Title: Serine integrase-assisted genome engineering (SAGE) enables efficient, iterative site-specific genome engineering in bacteria

Authors: Joshua Elmore[1]* (joshua.elmore@pnnl.gov), Gara Dexter[2], Ryan Francis[1], Lauren Riley[2], Jay Huenemann[2], Henri Baldino[1], Adam Guss[2], **Robert Egbert**[1]

Institutions: [1]Pacific Northwest National Laboratory, Richland, WA; [2]Oak Ridge National Laboratory, Oak Ridge, Tennessee

Website: https://genomicscience.energy.gov/research/sfas/pnnlbiosystemsdesign.shtml

Project Goals: The Persistence Control Science Focus Area at PNNL is focused on developing fundamental understanding of factors governing the persistence of engineered microbial functions in rhizosphere environments. From this understanding, we will establish design principles to control the environmental niche of native rhizosphere microbes for the model bioenergy crop sorghum through data-driven genome reduction and engineered metabolic addiction to plant root exudates. These principles will lead to secure plant–microbe biosystems that promote secure, stress-tolerant, and highly productive biomass crops.

Abstract text: Sustainable enhancements to crop productivity and increased resilience to adverse conditions are critical for modern agriculture, and application of plant growth-promoting rhizobacteria (PGPR) is a promising method to achieve these goals. However, many desirable PGPR traits are highly regulated in their native microbe, limited to certain plant rhizospheres, or insufficiently active for agricultural purposes. Synthetic biology can address these limitations, but its application is hampered by the lack of appropriate tools for sophisticated, high-throughput genome engineering for operating outside of the laboratory. Here we present an orthogonal genome engineering system, Serine integrase-assisted Genome Engineering (SAGE), which enables iterative, site-specific integration of up to 10 distinct heterologous DNA constructs (Figure 1). Transformation efficiencies with SAGE are frequently on par with or higher than those observed with common

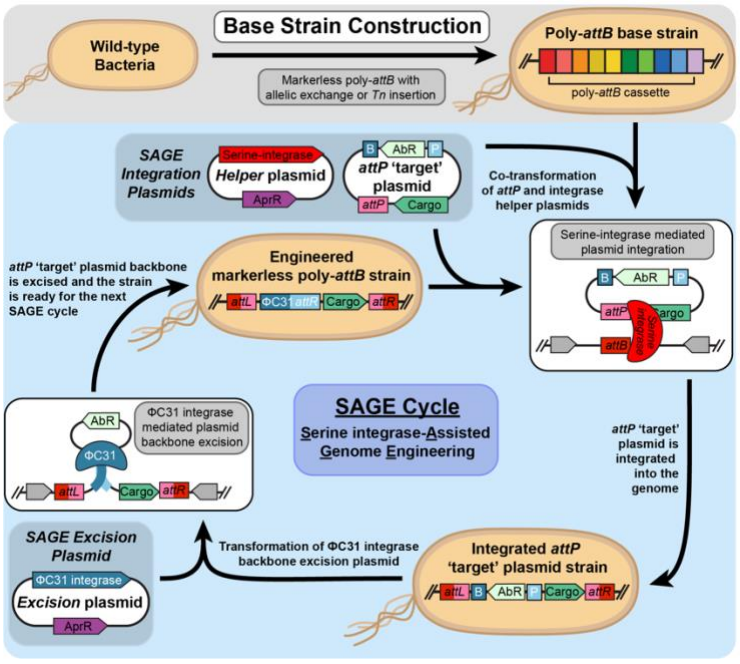

Figure 1. Overview of Serine integrase-assisted genome engineering

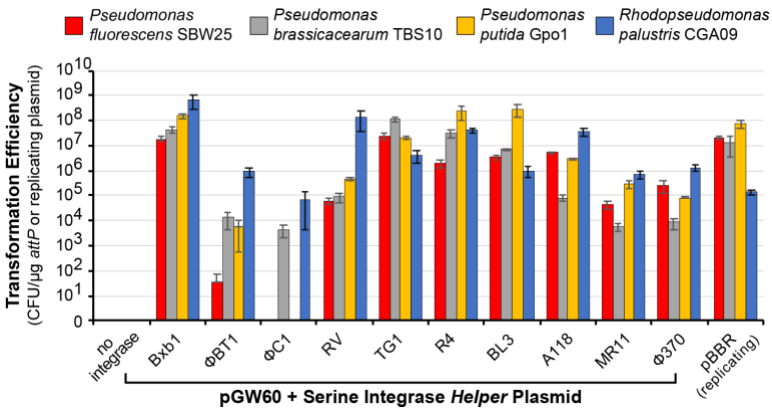

Figure 2. Serine recombinase performance in multiple bacteria.

replicating plasmids (Figure 2) – allowing rapid construction of large genome-integrated combinatorial strain libraries. Unlike other related genome engineering technologies, the machinery for SAGE only requires transient expression from a non-replicating *suicide helper* plasmid for maximal performance. Thus usable, plasmid-free strains are generated in a single step. A secondary serine integrase, ϕC31, can be utilized to excise undesired *E. coli* plasmid backbone DNA from the genome – allowing recycling of antibiotic markers for subsequent genome engineering steps. SAGE performs well in diverse bacteria. Currently, it has been established in a range of Gram-negative bacteria and is actively being developed for Gram-positive bacteria. Here we demonstrate the utility of SAGE by integrating a ~1400-member barcoded constitutive promoter library into 4 hosts. Promoter performance under diverse conditions for each bacteria was assessed using a custom variant of the barcode sequencing method developed by the Wang lab[2]. We identified a collection of *5' UTR insensitive* promoters with a ~40,000-fold dynamic range whose expression was consistent across all tested conditions. We intend to utilize SAGE to understand the genetic determinants of environmental persistence in phylogenetically diverse sorghum bacteria through the development of functional genomics tools and the rapid assessment of root exudate metabolism in heterologous hosts.

References/Publications

1. Elmore, J. R.; Dexter, G. N.; Francis, R.; Riley, L.; Huenemann, J.; Baldino, H.; Guss, A. M.; Egbert, R., The SAGE genetic toolkit enables highly efficient, iterative site-specific genome engineering in bacteria. *bioRxiv* **2020**.
2. Johns, N. I.; Gomes, A. L. C.; Yim, S. S.; Yang, A.; Blazejewski, T.; Smillie, C. S.; Smith, M. B.; Alm, E. J.; Kosuri, S.; Wang, H. H., Metagenomic mining of regulatory elements enables programmable species-selective gene expression. *Nat Methods* **2018,** *15* (5), 323-329.

Funding statement: *This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of BER's Genomic Science Program (GSP), and is a contribution of the Pacific Northwest National Laboratory (PNNL) Secure Biosystems Design Science Focus Area "Persistence Control of Engineered Functions in Complex Soil Microbiomes". A portion of this work was performed in the William R. Wiley Environmental Molecular Sciences Laboratory (EMSL), a national scientific user facility sponsored by BER and located at PNNL. PNNL is a multi-program national laboratory operated by Battelle for the DOE under Contract DE-AC05-76RL01830.*