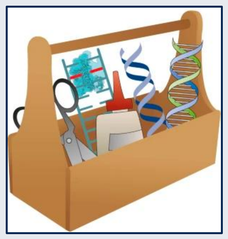
Design of a Phase Dependent Hybrid Promoter Library in *E. coli*

Jamiree Harrison[1], Jai Mehra[2], Annie Nguyen[2], Kevin Jay Chang[2], Enoch Yeung[1,3,4,5]

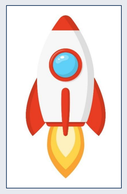
[1]Department of Mechanical Engineering, University of California, Santa Barbara
[2]Molecular, Cellular, and Developmental Biology, UCSB, Santa Barbara
[3]Center for Control, Dynamical Systems, and Computation, UCSB, Santa Barbara, CA
[4]Biomolecular Science and Engineering Graduate Program, UCSB, Santa Barbara, CA
[5]Center for Biological Engineering, UCSB, Santa Barbara, CA

BACKGROUND

Primary Motivation:
Expand toolbox for genetic engineering



Downstream Applications:
Optimize pathways to…

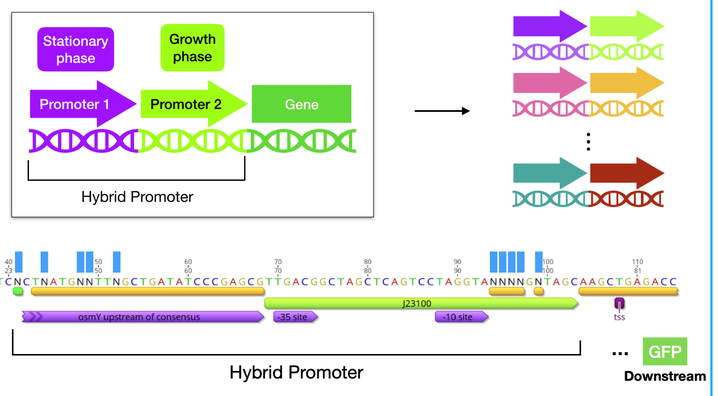


Produce Hydrazine

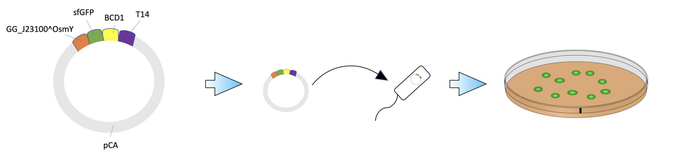
Degrade PET

BUILDING A HYBRID PROMOTER LIBRARY

Design Concept for Promoter Library



Experimental Design



Golden Gate Assembly — Transform into *E. Coli* — Grow and Pick Colonies

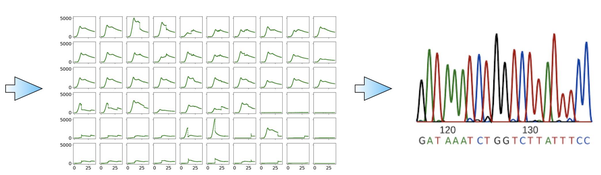


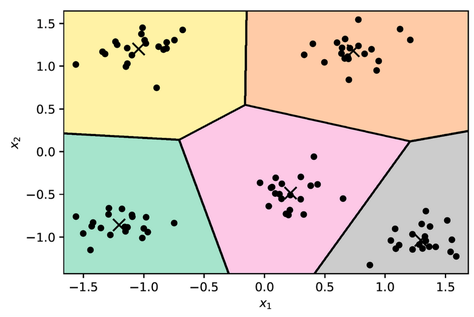
Plate Reader Experiment

Validate Via Sequencing

- Design and implementation of over 100 hybrid promoters which activate gene expression distinctly in both growth and stationary phase

CLUSTERING PHENOTYPE PROFILES

What is K-means?



K-means Phenotype Clustering

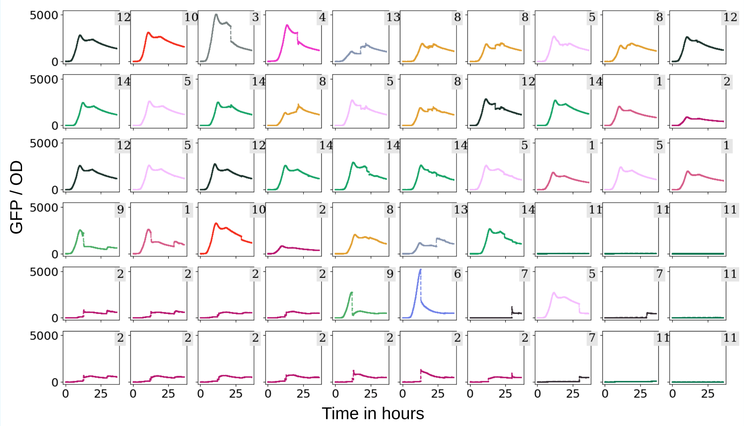


Fig: We cluster the phenotypes using k-means to draw similarities between gene expression profiles. Similar clusters are labeled and color-coded.

MODELING HYBRID PROMOTER BEHAVIOR

Parameter varying Toggle Switch Models

Time-invariant

$$\dot{m} = \alpha_m - \beta_m m$$

$$\dot{p} = \alpha_p m - \beta_p p$$

- $m, p \geq 0$ mRNA and protein concentration
- $\alpha \geq 0$ effective rates of synthesis
- $\beta \geq 0$ decay rates

Time-variant

$$\dot{m} = \alpha_m(t) - \beta_m m \qquad \alpha_m(t) = \begin{cases} \alpha_{m_1} & \text{if } t \in [0,\tau] \\ \alpha_{m_2} & \text{if } t \in (\tau,\infty] \end{cases}$$

$$\dot{p} = \alpha_p(t) m - \beta_p p \qquad \alpha_p(t) = \begin{cases} \alpha_{p_1} & \text{if } t \in [0,\tau] \\ \alpha_{p_2} & \text{if } t \in (\tau,\infty] \end{cases}$$

ALGORITHM FOR PARAMETER FITTING

Model Fitting from Data / System Identification

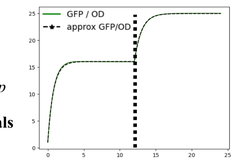
Algorithm Outline:

1. **Instantiate Model:** $m(t) = e^{-\beta_m t}\left(m_0 - \frac{\alpha_m}{\beta_m}\right)$

   $\frac{dX_{pred}}{dt}: \quad \dot{p} = \alpha_p(t)m(t) - \beta_p p$

2. **Detect switches and split data into two intervals**

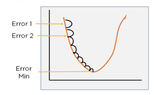
   $X_{data_1} \qquad X_{data_2}$



3. **Minimize error to obtain parameters for the first time interval**

   → **Numerically integrate:** $X_{pred_1}(t) = \int_0^t \frac{dX_{pred_1}}{dt} dt$

   → $\{\alpha_{m_1}, \alpha_{p_1}, \beta_m, \beta_p, m_{0_1}\}$ = **argmin** $\| X_{data_1} - X_{pred_1} \|$

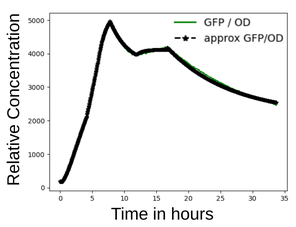


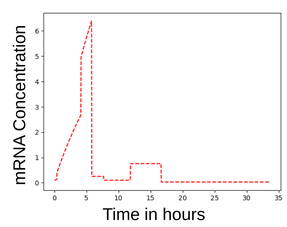
4. **Repeat step 3 for next time interval assuming we know static params**

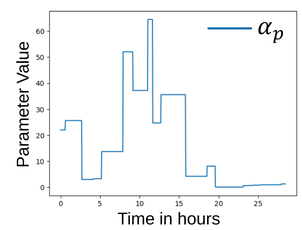
   $\Longrightarrow \{\alpha_{m_2}, \alpha_{p_2}, m_{0_2}\}$ = **argmin** $\| X_{data_2} - X_{pred_2} \|$

5. **Put the obtained parameters in the model and obtain n-step prediction**

Fitting with SciPy Optimization







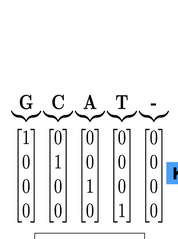
$$m(t) = e^{-\beta_m t}\left(m_0 - \frac{\alpha_m}{\beta_m}\right)$$

$$\dot{p} = \alpha_p(t)m(t) - \beta_p p$$

Thus, we have achieved a fully parametrized hybrid promoter library with ~450 distinct sequences and over 100 phenotypes (clustered groups).

NEXT STEPS

Genotype Clustering and Geno-Pheno Comparison

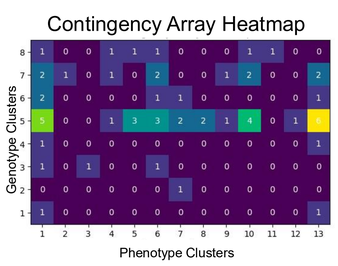


K-means

Cluster 1: Samples 10,3,5

Cluster 2: Samples 9,8,2

Cluster 3: Samples 1,4,6,7



kevinchang@ucsb.edu
jamiree@ucsb.edu

Links to our papers:

Details, references, and acknowledgements found here