

Adaptation as a genome-wide autoregulatory principle in the stress response of yeast

G. De Palo[†], F. Eduati[‡], M. Zampieri[†],
B. Di Camillo[‡], G. Toffolo[‡] and C. Altafini[†] *

[†] SISSA Int. School for Advanced Studies
via Beirut 2-4, 34014 Trieste, Italy

[‡] Department of Information Engineering,
University of Padova
via Gradenigo, 6/B, 35131 Padova, Italy

February 12, 2010

Keywords: Stress/stimulus response, ODE models, adaptation, gene expression time series, integral feedback

Abstract

The gene expression response of yeast to various types of stresses/perturbations shows a common functional and dynamical pattern for the vast majority of genes, characterized by a quick transient peak (affecting primarily short genes) followed by a return to the pre-stimulus level. Kinetically, this process of adaptation following the transient excursion can be modeled using a genome-wide autoregulatory mechanism by means of which yeast aims at maintaining a preferential concentration in its mRNA levels. The resulting feedback system explains well the different time constants observable in the transient response, while being in agreement with all the known experimental dynamical features. For example it suggests that a very rapid transient can be induced also by a slowly varying concentration of the gene products.

1 Introduction

Typically, at the level of gene expression, the response to a stimulus, or to a change in some environmental condition, or even to the substrate composition can be decomposed into a

*Corresponding author: altafini@sissa.it

rapid adaptation phase, occurring with a typical time constant of the order of the tens of minutes [1, 2, 3, 4, 5], superimposed to a long term permanent modification of the gene expression steady state (occurring, for example, at the diauxic shift, [6]). For *S.cerevisiae*, the rapid adaptation described in [3, 7, 8, 1, 5] consists essentially of a transient change in the mRNA concentration followed by a return to the basal pre-stimulus level for almost the entire population of genes. This massive adaptation phenomenon is observed in response to both temporary (such as the glucose pulses of [5]) and permanent (such as the environmental stresses of [3, 8, 1]) stimuli. A simple correlation analysis reveals that the responses to different types of stimuli have consistent similarities and are strongly correlated with the Half Life (HL) of the corresponding genes, as shown in Fig. 1 (a).

The aim of this paper is to propose a kinetic model for the genome-wide rapid adaptation of the transient response to stimuli able to explain the following features:

- (a) the response is highly stereotypical across many stimulations [9];
- (b) the response is to some extent graded (i.e., proportional to the magnitude of the stimulus, see e.g. the heat shock responses of [3]), and reciprocal (a temperature upshift induces a transient pattern which is roughly similar, except for the sign, to that of a temperature downshift) [3];
- (c) robust adaptation is observed to stimuli of various “order” i.e., both vanishing and persistent stimuli are reabsorbed;
- (d) whenever for a gene adaptation is not perfect, the deviation of the new steady state reached from the pre-stimulus one has the same sign of the peak of the transient excursion.
- (e) the transcriptional transient seems to be faster than the changes in other cellular quantities (such as growth rate [4] or protein concentration [7]);
- (f) the rise time of the transient peak is shorter than its decay time (i.e., the time needed to return to the pre-stimulus level);
- (g) the decay time constant is roughly of the same magnitude as the degradation time (HL) of the genes inferred from experimental data [10, 11, 12] although, if transcription is blocked after the onset of the stress response, the upregulated genes seem to degrade faster than expected, while repressed genes seem to degrade slower than expected [13];
- (h) the transient response typically does not induce oscillations (noticeable above the noise level);
- (i) both the maximal amplitude of the transient peak and the area under the transient response are roughly proportional to the HL of the genes, while the peaking time (i.e., instant at which a gene has its maximal excursion during the transient) is not significantly correlated with HL, as shown in Fig. 1 (a);

- (j) a large transient excursion is induced primarily on short genes;
- (k) for what is known, changes in the transcriptional response are maintained at the level of translation [14, 7].

Adaptation, intended as the mechanism by means of which a biological system is able to recover the “optimal” working level of a variable in spite of a persistent stimulus, is common to many biological systems. Examples are numerous: various signal transduction pathways [15], bacterial chemotaxis [16], sensory transduction [17]. This property is typically characterized by means of a negative feedback loop. In the present context, as the gene expression returns to its pre-stimulus level regardless of the amplitude of the stimulus, the system has to have an encoded robust regulatory mechanism as well as a memory of the nominal pre-stimulus concentration value for each gene.

Following [18], a control-theoretic interpretation of adaptation involves an integral feedback loop. In this scheme, the integral of the displacement from a nominal level (i.e., of the error) of a variable is fed back with negative sign. Adaptation is achieved as this variable returns to the nominal level (that is as the error tends to zero) in spite of a persistent stimulus that, in absence of feedback, would alter the steady state value. In the context of the present paper, the variable being integrated is the relative mRNA abundance, and its integral may be taken to represent the relative abundance of the corresponding “gene product”. We assume here that this quantity acts homeostatically on the mRNA transcription rate, reequilibrating the gene expression to the nominal level of concentration. Negative autoregulation of transcription or “autogenous control” [19, 20] is a mechanism that allows to reduce fluctuations around the steady state [21] and to decrease the rise time of a response [22], although it is often invoked for specific transcription factors repressing their own transcription. In our case, we shall assume that negative autoregulation works as a general ubiquitous homeostatic principle, opposing (permanent) changes into the mRNA levels with respect to an “optimal” working concentration for the transcripts, and affecting preferentially the short genes for which protein levels seem to fluctuate more. In the out of equilibrium scenario represented by the transient response, the autoregulatory action is meant to represent homeostasis in both the synthesis and the degradation components of the rate law. As such, the feedback action can modulate the “effective” degradation rate in presence of transcriptional blockage, coherently with data reported in [13].

Autoregulation, in practice, couples the dynamics of a gene and its product. On the dynamical model, this coupling results into a second time constant, which for each gene can be used in the description of the transient response. The aim of this paper is to show that using a negative integral feedback to describe this coupling can explain not only adaptation but also the very rapid transcriptional response to changes in other slower cellular variables (such as the gene products). As a result, even a simple linear ODE model can reproduce all the features listed above, provided that the modes (eigenvalues) of the system and the sign of the input response are appropriately chosen.

2 Results

The experiments here analyzed (a list of the time series used is in the Supplementary Notes) consist of two-channel microarrays in which the mRNA abundance during the transient is hybridized against a basal pre-stimulus mRNA concentration, so that a value approaching 1 (or 0 if a log scale is considered, as in Fig. 1) corresponds to a return to the pre-existing steady state. As can be seen from Fig. 1 (a), in each of the 5 time series of [3] (chosen among those providing a sufficiently fast kinetics, see Supplementary Notes for details), we observe that for almost 90% of the genes, the relative expression ranges within $[-\log_2(1.5), \log_2(1.5)]$ at the end of each transient (the percentage goes up to 95% if we consider an interval of $[-1, 1]$), while during the transient only $\sim 50\%$ of the genes remain inside the interval $[-1, 1]$ on each time series. Hence we can assume that globally the system undergoes a transient excursion in response to each stimulus, and that such an excursion is reabsorbed in a time scale of the order of the hour, meaning that the system has adapted in spite of a persistent stimulation.

Expanding on the concept of autogenous control [19, 20], the basic assumption underlying our model is that an increase of the abundance of a certain protein well above (resp. below) the normal “working” level disfavors (resp. favors) the transcription of the corresponding gene. Under such negative autoregulation, a basic model for transcription and translational kinetics [23, 19, 24, 22], derived in Material and Methods, is the following:

$$\begin{aligned}\frac{dm_i}{dt} &= -\delta_i(m_i - 1) - a_i(p_i - 1) + b_i u \\ \frac{dp_i}{dt} &= -\lambda_i(p_i - 1) + r(m_i - 1),\end{aligned}\tag{1}$$

where: $m_i = \frac{[\text{mRNA}]_i^{\text{red}}}{[\text{mRNA}]_i^{\text{green}}}$ is the mRNA concentration of the i -th gene relative to the basal level, equal to the ratio between the “red” and “green” channels associated to the stimulus response and the basal mRNA level, respectively; p_i is the relative concentration of the corresponding gene product; δ_i and λ_i are degradation rate constants; a_i is the strength of the negative feedback; r is a translational rate constant, assumed equal for all genes. The expressions $m_i - 1$ and $p_i - 1$ are meant to represent displacements from the basal levels in response to a stimulus u whose amplitude and sign (i.e., role as activator or repressor of m_i) are given by b_i . Model (1) assumes linear kinetics, but this does not affect the qualitative conclusions of the study, as shown in the Supplementary Materials. Moreover, model (1) does not account for translational delays, e.g. due to the export and localization of the mRNA, and/or the limited rates of translation initiation and peptidic chain elongation (see Supplementary material for a model with delay).

For the limited time horizon considered here, a couple of hours, the model can be further simplified. In fact, unlike the gene degradation rate δ_i , for which knowledge from several genome-wide datasets is available in the literature [10, 11, 12], knowledge of protein degradation rate λ_i is quite limited at a proteome-wide scale [25], but it is commonly accepted [23, 25] that the protein degradation dynamics are slower (or much slower) than the corresponding mRNA dynamics. For the stress response of yeast we have quantified this difference

using the recent data of [7] consisting of measurements of nine genes and their corresponding proteins in response to redox stress (DTT)¹. By fitting model (1) on these data, all model parameters were precisely estimated (Table 1). In particular it turns out that on average the ratio δ_i/λ_i is 20: while the genes have a HL ($= \ln(2)/\delta_i$) of approximately 35 min, in line with the known degradation time constants [10, 11, 12], the HL of the corresponding proteins is more than 11 hours. This means that in the time horizon considered here (a couple of hours), the contribution of λ_i is totally irrelevant. To confirm this observation, it is worth noticing that the protein time series of [7] show a growing front which is much slower than the corresponding genes, and no decline (i.e., the “transient” growth is not yet exhausted at the end of the recorded time series), see Fig. 1 (b). Thus, disregarding λ_i , the model equations become:

$$\begin{aligned}\frac{dm_i}{dt} &= -\delta_i(m_i - 1) - a_i(p_i - 1) + b_i u \\ \frac{dp_i}{dt} &= r(m_i - 1).\end{aligned}\tag{2}$$

The second equation of (2) can be integrated and the resulting integral, i.e., the area under the mRNA profile computed in the interval $[0, t]$, represents an estimate of the protein abundance $p_i(t)$. When fed back in the first equation of (2) with negative sign, it has the effect of achieving perfect adaptation in m_i , i.e. the mRNA abundance returns exactly to its basal level in spite of a persistent stimulus u , (Fig. 2). A second effect of the negative feedback is to speed up the transient response [22]. In order to understand how this is achieved, consider the state and input matrices A_i and B_i of the linear system (2) (see Material and Methods for details). A_i has two eigenvalues $s_{i,1}$ and $s_{i,2}$, meaning that the effect of coupling the gene m_i with the protein p_i is to introduce a second dynamical mode into the evolution of the system. These two eigenvalues are always stable and, if they are chosen real so that damped oscillations are excluded, the explicit solution for (2) to a step stimulus u is:

$$\begin{aligned}m_i(t) &= 1 + \frac{b_i}{\gamma_i} (e^{s_{i,2}t} - e^{s_{i,1}t}) \\ p_i(t) &= 1 + \frac{b_i r}{\gamma_i} \left(\frac{e^{s_{i,2}t} - 1}{s_{i,2}} - \frac{e^{s_{i,1}t} - 1}{s_{i,1}} \right)\end{aligned}\tag{3}$$

where $\gamma_i = \sqrt{\delta_i^2 - 4ra_i}$ (see Materials and Methods). The two eigenvalues give rise to two exponential modes: if $s_{i,2}$ is of the same order as the natural mRNA degradation time constant ($s_{i,2} \sim \ln(2)/\text{HL}_i$), and $s_{i,1} < s_{i,2}$, then, from the first equation in (3), the fast mode $s_{i,1}$ induces a sharp rising front in the transient but is rapidly exhausted, and is then followed by a more gentle decay to the pre-stimulus level which resembles a typical first order degradation, governed by the slow mode $s_{i,2}$ which is more long-lived. This is exemplified in Fig. 3 for a specific stress-inhibited gene and then extended to all genes: the difference in the

¹The dynamical response to this type of stimulation is known to be slower than for example the heat shock response [3, 26]. However, it is plausible to assume that the ratios δ_i/λ_i are similarly related.

two eigenvalues $s_{i,1}$ and $s_{i,2}$ induces a transient response which well reproduces the observed time courses. In the case of real eigenvalues, the lack of oscillatory behavior implies that $p_i(t)$ is typically monotonic and its displacement from the basal level has sign equal to that of the m_i transient (an effect similar to the “potentiation” described in [14]). Coherently with the experimental data of Fig. 1 (b), the dynamics for p_i are much slower than those of the corresponding m_i (Fig. 3, left). Nevertheless, this slow dynamics is crucial to speed up the transient response of m_i , see Supplementary notes for a further discussion on open-loop time constants. It is worth noticing that the embedding of a fast regulation loop into a slower one is a universal rule of thumb of an engineering control design requiring nested loops, because it minimizes the cross talk between the two loops and therefore also the possibility of spurious dynamical behaviors. Values of eigenvalues shown in Fig. 4 (top left panel) indicate that the fastest mode of A_i (dominating the rising front of the transient) is always much more negative than the slowest mode (dominating the decaying front): $s_{i,1} \ll s_{i,2} < 0$. For all 5 responses of Fig. 1 (a), the time at which the transient gene expression peaks, t_{peak} , is approximately 25 min. In (3), if $t_{\text{peak}} \sim 25\text{min}$, then $e^{s_{i,1}t_{\text{peak}}} < 0.2$ for 87% of the genes (while $e^{s_{i,2}t_{\text{peak}}} < 0.2$ for only 7%) meaning that indeed the transient response declines due to the exhaustion of the fast mode. Since the mean HL is $\approx 25 \pm 15$ min, in the time horizon of the 5 series (the final time points are $t_{\text{end}} = 80, 60, 160, 90, 120$ min for the 5 time courses), the transient has sufficient time to decay back at almost basal level for most genes. In order to evaluate the effect of different stimuli on a single gene, it is of interest to compare the sign of the parameter b_i across the five time series. Our results (Fig. 4 (top right panel)) say that for at least 50% of the genes the sign assignment is unanimous in the 5 series.

The model permits to predict the value of gene product p_i at the end of the observation period t_{end} , from the area under gene expression m_i . To reduce the effect of noise in the mRNA time series, it is convenient to lump together genes whose products form a protein complex (PC). In fact these genes are known to have similar dynamics [12], observation largely confirmed by our analysis. In Fig. 5 areas for the 5 time series of Fig. 1 are shown: it turns out that for most of the neatly up or down-regulated PCs the 5 values have identical sign. Similar results are obtained when instead of PC we consider co-localization on the same metabolic pathway. See Fig. S3 for a similar analysis on the metabolic pathways of the KEGG database. A good agreement is also found across different data sets (Fig. 6) still showing correlation in the PCs and KEGG pathways areas. The comparison between numerically computed areas under the m_i data and model-based estimated areas (p_i) is shown in Fig. 6 (values are averaged over all genes forming a PC). Panels in the main diagonal show the correlation between computed and estimated areas for all datasets which is good, as expected. Not only a strong degree of correlation in the responses to various inhibitory stimuli (such as thermal, oxidative, osmotic, acid stresses) can be reproduced by the model, but also the anticorrelation between responses to inhibitory and excitatory stimuli (such as the reciprocal stresses discussed in [3] or the nutrient inputs of [5]), see last row and column of Fig. 6.

The basic autoregulation mechanisms discussed so far neglects any gene-gene regulatory mechanism beyond co-participation in a protein complex or metabolic pathway. A popular

example of such regulatory mechanisms is the causal relationship between a transcription factor and its target genes. Having only mRNA profiles available, the only statistical test we can perform to evaluate this type of regulation is a significance analysis of the corresponding correlation coefficients on the transcriptional regulatory map of [27, 28]. For all of our time series, however, the correlation between transcription factors and corresponding target genes is always insignificant (Z-score test), see Fig. S2, Table S1 and the Supplementary Notes for more details. This suggests that the transient excursion must be triggered by post-transcriptional or post-translational modifications of the transcription factors which for the time being are largely unknown. However, even when we look at genes co-transcribed by the same transcription factor, we obtain that the correlation is still comparable to that of a random choice of genes, unlike for example the co-participation in a PC, see again Fig. S2. The approach of modeling the stress response in an “open-loop” fashion as the causal action of the transcription factors on their target genes is pursued for example in [28]. It has several drawbacks, like for example that the map of transcription factor-target genes, though largely incomplete, is already combinatorially complex, condition-specific [27, 28, 9], and the sign of the interactions (activator/repressor) is often unavailable. More importantly for us, this open-loop approach is unable to satisfy all of the kinetic constraints on the time series, like the sharpness of the rise front of the transient, and does not provide an explanation for the adaptation observed for which a form of feedback is required.

3 Discussion

In this work we have proposed a kinetic model aimed at describing different features characteristic of yeast transient response to stimuli, as listed in the Introduction. Here we discuss these features versus model behaviour in details.

Stereotypical response ((a) and (b)) Analysis of the 5 time series of Fig. 1 confirms that the similarities in their pattern are much more abundant than the stimulus-specific differences. This can be deduced from the sign concordances of the b_i values, see Fig. 4 (top right panel). If the 5 time series are compared with others from [8, 1, 29] also representing responses to prolonged stimuli, the pattern of up/down regulation is very similar. For the same PC as in Fig. 5, the comparison of average areas is shown in Fig. 6. In [5] instead, yeast is fed with pulses of glucose of different magnitude. In this case the sign of the responses is reciprocal for most genes, as can be seen in the scatter plots in the last row and column of Fig. 6 (see Supplementary Material). Sorting the PC complexes by the corresponding $p_i(t_{\text{end}})$, see Fig. 5, reveals that the downregulated categories (negative areas in the log scale of Fig. 5) are essentially all involved into transcriptional and translational processes, while in the most upregulated categories are respiratory metabolism and proteolysis. Notice that, coherently, also the ribosomal biogenesis is very different between the cytoplasmic and mitochondrial compartments. Analogous results are obtained for the KEGG pathways, see Fig. S2. Grouping the genes further according to KEGG hierarchy, we obtain the 15 macrocategories shown in Fig. S3 which give a general overview of the environmental stress

response strategy in agreement with e.g. [3, 9]. This consists in a reduction of the energy-consuming (cytoplasmic) ribosomal biosynthesis and RNA processing machinery in favor of energy-producing components such as the respiratory chain complexes and the mitochondrial compartment in general.

Notice on Fig. 5 and S3 the correlation with the empirical values of HL assignable to these categories. It is worth observing how the ordering of the categories found here resembles e.g. the ordering of the phases of the peaks in the so-called yeast metabolic cycle [30], suggesting the unfolding of a common gene expression program. See Fig. S4 and [31] for more details.

Robust adaptation ((c) and (d)) While a few genes required for reacting to a specific cellular stress might show a permanent change in the gene expression steady state [9], the vast majority of genes returns to their basal pre-stimulus level. It is for this category that we talk about adaptation. The rapid and massive transient excursion characterizing stress responses might help in activating immediate cellular reaction mechanisms (such as redistribution of energetic resources) while adaptation might be a mean to resume a mode of action as close as possible to optimal in spite of permanent environmental changes. The slower degradation rates for the gene products observed in real data [7] and confirmed by our observation (Table 1) guarantee that the proteins are long-lived and that the modified cellular response is sustainable for a long period. As sketched in Fig. 2 (d), when the protein degradation is not neglected, its effect on the model is to alter the steady state value of m_i . The sign of this modification agrees with that of the transient excursion.

Fast transcriptional response from slow feedback ((e) and (f)) The transcriptional response to stresses can be activated directly by the external perturbation through signaling mechanisms in an essentially open-loop fashion, or through changes in the cellular state (e.g. amount of biomass or energy or metabolite composition) that induce feedback reactions [4]. Some of these “internal variables”, such as growth rate, have been shown to happen at a slower pace than the transcriptional response [4]. One of the characteristics of our model is that slower dynamics are instrumental in inducing the fast transcriptional response, provided that there is coupling between quantities as happens in presence of feedback. While a feedback coupling between genes (faster) and gene products (slower) can optimize the stress recovery by speeding up the system response, the signs of the transcriptional transient excursions are instead directly correlated with the external perturbation and determine the strategy of the cellular response. Once the signature of the transient is identified, any cellular variable (for example the already mentioned growth rate) could in principle be used for the integral feedback in place of the gene products we use here. The fact that most genes show a synchronized peaking time for their transient may be a sign that a coordinated cellular mechanism is indeed responsible for the feedback. Experimental data for cellular quantities such as the growth rate are however too few and difficult to obtain [32]. See Supplementary Notes for more details and examples.

Stress response with transcriptional blockage ((g)) In [13], the stress response is studied in conjunction with a blockage of transcription (delayed in time with respect to the begin of the stressful stimulation). The mRNA profiling reveals that genes activated during the transient response seem to be destabilized when the stress is followed by the transcriptional arrest (i.e., the genes seem to degrade faster than expected by the known HL values) and, viceversa, genes that are repressed in the transient seem to be stabilized by the combination stress + transcriptional blockage. In our model, we assume that the transcriptional arrest occurs at the peaking time of the transient, t_{peak} . Blocking transcription means putting $c = 0$ in (5) and $b_i = 0$ in (2), i.e., the ODE for the system reduces to

$$\begin{aligned}\frac{dm_i}{dt} &= -\delta_i m_i - a_i(p_i - 1) \\ \frac{dp_i}{dt} &= r(m_i - 1).\end{aligned}$$

In our scheme, the negative autoregulatory feedback term has a dual role, influencing both the synthesis and the degradation rate, and predicts correctly the altered degradation rates in the perturbed system of [13]. Consider first the case of an upregulated gene. For it $m_i(t_{\text{peak}}) > 1$ and, from (2), $p_i(t_{\text{peak}}) > 1$, implying $-a_i(p_i(t_{\text{peak}}) - 1) < 0$. Hence we have

$$\underbrace{-\delta_i m_i - a_i(p_i - 1)}_{\text{degradation with negative autoregulation}} < \underbrace{-\delta_i m_i}_{\text{reference degradation}} < 0$$

i.e., when p_i is different from the nominal concentration ($p_i \neq 1$) the rate dm_i/dt is more negative than expected, meaning that the upregulated gene is destabilized. On the contrary, for a repressed gene we have $0 < m_i(t_{\text{peak}}) < 1$ and $0 < p_i(t_{\text{peak}}) < 1$, implying $-a_i(p_i(t_{\text{peak}}) - 1) > 0$, which leads to

$$\underbrace{-\delta_i m_i}_{\text{reference degradation}} < \underbrace{-\delta_i m_i - a_i(p_i - 1)}_{\text{degradation with negative autoregulation}} < 0.$$

In this case, the repressed gene is stabilized by the autoregulation.

Proportionality of transient peak amplitude and area of m_i with HL ((i)) From (3) we can obtain the following expression:

$$p_i(t_{\text{end}}) - 1 \simeq \frac{b_i r}{\gamma_i} \left(\frac{1}{s_{i,1}} - \frac{1}{s_{i,2}} \right) \simeq -\frac{b_i r}{\gamma_i s_{i,2}} = \frac{b_i r}{\gamma_i} \text{HL}_i.$$

This expression provides an explanation in terms of the model (2) of the roughly direct proportionality observed between the area and the values of HL, shown in Fig. 4 (bottom right panel) for the PCs of Fig. 5. Notice on the same Fig. 4 (bottom left panel) the tight relationship between the amplitude (i.e., the signed peak in m_i) and the area (computed either via the model or from the data). From (3), model-based area and amplitude share the same gene-specific multiplicative constant b_i/γ_i .

Transiently perturbed genes are short ((j)) The transient excursion is induced (for both upregulated and repressed genes) to a very large extent on short genes: 85% of the genes labeled as transiently perturbed ($|\log_2(p_i)| > 0.5$) have length ≤ 2 kbp (p-value 10^{-5} , hypergeometric test; genes with ORF (acronym for Open Reading Frame, i.e., gene-encoding DNA sequences) length ≤ 2 kbp form $\sim 70\%$ of the total), see Fig. 7 (a). The fact that also downregulated genes are relatively short excludes the scenario in which a marked transient excursion is only the consequence of an increased synthesis rate affecting more the short ORF than the long ones. It is interesting to compare the behavior observed on the transient with the average absolute abundance of the corresponding proteins, as estimated in [33] for non-stressed yeast. From Fig. 7 (b) (right lower plot), the induced genes (again, both up- and down-regulated) seem also to correspond to gene products having a low concentration in the “ordinary, stationary” conditions of [33] (p-value 0.05, t-test). The short length of the mRNAs certainly favours more rapid fluctuations which could induce more easily changes also at the level of gene products. More marked changes in protein abundances favour the feedback regulation we are hypothesizing. The model (2) does not explicitly include the length of a gene in its parameters. However, as can be seen on the lower left plot of Fig. 7 (b), it tends to associate to a consistent fraction of short genes a high value of the forcing parameter b_i , meaning that the impact of the stimulation on the kinetics on these genes is more pronounced.

4 Conclusions

Yeast reacts to a change of environmental conditions by means of a highly coordinated transcriptional response which is faster than it would be expected from the “natural” degradation time constant but which is only transient. In this paper we propose a model able to explain this quick response by means of a feedback mechanism aiming at adapting the system to the new condition. From a dynamical point of view, this can be formulated in terms of a second mode, faster than degradation, which dominates the transient excursion but which, being quickly exhausted, is not observable on standard turnover experimental curves. It is shown that this second mode can be induced by a feedback mechanism from a much slower dynamical variable, which could correspond to the concentration of gene products.

5 Materials and Methods

Model construction The changes in the relative concentration of mRNA with respect to its basal level can be described with a typical model for the transcription kinetics [23, 34, 35, 36, 5, 2, 37]:

$$\frac{dm_i}{dt} = -\delta_i m_i + f_i, \quad (4)$$

where the function f_i describes the transcription synthesis rate for the i -th gene and is usually zero-order in m_i , i.e., independent of the concentration of m_i . In the literature, f_i

is often expressed as a function of the transcription factor(s) w_i governing the expression of m_i , with various types of functional dependence like linear, Michaelis-Menten or of Hill type [35, 36, 38, 5, 2, 39], see Supplementary Notes for examples. Following this approach requires the knowledge of the transcription factor w_i acting on each gene. Even if this information is partially available for *S.cerevisiae* [27, 28], predicting the kinetics of the transient response from them is troublesome for the reasons explained at the end of Section 2, and also because in the literature the kinetic models mentioned above are mostly used for describing variations in the steady state following a perturbation, not for the transient dynamics itself. Moreover, under the assumption that the transcription synthesis rate f_i is of zero-order in m_i [35], the fast rising front of the transient cannot be explained in terms of a model like (4) at least for reasonable values of the degradation time constants δ_i (Fig. 2, see Supplementary Notes for a thorough analysis). On top of all these complications, modeling the effect of an external stimulus u on the transcriptional regulation means expressing w_i as a function of u . Nothing is known in general about this further functional dependence $w_i = w_i(u)$. Bypassing the transcription factors, the $f(u)$ can for example be represented as an open-loop impulse like in [26] or, more generally, as a finite width kernel, vanishing after some time, see Fig. 2 (a). As the m_i represent relative concentrations, these open loop models entail (without explicitey explaining) a form of memory of the “ideal” pre-stimulus absolute concentration, as well as a form of adaptation if one considers the stimulation u as a step (e.g. a permanent increase in temperature). Both elements are characterized in our model by means of a feedback term. In absence of such feedback, the transcription synthesis rate f_i consists for us only of a basal (constant) term plus a term linear in the stimulus u , of the form of a zero order kinetics in m_i :

$$\frac{dm_i}{dt} = -\delta_i m_i + c_i + b_i u. \quad (5)$$

The parameter c_i corresponds to the basal rate of transcription in absence of external stimuli ($u = 0$). Therefore, since for the unperturbed system the steady state must be $\bar{m}_i = c_i/\delta_i = 1$, we have $c_i = \delta_i$. The parameter b_i instead carries information about the activator/inhibitor effect of u on the mRNA concentration. When u is a persistent stimulus, e.g. $u(t) = 1, t \geq 0$, then in (5) the steady state value is modified to $m_i = (c_i + b_i u)/\delta_i = 1 + b_i u/\delta_i \neq 1$, i.e., the system (5) is not adapted to step-like inputs u and cannot recover its pre-stimulus mRNA level.

An increase in the transcription rate of the i -th gene induces an increase in the total quantity of mRNA produced over time

$$p_i(t) - \bar{p}_i = r \int_0^t (m_i(\tau) - \bar{m}_i) d\tau = r \int_0^t (m_i(\tau) - 1) d\tau \quad (6)$$

where, as above, $\bar{m}_i = 1$ is the pre-stimulus relative mRNA abundance, r is a rate constant (representing for example the ribosome density, and assumed to be the same for all genes) and \bar{p}_i is an integration constant (representing the basal level of p_i , see below). Differentiating this expression,

$$\frac{dp_i}{dt} = r(m_i - 1), \quad (7)$$

we see that the variable p_i represents a dynamical quantity “downstream” of transcription. In this paper p_i is taken to describe the concentration of the corresponding gene product relative to the basal level, hereafter fixed as $\bar{p}_i = 1$. This (very common [23, 25, 24]) choice is a simplification of the complex mechanisms characterizing translation and protein synthesis, involving for example changes in the translation initiation, in the ribosomal density or in the polysomal association [40, 14], all steps not well-characterized dynamically. From what is known experimentally, the dynamics at the polysomes level for example seems to be correlated with the transcriptional perturbation of the mRNAs (in [14] it is shown that the frequency of association with polysomes increases for upregulated genes and decreases for downregulated genes). An ODE like (7) for a gene product usually contains a degradation term. Given that the transcriptional perturbation propagates through the protein synthesis process with a time delay and that the protein turnover rate is typically considered slower than the corresponding mRNA turnover rate [23, 25], the influence of the protein degradation term on the dynamics becomes negligible for the time horizon of interest here, see Fig. 1 (b).

The homeostatic effect assumed in the paper consists of a feedback autoregulation acting in correspondence of a displacement from the basal level (i.e., for $p_i \neq 1$) and can be modeled as in the system (2). The model (2) predicts that the equilibrium is reached for p_i corresponding to $\tilde{p}_i = 1 + b/a$. In order to have $\tilde{p}_i > 0$, the parameters must therefore satisfy the consistency condition $b_i > -a_i$. *De facto*, the amplitude of p_i depends on the rate constant r . For all time series considered, a choice of $r = 0.01$ (motivated by the experimental data rather than by the dynamical model chosen, see Supplementary Notes) is sufficient to have biologically consistent values of \tilde{p}_i for the range of a_i, b_i required by the fitting procedure.

Since the model misses a degradation term in p_i , the protein concentration changes in response to the persistent stimulus from \bar{p}_i to \tilde{p}_i without ever returning to the basal level. Introducing such a term as in (1) typically leads to only minor differences, although exact adaptation in m_i and monotonicity in p_i are lost. For sufficiently high ratios of δ_i/λ_i (~ 5 or larger), the differences with respect to (2) are minimal, and we can talk about “quasi-adaptation” and of an autoregulatory feedback which behaves like a “leaky” integral, see Fig. 2 (d). The system still has two modes with distinct time constants and the dominant mode still affects primarily only the rising front of m_i .

Model identification and analysis To simplify calculations, it is convenient to change variables, shifting the steady state to the origin. Letting $x_i = \begin{bmatrix} m_i - 1 \\ p_i - 1 \end{bmatrix}$, and denoting the state and input matrices for the systems as

$$A_i = \begin{bmatrix} -\delta_i & -a_i \\ r & 0 \end{bmatrix}, \quad \text{and} \quad B_i = \begin{bmatrix} b_i \\ 0 \end{bmatrix}, \quad (8)$$

then for each gene we have the linear system (with input)

$$\dot{x}_i = A_i x_i + B_i u \quad (9)$$

whose solution for the step response is

$$x_i(t) = e^{A_i t} x_i(0) + \int_0^t e^{A_i(t-\tau)} B_i u(\tau) d\tau. \quad (10)$$

Since $\text{tr}(A_i) = -\delta_i < 0$ and $\det(A_i) = r a_i > 0$, the system is always stable and its eigenvalues are:

$$s_{i,1} = -\frac{\delta_i}{2} - \frac{\gamma_i}{2}, \quad \text{and} \quad s_{i,2} = -\frac{\delta_i}{2} + \frac{\gamma_i}{2} \quad (11)$$

where $\gamma_i = \sqrt{\delta_i^2 - 4r a_i}$. A visual inspection of the time series shows that for the vast majority of genes the large excursion corresponding to the transient is damped without inducing oscillatory behavior (at least above what can be considered measurement noise). Hence in the model fitting we assumed:

1. the two eigenvalues are real, i.e., $\delta_i^2 - 4r a_i > 0$;
2. the time constant of the fastest eigenvalue is shorter than that of the “free degradation” given by the HL alone.

The two conditions are compatible with each other and with the model structure. In order to agree also with the available HL measures, we shall assume the following:

$$s_{i,1} < s_{i,2} \sim -\frac{\ln(2)}{\text{HL}_i} < 0.$$

If we choose $s_{i,2} = -\frac{\ln(2)}{\text{HL}_i}$, then we obtain the following conditions:

$$\begin{cases} a_i = -s_{i,2}(\delta_i + s_{i,2})/r > 0 \\ \delta_i > -2s_{i,2} > 0 \end{cases} \quad (12)$$

In correspondence of a persistent stimulus, $u(t) = 1$ for $t \geq 0$, the system (10) can be solved explicitly. Since at $t = 0$ the system is at rest (i.e., in the basal state $x_i(0) = [0 \ 0]^T$ for all i , corresponding to $m_i(0) = 1$ and $p_i(0) = 1$), only the forced evolution (second term in (10)) matters and we obtain:

$$x_i(t) = \frac{b_i}{\gamma_i} \left[r \begin{pmatrix} e^{s_{i,2}t} - e^{s_{i,1}t} \\ \frac{e^{s_{i,2}t} - 1}{s_{i,2}} - \frac{e^{s_{i,1}t} - 1}{s_{i,1}} \end{pmatrix} \right] \quad (13)$$

i.e., Equation (3) for $m_i(t)$ and $p_i(t)$.

Notice that as $t \rightarrow \infty$ from the second equation of (3) we obtain that $p_i(t) > 0$ if $r b_i (1/s_{i,1} - 1/s_{i,2})/\gamma_i > -1$, i.e., for $b_i > -a_i$ as mentioned above.

The first equation of (13) can be used to fit the parameters in the dynamical model (9). For each gene, this corresponds to identifying the values of δ_i and b_i that optimize the fit of $m_i(t)$ to the experimental time series. With these parameters, (9) is completely determined. The second equation of (13) can then be used to compare the area predicted by the model with the area computed from the experimental data.

References

- [1] H C Causton, B Ren, S S Koh, C T Harbison, E Kanin, E G Jennings, T I Lee, H L True, E S Lander, and R A Young. Remodeling of yeast genome expression in response to environmental changes. *Mol Biol Cell*, 12(2):323–337, 2001.
- [2] Barrett C. Foat, S. Sean Houshmandi, Wendy M. Olivas, and Harmen J. Bussemaker. Profiling condition-specific, genome-wide regulation of mRNA stability in yeast. *Proc. Natl. Acad. Sci. U.S.A.*, 102(49):17675–17680, 2005.
- [3] Audrey P Gasch, Paul T Spellman, Camilla M Kao, Orna Carmel-Harel, Michael B Eisen, Gisela Storz, David Botstein, and Patrick O Brown. Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell*, 11(12):4241–4257, 2000.
- [4] S Levy, J Ihmels, M Carmi, A Weinberger, G Friedlander, and N Barkai. Strategy of transcription regulation in the budding yeast. *PLoS ONE*, 2(2), 2007.
- [5] Michal Ronen and David Botstein. Transcriptional response of steady-state yeast cultures to transient perturbations in carbon source. *Proc. Natl. Acad. Sci. U.S.A.*, 103(2):389–394, 2006.
- [6] Joseph L DeRisi, Vishwanath R Iyer, and Patrick O Brown. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science*, 278(5338):680–686, 1997.
- [7] Gal Chechik, Eugene Oh, Oliver Rando, Jonathan Weissman, Aviv Regev, and Daphne Koller. Activity motifs reveal principles of timing in transcriptional control of the yeast metabolic network. *Nat. Biotechnol.*, 26(11):1251–1259, 2008.
- [8] H Yoshimoto, K Saltsman, A P Gasch, H X Li, N Ogawa, D Botstein, P O Brown, and M S Cyert. Genome-wide analysis of gene expression regulated by the calcineurin/crz1p signaling pathway in *saccharomyces cerevisiae*. *J Biol Chem*, 277(34):31079–31088, 2002.
- [9] A. P. Gasch. The environmental stress response: a common yeast response to environmental stresses. In S. Hohmann and P. Mager, editors, *Yeast Stress Responses*, pages 11–70. Springer-Verlag, Heidelberg, 2002.
- [10] Jörg Grigull, Sanie Mnaimneh, Jeffrey Pootoolal, Mark D. Robinson, and Timothy R. Hughes. Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors. *Mol. Cell. Biol.*, 24(12):5534–5547, 2004.
- [11] Letian Kuai, Biswadip Das, and Fred Sherman. A nuclear degradation pathway controls the abundance of normal mRNAs in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.*, 102(39):13962–13967, 2005.
- [12] Yulei Wang, Chih Long Liu, John D. Storey, Robert J. Tibshirani, Daniel Herschlag, and Patrick O. Brown. Precision and functional specificity in mRNA decay. *Proc. Natl. Acad. Sci. U.S.A.*, 99(9):5860–5865, 2002.
- [13] O Shalem, O Dahan, M Levo, M R Martinez, I Furman, E Segal, and Y Pilpel. Transient transcriptional responses to stress are generated by opposing effects of mRNA production and degradation. *Mol Syst Biol*, 4:223–223, 2008.

- [14] Thomas Preiss, Julie Baron-Benhamou, Wilhelm Ansorge, and Matthias W Hentze. Homodirectional changes in transcriptome composition and mRNA translation induced by rapamycin and heat shock. *Nat. Struct. Biol.*, 10(12):1039–1047, 2003.
- [15] M Behar, N Hao, H G Dohlman, and T C Elston. Mathematical and computational analysis of adaptation via feedback inhibition in signal transduction pathways. *Biophys J*, 93(3):806–821, 2007.
- [16] U Alon, M G Surette, N Barkai, and S Leibler. Robustness in bacterial chemotaxis. *Nature*, 397(6715):168–171, 1999.
- [17] V Torre, J F Ashmore, T D Lamb, and A Menini. Transduction and adaptation in sensory receptor cells. *J Neurosci*, 15(12):7757–7768, 1995.
- [18] T M Yi, Y Huang, M I Simon, and J Doyle. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc. Natl. Acad. Sci. U.S.A.*, 97(9):4649–4653, 2000.
- [19] M. A. Savageau. Comparison of classical and autogenous systems of regulation in inducible operons. *Nature*, 252:546–549, 1974.
- [20] R. F. Goldberger. Autogenous regulation of gene expression. *Science*, 183:810 – 816, 1974.
- [21] A Becskei and L Serrano. Engineering stability in gene networks by autoregulation. *Nature*, 405(6786):590–593, 2000.
- [22] N Rosenfeld, M B Elowitz, and U Alon. Negative autoregulation speeds the response times of transcription networks. *J Mol Biol*, 323(5):785–793, 2002.
- [23] J I Hargrove and F H Schmidt. The role of mRNA and protein stability in gene expression. *FASEB J.*, 3(12):2360–2370, 1989.
- [24] M L Simpson, C D Cox, and G S Saylor. Frequency domain analysis of noise in autoregulated gene circuits. *Proc. Natl. Acad. Sci. U.S.A.*, 100(8):4551–4556, 2003.
- [25] Archana Belle, Amos Tanay, Ledion Bitincka, Ron Shamir, and Erin K O’Shea. Quantification of protein half-lives in the budding yeast proteome. *Proc. Natl. Acad. Sci. U.S.A.*, 103(35):13004–13009, 2006.
- [26] G. Chechik and D. Koller. Timing of gene expression responses to environmental changes. *Journal of Computational Biology*, 16(2):279–290, February 2009.
- [27] Nicholas M. Luscombe, M. Madan Babu, Haiyuan Yu, Michael Snyder, Sarah A. Teichmann, and Mark Gerstein. Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature*, 431:308–312, 2004.
- [28] Y Pilpel, P Sudarsanam, and G M Church. Identifying regulatory networks by combinatorial analysis of promoter elements. *Nat Genet*, 29(2):153–159, 2001.
- [29] I Tirosh, A Weinberger, M Carmi, and N Barkai. A genetic signature of interspecies variations in gene expression. *Nat. Genet.*, 38(7):830–834, 2006.

- [30] Benjamin P Tu, Andrzej Kudlicki, Maga Rowicka, and Steven L McKnight. Logic of the yeast metabolic cycle: Temporal compartmentalization of cellular processes. *Science*, 310(5751):1152–1158, 2005.
- [31] Nicola Soranzo, Mattia Zampieri, Lorenzo Farina, and Claudio Altafini. mRNA stability and the unfolding of gene expression in the long-period yeast metabolic cycle. *BMC Systems Biology*, 3:18, 2009.
- [32] E M Airoidi, C Huttenhower, D Gresham, C Lu, A A Caudy, M J Dunham, J R Broach, D Botstein, and O G Troyanskaya. Predicting cellular growth from gene expression signatures. *PLoS Comput Biol*, 5(1), Jan 2009.
- [33] S Ghaemmaghami, W K Huh, K Bower, R W Howson, A Belle, N Dephoure, E K O’Shea, and J S Weissman. Global analysis of protein expression in yeast. *Nature*, 425(6959):737–741, Oct 2003.
- [34] U. Alon. *An Introduction to Systems Biology - Design Principles of Biological Circuits*. Chapman & Hall/CRC;, 2006.
- [35] Paula M. Alepuz José E. Pérez-Ortín and Joaquín Moreno. Genomics and gene transcription kinetics in yeast. *Trends in Genetics*, 23(5):250–257, 2007.
- [36] M Ronen, R Rosenberg, B I Shraiman, and U Alon. Assigning numbers to the arrows: parameterizing a gene regulation network by using accurate expression kinetics. *Proc Natl Acad Sci U S A*, 99(16):10555–10560, 2002.
- [37] L Farina, A De Santis, S Salvucci, G Morelli, and I Ruberti. Embedding mrna stability in correlation analysis of time-series gene expression data. *PLoS Comput Biol*, 4(8), 2008.
- [38] R Khanin, V Vinciotti, and E Wit. Reconstructing repressor protein levels from expression of gene targets in Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.*, 103(49):18592–18596, 2006.
- [39] N E Buchler, U Gerland, and T Hwa. Nonlinear protein degradation and the function of genetic circuits. *Proc. Natl. Acad. Sci. U.S.A.*, 102(27):9559–9564, 2005.
- [40] K M Kuhn, J L DeRisi, P O Brown, and P Sarnow. Global and specific translational regulation in the genomic response of saccharomyces cerevisiae to a rapid transfer from a fermentable to a nonfermentable carbon source. *Mol. Cell. Biol.*, 21(3):916–927, 2001.

Figure 1: In (a) five time series from [3] showing the “step response” to different environmental stresses are plotted. For visualization purposes, the 5 series are shown sequentially one after the other. The time axis is in minutes, the relative mRNA abundance is in \log_2 basis. In the left panels, the 5153 genes are clustered in five groups according to the respective HL (in min). In the right panels the averages of the profiles in each cluster are computed. In all responses the trend followed during the transient is highly correlated, i.e. genes with similar HL behave similarly in the various responses. In particular, genes with short HL tend to be downregulated while genes with longer HL upregulated. In (b) the linear model (2) is fitted to the experimental time series of a pair gene/protein in response to treatment with DTT (data from [7], scales are normalized). For this class of redox stimulations, the time constants of the response are higher than in (a). Nevertheless, as predicted by the model, the protein time series resembles the integral of the gene time series, which corresponds to the adaptation scenario by means of integral feedback autoregulation. Parameter values estimated by fitting model (2) to these data, as well as to those of the other eight genes, are listed in Table 1.

Figure 2: Four different schemes for the step response in (4). (a): open-loop scheme with nonvanishing/vanishing transcriptional synthesis term $f(\cdot)$. When $f(w(u)) = f(w)u$ with $f(w)$ a Michaelis-Menten (MM) function of the transcription factors w , see Supplementary Notes for details, then the dotted red line is obtained for m_i . When instead $f(w(u))$ is a vanishing function then the response also vanishes (shown in red, solid). This model is equivalent to the so-called impulse model of [26]. However, if we think in absolute (rather than relative) terms, it entails an exact knowledge of the nominal level (hence, implicitly, a form of memory, like the one obtained here by means of integral feedback). Notice further how for a synthesis term f which is zero-order in m_i , the rising front has a limited slope regardless of the form of f (see Supplementary Notes for a more detailed analysis of open-loop time constants). (b): a regulation scheme with a proportional feedback, i.e., a feedback directly on the relative mRNA abundance. Adaptation is not achieved, rather, the mRNA level tends to “track” exactly the amplitude of the stimulus u . (c): regulation with integral feedback. Adaptation is achieved for any value of nominal concentration. Both feedback schemes (b) and (c) decrease the rise time of the response. In (c) this is achieved via the much slower dynamical variable p_i . In this case $p_i(t)$ represents the area under the $m_i(t)$, computed in the interval $[0, t]$. (d): “Quasi-adaptation” in presence of protein degradation terms. When a protein degradation term is added to the equations as in (1), then perfect adaptation is lost. However, for reasonable values of protein degradation rates, the new steady state is still close enough to the full recovery of the pre-stimulus level and the shape of the transient is essentially unchanged. Hence we can talk about “quasi-adaptation”. Notice how in this case the variation in the steady state has the same sign of the transient excursion.

Figure 3: Time response of the system (2) to a step-like input for the category of cytoplasmatic transcription initiation genes and for gene GCD2/YGR083C (subunit of the translation initiation factor eIF2B) in particular. The two modes have different real parts ($s_{i,1} = -0.11 < s_{i,2} = -0.04$), thus their difference typically shows a profile like that reproduced in the middle left plot. The sign of b_i then determines whether the gene is classified as up- or down-regulated by the stimulus (still middle plot). The area under the $m_i(t)$ time course, proportional to the gene product $p_i(t)$ shown in the bottom left plot, is monotonically growing with a much slower time constant, as expected. For the gene considered here, the experimental and reconstructed profiles are shown in the top right panel (blue and red respectively, both in \log_2 scale) while profiles and model-based predictions of the entire category of cytoplasmic transcription initiation genes are shown in the middle and bottom plots of the right column, respectively.

Figure 4: Relationship between eigenvalues $s_{i,1}$ and $s_{i,2}$ for all genes in the five time series of Fig. 1 (top left panel). Sum of the signs of the forcing term b_i (top right panel) in the five time series for all genes and for the most perturbed genes, selected according to a threshold k_p , i.e. satisfying $\max |log_2(m_i(t))| > k_p$. Notice that, choosing $k_p = 1$, for at least 50% of the genes the sign assignement is unanimous in the 5 series of [3] (more than 60% for $k_p = 1.5$). In the bottom row, for the PC complexes of Fig. 5 the area under the mRNA response measured on the data (i.e., in the log scale, $\log_2(p_i(t_{\text{end}}))$) is compared with the corresponding maximal signed amplitude observed during the transient (left) and with the HL (right). The high agreement between area and the sign of the peak of mRNA during the transient is confirming that most transient excursions are not oscillatory.

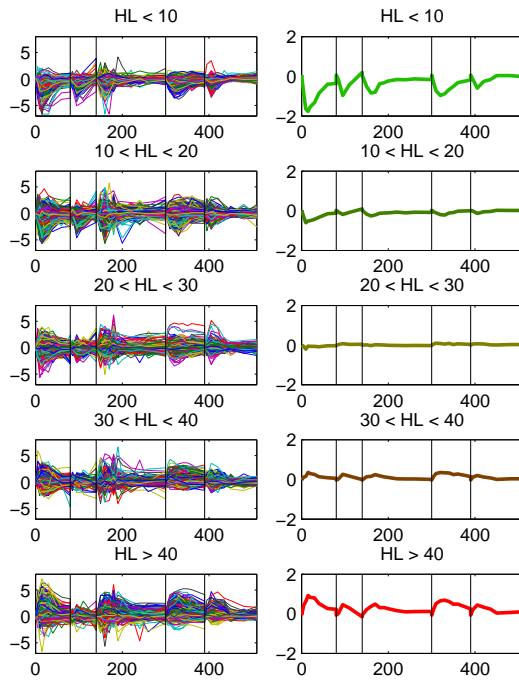
Figure 5: List of significant PC and corresponding areas (i.e., $\log_2(p_i(t_{\text{end}}))$ in the log scale used) for the 5 time series of Fig. 1 (a). The solid markers, representing the values in the 5 experiments, indicate that for most of the neatly up or down-regulated categories, the 5 values have identical signs. The color scale represents the value of HL associated to the protein complex: blue means short HL (minimum HL is 6 min) and red long HL (≥ 40 min).

Figure 6: Comparison of the average area under the curve for the time series of [3] (labeled “gasch”), [8] (“yoshimoto”), [1] (“causton”), [29] (“tirosh”) and [5] (“ronen”). In the upper triangular part, in blue, the scatter plots represent the area of one set of data against any of the other sets for the PC complexes of Fig. 5. In the bottom triangular part, in red, the scatter plots are for the KEGG pathways represented in Fig. S2. While “gasch”, “yoshimoto”, “causton” and “tirosh” are inhibitory stimuli (stresses), “ronen” are activatory pulses of nutrient. Hence the antidiagonal pattern in the areas shown in the bottom row and column. In the diagonal plots, the area predicted by the model (with parameters tuned on the 5 series of [3]) is shown against the corresponding measured area for PC (in green) and for the KEGG pathways (in magenta). In the simulation of the activatory stimuli of [5], the signs of the b_i are exchanged. This, together with the anticorrelated plots of the bottom row and column, validates the reciprocity property already observed in [3] for some classes of stresses.

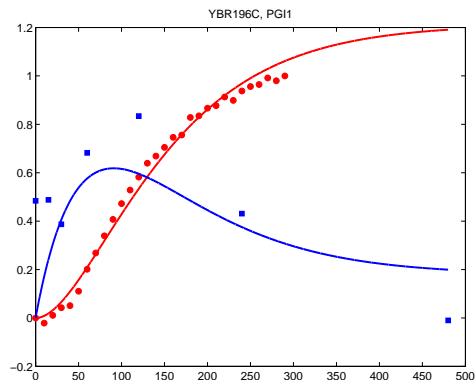
Figure 7: In (a), the average area under the transient response for the 5 time series is plotted against the length of the corresponding ORFs. Longer ORFs clearly tend to be perturbed less, while the genes significantly perturbed (both up- and down-regulated) are for the vast majority shorter than 2kbp (shown in red). There seems to be some degree of inverse correlation also between ORF length and (absolute) protein abundances estimated in [33] in ordinary (unperturbed) growth conditions, see (b) (top right), with, in particular, the really abundant proteins corresponding to short ORFs. Likewise, the correlations of both ORF lengths and protein abundances with HL seem to be to some extent skewed, with long lived mRNAs corresponding to short genes. The color code in (b) is the same as (a). From it (in particular the two plots on the right) we can deduce that most genes perturbed during the transient stress response (red dots) correspond to products having low/medium abundances. For a significant part of the short genes, the model identification returns a high value of the forcing parameter b_i (bottom left plot), indicating that their amplified response is attributed to an higher sensitivity to the stressful stimulation.

Table 1: Gene and protein parameters estimated by fitting model (1) to the time series of [7]. In this case, the translational rate constant r is also considered a gene specific parameter. The average value obtained for r corresponds to our choice in the rest of the paper.

| ORF | a_i | b_i | δ_i | λ_i | r |
|---------|-------|-------|------------|-------------|-------|
| YBR001C | 0.008 | 0.011 | 0.016 | 0.000 | 0.008 |
| YBR196C | 0.012 | 0.019 | 0.023 | 0.001 | 0.010 |
| YDR074W | 0.005 | 0.010 | 0.013 | 0.000 | 0.008 |
| YDR261C | 0.001 | 0.004 | 0.006 | 0.000 | 0.011 |
| YER003C | 0.013 | 0.019 | 0.023 | 0.001 | 0.009 |
| YGL253W | 0.013 | 0.016 | 0.023 | 0.001 | 0.010 |
| YHR163W | 0.005 | 0.012 | 0.033 | 0.000 | 0.013 |
| YKL127W | 0.009 | 0.012 | 0.017 | 0.000 | 0.008 |
| YNL241C | 0.010 | 0.016 | 0.020 | 0.001 | 0.009 |
| mean | 0.008 | 0.013 | 0.020 | 0.001 | 0.010 |



(a)



(b)

Figure 1:

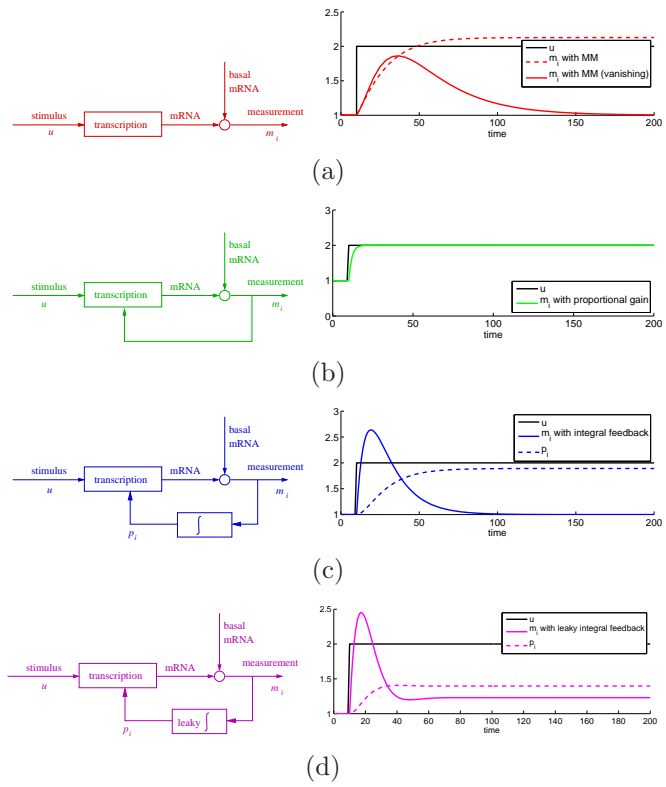


Figure 2:

YGR083C (Cytoplasmic translation initiation)

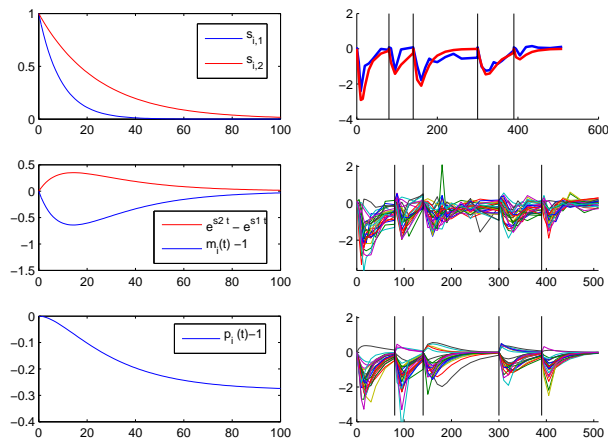


Figure 3:

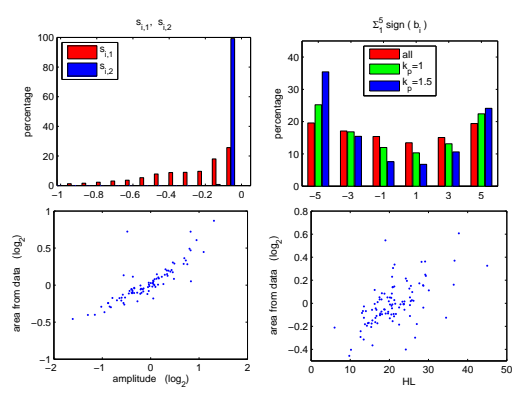


Figure 4:

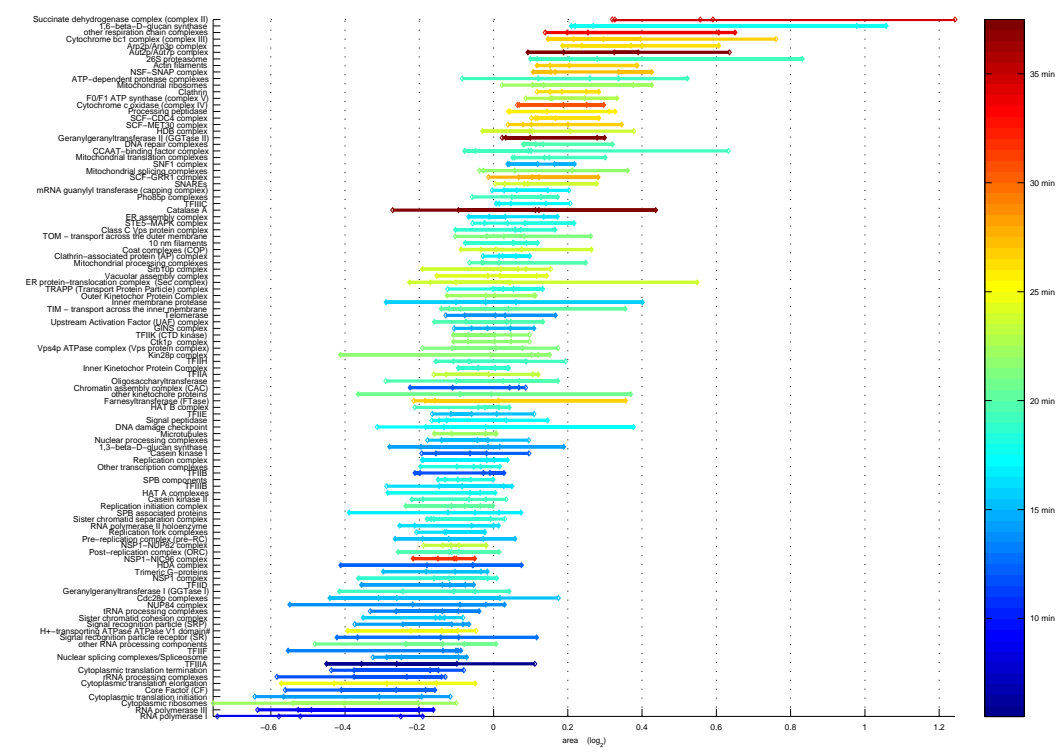


Figure 5:

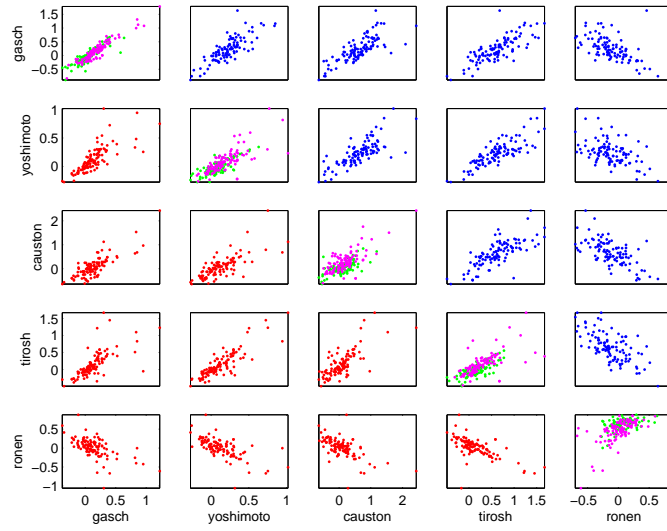


Figure 6:

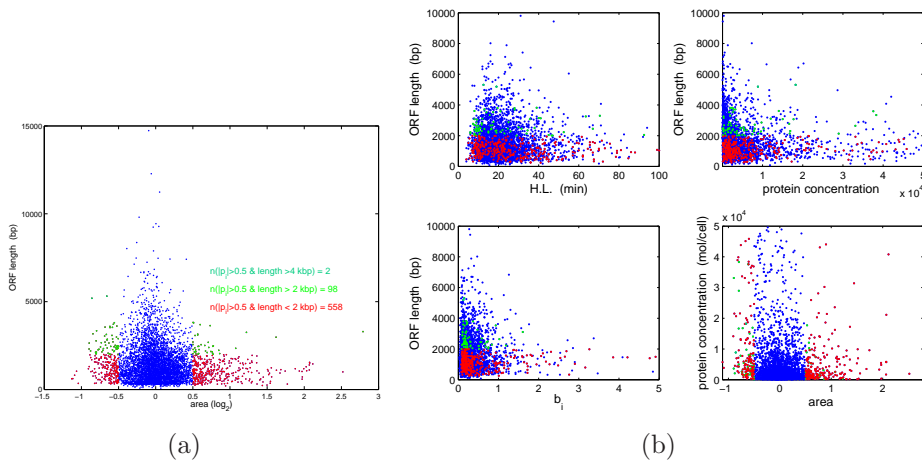


Figure 7: