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%

Barriers

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

Budget

<table>
<thead>
<tr>
<th></th>
<th>Total Costs Pre FY17**</th>
<th>FY 17 Costs</th>
<th>FY 18 Costs</th>
<th>Total Planned Funding (FY 21-Project End Date)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOE Funded</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$1350K</td>
</tr>
<tr>
<td>Project Cost Share*</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
</tr>
</tbody>
</table>

Partners: New Mexico Consortium (25%), NREL (0%)
Collaborator: Juergen Polle (CUNY)

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


- 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
Welcome to the Greenhouse

An omics knowledgebase for algal feedstocks

Largest eukaryotic algal genome collection available online
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
THE TEAM

<table>
<thead>
<tr>
<th>Personnel</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shawn Starkenburg</td>
<td>Science Direction, Administration</td>
</tr>
<tr>
<td>Blake Hovde</td>
<td>Website Maintenance &amp; Expansion</td>
</tr>
<tr>
<td>Joseph Msanne (NMC)</td>
<td>Annotation Curation and Phenotyping Studies</td>
</tr>
<tr>
<td>Yan Xu, Justin Zhang, Hajni Daligault</td>
<td>Bioinformatics, System admin.</td>
</tr>
<tr>
<td>Yuliya Kunde</td>
<td>Sequencing/Lab Technician</td>
</tr>
</tbody>
</table>

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.

Year 1
Task 1: Annotation Curation & Metabolic Pathway Construction

Year 2
Task 2: Phenotyping and Validation Studies

Year 3
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

**Proposed Model of Carbon Fixation**

*Scenedesmus obliquus* 152z

Monthly Average Productivity*

*Courtesy of MicroBio Engineering*
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/m$^2$/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
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 13C-CO2 or Bicarbonate

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

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


Publications, Patents, Presentations

Patents:

Presentations:


“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.