Published in final edited form as:

Integr Biol (Camb). 2016 April 18; 8(4): 556–563. doi:10.1039/c5ib00286a.

Performing selections under dynamic conditions for synthetic biology applications

Jessica M. Lindle and Mary J. Dunlop

School of Engineering, University of Vermont, Burlington, VT, 05405, USA

Abstract

As the design of synthetic circuits and metabolic networks becomes more complex it is often difficult to know *a priori* which parameters and design choices will result in a desired phenotype. To counter this, rational design can be complemented by library-based approaches where diversity is introduced and then coupled with screening or selection methods. Here, we used a model of competitive growth to show that selection can rapidly identify library variants with near-optimal phenotypes. Many synthetic biology applications require phenotypes that balance multiple objectives, such as responding to more than one chemical signal. In addition, desired traits may be time-dependent, for example changing with the growth phase. By applying dynamic inputs to the selection, we show that it is possible to select for traits that satisfy multiple goals. Furthermore, we demonstrate that the underlying diversity in a library is heavily influenced by the initial circuit design. Overall, our findings argue that rational synthetic circuit design, coupled with diversity generation and dynamic selection are powerful tools for many synthetic biology applications.

Introduction

Synthetic circuit design and metabolic engineering have benefited greatly from the integration of rational design with approaches that introduce diversity and variability. Tools from synthetic biology have enabled the construction of large libraries that contain many variations on a particular design. For example, in directed evolution error prone PCR or other mutagenesis methods can introduce mutations within specified regions and library members can be subsequently screened or selected for desirable phenotypes.¹ These methods have been enabled by advances in synthetic biology, including efficient cloning protocols and the decreasing cost of DNA synthesis.^{2–4} In addition, multiplexed genome engineering methods can rapidly introduce variability at a genome-wide scale.^{5–8} With recent advances in synthetic biology tools that have enabled the design and construction of libraries of genetic variants, our ability to screen or select for desired traits is critical to success.¹

In principle, given sufficient time, selection will identify the best candidate from a large library; however under practical engineering circumstances it is often preferable to rapidly identify a solution that is very close to the optimal while using less effort (time and resources). For example, the number of rounds of selection required may be dramatically

different to isolate a library variant in the top 1% as opposed to one in the top 10%. Because lengthy selections increase the possibility of mutation or adaptation effects that can confound analysis, it may be preferable for synthetic biology applications to find a solution that is “close enough” to optimal.

We highlight the use of biosensors here as a particularly relevant example where well-designed screening or selection procedures can have a large impact. Biosensors can be coupled with readable outputs, linked to an output providing a fitness advantage, or used directly in a metabolic pathway through feedback control.^{9,10} When biosensors are coupled with reporters, such as fluorescence, they provide a semi-quantitative way to rapidly screen many genetic variants (for recent reviews see ref. 9–11). Biosensors can also be used to control expression of genes that provide a fitness advantage.^{9–11} This growth-coupled assay approach has been used to select for improved L-lysine¹² and butanol¹³ production strains. In a recent study that has broad potential, a biosensor was used to control expression of TolC, where increased expression leads to higher levels of antibiotic resistance, providing a straightforward selection method.¹⁴ Critically, the authors made use of negative selection to eliminate false positives from their libraries. An alternative strategy uses a phage-assisted approach, where the target of interest is coupled with expression of a phage infection protein, allowing rapid evolution of novel synthetic systems for the production of biomolecules.¹⁵ Coupling a biosensor to a selectable output, especially when there is a way of performing negative selection, can be applicable for improving production of a broad variety of compounds. Here, we take this approach and assume that the phenotype of interest can be directly coupled to growth rate. This can be achieved by using a biosensor to control a gene that provides a growth advantage, or if the system of interest already provides a selective advantage it can be used directly.

Synthetic biology applications often have multi-factorial objectives where designs need to perform well under multiple conditions or conditions that change with time. For example, when designing a biofuel production strain, cells should produce fuel at high levels, but should not be burdened by the additional metabolic infrastructure to the point that they are under dramatic selective pressure to eliminate the engineered components. Another example comes from the design of biological sensors. Sensors need to respond to the presence a signal, but not the absence. This is especially critical if the sensor is driving an actuator as part of a feedback loop, where it might be costly to have the actuator on at all times. In addition, it might be desirable to select for a sensor that responds in a graded or linear fashion, while other applications may require a digital response. Furthermore, the desired phenotype may be one that needs to perform well under changing environments. For example, in metabolic engineering applications a toxic byproduct may build up over time. Therefore, strains need to be tolerant to ramping increases in the byproduct, but do not need to exhibit complete resistance in early stages of growth.

Synthetic systems can be designed to respond to dynamic conditions and efficient dynamic selection methods can aid in synthetic circuit optimization. Here, we explore a specific example using a biosensor to drive expression of a biofuel tolerance mechanism in a synthetic feedback loop. We compare the system without feedback (static control) to that with feedback (dynamic control).¹⁶ There have been several examples of circuit designs that

use synthetic feedback loops to achieve optimized outcomes, including production of lycopene,¹⁷ farnesyl pyrophosphate,¹⁸ and fatty acid based products.^{19–21} Importantly, all of these processes impose a burden on the cell if pathways are imbalanced due to accumulation of toxic intermediates or byproducts. Therefore, it is critical that the response not be static, but rather commensurate with the level of burden. These sensor–actuator designs typically require some tuning to achieve optimal performance, as it may not be clear what combination of parts will give the best performance. This can be achieved by constructing several variants and comparing them; for example, Zhang *et al.* built 30 combinations of their synthetic circuit design and measured fatty acid ethyl ester production directly in each individual.²¹ An alternative to testing individual designs is to pool the different variants, expose cultures to time-varying inputs that mimic the real system, and select for those that perform well. The ability to screen library variants for beneficial properties in an efficient fashion can dramatically improve throughput.

In evolution-guided design there may be features inherent to a synthetic circuit that impose constraints on the range of variability or shape of distribution that can be achieved by a library. For example, efflux pumps can provide biofuel tolerance, but their overexpression imposes a burden on cell growth.^{22,23} This tradeoff will become apparent in any design that introduces diversity *via* static pump expression: there will be a strong correlation between library members that provide tolerance to biofuel and those that do not grow as well in the absence of fuel. Introducing library diversity in a way that works around these constraints can make selection more straightforward. In the efflux pump example, introducing different constitutive promoters to control pumps will not get around the trade off between biofuel and pump toxicity, while designs that include feedback control can be used to circumvent this constraint.

Here, we show that dynamic selection protocols can identify library members that have complex phenotypes. For example, applying a time-varying input can select for library members that have beneficial properties under multiple different environmental conditions or that respond in a specific way to conditions that change with time. Given the prevalence of tools from synthetic biology that enable large-scale library construction, the ability to efficiently screen for complex traits represents an important area for optimization in synthetic biology. We envision that synthetic circuit design should include rational components, coupled with diversity generation, and dynamic screening or selection.

Experimental

Competitive growth model

We used the Lotka–Volterra competitive growth equation²⁴ to model the biomass of each of strain (N_i) over time during each selection cycle:

$$\frac{dN_i}{dt} = g_i N_i \left(1 - \sum_{j=1}^C A_{ij} N_j \right)$$

The growth rate of strain i is g_i , C is the total number of strains, and A_{ij} is the strain interaction matrix, which we define as $A_{ij} = \frac{g_j}{g_i}$ so that the competitive advantage that strain j has over strain i is proportional to the ratio of the growth rates. We note that other interaction matrix models could be used to incorporate more complex community dynamics.

The following conditions are identical across all simulations: Initial conditions were set to $0.01/C$ for all strains. Thus, the total initial biomass is 0.01 and all strains start with identical initial conditions. Simulated cultures were grown for 12 hours, then the biomass at 12 hours of all strains was divided by 100 (a 1:100 dilution) and these values were used as the initial conditions for the next selection cycle. Note that the carrying capacity of the total culture is normalized to one such that the final total biomass, which is the sum of the biomass of all C strains, at steady state is equal to one.

Growth rate distributions, number of strains, and number of selection cycles vary and are listed in the figure captions. Here we give a brief overview: for Fig. 1, 2D and E, the growth rate (g_i) values were sampled from normal distributions. Fig. 2F is similar, where g_i values were sampled from two normal distributions with switching between them as a function of the input. For Fig. 3, g_i values depend on the input and are a function of two normally distributed parameters, α_i and K_i , described below. In Fig. 4, the g_i values are calculated from the pump model and depend on the pump expression value, α_p , as described below.

All simulations were performed in Matlab (Mathworks, Inc.) using the ode45 differential equation solver.

Input-dependent growth rate model

In Fig. 3 we modeled the growth rate as a function of the input using the function

$$g_i = \alpha_i \frac{\left(\frac{I}{K_i}\right)^2}{1 + \left(\frac{I}{K_i}\right)^2}$$

where I is the input, α_i is the maximal growth rate, and K_i is the input concentration at which growth is half-maximal. Although the exact choice of this function is arbitrary, it represents situations where the input is required for growth, but where the growth rate eventually saturates, such as in a growth-coupled assay.¹⁴

Biofuel toxicity and efflux pump model

In Fig. 4A–C, we used a biofuel toxicity model we developed previously, which is based on fits to experimental data.²² Here, we briefly state the mathematical model; further details on its derivation and parameter values are available in ref. 22 and figures and additional background are available in ref. 25–27.

$$\frac{dN}{dt} = \mu_{\max} N \frac{S}{S+K_s} \frac{1}{1+\left(\frac{c_{\text{int}}}{K_c}\right)^{h_c}} \frac{1}{1+\left(\frac{p}{K_p}\right)^{h_p}}$$

$$\frac{dS}{dt} = -\frac{1}{\gamma} \mu_{\max} N \frac{S}{S+K_s}$$

$$\frac{dp}{dt} = \alpha_{p0} + \alpha_p - \beta p$$

$$\frac{dc_{\text{int}}}{dt} = \alpha_{c0} \frac{V_{\text{ext}}}{V_{\text{int}}} (c_{\text{ext}} - c_{\text{int}}) - \alpha_c c_{\text{int}} p$$

N is the biomass of the strain, S is the substrate level, p is the efflux pump level, and c_{int} is the intracellular biofuel concentration. The parameters and additional equations required for the simulations are: $\mu_{\max} = 3.0 \text{ h}^{-1}$, $K_s = 3.0$, $K_c = 0.25\% \text{ v/v}$, $h_c = 2$, $K_p = 0.8$, $h_p = 4.0$, $\gamma = 1.05$, $\alpha_{p0} = 0.092 \text{ proteins per h}$, $\beta = 1 \text{ h}^{-1}$, $\alpha_{c0} = 0.001 \text{ h}^{-1}$, and $\alpha_c = 30 \text{ h}^{-1}$. $V_{\text{cell}} = 1 \times 10^{-15} \text{ L}$, $m_{\text{cell}} = 9.5 \times 10^{-13} \text{ g}$, $V_r = 0.5 \times 10^{-3} \text{ L}$, $V_{\text{int}} = N \times (V_r/m_{\text{cell}}) \times V_{\text{cell}}$, $V_{\text{ext}} = V_r - V_{\text{int}}$, and $c_{\text{ext}} = (b_{\text{ext}} \times V_r - c_{\text{int}} \times V_{\text{int}})/V_{\text{ext}}$, where b_{ext} is the level of biofuel input.

We varied the pump expression rate α_p . For each pump expression rate, we simulated the effect of biofuel toxicity on growth using the four equations above. We then used the rate of biomass change over time, dN/dt , to calculate the mean growth rate in mid-log phase, which we define as between 2 and 4 hours. These growth rates, g_i , were then used for all subsequent simulations using the Lotka–Volterra equation described above.

Biofuel efflux pump model with feedback

The model in Fig. 4D and E is identical to the biofuel export model above with the exception that the pump expression depends on the level of intracellular biofuel. Thus, the dp/dt equation is replaced, as described in ref. 27, with:

$$\frac{dp}{dt} = \alpha_{p0} + \alpha_p \frac{\frac{c_{\text{int}}}{K_c}}{1 + \frac{c_{\text{int}}}{K_c}} - \beta p$$

Growth rates were calculated and used in the Lotka–Volterra competitive growth model in the same way as described in the previous section.

Results and discussion

For our library, we consider a set of strains, all of which are variants on each other, such as multiple instances of a synthetic circuit design where mutations have been introduced. When

grown separately, the strains have different growth rates, but show similar growth curves (Fig. 1A). When pooled in equal proportion and grown together, the differences in growth rate are amplified, with strains that grow well eventually dominating the culture. As our model of selection, we assume that the strains are grown in liquid media with 1:100 serial dilution every 12 hours. As an illustrative example, we simulated selection due to competitive growth of 10 strains with very minor differences in growth rates (mean = 0.5, standard deviation = 0.05 h⁻¹). Although the 10 strains start out with equal representation, with subsequent selection cycles the top ranked strains make up larger proportions of the biomass, rapidly eliminating the lower ranked strains (Fig. 1B).

We next considered a library with 1000 members and asked how long it would take to converge to the top ranked members. Because competitive growth amplifies differences in growth rate, our model suggests that any library member that has an appreciable biomass at the end of the selection should be very similar in its performance to the top ranked library member. In practical terms, this means that it may be possible to run the selection for a much shorter amount of time. To quantify this idea, we recorded the amount of biomass in the culture that was made up by the top 10%, 5%, and 1% of strains in the library (Fig. 1C). In all cases, as the number of selection cycles increases the top performers make up an increasing fraction of the biomass. As expected, the 10% case converges more quickly than the 5% case, which is faster than 1%. Importantly, these results indicate that within a small number of selection cycles, the top performers make up a large percentage of the biomass. For example, after six selection cycles, 99.8% of the biomass is made up of strains from the top 10% of the library. These results are in line with findings from a study targeting improved production of naringenin and glucaric acid, where four rounds of selection were sufficient to dramatically improve titers of these chemical compounds.¹⁴ Our findings suggest that convergence to near-optimal solutions is rapid and can occur within a small number of selection cycles, reducing effort and the potential for mutation.

The time to convergence depends upon library diversity. We recorded the fraction of the total biomass that is made up by the top 10% of the library after six dilution cycles as a function of the standard deviation of the library members' growth rates (Fig. 1D). We use standard deviation as a measure of diversity within the population. As expected, libraries that contain very similar members have low diversity and take longer to converge. As diversity increases, convergence is faster. Even though convergence to the top 10% is slow when library diversity is low, because the library members are very similar it is still reasonable to run the selection for a small number of cycles, as the solutions will all be similar and additional selection cycles will yield little improvement.

We next asked how the selection would be influenced by the type of selective pressure placed upon the library. We first considered a case where we were interested in optimizing growth rate in the presence of two inputs. This could represent a situation where there are two distinct inputs, for example selecting for a biosensor that responds to two different small molecules. Alternatively, the two inputs could be the same signal present at two different levels, such as selecting for strains that perform well in both low and high stress. We initially considered library members that had growth rates that were normally distributed about a mean value that depends on whether input A and B is present (growth rates are g_A and g_B).

We then subjected the libraries to conditions with only input A, only input B, and with switching between inputs A and B (Fig. 2A–C). The selection correctly sorts the library members given the selective pressure. With only input A, the library members that have the highest growth rate values (g_A) are selected preferentially; these results are agnostic to the growth rate in the presence of input B (g_B) (Fig. 2D). Equivalently, when growth is conducted with only input B, the selection optimizes for high g_B values (Fig. 2E). However, when there are multiple objectives, constant selective pressure can be insufficient to select for library members that function well under a variety of conditions. As an illustration of this point, we applied an input that varied between A and B. In this case, the selection returns library members that fall along a diagonal line, having good growth rates in the presence of both inputs A and B (Fig. 2F). The slope of this line is determined by the ratio of time strains are exposed to input A *versus* input B.

Dynamic selection methods can be used to identify library members that perform well under changing conditions. As an example, we considered a model where growth rate is a continuous function of input concentration. Our inspiration for this model comes from several recent examples of growth-coupled assays,^{12–15} where a sensor for a compound of interest is coupled with a selective marker. For our library, we used a sigmoidal function to describe growth as a function of the input. We allowed both the maximum growth rate and the input concentration where half-maximal growth is achieved to vary, such that each library member has a unique growth function associated with it (Fig. 3A). To summarize the results, we plot the maximum growth rate, $g_{\max}$, against growth with a constant low level of input, $g_{\min}$. The shape of the distribution is dependent upon how variation in the two parameters affects growth. As before, given a constant level of input, library members with high growth rates in the presence of that input will be selected (Fig. 3B and C). However, we also considered conditions where inputs changed with time. For example, such an approach might be useful if a sensed compound is likely to change with growth phase. In contrast to the constant input case, the selected library members now fall along a curved line that balances growth in the presence of changing levels of the input (Fig. 3D–G).

Tradeoffs can also impact the underlying distribution of a library, and synthetic circuit design choices can dramatically alter the distribution shape. As a representative example, we considered a model we developed previously for tradeoffs in the toxicity of biofuel efflux pumps.²² Efflux pumps can provide tolerance to biofuel, but overexpressing these pumps places a burden on the cell. These two competing sources of toxicity lead to a tradeoff, where growth rate is maximal at intermediate expression levels that balance biofuel and pump toxicity. Previous experimental and modeling work has characterized this tradeoff across a range of pump expression levels.²² In the context of this study, we simulated strains with different pump expression levels (Fig. 4A), competed them against each other, and identified the top performers under different biofuel inputs. When biofuel is constant, the strains that perform well are those that are optimized for that particular biofuel level (Fig. 4B and C). Here, we indicate the top performer and also those strains that have growth rates within 10% of this strain. We note that this metric differs from that used in Fig. 2 and 3, and is useful for showing how sensitive the results are to a specific pump expression level. As expected, when no biofuel is present, the top performing strains have low pump expression since there is no benefit to expressing pumps to counter biofuel toxicity (Fig. 4B and C). If

the biofuel levels vary with time, strains that balance the two types of toxicity are selected. If the likely fuel profile for a biofuel production strain is known, for example, if biofuel slowly ramps up over time, a similar pattern for exogenous addition of fuel can be tested to select for a strain with an optimized pump expression level.

The choice of synthetic circuit design can have a dramatic impact on the underlying distribution of library members, making it easier or more challenging to select for library variants that have ideal properties. To highlight this, we extended the biofuel tradeoff model to include feedback. We modeled this by allowing efflux pump expression to be controlled by intracellular biofuel concentrations.^{26,27} This design mirrors that of several recent studies that have used biosensors to control expression of toxic genes or the production of toxic pathway intermediates.^{17–21} As before, we introduced variability in the pump expression level. With the addition of feedback, the tradeoff between biofuel and pump toxicity changes, and broader pump expression ranges allow for good performance while mitigating biofuel toxicity (Fig. 4D). Using the same selection strategy as before, we identified those library variants within 10% of the optimal strain, testing different biofuel inputs (Fig. 4B and E). In all cases, the number of library members that met the 10% criterion was increased relative to the system without feedback (Fig. 4C and E). This finding highlights the importance of coupling random approaches to library generation with well-designed synthetic circuits. In the case of biofuel toxicity, there are feedback circuit designs that will dramatically outperform static control. Selection can be used to distinguish between many different options, but good synthetic circuit design will bias library members towards optimal solutions. These libraries can then be subjected to selection using dynamic inputs that mimic conditions of interest.

Conclusions

As designs of synthetic circuits and metabolic networks become more complex, it is often difficult to tell in advance which exact parameters will lead to desired phenotypes. The ability to couple rational design with diversity introduced through mutagenesis or multiplexed genome engineering strategies offers researchers the potential to screen many library variants for a desired phenotype. Recent studies have shown the great potential of coupling phenotype to growth-based assays, allowing for efficient selection of desired traits.^{12–14}

Here, we use a competitive growth simulation to show that it is possible to rapidly differentiate library variants based on subtle differences in growth rate. Only a small number of selection cycles are necessary to converge to near-optimal solutions. This offers great potential for engineering applications, where near-optimal solutions are often acceptable. In this work, we used a competitive Lotka–Volterra model, however other models, such as those that explicitly model competition for nutrients²⁸ or relative fitness benefits,^{29,30} could allow other effects to be incorporated.

The main contribution of this study is that we considered the possibility that the desired phenotype is a combination of different factors and may change over time. These are important considerations, as designs for biosensors and metabolic engineering applications

often involve tradeoffs and the ability to select based on more than one trait has the potential to dramatically improve how selections are performed. We considered differences in growth rate under two different inputs and showed that distinct library members are selected based on which inputs are applied. The shape of the underlying distribution is also important for determining what will result from a selection. We considered two examples: input-dependent growth and a biofuel export system. In each case, it is possible to use dynamic inputs to bias selections towards library members with traits that match changing inputs. In the biofuel model we showed that the underlying distribution of library members could be altered to bias the library towards ideal phenotypes. The ability to generate many variants and select is not a substitute for well-designed synthetic circuits; good circuit design can dramatically improve the changes of finding a variant that has a desired phenotype. This work highlights the potential of evolution-guided design for synthetic biology where desired phenotypes have multiple, dynamic objectives.

Acknowledgments

We thank Mads Almalkhi for helpful discussions. This work was supported by the Office of Science (BER) at the U.S. Department of Energy.

References

1. Dietrich JA, McKee AE, Keasling JD. Annu. Rev. Biochem. 2010; 79:563–590. [PubMed: 20367033]
2. Engler C, Kandzia R, Marillonnet S. PLoS One. 2008; 3:e3647. [PubMed: 18985154]
3. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO. Nat. Methods. 2009; 6:343–345. [PubMed: 19363495]
4. Quan J, Tian J. Nat. Protoc. 2011; 6:242–251. [PubMed: 21293463]
5. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM. Nature. 2009; 460:894–898. [PubMed: 19633652]
6. Jakobsen T, Bonde I, Herrgård M, Harrison SJ, Kristensen M, Pedersen LE, Jensen MK, Keasling JD. Metab. Eng. 2015; 28:213–222. [PubMed: 25638686]
7. DiCarlo JE, Conley AJ, Penttilä M, Jäntti J, Wang HH, Church GM. ACS Synth. Biol. 2013; 2:741–749. [PubMed: 24160921]
8. Ryan OW, Skerker JM, Maurer MJ, Li X, Tsai JC, Poddar S, Lee ME, DeLoache W, Dueber JE, Arkin AP, Cate JH, Izaurralde E. eLife. 2014; 3:e03703.
9. Ford TJ, Silver PA. Curr. Opin. Chem. Biol. 2015; 28:20–28. [PubMed: 26056951]
10. Liu D, Evans T, Zhang F. Metab. Eng. 2015:35–43. [PubMed: 26142692]
11. Zhang J, Jensen MK, Keasling JD. Curr. Opin. Chem. Biol. 2015; 28:1–8. [PubMed: 26056948]
12. Yang J, Seo SW, Jang S, Shin S-I, Lim CH, Roh T-Y, Jung GY. Nat. Commun. 2013; 4:1413. [PubMed: 23361004]
13. Dietrich JA, Shis DL, Alikhani A, Keasling JD. ACS Synth. Biol. 2012:47–58. [PubMed: 23656325]
14. Raman S, Rogers JK, Taylor ND, Church GM. Proc. Natl. Acad. Sci. U. S. A. 2014; 111:17803–17808. [PubMed: 25453111]
15. Esvelt KM, Carlson JC, Liu DR. Nature. 2011; 472:499–503. [PubMed: 21478873]
16. Holtz WJ, Keasling JD. Cell. 2010; 140:19–23. [PubMed: 20085699]
17. Farmer WR, Liao JC. Nat. Biotechnol. 2000; 18:533–537. [PubMed: 10802621]
18. Dahl RH, Zhang F, Alonso-Gutierrez J, Baidoo E, Batth TS, Redding-Johanson AM, Petzold CJ, Mukhopadhyay A, Lee TS, Adams PD, Keasling JD. Nat. Biotechnol. 2013:1039–1046. [PubMed: 24142050]

19. Liu D, Xiao Y, Evans BS, Zhang F. ACS Synth. Biol. 2015; 4:132–140. [PubMed: 24377365]
20. Xu P, Li L, Zhang F, Stephanopoulos G, Koffas M. Proc. Natl. Acad. Sci. U. S. A. 2014; 111:11299–11304. [PubMed: 25049420]
21. Zhang F, Carothers JM, Keasling JD. Nat. Biotechnol. 2012; 30:354–359. [PubMed: 22446695]
22. Turner WJ, Dunlop MJ. ACS Synth. Biol. 2015; 4:1056–1063. [PubMed: 25496359]
23. Dunlop MJ, Dossani ZY, Szmidt HL, Chu HC, Lee TS, Keasling JD, Hadi M, Mukhopadhyay A. Mol. Syst. Biol. 2011; 7:487. [PubMed: 21556065]
24. Strogatz SH. Nonlinear Dynamics and Chaos. 1994
25. Frederix M, Hütter K, Leu J, Batth TS, Turner WJ, Ruegg TL, Blanch HW, Simmons BA, Adams PD, Keasling JD, Thelen MP, Dunlop MJ, Petzold CJ, Mukhopadhyay A. PLoS One. 2014; 9:e101115. [PubMed: 24983352]
26. Harrison ME, Dunlop MJ. Front. Microbiol. 2012; 3:360. [PubMed: 23112794]
27. Dunlop MJ, Keasling JD, Mukhopadhyay A. Syst. Synth. Biol. 2010; 4:95–104. [PubMed: 20805930]
28. Weiße AY, Oyarzún DA, Danos V, Swain PS. Proc. Natl. Acad. Sci. U. S. A. 2015; 112:E1038–1047. [PubMed: 25695966]
29. Acar M, Mettetal JT, van Oudenaarden A. Nat. Genet. 2008; 40:471–475. [PubMed: 18362885]
30. Garcia-Bernardo J, Dunlop MJ. PLoS Comput. Biol. 2013; 9:e1003229. [PubMed: 24086119]

Insight, innovation, integration

In synthetic biology applications, researchers often rationally design gene circuits to perform a desired function. However, as their complexity increases, it can be difficult to predict exactly how a synthetic circuit will behave. To counter this, researchers can develop libraries and then apply evolution-guided selection to identify the best alternatives. However, synthetic biology applications often require that multiple different goals be met. Here, we use a mathematical model to show that applying dynamic inputs during selection is an efficient way of identifying optimal synthetic circuit designs. Furthermore, we show that the way diversity is introduced into a library can have a dramatic impact on how straightforward or challenging it is to select synthetic circuits with desired properties.

Fig. 1.

Selection using competitive growth quickly differentiates between strains with modest differences in growth rates. (A) Comparison of three strains growth separately in a non-competitive simulation and strains grown together in a competitive simulation. The biomass for the three strains over four selection cycles are shown. For the simulation with strains grown separately, the initial conditions of all but one strain were set to zero and this was repeated for each strain. For the simulation with strains grown together, initial conditions are as described in the Experimental section. The growth rates of the three strains were set to $g_1 = 0.8$, $g_2 = 0.75$, $g_3 = 0.6 \text{ h}^{-1}$. (B) Biomass at the end of each selection cycle of $C = 10$ strains as a function of the selection cycle number. Growth rates g_i are drawn from a normal distribution with mean = 0.5 and standard deviation = 0.05 h^{-1} . Strains are ranked by color, such that the strain with the highest growth rate is shown in red and the lowest is shown in blue. (C) Biomass of the top 10, 5, and 1% of strains as a function of selection cycle. For

these simulations $C = 1000$ strains, growth rates g_i are drawn from a normal distribution with mean = 0.75 and standard deviation = 0.02 h^{-1} . (D) Biomass of the top 10% of strains as a function of the standard deviation of the growth rates. $C = 1000$ strains were competed for six selection cycles. Growth rates g_i are drawn from a normal distribution with mean = 0.5 h^{-1} and standard deviation as indicated in the figure. Insets show representative histograms for two of the growth rate distributions.

Fig. 2.

Dynamic input affects which library members are selected. (A–C) Constant and dynamic inputs applied over six selection cycles. (D–F) Distributions of growth rates and top library members given selection with inputs shown in A–C. All strains are shown in blue; the top 10, 5, and 1% ranked strains are indicated in green, orange, and red. For all, $C = 1000$ strains and six selection cycles were performed. When the system is exposed to input A, the growth rates g_A are normally distributed with mean = 0.75 and standard deviation = 0.02 h^{-1} ; with input B, the growth rates g_B are normally distributed with mean = 0.5 and standard deviation = 0.05 h^{-1} .

Fig. 3.

Selections under dynamic inputs in a growth coupled assay. (A) Growth rate as a function of the input level. These curves are generated using normally distributed values for the maximal growth rate α_i (mean = 1, standard deviation = 0.25 h^{-1}) and half-maximal input levels K_i (mean = 0.5, standard deviation = 0.25). For all simulations $C = 1000$ strains; this figure shows 20 representative curves. (B, D, and F) Inputs applied over six selection cycles. (C, E, and G) Distributions of growth rates at the minimal and maximal input levels, where $I_{\min} = 0.1$ and $I_{\max} = 1$. All strains are shown in blue; the top 10, 5, and 1% ranked strains are indicated in green, orange, and red.

Fig. 4.

Biofuel export model shows that underlying distributions of growth rates may be governed by tradeoffs, but these can be changed with synthetic circuit design. (A) Growth rate as a function of the pump expression level α_p for different concentrations of biofuel. In all cases there is a maximal growth rate where biofuel and pump toxicity are balanced. (B) Biofuel input levels applied over six selection cycles. (C) We used $C = 50$ strains, each with a different pump expression level log-uniformly distributed between 0.1 and 10 h^{-1} . Strains were competed in a selection experiment using the biofuel inputs shown in (B). The green circles indicate strains that are within 10% of the top performing strain; blue circles are those that are not. (D) Dynamic feedback control of efflux pump expression. The growth rate depends on the pump expression rate for all but the case without biofuel, where pumps are never turned on. (E) Strains within 10% of the top performer are shown in green; others in blue.