

## Design, Synthesis, and Testing Toward a 57-Codon Genome

Nili Ostrov,<sup>1</sup> Matthieu Landon,<sup>1,2,3</sup> Marc Guell,<sup>1,4</sup> Gleb Kuznetsov,<sup>1,5</sup> Jun Teramoto,<sup>1,6</sup> Natalie Cervantes,<sup>1</sup> Minerva Zhou,<sup>7</sup> Kerry Singh,<sup>7</sup> Michael G. Napolitano,<sup>1,8</sup> Mark Moosburner,<sup>1</sup> Ellen Shrock,<sup>1</sup> Benjamin W. Pruitt,<sup>4</sup> Nicholas Conway,<sup>4</sup> Daniel B. Goodman,<sup>1,4</sup> Cameron L. Gardner,<sup>1</sup> Gary Tyree,<sup>1</sup> Alexandra Gonzales,<sup>1</sup> Barry L. Wanner,<sup>1,9</sup> Julie E. Norville,<sup>1</sup> Marc J. Lajoie,<sup>10</sup> **George M. Church**<sup>1,4\*</sup> ([church@genetics.med.harvard.edu](mailto:church@genetics.med.harvard.edu))

<sup>1</sup>Department of Genetics, Harvard Medical School, Boston, MA; <sup>2</sup>Program in Systems Biology, Harvard University, Cambridge, MA; <sup>3</sup>Ecole des Mines de Paris, Mines Paristech, Paris 75272, France; <sup>4</sup>Wyss Institute for Biologically Inspired Engineering, Boston; <sup>5</sup>Program in Biophysics, Harvard University, Boston; <sup>6</sup>Department of Biological Sciences, Purdue University, West Lafayette, IN; <sup>7</sup>Department of Biological Engineering, MIT, Cambridge, MA; <sup>8</sup>Program in Biological and Biomedical Sciences, Harvard University, Cambridge, MA; <sup>9</sup>Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA.; <sup>10</sup>University of Washington, Seattle, WA

<http://arep.med.harvard.edu>

### Project Goals:

We report progress towards assembly of a 3.97 MB, 57-codon *Escherichia coli* genome in which seven codons were replaced with synonymous alternatives across all protein coding genes. Target codons AGA (Arg), AGG (Arg), AGC (Ser), AGU (Ser), UUG (Leu), UUA (Leu) and UAG (Stop) were chosen by in-house designed software, guided by experimentally validated synonymous recoding rules<sup>1,2</sup>. Thus far, we have completed computational design and de novo synthesis of the recoded genome in 87 ~50kb segments, as well as functional testing of 85% of all recoded genes. We developed a robust pipeline for design, assembly, testing and troubleshooting of recoded segments, with only 25 of 3031 tested genes showing lethal effects. Here, we present our pipeline for the final genome construction step, combining multiple recoded segments into a single strain in a hierarchical fashion. This work underscores the feasibility of rewriting genomes and establishes a framework for assembly of synthetic organisms.

### Publications

1. Ostrov et al. (2016) Science 353(6301). doi: 10.1126/science.aaf3639.
2. Napolitano et al. (2016) PNAS 113(38). doi: 10.1073/pnas.1605856113

*Funding for this work was provided by U.S Department of Energy grant DE-FG02-02ER63445 and Defense Advanced Research Projects Agency grant BAA-12-64. B.L.W. was supported by NSF grant 106394; G.K. was supported by DOD NDSEG Fellowship; D.B.G. was supported by NSF Graduate Research Fellowship; E.S. was supported by the Origins of Life Initiative at Harvard University.*