



Published in final edited form as:

*Science*. 2013 December 6; 342(6163): 1188–1193. doi:10.1126/science.1242975.

## Genetic determinants and cellular constraints in noisy gene expression

Alvaro Sanchez<sup>1,\*</sup> and Ido Golding<sup>2,3,4,\*</sup>

<sup>1</sup>The Rowland Institute at Harvard, Harvard University, Cambridge MA 02142

<sup>2</sup>Verna & Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston TX 77030

<sup>3</sup>Center for Theoretical Biological Physics, Rice University, Houston TX 77005

<sup>4</sup>Center for the Physics of Living Cells, University of Illinois at Urbana-Champaign, Urbana IL 61801

### Abstract

In individual cells, transcription is a random process obeying single-molecule kinetics. Often, it occurs in a bursty, intermittent manner. The frequency and size of these bursts affect the magnitude of temporal fluctuations in mRNA and protein content within a cell, creating variation or “noise” in gene expression. It is still unclear to what degree transcriptional kinetics are specific to each gene and determined by its promoter sequence. Alternative scenarios have been proposed, where the kinetics of transcription are governed by cellular constraints and follow universal rules across the genome. Evidence from genome-wide noise studies and from systematic perturbations of promoter sequences suggest that both scenarios—namely gene-specific versus genome-wide regulation of transcription kinetics—may be present to different degrees in bacteria, yeast and animal cells.

---

The advent of rapid, inexpensive DNA sequencing methods allows scientists to map not only the protein-coding genes in the genomes of many organisms, but also the regulatory sequences present in those genomes. A key challenge for biologists in the next few decades is understanding how these regulatory sequences control the expression of every gene in the cell, and how they collectively determine the topology and dynamics of gene regulatory networks.

The regulation of gene expression has been traditionally studied in experiments that measured the average gene expression level in populations containing millions of cells. These studies relate the average rate of gene expression for a gene to its regulatory DNA sequence (the promoter architecture) (1). This approach has a major shortcoming, however, because averaging over populations masks differences in gene expression that may occur between individual cells (2). These differences may in turn have consequences for the whole

---

\*To whom correspondence should be addressed. [sanchez@rowland.harvard.edu](mailto:sanchez@rowland.harvard.edu), [golding@bcm.edu](mailto:golding@bcm.edu).

The authors declare no conflicts of interest.

multi-cellular community or organism, which makes it important to understand gene expression in single cells.

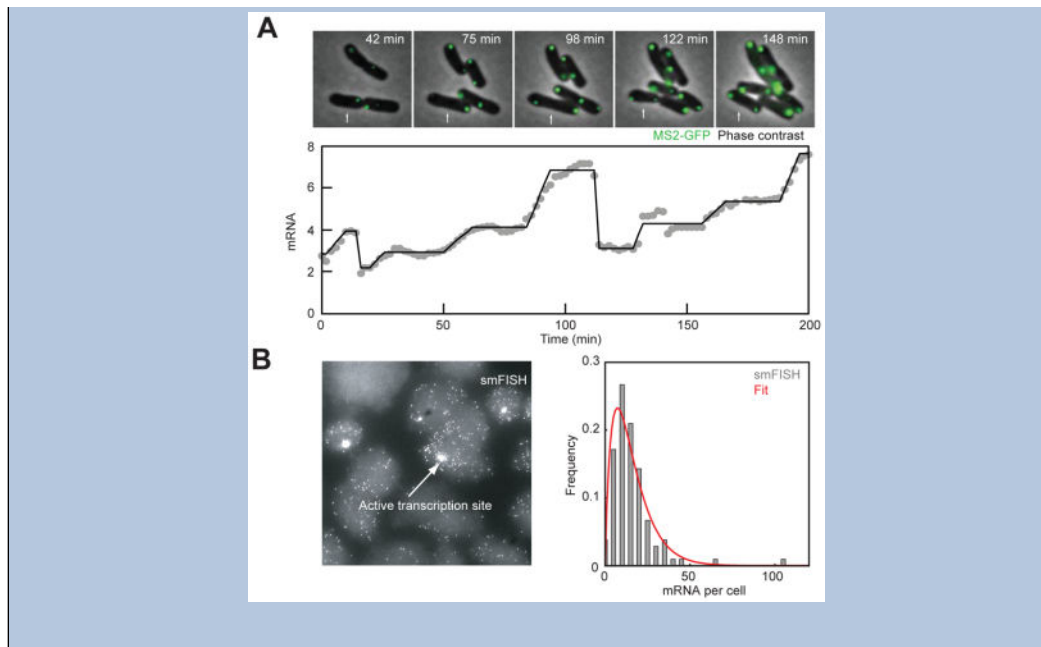
Within a single cell, gene expression is inherently stochastic, or random (2). Protein-coding genes are typically present in only one or two copies per cell. Whether a gene is transcribed at any given moment depends on the arrival, by diffusion, of multiple regulatory proteins to their designated binding sites, as well as the occurrence of multiple biochemical steps required for initiation of transcription (3). These biochemical reactions are all essentially single-molecule events and thus stochastic, resulting in significant randomness in the production of mRNA. Broadly, two stochastic kinetic modes of transcription have been observed in individual cells: “Poissonian”, in which mRNAs are synthesized in random, uncorrelated events, with a probability that is uniform over time (4, 5); and “bursty”, where mRNA is produced in episodes of high transcriptional activity (bursts) followed by long periods of inactivity (See Figure 1 and Box 1) (5–8). The kinetic features of mRNA production are in turn propagated to the proteins translated from them. The end result is temporal fluctuations, and corresponding cell-to-cell variability, in mRNA and protein numbers. This cell-to-cell variability is referred to as gene expression noise (2).

### BOX 1

#### Methods for probing gene expression at the single-cell level

Transcription can be followed in real time in live cells by labeling nascent mRNA with fluorescently tagged RNA-binding proteins that are strongly expressed in the cell (4–6) (Panel A; transcriptional kinetics in *E. coli*, L. So and I. Golding, adapted with permission from *Physical Biology of the Cell*, R. Phillips, J. Kondev and J. Theriot; Garland Science).

Distributions of the numbers of mRNA molecules per cell can be measured by single-mRNA counting techniques, notably single-molecule *fluorescent in situ hybridization* (smFISH) applied to fixed cells (8, 27, 35) (Panel B; mRNA distribution from animal cells (8)). These mRNA distributions carry the signature of the transcriptional kinetics.



Bursty transcription typically leads to higher noise than Poissonian transcription. In particular, the burst size controls the magnitude of the noise, and it is approximately proportional to the Fano factor (the ratio between the variance and the mean of the distribution of mRNA or protein numbers per cell). Thus, genes with large burst sizes are characterized by broader distributions of protein and mRNA, larger Fano factor and higher noise compared to genes with small burst sizes.

The inherent randomness associated with gene expression raises an important question: Are the stochastic kinetics of transcription—and therefore the resulting variability in mRNA and protein levels—encoded by the promoter regulatory sequence, just as the mean expression level of a gene appears to be? Two alternative answers have been put forward. One view is that the stochastic kinetics of gene activity are genetically determined by the promoter architecture, and governed by the binding and unbinding of various regulatory elements (histones, transcription factors) to their corresponding binding sites (9–15). In this view, it is the process of gene regulation, as it acts on each promoter individually, what causes bursting in some genes but not in others.

An alternative view is that transcription kinetics are dominated by genome-wide constraints that lead to general, as opposed to gene-specific, modulation of transcriptional kinetics (16–19). These constraints may reflect any number of physiological or biophysical mechanisms. Proposed mechanisms include cell-cycle-dependent regulation of promoter activity (20) as well as inherent features of the transcription process, such as the cooperative recruitment of RNA polymerases (21). Notwithstanding the specific details, this view implies that gene-specific transcriptional regulation acts on top of these gene-nonspecific constraints and has only a secondary effect on the observed kinetic features, such as transcription bursts.

In the first view described above, transcriptional kinetics (in particular, bursting) are gene-specific and free from global constraints. One consequence of this is that at both very low

and very high rates of transcription promoter activity is expected to be regular in time and well-described by Poisson statistics (Figure 1A). At intermediate rates of transcription, when genes are neither expressed at full capacity nor very infrequently, different genes may vary greatly in their temporal kinetics and exhibit either regular (Poissonian) or bursty behavior, depending on the particular mechanisms of transcriptional regulation for each gene (Figure 1A).

In contrast, global constraints on transcriptional kinetics that affect all genes (Figure 1B) result in a more limited space of possible kinetic behaviors. For instance, one such constraint that has been recently reported (16–18) is the presence of inherent and global bursting kinetics even in fully active gene loci. Thus, highly expressed genes are not Poissonian, but instead characterized by large burst sizes (Figure 1B) (16, 17) and therefore large cell-to-cell variability in mRNA and protein expression. Since this constraint operates globally, all genes in the cell follow a characteristic trend between the burst size and the mean amount of expression (Figure 1B).

The two views described above represent two limiting cases; it is possible that single-cell transcriptional kinetics are affected, to varying degrees, by both gene-specific (i.e. promoter architecture) and genome-wide processes. This review critically examines both views and the evidence supporting each of them. An organism for which there is strong evidence for a genetic origin of transcriptional noise is the yeast *S. cerevisiae*. In the prokaryote *E. coli*, as well as other eukaryotic systems, the picture is less definitive than in yeast but some evidence exists pointing to the presence of global constraints.

## Global studies of gene expression noise in yeast reveal no constraints and point to gene-specific transcriptional kinetics

The relation between noise and the mean expression level has been examined in several studies that measured gene expression at the single-cell level for a complete yeast genomic library (22) or a large set of promoters (14, 23, 24). As noted above, it is possible to estimate the burst size from the degree of cell-to-cell variability in protein concentration (Supplementary Materials). The results of all of these experiments are consistent in that they find no obvious trend between the mean expression level from a promoter and the estimated burst size (Figure 2A). The only global constraint observed in the noise measurements is the one corresponding to the limiting case of Poissonian transcription, followed by Poissonian translation of mRNA into protein (2).

These global noise experiments also reveal relationships between noise measured for a given promoter and the known properties of that promoter. The majority of low-noise promoters have a characteristic architecture (Depleted Proximal Nucleosome or DPN) that is defined by a nucleosome-free region immediately downstream of the initiation site. In contrast, the majority of high-noise promoters have a second type of architecture (Occupied Proximal Nucleosome or OPN), characterized by the lack of a nucleosome-free region. They are also enriched for strong TATA boxes (25, 26). The model emerging from these findings is that promoter switching between inactive (promoter occluded by nucleosome) and active (nucleosome free, pre-initiation complex formed) states may result in bursty transcription, in

turn leading to the higher noise observed in nucleosome-covered promoters. Within this picture, the strong TATA box would ensure that the promoter is expressed strongly when active, thereby increasing the burst size (11).

This model is further supported by single-mRNA counting experiments in fixed cells (27, 28) (Figure 3A). These studies examined the mRNA distributions for twelve different constitutive promoters that have low nucleosome density and were characterized as “low noise” in genomic studies. All of these promoters exhibited close-to-Poissonian mRNA copy-number distributions and sub-Poissonian (i.e. with lower variance than a Poissonian distribution of the same mean) nascent RNA copy number distributions, consistent with the absence of bursts. In contrast, promoters that are not constitutive, but regulated by nucleosomes and transcription factors, exhibited broader mRNA distributions consistent with bursty transcription (Figure 3A). Importantly, these single-cell mRNA counting experiments substantiated the notion that the observed differences in *protein* noise between promoters (as in the studies above (22–24)) reflect the fluctuations of the mRNA species (which in turn are driven by transcriptional kinetics). The mRNA-counting studies specifically predicted that in the low-noise, nucleosome depleted promoters, transcription events should occur regularly in time, in a simple Poissonian manner. This prediction was confirmed in experiments in which the kinetics of transcription were followed in real-time, using temporal correlation analysis of fluorescently labeled mRNA molecules in live yeast cells (4). Both the constitutive promoter *MDN1* and the cell-cycle activated promoter *POLI* were transcribed in random, uncorrelated events with a single rate of initiation that varied during the cell cycle (4)

The evidence above indicates that the kinetics of transcription for a yeast promoter are mainly encoded by the DNA sequence of the promoter. To substantiate this picture, investigators deliberately altered yeast promoter architecture and examined the resulting change in gene expression noise. The systematic alterations included the presence or absence of a TATA box and its strength (12, 24); the number (15, 29), location (15) and nucleosomal coverage (30) of transcription-factor binding sites; the presence of nucleosome disfavoring sequences (13); and the mode of action of a transcription factor (i.e. whether it was acting as an activator or as a repressor (14)). All of these architectural elements were found to strongly affect the relationship between the mean amount of expression and the burst size, in a way that is consistent with the expectation from simple models of transcriptional kinetics (10) (Figure 3).

In summary, the evidence indicates that the stochastic transcriptional kinetics for a given gene in yeast is mainly determined by its promoter architecture and no strong global constraints are observed. Promoter switching introduced by transcription factors and nucleosomes stochastically associating and dissociating leads to a bursty transcription and correspondingly higher degree of noise than that observed in constitutive promoters (11). Additional promoter features such as the strength (12, 13) and copy number (15, 29) of transcription-factor binding sites, their location within the promoter (15), and the specific mechanism of gene regulation by the transcription factors bound to them (14), all affect noise in gene expression.

## Evidence for global noise constraints in *E. coli*

A small number of studies have directly examined transcriptional kinetics in *E. coli*. Live-cell mRNA labeling (Box 1) has been used to visualize the synthesis of individual mRNA molecules in real-time from a synthetic *E. coli* promoter (6, 17). Analysis of how adding an inducer of gene expression affected the bursting parameters revealed that the burst frequency was modulated at low concentration of inducer, whereas the burst size was modulated at high concentration of inducer (17) (Figure 2B, panel iii). Transcription was bursty even for fully induced conditions (6). The  $P_{lac/ara}$  promoter investigated in this study is a derivative of the *lac* promoter ( $P_{lac}$ ), whose mechanism of regulation by *lac* repressor is thought to be well understood (1). However, simple theoretical models, based on these well-characterized mechanisms of gene regulation, failed to explain the observed effect of inducer on the bursting parameters (9). The transcriptional kinetics of the wild-type *lac* promoter have also been investigated as reflected in protein synthesis (31). These studies also indicate an increase in transcriptional burst size at high inducer concentration (31).

A predicted consequence of the observed modulation of the bursting parameters described above is that the Fano factor of the mRNA distribution should increase as the mean increases. This was observed for the *lac* promoter in the presence of various concentrations of inducer (17). Six other promoters were also studied under diverse conditions. Unexpectedly, it was found that all of them yielded values of the Fano factor that were very similar to those observed for the *lac* promoter (17) (Figure 2B, panel i). The authors interpreted this finding as indicating that the promoters investigated all modulate their bursting parameters in a similar fashion: modulation of the burst frequency when expression is low and of the burst size when expression is higher.

An examination of the protein copy-number distribution for a genomic library in *E. coli* (32) found that the Fano factor did indeed increase with the mean level of gene expression, in a pattern very reminiscent of the RNA data described above (17) (Figure 2B, panels i-ii), and consistent with the earlier protein distributions measured for  $P_{lac}$  (31). The authors offered an alternative interpretation for the observed relationship between the Fano factor and the mean protein concentration: that the increase in the Fano factor does not result from transcription bursts of increased size, but rather reflects the dominance of extrinsic noise (that caused by cell-to-cell differences in gene expression parameters) at high rates of expression. In support of this hypothesis, measurements of extrinsic noise coincided with the noise baseline observed at high rates of expression (32).

The difficulty to discriminate between alternative models on the basis of the relationship between the noise and the mean amount of expression underscores the potential pitfalls associated with over-interpreting static copy-number distributions. In addition to extrinsic noise, other effects may be confounded with stochasticity in transcription and translation. For instance, the statistics of protein and mRNA partitioning during cell division may in some cases lead to noise-mean scaling that is indistinguishable from that of stochastic transcription and translation (33). Effects due to the cell cycle may also be mistakenly attributed to stochastic transcription (20).

To achieve a mechanistic understanding of how promoter architecture affects transcriptional kinetics and noise in *E. coli*, the path previously taken in yeast will have to be emulated, by measuring gene expression noise in promoters whose architecture is systematically perturbed (12–15, 29). Such noise studies need to be accompanied by direct measurements of transcription kinetics in live cells.

## The case of mammalian cells and other eukaryotes

The studies reviewed above support the idea that constitutive genes in budding yeast are expressed in a Poissonian, non-bursty manner. However, it would be wrong to assume that this result is typical of all eukaryotic cells. Transcriptional kinetics of multiple genes in *Dictyostelium discoideum* (another single-celled eukaryotic organism), including housekeeping genes, were found to be bursty (7). Beyond protozoans, bursting kinetics seem to be the rule, rather than the exception, in many different types of animal cells, including cultured mammalian cells (5, 8, 16, 34), fly embryos (18, 35, 36), and in the mouse (37).

Real-time transcriptional kinetics of both natural and synthetic promoters inserted into mouse fibroblasts (34) revealed transcriptional bursting from these promoters, and a refractory period after the bursts. Synthetic promoters engineered to differ in the number of binding sites for a transcriptional activator or their binding strength showed that both of those architectural features can affect the burst size. This might be interpreted as evidence that promoter architecture can allow decoupling the mean amount of expression from the noise. However, when the bursting parameters were measured for these different promoters as a function of their mean level of expression, a single trend line was followed closely by all of the promoters regardless of their regulatory sequence (Figure 2C, panel i).

The existence of a non-trivial trend line between the noise and the mean amount of protein in the cell was also observed in a study analyzing single-cell gene expression from a weak viral promoter randomly inserted at over 8,000 genomic loci in a line of human T lymphocytes (16). Evidence of bursting was found across the whole set of loci where the lentiviral reporter was able to integrate. Similar to the modulation of bursting discussed above for *E. coli* (17), the data were consistent with modulation of burst frequency at low expression, and of burst size at high expression (Figure 2C, panel ii). A similar study of HIV promoters integrated at random locations in the genome of T-cells also showed a trend-line between the burst size and the mean expression level (Figure 2C, panel iii) (38).

## Are there cellular constraints that impose bursting?

The widespread observation of transcriptional bursting from bacteria to animal cells has prompted the idea that bursting may be a beneficial trait, possibly allowing an optimal allocation of resources or better processing of information from the environment (17). Bursting might also be an unavoidable kinetic feature of transcription, reflecting biophysical constraints that apply to most genes within a given organism (19, 21, 39).

But are transcription bursts really unavoidable? A good test case is the behavior of promoters acting at full capacity. In the absence of any constraints, such promoters are expected to exhibit Poissonian, non-bursty kinetics (Figure 1). In *E. coli*, ribosomal



promoters are transcribed at an average rate that is very close to the upper boundary imposed by RNAP elongation (40). This suggests that they can be expressed continuously and without large bursts at full capacity. On the other hand, the hypothesis that biophysical or cellular constraints impose bursting even in highly expressed promoters is consistent with the finding that fully induced mRNA-coding promoters display strong bursting (6) and highly non-Poissonian mRNA distributions (17). More experiments, where the architecture of a promoter is systematically varied, are needed for elucidating the mechanisms governing transcription kinetics in bacteria, in particular whether bursting is governed by universal or gene-specific constraints.

As for animal cells, bursting is prevalent in endogenous promoters (5, 8, 18, 34–36), but a few counter examples also exist. A highly expressed viral promoter inserted in mammalian cells was found to be expressed continuously in a burst-free manner (5). RNA counting in *Drosophila* embryos also suggests that highly expressed genes can act close to full transcriptional capacity and exhibit nearly Poissonian statistics (36) (but see also (18)). These counter-examples appear to indicate that bursting is not an unavoidable by-product of transcription in all animal cells. Only a limited number of promoters, many of which are synthetic, have been studied in animal cells. Moreover, most of experiments have been done *in vitro* with cell lines. In their natural context, animal cells are arranged in a complex multicellular environment, which may affect gene expression. Experimental advances that make it possible to count individual mRNA molecules in fixed animals or embryos (18, 35–37) offer a promising avenue to better understand transcriptional kinetics in multicellular environments.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We apologize to our colleagues whose work could not be cited due to space limitations. We thank the following for sharing data and insights with us: A. Bar-Even, N. Barkai, A. Boettiger, M. Dadiani, A. Depace, T. Gregor, J. Kondev, D. Larson, R. Philips, A. Raj, E. Segal and D. Zenklusen. IG thanks his lab members for help preparing the manuscript. Work in the Sanchez lab is supported by a Junior Fellowship from the Rowland Institute at Harvard. Work in the Golding lab is supported by the US National Institutes of Health grant R01 GM082837, US National Science Foundation grants 082265 (Physics Frontiers Center: Center for the Physics of Living Cells) and PHY-1147498 (CAREER), and Welch Foundation grant Q-1759.

## GLOSSARY (BOX 2)

<b>Transcription bursts</b>	The production of multiple mRNAs within a short time, followed by a period of promoter inactivity
<b>Gene expression noise</b>	Variability in the level of gene expression between genetically-identical cells, due in part to random fluctuations in the production of mRNA and protein

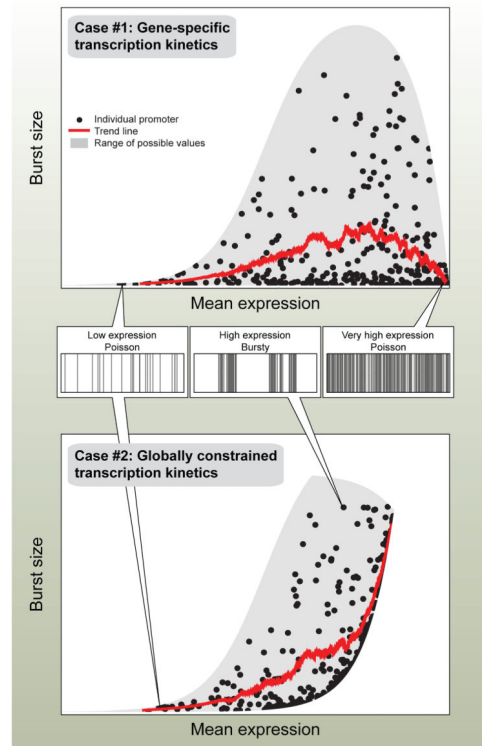


<b>Fano factor</b>	The ratio between the variance (standard deviation squared) and the mean of a measured quantity. The Fano factor of mRNA number per cell can serve as an estimate for the size of transcription bursts
<b>Poisson process</b>	The simplest random process, in which events occur with a constant probability over time
<b>Promoter</b>	A region of DNA that controls transcription from a gene. An important element of eukaryotic promoters is the TATA box, which strongly affects the strength of the promoter

## References

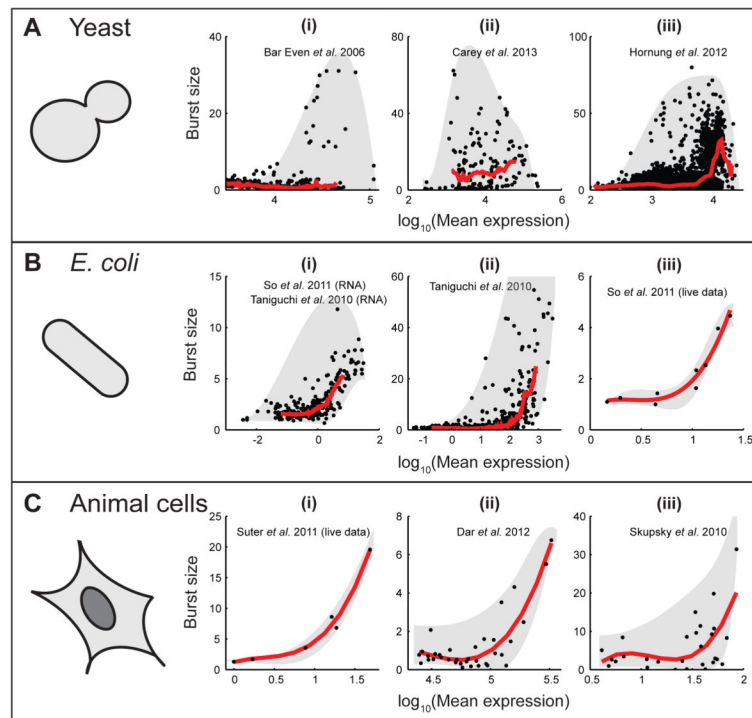
- Bintu L, et al. Transcriptional regulation by the numbers: applications. *Curr Opin Genet Dev.* 2005; 15:125–135. [PubMed: 15797195]
- Raj A, van Oudenaarden A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell.* 2008; 135:216–226. [PubMed: 18957198]
- Coulon A, Chow CC, Singer RH, Larson DR. Eukaryotic transcriptional dynamics: from single molecules to cell populations. *Nat Rev Genet.* 2013; 14:572–584. [PubMed: 23835438]
- Larson DR, Zenklusen D, Wu B, Chao JA, Singer RH. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science.* 2011; 332:475–478. [PubMed: 21512033]
- Yunger S, Rosenfeld L, Garini Y, Shav-Tal Y. Single-allele analysis of transcription kinetics in living mammalian cells. *Nat Methods.* 2010; 7:631–633. [PubMed: 20639867]
- Golding I, Paulsson J, Zawilski SM, Cox EC. Real-time kinetics of gene activity in individual bacteria. *Cell.* 2005; 123:1025–1036. [PubMed: 16360033]
- Muramoto T, et al. Live imaging of nascent RNA dynamics reveals distinct types of transcriptional pulse regulation. *Proc Natl Acad Sci U S A.* 2012; 109:7350–7355. [PubMed: 22529358]
- Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* 2006; 4:e309. [PubMed: 17048983]
- Sanchez A, Garcia HG, Jones D, Phillips R, Kondev J. Effect of promoter architecture on the cell-to-cell variability in gene expression. *PLoS Comput Biol.* 2011; 7:e1001100. [PubMed: 21390269]
- Sanchez A, Choubey S, Kondev J. Regulation of noise in gene expression. *Annu Rev Biophys.* 2013; 42:469–491. [PubMed: 23527780]
- Raser JM, O’Shea EK. Control of stochasticity in eukaryotic gene expression. *Science.* 2004; 304:1811–1814. [PubMed: 15166317]
- Blake WJ, et al. Phenotypic consequences of promoter-mediated transcriptional noise. *Mol Cell.* 2006; 24:853–865. [PubMed: 17189188]
- Dadiani M, et al. Two DNA-encoded strategies for increasing expression with opposing effects on promoter dynamics and transcriptional noise. *Genome Res.* 2013; 23:966–976. [PubMed: 23403035]
- Carey LB, van Dijk D, Sloot PMA, Kaandorp JA, Segal E. Promoter sequence determines the relationship between expression level and noise. *PLoS Biol.* 2013; 11:e1001528. [PubMed: 23565060]
- Murphy KF, Adams RM, Wang X, Balázsi G, Collins JJ. Tuning and controlling gene expression noise in synthetic gene networks. *Nucleic Acids Res.* 2010; 38:2712–2726. [PubMed: 20211838]
- Dar RD, et al. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proc Natl Acad Sci U S A.* 2012; 109:17454–17459. [PubMed: 23064634]
- So LH, et al. General properties of transcriptional time series in *Escherichia coli*. *Nat Genet.* 2011; 43:554–560. [PubMed: 21532574]

18. Little S, Tikhonov M, Gregor T. Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell*. 2013 In Press.
19. Salman H, et al. Universal protein fluctuations in populations of microorganisms. *Phys Rev Lett*. 2012; 108:238105. [PubMed: 23003996]
20. Zopf CJ, Quinn K, Zeidman J, Maheshri N. Cell-Cycle Dependence of Transcription Dominates Noise in Gene Expression. *PLoS Comput Biol*. 2013; 9:e1003161. [PubMed: 23935476]
21. Guptasarma P. Cooperative relaxation of supercoils and periodic transcriptional initiation within polymerase batteries. *BioEssays News Rev Mol Cell Dev Biol*. 1996; 18:325–332.
22. Newman JRS, et al. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature*. 2006; 441:840–846. [PubMed: 16699522]
23. Bar-Even A, et al. Noise in protein expression scales with natural protein abundance. *Nat Genet*. 2006; 38:636–643. [PubMed: 16715097]
24. Hornung G, et al. Noise-mean relationship in mutated promoters. *Genome Res*. 2012; 22:2409–2417. [PubMed: 22820945]
25. Tirosh I, Barkai N. Two strategies for gene regulation by promoter nucleosomes. *Genome Res*. 2008; 18:1084–1091. [PubMed: 18448704]
26. Field Y, et al. Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput Biol*. 2008; 4:e1000216. [PubMed: 18989395]
27. Zenklusen D, Larson DR, Singer RH. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat Struct Mol Biol*. 2008; 15:1263–1271. [PubMed: 19011635]
28. Gandhi SJ, Zenklusen D, Lionnet T, Singer RH. Transcription of functionally related constitutive genes is not coordinated. *Nat Struct Mol Biol*. 2011; 18:27–34. [PubMed: 21131977]
29. To TL, Maheshri N. Noise can induce bimodality in positive transcriptional feedback loops without bistability. *Science*. 2010; 327:1142–1145. [PubMed: 20185727]
30. Bai L, Charvin G, Siggia ED, Cross FR. Nucleosome-depleted regions in cell-cycle-regulated promoters ensure reliable gene expression in every cell cycle. *Dev Cell*. 2010; 18:544–555. [PubMed: 20412770]
31. Choi PJ, Cai L, Frieda K, Xie XS. A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science*. 2008; 322:442–446. [PubMed: 18927393]
32. Taniguchi Y, et al. Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science*. 2010; 329:533–538. [PubMed: 20671182]
33. Huh D, Paulsson J. Non-genetic heterogeneity from stochastic partitioning at cell division. *Nat Genet*. 2011; 43:95–100. [PubMed: 21186354]
34. Suter DM, et al. Mammalian genes are transcribed with widely different bursting kinetics. *Science*. 2011; 332:472–474. [PubMed: 21415320]
35. Paré A, et al. Visualization of individual *Scr* mRNAs during *Drosophila* embryogenesis yields evidence for transcriptional bursting. *Curr Biol CB*. 2009; 19:2037–2042.
36. Boettiger AN, Levine M. Rapid transcription fosters coordinate snail expression in the *Drosophila* embryo. *Cell Reports*. 2013; 3:8–15. [PubMed: 23352665]
37. Itzkovitz S, et al. Single-molecule transcript counting of stem-cell markers in the mouse intestine. *Nat Cell Biol*. 2012; 14:106–114. [PubMed: 22119784]
38. Skupsky R, Burnett JC, Foley JE, Schaffer DV, Arkin AP. HIV promoter integration site primarily modulates transcriptional burst size rather than frequency. *PLoS Comput Biol*. 2010; 6:1000952
39. Bremer H, Ehrenberg M. Guanosine tetraphosphate as a global regulator of bacterial RNA synthesis: a model involving RNA polymerase pausing and queuing. *Biochim Biophys Acta*. 1995; 1262:15–36. [PubMed: 7539631]
40. Liang S, et al. Activities of constitutive promoters in *Escherichia coli*. *J Mol Biol*. 1999; 292:19–37. [PubMed: 10493854]



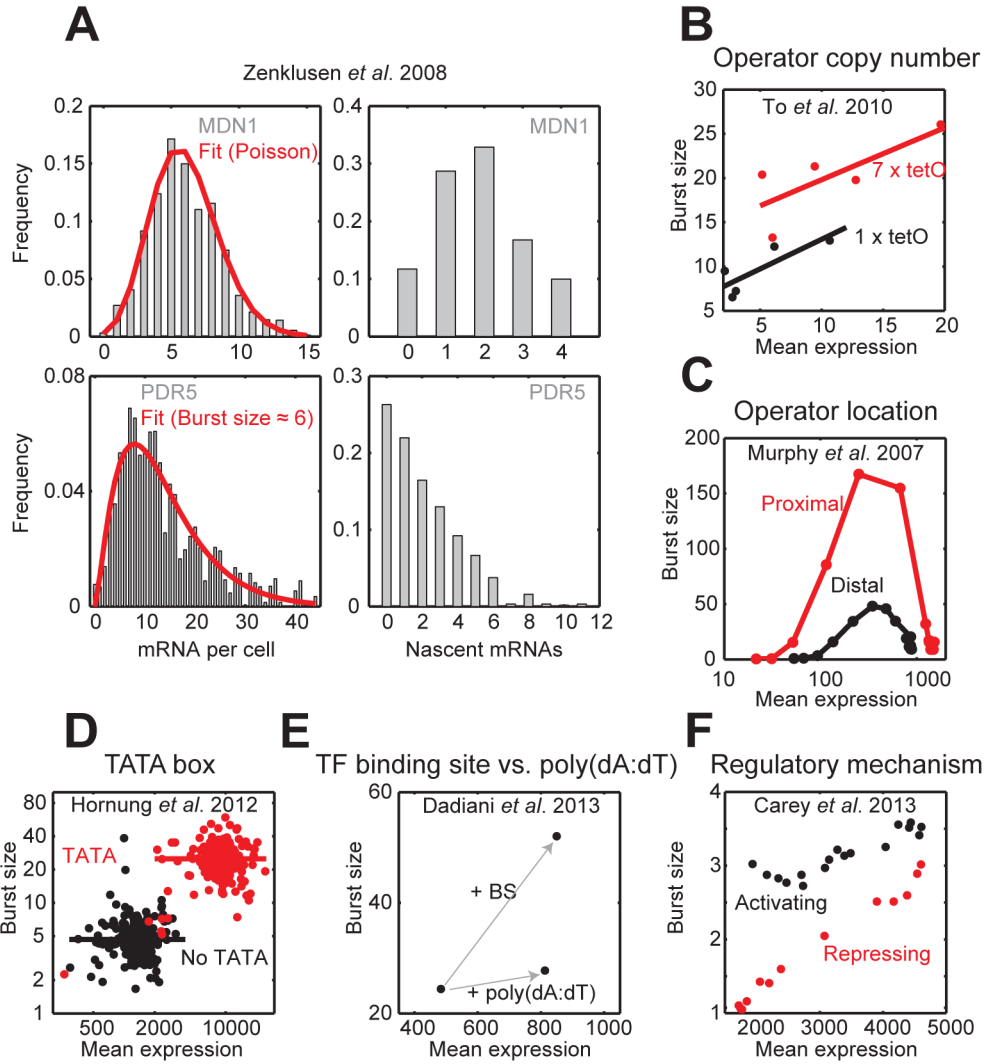
**Figure 1. Gene-specific versus global determinants of transcription kinetics**

Two alternative scenarios are presented, representing different ways in which stochastic promoter activity may be governed. In the first scenario (top), the rates of promoter activation and inactivation are controlled exclusively by gene-specific mechanisms. In the second scenario (bottom), gene-specific regulation occurs in the presence of a global constraint on the kinetics. We illustrate the consequences of the two scenarios using simple mathematical “toy models” of promoter kinetics (see Supplementary Materials). For each model, we calculate both the mean expression level and the size of transcription bursts for a large set of promoters and plot these two variables against each other. In the plot, black markers denote 300 randomly-chosen individual promoters, the red curve represents the smoothed behavior of  $10^4$  promoters, and the shaded region is a guide-to-the-eye that depicts the full range of possible burst sizes for a given mean expression (see Supplementary Materials).



**Figure 2. Evidence for genome-wide constraints on promoter kinetics in bacteria and mammalian cells, but not yeast**

Data from three studies in yeast ((14, 23, 24), panel **A**) show no obvious correlation between the mean expression level of a gene and the transcription burst size. In particular, both low-expression and high-expression genes can exhibit non-bursty behavior, consistent with a scenario of gene-specific regulation. In contrast, two studies in *E. coli* ((17, 32), panel **B**) and three in mammalian cells ((16, 34, 38), panel **C**) show that higher expression is accompanied by increased burstiness, consistent with the presence of a global cellular constraint on promoter kinetics. Excluding live-cell measurements (panels **B**-iii and **C**-i), the burst sizes were estimated using the Fano factor of the corresponding distribution (protein measurements, excluding panel **B**-i), after correcting for the level of extrinsic noise. Black circles designate the individual measurements. Red curves are calculated trend lines. The shaded area highlights the full range of burst sizes covered by each data set. For more details see Supplementary Materials.



**Figure 3. Bursting kinetics in yeast are promoter dependent, and not subject to strong constraints**

(A) Both Poissonian and bursty kinetics are found in endogenous yeast promoters. Copy-number statistics of mature (left) and nascent (right) mRNA are consistent with Poissonian promoter activity for MDN1 (above), while indicating bursty activity of PDR5 (below) (adapted by permission from Macmillan Publishers Ltd: Nature Structural & Molecular Biology (27) copyright 2008). (B–E) Manipulating promoter architecture leads to a different relation between the mean expression and the burst size, such that two promoters with the same mean expression level can exhibit different burst sizes. The burst size is plotted as a function of the mean amount of expression for various perturbations of promoter architecture, specifically: changing the number of operator sites from one to seven (Panel B (29)); changing the position of the operator site from a location proximal to the first transcribed nucleotide, to a more distant location (Panel C (15)); the presence (red dots) or absence of a TATA box (black dots) (Panel D (24)), including spontaneous mutations that delete it (red dots that overlap with the promoters lacking the TATA box); promoters engineered to yield the same mean level of expression by either adding a nucleosome

disfavoring poly dA:dT sequence or by increasing the strength of a transcription factor binding site (+BS) (Panel **E** (13)). The same transcription factor acting as a repressor (red dots) or as an activator (black dots) of a promoter leads to different burst sizes for the same mean level of expression (Panel **F** (14)). For more details see Supplementary Materials.