UNIVERSITY OF CALIFORNIA SAN DIEGO

Optimizing recombinant protein production in Chlamydomonas reinhardtii: turning a model organism into a commercially relevant biotechnological host

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

in

Biology

by

Yasin Torres-Tiji

Committee in charge:

Professor Stephen P. Mayfield, Chair
Professor Eric E. Allen, Co-Chair
Professor Michael D. Burkhart
Professor Susan S. Golden
Professor Jose Pruneda-Paz

2021

The Dissertation of Yasin Torres-Tiji is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

DEDICATION

This dissertation is dedicated to my mother Malika Tiji Joafari, who through no small effort emigrated from Africa to Europe seeking a better life and accomplished that goal. Then she proceeded to give everything to me, her son, so I could pursue my dreams in science on the other side of the world. Son of an immigrant become immigrant myself, my success is thanks to her success.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I must acknowledge first and foremost my supervisor, Dr. Stephen Mayfield for giving me the opportunity to perform the research I love in his lab and for continuing to believe in me through my mistakes and for pushing me to be a better scientist. The journey has been a long and arduous one in which I have matured both as a person and as a scientist, and the role my supervisor played in this was crucial and I acknowledge it as such. I remember like it was yesterday, when almost seven years ago I arrived to the United States as a naïve 20 year old with a head full of dreams and a hear full of passion. I came as an exchange student to the University of California San Diego to study my last year of Bachelor's Degree in Biotechnology. I was accepted to any campus of University of California, and I initially intended to attend the campus at Berkeley but that is when I saw the TedTalk that my supervisor has on Youtube about microalgal research. After that, I had to attend the campus at San Diego, meet him and volunteer in his lab somehow. After attending the first class I took from Dr. Stephen Mayfield, I went into his office during office hours, I introduced myself and asked for a volunteer position in his lab. He readily accepted me and offered me a volunteer assistant position under Dr. Miller Tran. After I completed my Bachelor's Degree, I was accepted in the PhD Program at UCSD, and he welcomed me back in the Mayfield Lab. The following 5 years I performed research under his supervision and learned invaluable lessons. In the end, I earned a PhD in Biology, but the most important lessons learned were not in the field of science but rather on how to lead a successful life.

I must now thank Dr. Miller Tran who was my mentor as an undergraduate performing research in the Mayfield Lab. Unfortunately, by the time I joined the lab as a graduate student he had left to work full time in his start-up biotech company, in which endeavour I wish him the best of luck, but the lessons I learned working with him during that initial year stuck with me until this day. I always remember what he taught me, and I made sure to pass on those lessons to the undegraduates who helped me during my journey in gradschool.

Next, I must acknowledge the undegraduates who assisted me with my research, Saumya Keremane, Natalia Avina-Ochoa, Ismael Plascencia Torres and Sydney Woo. I feel as if I learned as much from teaching them as they learned from being taught by me.

I must acknowledge all my teammates, Francis Fields, Anthony Berndt, Amr Badary, Bijie Ren and Tressa Smalley. Francis Fields, the lab manager and senior scientist in the Mayfield Lab, for always being there to support my research, to give professional and personal advice, and to make sure I had what I needed to perform my research without any problems. Anthony Berndt, for the immense scientific advice and from whom I learned so much, having someone so knowledgeable by my side in the lab has being crucial in my success in my PhD. Amr Badary and Bijie Ren, for being great teammates and being there to support each other. Tressa Smalley, for always being to do the work that sometimes does not get recognized but is essential, namely preparing media and cleaning dishes and helping grow out algal biomass.

Special acknowledgement for all the friends I made during this journey. As a foreigner living without family by my side, friends became my family away from home and their emotional support and the memories made push me forward to this day. I will never forget each one of you, as far as you might be living from me now and as much time might have past since the last time we talked, the memories live on with me. Of course, I must include my best friend and significant other, Rachel Borling, who endured by my side the last push in this journey and who offered invaluable emotional support during the times when I most needed. Thank you for always believing in me, even when I did not, and for being there for me, and for being the wonderful person that you are.

Finally, I must acknowledge my family for all the support I received from them. My uncle Said Tiji and my aunt Shawna Morrison, who live in the United States and helped a lost boy in a foreign land to establish himself and without whom I would have not made it. Thank you for being there for me. And my mother, Malika Tiji Joafari, who sacrificed everything to make sure I could have the best life I could. Thank you for striving me to study, you always made sure to let me know which was the path forward. We did not have much when I was a kid, but you recognized my talent for studying and relentlessly encouraged me to pursue the higher education you never had. All I have achieved was thanks to you.

Chapter 2, in full, is a reprint of the material as it appears in Biotechnology Advances, 2020. Torres-Tiji, Yasin; Fields, Francis J; Mayfield, Stephen P. The dissertation author was the primary investigator of this paper.

Chapter 3 is being prepared for publication. Torres-Tiji, Yasin; Mayfield, Stephen P. The dissertation author was the primary investigator of this paper.

Chapter 4 is being prepared for publication. Torres-Tiji, Yasin; Fields, Francis J; Mayfield, Stephen P. The dissertation author was the primary investigator of this paper.

Chapter 5 is being prepared for publication. Torres-Tiji, Yasin; Fields, Francis J; Jenkins, Hunter D; Mayfield, Stephen P. The dissertation author was the primary investigator of this paper.

.

VITA

Ph. D. Biology, University of California San Diego (2021)

B.S. Biotechnology, Autonomous University of Barcelona (2015)

Publications

Torres-Tiji, Y., Fields, F. J., & Mayfield, S. P. (2020). Microalgae as a future food source. *Biotechnology Advances, 41*, 107536. doi:https://doi.org/10.1016/j.biotechadv.2020.107536

Berndt, A., Smalley, T., Ren, B., Badary, A., Sproles, A., Fields, F., **Torres-Tiji, Y**., Heredia, V., Mayfield, S. (2021). Recombinant production of a functional SARS-CoV-2 spike receptor binding domain in the green algae Chlamydomonas reinhardtii. *bioRxiv*, 2021.2001.2029.428890. doi:10.1101/2021.01.29.428890

ABSTRACT OF THE DISSERTATION

Optimizing recombinant protein production in Chlamydomonas reinhardtii: turning a model organism into a commercially relevant biotechnological host

by

Yasin Torres-Tiji

Doctor of Philosophy in Biology

University of California San Diego, 2021

Professor Stephen P. Mayfield, Chair
Professor Eric E. Allen, Co-Chair

Human need for food and fuel has disturbed the balance of the biosphere thus triggering a catastrophe known as the Anthropocene Extinction. The use of microalgae in the biotechnology field offers multiple solutions that could alleviate the demand human activity imposes on the ecosystem. The most well studied microalga is the model organism *Chlamydomonas reinhardtii*, resulting in multiple genetic tools available in the alga for recombinant protein expression. However, the yields of recombinant protein expression are not high enough to be commercially viable, therefore this organism is usually not considered as a biotechnological host.

To boost the recombinant protein productivity in *Chlamydomonas reinhardtii* two milestones need to be achieved: increased recombinant protein expression at single cell level, and increased number of cells per unit of culture volume. To accomplish higher transgene expression in *C. reinhardtii* the GAL4/UAS system was adapted into algal protein expression vectors. This system showed a 10-fold improvement in recombinant mRNA and protein accumulation of a reporter gene under the control of a chimeric promoter 5XUAS-AR1. To accomplish higher number of cells, or biomass, per unit of culture volume an optimized algal fed batch bioreactor was designed. Through media optimization we achieved a 1.67-fold improvement in biomass accumulation which in turn yielded a 3-fold improvement over the highest recombinant protein concentration reported in the literature using *C. reinhardtii*. Finally, an extremophile green alga from the *Chlamydomonas* genus was isolated from the wild and used to express recombinant GFP. Said extremophile showed robust growth in open ponds thriving in media at pH 11 while continuing to express recombinant protein for the duration of the experiment.

These findings highlight the potential of *Chlamydomonas reinhardtii* to become a robust biotechnological host at commercial scale.

Chapter 1: Introduction

Anthropocene Extinction as the problem to be solved

Life on Earth is composed of a myriad of organisms, microscopic and macroscopic, multicellular and unicellular, coexisting in an extremely complex balance called the biosphere. In this system, fluxes of carbon and energy between individuals drive life and generate intricate food chains in which organic matter is recycled in a closed system while energy, mostly from the Sun, fuels this cycle. This arrangement has been evolving for approximately 4.2 billion years (Dodd, Papineau, Grenne, Slack, Rittner, Pirajno, O'Neil, & Little, 2017), yielding tremendously diverse forms of life that have fought to survive, perpetuate their species and pass on their genes on to future generations. This incredible process has resulted in the wonderful biodiversity currently present on Earth, passing through historic times of great diversification of life called mass speciations and times of great decline of diversification of life called mass extinctions. The human population has massively increased in the last 200 years, imposing an incredible tax on the ecosystem that has been so detrimental to other species, that experts are starting to call it the Holocene Extinction or Anthropocene Extinction (Raleigh, 1999).

The main reasons behind the Anthropocene Extinction are habitat destruction, overexploitation, introduction of non-native species, pollution and climate change (Turvey & Crees, 2019; Young, McCauley, Galetti, & Dirzo, 2016). Most of these are a direct consequence of human food and energy production, which are the two most resource intensive industries. In fact, human population massive increased was only possible due to intensive utilization of fossil fuels which were in turn used to boost agricultural production which allowed the human population to increase from 1 billion people to almost 8 billion people on Earth in a little over 200 years (Raleigh, 1999). This present mode of human life has proven to be not sustainable and is threatening the balance in the biosphere, which would not only destroy the biodiversity present on Earth but would also negatively impact humankind. It is projected that human population will peak at approximately 11.2 billion people by the end of the 21st century (United Nations, 2017). For this to happen and not further impact the ecosystem, we need to find ways to produce food and energy in a much more efficient and sustainable way.

A key element in this situation is that since humans are at the top of the food chain, food being the form by which energy and carbon flux through the biosphere, humans either need to greatly decrease their resource consumption, or all the levels below them need to boost their biomass production, or a combination of the two, so that a balance can be achieved that allows the whole structure to be sustainable. Energy is the force that drives the carbon cycle through the system, and since most of it comes from the Sun it ends up being that photosynthetic organisms are at the base of the food chain, as the main producers of organic carbon to be consumed directly or indirectly by all other organisms of the ecosystem. Therefore, if a large human world population is to be maintained in a sustainable way, we need to boost photosynthetic production of biomass. However, agricultural production with traditional crops (eg. rice, wheat, corn, soybean, etc) has not grown at a fast-enough rate to keep up with human demand increases (Fróna, Szenderák, & Harangi-Rákos, 2019). And so, from the need of sustainable, nutritious, and productive food sources, arises the opportunity for alternative crops to become mainstream agricultural products.

Introduction to microalgae

Microalgae are a group of eukaryotic unicellular microorganisms that can grow by utilizing sunlight to drive the reduction of carbon dioxide to carbohydrates while releasing oxygen; a process called oxygenic photosynthesis. Microalgae are naturally present in all oceans and bodies of water, and are responsible for the production of much of the oxygen in the atmosphere (Chapman, 2013). Cyanobacteria are bacteria that also perform oxygenic photosynthesis, and despite being prokaryotes they are often included in the term “microalgae”. Due to their ancient origin, their rapid growth and their robust photosynthetic apparatus, cyanobacteria are thought to be the primary organisms that allowed Earth’s atmosphere to become oxidative, thus allowing animal life to exist (Schirrmeister, Gugger, & Donoghue, 2015). Additionally, microalgae are a very biodiverse group, are capable of both sexual and asexual reproduction, and are genetically modifiable. For all these reasons, microalgae present a potentially very elegant solution to the problem of increasing demand of carbon and energy, that humans impose on the ecosystem.

The most important characteristic of microalgae in contributing to slowing down climate change and other consequences of the Anthropocene Extinction, is their ability for efficient photosynthesis (U. B. Singh & Ahluwalia, 2013). This enables microalgae to harness the virtually inexhaustible and plentiful energy from the Sun and transform it into chemical energy. By this method, microalgae can grow inexpensively, using very little resources to capture and fix atmospheric carbon dioxide, one of the main contributors to climate change. Additionally, algae can be grown using different types of water depending on the species: from fresh water, to brackish water, sea water, and even hypersaline water. This makes it so microalgal cultivation does not need to utilize drinkable water, thus not competing with humans nor animals, and it does not need freshwater like land plants used in agriculture do. Not only does agriculture require intensive freshwater utilization, but it also requires arable land to grow upon, potentially depleting the soil of nutrients and leading to desertification. Once that happens, new agricultural lands must be developed, often from virgin forests or grasslands, thus competing with the native flora and causing deforestation, another leading cause of the Anthropocene Extinction (U. B. Singh & Ahluwalia, 2013).

Microalgae potential in biotechnology

The unicellular nature of microalgae allows these photosynthetic organisms to grow much more rapidly than land plants used in agriculture for the simple fact that in unicellular organisms each cell is dividing as fast as possible without having to sacrifice photosynthetic capability to produce the biomass required for complex structural features like branches, stems or roots. Additionally, unicellular organisms have a significantly higher surface to volume ratio, thus allowing them to exchange nutrients and waste with their environment at much higher rate, enabling much faster growth (Brown, Gillooly, Allen, Savage, & West, 2004). Asexual reproduction is also an important contributor to the fact that microalgae excel as producing biomass in a short amount of time.

In spite of being simpler organisms than land plants and other multicellular organisms, microalgae possess the inherent complexity that comes with the fact that they are eukaryotes. This characteristic enables algae to synthesize complex bioproducts which in turn adds value to algal biomass (Georgianna & Mayfield, 2012). If the product of interest is biofuels, then the algae could be used to synthesize highly

valuable therapeutic proteins that could be used as co-products and the revenue of those would cheapen the cost of algal biofuels, potentially making them cost competitive with fossil fuels. The complexity derived from their eukaryotic nature makes possible that microalgae be utilized for multitude of bioproducts including, but not limited to: biofuels, food, feed, nutraceuticals and food additives, cosmetics, biopolymers, biofertilizers and therapeutics (Hannon, Gimpel, Tran, Rasala, & Mayfield, 2010). Additionally, microalgae can be employed for wastewater treatment, bioremediation and CO_2 capture, all of which would contribute to addressing some of the detrimental impacts of human activity on the environment (Rath, 2012).

The biodiversity within microalgae is vast, with estimates of total species ranging between 200,000 and several million, while there are 30,000 species already identified (Pulz & Gross, 2004). This natural diversity offers a great opportunity for bioprospection, encompassing both the potential to find microalgal species in which to biosynthesize high value products, useful microalgal derived molecules that could be produced using a different microorganism, or advantageous traits that could be transferred to other biotechnological hosts through genetic engineering. Over 15,000 novel compounds have been found by bioprospecting algae (Cardozo, Guaratini, Barros, Falcão, Tonon, Lopes, Campos, Torres, Souza, Colepicolo, & Pinto, 2007). This is but a fraction of the potential molecules with biotechnological relevance that could be found through bioprospection, due to microalgae's enormous genetic diversity and their presence in multitude of different and sometimes extreme ecosystems. Moreover, algae's nutrient rich composition makes it a desired prey for microzooplanktonic grazers, and thus in order to survive in a competitive environment, algae have created a multitude of structurally and chemically diverse compounds (Day, Gong, & Hu, 2017).

Nevertheless, bioprospecting microalgal species is not the only way to obtain a biotechnologically relevant microorganism. The concept of strain improvement is that of engineering a strain through a series of processes, all aimed at enhancing certain traits of interest. This can be done in two ways: generating unknown genetic diversity and selecting for the trait of interest using high throughput screening (HTS) or generating a known genetic alteration that yields the trait sought after. Owing to the fact that microalgae are capable of sexual reproduction, they can be bred and selected to obtain the

desired traits, just like it has been done in traditional agriculture, but a much faster rate thanks to their much shorter reproductive cycle. If the trait is easy to select through HTS (eg. using a fluorophore as marker) then random mutagenesis and breeding coupled with HTS becomes an extremely powerful tool for strain improvement (Fields, Ostrand, Tran, & Mayfield, 2019). If the genes that encode for a certain desired phenotype are known, then genetic engineering techniques can be applied. Many microalgal genomes have been sequenced and that number is constantly increasing. Additionally, recombinant gene expression of many different complex proteins has been successful in a number of different microalgae (Torres-Tiji, Fields, & Mayfield, 2020). Finally, targeted gene editing using CRISPR/Cas and other methods has been shown to work in some microalgae, which enables the precise deletion of unwanted genes and the precise addition of DNA sequences of interest (Patel, Soni, Prasad, Sapre, Dasgupta, & Bhadra, 2019).

It is for the aforementioned reasons that microalgae have a great potential in biotechnology. Despite most of its potential being untapped, there has been biotechnological use of microalgae for centuries, the most common as a food source, due to their highly nutritious nature (Wells, Potin, Craigie, Raven, Merchant, Helliwell, Smith, Camire, & Brawley, 2017). Besides its use as human food, microalgae are also currently used for feeding animals, which is the largest market as of 2014, grossing $4 billion USD globally per year (Spolaore, Joannis-Cassan, Duran, & Isambert, 2006). Another successful product derived from microalgae are food supplements, the most popular of them being omega-3 like Docosahexaenoic acid (DHA) and Eicosapentaenoic acid (EPA), which are also multi-million dollar markets (Koller, Muhr, & Braunegg, 2014). Biodiesel from algae is also produced but it has not been widely adopted in the market because it is not cost competitive with fossil fuels. However, all of these products are naturally present in algae, and there is a growing interest in synthesizing recombinant products in microalgae.

Genetic engineering of microalgae

Due to their rapid and inexpensive growth, amenable genetic manipulation, and single celled nature, microalgae have become a promising host for the expression of recombinant products. There has

been extensive research in this field, primarily trying to produce recombinant proteins in the model alga *Chlamydomonas reinhardtii*. This species of green algae is a single celled, flagellated microorganism that has been used to study a multitude of biological processes like photosynthesis, flagellar motility, algal genetics and cell cycle. The main reason it was adopted as a model organism was due to its ability to undergo sexual crosses, and to grow in the dark on a reduced carbon source, which allowed for photosynthetic mutants that were non-lethal (Harris, 2001). Due to all the work in molecular biology in *C. reinhardtii*, all three genomes (nuclear, chloroplast and mitochondrial) have been fully sequenced and are able to be transformed with recombinant genes. The chloroplast is a single cup shaped organelle that takes up two thirds of the cell's volume, with a prokaryote like environment due to their endosymbiotic origin from cyanobacteria and has a simple circular genome of approximately 200,000 base pairs. The chloroplast has been used for the production of recombinant proteins, and it was shown that complex mammalian proteins could be synthesized with the correct folding and di-sulfide bonds (Stephen P. Mayfield, Manuell, Chen, Wu, Tran, Siefker, Muto, & Marin-Navarro, 2007). Other post-translational modifications are available in the chloroplast, most notably the phosphorylation of proteins, but other post-translational modifications are lacking, the most important of which is glycosylation.

Recombinant protein expression in microalgae, via the nuclear genome, is gaining interest because of the ease of transformation as well as an ability to glycosylate proteins, and the possibility to target proteins to any cell compartment, including the chloroplast or the extracellular media (Beth A. Rasala, Chao, Pier, Barrera, & Mayfield, 2014). This is especially relevant for metabolic engineering, in which the strain is optimized to produce a specific molecule by adding enzymes that modify, add, or delete metabolic pathways inside the cell. Since each biochemical pathway step takes place in a specific subcellular location, being able to target recombinant metabolic enzymes to different organelles is a requirement. However, unlike in the chloroplast genome, integration of recombinant genes in the nuclear genome occurs randomly and targeted gene modification has not been shown to work very effectively. There has been some success using CRISPR/Cas9 in *C. reinhardtii*, but it is still far from the level of efficiency required to deliver engineered strains at a rate that could be called high throughput (Angstenberger, de Signori, Vecchi, Dall'Osto, & Bassi, 2020). Recombinant gene expression is extremely variable, depending on the loci in which the construct integrates, requiring a large number of

clones that must be screened to identify those with robust recombinant protein expression. Furthermore, nuclear recombinant gene expression in *C. reinhardtii* is subject to gene silencing, brought about by a strong immune mechanism against viruses by which foreign DNA expression is quickly suppressed through epigenetic gene silencing mechanisms. *C. reinhardtii* also has a strong codon bias in the nucleus, with a GC content of 66%, which allows it to differentiate native DNA from heterologous DNA. Additionally, native genes in this species possess a large number on introns, and addition of introns in heterologous genes has been shown to significantly boost expression, suggesting a mRNA processing mechanism that enhances gene expression. Lack of strong promoters and/or enhancers is also another problem to be overcome if *C. reinhardtii* is to become a biotechnological host for recombinant products (Schroda, 2019). There has been some progress made in that area by the incorporation of synthetic promoters, but further improvement is needed (Scranton, Ostrand, Georgianna, Lofgren, Li, Ellis, Carruthers, Dräger, Masica, & Mayfield, 2016). In this work, I will show how using a heterologous strong transcriptional activator can boost recombinant gene expression and therefore tremendously increase the yields of recombinant proteins.

Microalgae biomass production technologies

Robust recombinant gene expression at the single cell level is not the only requirement for *C. reinhardtii* to become a competitive microbial host for recombinant production of bioproducts, high biomass productivity yields also need to become a reality. There are two main ways in which algae can be grown: outdoors in open ponds, or in close axenic systems. The first option is the optimal for cheap biomass production, suitable for biofuels, biopolymers, food and feed production. The second option is more suitable for production of high value products, in which the price of the product offsets the higher cost of production involved in axenic contained systems where higher productivities are needed (Fields, Ostrand, & Mayfield, 2018). Closed systems also allow for a higher variety of algae to be cultivated, as contamination by other organisms is kept at bay via containment, while in open ponds the media is often selective (extreme salinity or pH) to avoid contamination, thus limiting the number of algal species that can be grown using such media. When it comes to closed systems, the main two options are

photobioreactors and heterotrophic bioreactors. The former ones have been employed in a limited number of large-scale algal cultivation sites and offer a variety of designs adapted to different strains, climates and products. However, since the source of energy is light, this imposes a limit on the highest biomass yield possible, since light will barely penetrate a dense culture in a photobioreactor (Acién, Molina, Reis, Torzillo, Zittelli, Sepúlveda, & Masojídek, 2017). In contraposition, heterotrophic bioreactors employ reduced carbon that can be readily dissolved in the growth media, reaching every cell without problems.

Many algae have been cultivated in heterotrophic bioreactors, some reaching biomass yields of 200 $g \cdot L^{-1}$ Dry Weight (DW) (Jin, Chuai, Li, Hou, Wu, Chen, Wang, Jia, Han, & Hu, 2021). However, those species that have reached high biomass yields in heterotrophic bioreactors, have not yet been engineered to produce high yields of recombinant protein expression at the single cell level. *C. reinhardtii* has been shown to grow well in heterotrophic fed-batch bioreactors, reaching biomass yields of up to 25 $g \cdot L^{-1}$ of Ash Free Dry Weight (AFDW) (Fields et al., 2018). This alga has also shown moderate yields of recombinant protein expression. Therefore, it seems possible that *C. reinhardtii* can achieve the best of both worlds and become a competitive recombinant protein host.

Conclusion

Once *C. reinhardtii* has both high recombinant protein expression on single cell level and high biomass yield per volumetric unit of culture, the opportunity to produce recombinant products at a competitive price, becomes a reality. There have been early attempts at producing different therapeutical recombinant proteins in *C. reinhardtii*, and the data shows that algal recombinant proteins can be folded correctly, can have complex post-translational modifications, and have the correct biological activity (Specht, Miyake-Stoner, & Mayfield, 2010). However, recombinant protein production would just be the beginning, with the mid-term goal of producing food and feed in algae. It must be remarked, that even though these technologies are being developed for the species like *C. reinhardtii*, the genetic tools and bioreactor mode of operation are designed to be applicable to other algal species that show beneficial features for synthesizing the bioproduct of interest.

Chapter 2: Microalgae as future food source

Introduction

Hunger is a problem that affects 1 out of 9 people in the world and the most important component of it, both in terms of the number of affected individuals as well as in terms of lethality, is the lack of appropriate caloric and protein intake; called Protein-Energy Malnutrition (PEM) (FAO, 2017). This problem will only worsen as human population increases, placing even greater demand on our limited protein supplies. With world population projected to reach 9.7 billion people by 2050 (United Nations, 2017), there is no current food solution that can be deployed to meet the expected increased demand for protein. During this same time, agriculture yields are predicted to be significantly impacted by climate change, and while some regions will benefit, other regions will see decline, with the overall balance expected to be negative (National Acadamy of Sciences, 2008). Overfishing of our oceans, and arable land decline, will also add to this problem, suggesting that we need to come up with a new sustainable and scalable option that can help feed a growing population without further taxing our oceans, forests, or fresh water supplies (Draaisma, Wijffels, Slegers, Brentner, Roy, & Barbosa, 2013). Microalgae have the potential to be a sustainable food and feed solution, but further development is needed to bring these organisms into mainstream food production.

The collection of organisms we refer to as algae is a polyphyletic group of very diverse species ranging from extremely large multicellular organisms like giant kelp to microscopic unicellular organisms like *Chlamydomonas sp*. The latter species can be included under the term microalgae, which is generally used to address eukaryotic unicellular organisms but in this review it will also include prokaryotic cyanobacteria for ease of nomenclature. Microalgae in particular will be the focus of this review. The fact that they are microscopic poses a number of advantages over their macroscopic counterparts such as simpler genetic manipulation, easier scale up processes, and higher protein content in general.

Microalgae have several attractive features for large scale sustainable production, like high biomass yields per unit area, and the ability to be grown on non-arable land using non-potable water, or even salt water. However, there are a number of improvements that will need to be made before microalgae can become a regular food source and an established crop. These improvements include

development of agricultural scale production, evaluating and perhaps enhancing nutritious content, optimizing yields, and developing improved organoleptic traits so that algae are appealing to the human or animal palate. To improve such traits, and even add new ones, it is necessary to understand that every phenotype depends on both the genotype and the environment in which the crop is grown. In terms of producing microalgae as a food source, improving the genotype requires choosing the appropriate algae strain and then genetically modifying the organism using both traditional breeding and genetic engineering techniques as needed. When it comes to improving the production environment, this requires choosing the best growth conditions in terms of infrastructure, climate, and media composition. Another way to view this is that the genotype determines the potential phenotype and the environment determines if and how much of that potential is achieved.

Historical use of algae as food

Algae have been consumed for thousands of years in different cultures (Dillehay, Ramírez, Pino, Collins, Rossen, & Pino-Navarro, 2008; Wells et al., 2017). The oldest known use of algae as a food sources is from Chile, in which archaeological records show consumption of algae dating to 14,000 years ago (Dillehay et al., 2008). Additionally, there are many written records showing algae's use as a food source throughout the world for the last several centuries (Aaronson, 1986; Miroslav & Zorica, 2008; Turner, 2003; Wells et al., 2017). There are even records from the Spanish conquistadors showing that the Aztecs harvested spirulina from the Lake Texcoco (Habib, 2008). However, none of these algae have been expanded as a major crop, likely due to a number of production and social acceptance issues, as algae are quite different from traditional crops.

Algae currently listed as Generally Recognized As Safe (GRAS) for human consumption

Generally Recognized As Safe (GRAS) is a status given by the Food and Drug Administration (U.S. Food and Drug Administration, 2018) to any substance or chemical, including sometimes whole organisms, that is considered safe for human consumption. There are two routes to obtaining GRAS status, the first is by years of documented consumption by humans, and the second by scientific evidence

that a substance is safe. Such status is required if an organism is going to be used as a food or food ingredient. If a GRAS organism is being used to produce a certain food additive or ingredient, the purification level of that product might be low or none, but if the organism is not GRAS then any ingredient purified from that organism will need to be purified to a high degree to ensure its safety. Purification can increase the cost of a product significantly.

There are only a few microalgae that have GRAS status as recognized by the FDA. These algae include: *Arthrospira platensis, Chlamydomonas reinhardtii, Auxenochlorella protothecoides, Chlorella vulgaris, Dunaliella bardawil,* and *Euglena gracilis*. Obtaining GRAS status requires time consuming and costly safety tests, which has limited the number of algal species with such status. This review will focus only on the organisms that have achieved GRAS designation.

However, the GRAS designation only applies to U.S. jurisdiction and it may differ from the regulations in other countries. In the European Union (EU), the European Food Safety Authority (EFSA) oversees the regulations pertaining to human food and animal feed within the EU. Taking a "precautionary principle", foods that have been significantly consumed within the EU territory prior to May of 1997 are deemed safe to be consumed, and any other food, excluding Genetically Modified Organisms (GMOs), are labelled as "novel food" and must undergo a safety assessment by the EFSA before being marketed (Authority, 2020b). A similar standard is held in Canada, where the Health Canada is the organization in charge of supervising food safety and stipulates that any food that is new or has changed compared to existing food products is classified as novel foods, including GMOs, and its safety must be assessed by Health Canada (Canada, 2020). Regulations in China, India and Japan have been consulted and the findings regarding the safety of microalgae as food are summarized. *Arthrospira platensis* has been found to be considered safe to consume in the Canada, China, EU, India, and Japan. Chlorella is also widely assumed as safe for human consumption, but the approved species of Chlorella varies among countries: *C. protothecoides* is approved in the U.S. and Japan, *C. pyrenoidesa* is approved in the EU and China, *C. vulgaris* is approved in Canada, EU and Japan, and *C. sorokiniana* and *C. regularis* are approved in Canada. *Chlamydomonas reinhardti*i is not considered safe in any of those countries. *Dunaliella bardawil* is not considered safe in any of those countries either, but *Dunaliella salina* is

considered safe in China and Canada. Finally, *Euglena gracilis* is approved in Canada, China and Japan (Authority, 2020a; Canada, 2019; India, 2016; Organization, 2011; PRC, 2016).

Algae nutritional composition

For algae to be considered as a potential new food sources, one crucial factor is their composition and nutritional content. The nutritional composition varies tremendously among algal species, and even within the same species nutritional content can vary significantly based on the growth environment, both the media composition as well as temperature and light regime. Important nutritional components to consider are protein and lipid content, as well as vitamins and mineral content, all of which are known to positively impact human health. The nutritional composition of the microalgae that have obtained the GRAS status to date can be found in Table 3.

Protein and amino acid composition

Protein is a crucial element in the human diet, providing most of the nitrogen humans need. At the same time, there are a subset of amino acids that cannot be synthesized by the human body and those need to be supplied in the diet. These essential amino acids are: histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan and valine (WHO, 2007). Additionally, there are other amino acids which may not be synthesized appropriately in certain conditions and are called conditionally essential, which consist of: arginine, cysteine, glutamine, glycine, proline and tyrosine. One of the most important factors in worldwide hunger is what is called Protein-Energy Malnutrition (PEM), which consist of a series of pathological conditions caused by deficient intake of essential amino acids and total energy intake (FAO, 2017). In order to solve this problem a new source of inexpensive and balanced protein is needed.

Some algae have a very high percentage of their dry biomass as protein. Species like *Arthrospira platensis* have been reported to have up to 70% of their biomass as protein content (Becker, 2007;

Griffiths & Harrison, 2009; Szabo, Matulka, & Chan, 2013). The rest of the GRAS species average about 40% protein content, which is relatively high if you compare it to other plant sources like soybean (38%), rice (~10%), pea (2.8%), or even animal sources like milk (4%) or eggs (13%) (United States Department of Agriculture, 2018; Chae, Hwang, & Shin, 2006). Not only do algae have a high protein content, but the composition of that protein is much richer in essential amino acids compared to common plant proteins (Matsuda, Hayashi, & Kondo, 2011; Miroslav & Zorica, 2008). In general, protein of plant origin is of lower quality than protein of animal sources. One of the main factors that determines such quality is whether a protein source contains all of the essential amino acids in adequate amounts, being those that do called complete proteins. Plants from certain groups are deficient in some essential amino acids, being corn deficient in tryptophan and lysine, cereals deficient in lysine and legumes deficient in methionine.(Friedman & Brandon, 2001; Gorissen, Crombag, Senden, Waterval, Bierau, Verdijk, & van Loon, 2018; Joy, Lowery, Wilson, Purpura, De Souza, Wilson, Kalman, Dudeck, & Jäger, 2013) The complete amino acid profile for each GRAS microalgae can be found in Table 4. Unlike most plant proteins, all of the GRAS algal species but *E. gracilis* contain all of the essential amino acids, making each of these algae a complete protein source.

Lipids and fatty acids

Lipids are an indispensable component of cells and are precursors of many essential molecules, and as such an appropriate intake of them is crucial for the human diet. Some algae can accumulate lipids to very high levels, for example *Auxenochlorella protothecoides* can accumulate up to 70% of dry biomass as lipids (Griffiths & Harrison, 2009; Miao & Wu, 2004). Just like essential amino acids, there are some lipids that are essential, including a-Linolenic acid and Linoleic acid. Additionally, there are certain lipids which have been proven to have a positive impact on human health, of which the most important are docosahexaenoic acid (DHA) and eicosapentaenoic acid (EPA) (Swanson, Block, & Mousa, 2012). The traditional source of such nutrients in human diets has been cold water fish, and seafood in general. However, fish are enriched in these omega-3 fatty acids because they consume plankton and algae as part of their diet, and it is these algae that actually produce these essential long chain

polyunsaturated fatty acids (PUFAs). Certain microalgae like *Phaeodactylum tricornutum* can accumulate up to 30% to 40% of total the fatty acids produced as EPA, and other species like *Schizochytrium sp*. can accumulate about 50% of the total lipids of the cell as DHA (Adarme-Vega et al., 2012; Wang et al., 2018; Ward and Singh, 2005).Therefore, algae can be an effective substitute for fish oil supplements, providing the healthy fatty acids that humans need in their diet without some of the drawbacks of using fish for these oils, such as odor, and the non-sustainability and non-vegetarian nature of omega-3 oils sourced from fish (A. L. M., A., P., Eileen, Gloria, Dror, Jacqueline, & Deanna, 2007).

Other nutrients

Besides protein and lipids, there are other nutrients that are crucial in human diet, like vitamins and minerals. Most of these are not synthesized by animals, but are produced by plants or other organisms, and are then provided to human and animals through their diet. Just like traditional vegetable foods, algae are very rich in vitamins and minerals. One example is the green alga *Dunaliella tertiolecta*, which has been shown to be a great source of vitamin A, vitamin B1, vitamin B9 and vitamin E. (Fabregas & Herrero, 1990). There are other nutrients that have a positive impact in human health that can be supplied by algae, like antioxidants (e.g., lycopene, b-carotene, and astaxanthin) or polysaccharides (β-Glucans) (Barsanti, Vismara, Passarelli, & Gualtieri, 2001; Fabregas & Herrero, 1990; Koller et al., 2014; Olaizola, 2000; Pulz & Gross, 2004). Additionally, the consumption of some microalgae has been correlated with health benefits including but not limited tocardiovascular health, immunomodulation, anti-aging and anticancerogenic(Caporgno & Mathys, 2018; Fallah, Sarmast, Habibian Dehkordi, Engardeh, Mahmoodnia, Khaledifar, & Jafari, 2018; Fields, Lejzerowicz, Schroeder, Ngoi, Tran, McDonald, Jiang, Chang, Knight, & Mayfield, 2020; Plaza, Herrero, Cifuentes, & Ibáñez, 2009). A recent study found that consumption of whole-cell Chlamydomonas reinhardtii mitigated weight loss in a murine model of acute colitis and positively impacted gastrointestinal health in humans (Fields et al., 2020). These are a few common examples, but we believe the vast and unexplored biodiversity of algae will reveal new bio-active molecules that could further supplement and improve the human diet (Koller et al., 2014; Pulz & Gross, 2004).

Current utilization of microalgae as food and nutrition products

The total global algae market is difficult to estimate due to the very large and diverse set of products that are sold, as well as the generally small scale for each individual product. Estimates for macroalgae products, mostly in the form of alginates and carragenaans, indicate an annual market value of approximately 6.7 Billion USD (Wells et al., 2017). Many seaweeds are traditionally used in Asian cuisine, like Nori used in Japanese sushi, which are produced from dried *Porphyra* leaves, with a market size of 1 billion USD per year (Pulz & Gross, 2004). Microalgae also have large markets as food ingredients, with a value well above 1.25 billion USD per year, although that number has been calculated only for dry biomass and not for the many processed products that can be obtained from these organisms like food supplements and food additives (Pulz & Gross, 2004). A summary of the top microalgae produced by weight can be found in Table 1.

Algal natural products

Algae are a highly diverse family of organisms, including both macro and microalgae, as well as cyanobacteria. Members of this polyphyletic group can be found in a wide range of environments, thus a diversified array of metabolisms can be found in algae. With such an extremely large pool of species and metabolites, it is safe to say that algae natural products constitute one of the largest untapped resources on the planet. It would be impractical to attempt to describe even a fraction of these natural products, so here we will focus on natural products only from the green algae, and mainly on those that are associated with food or nutritional products.

Algae are typically rich in protein, lipid, and many other components that are valuable for human health. One of the most prominent examples are the omega-3 fatty acids; DHA and EPA, which are usually introduced into the human diet through the consumption of fish. Cold water fish are rich in omega-3 oils not because they produce these oils themselves, but rather because they directly or indirectly obtain them from algae as part of their normal food chain (A. L. M. et al., 2007). Because many

algae naturally produce omega-3 oils, and because natural fish stocks are in decline, a market has developed for omega-3 fatty acids produced in algae, both as an aquaculture feed ingredient to introduce these oils into fish meal that lacks them, and also as nutritional supplements for direct human consumption. Together these products account for an approximately $500 million USD market today (Brennan & Owende, 2010; Koller et al., 2014).

Other products commercially produced in algae are pigments, such as B-carotene and astaxanthin. B-carotene is transformed in the human body into the essential vitamin, vitamin- A, and natural B-carotene is also used as food colorant, representing a market size of $200 million USD, some of which is produced in the green algae *Dunaliella salina* (Ami, Adriana, & Mordhay, 1982; Koller et al., 2014; Spolaore et al., 2006). Astaxanthin, also a carotenoid, has a distinctive red color that is currently used in animal feeds to confer a yellow or red color in the final product, with its biggest application in farmed-raised salmon and chicken egg yolks. With a global market size estimated to be around $200 million USD (Koller et al., 2014) most of this pigment is produced synthetically, but it is also able to be produced naturally by cultivating the microalgae *Haematococcus pluvialis*.

Some polysaccharides have been shown to significantly improve human health when supplemented in the diet. One such example are the β-Glucans, which are a group of polysaccharides comprised of D-glucoses bound together through β linkages and can be typically found in the cell walls of plant, fungi and bacteria. When fed to humans, β-glucans act as soluble fiber, which has been shown to decreases LDL cholesterol and reduce the risk for cardiovascular diseases (El Khoury, Cuda, Luhovyy, & Anderson, 2012). This kind of molecule is produced by many green algae, thus enhancing the nutritional value of products containing such algae (Spolaore et al., 2006). A summary of the most important microalgal products sold by global market size in USD can be found in Table 2.

Algal recombinant products

Although there are many existing genetic tools that make it possible to produce molecules of interest through recombinant DNA techniques in algae, there are few recombinant products from algae

being marketed today. There are however, a number of potential recombinant products that have been successfully made in algae, most of these have been produced in the alga *Chlamydomonas reinhardtii* (Anderson, Muff, Georgianna, & Mayfield, 2017; Barrera, 2013; Fields et al., 2018; Stephen P. Mayfield et al., 2007; Molino, de Carvalho, & Mayfield, 2018; Beth A. Rasala et al., 2014; F. Scott, Binh, Ekem, & P., 2002; M. Tran, Van, Barrera, Pettersson, Peinado, Bui, & Mayfield, 2013). These proof of concept products demonstrate that the technology can be deployed in the near future, if they are coupled with advances in algae growth technologies that enable low cost production. The vast biodiversity within algae, offers the potential for the biotechnological production of high valuable molecules, many of which are not easily produced in any other system (M. Tran et al., 2013). Most of the high value recombinant proteins that can be made in algae are therapeutic proteins, but there are some recombinant proteins that can be used to positively impact algae's nutritional value. Some proteins can retain their biological activity when they are ingested, thus affecting human health, and are referred to as "functional proteins". An example of them are colostrum proteins like osteopontin, which are naturally present in breast milk and have been shown to impact brain development and immune system function (Bo, Jin, Yue, & Feng, 2016; Lönnerdal, 2014). Another relevant functional protein found in breast milk is Immunoglobin A, which is an antibody found in most body secretions due to is antimicrobial activity (Vukavic, 1983). With additional advances in algae biotechnology, especially on low cost scaled production, the potential for recombinant protein production in algae will be realized.

All of these characteristics make algae a nutritional source that has enormous potential to provide a scalable solution to mankind's need to feed an ever-growing population in an affordable and complete manner. However, for this potential to be realized, algae must be domesticated with improved to reach the desired characteristics that traditional agriculture crops possess.

Techniques and methods to improve algae as a crop

Traditional agricultural crops have been improved over centuries by constant selection of strains that had more and more desirable traits. Such improvement strategies have ultimately yielded crops with key traits that are socio-economically viable in agriculture such as: high biomass productivity, appropriate

nutrient composition, lack of toxins and other adverse components, appealing organoleptic features, resistance to abiotic and biotic stresses, among many others. Even though algae have the genetic potential to display all of those characteristics at more than satisfactory levels, there has been no systematic selection of such traits in algae. This lack of domestication has significantly slowed the adoption of algae as a food source, but this also shows the potential of algae, as applying the well-established processes for domestication to algae is likely to yield strains highly suitable for large scale agriculture.

To improve such traits, and even add new ones, it is necessary to consider that every phenotype depends on both the underlying genotype and the environment the organism is grown in. Thus, if we intend to develop a microalga with a suitable phenotype to become a commodity food, then we need to improve both the genotype and the growth environment. In terms of producing microalgae as a food source, improving the genotype means to choose the appropriate algae starting strain and then improve this strain over time by selection following genome alteration either though mutagenesis, breeding, or genetic engineering, as appropriate. When it comes to improving the environment, this can translate into choosing the best growth conditions in terms of growth process, climate, and media composition. The genotype sets the potential of the environment determines how much of that potential is achieved.

Improving production processes

The production process under which algae are grown includes the growth media (i.e., salt or fresh water), cultivation system (i.e., open ponds or bioreactors), and relevant biotic and abiotic factors, like the temperature, light intensity, dark-light cycle, light wavelength composition and other components of the local climate, as well as co-cultivated organisms and pest and pathogens. Small variation in any of these environmental factors can result in large differences in algae productivity and biomass quality, especially when products like food, are the desired output (X. Feng, Walker, Bridges, Thornton, & Gopalakrishnan, 2014; Guccione, Biondi, Sampietro, Rodolfi, Bassi, & Tredici, 2014; Liang, Sarkany, & Cui, 2009).

Media composition

One of the key components of the growth process is the media used for cultivation. From providing a reduced carbon substrate to using minimal media without any organic carbon substrate, the composition of the media will greatly impact how the cells will grow and what products they will produce. For instance, when most algae are starved for nitrogen they are induced to accumulate lipids to very high levels (Binnal & Nirguna Babu, 2017; Z. T. Wang, Ullrich, Joo, Waffenschmidt, & Goodenough, 2009). Many other nutrients and minerals must also be supplied in the growth media, and the concentration of those is vital for obtaining the best growth yields. Optimizing media formulation is a classical approach to increase product yields in microbiology and is important regardless of the cultivation environment. Media optimization can be done with the classical one-variable-at-a-time method or with more complex statistical methods like the Box-Behnken design (Cheng, Ren, & Ogden, 2013; Kanaga, Pandey, Kumar, & Geetanjali, 2016).

Growth systems

The other key component of the production process that affects the phenotype is the growth system, or the infrastructure used to cultivate the microorganisms. For algae there are three primary options: photosynthetic growth in open ponds, photosynthetic growth in closed photobioreactors, and heterotrophic growth in bioreactors (Borowitzka, 1999a; Y.-K. Lee, 2001). For heterotrophic production, stainless steel fermentation equipment is often used, which can be very expensive, but yields from these systems can be 50 or 100 times greater than photosynthetic yields, and the biomass produced is often of higher quality in terms of higher cell density, as well as higher percentage of the molecule of interest per biomass unit (Borowitzka, 1999b; Davis, Aden, & Pienkos, 2011; Davis, Markham, Kinchin, Grundl, Tan, & Humbird, 2016; Sheehan, Dunahay, Benemann, & Roessler, 1998; Stephenson, Kazamia, Dennis, Howe, Scott, & Smith, 2010). The choice of which system to use often comes down to the nature of the product and the market price that can be obtained for that product. A higher value product will allow for a

more expensive system to be used, whereas a cheaper product must be grown in inexpensive open ponds.

It is of importance to mention that open pond cultivation still needs technological improvements to be competitive with traditional agriculture, but the biggest opportunity to reduce cost will likely come from the reduction in marginal costs as a result of a dramatic increase in the production scale, a concept commonly known as “economy of scale” (Davis et al., 2016). Just like traditional farms, producing cheap food products only becomes economically competitive at large scales, so it is not surprising that open pond cultivation has been modeled to be the most likely way of achieving algae mass production for commodity food and feed products (Norsker, Barbosa, Vermuë, & Wijffels, 2011). However, for specific products like nutritional supplement of high value, the options of photobioreactor and heterotrophic bioreactor are appealing for their controlled conditions and high yields. For example, for producing the valuable food colorant astaxanthin, the alga *Haematococcus pluviallis* has been economically produced in large photobioreactor systems (Olaizola, 2000).

Improving product yield

Another key trait desired in any agricultural crop is a robust and reliable yield. Under the right conditions, algae can have higher biomass productivity than any photosynthetic crop on Earth. On top of that, algae can be grown on non-arable land and using non-potable water (Dismukes, Carrieri, Bennette, Ananyev, & Posewitz, 2008) (Hannon et al., 2010), and since both of these resources can be inexpensive, the production cost of algae biomass can be extremely low.

However, algae grown photosynthetically do not accumulate to high cell densities due to the shading effect that the cells exert on each other, thus having to be grown in shallow ponds or with intensive mixing methods (Shigesada & Okubo, 1981). For that reason, algae strains have been engineered to have a truncated light-harvesting chlorophyll antennae, so that cells have less shading capacity, and thus the culture would show a higher photosynthetic productivity (Tetali, Mitra, & Melis, 2007). Another challenge to outdoor growth of algae is the natural variability of the local climate and

weather. Therefore, the selection of the appropriate strain that can deal with such varying conditions is important. Crop rotation is a potential technique that can be implemented, first selecting strains that have rapid growth in the summer and then a different strain that can tolerate colder conditions in the winter (Jorde, Leya, Thomas, Pereira, M Badenes, Santos, Costa, Verdelho, Friedl, & Kryvenda, 2017). This is possible since algae do not have flowering seasons that would restrict the harvesting of the algae crop to a certain time of the year.

One of the biggest challenges of outdoors growth is culture contamination, which can lower productivity and diminish the quality of the biomass. A common approach is to grow an extremophile alga in said conditions, so the number of possible contaminants is greatly reduced. Examples of that are *Dunaliella salina* which can grow in NaCl concentrations of up to 3 molar and *Arthospira platensis* which grows at alkaline pH of 9 to 11 (Varshney, Mikulic, Vonshak, Beardall, & Wangikar, 2015). Other production strains can be genetically engineered to survive in extreme conditions or other naturally extremophiles can be engineered to make a food product. A new potential approach would be the introduction of a phospite oxidoreductase gene, which encodes for an enzyme that transforms phosphite to phosphate, allowing algae to grow efficiently using phosphite as the only source of phosphorous. With this strategy, only the algae expressing said gene can transform phosphite into phosphate out-competing other possible contaminant species. This was demonstrated in *C. reinhardtii*, which was able to grow successfully to high cell density in wastewater full of bacteria and other biological contaminants (L. Q. M. M., Antonio, Gilberto, Lenin, Mauro, Damar, & Luis, 2016). An alternative approach to generate microalgal biomass is to culture a consortium of microorganisms rather than an axenic culture. This would present a robust mini ecosystem that would be more resistant to invading species due to the fact that all niches would already be occupied. However, it would be complicated to predict how different species would grow together and if there would be big fluctuations in the composition of the final product between batches. More research is needed in this area, but the potential benefits of culturing robust microalgal communities are very promising (Shurin, Abbott, Deal, Kwan, Litchman, McBride, Mandal, & Smith, 2013).

It is extremely important to choose the right media, as this will dictate growth rates, nutritional profile of the algae, and rate of contamination. In the case of bioreactor growth, that means autotrophic or heterotrophic growth are available choices. Although it is true that not all algae are able to uptake organic carbon sources, some researchers were able to enable strictly photosynthetic algae to take up glucose by adding a glucose transporter gene to the cells (Zaslavskaia, Lippmeier, Shih, Ehrhardt, Grossman, & Apt, 2001). Such strategies are important because even if some algae can grow with organic carbon, having an ability to modify what substrate can be consumed can enable algae to be grown cheaper, which may be essential to achieve commodity pricing for algae biomass. Additionally, the infrastructure system used to grow the algae is extremely important. In outdoor pond one of the most important factors is effective mixing with the lowest energy requirement, which will require further engineering efforts to improve (Borowitzka & Moheimani, 2013).

Finally, the biggest improvement in costs will come when improved yields and productivity are combined with scale to allow for economy of scale to drive the prices down. For that, the infrastructure needs to be developed and the crops deployed, and a steady production of algae biomass achieved at scale. Only then will the economics of this new crop become viable.

Improving the genotype

The first step in improving the genotype of the microorganism used in a bioproduction system is to choose the best starting strain through bioprospecting. Then, said strain can be improved through traditional breeding, mutagenesis and selection. Additionally,, there are a variety of molecular tools available to engineer the genomes of microalgae. These molecular tools can be used to strategically alter genetic information to appropriately fine tune algae in order to optimize production of a desired product. These techniques are currently limited by the availability of genomic and gene expression data of specific algal species, physiological barriers (i.e., cell size or rigid cell walls), and the availability of site-directed mutagenesis tools for the desired algal species.

Bioprospecting

When designing a production pipeline for a specific algae product, it is essential to choose an appropriate microalga as the starting strain. Choosing the wrong algae, simply because it is a traditional species and it is relatively well understood, can be a set up for failure. The organism of choice should always be the one that is closest to the final phenotype, both in terms of growth characteristics and metabolic pathways available. Fortunately, microalgae are an enormously diverse group of microorganisms which offers an untapped resource of biotechnologically relevant species, making bioprospecting the first step in the process of identify the optimal bioproduct production strain.

Traditional breeding, mutagenesis and selection

Improvements in traditional crops have historically come from breeding and selection, and more recently from molecular genetic technologies. Breeding approaches have yet to be used to develop improved microalgae phenotypes for any algae that are currently sold. There are two main barriers preventing successful breeding of microalgae: (1) most algae do not easily go through sexual cycles under controlled conditions, and (2) only certain phenotypes can be easily screened for in microorganisms. However, some species, like *Chlamydomonas reinhardtii,* have successfully been bred and selected for new phenotypes (Fields et al., 2019). A *Chlamydomonas reinhardtii* strain that could tolerate high salinity (up to 700 mM NaCl) was developed by mating two *C. reinhardtii* strains, that could only tolerate mild salinity (300 mM NaCl) and selecting for progeny with high salt tolerance (Takouridis, Tribe, Gras, & Martin, 2015). This technique, also called genome shuffling, shows that by introducing genetic variation and selecting for the right phenotype, new traits can be acquired.

Random mutagenesis is another technique available to use with many organisms and can effectively produce genetic variation that leads to phenotypic advances. Such approaches have been implemented in traditional agriculture for centuries, and this technique clearly has potential to be used in microalgae. Due to high throughput selection technologies, like Fluorescent Activated Cell Sorting (FACS), rare phenotypes can be detected (Mia, S., E., & C., 2015), thus mimicking what has been done

in traditional agriculture through thousands of years of selection, but with high throughput screening similar results can be achieved in just weeks or months. This procedure has the advantage that genetic knowledge is not required to identify the new trait, since the genetic variation introduced is random and the selection is applied directly on the phenotype. This is especially relevant in algae for which genetic information is limited, making this technology available for strain improvement that is broadly applicable to newly discovered species.

Sequenced and annotated genomes

The availability of a sequenced and annotated genome for a desired production species will facilitate any attempt to optimize the genome using any molecular or gene editing tool. Having a sequenced genome that is properly annotated lets researchers assess which genes and metabolic pathways are already present in the organism and makes it possible to predict which artificial pathways can be easily implemented. Not only is this genomic information useful for recognizing already present genes, but it can also be used to characterize the regulatory sequences in the genome, which will allow for the design of more efficient genetic tools.

Substantial effort has been put towards sequencing microalgal genomes. Since cyanobacterial genomes (generally less than 10 MegaBases) are smaller than eukaryotic algae, more of these bacterial genomes have been sequenced (Lü, Sheahan, & Fu, 2011; Qin, Lin, & Jiang, 2012; Radakovits, Jinkerson, Darzins, & Posewitz, 2010). To date, 85 cyanobacterial genomes belonging to over 30 different species have been completely sequenced, and a total of 376 genomes have been totally or partially sequenced. Due to the advancement of next-generation sequencing technologies, and their constant decreasing cost, the list of sequenced genomes is constantly increasing and can be accessed through the online database “Cyanobase” (Z. T. Wang et al., 2009).

Eukaryotic microalgae have 3 separate genomes which can be sequenced: nuclear, chloroplast, and mitochondria. Efforts in genetic engineering of those genomes have been focused mainly on the nuclear and the chloroplast genomes, since those have the most potential for genetic engineering.

Chloroplast, like cyanobacteria, have small genome size, making them much easier to sequence. Currently, up to 9 algae nuclear genomes have been completely sequenced including the species *Chlamydomonas reinhardtii*, *Chlorella variabilis, Micromonas pusilla, Ostreococcus lucimarinus, Ostreococcus tauri, Volvox carteri, Phaeodactylum tricornutum, Thalassiosira pseudonana and Cyanidioschyzon merolae.* When it comes to chloroplast genomes, 47 different species have been sequenced to date, and for the mitochondria genome, 42 species have been sequenced (Lü et al., 2011; Radakovits et al., 2010; Shrager, Hauser, Chang, Harris, Davies, McDermott, Tamse, Zhang, & Grossman, 2003; Wakasugi, Nagai, Kapoor, Sugita, Ito, Ito, Tsudzuki, Nakashima, Tsudzuki, Suzuki, Hamada, Ohta, Inamura, Yoshinaga, & Sugiura, 1997). As with the cyanobacteria genomes, the number of eukaryotic algae sequenced is always increasing, as new potential production strains are identified. Many of these species have not been tested for human consumption and would require costly downstream processing and refinement of a specific product to avoid any potential unwanted side effects. Focusing efforts on GRAS species will quickly bring cheaper products to market.

Genetic transformation techniques

Some cyanobacterial species are competent to be transformed with DNA that is present in their environment without any mechanical and/or chemical pretreatments (Johnsborg, Eldholm, & Håvarstein, 2007). However, for most microalgal species, especially the eukaryotic ones, cells are unable to be transformed so easily and they need to undergo some carefully designed techniques to enable uptake of heterologous DNA. All three different genomes present in eukaryotic algae can potentially be transformed.

The vast majority of DNA in any algae is found in the nuclear genome, and hence the potential for metabolic engineering is primarily in the nuclear genome. At present, more than 30 species from different divisions have been shown to be able to uptake and express foreign DNA. From the chlorophyte group these are; *Chlamydomonas reinhardtii*, *Chlorella ellipsoidea*, *Chlorella kessleri, Chlorella saccharophila, Chlorella sorokiniana, Chlorella vulgaris, Haematococcus pluvialis, Dunaliella salina and Dunaliella viridis*. From the rhodophyte group; *Cyanidioschyzon merolae* and *Porphyridium sp*, and from the

heterokontophytes; *Nannochloropsis oculata*. Diatoms, the most diverse group of algae, have the species *Chaetoceros sp, Cyclotella cliptica, Cylindrotheca fusiformis, Navicula saprophila, Phaeodactylum tricornutum, Thalassiosira pseudonana* and *Thalassiosira weissflogii* with transformation capability. Finally, from the dinoflagellates both *Amphidium sp* and *Symbiodinum microadriaticum* have been transformed, and from the eugleunids only *Euglena gracilis* can be transformed, so far (Anila, Chandrashekar, Ravishankar, & Sarada, 2011; Apt, Grossman, & Kroth-Pancic, 1996; Arisa, Takuma, Nozomu, Haruo, Kouhei, & Masao, 2011; Chow & Tung, 1999; Doetsch, Favreau, Kuscuoglu, Thompson, & Hallick, 2001; Falciatore, Casotti, Leblanc, Abrescia, & Bowler, 1999; S. Feng, Li, Xu, & Qi, 2014; Kawata, Yano, Kojima, & Toyomizu, 2004; Kilian, Benemann, Niyogi, & Vick, 2011; Lü et al., 2011; Nicole, M., & Nils, 2006; Radakovits et al., 2010; Shimogawara, Fujiwara, Grossman, & Usuda, 1998).

State of the art of genetic tools

If genetic engineering is going to be used for trait improvement, there are a number of requirements that will need to be met. First, techniques to introduce DNA into the genomes of specific algae will need to be developed. Basic tools already exist for many algae species under production today, including Spirulina, Chlorella, Chlamydomonas, Dunaliella and Euglena (Anila et al., 2011; Chow & Tung, 1999; Doetsch et al., 2001; Kawata et al., 2004). Similar tools will need to be developed for newly discovered algae, as they become commercially relevant. Additional advancements will need to be made, including the use of CRISPR and other genome modifying tools that allow for precise alterations of specific genes, leading to a more accurate recombinant or native gene expression. As the availability of sequenced and annotated genomes of microalgae increase, this should lead to the more efficient development of new genetic tools, as well as shed light on the metabolic pathways present in algae that could have significant biotechnological application. Finally, the development of genetic tools that allow for highly controlled gene expression will be essential.

“Genetic tools” is a term that comprises many different technologies needed for successful recombinant gene expression, as well as the tuning of endogenous gene expression. Genetic elements include promoters, UTRs, enhancers, selection markers, and knock-out and knock-down technology.

Having appropriate genetic tools dictates the extent and precision of the genetic modification that can be made, thus the availability of such tools for algae is a crucial element to enhance the value of algae as a food source. Algae genetics are different from other established model organisms like *Escherichia coli, Saccharomyces cerevisae* or CHO cells, such that the genetic tools used for these traditional biotechnological hosts do not work in algae. The only organisms to have shown some sort of compatibility in terms of genetic elements are plants, but even those are not efficient in algae.

The use of classic recombinant promoters, including viral promoters like CaMV35S and SV40, which have been shown to have some success in a very broad spectrum of organisms (Anila et al., 2011; Barrera, 2013; Bo et al., 2016; S. Feng et al., 2014; Qin et al., 2012; Tang, Qiao, & Wu, 1995), have failed to produce similar results in algal systems. However, endogenous promoters, or those from other taxonomically close algae, have proven to work in several algae species (Barrera, 2013; Qin et al., 2012). Examples of such promoters include the fucoxanthin-chlorophyll a/c binding protein (fcp) promoter from diatoms, which has been shown to work in other diatoms and marine algae (Apt et al., 1996; Barrera, 2013; Huang & Daboussi, 2017; Qin et al., 2012). Other algae promoters that have also been used successfully, are the promoters from two different violoxanthin/chlorophyll a-binding protein VCP1 and VCP2, that were shown to work in *Nannochloropsis sp*, for stable recombinant expression (Barrera, 2013; Kilian et al., 2011; Qin et al., 2012). For some green algae, exogenous recombinant promoters like the cauliflower mosaic virus (CaMV35S) or Ubi1 from maize, have been shown to work (Bo et al., 2016), while actin1 promoter from rice, was shown to work in Chlorella. Endogenous promoters also work, and often much better then exogenous promoters, including the nitrate reductase promoter (NR) which is repressed by ammonium and activated by nitrate, and the rbcS from *C. reinhardtii*, which also has light induction (Bo et al., 2016). In Dunalliela, the inducible endogenous promoters that have been successfully used are the nitrate reductase promoter, and the duplicated carbonic anhydrase 1 promoter (DAC1), which is responsive to the sodium chloride concentration (S. Feng et al., 2014; Jia, Li, Allen, Feng, & Xue, 2012; Qin et al., 2012).

Many genetic tools have been developed for the green algae *Chlamydomonas reinhardtii*, for both constitutive and inducible expression as well as nuclear and chloroplast expression (Barrera, 2013).

Recombinant protein expression has been best achieved in the chloroplast, where simpler genetic systems, including use of the endogenous constitutive promoters psbD, atpA and rbcL and the endogenous inducible promoter psbA (light induced) have resulted in reasonable levels of recombinant protein accumulation (Barrera, 2013; Stephen P. Mayfield et al., 2007). In Euglena, only moderate success was achieved at recombinant gene expression in chloroplast using the endogenous psbA promoter and 3'UTR (Doetsch et al., 2001). For nuclear encoded genes in *C. reinhardtii*, the endogenous promoters from rbcs2, B-2-tub, Nos, hsp70A, nia1 and cyc6, have all been shown to result in good recombinant gene expression (Barrera, 2013; Beth A. Rasala et al., 2014; Scranton et al., 2016). Even though recombinant gene expression in *C. reinhardtii* is the best understood of any microalgae, even this alga is far behind other systems like *E. coli* and yeast. Thus, improving these genetic tools, using molecular and synthetic biology approaches, has recently become a topic of interest, with some success to date. Another method to identify new promoters for recombinant expression is a technique called promoter trapping, in which a gene conferring a selectable phenotype is transformed into algae without a promoter, thus the gene is expressed only when it happens to land immediately downstream of a strong promoter (Vila, Díaz-Santos, de la Vega, Rodríguez, Vargas, & León, 2012). If a fluorescent protein is used, then the transformants can be sorted with a FACS machine and the strongest expressers isolated and identified (Scranton et al., 2016).

Other regulatory elements, like UTRs and enhancers, are also important for high recombinant gene expression. Unfortunately, these can be difficult to characterize, thus no solid knowledge on the field exist yet in microalgae. Selectable marker genes are also essential elements for proper recombinant gene expression, and marker genes for both cell survival from genes that confer antibiotic resistance and genes that compliment an auxotrophic mutation, have been identified and shown to work well in many algal species.

Finally, it is important to develop tools for targeted mutations, whether those are knock-in, knock-down or knock-out mutations, as all of these will be required to develop commercial algal phenotypes. The emerging technology of CRISPR-Cas has been successfully deployed in several algal species to date (Nymark, Sharma, Sparstad, Bones, & Winge, 2016; Qintao, Yandu, Yi, Li, Shi, & Jian, 2016; Shin,

Lim, Koh, Kim, Kang, Jeon, Kwon, Shin, Lee, Hwangbo, Kim, Ye, Yun, Seo, Oh, Kim, Kim, Jeong, Chang, & Jeong, 2016), and it is expected that the list of algae suitable for CRISPR modification will continue to expanded.

Improving the quality of algae as a food source

For algae to become an established agricultural crop it needs to be improved so it can be viable, nutritionally, socially, and economically. The economic and nutritional viability are obvious requirements, but scientist tend to forget that a product as important as food needs to have a social acceptance as well, or it will not be incorporated into the diet. For that reason, the traits that will need to be improved are not just productivity of the strain and its nutritional content, but also its organoleptic traits so they become appealing to human or animal palate.

Nutritional content improvement

Protein content and amino acid profile improvement

One of the most important nutritional characteristics to improve in algae is protein content and quality, which is determined by many genes. To increase the chances of obtaining the right phenotype, it will be essential to have the ability to use high throughput screening to obtain variants with the desired phenotype. There are several colorimetric assays for total protein content measurement, but having the ability to measure total protein content in living cells, and then be able to isolate the phenotype of interest in a cell sorter, would be a tremendous advantage. One such approach would be to use a Raman Activated Cell Sorter, which is a device that can characterize the composition of living cells based on their Raman Spectra, and then separate cells that present the phenotype of interest (Lau, Lee, & Chan, 2008; Song, Yin, & Huang, 2016; P. Zhang, Ren, Zhang, Shan, Wang, Ji, Yin, Huang, Xu, & Ma, 2015; Q. Zhang, Zhang, Gou, Mou, Huang, Yang, Xu, & Ma, 2015). It is unlikely that this method could detect

difference in protein quality, so other methods of selection will need to be developed for these qualitative traits, which can then be coupled with a method to generate genetic variation like UV mutagenesis or breeding, to obtain strains accumulating both high levels of protein as well as protein with the correct nutritional profiles.

In addition to selecting the correct high protein content strain, an appropriate production process will need to be developed that allows the high protein phenotype to be expressed. It will be essential to design suitable media for high protein production, which will require choosing the right nitrogen source and concentration, as well as other micronutrients. Additionally, growing algae in autotrophic or heterotrophic systems can have a dramatic effect on cell composition, so the most suitable growth system must also be chosen for each strain. Finally, it is important to have a good amino acid profile, in which essential amino acid are present in appropriate quantities. If the selected strain lacks any essential amino acid, this can theoretically be overcome by expressing a protein or peptides specially designed to be enriched in the needed amino acids. (Ma, Zhang, Xu, Zhang, Li, Fan, Xie, & Chen, 2017).

Lipid content and profile improvement

Lipid content in microalgae is generally very high, but depending on the strain and growth condition, those lipids can vary significantly. As with protein content, it is possible to use high throughput selection procedure that permits the isolation of strains with high lipid accumulation. One of the easiest ways to do this is to use a lipid specific fluorescent dye, coupled with Fluorescent Activated Cell Sorter (Mia et al., 2015). Lipid specific dyes include; Nile Red, which needs pre-treatment in some cases for the dye to penetrate the cell wall, or BODIPY 505/515 (Beacham, Macia, Rooks, White, & Ali, 2015; Mia et al., 2015). As in any high throughput selection method, genetic variation needs to be added to the population to increase lipid content, and that can easily be done with UV mutagenesis, breading, or other methods to increase genetic variation (Beacham et al., 2015). There are several mutations that are known to result in increased lipid content, including those with an inability to store energy in starch (Work, Radakovits, Jinkerson, Meuser, Elliott, Vinyard, Laurens, Dismukes, & Posewitz, 2010), thus shunting their energy into oils as the main source of energy storage. It is again very important to adjust the growth

environment to achieve the desired phenotype; lipids often accumulated to much higher levels when growth conditions are suboptimal. That means that total biomass productivity and optimal lipid accumulation do not necessarily occur under the same growth conditions. For that reason, it is common that algae are first grown in optimal conditions to produce maximum biomass accumulation, and when high cell density is achieved, the growth conditions are altered to promote lipid accumulation. Such conditions have traditionally consisted of nitrogen starvation, which triggers lipid synthesis and accumulation, and decrease in protein accumulation (Z. T. Wang et al., 2009).

Other improved nutrition opportunities

There are number of molecules with proven health benefits that could be engineered for production in algae. For instances, colostrum proteins, which are proteins present in breast milk, have been shown to be essential for proper development of infants, as well as enhancing the innate immune system in the gut (Jiang & Lönnerdal, 2016; Larson, Wei, Weber, Mack, & McDonald, 2003; Lönnerdal, 2014). Recombinant colostrum and milk proteins produced in algae could be very important as dietary supplements for infants in developing countries, since malnutrition can lead to poor brain development, while gut infections can lead to diarrhea, which is one of the main causes of infant mortality in the world (Black, Morris, & Bryce, 2003; Liu, Johnson, Cousens, Perin, Scott, Lawn, Rudan, Campbell, Cibulskis, Li, Mathers, & Black, 2012; Martorell, 1999). Colostrum proteins can be produced in algae, and due to their GRAS status, can be orally delivered from intact algae without purification, thus delivering the proteins to the gut where they will perform their action (L., Verónica, H., T., Miller, Annika, L., & P., 2007). Another example are carotenoids, which have shown great benefits as vitamin precursors and antioxidants, and are natural products in algae whose expression can be improved by metabolic engineering, as was demonstrated by overexpression of phytoene synthase gene from *Chlorella zofingiensis* in transgenic *Chlamydomonas reinhardtii* (Cordero, Couso, León, Rodríguez, & Vargas, 2011). Additionally, *Haematococcus pluvialis*, which naturally produces high yields of carotenoids, was genetically modified to express a modified phytoene desaturase from the same species that resulted in the strain accumulating astaxanthin to much higher levels (Steinbrenner & Sandmann, 2006). Research on this area has grown

tremendously in recent years and continues to do so as it is a very promising field. For further information on the topic readers can refer to the very complete review by M. Gong and A. Massi (Gong & Bassi, 2016).

The potential for improving the nutritional quality of algae by increasing the accumulation of already present nutrients, as well as the addition of new ones, in a system that is edible and easy to mass produces, is immense.

Improving organoleptic traits

The fact that algae are very nutritious is important, but just being nutritious is unlikely to be enough for humans to incorporate algae as a significant part of their diet. Organoleptic traits, like flavor, aroma, and texture, will be key factors for humans, or even animals, in accepting algae as part of their diet. Many algae today are generally unappealing to the human palate, as they lack qualities that have been selected in other plants as desirable as food by humans.

It is necessary to engineer algae so its taste and smell are more attractive and satisfying. However, those two characteristics are determined by such a large number of traits that it is currently unfeasible to fine tune each one of them to obtain the desired outcome. Flavor itself has been attributed to hundreds of different volatile compounds(Calkins & Hodgen, 2007). Due to its complexity, it is troublesome to screen for appealing flavors and/or aromas in a high-throughput manner. There are efforts being made towards and automated system that can taste and smell, providing an objective analytical tool that could potentially be implemented towards high-throughput screening of new pleasing organoleptic traits. An example are the so called Electronic Tongues, defined by the IUPAC as: “The electronic tongue is a multisensor system, which consists of a number of low-selective sensors and uses advanced mathematical procedures for signal processing based on the pattern recognition (PARC) and/or multivariate analysis” (Vlasov, Legin, Rudnitskaya, Di Natale, & D'Amico, 2005). These instruments are capable of measuring sweet, salty, sour, bitter and umami flavors in an objective and reproduceable manner but are not yet capable of assessing the hedonic aspect of taste without additional data derived

from human trials (Hayashi, Yamanaka, Toko, & Yamafuji, 1990; Lorenz, Reo, Hendl, Worthington, & Petrossian, 2009; Podrażka, Bączyńska, Kundys, Jeleń, & Witkowska Nery, 2017).

Even without high-throughput flavor and smell screening, there are ways to improve the organoleptic traits of algae. These organisms can have their taste modified by either direct alteration of their genetic information or by processing and cooking the derived food products. For example, through DNA recombinant techniques good tasting molecules can be engineered into microalgae. Proteins are an ideal target to use as simple flavor modifying molecules due to their capacity of being expressed introducing just a single gene. Some have being identified as intense sweeteners, having a sweetness hundreds and even thousands of times than that of sucrose on a weight to weight basis (Kant, 2005). A very interesting case is that of Brazzein, a thermostable protein made of 54 amino acids that was found in the African plant *Pentadiplandra brazzeana (Ming & Hellekant, 1994).* As a protein, it's caloric content per weight is very similar to that of carbohydrates, 4 kilocalories per gram. Since Brazzein has been assessed to be 2000 times sweeter than sucrose on a weight basis, Brazzein can be considered a non-caloric sweetener (Ming & Hellekant, 1994). Similar examples are Thaumatin and Monellin, both being approximately 3000 times sweeter than sucrose on a weight basis and both are proteins that originate from the fruit of tropical plants (Morris & Cagan, 1972; van der Wel & Loeve, 1972). Other proteic molecules that have relevant organoleptic traits are the non-essential amino acid Glutamic Acid and glutamate salts like monosodium glutamate (MSG). These related molecules elicit a flavor known as "umami" and has been used in the food industry as an additive for many years, especially in Asia where its use originated (Kurihara, 2009; Yeomans, Gould, Mobini, & Prescott, 2008).

Other molecules that are known to provide attractive organoleptic traits to foods are lipids. If a few target fatty acids were identified as good flavor sources, they could be engineered to be produced in algae as it has been done in plants before (C. Wang, Chin, Ho, Hwang, Polashock, & Martin, 1996; Yeomans et al., 2008). Some specific examples of lipids known to elicit appealing flavors are terpenes like vanillin and limonene, which are plant derived terpenes. Both of these molecules have been extensively used in the food industry as additives due to their pleasant aromas, and both have been successfully produced in microorganisms by genetically engineering *E. coli* to produce the enzymes that

lead to the desired terpenes (Jongedijk, Cankar, Buchhaupt, Schrader, Bouwmeester, & Beekwilder, 2016; Ni, Tao, Du, & Xu, 2015).

Another important part of achieving desirable organoleptic properties in algae as a food product, is the way it is processed and cooked. Important molecules that contribute to bad flavor could be removed in the processing, or other components could be added to improve organoleptic properties or at least mask the unappealing aromas (F. Gibbs, 1999). The cooking process is essential to achieving the desired end flavor, as the Maillard reaction needs high heat and the presence of amino acids and reducing sugars to yield the typical browning components found in cooked proteins. This can be observed in meat, as it only acquires its characteristic flavor after cooking, being quite insipid with only a blood-like taste when it is consumed raw (Calkins & Hodgen, 2007).

Conclusion

Algae show great potential to become a new crop that could significantly impact the worlds need for food and feed in the 21st Century. However, for that become a reality, algae need to undergo a series of improvements to enhance growth yields, nutritional quality, organoleptic traits, and perhaps most importantly, social acceptance of algae as food. A key to social acceptance may lie in the appeal of algae in terms of organoleptic traits, which may be essential to getting people to consume algae products, while at the same time they become aware of the documented healthy benefits of algae as food. The productivity and nutritional content of algae can be easily improved, and there are already many wild type algae strains that have desirable nutritional properties in terms of molecular composition. What remains a challenge is making algae that people want to eat, while at the same time improving the production processes so that algae can be economically viable as a commodity food and feed product.

Chapter 2, in full, is a reprint of the material as it appears in Biotechnology Advances, 2020. Torres-Tiji, Yasin; Fields, Francis J; Mayfield, Stephen P. The dissertation author was the primary investigator of this paper.

Figures

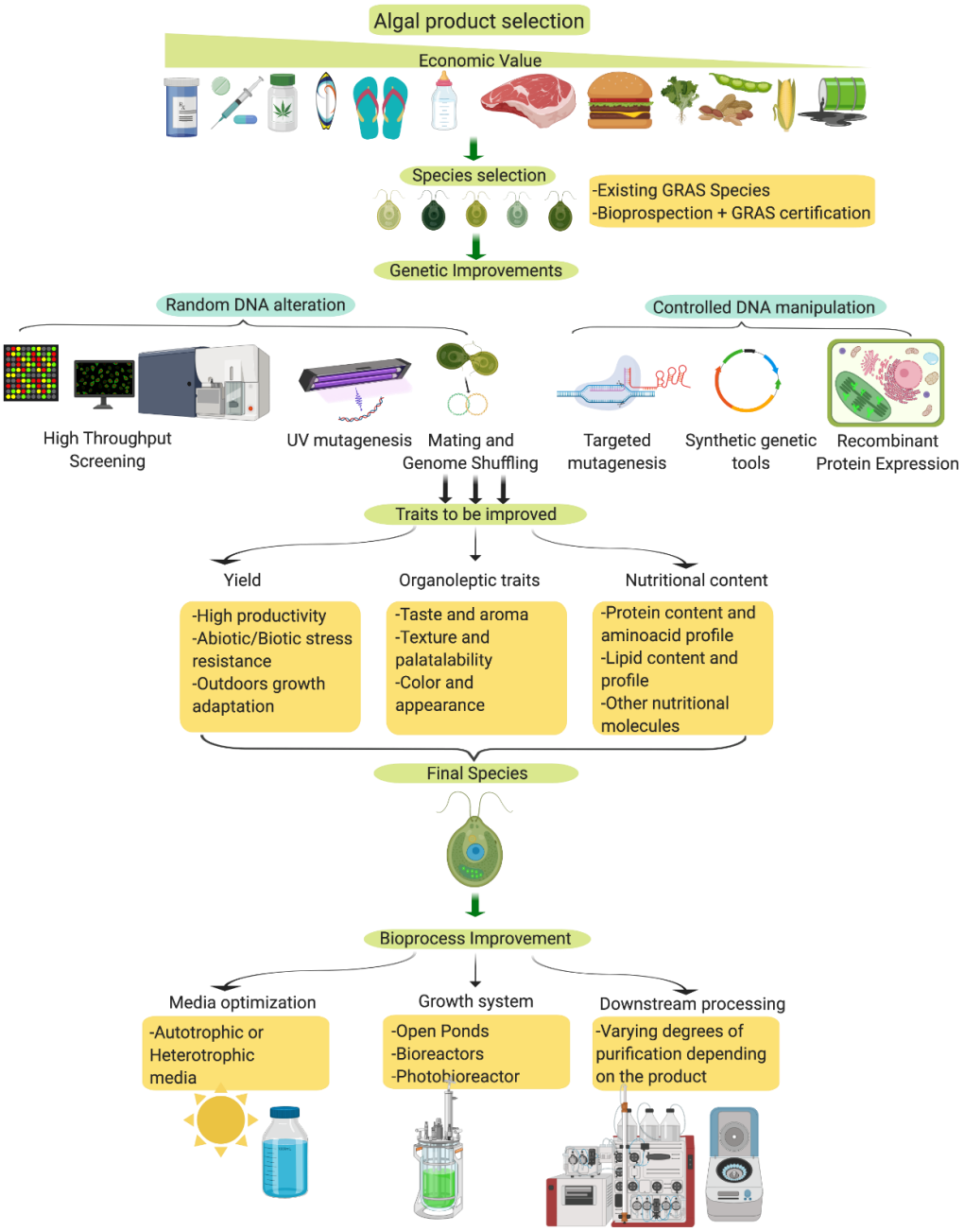

Figure 2.1. Outline of the procedure to design an algae-based bioprocess. The first step is to choose the product of interest, which in the case analyzed in this review is food, and then selecting a species that fits best the conditions required. The next step is to genetically improve the algal strain to optimize a set of traits of upmost importance, being those yields, organoleptic traits and nutritional content.
Finally, once the production strain has been developed, the bioprocess needs to be optimized, being the key features to be improved: media optimization, growth system and downstream processing.
This figure was created using Biorender.com.

Tables

Table 2.1. Estimates of global production of microalgae as of 2006. Adapted from (Spolaore et al., 2006).

References: (Hejazi and Wijffels, 2004; Lorenz and Cysewski, 2000; Pulz and Gross, 2004; Ratledge, 2004)

Microalgae	**Annual Production**	**Producer Country**	**Application & Product**
Spirulina sp.	3000 tons dry weight	USA, China, India, Myanmar, Japan	Human Nutrition Animal Nutrition Cosmetics Phycobiliproteins
Chlorella sp.	2000 tons dry weight	Taiwan, Germany, Japan	Human Nutrition Cosmetics Aquaculture
Dunaliella salina	1200 tons dry weight	Australia, Israel, USA, Japan	Human Nutrition Cosmetics B-carotene
Aphanizomenon flos-aquae	500 tons dry weight	USA	Human Nutrition
Haematococcus pluvialis	300 tons dry weight	USA, India, Israel	Aquaculture Astaxanthin
Crypthecodinium cohnii	240 tons DHA oil	USA	DHA oil
Shizochytrium	10 tons DHA oil	USA	DHA oil

Table 2.2. Estimates of global algae product production as of 2014 – Adapted from (Koller et al., 2014)

Product	Approximate market (USD/Kg)	Global market volume (USD)	References
Microalgal biomass for human nutrition	40–50	$1.25 \cdot 10^9$	(Pulz and Gross, 2004), (Spolaore et al., 2006)
Microalgal biomass for animal feed	10	$4 \cdot 10^9$	(Spolaore et al., 2006)
Microalgal nutraceuticals for human nutrition	120	$7 \cdot 10^7$	(Wijffels, 2008)
Biodiesel	0.5 (*general market price for biodiesel*)	$1 \cdot 10^9$	(Haas et al., 2006), (Sapci and Morken, 2014), (Sun et al., 2011)
	3–4 (*production price from algal origin; strongly fluctuating estimations!*)		
ß-Carotene	300–3000	$2 \cdot 10^8$	(Ben-Amotz, 2004), (Spolaore et al., 2006)
Astaxanthin	> 2000	$2 \cdot 10^8$	(Lorenz and Cysewski, 2000)
Phycobiliproteins	3000–25,000	$5 \cdot 10^7$	(Spolaore et al., 2006)
ß-1,3-Glucan	5–20	$1 \cdot 10^8$(USA)	(Spolaore et al., 2006)
Docosahexaenoic acid (DHA)	50	$4 \cdot 10^8$(USA)	(Brennan and Owende, 2010)
Eicosapentaenoic acid (EPA)	4600 (monoseptic cultivation of *Phaeodactylum tricornutum*, highly pure product)	$1.25 \cdot 10^3$(Japan)	(Belarbi et al., 2000), (Molina Grima et al., 2003)
	650 (*from fish oil*)		

Table 2.3. Characterization of the microalgae recognized by the US Food & Drug Administration as Generally Recognized As Safe (GRAS). Carbohydrate, protein, and lipid content are reported as percent of dry biomass. The availability of a sequenced genome, as well as which has been sequenced if there are more than one present in the cell. Recombinant gene expression assessed and reported for each of the genomes in the cell that show recombinant gene expression (N = nuclear genome, M = mitochondrial genome, C = chloroplastic genome)

Species	**FDA GRN No.**	**Products**	**Photoautotrophic**			**Heterotrophic**			**Genome**	**Recombinant gene expression**	**References**
			Carbs.	**Protein**	**Lipid**	**Carb.**	**Protein**	**Lipid**			
Arthospira platensis	417	Pigments, biomass	8-14%	46-68%	4-14%	n.a.	n.a.	n.a.	+	+	(Babadzhanov et al. 2004); (Becker, 2007); (Griffiths and Harrison, 2009)
Chlamydomonas reinhardtii	773	Recombinant proteins, biomass	17-50%	26%	19-21%	45%	22-48%	29%	N, C, M	N, C, M	(Becker, 2007); (Boyle and Morgan, 2009); (Griffiths and Harrison, 2009)
Auxenochlorella protothecoides	519	Biomass	11%	53%	13-23%	15-25%	10-48%	15-70%	-	-	(Feng et al., 2014); (Griffiths and Harrison, 2009); (Szabo et al., 2013); (Miao and Wu, 2004)
Chlorella vulgaris	396	Biomass	24-26%	24-44%	22-46%	23-44%	30-45%	22-36%	C	N	(Griffiths and Harrison, 2009); (Guccione et al., 2014); (Liang et al., 2009)
Dunaliella bardawil	351	Beta-carotene, biomass	11%	29%	10-19%	n.a.	n.a.	n.a.	M	N	(Ben-Amotz et al., 1982); (Griffiths and Harrison, 2009)
Euglena gracilis	697	Paramylon, biomass	30%	30-47%	20-35%	60-90%	25%	10%	C	C	(Barsanti et al., 2001); (Chae et al., 2006); (Griffiths and Harrison, 2009)

Table 2.4. Amino acid profile as percent of total protein. **= Essential amino acid, *=conditionally essential aminoacid, n.d.= not determined. Glutamine* and Asparagine data not available.

References: (Becker, 2007; Brown and Jeffrey, 1992; Kott, 1983; Nicolás Carcelén et al., 2017)

Amino Acid	*A. platensis*	*C. reinhardtii*	*A. protothecoides*	*C. vulgaris*	*D. bardawil*	*E. gracilis*
Alanine	*9.5*	*8.8*	*6.2*	*7.9*	*7.3*	*15.8*
Arginine*	7.3	7.2	13.4	6.4	7.3	3.4
Aspartic Acid	11.8	9.7	7.1	9	10.4	7.1
Cysteine*	0.9	n.d.	1.6	1.4	1.2	0.2
Glutamic Acid	10.3	11.3	10.3	11.6	12.7	9.5
Glycine*	5.7	5.7	5.5	5.8	5.5	7
Histidine**	2.2	2.3	3	2	1.8	2.2
Isoleucine**	6.7	4.4	3.7	3.8	4.2	0.2
Leucine**	9.8	9.8	5.6	8.8	11	3.7
Lysine**	4.8	6.6	4.9	8.4	7	4.9
Methionine**	2.5	2.7	2.1	2.2	2.3	0
Phenylalanine**	5.3	5.6	5.5	5	5.8	0.9
Proline*	4.2	5.6	5.6	4.8	3.3	0
Serine	5.1	4.3	5.1	4.1	5.4	10.6
Threonine**	6.2	5.1	4.9	4.8	5.4	4.5
Tryptophan**	0.3	2.8	0.5	2.1	0.7	1.7

Table 2.4. continued. Amino acid profile as percent of total protein. **= Essential amino acid, *=conditionally essential aminoacid, n.d.= not determined. Glutamine* and Asparagine data not available.

References: (Becker, 2007; Brown and Jeffrey, 1992; Kott, 1983; Nicolás Carcelén et al., 2017)

Amino Acid	*A. platensis*	*C. reinhardtii*	*A. protothecoides*	*C. vulgaris*	*D. bardawil*	*E. gracilis*
Tyrosine*	5.3	4.3	4.7	3.4	3.7	0.7
Valine**	7.1	6.5	5.2	5.5	5.8	8

Chapter 3: Strong transactivator GAL4 from Saccharomyces cerevisiae boosts recombinant protein expression in the green alga Chlamydomonas reinhardtii

Abstract

The GAL4/UAS system is a genetic tool used to increase transcription rates of specific target genes. This tool originates in the yeast Saccharomyces cerevisiae, but it has proven effective in various organisms ranging from other yeast, zebrafish, fruit flies and even tobacco plants. The green alga Chlamydomonas reinhardtii has been used to produce several recombinant proteins, but none at commercial levels. Due to its use as model organism, a variety of genetic tools exist for its use as biotechnological host, but none are as strong as the GAL4/UAS system has shown to be in other organisms. In this work we analyze the viability of the GAL4/UAS in C. reinhardtii. Moreover, both genes (GAL4 and reporter gene) are encoded in a single plasmid, which lets us assess transcription factor – target DNA sequence interactions and their consequences *in vivo* with a single transformation.

Introduction

Microalgae are a vastly biodiverse polyphyletic group comprising photosynthetic eukaryotic microorganisms. As such, they are responsible for the production of much of the oxygen in the atmosphere and hold extreme importance in the trophic chain as the main producers of energy and biomass (Chapman, 2013). They are generally capable of fast and inexpensive growth, producing high protein and lipid biomass leading to a growing interest in utilization of microalgae as food and fuel source, which is driving an increasing microalgal biotechnology, especially to produce omega 3 fatty acids, beta-carotenes and other nutritional supplements naturally present in microalgae (Demirbas, 2009; Lardon, Hélias, Sialve, Steyer, & Bernard, 2009; B. Singh, Guldhe, Rawat, & Bux, 2014; Spolaore et al., 2006). *Chlamydomonas reinhardtii* is a green alga from the family Chlorophyta that has been used a model organism for decades due to its ability to consume reduced carbon in the dark, allowing research to be done with photosynthetic mutants. Extensive research has been performed using this alga in the fields of photosynthesis, flagellar motility, cell cycle and algal genetics among others (Harris, 2001). As a result, many genetic tools are available for recombinant gene expression in *C. reinhardtii*, having successfully expressed several human recombinant proteins.

C. reinhardtii is capable of recombinant gene expression in all three genomes present in the cell (nuclear, chloroplast and mitochondrial) (Bonnefoy, Remacle, & Fox, 2007; Kindle, 1990; Kindle, Richards, & Stern, 1991). Traditionally, research has been focused on recombinant protein expression in the chloroplast due to its simpler genetics and its ability to integrate heterologous DNA in a targeted manner through homologous recombination. In contrast, nuclear gene expression is complex and heterologous genes are inserted through random integration(S P Mayfield & Kindle, 1990). Additionally, its strong codon bias, its resilient gene silencing and the lack of strong promoters and/or enhancers hinder heterologous gene expression (Schroda, 2006). There has been great progress towards the development of synthetic genes with customized codon bias and genetic tools that circumvent gene silencing, but there is still room for improvement if *C. reinhardtii* is to be competitive with other recombinant protein production hosts (Schroda, 2019).

A key element in the genetic toolkit necessary to achieve successful recombinant protein production is the promoter. A series of robust constitutive endogenous promoters have been used in *C. reinhardtii* to drive heterologous gene expression in the nucleus, including but not limited to: PSAD, RBCS2, β2TUB. Additionally, there has been some success in using endogenous inducible promoters like CAH1, CAH4, NIT1, HSP70A, FEA1, CYC6 and METE. The most used promoter to date in recombinant gene expression in *C. reinhardtii* is the chimeric promoter AR1, composed of the HSP70A promoter fused upstream of the RBCS2 promoter (Schroda, 2019). One synthetic promoter generated based on the cis-motifs elements, structure and sequence of the promoters pertaining to the top expressed genes in *C. reinhardtii*, named SAP11 (synthetic algal promoter 11), was shown to drive a higher gene expression than AR1 (Scranton et al., 2016). However, there have not been reported any attempts at expressing a heterologous transcriptional activator driving the expression of a recombinant gene of interest.

Strong transcriptional activators have been used to boost recombinant gene expression in a variety of hosts. These systems are characterized by the expression of a transcription factor that binds and upregulates transcription of the gene of interest driven by the target promoter (Rantasalo, Landowski, Kuivanen, Korppoo, Reuter, Koivistoinen, Valkonen, Penttilä, Jäntti, & Mojzita, 2018). An example is the lac operon from Escherichia coli, in which the activator catabolite activator protein (CAP) binds to the lac promoter and strongly activates its transcription as long as the lac repressor is not also bound to the promoter (Browning, Godfrey, Richards, Robinson, & Busby, 2019). Another example is the strong activator from herpes simplex virus (HSV) VP16 that forms a transcriptional regulatory complex with cellular factors and increases the transcription rate of the viral genes to very high levels (Sadowski, Ma, Triezenberg, & Ptashne, 1988). Gal4 is a strong transcriptional activator belonging to the GAL operon from Saccharomyces cerevisiae. This operon expression is driven by the transcriptional activation of Gal4, which is repressed by default through the binding of Gal80. In the presence of galactose, Gal3 is activated and binds Gal80 thus releasing Gal4, now free to activate transcription (Elliott & Brand, 2008).

Gal4 is known to bind a specific DNA sequence nearby target promoters (e.g. GAL1, GAL10, GAL7), thus recruiting transcription factors near the promoter and in this way activate transcription. This

sequence is called Upstream Activating Sequence and its consensus sequence is 5'-CGG-N_{11}-CCG-3' (where N is any nucleotide) (Elliott & Brand, 2008). Upon binding to this sequence, Gal4 recruits mediators of transcription like Gal11, Tra1 and TAF9 among others (Reeves & Hahn, 2005). This system, called Gal4/UAS system, has been used in many different organisms to drive recombinant gene expression, including but not limited to *Saccharomyces cerevisiae, other yeast, Drosophila melanogaster, Danio rerio, Xenopus tropicalis, Caenorhabditis elegans, Mus musculus*, etc. However, this system has not been tested in any algae to date. *Chlamydomonas reinhardtii* possesses homologs for some of the most relevant mediators of transcription involved in Gal4 transcriptional activation, so it could be possible that the Gal4/UAS system increased gene expression if the UAS sequence was placed upstream of a robust promoter (data not shown).

In this work we decided to clone the Gal4 gene driven by the promoter AR1 and to clone the UAS sequence upstream of another AR1 promoter driving the reporter gene DsRed. In order to simplify the generation on this recombinant algal line, we generate a single plasmid containing both the AR1-Gal4 ORF and the UAS-AR1-DsRed ORF.

Methods

Plasmid design, codon optimization and cloning

The control expression vector pAR1 was designed based on the sequences from pBR9 (Beth A. Rasala, Barrera, Ng, Plucinak, Rosenberg, Weeks, Oyler, Peterson, Haerizadeh, & Mayfield, 2013), pOpt (Lauersen, Kruse, & Mussgnug, 2015b) and pRMC (Berndt et al., 2021). A codon optimized DsRed was cloned between the ble gene and the rbcs2 3' UTR of the pAR1 vector. A codon optimized bovine myoglobin gene was cloned between the end of the DsRed gene, with a FMDV-2A ribosomal-skip motif in between, and the rbcs2 3' UTR. The pGAL4^{+} vector was generated by first making an intermediate vector (AR1-GAL4) by cloning a codon optimized GAL4 gene into pAR1 between the end of a *ble* gene followed by a FMDV-2A sequence, and the rbcs2 3' UTR. A second intermediary vector (5XUAS-AR1) was cloned

by inserting a 5XUAS sequence (GenBank: MN517552.1) immediately upstream of the AR1 promoter in pAR1. The insertion was performed using a Q5® Site-Directed Mutagenesis Kit (New England Biolabs, Ipswich, MA, USA, catalog#E0554S) and a 153 bp ultramer containing the 5XUAS sequence. pGAL4^{+} vector was assembled in a 3 fragment USER (New England Biolabs, Ipswich, MA, USA, catalog#M5505S) cloning reaction by assembling the part that contains the AR1 promoter until the 3' UTR from AR1-GAL4, the part that contains the 5XUAS until the AR1 promoter from 5XUAS-AR1 intermediary vector, and the part of pAR1 driving the expression of ble-DsRed-2A-myoglobin including the vector backbone. All the cloning, unless otherwise specified, was made using PCR-based amplification to add overlapping regions to the different fragments and use Gibson-style assembly methods (HiFi Assembly Kit, New England Biolabs, Ipswich, MA, USA). The pGAL4^{-} vector was generated by deleting the GAL4 gene, but leaving the ble gene intact, using a Q5® Site-Directed Mutagenesis Kit.

All genes were codon optimized following the procedures described by (Berndt et al., 2021). DsRed was codon optimized based on the amino acid sequence of DsRed T4 (Bevis & Glick, 2002). Bovine myoglobin was codon optimized based on the complete amino acid sequence found in the database Uniprot (Uniprot ID: P02192). Gal4 was codon optimized based on the complete amino acid sequence found in the database Uniprot (Uniprot ID: P04386). The codon optimized genes were ordered as gBlocks™ from Integrated DNA Technologies (Coralville, IA, USA). The GAL4 gene had to be split in two gBlocks due to its length.

Algal strain, growth conditions and transformation

The algal strain utilized in this study is a *Chlamydomonas reinhardtii* strain that was adapted to high nutrient concentration media in a bioreactor, called CR25. All the strains used were grown in TAP media following the methods described in (Beth A. Rasala, Lee, Shen, Briggs, Mendez, & Mayfield, 2012). CR25 was transformed following a modified electroporation protocol reported in (Beth A. Rasala et al., 2012). The only differences are the amount of linearized plasmid per transformation (3 µg instead of 1 µg), and the electroporator settings were 1250 V/cm, capacitance 50 µF and 800 Ω. After transformation, the product of each transformation was plated on five TAP/agar plates containing 15 µg/ml of Zeocin

(TAP-Zeo plates) (Thermo Fisher Scientific, Waltham, MA, catalog #R25001) to select for positive transformants (Beth A. Rasala et al., 2012).

Screening of DsRed expression through Fluorescent Activated Cell Sorting

Hundreds of transformants per plasmid were obtained and pooled into liquid TAP medium for two days. After that, the cells were diluted 1:2 into a new liquid TAP medium and allowed to grow for 24 hours. Then, 5 ml of cells were passed through a 40 µm cell strainer and analyzed through Fluorescence Activated Cell Sorting at the sorting facility at The Scripps Research Institute. The equipment used was MoFlo Astrios EQ (A) and the data was analyzed using FlowJo v.10.4. The population was gated using the strategy described in (Scranton et al., 2016). The DsRed fluorescence was then measured with a 561 nm yellow/green laser with a 579/16 nm filter set. At least 200,000 individual cells were analyzed per sample.

Fluorescence microplate reader assay

The fluorescence of cells was measured using a plate reader as described in (Beth A. Rasala et al., 2013). The excitation and emission wavelengths used to measure DsRed fluorescence were (excitation 555/9 nm, emission 585/20 nm). The best gain was determined using the optimization algorithm from the plate reader software, and that was a gain setting of 235. That gain was maintained throughout all the measurements. Measurements were normalized for chlorophyll fluorescence or absorbance at 750 nm (Optical Density 750nm). Chlorophyll measurement settings were (excitation 440/9 nm, emission 680/20 nm), and gain manually set at 160. To calculate the normalized DsRed fluorescence value we divided the DsRed fluorescence over the chlorophyll fluorescence value and multiplied that by 100. If the normalization was performed on Optical Density at 750nm, the normalization consisted on dividing the DsRed fluorescence value over the absorbance value at 750 nm. Statistical significance of the differences observed were calculated using a Brown-Forsythe and Welch ANOVA tests, performed using Prism 9 (GraphPad, San Diego, CA, USA).

Bradford and Western Blot

Samples were harvested through centrifugation, supernatant discarded, and cells were lysed by thoroughly resuspending in lysis buffer with a micropipette. Lysis buffer was made using BugBuster 10X (MilliporeSigma, Burlington, MA, USA, catalog #70921) diluted to 1X concentration in 1X TBS buffer (50 mM Tris-HCl, NaCl 150 mM, pH 7.5). Samples were then centrifuged at 10000G for 5 minutes and supernatants were used for protein analysis.

Total protein concentration determined using the Pierce Coomassie Protein Assay Kit following the manufacturer's protocol (Thermo Fisher Scientific, Waltham, MA, USA, catalog #23200). The assays were performed in Corning® 96-well plates (MilliporeSigma, Burlington, MA, USA, catalog #CLS9018BC), and measured using an Infinite® M200 PRO plate reader (Tecan, Männedorf, Switzerland).

Western blots were performed as described in the methods section of Chapter 4. The primary antibody used to measure DsRed was Living Colors® DsRed Monoclonal Antibody (Takara Bio USA, Mountain View, CA, USA, catalog# 632392) with a 1:2000 dilution in blocking solution. The secondary antibody used was anti-mouse IgG alkaline phosphatase conjugated (Sigma-Aldrich, St. Louis, MO, USA, catalog #A1682) with a 1:10000 dilution in blocking solution.

RT-qPCR

The total RNA from the cell was extracted using a modified version of the TRIzol reagent protocol. The composition of the lysis buffer was: Tris-HCl 50 mM pH 8.0, 200 mM NaCl, 20 mM EDTA, 2% SDS and Proteinase K at 20mg/mL. 1 mL of cells grown in liquid TAP cultures as described above were harvested at early-log growth phase ($1 \cdot 10^6$ cells/mL) in a fixed angle centrifuge chilled to 4°C at 3000G for 2 minutes. Cells were resuspended in 0.25 mL of lysis buffer. Lysing cells were incubated at 70°C for 2 minutes. 2 mL of TRIzol™ reagent (Thermo Fisher Scientific, Waltham, MA, USA, catalog

#15596026), incubated at room temperature for 5 minutes, 0.5 mL of chloroform was added and shaken vigorously for 15 seconds. Samples were incubated at room temperature for 5 minutes and then were centrifuged at 12000G for 15 minutes at 4°C. Aqueous phase was then transferred to a new tube and 1 mL of isopropanol was added. Mixture was incubated at 4°C for 10 minutes and then centrifuged at 12000G for 10 minutes at 4°C. The supernatant was discarded, and the pellets were resuspended in 2 mL of 75% ethanol. Samples were vortexed briefly and centrifuged at 7500G for 5 minutes at 4°C. The supernatant was discarded with a pipette and the pellet air dried for 10 minutes. Then pellets were resuspended with 25 µL of RNAse-free water and 3.5 µL of 10X DNAse buffer and 2 µL of DNAse (Turbo DNA-free™ Kit, Thermo Fisher Scientific, Waltham, MA, USA, catalog #AM1907).

The quality of the RNA was measured using a DU®730 Life Science UV/Vis Spectrophotometer (Beckman Coulter, Brea, CA, USA) and a NannoVette Microliter Cell (Beckman Coulter, Brea, CA, USA). RNA was quantified measuring the absorbance at 260 nm, and purity was assessed by the ratio of A_{260}/A_{280}. Reverse transcription was performed by using the SuperScript™ II Reverse Transcriptase (Invitrogen, Waltham, MA, USA, catalog #18064022) following manufacturer's protocol. Each reaction was run using 200 ng of RNA as template. Quantitative-PCR was performed targeting the gene *DsRed* and the endogenous control gene *lck1*. The reagent used was Power SYBR™ Green PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA, catalog #4367659). Reactions were run using 1 µL of cDNA template, duplicates using a 1:1 dilution of cDNA and duplicates using 1:10 dilution of cDNA, totaling 4 technical replicates per sample and target. Primers were used at a final concentration of 200 nM. Samples were analyzed in an Applied Biosystems StepOnePlus™ Real-Time PCR systems, using the default SYBR® Green Reagent and Comparative C_T. Relative mRNA levels were calculated using the $2^{-\Delta\Delta CT}$ method (Livak & Schmittgen, 2001).

Results

Vector design, assembly and transformation

The basis for the design of the overexpression vector is the vectors pBR9 (Beth A. Rasala et al., 2012) and pOpt (Lauersen et al., 2015b) in the manner described in (Berndt et al., 2021). The basal vector upon which improvements were made is called pAR1 and it features the chimeric promoter AR1 and the *rbcs2* 5' UTR driving the expression of the antibiotic resistance gene *ble* followed by the fluorescence reporter gene *DsRed* with intercalating introns to boost expression (Baier, Jacobebbinghaus, Einhaus, Lauersen, & Kruse, 2020; Baier, Wichmann, Kruse, & Lauersen, 2018; Lumbreras, Stevens, & Purton, 1998). This first protein is then separated by a foot and mouth disease 2A (FMDV-2A) ribosomal-skip motif placed downstream from the last codon of the DsRed coding sequence, and a second protein of interest follows. In this specific case the protein of interest is a chloroplast targeted bovine myoglobin which was encoded using the chloroplast targeting sequence from the native *psaE* gene and a codon optimized myoglobin gene (Uniprot ID: P02192). A FLAG® tag was added at the C-terminus of the myoglobin protein, which was then followed by a stop codon and 3' UTR *rbcs2* region. Finally, a second cistron immediately downstream from the first one, composed of the *beta-2 tubulin* promoter and 5' UTR followed by the APHVIII CDS and the *beta-2 tubulin* 3'UTR, confers antibiotic resistance to hygromycin and ensures intact delivery of the genetic payload.

The GAL4 CDS was codon optimized based on its amino acid sequence (Uniprot ID: P04386) and the codon usage in the nuclear genome of *C. reinhardtii*. It was cloned into a modified pAR1 vector (pAR1-GAL4) containing the AR1 promoter, *rbcs2* 5'UTR, *ble* gene with intercalating introns, a SV40 nuclear localization signal (Beth A. Rasala et al., 2014), FLAG® tag, GAL4 codon optimized CDS and *rbcs2* 3'UTR. For the UAS region, an ultramer containing 5 tandem repeats of the UAS sequence with an "AG" spacer nucleotide sequence between repeats was cloned immediately upstream of the AR1 promoter in pAR1 (pAR1-5XUAS). Then, through USER cloning (Nour-Eldin, Hansen, Nørholm, Jensen, & Halkier, 2006) the final vector was assembled by PCR amplification of fragment 1 from pAR1-GAL4 spanning from the AR1 promoter to the 3' UTR, PCR amplification of fragment 2 from pAR1-5XUAS spanning from the beginning of the 5XUAS sequence to the end of the AR1 promoter, and PCR amplification of fragment 3 from pAR1 spanning from the 5' end of the AR1 promoter to the end of the vector backbone. The resulting vector, pGAL4^{+}, contains in this order a AR1 promoter, *rbcs2* 5' UTR, *ble* gene with introns, FMDV-2A ribosomal skip-motif, SV40 NLS, FLAG® tag, GAL4 CDS, *rbcs2* 3' UTR,

5xUAS, AR1 promoter, *ble* and *DsRed* genes with introns, FMDV-2A ribosomal skip-motif, *psaE* chloroplast targeting sequence, myoglobin CDS, FLAG® tag, *rbcs2* 3' UTR, *beta-2 tubulin* promoter and 5' UTR, APHVIII CDS and *beta-2 tubulin* 3' UTR.

It is well known that a multiplicity of promoters and other genetic elements in the vicinity of a gene can alter their expression (Li, Wang, Geng, Li, Wang, Liang, & Qi, 2012; Wei, Xiang, & Tan, 2002). To ensure that the change in gene expression in pGAL4^{+} compared to pAR1 is due to GAL4 binding to the 5xUAS-AR1 promoter and not due to the fact that there are two AR1 promoters in pGAL4^{+}, we generated a third plasmid (pGAL4^{-}) by deleting the GAL4 gene from pGAL4^{+}, using a Q5 Site-Directed Mutagenesis kit, but leaving an intact AR1 promoter driving the expression of the *ble* gene.

The vectors were then miniprepped, digested with KpnI and XbaI to separate the payload from the backbone, purified with a DNA clean up kit (Promega, Madison, CI, catalog # A9282). *Chlamydomonas reinhardtii* was transformed with 3 µg of each of the purified plasmids (pAR1, pGAL4^{+}, pGAL4^{-}) and plated on TAP agar plates containing 15 µg/mL of Zeocin.

Analysis of gene expression through Fluorescent Activated Cell Sorting

The obtained transformants for each plasmid were pooled into 50 mL liquid TAP cultures and were analyzed through Fluorescence Activated Cell Sorting. Chlorophyll autofluorescence (excitation 488 nm, emission 710 nm) was employed as a measurement of cell size and DsRed fluorescence (excitation 561 nm, emission 579 nm) was used as a direct measurement of the recombinant gene expression level for each cell harboring one of the three plasmids. The top 5% individual cells showing the highest DsRed fluorescence were sorted into a 50 mL conical tube containing 5 mL of liquid TAP for further analysis.

From this FACS analysis we can see that the distribution of DsRed fluorescence among the population of cells harboring plasmids pAR1, pGAL4^{+} or pGAL4^{-} differs greatly (Figure 3.2). The median and standard deviation of DsRed fluorescence for each sample respectively was: negative control (*C. reinhardtii* CR25 wild-type) 41.4 and 29.3 relative fluorescent units (RFU), AR1 162 and 509 RFU, GAL4^{+} 867 and 2922 RFU, GAL4^{-} 66.5 and 907 RFU. We can see that GAL4^{+} shows much higher DsRed

fluorescence overall as a population, however there is great variation in DsRed fluorescence among individuals in a population. Said strong variation is to be expected since the strength of recombinant gene expression from a specific expression vector depends largely on the genomic loci in which it integrated, and *C. reinhardtii* integrates heterologous DNA mostly through random integration (Gumpel, Rochaix, & Purton, 1994).

Since the goal of this project was to design expression vectors that allowed the highest recombinant protein expression possible, we measured and compared the median DsRed fluorescence of the population of cells belonging to the top 5% DsRed fluorescence (Figure 3.2). As we can see, the control AR1 shows 8.22-fold higher DsRed fluorescence than the background fluorescence measured using wild-type. The construct GAL4$^-$ shows a 3.26-fold improvement over the control AR1 population even though the GAL4 transactivator is deleted. This can be due to having an upstream cistron driven by the constitutive promoter AR1 that could be helping recruit transcription factors in the downstream AR1 promoter. The highest expression is observed in the GAL4$^+$ population, boasting a 9.44-fold improvement over the AR1 control and a 2.89-fold improvement over the GAL4$^-$ population.

Down selection of top expressors and cross comparison using Western Blot

The top 5% DsRed expressing cells that were sorted for each plasmid were then plated on Tap-Zeo plates. Then we picked 96 individual colonies and those were transferred to 96-well plates containing 200 µL of liquid TAP per well. After 4 days, the DsRed fluorescence was measured using a plate reader and those values were normalized against chlorophyll fluorescence measured with the same instrument. As we can see in Figure 3.3, GAL4$^+$ shows the highest average value being 76.78 normalized RFU, whereas AR1 shows 58.65 normalized RFU and GAL4$^-$ shows 65.04 normalized RFU. A Brown-Forsythe and Welch ANOVA test was performed to assess whether the means of the different populations are significantly different. Based on the results (Figure 3.3), AR1 is significantly different from GAL4$^+$ (adjusted p-value <0.0001), but not from GAL4$^-$ (adjusted p-value 0.4301). GAL4$^+$ is significantly different from GAL4$^-$ (adjusted p-value 0.0233). Even though GAL4$^+$ shows a higher fluorescence than AR1, the difference is not nearly as drastic when compared to the data obtained through FACS.

From the 96 clones for each plasmid, we down-selected the top expressor for AR1, GAL4^{+} and GAL4^{-}. Cells were then grown on 50 mL liquid TAP cultures and after 4 days, 1 mL of culture was collected for western blot analysis. The cells were separated from the supernatant through centrifugation, lysed with lysis buffer and total protein was quantified using a Bradford assay. An anti-DsRed western blot was performed loading equal amounts of total protein, 10 µg per lane (Figure 3.4). The five clones obtained from GAL4^{+} showed much higher DsRed protein accumulation than AR1 and GAL4^{-}, which showed similar DsRed accumulation. As quantified using ImageStudio Lite, we determined that GAL4^{+} shows 11-fold higher DsRed accumulation than AR1 and 7.6-fold higher than GAL4^{-}.

mRNA fold change as result of Gal4 expression

Since it seems that GAL4 does indeed boost expression of recombinant DsRed when placed under the 5XUAS-AR1 promoter, we analyzed the levels of DsRed mRNA in the cell for each clone through Rt-qPCR. Cells were grown in liquid TAP cultures and harvested at early-log phase ($1 \cdot 10^{6}$ cells/mL). The total RNA from the cells was extracted and measured for purity. All the samples had a ratio of A_{260}/A_{280} ranging from 2.096 to 2.145, attesting to their purity. The RNA was retrotranscribed into cDNA using oligo $(dT)_{18}$ and qPCR were performed on housekeeping gene *lck1* and recombinant gene DsRed. The normalized DsRed mRNA levels were compared to those in the AR1 sample, and the results indicate that GAL4^{+} shows a 11.61-fold increase over AR1 whereas GAL4^{-} show 0.62-fold DsRed mRNA levels despite showing higher recombinant protein.

Discussion

The results presented in this work show a heterologous transcription factor from a very distantly related organism like *Saccharomyces cerevisiae* being cloned into a green alga resulting in the increase of both mRNA and protein levels of a recombinant reporter gene. The reporter gene in question was a codon optimized DsRed fused to the antibiotic resistance gene *ble* and this fusion is then directed to the nucleus due to ble's natural nuclear localization (Beth A. Rasala et al., 2012). The promoter AR1 has been used extensively in *C. reinhardtii* and is the arguably the most robust promoter for recombinant gene expression among the tested ones so far (Schroda, Beck, & Vallon, 2002; Schroda, Blöcker, & Beck, 2000). A study generated a suite of synthetic algal promoters among which some promoters proved to show almost twice as high recombinant protein expression as that obtained from AR1, while promoter strength was measured as the average mCherry fluorescence of the top 5% of cells analyzed in a Fluorescence Activated Cell Sorter (Scranton et al., 2016). In our study, we show that coupling the strength and robustness of the promoter AR1 with the GAL4/UAS system yields a 10-fold increase in recombinant protein expression measured as fluorescence of the top 5% of cells analyzed in a FACS, as recombinant protein accumulation measured through western blot and as mRNA levels measured through Rt-qPCR.

However, it is notable that the control vector GAL4^{-} also shows a 3-4 fold increase in recombinant protein expression as measured by the top 5% fluorescence of cells analyzed through FACS and by recombinant protein accumulation measured through western blot, but counter indicatively it shows a reduction of mRNA accumulation of 38% when compared to AR1. There are several hypotheses that could explain the discrepancy. First, the fact that increased expression is observed could be due to the fact that there are two AR1 promoters nearby, thus they could be helping recruit transcription factors to the area (Li et al., 2012; Wei et al., 2002). Second, this high expression is accompanied by a high variability, as indicated by the large error bars in figures 3.2, 3.3 and 3.4. This could be due to high expression induced by the duplicity of AR1 promoters coupled with the presence of a ble gene, a molecular tool to prevent gene silencing, could lead to gene silencing of the mRNA containing ble::DsRed in favor of the upstream mRNA that only contains ble (Beth A. Rasala et al., 2012; Schroda, 2006). In this

manner, the colonies of *C. reinhardtii* would maintain their resistance to Zeocin without having to express high levels of DsRed. Thefore, the plasmid design should be improved to avoid repeated DNA sequences and repeated antibiotic resistances. An optimized pGAL4^{+} plasmid could be used to detect transcription factor – promoter interactions and the effect of such interactions in *C. reinhardtii*. There have been efforts at characterizing transcription factor- promoter interactions in C. reinhardtii, but the use of a yeast-1-hybrid approach only sheds light on the physical protein-DNA interaction and does not inform of the consequences (Anderson et al., 2017). Additionally, one approach to yield a stable usage of GAL4/UAS system in *C. reinhardtii* would be to split the pGAL4^{+} vector into two: vector A would contain Gal4 being driven by a strong promoter with a different antibiotic and fluorescent protein as reporter gene, and vector B would contain a recombinant protein of interest being driven by a 5XUAS-AR1 promoter and using *ble* to prevent gene silencing and using a second fluorescent protein as reporter gene.

The GAL4/UAS system has being tried in many different organisms with different levels of success (Kakidani & Ptashne, 1988). In one example, it has been used in silkworms to increase recombinant protein production in a similar manner, and it has been showed that different genetic elements in combination with the GAL4/UAS system can yield vastly different results (Tatematsu, Kobayashi, Uchino, Sezutsu, Iizuka, Yonemura, & Tamura, 2010), which could be implemented in our algal system thus reducing the redundancy of elements employed. Additionally, this approach also employs a split system harboring GAL4 and 5XUAS in different plasmids. In another example, the GAL4/UAS system was used to drive GFP expression in *Arabidopsis thaliana*, in which an enhancer trap line was use to genetically manipulate lateral root development (Laplaze, Parizot, Baker, Ricaud, Martinière, Auguy, Franche, Nussaume, Bogusz, & Haseloff, 2005), albeit in this case they utilize a GAL4-VP16 fusion. Finally, the most successful GAL4 usage has been in Drosophila melanogaster, where the GAL4/UAS system has allowed a whole field of genetic manipulation to flourish (Fischer, Giniger, Maniatis, & Ptashne, 1988).

This work demonstrates that GAL4 can be used in microalgae, but optimization of the expression vectors is first needed before gene expression derived from such complicated plasmids can be reliable and robust.

Chapter 3 is being prepared for publication. Torres-Tiji, Yasin; Mayfield, Stephen P. The dissertation author was the primary investigator of this paper.

Figures

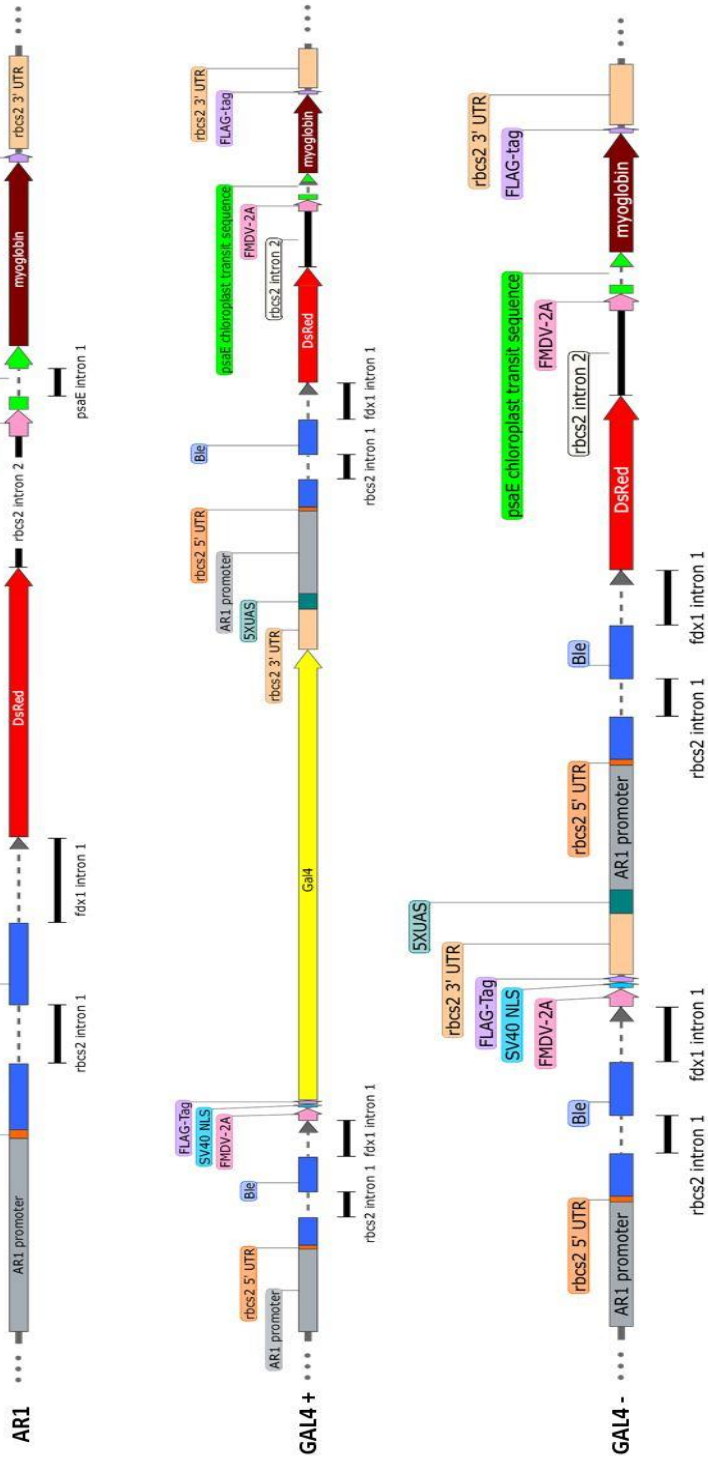

Figure 3.1. Vector map of all the expression vectors used in this study. AR1, GAL4+ and GAL4-, with all their respective elements. AR1 is the basic expression vector expressing DsRed, while also expressing the antibiotic resistance gene Ble. GAL4+ is a fusion of two AR1 vectors, the first part driving GAL4 and the second part includes a 5XUAS GAL4 binding region upstream of AR1, thus making DsRed expression responsive to GAL4. GAL4- was made by deleting GAL4 from GAL4+ and is a control for any increased expression derived from the altered vector architecture.

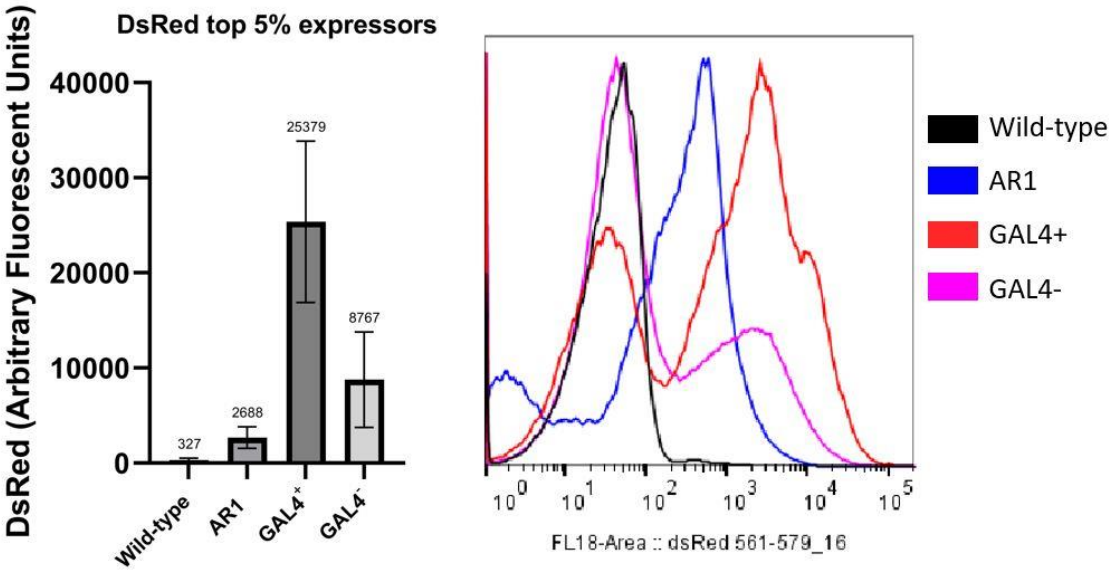

Figure 3.2**.** DsRed expression of four populations in a FACS: a negative control using *C. reinhardtii* CR25 wild-type, population of transformants using AR1 selected on TAP-Zeo plates, transformants from GAL4^{+} and transformants from GAL4^{-}. On the right panel, DsRed measured using a Fluorescence Activated Cell Sorter MoFlo Astrios EQ (A) using a 561 nm laser and a 579/16 nm filter, plotted as a histogram. On the left panel, median DsRed fluorescence of the top 5% individuals in each population with standard deviation calculated using the median. Analysis performed on FlowJo V10.

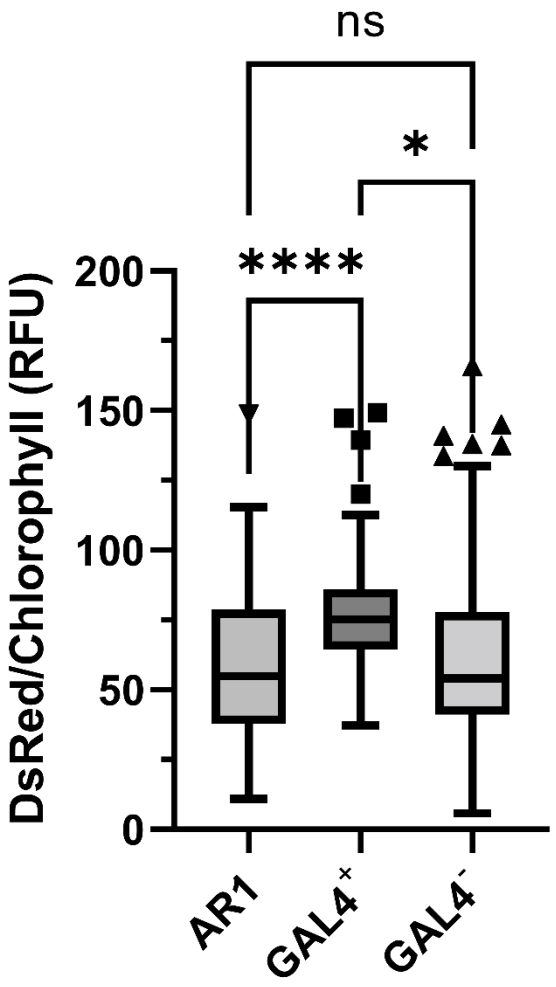

Figure 3.3**.** DsRed fluorescence of top 96 individuals from each population previously analyzed through FACS. DsRed values measured in a microplate reader with a manual gain of 235, normalized over chlorophyll fluorescence measured with the same instrument with a manual gain of 160. Data ploted as a Tukey's box plot. Statistical differences measured using Prism 9 (GraphPad), method used: Brown-Forsythe and Welch ANOVA test.

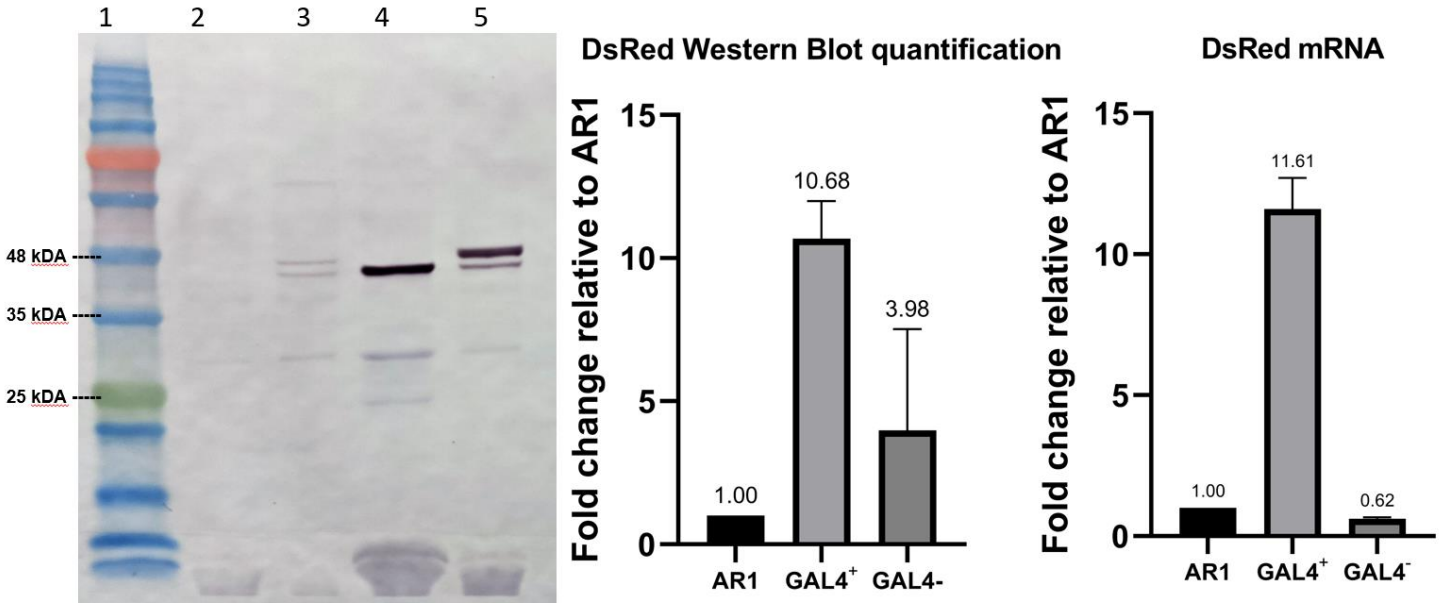

Figure 3.4. DsRed protein accumulation measure through western blot and mRNA accumulation measured through RT-qPCR. On the left panel, a western blot loaded with 20 µg of total protein as measured by Bradford assay, incubated with anti-DsRed monoclonal. Lane 1 protein ladder, lane 2 wild-type, lane 2 AR1 lysate, lane 3 GAL4+ lysate, lane 4 GAL4- lysate. Predicted protein size is 40 kDa, apparent protein size is approximately 42 kDa. Double band appears as the product of inefficient ribosomal-skipping from FMDV-2A sequence thus psaE chloroplast transit sequence is added (3.3 kDa). Mid panel shows quantification of western blot using ImageStudio Lite, normalized for AR1 intensity. Measurements done in duplicate, with different western blots performed on different days and on different algal cultures (same strains). Right panel shows DsRed mRNA quantification as measured, in quadruplicates, through Rt-qPCR. Relative mRNA levels were calculated using the $2^{-\Delta\Delta CT}$ method.

Chapter 4: *Chlamydomonas reinhardtii* grown at high cell density in bioreactor allows for high yields of secreted recombinant protein

Abstract

Microalgae have been identified as a viable resource for desirable bioproducts, such as biofuels, materials, nutritional additives, and high-value products like recombinant therapeutic proteins. Traditional biotechnological production platforms, such as *Escherichia coli*, *Saccharomyces cerevisae* and CHO cells, have dominated the market of recombinant proteins, but with novel research it is becoming more evident that microalgae can serve as an alternative. In the present study, we examine the potential of the microalga *Chlamydomonas reinhardtii* to produce a complex human recombinant protein in a fermenter. The recombinant protein, ICAM-1, was targeted for secretion to the extracellular media as the culture was grown using a fed-batch strategy to achieve high cell density. Ultimately, this resulted in biomass productivity of 40 g/L and yielded 50 mg/L of recombinant protein.

Introduction

Microalgae are a very diverse group of single celled eukaryotic organisms that can grow photosynthetically, and in some cases heterotrophically. These microorganisms are capable of rapid and inexpensive growth in outdoor ponds, as well as growth in heterotrophic bioreactors at high cell densities. They present a good balance between simplicity to be able to be cultured inexpensibly as well as high enough complexity to produce intricate bioproducts. This offers a tremendous potential for producing high value products, either recombinantly or naturally (Barrera, 2013; Hannon et al., 2010; Specht et al., 2010). *Chlamydomonas reinhardtii* is a green microalgae that has been used for decades a model organism to study a variety of biological phenomena including but not limited to: flagellar motility, photosynthesis, phototaxis, algal genetics and cell cycle. Among the reasons that this organism was adopted as model organism are its ease of genetic manipulation due to its haploid nature as well as its ability to grow heterotrophically in the dark, which allowed the production of photosynthetic mutants shedding light into its mechanism of action.

Due to all the molecular biology work in *C. reinhardtii*, multitude of genetic tools have been developed for this organism, with no other microalgae having near the amount of genetic manipulation options which ended up granting *C. reinhardtii* the nickname of “green yeast” (Rochaix, 1995). All three genomes (nuclear, chloroplastic and mitochondric) in *C. reinhardtii* have been fully sequenced and are capable of heterologous gene expression. Moreover, recombinant protein expression has been demonstrated in both the chloroplast and the nuclear genome by expressing complex mammalian proteins in multiple occasions. Recombinant expression in the chloroplast has traditionally been more successful due to its simpler genetics and the fact that this single organelle can take up to 70% of the cell’s volume. Several recombinant proteins of interest have been expressed in the chloroplast, reaching yields of up to 10.5% of total soluble protein (TSP) in the best case, but more typically ranging from 0.5 to 5% of TSP (Stephen P. Mayfield et al., 2007; Beth A Rasala, Muto, Lee, Jager, Cardoso, Behnke, Kirk, Hokanson, Crea, & Mendez, 2010). The downside of this recombinant protein production is that complex post-translational modifications are not added to the proteins, such glycosylation which can be crucial for the protein’s biological activity. In contrast, recombinant proteins expressed in the nucleus can be

targeted to any organelle, including the chloroplast, and thus receive complex post-translational modifications like glycosylation if they are targeted to the Endoplasmic Reticulum or Golgi Apparatus for later secretion (Beth A. Rasala et al., 2014). However, the yields of recombinant protein achieved through nuclear expression are typically much lower than those of chloroplast recombinant expression, approximately being 0.25% TSP. Alternatively, recombinant protein yields expressed through nuclear genome transformation are usually reported in mg/L, and those range from 0.7 to 15 mg/L (Lauersen, Huber, Wichmann, Baier, Leiter, Gaukel, Kartushin, Rattenholl, Steinweg, von Riesen, Posten, Gudermann, Lütkemeyer, Mussgnug, & Kruse, 2015a; Lauersen et al., 2015b).

Compared to other recombinant protein systems, the yields obtained using *C. reinhardtii* as host are lacking and need to be improved in order to be competitive. The yield of recombinant protein depends mainly on three factors: recombinant gene expression yields at the level of individual cells, total biomass yields in the culture, and stability of the recombinant protein overtime inside the cells or the culture media. To improve recombinant gene expression there needs to be genetic modifications made to the cells, either random or precise. Precise genetic modifications involve the design of optimized genetic tools (e.g. synthetic algal promoters, incorporation of enhancers through introns) or targeted mutagenesis of the algal genome to favor recombinant protein expression. Indirect genetic modifications involve the introduction of unknown and random genetic modifications and then screening for the desired outcome (e.g. random mutagenesis, mating and selection). These modifications can involve the downregulation of native proteases that impede the optimal accumulation of recombinant protein, or through the addition of charepones and other enzymes that facilitate the folding of recombinant proteins and impedes their aggregation and degradation. Prevention of recombinant protein degradation and/or aggregation can also be accomplished by modifying the culture growth conditions, such as lowering the temperature of the culture to prevent misfolding (Gopal & Kumar, 2013). Increasing total biomass yields is an effective way to increase recombinant protein yields as long as the fraction of biomass that corresponds to recombinant protein does not decrease proportionally to the increase of biomass achieved. This can be done through improved culture techniques, via modified medias, better outdoor pond design or intricate bioreactor design.

The simple media requirements for microalgal growth and their photosynthetic capability have allowed the growth of these microorganisms in large cheap outdoor ponds that have a big scale-up potential reaching agricultural scale in the best cases (Kumar, Mishra, Shrivastav, Park, & Yang, 2015). However, the legal limitations for genetically modified organisms to be grown outdoor, the low productivity yields and the unavoidable risk of biological contaminations have driven the growth of genetically engineered microalgae in contained systems. Such systems are comprised of inexpensive large hanging bags that grow algae with minimal cost harvesting the Sun's energy, intricated photobioreactors that are intensively and artificially illuminated to increase biomass yields, and heterotrophic bioreactors that grow algae using a reduced source of carbon. The biomass yields obtained through photosynthetic growth range from 0.27 g/L in simple hanging bags to 20 g/L in intensively illuminated and gassed shaken flasks (Freudenberg, Baier, Einhaus, Wobbe, & Kruse, 2021). Comparatively, heterotrophic bioreactor biomass yields range from 1-2 g/L in batch systems, where all nutrients are added at beginning of the growth, and 9 – 24 g/l in continually fed bioreactors (e.g., fed-batch, continuous bioreactor) (Chen & Johns, 1995, 1996).

In simple batches, the final biomass yield is limited by the amount of nutrients added at the beginning of the batch, but in fed-batches the final biomass yield can be very high since fresh nutrients are constantly added to the media. Using this approach, we increased the final biomass yield in a batch mode of operation from 0.45 g/L to 23.69 g/L in a fed-batch mode of operation (Fields et al., 2018). In this way, the cells cultured in a fed-batch never run out of substrate and grow until either an essential nutrient is depleted from the medium, a nutrient which is being overfed accumulates in the medium becoming toxic or a by-product of cell metabolism accumulates in the medium inhibiting growth (Fields et al., 2018; Stoffels, Finlan, Mannall, Purton, & Parker, 2019; Z. Zhang, Tan, Wang, Bai, Fan, Huang, Wan, & Li, 2019). Since we previously were able grow *C. reinhardtii* at high-cell density with no toxins being present in the medium (Fields et al., 2018), we decided that the main reason behind the cell growth arrest at the highest cell density obtained was nutrient imbalance. Through Inductively Coupled Mass Spectrometry (ICP-MS) we measured the concentration of all micronutrients overtime and discovered that sulfur was becoming limiting in the media. By supplying extra sulfur to the media, we achieved a new final biomass yield of 40 g/L.

At this point, we had two options moving forward: re-iterate the process and measure the concentration of each micronutrient overtime using ICP-MS and further improve the media, or accept the new final yields and try to produce a recombinant protein to ascertain the final recombinant product concentration, which is the end goal. We decided on the latter, and thus we engineered a *C. reinhardtii* strain that produced and secreted a commercially relevant recombinant protein. In our case, the relevant recombinant protein was a soluble truncated version of human ICAM-1 (Intercellular Adhesion Molecule 1). ICAM-1 is a transmembrane protein belonging to the immunoglobin superfamily and consists of a small intracellular domain at the C-terminus, a transmembrane domain, and five Ig-like extracellular domains in the N-terminus. The structure of this protein is characterized by heavy glycosylation on the extracellular domains, as well as multiple loops due to di-sulfide bridges. It serves as a cell surface receptor of endothelial cells in the upper airways mucosa used by leukocytes to transmigrate into the tissue. Additionally, this protein is exploited by rhinoviruses to infect the respiratory epithelium (Bella, Kolatkar, Marlor, Greve, & Rossmann, 1998). Due to its structural complexity and clear biological function, we chose this protein as a method to evaluate the potential of C. reinhardtii at producing complex human recombinant proteins at competitive yields.

In this manuscript we show how we took a fed-batch mode of operation that yielded 23.69 g/L of biomass, optimized the media to increase the biomass yields. Then we engineered a strain of *C. reinhardtii* to produce a human recombinant protein to assess the yields of recombinant protein that could be obtained in the bioreactor.

Methods

Strain of algae, transformations and media

The algae used in this study is a *Chlamydomonas reinhardtii* strain obtained after the progeny of the mating of multiple wild-type strains was placed in a bioreactor containing a concentrated nutrient media. Said fed-batch was allowed to proceed until a single strain represented the majority of the population. Then, the strain was plated on TAP agar plates and single colonies were selected. Said strain of algae (CR25) was chosen for further work in bioreactors. The algal strains were grown in the lab by culturing in Erlenmeyer flasks containing TAP (Tris-acetate-phosphate) medium placed on orbital shakers and constantly illuminated (Fields et al., 2018). CR25 was transformed following a modified electroporation protocol reported in (Beth A. Rasala et al., 2012). The only differences are the amount of linearized plasmid per transformation was 3 µg, and the electroporator settings were 1250 V/cm, capacitance 50 µF and 800 Ω. After transformation, the product of each transformation was plated on five TAP/agar plates containing 15 µg/mL of Zeocin (TAP-Zeo plates) (Thermo Fisher Scientific, Waltham, MA, catalog #R25001) to select for positive transformants (Beth A. Rasala et al., 2012).

Plasmid construction and expression screening

The amino acid sequence of the protein to be expressed was determined from (Bella et al., 1998). The protein D2 is a truncated soluble protein that corresponds to the first 185 aminoacids of the human protein ICAM-1 (Uniprot ID: P05362). The gene was codon optimized using the codon bias in *Chlamydomonas reinhardtii* nuclear genome (http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=3055) and the online OPTIMIZER tool (http://genomes.urv.es/OPTIMIZER/). Additionally, the sequence coding for the secretion peptide ARS1 (Beth A. Rasala et al., 2013) was added on the 5’ end of the gene and the *C. reinhardtii* optimized sequence coding for FLAG® tag, optimized in the same manner as D2, was added at the 3’ end. A XhoI cutting site was added at the 5’ end of the gene and a BamHI cutting site was added at the 3’ end of the gene. The resulting DNA sequence was ordered

as a Gblock™ from Integrated DNA Technologies (Coralville, IA, USA). The gene was inserted into the expression vector pBR9 (Beth A. Rasala et al., 2013) through traditional restriction digest cloning using the XhoI/BamHI sites and T4 ligase (New England Biolabs, Ipswich, MA, USA, catalog# M0202S)(Figure 4.3). After transformation and selection on TAP-Zeo plates, 200 individual transformants were selected and patched into TAP-Zeo plates.

Protein expression was assessed using a variation of a dot blot. In this dot blot, a nitrocellulose membrane (Bio-Rad Laboratories Inc., Hercules, CA, USA, catalog #1620122) was cut and fitted on top of a TAP-Zeo plate and said membrane can hold up to 100 individual transformants if a grid is carefully drawn on top of it with a soft pencil. Transformants were then picked up with a pipette from a Tap-Zeo plate and approximately 1 µL of cells were carefully placed on top of the nitrocellulose membrane. Cells were allowed to grow on top of the membrane for 48 hours and then the biomass was carefully stripped from the membrane using a plastic Pasteur pipette. Afterwards, the membrane was blocked using Blocking solution for 1 hour at room temperature in constant shaking. The blocking solution composition is: Tris 20mM at pH 7.5, NaCl 150 mM, 1% Tween-20 and 5% of Non-Fat Dry Milk Powder (Bio-Rad Laboratories Inc., Hercules, CA, USA, catalog #1706404XTU). Then it was incubated with the antibody Anti-Flag® conjugated with Alkaline Phosphatase (Sigma-Aldrich, St. Louis, MO, USA, catalog #A9469) at a 1:2000 dilution for 1 hour at room temperature. The membrane was then washed three times with TBST for ten minutes each wash, and then it was developed using an NBT/BCIP tablet following manufacturer's instructions (Roche, Penzberg, Germany, catalog #11697471001). TBST composition is: Tris 20mM at pH 7.5, NaCl 150 mM, 1% Tween-20.

Bioreactor set up and growth media

The experiments were conducted in a bioreactor Minifors 2 (Infors-HT, Basel, Switzerland) of 5 liters maximum volume, with an initial volume of 3 liters per batch. The batches were inoculated by scraping 3 grams of wet algal biomass from agar plates and introducing said biomass into the bioreactor under sterile conditions in a biological hood. The batch was then started, and the parameters of the batch were set as follows: temperature was kept constant at 30 degrees Celsius, pH at 6.5, pO_2 was kept over

20% by increasing the speed of the mechanical stirring of the bioreactor, stirring changed from 400 RPMs to 1200 RPMs based on the pO_2 value, airflow was constant at 1 l of air · l of culture^{-1} · minute^{-1}. The basal media employed, called T5, had a composition of: $NH_4CH_3CO_2$ 3 g/L, KH_2PO_4 5 .89 g/l, $MgSO_4 \cdot 7H_2O$ 0.8 g/l, $Ca(NO_3)_2$ solution 1 ml/l and Trace solution 4 ml/l (Kropat, Hong-Hermesdorf, Casero, Ent, Castruita, Pellegrini, Merchant, & Malasarn, 2011). $Ca(NO_3)_2$ solution is generated by dissolving 1.6 g of $Ca(NO_3)_2$ in 20 ml of water. Autoclave media at 121°C for 60 minutes. After, adjust pH with NH_4OH 28% solution in water until pH reaches 6.6. The T5 feed, which was automatically added through a pump every time the pH rose above 6.5 due to substrate consumption by the cells, was composed of: 77.08 g of $NH_4CH_3CO_2$, 543.1 g of CH_3COOH glacial, and add water until a final volume 1.1L is reached. The feed solution was filter sterilized, using a 0.45 µm vacuum filter, into an autoclaved bottle.

The new basal media employed, called T10, had a composition of: $NH_4CH_3CO_2$ 3 g/L, KH_2PO_4 5 g/l, $MgSO_4 \cdot 7H_2O$ 1 g/l, $Ca(NO_3)_2$ solution 1 ml/l, Trace Fe- solution 15 ml/l, $FeSO_4 \cdot 7H_2O$ 110 mg/l and yeast extract 5 g/l. $Ca(NO_3)_2$ solution is generated by dissolving 1.6 g of $Ca(NO_3)_2$ in 20 ml of water. Adjust pH with NH_4OH 28% solution in water until pH reaches 6.25. After, autoclave media at 121°C for 60 minutes. Trace Fe- solution is prepared by using the stock solutions described in(Kropat et al., 2011) and mixing them as follows: 700 ml of water into a 1 L bottle, add 175 mL of EDTA stock solution, add 2.3 g of sodium carbonate, separately mixing the following three solutions of 2.6 ml of Zinc solution with 10 ml of EDTA stock solution, 2.5ml of manganese solution with 10 ml of EDTA stock solution, 4ml of copper solution with 10 ml of EDTA stock solution, add the three solutions premixed to the previous mix in the bottle, add 3 ml of molybdenum solution and add 4.5 ml of selenium solution. The T10 feed, which was automatically added through a pump every time the pH rose above 6.5 due to substrate consumption by the cells, was composed of: 15 g of KH_2PO_4, 10 g of $MgSO_4 \cdot 7H_2O_1$, 3 ml of $Ca(NO_3)_2$, 77.08 g of $NH_4CH_3CO_2$, 543.1 g of CH_3COOH glacial, and add water until a final volume 1.1L is reached.

Sampling, biomass and total protein measurements

Samples were taken from the bioreactor using a Super Safe Sampler (Infors HT, Basel, Switzerland). Daily samples were taken, purging the first 10 ml to clear the lines and sampling 15 ml. Aliquots were taken for different analysis: biomass, total protein and recombinant protein.

For biomass measurements, samples were taken from the bioreactor at different time points. Then, 2 ml of sample were washed with 100 ml of Milli-Q filtered water and their Dry Weight (DW) were determined using the method described by Zhu & Lee (Zhu & Lee, 1997).

Samples were taken from the bioreactor and centrifuged at 7000 RPM for 7 minutes in a Centrifuge 5430 R using a F-35-6-30 rotor (Eppendorf, Hamburg, Germany). The supernatant was then collected, and total protein concentration determined using the Pierce Coomassie Protein Assay Kit following the manufacturer's protocol (Thermo Fisher Scientific, Waltham, MA, USA, catalog #23200). The assays were performed in Corning® 96-well plates (MilliporeSigma, Burlington, MA, USA, catalog #CLS9018BC), and measured using an Infinite® M200 PRO plate reader (Tecan, Männedorf, Switzerland).

SDS-PAGE, Western Blot, recombinant protein quantifications and Coomassie staining

Just like with total protein measurements, the samples were centrifuged and the supernatant analyzed using SDS Polyacrilamide Gel Electrophoresis (SDS-PAGE). Samples were loaded into a MiniProtean TGX 12% (Bio-Rad Laboratories Inc., Hercules, CA, catalog #4561035) gel. Samples were prepared by mixing biological sample and loading buffer in a 3:1 ratio. The loading buffer was prepared by mixing 4X Laemmli buffer (Bio-Rad Laboratories Inc., Hercules, CA, catalog #1610747) and β-Mercaptoetanol in a 9:1 ratio. Samples mixed with loading buffer were incubated at 70°C for 10 minutes. Equal volumes of sample were loaded into each well. A 10 µl of prestained protein ladder were added at the first lane of each protein gel (AccuRuler RGB Plus Prestained Protein Marker Broad Range, Biopioneer, San Diego, CA, USA). Gels were run at 200V for 30 minutes. Once the gel was run, the gel was incubated with Instablue staining (Thermo Fisher Scientific, Waltham, MA, catalog #LC6060) following the manufacturer instructions. For western blot, the proteins were transferred from an unstained

gel to a nitrocellulose membrane using a Mini Trans Blot Cell (Bio-Rad Laboratories Inc., Hercules, CA) and were transferred for 1 hour at 200mA. Membranes were blocked, incubated with antibody, washed and developed as previously described for Dot-Blot assays.

Recombinant protein quantifications were performed analyzing the supernatant samples through western blot loaded with known amounts of Flag® standard (Qoolabs, San Diego, CA, USA, catalog #915201-001), taking a picture of the developed western blot with a Pixel 4a (Google, Mountain View, CA, USA), and measuring the intensity of the bands with Image Studio Lite (LI-COR Biosciences, Lincoln, NE, USA). One sample, at batch time 465 hours, was directly quantified by loading equal volumes of four different dilutions in water (1:1, 1:2, 1:4 and 1:8) and four different concentration of the standard (1 mg/L, 5 mg/L, 10 mg/L and 20 mg/L). A standard curve for band intensity was then generated (R^2=0.994) and the sample was measured. Only the dilutions 1:4 and 1:8 fell within the range of the standard curve, and the average result was 46.60 mg/L ± 11.34 mg/L. Then using a western blot in which all daily samples are present and using the known concentration of sample t=465h and the relative intensity of the bands in the blot, the change in recombinant protein expression overtime was assessed.

Inductively Couple Mass Spectrometry

Inductively Coupled Plasma Mass Spectrometry (ICP-MS) analysis were performed by the Plasma Lab at the Scripps Isotope Geochemistry Laboratory. The samples were analyzed in a Thermo Fisher Scientific™ iCAP™ Q_cICP-MS instrument following the protocol described in (Paquet, Day, Udry, Hattingh, Kumler, Rahib, Tait, & Neal, 2021)

Protein Sequencing and analysis

Sample preparation Protein samples were diluted in TNE (50 mM Tris pH 8.0, 100 mM NaCl, 1 mM EDTA) buffer. RapiGest SF reagent (Waters Corp.) was added to the mix to a final concentration of 0.1% and samples were boiled for 5 min. TCEP (Tris (2-carboxyethyl) phosphine) was added to 1 mM (final concentration) and the samples were incubated at 37°C for 30 min. Subsequently, the samples

were carboxymethylated with 0.5 mg/ml of iodoacetamide for 30 min at 37°C followed by neutralization with 2 mM TCEP (final concentration). Proteins samples prepared as above were digested with trypsin (trypsin:protein ratio - 1:50) overnight at 37°C. RapiGest was degraded and removed by treating the samples with 250 mM HCl at 37°C for 1 h followed by centrifugation at 14000 rpm for 30 min at 4°C. The soluble fraction was then added to a new tube and the peptides were extracted and desalted using C18 desalting columns (Thermo Scientific, PI-87782). Peptides were quantified using BCA assay and a total of 1 µg of peptides were injected for LC-MS analysis.

LC-MS-MS: Trypsin-digested peptides were analyzed by ultra high pressure liquid chromatography (UPLC) coupled with tandem mass spectroscopy (LC-MS/MS) using nano-spray ionization. The nanospray ionization experiments were performed using a Orbitrap fusion Lumos hybrid mass spectrometer (Thermo) interfaced with nano-scale reversed-phase UPLC (Thermo Dionex UltiMate™ 3000 RSLC nano System) using a 25 cm, 75-micron ID glass capillary packed with 1.7-µm C18 (130) BEH™ beads (Waters corporation). Peptides were eluted from the C18 column into the mass spectrometer using a linear gradient (5–80%) of ACN (Acetonitrile) at a flow rate of 375 µl/min for 2 h. The buffers used to create the ACN gradient were: Buffer A (98% H_2O, 2% ACN, 0.1% formic acid) and Buffer B (100% ACN, 0.1% formic acid). Mass spectrometer parameters are as follows; an MS1 survey scan using the orbitrap detector (mass range (m/z): 400-1500 (using quadrupole isolation), 120000 resolution setting, spray voltage of 2200 V, Ion transfer tube temperature of 275 C, AGC target of 400000, and maximum injection time of 50 ms) was followed by data dependent scans (top speed for most intense ions, with charge state set to only include +2-5 ions, and 5 second exclusion time, while selecting ions with minimal intensities of 50000 at in which the collision event was carried out in the high energy collision cell (HCD Collision Energy of 30%), and the fragment masses where analyzed in the ion trap mass analyzer (With ion trap scan rate of turbo, first mass m/z was 100, AGC Target 5000 and maximum injection time of 35ms). Protein identification was carried out using Peaks Studio 8.5 (Bioinformatics solutions Inc.) (Guttman, Betts, Barnes, Ghassemian, van der Geer, & Komives, 2009; McCormack, Schieltz, Goode, Yang, Barnes, Drubin, & Yates, 1997).

Results

Establishing new biomass yields in a fed-batch bioreactor

We sought to reproduce the biomass yields previously obtained in (Fields et al., 2018) but in a complex bioreactor (Minivessel 2, Infors HT, Basel, Switzerland) that would allow for temperature control, live pO2 monitoring, increased aeration and stirring capabilities and foam formation monitoring. This complex bioreactor is also capable of automatically logging all the data collected for further analysis and improvement of the fed-batch strategy. Additionally, in order to reproduce the previous biomass yields of approximately 25 $g \cdot l^{-1}$ and be able to stop using Tris as pH buffering component, we discarded the media used in (Fields et al., 2018) and we designed a brand new basal media and feed called T5 media. The basal media composition was, per 1 litre of media: Ammonium Acetate 3g, Potassium Phosphate Monobasic 5.89g, Calcium solution (see Methods) 1ml, Magnesium Sulfate 0.8g and TRACE (see Methods) 4ml. The resulting media had a pH of 5.6, which was then autoclaved and after that 1.5 ml of Ammonium Hydroxide at 28% were added bringing the pH up to 6.6. The feed composition was:

After optimizing our fed-batch protocol in this new bioreactor with the new T5 media, we designed a fed-batch with the following parameters: temperature constant at 30 Celsius, air flow into the bioreactor was of 1L/min · L of culture, dynamic stirring to keep the pO2 above 30% and ranging from 400-1400 rpm, pH kept constant at 6.5 through the addition of T5 feed. After 190 hours of batch time we were at 15.4 g/L Dry Weight (DW) and we detected that cell growth was slowing down. To tackle this, we proceeded to perform a refill: a procedure in which we drain the bioreactor of its content and it's refilled with fresh media. The reason for this is because after a certain time of fed-batch the cell growth slows down and this can be due to two things: a nutrient being fed at a lower rate than it is consumed by the cells (causing depletion and starvation) or a nutrient being fed at a higher rate than it is consumed by the cells (causing accumulation and toxicity). By refilling the bioreactor in this manner, we dilute whatever nutrients could have accumulated and become toxic as well as we replenish whatever nutrients could have been depleted. In this occasion we drained the bioreactor down to 50% of its volume and we refilled it with an equal amount of fresh T5 Basal Medium. The cells inside the bioreactor continued to grow until

their growth slowed down and reached 24g/L after 303 hours of batch time, resulting in a productivity of 1.91 g/l·day (Figure 4.2). Now that we successfully replicated our previous highest biomass yield, we moved on to improve the T5 media.

Micronutrient utilization assessment in a fed-batch via Inductively Couple Plasma Mass Spectrometry

To generate new basal medium and feed formulations that would allow us to increase the biomass yields of C. reinhardtii grown in a fed-batch, we took daily samples of a new fed-batch run with the CR25 strain using T5 Basal Medium and T5 Feed, keeping all other fed-batch parameters equal. This run had two refills, occasions in which we drained the bioreactor down to 20% of its volume and refill it with fresh T5 media. We selected key samples during the fed-batch and we spun them in a centrifuge at 6000 G for 10 minutes and extracted the supernatant. The supernatant samples were shipped to a facility that analyzed the samples through Inductively Coupled Mass Spectrometry and the micronutrient concentration in each sample was measured (Figure 4.1). The results showed the cells were starved for sulfur. We can see that the concentration of sulfur diminishes rapidly as the cells grow and that growth slows down when sulfur is depleted, thus a refill is required, and sulfur is partially replenished only for it to be depleted quickly as the cells re-start growth. It is to be remarked that the error bars in the sulfur concentration measurements are big because of the measuring methodology employed. Additionally, we see another element being quite depleted, down to 18% from its starting concentration, is copper. Iron and calcium also seem to be moderately diminished in the media.. Both micronutrients, sulfur, copper and iron are heavily involved in the electron transport chain in photosynthesis. In spite on growing in a heterotrophic bioreactor, *C. reinhardtii* is well known for producing plenty of chlorophyll and other photosynthetic elements even when they are not being heavily utilized by the primary metabolism of the cell, thus requiring a steady supply of them while actively growing. Furthermore, sulfur is also heavily used in protein synthesis as two of the twenty universal amino acids contain sulfur. As a conclusion from this experiment, we decided to increase the sulfur, iron and trace element content of the basal media and add micronutrients to the previously macronutrients only feed. Additionally, the basal media was

supplemented by yeast extract as a method to supplement unknown nutrients that might contribute to achieving high cell densities.

Improved biomass yields by improving Basal Medium and Feed formulations

The new basal media employed, called T10, had a composition of: $NH_4CH_3CO_2$ 3 g/L, KH_2PO_4 5 g/l, $MgSO_4 \cdot 7H_2O$ 1 g/l, $Ca(NO_3)_2$ solution 1 ml/l, Trace Fe- solution 15 ml/l, $FeSO_4 \cdot 7H_2O$ 110 mg/l and yeast extract 5 g/l. $Ca(NO_3)_2$ solution is generated by dissolving 1.6 g of $Ca(NO_3)_2$ in 20 ml of water. The pH was adjusted with NH_4OH (28% NH_4OH dissolved in water) until pH reached 6.25. After, it was autoclaved at 121°C for 60 minutes. Trace Fe- solution is prepared by using the stock solutions described by (Kropat et al., 2011) and mixing them in a specific order. First, fill a 1 L bottle with 700 mL of Milli-Q filtered water and add 175 mL of EDTA stock solution and 2.3 g of sodium carbonate. Separately mix the following three solutions: 2.6 mL of Zinc solution with 10 ml of EDTA stock solution, 2.5 mL of manganese solution with 10 mL of EDTA stock solution and 4 ml of copper solution with 10 mL of EDTA stock solution. Add the three solutions premixed to the previous mix in the bottle and add 3 mL of molybdenum solution and 4.5 ml of selenium solution. The T10 feed was composed of: 15 g of KH_2PO_4, 10 g of $MgSO_4 \cdot 7H_2O_1$, 3 ml of $Ca(NO_3)_2$, 77.08 g of $NH_4CH_3CO_2$, 543.1 g of CH_3COOH glacial, and Milli-Q filtered water was added until a final volume of 1.1 L.

Utilizing this newly formulated media, a fed-batch was run with the same parameters as with the T5 fed-batch. The results showed significantly improved biomass yields, with a final biomass concentration of 40 g/L after 164 hour of batch time, resulting in a productivity of 5.85 g/L·day Compared to the productivity obtained from the fed-batch using T5 media we obtained a 3-fold improvement in biomass productivity (Figure 4.2). After having achieved such improvement, we had two choices: re-iterate the ICP-MS process further improving our media formulation striving for higher biomass yield and productivity or perform a fed-batch with a recombinant protein secreting strain and strive for the highest recombinant protein yields in *C. reinhardtii* in the literature to date. We chose the latter.

Creating a strain that secretes a human recombinant protein and measuring its recombinant protein yields in a bioreactor

A new strain of *C. reinhardtii* was engineered to produce a truncated version of ICAM-1 containing the first 185 aminoacids, which comprises two Ig-like domains situated in the extracellular medium and includes the binding sites for leukocytes, macrophages and rhinoviruses. This protein, namely D2, was engineered to be secreted to the extracellular medium by fusing it to the signal peptide of the natively secreted protein in *C. reinhardtii* Arylsulfatase (Figure 4.3). The D2 secreting strain was then grown in the bioreactor employing the fed-batch mode of operation that was optimized through the formulation of new basal medium and feed. The batch lasted for 21 days and it reached a maximum biomass concentration of 39g/L, resulting in a biomass productivity of 1.86 g/L·day (Figure 4.4). This is to be compared to the fed-batch in which the wild-type strain was grown, reaching a maximum biomass concentration of 40 g/L in roughly 7 days resulting in a biomass yield of 5.85 g/L·day. Both batches started at the same initial cell density of 1.4 g/L and reached a very similar final biomass concentration, but the recombinant strain took much longer to reach a growth plateau. This could be due to the fact that the previous strain had been growing in T10 media for a long time, whereas the recombinant strain was not so it did not have the same time to acclimate to such media. An alternative explanation could be that the recombinant strain had a bigger metabolic burden due to making recombinant proteins at high yields, this being detrimental to biomass accumulation in absence of antibiotics.

Recombinant protein yields of secreted ICAM-1 D2 truncated version in a fed-batch

The daily samples taken from the bioreactor were centrifuged in order to separate the cells from the extracellular media. From the supernatant, the concentration of recombinant protein was assessed via western blot against a Flag®-tag present at the C-terminus, thus tracking its evolution over time inside the bioreactor (Figure 4.5). Additionally, the absolute concentration of recombinant protein inside the bioreactor was determined by analyzing the samples via western blot alongside Flag® standards of known concentration. The recombinant protein is first detected at 161 hours after the batch was started,

with an approximate biomass concentration of 9g/L. The concentration of recombinant protein slowly accumulated in the supernatant overtime reaching a maximum of 46.60 mg/L ± 11.34 mg/L after approximately 465 hours of batch time. After that, recombinant protein concentration in the media drops drastically. It can be seen that recombinant protein production correlates with biomass accumulation, thus suggesting that recombinant protein secretion is a product coupled to primary metabolism.

Characterization of the secretome of *C. reinhardtii* in a fed-batch

The supernatant obtained from the daily samples collected from the bioreactor was analyzed through Bradford to measure the total protein concentration. The protein concentration in the supernatant correlated quite closely to the amount of biomass in the bioreactor. This could be due to two options: the cells secrete proteins at a constant rate while they are actively growing, or the sheer stress from active stirring in the bioreactor couple to strong aeration and foaming caused a percentage of the cells to lyse. The maximum yield of protein in the supernatant amounted to 9.7 g/L, which is approximately 25% of the total biomass of cells (Figure 4.5). This shows that *C. reinhardtii* has a great potential to produce recombinant proteins in a bioreactor and that the current yields of recombinant protein, albeit the highest reported in the literature to date, are but a tiny fraction of that of some native proteins. For that reason, we analyzed the supernatant through mass spectrometry to identify which proteins were the most abundant in the bioreactor supernatant. This will let us assess two things: which proteins are most actively secreted, and which purification protocol will be most suitable for our recombinant protein of interest.

From the mass spectrometry results, we determined which proteins were the most abundant in the supernatant by the spectral counts of each protein. We identified the top 100 most abundant proteins, and we assessed whether they had a signal peptide, as assessed by a signal peptide prediction software. From this analysis we discovered that only 23% of the 100 most abundant proteins in the supernatant had a signal peptide, being our recombinant protein the 85th most abundant protein in the supernatant. This could be explained by a significant rate of cell lysis that contaminates the supernatant with intracellular proteins or that most proteins secreted by *C. reinhardtii* do not have a conventional signal peptide. From the ones that do have a signal peptide, we extracted the signal peptide sequence from Uniprot and

performed a multiple protein alignment (Clustal Omega, default settings) and obtained the following consensus sequence (EMBOSS Cons, default settings): xxxxxxxxxxxxxxMxxxxxxxVLxxxxxLxVLxxxxAxAxxxxxxx. We can see that there is significant variation and there barely is conservation among sequences. Valine-Leucine pairs seem to be common, with a core of non-polar residues pre-vailing and a couple of Alanines with a random aminoacid in between towards the C-terminus.

Discussion

In this work we have shown how the final cell density of a *C. reinhardtii* fed batch could be improved by media optimization, almost doubling the previous biomass yields and tripling its biomass productivity in just one iteration. Extensive work has been performed on *C. reinhardtii* due to its use as model organism in algal molecular biology, but there have not been many efforts in improving the culturing techniques. Due to the substrate limitation of acetate, commercial interest has favored other algal strains that could consume glucose like *Chlorella sp.* (Shi & Chen, 2002), for which biomass yields of over 100 g/L have been achieved (Bumbak, Cook, Zachleder, Hauser, & Kovar, 2011). However, in this work we show that *C. reinhardtii* biomass yields are improvable, and those improvements translate into recombinant protein production improvements.

The methods described here illustrate an iterative process for media optimization that could be repeated until nutrient imbalance is no longer the limiting factor in cell growth in a fed-batch. Generally, once optimal media has been found then the oxygen transfer rate becomes limiting, specially at big scales, due to its low solubility in water (Garcia-Ochoa & Gomez, 2009). In this study, we increased the final biomass concentration from 24 g/L to 40 g/L. The highest biomass yields of *C. reinhardtii* reported in the literature to date are 25.44 g/L (Z. Zhang et al., 2019), thus we effectively increased the highest biomass yields reported by 57%.

Furthermore, the highest recombinant protein volumetric yields reported to date in the literature are 15 mg/L of a yellow fluorescent protein Venus (Ramos-Martinez, Fimognari, & Sakuragi, 2017) and 16 mg/L of GFP (Carrera Pacheco, Hankamer, & Oey, 2018). To achieve high recombinant protein titers, there are two important factors: that the cells produce recombinant protein at high levels, and that the recombinant protein produced does not degrade overtime. Fluorescent proteins like GFP are well known for their stability, so it can be misleading to establish high recombinant protein yields using those proteins if the intended final product will not be as stable (D. J. Scott, Gunn, Yong, Wimmer, Veldhuis, Challis, Haidar, Petrou, Bathgate, & Griffin, 2018). In this work we show the highest yields of recombinant protein produced in *C. reinhardtii* to date, reaching up to 46.6 mg/L. Moreover, the recombinant protein is a

truncated human protein that accurately reflects the complexity of recombinant proteins with therapeutical interest.

A remarkable result that was not expected were the 10 g/L of total protein found in the supernatant. The vast majority of this (99.5%) was composed of native C. reinhardtii protein. We identified the top 100 most abundant proteins in the supernatant, which amount to 48% of the total protein content. Additionally, we identified the proteins that presented a conventional signal peptide, which were 23/100. This could be due to two main reasons: cell lysis inside the bioreactor contaminates the supernatant with intracellular proteins and/or some proteins are secreted without a conventional signal peptide. We know there is cell lysis due to the supernatant having green color by the end of the batch, and also by the identification of RuBisCO large subunit and many Krebs cycle enzymes in the supernatant. However, it has been documented that 40-70% of the secreted proteins in plants lack a conventional signal peptide (Agrawal, Jwa, Lebrun, Job, & Rakwal, 2010; Alexandersson, Ashfaq, Resjö, & Andreasson, 2013; H. T. Tran & Plaxton, 2008). Having identified so many signal peptides that have allowed proteins to be present in the supernatant at such high concentrations opens the possibility to utilize those newly discovered signal peptides for efficient recombinant protein secretion in a high cell density bioreactor.

To conclude, we have presented in this work an iterative pipeline for media optimization that yielded significant increase in biomass accumulation in algal fed batch. Said process is not specific to the model organism used here, but it is applicable to any other organism that is grown in a culture mode that requires very stringent nutritional requirements. Furthermore, we have shown that *C. reinhardtii* produced a human recombinant protein at high levels and it is one step closer to becoming a biotechnological relevant host.

Chapter 4 is being prepared for publication. Torres-Tiji, Yasin; Fields, Francis J; Mayfield, Stephen P. The dissertation author was the primary investigator of this paper.

Figures

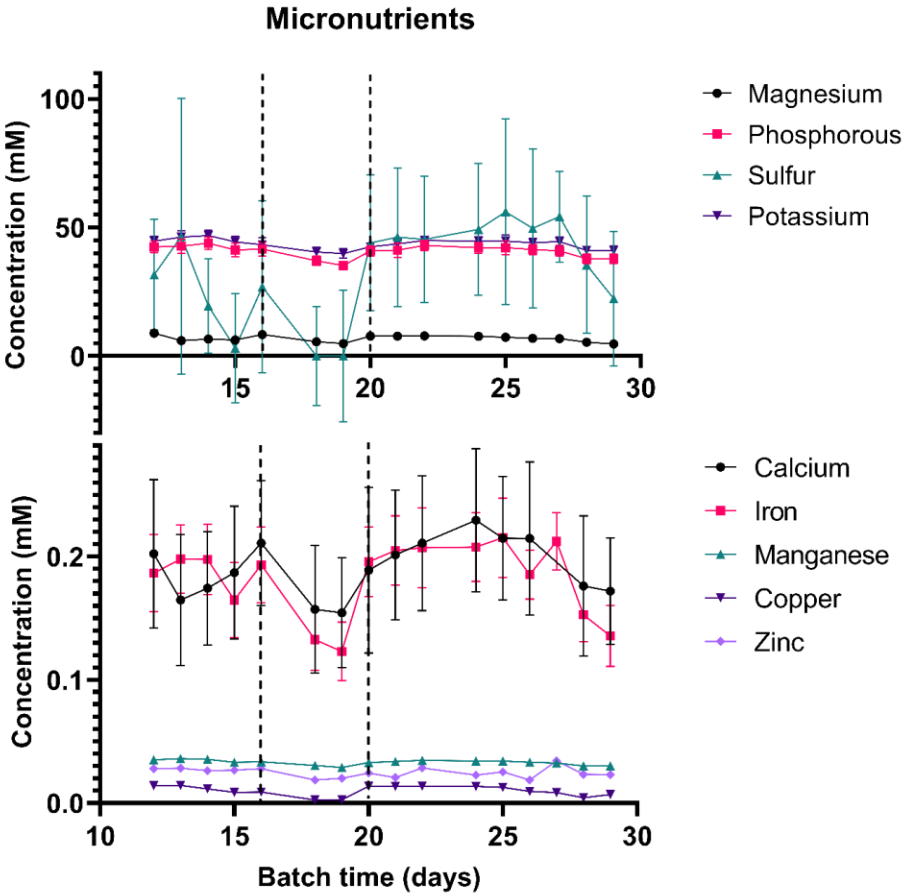

Figure 4.1. Micronutrient analysis performed using Inductively-Couple Plasma Mass Spectrometry (ICP-MS). Micronutrients were separated into two panels based on the scale of their concentration for ease of interpretation. Top panel shows the most abundant micronutrients: Magnesium, Phosphorous, Sulfur and Potassium. The bottom panel shows the trace micronutrients: Calcium, Iron, Manganese, Copper and Zinc. Daily samples are shown as points. Dashed lines indicate bioreactor knock-backs with fresh media.

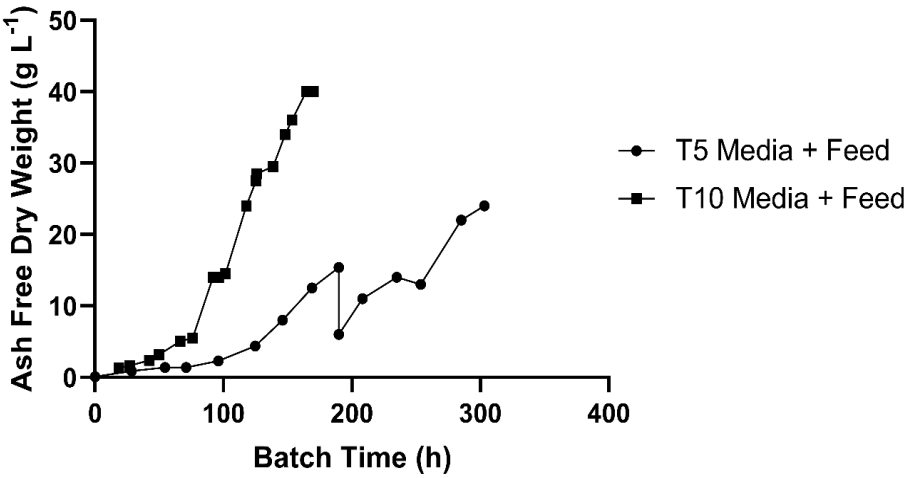

Figure 4.2. Comparison of biomass accumulation in a fed-batch using medias T5 and T10 with their corresponding feed. Biomass was measured as AFDW from approximately daily samples taken over course of the entire batch.

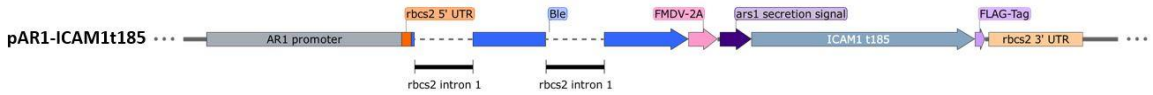

Figure 4.3. Vector map of the expression vector used to transform *Chlamydomonas reinhardtii* CR25. The vector features a robust AR1 promoter driving the expression of a single mRNA chain encoding for two separate proteins due to the presence of a FMDV-2A ribosomal-skip sequence. The first protein, Ble, confers Zeocin resistance and the second protein encodes the secreted ICAM-1t185. Said protein is secreted due to the addition of the ars1 secretion signal in the N-terminus and is Flag® tagged for ease of detection.

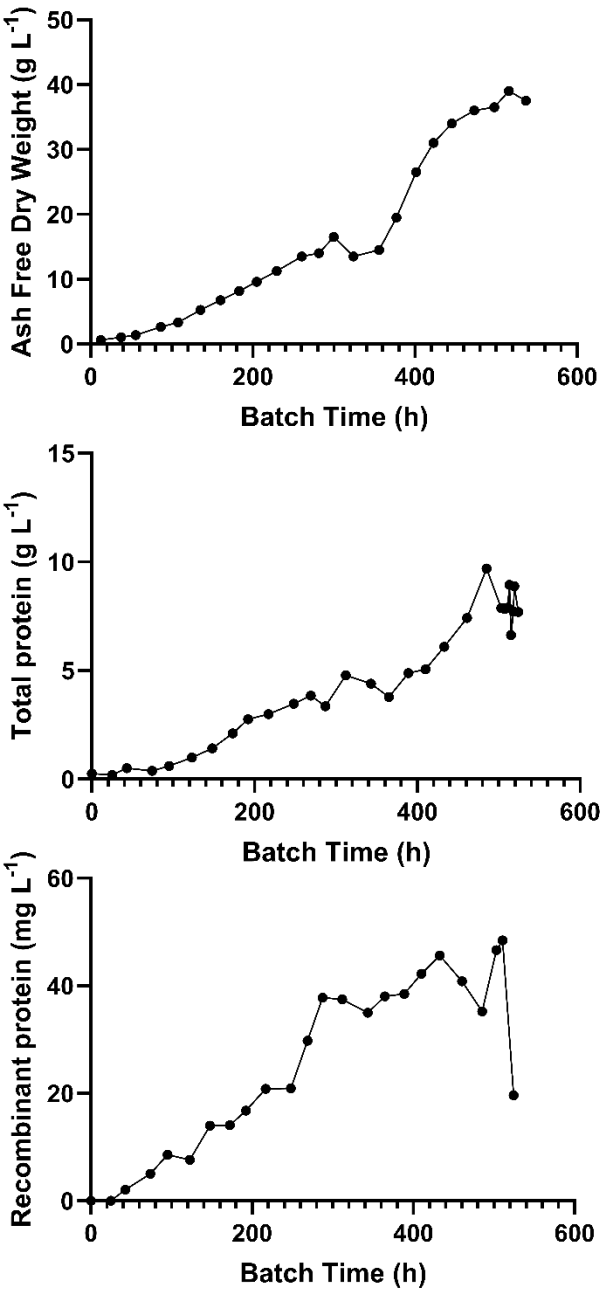

Figure 4.4. Comparison of biomass accumulation, total protein secreted and recombinant protein secreted. Data obtained from the same fed batch performed with the recombinant *C. reinhardtii* CR25 harboring the recombinant plasmid pAR1-ICAM1t185. Biomass accumulation measured using the AFDW method described in the methods section. Total protein as measure with a Bradford assay on the supernatant with dilutions ranging from 1:1 to 1:20 to stay within the dynamic linear range of the assay. Recombinant protein accumulation measured for one data point (t=511 h) with multiple dilutions and a 4-point calibration curve using Flag® Standards (R^2=0.994) and the use of Image Studio Lite (LI-COR Biosciences, Lincoln, NE, USA). The rest of the data points were estimated using relative abundance in a western blot loaded with equal volumes and using Image Studio Lite.

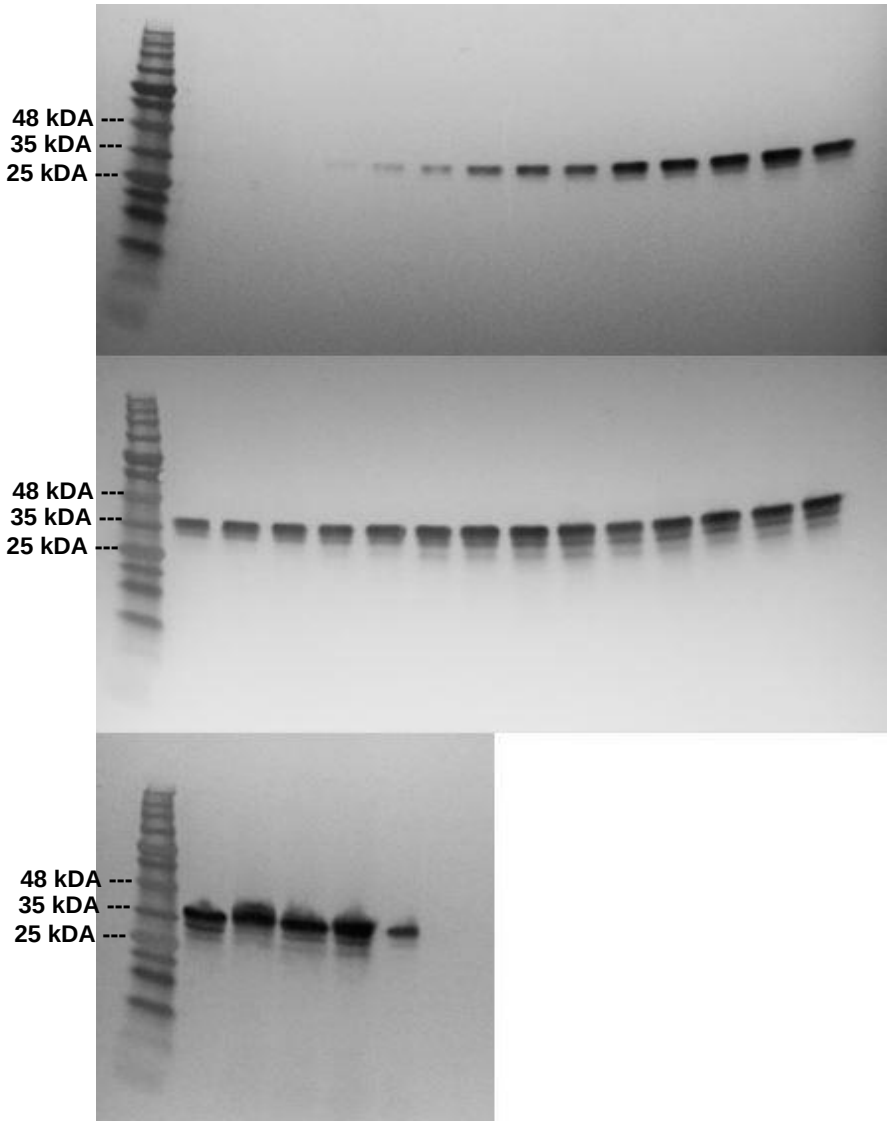

Figure 4.5. Western blot of all the supernatant samples from the fed batch using the recombinant strain. All samples were diluted 1:10 in TBS. Western blot was proved with anti-Flag® alkaline phosphatase conjugated. Predicted molecular weight is 24.8 kDa, but protein runs 10 kDa heavier on the Tris-Glycine 12% polyacrylamide gel, most likely due to post-translational modifications like glycosylation.

Tables

Table 4.1. Top 100 proteins identified in the supernatant through LC-MS-MS. Relative abundances calculated based on spectral counts (specific protein spectral counts/total spectral counts ·100)

Accession	Protein names	Protein relative abundance (%)	Avg. Mass (Da)	Signal peptide
A0A2K3DMI5	Uncharacterized protein	1.81	132090	MGAVLHLAARMARSTALLVALLGLAALGAANA
A8IXE0	Adenosylhomocysteinase (EC 3.3.1.1)	1.80	52718	NO
A0A2K3DSC2	Uncharacterized protein	1.61	101446	MLLPLLLLLLSAPFIASA
A0A2K3DSL2	Fructose-bisphosphate aldolase (EC 4.1.2.13)	1.19	40985	NO
A8J0E4	Oxygen-evolving enhancer protein 1 of photosystem II	1.10	30580	NO
A0A2K3D661	PsbP domain-containing protein	1.01	25899	NO
A8J244	Isocitrate lyase (EC 4.1.3.1)	1.00	45749	NO
A8J0N7	Phosphoenolpyruvate carboxykinase (ATP) (EC 4.1.1.49)	0.97	67537	NO
A0A2K3DLX7	S-adenosylmethionine synthase (EC 2.5.1.6)	0.92	42584	NO
A0A2K3DDA0	FAS1 domain-containing protein	0.92	156484	MGRLSVALAVAVALLAVLPGGVLS
A8JC04	Phosphoglycerate kinase (EC 2.7.2.3)	0.90	49032	NO
O48949	Protein disulfide-isomerase (EC 5.3.4.1)	0.84	58237	MNRWNLLALTLGLLLVAAPFTKHQFAHA
A8JEU4	Heat shock protein 70A (EC 3.6.1.3)	0.83	71215	NO
A0A2K3CTA8	Uncharacterized protein	0.81	60676	MGCLAILLLAASLGAVAA
A8I7T8	Binding protein 1	0.78	72493	MAQWKAAVLLLALACASYG
A8IAN1	Transketolase (EC 2.2.1.1)	0.76	77661	NO
P00877	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit) (EC 4.1.1.39)	0.76	52543	NO
A0A2K3E0I1	Uncharacterized protein	0.68	75289	NO
A0A2K3D1P1	Malate dehydrogenase (EC 1.1.1.37)	0.64	35139	NO
A8J3X4	Predicted protein	0.61	51220	MTLVVAVLLAALLVASAQA
A8JH68	Plastocyanin	0.61	14875	NO
A8J6Y8	Ferredoxin--NADP reductase,	0.61	38267	NO

Table 4.1. continued. Top 100 proteins identified in the supernatant through LC-MS-MS. Relative abundances calculated based on spectral counts (specific protein spectral counts/total spectral counts ·100)

Accession	Protein names	Protein relative abundance (%)	Avg. Mass (Da)	Signal peptide
	chloroplastic (FNR) (EC 1.18.1.2)			
A8JHX9	Elongation factor Tu, chloroplastic	0.59	94209	NO
A0A2K3D3L7	Glutamine amidotransferase type-2 domain-containing protein	0.58	174979	NO
A0A2K3DCL5	Uncharacterized protein	0.58	133118	MGLSLSWAALVLAATLLVVADA
A8I2V3	Peroxiredoxin (EC 1.11.1.24)	0.56	21642	NO
A8I980	Delta-aminolevulinic acid dehydratase (EC 4.2.1.24)	0.55	43034	NO
A8JHR9	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.-)	0.54	39760	NO
A8JH98	Phosphopyruvate hydratase (EC 4.2.1.11)	0.54	51603	NO
A0A2K3D151	Glutamine amidotransferase type-2 domain-containing protein	0.53	241844	NO
A0A2K3DDD4	Uncharacterized protein	0.52	70671	NO
A0A2K3DTU9	Thiamine thiazole synthase, chloroplastic (Thiazole biosynthetic enzyme) (EC 2.4.2.60)	0.52	36893	NO
A8HP84	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.-)	0.52	40304	NO
A8IYP4	Phosphoribulokinase (EC 2.7.1.19)	0.50	41892	NO
A2PZC2	UDP-arabinopyranose mutase (EC 5.4.99.30)	0.49	39358	NO
A8HMC0	Calreticulin	0.48	47328	MKWGVVAVLATLVVAASA
A0A2K3CTA9	Uncharacterized protein	0.48	60938	MCRIQLTIVLVAVVGLINA
A0A2K3D3E1	Elongation factor Tu, chloroplastic	0.48	55098	NO
A8HYV3	Heat shock protein 70B	0.47	71955	NO
A0A2K3DDB9	FAS1 domain-containing protein	0.47	159442	MGHIGLLLLALALLVVAPGALS

Table 4.1. continued. Top 100 proteins identified in the supernatant through LC-MS-MS. Relative abundances calculated based on spectral counts (specific protein spectral counts/total spectral counts ·100)

Accession	Protein names	Protein relative abundance (%)	Avg. Mass (Da)	Signal peptide
A0A2K3DTT8	Isocitrate dehydrogenase [NADP] (EC 1.1.1.42)	0.44	49284	NO
Q9LLL6	Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) (ADP-glucose pyrophosphorylase)	0.44	55345	NO
A8J0I0	Ribosomal protein L4	0.43	44911	NO
A0A2K3DY10	Uncharacterized protein	0.42	41446	NO
A0A2K3CZF3	Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) (ADP-glucose pyrophosphorylase)	0.41	55614	NO
A8J506	Citrulline--aspartate ligase (EC 6.3.4.5)	0.40	49021	NO
A0A2K3E7I5	Aconitate hydratase, mitochondrial (Aconitase) (EC 4.2.1.-)	0.40	89147	NO
A0A2K3DMT2	Uncharacterized protein	0.38	322939	NO
A0A2K3DKV5	Uncharacterized protein	0.38	54305	NO
A8IQU3	ATP synthase subunit beta (EC 7.1.2.2)	0.37	61822	NO
P93109	Carbonic anhydrase (EC 4.2.1.1) (Carbonate dehydratase)	0.35	27672	NO
P93108	Carbonic anhydrase (EC 4.2.1.1) (Carbonate dehydratase)	0.35	27688	NO
A0A2K3E094	Uncharacterized protein	0.35	81741	MRSSSACAGPLALAVLVAALACASA
A8J3L9	Methylmalonate-semialdehyde dehydrogenase (CoA acylating) (EC 1.2.1.27)	0.34	58161	NO
A0A2K3DKE9	Uncharacterized protein	0.33	42321	NO
A8JI07	Dual function alcohol dehydrogenase / acetaldehyde dehydrogenase	0.33	102233	NO
A8J5B8	Predicted protein	0.32	44199	NO

Table 4.1. continued. Top 100 proteins identified in the supernatant through LC-MS-MS. Relative abundances calculated based on spectral counts (specific protein spectral counts/total spectral counts ·100)

Accession	Protein names	Protein relative abundance (%)	Avg. Mass (Da)	Signal peptide
A8JCQ8	Acetate--CoA ligase (EC 6.2.1.1)	0.31	76677	NO
P93106	Malate dehydrogenase (EC 1.1.1.37)	0.31	36602	NO
A8J129	Aspartate aminotransferase (EC 2.6.1.1)	0.31	46532	NO
A8IKQ0	D-fructose-1,6-bisphosphate 1-phosphohydrolase (EC 3.1.3.11)	0.30	44444	NO
A8J8Z1	Phosphoglucomutase (alpha-D-glucose-1,6-bisphosphate-dependent) (EC 5.4.2.2)	0.30	64561	NO
A8J5Z0	60S acidic ribosomal protein P0	0.30	34575	NO
A0A2K3DKU5	Uncharacterized protein	0.29	54375	NO
A8IB85	Arylsulfatase-like protein	0.29	70124	MQRREARVGALWILAVVLLLAAGAEA
Q2VA40	Alpha-1,4 glucan phosphorylase (EC 2.4.1.1)	0.29	98296	NO
A8HVQ1	40S ribosomal protein S8	0.29	23896	NO
A0A2K3E7P1	Formate C-acetyltransferase (EC 2.3.1.54)	0.28	91105	NO
A0A2K3DER4	Uncharacterized protein	0.28	31197	MARQLSLRWALVALCAISLAASVQG
Q9FE86	Peroxiredoxin (EC 1.11.1.24)	0.28	25962	NO
A0A2K3DZ46	4Fe-4S ferredoxin-type domain-containing protein	0.28	90598	MSKRVLGLHCNSWIVALVVAGCVAAASA
A8HUK0	Peptidylprolyl isomerase (EC 5.2.1.8)	0.28	11669	NO
A8J3Y6	Formylglycinamide ribonucleotide amidotransferase (EC 6.3.5.3) (Formylglycinamide ribotide amidotransferase)	0.28	150652	NO
A8IYA1	Fe-assimilating protein 2	0.27	38779	MSRSPSIAIVLAAVALLGVCALA
A2PZC0	Zygote-specific Zys3 like protein	0.27	40407	MASKHFALVTLALVALLSGAAIA

Table 4.1. continued. Top 100 proteins identified in the supernatant through LC-MS-MS. Relative abundances calculated based on spectral counts (specific protein spectral counts/total spectral counts ·100)

Accession	Protein names	Protein relative abundance (%)	Avg. Mass (Da)	Signal peptide
A0A2K3E0A7	Uncharacterized protein	0.26	83460	NO
A8J5F7	6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)	0.26	61025	NO
A8ISZ1	Elongation factor EF-3	0.26	115308	NO
A8J9H8	Nucleoside diphosphate kinase (EC 2.7.4.6)	0.26	16541	NO
A8I528	Voltage-dependent anion-selective channel protein	0.26	28550	NO
A8IRU6	Peptidylprolyl isomerase (EC 5.2.1.8)	0.25	44695	NO
A0A2K3E587	Uncharacterized protein	0.24	43851	NO
A8IX80	Acetohydroxyacid dehydratase	0.24	64232	NO
A0A2K3D581	Oxoglutarate dehydrogenase (succinyl-transferring) (EC 1.2.4.2)	0.24	117324	NO
	ICAM-1 D2 Truncated version	0.24	23926	**MHARKMGALAVLAVACLAAV ASVAHA**
A0A2K3DRT1	Diadenosine tetraphosphate synthetase (EC 6.1.1.14)	0.24	76432	NO
A0A2K3DBQ5	Uncharacterized protein	0.24	931423	MIRYAGAIMAARSNAVLVILLTL APFAAC
A0A2K3DBP7	Uncharacterized protein	0.24	931577	MIRYAGAIMAARSNAVLVILLTL APFAAC
A0A2K3DBP8	Uncharacterized protein	0.24	930942	MIRYAGAIMAARSNAVLVILLTL APFAAC
A8JAV1	Actin	0.24	41836	NO
A8IVM9	Glycine cleavage system P protein (EC 1.4.4.2)	0.23	111351	NO
A0A2K3DB82	Uncharacterized protein	0.23	58358	MASRAFITFAVLLFAGSAFA
A8IH03	Phosphoserine aminotransferase (EC 2.6.1.52)	0.23	44130	NO
A0A2K3DW02	C-type lectin domain-containing protein	0.21	65209	NO
A8IDP6	Calmodulin	0.21	18296	NO
A0A2K3DZ72	Adenosine kinase (EC 2.7.1.20)	0.21	28203	NO

Table 4.1. continued. Top 100 proteins identified in the supernatant through LC-MS-MS. Relative abundances calculated based on spectral counts (specific protein spectral counts/total spectral counts ·100)

Accession	Protein names	Protein relative abundance (%)	Avg. Mass (Da)	Signal peptide
A8IYK1	Alpha-1,4 glucan phosphorylase (EC 2.4.1.1)	0.21	114088	NO
A8I596	Glyoxal or galactose oxidase	0.21	13939	NO
A8JFZ2	Predicted protein	0.20	96717	MRATILALVLGTLVLLANA
A8JFR9	Acetyl-coenzyme A synthetase (EC 6.2.1.1)	0.20	73224	NO

Chapter 5: Robust recombinant protein expression in an extremophile green alga grown in open ponds at a greenhouse

Abstract

Algal growth in outdoor open ponds is the best approach for massive and inexpensive production of algal biomass. However, not all algal species are suited for this mode of growth since biological contamination represents the single most important obstacle. A solution for this problem is to employ extremophile algae that can thrive in extreme medias that most biological contaminants will not be able to survive. In this work we present the identification of a novel green alga extremophile from the Chlamydomonas genus that can survive at extremely alkaline pH, as well as high salinity and high daily temperature oscillations. Said strain was temporarily named Chlamydomonas sp.-402 and it was successfully genetically engineered to express a recombinant protein, Green Fluorescence Protein (GFP) mClover variation. The recombinant strain was then grown in open ponds at a greenhouse for a month, and its growth and recombinant protein expression were characterized.

Introduction

Microalgae are a large polyphyletic group that includes a variety of eukaryotic photosynthetic microorganisms capable of living in almost all ecosystems on Earth. They represent one of the most important sources of oxygen in the world and are also responsible for a big fraction of the energy and carbon that enters the biosphere, thus being microalgae a key component for the biosphere. Their natural richness in nutrients coupled with their ability to grow fast and inexpensibly have made of microalgae a prime target to produce food, feed and nutraceuticals. Additionally, other recombinant products have been synthesized in microalgae, mostly recombinant therapeutical proteins in the model organism *Chlamydomonas reinhardtii*.

Chlamydomonas reinhardtii is a green alga that and its ability to grow photosynthetically or heterotrophically, as well as its ability to be genetically modified and express heterologous genes made it a prime target for becoming a model organism. This species has been used as a model organism for decades in the study of processes like photosynthesis, phototaxis, flagellar motility, algal genetics, etc (Harris, 2001). Due to the extensive work in this organism, all three genomes within the cell have been sequenced as well as transformed, and a variety of genetic tools for their genetic modification are available (Hannon et al., 2010; S P Mayfield & Kindle, 1990; Torres-Tiji et al., 2020). However, one of the most attractive features of microalgae when it comes to biotechnology is their capability to grow in open ponds using sunlight and minimal media. This mode of growth has one important drawback, biological contamination.

Extremophiles are organisms capable of surviving in extreme environments like near-boiling water or near-freezing cold, highly acidic or highly basic lakes, very salty or even contaminated with toxic metals (Varshney et al., 2015). It is usually a big metabolic burden for these organisms to adapt to such extreme environments, but the payoff is even bigger: they are capable of surviving in an environment with very little competition for resources. Therefore, using extremophile algae has become the main approach for growing algae in outdoor ponds for commercial purposes. An example of this practice is the culture of the green alga *Dunaliela salina*, a microorganism that thrives in extremely salty lakes and surviving in concentrations of up to 300 g/l of NaCl (Ami et al., 1982; Smith, Lee, Cushman, Magnuson, Tran, & Polle,

2010; Z. Wu, Duangmanee, Zhao, Juntawong, & Ma, 2016). This alga produces large amounts of β-carotene constituting its primary industrial source and it's used mostly as a food additive and food supplement. Another example would be the alkaliphilic cyanobacterium *Arthrospira platensis* that grows optimally between pH 8.5 and 10.5 and it's mainly used to produce food supplements and food and feed additives (Shimamatsu, 2004). The model organism *C. reinhardtii* is not an extremophile and even though it can be genetically modified, building an extremophile trait through genetic engineering is very complicated. There are naturally occurring extremophiles within the *Chlamydomonas* genus like the psychrophilic *C. nivalis* or the acidophil *C. acidophila (Malavasi, Soru, & Cao, 2020)*. However, we sought to find a green alga for which the genetic tools designed for *C. reinhardtii* would work while exhibiting an extremophile trait for alkaline pH and high salt so it would grow well in outdoor ponds using a proprietary media containing 15 g/L of Sodium Bicarbonate at pH 10.

Taking advantage of the vast biodiversity that microalgae showcase, bioprospection is a very powerful tool for finding the right strain for a bioprocess. For this project, we bioprospected green algae strains that would grow well in the special media designed for outdoor growth. After the selection and screening process of all the strains isolated, we found a green alga from the *Chlamydomonas* genus that presented all the qualities we were looking for: tolerates pH values up to 12 and salinity up to 20 g/L of Sodium Bicarbonate. We named that strain *Chlamydomonas sp.-402*, and in this manuscript we will describe how we achieved robust heterologous gene expression in this alga while growing in open ponds.

Methods

Algal strains and growth conditions in the lab

The algal strains utilized in this study are *Chlamydomonas reinhardtii* CR25, described in Chapter 4, and a novel species isolated from the wild from the *Chlamydomonas* genus named *Chlamydomonas sp.-402*. The algae were cultured in TAP liquid culture following the methods described in (Beth A. Rasala et al., 2012). Additionally, cells were also cultured in HSM media in a CO_2 enriched box as described in (Fields, Hernandez, Weilbacher, Garcia-Vargas, Huynh, Thurmond, Lund, Burkart, & Mayfield, 2021).

Plasmid design and assembly

The control expression vector pAR1 was designed based on the sequences from pBR9 (Beth A. Rasala et al., 2013), pOpt (Lauersen et al., 2015b) and pRMC (Berndt et al., 2021). The plasmid pAR1-mClover was generated by inserting a codon optimized mClover (GFP) protein between the end of the gene *ble* and the *rbcs2* 3' UTR. The mClover gene was obtained from pOpt_mclover (Lauersen et al., 2015b) and the pAR1 vector was obtained from pBle-GFP (Beth A. Rasala et al., 2012). The cloning was made using PCR-based amplification to add overlapping regions to the different fragments and use Gibson-style assembly methods (HiFi Assembly Kit, New England Biolabs, Ipswich, MA, USA).

The plasmid pAR1-ICAM1t185 was generated by cloning a codon optimized truncated ICAM-1 between the FMDV-2A sequence and the *rbcs2* 3' UTR. The details on the cloning procedure and the codon optimization of ICAM-1t185 can be found on the methods section of Chapter 4.

Transformation, selection of algae and genescreen

Transformation on algal cells were performed by electroporating algal cells on early log phase with 3 µg of linearized plasmid. More details on the transformation protocol can be found in methods section of Chapter 3.

Genescreens were performed by generating an algal lysate boiling a small colony of algae (1-2 mm of algae into a p10 pipette tip) in 20 µL of 10X TE (100 mM Tris-HCl, 10 mM EDTA, pH 8.0) at 95°C for 10 minutes. Use 2 µL from the lysate, trying not to pipette cell debris, as DNA template for a PCR reaction. PCR was performed using Q5® High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, USA, catalog# M0491S) following manufacturer's protocol.

PCR results were analyzed in agarose gels at 1% concentration in TAE buffer.

Fluorescent Microscopy and fluorescence measurement using a plate reader

Cells were culture on liquid TAP media until mid-log phase (1-4·10^6 cells/mL) on rotatory shaker under constant light. The cells were live mounted on conventional microscope slides in TAP media with cover glass coverslip. Images were captured on a Delta Vision (Applied Precision Inc., Issaquah, WA, USA) optical sectioning microscope system composed of an Olympus IX71 inverted microscope (Center Valley, PA, USA) equipped with an Olympus UPlanSAp0 100X/1.40 objective and a CoolSNAP HQ2/ICX285 camera (Photometrics, Tucson, AZ, USA). The Tetramethylrhodamine isothiocyanate (TRITC) filter (excitation 555/28 nm, emission 617/73 nm) was used to image chlorophyll autofluorescence. The GFP filter (emission 470/40 nm, emission 525/36 nm) was used to image mClover fluorescence. Image acquisition was performed using Resolve3D SoftWoRx-Acquire (Version 5.5.1, Applied Precision Inc, Issaquah, WA, USA). Brightness and contrast were adjusted identically across all images using FIJI software. Chlorophyll autofluorescence was false colored using Adobe Photoshop (San Jose, CA, USA). Methods describing microscopy image acquisition have been adapted from (Berndt et al., 2021).

The fluorescence microplate reader assay was performed as described in Chapter 3. To measure mClover fluorescence the wavelengths/bandwidth were (excitation 505/9 nm, emission 540/20 nm) and

gain settings set as optimal (gain 204). Chlorophyll fluorescence was measured using the wavelengths/bandwidth (excitation 440/9 nm, emission 680/20 nm) and gain settings set as optimal (gain 123). To calculate the normalized mClover fluorescence value we divided the mClover fluorescence value over the chlorophyll fluorescence value and multiplied that by 100.

Bradford Assay, PAGE-SDS and western blot

Samples were lysed following the methodology explained in Chapter 3. Total protein concentration in samples was determined using the Pierce Coomassie Protein Assay Kit as described in Chapter 3. Samples were analyzed through PAGE-SDS and western blotted following the methods described in Chapter 4. The antibody used to probe GFP was anti-GFP alkaline phosphatase conjugated (Abcam, Cambridge, UK, catalog# 6661), used at a dilution 1:2500 in blocking solution.

Open pond growth

Samples were grown in 75 L open ponds located at the CAL-CAB field station (Fields et al., 2021). The ponds were inoculated with a saturated 4 L algal culture, which was used to seed a 20 L culture in the ponds. After a week, the pond volume was raised to its operational volume (75 L). Ponds were grown in a semi-continuous fashion in alkaline media with high bicarbonate loads. No gaseous CO_2 was added. The ponds were knocked back every 48 hours with fresh media to maintain pH and cell density.

Results

Chlamydomonas sp.-402 shows extremophile traits of interest

Chlamydomonas sp.-402 was isolated from the wild in the field station at University of California San Diego, in which said strain was found in a rain puddle. The strain was found to be growing in a near monoculture state while thriving and mating, showing resilience to contamination. It was then isolated into axenic cultures and their growth characteristics in specialized media was assessed. This strain was found to grow much better on media containing 25g/L $NaHCO_3$ than C. reinhardtii (Data not shown). *C. sp.-402* was then grown in open ponds at our greenhouse at UC San Diego, were it tolerated pH values over 11 and temperatures over 40°C (Data not shown). Additionally, *C. sp.-402* was able to grow in the open ponds for an extended period of time without the culture ever crashing due to contamination, as opposed to *C. reinhardtii* which usually crashes in about a month (Data not shown). After confirming that *C. sp.-402* had resilience to contamination in outdoor growth, resisted extreme pH and was capable of sexual reproduction, we decided to attempt heterologous DNA expression in this strain.

Transformation of *Chlamydomonas sp.-402* with heterologous DNA

Chlamydomonas sp.-402 was isolated with the purpose of becoming a commercial strain that could be genetically engineered to produce commercially relevant bioproducts. We transformed this strain with an expression vector optimized for *C. reinhardtii* that confers antibiotic resistance to Zeocin and expresses a human recombinant protein targeted for secretion. The human recombinant protein is a truncated version of ICAM-1 (amino acids covered 1-185 starting from the N-terminus) including the signal peptide from the native *C. reinhardtii* protein Arylsulftase 1 (amino acids 1-26 starting from the N-terminus). We transformed *C. sp-402* following the electroporation-based protocol optimized for *C. reinhardtii* (Beth A. Rasala et al., 2012), in technical replicates. The transformations were plated on TAP agar plates containing 15 µg/mL of Zeocin, and after seven days Zeocin resistant colonies appeared.

From all three transformations, only 6 colonies showed and none for the negative controls (*C. sp.-402* transformed with 10 µL of DNA-free water).

Algal colony-PCR was performed on all 6 colonies against the recombinant gene encoding the truncated ICAM-1 and the PCR product was analyzed through DNA electrophoresis. All six colonies had a DNA band of the right size while the negative control (*C. sp.-402* wild-type) had none (Figure 5.2). We sequenced the DNA band and the sequence matched that of the target amplicon (Figure 5.2). We tested the six gene-positive clones for recombinant protein production through western blot, but no positive signal was obtained (Data not shown). In *C. reinhardtii* we typically observe that 5% of the clones from a transformation will show a positive signal for recombinant protein expression in a western blot. Therefore, we sought to increase the transformation efficiency to boost the number of clones thus increasing the chances at achieving recombinant protein production in *C. sp.-402*.

From routine work in the lab, we observed that *C. sp.-402* grew better in autotrophic conditions using HSM media (Sueoka, Chiang, & Kates, 1967) than in mixotrophic conditions using TAP media. Specifically, we observed that when *C. sp.-402* was grown in TAP media, the cells would form big clumps. If the cells are clumping, it could obstruct the DNA's access to the cell surface and it could also impede the isolation of genetically identical clones in agar plates. We substituted all the components of the transformation protocol that contained TAP media with HSM media; the cells were grown on HSM before transformation, they were recovered after electroporation with HSM media containing 40 mM sucrose and they were plated on HSM agar plates containing 15 µg/mL of Zeocin. Additionally, we attempted to express a simpler protein, a codon optimized Green Fluorescent protein (GFP), in a more optimized plasmid (Figure 5.3). The plasmid used contains an AR1 promoter followed by the *rbcs2* 5' UTR driving the expression of the *ble* gene fused to GFP, and the *rbcs2* 3'UTR to terminate the transcription. This plasmid was transformed following the optimized electroporation-based protocol using HSM instead of TAP. Three side-by-side transformations yielded a total of 39 clones, a 6.5-fold improvement over the previous result, although still orders of magnitude lower than the transformation efficiency of *C. reinhardtii (Kindle, 1990; Lumbreras et al., 1998; S P Mayfield & Kindle, 1990)*.

Recombinant Green Fluorescent Protein expression in *Chlamydomonas sp.-402*

We selected the top 9 clones that grew best on Zeocin containing plates and we grew them in 50 mL liquid HSM cultures. After 4 days, we measured the GFP fluorescence using a plate reader. Additionally, we analyzed GFP accumulation inside the cell by anti-GFP western blot. Of the 9 clones tested, 6 clones showed significantly higher mClover fluorescence than wild-type and 7 showed positive signal in a western blot using anti-GFP antibody (Figure 5.3). Based on this data, we down-selected from all the clones obtained to clone number 9 because of its high GFP expression and continued to do research with this strain named 402-GFP.

It has been proven that the ble protein is naturally localized in the nucleus, where it binds the antibiotic Zeocin and prevents is effects. Therefore, the fusion ble::mClover could be localized to the nucleus since the clones obtained were resistant to Zeocin. We grew the *C. sp.-402* and *C.sp.-402* GFP in 50 mL HSM liquid cultures and after 4 days we took a sample and we analyzed it through fluorescent microscopy. Using an inverted fluorescence microscope, we confirmed that the mclover GFP fluorescence localizes in the nucleus (Figure 5.3).

Robust growth in open ponds while expressing recombinant GFP

We inoculated two open ponds side-by-side with *C. sp-402* and *C.sp.-402* GFP, and we cultured them for 4 weeks. As can be seen in figure 5.4, both strains grew well and very similarly. The built-in probes monitored temperature, % O_2 saturation and pH. Temperature and % O_2 saturation showed great daily oscillations, temperature oscillations reaching a maximum of 28°C difference between day and night while % O_2 saturation oscillations reaching a maximum of 200% difference. There can be seen oscillations of pH as well, most likely due as a consequence of temperature and % CO_2 saturation oscillations. Furthermore, pH shows and upwards trend over the whole course of the experiment. The starting point is 9.5, dipping to pH 8.3 after the first 3 days, and then slowly rising overtime except when the pond would be knocked back which would lower the pH. During the last 4 days of the experiment the pH reached very basic values, between pH 10 and pH 11.

We took daily samples, and we measured the fluorescence normalized for optical density at 750 nm. As it was to be expected, the *C. sp.-402* wild-type did not show significant fluorescence throughout the culture. In contrast, *C. sp. -402* GFP showed initial high expression, 11.68-fold GFP normalized fluorescence over wild-type reading. GFP normalized fluorescence remained several folds higher when compared to that of wild-type, but it slowly declined over the course of the experiment to reach a third of the starting value.

Discussion

In this study we have isolate an extremophile green alga from the genus Chlamydomonas and identified it as a promising candidate for outdoor growth. The new species named *Chlamydomonas sp.-402* has shown robust growth at extreme pH (>11) and has tolerated high temperature daily oscillations, traits are especially relevant for outdoor growths (Malavasi et al., 2020; Varshney et al., 2015). The first one prevents algal culture crashing due to biological contamination as very few organisms can thrive in such extreme pH. The second one is relevant towards algal outdoors growth in desertic areas, where there is an overabundance of solar irradiance that can be harnessed to produce microalgal biomass while not competing with agriculture for the soil. Unfortunately, these areas present extreme oscillations in temperature between day and night, so an alga capable of resisting that abiotic stress is crucial (Urqueta, Jódar, Herrera, Wilke, Medina, Urrutia, Custodio, & Rodríguez, 2018).

We have shown that the novel algal strain is capable of incorporating heterologous DNA and expressing recombinant protein at high levels. Not only that, but the recombinant alga was capable of robust growth in our open ponds for a month without culture crash due to biological contamination while expressing the recombinant protein overtime. There is a decrease in recombinant protein production overtime, and this could be due to multiple factors. First, the recombinant alga was cultivated side by side with the wild-type version and some cross contamination could have occurred, therefore the wild-type algae could have started to take over the pong due to having a higher biological fitness. Second, the algae could have detected the expression vector as foreign DNA and silenced it (Schroda, 2006). The

vector employed was codon optimized for *C. reinhardtii*, but the codon usage between *C. reinhardtii* and *C. sp.-402* might be different thus triggering gene silencing.

As future work, we intend to sequence the genome of *C. sp.-402* and the transcriptome. With those two datasets we will be able to fully redesign the vectors employed: optimized codon usage, promoter sequences, UTRs and introns (Schroda, 2019). Additionally, one of the obstacles encountered while working with this newly discovered strain were the very low transformation efficiencies. The transformation protocol was customized for *C. sp.-402* but the electroporation step was not. Further optimization of the whole transformation protocol is likely to yield increased transformation efficiencies.

In conclusion, we have shown that an extremophile green alga was isolated from the wild that showed robust growth in outdoor ponds and it was successfully engineered to produce a recombinant protein. This makes *Chlamydomonas sp.-402* a prime candidate to become a biotechnological relevant host for low-cost production of recombinant proteins.

Figures

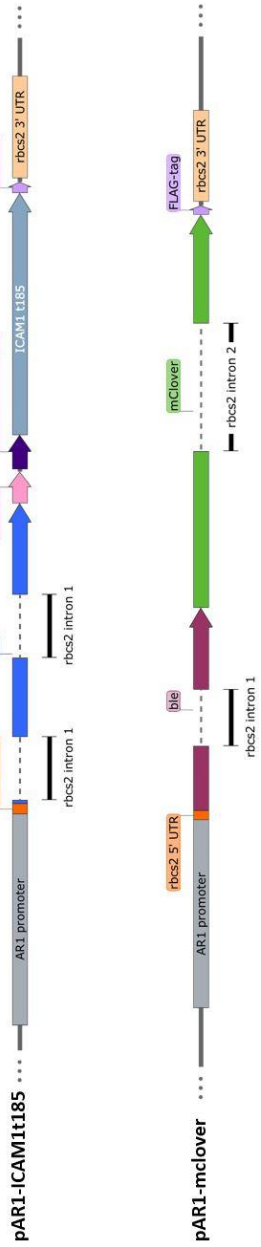

Figure 5.1. Vector map of the expression vector used to transform *Chlamydomonas sp.-402*. pAR1-ICAM1t185 presents an AR1 promoter driving the expression of the ble resistance gene which is then separated from the recombinant protein of interest, a codon optimized ICAM-1t185. pAR1-mClover presents an AR1 promoter driving the expression of the ble gene directly fused to a codon optimized mClover.

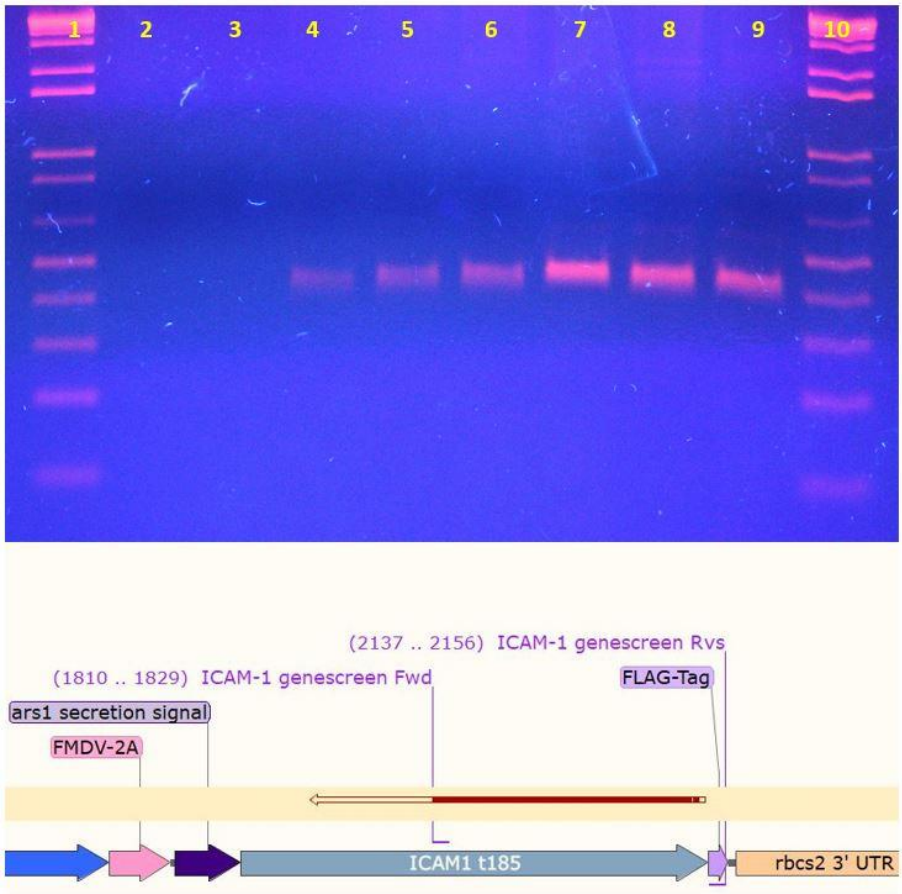

Figure 5.2. Genescreen of the integrated ICAM1-t185 into the *C. sp.-402.* Top panel, DNA electrophoresis showing the results of the genescreen. Lane 1 DNA 1Kb plus ladder, lane 2 wild-type C. sp.-402, lane 3 empty to avoid cross contamination, lane 4 clone 1, lane 5 clone 2, lane 6 clone 3, lane 7 clone 4, lane 8 clone 5, lane 9 clone 6, lane 10 DNA 1Kb plus ladder. Bottom panel shows the primers hybridization loci in the vector, and the sequenced amplicon matching the vector sequence.

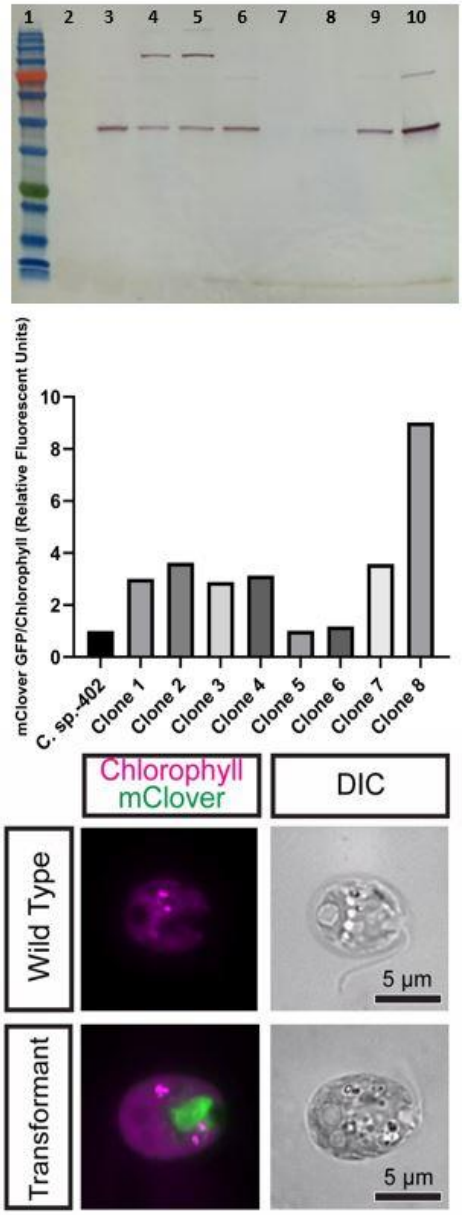

Figure 5.3. Recombinant GFP accumulation detected through western blot, microplate reader fluorescence and fluorescence microscopy. Top panel shows a western blot incubated with anti-GFP alkaline conjugated. Equal amounts of total protein were loaded as determined by Bradford assay (10 µg). Lane 1 contains protein ladder, lane 2 C. sp.-402 wild-type, lane 3 clone 1, lane 4 clone 2, lane 5 clone 3, lane 6 clone 4, lane 7 clone 5, lane 8 clone 6, lane 9 clone 7 and lane 10 clone 8. Expected band size is 40.5 kDa, apparent size is very similar. Higher molecular weight bands corresponding to dimers and trimers are visible in some lanes. Mid panel shows mClover fluorescence normalized over chlorophyl fluorescence for the same clones examine on the western blot. Bottom panel shows a comparison of fluorescence microscopy and differential interference microscopy. In green we can see mClover localized in the nucleus. In purple we can see chlorophyl falsely colored using Adobe Photoshop. Microscopy images taken by Ryan Symkovsky and edited by Anthony Berndt.

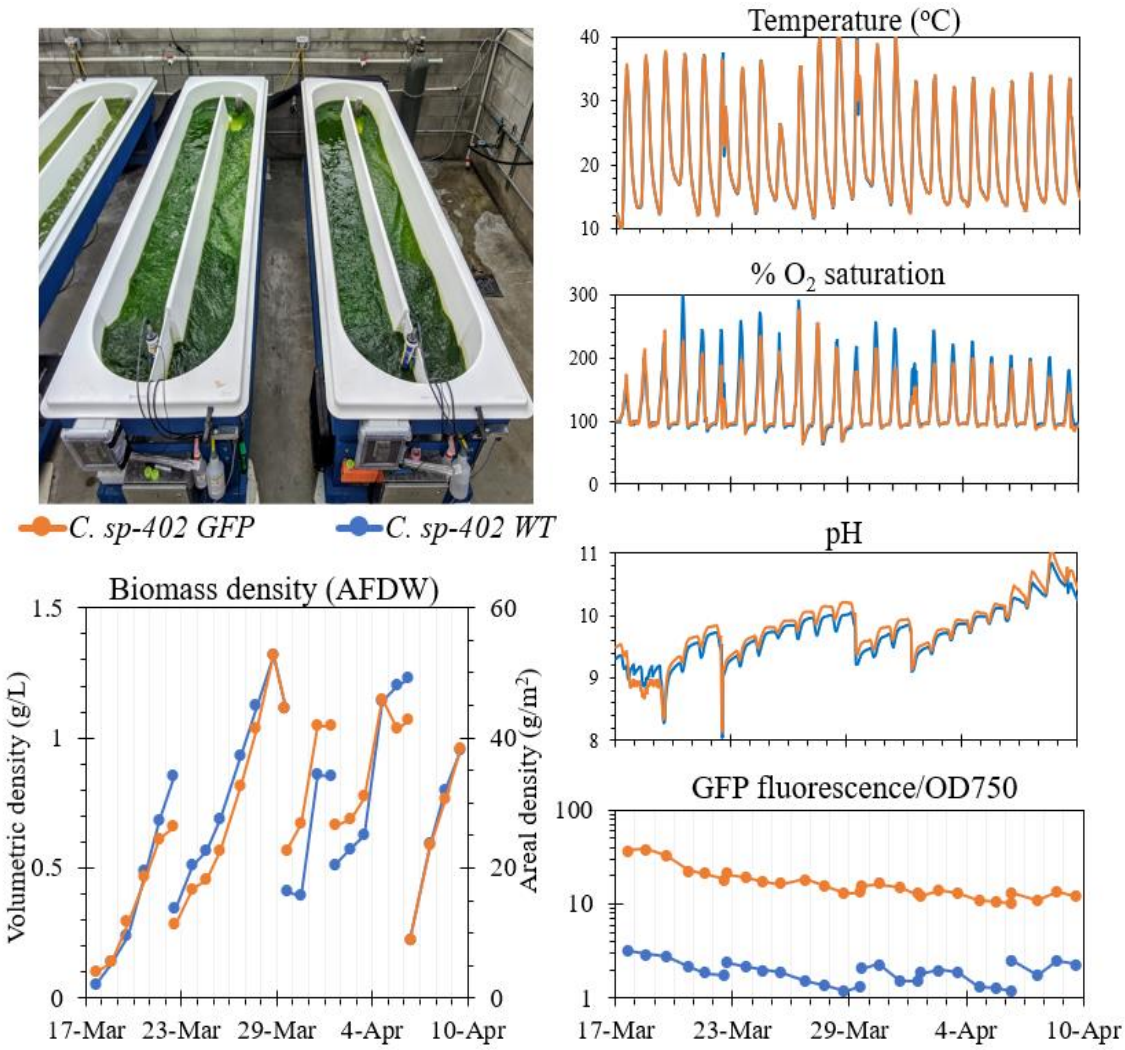

Figure 5.4**.** *C. sp.-402* wild-type and *C. sp.-402 GFP* growth in open ponds and GFP fluorescence overtime. Top left panel: strains growing in the open ponds at the field station located at UCSD. Bottom left panel shows volumetric density and areal density as measured by AFDW. Top right three panels show readings from constant monitoring of temperature, % O_2 Saturation and pH. GFP fluorescence was measured in a microplate reader from daily samples and normalized over Optical Density at 750 nm.

Chapter 6: Conclusions

Chlamydomonas reinhardtii has been used to produce recombinant proteins, first in chloroplast and later in the nucleus, with varying degrees of success. Complex human recombinant proteins showing proper folding and biological activity have been made, encoding the genes in both genomes. However, none of the proteins made reached the market for the simple reason that the product yields were just too low. In this work I focused on bridging that gap. The approach taken was to increase both the yields of recombinant protein expression at the single cell level, and the yields of biomass per unit of algal culture volume.

A novel approach in microalgae was undertaken by heterologous expression of trans-acting elements to boost transgene expression in the nuclear genome of C. reinhardtii. The element of choice was the GAL4/UAS system, in which a transcription factor from Saccharomyces cerevisiae binds to a specific 17-mer sequence (UAS) and increases transcription rates of the gene downstream from the UAS. Said system has proven successful in multitude of organisms, therefore it was possible that it could boost recombinant gene expression in C. reinhardtii. In this work I have shown that the GAL4/UAS system does increases recombinant gene expression of a reporter gene, DsRed. More specifically, the usage of the GAL4 system yielded 10-fold increase in DsRed mRNA levels and DsRed recombinant protein levels. Additionally, it also increased 10-fold the median fluorescence of the top 5% most fluorescent transformed cells as analyzed by FACS.

Biomass yields of *C. reinhardtii* grown in a heterotrophic bioreactor were increased 1.6-fold from the highest yields reported in the literature to date (Z. Zhang et al., 2019), reaching 40 g/L of AFDW. This in turn enabled the growth of a recombinant strain at very high cell densities, which yielded the highest recombinant protein concentration achieved in *C. reinhardtii*. The highest recombinant protein yields measured were 46.6 mg/L, a 3-fold increase over the highest reported yields in the literature to date (Carrera Pacheco et al., 2018; Ramos-Martinez et al., 2017). These increases were achieved by media optimization through ICP-MS. Moreover, the optimization pipeline is an iterative one, meaning it could be repeated to further improve the biomass yields although the increases obtained would most likely be smaller through the iterations. The increase achieved is significant and depending on the protein of

interest it could be considered commercial levels, but it would not be so for most recombinant proteins. Further media optimization, not only to achieve increased biomass but to achieve higher recombinant protein yields is needed. An example would be testing different nitrogen sources and characterize the effect that it has on protein synthesis and secretion in the bioreactor (D. Y. Lee, Park, Barupal, & Fiehn, 2012).

Additionally, I shown that an extremophile green alga from the Chlamydomonas genus could be transformed using genetic tools optimized for C. reinhardtii. Said strain, C. sp.-402, proved to grow robustly in an outdoors pond with highly alkaline media (pH >11) without culture crashing due to biological contamination for the duration of the experiment (1 month). The recombinant extremophile strain expressed GFP in the open ponds throughout the duration of the experiment, albeit with a slight decrease in recombinant protein yields overtime.

In conclusion, both of the goals set to bridge the gap from lab scale to commercial scale for the production of recombinant proteins in Chlamydomonas reinhardtii were achieved with satisfactory results. However, it is left to test whether the results can be stacked. If the results were to be perfectly additive, we would have recombinant protein yields of 466 mg/L in a fed-batch. That would mean commercial levels for significant number of recombinant proteins of interest. These kinds of improvements do not tend to add up in this manner, although it is very likely that they will increase the recombinant protein yields even further. As things stand right now, C. reinhardtii can be a commercially relevant host for the production of recombinant proteins for a small subset of proteins of interest. In order to expand the range of proteins that could be produced at commercial levels further optimization of the technology presented in this work is needed.

Appendix: Oral algae vaccine against SARS-Cov-2

Abstract

The potential emergence of a novel pathogen that could cause a pandemic for which the world was not ready has been prognosticated by a multitude of experts well over the last decade. Despite the warning, the world was not ready when a novel coronavirus outbreak in the Chinese city of Wuhan was detected in December of 2019. After only 4 months of its initial identification the WHO declared the global pandemic. An unprecedented collective effort to create vaccines to immunize the population against SARS-Cov-2 resulted in effective vaccines after merely 12 months of the virus detection. However, the vaccines created to date are very costly to manufacture, distribute and administer. Algal oral vaccines have proven to be successful in animal models in the past, therefore we sought to create an algae-based oral vaccine. The recombinant vaccine was designed as a recombinant protein fusion of the oral adjuvant Cholera Toxin Subunit B and the viral antigen Receptor Binding Domain (RBD) from the Spike protein of SARS-Cov-2. After successful recombinant expression, its effectivity was tested on mice.

Introduction

The threat of a global pandemic has been looming over humanity for the past years, and in late 2019 that threat became a reality when a novel coronavirus outbreak in humans was detected in the Chinese city of Wuhan. The genome of the virus was promptly sequenced (H. Wang, Li, Li, Zhang, Wang, Wu, & Liu, 2020) and due to its high similarity with the genome of the coronavirus that caused Severe Acute Respiratory Syndrome (SARS) in 2003, the novel coronavirus was named SARS-Cov-2. The virus quickly spread throughout the world during the year 2020 and by April of that year the WHO had classified the outbreak as global pandemic (Cucinotta & Vanelli, 2020). In spite of all the policies implemented like social distancing, mask wearing and lockdown, the virus has caused over 200 million infections and 4 million deaths worldwide by the time of the preparation of this manuscript (August 2021) (Hannah Ritchie, 2020). As the pandemic progresses and more epidemiologic data is collected it becomes apparent that the only way to overcome the SARS-Cov-2 pandemic is through herd immunity achieved via widespread vaccination (MacIntyre, Costantino, & Trent, 2021).

The SARS-Cov-2 virus is an airborne pathogen transmitted between people through direct, indirect or close contact with infected people that are asymptomatic, pre-symptomatic or symptomatic through infected secretions like saliva, respiratory secretions, respiratory droplets and aerosols (MacIntyre et al., 2021). Evidence shows that most of the infections occur in the nasal-lung axis via binding of the virus SARS-Cov-2 virus to the host entry factor angiotensin-converting enzyme 2 (ACE2) triggering the fusion of the viral and host cell membranes, which causes the release of the viral RNA into the host cell (Lan, Ge, Yu, Shan, Zhou, Fan, Zhang, Shi, Wang, Zhang, & Wang, 2020). The virus binds to the ACE2 receptors through the viral spike (S) protein which has two subunits, the subunit S1 containing the receptor binding domain (RBD) responsible for the binding to ACE2 and the subunit S2 which mediates the fusion of the membranes (Lan et al., 2020). Since the publishing of the SARS-Cov-2 genomic sequence on January 10th 2021 multiple of projects all around the world sought to create a vaccine against SARS-Cov-2 (F. Wu, Zhao, Yu, Chen, Wang, Song, Hu, Tao, Tian, Pei, Yuan, Zhang, Dai, Liu, Wang, Zheng, Xu, Holmes, & Zhang, 2020). After 18 months of that date, 21 vaccines have been approved for use in some countries, but most of the immunizations have been performed using the

vaccines Pzifer-BioNtech COVID-19 vaccine, Moderna COVID-19 vaccine, Oxford-AstraZeneca COVID-19 vaccine, Janssen COVID-19 vaccine, Sinopharm COVID-19 vaccine and Sputnik V (Burki, 2021). The Pfizer and Moderna vaccines are modRNA vaccines encoding the full length S protein, the AstraZeneca, Janssen and Sputnik V vaccines are adenovirus vaccines encoding the full length S protein and the Sinopharm vaccine is an inactivated virus vaccine (Caballero & Quirce, 2021; Livingston, Malani, & Creech, 2021; Rodriguez-Coira & Sokolowska, 2021; Walsh, Frenck, Falsey, Kitchin, Absalon, Gurtman, Lockhart, Neuzil, Mulligan, Bailey, Swanson, Li, Koury, Kalina, Cooper, Fontes-Garfias, Shi, Türeci, Tompkins, Lyke, Raabe, Dormitzer, Jansen, Şahin, & Gruber, 2020).The different vaccines present different efficiencies at preventing SARS-Cov-2 infection and they also present different advantages and disadvantages. Strict and expensive manufacture, distribution and administration procedures have hindered the prompt immunization of the global population, therefore enabling the SARS-Cov-2 virus to spread and mutate which has further difficulted the immunization of humanity against said virus. The changing nature of this coronavirus has forced the population into receiving booster shots of certain vaccines as the immunity against said virus appears to be transient. We believe that the discovery of an effective, cheap and easily distributable and administrable vaccine is essential in overcoming the SARS-Cov-2 pandemic. There are alternative vaccine approaches that could fulfill those requirements. Microalgae have been used in the past to generate inexpensive and stable vaccines that have been proven to be effective in animal models (Dreesen, Hamri, & Fussenegger, 2010; Gregory, Topol, Doerner, & Mayfield, 2013; Specht & Mayfield, 2014).

Microalgae are a vastly biodiverse group of microorganisms capable of oxygen photosynthesis that have gained notoriety in the biotechnological field in recent years. Several species of microalgae have their genomes fully sequenced, are amenable to genetic modification and have been proven to express recombinant proteins. Among them, the green alga and model organism *Chlamydomonas reinhardtii* is the one species with most genetic tools available for recombinant protein expression. This alga is capable of rapid growth in inexpensive minimal medias as well as production of complex human therapeutical proteins. Additionally, *C. reinhardtii* has been granted GRAS status by the FDA making them safe to eat (Murbach, Glávits, Endres, Hirka, Vértesi, Béres, & Szakonyiné, 2018). Generally, downstream processing of recombinant proteins represents most of the cost ranging from 45 to 92%

(Lowe, 2001; Straathof, 2011), therefore if the recombinant protein can be absorbed in the gut synthesizing it in algae and orally delivering whole cell algae with minimal downstream processing becomes a low-cost alternative.

Oral vaccines have been available since 1961, in which a reengineering of the Inactivated Polio Vaccine (IPV) resulted in the extremely successful Oral Polio Vaccine (OPV) (Onorato, Modlin, McBean, Thoms, Losonsky, & Bernier, 1991). Oral vaccines present a series of advantages over injectable ones: they are easier to administer, can be more affordable to manufacture and distribute, show a higher degree of compliance with immunization regimes, but perhaps the most important advantage is that they elicit both humoral and mucosal immunity (Specht & Mayfield, 2014). Mucosal immunity has proven to be essential to prevent person-to-person transmission, because most human pathogens' point of entry into the human body are the mucosal tissues a strong mucosal immunity can stop an infection before it reaches the bloodstream (Specht & Mayfield, 2014). An algal oral vaccine against *Staphylococcus aureus* has been shown to elicit both systemic IgG antibodies and mucosal IgA antibodies in mice and protect 80% of the vaccinated mice against a lethal challenge with *S. aureus* that killed all the control mice within 48 hours (Dreesen et al., 2010). The vaccine consisted of a recombinant protein fusion between the Cholera toxin subunit B (CtxB) and the *S. aureus*-specific epitope D2. Fusing the vaccine antigen with the CtxB allows for the fusion protein be internalized into the gut-associated lymphoid tissues via CtxB binding to GM1 receptor gangliosides, thus triggering the systemic and mucosal immune responses (Dreesen et al., 2010). Another algal oral vaccine was made by expressing the Pfs25 malarial antigen fused with CtxB in *C. reinhardtii*. The vaccine effectively prevented malaria transmission by disrupting the sexual cycle of Plasmodium falciparium in the midgut of mosquitos that were fed immunized mice sera. Additionally, a mucosal IgA response was elicited against Pfs25 but not an IgG response, whereas both IgA and IgG responses were elicited against CtxB. It has been discussed that this could have been due to addition for a furin protease cleavable linker between the two antigens (Gregory et al., 2013).

We decided to create an oral recombinant COVID-19 vaccine in *C. reinhardtii* by fusing the RBD domain of the SARS-Cov-2 S protein with the CtxB. Additionally, we decided to test three different subcellular localizations for the vaccine to be expressed: extracellular media, endoplasmic reticulum, and

chloroplast. After assessing which of the three options was the best vaccine candidate, we tested the vaccine in mice as an oral booster vaccine after they had been immunized through parenteral injection of recombinant SARS-2-Cov RBD protein.

Methods

Plasmid design and assembly

The plasmids were designed based on pRMC1, pRMC2 and pRMC3 plasmids (Berndt et al., 2021). The mClover was deleted from all 3 plasmids using a Q5® Site-Directed Mutagenesis Kit (New England Biolabs, Ipswich, MA, USA, catalog#E0554S). Afterwards, a codon optimized Cholera Toxin Subunit B (CtxB) was cloned between the end of the FMDV-2A sequence and the beginning of the SARS-Cov-2 RBD CDS using a HiFi Assembly Kit (New England Biolabs, Ipswich, MA, USA). The codon optimized CtxB was generated following the methodology described in (Berndt et al., 2021). The codon optimized gene included a 4 amino acid linker (GPGP) obtained from (Gregory et al., 2013).

Algal strains, culturing conditions and transformation

The strains used in this study are *Chlamydomonas reinhardtii* CR25 (see Chapter 4). Cells were cultured on TAP liquid media as described on (Beth A. Rasala et al., 2012). Algal cells in mid-log phase ($1\text{-}4\cdot10^6$ cell/mL) were transformed with 3 µg of linearized plasmid and plated on TAP-Zeo plates as described in Chapter 3.

Dot-blot and western blot screening

Dot-blots were performed to screen for CtxB::RBD secretion from transformants obtained from transformations using the GreenVaccine-2 plasmid. The methodology to perform dot-blot can be found in Chapter 4. The dot-blot assay was incubated with rabbit pAb anti-RBD (Sino Biological US Inc., Wayne, PA, USA, catalog# 40592-T62) at a 1:3000 dilution for 1 hour at room temperature. After the primary incubation, the dot-blot was incubated a second time using goat pAb anti-rabbit IgG conjugated with

alkaline phosphatase (Abcam, Cambridge, UK, catalog# 6722) at a 1:10000 dilution for 1 hour at room temperature.

Western blots were performed as described in Chapter 4. The western blot incubated with anti-RBD used the same primary and secondary antibodies as described above. The western blot incubated with anti-CtxB used a polyclonal antibody against cholera toxin B subunit conjugated with HRP. The membrane was developed using ECL™ Select Western Blotting Detection Reagent (MilliporeSigma, Burlington, MA, USA, catalog #GERPN2235) following the manufacturer's protocol. The membranes were imaged using a ChemiDoc Imaging System (Bio-Rad Laboratories Inc., Hercules, CA, USA), optimal exposure was chosen for all blots.

Immunization regime

The immunization experiment consisted of a preliminary vaccine trial in which mice would be injected with pure recombinant RBD (primary immunization, 2.4 µg of RBD in complete Freund's Adjuvant) and 24 days later they would be fed through oral gavage the oral algal vaccine (secondary immunization, 3 µg of CtxB::RBD contained in algae). Serums were collected on Day 0 (D0: 1 day before primary immunization), Day 21 (D21: 21 days post primary immunization) and Day 38 (D38: 38 days post primary immunization and 14 days post-secondary immunization). A total of 20 mice were split in different groups with the following arrangement: negative control without primary nor secondary immunization (2 female and 2 male mice), positive control with both primary and secondary immunization consisting of an RBD injection (2 female and 2 male mice), algal control group in which the primary immunization was an RBD injection and the secondary immunization was mice feeding with wild-type algae (3 female and 3 male mice) and the experimental group in which the primary immunization was an RBD injection and the secondary immunization was mice feeding with GV-3 algae (3 female and 3 male mice). For the feedings, mice were fed 500 µL of cell lysate, which in the experimental group contained 3 µg of recombinant CtxB::RBD protein, through oral gavage. Blood was collected at the indicated time points via submandibular bleeding. The methods herein described were adapted from (Yang, Wang, Chen, Lu, Yang, Bi, Bao, Mo, Li, Huang, Hong, Yang, Zhao, Ye, Lin, Deng, Chen, Lei, Zhang, Luo, Gao, Zheng,

Gong, Jiang, Xu, Lv, Li, Wang, Li, Wang, Wang, Yu, Qu, Yang, Deng, Tong, Li, Wang, Yang, Shen, Zhao, Li, Luo, Liu, Yu, Yang, Xu, Wang, Li, Wang, Kuang, Lin, Hu, Guo, Cheng, He, Song, Chen, Xue, Yao, Chen, Ma, Chen, Gou, Huang, Wang, Fan, Tian, Shi, Wang, Dai, Wu, Li, Wang, Peng, Qian, Huang, Lau, Yang, Wei, Cen, Peng, Qin, Zhang, Lu, & Wei, 2020)

ELISA

ELISA's were performed following the protocol described in (Amanat, Stadlbauer, Strohmeier, Nguyen, Chromikova, McMahon, Jiang, Arunkumar, Jurczyszak, Polanco, Bermudez-Gonzalez, Kleiner, Aydillo, Miorin, Fierer, Lugo, Kojic, Stoever, Liu, Cunningham-Rundles, Felgner, Moran, García-Sastre, Caplivski, Cheng, Kedzierska, Vapalahti, Hepojoki, Simon, & Krammer, 2020). The protein used to coat the wells was SARS-CoV-2 (2019-nCov) Spike/RBD Protein His Tag (Sino Biological US Inc., Wayne, PA, USA, catalog# 40592-V08H) for anti-RBD ELISA, or recombinant Cholera Toxin B subunit (Sigma-Aldrich, St. Louis, MO, USA, catalog #SAE0069). The 96-well plates used were Nunc MaxiSorp™ flat-bottom (Thermo Fisher Scientific, Waltham, MA, USA, catalog #44-2404-21). The optical density at 490 nm was measured using an Infinite® M200 PRO plate reader (Tecan, Männedorf, Switzerland). The AUC was calculated using Prism9 (GraphPad, San Diego, CA, USA). The baseline was set as 0.062. The untreated controls were measured in biological duplicates, the RBD+RBD positive control were measured in biological triplicates, and the remaining samples were measured in 6 biological replicates.

Results

Plasmid design

The general architecture of the plasmid is based on plasmids pBR9 and pOpt plasmids (Lauersen et al., 2015b; Beth A. Rasala et al., 2012). The plasmid encoding the vaccine is composed of the AR1 promoter and the RBCS2 5'UTR driving the expression of the antibiotic resistance gene BLE followed by the CtxB::RBD fusion and ending in the RBCS2 3' UTR. Downstream from this first cistron there is a second cistron that confers Paromomycin resistance and ensures the integrity of the genetic payload after nuclear transformation(Berndt et al., 2021). Additionally, in order to separate the BLE gene from the vaccine a foot and mouth disease 2A ribosomal-skip motif was placed in between (Beth A. Rasala et al., 2012).

The *Vibrio cholerae* CtxB amino acid sequence was obtained from Uniprot (ID: P01556, Thr22 - Asn124) and the CDS was codon optimized based on the codon usage in the *C. reinhardtii* nuclear genome. Similarly, the CDS for the recombinant RBD was based of the amino acid sequence obtained from Uniprot (ID: P0DTC2, Arg319 – Lys537) and codon optimized based on the codon usage in the *C. reinhardtii* nuclear genome. The two CDS were separated by short linker (Gly-Pro-Gly-Pro). This vector was named GreenVaccine (GV). From this vector, three final vectors were generated each one targeting the protein to a different subcellular localization. GreenVaccine-1 (GV-1) had a chloroplast targeting sequence from the native gene psaE between the 2A sequence and the CtxB CDS, GreenVaccine-2 (GV-2) had a secretion signal peptide from the native gene Pherophorin-C2, and GreenVaccine-3 (GV-3) had the same secretion signal peptide as GV-2 and an KDEL Endoplasmic Reticulum retention signal at the C-terminus of the protein fusion.

Screening of algal clones expressing the different GreenVaccines and analysis of the vaccine integrity

The three plasmids were transformed by electroporation using 3 µg of linearized plasmid. The electroporated cells were plated on TAP agar plates containing 15 µg /mL of Zeocin. The resulting clones

were screened for CtxB::RBD recombinant expression. GV-1 resulted in very poor transformation efficiencies, yielding 47 clones from 3 transformations. The 47 resulting clones were grown in 5 mL TAP cultures and cell lysates were analyzed through western blot. No positive signal was detected for none of the clones. In a similar experiment trying to express and target recombinant RBD::GFP to the chloroplast, very low number of transformants was obtained and no positive signal could be detected through western blot that showed RBD expression, suggesting that RBD could be toxic in the algal chloroplast. Hundreds of clones were recovered from GV-2 and GV-3 transformations. GV2 clones were screened through dot-blot (see Method section Chapter 4), culturing the cells on top of the dot-blot membrane for 72 h and incubating the membrane with anti-RBD antibody (anti-RBD antibiotic). As can be seen in Figure A.2, we found five clones expressing significant levels of CtxB::RBD. Those were grown in 5 mL TAP liquid cultures for 4 days and analyzed through western blot incubating with anti-RBD antibody. All the clones showed a positive band running at an apparent weight of 41 kDa, whereas the PHC2 signal peptide-CtxB-linker-RBD protein has a predicted weight of 38.8 kDa (Figure A.3). The difference in weight could be due to glycosylation performed on the recombinant protein during its journey through the secretory pathway. One of the clones was downselected for further analysis. GV3 clones accumulated the protein intracellularly and could not be screened through dot-blot, therefore they were directly screened through western blot performed on algal lysates from 5 mL TAP liquid cultures. From the approximately 200 clones screened, only one showed CtxB::RBD expression as indicated by a positive band of the same apparent size as the ones seen from GV-2 samples in a western blot incubated with anti-RBD (Figure A.3).

As can be seen in the western blots, the anti-RBD antibody showed significant unspecific binding to native proteins in *C. reinhardtii*. To make sure the positive band observed in the western blots was indeed the correct one and to ensure the integrity CtxB fragment was present we performed Mass Spectrometry protein identification on a partially purified sample from the supernatant of a GV2 liquid culture. The supernatant was separated from the cells by centrifugation, and it was purified using an Anionic Exchange Chromatography column (HiTrap Capto Q, GE, Boston, MA, USA). The protein was suspended in supernatant with the following composition: Tris 20 mM, pH 8.5, NaCl 0 M. The protein was eluted in an isocratic elution with the following buffer composition: Tris 20 mM, pH 8.5, NaCl 1M. The

proteins were then concentrated 10-fold using acetone precipitation. The protein was then analyzed through SDS-PAGE, stained with SimplyBlue (Invitrogen, Waltham, MA, USA) and the band corresponding to GV2 was analyzed through Electrospray Ionization Mass Spectrometry. The peptides identified matched perfectly the sequence of GV2 with a coverage of 34%, but with peptides matches on both the CtxB and RBD parts. This confirmed that the protein of interest was being successfully expressed.

Vaccine preparation and mice immunization regime

We analyzed through western blot the yields of each vaccine per gram of wet biomass. We obtained that GV2 had an approximate yield of 0.5 µg per wet gram of biomass, and GV-3 had an approximate yield of 3 µg per wet gram of biomass. Therefore, we decided to test the capacity of GV-3 to elicit a humoral response against RBD in a mice immunization experiment. Mice would be immunized on day 1 with a sub-cutaneous injection of 2.4 µg of RBD protein in Complete's Freud Adjuvant, and would be immunized a second time on day 24. Serum samples were collected on days 0, 21 and 38. Mice were split in four different groups: Untreated (negative control), RBD+RBD (positive control), RBD+Wt algae (algal booster negative control), and RBD+CtxB::RBD algae (experimental group).

The initial idea was to feed lyophilized whole cell algae resuspended in minimal volume of water to the mice and deliver the vaccine in this method, and the dosage required for the experiment was 3 µg of CtxB::RBD per mice and dose. However, there is a physical limitation on how much volume of algae can be administered through oral gavage at once (500 µL). Since the concentration of vaccine per volume of vaccine prep was 1 µg per mL of vaccine prep, we had to concentrate the vaccine 6-fold. For that, we collected cell pellets from 1 L TAP liquid cultures, washed the cell pellets with PBS and lysed the cells with sonication. After this, the cell debris was separated from the soluble protein and the soluble protein was lyophilized. Finally, the lyophilized protein was reconstituted with 1 mL of water to reach a final concentration of 6 µg per mL of vaccine prep (Figure A.4).

Antibody titers elicited in the mice immunization experiment

The sera from the immunized mice were collected as previously described. The levels of IgG against RBD and CtxB were measured through indirect ELISA. We coated 96-well plates with recombinant RBD or recombinant CtxB overnight at 4°C and we measured the levels of specific antibodies. The ELISA results show great variability that difficult the interpretation of the immunization success. The untreated group show only baseline signal on both D21 and D38 samples, which is exactly what was expected. The positive group that received RBD injections in both immunizations show expected results for 2 out of 3 mice tested, in which there are significant IgG titers in the D21 sample and even higher IgG titers in the D38 sample. However, one of the mice from the positive control showed only baseline levels of IgG titers, matching exactly those of the negative control group. Similarly, the mice from the algal control group show high titers of IgG in the D21 sample in 3 out of 6 mice and even higher IgG titers in the D38 sample. The other 3 mice show intermediate IgG titers. The experimental group follows a similar trend, in which 2 out of 6 mice show high titers on D21 and even higher titers on D38. However, the other 4 mice show very variable IgG titers that seem to not follow our predictions. Additionally, all the mice that received the same primary immunization (positive control, algae control and experimental groups) should present similar IgG titers on the D21 sample and this is not the case.

Discussion

In this work we have shown that an antigen for SARS-CoV-2 coupled with the adjuvant Cholera Toxin Subunit B could be made and targeted for secretion (GreenVaccine-2) or Endoplasmic Reticulum retention (GreenVaccine-3) in *Chlamydomonas reinhardtii*. No antigen could be produced and targeted to the chloroplast, indicating some potential toxicity of the RBD protein in the chloroplast (Berndt et al., 2021). It must be remarked that finding a clone that expressed GreenVaccine-2 was very easy thanks to the dot-blot assay. A clone expressing GreenVaccine-3 was much more complicated to find, and more than 200 clones had to be analyzed through western blot to find only one positive band.

In spite of that, the recombinant strains producing the CtxB::RBD fusion could not express high enough levels of antigen and we had to deviate from the original plan of feeding whole cells to the mice. Instead, we lysed the cells and lyophilized the solubilized proteins, thus concentrating the antigen 6-fold. When the vaccine was tested as an oral vaccine booster, the results obtained from ELISAs performed on mice sera post-immunization show extreme variability, with some mice in the positive control group not eliciting any antibodies against the antigen. Additionally, the fact that all the samples, but those from the negative control, on Day 21 do not show the same antibody titers indicates a major issue.

These results show that the immunization regime was not well designed and/or implemented. Similar studies have utilized one oral gavage per week for 5 consecutive weeks (Gregory et al., 2013) or 12 consecutive weeks (Dreesen et al., 2010). In order to obtain conclusive results, we need to utilize a larger sample population and, based on other algal vaccine trials, more dosages.

Figures

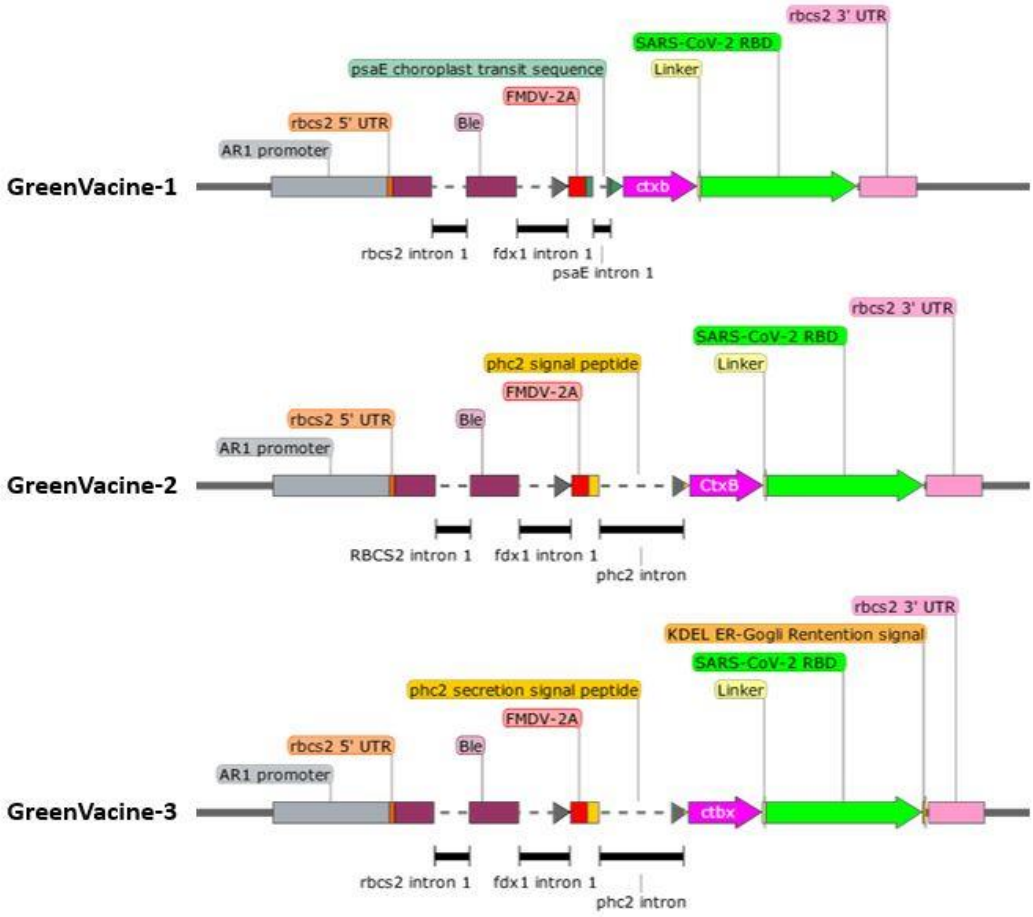

Figure A.1. Vector map showing the design of all three plasmids used in this study. They feature an AR1 promoter driving the expression of a ble gene, a FMDV-2A ribosomal-skip motif, Cholera Toxin Subunit B, a 4 AA linker, and the antigen SARS-CoV-2 RBD. The only differences are in the signal peptides used. GreenVaccine-1 has a psaE chloroplast transit sequence, GreenVaccine-2 has a phc2 signal peptide, and GreenVaccine-3 has a phc2 signal peptide and a KDEL ER-retention signal.

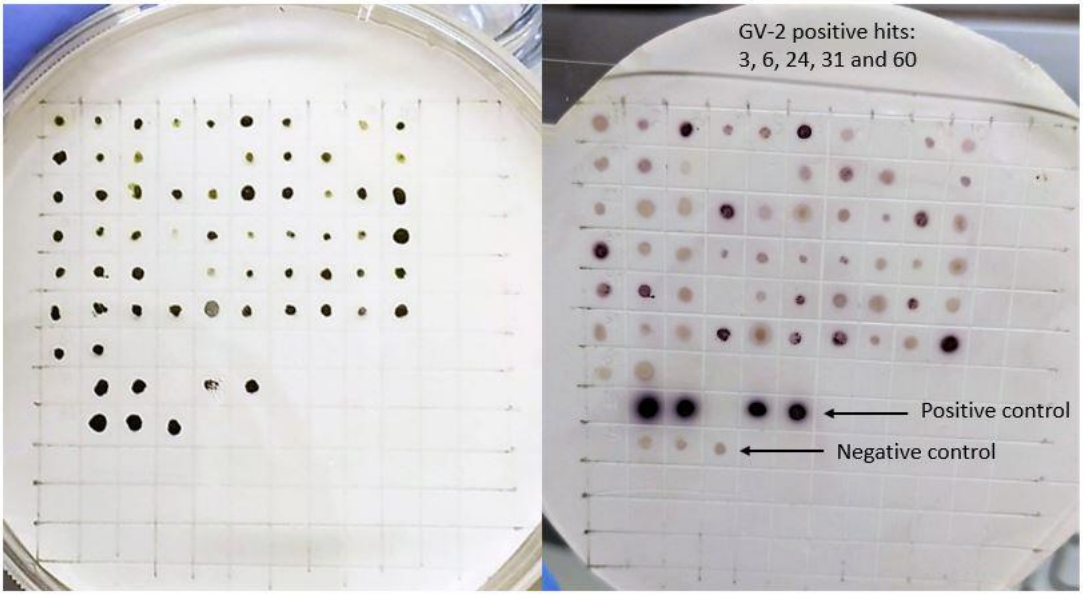

Figure A.2. Dot-blot using live algae. On the left panel we see the dot-blot after cells have grown on it for 72 hours before rinsing. On the right panel we see the same dot-blot after it has been incubated with anti-RBD, a secondary antibody conjugated with alkaline phosphatase and developed with BCIP/NBT substrate. Negative control were cells from C. reinhardtii CR23 wild-type and positive control were cells from RMC2 strain from (Berndt, Smalley, Ren, Badary, Sproles, Fields, Torres-Tiji, Heredia, & Mayfield, 2021).

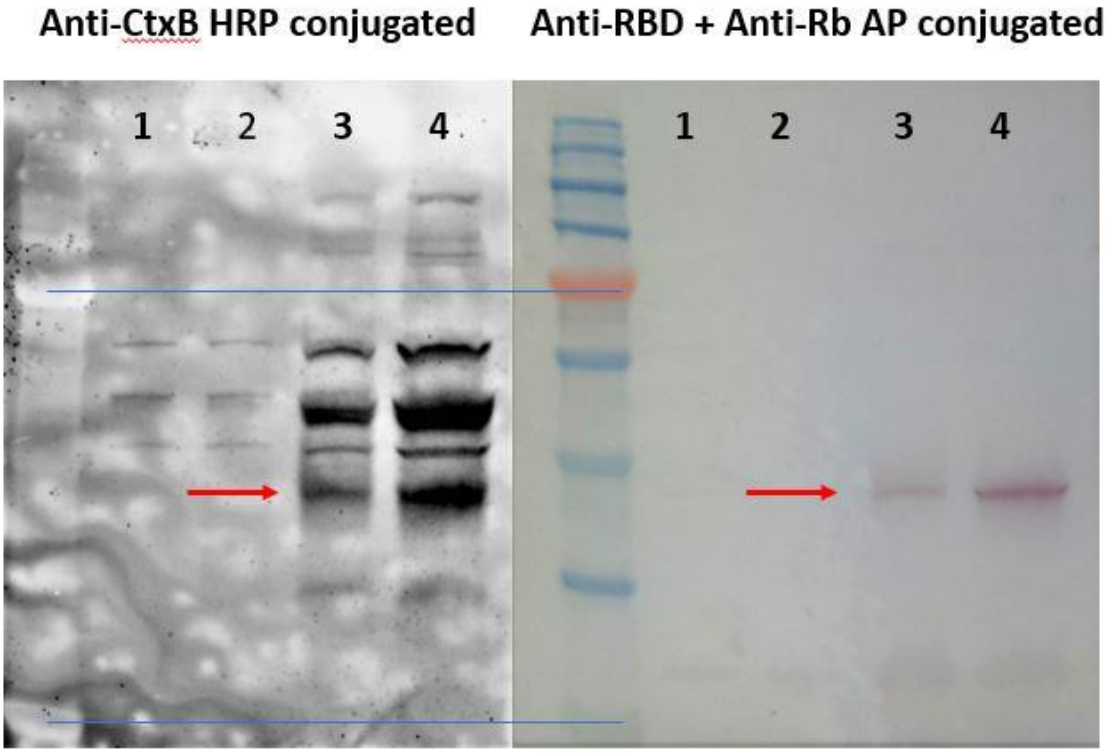

Figure A.3. Recombinant vaccine expression analysis through western blot. Western blot incubated with in the left panel anti-CtxB HRP conjugated. Lane 1 contains wild-type cell lysate, lane 2 contains a clone recovered from transforming with GV-1, lane 3 contains cell lysate of the top expressor of GV-2 and lane 4 contains cell lysate of the only GV-3 clone recovered. Image obtained from ChemiDoc Imaging System. On the right panel, same blot was stripped from antibody using stripping solution. Then it was re-incubated using anti-RBD and developed using Alkaline Phosphatase colorimetric method and imaged with a ChemiDoc Imaging System. Blue lanes indicate where the ladder is in the HRP blot, red arrows indicate the presence of our recombinant protein of interest.

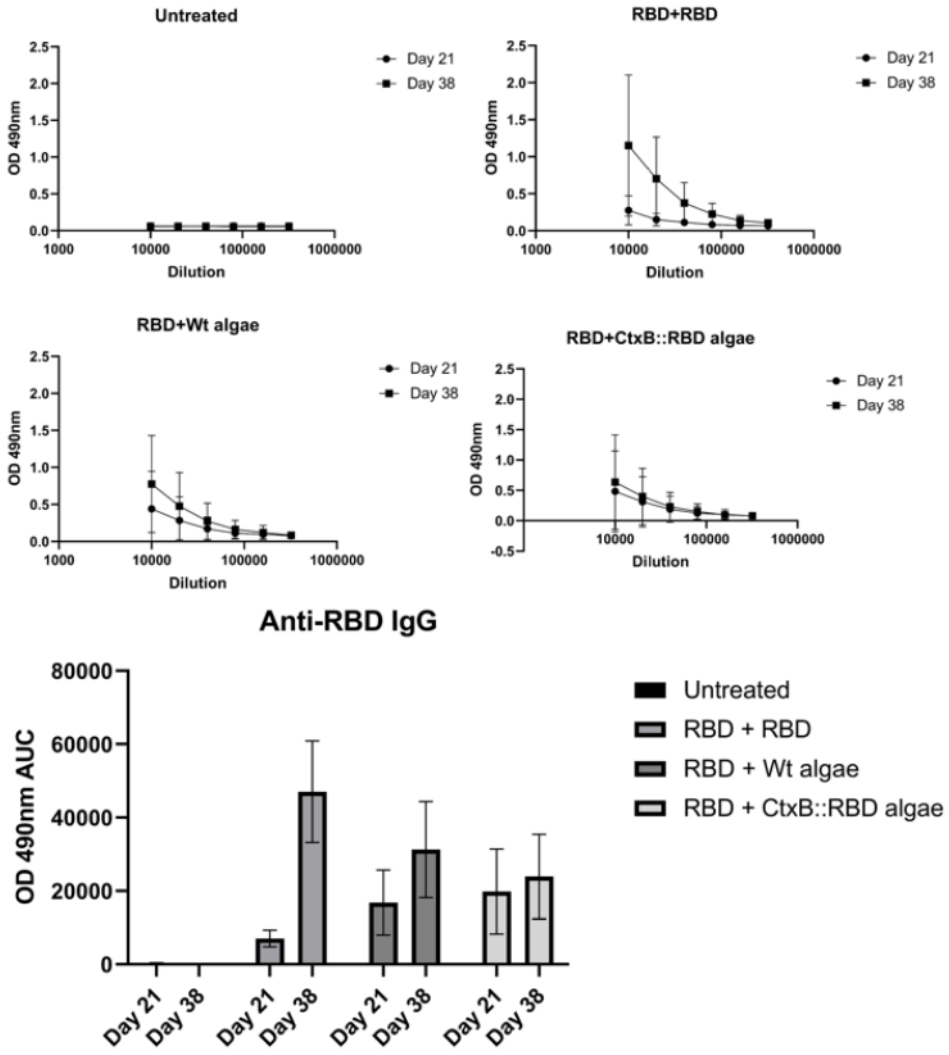

Figure A.4. Indirect ELISA performed on sera from mice previously immunized with different treatments. Untreated mice were injected with serological saline solution, RBD+RBD mice were injected with RBD in CFA on day 1 and on day 24, RBD+Wt algae mice were injected with RBD in CFA on day 1 and fed wild-type algae through oral gavage on day 24, RBD+CtxB::RBD algae mice were injected with RBD in CFA on day 1 and fed CtxB::RBD algae through oral gavage on day 24. Serum were collected on days 21 and 38, and indirect ELISA were performed on the serum to detect anti-RBD IgG titers in the mice serum. For the Untreated 2 mice were tested, for the RBD+RBD 3 mice were tested, for the RBD+CtxB::RBD algae and RBD+Wt algae 6 mice each were tested. The antibody used to detect mice IgG was HRP conjugated and the results were measured in a microplate reader as optical density at 490 nm. Results of those ELISA are shown on the top panels, values shown are average with standard deviations. Bottom panel shows the area under the curve as calculated on Prism9.

References

Aaronson, S. (1986). A role for algae as human food in antiquity. *Food and Foodways, 1*(3), 311-315. doi:https://doi.org/10.1080/07409710.1986.9961891

Acién, F. G., Molina, E., Reis, A., Torzillo, G., Zittelli, G. C., Sepúlveda, C., & Masojídek, J. (2017). 1 - Photobioreactors for the production of microalgae. In C. Gonzalez-Fernandez & R. Muñoz (Eds.), *Microalgae-Based Biofuels and Bioproducts* (pp. 1-44): Woodhead Publishing.

Administration, U. S. F. a. D. (2018). Generaly Recognized as Safe. Retrieved from https://www.fda.gov/Food/IngredientsPackagingLabeling/GRAS/

Agrawal, G. K., Jwa, N. S., Lebrun, M. H., Job, D., & Rakwal, R. (2010). Plant secretome: unlocking secrets of the secreted proteins. *Proteomics, 10*(4), 799-827.

Agriculture, U. S. D. o. (2018). USDA Food Composition Databases. Retrieved 6/26/2018 https://ndb.nal.usda.gov/ndb/search/list?home=true

Alexandersson, E., Ashfaq, A., Resjö, S., & Andreasson, E. (2013). Plant secretome proteomics. *Frontiers in plant science, 4*, 9.

Amanat, F., Stadlbauer, D., Strohmeier, S., Nguyen, T. H. O., Chromikova, V., McMahon, M., Jiang, K., Arunkumar, G. A., Jurczyszak, D., Polanco, J., Bermudez-Gonzalez, M., Kleiner, G., Aydillo, T., Miorin, L., Fierer, D. S., Lugo, L. A., Kojic, E. M., Stoever, J., Liu, S. T. H., Cunningham-Rundles, C., Felgner, P. L., Moran, T., García-Sastre, A., Caplivski, D., Cheng, A. C., Kedzierska, K., Vapalahti, O., Hepojoki, J. M., Simon, V., & Krammer, F. (2020). A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nature Medicine, 26*(7), 1033-1036. doi:10.1038/s41591-020-0913-5

Ami, B. A., Adriana, K., & Mordhay, A. (1982). Accumulation of β-carotene in halotolerant algae: Purification and characterization of β-carotene-rich globules from Dunaliella Bardawil(Chlorophyceae). *Journal of Phycology, 18*(4), 529-537. doi:https://doi.org/10.1111/j.1529-8817.1982.tb03219.x

Anderson, M. S., Muff, T. J., Georgianna, D. R., & Mayfield, S. P. (2017). Towards a synthetic nuclear transcription system in green algae: Characterization of Chlamydomonas reinhardtii nuclear transcription factors and identification of targeted promoters. *Algal Research, 22*, 47-55. doi:https://doi.org/10.1016/j.algal.2016.12.002

Angstenberger, M., de Signori, F., Vecchi, V., Dall'Osto, L., & Bassi, R. (2020). Cell Synchronization Enhances Nuclear Transformation and Genome Editing via Cas9 Enabling Homologous Recombination in Chlamydomonas reinhardtii. *ACS Synthetic Biology, 9*(10), 2840-2850. doi:10.1021/acssynbio.0c00390

Anila, N., Chandrashekar, A., Ravishankar, G. A., & Sarada, R. (2011). Establishment of Agrobacterium tumefaciens-mediated genetic transformation in Dunaliella bardawil. *European Journal of Phycology, 46*(1), 36-44. doi:https://doi.org/10.1080/09670262.2010.550386

Apt, K. E., Grossman, A. R., & Kroth-Pancic, P. G. (1996). Stable nuclear transformation of the diatom Phaeodactylum tricornutum. *Molecular and General Genetics MGG, 252*(5), 572-579. doi:https://doi.org/10.1007/bf02172403

Arisa, M. Y., Takuma, O., Nozomu, K., Haruo, Y., Kouhei, O., & Masao, A. (2011). Stable nuclear transformation of the diatom Chaetoceros sp. *Phycological Research, 59*(2), 113-119. doi:https://doi.org/10.1111/j.1440-1835.2011.00607.x

Authority, E. F. S. (2020a). EU Novel food catalogue (v.1.1).

Authority, E. F. S. (2020b). European Food Safety Authority

Baier, T., Jacobebbinghaus, N., Einhaus, A., Lauersen, K. J., & Kruse, O. (2020). Introns mediate post-transcriptional enhancement of nuclear gene expression in the green microalga Chlamydomonas reinhardtii. *PLOS Genetics, 16*(7), e1008944. doi:10.1371/journal.pgen.1008944

Baier, T., Wichmann, J., Kruse, O., & Lauersen, K. J. (2018). Intron-containing algal transgenes mediate efficient recombinant gene expression in the green microalga Chlamydomonas reinhardtii. *Nucleic Acids Research, 46*(13), 6909-6919. doi:10.1093/nar/gky532

Barrera, D. J. M., Stephen P. (2013). High-value Recombinant Protein Production in Microalgae. In R. a. Q. Hu (Ed.), *Handbook of Microalgal Culture*.

Barsanti, L., Vismara, R., Passarelli, V., & Gualtieri, P. (2001). Paramylon (β-1,3-glucan) content in wild type and WZSL mutant of Euglena gracilis. Effects of growth conditions. *Journal of Applied Phycology, 13*(1), 59-65. doi:https://doi.org/10.1023/a:1008105416065

Beacham, T. A., Macia, V. M., Rooks, P., White, D. A., & Ali, S. T. (2015). Altered lipid accumulation in Nannochloropsis salina CCAP849/3 following EMS and UV induced mutagenesis. *Biotechnology Reports, 7*, 87-94. doi:https://doi.org/10.1016/j.btre.2015.05.007

Becker, E. W. (2007). Micro-algae as a source of protein. *Biotechnology Advances, 25*(2), 207-210. doi:https://doi.org/10.1016/j.biotechadv.2006.11.002

Bella, J., Kolatkar, P. R., Marlor, C. W., Greve, J. M., & Rossmann, M. G. (1998). The structure of the two amino-terminal domains of human ICAM-1 suggests how it functions as a rhinovirus receptor and as an LFA-1 integrin ligand. *Proceedings of the National Academy of Sciences, 95*(8), 4140-4145. doi:10.1073/pnas.95.8.4140

Berndt, A., Smalley, T., Ren, B., Badary, A., Sproles, A., Fields, F., Torres-Tiji, Y., Heredia, V., & Mayfield, S. (2021). Recombinant production of a functional SARS-CoV-2 spike receptor binding domain in the green algae Chlamydomonas reinhardtii. *bioRxiv*, 2021.2001.2029.428890. doi:10.1101/2021.01.29.428890

Bevis, B. J., & Glick, B. S. (2002). Rapidly maturing variants of the Discosoma red fluorescent protein (DsRed). *Nature Biotechnology, 20*(1), 83-87. doi:10.1038/nbt0102-83

Binnal, P., & Nirguna Babu, P. (2017). Enhancement of lipid productivity of Chlorella protothecoides cultivated in secondary wastewater under nitrogen starvation through optimization of environmental factors. *Biofuels*, 1-16. doi:https://doi.org/10.1080/17597269.2017.1336351

Black, R. E., Morris, S. S., & Bryce, J. (2003). Where and why are 10 million children dying every year? *The Lancet, 361*(9376), 2226-2234. doi:https://doi.org/10.1016/S0140-6736(03)13779-8

Bo, Y., Jin, L., Yue, J., & Feng, C. (2016). Chlorella species as hosts for genetic engineering and expression of heterologous proteins: Progress, challenge and perspective. *Biotechnology Journal, 11*(10), 1244-1261. doi:https://doi.org/10.1002/biot.201500617

Bonnefoy, N., Remacle, C., & Fox, T. D. (2007). Genetic Transformation of Saccharomyces cerevisiae and Chlamydomonas reinhardtii Mitochondria. In *Methods in Cell Biology* (Vol. 80, pp. 525-548): Academic Press.

Borowitzka, M. A. (1999a). Commercial production of microalgae: ponds, tanks, and fermenters. In R. Osinga, J. Tramper, J. G. Burgess, & R. H. Wijffels (Eds.), *Progress in Industrial Microbiology* (Vol. 35, pp. 313-321): Elsevier.

Borowitzka, M. A. (1999b). Commercial production of microalgae: ponds, tanks, tubes and fermenters. *Journal of Biotechnology, 70*(1), 313-321. doi:https://doi.org/10.1016/S0168-1656(99)00083-8

Borowitzka, M. A., & Moheimani, N. R. (2013). Open Pond Culture Systems. In M. A. Borowitzka & N. R. Moheimani (Eds.), *Algae for Biofuels and Energy* (pp. 133-152). Dordrecht: Springer Netherlands.

Brennan, L., & 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*(2), 557-577. doi:https://doi.org/10.1016/j.rser.2009.10.009

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004). TOWARD A METABOLIC THEORY OF ECOLOGY. *Ecology, 85*(7), 1771-1789. doi:https://doi.org/10.1890/03-9000

Browning, D. F., Godfrey, R. E., Richards, K. L., Robinson, C., & Busby, S. J. W. (2019). Exploitation of the Escherichia coli lac operon promoter for controlled recombinant protein production. *Biochemical Society Transactions, 47*(2), 755-763. doi:10.1042/bst20190059

Bumbak, F., Cook, S., Zachleder, V., Hauser, S., & Kovar, K. (2011). Best practices in heterotrophic high-cell-density microalgal processes: achievements, potential and possible limitations. *Applied Microbiology and Biotechnology, 91*(1), 31-46. doi:10.1007/s00253-011-3311-6

Burki, T. K. (2021). Challenges in the rollout of COVID-19 vaccines worldwide. *The Lancet Respiratory Medicine, 9*(4), e42-e43. doi:10.1016/S2213-2600(21)00129-6

Caballero, M., & Quirce, S. (2021). Excipients as potential agents of anaphylaxis in vaccines: analyzing the formulations of the current authorized COVID-19 vaccines. *J Investig Allergol Clin Immunol, 31*(1), 92-93.

Calkins, C. R., & Hodgen, J. M. (2007). A fresh look at meat flavor. *Meat Science, 77*(1), 63-80. doi:https://doi.org/10.1016/j.meatsci.2007.04.016

Canada, H. (2019). Natural Health Products Ingredients Database.

Canada, H. (2020). About novel and genetically-modified (GM) foods.

Caporgno, M. P., & Mathys, A. (2018). Trends in Microalgae Incorporation Into Innovative Food Products With Potential Health Benefits. *Frontiers in Nutrition, 5*(58). doi:10.3389/fnut.2018.00058

Cardozo, K. H. M., Guaratini, T., Barros, M. P., Falcão, V. R., Tonon, A. P., Lopes, N. P., Campos, S., Torres, M. A., Souza, A. O., Colepicolo, P., & Pinto, E. (2007). Metabolites from algae with economical impact. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology, 146*(1), 60-78. doi:https://doi.org/10.1016/j.cbpc.2006.05.007

Carrera Pacheco, S. E., Hankamer, B., & Oey, M. (2018). Optimising light conditions increases recombinant protein production in Chlamydomonas reinhardtii chloroplasts. *Algal Research, 32*, 329-340. doi:https://doi.org/10.1016/j.algal.2018.04.011

Chae, S. R., Hwang, E. J., & Shin, H. S. (2006). Single cell protein production of Euglena gracilis and carbon dioxide fixation in an innovative photo-bioreactor. *Bioresource Technology, 97*(2), 322-329. doi:https://doi.org/10.1016/j.biortech.2005.02.037

Chapman, R. L. (2013). Algae: the world's most important "plants"—an introduction. *Mitigation and Adaptation Strategies for Global Change, 18*(1), 5-12. doi:10.1007/s11027-010-9255-9

Chen, F., & Johns, M. R. (1995). A strategy for high cell density culture of heterotrophic microalgae with inhibitory substrates. *Journal of Applied Phycology, 7*(1), 43-46. doi:10.1007/BF00003548

Chen, F., & Johns, M. R. (1996). Heterotrophic growth of Chlamydomonas reinhardtii on acetate in chemostat culture. *Process Biochemistry, 31*(6), 601-604. doi:https://doi.org/10.1016/S0032-9592(96)00006-4

Cheng, K.-C., Ren, M., & Ogden, K. L. (2013). Statistical optimization of culture media for growth and lipid production of Chlorella protothecoides UTEX 250. *Bioresource Technology, 128*, 44-48. doi:https://doi.org/10.1016/j.biortech.2012.09.085

Chow, K.-C., & Tung, W. L. (1999). Electrotransformation of Chlorella vulgaris. *Plant Cell Reports, 18*(9), 778-780. doi:https://doi.org/10.1007/s002990050660

Cordero, B. F., Couso, I., León, R., Rodríguez, H., & Vargas, M. Á. (2011). Enhancement of carotenoids biosynthesis in Chlamydomonas reinhardtii by nuclear transformation using a phytoene synthase gene isolated from Chlorella zofingiensis. *Applied Microbiology and Biotechnology, 91*(2), 341-351. doi:https://doi.org/10.1007/s00253-011-3262-y

Cucinotta, D., & Vanelli, M. (2020). WHO Declares COVID-19 a Pandemic. *Acta Biomedica Atenei Parmensis, 91*(1), 157-160. doi:10.23750/abm.v91i1.9397

Davis, R., Aden, A., & Pienkos, P. T. (2011). Techno-economic analysis of autotrophic microalgae for fuel production. *Applied Energy, 88*(10), 3524-3531. doi:https://doi.org/10.1016/j.apenergy.2011.04.018

Davis, R., Markham, J., Kinchin, C., Grundl, N., Tan, E. C., & Humbird, D. (2016). *Process Design and Economics for the Production of Algal Biomass: Algal Biomass Production in Open Pond*. Retrieved from

Day, J. G., Gong, Y., & Hu, Q. (2017). Microzooplanktonic grazers – A potentially devastating threat to the commercial success of microalgal mass culture. *Algal Research, 27*, 356-365. doi:https://doi.org/10.1016/j.algal.2017.08.024

Demirbas, A. (2009). Progress and recent trends in biodiesel fuels. *Energy Conversion and Management, 50*(1), 14-34. doi:https://doi.org/10.1016/j.enconman.2008.09.001

Dillehay, T. D., Ramírez, C., Pino, M., Collins, M. B., Rossen, J., & Pino-Navarro, J. D. (2008). Monte Verde: Seaweed, Food, Medicine, and the Peopling of South America. *Science, 320*(5877), 784-786. doi:https://doi.org/10.1126/science.1156533

Dismukes, G. C., Carrieri, D., Bennette, N., Ananyev, G. M., & Posewitz, M. C. (2008). Aquatic phototrophs: efficient alternatives to land-based crops for biofuels. *Current Opinion in Biotechnology, 19*(3), 235-240. doi:https://doi.org/10.1016/j.copbio.2008.05.007

Dodd, M. S., Papineau, D., Grenne, T., Slack, J. F., Rittner, M., Pirajno, F., O'Neil, J., & Little, C. T. S. (2017). Evidence for early life in Earth's oldest hydrothermal vent precipitates. *Nature, 543*(7643), 60-64. doi:10.1038/nature21377

Doetsch, N. A., Favreau, M. R., Kuscuoglu, N., Thompson, M. D., & Hallick, R. B. (2001). Chloroplast transformation in Euglena gracilis: splicing of a group III twintron transcribed from a transgenic psbK operon. *Current Genetics, 39*(1), 49-60. doi:https://doi.org/10.1007/s002940000174

Draaisma, R. B., Wijffels, R. H., Slegers, P. M., Brentner, L. B., Roy, A., & Barbosa, M. J. (2013). Food commodities from microalgae. *Current Opinion in Biotechnology, 24*(2), 169-177. doi:https://doi.org/10.1016/j.copbio.2012.09.012

Dreesen, I. A. J., Hamri, G. C.-E., & Fussenegger, M. (2010). Heat-stable oral alga-based vaccine protects mice from Staphylococcus aureus infection. *Journal of Biotechnology, 145*(3), 273-280. doi:https://doi.org/10.1016/j.jbiotec.2009.12.006

El Khoury, D., Cuda, C., Luhovyy, B. L., & Anderson, G. H. (2012). Beta Glucan: Health Benefits in Obesity and Metabolic Syndrome. *Journal of Nutrition & Metabolism*, 1-28. doi:https://doi.org/10.1155/2012/851362

Elliott, D. A., & Brand, A. H. (2008). The GAL4 System. In C. Dahmann (Ed.), *Drosophila: Methods and Protocols* (pp. 79-95). Totowa, NJ: Humana Press.

F. Gibbs, S. K. I. A. C. N. M. B. (1999). Encapsulation in the food industry: a review. *International Journal of Food Sciences and Nutrition, 50*(3), 213-224. doi:https://doi.org/10.1080/096374899101256

Fabregas, J., & Herrero, C. (1990). Vitamin content of four marine microalgae. Potential use as source of vitamins in nutrition. *Journal of Industrial Microbiology, 5*(4), 259-263. doi:https://doi.org/10.1007/bf01569683

Falciatore, A., Casotti, R., Leblanc, C., Abrescia, C., & Bowler, C. (1999). Transformation of Nonselectable Reporter Genes in Marine Diatoms. *Marine Biotechnology, 1*(3), 239-251. doi:https://doi.org/10.1007/pl00011773

Fallah, A. A., Sarmast, E., Habibian Dehkordi, S., Engardeh, J., Mahmoodnia, L., Khaledifar, A., & Jafari, T. (2018). Effect of Chlorella supplementation on cardiovascular risk factors: A meta-analysis of randomized controlled trials. *Clinical Nutrition, 37*(6, Part A), 1892-1901. doi:https://doi.org/10.1016/j.clnu.2017.09.019

FAO, I., UNICEF, WFP and WHO. (2017). *The State of Food Security and Nutrition in the World 2017. Building resilience for peace and food security.* Retrieved from Rome:

Feng, S., Li, X., Xu, Z., & Qi, J. (2014). Dunaliella salina as a novel host for the production of recombinant proteins. *Applied Microbiology and Biotechnology, 98*(10), 4293-4300. doi:https://doi.org/10.1007/s00253-014-5636-4

Feng, X., Walker, T. H., Bridges, W. C., Thornton, C., & Gopalakrishnan, K. (2014). Biomass and lipid production of Chlorella protothecoides under heterotrophic cultivation on a mixed waste substrate of brewer fermentation and crude glycerol. *Bioresource Technology, 166*, 17-23. doi:https://doi.org/10.1016/j.biortech.2014.03.120

Fields, F. J., Hernandez, R. E., Weilbacher, E., Garcia-Vargas, E., Huynh, J., Thurmond, M., Lund, R., Burkart, M. D., & Mayfield, S. P. (2021). Annual productivity and lipid composition of native microalgae (Chlorophyta) at a pilot production facility in Southern California. *Algal Research, 56*, 102307. doi:https://doi.org/10.1016/j.algal.2021.102307

Fields, F. J., Lejzerowicz, F., Schroeder, D., Ngoi, S. M., Tran, M., McDonald, D., Jiang, L., Chang, J. T., Knight, R., & Mayfield, S. (2020). Effects of the microalgae Chlamydomonas on gastrointestinal health. *Journal of Functional Foods, 65*, 103738. doi:https://doi.org/10.1016/j.jff.2019.103738

Fields, F. J., Ostrand, J. T., & Mayfield, S. P. (2018). Fed-batch mixotrophic cultivation of Chlamydomonas reinhardtii for high-density cultures. *Algal Research, 33*, 109-117. doi:https://doi.org/10.1016/j.algal.2018.05.006

Fields, F. J., Ostrand, J. T., Tran, M., & Mayfield, S. P. (2019). Nuclear genome shuffling significantly increases production of chloroplast-based recombinant protein in Chlamydomonas reinhardtii. *Algal Research, 41*, 101523. doi:https://doi.org/10.1016/j.algal.2019.101523

Fischer, J. A., Giniger, E., Maniatis, T., & Ptashne, M. (1988). GAL4 activates transcription in Drosophila. *Nature, 332*(6167), 853-856. doi:10.1038/332853a0

Freudenberg, R. A., Baier, T., Einhaus, A., Wobbe, L., & Kruse, O. (2021). High cell density cultivation enables efficient and sustainable recombinant polyamine production in the microalga Chlamydomonas reinhardtii. *Bioresource Technology, 323*, 124542. doi:https://doi.org/10.1016/j.biortech.2020.124542

Friedman, M., & Brandon, D. L. (2001). Nutritional and Health Benefits of Soy Proteins. *Journal of Agricultural and Food Chemistry, 49*(3), 1069-1086. doi:10.1021/jf0009246

Fróna, D., Szenderák, J., & Harangi-Rákos, M. (2019). The Challenge of Feeding the World. *Sustainability, 11*(20), 5816.

Garcia-Ochoa, F., & Gomez, E. (2009). Bioreactor scale-up and oxygen transfer rate in microbial processes: An overview. *Biotechnology Advances, 27*(2), 153-176. doi:https://doi.org/10.1016/j.biotechadv.2008.10.006

Georgianna, D. R., & Mayfield, S. P. (2012). Exploiting diversity and synthetic biology for the production of algal biofuels. *Nature, 488*(7411), 329-335. doi:10.1038/nature11479

Gong, M., & Bassi, A. (2016). Carotenoids from microalgae: A review of recent developments. *Biotechnology Advances, 34*(8), 1396-1412. doi:https://doi.org/10.1016/j.biotechadv.2016.10.005

Gopal, G. J., & Kumar, A. (2013). Strategies for the Production of Recombinant Protein in Escherichia coli. *The Protein Journal, 32*(6), 419-425. doi:10.1007/s10930-013-9502-5

Gorissen, S. H. M., Crombag, J. J. R., Senden, J. M. G., Waterval, W. A. H., Bierau, J., Verdijk, L. B., & van Loon, L. J. C. (2018). Protein content and amino acid composition of commercially available plant-based protein isolates. *Amino acids, 50*(12), 1685-1695. doi:10.1007/s00726-018-2640-5

Gregory, J. A., Topol, A. B., Doerner, D. Z., & Mayfield, S. (2013). Alga-Produced Cholera Toxin-Pfs25 Fusion Proteins as Oral Vaccines. *Applied and Environmental Microbiology, 79*(13), 3917-3925. doi:doi:10.1128/AEM.00714-13

Griffiths, M. J., & Harrison, S. T. L. (2009). Lipid productivity as a key characteristic for choosing algal species for biodiesel production. *Journal of Applied Phycology, 21*(5), 493-507. doi:https://doi.org/10.1007/s10811-008-9392-7

Guccione, A., Biondi, N., Sampietro, G., Rodolfi, L., Bassi, N., & Tredici, M. R. (2014). Chlorella for protein and biofuels: from strain selection to outdoor cultivation in a Green Wall Panel photobioreactor. *Biotechnology for Biofuels, 7*(1), 84. doi:https://doi.org/10.1186/1754-6834-7-84

Gumpel, N. J., Rochaix, J.-D., & Purton, S. (1994). Studies on homologous recombination in the green alga Chlamydomonas reinhardtii. *Current Genetics, 26*(5), 438-442. doi:10.1007/BF00309931

Guttman, M., Betts, G. N., Barnes, H., Ghassemian, M., van der Geer, P., & Komives, E. A. (2009). Interactions of the NPXY microdomains of the low density lipoprotein receptor-related protein 1. *Proteomics, 9*(22), 5016-5028. doi:https://doi.org/10.1002/pmic.200900457

Habib, M. A. B. P., M.; Huntington, T.C.; Hasan, M.R. (2008). *A review on culture, production and use of spirulina as food for humans and feeds for domestic animals and fish.* Retrieved from Rome:

Hannah Ritchie, E. M., Lucas Rodés-Guirao, Cameron Appel, Charlie Giattino, Esteban Ortiz-Ospina, Joe Hasell, Bobbie Macdonald, Diana Beltekian and Max Roser. (2020). Coronavirus Pandemic (COVID-19). Retrieved from https://ourworldindata.org/coronavirus

Hannon, M., Gimpel, J., Tran, M., Rasala, B., & Mayfield, S. (2010). Biofuels from algae: challenges and potential. *Biofuels, 1*(5), 763-784.

Harris, E. H. (2001). CHLAMYDOMONAS AS A MODEL ORGANISM. *Annual Review of Plant Physiology and Plant Molecular Biology, 52*(1), 363-406. doi:10.1146/annurev.arplant.52.1.363

Hayashi, K., Yamanaka, M., Toko, K., & Yamafuji, K. (1990). Multichannel taste sensor using lipid membranes. *Sensors and Actuators B: Chemical, 2*(3), 205-213. doi:https://doi.org/10.1016/0925-4005(90)85006-K

Huang, W., & Daboussi, F. (2017). Genetic and metabolic engineering in diatoms. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*(1728). doi:https://doi.org/10.1098/rstb.2016.0411

India, F. S. a. S. A. o. (2016). The Gazette of India.

Jia, Y., Li, S., Allen, G., Feng, S., & Xue, L. (2012). A Novel Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) Promoter for Expressing Transgenes in the Halotolerant Alga Dunaliella salina. *Current Microbiology, 64*(5), 506-513. doi:https://doi.org/10.1007/s00284-012-0102-y

Jiang, R., & Lönnerdal, B. (2016). Biological roles of milk osteopontin. *Current Opinion in Clinical Nutrition & Metabolic Care, 19*(3), 214-219. doi:https://doi.org/10.1097/mco.0000000000000275

Jin, H., Chuai, W., Li, K., Hou, G., Wu, M., Chen, J., Wang, H., Jia, J., Han, D., & Hu, Q. (2021). Ultrahigh-cell-density heterotrophic cultivation of the unicellular green alga Chlorella sorokiniana for biomass production. *Biotechnology and Bioengineering, n/a*(n/a). doi:https://doi.org/10.1002/bit.27890

Johnsborg, O., Eldholm, V., & Håvarstein, L. S. (2007). Natural genetic transformation: prevalence, mechanisms and function. *Research in Microbiology, 158*(10), 767-778. doi:https://doi.org/10.1016/j.resmic.2007.09.004

Jongedijk, E., Cankar, K., Buchhaupt, M., Schrader, J., Bouwmeester, H., & Beekwilder, J. (2016). Biotechnological production of limonene in microorganisms. *Applied Microbiology and Biotechnology, 100*(7), 2927-2938. doi:10.1007/s00253-016-7337-7

Jorde, F., Leya, T., Thomas, R., Pereira, S., M Badenes, S., Santos, E., Costa, L., Verdelho, V., Friedl, T., & Kryvenda, A. (2017). Algae Crop Rotation (ACR). In.

Joy, J. M., Lowery, R. P., Wilson, J. M., Purpura, M., De Souza, E. O., Wilson, S. M. C., Kalman, D. S., Dudeck, J. E., & Jäger, R. (2013). The effects of 8 weeks of whey or rice protein supplementation on body composition and exercise performance. *Nutrition Journal, 12*(1), 86. doi:10.1186/1475-2891-12-86

Kakidani, H., & Ptashne, M. (1988). GAL4 activates gene expression in mammalian cells. *Cell, 52*(2), 161-167. doi:https://doi.org/10.1016/0092-8674(88)90504-1

Kanaga, K., Pandey, A., Kumar, S., & Geetanjali. (2016). Multi-objective optimization of media nutrients for enhanced production of algae biomass and fatty acid biosynthesis from Chlorella pyrenoidosa NCIM 2738. *Bioresource Technology, 200*, 940-950. doi:https://doi.org/10.1016/j.biortech.2015.11.017

Kant, R. (2005). Sweet proteins – Potential replacement for artificial low calorie sweeteners. *Nutrition Journal, 4*(1), 5. doi:10.1186/1475-2891-4-5

Kawata, Y., Yano, S. i., Kojima, H., & Toyomizu, M. (2004). Transformation of Spirulina platensis Strain C1 (Arthrospira sp. PCC9438) with Tn5 Transposase–Transposon DNA–Cation Liposome Complex. *Marine Biotechnology, 6*(4), 355-363. doi:https://doi.org/10.1007/s10126-003-0037-1

Kilian, O., Benemann, C. S. E., Niyogi, K. K., & Vick, B. (2011). High-efficiency homologous recombination in the oil-producing alga *Nannochloropsis sp*. *Proceedings of the National Academy of Sciences, 108*(52), 21265-21269. doi:https://doi.org/10.1073/pnas.1105861108

Kindle, K. L. (1990). High-frequency nuclear transformation of Chlamydomonas reinhardtii. *Proceedings of the National Academy of Sciences, 87*(3), 1228-1232. doi:10.1073/pnas.87.3.1228

Kindle, K. L., Richards, K. L., & Stern, D. B. (1991). Engineering the chloroplast genome: techniques and capabilities for chloroplast transformation in Chlamydomonas reinhardtii. *Proceedings of the National Academy of Sciences, 88*(5), 1721-1725. doi:10.1073/pnas.88.5.1721

Koller, M., Muhr, A., & Braunegg, G. (2014). Microalgae as versatile cellular factories for valued products. *Algal Research, 6*, 52-63. doi:https://doi.org/10.1016/j.algal.2014.09.002

Kropat, J., Hong-Hermesdorf, A., Casero, D., Ent, P., Castruita, M., Pellegrini, M., Merchant, S. S., & Malasarn, D. (2011). A revised mineral nutrient supplement increases biomass and growth rate in Chlamydomonas reinhardtii. *The Plant Journal, 66*(5), 770-780. doi:https://doi.org/10.1111/j.1365-313X.2011.04537.x

Kumar, K., Mishra, S. K., Shrivastav, A., Park, M. S., & Yang, J.-W. (2015). Recent trends in the mass cultivation of algae in raceway ponds. *Renewable and Sustainable Energy Reviews, 51*, 875-885. doi:https://doi.org/10.1016/j.rser.2015.06.033

Kurihara, K. (2009). Glutamate: from discovery as a food flavor to role as a basic taste (umami). *The American Journal of Clinical Nutrition, 90*(3), 719S-722S. doi:https://doi.org/10.3945/ajcn.2009.27462D

L., M. A., Verónica, B. M., H., E. J., T., S. D., Miller, T., Annika, W., L., M. T., & P., M. S. (2007). Robust expression of a bioactive mammalian protein in Chlamydomonas chloroplast. *Plant Biotechnology Journal, 5*(3), 402-412. doi:https://doi.org/10.1111/j.1467-7652.2007.00249.x

Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., & Wang, X. (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature, 581*(7807), 215-220. doi:10.1038/s41586-020-2180-5

Laplaze, L., Parizot, B., Baker, A., Ricaud, L., Martinière, A., Auguy, F., Franche, C., Nussaume, L., Bogusz, D., & Haseloff, J. (2005). GAL4-GFP enhancer trap lines for genetic manipulation of lateral root development in Arabidopsis thaliana. *Journal of Experimental Botany, 56*(419), 2433-2442. doi:10.1093/jxb/eri236

Lardon, L., Hélias, A., Sialve, B., Steyer, J.-P., & Bernard, O. (2009). Life-Cycle Assessment of Biodiesel Production from Microalgae. *Environmental Science & Technology, 43*(17), 6475-6481. doi:10.1021/es900705j

Larson, M. A., Wei, S. H., Weber, A., Mack, D. R., & McDonald, T. L. (2003). Human serum amyloid A3 peptide enhances intestinal MUC3 expression and inhibits EPEC adherence. *Biochemical and Biophysical Research Communications, 300*(2), 531-540. doi:https://doi.org/10.1016/S0006-291X(02)02901-7

Lau, A. Y., Lee, L. P., & Chan, J. W. (2008). An integrated optofluidic platform for Raman-activated cell sorting. *Lab on a Chip, 8*(7), 1116-1120. doi:https://doi.org/10.1039/B803598A

Lauersen, K. J., Huber, I., Wichmann, J., Baier, T., Leiter, A., Gaukel, V., Kartushin, V., Rattenholl, A., Steinweg, C., von Riesen, L., Posten, C., Gudermann, F., Lütkemeyer, D., Mussgnug, J. H., & Kruse, O. (2015a). Investigating the dynamics of recombinant protein secretion from a microalgal host. *Journal of Biotechnology, 215*, 62-71. doi:https://doi.org/10.1016/j.jbiotec.2015.05.001

Lauersen, K. J., Kruse, O., & Mussgnug, J. H. (2015b). Targeted expression of nuclear transgenes in Chlamydomonas reinhardtii with a versatile, modular vector toolkit. *Applied Microbiology and Biotechnology, 99*(8), 3491-3503. doi:10.1007/s00253-014-6354-7

Lee, D. Y., Park, J.-J., Barupal, D. K., & Fiehn, O. (2012). System Response of Metabolic Networks in Chlamydomonas reinhardtii to Total Available Ammonium. *Molecular & Cellular Proteomics, 11*(10), 973-988. doi:10.1074/mcp.M111.016733

Lee, Y.-K. (2001). Microalgal mass culture systems and methods: Their limitation and potential. *Journal of Applied Phycology, 13*(4), 307-315. doi:https://doi.org/10.1023/a:1017560006941

Li, M., Wang, J., Geng, Y., Li, Y., Wang, Q., Liang, Q., & Qi, Q. (2012). A strategy of gene overexpression based on tandem repetitive promoters in Escherichia coli. *Microbial Cell Factories, 11*(1), 19. doi:10.1186/1475-2859-11-19

Liang, Y., Sarkany, N., & Cui, Y. (2009). Biomass and lipid productivities of Chlorella vulgaris under autotrophic, heterotrophic and mixotrophic growth conditions. *Biotechnology Letters, 31*(7), 1043-1049. doi:https://doi.org/10.1007/s10529-009-9975-7

Liu, L., Johnson, H. L., Cousens, S., Perin, J., Scott, S., Lawn, J. E., Rudan, I., Campbell, H., Cibulskis, R., Li, M., Mathers, C., & Black, R. E. (2012). Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. *The Lancet, 379*(9832), 2151-2161. doi:https://doi.org/10.1016/S0140-6736(12)60560-1

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods, 25*(4), 402-408. doi:https://doi.org/10.1006/meth.2001.1262

Livingston, E. H., Malani, P. N., & Creech, C. B. (2021). The Johnson & Johnson Vaccine for COVID-19. *JAMA, 325*(15), 1575-1575. doi:10.1001/jama.2021.2927

Lönnerdal, B. (2014). Infant formula and infant nutrition: bioactive proteins of human milk and implications for composition of infant formulas. *The American Journal of Clinical Nutrition, 99*(3), 712S-717S. doi:https://doi.org/10.3945/ajcn.113.071993

Lorenz, J. K., Reo, J. P., Hendl, O., Worthington, J. H., & Petrossian, V. D. (2009). Evaluation of a taste sensor instrument (electronic tongue) for use in formulation development. *International Journal of Pharmaceutics, 367*(1), 65-72. doi:https://doi.org/10.1016/j.ijpharm.2008.09.042

Lowe, C. R. (2001). Combinatorial approaches to affinity chromatography. *Current Opinion in Chemical Biology, 5*(3), 248-256.

Lü, J., Sheahan, C., & Fu, P. (2011). Metabolic engineering of algae for fourth generation biofuels production. *Energy & Environmental Science, 4*(7), 2451-2466. doi:https://doi.org/10.1039/C0EE00593B

Lumbreras, V., Stevens, D. R., & Purton, S. (1998). Efficient foreign gene expression in Chlamydomonas reinhardtii mediated by an endogenous intron. *The Plant Journal, 14*(4), 441-447. doi:https://doi.org/10.1046/j.1365-313X.1998.00145.x

M., A. L., A., O. H., P., H. J., Eileen, B. H., Gloria, C., Dror, R., Jacqueline, H., & Deanna, M. (2007). Bioequivalence of Docosahexaenoic Acid from Different Algal Oils in Capsules and in a DHA-Fortified Food. *Lipids, 42*(11), 1011. doi:https://doi.org/10.1007/s11745-007-3098-5

M., L. Q. M., Antonio, L. G. M., Gilberto, V. J., Lenin, S. C., Mauro, D. N., Damar, L. A., & Luis, H. E. (2016). A novel genetic engineering platform for the effective management of biological contaminants for the production of microalgae. *Plant Biotechnology Journal, 14*(10), 2066-2076. doi:https://doi.org/10.1111/pbi.12564

Ma, Q., Zhang, Q., Xu, Q., Zhang, C., Li, Y., Fan, X., Xie, X., & Chen, N. (2017). Systems metabolic engineering strategies for the production of amino acids. *Synthetic and Systems Biotechnology, 2*(2), 87-96. doi:https://doi.org/10.1016/j.synbio.2017.07.003

MacIntyre, C. R., Costantino, V., & Trent, M. (2021). Modelling of COVID-19 vaccination strategies and herd immunity, in scenarios of limited and full vaccine supply in NSW, Australia. *Vaccine*. doi:https://doi.org/10.1016/j.vaccine.2021.04.042

Malavasi, V., Soru, S., & Cao, G. (2020). Extremophile Microalgae: the potential for biotechnological application. *Journal of Phycology, 56*(3), 559-573. doi:https://doi.org/10.1111/jpy.12965

Martorell, R. (1999). The nature of child malnutrition and its long-term implications. *Food and nutrition Bulletin, 20*(3), 288-292. doi:https://doi.org/10.1177/156482659902000304

Matsuda, F., Hayashi, M., & Kondo, A. (2011). Comparative Profiling Analysis of Central Metabolites inEuglena gracilisunder Various Cultivation Conditions. *Bioscience, Biotechnology, and Biochemistry, 75*(11), 2253-2256. doi:https://doi.org/10.1271/bbb.110482

Mayfield, S. P., & Kindle, K. L. (1990). Stable nuclear transformation of Chlamydomonas reinhardtii by using a C. reinhardtii gene as the selectable marker. *Proceedings of the National Academy of Sciences, 87*(6), 2087-2091.

Mayfield, S. P., Manuell, A. L., Chen, S., Wu, J., Tran, M., Siefker, D., Muto, M., & Marin-Navarro, J. (2007). Chlamydomonas reinhardtii chloroplasts as protein factories. *Current Opinion in Biotechnology, 18*(2), 126-133. doi:https://doi.org/10.1016/j.copbio.2007.02.001

McCormack, A. L., Schieltz, D. M., Goode, B., Yang, S., Barnes, G., Drubin, D., & Yates, J. R. (1997). Direct analysis and identification of proteins in mixtures by LC/MS/MS and database searching at the low-femtomole level. *Analytical Chemistry, 69*(4), 767-776.

Mia, T., S., F. E., E., J. R., & C., J. M. (2015). A fluorescence-activated cell sorting-based strategy for rapid isolation of high-lipid Chlamydomonas mutants. *The Plant Journal, 81*(1), 147-159. doi:https://doi.org/10.1111/tpj.12682

Miao, X., & Wu, Q. (2004). High yield bio-oil production from fast pyrolysis by metabolic controlling of Chlorella protothecoides. *Journal of Biotechnology, 110*(1), 85-93. doi:https://doi.org/10.1016/j.jbiotec.2004.01.013

Ming, D., & Hellekant, G. (1994). Brazzein, a new high-potency thermostable sweet protein from Pentadiplandra brazzeana B. *FEBS Letters, 355*(1), 106-108. doi:10.1016/0014-5793(94)01184-2

Miroslav, G., & Zorica, S. (2008). Microalgae and Cyanobacteria: Food for Thought. *Journal of Phycology, 44*(2), 260-268. doi:https://doi.org/10.1111/j.1529-8817.2008.00469.x

Molino, J. V. D., de Carvalho, J. C. M., & Mayfield, S. P. (2018). Comparison of secretory signal peptides for heterologous protein expression in microalgae: Expanding the secretion portfolio for Chlamydomonas reinhardtii. *PLOS ONE, 13*(2), e0192433. doi:https://doi.org/10.1371/journal.pone.0192433

Morris, J. A., & Cagan, R. H. (1972). Purification of monellin, the sweet principle of Dioscoreophyllum cumminsii. *Biochimica et Biophysica Acta (BBA) - General Subjects, 261*(1), 114-122. doi:https://doi.org/10.1016/0304-4165(72)90320-0

Murbach, T. S., Glávits, R., Endres, J. R., Hirka, G., Vértesi, A., Béres, E., & Szakonyiné, I. P. (2018). A Toxicological Evaluation of Chlamydomonas reinhardtii, a Green Algae. *International Journal of Toxicology, 37*(1), 53-62. doi:https://doi.org/10.1177/1091581817746109

National Acadamy of Sciences, N. A. o. E., Institute of Medicine and National Research Council. (2008). *UNDERSTANDING AND RESPONDING TO CLIMATE CHANGE*. Retrieved from

Ni, J., Tao, F., Du, H., & Xu, P. (2015). Mimicking a natural pathway for de novo biosynthesis: natural vanillin production from accessible carbon sources. *Scientific Reports, 5*(1), 13670. doi:10.1038/srep13670

Nicole, P., M., C. P., & Nils, K. (2006). MOLECULAR GENETIC MANIPULATION OF THE DIATOM THALASSIOSIRA PSEUDONANA (BACILLARIOPHYCEAE)1. *Journal of Phycology, 42*(5), 1059-1065. doi:https://doi.org/10.1111/j.1529-8817.2006.00269.x

Norsker, N.-H., Barbosa, M. J., Vermuë, M. H., & Wijffels, R. H. (2011). Microalgal production — A close look at the economics. *Biotechnology Advances, 29*(1), 24-27. doi:https://doi.org/10.1016/j.biotechadv.2010.08.005

Nour-Eldin, H. H., Hansen, B. G., Nørholm, M. H. H., Jensen, J. K., & Halkier, B. A. (2006). Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Research, 34*(18), e122-e122. doi:10.1093/nar/gkl635

Nymark, M., Sharma, A. K., Sparstad, T., Bones, A. M., & Winge, P. (2016). A CRISPR/Cas9 system adapted for gene editing in marine algae. *Scientific Reports, 6*, 24951. doi:https://doi.org/10.1038/srep24951

Olaizola, M. (2000). Commercial production of astaxanthin from Haematococcus pluvialis using 25,000-liter outdoor photobioreactors. *Journal of Applied Phycology, 12*(3), 499-506. doi:https://doi.org/10.1023/a:1008159127672

Onorato, I. M., Modlin, J. F., McBean, A. M., Thoms, M. L., Losonsky, G. A., & Bernier, R. H. (1991). Mucosal Immunity Induced by Enhanced-Potency Inactivated and Oral Polio Vaccines. *The Journal of Infectious Diseases, 163*(1), 1-6. doi:10.1093/infdis/163.1.1

Organization, J. E. T. (2011). Specifications and Standards for

Foods, Food Additives, etc.

Under the Food Sanitation Act

(Abstract) 2010

Paquet, M., Day, J. M. D., Udry, A., Hattingh, R., Kumler, B., Rahib, R. R., Tait, K. T., & Neal, C. R. (2021). Highly siderophile elements in shergottite sulfides and the sulfur content of the martian mantle. *Geochimica et Cosmochimica Acta, 293*, 379-398. doi:https://doi.org/10.1016/j.gca.2020.10.024

Patel, V. K., Soni, N., Prasad, V., Sapre, A., Dasgupta, S., & Bhadra, B. (2019). CRISPR–Cas9 System for Genome Engineering of Photosynthetic Microalgae. *Molecular Biotechnology, 61*(8), 541-561. doi:10.1007/s12033-019-00185-3

Plaza, M., Herrero, M., Cifuentes, A., & Ibáñez, E. (2009). Innovative Natural Functional Ingredients from Microalgae. *Journal of Agricultural and Food Chemistry, 57*(16), 7159-7170. doi:10.1021/jf901070g

Podrażka, M., Bączyńska, E., Kundys, M., Jeleń, P. S., & Witkowska Nery, E. (2017). Electronic Tongue-A Tool for All Tastes? *Biosensors, 8*(1), 3. doi:10.3390/bios8010003

PRC, N. H. C. o. t. (2016). List of new food ingredients and general food.

Pulz, O., & Gross, W. (2004). Valuable products from biotechnology of microalgae. *Applied Microbiology and Biotechnology, 65*(6), 635-648. doi:https://doi.org/10.1007/s00253-004-1647-x

Qin, S., Lin, H., & Jiang, P. (2012). Advances in genetic engineering of marine algae. *Biotechnology Advances, 30*(6), 1602-1613. doi:https://doi.org/10.1016/j.biotechadv.2012.05.004

Qintao, W., Yandu, L., Yi, X., Li, W., Shi, H., & Jian, X. (2016). Genome editing of model oleaginous microalgae Nannochloropsis spp. by CRISPR/Cas9. *The Plant Journal, 88*(6), 1071-1081. doi:https://doi.org/10.1111/tpj.13307

Radakovits, R., Jinkerson, R. E., Darzins, A., & Posewitz, M. C. (2010). Genetic Engineering of Algae for Enhanced Biofuel Production. *Eukaryotic Cell, 9*(4), 486-501. doi:https://doi.org/10.1128/ec.00364-09

Raleigh, V. S. (1999). Trends in world population: how will the millenium compare with the past? *Human reproduction update, 5*(5), 500-505.

Ramos-Martinez, E. M., Fimognari, L., & Sakuragi, Y. (2017). High-yield secretion of recombinant proteins from the microalga Chlamydomonas reinhardtii. *Plant Biotechnology Journal, 15*(9), 1214-1224. doi:https://doi.org/10.1111/pbi.12710

Rantasalo, A., Landowski, C. P., Kuivanen, J., Korppoo, A., Reuter, L., Koivistoinen, O., Valkonen, M., Penttilä, M., Jäntti, J., & Mojzita, D. (2018). A universal gene expression system for fungi. *Nucleic Acids Research, 46*(18), e111-e111.

Rasala, B. A., Barrera, D. J., Ng, J., Plucinak, T. M., Rosenberg, J. N., Weeks, D. P., Oyler, G. A., Peterson, T. C., Haerizadeh, F., & Mayfield, S. P. (2013). Expanding the spectral palette of fluorescent proteins for the green microalga Chlamydomonas reinhardtii. *The Plant Journal, 74*(4), 545-556. doi:https://doi.org/10.1111/tpj.12165

Rasala, B. A., Chao, S.-S., Pier, M., Barrera, D. J., & Mayfield, S. P. (2014). Enhanced Genetic Tools for Engineering Multigene Traits into Green Algae. *PLOS ONE, 9*(4), e94028. doi:10.1371/journal.pone.0094028

Rasala, B. A., Lee, P. A., Shen, Z., Briggs, S. P., Mendez, M., & Mayfield, S. P. (2012). Robust Expression and Secretion of Xylanase1 in Chlamydomonas reinhardtii by Fusion to a Selection Gene and Processing with the FMDV 2A Peptide. *PLOS ONE, 7*(8), e43349. doi:10.1371/journal.pone.0043349

Rasala, B. A., Muto, M., Lee, P. A., Jager, M., Cardoso, R. M., Behnke, C. A., Kirk, P., Hokanson, C. A., Crea, R., & Mendez, M. (2010). Production of therapeutic proteins in algae, analysis of expression of seven human proteins in the chloroplast of Chlamydomonas reinhardtii. *Plant Biotechnology Journal, 8*(6), 719-733.

Rath, B. (2012). Microalgal bioremediation: current practices and perspectives. *Journal of Biochemical Technology, 3*(3), 299-304.

Reeves, W. M., & Hahn, S. (2005). Targets of the Gal4 Transcription Activator in Functional Transcription Complexes. *Molecular and Cellular Biology, 25*(20), 9092-9102. doi:doi:10.1128/MCB.25.20.9092-9102.2005

Rochaix, J.-D. (1995). Chlamydomonas reinhardtii as the photosynthetic yeast. *Annual review of genetics, 29*(1), 209-230.

Rodriguez-Coira, J., & Sokolowska, M. (2021). SARS-CoV-2 candidate vaccines - composition, mechanisms of action and stages of clinical development. *Allergy, 76*(6), 1922-1924. doi:https://doi.org/10.1111/all.14714

Sadowski, I., Ma, J., Triezenberg, S., & Ptashne, M. (1988). GAL4-VP16 is an unusually potent transcriptional activator. *Nature, 335*(6190), 563-564. doi:10.1038/335563a0

Schirrmeister, B. E., Gugger, M., & Donoghue, P. C. J. (2015). Cyanobacteria and the Great Oxidation Event: evidence from genes and fossils. *Palaeontology, 58*(5), 769-785. doi:https://doi.org/10.1111/pala.12178

Schroda, M. (2006). RNA silencing in Chlamydomonas: mechanisms and tools. *Current Genetics, 49*(2), 69-84. doi:10.1007/s00294-005-0042-1

Schroda, M. (2019). Good News for Nuclear Transgene Expression in Chlamydomonas. *Cells, 8*(12), 1534.

Schroda, M., Beck, C. F., & Vallon, O. (2002). Sequence elements within an HSP70 promoter counteract transcriptional transgene silencing in Chlamydomonas. *The Plant Journal, 31*(4), 445-455.

Schroda, M., Blöcker, D., & Beck, C. F. (2000). The HSP70A promoter as a tool for the improved expression of transgenes in Chlamydomonas. *The Plant Journal, 21*(2), 121-131.

Scott, D. J., Gunn, N. J., Yong, K. J., Wimmer, V. C., Veldhuis, N. A., Challis, L. M., Haidar, M., Petrou, S., Bathgate, R. A. D., & Griffin, M. D. W. (2018). A Novel Ultra-Stable, Monomeric Green Fluorescent Protein For Direct Volumetric Imaging of Whole Organs Using CLARITY. *Scientific Reports, 8*(1), 667. doi:10.1038/s41598-017-18045-y

Scott, F., Binh, N., Ekem, E., & P., M. S. (2002). Development of a GFP reporter gene for Chlamydomonas reinhardtii chloroplast. *The Plant Journal, 30*(6), 733-744. doi:https://doi.org/10.1046/j.1365-313X.2002.01319.x

Scranton, M. A., Ostrand, J. T., Georgianna, D. R., Lofgren, S. M., Li, D., Ellis, R. C., Carruthers, D. N., Dräger, A., Masica, D. L., & Mayfield, S. P. (2016). Synthetic promoters capable of driving robust nuclear gene expression in the green alga Chlamydomonas reinhardtii. *Algal Research, 15*, 135-142. doi:https://doi.org/10.1016/j.algal.2016.02.011

Sheehan, J., Dunahay, T., Benemann, J., & Roessler, P. (1998). *Look back at the US department of energy's aquatic species program: biodiesel from algae; close-out report*. Retrieved from

Shi, X.-M., & Chen, F. (2002). High-Yield Production of Lutein by the Green Microalga Chlorella protothecoidesin Heterotrophic Fed-Batch Culture. *Biotechnology Progress, 18*(4), 723-727. doi:https://doi.org/10.1021/bp0101987

Shigesada, N., & Okubo, A. (1981). Analysis of the self-shading effect on algal vertical distribution in natural waters. *Journal of Mathematical Biology, 12*(3), 311-326. doi:https://doi.org/10.1007/bf00276919

Shimamatsu, H. (2004). Mass production of Spirulina, an edible microalga. *Hydrobiologia, 512*(1), 39-44. doi:10.1023/B:HYDR.0000020364.23796.04

Shimogawara, K., Fujiwara, S., Grossman, A., & Usuda, H. (1998). High-efficiency transformation of Chlamydomonas reinhardtii by electroporation. *Genetics, 148*(4), 1821-1828.

Shin, S.-E., Lim, J.-M., Koh, H. G., Kim, E. K., Kang, N. K., Jeon, S., Kwon, S., Shin, W.-S., Lee, B., Hwangbo, K., Kim, J., Ye, S. H., Yun, J.-Y., Seo, H., Oh, H.-M., Kim, K.-J., Kim, J.-S., Jeong, W.-J., Chang, Y. K., & Jeong, B.-r. (2016). CRISPR/Cas9-induced knockout and knock-in mutations in Chlamydomonas reinhardtii. *Scientific Reports, 6*, 27810. doi:https://doi.org/10.1038/srep27810

Shrager, J., Hauser, C., Chang, C.-W., Harris, E. H., Davies, J., McDermott, J., Tamse, R., Zhang, Z., & Grossman, A. R. (2003). Chlamydomonas reinhardtii Genome Project. A Guide to the Generation and Use of the cDNA Information. *Plant Physiology, 131*(2), 401-408. doi:https://doi.org/10.1104/pp.016899

Shurin, J. B., Abbott, R. L., Deal, M. S., Kwan, G. T., Litchman, E., McBride, R. C., Mandal, S., & Smith, V. H. (2013). Industrial-strength ecology: trade-offs and opportunities in algal biofuel production. *Ecology Letters, 16*(11), 1393-1404. doi:10.1111/ele.12176

Singh, B., Guldhe, A., Rawat, I., & Bux, F. (2014). Towards a sustainable approach for development of biodiesel from plant and microalgae. *Renewable and Sustainable Energy Reviews, 29*, 216-245. doi:https://doi.org/10.1016/j.rser.2013.08.067

Singh, U. B., & Ahluwalia, A. S. (2013). Microalgae: a promising tool for carbon sequestration. *Mitigation and Adaptation Strategies for Global Change, 18*(1), 73-95. doi:10.1007/s11027-012-9393-3

Smith, D. R., Lee, R. W., Cushman, J. C., Magnuson, J. K., Tran, D., & Polle, J. E. W. (2010). The Dunaliella salina organelle genomes: large sequences, inflated with intronic and intergenic DNA. *BMC Plant Biology, 10*, 83-83. doi:https://doi.org/10.1186/1471-2229-10-83

Song, Y., Yin, H., & Huang, W. E. (2016). Raman activated cell sorting. *Current Opinion in Chemical Biology, 33*, 1-8. doi:https://doi.org/10.1016/j.cbpa.2016.04.002

Specht, E., & Mayfield, S. (2014). Algae-based oral recombinant vaccines. *Frontiers in Microbiology, 5*(60). doi:10.3389/fmicb.2014.00060

Specht, E., Miyake-Stoner, S., & Mayfield, S. (2010). Micro-algae come of age as a platform for recombinant protein production. *Biotechnology Letters, 32*(10), 1373-1383. doi:10.1007/s10529-010-0326-5

Spolaore, P., Joannis-Cassan, C., Duran, E., & Isambert, A. (2006). Commercial applications of microalgae. *Journal of Bioscience and Bioengineering, 101*(2), 87-96. doi:https://doi.org/10.1263/jbb.101.87

Steinbrenner, J., & Sandmann, G. (2006). Transformation of the Green Alga Haematococcus pluvialis with a Phytoene Desaturase for Accelerated Astaxanthin Biosynthesis. *Applied and Environmental Microbiology, 72*(12), 7477-7484. doi:https://doi.org/10.1128/aem.01461-06

Stephenson, A. L., Kazamia, E., Dennis, J. S., Howe, C. J., Scott, S. A., & Smith, A. G. (2010). Life-Cycle Assessment of Potential Algal Biodiesel Production in the United Kingdom: A Comparison of Raceways and Air-Lift Tubular Bioreactors. *Energy & Fuels, 24*(7), 4062-4077. doi:https://doi.org/10.1021/ef1003123

Stoffels, L., Finlan, A., Mannall, G., Purton, S., & Parker, B. (2019). Downstream Processing of Chlamydomonas reinhardtii TN72 for Recombinant Protein Recovery. *Frontiers in Bioengineering and Biotechnology, 7*(383). doi:10.3389/fbioe.2019.00383

Straathof, A. (2011). The proportion of downstream costs in fermentative production processes. In *Comprehensive biotechnology* (pp. 811-814): Elsevier.

Sueoka, N., Chiang, K. S., & Kates, J. R. (1967). Deoxyribonucleic acid replication in meiosis of Chlamydomonas reinhardi: I. Isotopic transfer experiments with a strain producing eight zoospores. *Journal of Molecular Biology, 25*(1), 47-66. doi:https://doi.org/10.1016/0022-2836(67)90278-1

Swanson, D., Block, R., & Mousa, S. A. (2012). Omega-3 Fatty Acids EPA and DHA: Health Benefits Throughout Life. *Advances in Nutrition, 3*(1), 1-7. doi:https://doi.org/10.3945/an.111.000893

Szabo, N. J., Matulka, R. A., & Chan, T. (2013). Safety evaluation of Whole Algalin Protein (WAP) from Chlorella protothecoides. *Food and Chemical Toxicology, 59*, 34-45. doi:https://doi.org/10.1016/j.fct.2013.05.035

Takouridis, S. J., Tribe, D. E., Gras, S. L., & Martin, G. J. O. (2015). The selective breeding of the freshwater microalga Chlamydomonas reinhardtii for growth in salinity. *Bioresource Technology, 184*, 18-22. doi:https://doi.org/10.1016/j.biortech.2014.10.120

Tang, D., Qiao, S.-Y., & Wu, M. (1995). Insertion mutagenesis of Chlamydomonas reinhardtii by electroporation and heterologous DNA. *Biochemistry and molecular biology international, 36*(5), 1025-1035.

Tatematsu, K.-i., Kobayashi, I., Uchino, K., Sezutsu, H., Iizuka, T., Yonemura, N., & Tamura, T. (2010). Construction of a binary transgenic gene expression system for recombinant protein production in the middle silk gland of the silkworm Bombyx mori. *Transgenic Research, 19*(3), 473-487. doi:10.1007/s11248-009-9328-2

Tetali, S. D., Mitra, M., & Melis, A. (2007). Development of the light-harvesting chlorophyll antenna in the green alga Chlamydomonas reinhardtii is regulated by the novel Tla1 gene. *Planta, 225*(4), 813-829. doi:https://doi.org/10.1007/s00425-006-0392-z

Torres-Tiji, Y., Fields, F. J., & Mayfield, S. P. (2020). Microalgae as a future food source. *Biotechnology Advances, 41*, 107536. doi:https://doi.org/10.1016/j.biotechadv.2020.107536

Tran, H. T., & Plaxton, W. C. (2008). Proteomic analysis of alterations in the secretome of Arabidopsis thaliana suspension cells subjected to nutritional phosphate deficiency. *Proteomics, 8*(20), 4317-4326.

Tran, M., Van, C., Barrera, D. J., Pettersson, P. L., Peinado, C. D., Bui, J., & Mayfield, S. P. (2013). Production of unique immunotoxin cancer therapeutics in algal chloroplasts. *Proceedings of the National Academy of Sciences, 110*(1), E15-E22. doi:https://doi.org/10.1073/pnas.1214638110

Turner, N. J. (2003). The ethnobotany of edible seaweed (Porphyra abbottae and related species; Rhodophyta: Bangiales) and its use by First Nations on the Pacific Coast of Canada. *Canadian Journal of Botany, 81*(4), 283-293. doi:https://doi.org/10.1139/b03-029

Turvey, S. T., & Crees, J. J. (2019). Extinction in the Anthropocene. *Current Biology, 29*(19), R982-R986.

United Nations, D. o. E. a. S. A., Population Division. (2017). *World Population Prospects: The 2017 Revision, Key Findings and Advance Tables.* . Retrieved from

Urqueta, H., Jódar, J., Herrera, C., Wilke, H.-G., Medina, A., Urrutia, J., Custodio, E., & Rodríguez, J. (2018). Land surface temperature as an indicator of the unsaturated zone thickness: A remote sensing approach in the Atacama Desert. *Science of The Total Environment, 612*, 1234-1248. doi:https://doi.org/10.1016/j.scitotenv.2017.08.305

van der Wel, H., & Loeve, K. (1972). Isolation and Characterization of Thaumatin I and II, the Sweet-Tasting Proteins from Thaumatococcus daniellii Benth. *European Journal of Biochemistry, 31*(2), 221-225. doi:10.1111/j.1432-1033.1972.tb02522.x

Varshney, P., Mikulic, P., Vonshak, A., Beardall, J., & Wangikar, P. P. (2015). Extremophilic micro-algae and their potential contribution in biotechnology. *Bioresource Technology, 184*, 363-372. doi:https://doi.org/10.1016/j.biortech.2014.11.040

Vila, M., Díaz-Santos, E., de la Vega, M., Rodríguez, H., Vargas, Á., & León, R. (2012). Promoter Trapping in Microalgae Using the Antibiotic Paromomycin as Selective Agent. *Marine Drugs, 10*(12), 2749. doi:https://doi.org/10.3390/md10122749

Vlasov, Y., Legin, A., Rudnitskaya, A., Di Natale, C., & D'Amico, A. (2005). Nonspecific sensor arrays ("electronic tongue") for chemical analysis of liquids (IUPAC Technical Report). In *Pure and Applied Chemistry* (Vol. 77, pp. 1965).

Vukavic, T. (1983). Intestinal Absorption of IgA in the Newborn. *Journal of Pediatric Gastroenterology and Nutrition, 2*(2), 248-251.

Wakasugi, T., Nagai, T., Kapoor, M., Sugita, M., Ito, M., Ito, S., Tsudzuki, J., Nakashima, K., Tsudzuki, T., Suzuki, Y., Hamada, A., Ohta, T., Inamura, A., Yoshinaga, K., & Sugiura, M. (1997). Complete nucleotide sequence of the chloroplast genome from the green alga Chlorella vulgaris: The existence of genes possibly involved in chloroplast division. *Proceedings of the National Academy of Sciences of the United States of America, 94*(11), 5967-5972. doi:https://doi.org/10.1073/pnas.94.11.5967

Walsh, E. E., Frenck, R. W., Falsey, A. R., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., Neuzil, K., Mulligan, M. J., Bailey, R., Swanson, K. A., Li, P., Koury, K., Kalina, W., Cooper, D., Fontes-Garfias, C., Shi, P.-Y., Türeci, Ö., Tompkins, K. R., Lyke, K. E., Raabe, V., Dormitzer, P. R., Jansen, K. U., Şahin, U., & Gruber, W. C. (2020). Safety and Immunogenicity of Two RNA-Based Covid-19 Vaccine Candidates. *New England Journal of Medicine, 383*(25), 2439-2450. doi:10.1056/NEJMoa2027906

Wang, C., Chin, C.-K., Ho, C.-T., Hwang, C.-F., Polashock, J. J., & Martin, C. E. (1996). Changes of Fatty Acids and Fatty Acid-Derived Flavor Compounds by Expressing the Yeast Δ-9 Desaturase Gene in Tomato. *Journal of Agricultural and Food Chemistry, 44*(10), 3399-3402. doi:https://doi.org/10.1021/jf960174t

Wang, H., Li, X., Li, T., Zhang, S., Wang, L., Wu, X., & Liu, J. (2020). The genetic sequence, origin, and diagnosis of SARS-CoV-2. *European Journal of Clinical Microbiology & Infectious Diseases, 39*(9), 1629-1635. doi:10.1007/s10096-020-03899-4

Wang, Z. T., Ullrich, N., Joo, S., Waffenschmidt, S., & Goodenough, U. (2009). Algal Lipid Bodies: Stress Induction, Purification, and Biochemical Characterization in Wild-Type and Starchless Chlamydomonas reinhardtii. *Eukaryotic Cell, 8*(12), 1856-1868. doi:https://doi.org/10.1128/ec.00272-09

Wei, W., Xiang, H., & Tan, H. (2002). Two tandem promoters to increase gene expression in Lactococcus lactis. *Biotechnology Letters, 24*(20), 1669-1672. doi:10.1023/A:1020653417455

Wells, M. L., Potin, P., Craigie, J. S., Raven, J. A., Merchant, S. S., Helliwell, K. E., Smith, A. G., Camire, M. E., & Brawley, S. H. (2017). Algae as nutritional and functional food sources: revisiting our understanding. *Journal of Applied Phycology, 29*(2), 949-982. doi:https://doi.org/10.1007/s10811-016-0974-5

WHO. (2007). *Protein and Aminoacid Requirements in Human Nutrition*. Retrieved from Singapore:

Work, V. H., Radakovits, R., Jinkerson, R. E., Meuser, J. E., Elliott, L. G., Vinyard, D. J., Laurens, L. M. L., Dismukes, G. C., & 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*(8), 1251-1261. doi:https://doi.org/10.1128/ec.00075-10

Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., Hu, Y., Tao, Z.-W., Tian, J.-H., Pei, Y.-Y., Yuan, M.-L., Zhang, Y.-L., Dai, F.-H., Liu, Y., Wang, Q.-M., Zheng, J.-J., Xu, L., Holmes, E. C., & Zhang, Y.-Z. (2020). A new coronavirus associated with human respiratory disease in China. *Nature, 579*(7798), 265-269. doi:10.1038/s41586-020-2008-3

Wu, Z., Duangmanee, P., Zhao, P., Juntawong, N., & Ma, C. (2016). The Effects of Light, Temperature, and Nutrition on Growth and Pigment Accumulation of Three Dunaliella salina Strains Isolated from Saline Soil. *Jundishapur Journal of Microbiology, 9*(1), e26732. doi:https://doi.org/10.5812/jjm.26732

Yang, J., Wang, W., Chen, Z., Lu, S., Yang, F., Bi, Z., Bao, L., Mo, F., Li, X., Huang, Y., Hong, W., Yang, Y., Zhao, Y., Ye, F., Lin, S., Deng, W., Chen, H., Lei, H., Zhang, Z., Luo, M., Gao, H., Zheng, Y., Gong, Y., Jiang, X., Xu, Y., Lv, Q., Li, D., Wang, M., Li, F., Wang, S., Wang, G., Yu, P., Qu, Y., Yang, L., Deng, H., Tong, A., Li, J., Wang, Z., Yang, J., Shen, G., Zhao, Z., Li, Y., Luo, J., Liu, H., Yu, W., Yang, M., Xu, J., Wang, J., Li, H., Wang, H., Kuang, D., Lin, P., Hu, Z., Guo, W., Cheng, W., He, Y., Song, X., Chen, C., Xue, Z., Yao, S., Chen, L., Ma, X., Chen, S., Gou, M., Huang, W., Wang, Y., Fan, C., Tian, Z., Shi, M., Wang, F.-S., Dai, L., Wu, M., Li, G., Wang, G., Peng, Y., Qian, Z., Huang, C., Lau, J. Y.-N., Yang, Z., Wei, Y., Cen, X., Peng, X., Qin, C., Zhang, K., Lu, G., & Wei, X. (2020). A vaccine targeting the RBD of the S protein of SARS-CoV-2 induces protective immunity. *Nature, 586*(7830), 572-577. doi:10.1038/s41586-020-2599-8

Yeomans, M. R., Gould, N. J., Mobini, S., & Prescott, J. (2008). Acquired flavor acceptance and intake facilitated by monosodium glutamate in humans. *Physiology & Behavior, 93*(4), 958-966. doi:https://doi.org/10.1016/j.physbeh.2007.12.009

Young, H. S., McCauley, D. J., Galetti, M., & Dirzo, R. (2016). Patterns, Causes, and Consequences of Anthropocene Defaunation. *Annual Review of Ecology, Evolution, and Systematics, 47*(1), 333-358. doi:10.1146/annurev-ecolsys-112414-054142

Zaslavskaia, L. A., Lippmeier, J. C., Shih, C., Ehrhardt, D., Grossman, A. R., & Apt, K. E. (2001). Trophic Conversion of an Obligate Photoautotrophic Organism Through Metabolic Engineering. *Science, 292*(5524), 2073-2075. doi:https://doi.org/10.1126/science.160015

Zhang, P., Ren, L., Zhang, X., Shan, Y., Wang, Y., Ji, Y., Yin, H., Huang, W. E., Xu, J., & Ma, B. (2015). Raman-Activated Cell Sorting Based on Dielectrophoretic Single-Cell Trap and Release. *Analytical Chemistry, 87*(4), 2282-2289. doi:https://doi.org/10.1021/ac503974e

Zhang, Q., Zhang, P., Gou, H., Mou, C., Huang, W. E., Yang, M., Xu, J., & Ma, B. (2015). Towards high-throughput microfluidic Raman-activated cell sorting. *Analyst, 140*(18), 6163-6174. doi:https://doi.org/10.1039/C5AN01074H

Zhang, Z., Tan, Y., Wang, W., Bai, W., Fan, J., Huang, J., Wan, M., & Li, Y. (2019). Efficient heterotrophic cultivation of Chlamydomonas reinhardtii. *Journal of Applied Phycology, 31*(3), 1545-1554. doi:10.1007/s10811-018-1666-0

Zhu, C. J., & Lee, Y. K. (1997). Determination of biomass dry weight of marine microalgae. *Journal of Applied Phycology, 9*(2), 189-194. doi:10.1023/A:1007914806640