



מכון ויצמן למדע

WEIZMANN INSTITUTE OF SCIENCE

Thesis for the degree
Master of Science

עבודת גמר (תזה) לתואר
מוסמך למדעים

Submitted to the Scientific Council of the
Weizmann Institute of Science
Rehovot, Israel

מוגשת למועצה המדעית של
מכון ויצמן למדע
רחובות, ישראל

By
Manor Yair

מאת
מנור יאיר

צימוד כלל-גנומי בין שעתוק ופירוק של רנ"א שליח,
ודעיכה מתואכת-פרומוטר של רנ"א שליח בשני מיני שמר מרוחקים.
Genome-wide coupling of mRNA transcription and degradation,
and promoter-mediated mRNA decay in two distant yeast species.

Advisor:
Prof. Yitzhak Pilpel

מנחה:
פרופ' יצחק פלפל

February 2012

שבט תשע"ב

Table of contents

Table of contents	2
Abstract:	3
Introduction:	4
Goals:	6
Methods:	7
Strains and plasmids used	7
Media and growth conditions	7
Expression calibration experiments	7
Optimization of time point selection	9
Measuring mRNA expression and decay rates in <i>S. pombe</i>	9
RNA extraction and microarray hybridization	10
Microarray data analysis and global scaling of decay data	10
Calculating Half-lives	11
Calculating inferred transcription rates	12
Calculating evolutionary conservation (CPA) scores	12
Optimizing colony scan	13
General procedures	14
Results:	15
Changes in mRNA abundance and decay in response to oxidative stress	15
Comparing intra-species and –mutant-strains parameters	17
Inter-Species comparison of expression parameters reveals that while transcription and decay rates vary, expression level remains similar	20
The effect of promoter architecture on mRNA decay	22
A potential effect of transcription rate and co-decay on evolutionary conservation of promoter architecture:	29
Measuring the effects of different promoters on the decay of identical transcripts using synthetic gene constructs	32
Discussion:	35
Literature:	39

Abstract:

Genes' expression levels are governed by the balance between mRNA transcription and degradation. Unlike mRNA transcription, regulation of mRNA decay in response to environmental stress is poorly studied. Recent studies suggest that regulation of mRNA decay might be directly coupled to transcription. In this research, we set out to examine whether the coupling between mRNA transcription and degradation is mediated by genes' promoter information and whether it is evolutionarily conserved for long periods of time.

We performed microarray experiments measuring mRNA expression and decay in the fission yeast *S. pombe*, in normal conditions and under oxidative stress, with and without transcription inhibition and compared results to those obtained in *S. cerevisiae*. Our results showed genome-wide coupling of mRNA transcription and degradation under stress, which becomes particularly pronounced in subsets of genes responding to the applied stress. This counter-intuitive coupling of increased transcription along with increased degradation affects the response kinetics, causing a fast response and allowing fast relaxation back into the ground state after the stress is relieved.

Comparing between *S. pombe* and *S. cerevisiae* revealed that evolutionary changes between the two species in transcription are counter-acted by changes in decay, thus maintaining relatively similar expression levels of orthologous genes in the two species. A mechanism coupling transcription to decay would facilitate co-evolution of the two parameters, stabilizing expression levels, and thus possibly facilitate evolution.

Thus, we show that the coupling between transcription and degradation is evolutionarily conserved between two distant yeast species, diverged circa 500 million years ago, both in unperturbed conditions and in response to stress.

Yet what regulates this observed coupling? An emerging possible coordinator of both transcription and degradation is the promoter. In order to further research the promoter's effects on mRNA decay we examined adjacent gene couples that share promoter architectures. We found modest correlations in the decay rates of promoter-sharing genes. Examination of subsets of evolutionarily conserved promoter architectures revealed that high decay similarity might impose an evolutionary pressure for the conservation of promoter architecture, indirectly implying an intriguing regulatory link to the effects of promoters on mRNA decay.

Introduction:

The cell's transcriptome changes drastically in response to environmental changes (1) as a large fraction of the genes' mRNA abundance increases or decreases. Even under constant conditions, when mRNA levels reach a steady-state, the transcriptome is still dynamic as mRNA transcription and degradation are constantly taking place. The balance between these two counter-acting forces governs the kinetics and expression levels of each gene (2), and as a consequence, the cell's state. A widely used method to monitor the cell's state and measure genome-wide mRNA levels is the use of micro arrays, throughout different time-courses, under different stresses, and in different species and mutant strains. A more modern emerging alternative is the deep sequencing of the transcriptome (3; 4). These methods are generally used to study transcription, as the measured changes in genes' mRNA abundance are often attributed predominantly to changes in transcription, ignoring the contribution of changes in decay to the overall abundance.

Contrary to the attention dedicated to study mRNA transcription, until recently mRNA decay has been poorly studied. Advancements in recent years have shed light on many mechanisms that promote degradation of aberrant mRNA, e.g. nonsense mediated decay (NMD), “no-go decay” and “non-stop decay” (5; 6), yet the exact regulation of standard mRNA decay that changes half-life times in response to a changing environment remains vague.

Recent studies show that these changes in half-life times are coupled to changes in transcription, and together affect the cellular response in very intriguing ways. Intuitively, an increase in the mRNA abundance can be obtained by increasing transcription, decreasing degradation, or a combination of both, and vice-versa for a decrease in mRNA abundance. Less intuitively, a large increase in transcription, together with a smaller *increase* in degradation, could also lead to the same total increase in mRNA abundance yet with different kinetics (2). As previously shown by our lab in yeast (7) and by others in additional species (8), many of the genes whose mRNA abundance increases when exposed to oxidative stress are also de-stabilized. This intriguing counter-intuitive coupling between transcription and degradation, working in opposite directions, and the faster kinetics derived from it (2), might be suitable for transient responses to oxidative stress. In contrast, the slower response, in yeast, induced by exposure to the DNA damaging agent, MMS, did not show this kind

of coupling, but rather the more intuitive coupling whereby most induced genes were also stabilized at the mRNA level.

This is not the only case in which a coupling between mRNA transcription and degradation was seen. A different research (9; 10) discovered that two sub-units of RNA polymerase II (Pol II), Rpb4p and Rpb7p, play a role not only in transcription, but also in specific and non-specific mRNA decay (respectively), possibly providing a physical link between degradation and transcription.

The coupling seen between transcription and degradation, and the existence of shared elements that govern the two processes led us to explore the possibility that the promoter might be responsible for regulating not only transcription, but also mRNA degradation, and thus facilitate coupling between mRNA transcription and degradation, possibly in a stress-specific manner. Most recently, two studies (11; 12) have shown that the promoter can indeed regulate mRNA degradation rates by showing that identical transcripts preceded by different promoters have different decay kinetics. These findings, albeit shown for only few genes, support the hypothesis that promoter elements can affect transcript decay rates, even when the transcript is already located in the cytoplasm.

To support this hypothesis, additional evidence is needed, to show that this phenomenon is not restricted to only few genes in one yeast species. For that we decided to take a genome-wide approach, and look at the entire transcriptome of an organism, rather than focusing on few genes. For that we examined two yeast species, *S. cerevisiae* and the fission yeast *Schizosaccharomyces Pombe*. The fact that the two species have diverged from one another some ~500 Million Years (MY) ago (13), yet share many orthologous genes, makes them perfect candidates to examine genome-wide promoters effects on mRNA decay and to examine whether this strategy of regulation is evolutionarily conserved.

Goals:

1. To examine whether the effects of the promoter on decay, which were shown for only few genes, can be detected on a genome-wide scale in both *S. cerevisiae* and *S. pombe*.
2. To further explore the intriguing counter-action coupling between mRNA transcription and degradation in *S. pombe*.
3. To examine the extent of evolutionary conservation of the coupling between transcription and degradation, and to further research the possibility that the promoter is capable of such coupling.
4. Use the special case of adjacent genes, which often share a promoter to examine the promoter effect on mRNA degradation.

Methods:

Strains and plasmids used

Microarray experiments were carried out on *Schizosaccharomyces pombe* 972 h⁻ wild-type. Work on synthetic gene constructs was done on *Saccharomyces cerevisiae* wild-type strain BY4741 (*MATa his3Δ1, leu2Δ0, met15Δ0, ura3Δ0*). In datasets taken from previous works (7; 14) experiments were carried out, by others, on *Saccharomyces cerevisiae* *rpb6*^{Q100R} mutant strain (*MATa ura3-52, his3Δ200, lys2Δ201, ade2, RPB6Δ::HIS3 pRPB6/CEN/LEU2*) (15), its isogenic parental wild-type strain WY37 (*MATa ura3-52, his3Δ200, lys2Δ201, ade2*), and *Saccharomyces cerevisiae* Y262 *rpb1-1* mutant (*MATa ura3-52, his4-539, rpb1-1*) (16).

Hygromycin resistance gene was PCR amplified from pFA6a-hph-NT1 plasmid (17). Synthetic gene constructs were cloned into pGEM-T vectors (PROMEGA) and integrated into *E. coli* DH5a strain. Plasmid were then purified using QIAGEN mini-prep kit according to manufacturer's protocol, and cleaved using *Pst*I or *Eco*RI (NEB) restriction enzymes.

Media and growth conditions

S. pombe cells were grown at 30°C in YES (5g/l yeast extract, 30 g/l glucose, and 225 mg/l amino acids supplements: adenine, histidine, leucine, uracil and lysine hydrochloride, as described in (18)). *S. cerevisiae* cells were grown at 30°C in YPD (10g/l yeast extract, 20g/l bacto peptone, 20g/l dextrose) or SD (6.7g/l yeast nitrogen base, 20g/l dextrose and supplemented with appropriate nutrients). Bacto Agar (18g/l) was added for solid media. Ura⁻ colonies were selected on SD complete medium with uracil (50mg/l) and 5-fluoroorotic acid (5-FOA, 0.8g/l) (19). Can^R colonies were selected on SD-Arg with canavanine (50mg/l). Hyg^R colonies were selected on YPD medium supplemented with hygromycin (0.6g/l).

Expression calibration experiments

Calibration experiments were done on several parameters prior to conducting the main experiment on *S. pombe*: The strength of the oxidative stress applied, the concentration of 1,10-phenanthroline used to arrest transcription, and the time gap between application of the stress and transcription arrest.

In order to calibrate the strength of oxidative stress, survival assays were performed. Mid-log cultures of *S. pombe* cells were treated with H₂O₂ in final concentrations of

0.33mM, 0.5mM or 1mM, from a 10mM stock solution. Samples were taken before applying the stress (t_0), and 15, 30, 60, and 90 minutes afterwards, diluted in double-distilled-water to two different cell concentrations, and sown in duplicates on solid media. After 2-3 days, colonies were counted, and survival rates in each H₂O₂ concentration were calculated as the fraction of the number of colonies from each sample normalized to pre-stress (t_0) colony count. Concentration of 0.33mM was selected due to similarity of its survival rates, 85%-90%, to those measured for *S. cerevisiae* in previous works ((7; 14), data not shown)

In order to quantify the exact concentration of 1,10-phenanthroline needed to arrest transcription in *S. pombe*, cell cultures were grown to mid-log phase, treated with 100, 150, or 300 µg/ml of 1,10-phenanthroline (from a freshly-made stock of 40 mg/ml, dissolved in ethanol 100%), and further grown in 96-well plates. Each 96-well plate contained 48 repetitions of samples treated with the drug, and 48 repetitions of non-treated samples, evenly dispersed throughout the plate, to avoid localization effects. Growth was monitored by frequent optical density measurements using Tecan Infinite® F500 plate-reader by the Hamilton robotic system. Cells treated with 100 µg/ml 1,10-phenanthroline exhibited slow yet constant growth, while cells treated with 300 µg/ml exhibited no growth at all, possibly indicative of immediate high cell-death ratios, and therefore too strong a stress. However, Cells treated with 150 µg/ml 1,10-phenanthroline showed minimal growth for a limited time-window, suggesting transcription arrest without causing immediate cell-death, leading us to choose this concentration of 1,10-phenanthroline for our main experiment.

In order to calibrate the time gap between application of oxidative stress and transcription arrest to match the time to reach circa 80% of the full expression reaction, the mRNA abundance levels of a few specific genes in response to oxidative stress was measured using real-time PCR. These specific genes were selected for showing a strong response to oxidative stress, in a data set taken from Chen et al. (20). A mid-log culture of cells was treated with 0.33mM H₂O₂ and 8ml samples were taken before (duplicate samples), and 5, 10, 15, 20, 30, 40, 50, 60, 90 and 120 minutes after application of oxidative stress, centrifuged and immediately snap-frozen in liquid nitrogen. RNA was extracted using MASTERPURE Yeast RNA purification kit, from which cDNA was prepared using random primers. The cDNA amounts were measured using the LightCycler 480 real time PCR machine (Roche diagnostics). LightCycler 480 SYBR green was used as the reagent at the volume of 10µl per

reaction. All samples were measured in duplicate wells. The mRNA levels of the selected oxidative-stress responsive genes were measured and normalized to the response profile of the reference gene, the known housekeeping gene ACT1. The normalized expression profiles peaked approximately 10-15 minutes after the stress. This short response-time, together with the desire to keep experimental conditions as close as possible to comparable previous experiments (14), led us to choose a 7-10 minutes time gap between oxidative stress and transcription inhibition, for our main experiment.

Optimization of time point selection

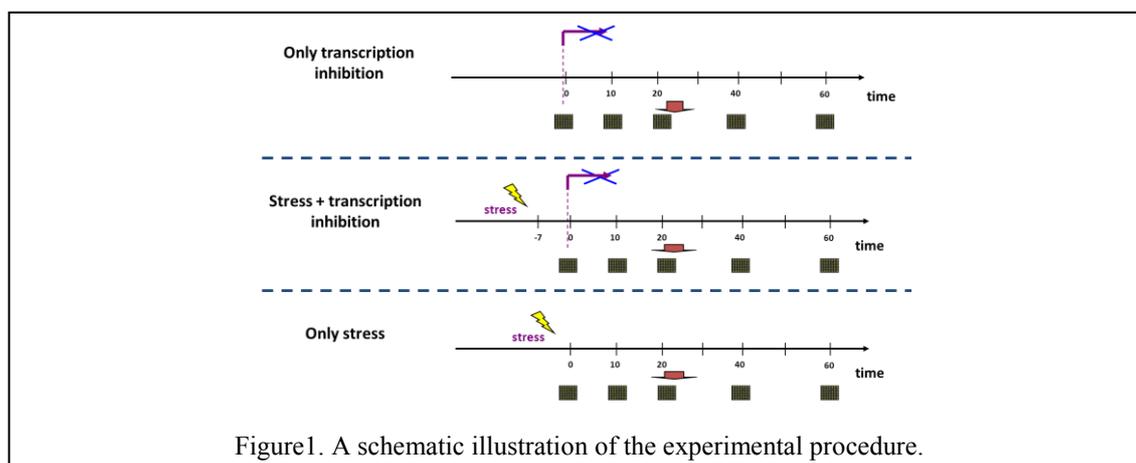
In order to choose the optimal configuration of time points to be taken, an optimization algorithm was written and implemented on data from a similar previous experiment performed in our lab in *S. cerevisiae* (14). For that raw data of a decay time-course containing seven time points were taken. Time-points were omitted one at a time, or two at a time, to create all possible time courses each lacking measurements on one, or two of the time points. Decay profiles were fitted for all genes in each of the time courses and their half-lives were calculated. All half-lives of each lacking time course were then compared to half-lives calculated from the complete time-course, and Pearson correlation coefficients of the half-lives over all genes were computed.

This analysis showed that the most important time points were the first and last, with average Pearson correlation of 0.94 and 0.97 between original half-lives and half-lives obtained after omitting the first and last points, respectively. Omitting all other time points hardly affected the computed half-life, with average Pearson correlation of 0.995 between original half-lives and half-lives obtained after omitting one or two of the other time points. Additional factors were taken into consideration, such as the median half-life (21) and the model exponential decay profile. Integrating all these factors, the following five-point time course was chosen: 0, 10, 20, 40 and 60 minutes, after transcription arrest.

Measuring mRNA expression and decay rates in *S. pombe*

A starter of *S. pombe* culture of 50ml that was grown over-night to a concentration of $\sim 2 \times 10^8$ cells/ml, was then diluted into 500ml of YES, and was further grown on 30°C for ~ 5 hrs until a concentration of 1×10^7 cells/ml was reached. Cells were counted, visually verified to be in mid-log state, and to contain no bacterial

contamination. The cell-culture was then separated into 3 vessels, one for each time-course: One to measure the expression profile in response to an oxidative stress (this sample is called “treated”), a second to measure the normal, non-treated decay profile, and a third to measure the decay profile after the same oxidative stress (Figure 1). Hydrogen peroxide (H₂O₂) in a final concentration of 0.33mM was administered to both the expression and decay treated profiles. For the decay experiments transcription was arrested by addition of the drug 1,10-Phenanthroline, to a final concentration of 150μM. The drug was added 7-10 minutes after the application of the Hydrogen peroxide in the case of the “treated” decay profile, and was also added for a decay experiment in the control non-treated population. Duplicate aliquots of 5ml were taken of each time course in five different time points, centrifuged and immediately snap-frozen in liquid nitrogen.



RNA extraction and microarray hybridization

RNA was purified from frozen samples using MASTERPURE™ yeast RNA purification kit (EPICENTER biotechnologies) according to the manufacturer instructions and its quality verified using Bioanalyzer 2100 platform (AGILENT). For each time-course, five time points were chosen (as described in "Choosing time points optimization" section) and sent to hybridization with Affymetrix yeast 2.0 microarrays. Both the Bioanalyzer and microarray hybridization procedures were done by the Biological Services department at the Weizmann Institute.

Microarray data analysis and global scaling of decay data

All microarray results were processed with the RMA preprocessing algorithm using MATLAB's bioinformatics tool box. As previously described (14; 22), during the time course, transcription is arrested and total mRNA levels are decreasing, but this

decrease is masked by the experimental protocol, as equivalent amounts of total RNA are extracted from each sample. Previous studies that used a PolII mutant strain (7) could circumvent this problem since mRNAs constitute only a minor fraction of the total RNAs in yeast cell, and the transcription of other RNAs by RNA polymerase I and III was not inhibited. However, 1,10-phenanthroline appears to inhibit all three RNA polymerases to approximately the same extent. We therefore scaled our decay profiles data at each time point according to an overall exponential decay with half-life of 30 minutes, consistent with previous measurements of *S. pombe* decay data (21). Accordingly, each time point in the decay profiles was scaled such that mean intensity over the time course would decay according to the expected reference profile. To improve the fit to exponential decay, scaled decay profiles were further normalized to the first time point and exponential fit was performed on these normalized profiles, to extract the decay coefficient of each gene. Unlike the decay profiles, the expression profile of each gene was simply normalized to its first time point, to reflect the changes in response to the stress.

Calculating Half-lives

The expression level of a gene is governed by the opposing forces of transcription and degradation. Since transcription is often assumed to be independent of the number of the mRNA transcripts, it is modeled as a zero-order reaction, whose rate coefficient we will term β . In contrast, mRNA decay is assumed to be linearly dependent of the number of the mRNA transcripts and is therefore modeled as a first-order reaction, whose rate coefficient we will term α . We can therefore write a differential equation for the rate of change of the mRNA levels of a gene X as a function of time:

$$\frac{dx}{dt} = \beta - \alpha X .$$

If we assume that the mRNA level reaches a steady state ($\frac{dx}{dt} = 0$) during the exponential growth phase, then according to this model mRNA level just before transcription arrest, termed X_0 , is equal to $\frac{\beta}{\alpha}$. After transcription inhibition, β becomes 0, and mRNA level start to decline. Integrating the above differential equation will give us the expression level of the gene as a function of time after transcription inhibition:

$$X(t) = X_0 \cdot e^{-\alpha t}$$

We can see that what governs the kinetics of the mRNA levels is the decay coefficient α , out of which the gene's half life time, $t_{\frac{1}{2}}$, can be easily calculated: $t_{\frac{1}{2}} = \frac{\ln(2)}{\alpha}$.

After pre-processing and normalizations of our data (as described above), we used MATLAB to fit each gene's decay profile to the exponential decay equation above, extracting the decay coefficient α for each gene, and using it, calculated the half-life time of each gene.

We then filtered the calculated half-lives, taking for further analysis only genes with R^2 goodness-of-fit values over 0.8, and whose half-life value was between 0 and 120 minutes, which constitute ~90% of all the genes.

Calculating inferred transcription rates

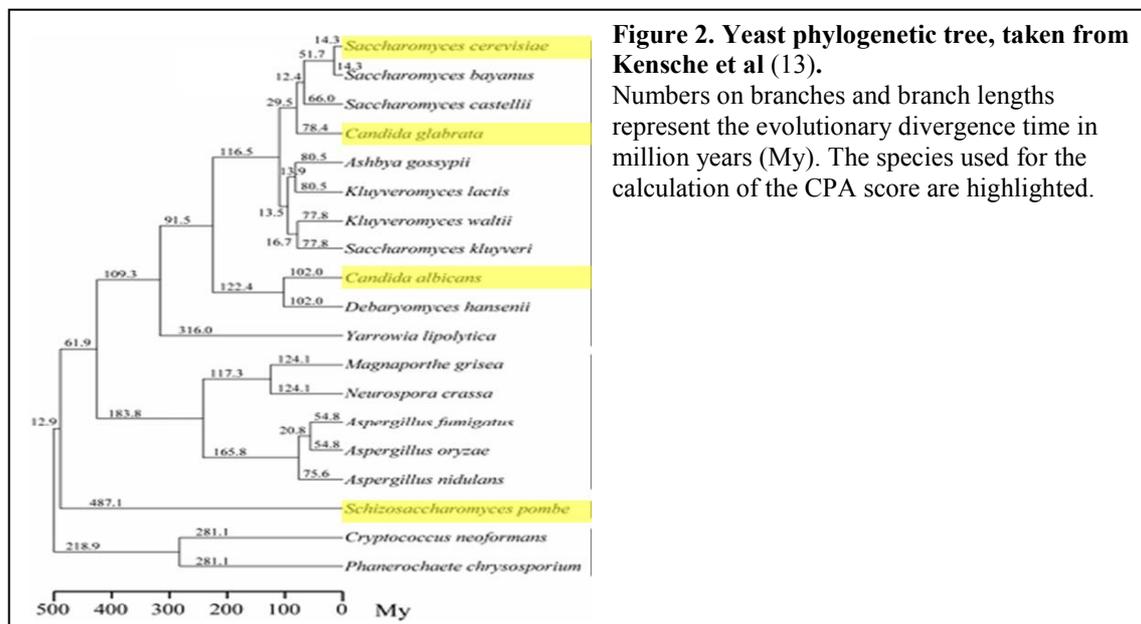
In order to better understand the interactions between mRNA transcription and degradation, and detect transcription-dependent phenomena otherwise masked out by expression levels, we wished to refine our analyses and distinguish between transcription and expression. This refinement is supported by a recent study (22) that showed stronger coupling between degradation and inferred transcription than between degradation and expression. For that reason inferred transcription rates, β , were calculated from the multiplication of decay rates, α , by steady-state expression levels, X_0 ($\beta = X_0 \cdot \alpha$), for all the strains, as derived above for steady-state conditions.

Calculating evolutionary conservation (CPA) scores

For each pair of promoter-sharing genes in each of the three architectures, we calculated a score to indicate the extent of preservation of architecture between species, hereby termed **Conserved Promoter Architecture** score, or **CPA** score. Genomic data of four yeast species was integrated in this score: *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Candida albicans*, and *Candida glabrata*. These species were selected for spanning across a broad phylogenetic range of yeast species, up to ~500My by common estimates (see Figure 2, taken from (13)). For each species, all protein-coding genes were divided into adjacent gene pairs, according to their genomic location, and further divided into one of three paired promoter architectures (divergent, convergent or tandem) according to their relative orientation to one another. Two published yeast orthologous gene tables were used: The first is *S. cerevisiae* -centric, in the sense that all orthology assignments started

from neighboring gene pairs in this species, taken from Man et al. (23). The second table is *S. pombe*-centric, taken from Wapinski et al. (24).

Using the orthologous gene tables and genomic data, we checked for each ortholog gene couple in *S. cerevisiae* and *S. pombe* whether this couple is still paired, and also conserves its paired promoter architecture in any of the other species mentioned. Gene couples that did not stay coupled or did not conserve their promoter architecture between two species got a CPA score of 0. For gene couples that did conserve their architecture in a different species, the evolutionary distance between the two species (in My) was added to their CPA score. Finally, in order to avoid over-estimation of the conservation distance, adding a third or fourth conserved specie, added to the CPA scores only the minimal distance between the added species and any of the other conserved species. Given the species currently included in this analysis, this caution was only necessary in the *S. pombe*-centric analysis. For example, a gene couple conserved between *S. pombe*, *C. albicans*, and *C. glabrata* will receive a score of 711.5My and not 974.2My. This score is comprised of 487.1My between *S. pombe* and one *Candida* species, plus 224.4My between the two *Candida* species, and not another 487.1My for conservation between *S. pombe* and the other *Candida* species.



Optimizing colony scan

During this study we encountered the need to optimize a screening of many yeast colonies by PCR for the presence of a construct, in order to minimize the number of PCR reactions performed. More specifically, after transformation of the synthetic constructs (see results), correct integration had to be verified by PCR. Initial PCR

verifications on a few colonies revealed a high number of false-positive colonies that grew despite the selective medium, forcing us to scan a large number of colonies. This large number was further multiplied by the number of different constructs and tripled again by the number of different primer designs, reaching a total of more than a thousand PCR reactions.

In order to minimize the number of reactions and reduce the time until a verified colony is found, colony pooling was used. Colony pooling is possible since positive results are not masked out by negative results, i.e. if one or more colonies in a certain "pool" were to contain the inserted construct the pooled PCR reaction will yield a product. Subsequently, in a second phase, PCR reactions were to be performed on individual colonies from the positive pool, in order to isolate the positive colony containing the construct.

But what is the optimal number of colonies to pool? There seems to be a simple trade-off: increasing the number of colonies pooled together lowers the number of PCR reactions in the first phase, but increases the subsequent, second phase of PCR reactions within the positive pool. In order to find the optimal pool size I have calculated the number of PCR reactions performed under each possible pool size and looked for the minimum of this function. The optimal setup, under assumption of constant pool size, was found to be five colonies in eight pools with respect to minimal number of reactions. The same minimal number of reactions was also obtained by a pool size of eight colonies in five pools, yet this alternative pooling would require additional invested working time. These two optimal pool sizes found are the reciprocal divisors closest to the square root of the number of colonies, and indeed for different initial conditions the optimal pool size converges to the square root of the number of colonies.

I have thus ended up pooling into eight pools of five colonies each.

General procedures

Standard molecular biology procedures such as restriction enzyme, bacterial transformation, were carried out as described in Sambrook et al. (25). Yeast media and molecular biology procedures (transformations, DNA purifications, etc.) were done as described in Sherman (26). Plasmids were purified using mini-prep kit supplied by QIAGEN following manufacturer instructions.

Results:

Changes in mRNA abundance and decay in response to oxidative stress

One of the common methods to measure decay rates is to stop transcription, e.g. using the drug 1,10-Phenanthroline. We wanted to measure the decay kinetics in the fission yeast *S. pombe*, both under normal, non-treated conditions and under an oxidative stress conditions. Furthermore, in order to investigate a possible connection between transcription and degradation, we also wished to measure the expression profile under these conditions, yet without stopping transcription. Therefore, our experiment was composed of three time courses: two in which we measured mRNA decay rates, of which one is under an oxidative stress and one in an untreated control condition. In a third time course we applied the same oxidative stress yet with no transcription arrest. Time point zero in this time course represents expression level in an unperturbed condition. The two decay time courses were treated with the drug 1,10-Phenanthroline at the time defined as time zero (Figure 1), and the treated decay time course was pre-treated with hydrogen peroxide (0.33mM final concentration) 7-10 minutes prior to the transcription arrest. The same hydrogen peroxide treatment was also given to the third time course in which we did not stop transcription, at the time defined (for this time course) as time zero.

Figure 3A shows the changes in each gene's half-life before and after treatment with oxidative stress, with a strong correlation of 0.8 between the two conditions. Two questions regarding the experimental setup arise from the high correlation: a. Was the oxidative stress strong enough? Perhaps the high correlation indicates a mild response to the stress given. b. Was the time gap (7-10 minutes) between oxidative stress and transcription arrest too short? It is possible that the cells were not given enough time to fully react to the stress before transcription was arrested. This possibility is supported by the observation that when looking at the expression time course more than 90% of the genes peak at 20-40 minutes (Figure 4). Note that the pre-experiment calibrations (using qRT-PCR) showed a shorter response peak time of 10-15 minutes according to which a time-gap of 7-10 minutes was chosen (see materials and methods). A combination of these two reasons is of course also possible.

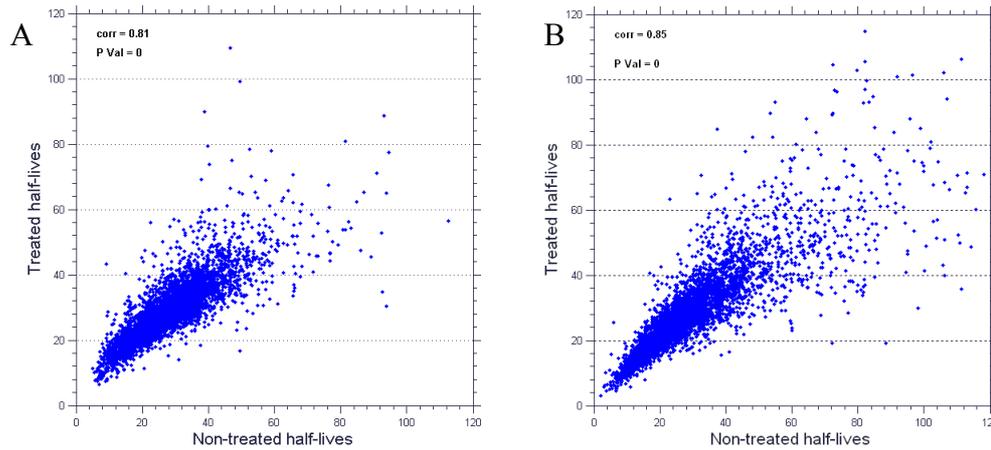


Figure 3. Considerable correlation between treated and non-treated half-lives. Scatter plot of the filtered half-lives calculated for the treated condition Vs. Non-treated condition in *S.pombe* (A) and *S.cerevisiae* (B, data taken from Shalem et al. (14)). Pearson correlation coefficient and p-value are annotated by 'corr' and 'P val', respectively, throughout this work.

Examining data from previous work (7; 14) in three *S. cerevisiae* strains (wild-type, *rpb1-1* mutant, and *rpb6* mutant) revealed similar Pearson correlations, of 0.8-0.85, under similar conditions (Figure 3B shows data from *S. cerevisiae* wild-type). This similarity between the correlations shows remarkable conservation of the extent of the decay response to stress in two such remote species.

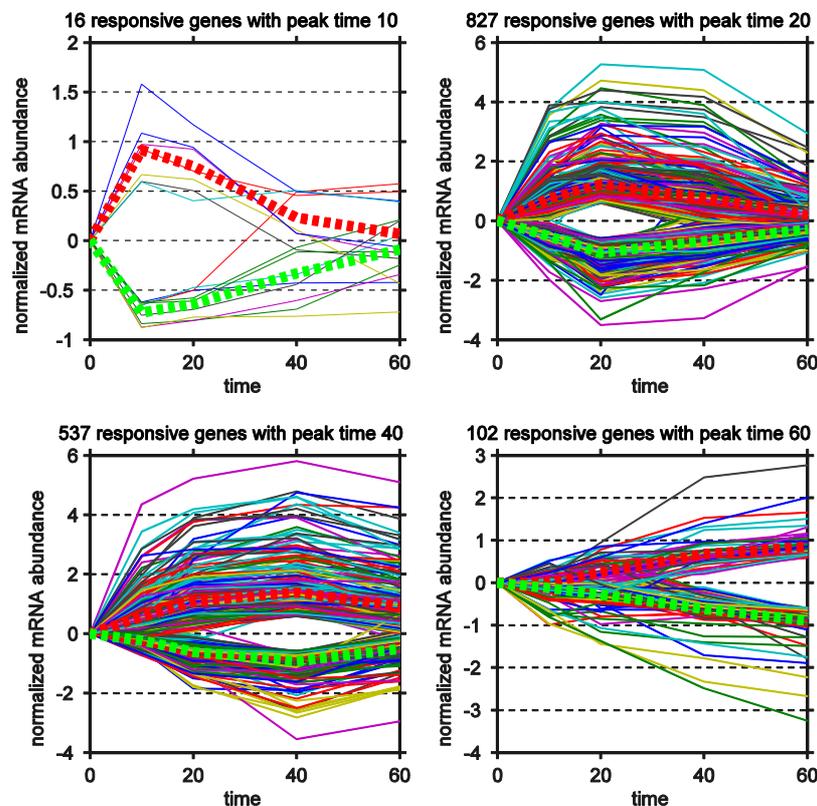


Figure 4. Expression responsive genes divided according to their peak-time. Expression profiles of responsive (>1.5 fold) genes in *S.pombe* in each peak-time, with mean profile of induced and repressed genes in red and green dashed bold lines, respectively.

Comparing intra-species and -mutant-strains parameters

In order to look for a possible correlation between transcription and degradation, we calculated each gene's maximal expression fold change in response to oxidative stress, and plotted that value against the change in half-lives in response to the stress (Figure 5A). A weak negative correlation was found that indicate that stress-induced genes have a low tendency to be de-stabilized in the stress, while stress-repressed genes show a weak tendency to be stabilized. We were looking for a potential sub-set of the genes that may show a stronger correlation. In particular we examined the most responsive genes only. Responsive genes were defined here as genes that change their expression over a certain threshold (1.5 fold) in at least one of the time points following the stress relative to time point zero. Indeed responsive genes show a somewhat higher correlation between mRNA level change and decay rate change compared to the rest of the genes (Figure 5B).

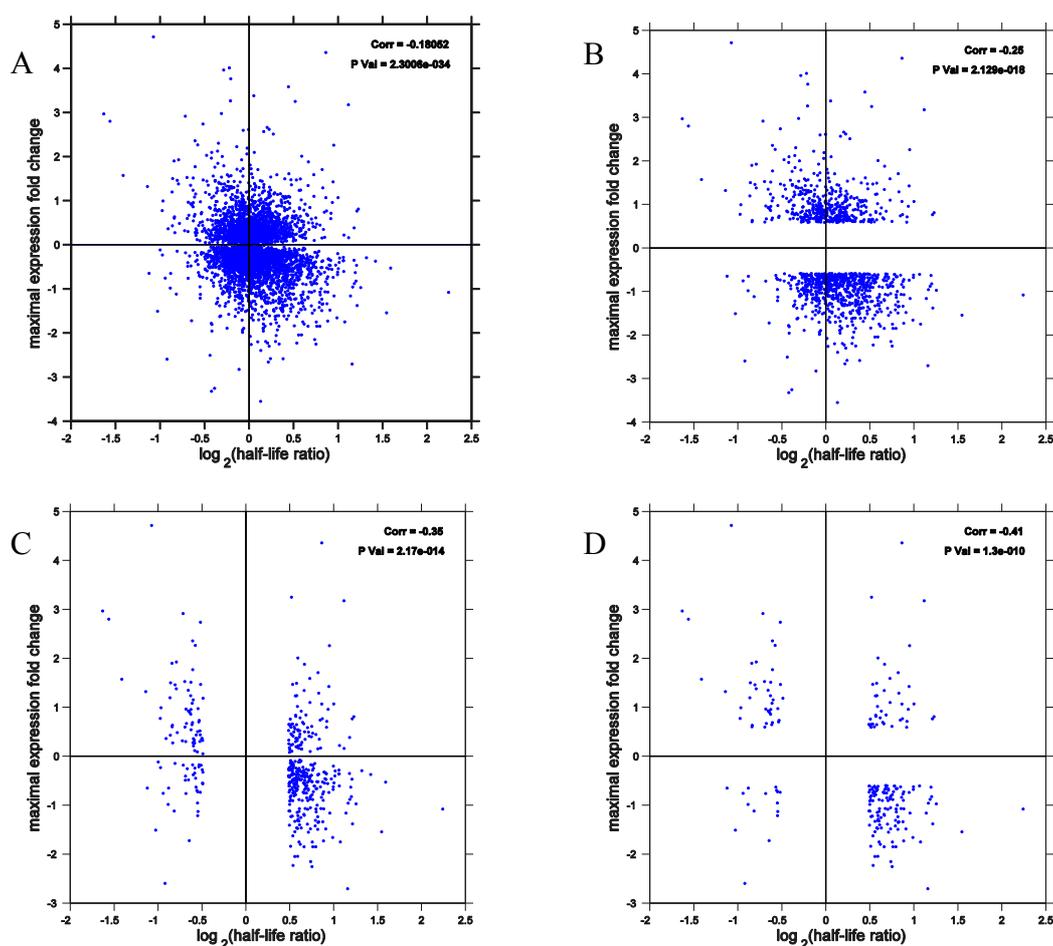
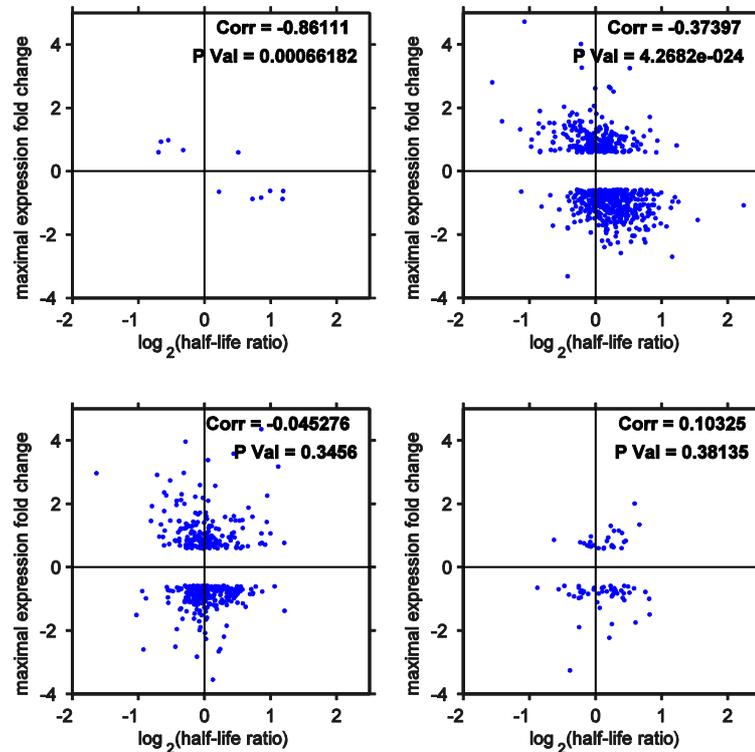


Figure 5. Weak coupling between transcription and degradation in all genes strengthens for responsive genes subsets. Maximal expression fold change Vs. Half-life ratio for all genes (A), expression responsive subset(B), decay responsive subset (C), and the intersection of both subsets (D)

We next looked at the subset of decay-responsive genes, i.e. genes that their half-life between treated and non-treated conditions, changes over a certain threshold (1.4 fold change). When looking at these decay-responsive subset of genes (Figure 5C), the negative correlation increases (Pearson correlation = -0.35). Finally, when looking at a subset of the genes that are responsive both in term of expression level and in terms of half-life (Figure 5D), the correlation increased only a little further (up to -0.41). We have next partitioned this later subset of responsive genes according to their time-to-peak, i.e. the time in which they reach their maximal absolute expression fold-change. This refinement reveals that fastest-peaking genes, those that reach their maximal response within 10-20 minutes after the stress, show strong negative correlation between expression and decay, while late-peaking genes show no significant correlation (Figure 6). We thus concluded that in similarity to the case in *S. cerevisiae* (7; 14), in *S. pombe* too fast peaking expression profiles are obtained by counter-acting the change in transcription with a change in stability of the mRNA and in particular that fast induction is often accompanied by stress-induced de-stabilization.

A



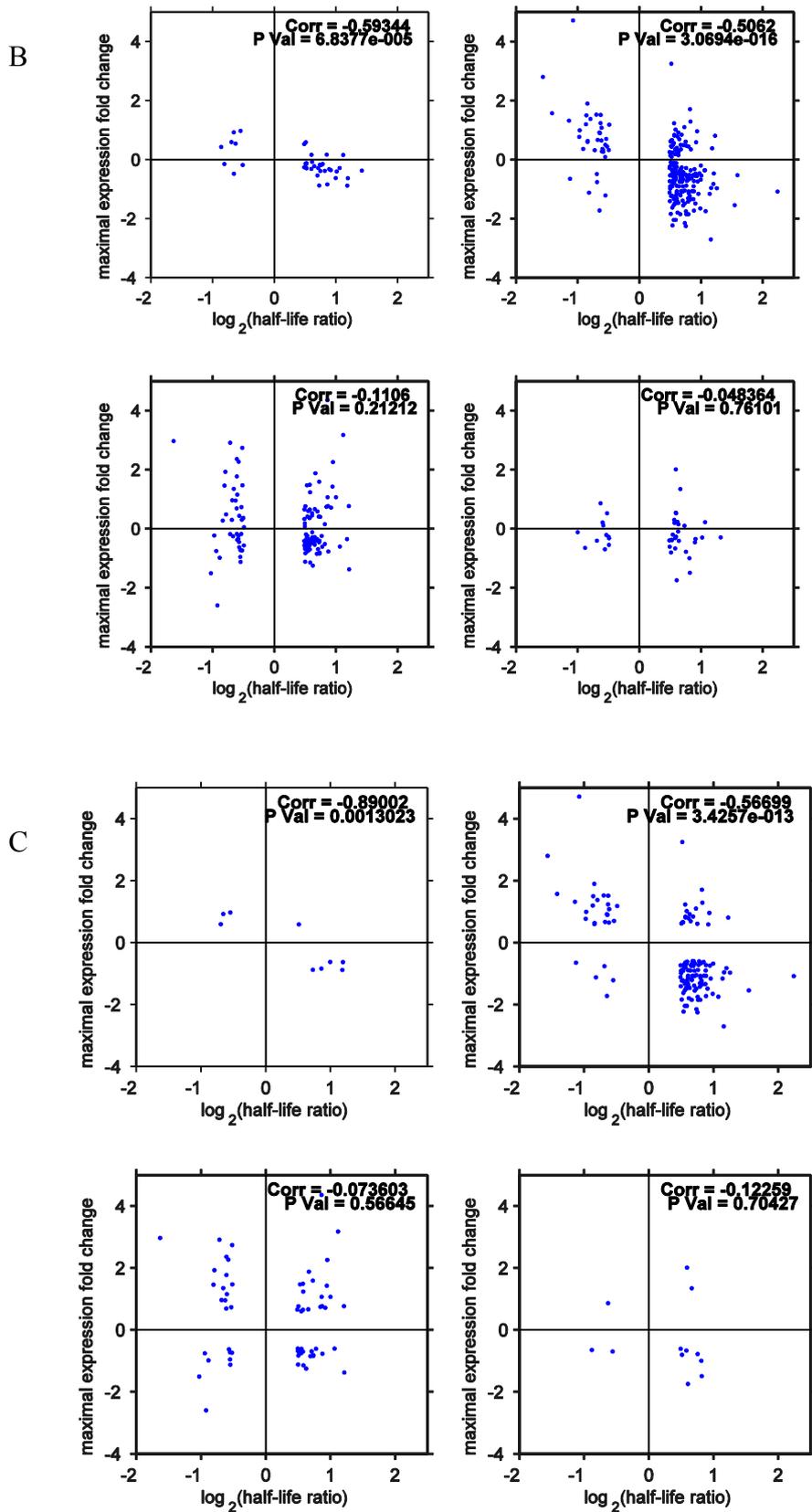


Figure 6. Fast-peaking genes show strong coupling between expression and half-life ratios. Maximal expression fold change Vs. Half-life ratio for genes in the expression responsive subset (A), decay responsive subset (B), and the intersection of both subsets (C), divided by their time-to-peak. In each plot, genes peaking at 10, 20, 40, and 60 minutes are plotted in the top-left, top-right, bottom-left and bottom-right subplots, respectively.

Inter-Species comparison of expression parameters reveals that while transcription and decay rates vary, expression level remains similar

It is conceivable that selection is acting on expression levels more than on rate of transcription and decay. Thus since steady state expression level is given by the ratio of transcription and degradation, looking at mRNA expression levels might mask out some differences between species at the transcription and decay rates. In particular, a counter-action coupling between transcription and degradation might be masked out. Indeed, a recent study (22) showed stronger coupling between degradation and (inferred) transcription than between degradation and expression. For that reason inferred transcription rates were calculated from the combination of decay data and steady-state expression level data, for all the strains (see Materials and Methods). In order to validate our inferred transcription rates we compared them to transcription rates obtained directly in *S. cerevisiae* using a recently developed metabolic labeling method (27). We used our non-treated data from *S. cerevisiae* wild-type in order to keep the comparison conditions as similar as possible. Interestingly, we obtained a high correlation of 0.75 between our inferred transcription rates and the directly measured rates (Figure 7). This is in contrast to the low correlation showed between the metabolic labeling-based inferred half-life and half-lives measured directly by several transcription arrest methods (see comparison in Cramer et al. (27)). We have thus concluded that we can reliably compute transcription rates given decay rates and steady-state levels.

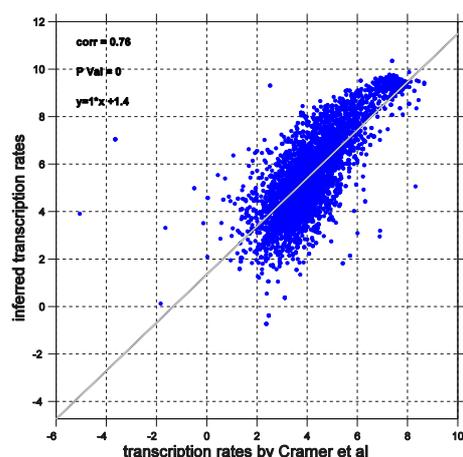


Figure 7. Strong correlation between our inferred transcription and Cramer's. Our transcription scores, inferred for *S. cerevisiae* under non-treated conditions, compared to transcription rates taken from Cramer et al. (27), also for *S. cerevisiae* in similar conditions. Both axes are in \log_2 scale. Least-square line was also plotted, and its equation appears on the graph

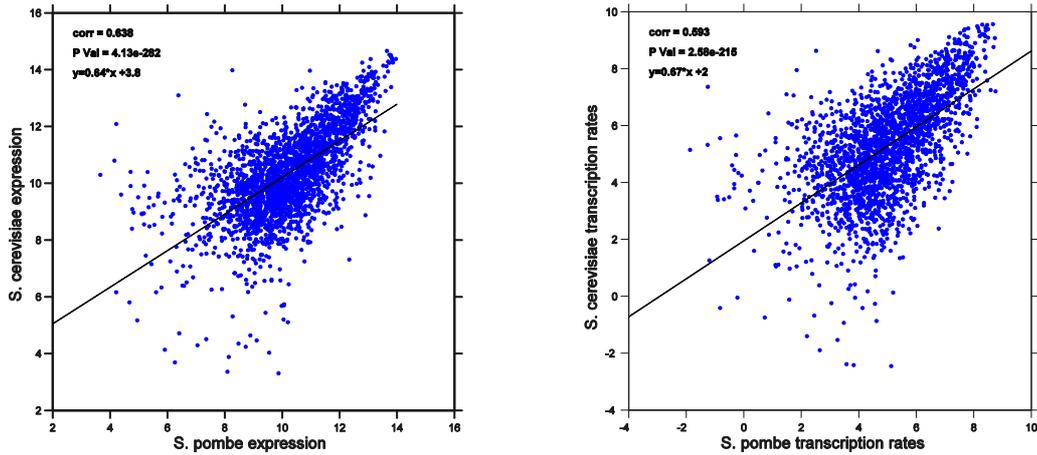


Figure 8. Strong correlation between *S. pombe* and *S. cerevisiae* in expression levels and transcription rates. Comparing steady state, untreated expression levels (left) and transcription rates (right) of *S. pombe* and *S. cerevisiae*. Only orthologs with no paralogs in either species were taken.

We could then proceed to a comparative evolutionary analysis and compare expression levels, transcription rates, and decay rates for orthologous genes of *S. cerevisiae* and *S. pombe* (to ensure true unequivocal orthology I examined only genes with no paralogs in either species). Both the expression level and transcription rates (β) of orthologous genes correlate considerably between the two species (Pearson = 0.64 and 0.6, respectively, Figure 8). For comparison, different strains of *S. cerevisiae*, e.g. *S. cerevisiae* wild-type and *rpb6* mutant, correlate in expression and transcription up to 0.99 and 0.97 (data not shown). The surprising finding is that despite this strong correlation between expression levels, the correlation of the α decay coefficients between those species, and strains, is considerably lower, around 0.33 when comparing the two species, and 0.82 when comparing the two *S. cerevisiae* strains (Figure 9).

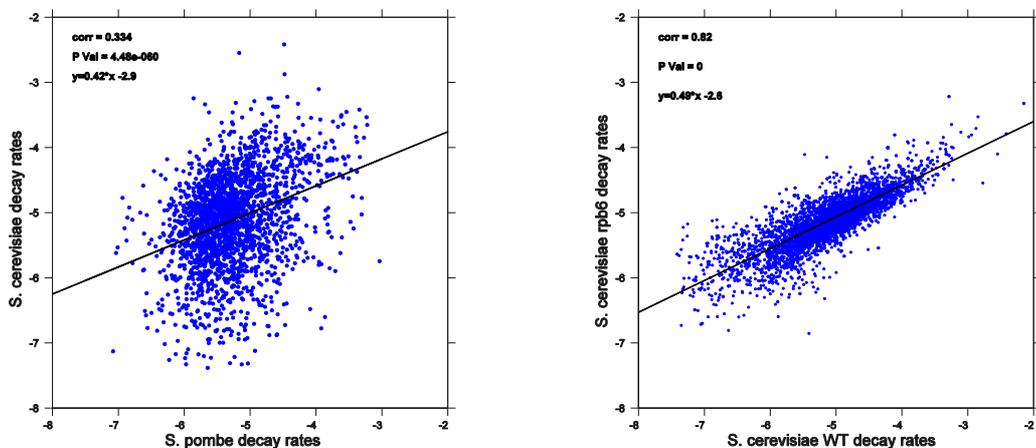


Figure 9. Correlation of decay rates for different species (left) and strains (right).

I thus concluded that despite changes in decay and transcription rates between strains and species, mRNA expression level at steady-state is relatively conserved. A counter-action between production and degradation may be the mechanism responsible for this conservation in expression – changes in production are counter-acted by opposite changes in decay rates. Indeed, when looking at the inter-species change in transcription parameters (β) for each gene, it is highly correlated to the inter-species change in decay parameters (α , Figure 10). It is tempting to speculate that such counter-action was selected evolutionarily as a built-in means to ensure constant steady-state level despite changes in transcription and decay rates.

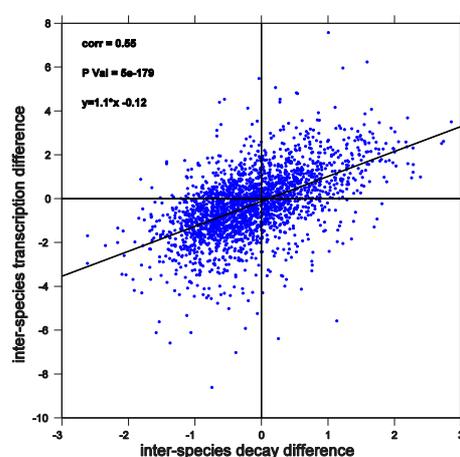


Figure 10. Changes in production are counter-acted by opposite changes in decay rates. Log2 of the inter-species ratio of transcription rates (*S. pombe* / *S. cerevisiae*) on the y-axis was plotted against the inter-species ratio of degradation rates (*S. pombe* / *S. cerevisiae*).

The effect of promoter architecture on mRNA decay

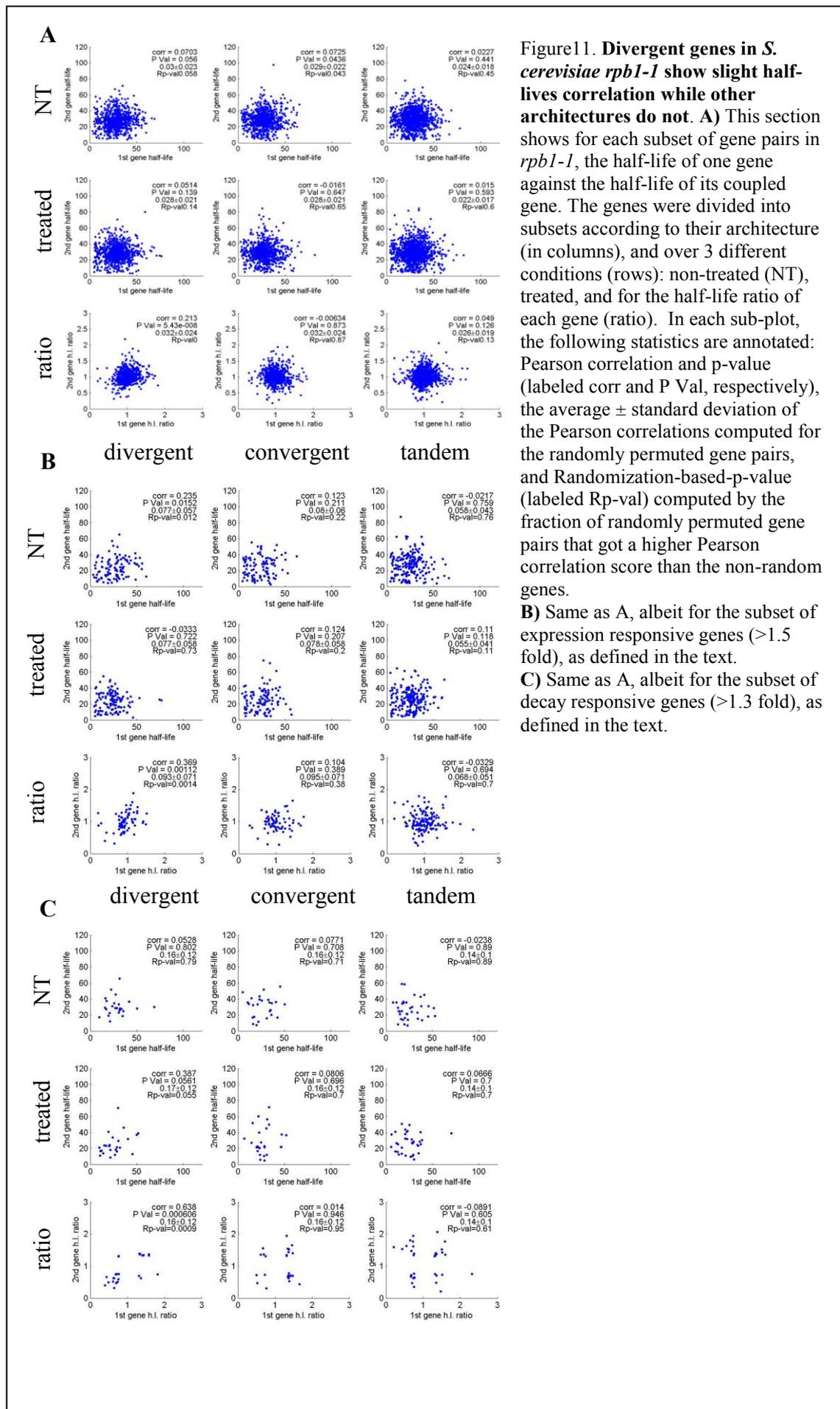
These results bring the question of how transcription and degradation are coupled in cells. It was shown recently that a likely possible coordinator of transcription and degradation is the promoter (11; 12). Yet these studies have demonstrated the role of promoters in controlling decay by inspection of a few genes only. Here we attempted to examine the potential effect of promoters on mRNA decay at the genome-wide scale. For that purpose we focused on neighboring pairs of genes in the genomes of *S. cerevisiae* and *S. pombe*. A sub-section of such genes the so-called “divergent promoters” share an intergenic region that controls the transcription of two divergent genes on opposite strands of the chromosome. It was previously shown that divergent gene pairs tend to have similar expression profiles (28; 13) compared to random pairs of genes, though this signal was relatively modest in its statistical significance. In addition were examined two alternative promoter architectures that connect between

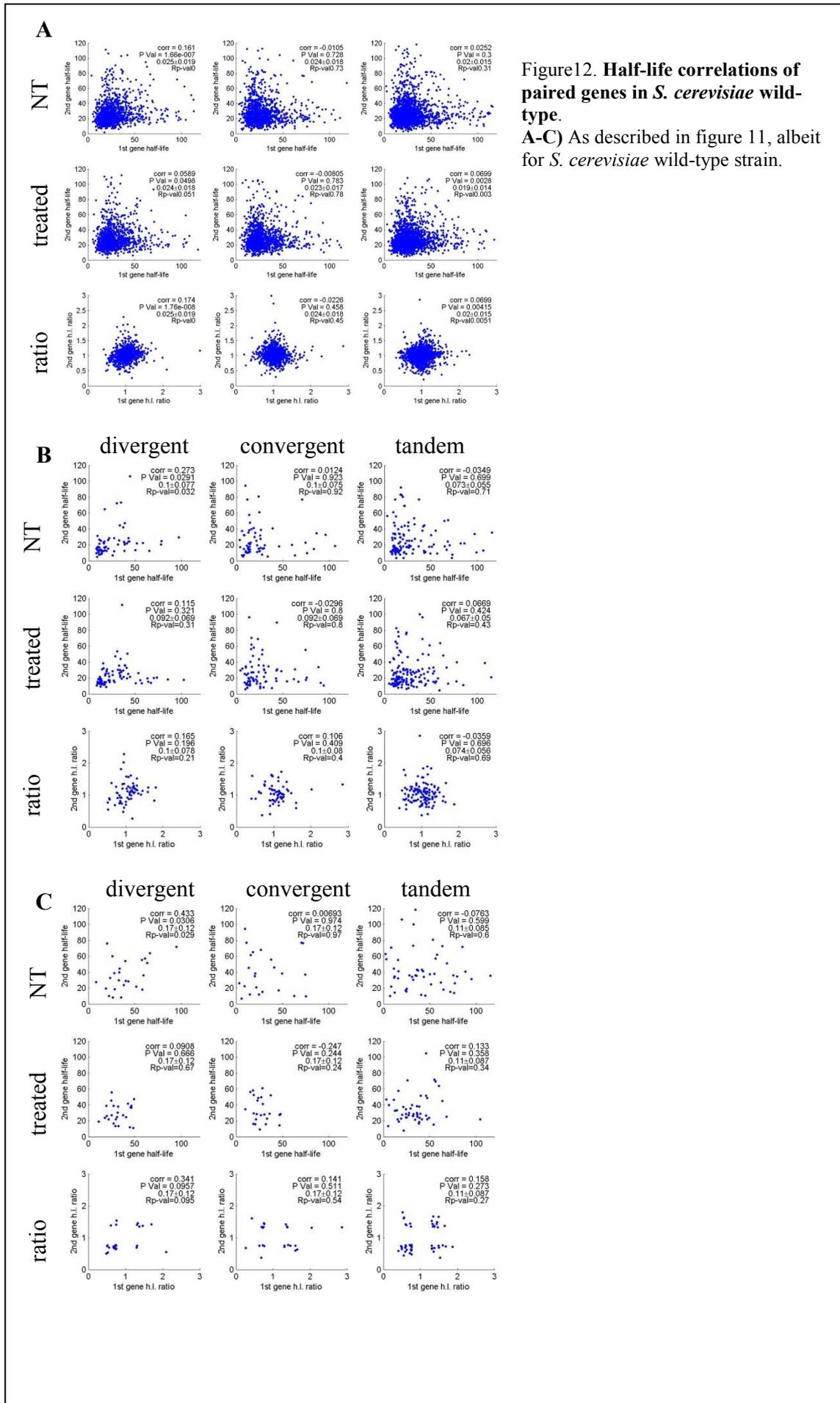
neighboring genes, the “convergent” in which two adjacent genes are transcribed on opposite strands yet the intergenic regions between them is at their 3’ end, and “tandem” in which the pair of adjacent genes are on the same strand. It was previously shown that a bit less than the divergent pairs, the tandem pairs, and to a lesser extent the convergent pairs, show similarity in expression compared to random pairs of genes (28). The fact that not only divergent pairs, but also tandem and convergent pairs show similarity in expression suggests that promoters can exert long distance effects or that entire chromosomal domains may affect expression. In this thesis we will use the term promoter architecture to refer to the possible relative orientations of adjacent gene couples, and mark them by: divergent ($\leftarrow \rightarrow$), convergent ($\rightarrow \leftarrow$), or tandem ($\rightarrow \rightarrow$).

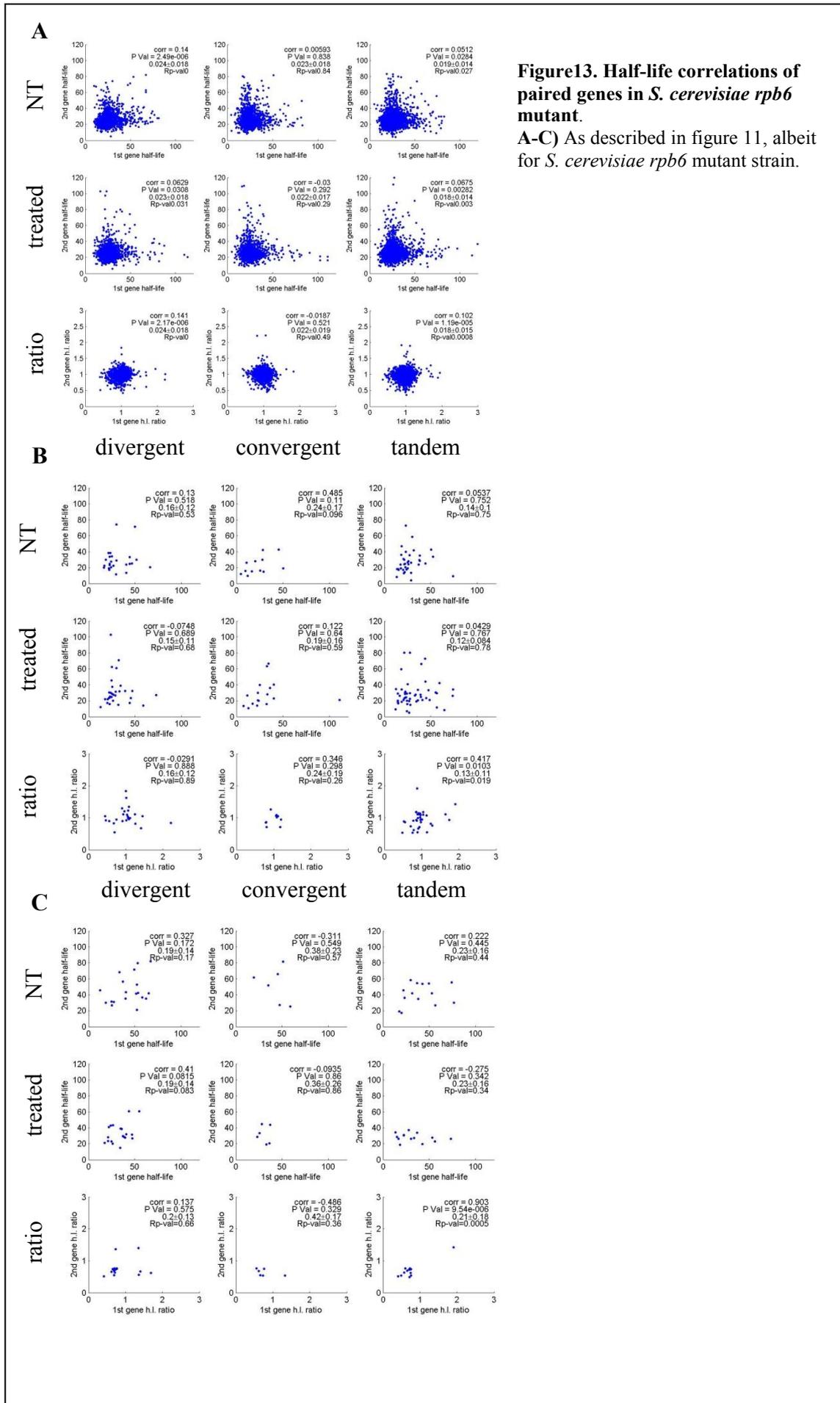
We hypothesized that if promoters do indeed affect mRNA decay then such pairs of genes, and in particular the divergent ones might exhibit similar mRNA decay dynamics.

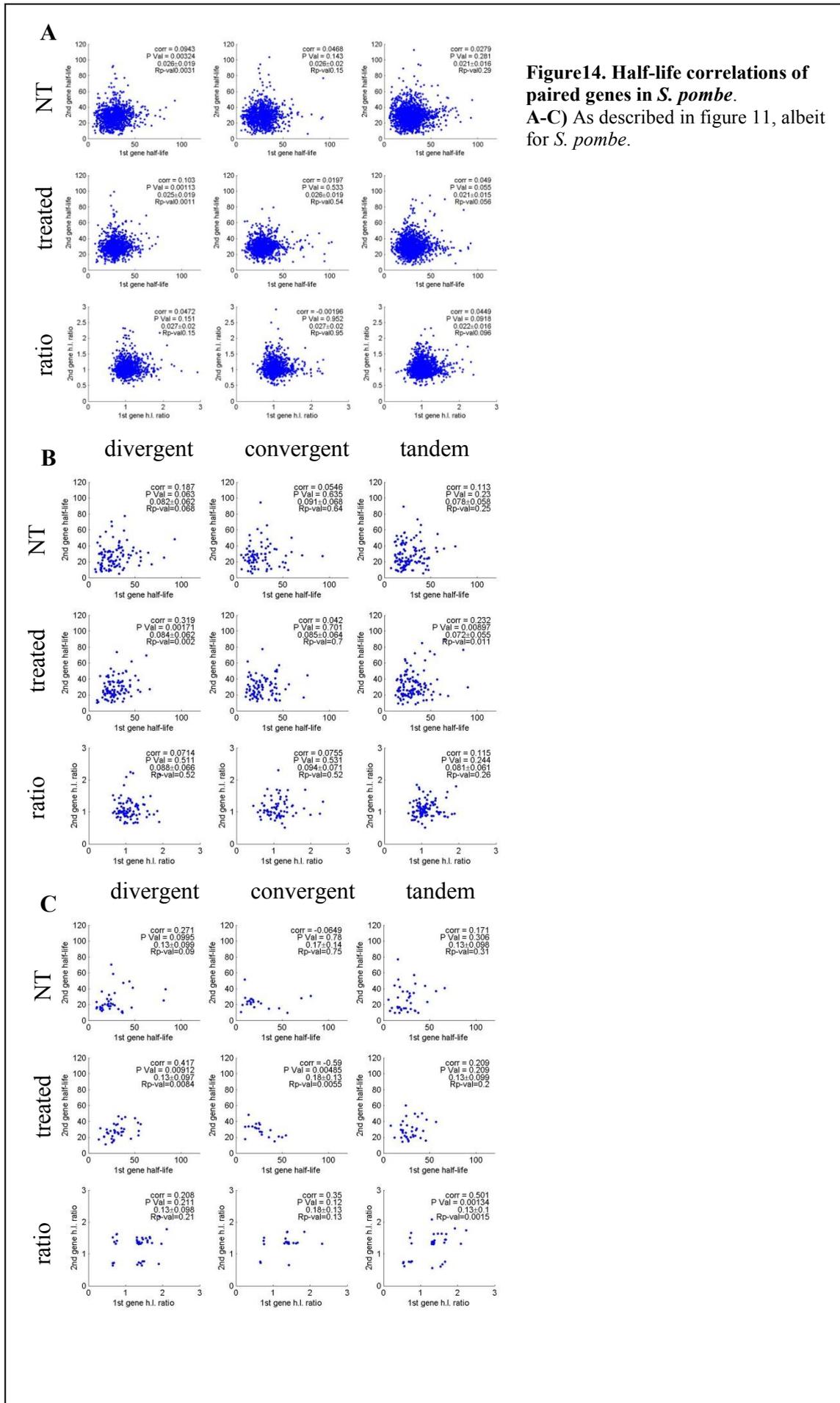
In order to check whether neighboring gene pairs that belong to each of the three architectures tend to co-decay more than control gene sets, we divided all the genes into subsets according to their architecture. We then plotted, for all the gene pairs in each architecture, the half-life of one gene against the half-life of its paired gene, and calculated their Pearson correlation. As control, we randomly paired genes in each architecture gene set by permutation of their order within a set, and calculated the average and standard deviation of the Pearson correlations of repeatedly permuted gene couples. We performed this analysis over three different conditions: non-treated, treated, and for the half-life ratio of each gene (treated to non-treated). We examined *S. cerevisiae* wild-type and the *rpb6* mutant that is defective in pausing transcription and mRNA decay (15) and in wild-type *S. pombe* (as described in materials and methods).

In the wild-type *S. cerevisiae* we had mRNA decay data obtained by two methods of transcriptional arrest, with a heat-sensitive *rpb1-1* mutant (16) and with the drug 1,10-Phenanthroline. Examination of decay profiles in the *rpb1-1* transcriptional arrest experiment revealed a modest yet significant correlation between pairs of divergent genes (Pearson correlation = 0.21, p-value = 5.4×10^{-8}). Such correlations were not observed in the two other architectures, nor were they observed in the averages of any of the randomly shuffled genes sets (Figure 11A).









Switching to the second method of drug-induced transcriptional arrest only partially reproduced these correlations. When looking for promoter architecture effects on all genes in *S. cerevisiae* wild-type, only a slight correlation (~ 0.15 , $p\text{-value} < 2 \times 10^{-7}$) was found between the half-lives of divergent gene couples, for both the non-treated half-life and half-life-ratio (Figure 12A), while gene couples from other promoter architectures did not show any significant correlation. The same analysis done on all *S. cerevisiae rpb6* mutant's genes, in which transcription was arrested with the drug, showed similar results to the wild-type (with same transcription arrest method) (Figure 13A), indicating that the mutation did not abolish the slight correlation seen in the wild-type.

Examining to *S. pombe* genes did not show any significant correlation in any coupled architecture when inspecting all the genes in the genome (Figure 14A).

Looking for a more subtle signal we decided to try portioning the genes into meaningful sections. As with the correlation between transcription and degradation, the correlation between adjacent couples might be dependent on their response to the stress – non responsive genes are less likely to show correlations. Therefore, we repeated the above analyses for subsets of expression- and decay-responsive genes. In order to compensate for the reduced number of genes in each architecture compared to the total number of genes, we lowered the decay-response threshold and set it to a 1.3 fold change hereafter. Indeed, For *S. cerevisiae rpb1-1* mutant increasing correlations of divergent couples' half-life ratios were observed in the responsive gene sets (Figure 11B, C). A smaller increase in correlations of responsive divergent couples was also seen in *S. cerevisiae* wild-type using 1,10-phenanthroline (Figure 12B, C), yet with marginal significance p-values (after Bonferroni correction for multiple hypotheses). It is possible that this signal might be masked out by the method of transcription-arrest. The same reason might also cause the lack of correlations in *S. cerevisiae rpb6* mutant responsive divergent genes. Interestingly, expression- and decay-responsive tandem genes in *S. cerevisiae rpb6* mutant show a medium-strong correlation of 0.41 – 0.9 respectively, between their half-life ratios (Figure 13B, C), yet with border-line significance.

Finally, similar trends were observed for responsive divergent and tandem couples in *S. pombe* (Figure 14B, C), with increasing correlations yet in changing levels of significance. Standing out are convergent genes in *S. pombe*, that showed a significant negative correlation of -0.59 (Figure 14C).

In summary of this section, examination of decay rates of neighboring gene pairs shows modest correlation in decay, that vary between strains, species and transcriptional arrest method. Selected gene sets may show some further modest increase in correlation, albeit with compromised statistics.

A potential effect of transcription rate and co-decay on evolutionary conservation of promoter architecture:

We next continued to an evolutionary analysis aimed at examining whether promoter sharing-mediated co-decay could have served as an evolutionary force that acted to keep a promoter architecture conserved in evolution. A common way to look at evolutionary conservation of a group of genes is microsynteny (29), an attribute of several genes that retain their genomic location relative to one another in different species. Synteny and microsynteny are sometimes used to address also the conservation of the gene order in the discussed group of genes, which others refer to as conserved-linkage (30) or co-linearity (31). Since we are looking at a special case of synteny, concerning the promoter-sharing genes that belong to each of the three architectures, we hereby term Conserved Promoter Architecture, or CPA, to indicate the extent of preservation of architecture between species. Since different species diverged in, and may or may not have conserved their promoter architecture for, different periods of time, a quantitative measure of conservation is required that will reflect the evolutionary time span in which an architecture was preserved for a pair of genes. Currently, we used the genomic data of four yeast species to calculate the CPA scores (see Materials and Methods for calculation), and while this analysis can be further expanded in the future by adding additional species, we believe it will enrich the results, but will not change them qualitatively, since the currently used species span across a broad phylogenetic range of yeast species (up to 500My by common estimates, see Figure 2, taken from (13)).

We calculated the CPA score for each gene couple in *S. cerevisiae* (using a published list of orthologous genes in yeast species (23)), and looked at the distribution histogram of those scores for each architecture (Figure 15A). Two observations can be made immediately: The first is that most of the *S. cerevisiae* gene couples do not conserve their promoter architecture even up to the closest species *Candida glabrata*, which is estimated at 78.4My apart from *cerevisiae*. Such genes are thus assigned here a CPA score of 0. On the other extreme only a handful of gene pairs in which both

pair members have orthologs, have conserved their promoter architecture between *S. cerevisiae* and *S. pombe*, with a CPA score of $\sim 500\text{My}$ or more.

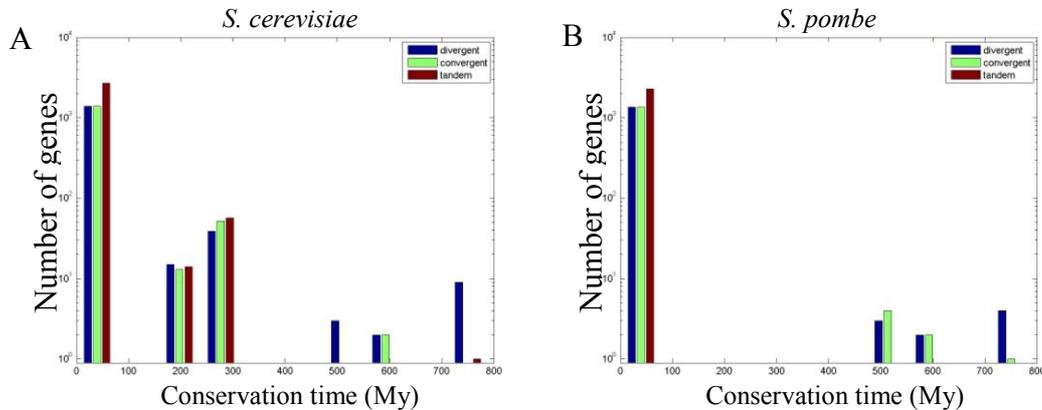


Figure 15. Divergent gene couples conserve their promoter architecture for longer periods of time than other architectures. This figure shows, for each architecture, A histogram of the number of gene pairs that conserved their architecture for the different evolutionary periods of time, in millions of years (My), for either a *S. cerevisiae* –centric (A) or *S. pombe* –centric (B) ortholog

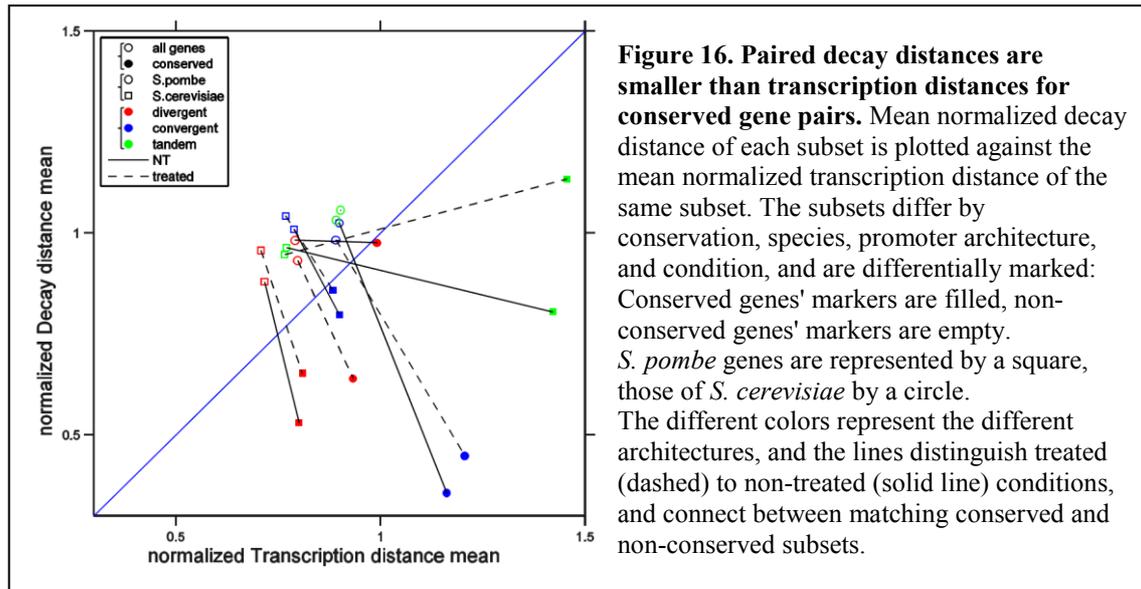
We collected a subset of 17 gene pairs in *S. cerevisiae* such that both pair members have an ortholog in *S. pombe* and that promoter architecture was preserved between the two remote species. We found that out of the 17 highly conserved gene couples, 14 are divergent couples (p-value $1.5 \cdot e^{-6}$, hyper geometric test for choosing 14 or more couples out of 17, given 1451 divergent couples out of a total of 5650 couples.) while only 2 are convergent and 1 tandem (hyper geometric p-value $2 \cdot e^{-4}$, for 1 couple or less). Thus, divergent couples conserve their promoter architecture a lot more than other architectures. This observation might imply an important role for the divergent promoter architecture in regulating the expression of the two genes in a coupled manner.

While the above analyses were *S. cerevisiae* –centric in the sense that all orthology assignments started from neighboring gene pairs in this species, we also computed a *S. pombe*-centric ortholog list, taken from Wapinski et al. (24). Naturally, even though most ortholog gene annotations are similar between the two lists used, minor differences between the genes common to both lists do exist.

Since *S. pombe* branches out separately in the yeast evolutionary tree (Figure 2), it shares the same last common ancestor with each of the other species, and therefore the evolutionary distance is equal between *S. pombe* and every one of the other species. That makes the minimal CPA score (other than zero) to be 487My for the *S. pombe*-centric analysis. Here too (Figure 15B), only a handful of gene couples

(sixteen couples) are highly conserved, and while divergent and convergent couples are almost equally present, 9 and 7 couples respectively (hyper geometric p-values 0.0138 for 9 or more and 0.1185 for 7 or more couples), tandem gene couples are noticeably absent (hyper geometric p-value $6.4 \cdot 10^{-5}$, for zero couples), taking into account that generally tandem gene couples are almost twice as prevalent as divergent or convergent gene couples throughout the genome.

We next asked what are the forces that preserve promoter architecture, and in particular whether gene pairs that show high decay similarity are more likely to retain their neighborhood compared to control genes. For that, we defined two scores for dissimilarity, or distance, of every gene couple. We defined decay distance for each couple as the absolute value of the difference between the half-lives of the couple, normalized to the standard deviation of the distribution of half-lives among all the gene pairs that belong to the same architecture. Transcription distance was defined similarly, using inferred transcription rates. When looking at the mean scores of all the genes of a specific architecture, the transcription distances are always smaller than decay distances, regardless of the condition, architecture, or even species. In other words, on average, adjacent gene couples co-transcribe similarly, more than they co-decay (Figure 16, empty markers). This might be logically expected from genes that share a chromosomal region. Yet the picture reverses for the subset of the gene pairs that have conserved architecture. When examining the subsets of genes with CPA scores larger than 200My, the average transcription distance is larger than the average decay distance. Moreover, conserved divergent couples in *S. cerevisiae* wild-type and conserved convergent couples in *S. pombe* show significantly smaller decay distances than their respective non-conserved sets (Figure 16, filled markers). Here too, the opposite is true for transcription distance scores, where the only significant result is that conserved tandem gene pairs in *S. cerevisiae* are significantly enriched for having higher mean of transcription distances compared to all non-conserved tandem genes. This indicates that high mRNA decay similarity between gene pairs may have served as an evolutionary pressure that preserved the architecture. Indirectly, this observation may serve as evidence that co-decay of divergent pairs is selected for, and raises the hypothesis that such co-decay may be due to promoter sharing.



Measuring the effects of different promoters on the decay of identical transcripts using synthetic gene constructs

Evolution may shuffle the order of genes in the genome. Such changes may be selected against, as shown above, if a transcript is removed from a natural regulatory region. To complement the picture we wished to create a synthetic library of genes in which such shuffling events will be created artificially and effects of expression would be measured. We thus attempted more straightforward, albeit low-throughput, approach in addition to the whole genome array-based and evolutionary analyses. In order to measure the direct effects of the promoter on mRNA degradation, several chimeric constructs were designed as follows: endogenous *S. cerevisiae* genes were selected according to their responsiveness to oxidative stress in a previous experiment (7). One group responded to oxidative stress and showed signs of coupling between transcription and degradation, while the second group did not react to the stress at all. The sequence of the selected endogenous genes was then divided into two parts. The first was their promoters, up to 1kb upstream of the transcription start site, and the second was their transcript, including the 5' un-translated region (UTR), the open reading frame (ORF), and the 3'UTR, all taken together. The two parts of each pair of selected genes were then re-assembled to create new couples of promoter and transcript. Together with the original endogenous sequences every two genes were represented in the four possible arrangements, hereby named a quartet. These quartets allow a clean comparison between the effects of the transcript and the promoter on mRNA decay, by measuring and comparing the decay rates of each construct. In order to exclude effects of genomic location, two fixed genomic loci were chosen as

integration sites of the chimeric constructs, the genomic loci of *CAN1* and *URA3* genes. Both these genes could be selected against, using canavanin and 5-FOA respectively, and could therefore be used as a negative selection marker for the successful integration of the chimeric construct.

This project was designed and started by other members in the lab, including Dr. Groisman and Dr. Shalem, and constructs were supplied by Ehud Shapiro's lab from the Weizmann Institute.

Lamentably, along the path from design to implementation, numerous technical difficulties were encountered, which ultimately prevented this project from coming into fruition.

The first issue regarded the selection of the *S. cerevisiae* master strain to harbor the constructs and upon which the experiments would be performed. Two different strains were selected, only to be found incompatible after some work had already been invested in them, until the BY4741 wild-type strain was finally selected. The second issue was *ura3* Δ 0 deletion in the selected strain BY4741 that had to be fixed. The exact deletion boundaries, not found in the literature, were sequenced and found to be at 218bp upstream to the transcription start site and 76bp downstream to the stop codon. The entire missing gene was picked up from a different strain and reinstated into BY4741. The third, most problematic issue was the integration of the chimeric constructs into *CAN1* and *URA3* loci. Unlike the positive selection with hygromycin successfully used for deletion of the endogenous transcripts of the chimeras, the negative selection against the presence of an active copy of *CAN1* or *URA3* was repeatedly fruitless. Initial scanning for positive integration products yielded no results. Transformations were repeated, each time changing one factor that might be hindering the reaction or scanning processes. Meanwhile, since the selection was negative, for the lack of an active product, any mutation in the ORF of the gene that leads to its inactivation would cause the harboring colony to circumvent the selection. Thus, negative selection greatly increased the number of background colonies that grew on the selective media without integrating the construct. Therefore, the number of colonies scanned for positive integration products had to be increased as well, to compensate for the increasing false positive background. To increase scanning efficiency, colony pooling was used (see chapter on optimizing colony scan). Despite several changes made to the transformation procedure, and more than 200 colonies

scanned, no colony was found positive for integrating any construct into the genome. As a result, it was decided to suspend further work on this project, pending a thorough redesign that will answer the problems and difficulties encountered.

The main problem that needs to be addressed in such a redesign is the lack of positive selection for the integration of the chimeric constructs. This issue should be addressed by integrating a selection marker at the 3' end of the chimeric transcripts, prior to its excision from the bacterial plasmid. This step was avoided in the previous design in order to keep the chimeric construct as short as possible, to facilitate the integration process, and out of expectations of the negative selection to work as well as the positive one.

The second major needed change in future redesign is the method of transcription arrest. Since each strain will harbor one chimeric construct to be measured, there is no need to stop transcription for the entire cell. A repressible promoter element, constant in all constructs, could be used instead of a general transcriptional arrest of the entire harboring cell. In addition, figuring out a way to make the construct building and measuring high-throughput processes would allow sampling of a bigger number of constructs and strengthen the results. In summation, a careful, *de-novo* design of the described project is needed and could overcome the difficulties that obstructed the project in its current format.

Discussion:

Recent studies supplied evidence that the promoter is a likely possible coordinator of both mRNA transcription and decay (11; 12), based on a few genes inspected. The most significant novel aspect of this thesis in view of these very recent publications is the demonstration that the effect of the promoter on decay applies at the genome-wide scale and that it is evolutionarily conserved. We also further explored the intriguing counter-action coupling between mRNA transcription and degradation, a phenomenon which might result from the effect of promoters on mRNA decay. While the phenomenon was originally observed in *S. cerevisiae* (7) we now demonstrate that it exists in the remote yeast species *S. pombe*. Furthermore, we explored the role of co-decay in shaping evolutionary dynamics of gene order and promoter sharing. Our results show a mild negative correlation between expression and degradation, which increases for subsets of genes that respond to the stress, be it by changing their expression or by changing their degradation. The correlation also increases for the subset of fast-peaking genes.

Several possible rationales for such coupling could be hypothesized. First, since half-life time is inversely correlated to the degradation coefficient, a coupled increase in both transcription and degradation allows the cell to reach the same final abundance levels with faster kinetics (2; 7). This faster response could explain the correlation strengthening for the fast-peaking genes. A different explanation for the lack of correlation in the late-peaking genes originates from the time gap between the stress and transcription arrest in our experiment. As can be seen from their expression profiles (Figure 3), the late-peaking genes did not yet start to express before transcription was arrested. Since the coupling between transcription and degradation is believed to be transcription-dependent it will not be detected for genes that were not yet transcribed. Conducting a similar experiment, while arresting transcription at a later point might reveal coupling between transcription and degradation also for late-peaking genes.

According to this hypothesis coupling is utilized by the cell as a regulatory feature that allows for a fast response e.g. to a stress, in a fashion that resembles the response ascribed to the negative-auto-regulation network motif (32). A second hypothesized role of coupling between transcription and degradation is to make the response to a stress endure only as long as the stress itself does. This could be achieved if mRNAs transcribed in response to the stress are co-transcriptionally labeled with a time stamp,

or an "expiry date" mark. Optimally, such a time stamp will be set to expire, i.e. accelerate the decay of its carrying mRNA, after a time equivalent to the expected duration of the stress. When the cell faces a transient stress, e.g. oxidative stress, it might mark the stress response transcripts with a rapid decay mark, whereas facing a more long-lasting stress, e.g. DNA damage stress, might label the nascent transcripts with a slower decay mark, or not labeling them at all, assuming a slow decay default. A possible candidate for such a cellular co-transcriptional decay-mark is the RNA Polymerase II subunit heterodimer Rpb4\7 (9; 10; 33; 14). The purpose of such a hypothetical mechanism is to facilitate the cell's rapid return to pre-stress homeostasis as soon as the stress is over.

Alternatively, instead of stress-dependent temporal response, in which certain mRNAs would be set to decay when the stress is over, the same mechanism could be utilized to determine cell-cycle-stage-dependent temporal regulation. In this scenario, certain mRNAs would be set to decay upon transitions between cell-cycle stages, as described in Trcek et al. (12).

It has been said that "Nothing in biology makes sense except in the light of evolution" (34). When examining coupling between mRNA transcription and degradation from an evolutionary point of view, another possible role emerges.

Our results show that steady-state mRNA abundance levels remained correlated even between species as distant as *S. pombe* and *S. cerevisiae* (Figure 8), while their decay parameters changed. Although the transcription parameters were also correlated between the two species (Figure 8), when looking at the inter-species change in transcription parameters (β) for orthologous genes, it is highly correlated to the inter-species change in decay parameters (α , Figure 10). Furthermore, the slope of the least square line fitted to the data indicates that the two parameters change on average almost in a 1:1 ratio. It appears that for many genes, α and β change together during evolution to maintain relatively constant steady-state expression levels. A coupling between transcription and degradation would facilitate such co-evolution and diminish the number of fatal or deleterious mutations that would change only one parameter and thus the gene's expression level. While we cannot exclude the possibility that the two parameters have been fine-tuned separately along evolution, this possibility would require a step-by-step incremental change of one parameter

after the other, resisting strong changes in one parameter only, and thus slowing down evolutionary change rate.

Dividing the genes into couples according to their promoter architecture reveals architecture-dependent differences in various parameters. One interesting parameter is the tendency to conserve the promoter architecture throughout evolution. Divergent couples are significantly enriched in the subset of highly conserved gene couples, while tandem couples are significantly depleted. Another parameter is the extent to which promoter architectures affect the co-decay of gene pairs. The mere fact that gene pairs from different promoter architectures differ in their co-decay, indicates the promoter's capability to regulate mRNA decay; the response-dependency of the extent of co-decay further strengthens this indication. The different extent of co-decay of evolutionarily conserved couples advocates for the surprising conclusion that co-decay might be a stronger cause of evolutionary conservation of promoter architectures than co-transcription.

Different promoter architectures stood out throughout the span of parameters, species and strains examined. This variety suggests that there is no specific promoter architecture that is "best" in all parameters, but rather that the different promoter architectures are by themselves a parameter, and could be used differentially by the cell as another tool in the cellular evolutionary toolbox.

When examining the *S. cerevisiae* data acquired using the two methods of transcription arrest we encountered some inconsistencies. Specifically, some of the results found using the *rpb1-1* heat sensitive mutant could not be fully recapitulated when using the drug 1,10-Phenanthroline, while other results could not be recapitulated at all. On one hand, this discrepancy sheds light on the importance of using the "right" technique for scientific measurements that would truly expose the investigated phenomenon without inserting false data or artifacts. On the other hand it exposes our ignorance as to what exactly happens inside the cell when we perturb its environment, e.g. for the purpose of transcription arrest. We here suggest a set of experiments, using novel techniques of decay-rate quantification such as 4sU labeling (27), to measure the exact effects of the two transcription arrest methods.

Transcription arrest during 4sU labeling should reveal the extent of transcription inhibition of each method, and will also allow measurement of the possible side-effects of each method, e.g. on the decay rates themselves. Should transcription not be entirely arrested, interesting questions arise: Do the remaining active RNA

polymerase molecules distribute equally between all genes? Or do some genes get the better share of the polymerase pool that is still active, presumably according to some importance hierarchy between the genes?

The choice of transcription arrest method is only one of the many choices made throughout this study. When designing the experiment, and perhaps even more so when analyzing the data, there exist many other degrees of freedom. A partial list includes examining various conditions and parameters, ranging from focusing on specific subsets under different thresholds, through correlation calculation method, and to plotting in different axes scales. On one hand, there must be a limit to the extent of tweaking and manipulating of the data, in order for the results and conclusions to be true to its biological meaning. On the other hand, the scientific approach is to check the whole range of possible parameters, and choose the best condition out of them, e.g. as done when calibrating an experiment. We tried to balance these factors by keeping high moral and scientific standards: making adequate controls, correcting for multiple hypothesis, staying objective about the results, adapting the hypothesis to the results and not vice-versa, in a constant attempt to reveal the true biological meaning of the data and unveil some of the secrets of life in our research.

Literature:

1. *Genomic expression programs in the response of yeast cells to environmental changes.* **Gasch, AP et al.** s.l. : Mol Biol Cell, 2000.
2. *Genomics and gene transcription kinetics in yeast.* **Perez-Ortin, JE, Alepuz, PM and Moreno, J.** s.l. : Trends in Genetics, 2007.
3. *A Global View of Gene Activity and Alternative Splicing by Deep Sequencing of the Human Transcriptome.* **Sultan, Marc et al.** s.l. : Science, 2008.
4. *The Transcriptional Landscape of the Yeast Genome Defined by RNA Sequencing.* **Nagalakshmi, Ugrappa et al.** s.l. : Science, 2008.
5. *The highways and byways of mRNA decay.* **Garneau, NL, Wilusz, J and Wilusz, CJ.** s.l. : Nat Rev Mol Cell Biol, 2007, Vol. 8.
6. *The Many Pathways of RNA Degradation.* **Houseley, J and Tollervey, D.** s.l. : Cell, 2009, Vol. 136.
7. *Transient transcriptional responses to stress are generated by opposing effects of mRNA production and degradation.* **Shalem, O et al.** s.l. : Mol. Syst. Biol., 2008.
8. *Major role for mRNA stability in shaping the kinetics of gene induction.* **Agami, Elkond and.** s.l. : BMC genomics, 2010.
9. *The RNA polymerase II subunit Rpb4p mediates decay of a specific class of mRNAs.* **Lotan, R et al.** s.l. : Genes Dev., 2005, Vol. 19.
10. *The Rpb7p subunit of yeast RNA polymerase II plays roles in the two major cytoplasmic mRNA decay mechanisms.* **Lotan, R et al.** s.l. : J Cell Biol., 2007, Vol. 178.
11. *Promoter Elements Regulate Cytoplasmic mRNA Decay.* **Bregman, Almog et al.** s.l. : Cell, 2011.
12. *Single-Molecule mRNA Decay Measurements Reveal Promoter- Regulated mRNA Stability in Yeast.* **Trcek, Tatjana et al.** s.l. : Cell, 2011.
13. *Conservation of divergent transcription in fungi.* **Kensche, Philip R. et al.** s.l. : Trends in Genetics, 2008.
14. *Transcriptome Kinetics Is Governed by a Genome-Wide Coupling of mRNA Production and Degradation - A Role for RNA Pol II.* **Shalem, O et al.** s.l. : Plos Genetics, 2011.
15. *Loss of the Rpb4/Rpb7 subcomplex in a mutant form of the Rpb6 subunit shared by RNA polymerases I, II, and III.* **Tan Q, Prysak MH, Woychik NA.** s.l. : Mol. Cell Biol., 2003.
16. *Eucaryotic RNA polymerase conditional mutant that rapidly ceases mRNA synthesis.* **M Nonet, C Scafe, J Sexton, and R Young.** s.l. : Mol. Cell Biol., 1987.
17. *A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes.* **Janke, Carsten et al.** s.l. : Yeast, 2004, Vol. 21.
18. [Online] <http://www-bcf.usc.edu/~forsburg/media.html>.
19. *5-Fluoroorotic acid as a selective agent in yeast molecular genetics.* **Boeke JD, Trueheart J, Natsoulis G, Fink GR.** s.l. : Methods Enzymol., 1987.
20. *Multiple Pathways Differentially Regulate Global Oxidative Stress Responses in Fission Yeast.* **Chen, Dongrong et al.** s.l. : MBC, 2008.
21. *Global coordination of transcriptional control and mRNA decay during cellular differentiation.* **Amorim, Maria J et al.** s.l. : MSB, 2010.
22. *Coupled Evolution of Transcription and mRNA Degradation.* **Dori-Bachash, Mally, Shema, Efrat and Tirosh, Itay.** s.l. : Plos Biology, 2011.
23. *Differential translation efficiency of orthologous genes is involved in phenotypic divergence of yeast species.* **Man, O and Pilpel, Y.** s.l. : Nature Genetics, 2007.

24. *Natural history and evolutionary principles of gene duplication in fungi.* **Wapinski, I et al.** s.l. : Nature, 2007.
25. **Sambrook.** *Molecular cloning: A laboratory manual.* 1989.
26. **Sherman, F.** Getting started with yeast. [book auth.] G R Fink and C Guthrie. *Guide to yeast genetics and molecular biology.* Cold-Spring Harbor, NY : Cold-Spring Harbor lab press, 1991.
27. *Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast.* **Miller, Christian and Cramer, Patrick.** s.l. : Molecular System Biology, 2011, Vol. 7:458.
28. *A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression.* **Cohen, BA et al.** s.l. : Nature Genetics, 2000.
29. *Genomic regulatory blocks underlie extensive microsynteny conservation in insects.* **Engström, PG et al.** s.l. : Genome Res, 2007.
30. [Online]
http://scienceblogs.com/evolgen/2008/06/synteny_a_semantic_debate.php.
31. [Online] <http://www.generationcp.org/genomics/index.php?page=1146>.
32. *Negative autoregulation speeds the response times of transcription networks.* **Rosenfeld N, Elowitz MB, Alon U.** s.l. : J Mol Biol., 2002.
33. *Rpb4 and Rpb7: subunits of RNA polymerase II and beyond.* **Choder, M.** s.l. : Trends Biochem Sci., 2004.
34. *Nothing in Biology Makes Sense Except in the Light of Evolution.* **Dobzhansky, Theodosius.** s.l. : American Biology Teacher, 1973, Vol. 35.
35. *Cerebral Autoregulation Dynamics in Humans.* **R Aaslid, KF Lindegaard, W Sorteberg and H Nornes.** s.l. : Stroke, 1989.
36. *Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast - - MSB 2011.* **Miller, Christian and Cramer, Patrick.** s.l. : Molecular System Biology, 2011, Vol. 7:458.