Supplementary Information

Comparative Dynamic Transcriptome Analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation

Mai Sun*, Björn Schwalb*, Daniel Schulz*,

Nicole Pirkl, Stefanie Etzold, Laurent Larivière, Kerstin C Maier, Martin Seizl,

Achim Tresch, and Patrick Cramer.

Gene Center Munich and Department of Biochemistry, Center for Integrated Protein Science CIPSM,

Ludwig-Maximilians-Universität München, Feodor-Lynen-Strasse 25, 81377 Munich, Germany.

*These authors contributed equally.

Correspondence should be addressed to

A.T. (tresch@LMB.uni-muenchen.de) or P.C. (cramer@LMB.uni-muenchen.de)

Contents

1 Supplementary Figures S1 (4tU labeling does not affect normal cell physiology and rate estimation)

In order to check that 4tU labeling does not perturb the gene expression pattern of *Sc* cells, we compared RNA intensities of wild type cells against the total RNA intensities of cells after 3, 6 and 12 min of 4tU labeling (all cells were grown in YPD media). Almost no significant folds above a factor of 2 (below a factor of 0.5) were detected, and the distributions were almost identical to that of replicate wild type measurements (Figure 1, Table 1).

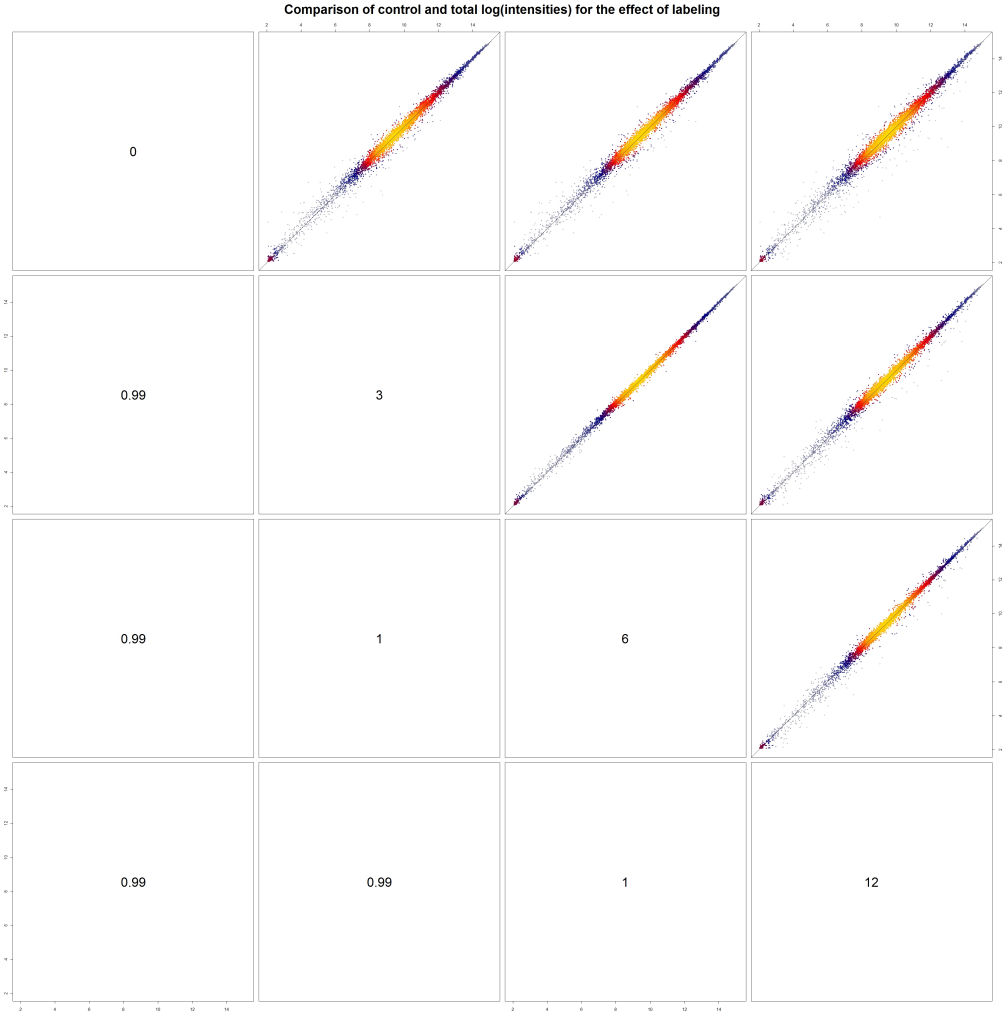


Figure 1: Pairwise scatterplots of log-intensities. The lower panel shows the respective Spearman correlations. The diagonal gives the length of the labeling period in minutes. Compared fractions are obtained by taking the gene-wise median over all intensities of replicate measurements.

repressed \ induced	0	3	6	12
0	-	8	11	21
3	1	-	0	7
6	0	0	-	4
12	0	7	0	-

Table 1: Counts of induced/repressed genes (upper triangle/lower triangle). Selection criteria were a multiple testing corrected significance level (local false discovery rate) of 0.05 and an expression fold of at least 2 between groups (length of the labeling period).

Note that differences in labeling efficiency p_{lab} do not affect decay and synthesis rate measures, as these efficiency biases only occur for transcripts with less than 500 uridine residues. Though these transcripts make up two thirds of all mRNAs, biases can be accurately removed. Therefore, different labeling biases estimated as in [8] can only adopt different curvatures, whereas altered synthesis and decay rates can only lead to shifted asymptotes of the respective curves:

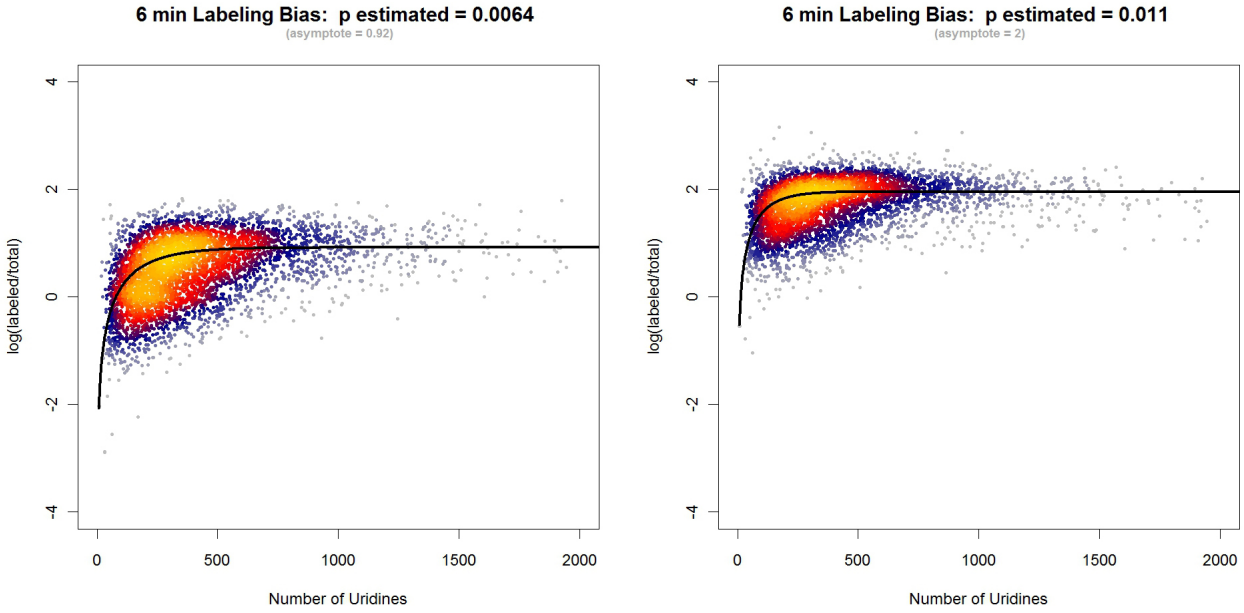


Figure 2: The number of uridines is plotted versus the log-ratio of L and T. The black line shows that the labeling bias curve estimated as in [8]. Left: labeling bias plot for the slow Pol II mutant. $p_{lab} = 0.0064$ means that approximately every 156th uridine residue is replaced by 4tU and afterwards attached to a biotin molecule. Right: labeling bias plot for the wild-type. $p_{lab} = 0.011$ means that approximately every 90th uridine residue is replaced by 4tU and afterwards attached to a biotin molecule. The shifted asymptotes indicate the observed fold of the decay rate comparing these two conditions (main text).

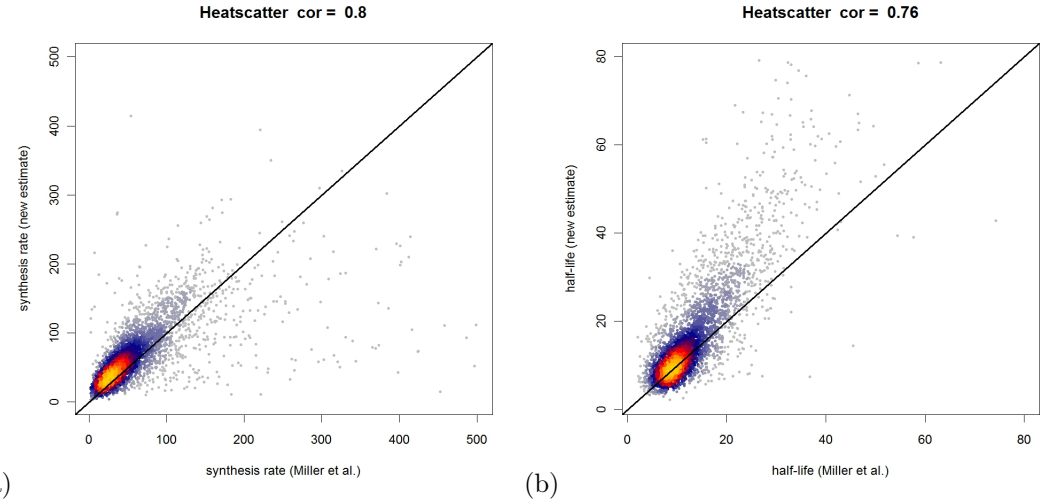
The calculated labeling efficiencies can vary across different conditions in *Sc* cells, whereas that of *Sp* cells should not, since they are not perturbed. Exactly that can be observed:

labeling bias p_{lab}	wild-type (rep.1)	wild-type (rep.2)	slow Pol II (rep.1)	slow Pol II (rep.2)
Sc	0.012	0.010	0.0064	0.006
Sp	0.0030	0.0029	0.0030	0.0028

Table 2: Estimated labeling biases p_{lab} for 2 replicate measurements of the wild-type and the slow Pol II mutant for *Sc* and *Sp*. Note the excellent agreement of the labeling efficiencies of the *Sp* aliquots (which should be identical, since they are taken from the same labeling experiment).

2 Supplementary Figures S2 (Comparison of cDTA measurements of RNA synthesis and decay with previous measurements)

In two *Sc* wild type samples, we measured the fraction of unlabeled mRNA in addition to the labeled and the total mRNA fraction. This made it possible to obtain an independent, second estimate of synthesis and decay rates. Note that the absolute half-life and synthesis rate levels show a remarkable agreement, despite the differences in cell cycle length and media. Both estimates were obtained using a new estimate of 60,000 transcripts per yeast cell [14], instead of the previously used old, four-fold lower estimate of 15000 [5], which was used in [8]:



(a) (b)

Figure 3: The scatterplots compare the *Sc* synthesis (a) and half-lives (b) as obtained by cDTA (y-axis) and those obtained from our data using the method from [8]. Spearman correlations are 0.8 and 0.76 respectively.

We also compared our *Sp* half-lives to that obtained in [1]. We were able to reproduce the half live estimates of this paper reasonably well, (correlation 0.95, linear fit: 2, R^2value: 0.94). The small discrepancy between the published half lives and our recalculated ones can be explained by the fact that we used only one replicate from the dye swap experiment. Moreover, since we are using open source software in the **R** developer environment instead of the commercial software Genepix, we did not apply Genepix-specific preprocessing steps to the raw data. When comparing the half lives from [1] to our cDTA estimates, they correlate moderately well (Spearman corelation 0.4 resp. 0.35 for the original resp. the recomputed half lives). This correlation increases to 0.58, if the labeling bias correction that has been applied to our measurements is also applied to the data of [1]. The pearson correlation coefficient raises to even from 0.33 to 0.48 if we further apply the the correction for exponential growth ($\alpha = \log(2)/116$ min, corresponding to a median *Sp* half live of 59 min, see Equation 75).

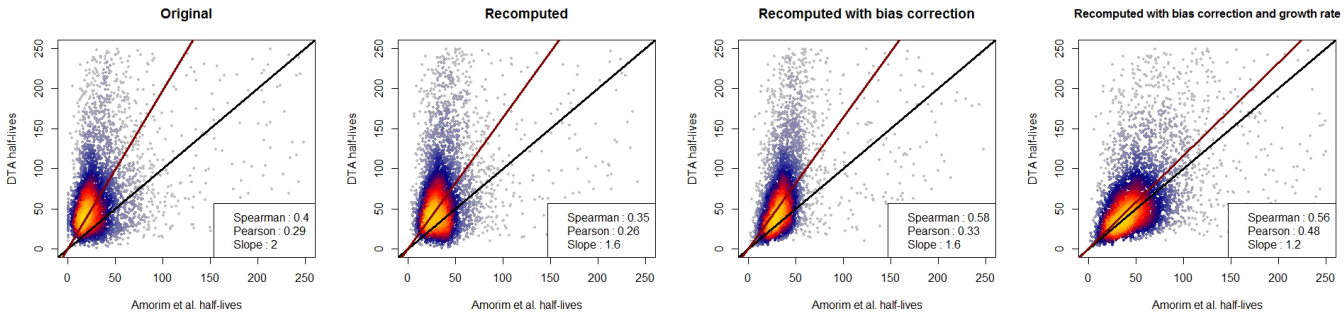


Figure 4: Comparison of the *Sp* transcript half-lives as obtained by us (y-axis; median half-life 59 min) and by [1] (x-axis). Spearman and pearson correlation coefficients are given in the legends. From left to right: comparison of our data to original estimiates [1], to recomputed estimates, to recomputed estimates with labeling bias correction and to recomputed estimates with labeling bias and growth rate correction.

While the half live estimates agree well up to a linear factor, they hugely disagree with respect to their absolute level (i.e., their median), by a factor of 2 (Figure 2). This is probably due to instabilities in the normalization procedure, which calibrates the contributions of the labeled, unlabeled, and total mRNA fractions. So far, three methods have been proposed: Linear regression (9.5.3), total least squares regression (9.5.3), bias-based regression (9.5.2). To illustrate this, we applied all three methods to the 9 wild type DTA samples in the [8] data. In theory, all 9 median half life estimates should be identical. As can be seen from (Figure 5) though, they vary strongly, in a range from ~5min to 105min, and no single method can be considered stable. This finding demonstrates the urgent need for an external normalization like proposed in the cDTA protocol.

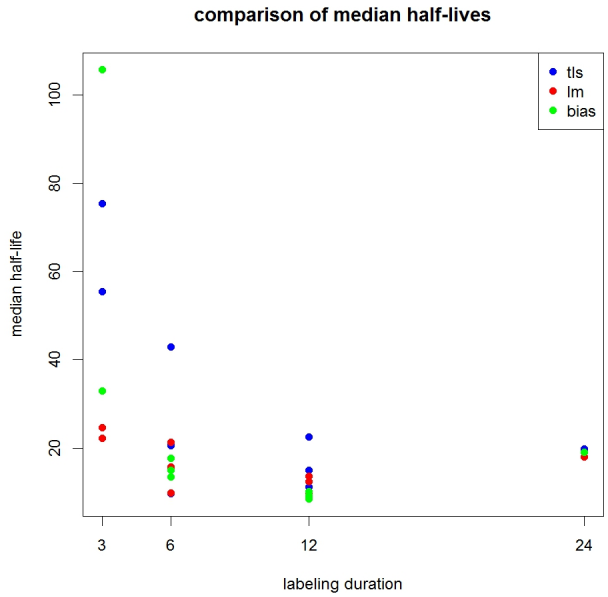


Figure 5: Median half-lives resulting from three different normalization methods (Linear regression (9.5.3), total least squares regression (9.5.3), bias-based regression (9.5.2)) of several wild-type replicate measurements of the labeled, unlabeled, and total mRNA fractions after labeling durations of 3, 6, 12 and 24 minutes.

We further compared our cDTA half-life estimates with *Sc* literature results (see Figure 7). The half-life estimates except for [9] were obtained by experiments using transcriptional arrest. The estimates of [6, 13, 4, 12] were obtained using a Yeast strain containing the RNA polymerase II temperature sensitive mutant rpb1-1. [3] uses the transcription inhibitor 1,10-phenanthroline. Decay rates can be measured after blocking transcription, but this requires a perturbing heat shock (cells have to be shifted to the nonpermissive temperature of 37°C), or is inherently

cell invasive (inhibitors). We suspect that this harsh treatment involves strong secondary effects that overlay the degradation process (e.g., the degradation enzymes themselves might alter their activity). The intensity of each mRNA species relative to that observed in a wild type cell gives their measure for mRNA decay, applying the usual first-order exponential decay model. All these data sets are compared to our estimates ('cDTA (this study)') which shows almost no correlation to any of the other data sets. We reproduced the experiments of [6, 13, 4, 12] using two different Yeast strains (two biological replicates each) also carrying the RNA polymerase II temperature sensitive mutant rpb1-1 ('rpb1-1 (this study)'), and wild-type cells exposed to heatschock ('heat shock (this study)'), measured 24 and 66 minutes after the temerature shift (Figure 6). These measurements are in good agreement to the data set obtained by [6, 13, 4, 12, 3]. This strongly indicates that the so obtained half-life estimates are heavily biased by the applied heatshock. To further investigate this, we also apply this method to the salt stress data [8] ('salt stress (Miller et. al. 2011)') which is also in a very good agreement with the estimates obtained by [6, 13, 4, 12, 3], indicating that the measured decay profiles strongly underlie a general stress response profile of *S.cerevisiae*.

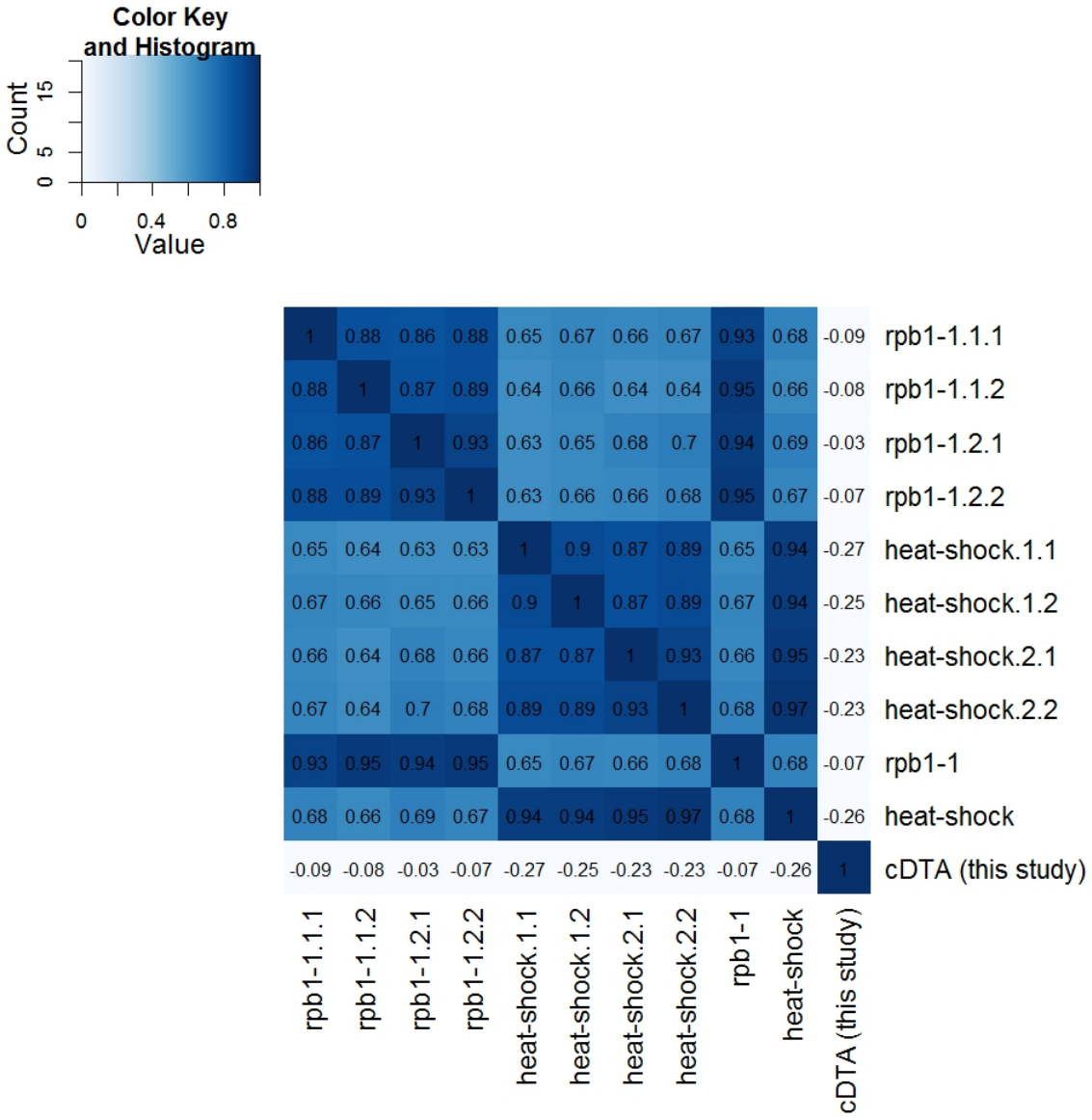


Figure 6: Heatmap showing the spearman correlation of half-life estimates obtained by linear regression of the rpb1-1 and heat-shock timecourse measurements equated to the timepoint of extraction. Suffix numbers indicate different Yeast strains or different biological replicates. 'rpb1-1' and 'heat-shock' were obtained from gene-wise medians of the measurements. 'DTA' gives our half-life estimates.

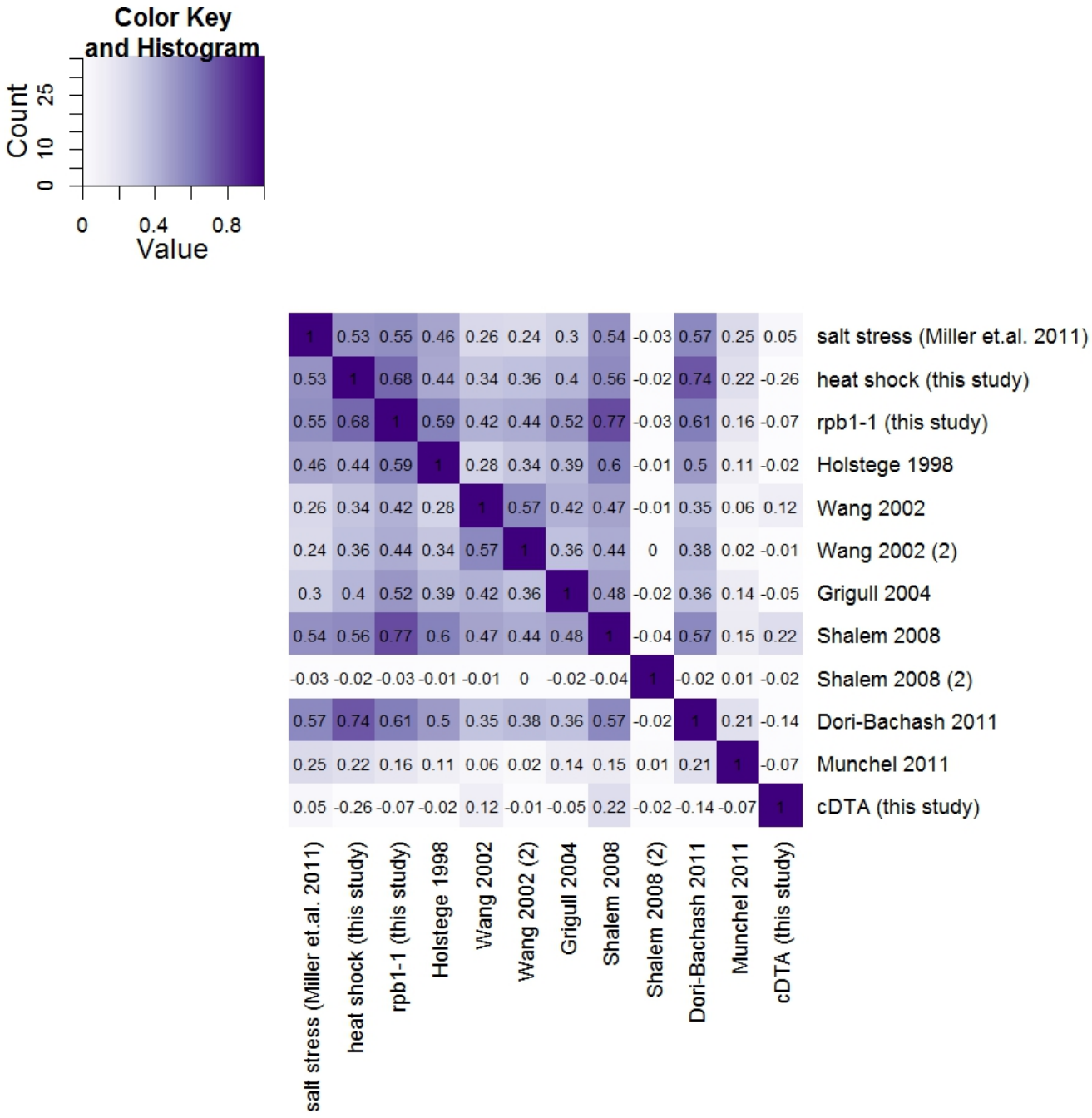


Figure 7: Correlation analysis of half-life measurements. Color-coded heatmap shows pairwise spearman correlation coefficients of half-life measurements. (white: negative or zero correlation to purple: pefect correlation).

3 Supplementary Figures S3 (Growth curves)

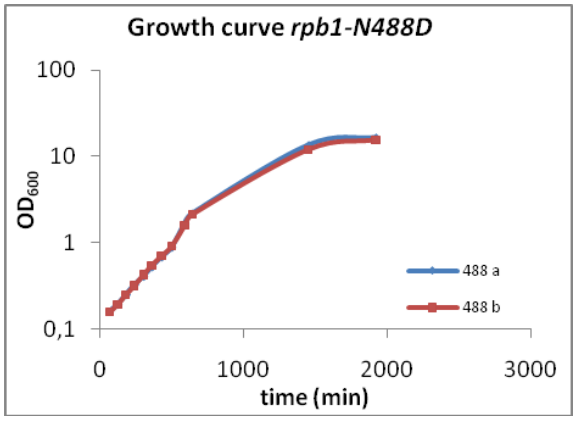


Figure 8: 2 replicate growth curves of the slow Pol II mutant. Growth rate: 150 min

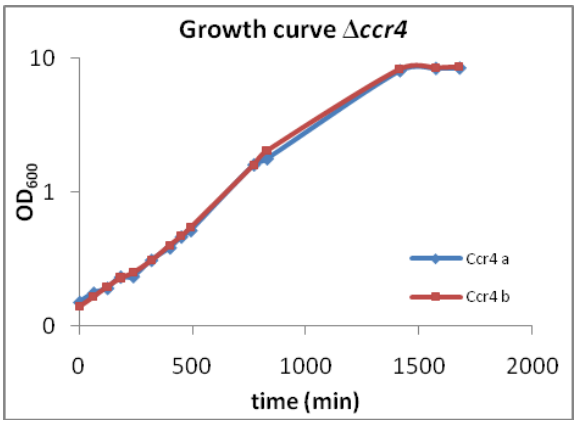


Figure 9: 2 replicate growth curves of the ccr4 deletion mutant. Growth rate: 219 min

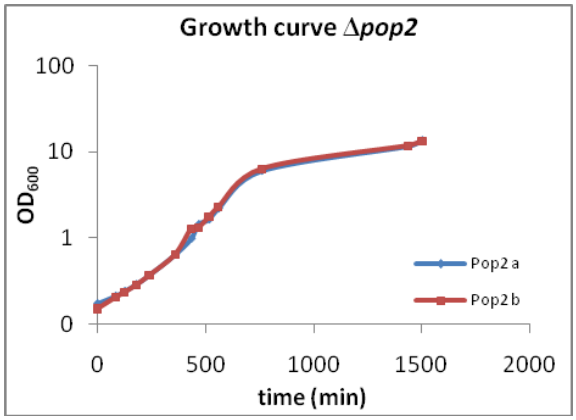


Figure 10: 2 replicate growth curves of the pop2 deletion mutant. Growth rate: 126 min

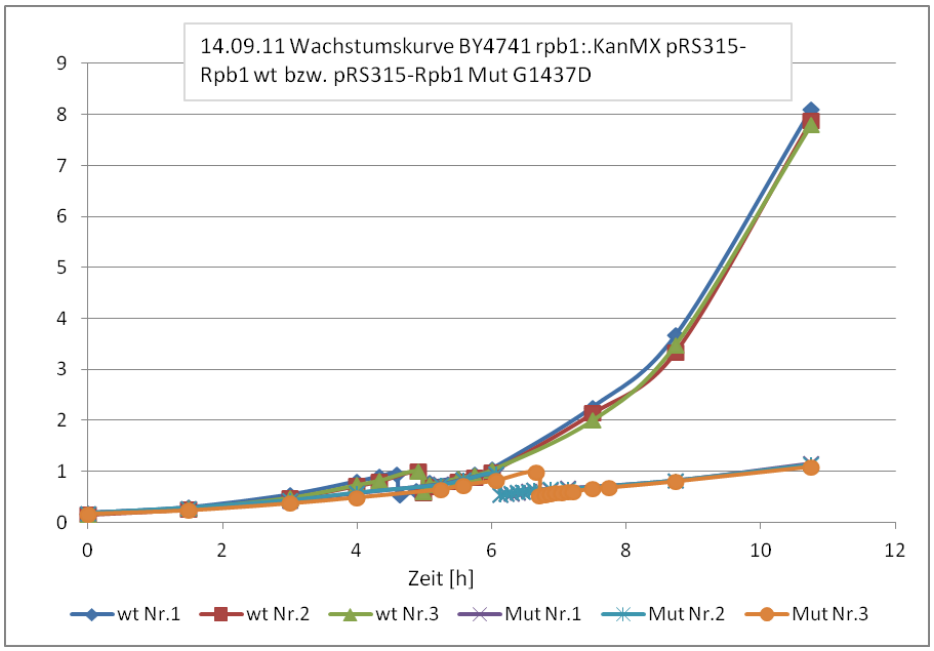


Figure 11: 2 replicate growth curves of the rpb1-1 mutant and its isogenic wild-type (BY4741 rpb1:.KanMX pRS315-Rpb1 wt bzw. pRS315-Rpb1 Mut G1437D). Growth rate: 153.8 min/107.6 min

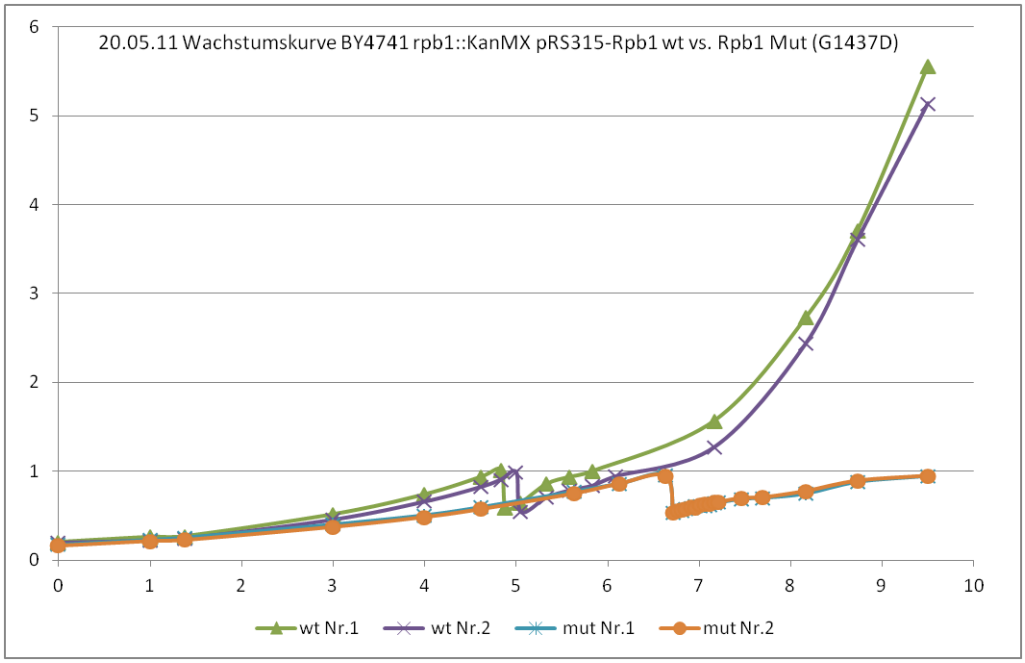


Figure 12: 2 replicate growth curves of the rpb1-1 mutant and its isogenic wild-type (BY4741 rpb1::KanMX pRS315-Rpb1 wt vs. Rpb1 Mut G1437D). Growth rate: 161.2 min/121.8 min

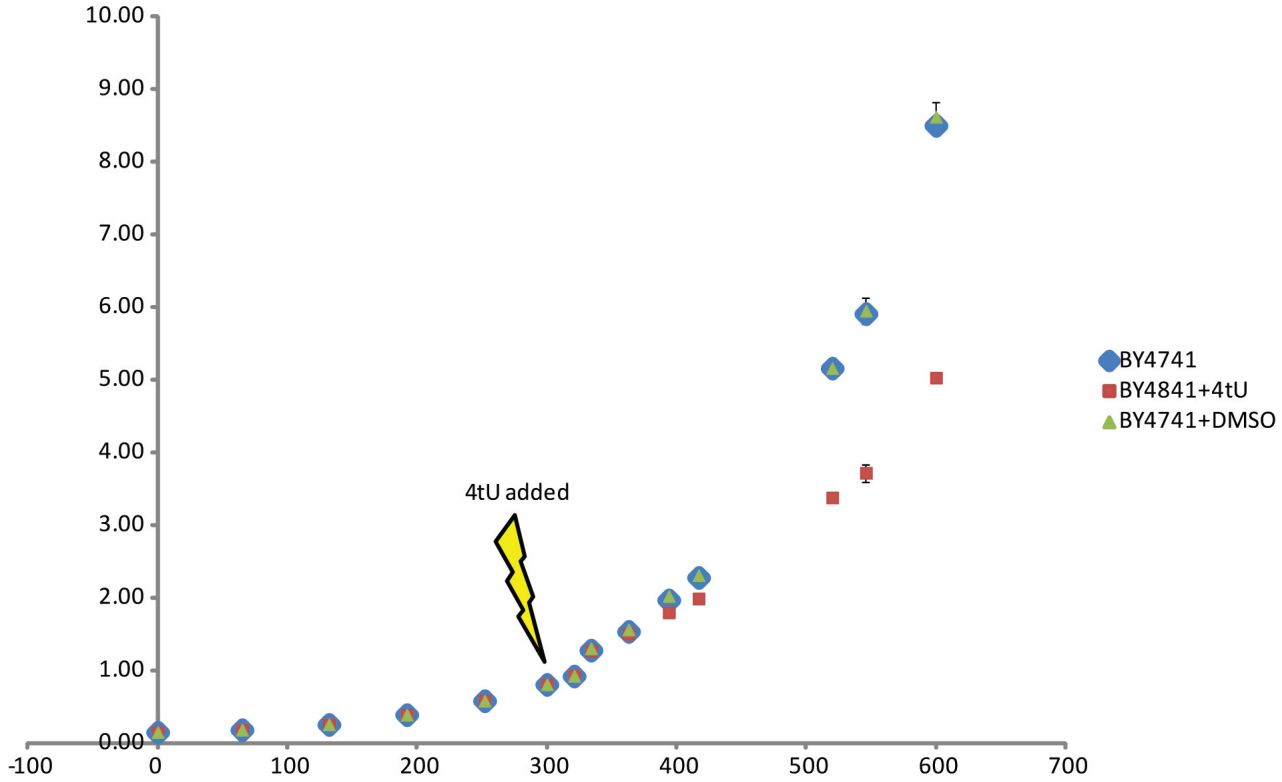
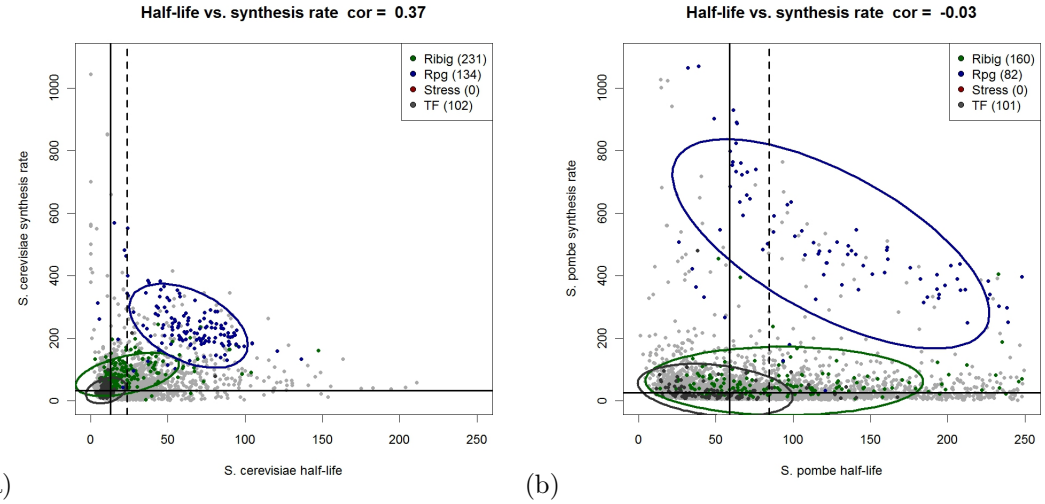


Figure 13: Growth curves of the wild-type (BY4741), wild-type with added DMSO (BY4741+DMSO) and wild-type with added 4-thiouracil (BY4741+4tU).

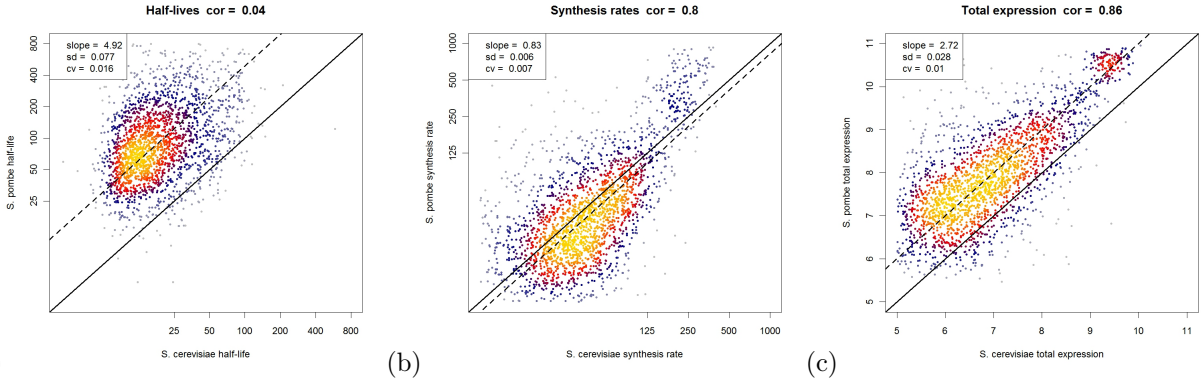
4 Supplementary Figures S4 (Differences in *Sp* vs. *Sc* total mRNA levels can essentially be explained by one global multiplicative factor)

After identifying orthologous transcripts in *Sc* and *Sp* [2], we compared *Sc* and *Sp* mRNA turnover of orthologous transcripts. Visual inspection suggests that most of the differences in total expression levels can be explained by a global multiplicative factor:



(a) (b)

Figure 14: Scatter plot of mRNA half-lives [min] and synthesis rates [counts/cell/90min] are shown for *S. cerevisiae* (left) and *S. pombe* (right). Colored points belong to the following gene sets: green, ribosomal biogenesis genes; violet, ribosomal protein genes; red, stress genes; dark gray, transcription factors (TFs). Assuming gene sets to be gaussian distributed, ellipses show the 75% regions of highest density for the respective sets. Black lines indicate median half-life/synthesis rate (*Sc*: 12/53, *Sp*: 59/44).



(a) (b) (c)

Figure 15: Scatter plot comparing mRNA half-life (a), synthesis rate (b) and total expression (c) of *S. pombe* vs *S. cerevisiae* orthologs (>25% protein sequence identity). The offset of dashed lines to (parallel) black lines indicate ratios of median half-life resp. synthesis rate resp. total mRNA of S.p. to S.c. (4.92/0.83/2.72).

5 Supplementary Figures S5 (Correlation of synthesis and decay rates in *Sc* and *Sp* with transcript length)

We compared the unspliced transcript length of orthologous genes in *Sc* and *Sp* to verify that they essentially agree:

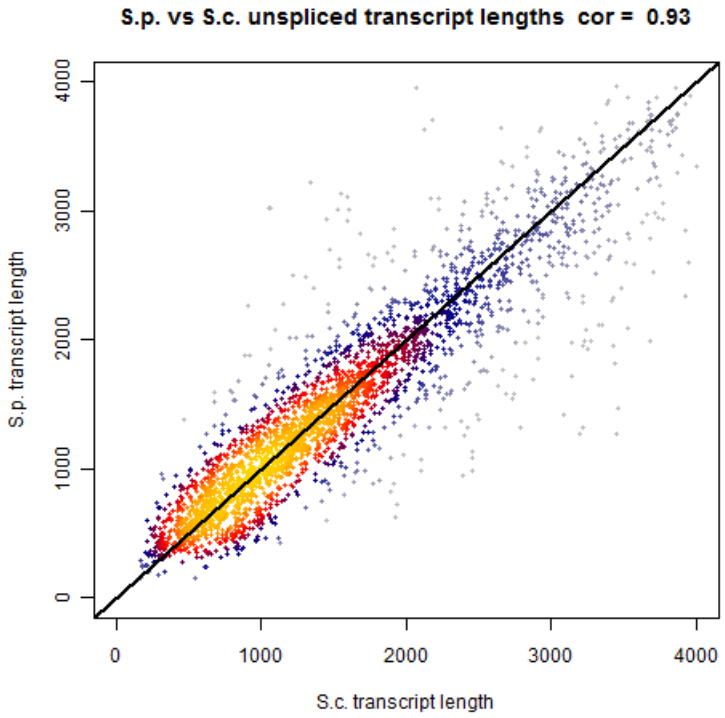
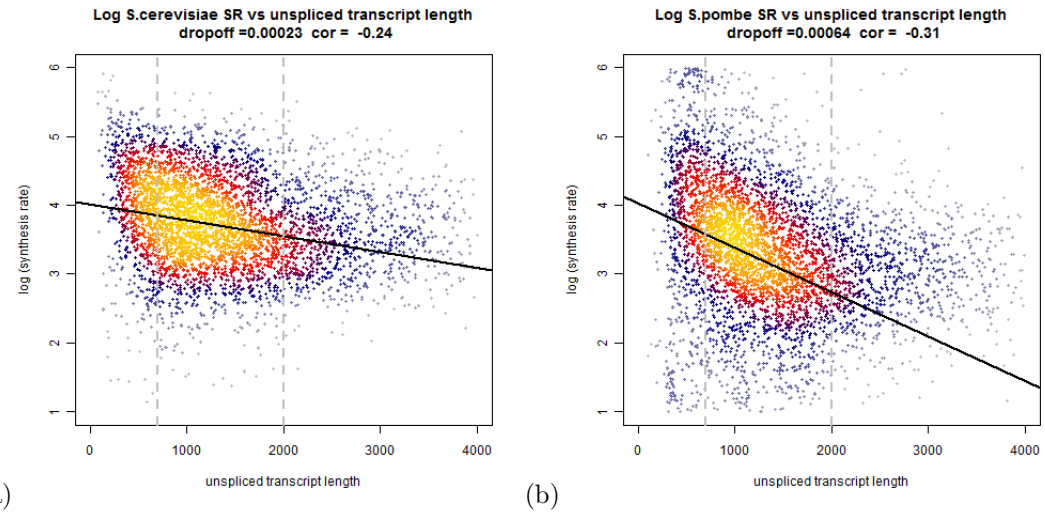


Figure 16: Comparison of the lengths of unspliced transcripts of *Sc*/*Sp* orthologs. There were 76 Transcripts of length greater than 5000 (in either *Sc* or *Sp*) that are not shown here for convenience. The diagonal line indicates identical transcript length of orthologs in *Sc* and *Sp*.

When plotting the mRNA synthesis rates versus their unspliced transcript length for both *Sc* and *Sp* (Figure 5), we observed a clear dependence. We therefore investigated if this dependence can be explained by Pol II drop off during elongation. Our model is based on two assumptions: 1) The transcription initiation rate is not dependent on the transcript length. 2) The Pol II drop off rate r is constant along the transcript. Based on these assumptions, one can predict the decrease of the average synthesis rate (SR) as a function of the transcript length l by

$$SR(l) \propto (1-r)^l \quad , \text{ or } \quad log\, SR(l) = l \cdot (1-r) + const \tag{1}$$

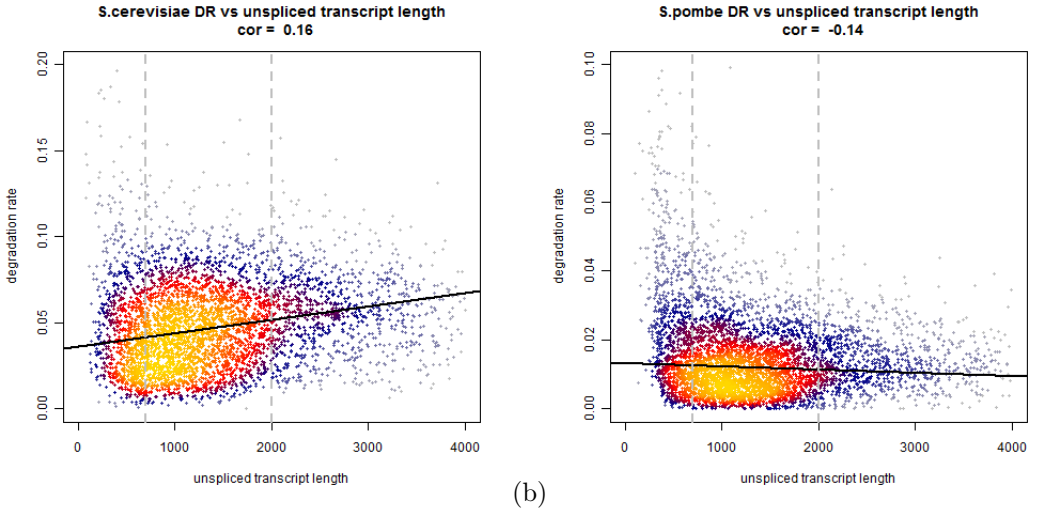
Equation (1) can be used to estimate the drop off rate r in a linear regression of $log\, SR(l)$ versus l. The regressions were performed using transcripts of a length between 700 and 2000 nucleotides. This was done to exclude potential estimation artifacts for long and short transcripts, which might arise from residual biases that have not been removed by our normalization. The results remain qualitatively unchanged if all genes are included (data not shown). It turned out that *Sc* (resp. *Sp*) synthesis rates slightly decrease with the length of the transcript, namely by a 21% (resp. 47%) [molecules/cell/cell cycle] per 1000 nucleotides transcript length (see also Table 7). The Null hypothesis ($r = 0$) could be rejected with a p-value of $< 10^{-6}$ in both cases.



(a) (b)

Figure 17: Correlation of (log) synthesis rates with length for *Sc* (a) and *Sp* (b). The linear regression and the Pearson correlation were calculated for the transcripts with a length between 700 and 2000 nucleotides.

We also assessed the transcript length dependence of the *Sc* resp. *Sp* decay rates. It turned out that the *Sc* decay rates increases with transcript length, by $7.7 \cdot 10^{-3}[1/min]$ per 1000 nucleotides, whereas the Sp decay rates decreases slightly by $0.9 \cdot 10^{-3}[1/min]$ per 1000 nucleotides. This amounts to a relative change of merely 0.017% per 1000 nucleotides (compared to the median decay rate in *Sc*). In *Sp*, this corresponds to a relative change of merely 0.0081% per 1000 nucleotides (compared to the median decay rate in *Sp*). We conclude that these changes are of no biological relevance (though they are statistically significant, p-value in a linear regression $< 10^{-6}$ for *Sc* resp. $< 10^{-4}$ for *Sp*).



(a) (b)

Figure 18: Correlation of decay rates with length. Pearson correlations were calculated for the transcripts with a length between 700 and 2000. a) Length dependence of *S.cerevisiae* decay rates b) Length dependence of *S.pombe* decay rates.

Without showing the results, we remark that these section's findings remain qualitatively unchanged if one restricts the analysis to orthologs with a certain minimum amount of sequence identity (e.g., of at least 25% - there are 2568 such transcripts).

6 Supplementary Figures S6 (Reproducibility)

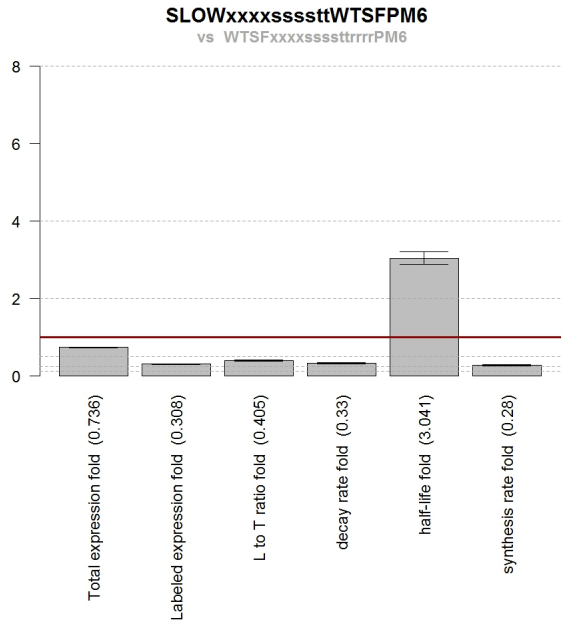


Figure 19: Barplot of measured global shifts (total expression, labeled expression, labeled to total ratio (decay proxy), decay rate, half-life and synthesis rate) of the slow Pol II mutant vs. its isogenic wild-type in log scale. Error bars indicate the standard deviation of pairwise comparison of respective replicates.

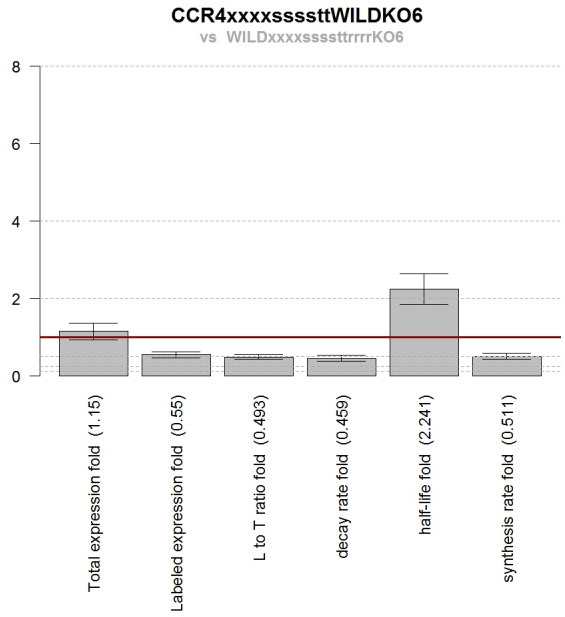


Figure 20: Barplot of measured global shifts (total expression, labeled expression, labeled to total ratio (decay proxy), decay rate, half-life and synthesis rate) of the Δccr4 mutant vs. wild-type in log scale. Error bars indicate the standard deviation of pairwise comparison of respective replicates.

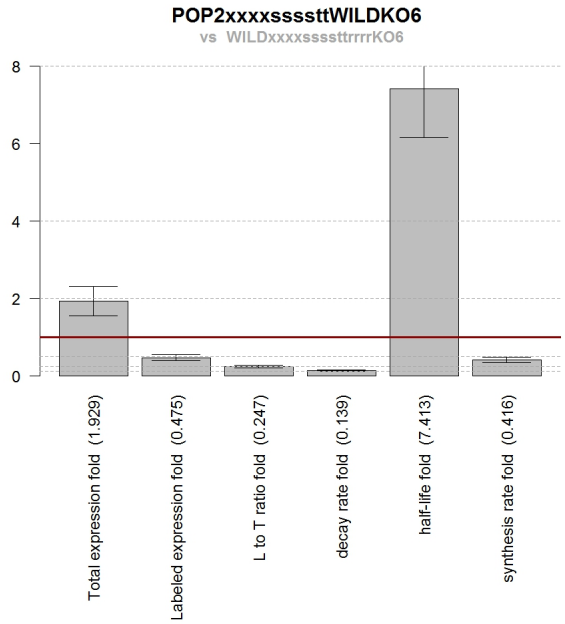


Figure 21: Barplot of measured global shifts (total expression, labeled expression, labeled to total ratio (decay proxy), decay rate, half-life and synthesis rate) of the Δpop2 mutant vs. wild-type in log scale. Error bars indicate the standard deviation of pairwise comparison of respective replicates.

7 Supplementary Figures S7 (Synthesis and decay rate regulation is independent of transcript length)

The discovery of a global regulation of mRNA transcription activity raises the question about the responsible mechanism. A straightforward hypothesis is that transcription inhibition is achieved by increasing the abort rate of Polymerase during transcription elongation. However, the abort rate of Pol II (as estimated in Supplementary Figures S4) in all three mutants (slow Pol II mutant, deadenylation mutants Δccr4 and Δpop2) is comparable to that of *Sc* wild type, see Figure 7, and Table 7. This is particularly remarkable for the slow Pol II mutant. It is thus likely that the feedback mechanism does not intervene at the elongation stage, but rather at the stage of transcription initiation or during the transition from transcription initiation to elongation.

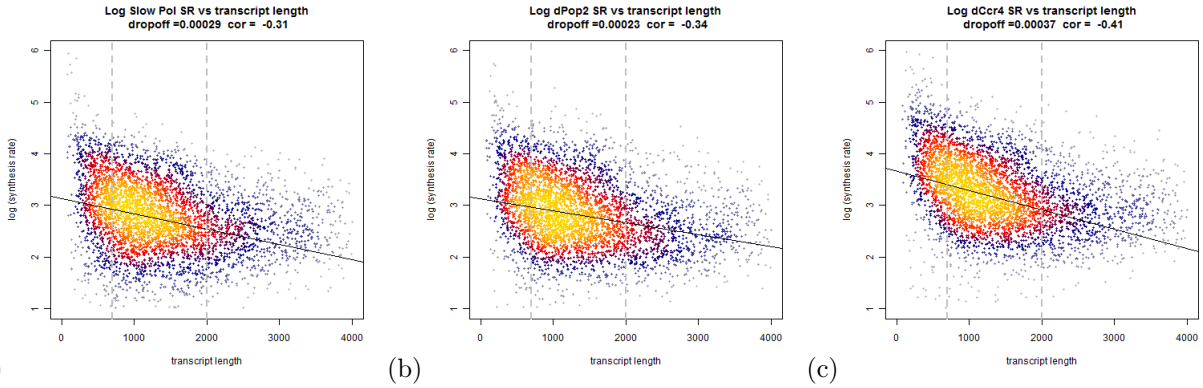


Figure 22: Dependence of synthesis rates on transcript length for the slow Pol II mutant (a), the deadenylation mutants Δccr4 (b) and Δpop2 (c).

	S.cerevisiae	S.pombe	Slow Pol II	Ccr4 mutant	Pop2 mutant
drop off rate per nucleotide	$2.3 \cdot 10^{-4}$	$6.4 \cdot 10^{-4}$	$2.9 \cdot 10^{-4}$	$3.7 \cdot 10^{-4}$	$2.3 \cdot 10^{-4}$
Pol II drop off per 1000 nucleotides	21%	47%	26%	31%	21%

Table 3: Pol II drop off rates estimated from Equation (1).

An alteration of the Pol II abort frequency affects the synthesis rates of longer transcripts stronger than those of shorter transcripts, which is indeed observed (Figure 7). According to our model in Supplementals S4, we are able to calculate the relative proportion of Pol II successfully terminating a hypothetical transcript of length 8000nt in the slow Pol mutant versus the *Sc* wild type. Doing so, we arrive at an estimate of 57%, i.e., for a given transcription initiation rate at this transcript, its synthesis rate in the slow Pol II mutant would be 57% of the synthesis rate in the *Sc* wild type. This is in excellent agreement with an estimate obtained by a direct experimental method in [7].

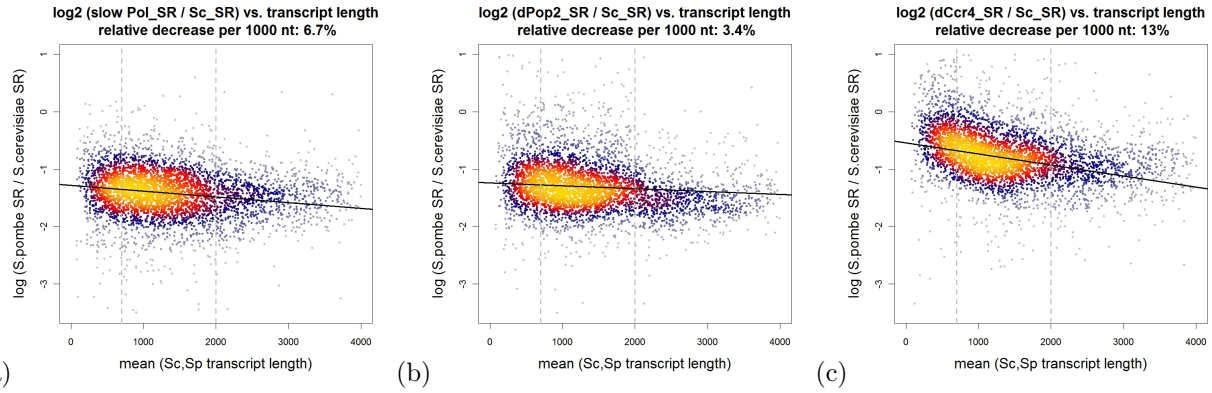


Figure 23: The (log2) quotient of mRNA synthesis rates of the slow Pol II mutant (a), the deadenylation mutants ΔPop2 (b) and ΔCcr4 (c) relative to the wild type synthesis rates, plotted as a function of the transcript length. The decrease (per 1000nt) in relative synthesis rates caused by different drop off rates is given in the header of the plots. Estimations were based on transcripts of length between 700nt and 2000nt, as indicated by dashed vertical lines.

8 Supplementary Method S8 (A model of regulated mRNA synthesis and degradation)

We model the dynamics of the mRNA concentration a gene G by a first order rate equation. It is determined by the synthesis rate $SR(G)$ at which G is produced, and the decay rate $DR(G)$ at which G is degraded:

$$\frac{dg}{dt} = SR(G) - DR(G) \tag{2}$$

$DR(G, D) = \lambda_G \cdot g \cdot h(d)$ For clarity of presentation, we explicitly state the assumptions that underlie our model.

Assumption 1: There are two proteins, S (the *transcription modulator*), and D (the *degradation modulator*), that globally modulate mRNA synthesis and degradation activity, respectively. The synthesis rate $SR(G, S)$ is a function of the (characteristics of) gene G and the transcription modulator S. More specifically, we assume that the synthesis rate $SR(G, S) = \mu_G \cdot f(s)$ decomposes into a gene-specific, constant term μ_G, the so-called *transcriptional efficiency* of G, and a gene-independent term $f(s)$ that depends on the expression level s of the gene corresponding to S. The degradation rate $DR(G, D)$ is a function of the (characteristics of) gene G and the degradation modulator D. More specifically, we assume that $DR(G, D) = \lambda_G \cdot h(d) \cdot g$ decomposes into a gene-specific, constant term λ_G, the so-called *degradation efficiency* of G, and a gene-independent term $h(d)$ that depends on the expression level d of the gene corresponding to D.

The differential equation describing the dynamics of the mRNA population of a gene G becomes

$$\frac{dg}{dt} = SR(G, S) - DR(G, D) = \mu_G f(s) - \lambda_G h(d) \cdot g \tag{3}$$

This is a very general model for regulated mRNA turnover. By choosing f and h the constant function with value 1, we obtain the classical model of RNA dynamics without regulation, henceforth called the *naive model*, i.e., $SR(G, S) = SR(G) = \mu_G$ and $DR(G, D) = DR(G) = \lambda_G \cdot g$. In the naive model, the synthesis rate equals the transcriptional efficiency, but these two quantities are distinct in general. The same holds for the constant λ_G which determines the degradation rate in the regulation-free model, but which is modulated by a factor $h(d)$ in the general case. Under steady state conditions ($\frac{dg}{dt} = 0$), we can solve (3) for the expression level of G,

$$g = \frac{\mu_G}{\lambda_G} \cdot \frac{f(s)}{h(d)} \tag{4}$$

Assumption 2: All meausrements were made under steady state conditions. This is clearly the case for the experiments in our paper.

The quotient g_1/g_2 of the mRNA concentration of two genes G_1 and G_2 is independent of s and d, since

$$\frac{g_1}{g_2} = \frac{\mu_{G_1}}{\mu_{G_2}} \cdot \frac{\lambda_{G_2}}{\lambda_{G_1}} \tag{5}$$

Observation 1: Equation (5) implies a convenient property of a gene regulatory system: Relative expression levels are kept constant, independent of the modulator levels s and d. Changes in modulator activity only affect the total mRNA level.

When comparing mutant strains with defects in the general transcription/degradation machinery, we assume that these defects affect the transcription/degradation efficiencies of all genes in the same way:

Assumption 3: In our slow Pol II mutant, the transcriptional efficiencies are decreased relative to wild type by a global factor, whereas degradation efficiencies are unchanged. In the *ccr4* and the *pop2* deletion mutants, the degradation efficiencies are decreased relative to wild type by a global factor, whereas transcription efficiencies are unchanged. This assumption can be justified by the fact that gene-specific transcriptional efficiencies resp. degradation efficiencies quantify a biochemical interaction between an enzyme (Polymerase II resp. Ccr4/Not deadenylation complex) and the substrate (the DNA gene sequence resp. the mRNA transcript of the gene). Since the substrates do not change in our experiment, a change in synthesis-/degradation efficiency can only result from a change in the enzyme (e.g., a point mutation in Pol II resp. a removal of a protein of the Ccr4/Not complex). That this change in the enzyme results in a proportional change in the efficiencies is highly likely, because, up to random variation, we do only observe global, proportional changes in synthesis-/decay rates.

Formally, we will consider steady state conditions C, C' (with mRNA concentrations denoted by g resp. g') in which the corresponding transcriptional efficiencies μ_G, μ'_G respectively degradation efficiencies λ_G, λ'_G are proportional, i.e., $\mu'_G = \beta\mu_G$ and $\lambda'_G = \gamma\lambda_G$ for positive constants β, γ, and for all genes G. In such a situation

$$\frac{g_1}{g_2} = \frac{\mu_{G_1}}{\mu_{G_2}} \cdot \frac{\lambda_{G_2}}{\lambda_{G_1}} = \frac{\beta\mu_{G_1}}{\beta\mu_{G_2}} \cdot \frac{\gamma\lambda_{G_2}}{\gamma\lambda_{G_1}} = \frac{\mu'_{G_1}}{\mu'_{G_2}} \cdot \frac{\lambda'_{G_2}}{\lambda'_{G_1}} = \frac{g'_1}{g'_2} \tag{6}$$

Observation 2: The relative mRNA concentrations of two genes are invariant under global proportional changes in transcriptional efficiency and degradation efficiency.

Let T, T' be the total mRNA concentrations in conditions C resp. C'. For any gene Y with concentration y we have

$$\frac{T'}{T} \quad = \quad \frac{\sum_{\text{all genes}} g'}{\sum_{\text{all genes}} g} = \frac{y'}{y} \cdot \frac{\sum_{\text{genes } G}(g'/y')}{\sum_{\text{genes } G}(g/y)} \overset{\text{(Obs.1)}}{=} \frac{y'}{y} \tag{7}$$

This leads us to

Observation 3: The quotient y'/y of any gene Y equals the relative total mRNA concentrations in C' and C. Our model therefore predicts global (proportional) changes in mRNA concentrations in response to global (proportional) changes in synthesis and degradation efficiency, in good agreement with our experimental data.

For convenience, we define the constant $c = s/d$. Be aware that by Observation 1, this quotient does not change across (steady state) conditions, since $s/d = s'/d'$ for different conditions C and C'. Letting $G = D$ in equation (4),

$$d = \frac{\mu_D}{\lambda_D} \cdot \frac{f(cd)}{h(d)} \text{ , and } d' = \frac{\mu'_D}{\lambda'_D} \cdot \frac{f(cd')}{h(d')} \tag{8}$$

Let $\alpha = \frac{\mu'_D}{\mu_D} \cdot \frac{\lambda_D}{\lambda'_D}$. In the naive model, the quotient α equals the relative total mRNA concentrations in C' and C (in this case, $\alpha \overset{(8)}{=} d'/d \overset{(7)}{=} T'/T$). The case $\alpha > 1$, e.g., occurs when comparing the *ccr4* or *pop2* knockout strains (condition C') to wild type (condition C). In condition C', the function of deadenylation complex is impaired, and hence the degradation efficiencies are globally decreased, $\lambda'_G = \alpha^{-1}\lambda_G$. At the same time, the transcription efficiencies are unchanged, $\mu'_G = \mu_G$, for all genes G. The case $\alpha < 1$ is observed for the comparison of the slow Pol II mutant (C') and wild type (C). The transcriptional efficiencies in the slow Pol II mutant are globally decreased, $\mu'_G = \alpha\mu_G$, while the degradation efficiencies are unaffected, $\lambda'_G = \lambda_G$, for all genes G.

Observation 4: The relative change in total mRNA concentration in C' and C is given by

$$\frac{T'}{T} \overset{(7)}{=} \frac{d'}{d} \overset{(8)}{=} \alpha \cdot \frac{f(cd')}{h(d')} \cdot \frac{h(d)}{f(cd)} \tag{9}$$

We do not observe a hypercompensation in our experiments, i.e., the direction of the observed total mRNA changes ($T' > T$ resp. $T' < T$) agrees with the changes predicted by the naive model ($\alpha > 1$ resp. $\alpha < 1$). A compensatory effect (buffering) occurs when the observed relative change in total mRNA concentration, T'/T, is closer to 1 than the change predicted from the naive model, α.

Let $d' > d$. Then $T' > T$, and by lack of hypercompensation $\alpha > 1$. By compensation, $T'/T < \alpha$, or equivalently, $\frac{f(cd')}{h(d')} \cdot \frac{h(d)}{f(cd)} \overset{(9)}{=} \frac{T'}{T}\alpha^{-1} < 1$. We conclude:

Observation 5: If buffering, but no hypercompensation, occurs,

$$\frac{h(d')}{f(cd')} \quad > \quad \frac{h(d)}{f(cd)} \quad \text{for } d' > d \tag{10}$$

In other words, a necessary condition for buffering to occur is that $h(d)/f(cd)$ be a monotonically increasing function in d. By symmetry reasons, the same monotonicity condition emerges when assuming $d' < d$ (exchange of conditions C and C' converts this case into the case $d' > d$).

Assumption 4: The functions f and h are monotonic. In other words, if the transcription modulator S is an activator (resp. an inibitor), $f = f(s)$ is a monotonically increasing (resp. monotonically decreasing) function. Similarly, if D stimulates degradation, $h = h(d)$ is monotonically increasing, and if D is a degradation inhibitor, h is monotonically decreasing. This seems a rather uncritical since natural restriction to f and h.

We will discuss the plausibility of different biological scenarios in the light of condition (10). Some properties of the functions f and h are determined by the mode of operation of the transcription resp. degradation modulators S and D. First note that f and h are positive functions by definition, i.e. they assume only positive values.

Case 1. S is a transcription activator, D is an degradation inhibitor (f increasing, h decreasing). Then, $h(d)/f(dc)$ is monotonically decreasing in d, which contradicts condition (10). We exclude this case.

Let us make some general considerations about the size of compensation effects. Sensibly, the larger the effects on total mRNA levels are (larger T'/T), the more adverse the effects to the cell. The quotient $\frac{T'}{T}\alpha^{-1} = \frac{f(cd')}{h(d')} \cdot \frac{h(d)}{f(cd)}$ measures the discrepancy between the observed ratio of total mRNA concentrations and the relative concentrations predicted by the naive model. The smaller $\frac{T'}{T}\alpha^{-1}$, the stronger the buffering. Thus the efforts to compensate for these changes / the buffering strength should increase along with T'/T. We will show that only two of the remaining cases meet these needs.

Case 2. S is a transcription activator, D is an degradation activator (f increasing, h increasing). Typically, activitor functions f and h show an asymptotic behaviour if their arguments tend to infinity. Thus it is sensible to assume that in the current case, $h(d)/f(cd)$ converges to a finite number for increasing d. Together with the fact that the degradation molecule D needs to be present abundantly (d, d' large), this in turn implies that the size of the compensation effect, $\frac{T'}{T}\alpha^{-1} = \frac{f(cd')}{h(d')} \cdot \frac{h(d)}{f(cd)}$, tends to one. In the light of the above considerations, such a mechanism is not well suited to achieve a fast and robust buffering.

Case 3. S is a transcription inhibitor, D is an degradation inhibitor (f decreasing, h decreasing). Although we cannot exclude this case, we argue that the relation of f and h needs to be properly fine-tuned in order to obtain an increasing quotient $h(d)/f(cd)$. For example, if $f(s) = s^{-a}$ and $h(d) = d^{-b}$ for some $a, b > 0$, $h(d)/f(cd) = c^a d^{a-b}$, and a needs to be substantially smaller than b in order to achieve a strong compensation effect $\frac{T'}{T}\alpha^{-1} = (\frac{T'}{T})^{a-b}$.

Case 4. S is a transcription inhibitor, D is an degradation activator. Then, f is monotonically decreasing, h is monotonically increasing, and the quotient $h(d)/f(cd)$ is monotonically increasing for any choice of h and f compatible with Case 4. More conveniently, the compensation effect, $\frac{T'}{T}\alpha^{-1}$ automatically increases for increasing T'/T, i.e. buffering is stronger the stronger the effect on total mRNA levels is. Since no exact tuning of the functions h and f is needed to achieve a strong buffering effect without hypercompensation, this is arguably the most natural and hence the most likely case.

Let us put the 4 alternative hypotheses to the test and collect experimental evidence:

Denote the synthesis- and degradation rates in condition C (wild type) and condition C' (e.g. a mutant) by $SR_C(G)$ and $SR_{C'}(G)$ respectively $DR_C(G)$ and $DR_{C'}(G)$. Then,

$$\frac{SR_{C'}(G,S)}{SR_C(G,S)} = \frac{\mu'_G f(s')}{\mu_G f(s)} \quad , \text{ and } \quad \frac{DR_{C'}(G,D)}{DR_C(G,D)} = \frac{\lambda'_G h(d') \cdot g'}{\lambda_G h(d) \cdot g} \tag{11}$$

For the *pop2* and *ccr4* mutants (condition C') relative to wild type (condition C), we have $\lambda'_G = \alpha\lambda_G$ for some $\alpha < 1$, and $\mu'_G = \mu_G$, for all genes G. From our measurements, we find that the total levels increase in the mutants (by a factor of 1.78 in *pop2* and a factor of 1.34 in *ccr4*). Thus, $T' > T$ and $s' > s$. It follows from (11) that $f(s')/f(s) = SR_{C'}(G)/SR_C(G)$ for all genes G. Thus, $f(s')/f(s)$ can be estimated reliably by

$$\frac{f(s')}{f(s)} = \text{median}_G\left(\frac{SR_{C'}(G,S)}{SR_C(G,S)}\right) = \begin{cases} 0.4 & \text{for pop2} \\ 0.6 & \text{for ccr4} \end{cases} < 1 \tag{12}$$

Observation 5. From $s' > s$ and $f(s') < f(s)$ it follows that f is monotonically decreasing, i.e. S is a transcriptional inhibitor.

For the slow Pol mutant (condition C') relative to wild type (condition C), we have $\lambda'_G = \lambda_G$, and $\mu'_G = \alpha\mu_G$, for some $\alpha < 1$ and for all genes G. The changes in total mRNA are given by $T'/T = g'/g$ for all genes G, and thus relative total mRNA levels can be estimated by

$$\frac{T'}{T} = \text{median}_G\left(\frac{g'}{g}\right) = 0.75 \tag{13}$$

Thus, $T' < T$ and $d' < d$. Again from (11), we have

$$\frac{h(d')}{h(d)} = \frac{SR_{C'}(G)}{SR_C(G)'} \cdot \frac{T}{T'} \quad \text{for all genes } G \tag{14}$$

and a reliable estimate of $h(d')/h(d)$ is given by

$$\frac{h(d')}{h(d)} = \text{median}_G \left(\frac{DR_{C'}(G,D)}{DR_C(G,D)} \right) \cdot \frac{T}{T'} = 0.28 \cdot 0.75^{-1} = 0.37 < 1 \tag{15}$$

Observation 6. From $d' < d$ and $h(d') < h(d)$ it follows that h is monotonically increasing, i.e. D is a degradation stimulator.

Conclusion. Our data provide support for a scenario in which a transcription inhibitor and a degradation stimulator establish a regulatory circuit for the global stabilization of both the absolute and the relative cellular mRNA levels.

We conclude with the discussion of some objections to our model.

1. One or both modulators might not be proteins, hence there is no term $f(s)$ resp. $h(d)$ that depends on mRNA concentrations of the modulators S and D. Or, although S and D are proteins, their activity might be subject to posttranscriptional / posttranslational modifications and therefore cannot be adequately modeled by a function of their corresponding mRNA levels. This is a justified objection, and we definitely and explicitely do not rule out alternative mechanisms of synthesis-/decay modulation that cannot be described in terms of mRNA abundances. In addition, it is well possible that there exist several independent mechanisms that establish a robust feedback.

2. The synthesis rates $SR(G,S)$ and the decay rates $DR(G,S)$ might not factor into a gene-specific and a modulator-specific term. This however would imply that the shift of synthesis rates would in general not be global but gene-specific. E.g., in the deadenylation mutants $\Delta ccr4$ or $\Delta pop2$ (in which $\mu'_G = \mu_G$ holds), we would expect from (11) that
$$\frac{SR_{C'}(G,S)}{SR_C(G,S)} \neq \frac{\mu'_G f(s')}{\mu_G f(s)} = \frac{f(s')}{f(s)} \tag{16}$$
and the left hand side term in equation (16) would become gene-dependent, in contrast to our model (right hand side). However, we *do* observe a strong global proportional shift of synthesis rates in three experimental conditions (slow Pol II, $\Delta ccr4$ and $\Delta pop2$ deletion mutants), and the deviations from this proportional shift are not larger than expected from random variations in the measurements. Future research will have to provide further experimental evidence for this proportional shift by measuring other strains with an altered transcriptional / degradation efficiency.

3. The changes in transcriptional resp. degradation efficiency might not be proportional (for all genes G). This would lead to the same conclusion as in the previous point, that the shift in synthesis rate resp. degradation rate is not proportional, and the same argument applies.

4. The functions f and h might not be monotonic. It is sensible and consistent with the assumption of monotonicity to believe that f and h are asymptotically constant, i.e. for large values, they converge to a constant value, respectively. However, to us there is absolutely no reason why the strength of modulation should suddenly increase above some modulator concentration when it was constantly decreasing below that concentration, or vice versa. We discard this case as unlikely, though we cannot definitely exclude it.

5. The model does not take into account spatial inhomogeneities in the mRNA- or modulator concentration within the cell, in particular one should at least formulate a compartment model including the nucleus and the cytosol. One could envisage to model the formation of mRNA processing bodies (P-bodies) with increased degradation modulator concentrations relative to other regions. Though desirable, these models require spatially resolved, genome-wide synthesis and degradation measurements which cannot be provided by cDTA, nor by any other current experimental method.

9 Supplementary Methods S9 (Modeling and estimation of synthesis and decay rates)

9.1 A model for mRNA synthesis and degradation

mRNA levels in a cell are the consequence of two opposing mechanisms, namely mRNA synthesis and mRNA degradation. DTA allows monitoring these contributions in a non-perturbing manner. The experimental setup for DTA requires culturing cells in the presence of a labeling substrate (e.g. 4 thiouridine (4sU) or 4 thiouracil (4tU)) for a certain amount of time. Until the extraction of the mRNA sample, the analogous (labeling) substrate will be incorporated into newly transcribed mRNA. This setup yields three types of mRNA fractions: total cellular mRNA, newly transcribed labeled mRNA and pre-existing unlabeled mRNA. All three fractions can subsequently be quantied through gene expression profiling on microarrays or next generation sequencing (RNAseq).

Let $r \in R$ be a sample. At time $t = 0$, we start the mRNA labeling. At the timepoint t_r, when the mRNA is extracted, the total mRNA amount $C_{gr}(t_r)$ of gene g in the sample r is composed of the amount $B_{gr}(t_r)$ of (pre-existing) mRNA that has been synthesized *before* $t = 0$ and the amount $A_{gr}(t_r)$ of mRNA that has been newly synthesized *after* $t = 0$,

$$\underbrace{C_{gr}(t_r)}_{\text{total RNA}} = \underbrace{A_{gr}(t_r)}_{\text{newly synthesized mRNA}} + \underbrace{B_{gr}(t_r)}_{\text{pre-existing mRNA}} \tag{17}$$

Let $N_r(t_r)$ denote the number of cells in the sample r at time t_r. The cells are grown and harvested during mid-log phase, i.e. the cell number follows an exponential law with growth rate,

$$\alpha = \frac{log(2)}{CCL} \geq 0 \tag{18}$$

with cell cycle length CCL, this means

$$\frac{dN_r(t)}{dt} = \alpha N_r(t) \tag{19}$$

and therefore

$$N_r(t_r) = N_r(0)e^{\alpha t_r}. \tag{20}$$

α is often refered to as dilution rate, i.e. the reduction of concentration due to the increase of cell volume during growth. Say we have a cellular expression level $m_g = m_g(t)$ (transcripts of gene g per cell), then

$$C_{gr}(t) = m_g(t)N_r(t) \tag{21}$$

and hence

$$\begin{aligned} \frac{dC_{gr}(t)}{dt} &= m_g(t)\frac{dN_r(t)}{dt} + \frac{dm_g(t)}{dt}N_r(t) \\ &= m_g(t)\alpha N_r(t) + \frac{dm_g(t)}{dt}N_r(t) \\ &= \left(m_g(t)\alpha + \frac{dm_g(t)}{dt}\right)N_r(t). \end{aligned} \tag{22}$$

We assume that the mRNA population of a gene g decays exponentially at a (relative) rate given by $\lambda_g = \lambda_g(t)$ if no other processes interfere. This means that the amount of $\lambda_g(t)C_{gr}(t)$ is degraded at time t. We further assume that the mRNA population of a gene g is synthesized at an absolute rate per cell given by $\mu_g = \mu_g(t)$. Hence the amount of $\mu_g(t)N_r(t)$ is synthesized at time t. This means for the total mRNA amount that

$$\begin{aligned} \frac{dC_{gr}(t)}{dt} &= \mu_g(t)N_r(t) - \lambda_g(t)C_{gr}(t) \\ &= \mu_g(t)N_r(t) - \lambda_g(t)m_g(t)N_r(t) \\ &= (\mu_g(t) - \lambda_g(t)m_g(t))\,N_r(t). \end{aligned} \tag{23}$$

Using Equations (22) and (23), we get

$$(\mu_g(t) - \lambda_g(t) m_g(t))\, N_r(t) = \left(m_g(t)\alpha + \frac{dm_g(t)}{dt} \right) N_r(t) \tag{24}$$

and hence

$$\mu_g(t) = m_g(t)\,(\alpha + \lambda_g(t)) + \frac{dm_g(t)}{dt}. \tag{25}$$

Additionally, we assume that the pre-existing mRNA population of a gene g decays exponentially at a (relative) rate given by $\lambda_g = \lambda_g(t)$. This means for the pre-existing mRNA fraction that

$$\frac{dB_{gr}(t)}{dt} = -\lambda_g(t) B_{gr}(t). \tag{26}$$

Consequently, the newly synthesized mRNA fraction can be given by

$$\frac{dA_{gr}(t)}{dt} = \frac{dC_{gr}(t)}{dt} - \frac{dB_{gr}(t)}{dt} = \mu_g(t) N_r(t) - \lambda_g(t) A_{gr}(t). \tag{27}$$

9.2 Constant synthesis and decay rates

To simplify our model, we assume genes to have a (time averaged) constant synthesis rates μ_g and decay rates λ_g during 4sU/4tU labeling, i.e. $\mu_g(t)$ and $\lambda_g(t)$ are given piecewise constant. For the newly synthesized mRNA, this means

$$\frac{dA_{gr}(t_r)}{dt} = \mu_g N_r(t_r) - \lambda_g A_{gr}(t_r) = \mu_g N_r(0) e^{\alpha t_r} - \lambda_g A_{gr}(t_r). \tag{28}$$

The solution of this differential equation yields

$$A_{gr}(t_r) = c e^{-\lambda_g t_r} + \frac{\mu_g N_r(0) e^{\alpha t_r}}{\alpha + \lambda_g} \tag{29}$$

with an initial value $A_{gr}(0) = 0$, and so

$$0 = c + \frac{\mu_g N_r(0)}{\alpha + \lambda_g}. \tag{30}$$

This finally leads to

$$A_{gr}(t_r) = \frac{\mu_g N_r(0)}{\alpha + \lambda_g} \left[e^{\alpha t_r} - e^{-\lambda_g t_r} \right] \tag{31}$$

and therefore

$$\mu_g = \frac{A_{gr}(t_r)(\alpha + \lambda_g)}{N_r(0) \left[e^{\alpha t_r} - e^{-\lambda_g t_r} \right]}. \tag{32}$$

For the total mRNA, this means

$$\frac{dC_{gr}(t_r)}{dt} = \mu_g N_r(t_r) - \lambda_g C_{gr}(t_r) = \mu_g N_r(0) e^{\alpha t_r} - \lambda_g C_{gr}(t_r). \tag{33}$$

The solution of this differential equation yields

$$C_{gr}(t_r) = c e^{-\lambda_g t_r} + \frac{\mu_g N_r(0) e^{\alpha t_r}}{\alpha + \lambda_g} \tag{34}$$

with

$$c = C_{gr}(0) - \frac{\mu_g N_r(0)}{\alpha + \lambda_g} \tag{35}$$

This finally leads to

$$C_{gr}(t_r) = C_{gr}(0) e^{-\lambda_g t_r} + \frac{\mu_g N_r(0)}{\alpha + \lambda_g} \left[e^{\alpha t_r} - e^{-\lambda_g t_r} \right] \tag{36}$$

and finally

$$B_{gr}(t_r) = C_{gr}(t_r) - A_{gr}(t_r) = B_{gr}(0)e^{-\lambda_g t_r} \tag{37}$$

with $B_{gr}(0) = C_{gr}(0)$. From (32) and (36) we can deduce

$$C_{gr}(t_r) = C_{gr}(0)e^{-\lambda_g t_r} + A_{gr}(t_r) \tag{38}$$

and rearrange it as follows

$$C_{gr}(t_r) - A_{gr}(t_r) = C_{gr}(0)e^{-\lambda_g t_r}. \tag{39}$$

So

$$e^{-\lambda_g t_r} = \frac{C_{gr}(t_r) - A_{gr}(t_r)}{C_{gr}(0)} \tag{40}$$

and finally

$$\lambda_g = -\frac{1}{t_r} log \left[\frac{C_{gr}(t_r) - A_{gr}(t_r)}{C_{gr}(0)} \right] \tag{41}$$

or

$$\lambda_g = -\frac{1}{t_r} log \left[\frac{B_{gr}(t_r)}{C_{gr}(0)} \right]. \tag{42}$$

As λ is assumed to remain constant in the above derivation, its value will be an average value along the labeling period.

9.3 Special case: steady state

Now we assume that the cells exhibit constant growth under constant environmental conditions. In particular, this implies that the amount of each mRNA population is constant over time, being the result of a dynamic equilibrium of a constant mRNA synthesis and decay, i.e. steady state. In this case we assume genes to have a (time averaged) constant cellular expression level m_g during 4sU/4tU labeling. Thus, equation (21) simplifies to

$$C_{gr}(t_r) = m_g N(t_r) \tag{43}$$

and leads to

$$C_{gr}(t_r) = m_g(0) N_r(0) e^{\alpha t_r} \tag{44}$$

as $m_g(t) = const. = m_g(0) = m_g$ and hence

$$C_{gr}(t_r) = C_{gr}(0)e^{\alpha t_r}. \tag{45}$$

So, equation (41) can be written as

$$\lambda_g = -\frac{1}{t} log \left[\frac{C_{gr}(t_r) - A_{gr}(t_r)}{C_{gr}(t_r)e^{-\alpha t}} \right] \tag{46}$$

and therefore

$$\lambda_g = -\alpha - \frac{1}{t} log \left[1 - \frac{A_{gr}(t_r)}{C_{gr}(t_r)} \right] \tag{47}$$

or

$$\lambda_g = -\alpha - \frac{1}{t} log \left[\frac{B_{gr}(t_r)}{C_{gr}(t_r)} \right]. \tag{48}$$

Steady state mRNA levels can now be derived, for $\alpha = 0$ from

$$\frac{dC_{gr}(t_r)}{dt} = \mu_g N_r(t_r) - \lambda_g C_{gr}(t_r) = m_g(t_r) \alpha N_r(t_r) = 0$$

and thus

$$\mu_g N_r(0) - \lambda_g C_{gr}(0) = 0.$$

Consequently

$$C_{gr}(0) = \frac{\mu_g N_r(0)}{\lambda_g} \tag{49}$$

For $\alpha \neq 0$

$$\frac{dC_{gr}(t_r)}{dt} = \mu_g N_r(t_r) - \lambda_g C_{gr}(t_r) = m_g(t_r)\alpha N_r(t_r)$$

which yields

$$m_g(0) = \frac{\mu_g}{\alpha + \lambda_g}$$

and so

$$C_{gr}(0) = \frac{\mu_g N_r(0)}{\alpha + \lambda_g} \tag{50}$$

gives the total mRNA level achieved by a dynamic equilibrium of a constant mRNA synthesis and decay. In this special case μ_g can be stated as

$$\mu_g = m_g(0)(\alpha + \lambda_g) \tag{51}$$

by implication.

9.4 Adaption to measured values

We now have to relate the measured levels of $L_{gr}(t_r)$, $U_{gr}(t_r)$ and $T_{gr}(t_r)$ to the levels of the mRNA fractions $A_{gr}(t_r)$, $B_{gr}(t_r)$ and $C_{gr}(t_r)$. Ideally, these fractions would respectively equal each other. There are however disagreements that are due to mRNA extraction efficiencies, amplification steps in the biochemical protocol and scanner calibration of the flourescence readouts. The amount $L_{gr}(t_r)$ of labeled mRNA for instance is proportional to the amount of labeled mRNA $A_{gr}(t_r)$ at the time t_r of sampling,

$$L_{gr}(t_r) = a_r A_{gr}(t_r) \ , \tag{52}$$

with an unknown array-specific constant a_r. Analogously, the measured amounts $T_{gr}(t_r)$ and $U_{gr}(t_r)$ depend on the actual amounts $C_{gr}(t_r)$ and $B_{gr}(t_r)$ respectively via

$$T_{gr}(t_r) = c_r C_{gr}(t_r) \ , \tag{53}$$

and

$$U_{gr}(t_r) = b_r B_{gr}(t_r) = b_r \left(C_{gr}(t_r) - A_{gr}(t_r)\right) \tag{54}$$

with unknown array-specific constants c_r and b_r.

There can also be disagreements that originate from the 4sU/4tU labeling efficiency (see Figure 24). Let p_r be the probability that during the labeling process of sample r, a Uridine is replaced by 4sU/4tU and afterwards attached to a Biotin molecule. Let l_{gr} represent the fraction of newly synthesized mRNAs of gene g in sample r that are biotinylated. We assume that all biotinylated mRNAs are captured by the Streptavidin beads. Denote by $\#u_g$ the number of Uridine residues present in the mRNA corresponding to gene g. We can calculate l_{gr} as

$$l_{gr} = l(p_r, \#u_g) = 1 - (1 - p_r)^{\#u_g} \tag{55}$$

l_{gr} is thus the probability that at least one Uridine is replaced by 4sU and afterwards attached to a Biotin molecule. Accordingly, we have to correct equations (52),(53) and (54) to that effect. As a consequence we have the dependencies:

$$L_{gr}(t_r) = l_{gr} a_r A_{gr}(t_r) \ , \tag{56}$$

and

$$T_{gr}(t_r) = c_r C_{gr}(t_r) \ , \tag{57}$$

and

$$U_{gr}(t_r) = b_r \left(C_{gr}(t_r) - l_{gr} A_{gr}(t_r)\right) \ . \tag{58}$$

(59) can also be stated as

$$U_{gr}(t_r) = b_r u_{gr} B_{gr}(t_r) \tag{59}$$

where u_{gr} gives the increase of pre-existing mRNAs by newly synthesized mRNAs of gene g in sample r that are not biotinylated. With equation (58) and (59) we have

$$C_{gr}(t_r) = l_{gr} A_{gr}(t_r) + u_{gr} B_{gr}(t_r) \ , \tag{60}$$

and therefore

$$u_{gr} B_{gr}(t_r) - B_{gr}(t_r) = A_{gr}(t_r) - l_{gr} A_{gr}(t_r) \tag{61}$$

with respect to equation (17). If we solve for u_{gr} we get

$$u_{gr} = 1 + \frac{A_{gr}(t_r)}{B_{gr}(t_r)} (1 - p_r)^{\#u_g} \tag{62}$$

and thus

$$u_{gr}\left(p_r, \frac{b_r}{a_r}, u_g\right) = 1 + \frac{b_r}{a_r}(1 - p_r)^{\#u_g} \quad \text{if} \quad \frac{u_{gr} L_{gr}(t_r)}{l_{gr} U_{gr}(t_r)} = 1 \ . \tag{63}$$

The proportional constants then relate as follows

$$\frac{c_r}{a_r} = \frac{\frac{b_r}{a_r}}{1 + \frac{b_r}{a_r}} \quad \text{if also} \quad \frac{L_{gr}(t_r)}{l_{gr} T_{gr}(t_r)} = 1 \ . \tag{64}$$

9.5 Parameter Estimation

Our model contains the parameters $\Theta = \{\alpha,\ p_r,\ a_r,\ b_r,\ c_r,\ \lambda_{gr},\ \mu_{gr} \mid r \in R,\ g \in G\}$. It contains an implicit normalization procedure, because the two sources of experimental bias are part of the model (the parameters a_r, b_r, c_r account for multiplicative bias introduced via sample preparation and array scanning, and p_r models the labeling bias (Figure 24)). We propose a 5-step procedure for the identification of the parameters Θ. There are often genes that can not be accurately measured, i.e. genes that are not expressed often lead to noisy measurement signals. These could lead to flawed parameters, and therefore all genes that are considered valid for the parameter estimation are aggregated in $G^{reliable}$. Since the doubling times of the cells are usually known or can be measured accurately, α is given by $\alpha = \log 2 / CCL$.

9.5.1 Estimation of the labeling probability p_r

The estimation of the sample-related parameters $\{p_r,\ a_r,\ b_r,\ c_r \mid r \in R\}$ is done on the basis of the reliable genes $G^{reliable}$. The quotient of observed total and labeled mRNA levels can be written as

$$\frac{L_{gr}}{T_{gr}} = \frac{l_{gr} a_r A_{gr}(t_r)}{c_r C_{gr}(t_r)} = l_{gr} \frac{a_r}{c_r} \left[1 - e^{-t_r(\alpha + \lambda_g)}\right] \tag{65}$$

The first equation follows by (56) and (57), the second by (36) and (31). We can visualize this dependence conveniently by plotting u_g versus $\log \frac{L_{gr}}{T_{gr}}$ (see Figure 24). If all decay rates were equal, all points would lie on the graph given by the relationship of u_g versus $\log l_{gr} + \log \frac{a_r}{c_r}$. The scatter around this graph is caused by measurement errors and differences in decay rates. We can also calculate the quotient

$$\frac{U_{gr}}{T_{gr}} = \frac{b_r u_{gr} B_{gr}(t_r)}{c_r C_{gr}(t_r)} = u_{gr} \frac{b_r}{c_r} \left[e^{-t_r(\alpha + \lambda_g)}\right] \tag{66}$$

This equation follows by (57) and (58). We will predominantly use equation (65) for the estimation of p_r. Taking logs in Equation (65) and rearranging terms, we obtain

$$\log \frac{L_{gr}}{T_{gr}} = \log \frac{a_r}{c_r} + \log l(p_r, u_g) + \log \left[1 - e^{-t_r(\alpha + \lambda_g)}\right] \tag{67}$$

$p_r < 0.95$ implies that for $\#u_g > 700$ say, the approximation $\log(l_{gr}) \approx 0$ is almost exact, as $(1-p_r)^{\#u_g}$ is a monotonic sequence with

$$\lim_{\#u_g \to \infty} (1-p_r)^{\#u_g} = 0$$

and $(1-p_r)^{700} < 2.5 \cdot 10^{-16}$. Hence Equation (65) simplifies to

$$\log \frac{L_{gr}}{T_{gr}} = \log \frac{a_r}{c_r} + \log \left[1 - e^{-t_r(\alpha+\lambda_g)}\right] \quad \text{for} \quad \#u_g > 700 \tag{68}$$

If we additionally assume that the distribution of decay rates do not depend on the number of the uridines, the right-hand side in (68) becomes a constant plus some error term with expectation 0. Thus, we estimate $asymptote_r^{L/T}$ by letting

$$asymptote_r^{L/T} = \text{median} \left\{ \log \frac{L_{gr}}{T_{gr}} \mid g \in G^{reliable},\ \#u_g > 700 \right\} \tag{69}$$

Given equation (69), it is relatively easy to compute a good estimate of p_r by finding an optimal fit to (67) (see Figure(24)), for all $g \in G$ with $\#u_g < 500$. So we optimize the value of p_r, $r \in R$, by minimizing the l_1-loss function

$$p_r^{est} = \underset{q \in (0,1)}{\text{argmin}}\, \text{loss}(q) \quad \text{with} \quad \text{loss}(q) = \sum_{g \in G,\, \#u_g < 500} \left| \log \frac{L_{gr}}{T_{gr}} - \log l_{gr}(q, u_g) - asymptote_r^{L/T} \right| \tag{70}$$

Here, 500 is an upper bound that ensures that the measurements are still responsive to changes in u_g.

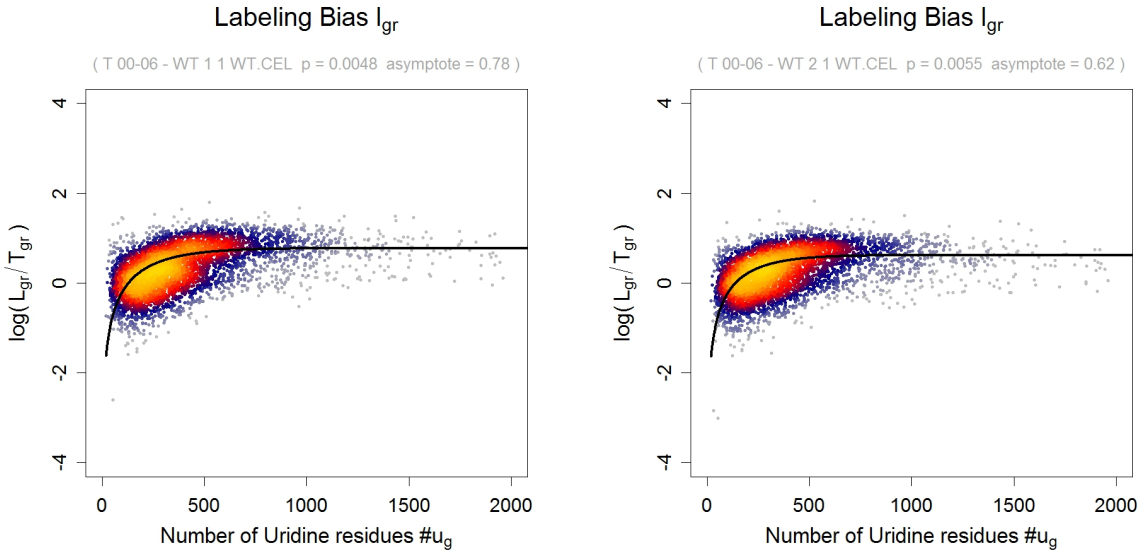


Figure 24: The number of Uridines $\#u_g$ is plotted versus the log-ratio of L_{gr} and T_{gr} for two replicates of *S.cerevisiae* at a labeling time of 6 minutes. The points of the scatterplot are colored according to the (estimated) point density in that region [10]. The labeling bias parameter $p_r^{est} = 0.0048$ and $p_r^{est} = 0.0055$ imply that approximately every 208th resp. 182th Uridine residue is replaced by 4sU. mRNAs which contain less than 500 Uridine residues (approx. 72% of all *S.cerevisiae* mRNAs) are not captured efficiently.

9.5.2 Estimation of the ratio of labeled to unlabeled mRNA $\frac{b_r}{a_r}$

Notice that for the purpose of synthesis and decay rate estimation it is sufficient to determine the quotients $\frac{b_r}{a_r}$, $\frac{a_r}{c_r}$ or $\frac{b_r}{c_r}$ instead of the individual constants a_r, b_r and c_r. As in described in section (9.5.1), we can use equation (66) to estimate

$$asymptote_r^{U/T} = \text{median} \left\{ \log \frac{U_{gr}}{T_{gr}} \mid g \in G^{reliable},\ \#u_g > 700 \right\} \tag{71}$$

and consequently optimize the value of $\frac{b_r}{a_r}$, $r \in R$, by minimizing the l_1-loss function

$$\frac{b_r}{a_r}^{est} = \underset{q\in(0,5)}{\operatorname{argmin}} \operatorname{loss}(q) \quad \text{with} \quad \operatorname{loss}(q) = \sum_{g\in G,\, \#u_g<500} \left| \log\frac{U_{gr}}{T_{gr}} - \log u_{gr}(p_r, q, \#u_g) - asymptote_r^{U/T} \right| . \tag{72}$$

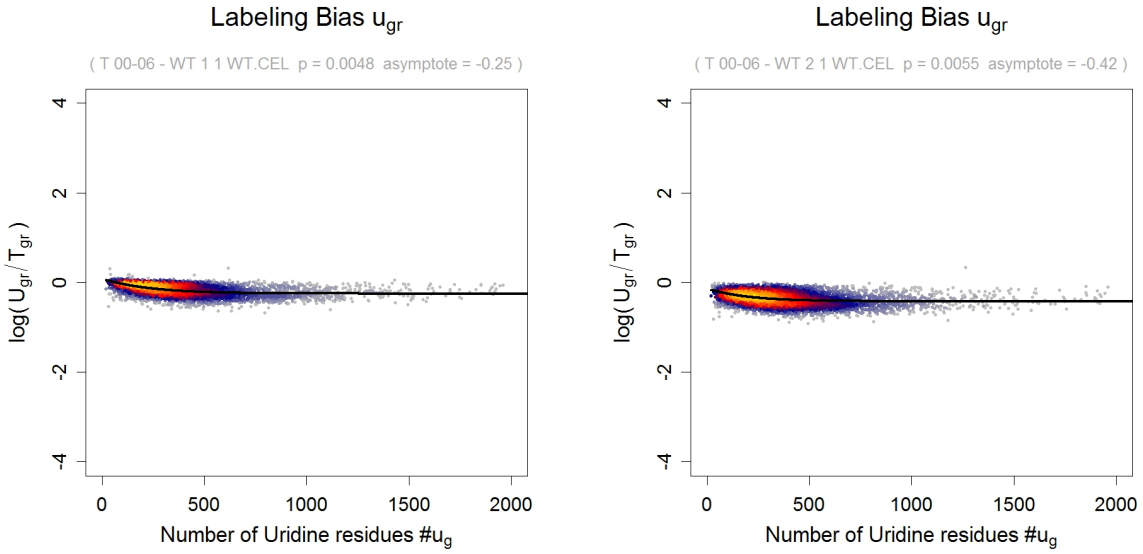


Figure 25: The number of Uridines $\#u_g$ is plotted versus the log-ratio of U_{gr} and T_{gr} for two replicates of *S.cerevisiae* at a labeling time of 6 minutes. The labeling bias parameter $p_r^{est} = 0.0048$ and $p_r^{est} = 0.0055$ estimated via the labeled mRNA fraction L_{gr} as described in section (9.5.1) reveal the ratio of newly synthesized to pre-existing mRNA given by $\frac{b_r}{a_r}^{est}$.

9.5.3 Estimation of the ratio of labeled to total mRNA $\frac{a_r}{c_r}$ and unlabeled to total mRNA $\frac{b_r}{c_r}$

In order to determine the quotients $\frac{a_r}{c_r}$ or $\frac{b_r}{c_r}$, we simply multiply equations (65) and (66) by the inverse of those quotients and add them up to obtain

$$\frac{c_r}{a_r}\frac{L_{gr}}{T_{gr}} + \frac{c_r}{b_r}\frac{U_{gr}}{T_{gr}} = 1 \quad \text{or} \quad T_{gr}(t_r) = \frac{c_r}{a_r}L_{gr} + \frac{c_r}{b_r}U_{gr} \tag{73}$$

Equation (73) describes a plane $\{(T_{gr}, L_{gr}, U_{gr}) \mid T_{gr} = \frac{c_r}{a_r}L_{gr} + \frac{c_r}{b_r}U_{gr}\}$ in a 3-dimensional Euclidean space. As the observed variables have measurement errors on both sides of Equation (73), we perform a total least squares regression of T_{gr} versus L_{gr} and U_{gr}, which accounts for a Gaussian error in the dependent variable T_{gr} and, in contrast to ordinary linear regression, also in the independent variables L_{gr}, U_{gr}. The total least squares regression minimizes the orthogonal distance of the datapoints to the inferred plane as opposed to a linear regression, which minimizes the distance of T_{gr} to the inferred linear function of L_{gr} and U_{gr}. We use a robust version of total least squares regression. After the first run, we remove the data points with the 5% largest residues to avoid the potentially detrimental influence of outlier values on the parameter estimation process (see Figure (26) and (27)).

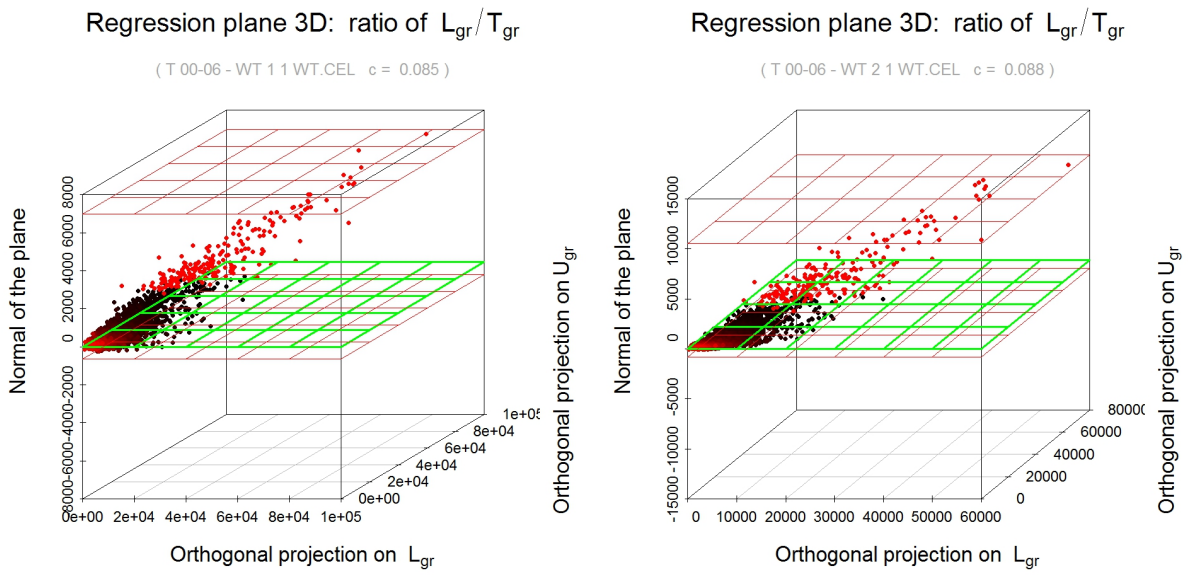


Figure 26: Two rounds of total least squares regression (*tls*) for two replicates of *S.cerevisiae* at a labeling time of 6 minutes. The resulting plane is colored green. The x-axis is chosen as the orthogonal projection on L_{gr}. The y-axis is the normal of the plane. The second round of *tls* is performed without the 5% largest residues of the 1^{st} round, depicted in red. Red planes indicate maximal residues.

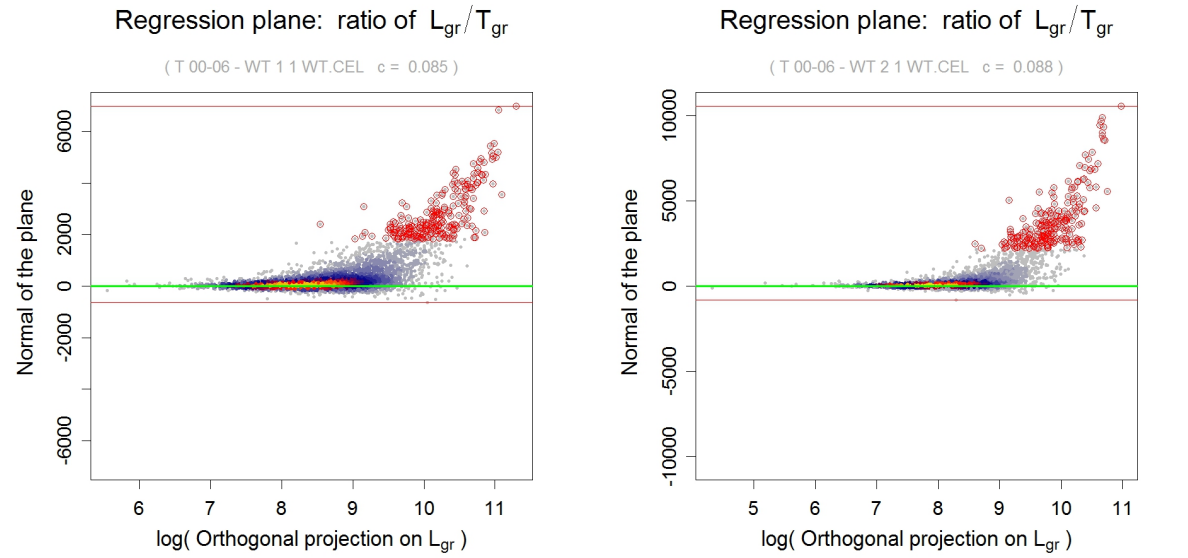


Figure 27: Two rounds of total least squares regression (*tls*) for two replicates of *S.cerevisiae* at a labeling time of 6 minutes in a different representation. The resulting plane is shown exactly from the side and is indicated by the green line. As in Figure (26), the x-axis is chosen as the orthogonal projection on L_{gr}, but in logarithmic scale. The y-axis is the normal of the plane. The 5% largest residues of the 1^{st} round are depicted in red. Red lines indicate maximal residues.

9.5.4 Estimation of the decay rate λ_{gr}

To estimate the decay rate λ_{gr}, we use equation (41) to yield

$$\lambda_{gr} = -\frac{1}{t_r} \log \left[\frac{T_{gr}(t_r) - \frac{c_r}{l_{gr} a_r} L_{gr}(t_r)}{T_{gr}(0)} \right] \tag{74}$$

for the dynamic case, and equation (47) to deduce

$$\lambda_{gr} = -\alpha - \frac{1}{t_r} \log \left[1 - \frac{\frac{c_r}{l_{gr} a_r} L_{gr}(t_r)}{T_{gr}(t_r)} \right] \tag{75}$$

for the steady state case. It is also possible to estimate the decay rate λ_g in an alternate way according to equation (48) as

$$\lambda_{gr} = -\alpha - \frac{1}{t_r} \log \left[\frac{\frac{c_r}{u_{gr} b_r} U_{gr}(t_r)}{T_{gr}(t_r)} \right] . \tag{76}$$

All three equations (74), (75) and (76) are subject to logarithmic calculation. The decay rate λ_g can not be calculated in cases where the quotient in the logarithm is ≤ 0. If this quotient reaches a certain size, the equations (74), (75) and (76) will yield negative decay rates which are then discarded. This circumstance is assessed in Figure (28). Each reasonable λ_g estimate can be used to calculate the half-life estimate

$$t_{1/2_{gr}} = \frac{\log(2)}{\lambda_{gr}} . \tag{77}$$

All measured samples r are combined to yield estimates

$$\lambda_g^{est} = \text{median} \left\{ \lambda_{gr} \mid r \in R \right\} , \tag{78}$$

and

$$t_{1/2_{gr}}^{est} = \text{median} \left\{ t_{1/2_{gr}} \mid r \in R \right\} . \tag{79}$$

The reproducibility of replicate measurements can be investigated by comparison of the quotient in the logarithm of equations (74), (75) and (76), see Figure (29).

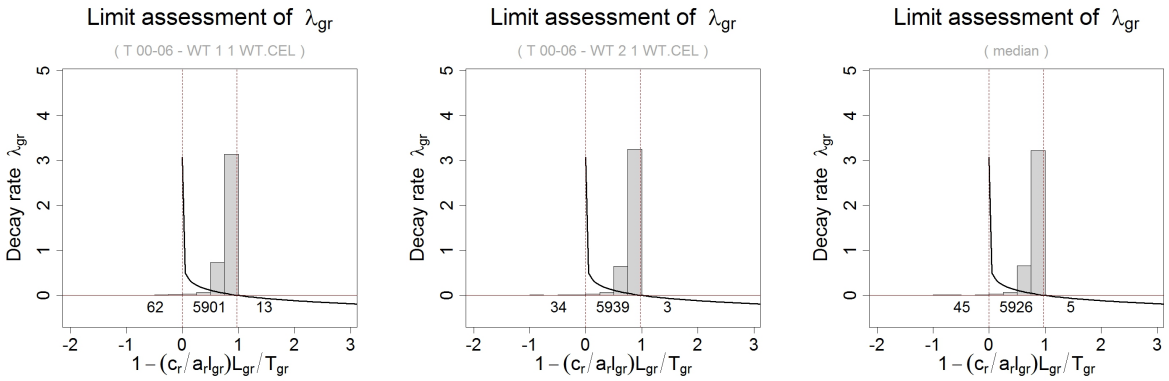


Figure 28: Dependency of the term $1 - \frac{c_r}{a_r l_{gr}} L_{gr}/T_{gr}$ to the resulting decay rate for two replicates of *S.cerevisiae* at a labeling time of 6 minutes. Reasonable decay rates can only be obtained for $1 - \frac{c_r}{a_r} L_{gr}/T_{gr}$ values between the two dashed lines.

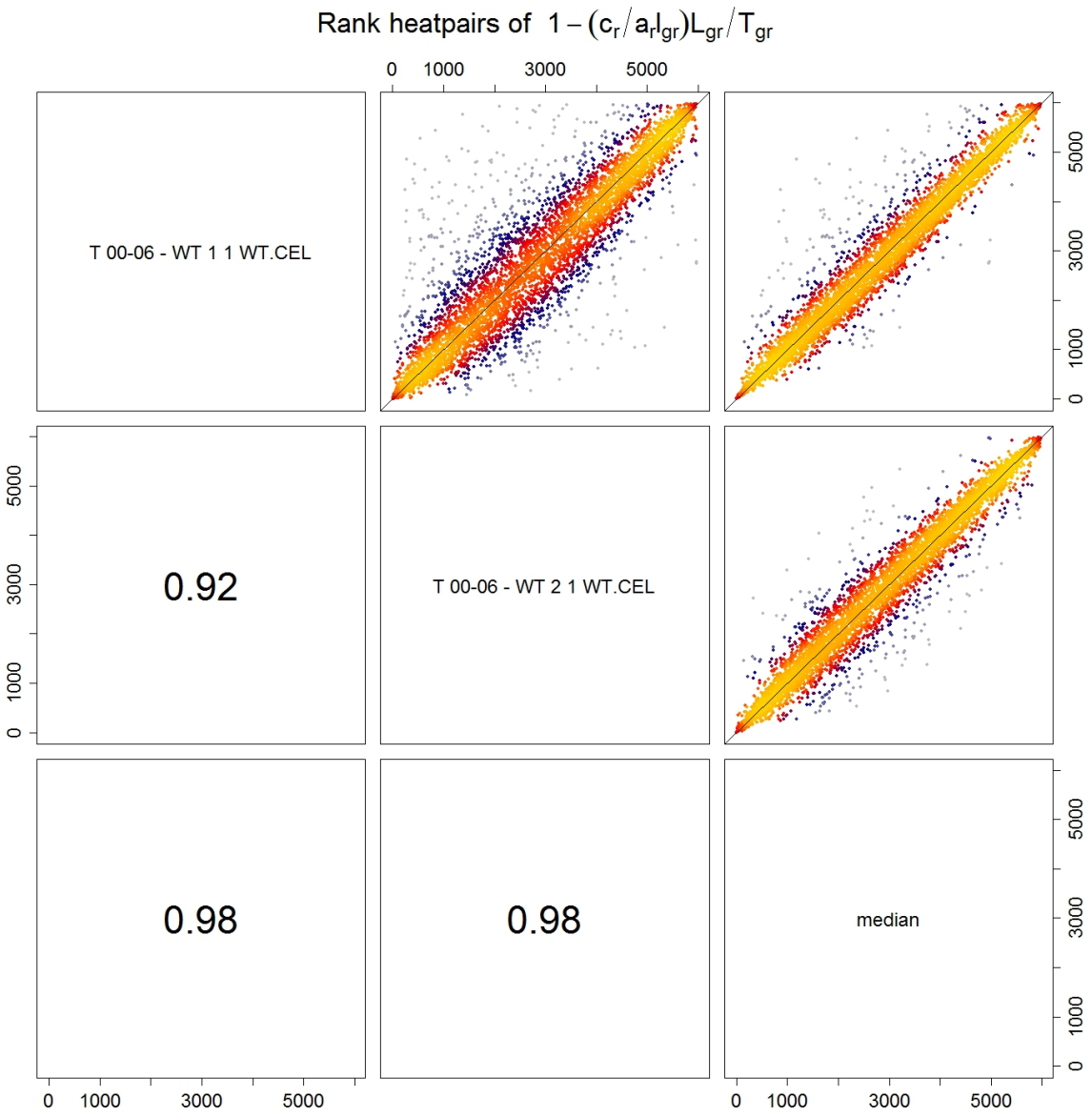


Figure 29: Pairwise heatscatter plots of the ranks of the respective $1 - \frac{c_r}{a_r l_{gr}} L_{gr}/T_{gr}$ value distributions for two replicates of *S.cerevisiae* at a labeling time of 6 minutes are shown in the upper panel. The lower panel shows the respective spearman correlations.

9.5.5 Estimation of the synthesis rate μ_{gr}

To estimate the synthesis rate μ_{gr}, we use equation (32) to yield

$$\mu_{gr} = \frac{\frac{c_r}{l_{gr} a_r} L_{gr}(t_r)(\alpha + \lambda_{gr})}{[e^{\alpha t_r} - e^{-\lambda_{gr} t_r}]} \ . \tag{80}$$

since $N_r(0)$ is set to 1, these values are on an arbitrary scale. For each replicate experiment r, the number m_{gr} of mRNA transcripts per cell of gene g is proportional to its total mRNA intensity value T_{gr}, $m_{gr} = d_r T_{gr}$. Assuming a total number of $\#mRNAs$ mRNAs per cell, this means that

$$\#mRNAs = \sum_{g \in G} m_{gr} = d_r \cdot \sum_{g \in G} T_{gr} \ , \tag{81}$$

and therefore

$$d_r = \frac{\#mRNAs}{\sum_{g \in G} T_{gr}} \ . \tag{82}$$

Together with equation (51), we may estimate μ_{gr} as

$$\mu_{gr}^{rescaled} = \frac{m_{gr}}{T_{gr}} \mu_{gr} \cdot CCL = d_r \mu_{gr} \cdot CCL \tag{83}$$

in molecules per cell and cell cycle. All measured samples r are combined to yield estimates

$$\mu_g^{est} = \text{median} \left\{ \mu_{gr} \ \mid \ r \in R \right\} \ . \tag{84}$$

To get an overview of the correlations of the measured and derived value distributions, see Figures (30) and (31).

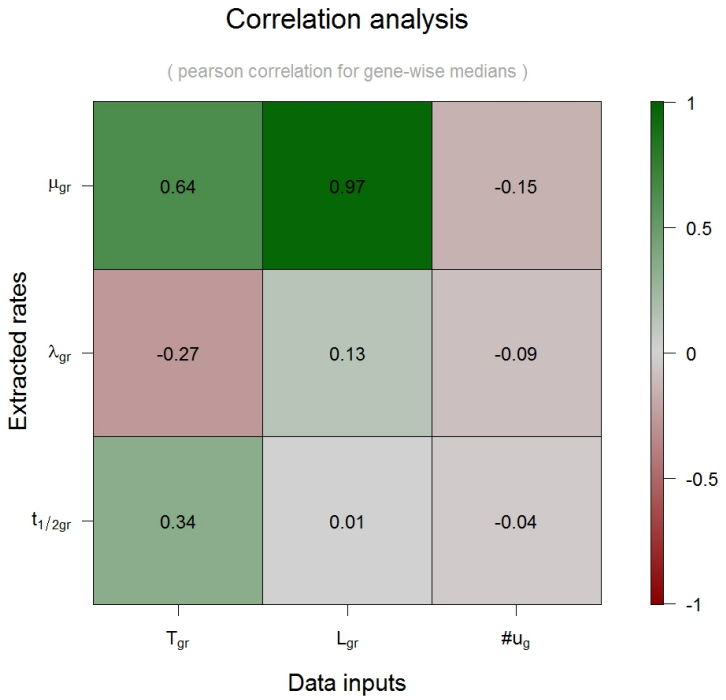


Figure 30: The pairwise correlations between the labeled expression values L_{gr}, total expression values T_{gr}, the number of uridines per transcript $\#u_g$ and the estimated synthesis rate μ_g^{est}, decay rate λ_g^{est} and half-life $t_{1/2_{gr}}^{est}$ is given in a color-coded image plot for the genewise median of two replicates of *S.cerevisiae* at a labeling time of 6 minutes .

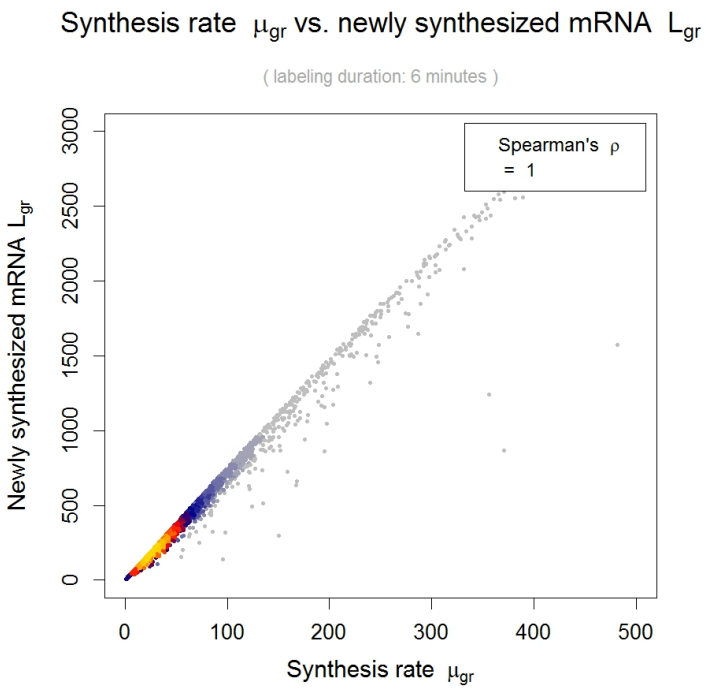


Figure 31: Comparison between the synthesis rate μ_{gr} and the labeled expression values L_{gr} of *S.cerevisiae* at a labeling time of 6 minutes. The Spearman correlation coefficient is 1. Given the short labeling time of 6 min, the differences are negligible, although we assume that the labeled RNA fraction is subject to degradation from the very time it is synthesized.

9.5.6 Variation analysis for the estimation of the median synthesis and decay rates with cDTA

Given the small number of replicate measurements per condition, we pooled the variance of the median absolute labeled and total RNA measurements respectively (4 wild-type, 2 ccr4, 2 pop2, 2 rpb1-N488D and 2 RPB1 measurements for each fraction). The calculated variance was subsequently used to resample the whole data set with medians randomly shifted according to this variance. This data was then re-analysed with cDTA (in particular Equations (75) and (80)). This procedure was performed 1000 times to obtain a distribution of median decay and synthesis rates. The 95% bootstrap confidence regions of the median synthesis and decay rates are shown in Figure 6D in the main text. Since these regions are clearly separated from each other, this proves that the observed coupling of decay rates and synthesis rates is not due to the variance of the estimators.

10 Supplementary Method S10 (Calculation of the relative 4tU incorporation efficiency)

We define three probabilities associated with mRNA labeling process: The incorporation efficiency p_{inc} is the probability that a 4tU nucleotide is incorporated into a nascent mRNA instead of a uridine. It is the product of the relative 4tU concentration (relative to the uridine concentration) in the nucleus and the relative affinity of the Polymerase for 4tU (compared to uridine). Since we have no handle to separate these two quantities, we merge them into p_{inc}. Secondly, the capture efficiency p_{cap} is the probability for a 4tU nucleotide which is included into a mRNA of being biotinylated, captured and recovered from streptavidin beads. Third, the probability that both events occur, assuming independence, is $p_{lab} = p_{inc} \cdot p_{cap}$. We call p_{lab} the labeling efficiency. A labeling efficiency substantially below 1 introduces a uridine-dependent labeling bias by letting newly transcribed, uridine-poor RNA have a higher probability to escape labeling. All three probabilities are sample- and strain-specific. The labeling efficiency can be estimated directly from cDTA data (see the folowing section). We can use cDTA to conclude from p_{lab} to the relative incorporation efficiencies. Note that for a sample x,

$$p_{cap}(x, Sc) = p_{cap}(x, Sp) \tag{85}$$

since the labeled *Sc* and *Sp* RNA from one sample are processed simultaneously. Moreover,

$$p_{inc}(x, Sp) = p_{inc}(y, Sp) \tag{86}$$

for two experiments x, y, since we use a common *Sp* standard for all experiments. Then,

$$\frac{p_{lab}(x, Sc)}{p_{lab}(x, Sp)} = \frac{p_{inc}(x, Sc) \cdot p_{cap}(x, Sc)}{p_{inc}(x, Sp) \cdot p_{cap}(x, Sp)} \overset{(85)}{=} \frac{p_{inc}(x, Sc)}{p_{inc}(x, Sp)} \tag{87}$$

Consequently, the relative incorporation efficiencies of two samples x and y is

$$\frac{p_{inc}(x, Sc)}{p_{inc}(y, Sc)} \overset{(86)}{=} \frac{p_{inc}(x, Sc)}{p_{inc}(x, Sp)} \cdot \frac{p_{inc}(y, Sp)}{p_{inc}(y, Sc)} \overset{(87)}{=} \frac{p_{lab}(x, Sc)}{p_{lab}(x, Sp)} \cdot \left(\frac{p_{lab}(y, Sc)}{p_{lab}(y, Sp)} \right)^{-1} \tag{88}$$

Equation (88) can be used to estimate the variation in the relative incorporation efficiency, a variataion estimate is given by

$$v = \text{std.dev.} \left(\frac{p_{inc}(x_1, Sc)}{p_{inc}(x_2, Sc)},\ x_j \in \text{replicates of group } j,\ j = 1, 2 \right) \tag{89}$$

cDTA reveals different in vivo 4tU labeling and incorporation efficiency upon mutation

It turns out that the relative incorporation efficiency $p_{inc}(mutant, Sc)/p_{inc}(wildtype, Sc)$ is less than 1 for all three mutants considered by us (Table 10).

Slow Pol II	Δccr4	Δpop2
0.51 ± 0.07	0.49 ± 0.05	0.71 ± 0.08

Table 4: Relative incorporation efficiencies were estimated by Equation (88), standard deviations were calculated using Equation (89).

Estimation of the labeling efficiency p_{lab} from the labeled and total mRNA expression data.

For the sake of self-containedness, we include a description of the estimation method for p_{lab} at this point. Actually, we applied a slightly improved version of the method presented here. The details are given in (Miller et al., Supplements Section 13). The probability that the mRNA of gene g containing $\#u_g$ uridines is *not* captured at all (again assuming independence of events for each uridine position) is $(1 - p_{lab})^{\#u_g}$. Conversely, the probability that a newly transcribed mRNA with $\#u_g$ uridines is captured in the labeled fraction is $1 - (1 - p_{lab})^{\#u_g}$. Let L_g and T_g be the intensities measured for the labeled resp. total mRNA of gene g (recall that cDTA normalization has been applied before, such that L_g and T_g equal the corresponding actual mRNA amounts in the resp. solution,

up to a common factor c. The amount of total mRNA in the sample is cT_g, the amount of newly transcribed mRNA is $cL_g\left[1-(1-p_{lab})^{\#u_g}\right]^{-1}$, where the factor $\left[1-(1-p_{lab})^{\#u_g}\right]^{-1}$ corrects for the mRNAs that escape the labeling/capture process. Let λ_g be the decay rate of gene g. Taking into account the growth rate α ($\alpha = \log 2$/cell cycle time) and the duration of the labeling process, the ratio of unlabeled mRNA over total mRNA is $\exp(-(\alpha+\lambda_g)\cdot t)$. It can be computed in a second way, using the amount of newly transcribed mRNA, namely

$$\exp(-(\alpha+\lambda_g)\cdot t) = \frac{cT_g - cL_g\left[1-(1-p_{lab})^{\#u_g}\right]^{-1}}{cT_g} = 1 - \frac{L_g}{T_g}\left[1-(1-p_{lab})^{\#u_g}\right]^{-1}$$

or, rearranging terms and taking logs,

$$\log\frac{L_g}{T_g} = \log\left[1-\exp(-(\alpha+\lambda_g)\cdot t)\right] + \log\left[1-(1-p_{lab})^{\#u_g}\right]$$

This equation holds for all genes. If we furthermore assume that there is no systematic dependence of the decay rate λ_g on the uridine content of g, we can model the L_g/T_g ratio as

$$\log\frac{L_g}{T_g} = \log\left[1-(1-p_{lab})^{\#u_g}\right] + \epsilon_g$$

with i.i.d errors ϵ_g, $g \in$ genes. It follows from this equation that a robust estimate of p_{lab} can be obtained as the minimizer

$$\hat{p}_{lab} = \underset{p}{\operatorname{argmin}} \sum_{g\in\text{genes}} \left|\log\frac{L_g}{T_g} - \log\left[1-(1-p)^{\#u_g}\right]\right|$$

Note

Plots have been produced using the open source R package LSD [10]

The analysis workflow has been carried out using the open source R/Bioconductor package DTA [11]

References

[1] Maria J. Amorim, Cristina Cotobal, Caia Duncan, and Juan Mata. Global coordination of transcriptional control and mRNA decay during cellular differentiation. *Molecular Systems Biology*, 6, June 2010.

[2] E. Birney, D. Andrews, M. Caccamo, Y. Chen, L. Clarke, G. Coates, T. Cox, F. Cunningham, V. Curwen, T. Cutts, T. Down, R. Durbin, X. M. Fernandez-Suarez, P. Flicek, S. Gräf, M. Hammond, J. Herrero, K. Howe, V. Iyer, K. Jekosch, A. Kähäri, A. Kasprzyk, D. Keefe, F. Kokocinski, E. Kulesha, D. London, I. Longden, C. Melsopp, P. Meidl, B. Overduin, A. Parker, G. Proctor, A. Prlic, M. Rae, D. Rios, S. Redmond, M. Schuster, I. Sealy, S. Searle, J. Severin, G. Slater, D. Smedley, J. Smith, A. Stabenau, J. Stalker, S. Trevanion, A. Ureta-Vidal, J. Vogel, S. White, C. Woodwark, and T. J. Hubbard. Ensembl 2006. *Nucleic acids research*, 34(Database issue), January 2006.

[3] Mally Dori-Bachash, Efrat Shema, and Itay Tirosh. Coupled Evolution of Transcription and mRNA Degradation. *PLoS Biol*, 9(7):e1001106+, July 2011.

[4] Jorg Grigull, Sanie Mnaimneh, Jeffrey Pootoolal, Mark D. Robinson, and Timothy R. Hughes. Genome-Wide Analysis of mRNA Stability Using Transcription Inhibitors and Microarrays Reveals Posttranscriptional Control of Ribosome Biogenesis Factors. *Mol. Cell. Biol.*, 24(12):5534–5547, June 2004.

[5] L. Hereford. Number and distribution of polyadenylated RNA sequences in yeast. *Cell*, 10(3):453–462, March 1977.

[6] F. C. Holstege, E. G. Jennings, J. J. Wyrick, T. I. Lee, C. J. Hengartner, M. R. Green, T. R. Golub, E. S. Lander, and R. A. Young. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell*, 95(5):717–728, November 1998.

[7] Silvia Jimeno-González, Line Lindegaard L. Haaning, Francisco Malagon, and Torben Heick H. Jensen. The yeast 5'-3' exonuclease Rat1p functions during transcription elongation by RNA polymerase II. *Molecular cell*, 37(4):580–587, February 2010.

[8] Christian Miller, Björn Schwalb, Kerstin Maier, Daniel Schulz, Sebastian Dümcke, Benedikt Zacher, Andreas Mayer, Jasmin Sydow, Lisa Marcinowski, Lars Dölken, Dietmar E. Martin, Achim Tresch, and Patrick Cramer. Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. *Molecular systems biology*, 7, January 2011.

[9] Sarah E. Munchel, Ryan K. Shultzaberger, Naoki Takizawa, and Karsten Weis. Dynamic profiling of mRNA turnover reveals gene-specific and system-wide regulation of mRNA decay. *Molecular Biology of the Cell*, 22(15):2787–2795, August 2011.

[10] Bjoern Schwalb, Achim Tresch, and Romain Francois. LSD Lots of Superior Depictions. *The Comprehensive R Archive Network*, 2010.

[11] Bjoern Schwalb, Benedikt Zacher, Sebastian Duemcke, and Achim Tresch. DTA Dynamic Transcriptome Analysis. *The Comprehensive R Archive Network and Bioconductor*, 2011.

[12] Ophir Shalem, Orna Dahan, Michal Levo, Maria R. Martinez, Itay Furman, Eran Segal, and Yitzhak Pilpel. Transient transcriptional responses to stress are generated by opposing effects of mRNA production and degradation. *Molecular Systems Biology*, 4(1), October 2008.

[13] Yulei Wang, Chih Long L. Liu, John D. Storey, Robert J. Tibshirani, Daniel Herschlag, and Patrick O. Brown. Precision and functional specificity in mRNA decay. *Proceedings of the National Academy of Sciences of the United States of America*, 99(9):5860–5865, April 2002.

[14] Daniel Zenklusen, Daniel R. Larson, and Robert H. Singer. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nature Structural & Molecular Biology*, 15(12):1263–1271, November 2008.