

# UC Santa Cruz

## UC Santa Cruz Previously Published Works

### Title

From Structural Variation of Gene Molecules to Chromatin Dynamics and Transcriptional Bursting

### Permalink

<https://escholarship.org/uc/item/2zs8g3f3>

### Journal

Genes, 6(3)

### ISSN

2073-4425

### Authors

Boeger, Hinrich  
Shelansky, Robert  
Patel, Heta  
et al.

### Publication Date

2015

### DOI

10.3390/genes6030469

Peer reviewed

Review

## From Structural Variation of Gene Molecules to Chromatin Dynamics and Transcriptional Bursting

Hinrich Boeger \*, Robert Shelansky, Heta Patel and Christopher R. Brown

Department of Molecular, Cell and Developmental Biology, University of California, Santa Cruz, CA 95064, USA; E-Mails: rshelans@ucsc.edu (R.S.); hppatel@ucsc.edu (H.P.); crb2012@gmail.com (C.R.B.)

\* Author to whom correspondence should be addressed; E-Mail: hboeger@ucsc.edu

Academic Editor: Jessica Tyler

Received: 23 April 2015 / Accepted: 24 June 2015 / Published: 30 June 2015

---

**Abstract:** Transcriptional activation of eukaryotic genes is accompanied, in general, by a change in the sensitivity of promoter chromatin to endonucleases. The structural basis of this alteration has remained elusive for decades; but the change has been viewed as a transformation of one structure into another, from “closed” to “open” chromatin. In contradistinction to this static and deterministic view of the problem, a dynamical and probabilistic theory of promoter chromatin has emerged as its solution. This theory, which we review here, explains observed variation in promoter chromatin structure at the level of single gene molecules and provides a molecular basis for random bursting in transcription—the conjecture that promoters stochastically transition between transcriptionally conducive and inconducive states. The mechanism of transcriptional regulation may be understood only in probabilistic terms.

**Keywords:** chromatin dynamics; transcriptional bursting; gene expression noise; stochastic process

---

### 1. Introduction

“There are great solutions, but a final solution does not exist.” (Karl Popper, 1994)

Thirty years ago, it was discovered upon mild digestion of chromatin with DNase I that transcriptional activation of the mouse mammary tumor virus coincided with a marked increase in the endonucleolytic

sensitivity of viral promoter sequences [1]. Similar observations of inducible chromatin changes were subsequently made for many other genes [2]. A classic example is the *PHO5* promoter of budding yeast [3]. The observed changes were generally attributed to a transformation between qualitatively distinct chromatin structures or states. Promoter chromatin was thought to be “closed” under repressing conditions and “open” when induced for transcription. Recent research on the *PHO5* promoter suggests that this static and deterministic view is mistaken. Rather, both activated and repressed promoter chromatin are dynamical systems, describable as stochastic processes that differ only in the numerical values of their probabilistic parameters, and not the range of structural states attainable by the promoter [4]. The transition between repressed and active promoter chromatin is a change in quantity (kinetic parameter values), and not quality (structure).

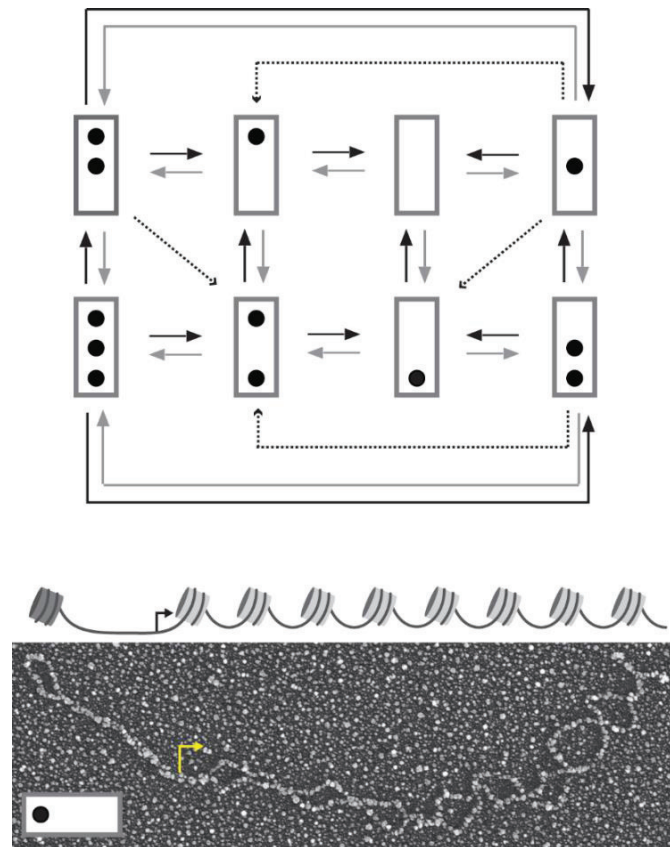
Analysis of single *PHO5* gene molecules by electron microscopy revealed the existence of alternative promoter nucleosome configurations in cells that expressed *PHO5* constitutively [4]. This observation had been anticipated theoretically [5–7]. However, contrary to prior expectation [5], all combinatorial possibilities of occupying three nucleosome positions were observed, including the fully nucleosomal and the nucleosome-free promoter (Figure 1). The same range of configurations was observed in cells that could not activate *PHO5*, but with different relative frequencies; the fully nucleosomal configuration was, by far, the most prominent under repressing conditions, whereas a comparatively small fraction of gene molecules retained all promoter nucleosomes under activating conditions [4].

How can this structural variation be explained? In the following, we review the basic assumptions of such an explanation, its experimental tests by analysis of single gene molecules, and then discuss its implications for gene expression and regulation. Section 3 covers some formal assumptions which underlie most stochastic gene expression models, but often go without mention. A major implication of this theory is transcriptional bursting—the notion that promoters randomly transition between transcriptionally conducive and inconducive states. Excellent reviews of the rich literature on transcriptional bursting have been recently published elsewhere [8,9].

## 2. Ergodic Hypothesis

Two classes of explanatory theories for the molecular variation of promoter chromatin may be distinguished, *viz.* static and dynamical hypotheses. Different promoter nucleosome configurations may correspond, for instance, to distinct states of cellular differentiation with no transitions between cellular states and thus nucleosomal configurations [10]. This static hypothesis has been refuted: nearly all possible pairs of nucleosome configurations have been observed for two *PHO5* promoter copies that shared the same cellular history (“conjugate reporter approach”); the two copies were found to be stochastically independent [11].

It is conceivable that the promoter randomly “freezes” in distinct nucleosome configurations, regardless of the state of its (intracellular) environment. However, such an assumption neither explains variation, nor does it seem physiologically reasonable. This leaves the following alternative explanation: single gene molecules visit each promoter nucleosome configuration over time, in some sequence (see below), with statistically distributed sojourn times in between transitions [12]; in other words, the variation across a population of molecules arises from the stochastic dynamical behavior of each single molecule. We shall refer to this conjecture as “ergodic hypothesis”. (The term is also used in statistical thermodynamics with a similar but more precise meaning, but this is of no concern here.)



**Figure 1.** Transition graph for the stochastic dynamics of *PHO5* promoter nucleosomes (top). The promoter is represented by a box and occupied nucleosome positions as dots. Nucleosome positions from top to bottom are called N-1 to N-3. N-1 contains the transcription start site. Black, gray, and dashed arrows indicate assembly, disassembly, and sliding transitions, respectively. At the bottom is the electron micrograph of an isolated *PHO5* molecule after psoralen-crosslinking, DNA denaturation, and heavy metal shadowing. Positions previously occupied by nucleosomes appear as single-stranded DNA “bubbles” (for psoralen crosslinks linker DNA, but not DNA within the nucleosome core particle). Nucleosome position N-3 is occupied, but not positions N-2 and N-1 (see reconstruction above electron micrograph with promoter nucleosomes in dark gray and nucleosomes over open reading frame in light gray) [4]. A bent arrow indicates the transcription start site.

### 3. Assumption of a Stationary Markov Process

Probabilistic theories make quantitative rather than qualitative predictions. Therefore, they have to be cast in mathematical form to be testable. Formal theories generally require seemingly artificial assumptions for mathematical convenience [13]. As with any other hypothesis, such assumptions are tested for their ability to provide reasonable approximations to the facts. A formalism which has proved surprisingly successful in this regard is the assumption of a time-homogeneous Markov process [13,14]—that the future is conditionally independent of the past, provided the present state is known. In short, it is assumed that the process is “memory-free”.

This lack of memory assumption is formalized as follows. The nucleosome configurations of the promoter are represented as nodes with directed edges (arrows) between them to indicate allowed

transitions between configurations (Figure 1). We call the set of nodes and directed edges a “transition graph” [12]. A stochastic process may be viewed as probability mass flowing between the nodes along the edges of the graph. This flow is described by the *transition functions* of the process  $p_{ji}(t+h, t)$ , which indicate the probability of nucleosome configuration  $j$  at time  $t+h$ , given the promoter’s present configuration,  $i$ , at time  $t$ . The transition functions of a time-homogeneous Markov process depend on the time interval  $h$  and the present value  $i$  alone, and not on  $t$  or the process’ history prior to  $t$ ; that is to say,  $p_{ji}(t+h, t) = p_{ji}(h, 0) \equiv p_{ji}(h)$ . This is a strong assumption; indeed, strong enough to completely specify the mathematical properties of the process [15]. Perhaps the most important of its properties is linearity: The flow of probability per unit time from node  $i$  to node  $j$  linearly depends on the probability of  $i$  (analogous to elementary rate equations in chemistry, but with probability replacing concentration [16]). The rate constant in this linear relationship, *i.e.*, the limit of  $p_{ji}(h)/h$  for  $h \rightarrow 0$ , exists for all  $i \neq j$  [15]. If there is no outgoing edge from  $i$  to  $j$ , this limit is zero. In principle, therefore, the stochastic process has as many rate constants or “kinetic parameters” as there are edges in its transition graph. Another salient property of time-homogeneous Markov processes should be mentioned here: the sojourn times between transitions are exponentially distributed.

One additional assumption is required to make testable predictions: the assumption of a stationary process. Note that a directed graph is called strongly connected if every node can be reached from any other node by a string of one or more edges. It can be proved that a time-homogeneous Markov process on a strongly connected graph tends toward a steady state where the probability currents into and out of each node are balanced; the probability at each node, then, is time-independent and uniquely defined and the process is said to be stationary [15,17]. This assumption may be experimentally satisfied, at least in good approximation [4].

#### 4. Simple Process Assumption

A “good” explanation (theory) is a prohibition—it prohibits certain things to happen—or else it cannot be refuted by experimental observation [18]. The refutability of mathematical models, *e.g.*, a stochastic process, decreases with the number of freely adjustable model parameters (degrees of freedom). Therefore, explanations with fewer degrees of freedom must be preferred over those with more [12].

The following assumption of a “simple process” dramatically limits the degrees of freedom. The assumption is this: the kinetic parameter for the transition between two nucleosome configurations only depends on the “kind of transition” [4,12]. We may distinguish between three different kinds of transitions: assembly, disassembly, and sliding. Accordingly, a transition either adds or removes nucleosomes, or redistributes nucleosomes between promoter positions. This conjecture implies that the process is a time-homogeneous Markov process and it limits the degrees of freedom to two—one of the three kinetic parameter values may be set to 1 on some suitable time scale (say, the parameter for nucleosome assembly).

The simple process assumption reduces the problem of finding an explanatory theory for the relative frequencies of the promoter nucleosome configurations to the comparative analysis of alternative transition graphs. The parameter values for each graph are chosen to maximize the probability of the data. Likewise, between competing transition graphs, the graph that is preferred is that which allows the experimental observations to attain greater probability. Dynamical conjectures, *i.e.*, transition graphs,

may thus be refuted given an alternative conjecture of greater likelihood [4,12]. Graphs that are not strongly connected may be excluded without calculation, because for such graphs only a subset of the experimentally observed configurations would have a probability larger than zero at steady state.

## 5. Promoter Chromatin Dynamics

It may be argued that violations of the simple process conjecture are quite plausible, that the simple process conjecture is too prohibitive or too bold. Yet, remarkably, a simple stationary process was discovered whose theoretical predictions closely correspond to microscopic observations, both for molecules isolated from induced and non-induced cells, albeit with different kinetic parameter values [4].

The transition graph of this process has the following features (Figure 1): each node of the graph has several outgoing edges, *i.e.*, the graph is “branched” [12]. Thus, single gene molecules visit the alternative promoter nucleosome configurations in random sequence. This does not imply, however, that all sequences are allowed; only sequences consistent with the transition graph are possible: nucleosomes are added, removed (by disassembly), or slid only one at a time. Sliding occurs from the central position into the outer positions of the promoter, but not *vice versa*, which explains why this position is more likely to be nucleosome-free than its neighbors. Unidirectional sliding contravenes the demand for detailed balance in thermodynamic equilibrium [19]. It follows that the steady state requires the constant input of free energy, *i.e.*, promoter chromatin is a dissipative system.

Earlier findings corroborated various aspects of this theory. The removal of nucleosomes from the *PHO5* promoter occurs by disassembly, rather than sliding away from the promoter [20]. All known remodeling enzymes accommodate one and only one nucleosome at a time [21–23], and they couple nucleosome transactions to ATP hydrolysis [24]. Removal of nucleosomes from the *PHO5* promoter can be largely dependent on the chromatin remodeling complex SWI/SNF [25]. As predicted by the assumption of a dynamical system of nucleosome removal and reformation, newly synthesized histones preferentially associate with promoter regions of both active and repressed promoters [26].

Our theory implies that ATP hydrolysis drives the process away from (thermodynamic) equilibrium rather than accelerates its approach toward equilibrium. This is a testable assertion: digestion of chromatin with micrococcal nuclease has suggested the presence of a “nucleosome-free region” at many promoters, in a location similar to the central (N-2) nucleosome position of the *PHO5* promoter. The free energy of nucleosome formation on the DNA of such nucleosome-free regions was found to hardly differ from a high affinity sequence (differences did not exceed one  $k_B T$  per molecule), suggesting that many nucleosome-free regions are the product of the continual removal of nucleosomes by chromatin remodelers rather than the result of exclusion by DNA sequence [27]. (This does not refute the well-corroborated conjecture that nucleosomes, once formed, preferentially occupy among closely related positions those that are thermodynamically favored [28]. The concepts of nucleosome occupancy and positioning should not be confused.)

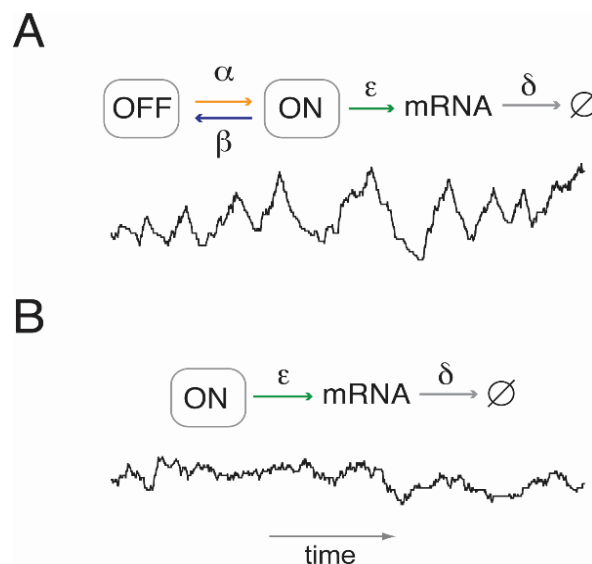
It should be noted here that possession of a nucleosome-free or (better) nucleosome-depleted region is not an intrinsic property of promoters, as the classification of promoters into those with and without such a region suggests [29]; the *PHO5* promoter, on average, would be classified as a promoter with and without a nucleosome-depleted region under-activating and repressing conditions, respectively, and single molecules may defy such a classification under any condition [4].



## 6. Transcriptional Bursting

Nucleosomes are general repressors of transcription, in part because they exclude transcription factors from their DNA binding sequences [30]. If single gene molecules visit alternative nucleosome configurations—some conducive to transcription and others not—in random sequence and with statistically distributed sojourn times in between transitions, then transcription is predicted to occur in stochastic bursts.

The simplest stochastic process that may be invoked to model this behavior assumes two promoter states—one conducive to transcription (ON), the other not (OFF)—between which the promoter jumps randomly, *i.e.*, with statistically distributed sojourn times between jumps (Figure 2) [31,32]. As indicated above, ON and OFF states may each consist of multiple “microstates”, e.g., nucleosome configurations [4]. Both the level of model granularity (the number of microstates) and the presence or absence of irreversible transitions between such microstates affect the sojourn time distribution of the ON and OFF state.



**Figure 2.** Transcript number fluctuations with and without transcriptional bursting. **(A)** An example or “sample path” (black line) for transcription with random transitioning between two promoter states, ON and OFF, that are conducive and inconducive to transcription, respectively. The sample path was calculated using Gillespie’s stochastic simulation algorithm [33]. The level of expression may be controlled by either one of four parameters: burst frequency,  $\alpha$ ; burst duration,  $\beta^{-1}$ ; the initiation rate of transcription in the ON state,  $\epsilon$ ; and the rate per molecule of mRNA degradation,  $\delta$ . **(B)** Sample path in the absence of transcriptional bursting (random birth-and-death process). The transcript level may be controlled by  $\epsilon$  and  $\delta$  alone.

This stochastic model of gene expression raises interesting new problems. On the assumption of a stationary Markov process with two promoter states, the mean number of mRNA molecules or “level of transcription”,  $M_R$ , is given by

$$M_R = \frac{\epsilon}{\delta} \frac{\alpha}{\alpha + \beta}$$

where  $\alpha$  and  $\beta$  are the kinetic parameters for the transition from OFF to ON, and ON to OFF, respectively;  $\varepsilon$  and  $\delta$  are the parameters for mRNA synthesis (in the ON state) and degradation, respectively [32].

In the absence of bursting (*i.e.*,  $\beta = 0$ ), the transcript level may be regulated either by controlling the kinetic parameter for mRNA synthesis,  $\varepsilon$ , or transcript stability,  $\delta$ . Bursting, furthermore, allows for regulation at the level of burst frequency,  $\alpha$ , and burst duration,  $\beta^{-1}$ .

By which of these mechanisms do transcriptional activators control transcription [34]? As detailed below, analysis of this problem, while providing novel insights into the mechanism of transcriptional regulation, also suggested experimental tests of the bursting conjecture and, by implication, of the stochastic process theory of promoter nucleosome dynamics.

## 7. Noise and the Mechanism of Transcriptional Regulation

Stochastic bursting contributes to the variation in the number of mRNA molecules or “noise” of gene expression. The magnitude of this noise may be expressed by the Fano factor,  $\Phi_R$ , the variance over the mean number of transcript molecules. For the stationary Markov process with two promoter states, the Fano factor is given by

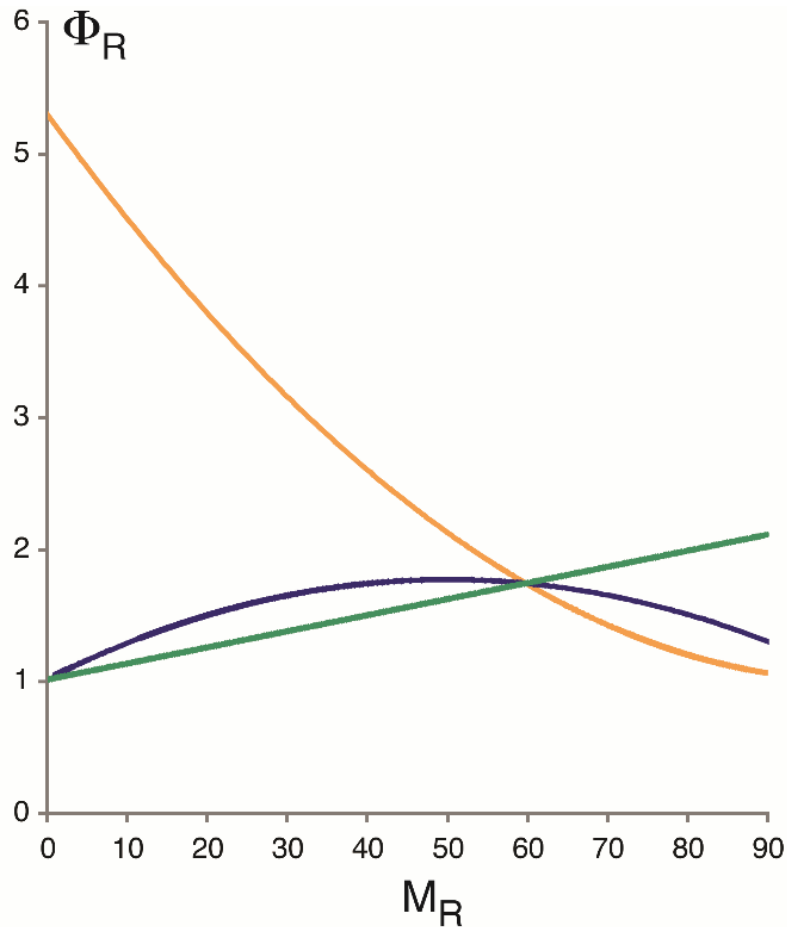
$$\Phi_R = 1 + \frac{\varepsilon\beta}{\kappa(\kappa + \delta)}$$

where  $\kappa = \alpha + \beta$  [32]. In the absence of bursting,  $\Phi_R$  invariably equals 1, regardless of expression level. Deviations from this expectation may be explained on the conjecture of transcriptional bursting. Thus,  $\Phi_R$  increases monotonically with increasing mean ( $M_R$ ) if transcription is controlled by tuning  $\varepsilon$  (Figure 3; green line) but decreases if regulated via  $\alpha$  (yellow line). If regulation occurs at the level of burst duration, then  $\Phi_R$  describes a convex arch when plotted as a function of  $M_R$  (blue line). The same qualitative conclusions hold for protein expression, and for promoter models with several ON and OFF states [4]. The different regulatory mechanisms may be tested, therefore, by measuring the mRNA or protein noise at different levels of expression of the same gene.

To test regulatory hypotheses, two noise components must be distinguished, which together comprise the total noise of gene expression: “extrinsic noise”, which results from the deterministic response of the gene expression process to variation in its intracellular environment (extrinsic variation), and “intrinsic noise”, which cannot thus be explained and must arise, therefore, within the gene expression process and not elsewhere [35].

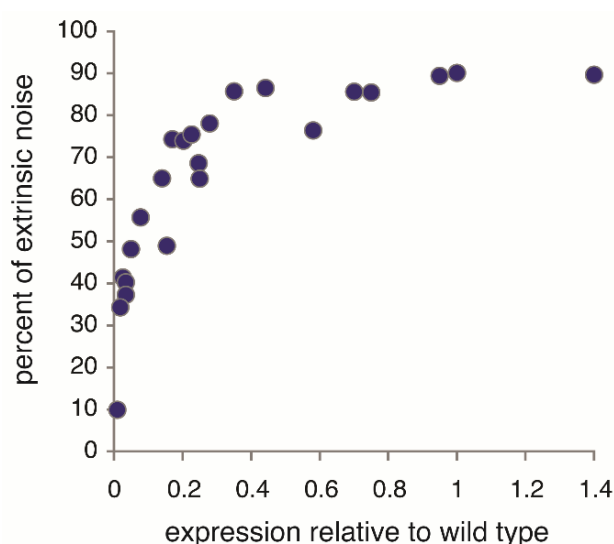
Experimentally, the noise components can be determined by employing the conjugate reporter approach: two gene copies that encode distinguishable gene products, but are otherwise identical, are observed within the same cell [35,36]. The deterministic response to extrinsic variation induces a correlation between the conjugate reporters from which the extrinsic noise is calculated. The remaining noise is intrinsic noise. That extrinsic noise arises from a deterministic response to extrinsic variation may be seen by considering gene expression in the absence of intrinsic noise. In this instance, both conjugate reporter genes would be perfectly correlated; the expression state of one reporter is completely determined by the expression state of the other reporter via their common environment, and hence, by the environment. It may be said, therefore, that extrinsic noise is “informative variation” [36].





**Figure 3.** Fano factor for RNA expression,  $\Phi_R$ , as a function of average transcript number,  $M_R$ . (Likewise, the ‘unit’ of  $\Phi_R$  is number of mRNA molecules.) Curves were calculated on the assumption of a stationary Markov process with two promoter states (Figure 2A), with transcription controlled either by tuning  $\alpha$  (burst frequency, yellow curve),  $\beta^{-1}$  (burst duration, blue curve), or  $\varepsilon$  (the initiation rate of transcription in the ON state, green curve). See main text for equations. The parameters for  $M_R = 60$  are  $\varepsilon = 6.75$  per minute,  $\alpha = 2.18$  per minute,  $\beta = 1.5$  per minute, and  $\delta = 0.066$  per minute and molecule [4]. In the absence of transcriptional bursting,  $\Phi_R$  invariably equals 1.

Under fully activating conditions, extrinsic noise is the dominant noise component of *PHO5* expression; *i.e.*, most variation in *PHO5* expression is explained by extrinsic variation (Figure 4). Surprisingly, the opposite is the case for the structural variation of promoter chromatin. The nucleosome configurations of two *PHO5* promoter copies were found to be stochastically independent [11]; that is to say, the variation in nucleosome configuration could not be explained whatever by extrinsic variation. It had to entirely arise “intrinsically”. The resulting noise in gene expression, then, is intrinsic noise. This was naively assumed previously [4,5,7], without experimental testing. The assumption was bold, for an effect of extrinsic variation on the promoter nucleosome configuration appeared likely, given it is the dominant influence on the noise of *PHO5* expression (Figure 4).



**Figure 4.** Extrinsic protein noise of *PHO5* promoter-driven gene expression as percentage of total noise at different levels of *PHO5* expression. Under fully activating conditions (relative expression of 1), extrinsic noise is the dominant noise component. Noise components were determined by the conjugate reporter approach (see main text) with CFP and YFP (derivatives of the green fluorescence protein or “GFP”) expressed under control of the *PHO5* promoter [7,37]. (These data have not been previously published.) *PHO5* expression was analyzed in *pho80Δ* cells, which express *PHO5* constitutively, and manipulated by mutating the activation domain of *PHO4*, the transcriptional activator of *PHO5* [38], or by *PHO5* mutation in the binding sequences of *PHO4* [7]. Protein expression was normalized by cell size, a significant contributor to extrinsic variation. Without this normalization, the extrinsic noise was close to 97% of the total noise for the wild type, consistent with earlier measurements [37].

Measurements of the intrinsic noise at different levels of *PHO5* protein expression (but not extrinsic or total noise) indicated an increasing Fano factor with decreasing expression level [4,7]; that is to say, the measurements fit the predictions of burst frequency control and not other regulatory conjectures (Figure 3), thus corroborating, by implication, the bursting hypothesis. Regulation of transcription by burst frequency is not limited to *PHO5* [39].

## 8. Accelerated Removal *versus* Exclusion

Burst frequency control has an important implication: the net loss of promoter nucleosomes observed upon transcriptional induction of *PHO5* must be due to accelerated disassembly rather than slowed reformation of nucleosomes—the first increases burst frequency, the latter burst duration. The hypothesis that loss of nucleosomes at the *PHO5* promoter is due to steric exclusion by promoter-bound transcription factors, *i.e.*, decelerated nucleosome formation, was independently refuted [4]. Thus, nucleosomes exclude transcription factors from their DNA binding sites, and not *vice versa*. In this context, it is of great interest that the SWI/SNF remodeling complex can dislodge transcription factors from DNA by catalysis of nucleosome sliding [40]. This may explain why the dwell time of transcription factors at their binding sites *in vivo* is markedly shorter than their dwell time on naked DNA *in vitro* [41].

Consistently, it is widely believed that activators stimulate transcription, among other mechanisms, by the recruitment of chromatin remodelers to promoters to relieve nucleosomal repression—in other words, by accelerating nucleosome removal. Chromatin immunoprecipitation experiments have supported this conjecture [42]. However, such experiments neither revealed the fate of nucleosomes, nor answered the question of whether nucleosome loss was attained by exclusion or accelerated removal; the theoretical and experimental analysis of molecular variation did because it enabled the decision between competing conjectures of cause and effect.

Because nucleosomes may form on any DNA sequence, albeit with different free energies [43], the conjecture of transcriptional stimulation by accelerated nucleosome removal predicts, together with the stochastic process theory of promoter chromatin, that all (eukaryotic) genes are regulated, at least in part, at the level of burst frequency. However, no such consensus has been reached across different studies. Some transgenes were found to be regulated by burst size rather than burst frequency [44]; and transcription of other genes did not appear to occur in bursts at all [45].

It has been argued that most promoter mutations affect the frequency rather than the size of transcriptional bursting [29]. However, this conclusion was based on the fitting of total (and not intrinsic) protein noise data to a Markovian model of gene expression which assumes that promoters are continually transcribed; that transcription does not occur in bursts [46,47]. Thus, the size of translational bursts (the average number of protein molecules translated per mRNA) was inferred from the data, and not transcriptional bursts. This may explain, perhaps, why most promoter mutations did not affect burst size, and those that did frequently generated upstream start codons [29]. The same approach was used to interpret the effects of loss-of-function mutations in genes for histone-modifying enzymes on total protein noise [48]. The implications of the reported data from these studies for transcriptional bursting are uncertain. To the extent that the applied model fit the data, the hypothesis of transcriptional bursting was not required to explain them.

## 9. Further Analysis is Warranted

Transcriptional bursting, the theory that promoters randomly transition between ON and OFF states, is an explanatory conjecture. It has been invoked to explain, essentially, two kinds of observations: a variable Fano factor with changing expression level (Figure 3), and the appearance of short and long gaps between transcription initiation events [4,7,37,39,44,49]. It should be kept in mind that the same observations may perhaps be explained otherwise. Explanatory theories can be tested, but not observed. (Explanations necessarily transcend the observations they explain.) Further testing is warranted: the above discussion shows that the notion of transcriptional bursting is closely linked to hypotheses of transcriptional regulation. This means that tests of the bursting hypothesis are essentially tests of regulatory conjectures. To this end, expression noise should be analyzed at different levels of expression of the same gene. Intrinsic and extrinsic noise must be experimentally distinguished (by employing the conjugate reporter approach), for pulsatile transcription may arise from both extrinsic and intrinsic variation. Genes that are strongly expressed naturally (such as *PHO5*) provide ideal models because their expression level may be manipulated over several orders of magnitude: for instance, by shortening the activation domain of their transcriptional activator, which addresses the question of how transcription is controlled physiologically [7], *viz.* by activators rather than chromosomal location [49,50].

Recent developments have made it possible to monitor nascent transcription at single gene loci in live cells over time [39,45]. This should allow for addressing the critical question of whether transcription occurs in bursts in real time, in instances where it previously had been inferred from population variation; in other words, the ergodic hypothesis may thus be tested. Temporal observations also provide the means to test the proper granularity of promoter models, *i.e.*, the number of promoter states required to explain noise. Thus, deviations from an exponential sojourn time distribution have hinted at two or more OFF states [51,52]. Consistently, it has been argued that several nucleosome configurations of the *PHO5* promoter are inconducive to transcription [4]. This predicts a refractory period prior to the reactivation of transcription, as the promoter passes through two or more OFF states before reentering an ON state.

In principle, it is possible to determine promoter model granularity by recording the autocorrelation function for RNA fluctuations in single cells, which indicates the correlation between two RNA measurements at time points  $t_1 \leq t_2$  as a function of  $\tau = t_2 - t_1$ . The autocorrelation function decreases monotonically from one to zero with increasing  $\tau$ , as the process “forgets” its increasingly distant past. Provided certain mathematical conditions (*viz.* that transcription is a stationary Markov process with a diagonalizable generator for promoter transitions), it can be proved that the autocorrelation function for an  $n$ -promoter state model is the sum of  $n$  exponential functions (H. Boeger, unpublished; but for solution of the analogous problem for transmembrane ion channels see [53]). However, the temporal resolution of current methods is likely insufficient to fully attain this goal; further technological advances are required.

Similar to Fano factor analysis, different regulatory modes of transcriptional bursts may be tested by autocorrelation analysis; different modes predict characteristic shifts in the autocorrelation time with changes in mean expression. Thus, if transcription occurs in bursts and is regulated by burst frequency, the autocorrelation function shifts toward longer times with decreasing mean expression, but it shifts toward shorter times if regulated by burst size. No shift is expected if transcription is controlled by the rate of RNA synthesis without bursting. This analysis may reveal bursting where it previously was not observed because autocorrelation measurements were conducted at one level of expression only. (Transcriptional bursting is masked if the promoter dynamics are significantly faster than the rate of RNA synthesis, if  $\varepsilon \ll \kappa$ ; see equation for  $\Phi_R$  above.)

Other critical tests of the bursting hypothesis are tests of its conjectured molecular basis, which is unknown. (For this reason alone, it should be clear that bursting is not observed, but hypothesized.) The random process theory of promoter chromatin expounded here provides such a conjecture, *i.e.*, a molecular explanation for random bursting. It currently is the most detailed and best-tested explanation, but other mechanisms likely exist: the observed range in nucleosome occupancy did not explain the observed range in *PHO5* transcription, suggesting that the frequency of bursting is not controlled at the level of promoter nucleosomes alone [4]. Furthermore, transgenes in mammalian cells appear to burst at a different time scale than genes in *Drosophila* and yeast [4,39,44,51,54], suggesting different underlying mechanisms. Noise measurements in mammalian cells may be affected by the propensity of transgenes to fall victim to gene silencing. The mechanism remains obscure, but chromatin appears to play a role [55]. The chromosomal integration site of transgenes has been found to affect both burst size and burst frequency [49,50,56]. The molecular basis of such locus effects is unknown. Again, chromatin structure may play a role [49].

## 10. Concluding Remarks

The analysis of both promoter chromatin structure and gene expression at the level of single gene molecules has indicated a surprising degree of intrinsically random behavior. Both appear mechanistically linked: the intrinsically stochastic dynamics of promoter chromatin provide an explanation for intrinsically stochastic bursting in transcription. Other, additional mechanisms for bursting likely exist. The regulation of transcription by control of burst frequency or duration is unique to the probabilistic theory of gene expression; gene regulation may not be understood in deterministic terms. However, the deterministic response of genes to environmental signals is an apparent requirement for biological function in many instances, such as embryonic development [54,57]. How is this determinism reconciled with the intrinsically random behavior of single gene molecules? This problem—of chance and necessity [58], of clouds and clocks, as it were [59]—is not new; it has been pondered by philosophers and biologists alike. Darwin's stroke of genius explained the emergence of purposive or teleonomic properties from random variation by selection. Randomness in gene expression, perhaps, may similarly be understood as a prerequisite for biological function rather than its impediment.

## Acknowledgments

We thank Stacy Harvey, Michael Doody, and Xiaoqing Wang for discussions and comments. This work was supported by funds from the National Science Foundation (NSF #1243957) to Hinrich Boeger.

## Author Contributions

Hinrich Boeger, Robert Shelansky, and Heta Patel wrote the manuscript; Christopher R. Brown generated Figure 1B.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. Zaret, K.S.; Yamamoto, K.R. Reversible and persistent changes in chromatin structure accompany activation of a glucocorticoid-dependent enhancer element. *Cell* **1984**, *38*, 29–38.
2. Gross, D.S.; Garrard, W.T. Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.* **1988**, *57*, 159–197.
3. Almer, A.; Rudolph, H.; Hinnen, A.; Horz, W. Removal of positioned nucleosomes from the yeast *PHO5* promoter upon *PHO5* induction releases additional upstream activating DNA elements. *Embo. J.* **1986**, *5*, 2689–2696.
4. Brown, C.R.; Mao, C.; Falkovskaia, E.; Jurica, M.S.; Boeger, H. Linking stochastic fluctuations in chromatin structure and gene expression. *PLoS Biol.* **2013**, *11*, e1001621.
5. Boeger, H.; Griesenbeck, J.; Kornberg, R.D. Nucleosome retention and the stochastic nature of promoter chromatin remodeling for transcription. *Cell* **2008**, *133*, 716–726.

6. Boeger, H.; Griesenbeck, J.; Strattan, J.S.; Kornberg, R.D. Nucleosomes unfold completely at a transcriptionally active promoter. *Mol. Cell* **2003**, *11*, 1587–1598.
7. Mao, C.; Brown, C.R.; Falkovskaia, E.; Dong, S.; Hrabeta-Robinson, E.; Wenger, L.; Boeger, H. Quantitative analysis of the transcription control mechanism. *Mol. Syst. Biol.* **2010**, *6*, doi:10.1038/msb.2010.83.
8. Sanchez, A.; Golding, I. Genetic determinants and cellular constraints in noisy gene expression. *Science* **2013**, *342*, 1188–1193.
9. Coulon, A.; Chow, C.C.; Singer, R.H.; Larson, D.R. Eukaryotic transcriptional dynamics: From single molecules to cell populations. *Nat. Rev. Genet.* **2013**, *14*, 572–584.
10. Small, E.C.; Xi, L.; Wang, J.P.; Widom, J.; Licht, J.D. Single-cell nucleosome mapping reveals the molecular basis of gene expression heterogeneity. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E2462–E2471.
11. Brown, C.R.; Boeger, H. Nucleosomal promoter variation generates gene expression noise. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 17893–17898.
12. Boeger, H. Nucleosomes, transcription, and probability. *Mol. Biol. Cell* **2014**, *25*, 3451–3455.
13. Feller, W. *An Introduction to Probability Theory and Its Applications*; John Wiley & Sons: Hoboken, NY, USA, 1968.
14. Van Kampen, N.G. *Stochastic Processes in Physics and Chemistry*; Elsevier: Amsterdam, The Netherlands, 2007.
15. Cinlar, E. *Introduction to Stochastic Processes*; Dover Publications: Mineola, NY, USA, 2013.
16. Gunawardena, J.; A linear framework for time-scale separation in nonlinear biochemical systems. *PLoS ONE* **2012**, *7*, e36321.
17. Mirzaev, I.; Gunawardena, J. Laplacian dynamics on general graphs. *Bull. Math. Biol.* **2013**, *75*, 2118–2149.
18. Popper, K.R. *Conjectures and Refutations*; Routledge Classics: New York, NY, USA, 1963.
19. Ahsendorf, T.; Wong, F.; Eils, R.; Gunawardena, J. A framework for modelling gene regulation which accommodates non-equilibrium mechanisms. *BMC Biol.* **2014**, *12*, doi:10.1186/s12915-014-0102-4.
20. Boeger, H.; Griesenbeck, J.; Strattan, J.S.; Kornberg, R.D. Removal of promoter nucleosomes by disassembly rather than sliding *in vivo*. *Mol. Cell* **2004**, *14*, 667–673.
21. Chaban, Y.; Ezeokonkwo, C.; Chung, W.H.; Zhang, F.; Kornberg, R.D.; Maier-Davis, B.; Lorch, Y.; Asturias, F.J. Structure of a RSC-nucleosome complex and insights into chromatin remodeling. *Nat. Struct. Mol. Biol. US* **2008**, *15*, 1272–1277.
22. Racki, L.R.; Yang, J.G.; Naber, N.; Partensky, P.D.; Acevedo, A.; Thomas, J.P.; Cooke, R.; Cheng, Y.; Narlikar, G.J. The chromatin remodeller ACF acts as a dimeric motor to space nucleosomes. *Nature* **2009**, *462*, 1016–1021.
23. Leschziner, A.E.; Saha, A.; Wittmeyer, J.; Zhang, Y.; Bustamante, C.; Cairns, B.R.; Nogales, E. Conformational flexibility in the chromatin remodeler RSC observed by electron microscopy and the orthogonal tilt reconstruction method. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 4913–4918.
24. Becker, P.B.; Workman, J.L. Nucleosome remodeling and epigenetics. *Cold Spring Harb. Perspect. Biol.* **2013**, *5*, doi:10.1101/cshperspect.a017905.
25. Brown, C.R.; Mao, C.; Falkovskaia, E.; Law, J.K.; Boeger, H. *In vivo* role for the chromatin-remodeling enzyme SWI/SNF in the removal of promoter nucleosomes by disassembly rather than sliding. *J. Biol. Chem.* **2011**, *286*, 40556–40565.



26. Dion, M.F.; Kaplan, T.; Kim, M.; Buratowski, S.; Friedman, N.; Rando, O.J. Dynamics of replication-independent histone turnover in budding yeast. *Science* **2007**, *315*, 1405–1408.
27. Lorch, Y.; Maier-Davis, B.; Kornberg, R.D. Role of DNA sequence in chromatin remodeling and the formation of nucleosome-free regions. *Genes Dev.* **2014**, *28*, 2492–2497.
28. Brogaard, K.; Xi, L.; Wang, J.P.; Widom, J. A map of nucleosome positions in yeast at base-pair resolution. *Nature* **2012**, *486*, 496–501.
29. Hornung, G.; Bar-Ziv, R.; Rosin, D.; Tokuriki, N.; Tawfik, D.S.; Tawfik, D.S.; Oren, M.; Barkai, N. Noise-mean relationship in mutated promoters. *Genome Res.* **2012**, *22*, 2409–2417.
30. Mao, C.; Brown, C.R.; Griesenbeck, J.; Boeger, H. Occlusion of regulatory sequences by promoter nucleosomes *in vivo*. *PLoS One* **2011**, *6*, e17521.
31. Peccoud, J.; Ycart, B. Markovian modeling of gene product synthesis. *Theor. Popul. Biol.* **1995**, *48*, 222–234.
32. Paulsson, J. Models of stochastic gene expression. *Phys. Life Rev.* **2005**, *2*, 157–175.
33. Gillespie, D.T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **1976**, *22*, 403–434.
34. Tyagi, S. Tuning noise in gene expression. *Mol. Syst. Biol.* **2015**, *11*, doi:10.15252/msb.20156210.
35. Swain, P.S.; Elowitz, M.B.; Siggia, E.D. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12795–12800.
36. Bowsher, C.G.; Swain, P.S. Identifying sources of variation and the flow of information in biochemical networks. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E1320–E1328.
37. Raser, J.M.; O’Shea, E.K. Control of stochasticity in eukaryotic gene expression. *Science* **2004**, *304*, 1811–1814.
38. Vogel, K.; Horz, W.; Hinnen, A. The two positively acting regulatory proteins *PHO2* and *PHO4* physically interact with *PHO5* upstream activation regions. *Mol. Cell Biol.* **1989**, *9*, 2050–2057.
39. Larson, D.R.; Fritsch, C.; Sun, L.; Meng, X.; Lawrence, D.S.; Singer, R.H. Direct observation of frequency modulated transcription in single cells using light activation. *Elife* **2013**, *2*, e00750.
40. Li, M.; Hada, A.; Sen, P.; Olufemi, L.; Hall, M.A.; Smith, B.Y.; Forth, S.; McKnight, J.N.; Patel, A.; Bowman, G.D.; *et al.* Dynamic regulation of transcription factors by nucleosome remodeling. *Elife* **2015**, *4*, doi:10.7554/eLife.06249.
41. Lickwar, C.R.; Mueller, F.; Hanlon, S.E.; McNally, J.G.; Lieb, J.D. Genome-wide protein-DNA binding dynamics suggest a molecular clutch for transcription factor function. *Nature* **2012**, *484*, 251–255.
42. Cosma, M.P.; Tanaka, T.; Nasmyth, K. Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell* **1999**, *97*, 299–311.
43. Thastrom A; Lowary PT; Widlund HR; Cao H; Kubista M; Widom, J. Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J. Mol. Biol.* **1999**, *288*, 213–229.
44. Raj, A.; Peskin, C.S.; Tranchina, D.; Vargas, D.Y.; Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **2006**, *4*, e309.
45. Larson, D.R.; Zenklusen, D.; Wu, B.; Chao, J.A.; Singer, R.H. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science* **2011**, *332*, 475–478.

46. Friedman, N.; Cai, L.; Xie, X.S. Linking stochastic dynamics to population distribution: An analytical framework of gene expression. *Phys. Rev. Lett.* **2006**, *97*, doi:10.1103/PhysRevLett.97.168302.
47. Thattai, M.; van Oudenaarden, A. Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8614–8619.
48. Weinberger, L.; Voichek, Y.; Tirosh, I.; Hornung, G.; Amit, I.; Barkai, N. Expression noise and acetylation profiles distinguish HDAC functions. *Mol. Cell* **2012**, *47*, 193–202.
49. Dey, S.S.; Foley, J.E.; Limsirichai, P.; Schaffer, D.V.; Arkin, A.P. Orthogonal control of expression mean and variance by epigenetic features at different genomic loci. *Mol. Syst. Biol.* **2015**, *11*, doi:10.15252/msb.20145704.
50. Dar, R.D.; Razoooky, B.S.; Singh, A.; Trimeloni, T.V.; McCollum, J.M.; Cox, C.D.; Simpson, M.L.; Weinberger, L.S. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 17454–17459.
51. Suter, D.M.; Molina, N.; Gatfield, D.; Schneider, K.; Schibler, U.; Naef, F. Mammalian genes are transcribed with widely different bursting kinetics. *Science* **2011**, *332*, 472–474.
52. Harper, C.V.; Finkenstadt, B.; Woodcock, D.J.; Friedrichsen, S.; Semprini, S.; Ashall, L.; Spiller, D.G. Mullins, J.J.; Rand, D.A.; Davis, J.R.E.; *et al.* Dynamic analysis of stochastic transcription cycles. *PLoS Biol.* **2011**, *9*, e1000607.
53. Colquhoun, D.; Hawkes, A.G. Relaxation and fluctuations of membrane currents that flow through drug-operated channels. *Proc. R. Soc. Lond. B Biol. Sci* **1977**, *199*, 231–262.
54. Bothma, J.; Garcia, H.G.; Esposito, E.; Schlissel, G.; Gregor, T.; Levine, M. Dynamic regulation of *eve* stripe 2 expression reveals transcriptional bursts in living *Drosophila* embryos. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 10598–10603.
55. Boeger, H.; Gruss, P. Functional determinants for the tetracycline-dependent transactivator tTA in transgenic mouse embryos. *Mech. Dev.* **1999**, *83*, 141–153.
56. Singh, A.; Razoooky, B.S.; Dar, R.D.; Weinberger, L.S. Dynamics of protein noise can distinguish between alternate sources of gene-expression variability. *Mol. Syst. Biol.* **2012**, *8*, doi:10.1038/msb.2012.38.
57. Atay, O.; Skotheim, J.M. Modularity and predictability in cell signaling and decision making. *Mol. Biol. Cell* **2014**, *25*, 3445–3450.
58. Monod, J. *Chance and Necessity*; Collins: London, UK, 1972.
59. Popper, K.R. *Objective Knowledge*; Clarendon Press: Oxford, UK, 1972.