Precise genome editing in new microbial species using SSAP libraries and broad-host recombineering methods

Gabriel T. Filsinger[1,2] (Filsinger@g.harvard.edu), Timothy M. Wannier[2,3], Felix B. Pedersen[4,b], Isaac D. Lutz[5,6,b], Julie Zhang[7,b], Devon A. Stork[2,8], Anik Debnath[3,9], Kevin Gozzi[10], Helene Kuchwara[3], Verena Volf[2,11], Stan Wang[2,3], Xavier Rios[3], Christopher J. Gregg[3], Marc J. Lajoie[3], Seth L. Shipman[12,13], John Aach[3], Michael T. Laub[10], **George M. Church[2, 3]**

[1]Department of Systems Biology, Harvard Medical School, Boston, MA; [2]Wyss Institute for Biologically Inspired Engineering, Harvard University, Cambridge, MA;[3]Department of Genetics, Harvard Medical School, Boston, MA; [4]Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense M, Denmark; [5]Institute for Protein Design, University of Washington, Seattle, WA; [6]Department of Bioengineering, University of Washington, Seattle, WA; [7]Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA; [8]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA; [9]Tenza Inc., Cambridge, MA; [10]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA; [11]Harvard University John A. Paulson School of Engineering and Applied Sciences, Cambridge, MA; [12]Gladstone Institutes, San Francisco, CA; [13]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA;

http://arep.med.harvard.edu

Project Goals: In this project we aim to develop efficient homologous recombination (HR) methods in diverse microbes. MAGE and recombineering are powerful tools that improve HR efficiency over 1000-fold in *E. coli*. These methods traditionally rely on SSAP protein, λ-Red β, which does not function in most in other species. Our work focuses on two distinct strategies of expanding recombineering to new chassis organisms. First, we studied the mechanism of SSAPs, including λ-Red β, in order to characterize the lack of portability of these proteins. Second, we developed a library of over 200 homologs of λ-Red β in order to screen diverse variants for functionality. With a reliable strategy for developing improved recombineering methods in species beyond *E. coli*, we plan to develop a suite of methods that will enable strain engineering in bacteria with unique capabilities for bioproduction.

Efficient genome editing methods are essential for biotechnology and fundamental research. Homologous recombination (HR) is the most versatile method of genome editing, but techniques that rely on host RecA-mediated pathways are inefficient and laborious. Phage-encoded ssDNA annealing proteins (SSAPs) improve HR 1000-fold above endogenous levels; however, they are not broadly functional. Using *Escherichia coli*, *Lactococcus lactis*, *Mycobacterium smegmatis*, *Lactobacillus rhamnosus*, and *Caulobacter crescentus* we first investigated the limited portability of SSAPs[1]. We find that these proteins specifically recognize the C-terminal tail of the host's single-stranded DNA-binding protein (SSB), and are portable between species only if compatibility with this host domain is maintained. Furthermore, we find that co-expressing SSAPs with a paired SSB can significantly improve activity, in some species enabling SSAP functionality even without host-compatibility. We used the improved portability of SSAP-SSB pairs to expand recombineering methods to new microbes.

In parallel, we designed and built a library of over 200 SSAP homologs in order to identify proteins that enable efficient genome editing across different prokaryotes[2]. We've screened and validated this library across *Escherichia coli*, *Lactococcus lactis*, *Mycobacterium smegmatis*, *Staphylococcus aureus*, *Agrobacterium tumefaciens*, *Corynebacterium glutamicum*, and *Caulobacter crescentus*. In each, we've identified and designed constructs with at least 10-fold higher genome editing efficiency than any previously developed recombineering methods. The library consistently enriches for a set of functional proteins, and provides the most reliable strategy yet developed for developing efficient HR methods in electrocompetent bacteria[3].

Finally, we demonstrate that high-efficiency HR using SSAPs far surpasses the mutational capacity of commonly used random mutagenesis methods, generating exceptional phenotypes inaccessible through sequential nucleotide conversions[1]. Error-prone methods of strain improvement, such as evolution and mutagenesis, do not efficiently generate key combination mutants that contribute to these expectational phenotypes. Additionally, HR focuses mutagenesis at a precise genomic locus, rather than across the entire genome. We demonstrate the use of high-efficiency HR to modulate protein function, but similar techniques could be used to diversify promoters or specifically disrupt target genes.

Publications

1) Filsinger, G. T., Wannier, T. M., Pedersen, F. B., Lutz, I. D., Zhang, J., Stork, D. A., ... & Church, G. M. (2021). Characterizing the portability of phage-encoded homologous recombination proteins. *Nature chemical biology*, 1-9.
2) Wannier, T. M., Nyerges, A., Kuchwara, H. M., Czikkely, M., Balogh, D., Filsinger, G. T., ... & Church, G. M. (2020). Improved bacterial recombineering by parallelized protein discovery. *Proceedings of the National Academy of Sciences*, *117*(24), 13689-13698.
3) Wannier, T. M., Ciaccia, P. N., Ellington, A. D., Filsinger, G. T., Isaacs, F. J., Javanmardi, K., ... & Church, G. M. (2021). Recombineering and MAGE. *Nature Reviews Methods Primers*, *1*(1), 1-24.

This project has been funded by DOE grant DE-FG02-02ER63445. Dr. Church is a founder of companies in which he has related financial interests: ReadCoor; EnEvolv; and 64-x. For a complete list of Dr. Church's financial interests, see also arep.med.harvard.edu/gmc/tech.html.