



Expression attenuation as a mechanism of robustness against gene duplication

Diana Ascencio^{a,b,c,d,e,f}, Guillaume Diss^{a,b,c,e,1}, Isabelle Gagnon-Arsenault^{a,b,c,d,e}, Alexandre K. Dubé^{a,b,c,d,e}, Alexander DeLuna^f, and Christian R. Landry^{a,b,c,d,e,2}

^aRegroupement Québécois de Recherche sur la Fonction, l'Ingénierie et les Applications des Protéines, Québec, QC G1V 0A6, Canada; ^bInstitut de Biologie Intégrative et des Systèmes, Université Laval, Québec, QC G1V 0A6, Canada; ^cCentre de Recherche en Données Massives de l'Université Laval, Université Laval, Québec, QC G1V 0A6, Canada; ^dDépartement de Biochimie, de Microbiologie et de Bio-informatique, Université Laval, Québec, QC G1V 0A6, Canada; ^eDépartement de Biologie, Université Laval, Québec, QC G1V 0A6, Canada; and ^fUnidad de Genómica Avanzada (Langebio), Centro de Investigación y de Estudios Avanzados, 36824 Irapuato, Guanajuato, Mexico

Edited by Michael Lynch, Arizona State University, Tempe, AZ, and approved December 24, 2020 (received for review July 10, 2020)

Gene duplication is ubiquitous and a major driver of phenotypic diversity across the tree of life, but its immediate consequences are not fully understood. Deleterious effects would decrease the probability of retention of duplicates and prevent their contribution to long-term evolution. One possible detrimental effect of duplication is the perturbation of the stoichiometry of protein complexes. Here, we measured the fitness effects of the duplication of 899 essential genes in the budding yeast using high-resolution competition assays. At least 10% of genes caused a fitness disadvantage when duplicated. Intriguingly, the duplication of most protein complex subunits had small to nondetectable effects on fitness, with few exceptions. We selected four complexes with subunits that had an impact on fitness when duplicated and measured the impact of individual gene duplications on their protein–protein interactions. We found that very few duplications affect both fitness and interactions. Furthermore, large complexes such as the 26S proteasome are protected from gene duplication by attenuation of protein abundance. Regulatory mechanisms that maintain the stoichiometric balance of protein complexes may protect from the immediate effects of gene duplication. Our results show that a better understanding of protein regulation and assembly in complexes is required for the refinement of current models of gene duplication.

gene duplication | gene expression | protein interaction | dosage balance hypothesis | fitness effects

Gene duplication and divergence are a primary source of functional innovation and diversity. During the last few decades, the long-term maintenance of gene duplicates through the gain or reciprocal loss of function has been studied extensively (1–4). However, we know relatively little about the immediate impact of duplications, which may have significant consequences on the preservation of paralogs. Genes with adverse effects on fitness upon duplication would have a reduced residence time in populations, thereby limiting their contribution to long-term evolution. For instance, even modest changes in gene dosage such as those caused by duplication sometimes produce significant phenotypic effects, both positive and negative (5–8). This is commonly referred to as dosage sensitivity. Understanding the immediate impact of duplication is, therefore, of paramount importance.

Several mechanisms have been suggested to explain dosage sensitivity. These include concentration dependency, promiscuous off-target interactions at high concentration, and dosage imbalance (8). The gene dosage balance hypothesis predicts that the single-gene duplication of protein complex subunits is harmful, as this can lead to an immediate stoichiometric imbalance with the rest of the subunits (8–10). There is indirect evidence supporting such a prediction: Complex subunits are less likely to be retained after small-scale duplication (SSD), are often coexpressed at similar levels, and are enriched among genes that reduce fitness when underexpressed (10, 11). For

instance, haploinsufficiency, a dominant phenotype in diploid organisms that are heterozygous for a loss-of-function allele, is more common among genes that encode protein complex subunits (12). However, other works have shown that genes that are toxic when overexpressed are not enriched as part of protein complexes (13, 14). This suggests that overexpression can be deleterious for reasons unrelated to complex stoichiometry.

While gene deletion and overexpression experiments inform us of how cells respond to lowered or increased abundance of proteins, they are fundamentally different from a naturally occurring duplication. For instance, the use of nonnative promoters and multicopy plasmids may cause the assayed genes to be overexpressed severalfold and may also alter the timing of expression. In addition, if dosage–fitness relationships are nonlinear, results from overexpression cannot be interpolated to gene duplication. For these reasons, and because fitness rather than complex assembly has been assayed in previous experiments (14, 15), we do not know what the impact of duplication on the assembly of protein complexes is. Experiments aimed at measuring the fitness benefits of increased gene dosage have been performed (7) but have not addressed how such dosage changes impact protein complex formation.

One reason why the overexpression of protein complex subunits can be less detrimental than a reduction of expression is

Significance

Many studies have focused on the mechanisms of long-term retention of gene duplicates, such as the gain of functions or reciprocal losses. However, such changes are more likely to occur if the duplicates are maintained for a long period. This time span will be short if duplication is immediately deleterious. We measured the distribution of fitness effects of gene duplication for 899 genes in budding yeast. We find that gene duplication is more likely to be deleterious than beneficial. However, contrary to previous models, in general, gene duplication does not affect fitness by altering the organization of protein complexes. We show that expression attenuation may protect complexes from the effects of gene duplication.

Author contributions: D.A., G.D., A.D., and C.R.L. designed research; D.A., G.D., I.G.-A., and A.K.D. performed research; A.D. contributed new reagents/analytic tools; D.A., G.D., and I.G.-A. analyzed data; and D.A. and C.R.L. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹Present address: Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland.

²To whom correspondence may be addressed. Email: christian.landry@bio.ulaval.ca.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2014345118/-/DCSupplemental>.

Published February 1, 2021.

dosage regulation (13). The correct dosage of protein subunits appears to be tightly regulated by the cell. For instance, members of multiprotein complexes are produced in precise proportion to their stoichiometry in both bacteria (16) and eukaryotes (17). Additionally, recent studies revealed that the abundance of members of multiprotein complexes is often attenuated or buffered in aneuploids (strains with extra copies of one or several chromosomes) (11, 18, 19). A study by Dephoure et al. (18) suggested that attenuation may be quite common since nearly 20% of the proteome is attenuated in aneuploids with more copy numbers. Moreover, attenuated genes (60 to 76%) are members of multiprotein complexes (18). In fact, Chen et al. (19) found that up to 50% of subunits with imbalanced gene copy numbers (compared with the rest of the complex) may be attenuated to normal protein abundance levels. Since attenuated genes often have mRNA transcript levels and protein synthesis rates proportional to their gene-copy number, the regulation of attenuated genes most often, but not exclusively, occurs posttranslationally (17, 18, 20). However, aneuploid cells may not be the best models to study the effect of small-scale gene duplications because a large proportion of their genome is duplicated at once. In addition, aneuploid cells often experience systemic effects like proteotoxic stress (21), which may cause pleiotropic consequences on protein synthesis and degradation rates that are challenging to disentangle from those of the duplication of a single gene. Furthermore, when a complete chromosome is duplicated, more than one subunit—or the complete set of subunits of a complex—may be duplicated, which could lessen or even prevent the effects of stoichiometric imbalance. Therefore, the extent and the nature of attenuation after SSD events are yet to be explored.

In this work, we sought to measure the immediate impact of gene duplication by experimentally simulating gene duplication of essential genes in yeast. We focused on this set because genes that are essential for growth are enriched in protein complexes (10). Genes that are essential are also less likely to be duplicated (22), which means that additional copies in the genome will not confound the results of duplication. We measured the fitness consequences of individual gene duplication for nearly 900 genes in individual strain competitions. Duplication of protein complex subunits is not more deleterious on average than that of other genes. We therefore also measured changes in protein–protein interactions (PPIs) in response to gene duplication for a subset of essential genes that are part of four large protein complexes. We finally estimated the extent to which the expression of proteasome subunits is attenuated as a response to gene duplication. Our results show that even though gene duplication often affects fitness, it has a small effect on the assembly of protein complexes. The apparent robustness of multiprotein complexes to gene duplication is likely to be a consequence of expression attenuation.

Results

An Important Percentage of Yeast Essential Genes Affect Fitness When Duplicated. We measured the fitness effect of gene duplication using high-resolution competition assays with fluorescently labeled cells in coculture (Fig. 1A) (23). We individually duplicated 899 essential genes using single-copy centromeric plasmids (pCEN) expressing the genes under their native promoter and 3' untranslated region (24). In parallel, we generated a distribution of control strains in which we competed a wild-type (*WT*) strain (also used as a reference) with itself. Each of the 192 replicates of this control set is an independent colony from a transformation.

We chose a conservative threshold of at least 1% of fitness effect ($-0.01 > s > 0.01$), that corresponds to a $|z \text{ score}| > 4.5$. For most genes (86%), duplication has little or no effect on relative fitness. However, around 9% (Fig. 1B) of the duplications have moderate to strong deleterious effects ($s < -0.01$, z

score < -4.5), while 4% have beneficial ($s > 0.01$, $z \text{ score} > 4.5$) but often modest effects. We validated the top 180 genes (top 180 absolute z scores) using the same growth conditions but monitoring the two populations by flow cytometry. We generated the strains *de novo* and tested three biological replicates per sample. While less scalable, this approach allows for a more accurate estimation of population ratios since fluorescence is measured at the single-cell level instead of the population level. There is a strong correlation between the two assays (Spearman's $r = 0.83$, $P < 1e-15$; Fig. 1C), giving us good confidence in the accuracy of our large-scale measurements. We validated most of the effects of the first competition assay (159/180; Fig. 1C). We compared our results with two other studies evaluating the relative fitness of strains harboring the same plasmids through pooled assays (7, 25). We found a weak but significant correlation between the selection coefficient (s) from these studies and ours (*SI Appendix*, Fig. S1). Although these two previous studies have used a similar pooling approach, they are only weakly correlated with each other ($r = 0.15$, $P < 1e-7$; *SI Appendix*, Fig. S1B). The weak correlations could come from the different media used (SC[MSG]-ura versus rich media, or nutrient limitation). Furthermore, pool assays create a complex and distinct environment compared with pairwise competition assays and may lead to noisier estimates, for instance, for strains that are rare in the pools.

The 4% of the duplications that have beneficial effects above 1% include genes such as *CDC25* and *IRA1* (Fig. 1B and C), which regulate the Ras-cAMP pathway (26, 27). In yeasts, growth and metabolism in response to nutrients, particularly glucose, are regulated to a large degree by this pathway. Since we used glucose as a carbon source, duplication of these two genes may modify the activity of the pathway in a way that increases growth rate. Adaptation in limited glucose conditions often involves mutations in these pathways (28). The duplication of some genes that encode central metabolism enzymes were also beneficial. For instance, duplication of *HIP1*, a histidine transporter (29), and *PDC2*, a pyruvate decarboxylase (30), may result in increased growth rate through metabolic activity in a similar manner to duplication of genes in the Ras-cAMP pathway.

The presence of beneficial duplications may appear surprising since the yeast lineage has a high rate of duplication (31) and has undergone a whole-genome duplication (WGD) (32), which would have provided the mutational input needed to fix any beneficial duplication. One possible reason for the lack of duplication of the genes we detect with beneficial effects is that the adaptive value of some gene duplications is highly dependent on environmental conditions and can even become deleterious in specific contexts (33). Such antagonistic pleiotropy has been observed in the study of aneuploidies (34) and of gene deletions (35). We, therefore, performed a parallel competition experiment in a condition of salt stress. We find a significant correlation in the selection coefficient between conditions (Spearman's $r = 0.48$, $P < 1e-15$), and a similar number of deleterious (8.5%) and beneficial (4.6%) effects (*SI Appendix*, Fig. S2). Many of the effects may, therefore, be general, while others are condition specific. The overlap in the identity of the deleterious genes between conditions is greater than for advantageous ones, suggesting that beneficial effects are more condition specific than deleterious ones. To further validate the condition specificity of the benefit of gene duplication, we measured the relative fitness of five strongly beneficial and three strongly deleterious duplications in five different conditions. We observed antagonistic pleiotropy for some beneficial duplications (*SI Appendix*, Fig. S3). For instance, *CDC25* and *IRA1* were beneficial in the standard condition, osmotic stress and with galactose as carbon source, but are strongly deleterious in 6% of ethanol, while

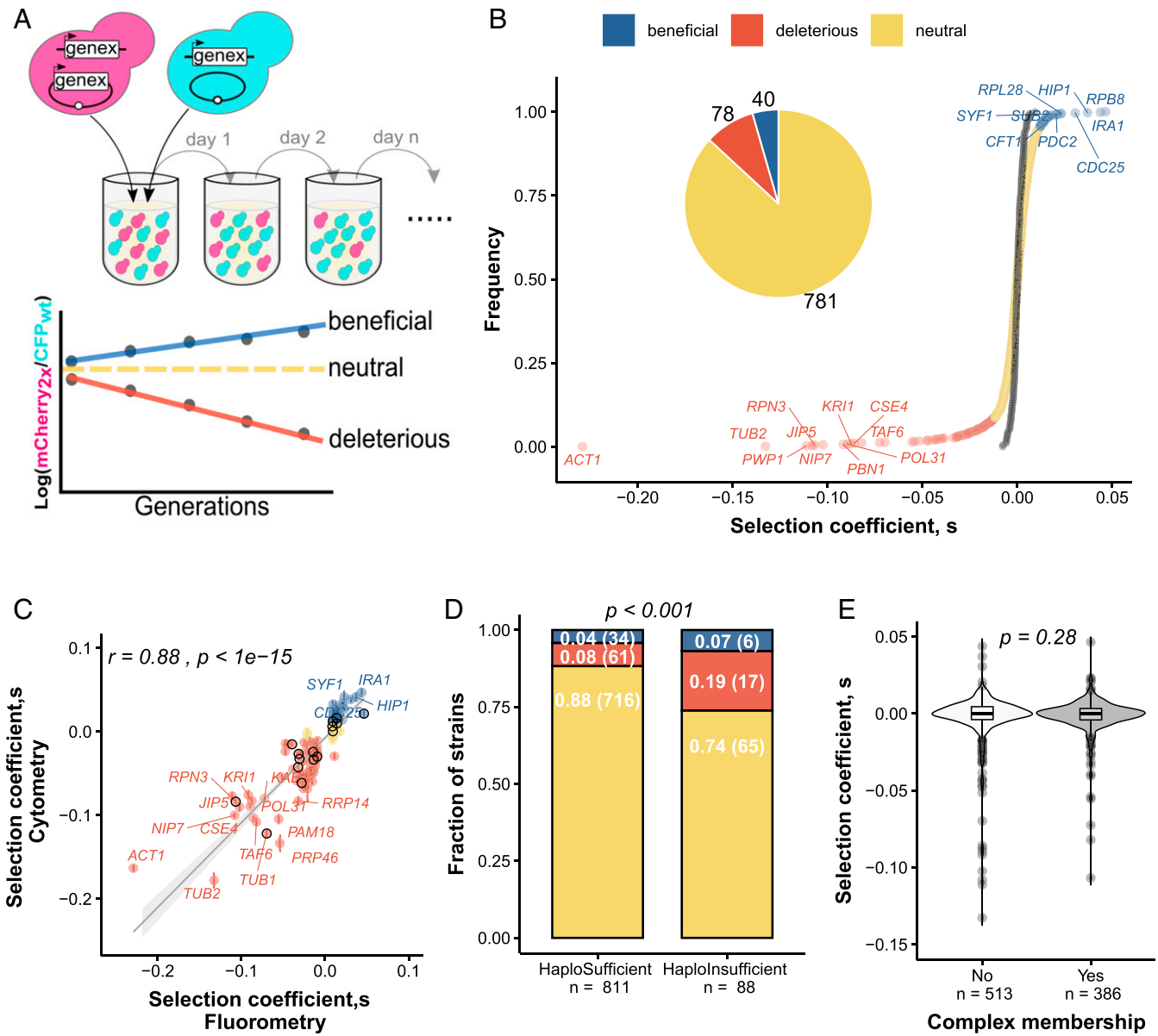


Fig. 1. More than 10% of yeast essential genes affect fitness when duplicated. (A) Relative fitness was measured using a high-resolution competition assay (23). We cocultured a mCherry-tagged strain carrying an extra copy of an essential gene on a centromeric plasmid with a CFP-tagged reference strain carrying a control plasmid. We followed the ratio of the two populations for up to 28 generations to calculate a slope, which corresponds to the selection coefficient (s). (B) Cumulative distribution of selection coefficients of all the 899 strains tested (Dataset S1). Each dot represents a strain expressing an additional gene copy. The black dots represent the distribution of 192 biological replicates of reference-versus-reference competition. The threshold used for deleterious (in red) or beneficial (in blue) effect is at least 1% ($-4.5 > z \text{ score} > 4.5$). (C) Selection coefficients for the validation of the 180 genes with significant effects measured by flow cytometry (Dataset S2). The labels are for genes with the strongest deleterious and beneficial effects. The bars indicate the SD of three biological replicates. The black circles highlight genes with haploinsufficient phenotypes. Spearman's correlation coefficient is indicated at the Top. (D) A comparison of fitness effects among haploinsufficient and haplo-sufficient genes (12). P value from a Fisher exact test is shown. The fraction and number of genes are indicated with white numbers. (E) Selection coefficients of genes that code for proteins that are members of complexes and proteins that are not. On Top, we show the P value from a Wilcoxon rank-sum test.

having no detectable effect in the presence of caffeine. On the other hand, *ACT1* and *TUB2* have similar deleteriousness across the five conditions tested when duplicated. The mechanisms by which an increase in gene dosage is sometimes beneficial remains to be determined. In the presence of antagonistic pleiotropy, it is possible that expression in any given condition is not optimal but rather represents a tradeoff in terms of adaptation across conditions. Indeed, a study looking at the fitness effects of changes in expression levels of several genes showed that the *WT* expression level in some conditions is often not optimal (36).

Deleterious duplications are more frequent and have stronger effects than advantageous ones and include genes such as *TUB1*, *TUB2*, and *ACT1*. The products of these genes are involved in the structural integrity of the cell cytoskeleton and have been previously shown to be highly sensitive to dosage increase (37–39). Our results confirm recent observations that doubling or halving the expression of *TUB1* and *TUB2* is enough to reduce fitness (36). These observations suggest that genes that are deleterious upon duplication could also be haploinsufficient. We indeed find that 19% of haploinsufficient genes identified by

Deutschbauer et al. (12) and tested here produce deleterious effects when duplicated (17/88; Fig. 1D), which is more than the 8% expected by chance (61/811; Pearson's χ^2 test, $P < 1e-3$; Fig. 1D). Conversely, Qian et al. (4) suggested that haploinsufficient genes may be beneficial when duplicated. We see a tendency in this direction, but it is not significant (6/88 and 34/811; χ^2 test, $P = 0.388$; Fig. 1D).

Our interpretation regarding the fitness effects of duplication depends on whether gene expression from a centromeric plasmid approximates the dosage effect of gene duplication, which we expect to commonly be a doubling of dosage [exceptions have been shown, with higher expression than expected (40)]. Centromeric plasmids usually segregate as yeast chromosomes and are on average found in one copy per haploid cell (41), but it is possible that the number of plasmids varies (42). Payen et al. (7) confirmed that in glucose-limiting conditions, the plasmids we used are typically found in only one copy per cell, so our results are overall representative of individual gene duplication events. We also examined if the plasmid-copy genes are regulated similarly as the genome-encoded copy. We reasoned that if pCEN are systematically present in multiple copies, a protein expressed from a plasmid would result in higher expression than when expressed from the genome in an equivalent genetic background where we have a genomic copy and a plasmid copy. We compared protein abundance of a strain with an endogenous green fluorescent protein (GFP) fusion and a copy of the gene on a pCEN plasmid, with the same strain but this time having the GFP fusion expressed from the pCEN plasmid. Only one gene out of five showed higher protein abundance when expressed on the plasmid, and one showed reduced expression (SI Appendix, Fig. S4A). For the two genes with different expression levels, the differences are rather modest (less than onefold) as opposed to orders in magnitude that are common in multicopy plasmids. We also directly measured the abundance of Act1p in the presence of an additional gene copy on a plasmid (pCEN-*ACT1*) by Western blot and found only a modest increase in abundance (SI Appendix, Fig. S4B), which is inconsistent with its deleterious effects being driven by severalfold expression changes. These observations suggest that our systematic strategy using pCEN constructions is a good experimental approximation of naturally occurring duplications in terms of dosage.

The Duplication of Individual Subunits Rarely Affects PPIs in Complexes. The dosage balance hypothesis predicts that both underexpression and overexpression of a protein complex subunit would cause deleterious phenotypes because dosage perturbation affects the stoichiometric balance between the subunits of the complex and compromises its assembly (9, 10). For instance, some subunits of the RNA polymerase II (Rpb2p) and of the proteasome (Rpn3p) are both haploinsufficient and deleterious when duplicated. This indicates that such proteins are sensitive to changes in gene dosage in both directions, just as the gene balance hypothesis predicts. However, when we mapped the fitness effects of gene duplication on annotated yeast protein complexes, we found no significant difference compared to genes that are not in complexes (Wilcoxon rank-sum test, $P = 0.28$; Fig. 1E). In fact, both gene categories share a similar percentage of genes with deleterious effects (8% and 9%, respectively). These findings are robust to different z-score thresholds (Dataset S3). Therefore, the duplication of members of protein complexes is not particularly associated with a decrease in fitness, in apparent contradiction with the dosage balance hypothesis.

Our results suggest that either doubling gene dosage does not affect the assembly of protein complexes, or it affects their assembly but without having particularly strong effects on fitness. Therefore, we next aimed at directly measuring if gene

duplication affects PPIs within complexes in vivo. We selected complexes that are sensitive to some but not all gene dosage changes (the 26S proteasome and the three RNA polymerases; SI Appendix, Fig. S5). We measured pairwise PPIs between all pairs of subunits with and without the duplication of each subunit using a protein-fragment complementation assay (PCA) based on the dihydrofolate reductase (DHFR) enzyme (DHFR-PCA) (43). The quantitative nature of PCA allows us to estimate a perturbation score (ps) as a direct measure of the effects of gene duplication on the PPI network of a complex (Fig. 2A).

Although nonessential retromer genes were not included in the fitness assays, we first tested our approach on this small and well-characterized complex as a proof of concept, since it has been reported to have PPIs sensitive to gene deletion (44). This complex is associated with endosomes and is required for endosome-to-Golgi retrieval of receptors (e.g., the Vps10 protein) that mediate delivery of hydrolases to the vacuole (45). Functionally, the retromer is divided into two subcomplexes: a cargo-selective trimer of Vps35p, Vps29p, and Vps26p, and a membrane-bending dimer of Vps5p and Vps17p. DHFR-PCA detects significant interactions from all the retromer subunits (SI Appendix, Fig. S6A). From our data, we detected same-subunit competition effects between DHFR-fused and nonfused copies of the proteins. Since the extra copy of the duplicated subunit is not tagged with a DHFR reporter fragment, it competes with the tagged copies for the same partners, resulting in a decreased colony size for all the interactions of this subunit (Fig. 2A). For instance, we see a reduction of the interaction of all Vps5p PPIs upon duplication of *VPS5* (blue row and blue column, Fig. 2A). We also find that the duplication increases the strength of the interactions between Vps35p-Vps26p, members of the other subcomplex, the cargo-selective trimer (Fig. 2A). These results show that our strategy has enough resolution to detect small perturbations in the PPIs after the duplication of a single gene coding for a complex subunit.

We next examined the effect of gene duplication on four complexes with proteins for which some gene duplications alter fitness, as identified in our previous analysis, namely the proteasome and the three RNA polymerases (SI Appendix, Fig. S5). The proteasome is a highly conserved and thoroughly described eukaryotic protein complex that is amenable to study by DHFR-PCA (43, 46). In yeast, the core complex (20S) is associated with the regulatory particle (19S) to form a large complex (26S) composed of 37 subunits (47). From hereinafter, we will refer to the 26S proteasome as the proteasome. We tested pairwise interactions among 21 subunits as baits and 16 subunits as preys that belong to either the regulatory particle or the core complex for a total of 305 combinations. We detected 47 PPIs between subunits in the *WT* strain (SI Appendix, Fig. S6B). The RNA polymerases are also well-described large complexes: RNApol I includes 14 subunits, RNApol II has 16, and RNApol III has 18 subunits (48–50). We tested all combinations between all available subunits since five subunits are shared between the three RNA polymerases. We observed 33 significant PPIs out of 689 combinations tested between 31 baits and 26 preys in a *WT* background (SI Appendix, Fig. S6C).

We observed same-subunit competition effects (Fig. 2B; Wilcoxon's rank-sum test, $P < 2e-16$), which validates that additional copies of the proteins are expressed and that we can measure quantitative changes in their PPIs. Indeed, 135/181 of cases where the duplicated subunit is involved in the PPI tested show a reduced interaction score. We observed that most subunit duplications have small to nondetectable effects on the interaction network of their complex and are weakly correlated with a fitness effect. For the proteasome, excluding competition combinations, only 46 out of 8,917 combinations tested were significantly

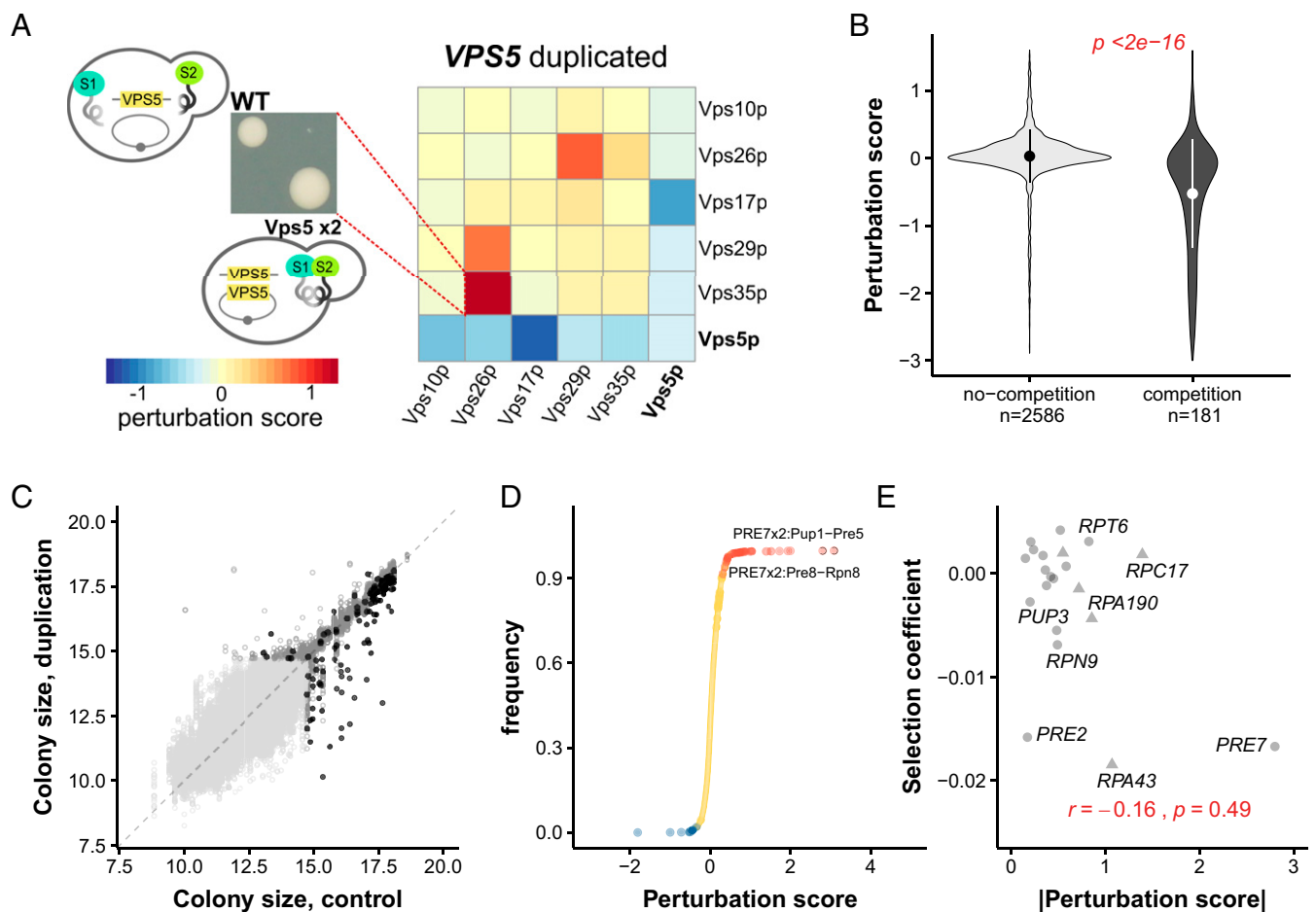


Fig. 2. Most duplications of subunits do not affect PPIs in large complexes. (A) DHFR-PCA-based strategy to measure perturbation of the pairwise physical interactions after a duplication. On the *Left*, in a DHFR-PCA, the colony size on selective media (MTX) is correlated with the stability and strength of the physical interaction between the two subunits S1 and S2 (shown in green). The perturbation score (ps) is defined as the colony size difference between the strain carrying a duplication (+pCEN-*VPS5*) and a *WT* (+pCEN) strain. Heatmap indicating ps values of the complete retromer PPI network due to the duplication of *VPS5*. (B) Distributions of ps for interactions with and without a competing subunit. Since the duplicated protein is not tagged with a DHFR fragment, it titrates PPI partners away from the tagged copy, decreasing colony size. We show violin plots of the distributions of all the interactions tested for five complexes (Dataset S5). On *Top*, we show the P value from a Wilcoxon rank-sum test. (C) Colony sizes of all strains carrying the duplication of a subunit compared with their control strain (the empty vector). Colony sizes of diploid strains carrying all tested combinations of preys, baits, and duplications for the proteasome, and the three RNA polymerases (Datasets S4 and S5). The dark gray circles indicate strains above our growth threshold indicative of physical interaction, while the light gray circles are strains below the growth threshold. The black circles indicate interactions with a competing subunit above the threshold. (D) Cumulative frequency of ps of the proteasome and the three RNA polymerases. All competition effects were excluded. Labeled are the prey–bait combinations that are perturbed by the duplication of *PRE7*. (E) Relationship between the selection coefficient and the average ps (absolute value) of duplicated subunits on PPIs. Only significant (FDR of 5%) and noncompetition combinations were used to calculate the averages. The circles represent duplications of proteasome subunits, while triangles represent subunits of any of the three RNA polymerases. In red, we show Spearman’s correlation coefficient.

different (false discovery rate [FDR] of 5%) from the *WT* interactions within the complex. For the three RNA polymerases, only 28 out of 21,341 combinations tested were significantly different (Fig. 2C). Overall, most of the significant perturbations are gains of PPIs (55/74), which may suggest that when a duplication alters the interaction dynamics within the complex it does so by increasing the strength or amount of PPIs of other subunits (Fig. 2C and D). The strongest effects are seen for the duplication of *PRE7*, especially for interactions Pup1p–Pre5p and Pre8p–Rpn8p (Fig. 2D). Pup1p, Pre5p, and Pre8p are part of the same subcomplex that includes Pre7p and they interact closely during the formation of the 20S proteasome. Pup1p is the $\beta 2$ subunit while Pre5p and Pre8p are the $\alpha 6$ and $\alpha 2$ subunits, respectively (51), and share close spatial proximity, ranging from 66 to 178 Å (46).

If changes in PPIs are associated with the fitness defects we measure, we hypothesized that we would see a correlation

between the perturbation of PPIs and fitness effects. We calculated the mean ps for each subunit (mean of the absolute values of significant perturbation scores) and compared it with the selection coefficient of strains containing a gene duplication of the same subunit (Fig. 2E). The correlation between ps and fitness is negative as predicted but not significant (Spearman’s $r = -0.16$, $P = 0.49$). These results suggest that subunit duplications typically have little or no effect on the protein interaction network within the complex and these effects are largely independent of the fitness effects.

Most Proteasome Subunits Have an Attenuated Expression Level When Duplicated. Our experiments indicate that most PPIs within the proteasome interaction network and RNA polymerases are not significantly perturbed after duplication of their subunits. This suggests that these protein complexes are largely

resilient to changes in gene dosage of their components. Because one of the strongest *ps* was observed in the proteasome, we focused on this complex to explore the underlying mechanisms of such robustness. It has been reported in multiple studies that transcription is usually correlated with gene copy number while protein abundance correlates more poorly (18). Therefore, regulatory mechanisms reducing the protein abundance of the proteasome subunits, also known as attenuation, could explain why duplication is not perturbing PPI between the subunits.

To test whether the proteasome subunits are attenuated, we looked for changes in protein abundance after gene duplication. We compared protein abundance in GFP-tagged strains (52) carrying a duplication of the gene or an empty vector as a control (Fig. 3A). Protein attenuation would lead to a reduction of protein levels of both copies, which would result in reduced fluorescence signal. Most subunits (17/19) have a significant decrease in GFP-fluorescent signal after duplication (Fig. 3B and Dataset S6). Next, we calculated an attenuation score: the difference between *WT* and the duplicated GFP-fluorescent signals normalized by the *WT* (Fig. 3C). If the expression of the fused copy was reduced by half to balance the additional copy in a plasmid (complete attenuation), the attenuation score would be 0.5. Interestingly, attenuation is similar for subunits belonging to the same subcomplex (Fig. 3D), suggesting a regulation that depends on complex assembly. Havugimana et al. (53) reported that the stoichiometry within each proteasome subcomplex is 1:1, while the stoichiometry among subcomplexes varies from 1:1 to 1:4 (Dataset S6). This observation could explain why subunits belonging to the same component have similar expression and attenuation patterns, while there are significant differences between components (Fig. 3D; $P = 0.002$, one-way ANOVA test).

Strikingly, all but two complex subunits appear to be attenuated. One of the most attenuated subunits Pre7p has a significant effect on both PPIs and fitness when duplicated (Fig. 2C). The duplication of *PRE7* perturbs PPIs between Pre5p, Pre8p, Pup1p, and Rpn8p (Fig. 3E). All of them share close spatial proximity with Pre7p in the proteasome core complex. We tested whether *PRE7* duplication affects the abundance of these subunits by measuring the GFP signal of all proteasome subunits with and without *PRE7* duplication. Most subunits are unaffected upon *PRE7* duplication but the four “perturbed” subunits (Pre5p, Pre8p, Pup1p, and Rpn8p) show a modest increase in their protein abundance (Fig. 3F). Even though the difference between the control and the *PRE7* duplicated background is small, it is highly reproducible and significant (SI Appendix, Fig. S7A), and appears specific to the subunits with altered interactions (SI Appendix, Fig. S7B). These modest but significant changes in protein abundance of perturbed subunits after the duplication of *PRE7* may explain the changes we observed in their PPIs. The duplication of *PRE7*, even if largely attenuated, affects the organization of the proteasome by affecting the abundance and interactions of a few other subunits. The dosage balance hypothesis may therefore apply to a limited number of subunits.

The proteasome subunits show different levels of attenuation, and recent studies (18, 19) suggest that attenuation occurs mostly at the posttranscriptional level across complexes. To examine this specifically for the proteasome, we retrieved data from Dephoure et al. (18) and looked at the mRNA abundance ratio of individual genes in disomic strains relative to *WT*. Most proteasome subunits roughly double their transcript levels when duplicated (relative to a log₂ mRNA ratio around 1; SI Appendix, Fig. S8A). This includes Pre3p and Pre7p, two of the most attenuated subunits in our experiments (Fig. 3C). Since aneuploidies can cause systemic-level changes on the transcriptome and proteome and confound the effects of an individual

duplication, we performed RT-qPCR analysis of four attenuated subunits (*PRE7*, *RPT2*, *RPT6*, and *SCL1*), including three not measured in Dephoure, a nonattenuated subunit (*PRE10*), and a gene that is not a part of the proteasome (*FAS2*). As expected, the nonattenuated genes *PRE10* and *FAS2* present no significant change in their transcript levels after their duplication (*t* test, $P = 0.6$ and 0.78 ; SI Appendix, Fig. S8B). Highly attenuated genes at the protein level exhibit varying responses: *RPT6* shows a significant mRNA attenuation (*t* test, $P = 0.01$), and *RPT2* and *SCL1* also show a marginally significant mRNA attenuation (*t* test, $P = 0.1$ and 0.07). *PRE7* displays no significant change in transcript levels (*t* test, $P = 0.24$). These results from individual gene duplication are therefore consistent with the data observed for disomic strains. Combined, these data suggest that attenuation can occur at the transcriptional level but more frequently at the posttranscriptional level (SI Appendix, Fig. S8C). In most cases, as for *PRE7*, attenuation appears to be posttranslational because transcription and translation rates are both maintained in disomic strains (SI Appendix, Fig. S8A) (17). We further explored the attenuation of the *PRE7* subunit by using a tunable expression system. We constructed a strain with tunable *PRE7* expression (54) and which contains two gene copies of *PRE7* tagged with different fluorescent proteins that were independently monitored by flow cytometry. There is a significant negative correlation (Pearson’s correlation $r = -0.25$, $P < 10e-15$) between the protein abundance of the copy expressed under the inducible promoter and abundance of the chromosome copy (SI Appendix, Fig. S9 and Dataset S8). These results suggest that attenuation is expression level dependent. Finally, we examined whether this posttranslational control depends simply on the presence of an additional gene copy or on that gene copy being transcribed and translated. For this purpose, we generated a centromeric construction that contains *PRE7*’s full promoter and open reading frame (ORF) but that cannot be translated because it lacks a start codon (two constructions) or the start codon is followed by a stop codon (SI Appendix, Fig. S10). Pre7p remains at the *WT* level of expression in the presence of these three constructions. Since all the *cis*-regulatory elements remain intact on these constructions, it is fair to assume that transcription is initiating but translation is not initiated or is terminated prematurely due to the insertion of the stop codons. This confirms that attenuation requires the translation of the mRNA.

Discussion

The long-term fate of gene duplicates has been studied in detail theoretically, experimentally, and by using genome, transcriptome, and proteome data. While less is known about the immediate impact of gene duplication, it has become clear that dosage sensitivity determines whether gene duplications have any chance to be retained and fixed in a population (55). By duplicating 899 genes individually and examining the distribution of fitness effects, we find that duplicates with a greater than 1% fitness effect are common (~12%). Furthermore, duplications are twice as likely to be deleterious than beneficial, and deleterious effects are larger in magnitude. Consistent with previous observations, deleterious duplications are more frequent among genes that are also sensitive to a reduction in gene dosage. However, duplications do not have more deleterious effects when they affect protein complexes, contrary to what is predicted from the dosage balance hypothesis. To elucidate this discrepancy, we looked at the effect of duplication at the PPI level. To test whether these deleterious effects impact fitness by affecting PPIs in protein complexes, we measured the perturbation of protein complexes *in vivo* as a response to gene duplication, focusing on the proteasome and three RNA polymerases. Overall, only 0.24% of the tested duplication–PPI combinations

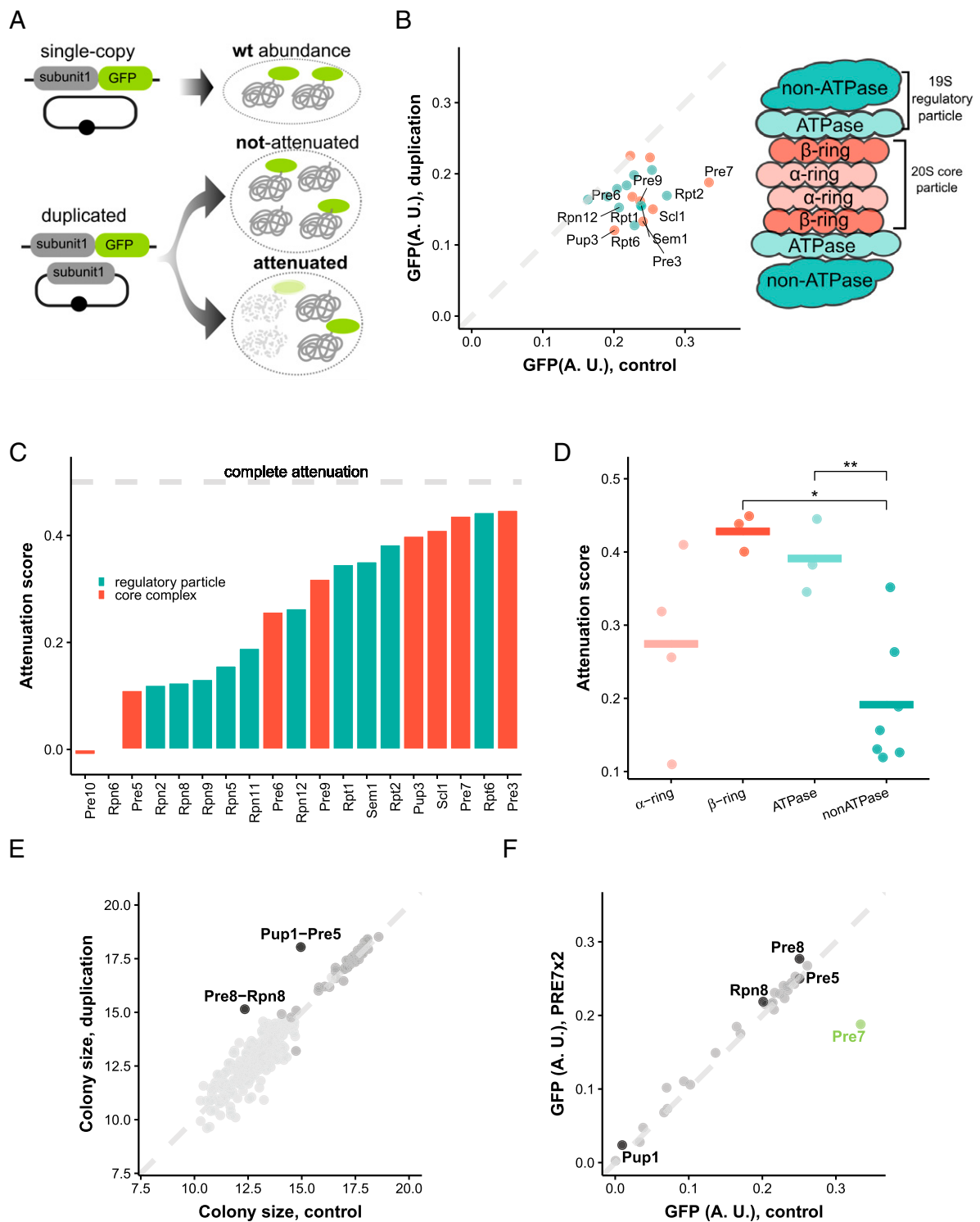


Fig. 3. Attenuation of protein abundance after duplication in most proteasome subunits. (A) Measure of attenuation with GFP-tagged proteins. Changes in abundance of each subunit can be detected by comparing fluorescent signals of GFP-tagged subunits before and after duplication. Upon attenuation, the abundance of the tagged copy will be reduced. (B) GFP signal comparison between strains carrying a duplication of the GFP-tagged subunit and their corresponding control. On the *Right*, a cartoon of the proteasome with its components. All GFP values are corrected for autofluorescence by subtracting the signal of the parental strain not expressing GFP and by cell size (Dataset S6). (C) Attenuation scores of all assayed proteasome subunits. The attenuation score is the difference between GFP fluorescent signals of the control strain (bearing a control plasmid) and the duplicated strain (bearing a centromeric plasmid with an extra copy of the subunit) divided by the GFP signal of the control. In the absence of attenuation, this value is 0. Upon complete attenuation, it is 0.5. (D) Attenuation scores of the proteasome subcomplexes. On the *Right*, the asterisks indicate significant differences between components calculated by correcting for multiple testing (Tukey's test; $*P \leq 0.05$ and $**P \leq 0.01$). (E) Colony sizes of all strains carrying the *PRE7* duplication (+pCEN-*PRE7*) compared with their control strain (the empty vector) indicating changes in PPI in the DHFR-PCA assay. The black dots highlight the subunits that have interactions disturbed after *PRE7* duplication. (F) GFP signal of proteasome subunits before and after *PRE7* duplication (+pCEN-*PRE7*). The black dots highlight the subunits that have interactions disturbed after *PRE7* duplication (Dataset S7). Replicate measurements are available in *SI Appendix*.

significantly perturbed their protein complexes, and a single subunit largely drives these results. By focusing on the proteasome, we further examined why so few PPIs were perturbed by changes in gene dosage and found that most of its subunits are attenuated at various extents, i.e., the protein level decreases close to a normal level even if the gene is duplicated or its expression is modified with a tunable promoter. Therefore, our results suggest that gene duplication is unlikely to have an impact on fitness through the perturbation of protein complex assembly at least partly because the extra copies of the genes show attenuated responses at the transcriptional, posttranscriptional, and posttranslational levels. Altogether, these observations challenge the dosage balance hypothesis. Our results rather bring support to a model in which decreased dosage, and not increased dosage, affects protein complexes (13) by identifying a potential mechanism for this asymmetry of effect. In the future, a better understanding of attenuation at the molecular level will allow us to manipulate it and test its causal role in buffering fitness and other molecular effects on the cell.

There is at least one gene that appears to be an exception, as despite being largely attenuated it is highly deleterious and affects PPIs. Further experiments suggest that these alterations are a result of changes in protein abundance of other subunits, which may occur through stabilizing interactions. In previous experiments where we combined gene deletion with the study of PPIs, we documented several cases of protein destabilization by the deletion of an interaction partner (44, 56). What we see here could be the reciprocal effect. Given that Pre7p is one of the most abundant subunits of its complex, attenuation may not be sufficient in this case to eliminate the effects of its duplication.

Duplication events of individual ORFs with their *cis*-regulatory region is less common than other mechanisms of duplication. Indeed, consecutive tandem duplications represent less than 2% of the yeast genome and are not conserved (57). For instance, when duplication occurs by retrotransposition, a single coding DNA sequence is duplicated without its *cis*-regulatory region (58). Most duplications occur due to recombination errors that lead to the duplication of long segments or even complete chromosomes, meaning that more than a single gene is duplicated at the same time (58, 59). These observations limit the generalization of our observations made on individual duplications. Nevertheless, it has been suggested that the adaptive value of some aneuploidies can be mapped to a single or a handful of loci (60, 61). This is the case for the duplication of the high-affinity sulfate transporter *SUL1* that confers an advantage under sulfate limitation condition (62). Even though our experimental strategy may not fully emulate the most common duplication events in nature, it is a powerful approach to systematically evaluate the impact of individual genes and the possible role of natural selection in the fixation or loss of newly arisen duplicated genes, independently of their origin. In addition, to understand how more complex duplication may impact cell biology and fitness, we need to understand how individual genes affect cell biology in the first place.

Attenuation of multiprotein complex members has been documented previously in overexpression experiments (20), in aneuploid yeast strains (18, 19), and to some extent in cancer cells (11), which may suffer from a general alteration of protein homeostasis and protein quality control (63). It is unclear whether attenuation is exclusive to complex subunits, but evidence from previous works indicate that it is more frequent among these proteins (18, 20). For instance, Dephoure et al. (18) found that in disomic strains, 76% of attenuated proteins are members of protein complexes. Here, we show that attenuation is also taking place with small copy-number variation affecting individual genes, such as in the case of gene duplication. This

feature appears to be part of the regulation of proteins in normal cells. Indeed, a recent study by Taggart et al. (64) suggested that nearly 20% of proteasome proteins are overproduced since more than half of the subunits synthesized are degraded in normal conditions. Therefore, our results suggest that mechanisms acting to regulate protein abundance in normal conditions provide, as a side effect, the extra advantage of protecting the cell against copy-number variation of some members of important multiprotein complexes. In the case of aneuploid cells and protein overexpression, several mechanisms for the attenuation of protein levels have been proposed such as autophagy, the *HSF1/HSP90* pathway (65), and the ubiquitin–proteasome system (18, 20). Since the proteasome itself may play an essential role in attenuation, studying the mechanism of attenuation in this complex may prove to be challenging. Nevertheless, the mechanisms protecting the cells of the proteotoxic stress caused by aneuploidy may not be the same that act in the case of individual duplications. The molecular mechanisms that attenuate protein abundance in the case of individual gene duplications are still an open question.

Our observations, along with those of previous studies, have an important impact on our understanding of the evolution of protein complexes. Complexes are often composed of multiple pairs of paralogs (66), particularly, but not exclusively, those coming from a whole-genome duplication (WGD). This has been suggested to come from the fact that WGD preserves the stoichiometry of protein complexes while SSDs do not. Our results show that many protein complexes may be resilient to increase dosage of individual genes, suggesting that other mechanisms than stoichiometric imbalance could be at play to explain the different retention rates of the two types of paralogs in complexes.

Evolutionary forces leading to regulatory mechanisms that attenuate protein abundance, therefore, could diminish the immediate molecular effects of gene duplications and could allow them to reach a higher frequency in populations by buffering negative effects. Such pressures to maintain gene dosage could come, for instance, from the requirement to assemble protein complexes in a stoichiometric fashion in the face of gene expression noise (67). Another pressure for the attenuation of extra subunits is the need to prevent spurious interactions between the unassembled subunits and other proteins through its exposed sticky interface (68) or aggregation (69). If selection for decreasing noise or to reduce spurious interactions led to the evolution of expression attenuation, it may have also contributed to the robustness of protein complexes to gene duplications. Questions that remain are to what extent expression attenuation is biased toward protein complexes and if attenuation could also affect proteins that are not part of protein complexes and thus could also influence the evolution of a much larger set of genes.

Materials and Methods

The detailed material and methods are available in *SI Appendix*. We measured competitive fitness using automated fluorescence-based competition assays (23). pCEN were obtained from the MoBY-ORF collection (24). The mCherry-tagged collection carrying individual “duplications” (Y8205-mCherry +MoBY-xxx) were competed with a universal CFP-tagged strain (Y7092-CFP + p5586). Saturated cultures were diluted 1:16 in fresh media every 24 h and monitored for ~28 generations (7 d) in a robotic system (Freedom EVO; Tecan) with a microplate multireader (Infinite Reader M100; Tecan). We estimated each strain selection coefficient as described by DeLuna et al. (23). We calculated z scores using the mean and the SD of a control distribution of 192 mCherry-tagged *WT* strains competed against the universal CFP-tagged reference. For the cytometry-based competition assays, the Y8205-mCherry + MoBY-xxx collection was cocultured with a universal YFP-tagged strain. Before each daily dilution, we took a sample of 10 μ L of saturated culture and made a 1:10 dilution in TE 2X to measure in the cytometer (LSRFortessa; BD) collecting up to 30,000 events per sample.

For the DHFR-PCA, we acquired strains with DHFR fragment fusions from the Yeast Protein Interactome Collection (43). Confirmed prey strains (DHFR-F [3] fusions) were transformed with MoBY plasmids available for their complex (Dataset S4). The DHFR-PCA screening was performed using standard methods used in previous works (56, 70). Haploid strains were assembled in arrays of 1,536 colonies on agar and were manipulated by a fully automated platform (BM5-SC1; S&P Robotics). The nine-plate prey collection (MAT α prey-DHFR-F[3] + MoBY-xxx, Hygromycin B, and G418 resistant) was mated with a set of 54 bait plates (MAT α bait-DHFR-F[1,2], NAT resistant). We replicated the set of 220 mating plates on diploid selection plates, followed by selection on DMSO and MTX media (Dataset S10). We monitored growth for plates of the last step of both DMSO and MTX selections every 24 h using a Rebel T5i camera (Canon) attached to the robotic system (S&P Robotics).

To measure changes in protein abundance of proteasome subunits, we extracted 33 strains from the Yeast GFP fusion Collection (52) and generated de novo GFP-fusions that were not present in this collection. The GFP

fluorescence of 5,000 cells in the exponential phase was measured using a Guava flow cytometer (Millipore). The GFP signal was normalized with cell size using the FSC-A value and corrected background autofluorescence using nonfluorescent parental strain.

Data Availability. All MatLab and R scripts used for data analysis and visualization have been deposited in the GitHub repository (https://github.com/Landrylab/AscencioETAL_2020).

ACKNOWLEDGMENTS. We thank E. Mancera and members of the C.R.L. and A.D. laboratories for discussions and comments on the manuscript. This research was supported by Canadian Institutes of Health Research Foundation Grant 387697 (to C.R.L.) and by Consejo Nacional de Ciencia y Tecnología México through Grant PN-2016/2370 (to A.D.). C.R.L. holds the Canada Research Chair in Cellular Systems and Synthetic Biology.

- M. Lynch, A. Force, The probability of duplicate gene preservation by sub-functionalization. *Genetics* **154**, 459–473 (2000).
- W. Qian, J. Zhang, Gene dosage and gene duplicability. *Genetics* **179**, 2319–2324 (2008).
- S. Ohno, *Evolution by Gene Duplication* (Springer, Berlin, 1970).
- W. Qian, B.-Y. Liao, A. Y.-F. Chang, J. Zhang, Maintenance of duplicate genes and their functional redundancy by reduced expression. *Trends Genet.* **26**, 425–430 (2010).
- R. C. Iskow, O. Gokcumen, C. Lee, Exploring the role of copy number variants in human adaptation. *Trends Genet.* **28**, 245–257 (2012).
- W. Qian, J. Zhang, Genomic evidence for adaptation by gene duplication. *Genome Res.* **24**, 1356–1362 (2014).
- C. Payen *et al.*, High-throughput identification of adaptive mutations in experimentally evolved yeast populations. *PLoS Genet.* **12**, e1006339 (2016).
- A. M. Rice, A. McLysaght, Dosage-sensitive genes in evolution and disease. *BMC Biol.* **15**, 78 (2017).
- J. A. Birchler, R. A. Veitia, The gene balance hypothesis: From classical genetics to modern genomics. *Plant Cell* **19**, 395–402 (2007).
- B. Papp, C. Pál, L. D. Hurst, Dosage sensitivity and the evolution of gene families in yeast. *Nature* **424**, 194–197 (2003).
- E. Gonçalves *et al.*, Widespread post-transcriptional attenuation of genomic copy number variation in cancer. *Cell Syst.* **5**, 386–398.e4 (2017).
- A. M. Deuschbauer *et al.*, Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics* **169**, 1915–1925 (2005).
- J. I. Semples, T. Vavouri, B. Lehner, A simple principle concerning the robustness of protein complex activity to changes in gene expression. *BMC Syst. Biol.* **2**, 1 (2008).
- R. Sopko *et al.*, Mapping pathways and phenotypes by systematic gene over-expression. *Mol. Cell* **21**, 319–330 (2006).
- H. Moriya, K. Makanae, S. Iwasaki, N. T. Ingolia, H. Moriya, Post-translational dosage compensation buffers genetic perturbations to stoichiometry of protein complexes. *PLoS Genet.* **13**, e1006554 (2017).
- A. B. Oromendia, S. E. Dodgson, A. Amon, Aneuploidy causes proteotoxic stress in yeast. *Genes Dev.* **26**, 2696–2708 (2012).
- X. He, J. Zhang, Higher duplicability of less important genes in yeast genomes. *Mol. Biol. Evol.* **23**, 144–151 (2006).
- A. DeLuna *et al.*, Exposing the fitness contribution of duplicated genes. *Nat. Genet.* **40**, 676–681 (2008).
- C. H. Ho *et al.*, A molecular barcoded yeast ORF library enables mode-of-action analysis of bioactive compounds. *Nat. Biotechnol.* **27**, 369–377 (2009).
- S. A. Morrill, A. Amon, Why haploinsufficiency persists. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 11866–11871 (2019).
- D. Broek *et al.*, The *S. cerevisiae* CDC25 gene product regulates the RAS/adenylate cyclase pathway. *Cell* **48**, 789–799 (1987).
- K. Tanaka, K. Matsumoto, A. Toh-E, IRA1, an inhibitory regulator of the RAS-cyclic AMP pathway in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **9**, 757–768 (1989).
- S. Venkataram *et al.*, Development of a comprehensive genotype-to-fitness map of adaptation-driving mutations in yeast. *Cell* **166**, 1585–1596.e22 (2016).
- J. Tanaka, G. R. Fink, The histidine permease gene (HIP1) of *Saccharomyces cerevisiae*. *Gene* **38**, 205–214 (1985).
- S. Velmurugan, Z. Lobo, P. K. Maitra, Suppression of *pdh2* regulating pyruvate decarboxylase synthesis in yeast. *Genetics* **145**, 587–594 (1997).
- M. Lynch *et al.*, A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9272–9277 (2008).
- K. H. Wolfe, D. C. Shields, Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**, 708–713 (1997).
- F. A. Kondrashov, Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proc. Biol. Sci.* **279**, 5048–5057 (2012).
- N. Pavelka *et al.*, Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature* **468**, 321–325 (2010).
- W. Qian, D. Ma, C. Xiao, Z. Wang, J. Zhang, The genomic landscape and evolutionary resolution of antagonistic pleiotropy in yeast. *Cell Rep.* **2**, 1399–1410 (2012).
- L. Keren *et al.*, Massively parallel interrogation of the effects of gene expression levels on fitness. *Cell* **166**, 1282–1294.e18 (2016).
- W. Katz, B. Weinstein, F. Solomon, Regulation of tubulin levels and microtubule assembly in *Saccharomyces cerevisiae*: Consequences of altered tubulin gene copy number. *Mol. Cell. Biol.* **10**, 5286–5294 (1990).
- C. Espinet, M. A. de la Torre, M. Aldea, E. Herrero, An efficient method to isolate yeast genes causing overexpression-mediated growth arrest. *Yeast* **11**, 25–32 (1995).
- D. Burke, P. Gasdaska, L. Hartwell, Dominant effects of tubulin overexpression in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **9**, 1049–1059 (1989).
- D. W. Loehlin, S. B. Carroll, Expression of tandem gene duplicates is often greater than twofold. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 5988–5992 (2016).
- L. Clarke, J. Carbon, Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature* **287**, 504–509 (1980).
- R. Gnügge, F. Rudolf, *Saccharomyces cerevisiae* Shuttle vectors. *Yeast* **34**, 205–221 (2017).
- K. Tarassov *et al.*, An in vivo map of the yeast protein interactome. *Science* **320**, 1465–1470 (2008).
- G. Diss, A. K. Dubé, J. Boutin, I. Gagnon-Arsenault, C. R. Landry, A systematic approach for the genetic dissection of protein complexes in living cells. *Cell Rep.* **3**, 2155–2167 (2013).
- A. S. Mukadam, M. N. J. Seaman, Retromer-mediated endosomal protein sorting: The role of unstructured domains. *FEBS Lett.* **589**, 2620–2626 (2015).
- A.-È. Chrétien *et al.*, Extended linkers improve the detection of protein-protein interactions (PPIs) by dihydrofolate reductase protein-fragment complementation assay (DHFR PCA) in living cells. *Mol. Cell. Proteomics* **17**, 373–383 (2018).
- M. Hochstrasser *et al.*, The *Saccharomyces cerevisiae* ubiquitin-proteasome system. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **354**, 1513–1522 (1999).
- J. Archambault, J. D. Friesen, Genetics of eukaryotic RNA polymerases I, II, and III. *Microbiol. Rev.* **57**, 703–724 (1993).
- A. Sentenac, Eukaryotic RNA polymerases. *CRC Crit. Rev. Biochem.* **18**, 31–90 (1985).
- P. Cramer, Multisubunit RNA polymerases. *Curr. Opin. Struct. Biol.* **12**, 89–97 (2002).
- L. Budenholzer, C. L. Cheng, Y. Li, M. Hochstrasser, Proteasome structure and assembly. *J. Mol. Biol.* **429**, 3500–3524 (2017).
- W.-K. Huh *et al.*, Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691 (2003).
- P. C. Havugimana *et al.*, A census of human soluble protein complexes. *Cell* **150**, 1068–1081 (2012).
- A. Aranda-Díaz, K. Mace, I. Zuleta, P. Harrigan, H. El-Samad, Robust synthetic circuits for two-dimensional control of gene expression in yeast. *ACS Synth. Biol.* **6**, 545–554 (2017).
- A. M. Rice, A. McLysaght, Dosage sensitivity is a major determinant of human copy number variant pathogenicity. *Nat. Commun.* **8**, 14366 (2017).
- G. Diss *et al.*, Gene duplication can impart fragility, not robustness, in the yeast protein interaction network. *Science* **355**, 630–634 (2017).

57. L. Despons *et al.*, Genome-wide computational prediction of tandem gene arrays: Application in yeasts. *BMC Genomics* **11**, 56 (2010).
58. B. Dujon, Yeast evolutionary genomics. *Nat. Rev. Genet.* **11**, 512–524 (2010).
59. J. Zhang, Evolution by gene duplication: An update. *Trends Ecol. Evol.* **18**, 292–298 (2003).
60. N. Pavelka, G. Rancati, R. Li, Dr Jekyll and Mr Hyde: Role of aneuploidy in cellular adaptation and cancer. *Curr. Opin. Cell Biol.* **22**, 809–815 (2010).
61. A. H. Yona *et al.*, Chromosomal duplication is a transient evolutionary solution to stress. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 21010–21015 (2012).
62. D. Gresham *et al.*, The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genet.* **4**, e1000303 (2008).
63. A. B. Oromendia, A. Amon, Aneuploidy: Implications for protein homeostasis and disease. *Dis. Model. Mech.* **7**, 15–20 (2014).
64. J. C. Taggart, H. Zauber, M. Selbach, G.-W. Li, E. McShane, Keeping the proportions of protein complex components in check. *Cell Syst.* **10**, 125–132 (2020).
65. N. Donnelly, V. Passerini, M. Dürrbaum, S. Stingele, Z. Storchová, HSF1 deficiency and impaired HSP90-dependent protein folding are hallmarks of aneuploid human cells. *EMBO J.* **33**, 2374–2387 (2014).
66. J. B. Pereira-Leal, E. D. Levy, C. Kamp, S. A. Teichmann, Evolution of protein complexes by duplication of homomeric interactions. *Genome Biol.* **8**, R51 (2007).
67. H. B. Fraser, A. E. Hirsh, G. Giaever, J. Kumm, M. B. Eisen, Noise minimization in eukaryotic gene expression. *PLoS Biol.* **2**, e137 (2004).
68. E. D. Levy, S. De, S. A. Teichmann, Cellular crowding imposes global constraints on the chemistry and evolution of proteomes. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 20461–20466 (2012).
69. C. M. Brennan *et al.*, Protein aggregation mediates stoichiometry of protein complexes in aneuploid cells. *Genes Dev.* **33**, 1031–1047 (2019).
70. A. Marchant *et al.*, The role of structural pleiotropy and regulatory evolution in the retention of heteromers of paralogs. *eLife* **8**, e46754 (2019).