From fast to furious phenotyping of microbes: advancing the understanding of microbial diversities and gene functions using droplet microfluidics and high throughput sequencing

F. Song[1,*] (fsong@lbl.gov), X. Lyu[2], M. Moore[2], A. Hung[2], J.V. Kuehl[1], H. Carlson[1], M. Price[1], V. Mutalik[1], A.M. Deutschbauer[1], A.P. Arkin[1,2] and **P.D. Adams[1,2]**

[1]Lawrence Berkeley National Lab, Berkeley; [2]University of California at Berkeley.

http://enigma.lbl.gov

Project Goals: ENIGMA -Ecosystems and Networks Integrated with Genes and Molecular Assemblies use a systems biology approach to understand the interaction between microbial communities and the ecosystems that they inhabit. To link genetic, ecological, and environmental factors to the structure and function of microbial communities, ENIGMA integrates and develops laboratory, field, and computational methods.

Knowing the phenotypes of diverse microbial species and mutations is essential for understanding their ecological roles and genetic functions. Previous methods for microbial phenotyping use fluorescence or optical density (OD_{600}) as the readout and 96 or 384 well plates as the containers. These methods became the standard and have been very popular. Although throughput has been dramatically increased by using lab automation, it does not meet the current demands. Thousands of microorganisms are now routinely isolated by environmental microbiology programs and are in need of basic phenotyping. Large-scale libraries of genetic variations of these, made to characterize and engineer their function, similarly are in dire need of rapid characterization. In addition, the well-plate-based method is expensive- consuming high quantities of disposable plasticware and reagents- and plagued by cross contaminations amongst the wells on the same plate. The contaminants could not be identified by the optical based readout. It is imperative to accelerate the phenotyping speed, decrease the cost and plastic utility, and avoid the effects of cross contamination.

To address problems of scale, we developed a droplet-based high throughput platform for microbial phenotyping using genetic marks as the readout and water-in-oil droplets as containers. The new method consists of droplet generation, droplet cultivation, cell pooling, spike-in controls addition for abundance quantification, DNA extraction, high throughput sequencing, and associated bioinformatic pipelines for data analysis. Briefly, single cells from a mix of hundreds of microbial species (or mutants) are captured in individual water-in-oil droplets. They are allowed to grow within these droplets for some period. The emulsion is then broken, and droplet contents are pooled along with a spike-in standard (a microbe or gene which is not in the pool of microbes). Sequencing of the population before and after this operation enables comparison of the relative abundances of each microbe at the beginning and in the end of the cultivation. In essence, the fitness of hundreds of microbes under the tested condition is obtained by amplicon sequencing in one set-up. Since the system is miniaturized and scalable, we were able to screen many different conditions in parallel using a multichannel syringe/pressure pump. We have demonstrated the efficacy of this method by comparing results from classical plate-based assays to our droplet method for seventeen well characterized organisms under the conditions of R2A and Xylose. We are then applying this technology to physiologically

characterize ENIGMA field isolates under conditions predicted to be important for field processes.

In addition to characterizing microbial isolates, we also combined this method with Random Barcode Transposon-site Sequencing (RB-TnSeq)[1] to study the gene functions of microbes during phage infection. As a proof of principle, we assayed an *E. coli* RB-TnSeq library with two *E. coli* phages (T4 and N4) to discover genes which confer resistance or sensitivity. Our droplet results are consistent with the findings from bulk assays[2] and also uncovered few new hits that were not seen in bulk assays. The droplet assays are useful in identifying medium strength resistance mutants that may be lost in bulk assays because of stronger selection pressure.

The droplet-based high throughput phenotyping method will enable us to link microbial genotypes with phenotypes, and further achieve a mechanistic understanding of field data leading to pointed hypotheses of the role of particular taxa, proteins, co-factors and their interactions in creating systems behavior.

References

1. Wetmore, Kelly M., et al. "Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons." MBio 6.3 (2015): e00306-15.
2. Mutalik, Vivek K., et al. "High-throughput mapping of the phage resistance landscape in E. coli." PLoS biology 18.10 (2020): e3000877.

This material by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies a Science Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231