Title: Phage Factory: Creating a phage for any bacterial species

Authors: Catherine M. Mageeney[1#*] (cmmagee@sandia.gov), Grant Rybnicky[2#], Michael Jewett[2], Kelly P. Williams[1], **Joseph Schoeniger[1]**

Institutions: [1] Sandia National Laboratories, Livermore, CA; [2] Northwestern University, Evanston, IL [#]equal contribution

Project Goals: The overall goal for the Intrinsic Control for Genome and Transcriptome Editing in Communities (InCoGenTEC) project funded under the Secure Biosystems Design initiative is to expand our mechanistic and practical understanding of horizontal gene transfer mechanisms in bacterial communities, and to harness mobile elements to create and deliver constructs to transform, control and detect the genetic and biochemical state of bacteria. Improved ability to engineer genomes of both isolatable and non-isolatable species will enable better scientific understanding of bacterial communities, and facilitate biotechnology applications that promote the growth of the bioeconomy. However, biocontainment and biosafety concerns must be addressed. We are creating modular synthetic genetic elements (SGEs) that can sense, and control altered states of microorganisms. These synthetic genetic elements can be used to transform community members *in situ* provided selective delivery and transformation mechanisms if delivery vectors can be created in an agile manner for new target strains. We have developed software that can identify numerous bacteriophages in near-neighbors of a target strain. These phages can be recovered and used to deliver modular SGEs to target species within a community, even non-isolable members.

Abstract Text:

In natural settings bacteriophages (phages) are extremely diverse, efficient shuttles of genetic information, and can infect all groups of bacteria currently discovered. We have developed a Phage Factory platform to computationally predict prophages precisely (phages integrated into bacterial genomes) and experimentally turn predicted phage sequences into productive virions capable of killing their host, or into phage-vectors, capable of cargo delivery.

We have computationally predicted 185,535 prophages in 207,083 bacterial and archaeal genomes obtained from NCBI, using our TIGER [1] and Islander [2] software. This creates the largest database of phage sequences to date, and at least one host of each phage is known. We have selected to initially focus on *Pseudomonas, Streptomyces*, and *Rhodococcus* species to produce phage using our Phage Factory. Our Phage Factory begins by mining our prophage database for prophages in close relatives of our target strain and annotating these prophage genomes. We experimentally verify the prophages through PCR-based methods and a computational method Juxtaposer [3]. We then create synthetic phage genomes through methods such as long PCR and Gibson assembly, but a key challenge is rebooting, or reconstituting the phage genomes into infectious phage particles [4]. Rebooting has been accomplished through four main methods: transformation, conjugation, L-form transformation, and cell-free protein synthesis. We will present results from initial work on applying the Phage Factory to *Pseudomonas,* a genome with many targets relevant to biomanufacturing or the rhizosphere, and results on developing cell-free phage rebooting methods.

We have identified 14 prophages in six *P. putida* strains of interest, with two clusters and 10 singletons. We verified 9 of these 14 prophages to be active after mitomycin C induction; however, no phage plaques or clearing were visible on any strain tested, indicating these are non-permissive hosts for these 9 active prophages. We further analyzed the genomic relationships

using fastANI to measure the average nucleotide identity (ANI) between our six strains. This revealed these six strains are phylogenetically diverse, with the closest sharing only 98% ANI. This indicated the genetic similarity for permissive prophage hosts is more than 98%.

We are now in the process of creating synthetic genomes for prophages harbored in close relatives of *P. putida* S12. We have two strategies to obtain synthetic phage genomes: 1) obtain the strain of the close relative, long PCR prophage fragments and Gibson Assemble, or 2) Assemble gBlocks for the prophage of interest and Gibson Assemble the fragments.

In order to convert these natural or synthetic phage genomes into infectious virions, we are pursuing cell-free rebooting techniques [5], which we will compare to established approaches such as electroporation. Cell-free protein synthesis uses cell extracts to perform transcription and translation *in vitro*, which can be directly manipulated to optimize rebooting viral genomes. Through screening a variety of biochemical parameters, we have found that adding crowding agents to the reaction and inhibiting exonuclease activity in the extract is essential to producing viable phage particles by cell-free protein synthesis. Likewise, the concentration of phage genome in a cell-free rebooting reaction positively impacts the final titer of phage particles produced beyond the stoichiometric amount that would be expected.

Using this biochemically optimized cell-free protein synthesis reaction, we intend to investigate the host range over which *E. coli* cell-free extracts can reboot phages as well as elucidate the phage-encoded determinants of cell-free rebooting. In determining extract host range, we will apply cell-free rebooting to multiple dsDNA phages per host organism in a phylogenetic walk-out strategy. Starting with *Pseudomonas*, we will evaluate the ability of phages to reboot in cell-free protein synthesis reactions and move further from *E. coli*, eventually to select *Actinobacteria* and *Cyanobacteria*. In parallel, we will test whether doping unoptimized cell-free extracts from the host organism is a viable strategy to extend the host range of *E. coli*-based cell-free rebooting reactions. To determine characteristics of phages that make them amenable to cell-free rebooting, we will assay a wide variety of *Pseudomonas aeruginosa* phages for their ability to reboot in cell-free reactions. Using *P. aeruginosa* phages gives sufficient diversity in genome molecule, genome length, particle morphology, and gene content such that we should be able to draw generalizable patterns from the dataset. Likewise, *P. aeruginosa* phages have been demonstrated to reboot in *E. coli* through electroporation previously. Through these investigations, we will inform the limits of cell-free phage rebooting and anticipate extending the range of bacteria that the Phage Factory can produce active phage against.

References/Publications

[1] C. M. Mageeney, B. Y. Lau, J. M. Wagner, C. M. Hudson, J. S. Schoeniger, R. Krishnakumar, and K. P. Williams, Nucleic Acids Res **48**, 4052 (2020).

[2] C. M. Hudson, B. Y. Lau, and K. P. Williams, Nucleic Acids Res **43**, D48 (2015).

[3] J. S. Schoeniger, C. M. Hudson, Z. W. Bent, A. Sinha, and K. P. Williams, Nucleic Acids Res **44**, 6830 (2016).

[4] C. M. Mageeney, A. Sinha, R. A. Mosesso, D. L. Medlin, B. Y. Lau, A. B. Rokes, T. W. Lane, S. S. Branda, and K. P. Williams, mSystems **5** (2020).

[5] M. Rustad, A. Eastlund, P. Jardine, and V. Noireaux, Synth Biol (Oxf) **3**, ysy002 (2018).

Funding Statement: *This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under the Secure Biosystems Design Initiative. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. SAND number: SAND2022-0152 A.*