Supplementary Information

Discriminating direct and indirect connectivities in biological networks

Taek Kang[1,2], Richard Moore[1,2], Yi Li[1,2], Eduardo Sontag[4], and Leonidas Bleris[*1,2,3]

[1]Bioengineering Department, University of Texas at Dallas, 800 West Campbell Road, Richardson TX 75080 USA

[2]Center for Systems Biology, University of Texas at Dallas, NSERL 4.708, 800 West Campbell Road, Richardson TX 75080 USA

[3]Electrical Engineering Department, University of Texas at Dallas, 800 West Campbell Road, Richardson TX 75080 USA

[4]Department of Mathematics and Center for Quantitative Biology, Rutgers-The State University of New Jersey, Hill Center, 110 Frelinghuysen Rd., Piscataway NJ 08854 USA

*Corresponding author: L.B. (bleris@utdallas.edu)

Keywords: Reverse Engineering, Benchmark Synthetic Circuits, Human Cells, Modular Response Analysis, Direct and Indirect Connectivities, Nonparametric Resampling, Bootstrap.

Supplementary Figures

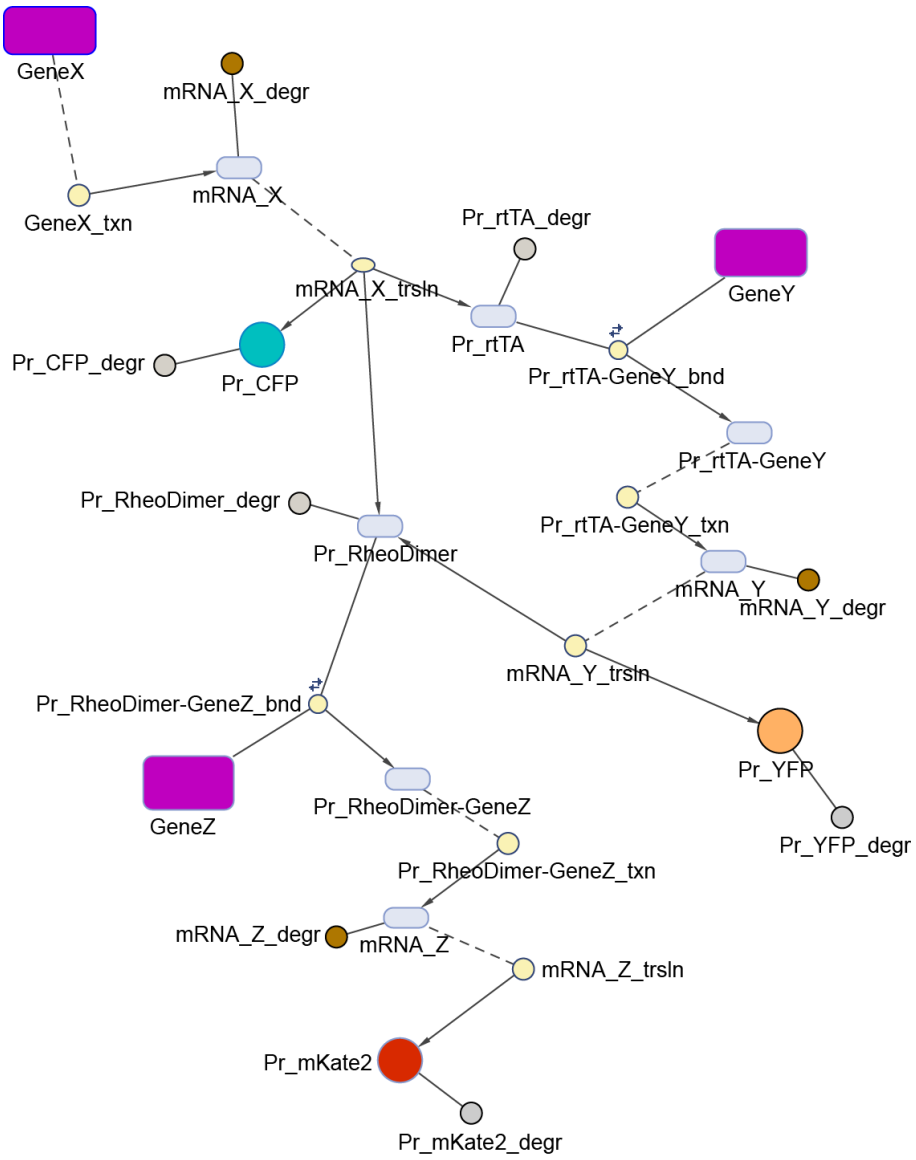

Figure S1. SimBiology model diagram of the synthetic feedforward architecture. To simulate the network behavior and their response to perturbations, the synthetic networks were recreated as mathematical models using MATLAB SimBiology. Each model is composed of distinct nodes that represent DNA, RNA and protein species and reactions that connect these nodes. The published kinetic parameters are listed in Table S4.

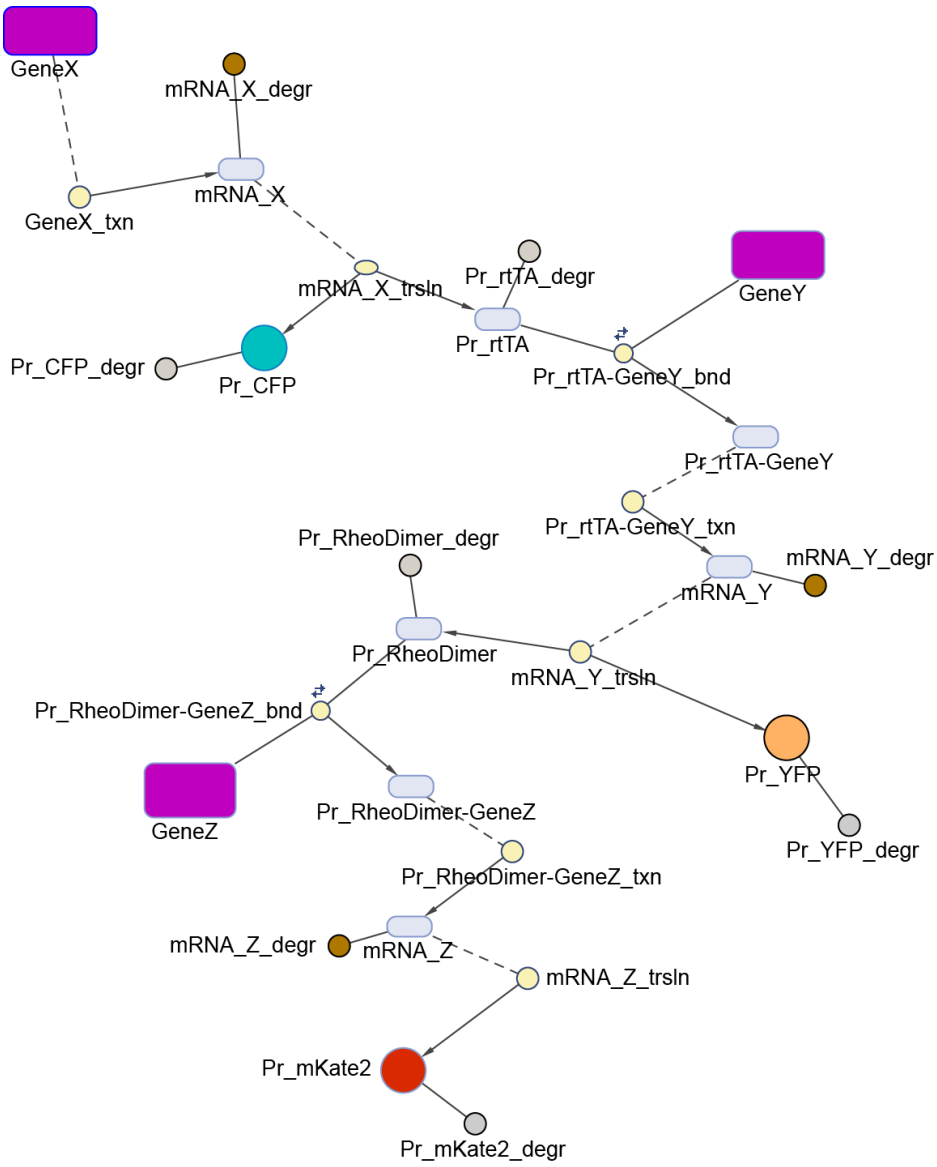

Figure S2. SimBiology model diagram of the synthetic cascade network. Mathematical model of synthetic cascade network was created using MATLAB SimBiology. See Supplementary Figure 1 for details.

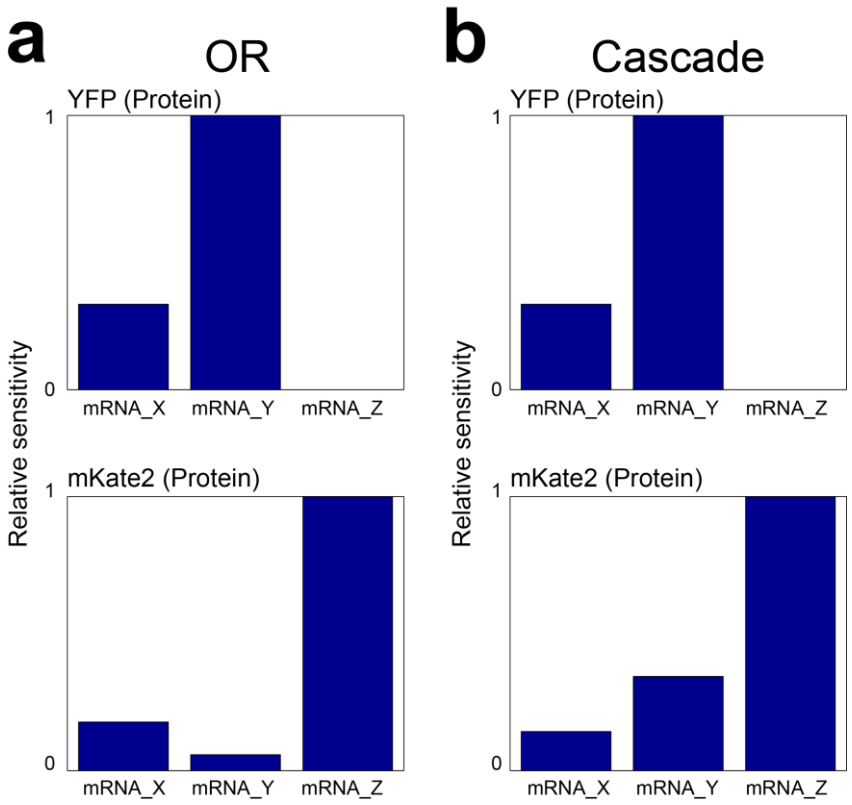

Figure S3. In silico sensitivity analysis of the model benchmark networks. Using the mathematical model of the synthetic networks as presented in Supplementary Figure 1, sensitivity analysis of YFP and mKate2 protein against mRNA species of each node was performed.

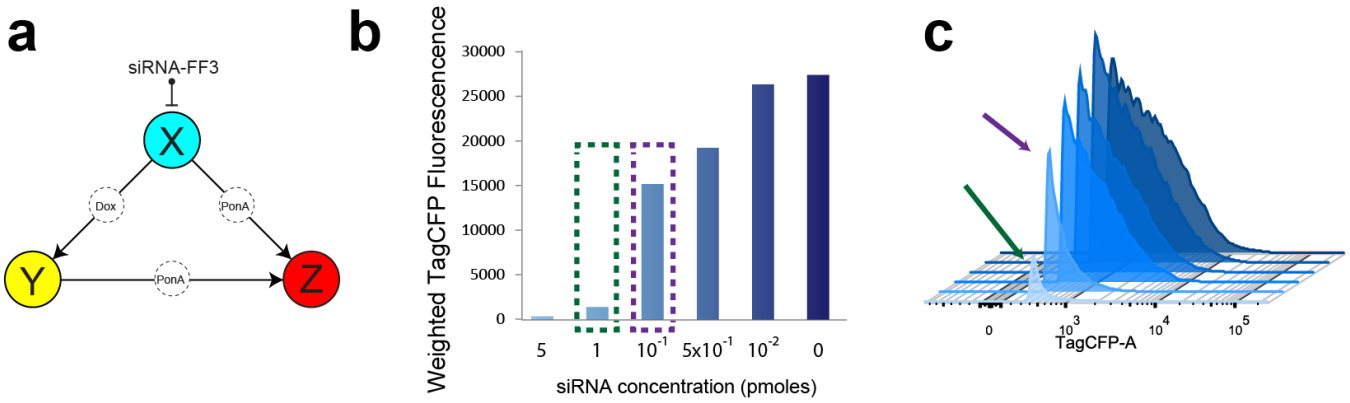

Figure S4. Node X response after perturbation with siRNA that target node X. Each node of the synthetic feedforward loop is perturbed via RNA interference. Ubiquitous suppression of nodes X and Y are achieved by incorporating synthetic siRNA target sites in the 3' UTR of the transcripts. For node Z, a synthetic siRNA that targets the mKate2 mRNA was used. To measure the efficacy of each siRNA, titrations were performed. (a) The activity of node X consists of fluorescent reporter TagCFP, the transactivator proteins rtTA and components of the RheoDimer (RheoActivator and RheoReceptor), and its activity is suppressed via siRNA FF3. (b) Weighted fluorescence of TagCFP obtained from fluorescence-positive cells after siRNA titration. Measurement surrounded by the purple box refers to the siRNA concentration that was included in a "weak" perturbation set, and the green box indicates siRNA concentration belonging to a "strong" perturbation set. (c) Histogram of the fluorescence-positive population, with the arrows indicating the corresponding population from (b).

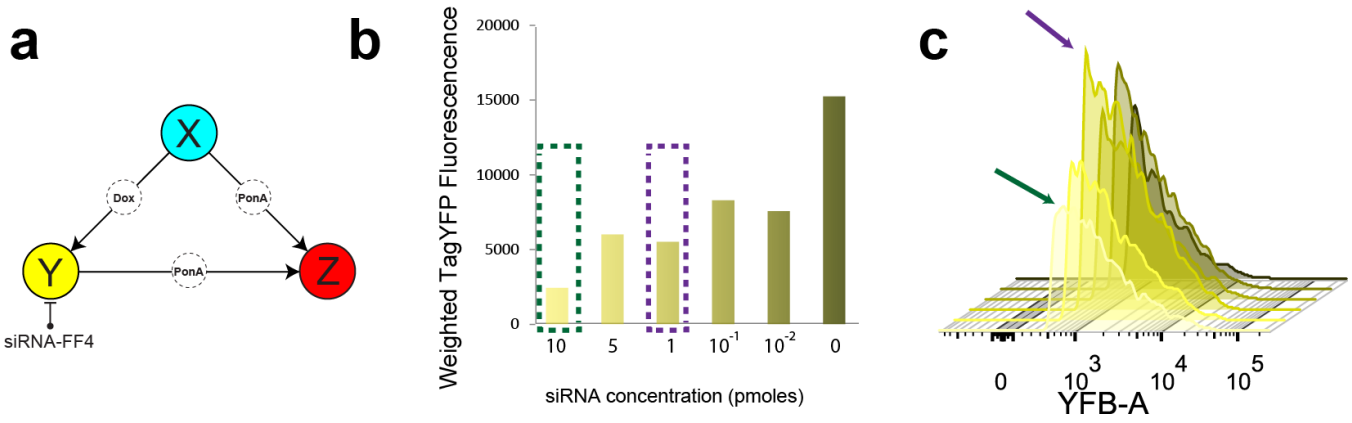

Figure S5. Synthetic network response following perturbation with siRNA. Node Y consists of the fluorescent reporter TagYFP and components of RheoDimer transactivator, and its activity is suppressed by FF4 siRNA. See Figure S4 for description of the experiment.

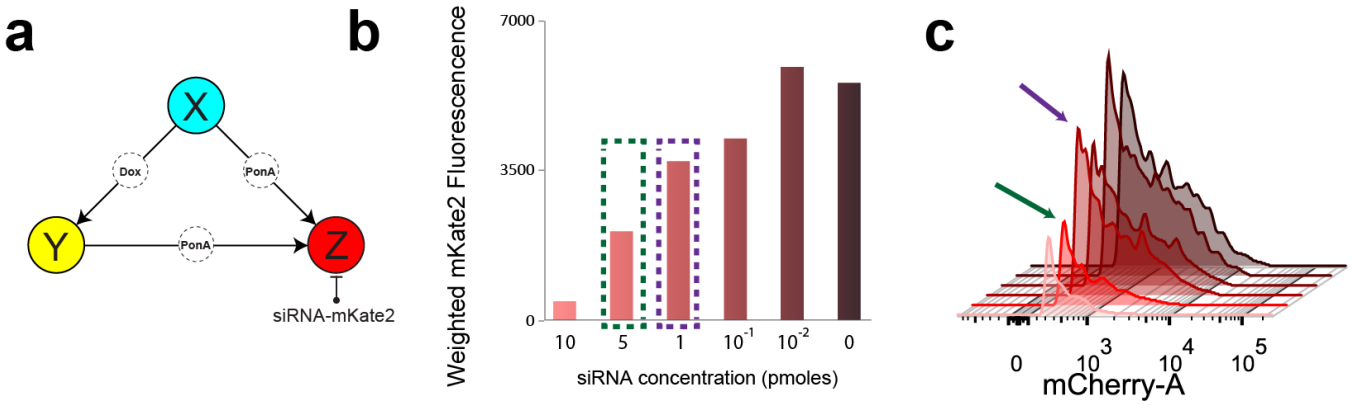

Figure S6. Synthetic network response following perturbation with siRNA. Node Z consists of fluorescent reporter mKate2, and its activity is suppressed by a custom siRNA that targets the mKate2 mRNA. See Figure S4 for description of the experiment.

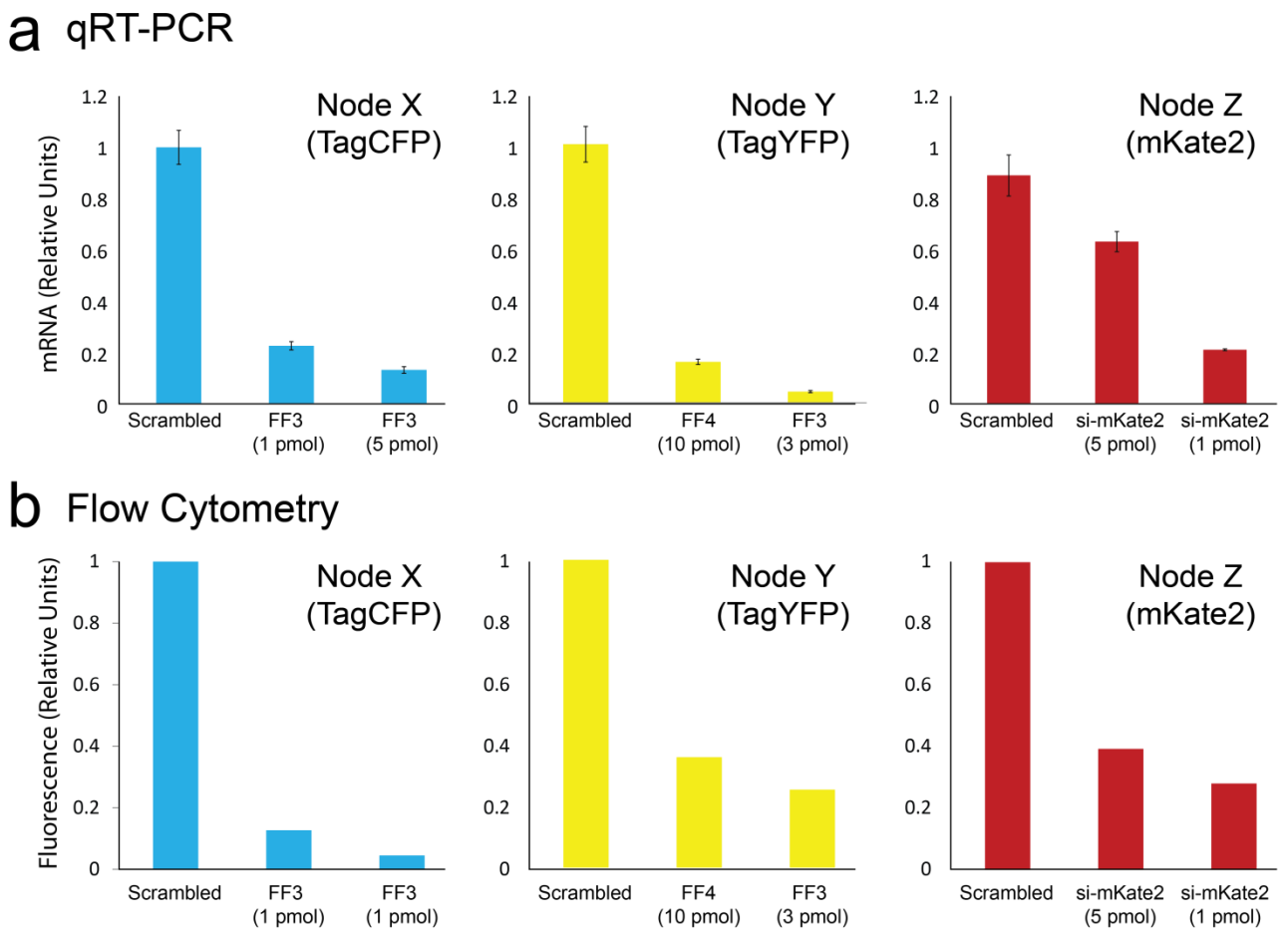

Figure S7. qRT-PCR of the synthetic network following systemic perturbation. (a) To directly test for the effectiveness of the siRNA on its target, qRT-PCR was performed 48 hours after co-transfection of the circuit cassette and the siRNA that target each node. (b) Flow cytometry was performed to confirm the relative protein expression after each perturbation condition. The protocol for qRT-PCR is described in SI Appendix, Supplementary Methods.

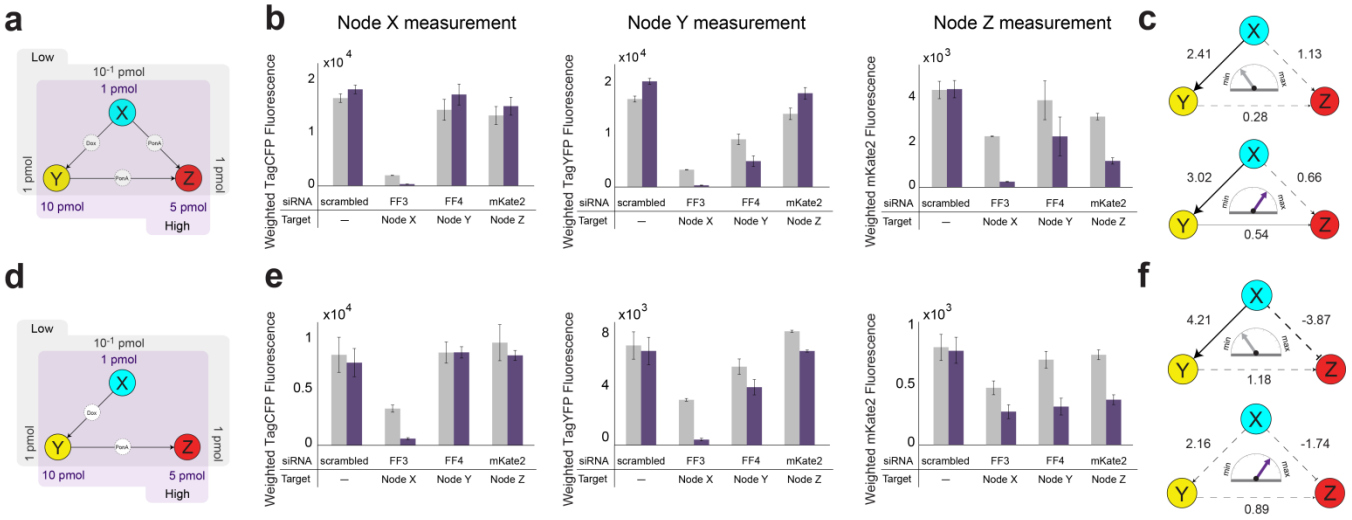

Figure. S8 Network reverse engineering using population data. The nodes X, Y, and Z of the benchmark synthetic networks are quantified by fluorescent reporters TagCFP, TagYFP and mKate2, respectively. (a) Graphical representation of the perturbation properties. Systematic perturbation to the network is performed twice, each with a different set of siRNA concentrations. Each siRNA concentrations are categorized as either "weak" or "strong" to indicate the relative strength of suppression. The set with lower overall concentration of siRNAs ("low") is in grey and the set with higher concentration ("high") is in purple. The same color scheme is used to distinguish these two experiments in subsequent analysis. (b) Perturbation responses of the gated population to perturbations to nodes X, Y and Z. To evaluate the fluorescence profile of the architecture post-perturbation, flow cytometry was performed 48 hours after transfection of the circuit. Weighted fluorescence corresponds to the percentage of the fluorescence-positive gated cells multiplied the mean fluorescence of the population. Error bars indicate standard deviation among triplicate experiments. (c) Graphical representation of the recovered circuit topology using population data. Network inference using MRA was performed with the weighted fluorescence profile shown in (b). Uncertainty of each local response coefficient was calculated by Monte Carlo simulations. Dotted lines are used to indicate that the 95% confidence interval of the corresponding distribution includes zero, which renders the connection as statistically insignificant. The corresponding results from the same experiment using the cascade architecture are shown in panels (d – f).

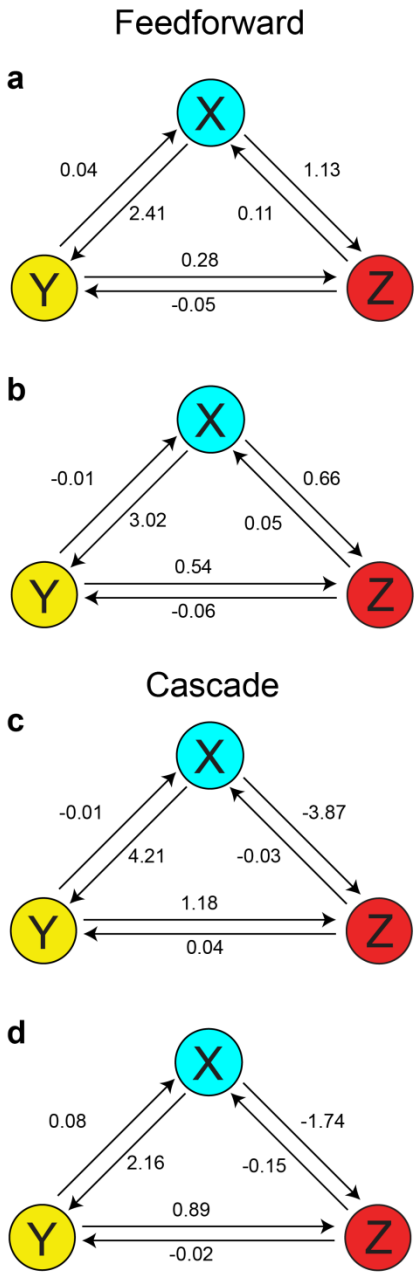

Figure S9. Graphical representation of the complete circuit topology derived from population-level statistics. Interaction coefficients for all regulatory links were calculated using modular response analysis, using mean fluorescent values of each population. Circuit topology of the feedforward architecture following weak and strong perturbation are shown in (a,b), respectively. Corresponding results from the cascade architecture are shown in (c,d).

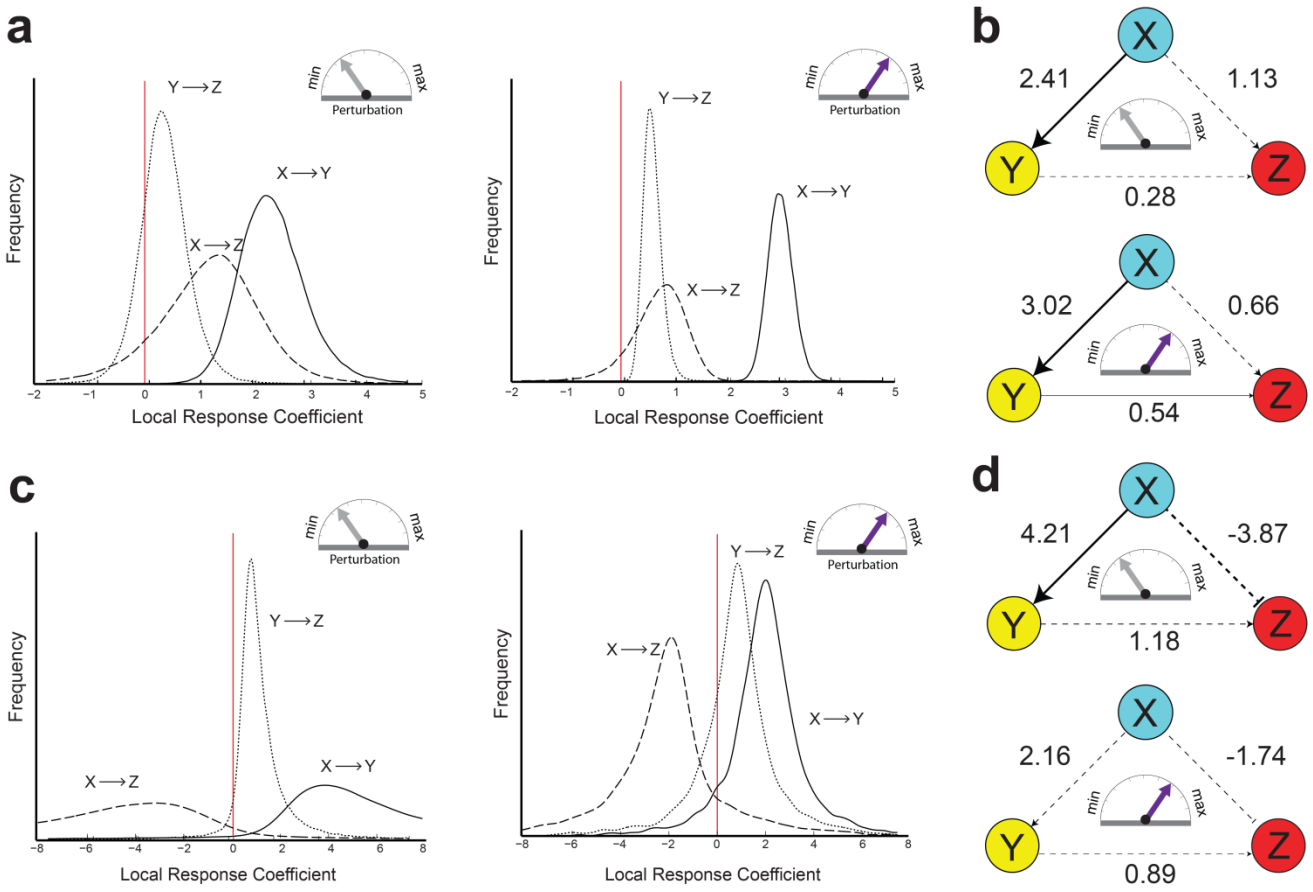

Figure S10. Monte Carlo error propagation analysis of modular response analysis. (a) To confirm the significance of the calculated local response coefficients, error propagation analysis using Monte Carlo simulation was performed. Each peak represents the probability distribution of calculated local response coefficients (LRCs) of the connections that make up the feedforward loop topology after weak perturbation (left) and strong perturbation (right). (b) Graphical representation of the recovered circuit topology, with mean LRCs from (a). (c) Monte Carlo simulation results after perturbation of the cascade networks and the corresponding recovered circuit topology (d).

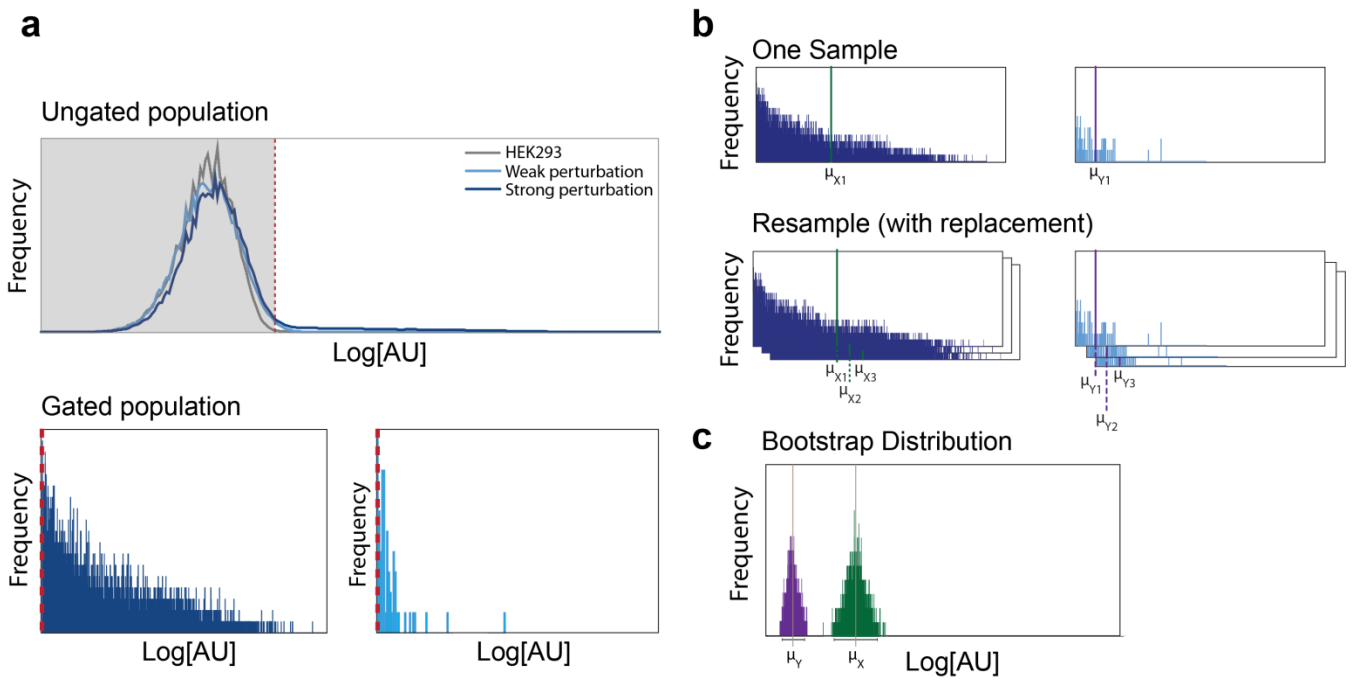

Figure S11. Bootstrap resampling workflow. To calculate the error of the obtained population statistic of the flow cytometry data, bootstrap distribution is used. (a) After gating for fluorescence-positive cells, we randomly sample (with replacement) to create a resampled dataset. (b) The size of a bootstrap resampled dataset is identical to the original population from which we sample from. We then calculate the desired population statistic (mean). (c) The repetition of this process produces a distribution and the confidence interval of means, which we can use as an estimator for the error of the calculated mean.

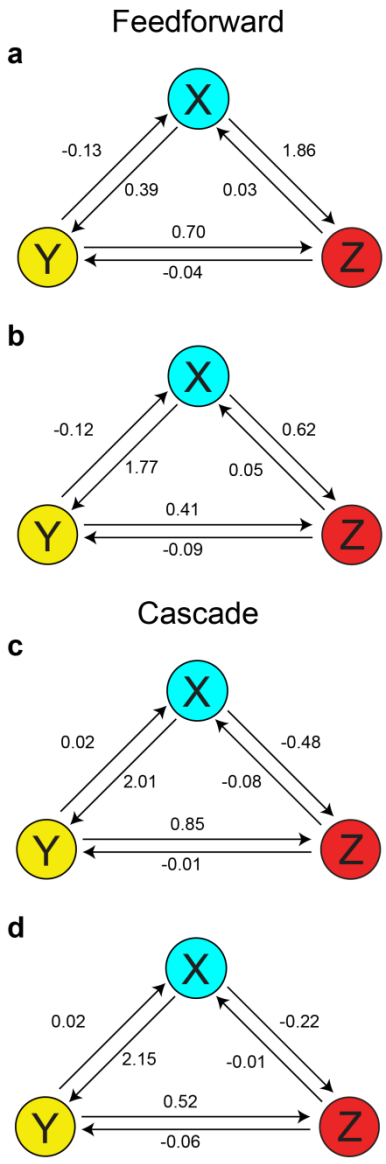

Figure S12. Graphical representation of the complete circuit topology derived from single cell data. Interaction coefficients for all regulatory links were calculated using modular response analysis using resampled fluorescence data. Circuit topology of the feedforward architecture following weak and strong perturbation are shown in (a, b), respectively. Corresponding results from the cascade architecture are shown in (c, d).

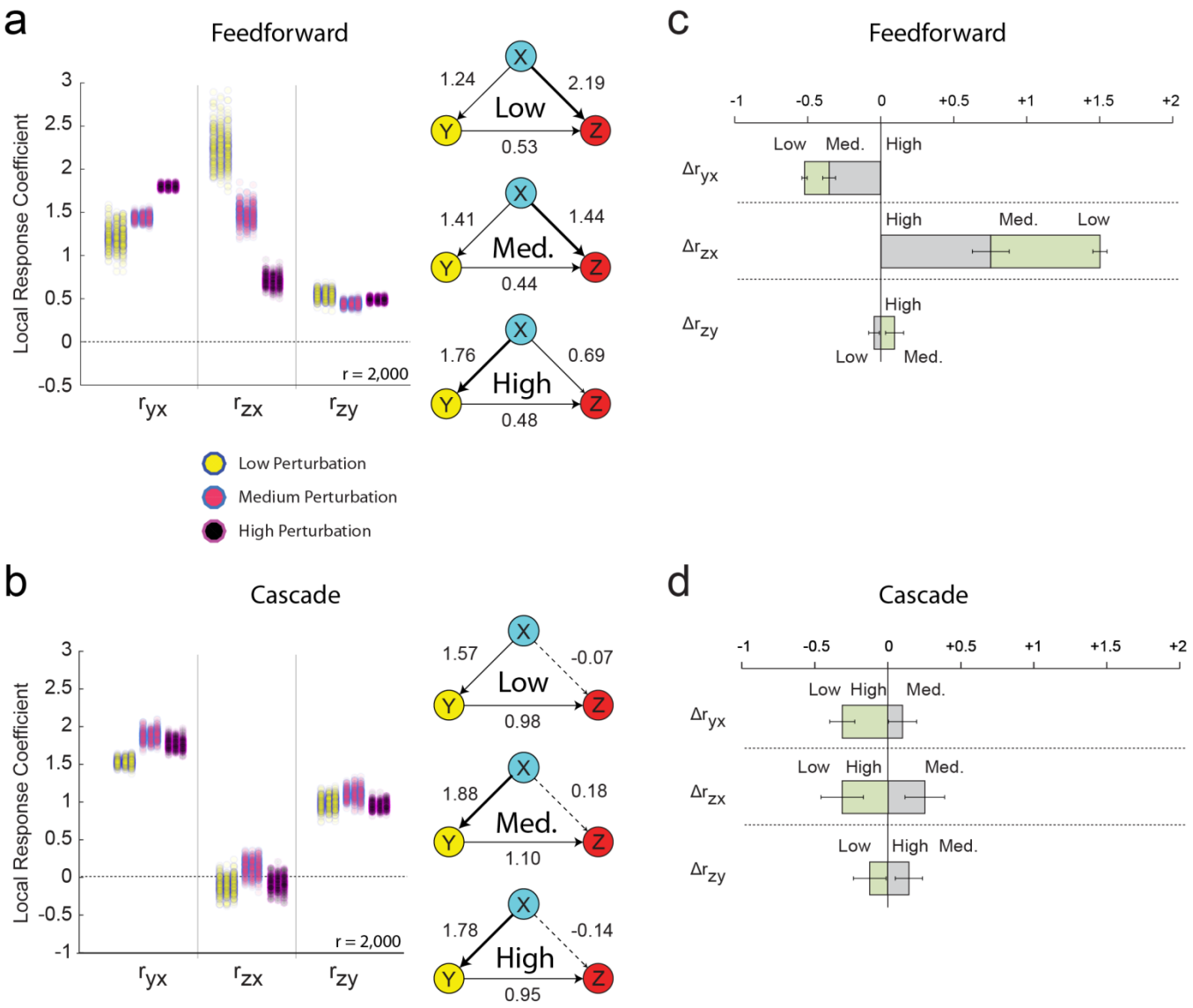

Figure S13. Reverse engineering of the benchmark topologies using resampled single-cell data. The complete reconstruction of the feedforward (a) and cascade (b) with modular response analysis performed after three different magnitudes of perturbations. For every set of subsampled means that make up the probability distributions, local response coefficients are calculated. This process cycle is performed 2,000 times, and the resulting local response coefficient distribution is plotted as a 1 dimensional scatter plot, and the corresponding graphical representation of the reconstructed synthetic networks with the mean values of these distributions are shown on the right (From top to bottom: low, medium and high perturbations, respectively). After reconstruction of the synthetic networks using three distinct sets of systemic perturbation, the change in response coefficients of equivalent edges are calculated after subsequent decrease in perturbation magnitude for feedforward (d) and cascade (d). The response coefficients recovered after the strongest perturbation sets ("High") are used as a reference point. The uncertainty associated with these values was obtained by error propagation based on the standard deviations of the original distributions.

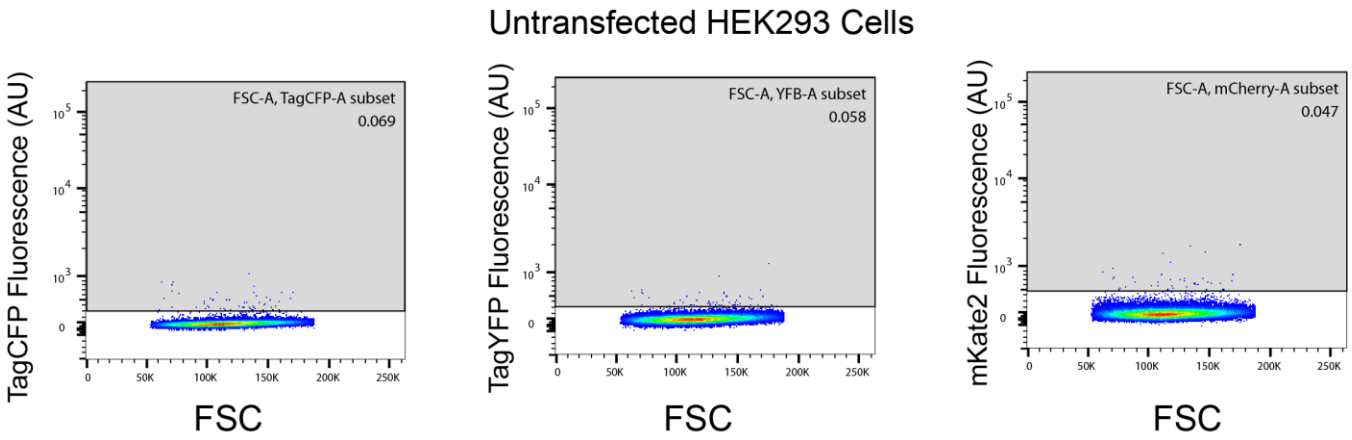

Figure S14. Fluorescence threshold used for gating. The threshold units used to select for fluorescence-positive cells were selected based on fluorescence profile of the untransfected HEK 293 cells.

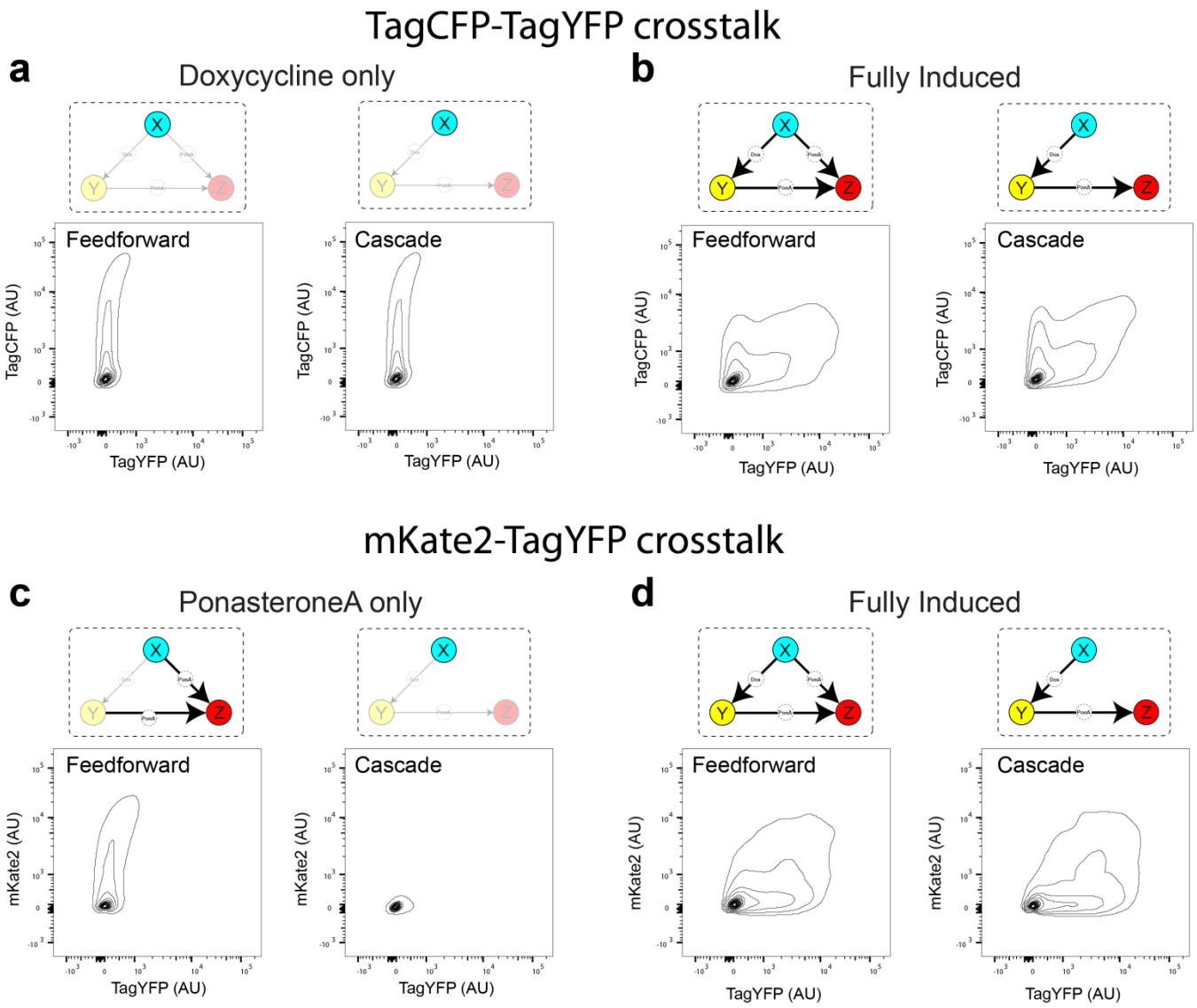

Figure S15. Crosstalk observed at working plasmid concentration. To check for potential cross-talk between the three fluorescent reporters, TagCFP and mKate2 proteins were expressed in the absence of TagYFP (a,c) and with active expression of TagYFP (b,d).

SimBiology Model

The SimBiology reaction array models for Supplementary Figures 1 and 2:

GeneX -> mRNA_X + GeneX

mRNA_X -> null

mRNA_X -> Pr_CFP + Pr_RheoDimer + Pr_rtTA + mRNA_X

Pr_CFP -> null

Pr_rtTA -> null

Pr_RheoDimer -> null

Pr_rtTA + GeneY <-> [Pr_rtTA-GeneY]

[Pr_rtTA-GeneY] -> mRNA_Y + [Pr_rtTA-GeneY]

mRNA_Y -> mRNA_Y + Pr_YFP + Pr_RheoDimer

mRNA_Y -> null

Pr_YFP -> null

Pr_RheoDimer + GeneZ <-> [Pr_RheoDimer-GeneZ]

[Pr_RheoDimer-GeneZ] -> [Pr_RheoDimer-GeneZ] + mRNA_Z

mRNA_Z -> Pr_mKate2 + mRNA_Z

mRNA_Z -> null

Pr_mKate2 -> null

Modular response analysis (MRA) method

We define the differential equations describing the system's dynamics as the following general form:

$$\dot{x}_i = F_i(x_1, \ldots x_N, p_i), \quad i = 1, \ldots, N$$

At steady state, the same equation can be written as:

$$\dot{x}_i = F_i(x_1, \ldots x_N, p_i) = 0$$

The MRA method amounts to the following procedure. We are interested in finding the direct dependencies of each variable x_i on each other variable x_j. First, we describe the system based on the implicit function theorem:

$$x_i = x_i(x_1(p_k), x_2(p_k), \ldots, x_{i-1}(p_k), x_{i+1}(p_k), \ldots x_N(p_k), p_i), \; k = 1, \ldots, N$$

When we take the derivative of the above with respect to p_j when $j \neq i$ and follow the multivariable chain rule, we obtain:

$$\frac{dx_i}{dp_j} = \sum_{k \neq i} \frac{\partial x_i}{\partial x_k} \cdot \frac{dx_k}{dp_j} + \frac{\partial x_i}{\partial p_i} \cdot \frac{dp_i}{dp_j} \tag{1}$$

Since p_i and p_j are mutually independent perturbations placed on different species, $\frac{dp_i}{dp_j} = 0$. Now (1) can be rewritten as:

$$\frac{dx_i}{dp_j} = \sum_{k \neq i} \frac{\partial x_i}{\partial x_k} \cdot \frac{dx_k}{dp_j} \qquad \forall\, j \neq i \tag{2}$$

Invoking the chain rule once again, we arrive at:

$$\frac{dx_i}{dx_j} \cdot \frac{dx_j}{dp_j} = \sum_{k \neq i} \frac{\partial x_i}{\partial x_k} \cdot \frac{dx_k}{dx_j} \cdot \frac{dx_j}{dp_j} \qquad \forall\, j \neq i \tag{3}$$

Since $\frac{dxj}{dpj}$ is usually non-zero, we divide this term from both sides of equation (3), which becomes:

$$\frac{dx_i}{dx_j} = \sum_{k \neq i} \frac{\partial x_i}{\partial x_k} \cdot \frac{dx_k}{dx_j} \qquad \forall\, j \neq i \tag{4}$$

In (4), the partial differential term $\frac{\partial x_i}{\partial x_k}$ indicates the "localized" response of species x_i when an infinitesimal perturbation is introduced on species x_k around its steady state, while the activity of other species are "clamped" and kept as constants. Thus, we define the local response coefficients as:

$$r_{ik} = \frac{\partial x_i}{\partial x_k}, \text{ and } r_{ii} = -1 \tag{5}$$

Moreover, the total differential term $\frac{dx_k}{dx_j}$ indicates the "global" response of species x_k when a perturbation is introduced on the species x_j around its steady state. Thus we define the global response coefficients as:

$$R_{kj} = \frac{dx_k}{dx_j}, \text{ and } R_{jj} = 1 \tag{6}$$

Now we plug the definitions (5) and (6) into equation (4), it becomes:

$$R_{ij} = \sum_{k \neq i} r_{ik} \cdot R_{kj} \qquad \forall\, j \neq i \tag{7}$$

Equation (7) relates the global response coefficients with local response coefficients, through which local response coefficients can be solved when the global response coefficients are available.

Reverse engineering of the benchmark topologies using population data

For each set of perturbation responses, the global response coefficients were calculated based on the weighted mean fluorescence of gated populations. The pairwise sensitivity coefficients were then obtained via calculating the local response coefficient (LRC) (Fig. S8C and S8F).

For both perturbation magnitudes, we were able to validate the positive regulatory link from node X to nodes Y and Z, as well as a positive link from node Y to node Z. All other edges had negligible interaction coefficients (Fig. S9). To determine the significance of the recovered LRC we performed error propagation analysis using Monte Carlo simulations, in order to probe whether the 95% confidence interval includes zero (Fig. S10). In this case, we observe that several regulatory links are not statistically significant (Fig. S10, dashed line edges).

In contrast to the feedforward loop where node Z is under the regulation of both nodes X and Y, the node Z in a cascade architecture has a single activation source (node Y). Therefore, for the cascade network, we hypothesize that node Z will be more sensitive to node Y disruption. Indeed, after performing the node-wise perturbation of the cascade using the siRNAs, we were able to confirm this pattern. We observe that the perturbation of node Y in the cascade architecture (Fig. S8E) has strong impact on mKate2 as compared to the same experiment on the feedforward loop (Fig. S8B).

The inference results point to higher importance of the connection between nodes Y and Z, as compared to the feedforward loop. Interestingly, MRA also recovered a direct inhibitory connection between nodes X and Z for both perturbation magnitudes which may be attributed to mild retroactivity effects. Finally, in this case, the error propagation analysis rendered most of the predicted regulatory connections as insignificant.

Monte Carlo simulation

For MRA performed using statistics obtained at a population level, the errors of calculated local response coefficient r_{ij} were projected using Monte Carlo simulations. In each cycle of the simulation, the global response coefficient was calculated using a hypothetical value of the steady-state concentrations x_i. This hypothetical value was obtained by randomly drawing from a normally distributed population with experimentally obtained mean and standard deviation values of the corresponding triplicate steady state measurement. The local response coefficient was calculated from the simulated global response matrix as described above. The simulation was repeated for 2,000 times, and the resulting distribution of the local response coefficients are plotted as a probability density function. Lastly, 95% confidence interval of each simulated local response coefficients, an interval bound by the 2.5^{th} and 97.5^{th} percentile values of the distribution, is calculated to determine whether a simulated local response coefficient is significantly different than zero.

Phenomenological Model

Consider the following set of equations to describe our synthetic three-node feedforward system:

$$\dot{x} = 1 - p \cdot x \tag{8}$$

$$\dot{y} = f(x) - q \cdot y \tag{9}$$

$$\dot{z} = g(x) + h(y) - z \tag{10}$$

The functions f, g and h represent rates of production for nodes x, y and z, respectively. In our specific experimental design, function f signifies activity by Tet-on transactivator and functions g and h indicate activities of RheoSystem dimer. We assume f(0) = g(0) = h(0), indicating that no production occurs if there is no substrate. Furthermore, p and q denote parameters for modulation of nodes x and y, which were experimentally achieved by the transient transfection of siRNAs. In other words, for an unperturbed system, p = 1 and q = 1, respectively. Therefore, a "weak" perturbation corresponds to p ≈ 1 and q ≈ 1, while a "strong" perturbation means that p and q will approach ∞. Notice that cascade motif is described by the special variation of this model where g = 0.

First, we compute the local response coefficient under weak perturbation. At the nominal steady state, $\bar{x} = 1$, $\bar{y} = f(1)$, and $\bar{z} = g(1) + h(f(1))$. Based on the definitions of local response coefficients provided in (5), we can obtain the following expression for local response coefficients:

$$r_{yx} = f'(1) \tag{11}$$

$$r_{zy} = g'(1) \tag{12}$$

$$r_{zx} = h'(f(1)) \tag{13}$$

To calculate the local response coefficient following a strong perturbation, we calculate the differences between the steady state coordinates when p = 1, q = 1 minus steady state coordinates when p = p' (p' approaches ∞), q = 1. Same rationale is applied to the perturbation of parameter q. The general expressions for the differences in each node are follows:

$$D_{xp} = \frac{1}{p'} - 1; D_{xq} = 0 \tag{14}$$

$$D_{yp} = f\left(\frac{1}{p'}\right) - f(1); D_{yq} = \frac{f(1)}{q'} - f(1) \tag{15}$$

$$D_{zp} = g(\frac{1}{p'}) + h(f\left(\frac{1}{p'}\right)) - g(1) - h\big(f(1)\big); D_{zq} = h(\frac{f(1)}{q'}) - h(f(1)) \tag{16}$$

As p' and q' both approach ∞, (14), (15) and (16) can be simplified:

$$D_{xp} = -1; D_{xq} = 0$$

$$D_{yp} = -f(1);\ D_{yq} = -f(1)$$

$$D_{zp} = -g(1) - h\big(f(1)\big);\ D_{zq} = -h(f(1))$$

Thus, the corresponding global response coefficients are:

$$R_{xx} = \frac{D_{xp}}{D_{xp}};\ R_{yx} = \frac{D_{yp}}{D_{xp}}; R_{zx} = \frac{D_{zp}}{D_{xp}};$$

$$R_{xy} = \frac{D_{xq}}{D_{yq}};\ R_{yy} = \frac{D_{yq}}{D_{yq}}; R_{zy} = \frac{D_{zq}}{D_{yq}};$$

From these we can obtain local response coefficients of the connections that make up our network:

$$r_{yx} = \frac{R_{yx}}{R_{xx}} = f(1) \tag{17}$$

$$r_{zx} = \frac{R_{zx} - r_{zy} \cdot R_{yx}}{R_{xx}} = g(1) \tag{18}$$

$$r_{zy} = \frac{R_{zy}}{R_{yy}} = \frac{h(f(1))}{f(1)} \tag{19}$$

Assume functions f, g and h can be described by Hill function:

$$\alpha(x) = \frac{x^n}{(K + x^n)}$$

For a faithful approximation of the system, we choose biologically relevant parameters. For example, for function f we choose $n_f = 2$, since the rtTA transactivator forms dimers. For functions g and h, which describe the activity of RheoDimer transactivator, we choose $n_g = n_h = 3$. Since rtTA is constitutively transcribed and upstream of RheoDimer proteins, we assume that $K_f > K_g = K_h$. We chose $K_f = 0.1$ and $K_g = K_h = 3$. With these parameters, we can estimate the local response coefficients. The results, alongside the values obtained from our experiments, are summarized in Table 1.

We emphasize that while the qualitative behavior in is the same, no quantitative agreement is sought nor is even possible, since we are using a phenomenological model. To highlight the qualitative changes we use color indicators in Tables S1: green indicates increase in calculated local response coefficient after decrease in perturbation magnitude, and red indicates decrease in calculated local response following the same change.

	Estimated			Experimental		
	Low Perturbation	High Perturbation	Difference (r_{low} - r_{high})	Low Perturbation	High Perturbation	Difference (r_{low} - r_{high})
r_{yx}	0.17	0.91	-0.74	0.39	1.77	-1.38 ± 0.17
r_{zx}	0.56	0.25	0.31	1.86	0.62	1.24 ± 0.52
r_{zy}	0.53	0.22	0.31	0.70	0.41	0.29 ± 0.24

Table S1. Local response coefficient following varying perturbation magnitudes for feedforward loop

We repeated the same procedure for estimation of local response coefficients in cascade motif. As mentioned previously, the cascade model is a special variation of the feedforward loop where (8) and (9) are the same, but g = 0 in (10). Thus, the only change in our local response coefficient after weak perturbation is in r_{zx}, as described in (11), which is now zero. Similarly, the differences in our local response coefficient following a strong perturbation are in (18), which also becomes zero.

	Estimated			Experimental		
	Low Perturbation	High Perturbation	Difference (r_{low} - r_{high})	Low Perturbation	High Perturbation	Difference (r_{low} - r_{high})
r_{yx}	0.17	0.91	-0.74	2.01	2.15	-0.14 ± 0.38
r_{zx}	0	0	0	-0.48	-0.22	0.26 ± 0.41
r_{zy}	0.53	0.22	0.31	0.85	0.52	0.33 ± 0.14

Table S2. Local response coefficient following varying perturbation magnitudes for cascade

Supplementary Methods

Mammalian Cell Culture and Transfections: HEK293 cell line was maintained at 37C, 100% humidity and 5% CO_2. The complete growth medium for the cells consists of Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 10% Fetal Bovine serum (Atlanta Biologicals), 0.1mM nonessential amino acids (Invitrogen) and 0.045units/mL of penicillin and streptomycin (Invitrogen). The cells were grown in a T-75 flask (Corning) and was passed when it reached 80% confluency. To pass the cells, the culture was washed with PBS (Life Technologies) and then trypsinized with 0.25% Trypsin-EDTA (Invitrogen). The cells were then reseeded at 40% density in a new flask. Circuit plasmid transfection was performed with jetPRIME (Polyplus) in 12-well plates at a plating density of 200,000 cells. Transfection was performed 24 hours after seeding, and each well received 10ng of plasmid containing node X and 25ng of plasmid containing nodes Y and Z, with 500ng of co-transfection junk DNA and varying amounts of siRNA. Inducer ligands doxycycline and ponasterone A, were added immediately following transfection. Perturbations to the synthetic network were applied by co-transfecting the network plasmid with three different siRNAs. For unperturbed control population, cells were co-transfected with a scrambled siRNA. Fluorescent reporter activity after various perturbation scenarios were acquired by flow cytometry.

Fluorescence Microscopy: The cells were grown on 12-well plates (Greiner Bio-One) in complete media for transfection. Approximately 48 hours after transfection of network plasmid, fluorescence images of live cells were captured using the Olympus IX81 microscope. For ambient temperature control, the entire apparatus was housed in a Precision Control environmental chamber. The images were captured using a Hamamatsu ORCA 03 Cooled monochrome digital camera. The filter sets (Chroma) are as follows: ET436/20x (excitation) and ET480/40m (emission) for TagCFP, ET500/20x (excitation) and ET535/30m (emission) for TagYFP, and ET560/40x (excitation) and ET630/75m (emission) for mKate2. Image acquisition and post-acquisition analysis was performed using Slidebook 5.0.

Flow Cytometry: All FACS experiments were performed 48 hours after transfection with BD LSRFortessa. To prepare live cells for flow cytometry, they were trypsinized with 0.25ml of 0.25% trypsin-EDTA for 3 minutes and pelleted by centrifugation at 4000 rpm for 2 minutes. The pellet was resuspended in 0.4 ml of PBS (Life Technologies). TagCFP protein was detected with a 445nm laser and a 470/20 band-pass filter, TagYFP with a 488-nm laser, 525 long-pass filter and 545/35 band-pass filter, mKate2 with a 561-nm laser 600 long-pass filter with 610/20 band-pass filter. At least 200,000 events were collected from each well. Data acquisition was performed using FACS Diva software. Subsequent analysis, including population gating, of the flow cytometry was performed with FlowJo (Treestar). The threshold fluorescence unit for selecting fluorescence-positive population was determined based on untransfected HEK293 cells (Figure S14). There was no compensation performed (Figure S15).

Quantitative reverse transcription PCR (qRT-PCR): 48 hours after co-transfection of circuit plasmid and node-specific siRNA, total RNA of the population was harvested using the RNeasy Mini kit (Qiagen) according to manufacturer's suggestion. 1 μg of total RNA was reverse transcribed to cDNA using QuantiTect Reverse Transcription kit (Qiagen). Quantitative PCR assays were performed with the Mastercycle ep realplex thermal cycler (Eppendorf) using the KAPA SYBR FAST qPCR kit (KAPA Biosystems). The relative mRNA expression levels of each node in the synthetic gene circuit were quantified with ΔΔCt method, using GAPDH as normalization factor. Primers used to amplify GAPDH and transcripts from each node are listed in Table S3. Amplification started with an enzyme activation step at 95°C for 3 minutes, followed by 45 cycles consisting of 3 seconds of denaturation step at 95°C and 20 seconds of annealing/extension at 60°C.

Recombinant DNA plasmids

All restriction enzymes are from New England Biolabs unless otherwise stated. The sequences for the synthetic parts and primers are provided in tables S3 and S4.

FF3-TagCFP-pBI-CMV4-rtTA-IRES-RheoReceptor-FF3 The backbone plasmid containing the constitutive bidirectional promoter pBI-CMV4 (Clontech) was digested with EcoRI and XbaI and gel purified. TagCFP open reading frame was amplified from pTagCFP-N (Evrogen) using primers P1 and P2 that contain EcoRI and XbaI restriction sites. This PCR product was digested with EcoRI and XbaI and re-purified. The digested insert TagCFP was ligated into the digested backbone to create pBI-CMV4-TagCFP. This plasmid serves as the backbone for the rtTA-IRES-RheoReceptor insert. To generate the insert, rtTA-IRES was amplified from pTRE-Tight-BI-LacI-IRES-Green-LacO-rtTA-IRES-dsRED (Bleris, unpublished) using P3 and P4. RheoReceptor open reading frame was amplified using P5 and P6 from pNEBR-R1 (New England Biolabs) and reamplified using P5 and P7 to miR-FF3 target at 3' end. To create rtTA-IRES-RheoReceptor insert, rtTA-IRES was combined with RheoReceptor via overlap PCR with P3 and P7. The overlap PCR product was digested with NheI and EcoRV and ligated with pBI-CMV-TagCFP digested with the same enzymes.

FF3-RheoActivator-IRES-FF3-TagCFP-pBI-CMV4-rtTA-IRES-RheoReceptor-FF3 The aforementioned plasmid FF3-TagCFP-pBI-CMV4-rtTA-IRES-RheoReceptor-FF3 was amplified using P8 and P9 and digested with AgeI and NotI to serve as the backbone. The same plasmid was also to purify IRES sequence using P10 and P11. In parallel, miR-FF3 target was added to the 3' of RheoAcceptor open reading frame from pNEB-R1 (New England Biolabs) using P12 and P13. Next, we performed overlap PCR with P10 and P13 to create the amplicon RheoAcceptor-IRES insert. This amplicon was digested with AgeI and NotI and ligated with the backbone.

FF3-RheoActivator-IRES-FF3-TagCFP-pBI-CMV4-rtTA-FF3-IRES-RheoReceptor-FF3 The miR-FF3 target was added to the open rtTA reading frame with P14 and P15, using FF3-TagCFP-pBI-CMV4-rtTA-IRES-RheoReceptor-FF3 as the template. As the common enzyme site for ligation only existed at 5' end (MluI), this amplicon was extended with another fragment purified with P16 and P17 using same template to add a common enzyme site at 3' end (BlpI). The amplicons were combined via overlap PCR with P14 and P17, and was digested with MluI and BlpI. FF3-TagCFP-pBI-CMV4-rtTA-IRES-RheoReceptor-FF3 digested with MluI and BlpI was used as a backbone and ligated with the combined fragment.

FF3-RheoActivator-IRES-FF3-TagCFP-pBI-CMV4-rtTA-FF3-IRES-Δ1255_1257del_RheoReceptor-FF3
Mutation to the RheoReceptor open reading frame was introduced in the same procedure as the miR-FF3 target site addition to rtTA open reading frame (see above). The miR-FF3 target was added to the open rtTA reading frame with P14 and P15, using FF3-TagCFP-pBI-CMV4-rtTA-IRES-RheoReceptor-FF3 as the template. As the common enzyme site for ligation only existed at 5' end (MluI), this amplicon was extended with another fragment purified with P16 and P18 using same template to add a common enzyme site at 3' end (BlpI). In addition to

extending the amplicon for ligation, this purification step introduced base pair additions in RheoReceptor open reading frame designed to induce nonsense frameshift mutation. The amplicons were combined via overlap PCR with P14 and P18, and was digested with MluI and BlpI. FF3-TagCFP-pBI-CMV4-rtTA-IRES-RheoReceptor-FF3 digested with MluI and BlpI was used as a backbone and ligated with the combined fragment.

FF3-Δ697_699del_RheoActivator-IRES-FF3-TagCFP-pBI-CMV4-rtTA-FF3-IRES-Δ1255_1257del_RheoReceptor-FF3 Mutation to the RheoActivator open reading frame was initiated by purification of RheoActivator with an extraneous base pair introduced by the 3' end primer. The PCR reaction was performed with P19 and P20, using FF3-RheoActivator-IRES-FF3-TagCFP-pBI-CMV4-rtTA-FF3-IRES-RheoReceptor-FF3 as the template. The amplicon was digested with AgeI and Bsu36I and ligated with the backbone plasmid FF3-RheoActivator-IRES-FF3-TagCFP-pBI-CMV4-rtTA-FF3-IRES-Δ1255_1257del_RheoReceptor-FF3, which was also digested with AgeI and Bsu36I.

5xGal4-mKate2 pNEB-X1 Hygro (New England Biolabs) digested with HindIII-HF and NheI-HF and was used as a backbone. Fluorescent protein mKate2 open reading frame was purified from pmKate2 (Evrogen) using P21 and P22. The amplicon was digested with HindIII and NheI and ligated with the backbone.

FF4-TagYFP-pTRE-TightBI Fluorescent protein TagYFP open reading frame was amplified from pTagYFP (Evrogen) using P23 and P24 and digested with EcoRI-HF and XbaI. pTRE-TightBI (Clonetech) was digested with XbaI and EcoRI and ligated with the digested TagYFP amplicon.

FF4-TagYFP-PTREBI-RheoAcceptor-FF4 We amplified RheoAcceptor from pNEB-R1 (New England Biolabs) using oligos PT and PU and digested with KpnI-HF and NotI-HF. FF4-TagYFP-PTREBI was cut with the same enzymes and ligated with the digested amplicon to.

FF4-TagYFP-PTREBI-RheoAcceptor-IRES-RheoReceptor-FF4—5xGal4-mKate2 FF3-TagCFP-BICMV4-rtTA-IRES-RheoReceptor-FF3 as template IRES-RheoReceptor was amplified with oligos P25 and P26 and digested with NotI-HF and EcoRV-HF. We digested FF4-TagYFP-PTREBI-RheoAcceptor-FF4 with NotI-HF and EcoRV-HF and ligated with the digested IRES-RheoReceptor amplicon to generate FF4-TagYFP-PTREBI-RheoAcceptor-IRES-RheoReceptor-FF4. We amplified FF4-TagYFP-PTREBI-RheoAcceptor-IRES-RheoReceptor-FF4 using oligos P27 and P28 (including Ampicillin resistance and origin of replication), digested with AgeI-HF and CIP treated. Next, we amplified 5xGal4-mKate2 with P29 and P30 and digested with AgeI-HF. Both digested PCR amplicons were ligated to yield FF4-TagYFP-PTREBI-RheoAcceptor-IRES-RheoReceptor-FF4—5xGal4-mKate2.

Table S3. Primers

P1	CAGTACACGCGTGCTAGCGCCACCATGTCTAG	TagCFP Forward
P2	CCAGTAGAATTCGCCACCATGAGCGGGGGCGAGGAGCTGTTCG	TagCFP Reverse
P3	CCAGTAGCTAGCATGTCTAGACTGGACAAGAGCAAAG	rtTA Forward
P4	ATAGAAGACAGTAGCTTCATGGTTGTGGCCATATTATCAT	IRES Rheo Receptor Overlap
P5	ATGATAATATGGCCACAACCATGAAGCTACTGTCTTCTAT	IRES Rheo Receptor Overlap (Reverse)
P6	TTTGTATTCAGCCCATATCGTTCTAGAGATTCGTGGGGGACTCGA	Rheo Receptor Reverse with FF3
P7	CCAGTAGATATCTTTGTATTCAGCCCATATCGTT	FF3 (Rheo Receptor Reverse)
P8	CCAGTAGCGGCCGCCTAGATTTGTATTCAGCCCATATCGTTTTAGCGG	TagCFP Reverse
P9	CCAGTAACCGGTAGGATCATAATCAGCCATACCACATTT	SV40 Poly A Reverse
P10	CCAGTAGCGGCCGCTAGCGCTACCGGACTCAGATCTCGAG	IRES Reverse
P11	CGCTTCTTTTTAGGGCCCATGGTTGTGGCCATATTATCAT	RheoActivator IRES Overlap
P12	ATGATAATATGGCCACAACCATGGGCCCTAAAAAGAAGCG	RheoActivator IRES Overlap (Reverse)
P13	CCAGTAACCGGTTTTGTATTCAGCCCATATCGTTTTATGAATCAGAAGGTGATT	RheoActivator Reverse with FF3
P14	CAGTACACGCGTGCTAGCGCCACCATGTCTAG	rtTA Forward
P15	TTTGTATTCAGCCCATATCGTTTTACCCGGGGAGCATGTCAA	rtTA reverse + FF3
P16	CCGGGTAAAACGATATGGGCTGAATACAAACCTGCATTAGCGCTACCGGA	rtTA_IRES_forward
P17	CAGTACGCTCAGCTGGTTCAGGATATAGATGCG	RheoR_reverse
P18	CAGTACGCTCAGCATGGTTCAGGATATAGATGCG	RheoR_forward
P19	ACCGGTTTTGTATTCAGCCC	RheoA_FF3_forward
P20	CCTGAGGAGCAATCATTCTGTTTAATCCAGAG	RheoA_reverse
P21	CCAGTAGCTAGCATGGTGAGCGAGCTGATTAAGGAGAACA	mKate2 FWD
P22	CCAGTAAAGCTTTTATCTGTGCCCCAGTTTGCTAGGGAGG	mKate2 REV
P23	CCAGTATCTAGATTTAATTAAAGACTTCAAGCGGTTAGCGGTACAGCTCGTCCATGC	YFP Fwd with FF4 (Rev comp)
P24	CCAGTAGAATTCATGGTTAGCAAAGGCGAGGAGCTGTTCG	YFP Rev (Rev comp)
P25	CCAGTAGCGGCCGCTAGCGCTACCGGACTCAGATCTCGAG	IRES Reverse
P26	CCAGTAGATATCTTTAATTAAAGACTTCAAGCGGCTAGAGATTCGTGGGGGACTCGA	Rheo Receptor FF4

		Reverse
P27	CCAGTAACCGGTGCGAGCGGTATCAGCTCACTCAAAGG	Yp Origin of Replication FWD
P28	CCAGTAACCGGTCGCAGCCGAACGACCGAGCGCAGCGAG	SV40 Poly A Reverse (Node Yp)
P29	CCAGTAACCGGTCACGACGTTGTAAAACGACGGACATG	5xGal4 FWD
P30	CCAGTAACCGGTACGCGTTAAGATACATTGATGAGTTTGGAC	mKate2 Poly A Reverse
P31	AACAACGCCAAGTCATTCCG	qRT-PCR: Node X Forward
P32	TCAGCGACAACGTGTACATC	qRT-PCR: Node X Reverse
P33	ACGTTCCTGATGGAGATGCTTG	qRT-PCR: Node Y Forward
P34	TCGACTGCAGAATTCGAAGC	qRT-PCR: Node Y Reverse
P35	AACCACCACTTCAAGTGCAC	qRT-PCR: Node Z Forward
P36	TTTGCTGCCGTACATGAAGC	qRT-PCR: Node Z Reverse
P37	AATCCCATCACCATCTTCCA	qRT-PCR: GAPDH Forward
P38	TGGACTCCACGACGTACTCA	qRT-PCR: GAPDH Reverse

Table S4.

Specific parameters used for the SimBiology simulations[1]:

Parameter	Value	Units
k_GeneX_txn	0.01	1/second
k_mRNA_X_trsln	0.0001	1/second
k_mRNA_X_degr	0.00002	1/second
kON_PrX-GeneY	0.00000001	1/(molecule*second)
kOFF_PrX-GeneY	0.0001	1/second
kON_PrX-GeneZ	0.00000001	1/(molecule*second)
kOFF_PrX-GeneZ	0.0001	1/second
k_mRNA_Y_trsln	0.0001	1/second
k_mRNA_Y_degr	0.00002	1/second
k_mRNA_Z_degr	0.00002	1/second
k_PrX-GeneY_txn	0.01	1/second
k_PrX-GeneZ_txn	0.01	1/second
k_mRNA_Z_trsln	0.0001	1/second
k_Pr_degr	0.00002	1/second

References

1. Bleris, L., Xie, Z., Glass, D., Adadey, A., Sontag, E., and Benenson, Y. (2011) Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template *Mol. Syst. Biol. 7*, 519.