Parallelized *in vivo* Construction of a Synthetic 57-Codon *E. coli* Genome

Nili Ostrov[1], Akos Nyerges[1*] (Akos_Nyerges@hms.harvard.edu), Maximilien Baas-Thomas[2], Shirui Yan[1], Alexandra Rudolph[2], Jenny Ahn[1], and **George M. Church**[1,3]

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

http://arep.med.harvard.edu

Project Goals: We are assembling a fully recoded, 3.97 Mb *Escherichia coli* genome, in which seven codons are replaced with synonymous alternatives in all protein-coding genes. For this aim, the full recoded genome was *de novo* synthesized and assembled *in vivo* into 87 segments. In the final steps of genome construction, we combine these 87 segments *in vivo* to assemble the fully recoded genome.

We present the synthesis of a fully recoded, 57-codon *Escherichia coli* genome, in which seven codons are replaced with synonymous alternatives in all protein-coding genes. To this aim, the entirely synthetic recoded genome was assembled into 50 kb episomal segments which were then individually tested for functionality. The genome is constructed by CRISPR/Cas9-mediated *in vivo* recombineering, in which each synthetic segment replaces its corresponding wild-type sequence. Multiplex Automated Genome Engineering (MAGE)[1] and directed evolution with random genomic mutations (DIvERGE)[2] are further used to identify alternative recoding schemes. Replacement efficiency was enhanced up to 100% by implementing a novel, three-plasmid CRISPR/Cas9 knock-in technique. Cycle time was reduced to 11 days by extensively streamlining the replacement procedure and accelerating sequencing-based quality-control steps. Importantly, no significant decrease in growth rate has been observed in eight recoded clusters (total up to 500 kb). In parallel with genome construction, we are optimizing conjugative assembly (CAGE)[3] for combining recoded clusters. As we approach the final assembly of a virus-resistant *E. coli* genome, intermediate strains are also used to implement dependency on non-standard amino acids and encode modules for self-destruction for stringent biocontainment of the final strain. Our work expands the toolkit available for large scale engineering in living cells and opens a new avenue for the bottom-up synthesis and refactoring of organismal genomes.

References

1. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM (2009) Programming cells by multiplex genome engineering and accelerated evolution. Nature, 460(7257):894–898. https://doi.org/10.1038/nature08187
2. 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

3. Isaacs FJ, Carr PA, Wang HH, Lajoie MJ, Sterling B, Kraal L, Tolonen AC, Gianoulis TA, Goodman DB, Reppas NB, Emig CJ, Bang D, Hwang SJ, Jewett MC, Jacobson JM, Church GM (2011) Precise Manipulation of Chromosomes in Vivo Enables Genome-Wide Codon Replacement. Science, 333(6040):348–353. https://doi.org/10.1126/science.1205822

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.