Construction of a Synthetic 57-Codon *E. coli* Chromosome and Tools for Microbial Genome-Scale Recoding

Akos Nyerges[1*] (akos_nyerges@hms.harvard.edu), Anush Chiappino-Pepe[1,4], Shirui Yan[1,3], Maximilien Baas-Thomas[2], Nili Ostrov[1], Alexandra Rudolph[2], Jenny Ahn[1], Huseyin Tas[1], Yue Shen[3], and **George M. Church**[1,4]

[1]Department of Genetics, Harvard Medical School, Boston, MA; [2]Program in Biological and Biomedical Sciences, Harvard University, Cambridge, MA; [3]BGI Research, Shenzhen, China, [4]Wyss Institute for Biologically Inspired Engineering, Boston, MA

http://arep.med.harvard.edu

Project Goals: We are finalizing the construction of a fully recoded 3.97 Mb *Escherichia coli* genome that relies on the use of only 57 codons. For this aim, the genome was previously computationally designed, synthesized, and assembled into 87 segments. In the final steps of genome construction, we combine and optimize these segments *in vivo* to assemble the fully recoded, viable genome.

We present the construction of a fully recoded, 57-codon *Escherichia coli* genome, in which seven codons are replaced with synonymous alternatives in all protein-coding genes. For this aim, the entirely synthetic recoded genome was assembled as 87 50-kb episomal segments, individually tested for functionality, and then integrated into the genome. Developing a specialized integration system and optimizing our workflow enhanced integration efficiency to 100% and resulted in an order of magnitude increase in construction speed. We are now combining recoded clusters with a novel technology that builds on our latest developments in recombineering and CRISPR-associated nucleases[1,2].

In parallel with genome construction, we developed novel experimental methods to identify fitness-decreasing changes and troubleshoot these cases. Leveraging massively parallel genome editing and accelerated laboratory evolution[3] allowed us to correct partially recoded strains' fitness within days.

As we approach the final assembly of a virus-resistant *E. coli* genome, we also implement dependency on non-standard amino acids and encoding modules for stringent biocontainment and enhanced virus resistance.

In sum, our work will soon I.) demonstrate the first 57-codon organism, II.) establish a tightly biocontained and virus-resistant chassis for new-to-nature protein production, and III.) open a new avenue for the bottom-up synthesis and refactoring of microbial genomes, both computationally and experimentally.

References

1. Wannier TM, Ciaccia PN, Ellington AD, Filsinger GT, Isaacs FJ, Javanmardi K, Jones MA, Kunjapur AM, Nyerges A, Pal C, Schubert MG, Church GM (2021) Recombineering and MAGE. *Nature Reviews Methods Primers*, 1(1):1–24. https://doi.org/10.1038/s43586-020-00006-x
2. Wannier TM, Nyerges A, Kuchwara HM, Czikkely M, Balogh D, Filsinger GT, Borders NC, Gregg CJ, Lajoie MJ, Rios X, Pál C, Church GM (2020) Improved bacterial recombineering by parallelized protein discovery. *Proceedings of the National Academy of Sciences*, https://doi.org/10.1073/pnas.2001588117
3. Nyerges Á, Csörgő B, Draskovits G, Kintses B, Szili P, Ferenc G, Révész T, Ari E, Nagy I, Bálint B, Vásárhelyi BM, Bihari P, Számel M, Balogh D, Papp H, Kalapis D, Papp B, Pál C (2018) Directed evolution of multiple genomic loci allows the prediction of antibiotic resistance. Proceedings of the National Academy of Sciences, 115(25):E5726–E5735. https://doi.org/10.1073/pnas.1801646115

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 64x Bio. For a complete list of Dr. Church's financial interests, see also arep.med.harvard.edu/gmc/tech.html.