Mapping the Path Forward to Next Generation Algal Technologies: Workshop on Understanding the Rules of Life and Complexity in Algal Systems

Authors: Junneng Wen [1], Kent Rapp[1]/ Lukas R. Dahlin[2]/ Chien-Ting Li[1]/ Jacob Sebesta[3], Amanda N. Barry[4], Michael T. Guarnieri[2], Christie Peebles[3], Michael Betenbaugh[1]

[1]Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD 21218. Corresponding author at beten@jhu.edu

[2]Biosciences Center, National Renewable Energy Laboratory, Golden, CO 80401

[3] Department of Chemical and Biological Engineering, Colorado State University, Fort Collins, CO 80523

[4]Department of Molecular and Microbiology, Sandia National Laboratories, Albuquerque, NM, 87185

Abstract

Algal technology offers enormous potential as a source of renewable biomass and bioproducts for fuels, biochemicals, nutrients, and numerous other applications. While tremendous opportunities exist, significant challenges are also present in harnessing and utilizing this valuable and potentially sustainable bioproduction platform. Funded by a joint partnership between the Bioenergy Technologies Office (BETO) at the Department of Energy (DOE) and the National Science Foundation (NSF), the recent virtual workshop, "Understanding the Rules of Life: Complexity in Algal Systems Workshop," was organized to bring together thought leaders from academia, national laboratories and other government agencies, and industry to identify key challenges and propose research strategies to overcome these barriers. Two parallel sessions were undertaken: 1) algae strain and toolkit development and 2) algal ecosystem and microbiome dynamics, in order to advance algae engineering and cultivation efforts. In each session, attendees discussed the most pressing challenges from the perspective of model systems and omics analysis. Themes such as genetic engineering toolboxes, translation from the laboratory to outdoor cultivation, functional genomics, metabolomics, algal ecology, consortial design, microbiome dynamics, mathematical models, and data analytics emerged from the discussion. Overall, this workshop has served to elucidate the key hurdles and many

opportunities present as we continue to advance algal science and technology from the laboratory and into the commercial realm.

1. Introduction

Over the past twenty years, the Department of Energy (DOE) Bioenergy Technologies Office (BETO) and the National Science Foundation (NSF) have invested in algal genome sequencing, multi-omics analysis, and microbiome interactions. In order to further improve our understanding of algal biology, biochemistry, and pond ecology, the recent algal workshop, entitled “Understanding the Rules of Life: Complexity in Algal Systems Workshop,” was held virtually in an effort to explore topics and identify challenges that are important for algal biology, algae cultivation at multiple scales, and translation into commercial processes. This workshop builds on previous efforts to identify research priorities that will enable algal biofuel commercialization, such as the “Algal Biology Toolbox Workshop” from 2016 in San Diego, California (U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Bioenergy Technologies Office, 2016).

This review summarizes the challenges and opportunities identified by workshop attendees and provides insights into a potential roadmap forward for the study and application of algal systems and communities in the coming decade. In regard to algal strain development, the attendees discussed several challenges from different perspectives including increased research on model systems, genetic engineering toolboxes for model and non-model organisms, and understanding the unique considerations when translating knowledge from the laboratory to outdoor cultivation settings. Attendees further discussed challenges in omics, including the need for additional systems level analyses, unified genomic datasets, and increased knowledge of functional genomics. For the ecosystem and microbiome area, attendees similarly discussed challenges in model systems and omics analysis. However, many of the challenges revolved around how to use omics analysis and computational models to understand the interactions across different microbial communities, as well as to characterize dynamics, and how to apply these methodologies at different scales from the laboratory to the open ponds systems.

Overall, this brainstorming effort has served to elucidate the many opportunities present as we continue to advance algal science and technology from the laboratory and into the commercial realm. Scientific and technical advances aimed at overcoming the challenges proposed here will enhance our understanding and knowledge of different algal strains and microbial communities and enable the community to better harness these systems for energy and a myriad of other applications.

2. Workshop Structure and Organization

To frame the purpose of the workshop, a virtual symposium (https://rulesoflifealgaesummersymposium.mystrikingly.com) was held in which 30 speakers discussed their current research in the areas of photosynthetic biomanufacturing, understanding and predicting phenotype, and microbiome theory and mechanisms (Appendix 1). This symposium informed attendees about the current breadth of knowledge within the NSF and BETO algae research portfolios. All presentations were pre-recorded and provided to workshop attendees prior to the September 25, 2020, workshop.

The workshop—during which challenges, opportunities, and research strategies were discussed—was one day and entirely virtual (see Table 1). Approximately 100 researchers distributed across academia, national laboratories, government, and industry (see estimated distribution of affiliations in Figure 1) participated. After a brief opening, the workshop was split into two different tracks, each of which hosted approximately half of the workshop attendees. One track emphasized algae strain development and molecular biology toolkits to gain fundamental knowledge about algal physiology for relevant algal species of both scientific and commercial interest. A second track focused on ecosystems and microbiome dynamics to better understand the complexity of algal microbiomes. For each track, the attendees discussed and identified challenges and barriers in a morning section. Then, they were encouraged to brainstorm R&D strategies to address these issues in an afternoon section (see Table 2). The goal was to compile the common challenges together with possible opportunities and solutions in order to map a path forward that integrates the needs of the algal industry with the capabilities of researchers and developers from academia, industry, and the government.

To organize the workshop participants' ideas, GroupMap™ was used. This software manages group brainstorming in a four-step process: brainstorm, group, validate, and vote. In the morning session, participants were first asked to brainstorm challenges they faced when applying algae technologies and systems to outdoor cultivation. Next, similar ideas were grouped together. Workshop attendees then validated that the grouped ideas were representative, captured sufficient detail, and—for the morning session—were appropriately positioned according to their likelihood of success and potential impact in the field. Each attendee then distributed votes across the proposed ideas according to importance. A similar process was followed for the afternoon session, but instead linked research strategies to the challenges identified in the morning session. Shown in Figure 2 is an example of the complete ideation process, spanning both the morning and afternoon breakout sessions.

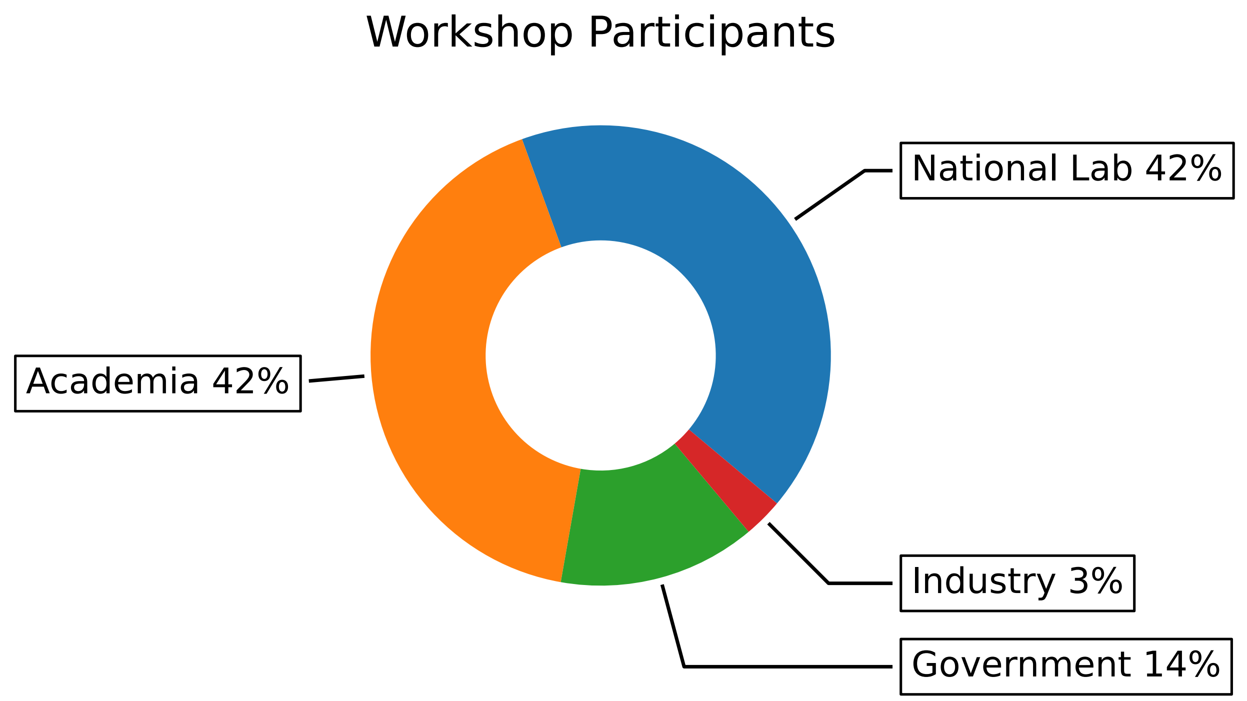


Figure 1. Donut chart showing the estimated distribution of affiliations for workshop attendees.

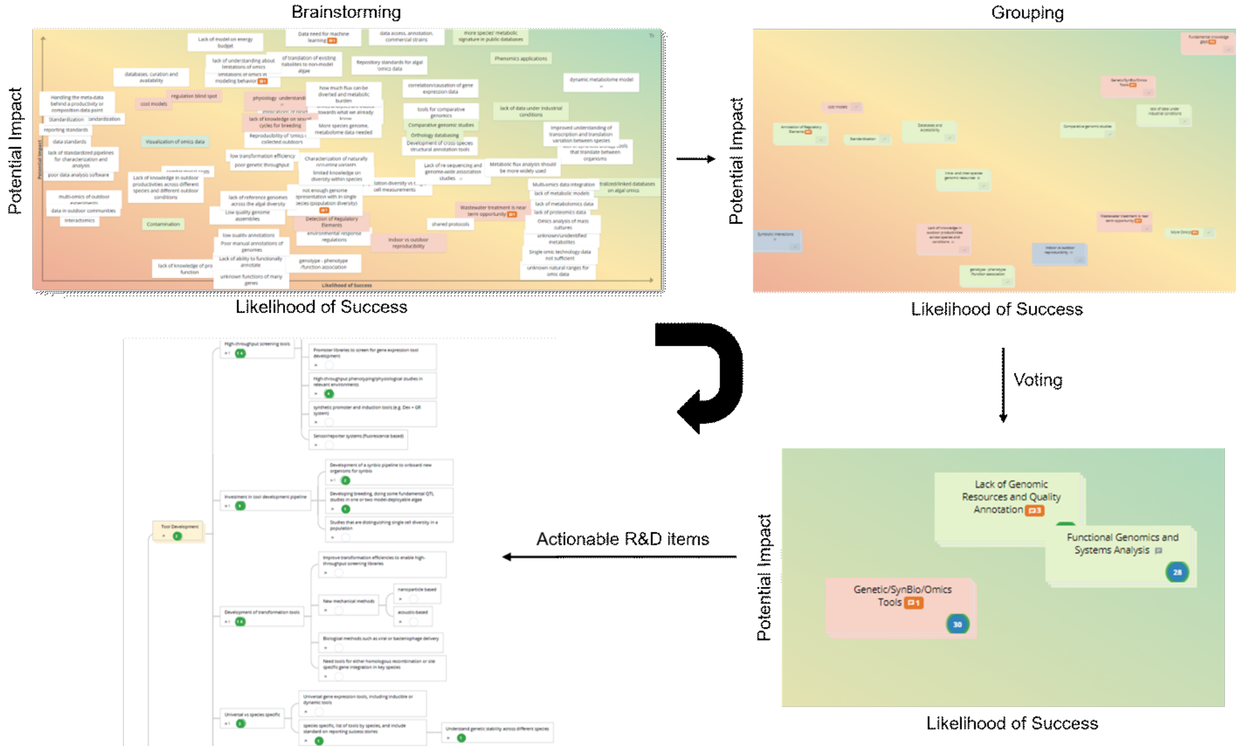


Figure 2. GroupMap output with an overview of the complete ideation process spanning the morning and afternoon sessions. The morning breakout session consisted of the first three steps: brainstorming, grouping, and voting. The afternoon breakout session developed actionable research and development ideas in response to the challenges identified in the morning breakout session.

Table 1: Agenda for the virtual workshop held on Friday, September 25, 2020.

Time (EDT)	Agenda Item
10:45 AM – 11:00 AM	Login and Networking
11:00 AM – 11:25 AM	Opening Remarks: Purpose and Virtual Symposia Synopsis
11:25 AM – 11:40 AM	Breakout Session Introduction
11:40 AM – 1:40 PM	Breakout Session 1: Challenges and Barriers *Track A: Ecosystem and Microbiome Dynamics* *Track B: Algae Strain Development and Toolkits*
1:40 PM – 2:05 PM	Break and Optional Networking
2:05 PM – 2:25 PM	Working Lunch: NSF & DOE Presentations
3:10 PM – 4:40 PM	Breakout Session 2: Strategies and Paths Forward *Track A: Ecosystem and Microbiome Dynamics* *Track B: Algae Strain Development and Toolkits*
4:40 PM – 5:00 PM	Breakout Session 2 Synopsis and Closing

Table 2: Topic areas identified in the workshop

Algae Strain Development and Toolkits		Ecosystem and microbiome dynamics	
Challenges and Opportunities in Model Systems	Challenges and Opportunities in omics and other challenges	Challenges and Opportunities in Model Systems	Challenges and Opportunities in omics and other challenges
Lack of Knowledge in Model Organisms	Functional Genomics and Systems Analysis	Understanding the role of metabolic interactions	Link omics to activities or phenotype

Genetic tool development	Genetic, synthetic biology, and omic tools	Contamination and colonization	High throughput/ AI and modeling infrastructure development
Indoor-Outdoor Translatability	Lack of Genomic Resources and Quality Annotation	Inadequacies in Observing the microbiome	Challenges in metabolomics
		Improve understanding of microbiome structural and functional dynamics	Translating laboratory scale interactions to large scale outcomes
		Engineering consortial design	

3. Algae Strain and Toolkit Development

3.1 Challenges and Opportunities in Model Systems

The Algal Strain and Toolkit Development breakout session focused on identifying the rules of life for algal strain and toolkit development. Among the challenges, areas identified as the highest impact and potential for success by workshop participants were as follows: Model system development, genetic tool development, and translatability of indoor systems to outdoor systems.

3.1.1 Model System Development

By virtue of having a more robust knowledgebase, several species of algae and cyanobacteria are recognized as model systems, including *Chlamydomonas,* and *Phaeodactylum* among eukaryotic algae and *Synechocystis* and *Synechococcus* among prokaryotic cyanobacteria, and workshop participants felt strongly that work must continue in these organisms. The extent of our knowledge about these species exceeds our knowledge of other algae and cyanobacteria, and this recognition encourages researchers to further investigate them. Although some model systems may have shortcomings for applied work, the general expectation is that more can be done in a model system in a timely manner, and that ultimately, these learnings can be translated to non-model systems that may be more appropriate for applied work, while also providing basic science insights into phototroph biology and evolution. An important point brought up by workshop participants is that phototrophic algal models lack the breadth of knowledge compared to heterotrophic models, such as *E. coli* and *S. cerevisiae*. There are likely significant lessons to be learned by looking at the history of scientific advances in these heterotrophic non-algae models in planning for the future of phototrophic models.

Among the more important factors discussed regarding model system development is the need for model systems that grow well outdoors and that can be used to better understand the unique problems faced when growing algae at an industrial scale (Figure 3). Indeed, recent work in the field has identified *Picochlorum spp.* as one such promising genus (Dahlin et al., 2019; Dahlin & Guarnieri, 2021; Krishnan et al., 2020; Weissman et al., 2018). Despite this progress, much work is needed before this genus has an equivalent knowledge base to canonical algal models, and additional models are necessary to account for the breadth of algal species across the tree of life. Additional models, could include hyper-saline and high lipid producing models, as significant work has gone into development of *Dunaliella* and *Nannochloropsis* in this regard (Ajjawi et al., 2017; Oren, 2005). Furthermore, models could be developed that target key lineages, as this would enable basic science research into algal evolution .

3.1.2 Genetic Tool Development

Genetic tools are needed for both basic and applied research for manipulating organisms either to probe function or to enhance it. The consensus of attendees was robust methods and a generalized workflow for transforming different species should be developed. This needs to be approached from the aspect of getting transgenic DNA into the cell using established technology and developing novel methodologies, and robust genetic elements for controlling gene expression. Once basic transformation protocols are established, a knowledge of functional genetic elements allowing for fine-tuned gene expression are necessary and currently lacking. For example, extensive promoter libraries, orthogonal gene expression components and ribosomal binding tunability available in *E.coli* and yeast systems, would greatly expand the possibilities for genetic modification if analogous systems were available in algal systems (Figure 3). Along the same lines, identification of DNA targeting approaches are necessary for tuning the expression and targeted knock-out of native genes. CRISPR technologies will likely be the most useful for eukaryotic algae in this regard (Krishnan et al., 2020). However, in cyanobacteria and eukaryotic chloroplasts, homologous recombination is highly prevalent and suitable for targeted integration. Special attention must be paid to the multiple copy numbers of genomic DNA, as integration must reach all genome copies. With the continued development of novel algal strains, it was also noted that development of universal tools and methodologies for genetic tools are needed. Breeding (sexual mating) strategies for algae should also be pursued, as to date, routine, controlled breeding has been limited to *Chlamydomonas.*

3.1.3 Indoor/outdoor Translatability

In general, there was a consensus that there must be development of small, laboratory scale systems that can effectively replicate outdoor environments (Figure 3). While considerable progress has been made in recent years by a variety of groups regarding this, many of these systems still fail to replicate carbon dioxide delivery, light intensity, light wavelengths, mixing regimes and particular media components used in outdoor cultivation. Scaled-down systems that can accurately predict effects at a larger scale are essential for controlled, high-throughput studies leading to rapid progress in the algal field. Thus, it is critical to invest in the development of these systems and push the field towards the adoption of such systems.

In order to inform the design of such small-scale systems, additional research needs to be placed on outdoor commercial systems. Learnings from this can then be used to iteratively inform small scale designs. Workshop participants noted the need for additional studies relating to basic science inquiries in outdoor commercial systems. Similarly, researchers have identified the need for additional outdoor testbed sites in geographically diverse locations, and improved access to current sites. Access of such testbed sites needs to be open to all researchers, and strains evaluated in such outdoor systems should be made publicly available to researchers for further evaluation and validation.

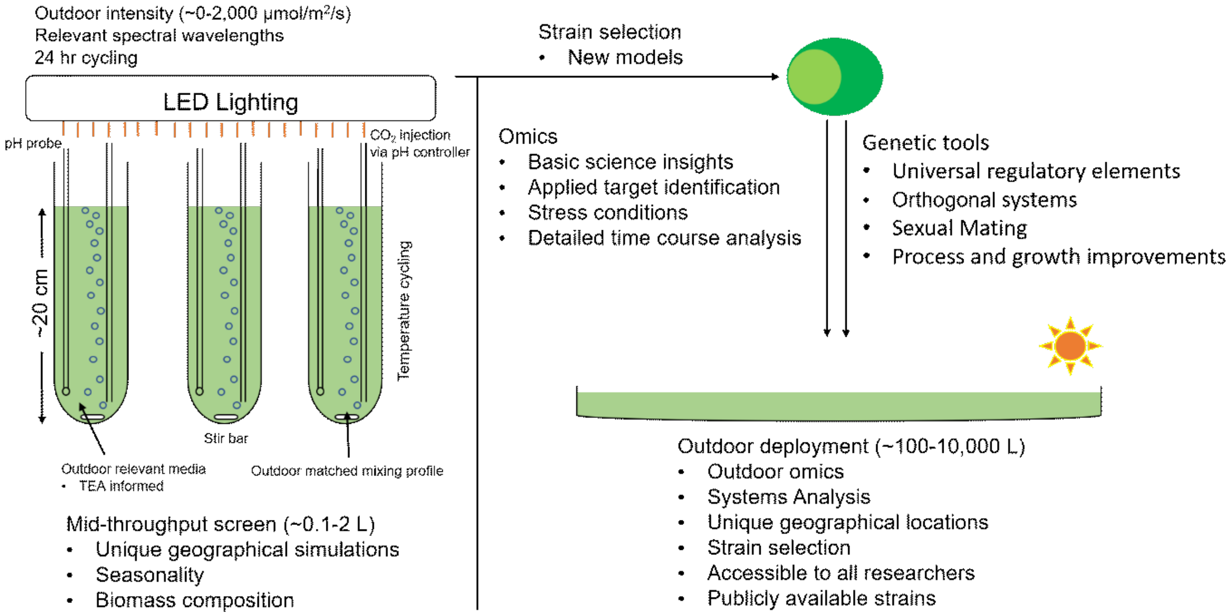


Figure 3. Integrated overview of various topics discussed in regard to algal strain and toolkit development. Bullet points highlight the areas where workshop participants remarked that further work and associated funding is needed.

3.2 Challenges and Opportunities in Omics Tools

This breakout session focused on identifying the role advanced -omics technologies should play in the continued development of algal systems. Among the challenges r, the areas identified as the highest impact and potential for success were: systems analysis, and genomics, with an emphasis on functional genomics.

3.2.1 Systems Analysis

The field of systems biology drives understanding of how the components of a system lead to the behavior of the system by integrating measurements of the components of the system ("omics") along with systems level modeling to capture the behavior (phenotype) of the system. This process eventually leads to models with the predictive power of phenotypes in response to perturbations. To deal with dynamic complexities, scientists have often simplified the system by performing omics experiments under well controlled conditions such as constant light, low light intensities, low cell densities, constant temperature, and well-mixed environments. Step changes in experimental conditions allows for some limited biological insight on the link between genotype and phenotype in these laboratory conditions. These conditions are limiting in replication of outdoor conditions reducing the relevance of the data and negatively impacting our translation of strains developed indoors to outdoor conditions. Thus, a major challenge identified was accurately replicating these dynamics, and performing systems level analyses in outdoor systems. Methodologies that allow for time course collection of multiple systems biology data sets on the same sample including transcriptomics, proteomics, metabolomics, and epigenetics are needed. Concurrently, comparative omics studies among variants of the same species can lead to unique insights, such as those found between *Synechococcus elongatus* PCC 7942 and UTEX 2973, wherein 3 mutations were found to account for vastly different growth rates (Ungerer et al., 2018; Yu et al., 2015). Further, all data collected in these efforts through government funding should be made available in digital repositories.

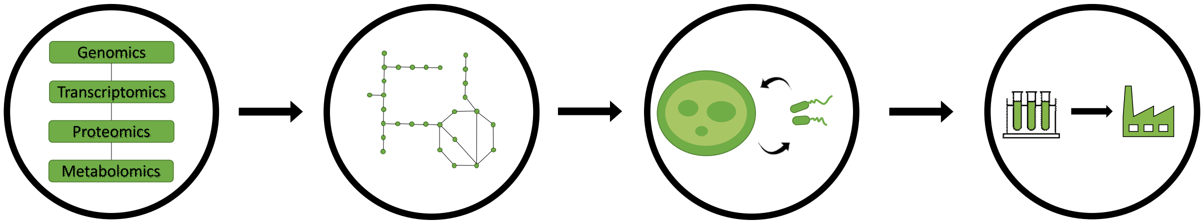
3.2.2 Mathematical Models

Due to the complexity of systems biology data sets, participants identified a considerable need for investment into computational tools that can accurately capture the time dependent variations inherent in photoautotrophic organisms where the pseudo-steady state hypothesis can break down. Participants agreed that these metabolic models are extremely important to identify breeding and engineering targets in order to improve productivity and concentrations of biobased fuels and chemicals generated by algae and cyanobacteria. These models would allow researchers to ask questions about the best way to divert carbon flux to the product of interest while allowing the cell to thrive in a dynamic environment or how to minimize the metabolic burden under changing conditions. With the increasing number of systems biology data sets, participants discussed the need for advances in the development of regulatory networks that allow researchers to understand the link between genotype, environmental cues, and the displayed phenotype of the organism. A more in-depth understanding of regulatory networks leads to a better understanding of the decision-making process cells go through in order to survive in their environment. Participants identified machine learning as one approach to address this. It was also noted that community models that account for population heterogeneity are needed.

3.2.3 Genomics and Functional Genomics

Another important area identified by participants is the area of algal genomics and functional genomics. Regarding algal genomics, many workshop participants discussed a need to increase our breadth and depth of sequence knowledge. More representative genomes are needed from across the algal tree of life to better understand the biological diversity in nature (Hanschen & Starkenburg, 2020). Closely linked to expanding our knowledge of new species is the need for improved genome annotation methods that are specific for algae/cyanobacteria and comprehensive functional genomics and related studies to build our understanding of protein function and the link between genotype and phenotype. Recent advances in high-throughput cloning methods, transposon libraries, gene-knockout libraries, etc. open avenues to explore high-throughput functional screening to determine protein functionality. Closely tied to functional genomics is the need to improve annotation of metabolite signatures in metabolomics data so that changes in metabolite data can be used to narrow down the function of hypothetical proteins.

To address some of these limitations, researchers should build on existing databases and unify those like cyanobase (genome.microbed.jp/cyanobase) and algaebase (https://www.algaebase.org). A lack of data quality and formatting standards makes comparisons difficult, though an attempt at setting standards has been made in the development of PhycoCosm (Grigoriev et al., 2021). As more genomes are published, associated annotation of regulatory elements (e.g. promoters, transcription factor binding sites, terminators, etc.), will greatly help researchers involved in the development and application of genetic tools. When publishing genomes, workshop participants suggested having set standards that would ensure genome quality and availability of annotations are met.



Omics Analysis
- Establish phenotype prediction standard
- Annotate pathways
- Develop high-throughput data collection techniques
- Decipher evolutionary adaptations
- Integrate metabolite data

Metabolic Network Modeling
- Capture spatiotemporal dynamics
- Incorporate external contamination
- Combine modeling with flux balance analysis

Microbiome Algae Interactions
- Model mechanisms behind algal microbiome ecosystem
- Develop engineered consortium
- Develop synthetic ecological pipelines

Translation to Large-scale System
- Model large-scale systems linking lab-pond disconnect
- Study at multiple scales
- Foster partnerships

Figure 4. Integrated overview of various topics discussed regarding ecosystem and microbiome dynamics. Bullet points highlight the areas where further work, and associated funding is needed.

4. Ecosystem and Microbiome Dynamics

In nature, organisms interact with other species to build microbial communities. While their potential in stabilizing and facilitating algal bioprocessing is enormous, key challenges exist due to their complex interacting structures and dynamics that hamper our ability to effectively leverage algal microbiomes for biofuels and other commercial applications. Workshop participants emphasized that understanding the complexity and dynamics of algal microbial communities is essential to answer many basic science and applied research questions, including how to prevent or minimize harmful contaminations and how to control these cocultures to enhance feasibility of algal biofuels and biochemicals production.

4.1 Challenges and Opportunities in Model Systems

4.1.1 Understanding the Role of Members and Interactions

A major hurdle in modeling algal systems identified by the workshop participants is the lack of fundamental ecological understanding of the algal microbiome. Researchers have achieved moderate successes in cultivations, but we lack a deeper understanding of the mechanisms driving these ecosystems. Mathematical models are being developed based on empirical functions but greater mechanistic insight is needed (Widder et al., 2016). While the metabolic interactions can sometimes be experimentally characterized using Nuclear Magnetic Resonance (NMR) and Gas Chromatography-Mass Spectrometry (GC-MS), an appreciation of the complex ecology behind these interactions is still sorely lacking (Zuñiga et al., 2020). Indeed, dynamic interactions between the microbiome participants, invasive microbes, and environment are at the heart of ecological communities. Lacking this fundamental understanding on the ecological role of microbiome makes it difficult to efficiently direct microbiome engineering efforts. While mathematical models are being developed based on empirical functions, greater mechanistic insight is needed to make these more predictive of algal microbiome behaviors (Widder et al., 2016).

Participants noted that the first step in understanding the role of different interactions is to identify the players, or the members of the microbiome including the algal host, partners, and relevant environmental factors. After compiling a well characterized list of the members involved, determining the function and fitness of the members in the group is critical. Computational models can be helpful to describe the behaviors and interactions of the key players in the algal microbiome system. While there have been limited efforts to model the molecular interactions in systems such as *C. vulgaris* UTEX 395 and *E. coli* to predict the growth and to track biomass concentration (Chan et al., 2017; Zuñiga et al., 2016), this area remains an opportunity for significant expansion. Furthermore, these modeling studies will guide data collection and experimental design in order to provide further insights around our understanding of microbiome, its behavior, and the environment.

One of the drivers identified by the participants for testing and improving our knowledge and understanding of the microbiome and enhancing our models is through the use of complementary analytical tools. For example, metabolomics information can help us to understand how the environment including probiotics can facilitate microbiome function and enable us to predict how probiotics will persist in the microbiome. Ultimately, a better understanding the interactions between the members of the microbial communities and their environment prepares users for engineering a superior microbiome with improved organismal interventions that enhance function of the overall system.

4.1.2 Contamination and Colonization

To have a valid model system, a major challenge that needs to be addressed is external contamination by pests. Representative algal systems, particularly in outdoor ponds, can face contamination and colonization by natural algae or bacteria. Foreign invaders can influence and crash the culture; in an outdoor environment, invaders such as grazers, fungi, or bacteria can alter algal productivity (Hoffman et al., 2008; Matthiessen et al., 2007) Not surprisingly, there is a lack of relevant model system that include pests as a means to describe the influence of these invaders. Conversely, there is a need to identify organisms that can protect algal systems from pests and in particular in relevant outdoor settings. Indeed, model organisms that are promising in the lab may prove to be less effective in outdoor cultivations. Understanding the mechanisms that underlie the protection afforded by these members is equally important; otherwise, researchers face significant challenges in implementing engineering interventions due to an opaque relationship between lab-scale microbial behaviors and outdoor behavior. An overarching strategy must be undertaken to build engineered consortia and synthetic ecological pipelines using genetic and eco-engineering tools together with mathematical models.

4.1.3 Inadequacies in Observing the Microbiome

Another challenge noted by workshop participants in model systems is the difficulty in observing or monitoring the microbiome. We now appreciate that the composition of the microbiome is variable and highly dynamic (Brenner et al., 2008). Currently there is a dearth of tools for tracking the spatiotemporal variability in partners as the microbiome changes and evolves over time. Knowledge of the rapid-changing spatiotemporal patterns can provide relevant insights into the ecological states of evolving microbiomes. As the complexity of the system expands, researchers are required to follow more contributing factors, and this increases the challenge of the undertaking. The algal research community must decide the level of complexity for model microbial ecosystems that is appropriate to sufficiently describe a real-world system while also maintaining stability and robustness. For these model systems, methodologies for automated, reliable and quantitative sample and data acquisition are needed. In particular, metabolomics data may be particularly useful and then combined with metabolic flux analysis in order to capture the complex interactions of consortium members (Sen & Orešič, 2019).

4.1.4 Improve Understanding of Microbiome Structural and Functional Dynamics

Furthermore, microbiomes are highly dynamic which lead to species diversity that also affects the overall state, health, and function of these ecosystems. Because of the evolving composition and structural dynamics that can occur in microbiomes over time, documenting or even identifying temporal microbiome profiles represents a significant challenge. These ever-changing population-level dynamics can dictate if a strain will be transient or if it will persist in the host and ultimately affect microbiome function. The influence of the environment represents another layer that can alter spatiotemporal dynamics of the microbiome composition and accompanying functionalities. Environmental factors such as duration of daylight and specific season among other factors can influence microbiome composition. As a result, there is an expanding need for analytical tools and datasets including omics profiles that can capture and describe the spatiotemporal dynamics of the microbiome composition and function given the variability in the environment.

4.1.5 Engineering Consortial Design

The design, development, and implementation of engineered consortia is an important step in making algae microbiomes an integral part of the algal bioprocessing landscape. This process of microbiome engineering involves the deliberate and/or random introduction of members as participants to inculcate the microbial consortia with desirable traits such as growth enhancement, pathogen resistance, or lipid production to name a few. Fortunately, there are different means available for consortia design ranging from trial-and-error approaches to rational design methods. For the first approach, consortia have been built by varying species combinations (Shurin et al., 2013). In order to increase the chances of success, the consortia design can be coupled together with mathematical models built based on metabolite exchange or other interactions (Ben Said & Or, 2017). Furthermore, the design and engineering of the microbial consortia may be guided in the future by incorporation of mathematical models built on metabolite exchange and other interactions in combination with flux balance analysis.

Another challenge for engineered consortia is long-term evolution in the lab. In some cases, the composition of the microbiome will not be stable and thus is not maintained over multiple transfers. Technologies to preserve the identity or ensure the longevity and reproducibility of the desired microbiome and its activity will be invaluable. By constructing an arsenal of synthetic ecology tools, researchers will be able to design radically improved and more effective engineered consortia with critical functionalities that will ensure long term sustained algal cultures.

4.2 Challenges and Opportunities in Omics Tools

4.2.1 Link Omics to Activities or Phenotype

Applying omics tools provides another major opportunity to enhance our understanding and ultimately the performance of algal microbiomes. Genomics, transcriptomics, proteomics, and metabolomics can each provide insights into distinct aspects regarding the characteristics and function of microbiomes. Combined, these and other omics fields of inquiry will yield a wealth of information that will build the foundation for generating a more detailed and nuanced understanding of algal microbiomes.

A thorough and quantitative deciphering of our microbiome omics data will be challenging and require sophisticate meta-omics tools. This is difficult for many reasons including a lack of knowledge concerning gene distribution across the microbiome. First, the gene relevance of some microbiome genomic and other omics datasets is unknown, as there will be insufficient mapping between the omics data and the functions of the many genes identified in meta-omics profiling. Further, the relationship between genomics, transcriptomics, proteomics, and metabolomics is not fully established. If the interconnection between omics profiles and specific metabolomic signatures can be identified, the role of specific gene products on the overall phenotype may be deduced even for complex algal microbiomes.

Accurate reflections of the active living and functional members is also needed as the presence of dead members represent another obstacle impacting the validity of omics data sets. Indeed, another goal in describing microbial communities will be to capture the spatiotemporal changes of the population members with assistance of mathematical models. Metabolic networks are inferred from the genomes by assembling enzyme networks and metabolites related to sequence and function (Covert, 2006), which requires proper gene annotation and determination of relative abundances of genes in the population. Even if the gene annotations are complete, another issue is to define the size and complexity of metabolic networks to be included, and the granularity selected can ultimately dictate a model's value. The inclusion of gene regulation activities may also be useful to accurately reflect actual system operation.

Moreover, deciphering evolutionary adaptations with omics tools represents another opportunity for the field. Algae and bacteria can co-evolve to change the ecosystem and this microbiome aspect, which can affect their physiology, metabolism and fitness, needs to be evaluated (Cullen et al., 2020; Ramanan et al., 2016). The co-evolutionary process may lead to beneficial traits for the algae industry, such as enhanced algal growth and increased biomass productivity, although applying omics tools to capture the evolutionary adaptations will be daunting especially for subtle phenotypic changes (Kim et al., 2014; Lee et al., 2013). Stable isotope tracking represents a useful complementary technology to enhance our understanding of the microbiome phenotype and partner interactions.

4.2.2 High Throughput/ AI and Modeling Infrastructure Development

Collecting and analyzing omics datasets to accurately describe algal systems is challenging and labor-intensive. While there is already infrastructure development such as Ensemble Genomes to organize genomics data, all relevant species and interactions in the algal microbiome may not be sufficiently captured through this interface (Howe et al., 2020). A fundamental challenge—extracting information about the relevant species in algal microbiota—underpins the need for high-throughput data collection and analysis techniques. Automated construction methods can be used to leverage omics information in generating mathematical models and the Systems Biology Markup Language (SBML) can integrate flux balance constraints to help standardize systems analysis for describing algal bioprocesses (Faria et al., 2018; Olivier & Bergmann, 2018). Likewise the algal community should continue to adopt artificial intelligence techniques that have already proven useful in areas such as species identification and system design (Teng et al., 2020).

4.2.3 Challenges in Metabolomics

Metabolomics, as one of the pillars of the omics technology platform, can lead to useful insights into the activities and phenotype of a single microbe, yet interpreting such data sets from algal consortia continues to challenge the field (Chong & Xia, 2017). Methods to identify intracellular, secreted, and shared metabolites in order to elucidate their roles are urgently needed. The limited extracellular accumulation of rapidly exchanged metabolites can further exacerbate challenges in detection. Data integration of metabolomics is further challenged by the need to identify the origin of the secreted metabolites and their relevance to the activities and functions of the community.

An approach to begin to address these challenges is to expand the genomic and biosynthetic pathway annotation of algal microbiome species to create a clearer picture of exported and imported metabolites. Metabolomics data is often used to inform and fine-tune computational models, an area in which our understanding is still developing when it comes to algal microbial communities. Moreover, initial stand-alone algal metabolic models can be embedded within complex multispecies microbiome metabolic models containing the key contributing members within the community. Building complex microbiome models and integrating multispecies metabolite data represents a unique opportunity to strengthen our understanding of the interactions present in communities as we seek to increase our knowledge about the functions of individual members and enhance overall effectiveness and utility of algal microbiomes.

4.2.4 Translating Laboratory Scale Interactions to Large Scale and Outdoor Outcomes

Research performed with microbiomes in the laboratory often yields results that can be different from their application at larger scales, especially in outdoor environments. Outdoor cultivations introduce complexities as well as extraneous and even unknown parameters absent in well-controlled laboratory settings. One approach to tackle this challenge is to study microbiome cultivations at multiple scales and then build a model that bridges the lab-pond disconnect. Taking the data from multiple sizes can begin to identify the factors influencing growth, health, members, and dynamics of the microbiome and then use the these finding to characterize the differences between small scale and large-scale cultivations.

Unfortunately, it is more difficult to collect data about interactions in large-scale, and particularly outdoor cultures compared to experiments done in the laboratory. The interpretation of experiments and integration of modeling and experiments in large-scale cultures can also be quite demanding since it is more difficult to control the variables that affect system performance at large scales and in outdoor ponds. While a limited number of variables may be successfully chosen, finely tuned, and manipulated within the laboratory scale, a wide assortment of variables, some of which may be difficult to manage, will exist for larger scaled outdoor cultures. In large-scale studies, two types of research are needed: 1) descriptive studies to help explain the behaviors of the large-scale system and 2) perturbation studies that turn findings from descriptive studies into actionable items to achieve a desired outcome. Data from large scale cultures can identify key variables critical to predicting behavior of microbiomes under various conditions. Large scale studies are costly and time-consuming, which makes partnerships and information sharing between researchers in the laboratory and demonstration scale critical to expanding our knowledge base, enhancing resource utilization, and ultimately producing effective algal microbiomes at all scales.

5. Concluding Remarks

Advances in algal biology and ecology have been considerable in recent years. However, as researchers learn more about algal systems and seek to expand their use and scale up production, many new challenges have emerged.

In order for algal cultivations to continue to expand and contribute to a sustainable bioeconomy, a number of scientific, engineering, and commercial hurdles must be overcome. Primary among these is the need to develop knowledge-bases and genetic tools for additional promising strains of algae in collaboration with academia and industry such that these strains may emerge as useful model systems. Emergence of an expanded portfolio of model systems across genera will ensure the capacity to pursue both basic science questions about algal growth, metabolism, function, and partnerships as well as applications-based inquiries such as the translation of learnings from controlled laboratory environments into cultivation in stochastic outdoor growth systems. In parallel, traditional model systems need to be sustained with continued investment as the current algal model systems are still lacking the breadth of knowledge present in heterotrophic non-algal systems. Continued study with traditional algal species systems is essential as these models often serve as the foundation and testing ground for demonstrating new advances in algal technology, guiding the development of novel strains and in particular those of critical industrial relevance.

To accomplish these goals the algal community must also embrace continued advances in genomic sciences, particularly regarding high quality annotation of genomes, functional genomics, and systems-based analysis. Understanding the relationship between genotype and phenotype in algal systems will be important to further develop applied systems.

In the area of algal ecosystem and microbiomes, there is a pressing need to connect genomics, transcriptomics, proteomics, and metabolomics through a global meta-omics approach in order to identify the functional roles and phenotypes of members that contribute to the overall community. Omics tools also present a potential solution to decipher spatiotemporal evolution and dynamic changes with the algal microbiome. In concert, high-throughput data collection techniques and testing together with a more widespread adoption of artificial intelligence and machine learning techniques offer significant potential opportunities in elucidating functional interactions and consortia system designs. The construction of genome-scale metabolic network models for algal microbiomes utilizing the omics datasets together with isotope labeling will help to delineate species interactions and accelerate algal biotechnology development. Integrating metabolite data into community models will be a useful step to fully utilize the power of metabolomics in order to more fully understand algal consortia behaviors. Also, progress in engineered consortium design that incorporates helper partners as well as invaders and external contamination present in these complex systems is essential for more effective algal communities especially in outdoor settings. Critically, studies addressing the differences present and the translatability when moving algal-based ecosystems from the controlled lab environment to large scale outdoor ponds must occur in order to ensure technologies utilizing algae and its partners become a commercial reality. Altogether, the workshop participants mapped out a path forward to next generation algal technologies, the development of which will help to overcome current challenges in algal science and make algal biotechnology a key part of our sustainable future.

6. Acknowledgement

We would like to thank all workshop participants for their discussions and valuable inputs. We would also like to thank all speakers for the symposium talks.

This work was supported, in part, under the LANL Technical Integration, Analysis, and Verification Support project awarded to A.N.B. by the Department of Energy's Bioenergy Technologies Office, Office of Energy Efficiency and Renewable Energy, under agreement numbers NL0035866 (WBS 1.3.5.011) and by Sandia National Laboratories. *Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. The workshop was also supported by US National Science Foundation grant 2013902*

Appendix 1: Titles of Symposium Talks

Titles of Symposium Talks

Area	Title	Presenter
Photosynthetic Biomanufacturing	Applying metabolic models for control in order to enhance algal growth and lipid production	Michael Betenbaugh
	SNRK protein kinases: Key regulators of microalgal photosynthesis and productivity	Sangeeta Negi
	Progress on controlling gene expression in diatoms	Graham Peers
	Challenges in Developing a Syn-Bio Toolkit for Microchloropsis gaditana	Anne Ruffing
	High-Throughput Selection and Directed Evolution of Microalgae for Improved Biomass and Bioproduct Yields	Matthew Posewitz
	Photomixotrophic Chemical Production in Cyanobacteria	Shota Atsumi
	SOFAST: Streamlined Optimization of Filamentous Arthrospira/Spirulina Traits	Damian Carrieri
	A new method for marker-less genome engineering in cyanobacteria	David Nielsen
	The Dynamics of Cyanobacteria Metabolism	Christie Peebles
	How to tell cyanobacteria to make biofuels and green chemicals	Wim Vermaas

Area	Title	Presenter
Understanding and Predicting Phenotype	Using synthetic biology and breeding to increase product accumulation in green algae	Steve Mayfield
	Sex in algae: a missing ingredient for domestication and strain improvement	James Umen
	Working towards Algae Phenomics: The green alga Scenedesmus obliquus	Juergen Polle
	The diversity of functional toolkits in algae	Erik R. Hanschen
	DISCOVR: Development of Integrated Screening, Cultivar Optimization, and Verification Research	Michael Huesemann
	The state of algal genome quality and diversity	Shawn Starkenburg
	Systematic genome-wide characterization of mutant phenotypes in a model alga	Martin Jonikas
	Understanding Algal Heat Responses by Combining Phenomics and Omics	Ru Zhang
	Building Algal Multi-omics	Igor Grigoriev

Area	Title	Presenter
Microbiome Theory and Mechanisms	Understanding the Phycosphere	Tisza Bell
	Starting with the fundamentals: microbiome engineering must rely on ecology	Chris Ward
	Pesticides and Algal Cultivation: Regulatory Considerations	Taylor Weiss
	Design and Testing of Algal and Algal-Bacterial Consortia for Outdoor Cultivation Systems	Alina Corcoran
	Ecological Engineering of Algal Communities for Maximum Productivity	Robin Gerlach
	Bacteria-algae interactions under the hot sun	Xavier Mayali
	Microalgae-methanotroph coculture: a synthetic microbiome to help understand the rules of life	Jin Wang
	Unraveling Community Interactions	Karsten Zengler
	Integrating microbial ‘omics, physiology, and environmental data with modeling	Greg Dick
	Algae Microbiomes: Natural and Engineered Systems	Zachary Johnson
	Impact of Microbial Community on Post-Harvest Algae Biomass	Brad Wahlen

References

Ajjawi, I., Verruto, J., Aqui, M., Soriaga, L. B., Coppersmith, J., Kwok, K., Peach, L., Orchard, E., Kalb, R., Xu, W., Carlson, T. J., Francis, K., Konigsfeld, K., Bartalis, J., Schultz, A., Lambert, W., Schwartz, A. S., Brown, R., & Moellering, E. R. (2017). Lipid production in Nannochloropsis gaditana is doubled by decreasing expression of a single transcriptional regulator. *Nature Biotechnology*, *35*(7), 647–652. https://doi.org/10.1038/nbt.3865

Ben Said, S., & Or, D. (2017). Synthetic Microbial Ecology: Engineering Habitats for Modular Consortia. *Frontiers in Microbiology*, *8*, 1125. https://doi.org/10.3389/fmicb.2017.01125

Brenner, K., You, L., & Arnold, F. H. (2008). Engineering microbial consortia: a new frontier in synthetic biology. *Trends in Biotechnology*, *26*(9), 483–489. https://doi.org/10.1016/j.tibtech.2008.05.004

Chan, S. H. J., Simons, M. N., & Maranas, C. D. (2017). SteadyCom: Predicting microbial abundances while ensuring community stability. *PLoS Computational Biology*, *13*(5), e1005539–e1005539. https://doi.org/10.1371/journal.pcbi.1005539

Chong, J., & Xia, J. (2017). Computational Approaches for Integrative Analysis of the Metabolome and Microbiome. In *Metabolites* (Vol. 7, Issue 4). https://doi.org/10.3390/metabo7040062

Covert, M. W. (2006). *Chapter 10 - Integrated Regulatory and Metabolic Models* (A. Kriete & R. B. T.-C. S. B. Eils (eds.); pp. 191–204). Academic Press. https://doi.org/https://doi.org/10.1016/B978-012088786-6/50029-0

Cullen, C. M., Aneja, K. K., Beyhan, S., Cho, C. E., Woloszynek, S., Convertino, M., McCoy, S. J., Zhang, Y., Anderson, M. Z., Alvarez-Ponce, D., Smirnova, E., Karstens, L., Dorrestein, P. C., Li, H., Sen Gupta, A., Cheung, K., Powers, J. G., Zhao, Z., & Rosen, G. L. (2020). Emerging Priorities for Microbiome Research. *Frontiers in Microbiology*, *11*, 136. https://doi.org/10.3389/fmicb.2020.00136

Dahlin, L. R., Gerritsen, A. T., Henard, C. A., Van Wychen, S., Linger, J. G., Kunde, Y., Hovde, B. T., Starkenburg, S. R., Posewitz, M. C., & Guarnieri, M. T. (2019). Development of a high-productivity, halophilic, thermotolerant microalga Picochlorum renovo. *Communications Biology*, *2*(1), 388. https://doi.org/10.1038/s42003-019-0620-2

Dahlin, L. R., & Guarnieri, M. T. (2021). Development of the high-productivity marine microalga, Picochlorum renovo, as a photosynthetic protein secretion platform. *Algal Research*, *54*, 102197. https://doi.org/https://doi.org/10.1016/j.algal.2021.102197

Faria, J. P., Rocha, M., Rocha, I., & Henry, C. S. (2018). Methods for automated genome-scale metabolic model reconstruction. *Biochemical Society Transactions*, *46*(4), 931–936. https://doi.org/10.1042/BST20170246

Grigoriev, I. V, Hayes, R. D., Calhoun, S., Kamel, B., Wang, A., Ahrendt, S., Dusheyko, S., Nikitin, R., Mondo, S. J., Salamov, A., Shabalov, I., & Kuo, A. (2021). PhycoCosm, a comparative algal genomics resource. *Nucleic Acids Research*, *49*(D1), D1004–D1011. https://doi.org/10.1093/nar/gkaa898

Hanschen, E. R., & Starkenburg, S. R. (2020). The state of algal genome quality and diversity. *Algal Research*, *50*, 101968. https://doi.org/https://doi.org/10.1016/j.algal.2020.101968

Hoffman, Y., Aflalo, C., Zarka, A., Gutman, J., James, T. Y., & Boussiba, S. (2008). Isolation and characterization of a novel chytrid species (phylum Blastocladiomycota), parasitic on the green alga Haematococcus. *Mycological Research*, *112*(1), 70–81. https://doi.org/https://doi.org/10.1016/j.mycres.2007.09.002

Howe, K. L., Contreras-Moreira, B., De Silva, N., Maslen, G., Akanni, W., Allen, J., Alvarez-Jarreta, J., Barba, M., Bolser, D. M., Cambell, L., Carbajo, M., Chakiachvili, M., Christensen, M., Cummins, C., Cuzick, A., Davis, P., Fexova, S., Gall, A., George, N., … Flicek, P. (2020). Ensembl Genomes 2020—enabling non-vertebrate genomic research. *Nucleic Acids Research*, *48*(D1), D689–D695. https://doi.org/10.1093/nar/gkz890

Kim, B.-H., Ramanan, R., Cho, D.-H., Oh, H.-M., & Kim, H.-S. (2014). Role of Rhizobium, a plant growth promoting bacterium, in enhancing algal biomass through mutualistic interaction. *Biomass and Bioenergy*, *69*, 95–105. https://doi.org/https://doi.org/10.1016/j.biombioe.2014.07.015

Krishnan, A., Cano, M., Burch, T. A., Weissman, J. C., & Posewitz, M. C. (2020). Genome editing using Cas9-RNA ribonucleoprotein complexes in the high-productivity marine alga Picochlorum celeri. *Algal Research*, *49*, 101944. https://doi.org/https://doi.org/10.1016/j.algal.2020.101944

Lee, J., Cho, D.-H., Ramanan, R., Kim, B.-H., Oh, H.-M., & Kim, H.-S. (2013). Microalgae-associated bacteria play a key role in the flocculation of Chlorella vulgaris. *Bioresource Technology*, *131*, 195–201. https://doi.org/https://doi.org/10.1016/j.biortech.2012.11.130

Matthiessen, B., Gamfeldt, L., Jonsson, P. R., & Hillebrand, H. (2007). EFFECTS OF GRAZER RICHNESS AND COMPOSITION ON ALGAL BIOMASS IN A CLOSED AND OPEN MARINE SYSTEM. *Ecology*, *88*(1), 178–187. https://doi.org/https://doi.org/10.1890/0012-9658(2007)88[178:EOGRAC]2.0.CO;2

Olivier, B. G., & Bergmann, F. T. (2018). SBML Level 3 Package: Flux Balance Constraints version 2. *Journal of Integrative Bioinformatics*, *15*(1). https://doi.org/doi:10.1515/jib-2017-0082

Oren, A. (2005). A hundred years of Dunaliella research: 1905–2005. *Saline Systems*, *1*(1), 2. https://doi.org/10.1186/1746-1448-1-2

Ramanan, R., Kim, B.-H., Cho, D.-H., Oh, H.-M., & Kim, H.-S. (2016). Algae–bacteria interactions: Evolution, ecology and emerging applications. *Biotechnology Advances*, *34*(1), 14–29. https://doi.org/https://doi.org/10.1016/j.biotechadv.2015.12.003

Sen, P., & Orešič, M. (2019). Metabolic Modeling of Human Gut Microbiota on a Genome Scale: An Overview. *Metabolites*, *9*(2), 22. https://doi.org/10.3390/metabo9020022

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

Teng, S. Y., Yew, G. Y., Sukačová, K., Show, P. L., Máša, V., & Chang, J.-S. (2020). Microalgae with artificial intelligence: A digitalized perspective on genetics, systems and products. *Biotechnology Advances*, *44*, 107631. https://doi.org/https://doi.org/10.1016/j.biotechadv.2020.107631

U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Bioenergy Technologies Office. (2016). *Algal Biology Toolbox Workshop Summary Report* (Issue May 2016). https://www.energy.gov/sites/prod/files/2016/09/f33/algal_biology_toolbox_workshop_summary_report.pdf

Ungerer, J., Wendt, K. E., Hendry, J. I., Maranas, C. D., & Pakrasi, H. B. (2018). Comparative genomics reveals the molecular determinants of rapid growth of the cyanobacterium Synechococcus elongatus UTEX 2973. *Proceedings of the National Academy of Sciences*, *115*(50), E11761 LP-E11770. https://doi.org/10.1073/pnas.1814912115

Weissman, J. C., Likhogrud, M., Thomas, D. C., Fang, W., Karns, D. A. J., Chung, J. W., Nielsen, R., & Posewitz, M. C. (2018). High-light selection produces a fast-growing Picochlorum celeri. *Algal Research*, *36*, 17–28. https://doi.org/https://doi.org/10.1016/j.algal.2018.09.024

Widder, S., Allen, R. J., Pfeiffer, T., Curtis, T. P., Wiuf, C., Sloan, W. T., Cordero, O. X., Brown, S. P., Momeni, B., Shou, W., Kettle, H., Flint, H. J., Haas, A. F., Laroche, B., Kreft, J.-U., Rainey, P. B., Freilich, S., Schuster, S., Milferstedt, K., … Fellows, I. N. I. (2016). Challenges in microbial ecology: building predictive understanding of community function and dynamics. *The ISME Journal*, *10*(11), 2557–2568. https://doi.org/10.1038/ismej.2016.45

Yu, J., Liberton, M., Cliften, P. F., Head, R. D., Jacobs, J. M., Smith, R. D., Koppenaal, D. W., Brand, J. J., & Pakrasi, H. B. (2015). Synechococcus elongatus UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO2. *Scientific Reports*, *5*(1), 8132. https://doi.org/10.1038/srep08132

Zuñiga, C., Li, C.-T., Huelsman, T., Levering, J., Zielinski, D. C., McConnell, B. O., Long, C. P., Knoshaug, E. P., Guarnieri, M. T., Antoniewicz, M. R., Betenbaugh, M. J., & Zengler, K. (2016). Genome-Scale Metabolic Model for the Green Alga Chlorella vulgaris UTEX 395 Accurately Predicts Phenotypes under Autotrophic, Heterotrophic, and Mixotrophic Growth Conditions. *Plant Physiology*, *172*(1), 589 LP – 602. https://doi.org/10.1104/pp.16.00593

Zuñiga, C., Li, T., Guarnieri, M. T., Jenkins, J. P., Li, C.-T., Bingol, K., Kim, Y.-M., Betenbaugh, M. J., & Zengler, K. (2020). Synthetic microbial communities of heterotrophs and phototrophs facilitate sustainable growth. *Nature Communications*, *11*(1), 3803. https://doi.org/10.1038/s41467-020-17612-8