A precisely-titratable, variation-suppressed transcriptional controller to enable genetic discovery

Aslı Azizoğlu[a], Roger Brent[b,1], and Fabian Rudolf[a,1]

[a]Computational Systems Biology and Swiss Institute of Bioinformatics, ETH Zurich D-BSSE, Mattenstr. 26, 4058 Basel, CH
[b]Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA
[1]Correspondence to rbrent@fredhutch.org or fabian.rudolf@bsse.ethz.ch

Abstract

For more than 65 years, means to express genes conditionally into phenotype have remained central to biological experimentation and discovery. Current experimental methods typically enable either on/off gene expression or growth-condition-specific, imprecisely controlled graded expression in response to exogenous inducers. Here we describe a "well-tempered" controller, one that allows precise and graded conditional expression of genes in *Saccharomyces cerevisiae*. This system, WTC_{846}, appropriates genetic and design elements from bacterial, eukaryotic, and engineered systems. Its main autorepressing circuitry relies on two identical instances of a strong, growth rate regulated promoter P_{TDH3} engineered to be repressible by the TetR protein. This allows for titration of inducer concentration to drive variation-reduced expression of regulated genes across the entire range of the proteome. To reduce basal expression to zero, a TetR-Tup1 fusion protein is expressed at a low, constitutive level. We showed that strains carrying conditional, WTC_{846} allelic forms of genes encoding stable low abundance proteins (eg. Cdc42), unstable low abundance proteins (eg. Ipl1), and highly expressed proteins including the glycolytic enzyme Tpi1 recapitulated known knockout and overexpression phenotypes. Strains bearing $WTC_{846:Cdc20}$ alleles allowed inducer controlled cell cycle synchronization of batch cultures and release. Chemical titration of WTC_{846} allelic strains of CDC28, TOR1, PBR1 and PMA1 brought about distinct, gene dosage dependent controlled growth rates, and highly penetrant morphological phenotypes expected for the different gene doses. The ability to generate WTC_{846} controlled, "expression-clamped" genes may operationally define a new kind of conditional allele whose cell-to-cell variation in expression is minimized and whose effective dosage is under the experimenter's control. We expect Well-tempered $Controller_{846}$ (WTC_{846}) strains to find use in assessment of phenotypes now incompletely penetrant due to variable dosage of the causative gene product, in targeted cell biological experimentation, and in genome-wide studies such as gene by gene epistasis screens. In higher cells, we hope that implementation of titratable, expression clamped control logic via mammalian specific transcription elements will enable experiments now impossible due to cell-to-cell variation and imprecise control.

Introduction

Since the first use of nonsense suppressing *Escherichia coli* strains to map and study function of genes in bacteriophage T4[1], methods to express genes conditionally into phenotype have remained central to biological discovery. After the advent of recombinant DNA methods, conditional expression of genes into proteins, for example by derepression of lac promoter derivatives[2], became important for production of industrial products, and conditional expression began to be developed for medical use in gene therapy[3]. In 2019, for biological discovery, approaches to conditional expression in wide use include but are not limited to targeted degradation of proteins already expressed[4], development of lines of cells and organisms containing transgenes activated by chimeric activators controlled by promoters whose expression is temporally and spatially restricted[5], hundreds of approaches based on production of DNA rearrangements by phage-derived site-specific recombination[6], and triggered induction of engineered genes by chimeric transcription regulators, including those that make use of derivatives of the tetracycline repressor encoded by the bacterial repressor Tn10[7, 8]. These approaches are typically all-or-none, in the sense that they were not developed to bring about expression of intermediate levels of protein and the observations they enable are often qualitative.

But it has long been recognized that the ability to experimentally titrate protein dosage can provide additional insight into gene function. For example, in *E. coli*, P_{lac} controlled expression of the bacteriophage l and cro gene products was key to understanding how changes in the level of those proteins regulated the phage's decision to undergo lytic or lysogenic growth[9]. In *S. cerevisiae*, widely used contemporaneous methods of titrating protein dosage rely on metabolite- and metal- induced promoters, such as P_{GAL1}, P_{MET3}, and the copper induced P_{CUP1}[10], for which expression is controlled by culture of cells in different media, some Tn10 based activator systems[8] and various DNA binding domain-hER-activation based systems[11, 12], for which expression is induced by adding a chemical to the cell culture that causes the chimeric activator[13] to bind a site on an engineered promoter. These inducible systems rely on "activation by recruitment", in which the activator binds DNA upstream of the transcription start, and recruits general transcription factors and regulators of the pre-initiation complex (PIC). Downstream of the bound activator, at the "core promoter", the proteins that comprise the PIC are assembled often at a conserved TATA sequence, and recruit RNA polymerase II to induce transcription[14].

Titratable activator expression systems in current use in *S. cerevisiae* suffer from a number of drawbacks. First, small molecule induced systems typically have high basal expression[15, 8, 11]. Second, all carry a DNA binding chimeric activator[13] as a key component, and this component itself has effects on cellular function, either due to off-target binding[16] or interaction with endogenous proteins[17]. These issues are addressed by using activators with DNA binding domains not native to the host organism, and/or weaker activators and more sensitive induced promoters, but all activators still impact cellular function when induced. Finally, no current inducible activation system can achieve precise titration of protein dosage, because in the population, cell-to-cell variability of protein expression in response to the inducer is very high[18, 19, 11].

Here, we envisioned that an ideal conditional expression system to support qualitative and quantitative experimentation would: 1) function in all growth media and growth conditions, 2) be inducible by an exogenous chemical with minimal other effects on the cell, 3) manifest no background expression of the gene product in absence of inducer, allowing generation of null phenotypes, 4) enable a very large range of titratable expression 5) drive very high maximum expression, allowing generation of overexpression phenotypes, and 6) exhibit low cell-to-cell variability at a given output, facilitating detection of threshold phenotypes and other inferences of single cell behaviors from population responses.

To meet these criteria, an alternative to the tunable activation approach would be to engineer repression of transcription and inactivate the repressing molecule by titratable addition of an inducer. In evolved eukaryotic systems, repression, silencing, suppression and other kinds of negative regulation of gene expression is quite common. Most of these negative regulatory mechanisms have in common interference with the formation of the PIC, via modification of DNA, nucleosomes and other chromatin associated proteins, by changes in nucleosome positioning[20], and by direct interaction with activator proteins[21]. *A priori*, many engineered systems based on controlled lifting of such repression might cause other changes in the cell. But there is a second, relatively less common mechanism of evolved eukaryotic repression; steric exclusion of activating transcription factors[22, 23] and PIC components[24] by repressor proteins. Repression of engineered eukaryotic promoters by prokaryotic repressors[25, 26, 27] is clearly different from the most common mechanisms of eu-

karyotic repression[13]. The mechanism of this engineered repression is not known. It may also operate by steric exclusion, or by direct competition with eukaryotic activators for components of the transcription machinery, or by interference with interaction between activators and the transcription machinery. In any case, for TATA containing promoters, binding of prokaryotic repressors near the TATA sequence can bring about repression[25, 28, 29], presumably by interference with formation of the PIC.

Here we describe the development of prokaryotic repressor-based transcriptional controller of gene expression, Well-tempered Controller$_{846}$ (WTC$_{846}$). At its heart lies two instances of an engineered version of the strong constitutive P_{TDH3} promoter that contains seven TetR binding sites. The new promoter was generated by a campaign of experimental and selective steps, to maximize repression while preserving powerful transcription. This new promoter is functional in cells grown in all commonly used carbon and nitrogen sources. In the WTC$_{846}$ main autorepression loop, one instance of this promoter drives the expression of TetR, while a second instance drives synchronized control of the regulated gene. To abolish basal leakiness completely, a weak, constitutive promoter drives expression of the zeroing repressor, a TetR-Tup1 fusion protein. By varying the extracellular concentration of the inducer, WTC$_{846}$ dependent expression of the regulated gene can be precisely titrated in different growth media, over time and cell cycle stage, with low cell-to-cell variability, from undetectable levels to the level of one of the highest expressed proteins in the cell. These results define WTC$_{846}$ controlled genes as a new kind of conditional allele, i.e. "expression clamped" genes whose level of expression can be titrated across wide dosage levels with an extracellular input. We expect WTC$_{846}$ strains to find use in assessment of phenotypes now incompletely penetrant due to variable dosage of the causative gene product, in targeted cell biological experimentation, and in genome-wide studies such as gene by gene epistasis screens.

Results

Construction of a repressible P_{TDH3} promoter

We used TDH3 promoter (P_{TDH3}) as our starting point, which is one of the strongest constitutive yeast promoters driving the synthesis of 800,000 molecules of the glyceraldehyde-3-phoshpate dehydrogenase protein per cell [30]. The main regulation of P_{TDH3} is by three activating transcription factor binding sites in its Upstream Activating Region (UAS), one for Rap1 and two for Gcr1 [31]. It also has a well-defined TATA sequence for PIC assembly (Figure 1A). Placing flanking binding sites for λ repressor cI tightly around this TATA sequence of P_{TDH3} represses activity of the promoter [29]. We therefore constructed versions of P_{TDH3} where binding of the bacterial repressor TetR could potentially repress transcription through steric exclusion of Rap1, Gcr1 and/or the PIC. We placed TetR binding sites (tetO$_1$) directly up-and downstream of the TATA sequence (P_{2tet}), or in close proximity to the binding sites of the Rap1 and Gcr1 transcription factors in UAS (P_{3tet}), or a combination of both (P_{5tet}). We quantified basal leakiness from these promoters in strains where they drove the expression of a genomically integrated Citrine Fluorescent Protein, and P_{ACT1} drove the expression of a genomically integrated TetR (Table S1, Px$_{tet}$-CR). Maximum expression was quantified in otherwise isogenic strains that lacked TetR (Table S1, Px$_{tet}$-const) (Figure 1B). TATA sequence flanking (P_{2tet}) configuration was enough to create a TetR repressible promoter with 12-fold repression. Three tetO$_1$ placed in the UAS (P_{3tet}) yielded only 1.5-fold repression (Figure S1), while the combination of both (P_{5tet}) showed the same 12-fold repression as P_{2tet}, but lowered basal leakiness considerably.

Even small changes to the sequence of Transcription Factor Binding Sites (TFBS) in the UAS and the TATA sequence of P_{TDH3} effect promoter activity[32]. We therefore decided to optimize these sequences to increase maximum activity without altering repressed expression, such that fold-repression of P_{5tet} would increase. For these optimizations we used P_{4tet} as our starting point, which is a variant of P_{5tet} where one TATA sequence tetO$_1$ was removed. P_{4tet} has a higher basal expression level than P_{5tet}, allowing us to better observe subtle differences in basal leakiness (Figure 1C). We employed three different strategies. Firstly, since even single base mutations in the TFBS of the UAS of P_{TDH3} can reduce its activity [32], we slided the tetO$_1$ sites away from TFBS$_{RAP1}$ and the upstream TFBS$_{GCR1}$ by one base pair ($P_{4tet.1}$ and $P_{4tet.2}$), with the assumption that the endogenous binding sites were longer than we initially predicted. Second, we replaced the downstream TFBS$_{GCR1}$ with the new upstream TFBS$_{GCR1}$ ($P_{4tet.3}$), since the upstream TFBS$_{GCR1}$ was closer to the reported consensus sequence[33] and was likely to be more active. Third, we tested alternative TATA sequences[34]. Extending the TFBS$_{RAP1}$, replacing the downstream

$TFBS_{GCR1}$ and the alternative TATA sequencees ($P_{4tet.5}$ more than $P_{4tet.4}$) lead to increased activity, whereas the extended upstream $TFBS_{GCR1}$ did not. Next we implemented these optimizations on the P_{5tet}, to generate $P_{5tet.1}$ (Figure 1D). While this increased both maximum and repressed expression compared to P_{5tet}, the fold repression of $P_{5tet.1}$ was higher than P_{2tet} (15 vs. 12-fold), and repressed expression was around the same level as P_{2tet}. Therefore careful placement of TetR binding sites in the UAS, combined with optimized TFBS and TATA sequences increased the fold repression observed in our promoter.

To further increase the maximum expression from $P_{5tet.1}$ and reach P_{TDH3} levels without compromising repression, we took advantage of previous work showing that increasing the number of transcription factor binding sites in a promoter could increase its strength [11]. Accordingly, we created P_{7tet} by adding one optimized Rap1 binding site, and $P_{7tet.1}$ by adding one optimized Rap1 and one optimized Gcr1 binding site to $P_{5tet.1}$. We kept the same $tetO_1$ placements in these duplicated segments. As expected, repressed level did not change, while maximum expression increased. $P_{7tet.1}$ showed 20-fold repression of Citrine expression, with maximum expression comparable to that driven by the endogenous P_{TDH3}. We therefore decided to build our controller using $P_{7tet.1}$.

Optimized controller architecture expands the titratable range and reduces basal leakiness

An Autorepression (AR) architecture can be used to expand the titratable range of a transcriptional controller (i.e. the range of doses where adding more inducer actually leads to more protein expression) compared to a Constitutive Repression (CR) architecture[35](Figure 2A). This reduces the steepness of the dose response curve and makes it easier to precisely titrate intermediate expression levels. We therefore constructed a strain where both TetR and Citrine are expressed from $P_{7tet.1}$ to create TetR autorepression (Table S1, $P_{7tet.1}$-AR). We recorded a steady state dose response of anhydrotetracycline (aTc) vs Citrine expression curve after 7 hours of induction and fitted the median of the data with a log-logistic model. We compared the obtained dose response curve to a CR architecture (Figure 2B), where Citrine is expressed from $P_{7tet.1}$ and TetR from the P_{ACT1} (Table S1, $P_{7tet.1}$-CR). As expected, the dose response curve of the AR architecture was flattened compared to the one of the CR. The titratable range can be seen by looking at the range of doses where the slope of the dose response curve is non-zero, and it is indeed much larger for the AR controller (Figure 2C). The resulting flattened dose-response curve allows for increased precision of titration.

One important advantage of the AR architecture is its reduced cell-to-cell variability (CCV) [36, 35]. Since large cells posses more proteins[37], CCV can be quantified by a measure that determines how similar protein expression is between two cells of comparable size. In order to do this, we fit a robust linear fit through the Size vs. Fluorescence data obtained by flow cytometry of any given population of cells (Figure S2 for an example). Size was estimated by the vector of the SSC-H and FSC-W signals. We then used the residual standard deviation (RSD) as an estimate of CCV. Bootstrapping was used to quantify the stability of our measurement. We showed that this RSD-based measure can be used to quantify variability of any promoter (Figure S3). The RSD of a strain where Citrine is expressed constitutively from P_{TDH3} is 0.18. This indicates that 82% of CCV is explained by size, and 18% is attributable to measurement error and unexplained variation between isogenic cells of same size. As expected, the CR architecture shows increased CCV variability (up to 0.5) at intermediate protein expression levels (Figure 2D and S4), which is a known limitation [35]. On the other hand the AR architecture shows consistently low CCV (around 0.2) comparable to the strain where Citrine is constitutively expressed from P_{TDH3} (0.18). Therefore the AR architecture maintains the CCV across a controller's titratable range at a level comparable to that of the natural CCV of the P_{TDH3} promoter.

Despite its advantages, basal leakiness is an intrinsic feature of this AR architecture. Since TetR represses its own expression, there is always a basal amount of TetR required for repression. And since both TetR and the Gene of Interest (GOI) are expressed from the same promoter, there will always be a basal amount of GOI as well. One way around this problem is to combine the AR and CR architectures, such that TetR is additionally expressed from a constitutive promoter outside the autorepression loop. We call this additional repressor a "zeroing" repressor, and the resulting architecture Autorepression with Constitutive Zeroing (AR_0) (Figure 2A). We built AR_0 (Table S1, $P_{7tet.1}$-AR_0) by integrating P_{ACT1} expressed TetR onto the AR strain, and compared it to the other architectures using a dose response experiment (Figure 2B). AR_0 showed a basal Citrine expression of 4.1-fold over autofluorescence level, which is lower compared to 6.3-fold for the AR

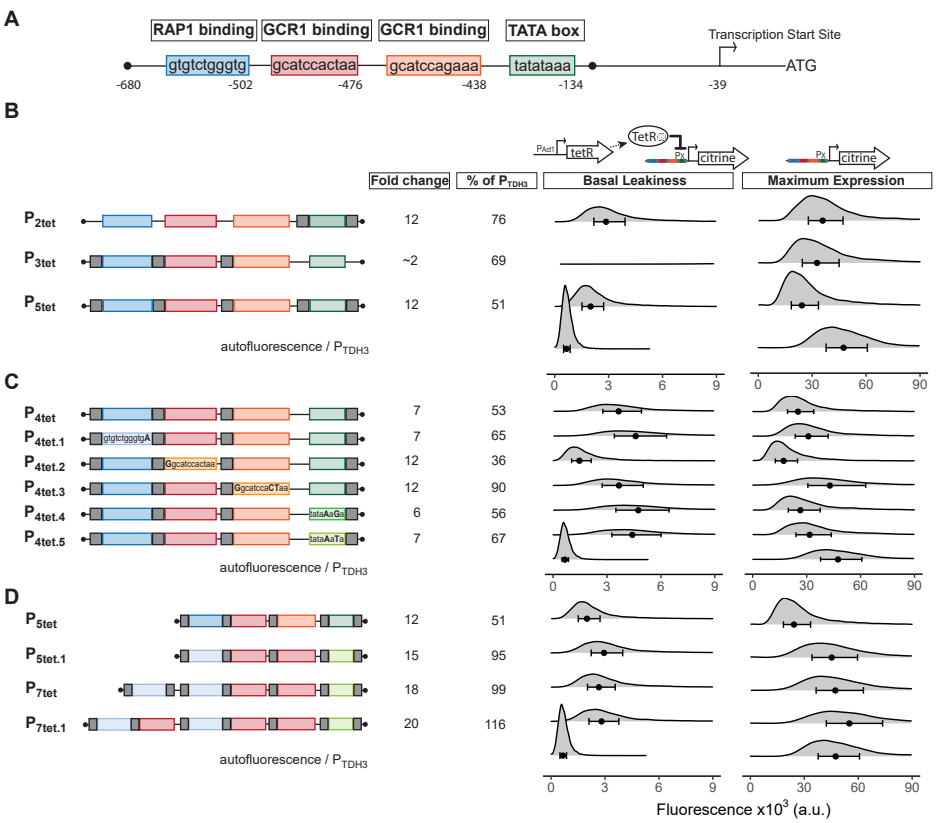

Figure 1: Optimized endogenous TF binding sites and better placement of TetR binding sites lead to higher maximum activity and increased fold-change of the TetR repressible versions of P_{TDH3}. A) Structure and main TF binding sites of P_{TDH3}. B) Different placements of $tetO_1$ sequences within P_{TDH3} C) Optimization of the endogenous transcription factor binding sites. The goal was to increase maximum expression from the promoter. D) The optimizations from (C) and additional binding sites for RAP1 and GCR1 were added to P_{5tet} to increase maximum expression. All promoter designs were assessed using flow cytometry. Maximum expression of Citrine is compared to a P_{TDH3} expressed Citrine strain. Basal Citrine level when TetR was expressed from P_{ACT1} is compared to autofluorescence of the parent strain. The median, Q1 and Q3 are indicated inside each density plot of ungated populations. Grey boxes indicate $tetO_1$ binding sites.

controller. Even though AR_0 has a higher induction threshold (first aTc concentration where fluorescence can be detected) compared to AR, it still had an increased titratable range compared to the CR architecture, and lower CCV (Figure 2C&D). In other words, the AR_0 architecture retained the favourable characteristics of an AR architecture, while reducing basal expression. Therefore we decided on the AR_0 architecture over the AR architecture for our final controller.

Hybrid repressor abolishes basal leakiness of $P_{7tet.1}$

We imagined two potential sources for the remaining basal leakiness from $P_{7tet.1}$. Either there isn't enough TetR for complete binding site occupancy, or once bound, TetR is inefficient at competing with endogenous TF. We therefore increased the nuclear concentration of TetR by adding extra NLS motifs, and created bigger repressors to increase competition efficiency. To this end we C-terminally fused the 26 kDA Glutathione S-Transferase (GST) and the 42 kDA Maltose Binding Protein (MBP) to TetR. We then measured basal Citrine expression from $P_{7tet.1}$ in presence of these repressor variants expressed by P_{ACT1} (strain names $P_{7tet.1}$-CR(X)) (Figure 3A). The results show that as the size of TetR increases, repression efficiency increases as well. Adding further NLS motifs reduces basal expression for smaller repressors, while it had no effect for larger variants. However, the basal leakiness of the best repressor, TetR-NLS-MBP, was still 2.2-fold above autofluorescence.

We reasoned that basal leakiness might be abolished by a hybrid repressor consisting of TetR and

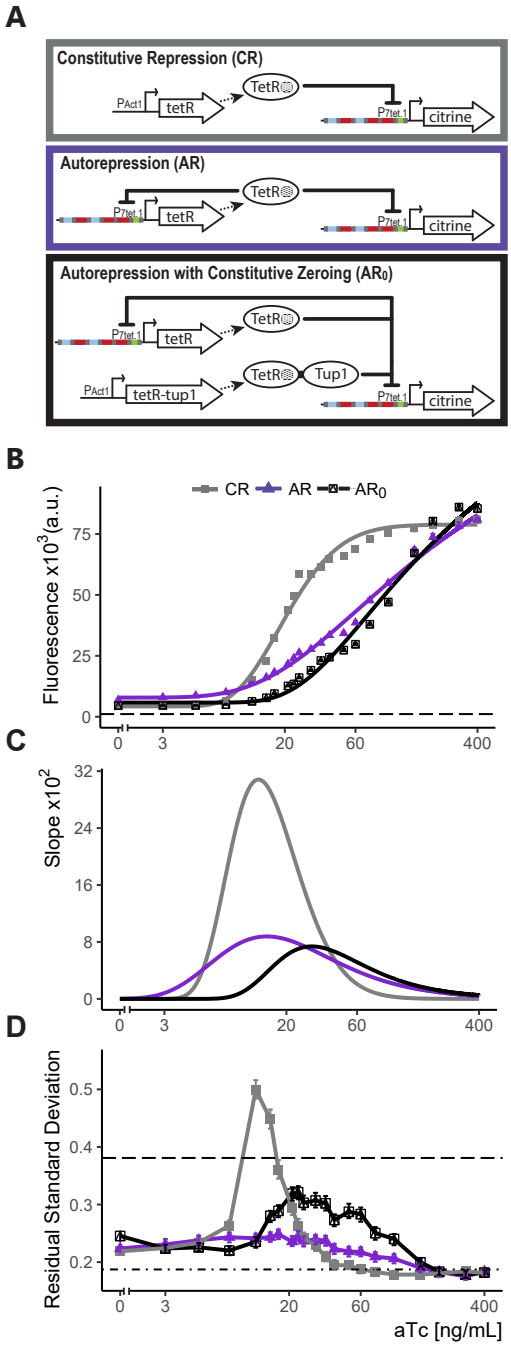

Figure 2: Comparison of different controller architectures shows that the AR_0 architecture is best in terms of low basal leakiness, high titratable range and low cell-to-cell variability. A) Schematics of different controller architectures. The chequered boxes indicate NLS. B) Dose response curves for the three different architectures at steady state, after 7 hours of induction. Symbols indicate the median of the experimentally measured population, lines are fitted. Dashed line indicates autofluorescence. C) Slopes of the dose response curves in (B). D) Cell-to-cell variability of the dose response curves in (B). Dot-dash line indicates cell-to-cell variability of the strain where Citrine is constitutively expressed from P_{TDH3}. All error bars indicate 95% confidence interval calculated using bootstrapping (n=1000).

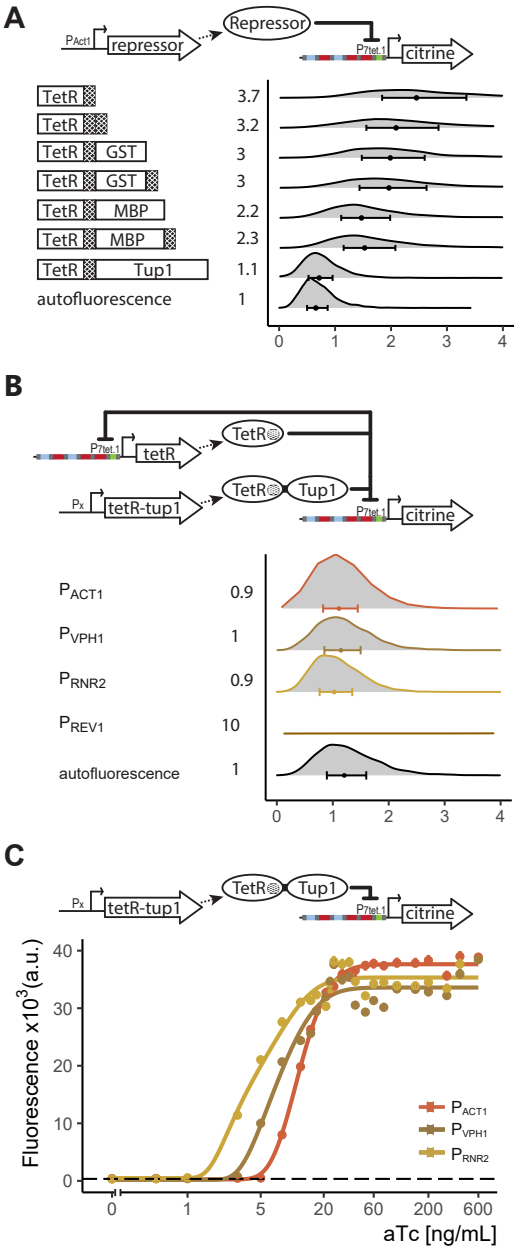

Figure 3: Optimized repressor TetR-nls-Tup1 abolishes basal expression from $P_{7tet.1}$. A) TetR was fused to proteins of increasing sizes, and/or additional NLS motifs and expressed from P_{ACT1}. Basal expression of Citrine from $P_{7tet.1}$ was measured in presence of these repressor variants using flow cytometry. Numbers to the left of the plot indicate fold expression over autofluorescence. Each chequered box indicates one NLS. B) Basal leakiness measurements for AR_0 strains where TetR-nls-Tup1 was expressed from promoters of decreasing strength. Numbers to the left of the plot indicate fold expression over autofluorescence. Symbols represent the median of the experimentally measured population, error bars span Q1 to Q3. P_{REV1} expressed population is not visible on the plot as it lies entirely above the upper limit of the x-axis. C) Dose response curves of strains where TetR-nls-Tup1 was expressed from different promoters within a CR architecture measured after 7 hours. Dashed line indicates autofluorescence control, symbols indicate the median of the experimentally measured population, lines are fitted. Error bars indicate 95% confidence interval calculated using bootstrapping (n=1000).

the 78kDa endogenous yeast transcriptional repressor Tup1. Tup1 complexes with Ssn6 (Cyc8) and this large corepressor complex represses not only by masking activation domains[38, 39], and excluding TBP[40], but also by actively changing chromatin structure and repositioning nucleosomes over the transcription start site[41, 39]. It is known that when Tup1 is artificially tethered to a minimal promoter, it can repress transcription[42]. Furthermore, it has been used to reduce basal leakiness in a dual TetR activator-repressor controller [15]. A TetR-nls-Tup1 fusion repressor abolished all basal expression from $P_{7tet.1}$ and achieved a tight controller, without increasing the time required for induction (Figure 3A and S5). Importantly, the TeR-nls-Tup1 fusion still requires all the $tetO_1$ sites in the UAS, as a P_{VPH1} expressed TetR-nls-Tup1 was unable to fully abolish basal leakiness in a from P_{2tet} (Figure S6).

Within the AR_0 architecture, the expression level of TetR-nls-Tup1 determines both the basal leakiness and the induction threshold. Induction above background occurs at lower aTc concentrations if the cell contains fewer TetR-NLS-Tup1 molecules to inactivate, but too few molecules would lead to basal leakiness. To test this point, we expressed TetR-nls-Tup1 from constitutive promoters of proteins with decreasing abundance[30] (P_{ACT1}, P_{VPH1}, P_{RNR2}, P_{REV1} and expressed it in strains bearing the $P_{7tet.1}$ expressed Citrine (Table S1, $P_{7tet.1}$-CR(X-Tup1)). We found that the P_{RNR2} driven expression gave the lowest induction threshold without basal leakiness, while the lower TetR-nls-Tup1 level in the P_{REV1} strain failed to meaningfully repress Citrine expression (Figure 3B&C and S7). Therefore our final controller was built such that $P_{7tet.1}$ controls the expression of the GOI (e.g. Citrine) and TetR, and P_{RNR2} controls the expression of TetR-nls-Tup1. Given its precise titratability, we call this system the Well-tempered Controller$_{846}$ (WTC_{846}) in honour of Bach's prelude&fugue No.1[1] (Figure 4A).

WTC_{846} fulfills all the criteria of an ideal transcriptional controller

We measured the time dependent dose response the $WTC_{846:Citrine}$ strain using flow cytometry to assess the basal expression, fold-change and dynamic behaviour of WTC_{846}. The results show that WTC_{846} behaves consistently across all time points. It induces rapidly within 30 minutes, a delay expected based on the maturation time of the fluorescent reporter, and reaches steady state within 7 hours, faster than our LexA-hER-activation domain system[11] (Figure 4A). Furthermore it has no detectable basal expression, a similar maximal expression as P_{TDH3}, and a large titratable range of around 1000-fold (Figure 4B,C). We confirmed using western blotting that both TetR and Citrine expression increase gradually in response to aTc, the controller has no detectable basal leakiness and has at least a 450-fold induction compared to the detection limit in YPD (Figure S8). We additionally confirmed with flow cytometry that Citrine expression can be precisely titrated independent of media condition (Figure S9). Importantly, WTC_{846} has no significant effect on normal cell physiology in any media condition tested, assessed using growth rates as a proxy (Figure S10).

We measured CCV of the $WTC_{846:Citrine}$ strain by flow cytometry as before using our residual standard deviation measure (Figure 4C). In YPD, CCV of WTC_{846} initially spikes up to 0.63, similar to P_{RNR2} expressed TetR-nls-Tup1 (0.67, Figure S11). From 8ng/mL aTc, it rapidly drops below autofluorescence and reaches the same level as P_{TDH3} expressed Citrine at 0.18. We interpret the two phases of the CCV as a reflection of the combined action of the two repressors in the cell. At low concentrations of inducer, the P_{RNR2} driven TetR-nls-Tup1 zeroing repressor predominates and CCV is high (Figure S12). This variability rapidly decreases upon induction of the AR controlled TetR repressor. Absolute RSD values between media conditions cannot be directly compared since the autofluorescence of cells grown in different media conditions differ greatly (Figure S9). However we observe similar trends in all conditions, such that CCV peaks around the same expression level in all conditions, and rapidly drops to the level of P_{TDH3} expressed Citrine (Figure S13). Overall WTC_{846} fulfills all the initially stated criteria for an ideal transcriptional control system.

WTC_{846} alleles represent a new class of conditional mutants with improved characteristics

We constructed strains in which WTC_{846} controlled expression of essential yeast genes (Figure 5A). The encoded proteins included stable and unstable low abundance cell cycle regulators (Cdc28, Cdc20, Cdc42), other low abundance proteins (Ipl1, Tor2, and the putative oxidoreducatase Pbr1), and a high abundance membrane protein (Pma1). Three strains where Cdc28, Tor2 and Pbr1 were being controlled survived without aTc, indicating a low amount of basal leakiness. We addressed this issue by lowering the translation efficiency of these genes by construct-

[1] Bach, Johann Sebastian, 1685-1750. The Well-Tempered Clavier. Book I : 24 Preludes and Fugues, BWV 846-857

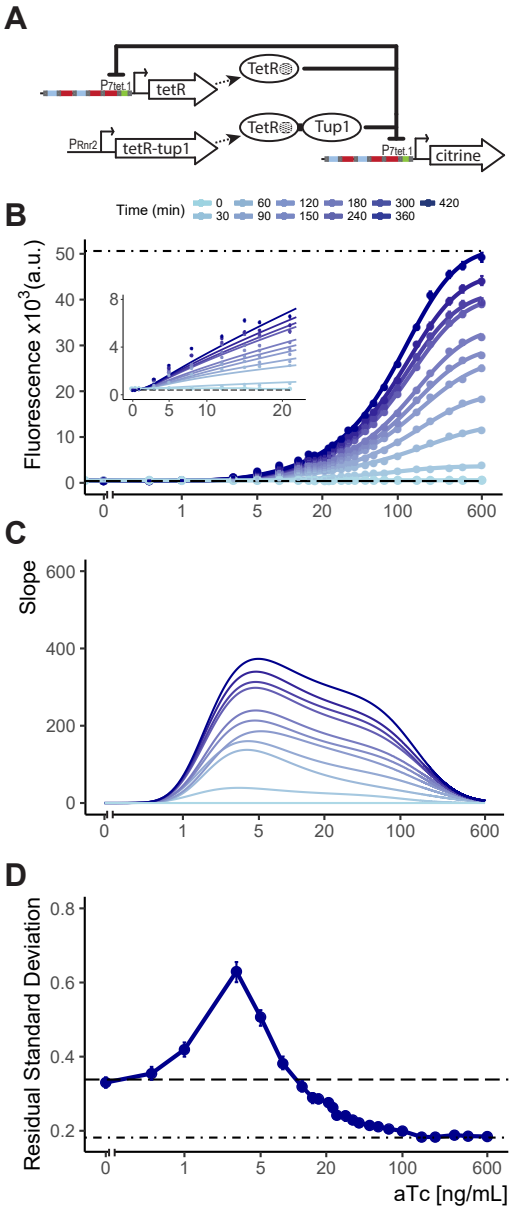

Figure 4: Characterization of the final WTC_{846} controller. A)TDDR of the final system. The median of autofluorescence (dashed line) and P_{TDH3} expression level (dot dash line) are marked. Symbols represent the median of the experimentally measured population, and the lines were fitted. Error bars indicate 95% confidence interval calculated using bootstrapping (n=1000). The inset zooms in to a part of the dose range tested. The chequered boxes indicate NLS. B) The slopes of the dose response curves in A. C) Cell-to-cell variability at t=7h.

ing different alleles with modified Kozak sequences (Table S1&S6))[43]. We then spotted these optimized Cdc28, Tor2 and Pbr1 strains along with the others on YPD, YPE, SD, SGlycerol and SDProline plates with and without aTc (Figure 5B and Figure S14). For all 7 strains, we observed complete null phenotypes without aTc, normal growth at low aTc concentrations, and for Ipl1 a decreased growth rate at higher levels of WTC_{846} induction.

Intermediate protein levels can be achieved with WTC_{846} alleles, which for essential or metabolic genes should yield gradually decreasing growth rates as protein level is decreased. To demonstrate this we used the alleles of Tor2, a low abundance, stable, essential protein necessary for both nutrient signalling and actin polarization [44], Pma1, an abundant, essential proton pump that regulates the internal pH of the cell[45], and Tpi1, a highly abundant, non-essential glycolytic enzyme[46]. We grew the strains $WTC_{846:\text{Tor2-K3}}$, $WTC_{846:\text{Tor2-K1}}$, $WTC_{846:\text{Pma1-K3}}$, and $WTC_{846:\text{Tpi1-K3}}$ with different levels of aTc, and measured growth with a growth reader (Figure 5C and Figure S15). We observed distinct growth rates per aTc concentration, and could reach the parent growth rate level in all cases. Importantly, $WTC_{846:\text{Pma1-K3}}$ strain showed not only a dose dependent growth regulation but also the expected, associated change in phenotype of cell separation deficiency at low concentrations as scored by DNA staining using flow cytometry as a readout (Figure S16). In the case of Tor2 expression, as low as 1ng/mL aTc was enough to induce some growth in the $WTC_{846:\text{Tor2-K3}}$ (Figure S15B), demonstrating the low induction threshold of WTC_{846}. At the same time, the $WTC_{846:\text{Tor2-K1}}$ strain with a more efficient Kozak sequence leads to a slow down in growth upon overexpression of Tor2 (Figure S15D), indicating that translation efficiency can be modulated to expand the already large titratable range of WTC_{846}.

Next, we confirmed that WTC_{846} alleles can demonstrate expected overexpression phenotypes using Ipl1. Previously, Ipl1 overexpression using the P_{GAL1} system did not show a phenotype[47], even though overexpression of the analogous Aurora kinase B in mammalian cells leads to increased mitosis duration and aberrant chromosome numbers [48]. However WTC_{846} mediated overexpression of Ipl1 is indeed lethal(Figure 5B). To check if the associated molecular phenotype of aberrant chromosome segregation and Spindle Assembly Checkpoint (SAC)-mediated arrest are indeed detectable, we grew $WTC_{846:\text{Ipl1-K2}}$ cells with a high concentration (400ng/mL) of aTc in YPD, and measured total DNA content to assess cell cycle progression.

Most cells arrested with 2n DNA content in the G2/M phase and a significant portion of the population showed aberrant chromosome numbers above 2n (Figure S17). This phenocopies the effect seen in mammalian cells, as well as the one seen in yeast cells when Ipl1 is co-overexpressed with another chromosome segregation machinery component, Sli15 [47].

We then attempted to verify that WTC_{846} steady state expression phenotypes showed low variability across its titratable range. To do so, we used $WTC_{846Whi5}$ haploid and diploid strains. Whi5 is a transcriptional repressor whose concentration needs to be diluted below a threshold for cells to commence synthesis of G1 specific cyclins, exit G1 arrest, and START [49]. The two strains and the parents were grown in SEthanol to late exponential phase with different aTc concentrations, and cell size was measured using flow cytometry. We selected the unbudded fraction using gaussian mixture modelling and calculated the coefficient of dispersion (CoD) on the size proxy of the obtained population (Figure S18A). The cell size monotonically increased over a 60-fold range of aTc input concentration for each strain (Figure S18). At the same time, the CoD showed a linearly increasing relationship with size for both haploid and diploid strains, which was in agreement with the CoD increase observed between the haploid and diploid parent strains (Figure 5D). Importantly, once the WTC_{846} regulated Whi5 strains reach the same size as either of the parent strains, CoD is at the same level as, or even lower than, the parent. This indicates that at intermediate dosages, WTC_{846} controlled Whi5 phenotype does not lead to increased CCV.

Finally we tested the utility of WTC_{846} for dynamic control of gene expression by using the $WTC_{846:Cdc20\text{-}K3}$ strain to synchronize cell cultures[50]. Cdc20 is an essential activator and specificity provider of Anaphase Promoting Complex (APC/C)[51]. Cells arrest in metaphase with large buds and 2n DNA content upon its depletion and go through the next cell cycle synchronously upon its reintroduction[52]. Synchronizing an entire batch culture of cells requires rapid and complete transcriptional shut-off in all cells within the population, followed by rapid and complete reintroduction of the gene with low variability in induction among different cells. We diluted exponentially growing cells into YPD media (pH 4) without aTc at time 0, and took samples for Sytox staining and flow cytometry analysis for DNA content. Within 480 minutes, the entire culture arrested at the G2/M phase with 2n DNA content (Figure 5E). Once all the cells in the culture had arrested with large buds, 600ng/mL aTc was added. This resulted in synchronous release into the next cell cycle within 35 minutes, demonstrating the rapid dynamics and low CCV of WTC_{846} alleles.

Discussion

The conditional expression of genes into phenotypes remains important for biological discovery. Many methods used historically, such as variation of protein activity via nonsense suppression and use of temperature sensitive alleles, do not facilitate titration of intermediate levels of gene activity. Many contemporary approaches for conditional expression work by placing transcription of the conditionally expressed genes under some kind of engineered regulation, but again, in these methods, conditional expression is typically all or nothing. Even titratable methods allowing variable expression in response to variable input have only a narrow input range and show high variation in protein expression in the individual controlled cells, and most also have known off-target effects including slowing of cell growth. In order to overcome these existing limitations, we developed and tested in *S. cerevisiae* Well-tempered $Controller_{846}$ (WTC_{846}), a "well-tempered", autorepression-based transcriptional controller of gene expression. It can precisely titrate protein levels across a large titratable range without any basal leakiness, has high maximum expression and low cell-to-cell variability, and is functional in different media conditions without any adverse effects on cell physiology.

Control of expression of genes into phenotypes remains important for biological discovery. Many methods used historically, such as variation of protein activity via nonsense suppression and use of temperature sensitive alleles, do not facilitate titration of intermediate levels of gene activity. Many contemporary approaches for conditional expression work by placing transcription of the conditionally expressed genes under some kind of engineered regulation, but again, in these methods, conditional expression is typically all or nothing. Even titratable methods allowing variable expression in response to variable input allow only a narrow range of controlling inputs and show high variation in protein expression in the individual controlled cells, and most also have known off-target effects including slowing of cell growth. In order to overcome these existing limitations, we developed and tested in *S. cerevisiae* WTC_{846}, a "well tempered" transcriptional controller of gene expression. WTC_{846}, is a complex autorepressing system that directs expression of regulated proteins to de-

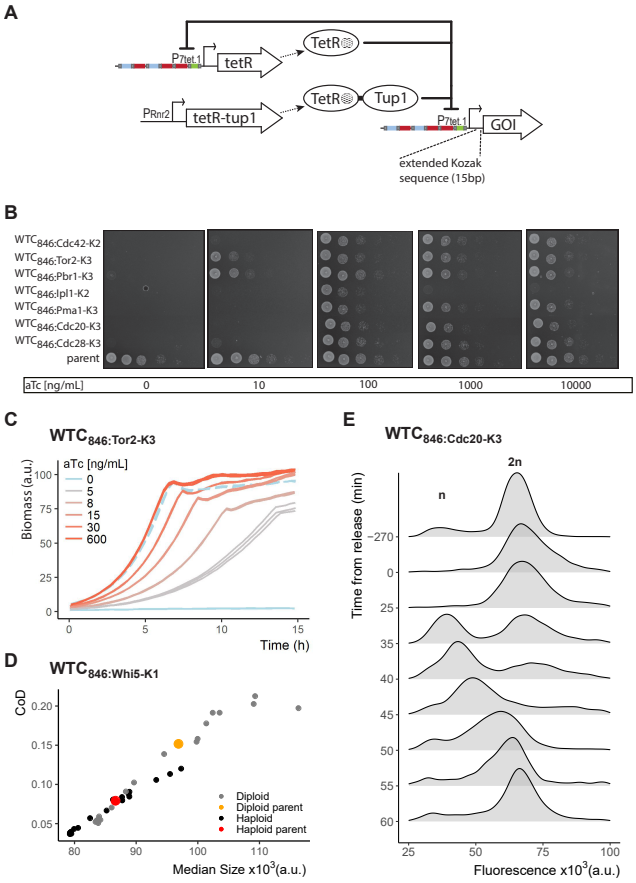

Figure 5: WTC_{846} can be used to reversibly abolish expression of, overexpress and titrate essential genes with low cell-to-cell variability. A) The chequered boxes indicate NLS. B) Spottings of conditional WTC_{846} alleles of essential genes without or with increasing concentrations of aTc, YPD plates. C) Growth curves (in triplicates per aTc concentration) of $WTC_{846:\text{Tor2-K3}}$ in YPD. D) Median cell size and Coefficient of Dispersion of the haploid $WTC_{846:\text{Whi5-K1}}$ and the diploid $WTC_{846:\text{Whi5-K1dip}}$ strain populations as Whi5 was titrated using aTc. Cells were analyzed using flow cytometry. E) Arrest&release using the $WTC_{846:\text{Cdc20-K3}}$ strain. 0 indicates the time of aTc addition to the arrested culture. Cells were stained with Sytox and analyzed with flow cytometry. 10000 cells per time point were recorded.

sired intracellular concentrations, with no basal expression, with tunable expression across a large range of inducer inputs, with high maximum expression, with low cell-to-cell variability, and with no adverse affects on cell growth.

WTC_{846}'s main autorepressing control circuit relies on two instances of an engineered version of the strong constitutive P_{TDH3} promoter. We engineered this promoter, $P_{7tet.1}$, to be repressible by derivatives of the bacterial TetR protein. In yeast, prokaryotic repressors including LexA[25], λcI[29] and Tn10 TetR[53] can repress transcription when bound to sites near a promoter's UAS, upstream of its TATA sequence, or flanking its TATA sequence. $P_{7tet.1}$ contains seven TetR binding sites, five near the binding sites for RAP1 and GCR1 in the UAS, two flanking the TATA sequence. The mechanism by which prokaryotic proteins repress engineered eukaryotic promoters is not known. It is possible that prokaryotic repressors might simply sterically block binding of proteins important for transcription, or compete with eukaryotic activators for interaction with components of the transcription machinery, or interfere with DNA looping mediated interaction between DNA bound activators and proteins near the transcription startpoint. Here, we found that while placement of repressor binding sites next to those for the UAS binding does result in repression, placement of two TetR binding sites flanking the TATA sequence was imperative for strong repression. Recent structural (cryo-EM) data of DNA bound transcription proteins[54] is certainly consistent with the idea that repressor bound near the TATA sequence might sterically block assembly of the activating transcription factors TBP and TFIID.

In WTC_{846}, $P_{7tet.1}$ drives expression of TetR, as well as expression of the controlled gene, so that TetR represses its own synthesis as well. This kind of autorepressing architecture is common in evolved bacterial systems (for example, in histidine utilization[55], in the LexA/ SOS regulon[56], and in the Tn10 transposon from which this TetR protein was derived[57]. It is less common in engineered eukaryotic systems, but, when used to repress a second controlled gene, autorepressing TetR control expands the dynamic titratable range of inducer input doses, and also reduces cell-to-cell variability[35]. In autorepressing TetR control, the ability to set gene expression depends on a "comparator-adjuster"[58], in which concentration of the inducer, anhydrotetracycline (aTc), defines a set-point, and molecules of aTc form complexes with some of the TetR molecules in the cell and inactivate them. Through this binding reaction, the concentrations of aTc and TetR and are continuously

compared, and total TetR expression continually adjusted by free TetR, so that total TetR concentration is maintained at the set-point determined by the aTc. This circuitry thus operationally defines WTC_{846} as an "expression clamp", one that results in low cell-to-cell variability in the cellular concentration of TetR and of the synchronized controlled protein, so that expression in single cells closely tracks the population average.

There is, of course, considerable cell-to-cell variability in absolute expressed protein level. Generally, cell to cell variation is reported to inversely correlate with abundance, is specific to the growth rate of the population, or it depends on the growth rate and the mitochondrial activity of the individual cells[59, 60, 61]. Most of the observed variation, between 50% and 80% depending on the growth condition, of WTC_{846} is due to difference in cell size: larger cells express more protein. We believe that some of the remaining variation may be a due to the diminished precision/diminished signal to noise of fluorescence intensity from small cells. Other contributions to the remaining variation may include growth rate and mitochondrial activity of individual cells, but do neither show a linear relationship with the population growth rate nor with the protein abundance. A Autorepression (AR) architecture is able to reduce cell-to-cell variability [36, 35]. We observed that the condition specific, minimal variation during a dose response with WTC_{846} is identical to that of the starting promoter P_{TDH3} measured in the same condition. Therefore, the natural variability of the starting promoter sets a lower bound which cannot be reduced further with the AR architecture.

In integral feedback control, the difference over time between the controlled variable (here, rate of TetR synthesis) from the desired target or set-point (here, aTc concentration) is corrected by negative feedback (here, repression) to counteract the deviation and reduce it to zero. Because TetR synthesis depends on a pool of intracellular TetR molecules that have accumulated over time, feedback control in WTC_{846} is by definition integral. Controlled systems that return to set-points when perturbed are said to exhibit perfect adaptation, and perfect adaptation shown by biological systems that operate via stochastic interactions among integer numbers of proteins is by definition robust to those random events[62]. By these definitions, it can be said that WTC_{846} behaves as an integral feedback controller that is likely to exhibit robust perfect adaptation.

In the absence of inducer, WTC_{846}-controlled genes show no basal expression. Suppression of basal expression depends on a second, constitutively expressed "zeroing" TetR derivative. This zeroing repressor is a fusion of TetR, a nuclear localization sequence, and the yeast transcriptional repressor Tup1 whose expression is driven by the weak constitutive promoter P_{RNR2}. Native Tup1 repressor is a 78kD protein that forms a complex comprised of four Tup1 subunits and a single 107kD Cyc8 subunit[63]. This large protein complex carries out repression by complexing with the histone modifying proteins Rpd3 and Hda1, may also directly position a nucleosome over the +1 transcription startpoint, or may spatially interfere with binding[64]. It is not clear whether the TetR-Tup1 mediated repression of in WTC_{846} works by these means and/or by the additional mechanisms by which prokaryotic TetR alone works, including sterically blocking access of necessary proteins to the DNA. However, zeroing repression by TetR-Tup1 depends on the same binding sites near the P_{TDH3} TATA sequence and its UAS, suggesting that part of its action may indeed be due to sterically blocking TBP binding. This creates the Autorepression with Constitutive Zeroing (AR_0) architecture, where there is no observable basal leakiness, but also a slight increase in cell-to-cell variability at low protein expression levels compared to the AR architecture, due to the lack of autorepression.

We used WTC_{846} to construct conditional alleles that are functional without basal leakiness in different growth conditions, and we used such alleles to generate graded, penetrant growth and morphological phenotypes for the low abundance proteins Cdc28, Ipl1, Tor2, Cdc20, and Pbr1. We made a strain with a $WTC_{846:Cdc20}$ allele and used it to bring about G2/M arrest followed by rapid and entirely synchronous release, with fast induction kinetics and low cell-to-cell variability in timing of the induction, thus demonstrating that WTC_{846} can be used for dynamic control of gene expression. We showed that WTC_{846} is a strong promoter by describing a predicted, but as of yet unobserved, overexpression phenotype for Ipl1. Importantly, we showed that different levels of WTC_{846} expression of Whi5 clamped cell size at different levels, decoupling that Whi5 size phenotype from its variable expression during the cell cycle. These results operationally define WTC_{846} controlled genes as a new type of conditional allele, useful in manipulating cells and in identifying gene dosage dependent phenotypes. We thus hope that WTC_{846} strains will find use in assessment of phenotypes now incompletely penetrant due to variable dosage of the causative gene product, in targeted cell biological experimentation, and in genomewide studies such as gene by gene and gene by chemical epistasis screens [65, 66].

We imagine that implementation of similar control logic in mammalian cells and organisms will enable types of experiments now impossible due to cell-to-cell variation and imprecise control. For example, in metazoans, phenotypic expression of many different alleles varies in penetrance, which is often affected by treatments that affect protein dosage[67]. Furthermore, the ability to have almost all the cells in a tissue or organ exhibiting a stable phenotype with low cell-to-cell variability might enable construction of more consistently behaved mammalian models of human disease. Finally, ample evidence supports the idea that a major component of key phenotypes such as tumor formation and aging is stochastic variation[68], including in gene expression. Although our understanding of different sources of variation in gene expression in single cells and organisms is expanding [69], the causes of such variation remain elusive. Use of WTC_{846} type, expression-clamped controllers to suppress variation in expression of key genes might help illuminate the importance of variable expression for these phenotypes.

Materials&Methods

Plasmids

Information on the contents of all plasmids, and main promoter and protein sequences used in this study can be found in Tables S2&S4. All plasmids with auxotrophic markers were constructed based on the pRG shuttle vector series [70] using either restriction enzyme cloning or isothermal assembly. Inserts were generated either by PCR or custom DNA synthesis (GeneArt, UK). All plasmids used to generate linear PCR products for tagging transformations were based on the pFA6 backbone [71]. All plasmids necessary to use WTC_{846} are available through Addgene along with a detailed protocol. The insert of all plasmids bearing TetR were cloned such that the insert promoter was closer to the T7 promoter and the terminator was near the T3 promoter of the backbone. In all plasmids bearing Citrine, the insert was flipped, such that the insert promoter was near the T3 promoter, and the terminator near the T7 promoter. This was done to avoid homologous recombination during subsequent integration of these plasmids into the same strain.

Strains

All strains used in this study can be found in Table S1. All strains used for fluorescent measurements and the $WTC_{846:Tpi1\text{-}K3}$ strain are based on a BY4743 derivative haploid background (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0). All other strains where $P_{7tet.1}$ replaced endogenous promoters were based on the haploid BY4741 background with the modifications whi5-mKok::His3 and myo1-mKate(3x)::Kan. The oligos used for promoter replacement can be found in Table S3. Correct replacement of the promoter of endogenous genes with $P_{7tet.1}$ was checked using colony PCR and subsequent sequencing (Microsynth, Switzerland). We used ORF specific oligos and the standard $P_{7tet.1}$ annealing oligo FRO4507 for these checks.

Chemicals & Media

All experiments were performed in YPD media, unless otherwise specified. YPD/YPE media were prepared with 1% yeast extract (Thermofisher, 212720), 2% bacto-peptone (Thermofisher, 211820) and 2% glucose (Sigma, G8270) / ethanol(Honeywell, 02860). All synthethic media contained 0.17% yeast nitrogen base (BD Difco, 233520) with ammonium (Sigma, 31119) as nitrogen source, complete complement of amino acids and adenine and uracil, except for SProline which contained 0.17% yeast nitrogen base without any nitrogen source, minimal complement of amino acids and 1mg/mL proline as the sole nitrogen source. The carbon source was 2% glucose for SD and SProline, 2% ethanol for SEthanol, 3% glycerol for SGlycerol (Applichem, A2957), 2% fructose for SFructose and 2% Raffinose for SRaffinose.

aTc was purchased from Cayman Chemicals (10009542) and prepared as a 4628.8ng/mL stock in ethanol for long term storage, and diluted in water for experiments as necessary. All oligos for cloning and for strain construction were synthesized by Thermofisher.

Plates for strain construction and for spotting assays were poured by adding agar (BD Sciences, 214040) to the media described above. To construct strains where $P_{7tet.1}$ replaces endogenous promoters, antibiotic markers were used. Cells were plated on YPD+antibiotic plates. Whenever the promoter of an essential gene was being replaced, transformations were plated on multiple plates with YPD+antibiotic and 10/50/100/500 ng/mL aTc.

Flow Cytometry

Cells were diluted from dense precultures and cultured to early exponential phase (5-10x10^6 cells/mL) in 96 deep-well plates at 30°C before induction with aTc if necessary. For aTc dose responses, samples were taken at times indicated. For experiments where no dose response was necessary, cells were measured at least 4 hours after dilution of precultures, but always before stationary phase. All samples were diluted in PBS and measured using a LSR-Fortessa LSRII equipped with a high-throughput sampler. PMT voltages for the forward and side scatter measurements were set up such that the height of the signal was not saturated. Citrine fluorescence was detected using a 488 nm excitation laser and a 530/30 nm emission filter. PMT voltage for this channel was set up such that the signal from TDH3 promoter (P_{TDH3}) expressed Citrine did not saturate the signal. Except for basal level measurements in Figure 3B, where PMT voltage for the

Citrine channel was increased to maximum. Side scatter was detected using the 488 nm excitation laser and 488/10 nm emission filter.

Spotting Assays

Cells were precultured in YPD media with 20ng/mL aTc (except for $WTC_{846:Ipl1\text{-}K2}$ which was precultured in 10ng/mL aTc) and the necessary antibiotic (Nourseothricin (Werner BioAgents, clonNAT) or Hygromycin (ThermoFisher,10687010) to stationary phase, and diluted into YPD+antibiotic without aTc at a concentration of $0.8x10^6$ cells/mL. 6 hours later, cells were spun down and resuspended in YPD. Cells were spotted on plates containing different media and aTc concentrations such that the most concentrated spot has $2.25x10^6$ cells, and each column is a 1:10 dilution. Pictures were taken after 24 hours for the YPD and SD plates, and 42 hours for Sproline, Sglycerol and YPE plates.

Western Blots

Cells were grown to stationary phase with the necessary aTc concentration. 5mL of cell culture was centrifuged and resuspended in 1mL 70% ethanol. Fixed cells were again centrifuged, and resuspended in 200μL Trupage LDS loading buffer (Merck, PCG3009) supplemented with 8M urea. Cells were broken using glass beads and a bead beater, and boiled at 95°C for 30 minutes. Proteins were separated using SDS-Page with Trupage precast 10% gels (Merck, PCG2009-10EA) and the associated commercial buffer, and transferred onto a nitrocellulose membrane (GE Healthcare Life Sciences, 10600008). We used mouse monoclonal primary antibodies for detecting TetR (Takara, Clone 9G9), and Citrine (Merck, G6539), both diluted 1:2000 in Odyssey Blocking buffer (PBS) (LI-COR Biosciences) + 0.2% Tween 20. The secondary antibody was IRDye 800CW Goat anti-Mouse IgG Secondary Antibody from Li-Cor (926-32210), diluted 1:5000 in the same manner. We used Chameleon Duo pre-stained Protein Ladder as our molecular weight marker (928-60000). We used the SNAP i.d.® 2.0 system and the associated protocol for antibody staining, and the Odyssey CLx (LI-COR) detector for imaging. Images were processed using the Fiji software[72].

Growth curves

Cells were precultured in YPD with aTc to stationary phase, then diluted into fresh media at a concentration of 50.000 cells per mL and induced with the necessary aTc concentrations, except for YPEthanol and SEthanol media where the concentration was 500.000 cells per mL. The Growth Profiler 960 (EnzyScreen) with 96 well plates and 250μm volume per well, or Biolector (m2p-labs) with 48 well plates and 1mL volume per plate was used to measure growth curves. Curves were aligned on the time axis to remove errors in dilution and variable lag phases.

Arrest & Release Assay and DNA Staining

Cells were precultured in YPD (pH 4) with 20ng/mL aTc to a concentration of $2x10^6$ cells/mL, then centrifuged and diluted 1:3 into YPD (pH 4) without aTc. To prevent the culture from becoming too dense, 25% of the cells were removed through filtration and the filtrated media was returned to the culture after 4 hours of arrest. After 8 hours of arrest, 600ng/mL aTc was added to the culture. Samples were taken every 90 minutes before, and every 5 minutes after the release, and fixed with 70% ethanol. The samples taken after the release were sonicated for 1 minute in a water bath before fixation.

All samples for DNA staining were digested with 5mg/mL proteinase K for 50 minutes at 50°C, followed by 2 hours of RNase A (Applichem, A2760,0500) treatment at 30°C. Samples were stained for DNA content using SYTOX Green (Thermofisher, S7020) diluted 1:5000 in PBS, and were sonicated in a water bath for 25 seconds before analysis using flow cytometry. Fluorescence was detected using a 488 nm excitation laser and a 525/15 nm emission filter. The PMT voltage was set up such that the sample with the highest expected ploidy did not saturate the signal.

Data analysis

All analysis was performed using R[73], and the packages Bioconductor[74], dplyr[75], drc[76], MASS[77], mixtools[78], and ggplot2[79]. All raw data is available upon request.

Flow cytometry data was not gated except when necessary to remove debris. Cell size was always calculated as the vector of the FSC-W and SSC-H signals. For aTc dose response experiments, median fluorescence of the entire population was used to fit a five-parameter dose response curve with the drm() command and the fplogistic formula $c + \frac{(d-c)}{1+exp(b(log(x+1))^{p1}+e(log(x+1))^{p2})}$ from the drc package. Parameters p_1 and p_2 were fixed individually for each curve, the rest of the parameters were estimated. All parameter values can be found in Table S5.

Robust linear fits for estimating cell-to-cell variability were performed using the rlm() command from the MASS package. A line was fit through the size vs fluorescence plot of each population, and the residual standard deviation given by the rlm() command was used as our measure of cell-to-cell variability. The RSD is a goodness of fit measure and increases as variability in fluorescence increases.

Where present, error bars for median fluorescence and the RSD were calculated using bootstrapping. The original set of data points were sampled with replacement 1000 times, and median fluorescence or RSD was calculated. 95% confidence intervals were calculated based on the 1000 repetitions and plotted as error bars.

For selecting the unbudded population using the SSC-W signal, we used the a normalmixEM() command from

the mixtools package to fit 3 gaussian distributions over the signal. The 99th percentile of the lower distribution was taken as the threshold that separates unbudded from budded cells.

Acknowledgements

We are grateful to Jörg Stelling, Gerry Smith, Steve Hahn, Hani Ahmad, William Peria, Steve Andrews, and Alejandro Colman-Lerner for useful discussions, and Kristina Elfström for early strain construction work. This work was supported by NIH NCI 5 R21 CA223901 (RB) and the NCCR Molecular Systems Engineering (AA, FR).

References

[1] R. H. Epstein, A. Bolle, C. M. Steinberg, E. Kellenberger, E. Boy de la Tour, R. Chevalley, R. S. Edgar, M. Susman, G. H. Denhardt, and A. Lielausis, "Physiological Studies of Conditional Lethal Mutants of Bacteriophage T4D," *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 28, no. 0, pp. 375–394, 1963.

[2] D. V. Goeddel, H. L. Heyneker, T. Hozumi, R. Arentzen, K. Itakura, D. G. Yansura, M. J. Ross, G. Miozzari, R. Crea, and P. H. Seeburg, "Direct expression in Escherichia coli of a DNA sequence coding for human growth hormone," *Nature*, vol. 281, no. 5732, pp. 544–548, 1979.

[3] M. A. Sochor, V. Vasireddy, T. G. Drivas, A. Wojno, T. Doung, I. Shpylchak, J. Bennicelli, D. Chung, J. Bennett, and M. Lewis, "An Autogenously Regulated Expression System for Gene Therapeutic Ocular Applications," *Scientific Reports*, vol. 5, 2015.

[4] K. Nishimura, T. Fukagawa, H. Takisawa, T. Kakimoto, and M. Kanemaki, "An auxin-based degron system for the rapid depletion of proteins in nonplant cells," *Nature Methods*, vol. 6, no. 12, pp. 917–922, 2009.

[5] A. H. Brand and N. Perrimon, "Targeted gene expression as a means of altering cell fates and generating dominant phenotypes," *Development*, vol. 118, no. 2, pp. 401–415, 1993.

[6] B. Sauer, "Functional expression of the cre-lox site-specific recombination system in the yeast Saccharomyces cerevisiae.," *Molecular and Cellular Biology*, vol. 7, no. 6, pp. 2087–2096, 1987.

[7] M. Gossen, S. Freundlieb, G. Bender, G. Müller, W. Hillen, and H. Bujard, "Transcriptional activation by tetracyclines in mammalian cells," *Science*, vol. 268, no. 5218, pp. 1766–1769, 1995.

[8] E. Garí, L. Piedrafita, M. Aldea, and E. Herrero, "A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in Saccharomyces cervisiae," *Yeast*, vol. 13, no. 9, pp. 837–848, 1997.

[9] R. Maurer, B. J. Meyer, and M. Ptashne, "Gene regulation at the right operator (OR) of bacteriophage λ. I. OR3 and autogenous negative control by repressor," *Journal of Molecular Biology*, vol. 139, no. 2, pp. 147–161, 1980.

[10] D. Maya, M. J. Quintero, M. De La Cruz Muñoz-Centeno, and S. Chávez, "Systems for applied gene control in Saccharomyces cerevisiae," 2008.

[11] D. S. Ottoz, F. Rudolf, and J. Stelling, "Inducible, tightly regulated and growth condition-independent transcription factor in Saccharomyces cerevisiae," *Nucleic Acids Research*, vol. 42, no. 17, 2014.

[12] R. S. McIsaac, B. L. Oakes, X. Wang, K. A. Dummit, D. Botstein, and M. B. Noyes, "Synthetic gene expression perturbation systems with rapid, tunable, single-gene specificity in yeast," *Nucleic Acids Research*, vol. 41, no. 4, 2013.

[13] R. Brent and M. Ptashne, "A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor," *Cell*, 1985.

[14] S. Hahn and E. T. Young, "Transcriptional regulation in saccharomyces cerevisiae: Transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators," *Genetics*, vol. 189, no. 3, pp. 705–736, 2011.

[15] G. Bellí, E. Garí, L. Piedrafita, M. Aldea, and E. Herrero, "An activator/repressor dual system allows tight tetracycline-regulated gene expression in budding yeast," *Nucleic Acids Research*, vol. 26, no. 4, pp. 942–947, 1998.

[16] R. S. McIsaac, B. L. Oakes, X. Wang, K. A. Dummit, D. Botstein, and M. B. Noyes, "Synthetic gene expression perturbation systems with rapid, tunable, single-gene specificity in yeast," *Nucleic Acids Research*, vol. 41, no. 4, 2013.

[17] G. Gill and M. Ptashne, "Negative effect of the transcriptional activator GAL4," *Nature*, vol. 334, no. 6184, pp. 721–724, 1988.

[18] M. Meurer, V. Chevyreva, B. Cerulus, and M. Knop, "The regulatable MAL32 promoter in Saccharomyces cerevisiae: characteristics and tools to facilitate its use," *Yeast*, vol. 34, no. 1, pp. 39–49, 2017.

[19] G. L. Elison, R. Song, and M. Acar, "A Precise Genome Editing Method Reveals Insights into the Activity of Eukaryotic Promoters," *Cell Reports*, vol. 18, no. 1, pp. 275–286, 2017.

[20] X. Wang, L. Bai, G. O. Bryant, and M. Ptashne, "Nucleosomes and the accessibility problem," 2011.

[21] K. Gaston and P. S. Jayaraman, "Transcriptional repression in eukaryotes: Repressors and repression mechanisms," 2003.

[22] J. B. Jaynes and P. H. O'Farrell, "Activation and repression of transcription by homoeodomain-containing proteins that bind a common site," *Nature*, vol. 336, no. 6201, pp. 744–749, 1988.

[23] Y. Ohkuma, M. Horikoshi, R. G. Roeder, and C. Desplan, "Binding site-dependent direct activation and repression of in vitro transcription by Drosophila homeodomain proteins," *Cell*, vol. 61, no. 3, pp. 475–484, 1990.

[24] Y. Ohkuma, M. Horikoshi, R. G. Roeder, and C. Desplan, "Engrailed, a homeodomain protein, can repress in vitro transcription by competition with the TATA box-binding protein transcription factor IID," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 87, no. 6, pp. 2289–2293, 1990.

[25] R. Brent and M. Ptashne, "A bacterial repressor protein or a yeast transcriptional terminator can block upstream activation of a yeast gene," *Nature*, vol. 312, no. 5995, pp. 612–615, 1984.

[26] M. C. Hu and N. Davidson, "The inducible Iac operator-repressor system is functional in mammalian cells," *Cell*, vol. 48, no. 4, pp. 555–566, 1987.

[27] M. Brown, J. Figge, U. Hansen, C. Wright, K. T. Jeang, G. Khoury, D. M. Livingston, and T. M. Roberts, "Lac repressor can regulate expression from a hybrid SV40 early promoter containing a lac operator in animal cells," *Cell*, vol. 49, no. 5, pp. 603–612, 1987.

[28] K. F. Murphy, G. Balázsi, and J. J. Collins, "Combinatorial promoter design for engineering noisy gene expression," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 31, pp. 12726–12731, 2007.

[29] H. Wedler and R. Wambutt, "A temperature sensitive lambda cl repressor functions on a modified operator in yeast cells by masking the TATA element," *Molecular and General Genetics*, vol. 248, no. 4, pp. 499–505, 1995.

[30] B. Ho, A. Baryshnikova, and G. W. Brown, "Unification of Protein Abundance Datasets Yields a Quantitative Saccharomyces cerevisiae Proteome," *Cell Systems*, vol. 6, no. 2, pp. 192–205.e3, 2018.

[31] S. Yagi, K. Yagi, J. Fukuoka, and M. Suzuki, "The UAS of the Yeast GAPDH Promoter Consists of Multiple General Functional Elements Including RAP1 and GRF2 Binding Sites," *Journal of Veterinary Medical Science*, vol. 56, no. 2, pp. 235–244, 1994.

[32] B. P. H. Metzger, D. C. Yuan, J. D. Gruber, F. Duveau, and P. J. Wittkopp, "Selection on noise constrains variation in a eukaryotic promoter," *Nature*, vol. 521, no. 7552, pp. 344–347, 2015.

[33] H. V. Baker, M. A. Huie, E. W. Scott, C. M. Drazinic, M. C. Lopez, I. A. N. K. Hornstra, T. P. Yang, and H. V. Bakerl, "Characterization of the DNA-Binding activity of GCR1 : in vivo evidence for two GCR1-binding sites in the upstream activating sequence of TPI of Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 12, no. 6, pp. 2690–2700, 1992.

[34] I. Mogno, F. Vallania, R. D. Mitra, and B. A. Cohen, "TATA is a modular component of synthetic promoters," *Genome Research*, vol. 20, no. 10, pp. 1391–1397, 2010.

[35] D. Nevozhay, R. M. Adams, K. F. Murphy, K. Josić, and G. Balázsi, "Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 13, pp. 5123–5128, 2009.

[36] A. Becskel and L. Serrano, "Engineering stability in gene networks by autoregulation," *Nature*, vol. 405, no. 6786, pp. 590–593, 2000.

[37] D. H. Williamson and A. W. Scopes, "The distribution of nucleic acids and protein between different sized yeast cells," *Experimental Cell Research*, vol. 24, no. 1, pp. 151–153, 1961.

[38] K. H. Wong and K. Struhl, "The Cyc8-Tup1 complex inhibits transcription primarily by masking the activation domain of the recruiting protein," *Genes and Development*, vol. 25, no. 23, pp. 2525–2539, 2011.

[39] Z. Zhang and J. C. Reese, "Redundant mechanisms are used by Ssn6-Tup1 in repressing chromosomal gene transcription in Saccharomyces cerevisiae," *Journal of Biological Chemistry*, vol. 279, no. 38, pp. 39240–39250, 2004.

[40] T. A. Mennella, L. G. Klinkenberg, and R. S. Zitomer, "Recruitment of Tup1-Ssn6 by Yeast Hypoxic Genes and Chromatin-Independent Exclusion of TATA Binding Protein," *Eukaryotic Cell*, vol. 2, no. 6, pp. 1288–1303, 2003.

[41] K. Chen, M. A. Wilson, C. Hirsch, A. Watson, S. Liang, Y. Lu, W. Li, and S. Y. Dent, "Stabilization of the promoter nucleosomes in nucleosome-free regions by the yeast Cyc8-Tup1 corepressor," *Genome Research*, vol. 23, no. 2, pp. 312–322, 2013.

[42] D. Tzamarias and K. Struhl, "Functional dissection of the yeast Cyc8-Tupl transcriptional co-repressor complex," *Nature*, vol. 369, no. 6483, pp. 758–761, 1994.

[43] M. Kozak, "Regulation of translation via mRNA structure in prokaryotes and eukaryotes," 2005.

[44] K. Bartlett and K. Kim, "Insight into Tor2, a budding yeast microdomain protein," 2014.

[45] A. Ambesi, M. Miranda, V. V. Petrov, and C. W. Slayman, "Biogenesis and function of the yeast plasma-membrane H+-ATPase," *Journal of Experimental Biology*, vol. 203, no. 1, pp. 155–160, 2000.

[46] D. G. Fraenkel, "The top genes: On the distance from transcript to function in yeast glycolysis," 2003.

[47] M. Muñoz-Barrera and F. Monje-Casas, "Increased Aurora B activity causes continuous disruption of kinetochore-microtubule attachments and spindle instability," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 38, pp. E3996–E4005, 2014.

[48] A. González-Loyola, G. Fernández-Miranda, M. Trakala, D. Partida, K. Samejima, H. Ogawa, M. Cañamero, A. de Martino, Á. Martínez-Ramírez, G. de Cárcer, I. Pérez de Castro, W. C. Earnshaw, and M. Malumbres, "Aurora B Overexpression Causes Aneuploidy and p21 Cip1 Repression during Tumor Development," *Molecular and Cellular Biology*, vol. 35, no. 20, pp. 3566–3578, 2015.

[49] K. M. Schmoller, J. J. Turner, M. Kõivomägi, and J. M. Skotheim, "Dilution of the cell cycle inhibitor Whi5 controls budding-yeast cell size," *Nature*, vol. 526, no. 7572, pp. 268–272, 2015.

[50] M. Angeles Juanes, "Methods of synchronization of yeast cells for the analysis of cell cycle progression," in *Methods in Molecular Biology*, vol. 1505, pp. 19–34, 2017.

[51] J. A. Pesin and T. L. Orr-Weaver, "Regulation of APC/C Activators in Mitosis and Meiosis," *Annual Review of Cell and Developmental Biology*, vol. 24, no. 1, pp. 475–499, 2008.

[52] M. P. Cosma, T. Tanaka, and K. Nasmyth, "Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter," *Cell*, vol. 97, no. 3, pp. 299–311, 1999.

[53] K. F. Murphy, G. Balázsi, and J. J. Collins, "Combinatorial promoter design for engineering noisy gene expression," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 31, pp. 12726–12731, 2007.

[54] E. Nogales, R. K. Louder, and Y. He, "Structural Insights into the Eukaryotic Transcription Initiation Machinery," *Annual Review of Biophysics*, 2017.

[55] G. R. Smith and B. Magasanik, "Nature and self-regulated synthesis of the repressor of the hut operons in Salmonella typhimurium.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 68, no. 7, pp. 1493–1497, 1971.

[56] R. Brent and M. Ptashne, "The lexA gene product represses its own promoter," *Proceedings of the National Academy of Sciences of the United States of America*, 1980.

[57] K. P. Bertrand, K. Postle, L. V. Wray, and W. S. Reznikoff, "Overlapping divergent promoters control expression of Tn10 tetracycline resistance," *Gene*, vol. 23, no. 2, pp. 149–156, 1983.

[58] S. S. Andrews, W. J. Peria, R. C. Yu, A. Colman-Lerner, and R. Brent, "Push-Pull and Feedback Mechanisms Can Align Signaling System Outputs with Inputs," *Cell Systems*, vol. 3, no. 5, pp. 444–455.e2, 2016.

[59] R. Dhar, A. M. Missarova, B. Lehner, and L. B. Carey, "Single cell functional genomics reveals the importance of mitochondria in cell-to-cell phenotypic variation," *eLife*, vol. 8, 2019.

[60] A. Bar-Even, J. Paulsson, N. Maheshri, M. Carmi, E. O'Shea, Y. Pilpel, and N. Barkai, "Noise in protein expression scales with natural protein abundance," *Nature Genetics*, vol. 38, no. 6, pp. 636–643, 2006.

[61] L. Keren, D. Van Dijk, S. Weingarten-Gabbay, D. Davidi, G. Jona, A. Weinberger, R. Milo, and E. Segal, "Noise in gene expression is coupled to growth rate," *Genome Research*, vol. 25, no. 12, pp. 1893–1902, 2015.

[62] C. Briat, A. Gupta, and M. Khammash, "Antithetic Integral Feedback Ensures Robust Perfect Adaptation in Noisy Bimolecular Networks," *Cell Systems*, vol. 2, no. 1, pp. 15–26, 2016.

[63] U. S. Varanasi, M. Klis, P. B. Mikesell, and R. J. Trumbly, "The Cyc8 (Ssn6)-Tup1 corepressor complex is composed of one Cyc8 and four Tup1 subunits.," *Molecular and Cellular Biology*, vol. 16, no. 12, pp. 6707–6714, 1996.

[64] A. B. Fleming, S. Beggs, M. Church, Y. Tsukihashi, and S. Pennings, "The yeast Cyc8-Tup1 complex cooperates with Hda1p and Rpd3p histone deacetylases to robustly repress transcription of the subtelomeric FLO1 gene," *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, vol. 1839, no. 11, pp. 1242–1255, 2014.

[65] M. Costanzo, B. VanderSluis, E. N. Koch, A. Baryshnikova, C. Pons, G. Tan, W. Wang, M. Usaj, J. Hanchard, S. D. Lee, V. Pelechano, E. B. Styles, M. Billmann, J. Van Leeuwen, N. Van Dyk, Z. Y. Lin, E. Kuzmin, J. Nelson, J. S. Piotrowski, T. Srikumar, S. Bahr, Y. Chen, R. Deshpande, C. F. Kurat, S. C. Li, Z. Li, M. M. Usaj, H. Okada, N. Pascoe, B. J. S. Luis, S. Sharifpoor, E. Shuteriqi, S. W. Simpkins, J. Snider, H. G. Suresh, Y. Tan, H. Zhu, N. Malod-Dognin, V. Janjic, N. Przulj, O. G. Troyanskaya, I. Stagljar, T. Xia, Y. Ohya, A. C. Gingras, B. Raught, M. Boutros, L. M. Steinmetz, C. L. Moore, A. P. Rosebrock, A. A. Caudy, C. L. Myers, B. Andrews, and C. Boone, "A global genetic interaction network maps a wiring diagram of cellular function," *Science*, vol. 353, no. 6306, 2016.

[66] X. Li, J. Lalic, P. Baeza-Centurion, R. Dhar, and B. Lehner, "Changes in gene expression predictably shift and switch genetic interactions," *Nature Communications*, vol. 10, no. 1, 2019.

[67] M. O. Casanueva, A. Burga, and B. Lehner, "Fitness trade-offs and environmentally induced mutation buffering in isogenic C. elegans," *Science*, vol. 335, no. 6064, pp. 82–85, 2012.

[68] L. A. Herndon, P. J. Schmeissner, J. M. Dudaronek, P. A. Brown, K. M. Listner, Y. Sakano, M. C. Paupard, D. H. Hall, and M. Driscoll, "Stochastic and genetic factors influence tissue-specific decline in ageing C. Elegans," *Nature*, vol. 419, no. 6909, pp. 808–814, 2002.

[69] A. Colman-Lerner, A. Gordon, E. Serra, T. Chin, O. Resnekov, D. Endy, C. G. Pesce, and R. Brent, "Regulated cell-to-cell variation in a cell-fate decision system," *Nature*, vol. 437, no. 7059, pp. 699–706, 2005.

[70] R. Gnügge and F. Rudolf, "Saccharomyces cerevisiae Shuttle vectors," 2017.

[71] C. Janke, M. M. Magiera, N. Rathfelder, C. Taxis, S. Reber, H. Maekawa, A. Moreno-Borchart, G. Doenges, E. Schwob, E. Schiebel, and M. Knop, "A versatile toolbox for PCR-based tagging of yeast genes: New fluorescent proteins, more markers and promoter substitution cassettes," *Yeast*, vol. 21, no. 11, pp. 947–962, 2004.

[72] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, "Fiji: An open-source platform for biological-image analysis," 2012.

[73] R Core Development Team, "A language and environment for statistical computing.," *Vienna, Austria*, vol. 1, 2013.

[74] B. Ellis, P. Haaland, F. Hahne, N. L. Meur, and N. Gopalakrishnan, "flowCore: flowCore: Basic structures for flow cytometry data," *http://bioconductor.org/*, pp. 1–33, 2009.

[75] H. Wickham, F. Romain, L. Henry, K. Müller, and RStudio, "dplyr: A Grammar of Data Manipulation," *R package version 0.8.0.1*, p. 75, 2018.

[76] C. Ritz, F. Baty, J. C. Streibig, and D. Gerhard, "Dose-response analysis using R," *PLoS ONE*, vol. 10, no. 12, 2015.

[77] K. Kafadar, J. R. Koehler, W. N. Venables, and B. D. Ripley, *Modern Applied Statistics with S-Plus*, vol. 53. 1999.

[78] T. Benaglia, D. Chauveau, D. R. Hunter, and D. S. Young, "Mixtools: An R package for analyzing finite mixture models," *Journal of Statistical Software*, vol. 32, no. 6, pp. 1–29, 2009.

[79] C. Ginestet, "ggplot2: Elegant Graphics for Data Analysis," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 174, no. 1, pp. 245–246, 2011.