

RESEARCH ARTICLE

Growth condition dependency is the major cause of non-responsiveness upon genetic perturbation

Saman Amini^{1,2}, Frank C. P. Holstege¹, Patrick Kemmeren^{1,2*}

1 Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands, **2** Center for Molecular Medicine, University Medical Centre Utrecht, Utrecht, The Netherlands

* p.kemmeren@prinsesmaximacentrum.nl



OPEN ACCESS

Citation: Amini S, Holstege FCP, Kemmeren P (2017) Growth condition dependency is the major cause of non-responsiveness upon genetic perturbation. PLoS ONE 12(3): e0173432. doi:10.1371/journal.pone.0173432

Editor: Michael Polymenis, Texas A&M University College Station, UNITED STATES

Received: December 23, 2016

Accepted: February 10, 2017

Published: March 3, 2017

Copyright: © 2017 Amini et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are available within the manuscript and corresponding Supporting Information.

Funding: Financial support was provided by the Dutch Organization for Scientific Research (NWO), grant 864.11.010 (SA & PK). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Investigating the role and interplay between individual proteins in biological processes is often performed by assessing the functional consequences of gene inactivation or removal. Depending on the sensitivity of the assay used for determining phenotype, between 66% (growth) and 53% (gene expression) of *Saccharomyces cerevisiae* gene deletion strains show no defect when analyzed under a single condition. Although it is well known that this non-responsive behavior is caused by different types of redundancy mechanisms or by growth condition/cell type dependency, it is not known what the relative contribution of these different causes is. Understanding the underlying causes of and their relative contribution to non-responsive behavior upon genetic perturbation is extremely important for designing efficient strategies aimed at elucidating gene function and unraveling complex cellular systems. Here, we provide a systematic classification of the underlying causes of and their relative contribution to non-responsive behavior upon gene deletion. The overall contribution of redundancy to non-responsive behavior is estimated at 29%, of which approximately 17% is due to homology-based redundancy and 12% is due to pathway-based redundancy. The major determinant of non-responsiveness is condition dependency (71%). For approximately 14% of protein complexes, just-in-time assembly can be put forward as a potential mechanistic explanation for how proteins can be regulated in a condition dependent manner. Taken together, the results underscore the large contribution of growth condition requirement to non-responsive behavior, which needs to be taken into account for strategies aimed at determining gene function. The classification provided here, can also be further harnessed in systematic analyses of complex cellular systems.

Introduction

Understanding the interplay between individual proteins and their role in various biological processes is critical for understanding cellular systems as well as understanding genotype to phenotype relationships. The role of individual proteins is often investigated by assessing the functional consequences of inactivation, often through genetic perturbation. With the advent

of efficient gene editing techniques such as CRISPR-Cas9, such reverse genetics approaches are now starting to be even more expansively applied [1]. Pioneering work by the yeast gene deletion consortium using homologous recombination has in the past yielded an exhaustive collection of gene deletion mutants for *Saccharomyces cerevisiae* [2]. Subsequent high-throughput studies have enabled systematic investigation of the functional consequences of deleting individual genes and have been conducted to investigate cellular systems in many different ways [3–6]. However, a large number of deletion strains do not show detectable phenotypic effects and the scale and cause of this non-responsive phenomenon has been the subject of considerable interest [7–14]. In *S. cerevisiae* over 66% of deletion mutants have little or no detectable effect on growth in a single rich medium [2,15]. Similar observations have been made for other organisms [16,17]. Growth is not necessarily the most sensitive assay and as an alternative, genome-wide gene expression can be used, also as a much more detailed phenotypic readout. A recent study investigating the consequences of almost 1,500 gene deletions revealed that 53% of these mutants have a genome-wide expression profile that is essentially the same as wildtype (WT) [3]. This confirms that a large proportion of gene deletions have no phenotype when assayed under a single condition. Lack of a phenotype is due to either redundancy mechanisms or due to growth condition dependency of the deleted gene. An important question addressed insufficiently so far, is the degree to which these different mechanisms contribute to non-responsiveness. Understanding the underlying causes and their relative contribution is important for understanding complex cellular systems and for improving reverse genetic strategies aimed at determining gene function.

There are at least two kinds of redundancy mechanisms that contribute to non-responsive behavior upon genetic perturbation. One factor is the presence of closely related paralogs (homology-based redundancy) [7,10,11,14]. Redundancy can also be achieved through the presence of alternative pathways that can facilitate the same biological process (pathway-based redundancy) [11,18]. Homology-based redundancy is attributed to gene duplicates or proteins with similar functional domains. Even though most close paralog pairs have diverged during evolution [19], there are still pairs that have retained the capability to compensate for each other's loss [10,11,14,20]. Previous estimates indicate that as much as 23% to 29% of non-responsiveness can be ascribed to duplicated genes [7,10]. These estimates are however based on growth-rate as the phenotypic read-out and involved analysis of only a single growth condition. It therefore remains to be seen whether these estimates are similar when applying a more sensitive phenotypic read-out and taking into account multiple growth conditions. To what degree pathway-based redundancy contributes to non-responsiveness is still unclear. This is mainly due to the fact that the degree to which negative genetic interactions reflect buffering between two genes in parallel pathways has remained unsolved. In addition to negative genetic interactions between parallel pathways, many negative genetic interactions have been detected between seemingly unrelated pathways [21–23]. It has also been suggested that most negative genetic interaction scores are not due to simple redundancy between pathways, but rather reflect network topology [24,25]. Providing clear estimates of the contribution of pathway-based redundancy to non-responsiveness solely based on negative genetic interactions is therefore unsatisfactory. Additional sources of information likely need to be included in order to provide reliable estimates of pathway-based redundancy.

Besides various redundancy mechanisms, condition dependency is another factor that contributes to non-responsiveness upon genetic perturbation. Some genes are almost certainly only required under specific growth conditions. It has been suggested that a large number of seemingly non-responsive mutants have important fitness roles under specific environmental conditions [8,18,26]. These estimates vary between 37% and 97%, using either focused subsets of genes, different assays or a combination of heterozygous and homozygous deletion strains.

The degree to which these different approaches affect condition dependency estimates is unclear. A uniform analysis, using a single resource of non-responsive deletion strains is therefore required for reliably estimating the contribution of condition dependency to non-responsiveness, also in combination with redundancy estimates.

In spite of much effort, the causes and mechanisms of non-responsiveness remains elusive [7,11,18,26]. Most studies have either focused on a single explanation or have insufficiently investigated the underlying mechanisms and their relative contributions. Here, using a sensitive phenotypic readout (gene expression) as a starting point, we systematically investigate the underlying causes of and their relative contribution to non-responsiveness upon gene deletion. This reveals that condition dependency is the major determinant of non-responsiveness, explaining approximately 71%. Homology-based and pathway-based redundancy contributes 17% and 12% respectively. For approximately 14% of protein complexes, just-in-time assembly provides a potential mechanistic explanation for how protein complexes can be regulated in a condition dependent manner. The classification of the relative contribution of underlying causes provided here, can be harnessed for other systematic analyses of cellular systems.

Results

A collection of non-responsive mutants

To systematically investigate the underlying causes of and their relative contribution to non-responsiveness observed upon gene deletion, we exploited the availability of a collection of 1,484 deletion mutants expression profiles in the yeast *Saccharomyces cerevisiae* [3]. To obtain the list of responsive and non-responsive mutants, the same definitions were applied as in the original study [3]. In short, a mutant is considered non-responsive (NR) if three or less transcripts are changing significantly ($FC > 1.7$ & $p\text{-value} < 0.05$) as a consequence of deleting a particular gene. A mutant is classified as responsive (R) if four or more transcripts are changing significantly upon gene deletion. Applying these thresholds, 784 mutants are classified as NR (53%) and 700 mutants are classified as R (47%) as indicated before [3].

To ensure that non-responsiveness is not restricted to a particular class of proteins, the relative contribution of non-responsiveness to different functional classes was investigated [3]. As expected, non-responsiveness is not limited to a specific functional class, but common to a wide range of functional classes (Fig 1). There are differences however in the relative contribution of non-responsiveness to various functional classes. Classes that have a central role in cellular systems such as chromatin factors, Pol II transcription and translation / ribosome biogenesis have a relatively low number of NR mutants. Other functional classes such as cell cycle regulation and meiosis have a high number of NR mutants. This indicates that there is a relationship between the relative importance of different functional classes and degree of non-responsiveness, but also suggests that different causes such as condition dependency (meiosis) and robustness (cell cycle regulation) may contribute to non-responsiveness.

Both redundancy and condition dependency contribute to non-responsiveness

Exclusive requirement of a gene under a specific growth condition is expected to contribute to the number of non-responsive mutations. As a starting point, differences in mRNA transcript and protein levels between responsive and non-responsive mutants were investigated. As also observed before [3], genes within the group of NR mutants have a slightly larger fraction of genes with low transcript [3] and protein [27] levels (Fig 2A and 2B). This also reveals that almost all genes within the group of NR mutants generally exhibit normal expression levels in

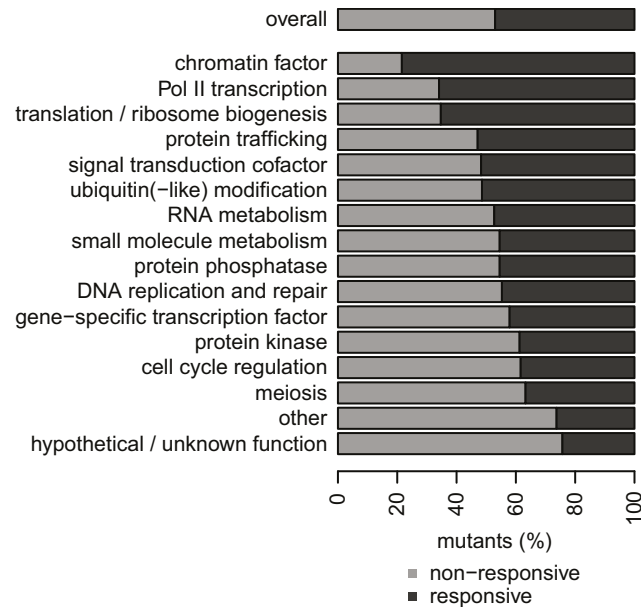


Fig 1. Distribution of non-responsive behavior. Percentage of non-responsive (NR; light grey) and responsive (R; dark grey) mutants within different functional categories, as well as the overall percentage. A mutant is classified as NR if three or less transcripts are changing compared to WT. A mutant is classified as R if four or more transcripts are changing. Functional categories as defined in the original study [3].

doi:10.1371/journal.pone.0173432.g001

wildtype (WT), both for transcript as well as protein levels. Their removal however, has little effect on gene expression and raises the question why so many genes are expressed while not being required under the condition investigated.

Another likely cause of non-responsiveness is redundancy. This is confirmed by a strong enrichment of close paralogs in NR mutants (Fig 2C). Interestingly, NR mutants exhibit markedly less negative genetic interactions [22] (Fig 2D, left panel). Negative genetic interactions do not necessarily reflect redundancy [10,11,24,28,29]. It is also suggested that redundant gene pairs have few negative genetic interactions as they are only expected to show a defect when both genes are absent [10], just as indicated here. NR mutants also have fewer physical interactions [30] and show less growth defects when assayed across many different conditions [8] (Fig 2D, middle and right panel). This is all confirmatory of more specialized, less centrally required roles in cellular biology and that both redundancy and condition dependency are contributing causes of non-responsiveness.

Close paralogs as a proxy for homology-based redundancy

To provide a reliable estimate of the relative contribution of complete redundancy to non-responsiveness, close paralog pairs within the non-responsive mutants are used here. Close paralog pairs are among the most recent evolutionary duplications and probably still have a similar function [32,33]. Close paralog pairs that show non-responsive behavior are therefore most likely the best starting point for estimating the relative contribution of redundancy to non-responsiveness. If two genes form a complete redundancy pair, it is also expected that removing either one of these genes does not result in any measurable defect, regardless of the condition investigated. The dataset from Hillenmeyer et al. [8] measures the relative growth defect of individual deletions in many different conditions. Pairs with a complete redundancy relationship are expected to have low sensitivity to different environmental conditions as

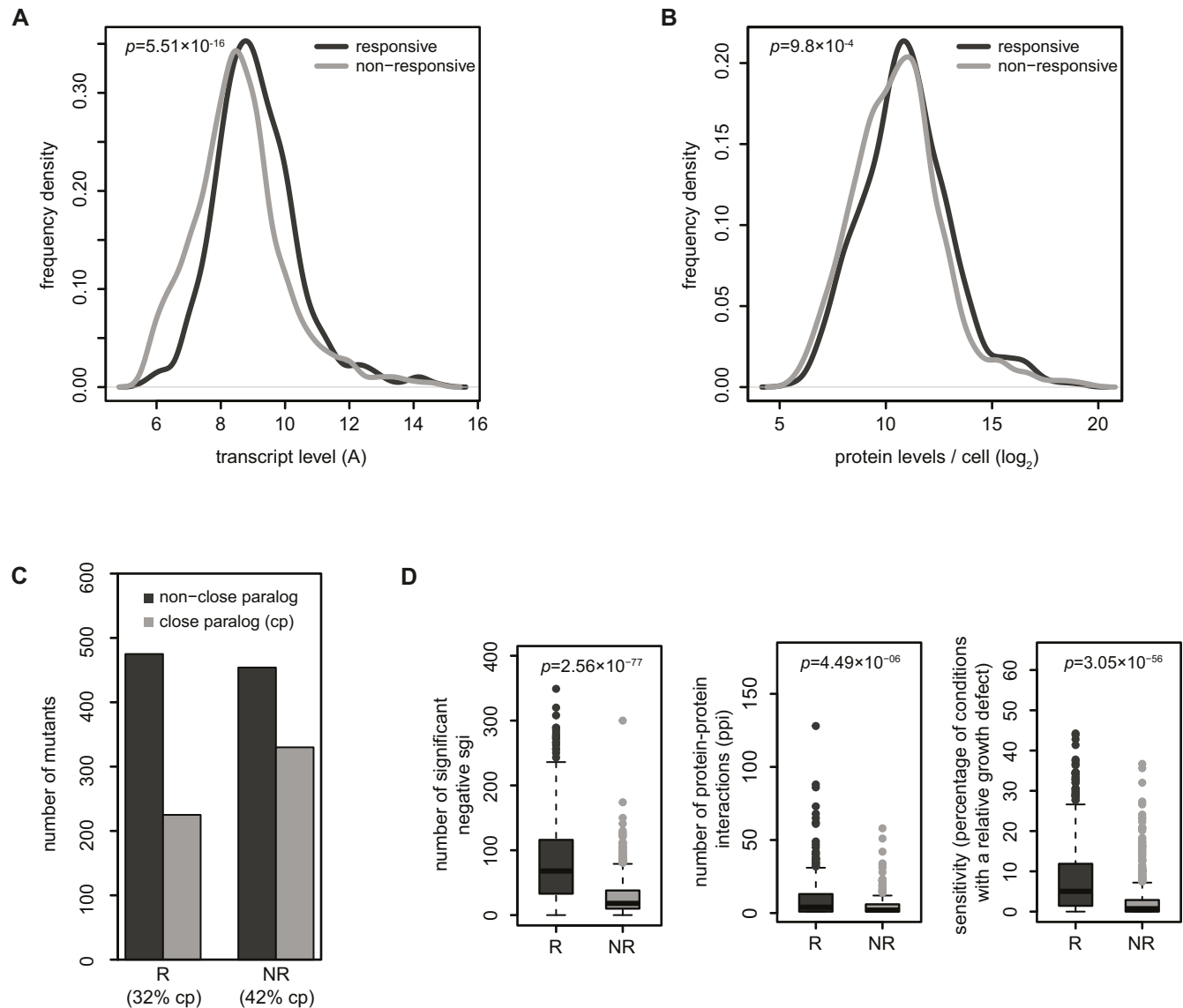


Fig 2. General characteristics of non-responsive and responsive mutants. (A) Frequency density distribution of wildtype (WT) mRNA transcript levels (A) obtained from the average of 200 WT strains, adapted from Kemmeren et al. [3]. *P*-value indicates the difference between transcript levels in the WT strains between NR and R mutants based on a two-sided Mann-Whitney test. (B) Frequency density distribution of the number of proteins per WT cell [27] for NR and R mutants, adapted from Kemmeren et al. [3]. Only detectable proteins are depicted. *P*-value indicates the difference between protein levels in NR and R mutants based on a two-sided Mann-Whitney test. (C) Number of NR and R mutants with a close paralog (CP) that arose from whole-genome duplication (WGD) [31] or small scale duplications (SSD) [20]. (D) Box plots showing the number of significant negative synthetic genetic interaction (SGI) scores [22] ($\epsilon \leq -0.08$, $p \leq 0.05$; left panel), number of protein-protein interactions [30] (PPI, middle panel) and sensitivity to different conditions [8] (percentage of conditions that deletion mutants show a relative growth defect; growth defect > 0, $p \leq 0.05$; right panel) for NR and R mutants. *P*-values are based on a two-sided Mann-Whitney test.

doi:10.1371/journal.pone.0173432.g002

measured in this dataset (Materials and Methods). Indeed, close paralog pairs between non-responsive mutants exhibit low sensitivity to different environmental conditions (Fig 3A, left panel). On the other hand, close paralog pairs between non-responsive and responsive mutants or between responsive mutants show a significant increase in sensitivity to different conditions (Fig 3A, middle & right panel; Fig 3B). This effect is stronger when only considering negative genetic interactions within the pairs depicted in Fig 3A (Fig 3C and 3D),

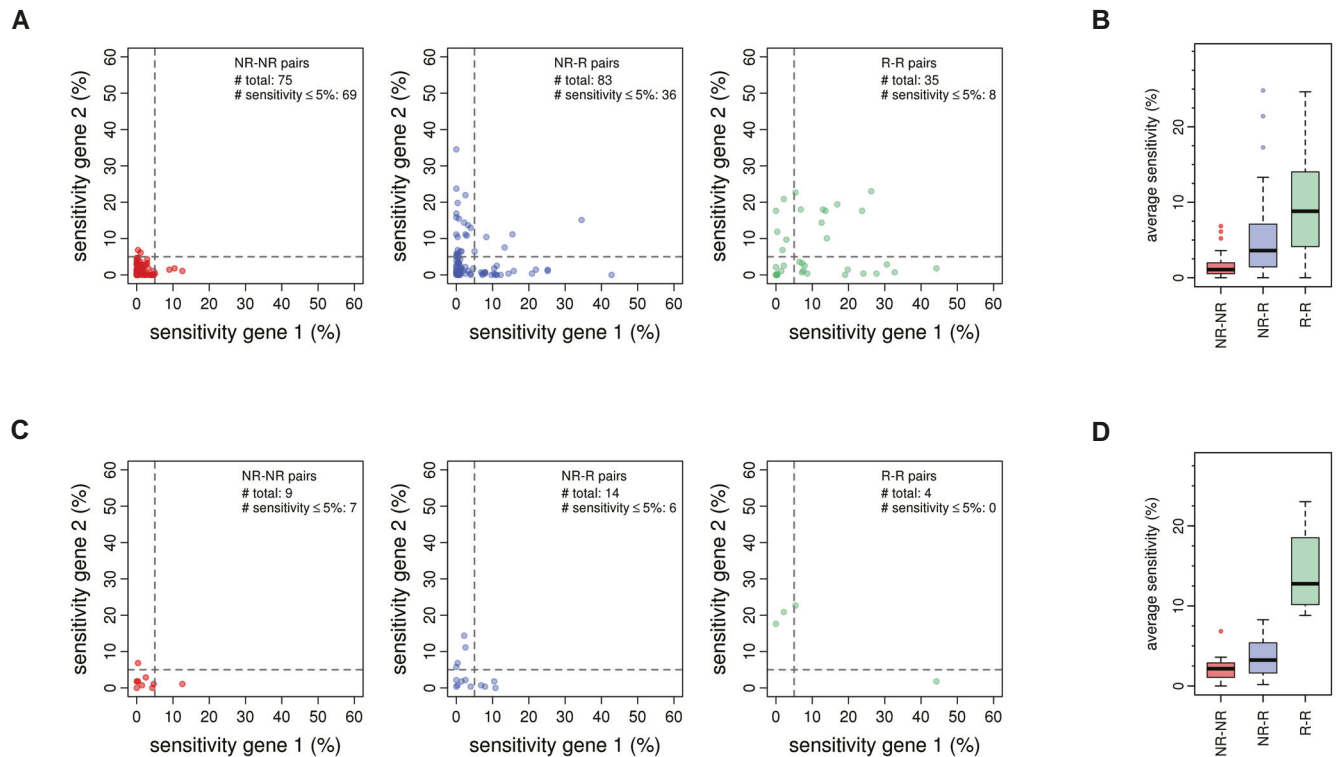


Fig 3. Homolog-based redundancy estimates. (A) Scatter plots depicting the sensitivity of gene 1 and gene 2 for close paralog pairs between non-responsive (NR-NR, left panel), between non-responsive and responsive (NR-R, middle panel) or between responsive (R-R, right panel) mutants. (B) Boxplots of the average sensitivity of the same groups as in A. The difference between all three groups is statistically significant (p -value (NR-NR vs. NR-R) = 3.14×10^{-9} , p -value (NR-NR vs. R-R) = 1.02×10^{-8} , p -value (NR-R vs. R-R) = 1.66×10^{-3}) based on a one-sided Mann-Whitney test. (C) Same as in A, except that only close paralog pairs with a significant negative SGI score are shown. (D) Boxplots of the average sensitivity of the same groups as in C (p -value (NR-NR vs. NR-R) = 0.128, p -value (NR-NR vs. R-R) = 1.4×10^{-3} , p -value (NR-R vs. R-R) = 3.27×10^{-3}), statistical test as in B.

doi:10.1371/journal.pone.0173432.g003

confirming that negative genetic interactions provide useful additional information when investigating redundancy.

Interestingly, a large difference in sensitivity between the various close paralog groups is observed (Fig 3A and 3B), suggesting that there might be a disparity in sequence divergence among the close paralog pairs. The number of non-synonymous substitutions per non-synonymous site (K_a) is often considered an indicator of sequence divergence [14]. Although there seems to be a slight difference in the K_a values for close paralog pairs between non-responsive (NR-NR), between non-responsive and responsive (NR-R) or between responsive mutants (R-R), this is not significant (S1A Fig). When considering percentage sequence identity, a significant difference is only observed when using the domains shared between close paralog pairs (Materials and Methods; S1B and S1C Fig). This effect is lost when using all pairs but close paralog pairs (S1D Fig), suggesting that compensation capability is conserved between pairs with the same ancestral gene.

To ensure that homology-based redundancy estimates are reliable, several sensitivity cutoffs and their corresponding false-discovery rate (FDR) were investigated (Materials and Methods). Selecting close paralog pairs between non-responsive mutants with 5% or less sensitivity leads to the lowest FDR ($\approx 10\%$; S2 Fig). Based on this, the contribution of homology-based redundancy to non-responsiveness is estimated at approximately 17% (Fig 3A, left panel; S1

Table). This encompasses 133 unique genes out of 784 NR mutants that are likely non-responsive due to a complete redundancy relationship with another gene.

Negative genetic interactions indicate pathway-based redundancy

In addition to homology-based redundancy (Fig 3A and 3B), where two genes can directly compensate for each other's loss, two genes can also operate in two redundant pathways. Negative genetic interactions have in the past been associated with redundant pathways [34]. Pairs of non-responsive mutants with negative genetic interactions show less sensitivity to different conditions compared to pairs between non-responsive and responsive or between responsive mutants (Fig 4A and 4B), as also observed before for close paralog pairs (Fig 3). However, taking into account the 5% sensitivity cutoff for estimating the contribution of pathway-based redundancy, the FDR rate is much higher (approximately 70%) compared to homology-based redundancy estimates. This clearly indicates that only using negative genetic interactions as a proxy for pathway redundancy is insufficient for estimating the contribution of pathway-based redundancy to non-responsiveness.

If two genes operate in a completely redundant manner, they should only exhibit a negative genetic interaction score with each other and not with any other gene [10]. Since the removal of a single gene can be completely compensated by the other gene, there should be no measurable phenotypic effect, except when both genes are deleted. The number of negative genetic interactions is indeed significantly lower for pairs between non-responsive mutants compared to responsive mutants (Fig 4C). Using a cutoff on the number of negative genetic interactions (40), a similar degree of reliability ($FDR \leq 10\%$) as obtained for homology-based redundancy estimates can be obtained (Fig 4D and 4E; Materials and Methods). Pathway-based redundancy estimates were therefore based on pairs between non-responsive mutants that fulfilled four criteria. First, pairs are not close paralogs of each other. Second, pairs have a significant negative genetic interaction. Third, pairs have 40 or less negative genetic interactions with other genes. And last, pairs show, on average, 5% or less sensitivity in other conditions. In total, 68 pairs, 97 unique genes, can be associated with pathway-based redundancy (Fig 4D, left panel; S2 Table), indicating that approximately 12% ($FDR \approx 8\%$) of non-responsiveness can be attributed to pathway-based redundancy.

Relative contribution of redundancy and condition dependency to non-responsiveness

By combining the estimates of homology-based redundancy (133 genes, 17%) and pathway-based redundancy (97 genes, 12%), the overall contribution of redundancy to non-responsiveness is approximately 29% (Fig 5A). This number is a rough estimate and the exact percentage also depends on the cutoffs used. Here, stringent cutoffs are used to avoid a high FDR, but therefore potentially interesting pairs might be missed that do not fulfill the stringent criteria. Nevertheless, the analysis indicates that the majority of non-responsiveness, as much as 71%, is attributable to condition dependency (554 genes; Fig 5A; S3 Table). Interestingly, different functional classes do show diverse behavior in the relative contribution of condition dependency and redundancy (Fig 5B). Non-responsiveness for classes such as meiosis, DNA replication and repair have a relatively high contribution of condition dependency. A few classes such as gene-specific transcription factors, protein kinases and protein phosphatases however, show a much higher contribution of homology-based redundancy, indicating that the degree of redundancy differs between different functional classes (Fig 5B).

Mechanistic explanations for homology-based as well as pathway-based redundancy have been proposed before [28,29]. For homology-based redundancy, this is based on two proteins

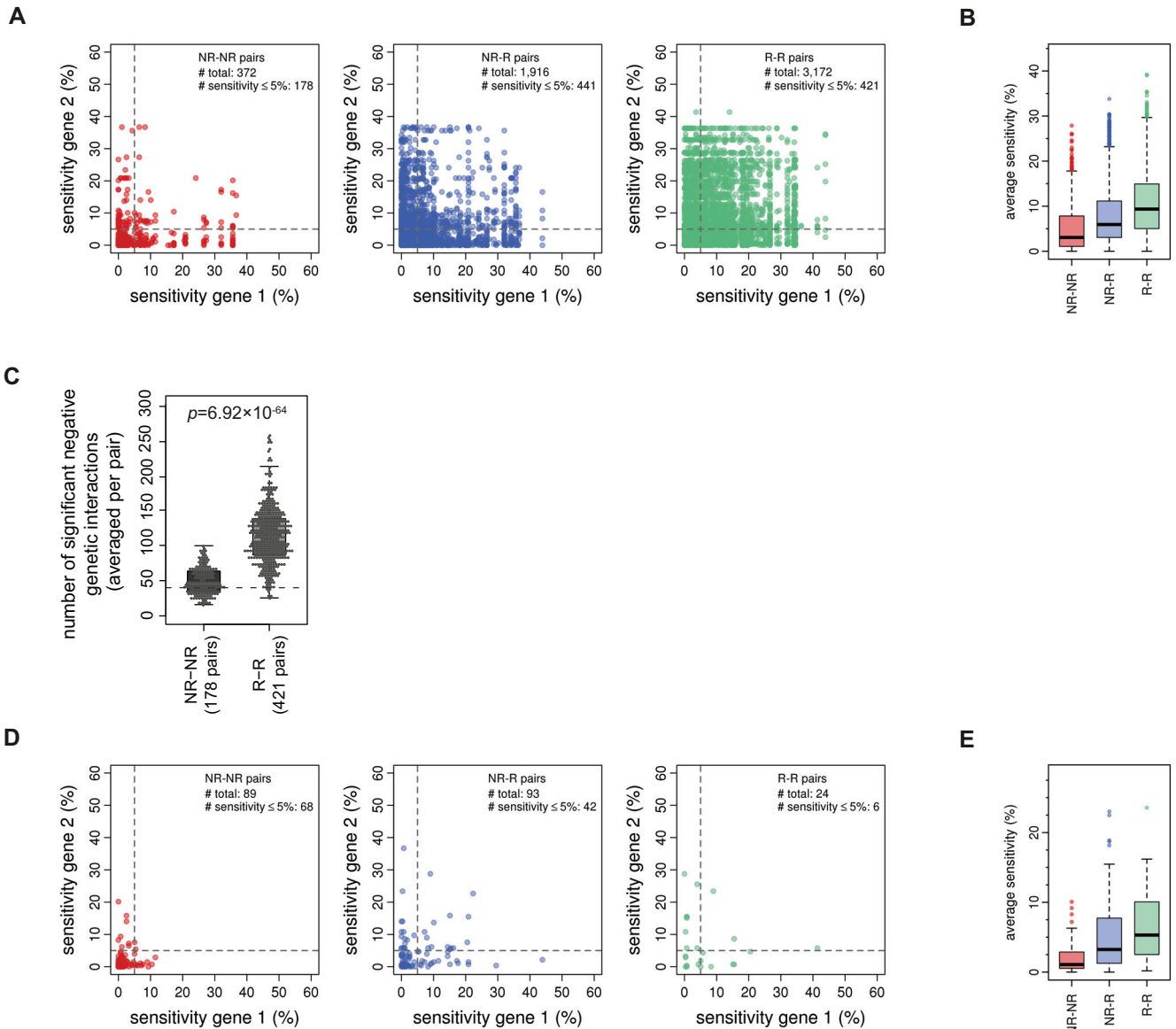


Fig 4. Pathway-based redundancy estimates. (A) Scatter plots depicting the sensitivity of gene 1 and gene 2 for pairs with a significant negative genetic interaction score between non-responsive (NR-NR), between non-responsive and responsive (NR-R) or between responsive (R-R) mutants. Close paralog pairs are excluded. (B) Boxplots of the average sensitivity of the same groups as in B, all three are significantly different (p -value $< 2.2 \times 10^{-16}$). P -values are based on a one-sided Mann-Whitney test. (C) Boxplot showing the number of significant negative genetic interactions for NR-NR and R-R pairs with 5% or less sensitivity. Close paralog pairs are excluded. The dashed line indicates the threshold (40) taken to obtain an FDR of 8%. (D) Scatter plots shown for the same pairs in A that in addition have, on average, less than 40 negative genetic interactions per pair. (E) Boxplots of the average sensitivity of the same groups as in D (p -value (NR-NR vs. NR-R) = 1.31×10^{-7} , p -value (NR-NR vs. R-R) = 3.64×10^{-6} , p -value (NR-R vs. R-R) = 0.064), statistical test as in B.

doi:10.1371/journal.pone.0173432.g004

that can completely replace each other's function (Fig 5A). For instance, the redundant kinases Ark1 and Prk1 that both phosphorylate the same consensus target amino acid sequence and are involved in regulating endocytosis and actin skeleton [24,35]. Another example includes Upc2 and Ecm22, two redundant gene-specific transcription factors that bind to the same set of target genes and regulate ergosterol biosynthesis [36,37]. A final example involves Nhp6a and Nhp6b, two highly homologous HMGB proteins involved in modulation of chromatin

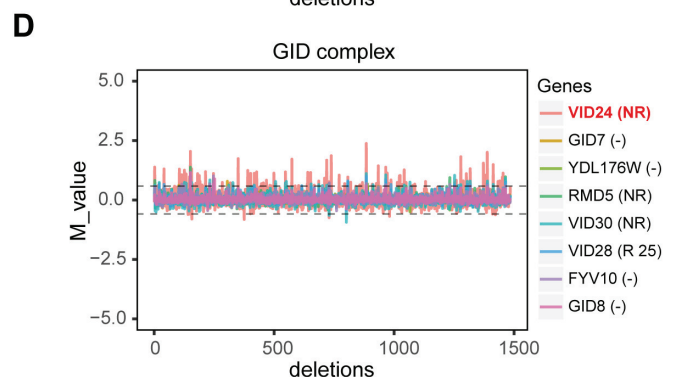
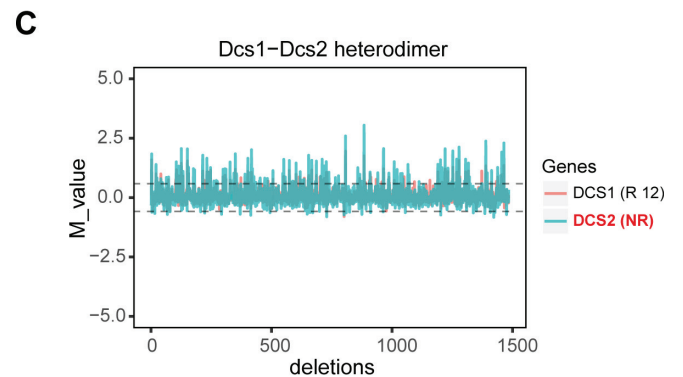
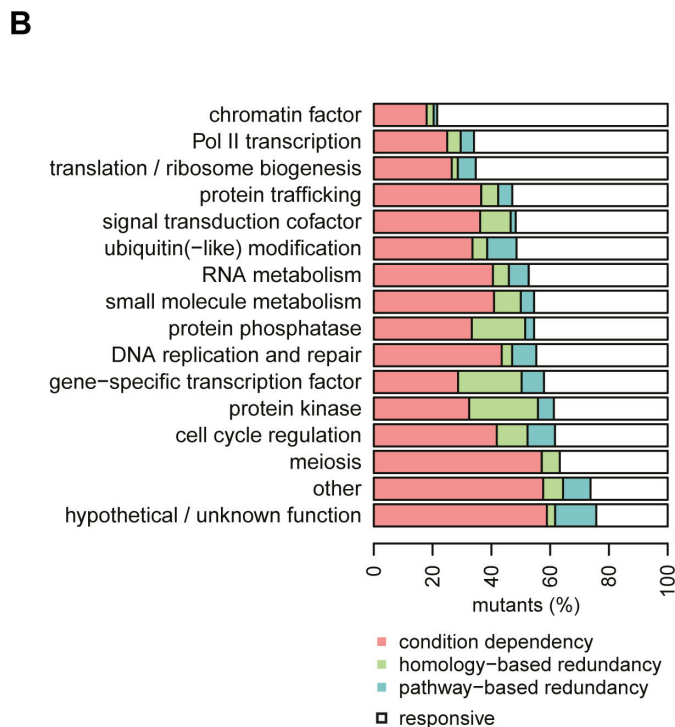
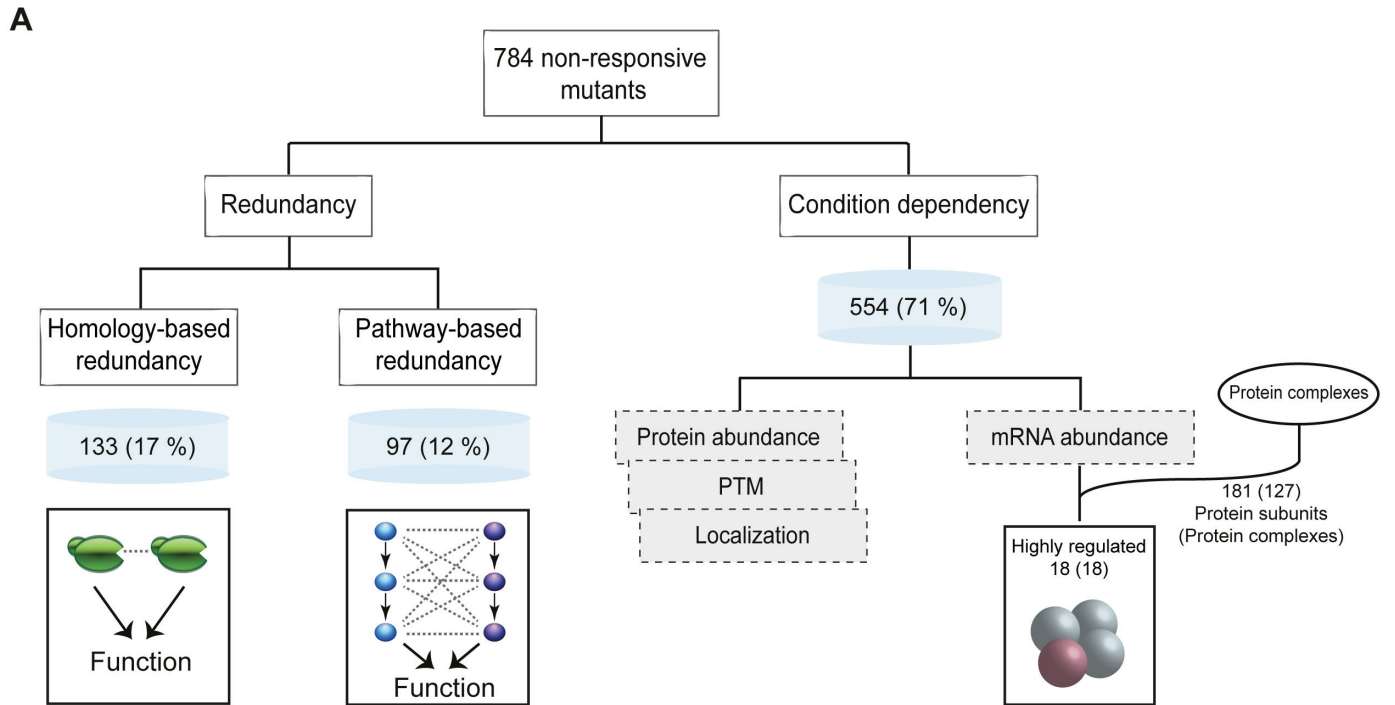


Fig 5. Relative contribution of redundancy and condition dependency and potential mechanisms. (A) Flowchart showing the relative contribution of redundancy and condition dependency to non-responsiveness. Potential mechanisms are indicated for homology-based redundancy, pathway-based redundancy and “just-in-time assembly”. (B) Relative contribution of redundancy and condition dependency to non-responsiveness within each functional category. (C) Line plot showing the mRNA expression changes across 1,484 deletion mutants [3] for the Dcs1-Dcs2 heterodimer. Individual lines indicate the Dcs1 and Dcs2 subunit. The highly regulated subunit is indicated in red. (D) Line plot as in C for the GID complex.

doi:10.1371/journal.pone.0173432.g005

structure [38]. These examples are all indicative of more direct relationships for homology-based redundancy pairs with proteins having the ability of taking over the exact same function.

For pathway-based redundancy, the concept is that two pathways or proteins can compensate each other's function in a more indirect fashion (Fig 5A). Although the exact relationship between many of these pairs is still unclear, they are often found to be involved in the same process and loss of both leads to significant unexpected defects. One example is the relationship between *Ecm1* and *Pom34*. *Ecm1* is a pre-ribosomal factor involved in pre-60S particle export and *Pom34* is a subunit of the nuclear pore complex (NPC). In the absence of *Pom34*, depletion of *Ecm1* leads to defects in pre-60S particle export [39]. Another example includes *Kar5* and *Prm3*, which are both involved in nuclear membrane fusion during karyogamy and loss of both genes leads to nuclear fusion defects [40]. Note that, as highlighted with these cases, the underlying mechanism for pairs associated with pathway-based redundancy is often unclear.

Potential mechanisms for condition dependent genes

The mechanistic explanations provided for redundancy are useful in understanding the interplay between different proteins and how this relates to cellular processes. For genes/proteins required only under certain conditions, regulation can occur at many different levels, such as regulation of mRNA and protein levels, post-translational modifications (PTMs) that change aspects of protein-protein interactions and cellular localization. Many genes that are non-responsive still have relatively high mRNA and protein expression levels (Fig 2A and 2B), suggesting that these are constitutively expressed despite not being required under the given condition. For cell-cycle related protein complexes, De Lichtenberg et al. [41] have proposed a mechanism termed "just-in-time assembly" to indicate that most members of these protein complexes are constitutively expressed and only a few key subunits need to be regulated in order to control the activity of the entire protein complex. Many proteins are expressed under the condition investigated here, but do not result in a detectable phenotypical defect when removed. This therefore raises the question of whether "just-in-time assembly" may also apply to condition dependency.

To investigate this, all proteins that are part of a protein complex were obtained using a pre-defined definition of protein complexes [42]. Out of 554 genes attributed to condition dependency 181 are part of 127 unique protein complexes (Fig 5A). To determine the degree of differential regulation within each protein complex, highly regulated subunits were detected based on the degree of variation in expression levels across different deletion mutants (Material and Methods). Out of the 127 protein complexes, 18 complexes contain a highly regulated subunit ($\approx 14\%$; Fig 5A; S3 Fig). To ensure that these results are not limited to the subset of proteins analyzed here, the same analysis was also performed using all protein coding genes in yeast. This revealed that 69 protein complexes out of 501 protein complexes [42] contain a highly regulated subunit ($\approx 14\%$), confirming the results obtained when using the non-responsive mutants attributed to condition dependency.

One example of a protein complex with a highly regulated subunit includes the *Dcs1-Dcs2* heterodimer that removes capped mRNA fragments left over from mRNA decay [43]. Previous reports indicated that *Dcs2* is most sensitive to transcriptional induction and suppresses *Dcs1* activity in response to nutrient stress [44,45]. Here, *Dcs2* is identified as a highly regulated subunit (Fig 5C), confirming its role as a modulator of the activity of the *Dcs1-Dcs2* heterodimer. Another example includes the *GID* complex, consisting of seven subunits [46–48]. The *GID* complex targets the key enzyme fructose-1,6-biphosphatase (FBPase) for proteosomal degradation during the switch from gluconeogenesis to glycolysis [48]. *Vid24*, the regulatory

subunit of the GID complex, initiates FBPase polyubiquitination and subsequent degradation [48]. Vid24 protein levels are undetectable under gluconeogenic conditions, but rapidly accumulate when glucose is added and decrease together with FBPase [48]. Our analysis also reveals Vid24 as a highly regulated subunit (Fig 5D), consistent with previous reports about the role of Vid24 in regulating the activity of the GID complex [46,48]. Besides these known examples, there are also putative protein complexes that have a highly regulated subunit (S3 Fig) that can be used as a starting point for further unraveling their function and potential regulatory mechanism. Taken together, approximately 14% of protein complexes are regulated similar to the “just-in-time assembly” model, but in a condition-dependent manner.

Discussion

Redundancy estimates

The high proportion of gene deletions that fail to exhibit a detectable defect under a specific condition has muddled efforts aimed at systematic phenotyping. Elucidating gene function requires the ability to elicit a detectable phenotypical readout, such as a growth defect or gene expression changes. Over 66% of deletion mutants in *Saccharomyces cerevisiae* have no detectable growth phenotype [15] and over 53% of non-essential gene deletions hardly show any gene expression changes [3]. Understanding the underlying causes of this non-responsive behavior will aid in elucidating gene function and understanding the complexity of cellular systems. Here, we estimate that approximately 29% of non-responsiveness can be attributed to either homology-based (17%) or pathway-based (12%) redundancy. Previous estimates have indicated that between 23% and 29% of negative genetic interactions can be ascribed to duplicated genes [7,10]. Our estimate of 17% is most likely lower because we require low sensitivity under multiple conditions (Fig 3A) whereas previous estimates are based on growth under a single condition. For complete redundancy pairs, deletion of a single gene should exhibit no detectable phenotype under any given condition. It is therefore expected that the redundancy pairs that we miss, but are included in previous estimates will fail to exhibit a redundancy relationship when investigated under different conditions.

Homology-based redundancy versus pathway-based redundancy

Estimates for the contribution of redundancy to non-responsiveness include both homology-based as well as pathway-based redundancy. The known examples found for pathway-based redundancy indicate a higher degree of uncertainty as they are found to be involved in the same process and loss of both leads to significant defects. The exact mechanism however, is often unknown. Since there is no direct sequence-based evidence for a redundancy relationship between these pairs, it remains to be seen to what degree pathway-based redundancy relationships can fully compensate the loss of one gene, or whether this is only partial. One can also question whether these should be considered as redundancy relationships, or whether these most likely reflect pathway connectivity or combinatorial effects as already suggested before [24,28]. The most likely candidate pairs providing redundancy are therefore those based on homology (estimated at 17%), whereas most other negative genetic interactions reflect pathway connectivity.

Partial redundancy

Partial redundancy is also expected to contribute to non-responsive behavior. Simplistically, two different types can be differentiated. For the first type, only one of the two genes involved elicits a response under the condition investigated. The other type consists of a pair whereby

deletion of either gene does not result in a detectable phenotype under the condition investigated. However, at least one of the genes would elicit a response when removed under a different (environmental) condition. We deliberately did not include these pairs within our redundancy estimates, as the purpose here is to investigate to what degree complete redundancy contributes to non-responsive behavior. Pairs that can fully compensate for each other's loss are expected to do so given any circumstance. It is therefore much more likely that partial redundancy is either an intrinsic property of pathway connectivity, just as for many pathway-based redundancy pairs, or is a different way to achieve condition dependency. The relationship might then indicate a manner to regulate condition dependency or ensure that genes with a similar function, but required under different environmental conditions can behave in a switch-like fashion so that the most efficient protein is available under the right circumstances.

Condition dependency and just-in-time assembly

While condition dependency can be regulated in many different ways, we focused here on a mechanism shown previously for cell cycle related protein complexes [41]. The just-in-time assembly mechanism may provide a mechanistic explanation how protein complexes can be regulated in a condition-dependent manner. Regulation of approximately 14% of all protein complexes can be explained in such a way. This still leaves many protein complexes unexplained and indicates that the just-in-time assembly is not such a widespread mechanism. Interestingly, many of the protein complexes with a highly regulated subunit are involved in metabolism. Almost 50% of these protein complexes are associated with metabolism and this percentage is even higher for heterodimers. Metabolic flux analyses have shown that the activity of metabolic pathways can be regulated in different ways and through several regulatory factors [49]. Although not necessarily the most predominant mode of action, transcriptional regulation of key regulatory factors has been shown as one way to achieve condition- and tissue-specific control of metabolic pathway activity [50]. The enrichment of heterodimers also fits the idea of enzymatic protein complexes consisting of a regulatory and catalytic subunit. The mechanism found here, might thus be more specific for metabolic pathways and protein complexes involved in metabolism.

The term just-in-time assembly proposed by De Lichtenberg et al. [41] to describe the dynamic formation of protein complexes during the cell cycle might not always be appropriate for the highly regulated subunits found here. The term implies that the entire complex is assembled just before it is needed. This doesn't necessarily need to be the case. For instance, most of the protein complex could already be pre-assembled and only the regulatory subunit joins the partly pre-assembled protein complex when needed. Since it is impossible to assess the exact mechanism by which a protein complex is activated using only gene expression data, we propose to use "just-in-time activation" as a more general term. This also allows other types of just-in-time activation to be included, such as protein phosphorylation or other post-translational modifications. Additional data exploring condition dependent protein expression or post-translational modifications are needed to gain more detailed mechanistic insights.

Materials and methods

Gene expression data and functional categories

The list of deletion mutants and corresponding gene expression profiles were obtained from Kemmeren et al [3]. Each gene expression profile consists of *p*-values and average mRNA transcript changes (*M* values; $\log_2(\text{fold-change})$) relative to wildtype (WT). Genes are considered significantly changed when fold-change (FC) > 1.7 and *p*-value < 0.05. Genes that were frequently changing regardless of the deletion mutant (WT variable genes) were excluded as

described in the original study [3]. The number of differentially expressed transcripts was used to classify mutants either as responsive (R) or non-responsive (NR) as done previously [3]. Mutants with less than 4 transcript changes relative to WT are considered NR and mutants with 4 or more changes are considered R. Mutants were grouped in different functional categories as before [3]. Both the overall ratio of NR and R mutants as well as for each individual functional category (Fig 1) was calculated.

mRNA transcript levels in WT

A wild-type pool consisting of 200 WTs [3] was used to obtain average mRNA transcript levels (A values; $\log_2(R^*G)$) for all genes corresponding to the different deletion mutants. mRNA transcript levels in WT were compared between NR and R mutants. *P*-values are calculated using a Mann-Whitney two-sided test to evaluate differences in mRNA transcript levels between NR and R mutants (Fig 2A).

Protein expression levels, protein-protein interactions and protein complexes

The number of protein molecules per WT cell was obtained from Ghaemmaghami et al. [27]. Proteins with undetectable signals, low signals or experimental problems were excluded. Protein expression levels were compared between NR and R mutants. *P*-values are calculated using a Mann-Whitney two-sided test to evaluate differences in protein levels between NR and R mutants (Fig 2B). Protein-protein interactions were obtained from Collins et al. [30]. The number of protein-protein interactions between NR and R mutants was compared. *P*-values are calculated using a Mann-Whitney two-sided test (Fig 2D, middle panel). Protein complexes (501) were obtained from the curated “consensus + GO” set from Benschop et al. [42].

Close paralog genes

For close paralog genes, both small scale duplications (SSD) as well as whole-genome duplications (WGD) are used. A list of SSD pairs was obtained from Guan et al. [20]. WGD (547 pairs) were obtained from Byrne and Wolf [31]. Global protein sequence alignment for each close paralog pair was performed using the Needleman-Wunsch algorithm [51] as implemented in Needle-EMBOSS version 6.6 [52] with default options. Only SSD pairs with 20% or higher identity were kept (751 pairs). Combining both WGD and SSD resulted in a total of 1,298 close paralog pairs. Close paralog pairs for which both genes were profiled [3] were used for further downstream analyses (205 pairs).

Genetic interaction data

Synthetic genetic interaction (SGI) scores were obtained from Costanzo et al. [22]. SGI scores were acquired for all pairwise combinations of NR and R mutants for which a score was calculated (Figs 3 and 4). For determining significant negative genetic interactions, the same cutoff was applied as used originally [22] ($\epsilon \leq -0.08$, $p \leq 0.05$).

Condition sensitivity

Growth rate of yeast homozygous gene deletions compared against WT under multiple conditions (418 experiments) was obtained from Hillenmeyer et al. [8]. A relative growth rate above zero is an indication of a growth defect of a deletion mutant relative to WT in a particular condition. A two-sided t-test was applied to assign *p*-values. Growth rates from duplicate conditions were averaged, resulting in 278 unique conditions. Only significant growth defects were

selected for downstream analysis ($p \leq 0.05$). The sensitivity of each gene is expressed as the percentage of significant relative growth defects within the 278 unique conditions.

Sequence divergence

The number of non-synonymous substitutions per non-synonymous site (Ka) was calculated based on a method introduced by Yang and Nielsen [53]. This method is implemented in the GenomeHistory 2.0 tool, which was used to calculate Ka values [54]. Non-default parameters chosen include: minimum translated ORF: 100; minimum number of aligned residues to accept pair: 100; accepting all BLAST hits with $e < = 1e-08$ and minimum percentage identity for analysis: 40% (S1A Fig). Protein sequence identity for close paralog pairs was calculated either across the whole protein sequence or only for protein domains using the Needleman-Wunsch algorithm [51] as implemented in Needle-EMBOSS version 6.6 [52] with default options (S1B Fig). InterProScan version 5.10–50 [55] was used to obtain a list of shared domains between genes. Only protein domains identified by Pfam [56] were included in downstream analyses. Sequence identity was calculated between two shared Pfam domains using the same tool as used for complete protein sequences. Gene pairs with at least one shared Pfam domain were included in the downstream analysis. Sequence identity for Pfam domains with the same identifier were averaged before averaging all distinct shared domains across two proteins.

Averages of sequence identity between shared Pfam domains for gene pairs were compared either for close paralog or non-paralog pairs. *P*-values are calculated using a Mann-Whitney two-sided test (S1C and S1D Fig).

Contribution of redundancy and condition dependency to non-responsiveness

Homology-based redundancy pairs had to fulfill the following criteria: the pair is a close paralog pair (either WGD or SSD), both genes are NR and both genes have a condition sensitivity of 5% or less. For calculating the false-discovery rate (FDR), NR-NR close paralog pairs below the sensitivity cutoff were used as true positives and R-R close paralog pairs below the sensitivity cutoff were used as false positives. The lowest false-discovery rate (FDR) was achieved when the applied cutoff was 5% sensitivity (S2 Fig). Pathway-based redundancy pairs had to fulfill the following criteria: the pair has a significant negative genetic interaction score, both genes are NR and both genes have a condition sensitivity of 5% or less. To obtain a similar degree of reliability for pathway-based redundancy estimates, the maximum allowed FDR was set to 10% (as obtained from the homology-based redundancy estimates). In addition to the abovementioned criteria, pathway-based redundancy pairs also have to have less than 40 significant negative genetic interaction scores with other genes in order to obtain a FDR lower than 10%. A pair that isn't classified as homology-based redundancy or pathway-based redundancy is considered condition dependent.

Highly regulated genes

Highly regulated genes were determined within the collection of deletion mutants [3]. For each transcript, the observed standard deviation across different deletion mutants or conditions was compared against 1,000,000 permuted transcript profiles. The permuted transcript profiles were generated by randomly selecting expression values within each column (deletion mutant) a million times. *P*-values were calculated as the fraction of permuted profiles with standard deviations equal or larger than the observed standard deviation. Transcripts with *p*-

values < 0.05 after multiple testing correction (Benjamini & Hochberg) are considered highly regulated. Wildtype variable genes [3] were excluded from the list of highly regulated genes.

Supporting information

S1 Fig. Sequence divergence in close paralog pairs. (A) Ka values as a measurement of sequence divergence are depicted for NR-NR (75), NR-R (83) and R-R (35) pairs. (B) Boxplots showing percentage sequence identity across the entire protein length for the same pairs as in A. (C) Boxplots showing percentage sequence identity within shared domains for the same pairs as in A. The difference between NR-NR and R-R pairs is statistically significant (p -value = 0.008). (D) Boxplots showing percentage sequence identity within shared Pfam domains (close paralogs excluded). No significant difference is observed between the three groups. The number of investigated pairs is 96,141, 196,672 and 100,128 for NR-NR, NR-R, and R-R mutants, respectively. Of these pairs, 455, 736 and 320 share at least one Pfam domain and are included in this figure.

(PDF)

S2 Fig. Selected sensitivity cutoff for homology based redundancy. Line plot showing the calculated FDR for different sensitivity cutoffs (Materials and Methods). The dashed line indicates the 5% sensitivity cutoff used to identify the contribution of homology-based redundancy to non-responsiveness (FDR \approx 10%). The same cutoff was also used to infer the contribution of pathway-based redundancy.

(PDF)

S3 Fig. Highly regulated protein complex subunits. Line plots showing the mRNA expression changes across 1,484 deletion mutants for protein complexes with a highly regulated subunit (18 in total). Each line shows the expression changes observed for the individual subunits. Subunits highlighted in red have been identified as highly regulated.

(PDF)

S1 Table. List of homology-based redundant pairs.

(TXT)

S2 Table. List of pathway-based redundant pairs.

(TXT)

S3 Table. List of genes attributed to condition dependency.

(TXT)

Acknowledgments

We would like to thank Philip Lijnzaad, Thanasis Margaritis, Wim de Jonge and Berend Snel for helpful discussions and advice throughout the project.

Author Contributions

Conceptualization: PK FCPH SA.

Formal analysis: PK FCPH SA.

Funding acquisition: PK.

Methodology: PK FCPH SA.

Software: SA.

Supervision: PK FCPH.

Visualization: SA.

Writing – original draft: PK FCPH SA.

Writing – review & editing: PK FCPH SA.

References

1. Wright AV, Nuñez JK, Doudna JA. Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. *Cell*. 2016; 164: 29–44. doi: [10.1016/j.cell.2015.12.035](https://doi.org/10.1016/j.cell.2015.12.035) PMID: [26771484](https://pubmed.ncbi.nlm.nih.gov/26771484/)
2. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, et al. Functional Characterization of the *S. cerevisiae* Genome by Gene Deletion and Parallel Analysis. *Science*. 1999; 285: 901–906. PMID: [10436161](https://pubmed.ncbi.nlm.nih.gov/10436161/)
3. Kemmeren P, Sameith K, van de Pasch LAL, Benschop JJ, Lenstra TL, Margaritis T, et al. Large-Scale Genetic Perturbations Reveal Regulatory Networks and an Abundance of Gene-Specific Repressors. *Cell*. 2014; 157: 740–752. doi: [10.1016/j.cell.2014.02.054](https://doi.org/10.1016/j.cell.2014.02.054) PMID: [24766815](https://pubmed.ncbi.nlm.nih.gov/24766815/)
4. Liberali P, Snijder B, Pelkmans L. Single-cell and multivariate approaches in genetic perturbation screens. *Nat Rev Genet*. 2015; 16: 18–32. doi: [10.1038/nrg3768](https://doi.org/10.1038/nrg3768) PMID: [25446316](https://pubmed.ncbi.nlm.nih.gov/25446316/)
5. Giaever G, Nislow C. The Yeast Deletion Collection: A Decade of Functional Genomics. *Genetics*. 2014; 197: 451–465. doi: [10.1534/genetics.114.161620](https://doi.org/10.1534/genetics.114.161620) PMID: [24939991](https://pubmed.ncbi.nlm.nih.gov/24939991/)
6. Usaj MM, Styles EB, Verster AJ, Friesen H, Boone C, Andrews BJ. High-Content Screening for Quantitative Cell Biology. *Trends Cell Biol*. 2016; 26: 598–611. doi: [10.1016/j.tcb.2016.03.008](https://doi.org/10.1016/j.tcb.2016.03.008) PMID: [27118708](https://pubmed.ncbi.nlm.nih.gov/27118708/)
7. Gu Z, Steinmetz LM, Gu X, Scharfe C, Davis RW, Li W-H. Role of duplicate genes in genetic robustness against null mutations. *Nature*. 2003; 421: 63–66. doi: [10.1038/nature01198](https://doi.org/10.1038/nature01198) PMID: [12511954](https://pubmed.ncbi.nlm.nih.gov/12511954/)
8. Hillenmeyer ME, Fung E, Wildenhain J, Pierce SE, Hoon S, Lee W, et al. The Chemical Genomic Portrait of Yeast: Uncovering a Phenotype for All Genes. *Science*. 2008; 320: 362–365. doi: [10.1126/science.1150021](https://doi.org/10.1126/science.1150021) PMID: [18420932](https://pubmed.ncbi.nlm.nih.gov/18420932/)
9. Hsiao T-L, Vitkup D. Role of Duplicate Genes in Robustness against Deleterious Human Mutations. *PLoS Genet*. 2008; 4: e1000014. doi: [10.1371/journal.pgen.1000014](https://doi.org/10.1371/journal.pgen.1000014) PMID: [18369440](https://pubmed.ncbi.nlm.nih.gov/18369440/)
10. Ihmels J, Collins SR, Schuldiner M, Krogan NJ, Weissman JS. Backup without redundancy: genetic interactions reveal the cost of duplicate gene loss. *Mol Syst Biol*. 2007; 3: 86. doi: [10.1038/msb4100127](https://doi.org/10.1038/msb4100127) PMID: [17389874](https://pubmed.ncbi.nlm.nih.gov/17389874/)
11. Wagner A. Robustness against mutations in genetic networks of yeast. *Nat Genet*. 2000; 24: 355–361. doi: [10.1038/74174](https://doi.org/10.1038/74174) PMID: [10742097](https://pubmed.ncbi.nlm.nih.gov/10742097/)
12. Li J, Yuan Z, Zhang Z. The Cellular Robustness by Genetic Redundancy in Budding Yeast. *PLoS Genet*. 2010; 6: e1001187. doi: [10.1371/journal.pgen.1001187](https://doi.org/10.1371/journal.pgen.1001187) PMID: [21079672](https://pubmed.ncbi.nlm.nih.gov/21079672/)
13. Zhang J. Genetic Redundancies and Their Evolutionary Maintenance. In: Soyer OS, editor. *Evolutionary Systems Biology*. Springer New York; 2012. pp. 279–300.
14. Plata G, Vitkup D. Genetic robustness and functional evolution of gene duplicates. *Nucleic Acids Res*. 2013; gkt1200.
15. Giaever G, Chu AM, Ni L, Connelly C, Riles L, Véronneau S, et al. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature*. 2002; 418: 387–391. doi: [10.1038/nature00935](https://doi.org/10.1038/nature00935) PMID: [12140549](https://pubmed.ncbi.nlm.nih.gov/12140549/)
16. Kamath RS, Fraser AG, Dong Y, Poulin G, Durbin R, Gotta M, et al. Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature*. 2003; 421: 231–237. doi: [10.1038/nature01278](https://doi.org/10.1038/nature01278) PMID: [12529635](https://pubmed.ncbi.nlm.nih.gov/12529635/)
17. Kobayashi K, Ehrlich SD, Albertini A, Amati G, Andersen KK, Arnaud M, et al. Essential *Bacillus subtilis* genes. *Proc Natl Acad Sci U S A*. 2003; 100: 4678–4683. doi: [10.1073/pnas.0730515100](https://doi.org/10.1073/pnas.0730515100) PMID: [12682299](https://pubmed.ncbi.nlm.nih.gov/12682299/)
18. Papp B, Pál C, Hurst LD. Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature*. 2004; 429: 661–664. doi: [10.1038/nature02636](https://doi.org/10.1038/nature02636) PMID: [15190353](https://pubmed.ncbi.nlm.nih.gov/15190353/)
19. Lynch M, Conery JS. The Evolutionary Fate and Consequences of Duplicate Genes. *Science*. 2000; 290: 1151–1155. PMID: [11073452](https://pubmed.ncbi.nlm.nih.gov/11073452/)
20. Guan Y, Dunham MJ, Troyanskaya OG. Functional Analysis of Gene Duplications in *Saccharomyces cerevisiae*. *Genetics*. 2007; 175: 933–943. doi: [10.1534/genetics.106.064329](https://doi.org/10.1534/genetics.106.064329) PMID: [17151249](https://pubmed.ncbi.nlm.nih.gov/17151249/)

21. Tong AHY, Lesage G, Bader GD, Ding H, Xu H, Xin X, et al. Global mapping of the yeast genetic interaction network. *Science*. 2004; 303: 808–813. doi: [10.1126/science.1091317](https://doi.org/10.1126/science.1091317) PMID: [14764870](https://pubmed.ncbi.nlm.nih.gov/14764870/)
22. Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, et al. The genetic landscape of a cell. *Science*. 2010; 327: 425–431. doi: [10.1126/science.1180823](https://doi.org/10.1126/science.1180823) PMID: [20093466](https://pubmed.ncbi.nlm.nih.gov/20093466/)
23. Costanzo M, VanderSluis B, Koch EN, Baryshnikova A, Pons C, Tan G, et al. A global genetic interaction network maps a wiring diagram of cellular function. *Science*. 2016; 353: aaf1420.
24. van Wageningen S, Kemmeren P, Lijnzaad P, Margaritis T, Benschop JJ, de Castro IJ, et al. Functional Overlap and Regulatory Links Shape Genetic Interactions between Signaling Pathways. *Cell*. 2010; 143: 991–1004. doi: [10.1016/j.cell.2010.11.021](https://doi.org/10.1016/j.cell.2010.11.021) PMID: [21145464](https://pubmed.ncbi.nlm.nih.gov/21145464/)
25. Tischler J, Lehner B, Fraser AG. Evolutionary plasticity of genetic interaction networks. *Nat Genet*. 2008; 40: 390–391. doi: [10.1038/ng.114](https://doi.org/10.1038/ng.114) PMID: [18362882](https://pubmed.ncbi.nlm.nih.gov/18362882/)
26. Harrison R, Papp B, Pál C, Oliver SG, Delneri D. Plasticity of genetic interactions in metabolic networks of yeast. *Proc Natl Acad Sci*. 2007; 104: 2307–2312. doi: [10.1073/pnas.0607153104](https://doi.org/10.1073/pnas.0607153104) PMID: [17284612](https://pubmed.ncbi.nlm.nih.gov/17284612/)
27. Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, et al. Global analysis of protein expression in yeast. *Nature*. 2003; 425: 737–741. doi: [10.1038/nature02046](https://doi.org/10.1038/nature02046) PMID: [14562106](https://pubmed.ncbi.nlm.nih.gov/14562106/)
28. Lehner B. Molecular mechanisms of epistasis within and between genes. *Trends Genet*. 2011; 27: 323–331. doi: [10.1016/j.tig.2011.05.007](https://doi.org/10.1016/j.tig.2011.05.007) PMID: [21684621](https://pubmed.ncbi.nlm.nih.gov/21684621/)
29. Boone C, Bussey H, Andrews BJ. Exploring genetic interactions and networks with yeast. *Nat Rev Genet*. 2007; 8: 437–449. doi: [10.1038/nrg2085](https://doi.org/10.1038/nrg2085) PMID: [17510664](https://pubmed.ncbi.nlm.nih.gov/17510664/)
30. Collins SR, Kemmeren P, Zhao X-C, Greenblatt JF, Spencer F, Holstege FCP, et al. Toward a comprehensive atlas of the physical interactome of *Saccharomyces cerevisiae*. *Mol Cell Proteomics MCP*. 2007; 6: 439–50. doi: [10.1074/mcp.M600381-MCP200](https://doi.org/10.1074/mcp.M600381-MCP200) PMID: [17200106](https://pubmed.ncbi.nlm.nih.gov/17200106/)
31. Byrne KP, Wolfe KH. The Yeast Gene Order Browser: combining curated homology and syntenic context reveals gene fate in polyploid species. *Genome Res*. 2005; 15: 1456–1461. doi: [10.1101/gr.3672305](https://doi.org/10.1101/gr.3672305) PMID: [16169922](https://pubmed.ncbi.nlm.nih.gov/16169922/)
32. Dean EJ, Davis JC, Davis RW, Petrov DA. Pervasive and Persistent Redundancy among Duplicated Genes in Yeast. *PLoS Genet*. 2008; 4: e1000113. doi: [10.1371/journal.pgen.1000113](https://doi.org/10.1371/journal.pgen.1000113) PMID: [18604285](https://pubmed.ncbi.nlm.nih.gov/18604285/)
33. Musso G, Costanzo M, Huangfu M, Smith AM, Paw J, Luis B-JS, et al. The extensive and condition-dependent nature of epistasis among whole-genome duplicates in yeast. *Genome Res*. 2008; 18: 1092–1099. doi: [10.1101/gr.076174.108](https://doi.org/10.1101/gr.076174.108) PMID: [18463300](https://pubmed.ncbi.nlm.nih.gov/18463300/)
34. Kelley R, Ideker T. Systematic interpretation of genetic interactions using protein networks. *Nat Biotechnol*. 2005; 23: 561–566. doi: [10.1038/nbt1096](https://doi.org/10.1038/nbt1096) PMID: [15877074](https://pubmed.ncbi.nlm.nih.gov/15877074/)
35. Cope MJ, Yang S, Shang C, Drubin DG. Novel protein kinases Ark1p and Prk1p associate with and regulate the cortical actin cytoskeleton in budding yeast. *J Cell Biol*. 1999; 144: 1203–1218. PMID: [10087264](https://pubmed.ncbi.nlm.nih.gov/10087264/)
36. Vik Á, Rine J. Upc2p and Ecm22p, Dual Regulators of Sterol Biosynthesis in *Saccharomyces cerevisiae*. *Mol Cell Biol*. 2001; 21: 6395–6405. doi: [10.1128/MCB.21.19.6395-6405.2001](https://doi.org/10.1128/MCB.21.19.6395-6405.2001) PMID: [11533229](https://pubmed.ncbi.nlm.nih.gov/11533229/)
37. Sameith K, Amini S, Groot Koerkamp MJA, van Leenen D, Brok M, Brabers N, et al. A high-resolution gene expression atlas of epistasis between gene-specific transcription factors exposes potential mechanisms for genetic interactions. *BMC Biol*. 2015; 13: 112. doi: [10.1186/s12915-015-0222-5](https://doi.org/10.1186/s12915-015-0222-5) PMID: [26700642](https://pubmed.ncbi.nlm.nih.gov/26700642/)
38. Stillman DJ. Nhp6: A small but powerful effector of chromatin structure in *Saccharomyces cerevisiae*. *Biochim Biophys Acta BBA—Gene Regul Mech*. 2010; 1799: 175–180.
39. Yao Y, Demoinet E, Saveanu C, Lenormand P, Jacquier A, Fromont-Racine M. Ecm1 is a new pre-ribosomal factor involved in pre-60S particle export. *RNA N Y N*. 2010; 16: 1007–1017.
40. Shen S, Tobery CE, Rose MD. Prm3p is a pheromone-induced peripheral nuclear envelope protein required for yeast nuclear fusion. *Mol Biol Cell*. 2009; 20: 2438–2450. doi: [10.1091/mbc.E08-10-0987](https://doi.org/10.1091/mbc.E08-10-0987) PMID: [19297527](https://pubmed.ncbi.nlm.nih.gov/19297527/)
41. Lichtenberg U de Jensen LJ, Brunak S, Bork P. Dynamic Complex Formation During the Yeast Cell Cycle. *Science*. 2005; 307: 724–727. doi: [10.1126/science.1105103](https://doi.org/10.1126/science.1105103) PMID: [15692050](https://pubmed.ncbi.nlm.nih.gov/15692050/)
42. Benschop JJ, Brabers N, van Leenen D, Bakker LV, van Deutekom HWM, van Berkum NL, et al. A consensus of core protein complex compositions for *Saccharomyces cerevisiae*. *Mol Cell*. 2010; 38: 916–928. doi: [10.1016/j.molcel.2010.06.002](https://doi.org/10.1016/j.molcel.2010.06.002) PMID: [20620961](https://pubmed.ncbi.nlm.nih.gov/20620961/)
43. Liu H. The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases. *EMBO J*. 2002; 21: 4699–4708. doi: [10.1093/emboj/cdf448](https://doi.org/10.1093/emboj/cdf448) PMID: [12198172](https://pubmed.ncbi.nlm.nih.gov/12198172/)
44. Malys N, Carroll K, Miyan J, Tollervey D, McCarthy JEG. The “scavenger” m7GpppX pyrophosphatase activity of Dcs1 modulates nutrient-induced responses in yeast. *Nucleic Acids Res*. 2004; 32: 3590–3600. doi: [10.1093/nar/gkh687](https://doi.org/10.1093/nar/gkh687) PMID: [15240832](https://pubmed.ncbi.nlm.nih.gov/15240832/)

45. Malys N, McCarthy JEG. Dcs2, a Novel Stress-induced Modulator of m7GpppX Pyrophosphatase Activity that Locates to P Bodies. *J Mol Biol.* 2006; 363: 370–382. doi: [10.1016/j.jmb.2006.08.015](https://doi.org/10.1016/j.jmb.2006.08.015) PMID: [16963086](https://pubmed.ncbi.nlm.nih.gov/16963086/)
46. Regelmann J, Schüle T, Josupeit FS, Horak J, Rose M, Entian K-D, et al. Catabolite Degradation of Fructose-1,6-bisphosphatase in the Yeast *Saccharomyces cerevisiae*: A Genome-wide Screen Identifies Eight Novel GID Genes and Indicates the Existence of Two Degradation Pathways. *Mol Biol Cell.* 2003; 14: 1652–1663. doi: [10.1091/mbc.E02-08-0456](https://doi.org/10.1091/mbc.E02-08-0456) PMID: [12686616](https://pubmed.ncbi.nlm.nih.gov/12686616/)
47. Pitre S, Dehne F, Chan A, Cheetham J, Duong A, Emili A, et al. PIPE: a protein-protein interaction prediction engine based on the re-occurring short polypeptide sequences between known interacting protein pairs. *BMC Bioinformatics.* 2006; 7: 365. doi: [10.1186/1471-2105-7-365](https://doi.org/10.1186/1471-2105-7-365) PMID: [16872538](https://pubmed.ncbi.nlm.nih.gov/16872538/)
48. Santt O, Pfirrmann T, Braun B, Juretschke J, Kimmig P, Scheel H, et al. The Yeast GID Complex, a Novel Ubiquitin Ligase (E3) Involved in the Regulation of Carbohydrate Metabolism. *Mol Biol Cell.* 2008; 19: 3323–3333. doi: [10.1091/mbc.E08-03-0328](https://doi.org/10.1091/mbc.E08-03-0328) PMID: [18508925](https://pubmed.ncbi.nlm.nih.gov/18508925/)
49. Metallo CM, Vander Heiden MG. Understanding Metabolic Regulation and Its Influence on Cell Physiology. *Mol Cell.* 2013; 49: 388–398. doi: [10.1016/j.molcel.2013.01.018](https://doi.org/10.1016/j.molcel.2013.01.018) PMID: [23395269](https://pubmed.ncbi.nlm.nih.gov/23395269/)
50. Cairns RA, Harris IS, Mak TW. Regulation of cancer cell metabolism. *Nat Rev Cancer.* 2011; 11: 85–95. doi: [10.1038/nrc2981](https://doi.org/10.1038/nrc2981) PMID: [21258394](https://pubmed.ncbi.nlm.nih.gov/21258394/)
51. Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol.* 1970; 48: 443–453. PMID: [5420325](https://pubmed.ncbi.nlm.nih.gov/5420325/)
52. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet TIG.* 2000; 16: 276–277. PMID: [10827456](https://pubmed.ncbi.nlm.nih.gov/10827456/)
53. Yang Z, Nielsen R. Estimating Synonymous and Nonsynonymous Substitution Rates Under Realistic Evolutionary Models. *Mol Biol Evol.* 2000; 17: 32–43. PMID: [10666704](https://pubmed.ncbi.nlm.nih.gov/10666704/)
54. Conant GC, Wagner A. GenomeHistory: a software tool and its application to fully sequenced genomes. *Nucleic Acids Res.* 2002; 30: 3378–3386. PMID: [12140322](https://pubmed.ncbi.nlm.nih.gov/12140322/)
55. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. *Bioinforma Oxf Engl.* 2014; 30: 1236–1240.
56. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 2016; 44: D279–D285. doi: [10.1093/nar/gkv1344](https://doi.org/10.1093/nar/gkv1344) PMID: [26673716](https://pubmed.ncbi.nlm.nih.gov/26673716/)