Multifunctionality and Robustness Trade-Offs in Model Genetic Circuits

Olivier C. Martin* and Andreas Wagner[†]

*Univ Paris-Sud, UMR8120, Laboratoire de Genetique Vegetale du Moulon, INRA, and CNRS, Gif-sur-Yvette, F-91190, France; and
[†]University of Zurich, Department of Biochemistry, CH-8057 Zurich, Switzerland

ABSTRACT   Most cellular systems, from macromolecules to genetic networks, have more than one function. Examples involving networks include the transcriptional regulation circuits formed by *Hox* genes and the *Drosophila* segmentation genes, which function in both early and later developmental events. Does the need to carry out more than one function severely constrain network architecture? Does it imply robustness trade-offs among functions? That is, if one function is highly robust to mutations, are other functions highly sensitive, and vice versa? Little available evidence speaks to these questions. We address them with a general model of transcriptional regulation networks. We show that requiring a regulatory network to carry out additional functions constrains the number of permissible network architectures exponentially. However, robustness of one function to regulatory mutations is uncorrelated or weakly positively correlated to robustness of other functions. This means that robustness trade-offs generally do not arise in the systems we study. As long as there are many alternative network structures, each of which can fulfill all required functions, multiple functions may acquire high robustness through gradual Darwinian evolution.

INTRODUCTION

Most quantitative models of cellular circuits are severely limited by many unknown biochemical parameters determining circuit behavior. Not only that, these parameters change constantly, because of nongenetic perturbations such as gene expression noise and environmental change, and because of mutations. This means that the regulatory topology of circuits—the who-interacts-with-whom—must become a focus of investigation, because much else about a circuit may be in constant flux.

We focus here on one aspect of circuit organization that receives ever-increasing attention: the robustness of cellular circuits to mutations and nongenetic change (1–10). Many cellular circuits are subject to constant perturbations, and they need to keep performing their function in the face of these perturbations. Studies of robustness in genetic circuits typically focus on one specific function of a biological circuit (1–4,11–13). For any one function, there may be many different network architectures or topologies that are equally capable of performing this function; these topologies may differ widely in their robustness (15,16), and high robustness may be evolvable through gradual stepwise changes of individual topologies (16).

Any one cellular circuit and its genes typically have more than one function in the organism. Prominent examples include the transcriptional regulation circuitry of *Hox* genes in organisms as different as fruit flies and mammals. For example, the mouse genome contains some 40 *Hox* genes, which influence each other's expression through transcrip-

tional cross-and autoregulation. This regulatory gene network plays a key role in patterning the main anteroposterior body axis. In addition, it is also centrally involved in a distinct developmental process, the patterning of the vertebrate limb. The network experiences different regulatory inputs in each of these two embryonic regions, and produces different gene expression outputs in response. Another example involves the *Drosophila* segment polarity genes, which include *wingless*, *engrailed*, and *hedgehog*. These genes are central to the segmentation of the *Drosophila* embryo, but they play equally important roles in later developmental processes, such as the development of the fly's wing (17,18).

The requirement to perform more than one function constrains the architectures of such networks. It is not clear whether the above observations about robustness of monofunctional circuits would also apply to circuits with more than one function. How strongly do additional functions constrain network topology? Do additional functions affect the extent to which a network can be robust to noise and mutations? Is a network topology that is robust with respect to one function also robust with respect to another? And, finally, is the gradual evolution of high robustness through stepwise architectural changes possible for circuits with more than one function?

We here make a small step toward answering these questions by studying a simple model of transcriptional regulation networks (Fig. 1). Despite being quite abstract, variants of this model have proved highly successful in explaining the regulatory dynamics of early developmental genes in the fruit fly *Drosophila*, as well as in predicting mutant phenotypes (19–22). The model has also helped elucidate why mutants often show a release of genetic variation that is cryptic in the wild-type, and how adaptive evolution of robustness occurs in genetic networks of a given topology (23–25). Most recently, it has also proved useful in explaining how sexual
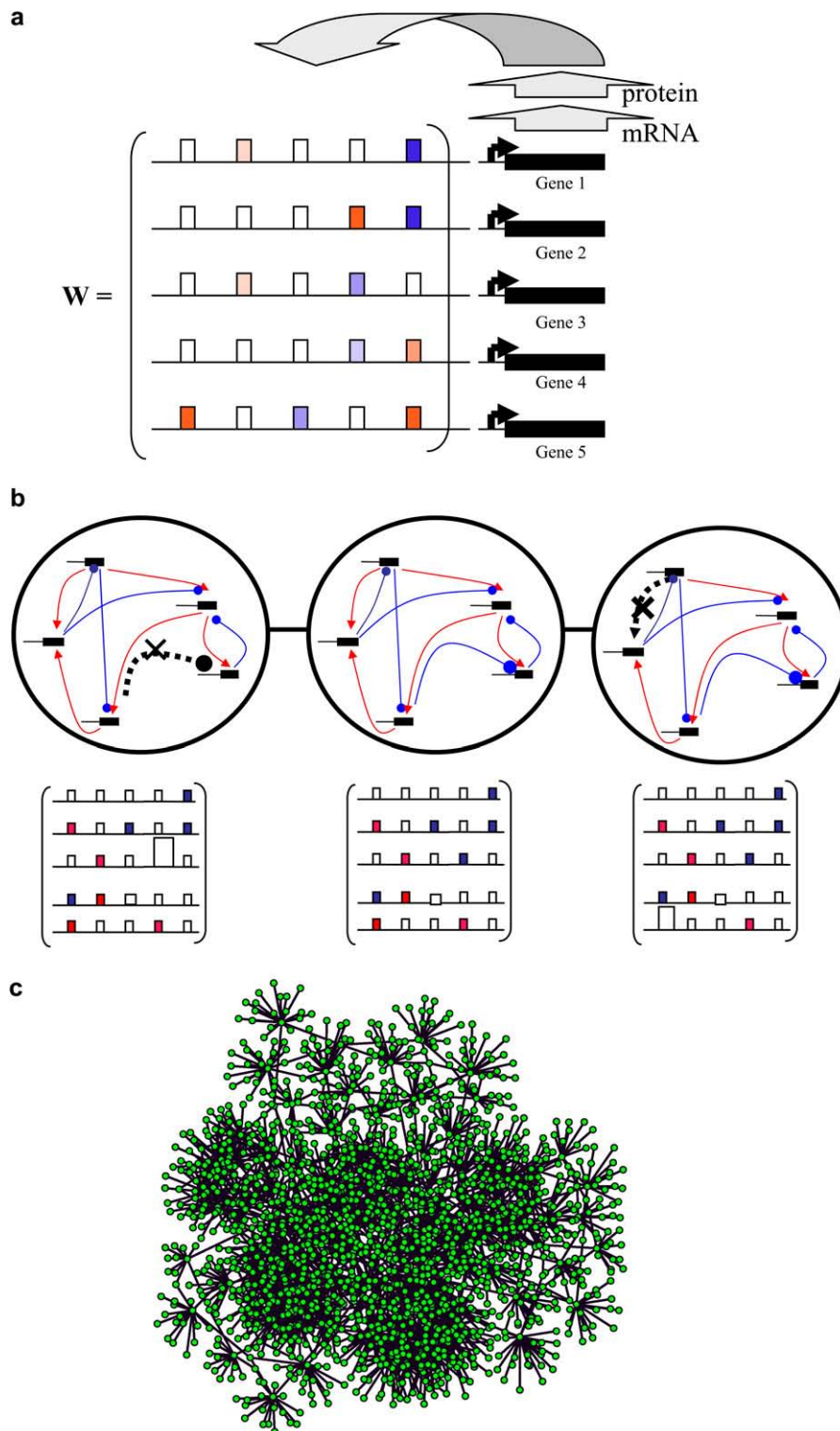
FIGURE 1 (*a*) Transcriptional regulation network (13). Solid black bars indicate genes that encode transcriptional regulators in a hypothetical five-gene network. Each gene is expressed at a rate that is influenced by the transcriptional regulators in the network. This influence is usually exerted by binding of a transcriptional regulator to a gene's regulatory region (*horizontal line*). The model represents the regulatory interactions between transcription factor j and genes i through a matrix $w = (w_{ij})$. A regulator's effect can be activating ($w_{ij} > 0$, *red rectangles*) or repressing ($w_{ij} < 0$, *blue rectangles*). Any given gene's expression may be unaffected by most regulators in the network ($w_{ij} = 0$, open rectangles). The different hues of red and blue correspond to different magnitudes of w_{ij}. The highly regular correspondence of matrix entries to binding sites serves the purpose of illustration and is not normally found, because transcription factor binding sites usually function regardless of their position in a regulatory region. (b) Gradual evolutionary changes and the metagraph. The middle panel shows a hypothetical network of five genes (*top*) and its matrix of regulatory interactions w (*bottom*), if genes are numbered clockwise from the uppermost gene. Red arrows indicate activating interactions and blue lines terminating in a circle indicate repressive interactions. The left-most network and the middle network differ in one repressive interaction from gene 4 to gene 3 (*dashed gray line, black cross, large open rectangle*). The right-most network and the middle network differ in one activating interaction from gene 1 to gene 5 (*dashed line, black cross, large open rectangle*). Each of the three network topologies corresponds to one node in a metagraph of network topologies, which is indicated by the large circle around the networks. These circles are connected because the respective networks are neighbors in the metagraph, i.e., they differ by one regulatory interaction. (*c*) Part of a metagraph for a network of $N = 4$ genes. Each node corresponds to a network of a given topology ($w_{ij} = \pm 1, 0$), and two nodes are connected by an edge if they differ at one regulatory interaction. ($8 \leq M \leq 9$ regulatory interactions, one input-target pair, and Hamming distance of $\mathbf{S}(0)$ and $\mathbf{S}_\infty$ of $d = 0.5$). The metagraph of this network is connected and the number of edges incident on a node is highly variable. The graph shown includes all viable networks that differ at no more than four regulatory interactions from an arbitrary node in the metagraph. Note that metagraphs typically have a huge number of nodes. The number of networks in a metagraph can be counted, because different nodes differ only in the signs of their regulatory interactions.

reproduction can enhance robustness to recombination (12). The model (23) is concerned with a regulatory network of N transcriptional regulators, which are represented by their expression patterns $\mathbf{S}(t) = (S_1(t), S_2(t), \ldots, S_N(t))$ at some time t during a developmental or cell-biological process and in one cell or domain of an embryo. These transcriptional regulators can influence each other's expression through cross-regulatory and autoregulatory interactions, which are

encapsulated in a matrix $w = (w_{ij})$. The elements w_{ij} of this matrix indicate the strength of the regulatory influence that gene j has on gene i (Fig. 1 a). This influence can be either activating ($w_{ij} > 0$), repressing ($w_{ij} < 0$), or absent. Put differently, the matrix w represents the (regulatory) genotype of this system, whereas the expression state is its phenotype. We model the change in the expression state of the network $\mathbf{S}(t)$ as time t progresses according to the difference equation $S_i(t+\tau) = \sigma\left[\sum_{j=1}^{N} w_{ij} S_j(t)\right]$, where τ is a constant and $\sigma(.)$ is a steep sigmoidal function whose values lie in the interval $(-1, +1)$. This equation reflects the regulation of gene i's expression by other genes. We are here concerned with networks whose expression dynamics starts from a prespecified initial state $\mathbf{S}(0)$ at some time $t = 0$ during development, and arrives at a prespecified stable equilibrium or "target" expression state $\mathbf{S}_\infty$. We will call such networks viable networks. The initial state can be thought of as being determined by regulatory factors upstream of the network, which may represent signals from the cell's environment or from other domains of an embryo. Transcriptional regulators that are expressed in the stable equilibrium state $\mathbf{S}_\infty$ may affect the expression of genes downstream of the network. We think of their expression as critical for the course of development. Thus, deviations from $\mathbf{S}_\infty$ are highly deleterious. To address the above questions about functional constraints, we will examine networks w that have two or more pairs of prespecified initial-target expression states. In the context of this model, we refer to these pairs as network functions. We denote these pairs as $(\mathbf{S}^{(1)}(0), \mathbf{S}_\infty^{(1)})$, $(\mathbf{S}^{(2)}(0), \mathbf{S}_\infty^{(2)})$, etc. We are acutely aware of the limitations of using an abstract model like ours. We are nonetheless compelled to use such a model, because there is a complete lack of empirical information about trade-offs in robustness, and because such information cannot be obtained with currently available experimental technologies.

RESULTS

Additional functions severely constrain network architecture

We first asked how the fraction of viable networks among all networks depends on the number of genes N and on the number (one or two) of network functions. To enumerate viable networks, we needed to focus on discrete genotypes ($w_{ij} = \pm 1, 0$), but we will show that our major conclusions hold also for networks with continuous interactions.

In this analysis, we focus on networks that have a number M of regulatory interactions within a given range.

Because there are 2^N possible equilibrium states, the probability that any network w arrives at any one single $\mathbf{S}_\infty$ should be of the order of $1/2^N$. In our numerical analysis (see Methods in the Supplementary Material), we find indeed an exponential scaling in N for this probability (13), which should decrease even more strongly as a function of N if we

require the network to arrive at more than one prespecified $\mathbf{S}_\infty$ from different initial states. Fig. 2 a shows that this is the case for networks with two input-target pairs. Open bars in the figure indicate the fraction p of viable monofunctional networks among all networks. Black bars indicate the fraction of bifunctional networks among all networks, averaged over random input-target pairs. Note the logarithmic scale on the vertical axis, and that the fraction of viable bifunctional
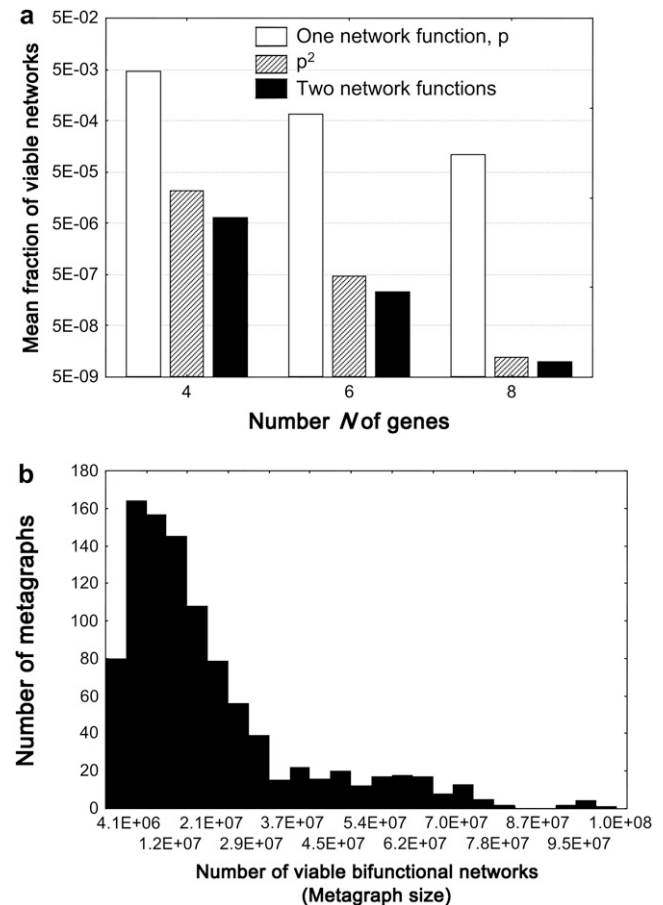

FIGURE 2 (a) Fraction p of viable monofunctional networks (*open bars*) and viable bifunctional networks (*solid bars*) for varying numbers of genes (*horizontal axis*). Shaded bars indicate p^2 and show that p^2 can serve as an approximation for the fraction of viable of bifunctional networks. Values for these fractions were obtained by random sampling, followed by averaging >1000 randomly chosen pairs of $(\mathbf{S}^{(1)}(0), \mathbf{S}_\infty^{(1)})$ and $(\mathbf{S}^{(2)}(0), \mathbf{S}_\infty^{(2)})$, as described in Methods (Supplementary Material). Standard errors of the means shown are 2.75×10^{-7}, 6.98×10^{-9}, and 6.32×10^{-10} for networks of sizes $N = 4, 6, 8$, respectively, and thus too small to be shown in the plot. $M \approx 0.5\, N^2$ nonzero regulatory interactions. (b) Metagraph sizes for bifunctional networks have a broad distribution. The horizontal axis indicates metagraph size (in number of networks) and the vertical axis indicates the number of metagraphs (out of 1000) with a size indicated on the horizontal axis. Even the largest metagraphs correspond to very small fractions of the set of all networks. For the networks analyzed here ($N = 6$ genes, $M \approx 0.5\, N^2$ nonzero regulatory interactions, 8.6×10^{13} total networks) a metagraph with 4.1×10^6 networks (*left end* of the *horizontal axis*) contains only a fraction 4.7×10^{-8} of all networks. The median of the distribution shown is 1.2×10^7 (compared to 5.9×10^{10} for monofunctional networks).

networks is orders of magnitude smaller than that of viable monofunctional networks. This means that it would be very difficult to find a viable bifunctional network through a random search in the space of all possible networks, even for only moderately sized networks. The hatched bars indicate p^2, the square of the fraction of viable monofunctional networks, which we find to be an order-of-magnitude approximation for the mean fraction of bifunctional networks. Finally, we also note that although the fraction of viable networks may be tiny, their absolute number is still very large. For example, for networks with as few as six genes, on average there are 1.96×10^7 viable bifunctional networks, and 5.92×10^{10} viable monofunctional networks. These large numbers stem from the very large total number of networks (e.g., 8.59×10^{13} for the networks of $N = 6$ used in Fig. 2 a).

We now define a graph that will aid in answering the questions we posed earlier. Each node in this graph corresponds to a viable network. Two networks (nodes) in this graph are connected if they differ in the value of only one regulatory interaction (Fig. 1, b and c). We call this graph a metagraph—a graph of graphs—because its nodes are networks, which could themselves be represented as graphs. These nodes differ in their topology of regulatory interactions. Neighboring networks in the metagraph can arise from one another by genetic changes that affect only one regulatory interaction. In the biological evolution of network topology, this graph could be traversed through a series of small genetic changes, each of which affects only one regulatory interaction.

The above analysis regarding the fraction of viable networks makes a statement about the mean size of metagraphs. For example, if we say that there are 3×10^7 viable bifunctional networks, we mean that the metagraph of these networks comprises 3×10^7 nodes. However, the size of a metagraph can vary widely, depending on the actual gene expression state pairs $(\mathbf{S}^{(1)}(0), \mathbf{S}_\infty^{(1)})$ and $(\mathbf{S}^{(2)}(0), \mathbf{S}_\infty^{(2)})$. Fig. 2 b shows an example, based on an analysis, in which we generated 1000 bifunctional phenotypes (two state pairs) at random, as described in the Supplementary Material, and estimated the metagraph sizes for each. The figure demonstrates that there is a large dispersion in the sizes of the metagraphs, but also that even for small networks ($N = 6$), metagraphs are typically very large. Specifically, the median metagraph size for the networks shown in this figure is 1.2×10^7 (whereas the median metagraph size for monofunctional networks of the same size is 5.9×10^{10}).

A broad distribution of robustness

When studying robustness, the network features we focus on are a network's equilibrium gene expression pattern(s), which we generically denote by $\mathbf{S}_\infty$. Robustness to mutations corresponds to robustness of $\mathbf{S}_\infty$ to changes in regulatory interactions, that is, to changes in network topology. Specifically, we define mutational robustness R_μ as the fraction

of a network's neighbors that differ in only one regulatory interaction, and that are still on the metagraph. Robustness to noise corresponds to robustness of $\mathbf{S}_\infty$ to changes in the initial expression pattern $\mathbf{S}(0)$. Specifically, we use two complementary measures of robustness to noise.

The first of them is the probability $R_{\nu,1}$ that a change in one gene's expression state in the initial expression pattern $\mathbf{S}(0)$ leaves the network's equilibrium expression pattern $\mathbf{S}_\infty$ unchanged. The second measure is the fraction $R_{\nu,*}$ of genes whose expression needs to change, such that the probability of attaining the equilibrium state falls below 1/2. Because we have shown previously that robustness to mutations and to noise is correlated, we here focus on mutational robustness, and show only selected results for robustness to noise (13).

For bifunctional phenotypes with sizeable metagraphs, we asked whether the mutational robustness of viable networks has a broad distribution (13). Fig. 3 a shows the distribution of mutational robustness R_μ for a sample of 1042 viable bifunctional networks with $N = 12$ genes. The distribution of robustness is clearly broad, spanning a factor 25 ($0.027 \leq R_\mu \leq 0.69$; Fig. 3 d). Similarly broad distribution are seen for robustness to noise $R_{\nu,1}$ and $R_{\nu,*}$, as well as for different numbers of genes and regulatory interactions (not shown). The breadths of these distributions increase with increasing network sizes. For example, for networks with $N = 16$ genes, mutational robustness in a smaller sample of 586 networks varies by more than two orders of magnitude ($0.0067 \leq R_\mu \leq 0.81$).

In assessing robustness thus far, we required that a network maintains both equilibrium gene expression states upon mutational change. In other words, we require that both network functions are preserved. This is clearly a more stringent requirement than asking for only one of the functions to be preserved. The distributions of mutational robustness, if we require that only function 1 or function 2 are preserved, are shown in Fig. 3, b and c, respectively. As we have shown previously (13), these distributions are also broad. Not unexpectedly, the likelihood that a mutation preserves both functions is substantially lower than the likelihood that it preserves only one function (median values of R_μ 0.25 rather than 0.5, Fig. 3 d, which shows the medians; minima and maxima of the distributions shown in a–c).

A design rule for robust multifunctional networks

The broad distribution of robustness among viable networks raises the question whether there are some principles underlying robust network design. We address this question by extending a previous design rule for networks with only one input-target pair (13). Briefly, this previous rule required that for every nonzero regulatory interaction w_{ij}, $w_{ij} = S_{i,\infty} S_{j,\infty}$ for any gene j whose expression is the same in the initial and equilibrium state. For genes j that are not of this type, the rule assigns the weights of nonzero interactions w_{ij} so that the sum $\sum_{j, S_j(0) \neq S_{j,\infty}} w_{ij} S_j(0)$ is zero or close to zero for every i.
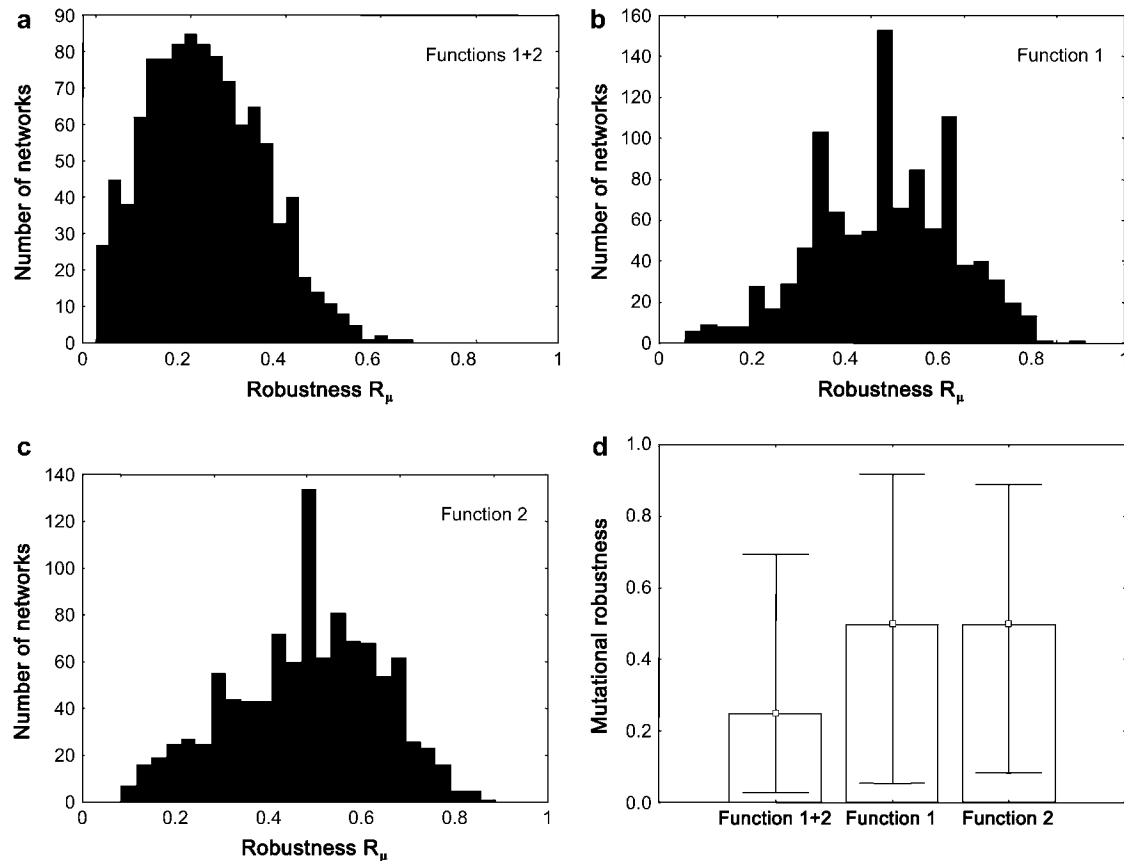
FIGURE 3 Robustness to mutations shows a broad distribution. Panels *a–c* show a histogram for the distribution of three different indicators of mutational robustness R_μ. In *a*, R_μ is defined as the fraction of a network's neighbors that preserve both functions, i.e., that attain $\mathbf{S}_\infty^{(1)}$ when presented with $\mathbf{S}^{(1)}(0)$ and that attain $\mathbf{S}_\infty^{(2)}(0)$. In *b* and *c*, respectively, R_μ is defined as the fraction of a network's neighbors that preserve function 1 and 2, respectively. (*d*) The medians (*bar*), as well as the maxima and minima (*whiskers*) of the distributions in *a–c*. All data are based on one random realization of two expression state pairs, and on a sample of 1042 viable bifunctional networks with $N = 12$ genes and $M \approx 0.25\, N^2$ nonzero regulatory interactions ($w_{ij} = \pm 1$).

(For sufficiently large N, choosing random values for these weights will achieve this goal.) This rule generally leads to rapid attainment of the equilibrium state from the initial state, and it is a sufficient criterion for high robustness.

For the case of multiple input-output pairs, we generalize this rule as follows. First, we apply the rule separately to each input output pair. Second, we average the matrices w thus obtained. This may result in matrices with too many regulatory interactions compared to the desired number. In a third step, we thus examine this list of interactions and keep only the M interactions that are largest in absolute value. Finally, for each regulatory interaction w_{ij}, we either take its sign to obtain a matrix of discrete regulatory interactions ($w_{ij} = \pm 1$), or we take a Gaussian random number of the same sign as w_{ij} to arrive at a matrix of continuous-valued regulatory interactions. Assuming that there exist viable networks for any given set of initial-target state pairs, this procedure is likely to produce the most robust of such networks. Our design rule shares important elements with a Hebb rule for storing information in artificial neural networks (26), an important difference being that biological networks show asymmetric regulatory interactions ($w_{ij} \neq w_{ji}$), which our rule can accommodate.

We next asked whether this prescription really produces highly robust networks. To this end, we defined an indicator (which we term Q, for network quality; see the Supplementary Material) of the extent to which the structure of an arbitrary network is similar to that prescribed by the design rule. Fig. 4 shows that mutational robustness is significantly correlated with Q, thus validating the design rule (Spearman's $s = 0.37$; $P < 10^{-17}$; Fig. 4 *a*). Robustness to noise is also significantly associated with Q; for instance, for the phenotype used in this figure, we find for $R_{\nu,1}$ that Spearman's $s = 0.48; P < 10^{-17}$; for $R_{\nu,*}$: Spearman's $s = 0.46; P < 10^{-17}$.

The metagraph of viable bifunctional networks is usually dominated by a giant component

Is the metagraph a connected graph? We first iterate an argument detailed elsewhere (13), which demonstrates that metagraph connectedness, if it is found, is not a trivial feature. Specifically, it does not hold for a ''random'' metagraph
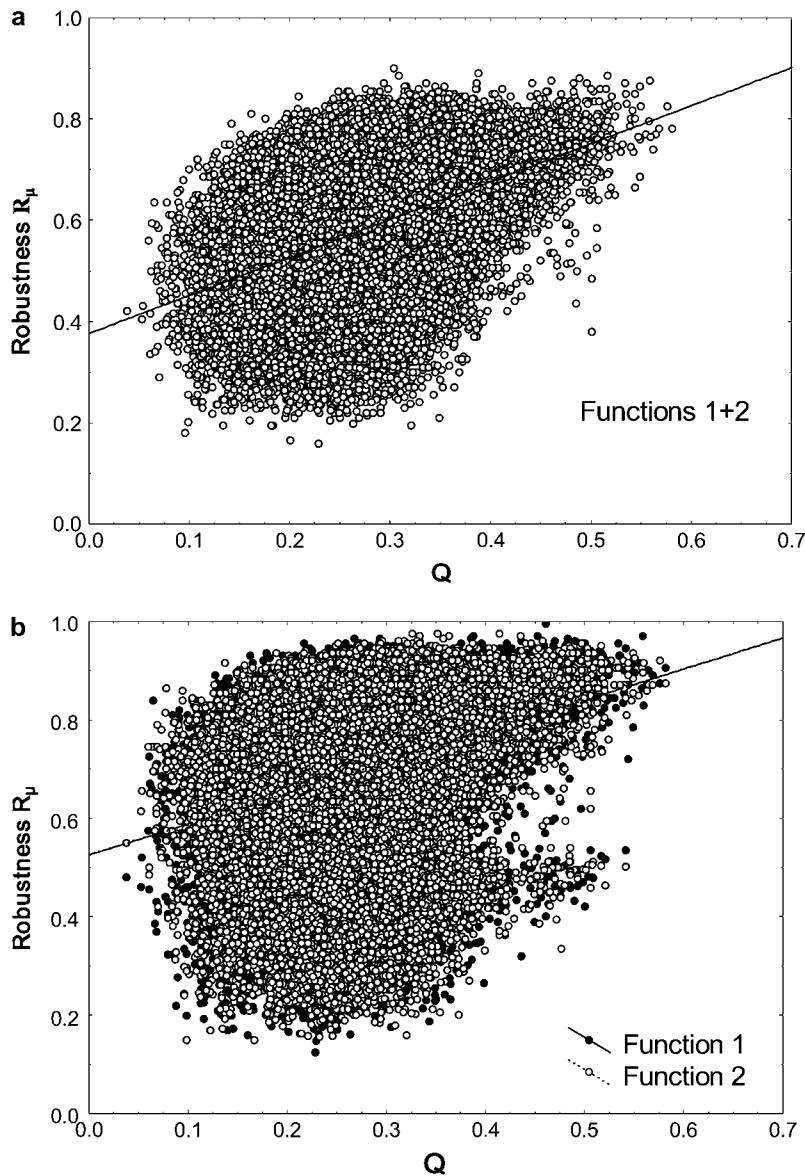
FIGURE 4 Network quality Q is associated with mutational robustness. (*a*) A scatter plot of Q (*horizontal axis*), an indicator of robust network design described in Methods (Supplementary Material), and mutational robustness R_{μ} of both functions of a bifunctional network (*vertical axis*) (Spearman's $s = 0.37$; $P < 10^{-17}$). (*b*) A scatter plot of Q (*horizontal axis*) against robustness R_{μ} to function 1 and 2, considered separately (Function 1: Spearman's $s = 0.28$; $P < 10^{-17}$; Function 2: Spearman's $s = 0.28$; $P < 10^{-17}$). The data sets are very similar and thus one of them (*solid circles*) is largely hidden behind the other. Two horizontal lines corresponding to linear regression lines are drawn, but because the lines are nearly identical, only one of them is visible. Data is based on randomly generated input-target pairs and a sample of 10^{5} viable bifunctional networks with $N = 20$ genes and $M \approx 0.25\ N^{2}$ nonzero regulatory interactions ($w_{ij} = \pm 1$). Significant positive associations are also observed with networks of different size and different numbers of regulatory interactions.

comprising the same number n_{v} of networks as the above metagraph of viable networks, where neighboring nodes (networks) differ in one regulatory interaction, but where the nodes need not be viable. Such a random metagraph consists mostly of isolated nodes, as we will now show. Let n be the total number of networks for a given number of genes and regulatory interactions. Consider an arbitrary node w of the random metagraph. It is easy to determine a lower bound for the probability that w is isolated in the random metagraph, i.e., that all the remaining $n_{v} - 1$ nodes in the random metagraph are distinct from w's K neighbors. This lower bound is

$$\left(1 - \frac{K}{n - n_{v} + 1}\right)^{n_{v}-1} \approx 1 - \frac{(n_{v} - 1)K}{n - n_{v} + 1} \approx 1.$$

The left approximation holds, because K is of order N^{2}, whereas the denominator is dominated by the total number of

networks n, which scales exponentially in N. In addition, $(n_{v} - 1)K/(n - n_{v} + 1) \ll 1$, because n_{v} is exponentially small compared to n, whereas K is no greater than N^{2}. Thus, the product $(n_{v} - 1)\ K$ divided by n is exponentially small.

In sum, the probability that an arbitrary network w in the random metagraph is isolated is very close to 1. It immediately follows that the average number of components of the random metagraph, given by n_{v} times the above probability, is only slightly smaller than the total number of networks n_{v}: only a negligible fraction of the nodes of the random metagraph are not isolated.

With this observation in mind, we numerically analyzed the connectivity of metagraphs comprising viable networks with two input-output pairs. Briefly, we estimated for a random sample of viable networks with two given input-output pairs the fraction of networks that lie in the same component

of the metagraph (see the Supplementary Material for details). As this fraction may depend on the input-output pairs, we repeated this approach for 100 different input-output pairs, which allowed us to collect statistics on the connectivity. We find that for a given N and c, the metagraph of viable networks is more often disconnected than when there is a single pair (13), but nevertheless a giant component dominates it when N increases. Table 1 shows statistical results of this procedure for small networks. For instance, the mean percentage of networks in the giant component increases from 63.3% to 82.9% as the number of genes increases from four to eight. Although we cannot generate similar statistics for networks much larger than these, the data suggest that for all but the smallest values of N, the great majority of bifunctional networks is contained in a giant connected component, as in the case of monofunctional networks.

Finally, we also asked whether networks near each other on the metagraph have similar robustness. If they do, then robustness changes smoothly on the metagraph and it could readily increase in a biased random walk (or through natural selection). If not, then the distribution of robustness on the metagraph shares properties with "rugged fitness landscapes" (27), where finding the near-global maximum of robustness would be very difficult. The question is best addressed by determining the autocorrelation function of robustness for a random walk of length L steps on a metagraph. This walk starts at some randomly chosen network on the metagraph. Denote by r_k the value of some observable (such as network robustness) at the kth step of the random walk. Then, the autocorrelation function $\rho(l)$ between two networks that are l steps apart is defined as

$$\rho(l) = \frac{\dfrac{1}{L - l - 1} \sum_{k=l+1}^{L} (r_k - \bar{r})(r_{k-l} - \bar{r})}{\dfrac{1}{L} \sum_{k=1}^{L} (r_k - \bar{r})^2}.$$

Fig. 5 shows the autocorrelation function $\rho(l)$ of mutational robustness R_μ, as well as for robustness $R_{\nu,1}$ and $R_{\nu,*}$ to noise, both for bifunctional networks (*upper panel*) and for monofunctional networks (*lower panel*). This function decays

exponentially in the lag l, but it is modestly large for small l. For example, for R_μ, $\rho(l = 1) = 0.56$ for bifunctional networks, and $\rho(l = 1) = 0.79$ for monofunctional networks. $\rho(l) > 0.25$ as long as $l < 10$ in the case of bifunctional networks, and as long as $l < 20$ in the case of monofunctional networks. These observations show that the metagraph is not very rugged with respect to mutational robustness.

No strong trade-offs between robustness in different functions

So far, we have shown that bifunctional networks have very large metagraphs in which the distribution of robustness is broad, and where most networks can be connected through single mutational changes. We now turn to the question whether there are trade-offs among different network functions with respect to robustness. That is, if a network has one function that is highly robust, does that mean that the other function has low robustness, and vice versa?

To get at this question, it is useful to take the following perspective. Consider only the first function (expression state pair), and call the metagraph formed by all networks that have this function M_1. Define analogously the metagraph M_2 for the second function. Next define the metagraph $M_{12} = M_1 \cap M_2$. The networks in M_{12} are networks that have both functions. Recall that the mutational robustness R_μ of a network with respect to one function is its degree k_1 (k_2) in M_1 (M_2) normalized to the interval $(0,1)$. Denote this indicator of robustness as R_μ^1 (R_μ^2). The mutational robustness with respect to both functions is its degree k_{12} in M_{12} normalized to $(0,1)$, which we will denote as R_μ^{12}. Fig. 6 *a* shows R_μ^{12} on the horizontal axis and R_μ^1 (R_μ^2) on the vertical axis. The upper-triangular shape of the plot is easily understood if one recalls that the fraction of a network's neighbors that carry out both functions cannot be greater than the fraction of neighbors that carry out only one function. Networks on the diagonal are networks whose degree in M_1 (M_2) is equal to their degree in M_{12}. There are few such networks, in line with our previous observation that M_1 and M_2 are much larger than M_{12}. The figure also shows that for any network with a given robustness with respect to both functions, there may be a broad distribution of robustness with respect to one or the other function, that is, there may be many networks at varying distances from the diagonal line.

Fig. 6 *b* plots R_μ^1 and R_μ^2 against each other for two given input-target pairs. Intriguingly, the figure shows no trade-off between the two measures of robustness, but a modest positive association (Spearman's $s = 0.19$; $P < 4.9 \times 10^{-10}$). This is not a fortuitous coincidence, resulting from the particular networks chosen for analysis. For example, among 10 identical analyses, using different randomly chosen gene expression state pairs (see the Supplementary Material), six analyses show a positive association that is significant at $P < 0.005$ (1042 networks). Four analyses show a nonsignificant positive association, and none shows a negative association.
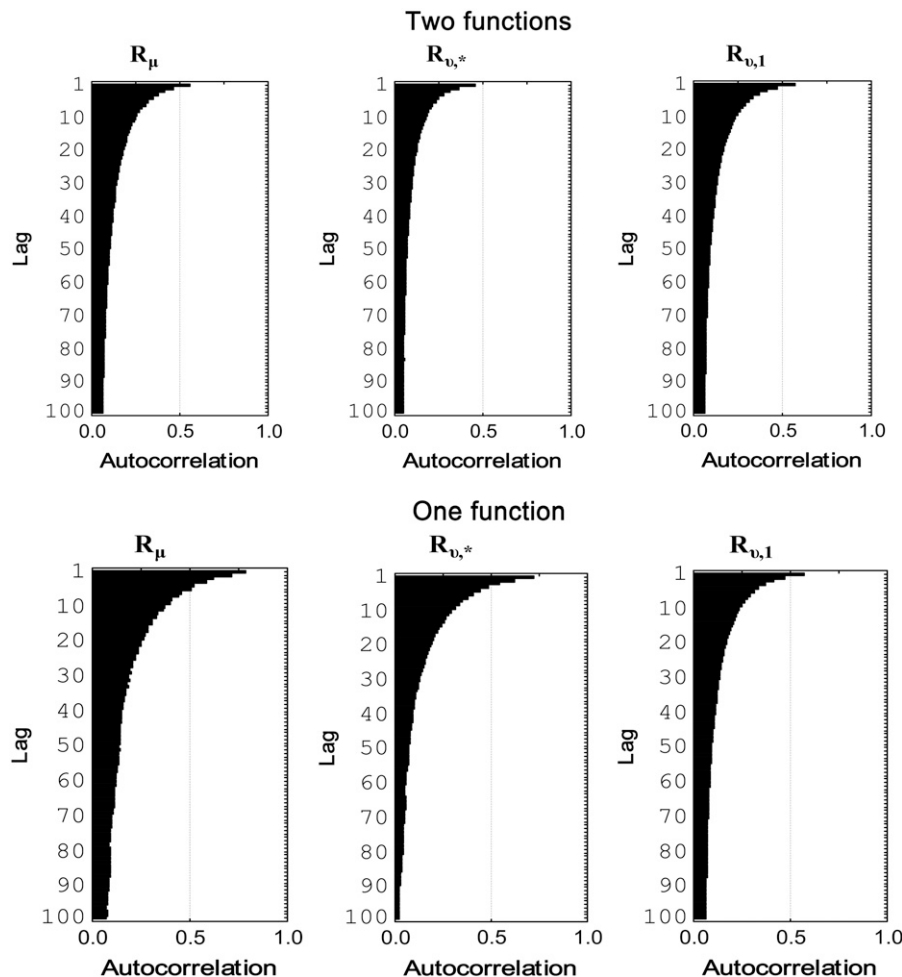
TABLE 1   Most networks in a metagraph are connected in a giant component

Network Size					
Genes	Mean	Mode	90th Percentile	Mean ± SE	$q*$
$N = 4$	63.3%	67%	93%	3.3	63
$N = 6$	69.1%	89%	94%	3.3	67
$N = 8$	82.9%	93%	98%	2.6	80

Table shows various statistics for the estimated fraction of networks contained in the giant component for bifunctional networks with $M \approx cN^2$ ($c = 0.5$) nonzero regulatory interactions ($w_{ij} = \pm 1$).

*q is the number such that for q% of the samples (each sample has two input-output pairs), the giant component of the metagraph contains at least q% of all networks.

FIGURE 5 Autocorrelation functions indicate that metagraphs are not rugged with respect to robustness. Each panel shows the autocorrelation function $\rho(l)$ as defined in the text (*horizontal axes*) for lag values up to $l = 100$ (*vertical axes*), and for various measures of robustness to mutations R_μ, as well as to noise $R_{\nu,*}$ and $R_{\nu,*}$. Results are shown for networks of $N = 12$ genes, $M \approx 0.25 \, N^2$ nonzero regulatory interactions ($w_{ij} = \pm 1$) and for random walks of $L = 9 \times 10^4$ steps.

If one pools data from 500 analyses with different input-target pairs, one also sees an overall positive association between R_μ^1 and R_μ^2 ($s = 0.17$; $P < 10^{-17}$). Networks with different numbers of genes show the same preponderance of positive associations.

This slightly positive association can be justified as follows. Consider a network that has both functions. Its mutational robustness R_μ^1 and R_μ^2 is proportional to the degree d_1 (d_2) the network has in M_1 (M_2). These degrees can be written as $d_1 = d_{12} + (d_1 - d_{12})$ and $d_2 = d_{12} + (d_2 - d_{12})$. In other words, d_{12} contributes to both R_μ^1 and R_μ^2. That this common contribution explains the positive association is shown by Fig. 6 c, which plots $(R_\mu^1 - R_\mu^{12})$ against $(R_\mu^2 - R_\mu^{12})$. Using this quantity, we eliminate the common contribution d_{12} from the analysis. The resulting association is then not positive but strongly negative (Spearman's $s = -0.68$; $P < 10^{-17}$).

We also observe a negative association if we bin data according to the value of R_μ^{12} and then determine the statistical association between R_μ^1 and R_μ^2 within each bin ($0 \le R_\mu^{12} < 0.25$: $s = -0.35$; $0.25 \le R_\mu^{12} < 0.5$: $s = -0.35$; $0.5 \le R_\mu^{12} < 0.75$: $s = -0.34$; $0.75 \le R_\mu^{12} \le 1$: $s = -0.26$; $P < 10^{-8}$ for each analysis.). Qualitatively, the same result is obtained

from a partial correlation analysis that estimates linear correlation coefficients between R_μ^1 and R_μ^2 while controlling for R_μ^{12} ($r = -0.77$; $P < 10^{-5}$).

Computational cost prevented us from carrying out much of our analysis for more than two expression state pairs. However, the qualitative finding that there is no trade-off in mutational robustness also holds in the limited number of analyses we have done for trifunctional networks, i.e., networks with three expression state pairs. For example, for networks with $N = 20$ ($c \approx 0.5$), we find only a very small association among $R_\mu^1 - R_\mu^2$, $R_\mu^1 - R_\mu^3$, and $R_\mu^2 - R_\mu^3$ ($s > -0.037$; $n = 10^5$). Partial correlation coefficients $R_\mu^i - R_\mu^j$, holding R_μ^k constant are even smaller ($-0.02 < s < 0.01$; $i \ne j \ne k$; $n = 10^5$). However, there is still a positive association between R_μ^{123} and R_μ^i, for $i = 1,2,3$ ($s > 0.49$), and a negative association between $(R_\mu^i - R_\mu^{123})$ and $(R_\mu^j - R_\mu^{123})$, where s < -0.24 for $i \ne j$. Thus, the main difference to the bifunctional case is that the moderately positive association between the robustness for each function disappears.

Although we reported most of our analysis above for networks with discrete regulatory interactions, we emphasize that qualitatively identical results hold for networks with
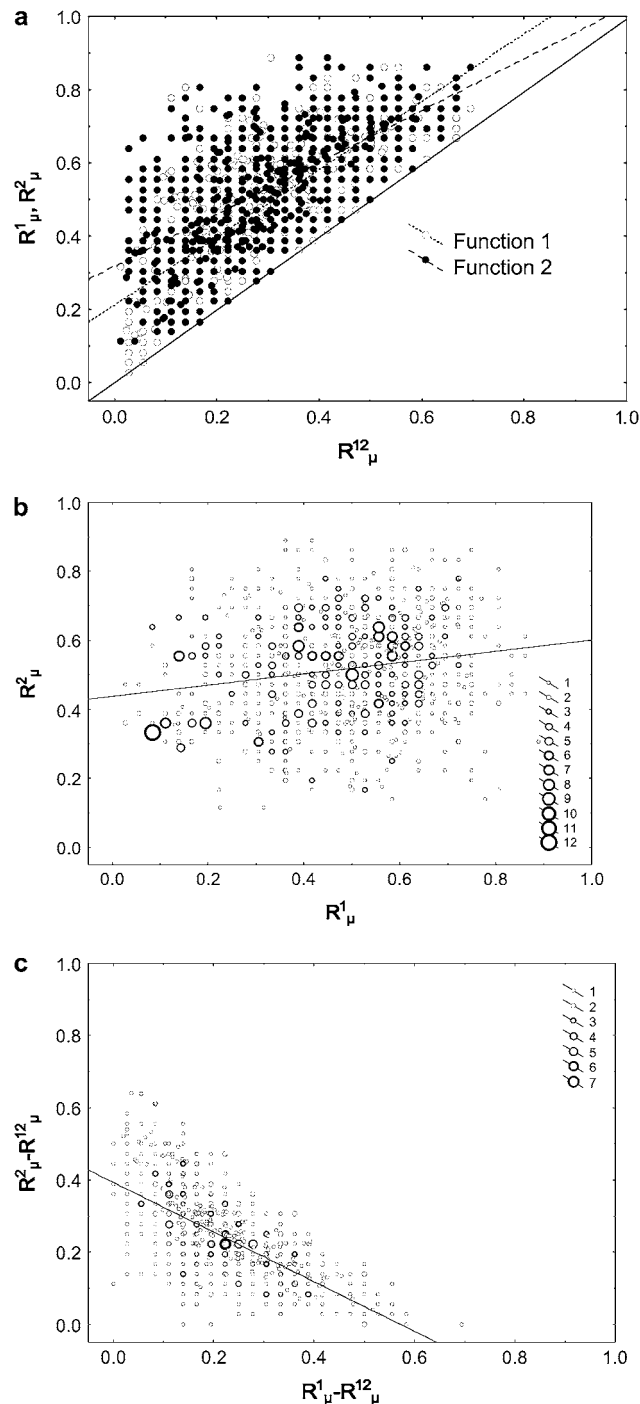
FIGURE 6 No negative association (trade-off) between different measures of robustness. (*a*) Scatter plot of R_μ with respect to both functions (*horizontal axis*) and R_μ with respect to only one function (*vertical axis*). See text for details. The solid diagonal line is the identity line, dashed lines indicate linear regressions. (*b*) Scatter plot of R_μ^1 versus R_μ^2. (Spearman's $s = 0.19$; $P < 4.9 \times 10^{-10}$). (*c*) Scatter plot of $(R_\mu^1 - R_\mu^{12})$ against $(R_\mu^2 - R_\mu^{12})$. (Spearman's $s = -0.68$; $P < \times 10^{-17}$). The solid lines in *b* and *c* indicate a linear regression. The circle sizes indicate the number of networks with the given robustness values, as indicated in the legend. (Circles of varying sizes have been omitted from panel *a* for clarity.) Data are based on a sample of at least 1000 viable bifunctional networks with $N = 12$ genes and $M \approx 0.25\,N^2$ nonzero regulatory interactions ($w_{ij} = \pm 1$).

continuous regulatory interactions. Fig. 7 *a* illustrates that R_μ^{12} is smaller than R_μ^1 and R_μ^2 also for networks with continuous regulatory interactions. Fig. 7 *b* shows that network quality, as defined above, still shows a positive association with robustness. Fig. 7 *c* shows that R_μ^1 and R_μ^2 are positively associated also for such networks.

DISCUSSION

In sum, we find that for regulatory networks with more than one function, the number of networks (topologies) that carry out all functions declines sharply with the number of functions. However, because the number of topologies carrying out one function is very large, there are still many bi-and trifunctional networks, even for the small network sizes we consider here.

In contrast to the constraints multifunctionality imposes on network architecture, we find no robustness trade-offs among functions. That is, if a network has one highly robust function, then other functions are not necessarily less robust. In our system, the maximally possible robustness in multifunctional networks tends to be lower than in monofunctional networks. Because most bifunctional networks are connected via the giant component of a metagraph, and because the autocorrelation function of random walks on this metagraph does not decay very rapidly, networks whose functions are all highly robust to mutations—within attainable limits—can readily evolve through small regulatory changes and gradual evolution. Although we focus for computational convenience on networks with discrete regulatory interactions, our main results also hold for continuously valued regulatory interactions (Fig. 7).

We note that a lack of a significant robustness, trade-off has recently also been reported for a completely different model (15). The networks studied in that work have only one function, but they contain highly conserved modules with clear subfunctions. Importantly, the overall robustness of the whole networks was positively correlated with the robustness of individual modules, and no robustness trade-off among the different subfunctions existed.

The concept of a metagraph is analogous to that of a "neutral set" or "neutral network" (28). In a neutral network, multiple RNA sequences that form the same secondary structure constitute the nodes of a graph. Two nodes are connected if they differ by one nucleotide. In a metagraph, multiple network topologies that have the same gene expression patterns $(\mathbf{S}(0), \mathbf{S}_\infty)$ form the nodes of the graph. Two nodes are connected if they differ in the sign of one regulatory interaction. We use the term metagraph (graph-of-graphs) because it contains a reminder that each of its nodes is itself a network that can be represented as a graph. For monofunctional networks, the existence of alternative topologies with different robustness that are connected in a metagraph has been shown for other systems, such as circadian oscillators (16).
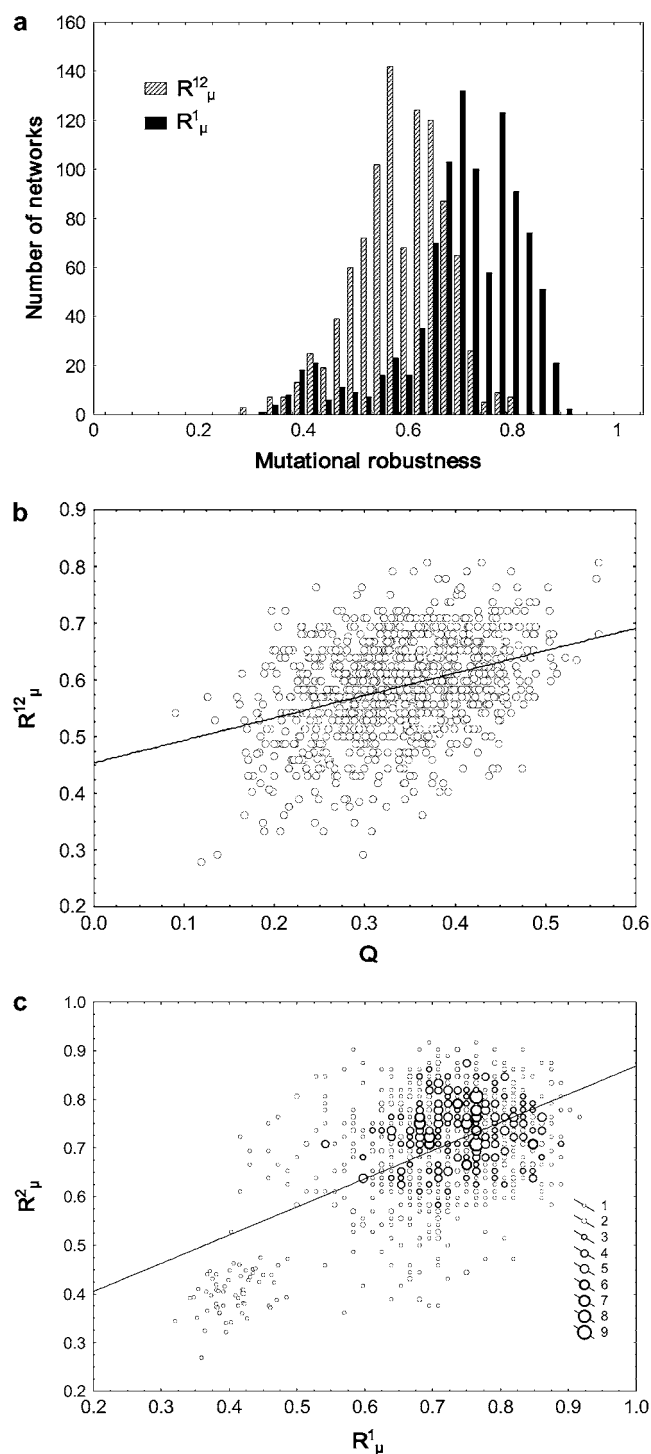
FIGURE 7 Networks with continuously valued regulatory interactions behave like discrete networks. (*a*) Distribution of R_μ^1 and R_μ^{12}, estimated for networks with continuously valued w_{ij}'s as explained in Methods (Supplementary Material (31,32)). (*b*) Positive association between Q and R_μ^{12} (Spearman's $s = 0.32$; $P < 10^{-17}$). (*c*) Positive association—no trade-off—between R_μ^1 and R_μ^2 (Spearman's $s = 0.33$; $P < 10^{-17}$). All data based on a sample of 1000 viable bifunctional networks with $N = 12$ genes and $M \approx 0.25\, N^2$ nonzero regulatory interactions with a continuous distribution.

The analogy between metagraphs and neutral networks has limitations. Whereas the structure of RNA sequence space is intrinsically discrete, regulatory gene interactions, however defined, can be continuously valued. To define a metagraph thus requires a discretization of the space of possible interactions by focusing on the signs of the interactions. We motivated such discretization and our focus on network topology in the introduction. Perhaps more important, however, is a second limitation of the analogy. Although some RNAs may have multiple, equally stable secondary structures, most biological RNA sequences adopt one well-defined RNA secondary structure (which may be a prerequisite for their biological function). In contrast, in the domain of networks, multiple stable gene expression states are the rule rather than the exception. The networks we study thus lend themselves ideally to the exploration of robustness trade-offs that do not have natural counterparts in RNA molecules.

We note that other reports of robustness trade-offs in biological systems, such as genome-scale cellular networks or physiological systems (29,30), use a fundamentally different notion of a trade-off. There, a system may have one function, but this function can experience common or random perturbation, as well as rare or targeted perturbations. Systems robust against common or random perturbations may be sensitive to rare or targeted perturbations, a phenomenon that could be viewed as a robustness trade-off. In contrast, we are here concerned with multifunctional systems, where all functions experience the same kinds of perturbations (mutations). This definitional difference means that our findings do not contradict earlier work. However, the relationship between these two kinds of trade-offs are worth exploring further.

We are acutely aware that the model we use is highly abstract, even though it may explain a wide variety of qualitative and quantitative information about transcriptional regulation networks and their evolution. (12,19–25). We use such a modeling approach because experimental observations that speak to the phenomenon of interest are sorely lacking. Not only that, it is not clear how these observations could be produced with available technology, as they would require the experimental analysis of thousands of network topologies, and systematic perturbations of each of them. Until the time that such technology becomes available, models such as this are needed to help shape our intuition. If the intuition we obtained here is correct, then architectural constraints are key features of multifunctional networks, but robustness trade-offs are not.

SUPPLEMENTARY MATERIAL

To view all of the supplemental files associated with this article, visit www.biophysj.org.

REFERENCES

1. von Dassow, G., E. Meir, E. Munro, and G. Odell. 2000. The segment polarity network is a robust development module. *Nature*. 406:188–192.

2. Alon, U., M. G. Surette, N. Barkai, and S. Leibler. 1999. Robustness in bacterial chemotaxis. *Nature*. 397:168–171.

3. Ingolia, N. T. 2004. Topology and robustness in the Drosophila segment polarity network. *PLoS Biol*. 2:805–815.

4. Espinosa-Soto, C., P. Padilla-Longoria, and E. R. Alvarez-Buylla. 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.

5. Stelling, J., E. D. Gilles, and F. J. Doyle. 2004. Robustness properties of circadian clock architectures. *Proc. Natl. Acad. Sci. USA*. 101:13210–13215.

6. Gonze, D., J. Halloy, and A. Goldbeter. 2002. Robustness of circadian rhythms with respect to molecular noise. *Proc. Natl. Acad. Sci. USA*. 99:673–678.

7. Freeman, M. 2000. Feedback control of intercellular signalling in development. *Nature*. 408:313–319.

8. Morohashi, M., A. E. Winn, M. T. Borisuk, H. Bolouri, J. Doyle, and H. Kitano. 2002. Robustness as a measure of plausibility in models of biochemical networks. *J. Theor. Biol*. 216:19–30.

9. Edwards, J. S., and B. O. Palsson. 2000. Robustness analysis of the *Escherichia coli* metabolic network. *Biotechnol. Prog*. 16:927–939.

10. Eldar, A., R. Dorfman, D. Weiss, H. Ashe, B. Shilo, and N. Barkai. 2002. Robustness of the BMP morphogen gradient in *Drosophila* embryonic patterning. *Nature*. 419:304–308.

11. Albert, R., and H. G. Othmer. 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.

12. Azevedo, R. B. R., R. Lohaus, S. Srinivasan, K. K. Dang, and C. L. Burch. 2006. Sexual reproduction selects for robustness and negative epistasis in artificial gene networks. *Nature*. 440:87–90.

13. Ciliberti, S., O. Martin, and A. Wagner. 2007. Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS. Comput. Biol*. 3:e15.

14. Reference deleted in proof.

15. Ma, W., L. Lai, Q. Ouyang, and C. Tang. 2006. Robustness and modular design of the Drosophila segment polarity network. *Molecular Systems Biology*. 2:70.

16. Wagner, A. 2005. Circuit topology and the evolution of robustness in two-gene circadian oscillators. *Proc. Natl. Acad. Sci. USA*. 102:11775–11780.

17. Carroll, S. B., J. K. Grenier, and S. D. Weatherbee. 2001. From DNA to Diversity. Molecular Genetics and the Evolution of Animal Design. Blackwell, Malden, MA.

18. Gilbert, S. F. 1997. Developmental Biology. Sinauer, Sunderland, MA.

19. Reinitz, J., and D. H. Sharp. 1995. Mechanism of eve stripe formation. *Mech. Dev*. 49:133–158.

20. Sharp, D. H., and J. Reinitz. 1998. Prediction of mutant expression patterns using gene circuits. *Biosystems*. 47:79–90.

21. Jaeger, J., S. Surkova, M. Blagov, H. Janssens, D. Kosman, K. Kozlov, Manu, E. Myasnikova, C. Vanario-Alonso, M. Samsonova, D. Sharp, and J. Reinitz. 2004. Dynamic control of positional information in the early *Drosophila* embryo. *Nature*. 430:368–371.

22. Mjolsness, E., D. H. Sharp, and J. Reinitz. 1991. A connectionist model of development. *J. Theor. Biol*. 152:429–453.

23. Wagner, A. 1996. Does evolutionary plasticity evolve? *Evolution Int. J. Org. Evolution*. 50:1008–1023.

24. Siegal, M., and A. Bergman. 2002. Waddington's canalization revisited: developmental stability and evolution. *Proc. Natl. Acad. Sci. USA*. 99:10528–10532.

25. Bergman, A., and M. Siegal. 2003. Evolutionary capacitance as a general feature of complex gene networks. *Nature*. 424:549–552.

26. Amit, D. J. 1989. Modeling Brain Function. The World of Attractor Neural Networks. Cambridge University Press, Cambridge, UK.

27. Kauffman, S. A. 1993. The Origins of Order. Oxford University Press, New York.

28. Schuster, P., W. Fontana, P. Stadler, and I. Hofacker. 1994. From sequences to shapes and back—a case-study in RNA secondary structures. *Proc. R. Soc. Lond. B. Biol. Sci*. 255:279–284.

29. Kitano, H., K. Oda, T. Kimura, Y. Matsuoka, M. Csete, J. Doyle, and M. Muramatsu. 2004. Metabolic syndrome and robustness tradeoffs. *Diabetes*. 53:S6–S15.

30. Albert, R., H. Jeong, and A. L. Barabasi. 2000. Error and attack tolerance of complex networks. *Nature*. 406:378–382.

31. Li, F. T., T. Long, Y. Lu, Q. Ouyang, and C. Tang. 2004. The yeast cell-cycle network is robustly designed. *Proc. Natl. Acad. Sci. USA*. 101:4781–4786.

32. Toulouse, G. 1977. Theory of the frustration effect in spin glasses: I. *Communications on Physics*. 2:115–119.