been successful in product extraction from wet biomass through subcritical co-solvent extraction (Chen et al., 2012), *in situ* transesterification (Haas et al., 2011) and acid-base hydrolysis (Laurens et al., 2015).

### 1.5.3. Stability of microalgae mass cultures

It has been determined that the successful production of LMVCs in HRPs depends on obtaining a microalga with the desired characteristics (e.g. high oil content, nutritionally rich), a well-engineered system (i.e. ideal operational parameters), ideal climate and geography (e.g. sufficient insolation, close to waste streams) and a low-cost extraction process. This review will now focus on the stability and productivity of large cultures. One of the considerable challenges of microalgal production in HRPs is the maintenance of the desired strain cultures and the desired levels of productivity. As in any industrial process fluctuations in performance increase waste, thus the ability to consistently supply a final product is key to the production of LMVCs. Factors that may affect the stability of microalgae cultures are invasion by other microalgal species, predation by grazers, fungal and viral infections, bacterial contaminations and changes in environmental conditions (Benemann, 2013). The current solutions are either based on growing cultures in extreme conditions (e.g. high salinity) or with chemical controls (e.g. ammonia, fungicides). Both these control routes still require frequent culture re-starts and extensive pond cleaning. Biological routes of dealing with contamination and culture crashes present themselves as interesting alternatives and are currently under development. These biological control routes are typically divided into single species strategy (i.e. strain improvement by random mutation or genetic modification) or multi-species strategy (i.e. community design).

### 1.5.4. Single species strategy: Strain improvement

One of the solutions proposed to overcome issues with culture stability and, simultaneously, increase productivity has been strain improvement. There are microalgal species that are known to survive in salt concentrations of 5M (e.g. *Dunaliella* spp.), have productivities of 30 g $m^{-2}$ $d^{-1}$ (e.g. *Chlorella vulgaris*), show hydrocarbon contents above 50% (e.g. *Botryococcus braunii*), produce valuable chemicals (e.g. *Haematococcus* spp.) and autoflocculate (e.g. *Scenedesmus dimorphus*). However, finding a natural strain that has all these desirable characteristics is extremely unlikely (Klein-Marcuschamer et al., 2013). Therefore, the field of strain improvement is seen as a more targeted and efficient alternative to bioprospecting, where thousands of natural strains are screened for biotechnologically relevant characteristics. There are three main routes for strain improvement: random mutagenesis followed by screening for desired characteristics (e.g. Work et al., 2010), metabolic engineering (e.g. Guarnieri and Pienkos, 2015) and adaptive evolution (e.g. Velmurugan et al., 2014). Whatever the choice of manipulation route, the final objective remains the creation of the strain with added biotechnological value. Strain improvement will most likely fall under three broad categories: strain resilience (e.g. resistance to viruses, toxic to predators), strain yield (e.g. maximum biomass and production rates, improved photosynthetic efficiency) and added value traits (e.g. omega-3 enrichment, lipids tailored for biofuels) (Gressel et al., 2013). While the widespread implementation of improved strains would undoubtedly bring added value to LMVC production there are potential risks associated to the large scale cultivation of such microalgae. Even with safety measures in place, there are bound to be releases of the culture, be it through human error, material fatigue or climatic events. This leads to risks of contamination of the surrounding environment and crossing with native strains with unpredictable consequences related to the introduction of non-indigenous microalgal species to an ecosystem (Snow et al., 2012). Care is needed to not extrapolate the risks (e.g. creation of a “superalga” is unlikely), however, a strain specific assessment, with the wild type as a baseline comparison, is necessary to ensure its safe use in an industrial setting (Gressel et al., 2013).

### 1.5.5. Multi-species strategy: Microbial communities

One of the more recent solutions proposed for biological control of contaminants is the multi-species approach. In their natural environment, microalgae live in close association with other organisms, many of which are typically seen as contaminants of large scale cultures (e.g. heterotrophic bacteria). Therefore, the design of multi-species artificial communities was proposed in order to protect the desired strain and, potentially, increase culture stability and productivity through cooperative behaviours. The designing of artificial microbial communities is called synthetic ecology and will be reviewed in the following section.

## 1.6. Synthetic Ecology

### 1.6.1. Background

Initially, community ecology was preoccupied with surveying the species found in particular environments and, through this, attempting to elucidate basic community patterns (Clements, 1916). These observations found that community changes, over time, are often quite repeatable. Thus, identification of the processes that drive temporal patterns of community change were fundamental to understand community assembly. These ideas were pervasive in contemporary ecology and the need to summarise them in graphical form led to the creation of the first food webs (Summerhayes et al., 1923). The original food webs, a collection of food chains within a community, were an attempt to organise trophic interactions to elucidate patterns in size distribution and feeding relations. Later, in 1942, Raymond Lindeman introduced the idea of ecological efficiency, a measure of the fraction of energy from one trophic level that is passed on to the next. He proposed that the inefficiency of the process (5 – 15% transfer between trophic levels) may ultimately limit the length of food chains (Lindeman, 1942). While food webs were widely recognised as descriptive devices, it wouldn't be until the 1970s that ecologists utilised quantitative approaches to study their statistical properties and reveal properties that repeated themselves across systems. This led to an increase in number and taxonomic resolution of the food webs under study (Morin, 2011). The expansion of systems studied under the food web theoretical framework eventually came to include aquatic microbial systems (Steele, 1974). However, these first views were simplified and incomplete. It was believed that only the production of large

phytoplankton (e.g. diatoms, dinoflagellates) contributed significantly to the aquatic carbon cycle and food web. Soon after, the role of nanophytoplankton (2 – 20 μm), as responsible for a significant part of aquatic primary production, was recognised. However, the role of bacteria, and their interactions with phytoplankton, was only starting to be explored (Hobbie et al., 1972; Cole, 1982). In 1983, Azam et al. summarised and connected a variety of discoveries made in the preceding decades and updated the classic view of the structure of marine plankton communities (Fenchel, 2008) coining the term "microbial loop" (Azam et al., 1983). The study concluded that a substantial part of the primary production was lost to the environment in the form of dissolved organic matter (DOM). The concept of the microbial loop as a model system will be discussed later in this review.

The major effect of the knowledge that organisms in aquatic microbial systems were interconnected, and formed a somewhat stable community, was an increase in interest in the diversity and functional properties of the organisms involved. In the first decades after the publication of the paper, unicellular eukaryotes in plankton were thoroughly described. This is mainly due to the fact that these organisms can be identified by light microscopy and can easily be cultivated in the lab. Bacteria, on the other hand, are difficult to culture in laboratory conditions and eluded characterisation for many years (Giovannoni et al., 2007). The progress of nucleic acid extraction and sequencing, in the 2000s, changed the field by revealing a huge diversity of "unculturable" bacteria and paving the way for new discoveries amongst archaea and viruses. With new genetic tools at their disposal the scientific community shifted their focus from microbial discovery and description to microbial manipulation. Thus, the field of synthetic biology was born (Stephanopoulos, 2012). Synthetic biology focused on individual organisms and components (e.g. gene circuits) and allowed researchers to learn more about the rules of gene expression and regulation. With this knowledge in hand the next logical step was to move from the individual to the community. The merging of disparate fields such as synthetic biology, community ecology and game theory created synthetic ecology.

The term "synthetic ecology" was first used by Maitreya Dunham to describe work developed by Shou et al. with yeast populations. Dunham described synthetic ecology as a new direction of synthetic biology where mixed cell populations are used to construct a synthetic simple obligatory cooperative ecology (Dunham, 2007). In the experiment by Shou et al. (2007), two yeast strains were engineered to overproduce either lysine or adenine to supplement the corresponding auxotrophy of the partner

strain. This created a mutually obligatory cooperative system that allowed for exploration of microbial mutualisms. Further systems have been created and explored (e.g. bacteria – amoeba in Kubo et al., 2013), but this review will focus on systems that involve eukaryotic microalgae and cyanobacteria and how these systems contribute to microalgal biotechnology.

### 1.6.2. Synthetic ecology for microalgal biotechnology

Synthetic ecology, in its application for microalgal biotechnology, is used to create functionally diverse microbial assemblages with complementary eco-physiological traits that can protect from pathogens and predators and increase productivity while maintaining stable yields (Nalley et al., 2014). This mission statement stems from the concept pioneered by Tilman (1982) that, given a similar environment, species-rich communities are more productive than species-poor communities (Tilman, 1982). This has been experimentally shown for several different habitats, and famously in a long term experiment on grasslands (Tilman et al., 2001). This 7-year experiment showed that a grassland area with 16 species attained 2.7 times greater biomass than the respective monocultures. Tilman later proposed to use these mixtures of native grassland perennials as biomass feedstock for carbon negative biofuels (Tilman et al., 2006). This progression from ecological theory to applied biotechnology is now influencing cultivation techniques (i.e. outdoor cultivation) in microalgal biotechnology. For example, Stockenreiter et al. (2012) showed that both, natural and constructed highly diverse microalgal communities, produced more biomass and more total neutral lipids than equivalent monocultures. These studies are encouraging indications that ecological theory can be successfully used to improve microalgal cultivation for LMVC production. The following section will now discuss the ecological theory behind the diversity – productivity relationship and how it is relevant to microalgal biotechnology.

### 1.6.3. Approaches to community design

In ecological terms, determining what interactions form a stable community is synonymous of determining the drivers of community assembly. However, in the context of microalgal biotechnology, most systems are short-lived and spatially constrained. Therefore, there is no interest in focusing on forces that operate over long temporal or large spatial scales nor on demographic drift or how dispersal influences local communities. The focus should be on the effect of biotic interactions and environmental factors that affect the composition and structure of communities at a small (local) scale (HilleRisLambers et al., 2012).

Two non-exclusive mechanisms have been proposed to explain the observed diversity – productivity relationship in small communities. The "sampling effect" states that a highly productive species in a pool of possible species has a higher likelihood of being present in a more diverse community, thus increasing the overall biomass (Hooper et al., 2005). Second, the "complementary effect" whereby an assortment of species occupies every possible niche ultimately maximising resource utilisation and, consequently, productivity (Hooper et al., 2005). The "sampling effect" may lead to a situation of "non-transgressive overyielding" where the species-rich community is more productive, on average, than the respective monocultures although does not out-produce the best monoculture. In the context of microbial communities for LMVC production, this situation is to be avoided because the most productive monoculture species would be selected over the community for mass production. Therefore the desired outcome is a situation of "transgressive overyielding" where facilitation and complementarity between community members increase productivity beyond the most productive monoculture species (Shurin et al., 2013).

As seen above the desired species-rich community must achieve a situation of "transgressive overyielding". For this end we must understand the notion of diversity and how to obtain the ideal diversity in an artificial system. The basic notion of diversity is synonymous with species richness, which is the total number of species in a system (Morin, 2011). However, to successfully assemble a microbial community one must look at other aspects of diversity such as trait diversity, functional diversity or the diversity of interspecific interaction. Here we define trait diversity as the potential function (e.g. the ability to fix $N_2$), functional diversity as the ecological role occupied (e.g. active $N_2$ fixation) and interspecific interaction as the relationships established with other community members (e.g. sharing the fixed $N_2$). The importance of

understanding these concepts is to avoid functional redundancy when assembling an artificial community. A community consisting of organisms with highly overlapping physiological characteristics (traits) would reduce the possibility for resource-use complementarity and, most likely, cause competition to the point of exclusion (two species competing for the same resource cannot coexist). For example, Nalley et al. (unpublished) assembled a community of three green microalgae and no significant increase in total lipid production was detected when compared to individual monocultures (Nalley et al., 2014). Therefore, to successfully establish cooperative diversity in microbial systems these need to be studied through a mix of trait-based and interaction-based approaches (Grosskopf et al., 2014; Nalley et al., 2014).

#### 1.6.3.1. Trait-based approach to community design

To design a community by a trait-based approach it is useful to look at community assembly from the perspective of contemporary (i.e. Chessonian) coexistence theory (Chesson, 2000). Under this framework, one of the hallmarks of a stable community is negative frequency-dependent population growth. This type of population growth describes a situation where, at any given point in time, the least abundant species has a higher growth rate thus avoiding competitive exclusion. This can only occur when the community is stabilised by niche differences (differences that cause species to more strongly limit themselves than others) (Chesson, 2000). Niche differences can include complementary use of resources (e.g. light, nutrients) or different environmental preferences (e.g. temperature, pH) so that, in combination, the community can outperform the best-performing monocultures (Nalley et al., 2014). For example, joining high and low temperature-adapted organisms may lead to a more efficient use of resources through coexistence (Edwards et al., 2011) (Fig. 1.1).

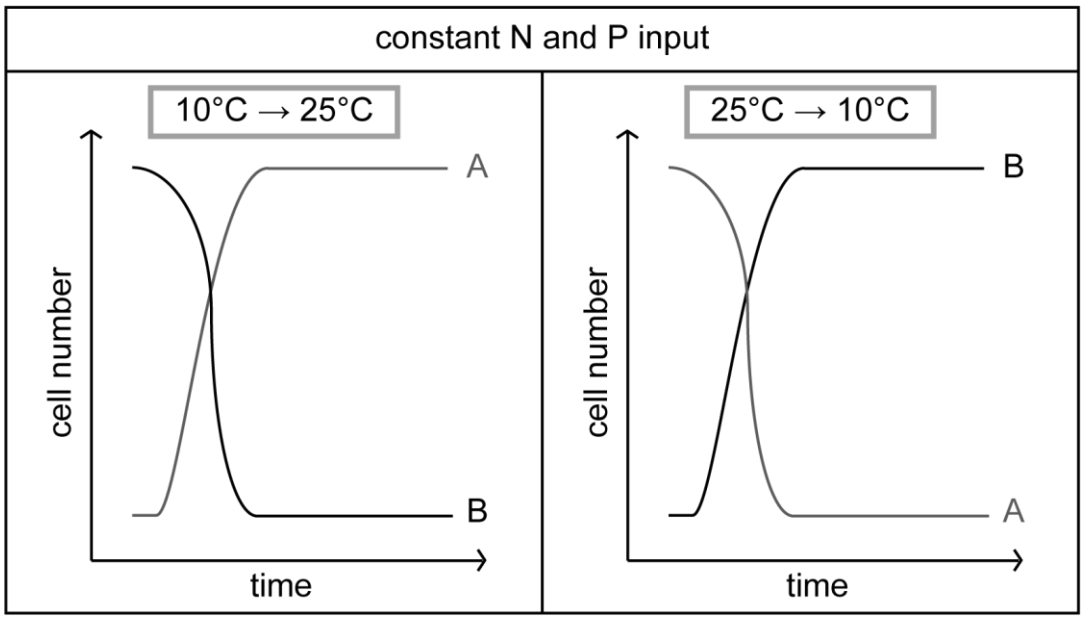


Figure 1.1. Hypothetical microbial system where organisms A and B can potentially coexist due to differences in temperature preference.

This approach can be beneficial when developing microbial communities for LMVC production. The identification of species that grow best under different environmental conditions, but produce the same product, would potentially allow cultures to withstand environmental fluctuation. In support of this hypothesis recent studies have concluded that species composition was crucial in predicting overall stability in systems experiencing environmental stress (Schabhüttl et al., 2013). An extension of this theory is that a diverse microalgal community, with complementary traits and several different occupied niches, may also be resistant to invasive species, disease and zooplankton grazing (Fargione et al., 2005). For example, a trait-based solution for zooplankton invasion could consist of introducing multiple ungrazeable microalgal species thus decreasing the foraging efficiency of invading zooplankton due to the increased energetic costs of finding the desired prey. This was investigated by Shurin et al. (2013) where communities containing one, two, five, or ten species of microalgae in various combinations were subjected to grazing by *Daphnia pulex*. The study concluded that *Daphnia* numbers decreased markedly when five or ten species of microalgae were grown together.

### 1.6.3.2. Interaction-based approach to community design

A complementary approach to trait-based community design is based on the identification of common interactions established between community species (Grosskopf et al., 2014; Escalante et al., 2015). To date, species-specific microbial interactions have attracted attention and have been extensively reviewed (Mitri et al., 2013; Werner et al., 2014). More recent studies have focused on how species interactions affect community assembly and have found that certain interactions can promote species coexistence (Bachelot et al., 2015). However, in the context of microalgal biotechnology, not every interaction is of interest.

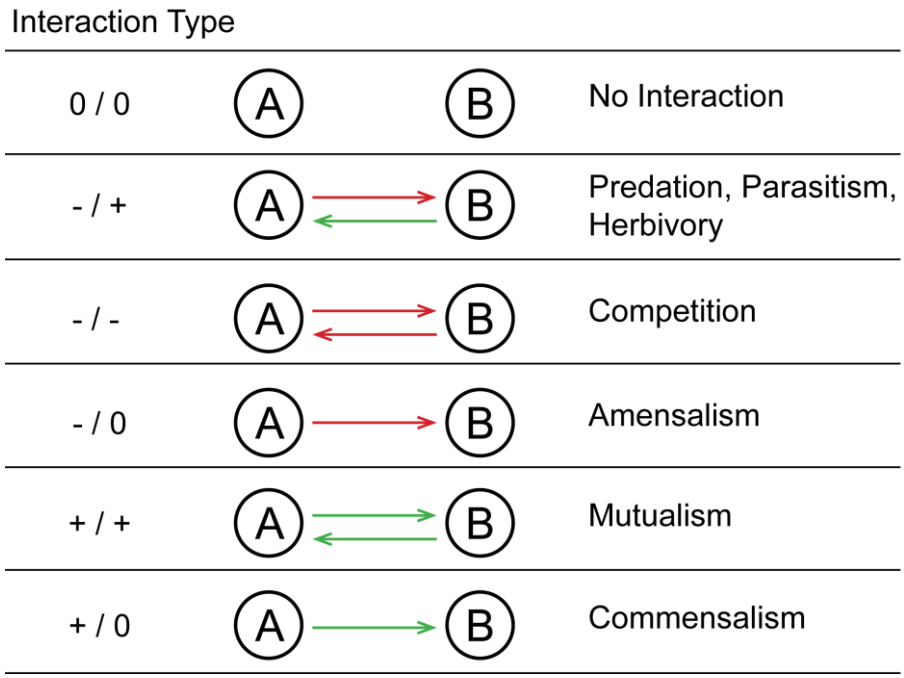


Figure 1.2. The seven basic types of interactions between organisms. Green arrows indicate a net positive effect, red arrows indicate a net negative effect and no arrow indicates no effect.

Interactions between organisms can be classified according to the net effect that organism A has on organism B: neutral, positive or negative (Fig. 1.2). Predation, parasitism, and herbivory all involve a (−/+) interaction between organisms A and B, where the net effect of the predator on the prey is negative and the effect of the consumed prey on the predator is positive. Competition involves a mutually negative (−/−) interaction between organisms A and B. Amensalism is a one-sided competitive (−/0) interaction where organism A has a negative effect on B, but where organism B

has no detectable effect on A. Mutualism involves a mutually positive (+/+) interaction between organisms A and B. Commensalism is a one-sided mutualistic (+/0) interaction, where organism A has a positive effect on B, but where organism B has no detectable effect on A (Morin, 2011). Interactions of the type (−/+), (−/−) and (−/0) have little interest for microalgal biotechnology. A predominance of this type of interactions will always lead to a situation of population unbalance. Thus, the focus will now be on interactions of the type (+/0) and (+/+).

Interactions of the type (+/0) (i.e. commensalism) occur when metabolites produced by organism A affect the growth of organism B without reciprocity. These interactions are characterised by resource partitioning and unidirectional cross-feeding (the phenomenon where one species lives off the products/waste of another species) and occur in nature because resource specialist genotypes are generally competitively superior to generalist genotypes when there is an abundance of resources. If both consumers are generalists, which feed on the same resources, they will compete. If both consumers become resource specialists, to a point where each require a different set of resources, then this can result in a reciprocal (+/+) or asymmetric (+/0) relationship (Morin, 2011). Figure 1.3 shows two examples of how these type of interactions can be relevant to microalgal mass cultivation. Briefly, bacteria are known to decompose microalgal exudate (Guenet et al., 2010). If two bacterial species, A and B, partition the microalgal exudate (e.g. bacterial species A specialises in amino-acid uptake and bacterial species B specialises in carbohydrate uptake) this would allow for coexistence of the microalga and both bacterial species. There are two potential advantages of this interaction: both bacterial species can provide the microalga with essential metabolites (Fig. 1.3 A) or the presence of more bacteria can increase the decomposition of organic matter which provides more carbon and nutrients to the microalga (Fig. 1.3 B).

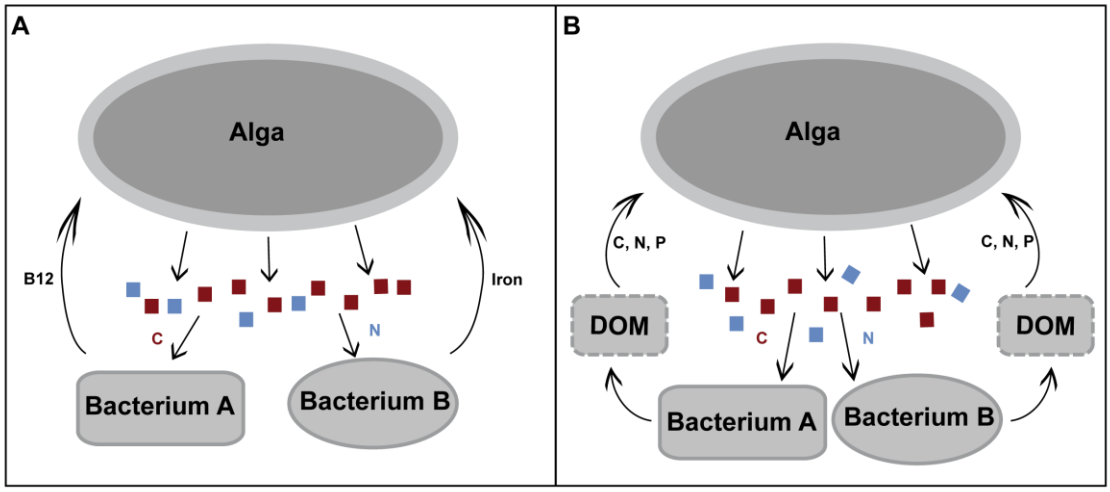


Figure 1.3. Hypothetic microbial community with bacterial coexistence. Species A specialises in amino-acid uptake and bacterial species B specialises in carbohydrate uptake which provides an advantage for the microalga by a) providing essential metabolites or b) increasing the decomposition of organic matter which provides more carbon and nutrients.

This last example introduces the concept of mutualism (+/+ type interaction). Establishing this type of interaction might be crucial for successful community design. If these interactions are obligate mutualists (where two organisms live in close association and depend on each other for survival) it is likely that this community will be highly resistant to invasion due to the specificity of the interaction (Kazamia et al., 2012). However, any type of mutualisms, obligate or facultative, have the potential to create synergistic relationships and increase productivity. Specifically, established mutualisms with particularly problematic contaminants (e.g. bacteria) can benefit mass cultivation in a number of ways. Several cases of successful microalgal – bacteria co-cultures have been reported. Kazamia et al. (2012) successfully co-cultured the $B_{12}$ - dependent green alga *Lobomonas rostrata* and the rhizobial bacterium *Mesorhizobium loti* without $B_{12}$ nor a carbon source indicating the bacterium supplies the vitamin to the microalga in exchange for carbon-rich exudate. Ortiz-Marquez et al. (2012) created a system consisting of a N-fixing bacterium (*Azotobacter vinelandii*) that was genetically modified to secrete ammonium into the surrounding medium. Co-cultures of the strain, in N-deplete medium, allowed for growth of *Chlorella sorokiniana*, *Pseudokirchineriella* sp. and *Scenedesmus obliquus*. Other cases, where the mutualistic mechanism is unknown, show just how widespread these interactions are. Park et al. (2008) reported that co-inoculation of eight bacterial strains, isolated from a long-term

laboratory culture of *Chlorella ellipsoidea*, resulted in 0.5 – 3 times greater microalgal growth than that of *Chlorella ellipsoidea* alone. Also, Do Nascimento et al. (2013) showed that inoculation of *Rhizobium* strain 10II into a culture of *Ankistrodesmus* sp. strain SP2-15 increased chlorophyll, biomass, and lipid accumulation, in the microalga, up to 30%.

As seen above, the screening of pre-existing cultures for possible microalgae – bacteria interactions can be successful. However, screening every microalgal strain (many of them kept in axenic conditions) in search for their bacterial partners would be a significant investment in time, materials, and manpower. Therefore, it would be advantageous to study natural systems where stable communities, and their respective interactions, have evolved over time. Thus, drawing upon the accumulated knowledge of ecological studies can aid microalgal biotechnology.

## 1.7. Microalgal blooms as model systems

One unique characteristic of aquatic systems is that primary production often occurs in blooms when microalgae rapidly multiply in the water column most commonly due to a rapid influx of nutrients (Khan et al., 2005). Thus, understanding microalgal blooms and their similarities to outdoor cultivation systems can simultaneously provide valuable information for microalgal biotechnology and bioremediation. The most common cause of microalgal blooms is a chronic excess of nutrients in the aquatic environment which is called eutrophication. Eutrophic environments originate when water bodies are enriched with nutrients that subsequently cause, sometimes irreparable, chemical and biological alterations. This process is very common in nature and occurs over a geological time scale. Lakes, if brought under natural succession, will take thousands of years to progress from oligotrophic (nutritionally poor), through a mesotrophic state, to become eutrophic (Rast et al., 1996). However, anthropogenic impacts on the environment cause severe imbalance in the natural biogeochemical cycles of C, N and P and drastically accelerate this process. This is known as cultural eutrophication and is considered the most widespread water quality problem on the Earth (Schindler, 2012). Freshwater eutrophication is responsible for damages in the value of £75.0-114.3 million in England and Wales (Pretty et al., 2003). It is estimated that the total UK area of eutrophic water corresponds to 1785 $km^2$ with over 80% of this resource in England, some 40% in Wales and

approximately 15% in Scotland in a eutrophic state. Worldwide 30-40% of freshwater lakes and reservoirs have been affected by eutrophication (Yang et al., 2008). Eutrophication can also affect the general public by making the water improper for consumption, spreading disease due to bacterial growth and decreasing property value and tourism due to unsafe and unappealing water bodies (USEPA, 1998; Shaw et al., 2002). Cyanobacteria, in particular are a group that deserve attention due to the production of freshwater toxins when they bloom (harmful microalgal blooms). Coming into contact with these toxins, through bathing or drinking, can cause the death of animals, gastrointestinal disorders, renal malfunction and severe allergic reactions (O'Neil et al., 2012). The freshwater toxins produced are generally classified as neurotoxins, hepatotoxins, cytotoxins, saxitoxins and endotoxins. Hepatoxins (e.g. microcystins), followed by neurotoxins, are globally the most prevalent (O'Neil et al., 2012).

The most visible manifestation of eutrophication is a microalgal bloom. Blooms are characterised by the uncontrolled growth of eukaryotic microalgae or cyanobacteria and disrupt entire ecosystems due to a severe impact on ecosystem structure and functionality (Shaw et al., 2002). In a healthy ecosystem, surface waters are clear and sunlight penetrates into the deeper waters where fish reproduce. An increase in nutrients causes a rapid multiplication of microalgae (bloom) which, in a first stage impedes light penetration and increases alkalinity (photosynthetic $CO_2$ absorption) (Smith et al., 1999). The decrease in light penetration will lead to macrophyte death. The death of the macrophyte population causes the lake to lose a large part of its nutrient buffering capacity and its source of sediment stabilization (Smith et al., 1999). These two factors are responsible for a continual increase in nutrient concentrations and, consequently, microalgal concentrations. The final stage of the bloom occurs when the nutrients are depleted. The microalgal population starts to decay and catalyses bacterial growth which leads to oxygen consumption and anoxic conditions in the water.

As microalgal blooms are a global problem, they have been the target of numerous studies since the 1970's. Until recently the data collected allowed a general understanding of food web dynamics and species interaction, at a macroscopic level, and phenotypical adaptation to environmental conditions (Carpenter et al., 1987). However, what is of interest for microalgal biotechnology are the microbial processes that underpin the formation of microalgal blooms (i.e. microbial loop). Understanding the microalgae – bacteria interactions, in a natural setting, and how these respond to nutrient enrichment can provide crucial information.

### 1.7.1. Microbial loop

Before the 1970s, the accepted model of aquatic nutrient flux was linear. Primary producers are consumed by zooplankton which in turn are consumed by predators which in turn are consumed by higher predators. This depiction, often called "the pelagic food web" assumed the bulk of primary production is consumed by zooplankton as particulate matter and little is dissolved and available for the microbial community. Many researchers also believed that bacteria live independently from particles and other organisms and are metabolically inactive thus underestimating their role in aquatic nutrient cycling (Grossart, 2010). In 1974, Lawrence Pomeroy proposed a new model for the aquatic food web where the microbial community is crucial to the carbon, energy and nutrient cycles. This new paradigm suggested that microbes were not only primary producers (e.g. microalgae) but also major consumers of DOM (Pomeroy, 1974). Azam and colleagues expanded on Pomeroy's ideas by describing an aquatic food web where the DOM released by primary producers is consumed by bacteria which, in turn, are consumed by ciliates. Microzooplankton consume ciliates and this constitutes the link to the traditional "pelagic food web" (Azam et al., 1983). The initial description of the microbial food web used the dichotomy of grouping microorganisms into photoautotrophs and organoheterotrophs. Microalgae and autotrophic bacteria (primary producers) take up inorganic materials to produce organic matter. This is released into the environment in the form of DOM (through exudation or upon cell lysis). Bacteria consume DOM and remineralise N and P which is supplied to primary producers. Predators, such as zooplankton and heterotrophic protists, consume primary producers and bacteria thus introducing the energy, carbon and nutrients to the higher levels of the food web. This is especially important in low nutrient environments where there is a constant need for nutrient recycling (Fenchel, 2008).

Recently, the importance of previously unknown metabolic modes (e.g. aerobic photoheterotrophy, mixotrophy) has made its way into our current understanding of the microbial food web (Fenchel, 2008; Mitra et al., 2014). However, Graham et al. (2014) suggests that the current depiction of the microbial loop still overlooks three ubiquitous bacterial groups:

1) N-fixing non-cyanobacterial groups. N fixation in cyanobacteria is well documented, however, the role of heterotrophic N fixation is mostly unknown. Heterotrophic N fixation has been found in unusual environments, (e.g. methane seeps

in Dekas et al., 2009), but, presently, there is no evolutionary explanation for their existence.

2) Particle and organism associated bacteria. The concept of bacteria living in close association (or attached) to organic particles and microalgae has been present for several decades (Cole, 1982). Only recently have researchers developed the tools to probe further into this relationship. It has been shown that bacteria which are "surface specialists" can increase protease activity 10 – 20 times after attachment (Grossart et al., 2007), improve environmental conditions for microalgal "hosts" (e.g. reducing oxygen levels to increase activity of the N-fixing enzyme nitrogenase (Ploug et al., 2011)) and degrade high molecular weight compounds that were thought to be inaccessible (Williams et al., 2013).

3) Predatory prokaryotes. Bacterial predation is one of the major controlling factors of nutrient recycling. Recent studies have focused on the importance of bacteriophage (Rohwer et al., 2009) and protist (Mitra et al., 2014) predation, however, the role of bacterial predation by bacteria is relatively unknown. The main effect of these predatory prokaryotes, on the microbial loop, is the provision of an alternative pathway of nutrient recycling. Viruses lyse bacterial cells and provide DOM to the environment, however, predatory prokaryotes retain the products of bacterial lysis inside themselves thus interrupting the flow of energy and nutrients. Presently, the control that predatory prokaryotes exert on the microbial loop, and how this compares to other bacterial predators, is largely unknown (Graham et al., 2014).

Research into the aforementioned aspects of microbial diversity is still ongoing and how they affect the microbial loop is still not fully understood. Based on these research areas a new microbial loop model can be proposed to incorporate new findings (Fig. 1.4).

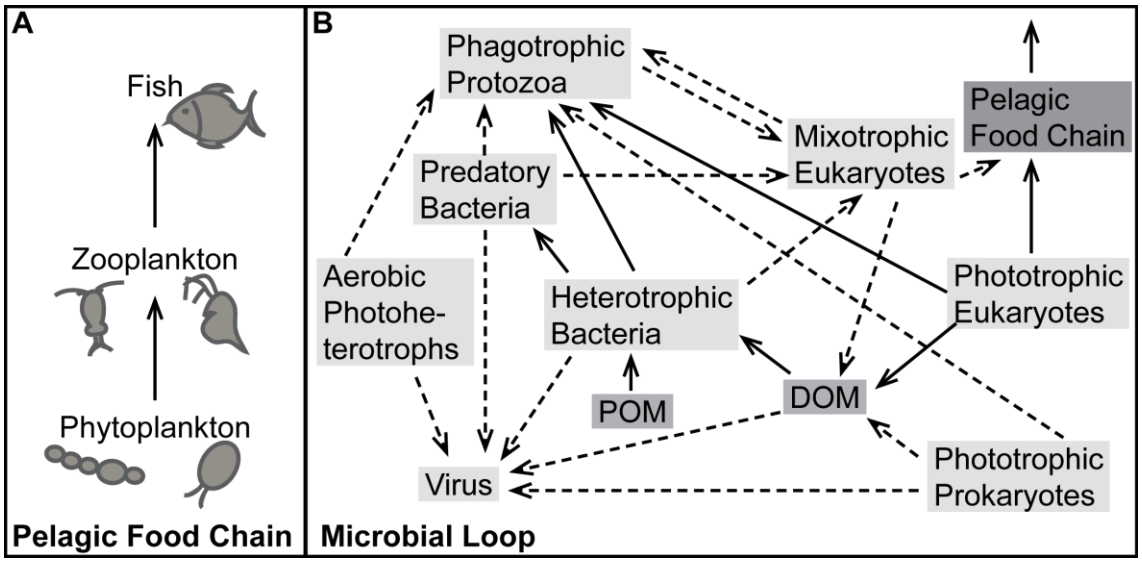


Figure 1.4. Evolution of the conceptual framework of the aquatic food chain. (A) A straightforward phytoplankton – zooplankton – fish aquatic food chain which preceded the description of the microbial loop. (B) A schematic representation of the microbial loop with relationships first proposed by Azam et al. (1983) in full lines and new relationships recently uncovered in dashed lines. Adapted from Furhman et al. (1999) and Fenchel (2008).

Ultimately, this model is speculative and environmental changes will play a crucial role in changing the relationships established in microbial communities. In fact, it has been argued that the "paradox of the plankton", which questions how plankton maintains a high diversity in a highly competitive low-resource environment, might be due to the dynamics of environmental changes and relationships established favouring certain species at certain time points (Scheffer et al., 2003; Paver et al., 2015). Therefore, a detailed study of the microbial loop, and associated phytoplankton groups, will allow us a better understanding of microbial interactions, how these are affected by environmental change and, ultimately, how this information can be applied to microalgal biotechnology.

## 1.8. Microbial metaproteomics

The first descriptions of the microbial loop were from a "black box" perspective due to the information being restricted to a small fraction of cultivable bacteria. In recent years, the boom in DNA sequencing allowed the field of genomics to move towards the analysis of environmental samples (Tyson et al., 2004). Thus, community genomics, or metagenomics, can thoroughly detail microbial genetic diversity and, with metatranscriptomics, identify the regulation and expression of said diversity (Abram, 2015). However, these techniques cannot explain how microbial diversity relates to biogeochemical processes on a whole ecosystem level. Proteins provide the cell framework and define its function, therefore a description of the proteome can provide a description of organism function and its current state. Therefore, metaproteomics allows us to identify the functional expression of the metagenome and elucidate the metabolic activities of a community at the moment of sampling (Wilmes et al., 2006). Metaproteomics was originally defined as "the large-scale characterization of the entire protein complement of environmental microbiota at a given point in time" (Wilmes et al., 2004). The field has evolved from the identification of 3 proteins, in 2004 (Wilmes et al., 2004), to the identification of 7000 in 2015 (Hultman et al., 2015) and has been applied to various environments and microbial consortia.

### 1.8.1. Lessons of aquatic metaproteomics for microalgal biotechnology

The finality of metaproteomics is to describe the behaviour of the microbial community in a specific system. Therefore, this section will focus on the discoveries in marine and freshwater systems which are relevant to microalgal biotechnology.

The vast majority of the studies undertaken have focused on bacterial proteomes in marine environments. This is due to the fact that between the Global Ocean Sampling and Sargasso Sea expeditions 7.5 billion base pairs of non-redundant sequences were generated, annotated and made publicly available (Venter et al., 2004; Rusch et al., 2007). These metagenomics datasets provided a platform for several future marine studies (e.g. Dong et al., 2014). Freshwater metaproteomic analyses have also successfully described ecosystem dynamics without the support of a metagenomic database (Hanson et al., 2014). However, these studies typically rely on very large databases (e.g. NCBI non-redundant bacterial database) which can lead to problems

with protein validation due to protein similarity between organisms. The next best approach is to determine community composition by sequencing of 16S and 18S rRNA and utilise this information to guide database assembly. This is a widely used workflow (e.g. Tang et al., 2014) that has the advantage of compensating for the inherent lack of robustness in proteomic taxonomic assignments. However, like the previous approach, it is biased towards organisms for which genome sequences are publicly available. In the case of organisms used in microalgal biotechnology many of the strains utilised have been sequenced thus providing a solid platform for metaproteomic studies.

Metaproteomics has given us a wealth of information regarding intracellular transport and, specifically, nutrient uptake. Sowell et al. (2009) was the first study to demonstrate the importance of nutrient uptake via high-affinity transporters and to identify transporters that are abundantly expressed. This study focused on the bacterioplankton community of the oligotrophic surface water of the Sargasso Sea. They found the metaproteome to be dominated by proteins belonging to ATP-binding cassette (ABC) transporters. These are high affinity transporters that utilise ATP to drive the uptake of organic (e.g. amino acids, simple sugars, phosphonates) and inorganic (e.g. iron and complexed phosphate) compounds. The synthesis of ABC transporters have a high metabolic cost therefore microbes only invest in the biosynthesis of ABC transporters that transport substrates that are limiting (Sowell et al., 2009). Determining which transporters are synthesised can also reveal the limiting substrates in high nutrient environments. Sowell et al. (2011) investigated the highly productive, nutrient-rich coastal upwelling system of the Oregon shelf (Pacific coast, USA) and found, again, a high abundance of ABC transporters. These transporters were specific for amino acids, taurine, and polyamines, which indicate that C and N were more limiting than phosphate in this environment. Another example, now regarding nutrient remineralisation, was shown in a study by Williams et al. (2012) in the Antarctic Peninsula coastal surface waters. One of the ways that $NO_3^-$ becomes bioavailable is through the process of nitrification, where $NH_4^+$ is oxidized to $NO_2^-$ and $NO_3^-$. This process has been known to be mediated by bacteria. However, Williams et al. (2012), in a summer and winter comparison, reported that ammonia-oxidizing Archaea proteins constituted almost a third of the detected proteins in the winter and were absent in the summer. This suggests that, in some environments, Archaea play a large role in nutrient remineralisation.

Metaproteomics is also a valuable tool to uncover the coexistence dynamics in microbial communities. In a metaproteomic study, in Newcomb Bay in East Antarctica,

of a phytoplankton bloom, *Flavobacteria* were found to utilise high affinity transporters and exoenzymes (TonB-dependent transporters and glycoside hydrolases) for the uptake and breakdown of complex carbohydrates, especially microalgal exudates. The breakdown of these complex substrates released simple substrates that were made available to the community allowing for a succession of other bacterial species as the bloom progressed (Williams et al., 2013). Georges et al. (2014) reported an occurrence of resource partitioning, by two closely related clades of *Gamma-proteobacteria* (ARCTIC96BD-19 and SUP05), in the Northwest Atlantic Ocean (Bedford Basin). The identification of ARCTIC96BD-19 proteins involved in the transport of organic compounds indicated a heterotrophic metabolism. In contrast, the identification of sulphur oxidation proteins from SUP05 indicated the use of reduced sulphur as energy source. This metabolic differentiation potentially allows both bacterial clades to coexist in the same environment.

In two freshwater metaproteomic studies Ng et al. (2010) and Lauro et al. (2011) have described in detail the ecophysiology of a green sulfur bacterium (*Chlorobiaceae*), a primary producer found in Ace Lake, Antarctica. Ng et al. (2010) identified the set of proteins that allow this bacterium to thrive under cold, nutrient-limited, oxygen-limited and extremely varied annual light conditions. The presence of chlorosomes (bacteriochlorophyll structures) allows growth at extremely low light intensities and several cold adaptation mechanisms were also identified. Finally, they reported the existence of a cross-feeding relationship between green sulphur bacteria and sulphate-reducing bacteria where an exchange of sulphur compounds allows for the survival of both organisms. Lauro et al. (2011) provided a broader view of the Ace Lake ecosystem by developing a competition model based on the metaproteomic analysis of the green sulphur bacteria. This model offered several possible outcome scenarios of local competition between green sulphur bacteria and cyanobacteria and how this relationship changes with shifts in seasonal light–dark cycles and virus predation.

The metaproteomic characterisation of the strains of interest, and associated bacteria, can give precious insight into the factors limiting growth at a specific point in time. This has the potential to be integrated into the development of growth medium for different strains and communities. For microalgal biotechnology community design, the discovery of new metabolic pathways and adaptation mechanisms could provide a list of alternative organisms to fulfil the same roles under different environmental conditions. It is of interest to utilise this information to create system models in order to describe and predict relationships between abiotic and biotic factors influencing the

microbial community. This tie between metaproteomics and modelling can be useful for predicting responses to environmental changes, biomass productivity and even community response to invasion and predation.

Over the next years further development can be expected and it remains to be seen if ecological principles, such as optimal nutrient supply and co-existence theory, can be expanded to microalgal biotechnology. If so, the utilisation of ecology within microalgal cultivation will need to tackle many challenges, including the feasible up-scaling of a production system and reduction of biomass production costs, to, ultimately, influence the future role of microalgae on the global market.

# Chapter 2: Multivariate analysis of the response of a freshwater microbial community under nutrient enrichment

## 2.1. Abstract

Anthropogenic eutrophication has caused widespread environmental problems in freshwater lakes, reducing biodiversity and disrupting the classic pelagic food chain. However, there is a lack of consensus regarding the exact role of nutrients and physicochemical variables on microbial dynamics and subsequent microalgal and cyanobacterial blooms. In an experimental microcosm study, we investigated the interaction among nutrients, physicochemical variables, microalgae and bacteria under oligo and eutrophic nutrient levels. We were specifically aiming to understand whether aquatic physicochemistry and nutrient dynamics act additively, synergistically or antagonistically on the biotic (eukaryotic microalgae-bacteria) community. We followed the growth of microalgae, cyanobacteria and heterotrophic bacteria from the microbial community. Using a combination of traditional univariate analyses and more modern multivariate approaches, we found that nutrients (nitrogen and phosphorus) and physicochemistry (pH, DO, temperature) act additively on the interactions between microalgae and bacteria, with the effects of physicochemistry elevated in eutrophic conditions. The data suggests that there is no synergistic interaction between nutrients and physicochemistry. Physicochemistry becomes a stronger predictor of biotic variation in nutrient enriched conditions and both nutrients and physicochemistry play independent roles in driving primary productivity. The application of multivariate methods to such data may prove useful for assessing a variety of human and natural perturbations in aquatic communities.

## 2.2. Introduction

Anthropogenic water pollution through nutrient enrichment is well known to disrupt aquatic biodiversity and the biogeochemical cycles of C, N and P. Nutrient enrichment, referred to as eutrophication, can lead to blooms of eukaryotic microalgae or cyanobacteria and associated shifts in the number of trophic levels supported and the diversity of species in lakes and ponds (Anderson et al., 2012). This is considered one of the most widespread water quality problems on Earth (Schindler, 2012) with 30-40% of lakes and reservoirs worldwide affected by eutrophication (Yang et al., 2008).

Understanding the effect of eutrophication on freshwater ecosystems requires an analysis of factors both internal and external to the system, as well as the complex interactions among them. Research into the process of eutrophication is typically centred on a univariate assessment of a subset of indicator variables including (i) nutrient concentrations (typically $NO_3^-$, phosphate ($PO_4^{3-}$) and $NH_4^+$), (ii) physicochemical factors (e.g. temperature, dissolved oxygen (DO), pH) and (iii) biological characteristics (e.g. relative abundances of microalgae and bacteria and higher trophic levels). These studies have regularly identified increasing nutrient concentrations as a key factor responsible for promoting blooms (Heisler et al., 2008; Smith et al., 2009).

However, more recent studies indicate that physicochemical factors, such as temperature, are also a key factor in microalgal bloom promotion (Johnk et al., 2008) and may even play a larger role than nutrient loading (Paerl et al., 2008; Kosten et al., 2012) Additionally, recent modelling and long-term data studies have suggested that nutrient concentrations and physicochemical factors may act synergistically and enhance the frequency and magnitude of blooms (Elliott, 2012). Thus, there is no consensus regarding the relative importance of nutrient concentrations and physicochemical factors in bloom promotion.

A feature of this historical research is the use of univariate statistical tools such as correlations and multiple linear regression. These tools may reveal an incomplete picture because they fail to capture potential interactions among nutrients, aquatic physicochemistry and the biotic community. For example, Rigosi et al. (2014), in an analysis of over 1000 USA lakes, used correlations and multiple linear regression to assess how cyanobacterial biovolume and chlorophyll *a* responded to nutrient availability and temperature. Whilst such univariate approaches hint at interactions (e.g.

chlorophyll *a* concentrations respond to both nutrients and temperature), a multivariate view of eutrophication, capturing simultaneous covariation among variables, is needed.

Path analysis (i.e. SEM) is a widely used technique for proposing and testing plausible sets of causal relations among three or more observed variables, whilst considering the simultaneous analysis of multiple pathways between these. As a multivariate technique, it deals explicitly with multiple testing and calculates partial correlations between all variables, controlling for all others (Arhonditsis et al., 2006). The models also allow the incorporation of multivariate indicator variables (i.e. latent variables), which provide the opportunity to statistically test multiple and overlapping regressions between variables of interest.

In this study we hypothesised three multivariate, "latent", variables: Biotic, Physicochemistry and Nutrients (Fig 2.1 A). Each of these is comprised of three potential variables that vary in space and or time. For example, the biotic community is comprised of eukaryotic microalgae, cyanobacteria and heterotrophic bacteria. This path-analytic framework allows evaluation of various hypotheses about the presence and absence of interactions among latent variables, potential direct and indirect interactions among them, and the variance, covariance and correlation among components of the latent variables (see Methods).

Microcosm experiments were carried out with traditional univariate analyses and multivariate SEM to determine the relative importance of nutrient concentrations ($NO_3^-$, $PO_4^{3-}$, $NH_4^+$) and physicochemical factors (temperature, DO and pH) in driving the dynamics of freshwater microbial groups (eukaryotic microalgae, cyanobacteria and bacteria) under low (oligotrophic) and high (eutrophic) nutrient treatments. We specifically address the following questions: (1) How do the univariate relationships found in the experimental microbial system compare to previous studies? (2) Do nutrients and physicochemistry act additively, synergistically or antagonistically to influence the biological component of the microbial community? and (3) Does the nature of the relationships among nutrients, physicochemistry and biological components of the system vary with nutrient enrichment?

## 2.3. Methods

### 2.3.1. Microcosm setup

Replicate experimental biological communities were constructed in 30 L white, opaque, polypropylene vessels, 42 cm high and with an internal diameter of 31 cm. The microcosms were housed in controlled environment facilities at the Arthur Willis Environmental Centre at the University of Sheffield, U.K. These were filled with 15 L of oligotrophic artificial freshwater growth medium (for detailed composition see Appendix 1). Over the course of the experiment the microcosms were kept at constant temperature, 23˚C, under 100 µmol $m^{-2}$ $s^{-1}$, provided by Hellelamp 400 watt IR Lamps HPS (Helle International Ltd., UK) with a 12:12 light dark cycle. A microbial community was introduced into each microcosm (detailed composition in Appendix 2). This inoculum was sourced from water samples collected at Weston Park Lake, Sheffield (53˚22'56.849'' N, 1˚29'21.235'' W). The inoculum was filtered with a fine mesh cloth (maximum pore size 200 µm) to exclude big particles, protists and grazer populations (Downing et al., 1999). The filtered sample was cultured for five days in the conditions described to allow acclimation to the controlled conditions. Subsequently, each 15 L media was inoculated with 2.5 L of this sample.

### 2.3.2. Experimental design

The experimental elevation of nutrient levels followed United States Environmental Protection Agency guidelines for oligotrophic and eutrophic conditions in freshwater lakes and reservoirs (USEPA, 1986): (1) non-enriched growth medium to simulate oligotrophic conditions ($NO_3^-$ = 0.42 mg $L^{-1}$ and $PO_4^{3-}$ = 0.03 mg $L^{-1}$) and (2) $NO_3^-$ and $PO_4^{3-}$ enriched growth medium ($NO_3^-$ = 4.20 mg $L^{-1}$ and $PO_4^{3-}$ = 0.31 mg $L^{-1}$) to simulate eutrophic conditions. Each treatment was replicated eighteen times, allowing for serial but replicated (n = 3) destructive sampling during the experiment. The experiment was run for 18 days. Three control microcosms, comprised of non-enriched growth medium, with no biological inoculum, were also followed to verify physicochemical variation in the absence of biological activity (Fig. 2.2).

### 2.3.3. Sampling of abiotic variables

Over the course of the experiment DO, pH, temperature, $NO_3^-$, $PO_4^{3-}$ and $NH_4^+$ were monitored in order to link the abiotic variation to the changes observed in the biological variables. DO, pH and temperature were measured at 12:00 and 18:00 daily with a Professional Plus Quatro (YSI, USA). 15 mL aliquots were collected and filtered (0.45 μm), daily, for the estimation of $NO_3^-$, $NH_4^+$ and $PO_4^{3-}$ concentrations. A total volume of 270 mL was removed from each microcosm, for nutrient analysis, over the course of the experiment. $NO_3^-$ concentrations were estimated with a Dionex ICS-3000 ion chromatograph (Thermo Fisher Scientific, USA) using an AG18 2x250 mm column with a 0.25 mL $min^{-1}$ flow rate and 31.04 mM potassium hydroxide as eluent. $NH_4^+$ concentrations were estimated with a Dionex ICS-3000 ion chromatograph (Thermo Fisher Scientific, USA) using a CS16 4x250 mm column with a 0.36 mL $min^{-1}$ flow rate and 48 mM methanesulfonic acid as eluent. $PO_4^{3-}$ concentrations were estimated according to the ammonium molybdate spectrometric method defined by British standard BS EN ISO 6878:2004 (BSI, 2004). Briefly, 10 mL of the filtered sample were mixed with 0.4 mL of 4.5 M sulphuric acid, 1 mL of 0.57 M ascorbic acid and 1 mL of 48.35 mM sodium thiosulphate pentahydrate and allowed to reduce for 10 min. After reduction, 2 mL of an acid molybdate solution (4.14 M sulphuric acid, 21.04 mM ammonium molybdate tetrahydrate and 1.14 mM antimony potassium tartrate hemihydrate) were added and the sample was incubated for 10 min. Finally, the absorbance of each solution was measured at 880 nm against water.

### 2.3.4. Sampling of biotic variables

Chlorophyll *a* and phycocyanin fluorescence were measured daily, *in situ*, with the AlgaeTorch (bbe Moldaenke GmbH, Germany) and utilised as proxies for microalgal and cyanobacterial biomass, respectively. By using specific excitation spectra of fluorescence, 450 nm for chlorophyll *a* and 610 nm for phycocyanin, the two spectral groups of chlorophyll based microalgae and cyanobacteria can be differentiated *in situ*. An internal algorithm is then used to calculate the relative amount of each class that must be present, expressed in terms of the equivalent amount of biomass per litre of water (μg $L^{-1}$) (Beutler et al., 2002).

Total heterotrophic bacteria were measured by using culturable heterotrophic bacteria as a proxy (Salvesen et al., 2000). Heterotrophic bacteria were enumerated

every three days by sampling 100 µL aliquots, in triplicate, plating on R2A agar (Oxoid, UK) and incubating for 24 h at 38˚C. A total volume of 1.8 mL was removed from each microcosm, for bacterial counts, over the course of the experiment. Colony forming units (CFU $mL^{-1}$) were counted using OpenCFU software (Geissmann, 2013). Because bacteria were only enumerated every three days, linear interpolation was used to generate a daily time series to obtain a uniform sample size across all variables. Interpolated values were calculated using the formula:

$$y = y_1 + (y_2 - y_1)\frac{x - x_1}{x_2 - x_1}$$

where $y$ is the missing value, $x$ is the missing time point, $y_1$, $y_2$ are the two closest measured bacterial counts and $x_1$, $x_2$ are the respective time points.

### 2.3.5. Effects of nutrient enrichment on experimental variables

Initially, t-tests were used to assess whether nutrient elevation from oligotrophic to eutrophic levels significantly increased the concentration of microalgae, cyanobacterial fluorescence and abundance of heterotrophic bacteria. The control (oligotrophic nutrient levels, no biological inoculum) treatment was not included in this comparison. The mean values of DO, pH and temperature among control, oligotrophic and eutrophic treatments were compared via one-way ANOVA.

Three metrics were defined to assess the effects of enrichment on the dynamics of the biotic community components (i.e. eukaryotic microalgae, bacteria): the peak value of the variable, the time point of the peak value and duration of elevated levels of the variables. The peak value corresponds to the mean of the three maximum values estimated from each of the three replicates of each variable. The time of the peak value corresponds to the mean of the three replicate times at which the maximum values were registered for each of the three replicates. The duration of the peak value was obtained using the formula:

$$D_p = t_2 - t_1$$

where $D_p$ is the duration of the peak, $t_1$, $t_2$ are the time points corresponding to the lowest values before and after the identified peak, respectively, for each replicate.

We estimated these values for the longest set of each time series (n = 3) and evaluated the differences between oligotrophic and eutrophic conditions using t-tests. All tests were Bonferroni corrected.

### 2.3.6. Pairwise correlations among variables

In order to compare the data to other published studies, where pairwise relationships were made, the Pearson correlation coefficients among all of the variables were estimated. P-values were adjusted with a Bonferroni correction to account for multiple testing and, after correction, differences were deemed significant for $p < 0.001$.

### 2.3.7. Quantifying direct and indirect effects among functional biology, nutrients and physicochemistry

SEM was employed to formally test the hypothesis that the latent variables of nutrients and physicochemistry act additively on the biotic community latent variable. The effort simultaneously estimates the strength and direction of covariation among the latent variables, and the covariation among component variables.

First, a baseline causal model that allows relationships among nutrients, physicochemistry and the biotic community during the process of eutrophication was constructed. Only variables that change through time are included; variables that are constant have no explanatory power. The baseline model (Fig. 2.1 A) is comprised of three latent variables. The physicochemical latent variable is comprised of pH, temperature, DO. Conductivity and salinity were excluded because they did not vary. The nutrient latent variable is comprised of $NH_4^+$, $NO_3^-$ and $PO_4^{3-}$. The biotic latent variable was comprised of colony forming bacterial units (CFU $mL^{-1}$), microalgae concentrations ($\mu g\ L^{-1}$) and cyanobacteria concentrations ($\mu g\ L^{-1}$). This model allowed the estimation of direct and indirect effects among all latent variables. It also allowed the estimation of variance and covariance among contributing variables within the latent variables. Two additional models were also specified to evaluate, against Fig. 2.1 A, the effect of each of the nutrients and physicochemistry latent variables on the biotic community. The models removed either the latent variable nutrients or physicochemistry (Fig. 2.1 B, C). Comparing each reduced model to the full model (Fig. 2.1 A) tests the hypotheses regarding the strength and importance of the removed latent variable.

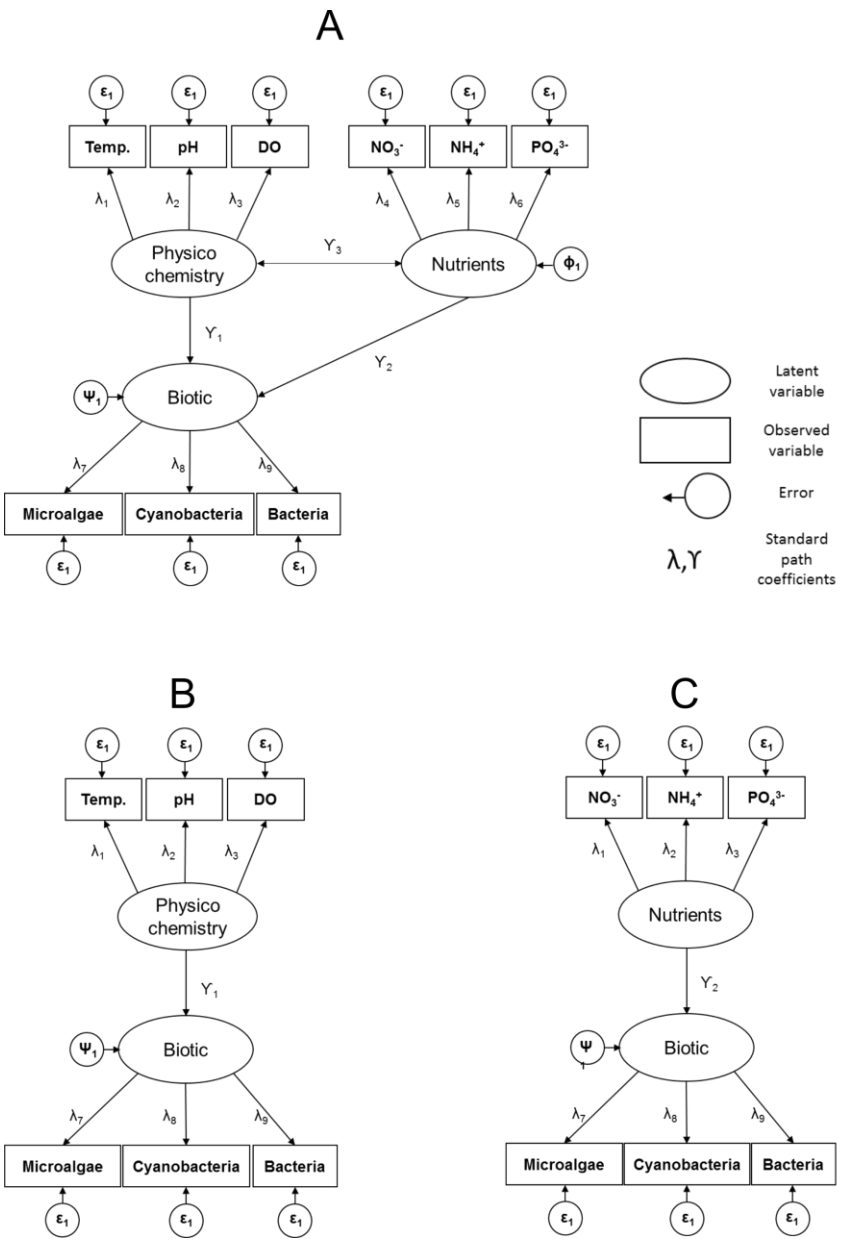


Figure 2.1. Structural equation models showing proposed relationships between latent variables physicochemistry, nutrients and biotic. Rectangles represent directly measured variables (e.g. DO). Ovals represent latent variables (e.g. biotic). In model (A) a full ecosystem model is proposed, incorporating all measured variables, with and without an interaction between physicochemistry and nutrients to test the hypothesis of additive effects. Models (B) and (C) were proposed to test the effect of the strength and importance of the removed latent variables.

Each of these assessments were made in the oligotrophic and in the eutrophic conditions. The estimated path correlations for each treatment allow insight into how the strength and relative role of the nutrients and physicochemistry latent variables vary

with enrichment. The SEMs were fit in the R Statistical Programming Environment (R Development Core Team, 2014) by employing the package "lavaan" (Rosseel, 2012) all variables were scaled to one standard deviation prior to analysis.

## 2.4. Results

### 2.4.1. Effects of nutrient enrichment on biological variables

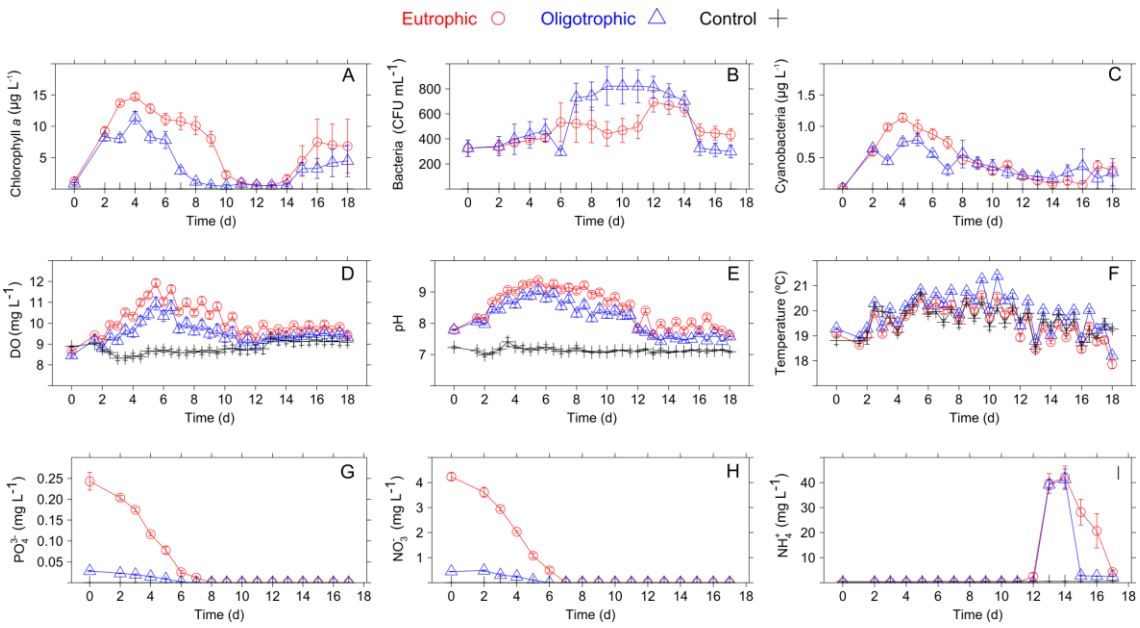


Figure 2.2. Time series of the measured variables in the oligotrophic and eutrophic treatments. (A) Microalgal fluorescence, (B) cyanobacterial fluorescence, (C) culturable heterotrophic bacteria, (D) DO, (E) pH, (F) temperature, (G) $NO_3^-$, (H) $PO_4^{3-}$ and (I) $NH_4^+$ concentrations. Legend is shown in the top left panel (triangles indicate the oligotrophic treatment, circles indicate the eutrophic treatment and crosses indicate the control). Error bars show standard errors (n = 3).

### 2.4.1.1. Temporal variation in the microcosms

The mean concentration of microalgae in the eutrophic treatment was significantly higher than in the oligotrophic treatment (Fig. 2.2 A; $t_{92.494} = -3.160$, $p = 0.002$). No significant difference was found in the mean concentrations of heterotrophic bacteria (Fig. 2.2 B; $t_{99.982} = -0.686$, $p = 0.494$) or cyanobacteria (Fig. 2.2 C; $t_{84.171} = 1.739$, $p = 0.086$).

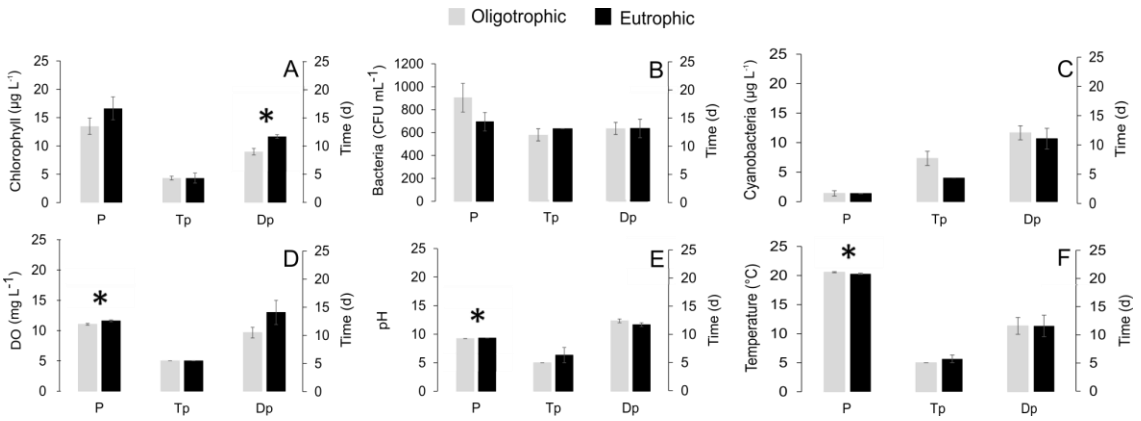


Figure 2.3. Comparison of peak abundance (P), time of peak abundance ($T_p$) and duration of peak abundance ($D_p$) of the measured variables in the oligotrophic and eutrophic treatments. (A) Microalgal fluorescence, (B) cyanobacterial fluorescence, (C) culturable heterotrophic bacteria, (D) DO, (E) pH and (F) temperature. Legend is shown in the top middle panel (grey bars indicate the oligotrophic treatment and black bars indicate the eutrophic treatment). Asterisks indicate significant differences ($p < 0.05$). Error bars show standard errors ($n = 3$).

### 2.4.1.2. Peak concentration

Peak microalgal concentrations were 23% higher in the eutrophic treatment. However, this increase was not significant (Fig. 2.3 A; $t_{3.611} = -1.280$, $p = 0.277$). Peak heterotrophic bacteria concentrations were also not different between the oligotrophic and eutrophic treatments (Fig. 2.3 B; $t_{3.416} = 1.405$, $p = 0.244$). There was no significant difference in peak cyanobacteria concentrations between the oligotrophic and eutrophic treatments (Fig. 2.3 C; $t_{2.022} = 0.074$, $p = 0.948$).

#### 2.4.1.3. Time of peak

There was no significant difference in the time of peak microalgae (Fig. 2.3 A; $t_{2.56} = 0$, $p = 1$), heterotrophic bacteria (Fig. 2.3 B; $t_2 = -1$, $p = 0.423$) and cyanobacteria (Fig. 2.3 C; $t_2 = 2.77$, $p = 0.109$) concentrations between the oligotrophic and eutrophic treatments.

#### 2.4.1.4. Peak duration

The duration of the peak microalgal concentrations was significantly shorter in the oligotrophic treatment than in the eutrophic treatment (Fig 2.3 A; $t_{3.2} = -4$, $p = 0.025$). There was no significant difference in the duration of peak heterotrophic bacteria (Fig 2.3 B; $t_{3.445} = 0$, $p = 1$) and cyanobacteria (Fig. 2.3 C; $t_{3.523} = 0.469$, $p = 0.667$) concentrations between the oligotrophic and eutrophic treatments.

### 2.4.2. Effects of nutrient enrichment on physicochemistry

#### 2.4.2.1 Temporal variation in the microcosms

The mean concentration of DO showed significant difference among the control, oligotrophic and eutrophic treatments (Fig. 2.2 D; $F_{2,150} = 39.36$, $p < 0.001$). A post hoc Tukey test showed that the mean concentrations of DO were significantly lower in the oligotrophic treatment than in the eutrophic treatment and both were significantly higher than in the control ($p < 0.001$). The mean values of pH showed significant difference among the control, oligotrophic and eutrophic treatments (Fig.2.2 E; $F_{2,150} = 106.5$, $p < 0.001$). A post hoc Tukey test showed that the mean values of pH were significantly lower in the oligotrophic treatment than in the eutrophic treatment and both were significantly higher than in the control ($p < 0.001$). The mean values of temperature showed significant difference among the control, oligotrophic and eutrophic treatments (Fig. 2.2 F; $F_{2,150} = 3.596$, $p = 0.03$). A post hoc Tukey test showed that the mean values of temperature were significantly higher in the oligotrophic treatment than in the eutrophic and control treatments ($p = 0.03$).

#### 2.4.2.2. Peak values

An ANOVA on the peak concentration of DO showed significant difference among the control, oligotrophic and eutrophic treatments (Fig. 2.3 D; $F_{2,6} = 58.97$, $p < 0.001$). A post hoc Tukey test showed that there is no significant difference between the oligotrophic and eutrophic treatments ($p = 0.094$), however, the peak concentration in both treatments is significantly higher than in the control ($p < 0.001$). An ANOVA on the peak values of pH showed significant difference among the control, oligotrophic and eutrophic treatments (Fig. 2.3 E; $F_{2,6} = 944.3$, $p < 0.001$). A post hoc Tukey test showed that there is no significant difference between the oligotrophic and eutrophic treatments ($p = 0.197$), however, the peak values in both treatments are significantly higher than in the control ($p < 0.001$). There was no significant difference in peak temperature values among the control, oligotrophic and eutrophic treatments (Fig. 2.3 F; $F_{2,6} = 0.8$, $p = 0.492$).

#### 2.4.2.3. Time of peak

There was no significant difference in the time of peak DO concentrations (Fig. 2.3 D; $F_{2,6} = 1.076$, $p = 0.399$), values of pH (Fig. 2.3 E; $F_{2,6} = 1.386$, $p = 0.320$) and temperature (Fig. 2.3 F; $F_{2,6} = 0.636$, $p = 0.562$) among the control, oligotrophic and eutrophic treatments.

#### 2.4.2.4. Peak duration

An ANOVA on the duration of peak DO concentrations showed significant difference among the control, oligotrophic and eutrophic treatments (Fig. 2.3 D; $F_{2,6} = 24.09$, $p = 0.001$). A post hoc Tukey test showed that there is no significant difference between the oligotrophic and eutrophic treatments ($p = 0.228$), however, the duration of peak concentrations in both treatments is significantly higher than in the control ($p = 0.006$ and $p = 0.001$, respectively). An ANOVA on the duration of peak pH showed significant difference among the control, oligotrophic and eutrophic treatments (Fig. 2.3 E; $F_{2,6} = 546$, $p < 0.001$). A post hoc Tukey test showed that there is no significant difference between the oligotrophic and eutrophic treatments ($p = 0.269$), however, the

duration of the peak in both treatments is significantly higher than in the control ($p < 0.001$). There was no significant difference in the duration of peak temperature among the control, oligotrophic and eutrophic treatments (Fig. 2.3 F; $F_{2,6} = 0.167$, $p = 0.85$).

### 2.4.3. Pairwise correlations among variables

12 of the 54 pairwise correlations, between physicochemical, nutrient and biotic variables (Table 2.1), were significant, after Bonferroni correction, at $p < 0.001$. In the oligotrophic treatment heterotrophic bacterial concentrations have negative correlations with $PO_4^{3-}$ ($r = -0.44$) and microalgae ($r = -0.45$) concentrations. In the eutrophic treatment heterotrophic bacterial concentrations have negative correlations with $NO_3^-$ ($r = -0.46$) and $PO_4^{3-}$ ($r = -0.45$) concentrations. Microalgal concentrations have positive correlations, in both the oligotrophic and eutrophic treatments, with DO ($r = 0.56$ and 0.49, respectively) and pH ($r = 0.58$ and 0.55, respectively). Also, microalgal concentrations have a positive correlation with cyanobacterial concentrations in the eutrophic treatment ($r = 0.61$). Microalgal concentrations have no significant correlations with $NH_4^+$, $NO_3^-$ or $PO_4^{3-}$ concentrations in either treatments. Cyanobacterial concentrations have a positive correlation with DO ($r = 0.65$), pH ($r = 0.74$) and temperature ($r = 0.50$) in the eutrophic treatment. Cyanobacterial concentrations have no significant correlations with $NH_4^+$, $NO_3^-$ or $PO_4^{3-}$ concentrations in either treatments.

Table 2.1. Correlation matrix of measured variables. The asterisks indicate significant values after Bonferroni correction ($p < 0.001$).

| Variable 1 | Variable 2 | Oligotrophic | Eutrophic |
|---|---|---|---|
| Bacteria | $NO_3^-$ | -0.40 | -0.46* |
| | $PO_4^{3-}$ | -0.44* | -0.45* |
| | $NH_4^+$ | 0.22 | 0.36 |
| | Temp. | 0.27 | -0.03 |
| | DO | -0.08 | 0.00 |
| | pH | 0.02 | -0.15 |
| Microalgae | $NO_3^-$ | 0.38 | 0.20 |
| | $PO_4^{3-}$ | 0.40 | 0.21 |
| | $NH_4^+$ | -0.31 | -0.37 |
| | Temp. | 0.17 | 0.29 |
| | DO | 0.56* | 0.49* |
| | pH | 0.58* | 0.55* |
| Cyanobacteria | $NO_3^-$ | 0.06 | 0.14 |
| | $PO_4^{3-}$ | 0.06 | 0.17 |
| | $NH_4^+$ | -0.23 | -0.41 |
| | Temp. | 0.37 | 0.50* |
| | DO | 0.33 | 0.65* |
| | pH | 0.41 | 0.74* |
| Temperature | $NO_3^-$ | 0.08 | 0.19 |
| | $PO_4^{3-}$ | 0.07 | 0.16 |
| | $NH_4^+$ | -0.37 | -0.48* |
| DO | $NO_3^-$ | -0.19 | -0.33 |
| | $PO_4^{3-}$ | -0.13 | -0.31 |
| | $NH_4^+$ | -0.06 | -0.22 |
| pH | $NO_3^-$ | 0.04 | 0.03 |
| | $PO_4^{3-}$ | 0.10 | 0.00 |
| | $NH_4^+$ | -0.36 | -0.57* |

### 2.4.4. Quantifying direct and indirect effects among latent variables biotic, nutrients and physicochemistry

No significant difference was found when comparing the baseline model, which allows direct and indirect effects among all latent variables, to the model where any interaction between nutrients and physicochemistry has been removed (oligotrophic ($\chi^2$ (df = 1) = 0.157); eutrophic ($\chi^2$ (df = 1) = 0.00)). This indicates that the reference hypothesis of additive effects can be accepted.

Comparing the full structural equation model to each of the reduced models, each omitting one of the latent variables, indicated that both nutrients and physicochemistry explained a significant component of variation in microbial dynamics in the experimental aquatic system; this was true for both the oligotrophic and eutrophic treatments (Table 2.2).

Table 2.2. SEM model comparison, reporting outcome of likelihood ratio test between full and reduced models. (A) Full ecosystem model. (B) Reduced model excluding the latent variable nutrients. (C) Reduced model excluding the latent variable water quality.

| Model comparison | Condition | DF difference | $\chi^2$ difference | p-value |
|---|---|---|---|---|
| A vs. B | Oligotrophic | 16 | 71.558 | < 0.001 |
| | Eutrophic | 16 | 73.305 | < 0.001 |
| A vs. C | Oligotrophic | 16 | 46.881 | < 0.001 |
| | Eutrophic | 17 | 76.334 | < 0.001 |

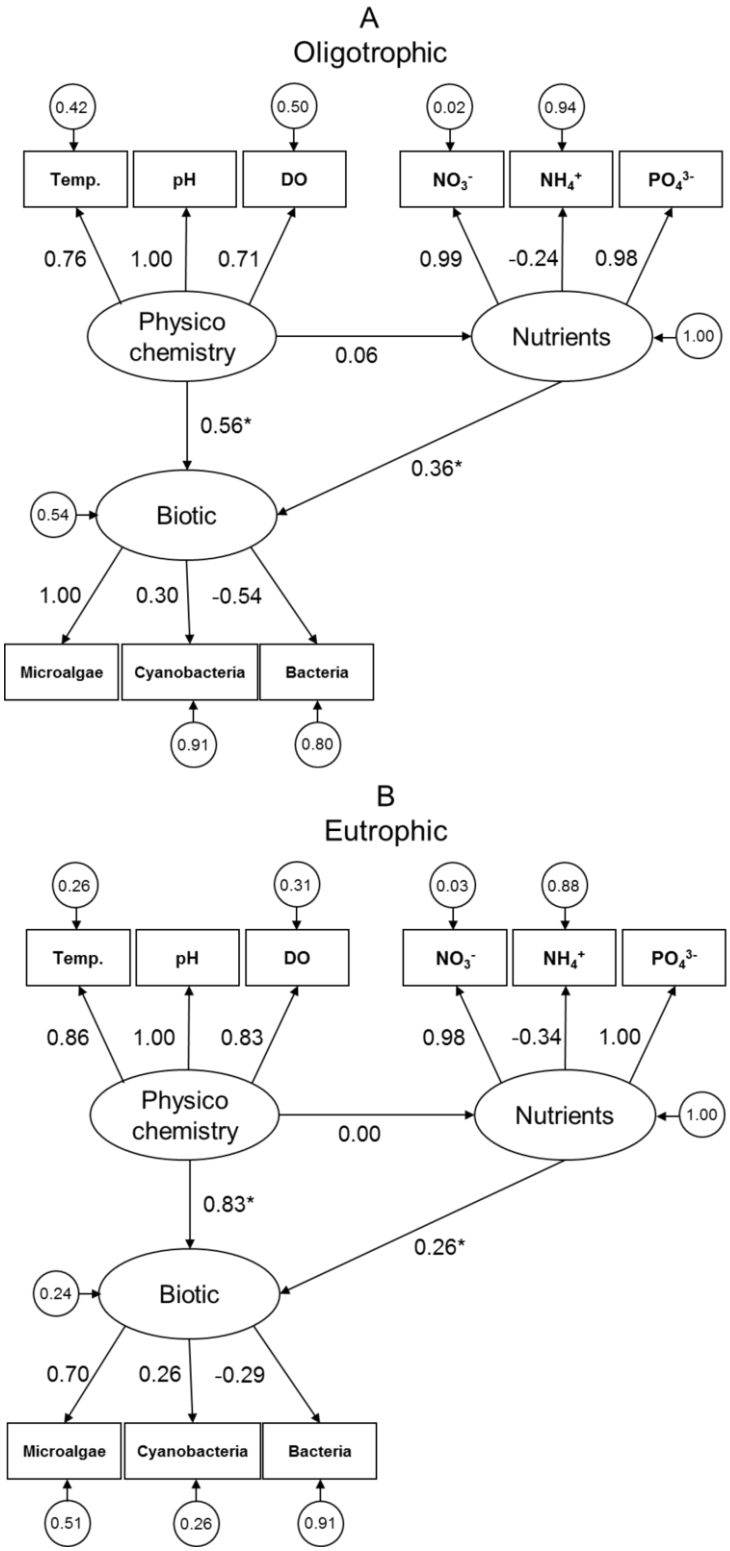


Figure 2.4. Structural equation models showing proposed relationships between latent variables physicochemistry, nutrients and biotic. (A) oligotrophic and (B) eutrophic treatments. The numbers in circles correspond to errors. All other numbers correspond to the standardised path coefficients. The asterisks indicate significant values ($p < 0.05$).

Several patterns emerge from the full model comprising inter-relationships among biotic, nutrients and physicochemistry (Fig. 2.4, asterisks indicate significant values). In both the oligotrophic and eutrophic treatments, the latent variable biotic is more strongly influenced by physicochemistry than by nutrients and the effect of physicochemistry increases in strength with eutrophication ($\rho_{xy}$ (partial correlation coefficient) = 0.56 ($p = 0.046$) in the oligotrophic treatment; $\rho_{xy} = 0.83$ ($p < 0.001$) in the eutrophic treatment). In contrast, the influence of the nutrient variables decrease from $\rho_{xy} = 0.36$ ($p = 0.063$) in the oligotrophic treatment to $\rho_{xy} = 0.26$ ($p = 0.009$) in the eutrophic treatment. In both the oligotrophic and eutrophic treatments, the partial-correlation between physicochemistry and nutrients was not significant ($p = 0.654$ and $p = 0.987$, respectively), indicating independent and thus additive effects of these two variables upon the biological variables and an absence of any indirect effects. Further insights into the role of specific variables arise from examining standardized path coefficients ($\lambda$) of indicator variables within each latent variable. The contribution of $NO_3^-$ and $PO_4^{3-}$ concentrations to the latent variable nutrients remained similar in both the oligotrophic (for $NO_3^-$ and $PO_4^{3-}$ $\lambda = 0.99$ and 0.98, respectively) and eutrophic treatments (for $NO_3^-$ and $PO_4^{3-}$ $\lambda = 0.98$ and 1.00, respectively). DO and temperature become stronger indicators of the latent variable physicochemistry ($\lambda_{DO} = 0.71$ and 0.83 in the oligotrophic and eutrophic treatments, respectively, and $\lambda_{temp} = 0.76$ and 0.86 in the oligotrophic and eutrophic treatments, respectively). Microalgal concentrations have the strongest correlation with the latent variable biotic in both the oligotrophic and eutrophic treatments, but is reduced in the eutrophic treatment ($\lambda = 1.00$ and 0.70, respectively).

## 2.5. Discussion

This microcosm study aimed to determine the relative importance of nutrient concentrations and physicochemical factors in explaining the variation observed in bacterial and microalgal concentrations over time under oligotrophic and eutrophic conditions. Presently, little work has been done to understand how the interconnectivity of abiotic factors is affected by nutrient enrichment and how nutrients and physicochemistry combine to drive ecosystem productivity. Given that changes in microalgal and cyanobacterial productivity are the dominant indicator of eutrophication, such an understanding is important for managing freshwater resources.

Nutrient concentrations and physicochemical factors were found to act additively on the experimental biotic community of eukaryotic microalgae and bacteria (Fig. 2.1), and physicochemical factors exerted a strong influence on both microalgal and bacterial concentrations, which became stronger in eutrophic conditions (Fig. 2.1 B–D). Furthermore, an assessment of the time series in the experimental microcosms indicated that eukaryotic microalgae and cyanobacteria responded to nutrient enrichment through a significant increase in the duration of peak concentrations but not in the peak concentrations themselves. In addition, both microalgal and cyanobacterial concentrations appear to be partially controlled by bacterial processes through the remineralisation and provision of inorganic $NH_4^+$. In the following sections these observations are scrutinised and ecological context is provided by linking them to previous work, with specific comparisons to the multivariate and univariate analyses undertaken in this study.

### 2.5.1. Nutrient enrichment did not affect the experimental variables equally

Peak microalgal concentrations, for both oligo- and eutrophic treatments, were not significantly different (Fig. 2.3 A). However, the eutrophic conditions generated an overall significant increase in the variation of microalgal concentrations over time and extended the duration of peak concentrations for an additional 52 hours (Fig. 2.2 A). While eutrophication did not lead to higher peak concentrations of microalgae, the increased duration of these peak concentrations appears to underpin impact on the overall system.

Observed microalgal and cyanobacterial dynamics were both as expected. Typically, growth is maintained in the presence of nutrients but net growth halts after their depletion (e.g. Fig. 2.2 A,G,H). The nutrient-limited autotrophic populations collapse, which provides organic matter for the heterotrophic bacterial population to degrade (Fig. 2.2 B). Heterotrophic bacteria are mostly stimulated by microalgal derived DOM hence, in a given body of water the peak of bacterial activity tends to follow the peak of primary production (Cole, 1982). This is partly due to the fact that microalgae and cyanobacteria, in an initial instance, outcompete bacteria for the pool of $NO_3^-$ and $PO_4^3$. This observation is supported by the fact that in the oligotrophic treatment, where the concentration of microalgae and cyanobacteria were lower, heterotrophic bacterial concentrations were significantly higher. This also reflects the

view that nutrient limitation increases microalgal exudation thus stimulating bacterial growth (Van den Meersche et al., 2004).

Also, two phases of exponential growth were observed in the variation of microalgal concentrations. The latter exponential growth phase coincides with the appearance of detectable $NH_4^+$ concentrations (Fig. 2.2 I) following the peak of bacterial activity. Microalgal cell lysis, after the nutrient-limited collapse, will release organic N into the medium, which, through the process of ammonification, is converted into $NH_4^+$ by heterotrophic bacteria. This hypothesis is supported by the observation of a longer $NH_4^+$ peak in eutrophic conditions (t = 1.6d; Fig. 2.2 I), which most likely arises from the increase in the microalgal-derived organic matter pool.

These results validate, with practical evidence, the findings from Li et al. (2014) where it was hypothesised that the bacterial component of the microbial loop plays an important role in modulating phytoplankton dynamics. This can occur through direct competition for common resource pools or breakdown of organic matter and recycling of inorganic nutrients. These processes regulate not only the quantity, but also the stoichiometry of N and P that is available to primary producers. Ultimately, the breakdown of organic matter, and recycling of nutrients, can contribute to chronic eutrophication and hypoxic conditions (i.e. “dead zones”).

Cyanobacterial concentrations followed a similar trend as microalgal concentrations. However they were not as sensitive to nutrient enrichment (Fig. 2.2 C). Although the temporal variation of cyanobacterial concentrations is generally higher in the eutrophic treatment there was no significant difference in any of the parameters assessed (Fig. 2.2 C, 2.3 C). This contrasts with previous field based studies where cyanobacteria dominated nutrient enriched environments (Paerl et al., 2008). There are several potential reasons for these results. First, the microcosms lack hydrological characteristics that promote cyanobacterial dominance, such as water column stratification. These characteristics favour cyanobacterial species with buoyancy regulation that can alter their position in the water column to obtain preferable growth conditions (Johnk et al., 2008). Also, temperature was not significantly different across treatments (Fig. 2.2 F) and it has been shown previously that, under enriched conditions, cyanobacterial species are favoured, in detriment of green microalgae, with an increase in water temperature (Kosten et al., 2012; Zhang et al., 2012). In the microcosms the temperatures did not exceed 21.7°C and, therefore, were still in the range facilitating microalgal maximum growth rates (Moss, 1973). Finally, there were higher values of DO and pH in the eutrophic treatments (4% and 1%, respectively) but

they were deemed insignificant (Fig. 2.3 D,E). Nevertheless, higher values are consistent with ecological theory stating that photosynthetic organisms will consume dissolved $CO_2$ (thus increasing pH) and release $O_2$ (thus increasing DO) during periods when photosynthesis dominates. This process is reversed during periods when respiration is the dominating metabolic process, resulting in the DO and pH diel cycles of photosynthetic communities (Talling, 1976). Their fluctuations in the eutrophic treatment mirrors natural conditions where highly productive lakes, as a result of nutrient enriched conditions, have diel cycles of a larger amplitude (Fig. 2.2 D,E) (Nimick et al., 2011).

### 2.5.2. Comparison of observed univariate relationships with previous studies

One of the objectives was to assess whether the abstracted experimental system and agenda of assessing a multivariate hypothesis still generated commonly found univariate relationships among variables. For example, research has indicated that biological variables are expected to have a strong positive correlation with $NO_3^-$ and $PO_4^{3-}$. General consensus is also that the magnitude and duration of the blooms increase with increasing nutrient loads (Cottingham et al., 1998; Elser et al., 2007; Teissier et al., 2011). In line with this, the results of this study show a significant increase in the temporal variation and peak of microalgal concentrations. However the same effect was not observed in cyanobacterial concentrations. Furthermore, in the pairwise analysis of both experimental conditions (Table 2.1), $NO_3^-$ and $PO_4^{3-}$ concentrations did not have a significant correlation with microalgal or cyanobacterial concentrations in either treatments.

There are several potential reasons for these contrasting results. Recent studies have shown that above certain nutrient thresholds (i.e. eutrophic conditions) microalgal concentrations may not correlate strongly with nutrient concentrations (Rigosi et al., 2014). This could be due to the fact that in eutrophic conditions, other environmental parameters (e.g. pH, water depth, temperature) can drive ecosystem processes such as the timing of bloom events, phytoplankton composition and bacterial mineralisation (Paerl et al., 2011; Elliott, 2012; Zhang et al., 2012). For example, recent limnological studies have focused on the strong correlation between cyanobacterial concentrations and temperature (Tadonléke, 2010; De Senerpont Domis et al., 2014). This relationship was confirmed by the pairwise analysis (Table 2.1) where a significant positive

correlation between cyanobacterial abundance and temperature was confirmed in the eutrophic treatment (r = 0.50). In freshwater ecosystems, temperature drives cyanobacterial concentrations directly, through increased growth rates, and indirectly through its influence on hydrological processes. Other characteristics such as high affinity for, and ability to store, P and the ability to fix inorganic N also enables them to have a competitive advantage over green microalgae in changing conditions (Carey et al., 2012).

Finally, based on existing literature it was also predicted that bacterial concentrations would correlate positively with $NO_3^-$ and $PO_4^{3-}$ concentrations (Bird et al., 1984; Haukka et al., 2006). In contrast to this, the results of this study show negative correlation between bacterial concentrations and both $NO_3^-$ and $PO_4^{3-}$ in both oligotrophic and eutrophic treatments. As discussed above, in an initial phase, heterotrophic bacteria are outcompeted by microalgae and cyanobacteria and in latter phases respond to an increase in DOM. Therefore, the peak of bacterial concentrations corresponded to the decline and exhaustion of the microcosm nutrient pool, which, consequently, results in a negative correlation.

### 2.5.3. Utilising SEM to quantify direct and indirect effects among functional biology, nutrients and physicochemistry

SEM provided a robust methodology to assess a substantial multivariate question: is the biological response to increasing nutrient enrichment an additive function of nutrients and physicochemistry, or does the effect of nutrient dynamics on the biotic community depend on physicochemistry (i.e. is there an interaction between them). A baseline model, which allowed the interaction among latent variables was constructed, and compared to one without interactions, leading to an unequivocal conclusion in the experimental system that the effects are additive in both the oligo- and eutrophic communities. There has been an increase in recent studies that have hypothesised that the interaction between nutrient concentrations and physicochemical factors may be synergistic. This opinion may stem from an overrepresentation of univariate studies that focus on eutrophic environments where, as established above, physicochemical factors are better predictors of microalgal and cyanobacterial concentrations (Rigosi et al., 2014).

Furthermore, the independent strength and significance of nutrients and physicochemistry was assessed by comparing this additive model to models without one or the other. This assessment, capturing several components of the nutrient dynamics and physicochemical components, revealed that nutrients and physicochemistry were both necessary to describe the dynamics of the biological variables and their response to enrichment. Finally, in addition to this high level, multivariate assessment of function, the partial correlation coefficients from the SEM allowed a comparison of specific pairwise variable relationships, which have been used extensively in the past.

As noted above, no evidence was found for an interaction between nutrients and physicochemistry. Furthermore, the direct effects (Fig. 2.4 A,B) of nutrients ($\Upsilon$ = 0.36 Oligo; $\Upsilon$ = 0.26 Eutrophic) were much weaker than physicochemistry ($\Upsilon$ = 0.56 Oligo; 0.83 Eutrophic). This generally agrees with published pairwise analysis and supports the ongoing hypothesis that, in eutrophic conditions, nutrients become less effective as a predictor of functional biology.

The application of an SEM based analysis to the multiple, potentially direct and indirect interactions, among several features of aquatic communities provides a holistic approach to understand ecosystem drivers. SEM, in particular, allows for the inclusion of multiple dependent variables and biologically meaningful collections of them, i.e. latent variables, to obtain a better overall picture of the system. Further studies, particularly with higher trophic levels (e.g. zooplankton), hydrological complexity (e.g. vertical mixing, thermal stratification) and environmental gradients (e.g. temperature and nutrients) can contribute to elucidate the network of interactions established in the process of eutrophication.

## 2.6. Conclusion

This study aimed to explore the array of direct and indirect effects among physicochemistry and nutrient concentrations and their influence on the microbial community, which, via productivity and nutrient recycling, define the phenomenon of eutrophication. The results show that the experimental community in this study was sensitive to nutrient enrichment, but only to a certain extent. Interestingly, it was found that the sensitivity of biological groups to both nutrient concentrations and physicochemical factors varied with trophic state. In the experimental eutrophic treatment physicochemical factors became superior predictors of microalgal and

bacterial concentrations. Lastly, in line with recent studies, the model results showed that nutrients and physicochemical factors have an independent but additive effect upon the biotic variables. Microalgal growth is, traditionally, believed to rely on simple cause-effect relationships (e.g. microalgae-nutrients paradigm). However, these relationships are confounded by changes in the physical environment. Microalgal production models need to address events on a short and long term scale and incorporate both nutrient and physicochemical parameters. If growth data can be integrated with data from routinely monitored physicochemical variables, this methodological framework has the potential to, successfully, be transferred from the field of ecology to the field of algal biotechnology.

# Chapter 3: A metaproteomic analysis of the response of a freshwater microbial community under nutrient enrichment

## 3.1. Abstract

Eutrophication can lead to an uncontrollable increase in microalgal biomass, which has repercussions for the entire microbial and pelagic community. Studies have shown how nutrient enrichment affects microbial species succession, however, details regarding the impact on community functionality are rare. Here, we applied a metaproteomic approach to investigate the functional changes to microalgal and bacterial communities, over time, in oligotrophic and eutrophic conditions, in freshwater microcosms. Samples were taken early during microalgal and cyanobacterial dominance and later under bacterial dominance. 1048 proteins, from the two treatments and two timepoints, were identified and quantified by their exponentially modified protein abundance index. In oligotrophic conditions, Bacteroidetes express extracellular hydrolases and TonB-dependent receptors to degrade and transport high molecular weight compounds captured while attached to the phycosphere. Alpha- and Beta-proteobacteria were found to capture different substrates from microalgal exudate (carbohydrates and amino acids, respectively) suggesting resource partitioning to avoid direct competition. In eutrophic conditions, environmental adaptation proteins from cyanobacteria suggested better resilience compared to microalgae in a low C, nutrient enriched environment. This study provides insight into differences in functional microbial processes between oligo- and eutrophic conditions at different timepoints and highlights how primary producers control bacterial resources in freshwater environments.

## 3.2. Introduction

Freshwater ecosystems are subjected to nutrient enrichment on a local, regional and global scale in a process known as eutrophication. Due to human activity, global aquatic fluxes of N and P have been amplified by 108% and 400%, respectively (Falkowski et al., 2000). These nutrient imbalances have led to a drastic increase in the occurrence of microalgal blooms, an event where photoautotrophic biomass may increase by several orders of magnitude (Elser et al., 2007). During a bloom, high amounts of organic C and nutrients are channelled through the bacterial community and made available for higher trophic levels in what is known as the microbial loop. The microbial loop plays a crucial role in the biogeochemical cycling of elements, such as C, P and N, as well as organic matter. It is ultimately responsible for a substantial fraction of aquatic nutrient and energy fluxes (Azam et al., 2007). Thus, a better understanding of how the microbial loop and associated microalgae respond to nutrient enrichment, can reveal important features of how ecosystem processes are affected by eutrophication.

The development and application of "omics" technologies has allowed for an unprecedented view of microbial dynamics and their role in driving ecosystem function, including biogeochemical cycling of elements and decomposition and remineralisation of organic matter. One approach is to obtain and sequence DNA from the microbial community in order to provide access to the genetic diversity of a microbial community (metagenomics). However, the genetic diversity gives us an incomplete view of what role these genes have in community processes. In contrast, metaproteomics can relate the intrinsic metabolic function by linking proteins to specific microbial activities and to specific organisms. Metaproteomics can thus address the long-standing objective in environmental microbiology of linking the identity of organisms comprising diversity in a community to ecosystem function (Hettich et al., 2013).

In the last few years metaproteomics has had a growing influence in aquatic environmental microbiology. It has been used to address questions about diversity, functional redundancy and provision of ecosystem services including nutrient recycling and energy transfer. For example, in one of the metaproteomic pioneering studies Giovannoni et al. (2005) demonstrated the ubiquity of proteorhodopsin-mediated light-driven proton pumps in bacteria. Later, a study by Sowell et al. (2011) was the first of its kind to demonstrate the importance of high affinity transporters for substrate acquisition in marine bacteria. Although most of the notable metaproteomic aquatic

studies have focused on marine environments, the tool has also been used in freshwater environments to examine, for example, the functional metaproteomes from the meromictic lake ecosystem in Antarctica (Ng et al., 2010; Lauro et al., 2011) or the microbes in Cayuga and Oneida Lake, New York (Hanson et al., 2014). The application of metaproteomics in such studies have successfully provided details regarding the importance of bacteriochlorophyll in the adaptation to low light (Ng et al., 2010), the metabolic traits that aid life in cold oligotrophic environments (Lauro et al., 2011) and nutrient cycling, photosynthesis and electron transport in freshwater lakes (Hanson et al., 2014).

In this chapter a comprehensive metaproteomic analysis of a microbial community under differing nutrient regimes is reported. Microcosms were inoculated with a freshwater microbial community subjected to two nutrient treatments to mimic oligotrophic and eutrophic conditions in freshwater lakes. Heterotrophic bacteria, cyanobacteria and microalgae abundances were quantified throughout the experiment as were physicochemical measurements. The microbial metaproteome was extracted from two nutrient treatments (oligotrophic and eutrophic) at two time points. The time points were selected to represent phases of microalgal/cyanobacterial dominance and, later, bacterial dominance. For each treatment the extracted proteome was analysed by nano-liquid chromatography–tandem mass spectrometry (LC-MS/MS). A meta-genetic community analysis of prokaryotic and eukaryotic diversity within the inoculum was used to generate a protein database and identified proteins were then grouped into taxonomic and functional categories to link identity with function. Changes in protein expression were analysed in individual phylogenetic groups, over time and in both nutrient concentrations, to give an insight into the functional attributes of the major microbial players in the experimental microcosm community.

## 3.3. Methods

### 3.3.1. Microcosm setup

Replicate experimental biological communities were constructed in 30 L white, opaque, polypropylene vessels, 42 cm high and with an internal diameter of 31 cm. The microcosms were housed in controlled environment facilities at the Arthur Willis Environmental Centre at the University of Sheffield, U.K. These were filled with 15 L of oligotrophic artificial freshwater growth medium (for detailed composition see Appendix 1). Over the course of the experiment the microcosms were kept at constant temperature, 23˚C, under 100 µmol $m^{-2}$ $s^{-1}$, provided by Hellelamp 400 watt IR Lamps HPS (Helle International Ltd., UK) with a 12:12 light dark cycle. A microbial community was introduced into each microcosm (detailed composition in Appendix 2). This inoculum was sourced from water samples collected at Weston Park Lake, Sheffield (53˚22'56.849'' N, 1˚29'21.235'' W). The inoculum was filtered with a fine mesh cloth (maximum pore size 200 µm) to exclude big particles, protists and grazer populations (Downing et al., 1999). The filtered sample was cultured for five days in the conditions described to allow acclimation to the controlled conditions. Subsequently, each 15 L media was inoculated with 2.5 L of this sample.

### 3.3.2. Sampling of abiotic variables

Over the course of the experiment DO, pH, temperature, $NO_3^-$, $PO_4^{3-}$ and $NH_4^+$ were monitored in order to link the abiotic variation to the changes observed in the biological variables. DO, pH and temperature were measured at 12:00 and 18:00 daily with a Professional Plus Quatro (YSI, USA). 15 mL aliquots were collected and filtered (0.45 µm), daily, for the estimation of $NO_3^-$, $NH_4^+$ and $PO_4^{3-}$ concentrations. A total volume of 270 mL was removed from each microcosm, for nutrient analysis, over the course of the experiment. $NO_3^-$ concentrations were estimated with a Dionex ICS-3000 ion chromatograph (Thermo Fisher Scientific, USA) using an AG18 2x250 mm column with a 0.25 mL $min^{-1}$ flow rate and 31.04 mM potassium hydroxide as eluent. $NH_4^+$ concentrations were estimated with a Dionex ICS-3000 ion chromatograph (Thermo Fisher Scientific, USA) using a CS16 4x250 mm column with a 0.36 mL $min^{-1}$ flow rate and 48 mM methanesulfonic acid as eluent. $PO_4^{3-}$ concentrations were estimated according to the ammonium molybdate spectrometric method defined by British

standard BS EN ISO 6878:2004 (BSI, 2004). Briefly, 10 mL of the filtered sample were mixed with 0.4 mL of 4.5 M sulphuric acid, 1 mL of 0.57 M ascorbic acid and 1 mL of 48.35 mM sodium thiosulphate pentahydrate and allowed to reduce for 10 min. After reduction, 2 mL of an acid molybdate solution (4.14 M sulphuric acid, 21.04 mM ammonium molybdate tetrahydrate and 1.14 mM antimony potassium tartrate hemihydrate) were added and the sample was incubated for 10 min. Finally, the absorbance of each solution was measured at 880 nm against water.

### 3.3.3. Sampling of biotic variables

Chlorophyll *a* and cyanobacterial fluorescence were measured daily, *in situ*, with the MicroalgaeTorch (bbe Moldaenke GmbH, Germany) and utilised as proxies for microalgal and cyanobacterial biomass, respectively. By using specific excitation spectra of fluorescence, 450 nm for chlorophyll *a* and 610 nm for phycocyanin, the two spectral groups of chlorophyll based microalgae and cyanobacteria can be differentiated *in situ*. An internal algorithm is then used to calculate the relative amount of each class that must be present, expressed in terms of the equivalent amount of biomass per litre of water (µg $L^{-1}$) (Beutler et al., 2002).

Total heterotrophic bacteria were measured by using culturable heterotrophic bacteria as a proxy (Salvesen et al., 2000). Heterotrophic bacteria were enumerated every three days by sampling 100 µL aliquots, in triplicate, plating on R2A agar (Oxoid, UK) and incubating for 24 h at 38˚C. A total volume of 1.8 mL was removed from each microcosm, for bacterial counts, over the course of the experiment. Colony forming units (CFU $mL^{-1}$) were counted using OpenCFU software (Geissmann, 2013). Because bacteria were only enumerated every three days, linear interpolation was used to generate a daily time series to obtain a uniform sample size across all variables. Interpolated values were calculated using the formula:

$$y = y_1 + (y_2 - y_1)\frac{x - x_1}{x_2 - x_1}$$

where $y$ is the missing value, $x$ is the missing time point, $y_1$, $y_2$ are the two closest measured bacterial counts and $x_1$, $x_2$ are the respective time points.

### 3.3.4. Experimental design

The experimental elevation of initial nutrient levels followed United States Environmental Protection Agency guidelines for oligotrophic and eutrophic conditions in freshwater lakes and reservoirs (USEPA, 1986): (1) non-enriched growth medium to simulate oligotrophic conditions ($NO_3^-$ = 0.42 mg $L^{-1}$ and $PO_4^{3-}$ = 0.03 mg $L^{-1}$) and (2) $NO_3^-$ and $PO_4^{3-}$ enriched growth medium ($NO_3^-$ = 4.20 mg $L^{-1}$and $PO_4^{3-}$ = 0.31 mg $L^{-1}$) to simulate eutrophic conditions. Each treatment was replicated eighteen times, allowing for serial but replicated (n = 3) destructive sampling during the experiment. The experiment was run for 18 days to allow the added $NO_3^-$ and $PO_4^{3-}$ to deplete and generate batch microbial growth curves (Fig. 2.2 A–C). Three control microcosms comprised of non-enriched growth medium, with no biological inoculum, were also set up in order to follow physicochemical variation in the absence of introduced biological activity (Fig. 2.2 D–I).

### 3.3.5. Protein preparation

Microcosm samples were concentrated, in triplicate, at days three and 12 of the time course using a Centramate tangential flow filtration (TFF) system fitted with three 0.1 µm pore size Supor TFF membranes (Pall Corporation, USA). After every use, the filter system was sanitised with a 0.5 M sodium hydroxide solution and flushed with deionised water. The permeate was then filtered with a 3 µm pore size polycarbonate isopore membrane (EMD Millipore, USA) in order to obtain fractions dominated by free-living bacteria (<3 µm in size) and algae/particle-associated bacteria (>3 µm in size) (Teeling et al., 2012). These fractions were harvested at 10 000 × *g* for 15 minutes at 4°C. The resulting cell pellets were further washed in 0.5 M triethylammonium bicarbonate buffer (TEAB) prior to storage at -20°C. Cells were defrosted and resuspended in extraction buffer (250 µL of 0.5 M TEAB, 0.1% sodium dodecyl sulfate (SDS)) and 1 µL of halt protease inhibitor cocktail (Fisher Scientific, USA)) incorporating a sonication bath step for 5 minutes with ice. The resulting suspension was submitted to five freeze-thaw cycles (each cycle corresponds to two minutes in liquid N and five minutes in a 37°C water bath). The lysed sample was centrifuged at 15,000 × *g* for 10 minutes at 4°C and the supernatant was transferred to a LoBind

microcentrifuge tube (Eppendorf, Germany). The remaining cell pellet was resuspended in extraction buffer (125 µL) and homogenised with glass beads (425 – 600 µm) for ten cycles (each cycle corresponds to two minutes homogenisation and two minutes on ice). The lysed sample was centrifuged at 15,000 × *g* for 10 minutes at 4°C and the supernatants from both extraction methods were combined. 1 µL of benzonase nuclease (Sigma-Aldrich, USA) was added to the collected supernatants. Extracted proteins were precipitated overnight, at -20°C, using four volumes of acetone. The dried protein pellet was resuspended in 100 µL of 0.5 M TEAB and quantified using the 230/260 spectrophotometric assay described by Kalb and Bernlohr (Kalb et al., 1977). Biological replicates were pooled before reduction, alkylation and digestion. This approach has been shown to be potentially valuable for proteomics studies where low amount of protein does not allow replication (Diz et al., 2009) whilst enhancing the opportunity to identify lower abundance proteins. Moreover, the small variances observed between replicate microcosms in terms of all biological and physiochemical measurements conducted (Fig. 2.2) gave further confidence to this approach. Protein samples (200 µg) were reduced with 20 mM tris-(2-carboxyethyl)-phosphine, at 60 °C for 30 minutes, followed by alkylation with 10 mM iodoacetamide for 30 minutes in the dark. Samples were digested overnight, at 37 °C, using trypsin (Promega, UK) 1:40 (trypsin to protein ratio) resuspended in 1 mM HCl. The samples were dried using a vacuum concentrator and stored at -20°C prior to fractionation.

### 3.3.6. Chromatography and mass spectrometry

The first dimensional chromatographic separation, off-line, was performed on a Hypercarb porous graphitic column (particle size: 3 µm, length: 50 mm, diameter: 2.1 mm, pore size: 5 µm) (Thermo-Dionex, USA) on an Ultimate 3000 UHPLC (Thermo-Dionex, USA). Peptides were resuspended in 200 µL of Buffer A (0.1% (v/v) trifluoroacetic acid (TFA) and 3% (v/v) HPLC-grade acetonitrile (ACN) in HPLC-grade water) and eluted using a linear gradient of Buffer B (0.1% (v/v) TFA and 97% (v/v) ACN in HPLC-grade water) ranging from 5 to 60% over 120 minutes with a flow rate of 0.2 mL $min^{-1}$. Peptide elution was monitored at a wavelength of 214 nm and with Chromeleon software, version 6.8 (Thermo-Dionex, USA). Fractions were collected every two minutes, between 10 and 120 minutes, using a Foxy Junior (Teledyne Isco, USA) fraction collector and dried using a vacuum concentrator. Dried fractions were

stored at -20°C prior to mass spectrometry analysis. The second dimensional chromatographic separation of each peptide fraction was performed on a nano-LC-CSI-MS/MS system. In this system a U3000 RSLCnano LC (Thermo-Dionex, USA), containing a trap column (300 µm × 5 mm packed with PepMap C18, 5 µm, 100Å wide pore, Dionex) followed by a reverse phase nano-column (75 µm × 150 mm packed with PepMap C18, 2 µm, 100Å wide pore, Dionex), was coupled to an ultra-high resolution quadrupole time-of-flight (UHR maXis Q-ToF 3G) mass spectrometer (Bruker, Germany) equipped with an Advance CaptiveSpray ion source. Peptide fractions were resuspended in loading buffer (0.1% (v/v) TFA and 3% (v/v) ACN in HPLC-grade water) and two injections were made. A 90 minute linear gradient elution was performed using buffer A (0.1% (v/v) formic acid (FA) and 3% (v/v) ACN in HPLC-grade water) and buffer B (0.1% (v/v) FA and 97% (v/v) ACN in HPLC-grade water), during which buffer B increased from 4 to 40% at a flow rate of 0.3 µL $min^{-1}$. On the mass spectrometer, the following settings were specified: endplate Offset -500 V, capillary voltage 1000 V, nebuliser gas 0.4 bar, dry gas 6.0 L $min^{-1}$, and dry temperature 150 °C. Mass range: 50-2200 *m/z*, at 4 Hz. Lock mass was used for enabling mass acquisition correction in real time, therefore high mass accuracy data were obtained. Data were acquired for positive ions in a dependent acquisition mode with the three most intense double, triple or quadruple charges species selected for further analysis by tandem mass spectrometry (MS/MS) under collision induced dissociation (CID) conditions where N was used as collision gas.

### 3.3.7. 16S and 18S rDNA gene sequencing of inoculum

#### 3.3.7.1. DNA extraction

Inoculum samples were lysed in 50 mM Tris-HCl (pH 8.0), 10 mM EDTA and 10% (w/v) SDS by vortexing with glass beads. DNA was extracted with a standard phenol-chloroform extraction protocol (Sambrook et al., 2001). The DNA was precipitated using sodium acetate (50 µL of 3 M stock solution, pH 4.8-5.2) and ice-cold ethanol. PCR amplification, product pooling, purification sequencing and bioinformatics and statistical analysis were performed by Research and Testing Laboratory (Texas, USA).

#### 3.3.7.2. PCR amplification

Markers were amplified from DNA extractions using adapted Illumina tagged primers. Forward primers were constructed with Illumina adapter i5 (AATGATACGGCGACCACCGAGATCTACAC) an 8-10bp barcode, a primer pad and either primer 28F (GAGTTTGATCNTGGCTCAG) or TAReukF (CCAGCASCYGCGGTAATTCC). Reverse primers were constructed with Illumina adapter i7 (CAAGCAGAAGACGGCATACGAGAT) an 8-10bp barcode, a primer pad and either primer 519R (GTNTTACNGCGGCKGCTG) or TAReukR (ACTTTCGTTCTTGATYRA). Primer pads were used to ensure a primer melting temperature of 63°C-66°C, as per the Schloss method (Schloss et al., 2009). Reactions were performed using corresponding primer pairs (i.e. 28F x 519R and TAReukF x TAReukR) using the Qiagen HotStar Taq master mix (Qiagen Inc, Valencia, California) adding 1 μL of each 5 μM primer, and 1 μL of template to make a final 25 μL reaction volume, with a thermal cycling profile of 95°C for 5 minutes, then 35 cycles of 94°C for 30 seconds, 54°C for 40 sec., 72°C for 1 minute., followed by one cycle of 72°C for 10 minutes. Amplified products were visualised with eGels (Life Technologies, Grand Island, New York) and pooled. Pools were purified (size selected) through two rounds of 0.7x Agencourt AMPure XP (BeckmanCoulter, Indianapolis, Indiana) as per manufacturer's instructions, before quantification with a Quibit 2.0 fluorometer (Life Technologies). Finally pools were loaded and sequenced on an Illumina MiSeq (Illumina, Inc. San Diego, California) 2x300 flow cell at 10 pM. The sequence data are available from the European Nucleotide Archive under Study Accession Number PRJEB12443, and Sample Accession Numbers ERS1037123 (16S DNA) and ERS1037124 (18S DNA).

#### 3.3.7.3. Bioinformatic and statistical analysis

Initially the forward and reverse reads were taken and merged together using the PEAR Illumina paired-end read merger (Zhang et al., 2014). Reads were then filtered for quality by trimming them once average quality dropped below 25 and prefix dereplication was performed using the USEARCH algorithm (Edgar, 2010). Sequences below 100bp were not written to the output file and no minimum cluster size restriction

was applied. Clustering was performed at a 4% divergence using the USEARCH clustering algorithm (Edgar, 2010). Clusters containing less than 2 members were removed. OTU selection was performed using the UPARSE OTU selection algorithm (Edgar, 2013). Chimeras were then checked for and removed from the selected OTUs using the UCHIME chimera detection software executed in *de novo* mode (Edgar et al., 2011). Reads were then mapped to their corresponding nonchimeric cluster using the USEARCH global alignment algorithm (Edgar, 2010). The denoised sequences were demultiplexed and the primer sequences removed. These sequences were then clustered into OTUs using the UPARSE algorithm (Edgar, 2013) which assigns each of the original reads back to their OTUs and writes the mapping data to an OTU table file. The centroid sequence from each OTU cluster was then run against the USEARCH global alignment algorithm and the taxonomic identification was done using a NCBI database as described in Bokulich, et al. (2015). Finally, the OTU table output from sequence clustering was collated with the output generated during taxonomic identification and a new OTU table with the taxonomic information tied to each cluster was created (Bokulich et al., 2015).

### 3.3.8. Protein identification and quantification

All MS and MS/MS raw spectra were processed using Data Analysis 4.1 software (Bruker, Germany) and the spectra from each Bruker analysis file were output as a mascot generic file (MGF) for subsequent database searches using Mascot Daemon (version 2.5.1, Matrix Science, USA). The peptide spectra were searched against a eukaryotic and a prokaryotic database created by collating all Uniprot entries (retrieved on 24 February 2015) from organisms with an abundance of > 1% in the 16 and 18S rDNA survey of our inoculum (Table 3.1, full list in Appendix 2). This search was undertaken utilising the two-step approach described in Jagtap et al. (2013). Briefly, the initial database search was done without any false discovery rate (FDR) limitation and then was followed by a second search with a 1% FDR threshold against a refined database created by extracting the protein identifications derived from the first search. FDRs for assigning a peptide match were determined from the ratio of the number of peptides that matched to the reversed sequence eukaryotic and prokaryotic databases to the number of peptides matched to the same databases in the forward sequence direction. The following search parameters were applied to both searches: up to one

missed cleavage with trypsin, fixed modification of cysteine residues by carbamidomethylation, variable modification of methionine by oxidation, instrument specification ESI Q-ToF, peptide charge: 2+, 3+ and 4+, precursor mass tolerance of ±0.2 Da and fragment-ion mass tolerance of ±0.02 Da. For the second search only matches above a 95% confidence homology threshold, with significant scores defined by Mascot probability analysis, and a 1% FDR cut-off were considered confidently matched peptides. 'Show sub-sets' and 'require bold red' were applied on initial Mascot results to eliminate redundancy. The highest score for a given peptide mass (best match to that predicted in the database) was used to identify proteins, which in turn were assigned a most probable host. Furthermore, only when two or more unique peptides, per protein, were matched did we consider a protein identified. Protein abundance was relatively estimated through the exponentially modified protein abundance index (emPAI) (Ishihama et al., 2005). emPAI is an approximate, label-free, relative quantitation of the proteins. This method is based on the protein abundance index (PAI) that calculates the number of different observed modified peptides divided by the number of observable peptides as a measure of abundance. This PAI value is then exponentially modified to derive the emPAI score. A protein abundance is then finally calculated after normalising the emPAI score for a protein by dividing it by the sum of the emPAI scores for all identified proteins (Ishihama et al., 2005).

## 3.4. Results and Discussion

### 3.4.1. Biological and physicochemical measurements

The time points chosen for metaproteomic analysis of our samples were based on biological and physicochemical variables measured in our microcosms. Microalgal and cyanobacterial abundance peaked at day three, and was maintained until $NO_3^-$ and $PO_4^{3-}$ concentrations were no longer in detectable range, but declined after their depletion between days six and eight (Fig. 2.2). The decline in abundance of microalgae and cyanobacteria was followed by a peak of bacterial abundance at day 12 (Fig. 2.2). Heterotrophic bacterial growth is known to be stimulated by an accumulation of DOM derived from senescent microalgae and cyanobacteria. Hence, in a given body of water the peak of heterotrophic bacterial activity tends to follow the peak of primary production.

Based on these patterns the samples selected for metaproteomic analyses were harvested at day three, the peak of microalgal and cyanobacterial concentrations, (early oligo- and eutrophic) and day 12, the peak of bacterial concentrations, (late oligo- and eutrophic). The comparative analysis of these biologically distinct time points can provide information regarding the activity of the microbial community during microalgal/cyanobacterial dominance and heterotrophic bacterial dominance under low and high nutrient conditions.

Similar patterns were observed in DO, pH and temperature measurements in both nutrient treatments and together with the low level of variation observed in biological measurements (Fig. 2.2), provided additional confidence in the sample pooling approach for metaproteomics analyses.

### 3.4.2. Metaproteomic database creation and search results

The 18S rDNA sequencing of the microcosm inoculum indicated that, at day 0, the eukaryotic community was predominantly composed of Chlorophyceae (e.g. *Chloromonas*) and Bacillariophyceae (e.g. *Stephanodiscus*) and Chrysophyceae (e.g. Chromulinaceae). These are typical unicellular freshwater eukaryotic microalgal species that are normally found in freshwater oligotrophic environments (Bailey-Watts, 1992).

The 16S rDNA sequencing of the microcosm inoculum showed that, at day 0, the prokaryotic community was predominantly composed of Alpha-proteobacteria (e.g. Brevundimonas), Beta-proteobacteria (e.g. *Rhodoferax*), Flaviobacteria (e.g. Flaviobacterium) and Cyanophyceae (e.g. *Anabaena*). Proteobacteria and Flaviobacteria are ubiquitous in freshwater environments with the latter being known to dominate eutrophic environments where microalgal population numbers are high (Eiler et al., 2007; Newton et al., 2011). *Anabaena* is a well-researched freshwater cyanobacterium that is known to occasionally be responsible for harmful microalgal blooms (Elser et al., 2007).

Table 3.1. List of the eukaryotic and prokaryotic organisms in the experimental freshwater microbial community inoculum, with an abundance higher than 1%, as determined by 16 and 18S rDNA sequencing. These organisms were used to guide creation of a protein database.

| Eukaryotic organisms | % | Prokaryotic organisms | % |
|---|---|---|---|
| *Chloromonas pseudoplatyrhyncha* | 26.93 | *Rhodoferax sp* | 21.94 |
| *Stephanodiscus* sp. | 18.17 | Unsequenced organisms | 17.84 |
| Unsequenced organisms | 17.87 | *Flavobacterium* sp. | 9.43 |
| *Chromulinaceae* sp. | 8.48 | *Anabaena* sp. | 8.85 |
| *Synedra angustissima* | 4.99 | *Brevundimonas diminuta* | 4.20 |
| *Ochromonadales* sp. | 3.01 | *Hydrogenophaga* sp. | 3.41 |
| *Chlamydomonas* sp. | 2.98 | *Runella limosa* | 2.47 |
| *Micractinium pusillum* | 1.62 | *Haliscomenobacter* sp. | 2.43 |
| *Chlorella* sp. | 1.08 | *Rhodobacter* sp. | 2.34 |
| *Pythiaceae* sp. | 1.07 | *Planktophila limnetica* | 2.13 |
| | | *Agrobacterium tumefaciens* | 2.11 |
| | | *Sphingobacterium* sp. | 2.03 |
| | | *Ochrobactrum tritici* | 1.98 |
| | | *Brevundimonas variabilis* | 1.83 |
| | | *Sphingomonas* sp. | 1.73 |
| | | *Curvibacter* sp. | 1.48 |
| | | *Phenylobacterium falsum* | 1.42 |
| | | *Roseomonas stagni* | 1.24 |
| | | *Oceanicaulis* sp. | 1.03 |

This list of organisms was utilised to create a eukaryotic and a prokaryotic protein database by collating all Uniprot entries from organisms with an abundance of > 1% in the 16 and 18S rDNA survey of the inoculum (Table 3.1; full list in Appendix 2). This approach was applied to limit the size of the resulting protein databases, which can lead to high false positive rates, and also in accordance with the nature of mass spectrometry based proteomics, where only the most abundant proteins are identified. As a result the eukaryotic and prokaryotic databases contained 86336 and 350356 sequence entries, respectively. These databases were utilised to identify proteins from peptide fragments in a two-step approach (Jagtap et al., 2013). This approach is

valuable when dealing with large metaproteomic database searches where the target and decoy identifications may overlap significantly and valuable identifications are missed out (Muth et al., 2015). Proteins of eight samples, representing the two time points selected under different nutrient concentrations (early and late oligo- and eutrophic) and two size separated fractions, (free-living bacteria (<3 µm in size) and microalgae/particle-associated bacteria (>3 µm in size)) were identified and an average of 131 ± 28 proteins, above a 95% confidence homology threshold, a 1% FDR cut-off and with two unique peptides, were identified per sample. Values were pooled by broad protein annotation and taxonomic categories to evaluate differences between early and late oligotrophic and eutrophic conditions.

### 3.4.3. Phylogenetic diversity according to the metaproteomic spectra

Identifying discrepancies between the phylogenetic classification of the identified proteins and the 16 and 18S rDNA sequencing used to create the metaproteomic database can indicate if any specific phylogenetic group is inadequately represented. rDNA sequencing was performed on the inoculum (i.e. at day 0 of the experiment) and therefore a direct comparison with the metaproteomes is not possible. Nevertheless, the 16 and 18S rDNA sequencing information provided a template to which the metaproteome could be compared. Proteins identified in the >3 µm fraction, across all four samples, were matched to three phyla: Chlorophyta (48-55%) followed by Heterokontophyta (9-27%) and finally Cyanobacteria (12-33%).

A more detailed look at the genus level of the phylogenetic distribution showed that *Chlamydomonas* sp. proteins are most abundant in the early part of the time series (oligotrophic (50%) and eutrophic (43%)), *Chlorella* sp. in late oligotrophic (39%) and *Anabaena* sp. in late eutrophic (37%) conditions. Generally, the taxonomic composition of the 18S rDNA and the metaproteome were in agreement. One notable exception was that identified proteins belonging to the *Chlorella* genus represented up to 39% of total protein while the 18S rDNA sequencing indicated only 1.08% of the initial inoculum. This is most likely due to an over representation of *Chlorella* sp. in the metaproteomic database due to it being a model genus with a large number of sequences available in Uniprot.

The phylogenetic distribution, based on proteins identified across the samples, mostly fitted with the biological measurements. Eukaryotic microalgae concentrations

were always higher than cyanobacteria concentrations over the course of the experiment (Fig. 2.2). Cyanobacteria were the most represented in late eutrophic (37%) due to the expression of highly abundant proteins related to C concentration mechanisms. It has been suggested that this is a mechanism of survival under adverse conditions that could, in the long term, favour cyanobacterial populations (Yeates et al., 2008).

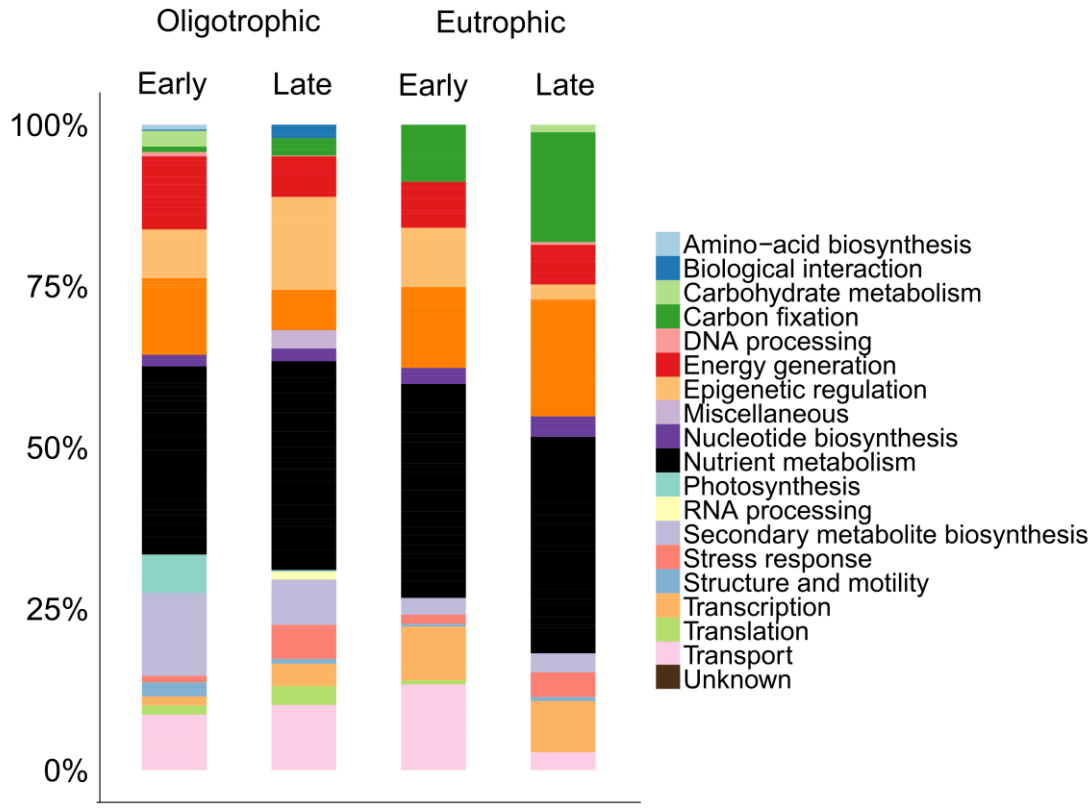


Figure 3.1. Comparison across samples of the distribution of identified proteins by their functional classification in the >3 µm fraction.

Proteins identified in the <3 µm fraction, across all four samples, were matched to two phyla (Fig. 3.1): Proteobacteria (60-73%) and Bacteroidetes (27-40%). Bacteroidetes proteins were more abundant in early oligotrophic conditions whereas Proteobacteria were more abundant in late oligotrophic and early eutrophic conditions. A more detailed look at the class level of the phylogenetic distribution showed that Flaviobacteria proteins were more abundant in the early phase (oligotrophic (29%) and eutrophic (30%)) while Alpha-proteobacteria were abundant in late oligotrophic (30%) and Beta-proteobacteria in late eutrophic (30%) conditions.

Again, the taxonomic community composition found by 16S rDNA sequencing and the metaproteome were in agreement and the phylogenetic distribution across the samples supports previous observations of these organisms. Flaviobacteria typically establish mutualistic relationships with microalgae on the cell surface and are more abundant when microalgal concentrations are high such as earlier in the time series (Fig. 2.2 A). Alpha-proteobacteria and Beta-proteobacteria, as opportunistic heterotrophs, therefore thrive in the presence of DOM derived from microalgal and cyanobacterial decay which was abundant later in the time series (Fig. 2.2 A,B) (Teeling et al., 2012).

### 3.4.4. Functional classification of proteins

The distribution of identified proteins by their functional classification was done utilising UniprotKB (http://www.uniprot.org/), Gene Ontology (http://geneontology.org/), Pfam database (http://pfam.xfam.org/) and EggNOG (http://eggnogdb.embl.de/) resulting in 20 distinct functional categories. The grouping of proteins identified in each fraction and nutrient condition can give an overview of how the community function differed over time and nutrient enrichment.

Overall, proteins involved in photosynthesis (25%) dominated the >3 µm fraction (Fig. 3.1). 9% of the total protein library were classified with unknown function. Proteins with assigned functions in each individual samples were dominated by photosynthesis (early oligotrophic, 21%; late oligotrophic, 25%; early eutrophic, 26%; late eutrophic, 30%). On the individual protein level, photosystem II (PSII) CP43 reaction centre proteins were the most abundant in early oligotrophic (8%), histone H2 proteins in late oligotrophic (14%), PSII CP43 reaction centre proteins and histone H4 proteins (8% each) in early eutrophic and microcompartment proteins (16%) in late eutrophic conditions.

In agreement with our findings, Hanson et al. (2014) observed that in both freshwater and marine surface samples (i.e. rich in primary production) there was widespread evidence of photosynthesis (e.g. PSII) and C fixation (e.g. ribulose-1,5-bisphosphate carboxylase oxygenase (RuBisCO). Although our samples were not rich in RuBisCO, the presence of microcompartment proteins are evidence of C fixation.

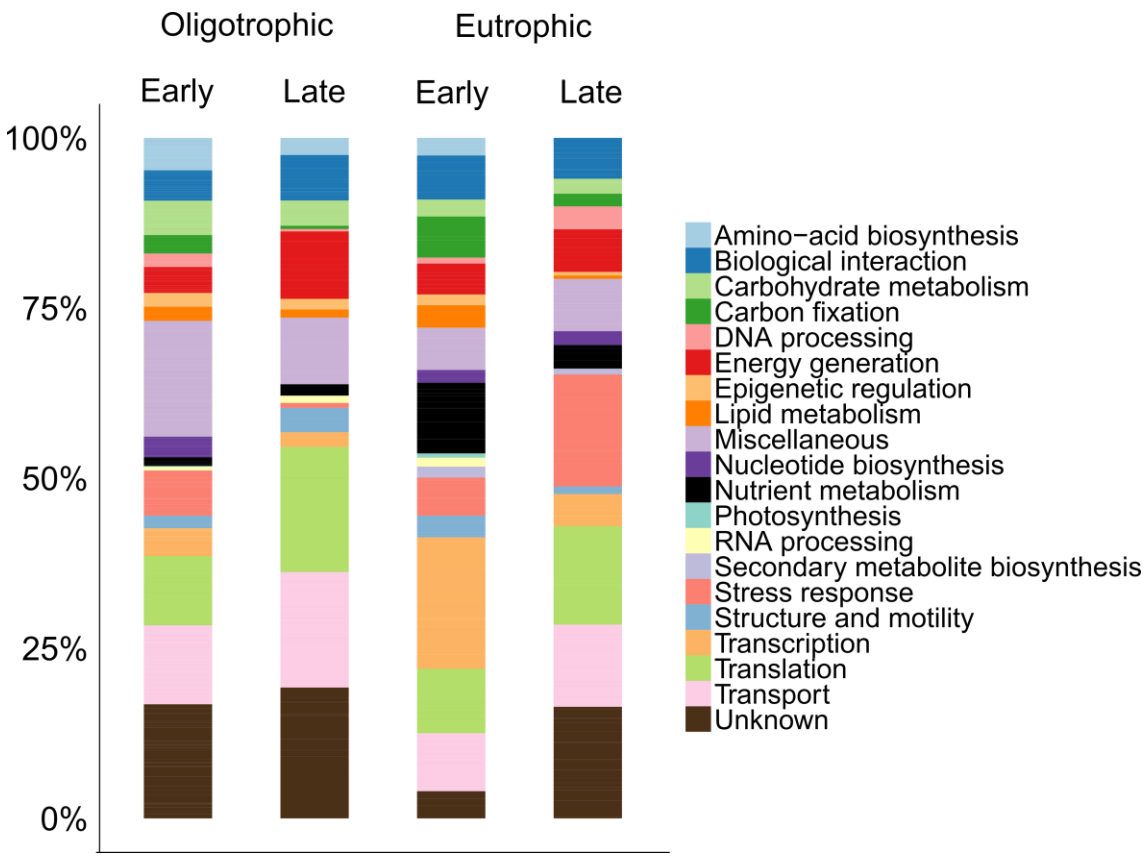


Figure 3.2. Comparison across samples of the distribution of identified proteins by their functional classification in the <3 µm fraction.

Transport (12%) and translation (12%) proteins were predominant in the <3 µm fraction (Fig. 3.2). A more detailed view showed early oligotrophic conditions dominated by transport proteins (12%), late oligotrophic by translation proteins (18%), early eutrophic by transcription proteins (19%) and late eutrophic by stress response proteins (16%). On the individual protein level the ABC transporter proteins were the most abundant in early oligotrophic (5%), elongation factor proteins in late oligotrophic (16%), DNA-directed RNA polymerase subunit beta in early eutrophic (19%) and ABC transporter proteins (7%) in late eutrophic conditions. Proteins involved in transport (e.g. ABC transporters), translation (e.g. elongation factors) and transcription (DNA-directed RNA polymerase subunit beta) are amongst the most commonly identified proteins in environmental samples (Ng et al., 2010; Sowell et al., 2011; Hanson et al., 2014).

### 3.4.5. Metaproteomic analysis of microcosm microbial activity

Having identified protein functional groups in eukaryotic and prokaryotic organisms throughout our samples, we can now assess functional differences between oligotrophic and eutrophic conditions, early and late in the time series. We found several patterns previously documented and several unexpected differences between time points and between oligotrophic and eutrophic conditions within each time point. Fig. 3.3 captures a summary of the functional differences among the times and treatments, and we now refer to this figure, and Figs. 3.1 and 3.2, to provide detail.

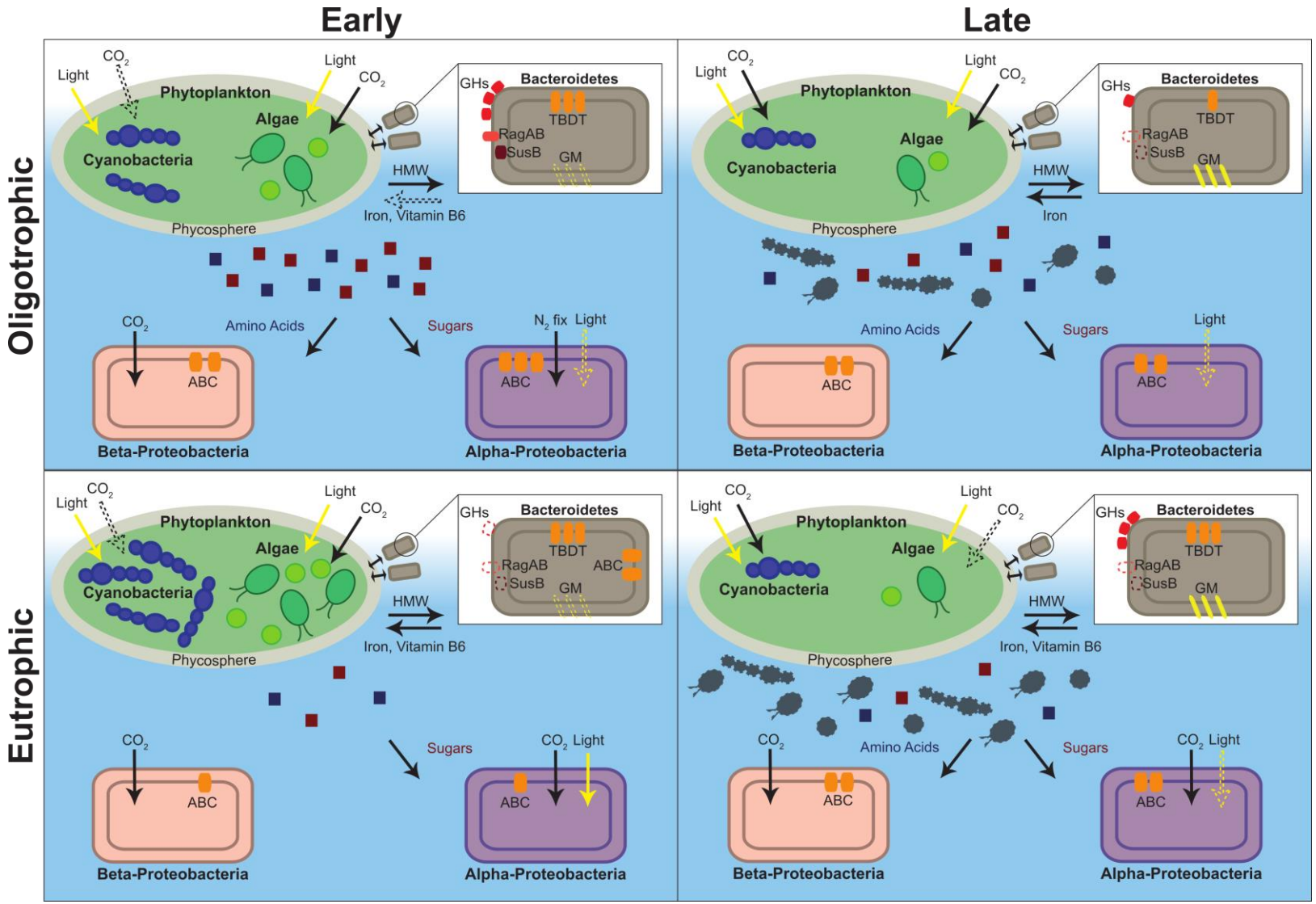


Figure 3.3. Depiction of the metabolic characteristics of oligotrophic and eutrophic communities inferred from the metaproteome.

Red and blue squares depict microalgal and cyanobacterial exudate (red, sugars; blue, amino acids). Grey microalgae and cyanobacteria depict senescent cells. Structures and processes that are hypothesised to be present, albeit with no direct evidence from our dataset, are depicted with a dashed line. ABC, ATP-binding cassette transporter; GHs, glycoside hydrolases; GM, gliding motility; HMW, high molecular weight compounds; N2 fix, N fixation; TBDT, Ton-B-dependent transporter.

First, virtually all the photosynthesis and C fixation proteins were identified in *Anabaena* sp., *Chlamydomonas* sp. and *Chlorella* sp. This is similar to previous metaproteomic studies where the freshwater surface is typically rich in photosynthetic organisms (Hanson et al., 2014). The most abundant of the two categories was photosynthesis (emPAI = 11.89) and it represented 40% of all proteins expressed by photoautotrophic organisms. The majority of the proteins were components of PSII (e.g. reaction centre components). This was expected because PSII proteins are 40% to 90% more abundant than PSI proteins and are the most abundant membrane proteins in microalgae and cyanobacteria (Nobel, 2005). Photosynthetic proteins were abundant in both timepoints (early, emPAI = 5.27 and late, emPAI = 5.61) and in both nutrient treatments (oligotrophic, emPAI = 5.31 and eutrophic, emPAI = 5.57), suggesting that the phototrophs are demanding a constant energy supply, even outside of the exponential growth phase.

Second, amongst the photosynthetic microbes, there is interest in identifying mechanisms that could potentially favour cyanobacteria in eutrophic conditions. The increase in the number of nutrient enriched water bodies has led to issues with freshwater quality and the proliferation of harmful cyanobacteria (O'Neil et al., 2012). There have been numerous proteomics studies of toxic bloom causing cyanobacteria that have focused on the molecular mechanisms of pure cultures. For example, a study of the proteomes of six toxic and nontoxic strains of *Microcystis aeruginosa* linked N regulation to toxicity (Alexova et al., 2011) and another study, of *Anabaena* sp. Strain 90, linked P starvation to the down regulation of the Calvin cycle and amino-acid biosynthesis (Teikari et al., 2015). Studies such as these provide valuable information regarding species in isolation, however, metaproteomics can go a step further and contextualise these findings within the microbial community structure and dynamics.

Our microcosm data showed that pigment proteins in *Anabaena* sp. were less abundant in oligotrophic than in eutrophic conditions (oligotrophic, emPAI = 0.42; eutrophic, emPAI = 0.96). A similar pattern was found for cyanobacterial proteins with roles in C fixation (oligotrophic, emPAI = 0.14; eutrophic, emPAI = 2.03). Cyanobacteria have the ability to adapt to different environments by adjusting their light harvesting abilities (i.e. increase in pigments) and C fixation mechanisms. However, these adaptation processes can be hampered by insufficient nutrient supply (Tilzer, 1987). Grossman et al. (1993) showed that during nutrient starvation, there is a rapid degradation of the phycobilisome. Phycobilisome degradation can provide nutrient-

starved cells with amino acids used for the synthesis of proteins important for their metabolism (Grossman et al., 1993). This suggests that nutrient enrichment would allow cyanobacteria to increase pigment numbers, thus increasing light harvesting ability, and outcompete microalgal species in eutrophic conditions (Tilzer, 1987).

Regarding C fixation, microcompartment proteins were identified in *Anabaena* sp. and were only found in in late eutrophic conditions (eutrophic, emPAI = 1.52). Microcompartments sequester specific proteins in prokaryotic cells and are involved in carbon concentrating mechanisms (CCMs) in low $CO_2$ conditions. The carboxysome, a bacterial microcompartment that is found in cyanobacteria and some chemoautotrophs, encapsulates RuBisCO and carbonic anhydrase. The carbonic anhydrase reversibly catalyses the conversion of bicarbonate into $CO_2$ within the carboxysome therefore acting both as a intracellular equilibrator and a $CO_2$ concentrating mechanism (Yeates et al., 2008). However, no carbonic anhydrases were identified in our dataset. A higher abundance of C fixation proteins in *Anabaena* sp., in eutrophic conditions, indicates that C requirement was higher, likely matching higher photosynthesis rates compared to the oligotrophic conditions, where low N and P concentrations are likely limiting factors and therefore, not allowing the population to reach a point of C limitation.

Finally, C fixation proteins in *Chlamydomonas* sp. were also more abundant in eutrophic conditions (oligotrophic, emPAI = 0.17; eutrophic, emPAI = 0.40). The proteins identified were mainly involved in the Calvin cycle (i.e. RuBisCO), however, unexpectedly, a low-$CO_2$ inducible protein (LCIB) was identified. The LCIB is located around the pyrenoid and traps $CO_2$, either from escaping from the pyrenoid or entering from outside the cell, into the stromal bicarbonate pool thus, functioning as a CCM (Wang et al., 2014). Wang and Spalding hypothesised that this system may reflect a versatile regulatory mechanism present in eukaryotic microalgae for acclimating quickly to changes in $CO_2$ availability that frequently occur in their natural environments. The possibility of switching between an energy-intensive bicarbonate transport system (low $CO_2$) and diffusion based $CO_2$ uptake system (high $CO_2$) that may be energetically less costly, would enable faster growth at a lower energy cost.

These observations suggest that microalgae and cyanobacteria both adapt to C limitation through an increase in C fixation proteins and the deployment of CCMs (e.g. carboxysomes). In a low-C lake, the microbial population may thus fix atmospheric $CO_2$ to correct the C deficiency and grow in proportion to existing N and P levels. This maps onto the hypothesis that C limitation may not be adequate for microalgal or cyanobacterial bloom mitigation (Schindler et al., 2008).

### 3.4.5.1. Bacterial photosynthesis and carbon fixation

Heterotrophic bacteria are known to be responsible for the bulk of sequestration and remineralisation of organic matter in bacterial assemblages associated with autotrophic organisms (Buchan et al., 2014). However, the role of photoheterotrophic and chemoautotrophic bacteria in these assemblages, and how they vary along environmental gradients, remains under-studied (Yutin et al., 2007; Ng et al., 2010). The observations to date suggest that these bacteria are ubiquitous but have a preference for C-limiting environments such as the DOM poor conditions found early in the time series, during microalgal and cyanobacterial dominance, in this study (Fig. 3.3).

In support of this hypothesis, bacterial photosynthesis (i.e. magnesium chelatase) and C fixation proteins (i.e., RuBisCO, carbonic anhydrase) were identified in in both treatments (Fig. 3.3) with predominance early in the time series (early, emPAI = 1.28; late, emPAI = 0.11) and eutrophic conditions (oligotrophic, emPAI = 0.57; eutrophic, emPAI = 0.82). Specifically, in Alpha- and Beta-proteobacteria, magnesium chelatase (emPAI = 0.03), which is involved in bacteriochlorophyll biosynthesis, was identified in early oligotrophic (emPAI = 0.03) and RuBisCO was present in both nutrient treatments. Alpha- and Beta-proteobacteria include several mixotrophic species that are known to perform aerobic and anaerobic respiration and use combinations of photo-, chemo-, auto- and heterotrophic metabolism to adapt to different environmental conditions. Some of these bacterial species perform anoxygenic photosynthesis, where light energy is captured and converted to ATP without the production of oxygen, and are described as photo(chemo)heterotrophs due to their requirement of organic C. It has been suggested that these bacteria grow chemoheterotrophically but utilise light as an additional energy source (Eiler, 2006).

The low levels of DOM in early oligotrophic conditions (i.e. microalgal and cyanobacterial dominance) provided a niche for phototrophy and autotrophy. Later, in the presence of DOM derived from microalgal and cyanobacterial cell lysis, the bacterial groups changed to a heterotrophic metabolism. This suggests that an increase in Proteobacterial metabolism depends more on the concentrations of DOM than on N and P, and that bacterial mixotrophy is ubiquitous in low DOM freshwater environments. This has consequences for biogeochemical models such as the microbial loop. The classic separation of primary and secondary producers into photoautotrophs

and organoheterotrophs, respectively, is no longer valid and may lead to the underestimation of bacterial biomass production and their importance to higher trophic levels (Eiler, 2006).

Finally, other bacterial groups found in our study, such as the Bacteroidetes, can also use non-photosynthetic routes of light-dependent energy generation. Previous metaproteomic studies have shown that proteorhodopsin, a light driven proton pump, is ubiquitous in marine and freshwater environments (Atamna-Ismaeel et al., 2008; Williams et al., 2013). Its expression has been linked to survival in situations where sources of energy are limiting and cells have to resort to alternative means of generating energy (González et al., 2008). However, proteorhodopsin was not detected either because of non-expression in the conditions tested, low abundance or low solubility of the protein; proteorhodopsin contains seven transmembrane helices and is imbedded in the plasma membrane thus making it difficult to solubilise and detect (Sowell et al., 2009).

#### 3.4.5.2. Bacteroidetes: A microalgal associated bacterial group

The Bacteroidetes phylum has been hypothesised to specialise in degrading high molecular weight (HMW) compounds and growing whilst attached to particles, surfaces and microalgal cells (Teeling et al., 2012; Fernandez-Gomez et al., 2013; Williams et al., 2013). Teeling et al. (2012) also observed that the bacterial response to a coastal microalgal bloom was characterised by an initial surge in Bacteroidetes abundance. Thus, it was hypothesised that this group colonises the microalgal surface and acts as "first responders" to microalgal blooms (Williams et al., 2013). Therefore, the identification of proteins that suggest a tight microalgae – bacteria relationship were expected to be found early in the time series. Also, the higher microalgal concentrations in eutrophic conditions would presumably provide a richer environment for the Bacteroidetes population.

As predicted, in both oligotrophic and eutrophic treatments, Bacteroidetes proteins were considerably more abundant in the early phase of the experiment (early, emPAI = 14.84, late, emPAI = 5.85) with several of the identified proteins suggesting a close association with microalgae (Fig. 3.3). First, several proteins attributed to the TonB-dependent transporter (TBDT) system were identified. TBDTs are involved in proton motive force-dependent outer membrane transport and once thought to be

restricted to iron-chelating compounds (i.e. siderophores) and vitamin B12 uptake. Recently TBDTs have been found to specialise in the uptake of HMW compounds that are too large to diffuse via porins (e.g. polysaccharides, proteins) (Blanvillain et al., 2007). In Bacteroidetes, the genes for the TBDT system are located in the same gene cluster as several of the polymer capture (e.g. starch utilisation system) and degradation genes (e.g. glycoside hydrolases (GHs), peptidases) suggesting an integrated regulation of capture, degradation and transport of complex substrates (Fernandez-Gomez et al., 2013). The proteins identified in our Bacteroidetes dataset support this suggestion.

Second, three starch utilisation system proteins (SusD/RagB) in Bacteroidetes were identified early in the time series (Fig. 3.3). SusD proteins are present at the surface of the cell and they mediate starch-binding before transport into the periplasm for degradation. RagAB is involved in binding exogenous proteins (Gilbert, 2008; Dong et al., 2014). GHs from several families (GH3, GH29, GH30 and GH92), together with three peptidases (methionine aminopeptidase, peptidase M16, peptidyl-dipeptidase) were also identified. As mentioned previously GHs are carbohydrate-active enzymes (CAZymes) specialised in the uptake and breakdown of complex carbohydrates, especially microalgal polysaccharides (Teeling et al., 2012; Mann et al., 2013). Together with peptidases these enzymes are responsible for extracellular breakdown of organic matter in order to be transported into the cytoplasm by the TBDT system.

Finally, the identification of proteins with cell adhesion functions (intimin, thrombospondin 1, gliding motility protein and YD repeat) provides further evidence that this bacterial phylum specialises in surface attachment. Intimin, thrombospondin and YD repeat protein are adhesive proteins that mediate cell-to-cell interactions and gliding mobility proteins allow exploration of solid surfaces (McBride, 2001). Other bacterial species utilise gliding motility for essential life cycle processes (e.g. swarming, predation) usually in coordinated groups but also as isolated adventurous individuals (Nan et al., 2011). In a similar way Bacteroidetes species may use gliding motility to follow microalgal exudate trails and to move to advantageous positions within the phycosphere, the microscale mucus region rich in organic matter that surrounds microalgal and cyanobacterial cells. This could confer a competitive advantage over free-floating bacterial species.

When contrasting oligo- and eutrophic treatments, Bacteroidetes associated proteins were, unexpectedly, more abundant in oligotrophic rather than eutrophic conditions (oligotrophic, emPAI = 14.02; eutrophic, emPAI = 6.67). In eutrophic conditions proteins attributed to transport, macromolecule degradation, outer membrane

capture and chemotaxis were virtually non-existent (Fig. 3.3). The fact that very little capture and degradation was occurring in eutrophic conditions suggests microalgal exudation was substantially lower. In the past, it has been hypothesised that nutrient limitation is a requirement for microalgal and cyanobacterial exudation (Wood et al., 1990; Guenet et al., 2010). Van den Meersche et al. (2004) determined that contribution of microalgal derived DOM to the experimental ecosystem C pool varied from ~2% (nutrient-replete early bloom) to 65% (nutrient-deplete mid-late bloom). Thus, the stimulation of DOM release, by nutrient limiting conditions, paradoxically provides C substrates for bacterial growth which then compete with the microalgae for nutrients (Van den Meersche et al., 2004). Therefore, the survival of Bacteroidetes populations seems to be linked to environmental conditions and the physiological state of neighbouring microalgae.

#### 3.4.5.3. ABC transporters reveal ecological niches

In Alpha- and Beta-proteobacteria, ABC transporters were the most prevalent transport proteins identified (Fig. 3.3). This is in agreement with previous freshwater and marine metaproteomic studies (Ng et al., 2010; Teeling et al., 2012; Georges et al., 2014). The majority of the ABC transporters were periplasmic-binding proteins (PBPs). The high representations of PBPs is commonly observed in aquatic metaproteomic studies. These subunits are far more abundant than the ATPase or permease components of ABC transporters in order to increase the frequency of substrate capture. Membrane proteins (e.g. permeases) are also inherently difficult to extract and solubilise therefore reducing the frequency of their detection (Williams et al., 2014).

In a metaproteomic comparison of Atlantic Ocean winter and spring microbial plankton, Georges et al. (2014) found ABC transporters were more abundant in low nutrient surface waters in mid-bloom and were mostly specific for organic substrates. Therefore, these type of transporters may be expected to more prevalent in the early oligotrophic conditions of our study where bacterial levels were higher (Fig. 2.2 B) and the environment was rich in microalgal and cyanobacterial exudate (discussed in previous section). As expected, transporter proteins in Alpha- and Beta-proteobacteria were more abundant in oligotrophic than eutrophic conditions (emPAI = 3.32 and emPAI = 1.73, respectively). They were predominant in early phase in oligotrophic (early, emPAI = 1.42 and late, emPAI = 0.9) and late phase in eutrophic conditions

(early, emPAI = 0.42 and late, emPAI = 0.88). Furthermore, in both treatments and timepoints the majority of ABC transporters were specific for organic substances (i.e. carbohydrates and amino acids). This suggests that both proteobacterial phyla are specialised in obtaining nutrients from DOM therefore investing more resources in the acquisition of organic rather than inorganic substrates and were favoured in early oligotrophic when the rate of microalgal exudation was potentially higher (Teeling et al., 2012).

Finally, another particularity of ABC transporters is that the expression of these transporters comes at an additional metabolic cost and therefore they are mainly synthesised to target substrates that are limiting in the environment. Thus, determining which transporters are being expressed can provide clues to which substrate is limiting. There was a clear difference in substrate preference between the two (Fig. 3.3); Alpha-proteobacteria predominantly expressed carbohydrate transporters (carbohydrate, emPAI = 0.61; amino acid, emPAI = 0.33) whereas in Beta-proteobacteria mainly amino acid transporter expression was observed (carbohydrate, emPAI = 0.10; amino acid, emPAI = 1.09). This has been previously observed (Schweitzer et al., 2001; Pérez et al., 2015) and is a case of resource partitioning, a mechanism through which two phylogenetic groups can co-exist in the same environment without leading to competitive exclusion (Morin, 2011).

## 3.5. Conclusions

A label-free comparative metaproteomics approach was applied on an experimental microcosm community under differing trophic states. The identification of proteins in early and late oligo- and eutrophic conditions allowed us to link function to phylogenetic diversity and reveal individual transitional niches. The results from this study also compared favourably with many *in situ* aquatic metaproteomic studies.

Microalgae and cyanobacteria predominantly expressed, as would be expected, proteins related to photosynthesis and C fixation. Interestingly, proteins involved in mechanisms of C concentration were abundant in virtually all samples, which indicated that C could be a limiting factor throughout the experiment. The fact that cyanobacteria, in eutrophic conditions, expressed several proteins related to environmental adaptation (e.g. microcompartment proteins) suggests that they may be better equipped than microalgal species to dominate nutrient enriched environments.

Proteins identified in all bacterial species suggested an alignment with oligotrophic environments. In early oligotrophic, Bacteroidetes showed characteristics that suggest a role as a fast-growing population that is specialised in cell and particle attachment and are the first to respond to microalgal growth. This ecosystem role can coexist with bacterial heterotrophs that live suspended in the water column and depend on microalgal exudate and decaying organic matter. ABC transporters were amongst the most abundant proteins identified. In a case of resource partitioning it was found that Alpha- and Beta-proteobacteria co-exist and metabolise microalgal/cyanobacterial exudate, but the former will preferentially uptake carbohydrates whereas the latter will prefer amino acid uptake thus avoiding direct competition. There is the evidence that bacterial metabolism controls primary production through the remineralisation of nutrients, however, here it is shown that primary producers can also be a driver of bacterial community composition and function.

This study successfully showed that microcosms can be used to observe microbial mechanisms that are typical of the natural environment. While these microcosm systems are simplified, and may not completely represent global biogeochemical cycles, they can accurately provide a snapshot of a microbial community in controlled conditions, and offer the potential to employ more manipulative experimentation to uncover functions and processes in oligo- and eutrophic conditions. The study also demonstrated that a community metagenetic analysis can provide a usable database for high mass accuracy metaproteomics studies. Ultimately, these data suggest that nutrient enrichment affected the dynamics of individual microbes and how they interact with others in their vicinity. Further manipulative experiments and associated ‘omics methodology will significantly contribute to our understanding of how microbial communities adapt to local environmental conditions.

# Chapter 4: Response surface methodology to determine the biotechnological value of a high-lipid *C. reinhardtii* mutant strain

## 4.1. Abstract

Microalgae are fast growing unicellular organisms that can quickly adapt to different environmental conditions and produce an array of bioactive molecules of commercial interest. Thus, they are a much sought after biomanufacturing platform for low- to medium-value compounds, such as animal feed and biodiesel. To date, economically viable large-scale production remains a challenge and, therefore, research into improving processes is vital. Typical approaches involve manipulation of metabolic pathways and growth conditions (i.e. nutrient stress). While these methods have been very successful in improving yields at laboratory scale, there are both environmental and industrial risks associated to the large-scale utilisation of mutant microalgae strains. These include escape of mutant microalgal strains into the natural environments and reduction in productivity as mutants potentially revert to wild type phenotypes. To assess the reality of both these potential risks we carried out a response surface methodology study to determine the competitive outcome of two *Chlamydomonas reinhardtii* strains, a wild-type (CC-124) and a high-lipid mutant (CC-4333), with differing levels of wild-type to mutant ratios and nitrogen source (ammonium chloride). Triglyceride concentrations were also used as a response variable in order to assess how competition between the wild type and mutant, and nutrient stress affect lipid production. Results show that the mutant strain is outcompeted by its wild type relative in every treatment, where the wild type was present, suggesting that there is little risk of *C. reinhardtii* mutant strains displacing the wild type in an "escape scenario". The highest total cell numbers were found at 200 mg $L^{-1}$ and 375 mg $L^{-1}$ of ammonium chloride, with no significant difference between them, which suggests that an optimal nitrogen source concentrations in cultivation need to be deciphered to minimise cultivation costs at industrial scale. Unexpectedly, triglyceride concentrations were found to be highest in the *C. reinhardtii* co-culture with 25% WT and in the pure mutant culture after just 24 hours. This suggests that, in co-culture, intraspecific competition enhanced lipid accumulation. In the pure culture this very short response period suggests there is potential to develop a small batch biomanufacturing system that can be scaled out (i.e. multiplication of small volume modules) instead of scaled up. This study paves the way for future industrial risk assessments with different strains of biotechnological interest.

## 4.2. Introduction

In recent years, the inevitable depletion of fossil fuels and concerns regarding climate change have led to increasing investment in searching for economically viable and environmentally friendly alternative sources of energy. Biofuels have been put forward as a sustainable and eco-friendly approach to solve this problem. Three generations of biofuel feedstocks have been developed. The first and second generation utilised food and non-food crops. However, given the inherent limitations of agricultural land availability and increasing food demand, they are not considered a long term and sustainable option. The third generation feedstock, microalgae, has been identified as a superior option. Microalgae can grow very rapidly, survive over a wide range of environmental conditions and produce a wide range of bioactive molecules of commercial interest. These characteristics make microalgae an appealing feedstock for LMVC production such as animal feed and biofuels (Benemann, 2013).

Research into microalgal biodiesel has progressed considerably in the last decade, however, large-scale production in outdoor raceway ponds, will not be economically and energetically viable until a number of challenges have been addressed (Leite et al., 2013). One important challenge is selecting a microalgal strain capable of high lipid productivity. Thousands of strains can be isolated from the environment and screened for desired properties such as the inhibition of starch production which leads to lipid over accumulation (Li et al., 2010). However, few will have the desired characteristics for large-scale production and ecological trade-offs suggest that an ideal strain will be extremely difficult to identify (Shurin et al., 2013). One way to improve lipid productivity is through random mutagenesis, followed by screening for desired characteristics, (e.g. Work et al., 2010; Guarnieri and Pienkos, 2015) Although mutagenesis and screening can be time consuming, costly and unspecific it can improve productivity several fold and mutants obtained this way are not, currently, considered GMOs.

A common strategy applied to increase lipid productivity is the manipulation of growth conditions. Many microalgal species have been found to have naturally high lipid content (30 – 50% dry weight) that can be increased by manipulating cultivation parameters. Environmental stress activates the neutral lipid biosynthetic pathway and the accumulation of neutral lipids, especially TAGs. This is believed to be a response to unfavourable growth conditions. In a review of the oleaginous microalgal literature, Hu

et al. (2008), reported that environmental stress increased total lipid content from 25.5% to 45.7%. The authors also concluded that N limitation is the single most critical stress affecting lipid metabolism in microalgae. Therefore, the use of transgenic strains subjected to N stress has been an active field of research for optimal lipid productivity (Work et al., 2010; Siaut et al., 2011).

While the widespread implementation of transgenic strains would undoubtedly bring added value to LMVC production, there are both environmental and industrial risks associated to the large scale cultivation of such microalgae. Environmentally, the risk stems from release of transgenic organisms, which could impact on the biodiversity of the surrounding environment by outcompeting native strains (Snow et al., 2012). Industrially, there is the risk that transgenic microalgal strains could revert to WT through natural mutation and through a clonal interference scenario dominating the outdoor raceway pond. Moreover, contamination can also occur through invasion by indigenous microalgal species, but this is not a focus of this study. Such risks would benefit from a strategic investigation of competitive dynamics where the fitness of modified strains are determined with their WT as a baseline. If the fitness of the transgenic strain is lower than that of the WT then it could be argued that any risk analysis could be based solely on the risks of cultivating the WT (Gressel et al., 2013).

To further this objective, this study focuses on the competitive outcome between two *Chlamydomonas reinhardtii* strains. *C. reinhardtii* is a popular model organism with a well-annotated genome (Merchant et al., 2007), powerful molecular tools (Jinkerson et al., 2015) and a strong increase of reserve molecules (e.g. lipid, starch) in response to N deprivation (Siaut et al., 2011). The CC-124 strain, with a cell wall and flagella, widely used in laboratory studies (Pröschold et al., 2005) (hereon referred to as WT strain), was competed against CC-4333, a low starch, no flagellum, cell wall deficient mutant, obtained through x-ray mutagenesis, that lacks the catalytic (small) subunit of ADP-glucose pyrophosphorylase (Ball et al., 1991) (hereon referred to as mutant strain). The residual starch accumulation has been shown to provide the mutant strain with the ability to accumulate up to 3.5 times more total lipids than the WT strain (Li et al., 2010).

Preliminary data on the growth of both strains (Fig. 4.1) in pure culture suggests the mutant strain consistently has a lower cell number during the lag phase, prior to exponential growth. Furthermore, Work et al. (2010) observed that these particular *C. reinhardtii* mutant strains have lower levels of photosynthetic $O_2$ evolution under

N-deprivation. Both these features suggest that the mutant strain will not compete well when in direct competition with the WT strain.

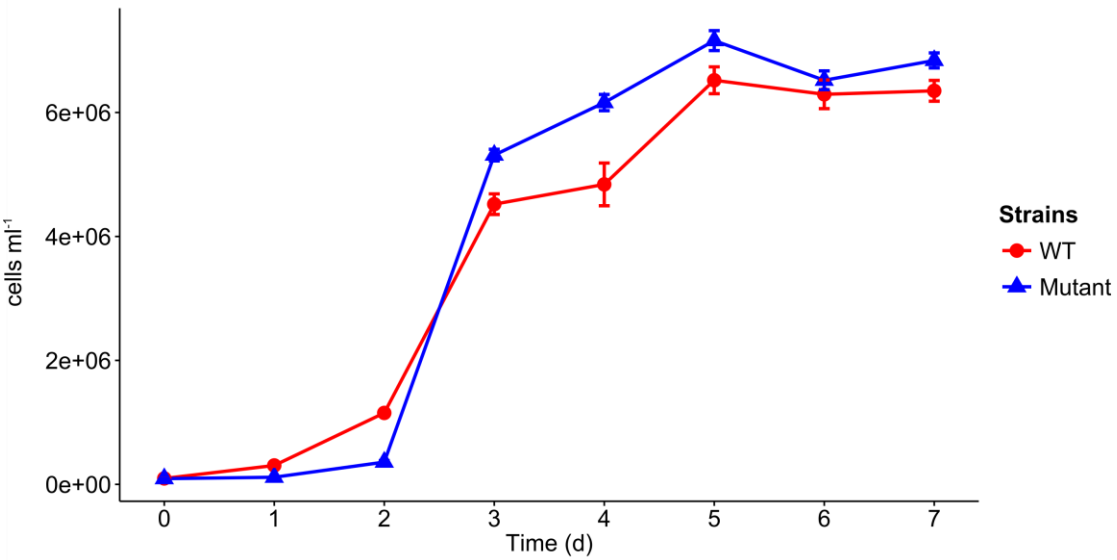


Figure 4.1. Growth curves, measured in cells $mL^{-1}$, of WT (CC-124) and mutant (CC-4333) *C. reinhardtii* strains. Triangles indicate the mutant strain and circles indicate the WT strain. Error bars show standard errors (n = 3).

A response surface competition experiment was developed, varying both the initial proportion of the WT and mutant strains, together with the N concentration (ammonium chloride ($NH_4Cl$)) in the media. This design was used to determine both the initial WT:mutant cell number ratio and $NH_4Cl$ concentrations, where extinction of the predicted slower growing mutant occurred, if at all. This allows an evaluation of whether there are points in time, initial WT:mutant cell number ratio and concentrations of $NH_4Cl$ where TAG concentrations are highest. Such data will allow an understanding of *C. reinhardtii* WT and high-lipid mutant competitive dynamics and whether, when and how intra-specific competition can affect TAG productivity. To date, this is the first study where the potential industrial impact of transgenic microalgae reversion to wild type or the environmental risk of mutant escape, have been assessed simultaneously.

## 4.3. Methods

### 4.3.1. Strains and culturing conditions

CC-124 (wild type) and CC-4333 (high lipid mutant) were obtained from the Chlamydomonas Resource Center (University of Minnesota, USA). The cultures were maintained on tris-acetate-phosphate (TAP) medium as described by Gorman and Levine (Gorman et al., 1965). Cells were grown in 50 mL centrifuge tubes, with 25 mL of culture, under 70 µmol $m^{-2}$ $s^{-1}$ constant illumination, on an orbital shaker at 110 rpm and with a horizontal orientation in order to maximise mixing and light penetration.

### 4.3.2. Competition experiment

The response surface design is an experimental design where two variables are varied simultaneously, and estimates made of response variables at several combinations. A standard analysis for such data is to fit a response surface model, the basic fitting a second order polynomial for each variable, and an interaction term between them. This is a flexible model that can estimate planes, ridges, valleys and peaks as a function of linear polynomial functions of each variable, and their interaction. It specifically allows the evaluation of whether there are additive or interactive (synergistic/antagonistic) effects of competition and nutrient enrichment on, for example, TAG production or long term dynamics (e.g. competitive outcomes).

In this response surface design, cultures were grown to late log phase and resuspended at $1 \times 10^5$ cells $mL^{-1}$, in parallel, in TAP medium with four different concentrations of $NH_4Cl$: 50, 100, 200 and 375 mg $L^{-1}$ and five different initial WT:mutant cell number ratio (100:0, 75:25, 50:50, 25:75 and 0:100). 200 µL aliquots were taken immediately after inoculation (0 h), and every 24 hours after that, for cell counts and TAG quantification. A total of 1.2 mL was removed, from each tube, over the course of the experiment. Cell counts were performed using a Bright-Line glass haemocytometer (Hausser Scientific, USA) on a BX 51 microscope (Olympus, Japan). From the cell counts time series the carrying capacities (K) and exponential growth rates (r) were estimated by fitting a logistic growth curve to each respective time series in GraphPad Prism 6.07 (GraphPad Software, Inc., USA). The WT and mutant strains are visually undistinguishable. Therefore, to separately count each strain the total number of cells were counted, samples were then incubated, for five minutes in 0.5%

Triton X-100 (Sigma, USA), to completely lyse the cell wall deficient mutant cells. Preliminary tests with all initial WT:mutant cell number ratios showed a correlation of 0.98 between theoretical and observed cell lysis values (Appendix 3). Afterwards, the total number of cells were recounted and the difference between both counts equalled the number of mutant cells present. Cells for TAG quantification were lysed by sonication on ice with a FB 15051 sonicator (Fischer Scientific, USA), at power level three and duty cycle 30%, for three cycles of 15 seconds. TAGs were then quantified with the commercially available Thermo Scientific Infinity TAG reagent kit (Thermo Scientific, USA) according to manufacturer's recommendations.

From the time series replicates of each $NH_4Cl$ and initial WT:mutant cell number ratio treatment combination the total WT cell number, total mutant cell number, K and TAG concentrations were estimated and used to fit and visualise the response surface model. The models and visualisations were fit in the R Statistical Programming Environment (R Development Core Team, 2014) by employing the package "rsm" (Lenth, 2009).

## 4.4. Results and Discussion

*C. reinhardtii* mutants, and other species produced in a similar way, have attracted interest as microalgal host cells in industrial biotechnology, not only due to over-accumulation of lipids, but also, due to the method used for their creation. The European definition of GM is determined by the Directive 2001/18/EC of the European Parliament on the deliberate release into the environment of genetically modified organisms. In this document "genetically modified organism" is defined as an organism, with the exception of human beings, in which the genetic material has been altered in a way that does not occur by natural recombination. Thus, the omission from this framework of mutants obtained through mutagenesis opens avenues of research that will allow the use of mutant strains for biotechnological purposes.

This study utilises response surface methodology to characterise the competitive outcome between a high-lipid accumulating *C. reinhardtii* mutant strain and its WT as a function of $NH_4Cl$ concentrations and initial WT:mutant cell number ratios. The aim is to assess whether the fitness of a *C. reinhardtii* high-lipid mutant is higher than that of its WT, important environmentally from an escape scenario perspective and, simultaneously, the industrial impact of WT reversion of the mutant culture. This

approach also evaluates whether there is a point in time or a combination of $NH_4Cl$ concentrations and initial WT:mutant cell number ratios where TAG production is increased due to the effects of intra-specific competition.

### 4.4.1. WT and mutant *C. reinhardtii* strains grown under optimal conditions in co-culture and in isolation

When grown under optimal conditions, as a pure culture, the WT strain has an r of 1.76 ± 0.26 $d^{-1}$ and achieves a K of 6.25 x $10^6$ cells after 5 days (Fig. 4.1). The mutant strain presents a similar growth pattern, but has an increased lag phase, characterised by lower cell numbers prior to exponential growth. The mutant strain also has a higher, albeit not significantly, final cell number, K of 6.68 x $10^6$ cells and a significantly higher r of 4.08 ± 0.52 $d^{-1}$ (Fig. 4.1). These results are in accordance with a previous study that compared growth and lipid accumulation in a *C. reinhardtii* WT and the same mutant strain as this study (Work et al., 2010). From this comparison we can conclude that the increased lag phase could be detrimental to the survival of the mutant strain when grown in co-culture with the WT strain. Work et al. (2010) also observed that, although the highest cell concentration and growth rate were achieved in the mutant strain, its cell diameter was smaller than the WT. This observation was confirmed in our study (Fig. 4.2) and in several publications which have demonstrated that cell size and growth rates are negatively correlated (Schlesinger et al., 1981).

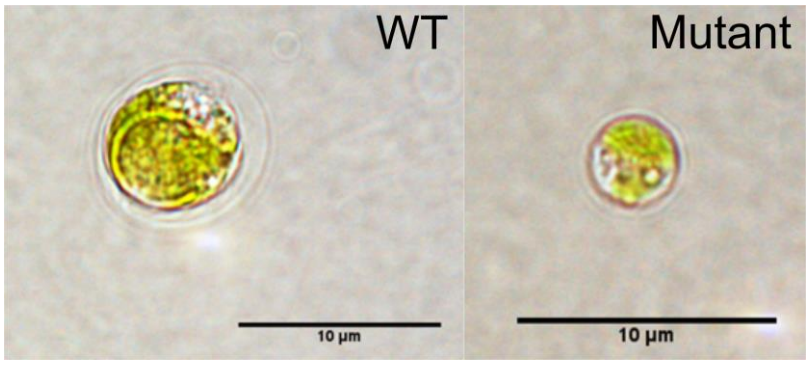


Figure 4.2. Optical microscopy pictures (1000x magnification) demonstrating the size difference between *C. reinhardtii* WT cells (left) and mutant cells (right).

### 4.4.2. WT and mutant *C. reinhardtii* strain competition as a function of $NH_4Cl$ concentrations and initial WT:mutant cell number ratio

Recent reports have hypothesised that genetic mutations of biotechnological interest (e.g. leading to increased lipid content) are all in the opposite direction of evolution (i.e. optimal growth under natural conditions) and therefore should be outcompeted by the respective WT strains (Flynn et al., 2010; Gressel et al., 2013). The results in this study support this theory. In every experimental treatment where the WT and mutant were mixed, the WT strain showed a significantly higher r and K than the mutant strain (Fig. 4.3). Gause's law of competitive exclusion states that, all other things being equal, two species competing for the same resource cannot coexist (Morin, 2011). This is especially true for closed systems such as the one in this study. Work et al., 2010 suggest several reasons this would happen: they found that, under N deprivation, starchless mutants metabolise less acetate and have severely attenuated levels of photosynthetic $O_2$ evolution than the WT. This indicates that the mutant strain responds to N deprivation by decreasing its overall anabolic processes (Work et al., 2010). If the cells are under similar stress, due to intraspecific competition, this could explain why the WT dominates over the mutant in every experimental scenario.

The results of this study suggest that the overall fitness of the mutant strain is inferior to that of its WT, when cultivated in co-culture in the initial ratios tested. In situations where both strains are sharing the same resource, the WT strains seems to maintain a physiological advantage which would prevent the mutant strain from achieving cell numbers that could put the existence of the WT at risk. However, this study was done in a closed system with uniform resource conditions, without the presence of predators (e.g. zooplankton), symbiotic and antagonistic organisms and without fluctuating environmental conditions. Any of these factors that would favour the mutant strain could create an unbalance and allow it to dominate over its WT, for example, differences in grazer palatability can have a dramatic effect on the competitive outcome. Van Donk et al. (1997) performed grazing experiments to determine the role of cell wall structure and nutrient limitation on the digestibility of the *C. reinhardtii* WT and cell wall deficient mutant. The study revealed that under nutrient limiting conditions, *Daphnia magna* clearance rates of *C. reinhardtii* WT cells were severely debilitated. However, under both non-limiting and limiting nutrient conditions, *C. reinhardtii* mutant cells were cleared equally (Van Donk, 1997). This supports our

conclusion that mutant cells, either stressed or non-stressed, remain unlikely to outcompete their WT strain.

Finally, both the data from this study and Van Donk et al. (1997) suggests that outdoor industrial cultivation of cell wall deficient mutant strains has many associated economic risks of failure to produce effective quantities of product. A reversion from the mutant strain to its WT, contamination of the mutant strain culture by its WT or invasion by grazers would result in rapid reduction or disappearance of the mutant. Therefore, while industrial outdoor cultivation of cell wall deficient mutant strains does not present a significant environmental risk to its WT in an escape scenario, the fragility of the mutant strain could impact on overall industrial productivity.

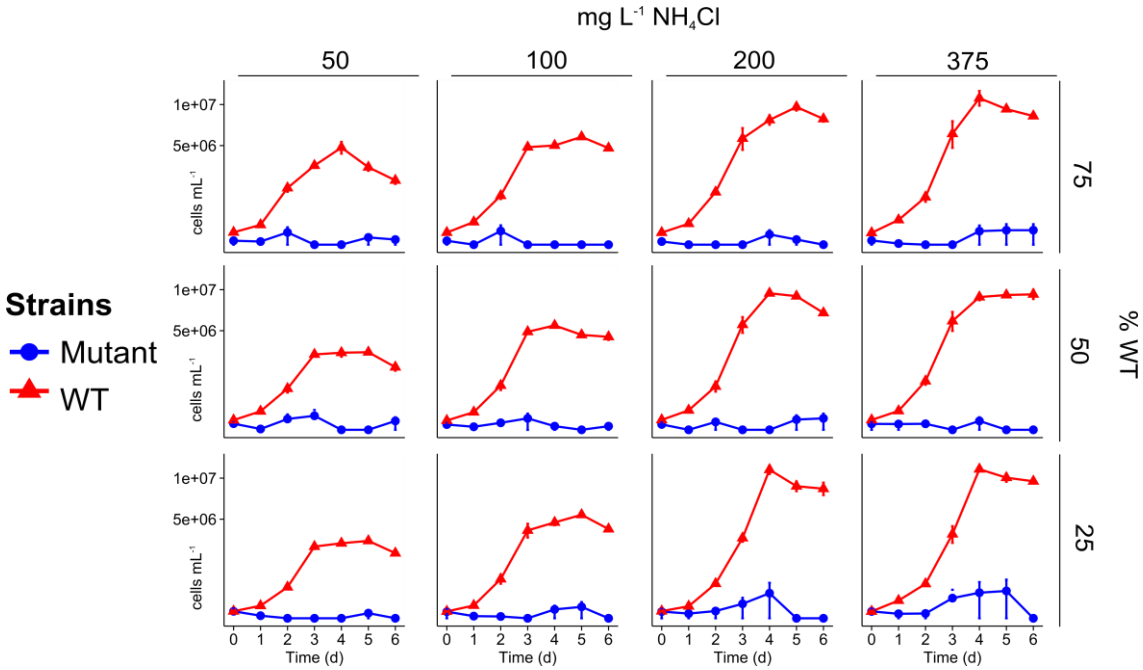


Figure 4.3. Time series of WT and mutant strain cell numbers as a function of $NH_4Cl$ concentrations (horizontal axis) and percentage of WT cells in initial culture (vertical axis). Data has been transformed with a square root transformation. Error bars are standard errors (n = 3).

### 4.4.3. Biomass production as a function of $NH_4Cl$ concentrations and initial WT:mutant cell number ratio

The analysis of the individual time series of both strains offered an understanding of the competitive dynamics in the experimental system and how this

could potentially affect natural ecosystems. However, in industrial cultivation, maximising biomass production is one of the major parameters of interest. A response surface model was fit to data on the maximum biomass (K, estimated from each replicate time series). K was estimated by fitting a logistic curve to each time series resulting in three estimates of K for each combination of $NH_4Cl$ concentration and initial WT:mutant cell number ratios.

K varied as a quadratic function of $NH_4Cl$ concentrations and initial WT:mutant cell number ratio ($F = 39.99$, $p < 0.001$) and the effect of nutrient levels depended on initial WT:mutant cell number ratios (interaction; $F = 35$, $p < 0.001$). The data support a ridge like pattern of K (Fig. 4.4). In accordance with Fig. 4.1 the maximum K observed was when the mutant strain was grown alone. The ridge of maximum K then decreases with increasing percentage of initial WT. As expected K is highest at higher nutrient levels with no significant difference between the K at 200 mg $L^{-1}$ $NH_4Cl$ and 375 mg $L^{-1}$ $NH_4Cl$. The standard recipe for TAP medium has a concentration of 375 mg $L^{-1}$ $NH_4Cl$ (Gorman et al., 1965), therefore this suggests that maximum K can be achieved with as little as 200 mg $L^{-1}$ $NH_4Cl$ and therefore, reduce N costs significantly in industrial cultivation.

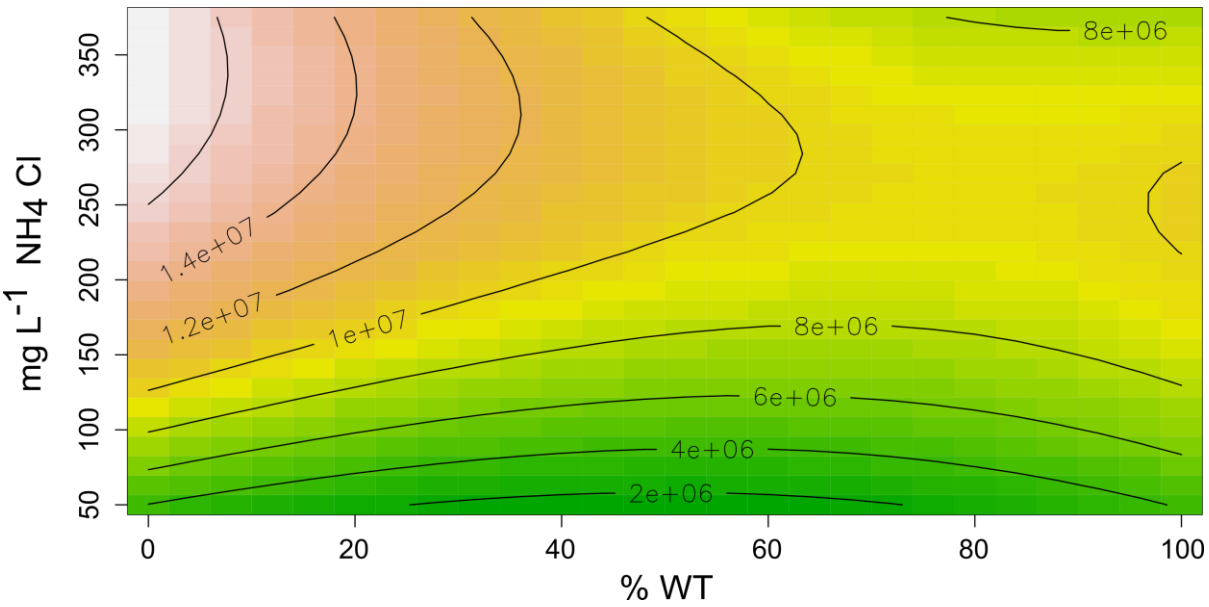


Figure 4.4. RSM analysis of carrying capacity (K) as a function of $NH_4Cl$ concentrations (vertical axis) and percentage of WT cells in initial culture (horizontal axis).

### 4.4.4. TAG production as a function of $NH_4Cl$ concentrations and initial WT:mutant cell number ratio

Lipid accumulation in *C. reinhardtii* strains has been the target of many studies in recent years. The lipids produced are normally composed of long-chain TAGs which can be converted to biodiesel by transesterification (Siaut et al., 2011) and concentrations per cell can be enhanced by nutrient deprivation (James et al., 2011). The mutant strain utilised in this study is especially interesting due to its inability to accumulate starch. Previous studies found that shutting down starch synthesis would lead to an increase in TAGs stored in the cell (Li et al., 2010; Work et al., 2010). Therefore, the production of TAGs in *C. reinhardtii* mutant strains under N deprivation has become a model procedure in the field of microalgal biodiesel research.

In this study, TAG concentrations were estimated every day to assess accumulation over time as a function of $NH_4Cl$ and initial WT:mutant cell number ratio. However, when accumulating TAGs in N-deplete conditions, overall biomass accumulation invariably stops. This is a cellular trade-off, where understanding the molecular mechanisms has been the target of many studies (Longworth et al., 2012; Yang et al., 2015) but remains a major challenge for scale up and commercial exploitation of N deprivation for TAG over-accumulation in *C. reinhardtii* cells. To shed light on this particular problem, TAG concentrations were plotted as TAGs per million cells in order to correlate TAG accumulation to biomass production (Fig. 4.5).

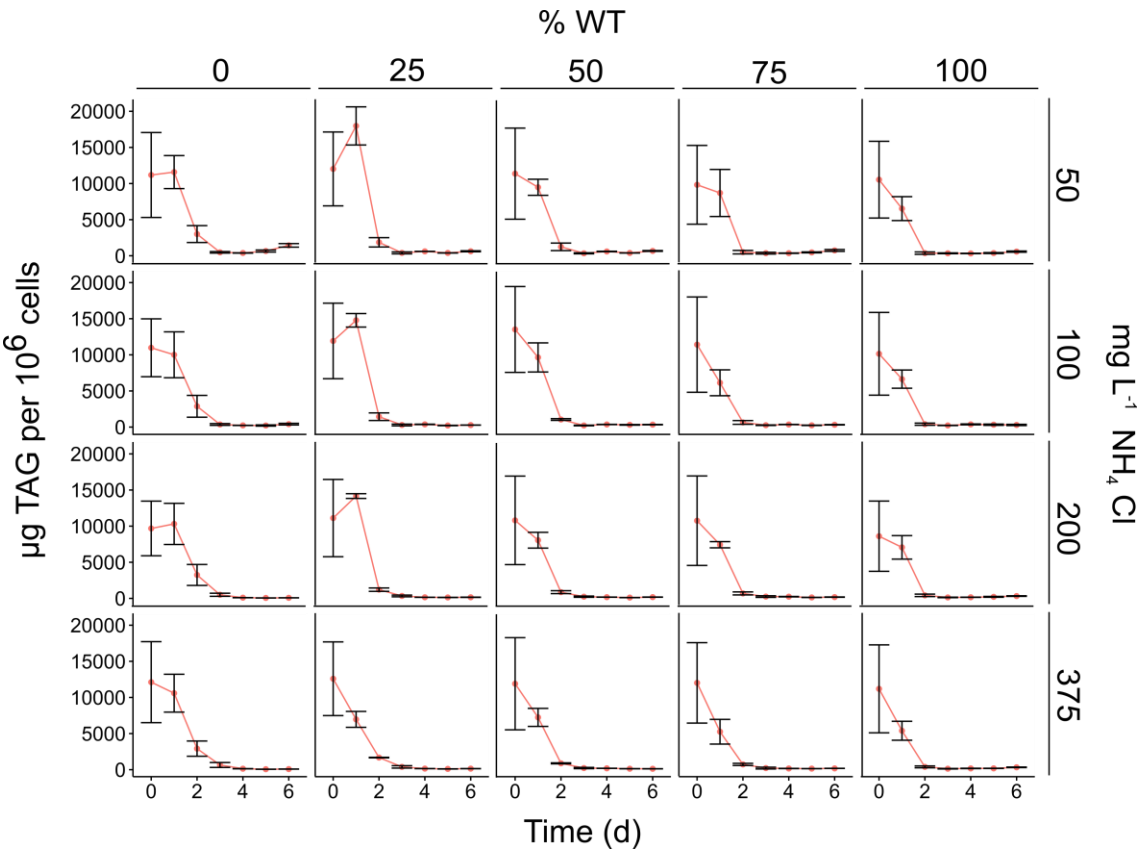


Figure 4.5. Time series of TAG concentrations per million cells (µg TAG per $10^6$ cells) for each combination of $NH_4Cl$ concentration and percentage of WT cells in initial culture. Error bars are standard errors (n = 3).

The trends observed in exponential (t2 – 4) and stationary phase (t4 – 6) are as expected. In exponential phase, lipid reserves decrease due to consumption in cell metabolism. In stationary phase, net growth halts and accumulation of reserve substances begins. This is in agreement with the extensive literature on lipid accumulation in *C. reinhardtii* (e.g. Siaut et al., 2011). Moreover, the highest TAG accumulation during stationary phase occurs with the 100% mutant strain at the lowest N concentration (50 mg $L^{-1}$ $NH_4Cl$) providing confidence in our data. However, unexpectedly, a sharp increase in TAG concentrations in the first 24 hours can be observed in the co-culture treatment with an initial percentage of WT cells of 25% and to a lesser extent in the mutant pure culture. A comparison between this co-culture and the mutant pure culture shows that the co-culture has a significantly higher concentration of TAG per million cells in 50 mg $L^{-1}$ $NH_4Cl$ ($p < 0.05$), but a significantly lower number of cells. Furthermore, when considering the total TAG values, independent of cell number, both co-culture and pure culture are not significantly different. This suggests that interspecific competition can lead to increased

lipid accumulation, on a per cell basis, however, the negative effect on net growth is detrimental to the total amount of lipids in the culture. This response is similar to the one under N stress where there is a trade-off between net growth and lipid accumulation. Ultimately, these results show stress through intraspecific competition does not lead to significant benefit in lipid production in *C. reinhardtii* mutant cultures.

To obtain further insight into lipid accumulation after 24 hours we fit a response surface model to data on the TAG concentrations on day 1 for each combination of $NH_4Cl$ concentrations and initial WT:mutant cell number ratios (Fig. 4.6).

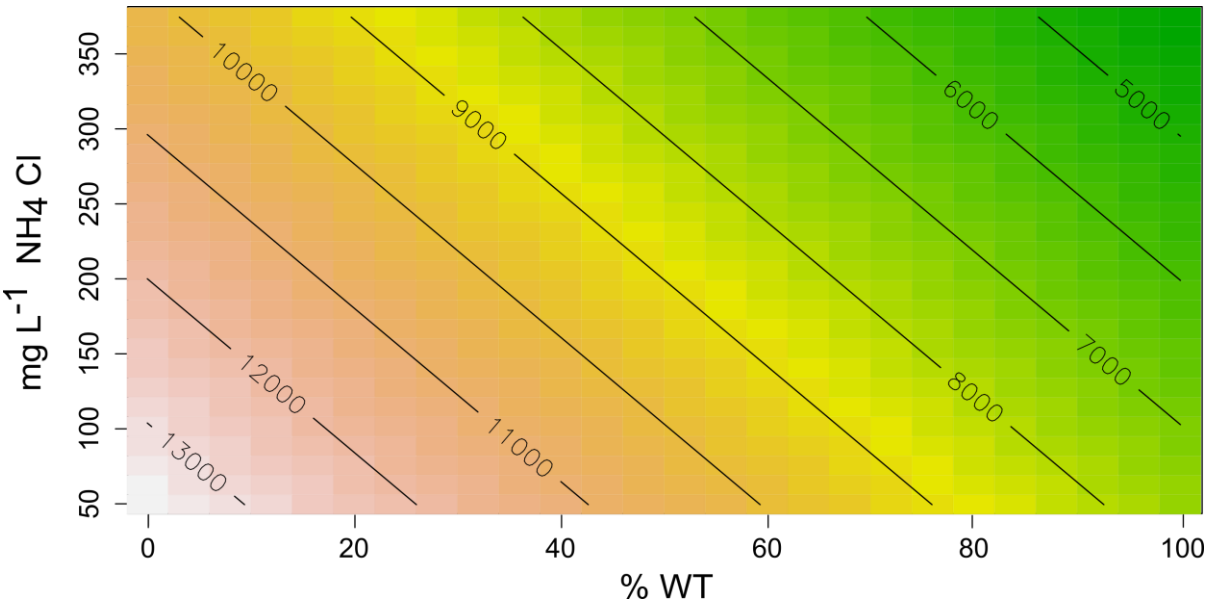


Figure 4.6. RSM analysis of TAG concentrations (µg TAG per $10^6$ cells) at day 1 as a function of $NH_4Cl$ concentrations (vertical axis) and initial WT percentage (horizontal axis).

TAG concentrations varied linearly with $NH_4Cl$ concentrations and initial WT:mutant cell number ratio (quadratic: F = 0.41, p = 0.66) and the effect of nutrient levels on TAG concentrations does not depend on cell number ratios (interaction: F = 0.54, p = 0.46). This indicates an additive effect of cell number ratio and nutrient levels on TAG concentrations. Finally, TAG concentrations decline linearly with increasing nutrients (slope:-10.4, t = -2.8, p = 0.006) and decline much more strongly and linearly with increasing WT (slope: -6010, t = -4.6, p < 0.001).The analysis of TAG concentrations at day 1 indicates that, as expected, TAG accumulation increases with the percentage of mutant strain cells and the decrease of $NH_4Cl$ concentrations. It also indicates that, in nutrient limiting conditions, there may be benefit in harvesting *C. reinhardtii* mutant cultures after 24 hours of nutrient stress.

Currently, stress induction for lipid accumulation is done in a two-phase approach where cells are grown to late exponential (biomass generation stage) and then transferred into a nutrient deplete medium for 72 – 96 hours (product generation stage) (Work et al., 2010). Alternatively, on a large scale, cultures can be grown until nutrient depletion thus entering a stress phase. The first approach is not cost effective due to double harvesting while the second approach is time consuming and difficult to control. Our findings suggest using a late exponential stage (high biomass) culture to inoculate multiple small batches with nutrient deplete medium. The *C. reinhardtii* mutant strain, in pure culture, at 50 mg $L^{-1}$ $NH_4Cl$, has a significantly higher ($p < 0.001$) TAG concentration, at day 1, 11583 ± 272 µg per $10^6$ cells, than at day 6, 1442 ± 29 µg per $10^6$ cells. In the same treatment, the cell number at day 1, $1.46 \times 10^5 \pm 1.54 \times 10^4$, is significantly lower than the cell number at day 6, $1.00 \times 10^6 \pm 1.14 \times 10^5$. Thus, the total lipid amount is higher, albeit not significantly, at day 1, 1613 ± 105 µg, than at day 6, 1376 ± 65 µg. This implies it is feasible to divide a late exponential stage culture into small batches in order to harvest after 24 hours. Due to the nature of the process each batch will have a low number of cells which forces a large increase in the number of modules. However, by keeping the batch scale small (e.g. < 1L) this allows for flexible small batch manufacturing, scale-out, rather than scaling up. Scaling out, also known as horizontal scaling, has the advantage of eliminating the risk of an organism changing its performance with an increase in vertical scale. It also allows for easy expansion, by simply adding more production units, and has the potential to be fully automatised in a biorefinery production line. Further work is needed to develop this 24 hour process across other *C. reinhardtii* strains and species and to prove its industrial and economic feasibility.

## 4.5. Conclusions

The aim of this study was to compare a scenario of a *C. reinhardtii* high-lipid mutant outbreak, and potential direct competition with its WT, to a scenario of WT contamination of an industrial *C. reinhardtii* high-lipid mutant production system. A response surface design was used to evaluate the outcome of competition across multiple levels of $NH_4Cl$ and initial WT:mutant cell number ratios. Finally, the same

experimental design also allowed evaluation of whether the competitive outcome would affect biomass and TAG concentrations across these treatments.

Based on previous studies, it was hypothesised that the mutant strain had a lower fitness than its wild type. This was confirmed in every treatment where the WT was present. Therefore, it was concluded that there is little risk in *C. reinhardtii* mutant strains outcompeting WT strains both in an artificial enclosure and in the wild. It was also hypothesised that higher nutrient concentrations would result in higher maximum cell numbers. This was true for the two highest nutrient treatments (200 mg $L^{-1}$ and 375 mg $L^{-1}$ $NH_4Cl$) albeit with no significant difference between them. Therefore, there is room to improve upon the traditional TAP medium recipe and decrease costs for its industrial formulation. Finally, the highest TAG concentrations were found after 24 hours at 50 mg $L^{-1}$ $NH_4Cl$, most notably in the *C. reinhardtii* co-culture with 25% WT and to a lesser extent in the high-lipid mutant pure culture. While this suggests that strong intraspecific competition led to increased TAG accumulation, on a per cell basis, the total TAG values between the *C. reinhardtii* co-culture and pure culture were not significantly different. Therefore, there is little apparent value in using intraspecific competition stress as a lipid trigger. Nevertheless, the results obtained with the pure mutant culture suggest a 24 hour growth period has the potential to be scaled-out in a small batch biomanufacturing system. These type of systems are more flexible and cheaper than the traditional scale-up systems. In summary, this study showed that the industrial cultivation of *C. reinhardtii* high-lipid mutant does not present a significant environmental risk to its WT in an escape scenario and has proposed novel routes to improve the economic and industrial feasibility of its cultivation.

# Chapter 5: Discussion

The overall aim of this thesis was to show how the application of ecological theory can be beneficial to the production of LMVCs in microalgae. This was subdivided into three chapters which approached the objective at hand with different ecological tools. In Chapters 2, 3 and 4 of this thesis, modelling and proteomics were applied to microbial microcosms. This section will review the main findings of each chapter and discuss their application to microalgal biotechnology.

## 5.1. Multivariate analysis of the response of a freshwater microbial community under nutrient enrichment

In Chapter 2, the interaction among nutrients, physicochemical variables, microalgae and bacteria was investigated under oligo- and eutrophic nutrient levels. The aim was specifically to characterise the joint effect of physicochemical variables (i.e. DO, pH and temperature) and nutrient variables (i.e. $NO_3^-$, $PO_4^{3-}$ and $NH_4^+$) on the microalgae-bacteria community. To this end multivariate (SEM) analysis was used to study all the interactions present in the system. It was concluded that physicochemical and nutrient variables had an additive, non-synergistic, effect on the microalgae-bacteria community and that physicochemistry became a stronger predictor, than nutrient concentrations, of biotic variation in nutrient enriched conditions. These results provided evidence towards the broader discussion of the relative importance of physicochemical and nutrient variables in microalgal and cyanobacterial bloom promotion. The data also highlighted the importance of utilising multivariate analysis to extract useful information from complex biological systems.

### 5.1.2. SEM for microalgal biotechnology

Mathematical models are useful tools to draw conclusions from existing data and utilise these to create predictive hypotheses. One of the more interesting applications of these models is their potential to use laboratory and field data to make predictions of how an organism responds to abiotic variables in order to design and upscale a growth system in a way that provides optimal performance. The conclusions from Chapter 2 suggest that different microalgal cultivation stages can be influenced by

different groups of abiotic variables. For example, a typical microalgal lipid production system consists of two stages: a nutrient-rich biomass generation stage and a nutrient-deplete product generation stage. Based on the SEM results, it is to be expected that nutrient variables will be stronger predictors of microalgal biomass variation in the nutrient-deplete stage (analogous to oligotrophic conditions). However, in the nutrient-rich stage (analogous to eutrophic conditions) microalgal biomass variation may not correlate strongly with nutrient concentrations. Thus, it is important to look at other abiotic parameters, such as temperature or light, as drivers of nutrient-rich microalgal cultures. This shows that utilising a SEM approach to determine the environmental drivers of natural blooms can provide useful insight for the management of pilot-scale microalgal cultures.

To date, SEM models have yet to be applied for microalgal cultures in biomanufacturing thus, it is important to look at other fields to aid model construction for microalgal cultivation. Wang et al. (2013) created a path diagram (i.e. structural equation model) to study the direct and indirect effects of water temperature, amplitude of water wave, chlorophyll concentration, pH and DO on the turbidity of water in settling tanks of a drinking water treatment plant. Furthermore, this analysis was done over four different periods of the day to test how the relationships change with diurnal variation in the settling tanks (Figure 5.1 A). Based on the results of the model, the turbidity of the water was found to be controlled, predominantly, by synergistic and antagonistic relationships between pH values in the settling tanks and temperature and chlorophyll *a* concentrations in the catchment area. Furthermore, the strength and direction of these relationships changed under different environmental conditions. The results allowed the creation of hypotheses to effectively regulate the quality of the water supply. The same exercise could, potentially, be adapted to understanding microalgal raceway operations. A hypothetical pilot-scale microalgal culture interaction model (Figure 5.1 B) can describe the direct effects of light, temperature (Temp), dissolved carbon dioxide ($dCO_2$), pH and nutrient concentrations (N/P) on the biomass productivity (DW). This type of model can be also extended and applied over different periods (e.g. day/night cycle), growth phases (e.g. lag, exponential, stationary) or environmental conditions (e.g. summer/winter) to summarise changes in the microalgal cultures and allow fine-tuning of growth parameters.

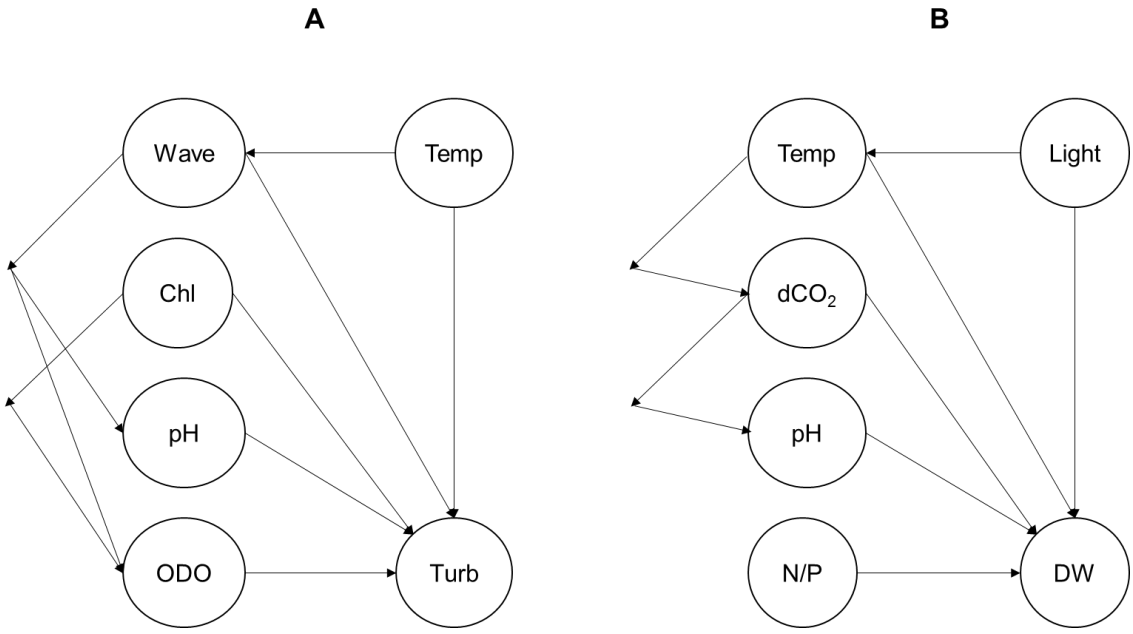


Figure 5.1. Comparison of path models with relevance to pilot-scale microalgal cultures. **A**: Path model from Wang et al. (2013) that describes the direct effects of amplitude of water wave (Wave), chlorophyll concentration (Chl), pH and optical dissolved oxygen (ODO) on the turbidity of the water (Turb) within the settling tanks of a drinking water treatment plant. **B**: Hypothetical path model, based on the design in Wang et al. (2013), that describes the effect of light, temperature (Temp), dissolved carbon dioxide ($dCO_2$), pH and nutrient concentrations (N/P) on the biomass productivity (DW) of a pilot-scale microalgal culture. Arrows represent direct effects between variables.

In summary, SEM is a tool to analyse a complex network of intercorrelated variables while aggregating and summarising the underlying causal linkages between those variables. However, this effort to summarise complex relationships results in a description of correlations (and partial correlations) among different variables that fails to touch upon the underlying mechanisms of interaction in biological systems. Tools that can provide insight into functional diversity are necessary to unravel the complex network of interactions. Therefore, a metaproteomic description of the system may be able to answer the questions raised by SEM.

## 5.2. A metaproteomic analysis of the response of a freshwater microbial community under nutrient enrichment

In Chapter 3, a metaproteomic approach was applied to investigate the functional changes to microcosm microalgal and bacterial communities, over time, in oligotrophic and eutrophic conditions. Samples were taken early during microalgal and cyanobacterial dominance and later under bacterial dominance. 1048 proteins, from the two treatments and two time points, were identified. The description of proteins expressed in oligo- and eutrophic conditions provided a snapshot of the whole community and confirmed several existing hypotheses regarding microbial ecophysiology under nutrient deplete and nutrient enriched conditions. Nevertheless, there were some limitations to this study. First, the samples processed in this study did not provide enough biomass to be analysed in replicate and had to be pooled. Second, many peptide spectra were not assigned to proteins because the proteins were not present within the created database. While these are major issues in the analysis of environmental samples, they are unlikely to hinder the analysis of biotechnologically relevant microbial communities. The production of large amounts of biomass is a given in pilot-scale cultivation and many model microalgae, and associated bacteria, are fully sequenced. Therefore, it would be easier to create a metagenomic or metatranscriptomic dataset and utilise this information to create a reference protein database.

The conclusions of this study show that metaproteomics can verify existing hypotheses regarding ecophysiological processes and shows promise for application in, pure microalgal and community-based growth systems. Some potential applications of meta'omics techniques in microalgal biotechnology will be discussed in the following section.

### 5.2.1. Protein expression can shed light on factors limiting growth in pure microalgal cultures

This comparative approach where oligotrophic and eutrophic conditions were put in direct contrast is directly analogous to studying two microalgal cultivation systems where metabolite production necessitates either nutrient poor (e.g. maximising lipids) or nutrient rich (e.g. maximising biomass) conditions. While some aspects of

these communities were only observed in one of the nutrient treatments, there were some patterns that were seen in both oligo- and eutrophic treatments. Several of the proteins identified in both nutrient treatments and time points are known to be expressed at an additional metabolic cost (i.e. only when strictly necessary). Therefore, the presence of these proteins provides clues regarding the physiological state of the organism and simultaneously, the characteristics of the surrounding environment. First, CCMs were found in both, cyanobacteria, carbxysomes found in *Anabaena* sp., and eukaryotic microalgae, a low-$CO_2$ inducible protein found in *Chlamydomonas* sp. The presence of these types of proteins suggests a C-limiting environment. Maximising biomass productivity was not the objective of this study, however, in a microalgal cultivation system the presence of these proteins would suggest the population would benefit from the addition of a C source. Second, ABC transporters were amongst the most abundant transport proteins identified. The particularity of these transporters is that they are mainly synthesised to enable the uptake of substrates that are limiting in the environment. Therefore, they can be very useful in identifying which compounds are limiting a microbial community. For example, Valladares et al. (2002) identified an ABC-type transporter with high-affinity for urea for which expression was controlled by N availability (i.e. expression was induced under N-limiting conditions). This type of analysis would be extremely useful for microalgal biotechnology, however, to date, not much is known about the ABC transporter superfamily in microalgae or cyanobacteria (Schulz et al., 2006; Andolfo et al., 2015). Im and Grossman (2002), demonstrated that an ABC transporter of the multidrug resistance-associated protein subfamily, in *C. reinhardtii*, is expressed in elevated light and low $CO_2$ levels under the control of Cia5, a protein that also regulates the low-$CO_2$ inducible protein and carbonic anhydrase identified in Chapter 3. In another study, Hanikenne et al. (2005) showed that a protein from the mitochondrion homolog subfamily of half-size ABC transporters was responsible for transporting phytochelatin-cadmium complexes from the cytoplasm to the vacuole thus, playing a major role in cadmium tolerance in *C. reinhardtii*. Further work is needed to identify and characterise the main ABC transporters in microalgae and exploit their potential for microalgal biotechnology.

### 5.2.2. Metaproteomics as a tool to develop microalgal cultivation with a community approach

The findings described above are of interest for pure microalgal cultures. However, several of the main findings in Chapter 3 can be useful to microalgal cultivation with a community approach. In oligotrophic conditions, Bacteroidetes expressed several proteins that suggest a role as fast-growing organisms specialised for the initial use of highly complex DOM (Fig. 3.3). Curiously, these proteins were not as prevalent in eutrophic conditions. This led to the hypothesis that nutrient limiting conditions stimulate the release of microalgal DOM which provides substrates for bacterial growth. Finally, it was also observed that Alpha- and Beta-proteobacteria captured different substrates from microalgal exudate (carbohydrates and amino acids, respectively) in what is, presumably, a mechanism to avoid direct competition (Fig. 3.3). This description of heterotrophic metabolism in the experimental microcosms provided several testable hypotheses for microalgal cultivation with a community approach.

First, Bacteroidetes, preferentially, attach themselves to the phycosphere and use this advantage to initially degrade HMW compounds. This is a relationship that has been described from a metaproteogenomic point of view (Williams et al., 2013) but how it could fit in a biotechnological scenario is still unclear. The fact that this bacterial group lives attached to the microalgal cell surface will make it very difficult to test how the same microalgal strain behaves with and without bacterial presence. Therefore, the available option for community design is to consider this bacterial group, and its microalgae host, as an inseparable biological unit. Second, if we consider that nutrient limiting conditions stimulate the release of microalgal DOM then this is a variable factor when designing a microbial community. The ideal microalgae to bacteria ratios will have to be determined both under nutrient rich and nutrient poor conditions in order to collect information about how this changes community dynamics. Furthermore, it is possible that the bacterial stimulation in nutrient limited conditions could lead to a point of imbalance and microalgal culture crash. Third, the observation that Alpha- and Beta-proteobacteria shared resources to maintain both populations opens up several avenues of possibilities for community design. One of the staples of a successful community are stabilising mechanisms that prevent competitive exclusion. Therefore, any two bacterial populations that are found to share resources are immediate candidates when designing a community. However, these relationships need to be thoroughly researched to detail

the mechanisms of co-existence. For example, changes in environmental conditions can lead to changes in the microalgal community and their exudation rate of specific dissolved organic monomers. This would inevitably alter the bacterial dynamics. Thus, individual bacterial group numbers have to be monitored over time and compared to substrate uptake patterns to understand how bacterial populations react to changes in the microalgal community. Pérez et al. (2015) used microautoradiography combined with catalysed reporter deposition fluorescent *in situ* hybridisation to assess resource partitioning among freshwater bacterial groups, from two oligotrophic lakes, using four types of dissolved organic substrates. The study concluded that substrate partitioning occurred between Actinobacteria, which were overrepresented in glucose and acetate utilisation, and Beta-proteobacteria, which dominated amino acid uptake. This illustrates another case of resource partitioning and suggests that this phenomenon might be ubiquitous between bacterial groups.

"Shotgun" 'omics studies have provided a wealth of information regarding algal-bacterial interactions, however, by expanding the use of microcosms to complex microbial communities, hypotheses can be tested in a controlled environment. For example, Pagaling et al. (2014) utilised pond sediment–water microcosms, based on the well-known Winogradsky column, to study microbial community history. The study found that when microbial communities colonise a new environment, final community composition is unpredictable. However, when the microbial community is pre-conditioned to its new habitat, community development is more reproducible. Due to its well-defined conditions, this study can easily be replicated and expanded upon in order to reduce unpredictability in the design of microbial communities.

Simultaneously, the field should focus on describing robust experimental model systems. Several of the model systems already in existence provide an insight into mechanisms such as bloom termination (*Emiliania huxleyi*/*Phaeobacter gallaeciensis* (Seyedsayamdost et al., 2011)), evolution of $N_2$ fixation (unicellular prymnesiophyte/N-fixing cyanobacterium (UCYN-A) (Thompson et al., 2012)), vitamin auxotrophy (*Lobomonas rostrata*/*M. loti* (Kazamia et al., 2012b)), programmed cell death (*Dunaliella salina*/*Halobacterium salinarum* (Orellana et al., 2013)) and mutualistic gene loss (*C. reinhardtii* $B_{12}$ auxotroph/$B_{12}$ producing bacteria (Helliwell et al., 2015)). With recently developed technologies such as single cell 'omics (Perkel, 2015) and stable isotope probing in combination with metaproteomics, exo-metabolomics (von Bergen et al., 2013) and metagenomics (Uhlik et al., 2013) an unprecedented view of nutrient and DOM fluxes can be achieved.

### 5.2.3. SEM and metaproteomics

This metaproteomic study provided evidence for many microbial processes and interactions under oligo- and eutrophic conditions, however, had little overlap with the findings from the SEM study in Chapter 2. There are several explanations for this. First, SEM confirmed that the experimental community in this study was sensitive to nutrient enrichment, however, the magnitude of the changes observed did not follow what was expected from a ten-fold increase in nutrient concentrations. The most likely explanation comes from the metaproteomic evidence of C limitation in the microbial community. If this was the case, the nutrient pool would have been assimilated as reserves instead of prioritising growth. In future studies, total and dissolved C concentrations must be considered essential parameters both from a modelling and metaproteomic point of view. Second, the SEM study concluded that, in the eutrophic treatment, physicochemical factors became superior predictors of microalgal and bacterial concentrations. In eutrophic conditions the metaproteomic data suggested that there was a shift in the microalgae-bacteria interaction as a result of the decrease in excretion of microalgal derived DOM. However, no evidence was found to support the direct effect of physiochemical parameters.

In conclusion, SEM and metaproteomics are complementary techniques that are able to look at a biological system from different perspectives without necessarily overlapping in their conclusions. The advantage of coupling both would come in the form of a continual update of the model based on conclusions derived from metaproteomic data. This could lead to the inclusion of more observed variables (e.g. C concentrations) or latent variables (e.g. separation of the latent variable bacteria into different phyla). The usage of SEM to study microbial systems is becoming increasingly popular (Lin et al., 2014; Schnecker et al., 2014; Li et al., 2015) and can contribute to providing a visual framework of microbial community function.

## 5.3. Response surface methodology to determine the biotechnological value of a high-lipid *C. reinhardtii* mutant strain

Studies such as the ones undertaken in Chapters 2 and 3 of this thesis can provide a better understanding of existing microbial systems and create novel

hypotheses for potential industrial applications in microalgal biotechnology. However, any move from laboratory-scale to pilot-scale must be accompanied by a feasibility study which encompasses not only a techno-economic study but also an environmental risk assessment. This is extremely relevant to pilot-scale LMVC production in microalgae where outdoor cultivation is prone to native organism invasion or non-native organism escape. Furthermore, the increasing utilisation of transgenic microalgal strains introduces more complexity to the regulatory assessment. The first step of any environmental risk assessment is to attempt to predict the ecological impact of a non-native microalgae spill in the local ecosystem. However, an accurate prediction would require in-depth knowledge of the intricacies of the ecological niches in the local ecosystem. This would be a near impossible task for most aquatic systems. Therefore, this discussion would be facilitated if it could be shown that the non-native alga has a severely reduced fitness when compared to its native counterpart. Or, in the case of this study, if it could be shown that the fitness of the transgenic alga was lower than the fitness of its WT. If this is confirmed then the environmental risk analysis needs only to consider the impact of cultivating the WT.

To further this objective, Chapter 4 focused on the use of a response surface to study the competitive dynamics between two *C. reinhardtii* strains, a wild-type (CC-124) and a high-lipid mutant (CC-4333), with differing levels of initial WT to mutant ratios and $NH_4Cl$ concentrations. This provided a comparison between the fitness of the mutant strain and its WT relative and, simultaneously, information on the combined effects of competition and nutrient stress on lipid production. The results showed that the *C. reinhardtii* mutant strain, in pure culture, at 50 mg $L^{-1}$ $NH_4Cl$, had a total amount of lipids of 1613 µg in $1.46 \times 10^5$ cells, after a stress phase of only 24 hours, compared to 1376 µg in $1.00 \times 10^6$ cells after six days of incubation. This suggested an opportunity for scaling-out, with a 24-hour stress period, to avoid loss of performance in scale-up and facilitate expansion and automation of the cultivation process. Scaling-out has been applied in the industry of fine chemicals (Pollak, 2011), vaccine production (Wen et al., 2014) and is now starting to be implemented in the microalgal industry (Lane, 2015).

Another objective of this study was to determine the competitive dynamics between both *C. reinhardtii* strains to assess the industrial risk of the mutant reverting to WT and the environmental risk of a mutant escape. From an industrial perspective reversion to WT can have a negative effect on product yields and, in the long term, lead to a complete loss of the mutation of interest. In this study, TAG yields, in the pure WT

culture, were 40% and 24% lower, at days 1 and 6, respectively, than in the pure mutant culture. Furthermore, results showed that, when grown together, the fitness of the mutant strain was lower than the fitness of its WT, in every treatment. Therefore, these results suggest that any percentage of reversion to WT can, initially, negatively affect overall productivity and ultimately, lead to a complete loss of the mutant strain. These observations should be taken into consideration when implementing a quality control system for industrial cultivation of *C. reinhardtii* high-lipid mutant strains. The next section will focus on the environmental risks of the escape of the *C. reinhardtii* mutant strain.

### 5.3.1. The risk of the escape of non-native microalgae

This study suggests that cultivating the *C. reinhardtii* high-lipid mutant should not pose a major risk for natural ecosystems, however, the dangers of releasing non-native microalgal species, whether transgenic or non-transgenic, are real and need to be addressed. Regarding the mass cultivation of non-transgenic strains (e.g. WT strains), these are typically model organisms that have been thoroughly described and regarding which accurate predictions can be made. However, transgenic strains will most likely present characteristics that facilitate cultivation and growth or traits that optimise metabolite production, both of which can affect strain fitness in unknown ways (Gressel et al., 2013). Generally, traits that increase culture robustness (e.g. secretion of algicides or bactericides, resistance to zooplankton) lead to a high potential environmental risk. On the other hand, traits that increase culture yields (e.g. enhanced lipid production during exponential growth) are less likely to pose environmental risks due to the nature of ecological trade-offs (e.g. species with higher cellular lipid concentrations show slower growth) (Shurin et al., 2013). Nevertheless, faced with the difficulties of predicting the impact of transgenic strains on natural ecosystems alternative solutions are necessary. For example, it is currently common practice to use cell wall-deficient strains for transformation due to a better transformation efficiency (Potvin & Zhang. 2010). This works in favour of ecosystem protection because studies have observed that grazer assimilation rates of cell wall-deficient strains can be up to four times higher than that of WT strains (Van Donk, 1997). Also, the *C. reinhardtii* high-lipid mutant strain, utilised in Chapter 4, lacks flagella. While this is not a negative aspect in pure culture it would affect its chances of survival in a natural ecosystem due to impaired phototaxis.

Finally, several authors have proposed gene mitigation strategies where genes are added or subtracted in order to severely reduce the fitness of the organism. This is currently a common technique in transgenic higher plants (Gressel, 1999; Al-Ahmad et al., 2006). An excellent example of this strategy in microalgae would consist in deleting essential C capture enzymes (e.g. carbonic anhydrase) which, as seen in Chapter 3, allow acclimation to low $CO_2$ environments. The absence of these enzymes would render the cell unable to survive in the ambient $CO_2$ conditions found in nature (Van et al., 2001; Brueggeman et al., 2012). In summary, this study lays the groundwork for the installation of regulatory procedures for pilot-scale microalgal cultivation.

### 5.3.2. The role of ecology in microalgal biotechnology

In recent years there have been many ecologists that have advocated for the application of ecological principles to improve operation and performance of LMVC production systems in microalgae (Shurin et al., 2013; Smith et al., 2015). One approach is based on microalgal trait analysis to understand ecophysiological trade-offs. For example, Shurin et al., 2013, found that amongst 16 microalgal species of broad taxonomic diversity the trade-offs were similar. A species which competes strongly for light would not compete strongly for nutrients and species that excelled in biomass production would not over-accumulate neutral lipids in the cells. These type of ecological trade-offs are widespread and can also be seen in the strains utilised in Chapter 4. When compared to its WT the *C. reinhardtii* mutant has a higher growth rate and carrying capacity but a smaller cell size and lower fitness. Also, the creation of this high-lipid mutant illustrates the consequences of ecological trade-offs for genetic engineering. The disruption of ADP-glucose pyrophosphorylase in the *C. reinhardtii* mutant conferred it the ability to over-accumulate lipids. However, this severely affected its ability to metabolise acetate and also decreased its photosynthetic activity which had an overall negative impact on fitness (Work et al., 2010). Another approach is to optimise the environment surrounding a microalgal strain to better fulfil its ideal requirements. For example, for several decades it has been known that microalgal growth not only depends on absolute concentrations of N and P but also on the ratio (N:P) between them (Smith, 1979; 1982). This is abundantly clear in the results from Chapter 4 of this dissertation. In both *C. reinhardtii* strains, K was significantly higher ($p < 0.05$) in 200 mg $L^{-1}$ $NH_4Cl$ than in 375 mg $L^{-1}$ $NH_4Cl$. Nevertheless, the popular

TAP medium recipe utilises $NH_4Cl$ concentrations of 375 mg $L^{-1}$ which shows that margin for optimisation exists even in well-established protocols. These examples pertain to laboratory scale *C. reinhardtii* cultures, however, the broad principles should be relevant to any mass microalgal culture.

To conclude, the overall aim of the thesis was to show how the application of ecological theory can be beneficial to the production of LMVCs in microalgae by utilising three different approaches with potential to be applied to pilot-scale microalgal cultivation. The integration of modelling, 'omics and traditional competition studies can provide a general overview of microbial dynamics in microalgal cultures. However, these approaches constitute only a subset of the potential techniques that can be transferred from aquatic ecology to microalgal biotechnology. As microalgal cultivation systems grow from pilot- to commercial-scale, the core principles of microalgal ecology and physiology can help contribute significantly to microalgal LMVC production.

# Chapter 6: Future directions

Based on the work developed in this thesis, there are several promising areas that can be further developed. As discussed in Chapters 2 and 5 (Fig. 5.1), SEM can be used to explore causal relationships in order to fully optimise conditions in microalgal cultures. This approach would also benefit from the inclusion of a high-throughput design of experiments (DOE) platform (e.g. response surface in Chapter 4) to screen the factors affecting biomass and target compound productivity. The results of such DOE experiments (i.e. what factors interact with each other) would then be fed back into SEM to improve the initial model. This feedback approach could significantly reduce the number of experiments required and tie both modelling and experimental approaches together.

Chapter 3 highlighted two microalgal-bacterial relationships that could be of use to synthetic ecology. First, several proteins were identified that suggest close ties between Bacteroidetes and microalgal species. Second, microalgal exudate in the system was supporting both Alpha- and Beta-proteobacteria, simultaneously, through resource partitioning. The strength of this relationship was determined by the nutrient status of the aquatic environment. While these insights derive from protein identification, the description of these relationships would benefit from a multi-omics characterisation. Genomics adds the potential to reveal the identities and gene pool of the microorganisms while transcriptomics determines which of these genes are expressed. These tie together with proteomics creating a detailed picture of the composition and activity of a microbial community. Furthermore, it would also be of interest to utilise ‘omics to delve deeper into the modes of microalgal-bacterial interaction. This is currently a very active field of research with studies targeting the exometabolome and volatile organic compounds in search of allelopathy (i.e. production of one or more biomolecules that affect the growth, survival, and/or reproduction of other organisms) (Silva et al., 2015; Tyc et al., 2015). If it is shown, for example, that bacterial allelochemicals can improve microalgal cultures these could, potentially, be characterised, purified (or synthesised) and utilised in microalgal cultures independently from the bacterial species that produce them.

Chapter 4 utilised RSM to study how intraspecific competition can affect lipid production but was also an attempt at introducing a biological safety study for pilot-scale cultivation of *C. reinhardtii* high-lipid mutants. Competing the high-lipid mutant against its WT revealed several fragilities inherent to the mutant strain, however, this simplistic study is still far from a full ecosystem risk assessment. Currently, there is much debate regarding the nature and extent of testing necessary to introduce GM

microalgal cultivation with some authors assuming an alarmist position based on *in silico* studies (Flynn et al., 2013) while others are proposing the creation of an *in vitro/in vivo* experimental framework for each GM strain (Gressel et al., 2013; Henley et al., 2013). Therefore, future work must consider developing experimental criteria for open pond microalgae testing that include potential ecological impacts and cultivation system/site specific protective measures. Preferably, this experimental framework will be developed to include all transgenic (GM and non-GM) and non-native microalgal cultivation systems that present risks to local populations and the natural environment.

# Chapter 7: References

Abram, F. (2015). Systems-based approaches to unravel multi-species microbial community functioning. *Computational and Structural Biotechnology Journal* 13**,** 24-32.

Adams, C., Godfrey, V., Wahlen, B., Seefeldt, L., and Bugbee, B. (2013). Understanding precision nitrogen stress to optimize the growth and lipid content tradeoff in oleaginous green microalgae. *Bioresource Technology* 131**,** 188-194.

Adarme-Vega, T.C., Lim, D.K., Timmins, M., Vernen, F., Li, Y., and Schenk, P.M. (2012). Microalgal biofactories: a promising approach towards sustainable omega-3 fatty acid production. *Microbial Cell Factories* 11**,** 96.

Al-Ahmad, H., and Gressel, J. (2006). Mitigation using a tandem construct containing a selectively unfit gene precludes establishment of *Brassica napus* transgenes in hybrids and backcrosses with weedy *Brassica rapa*. *Plant Biotechnology Journal* 4**,** 23-33.

Alexova, R., Haynes, P.A., Ferrari, B.C., and Neilan, B.A. (2011). Comparative protein expression in different strains of the bloom-forming cyanobacterium *Microcystis aeruginosa*. *Molecular & Cellular Proteomics* 10**,** M110 003749.

Andersen, R.A. (2005). *Algal Culturing Techniques.* Elsevier/Academic Press.

Anderson, D.M., Cembella, A.D., and Hallegraeff, G.M. (2012). Progress in understanding harmful algal blooms: paradigm shifts and new technologies for research, monitoring, and management. *Annual Review of Marine Science* 4**,** 143-176.

Andolfo, G., Ruocco, M., Di Donato, A., Frusciante, L., Lorito, M., Scala, F., and Ercolano, M.R. (2015). Genetic variability and evolutionary diversification of membrane ABC transporters in plants. *BMC Plant Biology* 15**,** 1-15.

Arhonditsis, G.B., Stow, C.A., Steinberg, L.J., Kenney, M.A., Lathrop, R.C., McBride, S.J., and Reckhow, K.H. (2006). Exploring ecological patterns with structural equation modeling and Bayesian analysis. *Ecological Modelling* 192**,** 385-409.

Atamna-Ismaeel, N., Sabehi, G., Sharon, I., Witzel, K.-P., Labrenz, M., Jurgens, K., Barkay, T., Stomp, M., Huisman, J., and Beja, O. (2008). Widespread distribution of proteorhodopsins in freshwater and brackish ecosystems. *The ISME Journal* 2**,** 656-662.

Azam, F., Fenchel, T., Field, J.G., Gray, J.S., Meyer-Reil, L.A., and Thingstad, F. (1983). The ecological role of water-column microbes in the sea. *Marine Ecology - Progress Series* 10**,** 257 - 263.

Azam, F., and Malfatti, F. (2007). Microbial structuring of marine ecosystems. *Nature Reviews Microbiology* 5**,** 782-791.

Bachelot, B., Uriarte, M., and McGuire, K. (2015). Interactions among mutualism, competition, and predation foster species coexistence in diverse communities. *Theoretical Ecology* 8**,** 297-312.

Bailey-Watts, A.E. (1992). Growth and reproductive strategies of freshwater phytoplankton. *Trends in Ecology & Evolution* 4**,** 359.

Ball, S., Marianne, T., Dirick, L., Fresnoy, M., Delrue, B., and Decq, A. (1991). A *Chlamydomonas reinhardtii* low-starch mutant is defective for 3-phosphoglycerate activation and orthophosphate inhibition of ADP-glucose pyrophosphorylase. *Planta* 185**,** 17-26.

Benemann, J. (2013). Microalgae for Biofuels and Animal Feeds. *Energies* 6**,** 5869-5886.

Beutler, M., Wiltshire, K.H., Meyer, B., Moldaenke, C., Lüring, C., Meyerhöfer, M., Hansen, U.-P., and Dau, H. (2002). A fluorometric method for the differentiation of algal populations *in vivo* and *in situ*. *Photosynthesis Research* 72**,** 39-53.

Bird, D.F., and Kalff, J. (1984). Empirical relationships between bacterial abundance

and chlorophyll concentration in fresh and marine waters. *Canadian Journal of Fisheries and Aquatic Sciences* 41**,** 1015-1023.

Blanvillain, S., Meyer, D., Boulanger, A., Lautier, M., Guynet, C., Denance, N., Vasse, J., Lauber, E., and Arlat, M. (2007). Plant carbohydrate scavenging through tonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. *PLoS One* 2**,** e224.

Bokulich, N., Rideout, J., Kopylova, E., Bolyen, E., Patnode, J., Ellett, Z., McDonald, D., Wolfe, B., Maurice, C., Dutton, R., Turnbaugh, P., Knight, R., and Caporaso, J. (2015). A standardized, extensible framework for optimizing classification improves marker-gene taxonomic assignments. *PeerJ PrePrints* 3.

Brennan, L., and Owende, P. (2010). Biofuels from microalgae—A review of technologies for production, processing, and extractions of biofuels and co-products. *Renewable and Sustainable Energy Reviews* 14**,** 557-577.

Brueggeman, A.J., Gangadharaiah, D.S., Cserhati, M.F., Casero, D., Weeks, D.P., and Ladunga, I. (2012). Activation of the carbon concentrating mechanism by $CO_2$ deprivation coincides with massive transcriptional restructuring in *Chlamydomonas reinhardtii*. *The Plant Cell* 24**,** 1860-1875.

BSI (2004). "BS EN ISO 6878:2004: Water quality. Determination of phosphorus. Ammonium molybdate spectrometric method". London. British Standards Institution.

Buchan, A., LeCleir, G.R., Gulvik, C.A., and Gonzalez, J.M. (2014). Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nature Reviews Microbiology* 12**,** 686-698.

Carey, C.C., Ibelings, B.W., Hoffmann, E.P., Hamilton, D.P., and Brookes, J.D. (2012). Eco-physiological adaptations that favour freshwater cyanobacteria in a changing climate. *Water Research* 46**,** 1394-1407.

Carpenter, S.R., Kitchell, J.F., Hodgson, J.R., Cochran, P.A., Elser, J.J., Elser, M.M., Lodge, D.M., Kretchmer, D., He, X., and Vonende, C.N. (1987). Regulation of Lake Primary Productivity by Food Web Structure. *Ecology* 68**,** 1863-1876.

Carvalho, A.P., Meireles, L.A., and Malcata, F.X. (2006). Microalgal reactors: a review of enclosed system designs and performances. *Biotechnology Progress* 22**,** 1490-1506.

Chen, M., Liu, T., Chen, X., Chen, L., Zhang, W., Wang, J., Gao, L., Chen, Y., and Peng, X. (2012). Subcritical co-solvents extraction of lipid from wet microalgae pastes of *Nannochloropsis* sp. *European Journal of Lipid Science and Technology* 114**,** 205-212.

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual Review of Ecology, Evolution, and Systematics* 31**,** 343-366.

Chisti, Y. (2007). Biodiesel from microalgae. *Biotechnology Advances* 25**,** 294-306.

Clements, F.E. (1916). *Plant succession: an analysis of the development of vegetation.* Washington. Carnegie Institution of Washington.

Cole, J.J. (1982). Interactions between bacteria and algae in aquatic ecosystems. *Annual Review of Ecology, Evolution, and Systematics* 13**,** 291-314.

Converti, A., Casazza, A.A., Ortiz, E.Y., Perego, P., and Del Borghi, M. (2009). Effect of temperature and nitrogen concentration on the growth and lipid content of *Nannochloropsis oculata* and *Chlorella vulgaris* for biodiesel production. *Chemical Engineering and Processing: Process Intensification* 48**,** 1146-1151.

Cottingham, K.L., and Carpenter, S.R. (1998). Population, community, and ecosystem variates as ecological indicators: Phytoplankton responses to whole-lake enrichment. *Ecological Applications* 8**,** 508-530.

Davies, D.R., and Plaskitt, A. (1971). Genetical and structural analyses of cell-wall formation in *Chlamydomonas reinhardi*. *Genetics Research* 17**,** 33-43.

De Senerpont Domis, L.N., Van de Waal, D.B., Helmsing, N.R., Van Donk, E., and Mooij, W.M. (2014). Community stoichiometry in a changing world: combined effects of warming and eutrophication on phytoplankton dynamics. *Ecology* 95**,** 1485-1495.

Diz, A.P., Truebano, M., and Skibinski, D.O. (2009). The consequences of sample pooling in proteomics: an empirical study. *Electrophoresis* 30**,** 2967-2975.

Dong, H.-P., Hong, Y.-G., Lu, S., and Xie, L.-Y. (2014). Metaproteomics reveals the major microbial players and their biogeochemical functions in a productive coastal system in the northern South China Sea. *Environmental Microbiology Reports* 6**,** 683-695.

Downing, J.A., Osenberg, C.W., and Sarnelle, O. (1999). Meta-analysis of marine nutrient-enrichment experiments: Variation in the magnitude of nutrient limitation. *Ecology* 80**,** 1157-1167.

Draaisma, R.B., Wijffels, R.H., Slegers, P.M., Brentner, L.B., Roy, A., and Barbosa, M.J. (2013). Food commodities from microalgae. *Current Opinion in Biotechnology* 24**,** 169-177.

Dunham, M.J. (2007). Synthetic ecology: a model system for cooperation. *Proceedings of the National Academy of Science* 104**,** 1741-1742.

Eckelberry, R. (2011). *Algae - Food or Chemical Grade?* [Online]. http://www.algaeindustrymagazine.com/algae-business-algae-food-or-chemical-grade/. [Accessed 01 October 2015].

Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26**,** 2460-2461.

Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* 10**,** 996-998.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27**,** 2194-2200.

Edwards, K.F., Klausmeier, C.A., and Litchman, E. (2011). Evidence for a three-way trade-off between nitrogen and phosphorus competitive abilities and cell size in phytoplankton. *Ecology* 92**,** 2085-2095.

Eiler, A. (2006). Evidence for the ubiquity of mixotrophic bacteria in the upper ocean: Implications and consequences. *Applied and Environmental Microbiology* 72**,** 7431-7437.

Eiler, A., and Bertilsson, S. (2007). Flavobacteria blooms in four eutrophic lakes: Linking population dynamics of freshwater bacterioplankton to resource availability. *Applied and Environmental Microbiology* 73**,** 3511-3518.

Elliott, J.A. (2012). Is the future blue-green? A review of the current model predictions of how climate change could affect pelagic freshwater cyanobacteria. *Water Research* 46**,** 1364-1371.

Elser, J.J., Bracken, M.E., Cleland, E.E., Gruner, D.S., Harpole, W.S., Hillebrand, H., Ngai, J.T., Seabloom, E.W., Shurin, J.B., and Smith, J.E. (2007). Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecology Letters* 10**,** 1135-1142.

Enzing, C., Ploeg, M., Barbosa, M., and Sijtsma, L. (2014). Microalgae-based products for the food and feed sector: An outlook for europe.

Escalante, A.E., Rebolleda-Gomez, M., Benitez, M., and Travisano, M. (2015). Ecological perspectives on synthetic biology: insights from microbial population biology. *Frontiers in Microbiology* 6**,** 143.

Falkowski, P., Scholes, R.J., Boyle, E., Canadell, J., Canfield, D., Elser, J., Gruber, N., Hibbard, K., Högberg, P., Linder, S., Mackenzie, F.T., Moore III, B., Pedersen, T., Rosenthal, Y., Seitzinger, S., Smetacek, V., and Steffen, W. (2000). The

global carbon cycle: A test of our knowledge of earth as a system. *Science* 290**,** 291-296.

Fargione, J.E., and Tilman, D. (2005). Diversity decreases invasion via both sampling and complementarity effects. *Ecology Letters* 8**,** 604 - 611.

Fenchel, T. (2008). The microbial loop – 25 years later. *Journal of Experimental Marine Biology and Ecology* 366**,** 99-103.

Fernandez-Gomez, B., Richter, M., Schuler, M., Pinhassi, J., Acinas, S.G., Gonzalez, J.M., and Pedros-Alio, C. (2013). Ecology of marine Bacteroidetes: a comparative genomics approach. *The ISME Journal* 7**,** 1026-1037.

Flynn, K.J., Greenwell, H.C., Lovitt, R.W., and Shields, R.J. (2010). Selection for fitness at the individual or population levels: modelling effects of genetic modifications in microalgae on productivity and environmental safety. *Journal of Theoretical Biology* 263**,** 269-280.

Flynn, K.J., Mitra, A., Greenwell, H.C., and Sui, J. (2013). Monster potential meets potential monster: pros and cons of deploying genetically modified microalgae for biofuels production. *Interface Focus* 3**,** 20120037.

Geissmann, Q. (2013). OpenCFU, a new free and open-source software to count cell colonies and other circular objects. *PLoS ONE* 8**,** e54072.

Georges, A.A., El-Swais, H., Craig, S.E., Li, W.K., and Walsh, D.A. (2014). Metaproteomic analysis of a winter to spring succession in coastal northwest Atlantic Ocean microbial plankton. *The ISME Journal* 8**,** 1301-1313.

Gilbert, H.J. (2008). Sus out sugars in. *Structure* 16**,** 987-989.

Giovannoni, S., and Stingl, U. (2007). The importance of culturing bacterioplankton in the 'omics' age. *Nature Reviews Microbiology* 5**,** 820-826.

González, J.M., Fernández-Gómez, B., Fernàndez-Guerra, A., Gómez-Consarnau, L., Sánchez, O., Coll-Lladó, M., del Campo, J., Escudero, L., Rodríguez-Martínez, R., Alonso-Sáez, L., Latasa, M., Paulsen, I., Nedashkovskaya, O., Lekunberri, I., Pinhassi, J., and Pedrós-Alió, C. (2008). Genome analysis of the proteorhodopsin-containing marine bacterium *Polaribacter* sp. MED152 (Flavobacteria). *Proceedings of the National Academy of Sciences* 105**,** 8724-8729.

Gorman, D.S., and Levine, R.P. (1965). Cytochrome f and plastocyanin: their sequence in the photosynthetic electron transport chain of *Chlamydomonas reinhardi*. *Proceedings of the National Academy of Sciences* 54**,** 1665-1669.

Graham, E.R., Chen, H., Dziallas, C., Fernández-Gómez, B., and Kirkpatrick, J.B. (Year). "A more inclusive loop: Examining the contribution of five bacterial specialists to nutrient cycling and the microbial loop", ed. P.F. Kemp. ASLO, 46 - 68.

Greenwell, H.C., Laurens, L.M., Shields, R.J., Lovitt, R.W., and Flynn, K.J. (2010). Placing microalgae on the biofuels priority list: a review of the technological challenges. *Journal of The Royal Society Interface* 7**,** 703-726.

Gressel, J. (1999). Tandem constructs: preventing the rise of superweeds. *Trends in Biotechnology* 17**,** 361-366.

Gressel, J., van der Vlugt, C.J.B., and Bergmans, H.E.N. (2013). Environmental risks of large scale cultivation of microalgae: Mitigation of spills. *Algal Research* 2**,** 286-298.

Grobbelaar, J.U. (2013). "Inorganic Algal Nutrition," in *Handbook of Microalgal Culture*. John Wiley & Sons, Ltd, 123-133.

Grossart, H.-P. (2010). Ecological consequences of bacterioplankton lifestyles: changes in concepts are needed. *Environmental Microbiology Reports* 2**,** 706-714.

Grossart, H.-P., and Simon, M. (2007). Interactions of planktonic algae and bacteria: effects on algal growth and organic matter dynamics. *Aquatic Microbial Ecology*

47**,** 163-176.

Grosskopf, T., and Soyer, O.S. (2014). Synthetic microbial communities. *Current Opinion in Biotechnology* 18**,** 72-77.

Grossman, A.R., Schaefer, M.R., Chiang, G.G., and Collier, J.L. (1993). Environmental effects on the light-harvesting complex of cyanobacteria. *Journal of Bacteriology* 175**,** 575-582.

Guenet, B., Danger, M., Abbadie, L., and Lacroix, G. (2010). Priming effect: bridging the gap between terrestrial and aquatic ecology. *Ecology* 91**,** 2850-2861.

Guiry, M.D. (2012). How many species of algae are there? *Journal of Phycology* 48**,** 1057-1063.

Haas, M.J., and Wagner, K. (2011). Simplifying biodiesel production: The direct or *in situ* transesterification of algal biomass. *European Journal of Lipid Science and Technology* 113**,** 1219-1229.

Hanson, B.T., Hewson, I., and Madsen, E.L. (2014). Metaproteomic survey of six aquatic habitats: discovering the identities of microbial populations active in biogeochemical cycling. *Microbial Ecology* 67**,** 520-539.

Hariskos, I., and Posten, C. (2014). Biorefinery of microalgae - opportunities and constraints for different production scenarios. *Biotechnology Journal* 9**,** 739-752.

Haukka, K., Kolmonen, E., Hyder, R., Hietala, J., Vakkilainen, K., Kairesalo, T., Haario, H., and Sivonen, K. (2006). Effect of nutrient loading on bacterioplankton community composition in lake mesocosms. *Microbial Ecology* 51**,** 137-146.

Heisler, J., Glibert, P.M., Burkholder, J.M., Anderson, D.M., Cochlan, W., Dennison, W.C., Dortch, Q., Gobler, C.J., Heil, C.A., Humphries, E., Lewitus, A., Magnien, R., Marshall, H.G., Sellner, K., Stockwell, D.A., Stoecker, D.K., and Suddleson, M. (2008). Eutrophication and harmful algal blooms: A scientific consensus. *Harmful Algae* 8**,** 3-13.

Helliwell, K.E., Collins, S., Kazamia, E., Purton, S., Wheeler, G.L., and Smith, A.G. (2015). Fundamental shift in vitamin $B_{12}$ eco-physiology of a model alga demonstrated by experimental evolution. *The ISME Journal* 9**,** 1446-1455.

Henley, W.J., Litaker, R.W., Novoveská, L., Duke, C.S., Quemada, H.D., and Sayre, R.T. (2013). Initial risk assessment of genetically modified (GM) microalgae for commodity-scale biofuel cultivation. *Algal Research* 2**,** 66-77.

Hettich, R.L., Pan, C., Chourey, K., and Giannone, R.J. (2013). Metaproteomics: Harnessing the power of high performance mass spectrometry to identify the suite of proteins that control metabolic activities in microbial communities. *Analytical Chemistry* 85**,** 4203-4214.

HilleRisLambers, J., Adler, P.B., Harpole, W.S., Levine, J.M., and Mayfield, M.M. (2012). Rethinking community assembly through the lens of coexistence theory. *Annual Review of Ecology, Evolution, and Systematics* 43**,** 227-248.

Ho, S.-H., Chen, C.-Y., and Chang, J.-S. (2012). Effect of light intensity and nitrogen starvation on $CO_2$ fixation and lipid/carbohydrate production of an indigenous microalga *Scenedesmus obliquus* CNW-N. *Bioresource Technology* 113**,** 244-252.

Hobbie, J.E., Holm-Hansen, O., Packard, T.T., Pomeroy, L.R., Sheldon, R.W., Thomas, J.P., and Wiebe, W.J. (1972). A study of the distribution and activity of microorganisms in ocean water. *Limnology and Oceanography* 17**,** 544-555.

Hooper, D.U., Chapin, F.S., Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J.H., Lodge, D.M., Loreau, M., Naeem, S., Schmid, B., Setala, H., Symstad, A.J., Vandermeer, J., and Wardle, D.A. (2005). Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological*

*Monographs* 75**,** 3-35.

Hu, Q., Sommerfeld, M., Jarvis, E., Ghirardi, M., Posewitz, M., Seibert, M., and Darzins, A. (2008). Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *The Plant Journal* 54**,** 621-639.

Hultman, J., Waldrop, M.P., Mackelprang, R., David, M.M., McFarland, J., Blazewicz, S.J., Harden, J., Turetsky, M.R., McGuire, A.D., Shah, M.B., VerBerkmoes, N.C., Lee, L.H., Mavrommatis, K., and Jansson, J.K. (2015). Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature* 521**,** 208-212.

Im, C.S., and Grossman, A.R. (2002). Identification and regulation of high light-induced genes in *Chlamydomonas reinhardtii*. *The Plant Journal* 30**,** 301-313.

Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., and Mann, M. (2005). Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Molecular & Cellular Proteomics* 4**,** 1265-1272.

Ismail, A. (2010). Marine lipids overview: markets, regulation, and the value chain. *OCL* 17**,** 205-208.

Jagtap, P., Goslinga, J., Kooren, J.A., McGowan, T., Wroblewski, M.S., Seymour, S.L., and Griffin, T.J. (2013). A two-step database search method improves sensitivity in peptide sequence matches for metaproteomics and proteogenomics studies. *Proteomics* 13**,** 1352-1357.

James, G.O., Hocart, C.H., Hillier, W., Chen, H., Kordbacheh, F., Price, G.D., and Djordjevic, M.A. (2011). Fatty acid profiling of *Chlamydomonas reinhardtii* under nitrogen deprivation. *Bioresource Technology* 102**,** 3343-3351.

Jinkerson, R.E., and Jonikas, M.C. (2015). Molecular techniques to interrogate and edit the Chlamydomonas nuclear genome. *The Plant Journal* 82**,** 393-412.

Johnk, K.D., Huisman, J.E.F., Sharples, J., Sommeijer, B.E.N., Visser, P.M., and Stroom, J.M. (2008). Summer heatwaves promote blooms of harmful cyanobacteria. *Global Change Biology* 14**,** 495-512.

Kalb, J., V. F., and Bernlohr, R.W. (1977). A new spectrophotometric assay for protein in cell extracts. *Analytical Biochemistry* 82**,** 362-371.

Kazamia, E., Aldridge, D.C., and Smith, A.G. (2012). Synthetic ecology – A way forward for sustainable algal biofuel production? *Journal of Biotechnology* 162**,** 163-169.

Kazamia, E., Czesnick, H., Nguyen, T.T.V., Croft, M.T., Sherwood, E., Sasso, S., Hodson, S.J., Warren, M.J., and Smith, A.G. (2012). Mutualistic interactions between vitamin $B_{12}$-dependent algae and heterotrophic bacteria exhibit regulation. *Environmental Microbiology* 14**,** 1462-2920

Khan, F.A., and Ansari, A.A. (2005). Eutrophication: An ecological vision. *Botanical Review* 71**,** 449-482.

Klein-Marcuschamer, D., Chisti, Y., Benemann, J.R., and Lewis, D. (2013). A matter of detail: Assessing the true potential of microalgal biofuels. *Biotechnology and Bioengineering* 110**,** 2317-2322.

Kosten, S., Huszar, V.L.M., Bécares, E., Costa, L.S., van Donk, E., Hansson, L.-A., Jeppesen, E., Kruk, C., Lacerot, G., Mazzeo, N., De Meester, L., Moss, B., Lürling, M., Nõges, T., Romo, S., and Scheffer, M. (2012). Warmer climates boost cyanobacterial dominance in shallow lakes. *Global Change Biology* 18**,** 118-126.

Lalibertè, G., and de la Noüie, J. (1993). Auto-, hetero-, and mixotrophic growth of *Chlamydomonas humicola* on acetate. *Journal of Phycology* 29**,** 612-620.

Lane, J. (2015). *Joule says "will go commercial in 2017": solar fuels on the way* [Online]. Biofuels Digest. Available:

http://www.biofuelsdigest.com/bdigest/2015/03/23/joule-says-will-go-commercial-in-2017-solar-fuels-on-the-way/ [Accessed 29 March 2016].

Laurens, L.M.L., Nagle, N., Davis, R., Sweeney, N., Van Wychen, S., Lowell, A., and Pienkos, P.T. (2015). Acid-catalyzed algal biomass pretreatment for integrated lipid and carbohydrate-based biofuels production. *Green Chemistry* 17**,** 1145-1158.

Lauro, F.M., DeMaere, M.Z., Yau, S., Brown, M.V., Ng, C., Wilkins, D., Raftery, M.J., Gibson, J.A.E., Andrews-Pfannkoch, C., Lewis, M., Hoffman, J.M., Thomas, T., and Cavicchioli, R. (2011). An integrative study of a meromictic lake ecosystem in Antarctica. *The ISME Journal* 5**,** 879-895.

Leite, G.B., Abdelaziz, A.E.M., and Hallenbeck, P.C. (2013). Algal biofuels: Challenges and opportunities. *Bioresource Technology* 145**,** 134-141.

Lenth, R.V. (2009). Response-surface methods in R, Using rsm. *Journal of Statistical Software* 1.

Li, J., Ma, Y.-B., Hu, H.-W., Wang, J.-T., Liu, Y.-R., and He, J.-Z. (2015). Field-based evidence for consistent responses of bacterial communities to copper contamination in two contrasting agricultural soils. *Frontiers in Microbiology* 6**,** 31.

Li, Y., Han, D., Hu, G., Sommerfeld, M., and Hu, Q. (2010). Inhibition of starch synthesis results in overproduction of lipids in *Chlamydomonas reinhardtii*. *Biotechnology and Bioengineering* 107**,** 258-268.

Lin, X., Tfaily, M.M., Steinweg, J.M., Chanton, P., Esson, K., Yang, Z.K., Chanton, J.P., Cooper, W., Schadt, C.W., and Kostka, J.E. (2014). Microbial community stratification linked to utilization of carbohydrates and phosphorus limitation in a boreal peatland at Marcell Experimental Forest, Minnesota, USA. *Applied and Environmental Microbiology* 80**,** 3518-3530.

Lindeman, R.L. (1942). The trophic-dynamic aspect of ecology. *Ecology* 23**,** 399-417.

Longworth, J., Noirel, J., Pandhal, J., Wright, P.C., and Vaidyanathan, S. (2012). HILIC- and SCX-based quantitative proteomics of *Chlamydomonas reinhardtii* during nitrogen starvation induced lipid and carbohydrate accumulation. *Journal of Proteome Research* 11**,** 5959-5971.

Mann, A.J., Hahnke, R.L., Huang, S., Werner, J., Xing, P., Barbeyron, T., Huettel, B., Stuber, K., Reinhardt, R., Harder, J., Glockner, F.O., Amann, R.I., and Teeling, H. (2013). The genome of the alga-associated marine flavobacterium *Formosa agariphila* KMM 3901T reveals a broad potential for degradation of algal polysaccharides. *Applied and Environmental Microbiology* 79**,** 6813-6822.

Marketsandmarkets.com (2014). "Synthetic biology market by tool (xna, chassis, oligos, enzymes, cloning kits), technology (bioinformatics, nanotechnology, gene synthesis, cloning & sequencing), application (biofuels, pharmaceuticals, biomaterials, bioremediation) - global forecast to 2018". India.

McBride, M.J. (2001). Bacterial Gliding Motility: Multiple Mechanisms for Cell Movement over Surfaces. *Annual Review of Microbiology* 55**,** 49-75.

Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Marechal-Drouard, L., Marshall, W.F., Qu, L.H., Nelson, D.R., Sanderfoot, A.A., Spalding, M.H., Kapitonov, V.V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S.M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C.L., Cognat, V., Croft, M.T., Dent, R., Dutcher, S., Fernandez, E., Fukuzawa, H., Gonzalez-Ballester, D., Gonzalez-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P.A., Lemaire, S.D., Lobanov, A.V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J.V., Moseley, J., Napoli, C., Nedelcu,

A.M., Niyogi, K., Novoselov, S.V., Paulsen, I.T., Pazour, G., Purton, S., Ral, J.P., Riano-Pachon, D.M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S.L., Allmer, J., Balk, J., Bisova, K., Chen, C.J., Elias, M., Gendler, K., Hauser, C., Lamb, M.R., Ledford, H., Long, J.C., Minagawa, J., Page, M.D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A.M., Yang, P., Ball, S., Bowler, C., Dieckmann, C.L., Gladyshev, V.N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R.T., Brokstein, P., et al. (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318**,** 245-250.

Milledge, J.J. (2012). Microalgae - commercial potential for fuel, food and feed. *Food Science & Technology* 26**,** 26-28.

Mitra, A., Flynn, K.J., Burkholder, J.M., Berge, T., Calbet, A., Raven, J.A., Granéli, E., Glibert, P.M., Hansen, P.J., Stoecker, D.K., Thingstad, F., Tillmann, U., Våge, S., Wilken, S., and Zubkov, M.V. (2014). The role of mixotrophic protists in the biological carbon pump. *Biogeosciences* 11**,** 995-1005.

Mitri, S., and Richard Foster, K. (2013). The genotypic view of social interactions in microbial communities. *Annual Review of Genetics* 47**,** 247-273.

Molina Grima, E., Belarbi, E.H., Acién Fernández, F.G., Robles Medina, A., and Chisti, Y. (2003). Recovery of microalgal biomass and metabolites: process options and economics. *Biotechnology Advances* 20**,** 491-515.

Morin, P.J. (2011). *Community Ecology*. Wiley.

Moss, B. (1973). The influence of environmental factors on the distribution of freshwater algae: An experimental study: III. Effects of temperature, vitamin requirements and inorganic nitrogen compounds on growth. *Journal of Ecology* 61**,** 179-192.

Muth, T., Kolmeder, C.A., Salojärvi, J., Keskitalo, S., Varjosalo, M., Verdam, F.J., Rensen, S.S., Reichl, U., de Vos, W.M., Rapp, E., and Martens, L. (2015). Navigating through metaproteomics data: A logbook of database searching. *Proteomics* 15**,** 3439-3453.

Nalley, J.O., Stockenreiter, M., and Litchman, E. (2014). Community ecology of algal biofuels: Complementarity and trait-based approaches. *Industrial Biotechnology* 10**,** 191-201.

Nan, B., and Zusman, D.R. (2011). Uncovering the mystery of gliding motility in the Myxobacteria. *Annual Review of Genetics* 45**,** 21-39.

Newton, R.J., Jones, S.E., Eiler, A., McMahon, K.D., and Bertilsson, S. (2011). A guide to the natural history of freshwater lake bacteria. *Microbiology and Molecular Biology Reviews* 75**,** 14-49.

Ng, C., DeMaere, M.Z., Williams, T.J., Lauro, F.M., Raftery, M., Gibson, J.A., Andrews-Pfannkoch, C., Lewis, M., Hoffman, J.M., Thomas, T., and Cavicchioli, R. (2010). Metaproteogenomic analysis of a dominant green sulfur bacterium from Ace Lake, Antarctica. *The ISME Journal* 4**,** 1002-1019.

Nimick, D.A., Gammons, C.H., and Parker, S.R. (2011). Diel biogeochemical processes and their effect on the aqueous chemistry of streams: A review. *Chemical Geology* 283**,** 3-17.

Nobel, P.S. (2005). "5 - Photochemistry of Photosynthesis," in *Physicochemical and Environmental Plant Physiology (Third Edition)*, ed. P.S. Nobel. Burlington. Academic Press, 219-266.

Norsker, N.H., Barbosa, M.J., Vermue, M.H., and Wijffels, R.H. (2011). Microalgal production - a close look at the economics. *Biotechnology Advances* 29**,** 24-27.

O'Neil, J.M., Davis, T.W., Burford, M.A., and Gobler, C.J. (2012). The rise of harmful cyanobacteria blooms: The potential roles of eutrophication and climate change.

*Harmful Algae* 14**,** 313-334.
Orellana, M.V., Pang, W.L., Durand, P.M., Whitehead, K., and Baliga, N.S. (2013). A role for programmed cell death in the microbial loop. *PLoS ONE* 8**,** e62595.
Oswald, W.J., and Gotaas, H.B. (1957). Photosynthesis in sewage treatment. *Transactions of the American Society of Civil Engineers* 2849**,** 73 - 105.
Paerl, H.W., Hall, N.S., and Calandrino, E.S. (2011). Controlling harmful cyanobacterial blooms in a world experiencing anthropogenic and climatic-induced change. *Science of The Total Environment* 409**,** 1739-1745.
Paerl, H.W., and Huisman, J. (2008). Blooms like it hot. *Science* 320**,** 57-58.
Pagaling, E., Strathdee, F., Spears, B.M., Cates, M.E., Allen, R.J., and Free, A. (2014). Community history affects the predictability of microbial ecosystem development. *The ISME Journal* 8**,** 19-30.
Park, Y., Je, K.-W., Lee, K., Jung, S.-E., and Choi, T.-J. (2008). Growth promotion of *Chlorella ellipsoidea* by co-inoculation with *Brevundimonas* sp. isolated from the microalga. *Hydrobiologia* 598**,** 219-228.
Paver, S.F., Youngblut, N.D., Whitaker, R.J., and Kent, A.D. (2015). Phytoplankton succession affects the composition of Polynucleobacter subtypes in humic lakes. *Environmental Microbiology* 17**,** 816-828.
Pérez, M.T., Rofner, C., and Sommaruga, R. (2015). Dissolved organic monomer partitioning among bacterial groups in two oligotrophic lakes. *Environmental Microbiology Reports* 7**,** 265-272.
Perkel, J.M. (2015). *Single-cell biology: The power of one* [Online]. Science. [Accessed 10 July 2016].
Ploug, H., Adam, B., Musat, N., Kalvelage, T., Lavik, G., Wolf-Gladrow, D., and Kuypers, M.M. (2011). Carbon, nitrogen and $O_2$ fluxes associated with the cyanobacterium *Nodularia spumigena* in the Baltic Sea. *The ISME Journal* 5**,** 1549-1558.
Pollak, P. (2011). *Fine Chemicals: The Industry and the Business.* Wiley.
Pomeroy, L.R. (1974). The ocean's food web, a changing paradigm. *BioScience* 24**,** 499-504.
Pretty, J.N., Mason, C.F., Nedwell, D.B., Hine, R.E., Leaf, S., and Dils, R. (2003). Environmental costs of freshwater eutrophication in England and Wales. *Environmental Science & Technology* 37**,** 201-208.
Pröschold, T., Harris, E.H., and Coleman, A.W. (2005). Portrait of a species: *Chlamydomonas reinhardtii*. *Genetics* 170**,** 1601-1610.
R Development Core Team (2014). "R: A language and environment for statistical computing". Vienna, Austria. R Foundation for Statistical Computing.
Rast, W., and Thornton, J.A. (1996). Trends in eutrophication research and control. *Hydrological Processes* 10**,** 295-313.
Redfield, A.C. (1958). The biological control of chemical factors in the environment. *American Scientist* 46**,** 230A-221.
Rigosi, A., Carey, C.C., Ibelings, B.W., and Brookes, J.D. (2014). The interaction between climate warming and eutrophication to promote cyanobacteria is dependent on trophic state and varies among taxa. *Limnology and Oceanography* 59**,** 99-114.
Rodolfi, L., Chini Zittelli, G., Bassi, N., Padovani, G., Biondi, N., Bonini, G., and Tredici, M.R. (2009). Microalgae for oil: strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. *Biotechnology and Bioengineering* 102**,** 100-112.
Rohwer, F., and Thurber, R.V. (2009). Viruses manipulate the marine environment. *Nature* 459**,** 207-212.
Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of*

*Statistical Software* 48**,** 1-36.

Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S., Wu, D., Eisen, J.A., Hoffman, J.M., Remington, K., Beeson, K., Tran, B., Smith, H., Baden-Tillson, H., Stewart, C., Thorpe, J., Freeman, J., Andrews-Pfannkoch, C., Venter, J.E., Li, K., Kravitz, S., Heidelberg, J.F., Utterback, T., Rogers, Y.H., Falcon, L.I., Souza, V., Bonilla-Rosso, G., Eguiarte, L.E., Karl, D.M., Sathyendranath, S., Platt, T., Bermingham, E., Gallardo, V., Tamayo-Castillo, G., Ferrari, M.R., Strausberg, R.L., Nealson, K., Friedman, R., Frazier, M., and Venter, J.C. (2007). The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biology* 5**,** e77.

Salvesen, I., and Vadstein, O. (2000). Evaluation of plate count methods for determination of maximum specific growth rate in mixed microbial communities, and its possible application for diversity assessment. *Journal of Applied Microbiology* 88**,** 442-448.

Sambrook, J., and Russel, D.W. (2001). *Molecular Cloning.* Cold Spring Harbor, NY, USA. Cold Spring Harbor Laboratory Press.

Schabhüttl, S., Hingsamer, P., Weigelhofer, G., Hein, T., Weigert, A., and Striebel, M. (2013). Temperature and species richness effects in phytoplankton communities. *Oecologia* 171**,** 527-536.

Scheffer, M., Rinaldi, S., Huisman, J., and Weissing, F. (2003). Why plankton communities have no equilibrium: solutions to the paradox. *Hydrobiologia* 491**,** 9-18.

Schindler, D.W. (2012). The dilemma of controlling cultural eutrophication of lakes. *Proceedings of the Royal Society B* 279**,** 4322-4333.

Schindler, D.W., Hecky, R.E., Findlay, D.L., Stainton, M.P., Parker, B.R., Paterson, M.J., Beaty, K.G., Lyng, M., and Kasian, S.E. (2008). Eutrophication of lakes cannot be controlled by reducing nitrogen input: results of a 37-year whole-ecosystem experiment. *Proceedings of the National Academy of Science* 105**,** 11254-11258.

Schlarb-Ridley, B., and Parker, B. (2013). "UK Roadmap for Algae Technologies". Algal Bioenergy Special Interest Group.

Schlesinger, D.A., Molot, L.A., and Shuter, B.J. (1981). Specific growth rates of freshwater algae in relation to cell size and light intensity. *Canadian Journal of Fisheries and Aquatic Sciences* 38**,** 1052-1058.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., and Weber, C.F. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* 75**,** 7537-7541.

Schnecker, J., Wild, B., Hofhansl, F., Eloy Alves, R.J., Bárta, J., Čapek, P., Fuchslueger, L., Gentsch, N., Gittel, A., Guggenberger, G., Hofer, A., Kienzl, S., Knoltsch, A., Lashchinskiy, N., Mikutta, R., Šantrůčková, H., Shibistova, O., Takriti, M., Urich, T., Weltin, G., and Richter, A. (2014). Effects of soil organic matter properties and microbial community composition on enzyme activities in cryoturbated Arctic soils. *PLoS ONE* 9**,** e94076.

Schulz, B., and Kolukisaoglu, H.Ü. (2006). Genomics of plant ABC transporters: The alphabet of photosynthetic life forms or just holes in membranes? *FEBS Letters* 580**,** 1010-1016.

Schweitzer, B., Huber, I., Amann, R., Ludwig, W., and Simon, M. (2001). Alpha- and beta-Proteobacteria control the consumption and release of amino acids on lake snow aggregates. *Applied and Environmental Microbiology* 67**,** 632-645.

Seyedsayamdost, M.R., Case, R.J., Kolter, R., and Clardy, J. (2011). The Jekyll-and-Hyde chemistry of *Phaeobacter gallaeciensis*. *Nature Chemistry* 3**,** 331-335.

Shaw, G.R., Moore, D.P., and Garnett, C. (2002). "Eutrophication and algal blooms", in: *Encyclopedia of Life Support Systems - Environmental and Ecological Chemistry*.

Sheehan, J., Dunahay, T., Benemann, J., and Roessler, P. (1998). "A look back at the U.S. Department of Energy's aquatic species program: biodiesel from algae". National Renewable Energy Laboratory, USA.

Shurin, J.B., Abbott, R.L., Deal, M.S., Kwan, G.T., Litchman, E., McBride, R.C., Mandal, S., and Smith, V.H. (2013). Industrial-strength ecology: trade-offs and opportunities in algal biofuel production. *Ecology Letters* 16**,** 1393-1404.

Siaut, M., Cuine, S., Cagnon, C., Fessler, B., Nguyen, M., Carrier, P., Beyly, A., Beisson, F., Triantaphylides, C., Li-Beisson, Y., and Peltier, G. (2011). Oil accumulation in the model green alga *Chlamydomonas reinhardtii*: characterization, variability between common laboratory strains and relationship with starch reserves. *BMC Biotechnology* 11**,** 7.

Silva, L.P., and Northen, T.R. (2015). Exometabolomics and MSI: deconstructing how cells interact to transform their small molecule environment. *Current Opinion in Biotechnology* 34**,** 209-216.

Smith, V.H. (1979). Nutrient dependence of primary productivity in lakes. *Limnology and Oceanography* 24**,** 1051-1064.

Smith, V.H. (1982). The nitrogen and phosphorus dependence of algal biomass in lakes: An empirical and theoretical analysis. *Limnology and Oceanography* 27**,** 1101-1111.

Smith, V.H., and Mcbride, R.C. (2015). Key ecological challenges in sustainable algal biofuels production. *Journal of Plankton Research* 37**,** 671-682.

Smith, V.H., and Schindler, D.W. (2009). Eutrophication science: where do we go from here? *Trends in Ecology & Evolution* 24**,** 201-207.

Smith, V.H., Tilman, G.D., and Nekola, J.C. (1999). Eutrophication: impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. *Environmental Pollution* 100**,** 179-196.

Snow, A.A., and Smith, V.H. (2012). Genetically engineered algae for biofuels: A key role for ecologists. *BioScience* 62**,** 765-768.

Sowell, S.M., Abraham, P.E., Shah, M., Verberkmoes, N.C., Smith, D.P., Barofsky, D.F., and Giovannoni, S.J. (2011). Environmental proteomics of microbial plankton in a highly productive coastal upwelling system. *The ISME Journal* 5**,** 856-865.

Sowell, S.M., Wilhelm, L.J., Norbeck, A.D., Lipton, M.S., Nicora, C.D., Barofsky, D.F., Carlson, C.A., Smith, R.D., and Giovanonni, S.J. (2009). Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *The ISME Journal* 3**,** 93-105.

Spolaore, P., Joannis-Cassan, C., Duran, E., and Isambert, A. (2006). Commercial applications of microalgae. *Journal of Bioscience and Bioengineering* 101**,** 87-96.

Steele, J.H. (1974). *The structure of marine ecosystems.* Harvard University Press.

Stephanopoulos, G. (2012). Synthetic Biology and Metabolic Engineering. *ACS Synthetic Biology* 1**,** 514-525.

Sterner, R.W., Andersen, T., Elser, J.J., Hessen, D.O., Hood, J.M., McCauley, E., and Urabe, J. (2008). Scale-dependent carbon:nitrogen:phosphorus seston stoichiometry in marine and freshwaters. *Limnology and Oceanography* 53**,** 1169-1180.

Stockenreiter, M., Graber, A.-K., Haupt, F., and Stibor, H. (2012). The effect of species

diversity on lipid production by micro-algal communities. *Journal of Applied Phycology* 24**,** 45-54.

Summerhayes, V.S., and Elton, C.S. (1923). Contributions to the ecology of Spitsbergen and Bear Island. *Journal of Ecology* 11**,** 214 - 286.

Tadonléke, R.D. (2010). Evidence of warming effects on phytoplankton productivity rates and their dependence on eutrophication status. *Limnology and Oceanography* 55**,** 973-982.

Takagi, M., Karseno, and Yoshida, T. (2006). Effect of salt concentration on intracellular accumulation of lipids and triacylglyceride in marine microalgae *Dunaliella* cells. *Journal of Bioscience and Bioengineering* 101**,** 223-226.

Talling, J.F. (1976). The depletion of carbon dioxide from lake water by phytoplankton. *Journal of Ecology* 64**,** 79-121.

Teeling, H., Fuchs, B.M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C.M., Kassabgy, M., Huang, S., Mann, A.J., Waldmann, J., Weber, M., Klindworth, A., Otto, A., Lange, J., Bernhardt, J., Reinsch, C., Hecker, M., Peplies, J., Bockelmann, F.D., Callies, U., Gerdts, G., Wichels, A., Wiltshire, K.H., Glockner, F.O., Schweder, T., and Amann, R. (2012). Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* 336**,** 608-611.

Teikari, J., Österholm, J., Kopf, M., Battchikova, N., Wahlsten, M., Aro, E.-M., Hess, W.R., and Sivonen, K. (2015). Transcriptomics and proteomics profiling of *Anabaena* sp. Strain 90 under inorganic phosphorus stress. *Applied and Environmental Microbiology*.

Teissier, S., Peretyatko, A., De Backer, S., and Triest, L. (2011). Strength of phytoplankton–nutrient relationship: evidence from 13 biomanipulated ponds. *Hydrobiologia* 689**,** 147-159.

Thompson, A.W., Foster, R.A., Krupke, A., Carter, B.J., Musat, N., Vaulot, D., Kuypers, M.M.M., and Zehr, J.P. (2012). Unicellular cyanobacterium symbiotic with a single-celled eukaryotic alga. *Science* 337**,** 1546-1550.

Tilman, D. (1982). *Resource Competition and Community Structure.* Princeton University Press Princeton, NJ.

Tilman, D., Hill, J., and Lehman, C. (2006). Carbon-negative biofuels from low-input high-diversity grassland biomass. *Science* 314**,** 1598-1600.

Tilman, D., Reich, P.B., Knops, J., Wedin, D., Mielke, T., and Lehman, C. (2001). Diversity and productivity in a long-term grassland experiment. *Science* 294**,** 843-845.

Tilzer, M.M. (1987). Light-dependence of photosynthesis and growth in cyanobacteria: Implications for their dominance in eutrophic lakes. *New Zealand Journal of Marine and Freshwater Research* 21**,** 401-412.

Tyc, O., Zweers, H., de Boer, W., and Garbeva, P. (2015). Volatiles in inter-specific bacterial interactions. *Frontiers in Microbiology* 6**,** 1412.

Tyson, G.W., Chapman, J., Hugenholtz, P., Allen, E.E., Ram, R.J., Richardson, P.M., Solovyev, V.V., Rubin, E.M., Rokhsar, D.S., and Banfield, J.F. (2004). Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* 428**,** 37-43.

Uhlik, O., Leewis, M.-C., Strejcek, M., Musilova, L., Mackova, M., Leigh, M.B., and Macek, T. (2013). Stable isotope probing in the metagenomics era: A bridge towards improved bioremediation. *Biotechnology Advances* 31**,** 154-165.

USEPA (1986). "Quality criteria for water". Washington DC. United States Environmental Protection Agency.

Valladares, A., Montesinos, M.L., Herrero, A., and Flores, E. (2002). An ABC-type, high-affinity urea permease identified in cyanobacteria. *Molecular Microbiology*

43**,** 703-715.

Van den Meersche, K., Middelburg, J.J., Soetaert, K., van Rijswijk, P., Boschker, H.T.S., and Heip, C.H.R. (2004). Carbon-nitrogen coupling and algal-bacterial interactions during an experimental bloom: Modeling a 13C tracer experiment. *Limnology and Oceanography* 49**,** 862-878.

Van Donk, E. (1997). Altered cell wall morphology in nutrient-deficient phytoplankton and its impact on grazers. *Limnology and Oceanography* 42**,** 357-364.

Van, K., Wang, Y., Nakamura, Y., and Spalding, M.H. (2001). Insertional mutants of *Chlamydomonas reinhardtii* that require elevated $CO_2$ for survival. *Plant Physiology* 127**,** 607-614.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D., Paulsen, I., Nelson, K.E., Nelson, W., Fouts, D.E., Levy, S., Knap, A.H., Lomas, M.W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannkoch, C., Rogers, Y.H., and Smith, H.O. (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304**,** 66-74.

Vieira, V. (2014). "Microalgae for fuel and beyond: Future trends", in: *International Society for Applied Phycology Newsletter*. International Society for Applied Phycology.

von Bergen, M., Jehmlich, N., Taubert, M., Vogt, C., Bastida, F., Herbst, F.-A., Schmidt, F., Richnow, H.-H., and Seifert, J. (2013). Insights from quantitative metaproteomics and protein-stable isotope probing into microbial ecology. *The ISME Journal* 7**,** 1877-1885.

Wang, L., Ma, F., Pang, C., and Firdoz, S. (2013). Multicausal analysis on water deterioration processes present in a drinking water treatment system. *Water Environment Research* 85**,** 232-238.

Wang, Y., and Spalding, M.H. (2014). Acclimation to very low $CO_2$: contribution of limiting $CO_2$ inducible proteins, LCIB and LCIA, to inorganic carbon uptake in *Chlamydomonas reinhardtii*. *Plant Physiology* 166**,** 2040-2050.

Weissman, J.C., Goebel, R.P., and Benemann, J.R. (1988). Photobioreactor design: Mixing, carbon utilization, and oxygen accumulation. *Biotechnology and Bioengineering* 31**,** 336-344.

Wen, E.P., Ellis, R., and Pujar, N.S. (2014). *Vaccine Development and Manufacturing.* Wiley.

Werner, G.D.A., Strassmann, J.E., Ivens, A.B.F., Engelmoer, D.J.P., Verbruggen, E., Queller, D.C., Noë, R., Johnson, N.C., Hammerstein, P., and Kiers, E.T. (2014). Evolution of microbial markets. *Proceedings of the National Academy of Sciences* 111**,** 1237-1244.

Williams, T.J., and Cavicchioli, R. (2014). Marine metaproteomics: deciphering the microbial metabolic food web. *Trends in Microbiology* 22**,** 248-260.

Williams, T.J., Wilkins, D., Long, E., Evans, F., DeMaere, M.Z., Raftery, M.J., and Cavicchioli, R. (2013). The role of planktonic Flavobacteria in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. *Environmental Microbiology* 15**,** 1302-1317.

Wilmes, P., and Bond, P.L. (2004). The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms. *Environmental Microbiology* 6**,** 911-920.

Wilmes, P., and Bond, P.L. (2006). Metaproteomics: studying functional gene expression in microbial ecosystems. *Trends in Microbiology* 14**,** 92-97.

Wood, A.M., and Van Valen, L.M. (1990). Paradox lost? On the release of energy-rich compounds by phytoplankton. *Marine Microbial Food Web* 4**,** 103-116.

Work, V.H., Radakovits, R., Jinkerson, R.E., Meuser, J.E., Elliott, L.G., Vinyard, D.J.,

Laurens, L.M., Dismukes, G.C., and Posewitz, M.C. (2010). Increased lipid accumulation in the *Chlamydomonas reinhardtii* sta7-10 starchless isoamylase mutant and increased carbohydrate synthesis in complemented strains. *Eukaryotic Cell* 9**,** 1251-1261.

Xu, L., Brilman, D.W.F., Withag, J.A.M., Brem, G., and Kersten, S. (2011). Assessment of a dry and a wet route for the production of biofuels from microalgae: Energy balance analysis. *Bioresource Technology* 102**,** 5113-5122.

Yang, D., Song, D., Kind, T., Ma, Y., Hoefkens, J., and Fiehn, O. (2015). Lipidomic analysis of *Chlamydomonas reinhardtii* under nitrogen and sulfur deprivation. *PLoS One* 10**,** e0137948.

Yang, X.E., Wu, X., Hao, H.L., and He, Z.L. (2008). Mechanisms and assessment of water eutrophication. *Journal of Zhejiang University Science B* 9**,** 197-209.

Yeates, T.O., Kerfeld, C.A., Heinhorst, S., Cannon, G.C., and Shively, J.M. (2008). Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nature Reviews Microbiology* 6**,** 681-691.

Yutin, N., Suzuki, M.T., Teeling, H., Weber, M., Venter, J.C., Rusch, D.B., and Beja, O. (2007). Assessing diversity and biogeography of aerobic anoxygenic phototrophic bacteria in surface waters of the Atlantic and Pacific Oceans using the Global Ocean Sampling expedition metagenomes. *Environmental Microbiology* 9**,** 1464-1475.

Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30**,** 614-620.

Zhang, M., Duan, H., Shi, X., Yu, Y., and Kong, F. (2012). Contributions of meteorology to the phenology of cyanobacterial blooms: Implications for future climate change. *Water Research* 46**,** 442-452.

# Appendix 1

Table A1.1. Complete composition of artificial freshwater growth medium.

| Component | Concentration (mg $L^{-1}$) |
|---|---|
| $NaHCO_3$ | 192 |
| $MnCl_2.4H_2O$ | 0.18 |
| $MgSO_47H_2O$ | 115 |
| KCl | 0.45 |
| $H_2SeO_3$ | 0.0016 |
| $Ca(NO_3)_24H_2O$ | 0.8 |
| $NH_4Cl$ | 1 |
| $KH_2PO_4$ | 0.025 |
| $K_2PO_4$ | 0.025 |
| $ZnSO_4.7H_2O$ | 0.022 |
| $Na_2EDTA.2H_2O$ | 0.5 |
| $H_3BO_3$ | 0.114 |
| $FeSO4.7H_2O$ | 0.05 |
| $CuSO_4.5H_2O$ | 0.016 |
| $CoCl_2.6H_2O$ | 0.016 |
| $(NH_4)_6Mo7O_{24}.4H_2O$ | 0.011 |

# Appendix 2

Table A2.1. Complete list of the eukaryotic organisms in the experimental freshwater microbial community inoculum as determined by 18S rDNA sequencing.

| Eukaryotic organisms | |
|---|---|
| | % |
| *Chloromonas pseudoplatyrhyncha* | 26.93 |
| *Stephanodiscus* sp. | 18.17 |
| Unsequenced organisms | 17.87 |
| *Chromulinaceae* sp. | 8.48 |
| *Synedra angustissima* | 4.99 |
| *Ochromonadales* sp. | 3.01 |
| *Chlamydomonas* sp. | 2.98 |
| *Micractinium pusillum* | 1.62 |
| *Chlorella* sp. | 1.08 |
| *Chloromonas cf schussnigii* | 0.58 |
| *Asterarcys quadricellulare* | 0.41 |
| *Cymbella minuta* | 0.38 |
| *Chloromonas insignis* | 0.30 |
| *Chlamydomonas reinhardtii* | 0.29 |
| *Chlamydomonas monadina* | 0.27 |
| *Chlorella fusca* | 0.26 |
| *Cyclotella meneghiniana* | 0.24 |
| *Paulschulzia pseudovolvox* | 0.16 |
| *Scenedesmus deserticola* | 0.11 |
| *Treubaria schmidlei* | 0.09 |
| *Choricystis* sp. | 0.01 |
| *Chloroidium saccharophila* | 0.01 |
| *Oocystis marssonii* | 0.01 |

Table A2.2. Complete list of the prokaryotic organisms in the experimental freshwater microbial community inoculum as determined by 16S rDNA sequencing.

| Prokaryotic organisms | |
|---|---|
| | % |
| *Rhodoferax* sp. | 21.94 |
| Unsequenced organisms | 17.84 |
| *Flavobacterium* sp. | 9.43 |
| *Anabaena* sp. | 8.85 |
| *Brevundimonas diminuta* | 4.20 |
| Hydrogenophaga sp | 3.41 |
| *Runella limosa* | 2.47 |
| *Haliscomenobacter* sp. | 2.43 |
| *Rhodobacter* sp. | 2.34 |
| *Planktophila limnetica* | 2.13 |
| *Agrobacterium tumefaciens* | 2.11 |
| *Sphingobacterium* sp. | 2.03 |
| *Ochrobactrum tritici* | 1.98 |
| *Brevundimonas variabilis* | 1.83 |
| *Sphingomonas* sp. | 1.73 |
| *Curvibacter* sp. | 1.48 |
| *Phenylobacterium falsum* | 1.42 |
| *Roseomonas stagni* | 1.24 |
| *Oceanicaulis* sp. | 1.03 |
| *Erythromicrobium* sp. | 0.89 |
| *Anabaena bergii* | 0.73 |
| *Rhodoluna lacicola* | 0.70 |
| *Inhella inkyongensis* | 0.60 |
| *Caulobacter* sp. | 0.53 |
| Sphingopyxis sp. | 0.51 |
| *Sphingopyxis alaskensis* | 0.43 |
| *Erythromicrobium ramosum* | 0.36 |
| *Achromobacter* sp. | 0.33 |
| *Alcaligenes* sp. | 0.32 |

| | |
|---|---|
| *Roseococcus* sp. | 0.31 |
| *Albidiferax ferrireducens* | 0.30 |
| *Phormidium* sp. | 0.29 |
| *Hyphomicrobium* sp. | 0.28 |
| *Delftia* sp. | 0.28 |
| *Bosea thiooxidans* | 0.28 |
| *Rhizobium* sp. | 0.28 |
| *Prosthecomicrobium pneumaticum* | 0.26 |
| *Stenotrophomonas* sp. | 0.24 |
| *Paucimonas lemoignei* | 0.21 |
| *Verrucomicrobium* sp. | 0.21 |
| *Reyranella massiliensis* | 0.18 |
| *Afipia* sp. | 0.18 |
| *Limnohabitans* sp. | 0.14 |
| *Pedobacter* sp. | 0.13 |
| *Polynucleobacter necessarius* | 0.13 |
| *Sandarakinorhabdus* sp. | 0.12 |
| *Polaromonas* sp. | 0.11 |
| *Rhodoferax antarcticus* | 0.11 |
| *Hahella* sp. | 0.11 |
| *Devosia ginsengisoli* | 0.08 |
| *Devosia* sp. | 0.06 |
| *Gluconobacter oxydans* | 0.06 |
| *Methylophilus* sp. | 0.05 |
| *Rheinheimera* sp. | 0.05 |
| *Synechococcus* sp. | 0.04 |
| *Acetobacter* sp. | 0.04 |
| *Planctomyces* sp. | 0.03 |
| *Aquabacterium* sp. | 0.03 |
| *Defluvibacter* sp. | 0.02 |
| *Acidovorax* sp. | 0.02 |
| *Mycobacterium* sp. | 0.01 |
| *Sediminibacterium salmoneum* | 0.01 |
| *Methylocella* sp. | 0.01 |

| | |
|---|---|
| *Mesorhizobium* sp. | 0.01 |
| *Pseudomonas* sp. | 0.01 |
| *Sediminibacterium* sp. | 0.01 |
| *Polynucleobacter* sp. | 0.01 |
| *Bacteriovorax* sp. | 0.01 |
| *Pseudomonas putida* | 0.01 |

# Appendix 3

## A3.1. Detergent-based lysis to differentiate *C. reinhardtii* WT from mutant strain

The WT and mutant *C. reinhardtii* strains are phenotypically very different (e.g. mutant strain lacks flagellum and cell wall), however, under an optical microscope, they are undistinguishable. Therefore, a simple and quick technique was devised in order to count cells for each strain separately in a mixed culture. The initial description of the *C. reinhardtii* cell wall-deficient strains describes them as flat, liquid colonies on agar, and with cells that lyse in a 1% solution of a non-ionic detergent NP-40 (Davies et al., 1971). Based on this, mutant and WT cells were mixed in known ratios (0:100, 0.25:0.75, 0.5:0.5, 0.75:0.25, 100:0) and an equal volume of Triton X-100 2% (v/v) was added to the cell suspensions in order to have a working concentration of 1%. The cell suspensions were counted and incubated for ten minutes with re-counts at five and ten minutes. The results (Fig. A3.1) indicate that there is complete lysis at five minutes with no improvement having been observed at ten minutes.

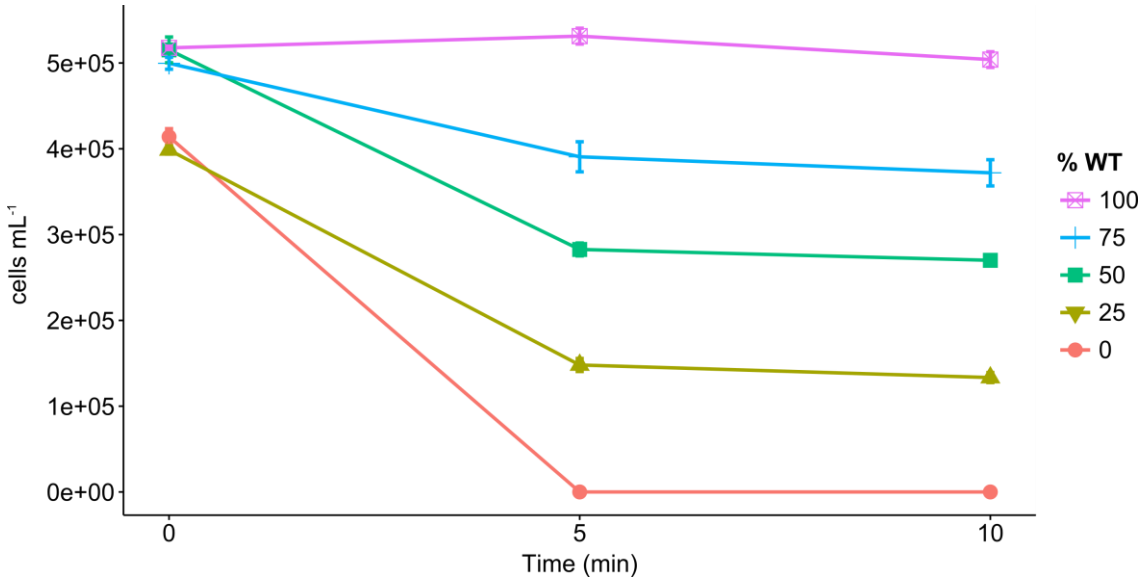


Figure A3.1. Number of cells ml$^{-1}$ after the addition of an equal volume of Triton X-100 1% (v/v) at time zero. Percentages represent the initial percentage of WT cells in the mixtures. Errors show standard error (n = 3).

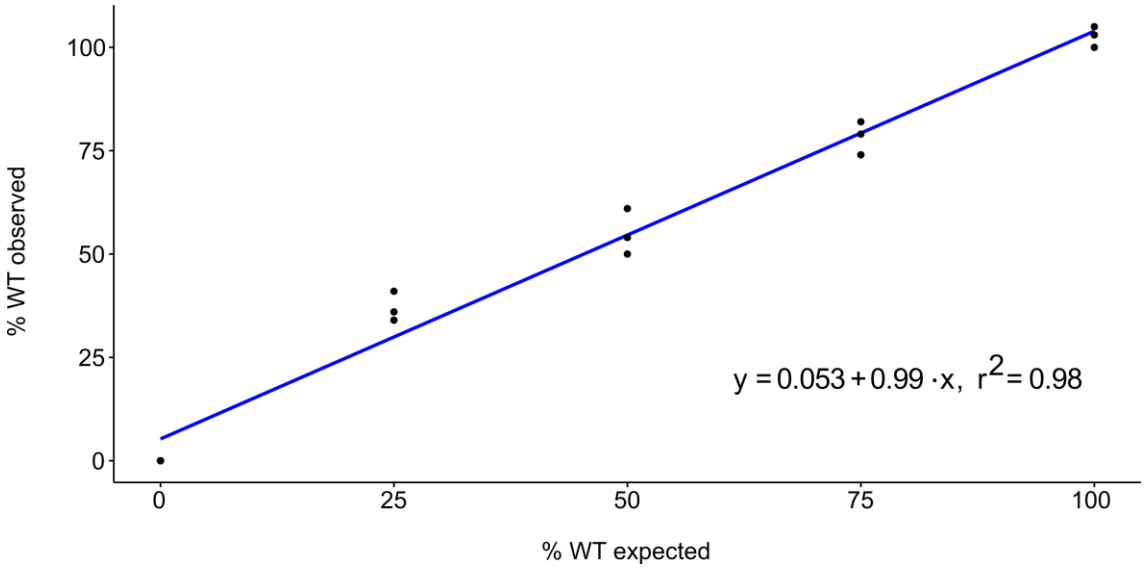


Figure A3.2. Plot of observed cell counts against expected cell counts after a five minute incubation period. The percentages on both axes represent the percentage of initial WT cells in the mixtures.

Based on the results at five minutes (Fig. A3.1) the observed cell counts were plotted against the expected cell counts to determine precision and accuracy. The observed cell counts (Fig. A3.2) are, for the most part, precise and accurate. The largest inaccuracy was registered at the point of 25% WT + 75% mutant. An explanation for this is the fact that cell counts become more inaccurate with lower numbers. This can be compensated by increasing the number of squares counted. The incubation time of five minutes was sufficient for complete cell lysis of the mutant strain without causing lysis of the WT strain. It also has the advantage of being quick and not leaving any traces of cell debris. This eliminates errors associated to, mistakenly, counting large debris as cells. In sum, the results show that outcome of competition, between the WT and mutant strain, can be determined both accurately and precisely detergent mediated cell lysis.