SUPPLEMENTARY INFORMATION

for

GENERAL PROPERTIES OF TRANSCRIPTIONAL TIME-SERIES IN

ESCHERICHIA COLI

Lok-hang So, Anandamohan Ghosh, Chenghang Zong, Leonardo A. Sepúlveda,

Ronen Segev, Ido Golding
# Table of Contents

## Supplementary Figures

| Figure S1. Comparing the measured burstiness $B$ and the Fano factor $b$. | 4 |
| Figure S2. Expression levels of the ribosomal promoter $rrnBP1$. | 5 |
| Figure S3. Expression levels of the $P_{mar}$ promoter. | 6 |
| Figure S4. Expression levels of the $P_{bioBFCD}$ promoter. | 7 |
| Figure S5. Expression levels of the $P_{RM}$ promoter. | 8 |
| Figure S6. Effect of gene copy-number on the transcriptional time-series. | 9 |
| Figure S7. Estimating the average gene copy-number in a population. | 10 |
| Figure S8. Effect of differences in mRNA lifetime on the degree of burstiness. | 11 |
| Figure S9. mRNA lifetime measurements. | 12 |
| Figure S10. Positive covariance between the activities of multiple gene copies in the cell. | 13 |
| Figure S11. Fitting the modulation schemes to smFISH data from individual promoters. | 15 |
| Figure S12. Fitting the modulation schemes to subsets of smFISH data separated by generation time. | 17 |
| Figure S13. $k_{off}$-modulation fails to describe the behavior of randomly generated time-series. | 19 |
| Figure S14. Comparing mRNA distributions from experiments and theory. | 20 |
| Figure S15. Comparing individual promoter data to the universal theoretical fit. | 21 |
| Figure S16. Mutual information calculated for three different promoters. | 22 |
| Figure S17. Similar behavior of mutual information for different promoters. | 23 |
| Figure S18. Simultaneous modulation of all three kinetic parameters. | 24 |

## Supplementary Tables

| Table S1. Sequences of smFISH probes. | 26 |
| Table S2. Bacterial strains used in this work. | 29 |
| Table S3. Phage strains used in this work. | 30 |
| Table S4. Characteristics of promoters used in this work. | 31 |
| Table S5. mRNA lifetimes of transcripts used in this work. | 35 |
| Table S6. Sequences of qRT-PCR primers. | 36 |

## Supplementary Note

| Growth media and conditions | 37 |
| Strains | 39 |
| Single-molecule fluorescence *in situ* hybridization (smFISH) | 40 |
Data analysis 43
Comparison of smFISH data to other assays 46
Quantitative reverse transcription polymerase chain reaction (qRT-PCR) 48
Measurement of gene activity in individual living cells 50
Modeling transcription kinetics 52
Mutual information calculation 56

References 58
Figure S1. Comparing the measured burstiness $B$ and the Fano factor $b$.

The burstiness of a simulated transcriptional time-series $B = \frac{\sigma_r - m_r}{\sigma_r + m_r}$ is plotted against the Fano factor $b = \frac{\sigma^2}{\langle n \rangle}$. Here $\sigma_r$ and $m_r$ are the standard deviation and mean of the inter-event times in the simulated time series, respectively, while $\sigma$ and $\langle n \rangle$ are the standard deviation and mean of mRNA per cell in the population. $B$ was calculated directly from the simulation data, while $b$ was calculated from the kinetic parameters using analytical expressions. Each data point is based on a specific combination of transcription kinetics parameters, and about 48,000 distinct combinations are represented in this plot. The monotonically increasing relationship demonstrates that $b$ is an appropriate measure of burstiness. The red curve is a fit to $y = \frac{A(x-1)^c}{B + (x-1)^c}$, giving $A = 0.95$, $B = 2.4$, and $C = 0.81$. $R^2 = 0.97$. 

Supplementary Figures
Figure S2. Expression levels of the ribosomal promoter *rrnBP1*.

Shown is a comparison between smFISH data and values from the literature. Expression levels as a function of the growth rate, obtained from smFISH experiments, were compared to the known numbers of ribosomes per cell \(^3\). In making the comparison, a correction factor was included, to compensate for the difference in lifetime between endogenous ribosomal RNA (which is stable) and the *lacZ* reporter used \(^4\). Red error bars represent standard deviation of data obtained from two independent repeats of smFISH experiments. The x-axis is the generation time normalized by the minimum generation time achievable by the strain studied.
Figure S3. Expression levels of the \( P_{marH} \) promoter.

Shown is a comparison between smFISH data and values obtained using a fluorescent reporter. Expression levels as a function of sodium salicylate concentration, obtained from smFISH experiments, were compared to YFP data (unpublished data from Lon Chubiz, Christopher Rao Lab, University of Illinois at Urbana-Champaign). Red error bars represent standard deviation of data obtained from two independent repeats of smFISH experiments.
Figure S4. Expression levels of the $P_{bioBFCD}$ promoter.

Shown is a comparison between smFISH data and values from the literature. Expression levels as a function of biotin concentration, obtained from smFISH experiments, were compared to numbers obtained from $\beta$-galactosidase (LacZ) activity assay given in $^5$. Error bars represent standard deviation of data obtained from two independent repeats of smFISH experiments.
Figure S5. Expression levels of the $P_{RM}$ promoter.

Shown is a comparison between smFISH data and values from the literature. Expression levels of lysogens carrying the c/D38N mutation and various $P_{RM}$ mutations in host strain JL2497 (RecA$^+$), obtained from smFISH experiments, were plotted against numbers obtained from β-galactosidase (LacZ) activity assay given in $^6$. Vertical error bars represent standard deviation of data obtained from two independent repeats of smFISH experiments. Grey line is a linear fit. Data for the NP6 mutant lies outside the plotted range.
Figure S6. Effect of gene copy-number on the transcriptional time-series.

Shown is the effect of gene copy number $X$ on mRNA number $n$ (A), transcriptional burst size $b$ (B), and burst frequency $r$ (C). Markers are values obtained from simulation. Each color represents a set of $k_{on}$, $k_{off}$, and $k_{TX}$ values, with one to six copies of the gene. Lines are guides to the eye, demonstrating that $n$ and $r$ scale linearly with gene copy number, while $b$ remains constant.
Figure S7. Estimating the average gene copy-number in a population.

To estimate the average gene copy number in a population \( \langle X \rangle \), the formula \( \langle X \rangle = \frac{C}{\tau} + \frac{D}{\tau} \) was used, where \( C \) and \( D \) are the durations of the \( C \)- and \( D \)-periods (time needed for replicating the genome and time between termination of replication and cell division, respectively) in minutes, \( \tau \) is the cell doubling time in minutes, and \( m' \) is the normalized distance of the gene from oriC. \( C \) and \( D \) as a function of \( \tau \) were taken from \(^3\). Values calculated for \( m' \) ranging from 0 to 1 are shown.
Figure S8. Effect of differences in mRNA lifetime on the degree of burstiness.

We calculated the burstiness $b$ for a range of mRNA lifetimes and transcription parameters corresponding to the data in this study. Specifically, mRNA lifetimes cover the range found in Table S5; the values of $k_{\text{on}}$, $k_{\text{off}}$, and $k_{\text{TX}}$ range from 10-fold below to 10-fold above typical values in our data set (see Figure 3). $b$ was calculated from analytical expressions (see Eq. (1) in Parameter Estimation under the section Modeling transcription kinetics) and was normalized by the value obtained for lacZ mRNA lifetime (Table S5). $b$ does not deviate by more than ~20% from the value for lacZ.
Figure S9. mRNA lifetime measurements.

The plots depict the relative mRNA level, measured using qRT-PCR, as a function of time after adding rifampicin. Each panel corresponds to a specific transcript, under specific growth conditions \((\tau_g = \text{cell doubling time})\); see TABLE S5 for details. Two independent repeats were performed in each case (red triangles and blue circles), as well as no-rifampicin controls (grey triangles and circles). The red and blue lines are fits to \(\ln(y) = Ax + B\), and mRNA lifetime was calculated from \(\tau_{RNA} = -1/A\). The values shown are mean ± standard deviation from the two independent repeats. See RNA lifetime measurement under the section Quantitative reverse transcription polymerase chain reaction (qRT-PCR) for details.
Figure S10. Positive covariance between the activities of multiple gene copies in the cell.

(A) mRNA statistics of cells having different gene copy-numbers can be obtained by examining cells at the beginning and end of the life cycle, as indicated by the cell length. Here, the mean number of mRNA per cell of the 20% longest cells $\langle n \rangle_{\text{long}}$ is plotted against that of the 20% shortest cells $\langle n \rangle_{\text{short}}$. Data (+) is from 115 independent experiments with different promoters and expression levels. The points fall on the $y = 2x$ line (grey line), consistent with the expectation that cells at the end of the life cycle will have twice the gene copies of newly divided cells \(^3\). The inset shows a histogram of $\frac{\langle n \rangle_{\text{long}}}{\langle n \rangle_{\text{short}}}$, with a mean 2.0 and standard error 0.04 (shaded region).
(B) The variance of mRNA number per cell of the 20% longest cells $\sigma_{\text{long}}^2$ is plotted against that of the 20% shortest cells $\sigma_{\text{short}}^2$. Fitting to $y = Ax$ gives $A = 3.1 \pm 0.1$ (grey line). This value is higher than the value of 2, which would be expected if the activities of individual gene copies were independent. In other words, we find a positive covariance between identical copies of the same gene. A positive covariance suggests that the two copies tend to turn “on” and “off” together. The inset shows a histogram of the normalized covariance, defined as $\frac{\sigma_{\text{long}}^2 - 2\sigma_{\text{short}}^2}{2\langle n \rangle_{\text{short}}^2}$.

The normalized covariance has a mean 0.4 and standard error 0.1 (shaded region).
Figure S11. Fitting the modulation schemes to smFISH data from individual promoters.

(A) Number of smFISH data points obtained from individual promoters.

(B) $R^2$ of fits obtained from fitting the three modulation schemes to smFISH data from individual promoters. $R^2$ of fitting the $k_{\text{off}}$ modulation scheme is consistently higher than $R^2$ for the other two schemes. The actual fits of the $k_{\text{off}}$ modulation scheme are shown in (E).
(C) $k_{on}$ values obtained from fitting $k_{off}$ modulation scheme to smFISH data from individual promoters. Black line is the value obtained from a fit to all data points, with 95% confidence interval shown in green shading. The actual fits of the $k_{off}$ modulation scheme are shown in (E).

(D) $k_{TX}$ values obtained from fitting $k_{off}$ modulation scheme to smFISH data from individual promoters. Black line is the value obtained from a fit to all data points, with 95% confidence interval shown in green shading. The actual fits of the $k_{off}$ modulation scheme are shown in (E).

(E) Fits of the $k_{off}$ modulation scheme to smFISH data from individual promoters. Black line is a fit to all data points.
Figure S12. Fitting the modulation schemes to subsets of smFISH data separated by generation time.

(A) Number of smFISH data points in each subset.
(B) $R^2$ of fits from fitting the three modulation schemes to subsets of smFISH data separated by generation time. $R^2$ of fitting the $k_{\text{off}}$ modulation scheme is consistently higher than $R^2$ for the other two schemes. The actual fits of the $k_{\text{off}}$ modulation scheme are shown in (E).

(C) $k_{\text{on}}$ values obtained from fitting $k_{\text{off}}$ modulation scheme to subsets of smFISH data separated by generation time. There is an obvious trend of $k_{\text{on}}$ increasing with generation time. Grey line is a linear fit. Black line is the value obtained from a fit to all data points, with 95% confidence interval shown in green shading. The actual fits of the $k_{\text{off}}$ modulation scheme are shown in (E).

(D) $k_{\text{TX}}$ values obtained from fitting $k_{\text{off}}$ modulation scheme to subsets of smFISH data separated by generation time. There is an obvious trend of $k_{\text{TX}}$ increasing with generation time. Grey line is a linear fit. Black line is the value obtained from a fit to all data points, with 95% confidence interval shown in green shading. The actual fits of the $k_{\text{off}}$ modulation scheme are shown in (E).

(E) Fits of the $k_{\text{off}}$ modulation scheme to subsets of smFISH data separated by generation time. Black line is a fit to all data points.
Figure S13. $k_{\text{off}}$-modulation fails to describe the behavior of randomly generated time-series.

(A) Grey markers are $\sigma^2/\langle n \rangle$ and $\langle n \rangle$ calculated from randomly generated $k_{\text{on}}$, $k_{\text{off}}$, and $k_{\text{TX}}$ using analytic expressions $^2, ^9, ^{10}$. $k_{\text{on}}$, $k_{\text{off}}$, and $k_{\text{TX}}$ were chosen randomly from ranges spanning 4 orders of magnitude, and data with $\langle n \rangle \geq 100$ were discarded. The black curve is a fit of the $k_{\text{off}}$ modulation curve to this data. $R^2$ obtained from this fit is 0.08, significantly lower than that obtained from fitting to smFISH data ($R^2 = 0.81$, see FIGURE 3A).

(B) Histogram of $R^2$ values obtained from fits of $k_{\text{off}}$ modulation curve to 10,000 sets of data from randomly generated $k_{\text{on}}$, $k_{\text{off}}$, and $k_{\text{TX}}$. Each set has the same number of points as in our smFISH data. The mean $R^2$ is 0.14, significantly lower than that obtained from fitting to smFISH data ($R^2 = 0.81$, see FIGURE 3A).
Figure S14. Comparing mRNA distributions from experiments and theory.

Distributions of mRNA per cell (red markers, lower panels) from eight smFISH experiments (marked I to VIII in top panel) were compared to the theoretically predicted distributions (black lines, lower panels)\(^2,9,10\). To take into account the effects of gene copy number \(X\), the theoretical distributions were calculated using scaled transcription parameters \(k_{\text{on}} \rightarrow k_{\text{on}} \times \langle X \rangle\), \(k_{\text{off}} \rightarrow k_{\text{off}} \times \langle X \rangle\), and \(k_{TX} \rightarrow k_{TX} \times \langle X \rangle\), where \(\langle X \rangle\) is the estimated average gene copy-number of the population.
Figure S15. Comparing individual promoter data to the universal theoretical fit.

(A) The ratio of experimental to theoretical burstiness for individual promoters. \( b_{\text{exp}} \) is the burstiness estimated from smFISH data. \( b_{\text{theory}} \) is the burstiness predicted from the theoretical, gene-independent fit in FIGURE 3A (at the same expression level). Grey crosses denote individual experiments. Black crosses mark mean ± standard deviation of the data from each promoter. The dashed horizontal line and horizontal grey patch denote the mean ± standard deviation for all the data sets pooled together. The average deviation of an individual promoter from the universal theoretical fit is ~33%.

(B) The correlation coefficient between the experimental and theoretical values of the burstiness. Data sets are as in Panel A. The dashed horizontal line marks correlation coefficient corresponding to all data sets pooled together. It can be seen that 6 out of 7 promoters exhibit a correlation coefficient >0.85 with the universal theoretical fit.
Figure S16. Mutual information calculated for three different promoters.

The mutual information between input and output, as a function of $\kappa$ and $\alpha$, was calculated for (A) $P_{\text{lac}}$ with IPTG concentration as input, (B) $P_{\text{bioBFCD}}$ with biotin concentration as input, and (C) $P_{\text{marII}}$ with sodium salicylate concentration as input. Calculations were based on the phenomenological form $\sigma^2/\langle n \rangle = 1 + \langle n \rangle^\alpha/\kappa$. The three mutual information maps have almost identical topographies with regards to the position of the “ridge” in $(\kappa, \alpha)$ plane.
Figure S17. Similar behavior of mutual information for different promoters.

The response of promoter activity to external stimulus can be schematically described by the red curve (e.g. $P_{lac}$) or the blue curve (e.g. $P_{bioBFCD}$) in the central panel. Given an input distribution $P(c)$ we can calculate the corresponding distribution of mRNAs by $P(n) = P(c) \left[ \frac{\partial f(c)}{\partial c} \right]^{-1}$. Although the blue response curve is a decreasing function of $c$ while the red response curve is an increasing function of $c$, $P(n)$ is insensitive to this difference due to the absolute value of the derivative. When evaluating the mutual information $I(m,n)$, n’s were randomly chosen from the distribution $P(n)$. The scaling rule $\sigma^2/\langle n \rangle = 1 + \langle n \rangle^\alpha \kappa$ was then applied and $P(m|n)$ was obtained. $I_{opt}(\kappa,\gamma)$ being promoter independent follows from the invariance of $P(n)$ with respect to the choice of promoter.
Figure S18. Simultaneous modulation of all three kinetic parameters.

(A) The burstiness parameter $b = \frac{\sigma^2}{\langle n \rangle}$ as a function of the mean expression level $\langle n \rangle$. Data (grey markers) is reproduced from FIGURE 3A. The data set is approximated by a Hill function (black line). Each point $(\sigma^2, \langle n \rangle)$ along the Hill curve (colored markers) defines a curve in $(k_{on}, k_{off}, k_{TX})$ space (panel B, curves of the same color). (B) Ensemble of trajectories in $k$-space.
yielding the observed \( (\sigma^2, \langle n \rangle) \) behavior. Each colored curve corresponds to a marker of the same color in panel (A). Black lines are trajectories traversing the colored lines, such that the trajectory is always perpendicular to a colored line when they intersect. Each black trajectory arises from a choice of an initial \( k_{\text{TX}} \) ranging from 0.003 s\(^{-1}\) to 10 s\(^{-1}\). (C) \( k_{\text{on}} \), \( k_{\text{off}} \), and \( k_{\text{TX}} \) as a function of \( \langle n \rangle \) given by each of the black trajectories in panel (B). For more details see \textbf{Modeling transcription kinetics} in the \textit{SUPPLEMENTARY NOTE}. 
# Supplementary Tables

Table S1. Sequences of smFISH probes.

| Transcript | Probe sequences (5’ to 3’) |
|------------|----------------------------|
| lacZ       | GTGAATCCGTAATCATGGTC       |
|            | TCACGACGTGTGAAAAACGAC      |
|            | ATTAAGTTGGTAAACGCGAGG      |
|            | ATTATCGGACTGGCGGAGGAT      |
|            | AAACCCAGCAAGCGCCCATT       |
|            | AGTATCGCGCTCGAGGAAGAT      |
|            | AACCCTGACATCTGGCAATG       |
|            | TAGTTCCACTGGTTGGTATTG      |
|            | GCTATTTTGTTGAGTCTGGTT      |
|            | TAAAGCCGAGTGCGAACATG       |
|            | AACTGGTACCCGTAGGATG        |
|            | ATATACTCACCGCGAAAAGG       |
|            | TTTCGCACCTGTACAGTGTAG      |
|            | ATAGAGTCTCGGAGTTCTGG       |
|            | TCTGGCTTCAATCCGCTGCG       |
|            | ACCATTTTCAAATCCCGACCT      |
|            | TAAACGCCCTGAAATCAGCA       |
|            | ATGCAGGAGTGAGTGCTGCTG      |
|            | TCTGCTACATGCTGACCTGA       |
|            | TTCACAGCAGATATCTCTGG       |
|            | CACGGGCTTTAAGATTGGTCT      |
|            | TGTTCCGATAATGCGAACA        |
|            | TTCATCCACCATATACAGGC       |
|            | TGGCCTGAGTTTCAATATTG       |
|            | ATCCGCTACGGGATTCTGGT       |
|            | TGTACAGACTGCGGATGTTA       |
|            | ATACGGCCGTTCTGGTAATG       |
|            | GATCGCACAGTATTGGACAG       |
|            | AAATAATATCGGTGGCGTGG       |
|            | TTTGATGGACCATTTCCGCA       |
|            | TATTCGCGAAAAGATACAGGG      |
|            | AAGACTGTTAACCAGTCCGCT      |
|            | TGCCAGTATTGCTGGAAGAC       |
|            | AAAACGCGGATACGTACGAGA      |
|            | TAATCAGCGACTGATCCACC       |
|            | GGTTGCCGTTTTTCACTATA       |
|            | TCGCGCTATCGCCAAAATCTA      |
|            | TCCATACAGAATCGGCAATC       |
|            | TGGTTCTGTTGACCTGCTGAG      |
|            | ACCGGACTGGAATACTGGTCT      |
|            | TATTCGCCTGCTCACTTCTG       |
|            | GTTATCGCTATGACGGAACA       |
|            | TTTACCTTGTGGAAGGCAAT       |
|            | GTTACGCCGATTTGCAATCACA     |
|            | TGGACTACGCTGAAACACCTCT     |
|            | AGCTCACGACGCTATCAGATG      |
|            | ACCCCACTGCAGTGCAAAAAT      |
|            | CGTGTAAAATAGCGAAACGCTTT    |
|            | CTGTGAAAGAAAGCGCTGACT      |
GGCGTCACGGCTGTTTTTT
TACGGCAATGCCTTATCCCA
TAAGGGTGTTCTCCTGTAGCT
ATCAATCCGTAGTTTTCCTCC
GTAATGCCCAATGGCGACAGTC
ATGTCGGACAAATGGCAGATG
ATCAATCCGTATGCTTCTCTAT
AATCCCCATATGGGAAACCGT
AGACCAACCTGTAATGGTGA

cl
GGTTTCTTTTCTTGCTCAT
CTCAAGCTGCTTGTGTTTA
AATTTGCTTTAAGGGCGACGTG
GGAGTAAAGGCAAGTTCTATT
ATCTTGTCTGCCAGACTGTC
AATAAAAGGACCAAGGCGCTG
GCGATAAGCGTAGCTTCGTTAT
CTTACAGCTAAACTTTTGGAG
CTTGCCGATTTGAGGGCTTAA
CTGTTCATACACATCTGTAG
TAAGTGACGCGCTGCACTACTA
ACAGGCTACTGATCACTACTA
CCCTGCTGAAACTGAGAAA
TTTACAGCTACAGGAGGGAAC
ATGTGCATCACTCCTGTGTA
GTGCTCAATCCGTCACCTG
AAGTGCGCAATGATGGCTG
AACTCATACCCCCAAGCT
CCTGATGACCTTTCTGGAAGG
GTAAAACACCTGACGCTA
TTGCGTACGTTGGTTTATG
CAACCTCTATTGCAAGGGAT
AGGGATAACTTTCCCCCACA
AAAGCGCTCTTTACAAGGCT
ATGTGTTACGCTATTGCGC
GGAGTGGAAAAATTCCCCA
CGTGTAAGCGCAATACAAA
GAGCGCTTATCTTCTCCCTTT
GCCAGCAGAGAATTAAGGGAA
CAACCTGAGCGGATGATTA
CTGAAAAGGACTTTGGCTAT
CAACTGAGGCTTATAGGCGA
GGGAGGCTGTTCTTAATATC
GGAATCCCAATGATCTGCA
GTGTCGCTCCTAACAAAAAC
CCGAAAAACCTTCTGCGAAA
AATACCTATGCCATAGGAC
CCTAGCTTTAATGACGATC
ATCAGTGGCTCTATTGCAAC
AACAATGCTTTTGGTGGTTT
TCGCGCGATGAAATGCTAT
TAGCTTGGCTCTACCTCTCA

cro
AGGGTTATGCGTTGTTCCAT
AAAGGCAATGGCATATCTT
CTTACGTCGTTGCTTGGTG
GCGCTTTGATATACGCCGAG
TGCATGAAATGGCGTGTGTA
TAGTTAAAATACCTTCTCG
TAAAACGCTTCATCAGCGT
GAAGGGCTTTACCCTTCCG
CTGGTTGTTTTTTGTACTC
GTCACGAGTGTCAAACGA
ATTAAGTGGTATTACGCAAG
TACGTACCAGCTCTGCGAAA
AGGTTCAGTTGGTGTAGATG
AATGTGAGCGAGTAACCC
GAACGCGATACCGGGAACCT
ATCTTCCAGATAACTGCGCT
AACGAGACGTACCGGAAAAT
GCTGATTTGTGTAGTCCGTT
TTAAACGGATATGCAACATG
AAACGTTTACCCTTAGGAGT
ATAATTTCACCAGCGGAAAAG
TTCCAGTTTACCAGTGATGT
ATAGAGATTCGGGATTTTCGG
ACCAATTTCTAACTCCGACCT
TTAACGCTCTCAATTCAGCAAA
TTTCATACGCAAGGATATCCTG
CAGGCGTTTTAAAGGTTTCT
GGTTCGGATAATGCGAACA
TTGGCTTACTTTACCACGATA
AGACGGATATGGCACCACAT
TGATCAGACCTCGGCGTATTGA
ATTGGCATTCAGCGAATACAGC
AAATATATTGGGCTGGCGTGG
TTTTGTTTGGATTTTCGGCA
ATTAGCGAGCCGACCCAGAG
AAACGCGGATACGCGGAAA
TTAACAGCGACTGATCCAC
ATACAGAAACTGGCCATCGTT
AAACTGCTGTGGTGTGTTG
TTGACGTTCTGTTACTTCCGAT
GTTACGCTATTACGGGAACA
TTAACGCTTTGGGGATTAGC
GTTCAACCGCTATTACCCA
Table S2. Bacterial strains used in this work.

| Strain   | Relevant genotype | Promoter | Reporter | Reference | Source              |
|----------|-------------------|----------|----------|-----------|---------------------|
| BW14894  | ΔlacIZYA          | -        | -        | 11        | CGSC #8280          |
| MG1655   | Wild-type, λ-     | -        | -        | 12        | Lab stock           |
| JL5902   | RecA-             | -        | -        | 12        | John Little         |
| JL2497   | RecA+             | -        | -        | 12        | John Little         |
| TK310    | ΔcyA ΔcpdA ΔlacY  | P_{lac}  | Endogenous lacZ mRNA | 13        | Terrence Hwa        |
| JT235S   | ΔlacIZY galE::lacZ F^+ gal^+ | P_{galETKM} | lacZ     | 14        | Sankar Adhya        |
| LC544    | ΔlacIZY P_{marII}lacZ | P_{marII} | lacZ     |           | Christopher Rao     |
| LC543    | ΔlacIZY           | -        | -        |           | Christopher Rao     |
| CF7753   | ΔlacZ(mluI) rrnBP1':lacZ-kan | rrnBP1 | lacZ     | 13,16     | Michael Cashel      |
| CY481    | ΔlacIZY bioF-lacZ | P_{bioBFCD} | lacZ     | 5,17      | John Cronan         |
| NC416    | luc-N''...cro+ cII'-lacIZYA | P_{RM}, P_{R} | cI, cro | 18        | Donald Court        |
Table S3. Phage strains used in this work.

| Strain   | Relevant genotype | Reference | Source     |
|----------|-------------------|-----------|------------|
| $\lambda_{IG831}$       | Wild-type         |           | Lab stock  |
| $\lambda_{IG2504}$       | cI T88C           | 19        | Lab stock  |
| $\lambda_{IG28061}$      | cI V36I           | 20        | Lab stock  |
| $\lambda_{IG28062}$       | cI L18V           | 20        | Lab stock  |
| $\lambda_{JL815}$         | cI D38N           |           | John Little |
| $\lambda_{NP2}$           | cI D38N           | 6         | John Little |
|                       | $P_{RM}$ -35 TAGA $\rightarrow$ GCTG   |
|                       | $P_{RM}$ -10 GATT $\rightarrow$ TATT    |
| $\lambda_{NP3}$           | cI D38N           | 6         | John Little |
|                       | $P_{RM}$ -35 TAGA $\rightarrow$ CATT   |
|                       | $P_{RM}$ -10 GATT $\rightarrow$ GAAT    |
| $\lambda_{NP4}$           | cI D38N           | 6         | John Little |
|                       | $P_{RM}$ -35 TAGA $\rightarrow$ CCTT   |
|                       | $P_{RM}$ -10 GATT $\rightarrow$ CCAT    |
| $\lambda_{NP5}$           | cI D38N           | 6         | John Little |
|                       | $P_{RM}$ -35 TAGA $\rightarrow$ CTAA   |
|                       | $P_{RM}$ -10 GATT $\rightarrow$ GAAT    |
| $\lambda_{NP6}$           | cI D38N           | 6         | John Little |
|                       | $P_{RM}$ -35 TAGA $\rightarrow$ CCCA   |
|                       | $P_{RM}$ -10 GATT $\rightarrow$ TGAT    |
| $\lambda_{NP7}$           | cI D38N           | 6         | John Little |
|                       | $P_{RM}$ -35 TAGA $\rightarrow$ TACC   |
|                       | $P_{RM}$ -10 GATT $\rightarrow$ TACT    |
| $\lambda_{NP8}$           | cI D38N           | 6         | John Little |
|                       | $P_{RM}$ -35 TAGA $\rightarrow$ GTGT   |
|                       | $P_{RM}$ -10 GATT $\rightarrow$ GTAT    |
| $\lambda_{NP10}$          | cI D38N           | 6         | John Little |
|                       | $P_{RM}$ -35 TAGA $\rightarrow$ CCAA   |
|                       | $P_{RM}$ -10 GATT $\rightarrow$ GAAT    |
| $\lambda_{NP11}$          | cI D38N           | 6         | John Little |
|                       | $P_{RM}$ -35 TAGA $\rightarrow$ CTCA   |
|                       | $P_{RM}$ -10 GATT $\rightarrow$ TGCT    |
Table S4. Characteristics of promoters used in this work.

| Promoter | Expression level (mRNA/cell)† | Molecular mode(s) of regulation‡ | Network motif | References |
|----------|------------------------------|---------------------------------|---------------|------------|
| $P_{lac}$ | $\sim 0.01 – 60$ | Activated by CRP$^a$ Repressed by LacI$^b$ | ![Network motif](lac.png) | Simple regulation with double negative feedback | $21,22$ |
| $P_{galETKM}$ | $\sim 0.1 – 10$ | Activated by CRP$^a$ Repressed by GalR$^c$ and GalS$^d$ | ![Network motif](gal.png) | Incoherent feed-forward loop | $22,23$ |
| $P_{marII}$ | $\sim 2 – 10$ | Activated by MarA$^e$, Rob$^f$ and SoxS$^g$ Repressed by MarR$^h$ | ![Network motif](mar.png) | Simple regulation | $24$ |
| Protein | Half-life (min) | Regulation | Notes |
|---------|----------------|------------|-------|
| *rrnBP1* | ~1 – 20 | Constitutive | Not fully characterized |
| **P_{bioBFCD}** | ~0.1 – 10 | Repressed by BirA | |
| **P_{RM}** | ~6 | Autoregulated by Cl | |
| **P_{RM} (cJV368)** | ~6 | Covalent dimerization of Cl | |
| **P_{RM} (c/K88C)** | ~4 | Misfolding of Cl | |
| **P_{RM} (c/L18V)** | ~9 | Misfolding of Cl | |
| **P_{RM} (c/D38N)** | ~2 | No positive autoregulation by Cl, reduced P_{RM} activity | |
| **P_{RM} (c/D38N)** | ~7 | No positive autoregulation by Cl, reduced P_{RM} activity | |
| **P_{RM} (c/D38N)** | ~1 | No positive autoregulation by Cl, reduced P_{RM} activity | |
| **P_{RM} (c/D38N)** | ~2 | No positive | |

**Diagram:**

Simple regulation with negative feedback

Two-node double negative feedback loop with single node positive autoregulation

Two-node double negative feedback loop
| Promoter | Description | Regulatory Elements | CRP: cAMP (3′-5′-cyclic adenosine monophosphate) receptor protein, activated in the presence of cAMP. | LacI: lac repressor, inactivated in the presence of lactose or lactose analogues. | GalR: gal repressor, inactivated in the presence of D-galactose or D-fucose. | GalS: gal isorepressor, inactivated in the presence of D-galactose or D-fucose. | MarA: mar activator. | Two-node double negative feedback loop with single node positive autoregulation |
|-----------|-------------|---------------------|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|
| CCTT P_{Ram}→CCAT | autoregulation by CI, reduced P_{RM} activity | P_{RM} (cI:IDSN P_{Ram}→35 TAGA→CTAA P_{Ram}→10 GATT→GAAT) | ~2 | No positive autoregulation by CI, reduced P_{RM} activity | 6 |
| P_{RM} (cI:IDSN P_{Ram}→35 TAGA→CCCA P_{Ram}→10 GATT→TGAT) | ~2 | No positive autoregulation by CI, reduced P_{RM} activity | 6 |
| P_{RM} (cI:IDSN P_{Ram}→35 TAGA→TACC P_{Ram}→10 GATT→TACT) | ~2 | No positive autoregulation by CI, reduced P_{RM} activity | 6 |
| P_{RM} (cI:IDSN P_{Ram}→35 TAGA→GTGT P_{Ram}→10 GATT→GTAT) | ~2 | No positive autoregulation by CI, reduced P_{RM} activity | 6 |
| P_{RM} (cI:IDSN P_{Ram}→35 TAGA→CCAA P_{Ram}→10 GATT→GAAT) | ~2 | No positive autoregulation by CI, reduced P_{RM} activity | 6 |
| P_{RM} (cI:IDSN P_{Ram}→35 TAGA→CTCA P_{Ram}→10 GATT→TGCT) | ~3 | No positive autoregulation by CI, reduced P_{RM} activity | 6 |
| P_{Ram} | ~0.1 – 6 | Autoregulated by Cro\(^k\) Repressed by CI\(^j\) | 26 |

\(^\dagger\) From this study.
\(^\ddagger\) All promoters are also regulated by nucleoid-associated proteins such as Fis and H-NS.

\(^a\) CRP: cAMP (3′-5′-cyclic adenosine monophosphate) receptor protein, activated in the presence of cAMP.

\(^b\) LacI: lac repressor, inactivated in the presence of lactose or lactose analogues.

\(^c\) GalR: gal repressor, inactivated in the presence of D-galactose or D-fucose.

\(^d\) GalS: gal isorepressor, inactivated in the presence of D-galactose or D-fucose.

\(^e\) MarA: mar activator.
Rob: Rob transcriptional activator.
SoxS: transcriptional dual regulator of superoxide response regulon.
MarR: mar repressor, inactivated in the presence of sodium salicylate and a variety of antibiotics and toxic chemicals.
BirA: biotin transcriptional repressor, activated in the presence of biotin, sequestered by unbiotinylated AccB.
Cl: λ repressor, activates $P_{RM}$ at lower concentrations and represses $P_{RM}$ at higher concentrations.
Cro: λ “anti-repressor”, represses $P_{RM}$ and $P_R$.
AccB: biotin carboxyl carrier protein.
Table S5. mRNA lifetimes of transcripts used in this work.

| #  | Strain     | Promoter | Transcript† | Expression level (mRNA/cell) | Doubling time (min) | RNA lifetime (min)‡ | Source     |
|----|------------|----------|-------------|-----------------------------|---------------------|---------------------|------------|
| E1 | TK310      | P_{lac}  | lacZ        | ~4.9                        | ~33                 | 2.2 ± 0.1           | This work  |
| E2 | TK310      | P_{lac}  | lacZ        | ~55                         | ~44                 | 2.020 ± 0.003       | This work  |
| E3 | TK310      | P_{lac}  | lacZ        | ~0.59                       | ~63                 | 1.9 ± 0.2           | This work  |
| E4 | LC544      | P_{novel} | lacZ        | ~4.1                        | ~40                 | 2.1 ± 0.1           | This work  |
| E5 | LC544      | P_{novel} | lacZ        | ~6.4                        | ~90                 | 2.4 ± 0.2           | This work  |
| E6 | CY481      | P_{novelCD} | lacZ      | ~1.4                        | ~34                 | 1.4 ± 0.1           | This work  |
| E7 | CF7753     | rrnBP1   | lacZ*       | ~22                         | ~28                 | 0.97 ± 0.05         | This work  |
| E8 | CF7753     | rrnBP1   | lacZ*       | ~1.9                        | ~64                 | 0.91 ± 0.02         | This work  |
| E9 | NC416      | P_{RM}   | cl          | ~10                         | ~34                 | 2.8 ± 0.8           |           |
| E10| JL5902 (λNP2) | P_{RM} (c/fD38N) | cl | ~6.9 | ~34 | 1.12 ± 0.05 | This work |
| E11| NC416      | P_{R}    | cro         | ~1.1                        | ~34                 | 1.9 ± 0.5           |           |

† lacZ* has an inefficient ribosome binding site to avoid toxicity from excessive β-galactosidase levels\textsuperscript{15,16} (Michael Cashel, personal communication).
‡ Mean ± standard deviation from 2 independent experiments.

Comments:

1. E1 through E3; E4 and E5; and E7 and E8:
Comparing mRNA lifetimes for the same transcript at different growth rates.

2. E1 through E3; E4 and E5; and E7 and E8:
Comparing mRNA lifetimes for the same transcript at different expression levels.

3. E1, E4, and E6:
Comparing mRNA lifetimes for the same transcript expressed from different promoters.

4. E1, E7, E9, and E11:
Comparing mRNA lifetimes of different transcripts.

5. E9 and E10:
Comparing mRNA lifetimes for the same transcript in RecA\textsuperscript{+} and RecA\textsuperscript{−} backgrounds.
Table S6. Sequences of qRT-PCR primers.

| Transcript | Primer sequences (5’ to 3’) |
|------------|-----------------------------|
| *lacZ*     | ACGACATTGGCGTAAGTGAAGCG     |
|            | CGTGAGCGGTCGTAATCAGC        |
| *cl*       | CAACAGCCTGCTCAGGGTCAAC      |
|            | GGTGATGCGAGAGATGGGTAAGC      |
Supplementary Note

Growth media and conditions

All strains were grown in M9 minimal media with thiamine and casamino acids, and glucose as a carbon source (M9CAgluc, Teknova, #M8010) unless otherwise stated. All strains were grown at 37 °C with shaking unless otherwise stated.

To achieve different expression levels from \( P_{\text{luc}} \), strain TK310 (gift of Terrence Hwa, University of California, San Diego) was grown with 0 to 1 mM of isopropyl \( \beta\)-D-1-thiogalactopyranoside (IPTG, Sigma, #I6758) and 0 to 10 mM of adenosine 3’,5’-cyclic monophosphate (cAMP) (Sigma, #A9501).

To achieve different expression levels from \( P_{\text{galETKM}} \), strain JT235S (gift of Sankar Adhya, National Institutes of Health) was grown with 10 mM cAMP and 0 to 30 mM of D-fucose (Sigma, #F8150).

To achieve different expression levels from \( P_{\text{marH}} \), strain LC544 (gift of Christopher Rao, University of Illinois, Urbana) was grown with 0 to 10 mM of sodium salicylate (Sigma, #S3007).

To achieve different expression levels from \( rrrnBP1 \), strain CF7753 (gift of Michael Cashel, National Institutes of Health) was grown in minimal media with different carbon sources, which resulted in different growth rates. Specifically, M9 minimal media (with thiamine) with glucose, glycerol, succinate, and acetate as carbon source were used. These are prepared following standard molecular biology protocols 28,29.

To achieve different expression levels from \( P_{\text{bioBFCD}} \), strain CY481 (gift of John Cronan, University of Illinois, Urbana) was grown in M9 minimal media with thiamine, vitamin-free casamino acids (Difco, #228830), and glucose, prepared following standard molecular biology protocols 28,29, with 0.0016 to 1.6 μM of biotin (Sigma, #B4501). To get rid of traces of biotin from the LB plate on which the cells grew, cells from overnight cultures were washed with media with the same conditions as the experimental conditions 3 times before diluted into fresh media.

To achieve different expression levels from \( P_{R\text{M}} \), lysogens of various \( P_{R\text{M}} \) and/or \( cI \) mutants were used (\( P_{R\text{M}} \) mutants are gift of John Little, University of Arizona). When strain JL5902 (RecA\(^-\)) (gift of John Little, University of Arizona) was used as host, cells were grown in LBGM (LB 28 with 1 mM MgSO\(_4\) and 0.2% glucose). When strain JL2497 (RecA\(^+\)) (gift of John Little, University of Arizona) was used as host, cells were grown in M9CAgluc.
To achieve different expression levels from $P_{RM}$ and $P_R$ with the temperature sensitive allele $cI857$, reporter strain NC416 was used (gift of Donald Court, National Institutes of Health). Cells were grown in LB at 30, 32, 34, 36, 38, and 40 °C.
Strains

1. Bacterial strains

Bacterial strains used are listed in TABLE S2. All of them are *E. coli* K-12 derivatives. BW14894 was used as a negative control for smFISH experiments using *lacZ* probes. MG1655 was used as a negative control for smFISH experiments using *cI* and *cro* probes. MG1655, JL5902, and JL2497 were used as hosts for bacteriophage λ in smFISH experiments using *cI* probes to study the \( P_{RM} \) promoter.

2. Phage strains

Phage strains used are listed in TABLE S3.

3. Promoters

TABLE S4 provides additional details of the promoters used in this study.
Single-molecule fluorescence in situ hybridization (smFISH)

The procedures are based on 30. Modifications were made to adapt the protocol to E. coli 27. Sterile, nuclease-free, aerosol-barrier pipette tips were used in all procedures. Nuclease-free reagents were used whenever available. Diethylpyrocarbonate (DEPC)-treated water (Ambion, #AM9922) was used wherever the protocol calls for water.

1. Probe design and labeling

DNA oligonucleotide probes were designed using the online program developed by Arjun Raj (van Oudenaarden Lab, Massachusetts Institute of Technology) (http://singlemoleculefish.com/designer.html). Sequences of smFISH probes used are listed in TABLE S1. DNA oligos with 3’-end amine modification were ordered from Biosearch Technologies. The 96-well cartridge in which the oligos were delivered was spun briefly to get all liquid to the bottom of the wells. The content of each well (100 μl) was then transferred to a microcentrifuge tube and stored at -20 °C.

To label the probes with fluorescent dyes, 10 μl of each DNA oligo was first pooled into a microcentrifuge tube. 1/9 volume of 1 M sodium bicarbonate solution (filter sterilized) (Fisher Scientific, #BP328) was added to the pool to get a final concentration of 0.1 M sodium bicarbonate. 2.5 mg of fluorescent dye (6-TAMRA, SE (Invitrogen, #C6123) for lacZ and cI probes, Cy5 Bis NHS ester (GE Healthcare Life Sciences #PA15000) for cro probes) was dissolved in 5 μl of dimethyl sulfoxide (Fisher Scientific, #BP231), and 50 μl of 0.1 M sodium bicarbonate solution was added to the solution. The DNA oligo solution was then added to the dye solution. The tube was wrapped in aluminum foil and incubated overnight at 37 °C in the dark.

The labeled probes were purified by ethanol precipitation. The solution was transferred to a 50 ml centrifuge tube. 1/9 volume of 3 M sodium acetate solution (pH 5.2, autoclaved) (Fisher Scientific, #BP333) was added to get a final concentration of 0.3 M sodium acetate. 2.5 volumes of 100% ethanol was added and mixed well. The solution was aliquoted to microcentrifuge tubes and incubated at -70 °C for at least 3 hours (up to overnight). It was then centrifuged in a bench-top microcentrifuge at maximum speed for 30 minutes. The supernatant was discarded and any remaining liquid was absorbed with Kimwipe. The pellet was dissolved in 90 μl of water. 10 μl of 3 M sodium acetate solution (pH 5.2) was added to obtain a final concentration of 0.3 M sodium acetate. This was followed by two more rounds of ethanol precipitation. Finally, the pellets were dissolved in a total of 500 μl Tris-EDTA (pH 8.0) (Fisher Scientific, #BP2473). The tube was wrapped in aluminum foil and stored at -20 °C.
2. Sample fixation and permeabilization

An overnight culture was diluted 100- to 1000-fold into 30 ml of growth medium in a baffled flask. The culture was incubated at 37 °C with shaking. When OD_{600} of the culture reached 0.3–0.4, it was cooled in an ice-water bath. A volume with number of cells equivalent to 15 ml of OD_{600} = 0.4 was transferred to an ice-cold 50 ml Corning centrifuge tube and the cells were pelleted by centrifugation (5 minutes, 4500×g, 4 °C). The supernatant was removed and the cells were resuspended in 1 ml freshly prepared 3.7% formaldehyde (Fisher, #BP531) in 1× PBS (diluted from 10× PBS, Ambion, #AM9625). The cells were then transferred to a microcentrifuge tube and mixed on a rotator at room temperature for 30 minutes. The cells were pelleted by centrifugation (8 minutes, 400×g). The supernatant was removed and the cells were washed in 1 ml 1× PBS twice (i.e. resuspended in 1 ml 1× PBS, centrifuged at 600×g for 3.5 minutes, and supernatant removed). The cells were resuspended in 300 μl water, and then 350 μl of 100% ethanol was added and mixed twice to get to a final concentration of 70% ethanol. The cells were left at room temperature with mixing on a rotator for at least 1 hour (or alternatively, at 4 °C for at least a week) to permeabilize the cell membrane.

3. Hybridization

After permeabilization, cells were centrifuged (7 minutes, 600×g) and the supernatant was removed. The cells were resuspended in 1 ml of 40% wash solution (see below) and the tube was left standing for a few minutes. An aliquot of 40% hybridization solution (see below) was warmed to room temperature and 50 μl was added to a microcentrifuge tube. 3 μl of a 10-fold diluted probe solution was added to the hybridization solution and mixed well. The cells were then centrifuged (7 minutes, 600×g) and the supernatant was removed. The cells were resuspended in the hybridization solution with probes and left at 30 °C overnight. Hybridized samples could be stored at 4 °C for at least 6 months.

10 ml of 40% wash solution contains 4 g of formamide (Ambion, #AM9342) and 1 ml of 20× SSC (Ambion, #AM9763).

10 ml of 40% hybridization solution contains 1 g of dextran sulfate (Sigma, #D8906), 4 g of formamide, 10 mg of *E. coli* tRNA (Sigma, #R4251), 1 ml of 20× SSC, 40 μl of 50 mg/ml BSA (Ambion, #AM2616), and 100 μl of 200 mM ribonucleoside vanadyl complex (New England Biolabs, #S1402S). The solution was filter sterilized, aliquoted, and stored at -20 °C.
4. Washing

10 μl of hybridized sample was transferred to a microcentrifuge tube. The rest was stored at 4 °C. 200 μl of 40% wash solution was added to the tube and mixed well. Cells were pelleted by centrifugation (3.5 minutes, 600×g) and the supernatant was removed. The cells were washed 3 more times (i.e. resuspended in 200 μl of 40% wash solution, incubated at 30 °C for 30 minutes, centrifuged at 600×g for 3.5 minutes, and supernatant removed). 4′,6-diamidino-2-phenylindole (DAPI, Fisher Scientific, #PI-46190) was added to the wash solution to a final concentration of 10 μg/ml in the last wash. The cells were resuspended in 10 μl of 2 SSC and imaged under the microscope.

5. Microscopy

1 μl of sample was pipetted onto a 24×50 mm #1 coverslip (Fisher Scientific, #12-545F). A 1 mm thick 1.5% agarose gel pad (in 1× PBS) was laid on the sample. A 22×22 mm #1 coverslip (Fisher Scientific, #12-545B) was placed on top of the agarose gel pad. The sample was imaged using an inverted epifluorescence microscope (Nikon Instruments Eclipse TE2000-E) and a cooled EMCCD camera (Photometrics Cascade 512). A 100× N.A. 1.40 oil immersion phase contrast objective (Nikon Instruments Plan Apo 100× / 1.40 Oil) was used in conjunction with a 2.5× lens in front of the camera. The microscope and camera were controlled using the Metamorph software (Molecular Devices). TexasRed filter set (Nikon Instruments, #96365) was used for imaging mRNA labeled by lacZ and cI smFISH probes, Cy5 filter set (Nikon Instruments, #96324) was used for imaging mRNA labeled by cro smFISH probes, and DAPI filter set (Nikon Instruments, #96310) was used to image DNA stained by DAPI. z-stacks with 9 slices and 250 nm spacing were acquired for phase contrast and TexasRed images. Each sample was imaged at multiple locations to get a total of at least 1000 cells.
Data analysis

All image processing and data analysis were performed using MATLAB (MathWorks).

1. Cell recognition

Cell recognition was performed on phase contrast images of cells using the Schnitzcell MATLAB module (gift of Michael Elowitz, California Institute of Technology). The program applies edge detection and other morphological operations, using the MATLAB Image Processing Toolbox. The output was checked and corrected using the manual interface offered by Schnitzcell. The output consisted of label matrices representing areas occupied by cells, which were then used for further analysis. From these matrices, quantities such as length, position, and area of cells were readily extracted.

2. Spot recognition

A spot recognition program developed in our lab was used to automatically identify and quantify localized fluorescence signals. A Gaussian filter was first applied to smooth out noise, and spots were recognized by the presence of a local maximum in both $x$- and $y$-directions. This was done at each $z$-position in the stack of images, and each spot was quantified at the $z$-position where it had the highest fluorescence intensity (where the spot is in focus).

3. Estimating mRNA numbers

A fluorescence spot could consist of multiple mRNAs in close proximity. The integrated intensity arising from a single mRNA needed to be estimated for each smFISH experiment so that fluorescence intensities could be normalized to give the absolute number of mRNAs. The typical intensity of “false positives” in an experiment was first estimated from the histogram of individual spot intensities of a negative control. Histograms of individual spot intensities from relatively low expression samples were then examined. Because most spots in these samples were expected to contain a single mRNA, the first peak that emerged above the false positive range in each of these histograms served as an estimate for the intensity of a single mRNA. The mean intensity of the first peaks from multiple such histograms was taken as the single mRNA intensity for that particular experiment. The sum of intensities of all spots in each cell was then normalized to give absolute number of mRNAs.
4. Estimating transcription kinetics parameters

In the approximation that mRNA production occurs in short rapid bursts, transcription kinetics is characterized by two parameters, the burst size $b$ and burst frequency $r$ \cite{9,10}. These parameters were estimated in two different ways: First, the histograms of mRNA copy number were fitted to a negative binomial distribution using the MATLAB Curve Fitting Toolbox, and $b$ and $r$ were calculated from the fitting parameters \cite{9,10}. Alternatively, $b$ and $r$ were also estimated using the relations $b = \sigma^2 / \langle n \rangle$ and $(b - 1)r = \langle n \rangle$, where $\sigma^2$ and $\langle n \rangle$ are the variance and mean of mRNA copy number. The two methods gave values of $b$ and $r$ that are in good agreement (data not shown).

5. Correcting for differences in gene copy number

In growing bacterial cells, the copy number of a given gene can vary in the range ~1–8, depending on growth conditions and cell age \cite{3}. Since we were interested in the kinetics of mRNA production from a single gene copy, we had to calibrate our data appropriately. To quantify the effect of gene copy number on transcription kinetics, we carried out Gillespie simulations of the two-state transcription model \cite{9,10}, with varying gene copy number. We found that, as the gene copy number $X$ changes, the burst size $b$ stays approximately constant (within 20% variation relative to value at one copy of the gene), while the burst frequency $r$ and the mean mRNA number $n$ scale linearly with $X$ (see FIGURE S6). Thus, in analyzing smFISH data we scaled mRNA numbers accordingly: $b \rightarrow b^X$, $r \rightarrow r/\langle X \rangle$, and $\langle n \rangle \rightarrow \langle n \rangle/\langle X \rangle$, where $\langle X \rangle$ is the estimated average gene copy number in the cell population (see below).

6. Estimating average gene copy number

To estimate the average gene copy number in a population $\langle X \rangle$, the formula $\langle X \rangle = 2^{\frac{C \cdot m'}{D} \cdot \tau}$ was used, where $C$ and $D$ are the durations of the $C$- and $D$-periods (time needed for replicating the genome and time between termination of replication and cell division, respectively \cite{3}) in minutes, $\tau$ is the cell doubling time in minutes, and $m'$ is the normalized distance of the gene from oriC \cite{7}. $C$ and $D$ as a function of $\tau$ were taken from \cite{3}, and $m'$ is known for the genetic constructs studied. The resulting $\langle X \rangle$ as a function of $\tau$ is shown in FIGURE S7.
7. Correcting for differences in mRNA lifetime

Different transcripts are subject to different degradation rates\(^{31}\). The mRNA degradation rate in turn affects the measured expression level. We wanted our analysis to reflect variations in mRNA production parameters alone. To achieve that, we corrected for differences in mRNA lifetimes, as follows: The mRNA degradation rates for the different transcripts used in this study were measured using quantitative RT-PCR following inhibition of transcription using rifampicin (see RNA lifetime measurement under the section Quantitative reverse transcription polymerase chain reaction (qRT-PCR) for details). The majority of promoters examined in our study (\(P_{\text{lac}}\), \(P_{\text{galETKM}}\), \(P_{\text{marII}}\), and \(P_{\text{bioBFCD}}\); see TABLES S2 and S5) expressed the full \(\text{lacZ}\) gene as a reporter. Thus, although the parameters of mRNA production varied largely between promoters, the actual transcripts produced were all identical, thus subject to similar post-transcriptional regulation, including mRNA degradation\(^{31}\). To verify that mRNA degradation rate was determined by the transcript rather than the promoter, we measured the mRNA lifetime of \(\text{lacZ}\) expressed from three promoters (\(P_{\text{lac}}\), \(P_{\text{marII}}\), and \(P_{\text{bioBFCD}}\)). The values found showed a standard deviation of \(~23\%\), significantly less than the differences in mRNA lifetime between different transcripts (see TABLE S5). We also verified that differences in growth rate did not have a strong effect on mRNA lifetime (\(~9\%\) standard deviation when comparing growth at doubling times of 40 minutes versus 90 minutes).

In addition to \(\text{lacZ}\), some strains used in this study expressed other transcripts (27 strains expressing variants of \(\text{cl}\) and \(\text{cro}\) genes; 1 strain expressing a modified version of the \(\text{lacZ}\) transcript in which the ribosome binding site has been mutated; see TABLES S2, S3, and S5). We measured the mRNA lifetime for these reporters as well (TABLE S5 and FIGURE S9). As expected, different transcripts exhibited different degradation rates, with more than a 3-fold range of mRNA lifetimes (from \(~0.9\) minutes to \(~2.8\) minutes; see TABLE S5).

Once the mRNA lifetimes (and hence the degradation rates \(k_d\)) were known, we were able to correct for the differences in \(k_d\) between different data sets, such that our examination of mRNA kinetics will only reflect differences on the production side. This normalization (calibration) was done by multiplying the mean expression level \(\langle n \rangle\) by the degradation rate \(k_d\) (relative to the value for \(\text{lacZ}\), chosen as an arbitrary standard). This calibration procedure allowed the comparison of data sets with different \(k_d\)’s. It is justified by the scaling properties of the two-state model, in which the mean level is proportional to the mRNA lifetime, while the burstiness exhibits only a weak dependence on \(k_d\) (see Eq. (1) and Eq. (2) in Parameter Estimation under the section Modeling transcription kinetics). We verified numerically that the normalization procedure above is justified for the range of parameters observed in our experiments (see FIGURE S8).
Comparison of smFISH data to other assays

For a subset of the promoters used in this study, we compared the mRNA levels measured using smFISH to numbers obtained using other well-established methods. We were thus able to establish the accuracy and dynamic range of the smFISH method.

1. $P_{\text{lac}}$

mRNA levels (molecules per cell) as a function of IPTG concentration, obtained from smFISH experiments, were compared to the expression levels measured using qRT-PCR as well as data from the literature (using $\beta$-galactosidase activity assay,\textsuperscript{13}). As seen in FIGURE 2 in the Main Text, the results of all three assays agree very well.

2. $rrnBP1$

mRNA levels (molecules per cell) as a function of the growth rate, obtained from smFISH experiments, were compared to the known numbers of ribosomes per cell.\textsuperscript{3} In making the comparison, a correction factor was included, to compensate for the difference in lifetime between endogenous ribosomal RNA (which is stable) and the $\text{lacZ}$ reporter used.\textsuperscript{4} As seen in FIGURE S2, the agreement between the values was very good.

3. $P_{\text{marII}}$

mRNA levels (molecules per cell) as a function of sodium salicylate concentration, obtained from smFISH experiments, were compared to the data obtained using a fluorescence reporter (unpublished data from Lon Chubiz and Christopher Rao, University of Illinois). As seen in FIGURE S3, the results of the two assays agree very well.

4. $P_{\text{bioBFCD}}$

mRNA levels (molecules per cell) as a function of biotin concentration, obtained from smFISH experiments, were compared to expression level given in literature.\textsuperscript{5} As seen in FIGURE S4, the results of the two assays agree very well.
5. $P_{RM}$

mRNA levels (molecules per cell) of lysogens carrying the c/D38N mutation and various $P_{RM}$ mutations in host strain JL2497 (RecA$^+$), obtained from smFISH experiments, were compared to expression level given in literature $^6$. As seen in FIGURE S5, the results of the two data sets agree.
Quantitative reverse transcription polymerase chain reaction (qRT-PCR)

Sterile, nuclease-free, aerosol-barrier pipette tips were used in all procedures. Nuclease-free reagents were used whenever available. DEPC-treated water (Ambion, #AM9922) was used wherever the protocol calls for water.

1. Total RNA extraction

Cells were grown in the same conditions as in smFISH experiments. When the OD$_{600}$ of the culture reached 0.3–0.4, a volume of culture with number of cells equivalent to 2 ml of OD$_{600}$ = 0.3 was added to 2 volumes of RNAprotect Bacteria Reagent (Qiagen, #76506). Total RNA was then extracted using the RNeasy Mini Kit (Qiagen, #74104) with the optional DNase I digestion step (RNase-free DNase Set, Qiagen, #79254). The eluate from the final step was passed through the spin column one additional time to increase total RNA concentration. 2 μl of each total RNA sample diluted 10-fold in 10 mM Tris·Cl, pH 7.5 (Fisher Scientific, #BP1757) was scanned on a Nanodrop 2000 (Thermo Scientific). The purity and concentration of the total RNA samples were estimated from the absorbance spectrum between 220 and 320 nm. Part of each total RNA sample was used as template in qRT-PCR reactions, and the rest was stored at -80 °C.

2. qRT-PCR experiment

All total RNA samples were diluted to 50 μg/ml. HotStart-IT SYBR Green One-Step qRT-PCR Master Mix Kit (USB, #75770) was used. For each reaction, 10 μl of HotStart-IT SYBR Green Master Mix (2×), 0.16μl of MMLV -RT, 0.16 μl of RNase inhibitor, 1 μl of 10 μM forward primer, 1 μl of 10 μM reverse primer, 2 μl of 50 μg/ml total RNA sample, and 5.68 μl of RNase-free water were mixed on ice. Sequences of primers used are listed in TABLE S6. Primers were ordered from Integrated DNA Technologies. White 8-tube strips (Bio-Rad, #TLS0851) and flat optical 8-cap strips (Bio-Rad, #TCS0803) were used. All reactions were carried out in duplicates. A dilution series of the sample with the highest expected expression level was included in the reactions for calibration. Negative controls with no reverse transcriptase and/or RNA template were also included in the reactions to ensure the reactions were free of contaminants. Thermal cycles and fluorescence measurements were carried out in a MiniOpticon Real-Time PCR System (Bio-Rad). Details of thermal cycles are as follows: 50 °C for 10 minutes; 95 °C for 2 minutes; and then 40 cycles of 95 °C for 15 seconds, 60 °C for 30 seconds, 72 °C for 40 seconds, followed by fluorescence measurement; and finally a melting curve from 90 °C through 70 °C at 0.3 °C intervals and 1 second dwell time. After the protocol was completed, melting curves were visually inspected to ensure no unexpected products were present.
3. qRT-PCR data analysis

The amplification curves were smoothed using a moving-average filter and then interpolated. The fluorescence threshold was chosen at a level where the amplification appeared exponential for every curve. The Ct value for each reaction is the cycle number at which the fluorescence equals the threshold. Ct values of the dilution series were plotted against the logarithm (base 2) of the dilution ratios. The plot was fitted to a linear curve, and the efficiency $E$ of the PCR reactions was calculated as $E = 2^{-1/slope}$. The coefficient of determination $R^2$ of the linear fit was verified to be at least 0.99, and $E$ was verified to be very close to 2. Finally, the relative expression levels of the samples were estimated as $E^{-Ct}$. 

4. RNA lifetime measurement

RNA lifetime measurements were carried out following the method of \textsuperscript{31}. Cells were diluted from overnight cultures and grown in a 25 ml volume, under the same growth conditions as in smFISH experiments. When OD$_{600}$ reached $\sim$0.3, the culture was separated into two halves ($\sim$12 ml + 12 ml), each half being transferred to a new culture flask, and grown at the same conditions for 10-15 minutes in a water-bath shaker. 1.5 ml of culture was then extracted from each flask and mixed with 3 ml of Qiagen RNAProtect Bacterial reagent to stabilize cellular RNA (see Total RNA extraction above for details). These samples were treated as $t = 0$ samples. Rifampicin (Fisher Scientific, #BP26795) was then added to one of the remaining cultures, at a final concentration of 500 $\mu$g/ml, to inhibit transcription. The culture without rifampicin served as a control. 1.5 ml of culture was extracted from each flask at regular intervals after addition of rifampicin (e.g. at $t = 2$, 4, 6, 8 minutes) and mixed with 3 ml of Qiagen RNAProtect Bacterial reagent to stabilize cellular RNA. Subsequent total RNA extraction and qRT-PCR were performed as described previously in this section. Target RNA level of the rifampicin-treated sample as a function of time was fitted to an exponential, $\ln y = Ax + B$, and RNA lifetime (in minutes) estimated as $\tau_{RNA} = -1/ A$. 

49
Measurement of gene activity in individual living cells

Protocols and genetic constructs were based on 32-34.

1. Strain and plasmid

The strain used was CZ071, harboring the two plasmids constructed by Golding et al. 32, pIG-K133(2cTG) and pIG-BAC(Plac/ara-mRFP1-96bs). CZ071 is wild type MG1655, into which the Z1 cassette 35 was integrated. pIG-K133(2cTG) is a ColE1-based plasmid, with an anhydrotetracycline (aTc) inducible promoter controlling expression of MS2d-GFP. pIG-BAC(Plac/ara-mRFP1-96bs) is an F-based plasmid, with an IPTG and arabinose inducible promoter controlling expression of mRFP1 followed by an array of 96 MS2d binding sites.

2. Time-lapse microscopy

Cells were grown overnight in LB with the required antibiotics at 37 °C with shaking, and diluted 1000-fold into LB with antibiotics and various amounts of IPTG and/or arabinose. Cells were grown at 37 °C with shaking until OD$_{600}$ reached about 0.2, and aTc was added to a final concentration of 10 ng/ml. The cells were then grown at 27 °C for one hour. 1 μl of cell culture was pipetted onto a 24×50 mm #1 coverslip. A 1 mm thick 1.5% agarose gel pad (in LB with 10 ng/ml aTc, and the same concentrations of IPTG and arabinose as in the culture) was laid on the sample. A 22×22 mm #1 coverslip was placed on top of the agarose gel pad. A time-lapse movie of the growing cells was recorded using an inverted epifluorescence microscope (Nikon Instruments Eclipse TE2000-E) and a cooled EMCCD camera (Photometrics Cascade 512). A 100× N.A. 1.40 oil immersion phase contrast objective (Nikon Instruments Plan Apo 100× / 1.40 Oil) was used in conjunction with a 2.5× lens in front of the camera. The microscope and camera were controlled using the Metamorph software (Molecular Devices). TexasRed filter set (Nikon Instruments, #96365) was used for imaging mRFP1. GFP filter set (Nikon Instruments, #96362) was used to image MS2d-GFP. 9 z-stacks at 250 nm spacing were obtained for phase contrast images. The cells were imaged at 2-minute intervals.

3. Image processing and data analysis

Images from time-lapse movies were processed and analyzed using MATLAB programs. Cell recognition was performed on phase contrast images of cells using the Schnitzcell MATLAB module as described in the Data analysis section. Cell lineages were reconstructed using a cell tracking program developed in our lab. RNA signals (localized MS2d-GFP fluorescence) at each
time point were detected and quantified using a spot recognition program developed in our lab as described in the **Data analysis** section. RNA time series of each cell lineage was analyzed by a step detection program developed in our lab, based on an algorithm from Sheyum Syed (Paul Selvin Lab, University of Illinois). This program identifies periods of constant RNA levels and “jumps” that connect these levels, from noisy experimental data using $t$-test $p$-values. Quantities such as gene “on” and “off” times, average mRNA per cell, and transcription burst sizes were then extracted.

### 4. Estimating transcription kinetics parameters

The measurables from time-lapse movies were: transcription burst size $B$ (in arbitrary fluorescence units), gene “on” time $\tau_{on}$ (in minutes), and gene “off” time $\tau_{off}$ (in minutes). To extract $k_{TX}$, the relationship $B \propto \tau_{on} k_{TX}$ was used, which could be written as $k_{TX} \propto B/\tau_{on}$. To extract $k_{TX}$, we used the relationship $k_{on} \propto 1/\tau_{off}$. To extract $k_{off}$, we noted that $k_{TX}$ was constant, and used the relationship $k_{off} \propto 1/\tau_{on} \propto 1/B$. 
Modeling transcription kinetics

1. Gillespie simulation of the two-state transcription model

The Gillespie algorithm 36 was used to simulate the stochastic kinetics of the two-state model 9,10 (FIGURE 1, Main Text), which can give rise to a “bursty” transcriptional time-series. In this model, each copy of a gene can either be in an active (“on”) or inactive (“off”) state. It switches from “on” state to “off” state with rate $k_{\text{off}}$, and from “off” state to “on” state with rate $k_{\text{on}}$. When it is “on”, mRNA is produced at a rate $k_{\text{TX}}$. mRNA is degraded at a rate $k_d$. Probability per unit time of each reaction occurring was calculated from these reaction rates, and the reaction trajectory was simulated stochastically. Cell division and gene replication were incorporated as optional components in the simulations. Reactant species segregate according to binomial statistics upon cell division. Gene copy number doubles at a time specified by the user. To mimic smFISH experiments, a random time point was chosen from each of 1000 trajectories at which number of mRNA is “measured”.

2. Parameter estimation

The analytical expressions for variance and mean of mRNA numbers, arising from the two-state model 2, are:

\[
\sigma^2 = \frac{k_{\text{on}}k_{\text{TX}}}{(k_{\text{on}} + k_{\text{off}})k_d} + \frac{k_{\text{on}}k_{\text{off}}}{(k_{\text{on}} + k_{\text{off}})^2} \frac{k_{\text{TX}}^2}{k_d(k_{\text{on}} + k_{\text{off}} + k_d)} \tag{1}
\]

and

\[
\langle n \rangle = \frac{k_{\text{on}}k_{\text{TX}}}{(k_{\text{on}} + k_{\text{off}})k_d}. \tag{2}
\]

These expressions were used to write $\sigma^2/\langle n \rangle$ as a function of $\langle n \rangle$. If we assume that $k_{\text{on}}$ is modulated, we get (by eliminating $k_{\text{on}}$ from Eq. (1) and Eq. (2))

\[
\frac{\sigma^2}{\langle n \rangle} = 1 + \frac{(k_{\text{TX}} - k_d\langle n \rangle)^2}{k_{\text{TX}}k_{\text{off}} + k_d(k_{\text{TX}} - k_d\langle n \rangle)}. \tag{3}
\]

Similarly, if we assume that $k_{\text{off}}$ is modulated, we get

\[
\frac{\sigma^2}{\langle n \rangle} = 1 + \frac{k_d\left(k_{\text{TX}} - k_d\langle n \rangle\right)\langle n \rangle}{k_{\text{on}}k_{\text{TX}} + k_d^2\langle n \rangle}. \tag{4}
\]

and if we assume $k_{\text{TX}}$ is modulated, we get
The MATLAB Curve Fitting Toolbox was used to fit these functions to data obtained from smFISH experiments. As described in the main text, the fit assuming $k_{\text{off}}$ modulation was significantly better than the others, thus supporting the assumption $k_{\text{off}}$ is modulated.

3. Simultaneous modulation of all three kinetic parameters

Here we consider a modulation scheme which allows all three kinetic parameters, $k_{\text{on}}$, $k_{\text{off}}$, and $k_{\text{TX}}$ to change. This scenario does not yield a single trajectory in $k$-space. Instead, an ensemble of possible trajectories can be defined. To form this ensemble, we first approximate the experimental $\sigma^2 / \langle n \rangle$ versus $\langle n \rangle$ data using a Hill function (FIGURE S18A, black line). Each point $(\sigma^2, \langle n \rangle)$ defines a curve in $(k_{\text{on}}, k_{\text{off}}, k_{\text{TX}})$ space. Thus, each colored marker in FIGURE S18A corresponds to a curve of the same color in FIGURE S18B. From Eq. (1) and Eq. (2), we can write down the equations of these curves with $k_{\text{TX}}$ as a parameter:

$$
\sigma^2 / \langle n \rangle = 1 + \frac{k_d k_{\text{off}} \langle n \rangle}{k_{\text{on}} (k_{\text{on}} + k_{\text{off}} + k_d)}.
$$

The direction tangent to this line at a given $k_{\text{TX}}$ is given by the derivative

$$
v = \frac{\partial k}{\partial k_{\text{TX}}} = \begin{pmatrix}
\frac{k_d \langle n \rangle}{k_{\text{TX}}} \left( \frac{\langle n \rangle^2}{\sigma^2 - \langle n \rangle} + 1 \right) \\
\frac{k_d^2 \langle n \rangle}{k_{\text{TX}}^2} \left( \frac{\langle n \rangle^2}{\sigma^2 - \langle n \rangle} + 1 \right)
\end{pmatrix}
$$

$$
= \begin{pmatrix}
\frac{\langle n \rangle}{\sigma^2 - \langle n \rangle} - \frac{k_d^2 \langle n \rangle}{k_{\text{TX}}^2} \left( \frac{\langle n \rangle^2}{\sigma^2 - \langle n \rangle} + 1 \right)
\end{pmatrix}.
$$

53
To go from a certain point along \( \mathbf{k} \) given by \((\sigma^2, \langle n \rangle)\) and \( k_{TX} \) to a point along an adjacent line \( \mathbf{k}' \) given by \((\sigma'^2, \langle n' \rangle)\) and \( k'_{TX} \), such that the line joining these two points is perpendicular to \( \mathbf{v} \), we require \((\mathbf{k}' - \mathbf{k}) \cdot \mathbf{v} = 0\). This leads to

\[
A k'_{TX} + B k'_{TX} + C = 0,
\]

where

\[
A = \frac{\langle n' \rangle}{\sigma'^2 - \langle n' \rangle} v_2 + v_3,
\]

\[
B = \frac{k_d \langle n' \rangle^2}{\sigma'^2 - \langle n' \rangle} (v_1 - 2v_2) - k_d v_2 - \mathbf{k} \cdot \mathbf{v},
\]

\[
C = k_d^2 \langle n' \rangle \left( 1 + \frac{\langle n' \rangle^2}{\sigma'^2 - \langle n' \rangle} \right) (v_2 - v_1),
\]

and \( v_1, v_2, \) and \( v_3 \) are the three components of \( \mathbf{v} \). If we start at a certain point along \( \mathbf{k} \) given by \((\sigma^2, \langle n \rangle)\) and \( k_{TX,i} \), and move to the adjacent line sequentially in this manner, we obtain a trajectory in \((k_{on}, k_{off}, k_{TX})\) space (FIGURE S18B, black lines). Such a trajectory gives a modulation scheme in which \( k_{on}, k_{off}, \) and \( k_{TX} \) are all changed, as shown in FIGURE S18C.

4. Estimating protein statistics from mRNA statistics

Experimental mRNA statistics was approximated by the form \( \sigma^2/\langle n \rangle = 1 + \langle n \rangle^\alpha/\kappa \), such that the deviation of \( \sigma^2/\langle n \rangle \) from the Poisson case goes like the mean \( \langle n \rangle \) to the power \( \alpha \) (FIGURE 4A). \( \kappa \) and \( \alpha \) were then scaled to represent the protein species as follows. We define

\[
B_{\text{mRNA}} = \langle n \rangle^{\alpha_{\text{mRNA}}} / \kappa_{\text{mRNA}}
\]

and

\[
B_{\text{protein}} = \langle p \rangle^{\alpha_{\text{protein}}} / \kappa_{\text{protein}},
\]

where the \( B \)’s can be considered estimates of mRNA and protein burst sizes, respectively \(^{32}\), and \( \langle p \rangle \) is the mean number of protein molecules. Note that
\[ B_{\text{protein}} = B_{\text{mRNA}} \times B_{\text{TL}} \]
\[ = \left( \frac{n}{\kappa_{\text{mRNA}}} \right)^{\alpha_{\text{mRNA}}} \times \left( k_{\text{TL}} \times \frac{\ln 2}{k_d} \right)^{\alpha_{\text{mRNA}}} \times \left( \frac{k_{\text{dp}}}{k_d} \right)^{\alpha_{\text{mRNA}}} \]  

(11)

where \( B_{\text{TL}} \) is the average number of proteins translated from one mRNA during its lifetime, \( k_{\text{TL}} \) is the translation rate, and \( \ln 2/k_d \) is the mRNA lifetime. Also, protein number follows

\[ \dot{p} = nk_{\text{TL}} - pk_{\text{dp}}, \]

where the first term represents translation and the second term represents degradation, with \( k_{\text{dp}} \) being the protein degradation rate. At steady state, \( \dot{p} = 0 \), so \( n = pk_{\text{dp}}/k_{\text{TL}} \), or

\[ \langle n \rangle = \langle p \rangle k_{\text{dp}}/k_{\text{TL}}. \]

(12)

Inserting Eq. (12) into Eq. (11), we get

\[ B_{\text{protein}} = \left( \frac{p}{\kappa_{\text{mRNA}}} \right)^{\alpha_{\text{mRNA}}} \times \left( k_{\text{TL}} \times \frac{\ln 2}{k_d} \right)^{1-\alpha_{\text{mRNA}}} \times \left( \frac{k_{\text{dp}}}{k_d} \right)^{\alpha_{\text{mRNA}}} \]  

(13)

Comparing Eq. (10) and Eq. (13), we get

\[ \frac{1}{\kappa_{\text{protein}}} = \frac{1}{\kappa_{\text{mRNA}}} \times \left( k_{\text{TL}} \times \frac{\ln 2}{k_d} \right)^{1-\alpha_{\text{mRNA}}} \times \left( \frac{k_{\text{dp}}}{k_d} \right)^{\alpha_{\text{mRNA}}}, \]

\[ \alpha_{\text{protein}} = \alpha_{\text{mRNA}}. \]

\( B_{\text{TL}} = k_{\text{TL}} \times \ln 2/k_d \) was estimated to be in the range \( \sim5–95 \) \(^{3,37}\). \( k_d \) was estimated to be \( 0.0008–0.008 \) s\(^{-1} \) \(^{3,31}\). \( k_{\text{dp}} \) was estimated to be equal to the cell growth rate, \( 0.0001–0.0004 \) s\(^{-1} \) from our experiments (cell generation time ranging from 30 to 100 minutes). The value at the marker in FIGURE 4B was obtained using the midpoint of the above ranges, and the shaded region around the marker represents the range of \( \kappa \) given by the above ranges and error of \( \alpha \) from the fit in FIGURE 4A.
Mutual information calculation

1. Calculating the mutual information between outside stimulus and gene activity

The information transfer from the input stimulus to the output protein level can be quantified by computing the Shannon mutual information. Shannon showed that the mutual information is independent of the details of the underlying mechanisms and can be interpreted as follows: if information is 1 bit, one can infer that the input is in an on or off state; if the information transmitted is about 3 bits, the output encodes \(2^3 = 8\) distinct states, etc.

However, the general optimization problem of determining the information transfer is hard and can only be solved for simple cases. In one such exact calculation was shown in the small-noise limit. In the problem current work, mRNA numbers follow a negative binomial distribution implying that the variance scales as mean. In this case, the small-noise assumption is not necessarily valid anymore. The negative binomial distribution makes it impossible to have a closed form expression for the mutual information and we have to estimate it numerically.

The mutual information, \(I(m,c)\), will give us an estimate of how much information is transmitted to the output level of mRNAs \((m,\text{ or similarly the protein level, see discussion above)}\) about changes in the input concentration, \(c\). We take the approximation that the input factors are drawn from a Gaussian distribution \(P(c)\) with mean \(\mu\) and variance \(\sigma^2\). From experiments (e.g. FIGURE 2A) we also know the dependence of mean mRNA number on \(c\), i.e. \(n = f(c)\), so by knowing \(P(c)\) we can evaluate the distribution of mean mRNA's as function of the input \(c\) using the formula

\[
P(n) = P(c)\frac{\partial f(c)}{\partial c}^{-1}.
\]

Equivalently we can compute the mutual information between the output expression level \(m\) and the mean \(n\),

\[
I(m,n) = \int dm \int dn \ P(m,n) \log_2 \left[ \frac{P(m,n)}{P(m)P(n)} \right].
\]

The experimental mRNA (and therefore protein) statistics follows the relation \(\sigma^2/\langle n \rangle = 1 + \langle n \rangle^\alpha / \kappa\). We can thus write the probability distribution above in terms of \(n\) alone. It is interpreted as the conditional distribution \(P(m|n)\), the probability of obtaining \(m\) mRNAs having their mean fixed at \(n\).

Using the identity \(P(m,n) = P(m|n)P(n)\) it is straightforward to numerically compute \(I(m,n)\) for varying mean \(\mu\) and variance \(\sigma^2\) of the input distribution and obtain the pair of \(\{\mu, \sigma^2\}\) that
maximizes the mutual information for this specific form of input distribution (Gaussian). The information transfer is the maximum if the input is Gaussian and the channel or the conditional distribution is also Gaussian. For the negative binomial distribution it is not known what input distribution will maximize the mutual information, however the Gaussian input is a good approximation in large $n$ limit.

The maximum of the mutual information thus estimated for different choice of $\{\kappa, \alpha\}$ is shown in FIGURE 4.

2. The promoter-independent behavior of the mutual information

The response of promoter activity to external stimulus can be schematically described as the red curve (e.g. $P_{lac}$) or the blue curve (e.g. $P_{bioBFCD}$) in the central panel of FIGURE S17. Given an input distribution $P(c)$ we can calculate the corresponding distribution of mRNAs using Eq. (14). Since the distribution $P(c)$ is sigmoidal, when mean $c$ is small or large the resulting mRNA distribution is delta function-like. For such signals information transfer will be minimal. The relevant regime is in the middle, where the input promoter distribution is multiplied by the inverse of the gain, which is a constant factor, to yield the mRNA distribution.

Although in FIGURE S17 the blue response curve is a decreasing function of $c$ while the red response curve is an increasing function of $c$, the distribution $P(n)$ is not sensitive to this difference due to the absolute value of the derivative in Eq. (14). However, it should be noted that the mutual information thus computed will be of same order but optimality will occur for a different set of input promoter distribution parameters $\{\mu, \sigma^2\}$.

When evaluating the mutual information $I(m,n)$, $n$’s were randomly chosen from the distribution $P(n)$. The scaling rule $\sigma^2/\langle n \rangle = 1 + \langle n \rangle^\alpha / \kappa$ was then applied and $P(m|n)$ was obtained. $I_{opt}(\kappa, \gamma)$ being promoter independent follows from the invariance of $P(n)$ with respect to the choice of promoter.
References

1. Goh, K.I. & Barabasi, A. Burstiness and memory in complex systems. *Epl* 81, - (2008).
2. Peccoud, J. & Ycart, B. Markovian Modeling of Gene-Product Synthesis. *Theoretical Population Biology* 48, 222-234 (1995).
3. Neidhardt, F.C. *Escherichia coli and Salmonella typhimurium: cellular and molecular biology*, (American Society for Microbiology, Washington, D.C., 1987).
4. Selinger, D.W., Saxena, R.M., Cheung, K.J., Church, G.M. & Rosenow, C. Global RNA half-life analysis in Escherichia coli reveals positional patterns of transcript degradation. *Genome Res* 13, 216-23 (2003).
5. Abdel-Hamid, A.M. & Cronan, J.E. Coordinate expression of the acetyl coenzyme A carboxylase genes, accB and accC, is necessary for normal regulation of biotin synthesis in Escherichia coli. *J Bacteriol* 189, 369-76 (2007).
6. Michalowski, C.B. & Little, J.W. Positive autoregulation of cl is a dispensable feature of the phage lambda gene regulatory circuitry. *J Bacteriol* 187, 6430-42 (2005).
7. Bremer, H. & Churchward, G. An examination of the Cooper-Helmstetter theory of DNA replication in bacteria and its underlying assumptions. *J Theor Biol* 69, 645-54 (1977).
8. Paulsson, J. Summing up the noise in gene networks. *Nature* 427, 415-8 (2004).
9. Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y. & Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol* 4, e309 (2006).
10. Shahrezaei, V. & Swain, P.S. Analytical distributions for stochastic gene expression. *Proc Natl Acad Sci U S A* 105, 17256-61 (2008).
11. Yakovleva, G.M., Kim, S.K. & Wanner, B.L. Phosphate-independent expression of the carbon-phosphorus lyase activity of Escherichia coli. *Appl Microbiol Biotechnol* 49, 573-8 (1998).
12. Little, J.W., Shepley, D.P. & Wert, D.W. Robustness of a gene regulatory circuit. *EMBO J* 18, 4299-307 (1999).
13. Kuhlman, T., Zhang, Z., Saier, M.H., Jr. & Hwa, T. Combinatorial transcriptional control of the lactose operon of Escherichia coli. *Proc Natl Acad Sci U S A* 104, 6043-8 (2007).
14. Tokeson, J.P., Garges, S. & Adhya, S. Further inducibility of a constitutive system: ultrinduction of the gal operon. *J Bacteriol* 173, 2319-27 (1991).
15. Potrykus, K. et al. Antagonistic regulation of Escherichia coli ribosomal RNA rnb P1 promoter activity by GreA and DksA. *J Biol Chem* 281, 15238-48 (2006).
16. Hernandez, V.J. & Bremer, H. Guanosine tetraphosphate (ppGpp) dependence of the growth rate control of rnb P1 promoter activity in Escherichia coli. *J Biol Chem* 265, 11605-14 (1990).
17. Barker, D.F. & Campbell, A.M. Use of bio-lac fusion strains to study regulation of biotin biosynthesis in Escherichia coli. *J Bacteriol* 143, 789-800 (1980).
18. Svenningsen, S.L., Costantino, N., Court, D.L. & Adhya, S. On the role of Cro in lambda prophage induction. *Proc Natl Acad Sci U S A* 102, 4465-9 (2005).
19. Sauer, R.T., Jordan, S.R. & Pabo, C.O. Lambda repressor: a model system for understanding protein-DNA interactions and protein stability. *Adv Protein Chem* 40, 1-61 (1990).
20. Lim, W.A. & Sauer, R.T. Alternative packing arrangements in the hydrophobic core of lambda repressor. *Nature* **339**, 31-6 (1989).
21. Müller-Hill, B. *The lac Operon : a short history of a genetic paradigm*, (Walter de Gruyter, Berlin ; New York, 1996).
22. Alon, U. *An introduction to systems biology : design principles of biological circuits*, xvi, 301 p., [4] p. of plates (Chapman & Hall/CRC, Boca Raton, FL, 2007).
23. Weickert, M.J. & Adhya, S. The galactose regulon of Escherichia coli. *Mol Microbiol* **10**, 245-51 (1993).
24. Alekshun, M.N. & Levy, S.B. The mar regulon: multiple resistance to antibiotics and other toxic chemicals. *Trends Microbiol* **7**, 410-3 (1999).
25. Liang, S. *et al*. Activities of constitutive promoters in Escherichia coli. *J Mol Biol* **292**, 19-37 (1999).
26. Ptashne, M. *A genetic switch : phage lambda revisited*, xiv, 154 p. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2004).
27. Zong, C., So, L.-h., Sepulveda, L.A., Skinner, S.O. & Golding, I. Lysogen stability is determined by the frequency of activity bursts from the fate-determining gene. *Mol Syst Biol* **6**(2010).
28. Ausubel, F.M. *Current protocols in molecular biology*, (John Wiley & Sons, Brooklyn, N. Y., 1987).
29. Sambrook, J. & Russell, D.W. *Molecular cloning : a laboratory manual*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2001).
30. Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A. & Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* **5**, 877-9 (2008).
31. Bernstein, J.A., Khodursky, A.B., Lin, P.H., Lin-Chao, S. & Cohen, S.N. Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. *Proc Natl Acad Sci U S A* **99**, 9697-702 (2002).
32. Golding, I., Paulsson, J., Zawilski, S.M. & Cox, E.C. Real-time kinetics of gene activity in individual bacteria. *Cell* **123**, 1025-36 (2005).
33. Golding, I. & Cox, E.C. Chapter 8: Spatiotemporal dynamics in bacterial cells: real-time studies with single-event resolution. *Methods Cell Biol* **89**, 223-51 (2008).
34. Golding, I. & Cox, E.C. RNA dynamics in live Escherichia coli cells. *Proc Natl Acad Sci U S A* **101**, 11310-5 (2004).
35. Lutz, R. & Bujard, H. 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**, 1203-10 (1997).
36. Gillespie, D.T. Exact Stochastic Simulation of Coupled Chemical-Reactions. *Journal of Physical Chemistry* **81**, 2340-2361 (1977).
37. Cai, L., Friedman, N. & Xie, X.S. Stochastic protein expression in individual cells at the single molecule level. *Nature* **440**, 358-62 (2006).
38. Shannon, C.E. A mathematical theory of communication. *Bell Syst Tech J* **27**, 379-423, 623-656 (1948).
39. Tkacik, G., Callan, C.G., Jr. & Bialek, W. Information flow and optimization in transcriptional regulation. *Proc Natl Acad Sci U S A* **105**, 12265-70 (2008).