

Bioverse: a comprehensive platform for designing biological materials

Andrea L. Halweg-Edwards*¹, Andrew Garst¹, Alaksh Choudhury¹, Cierra Walker², Adam Robbins-Pianka¹, Akash Gaonkar³, Gur Pines¹, Ryan Hockstad¹, Katia Tarasava⁴, Deanne Sammond⁵, Adam Arkin⁶, **Ryan T. Gill (PI)**¹

*Presenter: andrea.edwards@colorado.edu

¹Renewable and Sustainable Energy Institute, University of Colorado Boulder, Boulder, CO 80309

³Department of Computer Science, University of Colorado Boulder, Boulder, CO 80309

⁴Materials Science and Engineering Program, University of Colorado Boulder, Boulder, CO 80309

⁵Biosciences Center, National Renewable Energy Laboratory, Golden, Colorado, 80401

⁶Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

Project Goals: We aim to build a user-friendly platform for designing biological materials and exploring molecular adaptation. Our goal is to consolidate biological databases with structural, functional, and phylogenetics information to enhance the search speed and retrieval of relevant design features for a wide range of query types. As a result, the bioverse platform will allow advanced, google-esque query capabilities allowing engineers to design experiments that probe important processes including biophysical properties that confer disease phenotypes as well as design products ranging from secondary metabolites to biological packaging materials.

Abstract

The discovery of CRISPR-enabled genomic and episomal editing technology has sparked a surge in development of synthetic biology tools for editing DNA. Coupled with array-based DNA synthesis technology, biologists now have the opportunity to edit targeted sequences at unprecedented scales and throughputs. In response to this recent progress, our BioDesign Group is developing *bioverse*, a freely-available and open-sourced computational design platform. This web-based design tool will provide users with the capability to generate tens of thousands of rationally designed editing cassettes based on our patented CRISPR EnAbleD Trackable genome Engineering (CREATE) technology, which is compatible with array-based DNA synthesis platforms. Bioverse will incorporate phylogenetic informatics as well as structure-guided design principles powered by the Rosetta software suite for computational modeling and analysis of biological macromolecules. Furthermore, our design platform will be built upon active machine learning approaches that will allow users to input their test data back into the pipeline to improve mutant-phenotype predictive power and inform the next round of library synthesis. Additionally, we are working to integrate bioverse with the Department of Energy Systems Biology Knowledgebase (Kbase) to allow crowd-sourced feedback on macromolecular design principles and best practices as well as providing the community with a wide range of metabolic engineering tools.

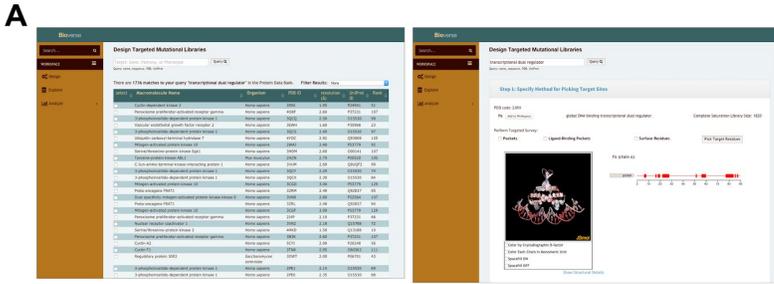
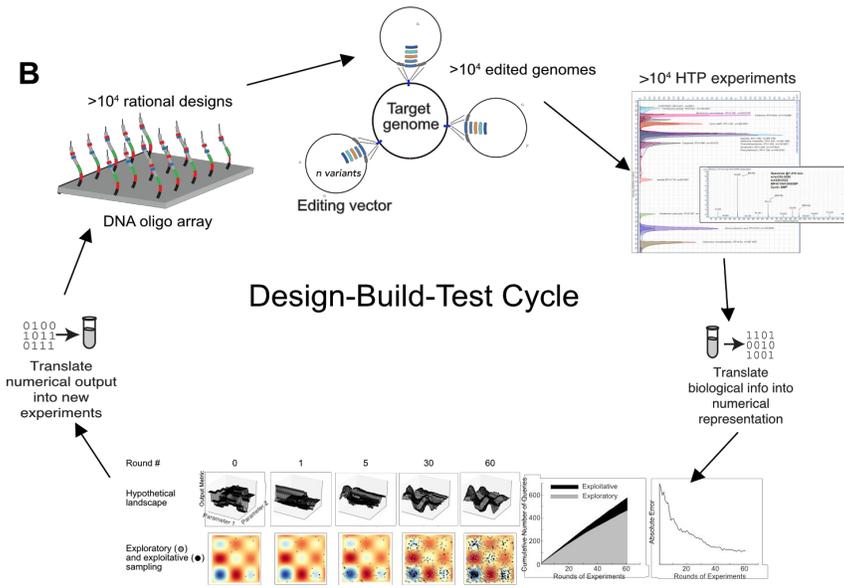


Figure 1. Synthesis Aided Design.

Affordable synthetic DNA oligomers combined with high throughput genome editing and testing technologies allow engineers to rapidly generate informative datasets regarding genotype-phenotype relationships. **A)** The bioverse platform will allow researchers to rapidly design biological materials for testing a wide range of phenotypes including small molecule production and molecular adaptation. **B)** Bioverse aims to design libraries that include a sufficient range of hypotheses and controls, allowing application of active machine learning to improve the predictive power of genotype-phenotype models during iterative rounds of Design-Build-Test-Learn. In this example, mass spectrometry based methods are used to collect data on a library of >10,000 genetic variants, and classification and regression trees are used to inform the next round of experiments. Exploration and exploitation of experimental datasets allows efficient optimization of the mathematical model describing the relationship between genotype and phenotype.



Funding Statement

Grant title: A Platform for Genome-scale Design, Redesign, and Optimization of Bacterial Systems; Project grant number (DE-SC008812) and FWP number (ERWER44).