Cite as: J. D. Boeke *et al.*, *Science* 10.1126/science.aaf6850 (2016).

# The Genome Project–Write

**Jef D. Boeke,*† George Church, * Andrew Hessel, * Nancy J. Kelley,* Adam Arkin, Yizhi Cai, Rob Carlson, Aravinda Chakravarti, Virginia W. Cornish, Liam Holt, Farren J. Isaacs, Todd Kuiken, Marc Lajoie, Tracy Lessor, Jeantine Lunshof, Matthew T. Maurano, Leslie A. Mitchell, Jasper Rine, Susan Rosser, Neville E. Sanjana, Pamela A. Silver, David Valle, Harris Wang, Jeffrey C. Way, Luhan Yang**

*These authors contributed equally to this work.
†Corresponding author. Email: jef.boeke@nyumc.org
The list of author affiliations is available in the supplementary materials.

We need technology and an ethical framework for genome-scale engineering

The Human Genome Project ("HGP-read") nominally completed in 2004 aimed to sequence the human genome and improve technology, cost, and quality of DNA sequencing (*1*, *2*). It was biology's first genome-scale project, and at the time was considered controversial by some. Now it is recognized as one of the great feats of exploration, one that has revolutionized science and medicine.

Although sequencing, analyzing, and editing DNA continue to advance at breakneck pace, the capability to construct DNA sequences in cells is mostly limited to a small number of short segments, restricting the ability to manipulate and understand biological systems. Further understanding of genetic blueprints could come from construction of large, gigabase (Gb)–sized animal and plant genomes, including the human genome, which would in turn drive development of tools and methods to facilitate large-scale synthesis and editing of genomes. To this end, we propose the Human Genome Project–Write (HGP-write).

## Responsible innovation

Genome synthesis is a logical extension of the genetic engineering tools that have been used safely within the biotech industry for ~40 years and have provided important societal benefits. However, recent technological advancements—e.g., standardized gene parts, whole-genome synthesis, and clustered regularly interspaced short palindromic repeats (CRISPR)–Cas9 genome editing technology (*3*, *4*)—are revolutionizing the field (*5*). Some applications are controversial; human germline editing in particular has raised intense moral debate (*6*). As human genome-scale synthesis appears increasingly feasible, a coordinated scientific effort to understand, discuss, and apply large-genome engineering technologies is timely. HGP-write will require public involvement and consideration of ethical, legal, and social implications (ELSI) from the start. Responsible innovation requires more than ELSI, though, and involves identifying common goals important to scientists and the wider public through timely and detailed consultation among diverse stakeholders.

We will enable broad public discourse on HGP-write; having such conversations well in advance of project implementation will guide emerging capabilities in science and contribute to societal decision-making. Through open and ongoing dialogue, common goals can be identified. Informed consent must take local and regional values into account and enable true decision-making on particularly sensitive use of cells and DNA from certain sources. Finally, the highest biosafety standards should guide project work, and safety for lab workers, research participants, and ecosystems should pervade the design process. A priority will be cost reduction of both genome engineering and testing tools to aid in equitable distribution of benefits—e.g., enabling research on crop plants and infectious agents and vectors in developing nations.

To ensure responsible innovation and ongoing consideration of ELSI, a percentage of all research funds could be dedicated to these issues, enabling inclusive decision-making on the topics mentioned above (*7*). In addition, there should be equitable distribution of any early and future benefits in view of diverse and pressing needs in different global regions. The broad scope and novelty of the project calls for consideration of appropriate regulation alongside development of the science and societal debates. National and international laws and regulations differ and, as in stem cell research, a major burden of responsibility of setting standards rests with the scientists and their community. Existing stem cell research guidelines (*8*) may serve as a useful template.

## From observation to action

The primary goal of HGP-write is to reduce the costs of engineering and testing large (0.1 to 100 billion base pairs) genomes in cell lines by over 1000-fold within 10 years. This will include whole-genome engineering of human cell lines and other organisms of agricultural and public health significance, or those needed to interpret human biological func-

tions—i.e., gene regulation, genetic diseases, and evolutionary processes.

This goal is necessarily ambitious, since building a human genome at today's prices would cost more than HGP-read (*9*) (see fig. S1). However, an expectation of HGP-write will be to catalyze a sharp price drop as new technology development occurs apace with advancement of the project, as with the cost of DNA sequencing in HGP-read. Small viral (*10*) and bacterial (*11*) genomes synthesized from scratch and organisms with recoded genomes (*12*) derived from large-scale genome editing (*13*) have demonstrated the feasibility and utility of synthetic genomes. By focusing on building the 3 Gb of human DNA, HGP-write would push current conceptual and technical limits by orders of magnitude and deliver important scientific advances.

HGP-write will aim to address a number of human health challenges. Potential applications include growing transplantable human organs; engineering immunity to viruses in cell lines via genome-wide recoding (*12*); engineering cancer resistance into new therapeutic cell lines; and accelerating high-productivity, cost-efficient vaccine and pharmaceutical development using human cells and organoids. The project could encourage broad intellectual property access via patent pooling. Extreme cost-reduction is feasible, as demonstrated by the $1000 genome grant program (*2*) as well as sharing of CRISPR tools from over 80 labs through addgene.org. Furthermore, because DNA synthesis, like sequencing and computation, is foundational technology, HGP-write could also facilitate biological engineering of many organisms, accelerating R&D across a broad spectrum of life sciences and supporting basic R&D of new bio-based therapies, vaccines, materials, energy sources, disease vector control, and nutrition.

**Pilot projects**

Similar to other Gb-scale genomic projects, including HGP-read, ENCODE (which aims to map genome functional elements), and Sc2.0 (which is synthesizing a heavily edited yeast genome) (*14*, *15*), HGP-write would be conducted in phases with explicit goals and metrics. Each of the earlier large-scale projects began with pilot projects focused on a fraction of the genome, typically ~1%. For HGP-write, the pilots should provide resources of immediate value for advanced biomedical research and/or biotech development. Technology development will likely also occur early in the project to propel large-scale genome design and engineering.

A series of pilot projects making use of very long DNA sequences that are nonetheless short of a full genome are anticipated: (i) synthesizing "full" gene loci with accompanying noncoding DNA to help explain still-enigmatic roles of noncoding DNA variants in regulating gene expression, and leading to more comprehensive models for the role of noncoding genetic variation in common human diseases and traits; (ii) constructing specific chromosomes—e.g., chromosome 21—or complex cancer genotypes to more comprehensively model human disease; (iii) producing specialized chromosomes encoding one or several pathways—e.g., all genes needed to make a prototrophic human cell, or pathways to transform the pig genome to make it more amenable as source for human organ transplantation; (iv) a potential transformation of gene therapy, with freedom to deliver many genes and control circuits to improve safety and efficacy, provided delivery challenges can be met. Indeed, many substantial and useful innovations may be realized in such "stepping stone" projects that are short of whole-genome re-engineering but require substantial improvement in synthesis capacity of Mb- to Gb-sized DNA. Both genome-wide and more modest changes could be tested for their impact on, e.g., organoid development and function in vitro, facilitated by ongoing progress in stem cell differentiation and "organ on chip" technologies. Novel cell culture technologies may, in some cases, be many times more cost-effective and accurate than current whole-organism testing.

**Box 1. Some properties of "ultrasafe" cells with a pervasively re-engineered genome.**

Virus resistance—can be conferred by systematically recoding certain codons across all genes. Subsequent deletion of tRNA genes would generate a cell line resistant to viruses.

Improved cancer resistance—tumor suppressor genes could be made multicopy; genes like p53 could be recoded to eliminate CpG dinucleotides that give rise to "hotspot" mutations.

Other useful traits—delete potentially deleterious genes such as prion genes.

Improved genome stability—comprehensively eliminate endogenous repetitive "selfish DNA" elements.

Fail-safe security—prevent formation of germ cells, e.g., by removing transcriptional regulators.

Applications— "go to" cell line for stem cell therapies; robust production of biologics.

Additional pilot projects being considered include (v) using induced pluripotent stem cells (*16*) to construct an "ultrasafe" human cell line via comprehensive recoding of protein-coding regions, and deletion of corresponding genome features to increase safety of such a cell line (see Box 1); and (vi) developing a homozygous reference genome

bearing the most common pan-human allele (or allele ancestral to a given human population) at each position to develop cells powered by "baseline" human genomes. Comparison to this baseline will aid in dissecting complex phenotypes such as disease susceptibility. The pervasive nature of the required changes makes whole- or partial-genome synthesis an efficient route to these goals.

## Project launch and administration

The goal is to launch HGP-write in 2016 with $100 million in committed support, from public, private, philanthropic, industry, and academic sources from around the world. The costs of the project lie not only in obtaining de novo synthesized DNA but in the assembly, integration, and functional assays required to evaluate and understand the modified genomes. Total project costs are difficult to estimate but would likely be less than the $3 billion cost of HGP-read.

HGP-write could be implemented through one or more centers [similar to Centers of Excellence in Genomic Science (CEGS) and the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) initiative centers] that will coordinate and support formation and work of multi-institutional and interdisciplinary research teams working in a highly integrated fashion responsive to and engaged with a broad public outreach.

We celebrate 2016—the 25th anniversary of HGP-read—as a major step forward for human knowledge and health. In this spirit, we look forward to the launch of HGP-write.

## REFERENCES AND NOTES


1. International Human Genome Sequencing Consortium, *Nature* **431**, 931 (2004). Medline doi:10.1038/nature03001
2. C. W. Fuller *et al*., *Nat. Biotechnol.* **27**, 1013 (2009). Medline doi:10.1038/nbt.1585
3. L. Cong *et al*., *Science* **339**, 819 (2013). Medline doi:10.1126/science.1231143
4. P. Mali *et al*., *Science* **339**, 823 (2013). Medline doi:10.1126/science.1232033
5. A. D. Haimovich, P. Muir, F. J. Isaacs, *Nat. Rev. Genet.* **16**, 501 (2015). Medline doi:10.1038/nrg3956
6. D. Baltimore *et al*., *Science* **348**, 36 (2015). Medline doi:10.1126/science.aab1028
7. Presidential Commission for the Study of Bioethical Issues, *New Directions: The Ethics of Synthetic Biology and Emerging Technologies* (Washington, DC, 2010).
8. J. Kimmelman *et al*., *Nature* **533**, 311 (2016). Medline doi:10.1038/533311a
9. P. A. Carr, G. M. Church, *Nat. Biotechnol.* **27**, 1151 (2009). Medline doi:10.1038/nbt.1590
10. K. J. Blight, A. A. Kolykhalov, C. M. Rice, *Science* **290**, 1972 (2000). Medline doi:10.1126/science.290.5498.1972
11. D. G. Gibson *et al*., *Science* **329**, 52 (2010). Medline doi:10.1126/science.1190719
12. M. J. Lajoie *et al*., *Science* **342**, 357 (2013). Medline doi:10.1126/science.1241459
13. F. J. Isaacs *et al*., *Science* **333**, 348 (2011). Medline doi:10.1126/science.1205822
14. N. Annaluru *et al*., *Science* **344**, 55 (2014). Medline doi:10.1126/science.1249252
15. J. S. Dymond *et al*., *Nature* **477**, 471 (2011). Medline doi:10.1038/nature10403
16. K. Takahashi, S. Yamanaka, *Cell* **126**, 663 (2006). Medline doi:10.1016/j.cell.2006.07.024


## ACKNOWLEDGMENTS


This paper is the result of meetings held at NYU Langone Medical Center October 31, 2015 and Harvard Medical School on May 10 2016. F.J.I. is a co-founder of enEvolv, Inc. G.C. has financial relationship with Gen9, Editas, Enevolv, and Egenesis (companies directly related to this article; for a full list of G.C. financial relationships see: arep.med.harvard.edu/gmc/tech.html). J.B. is on the Board of Directors, Neochromosome Inc., owns stock in Recombinetics, Inc., and Sample 6, Inc. A.H. has investments in Autodesk Inc. GC is an inventor on patents and patent applications filed by Harvard Medical School that covers synthesis, assembly and testing of large DNAs. This policy forum is the opinion of the authors and not that of their employers/institutions. The authors gratefully acknowledge the financial support of Autodesk, sponsor of the meetings.


## SUPPLEMENTARY MATERIALS

www.sciencemag.org/cgi/content/full/science.aaf6850/DC1
Author affiliations
Fig. S1


Published online 2 June 2016
10.1126/science.aaf6850

**The Genome Project–Write**

Jef D. Boeke, George Church, Andrew Hessel, Nancy J. Kelley, Adam Arkin, Yizhi Cai, Rob Carlson, Aravinda Chakravarti, Virginia W. Cornish, Liam Holt, Farren J. Isaacs, Todd Kuiken, Marc Lajoie, Tracy Lessor, Jeantine Lunshof, Matthew T. Maurano, Leslie A. Mitchell, Jasper Rine, Susan Rosser, Neville E. Sanjana, Pamela A. Silver, David Valle, Harris Wang, Jeffrey C. Way and Luhan Yang (June 2, 2016)
published online June 2, 2016 originally published online June 2, 2016

Editor's Summary



*Science* (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 