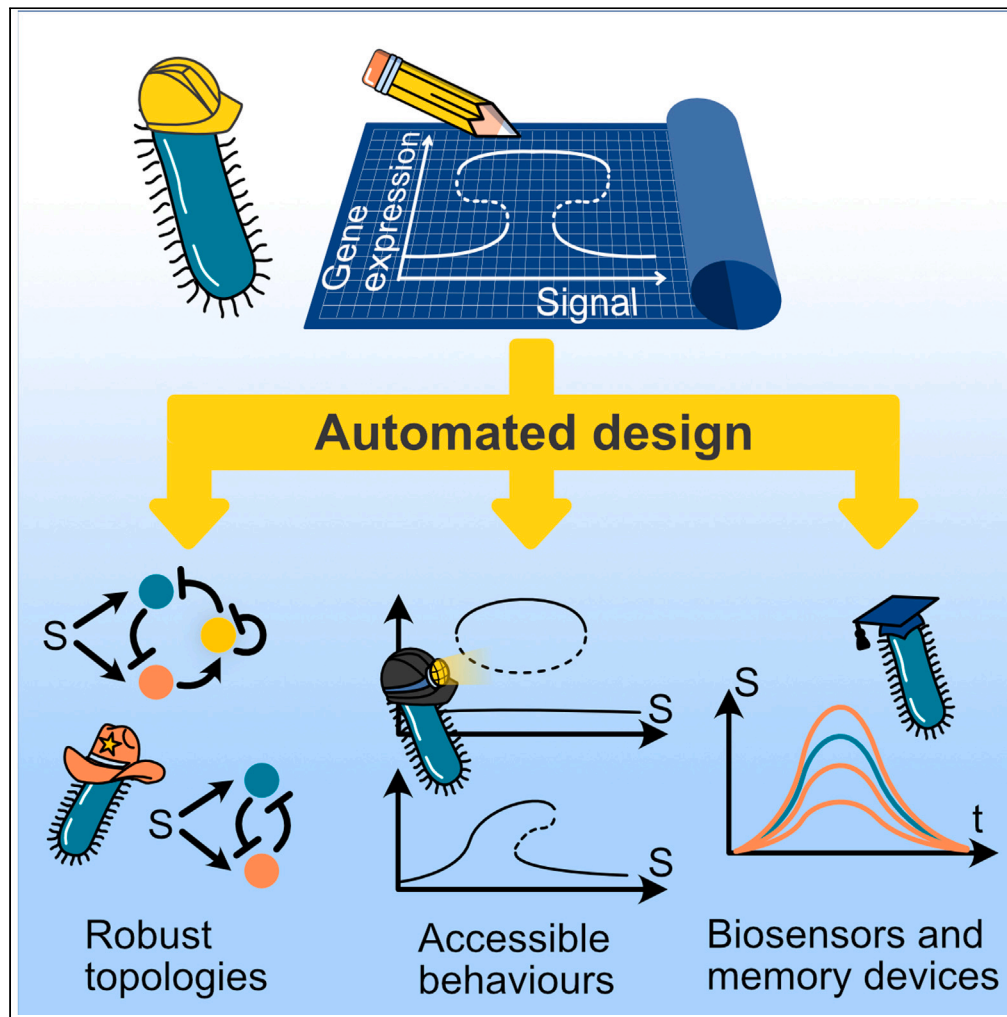**Article**

Automated design of gene circuits with optimal mushroom-bifurcation behavior

Irene Otero-Muras, Ruben Perez-Carrasco, Julio R. Banga, Chris P. Barnes

irene.otero.muras@csic.es (I.O.-M.)
r.perez-carrasco@imperial.ac.uk (R.P.-C.)

Highlights

Optimization enables gene circuit design with target bifurcation diagrams

The method finds robust topologies and kinetic parameters for bespoke circuit dynamics

Optimization on the mushroom bifurcation shows cross-repression as most robust topology

Robust mushroom topologies show capabilities as biosensor and memory devices

Article

Automated design of gene circuits with optimal mushroom-bifurcation behavior

Irene Otero-Muras,[1,*] Ruben Perez-Carrasco,[2,5,*] Julio R. Banga,[3] and Chris P. Barnes[4]

SUMMARY

Recent advances in synthetic biology are enabling exciting technologies, including the next generation of biosensors, the rational design of cell memory, modulated synthetic cell differentiation, and generic multifunctional biocircuits. These novel applications require the design of gene circuits leading to sophisticated behaviors and functionalities. At the same time, designs need to be kept minimal to avoid compromising cell viability. Bifurcation theory addresses such challenges by associating circuit dynamical properties with molecular details of its design. Nevertheless, incorporating bifurcation analysis into automated design processes has not been accomplished yet. This work presents an optimization-based method for the automated design of synthetic gene circuits with specified bifurcation diagrams that employ minimal network topologies. Using this approach, we designed circuits exhibiting the mushroom bifurcation, distilled the most robust topologies, and explored its multifunctional behavior. We then outline potential applications in biosensors, memory devices, and synthetic cell differentiation.

INTRODUCTION

Real-world and cutting-edge applications of synthetic biology are demanding circuit designs with increasingly complex behaviors. Toward building the synthetic cell from the bottom-up, *new developments are expected*, quoting,[1] *when designing cellular systems featuring complex behaviors, including division, cognition, and motility.*

One of the main challenges of synthetic biology is to design and implement gene regulatory circuits capable of complex behaviors in a near-optimal fashion while keeping a minimal design.[2,3] A milestone in gene circuit automated design is CELLO[4] enabling the design of circuits with pre-specified steady state and input-output behaviors, and for the first time proving good predictability in living cells of model-based automated design software. Aiming to address more complex (and dynamic) behaviors, tools based on mixed integer nonlinear programming have shown good flexibility and computational efficiency.[5,6]

The limited resources of the cell restrict the combination of multiple working circuits in the same organism.[7,8] This gives leading relevance to the design of multifunctionality, regulatory networks capable of distinct dynamical behaviors. But, how can different behaviors be integrated in the same circuit? How to endow a cell with the capacity to respond differently to a signal leading to complex dynamical behaviors? The bifurcation theory of dynamical systems provides powerful tools to answer these questions, enabling the mapping between the topology of the network (given by a set of parametrized differential equations) and the different dynamics available under a controllable input or signal. Each bifurcation of the system changes the number and/or nature of the long-term dynamics of the system, e.g., from monostable to bistable, or from a stable steady state to an oscillator. These phenomena can be represented by bifurcation diagrams, showing how the number, position, and dynamics of each steady state change under a set of controllable parameters. However, standard tools for bifurcation analysis are based on continuation algorithms which require precise *a priori* knowledge of parameters and steady-state solutions, hampering the integration of bifurcation diagrams within automated algorithms for circuit design. Alternative methodologies based on chemical reaction network theory[9–11] have paved the way for the integration of bifurcation theory into the automated design of biocircuits, an ambitious design framework that has not been addressed so far.

[1]Computational Synthetic Biology Group. Institute for Integrative Systems Biology (UV, CSIC), Spanish National Research Council, 46980 Valencia, Spain

[2]Department of Life Sciences. Imperial College London, London, UK

[3]Computational Biology Lab, MBG-CSIC, Spanish National Research Council, 36143 Pontevedra, Spain

[4]Department of Cell and Developmental Biology, University College London, London, UK

[5]Lead contact

*Correspondence:
irene.otero.muras@csic.es
(I.O.-M.),
r.perez-carrasco@imperial.ac.uk (R.P.-C.)

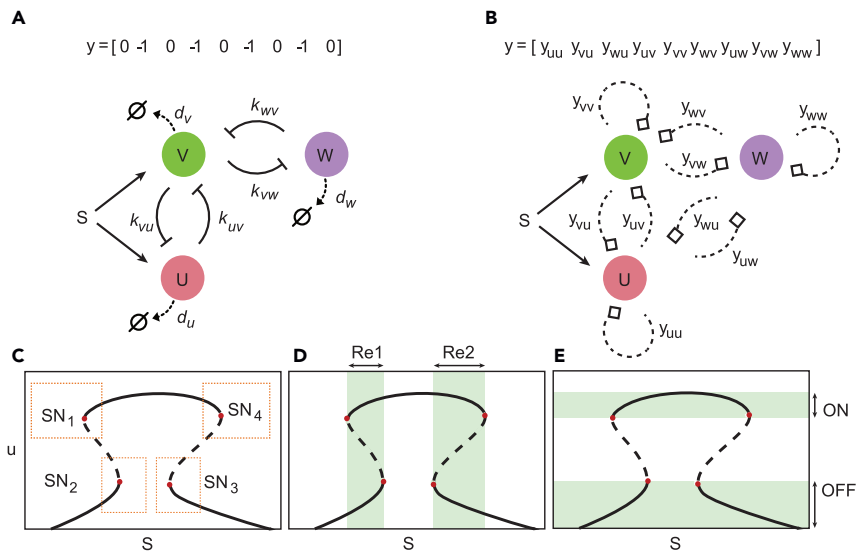https://doi.org/10.1016/j.isci.2023.106836

Figure 1. Searching for gene circuit topologies leading to mushroom bifurcation behavior

(A) Gene network containing two pairs of cross-repressing genes. The affinity parameter k_{ij} controls the strength of the regulatory interaction of gene i on gene j, and d_i is the degradation rate constant for each protein species i.

(B) The circuit superstructure employed in this work to search for mushroom behavior: super graph and vector encoding.

(C) The mushroom bifurcation diagram shows four saddle-node bifurcations (indicated by red dots). The dotted squares indicate possible constraints on the state-parameter space to be imposed on the search for a desired mushroom bifurcation behavior.

(D and E) (D) Two different bistability ranges (Re1, Re2) and (E) "ON state" and "OFF state" ranges (shadowed).

In this work we address this last topic and present a method for automated design of gene circuits with pre-specified bifurcation diagrams. The method extends and combines efficient global mixed integer nonlinear programming optimization methods with an innovative procedure for bifurcation detection, allowing the following novel features:

(i) Automated design of gene circuits that not only exhibit behaviors compatible with a target bifurcation diagram but are also optimized for specific additional criteria (like metabolic cost) given by sets of functions.

(ii) Systematic exploration of minimal topologies compatible with the required behavior.

(iii) Robust design i.e. finding robust topologies with respect to parameter perturbations (i.e., topologies that remain functional despite perturbations in the parameters).

(iv) More sophisticated design tasks, computing optimal trade-offs between design objectives and finding Pareto optimal topologies with respect to different opposing criteria (such as performance and metabolic cost, or robustness and topological complexity).

A multifunctional behavior of special interest for synthetic biology applications is the mushroom bifurcation, named after its mushroom-shaped bifurcation diagram, originally identified in the differentiation of neural stem cells.[12–14] The mushroom can result from the combination of two toggle switches, giving rise to four saddle-node bifurcations and a set of three disconnected loci of stable steady states (Figure 1). The intermediate steady state (termed "ON state") is only available for a window of intermediate values of the signal, while the other (termed "OFF states") are available for high and low values of the signal. Note that despite the "ON/OFF" nomenclature, we do not require a certain threshold in the expression levels to define a mushroom. Specifically, the expression levels along each locus can vary with the signal without compromising the bifurcation diagram structure. Similarly, the pattern of steady-states of the bifurcation diagram of the mushroom is a multidimensional feature and its categorization can be independent of the gene that is used to describe it.

In certain aspects, the phenotypic behavior of the mushroom is an extension of "band-detect" gene regulatory networks, which are constructed from incoherent feedforward loops.[15,16] These have recently been

applied to model French flag type pattern formation[2] and other programmed spatial behavior.[17] The mushroom-shaped locus of equilibria, leading to two different ranges of bistability provides the system with unique hysteresis properties where the state of the cell will be determined by the signal history. Bistability (and multistability in general) is in itself a property of interest in synthetic biology, microorganisms, and mammalian cells,[18–20] forming the basis of memory, cell decision-making systems, biological computation, and pattern formation.[21–24] In addition, we can explore the phenotypic space close to the mushroom bifurcation. In particular, mushroom topologies can also lead to isola bifurcation diagrams, which, though well-known in dynamical systems theory, have never been demonstrated experimentally. Together, these properties demonstrate how combining common motifs gives rise to an emergent range of dynamical behaviors not described by the individual components, a growing area of interest in systems biology.[3,25] Hence the contribution of this study is 2-fold: on the one hand incorporating for the first time target bifurcation behaviors in the automated design of biocircuits, while on the other exploring the capabilities of the identified robust mushroom topologies showing how automated exploration can be connected to functional design.

RESULTS

Minimal topologies leading to mushroom bifurcation

The method presented in this work allows for the general design of biocircuits with prescribed saddle-node bifurcations in target regions of the bifurcation diagram (pre-defined by the user). Here we apply this methodology for the design of biocircuits exhibiting mushroom bifurcations with particular gene expression levels for the saddle-node bifurcations. Given that the mushroom bifurcation diagram contains four saddle-node bifurcations, we expect that a network containing two pairs of cross-repressing genes (Figure 1A) can give rise to the target bifurcation diagram. Nevertheless, we hypothesized that simpler topologies with fewer connections can also yield the same functionality. In order to explore this idea we defined the superstructure in Figure 1B. Starting from this superstructure, which includes 3 genes (U, V, and W) with a signal S inducing genes U and V, we looked for circuits with a prescribed mushroom characteristic bifurcation diagram. There are 9 potential connection arrows between the nodes (genes) represented by a vector y of integers, such that $y_{ij} = -1$ if gene j inhibits gene i, $y_{ij} = +1$ if gene j activates gene i and zero otherwise. The mushroom characteristic bifurcation diagram is depicted in Figure 1C, with 4 saddle-node points (SN) delimiting the steady-state branches and generating two bistability regions ($Re1$ and $Re2$ in Figure 1D). This results in an "ON state" available only for intermediate values of the signal (see Figure 1E). Note that, by "ON state" and "OFF state" we denote branches of stable the steady states in the bifurcation diagram (where "ON" and "OFF" states correspond to intermediate values of the signal, and low and high values of the signal, respectively). The particular search employed in this manuscript allows for "ON" states with lowers expression levels than the "OFF" states i.e. *inverted mushrooms*. Nevertheless, the method introduced in this manuscript allows us to restrict also the particular location of the saddle nodes (see Figure 1C), which can be especially useful in the design of specific biosensors (see STAR Methods).

We search for topologies leading to a mushroom bifurcation diagram behavior using an optimization algorithm and a multistart strategy (as described in the STAR Methods section). In order to explore minimal topologies leading to the target behavior, we first impose a number of 2 genes and a number of 2, 3, and 4 connections.

From the potential 72 connected topologies, only 7 topologies were found leading to mushroom behavior, represented in Figure 2 and classified attending to the number of active connections. Topologies, A1, A2, and A3 were the topologies that appeared more frequently in our search. This suggests that the cross-repression motif is a robust way to obtain the mushroom bifurcation diagram. In particular, the minimal structure leading to a mushroom bifurcation (A2) only requires two nodes, resembling the paradigmatic cross-repressing topology encoding the bistable switch.[18] Note that symmetric structures such as A1 and A3 are considered different since the symmetry of the model might be broken in the particular conditions imposed. Here, we imposed a more strict range of possible values on gene u for the location of the saddle-node bifurcations (as indicated STAR Methods section).

In addition to the cross-repression motif, inspection of all the successful topologies revealed 2 more core topologies (for which no connection can be removed without losing the mushroom functionality), corresponding to structures A5 and A7. Interestingly, this global topology screen did not return the mushroom topology studied in[14] and[26] that is similar to structures A5 and A7 but without a self-activation. A specific
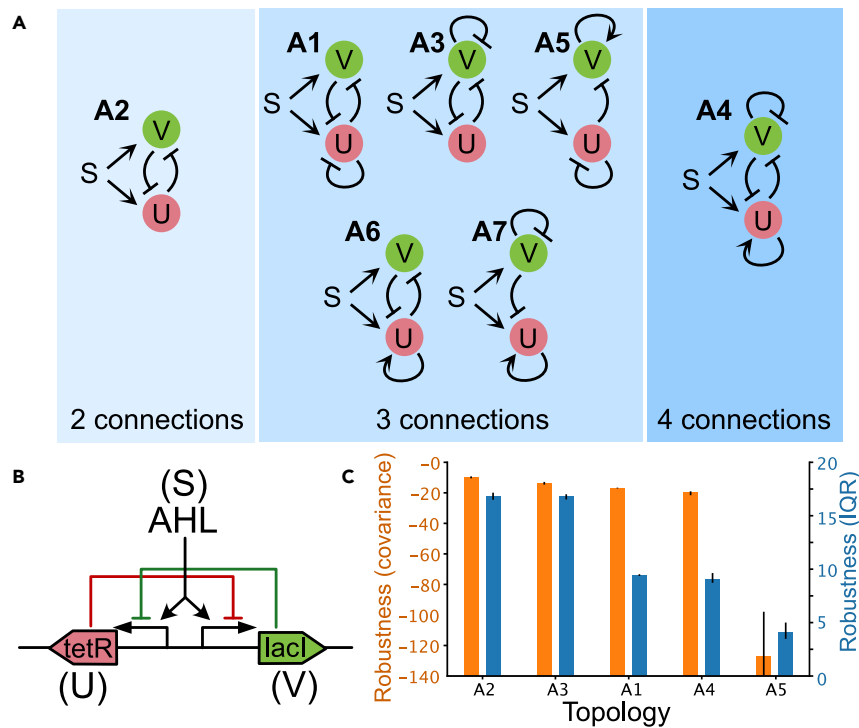
Figure 2. Gene regulatory circuits with 2 nodes with the capacity for mushroom bifurcation behavior

(A) All the 2-gene structures found by the algorithm leading to mushroom bifurcations, labeled according to their frequency within the set of solutions (A1 and A7 are the most and least frequent, respectively). The topologies are labeled according to their frequency within the set of solutions (*Mushroom2D*). Parameter bounds for the search are included in Table S1.

(B) One possible synthetic realization of the circuit with minimal topology (A2) leading to mushroom bifurcation.

(C) Top 5 topologies with highest robustness (see STAR Methods). Bars and errorbars correspond with the median and quartiles from bootstrapping.

screen targeting this topology confirmed that for the regulatory functions and parameter ranges used in our study, this topology was much less robust than the rest of the topologies found here (see Figure S4).

The key distinguishing feature for all topologies is the requirement of the signal activating both genes. A screening looking for the existence of mushroom bifurcation diagrams in the presence of a single activating input did not return any successful results, suggesting that the double activating role of the input is a requirement to reproduce the mushroom diagram.

With the same optimization strategy, we explore three-dimensional topologies leading to mushroom bifurcations by fixing the number of genes to three. The corresponding bounds for the parameters for the 3-gene network are included in Table S2. The algorithm detected through multiple optimization runs, more than 300 different 3 gene topologies leading to mushroom bifurcations from the potential 1728 topologies (the set of solutions we denote as *Mushroom3D*). Unlike exhaustive exploration strategies, our optimization-based approach finds structures fulfilling the target behavior very efficiently, in the order of seconds per run using a standard PC. The most frequent structures found are depicted in Table S3 and Figure S1. A selection of 3-gene structures which are not built up from 2-gene mushroom topologies are illustrated in Tables S4 and S5, and Figure S2.

Robust functionality vs. topological complexity

A circuit is considered functional if it shows a mushroom bifurcation. Nevertheless, it is important to in additionally assess its robustness by quantifying how the circuit functionality is kept with respect to perturbations in the parameters. With this aim, we defined different robustness scores based on the size of the parameter space occupied by successful parameter sets (see STAR Methods section). The most robust
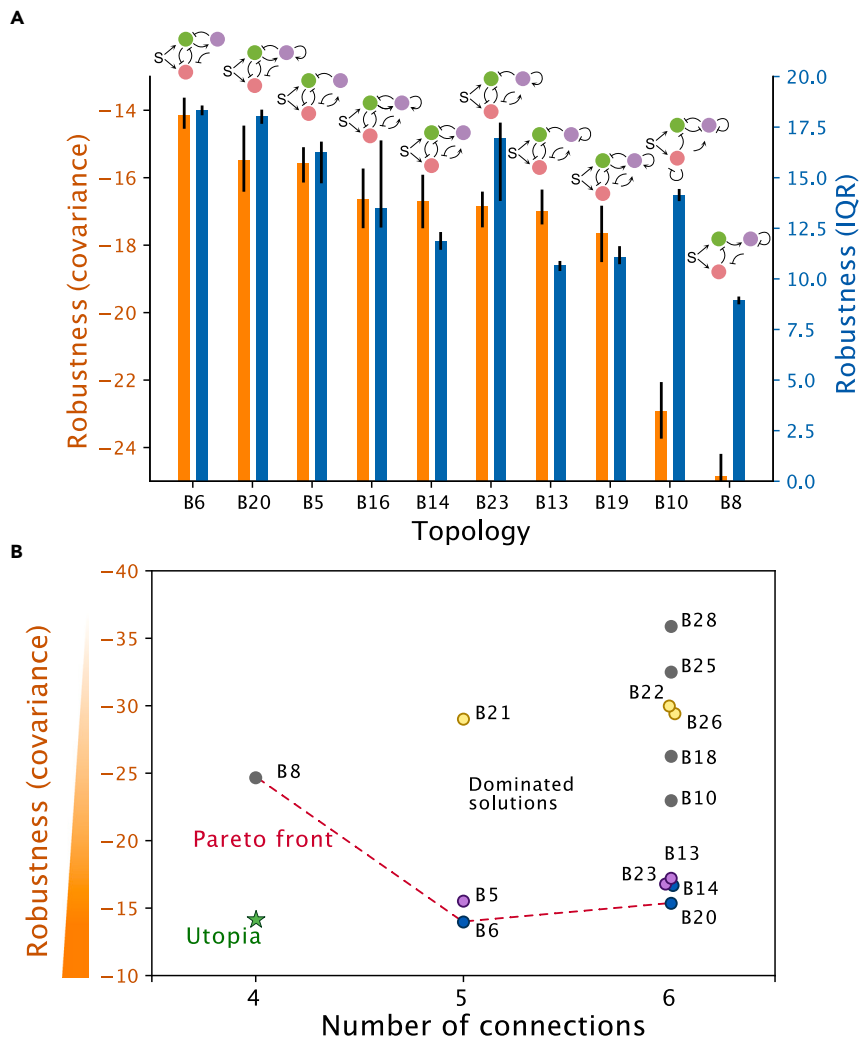
Figure 3. Robustness analysis for the 3-node topologies

(A) Top 10 topologies with highest robustness (see STAR Methods). Bars and errorbars correspond with the median and quartiles from bootstrapping.

(B) Relationship between number of connections and robustness for the most robust topologies. The Pareto front (red dashed line) is determined by the topologies with an optimal trade-off between robustness and complexity (in terms of the number of connections). Structures B8, B6, and B20 optimally trade-off robustness and complexity. Colored circles (blue, purple, and yellow) correspond with topologies that share the same 5-connection topologies (B6, B5, and B21) respectively.

structures are shown in Figures 2C and 3A. For the 2-node structures the most robust topologies show the expected cross-repression motif found in the bistable switch with marginal differences in robustness between topologies A1, A2, A3, and A4; highlighting the simplest topology A2 as a good potential candidate to construct a synthetic circuit in the lab (Figure 2B).

A similar analysis was performed for the successful topologies in 3-node networks. From the point of view of circuit implementation, we are interested in finding a good compromise between robustness and the number of connections. In order to select the best structures we build a Pareto front in the objective space (robustness vs. number of connections) as depicted in Figure 3B, identifying optimal topologies B8, B6, and B20 with 4, 5, and 6 connections, respectively. The structures in the Pareto front are shown in Figure 3B. Interestingly, the addition of connections to robust topologies did not show an improvement in the robustness of the network. In particular for the 3 most robust 5-connection topologies (B5, B6, and B21), the addition of connections just preserved or deteriorated the robustness of the network (see colored circles in Figure 3B).
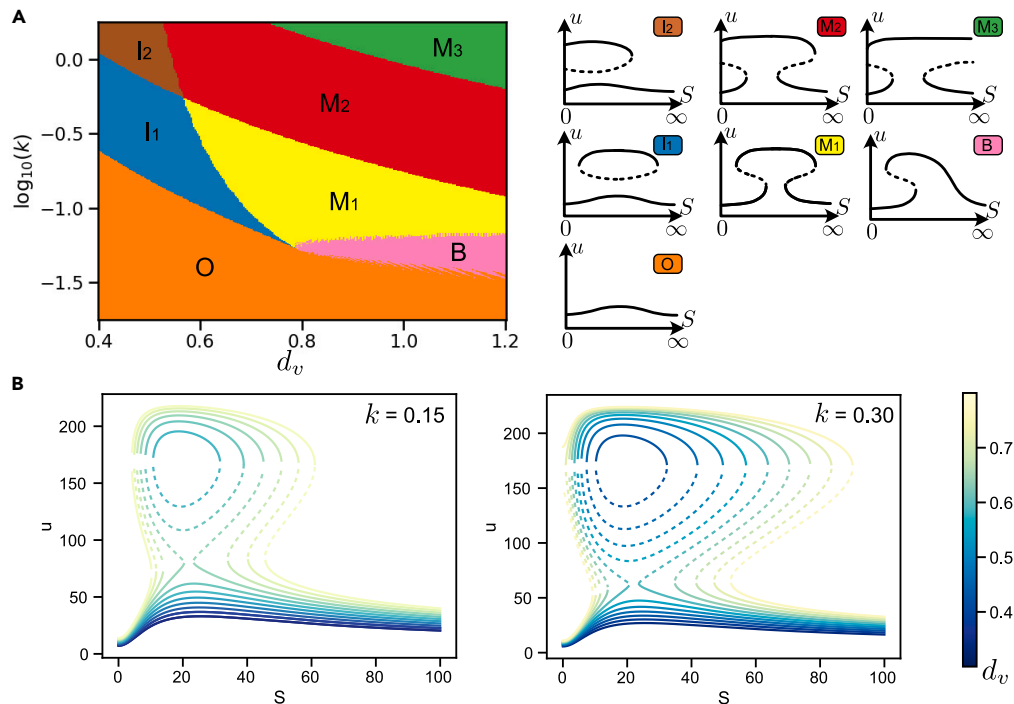
Figure 4. Bifurcation diagrams surrounding the mushroom behavior reveal a controllable rich functional landscape

(A) Phase diagram around a mushroom generated by varying the degradation rate d_v and the interaction strength k (see Figure 1A). Different colors show different bifurcation diagrams. A schematic of each bifurcation diagram is shown in the right panel indicating the different continuation combinations of stable (solid lines) and unstable states (dashed lines).

(B) Change in the mushroom bifurcation diagram by changing the degradation rate d_v for two different values of k. The different diagrams show the annihilation of the mushroom into an isola, followed by the collapse of the isola leading to monostability. Parameters used are: $p_0 = 230, p_1 = 50, p_2 = 1000, R_1 = 264, R_2 = 275, K_1 = 10, K_2 = 133, d_u = 1$ for topology A3.

Mushroom annihilation and isola formation

Following the analysis of the mushroom robustness, we explored the resulting bifurcation diagrams when the mushroom is lost. We observed three main transitions out of the mushroom bifurcation diagram (M_1 in Figure 4A): First, the mushroom head can cross values of saturation or absence or signal, resulting in incomplete mushrooms (M_2, M_3). Second, a pair of saddle nodes on one side of the mushroom collide, giving place to a cusp bifurcation giving rise to a bifurcation diagram similar to a bistable switch (B). Finally, the two saddle nodes forming the neck of the mushroom can collide, pinching the neck of the mushroom and producing an isola (I_1), a closed curve of equilibrium solutions delimited by two saddle nodes (Figure 4B). Strikingly, all these transitions are observed without the requirement to change the topology of the system, making this rich dynamical scenario exploration attainable in gene regulatory circuits.

While all the landscapes provide different dynamical properties, isolas are particularly compelling for their cell fate decision capabilities. Isolas are found in different contexts including chemical reactors, buckling of elastic shells and plasma physics.[27,28] Mechanisms of birth and annihilation of isolas have been classified via singularity theory,[28] being one of such mechanisms with particular interest in the transformation of a mushroom to an isola by reducing the number of saddle-node bifurcations.[12,29] It is interesting that despite being found both theoretically and experimentally in chemical systems, the isola bifurcation diagram has not yet been directly observed in a biological system. Understanding how to construct such a system in a gene network could therefore enable experimental verification of the underlying dynamical theory. In addition, as we will see below, the isola can form the basis of permanent switching systems, allowing new functionalities not achievable with a genetic toggle switch.

Building a genetic regulatory network with a functional isola requires us to be able to prescribe the range of signals of the isola. In our parameter exploration, we observed a controllable diversity on the level and
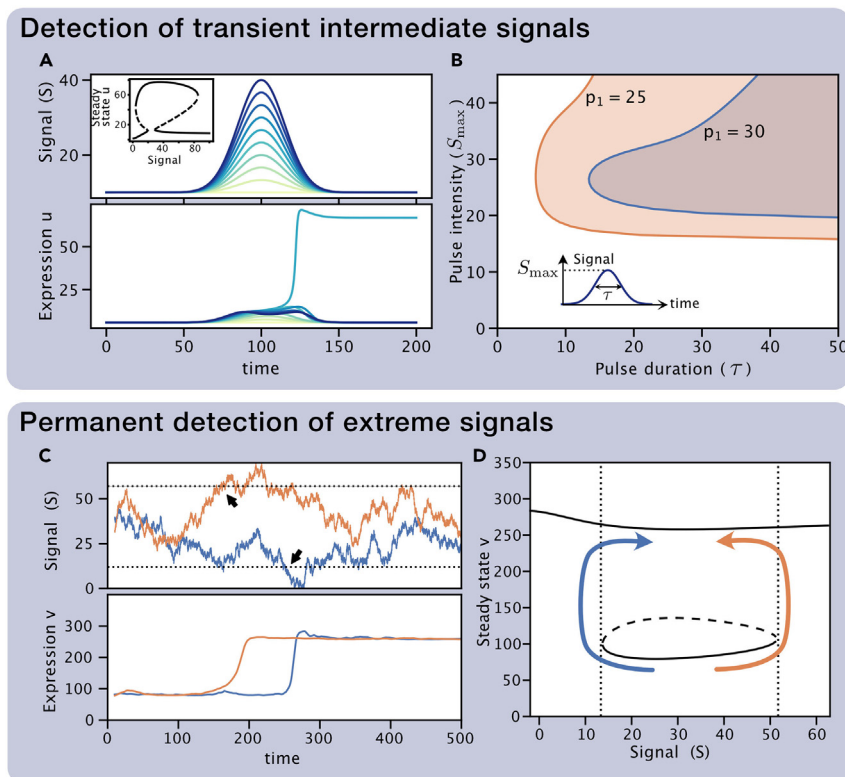
Figure 5. The mushroom circuit is able to discern intensities and durations of transient signals (top)

It can also be used to design a sensor of extreme signal values with memory using the isola regime (bottom).

(A) Response of the mushroom (inset) for different temporal signal pulses (top) of the same duration ($\tau = 20$) but different intensity (indicated by color). Expression of gene v (bottom) only becomes stably activated for pulses of intermediate maximum intensity (S_{max}).

(B) Response of the mushroom to signal pulses of different duration (τ) and maximum intensity (S_{max}) for two different values of the parameter b_2. Shaded regions indicate combinations of parameters for which node v is activated stably. Signal profiles follow the shape $S(t) = S_{min} + (S_{max} - S_{min})e^{-\frac{1}{2}((t-100)/\tau)^2}$ with $S_{min} = 10$. Results correspond to topology A1 with parameters $p_0 = 361, p_1 = p_3 = 30, p_2 = 411, k = 9.04 \cdot 10^{-2}, R_1 = 143, R_2 = 300, d_u = 1, d_v = 1.38, K_1 = 10.0, K_2 = 137$.

(C) Expression of the circuit (bottom) for two different realizations of a noisy signal (top).

(D) Isola bifurcation diagram showing the detection mechanism. When the signal reaches high or low levels determined by the isola boundary (dotted lines and black arrows in C and D), the system changes steady-state irreversibly. Schematic of the irreversible transitions is indicated by colored arrows. Results correspond to topology A1 with the same parameters than panel A but with $p_1 = 200$. Initial state was set by opening the mushroom setting parameter $p_1 = 20$. Noisy signal is a Wiener process.

range of signals delimiting the isola. In addition, we also detected different degrees of robustness as parameters varied (Figure 5B). To explore their prevalence, we explored the formation of isolas in all the identified mushrooms' 2D topologies and parameter sets by varying the degradation rate of the node V (d_v) and keeping the rest of the parameters constant. Strikingly, all the topologies tested exhibited the formation of isolas suggesting that the isola formation is a robust property of mushroom bifurcation. Alternatively, one can also employ the same automated method described in this paper to impose properties of the isola bifurcation diagram. In particular, these properties need to take into account, the location and curvature of the saddle-node points (see STAR Methods section for a more detailed description).

Design of biosensors and memory devices

The bifurcation diagrams identified in the manuscript are unexplored in synthetic biology and open the door to new functionalities. One of the main characteristics of the mushroom is that the "ON state" is only available for a reduced range of signals (the neck of the mushroom). This reveals the mushroom's capabilities as a biosensor, only allowing the activation (transition from OFF to ON) for precise levels of a target signal

located in the neck of the mushroom. Furthermore, since the neck of the mushroom is surrounded by two saddle-node bifurcations, the resulting dynamical ghosts will slow down the transition to the ON state.[23] Hence, activation is not immediate, filtering out transient signals, requiring the persistence of the target signal for a certain amount of time in order to reach the ON state (Figure 5A). In addition, the size of the head of the mushroom provides memory of the activation, preserving the ON state for a larger range of signals than the ones required for the activation (Figure 5A). As discussed in the previous section the mushroom bifurcation diagram is prone to pinching by controlling parameters such as the degradation rate of one of the genes. This provides a mechanism to regulate the size of the neck mushroom, opening the door to controlling the range of target signals as well as the required duration of such signals. This way the mushroom not only serves as an accurate signal detector but also as a timer (Figure 5B).

In addition to the mushroom bifurcation, the isola also provides compelling functionalities. In contrast to a genetic toggle switch, transitions out from the isola are irreversible. Hence, once in the isola state the system is able to detect a signal that goes outside of the range of the isola (either low or high), unable to reach back to the isola independently of future levels of the signal (Figures 5C and D). Thus the isola can serve as a sensor of extreme values with infinite memory, detecting if the signal has been high or low at any time point in the past. Since the isola bifurcation diagram is close to the mushroom diagram, reset of the system onto the isola state can be done by "opening" and "closing" the neck of the mushroom.

DISCUSSION

In this work, we have developed an approach to biosystem design based on the specification of the bifurcation structure (through the location of saddle nodes in the bifurcation diagram) and applied it to the case of the mushroom bifurcation of interest in developmental and synthetic biology. We found gene regulatory networks composed of two and three genes that displayed the desired behavior. In particular, we found that a system based on the genetic toggle switch incorporating a signal activating both genes was sufficient to reproduce a mushroom bifurcation diagram. We also explored the robustness of these circuits and built a Pareto front to capture the trade-off between the robustness and number of connections, identifying the best topologies to implement the circuit using synthetic biology tools. In addition, we explore specific dynamical functionalities of the mushroom that can be exploited to build biosensors with tuneable temporal and precision properties. We extended our exploration to the formation of different bifurcation diagrams, which occur in the parameter space close to the mushroom bifurcation. We showed that a variety of dynamical behaviors can be created that can provide interesting memory and signal-detecting capabilities unattainable with previous memory devices. Interestingly, this variety of behaviors was attained mainly by inducing differences in the degradation rates of the genes of the network, which is a common perturbation tool in synthetic biology,[30] and opens the door to the exploration of novel dynamical behavior when protein degradation is coupled to other synthetic circuits.[31,32] The most interesting bifurcation diagram found is the isola, which not only could be useful for error checking of deployed biosensors but could be useful in clinical applications, for example *in vivo* detection of inflammation[33] or metabolite levels.[34,35] Finally, the minimal mushroom topologies identified reveal a new way to create multistable systems, avoiding the need for loading a single bistable switch with additional autoregulation.[36–38]

Our results complement the small amount of other existing work on mushroom bifurcation. A recent study explored four incoherent feedforward networks with positive auto-regulation for their potential to display mushroom dynamics.[14] They found that all were capable of mushroom-like behavior, though the appearance depended on the incoherence of the networks and the strength of positive feedback. Even though their networks apparently depend on a unique input signal, their topologies depend on an upstream gene that imposes different regulation in the downstream genes, effectively recovering the double signal input we analyze in our study. By contrast, they included the possibility of repressing input signals, and a different regulatory formalism based on AND-type logic gate interaction in the activation and repression of genes. This reinforces how the results of such computational investigations depend somewhat on the modeling choice of the dynamics, a result demonstrated directly previously.[39] Further work in the experimental realization of these systems is required to understand the actual dynamics, and how these depend on the particular biological context. Another study modeled feedforward topologies with positive auto-regulation and demonstrated that only incoherent feedforward networks were capable of atypical bistabilities (mushrooms and isola's).[26] Although in this study we considered more complex topologies, we can see that when considering feedforward-like networks we also find that incoherence is required (e.g., A5 and A7 in Figure 2). Analysis of the topology studied in[26] using the thermodynamical formalism employed in our

study, revealed that the topology used by these authors can also show the mushroom bifurcation albeit in a smaller parameter region than topologies A1–A7. This manifests again the relevance that the choice of regulatory functions has on the resulting robustness of a topology.

Overall, designing biological systems based directly on bifurcation properties provides a natural tool to explore phenotype-genotype relationships, relating topologies of the network with the available dynamical behaviors. Such an approach allows for the specification of a set of functional requirements without precision of explicit integration of the differential equations. In addition, it provides a tool to design target dynamics that occur in a robust manner making these approaches key to the future of engineering of biological systems where uncertainty dominates. In future work, we will extend the method to encode different types of bifurcations (including the Hopf bifurcation) related to other complex nonlinear behaviors of interest in systems and synthetic biology.

Limitations of the study

Future research on this topic should take into account some limitations on the use of automated bifurcation design. Firstly, the choice of regulatory functions used in the study can significantly impact the results obtained, implying that it is necessary for case-by-case precise biochemical knowledge of the specific promoter and regulatory reactions used in experimental designs of the circuit. Additionally, the study does not account for stochastic effects such as intrinsic or extrinsic noise, which may influence the bifurcation behavior of the system. Finally, other cellular processes, such as competition for protein degradation, or metabolic overload; are not considered, requiring more tailored constraints in the optimization to real-world synthetic biology scenarios.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Gene network mixed integer modeling framework
  - Bifurcation conditions
  - Optimization-based automated design
  - Optimization strategy and solvers
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Analysis of robustness

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.106836.

AUTHOR CONTRIBUTIONS

Conceptualization, Investigation, Methodology, and Writing, I.O-M., R.P-C, J.R.B, and C.P.B.; Software, Formal Analysis, and Visualization I.O-M. and R.P-C.

REFERENCES

1. Schwille, P., Spatz, J., Landfester, K., Bodenschatz, E., Herminghaus, S., Sourjik, V., Erb, T.J., Bastiaens, P., Lipowsky, R., Hyman, A., et al. (2018). MaxSynBio: avenues towards creating cells from the bottom up. Angew. Chem. Int. Ed. Engl. *57*, 13382–13392.

2. Schaerli, Y., Munteanu, A., Gili, M., Cotterell, J., Sharpe, J., and Isalan, M. (2014). A unified design space of synthetic stripe-forming networks. Nat. Commun. *5*, 4905.

3. Pérez-Carrasco, R., Barnes, C.P., Schaerli, Y., Isalan, M., Briscoe, J., and Page, K.M. (2018). Combining a toggle switch and a repressilator within the AC-DC circuit generates distinct dynamical behaviors. Cell Syst. *6*, 521–530.e3.

4. Nielsen, A.A.K., Der, B.S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E.A., Ross, D., Densmore, D., Voigt, C.A., and Voigt, C.A. (2016). Genetic circuit design automation. Science *352*, aac7341.

5. Otero-Muras, I., Henriques, D., and Banga, J.R. (2016). SYNBADm: a tool for optimization-based automated design of synthetic gene circuits. Bioinformatics *32*, 3360–3362.

6. Otero-Muras, I., and Banga, J.R. (2017). Automated design framework for synthetic biology exploiting Pareto optimality. ACS Synth. Biol. *6*, 1180–1193.

7. Del Vecchio, D., Ninfa, A.J., and Sontag, E.D. (2008). Modular cell biology: retroactivity and insulation. Mol. Syst. Biol. *4*, 161. https://doi.org/10.1038/msb4100204.

8. Zhang, R., Goetz, H., Melendez-Alvarez, J., Li, J., Ding, T., Wang, X., and Tian, X.J. (2021). Winner-takes-all resource competition redirects cascading cell fate transitions. Nat. Commun. *12*, 853. https://doi.org/10.1038/s41467-021-21125-3.

9. Yordanov, P., Stelling, J., and Otero-Muras, I. (2020). BioSwitch: a tool for the detection of bistability and multi-steady state behaviour in signalling and gene regulatory networks. Bioinformatics *36*, 1640–1641.

10. Reyes, B.C., Otero-Muras, I., Shuen, M.T., Tartakovsky, A.M., and Petyuk, V.A. (2020). CRNT4SBML: a Python package for the detection of bistability in biochemical reaction networks. Bioinformatics *36*, 3922–3924.

11. Otero-Muras, I., and Banga, J.R. (2021). Synthetic gene circuit analysis and optimization. In Computational Methods in Synthetic Biology, Methods in Molecular Biology, *2189*, M.A. Marchisio, ed. (Humana).

12. Song, H., Smolen, P., Av-Ron, E., Baxter, D.A., and Byrne, J.H. (2006). Bifurcation and singularity analysis of a molecular network for the induction of long-term memory. Biophys. J. *90*, 2309–2325.

13. Sengupta, D., and Kar, S. (2018). Deciphering the dynamical origin of mixed population during neural stem cell development. Biophys. J. *114*, 992–1004.

14. Giri, A., and Kar, S. (2021). Incoherent modulation of bi-stable dynamics orchestrates the Mushroom and Isola bifurcations. Preprint at bioRxiv. https://doi.org/10.1101/2021.04.22.440901.

15. Basu, S., Gerchman, Y., Collins, C.H., Arnold, F.H., and Weiss, R. (2005). A synthetic multicellular system for programmed pattern formation. Nature *434*, 1130–1134.

16. Sohka, T., Heins, R.A., Phelan, R.M., Greisler, J.M., Townsend, C.A., and Ostermeier, M. (2009). An externally tunable bacterial band-pass filter. Proc. Natl. Acad. Sci. USA *106*, 10135–10140.

17. Kong, W., Blanchard, A.E., Liao, C., and Lu, T. (2017). Engineering robust and tunable spatial structures with synthetic gene circuits. Nucleic Acids Res. *45*, 1005–1014.

18. Gardner, T.S., Cantor, C.R., and Collins, J.J. (2000). Construction of a genetic toggle switch in Escherichia coli. Nature *403*, 339–342.

19. Litcofsky, K.D., Afeyan, R.B., Krom, R.J., Khalil, A.S., and Collins, J.J. (2012). Iterative plug-and-play methodology for constructing and modifying synthetic gene networks. Nat. Methods *9*, 1077–1080.

20. Zhu, R., del Rio-Salgado, J.M., Garcia-Ojalvo, J., and Elowitz, M.B. (2021). Synthetic multistability in mammalian cells. Preprint at bioRxiv. https://doi.org/10.1101/2021.02.10.430659.

21. Verd, B., Crombach, A., and Jaeger, J. (2014). Classification of transient behaviours in a time-dependent toggle switch model. BMC Syst. Biol. *8*, 1–19.

22. Dalchau, N., Szép, G., Hernansaiz-Ballesteros, R., Barnes, C.P., Cardelli, L., Phillips, A., and Csikász-Nagy, A. (2018). Computing with biological switches and clocks. Nat. Comput. *17*, 761–779.

23. Perez-Carrasco, R., Guerrero, P., Briscoe, J., and Page, K.M. (2016). Intrinsic noise profoundly alters the dynamics and steady state of morphogen-controlled bistable genetic switches. PLoS Comput. Biol. *12*, e1005154. https://doi.org/10.1371/journal.pcbi.1005154.

24. Barbier, I., Perez-Carrasco, R., and Schaerli, Y. (2020). Controlling spatiotemporal pattern formation in a concentration gradient with a synthetic toggle switch. Mol. Syst. Biol. *16*, e9361. https://doi.org/10.15252/msb.20199361.

25. Jiménez, A., Cotterell, J., Munteanu, A., and Sharpe, J. (2017). A spectrum of modularity in multi-functional gene circuits. Mol. Syst. Biol. *13*, 925.

26. Dey, A., and Barik, D. (2021). Emergent bistable switches from the incoherent feed-forward signaling of a positive feedback loop. ACS Synth. Biol. *10*, 3117–3128. https://doi.org/10.1021/acssynbio.1c00373.

27. Dellwo, D., Keller, H.B., Matkowsky, B.J., and Reiss, E.L. (1982). On the birth of isolas. SIAM J. Appl. Math. *42*, 956–963.

28. Avitabile, D., Desroches, M., and Rodrigues, S. (2012). On the numerical continuation of isolas of equilibria. Int. J. Bifurcation Chaos *22*, 1250277.

29. Ganapathisubramanian, N., and Showalter, K. (1984). On the numerical continuation of isolas of equilibria. J. Chem. Phys. *80*, 4177–4184.

30. Cameron, D.E., and Collins, J.J. (2014). Tunable protein degradation in bacteria. Nat. Biotechnol. *32*, 1276–1281. https://doi.org/10.1038/nbt.3053.

31. Cookson, N.A., Mather, W.H., Danino, T., Mondragón-Palomino, O., Williams, R.J., Tsimring, L.S., and Hasty, J. (2011). Queueing

up for enzymatic processing: correlated signaling through coupled degradation. Mol. Syst. Biol. *7*, 561. https://doi.org/10.1038/msb.2011.94.

32. Prindle, A., Selimkhanov, J., Li, H., Razinkov, I., Tsimring, L.S., and Hasty, J. (2014). Rapid and tunable post-translational coupling of genetic circuits. Nature *508*, 387–391. https://doi.org/10.1038/nature13238.

33. Riglar, D.T., Giessen, T.W., Baym, M., Kerns, S.J., Niederhuber, M.J., Bronson, R.T., Kotula, J.W., Gerber, G.K., Way, J.C., and Silver, P.A. (2017). Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. Nat. Biotechnol. *35*, 653–658.

34. Rutter, J.W., Ozdemir, T., Galimov, E.R., Quintaneiro, L.M., Rosa, L., Thomas, G.M., Cabreiro, F., and Barnes, C.P. (2019). Detecting changes in the Caenorhabditis elegans intestinal environment using an engineered bacterial biosensor. ACS Synth. Biol. *8*, 2620–2628.

35. Rutter, J.W., Dekker, L., Fedorec, A.J.H., Gonzales, D.T., Wen, K.Y., Tanner, L.E.S., Donovan, E., Ozdemir, T., Thomas, G.M., and Barnes, C.P. (2021). Engineered acetoacetate-inducible whole-cell biosensors based on the AtoSC two-component system. Biotechnol. Bioeng. *118*, 4278–4289.

36. Lou, C., Liu, X., Ni, M., Huang, Y., Huang, Q., Huang, L., Jiang, L., Lu, D., Wang, M., Liu, C., et al. (2010). Synthesizing a novel genetic sequential logic circuit: a push-on push-off switch. Mol. Syst. Biol. *6*, 350.

37. Wu, F., Su, R.Q., Lai, Y.C., and Wang, X. (2017). Engineering of a synthetic quadrastable gene network to approach Waddington landscape and cell fate determination. Elife *6*, e23702.

38. Leon, M., Woods, M.L., Fedorec, A.J.H., and Barnes, C.P. (2016). A computational method for the investigation of multistable system and its application to genetic switches. BMC Syst. Biol. *10*, 130.

39. Barnes, C.P., Silk, D., Sheng, X., and Stumpf, M.P.H. (2011). Bayesian design of synthetic biological systems. Proc. Natl. Acad. Sci. USA *108*, 15190–15195.

40. Otero-Muras, I., Banga, J.R., and Alonso, A.A. (2012). Characterizing multistationarity regimes in biochemical reaction networks. PLoS One *7*, e39194.

41. Otero-Muras, I., Yordanov, P., and Stelling, J. (2017). Chemical Reaction Network Theory elucidates sources of multistability in interferon signaling. PLoS Comput. Biol. *13*, e1005454.

42. Otero-Muras, I., and Banga, J.R. (2018). Optimization-based prediciton of fold bifurcations in nonlinear ODE models. In 18th IFAC Symposium on System Identification SYSID 2018.

43. Sendín, J.O.H., Exler, O., and Banga, J.R. (2010). Multi-objective mixed integer strategy for the optimisation of biological networks. IET Syst. Biol. *4*, 236–248.

44. Miettinen, K. (2012). Nonlinear Multiobjective Optimization, *12* (Springer Science & Business Media).

45. Otero-Muras, I., and Banga, J. (2019). Distilling robust design principles of biocircuits using mixed integer dynamic optimization. Processes *7*, 92.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Code used for analyses	This paper	https://doi.org/10.5281/zenodo.6024249
Software and algorithms		
Matlab	MathWorks	https://www.mathworks.com
Python 3	Python	https://www.python.org
Affinity Designer 2	Serif	https://affinity.serif.com

RESOURCE AVAILABILITY

Lead contact

Further information and requests related to code should be directed to and will be fulfilled by the lead contact, Ruben Perez-Carrasco (r.perez-carrasco@imperial.ac.uk).

Materials availability

This study is computational and did not generate new reagents.

Data and code availability

- Data files including the resulting topologies and parameters from the optimization are available online at Zenodo: https://doi.org/10.5281/zenodo.6024249.

- The code used in the analysis can be found in the same repository at Zenodo: https://doi.org/10.5281/zenodo.6024249.

- Any additional information required to reanalyse the data reported in this work is available from the lead contact upon reasonable request.

METHOD DETAILS

Gene network mixed integer modeling framework

The first step of the method proposed consists of encoding the dynamics of the gene regulatory superstructure in a mixed integer modeling framework. For models endowed with typical kinetics of the Hill or Shea-Ackers types, the state vector $\omega \in \mathbb{R}_{\geq 0}^{N}$ contains the concentrations of the protein species involved, and the system's dimension N is determined by the number of genes in the network. See[6] for mixed integer modeling encodings with different kinetics. The dynamics are expressed in a mixed integer framework in terms of the vectors ω, x, y in the following form as a set of Ordinary Differential Equations:

$$\frac{d\omega}{dt} = \xi(\omega, x, y)$$

(Equation 1)

where x, and y are the vectors containing respectively the real and integer design variables. In general terms, real variables might include kinetic parameters, degradation rates, promoter strengths etc, and integer variables determine the structure and connectivity of the network.

In this study, we use the superstructure in network in Figure 1B, encoded in a mixed integer framework as follows: the gene circuit topology is characterized by a vector y of 9 integer variables (y_{uu}, y_{vu}, y_{wu}, y_{uv}, y_{vv}, y_{wv}, y_{uw}, y_{vw}, y_{ww}) such that $y_{ij} = -1$ if j is repressed by i, $y_{ji} = 1$ if i is activated by j, and $y_{ij} = 0$ otherwise with $i,j = \{u,v,w\}$. Within this framework, a gene circuit is characterized by the vector y and a vector x containing 12 real variables coding for tunable biochemical parameters (including promoter strengths, leakiness, degradation rate constants, repression and activation affinities). We consider that the rate constants governing the strengths of the active interactions k_{ij} in Figure 1B are equal (k). As an illustrative example,

the three gene system in Figure 1A (for which $y_{uu} = y_{wu} = y_{vv} = y_{uw} = y_{ww} = 0$ and $y_{vu} = y_{uv} = y_{wv} = y_{vw} = -1$) is given by:

$$\frac{du}{dt} = p_0 \frac{p_1+q_1(S)}{1+p_1+q_1(S)+kv^2} - d_u u$$

$$\frac{dv}{dt} = p_2 \frac{p_3+q_2(S)}{1+p_3+q_2(S)+ku^2+kw^2} - d_v v$$

$$\frac{dw}{dt} = p_4 \frac{p_5}{1+p_5+kv^2} - d_w w,$$

where the state vector is denoted by $\omega = (u, v, w)$ being u, v, w the concentrations of the proteins expressed by genes U, V, W respectively; p_0, p_2, p_4 are the promoter strengths, p_1, p_3, p_5 the promoter leakiness, k are repression strengths and d_u, d_v, d_w the degradation rate constants. The functions $q_1(S), q_2(S)$ represent the concentration of activating transcription factors regulated by an input biochemical signal S, given by Hill functions:
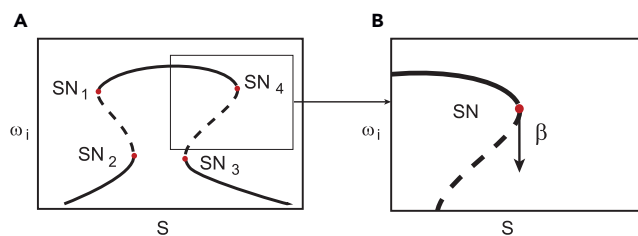
$$q_i(S) = R_i \frac{S^2}{K_i^2+S^2},$$

where R is the total concentration of transcription factor, S is the concentration of signal inducer, K is the dissociation constant, and the cooperativity is 2. We assume that repressor proteins bind as dimers (or equivalently there are two operator sites). Depending on the specific design scenario, a number of assumptions can be made to reduce the dimensionality of the parameter search space, here we consider $p_3 = p_1$ and $d_u = 1$. In this way, the vector of real design variables reads $x = (p_0, p_1, p_2, p_4, p_5, k, R_1, R_2, d_v, d_w, K_1, K_2)$, and the dynamics is expressed in terms of the vectors ω, x, y in the form given by Equation 1.

Importantly, previous to the formulation of the design as an optimization problem, we have to select an appropriate bifurcation parameter. We use the concentration of the inducer input S as bifurcation parameter and, for convenience, we express the dynamics in the compact form:

$$\frac{d\omega}{dt} = f(\omega, x, y, S). \qquad \text{(Equation 2)}$$

Bifurcation conditions

Once the dynamics are encoded in the mixed integer form (2), we formulate the conditions for the target bifurcation in terms of the design variables x, y and the bifurcation parameter S. Here, the bifurcation of interest is the saddle-node bifurcation (also denoted as fold or limit point bifurcation). In particular, there are 4 saddle-node bifurcations in the mushroom-shaped diagram (as indicated in figure below).

Bifurcation diagram structure used in the optimisation
(A and B) (A) Mushroom-shaped bifurcation diagram and (B) Saddle-Node bifurcation and tangent vector β at the bifurcation point.

A parametric condition for saddle-nodes in biochemical networks with mass action kinetics was demonstrated by.[40,41] This analytic condition is exploited in effective computational tools to detect, via optimization, saddle-node bifurcations and bistability in biochemical reaction networks with mass action kinetics.[9–11] For non-mass-action kinetics, as it is the case in this study, we use an alternative formulation of optimization problem (see[42]), in which the objective function to minimize is defined in terms of the extended Jacobian of the system, as follows.

Starting from Equation 2 we compute the following extended Jacobian:

$$Q = [D_\omega f(\omega, x, y, S), D_S f(\omega, x, y, S)] \quad \text{(Equation 3)}$$

with dimensions $N \times N+1$. This extended Jacobian can be obtained symbolically or approximated numerically by finite differences. The tangent vector at the bifurcation point can be computed as $\beta = null(Q)$. By definition, a saddle-node point occurs when $\beta_{N+1} = 0$, i.e., if the $N+1$ entry of the tangent vector β is equal to zero (see figure in bifurcation conditions). Therefore, the function:

$$J = \beta_{N+1}^2(\dot{\omega}, \omega, y, S) \quad \text{(Equation 4)}$$

reaches its minimum value (i.e. J = 0) at a saddle node point.

Optimization-based automated design

Mushroom-shaped bifurcation diagrams show four saddle-node bifurcation points, see figure in bifurcation conditions. in which SN1, SN2, SN3 and SN4 indicate saddle node bifurcations with coordinates (S_i, ω_i) for $i = 1, \ldots, 4$ in the bifurcation diagram. Following the reasoning of the previous subsection, the function:

$$\Psi = \sum_{k=1}^{4} J_k \text{ with } J_k = \beta_{kN+1}^2(\dot{\omega}_k, \omega_k, x, y, S_k)$$

reaches its minimum ($\Psi = 0$) when four saddle-node bifurcations occur at coordinates (S_k, ω_k) for $k = 1, \ldots, 4$ in the bifurcation diagram. Therefore, we can formulate the design of bio-circuits with mushroom bifurcation behavior (with saddle node bifurcations located within the desired regions in the bifurcation diagram) as the following optimization problem:

$$\min_{x, y, S} \Psi(\dot{\omega}, \omega, x, y, S)$$

subject to:

$$f(\omega, x, y, S) = 0 \quad \text{(Equation 5)}$$

$$\omega_L \leq \omega \leq \omega_U \quad \text{(Equation 6)}$$

$$\begin{aligned} x_L &\leq x \leq x_U \\ y_L &\leq y \leq y_U \\ S_L &\leq S \leq S_U \end{aligned} \quad \text{(Equation 7)}$$

Note that S here denotes a four dimensional vector, being S_L and S_U the lower and upper bounds of the target rectangular regions. The lower and upper bounds of the real and integer decision variables are encoded in vectors x_L, y_L and x_U, y_U, respectively. We can specify the (rectangular) regions in the bifurcation diagram where we want the bifurcation points to be located through the values of the lower and upper bounds of the inequality constraints (6) and the decision variable S, see for example the rectangular regions in Figure 1C. The choice of constraints can be defined according with the design specifications required by the user, related to the desired specific functionality (for example, if the desired mushroom might be inverted, i.e. upside down), the particular experimental constraints, etc. In this work, the goal was to find circuits leading to mushrooms (including the possibility of inverted ones), so we defined 2 instead of 4 rectangular regions as constraints, which led to a very efficient search for biocircuits with the target behavior). The constraints used are in Tables S1 and S2. Note that the constraints for u and v are different, resulting in a break of the symmetry of the model.

The optimization problem is a Mixed Integer Nonlinear Programming problem (MINLP). Due to its non-convex nature, global optimization solvers are needed to obtain adequate solutions.[43] It is important to remark that the solution of this optimization problem is not unique, since there are multiple combinations of topology and parameters leading to mushroom bifurcations. A single run of the algorithm will find a particular solution very efficiently (in the order of a few seconds for a standard PC), whereas a multistart strategy (running the optimization algorithm multiple times) will effectively find different topologies with mushroom bifurcation diagram.

For the screening of 2-gene and 3-gene topologies with mushroom behavior, we use a multistart strategy with 2000 and 10000 runs, respectively. As indicated in the results section, the corresponding bounds for the parameters for the 2-gene and 3-gene networks are included in the Tables S1–S3, and the solution sets denoted by *Mushroom2D* and *Mushroom3D*, respectively.

In advanced design applications, it is usually the case that multiple objectives need to be taken into account.[6] For example, during the design of next generation biosensors, we might be interested not only in a mushroom-shaped diagram, but also in maximizing the regions of bistability. In numerous occasions, it is also required to operate at low protein burden to avoid compromising the viability of the cells, being the minimal protein production cost another important objective to be considered. In order to find the best designs with respect to multiple criteria $(\psi_1, \ldots, \psi_M)$, we consider a multi-objective formulation of the optimization problem., i.e. we set the mushroom bifurcation condition $\Psi = 0$ as a constraint of the following Multiobjective Mixed Integer Nonlinear Programming (MO-MINLP) problem:

$$\min_{x,y,S} \psi_1(\dot{\omega}, \omega, x, y, S), \ldots, \psi_M(\dot{\omega}, \omega, x, y, S)$$

subject to:

$$f(\omega, x, y, S) = 0$$

$$\Psi(\dot{\omega}, \omega, x, y, S) = 0$$

$$\begin{aligned} x_L \leq x \leq x_U \\ y_L \leq y \leq y_U \\ S_L \leq S \leq S_U \end{aligned}$$

(Equation 8)

The solution of the above multicriteria mixed integer nonlinear problem is not unique, but a set of vectors representing the best compromises between the (usually conflicting) objective functions. This set of best trade-offs is usually known as the Pareto set.

The isola-behavior can be also designed automatically, taking into account that at the isola both bifurcation curves of steady-state solutions $u^*(S)$ exhibit two saddle-node bifurcations at the coordinates (S_1, u_1^*) and (S_2, u_2^*) with $S_1 < S_2$. The isola will fulfill the property $u_{SS}^{-1}(S_1) > 0$ and $u_{SS}^{-1}(S_2) < 0$, where u_{SS}^{-1} is the local curvature of the inverse of the function $u^*(S)$ that can be evaluated numerically. In contrast, for a bistable switch we would expect $u_{SS}^{-1}(S_1) < 0$ and $u_{SS}^{-1}(S_2) > 0$. These conditions are easily incorporated into the optimization problem in order to target the isola bifurcation.

Optimization strategy and solvers

Computing the Pareto set of optimal designs by solving the above problems efficiently and reliably can be a daunting task due to their non-convexity, arising from their highly constrained, partially discrete and nonlinear nature.

Here we transform the multicriteria mixed integer formulation (MO-MINLP) into a finite set of single-objective mixed integer (MINLP) problems by adopting an ε-constraint approach.[44] The resulting set of MINLPs is then solved using a hybrid strategy, eSS-MISQP, which combines a diversification phase (using a global optimization metaheuristic, eSS) with intensification steps (using an efficient local mixed integer optimization solver, MISQP). This hybrid strategy has been found to outperform other global approaches in terms of both efficiency and robustness.[6]

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of robustness

In order to assess the robustness of a topology, we used two different scores based on the covariance and the interquantile range of the successful parameter sets for that topology (see Figure 3). On the one hand we defined a robustness scores based on the hypervolume of the parameter space occupied by solutions that return a successful mushroom bifurcation. This is achieved by calculating the logarithm of the determinant of the covariance matrix of the parameter sets. On the other hand we used a score successfully employed for gene regulatory structures[45] calculated as the sum of the interquantile ranges (IQR) measuring the spread of the distributions of each of the parameters for the solution set. Both scores were regularized by using the standardized parameters of the aggregated parameter set for all the topologies. The robustness histograms for the extremes of the Pareto front in Figure 3 are included in the Figure S4.