HGP-write: Testing Large Genomes in Cells

Meeting Summary

Jef D. Boeke, George Church, Andrew Hessel, and Nancy J Kelley

MAY 10, 2016
BOSTON, MA

PREPARED BY FEINSTEIN KEAN HEALTHCARE

NYU Langone MEDICAL CENTER

WYSS INSTITUTE

AUTODESK.

NANCY J KELLEY + ASSOCIATES

Table of Contents

Executive Summary

The 25th Anniversary of the launch of the Human Genome Project (HGP-read) occurred on October 1, 2015. Successfully completed in 2003, HGP-read is now widely recognized as one of the great feats of exploration, one that sparked a global revolution in science and medicine, particularly in genomic-based diagnostics and therapeutics. Among the lessons learned from this historic Grand Challenge was the value of consortium-based research focused on the discovery of fundamental information and the value of large projects with daring goals for moving science and knowledge forward.

Francis Collins described the working draft of the human genome as "the first glimpse of our own instruction book." Despite the significant insights gained from large-scale studies designed to interpret this instruction book, including HapMap, Encyclopedia of DNA Elements (ENCODE), and genome-wide association studies (GWAS), our knowledge of the human genome remains far from complete. Thus, the full benefits of such knowledge have not been attained. Many scientists now believe that the way to truly understand our genetic blueprint is to "write" DNA and build human (and other) genomes from scratch. Such an endeavor will require research and development on a grand scale.

On October 31st, 2015, 26 leaders in the fields of genetics and synthetic biology convened at the Institute for Systems Genetics at NYU Langone Medical Center in New York City to conceptualize and discuss the feasibility of a bold quest: synthesizing the human genome (HGP-write). HGP-write represents the next chapter in our understanding of the blueprint of life.

As defined at this initial meeting, HGP-write will be an open, international scientific research project led by a group of scientific leaders from multiple disciplines, including biology, chemistry, computational biology, engineering, social science, and ethics. The primary goal of the project will be to synthesize, or write, human-sized genomes in a cell line within ten years.

A new Grand Challenge, HGP-write is expected to galvanize the scientific community, foster international collaboration, engage a wide network of educational institutions, inspire the next generation of scientists, and accelerate research and development across diverse areas. The ability to efficiently write DNA could pull us forward into a better future.

On May 10th, 2016, more than 130 international scientists, industry leaders, ethicists and policy makers steeped in human biology, health and synthetic biology came together at Harvard Medical School in Boston for an expanded organizing meeting of HGP-write. This meeting focused on the design and technical issues, ethical and social issues, and industry involvement of HGP-write. Although not every perspective on every topic discussed throughout the day could be captured, the most common themes discussed are represented in this meeting summary.

The October and May meetings are part of a series of scientific discussions amongst the community that have been unfolding over the past several years. Initially, these meetings were focused on the synthesis and testing of yeast and bacterial genomes, the future of synthetic biology, and more recently, the synthesis of variations within large genomes (on the scale of the 3 billion base pair human genome) as a pathway to addressing some of the many global challenges facing humanity, including healthcare and the environment.

The original intent of the May 10th meeting was to be highly open and transparent in order to catalyze broad community discussion, including media presence on-site, a live video feed, a web site, and real-time social media dialogue. These plans had to be changed because the associated commentary was undergoing peer review at the time, and the publication's guidelines strictly limited what could be discussed publicly.
(http://bit.ly/2dWR7nZ)

This meeting summary, meeting video, HGP-write white paper, as well as additional background materials on the project can be found here:
http://engineeringbiologycenter.org/resources/

We invite you to get involved in this project and to take part in the conversation.

Jef D. Boeke, Ph.D.
Director, Institute for Systems Genetics
Professor, Department of Biochemistry
and Molecular Pharmacology
NYU Langone Medical Center

George Church, Ph.D.
Professor of Genetics,
Harvard Medical School
Core Faculty Member,
Wyss Institute for Biologically Inspired E
ngineering at Harvard University
Professor of Health Sciences and Technology,
Harvard and the Massachusetts Institute
of Technology (MIT)
Associate Faculty Member,
Broad Institute

Andrew Hessel
Distinguished Researcher,
Bio/Nano Research Group
Autodesk

Nancy J Kelley, J.D., M.P.P.
President & CEO
Nancy J Kelley & Associates
Founding Executive Director,
New York Genome Center

Overview of Project and Current Synthetic Genomics Environment

This opening session provided an overview of HGP-write with a focus on responsible innovation and the benefits to science and humanity, as well as lessons learned from related efforts. HGP-write will be an open, international research project led by a multi-disciplinary group of scientific leaders who will oversee the synthesis, or "writing", of human-sized genomes in a cell line within ten years. The overarching goal of such a project is to understand the blueprint for life provided by the Human Genome Project (HGP-read). Such work will include whole-genome engineering of human and other cells necessary for interpreting and understanding human biological functions, such as gene regulation, genetic diseases and evolutionary processes.

The idea of sequencing the human genome (HGP-read) was considered radical in the 1980's, and the project generated vigorous debate about whether it should proceed prior to its launch in 1990. Yet HGP-read has dramatically advanced our knowledge base and continues to pay dividends today. The project has led us to the entire field of genomic-based diagnostics and therapeutics, and tremendous improvements in the treatment of cancer and other diseases. HGP-read also successfully placed science in the public view, launched careers, and inspired scientists to work in the area of genomics. Moreover, the project greatly advanced technology, reducing the cost of sequencing a human genome from $3 billion to less than $1,000 in ~8 years (between 2007 and 2015).

HGP-write will build on the knowledge and technological advances of HGP-read by utilizing genome-engineering technologies to responsibly synthesize human and other large genomes. Responsible innovation has been at the forefront in the consideration of the design of HGP-write, and will continue to remain at the forefront as the project plan advances. ***For example, HGP-write will be limited to cell lines only.***

The ability to efficiently write DNA could pull us forward into a better future, particularly with respect to human health challenges. Some potential applications include growing transplantable human organs, engineering universal T cells, revolutionizing gene therapy, and engineering virus, cancer and aging resistance into cell lines used for therapy. It can help us evaluate millions of "variants of unknown significance" arising from various genomes read and growing genomic diagnostics. Additionally, DNA synthesis and assembly are foundational technologies, which could accelerate research and development across a broad spectrum of areas, including new bio-based therapies, vaccines, materials, energy sources, disease vector control, and nutrition.

Sc2.0, the Synthetic Yeast Genome Project, and rE.coli, the design and construction of recoded E.coli genomes, are related whole genome engineering efforts that can offer lessons learned for HGP-write. A well-defined construct will need to be in place prior to launch of the project, which all participants must agree upon, including funding, space, personnel, QA/standards, material transfers, publication policy, intellectual property, software, ownership of the project, training and education, and compliance with local laws. Additionally, because HGP-write will involve global participation, it will be important

to maintain a collaborative and inclusive culture across international borders. The Sc2.0 consortium may serve as an excellent model for this.

Discussion of Pilot Projects

Similar to other large-scale genome projects, including HGP-read, Encyclopedia of DNA Elements (ENCODE), the RE.coli and the Synthetic Yeast Project (Sc2.0), HGP-write will be conducted in phases with explicit milestones, metrics, and assessments. Each of these earlier projects began with a pilot project that focused on a fraction of the genome, typically about 1%. For HGP-write, the pilot projects chosen will provide resources valuable for advanced biomedical research and/or biotechnology development.

This session described an initial set of proposed pilot projects for HGP-write. Whole genome design and engineering is central to each effort aimed at engineering "ultrasafe" properties into human cells, the synthesis of a prototrophic human genome, targeted efforts for cell and tissue therapies and the design and construction of E.coli with a radically new genetic code containing 57 codons. The diversity of these pilot projects reflects the breadth of HGP-write to include genome design and synthesis in human and model organisms.

Ultrasafe Cell Line

There is an unmet need for an "ultrasafe" human cell line designed to serve as a platform for many biomedical applications, from production of biologics, to modeling cell and tissue behaviors, to actual *ex vivo* and ultimately *in vivo* therapeutic applications. Such a cell line will be engineered to be ultrasafe from many distinct perspectives, including viral, prion and cancer resistance, as well as being engineered to minimize immune rejection, and senescence. This cell line would potentially be of great value to the pharmaceutical, vaccine, and biotechnology industries.

It is anticipated that many of the details of this cell line remain to be extensively discussed and the plans for exactly how to construct it will be jointly worked out by a group of human genetics and genomics experts as well as synthetic genomicists.

Synthesizing a Prototrophic Human Genome in Cell Lines

Humans are metabolically incapable of biosynthesizing 9 amino acids and at least 10 vitamins essential for life, all of which must be derived from diet. These deficiencies result in chronic malnutrition and food shortage in some populations, which may potentially be addressed using synthetic genomic approaches. The biosynthesis of essential amino acids and vitamins are generally well characterized and the genes and pathways responsible for them are known. This pilot project proposes to introduce these biosynthetic pathways into the human genome (and microbiome) to enable these cells to be less dependent on, or completely independent of, exogenous supplementation of these metabolites. In addition, this project could be of utility for understanding the biochemical milieu needed for mammalian organismic development, cell differentiation, and nutrition-associated aging

processes. Since current mammalian production cell lines (e.g. CHO cells) require expensive growth medium, one could imagine using these prototrophic cells with cheaper growth medium formulations at larger bio-production scales, leading to more economical biosynthesis of various drugs and biologics that require mammalian cell lines.

The 7 Signals Toolbox

The 7 signals toolbox proposal involves developing a synthetic approach to control the differentiation of a stem cell into any human cell-type or tissue. The diversity of cell and organ structures in the human body are largely generated from the action of just seven conserved developmental signals: the Hedgehog, Wnt, TGF-β, receptor tyrosine kinase, Notch, JAK/STAT and nuclear hormone pathways. These pathways are activated in many combinations and temporal patterns to drive the differentiation of stem cells into hundreds of cell types during embryogenesis. These pathways all have a tree-like structure, with many 'branches' accepting input information that is funneled to a core 'trunk' that transduces the signal, which is then relayed into many 'roots' leading to hundreds of transcriptional and cell biological changes. Synthetic methods will be developed to perturb the seven major developmental signaling pathways at their conserved core: the trunk of the tree. A genetic toolbox will be developed for the precise temporal modulation of the seven pathways at their 'trunk' in any combination using optogenetic and chemical controls. In addition, a complimentary toolbox of reporters will be built for the rapid phenotypic profiling of these cells, thereby enabling the mapping of signaling to phenotype in high-throughput.

This project will create a map for the artificial control of development and generate tools that facilitate the control cell of differentiation *in vitro*. Ultimately, a precise control of cell differentiation is a crucial step toward cell therapies, tissue replacement and even organ transplants either with patient-derived induced pluripotent stem cells (iPSCs) or an HGP-write "ultrasafe" cell line.

Recoding E.coli to 57 Codons

HGP-write provides an opportunity to synthesize a variety of highly modified genomes. Through the genome-wide elimination of a codon and the tRNA responsible for its translation, the creation of a recoded genome provides an opportunity for genetic isolation and new chemistry. Seven codons were synonymously replaced in all protein coding genes in the E.coli genome, which required the replacement of more than 62,000 codons. To date, 63% of all genes have been recoded and the genome appears amenable to radical changes in its genetic code as measured by cell viability and gene expression. *De novo* synthesis enables large-scale genome recoding, but it also requires robust computational design for successful codon replacement as well as parallelized experimental testing for both striking and subtle growth effects.

HGP-write welcomes new proposals for the pilot projects. Please download the Pilot Proposal form from the HGP-write website, and send proposals to Nancy J Kelley (info@engineeringbiologycenter.org).

Related Research Areas that Could Complement or Benefit HGP-write

Recognizing that HGP-write could have broad impact in other research endeavors, this session presented a series of topics that covered far-ranging fields, including constructing human organoids, organ transplantation, gene therapy and microbiome engineering. These research areas represent mature and emerging disciplines uniquely positioned to benefit from the advances in technologies and designer genomes that will emerge from HGP-write.

Human Organoids

The top down approach to producing human tissue involves recapitulating development through organoids. A significant challenge to this approach, however, lies in finding the right 'recipes' to make the thousands of human cell types from induced pluripotent stem cells (iPSCs). By screening the "human TFome," which includes 1,576 transcription factors, a shotgun approach was used to differentiate human iPSCs into many different cell types. Using this approach, approximately 80 new recipes were found (no combinations), including ones that could produce both homogenous human neurons and human blood vessel cells within four days. A "cell-chip" could enable rapid testing of genome function in diverse human cells. Another challenge with the top down approach is that a lack of vasculature within organoids can cause necrosis. However, organoid-editing technologies could add vasculature to these structures, thereby resulting in organ-sized organoids.

Transplantation

Organ transplantation remains an area of tremendous unmet need, particularly for patients with end-stage organ failure. Although pigs can serve as natural bioreactors for life-saving transplantation, there are significant challenges associated with this approach, including rejection of pig organs and endogenous porcine retrovirus (PERV) transmission. Xenotransplantation could become a clinical reality by using mammalian genome engineering tools and technologies to create modified pig cells that are compatible with the human immune system and free of PERVs. To achieve this, all 62 copies of PERVs are being eradicated and human immunological compatibility is being engineered into the pig genome. Ultimately, every cell within Pig2.0 will have a modified genome so that multiple tissue and organ types can be used.

Human Artificial Chromosomes

Human artificial chromosomes (HACs) are extra-mini chromosomes, ranging from 2-10 Mb, that are pure kinetochore, contained as a single copy per cell, are mitotically stable, and never integrate into chromosomes. HAC-based vectors also have unlimited cloning capacity. A HAC carrying a Tet-O protein recognition site (tetO-HAC) has been created, which can be used for a number of applications, including studying the organization and function of human kinetochores as a gene delivery vector, for measuring chromosome instability (CIN), and for a number of uses within synthetic biology, including the assembly of large genomic regions. The tetO-HAC could be used as a platform for HGP-write.

De Novo Chromosomes that Shuffle Between Yeast and Mammalian Systems

Human artificial chromosomes (HACs) are scored by segregation, compatibility, organization, replication, expression, and shuttle-ability. Building *de novo* HACs that are well defined and easy to manipulate is at the cutting edge of genome synthesis and assembly. A HAC containing a yeast-bacteria shuttle vector has been designed that contains selectable markers, an origin of replication, a self-perpetuating centromere, insulators for expression and chromatin structure regulation, and human telomeres. This HAC uses yeast to efficiently deliver megabases of assembled DNA to mammalian cells.

Microbiome Engineering

The human body is shared with trillions of commensal microbes that are involved in development, immunity, digestion, and mood/behavior. There exists a need for new tools and technologies that can modulate the human microbiome, for example, by engineering probiotics to detect colon cancer and provide local treatment. Such a technology would involve outfitting bacterial and human cells with sensors and controllers. But our current understanding of how gene regulatory events translate to biological function over time and space, especially *in situ*, is limited by existing technologies.

An improved ability to engineer the microbiome could open many opportunities for HGP-write, including the following:

- Domesticate all human-associated microbes to help decipher their functions
- Culture human-associated microbes and study their interactions in *in vitro* (organ-on-a-chip) and *in vivo* models of human physiology
- Intimate interactions between microbes and host to enable new interfaces for diagnostic and therapeutic intervention
- Optimize genetic memory and computing architectures in bacteria that are conceptually ported to human cells to enable *in situ* sensing, recording, and responses
- Unravel the mapping of combinatorial/high-order/complex genetic networks to biological phenotypes

Gene Therapy

In Duchenne Muscular Dystrophy (DMD), frame-disrupting mutations lead to loss of dystrophin protein. If these mutations could be re-coded to restore the reading frame, patients could be switched from a severe to mild form of the disease. A sensitive *in vivo* reporter system for CRISPR activity, AAV-CRISPR, was developed and used to simultaneously target multiple organs of therapeutic interest, which irreversibly corrected the *Dmd* reading frame and restored functional dystrophin protein expression in dystrophic mice.

Haploid Embryonic Stem Cells

Haploid genetics greatly facilitates functional genomics due to the absence of complementation. Haploid human embryonic stem cells (hESC) can be used for genetic modification. A haploid genome can be retained in neural progenitor cells and neurons

(ectoderm), cardiomyocytes (mesoderm), and pancreatic cells (endoderm), as well as in differentiated tissues *in vivo* (teratomas). The strength of the haploid ES cell system is stringent molecular and functional testing of genetic changes in any cell type of interest through differentiation. However, the National Institutes of Health (NIH) cannot support such a project due to restrictions resulting from the Dickey-Wicker amendment.

Is "HGP-write: Testing Large Genomes in Cells" a Wise Choice?

HGP-write will require significant consideration of the social, ethical, and legal implications from the start of the project. Specifically, responsible innovation in the form of deep conversations with diverse stakeholders will be key to identifying common goals of importance to scientists and the larger community. This highly interactive session examined whether HGP-write is a wise choice for a Grand Challenge at this point in time, and how to move forward with community outreach.

The participants did not share a single view on the utility or ethics of the project. Additionally, many thought-provoking questions were raised for which no clear-cut answers exist, including:

- Is the need for HGP-write strong enough to justify moving forward?
- Should other, non-human genomes be synthesized first?
- Just because we *can* do something, does that mean that we *should*?
- Where does the ethical boundary lie? Is it the synthesis of 10 genes, 1,000 genes, or a whole genome? And why does it change at that boundary?
- Do the operating rules of ethics really serve the human condition?
- What are the ethics of doing nothing?
- Is there a moral difference between natural and synthetic genomes?

There was general agreement that the answers to these questions would differ among individuals and that there would likely always be disagreement because individuals are compelled by different motivations.

Participants stressed that while it will be important to generate excitement for HGP-write among the larger community, such enthusiasm must be balanced by the need to avoid hype and overpromising the benefits. The HGP-write community can only build trust among the public by earning it through a long-term process of transparency, engagement, and addressing concerns. As part of this process, it should be clearly defined what the project is – and what it isn't. While there are numerous examples of scientific endeavors that have gained public trust, there is an unfortunate legacy of mistrust around GMOs, and in some cases an "us/them" attitude between scientists and the public that needs to be mitigated.

There was a great deal of discussion around community outreach and engagement. Participants warned that it would not be wise to engage with the larger community and

then ignore their comments and concerns about the project. To avoid this scenario, it will be very important to define the purpose of a community outreach program, how concerns will be addressed and/or feedback integrated into the project, and how the outcome of such a program will be evaluated.

Finally, the issue was raised that the precautionary principle – not doing something just in case something bad happens – is not scientific. Scientific advances can always be used for good and evil, and it is important not to let the potential for bad things happening prevent society from moving forward. Instead, the goal should be to maximize the probability of benefit to humanity, while minimizing risk.

The session ended with a reminder that the people in the room are shaping the world we all live in, so it is the responsibility of the HGP-write community to think this project through very carefully.

Safety Engineering

HGP-write will enhance our ability to design and build genetically modified organisms with increased sophistication and ease. In addition to the regulatory and bioethical implications, the development of technologies that endow biosafety measures, or 'safeguards', is essential for this project. This session discussed various safety measures, including biocontainment and genetic isolation, to secure the use of genetically modified organisms in closed systems (e.g., biofermentation processes) and enable safe applications in open systems.

Biocontainment measures include engineering host cells that cannot survive in the natural environment, thereby restricting the viability of the organisms to defined synthetic environments. Genetic isolation measures involve establishing barriers that limit the transfer of genetic material between genetically modified organisms and the environment, obstructing horizontal gene transfer, and engineering resistance to viruses.

Infrastructure, Technology Development and Technical Challenges

The development and use of large-scale DNA synthesis, genome engineering technologies and testing will be integral to the success of HGP-write. This session featured presentations by industrial and academic efforts pioneering computer aided design (CAD), DNA synthesis technologies, genome editing technologies (e.g., CRISPR/Cas systems/ libraries, alternatives to CRISPR -- integrases and beta-recombinases, engineered and custom cell and organoid models) and fully automated solutions to facilitate rapid and facile engineering of whole genomes using a biological foundry for advanced biomanufacturing (BioFAB).

Industry Engagement

Collaboration between academia and industry will be integral to achieve the lofty goals of HGP-write. Key industrial stakeholders to HGP-write will include companies at the forefront of developing low-cost sources of synthetic DNA and commercializing genome editing technologies for gene therapy and tissue engineering. Industry representatives discussed opportunities to contribute and bring to bear unique expertise to enhance the HGP-write endeavor.

There was a general consensus that HGP-write will help to advance current technologies by several orders of magnitude, and it is expected that HGP-write will do for DNA synthesis and assembly technologies what HGP-read did for DNA sequencing technologies. Additionally, we live in a far different time from when HGP-read was launched. For example, there is significantly more public-private collaboration and pre-competitive research collaboration, and the Internet has completely changed business models and made it much easier to collaborate than ever before. Collectively, these advancements will enhance the nature and speed of this endeavor.

Regulatory and Bioethical Implications of HGP-write

Science and technology govern life; therefore, with HGP-write comes great responsibility. Reflecting on the key scientific and technical aspects of HGP-write, a panel of experts in issues related to the ethical, legal and societal impacts (ELSI) of HGP-write discussed responsible innovation. This session was an important contribution towards what will be many transparent discussions to identify common goals and moral concerns of diverse stakeholders. Guided by ELSI experts, the HGP-write project will prioritize a culture of safety, together with a two-way dialogue with the public, as core project goals by establishing an inclusive decision-making partnership with humanities scholars, ethicists, legal scholars, scientists, and members of the lay public.

During this second ethics-focused discussion of the day, additional thought-provoking questions were raised, including:

- How open should this project be, and how should it be shared?
- Who gets to decide how the technologies are used?
- Should we focus less on ethics and more on regulations to protect against scientific risk?
- When does uncertainty (risk/benefit) about a project become a risk?

Participants discussed the notion of 'ripeness', meaning that there is a proper time when an issue is ready to be addressed. Many were in agreement that it is not beneficial to debate

what could potentially happen in 500 years, only what is known today. Keeping within these boundaries could help to keep far-flung issues out of the debate.

Additionally, there are many diverse communities, religions, and cultures throughout the world, each of which will have very different ideas about which biological questions should be answered through HGP-write. If these communities are to have a voice, then what is the right moral framework for reviewing the types of projects that should proceed? For example, despite its life-saving potential, xenotransplantation may not be considered universally ethical. There was general agreement that a single, moral framework may not be possible. Ultimately, the goal of the HGP-write community may be the same as that of the genome editing community, which is to come to a consensus that all international stakeholders will agree to: do no harm. The challenge, however, is that there is no way to enforce such an agreement.

Next Steps

Andrew Hessel, Distinguished Research Scientist at Autodesk and one of the leaders of this project, has obtained a leadership gift of $250,000 from Autodesk to seed the planning and launch of HGP-write. The project will require four matching gifts to cover the planning and launch effort.

Total project costs are difficult to estimate, but will likely be far less than the $3 billion cost of HGP-read. Indeed, recent and continued improvements in enabling technologies are not only improving the cost, quality and time period necessary to complete this project, but also most other fields of biomedicine. The goal is to support HGP-write in 2016 with $100 million in committed support. Initial funding will come from public, private, philanthropic, industry and academic sources from around the world and will build on current funding for E.coli, yeast and mammalian genome writing technology projects.

As funds are raised for project planning and cost reduction, working groups and support meetings will be organized to address governance and oversight, as well as the social and ethical implications of the project.

In the meantime, everyone is invited to become involved in the project and have his or her voice be heard. There are multiple ways to become involved in HGP-write:

- Suggest a new pilot project and/or improved technologies
- Join the conversation via social media
- Join a working group
- Fund the project

Twenty-six years ago, the launch of HGP-read represented a major leap forward for human knowledge and health. In 2016, the launch of HGP-write will represent another giant step for humanity.