The Transcription Factor Titration Effect Dictates Level of Gene Expression

Robert C. Brewster,[1,7] Franz M. Weinert,[1,7] Hernan G. Garcia,[2] Dan Song,[3,4] Mattias Rydenfelt,[5] and Rob Phillips[1,6,*]
[1]Department of Applied Physics, California Institute of Technology, Pasadena, CA 91125, USA
[2]Department of Physics, Princeton University, NJ 08540, USA
[3]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA
[4]Harvard Biophysics Program, Harvard Medical School, Boston, MA 02115, USA
[5]Department of Physics, California Institute of Technology, Pasadena, CA 91125, USA
[6]Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA
[7]Co-first author
*Correspondence: phillips@pboc.caltech.edu
http://dx.doi.org/10.1016/j.cell.2014.02.022

SUMMARY

Models of transcription are often built around a picture of RNA polymerase and transcription factors (TFs) acting on a single copy of a promoter. However, most TFs are shared between multiple genes with varying binding affinities. Beyond that, genes often exist at high copy number—in multiple identical copies on the chromosome or on plasmids or viral vectors with copy numbers in the hundreds. Using a thermodynamic model, we characterize the interplay between TF copy number and the demand for that TF. We demonstrate the parameter-free predictive power of this model as a function of the copy number of the TF and the number and affinities of the available specific binding sites; such predictive control is important for the understanding of transcription and the desire to quantitatively design the output of genetic circuits. Finally, we use these experiments to dynamically measure plasmid copy number through the cell cycle.

INTRODUCTION

Regulatory biology remains one of the most fertile areas for the quantitative dissection of biological systems, with two broad classes of examples coming from the study of cell signaling and gene regulation (Lim, 2002; Ptashne and Gann, 2002; Bhattacharyya et al., 2006; Kentner and Sourjik, 2010; Garcia et al., 2010). With increasing regularity, these systems are examined in tandem using both theoretical models with precise "governing equations" and precision measurements whose ambition is to explicitly test the validity of these models. The study of gene expression in bacteria has enjoyed a close interplay between the so-called thermodynamic models, which predict the mean level of expression as a function of architectural parameters

characterizing the regulatory motif of interest, and quantitative measurements, which can now even be performed at the single-cell level (Buchler et al., 2003; Vilar and Leibler, 2003; Dekel and Alon, 2005; Ozbudak et al., 2004; Kuhlman et al., 2007; Kinney et al., 2010; Daber et al., 2011; Garcia and Phillips, 2011).

Typically, such models rely on the assumption that the number of TFs is in excess with respect to the number of its binding sites in the cell. There are many situations in which this assumption might break down, such as those involving highly replicated viral DNA (Luria and Dulbecco, 1949), genes expressed on plasmids (Guido et al., 2006), genes existing in multiple identical copies on the chromosome (Bremer and Dennis, 1996; Wang et al., 1999; Navarro-Quezada and Schoen, 2002; Aitman et al., 2006; Hanada et al., 2011) or even just genes controlled by "overworked" TFs with many available target genes (Busby and Ebright, 1999). Additionally, this interplay between the number of TFs and the number of its binding sites provides yet another tuning parameter with which to test and refine theoretical models of transcriptional regulation as well as precisely control the output of synthetic genetic circuits (Endy, 2005; Voigt, 2006; Mukherji and van Oudenaarden, 2009; Elowitz and Lim, 2010). In fact, it is common to explore regulatory architecture in the context of multicopy plasmids (Guido et al., 2006; Cox et al., 2007; Kaplan et al., 2008; Kinney et al., 2010). As a result, precise knowledge of the role of plasmid copy number on the output levels of gene expression is required. This interdependence of a given gene's input-output relation with the external environment in which it exists has been termed "retroactivity" (Del Vecchio et al., 2008; Kim and Sauro, 2011) and is treated in analogy to impedance in electrical circuits. Some studies have explored this interplay typically in the context of activation, with binding competition stemming from molecular depletants (Ricci et al., 2011) or binding arrays (Lee and Maheshri, 2012).

Here, we dissect the interplay between TF copy number and the number of its target binding sites using the simple repression regulatory architecture. Simple repression is a ubiquitous motif in *E. coli* (Madan Babu and Teichmann, 2003; Gama-Castro et al., 2011), which consists of a promoter with a single proximal
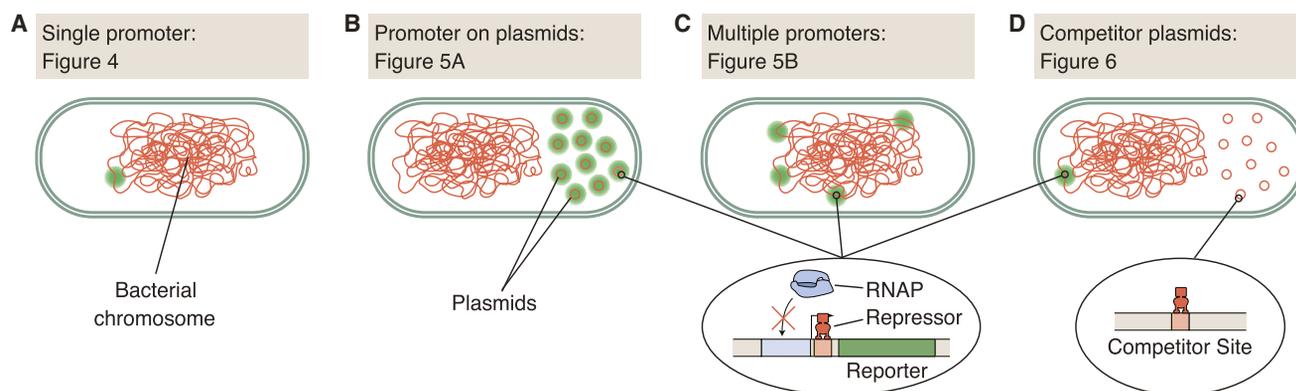
Figure 1. Examples Examined in This Study of Transcriptional Regulation with Competition for the TF
(A) Single chromosomal copy of the gene of interest.
(B and C) Competition from multiple identical genes in the simple repression regulatory architecture when the promoters are (B) placed on a high copy number plasmid or (C) integrated in multiple chromosomal locations.
(D) The chromosomal reporter construct competes with competitor plasmids that have binding sites for the repressor but do not code for the reporter gene. In this particular case, the competitor binding sites can have a different affinity than the regulated gene.

repressor binding site such that when a repressor is bound, no transcription ensues (Schlax et al., 1995; Rojo, 2001; Sanchez et al., 2011). In particular, we focus on simple repression by Lac repressor (LacI), which has been extensively studied in the context of theoretical models of transcriptional regulation (Oehler et al., 1994; Vilar and Leibler, 2003; Ozbudak et al., 2004; Bintu et al., 2005a; Kuhlman et al., 2007; Daber et al., 2011; Garcia and Phillips, 2011). Using video fluorescence microscopy, we simultaneously measure both the absolute number of repressors and the rate of expression of a reporter fluorescent protein in single cells as they progress through the cell cycle. This method is used to examine several cases of simple repression in which the TF-binding sites are placed in multiple locations, shown schematically in Figure 1. In particular, these include transcription from a plasmid at several distinct copy numbers (Figure 1B), transcription from multiple identical copies integrated in the chromosome (Figure 1C), and transcription from a single chromosomal copy that competes for the repressor with plasmids also containing a specific binding target (Figure 1D).

One major outcome of this study is that, when a TF is shared among many binding sites, either due to multiple identical copies of a gene regulated by that TF or due to unrelated genes that also independently bind the TF, the correlation in occupancy between the binding sites will lead to a complex dosage response to that TF (Lee and Maheshri, 2012; Rydenfelt et al., 2014). At low copy numbers (relative to the number of binding sites), this essentially buffers the transcriptional level to the presence of the TF, and at high copy numbers, the response of the fold-change is similar to that seen for a single isolated copy of the gene with no binding competition. The sharpness of the transition between these regimes is predicted to depend explicitly on the relative strength of the specific binding site on the gene of interest compared to the specific sites with which it competes. However, we find that the width of the plasmid distribution inside of the population of cells can also play a role in flattening the transition, and the distribution must be taken into account to accurately

predict gene expression when the plasmid distribution itself becomes wider than the transition region in the fold-change curve, which tends to occur for stronger binding operators.

Building on the success of the predictive model, we then exploit it as a tool for measuring plasmid copy number throughout the course of the cell cycle. The average number of plasmids per cell increases as the cell cycle progresses, with a time-averaged mean value that is consistent with our independent bulk qPCR measurements of the mean copy number.

RESULTS

Thermodynamic Model
Our results are based upon time-lapse fluorescence microscopy (Figure 2) in which we measure the level of gene expression by looking at the rate of production of a fluorescent reporter (i.e., dP/dt, where P is the fluorescent protein number per cell). Specifically, we measure the fold-change given by

$$\text{fold-change} = \frac{\frac{dP}{dt}(R \neq 0)}{\frac{dP}{dt}(R = 0)}, \tag{1}$$

which compares the rate of production in the presence of repressors R to the rate of production in their absence. This should be contrasted with bulk measurements, which typically measure the steady-state level of the gene product in cell populations. However, we can demonstrate the relationship between the fold-change data from steady-state measurements, in which expression is quantified as levels of fluorescence reporter, P, and that obtained using video microscopy by observing the rate of production of a fluorescent reporter, dP/dt. In the limit that degradation of the measured product is slow, the equivalence of these methods can be derived (see Extended Experimental Procedures section "Equivalence of fold-change in steady-state measurements and video microscopy" available online),
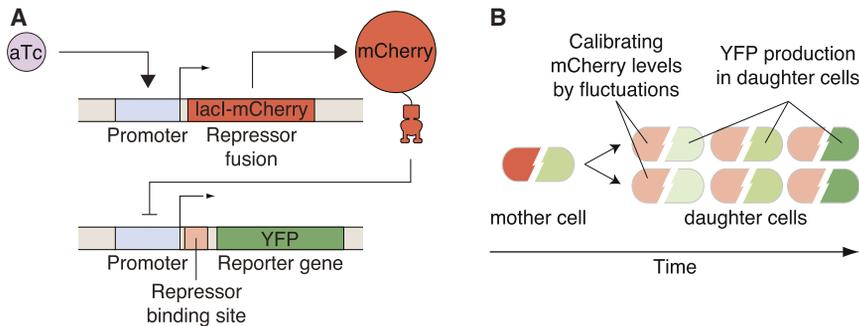
Figure 2. Experimental Methods for the Single-Cell Dissection of Regulatory Architectures

(A) Genetic circuit employed in this work. The expression of the LacI-mCherry fusion is induced by the small molecule aTc. The repressor acts on a promoter expressing a YFP reporter gene.

(B) Individual cells are observed through a division event. The fluctuations in the partitioning of the LacI-mCherry between the daughters are used to calibrate the signal such that the mCherry fluorescence measurement in each cell can be expressed as an absolute number of repressor molecules. In addition, the rate of YFP production is measured over the cell cycle.

$$\text{fold-change} = \overbrace{\frac{\frac{dP}{dt}(R \neq 0)}{\frac{dP}{dt}(R = 0)}}^{\text{experiments}} = \underbrace{\frac{P(R \neq 0)}{P(R = 0)}}_{\text{steady-state microscopy}} = \overbrace{\frac{p_{\text{bound}}(R \neq 0)}{p_{\text{bound}}(R = 0)}}^{\text{theory}},$$

$$\underbrace{\phantom{\frac{\frac{dP}{dt}(R \neq 0)}{\frac{dP}{dt}(R = 0)}}}_{\text{video microscopy}}$$

(2)

suggesting that a direct comparison between the bulk measurements and those presented here is admitted as is the comparison to thermodynamic models.

The basic idea of the thermodynamic model of transcriptional regulation is to enumerate the possible configurations of the molecular players among the available specific and nonspecific binding sites and calculate the probability of finding RNA polymerase bound at the promoter of interest. These models predict the fold-change in gene expression defined as the ratio of the level of gene expression in the presence of TF to the level of expression in its absence. In particular, the fold-change for simple repression in the case where the gene and corresponding TF specifically bind only at the reporter gene (Figure 1A) is (Bintu et al., 2005b)

$$\text{fold-change} = \frac{1}{1 + \frac{R}{N_{\text{NS}}} e^{-\Delta\epsilon/k_B T}},$$

(3)

where R is the number of repressors present in the cell, N_{NS} is the size of the nonspecific binding reservoir (which we take here to be the whole *E. coli* chromosome such that $N_{\text{NS}} = 5 \times 10^6$), and $\Delta\epsilon$ is the binding energy of repressor to its operator. In Rydenfelt et al. (2014), this model has been extended to the case of simple repression from multiple identical copies of the gene, schematically shown in Figures 1B and 1C. In this case, the fold-change is predicted to have the form,

$$\text{fold-change} = \frac{\sum_{m=0}^{\min(N,R)} \frac{R!}{(N_{\text{NS}})^m (R-m)!} \binom{N}{m} e^{-\beta m \Delta\epsilon} (N - m)}{N \sum_{m=0}^{\min(N,R)} \frac{R!}{(N_{\text{NS}})^m (R-m)!} \binom{N}{m} e^{-\beta m \Delta\epsilon}},$$

(4)

where the only new parameter is N, the copy number of the gene. Finally, the model predicts the regulatory outcome of a single

gene copy regulated by simple repression with a binding affinity $\Delta\epsilon$ in the presence of competing binding sites with a distinct affinity $\Delta\epsilon_c$ (Figure 1D). In this more complex case, the fold-change is given by

$$\text{fold-change} = \frac{Z_u}{Z_b + Z_u},$$

(5)

where Z_b and Z_u are the partition functions for the case where the repressor is bound or unbound to the chromosomal promoter, given by,

$$Z_u = \sum_{k=0}^{\min(N_c, R)} \frac{R!}{N_{\text{NS}}^k (R - k)!} \binom{N_c}{k} e^{-\beta \Delta\epsilon_c},$$

(6)

$$Z_b = \sum_{k=0}^{\min(N_c, R-1)} \frac{R!}{N_{\text{NS}}^k (R - k - 1)!} \binom{N_c}{k} e^{-\beta(k\Delta\epsilon_c + \Delta\epsilon)},$$

(7)

where N_c is the copy number of the plasmid containing the competing binding site and no reporter gene. The extension of this model to N copies of the gene with N_c competitors is detailed in the Extended Experimental Procedures section "Accounting for chromosome replication in competitor theory."

One feature of the theoretical predictions in Equations 4 and 5 is that, in the limit that $R \gg N$ (Equation 4) or $R \gg N_c$ (Equation 5), these expressions immediately simplify to Equation 3 (see Extended Experimental Procedures section "Thermodynamic model in the limit $R \gg N$" for details), meaning that the multiple promoters are independent in this limit. Between all of these situations, there are relatively few parameters: the number of TFs (R), the size of the nonspecific reservoir (N_{NS}), the strength of binding sites ($\Delta\epsilon, \Delta\epsilon_c$), and the copy number of the gene (N) or of the competing binding site plasmid (N_c). Interestingly, many of the same parameters arise within each of the different scenarios that we are considering, and a critical test of the theoretical understanding is the self-consistency of those results. Once these quantities are determined, the theory generates falsifiable predictions without any free parameters for all remaining experiments. In the following paragraphs, we discuss how these parameters were determined from independent measurements, with the ultimate objective of performing a stringent test of the thermodynamic models, in general, and of the impact of gene copy number on regulation, in particular.
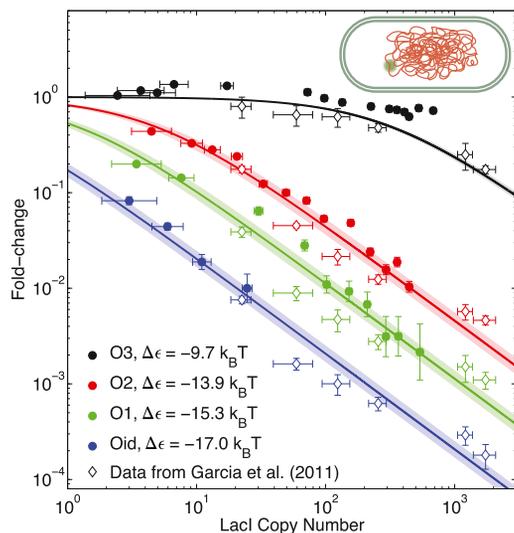
Figure 3. Simple Repression of a Single Chromosomal Construct
Fold-change of simple repression construct located on the chromosome as a function of Lac repressor copy number. The solid lines correspond to Equation 3, with values for $\Delta\epsilon$ from steady-state measurements of expression. The data from steady-state measurements (Garcia and Phillips, 2011) are shown as open symbols. The data from our experiments (filled symbols) are both consistent with the model with no free parameters (curves) and with expression data obtained from the same construct in steady-state measurements. The shaded regions on the curves represent the uncertainty from the errors in the measurement of the binding energies. For the solid points, error bars in fold-change measurements are SEM, and error bars in LacI copy number are the quadrature summed errors from the calibration factor and the inherent resolution limit of LacI detection.

Fluorescent Measurements of Gene Expression and Absolute TF Copy Number

We consider a number of distinct regulatory landscapes (Figure 1), all of which involve a rich interplay between the gene copy number and the copy number of the transcription factor controlling that gene. To test the expressions for fold-change given in Equations 3–5, we need to simultaneously measure both the rate of gene expression and the absolute number of TFs. To that end, as shown in Figure 2B (and in greater detail in Figure S1A), our cells harbor two important fluorescent proteins: one to mark the TF and one to mark the gene product.

We use the partitioning statistics of the repressor TF, a LacI-mCherry fusion, during cell division to determine the absolute TF copy number from the arbitrary mCherry fluorescence intensity in a given cell (Rosenfeld et al., 2005, 2006; Teng et al., 2010). We find that, at maximum induction, no secondary effects to the physiology (as measured by global transcription rate; Figure S2) are observed with 1,000 repressors per cell (Figure S3C). We also find that our lower resolution limit is 3−5 repressors per cell (Figure S3E). See Extended Experimental Procedures section "Calibrating LacI-mCherry intensity to absolute copy number" and Figure S3 for details on this method. Simultaneously, we determine the level of gene expression by measuring the rate of YFP production.

Gene Copy Number Measured by qPCR

We determine gene copy number using qPCR to measure the average number of plasmids in a cell. In this study, we use plasmids based off the ColE1 Δrom origin of replication from Lutz and Bujard (Lutz and Bujard, 1997; Lee et al., 2006a, 2006b). We also have made a version in which the Rom protein, responsible for regulating the plasmid copy number, is inserted back into the ColE1 Δrom origin to arrive at an origin functionally similar to the wild-type ColE1 origin (Twigg and Sherratt, 1980; Cesareni et al., 1982; Stueber and Bujard, 1982; Lutz and Bujard, 1997). Though previous measurements locate the copy number of ColE1 Δrom plasmid in the range of 50–70 (Lutz and Bujard, 1997), the addition of the Rom protein should result in a reduced average plasmid copy number (Twigg and Sherratt, 1980). We find that the ColE1 plasmid has an average copy number of 52 ± 5, whereas the ColE1 Δrom plasmid has a copy number of 64 ± 11 (error bars are SD from triplicates). These values for the copy number show up as either N or N_c in the predictions generated by Equations 4 and 5, respectively. One obvious naive aspect to this approach is that the plasmid copy number is treated as a single static value. In any population of cells, the copy number is subject to cell-to-cell variability, and thus the copy number is more accurately represented as a distribution rather than a single value (Wong Ng et al., 2010). Additionally, plasmid copy numbers are bound to increase as the cell progresses through its cycle under steady-state conditions (Paulsson and Ehrenberg, 2001). We will examine the consequences of these simplifications in a later section.

Determining Sequence-Dependent TF-Binding Energies

Finally, the affinities $\Delta\epsilon$ and $\Delta\epsilon_c$ of Lac repressor to its specific binding sites (Oid, O1, O2, and O3 from strongest to weakest) have been previously determined using bulk measurements (Oehler et al., 1994; Vilar and Leibler, 2003; Garcia and Phillips, 2011). Thus, we know all parameters in Equations 3, 4, and 5 in order to predict the fold-change in gene expression for every one of the regulatory cases considered in this paper (Figure 1). Effectively, this means that we can predict the fold-change as a function of the number of repressors without any free parameters at all.

Simple Thermodynamic Model Predicts Expression Level of Single Integrated Gene Copy

Our approach has several facets that require deeper examination. One possible confounding factor in comparing to other measurements on the same architecture is that the fusion of LacI to a fluorescent protein might affect its function as a TF, thus changing its binding properties with DNA. A second point is that it is not immediately clear that a comparison of expression rate from cells grown under a microscope on a flat surface is comparable to steady-state measurements grown in bulk media (Oehler et al., 1994; Kuhlman et al., 2007; Garcia and Phillips, 2011).

To assess these issues, we compare our video microscopy method against the outcome of previous bulk steady-state results performed using wild-type Lac repressor (Oehler et al., 1994; Garcia and Phillips, 2011; Brewster et al., 2012). In Figure 3, we show the result of measuring fold-change in expression of a
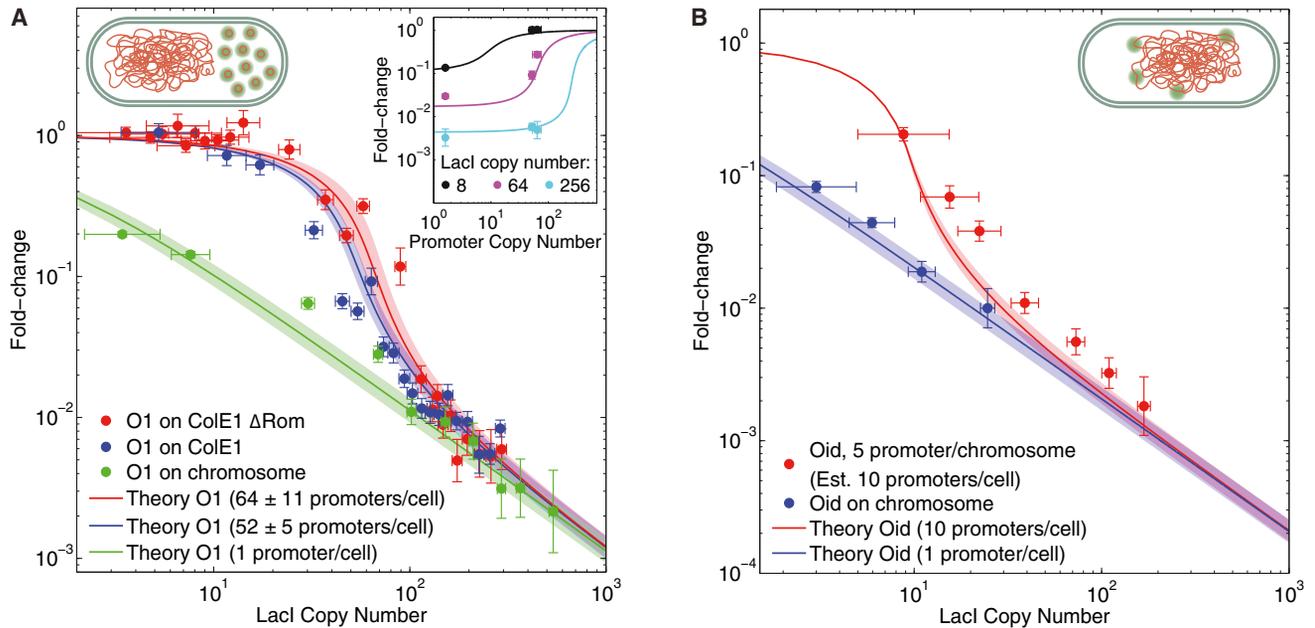
Figure 4. Fold-Change of Multiple Identical Gene Copies

(A) Fold-change as a function of Lac repressor copy number for two distinct plasmids with the O1 simple repression motif on a high copy number (ColE1) plasmid with (blue) and without (red) the Rom protein. Measurements are performed at the middle of the cell cycle. The blue and red solid lines are the theory from Equation 4 using the average copy number measured by qPCR and known binding energies from earlier steady-state measurements, as in Figure 3. The shaded regions represent the combined uncertainty in the copy number measurement and the binding energy measurement. For reference, the green symbols and line are the data and theory prediction from Figure 3 for simple repression with the O1 binding site for a single chromosomal copy. The inset shows the predicted scaling (lines) and measured fold-change (points) for three distinct repressor copy numbers, as the number of promoter copies is varied.

(B) Fold-change as a function of concentration of Lac repressor for multiple gene copies on the chromosome. The red symbols are measurements of the fold-change in expression at the end of the cell cycle of a strain with the Oid simple repression motif integrated into five unique sites on the chromosome. We expect ten copies of the gene at the end of the cell cycle. The red line is the theory prediction for multiple identical gene copies with N=10, from Equation 4. The shaded region represents the uncertainty from the measured value of $\Delta\epsilon$. The blue symbols and line are the data and theory prediction for simple repression with the Oid binding site from Figure 3. In both cases, the fold-change is ~1 when the copy number of the repressor is less than the copy number of the gene. At high repressor copy number, the curve coincides with simple repression from the chromosome, with a sharp transition between these two regimes.

Error bars in fold-change measurements are SEM. Error bars in LacI copy number are the quadrature summed errors from the calibration factor and the inherent resolution limit of LacI detection. Error bars in promoter copy number reflect uncertainty in the qPCR measurement of average plasmid copy number.

single chromosomal copy of our simple repression construct as a function of the number of repressors per cell for different binding sites (filled symbols) using the dilution method and video microscopy advocated here. The limits of our measurement both at low repressor numbers and at low fold-change (wherein repressed YFP production becomes small) are discussed in the Extended Experimental Procedures. The lines are the theory predictions from Equation 3 for each operator without any fit parameters, with a shaded region representing the uncertainty in $\Delta\epsilon$. One assumption in this simple theory of Equation 3 is that the copy number of the gene is exactly one. In reality, the copy number of our single integrated copy varies between 1 and 2 over the course of the cell cycle (Bremer and Dennis, 1996). However, the predicted expression for two chromosomal copies, Equation 4 with $N=2$, is identical to Equation 3 when $R\gg N$. Thus, the promoters will express independently, and we can ignore this small correction (see Figure S5). The data from Garcia and Phillips (2011) are shown as open symbols in the figure. These results lead to the interesting conclusion that single-cell measurement of the expression rates agrees precisely with previous bulk measurements of steady-state expression.

Predicting Expression Levels from Plasmid Constructs as a Function of Gene Copy Number

We now wish to compare the predictions of the thermodynamic theory against the more complicated cases involving TF binding. In this section, we compare the predictions of Equation 4 to measurements of expression from plasmid, as illustrated in Figure 1B.

To begin, we measure the expression of an O1 simple repression construct placed on either the ColE1 or ColE1 Δrom plasmids, akin to Figure 1B (for details on plasmid construction, see Figure S6). The fold-change in gene expression as a function of Lac repressor copy number is shown for both plasmids in Figure 4A. The data shown here are taken from the chronological middle of the cell cycle; the effect of the evolution of the copy number throughout the cell cycle on expression will be addressed later. The solid lines in the figure are plots of the predictions from Equation 4, with no adjustable parameters. The shaded region accounts for the standard deviation in N from our qPCR measurements of the average copy number and the uncertainty in the binding energy $\Delta\epsilon$. For reference, the green points and line are the chromosomal data and theory for the

O1 operator from Figure 3. The theory predicts the fold-change in expression and captures the major features observed in our data. When the repressor copy number is much larger than the gene copy number, the fold-change is relatively unchanged with respect to the single-copy chromosomal case, as predicted for the case $R \gg N$. However, when the repressor copy number is less than the gene copy number, the effect of the repressors is largely buffered away and the repressors have little effect on the fold-change. Between these two regions, the transition is sharp and switch-like. An alternative way to look at the data and its agreement with the theoretical predictions is to plot the fold-change as a function of the promoter copy number for a defined number of repressors. The inset to Figure 4A shows these data for three distinct repressor copy numbers (8 black, 64 violet, 256 cyan). The distinct values for promoter copy number, N, are obtained by taking data from the simple repression O1 construct taken at the end of the cell cycle ($N = 2$) in addition to the plasmid data with and without Rom ($N = 52$ and 64, respectively).

Simple Thermodynamic Model Predicts Expression Levels from Multiple Integrated Chromosomal Gene Copies

Plasmids can differ from the chromosome in their relative distribution throughout the cell, the accessibility to TFs, and their segregation mechanisms (Ebersbach and Gerdes, 2005; Ghosh et al., 2006; Kuhlman and Cox, 2012). However, the effect of the interplay between repressor and gene copy number is not exclusive to constructs located on plasmids. The same regulatory features can be seen at low gene copy numbers from multiple copies of the gene located on the chromosome. We measure gene expression using a strain that has the Oid simple repression construct integrated into the chromosome in five different locations, as schematically shown in Figure 1C. To avoid uncertainty in the copy number of the gene, we examine expression near the end of the cell cycle (the last 15 min before division). We expect that each of the five chromosomal copies will have fully replicated (the D period of the cell cycle, the time between replication completion and division, is roughly 27 min at our growth rate [Bremer and Dennis, 1996]). Therefore, the copy number of the gene should be ten, resulting from two sets of five copies, one on each completed chromosome (Bremer and Dennis, 1996; Michelsen et al., 2003). The resulting fold-change data for this construct are shown in Figure 4B as red symbols. The red curve is the prediction from Equation 4 for $N = 10$ and the Oid binding energy with, once again, no fit parameters. The shaded region in the fit comes from the uncertainty in $\Delta \epsilon$ for the Oid binding site. For reference, the blue curve and symbols are the theory and data (from Figure 3) for the Oid construct integrated at a single copy in the chromosome. The observed behavior of this construct is qualitatively comparable to that observed for genes on plasmids (Figure 4A). Additionally, the same theory predicts its quantitative features without any free parameters. Once again, there is a sharp drop in fold-change when the number of repressors equals the gene copy number before rejoining the predictions (and measurements) for fold-change from a single gene copy. The genetic locations of integration and a discussion of the distribution and uncertainty

in the measurement of gene copy number can be found in the Extended Experimental Procedures section "The copy number of multiple chromosomal integrations strain" and Figure S4.

Predicting Expression Levels in Complex TF-Binding Landscapes

Finally, a common situation that results in competition for TFs occurs when different genes share the same TF. For example, in the regulatory databases of RegulonDB for *E. coli* (Gama-Castro et al., 2011), three-quarters of TFs are listed as having specific interactions with more than one operon. In fact, many of these TFs target dozens of operons (and it is worth noting that these databases are far from complete and represent only a partial list of binding interactions, implying that these numbers will continue to grow). In this case, the competing specific binding sites that do not modulate the gene of interest may outcompete the gene copy when TFs are limiting.

To examine this competition scenario, we measure the expression from a single copy of the O1 simple repression construct integrated on the chromosome (identical to the O1 construct from Figure 3) in a cellular context containing competitor binding sites. These binding sites are carried in a high copy number ColE1 Δrom plasmid that does not express a gene product, illustrated schematically in Figure 1D. In Figure 5A, we show the measured fold-change of the chromosomal O1 construct in the case in which the competing plasmid has a weaker O2 binding site (green symbols), equal strength O1 binding site (red symbols), or a stronger Oid binding site (black symbols). The theory curves stemming from Equation 5 are shown in the corresponding color, with the shaded region corresponding to the uncertainty in the theory stemming from the uncertainty in $N_c, \Delta \epsilon$, and $\Delta \epsilon_c$ (see Extended Experimental Procedures section "Determining errors in theoretical predictions"). The stronger binding sites shift and sharpen the transition of the gene of interest with respect to LacI copy number. Once the repressor copy number exceeds the number of competitors, the gene finally gains access to the repressor and becomes regulated. In contrast, when the competitors are weaker, the position of the transition is shifted toward lower LacI numbers. This simple example illustrates how the regulatory behavior of a gene can be indirectly controlled through identical competitor plasmids. This effect is, however, more general, as in wild-type genes, the competition will come from a spectrum of binding sites each controlling a specific gene. The thermodynamic model can produce predictions for any such specific arrangement of binding sites with the theoretical infrastructure demonstrated here.

The Influence of Plasmid Distribution on the Repressor Titration Curves

One unsatisfactory feature of Figure 5A is that the observed transition for strong binding sites (Oid and, to a lesser extent, O1) is not as sharp as predicted by the theory. Furthermore, the blue bars in Figure 5C show the resulting fit values for the mean competitor plasmid copy number if Equation 5 is fit to the data in Figure 5A. The plasmids used in each of these three measurements are nearly identical. The only difference between them is the strength of their LacI-binding sites. Thus, we would expect
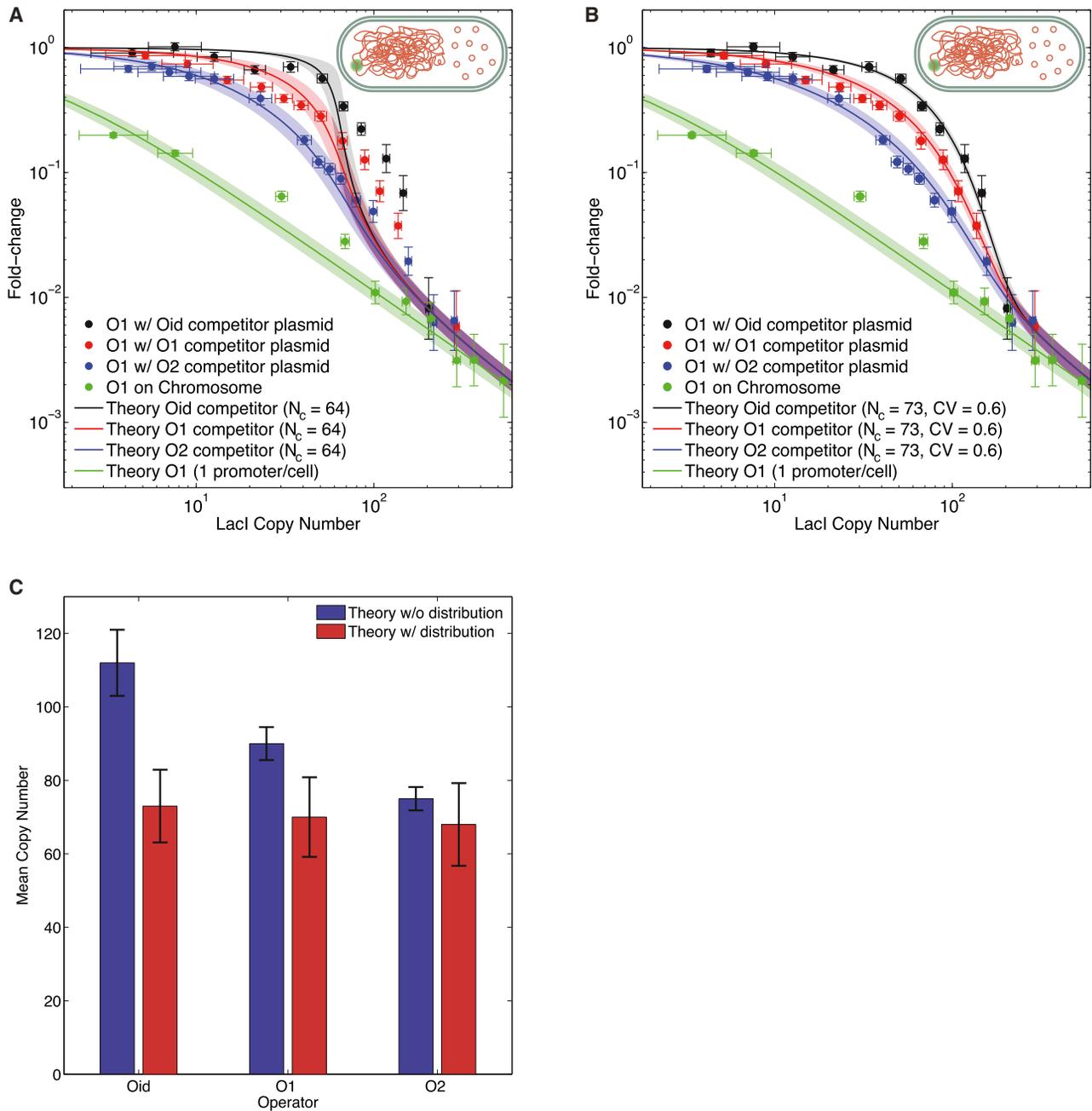
Figure 5. Effects of Repressor Competition on Expression

(A) Fold-change as a function of concentration of Lac repressor for the O1 simple repression construct integrated into the chromosome in competition with a ColE1 Δ*rom* plasmid containing a stronger (Oid; black symbols), equal (O1; red symbols), or weaker (O2; blue symbols) Lac repressor binding site. For reference, the green symbols and line are the data and theory prediction from Figure 3 for simple repression with the O1 binding site without the competitor plasmid.

(B) The same data from (A), but now the solid lines represent the plasmid distribution theory assuming a normal distribution. The parameters are found by fitting the Oid (black) data for N, the average copy number, and the CV, and these parameters are plotted for each binding energy, i.e., the red and blue curve are parameter free. Error bars in fold-change measurements are SEM. Error bars in copy number are the quadrature summed errors from the calibration factor and the inherent resolution limit of LacI detection.

(C) Fitted means of all three plasmid copy numbers for both the theory in Equation 5, which assumes a single static copy number for plasmid (blue bars, $N_c^{Oid} = 112, N_c^{O1} = 90, N_c^{O2} = 75$), and the same theory in which the plasmid distribution is normal with CV, as determined from fitting the Oid data (red bars, $N_c^{Oid} = 73, N_c^{O1} = 70, N_c^{O2} = 68$). All three plasmids in this case have the same origin of replication and differ only by a few bases, which alter Lac repressor affinity to their binding sites. As such, we expect all three plasmids to have identical copy numbers. This is observed for the theory with plasmid distribution (red bars); however, the theory without a plasmid distribution systematically overestimates the copy number for stronger binding (blue bars). Error bars in mean copy number are the results of bootstrap sampling the expression measurements of individual cells.

the mean copy number to be unchanged in the Oid, O1, or O2 strain. However, it is clear that stronger binding sites systematically predict a higher copy number, showing that, in this case, the model lacks internal consistency.

These discrepancies, both in the fit to the sharpness of the transition region and in the measured copy number between similar plasmids, likely stem from the theory not accounting for the fact that there is a distribution of plasmids in the population of cells (Guido et al., 2006; Ghozzi et al., 2010; Wong Ng et al., 2010). This distribution will result in a less sharp transition in the fold-change curve. To see intuitively how a distribution of plasmid copy numbers alters the fold-change curve, we imagine the situation of a simple distribution in which cells have a single chromosomal YFP gene, with half of the cells having N and the other half having $3N$ competing plasmids. Further, assume that the binding site on these competing plasmids has an extremely high affinity. The average plasmid copy number is $2N$, and it is at this value that the thermodynamic model predicts a sharp transition in the fold-change curve. However, when the number of repressors is $R = 2N$, all of the cells with N plasmids are repressed by free repressors and produce very little YFP because all of the competing plasmids are saturated. On the other hand, the cells with $3N$ plasmids will still not be repressed because the competing plasmid can buffer the $2N$ available repressors. Hence, for the entire population, the fold-change is no less than 1/2. Only when the repressor copy number reaches $3N$ will repression in every cell ensue and begin to show a steep drop on a log scale. This simple argument provides the intuition for why a distribution of plasmids is required in the theory when thinking about the strong operator limit on the competitor plasmids.

Generically, a distribution of plasmid copy number in a population of cells will move the location of the switch-like transition to repressor numbers above the mean plasmid copy number, and the transition will be softened. This effect is stronger as the copy number distribution becomes wider than the width of the transition region in the fold-change curve. Therefore, we expect that, the stronger the plasmid binding site, the worse the simple single copy number theory will fit. This is observed for the data in Figure 5A wherein O2 fits well to the simple theory, O1 fits worse, and Oid fits even worse.

It is relatively simple to account for a distribution of plasmids in the thermodynamic theory; however, the derivation is left to the SI. In the case of identical copies of the gene, Equation 4 must be modified such that the fold-change in the presence of a distribution, fold-change$_{dist}$, is related to the fold-change of the fixed copy number fold-change by,

$$\text{fold-change}_{dist} = \sum_{n=0}^{\infty} p(n) \frac{n}{\langle n \rangle} \text{fold-change}(n), \qquad (8)$$

where $p(n)$ is the probability that any cell in the population has n plasmids with $\langle n \rangle = \sum_{n=0}^{\infty} np(n)$ the average number of plasmids in the population. Finally, fold-change(n) is calculated using Equation 4 for any given value of n. Similarly, the theory for competitor binding sites on a plasmid can be adjusted for a plasmid distribution in a similar fashion. In this case, we find,

$$\text{fold-change}_{dist} = \sum_{n=0}^{\infty} p(n) \text{fold-change}(n), \qquad (9)$$

where now fold-change(n) refers to Equation 5 with n plasmids.

To exploit these ideas in the context of our data, we propose a simple phenomenological distribution for the plasmid copy number, a normal distribution with a fixed coefficient of variation (SD over mean). We have chosen a normal distribution for simplicity, as we do not expect the exact details of the distribution to have a major effect, given that a plasmid distribution will always dull the sharp transition of the fold-change repressor titration curve. We fit the Oid competitor data from Figure 5A to the distribution, treating the mean copy number (N_c) and coefficient of variation (CV), σ / N_c of the distribution as fit parameters. We find that the best fit mean is $N_c = 72$ with $CV = 0.6$; the resulting fit is shown in Figure 5B as the solid black line. We also plot the data from the O2 and O1 plasmid with the same values for the mean and standard deviation.

The theory, which was fit to the Oid data, now describes the data from all three operator sites very well. This is an important sanity check, as we do not expect the strength of a LacI-binding site far from the origin of replication to effect the plasmid copy number or its distribution. This is further demonstrated in Figure 5C, wherein we plot the mean plasmid copy number measured by either fitting the simple no distribution model to our data or fitting the mean copy number while holding the width of the distribution fixed, with $CV = 0.6$. The point is illustrated in these bar graphs. When the transition is not sharp, as in the O2 data (and, to a lesser extent, the O1 data), the fixed N_c single-parameter theory fits well and with a copy number consistent with what we expect. However, as the competitor binding gets strong and the fold-change response curve is expected to get sharp compared to the distribution of plasmids, the single parameter fixed N_c theory fits poorly, and N_c goes from being descriptive of the actual copy number to merely a phenomenological fit parameter. However, the fit to the mean copy number for the three operators remains consistent when the distribution of plasmids is accounted for in the model.

Cell-Cycle Dependence of the Plasmid Copy Number and the Resulting Expression

To this point, we have varied the copy number of competing binding sites or gene copy number by comparing the fold-change of different constructs at similar points in the cell cycle. However, it is clear that, over the course of the cell cycle, all genetic material in the cell must double. As a result, we examine the time dependence of the copy number by binning the data according to when in the cell cycle each measurement is made. In this metric, the time of birth of the cell is represented as 0, and the time of its subsequent division is 1. In Figure 6A, an example of the fold-change curves obtained are shown for the O1 simple repression chromosomal integration with the Oid competitor plasmid. In this case, each time bin is fit to Equation 9 for the copy number N_c, keeping the coefficient of variation fixed; the resulting copy number for that point in the cell cycle is written in the legend. As expected, the measured average plasmid copy number increases as the cell cycle progresses.
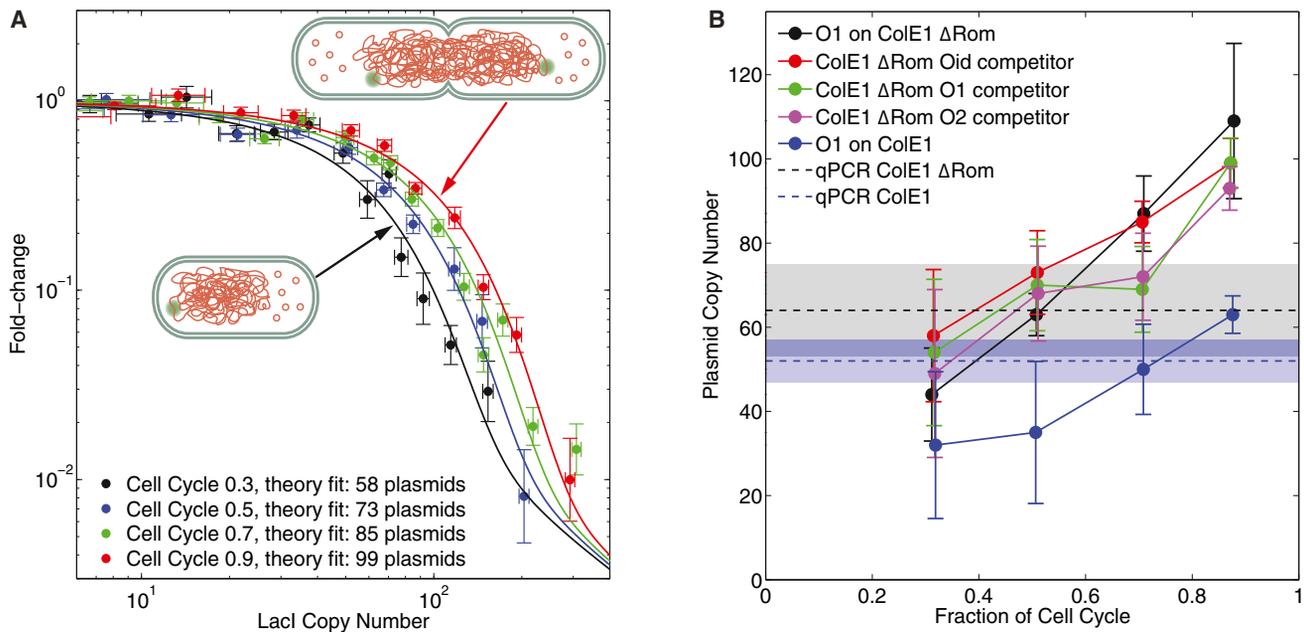
Figure 6. Variation of Plasmid Copy Number throughout the Cell Cycle

(A) Fold-change as a function of LacI copy number for the O1 simple repression construct integrated into the chromosome in the presence of a competing ColE1 Δ*rom* plasmid bearing an Oid site for different time points in the cell cycle. The "cell cycle" parameter is the average fraction of the total cell lifetime from which the binned data are taken, with 0 representing birth and 1 representing the cell division. The plasmid copy number is fit to Equation 9 at every time point, keeping CV = 0.6 fixed, and the resulting value for *N* is listed. Error bars in fold-change measurements are SEM. Error bars in LacI copy number are the quadrature summed errors from the calibration factor and the inherent limit of LacI detection.

(B) Plasmid copy number versus the cell cycle. The plasmid copy number is measured by fitting the copy number parameter in the theory to the data from all of our experiments binned by time in the cell cycle. Error bars in plasmid copy number are the results of bootstrap sampling the expression measurements of individual cells. The horizontal dashed lines and matching shaded regions are our qPCR measurements for average copy number of ColE1 (blue dashed line) and ColE1 Δ*rom* (black dashed line).

In addition, the copy on the chromosome will double. The operator associated with this copy will affect our measurement of N_c; however, the addition of one extra operator in the presence of dozens of copies on plasmid results in only a very small change to the predicted fold-change. The exact details of the size of this effect are discussed in the SI section "Accounting for chromosome replication in competitor theory." Repeating this process for all plasmids with the proper theory equations (Equation 8 for identical plasmids expressing YFP, Equation 9 for the competitor plasmid data), we plot in Figure 6B the measured plasmid copy number versus fractional cell cycle for each. The horizontal dashed lines and corresponding shaded region represent our qPCR measurements of the average copy number for ColE1 and ColE1 Δ*rom*. It should be noted that the cell-cycle parameter relates to when the measurement itself was made; due to fluorophore maturation, the actual measurement may represent a time period earlier in the cell cycle.

DISCUSSION

Recent experimental and theoretical efforts have focused on understanding the role of regulatory architecture in transcriptional decisions. In these studies, the details of isolated regulatory architectures (number, location, and strength of binding sites and TF copy number) are varied systematically, and the transcriptional output is compared to corresponding theoretical predictions (Buchler et al., 2003; Bintu et al., 2005a, 2005b). However, it is rarely the case that TFs act on only one promoter. As a result, the study of transcriptional decisions at individual promoters without taking into account the rest of the regulatory network might be insufficient. In particular, the presence of multiple targets for the same TF can result in a competition that reduces the available free TFs for the gene of interest (Rydenfelt et al., 2014).

In this paper, we explored the interplay between the binding sites for a transcription factor at a gene of interest and competing binding sites regulating other genes in the context of the well-studied simple repression architecture (Garcia and Phillips, 2011). We show that the presence of competing binding sites not only changes the effective amount of available TF, but also can affect the input-output relation by introducing a sharp transition. This transition separates the regime in which the repressor is depleted compared to the number of available binding sites and the regime in which the repressor is in excess of the number of binding sites. The width of this transition is controlled by the strength of the binding sites. The theory also predicts that the width of this transition, in a population of cells, depends on the size of cell-to-cell fluctuations in binding site copy number; larger fluctuations in the number of available competitor binding sites (or number of identical genes) tends to flatten this transition.

We find that, when a very strong transition is predicted, the measured transition is considerably dulled, which we attribute to the copy number variability in the population.

The quantitative consequences of binding site competition can be predicted using thermodynamic models without any free fitting parameters. Previously, these models had been successful in predicting transcription output for the simple repression architecture in the absence of binding site competition (Garcia and Phillips, 2011). By measuring binding site copy numbers using qPCR, we show that an extended version of these models accounting for the presence of multiple binding sites (Rydenfelt et al., 2014) describes our data precisely, with no fit parameters for a wide range of binding site copy numbers and strengths.

Building on the success of the theory in quantitatively predicting the regulatory outcome of the various architectures considered here, we fit the theory to the fold-change curve at different points in the cell cycle as a way to measure the time evolution of the plasmid copy number during the course of the cell cycle. One noteworthy feature of this method is that the reporter fluorescence molecule need not be expressed by the measured plasmid; it only requires that a TF be shared between an unrelated chromosomal copy and the plasmid for the copy number of the competitor to be measured. This can be of benefit, as it requires only inserting a binding site on the plasmid of interest. Additionally, this approach prevents possible changes in cell physiology due to starvation or phototoxicity from overexpression and measurement of a fluorescent reporter protein expressed from a high copy plasmid.

The ability to measure absolute numbers of both binding site and input TF copy number is key for contrasting our experimental data with the theoretical predictions stemming from thermodynamic models of transcriptional regulation (Bintu et al., 2005a, 2005a). In this paper, we have made use of a recently introduced fluctuation method for taking the repressor census and thereby checking the governing equation for the simple repression regulatory motif in a wide array of situations in which the TF was in demand from multiple sources (Rosenfeld et al., 2005; Teng et al., 2010). We find that this dilution method, a form of video microscopy, gives quantitatively compatible results to more traditional steady-state snapshots and bulk measurements (Garcia and Phillips, 2011). However, there are numerous advantages of the dilution method. This method provides a single-cell readout of the number of repressors in each cell. The fact that no new repressors are produced ensures that this "input" level of repressors is held constant through the entire measurement. The single-cell nature of this method certainly increases our resolution as compared to measuring a bulk sample in which the possible distribution of repressors from cell to cell can have a wider distribution than the feature we wish to illuminate—akin to the issues that we see when the plasmid distribution is wider than the sharp transitions that we wish to study, seen in Figure 5A. Finally, this method allows one to probe particular regions of the titration curve with varying degrees of resolution—a feature that was essential when trying to distinguish sharp features of the repressor titration curve at 10 repressors for the chromosomal case and >100 repressors in the plasmid case.

As our characterization of cellular decision making becomes more quantitative, so must our theoretical description of this process. The quantitative and predictive control of such decisions allows one to probe their molecular details at a level that escapes any qualitative description. In addition to expanding our understanding of regulation, quantitative models give us the ability to control regulatory output by predictive design. In fact, synthetic biology has focused on the development of standardized regulatory units with known input-output functions (Endy, 2005; Voigt, 2006). The modification of these input-output functions to have, for example, a particular shape usually requires the re-engineering of the regulatory architecture at the DNA level (Rosenfeld et al., 2005; Guido et al., 2006; Cox et al., 2007). Our work provides a complementary approach to controlling input-output functions, as the introduction of competing binding sites for a TF into a cell makes it possible to tune regulatory response in a predictive fashion, without the need for any modifications at the DNA of the gene of interest.

EXPERIMENTAL PROCEDURES

Gene Expression Measurements

Cultures are grown overnight in 2 ml of LB at 37°C and were diluted $\sim 1:10^4$ in M9 + 0.5% glucose minimal media with antibiotics and 1, 2, 3, 4, 6, 8, 100 ng/ml anhydrotetracycline (aTc) to induce the production of various levels of LacI-mCherry that cover the full repressor range (for induction curve, see Figure S1). The diluted cultures are grown at 37°C until they reach an OD600 $\approx 0.2 - 0.4$ and then they are washed twice with fresh M9 media (without aTc) to remove the inducer and halt the production of LacI-mCherry. They are then diluted to give several cells per field of view (at 100× magnification) when placed on a 2% low melting point M9 + 0.5% glucose agar pad. An automated fluorescent microscope simultaneously records multiple fields of view for each concentration of aTc. In addition, one pad contains cells without the repressor construct whose expression measurements serve as the denominator of our fold-change measurements. Growth of cells is observed by fluorescence microscopy at 37°C for 2.5 hr (at 7.5 min per frame) while measuring CFP (used for segmentation), YFP (reporter of gene expression), and mCherry (reporter of TF copy number) intensities.

Data Analysis

Data analysis was performed using the Matlab code "Schnitzcells" kindly provided by Michael Elowitz (Rosenfeld et al., 2005). This code segments cells in a movie and tracks their lineages, providing full information of fluorescent intensity of all three channels for each cell at each time point in the movie. The rate of gene expression is calculated between each time step by taking the difference in consecutive YFP signals (once corrected for autofluorescence, bleaching, and crosstalk). The TF concentration is calculated from the total mCherry signal in any given cell.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, six figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2014.02.022.

AUTHOR CONTRIBUTIONS

REFERENCES

Aitman, T.J., Dong, R., Vyse, T.J., Norsworthy, P.J., Johnson, M.D., Smith, J., Mangion, J., Roberton-Lowe, C., Marshall, A.J., Petretto, E., et al. (2006). Copy number polymorphism in Fcgr3 predisposes to glomerulonephritis in rats and humans. Nature *439*, 851–855.

Bhattacharyya, R.P., Reményi, A., Yeh, B.J., and Lim, W.A. (2006). Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. Annu. Rev. Biochem. *75*, 655–680.

Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., Kuhlman, T., and Phillips, R. (2005a). Transcriptional regulation by the numbers: applications. Curr. Opin. Genet. Dev. *15*, 125–135.

Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., and Phillips, R. (2005b). Transcriptional regulation by the numbers: models. Curr. Opin. Genet. Dev. *15*, 116–124.

Bremer, H., and Dennis, P.P. (1996). Modulation of chemical composition and other parameters of the cell by growth rate. In *Escherichia coli* and *Salmonella* Cellular and Molecular Biology (Washington, DC: ASM Press), pp. 1553–1569.

Brewster, R.C., Jones, D.L., and Phillips, R. (2012). Tuning promoter strength through RNA polymerase binding site design in *Escherichia coli*. PLoS Comput. Biol. *8*, e1002811.

Buchler, N.E., Gerland, U., and Hwa, T. (2003). On schemes of combinatorial transcription logic. Proc. Natl. Acad. Sci. USA *100*, 5136–5141.

Busby, S., and Ebright, R.H. (1999). Transcription activation by catabolite activator protein (CAP). J. Mol. Biol. *293*, 199–213.

Cesareni, G., Muesing, M.A., and Polisky, B. (1982). Control of ColE1 DNA replication: the rop gene product negatively affects transcription from the replication primer promoter. Proc. Natl. Acad. Sci. USA *79*, 6313–6317.

Cox, R.S., 3rd, Surette, M.G., and Elowitz, M.B. (2007). Programming gene expression with combinatorial promoters. Mol. Syst. Biol. *3*, 145.

Daber, R., Sochor, M.A., and Lewis, M. (2011). Thermodynamic analysis of mutant lac repressors. J. Mol. Biol. *409*, 76–87.

Dekel, E., and Alon, U. (2005). Optimality and evolutionary tuning of the expression level of a protein. Nature *436*, 588–592.

Del Vecchio, D., Ninfa, A.J., and Sontag, E.D. (2008). Modular cell biology: retroactivity and insulation. Mol. Syst. Biol. *4*, 161.

Ebersbach, G., and Gerdes, K. (2005). Plasmid segregation mechanisms. Annu. Rev. Genet. *39*, 453–479.

Elowitz, M., and Lim, W.A. (2010). Build life to understand it. Nature *468*, 889–890.

Endy, D. (2005). Foundations for engineering biology. Nature *438*, 449–453.

Gama-Castro, S., Salgado, H., Peralta-Gil, M., Santos-Zavaleta, A., Muñiz-Rascado, L., Solano-Lira, H., Jimenez-Jacinto, V., Weiss, V., García-Sotelo, J.S., López-Fuentes, A., et al. (2011). RegulonDB version 7.0: transcriptional

regulation of *Escherichia coli* K-12 integrated within genetic sensory response units (Gensor Units). Nucleic Acids Res. *39* (Database issue), D98–D105.

Garcia, H.G., and Phillips, R. (2011). Quantitative dissection of the simple repression input-output function. Proc. Natl. Acad. Sci. USA *108*, 12173–12178.

Garcia, H.G., Sanchez, A., Kuhlman, T., Kondev, J., and Phillips, R. (2010). Transcription by the numbers redux: experiments and calculations that surprise. Trends Cell Biol. *20*, 723–733.

Ghosh, S.K., Hajra, S., Paek, A., and Jayaram, M. (2006). Mechanisms for chromosome and plasmid segregation. Annu. Rev. Biochem. *75*, 211–241.

Ghozzi, S., Wong Ng, J., Chatenay, D., and Robert, J. (2010). Inference of plasmid-copy-number mean and noise from single-cell gene expression data. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. *82*, 051916.

Guido, N.J., Wang, X., Adalsteinsson, D., McMillen, D., Hasty, J., Cantor, C.R., Elston, T.C., and Collins, J.J. (2006). A bottom-up approach to gene regulation. Nature *439*, 856–860.

Hanada, K., Sawada, Y., Kuromori, T., Klausnitzer, R., Saito, K., Toyoda, T., Shinozaki, K., Li, W.H., and Hirai, M.Y. (2011). Functional compensation of primary and secondary metabolites by duplicate genes in *Arabidopsis thaliana*. Mol. Biol. Evol. *28*, 377–382.

Kaplan, S., Bren, A., Zaslaver, A., Dekel, E., and Alon, U. (2008). Diverse two-dimensional input functions control bacterial sugar genes. Mol. Cell *29*, 786–792.

Kentner, D., and Sourjik, V. (2010). Use of fluorescence microscopy to study intracellular signaling in bacteria. Annu. Rev. Microbiol. *64*, 373–390.

Kim, K.H., and Sauro, H.M. (2011). Measuring retroactivity from noise in gene regulatory networks. Biophys. J. *100*, 1167–1177.

Kinney, J.B., Murugan, A., Callan, C.G., Jr., and Cox, E.C. (2010). Using deep sequencing to characterize the biophysical mechanism of a transcriptional regulatory sequence. Proc. Natl. Acad. Sci. USA *107*, 9158–9163.

Kuhlman, T.E., and Cox, E.C. (2012). Gene location and DNA density determine transcription factor distributions in *Escherichia coli*. Mol. Syst. Biol. *8*, 610.

Kuhlman, T., Zhang, Z., Jr., Saier, M.H., Jr., and Hwa, T. (2007). Combinatorial transcriptional control of the lactose operon of *Escherichia coli*. Proc. Natl. Acad. Sci. USA *104*, 6043–6048.

Lee, T.H., and Maheshri, N. (2012). A regulatory role for repeated decoy transcription factor binding sites in target gene expression. Mol. Syst. Biol. *8*, 576.

Lee, C., Kim, J., Shin, S.G., and Hwang, S. (2006a). Absolute and relative QPCR quantification of plasmid copy number in *Escherichia coli*. J. Biotechnol. *123*, 273–280.

Lee, C.L., Ow, D.S., and Oh, S.K. (2006b). Quantitative real-time polymerase chain reaction for determination of plasmid copy number in bacteria. J. Microbiol. Methods *65*, 258–267.

Lim, W.A. (2002). The modular logic of signaling proteins: building allosteric switches from simple binding domains. Curr. Opin. Struct. Biol. *12*, 61–68.

Luria, S.E., and Dulbecco, R. (1949). Genetic Recombinations Leading to Production of Active Bacteriophage from Ultraviolet Inactivated Bacteriophage Particles. Genetics *34*, 93–125.

Lutz, R., and Bujard, H. (1997). Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. Nucleic Acids Res. *25*, 1203–1210.

Madan Babu, M., and Teichmann, S.A. (2003). Functional determinants of transcription factors in Escherichia coli: protein families and binding sites. Trends Genet. *19*, 75–79.

Michelsen, O., Teixeira de Mattos, M.J., Jensen, P.R., and Hansen, F.G. (2003). Precise determinations of C and D periods by flow cytometry in *Escherichia coli* K-12 and B/r. Microbiology *149*, 1001–1010.

Mukherji, S., and van Oudenaarden, A. (2009). Synthetic biology: understanding biological design from synthetic circuits. Nat. Rev. Genet. *10*, 859–871.

Navarro-Quezada, A., and Schoen, D.J. (2002). Sequence evolution and copy number of Ty1-copia retrotransposons in diverse plant genomes. Proc. Natl. Acad. Sci. USA *99*, 268–273.

Oehler, S., Amouyal, M., Kolkhof, P., von Wilcken-Bergmann, B., and Müller-Hill, B. (1994). Quality and position of the three *lac* operators of *E. coli* define efficiency of repression. EMBO J. *13*, 3348–3355.

Ozbudak, E.M., Thattai, M., Lim, H.N., Shraiman, B.I., and Van Oudenaarden, A. (2004). Multistability in the lactose utilization network of *Escherichia coli*. Nature *427*, 737–740.

Paulsson, J., and Ehrenberg, M. (2001). Noise in a minimal regulatory network: plasmid copy number control. Q. Rev. Biophys. *34*, 1–59.

Ptashne, M., and Gann, A. (2002). Genes and Signals (New York: Cold Spring Harbor Laboratory Press).

Ricci, F., Vallée-Bélisle, A., and Plaxco, K.W. (2011). High-precision, in vitro validation of the sequestration mechanism for generating ultrasensitive dose-response curves in regulatory networks. PLoS Comput. Biol. *7*, e1002171.

Rojo, F. (2001). Mechanisms of transcriptional repression. Curr. Opin. Microbiol. *4*, 145–151.

Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S., and Elowitz, M.B. (2005). Gene regulation at the single-cell level. Science *307*, 1962–1965.

Rosenfeld, N., Perkins, T.J., Alon, U., Elowitz, M.B., and Swain, P.S. (2006). A fluctuation method to quantify in vivo fluorescence data. Biophys. J. *91*, 759–766.

Rydenfelt, M., Cox, R.S., III, Garcia, H.G., and Phillips, R. (2014). Statistical mechanical model of coupled transcription from multiple promoters due to transcription factor titration. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. *89*, 012702.

Sanchez, A., Osborne, M.L., Friedman, L.J., Kondev, J., and Gelles, J. (2011). Mechanism of transcriptional repression at a bacterial promoter by analysis of single molecules. EMBO J. *30*, 3940–3946.

Schlax, P.J., Capp, M.W., and Record, M.T., Jr. (1995). Inhibition of transcription initiation by lac repressor. J. Mol. Biol. *245*, 331–350.

Stueber, D., and Bujard, H. (1982). Transcription from efficient promoters can interfere with plasmid replication and diminish expression of plasmid specified genes. EMBO J. *1*, 1399–1404.

Teng, S.W., Wang, Y., Tu, K.C., Long, T., Mehta, P., Wingreen, N.S., Bassler, B.L., and Ong, N.P. (2010). Measurement of the copy number of the master quorum-sensing regulator of a bacterial cell. Biophys. J. *98*, 2024–2031.

Twigg, A.J., and Sherratt, D. (1980). Trans-complementable copy-number mutants of plasmid ColE1. Nature *283*, 216–218.

Vilar, J.M., and Leibler, S. (2003). DNA looping and physical constraints on transcription regulation. J. Mol. Biol. *331*, 981–989.

Voigt, C.A. (2006). Genetic parts to program bacteria. Curr. Opin. Biotechnol. *17*, 548–557.

Wang, S., Liu, N., Peng, K., and Zhang, Q. (1999). The distribution and copy number of copia-like retrotransposons in rice (*Oryza sativa* L.) and their implications in the organization and evolution of the rice genome. Proc. Natl. Acad. Sci. USA *96*, 6824–6828.

Wong Ng, J., Chatenay, D., Robert, J., and Poirier, M.G. (2010). Plasmid copy number noise in monoclonal populations of bacteria. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. *81*, 011909.
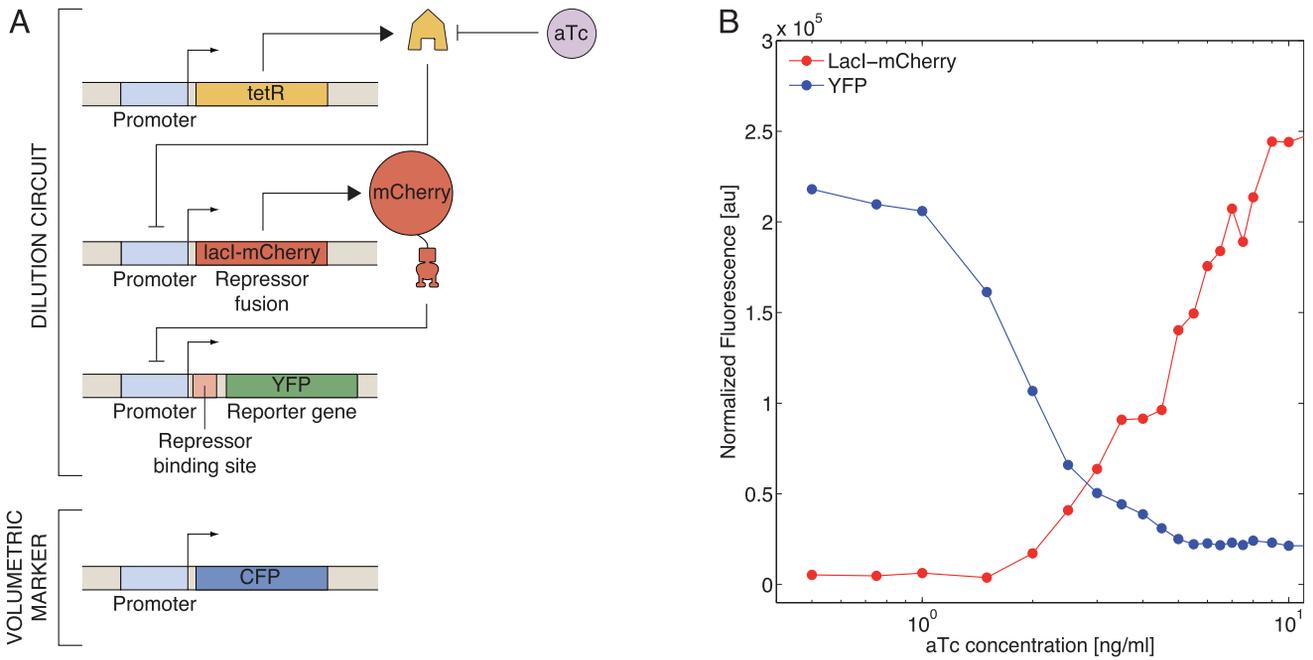
Figure S1. Experimental Circuit Schematic and Response to aTc, Related to Figure 2

(A) Full schematic of dilution circuit. LacI-mCherry regulates the target gene which expresses YFP. The LacI-mCherry is expressed from a P$_{LtetO-1}$ promoter (Lutz and Bujard, 1997). When the small inducer molecule aTc is not present, TetR, which is expressed from a P$_{N25}$ promoter, shuts off production of LacI-mCherry. When saturating amounts of aTc are present, the LacI-mCherry level can be induced to roughly 1000 dimers per cell. Finally, CFP is constitutively expressed from the chromosome and is used as a volume marker for segmentation.

(B) Induction of the dilution circuit. Steady state fluorescence level per cell of LacI-mCherry and YFP as a function of aTc concentration in the O2 single chromosomal copy strain (used for the green data in Figure 3). Fluorescence is measured in bulk using a plate reader as described in Garcia and Phillips (2011).
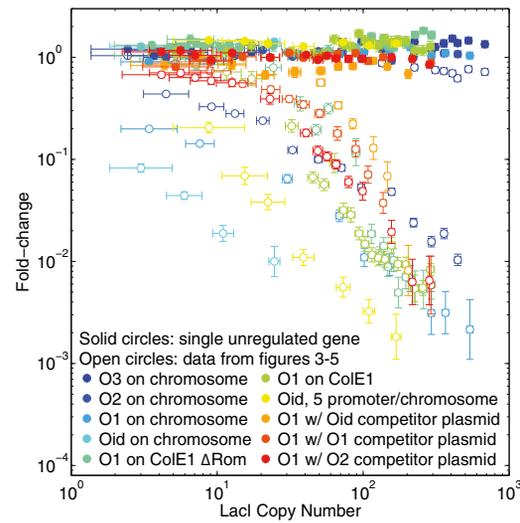
Figure S2. Comparison of Fold-Change of *lac*UV5 Promoter with and without Repressor Binding Site, Related to Figures 3, 4, and 5

Fold-change of unregulated CFP expression from a single integrated gene copy (solid circles) compared to the fold-change data from Figures 3, 4, and 5 (open circles). Like-colored data is taken simultaneously in the same population of cells and the promoter in both cases is *lac*UV5. The induction of repressor does not have any obvious secondary effects on the global transcription as demonstrated by the fold-change of the unregulated promoters staying constant at 1 while the same promoter with a repressor binding site exhibits orders of magnitude lower expression. Error bars in fold-change measurements are SEM. Error bars in LacI copy number are the quadrature summed errors from the calibration factor and the inherent resolution limit of LacI detection.
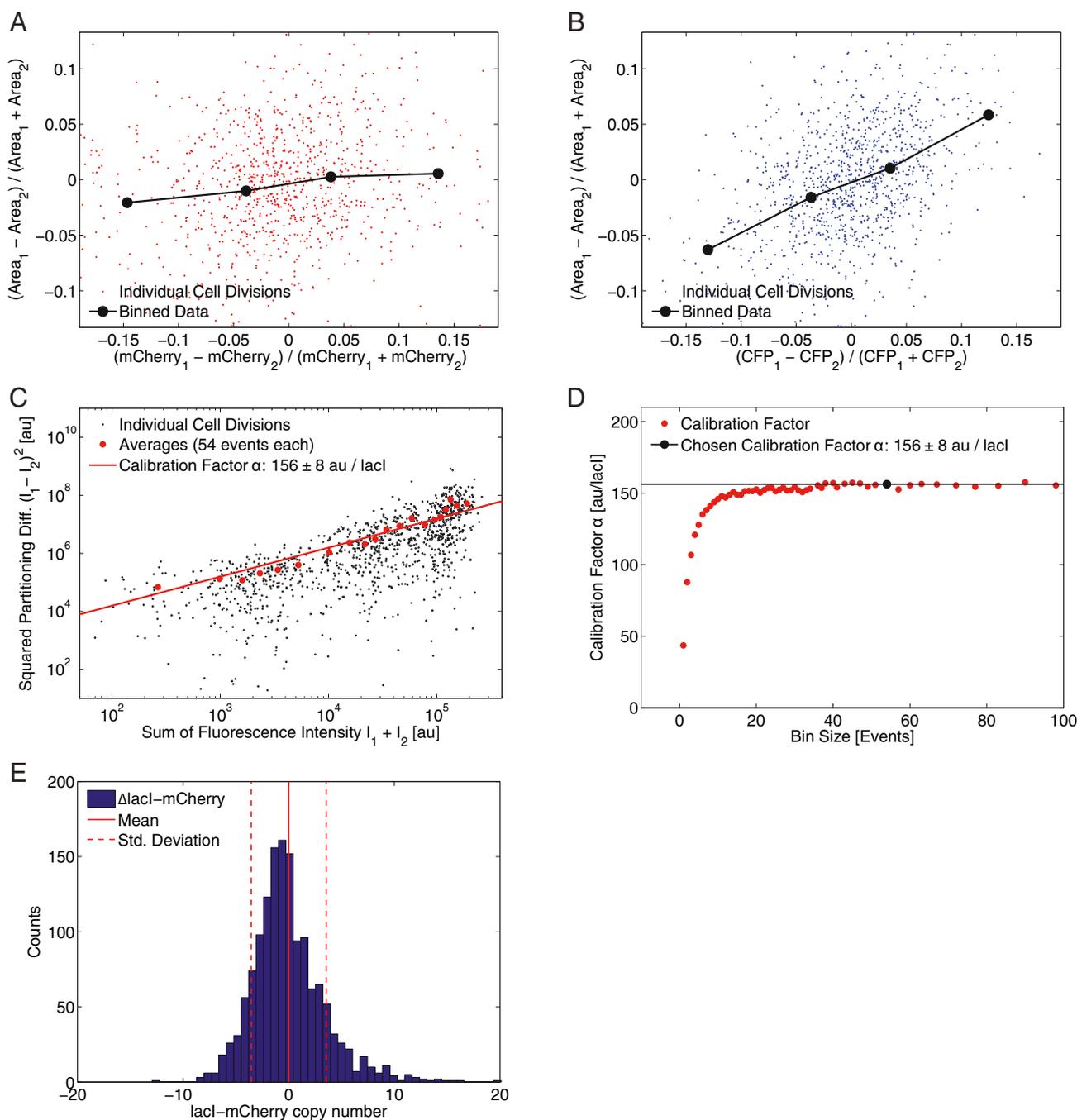
Figure S3. Determination of the Calibration Factor of LacI, Related to Figure 3

(A and B) Fraction of area partitioned between a daughter pair on division versus fraction of fluorescence signal partitioned between a pair of daughters for (A) LacI-mCherry and (B) CFP. The partitioning of the cytoplasmic CFP volume marker (B) is strongly influenced by area partitioning differences (more fluorescence partitions to larger cells with more volume). However, the LacI-mCherry (A) is only weakly influenced by volume area partitioning. This is expected since LacI tends to be DNA bound and the chromosomal DNA is partitioned in approximately equal measure independent of differences in cell volume of the daughters.

(C) Calibration of fluorescence of LacI-mCherry molecules. The square of the error in fluorescence partitioning between two daughter cells is plotted as a function of the fluorescence of the mother cell for a representative data set. Each black point represents a specific division event. These points are binned resulting in the red, averaged data points which are fitted to equation S4 in order to obtain the calibration factor α relating mCherry fluorescence to the absolute number of LacI-mCherry molecules inside the cell.

(D) Sensitivity of the calibration factor to data binning. The determined calibration factor as a function of the number of points in each total fluorescence bin. For sufficient averaging (bin sizes $\approx$ 15 or greater), the calibration factor is relatively insensitive to the binning.

(E) Histogram of mCherry measurement of the auto fluorescence strain in units of number of LacI-mCherry. This histogram depicts the inherent limit of detection for LacI-mCherry proteins calculated in an example experiment. The histogram is the mCherry fluorescence of a collection of cells with 0 LacI-mCherry but due to autofluorescence fluctuations are typically measured to have non-zero fluorescence. We use the width of this distribution to set the limit of where we can distinguish signal from autofluorescence fluctuations. This is calculated for every experiment and we ignore points in our data which have less than one standard deviation above 0 fluorescence.
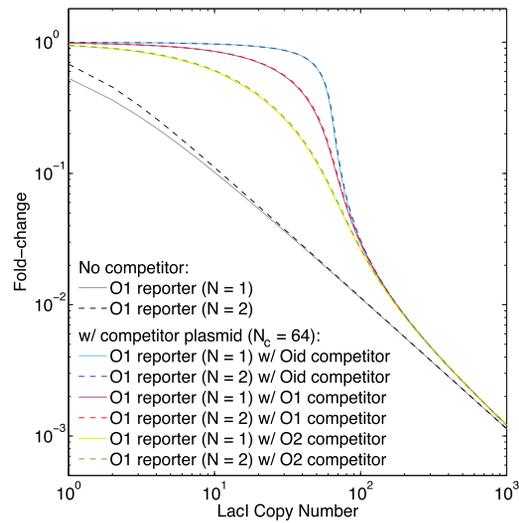
Figure S4. Comparison of Fold-Change Predictions for Multiple Chromosomal Gene Copies, Related to Figures 3 and 6

The fold-change predicted from 1 or 2 chromosomal copies with an O1 operator site either without a competing plasmid (gray solid and black dashed line) or competing with a high copy number plasmid $N_c = 64$ and bearing an O1 site (purple solid and red dashed line), O2 site (yellow solid and green dashed line), or Oid site (light blue solid and dark blue dashed line). In this case, simply going from 1 (solid lines) to 2 (dashed lines) copies of the chromosome does not change the expected fold-change significantly.
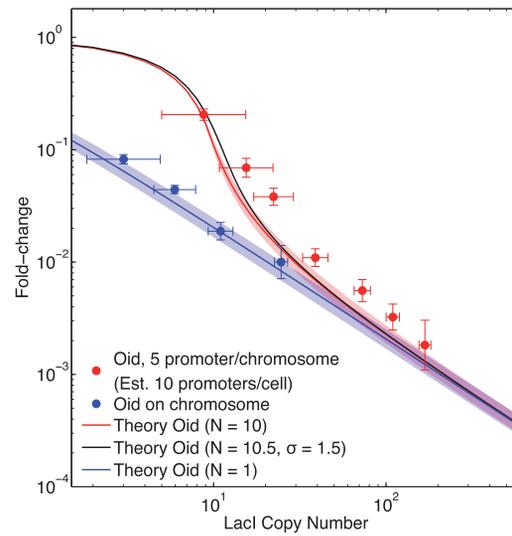
Figure S5. Effect of Copy Number Distribution on Multiple Chromosomal Integration Theory, Related to Figure 4B

Fold-change versus LacI copy number for the multiple integration strain (red points) overlayed with theoretical predictions from the thermodynamic model using a fixed value for the promoter copy number $N = 10$ (red line), as used in the main text, and using a normal distribution for promoter copy number centered around 10.5 with a standard deviation of 1.5 (black line). This choice of distribution is only meant to approximate the effect of chromosomal copy number variation within our expected copy number range of $9 - 12$ and demonstrate the relatively small effect of the distribution for this case. For reference, the theory predictions and data for the single integrated copy case are shown as blue. Error bars in fold-change measurements are SEM. Error bars in LacI copy number are the quadrature summed errors from the calibration factor and the inherent resolution limit of LacI detection.
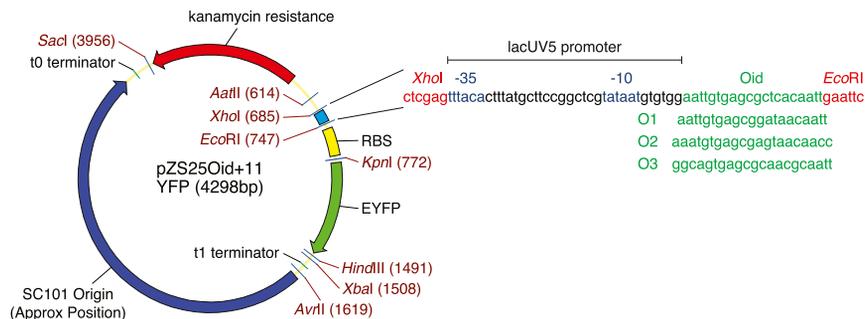
Figure S6. Plasmid Diagram and Promoter Sequence, Related to Figure 1

The main features of the plasmids pZS25O1+11-YFP are shown flanked by unique restriction sites. The particular promoter sequence based on the *lac*UV5 promoter is shown together with the sequences of the different Lac repressor binding sites used. This plasmid was used as a basis for creating the plasmids and chromosomal integration reporters for this study.