DOE Bioenergy Technologies Office (BETO) 2021 Project Peer Review

Advanced Algal Biofoundries for the Production of Polyurethane Precursors

March 11, 2021
Technology Area Session

Stephen Mayfield
UC San Diego
Nathan Hillson (LBL) & Jeremy Zucker (PNNL)

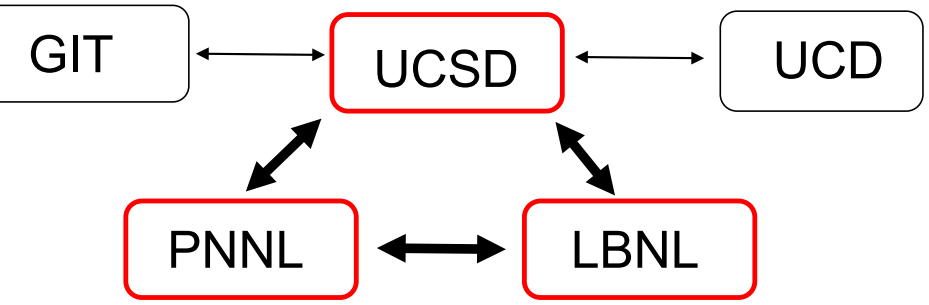
Project Overview

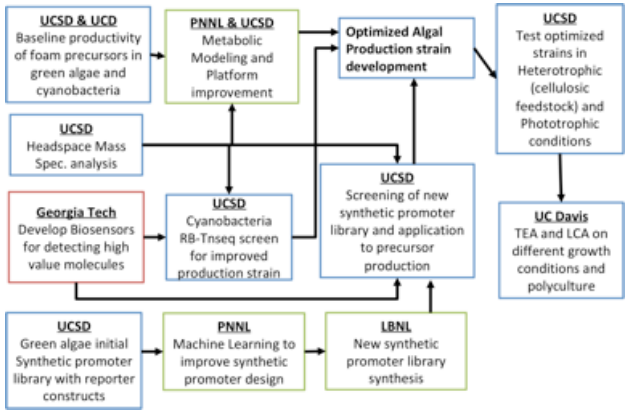
- **Mission of project**: Develop scalable, high-yield algae production platforms capable of producing polyurethane precursors for the sustainable manufacturing of bio-based, biodegradable, and recyclable foams and plastic products

- **Key Academic Partners:** UC San Diego, PNNL, LBL, UC Davis, & GT

- **Cost Share Partners:** Algenesis Materials, Reef, and Arctic Foam

1 – Management



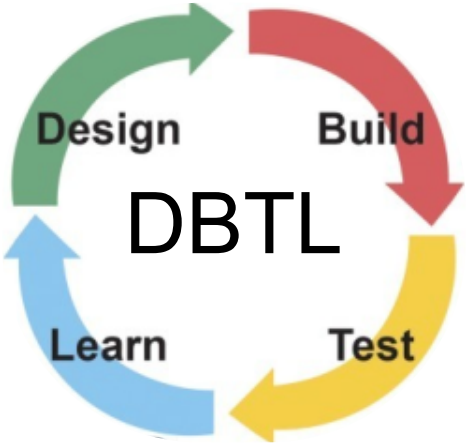


Key Tasks and Responsibility:

1. Perform multiple rounds of DBTL to optimize synthetic nuclear promoters for *Chlamydomonas* (PNNL, LBL, UCSD)
2. Metabolic modeling (PNNL) & subsequent engineering (UCSD & UCD) to optimize chemical production under heterotrophy or phototrophy
3. Develop & assess baseline production strains (UCSD & UCD)
4. Generate biosensors (GIT) and MS-based (UCSD) high-throughput screening systems
5. Scaled production (UCSD) & TEA/LCA assessments (UCD)

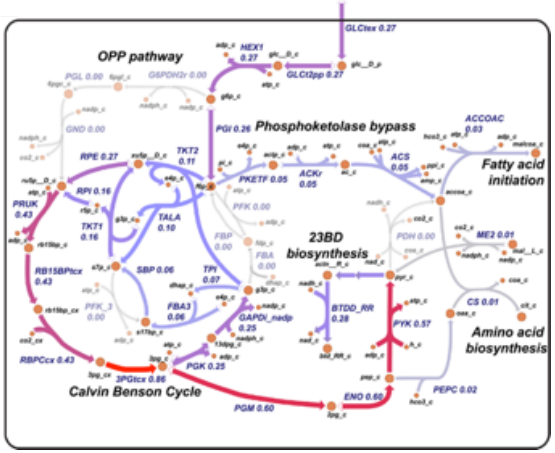
2 – Approach

Improve Synthetic Algal Nuclear Promoters



Perform 2+ rounds of DBTL:
D: PNNL designs library of promoters based on experimental data from **UCSD**
B: **LBNL** synthesizes & clones library
T: UCSD tests promoter strengths in algae
L: PNNL develop ML tools to analyze results for subsequent rounds

Predict ME Targets for Improved Production



PNNL leverages metabolic models for cyanobacteria & algae to predict heterologous pathways, KOs, or over-expressions to optimize yields
UCSD & UC Davis tests these predictions & feeds back info to the models for further optimization.

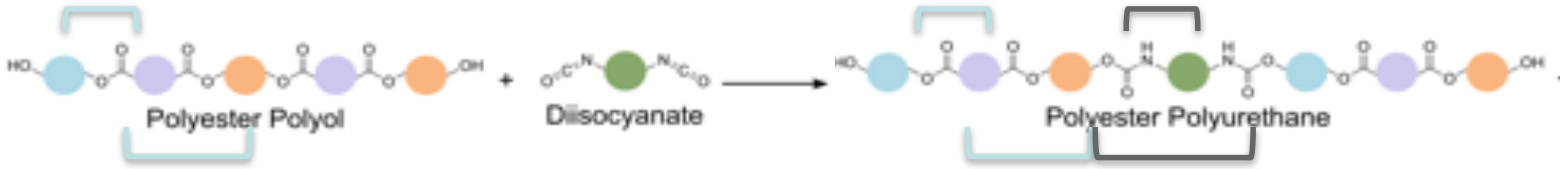
Scale Production

UCSD scale production in both heterotrophic and autotrophic growth for TEA/LCA and to meet performance metrics.

Go/No-Go decision points = 1 gm/liter PU precursor

3 – Impact

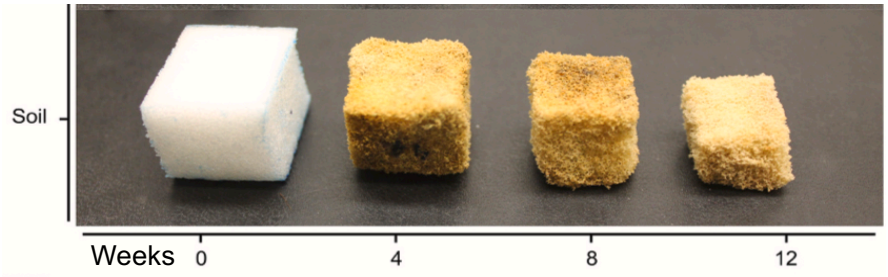
Biodegradable & Sustainable Polyurethanes



ALGENESIS

REEF

ARCTIC ALGAE



4 – Progress and Outcomes

Synthetic promoter design

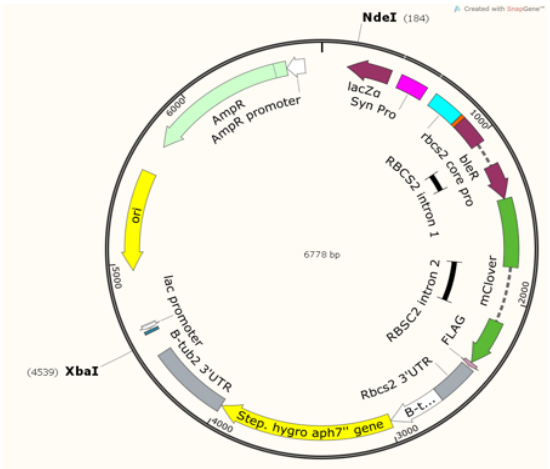
- **160bp promoters** trained on expression data from native promoter sequences
  - Neural network learned an embedding space mapping motifs and sequences to expression
  - AutoML learned RandomForest discriminator
- **260bp promoters** also contained predicted chromatin opening motifs in the upstream 100bp
  - Chromatin-opening motifs found empirically to be further from TSS
  - Motifs predicted from enrichment analysis of native expression data and native histone and RNAPol ChIP-Seq data.
- **Exploration**: designed low-confidence positive and negative promoters
- **Exploitation**: both methods used combinations of known high-expressing motifs

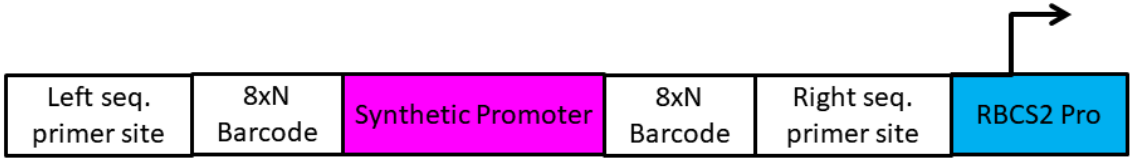
- 1500 X 160-bp and 1500 X 260-bp long synthetic promoters were designed at PNNL
- Promoter libraries were synthesized by Twist via the LBNL - ABL
- First round of libraries screened at UCSD and 2nd round designed and out for synthesis

4 – Progress and Outcomes

Promoter library construction

- LBL added barcodes to Twist-synthesized promoters and cloned oligos into a reporter vector designed by UCSD
- DNA library has been delivered to UCSD and transformations into *Chlamydomonas reinhardtii* are underway.

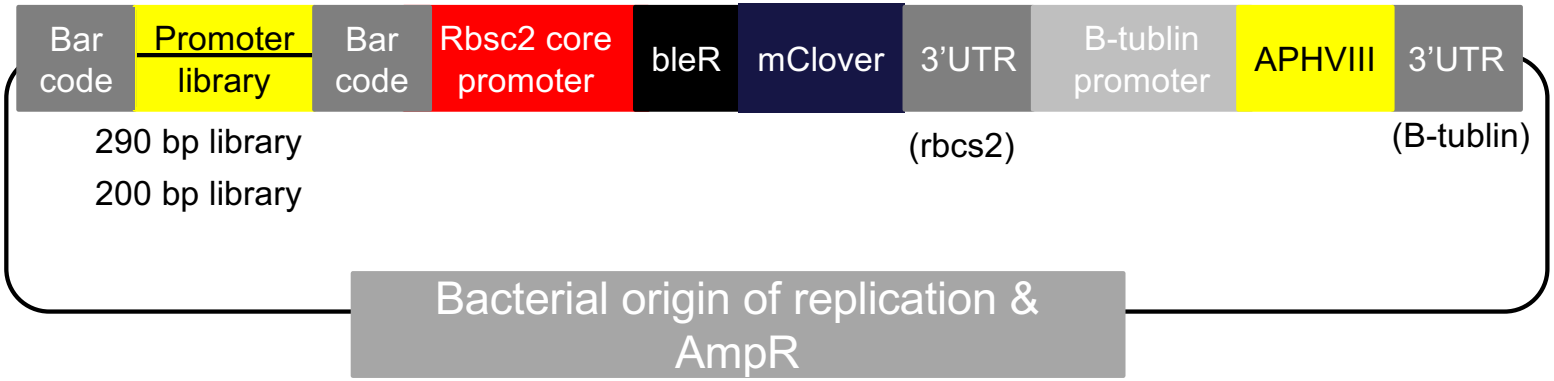




1500 290-bp and 1500 200-bp long synthetic promoters ordered

4 – Progress and Outcomes

Reporter vector design



Voltage: 500 V
Resistance: 800 Ohm
Capacity: 50 µF
Time constant: 3 ms

Electroporation into *C. reinhardtii* (CC-124)

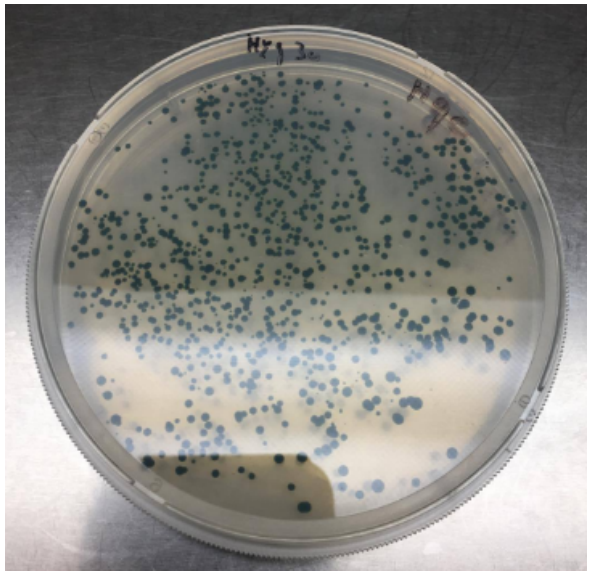
Hyg

Hyg+Zeo

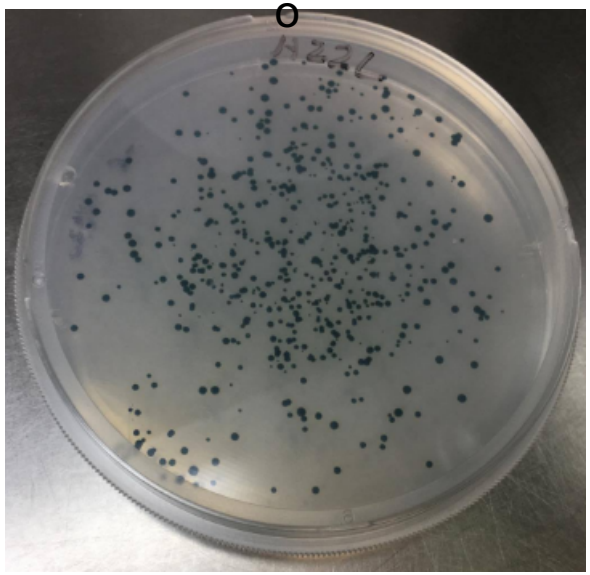
4 – Progress and Outcomes

Library	Hyg	Hyg+Zeo
200 bp	~4000 colonies	~600 colonies
290 bp	~4000 colonies	~450 colonies

Hyg

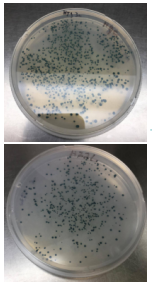


Hyg+Zeo



5 replicate transformations were made

4 – Progress and Outcomes



Recover
for 4-6 days

Harvesting &
getting
pellets

DNA
extraction

2 step PCR
- Adding
adapters
- Adding
indexes

Send for sequencing

FACS (Sorting Top 10%)

Harvesting &
getting pellets

DNA extraction

2 step PCR
- Adding adapters
- Adding indexes

Send for sequencing

Picking single colonies

Fluorescent intensity
measurement
Top expressors

Send for sequencing

1 complete round

4 – Progress and Outcomes

200 bp library (Tap + Light)

<>	1	2	3	4	5	6	7	8	9	10	11	12
A	0.95372	1.02403	0.37941	0.08195	0.13417	0.40463	0.28867	0.96536	0.85335	0.57334	0.07666	0.75164
B	0.09664	0.87577	0.54517	1.48784	0.49818	0.64185	0.26716	0.61128	0.64637	0.64539	0.58224	0.46919
C	0.57324	0.64768	0.13062	0.66614	0.22241	1.02571	0.89963	0.42023	0.52139	0.32952	0.80015	0.41149
D	0.48674	0.24893	0.8184	0.85876	0.60651	0.91873	0.69128	0.88772	0.43604	0.38825	0.34852	0.4104
E	0.96837	0.97287	0.94873	0.63096	0.64033	0.73714	1.07177	0.76428	0.55084	0.56026	0.56043	0.57189
F	0.48871	0.84526	0.87743	0.28122	0.7871	0.57909	0.80794	0.56353	1.09065	0.64586	0.54583	0.6262
G	0.09356	0.32825	0.59998	1.13067	0.58181	0.6	0.88035	0.3768	0.35914	0.46659	0.39118	0.63168
H	0.64894	0.61915	0.49459	0.65878	0.57559	0.41479	0.66962	0.65799	0.21425	0.55151	0.39487	0.60269

200 bp library (Tap + Dark)

<>	1	2	3	4	5	6	7	8	9	10	11	12
A	0.5557	0.53294	0.35469	0.09141	0.08944	0.37512	0.23739	0.75474	0.78011	1.21565	0.07495	0.47757
B	0.13666	1.07544	0.41297	0.53735	0.65454	0.62761	0.29731	0.63014	0.64233	0.31949	0.95421	0.43256
C	0.3472	0.62664	0.10997	0.3493	0.10791	0.23536	0.78908	0.19083	0.78242	0.47183	0.26175	0.48899
D	0.9644	0.1279	0.75141	0.70042	0.56498	0.83576	0.46123	1.0005	0.42966	0.56956	0.19506	0.3833
E	1.09312	1.35732	0.85387	0.52164	0.58951	0.59656	0.3477	0.96968	0.74225	0.75727	0.94672	0.75108
F	0.30109	0.45356	0.90955	0.1812	0.90095	0.49608	0.87134	0.41659	0.54724	0.53112	0.591	0.63898
G	0.11564	0.29795	0.68323	0.64283	0.38906	0.22245	0.53768	0.18266	0.27055	0.40303	0.32333	0.33541
H	0.62643	0.51684	0.4616	0.45602	0.366	0.15916	0.64246	0.41879	0.14356	0.29363	0.23338	0.30068

4 – Progress and Outcomes

Promoter assessment

- Separate Analyses for FACS-Sorted (enriched) and Unsorted Barcode Count Data
  - Survival Index Distribution (Gamma Distribution) – Hyg (expression) / HygZeo (high expression)
  - HygZeo Survival Probability (Binomial Distribution)
- Combined Analysis for FACS-Sorted and Unsorted Barcode Count Data
  - Fluorescence CDF Ratios (Non-Parametric Distribution) (protein accumulation)

$$H_i \sim \Gamma(\alpha_i, \beta_i)$$
$$Z_i \sim \Gamma(\gamma_i, \eta_i)$$
$$S_i = \frac{Z_i}{H_i}$$

$$H_i \sim B\left(n_i, p_i^H\right)$$
$$Z_i \sim B\left(H_i, p_i^Z\right)$$
$$\Rightarrow Z_i \sim B\left(n_i, p_i^H p_i^Z\right)$$

$$\int_0^x p_{F_i}(f_i)\, df_i = C_{F_i}(x)$$
$$\frac{1}{m}\sum_i C_{F_i}(x^*) = 0.9$$

$$E_i\left[\ln \frac{Z_i}{H_i}\right] = E[\ln Z_i] - E[\ln H_i]$$
$$= \psi(\gamma_i) - \psi(\alpha_i) + \ln(\beta_i) - \ln(\eta_i)$$

$$\int_{f_{thresh}}^{\infty} p_{F_i}(f_i)\, df_i = p_i^Z$$

$$\frac{Z_i Z_j^{FACS}}{Z_j Z_i^{FACS}} = \frac{1 - C_{F_i}(x^*)}{1 - C_{F_j}(x^*)}$$

4 – Progress and Outcomes

Prediction method

- Nucleotide - and Motif-Based Promoter Sequence Embeddings
- 5 Separate Predictors Per Embedding
  - 2 Methods (Gamma and Binomial Distributions) each for Standard and FACS data
  - 1 Combined Method (Non-parametric CDF)
- Nonlinear Regressions with Deep Learning Trained on Analysis Results
  - Random Promoter Generation
  - Average Performance Across Predictors

$$E_i\left[\ln\frac{Z_i}{H_i}\right] \qquad\qquad p_i^Z$$

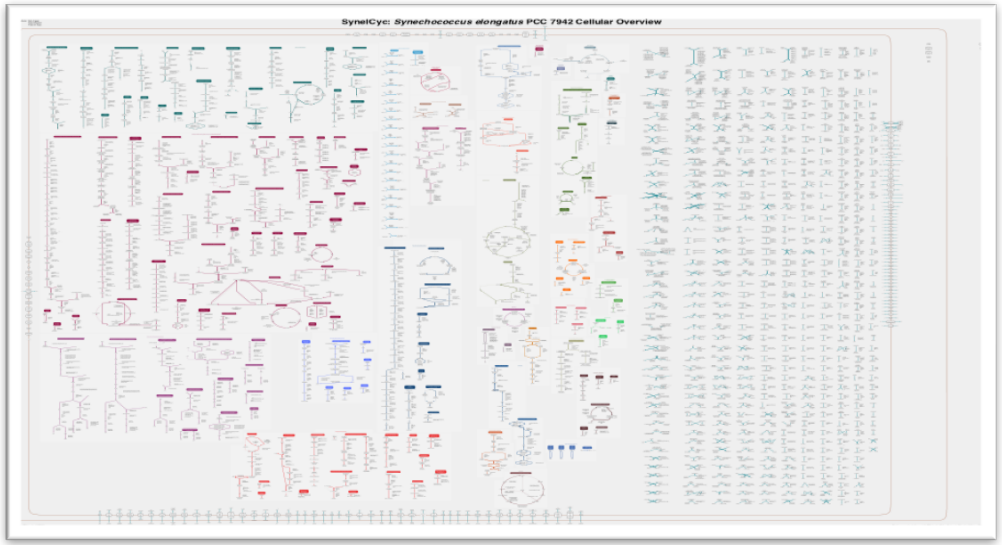
$$d_{ij} \equiv \log\left(1 - C_{F_i}\left(x^*\right)\right) - \log\left(1 - C_{F_i}\left(x^*\right)\right) = \log Z_i - \log Z_i^{FACS} - \left(\log Z_j - \log Z_j^{FACS}\right)$$

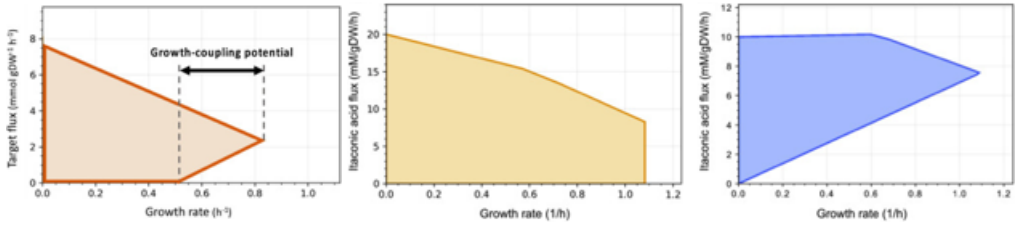
$$d_{ij} = -d_{ji}$$

$$\min_\theta \sum_{i,j} \left\| N\left(x_i;\theta\right) - N\left(x_j;\theta\right) - d_{ij}\right\|^2 + \epsilon \sum_i \left\| N\left(x_i;\theta\right)\right\|^2,\ \epsilon \ll 1$$

4 – Progress and Outcomes

Metabolic Models (MMs)





- Preliminary MMs have been coded for *Synechococcus elongatus* and *C. reinhardtii*, including essentiality and transcriptome data from UCSD.
- OptCouple, OptKnock, OptForce and **OptTilt** have been applied to these models to predict strains capable of improved production
- For example, 24 gene knockouts were predicted to improve diacid production in *S. elongatus.*

4 – Progress and Outcomes

Cell Factory Design: New Theorem (PNNL)

- OptTilt beats OptCouple and OptKnock
- ‘Tilting’ Term
  - Generates growth-coupled solution
  - Guarantees optimality
  - Ensures maximal production

$$\max_{y} v_{target}$$

$$y_j \in \{0,1\}$$

$$\sum_{j \in R_{native}} (1 - y_j) \leq K_{native}$$

$$\sum_{j \in R_{heterologous}} y_j \leq K_{heterologous}$$

$$\sum_{j \in R_{additions}} y_j \leq K_{additions}$$

$$\max_{v} v_{biomass} - \epsilon v_{target}$$

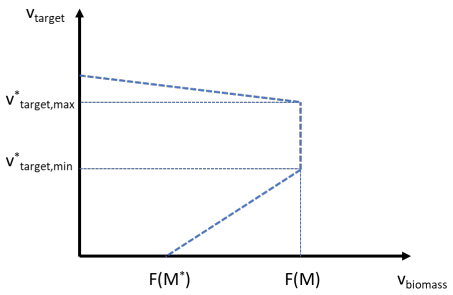
$$Sv = 0$$

$$yv_{min} \leq v \leq yv_{max}$$

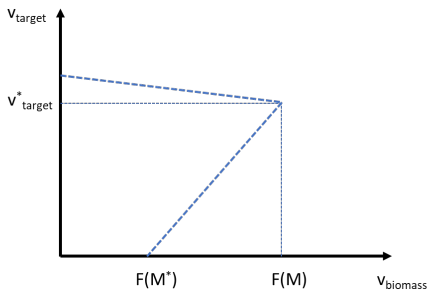
$$\tilde{v}_{target} \leq \bar{v}_{target}$$

$$\epsilon < \epsilon_{max},\ U(\epsilon, y) \subseteq U(0, y)$$

$$v^*_{target}(y^*_{tilt}) \geq v^*_{target}\left(y^*_{couple}\right)$$



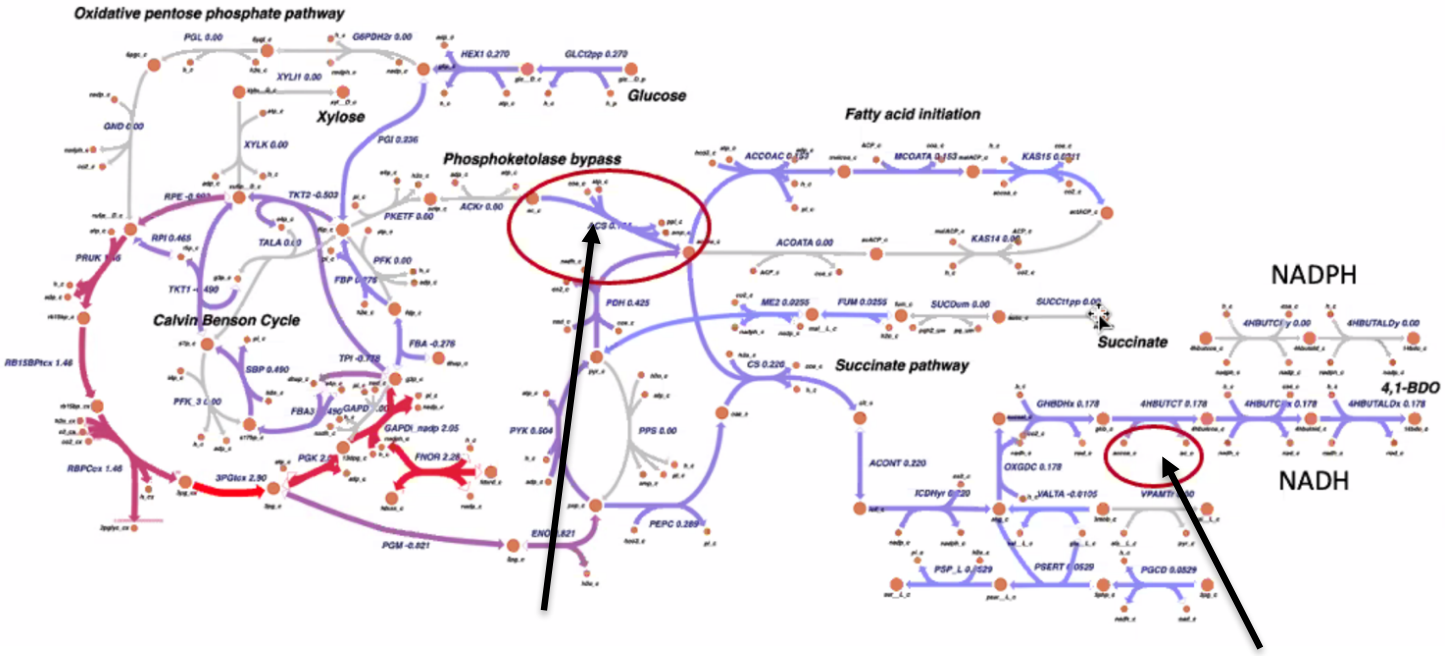
OptCouple



OptTilt

4 – Progress and Outcomes

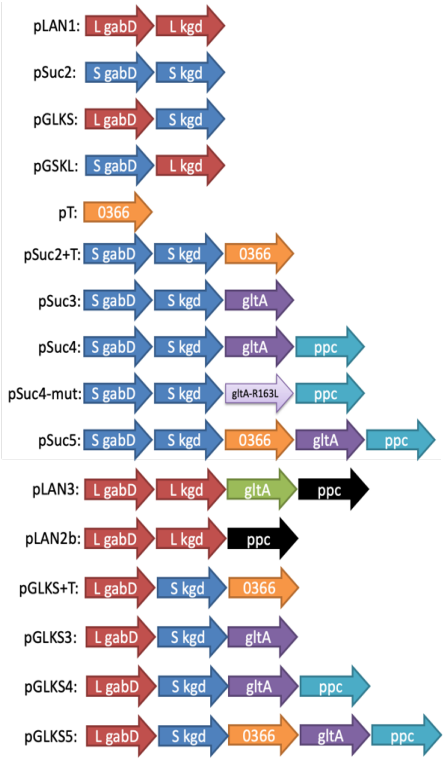
1,4-BDO Metabolic Model in cyanobacteria

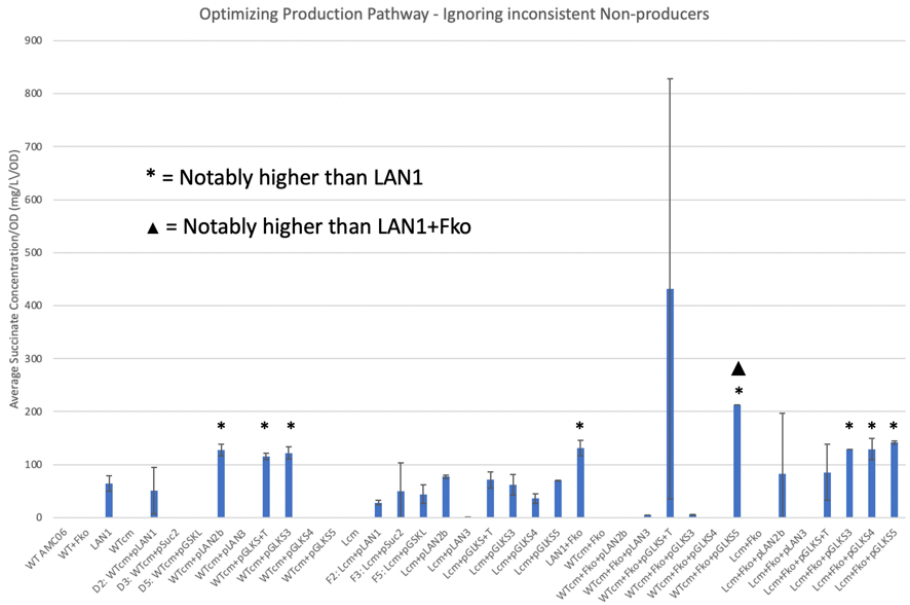


Prediction is to increase acetate production to overcome acetate deficit at 2nd step in 1,4-BDO synthesis. Overexpression vector for ACS and strain currently being generated in *S. elongatus*.

4 – Progress and Outcomes

Succinate: Improved production vectors for cyanobacteria

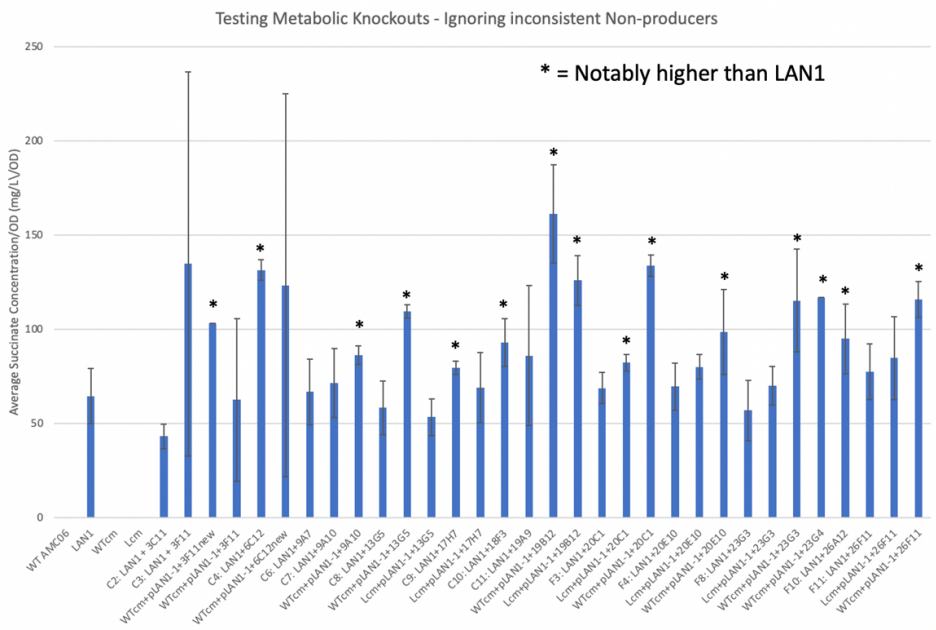




Generated vectors with increased productivity over baseline production, but in a background strain dependent manner. Currently investigating the 5 SNPs that differ between the background strains.

4 – Progress and Outcomes

Metabolic modeled mutants for succinate production in cyanobacteria



Over 30 KOs predicted to increase production. A number of knockouts increased production 2 – 3 fold. Currently combining multiple mutations and vector pathways into a single cell to test combined impacts.

4 – Progress and Outcomes

Phototrophic scaled production of succinic acid in cyanobacteria

LAN1 (baseline succinate production strain) grown at 20 L scale under continuous light conditions indoors

Succinate yields of:

- 178.8 mg/L
- or 336 mg/L/OD
- 2.7x improvement over smaller scale

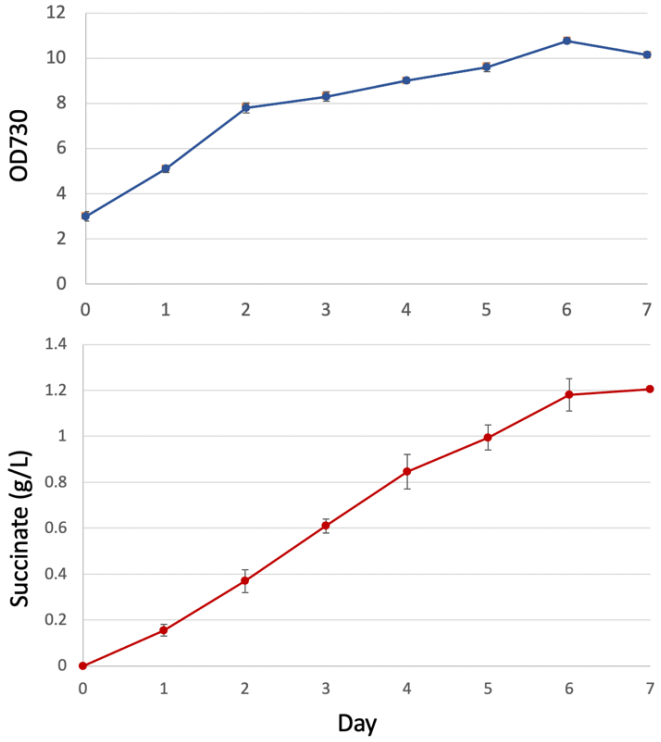
LAN1+sucD-KO grown at 20 L scale in outdoor greenhouse with natural light

Succinate yields of:

- 161 mg/L
- or 326 mg/L/OD
- Likely same as 2x less productive strain due to half the amount of light time

4 – Progress and Outcomes

Heterotrophic production succinic acid in cyanobacteria



Strain: *S. elongatus* PCC 7942 that can consume glucose and produce succinate (NSI pTrc: *galP-zwf-gnd* spec^R, NSIII pLlacO1: *gabD-kgd-ppc-gltA* gen^R)

Light: 30 mmol photons/m^2/s in the PAR range

Temperature: 30˚C
Production media: BG-11 + 50 mM $NaHCO_3$ + 10 g/L glucose

Every 24 h, 10% of the culture volume was removed, the pH was adjusted to 7.0 with 3.6N HCl and volume was replaced with production media containing 200 mM $NaHCO_3$.

Over 1 g/L yield reached in 5 days = Go/NoGo milestone reached

4 – Progress and Outcomes

Screening on glucose/xylose and complex corn stover hydrolysate substrate

		Mixotrophic on glucose	Heterotrophic on glucose	Indication of inhibition on xylose	Indication of inhibition on CSH
Imperial Valley isolates	IV-006	✓	✓		
	IV-031	✓	✓		
	IV-033	✓	✓		
	IV-055	✓	✓		
	IV-112	✓			✓
	IV-113	✓	✓		
	IV-118	✓	✓		
	IV-131	✓	✓		
	IV-132	✓	✓		
	IV-139	✓	✓		
	IV-157	✓	✓	✓	
	IV-233	✓			
	IV-238	✓	✓		
	IV-241	✓	✓	✓	
Chlamydomonas asymmetrica	CCAP 11/41	✓	✓		
	UTEX 227	✓	✓		
	NIES 2207	✓			
Chlamydomonas debaryana	UTEX 231	✓	✓	✓	✓
Chlamydomonas pseudagloe	CPA WT	✓	✓		
Chlamydomonas pseudococcum	CPC	✓	✓	✓	✓
Chlorella vulgaris	CV 25	✓	✓	✓	
Desmodesmus armatus	DA 25	✓			
(Chlorella sp.)	WHIT GREENS 7	✓	✓	✓	
Parachlorella kessleri	PK 25	✓	✓		

Screening on pure glucose/xylose:

- 54 strains screened on CSH components – separate glucose and xylose agar plates at 1% concentration
- 24 strains found to be capable to grow on glucose mixotrophically/heterotrophically
- No strain found to be capable to grow on xylose as sole carbon source; inhibition observed for a few strains

Screening on CSH (NREL) substrate:

- Mixotrophic/heterotrophic strains on glucose tested on CSH
- Goal: Confirm whether there is any toxicity visually observed on CSH (in a concentration equivalent to 1% glucose) compared to growth on pure 1% glucose
- No apparent toxicity observed for most of the strains

4 – Progress and Outcomes

Next Steps

- Promoter DBLT:
  - PNNL: Analyze data from the five replicate transformations and generate 2nd library
  - LBNL: Synthesize & clone 2nd library
  - UCSD: Transform 2nd library into C. reinhardtii and assess promoter strengths using FACS and NGS.
  - If budget allows, perform a third round of DBLT
- Metabolic Modeling:
  - UCSD & UC Davis: Optimize & scale production strains and provide yield & omics data for production to PNNL
  - PNNL: Adjust models based on data from UCSD & UC Davis; predict heterologous, knock down, or over-expression alternatives for boosting production in both S. elongatus and C. reinhardtii
- Pilot scale production:
  - ➢ UCSD down select, optimize, and scale production strains and provide yield and omics data for TEA/LCA to UCD

Summary

- We set out to develop a scalable, high-yield algae production platforms capable of producing polyurethane precursors
- This was a collaboration between academic labs working on algae biotechnology (UCSD, UCD, GT) and the ABF labs of PNNL (computation) and LBNL (DNA foundry)

- At first, there were delays due to paperwork and communication glitches, but we eventually learned each other's languages and capabilities, and are moving forward together at full speed now
- We have had some good early successes on both creating synthetic promoters for algae, and metabolic modeling of succinic acid production in cyanobacteria
- We have achieved our Go/NoGo milestone of 1 gm/L succinic acid
- We have a pretty far way to go to achieve the FOA milestone of 20gm/L … and that is probably still only half of the level required for commercial production – but we are only 18 months into the project ;-)

Quad Chart Overview (Competitive Project)

Timeline

- Project start date 10/1/2019
- Project end date 6/30/2022

	FY20 Costed	Total Award
DOE Funding	(10/01/2019 – 9/30/2020)	(negotiated total federal share)
	$859,559	$2,000,000
Project Cost Share	$301,610	$570,000

Project Partners*

- Partner 1 PNNL, LBNL
- Partner 2 UCD, GT

Project Goal

Develop algae as a production platforms capable of producing polyurethane precursors for the sustainable manufacturing of bio-based, biodegradable, and recyclable foams and plastic products

End of Project Milestone

Determination of yield and cost of heterotrophic versus photosynthetic production of polyurethane precursors in cyanobacteria and algae production systems

Funding Mechanism

DE-FOA-001916 Topic Area 2, Agile BioFoundry Industrial Partnership (ABF) 2018

*Only fill out if applicable.

Additional Slides

Responses to Previous Reviewers' Comments

- This project has not been peer reviewed previously
- Also provide highlights from any Go/No-Go Reviews

Publications, Patents, Presentations, Awards, and Commercialization

- CRADA signed between UCSD and ABF
- No Publications, Patents, or external Presentation to date