

RESEARCH ARTICLE

Intercellular Variability in Protein Levels from Stochastic Expression and Noisy Cell Cycle Processes

Mohammad Soltani¹, Cesar A. Vargas-Garcia¹, Duarte Antunes², Abhyudai Singh^{1,3,4,5*}

1 Electrical and Computer Engineering Department, University of Delaware, Newark, Delaware, United States of America, **2** Mechanical Engineering Department, Eindhoven University of Technology, Eindhoven, Netherlands, **3** Biomedical Engineering Department, University of Delaware, Newark, Delaware, United States of America, **4** Mathematical Sciences Department, University of Delaware, Newark, Delaware, United States of America, **5** Center for Bioinformatics and Computational Biology, University of Delaware, Newark, Delaware, United States of America

* absingh@udel.edu



Abstract

Inside individual cells, expression of genes is inherently stochastic and manifests as cell-to-cell variability or noise in protein copy numbers. Since proteins half-lives can be comparable to the cell-cycle length, randomness in cell-division times generates additional intercellular variability in protein levels. Moreover, as many mRNA/protein species are expressed at low-copy numbers, errors incurred in partitioning of molecules between two daughter cells are significant. We derive analytical formulas for the total noise in protein levels when the cell-cycle duration follows a general class of probability distributions. Using a novel hybrid approach the total noise is decomposed into components arising from i) stochastic expression; ii) partitioning errors at the time of cell division and iii) random cell-division events. These formulas reveal that random cell-division times not only generate additional extrinsic noise, but also critically affect the mean protein copy numbers and intrinsic noise components. Counter intuitively, in some parameter regimes, noise in protein levels can decrease as cell-division times become more stochastic. Computations are extended to consider genome duplication, where transcription rate is increased at a random point in the cell cycle. We systematically investigate how the timing of genome duplication influences different protein noise components. Intriguingly, results show that noise contribution from stochastic expression is minimized at an optimal genome-duplication time. Our theoretical results motivate new experimental methods for decomposing protein noise levels from synchronized and asynchronized single-cell expression data. Characterizing the contributions of individual noise mechanisms will lead to precise estimates of gene expression parameters and techniques for altering stochasticity to change phenotype of individual cells.

OPEN ACCESS

Citation: Soltani M, Vargas-Garcia CA, Antunes D, Singh A (2016) Intercellular Variability in Protein Levels from Stochastic Expression and Noisy Cell Cycle Processes. *PLoS Comput Biol* 12(8): e1004972. doi:10.1371/journal.pcbi.1004972

Editor: Suckjoon Jun, UCSD, UNITED STATES

Received: September 8, 2015

Accepted: July 29, 2016

Published: August 18, 2016

Copyright: © 2016 Soltani et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work is supported by the National Science Foundation Grant DMS-1312926. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Author Summary

Inside individual cells, gene products often occur at low molecular counts and are subject to considerable stochastic fluctuations (noise) in copy numbers over time. An important consequence of noisy expression is that the level of a protein can vary considerably even among genetically identical cells exposed to the same environment. Such non-genetic phenotypic heterogeneity is physiologically relevant and critically influences diverse cellular processes. In addition to noise sources inherent in gene product synthesis, recent experimental studies have uncovered additional noise mechanisms that critically effect expression. For example, the time within the cell cycle when a gene duplicates, and the time taken to complete cell cycle are governed by random processes. The key contribution of this work is development of novel mathematical results quantifying how cell cycle-related noise sources combine with stochastic expression to drive intercellular variability in protein molecular counts. Derived formulas lead to many counterintuitive results, such as increasing randomness in the timing of cell division can lower noise in the level of a protein. Finally, these results inform experimental strategies to systematically dissect the contributions of different noise sources in the expression of a gene of interest.

Introduction

The level of a protein can deviate considerably from cell-to-cell, in spite of the fact that cells are genetically-identical and are in the same extracellular environment [1–3]. This intercellular variation or noise in protein counts has been implicated in diverse processes such as corrupting functioning of gene networks [4–6], driving probabilistic cell-fate decisions [7–12], buffering cell populations from hostile changes in the environment [13–16], and causing clonal cells to respond differently to the same stimulus [17–19]. An important source of noise driving random fluctuations in protein levels is stochastic gene expression due to the inherent probabilistic nature of biochemical processes [20–23]. Recent experimental studies have uncovered additional noise sources that affect protein copy numbers. For example, the time take to complete cell cycle (i.e., time between two successive cell-division events) has been observed to be stochastic across organisms [24–32]. Moreover, given that many proteins/mRNAs are present inside cells at low-copy numbers, errors incurred in partitioning of molecules between two daughter cells are significant [33–35]. Finally, the time at which a particular gene of interest is duplicated can also vary between cells [36, 37]. We investigate how such noise sources in the cell-cycle process combine with stochastic gene expression to generate intercellular variability in protein copy numbers (Fig 1).

Prior studies that quantify the effects of cell division on the protein noise level have been restricted to specific cases. For example, noise computations have been done in stochastic gene expression models, where cell divisions occur at deterministic time intervals [33, 38, 39]. Recently, we have analyzed a deterministic model of gene expression with random cell-division events [40]. Building up on this work, we formulate a mathematical model that couples stochastic expression of a stable protein with random cell-division events that follow a general class of probability distributions. Moreover, at the time of cell division, proteins are randomly partitioned between two daughter cells based on a framework that allows the partitioning errors to be higher or lower than as predicted by binomial partitioning. For this class of models, we derive an exact analytical formula for the protein noise level as quantified by the steady-state squared Coefficient of Variation (CV^2). This formula is further decomposed into individual components representing contributions from different noise sources. A systematic

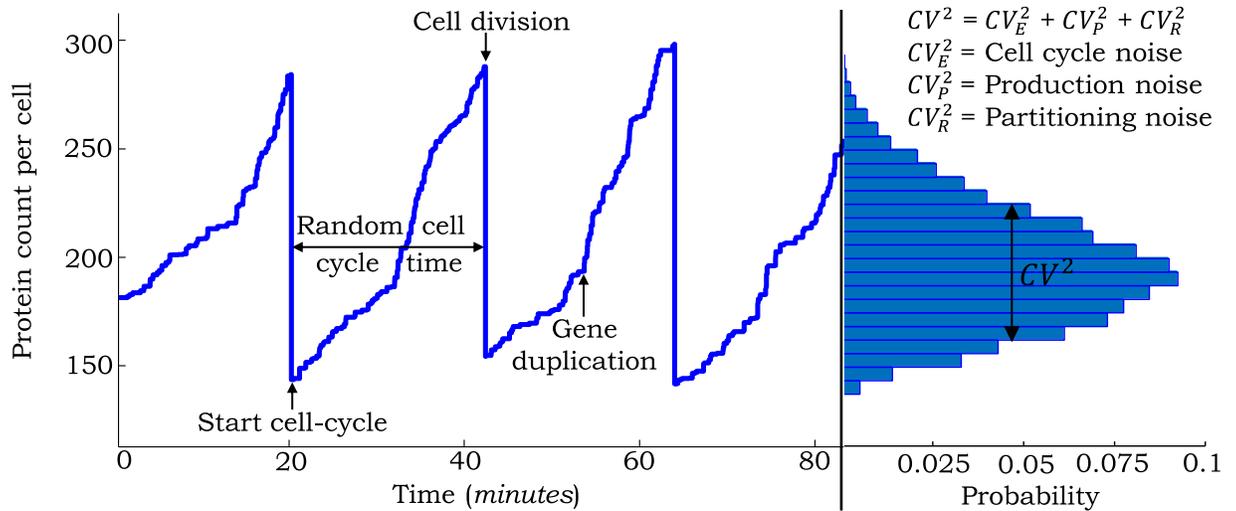


Fig 1. Sample trajectory of the protein level in a single cell with different sources of noise. Stochastically expressed proteins accumulate within the cell at a certain rate. At a random point in the cell cycle, gene duplication results in an increase in production rate. Stochastic cell-division events lead to random partitioning of protein molecules between two daughter cells with each cell receiving, on average, half the number of proteins in the mother cell just before division. The steady-state protein copy number distribution obtained from a large number of trajectories is shown on the right. The total noise in the protein level, as measured by the squared coefficient of variation (CV^2) can be broken into contributions from individual noise mechanisms.

doi:10.1371/journal.pcbi.1004972.g001

investigation of this formula leads to novel insights, such as identification of regimes where increasing randomness in the timing of cell-division events decreases the protein noise level.

Next, we extend the above model to include genome-duplication events that increase the gene’s transcription rate [36, 41]. To our knowledge, this is the first study integrating randomness in the genome-duplication process with stochastic gene expression. An exact formula for the protein noise level is derived for this extended model and used to investigate how the timing of duplication affects different noise components. Counter intuitively, results show that doubling of the transcription rate within the cell cycle can lead to smaller fluctuations in protein levels as compared to a constant transcription rate through out the cell cycle. Finally, we discuss how formulas obtained in this study can be used to infer parameters and characterize the gene expression process from single-cell studies.

Methods

Coupling gene expression to cell division

We consider the standard model of stochastic gene expression [42, 43], where mRNAs are transcribed at exponentially distributed time intervals from a constitutive gene with rate k_x . For the time being, we exclude genome duplication and the transcription rate is fixed throughout the cell cycle. Assuming short-lived mRNAs, each transcription event results in a burst of proteins [43–45]. The corresponding jump in protein levels is shown as

$$x(t) \mapsto x(t) + B, \tag{1}$$

where $x(t)$ is the protein population count in the mother cell at time t , B is a random burst size drawn from a positively-valued distribution and represents the number of protein molecules synthesized in a single-mRNA lifetime. Motivated by observations in *E. coli* and mammalian cells, where many proteins have half-lives considerably longer than the cell-doubling time, we assume a stable protein with no active degradation [46–48]. Thus, proteins accumulate within

the cell till the time of cell division, at which point they are randomly partitioned between two daughter cells.

Let cell division events occur at times $t_s, s \in \{1, 2, \dots\}$. The cell-cycle time

$$T := t_s - t_{s-1}, \tag{2}$$

follows an arbitrary positively-valued probability distribution with the following mean and squared coefficient of variation (CV^2)

$$\langle T \rangle = \langle t_s - t_{s-1} \rangle, \quad CV_T^2 = \frac{\langle T^2 \rangle - \langle T \rangle^2}{\langle T \rangle^2}, \tag{3}$$

where $\langle \cdot \rangle$ denotes expected value through out this paper. The random change in $x(t)$ during cell division is given by

$$x(t_s) \mapsto x_+(t_s), \tag{4}$$

where $x(t_s)$ denotes the protein levels in the mother cell just before division and $x_+(t_s)$ denotes the protein levels in one of the daughter cells just after division. Conditioned on $x(t_s)$, $x_+(t_s)$ is assumed to have the following statistics

$$\langle x_+(t_s) | x(t_s) \rangle = \frac{x(t_s)}{2}, \quad \langle x_+^2(t_s) - \langle x_+(t_s) \rangle^2 | x(t_s) \rangle = \frac{\alpha x(t_s)}{4}. \tag{5}$$

The first equation implies symmetric partitioning, i.e., on average each of the daughter cells inherits half the number protein molecules just before division. The second equation in [Eq \(5\)](#) describes the variance of $x_+(t_s)$ and quantifies the error in partitioning of molecules through the non-negative parameter α . For example, $\alpha = 0$ represents deterministic partitioning where $x_+(t_s) = x(t_s)/2$ with probability equal to one. A more realistic model for partitioning is each molecule having an equal probability of being in the each daughter cell [\[49–51\]](#). This results in a binomial distribution for $x_+(t_s)$

$$\text{Probability}\{x_+(t_s) = j | x(t_s)\} = \frac{x(t_s)!}{j!(x(t_s) - j)!} \left(\frac{1}{2}\right)^{x(t_s)}, \quad j \in \{0, 1, \dots, x(t_s)\}, \tag{6}$$

and corresponds to $\alpha = 1$ in [Eq \(5\)](#). Interestingly, recent studies have shown that partitioning of proteins that form clusters or multimers can result in $\alpha > 1$ in [Eq \(5\)](#), i.e., partitioning errors are much higher than as predicted by the binomial distribution [\[33, 39\]](#). In contrast, if molecules push each other to opposite poles of the cell, then the partitioning errors will be smaller than as predicted by [Eq \(6\)](#) and $\alpha < 1$.

The model with all the different noise mechanisms (stochastic expression; random cell-division events and partitioning errors) is illustrated in [Fig 2A](#) and referred to as the full model. We also introduce two additional hybrid models [\[52, 53\]](#), where protein production and partitioning are considered in their deterministic limit ([Fig 2B and 2C](#)). Note that unlike the full model, where $x(t)$ takes non-negative integer values, $x(t)$ is continuous in the hybrid models. We will use these hybrid models for decomposing the protein noise level obtained from the full model into individual components representing contributions from different noise sources.

Modeling the cell-cycle time

In order to quantify the steady-state protein mean and noise, we need to define the stochastic process that governs the timing of cell division. Variations in the duration of cell cycle can result from a variety of factors, such as cell physiology, growth rate, cell size and expression of

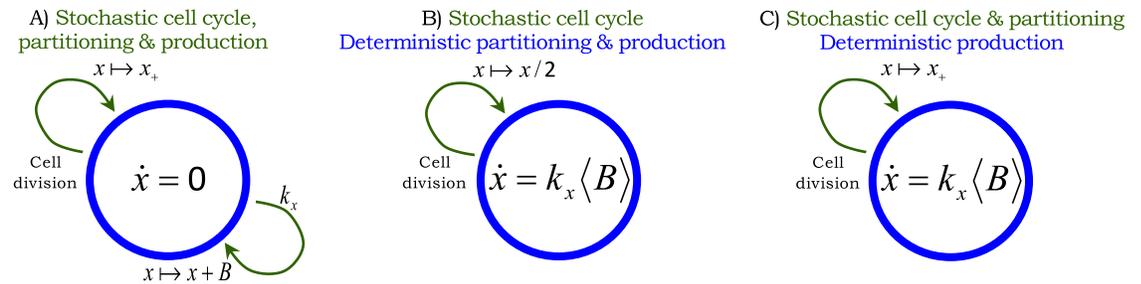


Fig 2. Stochastic models of gene expression with cell division. Arrows denote stochastic events that change the protein level by discrete jumps as shown in Eqs (1) and (4). The differential equation within the circle represents the time evolution of $x(t)$ in between events. **A)** Model with all the different sources of noise: proteins are expressed in stochastic bursts, cell division occurs at random times, and molecules are partitioned between the two daughter cells based on Eq (5). The trivial dynamics $\dot{x} = 0$ signifies that the protein level is constant in-between stochastic events. **B)** Hybrid model where randomness in cell-division events is the only source of noise. Protein production is modeled deterministic through a differential equation and partitioning errors are absent, i.e., $\alpha = 0$ in Eq (5). **C)** Hybrid model where noise comes from both cell-division events and partitioning errors. Protein production is considered to be deterministic as in Fig 2B. Since $x(t)$ is continuous here, $x_+(t_s)$ has a positively-valued continuous distribution with same mean and variance as in Eq (5)

doi:10.1371/journal.pcbi.1004972.g002

genes that affect cell-cycle time such as FtsZ [24–32]. Given these complexities, we take a phenomenological approach to modeling cell-cycle time, and assume it to be an independent and identically distributed random variable that is drawn from a mixture of Erlang distributions (also known as phase-type distribution). The motivation for choosing this distribution is two fold:

1. Mixture of Erlang distributions can be represented via a continuous-time Markov chain, allowing mathematical tractability in terms of deriving/analyzing time evolution of moments.
2. This class of distribution is fairly general, in the sense that, any positively-valued distribution with $CV \leq 1$ can be modeled via a mixture of Erlang distributions [54].

Consider a mixture of n Erlang distributions with mixing probabilities $p_i, i = \{1, \dots, n\}$. Recall that an Erlang distribution of order i is the distribution of the sum of i independent and identical exponential random variables. The cell-cycle time is assumed to have an Erlang distribution of order i with probability p_i and can be represented by a continuous-time Markov chain with states $G_{ij}, j = \{1, \dots, i\}, i = \{1, \dots, n\}$ (Fig 3). Let Bernoulli random variables $g_{ij} = 1$ if the system resides in state G_{ij} and 0 otherwise. The probability of transition $G_{ij} \rightarrow G_{i(j+1)}$ in the next

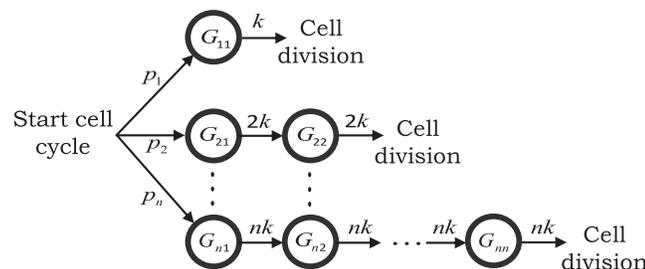


Fig 3. A continuous-time Markov chain model for the cell-cycle time. Left: The cell-cycle time is assumed to follow a mixture of Erlang distributions. At the start of cell cycle, a state $G_{i1}, i = \{1, \dots, n\}$ is chosen with probability p_i . The cell cycle transitions through states $G_{ij}, j = \{1, \dots, i\}$ residing for an exponentially distributed time with mean $1/ik$ in each state. Cell division occurs after exit from G_{ij} and the above process is repeated.

doi:10.1371/journal.pcbi.1004972.g003

infinitesimal time interval $[t, t + dt)$ is given by $ikg_{ij} dt$, implying that the time spent in each state G_{ij} is exponentially distributed with mean $1/ik$. To summarize, at the start of cell cycle, a state G_{i1} , $i = \{1, \dots, n\}$ is chosen with probability p_i and cell division occurs after transitioning through i exponentially distributed steps. Based on this formulation, the probability of a cell-division event occurring and a new cell cycle obtained from an Erlang distribution of size i starting in the next time interval $[t, t + dt)$ is given by $kp_i \sum_{j=1}^n (jg_{ij})dt$, and whenever the event occurs, the protein level changes as per Eq (4). Finally, the mean, the squared coefficient of variation, and the skewness of the cell-cycle time in terms of the Markov chain parameters are given by

$$\langle T \rangle = \frac{1}{k}, \quad CV_T^2 = \sum_{i=1}^n \frac{p_i}{i}, \quad \text{Skewness} = \frac{\langle T^3 \rangle - 3\langle T \rangle(\langle T^2 \rangle - \langle T \rangle^2) - \langle T \rangle^3}{(\langle T^2 \rangle - \langle T \rangle^2)^{3/2}} = 2 \sum_{i=1}^n \frac{p_i}{i^2} \quad (7)$$

[55], where $\langle T^3 \rangle$ is the third-order moment of the cell-cycle time. An important property of this class of distributions is that increasing CV_T^2 also makes the distribution highly skewed, because from Eq (7) both the CV and skewness are linear combinations of p_i , albeit with different linear coefficients that decrease with i . Considering that $\sum_{i=1}^n p_i = 1$, the only way to increase CV_T^2 is by increasing smaller-index components and decreasing larger-index components of the distribution (i.e. increasing p_i and decreasing p_j , where $i < j$). Since higher values of i are more penalized in the skewness equation, this would correspond to making the distribution more positively skewed. Hence high values of CV_T^2 also means high values of skewness, thus occurrences of longer cell cycles are more probable. As we will shortly see, this property leads to mean protein levels being dependent on CV_T^2 .

Results

Computing the average number of protein molecules

All the models shown in Fig 2 are identical in terms of finding $\langle x(t) \rangle$ and in principle any one of them could have been used. We choose to analyze the full model illustrated in Fig 2A. Time evolution of the statistical moments of $x(t)$ can be obtained from the Kolmogorov forward equations corresponding to the full model in Fig 2A combined with the cell-division process described in Fig 3. We refer the reader to [52, 56, 57] for an introduction to moment dynamics for stochastic and hybrid systems. Analysis in Appendix A in S1 Text shows

$$\frac{d\langle x \rangle}{dt} = k_x \langle B \rangle - \frac{k}{2} \sum_{j=1}^n (j \langle x g_{ij} \rangle). \quad (8)$$

Note that the time-derivative of the mean protein level (first-order moment) is unclosed, in the sense that, it depends on the second-order moment $\langle x g_{ij} \rangle$. Typically, approximate closure methods are used to solve moments in such cases [52, 57–62]. However, the fact that g_{ij} is binary can be exploited to automatically close moment dynamics. In particular, since $g_{ij} \in \{0, 1\}$

$$\langle g_{ij}^n x^m \rangle = \langle g_{ij} x^m \rangle, \quad n \in \{1, 2, \dots\} \quad (9)$$

for any non-negative integer m . Moreover, as only a single state g_{ij} can be 1 at any time

$$\langle g_{ij} g_{rq} x^m \rangle = 0, \quad \text{if } i \neq r \text{ or } j \neq q. \quad (10)$$

Using Eqs (9) and (10), the time evolutions of $\langle g_{ij} \rangle$ and $\langle xg_{ij} \rangle$ are obtained as

$$\frac{d\langle g_{i1} \rangle}{dt} = kp_i \sum_{j=1}^n \langle jg_{ij} \rangle - ik\langle g_{i1} \rangle, \tag{11}$$

$$\frac{d\langle g_{ij} \rangle}{dt} = ik\langle g_{i(j-1)} \rangle - ik\langle g_{ij} \rangle, j = \{2, \dots, i\}, \tag{12}$$

$$\frac{d\langle xg_{i1} \rangle}{dt} = k_x \langle B \rangle \langle g_{i1} \rangle + \frac{k}{2} p_i \sum_{j=1}^n \langle jxg_{ij} \rangle - ik\langle xg_{i1} \rangle, \tag{13}$$

$$\frac{d\langle xg_{ij} \rangle}{dt} = k_x \langle B \rangle \langle g_{ij} \rangle - ik\langle xg_{ij} \rangle + ik\langle xg_{i(j-1)} \rangle, j = \{2, \dots, i\} \tag{14}$$

and only depend on $\langle g_{ij} \rangle$ and $\langle xg_{ij} \rangle$ (see Appendix A in S1 Text). Thus, Eqs (8) and (11)–(14) constitute a closed system of linear differential equations from which moments can be computed exactly.

To obtain an analytical formula for the average number of proteins, we start by performing a steady-state analysis of Eq (8) that yields

$$\sum_{j=1}^n \overline{\langle jxg_{ij} \rangle} = \frac{2k_x \langle B \rangle}{k}, \tag{15}$$

where $\overline{\langle \cdot \rangle}$ denotes the expected value in the limit $t \rightarrow \infty$. Using Eq (15), $\overline{\langle xg_{i1} \rangle}$ is determined from Eq (13), and then all moments $\overline{\langle xg_{ij} \rangle}$ are obtained recursively by performing a steady-state analysis of Eq (14) for $j = \{2, \dots, i\}$. This analysis results in

$$\overline{\langle xg_{ij} \rangle} = \frac{k_x \langle B \rangle}{ik} p_i \left(1 + \frac{j}{i} \right). \tag{16}$$

Using Eqs (7), (16) and the fact that $\sum_{i=1}^n \sum_{j=1}^i g_{ij} = 1$ we obtain the following expression for the mean protein level

$$\overline{\langle x \rangle} = \overline{\left\langle x \sum_{i=1}^n \sum_{j=1}^i g_{ij} \right\rangle} = \sum_{i=1}^n \sum_{j=1}^i \overline{\langle xg_{ij} \rangle} = \frac{k_x \langle B \rangle \langle T \rangle (3 + CV_T^2)}{2}. \tag{17}$$

It is important to point that Eq (17) holds irrespective of the complexity, i.e., the number of states G_{ij} used in the phase-type distribution to approximate the cell-cycle time distribution. As expected, $\overline{\langle x \rangle}$ increases linearly with the average cell-cycle time duration $\langle T \rangle$ with longer cell cycles resulting in more accumulation of proteins. Consistent with previous findings, Eq (17) shows that the mean protein level is also affected by the randomness in the cell-cycle times (CV_T^2) [40, 63]. For example, $\overline{\langle x \rangle}$ reduces by 25% as T changes from being exponentially distributed ($CV_T^2 = 1$) to periodic ($CV_T^2 = 0$) for fixed $\langle T \rangle$. Next, we determine the noise in protein copy numbers, as quantified by the squared coefficient of variation.

Computing the protein noise level

Recall that the full model introduced in Fig 2A has three distinct noise mechanisms. Our strategy for computing the protein noise level is to first analyze the model with a single noise source, and then consider models with two and three sources. As shown below, this approach provides

a systematic dissection of the protein noise level into components representing contributions from different mechanisms.

Contribution from randomness in cell-cycle times. We begin with the model shown in Fig 2B, where noise comes from a single source—random cell-division events. For this model, the time evolution of the second-order moment of the protein copy number is obtained as

$$\frac{d\langle x^2 \rangle}{dt} = 2k_x \langle B \rangle \langle x \rangle - \frac{3k}{4} \sum_{j=1}^n \left(j \langle x^2 g_{jj} \rangle \right), \quad (18)$$

and depends on third-order moments $\langle x^2 g_{ij} \rangle$ (see Appendix B in S1 Text). Using the approach introduced earlier for obtaining the mean protein level, we close moment equations by writing the time evolution of moments $\langle x^2 g_{ij} \rangle$. Using Eqs (9) and (10)

$$\frac{d\langle x^2 g_{i1} \rangle}{dt} = 2k_x \langle B \rangle \langle x g_{i1} \rangle + \frac{k}{4} p_i \sum_{j=1}^n \left(j \langle x^2 g_{ij} \rangle \right) - ik \langle x^2 g_{i1} \rangle, \quad (19)$$

$$\frac{d\langle x^2 g_{ij} \rangle}{dt} = 2k_x \langle B \rangle \langle x g_{ij} \rangle - ik \langle x^2 g_{ij} \rangle + ik \langle x^2 g_{(i-1)j} \rangle, \quad j = \{2, \dots, i\}. \quad (20)$$

Note that the moment dynamics for $\langle x \rangle$ and $\langle x g_{ij} \rangle$ obtained in the previous section (Eqs (8), (13) and (14)) are identical for all the models in Fig 2, irrespective of whether the noise mechanism is modeled deterministically or stochastically. Eqs (8), (11)–(14) and (18)–(20) represent a closed set of linear differential equations and their steady-state analysis yields

$$\overline{\langle x^2 g_{ij} \rangle} = \frac{k_x^2 \langle B \rangle^2 \langle T \rangle (3 + CV_T^2)}{3ik} p_i + \frac{2k_x^2 \langle B \rangle^2}{i^2 k^2} \left(\frac{j^2 + 2ij + j}{2i} \right) p_i. \quad (21)$$

From Eq (21)

$$\overline{\langle x^2 \rangle} = \overline{\left\langle x^2 \sum_{i=1}^n \sum_{j=1}^i g_{ij} \right\rangle} = \sum_{i=1}^n \sum_{j=1}^i \overline{\langle x^2 g_{ij} \rangle} = k_x^2 \langle B \rangle^2 \frac{\langle T^3 \rangle + 4CV_T^2 \langle T \rangle^3 + 6\langle T \rangle^3}{3\langle T \rangle}, \quad (22)$$

$$\langle T^3 \rangle = \left(1 + 3CV_T^2 + 2 \sum_{i=1}^n \frac{p_i}{i^2} \right) \langle T \rangle^3. \quad (23)$$

Using Eq (22) and the mean protein count quantified in Eq (17), we obtain the following squared coefficient of variation

$$CV_E^2 = \frac{1}{27} + \frac{4 \left(9 \frac{\langle T^3 \rangle}{\langle T \rangle^3} - 9 - 6CV_T^2 - 7CV_T^4 \right)}{27(3 + CV_T^2)^2}, \quad (24)$$

where CV_E^2 represents the noise contribution from random cell-division events. Since cell division is a global event that affects expression of all genes, this noise contribution can also be referred to as *extrinsic noise* [49, 64–67]. In reality, there would be other sources of extrinsic noise, such as, fluctuations in the gene-expression machinery that we have ignored in this analysis.

Note that $CV_E^2 \rightarrow 1/27$ as T approaches a delta distribution, i.e., cell divisions occur at fixed time intervals. We discuss simplifications of Eq (24) in various limits. For example, if the time

taken to complete cell cycle is lognormally distributed, then

$$\frac{\langle T^3 \rangle}{\langle T \rangle^3} = (1 + CV_T^2)^3 \Rightarrow CV_E^2 = \frac{1}{27} + \frac{4(21CV_T^2 + 20CV_T^4 + 9CV_T^6)}{27(3 + CV_T^2)^2} \quad (25)$$

and extrinsic noise monotonically increases with CV_T^2 . If fluctuations in T around $\langle T \rangle$ are small, then using Taylor series

$$\langle T^3 \rangle / \langle T \rangle^3 \approx 1 + 3CV_T^2. \quad (26)$$

Substituting Eq (26) in Eq (24) and ignoring CV_T^4 and higher order terms yields

$$CV_E^2 \approx \frac{1}{27} + \frac{28CV_T^2}{81}, \quad (27)$$

where the first term is the extrinsic noise for $CV_T^2 \rightarrow 0$ and the second term is the additional noise due to random cell-division events.

Contribution from partitioning errors. Next, we consider the model illustrated in Fig 2C with both random cell-division events and partitioning of protein between the two daughter cells. Thus, the protein noise level here represents the contribution from both these sources. Analysis in Appendix C in S1 Text shows that the time evolution of $\langle x^2 \rangle$ and $\langle x^2 g_{ij} \rangle$ are given by

$$\frac{d\langle x^2 \rangle}{dt} = 2k_x \langle B \rangle \langle x \rangle + \frac{k}{4} \alpha \sum_{j=1}^n (j \langle x g_{jj} \rangle) - \frac{3k}{4} \sum_{j=1}^n (j \langle x^2 g_{jj} \rangle), \quad (28)$$

$$\frac{d\langle x^2 g_{i1} \rangle}{dt} = 2k_x \langle B \rangle \langle x g_{i1} \rangle + \frac{k}{4} p_i \sum_{j=1}^n (j \langle x^2 g_{ij} \rangle) + \frac{k}{4} \alpha p_i \sum_{j=1}^n (j \langle x g_{ij} \rangle) - ik \langle x^2 g_{i1} \rangle, \quad (29)$$

$$\frac{d\langle x^2 g_{i1} \rangle}{dt} = 2k_x \langle B \rangle \langle x g_{ij} \rangle - ik \langle x^2 g_{ij} \rangle + ik \langle x^2 g_{(i-1)j} \rangle, j = \{2, \dots, i\}. \quad (30)$$

Note that Eqs (28) and (29) are slightly different from their counterparts obtained in the previous section (Eqs (18) and (19)) with additional terms that depend on α , where α quantifies the degree of partitioning error as defined in Eq (5). As expected, Eqs (28) and (29) reduce to Eqs (18) and (19) when $\alpha = 0$ (i.e., deterministic partitioning). Computing $\overline{\langle x^2 \rangle}$ by performing a steady-state analysis of Eqs (28)–(30) and using a similar approach as in Eq (22) we obtain

$$\overline{\langle x^2 \rangle} = k_x^2 \langle B \rangle^2 \frac{\langle T^3 \rangle + 4CV_T^2 \langle T \rangle^3 + 6\langle T \rangle^3}{3\langle T \rangle} + \frac{2\alpha k_x \langle B \rangle \langle T \rangle}{3}. \quad (31)$$

Finding CV_R^2 of the protein level and subtracting the extrinsic noise found in Eq (24) yields

$$CV_R^2 = \frac{4\alpha}{3(3 + CV_T^2)} \frac{1}{\langle x \rangle}, \quad (32)$$

where CV_R^2 represents the contribution of partitioning errors to the protein noise level. Intriguingly, while CV_R^2 increases with α , it decrease with CV_T^2 . Thus, as cell-division times become more random for a fixed $\langle T \rangle$ and $\overline{\langle x \rangle}$, the noise contribution from partitioning errors decrease. It turns out that this dependence of CV_R^2 on CV_T^2 is a direct result of the second equation in Eq (5), where stochasticity in the partitioning process increases linearly with $x(t_s)$, the number of protein molecules just before division. Based on Eq (17), one needs to reduce k_x or $\langle B \rangle$ to

maintain a fixed $\langle x \rangle$ for increasing randomness in cell-division times. Since the average number of protein molecules just before division is $2k_x \langle B \rangle \langle T \rangle$ (see Appendix D in [S1 Text](#)), a reduction in k_x or $\langle B \rangle$ results in a lower number of protein molecules before division, and hence, lesser noise from partitioning as per [Eq \(5\)](#) and a smaller CV_R^2 . This reasoning is supported by the fact that if we redefine the noise in the partitioning process to make it independent of $x(t_s)$, i.e. modify [Eq \(5\)](#) as

$$\langle x_+(t_s) | x(t_s) \rangle = \frac{x(t_s)}{2}, \quad \langle x_+^2(t_s) - \langle x_+(t_s) \rangle^2 | x(t_s) \rangle = \alpha, \quad (33)$$

then the noise contribution from partitioning errors is given by

$$CV_R^2 = \frac{4\alpha}{3} \frac{1}{\langle x \rangle^2}, \quad (34)$$

and the dependency of CV_R^2 on CV_T disappears (Appendix D in [S1 Text](#)).

Contribution from stochastic expression. Finally, we consider the full model in [Fig 2A](#) with all the three different noise sources. For this model, moment dynamics is obtained as (see Appendix E in [S1 Text](#))

$$\frac{d\langle x^2 \rangle}{dt} = k_x \langle B^2 \rangle + 2k_x \langle B \rangle \langle x \rangle + \frac{k}{4} \alpha \sum_{j=1}^n (j \langle x g_{jj} \rangle) - \frac{3k}{4} \sum_{j=1}^n (j \langle x^2 g_{jj} \rangle), \quad (35)$$

$$\frac{d\langle x^2 g_{i1} \rangle}{dt} = k_x \langle B^2 \rangle \langle g_{i1} \rangle + 2k_x \langle B \rangle \langle x g_{i1} \rangle + \frac{k}{4} p_i \sum_{j=1}^n (j \langle x^2 g_{ij} \rangle) + \frac{k}{4} \alpha p_i \sum_{j=1}^n (j \langle x g_{ij} \rangle) - ik \langle x^2 g_{i1} \rangle, \quad (36)$$

$$\frac{d\langle x^2 g_{ij} \rangle}{dt} = k_x \langle B^2 \rangle \langle g_{ij} \rangle + 2k_x \langle B \rangle \langle x g_{ij} \rangle - ik \langle x^2 g_{ij} \rangle + ik \langle x^2 g_{(i-1)j} \rangle, \quad j = \{2, \dots, i\}. \quad (37)$$

Compared to [Eqs \(28\)–\(30\)](#), [\(35\)–\(37\)](#) have additional terms of the form $k_x \langle B^2 \rangle$, where $\langle B^2 \rangle$ is the second-order moment of the protein burst size in [Eq \(1\)](#). Performing an identical analysis as before we obtain

$$\langle x^2 \rangle = k_x^2 \langle B \rangle^2 \frac{\langle T^3 \rangle + 4CV_T^2 \langle T \rangle^3 + 6\langle T \rangle^3}{3\langle T \rangle} + \frac{2\alpha k_x \langle B \rangle \langle T \rangle}{3} + \frac{k_x \langle B^2 \rangle \langle T \rangle (3CV_T^2 + 5)}{2}, \quad (38)$$

which yields the following total protein noise level

$$CV^2 = CV_E^2 + CV_R^2 + CV_P^2 = CV_E^2 + \underbrace{\frac{4\alpha}{3(3 + CV_T^2)} \frac{1}{\langle x \rangle}}_{\text{Partitioning noise } (CV_R^2)} + \underbrace{\frac{3CV_T^2 + 5}{3(3 + CV_T^2)} \frac{\langle B^2 \rangle}{\langle B \rangle} \frac{1}{\langle x \rangle}}_{\text{Production noise } (CV_P^2)} \quad (39)$$

that can be decomposed into three terms. The first term CV_E^2 represents the contribution from random cell-division events and is given by [Eq \(24\)](#). The second term CV_R^2 is the contribution from partitioning errors determined in the previous section (partitioning noise), and the final term CV_P^2 is the additional noise representing the contribution from stochastic expression (production noise). A common approach to study gene expression noise is to decompose it into *intrinsic* and *extrinsic* components. These components are obtained experimentally using the dual-color assay that measures the correlation in the expression of two identical copies of the gene [\[49\]](#). As per this definition, CV_E^2 represents the extrinsic noise as random cell-division

events are common to all genes and makes expression levels more correlated in individual cells. In contrast, the contributions from noisy production and partitioning represent the intrinsic noise as they are specific to an individual gene and make expression levels less correlated.

An interesting observation from Eq (39) is that CV_T^2 has opposite effects on CV_R^2 and CV_P^2 (for fixed mean protein level). While CV_R^2 monotonically decreases with increasing CV_T^2 , CV_P^2 increases with CV_T^2 . Thus, if $\langle x \rangle$ is small and α is large, then the noise contributed from partitioning dominates the total noise, and making cell-cycle duration more random will reduce the total noise. However, since both CV_E^2 and CV_P^2 are monotonically increasing functions of CV_T^2 , the total noise will begin to increase with CV_T^2 once these noise sources become dominant. It turns out that in certain cases the intrinsic noise becomes invariant of CV_T^2 . For example, when $B = 1$ with probability one, i.e., proteins are synthesized one at a time at exponentially distributed time intervals and $\alpha = 1$ (binomial partitioning)

$$CV^2 = CV_E^2 + \frac{4}{3(3 + CV_T^2)} \frac{1}{\langle x \rangle} + \frac{3CV_T^2 + 5}{3(3 + CV_T^2)} \frac{1}{\langle x \rangle} = CV_E^2 + \frac{1}{\langle x \rangle}. \tag{40}$$

In this limit the intrinsic noise is always 1/Mean irrespective of the cell-cycle time distribution T [33]. Note that the average number of proteins itself depends on T as shown in Eq (17).

Another important limit is $CV_T^2 \rightarrow 0$, in which case Eq (39) reduces to

$$CV^2 \approx \underbrace{\frac{1}{27}}_{\text{Extrinsicnoise}} + \underbrace{\frac{4\alpha}{9} \frac{1}{\langle x \rangle} + \frac{5 \langle B^2 \rangle}{9 \langle B \rangle} \frac{1}{\langle x \rangle}}_{\text{Intrinsicnoise}}, \tag{41}$$

and is similar to the result obtained in [38] for deterministic cell-division times and binomial partitioning.

Fig 4 shows how different protein noise components change as a function of the mean protein level as the gene’s transcription rate k_x is modulated. The extrinsic noise is primarily determined by the distribution of the cell-cycle time and is completely independent of the mean. In contrast, both CV_R^2 and CV_P^2 scale inversely with the mean, albeit with different scaling factors (Fig 4). This observation is particularly important since many single-cell studies in *E. coli*, yeast and mammalian cells have found the protein noise levels to scale inversely with the mean across different genes [68–71]. Based on this scaling it is often assumed that the observed cell-to-cell variability in protein copy numbers is a result of stochastic expression. However, as our results show, noise generated thorough partitioning errors is also consistent with these experimental observations and it may be impossible to distinguish between these two noise mechanisms based on protein CV^2 versus mean plots unless α is known.

Quantifying the effects of gene duplication on protein noise

The full model introduced in Fig 2 assumes that the transcription rate (i.e., the protein burst arrival rate) is constant throughout the cell cycle. This model is now extended to incorporate gene duplication during cell cycle, which increases the burst arrival (transcription) rate by f times (Fig 5). Note that due to gene dosage compensation, doubling the genome does not always correspond to $f = 2$ [72–74]. If $f > 1$, then accumulation of proteins will be bilinear as illustrated in Fig 1. As before, we again take a phenomenological approach to model the timing of gene duplication. The cell-cycle time T is divided into two intervals: time from the start of cell cycle to gene duplication (T_1), and time from gene duplication to cell division (T_2). T_1 and T_2 are independent random variables, each drawn from a mixture of Erlang distributions (see

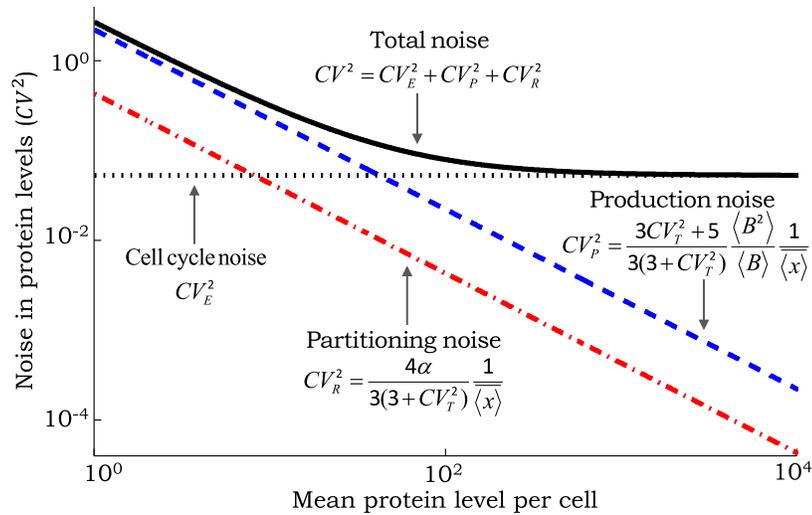


Fig 4. Scaling of noise as a function of the mean protein level for different mechanisms. The contribution of random cell-division events to the noise in protein copy numbers (extrinsic noise) is invariant of the mean. In contrast, contributions from partitioning errors at the time of cell division (partitioning noise) and stochastic expression (production noise) scale inversely with the mean. The scaling factors are shown as a function of the protein random burst size B , noise in cell-cycle time (CV_T^2) and magnitude of partitioning errors quantified by α (see Eq (5)). With increasing mean level the total noise first decreases and then reaches a baseline that corresponds to extrinsic noise. For this plot, B is assumed to be geometrically-distributed with mean $\langle B \rangle = 1.5$, $CV_T^2 = 0.05$ and $\alpha = 1$ (i.e., binomial partitioning).

doi:10.1371/journal.pcbi.1004972.g004

Fig. B in S1 Text). The mean cell-cycle duration and its noise can be expressed as

$$\langle T \rangle = \langle T_1 \rangle + \langle T_2 \rangle, \quad \beta = \frac{\langle T_1 \rangle}{\langle T \rangle}, \quad CV_T^2 = \beta^2 CV_{T_1}^2 + (1 - \beta)^2 CV_{T_2}^2, \quad (42)$$

where CV_X^2 denotes the squared coefficient of variation of the random variable X . An important variable in this formulation is β , which represents the average time of gene duplication normalized by the mean cell-cycle time. Thus, β values close to 0 (1) imply that the gene is duplicated

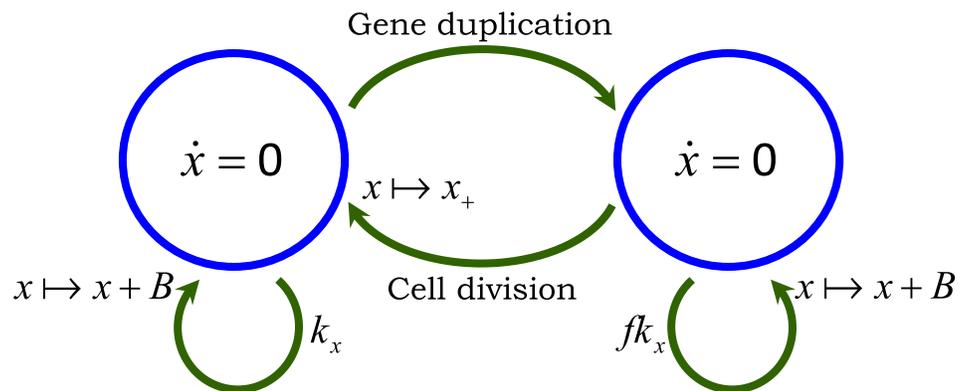


Fig 5. Model illustrating stochastic expression together with random gene-duplication and cell-division events. At the start of cell cycle, protein production occurs in stochastic bursts with rate k_x . Genome duplication occurs at a random point T_1 within the cell cycle and increases the burst arrival rate to $f k_x$ ($f > 1$). Cell division occurs after time T_2 from genome duplication, at which point the burst arrival rate reverts back to k_x and proteins are randomly partitioned between cells based on Eq (4).

doi:10.1371/journal.pcbi.1004972.g005

early (late) in the cell-cycle process. Moreover, the noise in the gene-duplication time is controlled via $CV_{T_1}^2$.

We refer the reader to Appendix F in [S1 Text](#) for a detailed analysis of the model in [Fig 5](#) and only present the main results on the protein mean and noise levels. The steady-state mean protein count is given by

$$\overline{\langle x \rangle} = \frac{k_x \langle B \rangle \langle T_1 \rangle (2f(1 - \beta) + 3\beta + \beta CV_{T_1}^2)}{2} + \frac{k_x \langle B \rangle \langle T_2 \rangle (3f(1 - \beta) + 4\beta + f(1 - \beta) CV_{T_2}^2)}{2}, \quad (43)$$

and decreases with β , i.e., a gene that duplicates early has on average, more number of proteins. When $\beta = 1$, then the transcription rate is k_x throughout the cell cycle and we recover the mean protein level obtained in [Eq \(17\)](#). Similarly, when $\beta = 0$ the transcription rate is $f k_x$ and we obtain f times of the amount as in [Eq \(17\)](#). As per our earlier observation, more randomness in the timing of genome duplication and cell division (i.e., higher $CV_{T_1}^2$ and $CV_{T_2}^2$ values) increases $\overline{\langle x \rangle}$.

Our analysis shows that the total protein noise level can be decomposed into three components

$$CV^2 = CV_E^2 + CV_R^2 + CV_P^2 \quad (44)$$

where CV_E^2 is the extrinsic noise from random genome-duplication/cell-division events, and the sum of the contributions from partitioning errors (CV_R^2) and stochastic expression (CV_P^2) is the intrinsic noise. We refer the reader to Appendix F in [S1 Text](#) for noise formulas for any f , and only present formulas for $f = 2$ here. In this case, the intrinsic noise is obtained as

$$CV_R^2 + CV_P^2 = \frac{\overbrace{4\alpha(2 - \beta)}^{CV_R^2}}{3 \left((\beta^2 - 4\beta + 6) + \beta^2 CV_{T_1}^2 + 2(1 - \beta)^2 CV_{T_2}^2 \right)} \frac{1}{\langle x \rangle} + \frac{\overbrace{(10 - 8\beta + 3\beta^2) + 6(1 - \beta)^2 CV_{T_2}^2 + 3\beta^2 CV_{T_1}^2}^{CV_P^2}}{3 \left((\beta^2 - 4\beta + 6) + \beta^2 CV_{T_1}^2 + 2(1 - \beta)^2 CV_{T_2}^2 \right)} \frac{\langle B^2 \rangle}{\langle B \rangle} \frac{1}{\langle x \rangle} \quad (45)$$

Note that for $\beta = 0$ and 1, we recover the intrinsic noise level in [Eq \(39\)](#) from [Eq \(45\)](#). Interestingly, for $B = 1$ with probability 1 and $\alpha = 1$, the intrinsic noise is always $1/\text{Mean}$ irrespective of the values chosen for $CV_{T_1}^2$, $CV_{T_2}^2$ and β . For high precision in the timing of cell-cycle events ($CV_{T_1} \rightarrow 0, CV_{T_2} \rightarrow 0$)

$$CV^2 \approx \underbrace{\frac{\overbrace{4 - 3(\beta - 2)^2 \beta^2}^{CV_E^2}}{3(\beta^2 - 4\beta + 6)^2}}_{\text{Extrinsicnoise}} + \underbrace{\frac{\overbrace{4\alpha(2 - \beta)}^{CV_R^2}}{3(\beta^2 - 4\beta + 6)} \frac{1}{\langle x \rangle}}_{\text{Intrinsicnoise}} + \underbrace{\frac{\overbrace{(10 - 8\beta + 3\beta^2) \langle B^2 \rangle}_{CV_P^2}}{3(\beta^2 - 4\beta + 6) \langle B \rangle} \frac{1}{\langle x \rangle}}_{\text{Intrinsicnoise}}, \quad (46)$$

where mean protein level is given by

$$\overline{\langle x \rangle} \approx \frac{k_x \langle B \rangle \langle T_1 \rangle (4 - \beta)}{2} + k_x \langle B \rangle \langle T_2 \rangle (3 - \beta). \quad (47)$$

We investigate how different noise components in [Eq \(46\)](#) vary with β as the mean protein

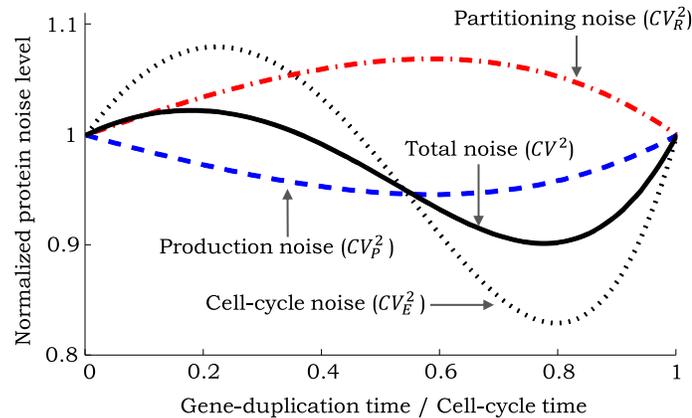


Fig 6. Contributions from different noise sources as a function of the timing of genome duplication for $CV_{T_1}^2 = CV_{T_2}^2 = 0.05$. Different noise components in Eq (46) are plotted as a function of β , which represents the fraction of time within the cell cycle at which gene duplication occurs. The mean protein level is held constant by simultaneously changing the transcription rate k_x . Noise levels are normalized by their respective value at $\beta = 0$. The noise contribution from partitioning errors is maximized at $\beta \approx 0.6$. In contrast, the contribution from stochastic expression is minimum at $\beta \approx 0.6$. The extrinsic noise contribution from random gene-duplication and cell-division events is maximum at $\beta \approx 0.2$ and minimum at $\beta \approx 0.8$. For this plot, the mean of the protein is 170 molecules per cell; and the bursts are geometrically distributed with $\langle B \rangle = 10$.

doi:10.1371/journal.pcbi.1004972.g006

level is held fixed by changing k_x . Fig 6 shows that CV_p^2 follows a U-shaped profile with the optima occurring at $\beta = 2 - \sqrt{2} \approx 0.6$ and the corresponding minimum value being $\approx 5\%$ lower than its value at $\beta = 0$. An implication of this result is that if stochastic expression is the dominant noise source, then gene duplication can result in slightly lower protein noise levels. In contrast to CV_p^2 , CV_r^2 has a maxima at $\beta = 2 - \sqrt{2}$ which is $\approx 6\%$ higher than its value at $\beta = 0$ (Fig 6). Analysis in Appendix F5 in S1 Text reveals that CV_r^2 and CV_p^2 follow the same qualitative shapes as in Fig 6 for any $CV_{T_1}^2$ and $CV_{T_2}^2$. Interestingly, when $CV_{T_1}^2 = CV_{T_2}^2$, the maximum and minimum values of CV_r^2 and CV_p^2 always occur at $\beta = 2 - \sqrt{2}$ albeit with different optimal values than Fig 6 (see Fig. C in S1 Text). For example, if $CV_{T_1}^2 = CV_{T_2}^2 = 1$ (i.e., exponentially distributed T_1 and T_2), then the maximum value of CV_r^2 is 20% higher and the minimum value of CV_p^2 is 10% lower than their respective value for $\beta = 0$. Given that the effect of changing β on CV_p^2 and CV_r^2 is small and antagonistic, the overall affect of genome duplication on intrinsic noise may be minimal and hard to detect experimentally.

While the above analysis is for a stable protein, a natural question to ask is how would these results change for an unstable protein? Consider an unstable protein with half-life considerably shorter than the cell-cycle duration. This rapid turnover ensures that the protein level equilibrates instantaneously after cell-division and gene-duplication events. Let γ_x denote the protein decay rate. Then, the mean protein level before and after genome duplication is $\overline{\langle x \rangle} = k_x \langle B \rangle / \gamma_x$ and $\overline{\langle x \rangle} = 2k_x \langle B \rangle / \gamma_x$, respectively. Note that in the limit of large γ_x there is no noise contribution from partitioning errors since errors incurred at the time of cell division would be instantaneously corrected. The extrinsic noise, which can be interpreted as the protein noise level for deterministic protein production and decay is obtained as (for analysis on general f see Appendix G in S1 Text)

$$CV_E^2 = \frac{(1 - \beta)\beta}{(2 - \beta)^2}, \tag{48}$$

which is similar to noise level reported in [75]. When $\beta = 0$ or 1, the transcription rate and the protein level are constant within the cell cycle and $CV_E^2 = 0$. Moreover, CV_E^2 is maximized at $\beta = 2/3$ with a value of $1/8$. Thus, in contrast to a stable protein, extrinsic noise in an unstable protein is strongly dependent on the timing of gene duplication. Next, consider the intrinsic noise component. Analysis in Appendix G in [S1 Text](#) shows that the noise contribution from random protein production and decay is

$$CV_p^2 = \frac{1}{2} \left(\frac{\langle B^2 \rangle}{\langle B \rangle} + 1 \right) \frac{1}{\langle x \rangle}, \quad \overline{\langle x \rangle} = \frac{k_x \langle B \rangle (2 - \beta)}{\gamma_x}. \quad (49)$$

While the mean protein level is strongly dependent on β , the intrinsic noise Fano factor = $CV_p^2 \times \langle x \rangle$ is independent of it. Thus, similar to what was observed for a stable protein, the intrinsic noise in an unstable protein is invariant of β for a fixed $\overline{\langle x \rangle}$.

Discussion

We have investigated a model of protein expression in bursts coupled to discrete gene-duplication and cell-division events. The novelty of our modeling framework lies in describing the size of protein bursts B , the time between cell birth and gene duplication T_1 , the time between gene duplication and cell division T_2 , and partitioning of molecules during cell division through general statistical distributions. Exact formulas connecting the protein mean and noise levels to these underlying distributions were derived. Furthermore, the protein noise level, as measured by the squared coefficient of variation, was decomposed into three components representing contributions from gene-duplication/cell-division events, stochastic expression and random partitioning. While the first component is independent of the mean protein level, the other two components are inversely proportional to it. Some important insights are as follows:

- The mean protein level is affected by both the first and second-order moments of T_1 and T_2 . In particular, randomness in these times (for a fixed mean) increases the average protein count. This increase can be attributed to the fact that increasing cell-cycle time variations leads to positively skewed distributions, making longer cell cycles (and hence higher protein accumulation) more likely.
- Random gene-duplication/cell-division events create an extrinsic noise term which is completely determined by moments of T_1 and T_2 up to order three. Interestingly, noise in the timing of these events also critically affects the intrinsic noise contributions from stochastic expression and partitioning. Hence, ignoring the effect of cell-cycle time variations, may lead to erroneous estimation of intrinsic noise.
- The noise contribution from partitioning errors decreases with increasing randomness in T_1 and T_2 . Thus, if $\overline{\langle x \rangle}$ is sufficiently small and α is large compared to B in [Eq \(45\)](#), increasing noise in the timing of cell-cycle events decreases the total noise level.

A key limitation of our approach is to model timing of gene-duplication/cell-division events through independent random variables. There is always non-zero correlation in the cell-cycle durations of mother and daughter cells [76–78]. Moreover, in the same cell cycle, times T_1 and T_2 could be dependent [79]. While our assumption on independence of timing maybe unrealistic, it played an important role in deriving exact analytical formulas for the protein mean and noise levels. We have used Monte Carlo simulations to investigate scenarios where T_1 and T_2 , or successive cell-division events, have memory and are dependent random variables (see Appendix H in [S1 Text](#)). Our analysis reveals that the results presented in [Figs 4 and 6](#) hold even when the assumption of independent timing is perturbed over biologically meaningful parameters.

Effect of gene duplication on noise level

In this first-of-its-kind study, we have investigated how discrete f -fold changes in the transcription rate due to gene duplication affect the intercellular variability in protein levels. Not surprisingly, the timing of genome duplication strongly affects the mean protein level— $\langle x \rangle$ can change up to f folds depending on whether the gene duplicates early ($\beta = 0$) or late ($\beta = 1$) in the cell cycle. Results show that genome duplication has counter intuitive effects on the protein noise level (Fig 6). For example, if stochastic expression is the dominant source of noise, then doubling of transcription due to duplication results in lower noise, as compared to constant transcription throughout the cell cycle. This is because for the same mean protein level, there are more burst (transcription) events in the case of genome duplication ($f = 2$) than constant transcription ($f = 1$). For example, consider deterministic timing ($CV_{T_1}^2 = CV_{T_2}^2 = 0$) and gene duplication in the middle of the cell cycle ($\beta = 0.5$). Then, for the case $\beta = 1$, there are on average $k_x \langle T \rangle$ burst events per cell cycle. For the same $\langle x \rangle$, there are $1.05 k_x \langle T \rangle$ production events in the case of gene duplication ($\beta = 0.5$). This slight increase in the number of transcription events leads to better averaging of bursty protein synthesis and lower noise levels. Overall, the effect of β on different noise component is quite modest: as β varies, noise components deviate at maximum $\approx 20\%$ from their values at $\beta = 0$ (Fig 6). These results are in contrast to the case of an unstable protein, where noise from the cell-cycle process is strongly dependent on β as shown in Eq (48).

Noise in synchronized cells

The mathematical framework introduced for modeling timing of cell division can be easily used to compute noise in synchronized cells. For example, let the cell-cycle duration be an Erlang distribution with shape parameter n and rate parameter nk (i.e., $p_n = 1$ in Fig 3), which can be biologically interpreted as cells moving through n cell-cycle stages $G_{n1}, G_{n2}, \dots, G_{nn}$. Statistical moments conditioned on the cell-cycle stage G_{nj} can be obtained using

$$\overline{\langle x^m | g_{nj} = 1 \rangle} = \frac{\overline{\langle g_{nj} x^m \rangle}}{\overline{\langle g_{nj} \rangle}}, \quad m \in \{1, 2\}. \tag{50}$$

Using Eq (50) and moments $\overline{\langle g_{nj} x^m \rangle}$ obtained from Eqs (16) and (35)–(37), yields the following conditional mean

$$\overline{\langle x | g_{nj} = 1 \rangle} = k_x \langle B \rangle \langle T \rangle \left(1 + \frac{j}{n} \right), \tag{51}$$

which increases with cell-cycle stage (i.e., higher values of j). The protein noise level given that cells are in stage G_{nj} is given by

$$\begin{aligned} CV^2|_{g_{nj}=1} &:= \frac{\overline{\langle x^2 | g_{nj} = 1 \rangle} - \overline{\langle x | g_{nj} = 1 \rangle}^2}{\overline{\langle x | g_{nj} = 1 \rangle}^2} \\ &= \underbrace{\frac{CV_E^2}{3(n+j)^2}}_{\text{Extrinsicnoise}} + \underbrace{\frac{2n\alpha}{3(n+j)} \frac{1}{\overline{\langle x | g_{nj} \rangle}}}_{\text{Intrinsicnoise}} + \underbrace{\frac{CV_P^2}{3(n+j)} \frac{\langle B^2 \rangle}{\langle B \rangle} \frac{1}{\overline{\langle x | g_{nj} \rangle}}}_{\text{Intrinsicnoise}}. \end{aligned} \tag{52}$$

Note that if n is large then the first term, which represents the noise contribution from the cell-cycle process, is negligible and can be dropped. Interesting, while the noise contribution from

partitioning errors CV_R^2 decreases with cell-cycle stage, the noise contribution from stochastic expression CV_P^2 increases with j . Moreover, for $B = 1$ with probability 1 and $\alpha = 1$, the intrinsic noise is always $1/\text{Mean}$ irrespective of j . Assuming high n , the noise at cell birth ($j = 1$) and division ($j = n$) are obtained as

$$CV^2|_{g_{n1}=1} = \underbrace{\frac{2\alpha}{3} \frac{1}{\langle x|g_{n1} \rangle}}_{\text{Intrinsic noise}} + \frac{1}{3} \frac{\langle B^2 \rangle}{\langle B \rangle} \frac{1}{\langle x|g_{n1} \rangle} \quad (53)$$

$$CV^2|_{g_{nn}=1} = \underbrace{\frac{\alpha}{3} \frac{1}{\langle x|g_{nn} \rangle}}_{\text{Intrinsic noise}} + \frac{2}{3} \frac{\langle B^2 \rangle}{\langle B \rangle} \frac{1}{\langle x|g_{nn} \rangle}, \quad (54)$$

respectively. Thus, measurements of Eqs (53) and (54) by synchronizing cells (or by using cell size as a proxy for cell-cycle stage) can be used to quantify α and $\langle B^2 \rangle / \langle B \rangle$, providing a novel way to separate these noise contributions. Next, we discuss how noise in asynchronous cell can be used to quantify these parameters.

Parameter inference in asynchronous cells

Simple models of bursty expression and decay predict the distribution of protein levels to be negative binomial (or gamma distributed in the continuous framework) [80, 81]. These distributions are characterized by two parameter—the burst arrival rate k_x and the average burst size $\langle B \rangle$, which can be estimated from measured protein mean and noise levels. This method has been used for estimating k_x and $\langle B \rangle$ across different genes in *E. coli* [47, 82]. Our detailed model that takes into account partitioning errors predicts (ignoring gene-duplication effects)

$$\text{Intrinsic noise} = \frac{4\alpha}{3(3 + CV_T^2)} \frac{1}{\langle x \rangle} + \frac{3CV_T^2 + 5}{3(3 + CV_T^2)} \frac{\langle B^2 \rangle}{\langle B \rangle} \frac{1}{\langle x \rangle}. \quad (55)$$

Using $CV_T^2 \ll 1$ and a geometrically distributed B [50, 83–85], Eq (55) reduces to

$$\text{Intrinsic noise} = \frac{4\alpha}{9} \frac{1}{\langle x \rangle} + \frac{5}{9} \frac{1 + 2\langle B \rangle}{\langle x \rangle}. \quad (56)$$

Given measurements of intrinsic noise and the mean protein level, $\langle B \rangle$ can be estimated from Eq (56) assuming $\alpha = 1$ (i.e., binomial partitioning). Once $\langle B \rangle$ is known, k_x is obtained from the mean protein level given by Eq (17). Since for many genes $\langle B \rangle \approx 0.5\text{--}5$ [47], the contribution of the first term in Eq (56) is significant, and ignoring it could lead to overestimation of $\langle B \rangle$. Overestimation would be even more severe if α happen to be much higher than 1, as would be the case for proteins that form aggregates or multimers [33]. One approach to estimate both $\langle B \rangle$ and α is to measure intrinsic noise changes in response to perturbing $\langle B \rangle$ by, for example, changing the mRNA translation rate through mutations in the ribosomal-binding sites (RBS). Consider a hypothetical scenario where the Fano Factor (intrinsic noise times the mean level) is 6. Let mutations in the RBS reduces $\langle x \rangle$ by 50%, implying a 50% reduction in $\langle B \rangle$. If the Fano factor changes from 6 to 4 due to this mutation, then $\langle B \rangle = 3.6$ and $\langle \alpha \rangle = 3.25$.

Our recent work has shown that higher-order statistics of protein levels (i.e., skewness and kurtosis) or transient changes in protein noise levels in response to blocking transcription provide additional information for discriminating between noise mechanisms [86, 87]. Up till now these studies have ignored noise sources in the cell-cycle process. It remains to be seen if such methods can be used for separating the noise contributions of partitioning errors and stochastic expression to reliably estimate $\langle B \rangle$ and α .

Integrating cell size and promoter switching

An important limitation of our modeling approach is that it does not take into account the size of growing cells. Recent experimental studies have provided important insights into the regulatory mechanisms controlling cell size [88–91]. More specifically, studies in *E. coli* and yeast argue for an “adder” model, where cell-cycle timing is controlled so as to add a constant volume between cell birth and division [78, 91–93]. Assuming exponential growth, this implies that the time taken to complete cell cycle is negatively correlated with cell size at birth. More importantly, cell size also affects gene expression—in mammalian cells transcription rates linearly increase with the cell size [94]. Thus, as cells become bigger they also produce more mRNAs to ensure gene product concentrations remains more or less constant. A main direction of future work would be to explicitly include cell size with size-dependent expression and timing of cell division determined by the adder model. This formulation will for the first time, allow simultaneous investigation of stochasticity in cell size, protein molecular count and concentration.

Our study ignores genetic promoter switching between active and inactive states, which has been shown to be a major source of noise in the expression of genes across organisms [95–104]. Taking into account promoter switching is particularly important for genome-duplication studies, where doubling the number of gene copies could lead to more efficient averaging of promoter fluctuations. Another direction of future work will be to incorporate this additional noise source into the modeling framework and investigate its contribution as a function of gene-duplication timing.

Supporting Information

S1 Text. Supplementary Material. Additional proofs and simulations. (PDF)

Author Contributions

Conceived and designed the experiments: MS DA AS.

Performed the experiments: MS CAVG.

Analyzed the data: MS CAVG.

Wrote the paper: MS AS.

References

1. Blake WJ, Kaern M, Cantor CR, Collins JJ. Noise in eukaryotic gene expression. *Nature*. 2003; 422:633–637. doi: [10.1038/nature01546](https://doi.org/10.1038/nature01546) PMID: [12687005](https://pubmed.ncbi.nlm.nih.gov/12687005/)
2. Raser JM, O’Shea EK. Noise in Gene Expression: Origins, Consequences, and Control. *Science*. 2005; 309:2010–2013. doi: [10.1126/science.1105891](https://doi.org/10.1126/science.1105891) PMID: [16179466](https://pubmed.ncbi.nlm.nih.gov/16179466/)

3. Neuert G, Munsky B, Tan RZ, Teytelman L, Khammash M, van Oudenaarden A. Systematic Identification of Signal-Activated Stochastic Gene Regulation. *Science*. 2013; 339:584–587. doi: [10.1126/science.1231456](https://doi.org/10.1126/science.1231456) PMID: [23372015](https://pubmed.ncbi.nlm.nih.gov/23372015/)
4. Libby E, Perkins TJ, Swain PS. Noisy information processing through transcriptional regulation. *Proceedings of the National Academy of Sciences*. 2007; 104:7151–7156. doi: [10.1073/pnas.0608963104](https://doi.org/10.1073/pnas.0608963104)
5. Fraser HB, Hirsh AE, Giaever G, Kumm J, Eisen MB. Noise Minimization in Eukaryotic Gene Expression. *PLOS Biology*. 2004; 2:e137. doi: [10.1371/journal.pbio.0020137](https://doi.org/10.1371/journal.pbio.0020137) PMID: [15124029](https://pubmed.ncbi.nlm.nih.gov/15124029/)
6. Lehner B. Selection to minimise noise in living systems and its implications for the evolution of gene expression. *Molecular Systems Biology*. 2008; 4:170. doi: [10.1038/msb.2008.11](https://doi.org/10.1038/msb.2008.11) PMID: [18319722](https://pubmed.ncbi.nlm.nih.gov/18319722/)
7. Losick R, Desplan C. Stochasticity and Cell Fate. *Science*. 2008; 320:65–68. doi: [10.1126/science.1147888](https://doi.org/10.1126/science.1147888) PMID: [18388284](https://pubmed.ncbi.nlm.nih.gov/18388284/)
8. Arkin AP, Ross J, McAdams HH. Stochastic Kinetic Analysis of Developmental Pathway Bifurcation in Phage λ -Infected *Escherichia coli* Cells. *Genetics*. 1998; 149:1633–1648. PMID: [9691025](https://pubmed.ncbi.nlm.nih.gov/9691025/)
9. Weinberger L, Burnett J, Toettcher J, Arkin A, Schaffer D. Stochastic Gene Expression in a Lentiviral Positive-Feedback Loop: HIV-1 Tat Fluctuations Drive Phenotypic Diversity. *Cell*. 2005; 122:169–182. doi: [10.1016/j.cell.2005.06.006](https://doi.org/10.1016/j.cell.2005.06.006) PMID: [16051143](https://pubmed.ncbi.nlm.nih.gov/16051143/)
10. Weinberger LS, Dar RD, Simpson ML. Transient-mediated fate determination in a transcriptional circuit of HIV. *Nature Genetics*. 2008; 40:466–470. doi: [10.1038/ng.116](https://doi.org/10.1038/ng.116) PMID: [18344999](https://pubmed.ncbi.nlm.nih.gov/18344999/)
11. Singh A, Weinberger LS. Stochastic gene expression as a molecular switch for viral latency. *Current Opinion in Microbiology*. 2009; 12:460–466. doi: [10.1016/j.mib.2009.06.016](https://doi.org/10.1016/j.mib.2009.06.016) PMID: [19595626](https://pubmed.ncbi.nlm.nih.gov/19595626/)
12. Dar RD, Hosmane NN, Arkin MR, Siliciano RF, Weinberger LS. Screening for noise in gene expression identifies drug synergies. *Science*. 2014; 344:1392–1396. doi: [10.1126/science.1250220](https://doi.org/10.1126/science.1250220) PMID: [24903562](https://pubmed.ncbi.nlm.nih.gov/24903562/)
13. Eldar A, Elowitz MB. Functional roles for noise in genetic circuits. *Nature*. 2010; 467:167–173. doi: [10.1038/nature09326](https://doi.org/10.1038/nature09326) PMID: [20829787](https://pubmed.ncbi.nlm.nih.gov/20829787/)
14. Veening JW, Smits WK, Kuipers OP. Bistability, Epigenetics, and Bet-Hedging in Bacteria. *Annual Review of Microbiology*. 2008; 62:193–210. doi: [10.1146/annurev.micro.62.081307.163002](https://doi.org/10.1146/annurev.micro.62.081307.163002) PMID: [18537474](https://pubmed.ncbi.nlm.nih.gov/18537474/)
15. Kussell E, Leibler S. Phenotypic diversity, population growth, and information in fluctuating environments. *Science*. 2005; 309:2075–2078. doi: [10.1126/science.1114383](https://doi.org/10.1126/science.1114383) PMID: [16123265](https://pubmed.ncbi.nlm.nih.gov/16123265/)
16. Balaban NQ, Merrin J, Chait R, Kowalik L, Leibler S. Bacterial persistence as a phenotypic switch. *Science*. 2004; 305:1622–1625. doi: [10.1126/science.1099390](https://doi.org/10.1126/science.1099390) PMID: [15308767](https://pubmed.ncbi.nlm.nih.gov/15308767/)
17. Sánchez-Romero MA, Casadesús J. Contribution of phenotypic heterogeneity to adaptive antibiotic resistance. *Proceedings of the National Academy of Sciences*. 2014; 111:355–360. doi: [10.1073/pnas.1316084111](https://doi.org/10.1073/pnas.1316084111)
18. Neildez-Nguyen TMA, Parisot A, Vignal C, Rameau P, Stockholm D, Picot J, et al. Epigenetic gene expression noise and phenotypic diversification of clonal cell populations. *Differentiation*. 2008; 76:33–40. doi: [10.1111/j.1432-0436.2007.00219.x](https://doi.org/10.1111/j.1432-0436.2007.00219.x) PMID: [17825084](https://pubmed.ncbi.nlm.nih.gov/17825084/)
19. Paldi A. Stochastic gene expression during cell differentiation: order from disorder? *Cellular and Molecular Life Sciences*. 2003; 60:1775–1778. doi: [10.1007/s00018-003-23147-z](https://doi.org/10.1007/s00018-003-23147-z) PMID: [14523542](https://pubmed.ncbi.nlm.nih.gov/14523542/)
20. Raj A, van Oudenaarden A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*. 2008; 135:216–226. doi: [10.1016/j.cell.2008.09.050](https://doi.org/10.1016/j.cell.2008.09.050) PMID: [18957198](https://pubmed.ncbi.nlm.nih.gov/18957198/)
21. Kaern M, Elston TC, Blake WJ, Collins JJ. Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics*. 2005; 6:451–464. doi: [10.1038/nrg1615](https://doi.org/10.1038/nrg1615) PMID: [15883588](https://pubmed.ncbi.nlm.nih.gov/15883588/)
22. Magklara A, Lomvardas S. Stochastic gene expression in mammals: lessons from olfaction. *Trends in Cell Biology*. 2013; 23:449–456. doi: [10.1016/j.tcb.2013.04.005](https://doi.org/10.1016/j.tcb.2013.04.005) PMID: [23689023](https://pubmed.ncbi.nlm.nih.gov/23689023/)
23. Munsky B, Trinh B, Khammash M. Listening to the noise: random fluctuations reveal gene network parameters. *Molecular systems biology*. 2009; 5:318. doi: [10.1038/msb.2009.75](https://doi.org/10.1038/msb.2009.75) PMID: [19888213](https://pubmed.ncbi.nlm.nih.gov/19888213/)
24. Wang P, Robert L, Pelletier J, Dang WL, Taddei F, Wright A, et al. Robust growth of *Escherichia coli*. *Current biology*. 2010; 20:1099–1103. doi: [10.1016/j.cub.2010.04.045](https://doi.org/10.1016/j.cub.2010.04.045) PMID: [20537537](https://pubmed.ncbi.nlm.nih.gov/20537537/)
25. Lambert G, Kussell E. Quantifying Selective Pressures Driving Bacterial Evolution Using Lineage Analysis. *Physical Review X*. 2015; 5:011016. doi: [10.1103/PhysRevX.5.011016](https://doi.org/10.1103/PhysRevX.5.011016) PMID: [26213639](https://pubmed.ncbi.nlm.nih.gov/26213639/)
26. Tsukanov R, Reshes G, Carmon G, Fischer-Friedrich E, Gov NS, Fishov I, et al. Timing of Z-ring localization in *Escherichia coli*. *Physical Biology*. 2011; 8:066003. doi: [10.1088/1478-3975/8/6/066003](https://doi.org/10.1088/1478-3975/8/6/066003) PMID: [22015938](https://pubmed.ncbi.nlm.nih.gov/22015938/)
27. Reshes G, Vanounou S, Fishov I, Feingold M. Cell shape dynamics in *Escherichia coli*. *Biophysical Journal*. 2008; 94:251–264. doi: [10.1529/biophysj.107.104398](https://doi.org/10.1529/biophysj.107.104398) PMID: [17766333](https://pubmed.ncbi.nlm.nih.gov/17766333/)

28. Reshes G, Vanounou S, Fishov I, Feingold M. Timing the start of division in *E. coli*: a single-cell study. *Physical Biology*. 2008; 5:046001. doi: [10.1088/1478-3975/5/4/046001](https://doi.org/10.1088/1478-3975/5/4/046001) PMID: [18997273](https://pubmed.ncbi.nlm.nih.gov/18997273/)
29. Roeder AH, Chickarmane V, Chuna A, Obara B, Manjunath BS, Meyerowitz EM. Variability in the control of cell division underlies sepal epidermal patterning in *Arabidopsis thaliana*. *PLOS Biology*. 2010; 8:e1000367. doi: [10.1371/journal.pbio.1000367](https://doi.org/10.1371/journal.pbio.1000367) PMID: [20485493](https://pubmed.ncbi.nlm.nih.gov/20485493/)
30. Zilman A, Ganusov VV, Perelson AS. Stochastic models of lymphocyte proliferation and death. *PLOS ONE*. 2010; 5:e12775. doi: [10.1371/journal.pone.0012775](https://doi.org/10.1371/journal.pone.0012775) PMID: [20941358](https://pubmed.ncbi.nlm.nih.gov/20941358/)
31. Hawkins ED, Markham JF, McGuinness LP, Hodgkin PD. A single-cell pedigree analysis of alternative stochastic lymphocyte fates. *Proceedings of the National Academy of Sciences*. 2009; 106:13457–13462. doi: [10.1073/pnas.0905629106](https://doi.org/10.1073/pnas.0905629106)
32. Stukalin EB, Aifuwa I, Kim JS, Wirtz D, Sun SX. Age-dependent stochastic models for understanding population fluctuations in continuously cultured cells. *Journal of The Royal Society Interface*. 2013; 101. doi: [10.1098/rsif.2013.0325](https://doi.org/10.1098/rsif.2013.0325)
33. Huh D, Paulsson J. Random partitioning of molecules at cell division. *Proceedings of the National Academy of Sciences*. 2011; 108:15004–15009. doi: [10.1073/pnas.1013171108](https://doi.org/10.1073/pnas.1013171108)
34. Gonze D. Modeling the effect of cell division on genetic oscillators. *Journal of Theoretical Biology*. 2013; 325:22–33. doi: [10.1016/j.jtbi.2013.02.001](https://doi.org/10.1016/j.jtbi.2013.02.001) PMID: [23434891](https://pubmed.ncbi.nlm.nih.gov/23434891/)
35. Lloyd-Price J, Tran H, Ribeiro AS. Dynamics of small genetic circuits subject to stochastic partitioning in cell division. *Journal of Theoretical Biology*. 2014; 356:11–19. doi: [10.1016/j.jtbi.2014.04.018](https://doi.org/10.1016/j.jtbi.2014.04.018) PMID: [24768865](https://pubmed.ncbi.nlm.nih.gov/24768865/)
36. Zopf CJ, Quinn K, Zeidman J, Maheshri N. Cell-Cycle Dependence of Transcription Dominates Noise in Gene Expression. *PLOS Computational Biology*. 2013; 9:e1003161. doi: [10.1371/journal.pcbi.1003161](https://doi.org/10.1371/journal.pcbi.1003161) PMID: [23935476](https://pubmed.ncbi.nlm.nih.gov/23935476/)
37. Narula J, Kuchina A, Lee DyD, Fujita M, Süel GM, Igoshin OA. Chromosomal Arrangement of Phosphorelay Genes Couples Sporulation and DNA Replication. *Cell*. 2015; 162:328–337. doi: [10.1016/j.cell.2015.06.012](https://doi.org/10.1016/j.cell.2015.06.012) PMID: [26165942](https://pubmed.ncbi.nlm.nih.gov/26165942/)
38. Schwabe A, Bruggeman FJ. Contributions of Cell Growth and Biochemical Reactions to Nongenetic Variability of Cells. *Biophysical Journal*. 2014; 107:301–313. doi: [10.1016/j.bpj.2014.05.004](https://doi.org/10.1016/j.bpj.2014.05.004) PMID: [25028872](https://pubmed.ncbi.nlm.nih.gov/25028872/)
39. Huh D, Paulsson J. Non-genetic heterogeneity from stochastic partitioning at cell division. *Nature Genetics*. 2011; 43:95–100. doi: [10.1038/ng.729](https://doi.org/10.1038/ng.729) PMID: [21186354](https://pubmed.ncbi.nlm.nih.gov/21186354/)
40. Antunes D, Singh A. Quantifying gene expression variability arising from randomness in cell division times. *Journal of Mathematical Biology*. 2015; 71: 437–463. doi: [10.1007/s00285-014-0811-x](https://doi.org/10.1007/s00285-014-0811-x) PMID: [25182129](https://pubmed.ncbi.nlm.nih.gov/25182129/)
41. Yu J, Xiao J, Ren X, Lao K, Xie XS. Probing Gene Expression in Live Cells, One Protein Molecule at a Time. *Science*. 2006; 311:1600–1603. doi: [10.1126/science.1119623](https://doi.org/10.1126/science.1119623) PMID: [16543458](https://pubmed.ncbi.nlm.nih.gov/16543458/)
42. Paulsson J. Model of stochastic gene expression. *Physics of Life Reviews*. 2005; 2:157–175. doi: [10.1016/j.plev.2005.03.003](https://doi.org/10.1016/j.plev.2005.03.003)
43. Shahrezaei V, Swain PS. Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences*. 2008; 105:17256–17261. doi: [10.1073/pnas.0803850105](https://doi.org/10.1073/pnas.0803850105)
44. Singh A, Hespanha JP. Optimal Feedback Strength for Noise Suppression in Autoregulatory Gene Networks. *Biophysical Journal*. 2009; 96:4013–4023. doi: [10.1016/j.bpj.2009.02.064](https://doi.org/10.1016/j.bpj.2009.02.064) PMID: [19450473](https://pubmed.ncbi.nlm.nih.gov/19450473/)
45. Jia T, Kulkarni RV. Intrinsic Noise in Stochastic Models of Gene Expression with Molecular Memory and Bursting. *Physical Review Letters*. 2011; 106: 058102. doi: [10.1103/PhysRevLett.106.058102](https://doi.org/10.1103/PhysRevLett.106.058102) PMID: [21405439](https://pubmed.ncbi.nlm.nih.gov/21405439/)
46. Alon U. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman and Hall/CRC; 2006.
47. Taniguchi Y, Choi PJ, Li GW, Chen H, Babu M, Hearn J, et al. Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science*. 2010; 329:533–538. doi: [10.1126/science.1188308](https://doi.org/10.1126/science.1188308) PMID: [20671182](https://pubmed.ncbi.nlm.nih.gov/20671182/)
48. Schwanhauser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global quantification of mammalian gene expression control. *Nature*. 2011; 473:337–342. doi: [10.1038/nature10098](https://doi.org/10.1038/nature10098) PMID: [21593866](https://pubmed.ncbi.nlm.nih.gov/21593866/)
49. Swain PS, Elowitz MB, Siggia ED. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*. 2002; 99:12795–12800. doi: [10.1073/pnas.162041399](https://doi.org/10.1073/pnas.162041399)

50. Berg OG. A model for the statistical fluctuations of protein numbers in a microbial population. *Journal of Theoretical Biology*. 1978; 71:587–603. doi: [10.1016/0022-5193\(78\)90326-0](https://doi.org/10.1016/0022-5193(78)90326-0) PMID: [96307](https://pubmed.ncbi.nlm.nih.gov/96307/)
51. Rigney DR. Stochastic model of constitutive protein levels in growing and dividing bacterial cells. *Journal of Theoretical Biology*. 1979; 76:453–480. doi: [10.1016/0022-5193\(79\)90013-4](https://doi.org/10.1016/0022-5193(79)90013-4) PMID: [439915](https://pubmed.ncbi.nlm.nih.gov/439915/)
52. Singh A, Hespanha JP. Stochastic hybrid systems for studying biochemical processes. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. 2010; 368:4995–5011. doi: [10.1098/rsta.2010.0211](https://doi.org/10.1098/rsta.2010.0211)
53. Daigle BJ, Soltani M, Petzold LR, Singh A. Inferring single-cell gene expression mechanisms using stochastic simulation. *Bioinformatics*. 2015; 31:1428–1435. doi: [10.1093/bioinformatics/btv007](https://doi.org/10.1093/bioinformatics/btv007) PMID: [25573914](https://pubmed.ncbi.nlm.nih.gov/25573914/)
54. Lagershausen S. Performance Analysis of Closed Queueing Networks. *Lecture Notes in Economics and Mathematical Systems*. Springer; 2013.
55. Buchholz P, Kriege J, Felko I. Input Modeling with Phase-Type Distributions and Markov Models. Springer; 2014.
56. Hespanha JP, Singh A. Stochastic Models for Chemically Reacting Systems Using Polynomial Stochastic Hybrid Systems. *International Journal of Robust and Nonlinear Control*. 2005; 15:669–689. doi: [10.1002/mc.1017](https://doi.org/10.1002/mc.1017)
57. Singh A, Hespanha JP. Approximate Moment Dynamics for Chemically Reacting Systems. *IEEE Transactions on Automatic Control*. 2011; 56:414–418. doi: [10.1109/TAC.2010.2088631](https://doi.org/10.1109/TAC.2010.2088631)
58. Gomez-Urbe CA, Verghese GC. Mass Fluctuation Kinetics: Capturing Stochastic Effects in Systems of Chemical Reactions through Coupled Mean-Variance Computations. *Journal of Chemical Physics*. 2007; 126:024109. doi: [10.1063/1.2408422](https://doi.org/10.1063/1.2408422) PMID: [17228945](https://pubmed.ncbi.nlm.nih.gov/17228945/)
59. Lee CH, Kim K, Kim P. A moment closure method for stochastic reaction networks. *Journal of Chemical Physics*. 2009; 130:134107. doi: [10.1063/1.3103264](https://doi.org/10.1063/1.3103264) PMID: [19355717](https://pubmed.ncbi.nlm.nih.gov/19355717/)
60. Goutsias J. Classical versus stochastic kinetics modeling of biochemical reaction systems. *Biophysical Journal*. 2007; 92:2350–2365. doi: [10.1529/biophysj.106.093781](https://doi.org/10.1529/biophysj.106.093781) PMID: [17218456](https://pubmed.ncbi.nlm.nih.gov/17218456/)
61. Gillespie CS. Moment-closure approximations for mass-action models. *IET systems biology*. 2009; 3:52–58. doi: [10.1049/iet-syb:20070031](https://doi.org/10.1049/iet-syb:20070031) PMID: [19154084](https://pubmed.ncbi.nlm.nih.gov/19154084/)
62. Soltani M, Vargas-Garcia CA, Singh A. Conditional moment closure schemes for studying stochastic dynamics of genetic circuits. *IEEE Transactions on Biomedical Systems and Circuits*. 2015; 9:518–526. doi: [10.1109/TBCAS.2015.2453158](https://doi.org/10.1109/TBCAS.2015.2453158)
63. Wang H, Yuan Z, Liu P, Zhou T. Division time-based amplifiers for stochastic gene expression. *Molecular BioSystems*. 2015; 11:2417–2428. doi: [10.1039/C5MB00391A](https://doi.org/10.1039/C5MB00391A) PMID: [26178011](https://pubmed.ncbi.nlm.nih.gov/26178011/)
64. Hilfinger A, Paulsson J. Separating intrinsic from extrinsic fluctuations in dynamic biological systems. *Proceedings of the National Academy of Sciences*. 2011; 108:12167–12172. doi: [10.1073/pnas.1018832108](https://doi.org/10.1073/pnas.1018832108)
65. Singh A, Soltani M. Quantifying Intrinsic and Extrinsic Variability in Stochastic Gene Expression Models. *PLOS ONE*. 2013; 8:e84301. doi: [10.1371/journal.pone.0084301](https://doi.org/10.1371/journal.pone.0084301) PMID: [24391934](https://pubmed.ncbi.nlm.nih.gov/24391934/)
66. Shahrezaei V, Ollivier JF, Swain PS. Colored extrinsic fluctuations and stochastic gene expression. *Molecular Systems Biology*. 2008; 4:196. doi: [10.1038/msb.2008.31](https://doi.org/10.1038/msb.2008.31) PMID: [18463620](https://pubmed.ncbi.nlm.nih.gov/18463620/)
67. Scott M, Ingalls B, Kaern M. Estimations of intrinsic and extrinsic noise in models of nonlinear genetic networks. *Chaos*. 2006; 16:026107. doi: [10.1063/1.2211787](https://doi.org/10.1063/1.2211787) PMID: [16822039](https://pubmed.ncbi.nlm.nih.gov/16822039/)
68. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A. Regulation of noise in the expression of a single gene. *Nature Genetics*. 2002; 31:69–73. doi: [10.1038/ng869](https://doi.org/10.1038/ng869) PMID: [11967532](https://pubmed.ncbi.nlm.nih.gov/11967532/)
69. Newman JRS, Ghaemmghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, et al. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature*. 2006; 441:840–846. doi: [10.1038/nature04785](https://doi.org/10.1038/nature04785)
70. Singh A, Razoooky B, Cox CD, Simpson ML, Weinberger LS. Transcriptional Bursting from the HIV-1 Promoter Is a Significant Source of Stochastic Noise in HIV-1 Gene Expression. *Biophysical Journal*. 2010; 98:L32–L34. doi: [10.1016/j.bpj.2010.03.001](https://doi.org/10.1016/j.bpj.2010.03.001) PMID: [20409455](https://pubmed.ncbi.nlm.nih.gov/20409455/)
71. Bar-Even A, Paulsson J, Maheshri N, Carmi M, O’Shea E, Pilpel Y, et al. Noise in protein expression scales with natural protein abundance. *Nature Genetics*. 2006; 38:636–643. doi: [10.1038/ng1807](https://doi.org/10.1038/ng1807) PMID: [16715097](https://pubmed.ncbi.nlm.nih.gov/16715097/)
72. Yunger S, Rosenfeld L, Garini Y, Shav-Tal Y. Single-allele analysis of transcription kinetics in living mammalian cells. *Nature Methods*. 2010; 7:631–633. doi: [10.1038/nmeth.1482](https://doi.org/10.1038/nmeth.1482) PMID: [20639867](https://pubmed.ncbi.nlm.nih.gov/20639867/)
73. Lott SE, Villalta JE, Schroth GP, Luo S, Tonkin LA, Eisen MB. Noncanonical Compensation of Zygotic X Transcription in Early *Drosophila melanogaster* Development Revealed through Single-Embryo RNA-Seq. *PLOS Biology*. 2011; 9:e1000590. doi: [10.1371/journal.pbio.1000590](https://doi.org/10.1371/journal.pbio.1000590) PMID: [21346796](https://pubmed.ncbi.nlm.nih.gov/21346796/)

74. Walker N, Nghe P, Tans SJ. Generation and filtering of gene expression noise by the bacterial cell cycle. *BMC Biology*. 2016; 14:1–10. doi: [10.1186/s12915-016-0231-z](https://doi.org/10.1186/s12915-016-0231-z)
75. Keren L, van Dijk D, Weingarten-Gabbay S, Davidi D, Jona G, Weinberger A, et al. Noise in gene expression is coupled to growth rate. *Genome Research*. 2015; 25:1893–1902. doi: [10.1101/gr.191635.115](https://doi.org/10.1101/gr.191635.115) PMID: [26355006](https://pubmed.ncbi.nlm.nih.gov/26355006/)
76. Siegal-Gaskins D, Crosson S. Tightly Regulated and Heritable Division Control in Single Bacterial Cells. *Biophysical Journal*. 2008; 95:2063–2072. doi: [10.1529/biophysj.108.128785](https://doi.org/10.1529/biophysj.108.128785) PMID: [18469083](https://pubmed.ncbi.nlm.nih.gov/18469083/)
77. Cerulus B, New AM, Pougach K, Verstrepen KJ. Noise and Epigenetic Inheritance of Single-Cell Division Times Influence Population Fitness. *Current Biology*. 2016; 26: 1138–1147. doi: [10.1016/j.cub.2016.03.010](https://doi.org/10.1016/j.cub.2016.03.010) PMID: [27068419](https://pubmed.ncbi.nlm.nih.gov/27068419/)
78. Taheri-Araghi S, Bradde S, Sauls JT, Hill NS, Levin PA, Paulsson J, et al. Cell-Size Control and Homeostasis in Bacteria. *Current Biology*. 2015; 25:385–391. doi: [10.1016/j.cub.2014.12.009](https://doi.org/10.1016/j.cub.2014.12.009) PMID: [25544609](https://pubmed.ncbi.nlm.nih.gov/25544609/)
79. Adiciptaningrum A, Osella M, Moolman MC, Lagomarsino MC, Tans SJ. Stochasticity and homeostasis in the *E. coli* replication and division cycle. *Scientific Reports*. 2015; 5:18261. doi: [10.1038/srep18261](https://doi.org/10.1038/srep18261) PMID: [26671779](https://pubmed.ncbi.nlm.nih.gov/26671779/)
80. Friedman N, Cai L, Xie XS. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Physical Review Letters*. 2006; 97:168302. doi: [10.1103/PhysRevLett.97.168302](https://doi.org/10.1103/PhysRevLett.97.168302) PMID: [17155441](https://pubmed.ncbi.nlm.nih.gov/17155441/)
81. Paulsson J, Ehrenberg M. Random Signal Fluctuations Can Reduce Random Fluctuations in Regulated Components of Chemical Regulatory Networks. *Physical Review Letters*. 2000; 84:5447–5450. doi: [10.1103/PhysRevLett.84.5447](https://doi.org/10.1103/PhysRevLett.84.5447) PMID: [10990965](https://pubmed.ncbi.nlm.nih.gov/10990965/)
82. Sherman MS, Cohen BA. A Computational Framework for Analyzing Stochasticity in Gene Expression. *PLOS Computational Biology*. 2014; 10:e1003596. doi: [10.1371/journal.pcbi.1003596](https://doi.org/10.1371/journal.pcbi.1003596) PMID: [24811315](https://pubmed.ncbi.nlm.nih.gov/24811315/)
83. Golding I, Paulsson J, Zawilski SM, Cox EC. Real-time kinetics of gene activity in individual bacteria. *Cell*. 2005; 123:1025–1036. doi: [10.1016/j.cell.2005.09.031](https://doi.org/10.1016/j.cell.2005.09.031) PMID: [16360033](https://pubmed.ncbi.nlm.nih.gov/16360033/)
84. McAdams HH, Arkin A. Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences*. 1997; 94:814–819. PMID: [9023339](https://pubmed.ncbi.nlm.nih.gov/9023339/)
85. Cai L, Friedman N, Xie XS. Stochastic protein expression in individual cells at the single molecule level. *Nature*. 2006; 440:358–362. doi: [10.1038/nature04599](https://doi.org/10.1038/nature04599) PMID: [16541077](https://pubmed.ncbi.nlm.nih.gov/16541077/)
86. Kumar N, Singh A, Kulkarni RV. Transcriptional bursting in gene expression: analytical results for general stochastic models. *PLOS Computational Biology*. 2015; 11:e1004292. doi: [10.1371/journal.pcbi.1004292](https://doi.org/10.1371/journal.pcbi.1004292) PMID: [26474290](https://pubmed.ncbi.nlm.nih.gov/26474290/)
87. Singh A. Transient Changes in Intercellular Protein Variability Identify Sources of Noise in Gene Expression. *Biophysical Journal*. 2014; 107:2214–2220. doi: [10.1016/j.bpj.2014.09.017](https://doi.org/10.1016/j.bpj.2014.09.017) PMID: [25418106](https://pubmed.ncbi.nlm.nih.gov/25418106/)
88. Osella M, Nugent E, Lagomarsino MC. Concerted control of *Escherichia coli* cell division. *Proceedings of the National Academy of Sciences*. 2014; 111:3431–3435. doi: [10.1073/pnas.1313715111](https://doi.org/10.1073/pnas.1313715111)
89. Robert L, Hoffmann M, Krell N, Aymerich S, Robert J, Doumic M. Division in *Escherichia coli* is triggered by a size-sensing rather than a timing mechanism. *BMC Biology*. 2014; 12:17. doi: [10.1186/1741-7007-12-17](https://doi.org/10.1186/1741-7007-12-17) PMID: [24580833](https://pubmed.ncbi.nlm.nih.gov/24580833/)
90. Kafri R, Levy J, Ginzberg MB, Oh S, Lahav G, Kirschner MW. Dynamics extracted from fixed cells reveal feedback linking cell growth to cell cycle. *Nature*. 2013; 494:480–483. doi: [10.1038/nature11897](https://doi.org/10.1038/nature11897) PMID: [23446419](https://pubmed.ncbi.nlm.nih.gov/23446419/)
91. Ghusinga KR, Vargas-Garcia CA, Singh A. A mechanistic stochastic framework for regulating bacterial cell division. *Scientific Reports*. 2016; 6: 30229. doi: [10.1038/srep30229](https://doi.org/10.1038/srep30229)
92. Amir A. Cell Size Regulation in Bacteria. *Physical Review Letters*. 2014; 112:208102. doi: [10.1103/PhysRevLett.112.208102](https://doi.org/10.1103/PhysRevLett.112.208102)
93. Campos M, Surovtsev IV, Kato S, Paintdakhi A, Beltran B, Ebmeier SE, et al. A Constant Size Extension Drives Bacterial Cell Size Homeostasis. *Cell*. 2014; 159:1433–1446. doi: [10.1016/j.cell.2014.11.022](https://doi.org/10.1016/j.cell.2014.11.022) PMID: [25480302](https://pubmed.ncbi.nlm.nih.gov/25480302/)
94. Padovan-Merhar O, Nair GP, Biaesch AG, Mayer A, Scarfone S, Foley SW, et al. Single Mammalian Cells Compensate for Differences in Cellular Volume and DNA Copy Number through Independent Global Transcription Mechanisms. *Molecular Cell*. 2015; 58:339–352. doi: [10.1016/j.molcel.2015.03.005](https://doi.org/10.1016/j.molcel.2015.03.005) PMID: [25866248](https://pubmed.ncbi.nlm.nih.gov/25866248/)
95. Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. Mammalian genes are transcribed with widely different bursting kinetics. *Science*. 2011; 332:472–474. doi: [10.1126/science.1198817](https://doi.org/10.1126/science.1198817) PMID: [21415320](https://pubmed.ncbi.nlm.nih.gov/21415320/)

96. Brown CR, Mao C, Falkovskaia E, Jurica MS, Boeger H. Linking Stochastic Fluctuations in Chromatin Structure and Gene Expression. *PLOS Biology*. 2013; 11:e1001621. doi: [10.1371/journal.pbio.1001621](https://doi.org/10.1371/journal.pbio.1001621) PMID: [23940458](https://pubmed.ncbi.nlm.nih.gov/23940458/)
97. Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. Stochastic mRNA synthesis in mammalian cells. *PLOS Biology*. 2006; 4:e309. doi: [10.1371/journal.pbio.0040309](https://doi.org/10.1371/journal.pbio.0040309) PMID: [17048983](https://pubmed.ncbi.nlm.nih.gov/17048983/)
98. Hornung G, Bar-Ziv R, Rosin D, Tokuriki N, Tawfik DS, Oren M, et al. Noise-mean relationship in mutated promoters. *Genome Research*. 2012; 22:2409–2417. doi: [10.1101/gr.139378.112](https://doi.org/10.1101/gr.139378.112) PMID: [22820945](https://pubmed.ncbi.nlm.nih.gov/22820945/)
99. Singh A, Razoooky BS, Dar RD, Weinberger LS. Dynamics of protein noise can distinguish between alternate sources of gene-expression variability. *Molecular Systems Biology*. 2012; 8:607. doi: [10.1038/msb.2012.38](https://doi.org/10.1038/msb.2012.38) PMID: [22929617](https://pubmed.ncbi.nlm.nih.gov/22929617/)
100. Dar RD, Razoooky BS, Singh A, Trimeloni T, McCollum J, Cox C, et al. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proceedings of the National Academy of Sciences*. 2012; 109:17454–17459. doi: [10.1073/pnas.1213530109](https://doi.org/10.1073/pnas.1213530109)
101. Corrigan AM, Chubb JR. Regulation of Transcriptional Bursting by a Naturally Oscillating Signal. *Current Biology*. 2014; 24:205–211. doi: [10.1016/j.cub.2013.12.011](https://doi.org/10.1016/j.cub.2013.12.011) PMID: [24388853](https://pubmed.ncbi.nlm.nih.gov/24388853/)
102. Bothma JP, Garcia HG, Esposito E, Schlissel G, Gregor T, Levine M. Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living *Drosophila* embryos. *Proceedings of the National Academy of Sciences*. 2014; 111:10598–10603. doi: [10.1073/pnas.1410022111](https://doi.org/10.1073/pnas.1410022111)
103. Chubb JR, Trcek T, Shenoy SM, Singer RH. Transcriptional Pulsing of a Developmental Gene. *Current Biology*. 2006; 16:1018–1025. doi: [10.1016/j.cub.2006.03.092](https://doi.org/10.1016/j.cub.2006.03.092) PMID: [16713960](https://pubmed.ncbi.nlm.nih.gov/16713960/)
104. Chong S, Chen C, Ge H, Xie XS. Mechanism of Transcriptional Bursting in Bacteria. *Cell*. 2014; 158:314–326. doi: [10.1016/j.cell.2014.05.038](https://doi.org/10.1016/j.cell.2014.05.038) PMID: [25036631](https://pubmed.ncbi.nlm.nih.gov/25036631/)