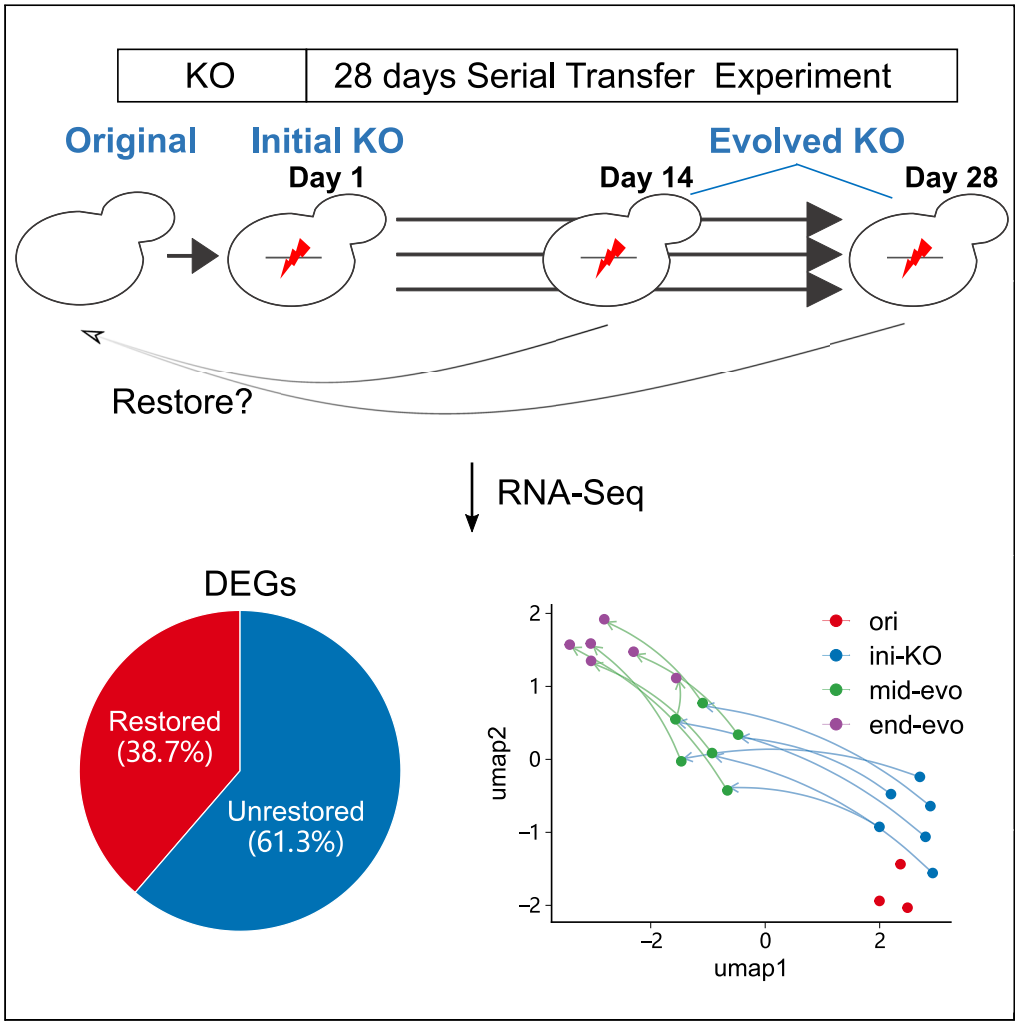
iScience

CellPress
OPEN ACCESS

Article

Progressive transcriptomic shifts in evolved yeast strains following gene knockout



Bei Jiang, Chuyao Xiao, Li Liu

liul47@mail.sysu.edu.cn

Highlights

Most knockout-responsive genes never restore to original expression levels

Global transcriptomes reveal distinct clusters for different evolution stages

Ribosomal genes play key role in adaptive evolution

Jiang et al., iScience 27, 111219
November 15, 2024 © 2024 The Author(s). Published by Elsevier Inc.
https://doi.org/10.1016/j.isci.2024.111219

iScience

CellPress
OPEN ACCESS

Article

Progressive transcriptomic shifts in evolved yeast strains following gene knockout

Bei Jiang,[1,2] Chuyao Xiao,[1,3,4] and Li Liu[1,5,*]

SUMMARY

Gene knockout disrupts cellular homeostasis, altering gene expression, and phenotypes. We investigated whether cells return to their pre-knockout transcriptomic state through adaptive evolution experiments on *hap4Δ* and *ade1Δ* yeast strains. Analysis revealed that genes with higher expression levels and more physical interaction partners in wild-type strains were more likely to be restored, suggesting that genes of significant functional importance have increased resilience to genetic perturbations. However, as the experiment progressed, most initially restored genes became unrestored. Over 60% of differentially expressed genes in knockout strains remained unrestored in evolved strains. Evolved strains exhibited distinct transcriptomic states, diverging from the original strain over time. Ribosome biogenesis components exhibited systematic sequential changes during the evolution. Our findings suggest the knockout strain transcriptomes struggle to return to the original state even after 28 days of culture. Instead, compensatory mechanisms lead to distinct suboptimal states, highlighting the complex transcriptomic dynamics following genetic perturbations.

INTRODUCTION

Gene knockout serves as a potent tool for probing gene functionality across a variety of organisms.[1] This method entails the removal or deactivation of a specific gene, enabling researchers to examine its influence on cellular processes and phenotypes. While gene knockout contributes valuable insights into gene functions, it simultaneously disrupts cellular homeostasis.[2–4] This disruption precipitates changes in gene expression networks, which subsequently affect cellular phenotypes, such as growth rates.[5,6]

Cells possess the capability to re-establish equilibrium through multiple mechanisms, including epigenetic modifications and genetic variations.[7,8] This phenomenon is frequently observed during adaptive evolution when organisms adjust to novel conditions.[9,10] The process encompasses the selection and accumulation of advantageous mutations, which provide a fitness edge in the novel conditions.[11,12] Researches have demonstrated that through adaptive laboratory evolution (ALE), cells can undergo various levels of change, and these changes ultimately result in phenotypic alterations, including improved growth rates and other adaptive traits.[13–15] Johnson et al. evolved 205 populations of *Saccharomyces cerevisiae* (*S. cerevisiae*) strain W303 for approximately 10,000 generations in three distinct environments, and almost all populations experienced rapid fitness increases in the first few hundred generations, followed by slower adaptation as time progressed.[16] A similar pattern of fitness variation was observed in another ALE study using genetically diverse asexual populations produced by segregants from the cross of two *S. cerevisiae* strains (Y55 and SK1) in four different environments.[17]

Focusing specifically on gene knockout scenarios, various studies illuminate the mechanisms that underpin post-knockout cellular responses and identify factors that influence the re-establishment of cellular homeostasis. For instance, research conducted by McCloskey et al. revealed that metabolic gene knockout engendered modifications in metabolic fluxes in *Escherichia coli* (*E. coli*), precipitating disturbances in metabolite concentrations and regulatory networks.[18] These perturbations were ultimately rectified through the acquisition of mutations during ALE. In a separate study, Hsu et al. found that the loss of SEF1 in *Lachancea kluyveri* yeast compelled cells to initiate a compensatory process to counteract the misexpression of TCA cycle genes, and two adaptive loss-of-function mutations of IRA1 and AZF1 were derived under disparate selective conditions during evolutionary repair experiments consequently.[19]

An important question arises regarding the transcriptome state of evolved cells. Several studies have uncovered patterns for the evolved states of adaptation to new environments. For example, Ho et al. examined 44 instances of yeast and *E. coli* adapting to new environments and found that changes in transcriptomes and fluxomes across various adaptations consistently indicated that genetic changes following

[1]MOE Key Laboratory of Gene Function and Regulation, State Key Laboratory of Biocontrol, Innovation Center for Evolutionary Synthetic Biology, School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China
[2]Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong SAR, China
[3]Greater Bay Area Institute of Precision Medicine (Guangzhou), Fudan University, Nansha District, Guangzhou 511400, China
[4]Institute of Life Sciences, Fudan University, Shanghai 200433, China
[5]Lead contact
*Correspondence: liul47@mail.sysu.edu.cn
https://doi.org/10.1016/j.isci.2024.111219

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

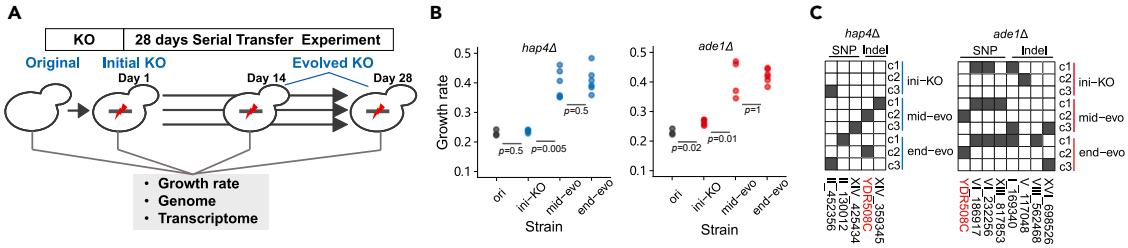


Figure 1. 28-day serial transfer experiments starting from the initial knockout strains of HAP4 and ADE1

(A) Experimental process of adaptive laboratory evolution countering gene knockout.

(B) Growth rates of the *hap4*Δ strains (left) and *ade1*Δ strains (right), respectively. Each point represents a sample, and the original strains are colored in black. The differences between adjacent time points were tested using the Wilcoxon test and the corresponding *p* values were labeled.

(C) Genome variations across each replicate (denoted as c1, c2, c3) for *hap4*Δ strains (left) and *ade1*Δ strains (right), respectively. The gray grid represents an SNP (left) or an indel (right). Roman numeral indicates the chromosome and the number connected by short line is the position in the chromosome. The SNP designated as II_452356 is situated 302 base pairs (bp) upstream of the SND3 gene's coding region. Similarly, XIV_359345 is found 251 bp upstream of the AAH1 gene's coding region, and XVI_698526 is located 343 bp upstream of the MRL1 gene's coding region.

initial plastic changes were reversed rather than reinforced.[7] Similarly, Sandberg et al. observed that the transcriptomes of evolved strains of *E. coli* in ALE to 42°C exhibited a general trend toward restoring the global expression state back toward pre-heat stressed levels.[20]

Interestingly, the pattern of restoration appears unclear when adapting to gene knockout. An ALE experiment involving *pgi* knockout *E. coli* strains with independent replicates led to eight phenotypically distinct endpoints, suggesting multiple optimal phenotypes exist.[21] This discrepancy may stem from varying research foci: some researchers may be interested in the trend of initially changed phenotypes, while others may focus on the global states of evolved strains compared to the initial strain. Despite these advances, it remains uncertain whether the post-knockout rebalanced state reflects the pre-knockout state or represents a distinct, suboptimal state.

In this report, we conducted a 28-day serial transfer experiment involving two gene knockout yeast strains until the growth rate stabilized (Figure 1A). The two gene knockout strains were freshly generated to avoid adaptive mutations that might have occurred during long-term storage. We started from the immediate response genes to the knockout events, and monitored their expression levels throughout the adaptive experiment. Our findings revealed that these response genes struggled to return to their original states, and the global transcriptomes varied across different stages of evolution.

RESULTS

Adaptive laboratory evolution experiment

HAP4 is a transcription factor that plays a wide role in diauxic shift, mitochondrial biogenesis, stress response and so on,[22,23] while ADE1 is involved in adenine biosynthesis.[24] Both genes are classified as non-essential, with their expression levels ranking in the top 25% (Figure S1). HAP4 serves as a regulatory element, while ADE1 is a biosynthetic element. We selected HAP4 and ADE1 to investigate whether the adaptive process varies depending on the gene's function. Gene knockout strains for HAP4 and ADE1 were constructed in *Saccharomyces cerevisiae* BY4741, respectively, which were denoted as initial knockout strains. The wild-type strain of BY4741 was defined as the original strain. We performed serial transfer experiments starting from the initial knockout strains for a period of 28 days under synthetic complete medium by providing all necessary nutrients. There were six replicates constituted by three biological replicates each with two technical replicates for each gene knockout strain. The evolved strains were examined at the midpoint (14 days) and the endpoint (28 days). The growth rates, whole genomes and transcriptomes were measured for the original strain (abbreviated as "ori"), initial knockout strains (abbreviated as "ini-evo") and evolved knockout strains at two time points (abbreviated as "mid-evo" and "end-evo", respectively) (Figure 1A).

Throughout the serial transfer experiment, both groups of knockout strains exhibited a swift augmentation in growth rate from the initial time point to the midpoint, followed by a deceleration in the trend (Figure 1B and Table S1). Between the midpoint and the endpoint of the experiment, no statistically significant differences in growth rates were observed within either group of knockout strains. This observation suggests that the strains had achieved adaptation to the prevailing conditions of cultivation. The growth rates among replicates in the original strains and initial knockout strains showed high consistency. However, the variability in growth rates increased over the duration of the cultivation process, particularly in the *hap4*Δ strains. This trend implies that the evolved strains became increasingly heterogeneous, even among biological replicates, thus allowing them to be considered as independent evolutionary lines.

The genomic analysis of the evolved strains further confirmed the heterogeneity, uncovering distinct temporal genomic variations in comparison to the original strain (Figure 1C and Table S2). Notably, different clones displayed unique genomic alterations, with both *hap4*Δ and *ade1*Δ strains exhibiting variations in single nucleotide polymorphisms (SNPs) and insertions/deletions (indels). In *hap4*Δ strains, we identified three SNPs and two indels, with one SNP and one indel located near the SND3 and AAH1 promoters, respectively, and another indel linked to gene YDR544C. The remaining alterations were found within intergenic regions. The *ade1*Δ strains presented four SNPs and four indels, including one SNP associated with gene YDR544C, one indel near the MRL1 promoter, with the others situated in intergenic areas. There was scarcely any overlap between the clones or strains, except for one indel in a *hap4*Δ strain and one SNP in an *ade1*Δ strain, both linked

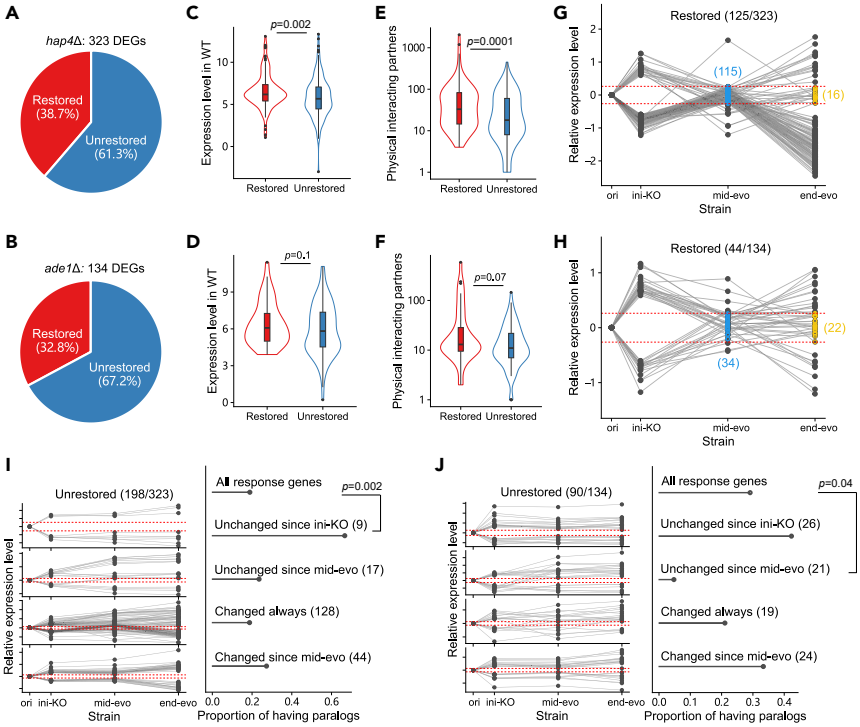


Figure 2. The expression changes of response genes during the ALE experiment

(A and B) Proportions of ever restored and unrestored response genes in evolved *hap4*Δ strains (A) and evolved *ade1*Δ strains (B), respectively.

(C and D) Violin plots of the expression levels in WT of restored and unrestored genes for *hap4*Δ (C) and *ade1*Δ (D), respectively. The values were derived from the mean value of replicates of the original strains. The solid boxes within the plots represent the interquartile range, spanning from the first quartile to the third quartile. The horizontal lines within these boxes indicate the median values. Statistical significance was assessed using the Wilcoxon test, and the corresponding *p* values were labeled.

(E and F) Violin plots of the number of physical interacting partners of restored and unrestored genes for *hap4*Δ (E) and *ade1*Δ(F). The solid boxes within the plots represent the interquartile range, spanning from the first quartile to the third quartile. The horizontal lines within these boxes indicate the median values. Statistical significance was assessed using the Wilcoxon test, and the corresponding *p* values were labeled.

(G and H) The relative expression changes of restored response genes in *hap4*Δ (G) and *ade1*Δ (H). The restored time point of response genes in *hap4*Δ and *ade1*Δ are colored by blue and yellow, respectively. The red dashed lines represent the range of $|\log_2 FC| < 0.263$, which is the scope of restoration.

(I and J) The relative expression changes of unrestored response genes in *hap4*Δ (I) and *ade1*Δ (J) can be classified into four categories, with the number of response genes in each category shown in parentheses and the proportion of genes having paralogs shown by stick charts. Statistical significances of the proportion of genes having paralogs between corresponding groups were tested by Chi-squared test and *p* values were labeled.

to gene YDR544C, which encodes a high-affinity glutamine permease. The annotations of these genomic variations could hardly link to the knockout genes directly. This might be explained by the fact that the deleted genes are non-essential and the experimental conditions did not impose any nutritional deficiencies. Interestingly, we observed a lack of continuity over time in the detection of four SNPs/indels (XIV_425434, II_452356, XIV_359345, and V_117048). The allele states for these four genomic variations were all found to be polymorphic. This suggested that the cell population at the examined time point may have exhibited a slight degree of heterogeneity. It also indicated that existing variations may disappear, with the possibility of new variations appearing at the next time point.

Constant expression changes of most response genes during ALE

Our initial focus was on the response genes of gene knockout, defined as the differentially expressed genes (DEGs) in the initial knockout strains compared to the original strains. These were identified using DESeq2[25] under a stringent statistical threshold ($|\log_2 FC| > 0.585$, $p_{adj} < 0.01$). We found that 323 genes responded to the HAP4 knockout and 134 genes to the ADE1 knockout at the expression level (Tables S3 and S4). Given the observed heterogeneity in genome and growth rate among the evolved strains, we sought to explore whether common patterns of expression changes for these response genes would emerge during adaptive evolution. We continued to monitor the expression levels of these genes in the evolved strains by analyzing their expression changes relative to the original strains.

We classified response genes that reverted to the original strain's range of expression levels at any time point ($|\log_2 FC| < 0.263$) as "ever restored". According to this definition, 61.3% (198/323) of *hap4*Δ response genes and 67.2% (90/134) of *ade1*Δ response genes were found to be never restored ("unrestored") (Figures 2A and 2B). To account for the heterogeneity in the evolved strains, we also compared each

replicate independently to the original strains. We defined the restored state as $|\log_2 FC| < 0.263$ in at least half of the replicates (Methods). Unrestored response genes remained dominant. The proportions of unrestored genes were 59.4% (192/323) for *hap4Δ* response genes and 73.1% (98/134) for *ade1Δ* response genes (Figure S2). Typically, genes that were restored exhibited higher expression levels in the wild-type strains and had a larger number of physical interaction partners compared to unrestored genes, particularly noticeable in the *hap4*Δ strains (Figures 2C–2F). These findings suggest that genes of differing functional importance display variable responses throughout the ALE process, with highly expressed and well-connected genes in the expression network showing resilience to perturbations. However, the restored state of response genes appeared unstable. In the evolved strains of *hap4*Δ, 110 out of 114 genes that were initially restored at 14 days reverted to an unrestored status at 28 days, and only 16 genes were restored at the endpoint (Figure 2G). Similarly, in the evolved strains of *ade1*Δ, 22 out of 34 initially restored genes experienced a similar reversion (Figure 2H).

In the unrestored category, nine genes in the *hap4*Δ strain were exclusively responsive immediately post-knockout and remained unchanged thereafter. Notably, six of these genes possess functionally redundant counterparts or paralogs, representing a 3.5-fold enrichment compared to all responsive genes (Chi-squared test $p = 0.002$, Figure 2I and Table S5). This pattern suggests that knocking out HAP4 may initiate a compensatory redistribution of expression among redundant genes, achieving a new equilibrium during the ALE process. Conversely, for *ade1*Δ strains, this compensatory signal was less pronounced among 26 response genes unchanged since initial knockout, showing only a 1.5-fold enrichment (Figure 2J and Table S6). However, a significant exclusion signal for *ade1*Δ strains was detected for unrestored genes whose expression levels remained stable since the midpoint of evolution; only one out of 21 such genes had a paralog, contrasting with 39 out of 134 response genes (Chi-squared test $p = 0.04$, Figure 2J and Table S6), indicating a distinct adaptive response mechanism for these genes compared to those discussed previously.

Notably, in *hap4Δ* strains, 128 response genes exhibited continuous changes in expression, accounting for 64.6% of the unrestored genes, with over 40% implicated in ribosomal biogenesis (Table S5). In *ade1*Δ strains, 19 response genes fell into this category, predominantly enriched in transmembrane transporter activity by Gene Ontology (GO) analysis (Table S7). Additionally, we identified distinct subsets of 44 and 24 genes in *hap4*Δ and *ade1*Δ strains, respectively, that exhibited no expression changes at 14 days but showed significant alterations at 28 days, both enriched in catalytic activity by GO analysis (Table S7). These findings underscore the nuanced and multifaceted nature of gene expression dynamics in response to different gene knockouts, and highlight a dynamic pattern of change as a central theme in the effects of knockouts during ALE.

Global transcriptomic states shift during ALE

A detailed examination of global transcriptomic dynamics of knockout strains across the culturing time was undertaken. We projected the transcriptomic data from all knockout strains, along with the original strains, onto a two-dimensional Uniform Manifold Approximation and Projection (UMAP)[26] for *hap4*Δ and *ade1*Δ, respectively (Figures 3A and 3B and Table S8). The resulting transcriptomic landscape revealed four distinct clusters, each characterized by unique time points. This was particularly pronounced in the *hap4*Δ strains, whose transcriptomes progressively diverged from the original strains in a pattern that correlated with the duration of culturing time. A similar pattern was observed in *ade1*Δ knockout strains, albeit with a higher degree of intra-replicate variability.

We identified DEGs between adjacent time points using the same statistical threshold as used for defining response genes (Tables S3 and S4). In concordance with the UMAP plot, a higher number of DEGs were detected in *hap4Δ* as well as in *ade1Δ*. This was observed both when comparing the evolved strains at the midpoint to the initial knockout strains (mid-evo vs. ini-evo), and when comparing the evolved strains at the endpoint to those at the midpoint (end-evo vs. mid-evo) (Figures 3C and 3D). Importantly, significant expression variations were observed across a large quantity of genes in evolved strains for both gene knockouts, compared to the number of response genes. This observation underscores a comprehensive shift in the transcriptomic landscape rather than changes confined to the knockout genes.

The intersection of DEGs from ini-evo vs. ori and mid-evo vs. ini-evo comprised 195 and 38 genes for *hap4*Δ and *ade1*Δ, respectively. The fold enrichments, compared to expected overlaps, were 2.4 and 1.8, respectively (Kolmogorov-Smirnov test $p = 0.01$ and 0.01, respectively, Methods). This level of enrichment effectively rules out the possibility of the overlaps occurring by chance. Furthermore, the overlaps between DEGs from mid-evo vs. ini-evo and end-evo vs. mid-evo was 655 and 166 for *hap4*Δ and *ade1*Δ, respectively, representing fold-enrichments of 3.9 and 2.1 over expected values (Kolmogorov-Smirnov test $p = 0.005$ and 0.03, respectively), which were also not coincidental. Intriguingly, the majority of these intersecting DEGs exhibited reverse directional changes in the two compared strain sets, confirming the dynamic shift between different time points (Figures 3E and 3F). Noticeably, ribosomal related genes occupied a considerable proportion in DEGs of knockout strains between adjacent time points. This observation prompted us to further investigate the dynamics of ribosome biogenesis.

Ribosomal related genes are deeply involved in the late stage of ALE

Ribosome biogenesis is intimately tied to growth rate.[27,28] We explored its dynamics in relation to the evolution process by analyzing the expression variances of 384 key genetic components involved in ribosome synthesis, assembly, and ribosomal RNA (rRNA) decay. These included genes which code 138 ribosomal proteins, 161 ribosomal assembly factors, 15 rRNAs, and 70 small nucleolar RNAs (snoRNAs) (Table S9). We found that ribosomal related genes made up 29.4% (95/323) of all initial response genes in *hap4*Δ, which might due to the function in the diauxic shift of HAP4 (Figure 4A). Contrastingly, in *ade1*Δ, ribosomal-related response genes comprised a mere 4.4% (6/134), highlighting a striking disparity (Figure 4B). The proportion of ribosomal related genes in DEGs increased in conjunction with the ALE process (Figures 4A and 4B). This suggests ribosome biogenesis has an influential role in the adaptive evolutionary process.

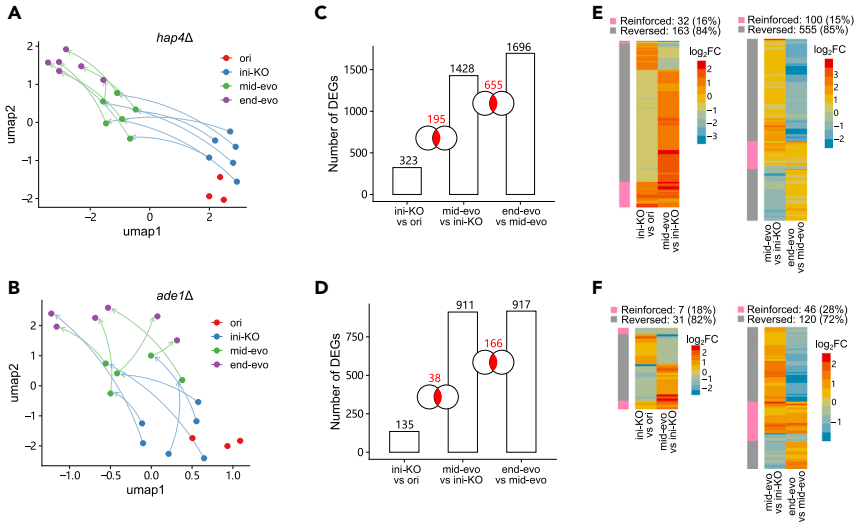


Figure 3. Global transcriptomic dynamics of knockout strains across the culturing time

(A and B) Two-dimensional UMAP plots of the transcriptome, estimated by log_2TPM, for *hap4*Δ (A) and *ade1*Δ (B) knockout strains, respectively. Arrows connect the states of the same replicate at different timepoints.

(C and D) The numbers of DEGs between adjacent timepoints were shown for *hap4Δ* (C) and *ade1Δ* (D). These comparing groups include: initial KO strains vs. original strains, evolved strains at the midpoint vs. initial KO strains, and evolved strains at the endpoint vs. evolved strains at the midpoint. Overlapping DEGs between neighboring groups are highlighted in red.

(E and F) Heatmaps represent the expression changes of these overlapping genes in each comparing group of *hap4*Δ (E) and *ade1*Δ (F), respectively. Overlapping DEGs with reverse directional changes in the two compared strain sets are designated as "reversed," while those with the same directional changes are designated as "reinforced." The numbers of these DEGs were counted and labeled.

The expression levels of most ribosomal protein genes changed very infrequently from the original strains to the midpoint evolved strains both in *hap4*Δ and *ade1*Δ, but a high proportion of these proteins, 67.4% (93/138) and 60.9% (84/138) respectively, were down-regulated in the evolved strains at the endpoint (Figures 4C and 4D). Regarding ribosomal assembly factors in evolved strains of *hap4*Δ, while there was an upregulation of 76.4% (123/161) of ribosomal assembly factors in midpoint evolved strains, 90.1% (145/161) of these genes were down-regulated at the endpoint (Figure 4E). The trends noted for evolved strains of *ade1*Δ were similar, albeit at a reduced scale (Figure 4F). Despite a smaller proportion of ribosomal assembly factors displaying changes in their expression at the midpoint, all of these genes were up-regulated, as seen in *hap4*Δ. However, at the endpoint, the domination of changed ribosomal assembly factors in expression persisted at 69.6% (112/161), with all of them marked by downregulation. These findings suggest that ribosome biogenesis may be active during the early adaptation stages of knockout strains but decelerate later on.

However, we noticed that the expression of rRNA genes, another core component of ribosome, showed a different pattern. For both gene knockouts, the rRNA genes were upregulated in the mid-evolved strains apparently (Figures 4G and 4H). A similar upregulation pattern was also noticed for snoRNAs in the midpoint evolved strains. Specifically, 71.4% (50/70) of snoRNAs in *hap4*Δ and 61.4% (43/70) in *ade1*Δ were up-regulated (Figures 4I and 4J). Since the transcriptome was measured according to the routine RNA-seq experiment, in which most rRNAs were removed by selection of polyadenylated RNA transcripts using oligo (dT) primers theoretically, the observed variations in rRNAs essentially reflect changes in polyadenylated forms of these RNAs, rather than the total rRNA pool. We further calculated proportions of reads aligned to rRNAs in the strains at each time point (Figure S3). Interestingly, we observed a shift in these proportions over the course of the experiment. In the original strains, 6.2 $\pm$ 0.5% of reads aligned to rRNAs, which is typical for standard RNA-seq experiments. However, in the evolved strains at the endpoint, these proportions changed dramatically to 48.1 $\pm$ 8% for *hap4*Δ and 42.8 $\pm$ 6% for *ade1*Δ. This substantial increase in the proportion of reads aligning to rRNAs suggests a significant change in the abundance of polyadenylated rRNA forms.

Similarly, the variation observed in snoRNA levels corresponds to this scenario, indicating alterations in their polyadenylated forms as well. The polyadenylation of non-coding RNAs, including rRNA and snoRNA, is linked to RNA quality control mechanisms,[29–31] suggesting that degradation processes might be active.

DISCUSSION

Gene loss, a common occurrence throughout evolution, is believed to enhance an organism's ability to adapt and evolve.[9,32] Numerous studies have explored the compensatory mechanisms and phenotypic outcomes of gene loss.[15,33,34] In this study, we conducted a 28-day serial transfer experiment involving two gene knockout yeast strains. We observed an increased growth rate among all strains, with rapid augmentation in both knockout strain groups until the midpoint of the experiment. The growth rates of the initial knockout strains were slightly faster than those of the original strains, although previous studies have found that the deletion of HAP4 and ADE1 can be detrimental

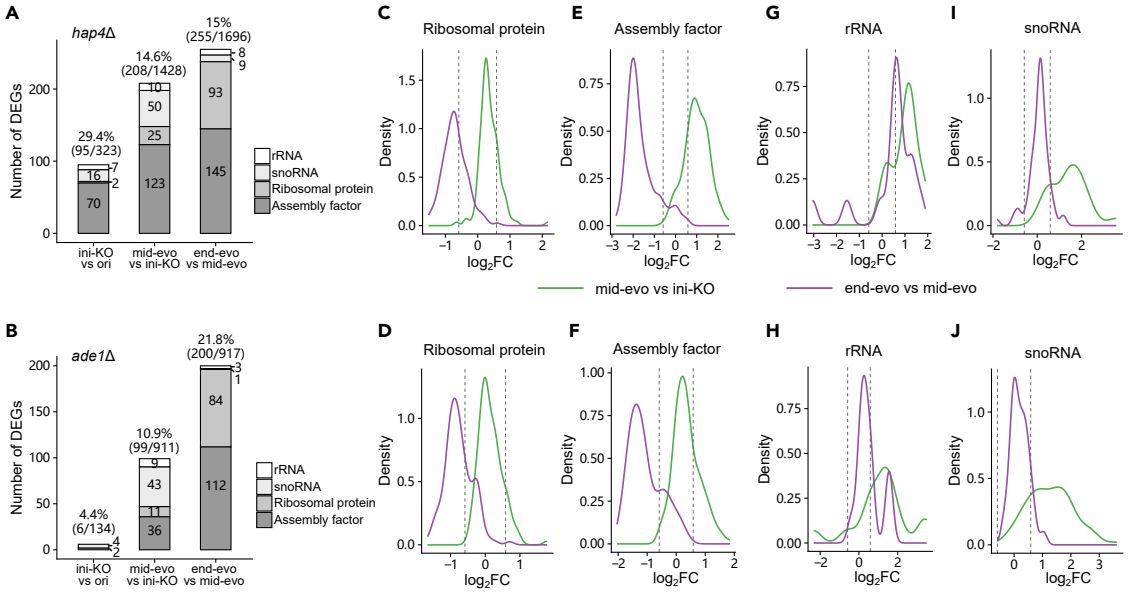


Figure 4. Dynamics of Ribosome-related genes during the ALE experiment

(A and B) The numbers and proportions of ribosome-related genes in DEGs between adjacent timepoints are shown for *hap4*Δ (A) and *ade1*Δ (B). Ribosome-related genes include rRNA, ribosomal protein genes, assembly factor genes, and snoRNA, which are indicated by different colors. The quantity of genes in each category is also clearly labeled.

(E–J) Density curves represent the expression changes for each group of genes derived from evolved strains at the midpoint compared to initial KO strains (green), and from evolved strains at the endpoint compared to evolved strains at the midpoint (purple). The upper row shows the expression changes in *hap4*Δ (C, E, G, I), and the lower row shows the expression changes in *ade1*Δ (D, F, H, and J). The gray dashed lines represent the range of $|\log_2 FC| < 0.585$, and values outside this range indicate significant changes.

to the growth rate in several *S. cerevisiae* strains.[35–37] This discrepancy may be attributed to several factors. Variations in experimental conditions and growth rate measurement methods may lead to inconsistent results. Previous studies often focused on specific contexts, such as the effects of gene deletion during fermentation or competitive growth among different knockout strains. Additionally, we cannot exclude the possibility of genomic variations in the original strains, potentially arising from long-term laboratory preservation, which might have altered their growth characteristics. Nevertheless, our parallel comparison experimental design ensured that these potential confounding factors did not affect our findings.

A significant discovery was the dynamic change in the transcriptome of evolved knockout strains, as opposed to maintaining a stable state in both knockouts. We identified this from two perspectives. From the viewpoint of response genes of gene knockout, only about one-third was ever restored to the expression levels in the original strains. The genes that did return to their original expression levels are those that are highly expressed and central within the expression network, suggesting a necessity to preserve the stability of these hub genes. The characteristics of the genes that were not restored vary depending on which gene was deleted, highlighting the specific nature of these responses. From the standpoint of the whole transcriptome, we noted a substantial quantity of DEGs of knockout strains between adjacent time points, accounting for 15.7%–27.6% of all genes. These results propose that the adaptive consequence of gene loss may not be unique and that there may exist multiple suboptimal states of transcriptome throughout evolution. Previous studies supporting diverse phenotypes after adaptation corroborate this suggestion.[15,21]

Another intriguing observation was the dynamic alteration in ribosome biogenesis, particularly regarding the polyadenylation of rRNA. Typically, the proportion of polyadenylated rRNA is exceedingly low, leading to the assumption that routine RNA sequencing offers minimal insight into rRNA characteristics. Kuai et al. revealed that 0.01%–0.1% of 25S rRNA was polyadenylated in the normal strain of *S. cerevisiae*.[29] However, studies in recent twenty years have found some rRNAs can be also polyadenylated in *Arabidopsis thaliana*, mice, human cells, and so on.[38–41]

Contrary to the role of the poly(A) tail in mRNA, polyadenylated rRNA is typically associated with rRNA degradation.[42–44] Remarkable polyadenylated rRNAs have been observed in cells defective in nuclear RNA degradation pathway, exosome in particular.[29,45] For instance, deleting RRP6, a gene encoding a riboexonuclease component of the exosome, led to a hundredfold increase in polyadenylated 25S rRNA in yeast compared to the wild-type strain.[29] Additionally, the polyadenylation has been identified as playing a critical role in the quality control of snoRNAs, at least in yeast.[42,46] To mitigate potential technical biases, we conducted an additional serial transfer experiment on the *hap4*Δ strain over two weeks, employing a different approach for cell collection and RNA extraction (Methods). This experiment also revealed a significant increase in polyadenylated rRNAs of evolved strains at 14 days (Figure S4).

Interestingly, a high proportion of upregulation of polyadenylated rRNAs and snoRNAs, as well as ribosomal assembly factors was observed in both knockout strains at the midpoint, coupled with downregulation ribosomal protein genes and ribosomal assembly factors at the endpoint. These findings provide valuable insights into the temporal coordination of ribosomal components, including the transcription of rRNA, the production of ribosomal proteins, and the synthesis of assembly factors. Ribosome biogenesis might be active during the early stages of the knockout strains' adaptation. However, maintaining such an elevated level of ribosome biogenesis could potentially disrupt the balance of protein synthesis within cells.[47] Over-proliferation might result in an excessively rapid depletion of environmental nutrients. As time progresses, it becomes necessary for cells to regulate ribosome biogenesis. A plausible regulatory method could involve reducing ribosome production and degrading rRNAs by polyadenylation. The pattern of accelerated ribosome biogenesis followed by a deceleration aligned with the growth curve we observed, suggesting a dynamic process that adjusts to environmental conditions and cellular needs over time. Furthermore, it is observed that the variations in the components involved in ribosome biogenesis are not synchronized, indicating meticulous regulation within the process.

In conclusion, we demonstrate that the transcriptomes of evolved strains at different stages fail to fully revert to their original state, and ribosome biogenesis emerges as a potentially crucial player in the adaptative response to gene knockout. These results significantly enhance our understanding of adaptive evolution in gene knockouts and offer valuable insights into how genetic perturbations affect various biological systems.

Limitation of the study

Contrary to typical expectations, we did not find any distinct genomic variations directly related to the knockout genes. However, recent studies have shown that mutations can stochastically accumulate in the evolutionary process,[48,49] indicating that each gene knockout presents unique circumstances and outcomes.[14] Considering that HAP4 and ADE1 are non-essential genes, and the culture conditions involved a complete nutrient medium with a limited cultivation period, there was probably not strong selective pressure to fix specific compensatory mutations. Our current data do not provide a comprehensive quantification of the abundance of polyadenylated rRNAs and snoRNAs relative to the entire pool of rRNAs and snoRNAs. This limitation restricts our ability to fully assess the magnitude and significance of the observed upregulation in the context of overall cellular RNA composition. While the study observes changes in gene expression related to ribosome biogenesis, we are unable to provide detailed molecular insights into these changes and their direct impact on cellular processes. Further investigation is needed to elucidate the specific mechanisms underlying the observed transcriptomic shifts.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact Li Liu: liul47@mail.sysu.edu.cn.

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The transcriptomic data and Whole genome sequencing data have been deposited at NCBI and is publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this work is available from the lead contact upon request.

ACKNOWLEDGMENTS

We are grateful to financial support from National Natural Science Foundation of China (32070687) and the Fundamental Research Funds for the Central Universities of Sun Yat-sen University (22lgqb28). We are grateful to helpful discussions with Dr. Xionglei He, Chang Ye, and Kehui Liu.

AUTHOR CONTRIBUTIONS

L.L. and B.J. designed the study; B.J. and C.Y.X. conducted the experiments and data analysis; L.L. and B.J. wrote the paper. All authors have read and agreed to the published version of the manuscript.

DECLARATION OF INTERESTS

The authors declare that they have no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Yeast gene deletions

- Serial transfer experiment
- Growth rate measurement
- Whole-genome sequencing and data analysis
- RNA extraction and sequencing
- RNA-seq data analysis
- Calculation of fold enrichment of the intersection of DEGs
- Gene annotations and features analysis

- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.111219.

Received: July 4, 2024
Revised: August 29, 2024
Accepted: October 17, 2024
Published: October 21, 2024

REFERENCES

1. Housden, B.E., Muhar, M., Gemberling, M., Gersbach, C.A., Stainier, D.Y.R., Seydoux, G., Mohr, S.E., Zuber, J., and Perrimon, N. (2017). Loss-of-function genetic tools for animal models: cross-species and cross-platform differences. Nat. Rev. Genet. *18*, 24–40. https://doi.org/10.1038/nrg.2016.118.
2. Kovacs, K., Farkas, Z., Bajic, D., Kalapis, D., Daraba, A., Almasi, K., Kintses, B., Bodi, Z., Notebaart, R.A., Poyatos, J.F., et al. (2021). Suboptimal Global Transcriptional Response Increases the Harmful Effects of Loss-of-Function Mutations. Mol. Biol. Evol. *38*, 1137–1150. https://doi.org/10.1093/molbev/msaa280.
3. Liu, L., Liu, M., Zhang, D., Deng, S., Chen, P., Yang, J., Xie, Y., and He, X. (2020). Decoupling gene functions from knockout effects by evolutionary analyses. Natl. Sci. Rev. *7*, 1169–1180. https://doi.org/10.1093/nsr/nwaa079.
4. Teng, X., Dayhoff-Brannigan, M., Cheng, W.-C., Gilbert, C.E., Sing, C.N., Diny, N.L., Wheelan, S.J., Dunham, M.J., Boeke, J.D., Pineda, F.J., and Hardwick, J.M. (2013). Genome-wide Consequences of Deleting Any Single Gene. Mol. Cell *52*, 485–494. https://doi.org/10.1016/j.molcel.2013.09.026.
5. Jakutis, G., and Stainier, D.Y.R. (2021). Genotype-Phenotype Relationships in the Context of Transcriptional Adaptation and Genetic Robustness. Annu. Rev. Genet. *55*, 71–91. https://doi.org/10.1146/annurev-genet-071719-020342.
6. Sztal, T.E., and Stainier, D.Y.R. (2020). Transcriptional adaptation: a mechanism underlying genetic robustness. Development *147*, dev186452. https://doi.org/10.1242/dev.186452.
7. Ho, W.C., and Zhang, J. (2018). Evolutionary adaptations to new environments generally reverse plastic phenotypic changes. Nat. Commun. *9*, 350. https://doi.org/10.1038/s41467-017-02724-5.
8. Levis, N.A., Isdaner, A.J., and Pfennig, D.W. (2018). Morphological novelty emerges from pre-existing phenotypic plasticity. Nat. Ecol. Evol. *2*, 1289–1297. https://doi.org/10.1038/s41559-018-0601-8.
9. Huelsmann, M., Hecker, N., Springer, M.S., Gatesy, J., Sharma, V., and Hiller, M. (2019). Genes lost during the transition from land to water in cetaceans highlight genomic changes associated with aquatic adaptations. Sci. Adv. *5*, eaaw6671. https://doi.org/10.1126/sciadv.aaw6671.
10. Lalejini, A., Ferguson, A.J., Grant, N.A., and Ofria, C. (2021). Adaptive Phenotypic Plasticity Stabilizes Evolution in Fluctuating Environments. Front. Ecol. Evol. *9*, 715381. https://doi.org/10.3389/fevo.2021.715381.
11. Persson, K., Stenberg, S., Tamás, M.J., and Warringer, J. (2022). Adaptation of the yeast gene knockout collection is near-perfectly predicted by fitness and diminishing return epistasis. G3-Genes Genom Genet *12*, jkac240. https://doi.org/10.1093/g3journal/jkac240.
12. Kavvas, E.S., Long, C.P., Sastry, A., Poudel, S., Antoniewicz, M.R., Ding, Y., Mohamed, E.T., Szubin, R., Monk, J.M., Feist, A.M., and Palsson, B.O. (2022). Experimental Evolution Reveals Unifying Systems-Level Adaptations but Diversity in Driving Genotypes. mSystems *7*, e0016522. https://doi.org/10.1128/msystems.00165-22.
13. Long, C.P., Gonzalez, J.E., Feist, A.M., Palsson, B.O., and Antoniewicz, M.R. (2018). Dissecting the genetic and metabolic mechanisms of adaptation to the knockout of a major metabolic enzyme in. Proc. Natl. Acad. Sci. USA *115*, 222–227. https://doi.org/10.1073/pnas.1716056115.
14. Helsen, J., Voordeckers, K., Vanderwaeren, L., Santermans, T., Tsontaki, M., Verstrepen, K.J., and Jelier, R. (2020). Gene Loss Predictably Drives Evolutionary Adaptation. Mol. Biol. Evol. *37*, 2989–3002. https://doi.org/10.1093/molbev/msaa172.
15. Szamecz, B., Boross, G., Kalapis, D., Kovács, K., Fekete, G., Farkas, Z., Lázár, V., Hrtyan, M., Kemmeren, P., Groot Koerkamp, M.J.A., et al. (2014). The Genomic Landscape of Compensatory Evolution. PLoS Biol. *12*, e1001935. https://doi.org/10.1371/journal.pbio.1001935.
16. Johnson, M.S., Gopalakrishnan, S., Goyal, J., Dillingham, M.E., Bakerlee, C.W., Humphrey, P.T., Jagdish, T., Jerison, E.R., Kosheleva, K., Lawrence, K.R., et al. (2021). Phenotypic and molecular evolution across 10,000 generations in laboratory budding yeast populations. Elife *10*, e63910. https://doi.org/10.7554/eLife.63910.
17. Ament-Velasquez, S.L., Gilchrist, C., Rego, A., Bendixsen, D.P., Brice, C., Grosse-Sommer, J.M., Rafati, N., and Stelkens, R. (2022). The Dynamics of Adaptation to Stress from Standing Genetic Variation and de novo Mutations. Mol. Biol. Evol. *39*, msac242. https://doi.org/10.1093/molbev/msac242.
18. McCloskey, D., Xu, S., Sandberg, T.E., Brunk, E., Hefner, Y., Szubin, R., Feist, A.M., and Palsson, B.O. (2018). Evolution of gene knockout strains of E. coli reveal regulatory architectures governed by metabolism. Nat. Commun. *9*, 3796. https://doi.org/10.1038/s41467-018-06219-9.
19. Hsu, P.C., Cheng, Y.H., Liao, C.W., Litan, R.R.R., Jhou, Y.T., Opoc, F.J.G., Amine, A.A.A., and Leu, J.Y. (2023). Rapid evolutionary repair by secondary perturbation of a primary disrupted transcriptional network. EMBO Rep. *24*, e56019. https://doi.org/10.15252/embr.202256019.
20. Sandberg, T.E., Pedersen, M., LaCroix, R.A., Ebrahim, A., Bonde, M., Herrgard, M.J., Palsson, B.O., Sommer, M., and Feist, A.M. (2014). Evolution of Escherichia coli to 42 degrees C and subsequent genetic engineering reveals adaptive mechanisms and novel mutations. Mol. Biol. Evol. *31*, 2647–2662. https://doi.org/10.1093/molbev/msu209.
21. McCloskey, D., Xu, S., Sandberg, T.E., Brunk, E., Hefner, Y., Szubin, R., Feist, A.M., and Palsson, B.O. (2018). Multiple Optimal Phenotypes Overcome Redox and Glycolytic Intermediate Metabolite Imbalances in Escherichia coli pgi Knockout Evolutions. Appl. Environ. Microbiol. *84*, e00823-18. https://doi.org/10.1128/AEM.00823-18.
22. Forsburg, S.L., and Guarente, L. (1989). Identification and characterization of HAP4: a third component of the CCAAT-bound HAP2/HAP3 heteromer. Genes Dev. *3*, 1166–1178. https://doi.org/10.1101/gad.3.8.1166.
23. Mao, Y., and Chen, C. (2019). The Hap Complex in Yeasts: Structure, Assembly Mode, and Gene Regulation. Front. Microbiol. *10*, 1645. https://doi.org/10.3389/fmicb.2019.01645.
24. Myasnikov, A.N., Sasnauskas, K.V., Janulaitis, A.A., and Smirnov, M.N. (1991). The Saccharomyces cerevisiae ADE1 gene: structure, overexpression and possible regulation by general amino acid control. Gene *109*, 143–147. https://doi.org/10.1016/0378-1119(91)90600-g.
25. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.

Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.
26. McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. Preprint at arXiv. https://doi.org/10.48550/arXiv.1802.03426.
27. Shore, D., Zencir, S., and Albert, B. (2021). Transcriptional control of ribosome biogenesis in yeast: links to growth and stress signals. Biochem. Soc. Trans. *49*, 1589–1599. https://doi.org/10.1042/BST20201136.
28. Lempiainen, H., and Shore, D. (2009). Growth control and ribosome biogenesis. Curr. Opin. Cell Biol. *21*, 855–863. https://doi.org/10.1016/j.ceb.2009.09.002.
29. Kuai, L., Fang, F., Butler, J.S., and Sherman, F. (2004). Polyadenylation of rRNA in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA *101*, 8581–8586. https://doi.org/10.1073/pnas.0402888101.
30. Shcherbik, N., Wang, M., Lapik, Y.R., Srivastava, L., and Pestov, D.G. (2010). Polyadenylation and degradation of incomplete RNA polymerase I transcripts in mammalian cells. EMBO Rep. *11*, 106–111. https://doi.org/10.1038/embor.2009.271.
31. Reinisch, K.M., and Wolin, S.L. (2007). Emerging themes in non-coding RNA quality control. Curr. Opin. Struct. Biol. *17*, 209–214. https://doi.org/10.1016/j.sbi.2007.03.012.
32. Sharma, V., Hecker, N., Roscito, J.G., Foerster, L., Langer, B.E., and Hiller, M. (2018). A genomics approach reveals insights into the importance of gene losses for mammalian adaptations. Nat. Commun. *9*, 1215. https://doi.org/10.1038/s41467-018-03667-1.
33. Farkas, Z., Kovács, K., Sarkadi, Z., Kalapis, D., Fekete, G., Birtyik, F., Ayaydin, F., Molnár, C., Horváth, P., Pál, C., and Papp, B. (2022). Gene loss and compensatory evolution promotes the emergence of morphological novelties in budding yeast. Nat. Ecol. Evol. *6*, 763–773. https://doi.org/10.1038/s41559-022-01730-1.
34. Koch, L. (2022). Turning gene loss into phenotypic gain. Nat. Rev. Genet. *23*, 393. https://doi.org/10.1038/s41576-022-00504-6.
35. Di Noia, M.A., Scarcia, P., Agrimi, G., Ocheja, O.B., Wahid, E., Pisano, I., Paradies, E., Palmieri, L., Guaragnella, C., and Guaragnella, N. (2023). Inactivation of HAP4 Accelerates RTG-Dependent Osmoadaptation in Saccharomyces cerevisiae. Int. J. Mol. Sci. *24*, 5320. https://doi.org/10.3390/ijms24065320.
36. Raghevendran, V., Patil, K.R., Olsson, L., and Nielsen, J. (2006). Hap4 Is Not Essential for Activation of Respiration at Low Specific Growth Rates in Saccharomyces cerevisiae. J. Biol. Chem. *281*, 12308–12314. https://doi.org/10.1074/jbc.M512972200.
37. Qian, W., Ma, D., Xiao, C., Wang, Z., and Zhang, J. (2012). The genomic landscape and evolutionary resolution of antagonistic pleiotropy in yeast. Cell Rep. *2*, 1399–1410. https://doi.org/10.1016/j.celrep.2012.09.017.
38. Slomovic, S., Fremder, E., Staals, R.H.G., Pruijn, G.J.M., and Schuster, G. (2010). Addition of poly (A) and poly (A)-rich tails during RNA degradation in the cytoplasm of human cells. Proc. Natl. Acad. Sci. USA *107*, 7407–7412. https://doi.org/10.1073/pnas.0910621107.
39. Lange, H., Holec, S., Cognat, V., Pieuchot, L., Le Ret, M., Canaday, J., and Gagliardi, D. (2008). Degradation of a polyadenylated rRNA maturation by-product involves one of the three RRP6-like proteins in Arabidopsis thaliana. Mol. Cell Biol. *28*, 3038–3044. https://doi.org/10.1128/MCB.02064-07.
40. Slomovic, S., Laufer, D., Geiger, D., and Schuster, G. (2006). Polyadenylation of ribosomal RNA in human cells. Nucleic Acids Res. *34*, 2966–2975. https://doi.org/10.1093/nar/gkl357.
41. Chang, Y., Stockinger, M.P., Tashiro, H., and Lin, C.l.G. (2008). A novel noncoding RNA rescues mutant SOD1-mediated cell death. Faseb. J. *22*, 691–702. https://doi.org/10.1096/fj.07-9532com.
42. Van Hoof, A., Lennertz, P., and Parker, R. (2000). Yeast exosome mutants accumulate 3′-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. Mol. Cell Biol. *20*, 441–452. https://doi.org/10.1128/Mcb.20.2.441-452.2000.
43. Fang, F., Hoskins, J., and Butler, J.S. (2004). 5-fluorouracil enhances exosome-dependent accumulation of polyadenylated rRNAs. Mol. Cell Biol. *24*, 10766–10776. https://doi.org/10.1128/Mcb.24.24.10766-10776.2004.
44. Win, T.Z., Draper, S., Read, R.L., Pearce, J., Norbury, C.J., and Wang, S.W. (2006). Requirement of fission yeast Cid14 in polyadenylation of rRNAs. Mol. Cell Biol. *26*, 1710–1721. https://doi.org/10.1128/Mcb.26.5.1710-1721.2006.
45. Kadaba, S., Wang, X., and Anderson, J.T. (2006). Nuclear RNA surveillance in Saccharomyces Cerevisiae: Trf4p-dependent polyadenylation of nascent hypomethylated tRNA and an aberrant form of 5S rRNA. RNA *12*, 508–521. https://doi.org/10.1261/rna.2305406.
46. Egecioglu, D.E., Henras, A.K., and Chanfreau, G.F. (2006). Contributions of Trf4p- and Trf5p-dependent polyadenylation to the processing and degradative functions of the yeast nuclear exosome. RNA *12*, 26–32. https://doi.org/10.1261/rna.2207206.
47. Shore, D., and Albert, B. (2022). Ribosome biogenesis and the cellular energy economy. Curr. Biol. *32*, R611–R617. https://doi.org/10.1016/j.cub.2022.04.083.
48. Jerison, E.R., Nguyen Ba, A.N., Desai, M.M., and Kryazhimskiy, S. (2020). Chance and necessity in the pleiotropic consequences of adaptation for budding yeast. Nat. Ecol. Evol. *4*, 601–611. https://doi.org/10.1038/s41559-020-1128-3.
49. Klim, J., Zielenkiewicz, U., Skoneczny, M., Skoneczna, A., Kurlandzka, A., and Kaczanowski, S. (2021). Genetic interaction network has a very limited impact on the evolutionary trajectories in continuous culture-grown populations of yeast. BMC Ecol. Evol. *21*, 99. https://doi.org/10.1186/s12862-021-01830-9.
50. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.
51. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
52. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
53. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.
54. Olivares-Marin, I.K., González-Hernández, J.C., Regalado-Gonzalez, C., and Madrigal-Perez, L.A. (2018). Saccharomyces cerevisiae Exponential Growth Kinetics in Batch Culture to Analyze Respiratory and Fermentative Metabolism. JoVE *25*, e58192. https://doi.org/10.3791/58192.
55. Zhu, A., Ibrahim, J.G., and Love, M.I. (2019). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Bioinformatics *35*, 2084–2092.
56. Wapinski, I., Pfeffer, A., Friedman, N., and Regev, A. (2007). Natural history and evolutionary principles of gene duplication in fungi. Nature *449*, 54–61. https://doi.org/10.1038/nature06107.
57. Ryu, H.Y., López-Giráldez, F., Knight, J., Hwang, S.S., Renner, C., Kreft, S.G., and Hochstrasser, M. (2018). Distinct adaptive mechanisms drive recovery from aneuploidy caused by loss of the Ulp2 SUMO protease. Nat. Commun. *9*, 5417. https://doi.org/10.1038/s41467-018-07836-0.
58. Monticolo, F., Palomba, E., and Chiusano, M.L. (2021). Translation machinery reprogramming in programmed cell death in Saccharomyces cerevisiae. Cell Death Dis. *7*, 17. https://doi.org/10.1038/s41420-020-00392-x.
59. Woolford, J.L., Jr., and Baserga, S.J. (2013). Ribosome biogenesis in the yeast Saccharomyces cerevisiae. Genetics *195*, 643–681. https://doi.org/10.1534/genetics.113.153197.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Transcriptomes of strains each timepoint	This paper	NCBI: BioProject ID PRJNA1060357
Whole genome sequencing data	This paper	NCBI: BioProject ID PRJNA1060357
Experimental models: Organisms/strains		
S. cerevisiae: strain BY4741	Xionglei He lab	N/A
Software and algorithms		
GATK (Version 4.2.0.0)	GATK	https://github.com/broadinstitute/gatk/
SnpEff (Version 4.3t)	SnpEff	https://pcingola.github.io/SnpEff/
STAR (Version 2.7.8a)	STAR	Dobin et al.[50]
FASTX Toolkit (Version 0.0.14)	FASTX Toolkit	http://hannonlab.cshl.edu/fastx_toolkit/
Trim galore (Version 0.6.4_dev)	Trim galore	https://github.com/FelixKrueger/TrimGalore
BWA (Version 0.7.17-r1198-dirty)	BWA	Li and Durbin[51]
Picard tools	Picard tools	http://picard.sourceforge.net
SAMtools (version 1.10)	SAMtools	Li et al.[52]
Featurecounts (version 1.6.2)	Featurecounts	Liao et al.[53]
DESeq2	DESeq2	Love et al.[25]
R studio (R version 4.0.3)	R studio	https://posit.co/downloads/

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

This study utilized *Saccharomyces cerevisiae* strain BY4741 (*MATa*, *his3*, *leu2*, *met15*, *ura3*), which was maintained in a −80°C freezer.

METHOD DETAILS

Yeast gene deletions

S. cerevisiae strain BY4741 (*MATa*, *his3*, *leu2*, *met15*, *ura3*) was included in this study. *HAP4* and *ADE1* were replaced respectively, by a *URA3* cassette in BY4741. The replacements were achieved by homologous recombination through standard polyethylene glycol (PEG)/LiAc-based method.[3] The transformation protocol was as follows: Yeast cells were cultured overnight in yeast extract peptone dextrose (YPD) medium at 30°C with continuous shaking at 200 RPM in darkness. Subsequently, 100 μL of the culture was inoculated into 50 mL fresh YPD medium and grown to mid-log phase (OD600 0.6–0.8). Cells were then harvested, washed with sterile water, and resuspended in 100 mM LiAc for a 10-min incubation at room temperature. For each transformation, 50 μL of cell suspension was combined with a transformation mix comprising 240 μL 50% PEG 3350, 30 μL 1M LiAc, 10 μL heat-denatured single-stranded carrier DNA, 30 μL sterile water, and 2–4 μg URA3 cassette DNA with homologous arms. This mixture underwent heat shock at 42°C for 25 min. Post-centrifugation, transformed cells were resuspended in sterile water. Then the transformants were spread on the plates of synthetic medium deprived of uracil (SC-URA, 2% glucose) and cultured at 30°C for 2 days. The positive clones were confirmed by polymerase chain reaction (PCR). For each gene deletion line, three independent clones were chosen for subsequent experiments.

Serial transfer experiment

A total of 12 samples were carried out, including 2 technical replicates for 3 positive biological clones of 2 gene deletion lines (*hap4*Δ and *ade1*Δ). All samples were incubated in 50mL conical tubes with screw caps containing 15mL of synthetic complete (SC) medium. Cultures were incubated at 30°C with continuous shaking at 200 RPM in a dark environment. The serial transfer proceeded for 28 days without stop. All samples were transferred to 15mL fresh SC medium at 50mL tube every 2 days. The dilution ratio for each transfer was $1:10^5$. The deleted loci of all samples in deletion lines were confirmed by PCR every 6 days to check status of contamination.

Growth rate measurement

Growth rate measurement were conducted for the wildtype strain, the initial knockout strains, and the evolved knockout strains at 14 days and 28 days. Strains (when saturated) at these measured days were diluted 1:200 to 15mL fresh SC medium in 50mL tube, which resulted in an

initially optical density OD600 = 0.05 (UNICOUV/VIS Spectrophotometer). All strains were incubated in the SC medium at 30°C with shaking. OD600 for each sample were measured per 2 h until the growth of samples achieved the platform period (Normally at OD600 ≥1.0). To calculate the growth rate, we focused on the values obtained during the logarithmic phase (0.1 < OD600 < 0.8). The natural logarithm of OD600 values (ln(OD600)) was plotted against time (t) for each strain. Linear regression analysis was performed using the *lm* function in R Studio to fit these data points. The growth rate for each strain was determined from the slope coefficient of the resulting linear fit.[54] The values of growth rate for each strain were listed in Table S1.

Whole-genome sequencing and data analysis

One technical replicate was randomly chosen for each biological replicate. In total 9 samples were subject to whole-genome sequencing to identify mutations. For each sample, genomic DNA was extracted from ~10^8 yeast cells by Omega Yeast DNA kit (D3370-01). The whole genome sequencing was performed using the paired-end strategy on Illumina HiSeq at Genwiz following the standard procedure.

Approximately 6 million reads were generated for each library, corresponding to an average sequencing depth of ~100×. After trimmed 5′-end 10 bp by FASTX Toolkit (Version 0.0.14; http://hannonlab.cshl.edu/fastx_toolkit/) with parameter settings (-f 11 -z) and cleaned by Trim galore (Version 0.6.4_dev; https://github.com/FelixKrueger/TrimGalore) with parameter settings (–length 100 –quality 30), remaining reads were aligned to the yeast genome by BWA[51] (Version 0.7.17-r1198-dirty) with default parameter settings, and duplicated reads were removed by Picard tools (http://picard.sourceforge.net). We used the genome of *S. cerevisiae* strain S288C as the reference (version R64-1-1; http://www.yeastgenome.org) with genome annotation file (version R64-1-1.104, genome-date 2011-09, genome-build-accession GCA_000146045.2, genebuild-last-updated 2018-10).

Single-nucleotide mutations (SNPs) and indels were called on the Genome Analysis Toolkit (GATK) platform (Version 4.2.0.0) with default settings (https://github.com/broadinstitute/gatk/). Variant annotation and effect prediction were done by SnpEff (Version 4.3t, https://pcingola.github.io/SnpEff/). SNPs were filtered by GATK with parameters (–cluster-window-size 10 –cluster-size 3 –missing-values-evaluate-*as*-failing -filter "QD < 2.0" –filter-name "QD2" -filter "QUAL <30.0" –filter-name "QUAL30" -filter "FS > 60.0" –filter-name "FS60" -filter "MQ < 40.0" –filter-name "MQ40" -filter "SOR >3.0" –filter-name "SOR3" -filter "MQRankSum < −12.5" –filter-name "MQRankSum-12.5" -filter "ReadPosRankSum < −8.0" –filter-name "ReadPosRankSum-8") and indels were filtered with parameters (-filter "QD < 2.0" –filter-name "QD2" -filter "FS > 200.0" –filter-name "FS200" -filter "QUAL <30.0" –filter-name "QUAL30" -filter "SOR >10.0" –filter-name "SOR10" -filter "MQRankSum < −12.5" –filter-name "MQRankSum-12.5" -filter "ReadPosRankSum < −8.0" –filter-name "ReadPosRankSum-8″).

RNA extraction and sequencing

For each sample involved in the serial transfer experiment, total RNA was extracted as following procedures. Strains (when saturated) were 1:100 diluted to 15mL fresh SC medium and incubated at 30°C with shaking. Cells of 8 mL culture at OD600 = 0.65–0.7 were harvested. Total RNA of all samples was extracted by QIAGEN RNeasy Plus mini kit (Cat No.74136). The sequencing was performed using the paired-end strategy on Illumina HiSeq platform at Genwiz by standard procedure.

To exclude the batch effect in RNA extraction, we performed another serial transfer experiment of *hap4*Δ line with three biological replicates. Ini-KO strains were maintained at −80°C in glycerol stocks, and were revived after 14 days. Total RNAs of these samples were extracted at the same batch and then sequenced as the above procedure.

RNA-seq data analysis

Approximately 20 million reads were generated for each sample, corresponding to an average sequencing depth of ~100×. Samples with total reads less than 15 million or detectable degradation were filtered (one replicate of the WT strains, and on replicate of *ade1*Δ at 14 days). After trimmed 5′-end 15 bp by FASTX Toolkit (Version 0.0.14; http://hannonlab.cshl.edu/fastx_toolkit/) with parameter settings (-f 16 -z) and cleaned by Trim galore (Version 0.6.4_dev; https://github.com/FelixKrueger/TrimGalore) with parameter settings (–length 100 –quality 30), remaining reads were mapped to reference yeast genomes using STAR[50] (Version 2.7.8a). We used the genome of *S. cerevisiae* strain S288C as the reference (version R64-1-1; http://www.yeastgenome.org) with genome annotation file (version R64-1-1.104, genome-date 2011-09, genome-build-accession GCA_000146045.2, genebuild-last-updated 2018-10).

After mapping, SAMtools[52] (version 1.10) was used to filter all secondary alignments from raw bam file by setting parameter ″-F 0x100″, leaving only primarily alignments. Modified bam file was taken into expression levels determination by Featurecounts[53] (version 1.6.2) with specific parameters (-p -T 20 -M –largestOverlap -t exon -g gene_id). We found the pre-transcript and mature transcript of rRNA were meanwhile annotated in the original version of genome annotation file (gtf), and we removed two features RDN37-1 and RDN37-2 in the gtf file in the following analysis. In total 5789 genes including verified genes and RNA genes related to ribosomal biogeneis were analyzed.

In order to detect the differentially expressed genes of knockout strains compared with the wild-type strains, and between knockout strains in adjacent timepoints, we used the negative binomial generalized linear models provided by the software package DESeq2.[25] At the same time, we used the apeglm method for effect size shrinkage.[55] Finally, genes of knockout strains meeting the following conditions were considered as differentially expressed genes: the expression fold change is greater than 1.5 or less than 0.67 ($|\log_2 FC| > 0.585$), and the adjusted p value is less than 0.01 ($p_{adj} < 0.01$). To account for the heterogeneity observed in the evolved strains, we implemented a replicate-level analysis. Each replicate of the evolved strains was separately compared to the original strains. This approach allowed us to capture the variability

between replicates that might be masked in a pooled analysis. For each gene in each replicate, we derived the value of log_2FC relative to the expression in the original strain.

TPM (Transcripts Per Million) of each gene were further analysis by R studio (R version 4.0.3). We performed dimensionality reduction with the UMAP algorithm by R package umap[26] to project transcriptomes (log_2TPM) of all strains into two dimensions using default parameters (Table S8).

Calculation of fold enrichment of the intersection of DEGs

We randomly selected a set of genes from the entire gene pool. The number of selected genes was equivalent to the number of DEGs in the two groups under comparison. This process was repeated 100 times to generate a robust average overlap value. The calculated average was then compared with the actual observed overlap value to evaluate the fold enrichment of our findings.

Gene annotations and features analysis

Gene annotations were downloaded from Saccharomyces Genome Database (SGD, http://sgd-archive.yeastgenome.org/curation/literature/gene_association.sgd). Physically interacting partners for each gene were extracted from the BioGRID database (https://thebiogrid.org/BIOGRID-ORGANISM-Saccharomyces_cerevisiae_S288c-3.5.169). The paralogous gene pairs were extracted from Wapinski et al.'s study.[56] The Gene Ontology (GO) analysis of response genes were performed by Gene Ontology Term Finder on SGD (https://www.yeastgenome.org/goTermFinder).

The list of 77 snoRNAs was downloaded from Ryu H-Y's study.[57] 138 ribosome protein genes were obtained from Monticolo's study.[58] A total of 161 assembly factors (3 overlaps between two subunits) that function in maturation of 40S (73) and 60S (91) ribosomal subunits in *S. cerevisiae* were acquired from Woolford's study.[59] These genes were listed in Table S9.

QUANTIFICATION AND STATISTICAL ANALYSIS

Growth rates of strains at each time point were compared pairwise. Statistical analyses were conducted using the Wilcoxon test in R Studio. The original strains had three replicates, while the knockout strains had six replicates. The Wilcoxon test in R Studio was used to assess statistical significance for expression levels and the number of physical interacting partners between restored and unrestored genes. The Chi-squared test in R Studio was employed to evaluate the statistical significance of the proportion of genes having paralogs between corresponding groups.