Global genetic rewiring during compensatory evolution in the yeast polarity network

Enzo Kingma[a] **and Liedewij Laan**[a,1]

[a]Department of Bionanoscience, Kavli Institute of Nanoscience Delft, Delft University of Technology, Delft 2629 HZ, the Netherlands.; [1]To whom correspondence should be addressed. E-mail: L.Laan@tudelft.nl

Functional defects resulting from deleterious mutations can often be restored during evolution by compensatory mutations elsewhere in the genome. Importantly, this process can generate the genetic diversity seen in networks regulating the same biological function in different species. How the options for compensatory evolution depend on the molecular interactions underlying these functions is currently unclear. In this study, we investigate how gene deletions compensating for a defect in the polarity pathway of *Saccharomyces cerevisiae* impact the fitness landscape. Using a transposon mutagenesis screen, we demonstrate that gene fitness has changed on a genome-wide scale in the compensated strain. An analysis of the functional associations between the affected genes reveals that compensation impacts cellular processes that have no clear connection to cell polarity. Moreover, genes belonging to the same process tend to show the same direction of gene fitness change, indicating that compensation rewires the fitness contribution of cellular processes rather than of individual genes. In conclusion, our results strongly suggest that functional overlap between modules and the interconnectedness of the molecular interaction network play major roles in mediating compensatory evolution.

1. Introduction

Cells must perform a series of complex tasks to stay alive and replicate. The successful execution of these tasks depends on a dense network of proteins, DNA, RNA and other small molecules that physically interact. For various cellular functions, research in model organisms has unraveled the order and timescale of these interactions. Important examples of elucidated pathways are the endocytocic (1, 2), cytokinesis (3) and cell polarity pathways (4). Results obtained from model organisms are often extrapolated to other biological systems based on the assumption that the molecular mechanisms underlying the same function are shared among different species. While molecular conservation has been demonstrated in several cases, it is becoming increasingly clear that components essential for a function in one species may be absent in another species that performs the same function (5–8). Thus, pathway composition is flexible during evolution, even when the same biological functionality is retained.

One of the possible mechanisms that can drive the emergence of different pathways performing the same cellular function is compensatory evolution. During compensatory evolution, the fitness effects of deleterious mutations, often in the form of gene loss or loss-of-function mutations, are alleviated by the acquisition of secondary mutations elsewhere in the genome. These secondary mutations effectively restore the perturbed pathway without reverting the original defect, thereby exploiting degenerate mechanisms to perform the same function. How the architecture of the molecular interaction network within the cell provides opportunities for genetic compensation of functional defects is unclear (9, 10).

A prominent view is that molecular interaction networks are organized into dynamic modules (11, 12). Here, a module regulates a single physiological function and consists of a group of proteins that interact extensively with each other, but sparsely with the rest of the network. Proteins may belong to different modules at different points in time, making module composition dynamic (12, 13). Importantly, the modular organization of the interaction network has been reasoned to improve cellular evolvability (11, 14). Constructing a larger network from smaller, modular, building blocks that interact sparsely would limit the pleiotropic effects of mutations, allowing each function to be optimized independently during evolution. In line with this role for modularity during adaptation, laboratory evolution experiments have shown that compensatory mutations preferentially arise in genes that are functionally related to the gene that was initially perturbed (15, 16). However, for modules to truly evolve independently, compensatory mutations must not alter the genotype-fitness relationship of other modules. Whether the effects of compensatory mutations on the genotype-fitness landscape are indeed only local is currently unexplored. In addition, doubts have arisen regarding the relevance and benefits of modules in network evolution due to findings from models, computer simulations, and phylogenetic analyses (17–19).

An appealing model system to study compensatory evolution is the comparatively well-studied cell polarization pathway (20–22). Cell polarization remains functionally conserved across the diverse branches of the tree of life (23–25). However, components essential for polarization in one species can lack orthologs in other species (5), suggesting that compensatory evolution occurs naturally within this pathway. As in most eukaryotes, cell polarization in *Saccharomyces cerevisiae* is regulated by the cycling of the protein Cdc42 between a cytosolic inactive and a membrane-bound active state. Extensive research has uncovered the following three interconnected functional modules that collectively form the polarity pathway (26–29): actin-based transport, spatial cues and a reaction-diffusion module (figure 1a). While all three modules contribute to cell polarization, the reaction-diffusion module has been identified as its primary driver in wild-type cells. This functional module critically depends on the protein Bem1 (30, 31), which serves as a scaffold that connects most other components of the module. Deletion of Bem1 is highly disruptive and nearly abolishes the ability of cells to polarize (32, 33). Intriguingly, an earlier study has shown that this defect can be almost completely compensated by the additional deletion of two other proteins, Bem3 and Nrp1 (33). Bem3 is a known component of the reaction-diffusion module (34, 35), while Nrp1 is a protein of unknown function for which a role in polarity establishment might still be uncovered.

In this study, we evaluate how compensatory gene deletions within the cell polarity pathway influence the global fitness landscape and explore the role of modules in shaping evolutionary dynamics. We utilize a transposon mutagenesis screen to determine changes in gene fitness, defined as the impact of a gene deletion on fitness, and demonstrate that compensatory evolution affects the fitness landscape on a genome-wide scale. Despite compensation being achieved solely by gene deletions, the number of genes with increased and decreased fitness in the compensated strain are nearly equal. This finding indicates that evolution through gene loss does not always reduce mutational robustness. Because the evolutionary dynamics of proteins are often linked to their topological role in the protein interaction network (36, 37), we attempted to explain our observed pattern of gene fitness changes using the structure of the protein-protein interaction network of *S. cerevisiae*. However, we were unable to find any meaningful correlation with the protein network structure. Instead, studying the functional associations among the affected genes revealed that

genes related to the same process tend to share the same direction of gene fitness change. Thus, although gene fitness changes do not remain isolated to the originally perturbed module, functionally related genes appear to respond similarly during adaptation. Cell polarity was amongst the biological processes enriched for genes with a negative fitness change, suggesting that redundancies between different modules within the polarity pathway mediate the capacity to compensate for gene loss. Collectively, our data demonstrates that mutations compensating for defects in a single functional module can lead to genome-wide changes in the gene-fitness relationship.

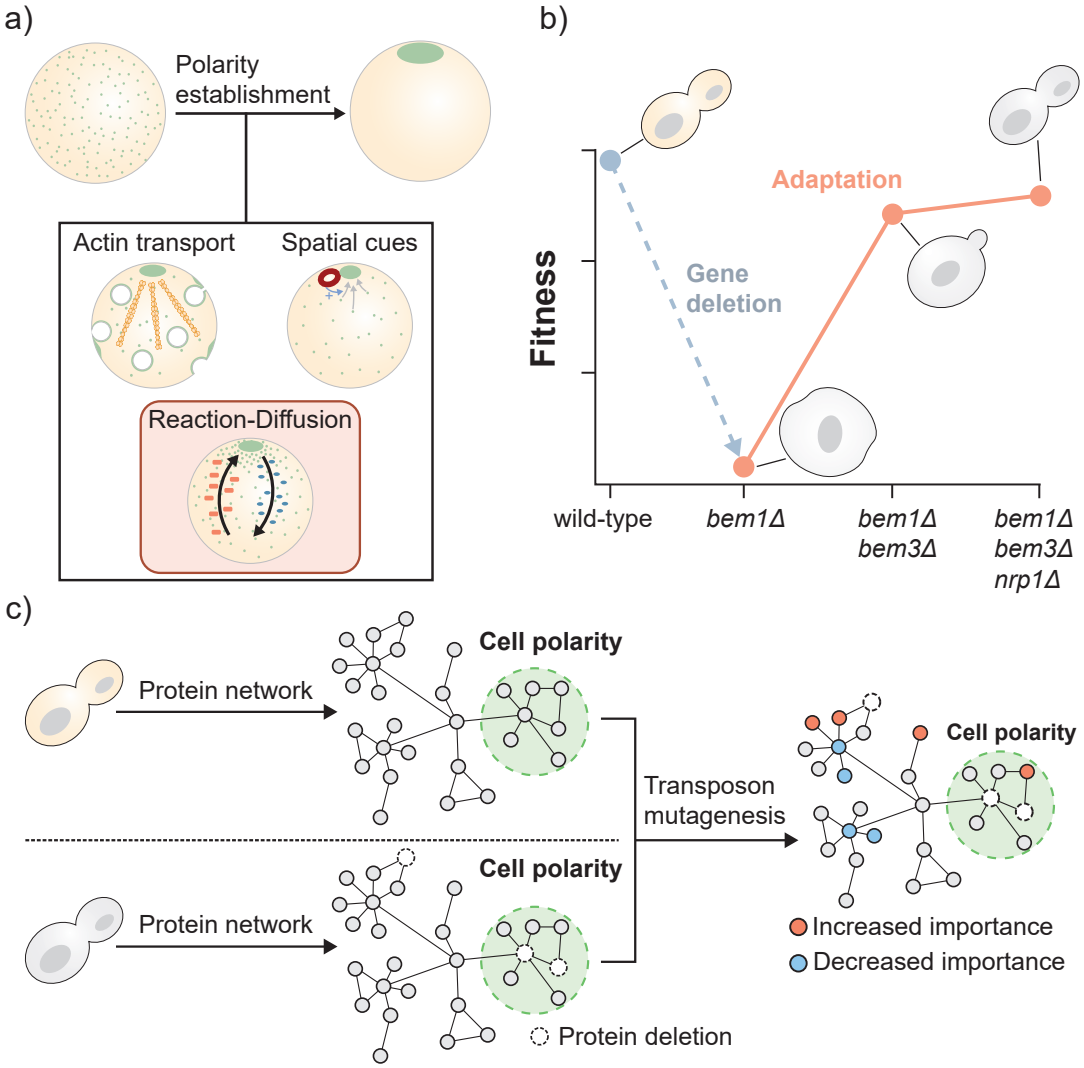

Fig. 1. Determining the consequences of compensatory evolution of the polarity pathway on global cellular physiology. (**a**) Polarity establishment is an evolutionary conserved function of nearly all cell types. In budding yeast, cell polarization is initiated by the formation of an asymmetric distribution of Cdc42 (shown in green). Actin-based transport, spatial cues and a Bem1-mediated reaction-diffusion module have all been suggested to be involved in Cdc42 polarization dynamics. In this study, we examine how the fitness landscape changes after evolutionary compensation for a genetic perturbation in the reaction-diffusion module. (**b**) The phenotypic effects resulting from the deletion of the polarity protein Bem1 can be compensated by the additional loss of Bem3 and Nrp1. The resulting polarity mutant has a fitness similar to a wild-type strain under standard laboratory conditions. The figure is based on previous data obtained by Laan et al. (33). (**c**) We use a transposon mutagenesis screen to determine the effects of compensatory evolution on the genotype-fitness relationship on a global scale. Genes that differ in fitness between the wild-type strain and the polarity mutant are anticipated to encode proteins with modified physiological roles.

2. Results

A. Compensatory mutations cause global changes in the fitness landscape. The unperturbed wild-type strain and the compensated *bem1*Δ*bem1*Δ*nrp1*Δ mutant, which we refer to as the polarity mutant, have similar fitness under laboratory conditions (33). However, whether the recovery of fitness after the loss of Bem1 is accompanied by restoration of the original fitness landscape has not been determined. The extent to which the fitness landscape is affected will reflect the molecular mechanism through which the gene deletions provide compensation. We considered the following two mechanisms for compensation to be likely.

The first mechanism relates to the argument that gene loss events during evolution often occur as a form of redundancy reduction (38). Similarly, modules can be redundant because they have largely overlapping functions. In the case of the polarity mutant, this implies that the reaction-diffusion module is redundant with one or more other modules that can drive polarity establishment. Deleting Bem1 causes the reaction-diffusion module to malfunction and interfere with these redundant modules. Consequently, further inactivation of the reaction-diffusion module through the deletion of *BEM3* and *NRP1* becomes advantageous. A recent study introduced a mathematical model that explains how functional redundancies between modules could be leveraged to form a latent pathway for polarity establishment (39). If compensation depends solely on the activation of these redundant mechanisms, we expect the fitness landscape to remain largely unchanged, with the exception of the decreased dispensability of the modules that are part of the latent pathway.

The second mechanism depends on the compensatory gene deletions facilitating novel interactions between other, non-mutated, proteins. According to the structure-function paradigm, a protein adopts a unique three-dimensional structure that ultimately determines its function. However, advancements in structural biology have revealed that protein structures are significantly more dynamic than previously believed (40). For example, up to 45% of residues in the eukaryotic proteome are found in intrinsically disordered regions of proteins (41), which lack a well-defined structure. Importantly, this structural flexibility enables protein function to adapt in response to varying cellular contexts and binding partner availability (42). Thus, the deletion of Bem3 and Nrp1 could compensate for the loss of Bem1 by modifying the functional properties of other proteins in the proteome, potentially leading to extensive rewiring of the interaction network. We anticipate that such rewiring of the protein-protein interaction network will result in global changes to the fitness landscape.

To uncover differences in gene fitness between the wild-type strain and the polarity mutant on a genome-wide scale, we performed the transposon mutagenesis screen SATAY (43). This screen utilizes the MiniDS transposon to generate a library of gene disruption mutants. In short, a plasmid carrying the MiniDS tranposon is transformed into the cell. By adding galactose to the growth media, the transposon is induced to migrate from the plasmid to a random location within the genome. This genomic integration disrupts the original sequence, typically abolishing gene function when it occurs within a coding region. By performing this process in a population of cells, a library of gene disruption mutants is created, with each mutant carrying a transposon at a different genomic location. The mutant library is subjected to a pooled fitness assay immediately after its generation. Subsequently, the genomic DNA is extracted from the mixed population and sequenced. Here, the genomic regions flanking the transposon insertion site serve as tags to distinguish the different mutants. The fitness of a mutant is reflected by the number of reads that align to the corresponding insertion site.

We generated six SATAY libraries of the wild-type strain and six libraries of the polarity mutant (figure 2a). Gene fitness was estimated by summing all read counts that map within a particular coding region. To compare datasets from independent experiments, the read count distributions were normalized (see section A) using a Beta-geometric normalization (figure S1), followed by a median of ratios normalization (figure S1). From the six replicates of each genetic background, we determined the mean (μ) and variance (σ) of the summed read read count for each gene (figure 2a).

The changes in the mean total read counts between the wild-type strain and polarity mutant are presented in the form of a volcano plot in figure 2b. Interestingly, we find that a large fraction of the annotated genes in the genome have a significant difference in fitness between the wild-type strain and the polarity mutant. Specifically, a total of 883 genes, which constitutes more than 13% of all annotated genes in the yeast genome, has an altered fitness (figure 2c). As a reference, a comparison of SATAY libraries obtained from the same genetic background revealed no statistically significant differences (figure S2). When considering the sign of the changes in gene fitness, we find that the set of genes with increased fitness (414 genes) is nearly as large as the set of genes with decreased fitness (468 genes). Moreover, both sets include several genes that have been annotated as essential (figure 2c).

These data show that compensatory evolution affects a substantial fraction of the genome. Despite the intention to limit the initial perturbation to a single functional module (33), the effects of compensation on the fitness landscape extend far beyond this module. These widespread changes in gene fitness suggests that the recovery of cell polarity cannot be accomplished by merely reallocating the lost functions of the disrupted module to other redundant modules. However, while many of the genes with an altered fitness lack a clear relation to cell polarity, the set also includes several genes that are known to encode for polarity regulators (for example, *SPA2*, *BNI1* and *AXL2*). Functional redundancies between modules may therefore still mediate compensation, but seemingly not without affecting gene fitness regulating other cellular processes. Another surprising finding is that compensation through gene deletions causes a nearly equal number of genes to display increased and decreased fitness. Intuitively, restoring a pathway through gene deletions is expected to shift dependency to other cellular components that take over the lost functions. This indicates that the compensatory gene deletions reorganize the gene-fitness relationship rather than simply redistribute the lost functions to other, possibly redundant, genes.

B. Changes in gene fitness cannot be predicted from the structure of the protein-protein interaction network. Are all genes equally likely to have a modified fitness in the compensated strain? If not, establishing a relationship between gene characteristics and features of the molecular interaction network could provide valuable insights for making evolutionary predictions. Earlier studies have indicated a correlation between a protein's evolutionary rate and its position within the protein interaction network (37). For example, highly connected proteins (hubs) are more likely to be essential (44) and tend to evolve slower (36, 45) than less-connected proteins, possibly because they more frequently participate in important interactions (46, 47). These findings suggest that the topology of protein interaction networks provides relevant information to understand the evolutionary dynamics of individual proteins.

We constructed a protein-protein interaction network (PPI) to see if we could identify a correlation between the topology of this network and the gene fitness changes in the compensated strain. Protein interactions were taken from the BioGrid database (48, 49). Because the ability to derive a correlation between evolutionary dynamics and PPI topology depends on network quality, we only included interactions from the Multi-Validated dataset (see Methods). The resulting PPI network consists of 3799 proteins (nodes) and 17205 undirected interactions (edges). While the majority of the proteins are part of a single, fully connected, network, a small portion is disconnected from the main graph (figure 3a). In addition, the network represents only 65% of the estimated 5800 proteins in the yeast proteome and is therefore incomplete. For the remaining 35% of the proteome, the protein interaction pattern has not been validated by at least two independent studies or experimental methods. Among the absent proteins is Nrp1, which is one of the proteins that provides compensation for the loss of Bem1 when deleted. Despite these limitations, we successfully validated that proteins in our network with more interactions are more likely to be essential (figure 3b-c), a relation formally referred to as the centrality-lethality rule (44)). This demonstrates that crucial structural features of the network can still be inferred.

The importance of a node for different topological features of a network can be described using network centralities (50). Here, we focus on three commonly used centralities: degree, betweenness and closeness. These centralities relate to the role of a protein in the PPI network. For example, hub proteins at the center of a module will typically have a high degree, while proteins that mediate connections between modules will have a high betweenness. We provide an overview of the distributions of these centralities in our network in figure S3b-d. Importantly, figure S3b shows that the degree distribution follows the sub-linear dependency characteristic of biological networks when plotted on a log-log scale. Furthermore, we show for 4 polarity proteins that those which strongly impact cellular viability (Cdc42 and Bem1) can clearly be distinguished from those which have a lesser effect (Bem3 and Bem2) based on their degree and betweenness (figure S3b and d). The relationship with closeness is, however, less clear (figure S3c).

Next, we attempted to derive a correlation between the three centrality measures described above and the likelihood that gene fitness has changed after compensatory evolution. However, the likelihood distributions are uniform for all centralities (figure 3d-f), indicating a lack of correlation. We therefore conclude that the changes in gene fitness are independent of the topology of the analyzed PPI network.

C. Genes with an altered fitness are associated with a diverse array of cellular processes. The lack of a correlation between PPI network structure and the likelihood of observing a change in gene fitness prompted us to see if our dataset can perhaps be structured based on a different feature. Importantly, the lack of dependence on the PPI network does not exclude the possibility that genes participate in the same cellular process, as this does not require proteins to physically interact. These indirect functional relationships between genes can be captured with functional association networks.

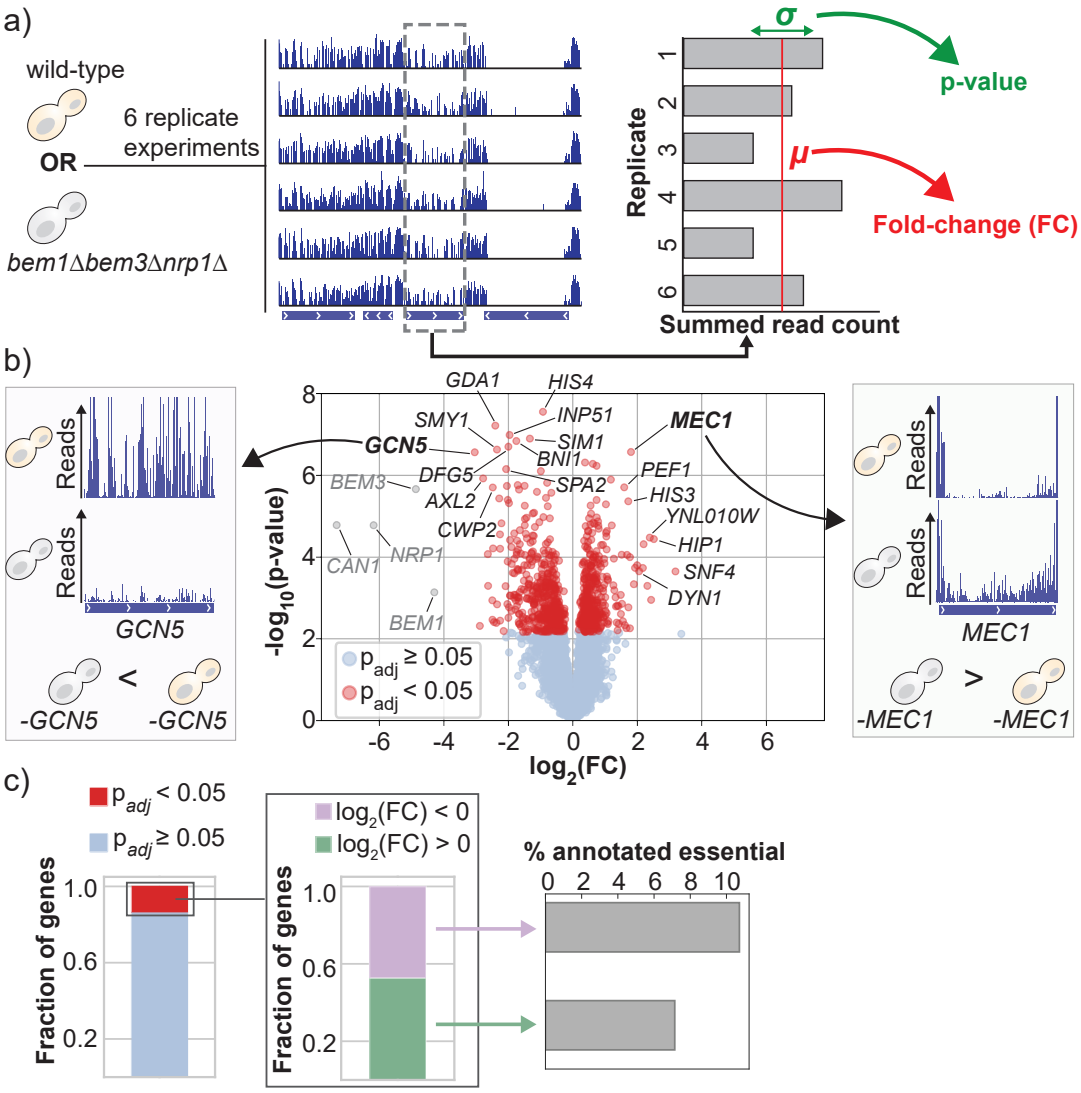

Fig. 2. Compensatory gene deletions change the gene-fitness relationship on a genome-wide scale. (**a**) We created six replicate SATAY libraries of the wild-type strain and six replicate libraries of the polarity mutant. To estimate gene fitness, we summed the read counts of transposon insertions that map to the same coding region. The mean μ and variance σ of the total read count for each gene were calculated from six replicate libraries for each genetic background. The mean and variance were subsequently used to determine the log fold-change and significance of the difference in total read count between the two genetic backgrounds. (**b**) Volcano plot of the changes in gene fitness showing the log fold-change (FC) and the significance value of the differences in gene fitness between the two strains. Positive log fold-changes relate to genes that have a higher fitness in the polarity mutant relative to the wild-type strain. Statistical significance was determined with Welch's t-test and corrected for multiple hypothesis testing with the Benjamini-Hochberg procedure. Genes that have been grayed out correspond to genes that were synthetically deleted from the genome of the polarity mutant during strain construction. (**c**) Approximately 13% of all annotated genes display a significantly altered tolerance to transposon disruptions between the two strains. An analysis of the sign of gene fitness changes in this set of genes shows that the ratio of genes that increase their relative fitness and those that decrease their relative fitness in the polarity mutant is nearly equal. Both the set of genes with increased and decreased fitness includes several genes that have been annotated as essential.

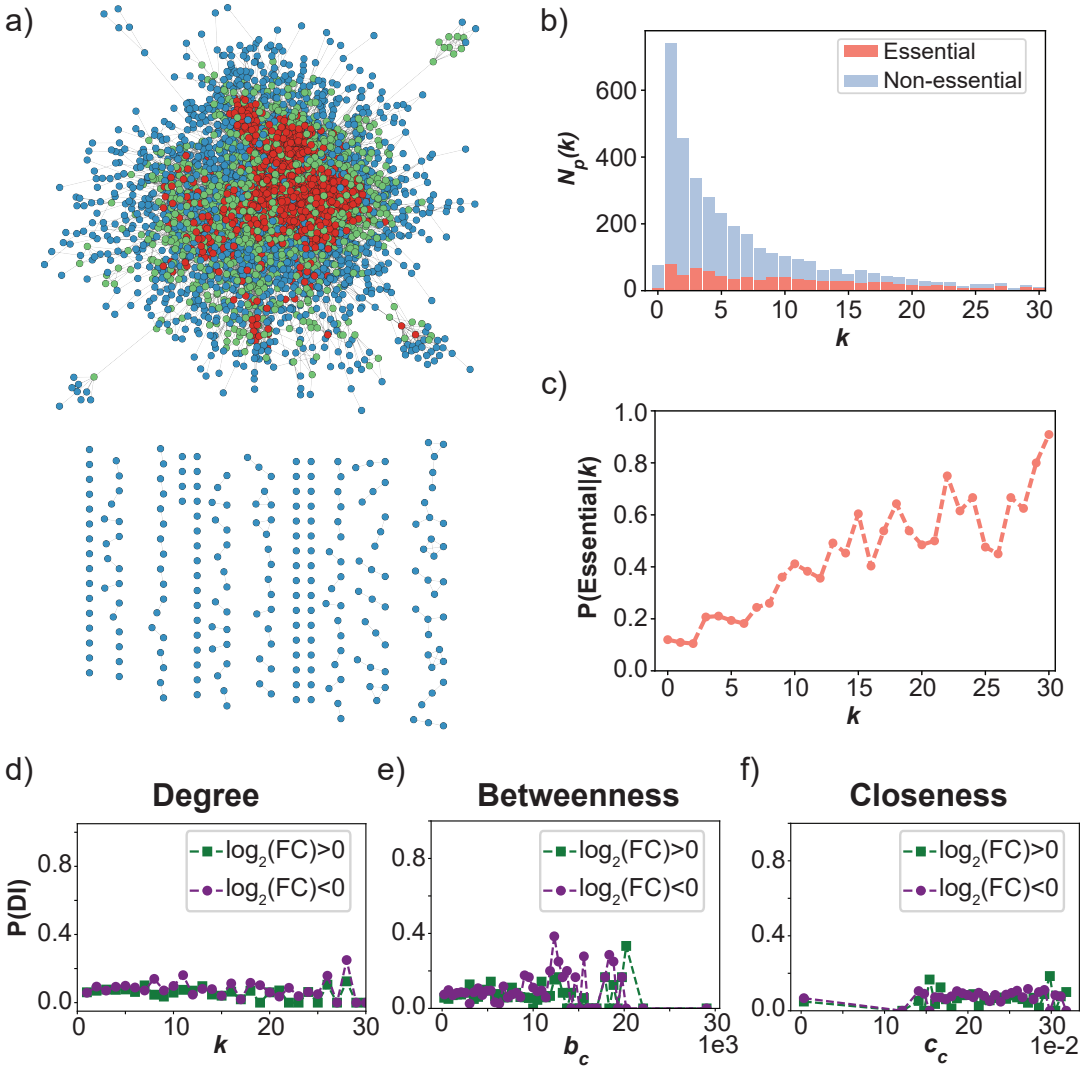

Fig. 3. Likelihood of gene fitness change does not correlate with centralities of the protein-protein interaction network. (**a**) Visualization of the protein-protein interaction network constructed from the interactions annotated by the BioGRID database. Nodes are colored according to their degree k (blue: $k \leq 3$, green: $4 \leq k \leq 10$, red: $k > 10$). (**b**) Stacked histogram of the number of proteins $N_p(k)$ with degree k. Proteins for which the gene is annotated as essential are shown in red, non-essential proteins are shown in blue. The histogram shows that the distribution for k is roughly uniform for essential proteins, while it is skewed towards lower values of k for non-essential proteins. (**c**) The probability that a node is essential as a function of k. The graph displays an increasing trend for the probability that a node is essential for larger values of k. Combined with the histogram shown in (b), it can be seen that this increasing trend is due to an enrichment of essential proteins among those with a high degree, and not because most essential proteins have a high degree. (**d-f**) The conditional likelihoods for changes in gene fitness as a function of the degree (d), betweenness (e) and closeness (f) centralities of their corresponding protein in the network shown in (a). For the three centralities shown in panels (d-f), the likelihood distribution is approximately uniform.

We used the STRING database to create a functional association network from the genes with a differential fitness between the two genetic backgrounds (51–53). In this database, functional relations between genes are inferred by integrating information from multiple sources, such as text mining, molecular complex annotations and (predicted) physical or genetic interactions. Each interaction is scored based on quality of the evidence, allowing more weight to be given to high-confidence interactions.

Similar to our PPI network, the functional associations resulted in most genes being part of a single, connected, network from which only few nodes are disconnected. Analyzing the degree distribution indicated a scale-free structure of the network, suggesting it can be partitioned into smaller sub-graphs. We ran the Markov Cluster (MCL) algorithm (54) to uncover these sub-graphs, assigning interaction confidence scores as weights. This approach uncovered 120 clusters, of which 49 consisted of three or more genes. For the seven largest clusters, we performed a gene ontology (GO) enrichment test for biological process GO terms (figure 4b and figure S4). All seven clusters were significantly enriched for at least one GO term. Five clusters were enriched for the following terms related to cellular homeostasis: translation, metabolism, signalling, chromatin remodeling and ribosome biogenesis. The remaining two clusters were associated with components of the cell division machinery: cell polarity and microtubule dynamics.

When assessing how genes with increased and decreased fitness are distributed among the clusters, we found that genes within the same cluster were often, but not always, more likely to share the same sign in fitness change than would be expected from a random partitioning (figure 4c). specifically, the clusters enriched for genes related to translation and cell polarity were also enriched for genes with an increased fitness. Alternatively, genes in the signalling and microtubule enriched clusters were more likely to have decreased in fitness.

Our analysis reveals widespread functional connections between the genes in our dataset. At the same time, by partitioning their functional association network into sub-graphs we were able to identify that these genes are involved in a diverse set of cellular processes. Intriguingly, the changes in fitness dependency seem to occur at the cellular process level rather than being attributed solely to individual genes. This uniform response of genes regulating the same process resembles the observed monochromatic behaviour of genetic interactions between modules of the yeast metabolic network (55). When comparing the cellular processes that are important for survival between the polarity mutant and wild-type strain, we observe several notable differences.

First, the general trend of genes involved in translation showing reduced fitness in the polarity mutant suggests that this strain is more sensitive to variations in protein copy numbers. This consequence of the compensatory mutations was also predicted by the mathematical model of Brauns et al. (39). Second, the finding that genes regulating microtubule dynamics become increasingly dispensable is surprising, given microtubules are required to perform vital cellular functions such as spindle positioning and nuclear migration (56, 57). Thus, compensatory evolution appears to profoundly affect the genetic wiring of the cell, changing the patterns of essentiality and dispensability beyond the genes related to the perturbed module.

D. Compensatory evolution is mediated by redundancies within the polarity pathway. The wild-type strain and polarity mutant are nearly identical with respect to their polarization efficiency (33). The ability to polarize efficiently despite the loss of three genes implies the existence of an alternative pathway that is active in the absence of the reaction-diffusion module. Because it can function without adding any components, this pathway is considered to be latent or hidden within the interaction network of the wild-type strain (39). While reconstruction of the polarity mutant has demonstrated the existence of this latent pathway, its specific structure remains to be determined. In the previous section, we found a cluster enriched with polarity genes amongst those affected by the mutations compensating for defects in the reaction-diffusion module. The presence of this cluster implies that functional redundancies between the different modules of the polarity pathway could play an important role during compensatory evolution.

To uncover how these redundancies can be exploited to constitute an alternative polarization mechanism, we annotated the genes in the cell polarity cluster according to their gene ontology (figure 5a). This revealed that the majority of the genes encode for proteins that are part of one the following three cellular components: (1) cortical actin patches, (2) the bud scar and (3) the polarisome. Interestingly, each of these components relates to a different module of the known polarity pathway. Specifically, the bud scar provides a spatial cue that directs the axis of polarity, while cortical actin patches and the polarisome regulate the transport of active Cdc42 via endo- and exocytic vesicles, respectively. For all three cellular components, a substantial fraction (>20%) of the complete set of annotated genes is represented in our dataset.

While all three cellular processes identified above are active in wild-type cells, their contribution to cell polarization

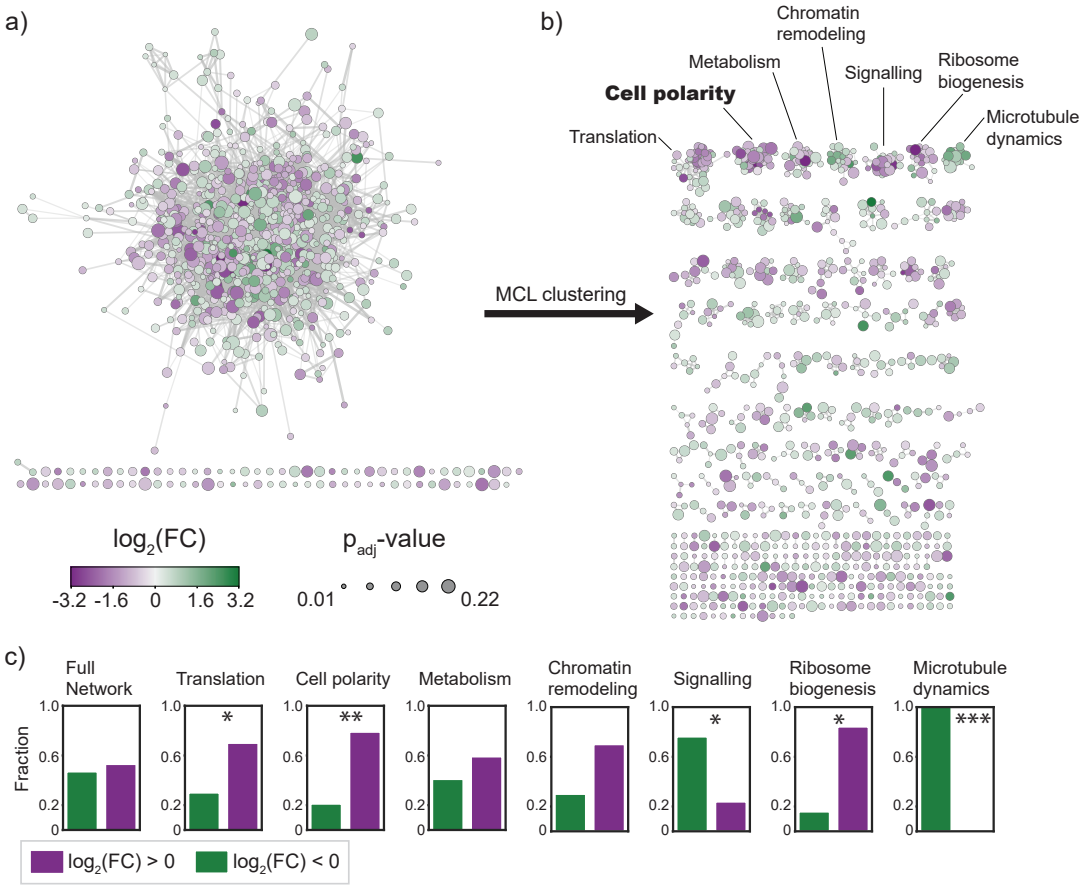

Fig. 4. Clustering of a functional association network reveals diversity of biological processes affected by compensatory evolution (**a**) Graph of the functional association network constructed from genes that have a differential fitness between the wild-type strain and the polarity mutant. Node color scales with the magnitude and sign of the fitness effect, node size scales according to the significance level (adjusted p-value). The graph shows that most genes are contained in a single, densely connected network. (**b**) The clusters formed from the graph shown in (a) by the Markov clustering algorithm (MCL). The clusters are arranged from large (top left) to small (bottom right). The biological process gene ontology enrichment is shown for the seven largest clusters. (**c**) Distribution of genes with a decreased (purple) and increased (green) fitness in the polarity mutant for the seven largest clusters. The clusters enriched for translation, cell polarity, and ribosome biogenesis are enriched for genes with decreased fitness. Conversely, the signalling and microtubule dynamics clusters are enriched for genes with increased fitness. *=$p < 0.05$, **=$p < 0.005$, ***=$p < 5 \cdot 10^{-4}$. The significance level was determined with Fisher's exact test.

remains inconspicuous when the Bem1-dependent reaction-diffusion module is active. This feature has been attributed to the hierarchical relationship that normally exists between modules in the polarity pathway (58). To determine if this hierarchy still exists within the latent pathway, we used the log-fold change distributions of genes belonging to each of the three modules. Specifically, the module that takes on the most prominent role in the latent pathway should show the greatest change in gene fitness.

However, quantitatively comparing log-fold changes in read counts summed on the gene level can give distorted results because their values depend on gene size. For example, if the region within a gene that exhibits a differential tolerance to transposon disruptions is small relative to the total gene size, the log-fold change value will be attenuated. To address this issue, we performed a more refined estimate of changes in gene fitness by using a moving average with a 300 bp window over the genes to obtain a fitness profile across the coding region (figure 5b). The choice for a window size of 300 bp was made based on the average size of a protein domain. Using this approach, we were able to identify the region within a gene that gives the largest log-fold change values between the wild-type strain and the polarity mutant. However, we note that while using a smaller window to compare genes allows a more accurate quantification of the log fold-change, it also increases the sensitivity to outliers.

When comparing the maximum log-fold changes in genes of the three modules described above, we observed no significant ($p<0.05$, t-test) changes between the means of their distributions (5c). Therefore, each module appears to be of equal importance for the functioning of the latent polarity pathway.

Although our results do not provide a complete molecular mechanism for the latent polarity pathway, they do offer important insights into its structure. The increased dependency on spatial cues we observe agrees with previous studies that have shown that the deletion of Bem1 is not lethal, but only if the bud scar remains intact (30, 32). This dependency on the bud scar therefore appears to persist even after mutations that compensate for the fitness effects caused by losing Bem1 have occurred. In contrast, the role of actin based transport in cell polarization has been a subject of ongoing debate (26, 59–63). Our results support the role of the actin network in promoting polarity establishment, at least within the latent pathway. However, the lack of hierarchy between the different modules in the latent pathway suggests that,while they contribute to polarity establishment, each individual module contribution is too weak to drive polarization.

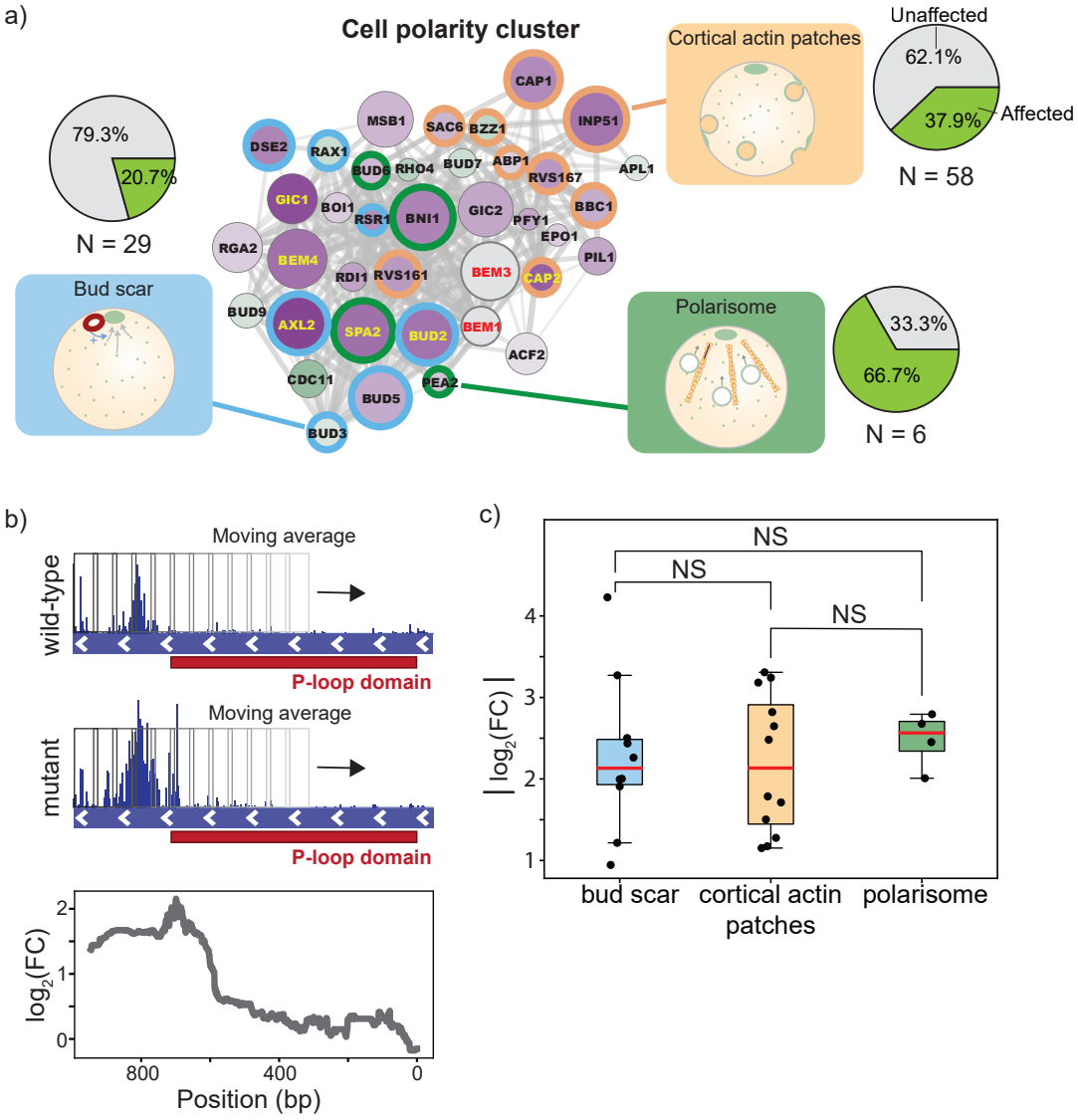

Fig. 5. Redundant polarity modules contribute to the latent polarity mechanism in a non-hierarchical manner. (**a**) Genes in the cluster enriched cell polarity factors are annotated based on whether they belong to the bud scar (blue), cortical actin patches (yellow) or polarisome (green). The fraction of all N genes annotated with each of these three cellular component GO terms that are present in our dataset is shown in the pie charts on the right. For all three components, at least 20% of the annotated genes are present in our dataset. The cortical actin patch has the lowest representation, but is also the least polarity-specific of the three cellular components. (**b**) An example of the sliding average procedure taken to obtain the domain with the largest log-fold change is shown for *CDC11*. In *CDC11*, the region downstream of the P-loop domain (relative to the start codon) has an increased tolerance to transposon disruptions, while the remainder of the gene shows a similar tolerance between the two strains. Using this method, we can accurately capture this domain-dependency. The largest absolute value of the log-fold change profile over a gene was used to generate and compare the distributions shown in c. (**c**) Log-fold change distributions for the genes in our dataset encoding for proteins of the bud scar, cortical actin patches and polarisome. The orange line denotes the median of the distribution, the whiskers show the interquartile range. The three distributions all had very similar medians, and the difference in their mean was not statistically significant (p>0.05, t-test).

3. Discussion

In this report, we examined the effect of mutations compensating for a perturbation targeting the polarity pathway on the genotype-fitness relation. Remarkably, our data show that two strains, which polarize with nearly identical efficiency and differ by only three genes, exhibit fitness differences in over 13% of the genes in the genome. This percentage is similar to what has been reported for the percentage of genes that display a background-dependent fitness effect between different natural isolates (64, 65). Given that the genomic divergence between natural isolates is generally larger than that between the two strains we studied, we anticipated finding a smaller number of genes with a background-dependent fitness effect. One explanation for the relatively high number of affected genes we identified is that our analysis method was aimed at detecting changes in gene fitness on a continuous scale, rather than categorizing genes simply as essential or non-essential. Our set therefore contains genes that would likely have remained undetected by other studies. However, despite the higher sensitivity of our method, the number of affected genes reported is likely still an underestimation. Importantly, because we sum read counts over the entire open reading frame, our analysis is relatively insensitive to domain effects on fitness when genes are large. Similarly, by only considering coding regions we are unable to identify changes in gene fitness that are related to changes in gene expression levels.

The impact of compensatory mutations on the fitness landscape may vary depending on the environment in which fitness is assessed. (66, 67). In a recent survey of changes in gene essentiality between natural isolates of *S. cerevisiae*, it was proposed that the majority of the observed variation was due to environmental factors (64). That is, it was argued that most genes that emerged as differentially essential in their screen were caused by strain-specific responses to growth in non-standard media. The mutagenesis screen we performed has a similar limitation. SATAY requires transposition to be induced in media in which galactose is the sole carbon source (43), while the phenotypic similarity between our two strains has only been verified for growth under standard conditions (33). We can therefore not exclude that we would have observed less changes in gene fitness if we had performed a screen in a different environment. However, we expect that the effect of the environment on our results is limited for the following two reasons. First, the overall genetic divergence between our strains should be smaller than what is typical for natural isolates that may have adopted different lifestyles to accommodate to their niche. A different study indeed showed that most genetic interactions are conserved across environments (68). Second, while we did identify a cluster enriched for metabolically related genes in our gene set, genes involved in galactose metabolism or respiration are not overrepresented in this cluster.

With the aim of making a first step towards predictive models, we assessed whether the genes that are affected by compensatory evolution can be identified based on their position in a PPI network (65). We did not observe any correlation between the likelihood that gene fitness changes and the degree, betweenness and connectedness centralities of the associated proteins. Thus, the genes that will have a differential fitness effect between the two strains cannot be inferred from these three centralities. While we were able to validate the presence of several features that are typically observed in PPI networks, we cannot reject the possibility that the lack of correlation is a consequence of poor quality of our PPI network. The accuracy of annotated physical interactions is know to substantially vary between datasets (69, 70). To exclude interactions with low confidence, we used the Multi-Validated dataset available from the BioGRID to construct our PPI network. However, network quality does not only depend on the confidence that an interaction exists, but also on the confidence that an interaction is absent (71). In addition, bias is a significant problem in the mapping of interaction networks. Weak and transient interactions are often discarded because they cannot be captured with high confidence by affinity based screens. These interactions could be of particular importance in cases where network structure needs to be recovered in response to gene loss, as recent work identified these interactions as the "glue that holds the cellular network together" (72).

The concept of modularity remains strongly embedded, both implicitly and explicitly, in cell and synthetic biology (16, 73). Using the functional associations available from the STRING database, we found that genes with differential fitness can be grouped into functionally enriched clusters, suggesting that some level of modular organization does exist within biological networks. This is further supported by our finding that several clusters are enriched for either genes with increased or decreased fitness. At the same time, our data demonstrates that the compensatory evolution of one module of the polarity pathway affects gene fitness in many other modules, including those that have no clear relation to cell polarization. These findings complement results from an earlier study by Harcombe et al. (74), which showed that gene loss can be compensated by mutations in functionally unrelated genes. Our results further demonstrate that, in addition to the compensatory mutations themselves, gene fitness across multiple cellular processes can change during compensatory evolution. It is conceivable that these changes in gene fitness will affect the further evolution of these modules. We therefore conclude that functional modules do not necessarily evolve

independently of each other and that evolvability does not require modularity.

From our set of genes with a differential fitness, we were able to identify the bud scar, the polarisome and actin patches as the major structures that contribute to the latent pathway for polarity. While earlier studies indicated that actin based transport can drive polarization of Cdc42 (31), their results were criticized to be largely caused by the unintended activation of the stress response pathway. Our data provides new evidence that actin based transport can contribute to cell polarization, but its activity is most likely too weak to achieve polarity establishment on its own. The bud scar could provide an additional activation mechanism for Cdc42 that, in conjunction with actin recycling, is able to maintain a polarized distribution of Cdc42. Indeed, the bud-scar associated GTPase Rsr1 is capable of recruiting Cdc24, the GEF for Cdc42, to the bud site (75). This recruitment would cause a locally increased rate of Cdc42 activation near the bud scar, which, combined with a possible saturation of Cdc42 GAPs caused by the loss of Bem3, could drive polarity establishment. While such a mechanism largely agrees with the finding obtained by the model of Brauns et al. (39), a key difference is that we do predict that Cdc24 is polarized prior to budding in the polarity mutant. In addition, we propose that this latent polarity pathway may also depend on protein interactions that have not been observed in wild-type cells. For example, it is unclear why Boi1, which depends on Bem1 to interact with Cdc24, has an increased importance in the polarity mutant. The same goes for several other genes that are too numerous too discuss individually. Microscopy, *in vitro* experiments and molecular models will likely be needed to obtain a complete picture of the latent polarity pathway.

We found that the number of genes that increase fitness and those that decrease fitness after compensatory evolution is nearly equal. Thus, gene loss does not appear to necessarily decrease mutational robustness, but rather cause a restructuring of the gene-fitness relationship. While it is somewhat counter-intuitive that gene loss can lead to the increased dispensability of other genes, a mechanistic explanation has been given for this phenomenon. Specifically, if the loss of a protein fully inactivates an associated pathway, the additional loss of other proteins on the same pathway does not incur additional fitness costs (76, 77). In this way, the loss of a protein can appear to buffer mutations in other proteins.

However, this explanation does not hold for the case where gene deletions lead to the functional restoration of the perturbed pathway. Similarly, the scale at which we see changes in gene fitness occur and the diversity of cellular processes that are affected suggests an alternative mechanism. How then, can the loss of a small number of genes cause changes in the gene-fitness relation of such a large part of the genome? Although our results do not provide direct evidence, we propose that gene loss can induce widespread rewiring of the physical interaction network based on the following reasoning. The fitness effect resulting from the loss of a gene is related to the physiological role of the protein encoded by that gene. However, it is increasingly recognized that a gene does not encode for a protein with a static structure, but rather for structural ensembles (40, 78). Which structure is preferred depends on the environment, making protein function heavily context dependent (79). We speculate that in a similar manner, protein structure can be affected by the presence or absence of specific binding partners (42, 80). Loss of one binding partner could in that case cause a conformational change that affects its affinity for other possible binding partners. In this way, the deletion of one protein can initiate a cascade that induces large-scale rewiring of the protein interaction network.

In conclusion, we show that modularity plays a minor role during evolution. Instead, compensatory evolution induces genome-wide rearrangements in interaction networks that cross module boundaries and which might be mediated by protein promiscuity. Because these genome-wide changes affect the adaptive potential of many different cellular processes, this could explain how gene loss can promote evolutionary novelty (81–86) and allow populations to escape fitness peaks (82). Despite the minimal role of modularity, some hope for making evolutionary predictions might be drawn from our finding that genes related to the same biological process occasionally share the same sign in fitness change after compensatory evolution. Our work also makes a case for the importance of protein ensembles and weak interactions in shaping adaptive pathways. Interactions that are weak in non-perturbed conditions could be driving the re-establishment of biological pathways in response to perturbations. Mapping these weak, non-specific interactions that are frequently perceived as irrelevant will be instrumental to bridge the gap between biological network structure and the evolutionary potential of biological processes.

4. Methods

A. Strains. All strains used in this study are from the W303 genetic background. The cas9 cassette was obtained from plasmid p414-TEF1p-cas9-CYC1t (87) and fused to the up- and downstream genomic sequences of the *HO*-locus and the *ScURA3* marker using overlap-extension PCR. The resulting genetic construct was transformed into the wild-type strain and the polarity mutant using according to a lithium-acetate transformation protocol (88). Correct integration was verified with colony PCR. Endogenous expression of Cas9 from the *HO*-locus had no significant effects on growth. Strains were made compatible with SATAY by removing the *ScURA3* marker and the endogenous *ADE2* locus using the CRISPR/Cas9 system according to the double guide-RNA (gRNA) method of Mans et al. (89). gRNA sequences targeting the *ScURA3* marker and *ADE2* locus were designed using the online toolbox CHOPCHOP (90). The repair fragment for removal of the *ScURA3* marker was constructed by PCR amplification of the genomic sequence upstream and downstream of the *ScURA3* marker and fusing the two fragments using overlap-extension PCR. Similarly, the repair fragment for removal of the *ADE2* locus was constructed by fusing the up- and downstream genomic sequences with overlap-extension PCR. Removal of *ScURA3* marker and the *ADE2* was verified by colony PCR using primers 5 and 6 and primers 7 and 8, respectively. Correct clones were cured from their gRNA plasmids by inoculating them in non-selective medium (YPD) and growing them until saturation (∼1.5 days) at 30 °C. After saturation was reached, the cultures were plated to single colonies on YPD agar plates and screened for the loss of growth on media selecting for the marker of the gRNA plasmid. Strains were stored at -80 °C as frozen stocks in 40% (v/v) glycerol. The double gRNA plasmids that were used as a template were a kind gift from Pascale Daran-Lapujade.

B. Media. Standard culturing and growth assays were performed in YPD (10g/L Yeast extract, 20 g/L Peptone, 20 g/L dextrose), SC (6.9 g/L Yeast nitrogen base, 0.75 g/L Complete supplement mixture, 20 g/L dextrose). For ade^- strains, standard growth media was supplemented with 20 mg/L adenine just before incubation. Liquid media for the preculture and induction steps of SATAY were prepared according to the recipe in table S1. After preparation, the media was filter sterilized using Rapid-Flow Sterile Disposable Filter Units (Nalgene) and stored at 4°C until use. Liquid media for the reseed step of SATAY was prepared by autoclaving 2.6 L of MiliQ water in a 5 L flask. 400 mL of a 7.5X concentrated solution of the nutrients was prepared separately and filter sterilized. To prevent the degradation of media components, this concentrate was stored in the dark at 4°C until used. On the day of reseed, the concentrate was aseptically added to the 5 L flask containing 2.6 L of MiliQ water and mixed. Solid media was prepared by adding 20 g/L agar and 30 mM Tris-HCl (pH 7.0) to the liquid media recipe and autoclaving the mixture for 20 minutes at 121°C. 20 mg/L adenine was aseptically added after autoclaving, unless plates were intended to be selective for adenine auxotrophy.

C. Transposon mutagenesis screens.

C.1. Library generation. SATAY libraries were generated based on the procedure described by (91), which is a modification of the original protocol (92) to allow transposition in liquid media. ade^- cells were transformed with plasmid pBK549 (91), which was a kind gift from Benoît Kornmann, according to a lithium acetate transformation protocol (88). To screen for clones transformed with the intact version of plasmid pBK549 (see (93) for details on the different species of pBK549), 12-24 colonies were picked from the transformation plate, re-streaked on fresh SD-ADE and SD-URA plates and incubated for 3 days at 30°C. For clones that showed full growth on SD-URA plates while producing a small number of colonies on SD-ADE plates, cells were scraped from the SD-URA plate and used to inoculate 25 mL of preculture media at an OD600 of 0.20-0.28. Precultures were grown on an orbital platform shaker at 160 rpm, 30 °C until the OD600 was between 5-7 (∼20h). The saturated precultures were used to inoculate 200 mL of induction media at an OD600 of 0.10-0.27 and grown for 52 hours to allow transposition to occur. The efficiency of transposition was monitored by plating samples of the liquid induction cultures on SD-ADE at T=0 and T=52 hours and scoring the number of colonies on these plates after 3 days of incubation at 30 °C. After 52 hours of induction, the resulting transposon mutagenesis libraries were reseeded in 3 liters of reseed media at an OD600 of 0.21-0.26. Typically, this meant that around 7 million transposon mutants were reseeded per library. Reseeded libraries were grown for 92 hours at 140 rpm, 30 °C. At the end of reseed, cells were harvested by centrifugation of the reseed cultures at 5000 xg for 30 minutes. Cell pellets were stored at -20 °C.

C.2. Genomic DNA extraction. A 500 mg frozen pellet was resuspended in 500 µL cell breaking buffer* and distributed into 280 µL aliquots. 300 µL of 0.4-0.6 mm glass beads (Sigma-Aldrich, G8772) and 200 µL of Phenol:Chloroform:isoamyl alcohol 25:24:1 (Sigma-Aldrich, P2069) were added to each aliquot and cells were lysed by vortexing the samples with a Vortex Genie 2 at maximum speed at 4°C for 10 minutes. 200 µL of TE buffer was added to each lysate, after which

*2% Triton X-100, 1% SDS, 100 mM NaCl, 100 mM Tris-HCl pH8.0, 1 mM EDTA

the samples were centrifuged at 16100x g, 4°C for 5 minutes. After centrifugation, the upper layer (∼400 µL) was transferred to a clean eppendorf tube. 2.5 volumes of 100% absolute ethanol was added to each sample and mixed by inversion to precipitate the genomic DNA. After precipitation, the DNA was pelleted by centrifugation at 16100x g, 20°C for 5 minutes. The supernatant was removed and the DNA pellet was resuspended in 200 µL of 250 µg/ml RNAse A solution (qiagen, Cat. No. 19101). the resuspended DNA pellets were incubated at 55°C for 15 minutes to allow digestion of the RNA. After digestion, 20 µL 3M, pH 5.2 sodium acetate (Merck) and 550 µL 100% absolute ethanol was added to each sample and mix by inversion. DNA was pelleted by centrifugation at 16100x g, 20°C for 5 minutes. Pellets were washed with 70% absolute ethanol and dried at 37°C for 10 minutes or until all ethanol had evaporated. The dried pellets were resuspended in a total volume of 100 µL MiliQ water and the concentration of the genomic DNA samples was quantified on a 0.6% agarose gel using the Eurogentec Smartladder 200bp-10kb as a reference. Prepared DNA samples were stored at -20°C or 4°C until used.

C.3. Library sequencing. To prepare genomic DNA samples for sequencing, 2x2 µg of DNA from each sample were transferred to non-stick microcentrifuge tubes and digested with 50 units of DpnII and NlaIII in a total volume of 50 µL for 17 hours at 37°C. After digestion, the restriction enzymes were heat-inactivated by incubating the samples at 65°C for 20 minutes. Digestion results were qualitatively assessed by visualization on a 1% agarose gel stained with Sybr-Safe. Successfully digested DNA samples were circularized in the same tube using 25 Weiss units of T4 DNA ligase (Thermo Scientific, Catalog #EL0011) at 22°C for 6 hours in a total volume of 400 µL. After ligation, the circularized DNA was precipitated using 1ml 100% absolute ethanol, 20 µL 3M, pH 5.2 sodium acetate (Merck) and 5 µg linear acrylamide (invitrogen, AM9520) as a carrier. DNA was precipitated for at least 2 days at -20°C. Precipitated DNA was pelleted by centrifugation for 20 minutes at 16100x g at 4°C and washed with 1 ml of 70% ethanol. After washing, the DNA was re-pelleted by centrifugation for 20 minutes at 16100x g at 20°C, the supernatant was removed and pellets were dried for 10 minutes a 37°C. Each dried pellet was resuspended in water and used as a template for 20 PCR reactions of 50 µL. Transposon-genome junctions were amplified using the barcoded primers 1 and 2 (table S2) for DpnII digested DNA or primers 3 and 4 (table S2) for NlaIII digested DNA on a thermal cycler (Bio-Rad C1000 Touch). PCR amplified samples were purified using the NucleoSpin Gel and PCR cleanup kit (Macherey-Nagel) and quantified on the NanoDrop 2000 spectrophotometer (Thermo Scientific). For each sample, equal ratios (w/w) of DpnII and NlaIII digested DNA were pooled. Library preparation and sample sequencing were performed by Novogene (UK) Company Limited. Sequencing libraries were prepared with the NEBNext Ultra II DNA Library Prep Kit, omitting the size selection and PCR enrichment steps. DNA libraries were sequenced on the Illumina NovaSeq 6000 platform using Paired-End (PE) sequencing with a read length of 150 bp.

C.4. Sequence alignment. FASTQ files obtained from the NovaSeq 6000 platform were demultiplexed into DpnII and NlaIII digested DNA samples based on the barcodes introduced during PCR amplification. Read pairs with non-matching barcodes were discarded. After demultiplexing, the forward read of each read pair was selected and the sequences upstream of primer 688_minidsSEQ1210 (43) and downstream of the DpnII (GATC) or NlaIII (CATG) restriction site were trimmed. All demultiplexing and trimming steps were performed with BBduk (94) integrated into a home-written pipeline written in Bash. After trimming, the forward reads were aligned to the S288C reference genome (version R64-2-1_20150113) with the Transposonmapper pipeline ((95) version v1.1.4) using the following settings:

- Data type: 'Single-end'
- Trimming software: 'donottrim'
- Alignment settings: '-t 1 -v 2'

D. Volcano plots. Read count distributions were corrected for spikes using the Beta-Geometric correction method and subsequently normalized for differences in transposon density and sequencing depth with the median of ratios normalization (see section A). P-values were generated using the unequal variance independent t-test available from the SciPy library in Python (96) and corrected for multiple hypothesis testing with the Benjamini-Hochberg procedure implemented in the TRANSIT software tool ((97), version 3.2.6). Genes were considered to be essential if they have been annotated as such by the Saccharomyces Genome Database ((98), accessed on 03/17/2006).

E. Physical interaction network analysis. Because the correlation between evolutionary dynamics and PPI topology depends on network quality, we only included interactions from the Multi-Validated dataset. The Multi-Validated (MV) dataset of physical interactions (release BIOGRID-4.4.214) was downloaded from the BioGRID in Tab 3.0 format, which contains all physical interactions that have been validated by at least two different experimental systems or publication sources. These experimental systems include high-throughput *in vitro* screens, and the network

therefore contains a mixture of *in vitro* and *in vivo* validated interactions. Moreover, it should be noted that the constructed PPI is a static representation of a dynamic network structure. The original set of physical interactions was filtered to obtain only those that correspond to interactions between proteins of *S. cerevisiae*. Interactions inferred from Affinity Capture-RNA or Protein-RNA were excluded from the dataset as these interactions do not represent direct physical interactions between proteins. Network visualizations were made with Cytoscape ((99), version 3.9.1). Network centralities (degree, betweenness and closeness) were calculated using the NetworkX package in Python ((100), version 2.8.4).

F. Functional interaction network analysis. Functional associations between gene products were retrieved from the STRING database (51) through the stringApp plugin ((52), version 1.7.1) of Cytoscape ((99), version 3.9.1). The confidence threshold of the imported interactions from STRING was set to medium (confidence level $\leq$0.40). Markov Clustering (MCL) was performed with the clusterMaker2 (version 2.2) plugin of Cytoscape using the confidence level as edge weight and setting the granularity parameter to 2.5.

5. Acknowledgements

We thank Agnès Michel and Benoît Kornmann for providing the plasmid pBK549 containing the MiniDS transposon system, their help and guidance when setting up the SATAY assays and insightful discussions. We are grateful to Melanie Wijsman and Nicole Bennis for providing us with the plasmids for CRISPR/Cas9 gene editing and their support during troubleshooting. We also thank Leila Iñigo de la Cruz and Gregory van Beek for establishing and maintaining the SATAY data analysis pipeline and their advice regarding data management. L.L. and E.K. gratefully acknowledge funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement 758132).

1. Helen E. Burston, Lymarie Maldonado-Báez, Michael Davey, Benjamen Montpetit, Cayetana Schluter, Beverly Wendland, and Elizabeth Conibear. Regulators of yeast endocytosis identified by systematic quantitative analysis. *Journal of Cell Biology*, 185(6):1097–1110, 6 2009. ISSN 1540-8140. . URL https://rupress.org/jcb/article/185/6/1097/35216/Regulators-of-yeast-endocytosis-identified-by.
2. Marko Kaksonen, Christopher P. Toret, and David G. Drubin. A Modular Design for the Clathrin- and Actin-Mediated Endocytosis Machinery. *Cell*, 123(2):305–320, 10 2005. ISSN 00928674. . URL https://linkinghub.elsevier.com/retrieve/pii/S0092867405009785.
3. Thomas D. Pollard and Ben O'Shaughnessy. Molecular Mechanism of Cytokinesis. *Annual Review of Biochemistry*, 88(1):661–689, 6 2019. ISSN 0066-4154. . URL https://www.annualreviews.org/doi/10.1146/annurev-biochem-062917-012530.
4. Wouter-Jan Jan Rappel and Leah Edelstein-Keshet. Mechanisms of cell polarization. *Current Opinion in Systems Biology*, 3:43–53, 6 2017. ISSN 24523100. . URL https://linkinghub.elsevier.com/retrieve/pii/S2452310016300166.
5. Eveline T Diepeveen, Thies Gehrmann, Valérie Pourquié, Thomas Abeel, and Liedewij Laan. Patterns of Conservation and Diversification in the Fungal Polarization Network. *Genome Biology and Evolution*, 10(7):1765–1782, 7 2018. ISSN 1759-6653. . URL https://academic.oup.com/gbe/article/10/7/1765/5040802.
6. Mohan K. Balasubramanian, Erfei Bi, and Michael Glotzer. Comparative Analysis of Cytokinesis in Budding Yeast, Fission Yeast and Animal Cells. *Current Biology*, 14(18):R806–R818, 9 2004. ISSN 09609822. . URL https://linkinghub.elsevier.com/retrieve/pii/S0960982204006980.
7. Jolien JE van Hooff, Eelco Tromer, Leny M van Wijk, Berend Snel, and Geert JPL Kops. Evolutionary dynamics of the kinetochore network in eukaryotes as revealed by comparative genomics. *EMBO reports*, 18(9):1559–1571, 9 2017. ISSN 1469-221X. . URL https://www.embopress.org/doi/10.15252/embr.201744102.
8. Elizabeth Conibear. Converging views of endocytosis in yeast and mammals. *Current Opinion in Cell Biology*, 22(4):513–518, 8 2010. ISSN 09550674. . URL https://linkinghub.elsevier.com/retrieve/pii/S0955067410000724.
9. Béla Szamecz, Gábor Boross, Dorottya Kalapis, Károly Kovács, Gergely Fekete, Zoltán Farkas, Viktória Lázár, Mónika Hrtyan, Patrick Kemmeren, Marian J. A. Groot Koerkamp, Edit Rutkai, Frank C. P. Holstege, Balázs Papp, and Csaba Pál. The Genomic Landscape of Compensatory Evolution. *PLoS Biology*, 12(8):e1001935, 8 2014. ISSN 1545-7885. . URL https://dx.plos.org/10.1371/journal.pbio.1001935www.plosbiology.orghttp://www.ncbi.nlm.nih.gov/pubmed/25157590.
10. Diana Blank, Luise Wolf, Martin Ackermann, and Olin K Silander. The predictability of molecular evolution during functional innovation. *Proceedings of the National Academy of Sciences of the United States of America*, 111(8):3044–3049, 2 2014. ISSN 00278424. . URL https://www.pnas.org/doi/abs/10.1073/pnas.1318797111.
11. Leland H Hartwell, John J Hopfield, Stanislas Leibler, and Andrew W Murray. From molecular to modular cell biology. *Nature*, 402(S6761):C47–C52, 12 1999. ISSN 0028-0836. . URL https://www.nature.com/articles/35011540http://www.nature.com/articles/35011540.
12. Jing-Dong J Han, Nicolas Bertin, Tong Hao, Debra S Goldberg, Gabriel F Berriz, Lan V Zhang, Denis Dupuy, Albertha J M Walhout, Michael E Cusick, Frederick P Roth, and Marc Vidal. Evidence for dynamically organized modularity in the yeast protein–protein interaction network. *Nature*, 430(6995):88–93, 7 2004. ISSN 0028-0836. . URL https://www.nature.com/articles/nature02555http://www.nature.com/articles/nature02555.
13. Xiao Chang, Tao Xu, Yun Li, and Kai Wang. Dynamic modular architecture of protein-protein interaction networks beyond the dichotomy of 'date' and 'party' hubs. *Scientific Reports*, 3(1):1691, 4 2013. ISSN 2045-2322. . URL https://www.nature.com/articles/srep01691.
14. Günter P. Wagner, Mihaela Pavlicev, and James M. Cheverud. The road to modularity. *Nature Reviews Genetics*, 8(12):921–931, 12 2007. ISSN 1471-0056. . URL https://www.nature.com/articles/nrg2267.
15. Béla Szamecz, Gábor Boross, Dorottya Kalapis, Károly Kovács, Gergely Fekete, Zoltán Farkas, Viktória Lázár, Mónika Hrtyan, Patrick Kemmeren, Marian J. A. Groot Koerkamp, Edit Rutkai, Frank C. P. Holstege, Balázs Papp, and Csaba Pál. The Genomic Landscape of Compensatory Evolution. *PLoS Biology*, 12(8):e1001935, 8 2014. ISSN 1545-7885. . URL www.plosbiology.orghttp://www.ncbi.nlm.nih.gov/pubmed/25157590https://dx.plos.org/10.1371/journal.pbio.1001935.
16. José I Rojas Echenique, Sergey Kryazhimskiy, Alex N Nguyen Ba, and Michael M Desai. Modular epistasis and the compensatory evolution of gene deletion mutants. *PLOS Genetics*, 15(2):e1007958, 2 2019. ISSN 1553-7404. . URL https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1007958.
17. Nadav Kashtan, Avi E. Mayo, Tomer Kalisky, and Uri Alon. An Analytically Solvable Model for Rapid Evolution of Modular Structure. *PLoS Computational Biology*, 5(4):e1000355, 4 2009. ISSN 1553-7358. . URL https://dx.plos.org/10.1371/journal.pcbi.1000355.
18. Kirsten H. ten Tusscher and Paulien Hogeweg. Evolution of Networks for Body Plan Patterning; Interplay of Modularity, Robustness and Evolvability. *PLoS Computational Biology*, 7(10):e1002208, 10 2011. ISSN 1553-7358. . URL https://dx.plos.org/10.1371/journal.pcbi.1002208.
19. Berend Snel and Martijn A. Huynen. Quantifying Modularity in the Evolution of Biomolecular Systems. *Genome Research*, 14(3):391–397, 3 2004. ISSN 1088-9051. . URL http://genome.cshlp.org/lookup/doi/10.1101/gr.1969504.
20. Erin Styles, Ji Young Youn, Mojca Mattiazzi Usaj, and Brenda Andrews. Functional genomics in the study of yeast cell polarity: moving in the right direction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1629), 11 2013. ISSN 09628436. . URL https://royalsocietypublishing.org/doi/10.1098/rstb.2013.0118.
21. Steven J Altschuler, Sigurd B Angenent, Yanqin Wang, and Lani F Wu. On the spontaneous emergence of cell polarity. *Nature*, 454(7206):886–889, 8 2008. ISSN 0028-0836. . URL https://www.nature.com/articles/nature07119http://www.nature.com/articles/nature07119.
22. Andrew B Goryachev and Marcin Leda. Many roads to symmetry breaking: molecular mechanisms and theoretical models of yeast cell polarity. *Molecular Biology of the Cell*, 28(3):370–380, 2 2017. ISSN 1059-1524. . URL https://www.molbiolcell.org/doi/10.1091/mbc.e16-10-0739http://www.ncbi.nlm.nih.gov/pubmed/28137950http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5341721.
23. Barry J Thompson. Cell polarity: models and mechanisms from yeast, worms and flies. *Development*, 140(1):13–21, 1 2013. ISSN 0950-1991. . URL https://journals.biologists.com/dev/article/140/1/13/76512/Cell-polarity-models-and-mechanisms-from-yeast.
24. Sandrine Etienne-Manneville. Cdc42 - the centre of polarity. *Journal of Cell Science*, 117(8):1291–1300, 2004. ISSN 0021-9533. . URL https://journals.biologists.com/jcs/article/117/8/1291/28257/Cdc42-the-centre-of-polarity.
25. W James Nelson. Adaptation of core mechanisms to generate cell polarity. *Nature*, 422(6933):766–774, 4 2003. ISSN 0028-0836. . URL https://www.nature.com/articles/nature01602http://www.nature.com/articles/nature01602.
26. Eugenio Marco, Roland Wedlich-Soldner, Rong Li, Steven J. Altschuler, and Lani F. Wu. Endocytosis Optimizes the Dynamic Localization of Membrane Proteins that Regulate Cortical Polarity. *Cell*, 129(2):411–422, 4 2007. ISSN 0092-8674. .
27. Benjamin Woods and Daniel J Lew. Polarity establishment by Cdc42: Key roles for positive feedback and differential mobility. *Small GTPases*, 10(2):130, 2019. ISSN 21541256. . URL /pmc/articles/PMC6380280//pmc/articles/PMC6380280/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6380280/.
28. Antonio Casamayor and Michael Snyder. Bud-site selection and cell polarity in budding yeast. *Current Opinion in Microbiology*, 5(2):179–186, 4 2002. ISSN 1369-5274. .
29. Tina Freisinger, Ben Klünder, Jared Johnson, Nikola Müller, Garwin Pichler, Gisela Beck, Michael Costanzo, Charles Boone, Richard A Cerione, Erwin Frey, and Roland Wedlich-Söldner. Establishment of a robust single axis of cell polarity by coupling multiple positive feedback loops. *Nature Communications*, 4(1):1807, 2013. ISSN 2041-1723. . URL https://www.nature.com/articles/ncomms2795.
30. Javier E Irazoqui, Amy S Gladfelter, and Daniel J Lew. Scaffold-mediated symmetry breaking by Cdc42p. *Nature Cell Biology*, 5(12):1062–1070, 12 2003. ISSN 1465-7392. . URL https://www.nature.com/articles/ncb1068http://www.nature.com/articles/ncb1068.
31. Ben Klünder, Tina Freisinger, Roland Wedlich-Söldner, and Erwin Frey. GDI-Mediated Cell Polarization in Yeast Provides Precise Spatial and Temporal Control of Cdc42 Signaling. *PLOS Computational Biology*, 9(12):e1003396, 12 2013. ISSN 1553-7358. . URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003396.
32. Lukasz Kozubowski, Koji Saito, Jayme M Johnson, Audrey S Howell, Trevin R Zyla, and Daniel J Lew. Symmetry-Breaking Polarization Driven by a Cdc42p GEF-PAK Complex. *Current Biology*, 18(22):1719–1726, 11 2008. ISSN 09609822. . URL https://linkinghub.elsevier.com/retrieve/pii/S0960982208012955.
33. Liedewij Laan, John H Koschwanez, and Andrew W Murray. Evolutionary adaptation after crippling cell polarization follows reproducible trajectories. *eLife*, 4, 2015. .
34. Yi Zheng, Richard Cerione, and Alan Bender. Control of the yeast bud-site assembly GTPase Cdc42. Catalysis of guanine nucleotide exchange by Cdc24 and stimulation of GTPase activity by Bem3. *Journal of Biological Chememistry*, 1994.
35. Michèle Knaus, Marie-Pierre Pelli-Gulli, Frank van Drogen, Sander Springer, Malika Jaquenoud, and Matthias Peter. Phosphorylation of Bem2p and Bem3p may contribute to local activation of Cdc42p at bud emergence. *The EMBO Journal*, 26(21):4501–4513, 10 2007. ISSN 0261-4189. . URL http://emboj.embopress.org/cgi/doi/10.1038/sj.emboj.7601873.
36. Hunter B Fraser, Dennis P Wall, and Aaron E Hirsh. A simple dependence between protein evolution rate and the number of protein-protein interactions. *BMC evolutionary biology*, 3:11, 5 2003. ISSN 1471-2148. . URL http://www.ncbi.nlm.nih.gov/pubmed/12769820http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC166126.
37. David Alvarez-Ponce, Felix Feyertag, and Sandip Chakraborty. Position Matters: Network Centrality Considerably Impacts Rates of Protein Evolution in the Human Protein–Protein Interaction Network. *Genome Biology and Evolution*, 9(6):1742–1756, 6 2017. ISSN 17596653. . URL https://academic.oup.com/gbe/article/9/6/1742/3922884.
38. Ricard Albalat and Cristian Cañestro. Evolution by gene loss. *Nature Reviews Genetics*, 17(7):379–391, 7 2016. ISSN 1471-0056. . URL https://www.nature.com/articles/nrg.2016.39.
39. Fridtjof Brauns, Leila Iñigo de la Cruz, Werner K.-G. Daalman, Ilse de Bruin, Jacob Halatek, Liedewij Laan, and Erwin Frey. Redundancy and the role of protein copy numbers in the cell polarization machinery of budding yeast. *Nature Communications*, 14(1):6504, 10 2023. ISSN 2041-1723. . URL https://www.nature.com/articles/s41467-023-42100-0.
40. Ruth Nussinov, Yonglan Liu, Wengang Zhang, and Hyunbum Jang. Protein conformational ensembles in function: roles and mechanisms. *RSC Chemical Biology*, 4(11):850–864, 2023. ISSN 2633-0679. . URL http://xlink.rsc.org/?DOI=D3CB00114H.
41. Bin Xue, A. Keith Dunker, and Vladimir N. Uversky. Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *Journal of Biomolecular Structure and Dynamics*, 30(2):137–149, 6 2012. ISSN 0739-1102. . URL http://www.tandfonline.com/doi/abs/10.1080/07391102.2012.675145.
42. Marton Miskei, Andrea Gregus, Rashmi Sharma, Norbert Duro, Fruzsina Zsolyomi, and Monika Fuxreiter. Fuzziness enables context dependence of protein interactions. *FEBS Letters*, 591(17):

2682–2695, 9 2017. ISSN 0014-5793. . URL https://febs.onlinelibrary.wiley.com/doi/10.1002/1873-3468.12762.

43. Agnès H. Michel, Riko Hatakeyama, Philipp Kimmig, Meret Arter, Matthias Peter, Joao Matos, Claudio De Virgilio, and Benoît Kornmann. Functional mapping of yeast genomes by saturated transposition. *eLife*, 6, 5 2017. ISSN 2050-084X. . URL https://elifesciences.org/articles/23570.
44. H Jeong, S P Mason, A.-L. Barabási, and Z N Oltvai. Lethality and centrality in protein networks. *Nature*, 411(6833):41–42, 2001. ISSN 0028-0836. . URL https://www.nature.com/articles/35075138http://www.nature.com/articles/35075138.
45. Hunter B. Fraser, Aaron E. Hirsh, Lars M. Steinmetz, Curt Scharfe, and Marcus W. Feldman. Evolutionary Rate in the Protein Interaction Network. *Science*, 296(5568):750–752, 4 2002. ISSN 0036-8075. . URL https://www.science.org/doi/10.1126/science.1068696http://www.ncbi.nlm.nih.gov/pubmed/11976460.
46. Nizar N Batada, Laurence D Hurst, and Mike Tyers. Evolutionary and Physiological Importance of Hub Proteins. *PLOS Computational Biology*, 2(7):e88, 2006. ISSN 1553-7358. . URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.0020088.
47. Xionglei He and Jianzhi Zhang. Why Do Hubs Tend to Be Essential in Protein Networks? *PLOS Genetics*, 2(6):e88, 2006. ISSN 1553-7404. . URL https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.0020088.
48. C. Stark. BioGRID: a general repository for interaction datasets. *Nucleic Acids Research*, 34(90001):D535–D539, 1 2006. ISSN 0305-1048. . URL https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkj109.
49. Rose Oughtred, Jennifer Rust, Christie Chang, Bobby-Joe Breitkreutz, Chris Stark, Andrew Willems, Lorrie Boucher, Genie Leung, Nadine Kolas, Frederick Zhang, Sonam Dolma, Jasmin Coulombe-Huntington, Andrew Chatr-aryamontri, Kara Dolinski, and Mike Tyers. The <scp>BioGRID</scp> database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Science*, 30(1):187–200, 1 2021. ISSN 0961-8368. . URL https://onlinelibrary.wiley.com/doi/10.1002/pro.3978.
50. Mikaela Koutrouli, Evangelos Karatzas, David Paez-Espino, and Georgios A Pavlopoulos. A Guide to Conquer the Biological Network Era Using Graph Theory. *Frontiers in Bioengineering and Biotechnology*, 8:34, 1 2020. ISSN 2296-4185. . URL http://www.ncbi.nlm.nih.gov/pubmed/32083072http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC7004966https://www.frontiersin.org/article/10.3389/fbioe.2020.00034/full.
51. Damian Szklarczyk, Annika L Gable, Katerina C Nastou, David Lyon, Rebecca Kirsch, Sampo Pyysalo, Nadezhda T Doncheva, Marc Legeay, Tao Fang, Peer Bork, Lars J Jensen, and Christian von Mering. The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research*, 49(D1):D605–D612, 1 2021. ISSN 0305-1048. . URL https://academic.oup.com/nar/article/49/D1/D605/6006194.
52. Nadezhda T Doncheva, John H Morris, Jan Gorodkin, and Lars J Jensen. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *Journal of Proteome Research*, 18(2):623–632, 2 2019. ISSN 1535-3893. . URL https://pubs.acs.org/doi/full/10.1021/acs.jproteome.8b00702http://www.ncbi.nlm.nih.gov/pubmed/30450911http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6800166https://pubs.acs.org/doi/10.1021/acs.jproteome.8b00702.
53. Damian Szklarczyk, Annika L. Gable, David Lyon, Alexander Junge, Stefan Wyder, Jaime Huerta-Cepas, Milan Simonovic, Nadezhda T. Doncheva, John H. Morris, Peer Bork, Lars J. Jensen, and Christian von Mering. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, 47(D1):D607–D613, 1 2019. ISSN 0305-1048. . URL https://pubmed.ncbi.nlm.nih.gov/30476243/https://academic.oup.com/nar/article/47/D1/D607/5198476http://www.ncbi.nlm.nih.gov/pubmed/30476243http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6323986.
54. Stijn Van Dongen. Graph Clustering Via a Discrete Uncoupling Process. *SIAM Journal on Matrix Analysis and Applications*, 30(1):121–141, 1 2008. ISSN 0895-4798. . URL https://doi.org/10.1137/040608635http://epubs.siam.org/doi/10.1137/040608635.
55. Daniel Segrè, Alexander DeLuna, George M Church, and Roy Kishony. Modular epistasis in yeast metabolism. *Nature Genetics*, 37(1):77–83, 1 2005. ISSN 1061-4036. . URL https://www.ncbi.nlm.nih.gov/pubmed/15592468http://www.nature.com/articles/ng1489.
56. Janet L. Carminati and Tim Stearns. Microtubules Orient the Mitotic Spindle in Yeast through Dynein-dependent Interactions with the Cell Cortex. *The Journal of Cell Biology*, 138(3):629, 8 1997. ISSN 00219525. . URL /pmc/articles/PMC2141630//pmc/articles/PMC2141630/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC2141630/.
57. Istvan R. Boldogh, Hyeong-Cheol Yang, and Liza A. Pon. Mitochondrial Inheritance in Budding Yeast. *Traffic*, 2(6):368–374, 6 2001. ISSN 1398-9219. . URL https://onlinelibrary.wiley.com/doi/10.1034/j.1600-0854.2001.002006368.x.
58. Werner Daalman, Els Sweep, and Liedewij Laan. The Path towards Predicting Evolution as Illustrated in Yeast Cell Polarity. *Cells*, 9(12):2534, 11 2020. ISSN 2073-4409. . URL https://www.mdpi.com/2073-4409/9/12/2534.
59. Benjamin Woods, Helen Lai, Chi Fang Wu, Trevin R Zyla, Natasha S Savage, and Daniel J Lew. Parallel Actin-Independent Recycling Pathways Polarize Cdc42 in Budding Yeast. *Current Biology*, 26(16):2114–2126, 8 2016. ISSN 09609822. . URL http://www.cell.com/article/S0960982216306881/fulltexthttp://www.cell.com/article/S0960982216306881/abstracthttps://www.cell.com/current-biology/abstract/S0960-9822(16)30688-1.
60. Javier E Irazoqui, Audrey S Howell, Chandra L Theesfeld, and Daniel J Lew. Opposing Roles for Actin in Cdc42p Polarization. *Molecular Biology of the Cell*, 16(3):1296–1304, 3 2005. ISSN 1059-1524. . URL https://www.molbiolcell.org/doi/10.1091/mbc.e04-05-0430http://www.ncbi.nlm.nih.gov/pubmed/15616194http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC551493.
61. Roland Wedlich-Soldner, Stephanie C. Wai, Thomas Schmidt, and Rong Li. Robust cell polarity is a dynamic state established by coupling transport and GTPase signaling. *Journal of Cell Biology*, 166(6):889–900, 9 2004. ISSN 00219525. . URL http://www.jcb.org/cgi/doi/10.1083/jcb.200405061.
62. Brian D Slaughter, Arupratan Das, Joel W Schwartz, Boris Rubinstein, and Rong Li. Dual Modes of Cdc42 Recycling Fine-Tune Polarized Morphogenesis. *Developmental Cell*, 17(6):823–835, 12 2009. ISSN 15345807. . URL https://linkinghub.elsevier.com/retrieve/pii/S1534580709004456.
63. Brian D. Slaughter, Jay R. Unruh, Arupratan Das, Sarah E. Smith, Boris Rubinstein, and Rong Li. Non-uniform membrane diffusion enables steady-state cell polarization via vesicular trafficking. *Nature Communications*, 4(1):1380, 1 2013. ISSN 2041-1723. . URL https://www.nature.com/articles/ncomms2370.
64. Elodie Caudal, Anne Friedrich, Arthur Jallet, Marion Garin, Jing Hou, and Joseph Schacherer. Loss-of-function mutation survey revealed that genes with background-dependent fitness are rare and functionally related in yeast. *Proceedings of the National Academy of Sciences*, 119(37):e2204206119, 9 2022. ISSN 0027-8424. . URL https://www.pnas.org/doi/abs/10.1073/pnas.2204206119https://pnas.org/doi/full/10.1073/pnas.2204206119.
65. Marco Galardini, Bede P Busby, Cristina Vieitez, Alistair S Dunham, Athanasios Typas, and Pedro Beltrao. The impact of the genetic background on gene deletion phenotypes in Saccharomyces cerevisiae. *Molecular Systems Biology*, 15(12):e8831, 12 2019. ISSN 1744-4292. . URL /pmc/articles/PMC6901017//pmc/articles/PMC6901017/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6901017/https://onlinelibrary.wiley.com/doi/full/10.15252/msb.20198831https://onlinelibrary.wiley.com/doi/abs/10.15252/msb.20198831https://.
66. Marie Filteau, Véronique Hamel, Marie-Christine Pouliot, Isabelle Gagnon-Arsenault, Alexandre K Dubé, and Christian R Landry. Evolutionary rescue by compensatory mutations is constrained by genomic and environmental backgrounds. *Molecular Systems Biology*, 11(10):832, 2015. ISSN 1744-4292. .
67. Martin N. Mullis, Takeshi Matsui, Rachel Schell, Ryan Foree, and Ian M. Ehrenreich. The complex underpinnings of genetic background effects. *Nature Communications*, 9(1):3548, 9 2018. ISSN 2041-1723. . URL https://www.nature.com/articles/s41467-018-06023-5.
68. Michael Costanzo, Jing Hou, Vincent Messier, Justin Nelson, Mahfuzur Rahman, Benjamin VanderSluis, Wen Wang, Carles Pons, Catherine Ross, Matej Ušaj, Bryan-Joseph San Luis, Emira Shuteriqi, Elizabeth N Koch, Patrick Aloy, Chad L Myers, Charles Boone, and Brenda Andrews. Environmental robustness of the global yeast genetic interaction network. *Science*, 372(6542), 2021. ISSN 0036-8075. . URL https://doi.org/10.1126/science.abf8424https://www.science.org/doi/10.1126/science.abf8424.
69. Christian von Mering, Roland Krause, Berend Snel, Michael Cornell, Stephen G. Oliver, Stanley Fields, and Peer Bork. Comparative assessment of large-scale data sets of protein–protein interactions. *Nature*, 417(6887):399–403, 5 2002. ISSN 0028-0836. . URL https://www.nature.com/articles/nature750http://www.nature.com/articles/nature750.
70. Jesse Gillis, Sara Ballouz, and Paul Pavlidis. Bias tradeoffs in the creation and analysis of protein–protein interaction networks. *Journal of Proteomics*, 100:44–54, 4 2014. ISSN 18743919. . URL /pmc/articles/PMC3972268//pmc/articles/PMC3972268/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC3972268/https://linkinghub.elsevier.com/retrieve/pii/S1874391914000384.
71. Teunis J P Van Dam and Berend Snel. Protein Complex Evolution Does Not Involve Extensive Network Rewiring. *PLOS Computational Biology*, 4(7):e1000132, 2008. ISSN 1553-7358. . URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1000132.
72. Marco Y. Hein, Nina C. Hubner, Ina Poser, Jürgen Cox, Nagarjuna Nagaraj, Yusuke Toyoda, Igor A. Gak, Ina Weisswange, Jörg Mansfeld, Frank Buchholz, Anthony A. Hyman, and Matthias Mann. A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances. *Cell*, 163(3):712–723, 10 2015. ISSN 00928674. . URL https://linkinghub.elsevier.com/retrieve/pii/S0092867415012702.
73. Russell M Gordley, Lukasz J Bugaj, and Wendell A Lim. Modular engineering of cellular signaling proteins and networks. *Current Opinion in Structural Biology*, 39:106–114, 8 2016. ISSN 0959440X. . URL https://linkinghub.elsevier.com/retrieve/pii/S0959440X16300690.
74. W. R. Harcombe, R. Springman, and J. J. Bull. Compensatory evolution for a gene deletion is not limited to its immediate functional network. *BMC Evolutionary Biology*, 9(1):1–11, 5 2009. ISSN 14712148. . URL https://bmcecolevol.biomedcentral.com/articles/10.1186/1471-2148-9-106.
75. Hay-Oak Park, Erfei Bi, John R. Pringle, and Ira Herskowitz. Two active states of the Ras-related Bud1/Rsr1 protein bind to different effectors to determine yeast cell polarity. *Proceedings of the National Academy of Sciences*, 94(9):4463–4468, 4 1997. ISSN 0027-8424. . URL https://pnas.org/doi/full/10.1073/pnas.94.9.4463.
76. Júlia Domingo, Pablo Baeza-Centurion, and Ben Lehner. The Causes and Consequences of Genetic Interactions (Epistasis). *Annual Review of Genomics and Human Genetics*, 20(1):433–460, 8 2019. ISSN 1527-8204. . URL https://www.annualreviews.org/doi/abs/10.1146/annurev-genom-083118-014857https://www.annualreviews.org/doi/10.1146/annurev-genom-083118-014857.
77. Jolanda van Leeuwen, Carles Pons, Charles Boone, and Brenda J. Andrews. Mechanisms of suppression: The wiring of genetic resilience. *BioEssays*, 39(7), 7 2017. ISSN 0265-9247. . URL https://onlinelibrary.wiley.com/doi/10.1002/bies.201700042.
78. Ruth Nussinov, Chung-Jung Tsai, and Hyunbum Jang. Protein ensembles link genotype to phenotype. *PLOS Computational Biology*, 15(6):e1006648, 6 2019. ISSN 1553-7358. . URL https://dx.plos.org/10.1371/journal.pcbi.1006648.

79. Petra Schwille and Béla P. Frohn. Hidden protein functions and what they may teach us. *Trends in Cell Biology*, 32(2):102–109, 2 2022. ISSN 09628924. . URL https://linkinghub.elsevier.com/retrieve/pii/S0962892421001859.
80. David D Boehr, Ruth Nussinov, and Peter E Wright. The role of dynamic conformational ensembles in biomolecular recognition. *Nature Chemical Biology*, 5(11):789–796, 11 2009. ISSN 1552-4450. . URL https://www.nature.com/articles/nchembio.232.
81. Zoltán Farkas, Károly Kovács, Zsuzsa Sarkadi, Dorottya Kalapis, Gergely Fekete, Fanni Birtyik, Ferhan Ayaydin, Csaba Molnár, Péter Horváth, Csaba Pál, and Balázs Papp. Gene loss and compensatory evolution promotes the emergence of morphological novelties in budding yeast. *Nature Ecology & Evolution*, 6(6):763–773, 4 2022. ISSN 2397-334X. . URL https://www.nature.com/articles/s41559-022-01730-1.
82. Jana Helsen, Karin Voordeckers, Laura Vanderwaeren, Toon Santermans, Maria Tsontaki, Kevin J Verstrepen, and Rob Jelier. Gene Loss Predictably Drives Evolutionary Adaptation. *Molecular Biology and Evolution*, 37(10):2989–3002, 2020. ISSN 0737-4038. . URL https://academic.oup.com/mbe/article/37/10/2989/5870834.
83. Cristina Guijarro-Clarke, Peter W. H. Holland, and Jordi Paps. Widespread patterns of gene loss in the evolution of the animal kingdom. *Nature Ecology & Evolution*, 4(4):519–523, 2 2020. ISSN 2397-334X. . URL https://www.nature.com/articles/s41559-020-1129-2.
84. Cristian Cañestro and John H. Postlethwait. Development of a chordate anterior–posterior axis without classical retinoic acid signaling. *Developmental Biology*, 305(2):522–538, 5 2007. ISSN 00121606. . URL https://linkinghub.elsevier.com/retrieve/pii/S0012160607001492.
85. Andrew W Murray. Can gene-inactivating mutations lead to evolutionary novelty? *Current Biology*, 30(10):R465–R471, 2020. ISSN 0960-9822. . URL http://www.cell.com/article/S0960982220304796/fulltexthttp://www.cell.com/article/S0960982220304796/abstracthttps://www.cell.com/current-biology/abstract/S0960-9822(20)30479-6.
86. Alison K Hottes, Peter L Freddolino, Anupama Khare, Zachary N Donnell, Julia C Liu, and Saeed Tavazoie. Bacterial Adaptation through Loss of Function. *PLoS Genetics*, 9(7):e1003617, 7 2013. ISSN 1553-7404. . URL https://www.ncbi.nlm.nih.gov/pubmed/23874220https://dx.plos.org/10.1371/journal.pgen.1003617.
87. James E. Dicarlo, Julie E. Norville, Prashant Mali, Xavier Rios, John Aach, and George M. Church. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Research*, 41(7):4336–4343, 4 2013. ISSN 0305-1048. . URL https://academic.oup.com/nar/article/41/7/4336/1075252.
88. R. Daniel Gietz and Robert H. Schiestl. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*, 2(1):31–34, 1 2007. ISSN 1754-2189. . URL https://www.nature.com/articles/nprot.2007.13http://www.nature.com/articles/nprot.2007.13.
89. Robert Mans, Harmen M. van Rossum, Melanie Wijsman, Antoon Backx, Niels G.A. Kuijpers, Marcel van den Broek, Pascale Daran-Lapujade, Jack T. Pronk, Antonius J.A. van Maris, and Jean-Marc G. Daran. CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in Saccharomyces cerevisiae. *FEMS Yeast Research*, 15(2):1–21, 3 2015. ISSN 1567-1356. . URL http://academic.oup.com/femsyr/article/doi/10.1093/femsyr/fov004/536897/CRISPRCas9-a-molecular-Swiss-army-knife-for.
90. Kornel Labun, Tessa G Montague, Maximilian Krause, Yamila N Torres Cleuren, Håkon Tjeldnes, and Eivind Valen. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Research*, 47(W1):W171–W174, 7 2019. ISSN 0305-1048. . URL https://academic.oup.com/nar/article/47/W1/W171/5491735.
91. Agnès H Michel, Sabine van Schie, Andreas Mosbach, Gabriel Scalliet, and Benoît Kornmann. Exploiting homologous recombination increases SATAY efficiency for loss- and gain-of-function screening. *bioRxiv*, page 866483, 12 2019. . URL https://www.biorxiv.org/content/10.1101/866483v1https://www.biorxiv.org/content/10.1101/866483v1.abstract.
92. Agnès H Michel, Riko Hatakeyama, Philipp Kimmig, Meret Arter, Matthias Peter, Joao Matos, Claudio De Virgilio, and Benoît Kornmann. Functional mapping of yeast genomes by saturated transposition. *eLife*, 6, 5 2017. ISSN 2050-084X. . URL https://elifesciences.org/articles/23570.
93. Agnès H Michel and Benoît Kornmann. SAturated Transposon Analysis in Yeast (SATAY) for Deep Functional Mapping of Yeast Genomes. In *Methods in molecular biology (Clifton, N.J.)*, volume 2477, pages 349–379. Humana Press Inc., 2022. . URL https://link.springer.com/protocol/10.1007/978-1-0716-2257-5_20http://www.ncbi.nlm.nih.gov/pubmed/35524127https://link.springer.com/10.1007/978-1-0716-2257-5_20.
94. Brian Bushnell, Jonathan Rood, and Esther Singer. BBMerge – Accurate paired shotgun read merging via overlap. *PLOS ONE*, 12(10):e0185056, 10 2017. ISSN 1932-6203. .
95. Leila Iñigo de la Cruz, Gregory van Beek, and Maurits Kok. transposonmapper, 1 2022. URL https://doi.org/10.5281/zenodo.5903081#.Y6L8bYdxJII.mendeley.
96. Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J van der Walt, Matthew Brett, Joshua Wilson, K Jarrod Millman, Nikolay Mayorov, Andrew R J Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E A Quintero, Charles R Harris, Anne M Archibald, Antônio H Ribeiro, Fabian Pedregosa, Paul van Mulbregt, Aditya Vijaykumar, Alessandro Pietro Bardelli, Alex Rothberg, Andreas Hilboll, Andreas Kloeckner, Anthony Scopatz, Antony Lee, Ariel Rokem, C Nathan Woods, Chad Fulton, Charles Masson, Christian Häggström, Clark Fitzgerald, David A Nicholson, David R Hagen, Dmitrii V Pasechnik, Emanuele Olivetti, Eric Martin, Eric Wieser, Fabrice Silva, Felix Lenders, Florian Wilhelm, G Young, Gavin A Price, Gert-Ludwig Ingold, Gregory E Allen, Gregory R Lee, Hervé Audren, Irvin Probst, Jörg P Dietrich, Jacob Silterra, James T Webber, Janko Slavič, Joel Nothman, Johannes Buchner, Johannes Kulick, Johannes L Schönberger, José Vinícius de Miranda Cardoso, Joscha Reimer, Joseph Harrington, Juan Luis Cano Rodríguez, Juan Nunez-Iglesias, Justin Kuczynski, Kevin Tritz, Martin Thoma, Matthew Newville, Matthias Kümmerer, Maximilian Bolingbroke, Michael Tartre, Mikhail Pak, Nathaniel J Smith, Nikolai Nowaczyk, Nikolay Shebanov, Oleksandr Pavlyk, Per A Brodtkorb, Perry Lee, Robert T McGibbon, Roman Feldbauer, Sam Lewis, Sam Tygier, Scott Sievert, Sebastiano Vigna, Stefan Peterson, Surhud More, Tadeusz Pudlik, Takuya Oshima, Thomas J Pingel, Thomas P Robitaille, Thomas Spura, Thouis R Jones, Tim Cera, Tim Leslie, Tiziano Zito, Tom Krauss, Utkarsh Upadhyay, Yaroslav O Halchenko, and Yoshiki Vázquez-Baeza. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272, 2020. ISSN 1548-7091. . URL https://www.nature.com/articles/s41592-019-0686-2http://www.nature.com/articles/s41592-019-0686-2.
97. Michael A DeJesus, Chaitra Ambadipudi, Richard Baker, Christopher Sassetti, and Thomas R Ioerger. TRANSIT - A Software Tool for Himar1 TnSeq Analysis. *PLOS Computational Biology*, 11(10):e1004401, 2015. ISSN 1553-7358. . URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004401https://dx.plos.org/10.1371/journal.pcbi.1004401.
98. J. Michael Cherry, Eurie L. Hong, Craig Amundsen, Rama Balakrishnan, Gail Binkley, Esther T. Chan, Karen R. Christie, Maria C. Costanzo, Selina S. Dwight, Stacia R. Engel, Dianna G. Fisk, Jodi E. Hirschman, Benjamin C. Hitz, Kalpana Karra, Cynthia J. Krieger, Stuart R. Miyasato, Rob S. Nash, Julie Park, Marek S. Skrzypek, Matt Simison, Shuai Weng, and Edith D. Wong. Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic Acids Research*, 40(D1):D700–D705, 1 2012. ISSN 0305-1048. . URL https://academic.oup.com/nar/article/40/D1/D700/2902744https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkr1029.
99. Paul Shannon, Andrew Markiel, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, Daniel Ramage, Nada Amin, Beno Schwikowski, and Trey Ideker. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11):2498–2504, 11 2003. ISSN 10889051. . URL http://genome.cshlp.org/lookup/doi/10.1101/gr.1239303.
100. Aric A Hagberg, Daniel A Schult, and Pieter J Swart. Exploring Network Structure, Dynamics, and Function using NetworkX. In Gaël Varoquaux, Travis Vaught, and Jarrod Millman, editors, *Proceedings of the 7th Python in Science Conference*, pages 11–15, Pasadena, CA USA, 2008.
101. Michael A. DeJesus and Thomas R. Ioerger. Normalization of transposon-mutant library sequencing datasets to improve identification of conditionally essential genes. *Journal of Bioinformatics and Computational Biology*, 14(03):1642004, 6 2016. ISSN 0219-7200. . URL https://www.worldscientific.com/doi/abs/10.1142/S021972001642004X.
102. Michael A DeJesus and Thomas R Ioerger. A Hidden Markov Model for identifying essential and growth-defect regions in bacterial genomes from transposon insertion sequencing data. *BMC Bioinformatics*, 14(1):303, 12 2013. ISSN 1471-2105. . URL https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-14-303.
103. Simon Anders and Wolfgang Huber. Differential expression analysis for sequence count data. *Genome Biology*, 11(10):R106, 10 2010. ISSN 1474-760X. . URL https://genomebiology.biomedcentral.com/articles/10.1186/gb-2010-11-10-r106http://www.ncbi.nlm.nih.gov/pubmed/20979621http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3218662.
104. Mark D Robinson and Alicia Oshlack. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, 11(3):R25, 2010. ISSN 1465-6906. . URL https://genomebiology.biomedcentral.com/articles/10.1186/gb-2010-11-3-r25http://www.ncbi.nlm.nih.gov/pubmed/20196867http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2864565http://genomebiology.biomedcentral.com/articles/10.1186/gb-2010-11-3.

6. Supplement

A. Read count correction and normalization procedure. The benefit of including read counts in our metric to identify changes in the fitness effect of gene disruptions across genetic backgrounds is that it allows for a more sensitive analysis than those based on insertion counts. For example, a gene may have a negligible difference in insertion counts between two genetic backgrounds if the change in gene disruption tolerance is not so drastic that it causes a large number of mutants to fall below the detection level in one of the two backgrounds. However, if a significant difference in gene disruption tolerance truly exists, we would expect that the read counts associated with each insertion site (which reflect mutant abundance) will be lower in one of the two strains. Thus, extending the analysis from a binary assessment of the occupancy of insertion sites to a gradual scale based on read counts can uncover differences that would otherwise remain undetected. However, a drawback is that read counts tend to be more variable between replicate experiments than insertion counts, often owing to read count spikes caused by technical and biological artifacts (101). This increased noise level can hinder the desired increase in sensitivity of the analysis. In addition, spikes in the read counts can significantly skew the read count distribution. This means that the small subset of insertion sites with an artificially high read count will strongly affect methods that aim to correct differences in sequencing depth between libraries by a linear transformation of the total read count.

To address these issues, we implemented the Beta-Geometric correction (BGC) method developed by DeJesus and Ioerger (102). This method is based on the observation that read count distributions obtained from transposon insertion sequencing (TIS) experiments resemble a geometric distribution where most insertion sites have a low read count, while only few sites contain many reads. During the BGC correction procedure, the quantiles of the empirical read count distribution are adjusted to match the quantiles of a fitted 'ideal' geometric distribution. To allow for greater flexibility of the fit, this ideal distribution is implemented as a mixture of geometric distributions with a Beta prior on the probability parameter. In practice, this procedure corrects for the 'skew' in the empirical read count distribution by suppressing spikes while inflating the insertion sites that have a low read count. The results in figure S1 show that the application of this procedure greatly improves the correlation between replicate datasets (figure S1b) compared to uncorrected datasets (figure S1a). Lastly, after applying the BGC correction, we normalized for differences in sequencing depth and library complexity across datasets using the median of ratios normalization (103, 104). In summary, the geometric mean across all samples was calculated for each gene:

$$\bar{X}_g = \exp \frac{1}{n} \cdot \sum_{a}^{n} \ln \left(R_{g,a}\right)). \qquad [1]$$

Where $\bar{X}_g$ is the geometric mean of the read counts mapping to gene g, n is the total number of datasets and $R_{g,a}$ is the number of reads that map to gene g in dataset a. Next, the ratio of the total read count to the geometric mean $\bar{X}_g$ was determined for each sample and the sample-specific normalization factor was taken to be the median of these ratios across all genes:

$$S_a = \text{median}(\frac{R_{g,a}}{\bar{X}_g}) \ \ , \text{ for all } a \in G \qquad [2]$$

With S_a the normalization factor for dataset a and G is the complete set of annotated genes used in this chapter. Finally, the read counts were linearly scaled by the normalization factor S_a:

$$R_{g,a,\text{norm}} = \frac{R_{g,a}}{S_a}. \qquad [3]$$

B. Supplemental figures.

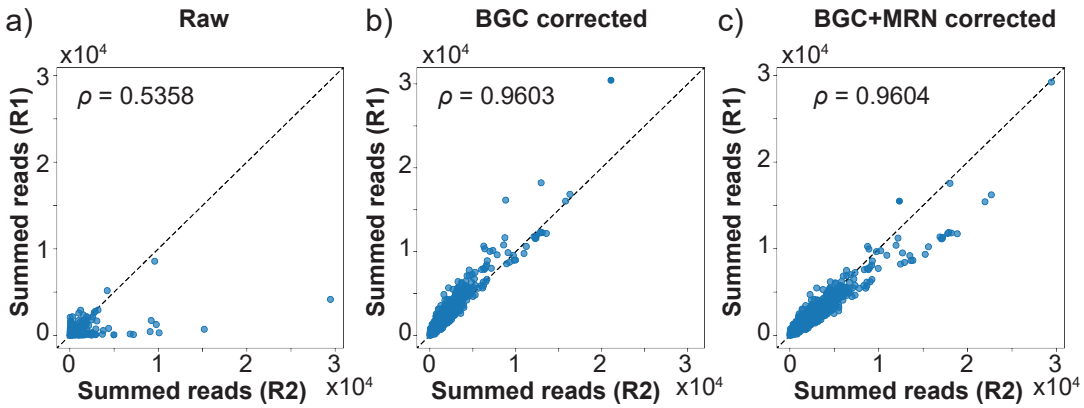

Fig. S1. Normalization procedure increases correlation in the summed read counts per gene between replicate experiments. In order to improve the correlation between replicate datasets derived from the same genetic background, the datasets are first aligned with their best fitting Beta-Geometric distribution (panel (a) to panel (b)) to correct for spikes in read counts, followed by a median of ratios normalization to account for differences in sequencing depth (panel (b) to panel (c)). In each panel, the value of Pearson's correlation coefficient (ρ) is shown.

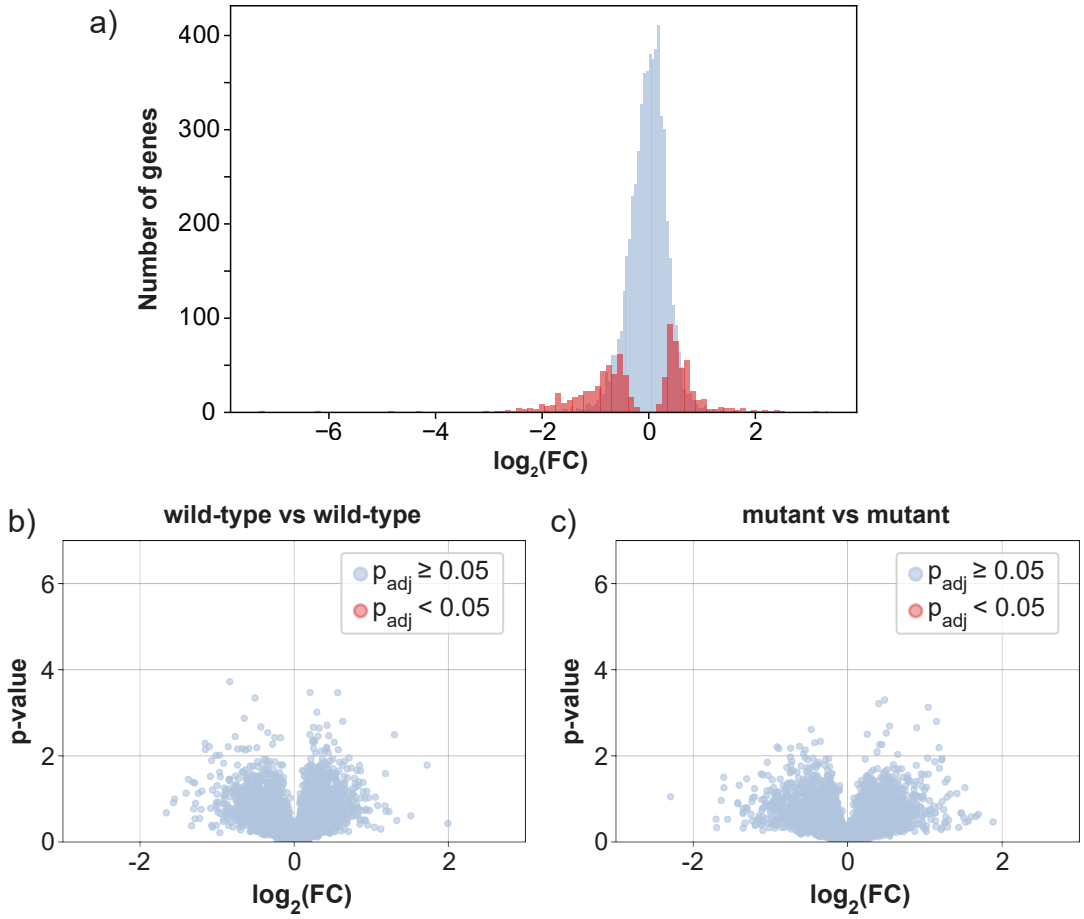

Fig. S2. Validity of the geneset that is identified as having differential fitness between the wild-type strain and the polarity mutant. (**a**) Significant genes have a larger effect size. Histogram of the $\log_2$-fold changes shown in figure 2. The plot shows that genes that are flagged to have a statistically significant difference between the two genetic backgrounds (red bars) typically occurs for larger fold-changes. (**b-c**) Comparing transposon mutagenesis libraries obtained from the same genetic background yields no significant differences in gene fitness. Volcano plots are shown for comparisons between wild-type and mutant datasets. The replicate datasets of each genetic background (6 in total) were split and compared 3 vs. 3. For both our wild-type strain and from our polarity mutant, no false positives are found at a significance threshold of $\mathrm{p_{adj}} < 0.05$.

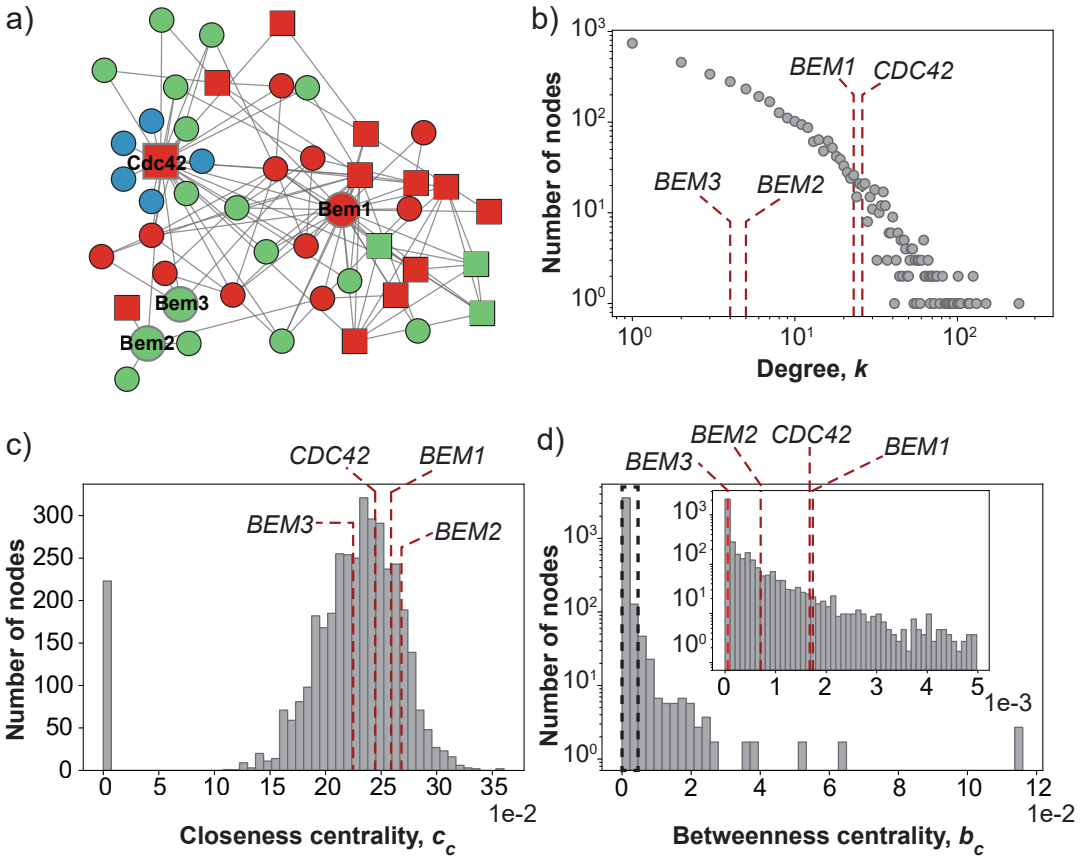

Fig. S3. Properties of the constructed protein-protein interaction network. (**a**) Sub-graph of of the protein-protein interaction network for Cdc42, Bem1, Bem2 and Bem3 and their first neighbours. Nodes are colored according to their degree k in the complete PPI network (figure 3a). Blue: $k \leq 3$, green: $4 \leq k \leq 10$, red: $k > 10$. Proteins are essential according to the SGD database are shown as squares, non-essential proteins are shown as circles. (**b**) The degree centrality of the nodes in the complete PPI network. The degree distribution shows the typical sub-linearity of PPI networks in biology when plotted on a log-log scale. (**c**) The closeness centrality distribution of the PPI network. (**d**) the betweenness centrality distribution of the PPI network. In the panels **c**-**d**, the degree values are indicated for two polarity proteins that have a strong negative effect on fitness when deleted (Bem1 and Cdc42) and for two polarity proteins that have a moderate to weak negative effect (Bem3 and Bem2). With respect to degree and betweenness, proteins with a similar gene fitness lie in proximity of each other on the distributions. For the closeness distribution, we find no relation with gene fitness.

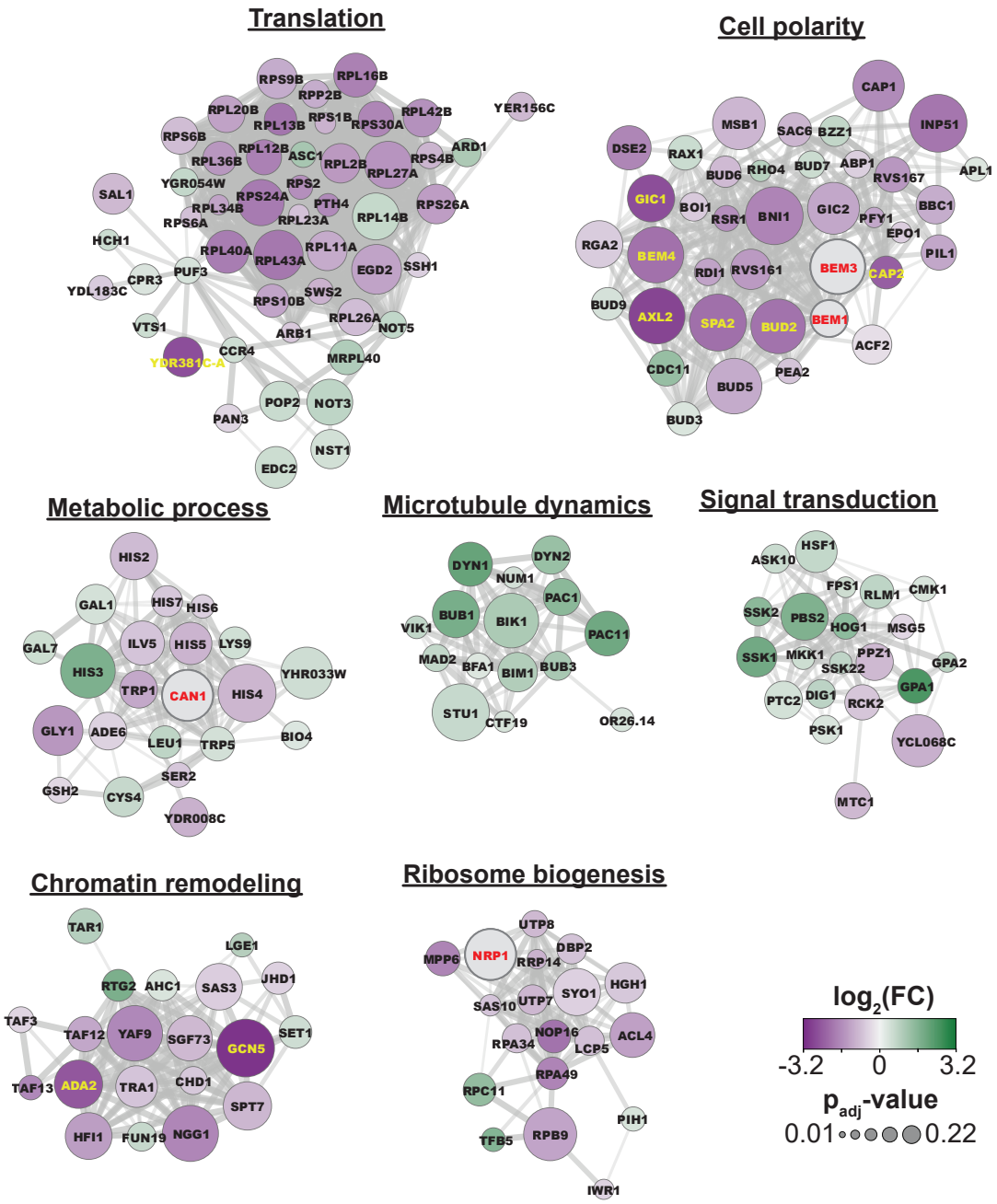

Fig. S4. Zoom-in of the seven largest clusters identified in our functional association network by the Markov clustering algorithm The biological process gene ontology enrichment is shown above each cluster.

C. Supplemental tables.

Table S1. Liquid media used for SATAY library generation. YNB: Yeast Nitrogen Base. CSM: Complete Supplement Mixture. Ura: Uracil. Ade: Adenine

Step	Media	Components
Preculture	SD-Ura+0.2% Glucose +2% Raffinose	• YNB w/o Amino Acids (6.8 g/L) • CSM -Ura (0.77 g/L) • Glucose (2 g/L) • Raffinose (20 g/L) • Adenine (20 mg/L)
Induction	SD-Ura+2% Galactose	• YNB w/o Amino Acids (6.8 g/L) • CSM -Ura (0.77 g/L) • Galactose (20 g/L) • Adenine (20 mg/L)
Reseed	SD-Ade+2% Glucose	• YNB w/o Amino Acids (6.8 g/L) • CSM -Ade (0.78 g/L) • Glucose (20 g/L)

Table S2. List of used primers

#	Name	Barcode	Sequence (5'-3')
1	HT50_688_minidsSEQ1210	HT50	GCC ACA TAT TTA CCG ACC GTT ACC GAC CGT TTT CAT CCC TA
2	E2_HT48_MiniDS_RV	HT48	AGG TCA GTC ACA TGG TTA GGA CGC AGA GCT GAA ACG AAA ACG AAC GGG ATA AA
3	HT60_688_minidsSEQ1210	HT60	TAG GAT GAT TTA CCG ACC GTT ACC GAC CGT TTT CAT CCC TA
4	E2_HT49_MiniDS_RV	HT49	AGG TCA GTC ACA TGG TTA GGA CGC AGA TAG ACA ACG AAA ACG AAC GGG ATA AA