Regulation of mRNA decay kinetics in response to environmental stimuli

Ophir Shalem

M.Sc. Thesis

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

Research conducted under the supervision of
Dr. Yitzhak Pilpel and **Dr. Eran Segal**

February 2007

Table of content

1. Abstract

The mRNA level of a gene reflects a balance between two opposing and highly regulated processes: transcription and transcript degradation. In contrast to the attention devoted towards the study of transcription, transcript degradation has been less studied and hence very little is known about the mechanisms controlling mRNA stability. Although decay rates for all genes, in several organisms, have been previously determined, those measurements were done only in permissive growth conditions providing no information on the extent to which mRNA decay is modulated in a response to various stimuli and stresses.

We have set to study the dynamical properties of mRNA decay by measuring the decay profiles of all genes in the genome of the yeast *S. cerevisiae* at different growth conditions following transcription inhibition. Our measured decay rates in permissive growth conditions are in accord with previous measurements done in such conditions both with respect to the actual rates and also in the observation that they are closely related for groups of genes defined by shared functional annotations.

We show that in response to stress major changes in mRNA degradation occur, and that these changes seem to be coordinated in functional groups of genes that display coherent decay profile within a condition but different across different conditions. Changes in decay kinetics across growth conditions are not only reflected in a constant rate change but also in qualitative difference in the decay kinetics between the conditions. To mathematically capture such changes we develop a kinetic model, based on first principle of interaction between RNA binding proteins and their target mRNAs, which is used in order to identify gene sets that undergo a coordinated change in decay modes in different conditions.

We then compared the change in mRNA abundance in response to the stress, before transcription is halted to changes in half life with respect to non-stress measurements. Interestingly, we see that increase in mRNA abundance, right after the stress is applied, is correlated with a reduced half life in the stresses measurements compared to control measurements. We propose that underlying this result is a putative molecular mechanism by which transcription is directly coupled to degradation. Thus, enhanced transcription would be immediately complemented by enhanced degradation, giving rise to an increased response time.

2. Acknowledgments

We thank Richard Young's lab for providing Yeast strain Y262, Dr. Shirley Horn-Saban from the microarray unit in Biological Services Department of the Weizmann Institute for performing the RNA sample processing and hybridization to affymetrix microarrays and Dr. Ron Ophir from the microarray unit for advice in the preprocessing steps of the microarray probe data.

I would like to personally thank my advisors, Dr. Tzachi Pilpel, Dr. Eran Segal and all members of Pilpel and Segal labs. Especially I thank Dr. Orna Dahan, for guidance in the wet lab and Zohar Bloom for help in the experiments.

3. Introduction

Microarray measurements of mRNA abundance following environmental perturbations show extensive modulation of transcript levels in response to external stimuli. Some of these changes are condition specific and some are similar across the different conditions[1,2]. Such high throughput mRNA abundance measurements are widely used to study transcription, generally by correlating sequence attributes in gene regulatory regions with changes in transcript abundance or by trying to learn more complex interactions using the correlations betweens genes to regulators [3-6]. Although such works give a global view of the combinatorial and dynamic properties of transcription control, they associate all changes in transcript levels to changes in transcription initiation, ignoring the fact that mRNA abundance reflects a balance between transcription initiation and transcript degradation.

In contrast to the attention devoted towards the study of transcription initiation, transcript degradation has been less studied and hence very little is known about the mechanisms controlling mRNA stability, the extent to which this process is modulated as a response to various stimuli and the relative contribution of transcript stability in the determination of mRNA abundance levels compared to transcription initiation.

3.1. Background on mRNA degradation

Mature eukaryotic mRNAs are created and exported into the cytoplasm with two integral stability determinants: the 5' cap structure and the poly(A) tail. Several pathways exist by which these structures are removed and the mRNA is degraded: deadenylation-dependent mRNA decay, deadenylation-independent decapping, endonuclease-mediated mRNA decay and mRNA surveillance pathways (nonsense, non-stop and no-go mediated decay) which were recently shown to take part in the degradation of non-faulty transcripts[7]. miRNA-mediated mRNA decay also exists, although there is no indication that these mechanisms exist in the yeast *S. cerevisiae*, the model organism used in this work.

In eukaryotes and especially in yeast, most mRNAs decay by the deadenylation-dependent mRNA decay pathway, and only a few mRNAs were shown to bypass this

route and decay by other pathways. The rates of deadenylation are highly correlated with the decay rates of total mRNA and this step is thought to be the rate limiting step of the whole decay process[8].

Deadenylation-dependent mRNA decay

The first step in deadenylation-dependent decay is the shortening of the poly(A) tail. Poly(A) shortening (deadenylation) can be catalyzed by several enzymes (CCR4-NOT is the main deadenylase in *S.cerevisiae*). This process was even shown to have varied kinetics ,defined by whether shortening requires one or multiple binding events of the deadenylase, which might affect the rates of degradation[9]. Once the poly(A) tail is shortened below a certain length, decay is proceeded via two possible pathways: the 5' → 3' decay begins by the removal of the 5' cap followed by 5' → 3' decay by an exonuclease. 3' → 5' decay is an alternative pathway in which the mRNA is degraded by a large complex of exonucleases known as the exosome. Figure 1 shows the main components of the whole deadenylation-dependent pathway[10]. Much of the enzymes involved in these two pathways were identified[11]. It is not clear why two distinct decay pathways were maintained in evolution and which of these two pathways is more active, although 5' → 3' decay seems to predominant in yeast [10].

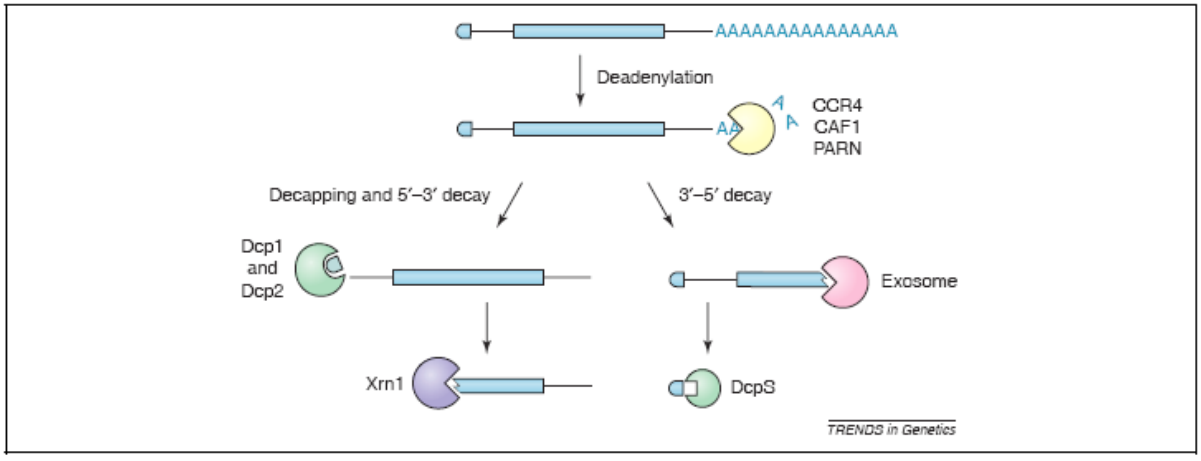

Figure 1: Outline of the deadenylation dependent mRNA decay pathway with the central identified protein factors

Recent studies indicate that mRNA decay is carried out in distinct cytoplasmatic compartments (P-bodies)[12]: These sites have been characterized in both human and yeast

and the enzymes known to be associated with the process of mRNA decay were shown to localize to these compartments together with mRNA decay intermediate products. Interestingly, it seems that at least in yeast, 3' → 5' and 5' → 3' decay pathways occur in distinct cytoplasmic sites[12]. P-bodies were shown to assemble when the amount of mRNA needed to be degraded is increased, this happens when the 5' → 3' decay pathway is overloaded with RNA substrates or when mRNA decay is impaired. Although the fact that the decay machinery together with mRNA decay intermediates are enriched in these sites suggests that P-bodies serve as sites for mRNA decay, P-bodies might have other functions rather than merely that. Evidence exists showing that P-bodies accumulate when mRNA decay is inhibited (following glucose deprivation). This observation, together with the possibility that P-bodies might only be formed when the 5' → 3' mRNA-decay pathway is disrupted or overloaded may suggest that they might be places to store mRNAs that are targeted for destruction but cannot be immediately degraded, in order to prevent them from reassociating with the polysomes and creating aberrant proteins. In addition, in some cases mRNAs were shown to be readenylated and return to the polysomes, suggesting a wider role for P-bodies as sites of regulation and sorting of mRNAs in the interplay between translation and degradation[7].

Regulation of deadenylation

Once the mRNA is exported from the nucleus to the cytoplasm the poly(A) tail and 5' cap structure are bound by the cytoplasmic proteins poly(A)-binding protein (PABP) and eIF4E respectively. During translation the mRNA is thought to be circularized by the interaction of these proteins and other translation and stability protein factors. This interaction enhances translation and prevents degradation by protecting the poly(A) tail from deadenylation[13].

Considering the above, a model is proposed for the effect of RNA binding proteins on mRNA stability: Stabilizing protein factors might act by enhancing the affinity of the PABP complex to the mRNA or by competing with binding sites for de-stabilizing elements, elements which might act by dissociating the PABP-mRNA complex either by competitive binding or by recruitment of deadenylases[7,13]. In budding yeast, it is estimated that ~570 different proteins have the ability to bind RNA, this number is considerably larger in higher eukaryotes[14]. Although other processes in the cell require

RNA binding activity (e.g. translation control, export etc.) there is accumulating evidence for cases of RNA binding proteins that have an effect on mRNA stability[10].

Various sequence elements, controlling decay rates, are known. These sequences are usually located in the 3'UTR region of mRNAs but have also been detected within ORFs and also at 5'UTRs[10]. The most studied sequence element that has been experimentally verified to control the stability of mRNAs is the AU-rich element (ARE) which is found in several eukaryotic systems including yeast and human[15]. AREs can have multiple roles ranging from stabilization to enhancing decay and differences between variants of this motif are proposed to account for the different roles[16]. In addition, sequences that flank this motif were also shown to influence the overall mRNA stability[7]. Many ARE-binding proteins have been identified (e.g. AU-rich binding factor (AUF1), KH splicing regulatory protein (KSRP) and more) by recent studies that showed that indeed the key to their de-stabilizing function is their ability to recruit elements of the mRNA-decay machinery (e.g. AUF1 can interact with the exosome). In some cases the AREs by themselves can perform this function (e.g. the exosome shows affinity for AREs)[7]. In parallel, evidence is accumulating in the literature, showing that diverse biological processes are regulated at the level of mRNA stability. For example, a recently published article shows that the zinc-finger antiviral protein (ZAP) recruits the exosome to degrade it's target mRNAs, suggesting that ZAP is a trans-acting factor that modulates mRNA stability[17].

When considering the regulation of mRNA stability by RNA binding proteins it is important to keep in mind that RNA recognition might require more than a linear sequence motif. RNA binding proteins where shown to recognize both sequence motifs and RNA secondary structures and in some cases a combination of both is required[18,19]. For example, members of the zinc-finger transcription factors were shown to also bind RNA, crystal structure reveals that this interaction requires recognition of both secondary structure and specific RNA single stranded sequence[20].

3.2. Review of previous works

Degradation profiles of all genes were measured in permissive conditions in several organisms[8,21,22]. The basic experimental procedure is to inhibit transcription and take

samples of decaying mRNA over a period of time following the transcriptional arrest. These samples are then hybridized to microarrays or cDNA arrays in order to get a degradation profile for each gene. Wang et al.[8], which preformed this experiment in *S.cerevisiae* showed that half lives are specified for individual genes ranging from ~3 min to more than 90 min. In addition they show that genes encoding for proteins that act together in stoichiometric complexes, or share a physiological function, tend to decay in similar rates proposing that the decay rates are used to tightly regulate the expression levels of these genes. Similar results were obtained by Yang et al.[21] measuring decay rates of human genes. In addition, these authors also found sequence motifs correlative with low and high half lives, one of which is the known AU-rich motif. Among several general observations, it was shown, both in human and in yeast, that transcription factors show fast decay rates while housekeeping genes, such as biosynthetic proteins, show slower decay rates compared to the genome average.

Extensive searches for regulatory sequence motifs were applied on 3'UTRs. In a previous work that was performed in our lab by Shalgi et al.[23] a catalog of 53 motifs corresponding to low or high half lives was derived using the decay data produced by Wang et al.[8] . A significant proportion of these 3'UTR motifs were found to be conserved. Also some of them correspond to known sites for RNA binding proteins. Xie et al.[24] preformed a comparative analysis of several mammalian genomes to find conserved sequence motifs both in promoters and 3'UTRs. They propose about 106 3'UTR motifs which are significantly conserved, half of which are claimed to be targets of miRNAs. The same was previously done in closely related yeast species by Kellis et al.[25], who in addition to promoter sequence motifs, identified 6 conserved motifs in 3' UTRs. Barrett et al.[26] also looked for sequence motifs in 3' UTRs; they used a multivariate model to explain mRNA abundance measurements using as explanatory variables occurrences of regulatory motifs in 3' UTRs. Six binding sites for mRNA stability regulators were discovered and characterized, two of them are claimed to belong to members of the Pumilio-homology domain (Puf) family of RNA binding proteins. For one of these motifs experimental validation was also provided.

Another way to study association between RNA binding proteins to mRNAs, which is equivalent to Chip-chip experiments used to study association between transcription factors to promoters[27], was used by Gerber et al.[28]. In this work the authors isolated, using affinity tags, each one of the five Puf proteins and used microarrays in order to

identify the targets of each one of the Puf proteins. Their results show that each Puf protein has a distinct group of targets of about 40-220 different mRNAs with common functions and sub-cellular localizations. The fact that two of the five Puf proteins were shown to affect stability of mRNAs suggests that co-regulation occurs also at the level of mRNA stability.

All the evidence described above indicates that mRNA decay is a property which varies between genes, it can be controlled by RNA binding proteins, it relates to the function of the gene and also has a role in determining the mRNA abundance levels. Still, it is not known to what extent mRNA degradation is a dynamic property which is modulated as a response to environmental stimuli. In other words, it is still not known what part of the mRNA abundance changes that are observed when cells encounter environmental perturbations can be explained by changes in transcription and what part of this response is due to changes in mRNA stability. Several studies have investigated condition specific modulation of mRNA decay on specific genes: For example, treatment by Rapamycin was shown to exert a destabilizing effect on multiple mRNAs[29]. In addition Hilgers et al.[30] showed that hyperosmolarity, heat shock, and glucose deprivation stabilize multiple mRNAs in yeast. Other works investigating the effect of different stimuli on mRNA stability are reviewed by Garneau et al. and others[7,10,13]. Yet a genome-wide perspective on changes in mRNA stability in response to various environmental challenges is still missing. A few studies did attempt to infer mRNA stability differences across different growth conditions, yet indirectly - using measurements of mRNA abundance and rates of transcription. A noticeable work is that of Garcia-Martinez et al.[31] who used a genomic run-on method to infer condition specific modulation of mRNA decay. These measurements were based on discrepancies between transcription rates and mRNA abundance measurements. Same method was used by others[32,33] on different organisms. All of these works show that only 40-50% of the changes observed in mRNA abundance can be explained by changes in transcription rates, implying a major effect of changes in mRNA stability. These works indeed give evidence to condition specific regulation of mRNA decay but they do not show any common attributes of genes showing similar effect neither on the level of functional annotation nor in the gene's sequences. This is probably due to the fact that indirect inference of mRNA decay kinetics gives only a vague idea to what extent of the genes mRNA abundance can be

explained by changes in transcription levels and there are no direct measurements of degradation kinetics and how they are modulated.

3.3. Motivation and work outline

The above reviewed works give the first indications that mRNA degradation is coordinated between functional groups and that regulation of mRNA degradation has a significant contribution to the observed changes in mRNA abundance in response to stress. Still, without condition-specific, direct, genome wide measurements of mRNA decay it is hard to study the principles of the regulation of decay kinetics (e.g. whether groups of genes are coordinately stabilized or destabilized under different environmental conditions and whether common regulators and sequence attributes can be found).

We have set to take the first steps in this direction by measuring directly mRNA degradation kinetics in response to different environmental conditions of all *S. Cerevisiae* genes. We use this data to show global modulation of mRNA stability: while some genes show a stabilization effect other show destabilization effect and this response is condition specific. Further, while it is customary to treat degradation profiles as exponential decay with a simple half life characteristic, we observed hundreds of genes that decay non-exponentially. We developed a simple kinetic model that is based on first principles that explains such kinetics and that is used in order to further investigate the data.

4. Methods

4.1. Experimental procedure

In order to measure degradation, transcription has to be inhibited. For this purpose we use a mutant strain of *S.cerevisiae* bearing a temperature sensitive mutation in the RNA polymerase II gene which is inactive in the non-permissive temperature of 37°. Thus, a heat shock is applied in order to stop transcription. We decided to wait a short period of time after applying the stress, and before inhibiting transcription. In this way the cells are given time to induce a transcriptional response to the stress and we hope that not all of the response will have been missed. In addition measuring mRNA levels before and after the stress (at the point of transcription inhibition) allows us to identify the responsive genes and study the relationship between the observed change in mRNA

levels to the later change in mRNA decay compared to the control time course. The length of the time gap between applying the additional stress to transcription inhibition was decided according to the following considerations: We wanted to stop transcription at the point when responsive genes reach their maximum response to the stress but not after. In order to estimate this time we used the average response of all responsive genes using published data[1] and also followed the mRNA level of known responsive genes in our strain using real time PCR as will be described below.

The general outline of each experiment is described in Figure 2: We apply an environmental stress and 25 minutes after it, transcription is halted by shifting to the mutant's non permissive temperature. Consequently, samples are taken for hybridization with the microarrays at the marked points. Two samples are taken at time point 0; this point is important both because the whole time course is zero transformed relative to that point and also it used in order to investigate changes in mRNA abundance in response to the stress.

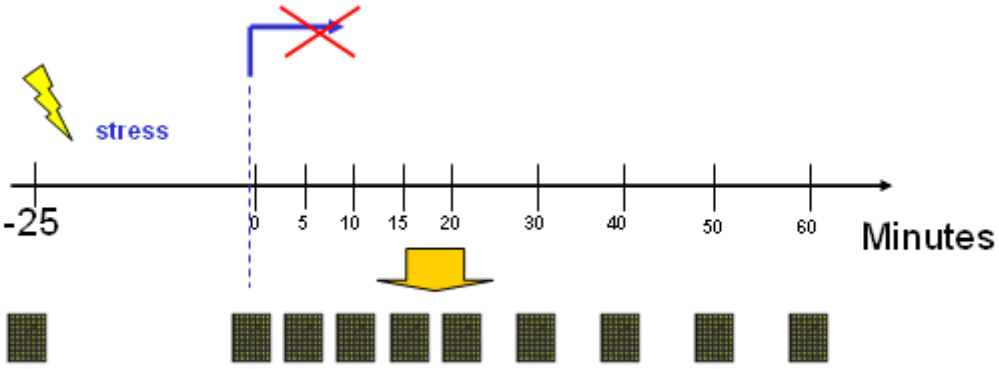

Figure 2: A sketch of the experimental procedure used to measure mRNA degradation in response to environmental perturbations. At time -25 a sample is taken and the stress is applied, at time 0 transcription is halted and samples are taken at time points (0, 5, 10, 15, 20, 30, 40, 50 and 60).

4.2. Strains and growth conditions

S.cerevisiae strain Y262 (MATa ura3–52 his4–939am rpb1-1)[34], carrying a temperature-sensitive mutation in RNA polymerase II was used in this study. Three separated time course experiments were preformed in the following procedure: Cells were grown in extract / peptone / dextrose (YPD) to the concentration on $2*10^7$ cells/ml (mid logarithmic phase) in a temperature of 26°. We carried out one "non-treated"

experiment in which we halted transcription without any additional stress, apart from the transcription-halting heat shock, and two "treated" experiments in which we exposed cells to one additional stress in each experiment. In the treated time courses stresses were applied (details are provided in table 1), and after 25 minutes (at a time point referred to as "zero") temperature was abruptly raised by adding an equal amount of medium that was warmed in advance to the temperature of 49°. Aliquots of the culture (15ml) were removed in the following time points: -25, 0, 0, 5, 10, 15, 20, 30, 40, 50, and 60. For each sample the medium was immediately removed and cells were frozen in liquid nitrogen. RNA was extracted using MASTERPURE Yeats RNA purification kit provided by EPICENTER. The quality of the RNA was assessed using bio-analyzer; the samples were then processed and hybridized to Affymetrix yeast 2.0 microarrays using the Affymetrix GeneChip system at the Biological Services unit at the Weizmann Institute.

	Concentration of stock	**Concentration used for treated sample**
H_2O_2	0.833M	0.3mM
Rapamycin	40 μg/ml	200 ng/ml

Table 1: treatments, amounts and concentrations

4.3. Real time PCR on specific genes

The mRNA abundance level of specific genes in calibration experiments and the main experiments was measured using real time PCR. RNA was extracted using MASTERPURE Yeats 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. Each sample was measured in duplicate wells. The standard curve was calculated using four dilutions of 1:5 starting from a mix of all cDNA samples for each primer. Samples were then diluted at 1:5 ratio so that the measured levels will not reside at the edge of the standard curve. This protocol was found optimal with respect to the variance between duplicate wells which was an important step in order to get a consistent measurement over a time course.

The mRNA level of a gene of interest over a time course was plotted using the following procedure: mRNA levels were calculated using a gene specific standard curve

for the gene of interest and a reference gene. The duplicated values for the reference gene were averaged, the gene of interest's values were normalized by the reference genes averaged values at the corresponding time points. The normalized values for the first time point were averaged and then the whole time course was divided by this value (zero transform). Then the mean and standard deviations of the normalized and zero transformed values were plotted. In order to calibrate the stresses, a known non responsive (housekeeping) gene was taken as the reference gene (ACT1). In order to validate transcription inhibition a fast decaying gene was taken as the gene of interest while a known slow decaying gene was taken as the reference (e.g. PGK1 and RPS6B as low and fast decaying genes respectively).

4.4. Data preprocessing

The first step in any microarray experiment, after hybridization and scanning, is data preprocessing in order to get from probe level data to expression values for each gene. Each gene is represented with a probe set containing several probes (11 in Affymetrix microarrays) which are complementary sequences to different regions of the target gene. The purpose of the preprocessing stage is to sum this data into one expression value per gene and to correct for technological deviations between samples that are a result of errors in the different stages of the experimental procedure: RNA extraction, synthesis of cDNA, labeling, *in vitro* transcription, hybridization and scanning. A standard Affymetrix preprocessing procedure is composed of the following steps:

- Reading in the probe data from the scanning results
- Sample specific background correction
- Normalization
- Probe specific background correction (e.g. subtracting MM values from PM values)
- Summing up probe data to expression values

The preprocessing steps are usually preformed by standard algorithms such as RMA or the MAS5 algorithm which is provided by Affymetrix.

Most preprocessing algorithms use a normalization step in order to bring all samples to have the same global distribution of intensity values. This is done under the assumption that global deviations, e.g., in the mean intensity between samples, represent artifacts that result from a difference in the processes that the samples undergo after RNA extraction till hybridization and scanning. The biological assumption is that the global distribution of mRNA in the cell does not change much and that the biological signal will be captured by changes in the relative rankings of genes within this distribution. A practical corollary of this assumption is that in regular procedures the mean intensity of all genes in each sample is subtracted from all the genes in that sample. This obviously results in samples that have the same means.

A unique aspect of our experiment is that the above assumption, that there is no global change in the distribution of mRNAs, is by definition not valid in our experimental setup: Due to transcription inhibition we do expect a global decrease in the total amount of mRNA as we advance in the time course. This critical point requires special considerations when approaching the normalization step; we used an approach similar to what was used by Wang et al.[8] for this purpose as will be described below.

We have used the following steps in order to preprocess our data; each step is described below. All the preprocessing steps were preformed using the "affy" package which is part of the bioconductor (http://www.bioconductor.org/) open source project implemented in the R programming language. Unique steps to our setup were implemented in R as well.

- Background correction using the RMA background correction algorithm.
- Summing up probe data to probe set data using the median polish algorithm.
- Spike normalization.
- Refinement
- Averaging and zero transform

(probe specific background correction using the mismatch probes is not used due to previous experience that showed that this correction is not very useful)

Background correction using the RMA background correction algorithm:

The goal of the background correction step, also termed signal adjustment, is to correct the effect that non biological factors in the experiment may have in the measures drawn from each array. Specifically it should correct background noise and processing effects, adjust for cross hybridization and calibrate the scales of each array, so that all arrays would have the same base level and the most linear relationship possible between intensity measurements to RNA concentration.

Not all algorithms available, which are referred to as background correction methods, deal with all of these points. Some only deal with correcting background noise and processing effects.

We chose to use the background correction method proposed by Bolstad et al.[35] and is implemented as part of the RMA algorithm. The motivation to this algorithm comes from observations of probe intensity distributions. The observed signal is modeled as a sum $S = X + Y$ where X is the signal and Y is the background. X is assumed to be distributed according to an exponential distribution $X \sim \exp(\alpha)$ and Y according to a normal distribution $Y \sim N(\mu,\sigma^2)$. Then the corrected signal is given by the expectation $E(X|S=s)$ after the parameters α, μ and σ are estimated from the observed intensities.

Summing up probe data to probe set data using the median polish algorithm:

The summation step is the process by which the different probe values for the same gene are summed up into one expression value. Summation methods are roughly divided into two groups: single and multi chip methods. Single chip methods use only the probe information of the individual array in order to compute the expression value while multi chip methods use the information in multiple arrays.

The motivation behind multi chip methods comes from the observation that intensity variation within a probe set is commonly larger than the variation of the same probe set between different arrays, yet this variation is correlated between the arrays. "Median polish" is an example of such an algorithm which is implemented as part of the RMA package[35,36]. In multi chip methods the probe set data across several arrays is fitted into a model that decomposes the observed intensity into several variables representing the probe effect, the sample effect and a random error (general or separated to the probe and the sample effect) and used the regressed value for the sample effect as the corrected

intensity value. Median polish does this by an iterative procedure that subtracts the median of rows and columns alternately, keeping the median value at each iteration separately for each row and column. The median of medians for each column is taken as the column (sample) effect.

Normalization using spike intensities:

The difference between the normalization step to the background correction step is that while the former is done independently for each array the normalization step is done using a group of array measurements in order to correct for non biological variations between them. As stated before, normalization methods assume that either only a small number of genes changes between samples or that genes with elevated expression levels are compensated by genes with reduced expression levels leaving the total distribution similar. This assumption which is clearly not true in our experimental settings required a different normalization method based on reference "spiked in" RNA which is known to be constant for all samples.

In order to normalize microarray samples with respect to each other, an internal standard was mixed with each RNA sample. The internal standard contains a pool of 4 "spikes" - *in vitro* transcribed *B.subtilis* genes, each in a different concentration (poly(A) control kit supplied by Affymetrix), each represented on the microarrays by several probe sets. Because the same amount is mixed with each one of the samples, the signal from the probe sets representing these genes, should ideally be identical, and in practice it can be used for normalization for deviations across the samples. Each one of the four genes comes in a different concentration (1:100,000; 1:50:000; 1:25,000 and 1:7,500) and is mixed with the samples in a concentration that will span the expected intensity distribution of the signal probes.

The intensity of probe sets representing the spiked in genes for two microarray samples is plotted against the known concentrations of these genes as can be seen in figure 3.

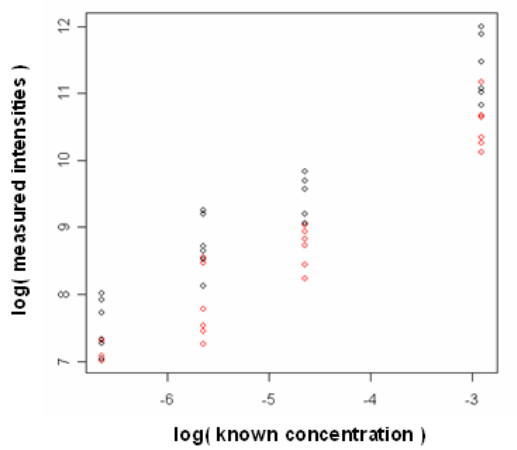

Figure 3: plot of the known concentrations vs. the measured intensities of all the probe sets representing the spiked in RNA genes. Each of the four genes is represented by several probe sets. The two colors represent two microarray samples.

Each one of the four control genes is represented by several probe sets representing different regions of the gene (3' end, middle and 5' end) and different methods for calculating the optimal probe sequences. The two colors represent two microarray samples. The variation of probe sets within each gene is probably the result of differences in the probe composition between the different probe sets. There is a clear consistent difference between these two microarray measurements which represent a global variation between the two microarrays that should be corrected for.

For each microarray measurement, a least square error linear fit is preformed (on a double logarithmic scale as in Fig. 3) on the spike intensities against the known concentrations. The intensity values of each microarray are divided by the slope of the fitted line to get a slope of unity for all microarrays; the differences in the intercept between every microarray to the microarray with the minimal intercept are then subtracted.

Figure 4A shows a scatter plot of two microarray measurements in two consecutive time points (0 and 15 minutes) in the non treated measurements before applying the above normalization. The red dots represent the spikes.It can be seen that the spikes deviate considerably from the diagonal, indicating a significant deviation between these two microarrays. Figure 4b shows the same plot after normalization, where spikes are brought to the same intensity values. The black dots in the two figures represent the probe sets for the *S. cerevisiae* mRNA. The normalization reveals the massive decrease

in these mRNA levels between time point 0 and time point 15 min, as evident by the accumulation of points below the diagonal. Similar picture is seen for the rest of all time points.

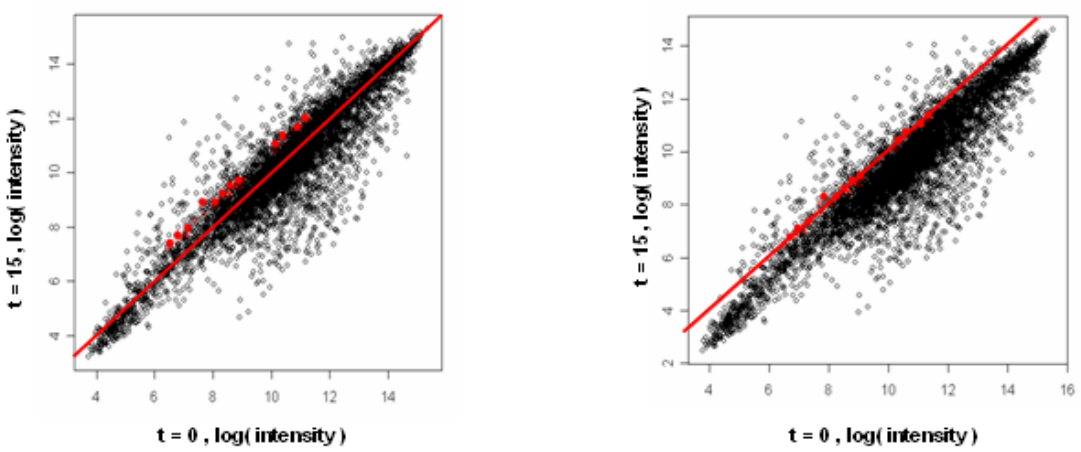

Figure 4: Illustration of the effect of spike in normalization. Two plots of probe set intensities, time point 0 vs. time point 15. Spike in RNA probe sets are marked in red. Right figure is the plot before spike normalization and the left plot after. Processing the data in order to put the spike-in RNA on the diagonal, reveals mRNA degradation.

Refinement:

Although the use of spiked-in RNA as an internal standard successfully reveals mRNA degradation, slight deviations in the amount of spikes mixed with each sample may cause imprecision in the normalization. Such errors in spikes concentrations may manifest themselves as deviations from the monotonic decrease in the mean sample intensities, which is expected due to the halt of transcription. Ideally, since most mRNAs are expected to decay exponentially the mean intensity of the entire transcriptome should show a linear decay on a log plot. Thus we added the step of Refinement in which sample mean log intensity of each time course is plotted as a function of time, and a linear fit is preformed and all samples are then fitted to this line. Figure 5 illustrates this transformation. This transformation is equivalent to what is done in microarray time course experiments where same average is assumed for all time course samples, except that here the average is assumed not to remain constant but rather to decay exponentially.

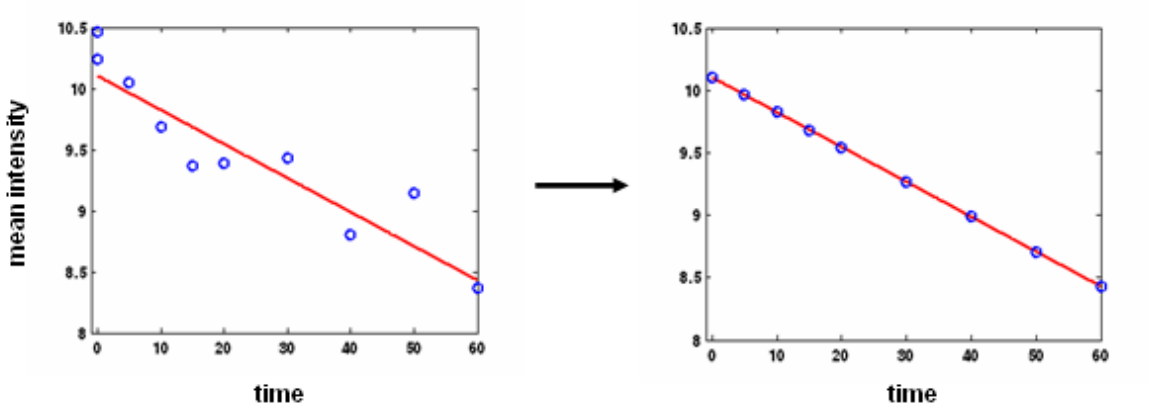

Figure 5: Illustration of the refinement step. For each time course experiment the sample mean intensity (blue dot) is plotted against time. A linear least square error is preformed and the data is scaled in order to have a mean value of corresponding to the fitted value at each time.

Averaging and zero transform:

After the above steps are preformed, for each gene the replicates of time point zero are averaged and all the data is divided by this point (subtracted in log scale).

4.5. Calculating gene specific half life

The first step in the analysis of such data is to assume that the rate of degradation is constant throughout the whole time course and to fit the data to a first order exponential model ($y(t) = y(0) \cdot e^{-kt}$), the rate constant k is then used in order to calculate for each gene a half life ($hl = \ln(2)/k$). Although we normalize all the data with respect to time point zero, hence all genes have the value of 1 for this time point, we do not fix the initial amount parameter ($y(0)$) to 1 but rather use it also as a free parameter. This is done because measurement errors are represented in all time points, including the zero time point, and constraining the fit to exactly match this point might allow for error measurements of the first point to bias the whole fit. The chosen parameters for each gene are those that minimize the least square error between the data to the model and are found using a non linear least square algorithm (using the Matlab fit function).

Fitting the data to an exponential model can be done either by fitting the data in linear scale to an exponentially decreasing decay model or by fitting the data in the natural logarithmic scale to a linear decreasing decay model. Although this seems as a

marginal issue it has an effect on the inferred parameters. Figure 6 illustrates the difference between the two choices. While low half life values in linear scale have an increased value in log scale the opposite happens in high values. This results in a narrower range for the half life distribution when the fit is preformed in log scale.

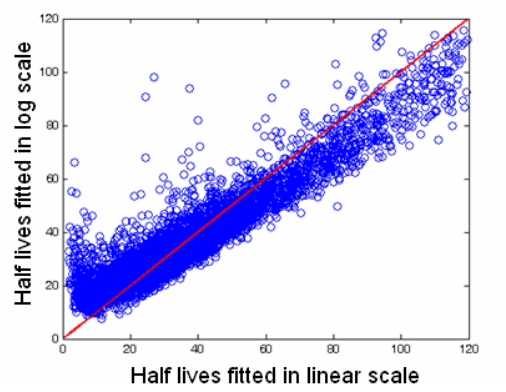

Figure 6: Half lives, inferred using a fit in log scale, are plotted against half lives inferred using a fit in linear scale. Fitting in log space produces a half life distribution that spans a narrower range.

We have decided to continue with fit results preformed in log scale because it makes the half life parameter more robust to perturbations in single data.

In addition, our ability to determine the half life of a gene depends on the precision of our measurements, which also varies between genes. This motivates the calculation of confidence bounds on the fit parameters, especially on the half life parameter. Calculating confidence bounds is done using the inverse R factor from the QR decomposition of the Jacobian, the degrees of freedom for error, and the root mean squared error (matlab confint function). For each gene a confidence bound of 68% is calculated on the rate constant k from which a confidence bound is directly calculated on the half life itself corresponding to the standard deviation of the half life.

4.6. Calculating half life coherence for groups of genes

We look for groups of genes that decay in similar rates in each one of the conditions separately. In order to find groups of genes (e.g. defined by a shared functional annotation) with similar half lives we use a method similar to what was used by Wang et al.[8] for the same purpose.

The standard deviation of a group of half life measurements is not independent of the mean half life of the group. In order to correct for this dependency 10^5 random groups of various sizes are selected without replacement and for each group the mean and standard deviation is calculated. A least square linear fit between the means and standard deviations is then calculated. Figure 6 shows the relationship between these two parameters and the fitted line.

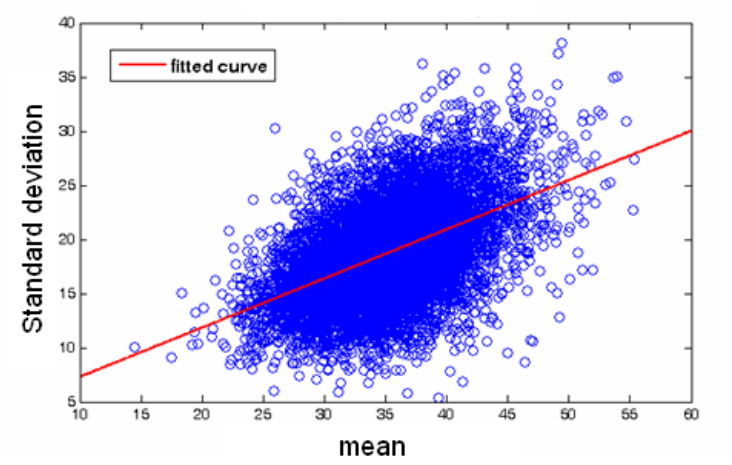

Figure 7: The half life standard deviation of random groups of genes with random sizes depends on the group mean half life. Each blue dot represents a random group of genes with x value as the mean half life and y value as the half life standard deviation. The red line represents the fit used to calculate the corrected standard deviation.

For each group of genes of size N that is typically defined by a shared annotation, a corrected standard deviation is calculated by dividing its actual standard deviation by the standard deviation expected for a group with that mean, using the above fitted graph. Then 10^4 random groups of size N are chosen without replacement and for each a corrected standard deviation is calculated. These 10^4 random corrected standard deviations are compared to the real corrected standard deviation in order to empirically calculate a p-value for each group. A set of genes will obtain a significant score if the corrected standard deviation of its members' half-lives is small compared to the corresponding values of random sets of the same size.

4.7. Assessing the significance of half life change

Our ability to determine that the degradation kinetics of a gene has changed between two measurements in two different conditions largely depends on our ability to

precisely evaluate the degradation kinetics of that gene in each one of those conditions separately. Using a constant rate decay model to describe the degradation profile of each gene, the change in half life can be used in order to evaluate whether a gene was stabilized or destabilized in response to a stress. The problem with this simplistic approach is that using the difference in half life directly may result in false discoveries of genes with a large change that is a result of noisy measurements and imprecise fit. Therefore we use the following distance measure that incorporates the goodness of fit for each gene in each condition: For gene g, the distance between two conditions (e.g. between non treated measurements to oxidative stress measurements) is the ratio between the absolute difference in half life to the sum of standard deviations on the half life fit parameter.

$$d_g = \frac{|\mu_{hl1} - \mu_{hl2}|}{\sigma_{hl1} + \sigma_{hl2}}$$

Note that the parameters do not come from a known distribution, we take as the mean the half life parameter that minimizes the sum of square error between the model to the data and as the standard deviation the confidence bound of 68% in the relevant direction (confidence bounds are non symmetric).

4.8. Finding groups of genes with a consistent change in the decay kinetics

The following procedure is used in order to compare each stress condition to the non treated measurements: For each group of genes of size N, the mean Euclidian distance between degradation profiles of all pairs in the group is calculated for the stress and non treated measurements separately. Then the ratio between the mean of these two values to the mean Euclidian distance of each gene's degradation profile in the two conditions is taken as the score.

$$score = \frac{\frac{1}{2\binom{N}{2}}\left(\sum_{i,j} d_{Euclidian}(v_i, v_j) + \sum_{i,j} d_{Euclidian}(u_i, u_j)\right)}{\frac{1}{N}\sum_i d_{Euclidian}(v_i, u_i)}$$

For gene i, v_i, u_i represents the degradation profiles in the non treated and treated conditions. To evaluate the significance of the score 10,000 groups of genes of size N are randomly selected without replacement and the score of these groups is used to empirically calculate a p-value.

4.9. Fitting each gene to a non constant decay kinetics model

In order to look for genes, or groups of genes, that show non constant decay kinetics we have developed a first principles kinetics model, which describes the effect of the interaction of RNA binding proteins to their target mRNAs, on the mRNA concentration as a function of time. The details of this model will be described below. Each gene is fitted to both a non constant decay model and a constant (exponential) decay model for comparison. The fit is preformed using gradient decent optimization scheme with the square error between the model to the data as the merit function. The software was implemented in C++ using GNU scientific library (GSL) for the optimization routine. The output contains parameters for the best fitted model for each gene, the squared error between the data and the model with the best parameters and a cross validation error which is calculated for each gene separately using leave one out cross validation method on the time points. In addition we calculate for each gene the r-square score (Coefficient of Determination) which is a measure to the amount of the variance in the data explained by the fitted model.

$$Rsq = 1 - \frac{\sum_i (y_i - f(x_i))^2}{\sum_i (y_i - \hat{y})^2}$$

Because we compare two models with different number of free parameters for each we also calculate an adjusted R-square which is an attempt to fix by adjusting both the numerator and the denominator by their respective degrees of freedom. In the following equation n stands for the number of observation (time points in our case) and k stands for the number of free parameters.

$$adjustedRsq = 1 - (1 - Rsq)\left(\frac{n-1}{n-k-1}\right)$$

5. Results

5.1. Response validation using real time PCR

Two aspects of the experiment needed to be validated before the real experiments could have been preformed and RNA was sent for hybridization to the arrays: first, the transcriptional response to the stress and second, transcription inhibition following the temperature shift. Calibration experiments were preformed to determine the stress dosage and exposure time, and for calibrating the transcriptional arrest. We used known responsive genes for each stress to measure their transcript levels under various induction and halting conditions. Two cultures were grown at 25°; one was treated with a stress while the other was not, and samples were taken every 5 minutes for one hour. mRNA abundance levels of known responsive genes and control genes were measured using real time PCR. These measurements were also used to validate that the chosen time gap between applying the stress to inhibiting transcription was sufficient, at least for a few known responsive genes using the consideration described in the experimental procedure. Figure 8 shows the mRNA abundance level of the gene GPX2, known to be induced by oxidative stress, in several samples along this time course. The mRNA levels shown are normalized by the mRNA levels of a housekeeping gene RPL25 as described in the methods. Induction was validated using GPX2 and TSA2 for the oxidative stress and using MEP2, GLN1 and GAP1 for the rapamycin stress.

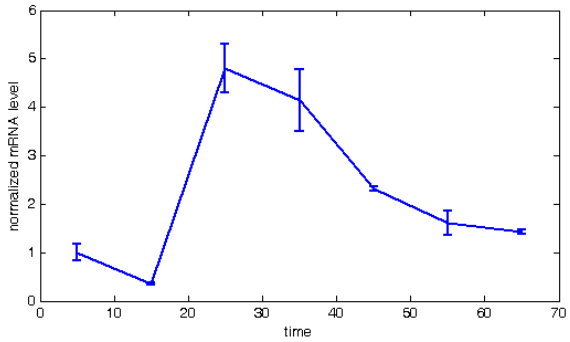

Figure 8: Real time PCR measurements of GPX2 mRNA normalized to RPL25 in a calibration experiment for hydrogen peroxide. This gene reaches it's peak at 25 minutes.

Transcription inhibition was validated (on both stresses) using mRNA levels of RPS6B normalized to the mRNA levels of PGK1 which are known to be fast and slow

decaying genes respectively. As seen in Figure 9, degradation of this gene is observed. This was shown for both stresses and only in the non-permissive temperature.

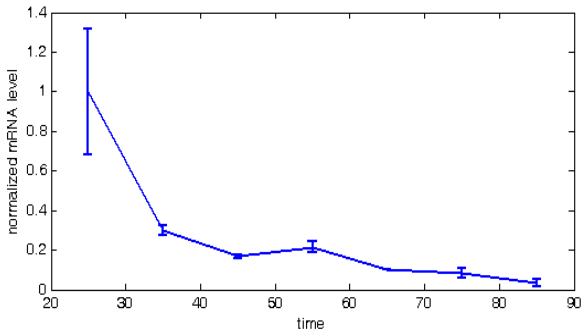

Figure 9: Real time PCR measurements of PGK1 (a fast decaying gene) following transcription inhibition.

5.2. Half life determination for each gene

Oxidative and non-treated measurements – global view:

In the first experiments we measured the decay kinetics of genes in non stressful and oxidative stress conditions. These two experiments were done simultaneously: A culture was grown to the concentration of $2*10^7$ cells/ml in a volume of 500 ml. The culture was then divided to half and the first was treated by hydrogen peroxide while the second was not. Samples were then taken from both cultures as described in the methods, the -25 minutes time point was taken before the samples were divided, thus, for this time point the intensity values for the two time courses are the same. After the experiments were preformed RNA was extracted from both time courses and all 21 samples were sent to hybridization together in the same processing and hybridization batch.

For the oxidative and non treated data, performing the preprocessing procedure described above reveals the global mRNA degradation of most of the genome. A linear decrease in the mean log intensity of samples over the time course together with a high goodness of fit score to an exponential decay model for most of the genes is observed. Figure 10 shows the mean intensity of both non treated and oxidative stress time courses before the refinement step of the preprocessing. The fitted lines used for refinement are

almost the same in both time courses indicating that in these two measurements the global decay of the whole genome is similar.

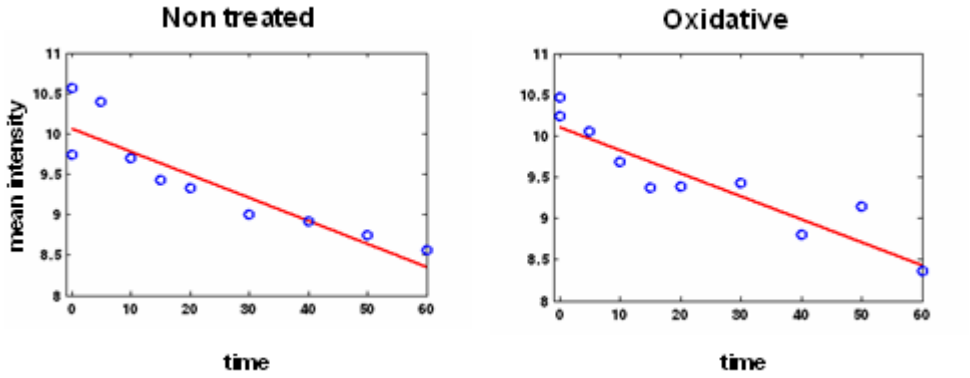

Figure 10: Sample mean log intensity vs. time for the two time courses (non treated and oxidative). The red line represents the fitted line used for refinement. Following spiked-in normalization a constant decrease of the mean intensity is observed which is similar to both time courses.

After a normalized expression measure is calculated for each gene at each time point, using the preprocessing procedure described, every gene's decay profile is fitted to an exponential model. Figure 11 shows a histogram of the r-square values for our measurements compared to the r-square histogram of previous measurements by Wang et al.[8](both data sets fitted using the same method and scripts). The data shows r-squares before the refinement step which evidently improves appreciably the goodness of the fit of the data to an exponential model. The figure shows that our data has a significantly better fit and is of higher quality compared to previous measurements. This is expected considering that our measurements were done using Affymetrix microarrays while previous measurements were done using cDNA arrays.

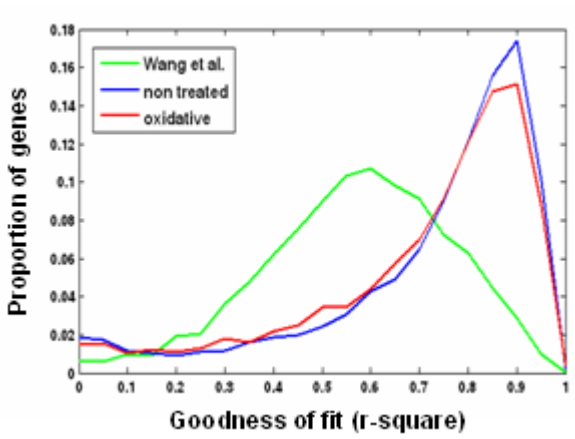

Figure 11: Histogram of goodness of fit scores (r-square) for our measurements compared to previous measurements by Wang et al.

When comparing the half life distribution between non treated measurements to oxidative stress treated measurements we discard all genes that have an r-square that is smaller than 0.6 (~20%). Figure 12 shows the half life histogram of the treated compared to non treated measurements after performing the complete preprocessing of the data (including the refinement step).

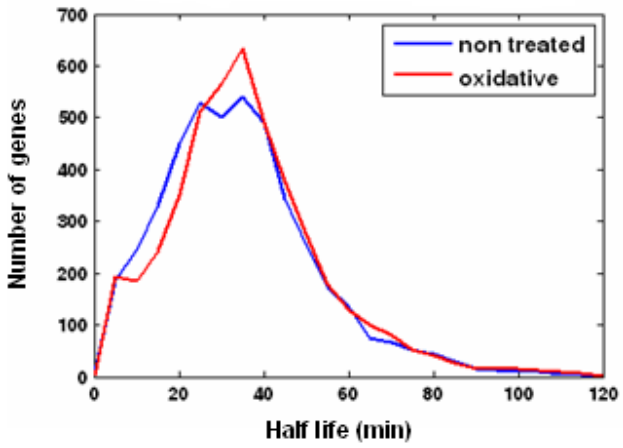

Figure 12: Histogram of half lives measured by non-treated and oxidative stress time courses.

Most genes have a half life between several minutes to around an hour. The mean half life is 34.5 minutes in the non treated measurements and 36 minutes in the oxidative stress measurements. These results are slightly higher than previous measurements by Wang et al.[8] (10 minutes in the mean half life). A slight but significant shift toward higher half lives is observed in the half lives distribution from non-treated to oxidative

stress ($p < 2.2438e-005$ using two sided T test) indicating to stabilization of a large set of genes upon the stress.

Oxidative and non-treated measurements – gene specific change in half life:

In order to follow the change in half life for each gene we plot for each gene the half life in the non treated measurement against the oxidative stress measurement as shown in figure 13.

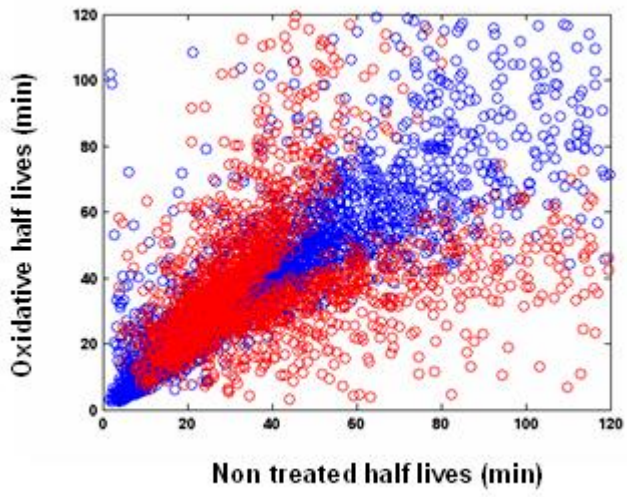

Figure 13: Half lives measured in the non-treated time course vs. half life measured in the oxidative stress time course. In red are genes whose significance score for the change in half life is larger than one, in blue are the rest of the genes.

Genes that are marked with red represents genes for which the significance score for the change in half life is larger than one, meaning that the difference in the half lives between the two measurements is larger than the sum of the 68% confidence bounds (in the relevant side).

The first observation from this figure is that for most genes the half life is correlated between the two measurements (correlation coefficient of 0.7186 after discarding genes with low r-square). Although for most genes the half life is correlated in the two conditions, there are 864 genes with a significant increase in the half life and 1181 genes with a significant decrease in the half life upon the stress. Interestingly, this difference is highly non symmetric: genes with high ½ life in the non-treated condition show de-stabilization, while genes with low ½ life values display stabilization.

Rapamycin measurements:

The Rapamycin experiment was preformed the same as the oxidative stress yet RNA extraction and hybridization were performed on a separate batch.

Unfortunately, in contrast to the former two time courses, in this case normalizing the data using the spikes did not work very well. The mean sample intensity did not show a nice linear decrease, as a function of time, and most genes did not show a high r-square value to the exponential model, as in the previous cases. We think that the reason for that is a difference in the ratio between the amounts of sample RNA and control RNA mixed together compared to the previous case. Indeed the spike intensities do not span the whole intensity distribution of the samples while in the previous measurements it did.

An attempt to discard several samples that seemed as outliers did not result in a significant improvement in the distribution of r-square. While this difference may reflect in part a global change in the genome decay profiles in response to rapamycin we cannot differentiate between real global changes in the mean mRNA levels on one hand, and errors in the normalization step that are a result of imprecise amount of spiked in RNA on the other.

In order to compare the rapamycin measurements with the two other time series we used the line fitted to the mean sample intensity as a function of time in the non-treated measurements and use this line in order to refine the rapamycin data. Doing this transformation we assume that the mean mRNA level at time t following transcription inhibition is equal in the rapamycin and the non-treated measurements. In other words we assume that the decay of the global amount of mRNA is independent of the stress which is similar to the assumption taken when microarray time course experiments are normalized to have the same mean. Thus, although global effects on the half life distribution are now missed genes that respond to the stress by changing their ranking within the distribution of decay profiles would still be discovered.

After performing this refinement the non-treated measurements and rapamycin measurements show similar distributions of both half lives and r-square as expected.

Figure 14 shows the scatter plot of gene half lives between non treated measurements and rapamycin measurements.

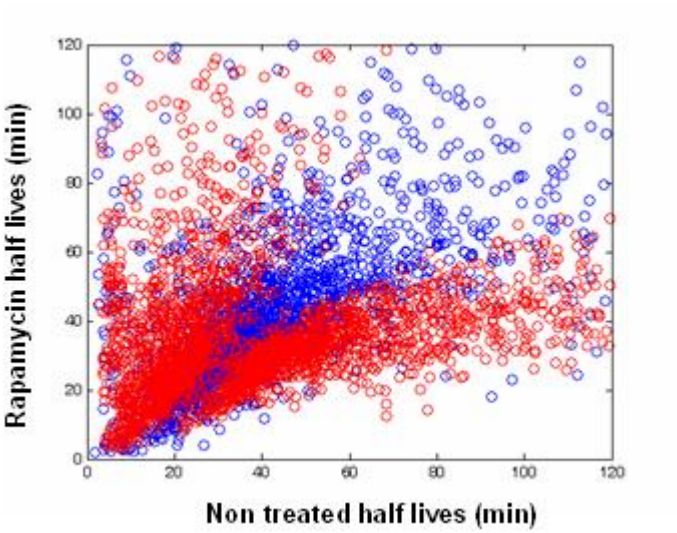

Figure 14: Half lives measured in the non-treated time course vs. half life measured in the rapamycin stress time course. In red are genes whose significance score for the change in half life is larger than one, in blue are the rest of the genes.

Similar to the comparison of the non-treated measurements to oxidative stress, a large set of genes show similar half life in both measurements (correlation coefficient of 0.5036 after discarding genes with low r-square < 0.6) with large outliers indicating to the modulation of half life of a large set of genes. More specifically 1129 genes show a significant increase and 1427 genes show a significant decrease in the half life upon the stress.

5.3. Groups of genes with coherent half life

Before comparing the change in half life of groups of genes we wanted to see if, similar to what was found by Wang et al.[8], groups of genes defined by a shared annotation tend to decay in similar rates and which groups show such behavior. Three sets of annotations were used: the gene ontology (GO), KEGG pathways and protein complexes (data received from Shoshana Wodak, personal communication), a p-value is calculated on each group's corrected standard deviation (as described in the methods) and for each set of annotations FDR[37] is used in order to find all the significant groups. Two FDR cut offs of 0.05 and 0.1 are used representing 5% or 10% false positive rates respectively. Results are summarized in figure 13;

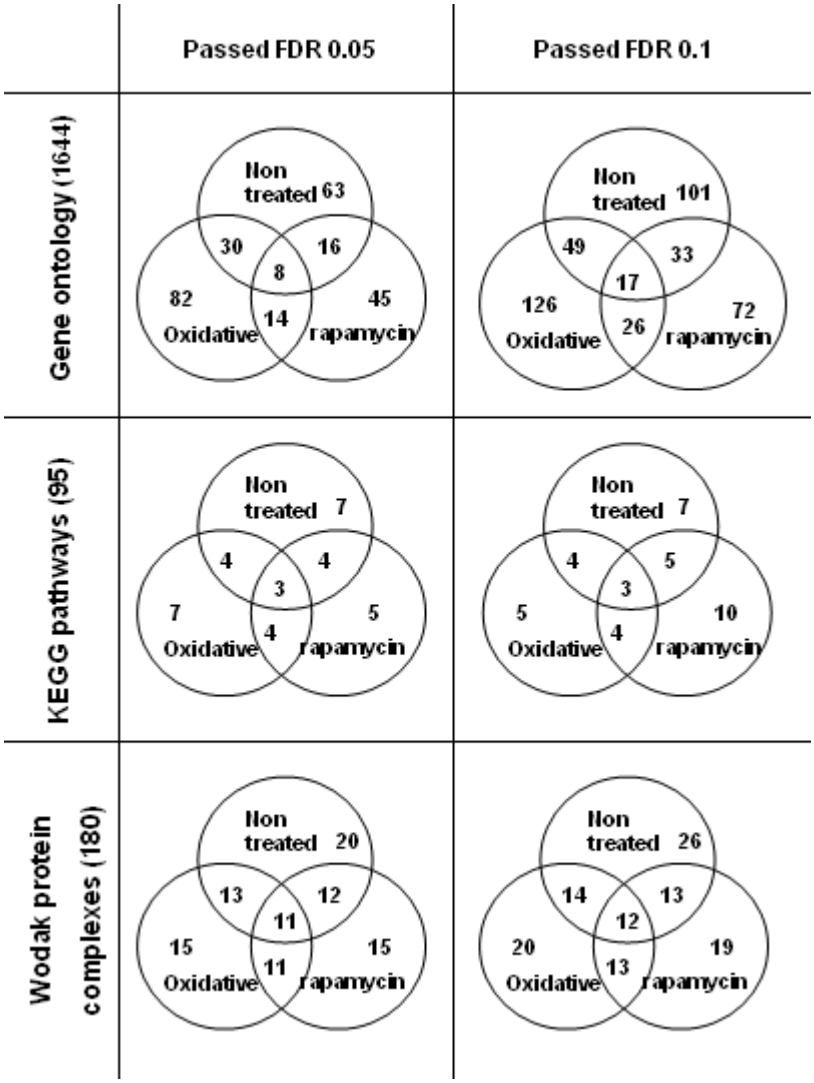

Figure 15: A table summarizing results of the test for half coherence of groups of genes. Each cell in the table contains a Venn diagram with the results for the three time course measurements.

consistent with the finding of Wang at al.[8] we find a significant number of groups that pass the FDR threshold, many of which are shared by all the 3 measurements but some are unique to a specific condition indicating that in some groups coordinated decay is condition specific.

In the KEGG pathway annotation, for example, genes which are part of the ATP synthesis, Ribosome and Proteasome pathways show coordinated degradation in all three measurements. Cell Cycle genes seem to be coordinated only in the oxidative stress measurements. In the protein complexes, groups that are related to Ribosome, ATP synthesis and subunits of the Proteasome are again coordinated in all three measurements. Also, the COPII complex (mediates export of protein cargo from the ER) and Coat-complexes (takes part in transport of proteins through the secretory pathway)

seem to have a coordinated decay which is shared by the three conditions. An example to a protein complex which is coordinated only in the oxidative stress measurements is the SNARE protein complex, which mediates fusion of cellular transport vesicles with the cell membrane. Actin-associated-proteins are an example of genes which are coordinated only in the rapamycin measurements. It is important to keep in mind that only whether the decay is coordinated is tested, thus, coordination might be shared by two conditions yet the response will be different in each of them, e.g. stable in one and unstable in the other.

5.4. Degradation kinetics is coordinately modulated in functional groups of genes in response to environmental changes

In order to check whether the decay kinetics of a group of genes is coordinately modulated in response to the two stress conditions, we chose a method that uses the Euclidian distance as a measure of distance between the entire decay profiles and not the difference in half lives which is biased to a predefined model of degradation (e.g. exponential vs. alternative models). This choice is motivated by the observation that a large set of genes show decay profiles that are not constant (as will be described below), in this case fitting a constant decay model and using the change in the half life parameter as a measure of the difference may be misleading due to imprecise fit.

For each group of genes (with n>2) in gene ontology, KEGG pathways and protein complexes we calculate the score representing the ratio between the mean distance within the group in each one of the conditions to the mean distance of genes in the group between the two conditions and assign a p-value for each group as described in the methods. Two FDR cut offs of 0.05 and 0.1 are used representing 5% or 10% false positive rates respectively. Intuitively, a set of genes will pass the test if it has coherent expression in each condition, provided that the decay profiles changed appreciably between the two conditions. Results are summarized in table 2: a large number of groups pass the FDR thresholds (~15-20%) indicating that the changes in half lives that were shown before tend to be similar in groups of genes that share a function or are part of a protein complex. Although we do not know the molecular mechanisms underling these changes, this observation might indicate that also at the level of mRNA decay there is co-

regulation by the same protein factors. More evidence to this idea is given by Gerber et al. that shown that each member of the Pumilio-homology domain (Puf) family of RNA binding proteins binds to a distinct set of about 40-220 different mRNAs with common functions and sub-cellular localizations.

	Passed FDR 0.05			Passed FDR 0.1		
	oxidative	rapamycin	shared	oxidative	rapamycin	shared
Gene ontology (n=1644)	97	103	35	123	149	47
KEGG pathways (n=95)	20	7	5	22	7	6
protein complexes (n=180)	28	38	15	36	48	19

Table 2: A table summarizing results of the test, for changes in decay kinetics, between the treated time courses to the non-treated time course.

Figure 16 shows some examples to these changes in decay kinetics. It can be seen that the change is not reflected in a constant rate change throughout the whole time course but in transient changes, especially at the beginning of the measured time course, that give rise to a qualitative difference in the decay kinetics between the two conditions. For examples, while the Sec 62-63 complex decays exponentially in oxidative stress it shows a clear deviation from a simple constant decay in the non-treated condition. This is a difference that would not be captured by a mere inspection of half life. These changes in kinetics might be a reflection of a transient nature of changes in stability or a result of the non physiological condition where transcription is halted and the mechanisms controlling mRNA stability are degraded as well.

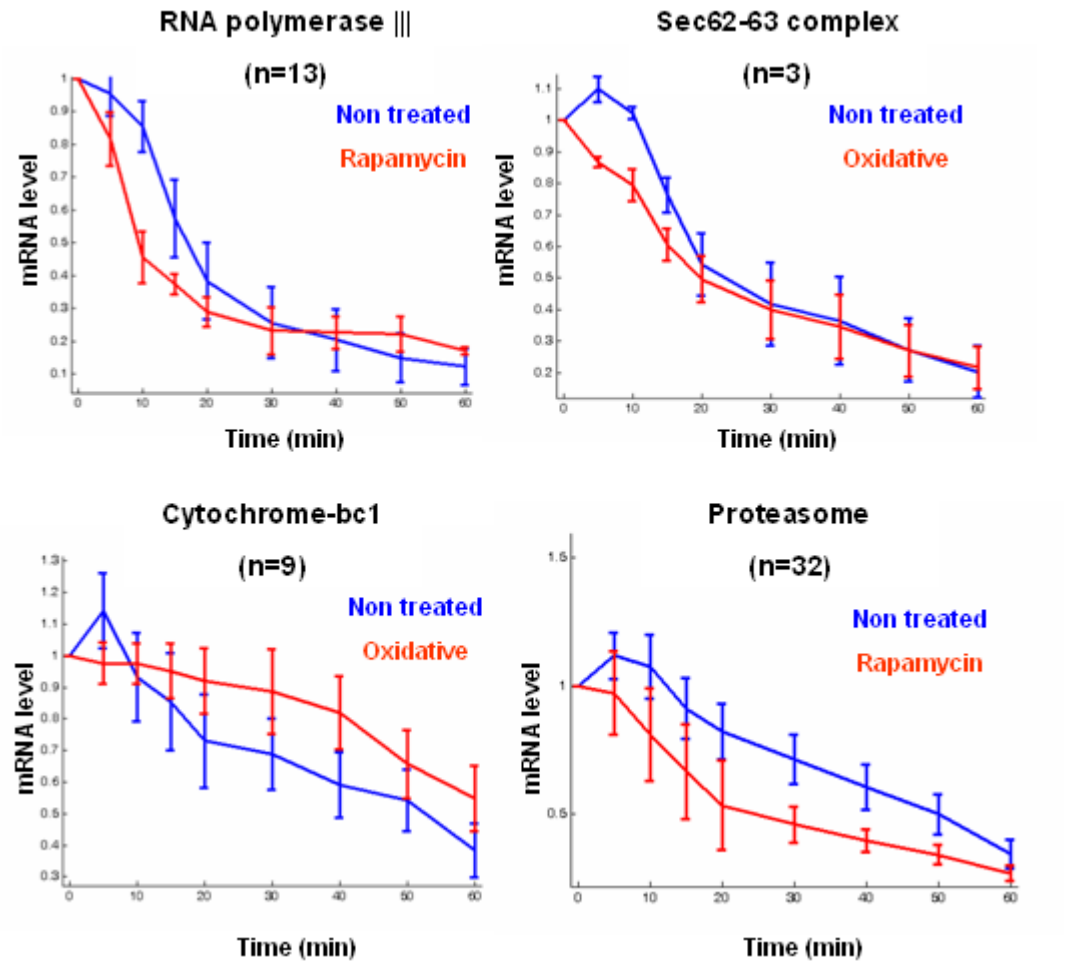

Figure 16: Examples of groups of genes with coordinated modulation in the decay kinetics between the non-treated time-course (blue) to a stress time-course (red). Line represents the mean of the group at each time point with error bars representing one standard deviation.

5.5. Opposite effect on mitochondrial and cytoplasmic ribosomal proteins

One of the most interesting observations that came up using the previous analysis is the effect of oxidative stress on the cytoplasmic and mitochondrial ribosomal proteins. While a lot is known on the regulatory mechanisms controlling the mRNA levels of cytoplasmic ribosomal proteins (CRPs)[38], no promoter motifs were found to explain the control on mRNA levels of mitochondrial ribosomal proteins(MRPs). Figure 15 shows that in response to oxidative stress, while CRPs are coordinately destabilized, MRPs are coordinately stabilized. Moreover, it seems that in oxidative stress, the two groups of ribosomal protein decay in a very similar rate while they are further apart in non treated conditions. This could mean that in oxidative stress these proteins are coordinately

regulated to have similar decay rates or that they are differentially regulated in normal condition, a regulation which is relaxed in response to stress.

In the rapamycin measurements the CRPs show a similar response of de-stabilization, but the change in MRPs is less apparent.

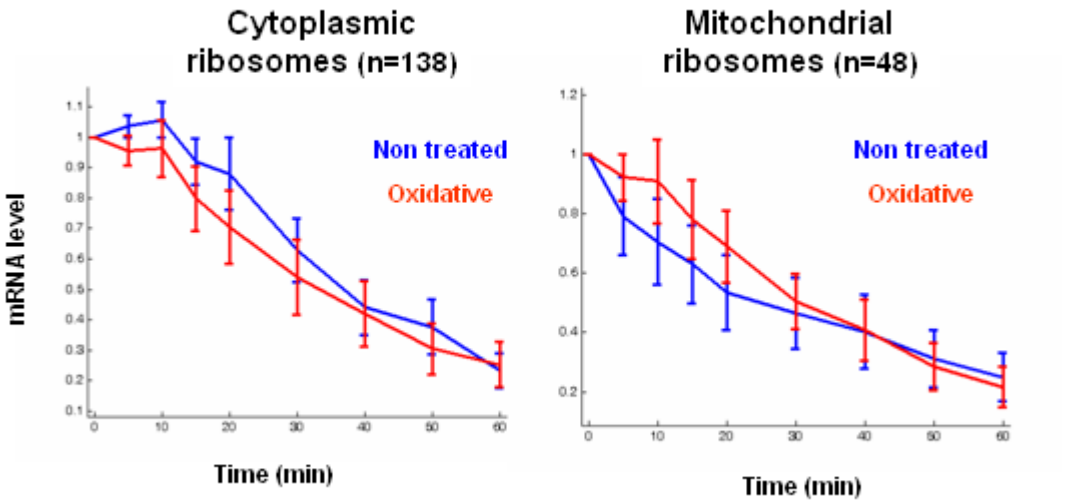

Figure 17: Mean decay profiles of cytoplasmic and mitochondrial ribosomal proteins between the non-treated time-course (blue) to the oxidative stress time-course (red). Line represents the mean of the group at each time point with error bars representing one standard deviation.

5.6. Direct coupling between transcription to degradation

In all three measured time courses, before transcription inhibition, a sample was taken for hybridization at time point -25, before the stress was applied. Comparing this sample to the sample taken at time point 0, concomitant with transcription inhibition, is used to evaluate changes in mRNA abundance in response to each stress. Of special interest is the relationship between changes in mRNA abundance in the treated measurements (e.g. genes that are induced or repressed in response to the stress) to changes in degradation kinetics between the non-treated to the treated time course. Suppose that a gene is induced by enhanced transcription in response to an external stimuli, will the decay rate of this gene be coordinately modulated, and in what direction? Will it be stabilized or de-stabilized? Intuitively one could think that the decay rate would be decreased for a gene that was induced in order to raise the level of mRNA as fast as possible, but is this really the best strategy to do so?

We have investigated the relationship between the changes in mRNA abundance between time points -25 to time point 0 in the oxidative treated measurements to the change in half life as a result of the oxidative stress. Figure 16 shows a scatter plot of this relationship, where on the x axis the logarithm of the half live ratio, between oxidative stress to non treated measurements, is plotted (negative values on this axis stand for genes that were destabilized). On the y axis the logarithm of the fold change in mRNA abundance levels between time point 0 to time point -25 in the oxidative stress measurements is plotted, positive values on this axis correspond to genes that were induced following the stress in this time interval. The figure shows a clear correlation between induction and de-stabilization: genes that are induced by the stress seem to be de-stabilized (with respect to the non-treated half life) and the proportions of these two responses seem to be related. No simple correlation between mRNA abundance levels to the half life is observed that might explain this.

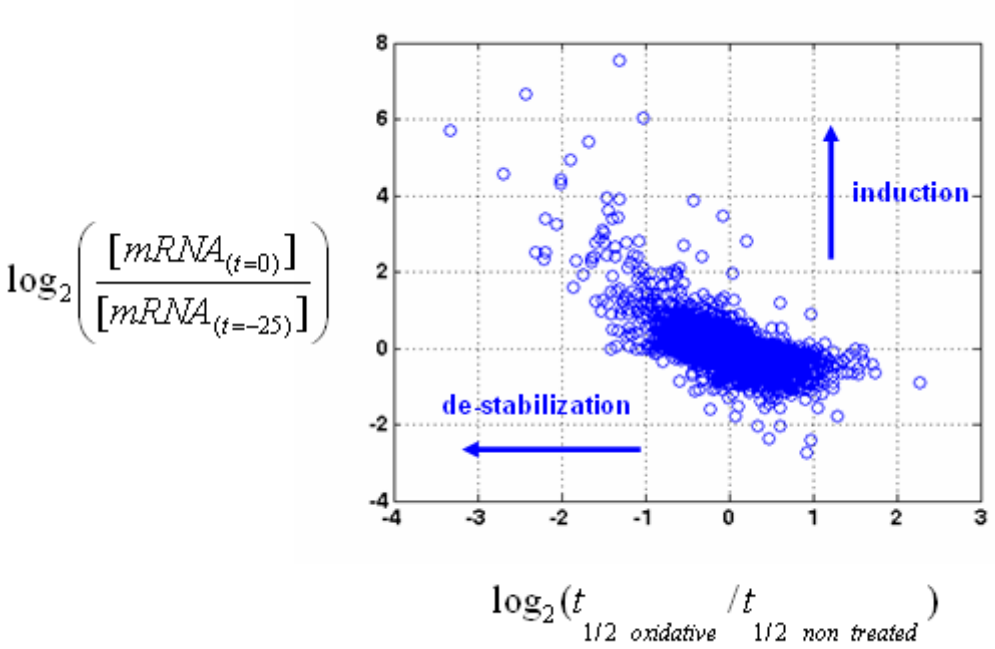

Figure 18: The change in half life between the oxidative stress time course to the non-treated time course vs. the change in mRNA abundance in the oxidative stress time course before transcription inhibition. Negative values on the x axis represent genes that were destabilized: have a reduced half life in the oxidative stress time course. Positive values on the y axis stand for genes that were induced in response to hydrogen peroxide.

What can the cell gain from such a correlation? The change in the mRNA level of a gene assuming a constant rate of transcription α, and a constant rate of degradation β is

given by: $\frac{d([mRNA])}{dt} = \alpha - \beta[mRNA]$. Thus, at steady state the mRNA level is given by the ratio of these two parameters α/β. Once a "decision" to raise the mRNA levels of the gene is made, it can be achieved using several strategies as shown in the simulations in Figure 17. One way is to increase the mRNA abundance level is to raise only transcription rate, this is represented in the figure by gene g_1. Raising the rate of transcription and degradation coordinately, while keeping the ratio between them constant (the rate of transcription would need to be higher than the rate of g_1), is represented by g_2. In such a case the same steady state is reached (as g_1) while the response time (time till half of the steady state level is reached) is shorter. Raising only the rate of degradation and keeping the rate of transcription as the rate of g_1, represented by g_3, will indeed shorten the response time but the steady state level would be lower, which might not be sufficient. g_4 represents the option when only the rate of transcription is higher than in the rate of g_1, in this case the steady state levels would be higher and the response time with respect to these levels would stay the same. It is true that in this case reaching half of the amount defined by the steady state level of g_1 would occur faster but eventually the steady state level reached would be considerably higher, a change which might not be needed. Thus, assuming that we would like to reach the steady state level defined by g_1, the best way (with respect to the response time) to get there would actually be to raise both the level of transcription and degradation as represented by g_2.

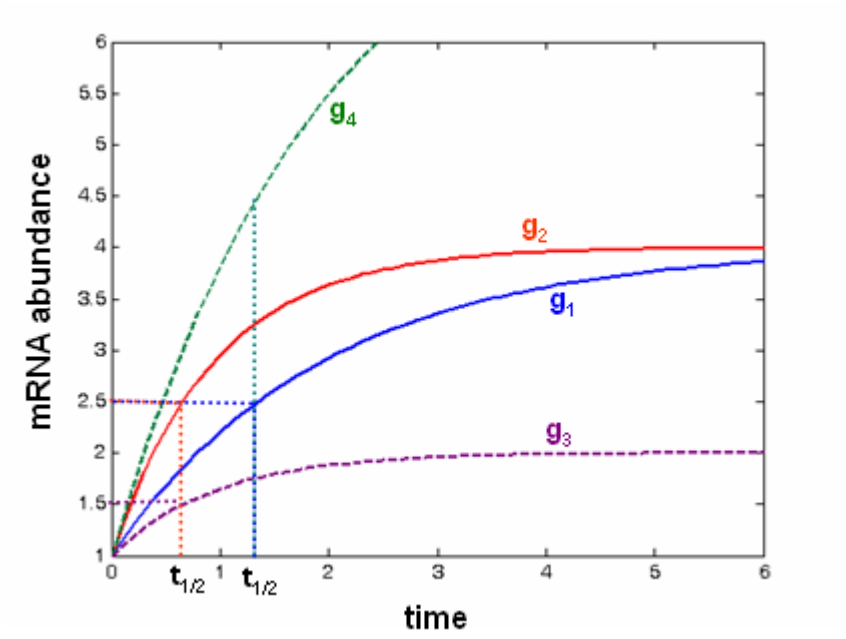

Figure 19: Examples of genes with different induction strategies, g1 is used as reference; in g2 both transcription and decay rates are increased keeping the ratio between them constant with respect to g1. g3 and g4 represent genes for which only the rates of degradation and transcription are increased respectively.

Enhancing the response time by coupling increased transcription with increased degradation has an effect not only on the way to reach a higher steady state but also on the way back to the initial (basal) level. Figure 18 shows a simulation of a full profile of mRNA level when transcription is increased and stopped completely after 6 minutes. It can be seen that g_2 reaches the higher level faster but also returns back to the basal level faster compared to g_1 while the both reach the same maximal level defined by the ratio between the rates of transcription to the rate of degradation in the first half of the time course.

Thus, the gain from coupling increased transcription with increased degradation is dual: First a new steady state is reached faster giving rise to faster response to stressful condition and higher fitness. Second, reaching back the basal level is also faster giving rise to faster adaptation.

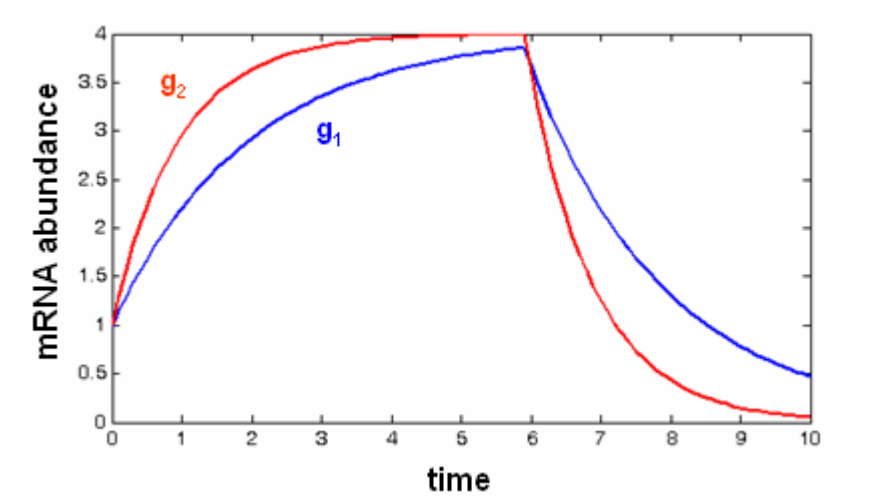

Figure 20: A time course of induction and adaptation for two genes with equal ratios of transcription to degradation but higher rates of both for g2. After 6 minutes transcription is stopped completely for both genes. Faster responses are observed for g2 on both induction and decay back to the basal level.

Figure 16 shows that genes that respond to oxidative stress with elevated levels of mRNA have also a reduced half life with respect to measurements in non stressful conditions. We propose that underlying this result is a molecular mechanism by which transcription is directly coupled to degradation. This mechanism is different than gene specific modulation of degradation kinetics by specific regulators on selected groups of genes with a functional relationship. Here we propose that once transcription is enhanced, degradation is coordinately enhanced as well by a mechanism that might involve protein factors that are part of the general molecular machinery involved in transcription initiation, export or degradation.

Throughout the life course of the mRNA it is escorted by a host of protein and RNA factors, some of which are stably bound while others are exchanged when moving from one stage of mRNA processing to the other. This complex of an mRNA with bound factors is denoted as the messenger ribonucleoprotein particle (mRNP) and individual components of this complex are proposed to interface with different intracellular machineries involved in processing, export, translation and decay of the mRNA[14]. This complex was proposed to serve as the ground for post transcriptional operons in eukaryotes and to link the different stages in the life course of each mRNA to each other. Taking this into account, and the fact that some protein factors were shown to be

exported from the nucleus while staying stably bound to the mRNA, it may be suggested how direct coupling between elevated transcription to elevated degradation might occur.

5.7. A kinetic model describing the effect of RNA binding proteins on degradation kinetics

The cases of mRNAs decaying non-exponentially motivated us to develop a first principle general models that would account for such kinetics, in which exponential decay will correspond to a special case. We have developed a kinetic model that will describe the interaction of mRNA binding proteins with mRNAs and how this interaction affects degradation kinetics. We have taken as the rate limiting step the removal of the poly(A) tail, thus an mRNA is considered completely degraded once it is deadenylated. Wilusz et al.[13] describes a scheme, based on several examples, on how stabilizing and destabilizing protein factors might act (figure 2). Immediately after an mRNA is exported from the nucleus to the cytoplasm it is safe from deadenylation due to the binding of PABP (poly(A) binding protein) to the poly(A) tail and cap binding proteins (as eIF4E) to the 5' cap. During translation the mRNA is thought to be circularized by the interaction of these proteins, this forms a complex that prevents deadenylation and keeps the mRNA intact for translation. A RNA binding protein, acting as a stabilizing factor, will bind the mRNA to stabilize this complex preventing deadenylation and increasing the time in which this mRNA is intact. A RNA binding protein, acting as a de-stabilizing factor, will bind the mRNA to interfere with this complex by competitive binding or by recruiting degradation factors as the exosome decreasing the time in which the mRNA is intact. Once an mRNA is free in the cytoplasm it is degraded in a rate which is independent of any regulatory sequences and assumed to be shared by all genes.

Assuming a stability factor S whose binding to an mRNA prevents its degradation, we obtained the following kinetic equation:

$$mRNA - S \underset{k_0}{\overset{k_1}{\rightleftarrows}} mRNA + S \xrightarrow{k_2} \Phi$$

We would like to derive an expression for [mRNA$_T$], the total amount of mRNA measured by the microarray.

$$[mRNA_T] = [mRNA] + [mRNA - S] (*)$$

We assume separation of time scales, namely that the binding and unbinding of the stability factor to the mRNA occurs much faster than the degradation of a naked mRNA. This assumption is equivalent to the assumption made on the binding to transcription factors to promoters when transcription is modeled in the same manner.

Using separation of time scales we can assume equilibrium in the binding and unbinding of the stability factor to get the following equation:

$$[mRNA - S]k_1 = [mRNA][S]k_0 (**)$$

Combining equation (*) and (**):

$$k_1([mRNA_T] - [mRNA]) = k_0[mRNA][S] \;\Rightarrow\; [mRNA] = \frac{k_1[mRNA_T]}{k_1 + k_0[S]}$$

$$\Rightarrow\; [mRNA] = \frac{[mRNA_T]}{1 + \frac{k_0}{k_1}[S]} (***)$$

We would like to derive an expression for the total rate to degradation. By the above kinetic equation this is equal to:

$$\frac{d([mRNA_T])}{dt} = -k_2[mRNA]$$

$$\xrightarrow{(***)} \quad \frac{d([mRNA_T])}{dt} = \frac{-k_2[mRNA_T]}{1 + \frac{k_0}{k_1}[S]}$$

The above equation is integrated in order to get an equation describing the mRNA concentration as a function of time:

$$[mRNA_T]_{(t)} = [mRNA_T]_{(t=0)} e^{\frac{-k_2}{1+\frac{k_0}{k_1}[S]} \cdot t}$$

Looking on the above equation and comparing it to the simple exponential model, commonly used to describe mRNA degradation, we can see that we got the same exponential decay model only with a decay constant which is decomposed into several parameters each representing a different part of our kinetic model.

$$k_{decay} = \frac{k_2}{1+\frac{k_0}{k_1}[S]}$$

k_2 is a parameter representing the rate of deadenylation of an mRNA which is naked in the cytoplasm thus this parameter can be assumed to be shared by all genes and independent of any regulatory sequences. The ratio $\frac{k_0}{k_1}$ represents the affinity of the stability factor to the mRNA, it is a function of the sequence and for the same gene it will remain the same over different conditions. The parameter [S] represents the concentration of the stability factor, it will be shared by genes that are subjected to the same regulation and will change as a function of the environmental condition. This is illustrated in figure 21.

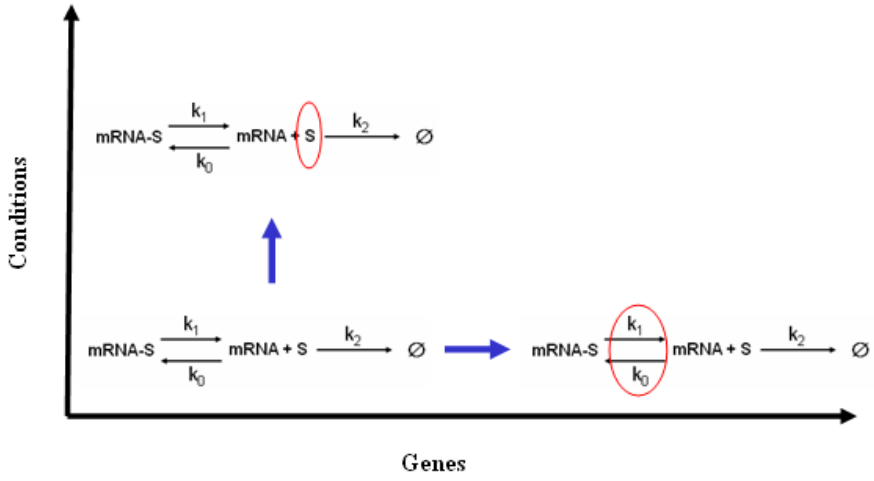

Figure 21: An illustration of the change in the model parameter values across the gene or condition dimensions. Marked in red are the parameters that change when we move in each one of the dimensions.

In order for the model to describe non constant degradation kinetics of an mRNA, which will be used to explain data instances that show such behaviour, we need to take [S] to be a function of time. We choose a logistic function to describe the concentration of [S], this describes a transient effect of stabilization: when stabilization is observed in the first minutes of the time course after which the rate of degradation increases.

$$[s]_{(t)} = 1 - \frac{1}{1 + e^{a-bt}}$$

Once mRNA decay is modelled using a logistic function for the concentration of [S] we can get a decay profile with non constant decay kinetics as illustrated in figure 22. Such decay profile simply describes cases when two decay rates are observed with a smooth transition between them as expected by the use of the logistic function.

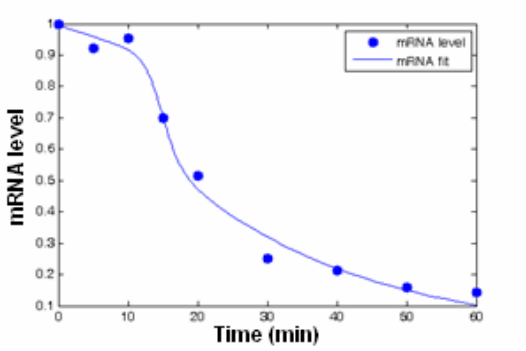

Figure 22: A decay profile of a gene that displays non constant decay kinetics with a function that represents the best fit to the model.

This model can then be used in order to detect data instances that show such behavior and also serve as a framework for further analysis where parameters are shared by groups of genes.

5.8. Analysis of the decay profiles using the non constant decay model

For each gene the best fit, with respect to the sum of square error, is calculated for both a constant rate decay model and our kinetic model. In order for the goodness of fit score to be comparable between the two models the number of free parameters has to be considered. We took two approaches in order to do so, one using a cross validation error

which measures the error in our predictions of unseen observation, and another by calculating the adjusted R-square that analytically corrects for the number of degrees of freedom in each model. The cross validation score is the sum, for all time points, of the square error between the predicted value, of the fit obtained by discarding that point, to the discarded point's real value. The adjusted R-square score is a function of the R-square, the number of observations and the number of free parameters as described in the methods.

Around 35-40% of the genes show an adjusted R-square which is higher for the non constant rate decay model compared to the constant decay model. This result is consistent for all three conditions and a similar result is obtained using the cross validation score. The results are summarized in the venn diagram displayed in figure 23. Although a large number of genes show non constant decay kinetics, only about half of these genes are shared by different conditions indicating that this property is condition specific.

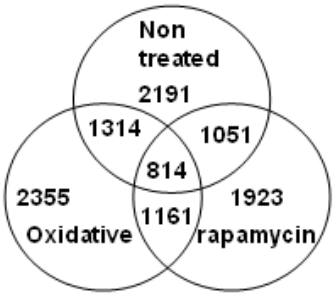

Figure 23: A venn diagram describing the number of genes that show non constant rate decay kinetics indicated by having a higher adjusted r-square in the non constant decay kinetics compared to a constant decay model.

In order to look for genes that display different decay kinetics we clustered the genes using the fitted parameters of the kinetic model in each one of the conditions (using k-means). Figure 24a shows the decay profile, in the oxidative stress time course, of two clusters of genes, one with parameters representing non constant decay kinetics and the other representing constant rate decay kinetics. A histogram of the adjusted R-square ratio between the non constant to the constant kinetic model is shown for the two clusters in figure 24b. An increase in the ratio for the cluster representing the non constant decay kinetics is observed indicating to the ability of this score to capture such behaviors.

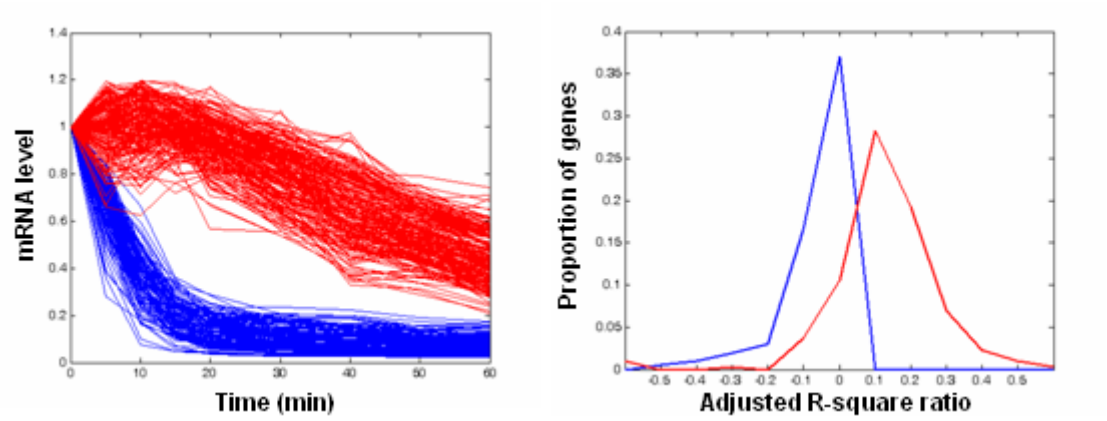

Figure 24: Example for cluster of genes displaying different decay kinetics in the oxidative stress time course with a histogram of the adjusted R-square ratio between non constant to constant decay kinetics.

Figure 24 shows different genes displaying different decay kinetics in the same condition. The same genes might display different decay kinetics in different conditions as apparent from the venn diagram in Figure 23. A nice example for a group of genes, for which the decay kinetics are changed between two conditions, are the genes for mitochondrial Ribosomal proteins changing between non treated to oxidative stress as described in chapter 5.5.

6. Discussion

The purpose of the present work was to assess the effect of environmental growth conditions on mRNA decay kinetics. While a lot is known about how the environment affects steady-state mRNA levels, little is known about the relative contribution of the two components that determine such levels, namely transcript production and degradation. An assessment of the relative extent of regulation of each of these levels is crucial not only because it will allow to better understand how degradation is regulated, but also because it will shed light on transcription production itself. Once degradation effect is "subtracted" from the mRNA concentration one may better estimate the extent of regulation acting at the transcript production level.

Our results show that in response to the two stresses investigated, massive changes in decay kinetics occur. Using a significance score that takes into account the goodness of fit, when comparing fit parameters between two different data sets, we show that the

difference is not a result of imprecise fit. This method can be used to separate between genes with significant to non significant changes at different levels of confidence.

As shown by others before us[8,21], we see that decay rates for groups of genes, defined by a shared functional annotation, are closely related suggesting that the control of their degradation is part of the regulatory mechanisms used to tune their expression levels. Although much of the coordinated groups show good coherence in several growth conditions, some are specific for only one or two of the conditions. More insight might be gained by a close inspection on the specific groups which have a coordinated decay in only one of the conditions suggesting that the decay of these groups is coordinately regulated in a condition specific manner.

To investigate whether the changes observed in decay rates are also coordinated for functional groups we chose a score for groups of genes which is not biased, as the exponential decay is, to a predefined model for the decay. Using this score we manage to identify groups of gene that show a coordinated change which is highly statistically significant. In these groups the changes are manifested not only in a general change in the decay rate over the whole measured time course but also in the kinetics of the decay profile itself as apparent from Figure 16. We realized that not all genes display a simple exponential decay. An alternative first principals kinetics model we developed uncovered a large portion of the genes whose decay display clear deviation from constant simple decay. Such genes may be candidates regulatees of a machinery that determines differential mRNA stability for the various genes in the genome.

In order to measure decay, we used a temperature sensitive mutant strain of *S.cerevisiae* bearing a temperature sensitive mutation in the RNA polymerase II. This raises two main problems: first, a heat shock is applied in order to stop transcription; second, if changes in mRNA stability require a transcriptional response (e.g. transcription of an RNA binding protein) we might overlook such changes due to transcription inhibition. Regarding the first point, it is true that our non treated time course, used as a control experiment, is actually treated by heat shock at the same time that transcription is stopped. Although the response to the other stresses may be partially overlapping with the response to the heat shock, our comparison is of two time courses in which the same method is used to halt transcription. Thus, we can relate all the differences to the additional stress. The second point actually reflects a trade-off between opposing parameters: early transcription inhibition might result in missing transcription dependent

responses, while late transcription inhibition will allow the cells to react to the stress, but will result in missing transient effects. This question is especially problematic in the case of degradation because nothing is known about the transient nature of changes in mRNA stability. We decided to wait a short period of time after applying the stress, and before inhibiting transcription. In this way the cells are given time to induce a transcriptional response to the stress and we hope that not all of the response will have been missed during this interval. Waiting a short period of time before transcription inhibition also relates to the first point mentioned above because while the cells are given time to respond to the additional stress they are not given time to respond to the heat shock due to the instantaneous halt of transcription.

It is conceivable that controlled degradation itself requires active transcription, in which case part of the decay kinetics will be distorted in the present experimental setup. One particular potential caveat is that stabilization factors are no longer produced following the temperature shift and the decay of some targets may be artificially enhanced upon depletion of these stabilizers. This could serve as an alternative explanation to the non-exponential decay that some genes were shown to display.

Another caveat of measuring mRNA decay during transcription inhibition is that the results may be relevant only to this non physiological condition. For example, our observations might reflect posttranscriptional responses to the reduction of transcription. A previous study measuring mRNA decay both following transcription inhibition and using *in vivo* radio-labeling showed that there is a correlation between the two measurements, yet this was demonstrated using a small number of genes[39].

The underlying biological variability of mRNA decay is another issue which is problematic when comparing measurements at different conditions. We have sampled the time course, over which transcription is inhibited and degradation occurs, densely enough in order to allow precise determination of degradation kinetics for most genes as they occurred in our experiment. Still no biological repetitions were preformed, implying that no indication to the biological variability of such data is provided. Microarray measurements of mRNA abundance reflect the mean mRNA level of a large population of cells, so single cell variability is not a concern. But the effect of small changes in the microenvironment, as slight differences in cell concentrations or in the medium composition, on the mean decay of genes in a large population of cells remains unknown. Nonetheless, the attempt to relate the observed changes to the environmental perturbation

applied comes from a function based analysis assuming that changes that reflect the biological variability would be randomly distributed among the genes.

The diverse composition of messenger ribonucleoprotein particles (mRNP) suggest that throughout the life of every mRNA, from transcription initiation, through processing, quality control and export till translation and degradation, it is subject to tight control which specifies the rates of each step and coordinates between them[14]. It is clear that the different stages of gene expression are tightly linked to each other: mRNA processing is linked to quality control and export, export is linked to degradation and translation, and in particular degradation is linked to transcription, which determines the mRNA steady state level for each gene. A comprehensive understanding of gene expression requires knowledge on how the life course of an mRNA is affected by this wide variety of regulatory mechanisms and how these mechanisms are coordinated and mutually affected. In particular, understanding the mechanisms behind changes in mRNA abundance requires understanding on how changes in transcription and degradation rates are related and coordinately regulated.

Indeed we find that genes that are induced in response to the stress have also a reduced half life with respect to the control measurements. We speculate that underlying this result is a general mechanism by which enhanced transcription is directly coupled to enhanced degradation. Such a mechanism will have a selective advantage by an increased response time as apparent from a simple inspection of different strategies for gene induction. Experimental evidence suggesting such mechanism exist, Lotan et al.[40] have recently shown that a subunit of RNA polymerase II, Rpb4p, is involved in mRNA decay by enhancing both deadenylation and decapping. It will be interesting to investigate whether such mechanism may account for the correlation observed here between induction and degradation.

References:

1. Gasch, A.P. et al. Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* **11**, 4241-57 (2000).
2. Spellman, P.T. et al. Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. *Mol Biol Cell* **9**, 3273-97 (1998).
3. Pilpel, Y., Sudarsanam, P. & Church, G.M. Identifying regulatory networks by combinatorial analysis of promoter elements. *Nat Genet* **29**, 153-9 (2001).
4. Lapidot, M. & Pilpel, Y. Comprehensive quantitative analyses of the effects of promoter sequence elements on mRNA transcription. *Nucleic Acids Res* **31**, 3824-8 (2003).
5. Segal, E. et al. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nat Genet* **34**, 166-76 (2003).
6. Beer, M.A. & Tavazoie, S. Predicting gene expression from sequence. *Cell* **117**, 185-98 (2004).
7. Garneau, N.L., Wilusz, J. & Wilusz, C.J. The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol* **8**, 113-26 (2007).
8. Wang, Y. et al. Precision and functional specificity in mRNA decay. *Proc Natl Acad Sci U S A* **99**, 5860-5 (2002).
9. Xu, N., Chen, C.Y. & Shyu, A.B. Modulation of the fate of cytoplasmic mRNA by AU-rich elements: key sequence features controlling mRNA deadenylation and decay. *Mol Cell Biol* **17**, 4611-21 (1997).
10. Wilusz, C.J. & Wilusz, J. Bringing the role of mRNA decay in the control of gene expression into focus. *Trends Genet* **20**, 491-7 (2004).
11. Parker, R. & Song, H. The enzymes and control of eukaryotic mRNA turnover. *Nat Struct Mol Biol* **11**, 121-7 (2004).
12. Sheth, U. & Parker, R. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* **300**, 805-8 (2003).
13. Wilusz, C.J., Wormington, M. & Peltz, S.W. The cap-to-tail guide to mRNA turnover. *Nat Rev Mol Cell Biol* **2**, 237-46 (2001).
14. Moore, M.J. From birth to death: the complex lives of eukaryotic mRNAs. *Science* **309**, 1514-8 (2005).
15. Vasudevan, S. & Peltz, S.W. Regulated ARE-mediated mRNA decay in Saccharomyces cerevisiae. *Mol Cell* **7**, 1191-200 (2001).
16. Graber, J.H. Variations in yeast 3'-processing cis-elements correlate with transcript stability. *Trends Genet* **19**, 473-6 (2003).
17. Guo, X., Ma, J., Sun, J. & Gao, G. The zinc-finger antiviral protein recruits the RNA processing exosome to degrade the target mRNA. *Proc Natl Acad Sci U S A* **104**, 151-6 (2007).
18. Chen, Y. & Varani, G. Protein families and RNA recognition. *Febs J* **272**, 2088-97 (2005).
19. Tian, B., Bevilacqua, P.C., Diegelman-Parente, A. & Mathews, M.B. The double-stranded-RNA-binding motif: interference and much more. *Nat Rev Mol Cell Biol* **5**, 1013-23 (2004).

20. Stefl, R., Skrisovska, L. & Allain, F.H. RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. *EMBO Rep* **6**, 33-8 (2005).
21. Yang, E. et al. Decay rates of human mRNAs: correlation with functional characteristics and sequence attributes. *Genome Res* **13**, 1863-72 (2003).
22. Bernstein, J.A., Khodursky, A.B., Lin, P.H., Lin-Chao, S. & Cohen, S.N. Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. *Proc Natl Acad Sci U S A* **99**, 9697-702 (2002).
23. Shalgi, R., Lapidot, M., Shamir, R. & Pilpel, Y. A catalog of stability-associated sequence elements in 3' UTRs of yeast mRNAs. *Genome Biol* **6**, R86 (2005).
24. Xie, X. et al. Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* **434**, 338-45 (2005).
25. Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E.S. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**, 241-54 (2003).
26. Foat, B.C., Houshmandi, S.S., Olivas, W.M. & Bussemaker, H.J. Profiling condition-specific, genome-wide regulation of mRNA stability in yeast. *Proc Natl Acad Sci U S A* **102**, 17675-80 (2005).
27. Elnitski, L., Jin, V.X., Farnham, P.J. & Jones, S.J. Locating mammalian transcription factor binding sites: a survey of computational and experimental techniques. *Genome Res* **16**, 1455-64 (2006).
28. Gerber, A.P., Herschlag, D. & Brown, P.O. Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast. *PLoS Biol* **2**, E79 (2004).
29. Albig, A.R. & Decker, C.J. The target of rapamycin signaling pathway regulates mRNA turnover in the yeast Saccharomyces cerevisiae. *Mol Biol Cell* **12**, 3428-38 (2001).
30. Hilgers, V., Teixeira, D. & Parker, R. Translation-independent inhibition of mRNA deadenylation during stress in Saccharomyces cerevisiae. *Rna* **12**, 1835-45 (2006).
31. Garcia-Martinez, J., Aranda, A. & Perez-Ortin, J.E. Genomic run-on evaluates transcription rates for all yeast genes and identifies gene regulatory mechanisms. *Mol Cell* **15**, 303-13 (2004).
32. Fan, J. et al. Global analysis of stress-regulated mRNA turnover by using cDNA arrays. *Proc Natl Acad Sci U S A* **99**, 10611-6 (2002).
33. Cheadle, C. et al. Control of gene expression during T cell activation: alternate regulation of mRNA transcription and mRNA stability. *BMC Genomics* **6**, 75 (2005).
34. Nonet, M., Scafe, C., Sexton, J. & Young, R. Eucaryotic RNA polymerase conditional mutant that rapidly ceases mRNA synthesis. *Mol Cell Biol* **7**, 1602-11 (1987).
35. Bolstad, B. Low Level Analysis of High-density Oligonucleotide Array Data: Background, Normalization and Summarization. *Dissertation. University of California, Berkeley.* (2004).
36. Jovaag, K.A. Explanation of Median Polish Algorithm. (1998).
37. Benjamini, Y.H., Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. Roy Stat Soc* **57**, 289-300 (1995).

38. Lieb, J.D., Liu, X., Botstein, D. & Brown, P.O. Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nat Genet* **28**, 327-34 (2001).
39. Herrick, D., Parker, R. & Jacobson, A. Identification and comparison of stable and unstable mRNAs in Saccharomyces cerevisiae. *Mol Cell Biol* **10**, 2269-84 (1990).
40. Lotan, R. et al. The RNA polymerase II subunit Rpb4p mediates decay of a specific class of mRNAs. *Genes Dev* **19**, 3004-16 (2005).