

Algal Translational Genomics (ATG)

**Bioenergy Technologies Office (BETO)
2019 Project Peer Review**

Advanced Algal Systems

Shawn Starkenburg

March 5, 2019





Overview

Timeline

- Project start: October 2018
- Project end : September 2021
- Percent complete: 10%

Budget

	Total Costs Pre FY17**	FY 17 Costs	FY 18 Costs	Total Planned Funding (FY 21-Project End Date)
DOE Funded	\$0	\$0	\$0	\$1350K
Project Cost Share*	\$0	\$0	\$0	\$0
Partners: New Mexico Consortium (25%), NREL (0%) Collaborator: Juergen Polle (CUNY)				

Barriers

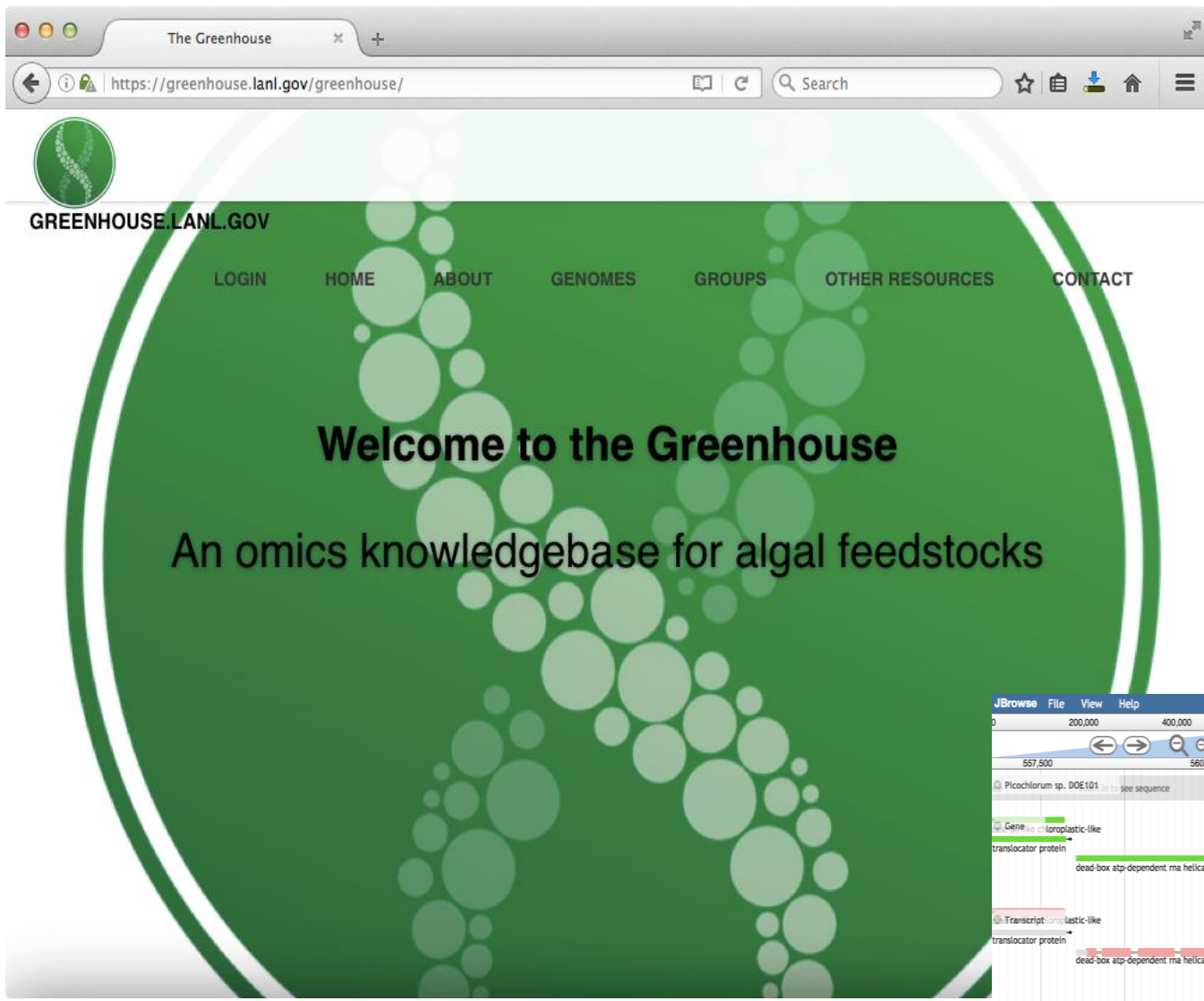
- **Lack of Biomass Genetics and Physiological Knowledge of Production Strains (AftC)**
- Working towards MYPP goals to improve productivity in mass cultivation

Objectives

Improve, curate, and validate genomic annotations of central carbon metabolism, carbon fixation, and organic carbon uptake to enable strain improvement and maximize carbon utilization

Overview- History

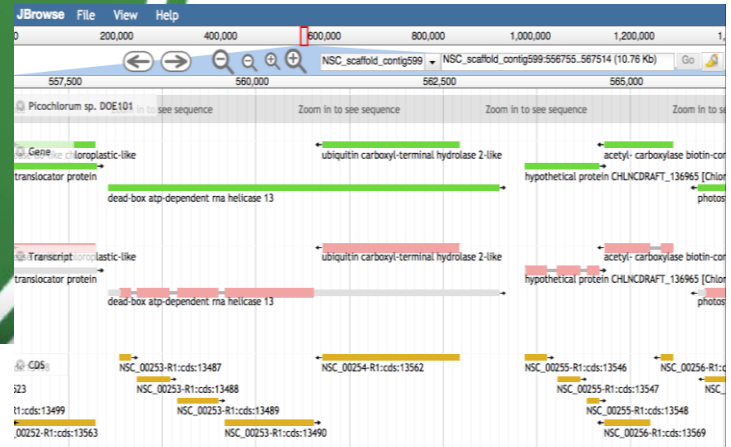
- Revamp of the Greenhouse AOP (2016-2018)
- Greenhouse Knowledgebase was built to provide genomics resources and analysis tools to improve the productivity and robustness of algae strains
- Create and disseminate fundamental knowledge of algal genomes to enable genetic manipulation work
- All Greenhouse milestones (+ stretch activity) completed



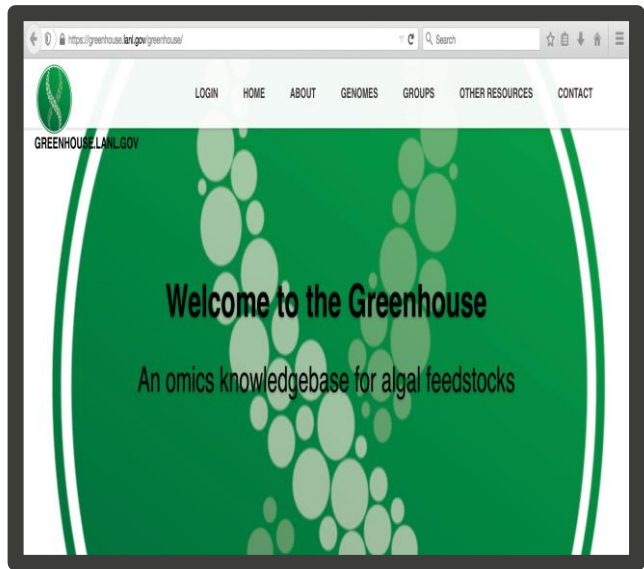
Public Web Portal

- Species Info
- Genome stats
- Download:
 - Assembly
 - Genes, Proteins
- Genome Viewer
- Blast Searching
- Other Metadata
- 6 'BETO' Genomes

Largest eukaryotic algal genome collection available online



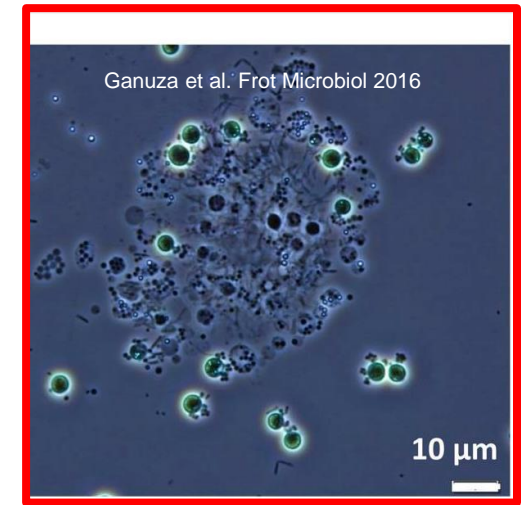
Genome Knowledgebase



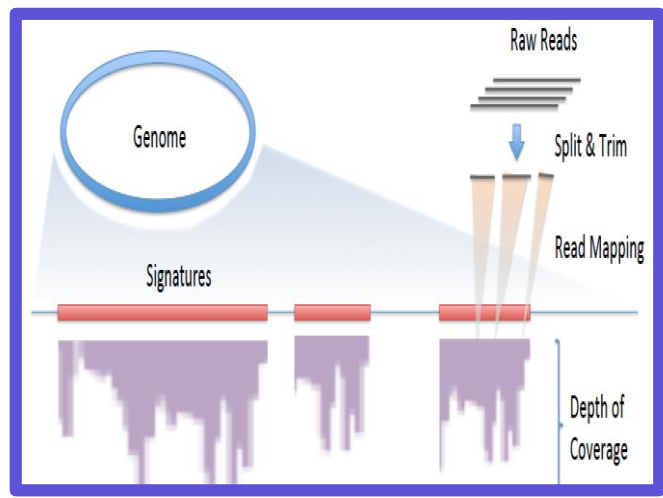
OMICS TOOLS DEVELOPMENT

FY16
FY17
FY18

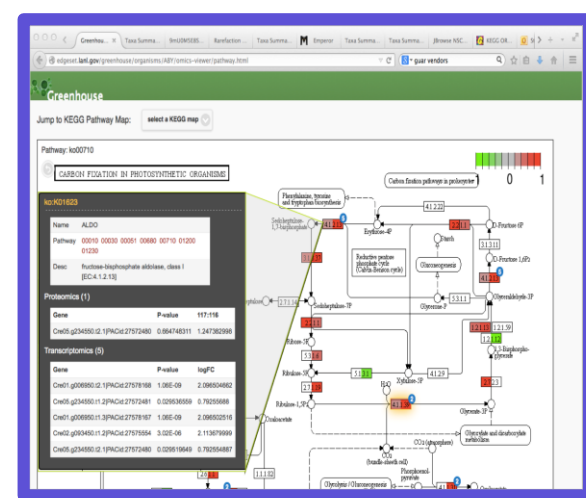
PREDATOR & PEST GENOMES (n=3)



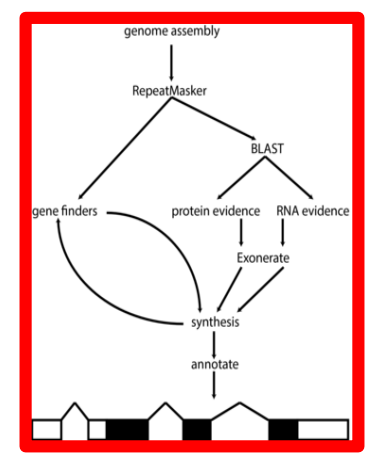
METAGENOMIC CULTIVATION DIAGNOSTICS



OPAVER: TIMECOURSE OMICS PATHWAY VIEWER



V2.0 GENOME ANNOTATION PIPELINE



ATG Project: Goals & Outcomes



Goal: Use genomics to accurately characterize the metabolic potential of leading production strains to enable performance improvements

Outcomes:

- curation of genome annotations to develop comprehensive models of canonical (and alternative) autotrophic and heterotrophic carbon assimilation
- experimental validation of both conserved and strain specific pathways for enhancements.
- lays the groundwork for improving biomass production rates under carbon- and light-limited conditions in large-scale ponds, and improve carbon input cost sustainability through discovery and utilization of efficient carbon metabolism pathways.

Motivation & Challenges

- **Few Metabolic Models**
 - IMPACT: impedes strain improvement and tailoring of cultivation conditions to maximize yield
- **Uncurated Annotations**
 - Genes misannotated, plus gaps in key metabolic pathways
- **Incomplete Genomes**
 - High degree of fragmentation- causes incomplete gene models
 - long repetitive DNA elements and allelic/structural variation in polyploid genomes confound assembly algorithms (only 1 euk. genome completed from telomere to telomere!)
 - IMPACT: hinders strain improvement via genetic modification, mating/breeding studies, and/or engineering of chromosomes to improve strain performance
- **Lack of Knowledge of Carbon Utilization from Waste Streams**



Objectives

Translate Genotype to Phenotype

- Develop and deploy new assembly methods to construct 100% complete genomes
- Curate carbon metabolism pathways for leading production strains
- Phenotype inorganic and organic carbon use to functionally validate genome annotations
- Maintain and expand the Greenhouse Knowledgebase



Approach

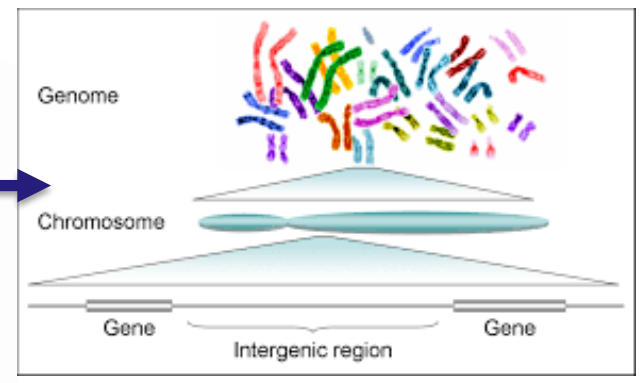
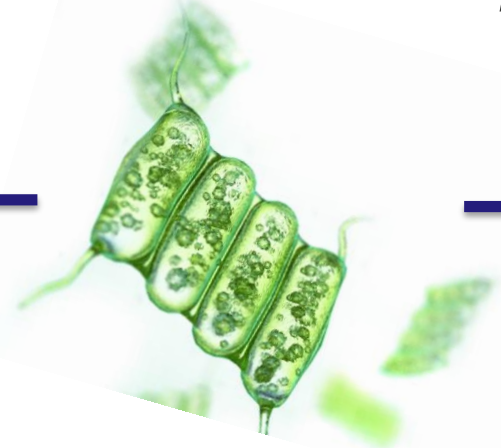
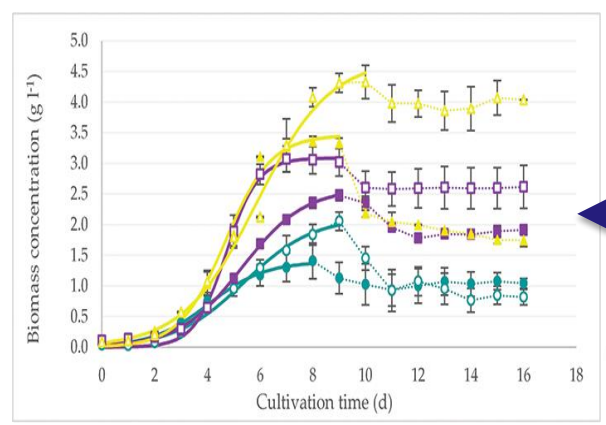
THE TEAM

Personnel	Role
Shawn Starckenburg	Science Direction, Administration
Blake Hovde	Website Maintenance & Expansion
Joseph Msanne (NMC)	Annotation Curation and Phenotyping Studies
Yan Xu, Justin Zhang, Hajni Daligault	Bioinformatics, System admin.
Yuliya Kunde	Sequencing/Lab Technician

MANAGEMENT- Weekly Team Meetings, One-on-One progress checks/consults as needed, Performance / Merit tied to Milestone Completion, and publications. Quarterly Communication with BETO (verbal and written).

Approach

Scenedesmus sp.



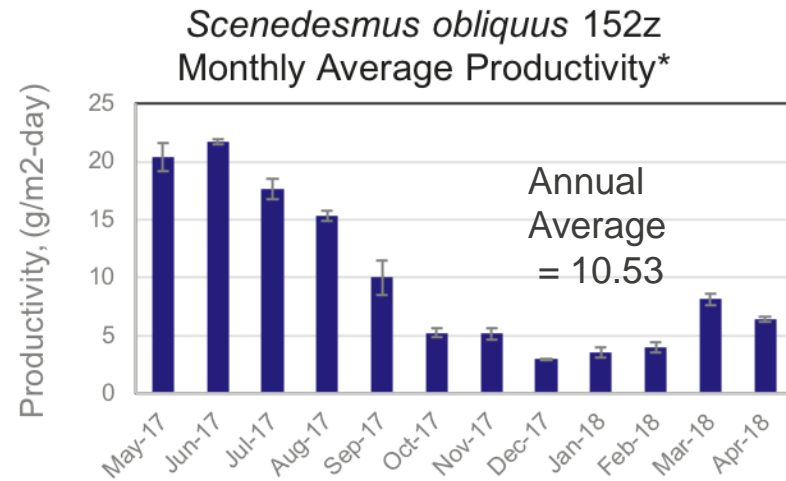
Task 1: Annotation Curation & Metabolic Pathway Construction

Task 2: Phenotyping and Validation Studies

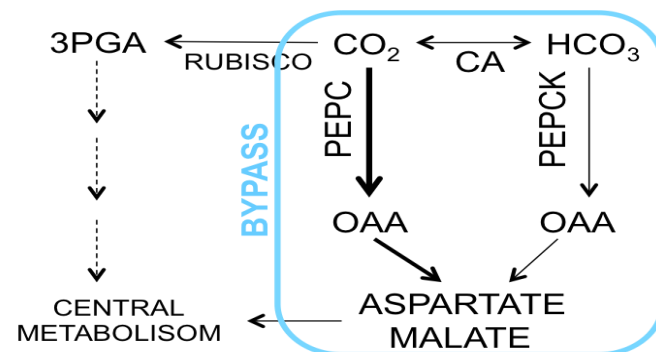
Task 3: Advancement of Genome Assembly & Analysis Tools

Why *Scenedesmus*?

- **Established Baseline Productivity**; contributes to SOT
- **Waste to Energy Potential**; high productivity and nutrient removal rates in waste water treatment systems
- **Genomic resources** for 4 strains to assess conserved and unique metabolic properties
 - Strain 46DB3 (Guarnieri, NREL)
 - *S. obliquus* 152Z, UTEX393
 - *S. quadricaudus*
- **Perfect Genome Size**; sufficiently large to challenge scaffolding methods



*Courtesy of MicroBio Engineering



Proposed Model of Carbon Fixation
In *Scenedesmus* (J. Polle, CUNY)



Technical Progress

Task 1: Annotation Curation & Metabolic Pathway Construction

COMPLETED ML1. Comprehensive literature search to identify carbon metabolism genes/proteins and known growth response(s) in wastewater sources

Report Delivered to BETO; will be adapted for publication as a minireview

LIT REVIEW HIGHLIGHTS

- Cultivated in Many Waste Streams
 - Municipal Secondary Effluent
 - Poultry
 - Food, Brewery, Dairy
 - Urban
 - Pharmaceuticals
 - Flue Gas CO₂ Compatible
- N&P removal (93-100%)
- COD removal (63-76%)
- TOC removal (60-65%)
- Highest productivity when grown mixotrophically
- Very few substrates have been tested for optimal heterotrophic or mixotrophic growth (acetate, glucose, molasses)
- Genes for “C4 metabolism” conserved in at least 3 strains
- Many enzymes for digestion of complex carbon molecules



Relevance

Develop technologies for mass cultivation that enable yields of 30 tons AFDW/ac/yr with yields of 80 GGE

Increase the summer areal productivity to 25 g/m²/d

- Optimizing productivity in organic rich wastewater sources would exceed MYPP productivity targets
- Helps achieve carbon input cost sustainability through discovery and utilization of efficient carbon metabolism pathways
- Disseminating genomics tools that enable academic and industrial entities to both contribute to the work and reap the benefits of BETO funded national laboratory programs
- Enabling genetic manipulation and breeding of production strains to improve biomass yields/traits



Future Work

Task 1: Annotation Curation & Metabolic Pathway Construction



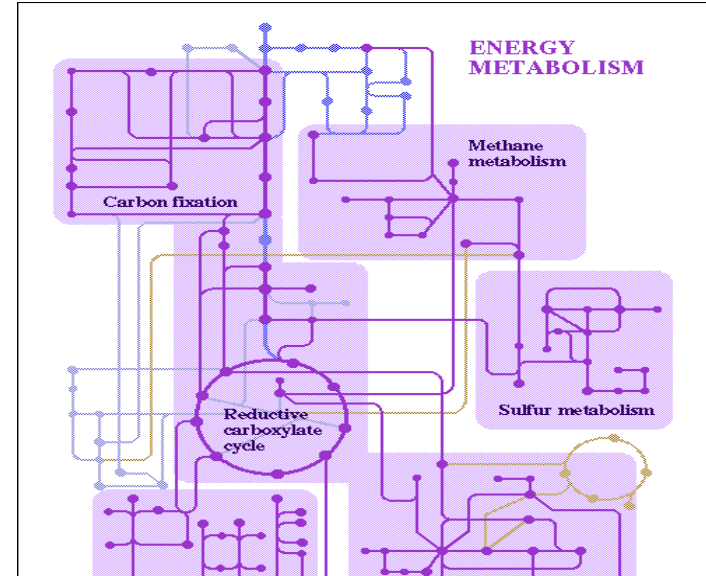
Literature Search
What is known about C metabolism in *Scenedesmus*?



Construct Custom C Metabolism Model



Curate Metabolic Maps with Phenotyping Data



Key Milestone: Custom, *in silico* model constructed for organic and inorganic carbon uptake, regulation, and metabolism conserved in all sequenced *Scenedesmus* strains with zero gaps in proposed pathways.

Future Work

Task 2: Phenotyping and Validation Studies

Carbon Utilization Studies
C2-C6, Polymers



Measure systems level
response in the best C
sources that boost productivity

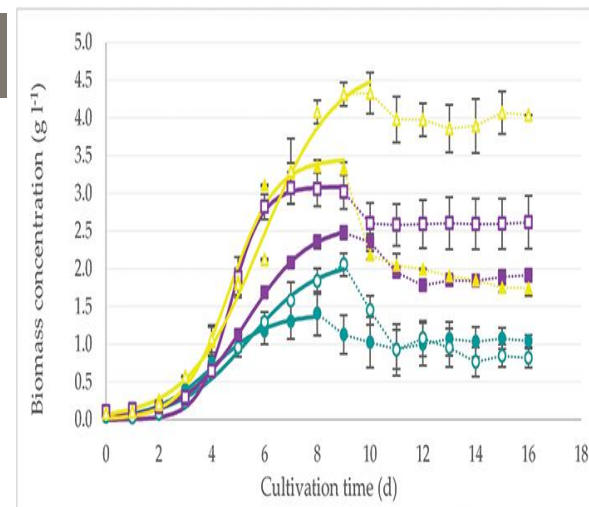


Targeted Metabolic Analysis
 ^{13}C -CO₂ or Bicarbonate



Curate Metabolic Maps with
Phenotyping Data

Validate models by evaluating the direct contribution of various organic compounds to the growth and biomass composition of *S. obliquus*



Biotechnology for Biofuels 2018 11:308
<https://doi.org/10.1186/s13068-018-1308-4>

Key Milestone: Studies completed to screen for improved growth in mixotrophic and heterotrophic growth on C2-C6 monomers and HMW polymers over autotrophic conditions.

GN 1. Replicated studies demonstrating 30% improved growth in mixotrophic vs autotrophic conditions. (Exceeds MYPP 2025 Productivity Targets)

Future Work

Task 3: Advancement of Genome Assembly & Analysis Tools

Extraction/Purification of HMW DNA

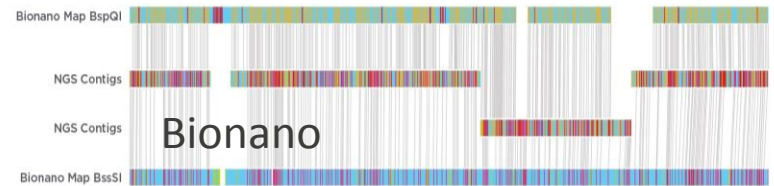
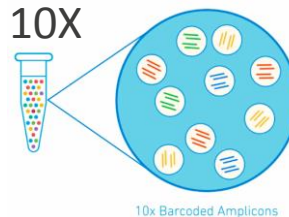
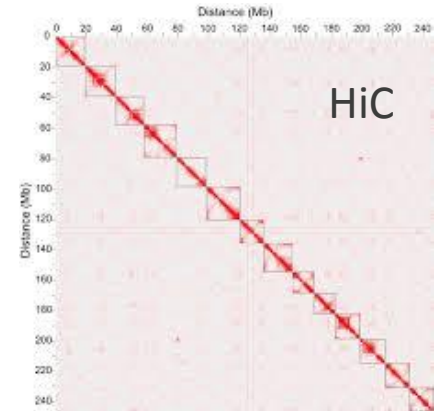


Construct Scaffolding Libraries

Scaffold S. obliquus UTEX 393
with 3 Technologies

Publish Results – Complete Genome + Curated Annotation

Key Milestone: Comparison of Contiguity (Telomere ends, N50, Gap Count, SNP analysis) from HiC, Bionano, and 10X Scaffolding Technologies.



Final Milestone: 100% complete assembly of an algal genome (establishes new “gold standard”)

Built In Mitigation- Comparing 3 Technologies, Conduct manual finishing if necessary



Los Alamos
NATIONAL LABORATORY
— EST. 1943 —

ADDITIONAL SLIDES

ALL FY19-FY2021 ATG Milestones	End Date
Comprehensive literature search to identify validated carbon metabolism genes/proteins and known growth response(s) in wastewater sources.	12/31/2018
Gene Inventory for all Sequenced Scenedesmus Genomes and Public Transcriptome Datasets Collected from NCBI	3/31/2019
DNA Extracted From Scenedesmus obliquus 152Z - 50 ug of High Molecular Weight (50 kbp average length) to support construction of 3 scaffolding libraries.	6/30/2019
Custom, in silico model constructed for organic and inorganic carbon uptake, regulation, and metabolism conserved in all sequenced Scenedesmus strains with zero gaps in proposed pathways.	9/30/2019
New publicly available genomes on NCBI deposited into Greenhouse (stretch)	3/31/2019
Replicated studies completed to screen for 30% improved growth in mixotrophic and heterotrophic growth on C2-C6 monomers over autotrophic conditions.	12/31/2019
New publicly available genomes on NCBI deposited into Greenhouse	3/31/2020
Replicated (n=3) experiments conducted to assess for growth on high molecular weight carbon molecules (e.g. lignocellulosics, pectin, starch).	6/30/2020
Comparison of Contiguity (Telomere ends, N50, Gap Count, SNP analysis) from HiC, Bionano, and 10X Scaffolding Technologies.	9/30/2020
Transcriptome differential expression analysis comparing growth under high vs. low CO₂ levels in triplicate.	12/31/2020
Genome Announcement published reporting 100% complete assembly of an algal genome	3/31/2021
¹³C labeling of C4 metabolites and biomass production rates (AFDW over 2 week time course in triplicate) are higher than N. salina under C limiting conditions	6/30/2021
Curated/validated metabolic models of inorganic and organic carbon uptake displayed on the Greenhouse website.	9/30/2021

Publications, Patents, Presentations

Blake Hovde, Erik Hanschen, Christina Steadman Tyler, Chienchi Lo, Yuliya Kunde, Karen Davenport, Hajnalka Daligault, Joseph Msanne, Stephanie Canny, Seong-il Eyun, Jean-Jack Riethoven, Juergen Polle, **Shawn Starkenburg**. 2018. Genomic characterization reveals significant divergence within *Chlorella sorokiniana* (Chlorellales, Trebouxiophyceae). *Algal Research*. Vol. 35, Nov 2018, 449-461. <https://doi.org/10.1016/j.algal.2018.09.012>

Brian W. Vogler, **Shawn R. Starkenburg**, Nilusha Sudasinghe, Jenna Y. Schambach, Joseph A. Rollin, Sivakumar Pattathil, Amanda N. Barry, Characterization of plant carbon substrate utilization by *Auxenochlorella protothecoides*, *Algal Research*, Vol. 34, 2018, Pages 37-48, ISSN 2211-9264, <https://doi.org/10.1016/j.algal.2018.07.001>.

CR Gonzalez-Esquer, SN Twary, BT Hovde, and **SR Starkenburg**. 2018. Nuclear, chloroplast, and mitochondrial genome sequence of the microalgal biofuel-prospect strain *Picochlorum soloecismus*. *Genome Announc*. 2018 Jan 25;6(4). pii: e01498-17. doi: 10.1128/genomeA.01498-17.

Starkenburg SR, Polle JEW, Hovde B, Daligault HE, Davenport KW, Huang A, Neofotis P, McKie-Krisberg Z. 2017. Draft nuclear genome, complete chloroplast genome, and complete mitochondrial genome for the biofuel/bioprocess feedstock species *Scenedesmus obliquus* strain DOE0152z. *Genome Announc* 5:e00617-17. <https://doi.org/10.1128/genomeA.00617-17>.

K.R. Fixen*, **S.R. Starkenburg***, **B.T. Hovde**, S.L. Johnson, C.R. Deodato, C. S. Harwood, and R.A. Cattolico. 2016. Genome sequences of 8 bacterial species found in co-culture with the haptophyte *Chrysochromulina tobin*. *Genome Announcements*. 2016 Nov 3;4(6). pii: e01162-16. doi: 10.1128/genomeA.01162-16. PubMed PMID: 27811091;

S.H. Park, **S.R. Starkenburg**, J. Kyndt, A. Angelova, O. Chertkov, and J. K. Brown. The Compact Chloroplast Genome of *Auxenochlorella protothecoides* Harbors Bacterial and Viral Sequences Suggestive of Horizontal Transfer. *Int. J. Genomics*. 2018. submitted

Blake Hovde, Seth Steichen, **Shawn R. Starkenburg** and Judy Brown. 2018. *Vampirovibrio chlorellavorus* draft genome sequencing, annotation, and a preliminary characterization of pathogenicity. *Microbial Genomics*. Submitted.

Publications, Patents, Presentations

Patents:

Negi, S., **Starkenburger, S.**, and R. Sayre. 2015. Improved Productivity and Bioproduct Formation in Phototropin Knock/Out Mutants in Microalga. Provisional Application (NMC0003-101-US PHOT K/O) filed June 4, 2015. Full application filed July 2016. *Most of the protein sequences claimed in this patent came from genomes sequenced through the Greenhouse AOP by BETO.*

Presentations:

'Greenhouse: an omics resource for algal feedstocks'. Algal Biomass, Biofuels and Bioproducts. San Diego, CA, June 27th, 2016

'Greenhouse: an omics resource for algal feedstocks'. Algal Biomass, Biofuels and Bioproducts. Orlando, FL 2017

'Greenhouse: an omics resource for algal feedstocks'. Algal Biomass Summit, Glendale AZ, Oct 25th, 2016

"Algal Omics Databases and Resources", Sharpening Our Tools: Algal Biology Toolbox Workshop, DOE-Bioenergy Technology Office, May 2016, San Diego, CA

"Advances in Algal Genomics", Lecturer, Spring ATP3 Workshop; Principles and Processes: Algae Culture Maintenance, Production and Downstream Processing, April 2016, Los Alamos, NM

Invited Speaker, 2018, Annual Meeting for the Society for Industrial Microbiology and Biotechnology, Chicago, IL, Title: "Deciphering a Chlorella Phycosphere"

Responses to Previous Reviewers' Comments

- This AOP was a new start in FY19 and was not apart of the 2015 Peer Review. However, components of a previous AOP (*'The Greenhouse': A Comprehensive Knowledgebase of Algal Feedstocks*) are carried through into this new AOP.