Spring 4-13-2022

Improving photosynthetic efficiency in microalgae through the genetic engineering of energy sensors and photoreceptors

Taylor L. Britton
University of New Mexico

Taylor Britton

Candidate

Biology

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

David Hanson , Chairperson

Donald Natvig

Sangeeta Negi

Matthew Posewitz

Improving photosynthetic efficiency in microalgae through the genetic engineering of energy sensors and photoreceptors

By Taylor Britton:

Bachelor of Science Biology and Chemistry

Master of Science Biomedical Engineering

DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of
Doctor of Philosophy in Biology

The University of New Mexico

Albuquerque, New Mexico

May 2022

IMPROVING PHOTOSYNTHETIC EFFICIENCY IN MICROALGAE THROUGH THE GENETIC ENGINEERING OF ENERGY SENSORS AND PHOTORECEPTORS

By

Taylor Britton

BACHELOR OF SCIENCE BIOLOGY AND CHEMISTRY

MASTER OF SCIENCE BIOMEDICAL ENGINEERING

DOCTOR OF PHILOSOPHY BIOLOGY

Abstract

Through photosynthesis microalgae can convert sunlight, water, and CO_2 into chemical energy that can be used to generate carbon neutral biofuels and biomass. With an ever-increasing demand and need for petroleum substitutes it is imperative that we improve the output of industrial-relevant crops such as microalgae. One important way of improving output in algae is by understanding the roles that stress and energy conversion is regulated in these organisms. Photosynthetic organisms fundamentally depend on light- and sugar-driven metabolic and signaling networks, which integrate environmental cues to govern and sustain growth and survival. SnRKs (SNF1-related protein kinases) and the photoreceptor neochrome are regulators of energy and stress metabolism, and coordinate energy balance and nutrient metabolism in plants. This work set to elucidate the role that these two sensors play in generating algal biomass. Further, we identified SnRKs in thirty-four algal species and analyzed their SNRK domain motifs and phylogenetic relationships. To determine the importance of the sensors in algal evolution.

Table of Contents

Part 1

Introduction

Over the past several decades there has been heavy investment into replacing fossil fuels via algal biofuels.[1] Through photosynthesis, plants and microalgae can efficiently convert sunlight, water, and CO_2 into chemical energy that can be used to generate carbon neutral biofuels and biomass. The increasing demand for food and the pressing need for "green" petroleum substitutes requires that we improve the output of industrial-relevant crops. However, an increase in food crop production is hampered by arable land and water availability.[2–5] Photosynthetic microbes are also recognized as potential biofuel production platforms that do not compete with food crops for arable land.[3] In general, microalgae and cyanobacteria possess favorable attributes such as: they do not partition biomass into support structures such as stems and roots, are considered "greener" because they do not require pesticides or herbicides, have high biomass per area, high photosynthetic efficiency, and higher lipid content than plants.[6–12]. Further, traditional crops must be adapted (through breeding or genetic engineering) to be more efficient at biomass production under constrained growth conditions and cultivation can lead to ecological degradation.[12]. Improving photosynthetic energy conversion efficiency has emerged as a promising strategy to increase crop yields due to the need to increase higher yields.[1,13–15].

The genetic tractability of microalgae is increasing due to the deployment of genetic engineering tools such as CRISPR/Cas9.[16–20] This is evident from recent research improving the efficiency of photosynthesis through genetic modification in micro-algae.[9,11,16–22] A lot of focus has been placed on the manipulation of central carbon metabolism. Even with improvements in the available genetic tools currently microalgal biofuels are not viable because of their cost. Gasoline from algae is currently between $12-14 per gallon with 75% of

this cost coming from algal cultivation and harvesting; thus, it is important that research focuses on increasing algal yields.[2,23–25] During algal cultivation several issues arise that lead to a suboptimal yield. These issues include self-shading, low yield during low light conditions, and natural regulatory suppression of photosynthesis in saturating light. This has been shown in photosynthetic microbes with reduced antenna size, which increased light penetration and photosynthetic efficiency.[6,26–28]

Microalgae come in many different forms ranging from single cells to multicellular structures (like filaments and colonies). They also can live in a wide range of environments including temperatures over 50°C and in pHs as low as 1.[29] This range expands potential areas in which microalgae can be used, including wastewater, bioremediation, thermal and textile industries.[29,30] Differences in the conditions in which algae live also increase the range of useful products that can be obtained from them including proteins, vitamins, pigments, nutraceuticals, and special oils.[30] Despite all these positive attributes there is still a need for further identification of species that produce high amounts of lipids (bioprospecting). Even less work has been done on identifying genes that positively effect biomass accumulation. This is especially true for genes that span multiple lineages of algae. We currently lack a comprehensive understanding of the mechanisms regulating metabolic pathways at the cellular level. Much of the research has focused on single pathways.[31] Many master regulators (genes) that control the expression of other genes are currently unknown. Further, many of the genetic engineering techniques have focused on a small number of algal species including Chlamydomonas, Chlorella, Synechocystis, Phaeodoactylum, and Nanochloropsis. While many of these species are good candidates for biofuels, many species remain with positive attributes yet to be studied for genetic manipulation.

The goal of this research was to increase photosynthetic efficiency and improve biomass production in microalgae by targeting regulatory genes and working with some non-traditional species to increase starch content, limit self-shading, deregulate suppression of photosynthesis in saturating light, and provide additional harvesting of the light spectrum. This will be accomplished through two experimental approaches in multiple species. First, overexpressing a known energy sensor sucrose non-fermenting related protein kinase (SnRK) in Chlorella sorokiniana 1412. Second, knocking out the photoreceptor phototropin and independently/subsequently knocking in the photoreceptor neochrome in the following micro-algal species; Picochlorum soloecismus and Scenedesmus obliquus, while also knocking in neochrome in Chlamydomonas reinhardtii phototropin knockout (PHOT KO) lines and wild-type. Third, knocking out a highly conserved photoreceptor that has been found to have regulatory effects upon photosynthesis. I will also be making use of CRISPR/Cas9 and CPF1 technologies in P. soloecismus and S. obliquus. I hypothesize that these genetic modifications will improve algal biomass accumulation and photosynthetic efficiency, lowering the cost of algal biofuel production.

Part 2

Engineering of the sugar signaling pathway regulator SnRK2 leads to improved photosynthesis and growth in *Chlorella sorokiniana*

Taylor Britton[1], C. Raul Gonzalez-Esquer[2], Carol Kay Carr[2], Nilusha Sudasinghe[2], Taraka Dale[2], Kyle Pittman[2], Shawn Starkenburg[2], Richard Sayre[2], Amanda Barry[2], David T. Hanson[1], Scott Twary[2] and, Sangeeta Negi[3, *]

1. Department of Biology, The University of New Mexico, Albuquerque, NM, 87104, USA

2. Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

3. New Mexico Consortium, Los Alamos, NM, 87544, USA

***** Corresponding author: Sangeeta Negi (sangeeta@lanl.gov)

Engineering of the sugar signaling pathway regulator SnRK2 leads to improved photosynthesis and growth in *Chlorella sorokiniana*

Abstract

Photosynthetic organisms fundamentally depend on light- and sugar-driven metabolic and signaling networks, which integrate environmental cues to govern and sustain growth and survival. SnRKs (SNF1-related protein kinases) are well-studied regulators of energy and stress metabolism, and coordinate energy balance and nutrient metabolism in plants. SnRK2 overexpression in plants enhances growth and biomass yield. We hypothesized that transgenic algae overexpressing SnRK2 will also increase their carbon assimilation rates, resulting in improved growth rates and productivity. To test this concept, a heterologous SnRK from the halotolerant microalgae *Picochlorum soloecismus* was engineered for high expression in the biofuel production strain *Chlorella sorokiniana*. SnRK2 transgenic lines show an increase in biomass, starch accumulation, and photosynthetic rates. These traits were further tested in environmental photobioreactors (ePBRs) during a semi-continuous harvesting experiment comparing growth and biomass accumulation. SnRK overexpression showed up to 3-fold increase in biomass and 2-fold in carbohydrates compared to wild type. These properties show that SnRK overexpression lines are prime candidates for potential biofuel crops

Keywords

SnRK, *Chlorella*, transgenic, biomass, carbohydrate, environmental photobioreactor

Introduction

Photosynthetic microorganisms are considered favorable platforms for the production of biofuels and renewable chemical feedstocks. Specifically, they do not compete with food crops for arable land, they have higher biomass yields per area than traditional plant feedstocks, and can accumulate high levels of useful fuel/product precursor molecules, such as certain carbohydrates within their cells.[2–5] Even with these promising traits, algae production at the industrial scale remains a challenge. As expected, the contrast between a laboratory-controlled and a highly dynamic exterior environment impedes the translation of indoor-to-outdoor yields, in part due to stress generated by fluctuations in light, temperature, nutrient availability, and contamination.[32–34] Therefore, targeted engineering of the mechanisms that regulate intracellular responses may help overcome abiotic stress factors that affect microalgae productivity in outdoor conditions.

In plants, sucrose non-fermenting related kinases (SnRKs) regulatory function is at the interface between energy signaling, stress signaling, and metabolism indicating that they are central within a network of signaling pathways. [35,36] Prevailing views indicate that SnRKs are evolutionarily conserved in eukaryotes including energy sensors in yeast (SNF), mammals (AMPK), and SnRK orthologs in microalgae.[16,35,37–42] In response to various abiotic and metabolic stress SnRKs activation signaling cascades via phosphorylation from a conserved Ser/Thr protein kinase domain near the N terminus of the protein.[35,38,43] Each pathway that is activated corresponds to one of the three subfamilies that make up the SnRK family. Specifically, SnRK2 family members link stress signaling with cellular metabolism and are involved in responses to nutrient limitation, drought, cold, salt, and other osmotic stresses, and ABA signaling.[16,36,37,44–47] While it is evident that SnRK2 kinases play an essential role in a plant's response to stress, they also play an important role in carbon metabolism. Several plant studies suggest that some of the SnRK2 family members are positive regulators of plant growth. [16,36,37,48] Studies overexpressing native SnRK2.6 and SnRK2.8 from *Arabidopsis* led to an increase in leaf expansion, biomass, and lateral branching in well-watered conditions, while the down regulation of SnRK2.8 during K^+, N, or P deprivation resulted in a reduction of growth during similar water regiments.[16,37] Overexpression of wheat SnRK2.8 also resulted in larger root biomass.[36] Research also shows that SnRK2 phosphorylation targets are proteins involved in energy metabolism such as glycolysis, as evidenced by increases in total soluble sugars and its regulation of sucrose metabolism.[16,37] While the role that SnRK2 has been explored in plants, little is known regarding their role in algae.[49] In *Chlamydomonas*, the SnRK2 (known as CKIN2) has shown involvement in abiotic stress response including sulfur limitation, nitrogen deprivation, and cold stress.[50–54] SnRK2.2 in *Chlamydomonas* was also shown to act as a positive regulator of DGTT1, which enhanced TAG synthesis under sulfur deprivation conditions.[55]

Figure 1: Sequence alignment of three *A. thaliana* SnRK2s and two potential SnRK2s from *C. sorokiniana* and *Picochlorum soloecismus*. Shared amino acid residues are indicated by black boxes and secondary structure elements are indicated including the red star representing the phosphoserine. The alignment was generated with the ClustalW program.

In this manuscript, we set to illustrate the effects of expressing the SnRK2 gene from the halotolerant microalga *P. soloecismus* in the freshwater green microalga *Chlorella sorokiniana* DOE 1412 (hereafter *Chlorella sorokiniana*). We demonstrate that recapitulating SnRK2 regulation of metabolism via supplementary expression of a heterologous SnRK gene improves growth, starch content and photosynthetic efficiency in the biofuel production platform *Chlorella sorokiniana.*

Results

SnRK2 Identification and molecular characterization in *C. sorokiniana* 1412 and *P. soloecismus*

To identify potential SnRK2s in *C. sorokiniana* we used the SnRK2.8 and SnRK2.6 protein sequences from *A. thaliana* as BLASTP query sequences. Candidates with an e value lower than e-^{30} were analyzed using InterPro to scan for the presence of the conserved Ser/Thr kinase domain (Pfam PF00069) found among *A. thaliana* SnRK2 members. Through this analysis we generated a list of 14 proteins found in *C. sorokiniana* that have the common SnRK kinase Pfam (Supplemental Figure 1). Protein sequence alignments generated with EMBOSS Clustal Omega show that the top candidate has close sequence similarity with *A. thaliana* SnRK2s (Figure1). To make screening of positive transformants more efficient we made use of a heterologous protein from the marine algae *P.* soloecismus, thus a similar analysis was conducted to identify an algal heterologous SnRK2 protein for

overexpression. The top candidates from *C. sorokiniana* and *P. soloecismus* were compared to each other and to the ten SnRK2 family members from *Arabidopsis* based on percent identities (Supplemental Figure 2). Our chosen protein sequence from *P. soloecismus* has a 75.38% sequence identity with the top hit from our *C. sorokiniana* (CSJ00008848), while having little nucleotide sequence similarity. Further, the GC content for *P. soloecismus* (40%) is significantly lower compared to *C. sorokiniana* (64%), which is similar to the overall GC content of the corresponding genomes (*P. soloecismus* (46%) and *C. sorokiniana* (64%). Sequence alignments show that the algal SnRK2s were thirteen to nineteen amino acids shorter than the *Arabidopsis* SnRK2s and that the SnRK2s identified in the two algal species have protein sequence structure similar to those found in Arabidopsis (Figure 1).[48] Further, based on the protein sequences from the two algal species the genetic structure similarity between the candidate SnRK2 from *P. soloecismus* is similar enough to the top hit from *C. sorokiniana*, while having no nucleotide sequence similarity, and lower GC content made it a prime candidate for overexpression.

Engineering of a *Chlorella sorokiniana* 1412 SnRK2 transgenic strain

After determining the presence of SnRK2 in *C. sorokiniana* and *P.* soloecismus, we designed an overexpression vector for the heterologous expression of SnRK2. A marker gene is used to determine if a successful genetic transformation has taken place. We used antibiotic selection as our marker gene of choice. We first tested the sensitivity of *C. sorokiniana* to various antibiotic concentrations. All the antibiotics tested (Paromomycin, Zeocin and G418) were effective at killing *C. sorokiniana* (Supplementary Figure 11). Zeocin was selected as our selective marker because it was effective at killing *C. sorokiniana*, but research has also shown that the addition of the zeocin resistance (sh-ble) gene can improve heterologous expression of transgenes and has been used to generate transgenic clones from various algal species.[56–58] Second, with this information we constructed an overexpression vector (PACE-SnrK2; Supplemental Figure 3) with zeocin selection as the antibiotic marker and placed the heterologous SnRK2 gene NSC_03857 from *P. soloecismus* under the control of a PsaD

promoter/terminator pair (Supplemental Figure 3). PsaD encodes for an abundant chloroplast protein located on the stromal side of the photosystem I complex and its promoter has widely been used for efficient gene expression in microalgae.[59] Lastly, the vector PACE-SnRK2 was linearized and transformed via electroporation. Electroporation has been shown to be an effective and efficient method for genetic transformations of algae.[60,61] We successfully generated > 100 colonies that successfully grew on the zeocin selective media after genetic transformation. These putative transgenic lines were first screened through PCR. Primers were designed with amplicons for a region within the PsaD promoter and the SnRK2 gene. This amplicon yielded an expected band of ~1 Kb in transformed algae, while no band was present in the wild type (Figure 2A); the band identity was further confirmed by Sanger sequencing. Five transgenic lines were selected to confirm SnRK2 gene expression via reverse transcriptase PCR (RT-PCR) and reverse transcriptase quantitative PCR (RT-qPCR) analysis. No SnRK2 transcripts were detected in wild-type cells, while the transgenic lines show SnRK2 expression above background levels with values ranging from a 15 to 1-fold increases (Figure 2B-C). Therefore, this indicated the successful generation of SnRK2 overexpression strains in *C. sorokiniana*.

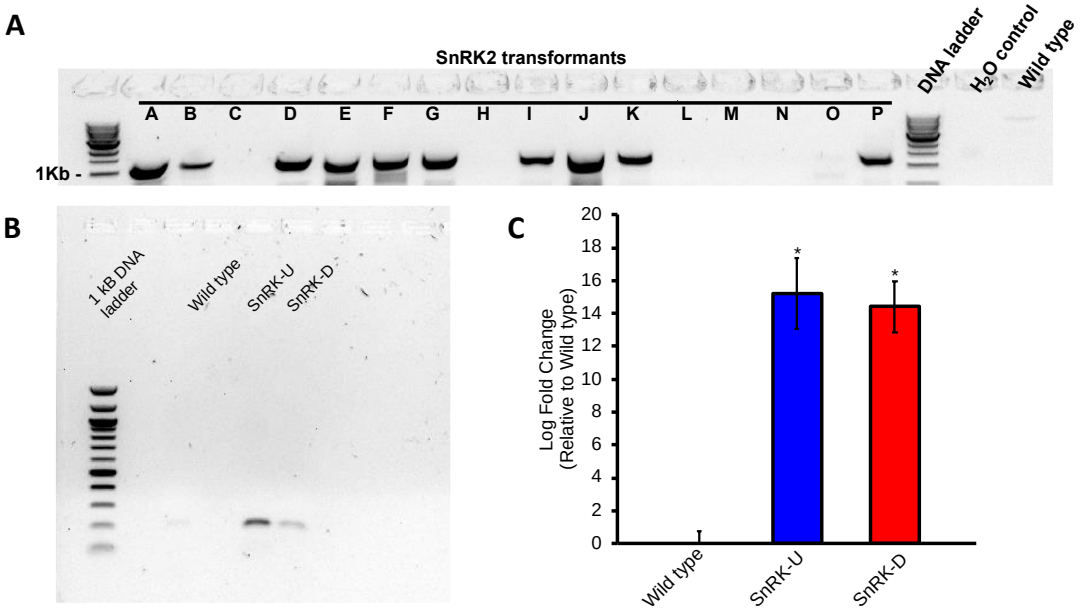


Figure 2: SnRK identification in transformants and gene expression confirmation. (A) Detection of the heterologous SnRK in transformants via PCR. PCR products had an expected size of 960bp. (B) PCR of cDNA generate from positive SnRK transformants using SnRK specific primers. PCR products had an expected size of 200bp. (C) qPCR of positive SnRK transformants cDNA using SnRK specific primers as in (B). Data was performed in triplicate. The asterisk (*) represents significant data (P-value <0.05).

SnRK2 overexpression increases batch mode growth, cell size, and starch accumulation in transgenic lines

Based on expression levels we decided to characterize the strains SnRK2-D and SnRK2-U, as they expressed SnRK2 >10-fold over the background (Figure 2C). A batch-mode comparison of growth between wild type and these two transgenic lines was conducted in both low light intensities (50 µmol m^{-2} s^{-1}) and high light intensities (550 µmol m^{-2} s^{-1}) in photoautotrophic conditions. Cultures exhibited very similar rates, albeit a short lag phase for the wild type control. In high light after the lag, wild type had 0.0452±0.004 doublings/day, while SnRK2-D was 1.246±0.004 doublings/day and SnRK2-U was 1.244±0.004 doublings/day. We determined the percent difference between the transgenic lines and the wild type on the bases of optical density (OD) on the final day of the growth experiment (day 7). SnRK2-D showed a 26% difference in low light and 30% increase in high light, while SnRK2-U showed 23% increase in low light and 20% increase in high light intensities. During these growth experiments we also observed cells under the microscope. Wild type cells had the characteristic

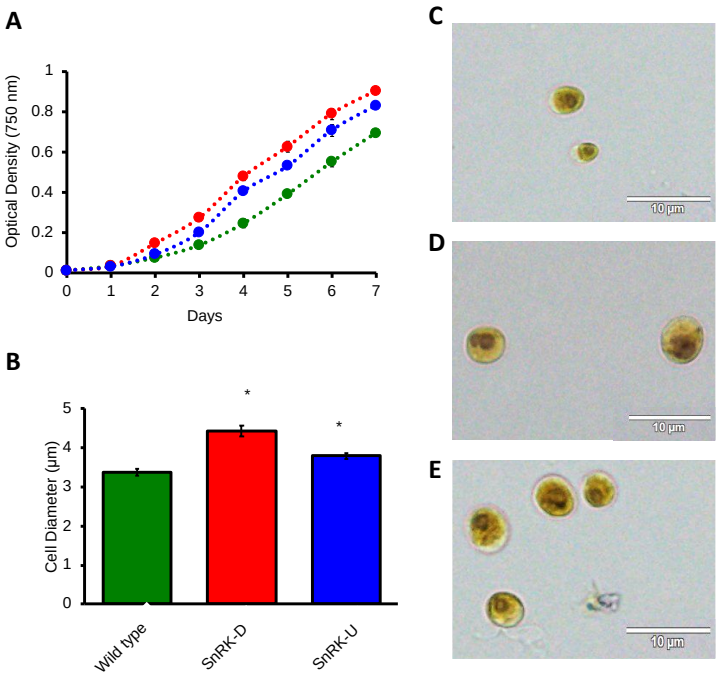


Figure 3: SnRK2 growth and cell size determination. (A) Growth curves of *C. sorokiniana* wild type and SnRK2 lines grown under photoautotrophic conditions. The growth curve is based on cell density measured at 750 nm. (B) Cell size values determined by measuring the diameter of individual cells. Data is the average of 50-100 measurements per cell line. The asterisk (*) represents significant data (P-value <0.05). Microscopy of wild type and SnRK cells. Cells were imaged after seven days of growth in photoautotrophic conditions. (C) Wild type (D) SnRK-D (E) SnRK-U

appearance of *C. sorokinana* (Figure 3C), with the transgenics being larger; the average diameter of wild type is at 3.4±0.6 µm, versus the diameters of SnRK2-D at 4.4 ± 0.9 µm and SnRK2-U at 3.8± 0.9 µm (Figure 3B). To analyze potential changes in starch accumulation cells were stained with Lugol's to observe intracellular starch by microscopy. This method has successfully been used in a *Chlorella* species.[62] Transgenic lines showed visible granularity in the cytoplasm (Figure 3D-E). After Lugol's staining of the cells, SnRK2-D and subsequently SnRK2-U qualitatively stained more with the reagent than the wildtype strain, putatively indicating that this granularity may be due to more starch present within the mutant cells.

SnRK2 effects total chlorophyll, chlorophyll a/b ratios, and flash fluorescence

Due to our observed increase in starch content in transgenic lines we wanted to observe the effect of SnRK2 overexpression on chlorophyll (Chl) amount, content, and photosystem II (PSII) regulation. First, we quantified total Chl; SnRK2 transgenic lines have a significantly lower amount of total Chl (0.42± 0.01 and 0.45± 0.01 for SnRK2-D and SnRK2-U, respectively) as compared to wild type cells 0.68± 0.03 (Figure 4A). Second, the Chl *a/b* ratio was also altered in the transgenic lines; 2.7±0.1 for wild type and 3.9±0.1 for SnRK2-D and 3.7±0.02 for SnRK2-U (Figure 4B). Smaller a/b ratios

correspond to a shrinking light harvesting complex. It has been shown before that the most effective strategy to increase photosynthetic light utilization efficiency is to reduce the size of the light-harvesting antenna, which reduces the optical cross section of the antenna complexes.[6,14,28,63] To determine the relationship between antenna size and PSII (Photosystem II) optical section we performed a quick Chl fluorescence induction measurement. Flash fluorescence induction was measured using a Chl concentration of ~ 2.5 µg Chl/mL. All the cells were dark adapted for 10 min prior to the experiment.

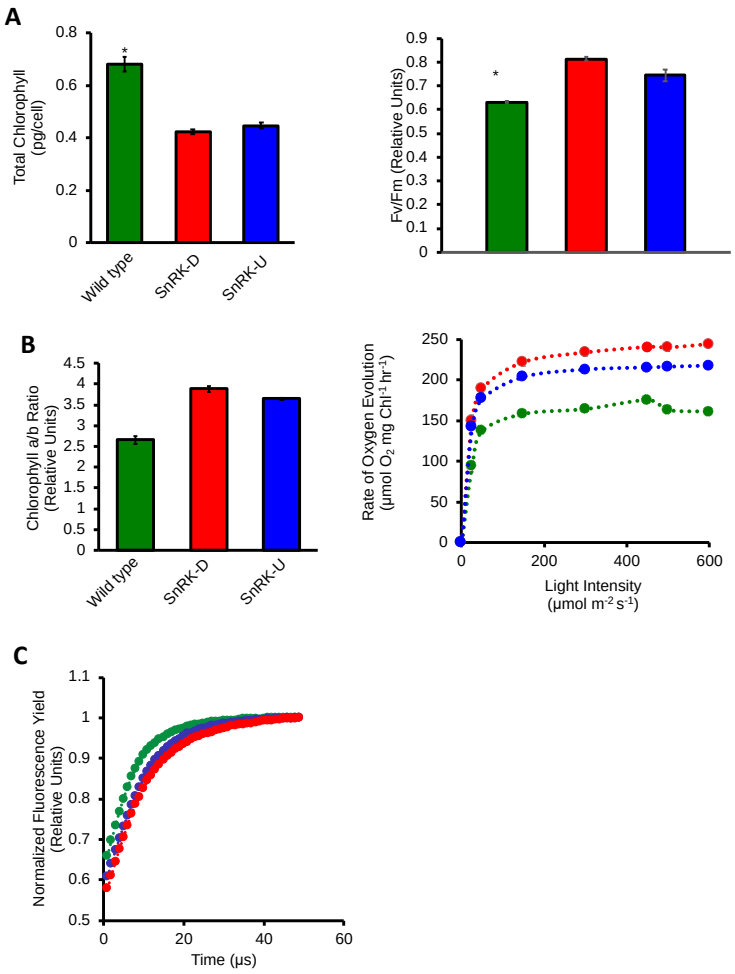


Figure 4: Chl analysis and light-dependent rates of photosynthesis for log-phase grown cultures in 50 µ mol photons $m^{-2}s^{-1}$ light. (A) Total Chl content measured on a per cells basis. (B) Chl a/b ratios of SnRK2 lines and wild type cells. (C) Analysis of maximum quantum efficiency (Fv/Fm) and Chl fluorescence induction kinetics for log-phase grown cultures in 50 µ mol photons $m^{-2}s^{-1}$ light. (D) Fv/Fm measurements shown were conducted at the end of a light curve with a final light intensity of 800 photons $m^{-2}s^{-1.}$ Chl induction kinetics of low-light grown wild type and SnRK lines. Chl levels were measured under continuous non-saturating illumination every 1 µ s. All data shown was performed in triplicate. The asterisk (*) represents significant data (P-value <0.05).(E) Oxygen evolution measurements of light response curves of photosynthesis in photoautotrophic conditions. Rates of oxygen evolution were measured in the presence of 10 mM $NaHCO_3$ and on a per Chl basis. All data shown was performed in triplicate. The asterisk (*) represents significant data (P-value <0.05).

The values of Chl fluorescence were normalized to the maximum achieved values for a given sample. Chl fluorescence analysis determined that both SnRK lines demonstrated slower saturation of PSII, confirming the smaller cross section of PSII antenna (Figure 4C).

Photosynthetic efficiency is increased in SnRK2 transgenics

Results obtained from our Chl analysis and flash fluorescence experiments along with the potential for increased starch accumulation, we preformed photosystem (II) efficiency (F_V/F_m) and photosynthetic

oxygen evolution measurements to determine the impact on photosynthetic output. Cells were harvested at mid-log phase which was the same culture age as the Chl content and flash fluorescent induction experiments. Samples for F_V/F_m measurements were subjected to single pulse measurement. Figure 4D shows a significant increase in F_V/F_m in the transgenic lines as compared to the wild type. Oxygen evolution rates were measured on a per Chl basis. As shown in Figure 4E, the SnRK2 transgenic lines show increased photosynthesis at light saturation conditions. SnRK2-U showed a 60% and SnRK2-D 80% increase compared the wild type. Light saturation rate of photosynthesis occurred at ~300 µmol m^{-2} s^{-1} which is consistent with *C. sorokiniana* grown in photoautotrophically conditions. [64]

SnRK2 transgenics preform significantly better in outdoor simulated conditions.

To test potential outdoor growth, we grew the microalgal cells in environmental photobioreactors (ePBRS) under sinusoidal light and temperature mimicking outdoor growth conditions of Mesa, AZ on October 1st. This experiment was performed semi-continuously over a thirty-day period with five consecutive harvests. Cells were harvested every five days. ePBRS were operated with sinusoidal light delivery with 12h light-dark period reaching peak intensity at 1:00pm. The cultures also experienced a range of temperatures from 14-27 °C over a 24 h period with low temperature of 14 C occurring at 4:00am and peak temperature 27 °C occurring at 4:00pm. Optimal growth for *C. sorokiniana* occurs between 30-40 °C thus our cultures experienced cold stress over the thirty day period.[65,66] Over the growth period, the transgenic lines exhibited increased growth rates compared to the wild type (Figure 5A), translating into higher recovered biomass of all harvests combined (wild type 2.17±0.47g m^{-2} d^{-1}, SnRK2-U 6.63±0.22g m^{-2} d^{-1}, and SnRK2-D 5.38±0.8222g m^{-2} d^{-1}) (Figure 5B). Statistical significance was detected for harvests two, four, and five. In addition to seeing an increase in total biomass an MBTH total carbohydrate which includes both structural and storage carbohydrates was quantified spectrophotometrically. Biochemical analysis showed that SnRK2-D mutants accumulate 15.5±1.3% carbohydrates, SnRK2-U mutants accumulate 12.5±0.7%, and that the wild type accumulates 7.0±1.9%

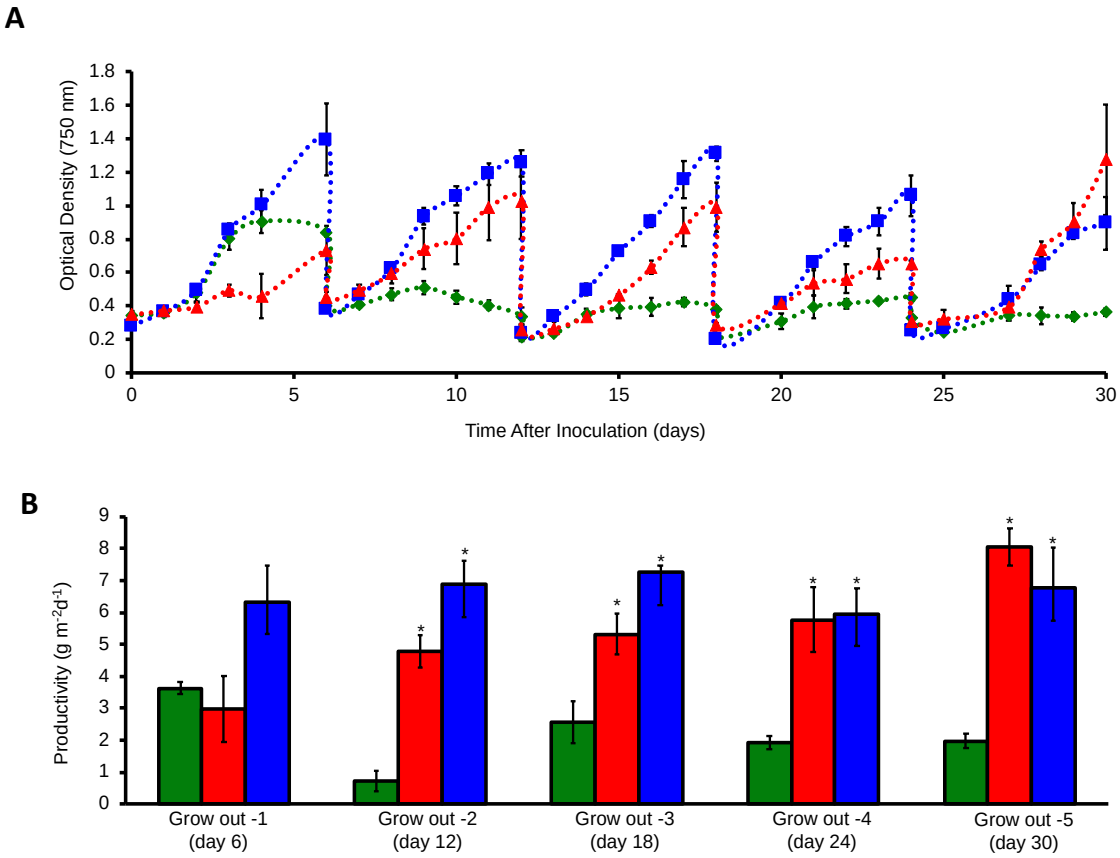


Figure 5: SnRK lines and wt growth and productivity during the semi-continuous harvesting experiment in the ePBRs. EPBR experiment was performed in triplicate or quadruplicate (depending on the strain). The asterisk (*) represents significant data (P-value <0.05). (A) OD measurements, taken at 750nm with a Bio-Rad SmartSpec Plus spectrophotometer, were used to measure culture growth over the course of the experiment, as we grew the 3-4 replicates of each strain up and then diluted back. (B) All biomass that was not used to re-inoculate the ePBR cultivation vessels for the continuation of the semi-continuous harvesting experiment was weighed, and productivity values were calculated. Values were compensated for ash content due to inorganics.

carbohydrates. Due to low biomass were not able to obtain carbohydrate data for the wild type during the second grow out. (Figure 6A). SnRK2-D and SnRK2-U exhibited significantly higher carbohydrate levels relative to the wildtype for harvests one, two, and four. Further we observed an 104% and 91% (AFDW basis) average increase in total carbohydrates for SnRK2-D and SnRK2-U respectively compared to the wild type. Neutral lipids were also quantified as fatty acid methyl esters (FAME) by gas chromatography-flame ionization detections (GC-FID). SnRK2-U exhibited significantly higher

FAME levels for harvest three and four, while a significant difference between the wild type and SnRK2-D was not detected (Figure 6B).

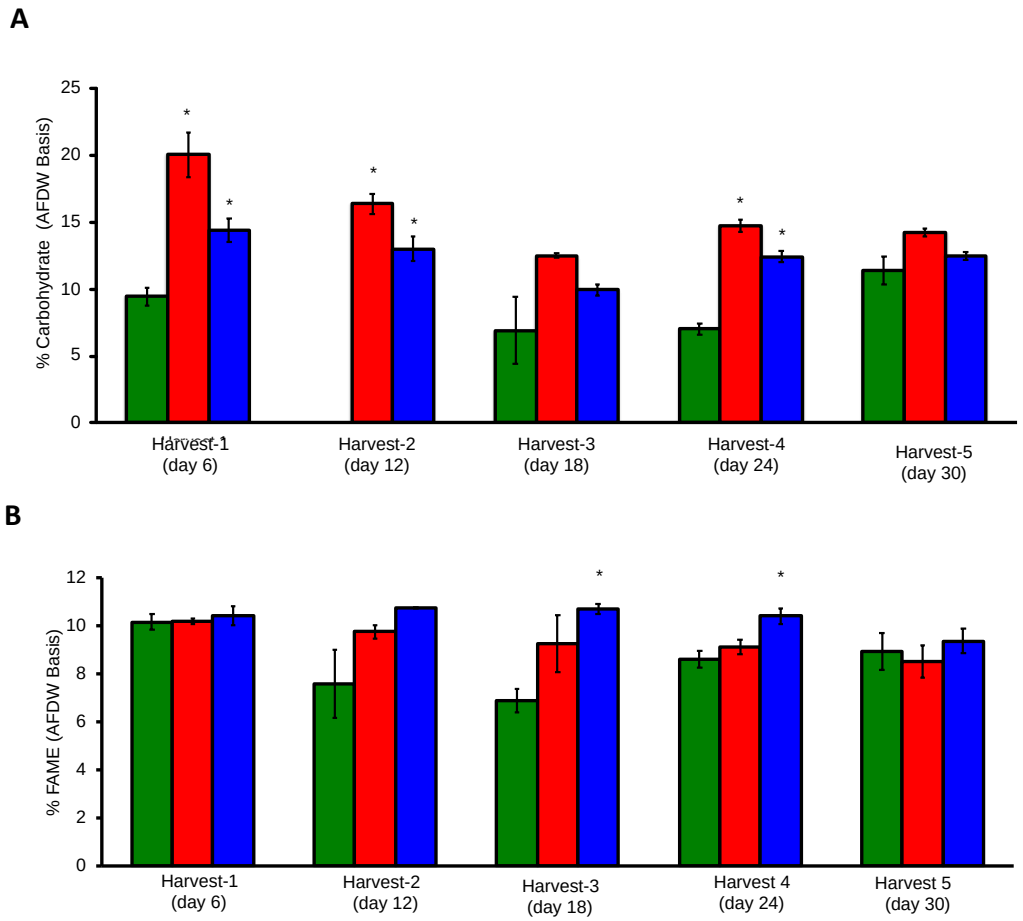


Figure **6**: SnRK lines and wt biochemical analysis conducted during the described EPBR experiment. Measurements were performed in triplicate or quadruplicate (depending on the strain). The asterisk (*) represents significant data (P-value <0.05). Gas chromatography- was used to calculate both % carbohydrate (A) and FAME (fatty methyl ester) (B). Both measurements were conducted on a ash free dry weight basis.

Discussion

Chlorella sorokiniana is a robust algal species that has attributes (especially its rapid growth and naturally high lipids) that make it a strong candidate for industrial applications that range from biodiesel feedstock to food additives. However, significant improvements in photosynthetic efficiency and biomass yield under outdoor −frequently demanding−growth conditions are required to progress

towards sustainability of a large-scale production system. Problems arise due to a number of biological limitations including inefficient conversion of solar energy into biomass along with stress conditions experienced during outdoor growth.[18] Genetic engineering has been applied towards the optimization of *Chlorella* as a biotechnological platform, but no work has overexpressed stress response genes such as SnRK genes despite analogs in plants.[18,61,61,67–71] We targeted SnRK for genetic engineering due to its potential influence on energy metabolism and stress response.

Here, we generated SnRK2 overexpression lines of *Chlorella sorokiniana* via electroporation. *Chlorella* transformation through electroporation has been shown to be easy and efficient, with the caveat that genome integration may be random and unstable.[60,61,72] In addition, the high GC content of *Chlorella sorokiniana* genome makes transgene screening by polymerase chain reaction (PCR) difficult. The top BLAST hits from the *Chlorella* and *Picochlorum* genomes were compared and at the protein level they are structurally similar while having low nucleotide similarity. As both of these organisms are part of the *Chlorellales* order we expected high protein sequence and structural similarities. These SnRKs however do differ from plant SnRKs in their overall structure. These differences include being shorter and lacking the ABA patch associated with one of the SnRK2 subfamilies.[49] Despite these structural differences we observed similar phenotypes as though found during SnRK2 overexpression.

While it has been demonstrated that SnRK2s play an important role in enhancing a plant's ability to survive various stresses including drought, osmotic, and freezing stresses the effect these kinases play on carbon metabolism has been understated. SnRK overexpression research in *Arabidopsis* has shown increases in growth under stress conditions including drought, salt, and freezing stresses. Increases in root length, above ground biomass, and total soluble sugar content has also been observed in non-stress conditions suggesting a prominent role in carbon metabolism with or without the presence of various stresses.[16,36,37,48] The increase in growth within our transgenic lines also indicates a potential role of SnRK2 in carbon metabolism in microalgae. Our SnRK overexpression lines showed increase in growth

in both low and high light intensities. One mechanism to explain increase in growth in non-stress conditions has been to elucidate a role between SnRKs and TOR (target of rapamycin). TOR kinases are evolutionary conserved master regulators that play an important role in regulating cell proliferation, development, protein synthesis, transcription and metabolism in plants.[73] It has been shown that SnRK2s in non-stress conditions can sequester signaling molecules involved in the inhibition of TOR, thus increasing its activity which would positively effect growth.[74]

The increase in growth we observed could potentially be due to an increase in photosynthetic output/efficiency we therefore analyzed various photosynthetic parameters. Further, researchers have been searching for algal transgenics with improved optical properties that will improve photosynthetic efficiency.[18] One such property is the truncation of the antenna complex. *C. sorokiniana's* antenna complex is composed of Chl a and b.[18] Where Chl b forms the periphery and is the main contributor to the size of the antenna complex. In our transgenic lines we observed a significantly lower amount of total Chl. Upon further investigation we found that SnRK2D and 2U have decreased amounts of Chl b which results in a higher Chl *a/b* ratios, both of which are indicative of smaller LHC.[6] Flash florescence was used to confirm this observation since the maximal level of fluorescence upon a pulse of saturating light is related to the number of Chl of active energy transfer to the PSII and can indicate the antenna size of the sample in question.[18] Studies demonstrate that reducing the antenna complex size improves photosynthetic efficiency, by increasing light penetration depth and enhancing photosynthetic yield by not saturating rate limiting downstream electron transfer processes.[6,18] Decreasing the antenna complex did not negatively affect photosynthetic output (O_2 yield) as we observed significantly enhanced photosynthetic yield during the linear phase of the light saturation curve. Since the samples all have saturation at similar light intensity (400 umol) it seems that a decrease in the size of antenna complex does not negatively affect output compared to the WT in terms of photooxidation (phototolerance is not affected).We further confirmed higher photosynthetic efficiency in SnRK2 lines by observing higher F_v/F_m values.[75–77] Increases in F_v/F_m have been observed in plant SnRK2 overexpression studies

when stress was induced. Therefore, we surmise that the increase in carbohydrate accumulation (carbon sink) is a direct result of improved photosynthetic efficiency (carbon source). Further analysis is needed to fully understand the mechanism behind the effects of SnRK expression on carbon metabolism and stress response.

As an approximation to outdoor growth, we assessed the capabilities of the SnRK transgenic lines by growing them in environmental photobioreactors (ePBRs), simulating a semi-continuous outdoor pond experiment. This experiment induced a simulated cold stress as *C. sorokiniana* has an optimal growth temperature at 30°C but has similar growth profiles between the range of 30-40°C.[65] Induced cold stress in *Chlamydomonas reinhardtii* has been shown to affect expression patterns of *Chlamydomonas* SnRKs along with increases in starch accumulation thus, we suspected greater carbohydrate accumulation in our transgenic lines compared to wild type cells with sustained cold stress.[53] Further, research has shown that overexpression of SnRK2s led to an increase in freezing tolerance without growth retardation in *Arabidopsis* and wheat.[36,46,78,79] Increase in survival rate percent ranged from ten to thirty percent higher than WT plants. Along with increases in survival there were also increases in growth which corresponded to increases in biomass.[36,78] As shown in Figure 5, both SnRK2 transgenic lines maintained higher growth rates and biomass productivities compared to the wild type throughout the experiment, except for the first grow out where the wild type performed slightly better relative to the SnRK2 D line. This could be due to an initial lag experienced by the transgenic line due to accumulation of starch within the cell. As shown in Figure 6A, SnRK2-D has the highest total carbohydrate content for all grow outs. Carbohydrate level of wildtype for grow out 2 is not shown due to low yield during this grow out. SnRK2-U exhibits significantly higher carbohydrate levels relative to the wildtype for harvests 1 and 4. Overall, we observe 124% and 79% (AFDW basis) average increase in total carbohydrates, for SnRK2-D and SnRK2-U, respectively compared to the wildtype. Neutral lipids were quantified as fatty acid methyl esters (FAME) by gas chromatography-flame ionization detection (GC-FID). Figure 6B shows the FAME contents determined for the two SnRK2

lines and the wildtype biomass collected from each harvest. The results indicate 11% and 22% (AFDW basis) average increase in neutral lipids for SnRK2-D and SnRK2-U, respectively. In plants, SnRK2 overexpression enhanced growth under both normal and stress conditions, by increasing biomass in roots and shoots leading to a corresponding increase in starch content.[33,36] Similarly, in our experiments, SnRK2 transgenic lines showed an increase in cell size, biomass yield and total carbohydrate content. In the future it will be important to explore the molecular roles of SnRK2s in algal carbohydrate metabolism.

Conclusion

This research demonstrates that overexpression of SnRK2 from *P. soloecismus* increases starch content, cell size, and photosynthetic efficiency in *Chlorella sorokiniana* (DOE1412). While the exact mechanism remains to be elucidated, we have demonstrated that the SnRK2 gene has an affect on photosynthetic efficiency, which increases the carbon sink strength as evidenced by the increase in starch content in our transgenic lines. Furthermore, overexpression of SnRK2 leads to an alteration of the light-harvesting complex. SnrK2 overexpression addresses one of the major bottle-necks in commercial algal growth; namely the increase of productivity yields, resulting in improved sustainability of *Chlorella sorokiniana* as a feedstock for biofuels and renewable chemicals.

Methods

Strains and culturing conditions. *Chlorella sorokiniana* (DOE1412) was obtained from Los Alamos National Laboratory (LANL). DOE 1412 was isolated by Dr. Juergen Polle at the Brooklyn College, NY.[45] Transgenic lines were maintained on HS media plates containing 25 μg/ml of Zeocin at 25°C and 50 µmoles photons $m^{-2}s^{-1}$.

Bioinformatic analysis. SnRK2 genes were identified in *C. sorokiniana* by using the ten SNRK2s from *A. thaliana* as BLASTP query sequences. The *C. sorokiniana* genome was obtained from (greenhouse.lanl.gov). The Arabidopsis SnRK2 sequences were obtained from Phytozome

(phytozome.jgi.doe.gov). We considered similar homology for e-values lower than 10^{-30} this generated an un-curated sequence list. Candidate sequences were analyzed using an Inter Pro Scan with default parameters to determine if domain structures common in *A. thaliana* were present in *C. sorokiniana*. This allowed us to filter out e sequences that did not have SnRK domain structures. Sequence alignments were generated using EMBOSS Clustal Omega and shading was conducted with (embnet.vital-it.ch).

Growth and morphology in batch mode. Cultures were grown in 250-ml Erlenmeyer flasks to the early to mid-log phase in trisacetate-phosphate (TAP) liquid media or in high salt (HS) minimal liquid media dependent on the experiment. Cells were grown in shaker flasks under constant light at an intensity of ~50 μmol m^{-2} s^{-1} photosynthetically active radiation (PAR) under constant illumination unless specified otherwise. Samples for experimentation were taken after the indicated times for each corresponding experiment. Cell counts and cell sizes were assessed using an Attune NxT acoustic focusing flow cytometer, Hausser Scientific Bright-Line hemocytometer, and Olympus BX51 microscope. Cells were also grown in environmental photobioreactors (ePBRs) with HS media supplemented with 10 mM $NaHCO_3$ with sinusoidal light conditions as indicated (Supplemental Figure 4) under diurnal light/dark cycle and sinusoidal temperature stimulating conditions in Mesa, AZ on October 1st. Semi continuous harvesting was conducted with harvesting taking place every five days.

Construction of the overexpression vector. Sequences flanking potential actin and PsaD genes were examined for potential transcription initiation and termination sites utilizing the Integrative Genome Viewer (https://www.broadinstitute.org/igv/). Promoter and terminator sequences were synthesized (Genewiz) and inserted into the *C. reinhardtii* vector pSL18 replacing the *C. reinhardtii* PsaD promoter/terminator after digestion with SapI and NotI. A codon optimized *sh-ble* gene was placed under control of the actin promoter/terminator, while the PsaD promoter/terminator is used to drive the expression of the SnRK2 gene.

Transformation by electroporation. Transformation of *C. sorokiniana* was preformed using electroporation with a Bio-Rad Gene Pulser Xcell electroporator. Cell suspensions were removed from HS liquid culture after three days of growth when cell counts were typically between 10^6 and 10^8 cells/mL. Cells were washed with HS media by centrifugation (3000RPM, five minutes) followed by two washes using 375mM sorbitol and re-suspended in 5mL of sorbitol. 300 μL of the final cell suspensions was placed in a 4-millimeter gap electroporation cuvette. Electroporation was achieved by applying three exponential-decay pules with twenty milliseconds time constants separated by thirty seconds interpulse gaps. The voltage applied was 1600 volts. Cells were allowed to recover for five minutes on ice before being plated on proteose agar plates containing the antibiotic zeocin at a concentration of 40 μg/mL. Colonies were re-plated on HS agar plates containing zeocin at a concentration 20 μg/mL for further analysis.

Gene presence and expression analysis. For screening of putative transformants, single colonies were grown on HS media plates containing 20 μg/mL of zeocin. DNA was isolated from the single colony grow ups using 6% Chelex in Tris that was boiled for thirty minutes. Isolated DNA was used to screen for the presence of zeocin and SnRK2 genes via PCR(Supplemental Figure 3).[47,48] PCR reactions were analyzed using gel electrophoreses. Positive PCR products were further confirmed by Sanger-sequencing (GeneWiz). Positive transformants were further screened for gene expression using RT-PCR and RT-qPCR. Cells were harvested by centrifugation at 3000RPM for five minutes, flash frozen in liquid nitrogen, and stored at -80°C prior beginning the RNA isolation processes. RNA was isolated using the TRIzol phenol chloroform method.[49] cDNA was prepared to a concentration of 50 ng/μL using a Clonetech RNA to cDNA EcoDry Premix. RT-PCR was performed on a T100 Bio-Rad thermal cycler and products were analyzed using gel electrophoreses. q-PCR was performed on an Agilent Technologies Mx3005P using 1 μL of actin primers that were used to target our housekeeping gene and SnRK2 1 μL gene specific primers at a concentration of 10 μM, 1μL of cDNA, and 10 μL of SYBR Green mix (Invitrogen SYBR Green ER qPCR). Water was used to bring the final volume to 20 μL.

The q-PCR protocol was performed as follows: denaturing at 95 °C for three minutes, and 40 cycles of denaturing at 95 °C for thirty seconds and annealing/elongation at 60 °C for thirty-five seconds. The specificity of PCR amplification was examined with a melting curve. Actin was used as the housekeeping gene and isolated RNA was examined for DNA contamination.

Chlorophyll content. To measure Chl content cells were grown in HS liquid media to early log phase under constant light of 300 μmol m^{-2} s^{-1}. Chl was extracted using methanol. 1 mL of culture was centrifuged at 15,000 X g for five minutes at room temperature. Pelleted cells were suspended in 100% methanol and vortexed. Samples were placed in the dark overnight. Cellular debris was pelleted using centrifugation at 15,000 X g for five minutes. Absorption was measured at 750 nm, 665.2 nm, and 652 nm using a Varian Cary 100 Bio UV/ Visible spectrometer. Calculations of total Chl, Chl A, and Chl B were performed as described previously.[50]

Measurement of chlorophyll fluorescence

For Chl fluorescence induction analysis, cell suspensions of the parent wild-type and SnRK *lines* were adjusted to a Chl concentration of ~ 2.5 μg/mL. FL-3500 fluorometer (Photon System Instruments) was used to measure quenching of Chl fluorescence. The cell suspensions were dark adapted for 15 minutes prior to the measurement. Chl fluorescence was induced using non-saturating continuous illumination. The actinic flash duration for this experiment was set to 50 μs and Chl fluorescence levels were measured every 1 μs using a weak pulse-modulated measuring flash.

Photosynthetic oxygen evolution rates. Oxygenic photosynthesis was measured using a Hansatech S1 Clark Type polarographic oxygen electrode. Cells were grown in HS liquid media to early log phase under constant light of 300 μmol m^{-2} s^{-1}. Cells were pelleted and re-suspended in 10mM sodium bicarbonate at an approximate concentration of 8 x 10^6 cells/mL. Calibration of the electrode signal was done at the beginning of each experiment using HS media sparged with air and nitrogen. Oxygen evolution was measured as a function of light intensity. The intensity was maintained for two minutes at the following eleven light intensities: 25, 50, 75, 100, 150, 300, 450, 600, 750, 850, and 1000 μmol

m^{-2} s^{-1} from a Heliospectra DYNA. Oxygen evolution rates were calculated from the change of O_2 concentration over the two-minute illumination phase. Rates were normalized based on the amount of Chl or cell count for each sample.

Photosynthetic parameters measurement using PAM Fluorometry. A Dual PAM 100 Fluorometry (Walz, Germany) was used to measure photosynthetic parameters of photosystem (II). The in-built software was used to measure photosystem (II) efficiency (Fv/Fm). Samples for both the transgenic lines and wild type were taken from cultures grown in shake flask to mid log phase. Using the internal in-built analytical software, a dark induction curve and light induction curve experiments were conducted. The values reported here are reported as an average of three replicates of the sample analyzed.

Growth in Phenometrics Environmental Photobioreactors. Phenometrics ePBRS were used to simulate climate conditions comparable to those at the AzCATI testbed site in Mesa, AZ during the month of October. The hardware portion of the ePBRs consists of a Peltier heat/cooling jackets,white LED lights , and pH feedback control using CO_2 (sparged in via solenoid in short bursts, though a glass sparger stone with pore sizes of 70-100um), and flat bottom, tapered culture vessels. EPBRs can be controlled via Phenometrics Algal Command software, which allows the user to run experiment specific parameters for culture temperature, light intensity, and pH set-points, enabling simulation of climate conditions in a particular region over a specific period of time.[80]

Chlorella Sorokiniana wild type, Chlorella Sorokiniana SnRK U, and Chlorella Sorokiniana SnRK D strains were started in ePBRS (from flask liquid cultures) with HS ammonia-based media. The cultures were allowed to grow up for five days before diluting them back, with HS media buffered with 10 mM $NaHCO_3$, and also containing Antifoam Agent A(~10ul/L), to a starting optical density (OD 750nm) of around 0.3. Throughout the experiment samples were collected daily for optical density measurements and microscopy. Cultures were harvested semi-continuously as the cultures reached stationary phase. Cultures were diluted back to the starting OD of around 0.3 after each grow/harvest

cycle, and harvested biomass was collected and lyophilized for dry weight, ash-free dry weights (AFDW), and biochemical analysis. This experiment was performed for five consecutive semi-continuous harvesting/growth cycles.

Biochemical Characterization. Lyophilized and homogenized biomass from the EPBR semi-continuous harvesting experiment were analyzed for its content of the storage macromolecules, carbohydrates and lipids. All measurement values are reported on a percent biomass basis normalized to the ash-free dry weight (AFDW). Neutral lipids in the biomass were converted into fatty acid methyl esters (FAMEs) by acid-catalyzed transesterification and analyzed by gas chromatography coupled with flame ionization detection (GC/FID).[81] 5-10 mg of the lyophilized biomass was combined with 25 µL of the 10 mg/mL internal standard (IS) which was methyl tridecanoate (C13:0ME). The biomass and the IS were then treated with 200 µL of chloroform:methanol (2:1, v/v) and 300 µL of 0.6M HCl:methanol, and incubated at 85 °C for 1 hr. FAMEs were extracted with 1 mL of hexane and the hexane layer, containing the dissolved FAMEs, was analyzed by an Agilent 7890A Series GC/FID. 2-µL injections (at a 10:1 split ratio) were loaded onto a DB-WAX column (30 m length x 0.25 mm inner diameter x 0.25 µm film thickness, (Agilent Technologies, Santa Clara, CA)). Helium was used as the carrier gas at a flow rate of 1 mL/min. The initial oven temperature was 100 °C and then increased at 25 °C/min to 200 °C, with a 1 min hold upon reaching the target temperature. Then the temperature was increased again at a rate of 1.5 °C/min, to 242 °C and held for 1 min. Inlet and detector temperatures were kept at 250 and 280 °C, respectively. The total run time of a single sample was 35 min. A GLC 461C 30-component FAME standard mix (Nu-Chek Prep, Inc., Elysian, MN) was used as the calibration standard mix and FAME quantification was performed using C13:0ME as the internal standard. A 3-methyl-2-benzothiazolinone hydrazone (MBTH) method was used to quantify the total carbohydrates in the biomass.[81] The method involved a two-step sulfuric acid hydrolysis to hydrolyze the polymeric carbohydrates in the biomass into their monomeric subunits followed by spectrophotometric quantification of the released monosaccharides. Briefly, 250µl of 72% (w/w)

sulfuric acid was added to ~25 mg of the lyophilized biomass and the samples were incubated for an hour in a 30°C (±3 °C) water bath with vortex-mixing every 5-10 minutes. Then, the samples were diluted to a concentration of 4% (w/w) H_2SO_4 with 18.2 MegaOhm water (0.2μM filtered), and then autoclaved at 121 °C for 1 hr. Samples were filtered to remove the remaining biomass solids from the sugar hydrolysate. The sugar hydrolysate was analyzed using the above mentioned MBTH method using D (+) glucose as a calibration standard.

Statistical Analysis

Unless noted all values reported are the average value of three replicates. Error is reported as the standard error. t-Test were used to test for significance difference between wild type cells and transgenic lines. P-values <0.05 represented significant data.

List of abbreviations

ABA (abcisic acid)

AFDW (ash free dry weight)

AMPK (AMP-activated protein kinase)

Chl (chlorophyll)

ePBR (electronic photobio reactor)

FAME (fatty acid methyl esters)

GC-FID (gas chromatography-flame ionization detection)

HS (high salt)

NPQ (non-photochemical quenching)

OD (optical density)

q-PCR (quantitative PCR)

RT-PCR (real time PCR

SnRK **(**sucrose non-fermenting related kinases)

SNF (sucrose non-fermenting)

TAG (triglyceride)

TAP (trisacetate phosphate)

TOR (target of rapamycin)

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

TB, CRGE, CKC, NS, TD, KP, SS, RS, AB, DH, ST, SN

Or the authors declare that they have no competing interests.

Funding

This material is based upon work supported as part of the PACE project, a DOE BETO Grant under Award # DE-EE0007089

Authors' contributions

TB, CRGE, CKC, NS, TD, KP, SS, RS, AB, DH, ST, SN

All authors read and approved the final manuscript.

Acknowledgements

XX

1. Varela Villarreal, J., Burgués, C. & Rösch, C. Acceptability of genetically engineered algae biofuels in Europe: opinions of experts and stakeholders. *Biotechnol. Biofuels* **13**, 92 (2020).
2. Slade, R. & Bauen, A. Micro-algae cultivation for biofuels: Cost, energy balance, environmental impacts and future prospects. *Biomass Bioenergy* **53**, 29–38 (2013).
3. Monari, C., Righi, S. & Olsen, S. I. Greenhouse gas emissions and energy balance of biodiesel production from microalgae cultivated in photobioreactors in Denmark: a life-cycle modeling. *J. Clean. Prod.* **112, Part 5**, 4084–4092 (2016).
4. Singh, S. P. & Singh, P. Effect of temperature and light on the growth of algae species: A review. *Renew. Sustain. Energy Rev.* **50**, 431–444 (2015).
5. Rodionova, M. V. *et al.* Biofuel production: Challenges and opportunities. *Int. J. Hydrog. Energy* **42**, 8450–8461 (2017).

6. Perrine, Z., Negi, S. & Sayre, R. T. Optimization of photosynthetic light energy utilization by microalgae. *Algal Res.* **1**, 134–142 (2012).
7. Sayre, R. Microalgae: The Potential for Carbon Capture. *BioScience* **60**, 722–727 (2010).
8. Menetrez, M. Y. An Overview of Algae Biofuel Production and Potential Environmental Impact. *Environ. Sci. Technol.* **46**, 7073–7085 (2012).
9. Georgianna, D. R. & Mayfield, S. P. Exploiting diversity and synthetic biology for the production of algal biofuels. *Nature* **488**, 329–335 (2012).
10. Kaye, Y. *et al.* Metabolic engineering toward enhanced LC-PUFA biosynthesis in Nannochloropsis oceanica: Overexpression of endogenous Δ12 desaturase driven by stress-inducible promoter leads to enhanced deposition of polyunsaturated fatty acids in TAG. *Algal Res.* **11**, 387–398 (2015).
11. Kang, N. K. *et al.* Increased lipid production by heterologous expression of AtWRI1 transcription factor in Nannochloropsis salina. *Biotechnol. Biofuels* **10**, 231 (2017).
12. Chowdhury, H. & Loganathan, B. Third-generation biofuels from microalgae: a review. *Curr. Opin. Green Sustain. Chem.* **20**, 39–44 (2019).
13. Long, S. P., Zhu, X.-G., Naidu, S. L. & Ort, D. R. Can improvement in photosynthesis increase crop yields? *Plant Cell Environ.* **29**, 315–330 (2006).
14. Ort, D. R. *et al.* Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 8529–8536 (2015).
15. Niinemets, Ü. *et al.* Photosynthesis: ancient, essential, complex, diverse … and in need of improvement in a changing world. *New Phytol.* **213**, 43–47 (2017).
16. Shin, R., Alvarez, S., Burch, A. Y., Jez, J. M. & Schachtman, D. P. Phosphoproteomic identification of targets of the Arabidopsis sucrose nonfermenting-like kinase SnRK2.8 reveals a connection to metabolic processes. *Proc. Natl. Acad. Sci.* **104**, 6460–6465 (2007).

17. Yang, B. *et al.* Genetic engineering of the Calvin cycle toward enhanced photosynthetic CO2 fixation in microalgae. *Biotechnol. Biofuels* **10**, 229 (2017).
18. Cazzaniga, S. *et al.* Domestication of the green alga Chlorella sorokiniana: reduction of antenna size improves light-use efficiency in a photobioreactor. *Biotechnol. Biofuels* **7**, 157 (2014).
19. Gimpel, J. A., Specht, E. A., Georgianna, D. R. & Mayfield, S. P. Advances in microalgae engineering and synthetic biology applications for biofuel production. *Curr. Opin. Chem. Biol.* **17**, 489–495 (2013).
20. Sharon-Gojman, R., Leu, S. & Zarka, A. Antenna size reduction and altered division cycles in self-cloned, marker-free genetically modified strains of Haematococcus pluvialis. *Algal Res.* **28**, 172–183 (2017).
21. Banerjee, A., Banerjee, C., Negi, S., Chang, J.-S. & Shukla, P. Improvements in algal lipid production: a systems biology and gene editing approach. *Crit. Rev. Biotechnol.* **38**, 369–385 (2018).
22. Banerjee, S., Banerjee, S., Ghosh, A. K. & Das, D. Maneuvering the genetic and metabolic pathway for improving biofuel production in algae: Present status and future prospective. *Renew. Sustain. Energy Rev.* **133**, 110155 (2020).
23. Vanthoor-Koopmans, M., Wijffels, R. H., Barbosa, M. J. & Eppink, M. H. M. Biorefinery of microalgae for food and fuel. *Bioresour. Technol.* **135**, 142–149 (2013).
24. Davis, R., Kinchin, C., Markham, J., Tan, E. C. D. & Laurens, L. M. Process Design and Economics for the Conversion of Algal Biomass to Biofuels: Algal Biomass Fractionation to Lipid- and Carbohydrate-Derived Fuel Products. *Renew. Energy* 110 (2014).
25. 2016 National Algal Biofuels Technology Review | Department of Energy. https://www.energy.gov/eere/bioenergy/downloads/2016-national-algal-biofuels-technology-review.

26. Neale, P. J. & Melis, A. Algal Photosynthetic Membrane Complexes and the Photosynthesis-Irradiance Curve: A Comparison of Light-Adaptation Response in Chlamydomonas Reinhardtii (chlorophyta). *J. Phycol.* **22**, 531–538 (1986).
27. de Mooij, T. *et al.* Antenna size reduction as a strategy to increase biomass productivity: a great potential not yet realized. *J. Appl. Phycol.* **27**, 1063–1077 (2015).
28. Negi, S. *et al.* Light regulation of light-harvesting antenna size substantially enhances photosynthetic efficiency and biomass yield in green algae†. *Plant J.* **103**, 584–603 (2020).
29. Saini, N., Pal, K., Sujata, Deepak, B. & Mona, S. Thermophilic algae: A new prospect towards environmental sustainability. *J. Clean. Prod.* **324**, 129277 (2021).
30. Saad, M. G., Dosoky, N. S., Zoromba, M. S. & Shafik, H. M. Algal Biofuels: Current Status and Key Challenges. *Energies* **12**, 1920 (2019).
31. Chen, H., Li, T. & Wang, Q. Ten years of algal biofuel and bioproducts: gains and pains. *Planta* **249**, 195–219 (2019).
32. Shuba, E. S. & Kifle, D. Microalgae to biofuels: 'Promising' alternative and renewable energy, review. *Renew. Sustain. Energy Rev.* **81**, 743–755 (2018).
33. Rizwan, M., Mujtaba, G., Memon, S. A., Lee, K. & Rashid, N. Exploring the potential of microalgae for new biotechnology applications and beyond: A review. *Renew. Sustain. Energy Rev.* **92**, 394–404 (2018).
34. Suparmaniam, U. *et al.* Insights into the microalgae cultivation technology and harvesting process for biofuel production: A review. *Renew. Sustain. Energy Rev.* **115**, 109361 (2019).
35. Barajas-Lopez, J. de D. *et al.* Upstream kinases of plant SnRKs are involved in salt stress tolerance. *Plant J.* **93**, 107–118.

36. Zhang, H., Mao, X., Wang, C. & Jing, R. Overexpression of a Common Wheat Gene TaSnRK2.8 Enhances Tolerance to Drought, Salt and Low Temperature in Arabidopsis. *PLOS ONE* **5**, e16041 (2010).

37. Zheng, Z. *et al.* The Protein Kinase SnRK2.6 Mediates the Regulation of Sucrose Metabolism and Plant Growth in Arabidopsis[W][OA]. *Plant Physiol.* **153**, 99–113 (2010).

38. Wu, P., Wang, W., Duan, W., Li, Y. & Hou, X. Comprehensive Analysis of the CDPK-SnRK Superfamily Genes in Chinese Cabbage and Its Evolutionary Implications in Plants. *Front. Plant Sci.* **8**, (2017).

39. Hardie, D. G. The AMP-activated protein kinase pathway – new players upstream and downstream. *J. Cell Sci.* **117**, 5479–5487 (2004).

40. Zhang, M. *et al.* AMP-activated Protein Kinase α1 Promotes Atherogenesis by Increasing Monocyte-to-macrophage Differentiation. *J. Biol. Chem.* jbc.M117.779447 (2017) doi:10.1074/jbc.M117.779447.

41. Robaglia, C., Thomas, M. & Meyer, C. Sensing nutrient and energy status by SnRK1 and TOR kinases. *Curr. Opin. Plant Biol.* **15**, 301–307 (2012).

42. Emanuelle, S., Doblin, M. S., Stapleton, D. I., Bacic, A. & Gooley, P. R. Molecular Insights into the Enigmatic Metabolic Regulator, SnRK1. *Trends Plant Sci.* **21**, 341–353 (2016).

43. Dong Xue-Fei *et al.* The SnRK Protein Kinase Family and the Function of SnRK1 Protein Kinase. *Int. J. Agric. Biol.* **14**, 196–200 (2012).

44. Su, Y. *et al.* SnRK2 Homologs in Gossypium and GhSnRK2.6 Improved Salt Tolerance in Transgenic Upland Cotton and Arabidopsis. *Plant Mol. Biol. Report.* **35**, 442–456 (2017).

45. Xie, T. *et al.* Molecular Mechanism for Inhibition of a Critical Component in the Arabidopsis thaliana Abscisic Acid Signal Transduction Pathways, SnRK2.6, by Protein Phosphatase ABI1. *J. Biol. Chem.* **287**, 794–802 (2012).

46. Mao, X., Zhang, H., Tian, S., Chang, X. & Jing, R. TaSnRK2.4, an SNF1-type serine/threonine protein kinase of wheat (Triticum aestivum L.), confers enhanced multistress tolerance in Arabidopsis. *J. Exp. Bot.* **61**, 683–696 (2010).

47. Coello, P., Hey, S. J. & Halford, N. G. The sucrose non-fermenting-1-related (SnRK) family of protein kinases: potential for manipulation to improve stress tolerance and increase yield. *J. Exp. Bot.* **62**, 883–893 (2011).

48. Kulik, A., Wawer, I., Krzywińska, E., Bucholc, M. & Dobrowolska, G. SnRK2 Protein Kinases—Key Regulators of Plant Response to Abiotic Stresses. *OMICS J. Integr. Biol.* **15**, 859–872 (2011).

49. Colina, F. *et al.* Genome-wide identification and characterization of CKIN/SnRK gene family in Chlamydomonas reinhardtii. *Sci. Rep.* **9**, 1–16 (2019).

50. Davies, J. P., Yildiz, F. H. & Grossman, A. R. Sac3, an Snf1-like Serine/Threonine Kinase That Positively and Negatively Regulates the Responses of Chlamydomonas to Sulfur Limitation. *Plant Cell* **11**, 1179–1190 (1999).

51. Gonzalez-Ballester, D., Pollock, S. V., Pootakham, W. & Grossman, A. R. The Central Role of a SNRK2 Kinase in Sulfur Deprivation Responses. *Plant Physiol.* **147**, 216–227 (2008).

52. González-Ballester, D. *et al.* RNA-Seq Analysis of Sulfur-Deprived Chlamydomonas Cells Reveals Aspects of Acclimation Critical for Cell Survival. *Plant Cell* **22**, 2058–2084 (2010).

53. Valledor, L., Furuhashi, T., Hanak, A.-M. & Weckwerth, W. Systemic Cold Stress Adaptation of Chlamydomonas reinhardtii. *Mol. Cell. Proteomics* **12**, 2032–2047 (2013).

54. Valledor, L., Furuhashi, T., Recuenco-Muñoz, L., Wienkoop, S. & Weckwerth, W. System-level network analysis of nitrogen starvation and recovery in Chlamydomonas reinhardtii reveals potential new targets for increased lipid accumulation. *Biotechnol. Biofuels* **7**, 171 (2014).

55. Sato, A., Matsumura, R., Hoshino, N., Tsuzuki, M. & Sato, N. Responsibility of regulatory gene expression and repressed protein synthesis for triacylglycerol accumulation on sulfur-starvation in Chlamydomonas reinhardtii. *Front. Plant Sci.* **5**, (2014).

56. Rasala, B. A. *et al.* 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**, e43349 (2012).

57. Rasala, B. A. *et al.* Expanding the spectral palette of fluorescent proteins for the green microalga Chlamydomonas reinhardtii. *Plant J.* **74**, 545–556 (2013).

58. Doron, L., Segal, N. & Shapira, M. Transgene Expression in Microalgae—From Tools to Applications. *Front. Plant Sci.* **7**, (2016).

59. Fischer, N. & Rochaix, J. D. The flanking regions of PsaD drive efficient gene expression in the nucleus of the green alga Chlamydomonas reinhardtii. *Mol. Genet. Genomics MGG* **265**, 888–894 (2001).

60. Run, C. *et al.* Stable nuclear transformation of the industrial alga Chlorella pyrenoidosa. *Algal Res.* **17**, 196–201 (2016).

61. Yang, B., Liu, J., Jiang, Y. & Chen, F. Chlorella species as hosts for genetic engineering and expression of heterologous proteins: Progress, challenge and perspective. *Biotechnol. J.* **11**, 1244–1261 (2016).

62. Takeshita, T., Takeda, K., Ota, S., Yamazaki, T. & Kawano, S. A Simple Method for Measuring the Starch and Lipid Contents in the Cell of Microalgae. *Cytologia (Tokyo)* **80**, 475–481 (2015).

63. Ort, D. R., Zhu, X. & Melis, A. Optimizing Antenna Size to Maximize Photosynthetic Efficiency. *Plant Physiol.* **155**, 79–85 (2011).

64. Li, T. *et al.* Assessment of photosynthesis regulation in mixotrophically cultured microalga Chlorella sorokiniana. *Algal Res.* **19**, 30–38 (2016).

65. Kumar, K., Dasgupta, C. N. & Das, D. Cell growth kinetics of Chlorella sorokiniana and nutritional values of its biomass. *Bioresour. Technol.* **167**, 358–366 (2014).

66. Cuaresma Franco, M., Buffing, M. F., Janssen, M., Vílchez Lobato, C. & Wijffels, R. H. Performance of Chlorella sorokiniana under simulated extreme winter conditions. *J. Appl. Phycol.* **24**, 693–699 (2012).

67. Přibyl, P., Cepák, V. & Zachleder, V. Production of lipids in 10 strains of Chlorella and Parachlorella, and enhanced lipid productivity in Chlorella vulgaris. *Appl. Microbiol. Biotechnol.* **94**, 549–561 (2012).

68. Norashikin, M. N., Loh, S. H., Aziz, A. & Cha, T. S. Metabolic engineering of fatty acid biosynthesis in Chlorella vulgaris using an endogenous omega-3 fatty acid desaturase gene with its promoter. *Algal Res.* **31**, 262–275 (2018).

69. Alishah Aratboni, H., Rafiei, N., Garcia-Granados, R., Alemzadeh, A. & Morones-Ramírez, J. R. Biomass and lipid induction strategies in microalgae for biofuel production and other applications. *Microb. Cell Factories* **18**, 178 (2019).

70. Spencer-Milnes, X. Towards engineering the microalga Chlorella sorokiniana for the production of tailored high-value oils. *Doctoral thesis, UCL (University College London).* (UCL (University College London), 2019).

71. Lin, W.-R. *et al.* Enhancing carbon capture and lipid accumulation by genetic carbonic anhydrase in microalgae. *J. Taiwan Inst. Chem. Eng.* **93**, 131–141 (2018).

72. Mosey, M., Douchi, D., Knoshaug, E. P. & Laurens, L. M. L. Methodological review of genetic engineering approaches for non-model algae. *Algal Res.* **54**, 102221 (2021).

73. Mao, X. *et al.* The Sucrose Non-Fermenting 1-Related Protein Kinase 2 (SnRK2) Genes Are Multifaceted Players in Plant Growth, Development and Response to Environmental Stimuli. *Plant Cell Physiol.* **61**, 225–242 (2020).

74. Belda-Palazón, B. *et al.* A dual function of SnRK2 kinases in the regulation of SnRK1 and plant growth. *Nat. Plants* **6**, 1345–1353 (2020).

75. Murchie, E. H. & Lawson, T. Chlorophyll fluorescence analysis: a guide to good practice and understanding some new applications. *J. Exp. Bot.* **64**, 3983–3998 (2013).

76. White, S., Anandraj, A. & Bux, F. PAM fluorometry as a tool to assess microalgal nutrient stress and monitor cellular neutral lipids. *Bioresour. Technol.* **102**, 1675–1682 (2011).

77. Santabarbara, S. *et al.* Comparative excitation-emission dependence of the FV/FM ratio in model green algae and cyanobacterial strains. *Physiol. Plant.* **166**, 351–364 (2019).

78. Zhang, H., Mao, X., Jing, R., Chang, X. & Xie, H. Characterization of a common wheat (Triticum aestivum L.) TaSnRK2.7 gene involved in abiotic stress responses. *J. Exp. Bot.* **62**, 975–988 (2011).

79. Tian, S. *et al.* Cloning and characterization of TaSnRK2.3, a novel SnRK2 gene in common wheat. *J. Exp. Bot.* **64**, 2063–2080 (2013).

80. Lucker, B. F., Hall, C. C., Zegarac, R. & Kramer, D. M. The environmental photobioreactor (ePBR): An algal culturing platform for simulating dynamic natural environments. *Algal Res.* **6**, 242–249 (2014).

81. Laurens, L. M. L., Quinn, M., Van Wychen, S., Templeton, D. W. & Wolfrum, E. J. Accurate and reliable quantification of total microalgal fuel potential as fatty acid methyl esters by in situ transesterification. *Anal. Bioanal. Chem.* **403**, 167–178 (2012).

Part 3

Characterization of sucrose nonfermenting-1-related kinase (SnRK) gene diversity across eukaryotic algal linages

Taylor Britton[1] and Blake Hovde[2, *]

1. Department of Biology, The University of New Mexico, Albuquerque, NM, 87104, USA

2. Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

Introduction

Algae are a diverse group of organisms with promising characteristics for biotechnology, including biofuels, nutraceuticals, food and other bioproduct feedstock applications.[1,2] Microalgae are also currently used as nutritional supplements and food colorants, and they are cultivated for high value products such as polyunsaturated fatty acids, carotenoids, phycobiliproteins, polysaccharides, and antioxidants.[2] With the ever increasing need for carbon neutral products, an increase in biomass production and photosynthetic output will become more of a necessity. These increases are likely to come from isolation and cultivation of new algal species/strains, improved culturing methods through engineering and genetic modification. Towards the latter point, improved curation of genes involved in the production and regulation of biomass production and photosynthesis may provide improved genetic engineering targets for overall algal biomass production in new and existing algal strains of interest. Efforts in enhancing plant productivity through overexpression of stress response mechanisms has been shown.[3–5]

During outdoor cultivation, microalgae cultures are constantly under pressure from multiple stressors that can lead to decreases in biomass accumulation. Identifying metabolic and regulatory networks involved in stress responses is important in minimizing losses to biomass accumulation.[1] Protein kinases involved in metabolic response to abiotic and cellular stresses are conserved amongst all eukaryotes including the sucrose nonfermenting-1 (SNF1) in yeast and AMP-activated protein kinases (AMPK) in mammals.[1,6,7] AMPKs exists as heterotrimeric complexes composed of one catalytic subunit (alpha), and two regulatory subunits (beta and gamma).[8] The structure of these kinases differs slightly in plants, where the two regulatory domains display equivalent beta gamma proteins and gamma-like subunits.[1] In plants, sucrose

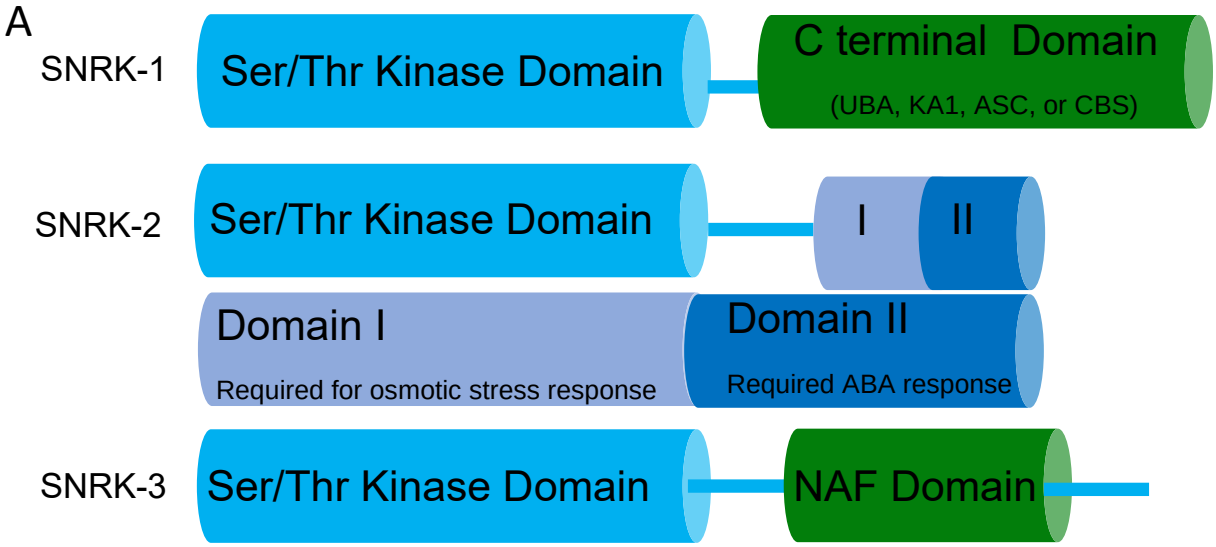
nonfermenting-1-related kinases (SnRK) have emerged as central components that interface stress signaling and energy metabolism.[1] Plant SnRKs can be subdivided into the three distinct families with all three having a conserved Serine/Threonine kinase domain (Figure 1).[1] However, the C-terminus differs amongst the three unique family members that make up the SnRK family.[1] SnRK1s contain UBA (ubiquitin-associated) and KA1 (kinase associated) domains, SnRK2s contain osmotic stress activation domains, and SnRK3s contain a NA/FISL domain and form a unique group.[1,7] While each of these family members are involved in stress singling and energy metabolism, their activation and downstream targets differ.

SnRK1s were first discovered in yeast, and orthologs of the regulatory sub-units have been discovered in the bacteria and archaea kingdoms.[1] SnRK1 family members are the most closely related to AMPKs and SNF1s and share approximately 47% amino acid sequence identity.[9] SnRK1s have been identified as key contributors to the regulation of global metabolism and energy status in plants in response to low glucose/high sucrose levels, dark period, hypoxia, salinity, pathogens, or herbivore attack.[1,6,10] In plants, SnRK1s will respond to these stress conditions by redirecting energy from growth and development to stress tolerance.[11] SnRK1s work by channeling carbon into the biosynthetic pathway through the modulation of sucrose synthase and ADP-glucose phyrophosphorylase.[9,12] Regulation of these processes through SnRK1 occurs at the transcriptional and post transcriptional levels in response to stress conditions.[6,9,13] It is thought that transcription reprograming is mediated by modifying transcription and chromatin remodeling. SnRK1s are regulated by high energy signals including the presence of trehalose-6-phosphate and sucrose.[11,13] SnRK1s and SnRK2s work together upstream during stress conditions to repress TOR signaling. TOR is a growth-stimulating and high energy activated kinase that has direct control of protein synthesis,

autophagy, and growth process in all eukaryotic organism.[11] It has also been found that during TOR repression SnRK1 can regulate some typical TOR target processes including autophagy.[11]

SnRK2s were first widely identified as important regulators of response to abiotic stresses in *Nicotiana tabacum*.[6] Additional research has shown that SnRK2 family members play an important role in ABA(abscisic acid)-mediated signaling pathways in response to abiotic stress.[3,7,14] In plants however, SnRK2s play many other roles, including the regulation of seed dormancy, germination, maturation, seedling growth, flowering time, and stomatal movement during drought stress.[14] Several overexpression studies have been conducted in various plant species that demonstrate increases in drought and salt tolerance and other abiotic stresses. This indicates that these genes are important in multi-environmental stress response and all of these family members have the potential to improve abiotic stress tolereance.[14] Several studies also show increases in sucrose and fatty acid metabolism in well-watered conditions suggesting that SnRK2s also plays an role in carbon metabolism.[14] Responses to these stresses and changes in energy level depend on the individual SnRK2 family member. SnRK2s can be divided into three groups: **group1** is not activated by ABA, **group 2** is not activated or activated weakly by ABA and **group 3** is strongly activated by ABA (Figure 1).[6,14] All of the SnRK2 family members consist of a highly conserved N-terminal kinase domain, which is related to those of SNF1 and AMPK in yeast and mammals respectively. The C-terminal domains of SnRK2s consist of two subdomains and is where the difference between family members lie. Domain 1 is universal of all SnRK2s and is required for activation of osmotic stress independently of ABA. Domain 2 is specific to ABA-dependent SnRK2s.[6]

SnRK3s are closely related to calcium dependent protein kinases (CDPKs).[9] It is thought that calcineurin B-like (CBL)- interacting protein kinases (CIPKS), also known and referred from here on as SnRK3, may be CDPKs themselves since they interact with CDPK.[10,15] Like SnRK1 and SNRK2, the SNRK3 subfamily is also involved in stress responses.[6,9,10] Calcium ions play an important role in a cell's response to stress conditions.[6,10] CDPKs and SnRK3s respond to and recognize specific calcium signatures induced from changes in environmental conditions leading to downstream effects including stomatal opening and closing and the activation of various membrane transporters or ion channels.[10,16] The overall domain structure of SnRK3 is similar to SnRK1s and 2s with kinase domains located near the N-terminus followed by C-terminal regulatory domains.[10] The C-terminal domain is characterized by the presence of a NAF domain (a novel 24 amino acid protein interaction domain) which is required for the interaction with CBL proteins in a calcium dependent manner.[15,17] The NAF domain interacts with CBL proteins via a hydrophobic helix.[15] SnRK3s interacts with various calcium binding proteins including CDPKs and CBLs.[6] In Arabidopsis SnRK3 family members can regulate plant growth, stomatal opening and closing, and seed germination under ABA treatment and are vital in the plant response to drought, cold, ABA, sugar, salt, and pH changes.[6,10]



B

SnRK Type	SnRK Function
SnRK1	Global energy regulator, interacts with bZip and is responsible for global changes in transcriptome and growth adaptation in response to energy deficiency (Jamsheer)
SnRK2	Important signaling network that responses to stress including drought, salt, and cold stress
SnRK3	SnRK3s or CIPKS are protein kinases that interact with CBL-type Ca2+ sensors. The interaction of CBLs and CIPKs make an important signaling network that regulates plant stress responses. (Tang)

	ABA activation
Group 1	-
Group 2	+
Group 3	+++

Figure 1: (A) The basic structure of each SnRK family member identified from vascular and seeded plants. Each family member consists of a conversed Ser/Thr kiase domains while the regulatory domain differs amongst each family member. (B) The function of each SnRK family member with the activation method of SnRK2s.

The ability to deal with stress is thought to be one of the critical steps in plant evolution, which can be traced from green algae, mosses, and ferns.[18] However, it is becoming evident that many of the stress response genes such as SnRKs are also being identified in ancestral lineages of plants. SnRK1s have been found across all eukaryotic lineages with subunits of the kinases also identified in prokaryotes.[1] The identification of SnRK2 and SnRK3 outside of plants has

been slower. Previous studies have postulated that plant SnRK2 and SnRK3 originated from a duplication event of SnRK and subsequently diverged during plant evolution - allowing for the plants to develop networks linking abiotic stress and ABA signaling with metabolic singaling.[7] Although research indicates that SnRK2-mediated ABA signaling most likely originated in land plants, SnRK2 homologs have been identified in seedless plants, including the mosses such as *Physcomitrella patens* and the fern-ally *Selaginella moellendroffii* as well as the green algae *Chlamydomonas*.[18–20] Seed plant SnRK2s differed from those found in mosses because theSnRK2s that were identified only belong to the SnRK2 group 3 (strongly activated by ABA). In addition, green algae also display distinct SnRK2s compared to higher plants due to differences in the C-terminal region.[20] These studies conclude that group 3 is the ancestral form, with group 1 being the most recent.[18,19] SnRK3 kinases have been found in all seed plants whose genomes have been sequenced as well as protozoan eukaryotes.[6,19] It was generally assumed that SnRK3 genes were specific to seed plants because of their absence in several green algal lineages including *Chlamydomonas* and *Volvox,* but recent studies have identified SnRK3 in ferns, mosses, and a few algal lineages including *Osterococcus* indicating an ancient origin.[15,21] SnRK3 genes were absent from *Porphyra* genome as well as several other red algal genomes. However, through phylogenetic analysis, SnRKs were identified in red algae and grouped into two clades that included *Arabidopsis* SnRK1 and SnRK3.[21] The authors suggest that while highly speculative, the identification of SnRKs in red algae suggest they play an important role in stress response, as they do in plants.[21]

While SnRKs have been identified as playing important roles in plant stress response, very little is still known about their role outside of the seed plant lineage including mosses and algae (Figure 2).[22] A widespread survey of SnRK identification across algal lineages has yet to be

shown. To remedy this shortcoming and shed light on the breadth of SnRK diversity in across eukaryotic algal linages, we identified SnRKs in thirty-four algal species and analyzed their SNRK domain motifs and phylogenetic relationships. During this analysis we found that SnRK1s are present across all algal lineages, SnRK2s were only found in the green algae, and SnRK3s were only found in algae with terrestrial species. A specific SnRK3 that was identified is one that activates a sodium/hydrogen exchanger that send excess sodium from cell into the soil suggesting that SnRK3s potentially played a role in the colonization of land.

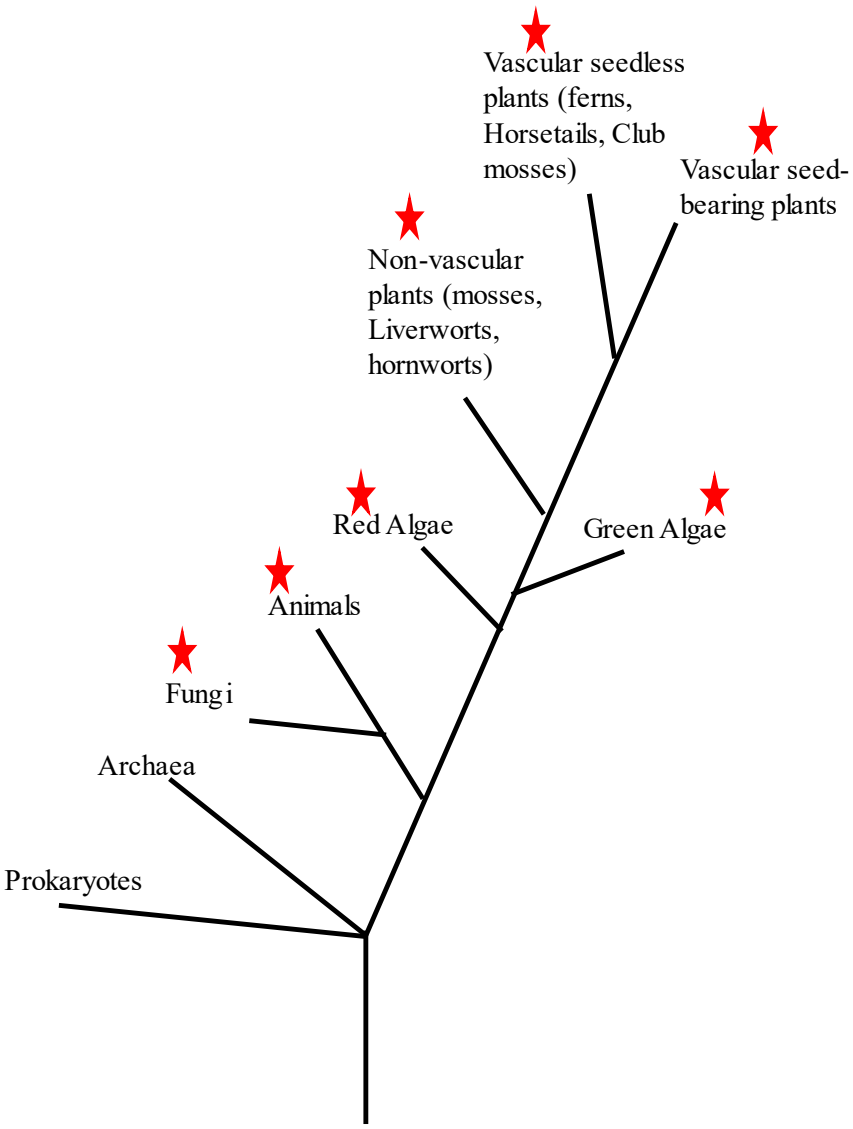


Figure 2: Basic phylogenetic tree. The red stars indicate where a SnRK or a version of a SnRK has been identified.

Material and Methods

Retrieval of sequences

Annotated genomes were downloaded from various databases as follows: Los Alamos National Laborites Greenhouse (greenhouse.lanl.gov), Joint Genome Institute (https://jgi.doe.gov/), Marine Genomics Unit (https://groups.oist.jp/mgu), REEF Genomics (http://reefgenomics.org/), Phytozome (https://phytozome.jgi.doe.gov/), and other academic websites as shown in Table 1.

Blast and Sequence Analysis

The *Arabidopsis* SnRK1.2, 2.8, 2.10, 3.1, 3.16, and 3.24 protein sequences were used as BLASTp queries to identify SnRKs in our thirty-four chosen algal species. Only sequence matches with an e-value threshold e^{-25} or lower were employed in this study. This approach generated a large list of potential SnRKs in each organism queried as expected. The Interpro protein classification tool (https://www.ebi.ac.uk/interpro/) was used to identify protein domains associated with SnRK1, 2, and 3. The PANTHER analysis utilized in this tool was used to identify domains. We first generated a list of all the *Arabidopsis* SnRKs as a comprehensive comparison. We used this same method to remove proteins that did not have canonical SnRK domain structure from the initial BLASTp for each algal species queried.

Phylogenetic Analysis and Motif Identification

Protein sequences with the characteristic SnRK domains were aligned with Muscle and manually trimmed using Jalview.[23,24] The phylogenetic tree was constructed using MEGA X.[25] Maximum-likelihood (ML) trees were constructed with 1000 bootstrap replications

(Figure 3). Genetic distance was also calculated with MEGA X. Motifs within each SnRK gene collected were identified by MEME with an optimum motif width of 6-200 bp and a motif maximum number set at 20 this value was determined based on previous studies observing SnRK motifs.[26]

Results

Identification of SnRK orthologs in thirty-four algal species

Initial BLAST searches against thirty-four algal genomes using *Arabidopsis* SnRK protein sequences as queries. This generated a large list of un-curated BLAST hits for each algal species. SnRKs along with other proteins with conserved Ser/Thr kinase domain were identified in this initial search. BLAST hits with an e-value $<10^{-25}$ were used for further analysis. Protein domain identification using the software Interpro was used to determine the distinction between SnRKS and other proteins identified as described in the material and methods. Domains specifically present in the SnRK family were searched for during this analysis which included the domains UBA, KA1, and NAF. The protein domain curation resulted in the identification of 352 putative SnRKs in the thirty-four algal species including 82 SnRK1s, 229 SnRK2s, and 41 SnRK3s. Polypeptides ranged in length from 256 to 1602 amino acids.

Interpro domains identified included PTHR24343 which corresponds to the SnRK catalytic subunit, along with the SnRK1 regulatory domains KA1 and ubiquitin associated domains, and the NAF domain associated with SnRK3s. Based on domain analysis SnRK2s fell into ten unique groups, with 2 groups making the majority of the SnRKs identified these

included ones corresponding to domains associated with group 2 SnRK2s. There were twenty-three matches that corresponded to group 3 SnRK2s.

Muscle alignments of the identified SnRKs were used to generate phylogenetic trees using MEGAX. A maximum likelihood tree was generated using the Dayhoff matrix(250 PAM matrix) and 1000 replicate bootstrap analyses. In Figure (X) we show the tree generated with all SNRK hits. With Figure 3 (Supplemental Figures 6,7, and 8) showing trees for each individual SnRK family member. SnRK1s grouped similarly to thatfound in terms of the 5 eukaryotic super groups. SnRK2s were only found in green algae and broadly diverged based on their corresponding sub-family. SnRK3s differed based on the absence of the NAF domain and species type. To further investigate phylogenetic relationships Markov Chain Monte Carlo (MCMC) in MrBayes version 3.2.7a was used to address potential bootstrap value issues(Supplemental Figure 5).[27,28]

Table 1. Algal genera used in this study sorted and color coded by phylogenetic clade.

Species Name	Order , Class
Symbiodinium goreaui	Suessiales, Dinophyceae[29,30]
Cryptophyceae sp.	NA, Cryptophyceae[31,32]
Euglena gracilis	Euglenales, Euglenoidea[33]
Cyanophora paradoxa	Glaucocystales, Glaucophyceae[34]
Phaeocystis globosa	Phaeocystales, Haptophyceae[32,35]
Emiliania Huxley	Isochrysidales, Haptophyceae[36]

Bigelowiella natans	Chlorarachniales, Chlorarachniophyceae[37]
Pyropia yezoensis	Bangiales,Bangiophyceae[38]
Galdieria sulphuraria	Cyanidiales,Cyanidiophyceae[39]
Chondrus crispus	Gigartinales,Florideophyceae[40]
Pseudo-nitzschia multiseries	Bacillariales, Bacillariophyceae[32]
Nannochloropsis oceanica	Eustigmatales,Eustigmatophyceae[41]
Microchloropsis salina	Eustigmatales, Eustigmatophyceae[42]
Nemacystus decipiens	Ectocarpales, Phaeophyceae[43]
Ectocarpus siliculosus	EctocarpalesPhaeophyceae[44]
Tetraselmis striata	Chlorodendrales, Chlorodendrophyceae[45]
Chlamydomonas reinhardtii	Chlamydomonadales, Chlorophyceae[46]
Dunaliella salina	Chlamydomonadales, Chlorophyceae[32,47]
Scenedesmus sp	Sphaeropleales, Chlorophyceae[48]
Chloropicon primus	Chloropicales, Chloropicophyceae[49]
Ostreococcus tauri	Mamiellales, Mamiellophyceae[50,51]
Bathycoccus prasinos	Mamiellales, Mamiellophyceae[52,53]
Caulerpa lentillifera	Bryopsidales, Ulvophyceae[54]

Chlorella sorokiniana	Chlorellales, Trebouxiophyceae[55]
Picochlorum soloecismus	NA, Trebouxiophyceae[56]
Asterochloris	Trebouxiales, Trebouxiophyceae[57]
Symbiochloris reticulata	Trebouxiales, Trebouxiophyceae[58]
Cocomyxa subellipsoidae	Choricystis/Brotococcus, Trebouxiophyceae[32]
Helicosporidium	Chlorellales, Trebouxiophyceae[59]
Auxenochlorella protothecoides	Chlorellales, Trebouxiophyceae[60]
Klebsormidium nitens	Klebsormidales, Klebsormidiophyceae[61]
Chara braunii	Charales, Charophyceae[32,62]
Mesotaenium endlicherianum	Zygnematales, Zygnematophyceae[63]
Spirogloea muscicola	Spirogloeales, Zygnematophyceae[63]

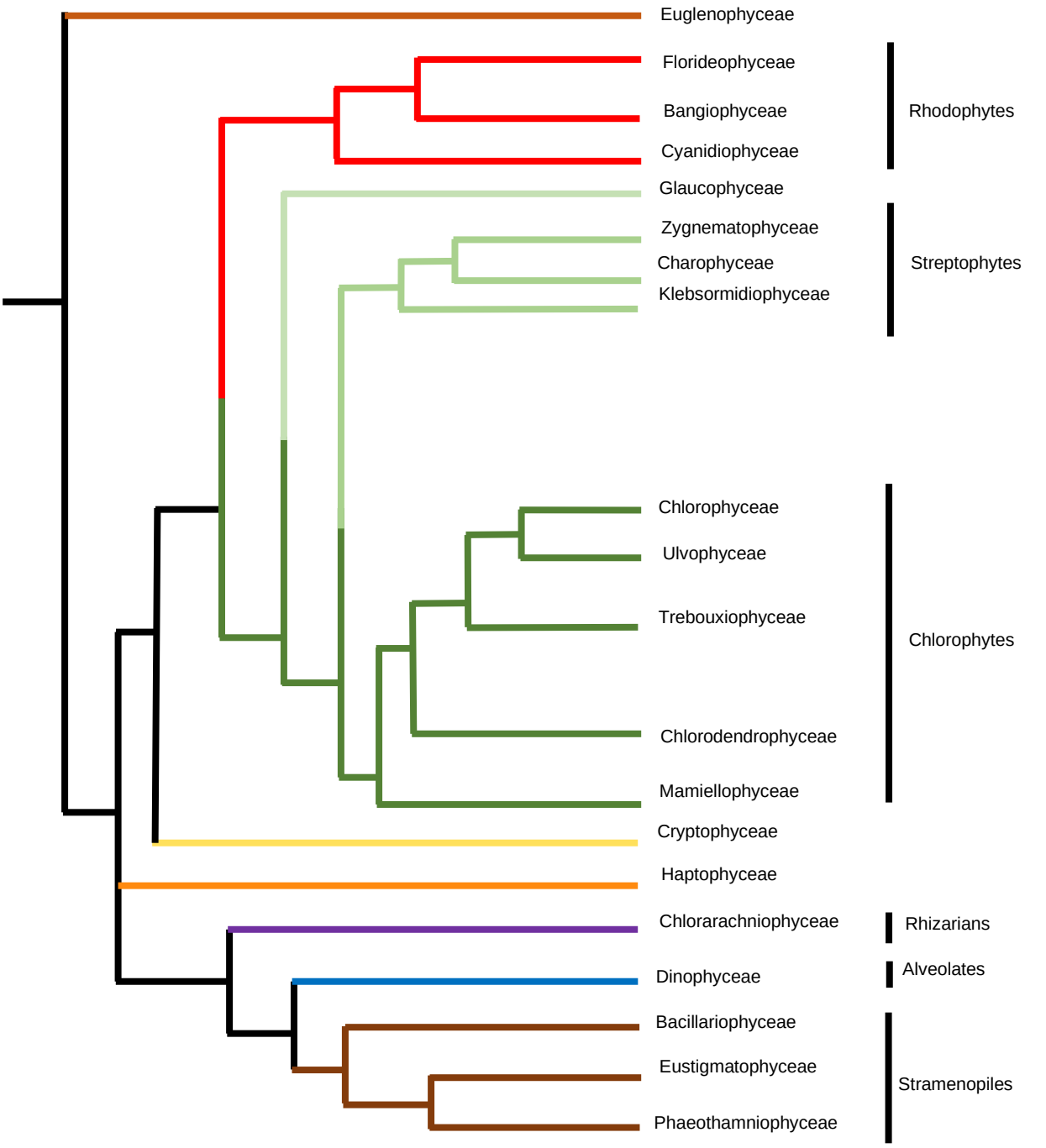


Figure 3: Phylogenetic tree of the algal clades used in this study. The colors used in this Figure match those of table 1. This figure was based on the recent study conducted on algal genomes.[78]

Analysis of SnRK motifs in algae

Using the MEME motif tool, we identified 20 individual motifs associated with SnRK1s, 19 associated with SnRK2s, and 20 associated with SnRK3s. Similar to plant SnRKs, alignments revealed a catalytic and regulatory domain within the algal SnRKs identified. Of the SnRK1s and SnRK3s identified in our algal species of interest we identified the conserved ser/thr kinase the defining charatersitic of SnRKs. Further, SnRK2s were broken down dependent upon group where the ser/thr residue were identified. Discerning a relationship between the SnRK2s is difficult due to low boot strap values.

Figure 4: Phylogenetic tree generated from the SnRKs identified in the thirty-four species of algae used in this study. The evolutionary history was inferred by using the maximum likelihood method and JTT matrix-based model. This analysis involved 232 amino acid sequences. There was a total of 3060 positions in the final dataset. Evolutionary analyses were conducted in MEGA X.

Discussion

Through the mining of thirty-four algal genomes using homology searchers, SnRKs were identified in all species that were queried in this analysis. SnRKs were identified using PANTHER domain analysis databases. Through this analysis we observed that almost all the algal species observed in this study had identifiable kinase domains found in SnRKs. Most of the algal species used in this analysis had domains corresponding to SnRK1. Through our analysis we were also able to identify SnRK2s and SnRK3s, but their identification was restricted to the green algal lineages and a few outliers.

When analyzing the SnRK families and SnRK2 specific sub-families, we observed a variation between species leading to low bootstrap values in our phylogenic analysis. To improve our boot strap values, we separated the C-terminus and N-terminus and conducted phylogenic analysis on each of those. We found that doing this analysis with just the N-terminus improved bootstrap values and resolution of the phylogenetic tree analysis. Looking at the sequence alignments, we also saw the highest alignment values when looking at the N-terminal region. This is to be expected as the N-terminus contains a highly conserved kinase domain.[6] This highly conserved kinase domain was found in all three SnRK families. The C-terminus was highly variable between sub-families as expected but surprisingly was also highly variable between species of the different algal lineages and even within algal species of the same lineages. When strictly analyzing the C-terminus, we saw low bootstrap values reflecting this variability. We thus believe this variability between algal species and variability of these protein sequences is leading to these low values. This suggests that methodologies beyond homology searchers are necessary for the characterization of protein families with atypical evolution such as the ones observed here.[1] These analyses include protein domain analysis and

motif analysis, methods used in this study. One of the main problems is due to the overall broad diversity of algal evolution due to endosymbiotic events that muddled the evolutionary history of algae leading to weak phylogenetic associations.[64,65]

SnRK1 findings:

Potential SnRK1s were only identified by the presence of a kinase, UBA, and KA1/alphaCTD domains which are indicative of these kinases. The UBA domain plays a role in the activation and maintenance of catalytic activity, while the KA1 domain plays an important role in the binding of the beta and gamma subunits and allows for the interaction of SnRK1 with phospahtes.[20,66] Some of the algal species used in this analysis didn't have these characteristic domains as they were missing either the UBA or KA1 domain. While the KA1 domain is not present in all lineages the UBA is a common feature of the SnRK family that is conserved across eukaryotic evolution.[1,67,68] A few of the algal species had more than one SnRK1, which is something not observed as of this time in organisms other than plants and humans.[1] Based on the gene balance hypothesis algal SnRKs should engage in most of the functions that land plant SnRKs are involved in because proteins belonging to signaling networks are more likely to be retained.[1,69,70] In plants, SnRK1s are involved in an important signaling network based on sugar signaling. Their function is to act as master regulators of carbon/energy ratios.[1] As mentioned, not all the algal species used in this analysis have a proper SnRK1 which suggest that pathways might have a different functionality compared to those found in plants.

SnRK2 findings:

During our analysis we observed that only green algae have SnRK2s. Our analysis revealed that there are seemingly 2 classes of algal SnRK2s identified – SNRK2s that are more closely

related to the plant SnRKs, and SNRK2s that were not strongly related to plant SnRKs in the phylogenetic analysis. When we analyzed the SnRK2 motifs and sequence alignments, we saw a distinct difference between these two groups. This could suggest a unique function for these novel, algal-specific SnRK2s. SnRK2s identified in this study were difficult to classify due to divergence from their land plant counterparts. All the green algal SnRK2s had the conserved Ser/Thr kinase domain and structure similar to domain I which is important in osmotic and salt stress in plants. They however did lack domain II as the SnRKs were mostly shorter than the land plant SNRKs.

Previous research has found that SnRK2s were identified in osmotic, drought, and cold stresses, but their direct connection to photosynthesis has yet to be discovered in algae. Several SnRK overexpression studies done in plants, however, have demonstrated a connection between stress and photosynthesis. Research has indicated that individual SnRK2s genes acquire various regulatory properties while responding to stresses.[71] In addition, potentially normal conditions as well could be even further misunderstood in algae because of a lack of research into the exact function of SnRKs. Further, SnRK2s play an important role in desiccation responses in plants however their role in desiccation response in algae has yet to be elucidated. One study did find when a Klebsormidium species was exposed to desiccation stress there was increase in SnRK2 expression.[72] Other species of algae have also been shown to possess a vegetative desiccationtolerance phenotype.[73]

SnRK3 findings:

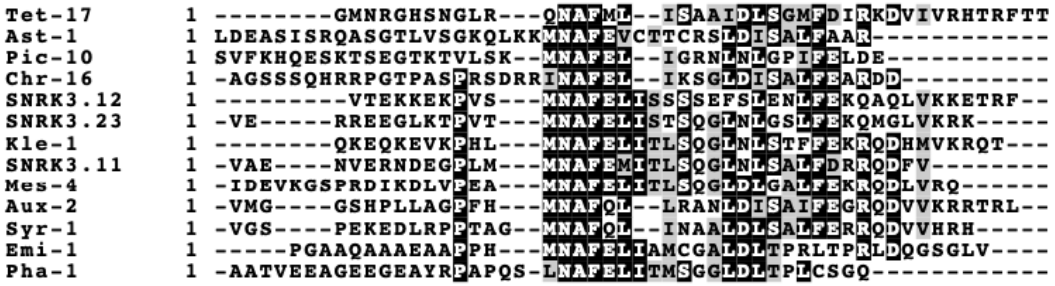


Figure 5: Sequence alignment of the identified SnRK3 family members from the various green algal species. As seen each candidate contains the "NAF" domain characteristics of SnRK3 family members.

We identified putative SnRK3s (CIPKs) in our domain analysis, but many of the algal species lacked SnRK3s with a defined NAF domain. As seen in plants the N-terminal kinase domain was conserved, while the C-terminal region is highly variant aside from the presence of the defined NAF domain.[74] We however, did identify SnRK3s in the green algae containing the NAF domain. It has been shown that some algal species have kinases similar to SnRK3s, but the lack of the NAF domain categorizes them differently than SnRK3s (Figure 5).[21] This approach however may be biased since our classification is based on using plant SnRKs as a reference, and not algae SnRKs, even though algae have a more ancient phylogenetic history and are the ancestors of terrestrial plants. Interestingly, all the SnRK3s identified in this analysis share sequence similarity to CIPK 24 an important Na transporter. CBL4-CIPK24 phosphorylates the sodium and hydrogen exchanger called SOS1 which sends excess Na+ from the cell back into the soil to improve salt tolerance.[74] Work has shown that CBL-CIPK signaling system is important in fine-tuning plant adaptive responses especially in response to adverse environmental conditions.[74] Based on the placement of these algae in the phylogenetic tree and the environments in which they live it may possible that this gene is a requirement for land adaption. Terrestrial algae include: Chlorophytes, Mesostigmatophyceae, Chlorokybophyceae, Klebsormidiophyceae, Zygnematophyceae, and the Coleochaetophyceae.[72,75] Our analaysis revlead that CIPKs are found in Chlorophytes,

Klebsormidiophyceae, and Zygnematophyceae. We did not have the ability to analyze genomes from the Chlorokybophtceae and Coleochaetophyceae due to the lack of available genome assemblies.

It has been suggested that SnRK3 is not only found in land plants, but it has not been explicitly found. Our research indicates that it has been found in iancestors to land plants, but has been lost in some of the core Chlorophytes.[1] It is generally accepted that the ancestors of terrestrial plants were closely related to charophytes.[61,76] The colonization of land requires a range of adaptive mechanisms to deal with new abiotic stresses including drought, osmotic, high-intensity light, and UV. [61,75] The reduced SnRK3 subfamily size in microalgae may explain this and may be compensated by the much larger and diverse group of SnRK2s. More research is needed to prove this assertion, which should include the utilization of more terrestrial algae and the analysis of their genomes when they become available.[77] A recent analysis has shown that of the 8605 known Chlorophytes only 124 have genome assemblies and of the 983 known Streptophyte algae only 7 have genome assmeblies.[78] That represents only 1.44% and 0.7% respectively. Even this small percentage is biased towards only a few genera.[78] This problem is exasperated by a decreasing quality of algal genome assemblies.[78] Currently there is a lack of terrestrial algae genomes to fully analyze their SnRK profile and their potential role on the evolution from aquatic environments to terrestrial ones.

Citations

1. Colina, F. *et al.* Genome-wide identification and characterization of CKIN/SnRK gene family in Chlamydomonas reinhardtii. *Scientific Reports* **9**, 1–16 (2019).
2. Subudhi, S. & Internationals, O. Bioprospecting for Algal Based Nutraceuticals and High Value Added Compounds. *Journal of Pharmacy and Pharmaceutics* **4**, 145–150 (2017).
3. Vannini, C. *et al.* Overexpression of the rice Osmyb4 gene increases chilling and freezing tolerance of Arabidopsis thaliana plants. *The Plant Journal* **37**, 115–127 (2004).
4. Cai, C. *et al.* Overexpression of a Novel Arabidopsis Gene SUPA Leads to Various Morphological and Abiotic Stress Tolerance Alternations in Arabidopsis and Poplar. *Frontiers in Plant Science* **11**, 1766 (2020).
5. Alexander, A., Singh, V. K. & Mishra, A. Overexpression of differentially expressed AhCytb6 gene during plant-microbe interaction improves tolerance to N2 deficit and salt stress in transgenic tobacco. *Sci Rep* **11**, 13435 (2021).
6. Kulik, A., Wawer, I., Krzywińska, E., Bucholc, M. & Dobrowolska, G. SnRK2 Protein Kinases—Key Regulators of Plant Response to Abiotic Stresses. *OMICS: A Journal of Integrative Biology* **15**, 859–872 (2011).
7. Tian, S. *et al.* Cloning and characterization of TaSnRK2.3, a novel SnRK2 gene in common wheat. *J Exp Bot* **64**, 2063–2080 (2013).
8. Wang, Y.-L., Wang, J., Chen, X., Wang, Z.-X. & Wu, J.-W. Crystal structure of the kinase and UBA domains of SNRK reveals a distinct UBA binding mode in the AMPK family. *Biochemical and Biophysical Research Communications* **495**, 1–6 (2018).
9. Hey, S. J., Byrne, E. & Halford, N. G. The interface between metabolic and stress signalling. *Ann Bot* **105**, 197–203 (2010).

10. Wu, P., Wang, W., Duan, W., Li, Y. & Hou, X. Comprehensive Analysis of the CDPK-SnRK Superfamily Genes in Chinese Cabbage and Its Evolutionary Implications in Plants. *Front. Plant Sci.* **8**, (2017).
11. Crepin, N. & Rolland, F. SnRK1 activation, signaling, and networking for energy homeostasis. *Current Opinion in Plant Biology* **51**, 29–36 (2019).
12. Robaglia, C., Thomas, M. & Meyer, C. Sensing nutrient and energy status by SnRK1 and TOR kinases. *Current Opinion in Plant Biology* **15**, 301–307 (2012).
13. Cho, Y.-H., Hong, J.-W., Kim, E.-C. & Yoo, S.-D. Regulatory Functions of SnRK1 in Stress-Responsive Gene Expression and in Plant Growth and Development. *Plant Physiology* **158**, 1955–1964 (2012).
14. Li, C. *et al.* Molecular Characterization and Co-expression Analysis of the SnRK2 Gene Family in Sugarcane (Saccharum officinarum L.). *Scientific Reports* **7**, 1–12 (2017).
15. Beckmann, L., Edel, K. H., Batistič, O. & Kudla, J. A calcium sensor – protein kinase signaling module diversified in plants and is retained in all lineages of Bikonta species. *Scientific Reports* **6**, 31645 (2016).
16. Shi, S. *et al.* The Arabidopsis Calcium-Dependent Protein Kinases (CDPKs) and Their Roles in Plant Growth Regulation and Abiotic Stress Responses. *Int J Mol Sci* **19**, (2018).
17. Albrecht, V., Ritz, O., Linder, S., Harter, K. & Kudla, J. The NAF domain defines a novel protein–protein interaction module conserved in Ca2+-regulated kinases. *EMBO J* **20**, 1051–1063 (2001).
18. Umezawa, T. *et al.* Molecular Basis of the Core Regulatory Network in ABA Responses: Sensing, Signaling and Transport. *Plant Cell Physiol* **51**, 1821–1839 (2010).

19. Hauser, F., Waadt, R. & Schroeder, J. I. Evolution of Abscisic Acid Synthesis and Signaling Mechanisms. *Curr Biol* **21**, R346–R355 (2011).
20. Jamsheer K, M., Jindal, S. & Laxmi, A. Evolution of TOR–SnRK dynamics in green plants and its integration with phytohormone signaling networks. *J Exp Bot* **70**, 2239–2259 (2019).
21. Brawley, S. H. *et al.* Insights into the red algae and eukaryotic evolution from the genome of Porphyra umbilicalis (Bangiophyceae, Rhodophyta). *PNAS* **114**, E6361–E6370 (2017).
22. Shinozawa, A. *et al.* SnRK2 protein kinases represent an ancient system in plants for adaptation to a terrestrial environment. *Communications Biology* **2**, 1–13 (2019).
23. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research* **47**, W636–W641 (2019).
24. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
25. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* **35**, 1547–1549 (2018).
26. Bailey, T. L. *et al.* MEME Suite: tools for motif discovery and searching. *Nucleic Acids Research* **37**, W202–W208 (2009).
27. Ronquist, F. *et al.* MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Systematic Biology* **61**, 539–542 (2012).
28. Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004).

29. Liu, H. *et al.* Symbiodinium genomes reveal adaptive evolution of functions related to coral-dinoflagellate symbiosis. *Communications Biology* **1**, 1–11 (2018).
30. Liew, Y. J., Aranda, M. & Voolstra, C. R. Reefgenomics.Org - a repository for marine genomics data. *Database* **2016**, (2016).
31. Lovejoy, C. *Small Planktonic single celled eukaryotes from the Arctic Ocean*. https://www.osti.gov/dataexplorer/biblio/dataset/1488049 (2010) doi:10.25585/1488049.
32. Nordberg, H. *et al.* The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Research* **42**, D26–D31 (2014).
33. Dobáková, E., Flegontov, P., Skalický, T. & Lukeš, J. Unexpectedly Streamlined Mitochondrial Genome of the Euglenozoan Euglena gracilis. *Genome Biology and Evolution* **7**, 3358–3367 (2015).
34. Price, D. C. *et al.* Analysis of an improved Cyanophora paradoxa genome assembly. *DNA Research* **26**, 287–299 (2019).
35. Allen, A. *Complete Genome Sequencing of the Prymnesiophyte Haptophyte Phaeocystis globosa*. https://www.osti.gov/dataexplorer/biblio/dataset/1488234 (2012) doi:10.25585/1488234.
36. Read, B. A. *et al.* Pan genome of the phytoplankton Emiliania underpins its global distribution. *Nature* **499**, 209–213 (2013).
37. Curtis, B. A. *et al.* Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. *Nature* **492**, 59–65 (2012).
38. Nakamura, Y. *et al.* The First Symbiont-Free Genome Sequence of Marine Red Alga, Susabi-nori (Pyropia yezoensis). *PLOS ONE* **8**, e57122 (2013).

39. Rossoni, A. W. *et al.* The genomes of polyextremophilic cyanidiales contain 1% horizontally transferred genes with diverse adaptive functions. *eLife* **8**, e45017 (2019).

40. Collén, J. *et al.* Genome structure and metabolic features in the red seaweed Chondrus crispus shed light on evolution of the Archaeplastida. *PNAS* **110**, 5247–5252 (2013).

41. Vieler, A. *et al.* Genome, Functional Gene Annotation, and Nuclear Transformation of the Heterokont Oleaginous Alga Nannochloropsis oceanica CCMP1779. *PLOS Genetics* **8**, e1003064 (2012).

42. Ohan, J. A. *et al.* Nuclear Genome Assembly of the Microalga Nannochloropsis salina CCMP1776. *Microbiol Resour Announc* **8**, (2019).

43. Nishitsuji, K. *et al.* Draft genome of the brown alga, Nemacystus decipiens , Onna-1 strain: Fusion of genes involved in the sulfated fucan biosynthesis pathway. *Scientific Reports* **9**, 4607 (2019).

44. Cock, J. M. *et al.* The Ectocarpus genome and the independent evolution of multicellularity in brown algae. *Nature* **465**, 617–621 (2010).

45. Tyler, C. R. S. *et al.* High-Quality Draft Genome Sequence of the Green Alga Tetraselmis striata (Chlorophyta) Generated from PacBio Sequencing. *Microbiol Resour Announc* **8**, (2019).

46. Merchant, S. S. *et al.* The Chlamydomonas Genome Reveals the Evolution of Key Animal and Plant Functions. *Science* **318**, 245–250 (2007).

47. Polle, J. E. W. *et al.* Draft Nuclear Genome Sequence of the Halophilic and Beta-Carotene-Accumulating Green Alga Dunaliella salina Strain CCAP19/18. *Genome Announc.* **5**, (2017).

48. ASM433591v1 - Genome - Assembly - NCBI. https://www-ncbi-nlm-nih-gov.libproxy.unm.edu/assembly/GCA_004335915.1/.

49. Lopes dos Santos, A. *et al.* Chloropicophyceae, a new class of picophytoplanktonic prasinophytes. *Scientific Reports* **7**, 14019 (2017).
50. Blanc-Mathieu, R. *et al.* An improved genome of the model marine alga Ostreococcus tauri unfolds by assessing Illumina de novo assemblies. *BMC Genomics* **15**, 1–12 (2014).
51. Palenik, B. *et al.* The tiny eukaryote Ostreococcus provides genomic insights into the paradox of plankton speciation. *PNAS* **104**, 7705–7710 (2007).
52. Vaulot, D. *et al.* Metagenomes of the Picoalga Bathycoccus from the Chile Coastal Upwelling. *PLoS One* **7**, (2012).
53. Moreau, H. *et al.* Gene functionalities and genome structure in Bathycoccus prasinos reflect cellular specializations at the base of the green lineage. *Genome Biol* **13**, R74 (2012).
54. Arimoto, A. *et al.* A siphonous macroalgal genome suggests convergent functions of homeobox genes in algae and land plants. *DNA Research* **26**, 183–192 (2019).
55. Hovde, B. T. *et al.* Genomic characterization reveals significant divergence within Chlorella sorokiniana (Chlorellales, Trebouxiophyceae). *Algal Research* **35**, 449–461 (2018).
56. Gonzalez-Esquer, C. R., Twary, S. N., Hovde, B. T. & Starkenburg, S. R. Nuclear, Chloroplast, and Mitochondrial Genome Sequences of the Prospective Microalgal Biofuel Strain Picochlorum soloecismus. *Genome Announc.* **6**, (2018).
57. Armaleo, D. *et al.* The lichen symbiosis re-viewed through the genomes of Cladonia grayi and its algal partner Asterochloris glomerata. *BMC Genomics* **20**, 1–33 (2019).
58. Škaloud, P., Friedl, T., Hallmann, C., Beck, A. & Grande, F. D. Taxonomic revision and species delimitation of coccoid green algae currently assigned to the genus Dictyochloropsis (Trebouxiophyceae, Chlorophyta). *Journal of Phycology* **52**, 599–617 (2016).

59. de Koning, A. P. & Keeling, P. J. The complete plastid genome sequence of the parasitic green alga Helicosporidium sp. is highly reduced and structured. *BMC Biology* **4**, 12 (2006).
60. Gao, C. *et al.* Oil accumulation mechanisms of the oleaginous microalga Chlorella protothecoides revealed through its genome, transcriptomes, and proteomes. *BMC Genomics* **15**, (2014).
61. Hori, K. *et al.* Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. *Nature Communications* **5**, 3978 (2014).
62. Nishiyama, T. *et al.* The Chara Genome: Secondary Complexity and Implications for Plant Terrestrialization. *Cell* **174**, 448-464.e24 (2018).
63. Cheng, S. *et al.* Genomes of Subaerial Zygnematophyceae Provide Insights into Land Plant Evolution. *Cell* **179**, 1057-1067.e14 (2019).
64. Stiller, J. W. *et al.* The evolution of photosynthesis in chromist algae through serial endosymbioses. *Nature Communications* **5**, 5764 (2014).
65. Morozov, A. A. & Galachyants, Y. P. Diatom genes originating from red and green algae: Implications for the secondary endosymbiosis models. *Marine Genomics* **45**, 72–78 (2019).
66. Emanuelle, S., Doblin, M. S., Stapleton, D. I., Bacic, A. & Gooley, P. R. Molecular Insights into the Enigmatic Metabolic Regulator, SnRK1. *Trends in Plant Science* **21**, 341–353 (2016).
67. Hardie, D. G. AMP-activated/SNF1 protein kinases: conserved guardians of cellular energy. *Nature Reviews Molecular Cell Biology* **8**, 774–785 (2007).
68. Broeckx, T., Hulsmans, S. & Rolland, F. The plant energy sensor: evolutionary conservation and divergence of SnRK1 structure, regulation, and function. *J Exp Bot* **67**, 6215–6252 (2016).

69. Birchler, J. A. & Veitia, R. A. Gene balance hypothesis: Connecting issues of dosage sensitivity across biological disciplines. *PNAS* **109**, 14746–14753 (2012).

70. Cheng, F. *et al.* Gene retention, fractionation and subgenome differences in polyploid plants. *Nature Plants* **4**, 258–268 (2018).

71. Liu, M. *et al.* Overexpression of NtSnRK2.2 enhances salt tolerance in Nicotiana tabacum by regulating carbohydrate metabolism and lateral root development. *Functional Plant Biol.* **47**, 537–543 (2020).

72. Holzinger, A. & Pichrtová, M. Abiotic Stress Tolerance of Charophyte Green Algae: New Challenges for Omics Techniques. *Front. Plant Sci.* **7**, (2016).

73. Terlova, E. F., Holzinger, A. & Lewis, L. A. Terrestrial Green Algae Show Higher Tolerance to Dehydration than Do Their Aquatic Sister-Species. *Microb Ecol* 1–13 (2021) doi:10.1007/s00248-020-01679-3.

74. Mao, J. *et al.* Mechanisms and Physiological Roles of the CBL-CIPK Networking System in Arabidopsis thaliana. *Genes* **7**, 62 (2016).

75. Vries, J. de & Archibald, J. M. Plant evolution: landmarks on the path to terrestrial life. *New Phytologist* **217**, 1428–1434 (2018).

76. Wang, S. *et al.* Genomes of early-diverging streptophyte algae shed light on plant terrestrialization. *Nature Plants* **6**, 95–106 (2020).

77. New Taxa of Streptophyte Algae (Streptophyta) from Terrestrial Habitats Revealed Using an Integrative Approach. *Protist* **169**, 406–431 (2018).

78. Hanschen, E. R. & Starkenburg, S. R. The state of algal genome quality and diversity. *Algal Research* **50**, 101968 (2020).

Part 4

Enhancing Algal Vision: The heterologous expression of the photoreceptor neochrome in the micro-alga *Chlamydomans reinhardtii*

Taylor Britton[1], Jackie Mettler[2], Sangeeta Negi[3], David T. Hanson[1], and Blake Hovde[2 *]

1. Department of Biology, The University of New Mexico, Albuquerque, NM, 87104, USA

2. Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

3. New Mexico Consortium, Los Alamos, NM, 87544, USA

Introduction

One of the major hurdles that is currently impeding the viability of an algae-based biofuel is cost. The majority of the costs of producing biofuels from algae is associated with cultivation and harvesting of the algae stock.[1,2] Therefore, research efforts focused on increasing algal yields will directly reduce the price of algal derived fuels. Low yields can arise due to suboptimal cultivation conditions such as: saturating light conditions, overcast days, low light conditions experienced during the early morning and evening, and during "self-shading" when cell densities in the culture vessel or pond become so high algal cells at the surface impede light from reaching other algae during cultivation. Photoreceptor engineering is one potential target to improve biomass yields, as such receptors are key regulators for many regulatory pathways. Photoreceptors are used to detect changes in light intensity and wavelength, which generate signal transduction cascades that result in physiological responses that play a critical role in cell development, biomass accumulation, and adaptation to changing environmental conditions.[3,4]

Many microalgae contain sophisticated light-sensing systems that are used for the monitoring of light.[3,5] These systems consists of regulators such as transcription factors, kinases, phosphatases, and protein degradation pathways.[6] These regulator pathways have key roles in mediating light signaling through the coordinated activation and repression of specific downstream genes. Despite the importance of photoreceptors, little work has been conducted in fully characterizing photoreceptors and their corresponding regulatory pathways in algae, particularly regarding the role that they play in carbon fixation and photosynthetic efficiency. One of these photoreceptors, and the target photoreceptor of the work presented here, is neochrome.

Neochrome is a type of photoreceptor that is prevalent in ferns. The main function of neochrome is the mediation of blue and re/far-red light dependent responses through a chimeric light-sensing

system. This system consists of a blue light phototropin regulatory domain located at the C-terminus and a red/far-red light phytochrome receptor located at the N-terminus.[7,8] Phototropins are highly conserved photoreceptors that are used by photosynthetic organisms to sense blue light.[9] When phototropins are activated by blue light it enables plants to maximize the capture of photosynthetically active radiation by the photosynthetic apparatus, promote leaf expansion, stomatal opening, chloroplast accumulation/movement, and leaf solar tracking to protect against damage under excessive light conditions (chloroplast avoidance).[9] The function of phototropins in algae, however, is less understood and has only been studied in the model green alga *Chlamydomonas reinhardtii*. Studies on *C. reinhardtii* phototropin have indicated that it is involved in the regulation of the reproductive life cycle and triggers major changes in expression of genes encoding enzymes involved in chlorophyll and carotenoid biosynthesis.[10,11] Moreover, little is known about phototropin's mode of action or cell-wide impacts on growth; however, recent research has shown that repressing phototropin appears to enhance photosynthetic capacity under high light conditions, resulting in improved biomass yields over wild type cells.

Phytochromes are unique among photoreceptors because they have the ability to interconvert between red absorbing conformation (absorption maximum of ~600nm) and a far-red absorbing conformation (absorption maximum of 730nm). This conversion is an important feature in plants because it allows for the conversion of light into biochemical signals as a measure of the ratio of red to far-red light allowing for plants to assess the quantity of PAR and trigger shade avoidance. Phytochromes also play a role in germination, chlorophyll synthesis, the timing of flowering, phototaxis, carbon allocation and biomass production in developing plants.[12,13] Research has shown that photochromes play an important part in growth and biomass yield [12,14] In algae the role that phytochromes play in key regulatory functions is in its infancy.[13]

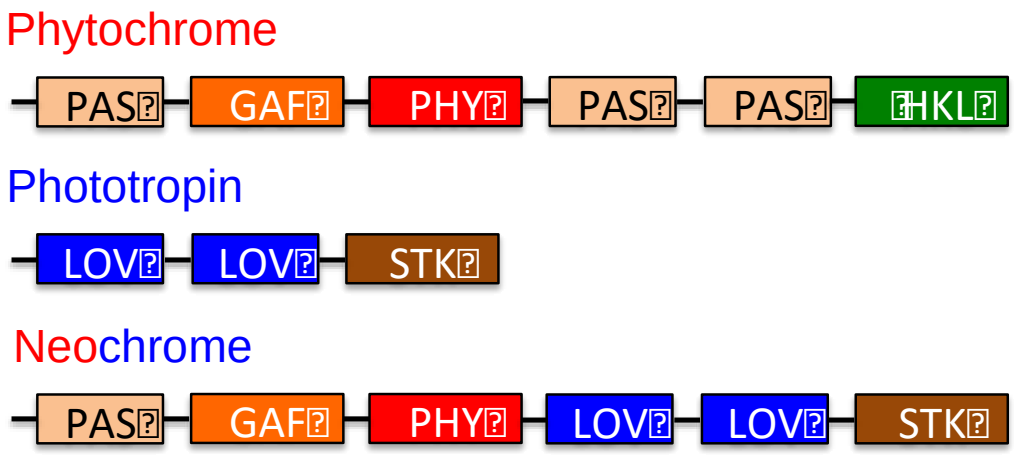


Figure 1: Protein domains of the photoreceptors phytochrome, phototropin, and neochrome. Neochrome is a chimera of as evident by it containing the photosensory domain of phytochrome and the regulatory domain of phototropin.

Neochromes chimeric nature enables organisms to dynamically sense and respond to the proportions of red and far-red light energy that mediate blue/far red-regulated phototrophic responses, which ultimately improves fitness under low-light canopy conditions (Figure 1).[15] Research in *Arabidopsis* has shown that neochrome can function as a phototropin in low-light conditions even though its main function is to perform as a phytochrome that mediates both the red/far-red and blue/far-red light responses.[15] As a defining feature, neochromes lack a conserved cysteine residue in the protropin region of the receptor that may impede some of its ability to act as a true phototropin.[16] Research into the algal neochrome is even more unclear, but it has been shown that some Streptophyte algae neochromes function within the cytosol to mediate changes in plastid orientation in response to changes light.[13] We hypothesize that incorporating this hybrid photoreceptor into algae will improve photosynthetic capacity and biomass production in low light conditions (early and late photoperiods, self-shading/dense cultures, and dense cloud cover) resulting in improved productivity compared to wild-type algae.

In this work we aimed to elucidate the role that neochrome plays in carbon metabolism and photosynthesis in microalgae. We expressed the neochrome gene from the fern *Adiantum capillus-veneris* and the green algae *Mougeotia scalaris* into lines of the model algae *C. reinhardtii* both

with and without a phototropin deletion. We demonstrate that expression of this transgene both improved growth and led to the secretion of carbohydrates. However, challenges with gene stability led to a loss of phenotype after longer term (months) of cultivation.

Methods

Strains and culturing conditions

Chlamydomonas reinhardtii UV4 and A4 were obtained from Peter Hegemann of the University of Berlin. Transgenic lines were maintained on HS and TAP media plates containing 6 μg/ml of the antibiotic zeocin. Cultures were kept at 25°C and 50 µmoles photons $m^{-2}s^{-1}$.

Construction of the expression vector of Neochrome gene for Chlamydomonas

The neochrome encoding genes were synthesized using the gene sequences from the fern *Adiantum capillus-veneris* and the algae *Mougeotia scalaris* as templates. The sequences were synthesized by the company Genewiz (https://www.genewiz.com/) and inserted into the *C. reinhardtii* vector pSL18. Our synthesized pSL18-neochrome vector contains an endogenous PsaD promoter/terminator to drive neochrome gene expression and a codon optimized sh-ble gene was placed under control of an actin promoter/terminator pair for colony selection.

Transformation

C. reinhardtii was transformed via a glass bead method.[17] Cell suspensions were removed from TAP liquid culture after three days of growth when cell counts were typically between 10^6 and 10^8 cells/mL. Cells were spun down and resuspended in TAP media containing 1M sucrose. This suspension was transferred to an Eppendorf tube containing 300mg of acid washed glass beads (Sigma, USA), PEG solution, and 2μg of linearized pSL18-neochrome DNA. This solution was vortexed for 15 seconds. 300uL of solution was transferred to a 15mL tube containing 5mL of

TAP media. Cells were allowed to recover for ~4 hours before being plated on TAP media containing zeocin at a concentration of 6 μg/mL and 50μg/mL of arginine. Transformed plates were incubated for 10 days under continuous light (50 µmoles photons $m^{-2}s^{-1}$). Colonies were re-plated on TAP agar plates containing zeocin at a concentration 6 µg/mL for further analysis.

Growth and morphology in batch mode. Cultures were grown in 250-ml Erlenmeyer flasks to the early to mid-log phase in trisacetate-phosphate (TAP) liquid media or in high salt (HS) minimal liquid media dependent on the experiment. Cells were grown in shaker flasks under constant light at an intensity of ~35 μmol m^{-2} s^{-1} photosynthetically active radiation (PAR) under constant illumination unless specified otherwise. Samples for experimentation were taken after the indicated times for each corresponding experiment. Cell counts and cell sizes were assessed using an Attune NxT acoustic focusing flow cytometer, Hausser Scientific Bright-Line hemocytometer, and Olympus BX51 microscope.

Gene presence and expression analysis

For screening of putative transformants, single colonies were grown on TAP media plates containing 6 µg/mL of zeocin. DNA was isolated from single colonies using 6% Chelex in Tris. The Chelex/Tris solution was boiled for fifteen minutes toisolate DNA to be used to screen for the presence of zeocin and neochrome genes via PCR.[47,48] PCR reactions were analyzed using gel electrophoreses. Positive PCR products were further confirmed by Sanger-sequencing (GeneWiz). Positive transformants were further screened for gene expression using RT-PCR. Cells were harvested by centrifugation at 2500RPM for five minutes, flash frozen in liquid nitrogen, and stored at -80°C prior beginning the RNA isolation processes. RNA was isolated using the TRIzol phenol chloroform method.[49] cDNA was prepared to a concentration of 50 ng/µL using a

Clonetech RNA to cDNA EcoDry Premix. RT-PCR was performed on a T100 Bio-Rad thermal cycler and products were analyzed using gel electrophoreses.

Chlorophyll content

To measure chlorophyll content cells were grown in HS liquid media to early log phase under constant light of 35 μmol m^{-2} s^{-1}. Chlorophyll was extracted using methanol. 1 mL of culture was centrifuged at 15,000 X g for five minutes at room temperature. Pelleted cells were suspended in 100% methanol and vortexed. Samples were placed in the dark for 4 hours. Cellular debris was pelleted using centrifugation at 15,000 X g for five minutes. Absorption was measured at 750 nm, 665.2 nm, and 652 nm using a Varian Cary 100 Bio UV/ Visible spectrometer. Calculations of total chlorophyll, chlorophyll A, and chlorophyll B were performed as described previously.[18]

Photosynthetic parameters measurement using PAM Fluorometry

A Dual PAM 100 Fluorometry (Walz, Germany) was used to measure photosynthetic parameters of photosystem (II) and photosystem (I). The in-built software was used to measure non-photochemical (NPQ) quenching, photosystem (II) efficiency (Fv/Fm), and electron transport. Samples for both the transgenic lines and WT were taken from cultures grown in shake flask to mid log phase. Using the internal in-built analytical software, a dark induction curve and light induction curve were experiments were conducted. The values reported here are reported as an average of three replicates of the sample analyzed.

Flowcytometry and BODIPY staining

Cell size and lipid content were analyzed using an Attune NXT flow cytometer. 250 μl of cells at a density of 10^6 cells/mL were placed into a well where 15 μl of BODIPY stain. Cells were

incubated at room temperature for 30 minutes. The stained algae were then analyzed using the Attune NXT flow cytometer.

Results

PCR Confirmation and Expression Confirmation

We designed an expression vector for the heterologous expression of neochrome. A marker gene was used to determine if a successful genetic transformation had taken place. We used the zeocin antibiotic selection as our marker gene of choice. Zeocin was selected as our selective marker because it is effective at killing *C. reinhardtii*. Research has also shown that the addition of the zeocin resistance (sh-ble) gene can improve heterologous expression of transgenes and has been used to generate transgenic clones from various algal species.[19–21] An expression vector (Supplemental Figure 10) with zeocin selection as the antibiotic marker and the heterologous neochrome gene from *Adiantum capillus-veneris* or the gene from the green algae *Mougeotia scalaris* under the control of a PsaD promoter/terminator pair (Supplemental Figure 10) were used for the genetic transformations of *C. reinardtii*. PSAD encodes for an abundant chloroplast protein located on the stromal side of the photosystem I complex and has widely been used for efficient gene expression in microalgae.[22] Lastly, the neochrome vector was linearized and transformed via glass beads. We successfully generated > 100 colonies that successfully grew on the zeocin selective media after genetic transformation these putative transgenic lines were first screened through PCR. Primers were designed with amplicons for a region within the neochrome transgene. This amplicon yielded an expected band of ~1 Kb in transformed algae, while no band was present in the wild type; the correct band identity was further confirmed by Sanger sequencing. 10 transgenic lines were selected to confirm neochrome gene expression via reverse transcriptase

PCR (RT-PCR) (Figure 2). No neochrome transcripts were detected in wild-type cells, while the transgenic lines showed neochrome expression using the RT-PCR amplification method.

Growth Measurements and Morphology

Based on expression we decided to characterize separate transgenic cell lines in two different backgrounds *C. reinhardtii* wild-type (UV4) and *C. reinhardtii* with phototropin knocked-out (A4) and two different versions of the neochrome gene, one from the fern *Adiantum capillus-veneris* (C) and the other from the green algae *Mougeotia scalaris* (M): A4C4, A4M1, A4M2, A4M3, A4C17, UV4C14, and UV4M3. A batch-mode comparison of growth between wild type and four transgenic lines was conducted in low light intensities (35 µmol m^{-2} s^{-1}) with white light in photoautotrophic conditions (Figure 1A). Cultures exhibited similar doubling rates, albeit an

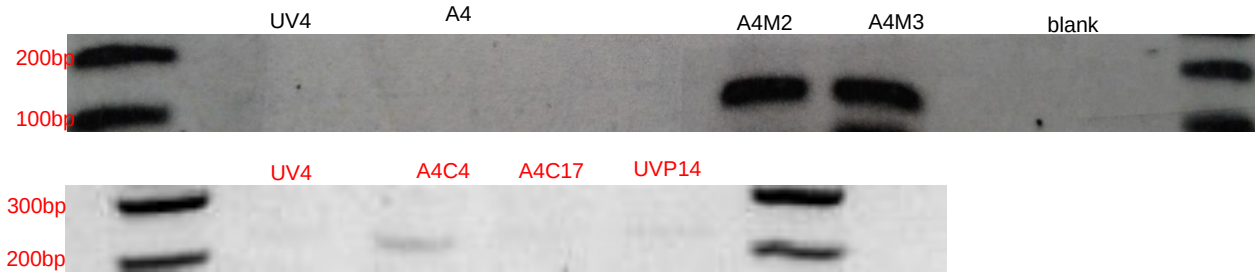


Figure 2: RT-PCR confirmation of neochrome expression in some of various lines. The band was only present in the transgenic lines and was absent in the wild-type (UV4) and the phototropin knockout line (A4).

increase in growth rates towards the end of the curve for one of the lines. In light, wild type-UV4 had 0.0869±0.04 doublings/day, while A4 had 0.0949±0.04 doublings/day A4M2 had 0.0869±0.05 doublings/day, A4C4 had 0.0816±0.05 doublings/day, A4C17 had 0.111±0.04 doublings/day, and A4M1 had 0.131 doublings/day. We determined the percent difference between the transgenic lines and the wild type on the bases of OD on the final day of the growth experiment (day 10). From this initial grow out A4M4 exhibited the fastest doubling time and highest culture density. We did a second grow out of A4C4, comparing this transgenic line with both wildtype cell lines

and saw similar results in improved growth rate (Figure 1C). UV4, A4, A4C4, A4M3, UV4C14, and UV4M3 were grown in photoautotrophic conditions at 35 µmol m^{-2} s^{-1} these however were grown with light containing wavelengths close to the excitation values of neochrome (a far-red light component) (Supplemental Figure 9). This was absent in our previous white light experiments (Figure 2A). Cultures were also grown in a greenhouse kept at 23 C. The cultures were subjected to natural light conditions

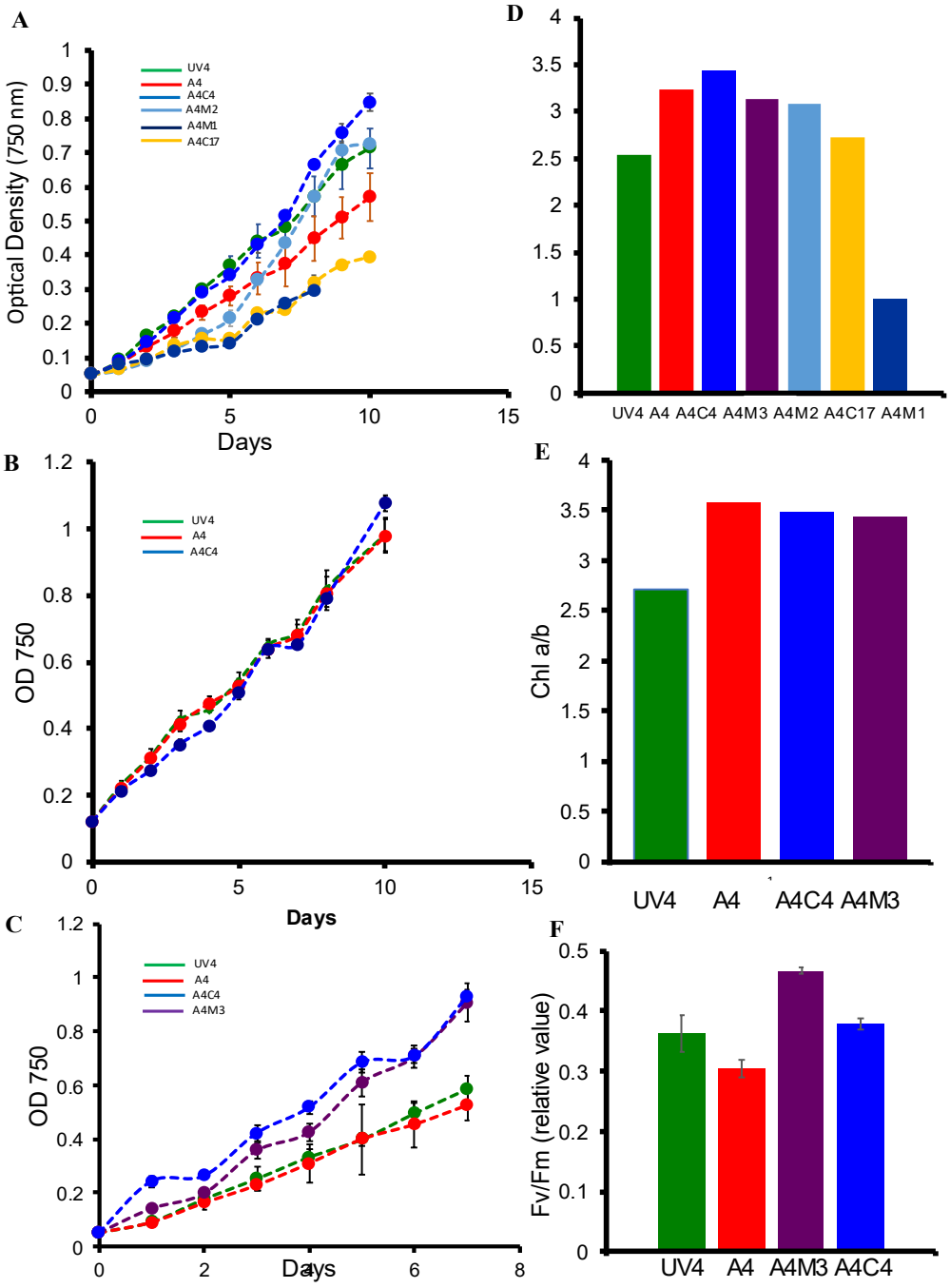


Figure 3: (A) Initial growth curve of various neochrome transgenic lines. Samples were grown in 35 µmol m^{-2} s^{-1} of white light. Candidates that showed promising results were grown in triplicate in light containing wavelengths to excite neochrome with an intensity of 35 µmol m^{-2} s^{-1} (B and C). B and C show significant difference compared to the wildtype and background line. (D) chlorophyll a/b ratio corresponding to the (A) growth curve. Transgenic lines had a high ratio compared to the wild type. (E) is the chlorophyll a/b ratio corresponding to the (C) growth curve. Transgenic lines have higher ratio compared to the wild type indicating a shrunken antenna complex. (F) Corresponds to the (C) growth curve. Photosystem 2 efficiency represented by Fv/FM ratio. The transgenic lines were significantly more efficient than the wild type and phototropin knockout line at these light intensities.

experienced during December 20-30 in Albuquerque, NM. We also grew the mutant lines and WTs in red light with a light intensity of 50 µmol $m^{-2}s^{-1}$. This growth curve showed a 15% difference between the mutant lines and the WTs. These samples along with ones grown at 35 µmol m^{-2} s^{-1} with a far-red light component were grown in nitrogen starved conditions.

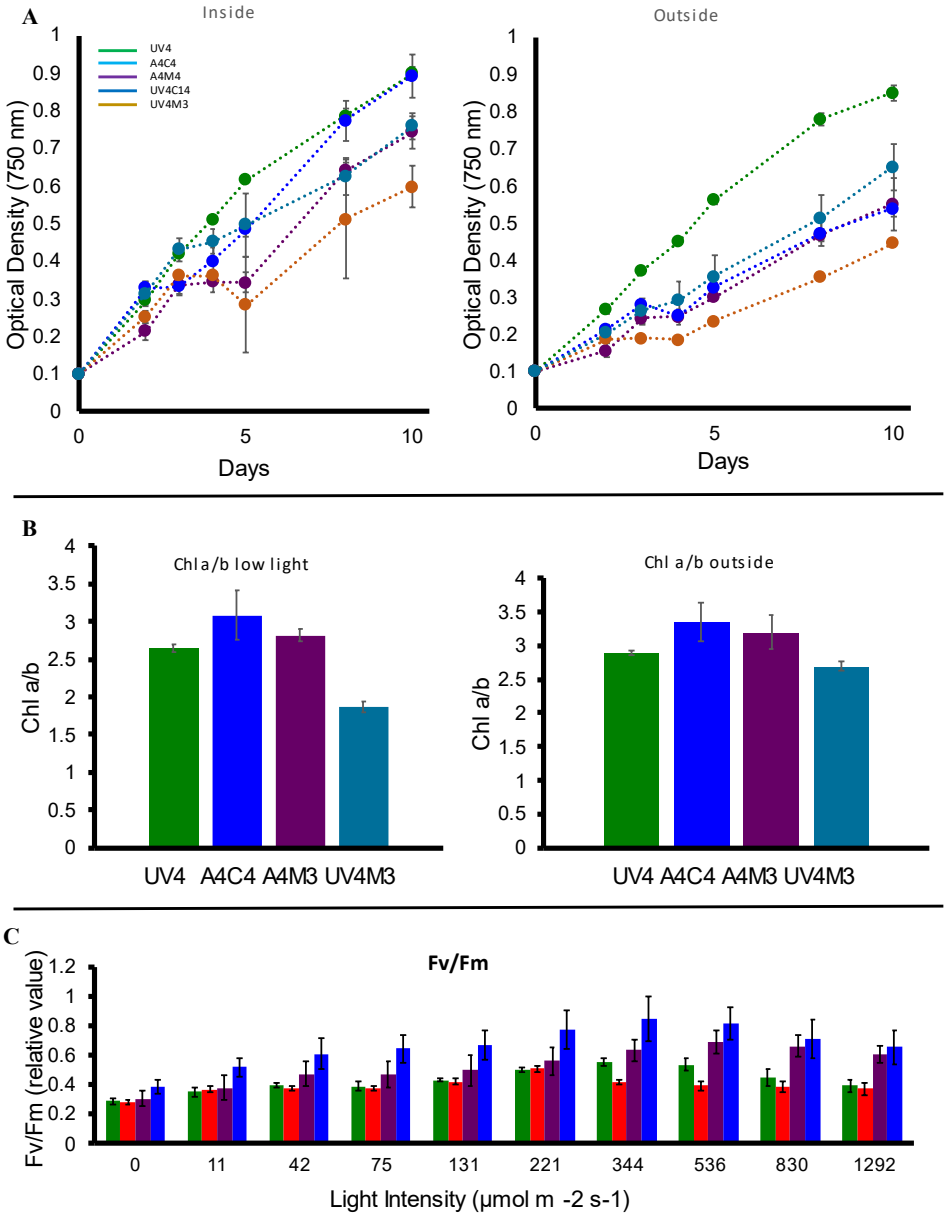


Figure 4: Growth curves, chlorophyll a/b ratios, and photosynthetic analysis of samples that generated an oily phenotype. (A) Growth curves of transgenic lines grown in 35 µmol m^{-2} s^{-1} of light with the neochrome excitation. The samples were also grown outside in a greenhouse kept at constant temperature. (B) Chlorophyll a/b ratios of these two experiments. Two of the three transgenic lines had ratios higher than the wild type. (C) Fv/Fm (photosystem II efficiency) was higher than the transgenic lines at higher light intensities.

Neochrome affects total chlorophyll and chlorophyll a/b ratios Due to our observed changes in growth in the transgenic lines we wanted to observe the effect of neochrome expression on

chlorophyll amount, content, and photosystem II (PSII) regulation. First, we quantified the chlorophyll *a/b* ratio with the transgenic lines showing a high ratio compared to the wild type-UV4 expect for U4M3 in all of our growth experiments (Figure 1B, 1D, 2C, and 2D). Higher a/b ratios correspond to a shrinking light harvesting complex. It has been shown before that the most effective strategy to increase photosynthetic light utilization efficiency is to reduce the size of the light-harvesting antenna, which reduces the optical cross section of the antenna complexes.[25–28] To determine the relationship between antenna size and PSII (Photosystem II) optical section we performed a quick chlorophyll fluorescence induction measurement.

BODIPY staining

Chlamydomonas cells that had confirmed expression for neochrome were further analyzed using Dual-PAM and BODIPY experiments. BODIPY treatment allows the staining and subsequent detection of neutral lipids in living cells. *C. reinhardtii* cells for the Dual-Pam and BODIPY analysis were grown in shake flask with HS media. Cells were grown for 7 days experiencing continuous light at an intensity of 25umol m-2 s-1. In Figure 1, we show BODIPY analysis done on mutants in the A4 background. Our mutants have a larger number of BODIPY events as compared to the WTs. Flash fluorescence induction was measured using a chlorophyll concentration of ~ 2.5 μg chlorophyll/mL. All the cells were dark adapted for 10 min prior to the experiment. The values of chlorophyll fluorescence were normalized to the maximum achieved values for a given sample. Chlorophyll fluorescence analysis determined that both transgenic lines

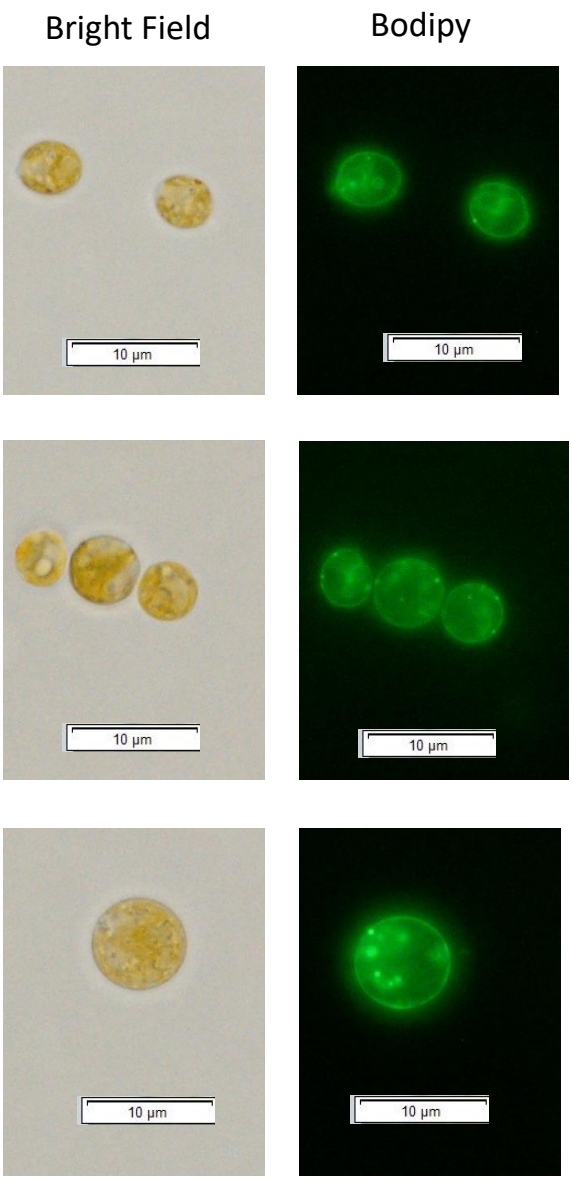


Figure 5: Bodipy (lipid) stain of the wild type and one of our transgenic lines. More lipid content is indicated by more fluorescent bodies and intensity.

that had experiments conducted on them had higher Fv/Fm ratios indicating higher PSII efficiency (Figure 1E and Figure 2E).

Discussion

Efforts are needed to improve algae's photosynthetic output so as to make them more economically competitive as a source of biofuels and food additives. Significant improvements are needed to increase biomass yield under non-optimal growth conditions (shady days, dense cultures, and during dawn and dusk). Genetic engineering has been applied towards the optimization of *Chlamydomonas* as a biotechnological platform. We targeted neochrome for genetic engineering due to its potential to expand the spectrum available to *Chlamydomonas* and improve photosynthetic output during times of sub-optimal growth. Here, we generated a neochrome expression line of *Chlamydomonas* via the glass bead method. C*hlamydomonas* transformation through electroporation has been shown to be easy and efficient, with the caveat that genome integration may be random and unstable.[17]

The expression of a heterologous neochrome transcripts in the *Chlamydomans* genetic background allowed us to successfully screen and differentiate from native transcripts, resulting in tractable neochrome expression lines. We selected the strains A4C4, A4M1, UV4C14, and UV4M3 for further characterization, due to their improved growth compared to the rest of the transgenic lines. Neochrome expression lines showed decreased amounts of chlorophyll *b* and higher chlorophyll *a/b* ratios, which are both indicative of smaller antenna size.[27] Studies demonstrate that reducing the antenna complex size improves photosynthetic efficiency, by increasing light penetration depth.[27]

We also observed higher photosynthetic efficiency in neochrome lines supported by high *Fv/Fm* values. During one of our growth experiments conducted with wavelengths close to the neochrome excitation peak and the lines grown outside we observed the development of an oily phenotype (Figure 3A-D). The cells of the transgenic lines would not pellet where as the wild type cells did (Figure 6). The oily portion was not water soluble as it would settle to the bottom of the tube without mixing with the culture media. Further analysis is needed to identify the exact chemical structure of this compound. Nitrogen content was tested on these samples as it is known that nitrogen depletion can induce the accumulation of oil in microalgae.[20,21] We also grew samples in the same light conditions in nitrogen deplete conditions to determine if this would induce the observed oily phenotype. We, however, did not observe the phenotype as before. We tried to replicate the experiment that led to the oily phenotype multiple times, but we were not able to observe the phenotype again. Over the course of this time, we tested for the presence of the gene via PCR. However, due to the pandemic it was some time before gene expression could be tested. Over the eight months of trying to replicate the phenotype sometime along the way we lost detection of transgene expression across our cell lines.

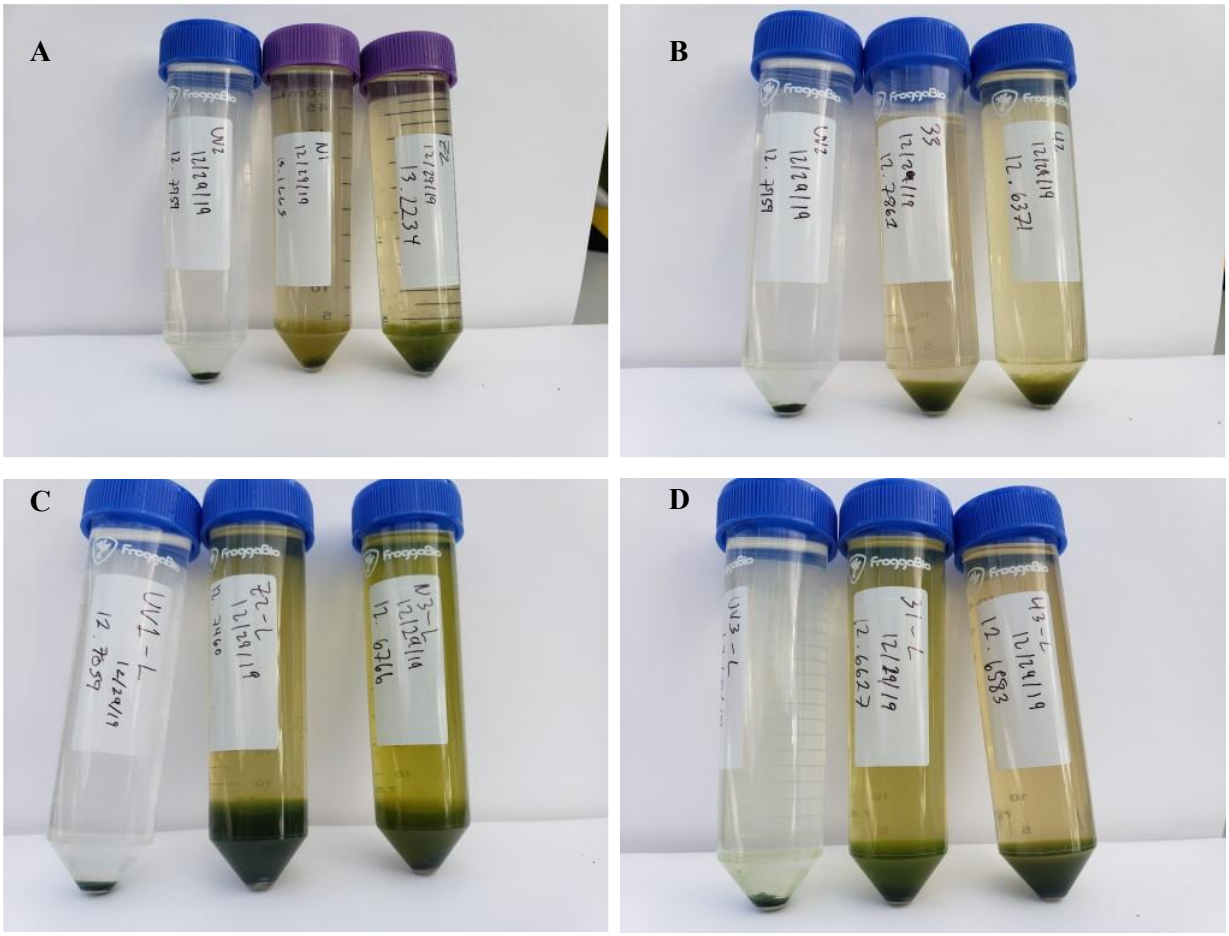


Figure 6: The oily phenotype observed during the experiment corresponding to Figure 4. In each panel the wild type is shown on the left with two transgenic lines. (A and B) were samples grown outside in a green house. (C and D) correspond to the samples grown inside at a light intensity of 35 µmol m^{-2} s^{-1}. (A) From left to right wild type, UV4C14, and UV4M3. (B) From left to right wild type, A4M4, and A4C4. (C) From left to right wild type, UV4M3, and UV4C14. (D) From left to right wild type, A4M4, and A4C4.

Gene expression in algae is not without its pitfalls. It has been shown that gene silencing occurs at transcriptional and post-transcriptional levels.[29,30] Silencing of exogenous DNA can make transgene expression very inefficient.[31,32] Genetic screens in the model alga *Chlamydomonas* have demonstrated that transgene silencing can be overcome by mutations in unknown gene(s), thus producing algal strains that can express foreign genes.[30,33,34] New research has identified a potential gene conferring the exogenous silencing and shown to be a sirtuin-type histone deacetylase.[34] The data suggest that transgenic DNA is recognized and then quickly inactivated by the assembly of a repressive chromatin structure composed of deacetylated histones. The

researcher's hypothesis that the silencing of foreign DNA is a defense mechanism against viruses and other deleterious types of environmental DNA sequences.[34] Several strategies have been studied to overcome gene silencing, which includes the construction of hybrid promoters, the inclusion of endogenous introns in the expression cassette, and codon optimization of the foreign gene. These strategies have alleviated the problem to some extent however, the expression of common transgenes such as GFP and YFP still remains challenging in microalgae. Another potential solution is increased transgene transcript levels as it has been shown that enhanced transgene expression correlates with greatly increased transcript levels.[30]

Neochrome did increase doubling time when cells were grown in low light conditions (with white light). With the addition of the far-red light, we initially observed an oily phenotype, but were not able to repeat the observation due to the loss of gene expression. Based on our preliminary results we can conclude that the expression of neochrome in algae offers a potential way to improve growth of microalgae in low light conditions like those found at dusk and dawn.

References

1. National Algal Biofuels Technology Review. *Energy.gov* https://www.energy.gov/eere/bioenergy/downloads/national-algal-biofuels-technology-review.
2. Alishah Aratboni, H., Rafiei, N., Garcia-Granados, R., Alemzadeh, A. & Morones-Ramírez, J. R. Biomass and lipid induction strategies in microalgae for biofuel production and other applications. *Microb Cell Fact* **18**, 178 (2019).
3. Kianianmomeni, A. & Hallmann, A. Algal photoreceptors: in vivo functions and potential applications. *Planta* **239**, 1–26 (2014).
4. Jaubert, M., Bouly, J.-P., Ribera d'Alcalà, M. & Falciatore, A. Light sensing and responses in marine microalgae. *Current Opinion in Plant Biology* **37**, 70–77 (2017).
5. Hegemann, P. Algal Sensory Photoreceptors. *Annu. Rev. Plant Biol.* **59**, 167–189 (2008).
6. Jiao, Y., Lau, O. S. & Deng, X. W. Light-regulated transcriptional networks in higher plants. *Nature Reviews Genetics* **8**, 217–230 (2007).
7. Kanegae, T., Hayashida, E., Kuramoto, C. & Wada, M. A single chromoprotein with triple chromophores acts as both a phytochrome and a phototropin. *Proc Natl Acad Sci U S A* **103**, 17997–18001 (2006).
8. Kanegae, T. Intramolecular co-action of two independent photosensory modules in the fern phytochrome 3. *Plant Signal Behav* **10**, (2015).
9. Christie, J. M. Phototropin Blue-Light Receptors. *Annu. Rev. Plant Biol.* **58**, 21–45 (2007).
10. Huang, K. & Beck, C. F. Phototropin is the blue-light receptor that controls multiple steps in the sexual life cycle of the green alga Chlamydomonas reinhardtii. *Proc Natl Acad Sci U S A* **100**, 6269–6274 (2003).

11. Im, C.-S., Eberhard, S., Huang, K., Beck, C. F. & Grossman, A. R. Phototropin involvement in the expression of genes encoding chlorophyll and carotenoid biosynthesis enzymes and LHC apoproteins in Chlamydomonas reinhardtii. *The Plant Journal* **48**, 1–16 (2006).
12. Yang, D., Seaton, D. D., Krahmer, J. & Halliday, K. J. Photoreceptor effects on plant biomass, resource allocation, and metabolic state. *PNAS* 201601309 (2016) doi:10.1073/pnas.1601309113.
13. Rockwell, N. C. & Lagarias, J. C. Phytochrome diversification in cyanobacteria and eukaryotic algae. *Current Opinion in Plant Biology* **37**, 87–93 (2017).
14. Duanmu, D., Rockwell, N. C. & Lagarias, J. C. Algal light sensing and photoacclimation in aquatic environments. *Plant Cell Environ* **40**, 2558–2570 (2017).
15. Kanegae, T. & Kimura, I. A phytochrome/phototropin chimeric photoreceptor of fern functions as a blue/far-red light-dependent photoreceptor for phototropism in Arabidopsis. *The Plant Journal* **83**, 480–488 (2015).
16. Li, F.-W. & Mathews, S. Evolutionary aspects of plant photoreceptors. *J Plant Res* **129**, 115–122 (2016).
17. Kindle, K. L. High-frequency nuclear transformation of Chlamydomonas reinhardtii. *Proc Natl Acad Sci U S A* **87**, 1228–1232 (1990).
18. Porra, R. J., Thompson, W. A. & Kriedemann, P. E. Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **975**, 384–394 (1989).

19. Rasala, B. A. *et al.* 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**, e43349 (2012).
20. Rasala, B. A. *et al.* Expanding the spectral palette of fluorescent proteins for the green microalga Chlamydomonas reinhardtii. *Plant J* **74**, 545–556 (2013).
21. Doron, L., Segal, N. & Shapira, M. Transgene Expression in Microalgae—From Tools to Applications. *Front Plant Sci* **7**, (2016).
22. Fischer, N. & Rochaix, J. D. The flanking regions of PsaD drive efficient gene expression in the nucleus of the green alga Chlamydomonas reinhardtii. *Mol. Genet. Genomics* **265**, 888–894 (2001).
23. Run, C. *et al.* Stable nuclear transformation of the industrial alga Chlorella pyrenoidosa. *Algal Research* **17**, 196–201 (2016).
24. Yang, B., Liu, J., Jiang, Y. & Chen, F. Chlorella species as hosts for genetic engineering and expression of heterologous proteins: Progress, challenge and perspective. *Biotechnology Journal* **11**, 1244–1261 (2016).
25. Ort, D. R., Zhu, X. & Melis, A. Optimizing Antenna Size to Maximize Photosynthetic Efficiency. *Plant Physiology* **155**, 79–85 (2011).
26. Ort, D. R. *et al.* Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *PROC. NAT. ACAD. OF SCI. (U.S.A.), Proceedings - National Academy of Science, U.S.A., Proceedings - National Academy of Sciences, USA, Proceedings of the National Academy of Sciences of the United States of, Proceedings of the National Academy of Sciences of the United States of America.* **112**, 8529–8536 (2015).

27. Perrine, Z., Negi, S. & Sayre, R. T. Optimization of photosynthetic light energy utilization by microalgae. *Algal Research* **1**, 134–142 (2012).

28. Negi, S. *et al.* Light regulation of light-harvesting antenna size substantially enhances photosynthetic efficiency and biomass yield in green algae†. *The Plant Journal* **103**, 584–603 (2020).

29. Fakhry, E. M. & El Maghraby, D. M. Lipid accumulation in response to nitrogen limitation and variation of temperature in Nannochloropsis salina. *Bot Stud* **56**, 6 (2015).

30. Yang, D. *et al.* Lipidomic Analysis of Chlamydomonas reinhardtii under Nitrogen and Sulfur Deprivation. *PLOS ONE* **10**, e0137948 (2015).

31. Cerutti, H., Ma, X., Msanne, J. & Repas, T. RNA-Mediated Silencing in Algae: Biological Roles and Tools for Analysis of Gene Function. *Eukaryotic Cell* **10**, 1164–1172 (2011).

32. Tran, N. T. & Kaldenhoff, R. Achievements and challenges of genetic engineering of the model green alga Chlamydomonas reinhardtii. *Algal Research* **50**, 101986 (2020).

33. Banerjee, S., Banerjee, S., Ghosh, A. K. & Das, D. Maneuvering the genetic and metabolic pathway for improving biofuel production in algae: Present status and future prospective. *Renewable and Sustainable Energy Reviews* **133**, 110155 (2020).

34. Rasala, B. A., Chao, S.-S., Pier, M., Barrera, D. J. & Mayfield, S. P. Enhanced Genetic Tools for Engineering Multigene Traits into Green Algae. *PLOS ONE* **9**, e94028 (2014).

35. Schroda, M. RNA silencing in Chlamydomonas: mechanisms and tools. *Curr Genet* **49**, 69–84 (2006).

36. Neupert, J. *et al.* An epigenetic gene silencing pathway selectively acting on transgenic DNA in the green alga Chlamydomonas. *Nat Commun* **11**, 6269 (2020).

Part 5 Additional Significant Projects

My work during my PhD expanded beyond the chapters presented previously. Each piece of work I will expand on and present in its own section in this chapter.

Transformation of *Chlorella sorokiniana* 1412 with a SnRK1 from *Chlorella sorokiniana* 1230 and the development of a reliable *Chlorella sorokiniana* method

As shown in chapters 1 and 2 we have seen that SnRKs play an important role in plant carbon metabolism. My work has also shown that the overexpression of an algal SnRK2 increased growth and starch content in *Chlorella sorokiniana* 1412. At the beginning of this project, I also transformed *C. sorokiniana* 1412 with a SnRK1 from *Chlorella sorokiniana* 1230. These are two different strains of the same species. SnRK1s in plants are key contributors to regulating global metabolism and energy status. Another important part of this work was developing a reliable method for *C. sorokiniana* 1412 transformation. *C. sorokiniana* 1412 offers many advantages such as rapid growth, low cost, ease of culture , achieve high cell density under various culture methods, and broad industrial applications that range from biodiesel feedstock to food additives. [1]. Chlorella spp. have shown to produce high value products such as lipids, proteins, carotenoids, vitamins, and minerals.[2,3]. This is due Chlorella's naturally high lipid content, which allows them to be one of the best candidates for biofuel production.[4]

However, the biggest disadvantage in using *Chlorella* is that transformation methods have yet to be optimized or attempted in many species including *C. sorokiniana*.[1,5] This hinders the ability to engineer promising Chlorella strains for biotechnological purposes. Previous methods include glass bead, *Agrobacterium tumefaciens* mediated method, polyethylene glycol(PEG) mediated method, and electroporation.[6] Each of these methods has its advantages, but research has shown that electroporation is the preferred method for *Chlorella* spp because of its ease of use and

efficiency. [1,6] However, this method has been applied to only a few species of commercial interest. Other disadvantages in using *Chlorella* are the low efficacy of transformation methods and low transgene expression levels from the nuclear genome. These disadvantages limit the use of *Chlorella* species as an expression platform.[5,7] Research has shown that the addition of the bleomycin/zeocin-resistance sh-ble gene can improve heterologous expression of transgenes. [7,8]. Furthermore, the sh-ble gene has been used to generate transgenic clones from different algae species, but this has yet to be demonstrated in *Chlorella sorokiniana*. [5]. Therefore, we aimed to develop an efficient method to transform *C. sorokiniana* 1412 using electroporation.

We have successfully confirmed the integration of the Zeocin resistant gene (*Sh-ble*) and *SNRK-1* gene overexpression cassette into *C. sorokiniana* 1412 by PCR and DNA sequencing the PCR products. We further confirmed the overexpression of SNRK-1 using qPCR analysis.

Once the transformation protocol was established it was deployed to generate our transgenic lines. In Figure 1 we show the successful integration of our SnRK1 gene. Sequence analysis was done for both the selection marker and gene of interest. This analysis confirmed that our gene was integrated intact and was not fragmented. As shown below in Figure 2 the two transgenic lines show 2-8 fold increase in gene expression. Once our gene was confirmed to being expressed, we moved into analyzing its phenotype. Like SnRK2 (Chapter 2) we observed that SnRK1 overexpression led to an increase in cell size as well as starch accumulation (Figure 3). Further similarly to SnRK2 expression we observed a change in chlorophyll content as seen by a lower amount of total chlorophyll and a higher chlorophyll a/b ratio compared to the wild type (Figure 4). SnRK 1 transgenic lines also had a higher oxygen evolution rate (Figure 4). These results imply that SnRK 1 could potentially play a role in carbon metabolism and chlorophyll synthesis with *C. sorokiniana* 1412.

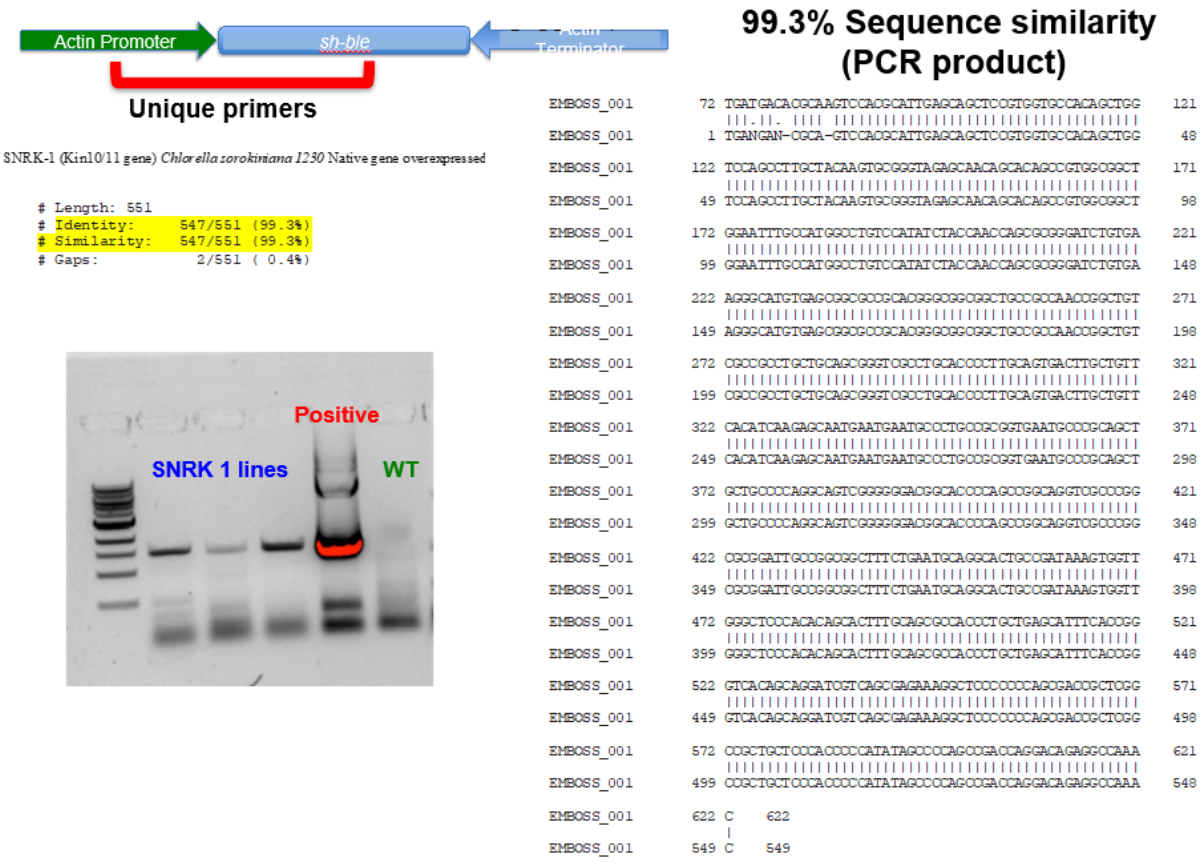


Figure 1: PCR confirmation, sequence confirmation and q-PCR confirmation of SnRK1 insertion into *Chlorella sorokiniana* 1412.

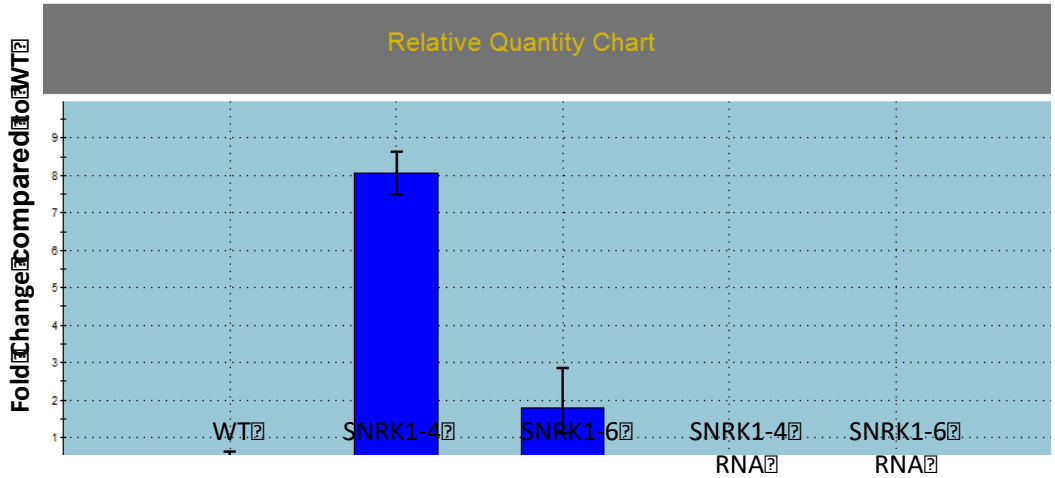
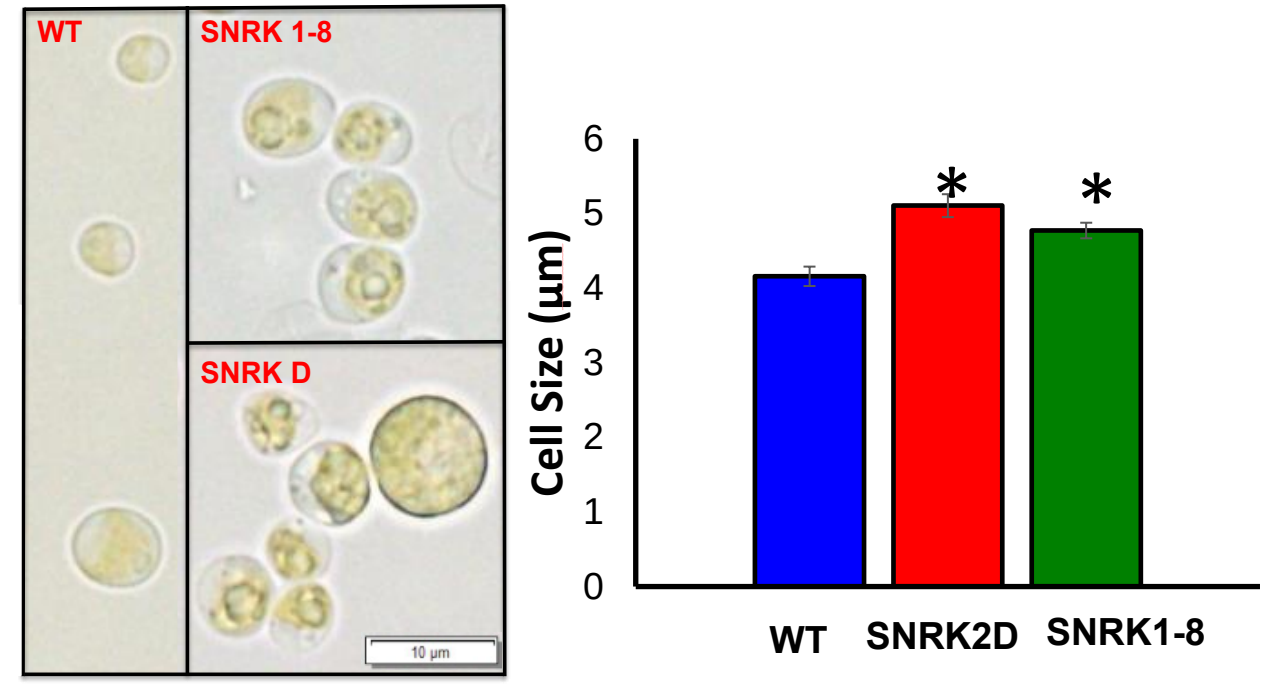


Figure 2: QPCR confirmation of two SNRK1 transgenic lines. Both transgenic lines showed higher expression of SNRK1 compared to wild type. To ensure there was no DNA contaminations in RNA samples we also ran a expression analysis with RNA samples and found no SNRK1 expression. These experiments were performed in triplicates.



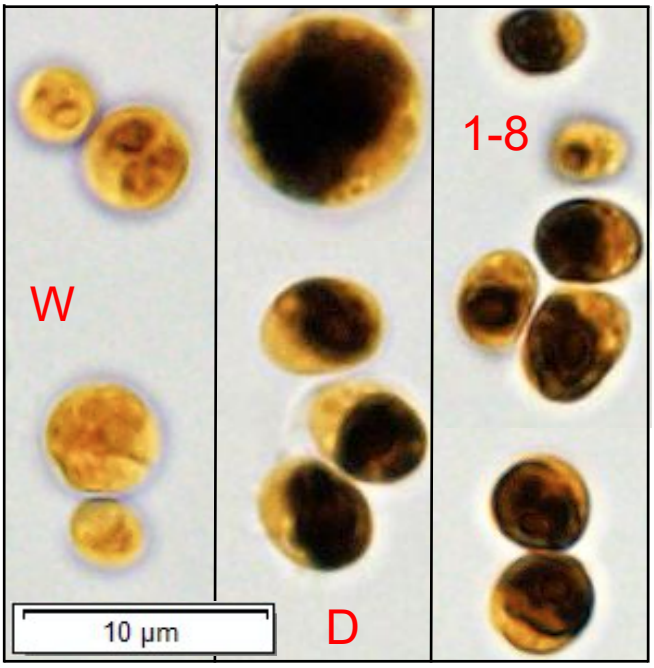


Figure 3: Increase cell size observed with microscopy along with a transgenic lines being stained darker with Lugol's stain.

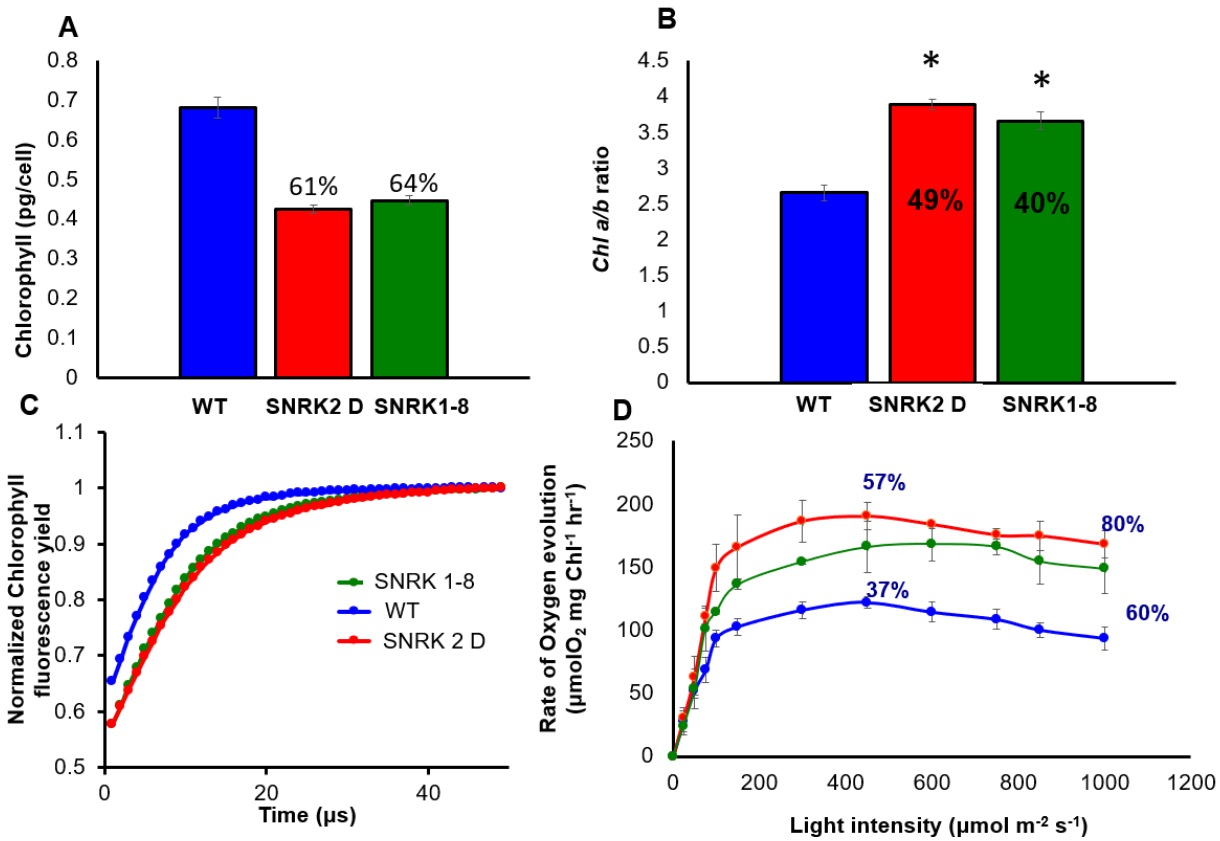


Figure 4: (A) SnRK 1-8 like SnRK2D (Part 2) has a lower amount of total chlorophyll while having a higher chlorophyll a/b ratio (B). This indicates a shrinking of the antenna complex. This is further supported by our flash fluorescence induction experiment (C). SNRK1-8 also showed that it generated more oxygen over a light curve compared to the wild type. Demonstrating a higher photosystem II efficiency.

An effective transformation method for *Scenedesmus obliquus* 393 and the development of neochrome and Cas9 expression lines

To be able to move forward with expressing neochrome and the gene coding for Cas9 we needed to have an effective transformation method for the *S. obliquus* 393 species. There was a previously published method on *S. obliquus* transformations via electroporation.[9] Our initial attempts to replicate these methods were unsuccessful as we had no successful transformants (Figure 5). Each attempt using this protocol yielded colonies with no vector integration as well as many colonies growing on the control plates (wild type cells plated after electroporation on selection media). The plasmid we used to test this method contained a fluorescent protein (mPapaya) and the selection marker paromomycin. After these initial failures we decided to move forward developing our own protocol.

The first thing we did was create an antibiotic kill curve with paromomycin, zeocin, and blasticidin to find the optimal antibiotic concentration for selection (Table 1). Using knowledge gained from working on the *C. sorokiniana* 1412 transformation method we began to systematically change a some of the parameters associated with electroporation. These included the voltage amount, the time constant (the amount of time the voltage is applied), and the amount of vector DNA (Tables 2-4). Experiments were conducted in triplicate or duplicate the numbers shown in the tables are the averages of the colonies counted from the selection plates. All cells were electroporated with a Bio-Rad Gene Pulser Xcell electroporator, in a 2mm cuvette with Maxx efficiency electroporation buffer. Colonies were grown for 25 days and then screened for the presence of the selection vector (paromomycin or zeocin). The only successful colonies were observed with these conditions: 2.5kv, 5ms (time constant), 2 or 5µg of DNA, and plated on Zeocin 30, 40 or Paromomycin 80 (not shown). Using 5µg of DNA yielded more DNA, so decided to move forward with these conditions.

Once this method was confirmed to work, we used it to generate *S. obliquus* neochrome expression lines as well as a Cas9/GFP expression line. Our neochrome expression vector has zeocin for its selection marker while paromomycin is used in the Cas9/GFP expression vector. Colonies were screened for the presence of these genes. Further RT-PCR was used to confirm expression in the neochrome lines this has yet to be conducted in the Cas9 line. We did however, run Cas9 colonies through a Accuri flow cytometer to screen for GFP signal. We did identify one colony with the correct signal. PCR was used to confirm the presence of the Cas9 gene. Our vector was created using Gibson assembly. The Cas9 gene was GFP linked on the 3' end. We also tested to confirm that zeocin selection was maintained on the cultures during growth. This was tested by adding zeocin to liquid cultures of the transgenic lines as well as the wild type. This was also conducted

on plates. A few initial growth curves were conducted on the *S. obliquus* lines and cells were observed under the microscope. Our initial findings show that the neochrome expression lines have larger cells compared to the wild type cells.

Table 1: Antibiotics used to determine the best conditions for *Scenedesmus* selection.

Antibiotic	Organism	Concentration (µg/mL)	Kill Concentration (µg/mL)
Paromomycin	Scenedesmus obliquus	5, 10, 20, 30, 40, 50, 60, 70, and 80	60
Blasticidin	Scenedesmus obliquus	5, 10, 20, 30, and 40	30
Zeocin	Scenedesmus obliquus	1, 5, 10, 20, 30, 40, 50, 60, and 70	70

Table 2: Various transformation parameters tested to determine the optimal conditions for transformation. The values shown below represent the average number of colonies found on three plates for each condition indicated.

Plated on Paromomycin 60µg/mL	45 µg DNA 2.5 kv 3ms	15 µg DNA 2.5 kv 5ms	15 µg DNA 2.5 kv 10ms	15 µg DNA 2.5 kv 20ms
Transgenic	11±2	11±3	8±1	5±1
WT	12±1	11±2	7±1	7±2

Table 3: Various transformation parameters tested to determine the optimal conditions for transformation. The values shown below represent the average number of colonies found on three plates for each condition indicated.

Plated on Zeocin 30 µg/mL	2 µg DNA 2.5 kv 5ms	5 µg DNA 2.5kv 5ms
Transgenic	10±2	21±3
WT	3	3

Table 4: Various transformation parameters tested to determine the optimal conditions for transformation. The values shown below represent the average number of colonies found on three plates for each condition indicated.

Plated on Zeocin 40 µg/mL	2 µg DNA 2.5 kv 5ms	2 µg DNA 2.5 kv 10ms	2 µg DNA 2.5 kv 20ms	5 µg DNA 2.5kv 5ms
Transgenic	0	0	0	4±1
WT	0	0	0	0

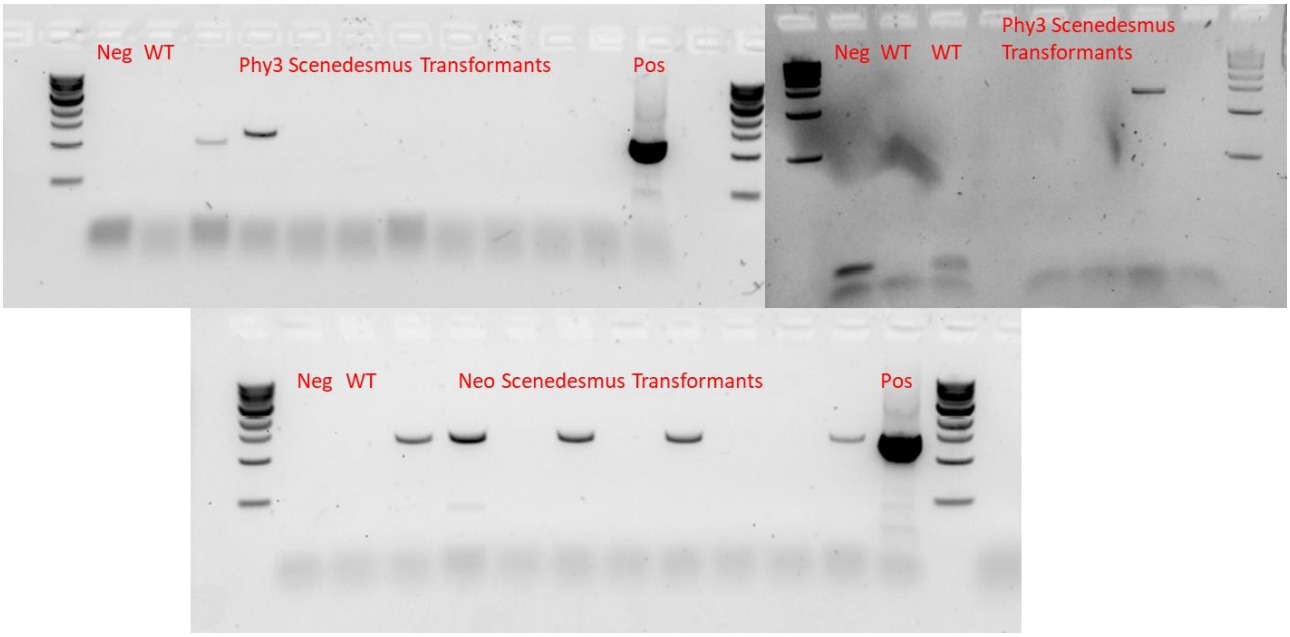


Figure 5: RT-PCR used to confirm the expression of neochrome in *Scenedesmus*. The band was only found in the transgenic lines and was absent in the wild type cells.

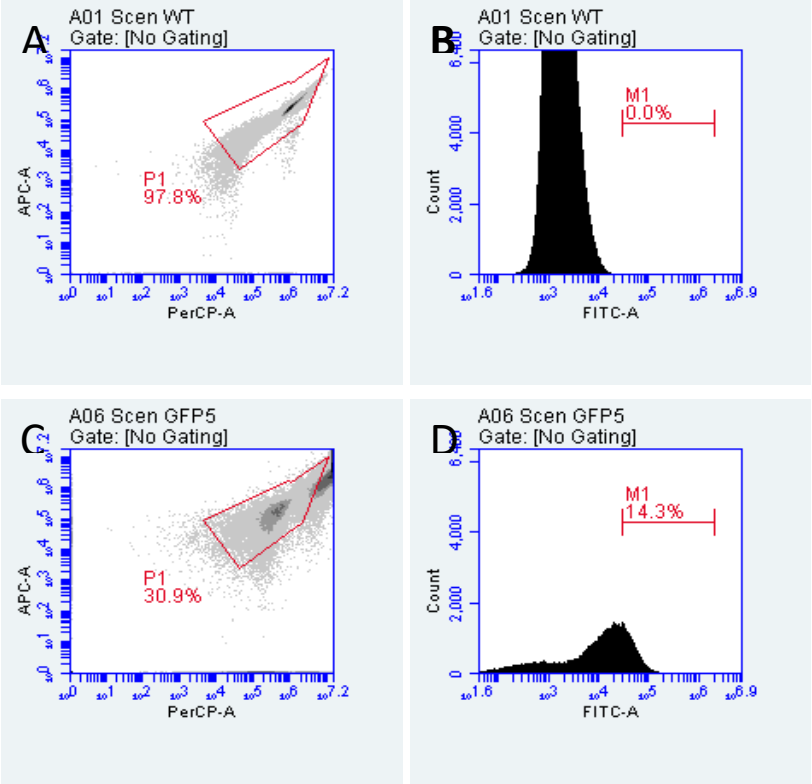


Figure 6: WT *Scenedesmus* and GFP expression line observed under flow cytometry.

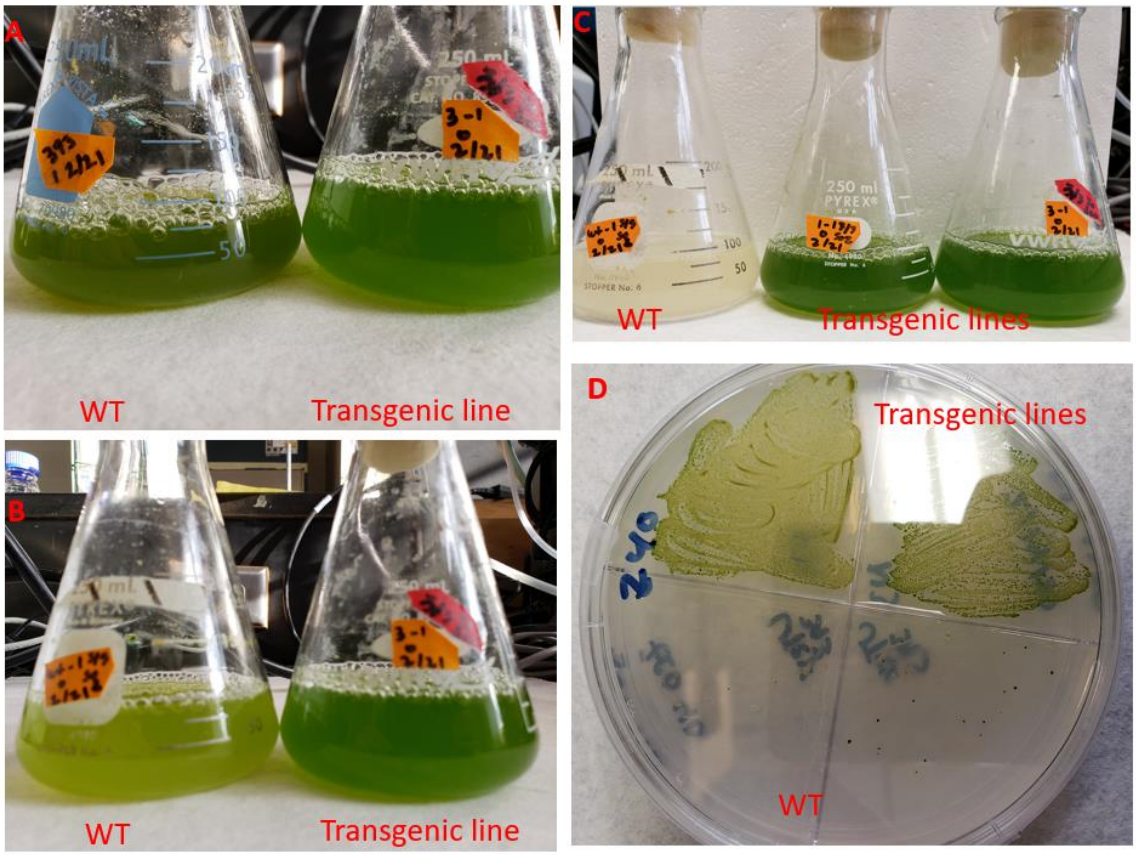


Figure 7: Wild type and transgenic *Scenedesmus* grown in the presence of the zeocin antibiotic. (A) Wild type and one transgenic line before zeocin treatment, (B) The same samples from (A) after 1 on week of treatment. (C) 2 weeks after zeocin treatment. (D) wild type and transgenic lines grown on media contain the zeocin antibiotic.

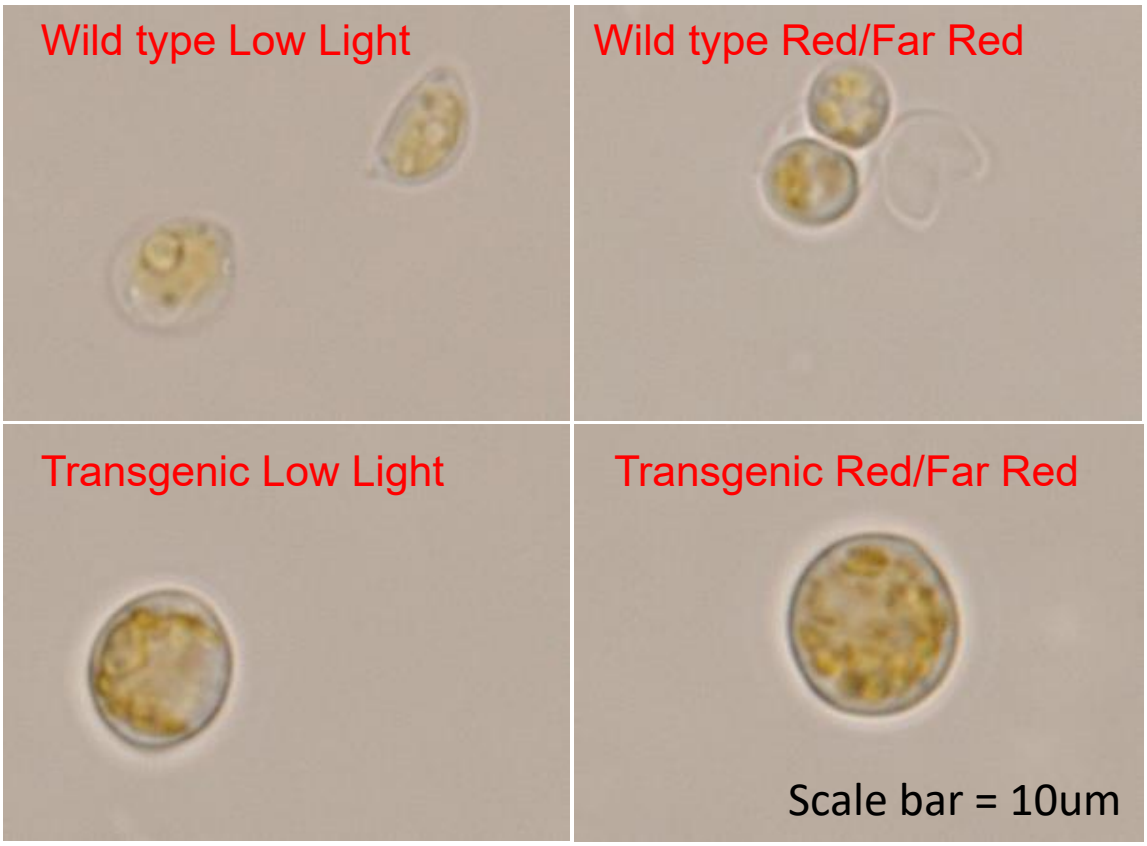


Figure 8: Wild type *Scenedesmus* and transgenic lines expression neochrome after being grown at a light intensity of 35 µmol m^{-2} s^{-1} with the wavelength necessary to excite neochrome. This single experiment demonstrates that neochrome might a have positive effect on increasing *Scenedesmus* size.

The generation of a *Picochlorum soloecismus* CPF1 expression line and the generation of phototropin knockout lines

I also worked on generating a *Picochlorum soloecismus* CPF1 expression line. In Figure 9 we show the plasmid used to generate these expression lines. Transformation of our vector was preformed using electroporation. Cells were spun down then washed two times with MAXX efficiency electroporation buffer and were finally re-suspended in the same buffer. Approximately 5ug of plasmid DNA was added to the cells. Cells were electroporated at a voltage of 1.6 KV with a time constant of 20 milliseconds. Cells were plated on 40ug/mL zeocin containing F/2 agar plates for antibiotic screening. Zeocin resistant transgenic colonies were re-streaked onto fresh zeocin containing plates. Colonies were subjected to PCR validation for CPF1 plasmid insertion using

primers specific to sfGFP gene present within the plasmid. Once the presence of the gene was confirmed via PCR and gene sequencing RT-PCR was used to determine if the gene was being expressed Figure 10. We also confirmed that zeocin was being expressed by adding the antibiotic to our liquid cultures Figure 6. With these confirmations we moved forward with trying to generate phototropin knockouts in *P. soloecismus*. To generate these knockouts, we began to transform our CPF1 expression line with guide RNAs and repair template.

Our repair template was generated using Gibson assembly. The assembly was confirmed via sequencing of the *Picochlorum* repair template. The template is composed of: PUC 19 backbone (Ampicillin resistance), RG007 zeocin resistance, and homology arms (both 5'and 3') for Cpf1-directed homologous repair within the phototropin gene of *Picochlorum*. Potential phototropin knockout colonies generated with our cpf1 and *Picochlorum* repair template were screened for the insertion of the repair template or small deletion of the gene. We however were unable to identify a successfully generated knockout. We hypothesized that maybe our guide RNAs were being degraded before they could bind to the target DNA. With this information we begin to create a vector that would express our guide RNAs. This work was unable to be completed due to shutdowns associated with COVID 19.

P. soloecismus transformations were also conducted to generate lines expressing the neochrome gene. Cells were transformed as mentioned above. Selection occurred on zeocin antibiotic However, our initial colonies were not able to be screened due to the onset of the pandemic.

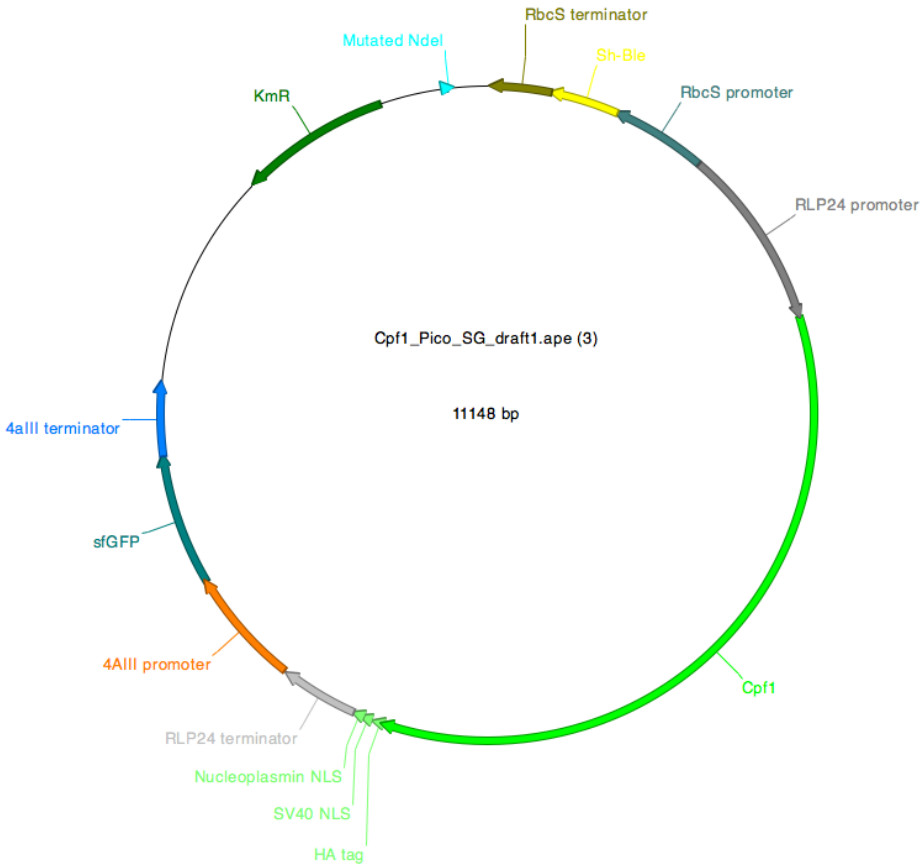


Figure 9: *Picochlorum* CPF1 expression vector map. CPF1 expression is under control of the RLP24 promoter/terminator pair, while CPF1 is under control of a 4AIII promoter and terminator pair. Zeocin was used as our selection antibiotic.

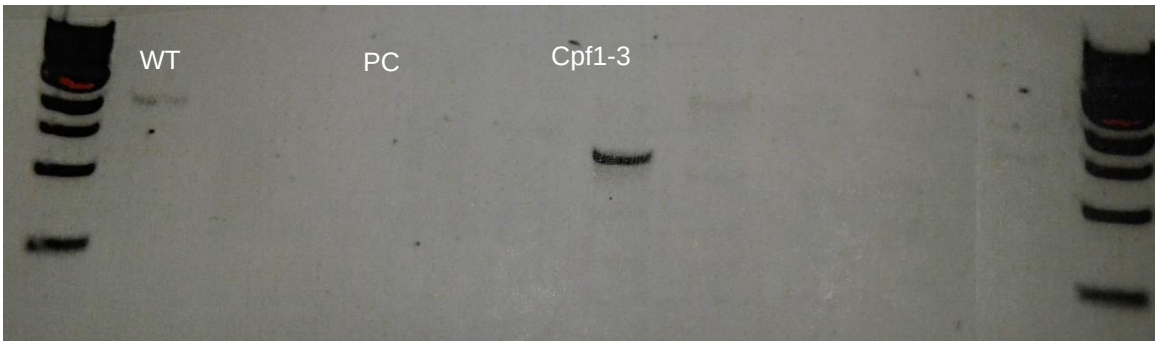


Figure 10: RT-PCR confirming the expression of CPF1. This band is absent in the wild type.

Figure 11: *Picochlorum* wild type and CPF1 expression line subjected to antibiotic treatment. From left to right *Picochlorum* wild type grown for 2 weeks without treatment, *Picochlorum* wildtype grown for 2 weeks with treatment, and CPF1 expression line grown for 2 weeks with treatment.

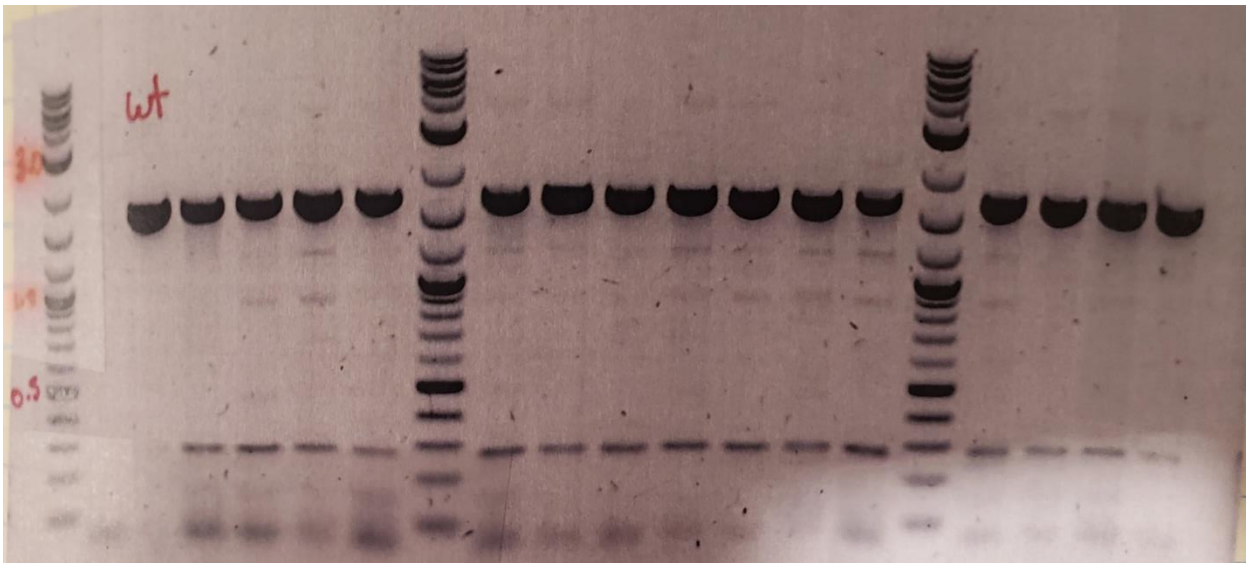


Figure 5: Attempt to generate phototropin knockout lines in *Picochlorum*. Bands observed match those found in the wild type indicating our repair template was not inserted nor was there deletions as this was confirmed with sequence analysis.

The development of phototropin knockout lines in *Scenedesmus obliquus* 393

To develop these phototropin knockout (PHOT KO) lines we aimed to make use of ribonucleoprotein (RNP) technology. RNPs consist of the Cas9 protein in a complex with a guide RNA. We designed three different guides to target the phototropin gene within *S. obliquus*. Our transformations took place with and without the presence of a repair template. PHOT KO lines if successfully transformed will have our dsDNA repair template inserted at the PAM cut site of cas9. This will lead to a detectable insertion of approximately 3KB. As seen in Figure 12 we have been unable to generate any PHOT KO lines as our transformant bands are the same size as the WT. We attempted many different methods and ratios of Cas9 protein to guide RNA as show in Table 5. We thus set out to determine what maybe inhibiting Cas9 form creating a cut in the protropin gene.

Using confocal microscopy, *S. obliquus* 393 cells were imaged with the goal to visualize whether the Cas9 protein (as we have previously been using) is in fact entering the algal cells. We used a GFP labeled Cas9 protein to investigate. We found that GFP signal was aggregating on the exterior of the algal cells. None had localized fluorescence inside. The nuclear localization signal should have yielded a bright signal concentrated at the nucleus, but that was not the case. Second imaging used to disassociate exteriorly aggregated gfp-cas9 from the cells and yield a clearer image. This screening also did not exhibit any algal cells with internal GFP fluorescence. In conclusion, Cas9 was not entering UTEX 393 cells as a full protein and our efforts shifted to creating a Cas9 line via plasmid electroporation (see above).

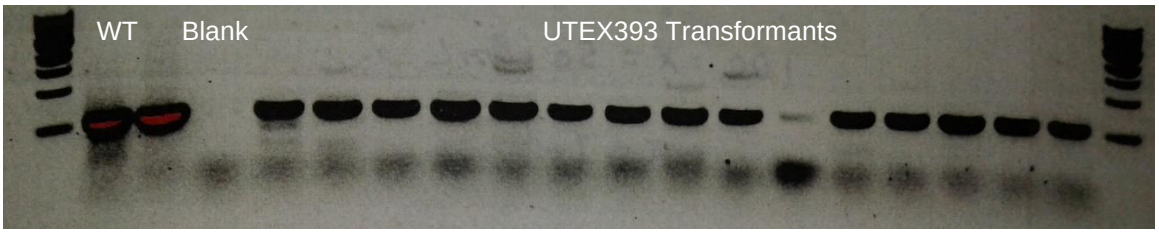


Figure 13: Attempt to generate phototropin knockout lines in *Scenedesmus.* Bands observed match those found in the wild type indicating our repair template was not inserted nor was there deletions as this was confirmed with sequence analysis.

Table 5: Colonies generated for various conditions using CAS9 RNP complex with two different guide ratios.

	3:1	**9:1**
No guides	0	0
Pooled guides no repair	0	0
Pooled guides with repair	0	1
Sg 1	0	0
Sg2	0	0
Sg3	1	3
Control	0	0

Dual-Pam Analysis

During my time doing various PAM (pulse amplitude modulated fluorometers) measurements using a Walz Dual-PAM I noticed varying values when my stir speed changed slightly. I decided to investigate this a bit more and to test a few different methods of agitation. Using *Chlamydomonas* wild type cells I tested stirring the samples, shacking the whole apparatus, and

use no agitation. To my surprise the values differed significantly especially at higher light intensities experienced during the light curve.

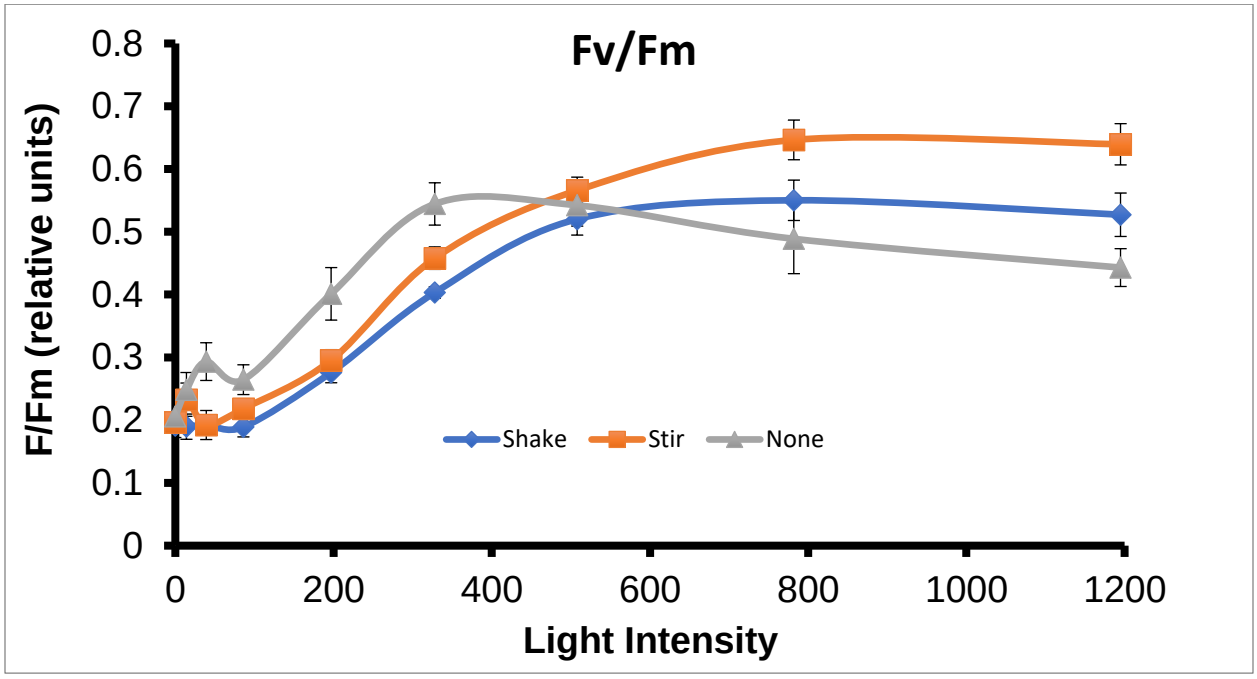


Figure 14: Results of triplicate measurements of *Chlamydomonas* wild type when measuring Fv/Fm over a light curve with two different agitation methods.

Table 6: P-values calculated between 3 points on the light curve from Figure 12. There was a significant difference at high light intensities.

	508	**782**	**1195**
P-value between none and stir	N/A	0.0389	0.0023
P-value between shake and stir	N/A	N/A	0.0467

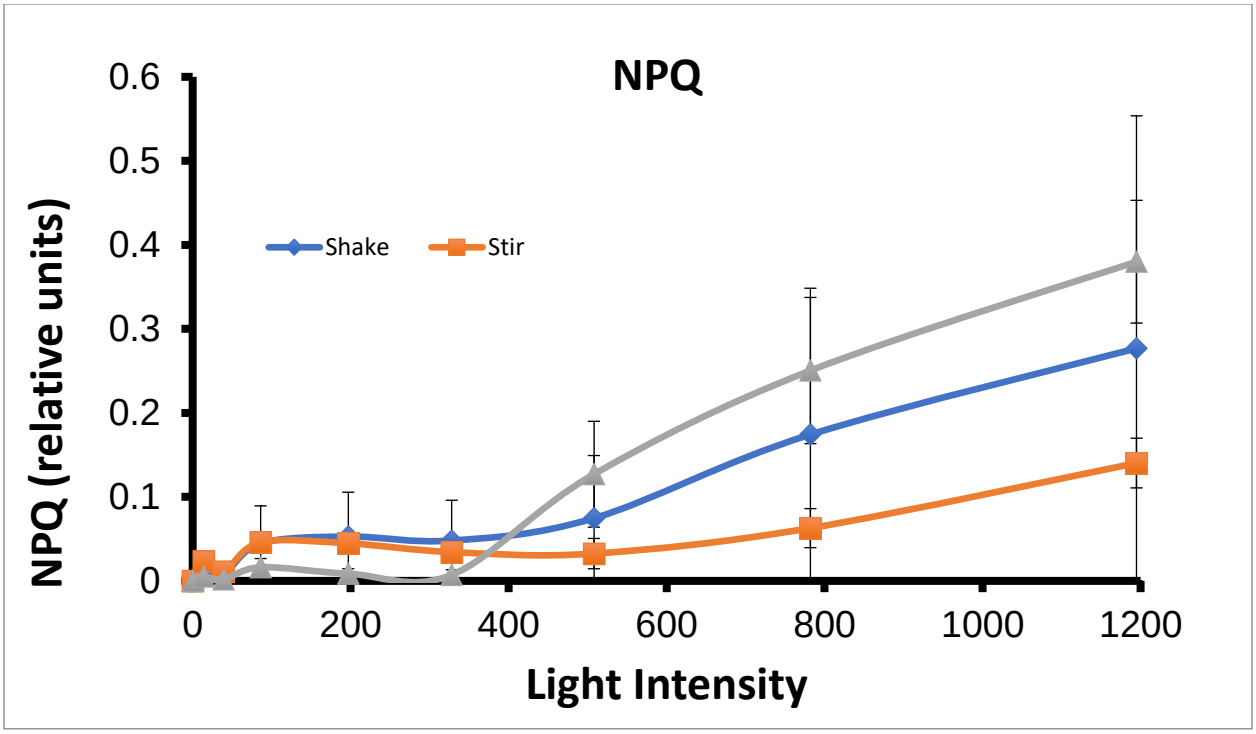


Figure 15: Results of triplicate measurements of *Chlamydomonas* wild type when measuring NPQ over a light curve with two different agitation methods.

Table 7: P-values calculated between 3 points on the light curve from Figure 12. There was a significant difference at high light intensities

	508	**782**	**1195**
P-value between none and stir	N/A	0.0707	0.0148
P-value between shake and stir	N/A	0.0456	0.0257

Conclusions and Future Directions

Across my PhD I worked on various transformation methods for *Scenedesmus obliquus* 393 and *Chlorella sorokiniana* 1412 as well as generating various expression lines. These lines were developed to express SnRKs, neochrome, CPF1, and CAS9. Some of these were more successful than others. To perform these experiments, I had to learn about various algal transformation

methods while focusing on electroporation and glass bead methods. I also had to learn how to optimize these methods and various factors that can affect efficiency such as cell count, cell health, antibiotic concentration, media selection, etc... I then had to learn how to assess the success of these transformations. This entailed PCR (hundreds and hundreds of PCRs), quantitative PCR methods, sequencing analysis, and western blots. Once expressed phenotype analysis was done to concluded weather or not our transformations yielded any results. I learned how to conduct various photosynthetic measurements including oxygen evolution, PAM, chlorophyll analysis, Lugol staining, BODIPY staining, flow cytometry, and observing cells under the microscope. Along the way I learned about algal culturing techniques and how these can be optimized for individual species of algae. I learned a lot about failure as many of these techniques had not been optimized for some of these algal species. Sometimes it is taken for granted but something as easy as DNA extraction was quite difficult on these agal species and optimizing these methods took many iterations most times with no success.

Due to COVID some of my work had to switch to being remote which allowed me to gain a more in-depth knowledge about algal phylogenetics. I was also able to apply knowledge I gained from various classes to build my own phylogenetic trees based on the SnRK protein families. I learned how to annotate genomes, analyzing protein domains, and identify proteins in various algal lineages. All this allowed me to identify SnRK family members in thirty-four agal species. Hopefully providing us insight on the genes needed for land adaptation.

The most promising thing I would do next would be to further identify genes involved in land adaption of algae and the evolution of algae to plants. This would entail analyzing more terrestrial algae. This would include genome assemblies and annotations something that is currently lacking in the field. Further, there is large number of datasets currently available, but they lack the quality

for any formal analysis further work could include improving some of these current assemblies. I would also test if the identified SnRK3s from part 3 respond to salt or other stress conditions. This could be done with transcriptome analysis and/or q-PCR analysis to observe how SnRK3 expression changes over time in response to these stress conditions.

Citations

1. Yang, B., Liu, J., Jiang, Y. & Chen, F. Chlorella species as hosts for genetic engineering and expression of heterologous proteins: Progress, challenge and perspective. *Biotechnol. J.* **11**, 1244–1261 (2016).
2. Plaza, M., Herrero, M., Cifuentes, A. & Ibáñez, E. Innovative Natural Functional Ingredients from Microalgae. *J. Agric. Food Chem.* **57**, 7159–7170 (2009).
3. Pulz, O. & Gross, W. Valuable products from biotechnology of microalgae. *Appl. Microbiol. Biotechnol.* **65**, 635–648 (2004).
4. Liu, L. *et al.* Development of a New Method for Genetic Transformation of the Green Alga Chlorella ellipsoidea. *Mol. Biotechnol.* **54**, 211–219 (2013).
5. Doron, L., Segal, N. & Shapira, M. Transgene Expression in Microalgae—From Tools to Applications. *Front. Plant Sci.* **7**, (2016).
6. Run, C. *et al.* Stable nuclear transformation of the industrial alga Chlorella pyrenoidosa. *Algal Res.* **17**, 196–201 (2016).
7. Rasala, B. A. *et al.* 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**, e43349 (2012).
8. Rasala, B. A. *et al.* Expanding the spectral palette of fluorescent proteins for the green microalga Chlamydomonas reinhardtii. *Plant J.* **74**, 545–556 (2013).
9. Guo, S.-L. *et al.* Establishment of an efficient genetic transformation system in Scenedesmus obliquus. *J. Biotechnol.* **163**, 61–68 (2013).

Supplemental Figures

A

Chlorella sorokiniana 1412 Accession	E-value	Pfam (Clan)
CSJ00008848	E-108	CL0016
CSJ00006659	1E-79	CL0016
CSJ00011539	4E-75	CL0016
CSJ00000593	5E-74	CL0016
CSJ00002346	1E-70	CL0016
CSJ00001882	2E-66	CL0016
CSJ00002345	2E-56	CL0016
CSJ00011026	1E-55	CL0016
CSJ00004407	2E-55	CL0016
CSJ00005514	1E-52	CL0016
CSJ00005719	1E-50	CL0016
CSJ00009858	2E-48	CL0016
CSJ00004491	4E-41	CL0016
CSJ00000848	1E-37	CL0016

B

Picochlorum soloecismus Accession	E-value	Pfam (Clan)
NSC_03857	8.37E-151	CL0016
NSC_00326	1.30E-106	CL0016
NSC_00324	6.69E-92	CL0016
NSC_00414	7.15E-83	CL0016
NSC_01618	7.32E-72	CL0016
NSC_03481	3.33E-65	CL0016
NSC_06295	3.65E-65	CL0016
NSC_04901	1.72E-62	CL0016
NSC_06478	1.98E-52	CL0016
NSC_06572	7.37E-51	CL0016

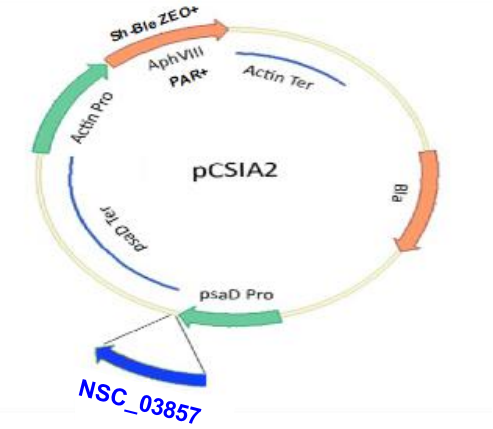
Figure 1: Pfam hits for the identified SnRKs in *Chlorella sorokiniana* 1412 and *Picochlorum soloecismus*.

	Pico SnRK	1412 SnRK
SnRK2.1	65.16%	65.29%
SnRK2.2	61.49%	61.82%
SnRK2.3	62.73%	63.53%
SnRK2.4	65.13%	63.58%
SnRK2.5	64.35%	63.86%
SnRK2.6	64.29%	63.41%
SnRK2.7	63.29%	60.57%
SnRK2.8	65.18%	61.32%
SnRK2.9	63.17%	60.31%

SnRK2.10	64.04%	64.80%
Pico SnRK	100%	75.38%

Figure 2: Percent identities of the top candidates corresponding to SnRK2 identified in *Chlorella sorokiniana* 1412 and *Picochlorum soloecismus* when compared to those found in *Arabidopsis*.

A



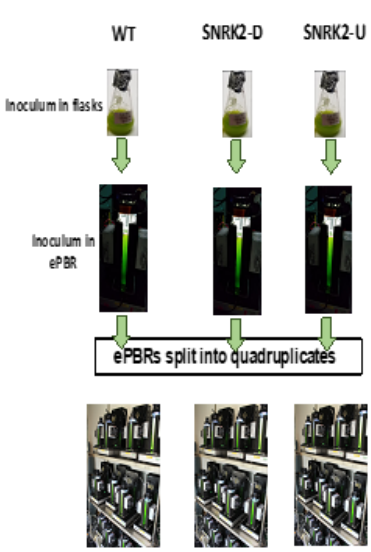
B

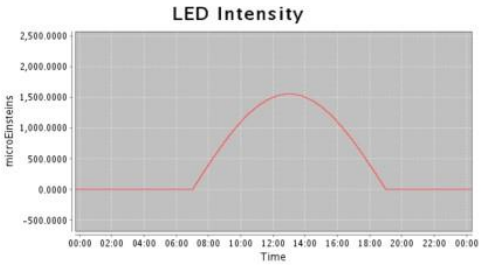
	FWD	REV	Amplicon Length (bp)
PSAD-SnRK2	5' GCTGTTCACATCAAGAGCAATGAATGAATG 3'	5' CTAGGATCGTTTGCATCTTCAAAGGGATATG 3'	960
SNRK-RT	5' AACCATCCAACGGATCATGG 3'	5' CACTTCTCCGATAATTTCTTTCTCATC 3'	200
Actin-RT	5' GGCTTTGCGGGCGATGACGCT 3'	5' CATGATGGAGTTGAAGGTGGTCTC 3'	199

Figure 3: (A) Vector used for the expression of SnRK. The SnRK gene was under control of the PSAD promoter/terminator pair. (B) The primer sequences used for confirmation of expression and insertion of the SnRK gene.

Outdoor Growth Simulation

ePBR Software Settings	
Algal command version used	2.4
Script used	N/A
Light intensity	1559
Peak intensity	13:00/1PM
Day length	12hours
Light delivery	sinusoidal
Temperature set point(°C)	21
Temperature amplitude	6
Temperature peak	16:00/4PM
Temperature delivery	sinusoidal
pH set point	7.9
tolerence	0.05
CO2 delivery	pH controlled
Spin bar RPM	200





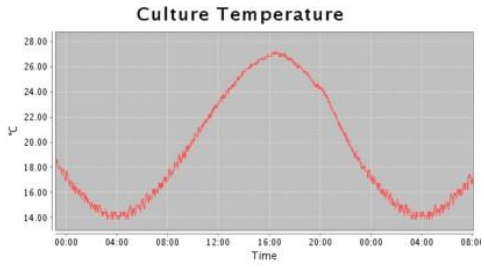


Figure 4: Parameters and experimental setup for the EPBR outdoor growth simulation.

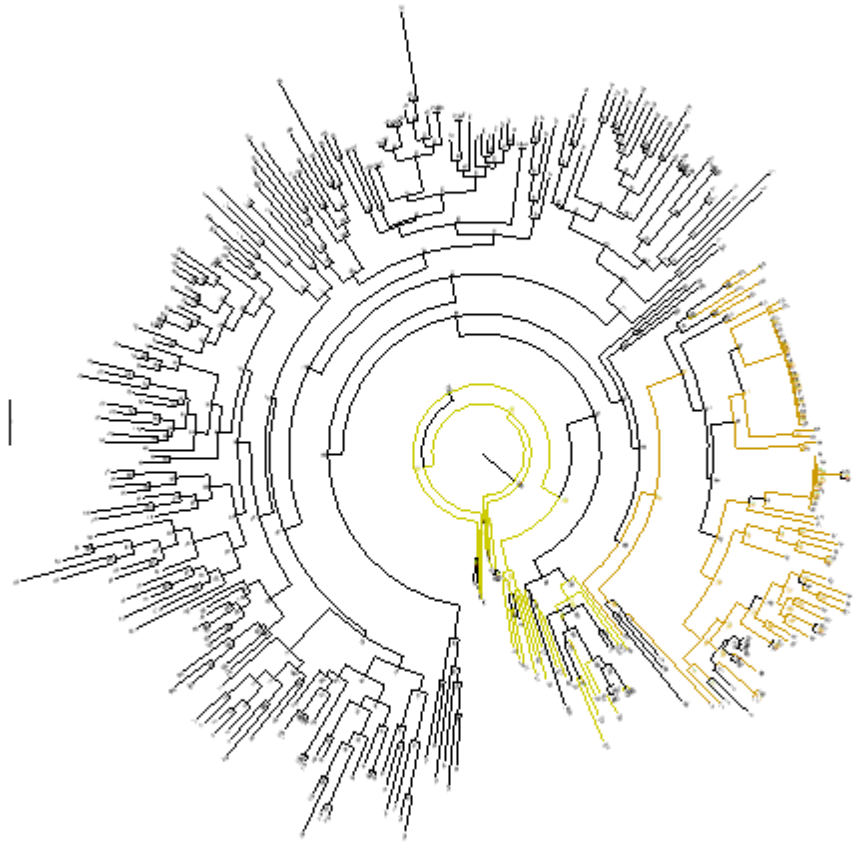


Figure 5: To further investigate phylogenetic relationships Markov Chain Monte Carlo (MCMC) in MrBayes version 3.2.7a was used. This method yielded different results compared to the maximum likelihood method, but the overall trends were the same.

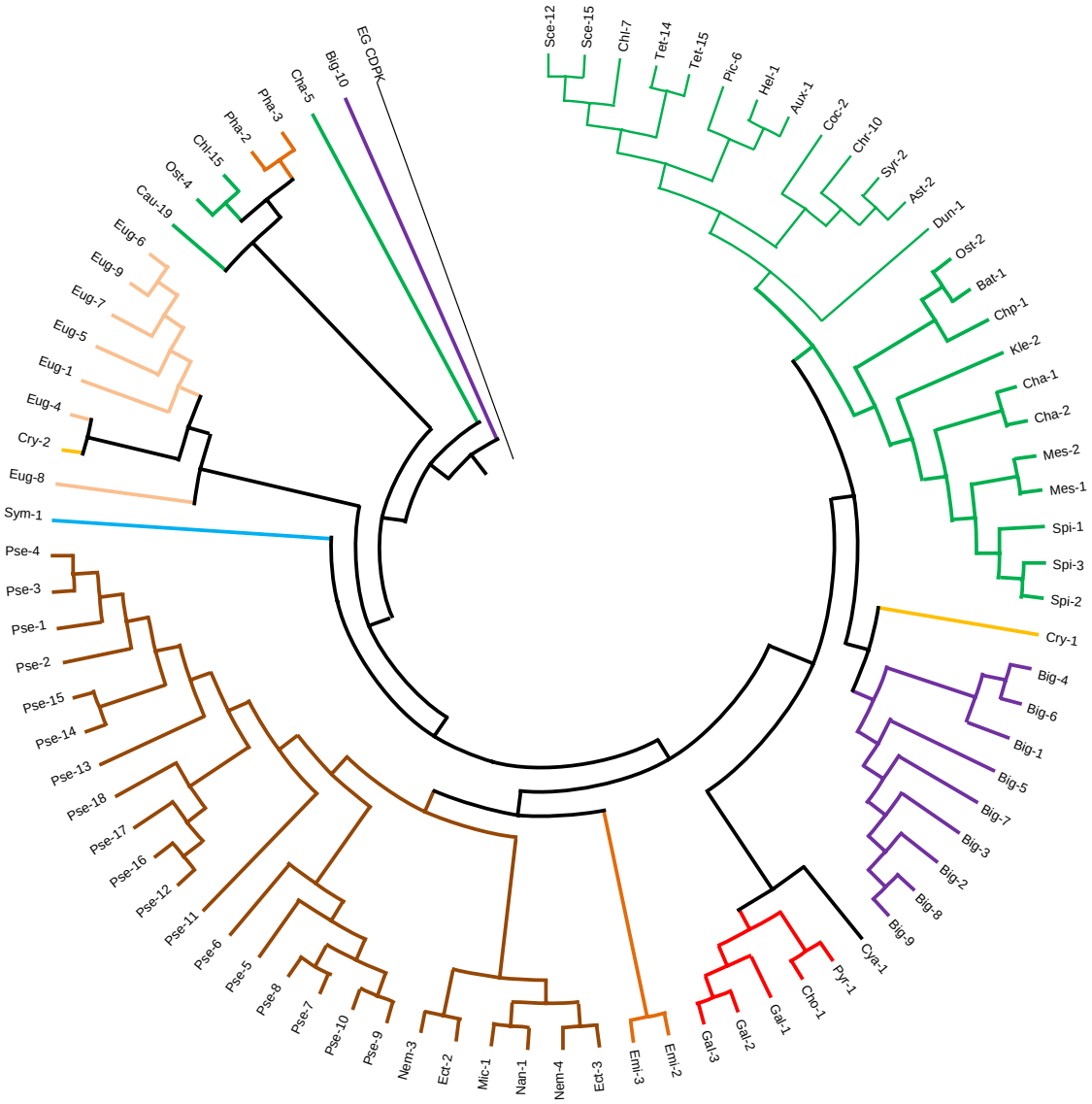


Figure 6: Maximum likelihood tree of just the identified SnRK1s.

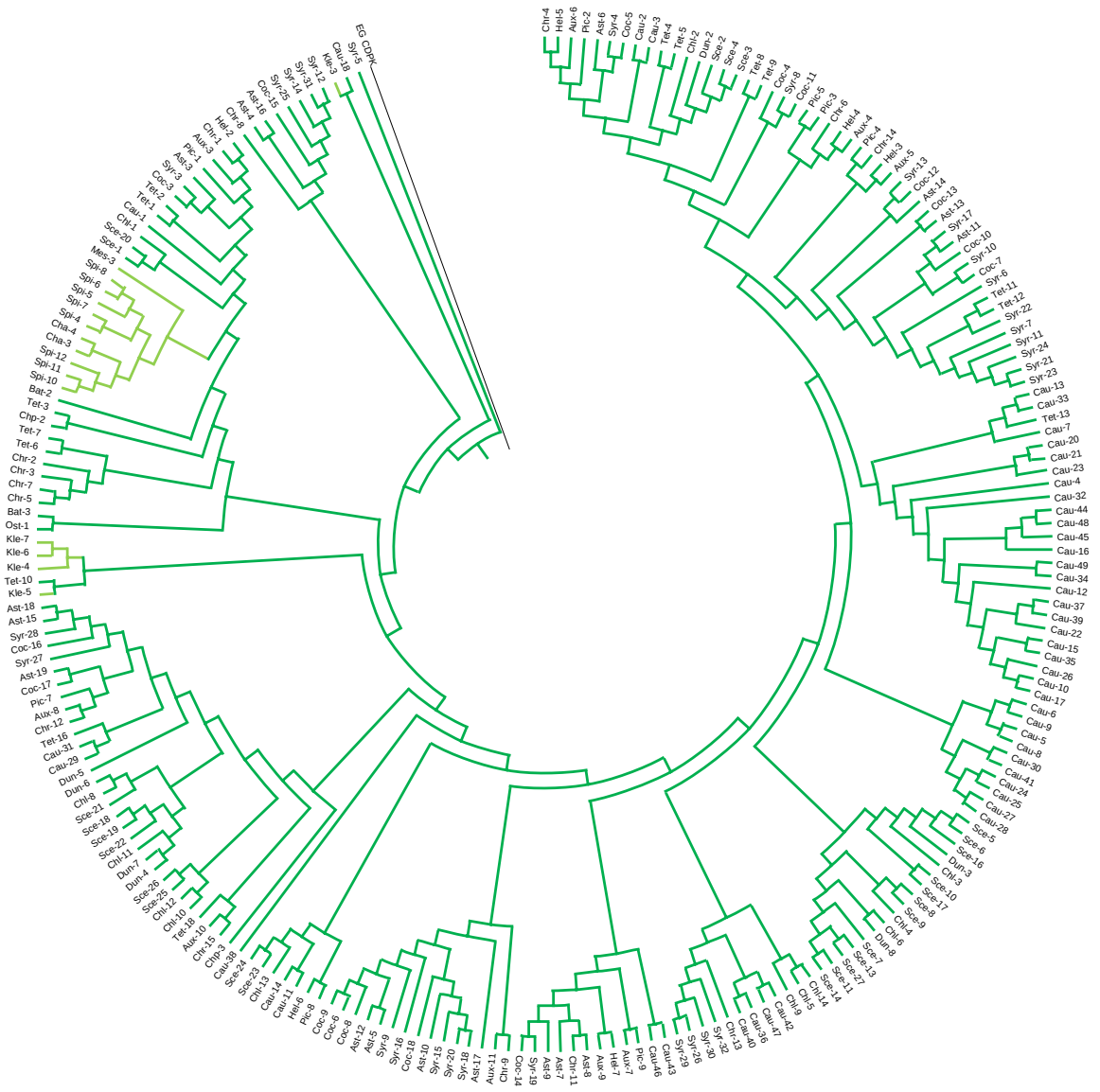


Figure 7: Maximum likelihood tree of just the identified SnRK2s.

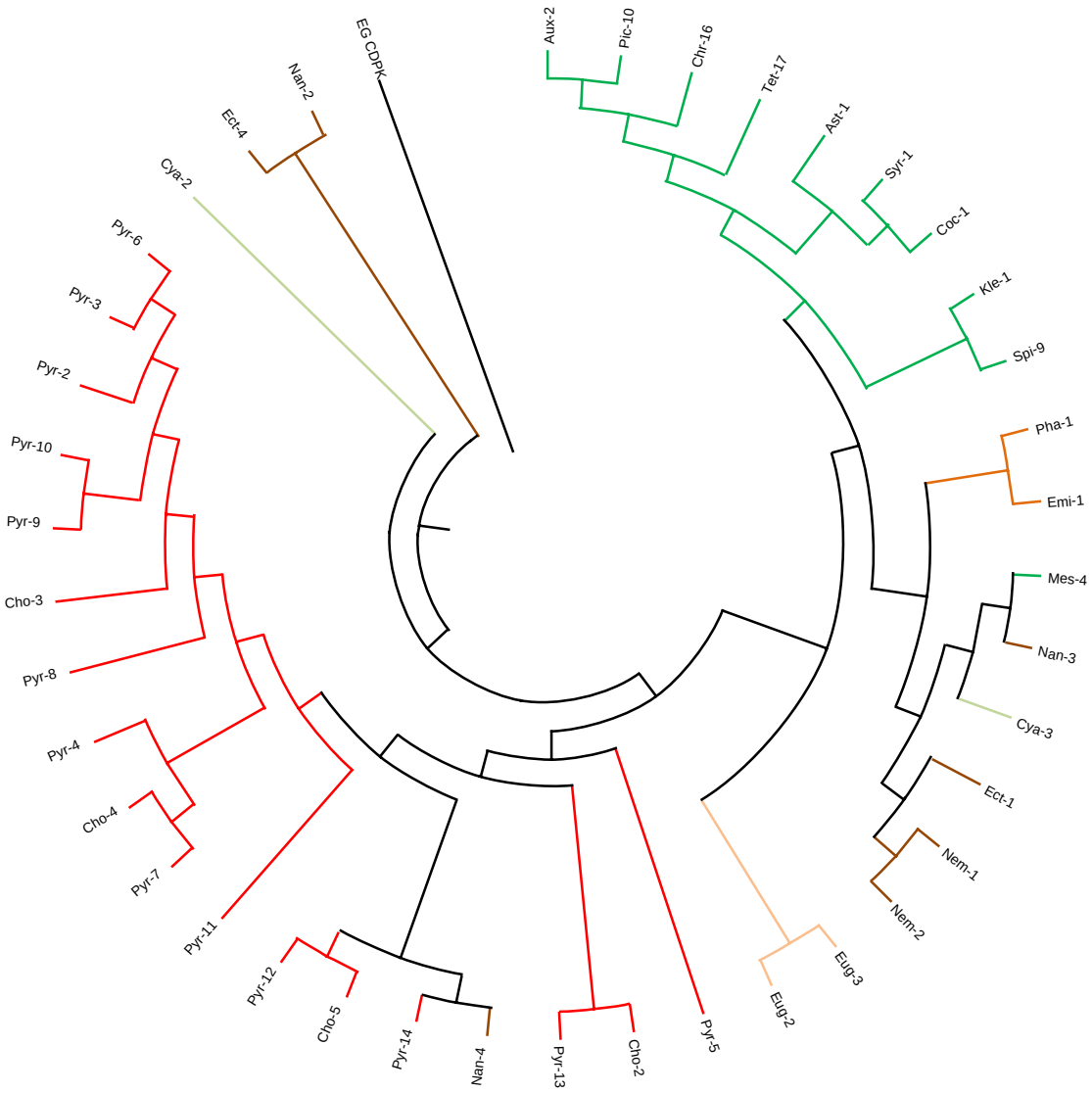


Figure 8: Maximum likelihood tree of just the identified SnRK3s.

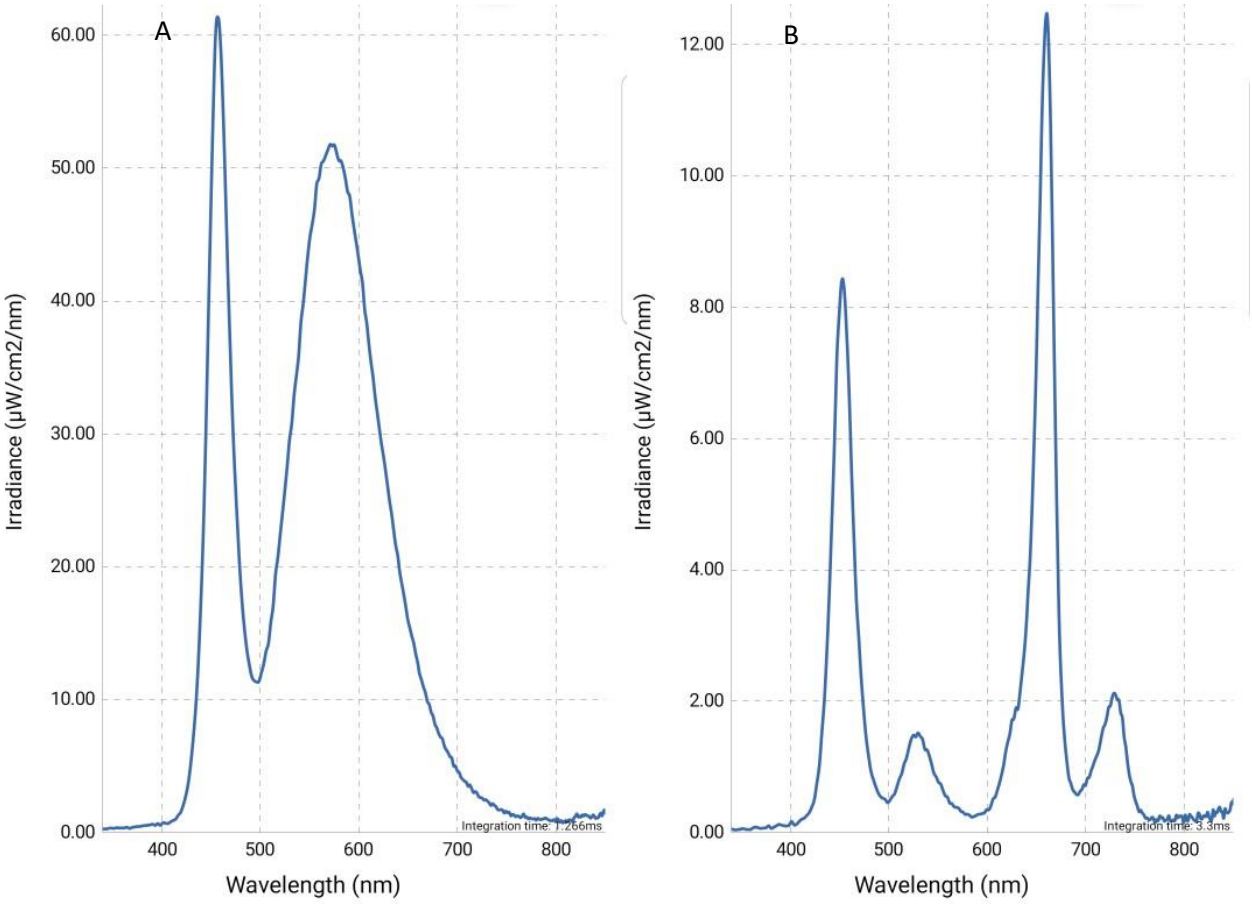


Figure 9: Spectrums of the lights used during the neochrome growth experiments. (A) Was the "white" light used during the initial growths. (B) Was the lights containing the wavelengths needed to excite neochrome (~735nm).

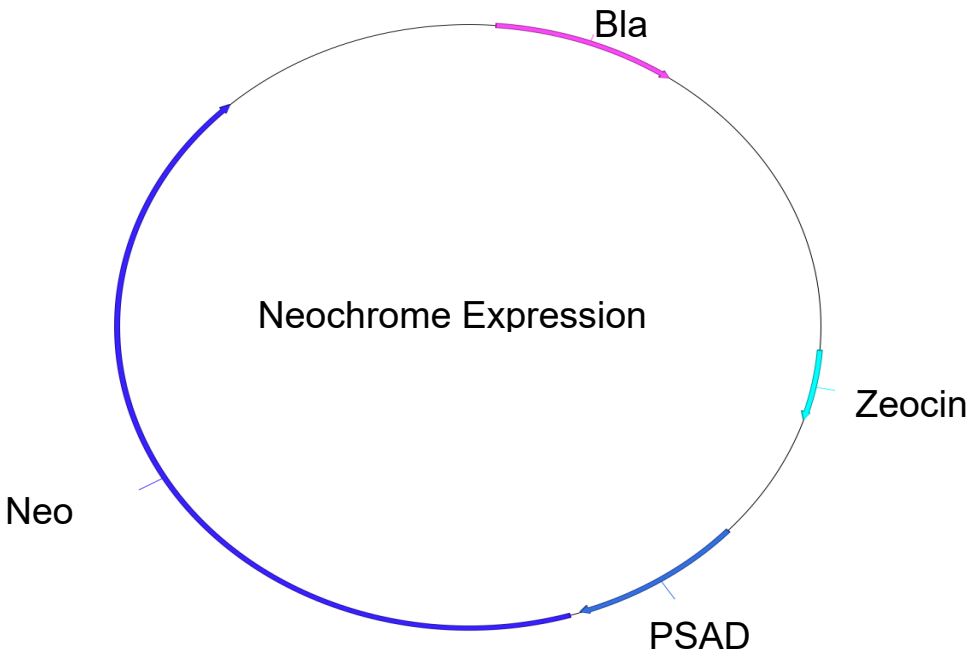


Figure 10: Neochrome expression plasmid.

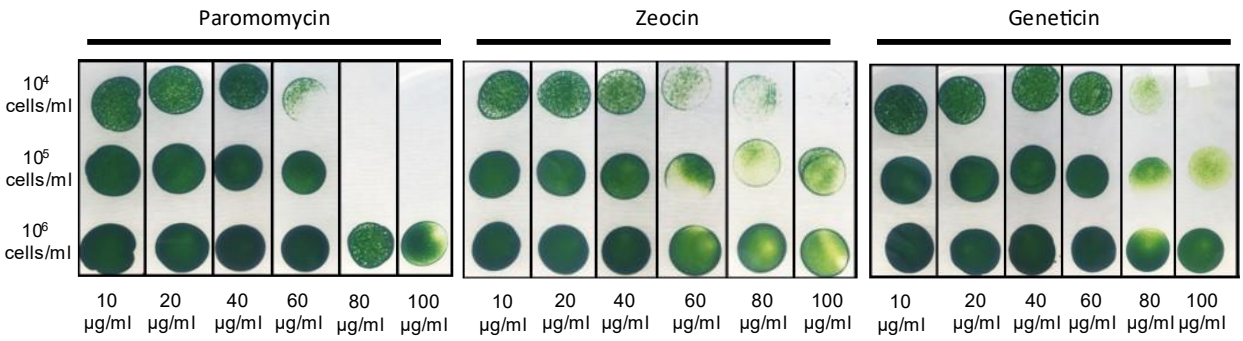


Figure 11: *Sorokiniana* cells were gown in HS media to approximately $\mathbf{10^7}$ cells/mL. Cell count measurements were conducted using a using a Hausser Scientific Bright-Line hemocytometer. Cells were diluted to $\mathbf{10^6}$,$\mathbf{10^5}$, and $\mathbf{10^4}$ cells/mL respectively. One milliliter of each of these dilutions was plated onto HS media plates containing antibiotic concentrations ranging from 10-100 $\boldsymbol{\mu g/mL}$. The antibiotics used for this kill curve were Paromomycin (A), Zeocin (B), and G418 (geneticin) (C).