Simulated Evolution for Synthetic Biology: A Case Study on Biological Counters

KEYWORDS: *digital logic, evolutionary circuit design, genetic models, inversion, recombination, systems biology, transcriptional networks.* ***This project is original and of my own design.***

BACKGROUND: Biological systems self-regulate via networks of interacting transcription factors. Such networks produce complex behavior; some theorists have argued that they can produce any behavior desired.[1] However, existence proofs do not guarantee that *practical* solutions exist.

One method to test if a desired behavior can be achieved is to run a computer simulation. The computer can test thousands of designs quickly, converging on the best matches. This process has been used to evolve designs for oscillators, latches, and other interesting behaviors.[2,3]

To demonstrate and extend the power of simulated evolution for biological network design, I will evolve *a genetically encoded binary counter*. Synthetic biology has long used binary counting as a model system. It is perhaps the simplest behavior which requires a full implementation of digital logic; this is significant because of the robustness and simplicity of digital circuits. Counters have inspired many entries in the iGEM design contests[4] as well as high profile experimental attempts, such as the unary (linear) counter in last year's *Science*.[5] However, to my knowledge, all these designs have used nonstandard elements such as recombination-based genetic switches.

RESEARCH PLAN: I will perform simulated evolution and compare counter designs with and without switches. Viable designs will be refined by hand and ultimately become physical DNA constructs for testing in vivo. *I predict that only designs evolved with switches will be viable.* These results will be significant in themselves; they will also inform the synthetic biology design process and improve the modeling infrastructure for future research.

ON COUNTERS: Briefly, a counter element changes state every n^{th} time it is triggered. For a binary counter, $n = 2$. One could imagine using a binary counter to produce yeast that turn green for every 2^{nd} cell division. More seriously, one could imagine a researcher in aging or cancer connecting five such counters to make cells turn green when they have divided $2^5 = 32$ times.

AIM 1: Perform selection for binary counting in simulated genetic networks.

Existing software can simulate evolution in genetic networks. This work will use Genetdes,[6] free open-source software which natively uses SBML format[7] to describe the networks it acts upon. As a former metabolic engineer, I am proficient in SBML and with network modeling.

HYPOTHESIS 1: *Electronic circuits will evolve into binary counters.* The field of evolutionary circuit design began in electronics, and a mature literature exists on the topic.[8,9] Counter design is a common test case, and has been demonstrated a number of times. This will serve as a positive control; I will test and refine the selection methods on a target known to exist.

HYPOTHESIS 2: *Under validated selection conditions, networks of interacting transcription factors will NOT evolve into binary counters.* Selection will be performed with Genetdes using a proven fitness function. However, success appears unlikely. No human designer has produced a counter network from purely transcriptional logic, nor do known examples exist in nature.

AIM 2: Improve simulated networks by including recombination-based genetic switches.

To test these switches in the context of a modeled network, I will make several changes to the Genetdes simulator and its underlying SBML representations. Current SBML standards permit embedded triggers for functions and discontinuous events; these are key features for implementing recombination. In this Aim, I will therefore extend Genetdes to full SBML-2 compliance.

To update the model, I will first create new parts: matched pairs of recombinase and target site (*Rec* and INT), using available kinetic data for the *fim* system.[10] *Rec* binds INT, then complexes

with another bound *Rec*. This complex triggers a discontinuous event, in which the paired INTs exchange locations. INTs will be context aware, storing the connections made at their genetic location (*cis* interactions). I will also create a new cell compartment where active INTs and their neighbors will be hidden during this exchange, mimicking DNA blocked by *Rec* complex.

HYPOTHESIS 3: *Networks containing switches will outperform transcription factors alone.* Recombination is the favored mechanism for natural behaviors with periodic state changes, such as *E. coli* virulence[10] and *S. cerevisiae* mating.[11] The switches are discrete, leak-free, and fairly efficient - good traits for a counter.[5] I expect that switching will improve the counter designs.

AIM 3: Convert evolved counter models to physical form and confirm activity in vivo.

GOAL: *Produce functional networks in a living cell with simulated evolution.* For high-scoring designs, I will perform stochastic simulations on their network to assess noise tolerance. Designs which survive this test will be subject to sensitivity analyses, determining their robustness to kinetic parameters. These analyses will be performed with the SimBiology toolkit in MATLAB.

For the most promising designs, I will manually ensure that each model element has a corresponding physical part with the right kinetics. I will also choose the most informative elements to tag with fluorescent reporter proteins. Designs will be built with characterized parts libraries.[12,13]

Testing and debugging the designs will require dynamic analysis of multiple reporter proteins, ideally in single cells. *I plan to work with Prof. Hasty of UCSD, a pioneer of this technique.* His group has performed similar work for oscillators and circadian clocks, showing feasibility.[14]

AIM 4: Produce educational software using ideas from network evolution models.

The algorithms that evolve SBML models in Genetdes can be generalized to use any structured input.[6] I will refactor existing software to develop a general-purpose evolution simulator; this simulator will be used to create programs in Turtle, a simple graphics language for children.

The fitness of a Turtle program will be rated by humans comparing its output to a target image. This will be a web game, which will incorporate game dynamics such as allowing users to compete on how well their scores match the consensus. I will create time-lapse videos of the evolutionary process; these will be posted on YouTube to provide visually appealing tools for education.

RESOURCES: Pilot studies for Aims 1 and 4 will be conducted on a small Beowulf cluster[15] planned for DIYbio-Boston. Next fall, I plan to enroll in graduate school; I will use university resources (or ideally, TeraGrid) for Aim 2. Labs for Aim 3 can be found at UCSD and elsewhere.

BROADER IMPACTS: I have dedicated a full Aim to public outreach. For the Turtle project, the tools will be developed in partnership with amateur scientists in DIYbio. Data will be collected through public participation, and results will be packaged to reach the widest possible audience.

I take scientific impacts just as seriously. I will continue to publish and give talks, making sure to reach the broader group which could expand on my work. For instance, the parts designers would want to know if I was limited by a specific gap in the parts libraries, and the modelers want a formalism for recombination. Within my subfield, I also hope to use interesting results to make the case for design practices like simulated evolution. Synthetic biology wants to be transformative; it could be, if we were better at it. By learning to design biological systems better, smarter, and faster, and by sharing that knowledge, we build the infrastructure that will allow it to reach its potential.

REFERENCES: (1) Buchler *et al*. 2003. *Proc Natl Acad Sci USA*. (2) Rodrigo G & Jaramillo A. 2007. *Syst Synth Biol*. (3) Cao H *et al*. 2010. *Syst Synth Biol*. (4) igem.org (5) Friedland AE *et al*. 2009. *Science*. (6) Rodrigo G *et al*. 2007. *Bioinformatics*. (7) sbml.org (8) Beielstein T *et al*. 2002. *IEEE-CEC*. (9) Shanthi AP *et al*. 2005. *IEEE-EH*. (10) Ham TS *et al*. 2008. *PLoS ONE*. (11) Haber JE. 1998. *Annu Rev Genet*. (12) partsregistry.org (13) biofab.org (14) Bennett MR & Hasty J. 2009. *Nat Rev Genet* (15) beowulf.org