Science Signaling
AAAS

Rewiring of Transcriptional Regulatory Networks: Hierarchy, Rather Than Connectivity, Better Reflects the Importance of Regulators

Nitin Bhardwaj, Philip M. Kim and Mark B. Gerstein (2 November 2010)
Science Signaling **3** (146), ra79. [DOI: 10.1126/scisignal.2001014]

SYSTEMS BIOLOGY

Rewiring of Transcriptional Regulatory Networks: Hierarchy, Rather than Connectivity, Better Reflects the Importance of Regulators

Nitin Bhardwaj,[1] Philip M. Kim,[2,3,4,5]* Mark B. Gerstein[1,6,7]*

(Published 2 November 2010; Volume 3 Issue 146 ra79)

Network connectivity has been related to essentiality: Highly connected proteins (hubs) are more important for cell growth and survival. Although this is intuitively reasonable, another way to assess the role of a regulator is to assign it to a level within a "chain-of-command" hierarchy. Here, we analyzed the effects of network rewiring events on transcriptional regulatory hierarchies in two species. First, we superimposed the phenotypic effects of tampering with specific genes and their regulatory connections directly onto the hierarchies. To study second-order effects, which involved changes in the level of regulators within the hierarchy upon deletions or insertions of other regulators or connections, we reconstructed modified hierarchies. We found that rewiring events that affected upper levels had a more marked effect on cell proliferation rate and survival than did those involving lower levels. Moreover, we showed that the hierarchical level and type of change better reflected the phenotypic effect of rewiring than did the number of changes. We also investigated other features connected to the importance of upper-level regulators: In particular, relative to lower-level regulators, upper-level regulators exhibited a greater range of expression values across species, had fewer functionally redundant copies, and had a shorter half-life. Overall, our analysis shows that broadly constructed hierarchies may better reflect the importance of regulators for cell growth than classifications based on the number of connections (hubbiness).

INTRODUCTION

The transcriptional regulatory machinery consists of a set of transcription factors (TFs) responsible for regulating the spatial and temporal expression of genes in any given genome (*1*, *2*). Such functional linkages can be visualized as a network with edges pointing from TFs to their target genes. This regulatory network, by analogy to various other kinds of networks, has been the focus of efforts to understand the system-level features of biological organization in the postgenomic age (*3*, *4*). Other kinds of biological networks, such as protein-protein interactions (*5*, *6*) and genetic interactions (*7*), along with other more commonplace ones, such as social interactions, the Internet, and food webs (*4*, *8*–*10*), have also been studied to obtain more intuition into network organization. Analysis of these diverse networks has revealed various properties, such as small-world property (*11*), scale-free nature (*10*, *12*), modularity (*13*–*15*), and disassortativeness (*16*). Many of these properties are believed to increase the tolerance of networks to errors and deletions (*10*, *16*–*18*).

The analysis of network robustness to perturbations, such as the deletion and addition of nodes and edges, from a dynamic perspective has emerged as a new frontier. There have been systematic studies to estimate the effects of tinkering with networks (*19*–*23*). In a biological network, these perturbations represent such changes as deletion or overexpression of genes (nodes), or the addition or removal of regulatory connections (edges) between nodes. For example, Isalan *et al.* determined the tolerance of *Escherichia coli* to the addition of new regulatory edges between a subset of regulators by measuring deviations in mutant proliferation rates from the wild-type rates (*21*). The accurate determination of proliferation rate and fraction of surviving colonies upon deletion of every single gene has been accomplished for *E. coli* and *Saccharomyces cerevisiae* (*19*, *20*). For yeast, synthetic genetic arrays have identified genetic interactions by measuring double mutant fitness through the effect of the mutations on proliferation rates (*7*, *22*, *24*). These studies have identified the relative importance of genes for cell survival and proliferation by measuring the phenotypic effect of network rewiring.

In agreement with intuition, the effects of gene perturbations on cell survival and proliferation fall on a continuum (Fig. 1). On one extreme are genes for which single knockout mutants are not viable (the cell dies). On the other extreme, there are genes that can be deleted without any loss of viability or any defect in proliferation. In the middle, there are genes that when perturbed produce a range of effects; removal of certain genes will affect only the expression of a few other genes, whereas others are synthetically lethal (simultaneous removal of certain nonessential gene pairs causes cell death), and the loss of yet others causes a conditional phenotype (where the cell grows only under certain conditions). Thus, the "essentiality" of genes should be considered as a continuous metric that has been termed "conditional benefit" or "marginal essentiality" (*25*, *26*). A given gene's position along this continuum can be estimated from the deviation of a deletion mutant from the wild type in terms of proliferation, the number of colonies with correct structure, or cell survival in a specific environment.

[1]Program in Computational Biology and Bioinformatics, Yale University, Bass 426, 266 Whitney Avenue, New Haven, CT 06520, USA. [2]Terrence Donnelly Centre for Cellular and Biomolecular Research, Department of Medical Research, University of Toronto, Toronto, Ontario, Canada M5S 3E1. [3]Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario, Canada M5S 3E1. [4]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada M5S 3E1. [5]Department of Computer Science, University of Toronto, Toronto, Ontario, Canada M5S 3E1. [6]Department of Molecular Biophysics and Biochemistry, Yale University, Bass 432, 266 Whitney Avenue, New Haven, CT 06520, USA. [7]Department of Computer Science, Yale University, Bass 432, 266 Whitney Avenue, New Haven, CT 06520, USA.
*To whom correspondence should be addressed. E-mail: pi@kimlab.org (P.M.K.); mark.gerstein@yale.edu (M.B.G.)

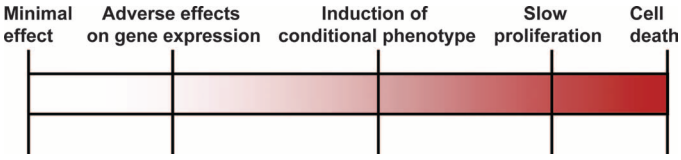

Fig. 1. Phenotypic effect spectrum. Range of possible phenotypic effects arising from perturbation of certain genes or network tinkering.

Some previous studies suggest that connectivity is related to essentiality—that is, hubs are often essential parts of the network (*18*, *26*–*28*). Although at first this may be intuitively reasonable, connectivity distribution is only one of the numerous properties of networks. A highly connected protein can be more indispensable for cell survival, but stipulating this as a requirement or definite signature of essentiality may be an oversimplification. Whether the connectivity-essentiality relationship is true or whether it might be due to a bias resulting from incomplete data remains an open issue (*29*). An alternative way to gauge the importance of regulators would be to assign them to "chain-of-command" hierarchical levels, which are analogous to the positions held by employees within a corporate or governmental hierarchy. The representation of networks into intuitive hierarchical structures also provides the potential for greater insight into the organizational principles governing regulatory networks. For example, these networks may be arranged into pyramidal hierarchies that share many common characteristics with corporate and government hierarchies (*30*). Arrangement into hierarchies has also been used to obtain insights into the evolutionary history of these networks (*31*) and to identify modules in the network by a top-down approach—starting from the whole network structure and identifying subsystems or modules by network decomposition (*32*). The representation of networks as hierarchies has also been used to identify topological units (subhierarchies, called "origons") that originate at the top of a hierarchy and are responsive to specific environmental signals (*33*). The intermediate levels of these regulatory hierarchies segregate into distinct densely interlinked subnetworks and integrate signals under specific conditions (*34*). Furthermore, within a regulatory hierarchy of TFs, top-level TFs are relatively abundant, long-lived, and noisy in terms of their abundance in a cell population, whereas middle-level TFs are involved in more biological processes and have a larger number of targets (*35*). Although there has been a limited analysis of the essentiality of regulators in different hierarchical levels in the context of a static network without any perturbations (*30*), examination of the phenotypic effects of network rewiring should provide additional insight into the relationship between hierarchical rearrangement and cell proliferation and survival in a dynamic context.

We report the phenotypic effects of tinkering with *E. coli* and *S. cerevisiae* transcriptional regulatory networks organized into a hierarchical arrangement in the context of static networks, as well as in networks that undergo rewiring in response to the introduced changes. We mapped the phenotypic effects of tinkering with different levels to elucidate which levels were more tolerant to perturbations. To examine how tinkering influenced phenotype in a dynamic context, we rebuilt the hierarchies in response to the introduced perturbations and correlated the phenotypic effects to the locations of changes in modified hierarchies. We studied two kinds of perturbations: addition of new edges to the existing network and deletion of nodes. We primarily focused on genes that are in the middle of the "phenotypic effect spectrum"—those genes that, when perturbed, do not cause cell death but have other intermediate effects.

RESULTS

Construction of regulatory hierarchies

We used a breadth-first search (BFS) algorithm to build hierarchies in a bottom-up approach (*30*), which is reasonable because all of the regulatory interactions point downward in a chain-of-command fashion. With the BFS algorithm, the TFs that do not regulate any other TF were assigned to the bottom level (level 1) along with TFs that only regulate themselves (only autoregulators). Beginning from each level 1 TF, we searched for their immediate regulators, which were assigned to the level immediately above (level 2). To handle loops, we placed the targets of TFs and unassigned targets in the same level, implying that loops will only be placed in the same level. Continuing in the same way, we built a breadth-first tree for both *E. coli* (list S1) and *S. cerevisiae* (list S2) consisting of four levels in a pyramidal shape (with level 4 at the top and level 1 at the bottom). These four levels are composed of regulators only; all nonregulators are placed in the lowest level (level 0, below level 1). This approach unambiguously placed the TFs into different levels, with all regulatory edges pointing downward, or horizontally for loops (Fig. 2). This means that no lower-level TF regulates an upper-level TF. This type of hierarchy is similar to any social or corporate hierarchy, in which the chain of command points downward. With the BFS approach, the resulting hierarchies were pyramidal, which are intuitive and allow analogies to be drawn with other more commonplace hierarchies from corporate or government settings (*30*). This also makes sense from a biological context, because there are a handful of master regulators at the top that initiate most of the downstream processes. These top-level regulators control a larger number of middle managers, which, in turn, control an even larger set of lower-level regulators. At the bottom of the hierarchy is the pool of nonregulators that perform their stand-alone function(s). In addition, most previous studies related to regulatory hierarchies have resulted in pyramidal or near-pyramidal arrangements (*21*, *30*, *33*–*39*). The BFS approach is only one way to build a hierarchy; there are many other methods (*30*, *35*).

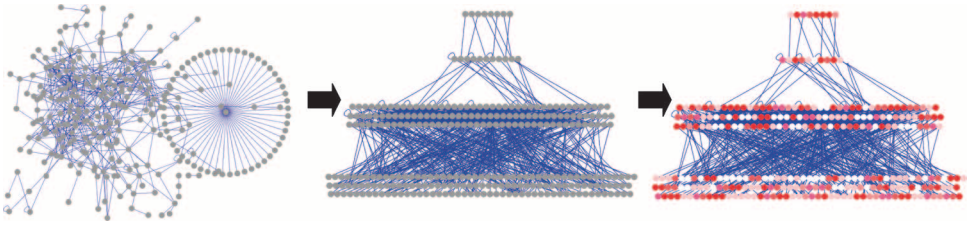

Fig. 2. Analysis of first-order effects. First, the network is arranged into a hierarchy, with the regulatory edges pointing downward only. Next, the phenotypic effects of tinkering with various nodes are overlaid onto the hierarchy. The color of the nodes (scaling from white to red) indicates the effect of tinkering on cell growth: Tampering with white nodes has a minimal effect, whereas red nodes are the genes that, upon deletion, affect the cell growth adversely. This figure serves as a schematic representation only and does not reflect the actual data.

First-order effect: Phenotypic effects of perturbations in the absence of network rewiring (static analysis)

We mapped the phenotypic effects of perturbing the TFs (their deletion or introducing new regulatory edges into them) of different levels of the hierarchy (Fig. 2). In the *E. coli* and *S. cerevisiae* TF networks, we found that perturbations affecting the

higher-level TFs had a greater effect on cell proliferation rate or survival than did perturbations to lower-level TFs (Fig. 3). We determined the Pearson correlation coefficient (PCC) and then calculated *P* values by generating the null model, using Fisher's transformation with the null hypothesis that there is no correlation (PCC = 0). PCC ranges between −1 and 1, indicating that there is a strong negative and positive correlation between the paired samples, respectively, and PCC = 0 indicates no correlation.

For *E. coli*, deletion of upper-level TFs (data S1) or addition of regulatory edges into the upper levels (data S2) resulted in a slower growth rate or higher fraction of colonies with incorrect structure (Fig. 3A; PCC = 0.39 with $P = 7.9 \times 10^{-4}$ for the left plot, PCC = 0.34 with $P = 8.2 \times 10^{-3}$ for the right plot). Similarly, in the yeast network, deletion of higher-level TFs (data S3 and S4) led to a lower cell fitness (as measured by proliferation rate) than did the deletion of lower-level TFs (Fig. 3B; PCC = 0.28 with $P = 1.6 \times 10^{-4}$ for the left plot, PCC = 0.36 with $P = 9.2 \times 10^{-4}$ for the right plot).

We built similar hierarchies based on a specific subset of the data: operon-based hierarchies for *E. coli* (fig. S1 and list S5) and condition-specific hierarchies for *S. cerevisiae* (fig. S2 and list S6). We grouped genes according to their operon organization in *E. coli* [data from RegulonDB (*40*)] and rebuilt the hierarchy, using the same algorithm as that described above. Because the data set was smaller (one operon represented all the component genes), there were only three levels in this hierarchy (fig. S1). Similarly, for *S. cerevisiae*, we used subnetworks that were responsive only under certain conditions to build smaller hierarchies with only three levels using the same method (*41*). We used the subset of the transcriptional networks that is active under four conditions: cell cycle, sporulation, diauxic shift, and DNA damage (list S6). For many of these subhierarchies (except the sporulation and cell cycle hierarchies), we found that the deletion of upper-level nodes had a larger effect on cell growth and survival. Similar results were obtained when subnetworks reported by individual studies were analyzed (figs. S4 and S5). These observations in both species further reinforce our results above.

Apart from the removal of nodes and the addition and deletion of edges, we also studied another kind of perturbation: changing the strength of interactions between regulators and targets through overexpression of regulatory nodes. Sopko *et al.* studied the phenotypic effects of gene overexpression systematically covering >80% of the *S. cerevisiae* genome and compiled a list of ~15% of overexpressed genes (called "toxic genes") that reduced cell growth rate through colony defects (*23*). We mapped these toxic genes onto the regulatory hierarchy and found that the top level had the highest fraction and the bottom level had the lowest fraction of toxic genes, with the middle levels having intermediate fractions (fig. S3). This result agrees with our above observation that tinkering with upper-level regulators had larger effects on cell growth and survival.

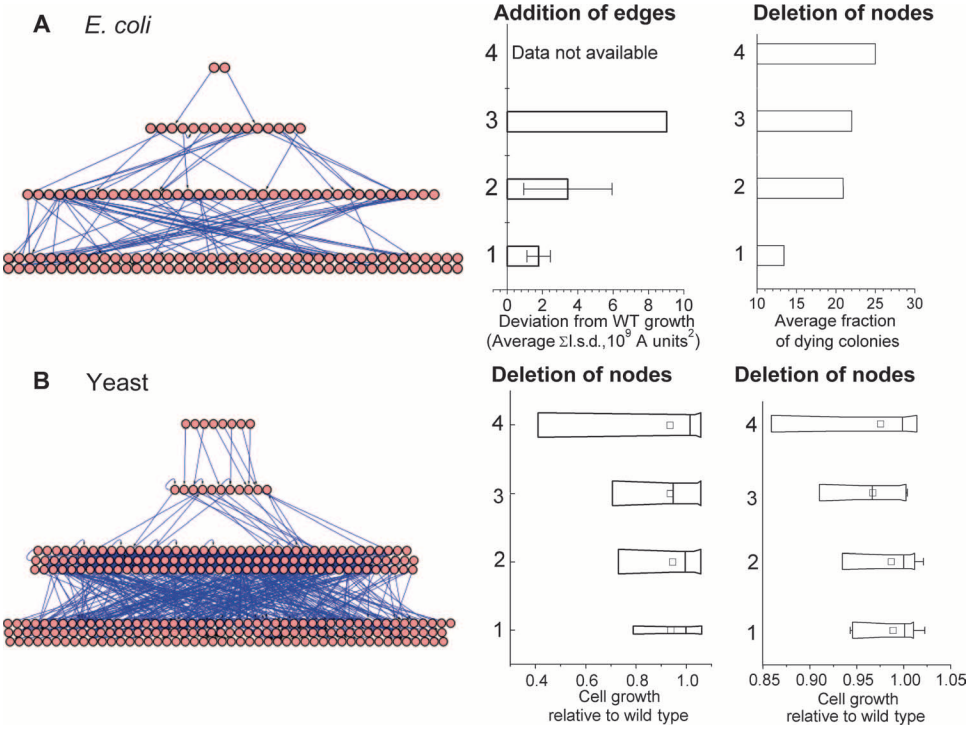

Fig. 3. Phenotypic effect of node tampering from various levels in *E. coli* and *S. cerevisiae*. Level 4 (at the top of the hierarchy) has the fewest nodes and level 1 (at the bottom) has the most. (**A**) For the *E. coli* network, the left graph plots the deviation from wild-type (WT) growth upon addition of new edges to nodes of different levels [as measured by Isalan *et al.* (*21*)], and the right graph displays the fraction of dying colonies upon deletion of genes [data from Baba *et al.* (*19*)]. The average ∑I.s.d. in the left graph measures the "least-squares difference" between mutant and WT growth rates as monitored by measuring the absorbance at 600 nm (A_{600}) or simply *A*. (**B**) For the *S. cerevisiae* network, both graphs plot cell growth relative to WT upon deletion of nodes. Phenotypic data were obtained from Deutschbauer *et al.* (*20*) (left plot) and Costanzo *et al.* (*7*) (right plot).

Second-order effect: Changes in the hierarchy upon network rewiring

To study the second-order effects of tampering with the network, we allowed the hierarchy to rearrange in response to rewiring of the network, such as the addition or deletion of edges or the removal of nodes. For example, in many cases, the deletion of nodes from the middle level sometimes resulted in the migration of one or more TFs to a lower level (Fig. 4A) or an upper level (Fig. 4B). Similarly, the addition of a new edge from a lower-level TF to an upper-level one sometimes resulted in the reassignment of one or more TFs to lower levels (Fig. 4C). For every rewiring of the network, we rebuilt the corresponding modified hierarchies and compared them to the wild-type hierarchy (tables S2 to S5 and data S1 to S4). We computed the changes in the hierarchy by counting the absolute number of regulators that changed their levels in the rearranged hierarchy. The extent of rewiring (the number of nodes deleted or edges added) did not correlate directly with the number of changes in the hierarchy; the deletion of one node may lead to any number of changes in the hierarchy (including zero), and the addition of one or more edges may result in an unchanged hierarchy (tables S2 to S5 and data S1 to S4; http://info.gersteinlab.org/Rewiring_hierarchy).

We found that the number of changes did not correlate well with the phenotypic effect either (Fig. 5); there were no con-

sistent trends between the number of genes that changed their levels and the phenotypic effects of these changes. For example, for *S. cerevisiae*, cell proliferation rates from both node-deletion studies (*7*, *20*) were poorly correlated with the number of changes in the hierarchy (Fig. 5, C and D). For *E. coli*, the highest deviation from wild-type proliferation occurred when there was only one change. Upon further inspection of these network changes in *E. coli*, we found that all of these changes resulted from the addition of new regulatory edges into *rpoE*, encoding a TF in level 3 of the *E. coli* regulatory hierarchy, and more than 75% of these changes involved reassignment of *rpoE* within the hierarchy. Thus, we hypothesized that the location of changes within the hierarchy, rather than the absolute number of changes, might be more closely related to the phenotypic effect of rewiring.

We plotted the phenotypic effect of each kind of statistically significant change (a change that had five or more instances) that occurred in the hierarchy upon rewiring. Higher-level changes had a greater effect on cell proliferation and survival than did those in the lower levels (Fig. 6). For example, the changes when a TF moved from level 4 or 3 to level 1 had a greater effect on cell proliferation and survival than the effect of a TF moving from level 2 to level 1. This suggests that the location of the change correlates better with the phenotypic effect.

Properties of upper-level regulators

A previous study has shown that upper-level deletions lead to larger changes in gene expression and affect a higher number of genes (*30*). We further analyzed the different levels of the hierarchy for other characteristics that might rationalize their relative importance in terms of phenotypic effects. First, we investigated the extent of expression divergence of regulators from different levels. Expression divergence quantifies the difference in patterns of expression profiles of homologous (similar) genes in closely related species. It is often caused by the adaptive evolution of regulatory sequences and frequently results in different phenotypes. Thus, a higher level of divergence of gene indicates its role in phenotypic variations. Using data from a previous study on yeast (*42*), we mapped the expression divergence onto the hierarchy and found that upper-level regulators display a higher degree of divergence (Fig. 7A).

Using a large-scale genetic interaction map from a previous study, we also studied the patterns of genetic interactions between and within different levels (*7*). Genetic interactions occur between genes for which the phenotypic effect of simultaneous deletion is much stronger than would be expected from the two single deletions, indicating that the two genes compensate for each other in the event of single gene deletion (*43*). We define "genetic interaction propensity" to quantify the tendency to have genetic interactions between two different genes from the same or different levels ($L_i - L_j$) as $(I_{ij} \times 100)/(|L_i| \cdot |L_j|)$ when $L_i \neq L_j$, and $(I_{ij} \times 2 \times 100)/[|L_i| \cdot (|L_j| - 1)]$ when $L_i = L_j$, where $|L_a|$ is the size of level L_a (the number of regulators), and I_{ij} is the number of interactions between the gene pairs from L_i and L_j. In other words, it is the percentage ratio of

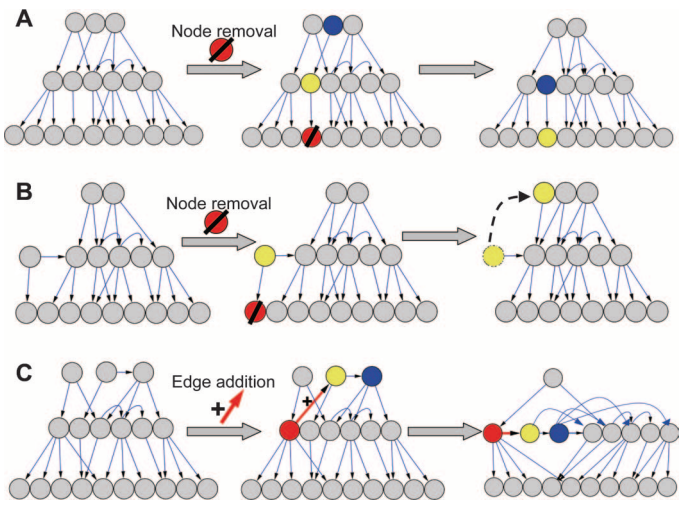

Fig. 4. Change in a hierarchy upon the deletion of nodes and the addition of new edges. (**A**) When the only target (red node) of a certain node (yellow node) is deleted, it (yellow) may change its levels, because it occupied a higher level only by virtue of its regulation of a lower-level gene. Once its target is deleted, it slides down (along with its regulator, in blue). (**B**) Node deletion may result in the migration of some genes up the hierarchy. When the red node is deleted, its regulator (yellow), which had occupied the middle level only by virtue of its regulation of the deleted node, moves up the hierarchy, because its only remaining target is in the middle level. (**C**) When a new edge is added (in red) from a lower-level gene (red node) to an upper-level gene (in yellow), the upper-level gene moves to the lower level, because regulatory edges from lower- to higher-level genes are not allowed. In certain cases, this condition requires more than one node to change level. For example, the blue node also moves down because it is also regulated by the yellow node.

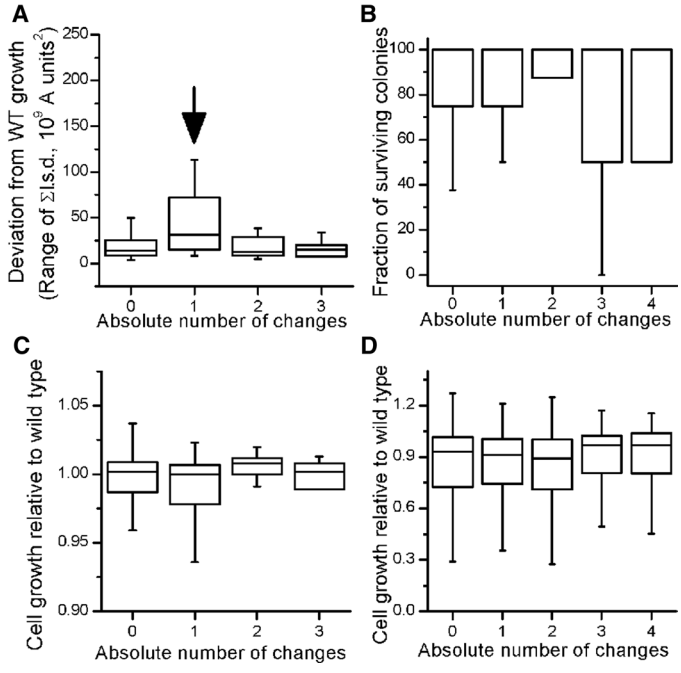

Fig. 5. Phenotypic effect does not correlate with the absolute number of changes in the two regulatory hierarchies of either *E. coli* or *S. cerevisiae*. Distribution of various cell fitness parameters is plotted for different numbers of changes in the hierarchy. (**A**) ∑I.s.d., a measure of deviation of *E. coli* mutant cell growth from WT values, was obtained from Isalan *et al.* (*21*). There is a large variability in cell growth when exactly one change occurs in the hierarchy (indicated by the arrow). All those changes correspond to the addition of new edges to *rpoE*. (**B**) The fraction of surviving *E. coli* colonies, as measured by Baba *et al.* (*19*), is plotted against the number of changes in the hierarchy. (**C** and **D**) *S. cerevisiae* growth relative to WT, as measured by Deutschbauer *et al.* (*20*) and Costanzo *et al.* (*7*), respectively, is plotted for different numbers of changes in the hierarchy.

the actual number of genetic interactions to the total number of possible interactions, normalized by the size of the two levels. We found that, although there were some interactions between and within lower-level regulators (from levels 1 and 2), there were no genetic interactions between upper-level regulators (from levels 3 and 4; Fig. 7B). This shows that these regulators have no backup genes within the same level, suggesting that in the event of deletion of these nodes, there are no upper-level regulators that can compensate for their loss.

We also analyzed the expression of regulators in different hierarchical levels. These quantities were obtained from previous studies in yeast that measured absolute abundance of transcripts in normal growth conditions (*44*) and messenger RNA (mRNA) half-life (*45*). We found that regulators higher in the hierarchy had a lower amount of expression (less abundant transcripts) and a shorter half-time of the corresponding mRNAs (Fig. 7, C and D). This shows that, relative to the lower-level regulators, the expression of upper-level genes is under tighter control, presumably because of their higher importance. Hence, an artificial perturbation to these regulators has a more marked effect on cell growth and survival than does perturbation of the lower-level ones.

DISCUSSION

We have shown that the phenotypic effects of rewiring in two transcriptional networks are well reflected in the position of TFs within the hierarchy, both in a static picture as well as upon network rewiring and rearrangement. In a static context, our results show that, whereas changes in the lower-level TFs had little effect on proliferation, tampering with nodes in higher levels had a greater effect on cell proliferation and survival. Upon rearranging the hierarchies in response to rewiring events, we showed that upper-level changes had more influence on cell fitness than did changes in lower levels. The results were consistent for two different species and for different kinds of rewiring events. We also showed that similar results were obtained for two condition-specific hierarchies in *S. cerevisiae* and operon-based hierarchies in *E. coli*. Both of these types of hierarchies are derived from a subset of the entire network. Finally, we observed that upper-level TFs had greater influence when smaller subsets of the regulatory network based on different individual studies were analyzed (figs. S4 and S5). Consistent results obtained for a smaller subset shows that our results are unlikely to be affected by the incompleteness of the data, an issue often addressed by repeating the analysis with a subset of the data. A previous study reported that lower levels in regulatory hierarchies in *E. coli* and *S. cerevisiae* had higher fractions of essential genes (*30*) than did upper levels. Although that observation was pertinent only to essential genes (defined as those that result in cell death when deleted), in this study, we examined genes in the intermediate region on the phenotypic effect spectrum (Fig. 1); their deletion does not result in cell death but affects the rate of cell proliferation, gene expression, conditional fitness of the cell, or some combination thereof.

In our analysis, the placement of regulators into different levels was determined not by the number of regulatory interactions with other nodes (in- and out-degrees) but by the other regulators that they control. Level position is thus a global, rather than a local, property of the network. Previous work has shown that there is no direct correlation between the local connectivity of a TF (its degree) and its position in the hierarchy (*30*). We calculated the number of direct and indirect targets (the ones that might be affected downstream, such as the targets of targets) and found that this number was not directly related to the position of a TF in a hierarchy (figs. S6 and S7). This fact, combined with our result that upper-level changes in the hierarchy had a larger effect on cell proliferation and survival, suggests that the relative importance of a TF is determined not just by its local connectivity but also by the other TFs it regulates. For example, *dnaA*, which encodes a protein that initiates chromosomal replication, has only nine direct targets, but its deletion reduces the cell survival rate to 10%. However, this gene has more than 200 indirect targets. This suggests that a "cascade effect" might occur such that the effect of tinkering with an upper-level gene is propagated downstream to the lower levels through all its direct and indirect targets, amplifying further

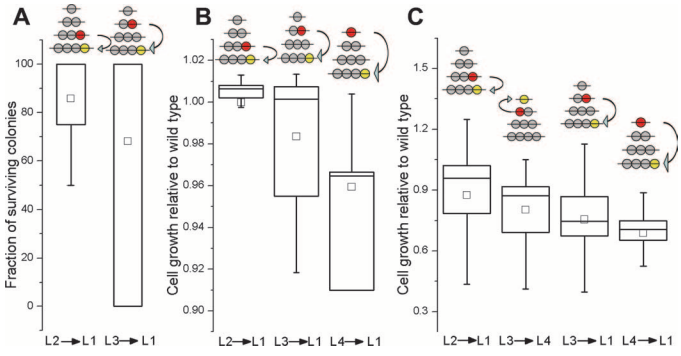

Fig. 6. Phenotypic effect plotted as a function of the type of change in the hierarchy. (**A**) For the *E. coli* regulatory hierarchy, the fraction of surviving colonies [as measured by Baba *et al.* (*19*)] is plotted as a function of the type of changes. (**B** and **C**) For the *S. cerevisiae* hierarchy, cell growth relative to WT [as measured by Deutschbauer *et al.* (*20*) and Costanzo *et al.* (*7*), respectively] is plotted against different kinds of changes in the hierarchy. Perturbation of the higher-level regulators markedly affects cell growth and survival.

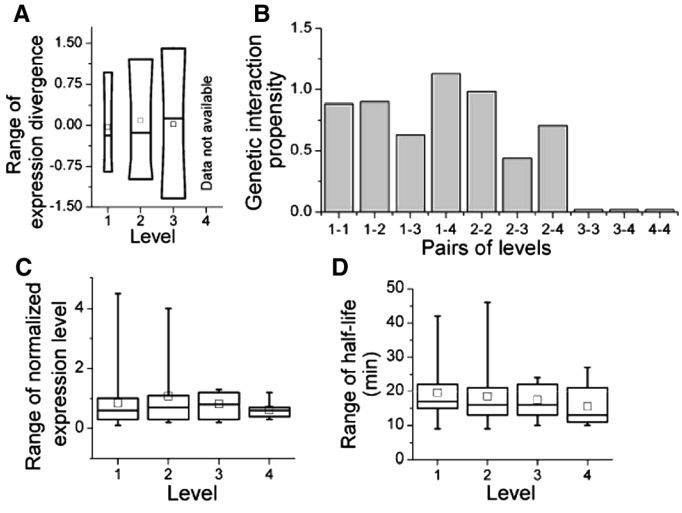

Fig. 7. Properties of regulators from different levels. (**A**) Range of expression divergence for different levels in yeast regulatory hierarchies. A higher value indicates that the genes display higher expression divergence across species. The width of the boxes is proportional to the SD of the data. (**B**) Genetic interaction propensity between different levels and within the same levels. Level pairs are indicated on the *x* axis (for example, the bar corresponding to 1-2 indicates the propensity between levels 1 and 2). (**C**) Normalized gene expression. (**D**) Half-life of the corresponding mRNA, in minutes.

with each level. However, the cascade effect does not always explain the relative influence of a gene on cell survival. For example, although the transcriptional activator encoded by *modE*, which is a component of the molybdenum transport operon *modABC* in *E. coli*, has 35 direct targets and 73 indirect targets, its deletion reduces the cell survival rate to only 88%. This shows that connectivity alone is not sufficient, and a more diverse set of properties must be integrated to better assay the relative importance of regulators for cell growth.

We hope that the findings presented in this study can be related to (and reveal more about) other networks and hierarchies for which this kind of dynamic testing is not feasible. Conventional mining studies of social networks and Web graphs have been limited to observational analysis, such as determining the connectivity distribution and the network diameter, because it is infeasible to rigorously test these types of networks by the experimental analysis, such as rewiring and deletion or addition of nodes. Because this kind of dynamic testing is possible in biological networks, consideration of analogies between biological regulatory networks and social hierarchies may reveal fresh insights into network behavior in general. The architecture of TF regulatory networks can be thought of as a hierarchical corporation with various levels of management. Just as in biological hierarchies, upper-level officials or managers in government or corporate settings have more responsibility and a larger impact than do lower-level officials. As such, upper-level changes in these social hierarchies tend to be more influential than would be lower-level changes. Such analogies between cellular and social hierarchies may aid our understanding of the architecture and dynamics of regulation and the relative importance of different regulators.

MATERIALS AND METHODS

We focus on transcriptional regulation in two well-characterized species, *E. coli* and *S. cerevisiae*. For *E. coli*, the transcriptional regulatory network was obtained from RegulonDB (as of July 2009) (*40*) (list S3). For *S. cerevisiae*, the data were obtained from various biochemical and genetic experiments (*1*, *46*–*48*) (list S4) as used in many previous studies (*37*, *41*, *49*–*51*). The networks (lists S3 and S4), their size and properties (table S1), and the hierarchical assignments (lists S1 and S2) are detailed in the Supplementary Materials and also at the following Web site (http://info.gersteinlab.org/Rewiring_hierarchy).

Phenotypic data quantifying the effect of tinkering with the networks were obtained from previous work (*7*, *19*–*21*). In these studies, individual nodes (genes) were deleted (for either *E. coli* or *S. cerevisiae*) with single-gene knockouts, such as genome-wide fitness profiling (*19*, *20*), and double-gene knockout experiments, such as synthetic genetic arrays for *S. cerevisiae* (*7*). In addition, for *E. coli*, new edges were introduced to the wild-type network by artificially constructing different combinations of promoters with different TFs or σ factor genes (*21*). In each case, the data were used to determine the effect of network rewiring on cell fitness (data S1 to S4 and tables S2 to S5).

SUPPLEMENTARY MATERIALS

www.sciencesignaling.org/cgi/content/full/3/146/ra79/DC1
Text
Fig. S1. Operon-based hierarchy in *E. coli*.
Fig. S2. Condition-specific hierarchies in yeast.
Fig. S3. Effects of overexpression, a means of increasing the strength of regulatory interactions.
Fig. S4. Phenotypic effect of deletion of nodes from various levels in the yeast regulatory subnetworks coming from individual studies.
Fig. S5. Phenotypic effect of deletion of nodes from various levels in yeast regulatory subnetworks coming from individual studies.
Fig. S6. Indirect targets and number of targets versus position in the hierarchy for *E. coli* and *S. cerevisiae*.
Fig. S7. Phenotypic effect versus out-degree.
Table S1. Properties of the regulatory networks used.
Table S2. Description of data S1.
Table S3. Description of data S2.
Table S4. Description of data S3.
Table S5. Description of data S4.
List S1. *E. coli* transcriptional regulatory network hierarchy.
List S2. Yeast transcriptional regulatory network hierarchy.
List S3. *E. coli* transcriptional regulatory network.
List S4. *S. cerevisiae* transcriptional regulatory network.
List S5. Regulatory network based on grouping a subset of the genes in *E. coli* by operon.
List S6. Regulatory network based on specific conditions in *S. cerevisiae*.
Data S1 to S4. The rewiring events and effects of gene deletion or edge addition to *E. coli* and *S. cerevisiae* transcriptional networks.
References

REFERENCES AND NOTES

1. T. I. Lee, N. J. Rinaldi, F. Robert, D. T. Odom, Z. Bar-Joseph, G. K. Gerber, N. M. Hannett, C. T. Harbison, C. M. Thompson, I. Simon, J. Zeitlinger, E. G. Jennings, H. L. Murray, D. B. Gordon, B. Ren, J. J. Wyrick, J. B. Tagne, T. L. Volkert, E. Fraenkel, D. K. Gifford, R. A. Young, Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**, 799–804 (2002).
2. E. Pérez-Rueda, J. Collado-Vides, The repertoire of DNA-binding transcriptional regulators in *Escherichia coli* K-12. *Nucleic Acids Res.* **28**, 1838–1847 (2000).
3. X. Zhu, M. Gerstein, M. Snyder, Getting connected: Analysis and principles of biological networks. *Genes Dev.* **21**, 1010–1024 (2007).
4. S. Wuchty, E. Ravasz, A.-L. Barabási, in *Complex Systems Science in Biomedicine*, T. S. Deisboeck, J. Y. Kresh, Eds. (Springer, Berlin, 2006).
5. T. Ito, T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, Y. Sakaki, A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4569–4574 (2001).
6. P. Uetz, L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson, J. R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, J. M. Rothberg, A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* **403**, 623–627 (2000).
7. M. Costanzo, A. Baryshnikova, J. Bellay, Y. Kim, E. D. Spear, C. S. Sevier, H. Ding, J. L. Koh, K. Toufighi, S. Mostafavi, J. Prinz, R. P. St Onge, B. VanderSluis, T. Makhnevych, F. J. Vizeacoumar, S. Alizadeh, S. Bahr, R. L. Brost, Y. Chen, M. Cokol, R. Deshpande, Z. Li, Z. Y. Lin, W. Liang, M. Marback, J. Paw, B. J. San Luis, E. Shuteriqi, A. H. Tong, N. van Dyk, I. M. Wallace, J. A. Whitney, M. T. Weirauch, G. Zhong, H. Zhu, W. A. Houry, M. Brudno, S. Ragibizadeh, B. Papp, C. Pal, F. P. Roth, G. Giaever, C. Nislow, O. G. Troyanskaya, H. Bussey, G. D. Bader, A. C. Gingras, Q. D. Morris, P. M. Kim, C. A. Kaiser, C. L. Myers, B. J. Andrews, C. Boone, The genetic landscape of a cell. *Science* **327**, 425–431 (2010).
8. R. Albert, A.-L. Barabási, Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97 (2002).
9. R. Albert, H. Jeong, A.-L. Barabási, Internet: Diameter of the World-Wide Web. *Nature* **401**, 130–131 (1999).
10. R. Albert, H. Jeong, A. L. Barabási, Error and attack tolerance of complex networks. *Nature* **406**, 378–382 (2000).
11. D. J. Watts, S. H. Strogatz, Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998).
12. H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, A. L. Barabási, The large-scale organization of metabolic networks. *Nature* **407**, 651–654 (2000).
13. J. D. Han, N. Bertin, T. Hao, D. S. Goldberg, G. F. Berriz, L. V. Zhang, D. Dupuy, A. J. Walhout, M. E. Cusick, F. P. Roth, M. Vidal, Evidence for dynamically organized modularity in the yeast protein–protein interaction network. *Nature* **430**, 88–93 (2004).
14. J. Ihmels, G. Friedlander, S. Bergmann, O. Sarig, Y. Ziv, N. Barkai, Revealing modular organization in the yeast transcriptional network. *Nat. Genet.* **31**, 370–377 (2002).
15. E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, A. L. Barabási, Hierarchical organization of modularity in metabolic networks. *Science* **297**, 1551–1555 (2002).
16. S. Maslov, K. Sneppen, Specificity and stability in topology of protein networks. *Science* **296**, 910–913 (2002).
17. S. Ciliberti, O. C. Martin, A. Wagner, Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Comput. Biol.* **3**, e15 (2007).
18. H. Jeong, S. P. Mason, A. L. Barabási, Z. N. Oltvai, Lethality and centrality in protein networks. *Nature* **411**, 41–42 (2001).
19. T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, H. Mori, Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).

20. A. M. Deutschbauer, D. F. Jaramillo, M. Proctor, J. Kumm, M. E. Hillenmeyer, R. W. Davis, C. Nislow, G. Giaever, Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics* **169**, 1915–1925 (2005).
21. M. Isalan, C. Lemerle, K. Michalodimitrakis, C. Horn, P. Beltrao, E. Raineri, M. Garriga-Canut, L. Serrano, Evolvability and hierarchy in rewired bacterial gene networks. *Nature* **452**, 840–845 (2008).
22. A. H. Tong, M. Evangelista, A. B. Parsons, H. Xu, G. D. Bader, N. Pagé, M. Robinson, S. Raghibizadeh, C. W. Hogue, H. Bussey, B. Andrews, M. Tyers, C. Boone, Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* **294**, 2364–2368 (2001).
23. R. Sopko, D. Huang, N. Preston, G. Chua, B. Papp, K. Kafadar, M. Snyder, S. G. Oliver, M. Cyert, T. R. Hughes, C. Boone, B. Andrews, Mapping pathways and phenotypes by systematic gene overexpression. *Mol. Cell* **21**, 319–330 (2006).
24. A. H. Tong, G. Lesage, G. D. Bader, H. Ding, H. Xu, X. Xin, J. Young, G. F. Berriz, R. L. Brost, M. Chang, Y. Chen, X. Cheng, G. Chua, H. Friesen, D. S. Goldberg, J. Haynes, C. Humphries, G. He, S. Hussein, L. Ke, N. Krogan, Z. Li, J. N. Levinson, H. Lu, P. Menard, C. Munyana, A. B. Parsons, O. Ryan, R. Tonikian, T. Roberts, A. M. Sdicu, J. Shapiro, B. Sheikh, B. Suter, S. L. Wong, L. V. Zhang, H. Zhu, C. G. Burd, S. Munro, C. Sander, J. Rine, J. Greenblatt, M. Peter, A. Bretscher, G. Bell, F. P. Roth, G. W. Brown, B. Andrews, H. Bussey, C. Boone, Global mapping of the yeast genetic interaction network. *Science* **303**, 808–813 (2004).
25. J. W. Thatcher, J. M. Shaw, W. J. Dickinson, Marginal fitness contributions of nonessential genes in yeast. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 253–257 (1998).
26. H. Yu, D. Greenbaum, H. Xin Lu, X. Zhu, M. Gerstein, Genomic analysis of essentiality within protein networks. *Trends Genet.* **20**, 227–231 (2004).
27. M. W. Hahn, A. D. Kern, Comparative genomics of centrality and essentiality in three eukaryotic protein-interaction networks. *Mol. Biol. Evol.* **22**, 803–806 (2005).
28. N. N. Batada, L. D. Hurst, M. Tyers, Evolutionary and physiological importance of hub proteins. *PLoS Comput. Biol.* **2**, e88 (2006).
29. S. Coulomb, M. Bauer, D. Bernard, M. C. Marsolier-Kergoat, Gene essentiality and the topology of protein interaction networks. *Proc. Biol. Sci.* **272**, 1721–1725 (2005).
30. H. Yu, M. Gerstein, Genomic analysis of the hierarchical structure of regulatory networks. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14724–14731 (2006).
31. M. Cosentino Lagomarsino, P. Jona, B. Bassetti, H. Isambert, Hierarchy and feedback in the evolution of the *Escherichia coli* transcription network. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 5516–5520 (2007).
32. H. W. Ma, J. Buer, A. P. Zeng, Hierarchical structure and modules in the *Escherichia coli* transcriptional regulatory network revealed by a new top-down approach. *BMC Bioinformatics* **5**, 199 (2004).
33. G. Balázsi, A. L. Barabási, Z. N. Oltvai, Topological units of environmental signal processing in the transcriptional regulatory network of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7841–7846 (2005).
34. I. J. Farkas, C. Wu, C. Chennubhotla, I. Bahar, Z. N. Oltvai, Topological basis of signal integration in the transcriptional-regulatory network of the yeast, *Saccharomyces cerevisiae*. *BMC Bioinformatics* **7**, 478 (2006).
35. R. Jothi, S. Balaji, A. Wuster, J. A. Grochow, J. Gsponer, T. M. Przytycka, L. Aravind, M. M. Babu, Genomic analysis reveals a tight link between transcription factor dynamics and regulatory network architecture. *Mol. Syst. Biol.* **5**, 294 (2009).
36. G. Balázsi, A. P. Heath, L. Shi, M. L. Gennaro, The temporal response of the Mycobacterium tuberculosis gene regulatory network during growth arrest. *Mol. Syst. Biol.* **4**, 225 (2008).
37. N. Bhardwaj, K. K. Yan, M. B. Gerstein, Analysis of diverse regulatory networks in a hierarchical context shows consistent tendencies for collaboration in the middle levels. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6841–6846 (2010).
38. H. W. Ma, B. Kumar, U. Ditges, F. Gunzer, J. Buer, A. P. Zeng, An extended transcriptional regulatory network of *Escherichia coli* and analysis of its hierarchical structure and network motifs. *Nucleic Acids Res.* **32**, 6643–6649 (2004).
39. K. K. Yan, G. Fang, N. Bhardwaj, R. P. Alexander, M. Gerstein, Comparing genomes to computer operating systems in terms of the topology and evolution of their regulatory control networks. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 9186–9191 (2010).
40. S. Gama-Castro, V. Jiménez-Jacinto, M. Peralta-Gil, A. Santos-Zavaleta, M. I. Peñaloza-Spinola, B. Contreras-Moreira, J. Segura-Salazar, L. Muñiz-Rascado, I. Martínez-Flores, H. Salgado, C. Bonavides-Martínez, C. Abreu-Goodger, C. Rodríguez-Penagos, J. Miranda-Ríos, E. Morett, E. Merino, A. M. Huerta, L. Treviño-Quintanilla, J. Collado-Vides, RegulonDB (version 6.0): Gene regulation model of *Escherichia coli* K-12 beyond transcription, active (experimental) annotated promoters and Textpresso navigation. *Nucleic Acids Res.* **36**, D120–D124 (2008).
41. N. M. Luscombe, M. M. Babu, H. Yu, M. Snyder, S. A. Teichmann, M. Gerstein, Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature* **431**, 308–312 (2004).
42. I. Tirosh, A. Weinberger, M. Carmi, N. Barkai, A genetic signature of interspecies variations in gene expression. *Nat. Genet.* **38**, 830–834 (2006).
43. J. Ihmels, S. R. Collins, M. Schuldiner, N. J. Krogan, J. S. Weissman, Backup without redundancy: Genetic interactions reveal the cost of duplicate gene loss. *Mol. Syst. Biol.* **3**, 86 (2007).
44. S. Ghaemmaghami, W. K. Huh, K. Bower, R. W. Howson, A. Belle, N. Dephoure, E. K. O'Shea, J. S. Weissman, Global analysis of protein expression in yeast. *Nature* **425**, 737–741 (2003).
45. Y. Wang, C. L. Liu, J. D. Storey, R. J. Tibshirani, D. Herschlag, P. O. Brown, Precision and functional specificity in mRNA decay. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5860–5865 (2002).
46. C. T. Harbison, D. B. Gordon, T. I. Lee, N. J. Rinaldi, K. D. Macisaac, T. W. Danford, N. M. Hannett, J. B. Tagne, D. B. Reynolds, J. Yoo, E. G. Jennings, J. Zeitlinger, D. K. Pokholok, M. Kellis, P. A. Rolfe, K. T. Takusagawa, E. S. Lander, D. K. Gifford, E. Fraenkel, R. A. Young, Transcriptional regulatory code of a eukaryotic genome. *Nature* **431**, 99–104 (2004).
47. C. E. Horak, N. M. Luscombe, J. Qian, P. Bertone, S. Piccirrillo, M. Gerstein, M. Snyder, Complex transcriptional circuitry at the G1/S transition in *Saccharomyces cerevisiae*. *Genes Dev.* **16**, 3017–3033 (2002).
48. V. V. Svetlov, T. G. Cooper, Review: Compilation and characteristics of dedicated transcription factors in *Saccharomyces cerevisiae*. *Yeast* **11**, 1439–1484 (1995).
49. S. Balaji, M. M. Babu, L. M. Iyer, N. M. Luscombe, L. Aravind, Comprehensive analysis of combinatorial regulation using the transcriptional regulatory network of yeast. *J. Mol. Biol.* **360**, 213–227 (2006).
50. S. Balaji, L. M. Iyer, L. Aravind, M. M. Babu, Uncovering a hidden distributed architecture behind scale-free transcriptional regulatory networks. *J. Mol. Biol.* **360**, 204–212 (2006).
51. S. A. Teichmann, M. M. Babu, Gene regulatory network growth by duplication. *Nat. Genet.* **36**, 492–496 (2004).

52. **Acknowledgments:** We thank H. Yu for various beneficial discussions. **Funding:** M.B.G. acknowledges the support from the NIH and from the AL Williams Professorship funds. **Author contributions:** N.B., P.M.K., and M.B.G. designed the research; N.B. performed the research; P.M.K. contributed new reagents and analytic tools; N.B. analyzed the data; and N.B., P.M.K., and M.B.G. wrote the paper. **Competing interests:** The authors declare that they have no competing interests.

Submitted 19 March 2010
Accepted 14 October 2010
Final Publication 2 November 2010
10.1126/scisignal.2001014
Citation: N. Bhardwaj, P. M. Kim, M. B. Gerstein, Rewiring of transcriptional regulatory networks: Hierarchy, rather than connectivity, better reflects the importance of regulators. *Sci. Signal.* **3**, ra79 (2010).