Experimental Request: Novel Chassis IcarR/PhlF NAND circuit (1.0), Second Iteration

Primary point of contact: Enoch Yeung, UC Santa Barbara
Modifications:
2018-11-06 -- Request Document Created

...

Goal:

Iterate on the original IcarR/PhlF NAND circuit experiment, with the goal of working around sources of variation and honing in on surprises discovered in the last round.

Subgoals:

- Analyze the day-to-day variability of the experiment by conducting two separate runs of the experiment
- Identify whether the absence of IcarR in the RNAseq is due to biology, sample prep, or analysis.
-

If possible, divide overall goal into more specific subgoals . Connect to the overall goal if not obvious. Order by priority. When possible, add deadlines or estimated timeline for each subgoal. Ex.
1) Determine robustness of circuit function from 37 C by measuring output at the single cell level using FACS.
2) Use RNAseq and targeted proteomics to identify which intermediate circuit elements fail when overall circuit performance degrades
3) (In)validate a predictive model for how relative protein levels of circuit components and the circuit output should change at various temperatures using absolute targeted proteomics and FACS
4) Validate a predictive model that suggests circuit configuration A should lead to levels of protein X and transcriptional activity of promoter Y being less sensitive to temperature versus circuit configuration B
5) Use [differential analytical method of choice] to identify groups of genes/proteins that are affected by two factors: temperature, IPTG level, for circuit X.

Rationale:

Provide a short rationale for the goal (or subgoals). Include references to related experiments/data/literature that are relevant.

Ex: We previously found that Circuit Y performance degrades at slower growth rates. Design of Circuit X employs feedback control to lead to lower sensitivity to perturbations of transcriptional/translation rates of all circuit components and the overall growth rate of the cell (as compared to Circuit Y). Varying temperature allows us to make significant changes in multiple circuit and host parameters and measure ensuing circuit performance. Because it is important that the majority (>80%) of the cell population behaves as we would like, we will measure circuit performance (ON or OFF with respect to some inputs) using flow cytometry.

Experimental and Analytical Approach:

Expected data and analysis:

What's the experimental design? What samples will be compared? Relate to the goal (or subgoal).

Where might it be performed? By who?

Measurement type	**Ginkgo**	**UW Biofab**	**Transcriptic**
RNAseq	All conditions	N/A	N/A
Plate reader time series	None	N/A	N/A
Global, relative Proteomics	All conditions	N/A	N/A
Absolute Proteomics	All conditions	N/A	N/A
Flow	All conditions	N/A	N/A
Genome seq	None	N/A	N/A

What data are you expecting to collect/receive? Is there an ETL process in place and who is responsible? Are there expected results, and if so what are they? State any assertions about the dependence/independence of measurements. If possible, estimate how much data expected.

How will the data be analyzed? By what groups/stakeholders?

How was the number of replicates chosen?

Are there subsequent planned experiments based on outcomes and will they (re)use existing TA3 protocols? What's their expected throughput?

Are there analysis that can be placed in ETL which would automate the sending of future experiments?

Timeline

2018-11-06	Presented analyses of first experiment
2018-11-12	Experiments Start
2018-12-31	Data available on TACC, except absolute proteomics
2019-04-22	Absolute proteomics available on TACC

Experimental protocol:
Does the experiment use an already onboarded TA3 protocol (or is it closely related to an existing protocol)?

The protocol is based on the original NC NAND protocol. This the planning document, including the detailed protocol.

Here is a graphical representation of this experiment.

High-level summary:

1. Strains are grown up in flasks and used to fill 96-well glycerol plates, which are then frozen
2. A plate is thawed and stamped into a plate for overnight growth, where it reaches stationary phase.
3. The OD is taken after the overnight and the plate is stamped into another plate at higher volume to recover from stationary back into exponential growth.

4. The OD is taken after recovery and the plate is stamped into several plates at higher volume with the inducing chemicals; copies are made of each well for each timepoint.
5. At each timepoint, a copy of the plate is harvested for OD, flow cytometry, proteomics, and RNAseq.
    a. In the first run, 3 timepoints were taken (5h, 8h, 18h). In the second run, a 6.5h timepoint was added.

Of these measurements, OD and flow is available right away, while RNAseq and proteomics are generally available within a month.

Strains

10 strains of *E. coli* MG1655 created at MIT.

#	Common Name		Expected Behavior
1	MG1655_WT	Wildtype	negative control, genomic, always OFF
2	MG1655_PhlF_Gate	Genomic	ON without IPTG, OFF with IPTG, genomic (low level)
3	MG1655_IcaR_Gate	Genomic	ON without IPTG, OFF with IPTG, genomic (low level)
4	MG1655_NAND_Circuit	Genomic	OFF with both IPTG and arabinose, ON otherwise, genomic (low level)
5	MG1655_pBADmin	Genomic	OFF without arabinose, ON with arabinose, genomic (low level)
6	MG1655_pTACmin	Genomic	OFF without IPTG, ON with IPTG, genomic (low level)
7	MG1655_pJS007_LALT__backbone	Plasmid	negative control, plasmid, always OFF
8	MG1655_pJS007_LALT__P3__PhlF	Plasmid	ON without IPTG, OFF with IPTG, plasmid (high level)
9	MG1655_pJS007_LALT__I1__IcaRA	Plasmid	ON without IPTG, OFF with IPTG, plasmid (high level)
10	MG1655_empty_landing_pads	Genomic	negative control, genomic, always OFF

Strain 7 is not a good control, bad quality data? Yuval mentioned something

For strain 10, Yuval mentioned that the sensor pad actually contained sensors e.g. yfp etc.

Chemical Conditions

Each strain was induced with 25 mM Arabinose, 0.25 mM IPTG, both, and neither.

Automation of experimental requests and data return:

What metadata needs to be provided where and by who for analysis?

Does an ETL pipeline exist for all analyses? If not, what is the workaround and by who/on what timeline should it be implemented?.

Potential / expected challenges / prioritization :

What are some potential ways the experiment or analysis could fail or fail to provide answers to the goal?

Are their other routes?

Are there some parts / subgoals that should be prioritized over others? Related -- to what degree are some of these subgoals relevant across multiple challenge problems?

Identified stakeholders:

CP WG: Novel chassis
- TA1: Wyss Institute, PNNL, Netrias,
- TA2: MIT
- TA3: Ginkgo

TA2 ingest: SIFT
TA3: Ginkgo
TA4: TACC
DARPA: Susan, Jen, Conrad
OTHER:

Data Upload Progress:

Please use the table below to track progress of data that was uploaded. We are not going to be computer parsing from here, so keep things simple and unstructured. It's just a touch point for folks like me so that we don't lose it in Slack. Please list out the conditions informally (eg. *All plasmid NAND circuit samples* is just fine). We will just ask Qs if we don't get something but this will likely minimize Qs from TA1. This is NOT meant to map files to metadata attributes, so if you are spending more than 1 min per entry, then something is wrong :)

Measurement type	**Ginkgo**
RNAseq	all -- 2018-12-31 -- /sd2e-community/uploads/ginkgo/201811/Novelchassis-Nand-Gate
Plate reader	all -- 2018-12-31 -- /sd2e-community/uploads/ginkgo/201811/Novelchassis-Nand-Gate/attachments/biotek-multi-csv
Global, relative	all -- 2018-12-31 --

Proteomics	/sd2e-community/uploads/ginkgo/201811/Novelchassis-Nand-Gate
Absolute Proteomics	all -- 2019-04-22 -- /sd2e-community/uploads/ginkgo/201901/NovelChassis-NAND-Ecoli-Titration/attachments/proteomics/exp11000*
Flow	all -- 2018-12-31 -- /sd2e-community/uploads/ginkgo/201811/Novelchassis-Nand-Gate/attachments/fcs
Genome seq	N/A