



CENTER of EXCELLENCE
for ENGINEERING BIOLOGY



Pilot Project Proposal

Name of Project:

AN *IN-SITU* DIGITAL ANNOTATION SYSTEM TO DOCUMENT AND SAFEGUARD GP-WRITE APPLICATIONS

Proposer and Contact Information: Stephan Züchner*, MD, PhD, Professor and Chair and Rosario Isasi, JD, MPH, Assistant Professor and Director Health Policy, Genetics, Ethics, & Policy Program, Institute for Bioethics, Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami Leonard M. Miller School of Medicine

*Email: szuchner@med.miami.edu

Background: The Genome Project (GP)-write aims for the widespread and routine application of synthetic DNA applications in disparate fields such as clinical medicine, agriculture, biotech, and data storage. The future genome engineer will work similar to a computer programmer, where the ‘commenting’ of code and ‘code signage’ are standard quality assuring procedures. It is now standard to keep developing computer code in GIT repositories (sometimes referred to as ‘global information trackers’). Further, software based products often have ‘watermarks’ that allow for the tracking of copies, authentication of originality, and enforcement of access and copy rights. Will DNA engineering adopt similar procedures? Is there a role for DNA watermarks in GP-write products? This proposal will outline an innovative platform that will allow for a digital global annotation system to document and safeguard GP-write applications. In addition, it will contribute to identifying the regulatory issues and challenges that would need to be resolved in order to implement such a global annotation system.

Technical Idea: Similar to computer code, GP-write products will be more useful, safer, and traceable, even over very long time spans, when documented ‘in situ’ accordingly. We propose a global, open, standardized system of documenting GP-write edits and products that could become part of a regulatory frame work. Edited genomes or parts thereof will contain for example, a 16-digit nucleotide based bar code (16 bits) providing 10^{16} unique tokens, encoded in a degradation resistant fashion, which can be read out by standard sequencing technology – similar to technology described in Church et al¹. Unique nucleotide signatures will be linked to an open, web-based database, containing structured information elements such as date, user, detailed feature and product description, release dates, etc. This approach may be used to ‘comment out’ parts of a GP-write product, such as describing unique features in an artificial chromosome, regulatory regions, coding regions, etc. APIs will allow for seamless integration of this database into future software platforms for GP-write engineering. There could also be a closed (commercial) section for such a database system, and even a two-way authentication requirement that is DNA based (for instance the user may need knowledge of a second DNA-encoded code to unlock a set of protected signatures). Finally, the system can be used as a shared GIT repository, detailing different developing versions of an engineered genome.

Utility: Usability, safety, and regulatory compliance of GP-write products will ultimately rely on standardized documentation. There is a unique opportunity to create an ‘in situ’ standard for documenting artificial DNA products as the field is in its infancy. Such standardized annotation of synthetic DNA products will satisfy regulatory demands and conciliate public concerns. With widespread GP-write applications, such barcodes will appear in the environment as GP-write products are released (agriculture), become part of medicine (one’s blood may literally contain such annotations), and more. With deep sequencing it is relatively straight forward to detect these unique signatures (at small relative amounts), quantify, and read such codes; allowing for data driven decisions (environmental, medicine, etc).

“Fit” For GP-write: GP-write in general, and the Human Genome Project-write (HGP-write) in particular, will be scrutinized for safety and global compliance. Developing early an open standard to track deeply annotated GP-write products with unique ‘in situ’ signature barcodes will secure public trust, improve quality, and ensure long-term traceability of engineered DNA independent of the originating laboratory.

Reference

1: Church GM, Gao Y, Kosuri S. Next-generation digital information storage in DNA. *Science*. 2012 Sep 28;337(6102):1628. Epub 2012 Aug 16. PubMed PMID: 22903519.