

RESEARCH ARTICLE

Altered expression response upon repeated gene repression in single yeast cells

Lea Schuh^{1,2,3,4}, Igor Kukhtevich², Poonam Bheda², Melanie Schulz^{1,3}, Maria Bordukova¹, Robert Schneider^{2,5,6*}, Carsten Marr^{1,3*}

1 Institute of AI for Health, Helmholtz Zentrum München—German Research Center for Environmental Health, Neuherberg, Germany, **2** Institute of Functional Epigenetics, Helmholtz Zentrum München—German Research Center for Environmental Health, Neuherberg, Germany, **3** Institute of Computational Biology, Helmholtz Zentrum München—German Research Center for Environmental Health, Neuherberg, Germany, **4** Department of Mathematics, Technical University of Munich, Garching, Germany, **5** German Center for Diabetes Research (DZD), Neuherberg, Germany, **6** Faculty of Biology, Ludwig-Maximilians University of Munich, Planegg-Martinsried, Germany

* robert.schneider@helmholtz-muenchen.de (RS); carsten.marr@helmholtz-muenchen.de (CM)



OPEN ACCESS

Citation: Schuh L, Kukhtevich I, Bheda P, Schulz M, Bordukova M, Schneider R, et al. (2022) Altered expression response upon repeated gene repression in single yeast cells. *PLoS Comput Biol* 18(10): e1010640. <https://doi.org/10.1371/journal.pcbi.1010640>

Editor: Pedro Mendes, University of Connecticut School of Medicine, UNITED STATES

Received: July 4, 2022

Accepted: October 5, 2022

Published: October 18, 2022

Copyright: © 2022 Schuh et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data and MATLAB code corresponding to this manuscript are available under <https://github.com/marrlab/Gal1repression>.

Funding: C.M. acknowledges funding from the Bundesministerium für Bildung und Forschung (project TIDY, 031L0170B) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 866411). The work in R.S. laboratory was supported by the Deutsche

Abstract

Cells must continuously adjust to changing environments and, thus, have evolved mechanisms allowing them to respond to repeated stimuli. While faster gene induction upon a repeated stimulus is known as reinduction memory, responses to repeated repression have been less studied so far. Here, we studied gene repression across repeated carbon source shifts in over 1,500 single *Saccharomyces cerevisiae* cells. By monitoring the expression of a carbon source-responsive gene, galactokinase 1 (*Gal1*), and fitting a mathematical model to the single-cell data, we observed a faster response upon repeated repressions at the population level. Exploiting our single-cell data and quantitative modeling approach, we discovered that the faster response is mediated by a shortened repression response delay, the estimated time between carbon source shift and Gal1 protein production termination. Interestingly, we can exclude two alternative hypotheses, i) stronger dilution because of e.g., increased proliferation, and ii) a larger fraction of repressing cells upon repeated repressions. Collectively, our study provides a quantitative description of repression kinetics in single cells and allows us to pinpoint potential mechanisms underlying a faster response upon repeated repression. The computational results of our study can serve as the starting point for experimental follow-up studies.

Author summary

Cells have to continuously adjust to their environment and cope with changing temperature, stress conditions, or metabolic resources. Yeast cells exposed to repeated carbon source shifts have shown to be “primed” by their first exposure, exhibiting enhanced gene expression of specific genes later on. However, how cells respond to a repeated repressive stimulus, e.g., withdrawal of metabolic resources, has been so far much less studied.

For this, we investigated the expression kinetics of a carbon source-responsive gene across repeated repressions. We measured single-cell expression and used mathematical

Forschungsgemeinschaft (DFG, German Research Foundation) through SFB 1064 (Project-ID 213249687) and SFB 1309 (Project-ID 325871075). The funders played no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

modeling to evaluate potential causes underlying an observed faster repression response upon a repeated stimulus. Specifically, we investigated whether i) an increased dilution due to e.g., proliferation, ii) an increased fraction of repressing cells, or iii) different kinetic parameters in the repeated repression cause the observed faster response in the second repression. Leveraging quantitative mathematical model comparison, we discovered that the faster response is mediated by a shortened estimated time between carbon source shift and protein production termination at the single-cell level.

Our study provides a quantitative description of repression kinetics in single cells and allows us to pinpoint potential mechanisms underlying a faster response upon repeated repression.

Introduction

Cells receive and process external signals to optimally adjust to changing environments. Repeated stimulation from the same external signal induces an altered transcriptional response, a phenomenon termed transcriptional memory [1], with implications for a broad range of cellular functions, including the human adaptive immune system [2,3], disease development in diabetes [4,5], and aging [6]. So far, transcriptional memory has primarily been studied with respect to gene induction, despite gene repression playing an essential role in gene regulation [7,8].

The adjustment of *Saccharomyces cerevisiae* (budding yeast) to carbon sources is among the most well-studied eukaryotic signal integration systems. Whereas glucose directly enters glycolysis, a vital metabolic route providing cells with energy, galactose is first converted to glucose-6-phosphate [9,10], necessitating the production of Gal gene-encoded enzymes [11]. Repeated alternations between glucose and galactose media revealed that yeast cells are primed by their carbon source history, exhibiting transcriptional memory: repeated galactose induction results in enhanced Gal gene expression [12–16]. In Bheda et al. we examined the expression of the galactokinase 1 gene (*Gal1*) in single cells, for which reinduction memory has been well characterized, and discovered that a shorter delay, rather than an increased expression rate, contributed to the observed increase in Gal1 levels [17].

In contrast, multiple rounds of transcriptional repression have been much less studied. Lee et al. identified a stronger decrease in transcription levels and a faster response upon repeated galactose exposure in bulk experiments [18]. Here, we systematically analyzed short-term transcriptional repression kinetics in single cells. For this, we studied the expression kinetics of a carbon source-responsive gene across repeated repressions. We used mathematical modeling and the single-cell expression information to evaluate potential causes underlying a faster response. We investigated whether i) an increased dilution due to e.g., proliferation, ii) an increased fraction of repressing cells, or iii) different kinetic parameters in the repeated repression cause the observed faster response in the second repression. Specifically, we used computational methods to deconvolute protein kinetics from dilution effects in single cells and developed a mathematical model to quantitatively describe single-cell repression kinetics and to determine the subpopulation of repressing cells. Additionally, we used the estimated single-cell parameters of protein production, degradation and repression response delay to identify changes in kinetic parameters.

Results

Faster Gal1 response upon repeated repression

To study Gal1 protein kinetics over multiple short-term repressions, we re-analyzed images/movies of budding yeast cells from Bheda et al., which were alternatingly exposed to glucose or galactose media (Fig 1A) [17]. For this, cells had been cultured in custom-made microfluidic devices to ensure precise media shifts and long-term tracking. Gal1 expression levels in single yeast cells had been monitored using a Gal1-GFP fusion, a standard reporter to study gene expression in time-lapse microscopy and images from the microfluidics chambers had been taken every 3 min resulting in 320 images per chamber during a 16 h experiment (see Bheda et al., 2020 for details). Semi-automatic segmentation, tracking of the yeast cells and extraction of the total Gal1-GFP fluorescence signal per cell and time point using Autotrack and Phylo-Cell [19], YeaZ [20] and Cell-ACDC [21] (see [Materials and Methods](#)) yielded in over 1,500 single-cell Gal1 expression traces (Fig 1B). To obtain an independent biological replicate, we repeated the experiment and analyzed it according to Bheda et al. (S1 Fig). As expected, the total Gal1-GFP fluorescence signal of single cells recapitulates Gal1 inductions and repressions during galactose and glucose, respectively, and increased overall Gal1 levels in induction i2 in both the original data as well as in our new experiments (Figs 1B and S1A). We also visually observed a faster response of Gal1-GFP fluorescence at the population level upon repression r2 (Fig 1C), which is in line with the findings of Lee et al. [18].

Dilution effects cannot explain faster response upon repeated repression

During budding, cytoplasmic proteins are disseminated between mother and daughter cells. Assuming a constant Gal1 protein amount, this decreases the total Gal1-GFP fluorescence signal in the mother cell (Fig 2A top), a phenomenon called dilution. Could an increased dilution because of e.g., increased proliferation in repression r2 compared to r1, explain the faster repression response observed upon repeated carbon source shifts (Fig 1C)? The single-cell resolution of our data and asymmetric budding allowed us to identify mother-daughter relationships and to deconvolute dilution and repression kinetics. For this, we calculated the sum of the total Gal1-GFP fluorescence signal of the mother cell, i.e., cell present at the start of repression, and its progeny, i.e., daughter cell(s) detected during repression (Fig 2A bottom, and [Materials and Methods](#)) to account for dilution effects due to cell divisions. In the following, the adjusted sum of the total Gal1-GFP fluorescence signal of the mother cell and its progeny

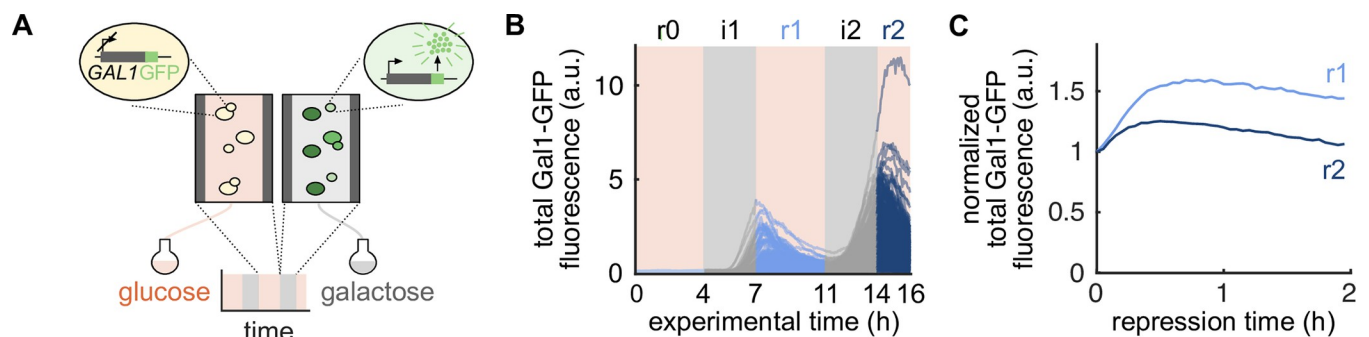


Fig 1. Faster Gal1 response upon repeated repression. (A) Budding yeast cells were grown in microfluidic chambers and alternatingly exposed to a medium containing either glucose (orange) or galactose (gray) as carbon source. Galactokinase 1 (*Gal1*) is induced in cells exposed to galactose and repressed in cells exposed to glucose. Gal1 expression was monitored via a Gal1-GFP fusion and time-lapse microscopy. (B) Single-cell traces of total Gal1-GFP fluorescence signal across two inductions i1 and i2 (gray) and repressions r0, r1, and r2 (blue). (C) Population means of the normalized total Gal1-GFP fluorescence signals of repressions r1 and r2, shifted to the respective start of the repression time.

<https://doi.org/10.1371/journal.pcbi.1010640.g001>

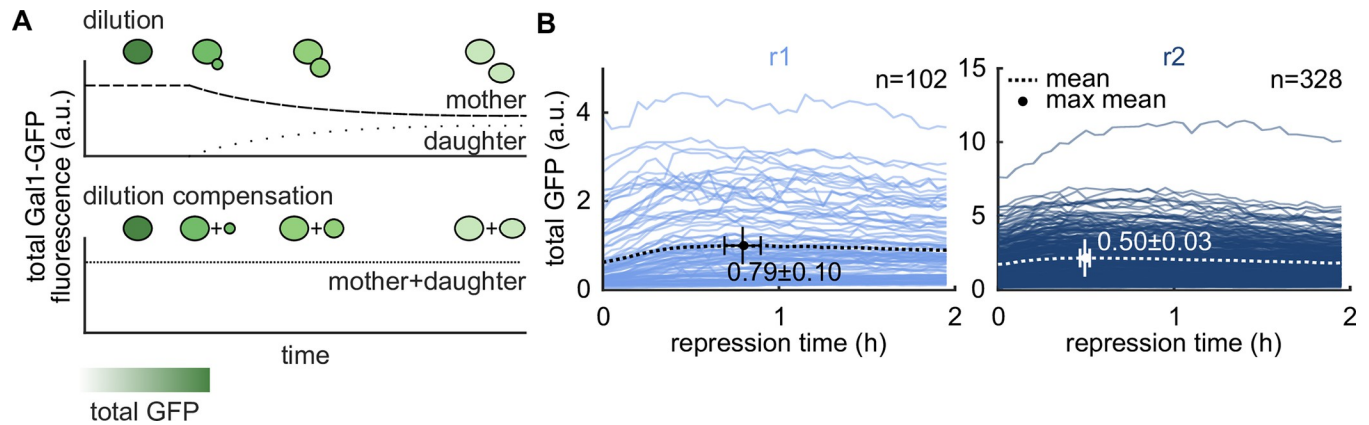


Fig 2. Dilution effects cannot explain the faster response upon repeated repression. (A) Budding decreases the total Gal1-GFP fluorescence signal in mother cells and increases the total Gal1-GFP fluorescence signal in daughter cells (top). To compensate for this dilution, we summed up the total Gal1-GFP fluorescence signal of each mother cell present and its progeny during one repression (bottom). (B) Single-cell traces of total GFP signal adjusted for dilution (see (A)) for the first two hours of repressions r1 (left) and r2 (right). Time to maximal mean total GFP is 17 min shorter in repression r2, where mean total GFP is indicated by the dotted line, the maximal mean total GFP is highlighted by the dot. Bootstrap (10^5) samples were drawn to generate mean \pm std.

<https://doi.org/10.1371/journal.pcbi.1010640.g002>

is referred to as total GFP. We applied the same dilution compensation to repressions r1 and r2 (Fig 2B). Note that the increase in total GFP traces from $n = 102$ in r1 to $n = 328$ in r2 is due to cellular proliferation during the 4 hours of repression r1.

Following a galactose-glucose shift, total GFP intensities initially rise before decreasing. To obtain quantitative insights, we calculated the time to maximal mean total GFP for repressions r1 and r2. This time reduced from 0.79 ± 0.10 h (mean \pm std, $n = 102$ cells) in r1 to 0.50 ± 0.03 h ($n = 328$ cells) in r2, where the time point of the maximal mean was calculated with 10^5 bootstraps (Fig 2B). A shorter time to maximal mean total GFP suggests a faster response upon repeated repression, even after accounting for dilution effects due to proliferation. Hence, dilution effects cannot explain the faster response in repression r2 at the population level.

A computational model identifies cells with and without repression kinetics

At the single-cell level, *Gal1* induction delay varies significantly. Zacharioudakis et al. [16] showed that *Gal1* induction caused by a glucose–galactose media shift results in a bimodal population distribution, with only a subset of cells inducing *Gal1* even after several hours of galactose exposure. As in our experiments repression was preceded by 3 h of galactose induction, we expected that our cell population at the start of repression contained induced and uninduced cells, which thus show and do not show protein repression kinetics, respectively. Could a larger fraction of repressing cells in repression r2 explain the faster response observed upon repeated repression (Fig 2B)? To determine the subpopulation of cells with protein repression kinetics, information on single-cell kinetics is required. As it is difficult to distinguish between cells showing and not showing protein repression kinetics from the total GFP traces alone (Fig 2B), we used computational modeling and model selection to systematically describe and classify the kinetics of single total GFP traces. In comparison to other classification methods, model selection allows us to classify total GFP traces without determining arbitrary thresholds for e.g. initial total GFP when classifying induced and uninduced cells at repression start. Since *Gal1* induction results in an approximate 1000-fold change in Gal1 expression [22], we assumed that stochasticity inherent to gene expression was insignificant and that a deterministic modeling approach was sufficient in describing the kinetics of the

total GFP traces. To discriminate between cells showing and not showing protein repression kinetics, we defined two models.

The model not accounting for repression assumes a constant basal GFP production and degradation over time with rates r_{basal} and r_{deg} (Fig 3A left), since we observed a gradual increase in total GFP signal in cells visually identified as not showing repression kinetics (Fig 3B top right). Note that r_{basal} encompasses all reaction steps that lead to the production of a detectable GFP protein, including mRNA transcription, Gal1-GFP translation, and GFP maturation. Similarly, r_{deg} comprises active GFP degradation and potential effects of photo-bleaching. Dilution effects are compensated prior to modeling and are not part of r_{deg} (see Methods).

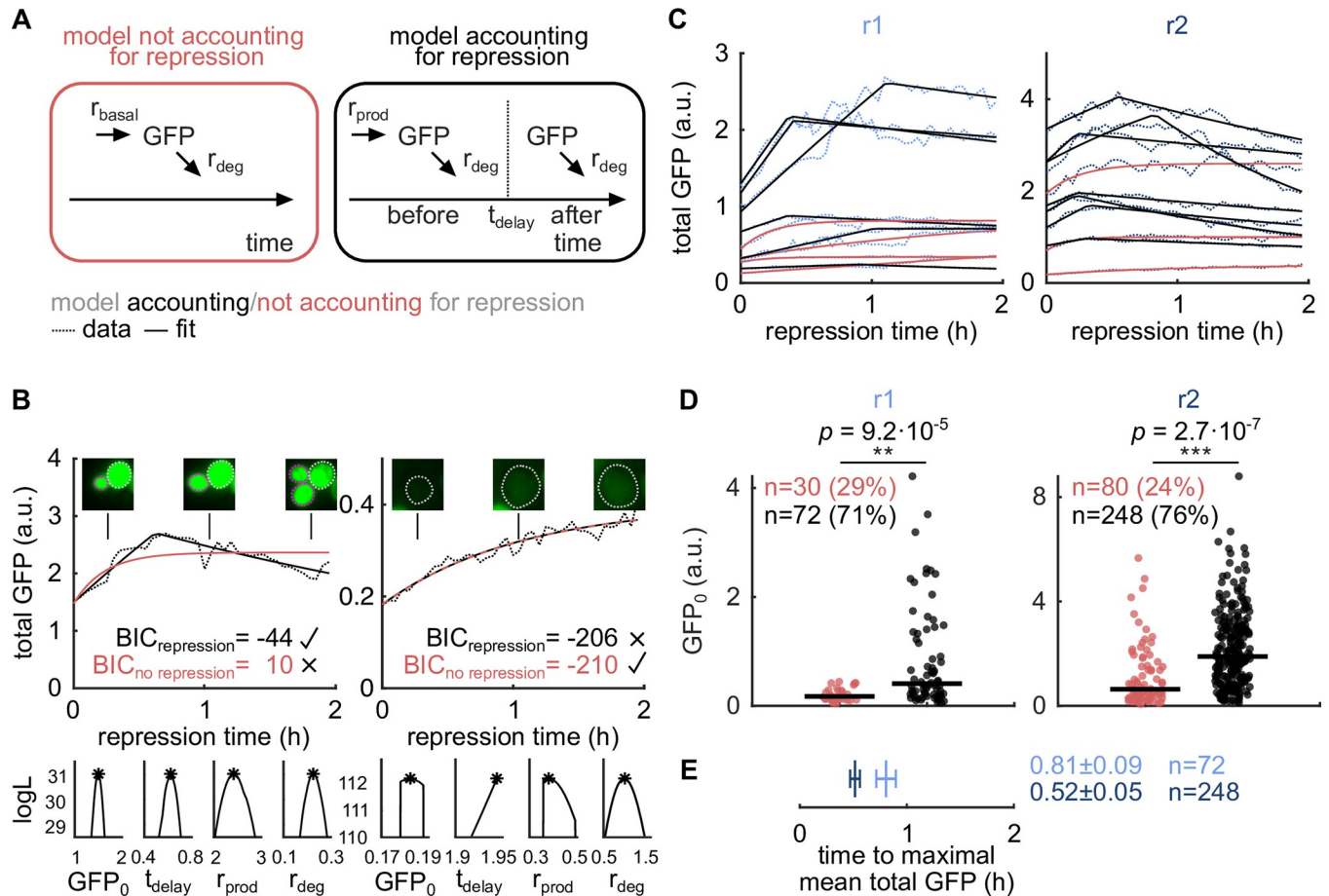


Fig 3. Fraction of cells with repression kinetics cannot explain the faster response upon repeated repression. (A) Left: a model for cells without repression kinetics composed of basal GFP production (r_{basal}) and degradation (r_{deg}). Right: a model for cells with repression kinetics composed of an initial constant and active GFP production (r_{prod}) and degradation (r_{deg}) until a delayed repression onset (t_{delay}) where GFP production is switched off. (B) Top: two exemplary total GFP traces (dotted line) and fits of the model not accounting for repression (red solid line) and model accounting for repression (black solid line). Exemplary images of the cell(s) at three different time points are shown above. Mother cells are circled in gray, progeny in pink. The better fitting model was selected according to the Bayesian information criterion (BIC). Left: total GFP trace better fitted by the model accounting for repression. Right: total GFP trace fitted equally well by both models. Due to the higher model complexity of the model accounting for repression, the model accounting for repression is rejected. Bottom: profile likelihoods of the model accounting for repression corresponding to the two exemplary total GFP traces above endorse parameter identifiability. Asterisks represent optimized parameters and corresponding log-likelihood (logL) values. (C) Ten exemplary total GFP traces (dotted lines) and best fits (solid lines) for repressions r1 (left) and r2 (right). GFP traces best fitted with the model accounting for repression are shown in black: total GFP traces best fitted with the model not accounting for repression are shown in red. (D) The median initial total GFP, GFP_0 , is significantly higher ($p = 9.2 \cdot 10^{-5}$ and $p = 2.7 \cdot 10^{-7}$, two-sided median test corrected for multiple testing with Bonferroni correction, $m = 12$) in traces better fitted by the model accounting for repression (black) than in traces better fitted by the model not accounting for repression (red). This confirms that the model accounting for repression fits induced cells better, while the model not accounting for repression fits uninduced cells. The number of cells and percentages of all GFP traces best fitted by the model accounting and not accounting for repression are shown. (E) Time to maximal mean total GFP is decreased in repression r2 for cells with repression kinetics ($0.81 \text{ h} \pm 0.09 \text{ h}$ vs. $0.52 \text{ h} \pm 0.05 \text{ h}$). Bootstrap (10^5) samples of the cells with repression kinetics were drawn to generate mean \pm std.

<https://doi.org/10.1371/journal.pcbi.1010640.g003>

The temporal variation of the total GFP signal described by the model not accounting for repression is summarized using the following ordinary differential equation:

$$\frac{\partial \text{GFP}(t)}{\partial t} = r_{\text{basal}} - r_{\text{deg}} \text{GFP}(t).$$

This is solved by

$$\text{GFP}(t) = \frac{r_{\text{basal}}}{r_{\text{deg}}} (1 - e^{-r_{\text{deg}}t}) + \text{GFP}_0 e^{-r_{\text{deg}}t},$$

where $\text{GFP}_0 = \text{GFP}(0)$, the initial total GFP at time point 0.

According to the model accounting for repression, cells that induced Gal1 during galactose induction require time to reestablish glucose-mediated repression. Hence, GFP is actively produced at rate r_{prod} , where r_{prod} comprises a basal production rate r_{basal} and an additional production rate accounting for the active production of GFP, till a time point t_{delay} . GFP production is switched off ($r_{\text{prod}} = 0$) and GFP is degraded with rate r_{deg} after this estimated repression response delay t_{delay} (Fig 3A right). Again, r_{prod} includes mRNA transcription, Gal1-GFP translation, and GFP maturation, while r_{deg} comprises active GFP degradation, and photo-bleaching, but no dilution. The temporal change of total GFP over time described by the model accounting for repression is summarized by the following ordinary differential equations:

$$\text{before } t_{\text{delay}} : \frac{\partial \text{GFP}(t)}{\partial t} = r_{\text{prod}} - r_{\text{deg}} \text{GFP}(t)$$

$$\text{after } t_{\text{delay}} : \frac{\partial \text{GFP}(t - t_{\text{delay}})}{\partial t} = -r_{\text{deg}} \text{GFP}(t - t_{\text{delay}})$$

with solutions

$$\text{before } t_{\text{delay}} : \text{GFP}(t) = \frac{r_{\text{prod}}}{r_{\text{deg}}} (1 - e^{-r_{\text{deg}}t}) + \text{GFP}_0 e^{-r_{\text{deg}}t}$$

$$\text{after } t_{\text{delay}} : \text{GFP}(t - t_{\text{delay}}) = \text{GFP}(t_{\text{delay}}) e^{-r_{\text{deg}}(t - t_{\text{delay}})},$$

where

$$\text{GFP}(t_{\text{delay}}) = \frac{r_{\text{prod}}}{r_{\text{deg}}} (1 - e^{-r_{\text{deg}}t_{\text{delay}}}) + \text{GFP}_0 e^{-r_{\text{deg}}t_{\text{delay}}}.$$

An example of a cell visually identified as showing repression kinetics is shown in Fig 3B top left. Right next to it, we show the fluorescent microscopy images and total GFP trajectory of a cell visually identified as not showing repression kinetics. Until t_{delay} , where $t_{\text{delay}} < 2$ h, the model accounting for repression equals the model not accounting for repression. However, these model definitions allow for model selection and circumvent the usage of arbitrary thresholds for, e.g., the estimated value of t_{delay} . Both models comprise four and five model parameters, respectively: initial total GFP, GFP_0 , basal GFP production rate, r_{basal} , or GFP production rate, r_{prod} , GFP degradation rate r_{deg} , a noise parameter σ determining the width of the Gaussian noise distribution (see Materials and Methods), and the repression response delay t_{delay} for the model accounting for repression. For repressions r1 and r2, respectively, we performed multi-start maximum likelihood optimization to estimate the single-cell parameters, and model selection on both models for each total GFP trace (Fig 3B top and 3C). Calculating the

profile likelihoods of exemplary total GFP traces, we found the model parameters of the model accounting for repression to be identifiable (Fig 3B bottom). We then determined whether the model accounting for repression was required to explain a total GFP trace using the Bayesian information criterion (BIC). A BIC difference of ten between the model accounting for and not accounting for repression ($\text{BIC}_{\text{repression}} < \text{BIC}_{\text{no repression}} - 10$) was considered an appropriate threshold to reject the model not accounting for repression with fewer model parameters (see Materials and Methods and Fig 3B and 3C), as commonly done in model selection using BIC as selection criterion [23–25]. GFP traces of cells requiring the model accounting for repression are henceforth referred to as “cells with repression kinetics”. The median initial total GFP, GFP_0 , was significantly higher in cells with repression kinetics than in cells without repression kinetics ($p = 9.2 \cdot 10^{-5}$ for r1 and $p = 2.7 \cdot 10^{-7}$ for r2), meaning that the model accounting for repression mainly describes cells with higher Gal1 induction levels (Fig 3D). This implies that our models can discriminate between cells that were repressing Gal1 and cells uninduced at the beginning of repression. The overlapping ranges of initial total GFP between cells with and without repression kinetics (Fig 3D) reveal how simple thresholding could result in wrong classification of cells with and without repression kinetics.

Fraction of cells with repression kinetics in r2 cannot explain faster response upon repeated repression

Using our modeling approach to classify total GFP traces into cells with and without repression kinetics, we found that of all total GFP traces, 71% and 76% of r1 and r2, respectively, require the model accounting for repression (Fig 3D). Since the fraction of cells with repression kinetics is only slightly increased in repression r2, this effect cannot explain the faster response upon repeated repression. Hence, we determined the time to maximal mean total GFP again, this time constraining the analysis to the subpopulation of cells with repression kinetics. We again found a shortened time to maximal mean total GFP in r2, with $0.81 \text{ h} \pm 0.09 \text{ h}$ (mean \pm std, $n = 72$ cells) and $0.52 \text{ h} \pm 0.05 \text{ h}$ ($n = 248$ cells) for r1 and r2, respectively (Fig 3E), demonstrating that the faster response in r2 at the population level is not due to the increased fraction of cells with repression kinetics.

Shortened repression response delay mediates faster response upon repeated repression at the single-cell level

So far, we excluded major contributions of dilution effects and an increased fraction of cells with repression kinetics in r2 to underlie the faster response upon repeated repression. Could altered kinetics of individual cells between repressions r1 and r2 lead to the observed change in repression response? In the previous section, we estimated the kinetic parameters GFP_0 , t_{delay} , r_{prod} , and r_{deg} for cells with repression kinetics. To compare protein kinetics between repression r1 and r2 at the single-cell level, we used these estimated single-cell parameters of cells with repression kinetics present in both repressions r1 and r2. We discovered that the median initial total GFP, GFP_0 , and median repression response delay, t_{delay} , are significantly different ($p = 3.3 \cdot 10^{-9}$ and $p = 1.5 \cdot 10^{-3}$, respectively) between both repressions using a two-sided paired sign test with multiple testing correction (Fig 4A and 4B). Median GFP_0 is increased (median values of 0.59 and 2.05 for r1 and r2, respectively, Fig 4A), while median t_{delay} is shortened in r2 (median values of 0.50 and 0.38 for r1 and r2, respectively, where 72% of paired cells showed a decrease in t_{delay} , Fig 4B). The increased GFP_0 in r2 conforms to the transcriptional reinduction memory of Gal1 and results in higher GFP_0 at the start of r2. The median r_{prod} and median r_{deg} between the two repressions were comparable ($p = 0.22$ and $p = 0.34$, respectively) (Fig 4C and 4D). In line with our findings, Bheda et al. reported similar

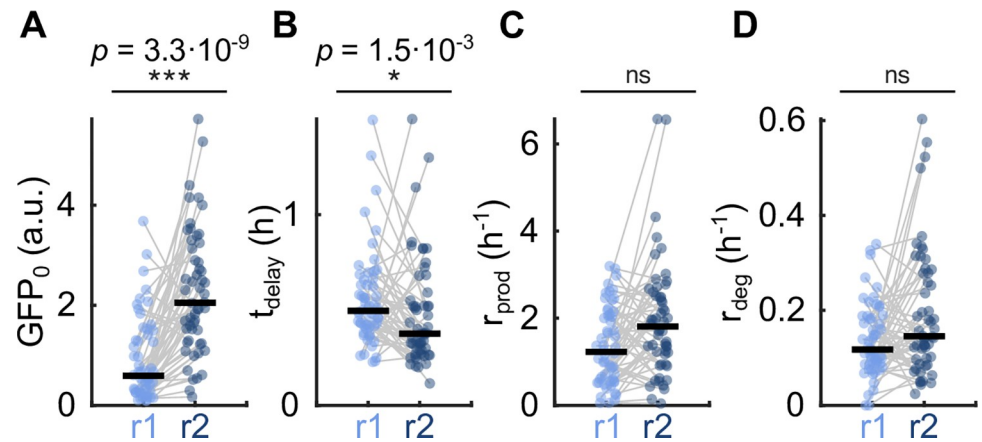


Fig 4. Shortened repression response delay mediates faster response upon repeated repression at single-cell level. Comparison of paired estimated single-cell parameters of cells with repression kinetics of repression r1 and r2 shows that the median initial total GFP, GFP_0 , and median repression delay, t_{delay} , are significantly different ($p = 3.3 \cdot 10^{-9}$ and $p = 1.5 \cdot 10^{-3}$, respectively, two-sided paired sign test correcting for multiple testing with Bonferroni correction, $m = 12$, and the number of paired cells = 54), with median GFP_0 increased and median t_{delay} decreased (median t_{delay} values of 0.50 and 0.38 for r1 and r2, respectively). GFP degradation is not different between repressions r1 and r2 ($p = 0.34$).

<https://doi.org/10.1371/journal.pcbi.1010640.g004>

production rates [17]. We repeated the entire analysis based on data from independent experiments (see [Methods](#) for details). While the exact parameters vary between replicates due to the heterogeneity in Gal1 induction at the single-cell level, our replicate analysis confirms the conclusions, in particular that a shortened repression response delay in r2 mediates the observed faster response upon repeated repression ([S1 Fig](#)). We also tested a more complex, moment equation derived noise model, explicitly accounting for the stochastic nature of single-cell GFP production and degradation (see [Methods](#) for details). Due to the increased flexibility of this noise model, we detected fewer cells with repression kinetics for repressions r1 and r2. The set of cells with repression kinetics detected using this revised noise model is, however, largely overlapping with the previously reported set of cells with repression kinetics for both repressions r1 and r2 ([S2A and S2B Fig](#)). More importantly, the repression response delays of the cells with repression kinetics detected by both models are estimated to comparable values, suggesting that a more complex noise model does not influence the main conclusions of this work (see [S2C and S2D Fig](#)). Of note, a comparison between the estimated single-cell parameters of cells with repression kinetics of the two different models is uninformative, due to the low statistical power of the small set of paired cells with repression kinetics across both models for repressions r1 and r2. Finally, using a simulation study, we tested whether the GFP degradation rate influences our results ([S3 Fig](#)). We found that too small degradation rates might lead to misclassifying cells as cells without repression kinetics, hence potentially leading to a loss of information. However, this does not affect our conclusions, since degradation rates are estimated correctly and reliably ([S3 Fig](#)).

Discussion

Here, we described a faster Gal1 response upon repeated short-term repression at the population level in support of what had been shown by Lee et al. [18]. By using single-cell Gal1 expression data and mathematical modeling, we demonstrated that a faster response upon repeated repression is not simply due to increased dilution effects or a larger fraction of repressing cells, but is mediated by a shortened repression response delay at the single-cell level which could suggest the existence of transcriptional repression memory.

We would like to note that we assessed the kinetics of the Gal1-GFP fusion protein, and hence also modeled Gal1-GFP kinetics. Although this reporter has been successfully employed in several studies involving Gal1 gene expression [16,26–29], Gal1-GFP does not allow us to exclude alternative molecular mechanisms driving the shortened repression response delay $t_{\text{de-lay}}$ in repression r2 such as an accelerated mRNA degradation. Altered GFP maturation was accounted for in the production rate r_{prod} . Of note, GFP stability is estimated by the degradation rate, r_{deg} , which we found to be not significantly different between repressions r1 and r2 ($p = 0.34$). While the degradation rates are estimated to be quite low, they are high enough such that repression kinetics are observable (see Fig 3C). The stability of GFP may only lead to a loss of information, i.e., repression kinetics are overshadowed by the accumulation of GFP, such that cells with repression kinetics might no longer be detected as such (S3 Fig). GFP stability does not influence the estimated repression response delays and hence our conclusions. Potential effects due to photo-bleaching would increase r_{deg} but not affect the estimation of $t_{\text{de-lay}}$ and hence our conclusions.

Due to the rather simple kinetics of the total GFP traces, our model is sufficiently complex to describe the data (Fig 3C). More detailed descriptions of this process e.g., shut off kinetics of protein production, mRNA/promoter kinetics, or histone occupancy, would be desirable but inevitably lead to unidentifiable model parameters and, hence, to little or none additional insight. Similarly, we did not consider cell-size and cell-cycle aware stochastic models, as for example successfully applied by Song et al. [30].

During model design, we also tested mixed effect modeling frameworks, where single-cell parameters are assumed to underlie a certain population distribution and parameter correlations can be accounted for. However, due to the large and unconventional noise distributions underlying the model parameters of our data, mixed effect modeling did not describe the data well enough at the population level. Hence, we believe that for the given data, our simple modeling framework was the most informative approach possible.

Throughout this study, we mainly assumed an additive Gaussian noise model with constant variance to correctly reflect measurement noise. However, other noise models might also be applicable. For example, the Laplacian noise model has shown to be more robust against outliers and, hence, more suitable for corrupted data [31]. We also investigated time-dependent noise from a moment equation based model, taking into account not only measurement but also intrinsic noise emerging from the stochasticity of the underlying molecular processes. While we show that these different noise models lead to the same conclusions in our study, we want to highlight here that identifying a suitable noise model is generally important for both parameter inference as well as model selection [32,33]. The similarity between the two approaches is likely due to the large molecule numbers of Gal1-GFP expression in our study.

In this study, we focused on the altered *Gal1* expression response upon repeated gene repression. Our simple and linear repressor model used to identify the repression delay is basic data analysis, such that its results could be understood as empirical evidence for altered gene repression kinetics. Whether this phenomenon is a unique characteristic of *Gal1* or if this is universal to other non-galactose induced memory genes of, e.g, the MAL or INO system, will be interesting to investigate in the future. Moreover, *Gal1* reinduction memory is inherited by daughter cells [17]. Whether altered repression kinetics can also be maintained through multiple cell divisions is still unknown. Due to yeast cells mainly arresting their cell-cycles in the initial hours of galactose exposure (inductions i1 and i2), dividing cells are rarely detected in i2 and thus we cannot investigate the effect of cell divisions on altered repression kinetics in r2.

Together, our work lays the foundation for further mechanistic studies using single-cell mRNA kinetics or time-resolved information regarding the chromatin architecture at the *Gal1*

promoter to investigate the molecular workings underlying a shortened repression response delay upon repeated short-term repression.

Materials and methods

Data acquisition and sources

For the analysis for Figs 1 and 2, we used microscopy images and initial segmentation, mapping, and tracking information from a microfluidics experiment from Bheda et al. [17], which contained 13 positions. The images from the first two hours of repression r1 were rectified, and the segmentation, mapping, and tracking were extended to the entire two hours of repression r2. Bheda et al. only segmented r2 partially since they were primarily interested in galactose induction, and did not adjust the final repression frames. Using the software PhyloCell [19], we manually corrected the segmentation, mapping, and tracking of r1 and r2. For the replicate analysis (S1 Fig), we repeated the induction-repression experiment as described in Bheda et al. [17] (Fig 1A). Due to low cell numbers, we pooled data from three independent experiments, totalling 13 positions, respectively. Using the softwares YeaZ [20] and Cell-ACDC [21] for cell segmentation, mapping and tracking, we extracted the relevant single-cell information of the live-cell images for both repressions r1 and r2. During glucose repression, the yeast cells proliferated, increasing the cell numbers within the microfluidic chambers. However, filled microfluidic chambers no longer assure that all the progeny of a cell is recorded, and mapping and tracking of cells become infeasible. To ensure mapping and tracking of single yeast cells within the microfluidic chambers, the glucose repressions were limited to a maximum of 4 h and the overall experiment was limited to 16 h (4 h in glucose (r0), 3 h in galactose (i1), 4 h in glucose (r1), 3 h in galactose (i2), 2 h in glucose (r2)).

Data preprocessing

We extracted the single-cell information relevant for our analysis, namely cell ID, mother cell ID, detection frame (first frame in which a cell is detected), last frame (last frame a cell is detected), relative GFP intensities per time (mean GFP intensity of a segmented cell) and cell area per time. As the data regarding the relative GFP intensities and cell area was not sorted over time, we first sorted it and then calculated the total GFP fluorescence per time given by

$$\text{total GFP fluorescence} = \text{relative GFP fluorescence} \times \text{cell area}.$$

Finally, cells that were not imaged till the end of the experiment, cells with missing relative GFP and/or cell area values, and cells that were supposedly detected before their mother cells (segmentation error) were discarded.

Dilution compensation

Cytoplasmic proteins are disseminated between the mother and daughter cells during budding. Assuming that Gal1 is not produced or degraded, protein redistribution causes a drop in total Gal1-GFP fluorescence in the mother cell and a rise in total Gal1-GFP fluorescence in the daughter cell till the mother and daughter cells split (Fig 2A top). As a result, regardless of repression, dilution causes variations in total Gal1-GFP fluorescence. The daughter cell grows to about $\frac{1}{3}$ of the size of the mother cell [34] such that the decrease in total Gal1-GFP fluorescence due to dilution was expected to be $\frac{1}{3}$ of the initial total Gal1-GFP fluorescence of the mother cell. To ensure that dilution does not overshadow potentially more subtle repression kinetics, we created artificial non-dividing cells compensating for dilution by adding the total Gal1-GFP fluorescence of the progeny of a cell present at the start of glucose repression, which we called mother cell, to the total Gal1-GFP fluorescence of that mother cell during the first 2

h of repression (Figs 2A bottom, 2B and S1C). For mother cells with a bud at the beginning of a repression period, we additionally added the bud to the total Gal1-GFP fluorescence of that mother cell. The GFP traces of all computed non-dividing cells can be found under <https://github.com/marrlab/Gal1repression>. As we found the maximal mean total GFP to be attained before 2 h of glucose exposure, we restricted our analysis to the first 2 h of repression.

Models

During the first two hours of glucose repression, we modeled the kinetics of the total GFP of every single cell. Due to the high variability in galactose induction, we assumed that our initial cell population at the beginning of repression contained induced and uninduced cells, which show and do not show repression kinetics, respectively. We developed two models, with and without repression, to account for both total GFP kinetics during repression.

Model not accounting for repression

For more information regarding the model not accounting for repression, see the main text.

Model accounting for repression

For more information regarding the model accounting for repression, see the main text.

Noise models

Time-independent noise model taking into account only measurement noise:

Experimental data, such as total GFP per cell per time, is noise corrupted. As a result, we used an underlying additive Gaussian noise model with a constant variance σ^2 throughout time to test our models. The single-cell specific model parameters are comprised in the parameter vector Θ_i for cell i and the experimental measurement at time point k for cell i is denoted by \bar{y}_i^k . The log-likelihood for the Gaussian noise model is given by

$$\log L(\theta_i) = \frac{-1}{2} \sum_k \log(2\pi\sigma_i^2) + \frac{(\bar{y}_i^k - y(t_k, \theta_i))^2}{\sigma_i^2}.$$

Time-dependent noise model taking into account measurement and intrinsic noise:

Due to the stochastic nature of GFP production and degradation, the single-cell GFP measurements also underlie, in addition to measurement noise, intrinsic noise. The variance of the more sophisticated noise model is hence composed of (1) the constant measurement noise σ^2 and (2) the time-dependent variance of the stochastic process underlying the corresponding model. According to the moment equations, the time evolution of the variance of the stochastic process underlying the model not accounting for repression is given by

$$\frac{\partial \text{Var}(t)}{\partial t} = -2r_{\text{deg}} \text{Var}(t) + r_{\text{basal}}(2 - e^{-r_{\text{deg}}t}) + \text{GFP}_0 r_{\text{deg}} e^{-r_{\text{deg}}t},$$

with solution

$$\text{Var}(t) = \frac{1}{r_{\text{deg}}} e^{-2r_{\text{deg}}t} (r_{\text{deg}} (\text{GFP}_0 (e^{r_{\text{deg}}t} - 1) + \text{Var}_0) + r_{\text{basal}} e^{r_{\text{deg}}t} (e^{r_{\text{deg}}t} - 1)),$$

where $\text{GFP}_0 = \text{GFP}(0)$, the initial total GFP at time point 0, $\text{Var}_0 = \text{Var}(0)$, the initial stochastic variance at time point 0, and $\text{GFP}(t)$ as derived in the main text.

For the model accounting for repression, we need to consider the piecewise stochastic processes before and after t_{delay} . The time evolutions of the variances of the stochastic processes

are given by

$$\text{before } t_{\text{delay}} : \frac{\partial \text{Var}(t)}{\partial t} = -2r_{\text{deg}} \text{Var}(t) + r_{\text{prod}}(2 - e^{-r_{\text{deg}}t}) + \text{GFP}_0 r_{\text{deg}} e^{-r_{\text{deg}}t}$$

$$\text{after } t_{\text{delay}} : \frac{\partial \text{Var}(t - t_{\text{delay}})}{\partial t} = -2r_{\text{deg}} \text{Var}(t - t_{\text{delay}}) + r_{\text{deg}} \text{GFP}(t_{\text{delay}}) e^{-r_{\text{deg}}(t - t_{\text{delay}})},$$

with solutions

$$\text{before } t_{\text{delay}} : \text{Var}(t) = \frac{1}{r_{\text{deg}}} e^{-2r_{\text{deg}}t} (r_{\text{deg}} (\text{GFP}_0 (e^{r_{\text{deg}}t} - 1) + \text{Var}_0) + r_{\text{prod}} e^{r_{\text{deg}}t} (e^{r_{\text{deg}}t} - 1))$$

$$\text{after } t_{\text{delay}} : \text{Var}(t - t_{\text{delay}}) = e^{-2r_{\text{deg}}(t - t_{\text{delay}})} (\text{GFP}(t_{\text{delay}}) (e^{r_{\text{deg}}(t - t_{\text{delay}})} - 1) + \text{Var}(t_{\text{delay}})),$$

where

$$\text{Var}(t) = \frac{1}{r_{\text{deg}}} e^{-2r_{\text{deg}}t_{\text{delay}}} (r_{\text{deg}} (\text{GFP}_0 (e^{r_{\text{deg}}t_{\text{delay}}} - 1) + \text{Var}_0) + r_{\text{prod}} e^{r_{\text{deg}}t_{\text{delay}}} (e^{r_{\text{deg}}t_{\text{delay}}} - 1))$$

$\text{Var}_0 = \text{Var}(0)$, the initial stochastic variance at time point 0, and $\text{GFP}(t)$, $\text{GFP}(t_{\text{delay}})$ as derived in the main text.

Together, the total variance of the Gaussian distribution accounting for measurement and intrinsic noise for time point t is given by

$$\text{Var}_{\text{total}}(t) = \sigma^2 + \text{Var}(t).$$

The log-likelihood for the more complex, time-dependent, Gaussian noise model is then given by

$$\log L(\theta_i) = \frac{-1}{2} \sum_k \log(2\pi \text{Var}_{\text{total}i}(t)) + \frac{(\bar{y}_i^k - y(t_k, \theta_i))^2}{\text{Var}_{\text{total}i}(t)}.$$

We obtained the optimal model parameters of both models and both noise models for the total GFP traces for each cell by performing maximum likelihood estimation.

Optimization and parameter estimation

For each total GFP trace separately for r_1 and r_2 , we computed the model parameters for the models accounting and not accounting for repression. The initial total GFP GFP_0 , the basal production rate r_{basal} , and the degradation rate r_{deg} are the model parameters for the model not accounting for repression. Instead of a basal production rate, r_{basal} , we have a production rate, r_{prod} , for the model accounting for repression. Also, we discovered the time point of the repression response delay t_{delay} for the model accounting for repression. For both models, we also inferred the noise parameter σ determining the spread of the Gaussian noise model and potentially the initial stochastic variance Var_0 when testing the sophisticated noise model. Overall, this results in 4 (5) inferrable parameters for the model not accounting for repression and 5 (6) inferrable parameters for the model accounting for repression, assuming a time-(in) dependent noise model. We assumed that all parameters are constant over time. For numerical reasons we optimized the parameters in \log_{10} scale [35] and rescaled the data by 10^7 . As total Gal1-GFP fluorescence signal and total Gal1-GFP molecules are (linearly) mapped by an unknown constant, the number of total Gal1-GFP molecules is always scalable by that unknown constant that we exploit to increase convergence. The lower and upper bounds for

all initial, rate, and noise parameters are -10 and 1 in \log_{10} scale, respectively, assuring that the whole range of biologically plausible parameter values is covered. The lower and upper bounds for the repression response delay are given by 36 s and 2 h (corresponding to -2 and $\log_{10}(2)$ in \log_{10} scale). As we only considered 2 h of glucose repression, we did not allow the time delay to take on larger values. We performed multi-start local optimization of the negative log-likelihood using the parameter estimation toolbox PESTO [36]. For each model and total GFP trace, we performed local optimization runs from at least 20 different Latin-hypercube-sampled starts. If less than five starts converged, i.e. the objective function values of the starts differ less than 0.1 to the best start, we re-ran the optimization with 50, 100, and 200 starts until at least five starts converged for each GFP trace.

Model selection

We used the Bayesian information criterion (BIC) [37] for comparing the model accounting and not accounting for repression per total GFP trace. The BIC is calculated by

$$\text{BIC} = \log(n) k - 2 \log L,$$

where n is the number of data points, k is the number of estimated parameters and $\log L$ is the log-likelihood value for the maximum likelihood estimate of the model parameters. Here, the number of estimated parameters is either four for the model not accounting for repression or five for the model accounting for repression. The BIC rewards high likelihood values and penalizes the model complexity in the form of additional model parameters. We considered the model accounting for repression to fit a given total GFP trace considerably better than the model not accounting for repression if $\text{BIC}_{\text{repression}} < \text{BIC}_{\text{no repression}} - 10$ (Figs 3B–3D, S1D and S1E).

Statistical analysis

Comparison of initial total GFP of total GFP traces. On the estimated initial total GFP, GFP_0 , of all total GFP traces significantly better fitted by a model accounting for repression and all total GFP traces better fitted by a model not accounting for repression, we did a two-sided median test. To avoid false-positive results, we used the Bonferroni correction, which adjusts the significance-level $\alpha = 0.05$ by the total number of investigated null hypotheses m , such that $\alpha' = \alpha/m$. In this study, the total number of null hypotheses for the analysis is $m = 12$:

- Two hypothesis tests comparing initial total GFP between fits of the models accounting and not accounting for repression for repressions r1 and r2 (main analysis),
- Four hypothesis tests comparing estimated single-cell parameters between repressions r1 and r2 (main analysis),
- Two hypothesis tests comparing initial total GFP between fits of the models accounting and not accounting for repression for repressions r1 and r2 (replicate analysis),
- Four hypothesis tests comparing estimated single-cell parameters between repressions r1 and r2 (replicate analysis).

Comparison of estimated single-cell parameters between repressions r1 and r2

We ignored all traces that were well described by a model not accounting for repression (uninduced cells) and focused the statistical analysis on the total GFP traces for which the model accounting for repression gave a considerably better fit (see the Model selection, Figs 3B–3C

and S1D). For those total GFP traces, we compared the estimated single-cell parameters of initial total GFP, GFP_0 , repression response delay t_{delay} , production and degradation rates, r_{prod} and r_{deg} , for repressions r1 and r2. We performed a two-sided paired sign test on the estimated single-cell parameters of paired mother cells in repressions r1 and r2 (Figs 4 and S1G). To avoid false-positive outcomes, we used the Bonferroni correction, which modified the significance level of $\alpha = 0.05$ by the total number of tested null hypotheses m to $\alpha' = \alpha/m$, with $m = 12$.

Comparison of estimated single-cell parameters between the time-independent and time-dependent noise models

We compared the estimated single-cell parameters using a paired-sample t-test and corrected for false-positive outcomes by applying the Bonferroni correction with $m = 8$ (one null hypothesis for each estimated single-cell parameter = 4, for both repressions r1 and r2 = 8).

Implementation and data/code availability

The data and MATLAB code corresponding to this manuscript are available under <https://github.com/marrlab/Gal1repression>.

Supporting information

S1 Fig. Shortened repression delay in repression r2 at the single-cell level for replicate experiment.

(A) Single-cell traces of total Gal1-GFP fluorescence signal across two inductions i1 and i2 (gray) and repressions r0, r1, and r2 (blue). (B) Comparison of means of the normalized total Gal1-GFP fluorescence signals of repressions r1 and r2. (C) Single-cell traces of total GFP signal adjusted for dilution for the first two hours of repression r1 (left) and repression r2 (right). Time to maximal mean total GFP is 24 min shorter in repression r2, where mean total GFP is indicated by the dotted line and the maximal mean total GFP is highlighted by the dot. Bootstrap (10^5) samples were drawn to generate mean \pm std. (D) Ten exemplary total GFP traces (dotted lines) and best fits (solid lines) for repressions r1 (left) and r2 (right). GFP traces best fitted with a model accounting for repression are shown in black and fits of total GFP traces best fitted with a model not accounting for repression are shown in red. (E) The median initial total GFP, GFP_0 , is higher in traces better fitted by the model accounting for repression (black) than in traces better fitted by the model not accounting for repression (red). This confirms that the model accounting for repression fits induced cells better, while the model not accounting for repression fits uninduced cells. The number of cells and percentages of all GFP traces best fitted by the model accounting for repression and model not accounting for repression are shown. (F) Time to maximal mean total GFP is decreased in repression r2 for cells with repression kinetics (0.97 ± 0.08 vs. 0.88 ± 0.04). Bootstrap (10^5) samples of the cells with repression kinetics were drawn to generate mean \pm std. (G) Comparison of paired estimated single-cell parameters of cells with repression kinetics of repression r1 and r2 shows that the median initial total GFP, GFP_0 , and median repression delay, t_{delay} , are significantly different ($p = 4.6 \cdot 10^{-7}$ and $p = 3.3 \cdot 10^{-3}$, respectively, two-sided paired sign test correcting for multiple testing with Bonferroni correction, $m = 12$, and the number of paired cells = 31), with median GFP_0 increased and median t_{delay} decreased (median t_{delay} values of 0.83 and 0.53 for r1 and r2, respectively). (TIF)

S2 Fig. Estimated repression response delays are comparable between time-independent and time-dependent noise models for repressions r1 and r2. (A-B) The time-dependent

noise model detects fewer cells with repression kinetics in repressions r1 (A) and r2 (B). However, the set of repressor cells detected across the two noise models is largely overlapping. (C-D) The estimated single-cell repression response delays are comparable (paired-sample t-test with Bonferroni correction, $m = 8$, and the number of paired cells = 40 for r1 and = 162 for r2) across both noise models for both repressions r1 (C) and r2 (D). (TIF)

S3 Fig. Estimated repression response delay is independent of GFP degradation rate. Total GFP traces of four simulated cells with varying degradation rates, $r_{\text{deg}} = 0.5 \text{ h}^{-1}$ (fast), $r_{\text{deg}} = 0.2 \text{ h}^{-1}$ (slow), $r_{\text{deg}} = 0.04 \text{ h}^{-1}$ (super slow), and $r_{\text{deg}} = 0 \text{ h}^{-1}$ (none) using Gillespie's stochastic simulation algorithm and simulation parameters $\text{GFP}_0 = 100$, $t_{\text{delay}} = 0.5 \text{ h}$, $r_{\text{prod}} = 100 \text{ h}^{-1}$, and $\sigma = 5$. Performing parameter estimation and model selection, only the cell with no degradation (right) is misclassified as a cell without repression kinetics (with $\text{BIC}_{\text{no repression}} < \text{BIC}_{\text{repression}}$). The repression response delay is correctly estimated to approximately 0.5 h for all simulated cells independent of the GFP degradation rate. (TIF)

Acknowledgments

L.S. is especially grateful to the Technical University of Munich's Department of Mathematics, whose Entrepreneurial Award (within the Program "Global Challenges for Women in Math Science") contributed to the completion of this project. L.S. acknowledges further support by the Add-on Fellowship for Interdisciplinary Life Science of the Joachim Herz Foundation.

Author Contributions

Conceptualization: Poonam Bheda, Robert Schneider, Carsten Marr.

Data curation: Lea Schuh, Maria Bordukova.

Formal analysis: Lea Schuh.

Funding acquisition: Robert Schneider, Carsten Marr.

Investigation: Lea Schuh, Igor Kukhtevich.

Methodology: Lea Schuh.

Project administration: Robert Schneider, Carsten Marr.

Resources: Robert Schneider, Carsten Marr.

Software: Lea Schuh.

Supervision: Poonam Bheda, Robert Schneider, Carsten Marr.

Validation: Lea Schuh.

Visualization: Lea Schuh.

Writing – original draft: Lea Schuh.

Writing – review & editing: Lea Schuh, Melanie Schulz, Robert Schneider, Carsten Marr.

References

1. Acar M, Becskei A, van Oudenaarden A. Enhancement of cellular memory by reducing stochastic transitions. *Nature*. 2005; 435: 228–232. <https://doi.org/10.1038/nature03524> PMID: 15889097

2. Liu K, Li Y, Prabhu V, Young L, Becker KG, Munson PJ, et al. Augmentation in Expression of Activation-Induced Genes Differentiates Memory from Naive CD4 T Cells and Is a Molecular Mechanism for Enhanced Cellular Response of Memory CD4 T Cells. *The Journal of Immunology*. 2001. pp. 7335–7344. <https://doi.org/10.4049/jimmunol.166.12.7335> PMID: 11390484
3. Araki Y, Wang Z, Zang C, Wood WH 3rd, Schones D, Cui K, et al. Genome-wide analysis of histone methylation reveals chromatin state-based regulation of gene transcription and function of memory CD8+ T cells. *Immunity*. 2009; 30: 912–925. <https://doi.org/10.1016/j.immuni.2009.05.006> PMID: 19523850
4. Mobbs CV. Glucose-Induced Transcriptional Hysteresis: Role in Obesity, Metabolic Memory, Diabetes, and Aging. *Front Endocrinol*. 2018; 9: 232.
5. Villeneuve LM, Reddy MA, Natarajan R. Epigenetics: deciphering its role in diabetes and its chronic complications. *Clin Exp Pharmacol Physiol*. 2011; 38: 451–459. <https://doi.org/10.1111/j.1440-1681.2011.05497.x> PMID: 21309809
6. Mobbs CV. Genetic influences on glucose neurotoxicity, aging, and diabetes: a possible role for glucose hysteresis. *Genetica*. 1993; 91: 239–253. <https://doi.org/10.1007/BF01436001> PMID: 8125273
7. Courey AJ, Jia S. Transcriptional repression: the long and the short of it. *Genes Dev*. 2001; 15: 2786–2796. <https://doi.org/10.1101/gad.939601> PMID: 11691830
8. Maldonado E, Hampsey M, Reinberg D. Repression: targeting the heart of the matter. *Cell*. 1999; 99: 455–458. [https://doi.org/10.1016/s0092-8674\(00\)81533-0](https://doi.org/10.1016/s0092-8674(00)81533-0) PMID: 10589673
9. Kosterlitz HW. The fermentation of galactose and galactose-1-phosphate. *Biochemical Journal*. 1943. pp. 322–326. <https://doi.org/10.1042/bj0370322> PMID: 16747644
10. Leloir LF. The enzymatic transformation of uridine diphosphate glucose into a galactose derivative. *Archives of Biochemistry and Biophysics*. 1951. pp. 186–190. [https://doi.org/10.1016/0003-9861\(51\)90096-3](https://doi.org/10.1016/0003-9861(51)90096-3) PMID: 14885999
11. Douglas HC, Hawthorne DC. Enzymatic expression and genetic linkage of genes controlling galactose utilization in *Saccharomyces*. *Genetics*. 1964. pp. 837–844. <https://doi.org/10.1093/genetics/49.5.837> PMID: 14158615
12. Brickner DG, Cajigas I, Fondufe-Mittendorf Y, Ahmed S, Lee P-C, Widom J, et al. H2A.Z-Mediated Localization of Genes at the Nuclear Periphery Confers Epigenetic Memory of Previous Transcriptional State. *PLoS Biology*. 2007. p. e81. <https://doi.org/10.1371/journal.pbio.0050081> PMID: 17373856
13. Kundu S, Horn PJ, Peterson CL. SWI/SNF is required for transcriptional memory at the yeast GAL gene cluster. *Genes & Development*. 2007. pp. 997–1004. <https://doi.org/10.1101/gad.1506607> PMID: 17438002
14. Sood V, Cajigas I, D'Urso A, Light WH, Brickner JH. Epigenetic Transcriptional Memory of GAL Genes Depends on Growth in Glucose and the Tup1 Transcription Factor in *Saccharomyces cerevisiae*. *Genetics*. 2017. pp. 1895–1907. <https://doi.org/10.1534/genetics.117.201632> PMID: 28607146
15. Stockwell SR, Landry CR, Rifkin SA. The yeast galactose network as a quantitative model for cellular memory. *Mol Biosyst*. 2015; 11: 28–37. <https://doi.org/10.1039/c4mb00448e> PMID: 25328105
16. Zacharioudakis I, Gligoris T, Tzamaras D. A yeast catabolic enzyme controls transcriptional memory. *Curr Biol*. 2007; 17: 2041–2046. <https://doi.org/10.1016/j.cub.2007.10.044> PMID: 17997309
17. Bheda P, Aguilar-Gómez D, Becker NB, Becker J, Stavrou E, Kukhtevich I, et al. Single-Cell Tracing Dissects Regulation of Maintenance and Inheritance of Transcriptional Reinduction Memory. *Mol Cell*. 2020; 78: 915–925.e7. <https://doi.org/10.1016/j.molcel.2020.04.016> PMID: 32392469
18. Lee BB, Choi A, Kim JH, Jun Y, Woo H, Ha SD, et al. Rpd3L HDAC links H3K4me3 to transcriptional repression memory. *Nucleic Acids Res*. 2018; 46: 8261–8274. <https://doi.org/10.1093/nar/gky573> PMID: 29982589
19. Goulev Y, Morlot S, Matifas A, Huang B, Molin M, Toledano MB, et al. Nonlinear feedback drives homeostatic plasticity in HO stress response. *Elife*. 2017;6. <https://doi.org/10.7554/eLife.23971> PMID: 28418333
20. Dietler N, Minder M, Gligorovski V, Economou AM, Joly DAHL, Sadeghi A, et al. A convolutional neural network segments yeast microscopy images with high accuracy. *Nat Commun*. 2020; 11: 5723. <https://doi.org/10.1038/s41467-020-19557-4> PMID: 33184262
21. Padovani F, Mairhormann B, Falter-Braun P, Lengfeld J, Schmoller KM. Cell-ACDC: a user-friendly toolset embedding state-of-the-art neural networks for segmentation, tracking and cell cycle annotations of live-cell imaging data. <https://doi.org/10.1101/2021.09.28.462199>
22. Johnston M. A model fungal gene regulatory mechanism: the GAL genes of *Saccharomyces cerevisiae*. *Microbiological Reviews*. 1987. pp. 458–476. <https://doi.org/10.1128/mr.51.4.458-476.1987> PMID: 2830478

23. Schuh L, Loos C, Pokrovsky D, Imhof A, Rupp RAW, Marr C. H4K20 Methylation Is Differently Regulated by Dilution and Demethylation in Proliferating and Cell-Cycle-Arrested *Xenopus* Embryos. *Cell Syst.* 2020; 11: 653–662.e8. <https://doi.org/10.1016/j.cels.2020.11.003> PMID: 33296683
24. Alabert C, Loos C, Voelker-Albert M, Graziano S, Forné I, Reveron-Gomez N, et al. Domain Model Explains Propagation Dynamics and Stability of Histone H3K27 and H3K36 Methylation Landscapes. *Cell Rep.* 2020; 30: 1223–1234.e8. <https://doi.org/10.1016/j.celrep.2019.12.060> PMID: 31995760
25. Bast L, Buck MC, Hecker JS, Oostendorp RAJ, Götze KS, Marr C. Computational modeling of stem and progenitor cell kinetics identifies plausible hematopoietic lineage hierarchies. *iScience.* 2021; 24: 102120. <https://doi.org/10.1016/j.isci.2021.102120> PMID: 33665548
26. Li J, Wang S, VanDusen WJ, Schultz LD, George HA, Herber WK, et al. Green fluorescent protein in *Saccharomyces cerevisiae*: real-time studies of the GAL1 promoter. *Biotechnol Bioeng.* 2000; 70: 187–196. [https://doi.org/10.1002/1097-0290\(20001020\)70:2<187::aid-bit8>3.0.co;2-h](https://doi.org/10.1002/1097-0290(20001020)70:2<187::aid-bit8>3.0.co;2-h) PMID: 10972930
27. Lang GI, Botstein D. A test of the coordinated expression hypothesis for the origin and maintenance of the GAL cluster in yeast. *PLoS One.* 2011; 6: e25290. <https://doi.org/10.1371/journal.pone.0025290> PMID: 21966486
28. Niedenthal RK, Riles L, Johnston M, Hegemann JH. Green fluorescent protein as a marker for gene expression and subcellular localization in budding yeast. *Yeast.* 1996; 12: 773–786. [https://doi.org/10.1002/\(SICI\)1097-0061\(19960630\)12:8%3C773::AID-YEA972%3E3.0.CO;2-L](https://doi.org/10.1002/(SICI)1097-0061(19960630)12:8%3C773::AID-YEA972%3E3.0.CO;2-L) PMID: 8813764
29. Biggar SR. Cell signaling can direct either binary or graded transcriptional responses. *The EMBO Journal.* 2001. pp. 3167–3176. <https://doi.org/10.1093/emboj/20.12.3167> PMID: 11406593
30. Song R, Peng W, Liu P, Acar M. A cell size- and cell cycle-aware stochastic model for predicting time-dynamic gene network activity in individual cells. *BMC Syst Biol.* 2015; 9: 91. <https://doi.org/10.1186/s12918-015-0240-5> PMID: 26646617
31. Maier C, Loos C, Hasenauer J. Robust parameter estimation for dynamical systems from outlier-corrupted data. *Bioinformatics.* 2017; 33: 718–725. <https://doi.org/10.1093/bioinformatics/btw703> PMID: 28062444
32. Komorowski M, Finkenstädt B, Harper CV, Rand DA. Bayesian inference of biochemical kinetic parameters using the linear noise approximation. *BMC Bioinformatics.* 2009; 10: 343. <https://doi.org/10.1186/1471-2105-10-343> PMID: 19840370
33. Munsky B, Li G, Fox ZR, Shepherd DP, Neuert G. Distribution shapes govern the discovery of predictive models for gene regulation. *Proc Natl Acad Sci U S A.* 2018; 115: 7533–7538. <https://doi.org/10.1073/pnas.1804060115> PMID: 29959206
34. Kukhtevich IV, Lohrberg N, Padovani F, Schneider R, Schmoller KM. Cell size sets the diameter of the budding yeast contractile ring. *Nat Commun.* 2020; 11: 2952. <https://doi.org/10.1038/s41467-020-16764-x> PMID: 32528053
35. Hass H, Loos C, Raimúndez-Álvarez E, Timmer J, Hasenauer J, Kreutz C. Benchmark problems for dynamic modeling of intracellular processes. *Bioinformatics.* 2019. pp. 3073–3082. <https://doi.org/10.1093/bioinformatics/btz020> PMID: 30624608
36. Stapor P, Weindl D, Ballnus B, Hug S, Loos C, Fiedler A, et al. PESTO: Parameter ESTimation TOolbox. *Bioinformatics.* 2018; 34: 705–707. <https://doi.org/10.1093/bioinformatics/btx676> PMID: 29069312
37. Schwarz G. Estimating the Dimension of a Model. *The Annals of Statistics.* 1978. <https://doi.org/10.1214/aos/1176344136>