2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

On the implications of controller resource consumption for the long-term performance of synthetic gene circuits

Daniel P. Byrom and Alexander P. S. Darlington

***Abstract*— Loss-of-function due to mutation presents a fundamental roadblock to the widespread application of engineered biological systems. The onset of mutations in synthetic gene regulatory networks creates parametric uncertainty while the resulting growth competition between mutant populations creates perturbations to the uptake of nutrients. Negative feedback is therefore an attractive strategy to extend the performance of gene circuits over evolutionary time. Here, we develop a mathematical model to evaluate the performance of a simple gene circuit within an evolving population. We show that negative feedback can improve evolutionary longevity. However, when we account for additional host resource consumption by the controller, this benefit can be reversed.**

I. Introduction

Through the engineering of novel gene regulatory networks in cells, it is now possible to design synthetic biological systems with increasing variety, complexity and industrial utility. Modern applications include the production of biofuels via novel metabolic pathways [1], the design of materials capable of sensing and responding to environmental signals [2], and the creation of living therapeutics [3]. These processes, referred to here as 'gene circuits', are encoded in DNA molecules and engineered into cells where this new genetic program can be 'executed' by exploiting the transcription and translation machinery already present in the cells. This co-option of the host's resources towards the expression of the synthetic genes and away from existing host processes typically results in engineered cells having slower growth than their non-engineered counter-parts. This phenomenon is often described as 'burden' and is reviewed extensively in [4] and [5].

As cells reproduce, their DNA is replicated. This is an inherently error-prone process which introduces random mutations into the DNA encoding the engineered process. These mutations can disrupt or even abolish engineered function which can simultaneously reduce burden. Cells with mutated circuits are likely to have a higher growth rate than their ancestral engineered strain. As a result, they can 'outcompete' the ancestral strain and dominate the population. Therefore, over long periods of time, circuit function is eliminated (Fig. 1) [6], [7]. This is a fundamental roadblock to the widespread adoption of engineering biology [8], [9], presenting a new challenge for control engineers [10].

DPB is funded by EPSRC Standard Research Studentship (EP/W52455/1). APSD is funded by the Royal Academy of Engineering under the Research Fellowship scheme. The authors are with the School of Engineering, University of Warwick, Coventry, UK. (correspondence to a.darlington.1@warwick.ac.uk)

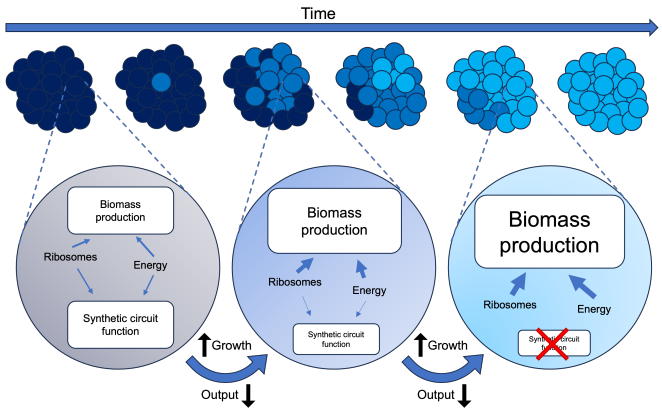


Fig. 1. A simple schematic of an evolving population of engineered cells. Over time, mutants arise in the population with reduced synthetic circuit function and enhanced growth. Faster growing cells come to dominate the population and circuit function is eventually eliminated.

To measure the long-term performance or 'evolutionary longevity' of a gene circuit, we can consider its half-life, defined as the time taken for its function (e.g. protein output) to halve within a population [8]. Many experimental approaches have been suggested to extend the half-life of synthetic constructs [11]. However, even when these are accounted for during the circuit design process, gene circuits are still prone to evolutionary failure [12]. There is increasing interest in developing theoretical frameworks to create circuit typologies that are robust to evolutionary failure. Castle *et al.* recently proposed the concept of the 'evo-type', which describes the 'evolutionary dispositions' of a circuit [13]. New host-aware modeling frameworks which describe host-circuit interactions enable burden to be incorporated into the circuit design cycle [14]–[16]. Combining these tools with models of mutational dynamics can enable the consideration of evolutionary longevity [17], [18].

Loss-of-function mutations primarily affect the promoter sequences of circuit genes [7]. Mechanistically, these mutations affect a promoter's ability to recruit RNA polymerase by altering its DNA sequence, leading to weaker binding by the RNA polymerase and subsequently decreased transcription of circuit genes. From a systems point of view, the introduction of such mutations creates uncertainty in the parameters governing a circuit's mRNA birth rate. As a gene circuit mutates, the host strain becomes subject to a different level of burden, leading to changes in the distribution of the cell's cellular resource economy. This alters the uptake of nutrients between strains, perturbing

the shared environment. We therefore consider the use of negative feedback to increase the long-term performance of gene circuits, due to its ability to provide robustness to both parametric uncertainties and perturbations.

In this paper, we develop a multi-scale model of gene circuit function in the context of host resource utilization, mutation and competition between strains. We apply this model to investigate the potential performance benefits of negative feedback on evolutionary longevity. In Sections II and III, we outline the modeling framework and apply it to a simple gene circuit. We demonstrate how mutation results in loss-of-function at the population level over time. In Section IV, we consider the impact of a simple feedback control motif on the long-term performance of the circuit. In Section V, we show that the additional burden created by controller resource consumption can abolish the benefits of feedback in the long-term.

II. Modeling an evolving population of engineered cellular processes

A. Modeling the impact of gene circuits on their host

Here, we consider a simple synthetic gene circuit consisting of a single, protein-producing gene, which we denote A. This is expressed constitutively, and the quantity of the synthetic protein p_A defines the output of the circuit. We embed this circuit in an established model of microbial metabolism, gene expression and growth to capture the impact of host-process interactions and burden [19]. The model consists of 19 ordinary differential equations (ODEs) describing the time-evolution of each cell's internal components via the key mechanisms of metabolism, transcription and translation. Proteins are distinguished based on their function, divided into the following five coarse-grained types/functions: transport proteins p_T, enzymes p_E, ribosomal proteins (as nonfunctional p_R and functional R), other host proteins p_H and synthetic protein p_A. Model parameters for the host are the same as those used in [19]. Table I outlines each of the 19 variables and parameters for the circuit process. All internal cell variables are measured in molecules per cell (mc) and undergo dilution proportional to the growth rate λ, describing their distribution to daughter cells. The growth rate λ is dynamically calculated within the model as a function of the number of translating mRNA-ribosome complexes c_x and the abundance of the cell's internal energy e (see [19]):

$$\lambda = \frac{\gamma_{\max} e}{M(K_\gamma + e)} \sum_x c_x. \tag{1}$$

Here, $\gamma_{\max}$ and K_γ represent respectively the maximal elongation rate and the elongation energy threshold for transcription. M is the total cell mass, assumed constant.

Metabolism is modeled via an external substrate s, internalized substrate s_I and intermediate energy species e. The external substrate s acts as a nutrient supply and is imported to each cell at a rate $v_{\text{imp}}(p_T, s)$ to generate internal substrate s_I, proportional to the quantity of transport proteins p_T. Internal substrate s_I is subsequently converted to energy e at a rate $v_{\text{cat}}(p_e, s_I)$ scaled by the nutrient efficiency ϕ_e. Both v_{imp} and v_{cat} have a Michaelis-Menten form. The energy species e is consumed by translation at a rate $T_{L_x}(c_x, e)$ (defined in (5)). This yields the following equations:

$$\dot{s_I} = v_{\text{imp}}(p_T, s) - v_{\text{cat}}(p_E, s_I) - \lambda s_I \tag{2}$$

$$\dot{e} = \phi_e v_{\text{cat}}(p_E, s_I) - \sum_x [n_x T_{L_x}(c_x, e)] - \lambda e \tag{3}$$

Gene expression is modeled as follows: for each protein type $x \in \{T, E, R, H, A\}$, there are three variables m_x, c_x and p_x representing respectively mRNA transcripts, mRNA-ribosome translation complexes and complete proteins. mRNA transcripts are spawned at a transcription rate $T_{X_x}(e)$ dependent on the availability of energy e and scaled by the maximal transcription rate ω_x via the energy threshold θ_x. These combine with free ribosomes R to form translating complexes c_x. Translation occurs at a translation rate $T_{L_x}(c_x, e)$ dependent on the availability of energy e and scaled by protein length n_x. The transcription and translation rates are given by:

$$T_{X_x}(e) = \frac{\omega_x e}{\theta_x + e}, \tag{4}$$

$$T_{L_x}(e) = \frac{c_x}{n_x} \frac{\gamma_{\max} e}{K_\gamma + e}. \tag{5}$$

Upon translation, complexes c_x are converted to proteins p_x and the occupied mRNA and functional ribosomes are released. mRNA transcripts degrade at a rate $\delta_{m_x} = 0.1$ min^{-1}. Host proteins do not degrade. Ribosomal rRNA r is produced in the same way as mRNA and combines with ribosomal protein p_R to produce functional ribosomes R which mediate transcription. Applying the law of mass action to this scheme yields the following dynamics, for $x \in \{T, E, R, H, A\}$:

TABLE I

Variables and circuit parameters

Variable	Description
s_I	Internal substrate
e	Energy
${m_x}^*$	mRNA
${c_x}^*$	mRNA-ribosome complex
${p_x}^*$	Protein
r	Ribosomal rRNA
R	Functional ribosome

Parameter	Description	Default Value
ω_A	Maximal transcription rate	50 mc
θ_A	Transcription energy threshold	4.38 mc
n_A	Average protein length	300 aa
b_A	mRNA-ribosome binding rate	0.1 mc^{-1}min^{-1}
u_A	mRNA-ribosome unbinding rate	0.01 mc^{-1}min^{-1}
δ_{m_A}	mRNA degradation rate	0.1 min^{-1}
δ_{p_A}	Protein degradation rate	log(2)/4 min^{-1}

* $x \in \{T, E, R, H, A\}$, representing respectively transport proteins, enzymes, ribosomal proteins, other host proteins and synthetic circuit protein.

$$\dot{m}_x = T_{X_x}(e) + T_{L_x}(c_x, e) - b_x R m_x + u_x c_x - (\lambda + \delta_{m_x}) m_x \quad (6)$$

$$\dot{c}_x = -T_{L_x}(c_x, e) + b_x R m_x - u_x c_x - \lambda c_x \quad (7)$$

$$\dot{p}_{x\backslash R} = T_{L_x}(c_x, e) - (\lambda + \delta_{p_x}) p_x \quad (8)$$

$$\dot{p}_R = T_{L_R}(c_r, e) - \beta_r p_R r + \mu_r R - (\delta_{p_R} + \lambda) p_R \quad (9)$$

$$\dot{r} = T_{X_r}(e) - \beta_r p_R r + \mu_r R - (\delta_r + \lambda) r \quad (10)$$

$$\dot{R} = -\sum_x [T_{X_x}(c_x, e) - b_x R m_x + u_x c_x] + \beta_r p_R r - \mu_r R - (\delta_R + \lambda) R \quad (11)$$

B. Modeling population dynamics

We embed the above host-process model into a simple model describing a population of cells in a chemostat. The population size N obeys simple birth-death dynamics, with births proportional to the growth rate λ and deaths set by the chemostat dilution rate δ. The extracellular substrate s is consumed by each cell at the nutrient import rate v_{imp}. s is replenished at an influx rate k_{in} and diluted at the chemostat dilution rate δ. This yields the additional dynamics:

$$\dot{N} = \lambda N - \delta N \quad (12)$$

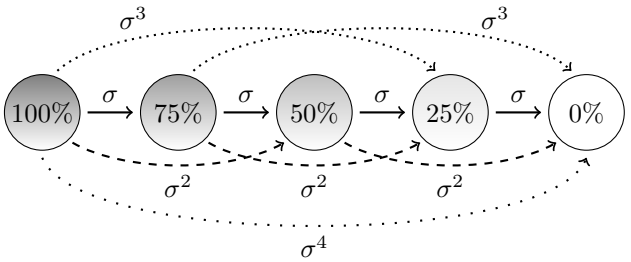
$$\dot{s} = k_{\text{in}} - v_{\text{imp}}(p_T, s) N - \delta s \quad (13)$$

C. Modeling competition between mutant strains

Biological mutation of a gene is the result of a change in its DNA sequence. Among synthetic circuit components, promoters are the most vulnerable to loss-of-function mutations [7]. Such mutations influence the promoter's ability to recruit RNA polymerase and so reduce the production rate of mRNA transcripts. We therefore implement mutation in the model by altering the maximal transcription rate of the process gene ω_A. Possible sequence changes are numerous, each capable of affecting gene expression to a different degree. To account for such mutation heterogeneity without requiring an approach of unbounded complexity, we discretize the possible levels of gene expression into a set of n distinct 'mutation states'. Each mutation state is defined as a separate population using the described model and represents a particular strain of engineered cells. The overall system then consists of one mutation state with designed levels of transcription (100% - fully functional), one mutation state with no transcription (0% - complete loss of function) and $n-2$ mutation states with intermediate levels of transcription. Mutation is modeled by allowing cells to transition between these mutation states. We define transition rates between each pair of states which provide a measure of the likelihood of that particular mutation occurring. If state i mutates into state j at a rate σ_{ij}, then we update the equations for their population size:

$$\begin{aligned} \dot{N}_i &= \lambda_i N_i - \delta N_i - \sigma_{ij} N_i \\ \dot{N}_j &= \lambda_j N_j - \delta N_j + \sigma_{ij} N_i. \end{aligned} \quad (14)$$

To capture all potential transitions, we define an $n \times n$ matrix describing the rates of transition between each pair of mutation states. We assume that mutations can only occur in one direction, reducing the rate of transcription. (If mutations which increase w_A were to occur, such mutants with increased transcription would exhibit slower growth, so would be out-competed by existing strains.) To maintain model simplicity, we define a mutation scheme using a single parameter σ which we call the *global mutation rate*. This describes the rate of transition from state i to state $i+1$. We assume that more extreme mutations are less likely, so that mutations from state i to $i+j$ occur at a rate σ^j. We assume an even spacing of functionality between mutation states. Fig. 2 presents the mutation scheme and mutation matrix for $n = 5$. For this initial study, we set $\sigma = 10^{-5}\text{min}^{-1}$. For the simple systems considered in this paper, varying σ only serves to stretch any of the time-series simulations to occur over shorter or longer time periods, without altering the nature of the results.



$$\Sigma = \begin{pmatrix} 0 & \sigma & \sigma^2 & \sigma^3 & \sigma^4 \\ 0 & 0 & \sigma & \sigma^2 & \sigma^3 \\ 0 & 0 & 0 & \sigma & \sigma^2 \\ 0 & 0 & 0 & 0 & \sigma \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Fig. 2. The mutation scheme used for the model with $n = 5$. Each circle represents a 'mutation state' - a strain of cells distinguished by different maximal transcription rates ω_A. Percentages describe the level of ω_A relative to the designed level. Mutations are represented by arrows between states, with labeled values describing the rate of transition between states. Only mutations which reduce function are allowed. Mutations from state i to state j can be represented by the transition matrix Σ.

Because we are considering n strains within the same chemostat, the modeled mutation states must compete for a single pool of nutrients. We define a single external substrate variable s so that it consumed by each of the n populations:

$$\dot{s} = k_{\text{in}} - \sum_{i=1}^{n} [v_{\text{imp}}(p_{t_i}, s) N_i] - \delta s. \quad (15)$$

This results in a multi-scale model of $20n+1$ ODEs, one per variable per mutation state plus s. A similar approach was recently proposed by Ingram and Stan [18]. In this paper, we incorporate a greater number of mutation states and describe transitions between them via a new mutation scheme.

III. Circuit dynamics with mutation

To understand the impact of mutation on the synthetic gene A, we simulated an open-loop system consisting of a single gene circuit within an evolving population. We considered a circuit with maximal transcription rate $\omega_A = 50$ mc min^{-1}

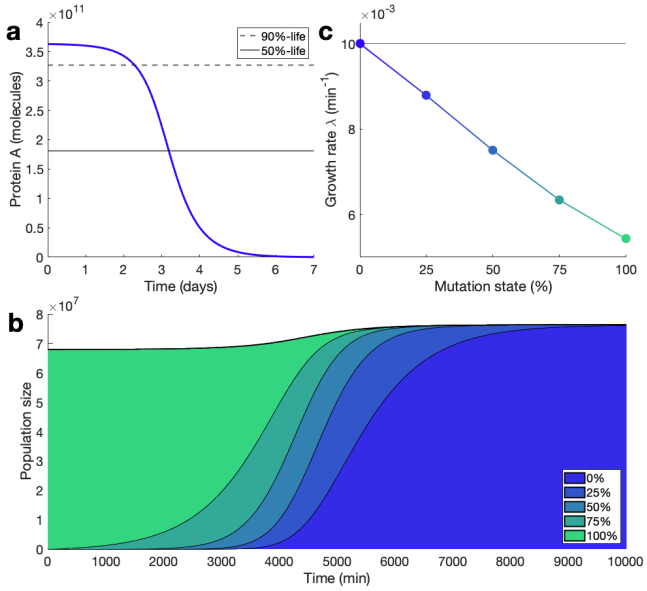


Fig. 3. Simulation of a single gene circuit within an evolving population. Circuit parameters given in Table I. (a) Time-series of synthetic protein A in the population. Horizontal lines display 50% (solid) and 90% (dashed) of the initial output. (b) Time-series of population sizes of the different mutation states. (c) Growth rate of cells at steady state. Possible mutations can be thought of as leftwards movements between points. The growth rate λ of the population as a whole is defined by $\delta = 0.01\text{min}^{-1}$, marked by a horizontal line. At steady state, the population is dominated by the non-functional mutation state which grows at this rate.

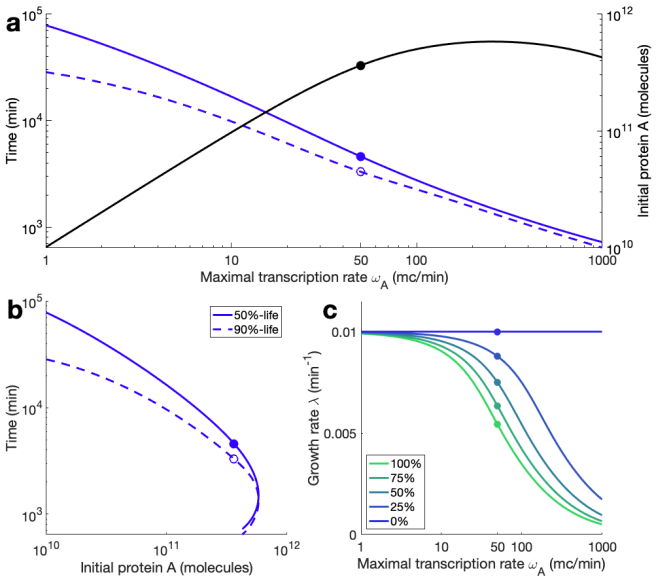


Fig. 4. Outputs of a simple circuit process within an evolving population for a range of maximal transcription rates ω_A. (a) Blue lines display the 50%- (solid) and 90%-lives (dashed) against ω_A. The black line displays the initial population-level quantity of synthetic protein A. (b) 50%- (solid) and 90%-lives (dashed) against initial quantity of synthetic protein A in the population. (c) Steady-state growth rates of the different mutation states across the range of ω_A. In each plot, markers correspond to the system with $\omega_A = 50$ mc min^{-1} from Fig. 3.

and discretized the mutation space into five states (as in Fig. 2), corresponding to maximal transcription rates of 0% (non-functional), 25%, 50%, 75% and 100% (fully functional) of the designed level. We establish the system's initial condition by simulating a population consisting of only fully functional cells to steady state without mutation (time $t < 0$). From time 0, mutation is introduced and the model is simulated for finite time. Over time, the amount of synthetic protein in the population (our process output) falls to zero (Fig. 3a) as the population becomes dominated by non-functional cells. Populations of intermediate function arise and die out over time (Fig. 3b). While δ sets the growth rate of the combined population, the individual mutation states have different growth rates (Fig. 3c). Mutation states with lower protein output display faster growth and so outgrow those with higher protein output over time.

To quantify evolutionary longevity we introduce two metrics: 50%-life (half-life) and 90%-life. The 50%-life is the time taken for the population-wide quantity of synthetic protein p_A to fall below 50% of its original value and is widely used as a measure of long-term performance [11]. The 90%-life is the time taken for the circuit output to fall below 90% of its original value. By considering both the 50%-life and the 90%-life, we can better understand the suddenness with which function is lost. The 90%-life is worth considering in its own right because maintaining high levels of functionality can be important, particularly when considering more complex multi-component systems such as logic gates or biochemical pathways.

We simulated the system for a range of circuit designs by varying the maximal transcription rate ω_A of the fully functional (100%) mutation state across its biologically feasible range. As the rate of transcription increases, the initial quantity of synthetic protein A in the population increases, up to a threshold where additional transcription and translation overburden the cell. As transcription increases, there is a corresponding reduction in both the 50%- and 90%-lives and the difference between them falls, suggesting a more sudden reduction in function (Fig. 4a). There is a trade-off between the initial output and long-term circuit performance (Fig. 4b), recapitulating known experimental results (e.g. [7]). For larger transcription rates, transitions to mutation states with reduced function yield a more significant improvement in growth (Fig. 4c).

IV. Negative Feedback Improves Long-Term Circuit Performance

A. Introducing feedback control

Negative feedback increases a system's robustness to parametric uncertainty. Given mutations alter the transcription rate ω_A, they represent the introduction of parametric uncertainty, making negative feedback an attractive strategy to improve evolutionary longevity. Here, we introduce feedback using a simple auto-inhibition scheme. We modified the original single-gene circuit model so that protein p_A inhibits the transcription of its own mRNA transcripts m_A, creating an autoregulatory negative feedback loop. We updated the ODE for m_A ((6), $x = A$) by scaling the mRNA birth

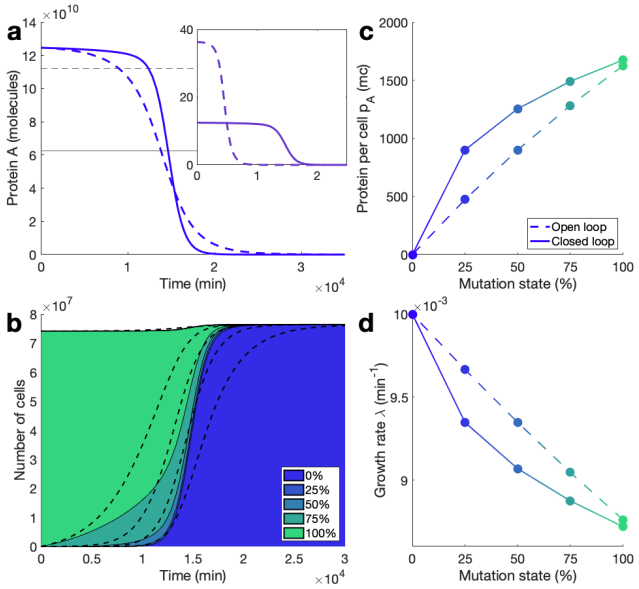


Fig. 5. Impact of negative autoregulation on a single gene process. (a) Time-series of protein A in the population for closed-loop (solid) and open-loop (dashed) systems. Horizontal lines display 50% (solid) and 90% (dashed) of the initial output. The open-loop system is generated by setting the ribosome binding rate to $b_A = 0.014\text{mc}^{-1}\text{min}^{-1}$ so that initial outputs align. The inset shows outputs without altering b_A. (b) Time-series of population sizes of the different mutation states for closed-loop (solid, colored) and open-loop (dashed) systems. (c) Growth rate and (d) per cell protein output p_A for each mutation state at steady state for closed-loop (solid) and open-loop (dashed) systems.

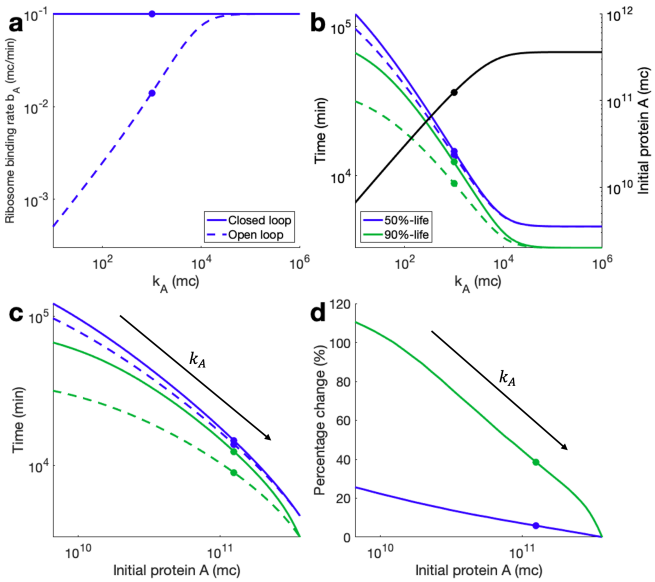


Fig. 6. Designing the negative autoregulation feedback loop by tuning the dissociation constant k_A. (a) Ribosome binding rate b_A for closed-loop (solid) and open-loop (dashed) systems. b_A is tuned in the open-loop system so that initial protein outputs align. (b) Blue and green lines display respectively the 50%-life and 90%-life against k_A for closed-loop (solid) and open-loop (dashed) systems. The black line presents the initial quantity of protein A in the population, set to be equal for both closed- and open-loop systems. (c) 50%-life (blue) and 90%-life (green) against initial quantity of protein A in the population for closed-loop (solid) and open-loop (dashed) systems. (d) Percentage change in 50%-life (blue) and 90%-life (green) against initial protein A output for the closed-loop system relative to the open-loop system.

rate $T_{X_A}(e)$ by an inhibitory Hill function with dissociation constant k_A and Hill constant h_A:

$$\mathcal{R}(p_A) = \left[1 + \left(\frac{p_A}{k_A}\right)^{h_A}\right]^{-1}, \tag{16}$$

$$\begin{aligned}\dot{m}_A = {} & T_{X_A}(e)\mathcal{R}(p_A) + T_{L_A}(c_A, e) \\ & - b_A R m_A + u_A c_A - (\lambda + \delta_{m_A}) m_A\end{aligned} \tag{17}$$

For the circuit design $\omega_A = 50$ mc min^{-1}, we initially chose a nominal controller design of $k_A = 1000$ mc and $h_A = 2$. Introduction of the controller vastly reduces the system's protein output while significantly increasing both the 50%-life and 90%-life (Fig. 5a, inset). Noting the relationship between evolutionary performance and initial protein (Fig. 5b), we chose to assess the performance of the controller against an open-loop system of equivalent output. This is achieved by setting b_A, the ribosome binding rate, to $b_A = 0.014\text{mc}^{-1}\text{min}^{-1}$. The controller yields improvements of 39% in the 90%-life and 6% in the 50%-life when compared to the open-loop system with the same initial output. However, complete loss-of-function (i.e. zero protein output) occurs sooner (Fig. 5a).

Introducing the controller extends the lifespan of the fully functional (100%) mutation state. However, the non-functional (0%) state arises sooner, with the intermediate states making up a greatly reduced fraction of the overall population (Fig. 5b). The autoregulation motif drives gene expression in proportion to the protein output (with a set point created by k_A). Intermediate mutation states (with lower ω_A) therefore show increased protein production compared with the open-loop system, as the feedback controller forces the process to reach the set point (Fig. 4c). This reduces the growth rates of the intermediate mutation states compared with the open-loop system (Fig. 5d). Therefore, mutating into these intermediate states provides a smaller growth benefit, while mutating to the final non-functional state from any other state yields a greater boost in growth. This explains how the 90%-life can be improved while complete loss-of-function is expedited.

B. *Varying controller design*

Having demonstrated that a nominal feedback controller can improve the long-term function of a simple circuit, we next considered how to design the controller. h_A represents the co-operativity of the controller protein binding to the promoter of gene A. As this is difficult to engineer *in vivo*, we maintain $h_A = 2$ throughout. k_A represents the dissociation constant of the controller protein binding the promoter of gene A and can be varied experimentally by changing the DNA sequence of the binding site. We varied k_A across a wide, biologically feasible range between 10 and 10^6 mc and considered the impact on three metrics: (i) initial quantity of protein A in the population, (ii) 50%-life and (iii) 90%-life. In each case, b_A was tuned so that the initial open-loop and closed-loop protein outputs align (Fig.

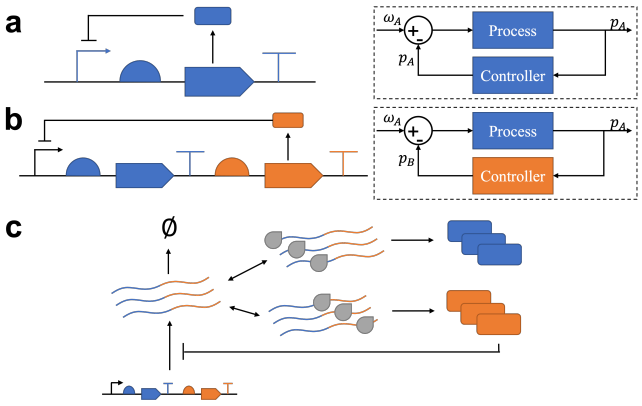


Fig. 7. Implementing negative feedback biologically. Blue components correspond to process gene A. Orange components correspond to controller gene B. Controller proteins inhibit the transcription of the co-produced genes. Insets, block diagrams for the corresponding systems. (a) Autoregulation. Here the process gene functions as its own controller. (b) More conventional negative feedback. Here the process and controller components are separate but share a promoter. (c) Resource consumption in the two gene process-controller system. m_{AB} (blue-orange lines of mRNA), c_A (mRNA with ribosome connected to process RBS), c_B (mRNA with ribosome connected to controller RBS), p_A (blue boxes) and p_B (orange boxes).

6a). There is a trade-off between initial output and long-term performance for both the open- and closed-loop systems (Fig. 6b,c). For any choice of k_A, feedback yields an improvement in both the 50%- and 90%-lives. The largest performance increases are achieved with low k_A, with the 50%-life improving by 20% and the 90%-life more than doubling (Fig. 6d). These controller designs correspond to more tightly binding repressor proteins. Across all possible controllers, autoregulatory negative feedback makes our simple circuit robust to evolutionary loss-of-function.

V. Accounting for Controller Burden

A. Controller burden reduces performance

Autoregulation (above) relies on the synthetic process being able to repress itself. Biologically, the process gene *A is its own controller*. This is achievable only if the circuit output is an inhibitory transcription factor (Fig. 7a). This severely limits process topology and narrows the breadth of potential applications. Therefore, we consider an alternative design consisting of biologically separate process and control components (Fig. 7b). Here, process output gene A and controller protein gene B are co-produced from a single promoter. Protein p_B acts as an inhibitory transcription factor on the shared promoter, completing a feedback loop. mRNA transcripts contain genetic information for both proteins. Ribosomes can attach to either of two ribosome binding sites (RBSs) and translate the corresponding protein (Fig. 7c).

We model this as follows: transcription produces an mRNA transcript m_{AB} that encodes both process output gene A and controller protein gene B. The mRNA binds with ribosomes R at one of two RBSs to produce mRNA-ribosome complexes c_A and c_B. Translation yields the corresponding proteins p_A and p_B. Controller protein p_B inhibits the transcription of m_{AB} via a Hill function $\mathcal{R}(p_B)$ with parameters k_B and h_B as in (16). ODEs describing c_A and p_A are unchanged from the original system ((7), (8), $x = A$) and c_B and p_B obey equations of the same form. We replace (17) for m_A with the following:

$$\begin{aligned}\dot{m}_{AB} = T_{X_{AB}}(e)\mathcal{R}(p_B) + T_{L_A}(c_A, e) + T_{L_B}(c_B, e) \\ - b_A R m_A + u_A c_A - b_B R m_B + u_B c_B \\ - (\lambda + \delta_{m_A}) m_A\end{aligned} \tag{18}$$

Additional controller parameters are set to correspond to those in Table I. This control mechanism exerts an additional drain on the host's resources, because additional proteins p_B must be translated.

Maintaining a process $\omega_A = 50$ mc min^{-1} with protein length $n_A = 300$ amino acids (aa), we considered a nominal controller with $k_B = 1000$ mc, $h_B = 2$. To tune the controller resource consumption, we varied the protein length n_B between 0 (equivalent to autoregulation) and 600 aa (twice the process length). Varying n_B does not qualitatively impact the shape of the time-series output curve; the closed-loop systems initially maintain a more consistent, close-to-maximal output, but see a steeper decline (Fig. 8a). As n_B increases, the fall in protein output occurs sooner, resulting in a reduction of both the 50%- and 90%-lives. The 50%-life falls for all biologically feasible protein lengths ($n_B > 15$ aa), but the 90%-life improves when the controller size is less than half that of the process ($n_B < 150$ aa). This suggests that, although burden-free feedback always improves performance, additional controller resource consumption can in fact reduce the longevity of a process.

B. Varying controller design

We investigated the relationship between the initial protein output and the 50%- and 90%-lives by varying the dissociation constant k_B, again maintaining $h_B = 2$. Both the 50%- and 90%-lives decline across the design space as n_B increases (Figs. 8b,c). As in Section IV, controllers with reduced output (smaller k_B, tighter binding repressor proteins) perform better relative to the open-loop system. However, the addition of controller resource consumption can cause the 50%- and 90%-lives to fall below those of the open-loop system (Figs. 8d,e). The 50%-life can be improved for $n_B < 150$ aa, but only for designs where little output is produced. With initial outputs smaller than 1.9×10^{10} mc ($k_B < 75$ mc), the 90%-life increases for all controller sizes. For the largest of protein outputs ($k_B > 10^4$ mc), improvement is only possible if the controller size is close to zero. For many controller designs, it is possible to achieve an improvement in the 90%-life alongside a reduction in the 50%-life. The effectiveness of feedback control is dependent on the design of the controller, via both its resource consumption (n_B) and the strength of inhibition of the controller proteins (k_B). If these aren't selected carefully, the inclusion of feedback control can negatively impact long-term function.

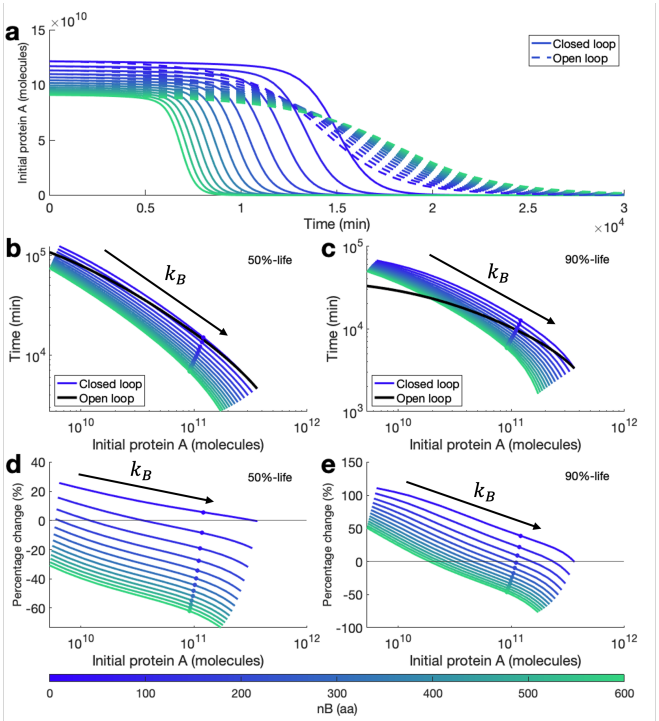


Fig. 8. Impact of controller resource consumption on evolutionary performance. Controller protein length n_B is varied between 0 and 600 aa. Circles in (b,c,d,e) correspond to the system in panel (a). (a) For a controller with $k_B = 1000$ mc, $h_B = 2$, time-series of population-level protein A. Solid lines represent closed-loop controllers. Dotted lines represent open-loop systems where b_A is tuned so that both systems have the same initial protein. Color represents n_B. (b,c) For a range of controllers ($10 \leq k_B \leq 10^6$ mc), initial protein A in the population against (b) 50%-life and (c) 90%-life. Black lines represent relationships for the open loop system. (d,e) Plots of initial quantity of protein A in the population against percentage change relative to open-loop in (d) 50%-life and (e) 90%-life.

VI. Conclusions

In this paper, we developed a multi-scale modeling framework incorporating gene expression, host constraints, mutation (as discrete mutation states) and competition between strains to assess the evolutionary longevity of a simple gene circuit. This framework identifies the trade-off between the expression of a simple synthetic circuit and long-term function previously identified experimentally. We show that negative feedback increases the evolutionary longevity of gene circuits (i.e. increases in 90%-life and 50%-life) but that function is abolished sooner. When we account for controller resource consumption, both the 50%-life and 90%-life steadily decrease as resource consumption increases. At extremes, this resource consumption can result in closed-loop performance that is *worse* than open-loop. For our circuit, increases in the 90%-life are greater than increases in the 50%-life (compared with an equivalent open-loop system) and the 90%-life is less sensitive to increasing controller resource consumption. This suggests that, through the engineering of negative feedback, it is easier to boost short-term performance than long-term performance. Whilst feedback confers evolutionary robustness to gene circuits, careful attention must be paid to controller resource consumption during design. We are now determining how to design controllers with minimal resource consumption to ensure controller production does not reduce 50%-life, and applying our approach to more complex gene circuits. Our work here shows that host-aware models can be used to design 'host-friendly' control schemes which maximize performance but minimally affect the host [20]. Our results have implications for the design of robust gene circuit devices where long-term performance over several weeks may be crucial, with applications in biomedicine and industrial biotech.

References

[1] Z. Liu, J. Wang, and J. Nielsen, "Yeast synthetic biology advances biofuel production," *Current Opinion in Microbiology*, vol. 65, pp. 33–39, 2022.

[2] T.-C. Tang, B. An, Y. Huang, *et al.*, "Materials design by synthetic biology," *Nature Reviews Materials*, vol. 6, no. 4, pp. 332–350, 2021.

[3] A. Cubillos-Ruiz, T. Guo, A. Sokolovska, *et al.*, "Engineering living therapeutics with synthetic biology," *Nature Reviews Drug Discovery*, vol. 20, no. 12, pp. 941–960, 2021.

[4] O. Borkowski, F. Ceroni, G.-B. Stan, and T. Ellis, "Overloaded and stressed: Whole-cell considerations for bacterial synthetic biology," *Current opinion in microbiology*, vol. 33, pp. 123–130, 2016.

[5] A. Boo, T. Ellis, and G.-B. Stan, "Host-aware synthetic biology," *Current Opinion in Systems Biology*, vol. 14, pp. 66–72, 2019.

[6] S. C. Sleight, B. A. Bartley, J. A. Lieviant, and H. M. Sauro, "Designing and engineering evolutionary robust genetic circuits," *Journal of biological engineering*, vol. 4, no. 1, pp. 1–20, 2010.

[7] S. C. Sleight and H. M. Sauro, "Visualization of evolutionary stability dynamics and competitive fitness of escherichia coli engineered with randomized multigene circuits," *ACS synthetic biology*, vol. 2, no. 9, pp. 519–528, 2013.

[8] B. Zakeri and P. A. Carr, "The limits of synthetic biology," *Trends in Biotechnology*, vol. 33, no. 2, pp. 57–58, 2015.

[9] S. Kumar and J. Hasty, "Stability, robustness, and containment: Preparing synthetic biology for real-world deployment," *Current Opinion in Biotechnology*, vol. 79, p. 102 880, 2023.

[10] D. Del Vecchio, Y. Qian, R. M. Murray, and E. D. Sontag, "Future systems and control research in synthetic biology," *Annu. Rev. Control*, vol. 45, pp. 5–17, 2018.

[11] B. A. Renda, M. J. Hammerling, and J. E. Barrick, "Engineering reduced evolutionary potential for synthetic biology," *Molecular BioSystems*, vol. 10, no. 7, pp. 1668–1678, 2014.

[12] A. Hossain, E. Lopez, S. M. Halper, *et al.*, "Automated design of thousands of nonrepetitive parts for engineering stable genetic systems," *Nature biotechnology*, vol. 38, no. 12, pp. 1466–1475, 2020.

[13] S. D. Castle, C. S. Grierson, and T. E. Gorochowski, "Towards an engineering theory of evolution," *Nature Communications*, vol. 12, no. 1, p. 3326, 2021.

[14] A. Y. Weiße, D. A. Oyarzún, V. Danos, and P. S. Swain, "Mechanistic links between cellular trade-offs, gene expression, and growth," *Proceedings of the National Academy of Sciences*, vol. 112, no. 9, E1038–E1047, 2015.

[15] C. Liao, A. E. Blanchard, and T. Lu, "An integrative circuit–host modelling framework for predicting synthetic gene network behaviours," *Nature microbiology*, vol. 2, no. 12, pp. 1658–1666, 2017.

[16] J. Kim, A. P. Darlington, D. G. Bates, and J. I. Jimenez, "The interplay between growth rate and nutrient quality defines gene expression capacity," *BioRxiv*, pp. 2021–04, 2021.

[17] T. Ellis, "Predicting how evolution will beat us," *Microbial biotechnology*, vol. 12, no. 1, p. 41, 2019.

[18] D. Ingram and G.-B. Stan, "Modelling genetic stability in engineered cell populations," *Nature Communications*, vol. 14, no. 1, p. 3471, 2023.

[19] A. P. Darlington, J. Kim, J. I. Jiménez, and D. G. Bates, "Dynamic allocation of orthogonal ribosomes facilitates uncoupling of co-expressed genes," *Nature communications*, vol. 9, no. 1, p. 695, 2018.

[20] A. P. Darlington and D. G. Bates, "Performance and robustness analysis of control strategies for ameliorating cellular host-circuit interactions," in *2022 European Control Conference (ECC)*, IEEE, 2022, pp. 233–239.