Co-occurrence of essential gene dispensability and bypass suppressor mutations across species

Carles Pons[1,*], Patrick Aloy[1,2], Jolanda van Leeuwen[3]

[1]Institute for Research in Biomedicine (IRB Barcelona), the Barcelona Institute for Science and Technology, Barcelona, Catalonia (Spain).
[2]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia (Spain).
[3]Center for Integrative Genomics, Bâtiment Génopode, University of Lausanne, 1015 Lausanne, Switzerland.

[*]Corresponding author: carles.pons@irbbarcelona.org

ABSTRACT

Essential genes that become dispensable upon specific changes in the genetic background represent a valuable model to understand the incomplete penetrance of loss-of-function mutations and the emergence of drug resistance mechanisms. Systematic identification of dispensable essential genes has recently challenged the canonical binary categorization of gene essentiality. Here, we compiled data from multiple studies on essential gene dispensability in *Saccharomyces cerevisiae* to comprehensively characterize these genes. In analyses spanning different evolutionary time-scales, ranging from *S. cerevisiae* strains to human cell lines, dispensable essential genes exhibited distinct phylogenetic properties compared to other essential and non-essential genes. Integration of interactions with suppressor genes unveiled the emergent high functional modularity of the bypass suppression network. Also, dispensable essential and suppressor gene pairs reflected simultaneous changes in the mutational landscape of *S. cerevisiae* strains. Importantly, species in which dispensable essential genes were non-essential tended to carry bypass suppressor mutations in their genomes. Overall, our study offers a comprehensive view of dispensable essential genes and illustrates how their interactions with bypass suppressor genes reflect evolutionary outcomes.

INTRODUCTION

Identification of the genes required for viability is key for research in biomedicine and biotechnology. Essential genes constrain genome evolution (Jordan *et al.*, 2002; Bergmiller, Ackermann and Silander, 2012; Luo, Gao and Lin, 2015), identify core cellular processes (Wang *et al.*, 2015), make efficient drug targets in pathogens and tumors (Roemer *et al.*, 2003; Behan *et al.*, 2019), and are the starting point to determine minimal genomes (Juhas, Eberl and Glass, 2011; Hutchison *et al.*, 2016). The fraction of essential genes within a genome reflects its complexity and redundancy, and anticorrelates with the number of encoded genes (Rancati *et al.*, 2018). For instance, 80% of 482 genes in *M. genitalium* (Glass *et al.*, 2006), 18% of ~6,000 genes in *S. cerevisiae* (Giaever *et al.*, 2002), and only 10% of the encoded ~20,000 genes in human cell lines (Blomen *et al.*, 2015; Hart *et al.*, 2015; Wang *et al.*, 2015) are essential for viability. Essential genes tend to code for protein complex members (Dezso, Oltvai and Barabási, 2003; Hart, Lee and Marcotte, 2007), play central roles in genetic networks (Costanzo *et al.*, 2010), have few duplicates (Giaever *et al.*, 2002), and share other properties (Deng *et al.*, 2011; Hart *et al.*, 2015) that differentiate them from non-essential genes, enabling their prediction (Hwang *et al.*, 2009; Lloyd *et al.*, 2015; Zhang, Acencio and Lemke, 2016). Although gene essentiality is significantly conserved, essentiality changes are not rare across species and even between individuals. For instance, 17% of the 1:1 orthologs between *S. cerevisiae* and *S. pombe* have different essentiality (Kim *et al.*, 2010). Also, 57 genes differ in essentiality between two closely related *S. cerevisiae* strains (Dowell *et al.*, 2010), and a systematic analysis of 324 cancer cell lines from 30 cancer types found that only ~40% of the essential genes were shared across cell lines (Behan *et al.*, 2019). Thus, essentiality is not a static property, and modifications in the environment and the genetic background can change the essentiality of a gene (Rancati *et al.*, 2018)**.**

Recently, we and others have systematically identified essential genes that become non-essential in the presence of suppressor mutations in *S. cerevisiae* (Liu *et al.*, 2015; van Leeuwen *et al.*, 2020) and *S. pombe* (Li *et al.*, 2019; Takeda *et al.*, 2019). Dispensable essential genes and their bypass suppressors (i.e. the genetic changes enabling the bypass of gene essentiality) represent an extreme case of suppression interaction, a positive genetic interaction in which the resulting double mutant is healthier than the sickest individual mutant (Mani *et al.*, 2008), in this case an unviable cell. Both dispensable essential and suppressor genes exhibit specific features that differentiate them from other essential genes (i.e. core essential genes) and passenger mutations (i.e. randomly acquired mutations without an effect on fitness), respectively, which we previously exploited for their successful prediction (van Leeuwen *et al.*, 2020). For instance, dispensable essential genes are more likely to have paralogs, to be non-essential in other *S. cerevisiae* strains, to be absent in other species, and to not code for protein complex members (Liu *et al.*, 2015; van Leeuwen *et al.*, 2020), whereas suppressor genes tend to be functionally related to the dispensable essential gene (van Leeuwen *et al.*, 2020).

Though dispensable essential genes can be identified without determining their suppressor, knowledge of the suppressor genes is important to dissect the function of both genes (van

Leeuwen *et al.*, 2016), to expose the genetic architecture of the cell, and to understand drug resistance mechanisms (Woodford and Ellington, 2007) and the existence of presumably very detrimental genetic variants in natural populations (Jordan *et al.*, 2015; Narasimhan *et al.*, 2015; Chen *et al.*, 2016). Indeed, genetic suppression could explain why variants that are pathogenic in human are fixed in other mammalian species without obvious deleterious consequences (Jordan *et al.*, 2015) and the fact that some individuals are healthy despite carrying highly penetrant disease-associated mutations (Chen *et al.*, 2016), enabling a broader evolutionary landscape for otherwise highly constrained genes. Thus, suppressor mutations restoring the viability of deleterious gene variants should be fixed in species in which the deleterious variant is prevalent. Certainly, genetic interactions are known to partially constrain the outcome of controlled evolutionary experiments (Tenaillon *et al.*, 2012; Good *et al.*, 2017; Vignogna, Buskirk and Lang, 2021). In spite of the differences with suppression screens (LaBar *et al.*, 2020), evolutionary repair experiments also aim to identify the genetic changes that overcome the fitness defect of a particular mutant strain. In these experiments, compensatory evolution usually takes place through loss-of-function (LOF) mutations (Szamecz *et al.*, 2014), evolutionary trajectories are similar for mutants in functionally similar genes (Rojas Echenique *et al.*, 2019; Fumasoni and Murray, 2020), and the acquired mutations are enriched in functional associations with the initially mutated gene (Szamecz *et al.*, 2014). Although these properties are mostly shared with suppressor screens (van Leeuwen *et al.*, 2020), there are no reports of systematic agreement between the initially compromised genes and the acquired mutations identified in repair experiments with the positive genetic or suppression interaction network (Klim *et al.*, 2021). Thus, whether suppression interactions are reflected in the real evolutionary landscape and explain the existence of puzzling genetic variants is still an open question.

Here, we compiled a comprehensive set of dispensable essential genes in *S. cerevisiae* identified across different studies to exhaustively compare their properties to core essential and non-essential genes, with a particular focus on phylogenetic features. We integrated bypass suppressor genes into an interaction network with dispensable essential genes to identify prevalent interaction motifs and to analyze the relationship of bypass suppression pairs in other species. This work presents a systematic characterization of dispensable essential genes and explores how evolution reflects their interactions with suppressors.

RESULTS

Dispensable essential genes

We compiled a comprehensive list of dispensable essential genes in *S. cerevisiae* from two large-scale studies (Liu *et al.*, 2015; van Leeuwen *et al.*, 2020) and from individual cases described in the literature (van Leeuwen *et al.*, 2020) (**Figure 1A**), totalling 205 dispensable genes and representing ~20% of the tested essential genes (**Figure 1B**). In spite of the different experimental techniques used, the dispensable essential genes identified in the three datasets significantly overlapped (p < 0.001). Except in the case of the literature dataset, dispensable essential genes were more likely than core essential genes to be non-essential in the closely related *S. cerevisiae* strain Sigma1278b (**Figure 1C**), and to be absent in the *S. cerevisiae* core pangenome (**Figure S1A**). The three datasets presented similar functional enrichments (**Figure S1B**), with central cellular processes like RNA processing or translation depleted for dispensable essential genes, and more peripheral functions related to signaling or transport enriched for dispensable essentiality, as previously reported (van Leeuwen *et al.*, 2020). The combined dataset contained 14 protein complexes with all essential subunits dispensable (**Figure S1C**), more than expected (**Figure S1D**) and mostly driven by cases identified in our recent study (van Leeuwen *et al.*, 2020). Interestingly, protein complexes tended to be either completely dispensable or indispensable (**Figure S1D**). Since the quality and basic properties of the combined and individual datasets were similar, we focused the following analyses on the combined dataset.

Properties of dispensable essential genes

By querying an extensive panel of gene features, we compared the properties of dispensable and core essential genes. Dispensable essential genes tended to be more multifunctional, to exhibit more stable gene expression levels and lower transcript counts, to be less conserved across species, to have more gene duplicates and higher evolutionary rates, to be coexpressed with fewer genes, and to code for proteins more often localized to the membrane, not present in complexes, without structural domains, and with fewer protein-protein interactions (PPIs), lower abundance, and shorter half-life (**Figure 1D**). Since the observed differences between dispensable and core essential genes resembled the differences between non-essential and essential genes (**Figure S1E**), we asked whether dispensable essential and non-essential genes shared the same properties, and found they comprised two distinct classes of genes with clearly distinct features (**Figure 1D**). Broadly, features of dispensable essential genes laid between those of core essential and non-essential genes, consistent with and extending previous findings (Liu *et al.*, 2015).

Phylogenetic analysis of dispensable essential genes

We found the differences in gene conservation across other yeasts and distant species intriguing and decided to further explore the phylogeny of *S. cerevisiae*, starting with a large panel of sequenced *S. cerevisiae* strains (Peter *et al.*, 2018). Dispensable essential genes were more likely than core essential genes, but less than non-essential genes, to harbor deleterious mutations disrupting protein sequences (**Figure S2A**), to present higher local evolutionary rates (**Figure S2B**), and to undergo copy number loss events in other *S. cerevisiae* strains (**Figure**

S2C). Interestingly, the relative rate of copy number loss events affecting core and dispensable essential genes increased with respect to non-essential genes in strains carrying more gene loss events, suggesting that broader changes in the genetic background are required to enable variation in essential genes (**Figure S2D**).

To further investigate differences in the evolutionary pressure on dispensable essential and core essential genes, we analyzed essentiality data and orthology relationships in *Candida albicans*, *Schizosaccharomyces pombe*, and human cell lines (**Figure 2A, S2E, S2F**). Genes that were dispensable essential in *S. cerevisiae* were more often absent than core essential genes in each of the analyzed species (**Figure 2B**) which reflects either: i) genes not present in their common ancestor and specific to the *S. cerevisiae* phylogenetic branch, or ii) genes present in their common ancestor but lost in the phylogenetic branch of the analyzed species. To determine the contribution of each factor, we calculated the age of each *S. cerevisiae* gene by identifying the farthest species with an orthologous gene. We found dispensable essential genes to be enriched for younger genes with respect to core essential genes (**Figure 2C**), particularly for genes with no ortholog in any other species (i.e. the most specific to *S. cerevisiae;* **Figure 2D**). Next, for each species we defined lost genes as those absent in that species but present in its common ancestor with *S. cerevisiae* and found dispensable essential genes more often lost than core essential genes (**Figure 2E**). Thus, the absence of dispensable essential genes in other species can be explained both by gene loss events in those species and by *S. cerevisiae* specific genes.

Dispensable essential genes present in other species were more frequently duplicated and had more N:1 orthology relationships (**Figure 2B**) than core essential genes. For genes with a 1:1 ortholog in other species, dispensable essential genes were enriched for non-essential orthologs (**Figure 2B**), also in the closely related *S. uvarum* species (**Figure S2G**). We show the comparison between essential and non-essential genes, and dispensable essential and non-essential genes to contextualize the observed differences (**Figure S2H-J**). Additionally, fitness data from a panel of 1,070 cancer cell lines (Meyers *et al.*, 2017) revealed that knockout of dispensable essential genes led to less severe proliferation defects than knockout of core essential genes (**Figure 2F**).

Finally, we compared sequences of *S. cerevisiae* proteins and their 1:1 orthologs in *S. pombe* and *C. albicans*. Sequence divergence for dispensable essential and non-essential genes was indistinguishable, but products of dispensable essential genes had lower sequence identity and differed more in sequence length than core essential proteins (**Figure 2G, S2K-M**), in line with the dN/dS data (**Figure 1D, S2B**). Overall, orthology relationships, phenotypic changes, and sequence divergence reflect that the evolutionary pressure on dispensable essential genes is more lenient than on core essential genes but more strict than on non-essential genes (**Figure 2B, S2I**).

The bypass suppressor interaction network
Identification of the relevant genetic changes (i.e. suppressors) required to tolerate the deletion of an essential gene is key to interpreting the presence of deleterious genetic variants in natural

populations. To improve our knowledge on the mechanisms of genetic suppression, we built an interaction network between dispensable essential genes and their bypass suppressors by combining data from our recent systematic study (van Leeuwen *et al.*, 2020) and the literature (van Leeuwen *et al.*, 2020). This network included a total of 319 unique bypass suppression gene pairs (**Figure 3A**), corresponding to 243 suppressors and 137 dispensable essential genes out of the 205 dispensable essential genes known. For the remaining dispensable essential genes (33% of the dataset), the suppressor variants were not identified. The two individual suppression interaction networks overlapped significantly ($p < 0.001$) and were similarly enriched in functional associations (**Figure S3A**). In the combined dataset, both dispensable essential and suppressor genes tended to be functionally related (**Figure S3B**), particularly for close functional relationships like cocomplex or copathway membership (**Figure S3A**), and suppressors related to translation and transcription processes were prevalent (**Figure S3B**). For 50% and 26% of the dispensable genes, only LOF and GOF suppressors were isolated, respectively, and in 15% of the cases, both types of suppressors were identified (**Figure S3C**). For the remaining cases, the nature of the suppressor could not be determined. In agreement with their suppression mode, deletion or TS mutants of LOF suppressor genes showed more positive than negative genetic interactions with TS alleles of the corresponding dispensable essential gene. Conversely, deletion or TS mutants of GOF suppressor genes had more negative than positive genetic interactions with TS mutants of the corresponding dispensable essential gene (**Figure 3B**).

Structure of the bypass suppression interaction network
Network density (i.e. the percentage of gene pairs with an interaction) ranged from 0.007% to 0.96% depending on whether we considered all tested gene pairs or only the identified genes in the network, respectively. In spite of the network sparsity, several patterns emerge showing its structure and modularity. For instance, all dispensable essential genes in the same protein complex tended to interact with either GOF or LOF suppressors. These monochromatic interactions affected 13 out of 17 non-redundant complexes with at least two dispensable essential subunits in our dataset (**Figure 3C**; $p < 0.028$), suggesting similar suppression types apply for functionally related genes. Importantly, both individual suppression networks contributed to this result (**Figure S3D**), discarding the potential bias from specific hypothesis-driven experiments in the literature dataset. We analyzed the topology of the network and found 45% of the dispensable genes had multiple suppressors (**Figure S3E**). These dispensable essential genes were more likely to have lower expression variance and coexpression degree (i.e. share similar expression with fewer other genes), to be multifunctional, and to code for proteins with more structural domains and lower abundance than dispensable genes with a single suppressor (**Figure S3F**). Suppressors were more specific than dispensable essential genes, and only 23% interacted with multiple genes (**Figure S3E**). Next, we explored the relationship between gene functional similarity and connectivity patterns, and found 52% of the dispensable essential genes belonging to the same complex shared suppressor genes, and 70% of the suppressor genes in the same complex shared dispensable essential genes (**Figure 3D**), more than expected by chance ($p < 0.05$).

To illustrate the underlying modular structure of the bypass suppression interaction network, we explored the connectivity of *NCB2* and *BUR6*, both dispensable essential genes with known suppressors and the only two members of the negative cofactor 2 complex (ID CPX-1662 in the Complex Portal (Meldal *et al.*, 2021)). *NCB2* and *BUR6* have seven and ten suppressor genes, respectively, six of which are in common, showing that functionally related dispensable genes tend to have overlapping suppressors (**Figure 3E**). Two of these common suppressors belong to the core Mediator complex (CPX-3226), reflecting that interactors of the same dispensable gene tend to be functionally related, and the rest to the transcription factor TFIIA complex (CPX-1633), the general transcription factor complex TFIIH (CPX-1659), and the DNA-directed RNA polymerase II complex (CPX-2662). Interestingly, the *NCB2* specific suppressor, *TOA2*, also belongs to TFIIA, and three of the four *BUR6* specific suppressors to RNA pol II or Mediator, further illustrating the modularity of the network. In another example (**Figure S3G**), members of the RPD3L histone deacetylase complex (CPX-1852) suppress two different protein complexes. *DEP1*, *SAP30*, and *SDS3* suppress the two essential subunits of piccolo NuA4 histone acetyltransferase complex (CPX-3185), whereas *RPD3*, *SIN3*, and *SDS3* interact with the Rer2 subunit of the dehydrodolichyl diphosphate synthase complex (CPX-162). This modularity in the suppression interaction pattern of RPD3L subunits suggests a functional modularity within the complex which is in fact supported by its modeled structure (Sardiu *et al.*, 2009) and genetic interaction patterns (**Figure S3H**).

Mutational landscape of *S. cerevisiae* strains reflects bypass suppression relationships

We wondered if the suppression interaction network could capture the variability observed in the genomes of the natural population. For that, we first evaluated if bypass suppression gene pairs in our network reflected simultaneous gene copy number changes across *S. cerevisiae* strains. For each gene pair and strain, we evaluated if the dispensable essential gene presented a copy number loss (CNL) event (partially resembling the effect of a gene deletion) and the suppressor gene either a CNL or a copy number gain (CNG) event (equivalent to LOF and GOF suppressor mutations, respectively). Remarkably, the co-loss of both dispensable essential and suppressor genes was enriched for LOF bypass suppression pairs (**Figure 4A**). In contrast, loss-gain events of dispensable essential and suppressor genes, respectively, were enriched for GOF bypass suppression (**Figure 4A**), also when normalizing by strain (**Figure S4A**). Next, we asked whether deleterious coding mutations in dispensable essential genes and in identified bypass suppressor genes co-occurred in *S. cerevisiae* isolates. We only considered haploid strains so the deleterious effects of mutations would not be masked by other alleles. In LOF bypass suppression pairs, we found 18 cases in which both the dispensable essential gene and the suppressor gene carried deleterious mutations in at least one of the haploid strains (**Figure 4B**), significantly more than in randomized gene pairs (**Figure 4C**; p<0.014). As expected, we did not observe a similar enrichment in diploid strains (**Figure S4B**) nor for GOF bypass suppression gene pairs (**Figure S4C**). Thus, the bypass suppression network reflects evolutionary outcomes in natural *S. cerevisiae* strains.

Co-occurrence of viability changes and fixed bypass suppressor mutations

We have shown that dispensable essential genes are often non-essential in other species (**Figure 2B**). Differences in the environment or the genetic background may be responsible for

these changes in essentiality. Here, we hypothesized that the genetic changes that bypass the essentiality of a gene in *S. cerevisiae* should be reflected in the genome of species in which the gene is also dispensable (i.e. non-essential or absent). To test this, we evaluated whether changes in essentiality for dispensable essential genes in a given target species co-occurred with bypass suppressor mutations that were fixed in the genome. Given that genome-scale essentiality data is scarce, we focused our analysis on *S. pombe*, for which high quality essentiality data is available for most genes (Harris *et al.*, 2022). It should be noted that, usually, the identified suppressor mutation is not the only possible change in the suppressor gene to bypass the essentiality of the dispensable gene. For instance, the essentiality of gene *ECM9* can be bypassed by eight different nonsynonymous substitutions in gene *YNL320W*, suggesting that other mutations may have the same effect. Thus, the lack of saturation in the suppressor sequence space makes it unlikely to identify the exact same *S. cerevisiae* bypass suppressor mutation fixed in the target species. This task is even more challenging when considering the changes in genetic context of the target species, which may enable novel suppression variants. However, in most cases knowledge of the specific bypass suppressor mutation is not paramount, and the relevant information lies in the identity of the bypass suppressor gene and the suppression mode. For instance, gene deletion and overexpression of bypass suppressor genes usually mirrored the phenotypic suppression of more specific and local mutations, such as missense variants (van Leeuwen *et al.*, 2020). This genotype-to-phenotype funneling is key in defining a set of coarse-grained rules to identify equivalent bypass suppressor variants in other species.

Next, we briefly describe our approach (see Methods). First, we annotated the orthology relationships of dispensable essential genes and bypass suppressor genes in *S. pombe*. For genes with 1:1 orthologs, we annotated the essentiality of the dispensable essential genes in that species, and compared the sequences of suppressors to their orthologs. Then, we evaluated whether the *S. pombe* genome carried known equivalent bypass suppressor mutations by analyzing orthology relationships and protein sequences. We considered as equivalent bypass mutations those that could reduce or increase the gene activity, for LOF and GOF suppressors, respectively. For instance, we considered an absent ortholog in *S. pombe* as equivalent to a LOF bypass suppressor variant since it mirrors a gene deletion in *S. cerevisiae*. Conversely, we considered a duplicated ortholog as equivalent to a GOF bypass suppressor variant since it resembles gene overexpression. Also, even if rare, we considered an ortholog harboring exactly the same suppressor missense substitution as an equivalent bypass suppressor. Finally, we grouped dispensable essential genes by their dispensability in *S. pombe* into genes with equivalent phenotypes (i.e. absent or have a 1:1 non-essential ortholog in *S. pombe*), and genes with non-equivalent phenotypes (i.e. 1:1 essential ortholog in *S. pombe*), expecting dispensable essential genes with equivalent phenotypes in *S. pombe* to more often co-occur with equivalent bypass suppressors than dispensable essential genes that are essential in *S. pombe*.

We found that 67% (18/27) of the dispensable essential genes that are non-essential in *S. pombe* co-occurred with bypass suppressor mutations in that species, whereas this happened only in 27% (12/44) of the dispensable essential genes that were essential in *S. pombe* (2.4 fold

enrichment; $p < 0.05$; **Figure 5A, 5B**). A similar trend (53%) was observed for dispensable essential genes that were absent (i.e. without an ortholog) in *S. pombe*, although this difference was not significant compared to the set of essential orthologs (**Figure 5B**). In order to increase the statistical power of our analyses, we combined the non-essential and absent genes in *S. pombe* into a single set (i.e. dispensable essential genes with equivalent phenotype in *S. pombe*) and observed a clear difference with the essential orthologs (2.2 fold enrichment; $p < 0.05$).

We controlled for potential biases to ensure the robustness of our observation (**Figure 5B**), starting with gene degree by generating 1,000 randomized bypass suppression networks while respecting the original topology (**Figure 5C**) and by considering only dispensable essential genes with a single bypass suppressor to avoid the potential bias introduced by gene degree (**Figure S5A**). Additionally, we removed bypass suppression interactions from the literature which may have been potentially identified because of phylogenetic properties (**Figure S5B**), functionally related bypass suppression pairs which may be prone to present similar evolutionary patterns (**Figure S5C**), and every node in the network to discard dependence on a single gene (**Figure S5D**). Also, we applied three alternative orthology mappings (**Figure S5E**), and used essentiality annotations and orthology mappings from *C. albicans* (**Figure 5D, 5E, S5F**). In all these analyses, bypass suppressor mutations co-occurred more often with absent and non-essential orthologs than with essential orthologs. Conversely, switching LOF and GOF annotations resulted in a non-significant difference, as expected (**Figure S5G**), further showing the specificity of the bypass suppression associations and the set of equivalence rules we defined.

DISCUSSION

Differences between essential and non-essential genes have been widely characterized (**Figure S1E, S2H, S2J**) and a myriad of machine learning algorithms have exploited this information for the successful prediction of gene essentiality (Hwang *et al.*, 2009; Lloyd *et al.*, 2015; Zhang, Acencio and Lemke, 2016). Recently, we and others have identified a subset of *S. cerevisiae* essential genes that become dispensable under specific genetic conditions (Liu *et al.*, 2015; van Leeuwen *et al.*, 2020). Here, we have combined these datasets of dispensable essential genes, after showing they exhibit similar properties (**Figure 1**), for the comprehensive characterization of these genes. While recapitulating previously reported features, we have also unveiled properties not identified before (**Figure 1D, 2**), probably because of the limited statistical power of the smaller individual datasets, such as low cocomplex degree, low expression variance, high multifunctionality, 1:N and N:1 orthology relationships, and specific mutational signatures, which can complement existing methods for the prediction of essential gene dispensability (van Leeuwen *et al.*, 2020). Since properties of dispensable essential genes are highly conserved (van Leeuwen *et al.*, 2020), predictions could potentially target other species. Though the differences between dispensable essential and core essential genes resemble the differences between essential and non-essential genes (**Figure 1D, S1E, F2B, S2J**), dispensable essential and non-essential genes also make up two clearly distinct groups (**Figure 1D, S2I**). Thus, besides the classical binary classification of genes based on their essentiality, three different sets of genes emerge with specific properties: non-essentials, dispensable essentials, and core essentials, as previously suggested (Liu *et al.*, 2015).

Importantly, we presented extensive evidence of the distinct evolutionary pressure exerted on these gene sets by performing phylogenetic analyses spanning very different evolutionary time-scales, from *S. cerevisiae* strains to human cancer cell lines. Dispensable essential genes were frequently absent, non-essential, or mutated in other *S. cerevisiae* strains (**Figure S2**). By extending our phylogenetic analysis to other species (**Figure 2**), dispensable essential genes showed distinct orthology relationships, gene age, gene loss patterns, essentiality changes, and sequence divergence, further expanding previous observations focused mostly on gene absence profiles and essentiality changes across species (Liu *et al.*, 2015; van Leeuwen *et al.*, 2020). The observed differences in *S. uvarum*, *C. albicans*, *S. pombe*, and even human, which diverged from *S. cerevisiae* ~1 billion years ago, reflect the deep evolutionary footprint of essential gene dispensability.

For a better characterization of the mechanisms associated with the tolerance of highly deleterious mutations, we integrated data from multiple studies to build a bypass suppression interaction network between dispensable essential genes and their suppressors. In spite of the low density of this network, several properties emerge reflecting its modularity and structure. Complexes tended to be either composed of only dispensable essential subunits or of only core essential subunits (**Figure S1D**), mirroring the essentiality composition bias previously described (Hart, Lee and Marcotte, 2007) and the functional modularity that complexes encapsulate. Dispensable essentiality, thus, would be a modular feature of protein complexes (Li *et al.*, 2019), as is essentiality. Also, protein complexes exhibited monochromaticity of

10

suppressor type (**Figure 3C**), with members of the same complex being all suppressed by either LOF or GOF mutations and suggesting that similar suppression mechanisms apply for functionally related genes. Last, members of the same complex exhibited interaction coherence, with cocomplexed dispensable essential genes sharing suppressors and cocomplexed suppressor genes interacting with the same dispensable essential genes (**Figure 3D**), as illustrated in **Figure 3E and S3G**. All these observations expose the inherent modularity of the bypass suppression network, which can lead to the identification of new dispensable essential and suppressor genes, and to the characterization of their suppression mode. Certainly, network modularity is not restricted to the bypass suppression network, and it is in fact a hallmark of the genetic interaction network (Costanzo *et al.*, 2016), but it is particularly relevant here given its directionality, small size, and low interaction density, reflecting the strong functional relationships it encapsulates.

Evolution experiments, in which the fitness defect of an initially compromised gene is overcome by the acquisition of additional mutations, are closely related to suppression screens (LaBar *et al.*, 2020). Although the relationship between the acquired mutations and the compromised gene is similar to suppression interactions (Szamecz *et al.*, 2014), no systematic overlap has been reported between both approaches. Thus, the potential role of genetic suppression in explaining the existence of deleterious variants among natural populations (Chen *et al.*, 2016) is still not fully understood. To address this knowledge gap, we evaluated how bypass suppression gene pairs reflected simultaneous changes across evolution. Remarkably, we found co-occurrence of copy number changes and deleterious mutations in both the dispensable essential and the suppressor gene across *S. cerevisiae* strains, suggesting that within-species variability can follow the same evolutionary paths as spontaneous mutations in a laboratory environment. Importantly, dispensable essential genes that were absent or non-essential in *S. pombe* were more likely to co-occur with a bypass suppressor mutation in the *S. pombe* genome (**Figure 5**). This co-occurrence was robustly found also after controlling for several potential biases, using different orthology mappings and, importantly, in *C. albicans*. Nevertheless, several cases were not consistent with our hypothesis that equivalent phenotypes co-occur with equivalent suppressor mutations. Cases of dispensable essential genes with equivalent phenotypes in *S. pombe* (i.e. absent or with a non-essential ortholog) but without an equivalent bypass suppressor in *S. pombe* may reflect that the suppressor sequence space is not saturated and other suppressor missense mutations may be present in the ortholog. Also, the suppressor gene space is not saturated in the systematic dataset (van Leeuwen *et al.*, 2020), and probably even less in the literature dataset, suggesting that other unreported bypass suppressor genes could have an equivalent mutation fixed in *S. pombe*. Conversely, cases of dispensable genes with essential orthologs and bypass suppressors fixed in the target genome may be related to rewiring of the underlying genetic network creating further genetic dependencies. Still, and in spite of the limitations of the approach and the inherent noise introduced by evolution, the coarse-grained rules defined here robustly identified equivalent bypass suppressor mutations enriched for dispensable essential genes with an equivalent phenotype in other species, illustrating the constraints genetic networks may impose on evolutionary outcomes.

CONCLUSION

We compiled a comprehensive list of dispensable essential genes across different studies, which enabled the identification of new features of these genes. Phylogenetic analyses across *S. cerevisiae* strains and *S. pombe*, *C. albicans*, and human, illustrated the more lenient evolutionary pressure affecting dispensable essential genes compared to core essential genes. The interaction network between dispensable essential and bypass suppressor genes exhibited a strong functional modularity. Importantly, the mutational landscape of *S. cerevisiae* strains reflected the bypass suppression relationships. Integration of phenotypic data from other species revealed that changes in essentiality across species co-occur with the presence of fixed bypass suppressor mutations. Overall, our study provides an in-depth characterization of dispensable essential genes and unveils how bypass suppression relationships reflect on the evolutionary landscape.

METHODS

Dispensable essential gene analyses

Dispensable essential gene datasets

We retrieved dispensable essential genes in *S. cerevisiae* from two systematic experimental datasets (Liu *et al.*, 2015; van Leeuwen *et al.*, 2020) and from two studies that compiled data from the literature (van Leeuwen *et al.*, 2020), which we filtered to select only essential gene deletions rescued in standard conditions. The set of tested genes are explicitly mentioned in the systematic studies, whereas for the literature set they are unknown and, therefore, we used all essential genes in *S. cerevisiae*. The combined dataset contained the dispensable essential genes identified in any of the three individual datasets. As tested genes, we considered all tested genes in the systematic studies and the dispensable genes identified in the literature set. We randomly generated 1,000 sets of genes of the same sizes as the individual datasets, sampling from the corresponding set of tested genes.

We calculated the overlap between the different datasets by counting the number of dispensable genes found across two and three datasets. We repeated the same process in the randomly generated datasets to derive empirical p-values.

Essentiality data

In our analyses, we used essentiality data from *S. cerevisiae* (van Leeuwen *et al.*, 2020), *S. uvarum* (Sanchez *et al.*, 2019), *C. albicans* (Segal *et al.*, 2018), *S. pombe* (downloaded in November 2021 from Pombase (Harris *et al.*, 2022)), and human cell lines (Hart *et al.*, 2015). We considered human essential genes those that were required for viability in at least three of the five cell lines tested. In *C. albicans*, genes with essentiality confidence scores above 0.5 were classified as essential, and the remaining genes as non-essential.

Orthology mappings

We used PantherDB 16.1 (Mi *et al.*, 2021) to identify orthology relationships. When indicated, we also used OrthoMCL (Li, Stoeckert and Roos, 2003), SonicParanoid (Cosentino and Iwasaki, 2019), based on the popular Inparanoid (Sonnhammer and Östlund, 2015), and Pombase (Wood *et al.*, 2012) orthology mappings.

Functional enrichment of dispensable essential genes

For each dispensable essential gene set and functional class, we calculated the fold enrichment as the fraction of dispensable essential genes annotated to that functional class with respect to the corresponding fraction of core essential genes. Statistical significance was calculated with two-sided Fisher's exact tests.

Enrichment for non-essential genes in the Sigma 1278b strain

For each dispensable gene set, we calculated the fold enrichment as the ratio of dispensable essential genes identified as non-essential in the Sigma 1278b strain divided by the analogous ratio of core essential genes. P-values were calculated using two-sided Fisher's exact tests.

13

Complex dispensability bias

For each dispensable essential gene set, we counted the number of complexes (Meldal *et al.*, 2021) in which all essential subunits were identified either as dispensable or core essential genes. We repeated the same process using the randomly generated datasets to derive empirical p-values.

Properties of dispensable essential genes

We queried a panel of gene features previously defined (van Leeuwen *et al.*, 2020). For numerical features, values of dispensable genes were z-score normalized using the median and standard deviation of the core essential genes. Dot size in plots is proportional to the median z-score value. We calculated the statistical significance by means of Mann-Whitney U tests. Only dots of significant differences (two-sided p-value < 0.05) are shown. For boolean features, we calculated the fold enrichment as the ratio of dispensable essential genes with that feature divided by the equivalent ratio of core essential genes. We calculated the p-values with Fisher's exact tests. We followed the same approach to characterize: i) dispensable essential vs non-essential genes; ii) essential vs non-essential genes; iii) dispensable essential genes with multiple suppressors vs single suppressors; iv) dispensable essential genes with LOF bypass suppressors vs GOF bypass suppressors.

Analyses on *S. cerevisiae* strains

We downloaded gene presence/absence data for a large panel of *S. cerevisiae* strains (Li, Ji and Nielsen, 2019) and defined several core pangenome gene sets at different stringency levels (see x-axis in **Figure S1A**). For instance, a threshold of ten identifies the core pangome composed of all genes absent only in ten strains or less. For each dispensable essential gene dataset, we calculated the fraction of dispensable essential genes missing from the pangenome and the corresponding fraction for core essential genes, from which we calculated the fold enrichment. P-values were calculated with two-sided Fisher's exact tests.

We retrieved precomputed loss-of-function data for *S. cerevisiae* strains (Peter *et al.*, 2018) from http://1002genomes.u-strasbg.fr/files/, including frameshift mutations and missense mutations predicted to be deleterious by SIFT (Ng and Henikoff, 2001), and added indel mutations calculated from the sequencing data. We calculated the number of strains in which these mutations affected each gene and aggregated the results per gene set (i.e, dispensable essentials, core essentials, and non-essentials). P-values were calculated using Fisher's exact tests.

For each strain, we counted the genes affected by copy number loss (CNL) events in a panel of *S. cerevisiae* strains (Peter *et al.*, 2018) and aggregated the result per gene set. P-values were calculated using Fisher's exact tests. Then, we grouped the strains by their number of CNL events, which we used as a measure of distance to the reference strain, in three sets of equal size. For each set of strains, we calculated the proportion of CNL events that corresponded to each of the gene sets. P-values were calculated by Fisher's exact tests comparing the fraction of CNL events corresponding to dispensable essential genes across strain sets. Finally, we

14

retrieved dN/dS data for the same panel of *S. cerevisiae* strains and grouped them by gene set. P-values were calculated using Mann-Whitney U tests.

Orthology relationships of dispensable essential genes

For each gene, we calculated its orthology relationships in *C. albicans*, *S. pombe*, and human. Specifically, we considered gene absence, gene duplication (including 1:N and N:M orthology relationship), N:1 relationships, and 1:1 orthologs. For 1:1 orthologs, we evaluated the essentiality in the target species. For each species and property, the fold enrichment was calculated as the fraction of dispensable essential genes with respect to the fraction of core essential genes with that property. P-values were calculated by Fisher's exact tests. We used the same approach to compare dispensable essential to non-essential genes, and non-essential to essential genes.

Gene age

For each gene, we calculated its age by identifying the farthest species from *S. cerevisiae* with a present ortholog. We used orthology relationships for 98 species from PantherDB (Mi *et al.*, 2021). The phylogenetic tree to calculate species relationships was downloaded from Uniprot (UniProt Consortium, 2021), and for each species we calculated the distance to *S. cerevisiae* as the number of main branches separating them. Thus, genes with age 0 are specific to *S. cerevisiae* and not present in any other of the 98 species, whereas age 5 corresponds to genes present in the most distantly related species. We grouped gene ages for each gene set (core, dispensable, and non-essentials) and calculated p-values with Mann-Whitney U tests.

Gene loss

For each gene of age X, we calculated the fraction of species closer to *S. cerevisiae* (distance < X) in the phylogenetic tree with that gene absent from their genome. For instance, for a given gene of age 3, we calculated the fraction of species at distance 1 or 2 to *S. cerevisiae* with the gene of interest absent. We aggregated data for each gene set (core essentials, dispensable essentials, and non-essentials) and calculated p-values by means of Mann-Whitney U tests. Also, we specifically evaluated gene loss taking place only in *S. pombe* and *C. albicans*, by considering genes with ages higher than their distance to *S. cerevisiae* (i.e. genes found in any of their common ancestors). P-values were calculated using Fisher's exact tests.

Cancer cell lines

We used fitness data from 1,070 cancer cell lines from DepMap (Meyers *et al.*, 2017). For each gene, we calculated the median fitness and standard deviation across all cell lines. P-values were calculated using Mann-Whitney U tests.

Sequence analysis

For all 1:1 ortholog pairs between *S. cerevisiae* and *S. pombe*, we calculated their protein sequence identity. Sequence length similarity was calculated as the length ratio between the shortest of the sequences with respect to the longest. Thus, values closer to 1 describe sequence pairs of similar length, whereas values closer to 0 correspond to sequences of very

different lengths. P-values were calculated using Mann-Whitney U tests. We followed the same approach to compare *S. cerevisiae* and *C. albicans* sequences.

Suppression network analyses

Interaction data

We combined suppression interactions from our recent study (van Leeuwen *et al.*, 2020) with interactions found in the literature (van Leeuwen *et al.*, 2016) including only deletions of essential genes suppressed in standard conditions. We generated 1,000 randomized networks respecting the topology (i.e. maintaining the total number of connections of each gene) using the BiRewire R package (Iorio *et al.*, 2016). We calculated the number of bypass suppression pairs present in both datasets and compared that value to the number of overlapping pairs in randomized networks to derive an empirical p-value.

Functional overlaps

We calculated the fraction of bypass suppression gene pairs that coded for members of the same complex (Meldal *et al.*, 2021), belonged to the same molecular pathway (Kanehisa *et al.*, 2016), had MEFIT (Huttenhower *et al.*, 2006) coexpression scores above 1.0, localized in the same subcellular compartment (Huh *et al.*, 2003), and were annotated to the same biological process GO term (Myers *et al.*, 2006; Costanzo *et al.*, 2016). We repeated this calculation with the non-interacting gene pairs in the bypass suppression network, and derived fold enrichments and p-values using Fisher's exact tests. We applied this approach to the individual and the combined datasets, and to the LOF and GOF suppression pairs of the combined dataset.

Complex monochromaticity by suppression mode

We selected a non-redundant set of 17 complexes with at least two dispensable essential subunits in the bypass suppression network. We only kept one representative complex when several complexes had the same set of dispensable essential genes. For each complex, we calculated if all dispensable essential subunits could be suppressed by the same suppressor mode (LOF or GOF). Note that in one complex, all subunits could be suppressed by LOF suppressors but also by GOF suppressors (indicated by "LOF & GOF" in the panel). We counted all complexes with this monochromaticity in suppression mode and compared that value to the number of monochromatic complexes in a set of 1,000 randomized bypass suppression networks to derive an empirical p-value. We applied the same approach to the two individual suppression networks to discard a bias in the literature dataset.

Network modularity based on cocomplex relationships

We counted the number of dispensable essential genes within the same protein complex (Meldal *et al.*, 2021) that shared at least a suppressor. We repeated the same calculation using pairs of dispensable essential genes belonging to different complexes to derive a fold enrichment and a p-value calculated with a Fisher's exact test. We followed the same approach querying for interactors of bypass suppressors instead of the interactors of dispensable essential genes.

16

Functional preferences

We annotated the dispensable essential and suppressor genes in the network using 14 broad functional classes (Costanzo *et al.*, 2016). We then calculated the number of bypass suppression gene pairs within each pair of classes and repeated the process in randomized bypass suppression networks to derive empirical p-values. We used the median values of the randomized set to calculate the fold enrichments. Only fold enrichments of significant associations are shown.

We calculated fold enrichments for dispensable essential genes as the fraction of those genes in each class with respect to the corresponding fraction of core essential genes, from which we calculated a p-value using Fisher's exact test. We followed the same approach for the suppressors.

Agreement in copy number changes and suppression mode across *S. cerevisiae* strains

We defined copy number loss (CNL) and copy number gain (CNG) events as having a copy number below 1 or above 1, respectively. For each bypass suppression gene pair, we calculated the number of strains in which both genes had a CNL (i.e. co-loss events) and a copy number loss event for the dispensable essential gene and copy number gain for the suppressor gene (i.e. loss-gain events). We disregarded 17 hypermutated strains with copy-number changes in >33% of the genes, and aggregated co-loss and loss-gain events for all bypass suppression gene pairs after splitting pairs by their suppression mode. We compared the proportion of co-loss events to the proportion of loss-gain events overlapping with LOF bypass suppression pairs, and calculated the statistical significance by a two-sided Fisher's exact test. Finally, we counted the number of strains in which the proportion of co-loss events overlapping with LOF bypass suppression pairs was higher than that of loss-gain events, and viceversa, calculating the statistical significance by a two-sided binomial test. We repeated the same process to calculate the overlap with GOF bypass suppression pairs.

Co-mutation in *S. cerevisaie* strains

For each dispensable essential gene, we evaluated the strains in which it had a deleterious mutation, including missense mutations predicted as deleterious by SIFT, and indel and frameshift mutations. Next, we checked if any of its bypass suppressor genes was also mutated in any of those strains. We counted the number of dispensable essential genes co-mutated in any strains with any of their suppressor genes, the number of dispensable genes mutated alone, and the number of dispensable genes not mutated in any strain. We repeated the same process using 1,000 randomized bypass suppression networks. We performed this calculation using: 1) LOF bypass suppression pairs and haploid strains; 2) LOF bypass suppression pairs and diploid strains; and 3) GOF bypass suppression pairs and haploid strains.

Phenotypic changes across species and presence of bypass suppressor mutations

We annotated the orthology relationship of each dispensable essential gene in *S. pombe*. We only considered dispensable essential genes absent in *S. pombe* or with a 1:1 ortholog. For genes with 1:1 orthologs, we annotated the essentiality of the ortholog in that species. We also annotated the orthology relationships of bypass suppressor genes in *S. pombe*. For

suppressors with 1:1 orthologs, we performed a sequence alignment between the protein sequences of both species.

We next describe the set of rules that we evaluated to identify cases with equivalent bypass mutations in *S. pombe*. Briefly, in LOF suppressors we seeked for orthologs with decreased activity with respect to the *S. cerevisiae* gene, whereas in GOF suppressors, orthologs with increased activity. The first set of rules was based on orthology relationships. We considered *S. pombe* to have a LOF bypass mutation if the suppressor gene had an absent or N:1 ortholog, which could be similar to a copy number decrease and, thus, a decrease in activity. Suppressors with more than one ortholog in *S. pombe* or with a 1:1 ortholog were considered non-equivalent LOF bypass mutations, since their copy number did not decrease. Conversely, we considered as GOF bypass mutations cases in which the suppressor gene had more than one ortholog, similar to increasing their copy number and their activity, and non-equivalent GOF bypass mutations cases in which there was a N:1, 1:1, or absent ortholog in *S. pombe*.

The second set of rules we used to evaluate equivalent mutations was based on protein sequences. We only considered frameshift, nonsense, and missense mutations of suppressor genes with 1:1 orthologs. For the rest of cases, only the orthology rule was applied. The position of the nonsense and frameshift suppressor mutations identifies the part of the protein that should remain functional. Functionality encoded beyond that residue is compromised. Thus, we considered 1:1 orthologs in *S. pombe* with a shorter sequence than the position of the nonsense or frameshift suppressor mutation as LOF bypass mutations. Conversely, we considered cases in which the ortholog sequence was equal or longer than the position of the nonsense or frameshift suppressor mutation as non-equivalent LOF bypass mutations. In cases with missense mutations, we performed a sequence alignment between the *S. cerevisiae* suppressor gene and its 1:1 ortholog in *S. pombe*. We considered the ortholog to have an equivalent LOF bypass mutation if the same mutated residue or a gap was found in the aligned mutated position of the ortholog sequence. If the aligned mutated residue was the same as in the wildtype *S. cerevisiae* sequence (i.e. unmutated), we considered the ortholog to have a non-equivalent LOF bypass mutation. Cases in which the aligned position had different residues in *S. pombe* (not the wildtype and not the suppressor mutation) could not be classified as either equivalent or non-equivalent LOF bypass mutations. For GOF suppressors with a missense mutation and a 1:1 ortholog in *S. pombe*, we also performed a sequence alignment between the suppressor gene and its 1:1 ortholog. We considered the ortholog to have an equivalent GOF bypass mutation if the same mutated residue was found in the aligned mutated position of the ortholog sequence. If the aligned mutated residue was the same as in the wildtype *S. cerevisiae* sequence (i.e. unmutated), we considered the ortholog to have a non-equivalent GOF bypass mutation. The rest of cases could not be classified as either equivalent or non-equivalent GOF bypass mutations. Importantly, in suppressor genes with a frameshift, nonsense, or missense mutation, and with a 1:1 ortholog in *S. pombe*, the sequence based assessment took precedence over the orthology based evaluation.

Finally, we considered a dispensable gene to have an equivalent bypass suppressor in *S. pombe* if any of its suppressors satisfied that criteria. We grouped dispensable essential genes

by their essentiality in *S. pombe*, expecting dispensable essential genes with equivalent phenotypes in *S. pombe* (i.e absent or 1:1 non-essential orthologs) to have equivalent bypass suppressors more often than dispensable essential genes with a 1:1 essential ortholog. We calculated the fraction of genes with equivalent bypass suppressors for both gene sets to derive a fold enrichment and the p-value with a one-sided Fisher's exact test. We compared the fold enrichment of the bypass suppression network to a set of randomized bypass suppression networks, which we used to derive an empirical p-value.

We repeated the exact same process: i) using *C. albicans* sequences, orthology relationships, and essentiality annotations; ii) using orthoMCL (Li, Stoeckert and Roos, 2003), SonicParanoid (Cosentino and Iwasaki, 2019), and Pombase (Wood *et al.*, 2012) as alternative orthology mappings; iii) considering only dispensable essential genes with a single bypass suppressor to control for the bias introduced by gene degree; iv) removing bypass suppression pairs from the literature which may have been potentially identified by phylogenetic approaches; v) removing cocomplex and copathway bypass suppression pairs which may be more prone to present similar phylogenetic patterns; vi) switching LOF and GOF annotations to demonstrate the specificity of our sets of rules; vii) removing every node in the network to discard dependence on a single gene.

FIGURE LEGENDS

Figure 1.- Properties of dispensable essential gene sets
A: Number of dispensable essential genes per individual dataset and their overlap.
B: Fraction of dispensable and core essential genes in the combined dataset.
C: Fold enrichment of each dataset of dispensable essential genes for non-essential genes in the Sigma1278b *S. cerevisiae* strain.
D: Enrichment of dispensable essential vs core essential genes (left columns) and vs non-essential genes (right columns) for a panel of gene features. Left and right panels include binary and numeric features, respectively. Dot size is proportional to the fold enrichment and only enrichments with p-value < 0.05 are shown. The red arrows indicate properties of dispensable essential genes not previously identified.

Figure S1
A: Overlap between dispensable essential genes and missing genes in the core pangenome, defined at different thresholds. Fold enrichment with respect to the core essential genes is shown.
B: Functional enrichments of each dataset of dispensable essential genes.
C: The 14 complexes for which all essential genes are dispensable in the combined dataset. The subunits are colored by their dispensability in each individual dataset. Note that for the literature dataset we consider all essential genes as tested.
D: (left) Number of dispensable essential (i.e. all essential subunits are dispensable) and (right) core essential complexes (i.e. all essential subunits are core essential genes) for each dispensable gene set. The number of complexes in randomly selected gene sets of the same size are indicated in grey.
E: Enrichment of non-essential genes vs essential genes for a panel of gene features. Left and right panels include binary and numeric features, respectively. Dot size is proportional to the fold enrichment and only enrichments with p-value < 0.05 are shown.

Figure 2.- Phylogenetic analysis of dispensable essential genes
A: Orthology relationships in *S. pombe* of dispensable and core essential genes. The fraction of absent, duplicated, N:1, and essential and non-essential 1:1 orthologs is shown for each gene set.
B: Fold enrichment of dispensable essential genes with respect to core essential genes for absence, duplication, N:1 relationships, and essential 1:1 orthologs in *S. pombe*, *C. albicans*, and human. Purple and orange bars identify significant enrichments for dispensable essential and core essential genes, respectively. Grey bars identify non-significant enrichments.
C: Fraction of genes within each age group, ranging from zero (found only in *S. cerevisiae*) to five (found in the farthest ancestor), for the three sets of genes.
D: Fraction of genes with age zero (*S. cerevisiae* specific) for each gene set.
E: Fraction of gene loss events per gene across species grouped by gene set.
F: Median fitness per gene knockout across a panel of cancer cell lines. Genes are grouped by their essentiality in *S. cerevisiae*, and the density is shown

G: Protein sequence identity between gene products in *S. cerevisiae* and 1:1 orthologs in *S. pombe*.

Figure S2

A: Fraction of loss-of-function mutations across *S. cerevisiae* strains for non-essential, dispensable essential, and core essential genes.

B: Evolutionary rate (dN/dS) across *S. cerevisiae* strains for the three gene sets.

C: Fraction of copy number loss events across *S. cerevisiae* strains for the three gene sets.

D: Proportion of copy number loss events for non-essential, dispensable essential, and core essential genes among *S. cerevisiae* strains, grouped in sets of strains of the same size by their number of copy number loss events.

E: Orthology relationships in *C. albicans* of dispensable essential and core essential genes. The fraction of absent, duplicated, N:1, and essential and non-essential 1:1 orthologs is shown for each gene set.

F: Orthology relationships in human of dispensable essential and core essential genes. The fraction of absent, duplicated, N:1, and essential and non-essential 1:1 orthologs is shown for each gene set.

G: Fraction of essential genes in *S. cerevisiae* that are non-essential in *S. uvarum*.

H: Orthology relationships in *S. pombe* of non-essential and essential genes. The fraction of absent, duplicated, N:1, and essential and non-essential 1:1 orthologs is shown for each gene set.

I: Fold enrichment of dispensable genes with respect to non-essential genes for absence, duplication, N:1 relationships, and essential 1:1 orthologs in *S. pombe*, *C. albicans*, and human. Purple and green bars identify significant enrichments for dispensable and non-essential genes, respectively. Grey bars identify non-significant enrichments.

J: Fold enrichment of non-essential genes with respect to essential genes for absence, duplication, N:1 relationships, and essential 1:1 orthologs in *S. pombe*, *C. albicans*, and human. Green and orange bars identify significant enrichments for non-essential and essential genes, respectively. Grey bars identify non-significant enrichments.

K: Ratio between the protein sequence length in *S. cerevisiae* and the 1:1 ortholog in *S. pombe*. The shorter length is divided by the longer one.

L: Protein sequence identity between gene products in *S. cerevisiae* and 1:1 orthologs in *C. albicans*.

M: Ratio between the protein sequence length in *S. cerevisiae* and the 1:1 ortholog in *C. albicans*. The shorter length is divided by the longer one.

Figure 3.- Suppression interaction network

A: Number of bypass suppression gene pairs in each individual dataset and their overlap.

B: Fraction of loss-of-function (LOF) and gain-of-function (GOF) bypass suppression pairs that overlap with negative and positive genetic interactions.

C: (left) Fraction of monochromatic complexes in which all dispensable essential genes are suppressed by either loss-of-function (LOF) or gain-of-function (GOF) bypass suppressors. Only complexes with 2 or more dispensable essential subunits with known bypass suppression mode

are shown. (right) Number of monochromatic complexes in the suppression bypass network (blue) and in 1,000 randomized networks (grey).

D: Fraction of gene pairs in the same complex and in different complexes that share an interactor. Dispensable essential gene pairs are shown on top, bypass suppressor gene pairs at the bottom.

E: Bypass suppression network for the negative cofactor 2 complex (CPX-1662).

Figure S3

A: Functional enrichment for cocomplex, copathway, coexpression, colocalization, and connotation functional relationships of bypass suppression gene pairs in the combined and the two individual bypass suppression datasets.

B: (central) Functional enrichment of interacting pairs in the bypass suppression network for 14 broad functional categories. Only significant enrichments are shown. Functional bias among dispensable essential (right) and suppressor genes (bottom) is also shown, in which orange and grey bars identify significant and non-significant associations, respectively.

C: Fraction of dispensable essential genes with only loss-of-function suppressors (LOF), only gain-of-function suppressors (GOF), both loss-of-function and gain-of-function suppressors (LOF & GOF), and with suppressors of unknown type.

D: Fraction of monochromatic complexes supported by each bypass suppression dataset.

E: (left) Number of suppressors per dispensable essential gene. (right) Number of dispensable essential genes per suppressor

F: Enrichment of dispensable essential genes with multiple suppressors vs dispensable essential genes with single suppressors for a panel of gene features. Dot size is proportional to the fold enrichment and only enrichments with p-value < 0.05 are shown.

G: Interaction modularity of the bypass suppressor genes coding for members of the RPD3L histone deacetylase complex (CPX-1852).

H: Genetic interaction profiles of the bypass suppressor genes in (G): (left) network showing genetic interaction profile similarities above 0.2; (right) hierarchical clustering of the genetic interaction profiles.

Figure 4.- Co-occurring mutations in *S. cerevisiae* strains

A: Proportion of copy number co-loss and loss-gain events across a panel of *S. cerevisiae* strains for LOF and GOF bypass suppression pairs.

B: Fraction of dispensable essential genes with no deleterious mutation across haploid *S. cerevisiae* strains, with a mutation in any of the strains but not simultaneously with any of its bypass suppressor genes, and with at least one strain in which it has a deleterious mutation together with one of its bypass suppressors. Only LOF bypass suppression gene pairs are considered.

C: Number of dispensable essential genes with a deleterious mutation in any of the haploid *S. cerevisiae* strains simultaneously with at least one of its bypass suppressors using the bypass suppression network (pink) and a set of 1,000 randomized networks. Only LOF bypass suppression gene pairs are considered.

Figure S4

22

A: Number of *S. cerevisiae* strains in which bypass suppression gene pairs overlap more with copy number co-loss events than copy number loss-gain events (purple), and vice versa (green). Gene pairs are grouped by the bypass suppressor type: loss-of-function (LOF) and gain-of-function (GOF).

B: Number of dispensable essential genes with a deleterious mutation in any of the diploid *S. cerevisiae* strains simultaneously with at least one of its bypass suppressors using the bypass suppression network (pink) and a set of 1,000 randomized networks. Only LOF bypass suppression gene pairs are considered.

C: Number of dispensable essential genes with a deleterious mutation in any of the haploid *S. cerevisiae* strains simultaneously with at least one of its bypass suppressors using the bypass suppression network (pink) and a set of 1,000 randomized networks. Only GOF bypass suppression gene pairs are considered.

Figure 5.- Changes in essentiality co-occur with bypass suppressor mutations
A: Dispensable essential genes with an absent or 1:1 ortholog in *S. pombe*, and their loss-of-function and gain-of-function suppressors. Color code reflects whether dispensable essential and suppressor genes have equivalent phenotypes and mutations, respectively, in *S. pombe*. The circle indicates, for each dispensable essential gene, whether any of the bypass suppressor genes is mutated in *S. pombe*.

B: Fraction of dispensable essential genes with at least one bypass suppressor mutation in *S. pombe*. Dispensable essential genes are grouped by the phenotype of their 1:1 ortholog in *S. pombe* (E: essential; NE: non-essential; absent: without an ortholog).

C: Bypass suppressor mutations in *S. pombe* for dispensable essential genes with absent or 1:1 non-essential orthologs in *S. pombe* vs dispensable essential genes with 1:1 essential orthologs. Fold enrichments for the suppression interaction network (blue) and 1,000 randomizations (grey) are shown.

D: Like (B) but using *C. albicans* orthology relationships and phenotype data.

E: Like (C) but using *C. albicans* orthology relationships and phenotype data.

Figure S5
A: Like (5B) but considering only dispensable essential genes with a single bypass suppressor.

B: Like (5B) but removing bypass suppression pairs from the literature.

C: Like (5B) but removing bypass suppression pairs belonging to the same protein complex or pathway.

D: Bypass suppression subnetworks grouped by the enrichment p-value of the co-occurrence of dispensable essential genes with absent or 1:1 non-essential orthologs, and bypass suppressor mutations in *S. pombe*. Each bypass suppression subnetwork has a different gene removed.

E: Like (5B, 5C) but using different orthology mappings for *S. pombe*: orthoMCL (top), SonicParanoid (middle), and Pombase (bottom).

F: Like (5D, 5E) but using different orthology mappings for *C. albicans*: orthoMCL (top), SonicParanoid (bottom).

G: Like (5B) but with loss-of-function and gain-of-function switched annotations.

REFERENCES

Behan, F.M. *et al.* (2019) 'Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens', *Nature*, 568(7753), pp. 511–516.

Bergmiller, T., Ackermann, M. and Silander, O.K. (2012) 'Patterns of evolutionary conservation of essential genes correlate with their compensability', *PLoS genetics*, 8(6), p. e1002803.

Blomen, V.A. *et al.* (2015) 'Gene essentiality and synthetic lethality in haploid human cells', *Science*, 350(6264), pp. 1092–1096.

Chen, R. *et al.* (2016) 'Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases', *Nature biotechnology*, 34(5), pp. 531–538.

Cosentino, S. and Iwasaki, W. (2019) 'SonicParanoid: fast, accurate and easy orthology inference', *Bioinformatics* , 35(1), pp. 149–151.

Costanzo, M. *et al.* (2010) 'The genetic landscape of a cell', *Science*, 327(5964), pp. 425–431.

Costanzo, M. *et al.* (2016) 'A global genetic interaction network maps a wiring diagram of cellular function', *Science*, 353(6306). Available at: https://doi.org/10.1126/science.aaf1420.

Deng, J. *et al.* (2011) 'Investigating the predictability of essential genes across distantly related organisms using an integrative approach', *Nucleic acids research*, 39(3), pp. 795–807.

Dezso, Z., Oltvai, Z.N. and Barabási, A.-L. (2003) 'Bioinformatics analysis of experimentally determined protein complexes in the yeast Saccharomyces cerevisiae', *Genome research*, 13(11), pp. 2450–2454.

Dowell, R.D. *et al.* (2010) 'Genotype to phenotype: a complex problem', *Science*, 328(5977), p. 469.

Fumasoni, M. and Murray, A.W. (2020) 'The evolutionary plasticity of chromosome metabolism allows adaptation to constitutive DNA replication stress', *eLife*, 9. Available at: https://doi.org/10.7554/eLife.51963.

Giaever, G. *et al.* (2002) 'Functional profiling of the Saccharomyces cerevisiae genome', *Nature*, 418(6896), pp. 387–391.

Glass, J.I. *et al.* (2006) 'Essential genes of a minimal bacterium', *Proceedings of the National Academy of Sciences of the United States of America*, 103(2), pp. 425–430.

Good, B.H. *et al.* (2017) 'The dynamics of molecular evolution over 60,000 generations', *Nature*, 551(7678), pp. 45–50.

Harris, M.A. *et al.* (2022) 'Fission stories: using PomBase to understand Schizosaccharomyces pombe biology', *Genetics*, 220(4). Available at: https://doi.org/10.1093/genetics/iyab222.

Hart, G.T., Lee, I. and Marcotte, E.R. (2007) 'A high-accuracy consensus map of yeast protein complexes reveals modular nature of gene essentiality', *BMC bioinformatics*, 8, p. 236.

Hart, T. *et al.* (2015) 'High-Resolution CRISPR Screens Reveal Fitness Genes and

Genotype-Specific Cancer Liabilities', *Cell*, 163(6), pp. 1515–1526.

Huh, W.-K. *et al.* (2003) 'Global analysis of protein localization in budding yeast', *Nature*, 425(6959), pp. 686–691.

Hutchison, C.A., 3rd *et al.* (2016) 'Design and synthesis of a minimal bacterial genome', *Science*, 351(6280), p. aad6253.

Huttenhower, C. *et al.* (2006) 'A scalable method for integration and functional analysis of multiple microarray datasets', *Bioinformatics* , 22(23), pp. 2890–2897.

Hwang, Y.-C. *et al.* (2009) 'Predicting essential genes based on network and sequence analysis', *Molecular bioSystems*, 5(12), pp. 1672–1678.

Iorio, F. *et al.* (2016) 'Efficient randomization of biological networks while preserving functional characterization of individual nodes', *BMC bioinformatics*, 17(1), p. 542.

Jordan, D.M. *et al.* (2015) 'Identification of cis-suppression of human disease mutations by comparative genomics', *Nature*, 524(7564), pp. 225–229.

Jordan, I.K. *et al.* (2002) 'Essential Genes Are More Evolutionarily Conserved Than Are Nonessential Genes in Bacteria', *Genome Research*, pp. 962–968. Available at: https://doi.org/10.1101/gr.87702.

Juhas, M., Eberl, L. and Glass, J.I. (2011) 'Essence of life: essential genes of minimal genomes', *Trends in cell biology*, 21(10), pp. 562–568.

Kanehisa, M. *et al.* (2016) 'KEGG as a reference resource for gene and protein annotation', *Nucleic acids research*, 44(D1), pp. D457–62.

Kim, D.-U. *et al.* (2010) 'Analysis of a genome-wide set of gene deletions in the fission yeast Schizosaccharomyces pombe', *Nature biotechnology*, 28(6), pp. 617–623.

Klim, J. *et al.* (2021) 'Genetic interaction network has a very limited impact on the evolutionary trajectories in continuous culture-grown populations of yeast', *BMC ecology and evolution*, 21(1), p. 99.

LaBar, T. *et al.* (2020) 'Evolutionary Repair Experiments as a Window to the Molecular Diversity of Life', *Current biology: CB*, 30(10), pp. R565–R574.

van Leeuwen, J. *et al.* (2016) 'Exploring genetic suppression interactions on a global scale', *Science*, 354(6312). Available at: https://doi.org/10.1126/science.aag0839.

van Leeuwen, J. *et al.* (2020) 'Systematic analysis of bypass suppression of essential genes', *Molecular systems biology*, 16(9), p. e9828.

Li, G., Ji, B. and Nielsen, J. (2019) 'The pan-genome of Saccharomyces cerevisiae', *FEMS yeast research*, 19(7). Available at: https://doi.org/10.1093/femsyr/foz064.

Li, J. *et al.* (2019) 'Systematic analysis reveals the prevalence and principles of bypassable gene essentiality', *Nature communications*, 10(1), p. 1002.

Li, L., Stoeckert, C.J., Jr and Roos, D.S. (2003) 'OrthoMCL: identification of ortholog groups for

eukaryotic genomes', *Genome research*, 13(9), pp. 2178–2189.

Liu, G. *et al.* (2015) 'Gene Essentiality Is a Quantitative Property Linked to Cellular Evolvability', *Cell*, 163(6), pp. 1388–1399.

Lloyd, J.P. *et al.* (2015) 'Characteristics of Plant Essential Genes Allow for within- and between-Species Prediction of Lethal Mutant Phenotypes', *The Plant cell*, 27(8), pp. 2133–2147.

Luo, H., Gao, F. and Lin, Y. (2015) 'Evolutionary conservation analysis between the essential and nonessential genes in bacterial genomes', *Scientific reports*, 5, p. 13210.

Mani, R. *et al.* (2008) 'Defining genetic interaction', *Proceedings of the National Academy of Sciences of the United States of America*, 105(9), pp. 3461–3466.

Meldal, B.H.M. *et al.* (2021) 'Analysing the yeast complexome-the Complex Portal rising to the challenge', *Nucleic acids research*, 49(6), pp. 3156–3167.

Meyers, R.M. *et al.* (2017) 'Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells', *Nature genetics*, 49(12), pp. 1779–1784.

Mi, H. *et al.* (2021) 'PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API', *Nucleic acids research*, 49(D1), pp. D394–D403.

Myers, C.L. *et al.* (2006) 'Finding function: evaluation methods for functional genomic data', *BMC genomics*, 7, p. 187.

Narasimhan, K. *et al.* (2015) 'Mapping and analysis of Caenorhabditis elegans transcription factor sequence specificities', *eLife*, 4. Available at: https://doi.org/10.7554/eLife.06967.

Ng, P.C. and Henikoff, S. (2001) 'Predicting deleterious amino acid substitutions', *Genome research*, 11(5), pp. 863–874.

Peter, J. *et al.* (2018) 'Genome evolution across 1,011 Saccharomyces cerevisiae isolates', *Nature*, 556(7701), pp. 339–344.

Rancati, G. *et al.* (2018) 'Emerging and evolving concepts in gene essentiality', *Nature reviews. Genetics*, 19(1), pp. 34–49.

Roemer, T. *et al.* (2003) 'Large-scale essential gene identification in Candida albicans and applications to antifungal drug discovery', *Molecular microbiology*, 50(1), pp. 167–181.

Rojas Echenique, J.I. *et al.* (2019) 'Modular epistasis and the compensatory evolution of gene deletion mutants', *PLoS genetics*, 15(2), p. e1007958.

Sanchez, M.R. *et al.* (2019) 'Transposon insertional mutagenesis in reveals -acting effects influencing species-dependent essential genes', *Genome research*, 29(3), pp. 396–406.

Sardiu, M.E. *et al.* (2009) 'Determining protein complex connectivity using a probabilistic deletion network derived from quantitative proteomics', *PloS one*, 4(10), p. e7310.

Segal, E.S. *et al.* (2018) 'Gene Essentiality Analyzed by Transposon Mutagenesis and Machine

Learning in a Stable Haploid Isolate of', *mBio*, 9(5). Available at: https://doi.org/10.1128/mBio.02048-18.

Sonnhammer, E.L.L. and Östlund, G. (2015) 'InParanoid 8: orthology analysis between 273 proteomes, mostly eukaryotic', *Nucleic acids research*, 43(Database issue), pp. D234–9.

Szamecz, B. *et al.* (2014) 'The genomic landscape of compensatory evolution', *PLoS biology*, 12(8), p. e1001935.

Takeda, A. *et al.* (2019) 'Identification of 15 New Bypassable Essential Genes of Fission Yeast', *Cell structure and function*, 44(2), pp. 113–119.

Tenaillon, O. *et al.* (2012) 'The molecular diversity of adaptive convergence', *Science*, 335(6067), pp. 457–461.

UniProt Consortium (2021) 'UniProt: the universal protein knowledgebase in 2021', *Nucleic acids research*, 49(D1), pp. D480–D489.
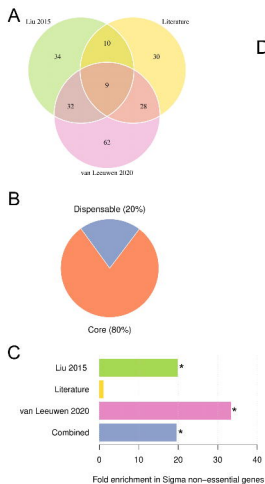
Vignogna, R.C., Buskirk, S.W. and Lang, G.I. (2021) 'Exploring a Local Genetic Interaction Network Using Evolutionary Replay Experiments', *Molecular biology and evolution*, 38(8), pp. 3144–3152.

Wang, T. *et al.* (2015) 'Identification and characterization of essential genes in the human genome', *Science*, 350(6264), pp. 1096–1101.
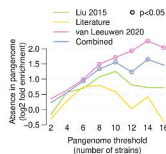
Woodford, N. and Ellington, M.J. (2007) 'The emergence of antibiotic resistance by mutation', *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 13(1), pp. 5–18.

Wood, V. *et al.* (2012) 'PomBase: a comprehensive online resource for fission yeast', *Nucleic acids research*, 40(Database issue), pp. D695–9.
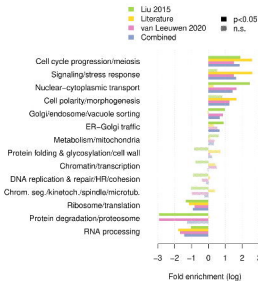
Zhang, X., Acencio, M.L. and Lemke, N. (2016) 'Predicting Essential Genes and Proteins Based on Machine Learning and Network Topological Features: A Comprehensive Review', *Frontiers in physiology*, 7, p. 75.
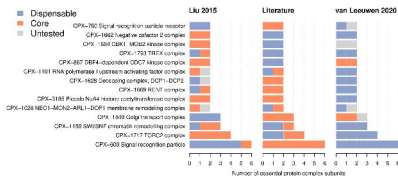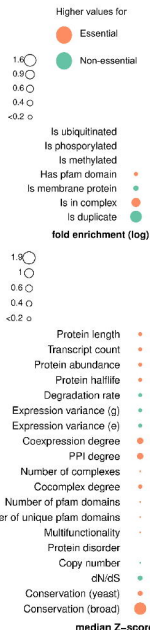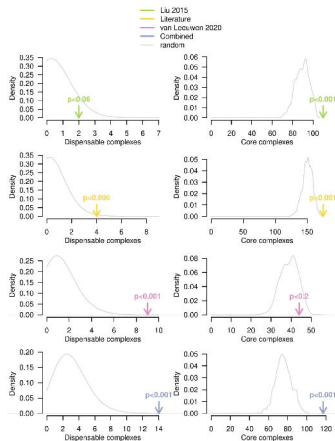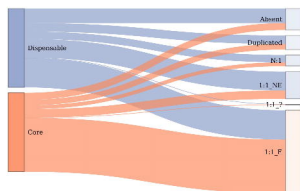
F1

A

Liu 2015 — 34, 10, Literature — 30
9
32, 28
62
van Leeuwen 2020

B

Dispensable (20%)
Core (80%)

C

Liu 2015 *
Literature
van Leeuwen 2020 *
Combined *

Fold enrichment in Sigma non-essential genes

D

Higher values for
● Core
● Dispensable
● NE

→ Novel properties

fold enrichment (log)

	Disp vs core	Disp vs NE
Is ubiquitinated		
Is phosporylated		
Is methylated		
→ Has pfam domain	●	●
Is membrane protein	●	●
Is in complex	●	●
Is duplicate	●	●

1.4 ○
0.8 ○
0.4 ○
0.2 ○
<0.1 ○

median z-score

	Disp vs core	Disp vs NE
Protein length		●
→ Transcript count	●	●
Protein abundance	●	●
Protein halflife		
Degradation rate		●
→ Expression variance (g)	●	●
→ Expression variance (e)	●	●
Coexpression degree	●	●
PPI degree	●	●
Number of complexes		
→ Cocomplex degree	●	●
Number of pfam domains		●
Number of unique pfam domains		●
→ Multifunctionality	●	●
Protein disorder		
Copy number	●	●
Conservation (yeast)	●	●
Conservation (broad)	●	●

1 ○
0.6 ○
0.4 ○
0.2 ○
<0.1 ○

F2

F3

A Literature / MSB_20 Venn diagram: 172, 31, 116

B Overlap w/ GI bar chart (LOF, GOF)

C Fraction of complexes stacked bar (Suppressor monochromaticity: None, LOF, LOF & GOF, GOF); Number of randomized networks histogram (real=13, p<0.028), Number of complexes with monochromatic suppressors

D Fraction with shared interactors (Same complex, Different complexes); Dispensable essential pairs, Suppressor pairs

E Suppressor: NCB2 & BUR6, BUR6, NCB2; Dispensable essential gene; CPX-3226, CPX-1662 (NCB2, BUR6), CPX-2662, CPX-1633, CPX-1659

S3

A

B

Suppressor gene

C

Suppressor type

D

E

F

Higher values for

● Single suppressors

● Multiple suppressors

G

H

Clustering of GI profile similarities

F4

A

CNL-CNG
CNL-CNL

Fraction of the overlapping co-CNC events

LOF   GOF

B

Fraction of dispensable genes

Dispensable gene
Not mutated
Mutated alone
Co-mutated with suppressor

C

Number of randomized networks

real=18
(p<0.014)

Number of dispensable genes with
simultaneous LOF mutation in suppressor

S4

F5

S5

A

Dispensable gene in Spombe: E *, NE, Absent
Fraction with equivalent suppressors

D

Dispensable–suppressor subnetworks

p=0.001 (68.3%)
p=0.002 (31.0%)
p=0.003 (0.7%)

F

Dispensable gene in Calbicans: E *, NE, Absent
Fraction with equivalent suppressors

B

Dispensable gene in Spombe: E *, NE, Absent
Fraction with equivalent suppressors

E

Dispensable gene in Spombe: E *, NE, Absent
Fraction with equivalent suppressors

Dispensable gene in Calbicans: E *, NE, Absent
Fraction with equivalent suppressors

C

Dispensable gene in Spombe: E *, NE, Absent
Fraction with equivalent suppressors

Dispensable gene in Spombe: E *, NE, Absent
Fraction with equivalent suppressors

G

Real *, Switched
Fraction with equivalent suppressor
E
Abs/NE

Dispensable gene in Spombe: E *, NE, Absent
Fraction with equivalent suppressors