Stability Depends on Positive Autoregulation in Boolean Gene Regulatory Networks

Ricardo Pinho[1,2]*, Victor Garcia[3], Manuel Irimia[4]*, Marcus W. Feldman[1]

1 Department of Biology, Stanford University, Stanford, California, United States of America, 2 PhD Program in Computational Biology, Instituto Gulbenkian de Ciência, Oeiras, Portugal, 3 Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland, 4 EMBL/CRG Research Unit in Systems Biology, Centre for Genomic Regulation (CRG), Barcelona, Spain

Abstract

Network motifs have been identified as building blocks of regulatory networks, including gene regulatory networks (GRNs). The most basic motif, autoregulation, has been associated with bistability (when positive) and with homeostasis and robustness to noise (when negative), but its general importance in network behavior is poorly understood. Moreover, how specific autoregulatory motifs are selected during evolution and how this relates to robustness is largely unknown. Here, we used a class of GRN models, Boolean networks, to investigate the relationship between autoregulation and network stability and robustness under various conditions. We ran evolutionary simulation experiments for different models of selection, including mutation and recombination. Each generation simulated the development of a population of organisms modeled by GRNs. We found that stability and robustness positively correlate with autoregulation; in all investigated scenarios, stable networks had mostly positive autoregulation. Assuming biological networks correspond to stable networks, these results suggest that biological networks should often be dominated by positive autoregulatory loops. This seems to be the case for most studied eukaryotic transcription factor networks, including those in yeast, flies and mammals.

Introduction

Gene regulatory networks (GRNs) are believed to play a central role in organismal development and evolution [1–3]. Recent theoretical and experimental studies have revealed that GRNs have many interesting quantitative and qualitative features, including scale-free structure [4], recurring motifs [5], robustness [6], and evolvability [7]. Here we focus on a very specific and common network motif, autoregulation [8], and its contribution to stability and mutational robustness [9].

A direct autoregulation motif in transcriptional GRNs consists of a regulator that binds to the promoter region of its own gene, thus regulating its own transcription. It constitutes the simplest case of a feedback mechanism. Two thirds of *E. coli*'s transcriptional factors (TFs) are believed to be autoregulated [10]. The fraction of autoregulated TFs is lower for yeast (10% [11]), but extensive autoregulation at the post-transcriptional level has been suggested [12]. Two rules relating the presence of feedback loops in GRNs to their dynamical properties have been proposed [13]: (i) a necessary condition for multistability (i.e., the existence of several stable fixed points in the dynamics) is the existence of a positive circuit in the regulatory network (the sign of a circuit being

defined as the product of the signs of its edges); and (ii) a necessary condition for the existence of an attractive cycle in the dynamics is the existence of a negative circuit.

These two types of dynamical properties have been associated with important biological phenomena: cell differentiation and stochastic switching in the first case [14], homeostasis [9] and periodic behaviors (e.g., cell cycle [15] and circadian rhythms [16]) in the second. Although these conditions are necessary, they are often not sufficient to define network dynamics, which can depend on other details of the GRN model [13]. For example, negative autoregulation (NAR), the shortest negative circuit possible, has been traditionally associated with robustness of gene expression to noise [9]. However, if the NAR feedback contains a long delay, noise may be amplified [17]. Moreover, both positive and negative feedback circuits are usually embedded in larger networks, and the relative contributions of multiple positive and negative feedback loops to the dynamics of a whole network are largely unknown [13,14,18–20].

Here, we investigate the relationship between the sign of autoregulation and the stability and mutational robustness of genetic networks. We study this in the context of a widely used gene network model [21–23], related to the modeling framework

Author Summary

Multicellular organisms show an incredible diversity of cell types in their different tissues. Functional classes of cells can be attributed to the activation and repression of genes, which enable each cell type to support different functions within the organism. These patterns of activity have been studied by means of gene regulatory networks (GRNs). How these gene networks generate stable phenotypic states is thought to underlie the development and evolution of organisms. The pathways to these states are influenced by the autoregulatory properties of these networks. The stability and robustness of gene networks are used to investigate how such states are maintained. This study sheds light on how these properties relate to one another. By simulating the evolution of these networks, we show that genes depend on positive self-regulation to remain stable and robust when faced with random mutations or environmental perturbations. Assuming biological networks correspond to stable networks, our results suggest that biological networks should often be dominated by positive autoregulatory loops. This seems to be the case for most studied eukaryotic transcription factor networks, including those in yeast, flies and mammals.

of Boolean networks [24]. We find that stability and robustness are highly correlated with the sign of autoregulation, and that selection for stability leads to positive autoregulation. Despite these positive associations, we show that selection does not maximize robustness and that it is possible to engineer networks with higher robustness by manipulating their diagonal and off-diagonal elements. We also show that autoregulation is conserved over time and that evolved networks are a special subset of stable networks (networks that show fixed point dynamics) with high robustness. Finally, we discuss some implications for biological systems and compare our results with biological networks of different organisms.

Methods

Developmental model

To study how stability, robustness and autoregulation change during evolution, we use a standard model for GRN. In one generation, we assume that that the *phenotype* of an organism $S(t)$ develops over time t, starting from an initial phenotype $S(t=0)$, under the influence of a gene-interaction network W. In general, phenotypes are thought of as expression levels of the genes of the organism at time t. Thus, they are vectors of dimension N, $S(t) = (s_1(t), \ldots, s_N(t))$, with binary entry values $s_i \in \{-1, 1\}$, where N is the number of genes of the organism.

Phenotypes $S(t)$ change by the action of a gene-interaction network that drives their *development*, and is represented by an $N \times N$ matrix, W, whose elements, w_{ij}, denote the effect on gene i of the product of gene j. These interaction weights w_{ij} are nonzero and binary, $w_{ij} \in \{-1, 1\}$. Thus, all genes either repress or activate each other's expression.

In this study, we assume that size of the gene interaction network is $N = 10$ genes. The matrix W is not necessarily symmetric. Diagonal elements, w_{ii}, represent autoregulation, i.e., the action of the ith gene on itself.

Each network W determines the dynamics of the phenotype $S(t)$ in a series of *development steps*. The repeated application of such development steps on a phenotype results in deterministic, discrete-time dynamics of $S(t)$, modeled by the set of nonlinear coupled difference equations:

$$s_i(t+1) = sgn\left(\sum_{j=1}^{N} w_{ij}s_j(t)\right), \qquad (1)$$

where $sgn(0) = 1$. This spin glass or neural network-type model [22] represents a subclass of Random Boolean Networks [24] known as Random Threshold Networks [25].

When simulating development, the network is updated synchronously, that is, only values of s_i from time step t are used for the calculation of s_i $(t+1)$ (see [26–28] for asynchronous updates.) We refer to Equation (1) as the *development process* (see [23,29] for model illustration, biological motivation and assumptions).

The development process can be extended to include *sparse networks G*. Sparse networks are used to model gene interactions in which only a fraction of the genes repress or activate a fraction of all the other genes, in contrast to fully connected networks W, where all genes have some effect on all other genes.

Let G denote an interaction network represented by a $N \times N$, $N = 10$ square matrix whose entries, g_{ij}, take the values of $\{-1,0,1\}$. The parameter c, the *density* of the network, determines the proportion of non-zero matrix elements. When simulating sparse networks (see Results), we chose $c = 0.2$ (due to similarity to the biological networks in Table 1) and a regular, directed graph topology, where all genes have degree 2. This means that all genes in a network are regulated by two genes and also regulate two other genes.

Examining model behavior. Starting from an initial gene expression state the system described in Equation (1) will eventually reach an *attractor*. Such an attractor may be a fixed point or a limit cycle. In a biological context, a fixed point can be interpreted as one mature phenotype of the organism after the completion of development.

Simulation experiment setups. To investigate how specific network features change within populations under development as well as under evolution, we devised two main simulation experiment setups.

Each organism is represented by a network W and an initial state $S(0)$. In this study we limit ourselves to two experimental setups: pairs of randomly chosen networks and random initial conditions (*RNRC setup*), and n randomly chosen networks that each act on one single randomly chosen initial condition (termed *RNIC setup*). We use RNRC for populations that don't evolve, and RNIC for populations subject to evolution (as explained below). The *random initial conditions* for the simulation experiments are generated by sampling one phenotype with uniform probability from the entire phenotype space ($s_i(0) = 1, -1$ with probability 0.5).

Random and stable networks

To generate a *random* network W_r the matrix-elements w_{ij} are sampled from $\{-1,1\}$ with equal probability (0.5 per element and entry). Additionally, we can generate *stable networks* W_s with a pre-selection procedure. In this procedure, a random network W_r and random initial state pair are first generated. This pair undergoes the development process. If no fixed point is attained, a new pair is sampled and developed. This step is repeated until some $(W, S(0))$- pair generates a fixed point. The final network W_s, is a stable network. This notion of stability refers to an *individual level stability*, which differs from a notion of *population level stability* that will be introduced below.

Table 1. Most studied eukaryotic transcription factor networks (including yeast, flies and mammals) show values of p ranging from 0.76 to 1.

Species	TF System	p	# autoregulated	# TFs	Reference
Mammals	core pluripotency network	1	5	5	[42]
Drosophila	gap genes	1	4	4–7	[48,49]
Drosophila	segment polarity	1	5	10	[37]
Drosophila	circadian clock	1	5	6	[50]
Arabidopsis	circadian clock	1	4	4	[51]
Arabidopsis	flower morphogenesis	1	2	10	[38]
Mouse	blood stem cells	1	7	11	[52]
Human	genome-wide	0.76	21	301	[46]
Drosophila melanogaster	genome-wide	0.78	14	87	[46]
Saccharomyces cerevisiae	genome-wide	0.75	12	169	[46]
E. coli	genome-wide	0.26	109	182	[45]
Sea urchin	endomesoderm development	0.25	16	50	[53]

Description of different regulatory networks of different organisms in terms of size, ubiquity and sign of autoregulation.
doi:10.1371/journal.pcbi.1003916.t001

In this study, we refer to stability as the property of a network, while strictly speaking, it is a property of a W, $S(0)$ pair. However, we have previously shown [30] that the network is by far the most important determinant of stability. If a network is stable/unstable with a random initial state, it most likely remains stable/unstable with any other initial state. For this reason, we classify networks as stable or unstable, even if we just solve Equation (1) for one possible initial state.

Evolved and non-evolved networks

Two types of simulation experiments are our primary focus in this study. First, experiments in which a population of organisms undergoes the development process only, which we refer to as *non-evolved*. Secondly, experiments with multiple generations (*evolved*), where after each development process (one generation) the composition of the population of organisms is additionally altered by evolutionary mechanisms. The development process is completed after all organisms have reached some stage of development: either a fixed point or a cycle. We implemented standard evolutionary mechanisms, such as selection, mutation and recombination. After these evolutionary forces have acted on the population, a new development process starts in the next generation with identical initial phenotypes for each organism.

In this study, we set the population size to $n = 500$ across all experiments (unless otherwise noted). This population size remains constant during evolutionary simulations (Wright-Fisher model with sampling with replacement).

Selection

To study how selection affects evolving populations, we implemented different *types of selection* or *selection models*. The mutation and recombination mechanisms applied were the same for all evolved populations.

Selection mechanisms modify the number of copies of one specific network within the population depending on the fitness of the phenotype that specific network has generated through development. In a selection mechanism, one phenotypic state can be marked as the optimal state, with the highest possible fitness. If such an optimal state, $S^{opt}(\infty)$ is specified, the fitness of a network with attractor $S(\infty)$ is given by:

$$\text{fitness}(S) = \exp\left(-\frac{d(S(\infty), S^{opt}(\infty))^2}{\sigma}\right), \qquad (2)$$

where d is the normalized Hamming distance and $\sigma > 0$ determines selection strength [23].

Small values of σ imply strong selection against deviations from the optimal state. Large values minimize the fitness difference between phenotypes. The Hamming distance d corresponds to the number of differing expression states of individual genes between two phenotypic states [31], subsequently normalized to the interval [0,1] in this study.

Equation (2) is valid under the assumption that $S^{opt}(\infty)$ and $S(\infty)$ are attractors with identical, optimal cycle lengths l_{opt}. The cycle length of the attractor with highest fitness is denoted with l_{opt}. The fitness of attractors of length $l \neq l_{opt}$ depends on the selection model. We use attractor length and cycle size or period as synonyms.

Selection models

We implemented selection models similar to those used by other authors [23,29] and also introduced new ones. In these models, the fitness of a developed organism depends on two parameters: selection strength, σ, and optimal period, l_{opt}.

Selection model 1 (selecting for stability): $l_{opt} = 1$, fitness($l \neq l_{opt}$) = 0

$\sigma = 0.1$ 'target' model

Selects for fixed points and an optimal gene expression state. Fitness is given by Equation (2) for fixed points and is 0 for cycles.

$\sigma = \infty$ 'no target' model

Fitness is 1 for all fixed points and 0 for cycles.

Selection model 2 (selecting against stability): $l_{opt} > 1$, fitness($l \neq l_{opt}$) = 0

$\sigma = \infty$, $l_{opt} = 2,3,...,7$ (cycles)

We generalize the 'no target' model to select for cycles. Fitness is 1 for cycles of length $l = l_{opt}$ and 0 otherwise, including fixed points. We try different $l_{opt} > 1$.

Selection model 3 (neutral for stability):

$\sigma = 0.1$, fitness = max(fitness(S)) for all S in $S(\infty)$, any l (S represents any state in the attractor $S(\infty)$)

We generalize the 'target' model to not require stability. When $l = 1$, we have the 'target' model as a special case and fitness is given by Equation (2). When $l > 1$, fitness is the maximum fitness given by Equation (2) for all states in the cycle. The attractor $S(\infty)$ can be a fixed point or a cycle.

Selection model 4 (random sampling):
$\sigma = \infty$, fitness $= 1$, any l

No selection. We take this as the null model.

For each selection model, we generated between $z = 100$ and $z = 300$ independent populations (depending on the model). Specifically: $z = 200$ for the 'target' model; $z = 300$ for the 'no target' model; $z = 200$ for Selection model 2; $z = 100$ for Selection models 3 and 4. We denote such an aggregation of populations as a *set of populations* and z as its *set size*. Each evolved population has a different initial state, but all individuals within the same population have the same initial state.

Mutation

Mutations randomly change the sign of w_{ij} at a rate $\mu = 0.1$ per network per generation. All matrix entries, w_{ij}, including diagonal elements, w_{ii}, have equal probability of changing sign, namely $\mu / N^2 = 0.001$ per generation. For sparse networks we use a probability for changing sign of $\mu/(c\, N^2)$.

Recombination

To model recombination we follow the methods in [23], where full chromosome segregation (no crossover) is implemented. The two offspring of a randomly chosen pair of recombinant parents are generated by randomly taking half the rows from each parent matrix. This procedure is performed on the entire population.

Population metrics

We define a population-level *stability* (henceforth referred to as *stability* if not otherwise stated) as the fraction of networks that are stable (individual-level) in a given population [30]:

$$stability = n_f / n \qquad (3)$$

where $n_f \leq n$ is the number of times the attractor is a fixed point, and n is population size, that is, the number of network matrices. Stability takes values between 0 and 1.

Similarly, we define the *robustness* of a population as the fraction of all possible mutated networks in a population that reach the same fixed point attractor as their un-mutated originals [23], conditional on the fact that the attractor did not become a limit-cycle. Specifically, we estimate robustness by looping through the population of networks and mutating every element of each network matrix W (changing the sign of w_{ij} for binary matrices), thus generating N^2 single-mutants per network. Then, the $n \times N^2$ networks undergo the development process starting from the same initial phenotypes as their originals, and are further analyzed.

For a single network, we define individual-level viability as the fraction of single-mutants that attain a fixed point:

$$viability_{ind} = n_{fixed} / N^2 \qquad (4)$$

where $n_{fixed} < N^2$ is the number of times the N^2 single-mutants have still generated a fixed point. With this metric, we can now define individual-level robustness:

$$robustness_{ind} = n_= / n_{fixed} \qquad (5)$$

and $n_= \leq n_{fixed}$ is the number of times the same attractor state as the one attained by the un-mutated original is reached starting with identical initial conditions (i.e., the mutant has the same phenotype as its wildtype). The *population-level viability* and *robustness* measures are computed from the averages of all networks in the entire set of populations.

Both robustness and viability take values between 0 and 1. In normalizing robustness by n_{fixed} instead of N^2, we attempt to decouple the effects of stability and robustness. In the vast majority of cases, mutations that change the stability of a network do not affect its robustness score. Exceptions to this are the rare occasions when $n_{fixed} = 0$ (robustness is not defined), or when n_{fixed} is low (robustness can only take a few specific values).

In an extension of the fraction of activating connections-statistic [32], we found it useful to measure properties of diagonal and off-diagonal elements of a matrix W separately, thereby decoupling the effects of direct autoregulation and off-diagonal regulation. For a single network matrix W, we define:

$$p = N_p^+ / N$$
$$q = N_q^+ / N(N-1) \qquad (6)$$

where N^+_p and N^+_q are the number of positive diagonal and off-diagonal elements of W, respectively. Both p and q are always positive and take values between 0 and 1. We call p the *sign of autoregulation*, because autoregulation is predominantly positive when $p > 0.5$ (we call this *positive autoregulation*), and mostly negative when $p < 0.5$ (we call this *negative autoregulation*).

The metrics p and q measure direct regulatory influence. However, network dynamics can also be affected by long-range interactions. To assess the role of such long-range regulation, we introduce a metric of *indirect positive autoregulation* r, which measures the fraction of autoregulatory paths over two genes that are positive (i.e., gene A activates gene B, which activates gene A). For a single network W, we define:

$$r = N_r^+ / N(N-1) \qquad (7)$$

where N^+_r is the number of positive off-diagonal elements of WW^T (W^T is the transpose of W). Because WW^T is symmetrical, it suffices to count the fraction of positive entries in either of the triangles of the matrix.

We also define metrics to assess the population-average of gene interaction strengths:

$$o_{ij} = n^+_{ij} / n$$
$$conservation_{ij} = 1 - abs\big(o_{ij}(t_1) o_{ij}(t_2)\big), \qquad (8)$$

where n^+_{ij} is the number of positive elements in position i,j across all n networks in a population, abs is the absolute value function, and t_1 and t_2 are two different evolutionary time points. Here, o_{ij}, is referred to as *average positive ij- interaction strength*, and measures how much gene i activates gene j on average, whereas the conservation statistics between population-averaged gene interaction strengths measures how much these interaction strengths are maintained over time. In evolutionary experiments, conservation, p, q, and r are averaged over all individuals and all populations.

The code utilized in this paper can be downloaded from https://github.com/rpinho/phd.

Results

Individual- and population-level stability and autoregulation are correlated in Boolean GRNs

To study the relationship between the sign of autoregulation (p) and stability during development, we devised two experiments. First, for each $p = 0, 0.1, \ldots, 1$, a pair consisting of one random network and one random initial condition was sampled (RNRC setup; see Methods). Equation (1) was then evaluated for each pair: if the attractor was a fixed point, the network was considered stable. Instead, if the solution to Equation (1) was a limit cycle, the network was considered unstable. This process was repeated $n = 10^5$ times for each p.

Figure 1A shows that individual-level stability is strongly associated with p. Stable networks have significantly higher values of p than unstable networks (median of 0.9 compared to 0.4 for unstable networks; Mann-Whitney U p-value ~ 0, Figure 1A).

This positive association was also observed for population-level stability in a second experiment. We subdivided the networks generated in the first experiment into populations of identical p, and measured the average *population-level stability* for each p. We observed that the fraction of stable networks increases rapidly with higher values of p (Figure 1B). These results also indicate that p and stability are strongly associated.

Evolution of autoregulation when selecting for stability is non-linear

We next studied how p changes when explicitly selecting for and against individual-level stability in evolutionary simulations. To this end, we founded sets of populations with random networks

and the same initial state for each population (RNIC setup; see Methods) [23,29]. The average p was set to $p = 0.5$ at generation 0. We then evolved all populations under the six different selection models (including mutation and recombination) described in the Methods. For all of these selection models, we followed the evolution of the sign of autoregulation p over 10^6–10^7 generations (until equilibrium was attained).

Consistent with the observations for non-evolving networks, positive autoregulation is strongly favored during evolution, both under the 'target' and 'no target' models (Figure 2A). However, the evolution of p follows a complex, non-linear pattern. After a sharp initial increase over the first ~ 50 generations, p reaches its maximum when population-level stability is above 95% (stability-metric not shown in the Figure), and starts to decay slowly to a stable evolutionary equilibrium of $p \sim 0.8$ from $t \sim 10^3$ generations for both models (Figure 2A). At the peak, a fraction of up to $p \sim 0.95$, or 19/20 surviving matrices show positive autoregulation for all genes under the 'no target' selection model.

Interestingly, selecting for cycles of length $l = 2$ (i.e. against individual-level stability), has the opposite effect on evolving networks: p decreases sharply down to ~ 0.3 (Figure 2B), leading to negative autoregulation. A similar, but less pronounced pattern is observed when selecting for longer cycles with lengths $l > 2$ (Figures 2B and S2).

As expected, a neutral model with no selection for individual-level stability or a specific target produces random networks, with values of p centered on $p \sim 0.5$ (Figure 2C). However, mean values of $p < 0.5$ also evolve when not selecting for any particular attractor length, but still selecting for a specific target (Figure 2C).

Thus, the selection for stability leads to positive autoregulation.
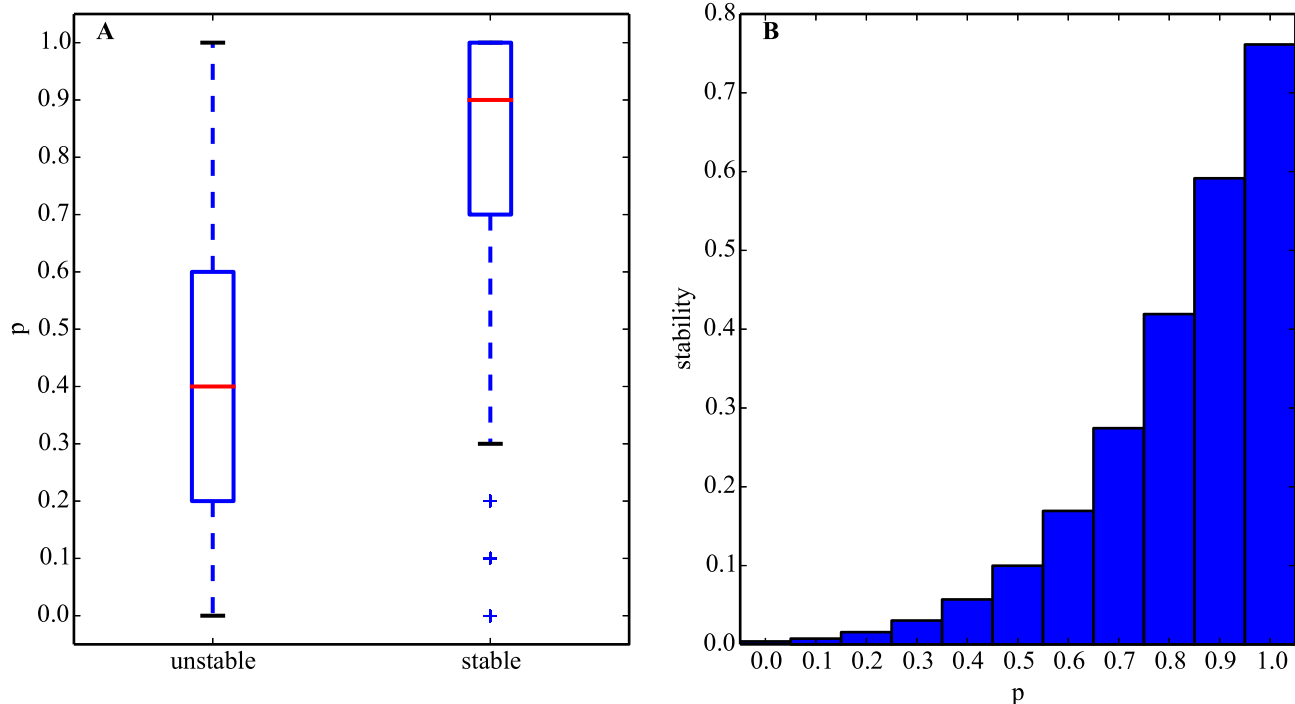
Figure 1. Positive autoregulation favors stability. A. Boxplots of the sign of autoregulation (p in Equation (6)) for two network classes: stable (fixed points) and unstable (cycles). Whiskers are 1.5 times the inter-quartile range (the difference between the first and third quartiles). All outliers are represented. B. Histogram of the values of p for stable matrices, i.e., stability (Equation (3)) as a function of p. The dotted line is an exponential function of p. Networks are random and non-evolved. Panels A and B represent the same data in different form. Equation (1) is solved $n = 10^5$ times for each $p = 0, 0.1, \ldots, 1$ (11 bins).
doi:10.1371/journal.pcbi.1003916.g001
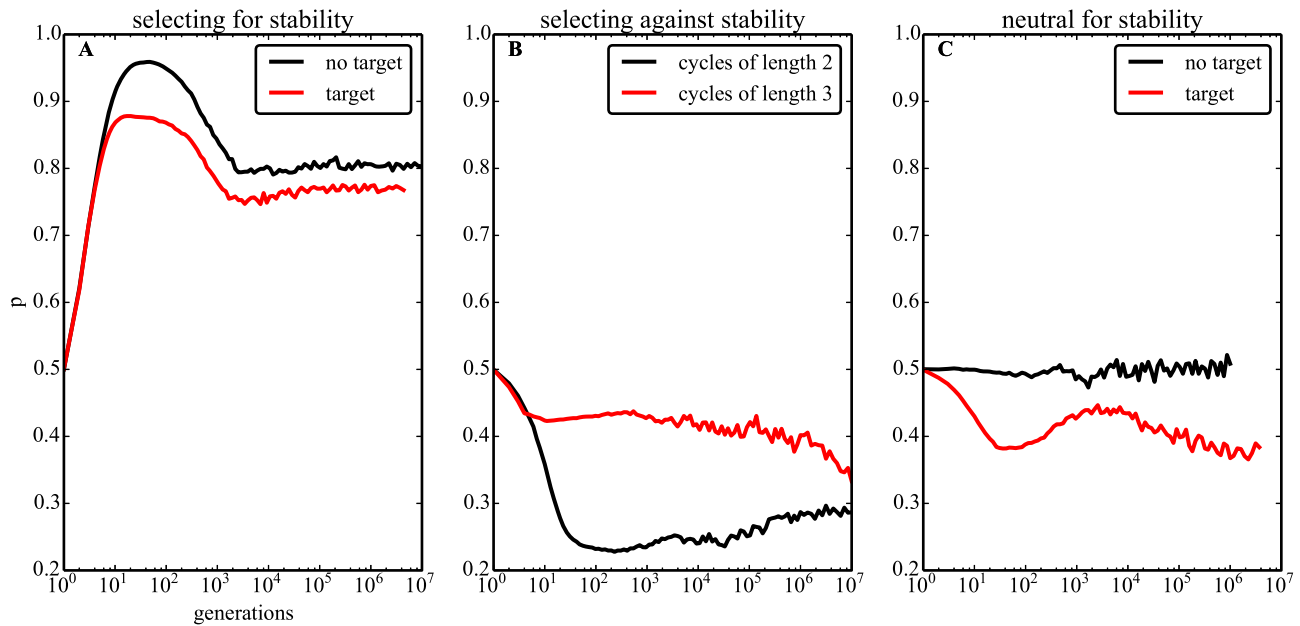
Figure 2. Stability selects for positive autoregulation. Time evolution of p (Equation (6)) on the vertical axis for different selection models. A. Selecting for stability: only fixed points are allowed to survive. 'no target' means all fixed points have fitness of 1 and all cycles have fitness of 0. 'target' means selection for a specific fixed point, with fitness given by Equation (2) as a function of the hamming distance from the target phenotype. B. Selecting against stability: only cycles of specific length are allowed to survive. Shown are selection for cycles of size 2 and 3. C. Neutral for stability: fitness does not depend on stability. 'no target' means random sampling. 'target' means selection for a specific fixed point, with fitness given by the maximum value of Equation (2) applied to all points of the attractor as explained in Methods (allows cycles of all sizes including fixed points). Sample size is $n \sim 10^5$ evolved networks per generation.
doi:10.1371/journal.pcbi.1003916.g002

Autoregulatory motifs are highly conserved over time

These results suggest that selection may act over direct autoregulatory motifs (i.e. the diagonal elements of the GRN matrix) to promote individual-level stability. If this is the case, positive diagonal elements should be overrepresented across evolved populations relative to off-diagonal elements, since selection for individual-level stability could be achieved by maximizing autoregulation p.

To test this hypothesis, we calculated the average interaction weights $\langle\langle w_{ij}\rangle_{n \times z}\rangle_T$ in populations evolved from a RNIC setup (the number of populations, z, was 100, and the population size $n = 500$). The metric $\langle\langle w_{ij}\rangle_{n \times z}\rangle_T$ is the average value of the w_{ij} entry of individual networks W, taken across a set of populations (including individuals within the populations) as well as evolutionary time. Extreme values of $\langle\langle w_{ij}\rangle_{n \times z}\rangle_T = -1$ or 1 indicate that the matrix element i,j is identical across individuals, different populations and different generations, whereas a neutral value of $\langle\langle w_{ij}\rangle_{n \times z}\rangle_T = 0$ means the matrix element i,j fluctuates randomly in individuals and populations and is not conserved over time.

We found that, in our evolution experiments, the averaged diagonal elements $\left(\langle\langle w_{ij}\rangle_{n \times z}\rangle_T \text{ for each } i\right)$ attained higher values than the off-diagonal elements, consistent with stronger selection for positive autoregulation acting on the diagonal elements (Figure S3). This is further supported by the observation that selecting for individual-level stability leads to positive values of $\langle\langle w_{ij}\rangle_{n \times z}\rangle_T$, whereas selecting against individual-level stability leads to negative $\langle\langle w_{ij}\rangle_{n \times z}\rangle_T$ values. Selecting neither for nor against stability, but still selecting for a specific target, also yields negative $\langle\langle w_{ij}\rangle_{n \times z}\rangle_T$.

To increase our confidence that the value of $p = 0.8$ emerging under the no-target model is maintained by selection for stability,

we compared the effects of diagonal and off-diagonal mutations on individual-level stability in the evolved networks. We hypothesized that since single mutations in diagonal elements could result in unstable networks more often than in off-diagonal elements, networks carrying diagonal entry mutations would therefore be weeded out at higher rates. Because p and stability are correlated, p would then be maintained at a high value.

In a first approach, we sampled $n = 10^5$ stable networks at $t = 10^6$ generations (equilibrium reached) evolved under the 'no-target' model. The overall fraction of networks that survive after acquiring single mutations (viability in Equation (4)) is high (95%). Consistent with our hypothesized maintenance mechanism, the population-level viability of networks is significantly higher for off-diagonal elements (median of 0.97 compared to 0.93 for diagonal; Mann-Whitney U p-value$\sim$0, Figure 3A). Intriguingly, the difference in viability is even higher for $t = 48$ generations, when p is close to maximum (Figure 3B), and the evolutionary dynamic has not yet reached equilibrium. These results suggest that a mutation of a diagonal element is more likely to lead to a cycle than mutation of off-diagonal elements.

In a second approach, we studied the conservation of values in diagonal versus non-diagonal elements between two given time points. To this end, we sampled $n = 10^5$ evolved stable networks at $t_1 = 10^5$ and $t_2 = 10^6$ generations from the 'no-target' and the random models. Subsequently, we counted the fraction of positive elements o_{ij}, (Equation (8)) across all networks and for each position in the gene regulatory matrix (w_{ij}) for both t_1 and t_2, and estimated conservation values as in Equation (8).

For the no-target model, we found that diagonal entries w_{ii} with positive sign are significantly more conserved over time than off-diagonal matrix elements (medians of 0.86 and 0.70, respectively; Mann Whitney U p-value$\sim$0, Figure 4A). These diagonal
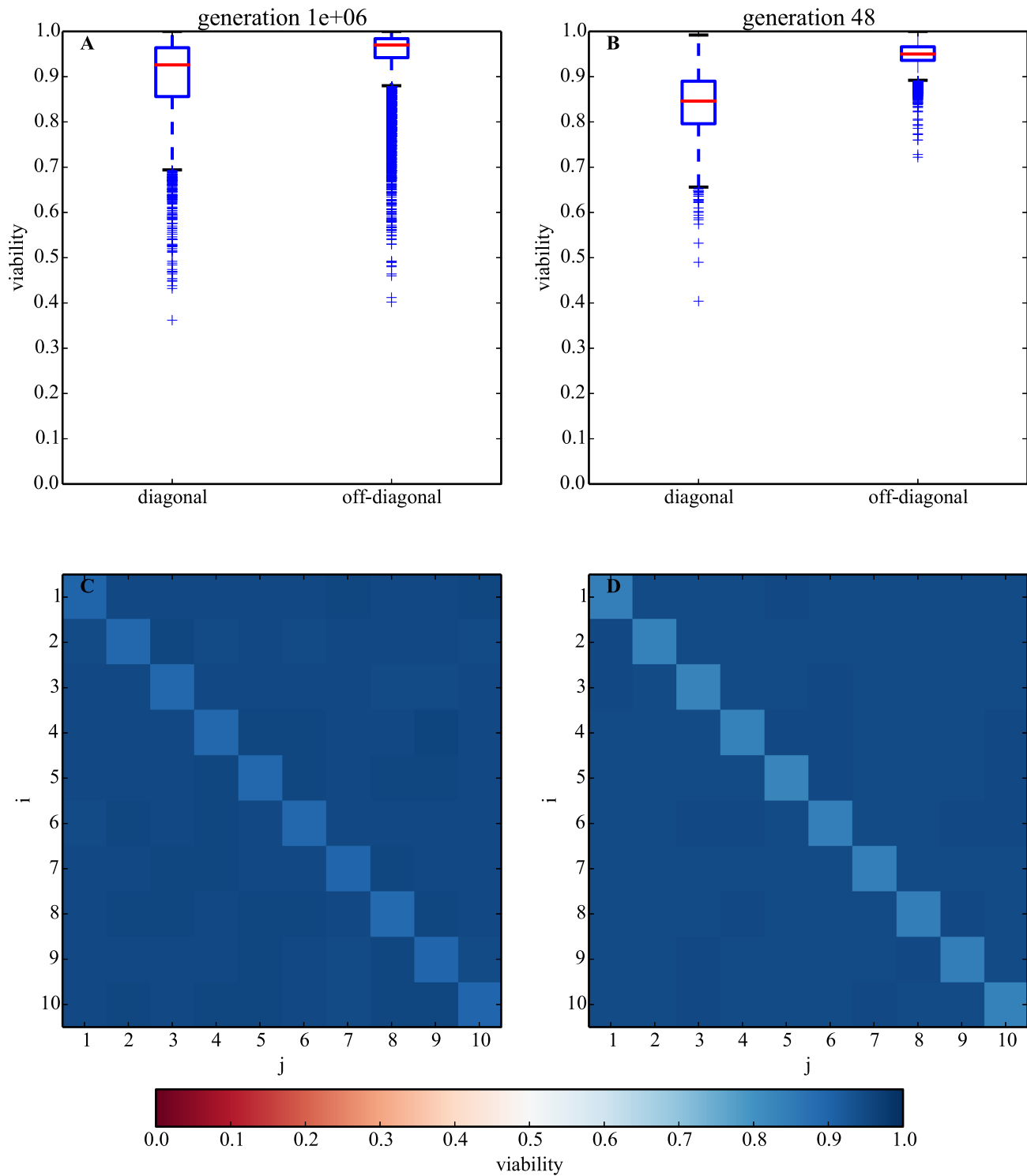
Figure 3. Changing the sign of an autoregulatory element is more likely to kill a network than a random mutation. A. Boxplots of average viability (Equation (4)) for diagonal and off-diagonal elements of the network matrices at evolutionary equilibrium ($t = 10^6$ generations). B. Same as (A) but at $t = 48$, when p (Equation (6)) is largest. C. Heat map of average viability for all matrix elements at $t = 10^6$ generations. Darker blue (off-diagonal elements) means higher viability after mutation than lighter blue (diagonal elements). D. Same as (C) but at $t = 48$. Sample size is $n \sim 10^5$ evolved networks per generation.
doi:10.1371/journal.pcbi.1003916.g003

elements under the no-target model are also more conserved when compared to diagonal elements evolved under the random model (medians of 0.86 and 0.68 respectively; Mann-Whitney U

p-value$\sim$0, Figure 4A). We found no significant differences in conservation of the off-diagonal entries between the 'no-target' and the random models (one-sided Mann-Whitney U
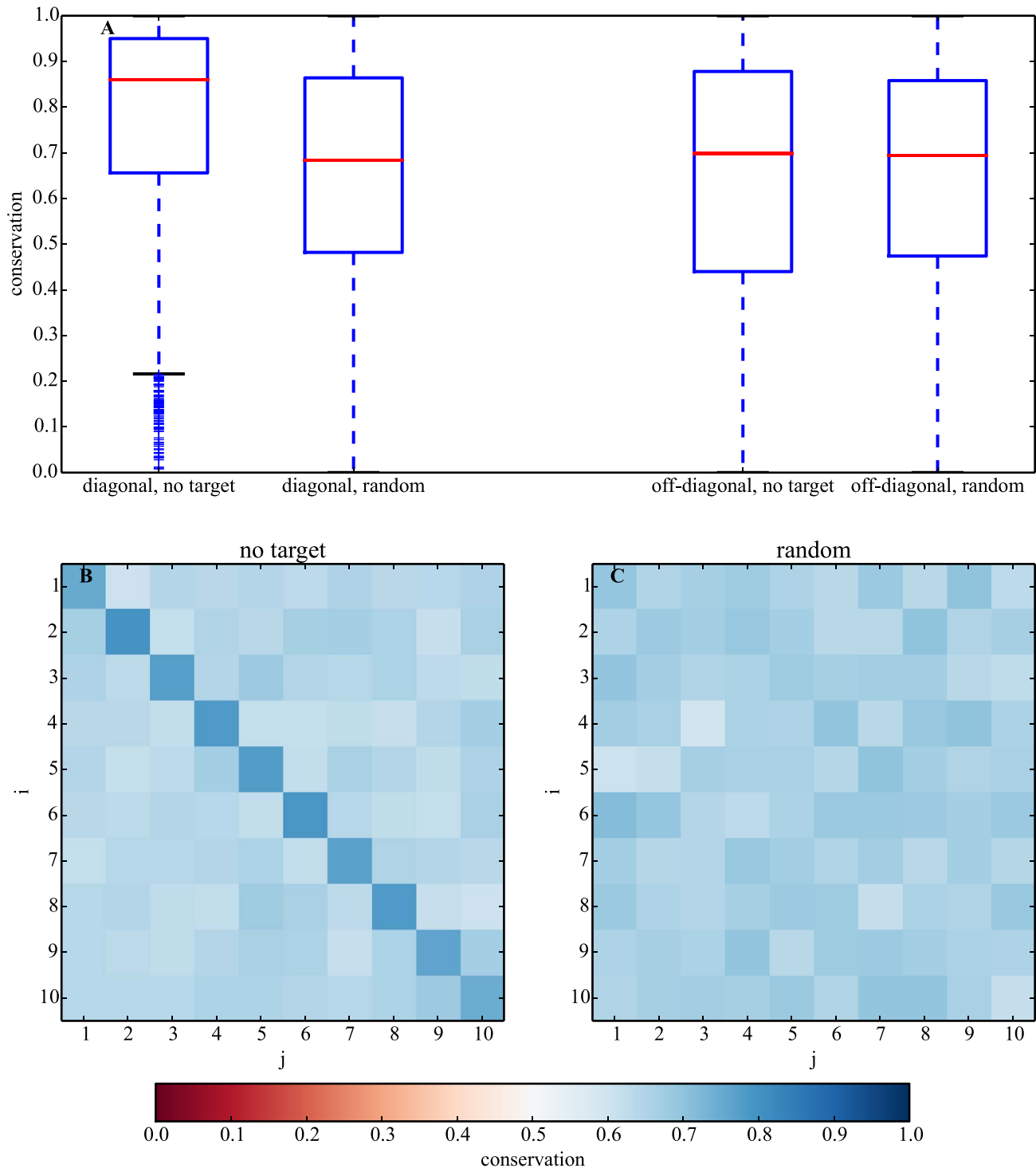
Figure 4. Autoregulation is conserved over time. A. Boxplots of average conservation (Equation (8), $t_1 = 10^5$ and $t_2 = 10^6$ generations) for diagonal and off-diagonal elements of random and evolved 'no target' networks. B. Heat map of average conservation for all matrix elements under the 'no-target' model. Darker blue (diagonal elements) means higher conservation than lighter blue (off-diagonal elements). C. Same as (B) but for random networks. Sample size is $n \sim 10^5$ evolved networks per generation.
doi:10.1371/journal.pcbi.1003916.g004

p-value~0.1, Figure 4A). This finding provides further evidence that positive autoregulation is maintained by selection for stability.

## Stability, robustness and autoregulation coevolve

The time course of p displays an intriguing complexity. After the stability metric reaches its maximum and ceases to change, p keeps evolving and decreases to a lower equilibrium value (Figure 2A). To investigate this behavior, we asked which other network parameters may also affect the evolution of p. Since Wagner [23] has previously shown that during network evolution robustness is also (indirectly) selected for when selecting for stability, we studied how this robustness compares with p over the course of evolution.

Intriguingly, during the simulation experiments under the no-target model, we observed that robustness (Equation (5)) increases with time and appears to coevolve with p, reaching its maximum at the same time point at which p reaches equilibrium (Figure 5). This association, however, is also not linear: shortly after stability has reached equilibrium, robustness still increases despite the fact that p has started decreasing.

Robustness and autoregulation are associated, but this relationship is dynamic during evolution

Because robustness and p reach equilibrium at about the same time, we hypothesized that p could be adapting under indirect selection for robustness. In that case, the equilibrium value of p~0.8 would favor higher robustness (or perhaps: "maximize robustness").

To test this, we generated groups of 10^5 random stable networks for $p = \{0.1, 0.2, 0.3, \ldots, 1\}$, and calculated the robustness for each group. Robustness was assessed after running one single develop-
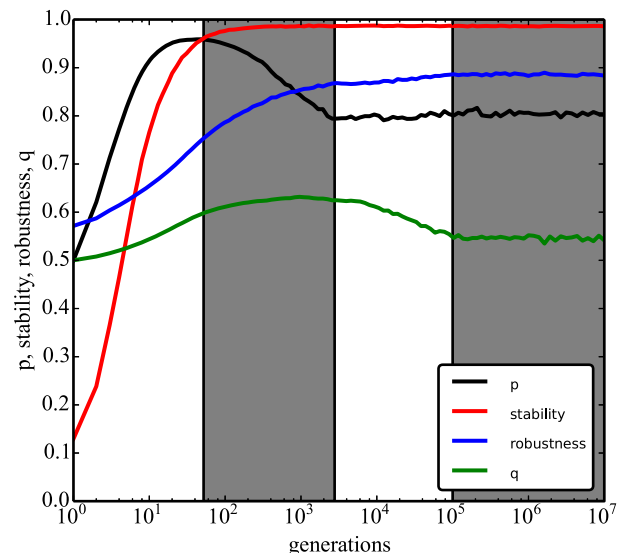
Figure 5. Positive/negative regulation evolves non-linearly with robustness. Coevolution of stability (Equation (3)), robustness (Equation (5)), p and q (Equation (6)) under selection for stability ('no-target' model). Black line (p) is the same as in Figure 2A. White and grey regions serve as guides to the eye signaling different phases of evolution. From $t = 1$ to $t \sim 48$, p reaches its maximum and stability reaches 0.96. From $t \sim 48$ to $t \sim 10^3$, p approaches equilibrium and q reaches its maximum. From $t \sim 10^3$ to $t \sim 10^5$, robustness and q approach equilibrium. From $t \sim 10^5$ onwards, all represented quantities are at evolutionary equilibrium. Sample size is $n \sim 10^5$ evolved networks per generation.
doi:10.1371/journal.pcbi.1003916.g005

ment process for each network in the RNRC setup. Only stable networks were considered for this analysis.

Surprisingly, we found that, similarly to stability, robustness is also positively associated with p and does not have a maximum at an intermediate p~0.8 (Figure 6A). Contrary to our hypothesis, robustness of stable, non-evolved networks is maximized by $p = 1$ (see Figure S4 for random networks not pre-selected for stability). This positive association was also observed when representing the data differently: stable networks binned by increasing average values of robustness also show increasing p (Figure 6B).

This general positive association is inconsistent with the hypothesized relationship between p and robustness after stability reaches its maximum. However, a more in-depth analysis of robustness of *evolving* networks at different time points reveals a completely different picture to the situation in non-evolved matrices (Figures 6C and D).

At both early ($t = 48$, $p_{\max}$) and late ($t = 10^6$, $p_{\text{equilibrium}}$) evolutionary stages, robustness is maximal in evolving matrices with values of $p < 1$. At early time points, when p has reached its maximum, the relationship between p and robustness is fully inverted compared to non-evolved networks, with lower p having significantly higher robustness, thus suggesting strong selection for lower p to increase robustness (Figure 6C). Strikingly, at equilibrium values, robustness is non-monotonic in p and is maximal for p~0.7–0.8, coinciding with the equilibrium value of p (Figure 6D). Therefore, these results strongly suggest that, as we hypothesized, it is the maximization of robustness during evolution that determines the equilibrium value of p.

## Evolved networks are a distinct subset of stable networks

The above results may seem contradictory: whereas p and robustness show a positive association in stable networks generated completely at random (i.e. for all non-evolved stable networks), this association is non-monotonic for stable networks selected by evolution (i.e. evolved networks), where robustness is maximized for more intermediate values of p (0.7–0.8). The solution to this apparent paradox might lie in the fact that evolved networks constitute only a subset of all stable networks.

Remarkably, the average robustness for the subset of evolved networks is twice as high as that of non-evolved networks with similar p (compare Figures 6A and 6D), suggesting that the relationship between p and robustness is modulated by other matrix characteristics on which selection can act.

In order to investigate this possibility, we studied the evolution of the sign of off-diagonal elements (q). There are more off-diagonal than diagonal elements; thus the former offer many more targets for mutation. However, a mutation in the off-diagonal has a smaller effect on q than a mutation on the diagonal has on p. For this reason, q may seem to evolve at lower rates than p. More importantly, off-diagonal elements represent regulation of other genes and can form larger and more complex motifs than autoregulatory loops of size one. For this reason, they are harder to study and to interpret. However, under a random model, it is clear that the expected average value of q equals 0.5.

The evolution of q occurs over a much smaller range than that of p, with values spanning from 0.5 to ~0.63. However, it also shows a non-linear pattern of co-evolution with the other variables (Figure 5). q increases up to its maximum approximately until p stabilizes, and then it starts to slowly decrease to its equilibrium value at $t \sim 10^5$. It reaches equilibrium around the same time as robustness.

These co-evolutionary dynamics suggests that stability and robustness may not only depend on p, but also on q. Therefore,
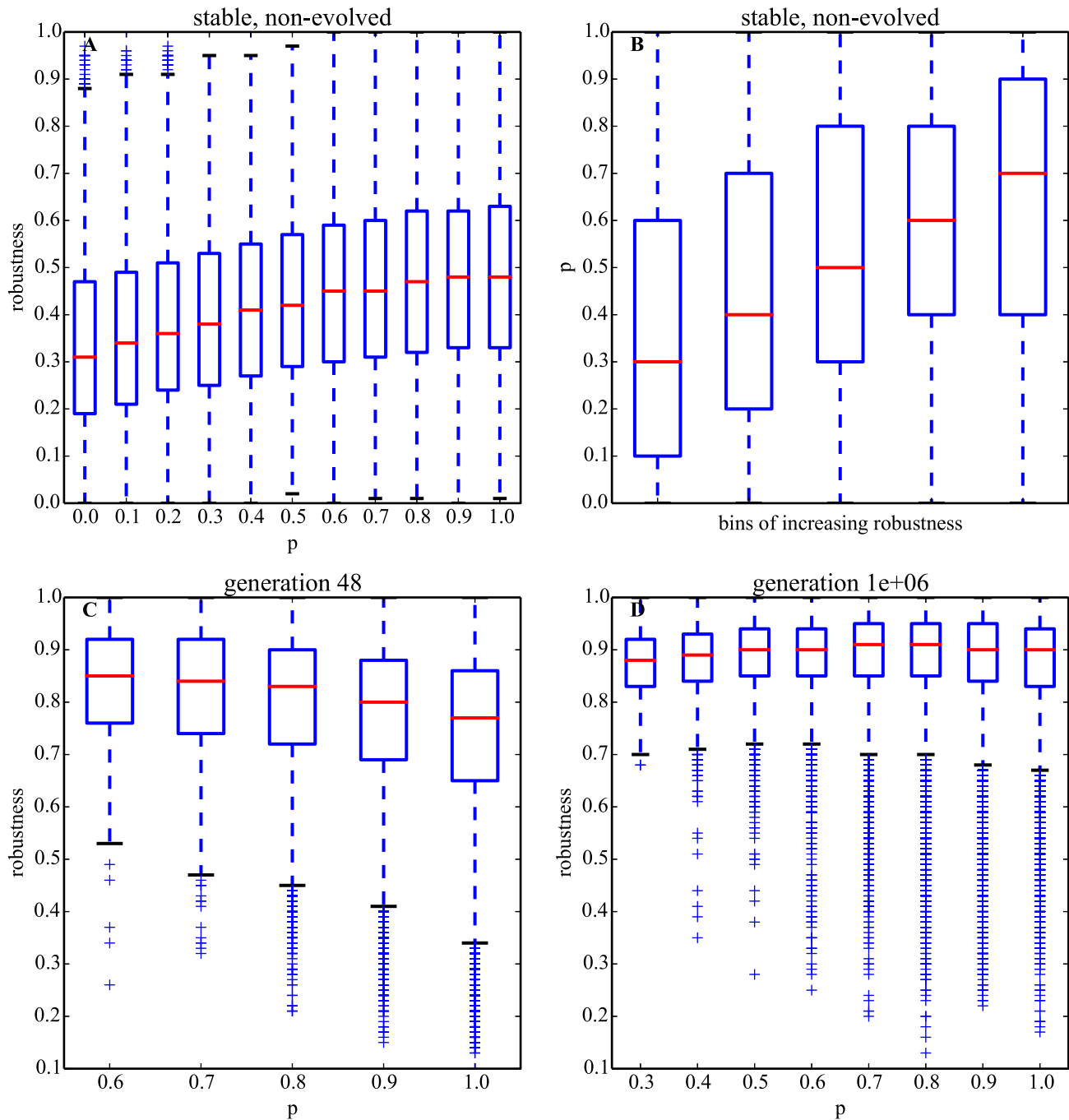
Figure 6. Maximization of robustness during evolution determines the equilibrium values of *p*. The relationship between robustness and autoregulation changes during evolution. A. Boxplots of robustness (Equation (5)) for each $p = 0, 0.1, \ldots, 1$ (11 bins) for stable, non-evolved networks ($q = 0.5$). B. The inverse of (A), i.e., boxplots of *p* (Equation (6)) for 5 bins of increasing robustness, each of size 0.2, from 0 to 1. C. Same as (A) but for evolved networks, at $t = 48$, when *p* is largest. At this stage of evolution, there are no significant samples of networks with $p < 0.6$. D. Same as (C) but at evolutionary equilibrium ($t = 10^6$ generations); there are no significant samples of evolved networks with $p < 0.3$. $q = 0.5$ (Equation (6)) for panels A and B but *q* is not constant or controlled for in panels C and D. (It varies with *p* in a non-linear fashion in evolved networks.) Sample size is $n = 10^4$ networks for each *p* (total n~10^5) for (A) and (B). Sample size of evolved networks varies with *p* but total size is n~10^5 for (C) and (D). Robustness is measured by Equation (5) with $N^2 = 100$ mutations per network.
doi:10.1371/journal.pcbi.1003916.g006

although robustness is maximized by high p for stable networks with $q = 0.5$, the same is not necessarily true when $q > 0.5$. This is the case at $q > 0.65$, for which robustness is higher for $p = 0.7$–0.8 than for $p = 1$ (Figure S5). Interestingly, these values correspond closely to ($p = 0.8$, $q = 0.63$) of evolved networks at generation ~1000, when both stability and p reach their equilibrium values.

Engineering super-robust networks: robustness is not fully maximized in evolution

A prediction from our results is that certain combinations of p and q are more likely to provide stable networks.

To test this, we combined the off-diagonal elements (determining q) of stable networks with low p and a q similar to that of equilibrium ($q = 0.53$ or 0.54), with diagonal elements of high p ($p = 0.9$ or 1.0). We call these networks "engineered". As before, we generated groups of 10^5 random stable networks for different values of (p,q) and calculated the robustness for each group. Robustness was assessed after running one single development process for each network in the RNRC setup. Only stable networks were considered for this analysis.

The engineered networks resulted in extremely stable and robust networks (Figure S6A); importantly, randomly sampled stable networks with the same average p and q are not nearly as stable and robust to mutations (Figure S6B). These observations support the idea that features in the topology of off-diagonal elements of these matrices (i.e., how genes regulate one another) may buffer the destabilizing effects of mutations.

These findings also show that it is possible to engineer networks more robust than those evolved under selection for stability. Thus, robustness is not fully maximized during evolution. In fact, when the founding populations ($t = 0$) have a mean of $q = 0.9$ (rather than being normally distributed around 0.5, see Methods), robustness decreases throughout evolution (Figure S7).

Positive autoregulatory motifs of length two are also associated with stability and robustness

Direct regulatory influence of genes on one another can explain qualitatively why p and q are being maximized at the beginning of evolution experiments. Their subsequent decline below the maximum values is related to constraints imposed by the indirect selection on robustness. To further elucidate how these constraints operate, we investigated how long-range interactions embedded in the matrix of direct influences (direct interaction, i.e., W) of an organism could contribute to the settling of p and q below their maximum values.

To this end, we repeated the evolutionary experiments tracking the measure for length-2 autoregulatory interaction r, which measures the frequency of networks that contain self-reinforcing interaction loops (gene A activates gene B, which reactivates A, see Methods). We found that r is maximized early and attains values above 0.5 throughout evolution, indicating positive long-range autoregulation (Figure S8). Additionally, r lags behind the evolution of p and q, adapting to selective pressures at lower rates.

To understand why engineered networks are more robust to mutations than random stable networks with the same values of p and q, we measured r for networks similar to the ones shown in Figure S6. We find that r is larger for the engineered networks (Figure S9), which explains why engineered networks are more robust for the same values of p and q.

Discussion

In this study, we have shown that stability and robustness positively correlate with autoregulation in a Boolean network model of gene regulation, where stable networks have mostly positive autoregulation ($p > 0.5$). During evolution in the no-target model, selecting for stability leads to indirect selection for robustness. Strong selection for stability is expressed in the adaptation of direct autoregulatory network properties summarized by p, which is maximized early in evolution. The subsequent decline of p is explained by additional autoregulatory effects stemming from long-range gene interactions that allow maintenance of high stability values, while simultaneously increasing robustness.

We have limited this study to small networks of 10 genes, comparable to some sub-circuits in genomes found in organisms, summarized in Table 1. We hypothesize that larger gene numbers would lead to similar results. In previous work we have shown that stability decreases with network size, which we simulated for up to $N = 10^4$ for sparse networks ($c = 0.2$) with scale-free topology [30].

We expect such a decrease in stability with N to increase the direct selective pressure on stability, as well as the indirect selective pressure on robustness in an evolutionary experiment. This is supported by the finding that large networks show a increase in robustness after selection for a target compared to small networks [23].

We have also neglected the role of bistability in the evolution of the networks. In other models of gene-regulatory networks it has been shown that mutational robustness correlates with the robustness of phenotypes to changes in initial conditions of the networks R_i [33]. If a similar correlation exists for the model presented in this study, we would expect indirect selection for less multi-stable networks due to the indirect selection for higher R_i. Networks with high R_i are expected to have large basins of attraction, decreasing the number of possible fixed points and thus multi-stability.

The model presented here deviates in some important aspects from Wagner's model [23] and Siegal & Bergman's model [29] by which it is inspired. In particular, our model only includes binary matrix elements, whereas [23,29] allow for real valued entries. Also, in contrast to [29], the normalization function used is not a sigmoidal, but a sign function. That results in the state vectors having real values in [29], whereas in our model we only allow for binary states.

The motivation to deviate from [23] lies in the focus on the sign of autoregulation. Since previous work [30] has shown that the behaviors of real-valued and binary-valued networks show little or no qualitative difference in the context of the questions asked here, it is technically more feasible to implement the easier, binary form of networks. Furthermore, most of the knowledge about gene regulatory networks exists in binary form, given as qualitative information about activation or repression interactions between genes. Thus, to make comparisons with the available data calculated on the basis of binary data, it was justified to limit the study to binary networks.

The assumption that a population of organisms consists of random networks or has random initial conditions is unrealistic. We use random samples of networks or initial phenotypic states because we are interested in the general, overall behavior of populations with respect to some metrics. Random sampling allows us to obtain an unbiased sample of all possible networks, and to capture a part of the heterogeneity in their behavior. To satisfy more realistic assumptions, a sub-space of phenotypes that corresponds to more realistic biological phenotypes needs to be specified. How to achieve this is currently unknown, and such a restriction would have amounted to studying random initial conditions.

The Boolean network model of gene regulation has recently been shown to predict specific patterns of protein and gene activity observed in a wide diversity of biological systems, including yeast [34,35] and mammalian [36] cell cycles, embryonic segmentation in D. melanogaster [36,37], and flower development in A. thaliana [38–40].

Assuming biological networks correspond to stable networks [23,29,34], our results suggest that biological networks should often be dominated by positive autoregulatory loops (i.e. have high p). This seems to be the case for most eukaryotic transcription factor networks (including yeast, flies and mammals), with various studies showing values of p ranging from 0.76 to 1 (Table 1; with the exception of early sea urchin developmental gene regulatory networks), and with autoregulatory loops being highly conserved across vertebrates [41].

Moreover, in some cases, the presence of strong positive autoregulatory loops seems to be crucial to achieving a stable biological state. For example, in mammalian embryonic stem cells, the core pluripotency network of Oct4, Sox2 and Nanog (plus Klf4 and Esrrb [42]) forms a tight autoregulated circuit, in which each gene activates its own expression as well as the expression of the others, and these interactions are crucial to maintaining a stable pluripotent state [43]. Furthermore, this autoregulatory circuit is likely behind the capacity of somatic cells to be reprogrammed into induced pluripotent stem (iPS) cells when reprogramming factors are expressed exogenously [44].

On the other hand, negative autoregulation seems to dominate in the bacterium E. coli ($p = 0.26$) [45]. Stewart and coworkers [46] have recently suggested that this difference may be due to the presence/absence of sexual reproduction. To test this hypothesis, we reproduced our simulations for evolution without recombination (see Methods) under the no-target model, as a proxy for a model with asexual reproduction, but obtained essentially the same equilibrium values of p, despite divergent intermediate evolutionary dynamics and robustness at equilibrium (Figure S10).

Another caveat may lie in the density of the networks employed in our simulations. Biological networks are often sparse [47], and may vary between species as well as for different gene regulatory subcircuits within species; however, we have used fully connected networks in our analyses. Thus, we tested Boolean networks with the same average connectivity as some biological networks (average degree of 2, [47]) (see Methods, Sparse Networks). The evolutionary simulations were conducted under the *RNIC setup* (no pre-selected networks, see Methods), in a similar fashion to the previous simulations. We obtained similar results for the long-term evolution of q, and for p in sparse networks without recombination, while we obtained even larger values of p at equilibrium for sparse networks with recombination (Figure S11), suggesting that connectivity density has a minor impact on the evolution of these parameters. These results are aligned with our previous study showing that network density and topology have only a small effect on the stability of networks of 10 genes [30].

Finally, the differences between eukaryotic and bacterial autoregulation values may also relate to the distinct regulatory processes of bacteria (e.g. common presence of operons) and eukaryotes (e.g. more widespread post-transcriptional regulation). As new circuits of transcription factor networks are elucidated in detail, the roles of negative and positive autoregulation in organismal development and evolution should be more clearly understood.

Supporting Information

Figure S1   Stability favors positive autoregulation. Error bar chart of sign of autoregulation (p, Equation (6)) as a function of

stability (Equation (3)) for two network sizes $N = 4$ and 10. This represents the inverse relationship of Figure 1B. Error bars represent standard deviations.
(EPS)

Figure S2   Stability and cycles select for opposite signs of autoregulation. Time evolution of p (Equation (6)) for different selection models. This reproduces Figure 2 with the added results of cycles of length up to 7, all in the same panel. Symbols and selection models are explained in Methods.
(EPS)

Figure S3   Stability and cycles select for opposite signs of regulation. Heat maps of the average matrix for different selection models. This is a different way of representing Figures 2 and S2, with the added information of the sign of off-diagonal elements. The matrices were averaged across individuals, populations and evolutionary time as explained in Results.
(EPS)

Figure S4   Positive autoregulation favors robustness. Same as Figures 6A and B but for random networks not pre-selected for stability. A. Boxplots of robustness (Equation (5)) for each $p = 0, 0.1,…, 1$. B. The inverse of (A), i.e., boxplots of p (Equation (6)) for 5 bins of increasing robustness, each of size 0.2, from 0 to 1.
(EPS)

Figure S5   The sign of off-diagonal elements changes the relationship between robustness and autoregulation. Robustness is not always maximized by $p = 1$ in stable networks. Here we show maximization at ($p = 0.9$, $q = 0.65$) and ($p = 0.6$, $q = 0.70$), for example. Networks are non-evolved. Lines serve as guides to the eye as p takes discrete values between 0 and 1.
(EPS)

Figure S6   Engineering super-robust networks: there's more to the topology of robust networks. Boxplots of robustness for specific values of p and q. A. Networks were engineered by combining the off-diagonal elements of stable networks with low p and a q similar to that of the equilibrium ($q = 0.53$ or 0.54), with a diagonal with high p ($p = 0.9$ or 1.0). B. Randomly sampled stable networks with same average p and q as the engineered ones.
(EPS)

Figure S7   Robustness is not fully maximized in evolution. Coevolution of p (A), q (B), stability (C), and robustness (D) under selection for stability ('no-target' model). Same as Figure 5 but for five different starting conditions: (p_0, q_0) = (0.5, 0.5), (0.5, 0.9), (0.5, 0.1), (0.9, 0.9), (0.9, 0.1). All networks in a population have the same (p_0, q_0) at the start of evolution.
(EPS)

Figure S8   Positive long-range autoregulatory motifs are also associated with stability and robustness. Same as Figure 5 with the inclusion of r given by Equation (7).
(EPS)

Figure S9   Engineered networks have positive long-range autoregulatory interactions. Boxplots of r (Equation (7)) for a subset of the networks shown in Figure S6. A. Stable networks with $p = 0.1$ and $q = 0.54$ ("engineered"). B. Stable networks with $p = 0.9$, $q = 0.53$ ("stable"). Population size is $n = 10^4$ for each.
(EPS)

Figure S10   Sign of regulation evolves to the same values with or without recombination. Coevolution of p (A), q (B),

stability (C), and robustness (D) with and without recombination. Black lines (recombination) are the same as Figure 5. Recombination is modeled as full chromosome segregation (no crossover) as described by Wagner [23].
(EPS)

Figure S11 Autoregulation is even more positive for sparse networks with recombination. Coevolution of p (A), q (B), stability (C), and robustness (D) for dense and sparse networks with and without recombination. Black lines (dense networks with recombination) are the same as Figure 5. Dense networks are fully connected (all genes regulate themselves and each other). Each gene in sparse networks regulates and is

regulated by two other genes (average network degree of 2, as in biological networks [47]).
(EPS)

References

1. King M-C, Wilson A (1975) Evolution at two levels in humans and chimpanzees. Science 188: 107–116. doi:10.1126/science.1090005.
2. Carroll SB (2005) Evolution at two levels: on genes and form. PLoS Biol 3: e245. doi:10.1371/journal.pbio.0030245.
3. Davidson EH (2006) The regulatory genome: gene regulatory networks in development and evolution. Academic Press.
4. Jeong H, Tombor B, Albert R, Oltvai ZN, Barabasi A-L (2000) The large-scale organization of metabolic networks. Nature 407: 651–654.
5. Milo R, Shen-Orr SS, Itzkovitz S, Kashtan N, Chklovskii D, et al. (2002) Network motifs: simple building blocks of complex networks. Science 298: 824–827. doi:10.1126/science.298.5594.824.
6. Barkai N, Leibler S (1997) Robustness in simple biochemical networks. Nature 387: 913–917. doi:10.1038/43199.
7. Wagner GP, Altenberg L (1996) Perspective: Complex Adaptations and the Evolution of Evolvability. Evolution (N Y) 50: 967–976.
8. McAdams HH, Arkin AP (1997) Stochastic mechanisms in gene expression. Proc Natl Acad Sci U S A 94: 814–819.
9. Becskei A, Serrano L (2000) Engineering stability in gene networks by autoregulation. Nature 405: 590–593. doi:10.1038/35014651.
10. Shen-Orr SS, Milo R, Mangan S, Alon U (2002) Network motifs in the transcriptional regulation network of Escherichia coli. Nat Genet 31: 64–68. doi:10.1038/ng881.
11. Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, et al. (2002) Transcriptional Regulatory Networks in Saccharomyces cerevisiae. Science 298: 799–804. doi:10.1126/science.1075090.
12. Hogan DJ, Riordan DP, Gerber AP, Herschlag D, Brown PO (2008) Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. PLoS Biol 6: e255. doi:10.1371/journal.pbio.0060255.
13. Thomas R (1981) On the Relation Between the Logical Structure of Systems and Their Ability to Generate Multiple Steady States or Sustained Oscillations. Numer methods study Crit Phenom 9: 180–193.
14. Thomas R, Thieffry D, Kaufman M (1995) Dynamical behaviour of biological regulatory networks—I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state. Bull Math Biol 57: 247–276. doi:10.1007/BF02460618.
15. Bateman E (1998) Autoregulation of eukaryotic transcription factors. Prog Nucleic Acid Res Mol Biol 60: 133–168. doi:10.1016/S0079-6603(08)60892-2.
16. Aronson B, Johnson K, Loros J, Dunlap J (1994) Negative feedback defining a circadian clock: autoregulation of the clock gene frequency. Science 263: 1578–1584. doi:10.1126/science.8128244.
17. Alon U (2007) Network motifs: theory and experimental approaches. Nat Rev Genet 8: 450–461. doi:10.1038/nrg2102.
18. Snoussi EH, Thomas R (1993) Logical identification of all steady states: The concept of feedback loop characteristic states. Bull Math Biol 55: 973–991. doi:10.1016/S0092-8240(05)80199-5.
19. Thomas R (1998) Laws for the dynamics of regulatory networks. Int J Dev Biol 42: 479–485.
20. Comet J-P, Noual M, Richard A, Aracena J, Calzone L, et al. (2013) On Circuit Functionality in Boolean Networks. Bull Math Biol. doi:10.1007/s11538-013-9829-2.
21. McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 5: 115–133. doi:10.1007/BF02478259.
22. Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. PNAS 79: 2554–2558.
23. Wagner A (1996) Does evolutionary plasticity evolve? Evolution (N Y) 50: 1008–1023.
24. Kauffman SA (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. J Theor Biol 22: 437–467. doi:10.1016/0022-5193(69)90015-0.
25. Kurten KE (1988) Correspondence between neural threshold networks and Kauffman Boolean cellular automata. J Phys A Math Gen 21: L615–L619. doi:10.1088/0305-4470/21/11/009.
26. Greil F, Drossel B (2005) Dynamics of critical Kauffman networks under asynchronous stochastic update. Phys Rev Lett 95: 3–6. doi:10.1103/PhysRevLett.95.048701.
27. Greil F, Drossel B, Sattler J (2007) Critical Kauffman networks under deterministic asynchronous update. New J Phys 9: 373–373. doi:10.1088/1367-2630/9/10/373.
28. Klemm K, Bornholdt S (2005) Stable and unstable attractors in Boolean networks. Phys Rev E 72: 1–4. doi:10.1103/PhysRevE.72.055101.
29. Siegal ML, Bergman A (2002) Waddington's canalization revisited: developmental stability and evolution. Proc Natl Acad Sci U S A 99: 10528–10532. doi:10.1073/pnas.102303999.
30. Pinho R, Borenstein E, Feldman MW (2012) Most Networks in Wagner's Model Are Cycling. PLoS One 7: e34285. doi:10.1371/journal.pone.0034285.
31. Hamming RW (1950) Error detecting and error correcting codes. Bell Syst Tech J 29: 147–160.
32. McDonald D, Waterbury L, Knight R, Betterton MD (2008) Activating and inhibiting connections in biological network dynamics. Biol Direct 3: 49. doi:10.1186/1745-6150-3-49.
33. Ciliberti S, Martin OC, Wagner A (2007) Robustness can evolve gradually in complex regulatory gene networks with varying topology. PLoS Comput Biol 3: 164–173. doi:10.1371/journal.pcbi.0030015.
34. Li F, Long T, Lu Y, Ouyang Q, Tang C (2004) The yeast cell-cycle network is robustly designed. PNAS 101: 4781–4786. doi:10.1073/pnas.0305937101.
35. Davidich MI, Bornholdt S (2008) Boolean network model predicts cell cycle sequence of fission yeast. PLoS One 3: e1672. doi:10.1371/journal.pone.0001672.
36. Fauré A, Naldi A, Chaouiya C, Thieffry D (2006) Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. Bioinformatics 22: e124–31. doi:10.1093/bioinformatics/btl210.
37. Albert R, Othmer HG (2003) The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in Drosophila melanogaster. J Theor Biol 223: 1–18. doi:10.1016/S0022-5193(03)00035-3.
38. Mendoza L, Thieffry D, Alvarez-Buylla ER (1999) Genetic control of flower morphogenesis in Arabidopsis thaliana: a logical analysis. Bioinformatics 15: 593–606.
39. Espinosa-Soto C, Padilla-Longoria P, Alvarez-Buylla ER (2004) A gene regulatory network model for cell-fate determination during Arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles. Plant Cell 16: 2923–2939. doi:10.1105/tpc.104.021725.
40. Thum KE, Shasha DE, Lejay L V, Coruzzi GM (2003) Light- and carbon-signaling pathways. Modeling circuits of interactions. Plant Physiol 132: 440–452. doi:10.1104/pp.103.022780.
41. Kiełbasa SM, Vingron M (2008) Transcriptional autoregulatory loops are highly conserved in vertebrate evolution. PLoS One 3: e3210. doi:10.1371/journal.pone.0003210.
42. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, et al. (2013) Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell 153: 307–319. doi:http://dx.doi.org/10.1016/j.cell.2013.03.035.
43. Young RA (2011) Control of the embryonic stem cell state. Cell 144: 940–954. doi:10.1016/j.cell.2011.01.032.
44. Jaenisch R, Young R (2008) Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell 132: 567–582. doi:10.1016/j.cell.2008.01.015.
45. Salgado H, Peralta-Gil M, Gama-Castro S, Santos-Zavaleta A, Muñiz-Rascado L, et al. (2013) RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. Nucleic Acids Res 41: D203–13. doi:10.1093/nar/gks1201.
46. Stewart AJ, Seymour RM, Pomiankowski A, Reuter M (2013) Under-Dominance Constrains the Evolution of Negative Autoregulation in Diploids. PLoS Comput Biol 9: e1002992. doi:10.1371/journal.pcbi.1002992.
47. Leclerc RD (2008) Survival of the sparsest: robust gene networks are parsimonious. Mol Syst Biol 4: 213. doi:10.1038/msb.2008.52.

48. Jaeger J, Blagov M, Kosman D, Kozlov KN, Manu, et al. (2004) Dynamical analysis of regulatory interactions in the gap gene system of Drosophila melanogaster. Genetics 167: 1721–1737. doi:10.1534/genetics.104.027334.

49. Azevedo RBR, Lohaus R, Srinivasan S, Dang KK, Burch CL (2006) Sexual reproduction selects for robustness and negative epistasis in artificial gene networks. Nature 440: 87–90. doi:10.1038/nature04488.

50. Bell-Pedersen D, Cassone VM, Earnest DJ, Golden SS, Hardin PE, et al. (2005) Circadian rhythms from multiple oscillators: lessons from diverse organisms. Nat Rev Genet 6: 544–556. doi:10.1038/nrg1633.

51. Locke JCW, Southern MM, Kozma-Bognár L, Hibberd V, Brown PE, et al. (2005) Extension of a genetic network model by iterative experimentation and mathematical analysis. Mol Syst Biol 1: 2005.0013. doi:10.1038/msb4100018.

52. Bonzanni N, Garg A, Feenstra KA, Schütte J, Kinston S, et al. (2013) Hard-wired heterogeneity in blood stem cells revealed using a dynamic regulatory network model. Bioinformatics 29: i80–8. doi:10.1093/bioinformatics/btt243.

53. Cameron RA, Samanta M, Yuan A, He D, Davidson EH (2009) SpBase: the sea urchin genome database and web site. Nucleic Acids Res 37: D750–4. doi:10.1093/nar/gkn887.