

Investigation of RNA metabolism through large-scale genetic interaction profiling in yeast

Laurence Decourty^{1,2}, Christophe Malabat³, Emmanuel Frachon⁴, Alain Jacquier^{1,2} and Cosmin Saveanu^{1,2,*}

¹Unité de Génétique des Interactions Macromoléculaires, Département Génomes et Génétique, Institut Pasteur, 75015 Paris, France, ²UMR3525, Centre national de la recherche scientifique (CNRS), 75015 Paris, France, ³Hub Bioinformatique et Biostatistique, Département de Biologie Computationnelle, Institut Pasteur, 75015 Paris, France and ⁴Plate-forme Technologique Biomatiériaux et Microfluidique, Centre des ressources et recherches technologiques, Institut Pasteur, 75015 Paris, France

Received November 01, 2020; Revised July 19, 2021; Editorial Decision July 19, 2021; Accepted August 02, 2021

ABSTRACT

Gene deletion and gene expression alteration can lead to growth defects that are amplified or reduced when a second mutation is present in the same cells. We performed 154 genetic interaction mapping (GIM) screens with query mutants related with RNA metabolism and estimated the growth rates of about 700 000 double mutant *Saccharomyces cerevisiae* strains. The tested targets included the gene deletion collection and 900 strains in which essential genes were affected by mRNA destabilization (DAMP). To analyze the results, we developed RECAP, a strategy that validates genetic interaction profiles by comparison with gene co-citation frequency, and identified links between 1471 genes and 117 biological processes. In addition to these large-scale results, we validated both enhancement and suppression of slow growth measured for specific RNA-related pathways. Thus, negative genetic interactions identified a role for the OCA inositol polyphosphate hydrolase complex in mRNA translation initiation. By analysis of suppressors, we found that Puf4, a Pumilio family RNA binding protein, inhibits ribosomal protein Rpl9 function, by acting on a conserved UGUA-cauUA motif located downstream the stop codon of the *RPL9B* mRNA. Altogether, the results and their analysis should represent a useful resource for discovery of gene function in yeast.

INTRODUCTION

The process of assigning function to a gene involves switching it off, partially or totally, and evaluating a phenotype. A major limitation of this approach is that genes do not function in isolation and evolved from other genes, sometimes

following cataclysmic events, such as whole genome duplication, or more restricted chromosome segment duplication (1,2). As a consequence, removal or alteration of a gene from a duplicated pair might show no effect under standard culture conditions. The presence of duplicated genes can increase fitness, a phenomenon that was confirmed by testing single gene deletion mutants in yeast (3), by performing experimental evolution under random mutagenesis conditions (4) or by comparing the effect of duplicated gene pair deletion in comparison with singletons in *Saccharomyces cerevisiae* versus *Schizosaccharomyces pombe* (5). Gene duplication is just the simplest illustration of how cells can adapt to mutations. In many other cases, the flexibility and robustness of cellular pathways allows adaptation of cells to gene loss.

A way to identify and study gene redundancy and robustness against mutations is to combine perturbations for several genes in a single strain and look at the resulting phenotype. This strategy worked well for studies such as the identification of genes involved in the secretory pathway (6). It only became a systematic way to study gene function when methods to identify and quantify growth of combinations for thousands of mutants became available (7–10), as reviewed in (11). Simultaneous perturbation of two genes can result in various effects on growth. Sometimes, the combination is neutral, sometimes it leads to a strong growth inhibition (synthetic lethality) and sometimes one mutation can hide or overcome the effects of the other (12). Altogether, these effects are covered by the convenient umbrella term of ‘genetic interactions’ (GIs).

The behaviour of a gene variant over many screens establish a GI profile (8,10,13). A similarity of GI profiles can predict physical interactions of the corresponding proteins in complexes and subcomplexes. For example, the analysis of proteasome component mutants, allowed to correctly assign proteins to the corresponding proteasome subcomplexes (14). Large scale double mutant screens can also as-

*To whom correspondence should be addressed. Tel: +40 01 44 38 92 80; Email: cosmin.saveanu@pasteur.fr

sociate previously uncharacterised genes with specific pathways. For example, the RNA exosome co-factor Mpp6 was identified on the basis of the observed synthetic lethality between its gene deletion and the absence of the nuclear exosome component Rrp6 (15). Thus, description of GIs serves several goals. It can identify the potential function of genes and find combinations of mutants that uncover phenotypes otherwise hidden by gene redundancy. It can also help in understanding the evolutionary trajectory of duplicated genes towards redundancy or towards unrelated cellular processes (16, for an example). These goals require high quality and validated large scale results, based on independent studies performed under a variety of culture conditions.

Early systematic gene deletion combination screens were restricted to the study of non-essential genes. To investigate essential gene mutants, several strategies have been used, including mRNA destabilization, the study of mutations leading to thermosensitivity, CRISPR genome editing and transposon insertion analysis. The ‘decreased abundance by mRNA perturbation’, DAmP, strategy was the first to be used for systematic investigation of hypomorphic alleles of essential genes in yeast and is based on the addition of a long extension downstream the stop codon position of targeted genes. This extension leads to mRNA destabilization through nonsense-mediated mRNA decay, NMD (10). Three independent systematic yeast libraries were built using variations of this strategy, for large scale genetic or chemogenomic screens. One did not include molecular barcodes in the strains (14) and can not be directly used for growth estimation in pooled mutant assays. For such assays, a second collection was generated in which ‘molecular barcodes’, unique artificial short sequences flanked by universal sequences allowing their amplification, were included at a specific genomic locus for each strain (17). However, since the modified locus and the barcode are not physically linked, this second collection was not usable for genetic interaction mapping (GIM) screens, which depend on co-segregation of mutant and barcodes in a pooled population of mutants (8). To solve this problem, we generated a third DAmP collection, where barcodes are present at the modified locus. These strains can be used both for measuring cell numbers in GIM screens and for transcript quantitation, by reverse transcription of RNA and barcode amplification (18).

In addition to DAmP essential gene perturbations, recent methods that are able to generate collections of mutants analysed by DNA sequencing became available. For example, CRISPR interference was used to generate new collections of yeast mutants (19) and was adapted to the study of genetic interactions under several growth conditions (20). Alternatively saturated transposon insertion coupled with sequencing allows the exploration of a broad spectrum of mutations, including protein truncation or transcription deregulation, and can be used to characterize the function of essential genes (21). These new methods remain technically challenging and have not yet been used on a large scale. Thus, results about GIs from systematic large-scale studies using essential gene variants in yeast are, for the moment, restricted to thermosensitive (TS) and DAmP alleles under the specific conditions of the synthetic genetic array, SGA (7), screens (22).

The modest overlap between SGA results and those obtained for the same pairs of mutant genes by CRISPRiSeq (20) confirmed previous demonstrations that culture conditions might be crucial for the detection and measurement of GIs (23,24). In this respect, the GIM screens (8) performed under selection with antibiotics that affect mRNA translation, are particularly good at detecting GIs for factors involved in RNA metabolism. For example, the strong effect observed in GIM screens for double deletions involving components of the ribosome quality control complex and the SKI complex (25), was validated on individual strains only in the presence of low concentrations of hygromycin B, a translation inhibitor (26).

The specific conditions of GIM screens that had the potential to identify new GIs, and the availability of the barcoded DAmP collection, compatible with these screens (8,18), motivated us to generate a new set of large-scale GIs in yeast. We selected 154 genes, mostly related with RNA metabolism, and tested their GIs when combined with the 5500 non-essential gene deletions and 900 DAmP alleles for essential genes.

The size of the generated data set and the fact that RNA metabolism perturbation directly or indirectly affects most cellular processes ensured that our results cover a large variety of functions. A major challenge in analyzing the new large data set was extracting meaningful information allowing the discovery of links between genes and specific cellular processes. SAFE, a recently developed method that is specific to GI networks (27), uses the local neighbourhood in complex networks to identify enrichment for specific annotations. We present here a different approach, called RECAP for ‘Rational Extension of Correlated Annotations and GI Profiles’ that starts instead from links between genes inferred from co-occurrence in publications, based on the set of scientific articles curated by the *Saccharomyces Genome Database* (28). This approach uses well annotated gene groups in combination with GI profile similarity to find which mutants behave ‘as expected’ from previous studies. Only the validated mutants were then used to extend the network of related genes and predict the potential association of hundreds of genes with specific cellular processes or multiprotein complexes. RECAP is not limited to networks based on GI profiles and is potentially useful for any type of large-scale results that establish links between genes.

The presented results complement and extend previous studies and, together with the RECAP data analysis strategy, provide new predictions about the function of hundreds of yeast genes.

MATERIALS AND METHODS

GIM screens

GIM screens were performed as described originally (8), and following the protocol described in detail in (29), using custom-made turbidostat devices that allowed performing 16 cultures in parallel. Briefly, MAT α query strains were obtained by replacing the KanMX resistance cassette in strains from the gene deletion collection (30) with a Pr α -Nat cassette that expresses the nourseothricin resistance gene only in the context of a haploid MAT α strain. Hygromycin

B resistance was also added to the query strain using a centromeric plasmid to allow selection of diploid strains after mating. Pools of deletion (30) and DAmP strains (18) were recovered from stock maintained at -80°C and left to recover in rich medium for 30 minutes by incubation at 30°C , then mixed with fresh query strain culture for 5 hours of mating on a GNA plate (5% glucose, 3% Difco nutrient broth, 1% Difco yeast extract). Recovered diploids were incubated overnight at 30°C in 100 ml liquid GNA medium in the presence of 0.2 mg/ml hygromycin B and 0.2 mg/ml G418. Sporulation was induced by incubating about 10^9 diploid cells in flasks containing 150 ml potassium acetate medium (1% potassium acetate, 0.005% zinc acetate, supplemented with 2 mg uracil, 2 mg histidine and 6 mg leucine for 100 ml). After sporulation, cells were recovered in YPD medium (2% glucose, 1% Difco yeast extract, 1% Difco Bactopeptone), incubated for 6 hours without antibiotics and grown for 45–60 h in the presence of 0.2 mg/ml G418 and 20 $\mu\text{g}/\text{ml}$ nourseothricin. For each batch of 16 screens, the reference against which the screens were compared was a mix of cells from all the final cultures. For *Tet-off* query strains and screens, doxycyclin at 10 $\mu\text{g}/\text{ml}$ was included in the dual antibiotic haploid selection step for 16–24 h in liquid culture.

DNA was extracted from the final cell pellets (about 10^8 cells) and used to amplify upstream and downstream barcodes. Barcode DNA relative levels were measured using custom microarrays (Agilent Technologies, California, USA) designed accordingly to the updated barcode sequences (31) and the collected images were processed with GenePix Pro 6 (Molecular Devices, California, USA) and analysed using R (32). Data analysis consisted of normalization of the Cy3/Cy5 using the *loess* algorithm, aggregation of results for upstream and downstream barcodes and normalization of the aggregated results. Each screen result was examined for the presence of the expected signal around the query gene locus that corresponds to the decrease in recombination frequency during meiosis due to physical proximity on the same chromosome. For 42 query genes out of 196 tested, the screen results were eliminated from further analysis either because additional peaks were present, indicating secondary suppression mutations in the used strains (33), or because no exclusion peaks were detected. For the analysed screens, the exclusion peaks were corrected using estimates of recombination frequency based on the observed signal (8).

Finally, results from at least two independent screens were expressed as the \log_2 of the ratio between the screen of interest and the reference (Q/R) and combined to obtain GI estimates. Results were corrected for pleiotropic effects by counting the fraction of screens in which a given mutant showed a $\log_2(Q/R)$ value inferior to the arbitrary threshold of -1.25 , named pleiotropic index (PI). Each initial $\log_2(Q/R)$ value was multiplied with $(1 - \text{PI})^3$ to decrease the weight of mutants showing a response in most screens. Adjusted values of $\log_2(Q/R)$ were used to compute Pearson correlation coefficients for all the GI profiles pairs. To assess the reciprocity of the observed GI profiles, we ranked for each mutant the similarity of profiles for all the other mutants in decreasing order of the corresponding Pearson correlation coefficient. If the GI profile for mutant

A was among the top 20 profiles for mutant B and, conversely, the profile for mutant B was in the top 20 profiles for mutant A, we considered that A and B were linked.

RECAP data analysis

To annotate the observed GI profile links we used the curated database of yeast literature from the Saccharomyces Genome Database (28). The table associates genes with publications. Only articles dealing with <100 genes were selected, and only the half most cited yeast genes were used to build a network of co-citations. Links in this network were based on the presence of the two genes in the same publications. At least 20% of articles citing a gene had also to cite the other one to establish a connection. The obtained co-citation network for 1847 genes showed strong connections among 439 isolated groups of genes. Within each group, we analyzed the presence of genes linked by GI profile similarity and selected 117 cases in which at least one such connection was present.

Genes connected by GI profile similarity and by co-citation were considered valid in terms of GIM screens and served to extend the network using the current GIM data set of adjusted $\log_2(Q/R)$ and the computed links based on reciprocal GI profile similarity. The links based on GI profile similarity were then used to associate genes with biological processes or cellular components. To this end we used the list of genes of each group to interrogate the g:Profiler web server <https://biit.cs.ut.ee/gprofiler/gost> for gene ontology term enrichment, using the gprofiler2 R package (34) and selected the top entry for biological process and cellular component in each case. To test various configurations for RECAP, we used the R igraph (35) and RCy3 (36) packages, together with visualization and network analysis in Cytoscape (37).

Strains and plasmids

The generation and details of the DAmP strains collection were previously published (18). Briefly, DNA from diploid strains from the deletion collection (30) for the genes of interest was used to amplify the Kan^{R} cassette flanked by upstream and downstream barcode sequences. The cassette was amplified using oligonucleotides that targeted the insertion of the cassette in the genome of a BY4741 yeast strain downstream the stop codon of the same gene. Individual clones from each transformation were tested by PCR amplification with specific oligonucleotides.

For the situations when the DAmP strain was used to perform GIM screens, the Kan^{R} cassette was transferred by amplification and transformation in the BY4742 $\text{MAT}\alpha$ background. In these strains, the G418 resistance cassette was next replaced by the $\text{Pr}\alpha\text{-Nat}$ cassette. For the *Tet-off* strains, we used the pCM224 vector (38) to amplify the G418 resistance cassette and place the tetO2 sequence upstream the start position for the coding sequence of selected genes. Individual clones were tested by PCR on genomic DNA to test the correct insertion of the cassette. G418 resistance was next changed to $\text{MAT}\alpha$ -specific nourseothricin resistance using the pGID3 vector (8,29).

SnoRNA deletion strains for box H/ACA snoRNPs were derived from yeast strains available in our laboratory (39). For box C/D snoRNAs deletion strains, we used the KanMX6 cassette from the pFA6 vector (40) with the oligonucleotides listed in Supplementary Table S10. Double deletion mutant strains were built by mating of G418-resistant and nourseothricin resistant strains. After sporulation and selection of haploid clones, we ensured that the obtained strains had the same panel of auxotrophy markers as the parental strain BY4742. The strains for individual validation of screen results are listed in Supplementary Table S11.

Plasmids to test the effects of the RPL9B 3' UTR region on a reporter gene expression were built by ligase-less Gibson assembly (41) starting from pCM189 (42), digested with BamHI and NotI. HA-ALAI coding sequence was amplified with oligonucleotides CS1635 (TAC ACA CACT AAA TTA CCG GAT CAA TTC GGG GGA TCC AGC GAT GTC GTA TC) and CS1636 (CTA AGC GCA AAG ACA CTT TTT TGT TCC ATT TAC ATG TCC AAC TTA ATT AAA CCA ACA TTT TTC C) from a previously described NMD reporter (43). The last two amino acid codons, the stop codon and the following 103 nucleotides of the 3' UTR of RPL9B were amplified with oligonucleotides CS1631 (TGT AAA TGG AAC AAA AAA GTG TCT TTG CGC TTA GAG GCG CTT AAA TAT GTA CAT TAT TCC TAC, native Puf4 binding motif) or CS1632 (TGT AAA TGG AAC AAA AAA GTG TCT TTG CGC TTA GAG GCG CTT AAA TAG ATT CAT TAT TCC TAC, mutated Puf4 binding motif) and CS1634 (TGA TGC GGC CCT CCT GCA GGG CCC TAG CGG CCG CCC AAT GAC AAA GAA CCT CAG TCT TTA C) from a plasmid containing the 3' UTR region of RPL9B (44). A version with a longer region of RPL9B in the 3' UTR of the reporter mRNA was obtained by using oligonucleotides CS1631 or CS1632 and CS1633 (TGA TGC GGC CCT CCT GCA GGG CCC TAG CGG CCG CTC ATT TGT AAC GTA TTA GAG AAT ATT TAT TTT TTA G). This vector variant contained a 308 nucleotides region of the RPL9B 3' UTR region, which was shown to ensure auto-regulation to CUP1, when used as a reporter gene (44). The obtained plasmids were verified by sequencing and immunoblot against the N-terminal HA tag. Anti-HA antibodies, coupled with the horseradish peroxidase (Roche, rat monoclonal 3F10), used at a dilution of 1 in 1000, were revealed using a chemiluminescence substrate (Clarity, BioRad).

Beta-galactosidase activity

Beta-galactosidase activity was measured on total cell extracts obtained by lysis by vortexing with glass beads. The assay buffer contained 0.5 mM chlorophenol red- β -D-galactopyranoside (CPRG), 100 mM sodium phosphate, 10 mM KCl, 1 mM MgSO₄, 5 mM dithiothreitol (DTT), pH 7. After incubation at 37°C, the absorbance change at 574 nm was normalized to protein concentration measured using the Bradford assay. The reported values are relative to the beta-galactosidase activity of a wild-type strain processed in parallel.

RNA and RT-QPCR

Cells were grown in synthetic complete medium to log phase and collected. Total RNA was obtained using hot phenol extraction (45) and DNA was removed with DNase I (Ambion TURBO DNA-free kit) before reverse-transcription (RT) and PCR amplification. For each experiment for the evaluation of OCA complex effects, 500 ng of total RNA were used in a RT reaction with Superscript III (Invitrogen) using a mix of the following oligonucleotides: SNO1rv AAC TCC TGA GGA TCT AGC CCA GTG, ARG1rv ACC ATG AGA GAC CGC GAA ACA G and RIM1rv ACC CTT AGA ACC GTC GTC TCT C. Quantitative PCR reactions used the same oligonucleotides coupled with the following forward primers (one pair for each target): SNO1fw AAC TCC TGA GGA TCT AGC CCA GTG, ARG1fw GCA AGA CCT GTT ATT GCC AAA GCC and RIM1fw GCG CTT TGG TAT ATG TTG AAG CAG. For each experiment, the ARG1 and SNO1 signal was normalized to the RIM1 signal and all the results were compared with the wild type strain.

For the HA-ALAI-RPL9B reporters, the pairs of oligonucleotides used to estimate mRNA levels by Q-PCR were: CS1020fw (AGG CTT TGG AGA ACC CAA CCA G) and CS1021rv (TTG TCA GCA ACC TTG GCA CCA C) for *PGK1*; CS1127fw (ACA CAC ACT AAA TTA CCG GAT C) and CS1128rv (GCA TAA TCT GGA ACA TCA TAT G) for the HA part of the reporter; CS1038fw (AAG GTC GTG TTG CTC ACG GTT G) and CS1039rv (GGC AAG CGC AGA ACC ATC AAT AC) for endogenous *ALAI* mRNA. Normalization was done against *PGK1* mRNA.

RESULTS

Correcting for pleiotropic behavior improves the specificity of GIs

The choice of the 154 perturbed genes, called 'query genes', used to test their effect in combination with the collections of gene deletion and DAmP strains was guided by published GI results and gene annotations, with a focus on RNA-related processes, as summarized in Figure 1A and listed in Supplementary Table S1. In addition to factors affecting RNA transcription, export and degradation, a set of 25 mutants included genes that were either directly or indirectly related with proteasome function, or had shown genetic interactions with proteasome deficiency. To limit the bias induced by the choice of tested mutants, we included 14 metabolism genes, 21 genes affecting other processes and 10 unannotated or poorly annotated genes. Gene deletion strains from the collection of haploid strains (30) were modified to be suitable for GIM screens by changing the geneticin resistance cassette with the MAT α haploid specific nourseothricin resistance cassette (8,29). The set of query mutants (Figure 1B) included also 11 strains modified by the DAmP strategy and 15 deletion mutants affecting individual snoRNA genes, involved in targeting 2'O-methylation and pseudouridylation of rRNA (46). For four essential genes, we decided to test the flexibility of the GIM approach, and evaluate if a *Tet-off* system (47) could

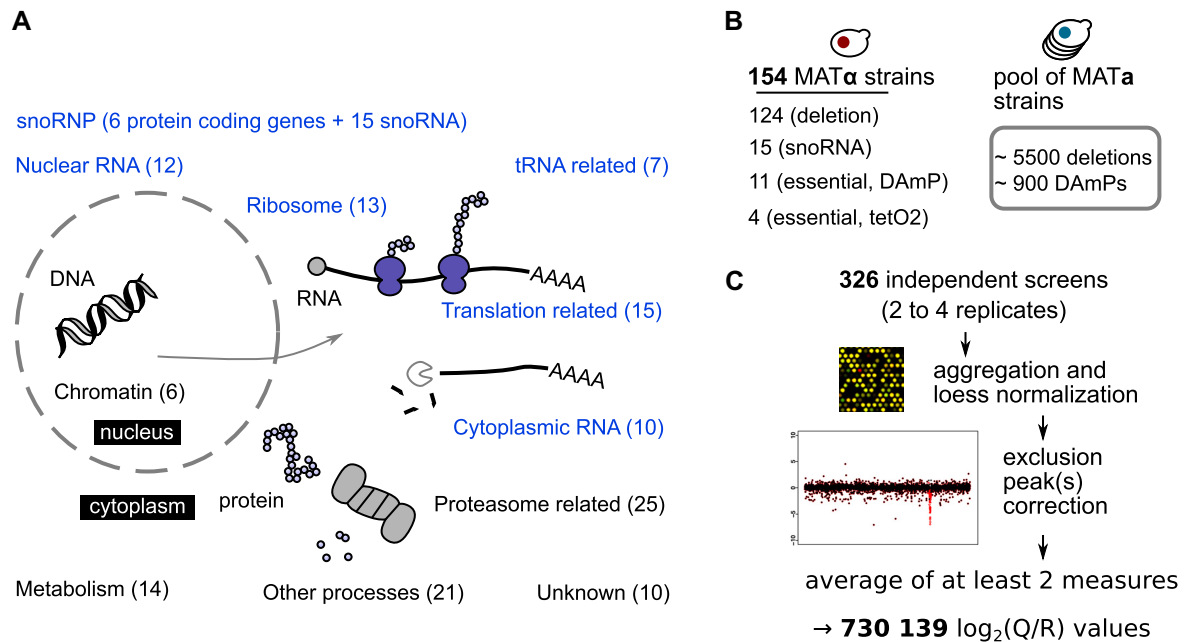


Figure 1. Overview of the cellular functions of query genes tested in GIM screens. (A) Classification of the tested mutants in broad groups associated with major cellular processes, including mRNA translation, protein degradation and ribosome function and biogenesis. The number of genes for which we performed GIM screens from each class is indicated, with RNA-related processes highlighted in blue. (B) Three types of mutants were used in screens, mostly gene deletion, but also DAMP and regulated expression strains (left). The pool of barcoded deletion strains used in each screen was supplemented with our collection of DAMP strains for essential genes (right). (C) The workflow for analysing the microarray results involved normalization, correction of the signal peaks that indicate the low frequency of meiotic recombination that occurs for loci situated close on the same chromosome and averaging of values from independent screens. The initial signal and corrected version for each of the screens are presented in Supplementary Data Set 1.

be used to study GIs of essential genes, as an alternative to DAMP or TS mutants. In these cases, the screen protocol, schematically depicted in Supplementary Figure S1, included the addition of doxycyclin in the final culture where double mutant haploid strains are selected in the presence of geneticin (G418) and nourseothricin. The effect of doxycyclin on the growth of *Tet-off* strains was estimated on plates, with strong effects in the case of *RAT1* and *RNT1*, a milder effect on *SSU72* and no clear effect on *RRP6*. However, in all strains, a strong decrease of the corresponding mRNA levels was observed, to values that were, after only 4 h of doxycyclin addition, lower than 50% of the normal mRNA levels for each gene (Supplementary Figure S2).

Each screen was performed at least twice, leading to results for 326 independent experiments (list of the experiments in Supplementary Table S2). DNA extracts from pools of double mutants, including a query gene and a target gene perturbation, were labeled and used for microarray hybridization. The obtained microarray data were normalized and the results for the two barcodes that are characteristic for each mutant were aggregated. For each group of two or more screens performed with the same query strain, we computed the average correlation coefficients between the results from independent replicates. The Pearson correlation coefficient ranged from 0.05 for *nop6* Δ to 0.95 for *ubp6* Δ (Supplementary Figure S3). The lack of correlation between the screens performed with *nop6* Δ , for example, originated from the very low amplitude of the measured GI in the replicate screens. On the contrary, in the case of *ubp6* Δ , where a large number of negative GIs were detected,

the obtained values showed high levels of correlation between replicated experiments.

Each screen showed a peak of signal that corresponds to the decreased meiotic recombination frequency for genes located close to the ‘query’ gene locus (8,48), which was corrected (Figure 1C and Supplementary Data S1 for the corrections of each screen). The raw results of query versus reference ratios (Q/R) were normalized across target genes and screens, to obtain a primary table of 730 139 ratios between the levels of a mutant in a given screen and its levels in a control population (\log_2 transformed values, Supplementary Table S3). Negative values of $\log_2(Q/R)$ correspond to a depletion of a given mutant when combined with the specific ‘query’ allele, null values indicate no interaction, while positive values suggest an epistatic relationship or suppression of a slow growth phenotype for the query strain.

When looking at the distribution of the obtained values for each target mutant, we observed that several strains showed large relative growth defects in screens performed with unrelated mutants. For example, the distribution of the scores observed for *VPS63* was different from the average cumulative distribution of the measured growth defects, with a much larger spread (Figure 2A). For comparison, the distribution of scores for the nuclear exosome factor *MPP6*, known to show a very specific response to perturbations of the nuclear exosome, showed a very steep slope (Figure 2A). To identify other mutants following the same trend as *VPS63*, we took the number of screens in which the $\log_2(Q/R)$ of a target mutant was inferior to -1.25 , and expressed it as the ratio of the total number of screens in

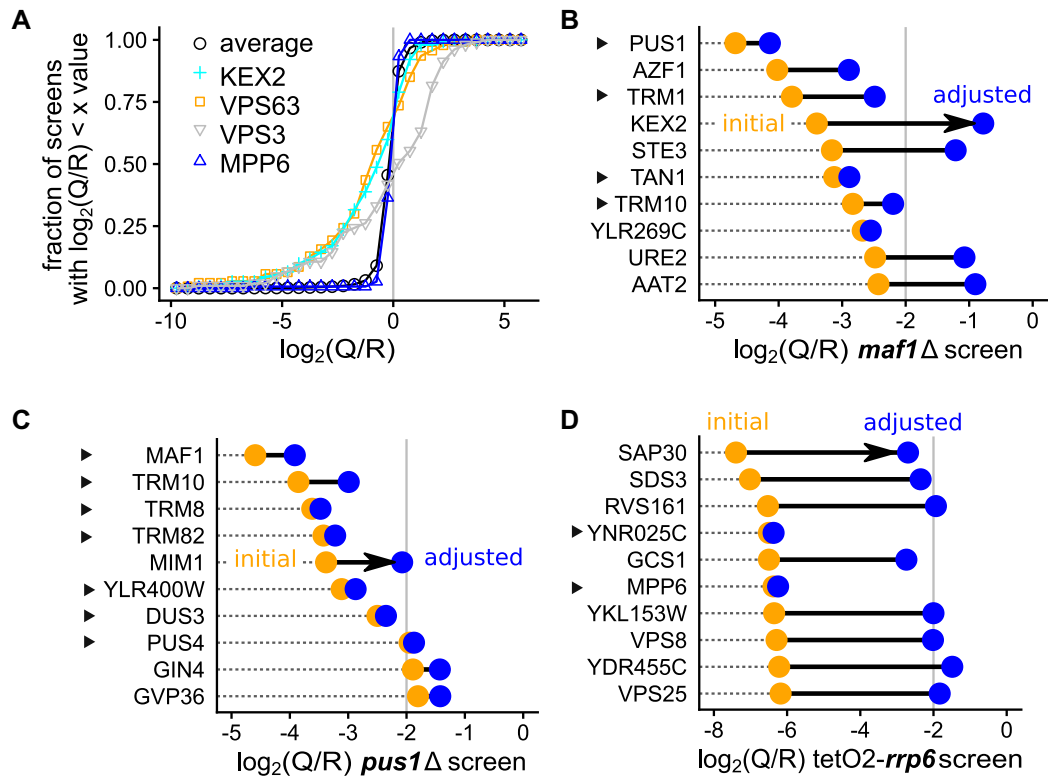


Figure 2. Correcting for pleiotropy improves ranks of genes functionally related with the tested mutant. (A) For each measured effect of a mutant in the 154 screens, we evaluated the cumulative distribution of the $\log_2(Q/R)$ values. Results for genes having an unusual behavior are displayed, including *KEX2* (blue cross), *VPS63* (orange square), *VPS3* (downside gray triangle) compared with the average for all screens (black circle) or for a mutant showing highly specific interactions, *MPP6* (upside dark blue triangle). Examples of applying a correction based on pleiotropy to the ranks of the 10 best hits for the screens performed with *maf1*Δ (B), *pus1*Δ (C) and *tetO2-rrp6* (D). Initial scores are indicated by orange dots and adjusted values are illustrated as blue dots. Genes marked with a triangle correspond to mutants that are known to affect the same pathway (tRNA synthesis for *MAF1* and *PUS1* and RNA degradation in the nucleus for *RRP6*).

which the mutant was measured. The calculated value is a ‘pleiotropy index’ (PI) specific to our data set and has values between 0 and 1, with higher values indicating a broader shoulder of the values distribution. For example, a PI value of 0.5 would indicate that a gene deletion was seen deleterious for growth in combination with half of the query genes used in the 154 screens. The values of PI were 0.44 for *VPS63*, the maximum in our results, 0.39 for *KEX2*, 0.26 for *VPS3*, and 0.01 for *MPP6* (Supplementary Table S4). Only about 11% of the tested strains had PIs higher than 0.1 (539 out of 5063 measured strains). When ranking genes in decreasing order of measured PI, we observed an enrichment of functions related to intracellular vesicular transport. Thus, 15 of the top 32 genes were annotated with the GO term for biological process ‘16192’, ‘vesicle-mediated transport’, with an adjusted *P*-value for functional enrichment of 5.4×10^{-6} , as tested using the g:Profiler tool (34). Correlated with this result, affecting the function of genes from this group leads to pleiotropic morphological changes of yeast cells (49). Perturbation of the intracellular transport of macromolecules or metabolites can affect a relatively large number of different cellular processes, which probably explains these results.

Since mutants showing pleiotropic effects are not informative and their effects can mask more interesting func-

tional interactions, we adjusted the screen result values by multiplication with a correction factor derived from the PI. Among several possible transformations, we applied one that improved the identification of known GIs in the screens performed with *maf1*Δ, *pus1*Δ and temporary depletion of RRP6 (*tetO2-rrp6* strain). After testing several transformations, we chose to multiply the original $\log_2(Q/R)$ values with $(1 - \text{PI})^3$, which had little effect on most results, but diminished the relative contribution of the highly variable mutants. More than 80% of the results were only slightly corrected, by factors between 0.8 and 1, and only 3% of the results were decreased with a factor of more than 2 (Supplementary Figure S4).

When looking at the top 10 hits of the screens mentioned above, we observed the effects of the applied corrections. For example, *Maf1* is a major regulator of RNA polymerase III activity (50, for a review) and is thus tightly linked with tRNA metabolism. The first and third most affected gene deletions affected by *maf1*Δ, *PUS1* and *TRM1*, are both linked with tRNA modification (51, for a review). However, the fourth and fifth values in the *maf1*Δ screen correspond to *KEX2*, coding for a protease involved in the secretory pathway, and *STE3*, a membrane receptor. The pleiotropy correction effectively filtered out these results, while improving the ranks of *TAN1* and *TRM10*, linked with tRNAs

(Figure 2B) which were promoted in the top 5 of the adjusted results. In the screen using the *pus1*Δ strain, affecting tRNA modification, the fifth hit was most affected by the pleiotropy correction (Figure 2C). *MIMI*, the corresponding gene, has no known link with tRNA. However, both the first four hits and the next three correspond to genes affecting tRNA modification or synthesis: *MAF1*, *TRM10*, *TRM8*, *TRM82*, *DUS3*, YLR400W, which partially overlaps *DUS3*, and *PUS4*.

The effects of the applied corrections were occasionally impressive, as seen with the screen in which the expression of *RRP6* was blocked by doxycyclin addition during the growth of the double mutant strains. *RRP6* is a 3' to 5' exonuclease that associates with the nuclear exosome and is involved in RNA synthesis, maturation and degradation (52, for a review). However, 8 out of 10 top hits in the corresponding screen did not match known genetic interactions for this factor (Figure 2D). These 8 factors were among those highly variable in many screens and the corresponding values were strongly reduced by applying the pleiotropy correction. The remaining genes, *MPP6* and deletion of overlapping YNR025C, were the top hits of the tetO2-*rrp6* screen, in agreement with previous results obtained with an *rrp6*Δ strain (15). Thus, correction for pleiotropic effects can help in recovering important functional information, with variable efficiency depending on each screen particular conditions. Among the corrected results, there were 2356 gene deletions and 402 DAmP mutants with at least one adjusted log₂(*Q/R*) value lower than -1 (Supplementary Table S5). Next, we wondered how good the measured GIs were, and, for this task we used several criteria, as described below.

We were confident that the identified GIs were meaningful, since the screen results were compatible with current known annotations and with previously published data sets (annotated examples in Figure 2B–D). However, we wanted to assess the quality of the measured GIs globally. We thus took advantage of the fact that GI profiles, the set of values obtained for a given mutant, provide more information than direct GIs for inferring gene function (8). Thus, we used correlation of GIs to test the validity of the newly obtained data set. For a first validation of the adjusted results we looked for the similarity of GIs for the same gene mutant when tested in the query MATα strain (Nat^R marker) or in the tested pool of target MATα strains (Kan^R marker). Pearson correlation values of the GI profiles for the 127 pairs of genes tested independently were clearly skewed towards positive values, as expected. In contrast, correlation for all the possible pairs of GIs in the data set showed a bell-shaped distribution centred on zero (Figure 3A). Thus, the phenotype of mutating the same gene was similar, whether the mutant was present in the query strain or in the pool of tested strains. We performed a similar analysis for cases of overlapping gene deletions to analyze the correlation between the effect of independent mutations affecting the same locus. Since two deletions affect the same gene, the two strains should behave similarly in the screens. For the available pairs of overlapping gene deletions, we observed a strong positive correlation for their GI profiles (Figure 3B, right). In conclusion, correlated GI profiles for mutants affecting the same gene (Figure 3A) and for overlapping mu-

tants (Figure 3B) globally validated the quality of our large-scale screen results.

The GI profiles from the GIM screens were compared with results obtained with similar mutants using the SGA approach (13). There were 52 screens that were done with mutants affecting the same gene in the two sets. For many screens, a positive correlation coefficient between the SGA and GIM results indicated that part of the observed GIs were similar among these two independent assays. On other occasions, no correlation could be detected. This lack of correlation was a characteristic of GIM screens with low variance of the results. However, in specific instances, such as the deletion of *GCN2*, relatively strong GIs were observed in both the GIM and SGA assays, even if the identified gene deletions were different (Supplementary Figure S5). It is unclear if GIM results, SGA results, or both are correct in this case. However, the strongest negative GI that we observed with the GIM screen corresponded to deletion of *TAT2*, a gene involved in aromatic amino acid transport (53). Since *GCN2* is involved in the adaptation of yeast to amino acid starvation, this result is compatible with the current knowledge about *GCN2* and *TAT2* function.

We focused on the situations where SGA and GIM results were correlated, as these cases depend on GIs that were robustly detected across assays and laboratories. As examples of GIs responsible for the observed correlations, we present the comparison of the GIM and SGA screens performed with *mafl*Δ (Figure 3C) and the comparison of the SGA *rrp6*Δ screen with *RRP6* depletion in the GIM screen using a *Tet-off* system (Figure 3D). The correlation between *RRP6* deletion and its depletion by *Tet-off* shows that the GIM protocol can be adapted to new ways to affect gene function. The mRNA depletion by transcription inhibition using the *Tet-off* system is particularly appealing, since this strategy should work for any essential gene.

While some results were specific to either GIM or SGA assays, thus being condition and assay-dependent, a number of GIs were detected in both types of screen at the arbitrary threshold of -1 for GIM screens (3951 GIs) and -0.1 for SGA screens (7947 GIs). We thus generated, from the about 180 000 GIs reported by both assays, a list of 479 GIs detected by both GIM and SGA screens (adjusted log₂(*Q/R*) score lower than -1 and SGA fitness score lower than -0.1, Supplementary Table S6). For example, the negative effect of combining the deletion of *REX2* and *MYG2*, two 3' to 5' RNA exonucleases that have both nuclear and mitochondrial localization (54,55), might indicate a cooperation of these enzymes for the degradation of an important RNA substrate or in the maturation of ribosomal RNA. Another interesting example is the strong interaction observed between *RSA3*, a gene involved in ribosome biogenesis (56), and YCR016W, which is uncharacterised but shows a regulation pattern compatible with a similar role (57). Thus, the list of GIs that were observed in very different assay conditions could serve to benchmark future large-scale GI screen results and to launch new research projects.

The results presented here were thus validated by the correlation of GI profiles for the same gene mutation either as query or target in screens and by overlapping target gene deletion analysis. In combination with previous SGA

results, our data also validate several hundred GIs that can be considered robust.

The amplitude of GIs for DAmP alleles correlates with specific gene features

Since the DAmP mutant strains were used in addition to gene deletion strains in our screens, we wondered if we could obtain insights into global differences between these two types of gene perturbation. DAmP perturbation of gene function depends on the effect of a long 3' untranslated region on RNA stability. The NMD degradation pathway, responsible for destabilization of DAmP RNAs, is potentially more efficient for mRNAs with relatively short coding sequences (14,18). Thus, its impact on the essential gene function, and the profile of GIs for the corresponding mutant, was likely to vary and could be correlated to various RNA features, such as abundance or coding sequence length. We thus looked for a correlation between original mRNA abundance for essential genes and the frequency at which the corresponding DAmP alleles showed a GI in the results. To this end, we arbitrarily defined **screen-responsive** gene perturbations as those in which we observed at least a variation by a factor of 2 for a given gene in at least one of the GIM screens. Mutants that showed no effect in combination with any of the 154 query gene perturbations would be included in the non-responsive category. We calculated which fraction of the tested mutants was in the responsive or non-responsive category in correlation with RNA abundance and coding sequence length.

Interestingly, the percent of screen-responsive DAmPs increased with the abundance of the corresponding mRNAs (Figure 4A). As background, and for comparison, we applied the same analysis to gene deletions, for which the effect of mRNA abundance on the frequency of response in GIM screens was less marked. However, in both cases deletion or DAmP perturbation were most correlated with an effect in GIM screens for the most abundant mRNAs. We also looked at the relation between screen responsiveness and the length of the coding sequence for the affected gene, which is linked with the destabilization of DAmP modified mRNAs (18). In this case, the fraction of screen-responsive DAmP mutants decreased as the length of the initial gene coding sequence increased. This effect of coding sequence length was not found for gene deletions, where, on the contrary, the highest proportion of screen-responsive mutants was found in the group of long genes (Figure 4B). Thus, features associated with an effect visible in the GIM screen conditions were high expression level and large gene size for gene deletion and high abundance mRNA and short coding sequence size for DAmP modification.

To illustrate how useful the new results on DAmP strains can be and further validate the obtained result on a large scale, we focused on a group of 22 DAmP mutants affecting proteasome-related genes, which are highly expressed and can be relatively short. For example, 11 out of the 22 selected genes have coding sequences shorter than 1176 nucleotides, which places them in the first two bins represented in Figure 4B. We ranked the screens to find those in which the median of the adjusted $\log_2(Q/R)$ values for this group of proteasome DAmP mutants was lowest. The top 5 screens

showing GIs with proteasome-related genes were, in order, those using as query genes the deletion of *RPN10*, the depletion of *RAT1*, and deletions of *RPN4*, *PRE9* and *POC4*. Four out of the five screens in which DAmP proteasome mutants were most affected corresponded thus to perturbation of proteasome components *RPN10* and *PRE9*, of a regulator of proteasome formation, *RPN4*, and of the proteasome assembly factor *POC4*. Values for the DAmP proteasome mutants in those screens were clear outliers, when compared with the overall distribution of adjusted $\log_2(Q/R)$ values in each screen (Figure 4C). In the screens performed with deletions of *RPN10* and *RPN4*, DAmPs for proteasome-related genes represented the majority of strong negative measured GIs, as illustrated in Figure 4D.

A fifth screen showing a strong global effect in combination with proteasome DAmP mutants involved the temporary depletion of the major nuclear 5' to 3' exonuclease *RAT1* using the *Tet-off* system. This result was surprising but compatible with the various roles of the proteasome in transcription (58) and the described role of Rat1 in RNA polymerase II transcription termination (59). Alternatively, it could be an illustration of the deregulation of protein homeostasis following Rat1 depletion, which might require compensation by an increase in proteasome activity (60,61). Thus, the use of DAmP mutants in combination with the *Tet-off* strategy for query gene perturbation uncovered GIs that are functionally relevant and potentially important.

In addition to the proteasome analysis detailed in this section, DAmP modification led to the identification of other new links between genes involved in RNA metabolism. Thus, for example, the DAmP modification of the 3' to 5' RNA degradation exosome complex component *RRP46*, and of the rRNA modification complex component *NOP56* were synthetic sick with the deletion of several RNA-related genes (Supplementary Figure S6), including the poorly characterized locus YCL001W-A and the recently identified SKI complex associated protein Ska1 (62). Altogether, these results illustrate the value of including mutants of essential genes in GIM screens.

Predicting function based on GI profiles by using RECAP

Associating genes with a cellular pathway is often based on the observation of a specific phenotype when the gene function is affected by deletion, down-regulation or mutation. A different type of phenotype, tested in large-scale genetic screens, is represented by the constellation of gene perturbations that, when combined with a yeast gene mutant of interest, have an effect on the strain's growth rate. This profile of response of a mutant to combinations with the query gene perturbation, also called GI profile, can be used to find functional relationships from screens data (8,10,13). GI profile similarity is thus an important type of result derived from large-scale genetic screens. We wondered whether we could use GI profile similarity from our results and integrate it with curated literature data on yeast genes to understand : (a) what fraction of known functional interactions can be reached with our GI data set and (b) whether we can use the similarity of GI profiles to assign new genes to known cellular pathways. To answer these questions, we developed a data analysis strategy called RECAP

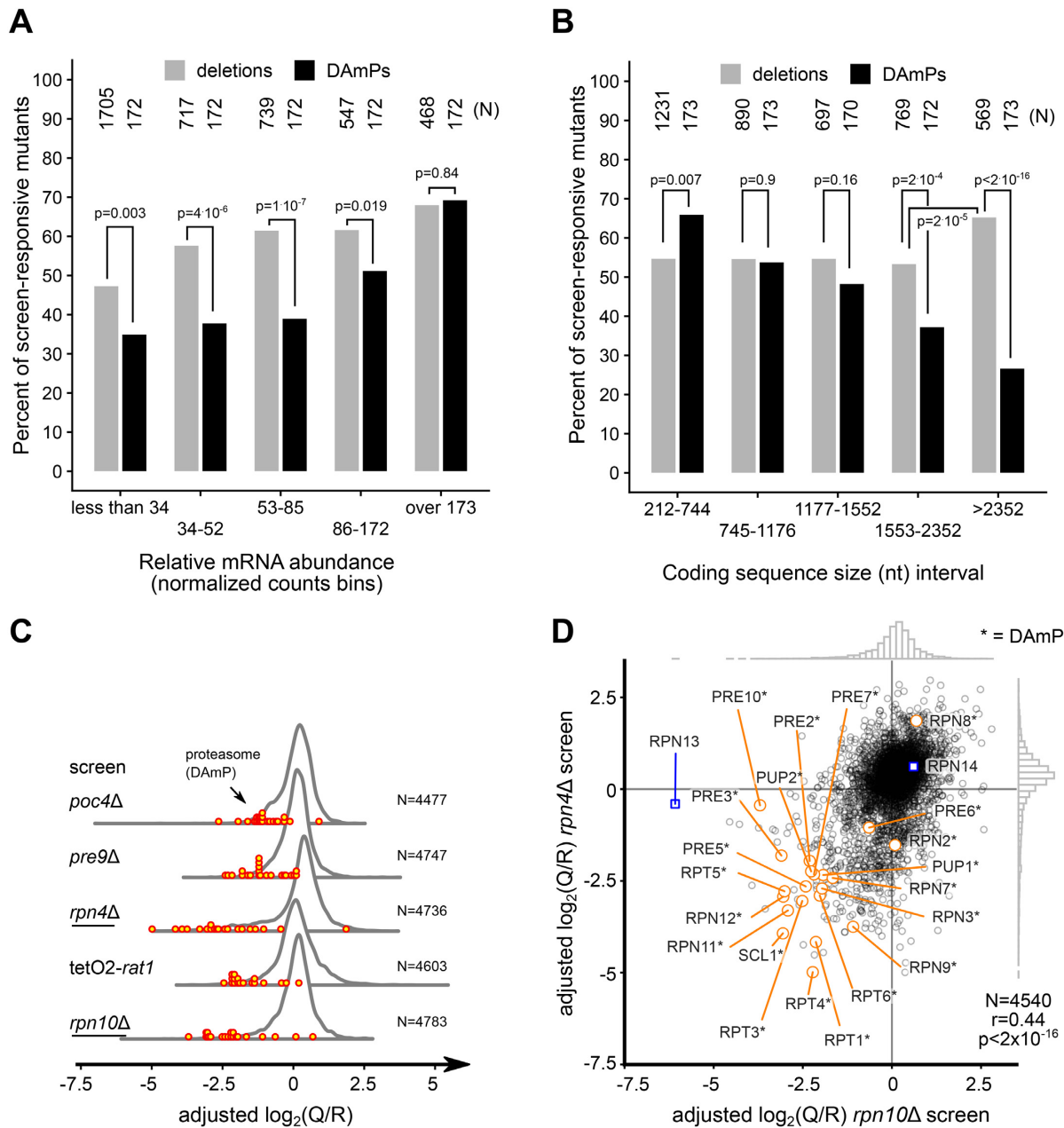


Figure 4. DAmP perturbation has effects correlated with mRNA abundance and coding sequence length and is valuable for the study of major cellular functions. (A) We arbitrarily assigned the various mutants from this study in two classes: screen-responsive, if the corresponding mutant showed a growth defect score of at least 2 ($\log_2(Q/R) < -1$) in at least one of the 154 screens, and screen-neutral if the mutant was not affected in any of the screens. The percent of screen-responsive deletion (light gray) and DAmP (black) strains was plotted as a function of relative mRNA abundance (104), with transcripts grouped in five bins having identical numbers of DAmP mutants. The differences between the numbers of DAmP and deletion mutants in each bin were evaluated with a chi-squared test (the *P* value for the null hypothesis of identical percentages is indicated). The number of genes in each bin is indicated in the upper part of the panel. (B) Equal sized bins of DAmP mutants were created based on the coding sequence length and the percentage of screen-responsive strains was compared with the results for deletion mutants for genes having similar sizes of coding sequences. The number of mutants in each bin is indicated. (C) We used the median of the relative rank for 22 DAmP mutants affecting proteasome and proteasome-related genes to identify the five screens in which these mutants were most affected (in increasing rank order from bottom to top). The distribution of all adjusted $\log_2(Q/R)$ values in the five selected screens is indicated. Red dots indicate the position of the adjusted $\log_2(Q/R)$ scores for proteasome DAmP mutants. (D) Specific DAmP effects are illustrated by a scatter plot showing the correlation between the adjusted $\log_2(Q/R)$ scores obtained when the screen was done with the deletion in each *RPN10* proteasome component gene (horizontal axis) compared with the deletion of the *RPN4* proteasome regulator (vertical axis). DAmP proteasome related mutants are indicated in orange and two non-essential proteasome gene deletions are indicated in blue.

(Rational Extension of Correlated Annotations and GI Profiles, summarized in Supplementary Figure S7), which, instead of focusing on the GI network, starts from published data curated by the Saccharomyces Genome Database maintainers (28). The main objective of this strategy was to validate target gene behaviour in GIM screens on the basis of previous knowledge on gene function. Once validated, the GI profiles could serve to find related genes by GI profile similarity.

To establish links between genes from the table that associates genes and publications, we first removed publications associated with more than 100 genes, since we considered that such publications are too general to be informative. The remaining literature corpus consisted of 76 160 publications. We restricted our analysis to the upper half of the most cited yeast genes, leading to a selection of 3575 genes or genomic features cited in at least 31 scientific publications. Among these well studied genes, 1847 were present in our GIM data set of 5 063 genetic interaction profiles. For these 1847 genes, we identified 4072 linked gene pairs. We considered two genes, A and B, to be linked by co-citation if, for each gene A, at least 20% of its publications mentioned gene B and reciprocally, if 20% of citations for B also contained A. To identify groups of related genes we used the Louvain algorithm (63), designed to define groups of nodes in a network by maximizing modularity, a value that reflects the difference between the density of links between nodes inside a group and the number of links between groups.

The set of 4072 gene pairs was thus split in 439 communities of related genes corresponding either to well-studied complexes or to well-known genes involved in the same cellular pathway (Supplementary Table S7).

In the next step of the RECAP approach, we wanted to combine the newly defined literature-based clusters with the information available from the similarity of GI profiles in the GIM data. We calculated Pearson correlation for each pair of the 5063 GI profiles of target genes in our data set. For each mutant, we sorted the obtained correlation coefficients in decreasing order and arbitrarily considered two mutant profiles, X and Y, to be linked if the correlation coefficient of X with Y and of Y with X were in the top 20 of correlation coefficients for both X and Y. This choice increased the specificity of the method and avoided situations in which spurious correlations would pollute the results.

Having established literature and GIM-based links between genes, we wondered how to combine these results. Among the 439 communities of related genes identified from the literature, 117 had at least two genes linked by GI profiles similarity in our GIM data set (Supplementary Table S8). To visualize these results, we made a selection of 35 groups that had at least four genes from the GIM data showing correlated profiles to other genes from the same sub-network (Figure 5A). These groups comprised 550 genes and 2393 co-citation links and involved data for 270 genetic interaction profiles, including those of 67 DamP mutants. The purpose of this visualization was to show that cellular functions that corresponded to the selected gene groups covered a wide range of cellular functions, from DNA transcription to vesicular transport and mitochondrial function (Figure 5A, see annotations).

Our literature-based analysis of GI profiles indicated which mutant strains had phenotypes specifically correlated with the function of the corresponding gene. This knowledge allows focusing on these mutants first, since they were independently validated to provide functional information. Using this knowledge is essential to avoid conclusions about gene function that would come, for example, from perturbing an unrelated gene that is physically close on the chromosome (64). We thus used the validated GI profiles from each of the established 117 co-citation based clusters and used GI similarity to find related GI profiles. Importantly, if genes for which we had profile data were present in co-citation clusters but were not linked by GI profile similarity to other genes in the same cluster, these profiles were ignored.

An example of the performance of this approach is shown in Figure 5 for the group of literature-linked genes *MAD1*, *MAD2*, *MAD3* and *BUB1* (a paralogue of *MAD3*). The MAD genes contribute to the spindle assembly checkpoint in relation with kinetochores, and are required for cell division (for a review, 65). This group of four genes included the correlated GI profiles for the *MAD1/MAD2* and for the *MAD1/MAD3* pairs (Figure 5B). The RECAP-extended network based on the MAD gene group included other genes with roles in the spindle assembly checkpoint, such as *KAR9*, *CTF19* and *BFA1*, but also many genes related with microtubule cytoskeleton organization and function, a process that is linked directly with spindle assembly and function (Figure 5C). A few genes were not annotated directly to spindle assembly or microtubule function, such as *RAD53* (DamP modification), *SET2*, *ZIP2* or *MCM21*. However, *MCM21* is a component of the COMA kinetochore sub-complex (66), *RAD53*, a gene with multiple roles in DNA repair, is also linked with mitotic checkpoints (for a review, 67) and *ZIP2* is involved in homologous chromosome pairing in meiosis (68). Thus, the large majority of the genes identified by RECAP starting from just a few components of the spindle assembly machinery had functions in relation with this process.

We used GO term enrichment analysis (34) to establish the biological process and cellular component that were predominant in each of the 117 literature-defined communities and associated the genes from the extended RECAP network to these processes (Supplementary Table S9). A total of 1471 genes involved in 3893 gene pairs were finally associated with defined processes, allowing new hypotheses about these genes function to be tested by future oriented experiments. To assess the robustness of these results, we used the Curated Yeast Complexes (CYC2008) data set (69) as a well established standard for links between genes. We compared first this standard with the co-citation table that we built from literature data and found that about 33% of the links from this table, used to validate GI profiles, was present in the CYC2008 data set. This percentage was lower for pairs of GI profiles based on reciprocal top 20 rank (5%). However, validation through RECAP analysis led to an increase of this percentage to about 17%. As a control, we randomized the CYC2008 data set and estimated, once again, the percentage of links that are likely to occur by chance. This percentage was lower than 1% in all cases (Supplementary Figure S8). This analysis also allowed us to look at the RECAP-based results and how they correlated with

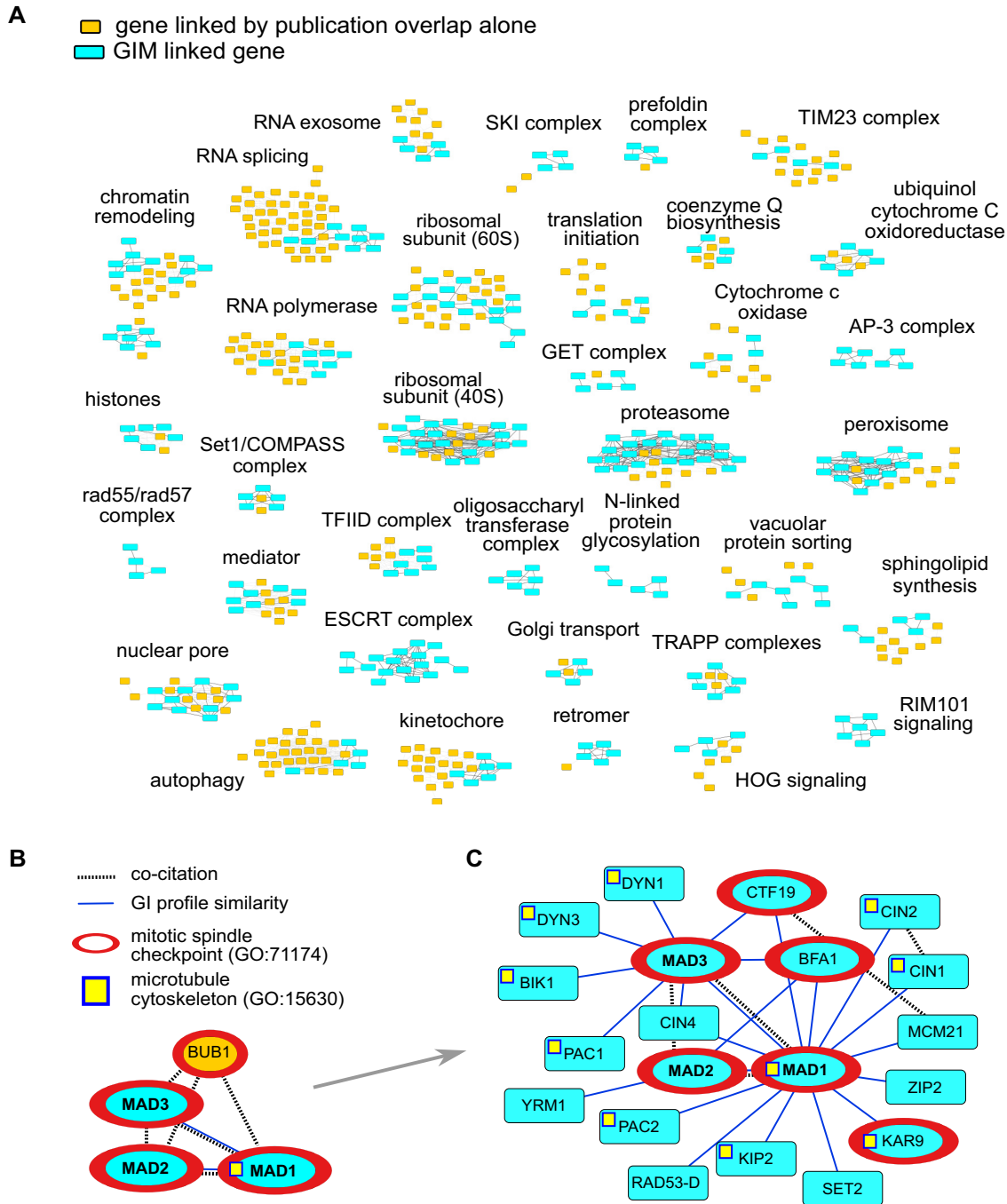


Figure 5. Integration of literature data and GIM profile similarity allows extension of known functional networks. (A) Links between genes were inferred from literature data obtained from the *Saccharomyces* Genome Database and were used to define functionally related gene groups. If two genes in a group were, in addition to a literature link, also connected by correlated GI profiles from our data set, the two genes were highlighted in cyan, otherwise, the colour used was orange. Only a selection of 35 gene groups in which at least 4 genes showed correlated GI profiles is shown. Each gene group was annotated manually, either in terms of a protein complex or based on known cellular or molecular function. (B) Example of a literature-based gene group, not shown in A, bringing together several genes involved in mitotic spindle checkpoint. The GI profile similarity between deletions of *MAD1* and *MAD3* and *MAD1* and *MAD2* are shown as solid lines, while co-citation links are shown as dotted lines. Extending this network using only the similarity of GIM profiles led to the network shown in C. (C) Starting from *MAD1*, *MAD2* and *MAD3*, the GI similarity-based functional network adds supplementary genes involved in mitotic spindle checkpoint, such as *KAR9*, *BFA1* and *CTF19* (marked with a red outline) and genes involved in the dynamics of the microtubule cytoskeleton (marked with a yellow square). All the shown links, depicted as blue lines, represent GI profile similarity above an arbitrary threshold. Dashed lines represent co-citation.

co-citation and CYC2008 links. Four examples are shown in Supplementary Figure S8, to depict new GI profile similarity links that might lead to the discovery of new genes and complexes with roles in mRNA splicing, NMD, RNA deadenylation or unconventional cytoplasmic splicing.

The RECAP strategy is not limited to GI profile similarity, but can be adapted to filter other large-scale data sets in which links between genes have been established.

Linking inositol polyphosphate metabolism and translation initiation

In addition to GI profile similarity, the discovery of individual GIs can also be informative for gene function. We explored in detail the synthetic lethality between the deletion of *LOS1*, a gene involved in tRNA export from the nucleus to the cytoplasm (51, for a review), and several OCA genes. *OCA1*, the founding member for the OCA nomenclature, was initially identified as an Oxidant induced Cell-cycle Arrest factor (70). The other five members of this protein family were identified based on protein sequence similarity (71,72). Only recently a biochemical role was attributed to *OCA3* in the hydrolysis of specific inositol-polyphosphate species (73). Modification of inositol-polyphosphate levels in OCA mutants is probably responsible for the observed phenotypes when OCA genes are deleted, ranging from changes in replication of an RNA virus (74) to effects on yeast prion propagation (75).

We have previously observed a strong growth defect when *OCA2* and *LOS1* deletions were combined (8). This link was confirmed when the deletion of *OCA4*, another OCA gene, was tested by SGA, although *LOS1* deletion was only found as the 101st most affected hit (22). GI profiles for deletions of *OCA1*, *OCA2*, *OCA3* (*SIW14*) and *OCA5* were highly similar in the SGA data (22). All the OCA deletion mutants also showed a coordinated response to a set of chemical compounds (76), indicating that loss of these proteins leads to a similar cellular response. In view of the similarity between OCA deletion profiles, we compiled physical interaction results about OCA proteins from the BioGrid database (77). Except for Oca6, we found evidence for physical interactions between OCA proteins, potentially in a multimeric complex (Supplementary Figure S9).

In correlation with the previous data on OCA genes and *LOS1*, we found that *LOS1* deletion showed, among our 154 GIM screens, the strongest negative effect on growth when combined with deletions of *OCA2*, *OCA4*, *OCA6* and *OCA1* (Figure 6A). The sixth screen in which deletion of *LOS1* strongly affected growth involved the deletion of *RIT1*, which modifies initiator tRNA and renders it incompetent for translation elongation (78). This result was compatible with the role of *LOS1* in tRNA export, including tRNA_i^{Met}, and suggested that the other observed GIs for *LOS1* had high confidence. To further validate the GIs between OCA genes and *LOS1* deletion, we tested the growth of single and double mutant strains in various media. We found that moderate doses of lithium chloride in rich medium sensitized the growth assay for the *los1Δ*, *oca2Δ* strain. Under these condition, expression of *LOS1* and *OCA2* from plasmids (79) partially restored growth of the double mutant strain, in support of the screen results (Figure 6B).

To further explore OCA genes role and how they could be connected with *LOS1* function, we analyzed results obtained during large-scale transcriptome profiling of deletion mutants (80). Among the tested deletion mutants, the data set contained the transcriptome measures for the effects of deleting *OCA5*. We noticed that the transcriptome changes in this *OCA5* mutant were inversely correlated with those observed when the translation-regulated transcription factor *GCN4* (or its overlapping gene YEL008W) were deleted.

GCN4 is one of the best studied example of translation regulation and plays a crucial role in adaptation of yeast cells to amino acid starvation (for reviews, 81, 82). Activation of *GCN4*, whose mRNA contains four short open reading frames upstream the start codon, occurs when translation is inhibited and leads to transcription of hundreds of targets, including many genes involved in amino acid synthesis (83). Such *GCN4* targets were responsible for the strong inverse correlation between the transcriptome results in the *gen4Δ* strain compared with *oca5Δ* (Figure 6C). To validate this observation, we measured the changes in the levels of two representative transcripts, *ARG1* and *SNO1*, in strains deleted for *OCA2* and *GCN4*, by reverse-transcription and quantitative PCR. Correlated with the published results, *OCA2* absence led to an increase, while *GCN4* absence led to a decrease in their levels (Figure 6D and E). The increase in the levels of these transcripts in the absence of *OCA2* was further enhanced in double mutant strains combining the deletion of *OCA2* with the deletion of *LOS1* (Figure 6D and E). Altogether, these results suggested a link between OCA genes and translation regulation that was potentiated in the absence of the tRNA export factor *LOS1*.

A possible explanation for the observed results was that the loss of OCA genes leads to *GCN4* activation. To test this hypothesis, we used a reporter system in which β-galactosidase is expressed in a *GCN4*-like configuration, with its coding sequence fused with the 5' untranslated region of *GCN4* (84). In this system, we observed a clear increase in the β-galactosidase activity when *LOS1* deletion was present in an *oca2Δ* strain (Figure 7A). Thus, the perturbation of OCA function coupled with a tRNA export deficiency led to *GCN4* activation, most likely through inhibition of translation initiation.

A central factor in translation initiation is the tRNA_i^{Met} and, since *LOS1* is involved in its nuclear export, we wondered if it could be involved in the synthetic sick effect observed when combining *LOS1* and OCA gene deletions (Figure 6A and B). Over-expression of tRNA_i^{Met} led to a reversal of the slow growth phenotype of both *los1Δ*, *oca2Δ* and *los1Δ*, *oca5Δ* strains (Figure 7B), indicating that initiator tRNA shortage becomes limiting in the absence of OCA genes. How the inositol-polyphosphate imbalance generated in such strains affects translation initiation remains an interesting question for future research.

Suppression results identify a role for Puf4 in regulating a ribosomal protein function

GIM screens allow measurements of both negative and positive GIs. We thought that it would be interesting to highlight and validate such a situation and, to this end, we focused on *PUF4*. Deletion of this target gene showed a par-

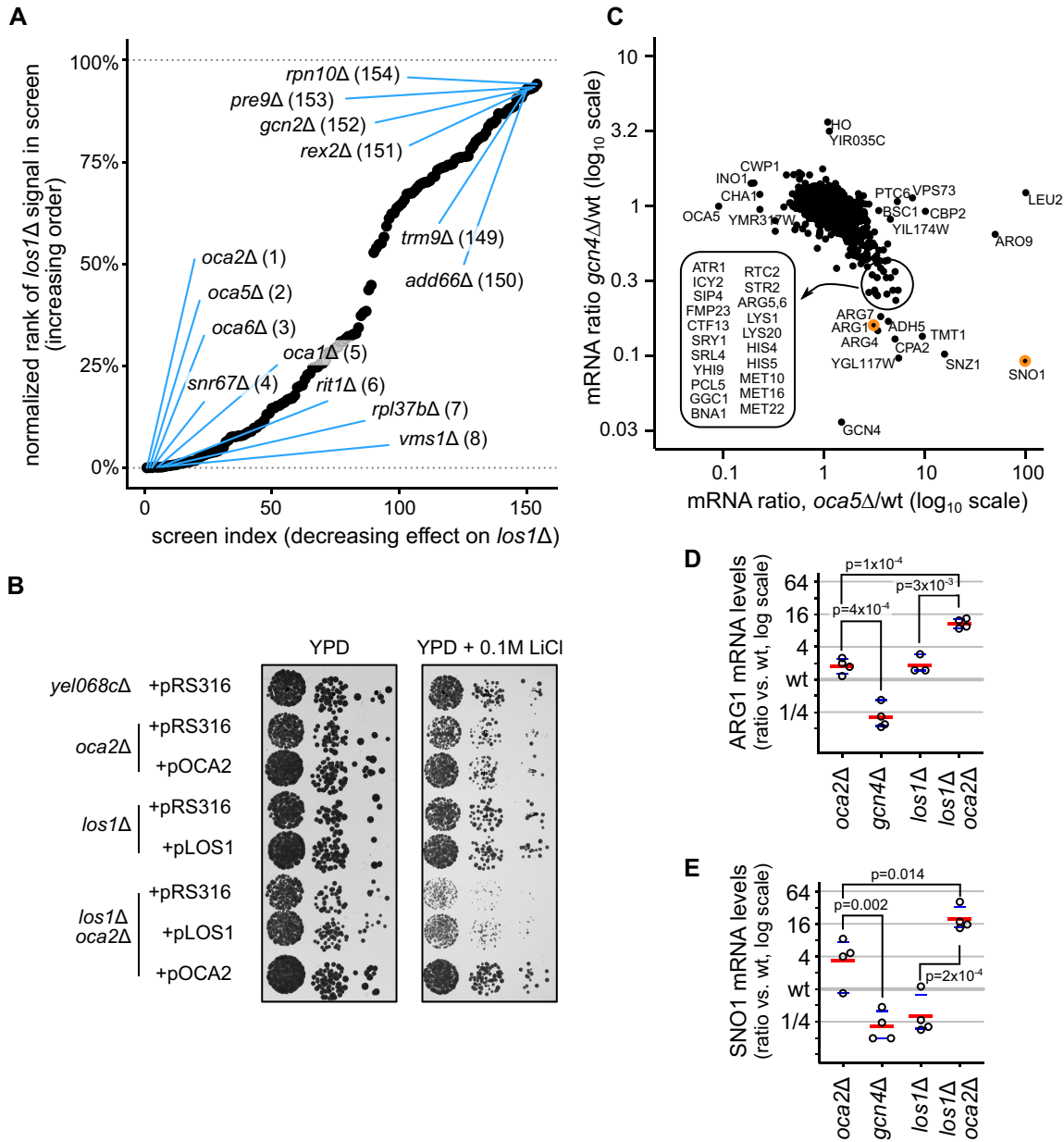


Figure 6. Deletion of *LOS1* is functionally related to a defective OCA complex. (A) Rank analysis of *los1Δ* results for GIM screens highlighting synthetic slow growth (lower left) and potential epistasis (upper right). (B) The double deletion strains combining *los1Δ* and *oca2Δ* were strongly affected by the presence of 0.1M LiCl in the medium. Complementation of growth defect by empty vector (pRS316) or by centromeric plasmids expressing *OCA2* and *LOS1* was estimated by serial dilutions and observation of colonies after 48 hours of growth at 30°C. Deletion of YEL068C served as the equivalent of a wild type control. (C) The inverse correlation between the transcriptome changes in *oca5Δ* and *gcn4Δ* shows transcripts that were up-regulated in the absence of OCA5, while being targets of *GCN4* activation, including many mRNAs that code for amino acid biosynthesis proteins. The position of the signal for mRNA of *ARG1* and *SNO1*, chosen for validation of the transcriptome results, are indicated by orange dots. (D) and (E) Validation by RT-qPCR of mRNA level changes in an *oca2Δ* strain, in comparison with *gcn4Δ*, *los1Δ*, and the combination *oca2Δ los1Δ* for *ARG1* and *SNO1* mRNA, using *RIM1* mRNA levels as reference. Individual measurements for three to four independent experiments are shown, with the red bar indicating the mean and the blue bars indicating limits of the 99% confidence interval (non-parametric bootstrap). The indicated *P*-values correspond to results of single sided *t*-tests.

ticularly strong positive effect when combined with the deletion of the ribosomal protein gene *RPL9A* (Figure 8A). Puf4 is a member of the ‘PUF’ or ‘Pumilio’ family of RNA-binding proteins that participate in post-transcriptional control of gene expression by interacting with specific mRNAs through sequence motifs that are present in the 3′ untranslated region of the affected RNA. PUF proteins affect

mRNA localization, translation or degradation (reviewed in 85, 86). Puf4 targets, defined by *in vivo* tagging and sequencing, correspond to a subset of mRNAs enriched for transcripts coding for proteins involved in ribosome biogenesis (87).

To see if the positive effect of *PUF4* deletion was specific for the strain lacking *RPL9A*, we looked at the re-

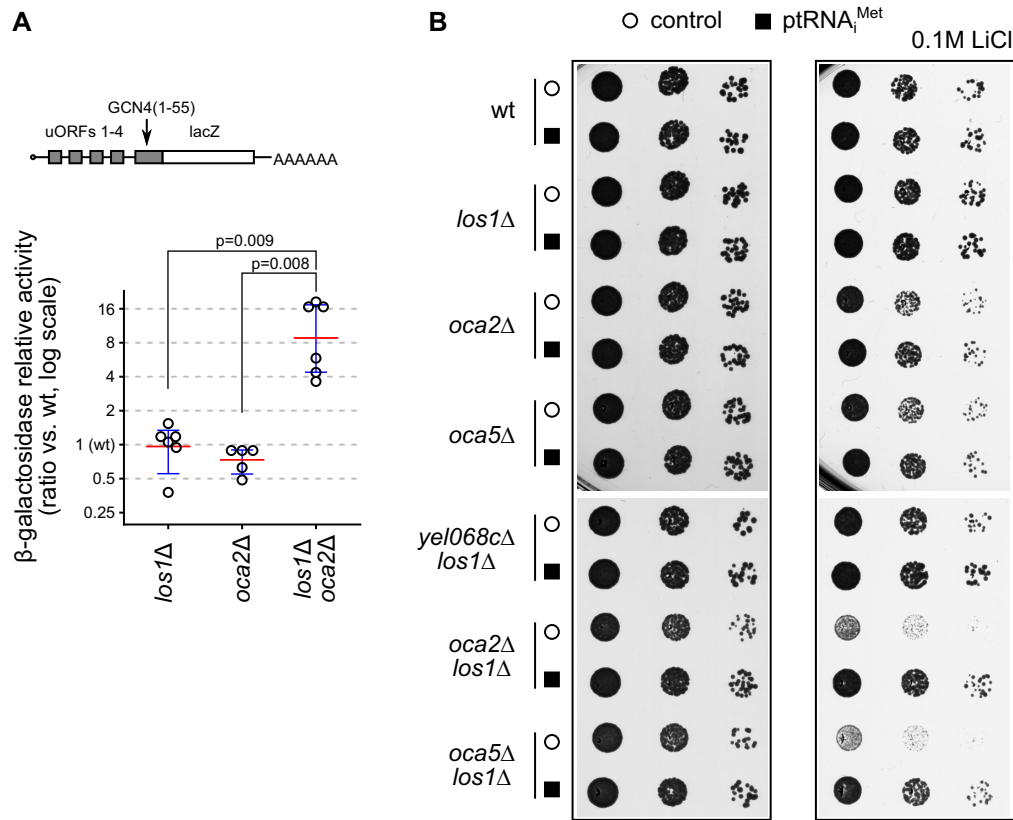


Figure 7. Initiator tRNA limits cell growth when the OCA complex is defective. (A) Double mutant strains show translation initiation defects, as measured using a *GCN4* uORF-lacZ reporter, schematically represented in the upper part of the panel (p180 plasmid, 84). Wild type strain was used as reference and the variation in the amounts of produced beta-galactosidase were measured in at least five independent experiments. The *P*-values of single-sided t-tests for differences between the different conditions are indicated. (B) Over-expression of the $\text{tRNA}_i^{\text{Met}}$ (p1775 plasmid, 105) allows better growth of *oca2/los1* and *oca5/los1* double deletion strains under stress conditions (0.1 M LiCl). Serial dilutions of fresh cells were grown on plates for 48 hours. The presence of an empty vector (empty circle) or of the plasmid over-expressing the $\text{tRNA}_i^{\text{Met}}$ (black square) are indicated. Deletion of *YEL068C* was used as the equivalent of a wild type control.

results for nine other GIM screens in which ribosomal protein genes were deleted. While the strongest positive effect of *PUF4* deletion was found in the *RPL9A* screen, a relatively strong positive GI for *PUF4* was also detected in the case of *RPL7B* deletion. For *RPL37A*, in contrast, we measured a strong negative GI with *PUF4*, while for other ribosomal protein genes the results were intermediary (Figure 8A).

Yeast ribosomal proteins are often coded by duplicated genes. In the case of Rpl9, the A form has the strongest expression both in term of mRNA amount and protein (88). This expression difference is correlated with a growth defect that is more pronounced in the absence of *RPL9A* than in the absence of *RPL9B*, as estimated from previous liquid culture experiments (89). Thus, *RPL9B* mRNA, despite coding for the same Rpl9 protein as *RPL9A*, is less able, under normal conditions, to support yeast growth. To validate the large scale results, we built a double mutant strain combining the deletion of *PUF4* and *RPL9A*. The *rpl9a* Δ strain displayed a slow growth phenotype that could be compensated by expressing Rpl9 from either *RPL9A* or *RPL9B* single-copy plasmids. The slow growth of *rpl9a* Δ was suppressed by deleting *PUF4*, confirming the screen results. Moreover, we could even observe a slight toxicity of a plas-

mid expressing *PUF4* in this double mutant strain (Figure 8B). These results indicate that, in the absence of *RPL9A*, *PUF4* inhibits the function of the *RPL9B* gene.

We wondered if there were elements in the *RPL9B* mRNA that could explain this finding. A consensus motif for Puf4 binding to mRNAs has been experimentally established based an enrichment analysis for co-purified (90) or Puf4-based RNA tagging (87) experiments, further polished by an in-depth comparative analysis of various Puf targets in multiple species (91). The motif recognized by Puf4, 'UGUXXXUA', where X represents any nucleotide, was identified in the 3' UTR of *RPL9B* mRNA and is absent from the same region of *RPL9A* mRNA. Yeast species that are relatively close to *S. cerevisiae* have a similar genome architecture, with two copies for the *RPL9* gene. A conserved element in the 3' UTR of *RPL9B*-like genes was previously shown in our laboratory to be involved in a nuclear regulation mechanism for *RPL9B* expression (44). We searched for consensus Puf4 binding motifs in the potential 3' UTR regions of *RPL9B* equivalents from several yeast species (*Kluyveromyces lactis*, *Lachancea thermotolerans* and *Lachancea kluyverii*) and found that these motifs were conserved and situated at 22 to 42 nucleotides downstream the stop codon (Figure 8C).

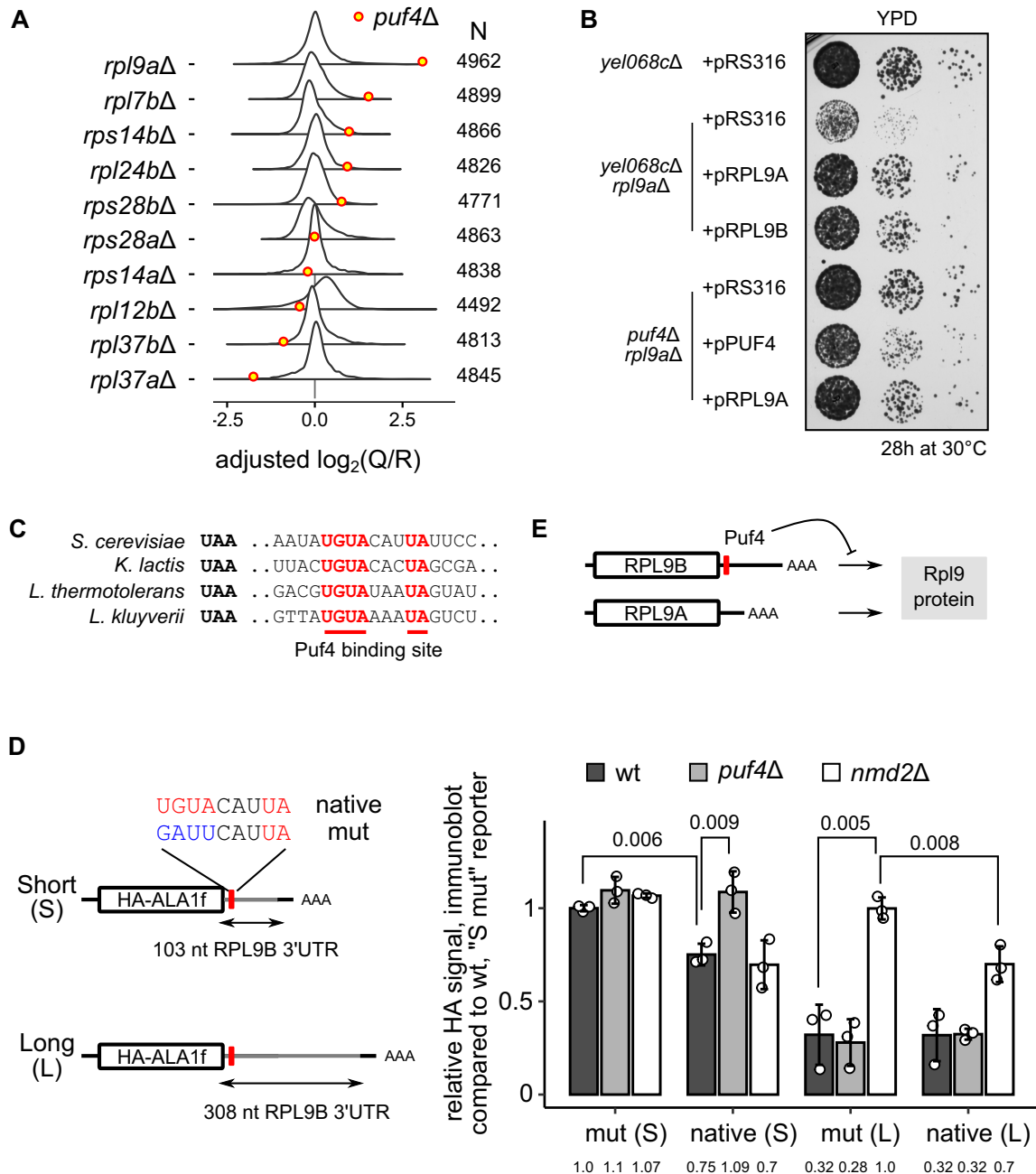


Figure 8. Puf4 represses *RPL9B* function through a conserved UGUAcAU motif. (A) Results of screens performed with query strains having deletions of ribosomal protein genes were ranked in decreasing order of the size of GI with the deletion of *PUF4*, from positive, upper part of the panel, to negative values. The position of the value measured for *PUF4* is indicated by a yellow dot in comparison with the overall distribution of all measured values for each screen (continuous line). (B) Adding a deletion of *PUF4* to a slow-growing *rpl9aΔ* strain improved growth on plates. Growth of the *rpl9aΔ* strain was tested in the presence of centromeric plasmid with *RPL9A* and *RPL9B*. Expression of *PUF4* from a centromeric plasmid was slightly detrimental for growth of the *rpl9aΔ/puf4Δ* strain. YEL068C deletion allowed the presence of equivalent antibiotic resistance markers in the different strains. YPD plates with serial dilutions of cells were grown for 28 hours at 30°C. (C) Puf4 consensus binding motifs, in red, are present and conserved in the 3' UTR of *RPL9B* in *S. cerevisiae* and related yeast species: *Kluyveromyces lactis*, *Lachancea thermotolerans* and *Lachancea kluyverii*. (D) The presence of the conserved Puf4 binding motif in the 3' UTR of *RPL9B* limits the expression of a reporter protein, as measured by immunoblots. Regions of 1 to 103 nucleotides (S) and 1 to 308 nucleotides (L) of the *RPL9B* 3' UTR region were added downstream a coding sequence that consists of a fusion between the HA epitope and a fragment of the *ALA1* coding sequence (nucleotides 4 to 492, the fusion protein is unstable and its levels mirror mRNA translation and stability). The potential Puf4 binding motif, indicated as a red rectangle, was changed from UGUAcAU (native) to GAUUcAU (mut) and the corresponding plasmids were introduced in control (wt), *puf4Δ* and *nmd2Δ* strains. Anti-HA immunoblot results were calibrated using successive binary dilutions of one of the extracts (see Supplementary Figure S10A for a representative immunoblot example). The results represent average and standard deviation corresponding to three independent experiments. Comparisons between conditions were performed using the Welch two sample t-test, and the indicated *P*-values correspond to one-sided difference results. (E) The inhibitory action of Puf4 on *RPL9B* mRNA (binding motif depicted as red rectangle) is revealed in the absence of *RPL9A* mRNA, when the only source of Rpl9 protein is *RPL9B*. Coding sequences are depicted as white rectangles, with 5' UTR and 3' UTR regions drawn as lines to the same relative scale.

To test whether the Puf4 motif present in the 3' UTR of *RPL9B* was required and sufficient for the proposed repression mechanism, we tested the importance of this motif when regions of the *RPL9B* gene were transferred to a reporter system and the motif was modified. The reporter protein was a fusion between the HA epitope and a fragment of the Ala1 yeast protein, that we have previously used to test NMD efficiency (43). The instability of the synthesized HA-ALA1 fragment protein makes the system sensitive to variations in mRNA stability and mRNA translation. We chose to test two different situations for this 3' UTR transfer experiment. In the first, we transferred a short, 103 nucleotides long, region from the *RPL9B* 3' UTR to the reporter gene upstream the *CYC1* terminator of the pCM189 vector (42). This relatively short 3' UTR did not include the hairpin region that is involved in the autoregulation of transcription termination for *RPL9B* (44) and mimics the situation in which Rpl9 is not in excess, leading to an early termination and a relatively short 3' UTR (Short 3' UTR, reporter denoted 'S'). We also tested the insertion of a 308 nucleotides fragment from the region downstream the stop codon of *RPL9B* (Long 3' UTR, reporter denoted 'L'), since this region confers Rpl9-dependent regulation to a reporter system (44). In both cases we built either a version of the 3' UTR with the native Puf4 binding motif or a version in which the first four nucleotides of the motif were shuffled from UGUA (native) to GAUU (mutated, mut). As expected, the levels of HA-Ala1 were lower when the 3' UTR contained the native motif when compared with the mutated motif (Figure 8D). This difference was no longer present if the plasmids were tested in the absence of *PUF4*. The levels of mRNA, as tested by reverse transcription and Q-PCR were not significantly different among these conditions (Supplementary Figure S10).

The longer version of the 3' UTR led to strongly decreased levels of HA-Ala1, compared with the shorter 3' UTR and this difference was fully reversed by inactivating NMD in a strain lacking the core NMD factor *NMD2* (*UPF2*). Interestingly, when the mRNA of the reporter system was stabilized in the absence of Nmd2, the levels of HA-Ala1 followed the same pattern as for the short 3' UTR version of the reporter. The presence of the native Puf4-binding motif led to a relative decrease in the levels of HA-Ala1, in comparison with the mutated version of the motif (Figure 8D). Altogether, these data demonstrate that the identified Puf4-binding motif can modulate *RPL9B* expression and that this effect is mediated by *PUF4*.

The presented results are consistent with a conserved role of Puf4 in limiting *RPL9B* expression under normal conditions (Figure 8E). Thus, the suppression result obtained by large-scale GIM screens allowed to pinpoint a major role for Puf4 in Rpl9 function. Why Rpl9 expression requires at least two different levels of control, in both nucleus, at the level of transcription termination (44) and cytoplasm, at a post-transcriptional level, as shown here, remains unknown.

DISCUSSION

The set of about 700 000 GIs described in our study, together with multiple validations of the obtained results

(Figures 2 and 3), establish a new resource for functional genomics in yeast, along and complementary to previous large-scale GI results (for example, 13, 22). Individual screens performed with temporary transcription repression of query genes also demonstrated the value and flexibility of GIM screens for the study of essential gene function (see, for example the case of *RRP6*, Figure 2D and *RATI*, Figure 4A). One of the advantages of GIM screens is that they do not require any robotic devices. The detection of barcodes, originally done with DNA microarrays can be switched to DNA sequencing, as shown for chemogenomic screens (31). Thus, GIM screens are an alternative to SGA for large scale GI tests. The results presented here identify novel GIs for essential and non-essential genes involved in RNA metabolism in yeast and bring an independent validation for hundreds of previously observed GIs. As demonstrated through the analysis of the correlated GI profiles, this new data set explores a large variety of cellular processes and macromolecular complexes, well beyond the function of the 154 screen query genes (Figure 5).

One of the goals of performing large-scale genetic screens is to establish new functional links between genes and cellular processes. To this end, gene set enrichment analysis (92) can be applied to groups of genes that share similar GI profiles. A refinement of this approach, as implemented in the spatial analysis of functional enrichment (SAFE) method (27), includes in the enrichment analysis the topology of the gene network, built, most frequently, from the similarity of GI profiles. This method has the advantage of providing a map for how various enriched GO terms distribute across network and allows a visually rich inspection of the results. We demonstrate here a complementary approach, called RECAP, that combines gene co-citation links with GI profile similarity and identifies pairs of genes that are both related by the literature data and by the experimental results (overview of the method in Supplementary Figure S7). The inclusion of literature information in the analysis of GI profiles highlights mutants that behave as expected in the GI data set. This selection process validates hundreds of GI profiles and allows the identification of linked genes and their association with well described biological processes or macromolecular complexes.

The originality of RECAP consists in the use of published results to find GI profiles of high confidence, serving as anchoring points to extend the network of functional links. The reason we used co-citation to build the initial functional network is that it suggests links between genes and groups of genes in a manner that is quite natural and independent from the hierarchical gene ontology terms annotations (93,94). Co-citation is used by major gene annotation and protein interaction databases such as STRING (95). However, it has not been used until now for the analysis of GI networks as a yardstick in the initial validation of experimental results.

A potential problem of co-citation is that its quality depends on the availability of a high-quality curated database that associates genes and publications, such as the one maintained by the Saccharomyces Genome Database (28). Since the version of RECAP presented here depends on such manually curated database, extending it to other organisms depends on the presence of equivalent resources. It

is likely that full text mining, such as the one implemented by Textpresso (96) would allow automatic building of such databases for other organisms. Alternatively, association by GO term similarity and by other ways to link genes, such as the network extracted ontology (97), could be also effective as the first step of a RECAP analysis.

RECAP is useful to confidently extend functional interaction networks with GI results (for examples see Figure 5B, C and Supplementary Figure S8). Among the mutants validated by this analysis we were particularly interested in those affecting essential genes through the DAmP modification, since functional interactions, as detected by large-scale genetic screens, are scarcer for essential than for non-essential genes. Many DAmP strains display normal growth rates under standard culture conditions (14,17). However, testing a few conditions might miss specific phenotypes associated with DAmP perturbation of genes. The GIM screens performed in this study are equivalent to testing 154 different stress conditions for each of the tested DAmP strain. Similar to the observation that deletion of most non-essential genes does not affect growth under standard culture conditions, but can be limiting in the presence of a chemical (98), we view the set of GIM screens we performed as a series of highly diverse stress conditions. Thus, the screens probed a panorama of conditions for the 900 DAmP mutant strains and allowed the identification of global trends, such as the correlations of DAmP modification effect with short coding sequence length and high gene expression (Figure 4). Together with the previously published results on DAmP mutants (13,22), the large-scale characterization of these collections of strains is a useful resource for anyone interested in the study of specific essential genes.

In addition to the 154 different tests for each of the ~5000 tested mutants, the presence of translation inhibitors in the GIM screens introduced an additional stress common to all our screens. This stress allowed the identification of GIs that would have been otherwise of much lower amplitude or undetectable. An example is the link between the OCA complex genes and the tRNA export factor *LOSI*, which was weak in SGA results (22), but among the highest-ranking ones in the GIM data (Figure 6A). We validated this link on individual double mutant strains and showed that it is dependent on the availability of the initiator tRNA_i^{Met} (Figures 6 and 7). Since the OCA complex affects inositol polyphosphate metabolism, this result adds a new element in the complex puzzle of the influence of inositol polyphosphates on cellular processes and highlights the usefulness of measuring genetic interactions under a variety of conditions, as previously suggested (20,23).

As a further demonstration of the importance of performing independent experiments for genetic interactions, we validated the discovery that *PUF4* deletion suppresses the slow growth phenotype of cells lacking *RPL9A*. Moreover, we could identify a conserved Puf4 binding motif in the *RPL9B* mRNA that is responsible for this suppression effect (Figure 8). It is not clear what Puf4 does to the *RPL9B* mRNA since previous studies demonstrated that the absence of this protein alone has no measurable impact on the half-life of target mRNAs (99). Moreover, a large-scale study showed that deletion of *PUF4* had no impact

on *RPL9B* steady state mRNA levels (100). It is thus possible that Puf4 inhibits the translation of *RPL9B* mRNA, since interactions between PUF proteins and translation have been documented (101–103). It is interesting that this effect of Puf4 is likely to be prevalent especially when relatively low Rpl9 levels allow early transcription termination at the *RPL9B* locus (44). When Rpl9 is abundant, inefficient transcription termination leads probably to long 3' UTR transcripts that are either retained in the nucleus or become targets of the NMD pathway. Since we did not identify NMD factors deletions as suppressors of *RPL9A* deletion, transcription termination at the *RPL9B* locus probably occurs almost exclusively early in *rpl9a*Δ strains and the major produced form of the *RPL9B* mRNA is not NMD sensitive. Under these conditions post-transcriptional regulation by *PUF4* becomes dominant. The demonstration of a strong phenotypic effect due to *PUF4* deletion alone, that could be traced back to a specific ribosomal protein mRNA, is compatible with the previously proposed role of this RNA binding protein in ribosome biogenesis (90) and provides a new experimental model for further studies on PUF proteins molecular mechanism of action. Such studies might be important for the understanding of the mechanism of action of this conserved family of proteins in yeast and other eukaryotes.

The new GI resource together with RECAP and the associated validation experiments will be useful for further exploration of gene function in yeast and other organisms.

DATA AVAILABILITY

Raw and normalized microarray data were deposited in GEO (GSE119174, 312 samples), and ArrayExpress (E-MTAB-7191, 16 samples). Aggregated, normalized and pleiotropy adjusted results, including correlations of GI profiles can be explored at http://hub05.hosting.pasteur.fr/GIM_interactions/.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Lucia Oreus and Antonia Doyen for their contribution to performing the GIM screens, Thomas Moncion for early work on data analysis, Gwenaël Badis-Breard, Micheline Fromont-Racine and Frank Feuerbach for critically reading of the manuscript and our other colleagues from the GIM laboratory for help, advice, discussions and criticism. We thank Jean-Yves Coppée, Institut Pasteur, for his assistance in microarray hybridization and scanning. We thank Alan Hinnebusch and Thomas Dever (National Institutes of Health, Bethesda, MD, USA) for sharing plasmids and antibodies.

FUNDING

ANR GENO-GIM, CLEANMD and DEFineNMD [ANR-08-JCJC-0019, ANR-14-CE10-0014, ANR-18-CE11-0003] from the French 'Agence Nationale de la

Recherche'; Institut Pasteur and Centre National de Recherche Scientifique, France. Funding for open access charge: Institut Pasteur.

Conflict of interest statement. None declared.

REFERENCES

- Dujon, B. (2010) Yeast evolutionary genomics. *Nat. Rev. Genet.*, **11**, 512–524.
- Marsit, S., Leducq, J.-B., Durand, E., Marchant, A., Filteau, M. and Landry, C.R. (2017) Evolutionary biology through the lens of budding yeast comparative genomics. *Nat. Rev. Genet.*, **18**, 581–598.
- Gu, Z., Steinmetz, L.M., Gu, X., Scharfe, C., Davis, R.W. and Li, W.-H. (2003) Role of duplicate genes in genetic robustness against null mutations. *Nature*, **421**, 63–66.
- Keane, O.M., Toft, C., Carretero-Paulet, L., Jones, G.W. and Fares, M.A. (2014) Preservation of genetic and regulatory robustness in ancient gene duplicates of *Saccharomyces cerevisiae*. *Genome Res.*, **24**, 1830–1841.
- Qian, W. and Zhang, J. (2014) Genomic evidence for adaptation by gene duplication. *Genome Res.*, **24**, 1356–1362.
- Kaiser, C.A. and Schekman, R. (1990) Distinct sets of SEC genes govern transport vesicle formation and fusion early in the secretory pathway. *Cell*, **61**, 723–733.
- Tong, A.H.Y., Lesage, G., Bader, G.D., Ding, H., Xu, H., Xin, X., Young, J., Berriz, G.F., Brost, R.L., Chang, M. *et al.* (2004) Global mapping of the yeast genetic interaction network. *Science*, **303**, 808–813.
- Decourty, L., Saveanu, C., Zemam, K., Hantraye, F., Frachon, E., Rousselle, J.-C., Fromont-Racine, M. and Jacquier, A. (2008) Linking functionally related genes by sensitive and quantitative characterization of genetic interaction profiles. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 5821–5826.
- Pan, X., Yuan, D.S., Xiang, D., Wang, X., Sookhai-Mahadeo, S., Bader, J.S., Hieter, P., Spencer, F. and Boeke, J.D. (2004) A robust toolkit for functional profiling of the yeast genome. *Mol. Cell*, **16**, 487–496.
- Schuldiner, M., Collins, S.R., Thompson, N.J., Denic, V., Bhamidipati, A., Punna, T., Ihmels, J., Andrews, B., Boone, C., Greenblatt, J.F. *et al.* (2005) Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. *Cell*, **123**, 507–519.
- Dixon, S.J., Costanzo, M., Baryshnikova, A., Andrews, B. and Boone, C. (2009) Systematic mapping of genetic interaction networks. *Annu. Rev. Genet.*, **43**, 601–625.
- Costanzo, M., Kuzmin, E., van Leeuwen, J., Mair, B., Moffat, J., Boone, C. and Andrews, B. (2019) Global genetic networks and the genotype-to-phenotype relationship. *Cell*, **177**, 85–100.
- Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L.Y., Toufighi, K., Mostafavi, S. *et al.* (2010) The genetic landscape of a cell. *Science*, **327**, 425–431.
- Breslow, D.K., Cameron, D.M., Collins, S.R., Schuldiner, M., Stewart-Ornstein, J., Newman, H.W., Braun, S., Madhani, H.D., Krogan, N.J. and Weissman, J.S. (2008) A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nat. Methods*, **5**, 711–718.
- Milligan, L., Decourty, L., Saveanu, C., Rappilber, J., Ceulemans, H., Jacquier, A. and Tollervy, D. (2008) A yeast exosome cofactor, Mpp6, functions in RNA surveillance and in the degradation of noncoding RNA transcripts. *Mol. Cell. Biol.*, **28**, 5446–5457.
- Kuzmin, E., VanderSluis, B., Ba, A.N.N., Wang, W., Koch, E.N., Usaj, M., Khmelinskii, A., Usaj, M.M., Leeuwen, J., Kraus, O. *et al.* (2020) Exploring whole-genome duplicate gene retention with complex genetic interaction analysis. *Science*, **368**, eaaz5667.
- Yan, Z., Costanzo, M., Heisler, L.E., Paw, J., Kaper, F., Andrews, B.J., Boone, C., Giaever, G. and Nislow, C. (2008) Yeast Barcoders: a chemogenomic application of a universal donor-strain collection carrying bar-code identifiers. *Nat. Methods*, **5**, 719–725.
- Decourty, L., Doyen, A., Malabat, C., Frachon, E., Rispal, D., Séraphin, B., Feuerbach, F., Jacquier, A. and Saveanu, C. (2014) Long open reading frame transcripts escape nonsense-mediated mRNA decay in yeast. *Cell Rep*, **6**, 593–598.
- Smith, J.D., Schlecht, U., Xu, W., Suresh, S., Horecka, J., Proctor, M.J., Aiyar, R.S., Bennett, R.A.O., Chu, A., Li, Y.F. *et al.* (2017) A method for high-throughput production of sequence-verified DNA libraries and strain collections. *Mol. Syst. Biol.*, **13**, 913.
- Jaffe, M., Dziulko, A., Smith, J.D., St Onge, R.P., Levy, S.F. and Sherlock, G. (2019) Improved discovery of genetic interactions using CRISPRiSeq across multiple environments. *Genome Res.*, **29**, 668–681.
- Michel, A.H., Hatakeyama, R., Kimmig, P., Arter, M., Peter, M., Matos, J., De Virgilio, C. and Kornmann, B. (2017) Functional mapping of yeast genomes by saturated transposition. *Elife*, **6**, e23570.
- Costanzo, M., VanderSluis, B., Koch, E.N., Baryshnikova, A., Pons, C., Tan, G., Wang, W., Usaj, M., Hanchard, J., Lee, S.D. *et al.* (2016) A global genetic interaction network maps a wiring diagram of cellular function. *Science*, **353**, aaf1420.
- Martin, H., Shales, M., Fernandez-Piñar, P., Wei, P., Molina, M., Fiedler, D., Shokat, K.M., Beltrao, P., Lim, W. and Krogan, N.J. (2015) Differential genetic interactions of yeast stress response MAPK pathways. *Mol. Syst. Biol.*, **11**, 800.
- St Onge, R.P., Mani, R., Oh, J., Proctor, M., Fung, E., Davis, R.W., Nislow, C., Roth, F.P. and Giaever, G. (2007) Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. *Nat. Genet.*, **39**, 199–206.
- Brown, J.T., Bai, X. and Johnson, A.W. (2000) The yeast antiviral proteins Ski2p, Ski3p, and Ski8p exist as a complex in vivo. *RNA*, **6**, 449–457.
- Defenouillère, Q., Yao, Y., Mouaikel, J., Namane, A., Galopier, A., Decourty, L., Doyen, A., Malabat, C., Saveanu, C., Jacquier, A. *et al.* (2013) Cdc48-associated complex bound to 60S particles is required for the clearance of aberrant translation products. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 5046–5051.
- Baryshnikova, A. (2016) Systematic functional annotation and visualization of biological networks. *Cell Syst*, **2**, 412–421.
- Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R. *et al.* (2012) Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res.*, **40**, D700–D705.
- Malabat, C. and Saveanu, C. (2016) Identification of links between cellular pathways by genetic interaction mapping (GIM). *Methods Mol. Biol.*, **1361**, 325–343.
- Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B. *et al.* (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature*, **418**, 387–391.
- Smith, A.M., Heisler, L.E., Mellor, J., Kaper, F., Thompson, M.J., Chee, M., Roth, F.P., Giaever, G. and Nislow, C. (2009) Quantitative phenotyping via deep barcode sequencing. *Genome Res.*, **19**, 1836–1842.
- R Core Team (2019) In: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Leeuwen, J., Pons, C., Mellor, J.C., Yamaguchi, T.N., Friesen, H., Koschwanez, J., Ušaj, M.M., Pechlaner, M., Takar, M., Ušaj, M. *et al.* (2016) Exploring genetic suppression interactions on a global scale. *Science*, **354**, aag0839.
- Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H. and Vilo, J. (2019) g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.*, **47**, W191–W198.
- Csardi, G. and Nepusz, T. (2006) The igraph software package for complex network research. *InterJournal*, 1695.
- Gustavsen, J.A., Pai, S., Isserlin, R., Demchak, B. and Pico, A.R. (2019) RCy3: network biology using cytoscape from within R. *F1000Res*, **8**, 1774.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.
- Belli, G., Gari, E., Aldea, M. and Herrero, E. (1998) Functional analysis of yeast essential genes using a promoter-substitution cassette and the tetracycline-regulatable dual expression system. *Yeast*, **14**, 1127–1138.

39. Torchet,C., Badis,G., Devaux,F., Costanzo,G., Werner,M. and Jacquier,A. (2005) The complete set of H/ACA snoRNAs that guide rRNA pseudouridylations in *Saccharomyces cerevisiae*. *RNA*, **11**, 928–938.
40. Longtine,M.S., McKenzie,A., Demarini,D.J., Shah,N.G., Wach,A., Brachat,A., Philippsen,P. and Pringle,J.R. (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast*, **14**, 953–961.
41. Fu,C., Donovan,W.P., Shikapwashya-Hasser,O., Ye,X. and Cole,R.H. (2014) Hot fusion: an efficient method to clone multiple DNA fragments as well as inverted repeats without ligase. *PLoS ONE*, **9**, e115318.
42. Gari,E., Piedrafita,L., Aldea,M. and Herrero,E. (1997) A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in *Saccharomyces cerevisiae*. *Yeast*, **13**, 837–848.
43. Dehecq,M., Decourty,L., Namane,A., Proux,C., Kanaan,J., Le Hir,H., Jacquier,A. and Saveanu,C. (2018) Nonsense-mediated mRNA decay involves two distinct Upf1-bound complexes. *EMBO J.*, **37**, e99278.
44. Gudipati,R.K., Neil,H., Feuerbach,F., Malabat,C. and Jacquier,A. (2012) The yeast RPL9B gene is regulated by modulation between two modes of transcription termination. *EMBO J.*, **31**, 2427–2437.
45. Collart,M.A. and Oliviero,S. (2001) Preparation of yeast RNA. *Curr. Protoc. Mol Biol*, **Chapter 13**, Unit13.12.
46. Kiss,T. (2002) Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell*, **109**, 145–148.
47. Wishart,J.A., Osborn,M., Gent,M.E., Yen,K., Vujovic,Z., Gitsham,P., Zhang,N., Ross Miller,J. and Oliver,S.G. (2006) The relative merits of the tetO2 and tetO7 promoter systems for the functional analysis of heterologous genes in yeast and a compilation of essential yeast genes with tetO2 promoter substitutions. *Yeast*, **23**, 325–331.
48. Baryshnikova,A., VanderSluis,B., Costanzo,M., Myers,C.L., Cha,R.S., Andrews,B. and Boone,C. (2013) Global linkage map connects meiotic centromere function to chromosome size in budding yeast. *G3 (Bethesda)*, **3**, 1741–1751.
49. Mattiazzi Usaj,M., Sahin,N., Friesen,H., Pons,C., Usaj,M., Masinas,M.P.D., Shuteriqi,E., Shkurin,A., Aloy,P., Morris,Q. *et al.* (2020) Systematic genetics and single-cell imaging reveal widespread morphological pleiotropy and cell-to-cell variability. *Mol. Syst. Biol.*, **16**, e9243.
50. Boguta,M. (2013) Maf1, a general negative regulator of RNA polymerase III in yeast. *Biochim. Biophys. Acta*, **1829**, 376–384.
51. Hopper,A.K. (2013) Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast *Saccharomyces cerevisiae*. *Genetics*, **194**, 43–67.
52. Fox,M.J. and Mosley,A.L. (2016) Rrp6: Integrated roles in nuclear RNA metabolism and transcription termination. *Wiley Interdiscip. Rev. RNA*, **7**, 91–104.
53. Schmidt,A., Hall,M.N. and Koller,A. (1994) Two FK506 resistance-conferring genes in *Saccharomyces cerevisiae*, TAT1 and TAT2, encode amino acid permeases mediating tyrosine and tryptophan uptake. *Mol. Cell Biol.*, **14**, 6597–6606.
54. van Hoof,A., Lennertz,P. and Parker,R. (2000) Three conserved members of the RNase D family have unique and overlapping functions in the processing of 5S, 5.8S, U4, U5, RNase MRP and RNase P RNAs in yeast. *EMBO J.*, **19**, 1357–1365.
55. Grover,R., Burse,S.A., Shankrit,S., Aggarwal,A., Kirty,K., Narta,K., Srivastav,R., Ray,A.K., Malik,G., Vats,A. *et al.* (2019) Myg1 exonuclease couples the nuclear and mitochondrial translational programs through RNA processing. *Nucleic Acids Res.*, **47**, 5852–5866.
56. Rosado,I.V., Dez,C., Lebaron,S., Caizergues-Ferrer,M., Henry,Y. and de la Cruz,J. (2007) Characterization of *Saccharomyces cerevisiae* Npa2p (Urb2p) reveals a low-molecular-mass complex containing Dbp6p, Npa1p (Urb1p), Nop8p, and Rsa3p involved in early steps of 60S ribosomal subunit biogenesis. *Mol. Cell Biol.*, **27**, 1207–1221.
57. Wade,C.H., Umbarger,M.A. and McAlear,M.A. (2006) The budding yeast rRNA and ribosome biosynthesis (RRB) regulon contains over 200 genes. *Yeast*, **23**, 293–306.
58. Durairaj,G. and Kaiser,P. (2014) The 26S proteasome and initiation of gene transcription. *Biomolecules*, **4**, 827–847.
59. Kim,M., Krogan,N.J., Vasiljeva,L., Rando,O.J., Nedeia,E., Greenblatt,J.F. and Buratowski,S. (2004) The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature*, **432**, 517–522.
60. Tye,B.W., Commins,N., Ryazanova,L.V., Wühr,M., Springer,M., Pincus,D. and Churchman,L.S. (2019) Proteotoxicity from aberrant ribosome biogenesis compromises cell fitness. *Elife*, **8**, e43002.
61. Albert,B., Kos-Braun,I.C., Henras,A.K., Dez,C., Rueda,M.P., Zhang,X., Gadal,O., Kos,M. and Shore,D. (2019) A ribosome assembly stress response regulates transcription to maintain proteome homeostasis. *Elife*, **8**, e45002.
62. Zhang,E., Khanna,V., Dacheux,E., Namane,A., Doyen,A., Gomard,M., Turcotte,B., Jacquier,A. and Fromont-Racine,M. (2019) A specialised SKI complex assists the cytoplasmic RNA exosome in the absence of direct association with ribosomes. *EMBO J.*, **38**, e100640.
63. Blondel,V.D., Guillaume,J.-L., Lambiotte,R. and Lefebvre,E. (2008) Fast unfolding of communities in large networks. *J. Stat. Mech.*, **2008**, P10008.
64. Atias,N., Kupiec,M. and Sharan,R. (2016) Systematic identification and correction of annotation errors in the genetic interaction map of *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **44**, e50.
65. Yamagishi,Y., Sakuno,T., Goto,Y. and Watanabe,Y. (2014) Kinetochores composition and its function: lessons from yeasts. *FEMS Microbiol. Rev.*, **38**, 185–200.
66. De Wulf,P., McAinsh,A.D. and Sorger,P.K. (2003) Hierarchical assembly of the budding yeast kinetochore from multiple subcomplexes. *Genes Dev.*, **17**, 2902–2921.
67. Lanz,M.C., Dibitetto,D. and Smolka,M.B. (2019) DNA damage kinase signaling: checkpoint and repair at 30 years. *EMBO J.*, **38**, e101801.
68. Chua,P.R. and Roeder,G.S. (1998) Zip2, a meiosis-specific protein required for the initiation of chromosome synapsis. *Cell*, **93**, 349–359.
69. Pu,S., Wong,J., Turner,B., Cho,E. and Wodak,S.J. (2009) Up-to-date catalogues of yeast protein complexes. *Nucleic Acids Res.*, **37**, 825–831.
70. Alic,N., Higgins,V.J. and Dawes,I.W. (2001) Identification of a *Saccharomyces cerevisiae* gene that is required for G1 arrest in response to the lipid oxidation product linoleic acid hydroperoxide. *Mol. Biol. Cell*, **12**, 1801–1810.
71. Wishart,M.J. and Dixon,J.E. (1998) Gathering STYX: phosphatase-like form predicts functions for unique protein-interaction domains. *Trends Biochem. Sci.*, **23**, 301–306.
72. Romá-Mateo,C., Sacristán-Reviriego,A., Beresford,N.J., Caparrós-Martín,J.A., Culiáñez-Maciá,F.A., Martín,H., Molina,M., Tabernero,L. and Pulido,R. (2011) Phylogenetic and genetic linkage between novel atypical dual-specificity phosphatases from non-metazoan organisms. *Mol. Genet. Genomics*, **285**, 341–354.
73. Steidle,E.A., Chong,L.S., Wu,M., Crooke,E., Fiedler,D., Resnick,A.C. and Rolfes,R.J. (2016) A Novel Inositol Pyrophosphate Phosphatase in *Saccharomyces cerevisiae*: Siw14 protein selectively cleaves the β -phosphate from 5-diphosphoinositol pentakisphosphate (SPP-IP5). *J. Biol. Chem.*, **291**, 6772–6783.
74. Kushner,D.B., Lindenbach,B.D., Grdzlishvili,V.Z., Noueiry,A.O., Paul,S.M. and Ahlquist,P. (2003) Systematic, genome-wide identification of host genes affecting replication of a positive-strand RNA virus. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 15764–15769.
75. Wickner,R.B., Edskes,H.K., Bezsonov,E.E., Son,M. and Ducatez,M. (2018) Prion propagation and inositol polyphosphates. *Curr. Genet.*, **64**, 571–574.
76. Hoepfner,D., Helliwell,S.B., Sadlish,H., Schuierer,S., Filipuzzi,I., Brachat,S., Bhullar,B., Plikat,U., Abraham,Y., Altorfer,M. *et al.* (2014) High-resolution chemical dissection of a model eukaryote reveals targets, pathways and gene functions. *Microbiol. Res.*, **169**, 107–120.
77. Oughtred,R., Stark,C., Breitkreutz,B.-J., Rust,J., Boucher,L., Chang,C., Kolas,N., O'Donnell,L., Leung,G., McAdam,R. *et al.* (2019) The BioGRID interaction database: 2019 update. *Nucleic Acids Res.*, **47**, D529–D541.
78. Aström,S.U. and Byström,A.S. (1994) Rit1, a tRNA backbone-modifying enzyme that mediates initiator and elongator tRNA discrimination. *Cell*, **79**, 535–546.

79. Ho, C.H., Magtanong, L., Barker, S.L., Gresham, D., Nishimura, S., Natarajan, P., Koh, J.L.Y., Porter, J., Gray, C.A., Andersen, R.J. *et al.* (2009) A molecular barcoded yeast ORF library enables mode-of-action analysis of bioactive compounds. *Nat. Biotechnol.*, **27**, 369–377.
80. Hughes, T.R., Marton, M.J., Jones, A.R., Roberts, C.J., Stoughton, R., Armour, C.D., Bennett, H.A., Coffey, E., Dai, H., He, Y.D. *et al.* (2000) Functional discovery via a compendium of expression profiles. *Cell*, **102**, 109–126.
81. Hinnebusch, A.G. (2005) Translational regulation of GCN4 and the general amino acid control of yeast. *Annu. Rev. Microbiol.*, **59**, 407–450.
82. Hinnebusch, A.G., Ivanov, I.P. and Sonenberg, N. (2016) Translational control by 5'-untranslated regions of eukaryotic mRNAs. *Science*, **352**, 1413–1416.
83. Natarajan, K., Meyer, M.R., Jackson, B.M., Slade, D., Roberts, C., Hinnebusch, A.G. and Marton, M.J. (2001) Transcriptional profiling shows that Gcn4p is a master regulator of gene expression during amino acid starvation in yeast. *Mol Cell Biol*, **21**, 4347–4368.
84. Mueller, P.P. and Hinnebusch, A.G. (1986) Multiple upstream AUG codons mediate translational control of GCN4. *Cell*, **45**, 201–207.
85. Miller, M.A. and Olivas, W.M. (2011) Roles of Puf proteins in mRNA degradation and translation. *Wiley Interdiscip. Rev. RNA*, **2**, 471–492.
86. Quenault, T., Lithgow, T. and Traven, A. (2011) PUF proteins: repression, activation and mRNA localization. *Trends Cell Biol.*, **21**, 104–112.
87. Lapointe, C.P., Preston, M.A., Wilinski, D., Saunders, H.A.J., Campbell, Z.T. and Wickens, M. (2017) Architecture and dynamics of overlapped RNA regulatory networks. *RNA*, **23**, 1636–1647.
88. Ghulam, M.M., Catala, M. and Abou Elela, S. (2020) Differential expression of duplicated ribosomal protein genes modifies ribosome composition in response to stress. *Nucleic Acids Res.*, **48**, 1954–1968.
89. Peyroche, G., Saveanu, C., Dauplais, M., Lazard, M., Beuneu, F., Decourty, L., Malabat, C., Jacquier, A., Blanquet, S. and Plateau, P. (2012) Sodium selenide toxicity is mediated by O₂-dependent DNA breaks. *PLoS ONE*, **7**, e36343.
90. Gerber, A.P., Herschlag, D. and Brown, P.O. (2004) Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast. *PLoS Biol.*, **2**, E79.
91. Hogan, G.J., Brown, P.O. and Herschlag, D. (2015) Evolutionary Conservation and Diversification of Puf RNA Binding Proteins and Their mRNA Targets. *PLoS Biol.*, **13**, e1002307.
92. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *PNAS*, **102**, 15545–15550.
93. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T. *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.
94. The Gene Ontology Consortium (2019) The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.*, **47**, D330–D338.
95. Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P. *et al.* (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.*, **45**, D362–D368.
96. Müller, H.-M., Van Auken, K.M., Li, Y. and Sternberg, P.W. (2018) Textpresso Central: a customizable platform for searching, text mining, viewing, and curating biomedical literature. *BMC Bioinformatics*, **19**, 94.
97. Dutkowski, J., Kramer, M., Surma, M.A., Balakrishnan, R., Cherry, J.M., Krogan, N.J. and Ideker, T. (2013) A gene ontology inferred from molecular networks. *Nat. Biotechnol.*, **31**, 38–45.
98. Hillenmeyer, M.E., Fung, E., Wildenhain, J., Pierce, S.E., Hoon, S., Lee, W., Proctor, M., St Onge, R.P., Tyers, M., Koller, D. *et al.* (2008) The chemical genomic portrait of yeast: uncovering a phenotype for all genes. *Science*, **320**, 362–365.
99. Fischer, A.D. and Olivas, W.M. (2018) Multiple Puf proteins regulate the stability of ribosome biogenesis transcripts. *RNA Biol*, **15**, 1228–1243.
100. Kemmeren, P., Sameith, K., van de Pasch, L.A.L., Benschop, J.J., Lenstra, T.L., Margaritis, T., O'Duibhir, E., Apweiler, E., van Wageningen, S., Ko, C.W. *et al.* (2014) Large-scale genetic perturbations reveal regulatory networks and an abundance of gene-specific repressors. *Cell*, **157**, 740–752.
101. Wang, Z., Sun, X., Wee, J., Guo, X. and Gu, Z. (2019) Novel insights into global translational regulation through Pumilio family RNA-binding protein Puf3p revealed by ribosomal profiling. *Curr Genet*, **65**, 201–212.
102. Chritton, J.J. and Wickens, M. (2010) Translational repression by PUF proteins in vitro. *RNA*, **16**, 1217–1225.
103. Cridge, A.G., Castelli, L.M., Smirnova, J.B., Selley, J.N., Rowe, W., Hubbard, S.J., McCarthy, J.E.G., Ashe, M.P., Grant, C.M. and Pavitt, G.D. (2010) Identifying eIF4E-binding protein translationally-controlled transcripts reveals links to mRNAs bound by specific PUF proteins. *Nucleic Acids Res.*, **38**, 8039–8050.
104. Lipson, D., Raz, T., Kieu, A., Jones, D.R., Giladi, E., Thayer, E., Thompson, J.F., Letovsky, S., Milos, P. and Causey, M. (2009) Quantification of the yeast transcriptome by single-molecule sequencing. *Nat. Biotechnol.*, **27**, 652–658.
105. Dever, T.E., Yang, W., Aström, S., Byström, A.S. and Hinnebusch, A.G. (1995) Modulation of tRNA(iMet), eIF-2, and eIF-2B expression shows that GCN4 translation is inversely coupled to the level of eIF-2.GTP.Met-tRNA(iMet) ternary complexes. *Mol. Cell. Biol.*, **15**, 6351–6363.