# Design and engineering of native regulatory networks

Emily Freed[1][*] (emily.freed@colorado.edu), Alaksh Choudhury[1], Rongming Liu[1], Liya Liang[1], Adam Arkin[2], Christopher Voigt[3], Carrie Eckert[1,4], and **Ryan Gill**[1]

[1]Renewable and Sustainable Energy Institute, University of Colorado, Boulder CO; [2]Lawrence Berkeley National Laboratory, Berkeley, CA; [3]Massachusetts Institute of Technology, Cambridge, MA; [4]National Renewable Energy Laboratory, Golden, CO

http://www.gillgroup.org/research/

**Project Goals: Our objective is to develop a new standard for the engineering of microbial systems based on rational design, engineering, and optimization of hybrid regulatory networks. We envision a future biorefinery that is based on the development of designer organisms that have exquisite and predictable control architectures governing the expression of a range of valuable traits. Computer aided design platforms will guide the assembly of synthetic constructs containing orthogonal heterologous circuits to recode native regulatory networks. Together, these will enable predictable and dynamic control of multiple designer phenotypes such as: i) growth on various feedstocks in consolidated bioprocesses, ii) feedback control to mitigate accumulation of toxic metabolites, iii) production of target molecules (C3-C4 alcohols), and/or iv) robustness to process upsets (e.g. temp., phage). The focus of this proposal is to develop the technical and computational infrastructure to enable this vision. We will develop this platform first in the model organism *E. coli* and then in DOE relevant non-model organisms.**

Advances in DNA reading and writing technologies are driving adoption of new paradigms for engineering biological systems. The rate limiting step is no longer our ability to construct designer strains, but rather how to design and engineer increasingly complex networks of combinatorial phenotypes required for the economic and sustainable production of biofuels and bioproducts. Regulation plays a special role in the evolution of such combinatorial structures, where mutations affecting the structure and function of regulatory networks are known to account for much of life's diversity. For example, major evolutionary innovations often arise through mutations in regulatory networks and organisms with larger regulatory networks are both more evolvable and more robust. A challenge, however, is that very little is known about the more detailed design rules underlying these observations. This lack of understanding generally undermines efforts to forward design and engineer native regulatory networks, and as such prevents us from taking advantage of such networks in the same way evolution does to go after complex phenotypes.

To address these challenges, we previously developed a technology, CRISPR Enabled Trackable Genome Engineering (CREATE), for constructing and tracking (by barcode) 100,000's of designer mutations in microbial genomes. We used CREATE to construct a regulator library in *E. coli* that contains 170,000 designer mutations spanning 131 *E. coli* regulators. These regulators were sub-divided into the following groups: "global" regulators (50+ regulated genes), regulators involved in core transcriptional process, and "high-level" regulators (20-49 regulated genes). We designed mutations to target known and predicted functionally important sites. The functions include active site, ligand binding sites, protein dimerization interfaces and DNA/RNA binding sites.

We further designed a pipeline that uses bioinformatics mining, protein structural analysis and homology modelling to find sites with the highest likelihood of functional importance. Using the pipeline, we were able to identify ~8800 amino acid residues across the 131 proteins, with an average coverage of ~30% of protein sequence space. We then designed CREATE cassettes to generate site-saturation mutagenesis libraries of these sites. We exposed these libraries to growth selections in several environments relevant to biofuels production, in all cases rapidly identifying mutants with improved fitness. We similarly were able to identify regulators in *S. cerevisiae* that, when mutated, confer a fitness advantage in environments relevant to biofuels production.

In future work, we will expand upon these early studies to now target all regulators in the *E. coli* genome (~300). We expect this will require a library of ~500,000 mutants. In addition to substitution libraries targeting the DNA binding pocket of regulatory proteins, we also have the ability to modify the promoter regions of any gene in the genome by: i) inserting synthetic promoters to turn "on" or "off" downstream genes, ii) inserting DNA binding consensus sequences to introduce novel "on" of "off" native regulation, or (iii) removing DNA binding sequences to remove native regulation. We propose to do so in combination with quantitative measurements of global expression (e.g. RNAseq) and DNA binding patterns (e.g. ChIPseq) to not only decipher how selected regulatory protein mutants affect overall regulatory network function but also to codify this understanding into design criteria for improved engineering of native regulation.

## References

1. Gompel, N., Prud'homme, B., Wittkopp, P. J., Kassner, V. A. & Carroll, S. B. Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in Drosophila. *Nature* **433**, 481–487 (2005).

2. Payne, J. L. & Wagner, A. The robustness and evolvability of transcription factor binding sites. *Science* **343**, 875–877 (2014).

3. Isalan, M. et al. Evolvability and hierarchy in rewired bacterial gene networks. *Nature* **452**, 840–845 (2008).

4. Alper, H., Moxley, J., Nevoigt, E., Fink, G. R. & Stephanopoulos, G. Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science* **314**, 1565–1568 (2006).

5. Garst, A. D. et al. Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nat Biotechnol* **35**, 48–55 (2017).

6. Payne, J. L. & Wagner, A. Mechanisms of mutational robustness in transcriptional regulation. *Front Genet* **6,** 322 (2015).

7. Wagner, A. The molecular origins of evolutionary innovations. *Trends Genet* **27,** 397–410 (2011).

8. Warner, J. R., Reeder, P. J., Karimpour-Fard, A., Woodruff, L. B. A. & Gill, R. T. Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. *Nat Biotechnol* **28**, 856–862 (2010).