Modular, robust and extendible multicellular circuit design in yeast

Alberto Carignano¹, Dai Hua Chen¹, Cannon Mallory¹, Clay Wright³, Georg Seelig¹,², Eric Klavins¹

¹Department of Electrical & Computer Engineering, University of Washington; ²Paul G. Allen School of Computer Science & Engineering, University of Washington; ³Department of Biological Systems Engineering, Virginia Tech

Abstract Division of labor between cells is ubiquitous in biology but the use of multi-cellular consortia for engineering applications is only beginning to be explored. A significant advantage of multi-cellular circuits is their potential to be modular with respect to composition but this claim has not yet been extensively tested using experiments and quantitative modeling. Here, we construct a library of 24 yeast strains capable of sending, receiving or responding to three molecular signals, characterize them experimentally and build quantitative models of their input-output relationships. We then compose these strains into two- and three-strain cascades as well as four-strain bistable switch and show that experimentally measured consortia dynamics can be predicted from the models of the constituent parts. To further explore the achievable range of behaviors, we perform a fully automated computational search over all two-, three- and four-strain consortia to identify combinations that realize target behaviors including logic gates, band-pass filters and time pulses. Strain combinations that are predicted to map onto a target behavior are further computationally optimized and then experimentally tested. Experiments closely track computational predictions. The high reliability of these model descriptions further strengthens the feasibility and highlights the potential for distributed computing in synthetic biology.

Introduction

The leading paradigm for genetic circuit design is to combine biological parts in a delicate balance within the same cell [Ellis et al. (2009), Kosuri et al. (2013), Ottoz et al. (2014)]. This approach has resulted in increasingly large genetic circuits that realize functions such as logic gates and circuits [Moon et al. (2012), Bonnet et al. (2013), Nielsen et al. (2016), Gander et al. (2017)], time pulses (Gao et al. (2018), Guo and Murray (2019)), incoherent feed-forward loops (Entus et al. (2007), Ellis et al. (2009)), bistable switches (Chen et al. (2012), Huang et al. (2012), Yang et al. (2019), Barbier et al. (2020), Grant et al. (2020)) or oscillators (Elowitz and Leibler (2000), Tigges et al. (2009), Tigges et al. (2010)). Albeit very successful, this approach shows its limitations when it comes to scalability and robustness. Albeit very successful, single cell circuit engineering has limited scalability and robustness because parts cannot be reused, the genetic burden on the cell grows with circuit size Nikolados et al. (2019), and retroactivity Del Vecchio et al. (2008) and component crosstalk Del Vecchio (2015) interfere with expected behavior.

An alternative approach is to generate complex behaviors using consortia of cells wherein different cell types perform distinct functions and communicate with each other through chemical signals to realize more complex behaviors. Such division of labor is common in nature and interest has recently emerged in engineering multi-cellular synthetic systems (Brenner et al. (2008)).
Although developed later, synthetic consortia have caught up with single-cell circuits in terms of the complexity of functions that have been realized, generating behaviors such as multicellular time pulses (Basu et al. 2004), oscillations (Chen et al. 2015), and logic gates (Regot et al. 2011, Tamsir et al. 2011), bioproduction Egbert et al. (2017), or circuits that use quorum sensing to define social interactions (Kong et al. 2018). Furthermore, the emergence of novel orthogonal cross-species signaling molecules (Billerbeck et al. 2018, Du et al. 2020) has opened the path for engineering more complex and precise multicellular behaviors.

To date, synthetic multi-cellular systems were largely designed with specific target behaviors in mind, rather than optimizing modularity of components to make them usable in a large number of contexts. Still, mathematical modelling has demonstrated that multi-cellular computation should easily access a larger space of behaviors: for instance, just three different cell populations can, in theory, generate up to 100 different logic functions (Regot et al. 2011) and even bimodality and cell-synchronization (Thurley et al. 2018). Furthermore, the intuitive modular nature of cell-to-cell communication should provide a useful tool to rationally design synthetic circuits with predetermined performance. Rational design significantly speeds up circuit assembly (Chen et al. 2012, Chen et al. 2020) and allows design of global functions from local behaviors (Salis et al. 2009, Corrothers et al. 2011). As a matter of fact, mathematical models have been successful to bridge the gap between individual processes and collective behaviors when applied to ecological interactions between distinct populations of communicating cells [Shou et al. (2007), Momeni et al. (2013), Egbert et al. (2017)]. However, an example of a model-driven strain selection for multicellular circuit design is currently missing.

Here we propose a large vocabulary of yeast strains that use chemical signals for cell-to-cell communication and that can be modularly combined to realize a large number of functions (1A). Each strain senses one or two inputs and produces a single output. The transfer function relating inputs and outputs can be either activating or repressing. The output is a fluorescent protein which can be used to read out circuit behavior, another signaling molecule which can be used to connect strains or an enzyme that sequesters or degrades a signal thus disrupting communication. We experimentally characterize each strain, model their dynamics using differential equations and then use these models to predict the behavior of strain combinations (1B). Using a library of 24 strains, we rationally design complex multicellular behaviors, such as bandpass filters, negative and positive feedbacks, time pulses, logic gates, and bistable switches. These behaviors are common in nature and they underlie cell decisions on metabolism (for instance, responding to environmental signals, Lee et al. 2002, utilization of carbon sources, Brandman et al. 2005) or cell differentiation (Kueh et al. 2016, Duddu et al. 2020), and have been repeatedly connected to multicellular organization (signal gradients in development) and signaling (auxin pulses in plants). The modularity of composition achieved by our multicellular circuits allow us to design these non-trivial behaviors based on model simulations alone. This qualitative and quantitative precision demonstrated in both time and steady-state experiments further extend circuit design automation (Nielsen et al. 2016, Chen et al. 2020).

Results

Engineering yeast strains for signal sensing, synthesis and depletion

As a first step towards the construction of a vocabulary of modular yeast strains we selected a set of signals to enable cell-to-cell communication and then optimized chassis strains for signal sensing, synthesis and depletion. We selected the plant hormone auxin (we will also refer to it as IAA, indole-3-acetic acid) and the yeast hormone α-factor as signaling molecules and additionally used the mammalian hormone β-estradiol as an inducer.

The α-factor pathway has been subject of extensive studies and mechanisms for sensing through the surface receptor STE2, synthesis by the MFα1 gene and degradation through the BAR1 protease are well understood and have been engineered to generate a wide range of behaviors (Youk and
To have a baseline strain that does not interfere with signaling, we knocked out the native BAR1 gene as in Youk and Lim (2014) to prevent α-factor degradation. Moreover, to account for growth-arrest induced by α-factor, which affects gene expression on a large scale, we knocked out FAR1 a protein that contributes to arresting the cell cycle at G1 Chang and Herskowitz (1990), and constitutively expressed POG1, a protein that promotes growth-arrest recovery Leza and Elion (1999). Unexpectedly, we detected no α-factor output from strains that both sense and secrete α-factor. We suspected that this was caused by the surface receptor STE2 internally binding to α-factor. It has been shown that, upon binding to α-factor, STE2 undergoes endocytosis and then shares the secretory pathway of α-factor itself Schandel and Jeness (1994). We solved this problem by overexpressing STE2 on a pGPD promoter Sun et al. (2012) in these strains, aiming to have some protein copies escaping this interaction.

Next, we turned to the optimization of components for Auxin sensing, synthesis and elimination. In prior work from our group, we (Khakhar et al. (2016), Pierre-Jerome et al. (2014)) developed an auxin-responsive transcription factor using a chimeric dCas9-Aux/IAA protein regulation. In the presence of Auxin, the Aux/IAA degron part of the protein binds to the auxin signaling F-box protein (a modified TIR1 for this study) and acts as part of an E3 ubiquitin ligase to catalyze ubiquitination and degradation of the Aux/IAA-degron-containing protein. Similarly, we constructed a synthetic auxin synthesis pathway in yeast Khakhar et al. (2016). We demonstrated conversion of the precursor molecule IAM (indole-3-acetamide) into IAA through expression of the IaaH gene in yeast. Here, we amplified Auxin secretion and thus effectively the strength of the signal produced, through integration of the auxin-efflux pump PGP1 from A.thaliana into our yeast strains, as previously reported in Geisler et al. (2005). We measured a significant increase in the auxin-synthesis yield as measured by a neighbouring IAA-detecting cell when PGP1 was integrated (SI Figure1).

Unlike for α-factor, a depletion mechanism for Auxin had not yet been reported in yeast. To create an IAA depletion mechanism, we thus selected the plant protein GH3.3 that has been shown to conjugate IAA to aspartic acid, forming the signaling-inactive IAA-Asp. GH3.3 is part of a family of proteins that encode IAA-amino synthetases, which have been reported to control auxin homeostasis Staswick et al. (2005). To test whether GH3.3 or related proteins could be used to inactivate Auxin in yeast, we first expressed codon-optimized versions of GH3.3 and GH3.6 from A.thaliana and C.papaya from a highly-expressed pGPD promoter. We then tested the IAA to IAA-Asp conversion rate using mass spectrometry and found that GH3.3 from A.thaliana had the higher efficiency (SI Figure2, SI Figure3). Finally, we tested that IAA-Asp does not activate the IAA-mediated degradation pathway in yeast, by adding 10uM of IAA-Asp in an auxin-sensor yeast culture and detecting no variation in fluorescence (SI Figure2).

Establishing a vocabulary of parts for cell to cell communication

Having established conditions for efficient signal sensing, synthesis and inhibition, we combined signals with activating and repressive transfer functions to create a vocabulary of strains. Transfer functions that use α-factor as input are mediated by the transcription factor (STE2) that either directly induces expression of the gene of interest (activation) or induces expression of a repressor that inhibits the output (repression). Similarly, β-estradiol binds and activates a transcription factor (ZEV4) that either directly activates gene expression (activation) or induces expression of a repressor and inhibits output synthesis (repression). For both signaling molecules, we chose dCas9 fused with the repressor domain Mxi1 (Gander et al. (2017)) as repressor (SI Figure 4 for the full pathways).

To induce activation or repression with auxin, we build on the same auxin-mediated degradation pathway used for auxin sensing above. Specifically, activation results from degradation of a repressor, while repression results from degradation of an activator. We chose a dCas9-Mxi1-αuxin degron Khakhar et al. (2016) fusion as a repressor and used a dCas9-VP64-αuxin degron fusion as the activator (SI Figure 4).

We conducted extensive pathway optimization to increase the separation between high and low
Figure 1. Modular components for engineering multi-cellular signaling circuits. (a) Three input signals, two transfer functions and five outputs are used to assemble 24 distinct strains. (b) Differential equations are used to model and predict behavior of all strains and strain combinations in the paper. In this example, a model (Model 1) and symbolic representation for a fluorescent reporter repressed by high auxin concentration are shown. (c) Left: Time course fluorescence data for different auxin concentrations for the sensor strain shown in (b). Full lines are model simulations. Right: End point fluorescence data is shown as a function of auxin concentration color-matching the time series on the left. The steady-state simulation is shown in orange. (d) Symbolic representation, steady state data and model fit for all other single input reporter strains. (e) Modeling framework (Model 2) and fluorescence kinetics data for two different two-input reporter strains.

expression levels and IAA sensitivity, using a mechanistic model to explore the parameter space and guide the genetic engineering (SI Figure 5). We adopted the model proposed in Pierre-Jerome et al. (2014), where each parameter easily translates to a biological function, and performed parameter sensitivity analysis with respect to fold change between the baseline and the Aux/IAA-induced fully-repressed state. We then selected the top six highest scoring parameter perturbations, designed circuit variants that reflected those changes, and tested them resulting in a good agreement with our predictions (SI, Figure 5). Guided by the model, we combined five of the tested circuit variants to increase the fold change from 1.3-fold (sensor from Khakhar et al. (2016)) to 3.1-fold: we used this final circuit for all the repressive strains, swapping the fluorescent reporter gene with the output gene for non-sensor strains. For the auxin activating strains, we used a similar model-driven
Combining the three different input signals, the activation/repression circuits and output secretion, we built and tested all the possible combinatorial designs presented in 1A. The strains that sense β-estradiol and repress expression of BAR1 and GH3.3, the strain that senses α-factor and represses expression of GH3.3, and the strain that senses IAA and represses BAR1 expression are shown in the SI (SI Figure 6) and not used below, since their response was too slow to produce meaningful results. Of the remaining twenty-four strains, six sensor strains express GFP in response to the three input signals (three activators and three repressors), twelve strains (six activators and six repressors) synthesize a signaling molecule and six strains act as signal attenuators (expressing BAR1 or GH3.3). Four of these twenty-four strains sense and secrete the same signaling molecule (α-factor or IAA, ‘positive’ or ‘negative’ feedback strains). Finally, two strains express repressors of their own input (α-factor expressing BAR1 and IAA expressing GH3), also describing a negative feedback topology (SI Figure 6).

**Sensor strain characterization**

For sensor strain characterization, we collected time series data for eight different input concentrations. Input concentrations were selected to fully cover the sensor dynamic range for model fitting. We normalized fluorescence data by cell size and took the mean of the histogram as in Groves et al. (2016) and then subtracted background fluorescence. Each measurement in the figure is an average of three experimental repeats (error bars representing the standard deviation).

With a scalable and modular system in mind, we fitted a set of three ordinary differential equations (ODEs) with eight parameters for each strain to describe input sensing, signal processing and fluorescence output synthesis (1B, Model 1). Signal processing (activation or repression) is modelled with a simple Hill function (Model 1b), which naturally incorporates signal saturation. Input sensing and output synthesis are modelled as linear ODEs (Model 1a and c). A constant term in the last ODE accounts for background promoter activity. These simple models capture the system dynamics, with the benefit of being easy to fit and analytically approachable. Parameters were fitted independently for each experimental repeat to obtain a mean and a standard deviation for the Hill coefficient. We also separately fitted a Hill curve to an average of the three experimental repeats and the resulting Hill coefficient (n) is reported in the figures and used for simulations (1B). The sensor strains range in sensitivity depending on the input. The β-estradiol sensors respond to inputs concentration ranging from 0.1 to 100nM with an EC50 of 0.5±0.0nM for repression and 12.6±0.9nM for activation. α-factor sensors are sensitive to input concentrations ranging from 1-500nM range with an EC50 ~ 6.0±0.4nM for activation and 89.0±6.6nM) for repression. Finally the IAA sensor is least sensitive and responds to inputs ranging from 5nM-10μM with an EC50 ~ 964.8±93.8nM for activation and 276.5±8.8nM for repression. The Hill coefficients for these sensors vary between 0.8 (3-repeat interval: 0.8 ± 0.0, α-factor repression) and 1.0 (1.0 ± 0.0, α-factor activation), 2.2 (2.1 ± 0.3, β-estradiol repression) and 1.3 (1.25 ± 0.0, β-estradiol activation), 1.0 (1.0 ± 0.4, auxin repression) and 0.8 (0.8 ± 0.0, auxin activation) consistent with previously reported values for β-estradiol repression Gander et al. (2017) and α-factor activation Groves et al. (2016). Finally, the sensors achieve an ON/OFF fold-change that ranges between a minimum of 3 (IAA activating GFP) and a maximum of 42 (β-estradiol activating GFP).

In addition to single-input sensors, we constructed two sensor strains that respond to both α-factor and IAA. In the first strain, α-factor induces fluorescence expression and IAA represses it, while in the second strain, IAA induces fluorescence expression and α-factor represses it. We collected data points for eight different input combinations taken at five time points up to saturation. The output signal is monotonic with respect to each input and the input functional range is similar to the one measured for the correspondent strains that respond to only one of the two inputs.

(1E).
Figure 2. Multi-strain signaling cascades. (a) Symbolic representation of a two-strain cascade that uses α-factor as the input and auxin as an intermediate signal. To model multi-strain cascades, models for individual strains are concatenated and only the last differential equation is fit. End-point fluorescence data and fit are shown for the example cascade. (b) By varying the concentration of the upstream strain, the strength of the signal seen by the downstream strain can be predictably controlled. This change in signal is modeled with a single parameter K. Right: data and model predictions for three experiments with the same strains but varying concentrations of the upstream strain. (c) Symbolic representation, model and data for a two strain cascade wherein the upstream strain removes a signal rather than secreting it. (d) Symbolic representation, model and data for all two strain cascades where the upstream input is activating the production of the intermediate signal. (e) Symbolic representation, model and data for all two strain cascades where the upstream input is repressing the production of the intermediate signal. (f) Symbolic representation, model and data for three-layer signaling cascades.
To model these strains, we used a simple extension of our previously introduced models with five ODEs (1E, Model 2): two ODEs model input sensing (a), two model input processing (b), and one ODE (c) combines the signals according to their activating or repressing nature and defines output synthesis. These models fit the experimental data even when the output is non-monotonic over time.

**Assembling modular, tunable and easily-extendable circuits using cell-to-cell communication**

To determine the potential to build biological circuits using cell-to-cell communication, we experimentally tested if communication occurs between strains that secrete an output and their corresponding sensor strains. For example, a strain that produces auxin in response to \(a\)-factor sensing was grown in coculture with a sensor that switches off GFP expression in response to auxin(1B).

The two strain populations were added at the same concentration and the fluorescent output was measured at steady state (10 hours after induction). The experimental data is consistent with a negative \(a\)-factor sensor as expected (the more \(a\)-factor, the lower the fluorescent signal, as seen in 2A).

Since the core genetic circuit of this sender strain is identical to the \(a\)-factor repressing sensor, we tested if the mathematical model previously fit to the sensor data preserves its predictive power. We re-fit only the 3 output parameters to account for the fact that the output is now auxin rather than GFP. The output of the sender cell model was used directly as input of the sensor cell model. To test the hypothesis that strains that have common input/processing parts share parameters and model structure, we collected two datasets composed of four data points each: we used one for fitting (yellow dots), and the other for validation (orange dots, 2A). A time-series at EC50 input concentration was also used for fitting (not present in the figure).

Intuitively, we expected that higher initial sender cell concentration would result in an overall higher concentration of their output signal over the same time scale. Most importantly, we wanted to know if this output ‘gain’ can be predicted by our models so that we could use it to tune circuit behavior. We modelled this effect with a factor \(K\) multiplied by the output signal, where \(K\) is the fold-change with respect to the standard initial concentration: we represented this gain using the same iconography used in electronic circuits (2B). Our predictions matched closely with data collected using 5X and 10X the original concentration of sender cells (green and blue lines) using the same strain co-culture as in the previous panel (orange line).

We further tested if we could modulate the concentration of signaling molecules by removing it from the system through BAR1 and GH3.3-expressing strains (2C). We modelled signal degradation as a first-order Hill repressing function, where the output of the sender cell acts as a negative regulator. As before, only the three output parameters of the sender strain were fitted using receiver cell fluorescence data. Finally, we tested all of the sender-receiver pairs in our vocabulary with the exception of those generating positive or negative feedback. All activating strains function within the sensor range of their receiver strains (2D), and the models correctly fit or predict the data. On the other hand, the output of repressing strains (2E) did not fully cover the input range of their receiver strains, as seen in the limited fold-change of the fluorescent signal. Models suggested that increasing the sender strain population to ten-fold its original value would produce a more noticeable response, which we successfully verified experimentally (purple dots and lines, 2D and E).

After the 2-strain combinations, we also verified the predictive power of our model on two 3-strain chains with different topologies and different strain stoichiometries (2F). Here too, the simple model strategy we outlined earlier captured the overall dynamics even when different strain concentrations were used. These results support the hypothesis that multicellular circuits behave like a sum of their individual parts (modularity), are easy to modulate (whether through altering initial strain concentrations or signal degradation) and can be extended to longer chains.
Increasing nonlinear response using external positive feedback

Thus far, we explored ways to simulate and design multicellular circuits with tunable gains to obtain monotonic, quasi-linear dynamic systems with a single equilibrium point. To extend the range of observable behaviors and generate non-linear responses to the inputs, we used positive feedback strains that sense and secrete the same signaling molecule (3A).

To highlight the increase of nonlinear response, on top of fitting models to the positive feedback strains (as in 2A), we also re-fitted the sensor strains to estimate a new Hill coefficient (much as in Shaw et al. (2019)). We then plotted these fitted curves, the data points and the two Hill coefficients for both the feedback system and the nominal response (in this case, the sensor alone). In both cases, the positive feedback increased the non-linearity of the response (from 0.8±0.0 to 1.5±0.1 and from 1.0±0.4 to 1.2±0.3 for the α-factor and the auxin case respectively). We also considered positive feedback circuits that operate through double repression (3B). In this case, we tested topologies where either α-factor represses BAR1 synthesis, or auxin represses GH3.3 synthesis. Ideally, at low signaling molecule concentration, signal degradation through BAR1 or GH3.3 is predominant so there is no fluorescence response in the receiver cells. But at high input concentration, the degradation pathway is switched off and the input is free to reach the receiver cells. As in 3A, we measured the Hill coefficients of these circuits and reported an increase in nonlinear response (from 0.8±0.0 to 1.0±0.0 and from 1.0±0.4 to 1.4±0.6 for α-factor and auxin circuits respectively).
We next turned to the construction of a bistable switch circuit. We opted to use a mutual-repression topology to generate bistability (Gardner et al. (2000), Oyarzún and Chaves (2015)). For a first design, we combined the strain that senses α-factor and represses auxin synthesis with the strain that senses auxin and represses α-factor synthesis. Here, the main state variables are signaling molecule concentration in the media rather than the internal state of the cells (specifically, high auxin/low α-factor and low auxin/high α-factor). However, both model simulation and experimental data showed that this circuit can generate only a single equilibrium, independently of the strain stoichiometry because of a lack of non-linearity in the system (SI, Figure 7).

To boost non-linearity, we added two strains to the mix that induce signal degradation: a strain that senses α-factor and synthesizes GH3 and one that senses auxin and synthesizes BAR1 (3C). The resulting system can still be seen as two modules that repress each other’s activity: α-factor lowers auxin concentration (through pathway repression and GH3 expression) and auxin reduces α-factor concentration (through pathway repression and BAR1 expression). We studied this new system with a steady-state model for auxin and α-factor concentrations (3C, Model 3 and SI Figure 8 for model derivation). The model suggests that the Hill coefficient is higher than 2, a necessary condition for the existence of more than one equilibrium.

We investigated the existence and properties of the equilibria while varying the individual concentrations of the four strains in the circuit: these variations are captured by the four $K$ parameters in the model, representing the gains of the four strains in the circuit as explained earlier. Using the model, we identified a range of concentrations predicted to result in multiple equilibria. For further investigation, we picked a set concentrations that maximize the distance between equilibria such that the equilibria are robust to small variations in strain concentrations (red diamond, 3D). This solution generates two nullclines that intersect three times, resulting in 2 stable (red crosses) and 1 unstable (black empty circle) equilibrium as expected (3E).

We tested this model-guided design experimentally (3E). Strains were mixed according to the chosen concentrations and an auxin negative sensor strain was added. Upon reaching steady state (Time 0 in 3F), samples were diluted at regular intervals to prevent the cells from saturating. To alternate between the states, we exogenously added first auxin (at 3 hours) and then α-factor (at 15 hours) and let dilution reduce their concentration to below detectable levels for our sensor (at 12 and 24 hours, respectively).

Model 3 simulation and data largely agree, although there is a lag between the two. We hypothesize that the lag is due to a signaling delay between the strains that was not fully captured by the models in this five-strain circuit. More importantly, the ‘high’ equilibrium is stable in the first 3 hours of the experiment, but seems to become unstable after 25 hours, as shown by a negative trend in the data toward the ‘low’ equilibrium. Modelling suggests that a small increase in the $K_a$ gain (strain: IAA expressing BAR1) would lead the system to a single ‘low’ equilibrium (SI Figure 9). This could occur if the corresponding strain grows slightly faster than the other strains (although, in practice, we could not detect any growth difference between the single-input/single-output strains, SI Figure 10).
Figure 4. Model-generated implementation of Boolean logic gates. (a) We used an automated search algorithm to screen all possible strain combinations up to size four (and including exactly one sensor strain) for their ability to realize a set of logic functions. In the example, each strain combination is scored according to how close the combination is to realizing an ideal Boolean AND gate. Each colored diamond corresponds to one strain combination. Circuits consisting of two strains (+ a sensor) are shown in blue, three-strain gates are shown in red, and four-strain gates in yellow. All strains in these simulations are at equal stoichiometry. Higher scores indicate more AND-like behavior. The top-performing strains (score > 1, teal box) are selected for further computational optimization of strain stoichiometry. Optimized strain combinations have higher AND scores. Optimized circuits consisting of two strains are shown in light blue, three-strain gates are shown in green, and four-strain gates in purple. Simulations are separated according to which sensor strain is used. A specific high-scoring four-strain combination was chosen for experimental testing (red box). The experimental data (blue bars) show good quantitative agreement with the predictions (orange bars). (b) Similar optimization procedures to the ones shown in (a) were used to identify strain combinations that realize NOR, NAND and OR logic functions. Example implementations, model predictions and experimental data are shown for all three logic functions.

Automated design of strain circuits to generate logic gates

In order to expand our target behaviors, we developed an automated approach to select circuits using the twenty-four strains introduced above. Specifically, we simulated all possible strain combinations for networks of size 2, 3 and 4 strains using Models 1 and 2. For each network, we simulated the system response over 12 hours to all possible combinations of α-factor (in the discretized...
range of [0 1 5 10 50 200 1000] nM), auxin ([0 100 500 1000 5000] nM), and β-estr ([0 1 5 10 100] nM). The ranges were chosen to properly sample the operational range of the strains (see 1C, D, E).

Next, we screened our simulation space for steady-state behaviors whose profiles resemble AND, OR, NAND or NOR logic gates. For each strain combination, we restricted the steady-state behavior space to simulations obtained using combinations of the highest input concentrations (α-factor=1000uM, auxin= 5000µM, β-estr=100nM) or no input to match the [01] logic table input entries. This selection resulted in 120 combinations from the 2-node networks, 560 from the 3-node networks, and 1820 from the 4-node networks. After normalization to allow for comparison between different sensors (SI Figure 11), each of these combinations was scored according to how well they match one of the logic gate truth tables. If the predicted output for one of the expected OFF states was higher than the output for one of the expected ON states, then the metric would return a value below 1 labeling that strain combination to be a poor logic gate realization. On the other hand, values above 1 imply a match between the output vector and the target truth table: the higher the value, the better the separation between the ON and OFF state.

We first applied this method to identify circuit topologies that generate AND gate profiles (4A). Each network topology is represented as a diamond, color-coded according to the network size (blue for 2-node, red for 3-node and yellow for 4-node networks), and divided in 4 groups according to the sensor strain reporter. Each network was scored, and its value reported on the x-axis.

Next, we selected all the network topologies that scored higher than 1.0 and performed an optimization step. Optimization aimed to maximize the target metric using strain concentration (gains) as optimization parameters (4A, left panel). Finally, we picked the highest scoring topology for experimental testing. The best AND gate was a 4-node network topology that used α-factor and β-estradiol as inputs and a negative auxin sensor to determine the output. This strain combination included two strains repressing auxin output in the presence of α-factor and β-estradiol respectively. These strains by themselves should ideally generate an AND gate in concert with the negative auxin sensor. The other two strains improve performance of the AND gate, likely by reducing the effect of leaky auxin synthesis from the α-factor-sensing/auxin-repressing strain. The experimental realization of this circuit shows separation between the ON and OFF states. In fact, the data (blue bars) seem to slightly outperform the predictions (4A, right panel).

We repeated this same procedure for NOR, NAND and OR gates (4B). The optimal NOR gate is also a four-node network, while the optimal NAND gate is a three-node network, and both use the same auxin sensor as the optimal AND gate. The bar separating ON and OFF states as defined for the AND gate holds for all the gates sharing the same sensor. The optimal OR gate is the naive realization with two strains synthesizing auxin from different inputs and an activating an auxin sensor. All the gate profiles are close to their predicted values from simulation, showing the high degree of modularity of our vocabulary of strains, even for complex systems with internal feedback as the NOR gate architecture. For each gate, we also tested the optimal realization of the strain combinations for the two remaining network sizes with similarly positive results (SI Figure 12).
Figure 5

A) Time Series behaviors

B) Steady State behaviors

C) Optimal circuit identification

D) Stoichiometry optimization

E) IAA (nM)

F) \( \alpha \)-factor (nM)
Identification of circuit designs for time pulses and band pass filters

We next extended our automated design strategy to circuits that either generate time pulses or that acts as band pass filters on the input signal concentrations. Starting from the simulation dataset of all possible strain combinations, we selected all those that displayed non-monotonic behaviors (having at least one local maximum/minimum) as a function of time or as a function of the steady state input concentration. We defined a non-monotonicity metric as the distance between the local maximum/minimum and the maximum/minimum between the initial value and the final value of the series. Higher values of the metric hint to more pronounced non-monotonic behaviors, while 0 implies that no local minimum/maximum is present (5A). We then selected the top six candidates (one for each sensor type in 1B, D and E, excluding the β-estradiol ones), and performed optimization to maximize the metric using cell concentrations and input concentration as parameters. Finally, we experimentally tested the top two topologies overall for both time pulses and steady-state band-pass filters. It is worth noticing that only 0.32% of all possible topologies across network sizes generated non-monotonic behaviors in time (of about $10^6$ in total, 5A) and only 0.32% at steady state (of about $10^5$ in total, 5B). Hence, a ‘brute force’ experimental approach to test all possible strain combinations would be evidently out of reach.

The highest performing time pulse topology (5C) is induced by α-factor that activates both fluorescence expression and auxin synthesis. In turn, auxin induces BAR1 production, which degrades the exogenous α-factor signal: unsurprisingly, this is an incoherent feed-forward loop, type 1 (as in Mangan and Alon (2003)). The optimal ‘reversed’ time pulse, i.e. a dip in the output at intermediate times (5D), responds to α-factor induction and implements a modified incoherent feed-forward loop type 3, where α-factor both represses and activates fluorescences. Model predictions suggested that three different α-factor concentrations would generate this behavior at different capacities (0.1nM, 1nM and 5nM α-factor). According to the models, both these nonlinear behaviors are a consequence of delay between an activator and an inhibitor pathway.

The low-pass concentration filter (5E) uses very similar components but the sensor with opposite topology. In this case, α-factor activates fluorescence expression but also auxin synthesis, which then activates its own expression in a positive feedback. This topology is also an incoherent feed-forward loop (type 1) and results in a band-pass filter output. Finally, the band-stop filter (5F) uses both β-estr and α-factor as inputs, and the circuit topology presents a combination of negative feedback and an incoherent feed-forward loop (type 3). The circuit operates as a ‘reversed’ band-pass filter over α-factor concentration, while β-estr is kept constant and only adds a baseline of α-factor and auxin through the two top strains in the circuit.

The experimental data qualitatively agreed with the predictions, although we notice a time delay for the first two circuits and a shift in baseline for the latter two. The time delay is caused by underestimation of sensor strain activation due to difficulty in cytometrically isolating the sensor.
strain from the other strains (5C) and possibly by an undermodelled delay of auxin synthesis (5D). The baseline shift is consistent with a higher concentration of auxin in the system in both cases. As a matter of fact, these two experiments (5E,F) ended at a higher cell concentration than usual (running time of 12 and 14 hours respectively), and it is known that yeast natively synthesizes auxin at saturation (Rao et al. 2010).

Discussion

Here, we demonstrated the potential of synthetic multicellular circuits to generate a wide range of behaviors starting from simple activating or repressing individual strains. Instead of implementing complex circuits in isogenic populations, we designed simple monotonic circuits in different strains and allowed them to communicate using just two signaling molecules (α-factor and auxin) or to affect their environment by attenuating those signals (through BAR1 and GH3.3). These simple constructs alone were capable of recapitulating many behaviors previously realized with synthetic gene circuits such as bistability, band-pass filters, pulses and logic gates. Moreover, our last results on automated identification of circuit topologies that realize a target behavior (5) hint that the space of possible functional circuit architectures is larger than we explored.

We demonstrated through an extensive use of mathematical models, that these synthetic multicellular circuits are modular, easy-to-tune and extendable. Modularity is achieved through cell-cell communication that avoids cross-talk, and is demonstrated by combining input and output of our simple models. Multi-cellular circuits are tuned using different cell concentrations or positive-feedback architectures. Finally, we realized that we could tune the circuit behavior by extending or shortening the length of the signaling chain by adding or removing intermediate strains. We exploited this property to build the two time-pulses (Figure5, A and B): a slower repression or activation dynamic allows for the opposite signal to operate first, generating the non-monotonic behaviors we observed. We imagine this property will gain more practical applications when the number of signaling molecules increases, which is the current major limitation in our system.

We reported that our most complex circuits display some discrepancies between simulations and experimental data, whose explanation is not always clear. It is possible that the delay between sending and receiving is not adequately modelled, or that factors that were not modelled, such as individual cell differences within cell populations, are affecting those circuits. Future work should account for these factors at the modelling steps, for instance using distributions to describe input-output relationships Thurley et al. (2018) or more complex models.

Finally, we leveraged the mathematical description to define an automated method to design behaviors according to performance specifications. Computationally, the method simulates all the possible strain combinations given a fixed number of nodes, and then scores them according to how well they reproduce the behavior of interest. As it is defined, the method is not easily scalable because the number of total simulations increases exponentially with the number of strains and does not account for differences in the initial strain populations. However, we found no catastrophic failures in the experimental validations, such as a qualitatively different behavior, owing to the modularity of our system. Future efforts should be directed towards more efficient ways to simulate the networks, for instance by training a neural network on the current simulation sets to predict the output of interactions.

Method

Construction of yeast strains

Yeast transformations were carried out using a standard lithium acetate protocol used by (Gander et al. 2017). Yeast cells were made competent by growing 50ml cultures in rich media to log growth phase, then spinning down the cells and washing with H2O. Next, linearized DNA, salmon sperm donor DNA, 50% polyethylene glycol and 1M LiOAc were combined with cell pellet and the mixture was heat shocked at 42° for 15 min. The cells were then spun down, supernatant was removed and
they were resuspended in H₂O and then plated on selective agar media. Transformations were done into MATa W303 – 1A.

**Cytometry**

Fluorescence intensity was measured with a BD Accuri C6 flow cytometer equipped with a CSampler plate adapter using excitation wavelengths of 488 and 640nm and an emission detection filter at 533nm (FL1 channel). A total of 10,000 (20,000 for the bistable switch samples) events above a 400,000 FSC-H threshold (to exclude debris) were recorded for each sample using the Accuri C6 CFlow Sampler software. Cytometry data were exported as FCS 3.0 files and processed using a custom Python script to obtain the mean FL1-A value for each data point.

**Data collection for sensor strains**

Synthetic complete growth medium was used to grow the cells overnight from glycerol stock, while 300μM IAM (Indole-3-acetamide) was added in all the experimental medium (also synthetic complete). Experiments involving time course data were taken during log phase via the following preparation: 16 hrs of overnight growth in the synthetic complete medium in a 30° shaker incubator followed by dilution to 30 events/μL into fresh, room-temperature medium. After 10 hrs of growth at 30°, we performed a new dilution to 30 events/μL in 3ml of medium, added the inducers, and started collecting 100μL samples for measurements periodically until the completion of the experiment. Ten thousand events were collected for each condition.

**Data collection for co-culture experiments**

Sample preparation was conducted as described above with each strain in a separate test tube. At the start of the experiment, each strain concentration was first measured and then strains were combined at the concentration specified by the experiment. We considered the concentration of 30 events/μL as the baseline concentration for all the experiments and all concentrations are expressed as multiples of this reference concentration. Once the samples were added together at the desired concentrations, inducers were added and measurements were taken periodically until the completion of the experiment as described for experiments with sensor strains. For the duration of the experiment, the samples were kept in a 30° shaker.

**Data collection for Bistable Switch data**

The sample preparation was conducted as described above, each strain in individual test tubes. Then, each of the five strains in the bistable switch were combined together at the concentration as outlined in the main text in a 3 ml test tube kept in a 30°C shaker for the duration of the experiment. Every 40 minutes, we performed a 1/3 dilution with fresh media with 300μM concentration of IAM, through manual pipetting. Samples of 120μL were collected for the duration of the experiment approximately every 3 hours.

**Model fitting procedure and simulations**

Model parameters were estimated using Matlab© fminsearch function to minimize the L2-norm of the difference between observations and simulations. For sensor strains in 1, each parameter was estimated three times on three different experimental repeats to identify mean and standard deviation. Then, the parameters used for all the simulations in this study were estimated by fitting the average of the three measurements. Model parameters for the non-sensor strains were estimated by minimizing the L2-norm of the difference between the simulations and the average of three experimental repeats as data points. Models were simulated using the Matlab© ode15s function.

**Codes and data availability**

Codes and data are available at https://github.com/Alby86/MulticellularYeast.git.
Acknowledgments

We are grateful to Cameron Cordray, Samer Halabiya, and Klavins lab technicians for technical support and to Mitchell Szeto and Alex Carr for help with python scripts.

Funding

This work was supported by ONR award N00014-16-1-3189 to G.S and E.K.

References


Youk H, Lim WA. Secreting and sensing the same molecule allows cells to achieve versatile social behaviors. *Science*. 2014; 343(6171).