RESEARCH ARTICLE

Origins of Transcriptional Transition: Balance between Upstream and Downstream Regulatory Gene Sequences

Adrien Sala,[a] Muhammad Shoaib,[a] Olga Anufrieva,[a] Gnanavel Mutharasu,[a] Olli Yli-Harja,[a,b] Meenakshisundaram Kandhavelu[a]

Molecular Signaling Laboratory, Computational Systems Biology Research Group, Signal Processing Department, Tampere University of Technology, Tampere, Finland[a]; Institute for Systems Biology, Seattle, Washington, USA[b]

ABSTRACT **By measuring individual mRNA production at the single-cell level, we investigated the *lac* promoter's transcriptional transition during cell growth phases. In exponential phase, variation in transition rates generates two mixed phenotypes, low and high numbers of mRNAs, by modulating their burst frequency and sizes. Independent activation of the regulatory-gene sequence does not produce bimodal populations at the mRNA level, but bimodal populations are produced when the regulatory gene is activated coordinately with the upstream and downstream region promoter sequence (URS and DRS, respectively). Time-lapse microscopy of mRNAs for *lac* and a variant *lac* promoter confirm this observation. Activation of the URS/DRS elements of the promoter reveals a counterplay behavior during cell phases. The promoter transition rate coupled with cell phases determines the mRNA and transcriptional noise. We further show that bias in partitioning of RNA does not lead to phenotypic switching. Our results demonstrate that the balance between the URS and the DRS in transcriptional regulation determines population diversity.**

IMPORTANCE **By measuring individual mRNA production at the single-cell level, we investigated the *lac* promoter transcriptional transition during cell growth phases. In exponential phase, variation in transition rate generates two mixed phenotypes producing low and high numbers of mRNAs by modulating the burst frequency and size. Independent activation of the regulatory gene sequence does not produce bimodal populations at the mRNA level, while it does when activated together through the coordination of upstream/downstream promoter sequences (URS/DRS). Time-lapse microscopy of mRNAs for *lac* and a *lac* variant promoter confirm this observation. Activation of the URS/DRS elements of the promoter reveals a counterplay behavior during cell phases. The promoter transition rate coupled with cell phases determines the mRNA and transcriptional noise. We further show that bias in partitioning of RNA does not lead to phenotypic switching. Our results demonstrate that the balance between URS and DRS in transcription regulation is determining the population diversity.**

Received 22 October 2014 **Accepted** 18 December 2014 **Published** 27 January 2015

Citation Sala A, Shoaib M, Anufrieva O, Mutharasu G, Yli-Harja O, Kandhavelu M. 2015. Origins of transcriptional transition: balance between upstream and downstream regulatory gene sequences. mBio 6(1):e02182-14. doi:10.1128/mBio.02182-14.

Editor George L. Drusano, University of Florida

Address correspondence to Meenakshisundaram Kandhavelu, meenakshisundaram.kandhavelu@tut.fi.

Stochasticity in gene expression increases the probability of survival of microorganisms under stress conditions (1). However, it still remains elusive how stochasticity in gene expression is involved in switching of phenotypes during metabolic demand. Individual bacteria have to cope with a fluctuating environment by altering gene expression for their survival. Typically, bacterial cells undergo the following stages during their growth, depending on the availability of nutrients and the levels of metabolites present in the surrounding environment: lag, acceleration, exponential, retardation, stationary, and death phases (2). Transition to a different cell phase leads to numerous changes in the levels of intracellular molecules and hence altered reaction rates, which leads to noise in gene expression (3–5). Such variables include the concentrations of RNA polymerase (RNAP), RNase, ribosome, and protein, the mRNA degradation rate, the presence of other intracellular molecules, and the environment (6). In particular, cell phase has a pronounced global role in the regulation of transcription and translation, which gives fitness to the population in changing environments (4, 7).

The production of mRNA is a kinetic process that essentially depends on the characteristics of the promoter (8–10). Fluctuation in the concentrations of molecules involved and their binding frequencies determine the rate of transcription initiation (8, 11). It is known that transitions between promoter states are the crucial mechanisms that influence the noise of transcripts. In the active state, the promoter allows the production of an increased number of mRNAs. Rapid successive binding of RNAP produces the burst of mRNA, and the burst size is led by the duration of the promoter's active and inactive states. Consequently, the promoter state determines the number of mRNAs and, hence, protein molecules produced (12), either low or high, producing a bimodal or mixed distribution of these molecules in the population. Promoter gene sequences located in the key regions of DNA have an important regulatory role in the transcription mechanism. However, the transcription transition mechanism of any promoter region, including the well-studied *lac* promoter in *Escherichia coli*, remains elucidative (13).

The *lac* promoter controls the expression of the *lacZYA* struc-

tural genes, which are effectively involved in lactose metabolism (14). Their expression is positively controlled by a catabolite activator protein (CAP), which binds to the upstream activation region of the promoter sequence (URS), and it is negatively controlled by a Lac repressor protein, which is known to bind to operator 1 of the downstream derepression region of the promoter sequence (DRS) (14). Although reports show that the bistable transition of the *lac* promoter leads to phenotypic switching, the roles of the URS and DRS during transition are poorly understood (12, 15, 16). Recently, it was hypothesized that a small burst of proteins was produced from a single mRNA, which burst with a large number of proteins produced from multiple mRNAs (12). It was speculated that the variability in the production of mRNA may be due to differences in the levels of binding of the repressor to the operator. Small and large numbers of mRNAs are produced from partial and full disassociations of the *lac* repressor protein from the URS and DRS, respectively.

It is also unclear whether the promoter gene sequence or biochemical nature of the cellular environment leads to the variation in mRNA numbers (17). The intracellular environment is tightly controlled by the cell phase, during which the transition of the promoter state is a crucial mechanism for gene expression (6, 18). Tracking the transition of the promoter state across phases would provide new insights and help to provide a better understanding of the dynamic nature of gene expression. Notably, most genetic expression studies attempt to explain gene regulation with reference to specific cell phases. The present study aimed to understand how promoter transition occurs from one phase to another, thus regulating the production of mRNA. More specifically, we address how transcriptional bursts are regulated by the transition of the *lac* promoter state at the single-cell/single-molecule resolution level. To generalize our observations, we used $P_{lac/ara\text{-}1}$ (9, 19, 20), a variant of the *lac* promoter with elements of the *ara-1* promoter which allows induction of the URS and DRS independently (19, 20). Here, the upstream activation of $P_{lac/ara\text{-}1}$ is achieved by AraC and L-arabinose instead of cyclic Amp (cAMP) receptor protein (CRP) and cAMP. Upon independent activation of the URS or DRS, it allows us to observe the defined transcription activity induced by URS or DRS activation. Below, we discuss the results of the observed mechanism in detail.

RESULTS

Single-molecule resolution of the transcriptional response. Here, we study the *lac* promoter transition mechanisms that regulate the transcriptional events and mRNA numbers of the target gene at the single-cell level in distinct growth phases (lag, acceleration, exponential, retardation, and stationary) (see Fig. S1 and S2 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf) (21). In the presence of inducers, the *lac* repressor protein falls off from the operator region of the promoter, allowing the successive binding of RNAP to initiate transcription. mRNA production increases with the increase in the concentration of the inducer. Despite the varying inducer concentrations, we used the maximum concentration (5, 9, 12, 22) (full induction) to activate the *lac* promoter (CAP-P_{lac}) in different phases of *E. coli* cell growth (see Fig. S1 and Table 1 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf). Using the MS2-FP (fluorescent protein) method, we also quantified single-mRNA production and analyzed the distribution of the mRNA production (see Fig. S3 and S4 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf) (9, 23). With full induction of *lac*, we observed a unimodal distribution of mRNA molecules produced in all the phases, except the exponential phase (Fig. 1a). The cells at the exponential phase showed the emergence of two mixed populations in terms of RNA production, even with high concentrations of the inducer. In detail, a certain fraction of cells produced 1 to 2 molecules (a low number of mRNAs), and the remaining population produced >2 molecules (a high number of mRNAs), which appears to be like a bimodal distribution of mRNA that was not reported previously at the mRNA level. It suggests that in exponential phase, promoter transition is not stable, which implies that it is either open for a long time to produce more mRNAs or open for a short time to produce a smaller number of mRNAs.

To further validate this mechanism independently, we used a variant *lac* promoter ($P_{lac/ara\text{-}1}$) (9, 19, 20). Figure 1b shows the existence of a bimodal pattern in mRNA distribution in the exponential phase, which is not observed in other phases; this is in agreement with the observed mRNA distribution of the *lac* promoter (23). To determine whether the bimodal behavior is regulated by the URS/DRS of the promoter, we activated the URS and DRS of the *lac* variant independently (see Fig. S1 and Table 1 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf). Figure 1c shows a unimodal distribution of mRNA in all phases, during which the independent activation of the URS is capable of producing low numbers of mRNA (1 to 2 molecules), apparently in a single-grade mode. Moreover, the DRS produced a broad range of mRNAs (1 to 20 molecules); we also observed the existence of a single-grade mode alone (Fig. 1c). We also noticed that in a population, the activated URS and DRS tend to produce low and high numbers of mRNA, respectively (Fig. 1c). It is possible that the URS has a stronger RNAP binding site and a weaker promoter than the DRS, which is in close agreement with previous observations (19, 20, 24). Notably, when cells receive inducers to activate both the URS and the DRS, the promoter regulates the production of a small number (1 to 2 molecules) and a large number (>2 molecules) of mRNAs at low and high frequencies, respectively (Fig. 1b). Activation appears to occur in a binary mode, confirming the existence of two distinct populations (Fig. 1b). Thus, we conclude that the bimodal mRNA distribution results from the coordinated regulation by URS/DRS gene activation. We suggest that there are multiple stochastic events in promoter transition that trigger phenotypic switching (1); this may be due to the competitive binding of RNAP at these sites (24). Moreover, we see the same general features of mRNA distribution in the *lac* variant as in endogenous *lac*, which further supports our conclusion regarding the bimodal behavior due to the complexity of promoter transition (Fig. 1a and b). We also observed that bias in the partitioning of RNAs does not lead to phenotypic switching (Fig. 1e), despite the change in the frequency of mRNA production (Fig. 1f), which in turn triggers the bimodality.

Figure 1d shows the trend of mean numbers of mRNAs ($\langle m \rangle$) per cell from endogenous *lac* and the *lac* variant; i.e., means of 10 and 8 RNA molecules per cell were produced from *lac* and the *lac* variant, respectively. The mean number of mRNAs produced increases and decreases over the phases. Figure 1d also supports the possibility that the independently activated URS results in a low $\langle m \rangle$ per cell compared to the DRS. Further, we calculated the cell-to-cell heterogeneity using the Fano factor, $\langle b \rangle$ [*variance* (σ^2)/

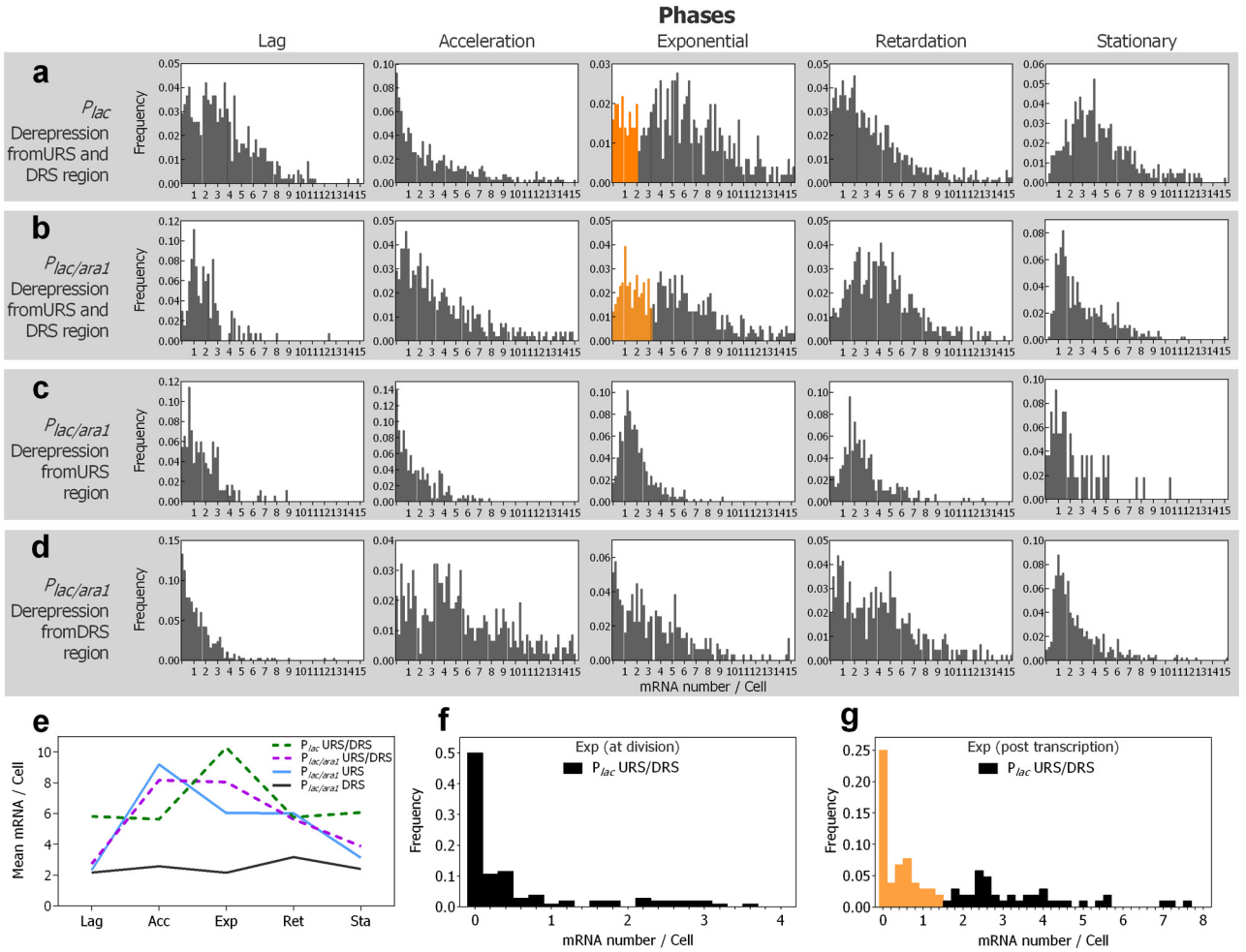

FIG 1 Transcriptional responses of *lac* and the *lac* variant in multiple cell growth phases of *E. coli*. (a) Probability distributions of mRNA numbers in the population, where the P_{lac} URS was activated by cAMP and the DRS was activated by IPTG. A bimodal distribution of mRNA numbers was observed only in exponential phase (1 to 2 mRNAs is considered one population [orange], and more than 2 mRNAs is considered another population [black]). (b) The $P_{lac/ara\text{-}1}$ URS is activated by L-arabinose, and the DRS is activated by IPTG. The bimodal distribution seen in exponential phase is similar to that observed with P_{lac}. (c) The $P_{lac/ara\text{-}1}$ URS is activated by cAMP (cyan), and the DRS is activated by IPTG (black) independently. An exponential distribution was observed here, unlike the bimodal distribution when both the URS and the DRS were activated. Over 100 cells were analyzed in each induction (see Table S2 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf). (d) Mean numbers of mRNA/cell over the phases of all conditions, with P_{lac} activated from the URS and DRS (green), $P_{lac/ara\text{-}1}$ activated from the URS and DRS (red), $P_{lac/ara\text{-}1}$ activated from the URS (black), and $P_{lac/ara\text{-}1}$ activated from the DRS (blue). The trends of P_{lac} and $P_{lac/ara\text{-}1}$ are similar over the phases, and the mean number of mRNAs/cell with $P_{lac/ara\text{-}1}$ is higher with activation from the DRS than with activation from the URS. (e) Time-lapse microscopy captures the partitioning of mRNA molecules produced from the activated *lac* promoter. The distribution of mRNA numbers does not show bimodality. Acc, acceleration phase; Exp, exponential phase; Ret, retardation phase; Sta, stationary phase. (f) Time-lapse microscopy of the same cells captured a phenotypic switch at the single-event level. Induction of the *lac* promoter produced mRNAs in bursts, and the frequency of production events determined the phenotypic switches. A bimodal mRNA distribution shows a fraction of the population, with one to two mRNA molecules, and another fraction with a higher number of mRNAs (>2). These data are from 54 random cells which were dividing in exponential phase (Exp). (g) From the 120-min time-lapse microscopy images, divided cells (108 sister cells) were observed and followed over the time. At 60 min after the division, the number of mRNAs/cell was calculated and plotted. A strong bimodal distribution of mRNAs in a large population is shown in panel a (middle graph).

mean ($\langle n \rangle$)]. Figure 2a and b show the increase and decrease in mRNA noise behavior over the phases, calculated from variance (σ^2) and mean ($\langle n \rangle$) ($\langle b \rangle = \sigma^2/\langle n \rangle$). If mRNA production is non-bursty (Poissonian), $\langle b \rangle$ is equal to 1. For *lac* and the *lac* variant when they are fully activated, a [$\langle b \rangle$] of >1 indicates that mRNA molecules may be produced from bursty transcriptional events (Fig. 2a). The independently activated URS and DRS have a [$\langle b \rangle$] that is >1 (Fig. 2b). The highest value in $\langle b \rangle$ not only is an indication of bursty transcription but also might be the result of biased portioning of mRNA molecules being produced. Figure 1f shows that activating the promoter from the URS and/or DRS increases the mRNA noise. We also observed an increase in $\langle b \rangle$ when gene expression was induced. The trend of the Fano factor follows cell division (Fig. 2a and b). From the Pearson correlation coefficient, it is clear that the trend of $\langle b \rangle$ follows a trend similar to that of $\langle m \rangle$ under all conditions and phases. We next performed Kolmogorov-Smirnov tests to assess the distinctiveness between the trend of induction conditions of the two promoters over the phases; the resulting changes in mean mRNA production and Fano factor are statistically distinguishable (see Table S3 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf).

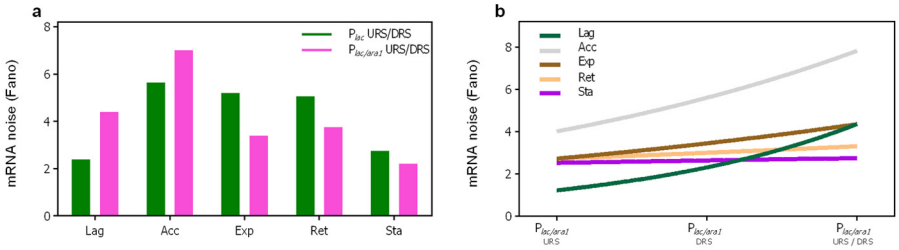

FIG 2 mRNA noise of the promoter. (a) Trend of mRNA numbers (Fano factors) over the gene activation states and cell phases. The Fano factors of P_{lac} and $P_{lac/ara-1}$ activated from the URS and DRS over the phases show similar trends in variation. The factor first increases and then decreases. (b) In the case of independent activations of $P_{lac/ara-1}$, the Fano factor increases when gene activation increases. The higher mRNA noise in the acceleration and exponential phases is indicated.

Real-time dynamics of gene regulation. To further investigate whether the observed $\langle b \rangle$ is an output of the slow or fast transition of the promoter in specific cell phases, we monitored the kinetics of mRNA production at single-event levels of individual cells for all inductions and phases. Analysis of the data allows us to address the following questions. Is the burst of mRNAs controlled by the URS or the DRS or both? Does a cell phase change the promoter transition and thus burst size? What is the burst behavior of the URS/DRS during cell-phase transition? Is gene-specific transcription transition regulated by cell growth phase?

To address these questions, we tracked individual cells from time-lapse microscopy; the moments of appearance of new RNA molecules (see Fig. S5 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf and Movie 1 at http://www.cs.tut.fi/~kandhave/supplementary/supporting_movie_1.avi) were obtained by fitting a monotonic piecewise-constant curve, in a least-squares sense, to the corrected intensity of signals obtained from a cell (25). Figure 3a to d show the probabilities of mRNA production from each burst as a function of induction (of the URS, the DRS, and both) and phases. The *lac* gene and the *lac* variant produce the mRNA molecules in burst at full induction with identical burst frequency distributions (Fig. 3a and b). Independent activations of the URS (Fig. 3c) and DRS (Fig. 3d) of the *lac* variant produce single mRNAs with high probability, whereas the DRS produces a higher number of mRNA molecules with the lowest probability (Fig. 3d). Comparison of burst behavior revealed that the burst size of the DRS is greater than that of the URS (Fig. 3e). This implies that the burst frequency is modulated by the URS and that its size is modulated by the DRS. Figure 2c and d also show the burst behavioral changes over the phases. They show that, in the lag, acceleration, and exponential phases, the URS is more active than the DRS. This activation trend is in contrast to those of later phases (retardation and stationary). It indicates that during the early phases of cell growth, the promoter transition is regulated by the URS and later by the DRS. It may have some relation to the metabolic demand and activates the DRS when cells reach a stressful environment. Since our aim is restricted to studying production kinetics, we did not consider how produced mRNA molecules diffuse once they are produced.

Several features are revealed from the time series data analysis of transcription. Figure 3e shows the burst behavior of the endogenous *lac* promoter and its variant where both the promoters produce larger $\langle m \rangle$ per burst. Markedly at full induction, the probability of burst size and frequencies (Fig. 3a and b) are modulated over the phases due to the combinatorial activation of the URS and DRS. In exponential phase, the $\langle m \rangle$ per burst is larger than during other phases. To be precise, the independent activation of the URS/DRS produced a low $\langle m \rangle$ per burst (1 to 1.4 mRNAs) compared to that of the fully active promoter (2.5 to 4 mRNAs). From all of the above evidence, we conclude that the transcription kinetics is controlled by sequences both upstream and downstream of the promoter by changing the burst size and frequency depending on the cell phase. We also point out the possible relation between our findings and the results of Golding et al. (9), who describe the burst in transcription initiation of the variant *lac* promoter at exponential phase (9). Burst may arise from the initiation of many transcriptional events in a short time window, indicating the fast and consecutive binding of multiple RNAPs. It may also arise from the pausing of RNAP during transcription (9, 26, 27).

To understand the role of promoter architecture in the kinetics of transcription and noises, burst production intervals over the phases were analyzed (see Fig. S6 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf). At first, we calculated the mean burst intervals of *lac* and the *lac* variant under full induction. Figure 3f shows that the kinetics of mean intervals are identical. The production interval becomes longer when cells transit from lag phase to exponential phase; then it becomes shorter during the retardation phase and becomes longer again in the stationary phase. Nonidentical behaviors are noticed when the gene is activated by the URS and DRS, independently. In detail, Fig. 2f shows that in the early stages of the cells, an independent activation of the URS produced the burst with longer time intervals but that the DRS does the opposite. This is in agreement with our previous observations demonstrating that longer and shorter intervals allow the burst of larger and smaller numbers of mRNAs, respectively. In exponential phase, both regions actively produce multiple mRNAs (Fig. 3f). In the later phases, the promoter is expressed in the opposite manner; i.e., the DRS shows a longer interval than the URS, which agrees with the observed burst size and frequency (Fig. 3c and d). This proves that the URS and DRS behaviors play against each other over the phases. From the above findings, it is possible to assume the occurrence of similar regulations of *lac* and the *lac* variant in full induction (19, 20). In the early stages of cells, the URS is actively involved, and in the later stages, the DRS is actively involved. In exponential phase, the URS and DRS work in a coordinated fashion in transcribing the target gene. We note that attached RNAP may transcribe the gene from the upstream promoter alone (28); it can also bind down-

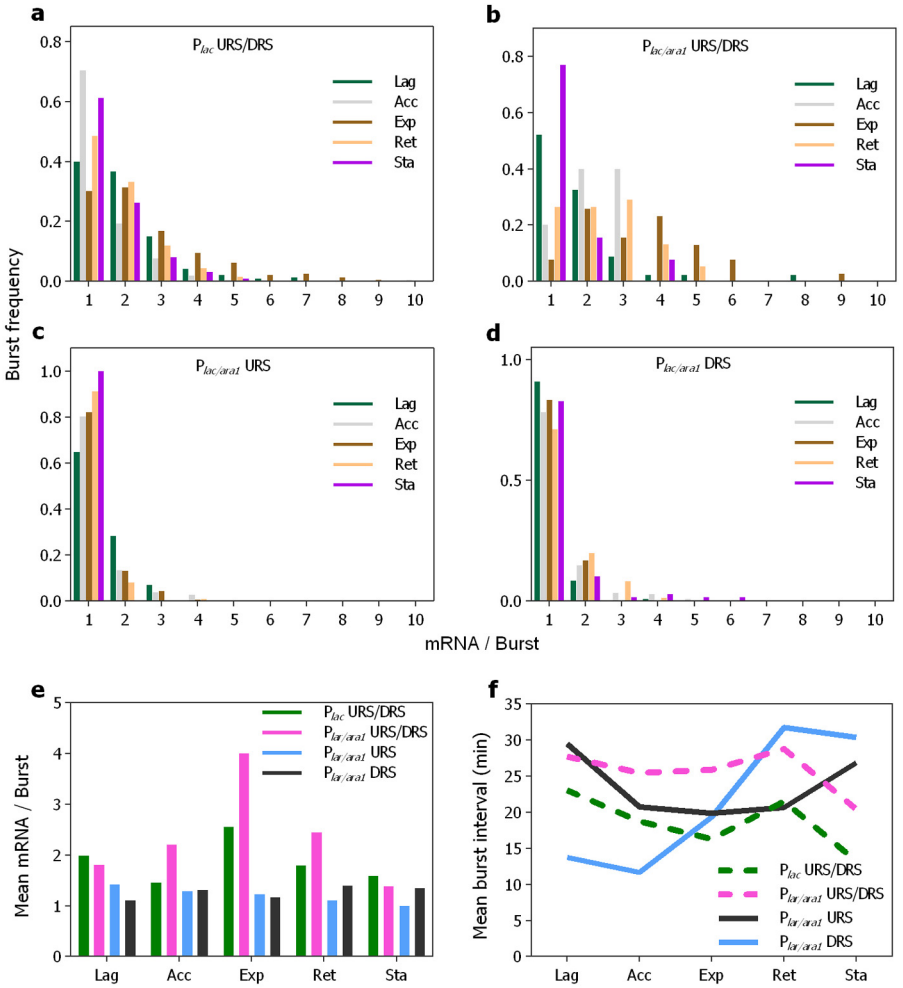

FIG 3 Tracking of real-time dynamics of gene regulation. Probability distributions of mRNA burst size from the endogenous P_{lac} (a) and $P_{lac/ara-1}$ (b) promoters in all phases. P_{lac} and $P_{lac/ara-1}$ show similar trends in the exponential phase. (c) Independent activation of the URS by $P_{lac/ara-1}$. (d) Independent activation of the DRS by $P_{lac/ara-1}$. In both cases of URS and DRS activation, the probability of producing a single RNA is higher. These distributions are different from those with the URS/DRS activated from P_{lac} and $P_{lac/ara-1}$. Over 100 cells were analyzed in each induction (see Table S2 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf). (e) Mean numbers of mRNAs/burst in all phases. (f) Mean burst intervals of all activation states over the phases. The trends of endogenous P_{lac} and $P_{lac/ara-1}$ are similar, where URS and DRS regulation shows inverse trends, with the midpoint of equal activations being in exponential phase.

stream of the gene in parallel to transcribe the target gene (24, 26, 27).

From the time series, we calculated the transcriptional "noise," quantified by the proportion of the squared coefficient of variance ($\eta^2 = \sigma^2/\langle n \rangle^2$) and mean. Figure 4a and b show that the noise increases monotonically with the mean; the higher the mean, the

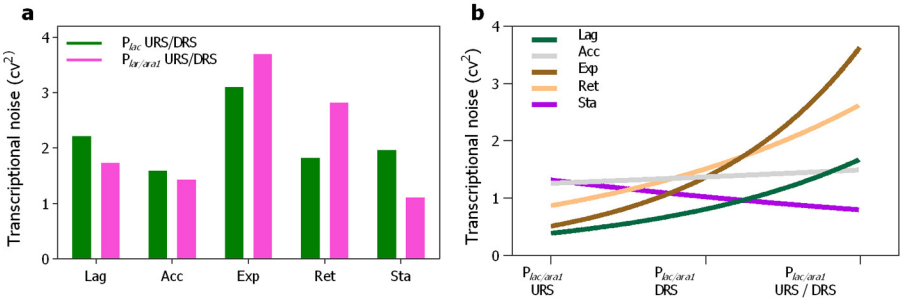

FIG 4 Promoter dynamics noise. (a) Transcriptional noise of P_{lac} and $P_{lac/ara-1}$ showing similar trends. (b) Trends of transcriptional noise in different modes of gene activation by $P_{lac/ara-1}$. The noise increases as gene activation increases. Also, accelerating cells exhibit greater noise than cells in other phases.

greater the noise reported as characteristics of extrinsic noise (5). With full induction of *lac* and the *lac* variant, an η^2 of >1 indicates the super-Poissonian behavior of transcription in all phases. Moreover, the transcriptional noise is proportional to the mRNA noise of the population. At the same time, mRNA noise is higher than expected; it may be an added variable of error in partitioning (29). Also, the observed result is in close agreement with previous observations of *lac* promoter expression (9) in exponential phase ($\sigma^2/\langle n\rangle = 3.6$). The Kolmogorov-Smirnov test reveals that the trends of observed transcriptional noise under all conditions over the phases (see Table S4 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf) are statistically distinguishable (Fig. 4a). We also note that the URS is less noisy than the DRS when it is independently activated; it further increases when fully activated (Fig. 4b). This also shows that transcriptional noise of fast-dividing and slow-dividing cells has registered an increasing and decreasing trend, respectively; these results are in close agreement with the observed trend of the $\langle m\rangle$ per burst (Fig. 3e). The observed results provide evidence that the cell phase and stochastic transcription play vital roles in transcription and mRNA noise in a combinatorial fashion to regulate the level of heterogeneity.

DISCUSSION

Although it is known that a high concentration of inducer can actively pull off the repressor from the operator, as described by Jacob and Monod (2), the role of promoter transition in bimodal distribution is not clear. Novick and Weiner hypothesized that a single rate-limiting step determines the phenotypic switching of the population (30). This is also supported by Xie et al., as single-molecular events trigger the phenotypic switch (12). They further show that one molecular permease protein is not enough to switch the phenotype. If the induction of switching is due to a single rate-limiting step, as argued by the Choi and Novick groups, there must be a range of small and a large number of mRNA molecules produced at high concentrations of inducer, and it has to appear as a single distribution. Since we observed a bimodal distribution of mRNAs (Fig. 1a) in the population, this suggests the existence of more than one rate-limiting step in promoter transition similar to that of the model proposed (16, 24). Further, the study explains a coordinated functioning of intrinsic and extrinsic noise. Such mechanisms are observed mostly in eukaryotic transcription, which produces an mRNA burst coupled with chromatin modification (1, 26, 27). Recent single-cell studies hypothesize that in prokaryotes, DNA unlooping triggers the stochastic events in transcription that determine phenotype switching (12, 16). Our experiments provide evidence that the phenotypic switch occurs due to the coordinated behaviors of the URS and DRS, and their activities are also coupled with cell phase to determine the promoter transition.

Previous *in vivo* and *in vitro* (2, 12, 30) studies hypothesize that bimodality may arise from a single transcript. Our results support this phenomenon but are not restricted to it. Our primary result is the complete characterization of the *in vivo* kinetics of transcription transition over the phases. It extends to capture the nature of transition activated from either the upstream or the downstream promoter region. No previous work has shown the details of promoter transition and its importance in transcriptional regulation. Also, there has been no report of bimodality in mRNA distribution or how mRNAs are regulated at the population level *in vivo*. These results thus allow us to characterize the plasticity and kinetics of this promoter *in vivo* from the measurements of single mRNA molecules at the single-cell level. The observed experimental and computational models provide evidence that promoter sequences have a significant role in the transitioning of the promoter, which determines the stochasticity in mRNA burst frequency and size during cell growth. Continuous utilization of given nutrition leads to a depletion state and thus to a stressful environment in which survival depends on phenotypic characteristics. Previously, it was also noted that *E. coli* adjusts its gene expression continuously in Luria-Bertani broth by using diverse nutrients over time, which also agrees with our observations (31). Our study shows that, even though cells receive sufficient nutrition, they favor the activation of the URS region through CRP/cAMP complex formation in the early phases of cell growth, while derepression of OR-1 (operator 1 region) appears mainly to regulate transcription kinetics in later phases. It is possible that transcription factor concentration determines the transcription kinetics at a given stage of cell growth. Thus, the roles of the two transcription regulatory mechanisms seem to vary depending on cell state and extracellular environment. This also gives us a clue that bacterial cells may reserve the DRS region of DNA for activation in the stressed environment.

The $P_{lac/ara-1}$ promoter is a well-characterized promoter with a logical structure very similar to that of P_{lac}. Their identical repression systems, their binding regions for the LacI repressor protein, are localized in the downstream region of the promoter. Moreover, the lack of an operator 3 region in both promoters prevents the DNA looping mechanism from operating in the promoter region. Finally, a comparison of the gene sequences of both promoters, from the upstream regulatory sequence to the end of the operator 1 region, a region of 96 nucleotides, indicates 65.5% similarity. The key region differentiating these two promoters is the URS, which in the case of the Lac variant is an AraC binding region, allowing tight control of its activation, unlike in P_{lac}, where URS activation is controlled by endogenous cAMP. The observed biophysical mechanism possibly exists in other structurally different promoters. However, multiple variables responsible for transcription initiation, including activator and repressor molecules, need to be considered in order to understand the molecular mechanisms over the phases. In the present work, we consider only one variable change in the upstream region. Overall, the tendency for stochastic promoter transition in gene activation and changes during the growth phases of the cells are crucial to understand the process. Gene regulation depends upon the environment, which in turn determines phenotypic diversity.

MATERIALS AND METHODS

Strains and media. Strains were constructed using standard molecular biology techniques. We used the *Escherichia coli* Dh5α-PRO strain as a host to study gene expression over the phases. At first, to study mRNA transcription at the single-cell level, we created two strains (see Fig. S1 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf). The first system is cotransformed with two plasmids. The first is a single-copy bacterial artificial chromosome, pTRUEBLUE-BAC2 (resistant to chloramphenicol), expressing the mRFP1 gene and 96 target binding sites (96xbs) for MS2 protein. It is under the control of the wild-type *lac* promoter, with the catabolic activator protein (CAP) site in the upstream region for the binding of CAP and an operator 1 site in the downstream region for the binding of the *lac* repressor (23) (generously provided by Ido Golding, University of Illinois). The CAP and Lac re-

pressors are inducible by cyclic AMP (cAMP) and isopropyl-D-1-thiogalactopyranoside (IPTG), respectively. The second is a medium-copy-number reporter plasmid, pPROTET.E (resistant to kanamycin), expressing the bacteriophage MS2 coat protein tagged with the green fluorescent protein (GFP)-coding gene (*mut3*) under the control of the promoter of operator 1 of *tet* (P_{teto-1}) (23) (see Fig. S1 and Table 1 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf).

The second system also consists of two plasmids as described above, one of which (the reporter plasmid) is identical to that described above. The target plasmid for this system expresses 96xbs for the MS2 protein under the control of a variant *lac* promoter, $P_{lac/ara-1}$, with the I1-I2 binding site of *araC* in the upstream region and the *lac* operator 1 binding site for the *lac* repressor in the downstream region (generously provided by Ido Golding) (9) (see Fig. S1 and Table 1 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf).

All strains were grown in liquid LB medium, composed of 10 g/liter of tryptone (T7293; Sigma Aldrich, USA), 5 g/liter of yeast extract (MC 001; Lab M, United Kingdom), and 10 g/liter of NaCl (S3014; Lab M, United Kingdom). Antibiotics were added according to the antimicrobial resistance of the respective strain.

Cell phase determination and induction. We first obtained the generation time of the *E. coli* DH5a-PRO strain to determine the phases of the cells. For this purpose, cells were grown overnight in an orbital shaker (Labnet) at 30°C with aeration at 250 rpm. Following overnight culture, cells were diluted in fresh medium to reach an optical density at 600 nm (OD_{600}) of 0.05 as determined with a spectrophotometer (Ultrospec 10; Amersham Biosciences) and grown at 37°C with aeration at 250 rpm. Cell growth was measured every 30 min, and OD_{600} values were used to calculate the generation time as described in earlier literature (32) (see Fig. S2 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf). Cell phases were determined as suggested from the Jacob-Monod study based on the state of the cell division time, i.e., for the lag, acceleration, exponential, retardation, and stationary phases (21).

To study the regulation of *lac* and the variant *lac* promoter over the phases, cells were induced into the state of a particular phase. To detect a single mRNA molecule, the reporter gene was activated with 50 ng/ml of anhydrotetracycline (aTc) (lot no. 2-0401-001; IBA GmbH, Germany) for a 20-min incubation. Full induction of the P_{lac} target was achieved by the activation of the DRS with 1 mM IPTG (L6758; Sigma Aldrich, USA) and of the URS with endogenous cAMP (23). For the induction of $P_{lac/ara-1}$, 0.1% L-arabinose (A3256; Sigma Aldrich, USA) and 1 mM IPTG were used for the URS and DRS, respectively (9). In the case of $P_{lac/ara-1}$ induction, depending on the regulation scheme, the promoter was activated either from the URS or the DRS using corresponding inducers (see Table S1 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf). In the population study, to observe single mRNAs, the cells were incubated with inducers for 15 min, depending on the induction scheme in each phase. The cells were observed under a microscope immediately after the division time of every phase. For the time trace experiments, cells were induced with aTc as explained above and corresponding inducers for 10 min according to the induction scheme of the target promoter (see Table S1 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf). Cells were taken prior to the division time and observed under a microscope for 120 min. For the population and time-lapse microscopy study, over 100 cells were analyzed in each induction (see Table S2 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf).

Fluorescence microscopy of live cells. For the total population analysis, *E. coli* DH5a-PRO cells were pelleted by centrifugation at 6,000 rpm and then suspended in the required volume of LB broth to facilitate the observation of multiple cells under a microscope. Suspended cells were then placed on a microscope slide between a 1% LB broth-agarose gel pad and a coverslip.

For the image acquisition of the produced mRNA molecules in the cells, we used an inverted fluorescence microscope (Eclipse Ti-E; Nikon) with a 100× 1.49-numerical-aperture (NA) oil immersion objective. The microscope is equipped with a hardware autofocus module, motorized z drive, and Nikon's Perfect Focus System (PFS) to maintain the cells in focus during image acquisition. The PFS function automatically corrects small changes in focus to capture the cell. Built-in microscope software (NIS-Elements C; Nikon) was used to acquire images. Fluorescence was measured using a 488-nm laser (Melles Griot) and a 515/30-nm detection filter.

For the real-time monitoring of mRNA production, reporter and target promoters were activated as described above. Cells were then placed between a 1% LB broth-agarose gel pad and a coverslip in an FCS2 temperature-controlled perfusion chamber (Bioptechs), maintained at 37°C during the measurement. The pad was supplemented with required inducers to maintain the induction conditions for time-lapse microscopy. To monitor the synchronously dividing cells, image acquisition was initiated prior to the division time and captured under a microscope for a period of 120 min.

Cell segmentation and image processing. From the images acquired by microscopy, we selected cells that were producing mRNA molecules. Bacterial cell detection was performed using a semiautomated method consisting of manual masking of the region of the images displaying cells containing fluorescent spots (see Fig. S3 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf). Application of principal-component analysis (PCA) allowed us to extract information about the dimensions, locations, and orientations of masked cells. PCA assumes that the fluorescence intensity is relatively uniformly distributed within cells and approximately zero in the background (25, 33). Further, segmentation of fluorescent spots within the cells was implemented by subcell object detection methods (25, 34). Particularly, we used kernel density estimation (Gaussian kernel) along with Otsu's threshold method (as noted in reference 25). Notably, determining precise locations of segmented fluorescent spots reduces false negatives or missed spots and, thus, avoids the enlarged noise of intensity signals of spots in the cells. The threshold value was selected so that intraclass variance was minimized.

Next, segmented cells were connected in lineage across time by matching their closest centroids from one frame to the following, as determined by PCA, in order to quantify the mRNA production kinetics over time by using time-lapse microscopy images. As quenching does not occur throughout the experiment, absolute total intensities from background-corrected spot intensities were fit to a monotonic piecewise-constant curve (25, 34). From there, by tracking significant increases in intensity visible in the monotonic step-like curve, the times of appearance of new mRNAs and intervals between productions of consecutive RNA molecules were determined. Jumps in the curve indicate the appearance of mRNA molecules and the transcription interval between two consecutive events. However, such a jump detection method does not quantify mRNA molecules produced in the event. To determine the quantity of mRNA molecules produced from each event, the mean total intensity of each intensity jump was used and mRNA numbers were calculated as described below. The time interval represents the time necessary to produce the mRNA as a burst. From intensity analysis, we were able to discriminate whether they were produced as single mRNAs or as multiple mRNAs in a burst.

Quantification of mRNAs and burstiness at single-molecule resolution. To quantify single-mRNA intensity, we followed the production of mRNA under weak induction. For this, cells were grown at 37°C with aeration at 250 rpm, with aTc for reporter activation and without inducers for the target-mRNA production. From time series images, we sought the first appearance of a fluorescent spot in 50 random cells. We then extracted the fluorescence intensity of each mRNA spot and calculated the mean intensity (see Fig. S4 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf), which is defined as the mean intensity value of a single mRNA molecule. Combining information

from the jump detection method, determining cell background-corrected spot intensities and the mean intensity of a single mRNA molecule allowed us to calculate the number of mRNA molecules produced in the transcription events.

Further, it is known that P_{lac} mRNA molecules are produced in burst-like transcription (6), and we also observed the same behavior. In order to study the transcriptional bursting behavior over phases, we investigated the transcription events of individual cells. The production interval between the burst was also calculated from the intensity jump. The model for its examination was implemented in MATLAB R2013a software. From the data, we calculated the number of mRNAs produced in a burst (burst size) and the occurrence of such an event in a population (frequency).

Transcription noise. From the population of cells, to measure the noise as a function of mRNA number, we used the ratio of the variance to the mean of the distribution of mRNA number per cell, i.e., the Fano factor ($\langle b\rangle = \sigma^2/\langle n\rangle$). Quantification of the total noise in transcriptional events was determined using the square root of variance over the mean (i.e., $\eta^2 = \sigma^2/\langle n\rangle^2$). For this, we used the transcription production interval observed from time-lapse microscopy. This allowed us to characterize the transcriptional noise in the different phases over the induction cases.

Statistical tests. To determine whether the trend of cell phase follows the trend of mean mRNA distributions and Fano factors, we used the Wilcoxon-Mann-Whitney test with a threshold of 0.05. To study the correlation between the mean number of mRNAs and production interval, we used the Pearson correlation coefficient test. The fold change in numbers of mRNAs and total noise under different induction conditions were measured using the Kolmogorov-Smirnov test, with a threshold of 0.05. Using this test, we tested the significant changes in mean numbers of mRNAs between the induction conditions of the two promoters over the phases. The test shows significance levels for the observed changes in mean numbers of mRNAs for P_{lac} and $P_{lac/ara\text{-}1}$ (see Fig. S3 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf). Further, we tested the significance levels of changes in mRNA noise over the conditions. Results are shown in Table S4 at http://www.cs.tut.fi/~kandhave/supplementary/Supplementary%20Material.pdf.

ACKNOWLEDGMENTS

We thank the A. S. Ribeiro team for helpful discussions and technical help, I. Golding for the generous gift of *lac* and the *lac* variant MS2-GFP system and C. Balasundaram for helpful discussion and his critical reading of the manuscript.

This work was supported by the Academy of Finland.

REFERENCES

1. **Kaern M, Elston TC, Blake WJ, Collins JJ.** 2005. Stochasticity in gene expression: from theories to phenotypes. Nat Rev Genet **6:**451–464. http://dx.doi.org/10.1038/nrg1615.
2. **Jacob F, Monod J.** 1961. Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol **3:**318–356. http://dx.doi.org/10.1016/S0022-2836(61)80072-7.
3. **Raj A, van Oudenaarden A.** 2008. Nature, nurture, or chance: stochastic gene expression and its consequences. Cell **135:**216–226. http://dx.doi.org/10.1016/j.cell.2008.09.050.
4. **Sanchez A, Golding I.** 2013. Genetic determinants and cellular constraints in noisy gene expression. Science **342:**1188–1193. http://dx.doi.org/10.1126/science.1242975.
5. **So L-H, Ghosh A, Zong C, Sepúlveda LA, Segev R, Golding I.** 2011. General properties of transcriptional time series in Escherichia coli. Nat Genet **43:**554–560. http://dx.doi.org/10.1038/ng.821.
6. **Swain PS, Elowitz MB, Siggia ED.** 2002. Intrinsic and extrinsic contributions to stochasticity in gene expression. Proc Natl Acad Sci U S A **99:**12795–12800. http://dx.doi.org/10.1073/pnas.162041399.
7. **Acar M, Mettetal JT, van Oudenaarden A.** 2008. Stochastic switching as a survival strategy in fluctuating environments. Nat Genet **40:**471–475. http://dx.doi.org/10.1038/ng.110.
8. **Kierzek AM, Zaim J, Zielenkiewicz P.** 2001. The effect of transcription and translation initiation frequencies on the stochastic fluctuations in prokaryotic gene expression. J Biol Chem **276:**8165–8172. http://dx.doi.org/10.1074/jbc.M006264200.
9. **Golding I, Paulsson J, Zawilski SM, Cox EC.** 2005. Real-time kinetics of gene activity in individual bacteria. Cell **123:**1025–1036. http://dx.doi.org/10.1016/j.cell.2005.09.031.
10. **Sanchez A, Garcia HG, Jones D, Phillips R, Kondev J.** 2011. Effect of promoter architecture on the cell-to-cell variability in gene expression. PLoS Comput Biol **7:**e1001100. http://dx.doi.org/10.1371/journal.pcbi.1001100.
11. **McAdams HH, Arkin A.** 1997. Stochastic mechanisms in gene expression. Proc Natl Acad Sci U S A **94:**814–819. http://dx.doi.org/10.1073/pnas.94.3.814.
12. **Choi PJ, Cai L, Frieda K, Xie XS.** 2008. A stochastic single-molecule event triggers phenotype switching of a bacterial cell. Science **322:**442–446. http://dx.doi.org/10.1126/science.1161427.
13. **Hammar P, Walldén M, Fange D, Persson F, Baltekin O, Ullman G, Leroy P, Elf J.** 2014. Direct measurement of transcription factor dissociation excludes a simple operator occupancy model for gene regulation. Nat Genet **46:**405–408. http://dx.doi.org/10.1038/ng.2905.
14. **Dickson RC, Abelson J, Barnes WM, Reznikoff WS.** 1975. Genetic regulation: the lac control region. Science **187:**27–35. http://dx.doi.org/10.1126/science.1088926.
15. **Earnest TM, Roberts E, Assaf M, Dahmen K, Luthey-Schulten Z.** 2013. DNA looping increases the range of bistability in a stochastic model of the lac genetic switch. Phys Biol **10:**026002. http://dx.doi.org/10.1088/1478-3975/10/2/026002.
16. **Choi PJ, Xie XS, Shakhnovich EI.** 2010. Stochastic switching in gene networks can occur by a single-molecule event or many molecular steps. J Mol Biol **396:**230–244. http://dx.doi.org/10.1016/j.jmb.2009.11.035.
17. **Taniguchi Y, Choi PJ, Li G-W, Chen H, Babu M, Hearn J, Emili A, Xie XS.** 2010. Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. Science **329:**533–538. http://dx.doi.org/10.1126/science.1188308.
18. **Elowitz MB, Levine AJ, Siggia ED, Swain PS.** 2002. Stochastic gene expression in a single cell. Science **297:**1183–1186. http://dx.doi.org/10.1126/science.1070919.
19. **Lutz R, 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. http://dx.doi.org/10.1093/nar/25.6.1203.
20. **Lutz R, Lozinski T, Ellinger T, Bujard H.** 2001. Dissecting the functional program of Escherichia coli promoters: the combined mode of action of lac repressor and AraC activator. Nucleic Acids Res **29:**3873–3881. http://dx.doi.org/10.1093/nar/29.18.3873.
21. **Monod J.** 1949. The growth of bacterial cultures. Annu Rev Microbiol **3:**371–394. http://dx.doi.org/10.1146/annurev.mi.03.100149.002103.
22. **Elf J, Li GW, Xie XS.** 2007. Probing transcription factor dynamics at the single-molecule level in a living cell. Science **316:**1191–1194. http://dx.doi.org/10.1126/science.1141967.
23. **Golding I, Cox EC.** 2004. RNA dynamics in live Escherichia coli cells. Proc Natl Acad Sci U S A **101:**11310–11315. http://dx.doi.org/10.1073/pnas.0404443101.
24. **Malan TP, McClure WR.** 1984. Dual promoter control of the Escherichia coli lactose operon. Cell **39:**173–180. http://dx.doi.org/10.1016/0092-8674(84)90203-4.
25. **Kandhavelu M, Häkkinen A, Yli-Harja O, Ribeiro AS.** 2012. Single-molecule dynamics of transcription of the lar promoter. Phys Biol **9:**026004. http://dx.doi.org/10.1088/1478-3975/9/2/026004.
26. **Guptasarma P.** 1995. Does replication-induced transcription regulate synthesis of the myriad low copy number proteins of *Escherichia coli*? Bioessays **17:**987–997. http://dx.doi.org/10.1002/bies.950171112.
27. **Guptasarma P.** 1996. Cooperative relaxation of supercoils and periodic transcriptional initiation within polymerase batteries. Bioessays **18:**325–332. http://dx.doi.org/10.1002/bies.950180411.
28. **Benoff B, Yang H, Lawson CL, Parkinson G, Liu J, Blatter E, Ebright YW, Berman HM, Ebright RH.** 2002. Structural basis of transcription activation: the CAP-αCTD-DNA complex. Science **297:**1562–1566. http://dx.doi.org/10.1126/science.1076376.
29. **Lloyd-Price J, Lehtivaara M, Kandhavelu M, Chowdhury S, Muthukrishnan A-B, Yli-Harja O, Ribeiro AS.** 2012. Probabilistic RNA partitioning generates transient increases in the normalized variance of RNA numbers in synchronized populations of *Escherichia coli*. Mol Biosyst **8:**565–571. http://dx.doi.org/10.1039/c1mb05100h.

30. **Novick A, Weiner M.** 1957. Enzyme induction as an all-or-none phenomenon. Proc Natl Acad Sci U S A **43:**553–566. http://dx.doi.org/10.1073/pnas.43.7.553.
31. **Sezonov G, Joseleau-Petit D, D'Ari R.** 2007. *Escherichia coli* physiology in Luria-Bertani broth. J Bacteriol **189:**8746–8749. http://dx.doi.org/10.1128/JB.01368-07.
32. **Powell EO.** 1956. Growth rate and generation time of bacteria, with special reference to continuous culture. J Gen Microbiol **15:**492–511. http://dx.doi.org/10.1099/00221287-15-3-492.
33. **Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA, Moffat J, Golland P, Sabatini DM.** 2006. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol **7:**R100. http://dx.doi.org/10.1186/gb-2006-7-10-r100.
34. **Kandhavelu M, Lloyd-Price J, Gupta A, Muthukrishnan AB, Yli-Harja O, Ribeiro AS.** 2012. Regulation of mean and noise of the *in vivo* kinetics of transcription under the control of the lac/ara-1 promoter. FEBS Lett **586:**3870–3875. http://dx.doi.org/10.1016/j.febslet.2012.09.014.