Moving Recombineering Beyond *Escherichia Coli*

Christopher Gregg,[1] Xavier Rios,[1] Marc Lajoie,[2] Timothy Wannier,[1]* (timothy_wannier@hms.harvard.edu), Gabriel Filsinger,[1] and **George M. Church**[1]

[1]Harvard Medical School, Department of Genetics, Boston, MA; and [2]University of Washington, Seattle, WA

http://arep.med.harvard.edu (Church lab)

Project Goals: We aim to develop the ability to make multiplexed, high-throughput gene edits in a wide array of microbial species. By improving our mechanistic knowledge of recombineering and developing the ability to screen large libraries of protein variants, we endeavor to greatly expand the reach of Multiplex Automated Genome Engineering (MAGE). Finally, we hope to demonstrate the utility of this platform technology in applications ranging from improving biomolecule yield in Streptomyces to studying photosynthesis in cyanobacteria.

It has been demonstrated in our laboratory that Multiplex Automated Genome Engineering (MAGE) is an easy and efficient mechanism for allowing multiplexed genomic mutations in Escherichia coli.[1] The MAGE process has enabled rapid improvement of metabolic synthesis pathways, genome-wide recoding, and bio-containment applications.[2-4] The protein at the core of MAGE is the λ Red recombinase, λ β (Bet), a phage recombinase, that when expressed improves recombination efficiency of single-stranded DNA oligonucleotides into the bacterial genome ~1E5-fold.[5] Here, we present a platform technology, Serial Evolutionary Enrichment for Recombinases (SEER), that enables the rapid discovery of Bet variants that work well in intractable prokaryotic strains. We queried metagenomic space for Bet homologues, and synthesized a library of putative recombinases with representatives from all families of prokaryotic single stranded annealing proteins. This library (Bet included) was then subjected to successive rounds of enrichment in E. coli for functional recombinases. Surprisingly, we found library members that recombined at or above the frequencies seen for Bet in Escherichia coli, including homologues from Lactobacillus reuteri and Corynebacterium glutamicum.

We further investigated the molecular mechanism of Bet-mediated recombination, and demonstrate an interaction with single-stranded binding protein (SSB). Based on in vitro annealing assays, Bet acts to specifically unload SSB from SSB-coated single-stranded DNA (ssDNA). Removing SSB from ssDNA (like the lagging strand of replication) then enables strand-strand annealing, which is the mechanism by which mutagenic ssDNA is incorporated into the replication fork in Bet-mediated recombineering.[5] Furthermore, we show that an SSB taken from the same host organism as the Bet recombinase homolog can improve the functioning of said Bet homolog in other organisms. The SEER platform will be improved by this knowledge, as Bet fusion proteins will be introduced to take advantage of high recombineering efficiency and organism-specific SSB interaction. High-throughput evolutionary screening will

allow us to test thousands of protein variants in parallel to quickly and robustly identify high-efficiency recombinase mutants for any prokaryotic microbe with the requisite biological toolkit, which includes competency, known plasmids, and inducible promoters.

References

1. Wang, H.H. et al. 2009. Programming cells by multiplex genome engineering and accelerated evolution. *Nature*. 460, 7257 (Aug. 2009), 894–8.

2. Wang, H.H. et al. 2012. Genome-scale promoter engineering by co-selection MAGE. *Nature methods*. 9, 6 (Jun. 2012), 591–3.

3. Lajoie, M.J. et al. 2013. Probing the limits of genetic recoding in essential genes. *Science (New York, N.Y.)*. 342, 6156 (Oct. 2013), 361–3.

4. Lajoie, M.J. et al. 2013. Genomically recoded organisms expand biological functions. *Science (New York, N.Y.)*. 342, 6156 (Oct. 2013), 357–60.

5. Lajoie, M.J. et al. 2012. Manipulating replisome dynamics to enhance lambda Red-mediated multiplex genome engineering. *Nucleic acids research*. 40, 22 (Dec. 2012), e170.

This project was funded by DOE grant DE-FG02-02ER63445.