Functional resilience of mutually repressing motifs embedded in larger networks

Running title: Resilience of toggle switch and triad dynamics

Pradyumna Harlapur[1], Atchuta Srinivas Duddu[2], Kishore Hari[2], Mohit Kumar Jolly[2,*]

1 Department of Biological Sciences, Indian Institute of Science Education and Research, Bhopal 462066
2 Center for BioSystems Science and Engineering, Indian Institute of Science, Bangalore 560012

*Author to whom correspondence should be addressed: mkjolly@iisc.ac.in (M.K.J.)

Keywords: Toggle switch, Toggle Triad, Bimodality, Dynamical robustness, Multistability

Abstract

Elucidating the design principles of regulatory networks driving cellular decision-making has important implications in understanding cell differentiation and guiding the design of synthetic circuits. Mutually repressing feedback loops between 'master regulators' of cell-fates can exhibit multistable dynamics, thus enabling multiple "single-positive" phenotypes: (high A, low B) and (low A, high B) for a toggle switch, and (high A, low B, low C), (low A, high B, low C) and (low A, low B, high C) for a toggle triad. However, the dynamics of these two network motifs has been interrogated in isolation *in silico*, but *in vitro* and *in vivo*, they often operate while embedded in larger regulatory networks. Here, we embed these network motifs in complex larger networks of varying sizes and connectivity and identify conditions under which these motifs maintain their canonical dynamical behavior, thus identifying hallmarks of their functional resilience. We show that the in-degree of a motif - defined as the number of incoming edges onto a motif - determines its functional properties. For a smaller in-degree, the functional traits for both these motifs (bimodality, pairwise correlation coefficient(s), and the frequency of "single-positive" phenotypes) are largely conserved, but increasing the in-degree can lead to a divergence from their stand-alone behaviors. These observations offer insights into design principles of biological networks containing these network motifs, as well as help devise optimal strategies for integration of these motifs into larger synthetic networks.

Introduction

Gene Regulatory Networks (GRNs) are an integral part of the control structure involved in various cellular processes such as cell-fate decisions made during embryonic development, cellular reprogramming, and phenotypic switching among two or more cell types. A pluripotent cell is capable of differentiating to more than one cell type in response to varying stimuli. This property of coexistence of more than one stable steady state (phenotypes) is referred to as multi-stability and it underlies the dynamics of many GRNs involved in decision-making during differentiation [1]. Such multi-stability has been seen during cellular reprogramming as well as phenotypic switching under many circumstances. Thus, elucidating the dynamical principles of multi-stable GRNs and network motifs holds promise for understanding many biological processes and control applications in synthetic biology[2–5].

One of the most frequently observed and extensively investigated network motifs is the 'Toggle Switch' (TS), i.e., two mutually repressing regulators A and B, each driving a different cell fate[6–8]. The TS enables two mutually exclusive "single-positive" outcomes - (high A, low B) and (low A, high B), thus showing bistable dynamics and allowing a pluripotent cell to choose from two cell fates [2,8,9]. For instance, PU.1 and GATA1 form a TS that drives hematopoietic stem cells to either a common myeloid progenitor (PU.1 high, GATA1 low) or an erythroid one (PU.1 low, GATA1 high)[2,10]. Also, in *Escherichia coli*, the construction of a TS exhibiting bistability and switching between the two states in response to external signals has driven an extensive design of synthetic genetic circuits [8,11,12]. Another network motif is a 'Toggle Triad' (TT), i.e., three mutually repressing regulators A, B, and C, each driving a respective cell fate [13–16]. TT can enable a progenitor cell to differentiate into three distinct cell fates; for instance, the case of naïve CD4+ T helper cells differentiating to Th1, Th2, and Th17 cells [13,17,18]. The three canonical "single-positive" states enabled by TT are: (high A, low B, low C), (low A, high B, low C) and (low A, low B, high C).

While the dynamics of TS and TT have been investigated extensively deterministically and stochastically, most such investigations have considered them in isolation, i.e. TS or TT are assumed to be not connected to any other network components [13–15,19–21]. However, in reality, a TS or TT is only a small part embedded in a large network of interconnected proteins and signaling components. Here, we investigate and quantify the behavior of TS and TT network motifs when embedded in much larger networks, using three properties: bimodality of individual nodes, pairwise correlation coefficient between nodes, and fraction of canonical "single-positive" states. We noticed that for a TS, an increase in the number of incoming edges on the two nodes of a TS (i.e. in-degrees for A and B) resulted in deviation from stand-alone behavior, as captured by changes in all the three abovementioned properties. Further, an asymmetry in the in-degrees for both nodes also compromised bistability. For a TT, however, only the fraction of "single-positive" states is a reliable metric to quantify its deviation from stand-alone dynamics. We observed that as the net in-degree of a TT increased, "single-positive" steady states (e.g. (high A, low B, low C)) were replaced by "double-positive" ones (e.g. (high A, high B, low C)). These observations suggest that in addition to the previously studied factors influencing the dynamics of TS or TT, the local density around these motifs, when embedded in larger networks, could also influence their functional properties.

Results

Stand-alone properties of toggle switch and toggle triad

We first investigated the stand-alone properties of a toggle switch (TS) and toggle triad (TT), before embedding them in larger networks. RACIPE formalism [22] was used to simulate these networks in triplicates for 10,000 randomized parameter sets sampled from a predetermined parameter space. The resultant output, which contained the steady state values of the nodes and their corresponding kinetic parameters, was used to calculate the pairwise correlation coefficient (CC) between steady state values of the nodes in a TS or TT, the bimodality coefficient (BC) [23] of the steady state distributions of the individual nodes within TS or TT and the frequency of the different steady states of TS or TT.

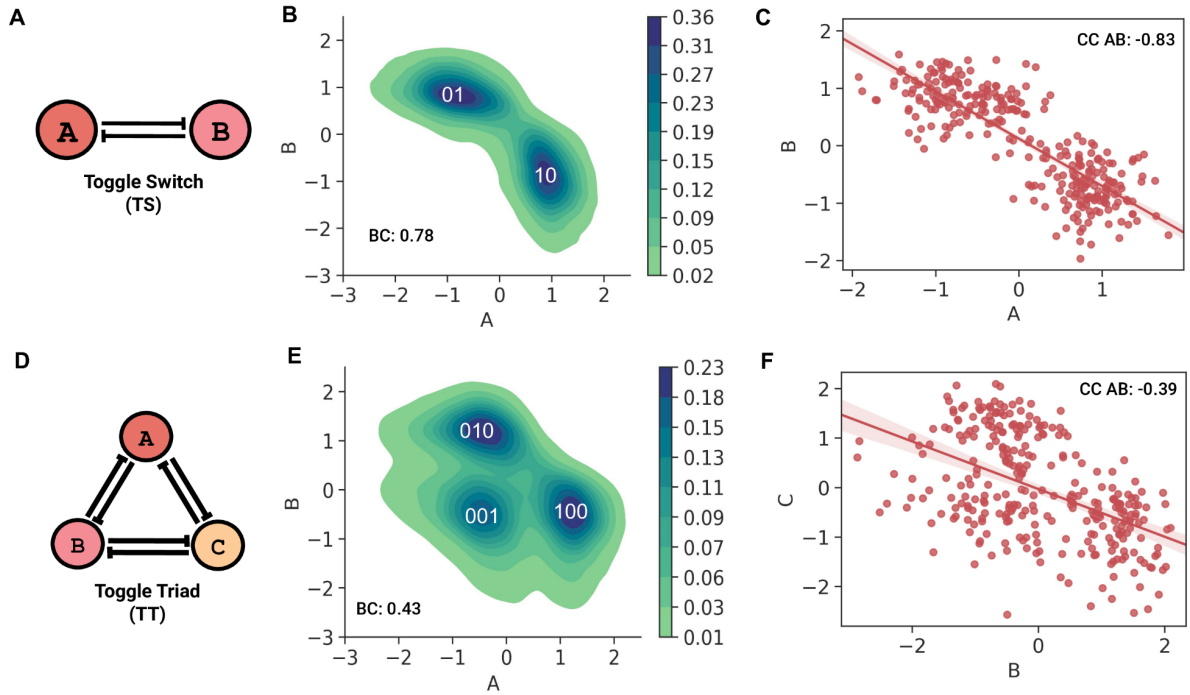

Figure 1: Stand-alone dynamics of two-node and three-node mutually repressing motifs. *(A) A toggle switch (TS) motif comprises two nodes A and B that mutually inhibit each other. (B) Probability density plot of steady state values of nodes in a TS. The two dense clusters correspond to "single-positive" 01 and 10 steady states of a TS. (C) Regression plot between the steady states values of two nodes, A and B of a TS. The correlation coefficient (CC AB) between them is -0.83. (D) A toggle triad (TT) motif comprises three mutually repressing nodes A, B, and C. (E) Probability density plot of steady-state values of two TT nodes A and B; the three clusters represent three "single-positive" steady states, 001, 010, and 100. (F) Regression plot between steady state values of nodes B and C of a TT. Correlation coefficient (CC BC) = -0.39.*

The TS motif (**Fig 1A**) mainly showed two "single-positive" steady states: ((low A, high B); (A, B) = (0, 1)) and ((high A, low B); (A, B) = (1, 0)), as observed in the bivariate plot (**Fig 1B**). The steady state values of both nodes in a TS showed a bimodal distribution (bimodality coefficient =

0.78), with the peaks representing the corresponding high and low steady state values. Because the two nodes of a TS repress each other, the correlation coefficient between steady state values of the nodes (CC AB) was strongly negative (r = - 0.83) (**Fig 1C**). The steady state values for a TT motif (**Fig 1D**), for any given pair of TT nodes, had three distinct clusters, which represent the three "single-positive" stable steady states namely ((low A, low B, high C); (A, B, C) = (0, 0, 1)) state, ((low A, high B, low C); (A, B, C) = (0, 1, 0)) state and ((high A, low B, low C); (A, B, C) = (1, 0, 0)) state. One node shows higher expression in these states while the other two nodes have repressed expression (**Fig 1E, S1A, S1B**). In contrast to TS, the average BC of a node in TT is 0.43 (standard deviation = 0.004), implying a more unimodal-like distribution with little difference in the high and low steady state values of a node in a TT. Although negative, the magnitude of pairwise correlation coefficient between the steady state values of any two nodes of a TT was less than that of a TS (**Fig 1F, S1C, S1D**). This decrease could be because although any pair of two nodes mutually repress, due to the dominance of single-positive steady states, two nodes of TT can still show similar low-expression steady state values, thus leading to a relatively lower magnitude of correlation coefficient between them. On the other hand, two nodes of a TS are strictly confined to having opposing expression profiles, leading to a strongly negative correlation.

Functional traits of toggle switch depend on density rather than the size of the larger networks it is embedded in

Next, we embedded TS and TT motifs in twelve different types of large networks having four different network orders, each having three distinct densities (mean connectivity) to understand how the abovementioned dynamic traits of TS and TT change. The four network orders were 5N, 10N, 15N and 20N, where N is the number of nodes in a network in which these motifs were embedded. The three mean connectivity considered were E:2N, E:4N, and E:6N, where E:xN signifies that the number of edges (E) is x times the number of nodes (N). These resulted in twelve groups of networks, for each of which n=100 random network topologies were generated. Thus, in a total of 1200 networks, we embedded a TS and in another 1200 ones, a TT was embedded. We used RACIPE to simulate the dynamics of these networks, and the generated output was analyzed to identify the bimodality coefficients, correlation coefficients, and fraction of canonical "single-positive" states obtained in a TS or TT. These three metrics were then compared to quantify the dynamic resilience of TS and TT motifs when embedded in larger networks **(Fig 2)**. For statistical comparisons, three replicates (three different ensembles of parameter sets and initial conditions) were used for each of the 1200 networks considered for either TS or TT.

To evaluate how network size and density influence the behavior of a TS when embedded in larger networks of varying orders and mean connectivity, we compared the behavior for networks sharing the same mean connectivity but having different network orders or *vice versa*. Interestingly, for networks with same mean connectivity, correlation coefficient between the two nodes in a TS (CC AB) did not show any significant variation for varied network orders (**Fig 3A, i**). However, when controlling for network order, CC AB reduced as the mean connectivity increased (**Fig 3A, ii**). A similar trend, i.e., dependence on network density rather than on network size and a decrease in magnitude with increasing network density, was also observed in

distributions of bimodality coefficient values: BC A and BC B **(Fig 3B, i-ii; S1E-F)**. Next, we quantified the fraction of “single-positive” - 01 (low A, high B) and 10 (high A, low B) - steady states across all these networks. Again, we observed that the trends depended on mean connectivity of the network rather than network orders. The higher the mean connectivity, the lower the frequency of these canonical states observed in a TS (**Fig 3C, i-ii**).

In certain cases, the trends seen when comparing between E:4N and E:6N mean connectivity were not significantly different, potentially because the TS dynamics had been compromised enough in the E:4N case but not in the E:2N case. Put together, all the three metrics considered here to capture the behavior of a TS - correlation coefficient (how strongly are the two nodes in a TS anti-correlated), bimodality coefficient (how clearly the high and low levels of a node are segregated), and the fraction of “single-positive” 10 and 01 states together (how strong is the dominance of mutually exclusive trend) - tend to show trends indicating a weakening of the dynamical behavior of a TS, as it is embedded in increasingly dense large networks.

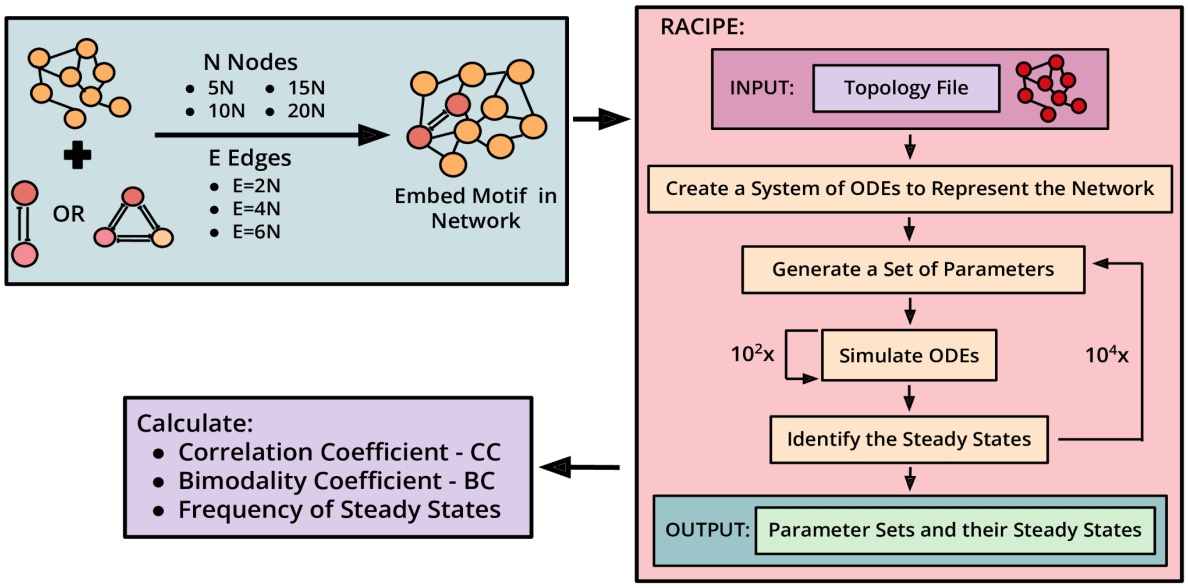

Figure 2: Schematic of the pipeline used to generate, simulate and analyze the motifs embedded in random networks.

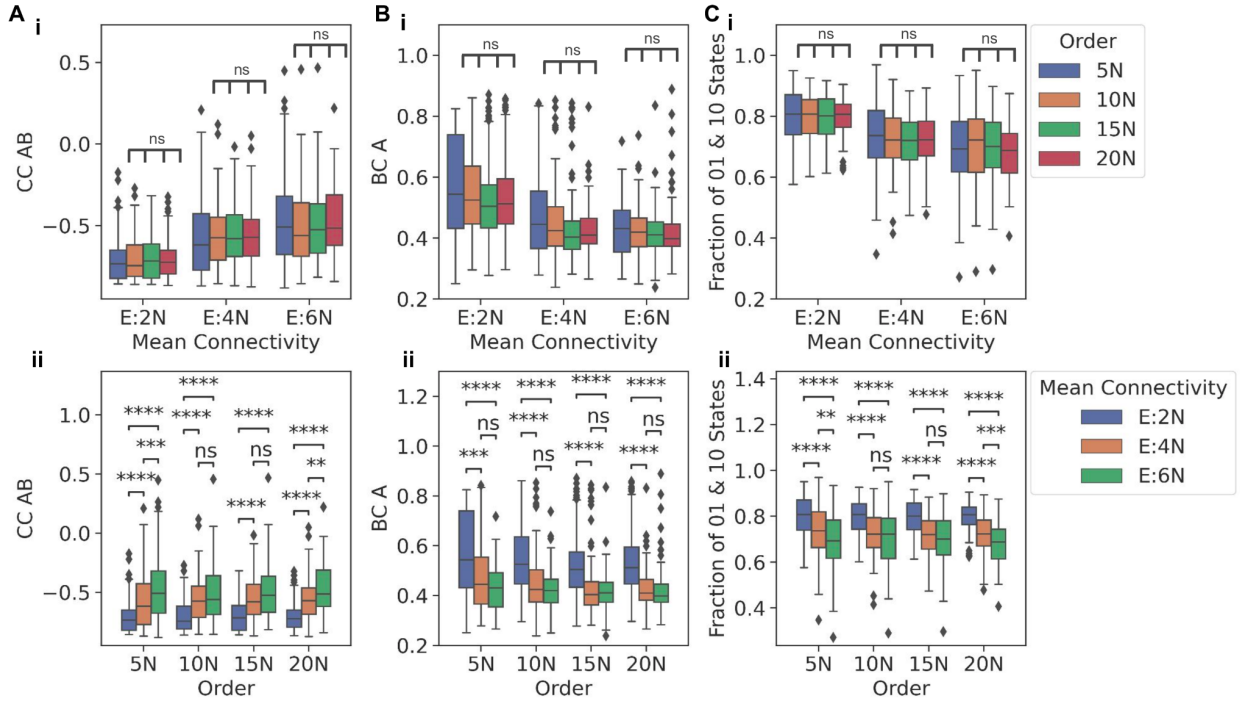

Figure 3: Functional traits of TS embedded in large networks. *(A) Comparison between the distributions of CC AB for TS embedded in: (i) networks of same mean connectivity but having different orders and (ii) networks of same order having different mean connectivity. (B) Comparison between the distributions of BC A for TS embedded in: (i) networks of same mean connectivity but having different orders and (ii) networks of same order having different mean connectivity. (C) Comparison between the distributions of the fraction of 01 & 10 states for TS embedded in: (i) networks of same mean connectivity but having different orders and (ii) networks of same order having different mean connectivity. p-values of pairwise Mann-Whitney U tests are denoted by: ns - p <= 1, * - 0.01< p <= 0.05 , ** - 0.001 < p <= 0.01, *** - 0.0001 < p <= 0.001, **** - p <= 0.0001*

Local density around a toggle switch impacts its dynamic behavior

Mean connectivity of a network is the ratio of the total number of edges to the number of nodes in the network, i.e. a measure of global network density. Thus, assuming that the network, on average, is equally sparse or dense, with an increase in mean connectivity of the network, the average in-degree of the nodes of a TS embedded in the network also increases. To ascertain whether this increase in the in-degree (In A and In B) for the TS nodes (as a consequence of the increased mean connectivity of the network) contributed to divergence from stand-alone TS dynamics, we analyzed the variation in the three metrics (CC, BC, fraction of 10 and 01 states) with the change in the in-degrees of the TS nodes. We observed that as the in-degree of both the nodes of a TS increased, the mean CC AB values decreased in magnitude (**Fig 4A, i**). For an in-degree of one for both the nodes in a TS, i.e. the case when these two nodes inhibited each other but had no other incoming edges from the larger network, we noticed a mean CC AB value of -0.83 (**Fig 4A, i**), the same as that of an isolated TS motif (**Fig 1B**). Furthermore, the magnitude of CC AB showed the fastest decline when both the nodes had equally increased in-degrees

(along the diagonal of the heatmap shown in **Fig 4A, i**). A similar trend was seen in the fraction of the "single-positive" 01 and 10 steady states: as in-degrees of both the nodes of a TS increased, the fraction of 01 and 10 steady states decreased steadily (**Fig 4A, ii**). On the other hand, for BC of a given node in a TS, only the in-degree for that node, but not the overall in-degree for a TS, had an effect (**Fig 4A, iii; S2A**). Additionally, the mean BC values for nodes with in-degree greater than two were lower than the typical cut-offs considered for bimodality, indicating that those nodes showed more unimodal distributions than the canonical bimodal distributions observed in nodes of an isolated TS.

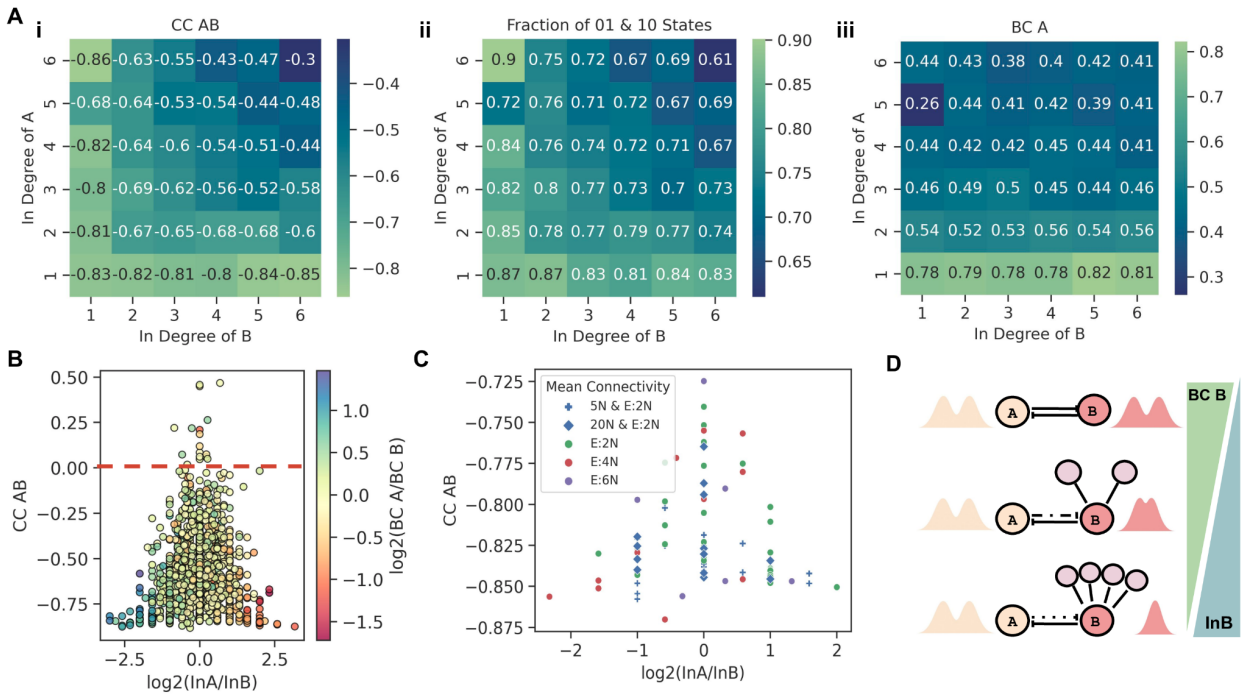

Figure 4: Influence of in-degree of TS on its functional traits. *(A) Heatmaps of values of (i) CC AB, (ii) fraction of 01 and 10 states, and (iii) BC A for varying in-degree of the two nodes in a TS. (B) Scatter plot of log_2(In A/In B) values against their CC AB values, with points colored according to their log_2(BC A/BC B) values. (C) Scatterplot of log_2(In A/In B) values against their CC AB values, with points colored according to the mean connectivity of the network in which the TS is embedded. All points in the plot represent toggle switches in which both the nodes showed bimodality. (D) Summary of the effect of local density on the dynamics of an embedded TS motif.*

Further, we quantified the impact of asymmetry in terms of incoming edges on a TS by considering the impact of the ratio of in-degree of A to in-degree of B (log_2(In A/In B)) on CC AB and relative BC values simultaneously. We noticed that the higher the asymmetry in terms of in-degree (log_2(In A/In B) > 1 or log_2(In A/In B) < -1), the stronger the negative correlation between the two TS nodes (CC AB < -0.5) (**Fig 4B**). Importantly, as the magnitude of log_2(In A/In B) increased, the range of CC AB values, which started from a wide range spanning even positive values, constricted to highly negative values closer to values observed in isolated TS nodes. Similar behavior was seen in the fraction of 01 and 10 steady states: as the magnitude of log_2(In A/In B) increased, the range of values of the fraction started to constrict to higher values tending towards those seen in an isolated TS (**Fig S2B**). Also, we noted that the higher the in-degree of a node, the smaller its

corresponding BC ($\log_2$(BC A/BC B) < -1 for $\log_2$(In A /In B) > 1, $\log_2$(BC A/BC B) > 1 for $\log_2$(In A /In B) < -1) **(Fig 4B)**. Therefore, when there is an asymmetry in the in-degrees of nodes in a TS, the node with the higher in-degree starts to lose its switch-like behavior and shows more unimodal distributions of its steady state values. However, in such scenarios, a high negative correlation seen in isolated TS is largely maintained. As the asymmetry increases, the node with the lower in-degree is still being affected by the expression of the higher in-degree node, showing an expression profile opposite to that of the higher in-degree node, thereby leading to a higher fraction of 01 and 10 states as well, as compared to a more symmetric in-degree case.

We next zoomed on to cases where both nodes in a TS showed distinct bimodal distributions (i.e. BC A and BC B > 0.55) and plotted their CC AB values against their $\log_2$(In A/In B) values (**Fig 4C**), because these sub-networks maintain the bimodal nature of both nodes as seen in an isolated TS. In such cases, the CC AB values and the fraction of single-positive states show values similar to those for an isolated TS (**Fig 4C, S2C**). A constriction in the range of $\log_2$(BC A/BC B) to values with a magnitude less than 0.6 was also seen, indicating a decreased skew in the BC ratios (**Fig S2D**). Thus, this subset of networks featured all canonical features of a TS: high fraction of "single-positive" 01 and 10 states, bimodal distributions of A and B, and strongly negative correlation between A and B. Therefore, the condition of both the nodes in an embedded TS showing bimodal distributions (BC A, BC B) for their steady state values could be used to distinguish toggle switches that display isolated TS-like dynamics. Furthermore, in this subset of networks, we observed networks across all three mean connectivity, reiterating that the local density around the TS motif (given by the in-degree of the TS nodes) leads to divergence from TS-like behavior rather than the network order or even the network's mean connectivity (**Fig 4D**).

In a toggle triad, fraction of single-positive states capture its functional resilience

After investigating the patterns seen in a TS embedded in large networks, we focused our attempt to understand the functional resilience of TT motif. We embedded it into the previously described twelve types of large random networks. Similar to observations in TS, the distributions of both CC and BC for all three cases - CC AB, CC BC, CC AC and BC A, BC B and BC C - did not show any significant consistent variation when they were grouped by mean connectivity and compared across the different network orders (**Fig 5A, i; 5B, i; S3A i-iv**). Intriguingly, unlike the observations in TS, we did not observe any significant differences when they were grouped by their network orders and compared across the three mean connectivity either (**Fig 5A, ii; 5B, ii; S3B i-iv**). On the other hand, reminiscent of observations in TS, the fraction of "single-positive" (010, 100, and 001) steady states decreased overall when comparisons were made across their mean connectivity (**Fig 5C, i**), but not across their network orders (**Fig 5C, ii**).

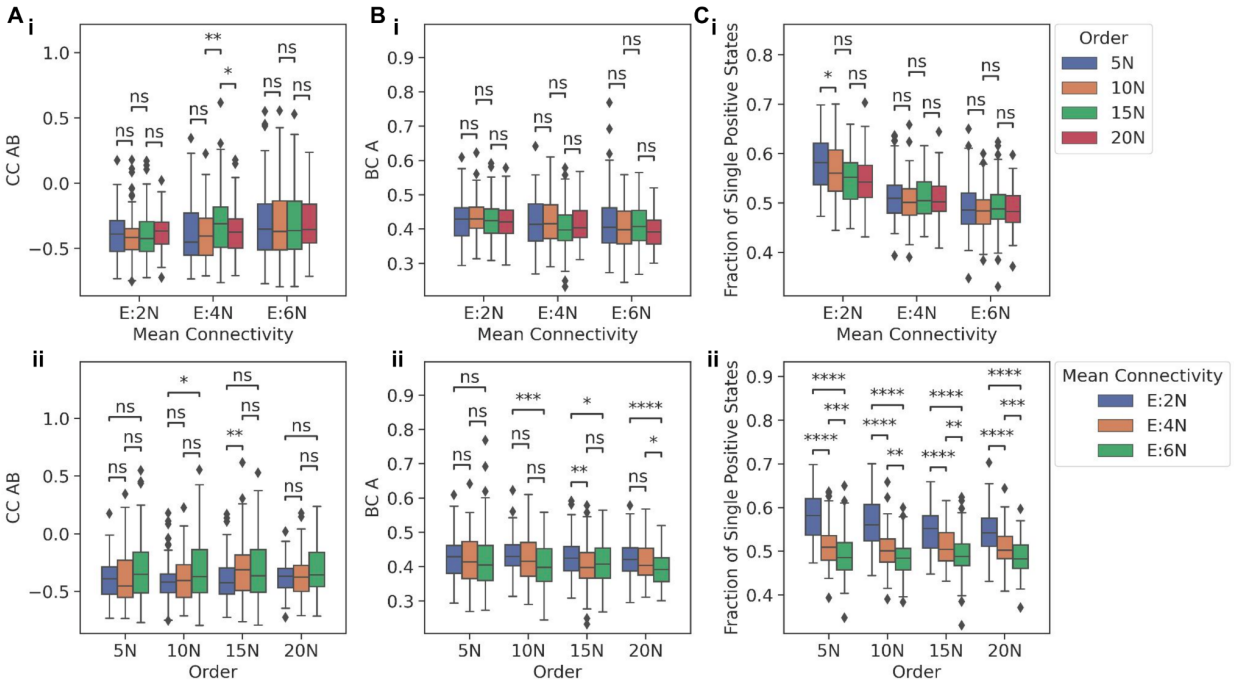

Figure 5: Functional traits of TT embedded in larger networks. *(A) Comparison between distributions of CC AB for TT embedded in (i) networks of same mean connectivity but having different orders and (ii) networks of same order having different mean connectivity. (B) Comparison between the distributions of BC A for TT embedded in (i) networks of same mean connectivity but having different orders and (ii) networks of same order having different mean connectivity. (C) Comparison between the distributions of the fraction of "single-positive" states (across all steady states observed) for TT embedded in (i) networks of same mean connectivity but having different orders and (ii) networks of same order having different mean connectivity. p-values of pairwise Mann-Whitney U tests are denoted by: ns - p <= 1, * - 0.01< p <= 0.05 , ** - 0.001 < p <= 0.01, *** - 0.0001 < p <= 0.001, **** - p <= 0.0001*

We next investigated how the BC and CC metrics are associated with varying in-degrees for the nodes in a TT. Unlike observations for TS, the CC and BC metrics, again, had no consistent trend (**Fig 6A,i-ii; S4; S5**). In a TS embedded in larger networks, the BC and CC metrics showed significant changes in their distributions when the mean connectivity of the networks and the in-degrees of the motif nodes were varied, but CC and BC metrics in TT due to their insignificant variance in distribution were not suitable for further analysis of the TT motif. Meanwhile, the fraction of single-positive states decreased as the in-degrees of any two nodes in a TT were increased (**Fig 6A, iii; S6A-B**). Thus, we took the ratio of the single-positive (100, 010, 001) steady states to that of the double-positive steady states (110, 011, 101) - (f1/f2) - to investigate which type of steady states were getting preferentially enhanced due to an increasing in-degree of the TT (**Fig 6B, i**). We observed that the f1/f2 ratio decreased with an increase in the in-degree of the TT motif, which meant that as the fraction of single-positive steady states decreased, they were being replaced by double-positive ones. As the in-degree of TT was increased, the f1/f2 ratios tended towards a value of 1, indicating an equivalent fraction of single-positive and double high steady states. The mean of f1/f2 distributions grouped by different mean connectivity also showed a decreasing trend when the mean connectivity of the network was increased. When the

distributions of f1/f2 values of different network orders having the same mean connectivity of E:2N were compared, the distributions were very similar. This showed that the previously observed pattern of only the in-degree affecting the dynamics of a motif was also applicable for the f1/f2 ratio. Similar analysis done with in-degree of the TT and the other metrics did not show this pattern, and there was not a substantial variation in the distributions of the metrics across either their in-degree of TT, mean connectivity, or network orders (**Fig 6B, ii-iii; S6C, i-iv; Table S1-2**).

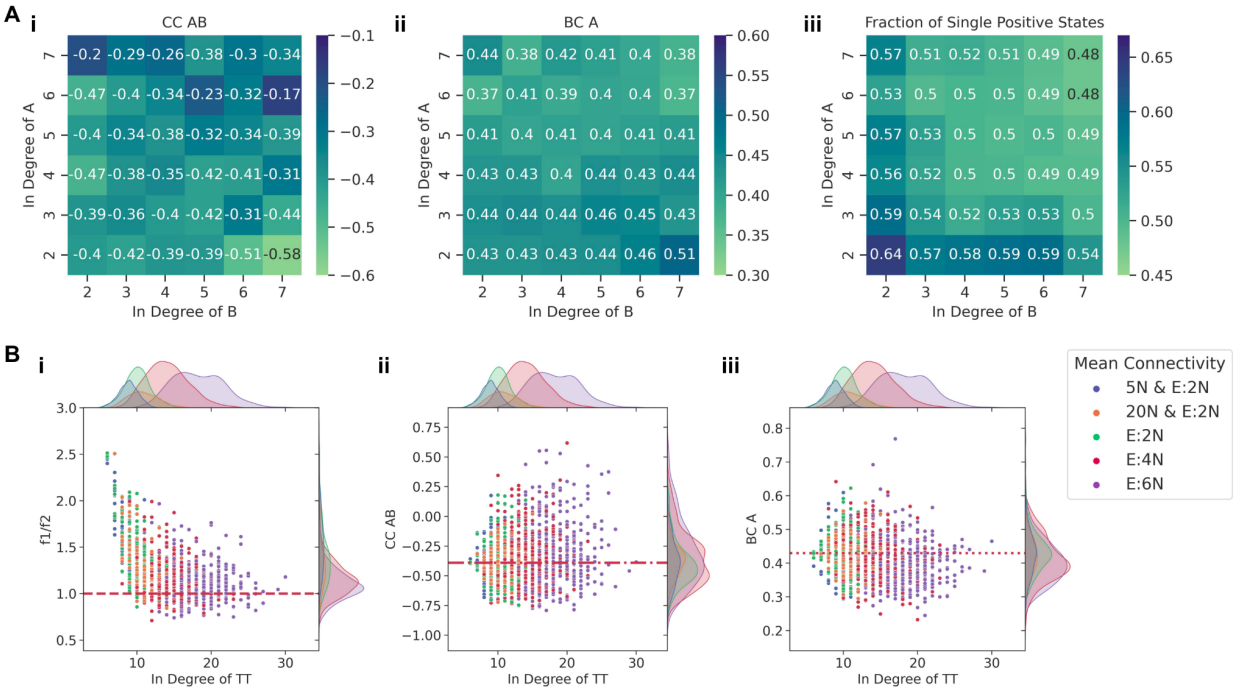

Figure 6: Influence of in-degree of TT on its dynamics. *(A) Heatmaps depicting the change in metrics: (i) CC AB, (ii) BC A, and (iii) fraction of single-positive states with an increase in in-degree of A and B nodes of a TT. (B) Plots showing the dependence of change in the distribution of the metrics: (i) f1/f2, (ii) CC AB and (iii) BC A with changing in-degree of a TT motif.*

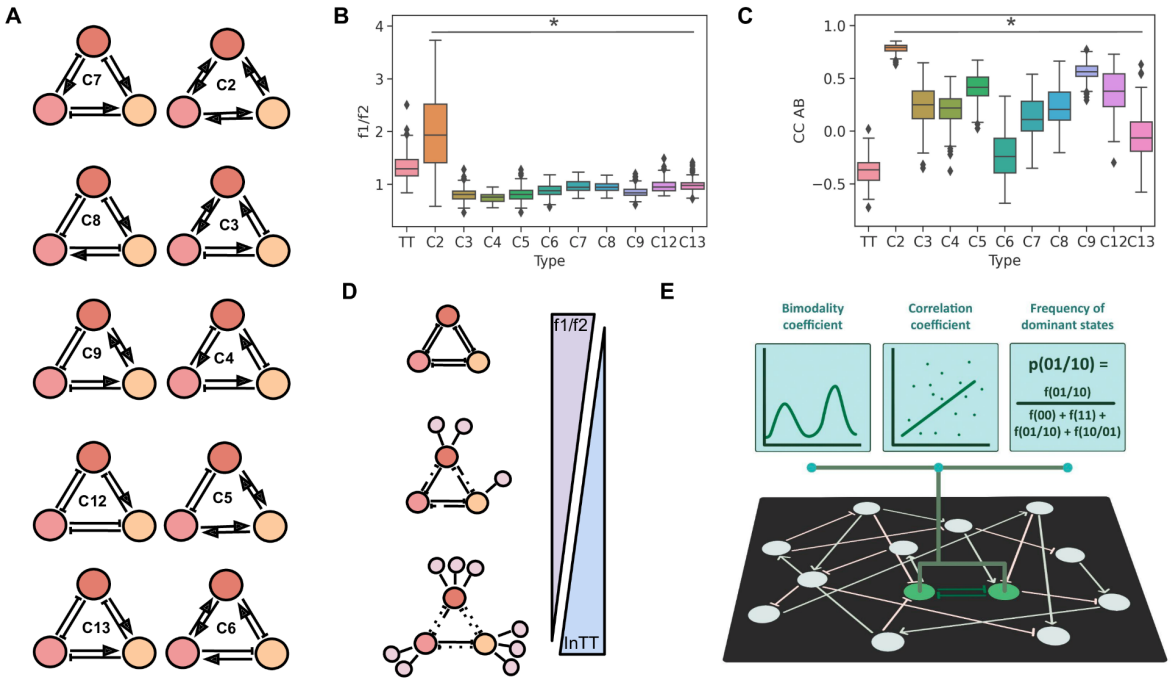

Figure 7: Above-mentioned dynamical metrics for other three-node networks (triads) *(A) Triads generated where one or more negative edges between a pair of nodes in a TT is replaced by activating links. (B) Comparison of the distribution of f1/f2 ratio of an embedded TT against the triads shown in A when embedded. (C) Comparison of the distribution of CC AB values of an embedded TT against the other triads shown in A when embedded. Bar with * implies a significantly different distribution of the corresponding triad than that of a TT. (D) Summary of the change in dynamics of TT with respect to changes in its local density. (E) Summary of the effect of a larger network on the dynamics of an embedded mutually repressing motif.*

Because the CC and BC metrics of the TT showed no significant variations upon being embedded in larger networks, we investigated other triads where one or more inhibitory links in a TT were replaced by the activatory ones (**Fig 7A**). These triads were then embedded into 100 randomly generated networks of mean connectivity of E:2N and network order of 20N, and compared distributions of correlation coefficients, bimodality coefficients, and f1/f2 values were then compared to those of a TT embedded in 100 random networks of the same mean connectivity and network order. We found that all the three metrics for the case of TT were significantly different from those of the other triads (**Fig 7B-C, S7**). In a few cases, some metrics did show similar distributions to that of a TT (for instance, BC B for network C5), but the other metrics of those triads had significantly different distributions from those of TT. Thus, although CC and BC metrics do not show statistically significant variations in the case of an embedded TT, the distributions themselves are unique properties of the TT motif and not due to a complete breakdown of the TT dynamics. In conclusion, for a TT, the f1/f2 ratio was found to be a good metric to gauge the dynamic resilience of the motif when embedded in larger networks. It was also observed that similar to a TS, as the in-degree of a TT increased, the probability of a TT maintaining its ideal properties decreased (**Fig 7D**), independent of the sign distribution of the incoming edges on a node in a TS or TT (**Fig S8**).

Discussion

Investigating the modularity of biological networks has been an active area of research. Modularity has been loosely defined as corresponding to a highly interconnected set of nodes such that density of connections within that module is significantly higher than that of density of connections of this module with other modules [24]. Thus, modularity has been mostly studied from a network topology perspective, rather than a functional one. Similarly, the concept of network motifs - recurring sets of regulatory interactions that appear more frequently than expected in a given network - also highlights network sub-structures based on their topology [25]. While the dynamics of such motifs has received extensive attention [25]; how insular or intact the dynamics of these network motifs are when embedded in a large network, remains largely underexplored.

Here, we investigated how 'modular' the behavior of a TS or TT is when embedded in random large networks of varying sizes and densities. For both these motifs, we observed that an increase in local density around the motifs (i.e., number of incoming edges) was capable of changing their dynamical behavior rather than any global topological properties associated with the large network (**Fig 7E**). Although we witnessed how increasing mean connectivity changed the distributions of the metrics we have used to characterize the dynamical behavior of TS or TT, further analysis revealed that it was the increasing in-degree of the motif (reflected, in part, by mean connectivity as well) that was driving this change in motif behavior. For the TS, it was found that all the three metrics, BC, CC, and the fraction of 01 and 10 states were suitable to gauge its change in dynamics. On the other hand, for a TT motif, CC and BC metrics were not suitable to gauge its behavior, as they did not show extensive variation upon being embedded in large networks. This difference could be because CC, being a pairwise metric, is not optimal to capture the variations in the steady state values of all three nodes. Similarly, for BC, because the steady state distribution of a node in TT does not have a very high value of BC even in isolation, using the variations in BC of a node was not suitable to understand the changes in the properties of TT embedded in larger networks. Hence, only the fraction of single-positive states was found to be adequate to observe the variations in the dynamics of an embedded TT. We found that as the in-degree of an embedded TT increased, the single-positive states and double-positive states were almost comparable in frequency. This observation suggests that in the case of CD4+ T-helper cell differentiation - a case study of TT dynamics - the "double positive" cell-states (hybrid Th1/Th2, Th1/Th17, and Th2/ Th17 phenotypes) seen experimentally [26,27] could exist due to the TT between GATA3, RORγT and T-bet being driven by various other stimuli that impinge on these nodes via activation or repression. Therefore, besides self-activation [2,28], embedding in large networks can be an additional way to enrich such 'hybrid' states, as being increasingly reported in various biological systems [29–31], both in cases of TS and TT.

Many studies have been carried out to study the bistable dynamics of a TS motif. It has been shown that the nodes of a TS motif show a bimodal distribution [32–34] of their steady states and this behavior is dependent upon a balance between the kinetic parameters of the two nodes [14,35]. An asymmetry in these parameters such that effective inhibition of one node on another is much stronger than the other way around can lead to a one-way cause-effect relation between the two TS nodes, with only one node showing bimodality while the other node showing unimodal behavior [14,19,32,35]. Here, we show that in addition to kinetic parameters, the relative in-degrees of

the nodes, when embedded in a large network, can also contribute to this behavior. A skew in the in-degrees led to a divergence from the bistable behavior of a TS, much like the skew in the kinetic parameters. Thus, our analysis uncovers an important design principle of gene regulatory networks (GRNs) that in order to maintain bistable features for a TS, the in-degree, which represents the number of regulators acting at a given point of time, should be minimal. This observation is reminiscent of previous studies demonstrating that in an *E. coli* transcriptional network, no transcription factor had an in-degree or out-degree greater than two, and this feature played a key role in enabling the robustness of the network [36]. Thus, the in-degrees we report here for robustness are in good agreement with those seen in networks for GRNs of various organisms [36]. Similarly, another recent study associated dynamical robustness of networks with their low in-degree variance [37]. Together, these results can help us design optimal strategies to design and integrate synthetic circuits into the GRN of a cell rather than developing stand-alone modules that work in isolation.

Often, various cellular processes are viewed as generally being robust to noise and perturbations, but in diseased states such as cancer, due to a change in network topology, protein production rates or short decay times can potentially perturb these steady states and encourage the progression of disease into states which are more robust and hard to reverse [35,38–40]. Our results shown here can hence help us develop better algorithms to identify potential drug targets and treatment strategies to escape these diseased states [34,41]. The ability to understand the behavior of these multistable states and modulate them can help us potentially traverse through the trajectories set by cellular differentiation branching tree to reprogram cells into the desired fate [4,34,42].

Materials and Methods

Random Network Generation

A total of 12 types of randomized networks were generated for each motif. These network types belong to four network orders with the number of nodes (N) equal to 5, 10, 15, and 20 and three mean connectivity E:2N, E:4N and E:6N with edge to node ratios E:xN, where the number of edges, E, is x times the number of nodes N. One hundred random networks were generated for each class of network. Thus, 1200 networks were simulated for each motif (TS, TT) to characterize its properties. This analysis was repeated in triplicate for statistical tests/comparisons.

A custom python3 script was written to generate random networks of a given size and mean connectivity by creating square null matrices (of order N+y, where N is the order of the network and y the number of nodes in the embedded motif) and later populating it with the motif edges, i.e. for a TS embedded in a 5N, E:2N networks, there were total 7 (= 5 + 2) nodes and 12 (=5*2 + 2) edges. Random edges were then added depending on the mean connectivity of the network. It was also ensured that no self-activating links were formed in any of the network nodes. Generated networks were checked for their connections using the *isconnected* function of the

networkx library of python3. When duplicates were found, they were replaced by newly generated networks to avoid any skew in data due to the repetition of network topologies.

Random Circuit Perturbation (RACIPE)

Random Circuit Perturbation (RACIPE) formalism generates a system of ordinary differential equations (ODEs) for a given network topology and simulates the ODEs by pooling parameters from a randomized predetermined range to identify dynamical properties of a network topology[22].

For a node T in the network having P_i activating and N_i inhibiting nodes with incoming edges, the ODE generated by RACIPE to represent the node T will be given by:

$$\frac{dT}{dt} = G_T * \prod_i \frac{H^S\left(P_i, P_{i\ T}^0, n_{P_iT}, \lambda_{P_iT}\right)}{\lambda_{P_iT}} * \prod_i H^s\left(N_j, N_{j\ T}^0, n_{N_jT}, \lambda_{N_jT}\right) - (k_T * T)$$

where the terms T, P_i and N_i are the concentrations of the nodes as time t, n is the Hill coefficient showing the influence of P_i or N_i on T, λ is the fold change in expression caused by node P_i or N_i upon acting on node T, P_i^0 or N_i^0 are the threshold values of Hill function, G_T is the production rate, and k_T is the degradation rate of the node T.

H^S represents the shifted Hill equation and is defined by:

$$H^S\left(B, B_A^0, n_{BA}, \lambda_{BA}\right) = \frac{B_A^{0\ n_{BA}}}{B_A^{0\ n_{BA}} + B^{n_{BA}}} + \lambda_{BA} * \frac{B^{n_{BA}}}{B_A^{0\ n_{BA}} + B^{n_{BA}}}$$

For a particular topology file, RACIPE generates multiple randomized parameter sets and simulates them over multiple initial conditions to identify the steady state levels of the nodes. The parameters are randomized by sampling from their respective pre-defined ranges given below:

Table 1: Ranges of randomized parameters in RACIPE

Parameter	Minimum Value	Maximum Value
Production Rate (G)	1	100
Degradation Rate (k)	0.1	1
Inhibition Fold Change (λ^-)	0.01	1
Activating Fold Change (λ^+)	1	100
Hill's Coefficient (n)	1	6
Threshold	Half-Functional Rule: Sets the threshold depending on the in-degree of the node.	

Simulations for all the networks were done in triplicates, with 10000 parameter sets per replicate and 100 initial conditions for each parameter set.

Normalizing the steady state values

The following formula first normalized the steady state values obtained from RACIPE simulations for each network:

$$S_i^N = S_i / (G_i/k_i)$$

Here, S_i^N is the normalized steady state value of the ith node; Si is the steady state value of ith node given in RACIPE output; G_i is the production rate, and k_i is the degradation rate parameter value of the ith node. The normalized values were then converted to z-scores using the *zscore* function of the SciPy library of python3. If a node showed a z-score above zero, it was considered to be showing higher expression and was considered to be in "ON" state represented by 1 and if it showed a z-score below zero, it was inferred to be in "OFF" state represented by 0. This criteria was then used to convert the steady state values into a string of zeros and ones representing the binarized steady state shown for a particular parameter set.

Sarle's Bimodality Coefficient (BC)

Sarle's bimodality coefficient (BC) [23] was used to identify the nature of the distribution of nodes' steady state values as either bimodal or not. The formula used to calculate bimodality is given by:

$$BC = \frac{m_3^2 + 1}{m_4 + 3 \cdot \frac{(n-1)^2}{(n-2)(n-3)}}$$

Where m_3 is the skew of the distribution, m_4 is the excess kurtosis, and the sample size is denoted by *n*. BC varies from 0 to 1, with values above 0.55 (i.e. 5/9) representing bimodality in the distribution and those below this threshold indicating a unimodal distribution [23]. Skew and excess kurtosis were calculated using functions of the SciPy library of python3. The values were then plugged into the above equation to obtain the bimodality coefficients.

Correlation Coefficient (CC)

Correlation coefficients (CC) were calculated using the functions *spearmanr* and *pearsonr* of the SciPy library of python3 wherever they were applicable.

Data and Code Availability

The Github repository for RCAIPE-1.0 [22] can be found at https://github.com/simonhb1990/RACIPE-1.0.

Custom python3 (python3 version 3.8.10) scripts were also written to analyze the RACIPE output files further and can be found at https://github.com/MoltenEcdysone09/Modularity.

Author contributions MKJ conceived and supervised research, PH performed research. All authors contributed to data analysis and writing of the manuscript.

Funding This work was supported by Ramanujan Fellowship (SB/S2/RJN-049/2018) by the Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India, awarded to MKJ.

References

(1) Guantes, R.; Poyatos, J. F. Multistable Decision Switches for Flexible Control of Epigenetic Differentiation. *PLOS Comput. Biol.* **2008**, *4* (11), e1000235. https://doi.org/10.1371/journal.pcbi.1000235.

(2) Zhou, J. X.; Huang, S. Understanding Gene Circuits at Cell-Fate Branch Points for Rational Cell Reprogramming. *Trends Genet.* **2011**, *27* (2), 55–62. https://doi.org/10.1016/j.tig.2010.11.002.

(3) Qian, Y.; McBride, C.; Del Vecchio, D. Programming Cells to Work for Us. *Annu. Rev. Control Robot. Auton. Syst.* **2018**, *1* (1), 411–440. https://doi.org/10.1146/annurev-control-060117-105052.

(4) Agozzino, L.; Balázsi, G.; Wang, J.; Dill, K. A. How Do Cells Adapt? Stories Told in Landscapes. *Annu. Rev. Chem. Biomol. Eng.* **2020**, *11*, 155–182. https://doi.org/10.1146/annurev-chembioeng-011720-103410.

(5) Santos-Moreno, J.; Tasiudi, E.; Stelling, J.; Schaerli, Y. Multistable and Dynamic CRISPRi-Based Synthetic Circuits. *Nat. Commun.* **2020**, *11* (1), 2746. https://doi.org/10.1038/s41467-020-16574-1.

(6) Chang, H. H.; Oh, P. Y.; Ingber, D. E.; Huang, S. Multistable and Multistep Dynamics in Neutrophil Differentiation. *BMC Cell Biol.* **2006**, *7* (1), 11. https://doi.org/10.1186/1471-2121-7-11.

(7) Ferrell, J. E. Bistability, Bifurcations, and Waddington's Epigenetic Landscape. *Curr. Biol.* **2012**, *22* (11), R458–R466. https://doi.org/10.1016/j.cub.2012.03.045.

(8) Gardner, T. S.; Cantor, C. R.; Collins, J. J. Construction of a Genetic Toggle Switch in Escherichia Coli. *Nature* **2000**, *403* (6767), 339–342. https://doi.org/10.1038/35002131.

(9) Celià-Terrassa, T.; Bastian, C.; Liu, D. D.; Ell, B.; Aiello, N. M.; Wei, Y.; Zamalloa, J.; Blanco, A. M.; Hang, X.; Kunisky, D.; Li, W.; Williams, E. D.; Rabitz, H.; Kang, Y. Hysteresis Control of Epithelial-Mesenchymal Transition Dynamics Conveys a Distinct Program with Enhanced Metastatic Ability. *Nat. Commun.* **2018**, *9* (1), 5005. https://doi.org/10.1038/s41467-018-07538-7.

(10) Huang, S.; Guo, Y.-P.; May, G.; Enver, T. Bifurcation Dynamics in Lineage-Commitment in Bipotent Progenitor Cells. *Dev. Biol.* **2007**, *305* (2), 695–713. https://doi.org/10.1016/j.ydbio.2007.02.036.

(11) Atkinson, M. R.; Savageau, M. A.; Myers, J. T.; Ninfa, A. J. Development of Genetic Circuitry Exhibiting Toggle Switch or Oscillatory Behavior in Escherichia Coli. *Cell* **2003**, *113* (5), 597–607. https://doi.org/10.1016/S0092-8674(03)00346-5.

(12) Oyarzún, D. A.; Chaves, M. Design of a Bistable Switch to Control Cellular Uptake. *J. R. Soc. Interface* **2015**, *12* (113), 20150618. https://doi.org/10.1098/rsif.2015.0618.

(13) Duddu, A. S.; Sahoo, S.; Hati, S.; Jhunjhunwala, S.; Jolly, M. K. Multi-Stability in Cellular Differentiation Enabled by a Network of Three Mutually Repressing Master Regulators. *J. R. Soc. Interface* **2020**, *17* (170), 20200631. https://doi.org/10.1098/rsif.2020.0631.

(14) Leon, M.; Woods, M. L.; Fedorec, A. J. H.; Barnes, C. P. A Computational Method for the Investigation of Multistable Systems and Its Application to Genetic Switches. *BMC Syst. Biol.* **2016**, *10* (1), 130. https://doi.org/10.1186/s12918-016-0375-z.

(15) Faucon, P. C.; Pardee, K.; Kumar, R. M.; Li, H.; Loh, Y.-H.; Wang, X. Gene Networks of Fully Connected Triads with Complete Auto-Activation Enable Multistability and Stepwise Stochastic Transitions. *PloS One* **2014**, *9* (7), e102873. https://doi.org/10.1371/journal.pone.0102873.

(16) Yang, L.; Sun, W.; Turcotte, M. Coexistence of Hopf-Born Rotation and Heteroclinic Cycling in a Time-Delayed Three-Gene Auto-Regulated and Mutually-Repressed Core

Genetic Regulation Network. *J. Theor. Biol.* **2021**, *527*, 110813. https://doi.org/10.1016/j.jtbi.2021.110813.
(17) O'Shea, J. J.; Paul, W. E. Mechanisms Underlying Lineage Commitment and Plasticity of Helper CD4+ T Cells. *Science* **2010**. https://doi.org/10.1126/science.1178334.
(18) Fang, D.; Zhu, J. Dynamic Balance between Master Transcription Factors Determines the Fates and Functions of CD4 T Cell and Innate Lymphoid Cell Subsets. *J. Exp. Med.* **2017**, *214* (7), 1861–1876. https://doi.org/10.1084/jem.20170494.
(19) Strasser, M.; Theis, F. J.; Marr, C. Stability and Multiattractor Dynamics of a Toggle Switch Based on a Two-Stage Model of Stochastic Gene Expression. *Biophys. J.* **2012**, *102* (1), 19–29. https://doi.org/10.1016/j.bpj.2011.11.4000.
(20) Jaruszewicz, J.; Lipniacki, T. Toggle Switch: Noise Determines the Winning Gene. *Phys. Biol.* **2013**, *10* (3), 035007. https://doi.org/10.1088/1478-3975/10/3/035007.
(21) Ribeiro, A. S. Dynamics and Evolution of Stochastic Bistable Gene Networks with Sensing in Fluctuating Environments. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **2008**, *78* (6 Pt 1), 061902. https://doi.org/10.1103/PhysRevE.78.061902.
(22) Huang, B.; Lu, M.; Jia, D.; Ben-Jacob, E.; Levine, H.; Onuchic, J. N. Interrogating the Topological Robustness of Gene Regulatory Circuits by Randomization. *PLOS Comput. Biol.* **2017**, *13* (3), e1005456. https://doi.org/10.1371/journal.pcbi.1005456.
(23) Pfister, R.; Schwarz, K.; Janczyk, M.; Dale, R.; Freeman, J. Good Things Peak in Pairs: A Note on the Bimodality Coefficient. *Front. Psychol.* **2013**, *4*, 700. https://doi.org/10.3389/fpsyg.2013.00700.
(24) Alcalá-Corona, S. A.; Sandoval-Motta, S.; Espinal-Enríquez, J.; Hernández-Lemus, E. Modularity in Biological Networks. *Front. Genet.* **2021**, *12*, 701331. https://doi.org/10.3389/fgene.2021.701331.
(25) Alon, U. Network Motifs: Theory and Experimental Approaches. *Nat. Rev. Genet.* **2007**, *8* (6), 450–461. https://doi.org/10.1038/nrg2102.
(26) Peine, M.; Rausch, S.; Helmstetter, C.; Fröhlich, A.; Hegazy, A. N.; Kühl, A. A.; Grevelding, C. G.; Höfer, T.; Hartmann, S.; Löhning, M. Stable T-Bet(+)GATA-3(+) Th1/Th2 Hybrid Cells Arise in Vivo, Can Develop Directly from Naive Precursors, and Limit Immunopathologic Inflammation. *PLoS Biol.* **2013**, *11* (8), e1001633. https://doi.org/10.1371/journal.pbio.1001633.
(27) Chatterjee, S.; Daenthanasanmak, A.; Chakraborty, P.; Wyatt, M. W.; Dhar, P.; Selvam, S. P.; Fu, J.; Zhang, J.; Nguyen, H.; Kang, I.; Toth, K.; Al-Homrani, M.; Husain, M.; Beeson, G.; Ball, L.; Helke, K.; Husain, S.; Garrett-Mayer, E.; Hardiman, G.; Mehrotra, M.; Nishimura, M. I.; Beeson, C. C.; Bupp, M. G.; Wu, J.; Ogretmen, B.; Paulos, C. M.; Rathmell, J.; Yu, X.-Z.; Mehrotra, S. CD38-NAD+Axis Regulates Immunotherapeutic Anti-Tumor T Cell Response. *Cell Metab.* **2018**, *27* (1), 85-100.e8. https://doi.org/10.1016/j.cmet.2017.10.006.
(28) Jia, D.; Jolly, M. K.; Harrison, W.; Boareto, M.; Ben-Jacob, E.; Levine, H. Operating Principles of Tristable Circuits Regulating Cellular Differentiation. *Phys. Biol.* **2017**, *14* (3), 035007. https://doi.org/10.1088/1478-3975/aa6f90.
(29) Jolly, M. K.; Boareto, M.; Lu, M.; Onuchic, J. N.; Clementi, C.; Ben-Jacob, E. Operating Principles of Notch–Delta–Jagged Module of Cell–Cell Communication. *New J. Phys.* **2015**, *17* (5), 055021. https://doi.org/10.1088/1367-2630/17/5/055021.
(30) Calcium Signaling Induces a Partial EMT. *EMBO Rep.* **2021**, *22* (9), e51872. https://doi.org/10.15252/embr.202051872.
(31) MacLean, A. L.; Hong, T.; Nie, Q. Exploring Intermediate Cell States through the Lens of Single Cells. *Curr. Opin. Syst. Biol.* **2018**, *9*, 32–41. https://doi.org/10.1016/j.coisb.2018.02.009.
(32) Yong, C.; Gyorgy, A. Stability and Robustness of Unbalanced Genetic Toggle Switches in the Presence of Scarce Resources. *Life* **2021**, *11* (4), 271.

https://doi.org/10.3390/life11040271.
(33) Shu, C.-C.; Chatterjee, A.; Dunny, G.; Hu, W.-S.; Ramkrishna, D. Bistability versus Bimodal Distributions in Gene Regulatory Processes from Population Balance. *PLOS Comput. Biol.* **2011**, *7* (8), e1002140. https://doi.org/10.1371/journal.pcbi.1002140.
(34) Shu, C.-C.; Yeh, C.-C.; Jhang, W.-S.; Lo, S.-C. Driving Cells to the Desired State in a Bimodal Distribution through Manipulation of Internal Noise with Biologically Practicable Approaches. *PLoS ONE* **2016**, *11* (12), e0167563. https://doi.org/10.1371/journal.pone.0167563.
(35) Kim, K.-Y.; Wang, J. Potential Energy Landscape and Robustness of a Gene Regulatory Network: Toggle Switch. *PLOS Comput. Biol.* **2007**, *3* (3), e60. https://doi.org/10.1371/journal.pcbi.0030060.
(36) Albergante, L.; Blow, J. J.; Newman, T. J. Buffered Qualitative Stability Explains the Robustness and Evolvability of Transcriptional Networks. *eLife* **2014**, *3*, e02863. https://doi.org/10.7554/eLife.02863.
(37) Ichinose, N.; Kawashima, T.; Yada, T.; Wada, H. Dynamical Robustness and Its Structural Dependence in Biological Networks. *J. Theor. Biol.* **2021**, *526*, 110808. https://doi.org/10.1016/j.jtbi.2021.110808.
(38) Perez-Carrasco, R.; Guerrero, P.; Briscoe, J.; Page, K. M. Intrinsic Noise Profoundly Alters the Dynamics and Steady State of Morphogen-Controlled Bistable Genetic Switches. *PLOS Comput. Biol.* **2016**, *12* (10), e1005154. https://doi.org/10.1371/journal.pcbi.1005154.
(39) Ciliberti, S.; Martin, O. C.; Wagner, A. Robustness Can Evolve Gradually in Complex Regulatory Gene Networks with Varying Topology. *PLoS Comput. Biol.* **2007**, *3* (2), e15. https://doi.org/10.1371/journal.pcbi.0030015.
(40) Chaves, M.; Albert, R.; Sontag, E. D. Robustness and Fragility of Boolean Models for Genetic Regulatory Networks. *J. Theor. Biol.* **2005**, *235* (3), 431–449. https://doi.org/10.1016/j.jtbi.2005.01.023.
(41) Wu, M.; Liu, L.; Chan, C. Identification of Novel Targets for Breast Cancer by Exploring Gene Switches on a Genome Scale. *BMC Genomics* **2011**, *12* (1), 547. https://doi.org/10.1186/1471-2164-12-547.
(42) Heiderscheit, E. A.; Eguchi, A.; Spurgat, M. C.; Ansari, A. Z. Reprogramming Cell Fate with Artificial Transcription Factors. *FEBS Lett.* **2018**, *592* (6), 888–900. https://doi.org/10.1002/1873-3468.12993.