Engineered evolvability: *in vivo* optimization of gene networks

Chris N. Takahashi[1], Robert G. Egbert[1], Rahul Brito[1], and Eric Klavins[1]

***Short Abstract* — Directed evolution is a powerful technique to engineer organisms with novel or useful properties, but its utility can be constrained by a limited search in time over a vast mutational landscape that primarily consists of non-beneficial mutations. Here, we have developed a platform to increase the evolvability of gene networks in *E. coli* by engineering with hypermutable DNA sequences that focus mutations to highly tunable genetic elements. We select for these optimized networks in continuous culture environments for which we have created a model to predict the fitness of cell populations to maximize evolvability. To demonstrate the utility of our platform, we have engineered a strain of *E. coli* that is deficient in lactose utilization that rapidly evolves in selective environments. In addition to its application to directed evolution, we believe this platform will be a useful tool to probe natural gene networks to determine optimal environments for which they have evolved.**

***Keywords* — evolvability; continuous culture; simple sequence repeats; optimal gene expression**

I. Background

Engineering increasingly complex gene networks is enabled by our understanding of natural biological processes and advances in DNA synthesis and assembly techniques. However, engineering a functional gene network that meets design specifications often requires a long development period or is never realized. Directed evolution techniques use selective pressure to enrich populations of organisms for variants with high fitness in mutationally diverse pools. Unfortunately, using evolution to optimize networks is often slow, because generally only a small pool of beneficial mutations exists across a vast mutational landscape. One way to create more evolvable gene networks would be to sample beneficial mutations more frequently by focusing variability to sequences that affect phenotype.

Simple sequence repeats (SSRs) - tandem repeats of short nucleotide sequences - are hypermutable DNA sequences known to accelerate the rate of evolution in regulatory and coding regions by generating phenotypic variability across a broad range of organisms [1]. When engineered in the spacer region of the bacterial ribosome binding site, insertion/deletion (indel) mutations to the SSR region result in incremental and predictable changes to translation rates [2]. These repeat sequences can be used to fine-tune gene networks encoded on plasmids or genomic loci. Since SSRs undergo indel mutations at rates 10^4-10^5 higher than arbitrary sequences of the same length, they are ideal genetic tunable elements for optimizing gene networks through directed evolution.

II. Results

A tightly controlled environment is necessary to quantify and control the evolutionary process, and allows us to make simplified modeling assumptions and change selective pressures dynamically [3]. To better quantify experimental results, evolution experiments must be performed in replicate. Multiple experimental replicates not only allow us to find the most common mutations, but also the distribution of evolutionary outcomes. To accomplish this, we have built a multiplexed turbidostat with mixture control that simultaneously controls the optical density and chemical environment of eight continuously growing cultures.

To inform our choice of environment and predict the experimental outcome, we have developed a mathematical model that captures the growth and mutation dynamics of a mixed population that varies only in SSR length. In this model, growth rate is a function of the chosen environment and gene expression, which depends only on SSR length.

To test our platform we have de-tuned the native lactose utilization network in *E. coli* by inserting SSR sequences into the native RBS sequences for the lactose permease (*lacY)* and β-galactosidase (*lacZ*) genes. Our strains are also type I fimbriae knockouts and form biofilms at much lower rates than wild-type cells in the turbidostat. Starting with the suboptimal strain, we experimentally evolved growth phenotypes equivalent to wild-type cells through standard culture in minimal lactose media and optimized growth conditions in our turbidostat. This work has implications both for further elucidating the environments for which the lactose utilization network has been evolved and for rapidly optimizing novel strains for selectable phenotypes.

References

[1] Gemayel R, Vinces M, Legendre M, and Verstrepen K (2010). Variable Tandem Repeats Accelerate Evolution of Coding and Regulatory Sequences. *Annu. Rev. Genet.* **44**: 445-4477.

[2] Egbert, RG and Klavins E (2012). Fine-tuning gene networks with simple sequence repeats. *Submitted*.

[3] Toprak et al. (2012). Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nat. Genetics*. **44**(1): 101-105.

Acknowledgements: This work is supported by NSF Grant #0832773.

[1]Department of Electrical Engineering, University of Washington, Seattle, WA 98195. E-mail: {cnt, robere, britor2, klavins}@uw.edu