Research
A SCIENCE PARTNER JOURNAL

RESEARCH ARTICLE

Regulatory Mechanisms for Transcriptional Bursting Revealed by an Event-Based Model

Renjie Wu[1], Bangyan Zhou[1], Wei Wang[1,2*], and Feng Liu[1,2*]

[1]National Laboratory of Solid State Microstructures, Department of Physics, and Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing 210093, P. R. China. [2]Institute for Brain Sciences, Nanjing University, Nanjing 210093, P. R. China.

*Address correspondence to: wangwei@nju.edu.cn (W.W.); fliu@nju.edu.cn (F.L.)

Citation: Wu R, Zhou B, Wang W, Liu F. Regulatory Mechanisms for Transcriptional Bursting Revealed by an Event-Based Model. *Research* 2023;6:Article 0253. https://doi.org/10.34133/research.0253

Submitted 25 May 2023
Accepted 1 October 2023
Published 24 October 2023

Copyright © 2023 Renjie Wu et al. Exclusive licensee Science and Technology Review Publishing House. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY 4.0).

Gene transcription often occurs in discrete bursts, and it can be difficult to deduce the underlying regulatory mechanisms for transcriptional bursting with limited experimental data. Here, we categorize numerous states of single eukaryotic genes and identify 6 essential transcriptional events, each comprising a series of state transitions; transcriptional bursting is characterized as a sequence of 4 events, capable of being organized in various configurations, in addition to the beginning and ending events. By associating transcriptional kinetics with mean durations and recurrence probabilities of the events, we unravel how transcriptional bursting is modulated by various regulators including transcription factors. Through analytical derivation and numerical simulation, this study reveals key state transitions contributing to transcriptional sensitivity and specificity, typical characteristics of burst profiles, global constraints on intrinsic transcriptional noise, major regulatory modes in individual genes and across the genome, and requirements for fast gene induction upon stimulation. It is illustrated how biochemical reactions on different time scales are modulated to separately shape the durations and ordering of the events. Our results suggest that transcriptional patterns are essentially controlled by a shared set of transcriptional events occurring under specific promoter architectures and regulatory modes, the number of which is actually limited.

Introduction

Gene transcription is a complicated dynamic process integrating regulatory signals with genetic information. mRNAs can be produced at constant rates or in short intervals followed by long periods of inactivity, termed transcriptional bursting [1–3]. Although burst traces are qualitatively similar across organisms, the duration of the burst cycle varies remarkably, ranging from minutes to hours [4]. Does this great variability reflect diverse molecular mechanisms or stem from a shared set of highly adjustable molecular events? Similarly, molecular processes involved in transcription typically last from milliseconds to hours [3–5]. How can they be affected by regulators to underlie transcriptional bursting? How can bursting kinetics be modulated across such a wide range of time scales? These topics are fundamental to comprehending the regulation of gene transcription.

Live-cell fluorescence measurements have yielded large amounts of single-cell data [3,6,7]. Durations of transcriptionally active and inactive states and burst size were extracted to characterize transcriptional bursting and expression noise. To theoretically interpret them, phenomenological models have been built upon independent effective processes [8–14], enabling the association between regulatory signals and experimental observables. By incorporating more transcriptional details such as multistep initiation and multistep degradation of mRNA, newly developed models gain more insights into regulatory mechanisms [15–22]. To unravel gene regulation and promoter configurations through the lens of input–output relationships, however, precise knowledge of biochemical reaction mechanisms is essential; yet, such knowledge is often unavailable and must be inferred on the basis of assumptions. This could result in discrepancy between model predictions and experimental data, ignorance of global constraint on transcriptional noise, and inaccurate mapping between rate-limiting steps and effective processes under limiting conditions when deducing the underlying regulatory mechanism [23,24]. A different modeling framework is required.

On the basis of a set of common biochemical reactions essential to mRNA synthesis, a generic model has been developed, which is compatible with those simplified models and fills in their gaps [25–27]. Nevertheless, there always exist additional gene-specific reactions; some common reactions actually comprise multiple steps, and the transition process often occurs through distinct paths, leaving a memory in the transcriptional path [9,28]. While these details could be incorporated to construct more intricate models, we may develop a simpler and more computationally efficient method using queuing theory, which is widely leveraged to understand the behavior of systems with queues. Indeed, queuing models [28–30] were proposed to analyze empirical data on mRNA copy numbers and waiting-time distributions of biochemical reactions while

neglecting the underlying molecular processes. To elucidate the molecular mechanism of gene regulation, we may focus on essential transcriptional events, with each comprising multiple molecular processes, and simplify the transcription into a sequence of events.

Gene transcription is modulated by various regulatory factors, such as transcription factors (TFs) that bind DNA to activate transcription and modifiers and remodelers that induce changes to histone markers and nucleosome arrays to overcome structural barriers [31]. They may exert a wide influence on transcription by altering the duration and directionality of molecular processes [32–34]. To unravel the roles for regulators in gene expression, it is essential to differentiate their impacts on rate-limiting steps and probability flows of event transitions in transcription cycle.

Here, we first propose a network of first-order Markovian biochemical reactions. Because transcriptional bursting is composed of inactive and active phases involving 3 major stages, i.e., chromatin opening for promoter access, assembly of the preinitiation complex (PIC), and mRNA production, all gene states can be divided into 5 sets to separate the stages, with functional state transitions classified into 9 categories. A series of state transitions is combined into an event, and a 6-event model is developed for transcriptional bursting. This model is generic and requires no free parameters. Transcriptional commonality is reflected in the sequential order of the 4 events, while gene specificity is embodied in the duration distributions and ordering of the events. The kinetic features of transcription, such as the mean mRNA number, burst size, burst frequency, and duration of the (in)active phase, are interconnected and functions of 6 model parameters depending on the regulator concentration (i.e., 4 average event durations and 2 repetition probabilities).

On the basis of this event model, we clarify typical features of and intrinsic constraints on transcriptional bursting, illuminate how the regulatory mechanism can be inferred by fitting experimental data at the single-gene and genome levels, and explore the transient transcriptional dynamics upon stimulation. Both numerical and analytical results are presented to provide an integrated view of how reactions on different time scales are modulated to underpin specific burst profiles. The pivotal regulatory modes and their functional implications are elaborated in detail.

Results

Establishment of the event model

Although gene transcription exhibits heterogeneity, its common features allow for the development of a generic model for transcriptional dynamics. To reveal the role for regulatory factors in transcriptional modulation, we need establish a model based on transcriptional events, which is both accurate and simple enough for theoretical analysis.

We first proposed a network model of inducible eukaryotic transcription (Fig. 1A). The transcription begins with activators binding to the promoter/enhancer [35], which promotes the recruitment of cofactors, chromatin remodeling [34,36] and histone modifications [31,37]. General TFs (GTFs; including TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH) and RNA polymerase II (Pol II) are then recruited to the core promoter, forming the PIC. The enhancer-bound activator and the PIC are connected by the mediator complex [38,39]

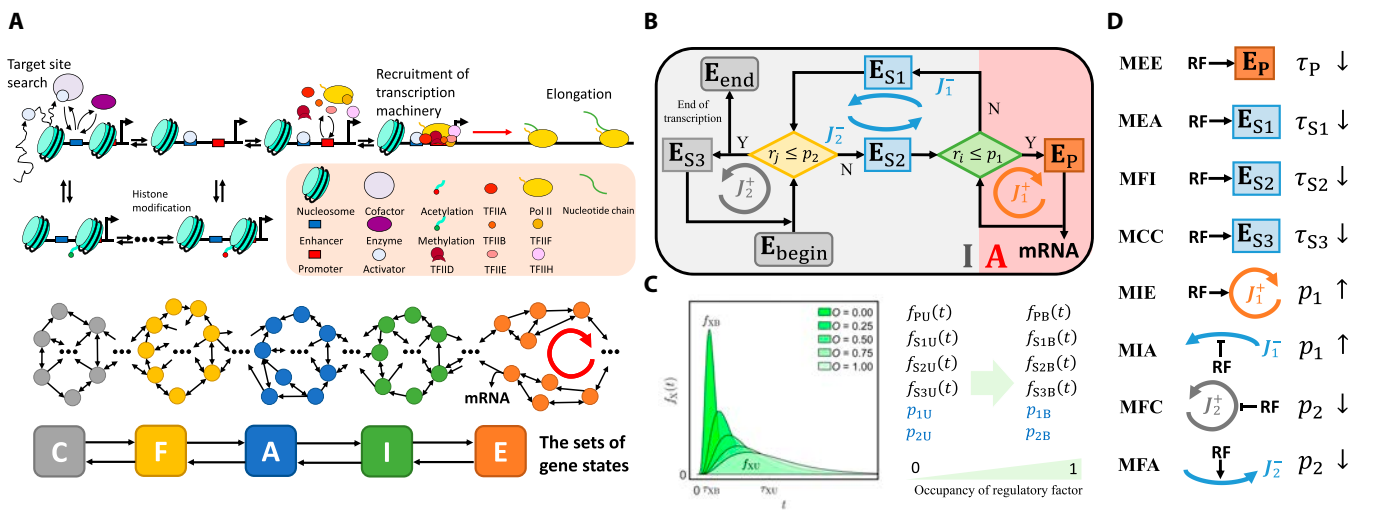


Fig. 1. Model of the transcriptional cycle and transcriptional regulation. (A) Schematic of biochemical reactions underlying the gene states (top) and state transitions (middle; large circles denote gene states, while small circles denote omitted states). According to whether the core promoter is open or activated, all states are divided into 5 subsets (*C*, *F*, *A*, *I*, and *E*; bottom). (B) Flow-process diagram of the event-based model simplified from (A). There are 6 non-Markovian events (rectangles) and 2 checkpoints (diamonds). "Y" and "N" separately represent whether the criterion is satisfied or not; r_i and r_j are random numbers from the uniform distribution on the unit interval. Transcription begins with E_{begin}, goes through E_{S2}, E_P, E_{S1}, and E_{S3}, and ends with E_{end}. The ordering of the events is determined by p_1 and p_2, which are decided by the event flows (J_1^+, J_1^-, J_2^+, and J_2^-). mRNAs are synthesized in the active phase (*A*, red area) comprising a series of E_P, while no mRNA is produced in the inactive phase (*I*, gray area) consisting of sequences of E_{S1}, E_{S2}, and E_{S3}. (C) Gene transcription is simulated based on the distribution functions f_X ($X = P, S1, S2, S3$), p_1 and p_2, which depend on the regulator occupancy rate (*O*). "B" denotes regulators bound and "U" regulator unbound. The left panel schematically shows f_X for different *O* values. When *O* rises from 0 to 1, (f_{XU}, p_{1U}, p_{2U}) turns into (f_{XB}, p_{1B}, p_{2B}). (D) Eight basic regulatory modes in which regulator binding can promote transcription by decreasing the mean event duration τ_X and p_2 or by increasing p_1. Only 1 of the 6 bursting parameters is adjustable in each mode. The first "M" in the leftmost column stands for modulation, while "RF" stands for regulatory factor. The black arrows signify that RF binding accelerates the events or increases the event flows (J_1^+ or J_2^-), whereas the black lines with a flat head signify that RF binding decreases the event flows (J_1^- or J_2^+).

via enhancer-promoter proximity. The PIC turns into the open complex after the DNA template strand is positioned into the active center cleft of Pol II. With the preparations completed, Pol II gets away into elongation to synthesize nucleotide chains [40], while the remaining scaffold complex (composed of activators, TFIIA, TFIID, TFIIE, TFIIH, and mediator) on the promoter facilitates transcriptional reinitiation. During early transcript synthesis, the early elongation complex (EEC) retains a measurable tendency to undergo backtracking until Pol II pauses about 30-base-pair downstream of the transcription start site (i.e., promoter-proximal pausing) [41] and then is released for productive elongation, restricting the frequency of transcriptional initiation [42].

Each node in the network represents a gene state, and the transition between adjacent states is a Markov process (Fig. 1A). The same reactions may occur at diverse sites around the promoter, and some reactions are not exclusive with each other, leading to various transition paths between 2 states. All states are divided into 5 sets (Table 1), separately responsible for maintaining chromatin inaccessibility (C), chromatin remodeling and histone modification when the core promoter is a nucleosome-free region (F), PIC assembly (A), initiation of mRNA synthesis by Pol II and promoter escape (I), and promoter-proximal elongation and pausing of Pol II (E). When a Pol II enters productive elongation (P), the gene state returns to set I, and another Pol II is recruited for transcription reinitiation. Each state can transition to a state in the same or adjacent set.

This network model is too complicated to be analyzed easily. To simplify it, we exploited queueing theory [28,43] to turn the network model into an event-based model. First, 9 types of functional transitions are specified: C_F, F_C, F_A, A_F, A_I, I_A, I_E, E_I, and E_I^P, each comprising a series of reactions with the same function (Table 2). $X_{X'}$ ($X, X' \in C, F, A, I$, or E) represents the process where the initial state stems from set X and all state transitions occur in X until the last state belongs to X'. E_I^P involves Pol II released into productive elongation, such that a nascent full-length transcript can be produced, whereas E_I does not. The time scales of (I_E, E_I, E_I^P), (F_A, A_I, I_A, A_F), and (C_F, F_C) are in the order of seconds to minutes, minutes, and minutes to hours, respectively.

Second, we defined 6 events, a detailed description of which is presented in Fig. S1. E_P refers to a series of I_E and E_I transitions plus the ending transition E_I^P, resulting in synthesis of mRNA transcripts. E_{S1} comprises a series of I_E and E_I plus I_P, responsible for disassembly of the scaffold complex. E_{S2} consists of a series of A_F and F_A plus A_I, enabling assembly of the PIC. E_{S3} comprises a series of A_F, F_C, C_F, and F_A, maintaining the inaccessibility of GTFs to the core promoter and delaying the entry into E_{S2}. As the beginning event of the entire transcriptional process, E_{begin} comprises a series of C_F and F_C plus F_A, contributing to chromatin opening; as the ending event, E_{end} is composed of a series of A_F and F_A plus F_C, resulting in termination of transcription. Notably, each event is not fixed in composition but amenable to modulation; E_{begin} and E_{end} are governed by the initial and final states in set C, respectively. A single transcriptional burst undergoes 2 stages: The inactive phase refers to no mRNA production, including E_{S1}, E_{S2}, and E_{S3}, while the active phase corresponds to mRNA synthesis, comprising a sequence of E_P.

Last, we developed an event-based model for transcription (Fig. 1B). E_{begin} first appears, and then E_{S2} follows, or E_{S3} occurs once or repetitively until E_{S2} arises. Subsequently, different paths lead to the (repetitive) occurrence of E_P, with a burst of mRNAs generated. After bursting, E_{S1} emerges, then E_{S2} appears with or without E_{S3} preceding, and a new burst may ensue. Notably, there may exist a refractory period in E_{S1} where immediate reactivation of transcription is prohibited. The above process repeats until E_{S1} and E_{end} appear successively, with the gene returning to the silent state. Thus, the transcription is characterized by a sequence of ordered events. The process of mRNA splicing, nuclear export, and degradation of mature mRNA is simplified as an effective Poisson process with rate constant $1/\tau_m$. τ_m is set to 5 min unless specified otherwise. The biochemical master equation describing mRNA production is presented in Text S1.

Given the stochasticity in transcription, we focused on the distribution function f_i of the duration of event i (i = P, S1, S2, S3) in steady state with the mean τ_i (Text S2 and Figs. S2 and S3). The probability of repeated occurrence of E_P (E_{S3}) is p_1 (p_2) (Fig. S4). In general, f_P, f_{S1}, f_{S2}, and f_{S3} are primarily controlled by Pol II-dependent reactions, disassembly of the scaffold complex, GTF-dependent reactions, and nucleosome-associated reactions near the core promoter, respectively. f_i is an exponential or a unimodal function in most cases [11,44]. τ_i is mainly determined by rate-limiting steps, whose durations in a burst are the main constituent of the burst period. p_1 is elevated via accelerating such reactions as recruitment, phosphorylation, and release of Pol II, which promote the transition of $A \rightarrow I \rightarrow E \rightarrow P$, while p_2 is boosted by accelerating reactions that promote $I \rightarrow A \rightarrow F \rightarrow C$.

Once f_i, p_1 and p_2 were given or derived from experimental data, numerical simulation was performed to depict mRNA

Table 1. Gene state sets.

State set	Description of state sets
C	States in which the access of GTFs to the core promoter is inhibited by occluded canonical nucleosomes or other macromolecules.
F	States in which the core promoter is a nucleosome free region and ready for GTF recruitment.
A	States in which the core promoter is occupied by GTFs but the PIC has not yet formed completely.
I	States in which the scaffold complex is at the promoter without the EEC formed.
E	States in which Pol II undergoes promoter-proximal elongation until it enters productive elongation.

Table 2. Reactions in functional transitions.

Transition	Reactions during state transitions	
C_F	Reactions that make the core promoter nucleosome-free to promote the recruitment of GTFs, e.g., removal of nucleosomes at the promoter, altering the accessibility of DNA on the surface of nucleosomes by remodelers, replacement with certain histone variants, or destabilization of internucleosome contacts by acetylation of lysine residues. Reactions that change the microenvironment when the promoter is occluded by canonical nucleosomes or other macromolecules different from GTFs, e.g., formation or disruption of topologically associating domain boundaries and changes in distribution modes and interactions of chromatin marks [36,75].	
F_C	Reactions that change the microenvironment when the promoter is a nuclosome-free region [75].	The core promoter is wrapped around histone octamers to form canonical nucleosomes or occupied by other macromolecules [36].
F_A		Binding of GTFs, especially TFIID [75–77].
A_F	Reactions that change the microenvironment when the promoter is occupied by GTFs; enhancer-promoter proximity (hubs or looping) [75].	Dissociation of GTFs bound to the core promoter [75–77].
A_I		Formation of the PIC scaffold [78–80].
I_A	Reactions that change the microenvironment when the PIC scaffold is formed without the EEC on the DNA template; enhancer-promoter proximity [75,81].	Dissociation of the PIC scaffold components [78–80].
I_E		Stabilizing the initially transcribing complex to form the EEC via promoter escape, e.g., the B-finger of TFIIB inserts into the polymerase active site to complete escape commitment [81].
E_I	Reactions that change the microenvironment when the EEC begins elongation; enhancer-promoter proximity [75,81].	Transcript slippage and backtracking [81].
E_I^P		Transcript synthesis; recovery of the elongation competency of arrested Pol II via TFIIS; release of paused Pol II to enter productive elongation [81].

production using the Gillespie algorithm (Text S3). The default setting of f_i, p_1, and p_2 is denoted as (f_{iU}, p_{1U}, p_{2U}), corresponding to the situation without regulators binding to cognate sites, and is determined by the gene specificity [45] and local environment; it changes into (f_{iB}, p_{1B}, p_{2B}) with bound regulators (Fig. 1C). At various regulator concentrations, its rate of occupancy at the regulatory site varies between 0 and 1, and the resulting (f_i, p_1, p_2) lies between (f_{iU}, p_{1U}, p_{2U}) and (f_{iB}, p_{1B}, p_{2B}). That is, the regulation of transcription is mapped to changes in f_i, p_1, and p_2. We found that the main conclusions do not rely heavily on the concrete form of f_i provided that it is exponential or unimodal (Figs. S5 and S6). All the results presented here are based on the exponential distributions. Since an exponential distribution is determined by its mean value, the roles for regulators in gene transcription are classified on the basis of their influence on τ_i, p_1, and p_2. There exist 255 ($2^8 - 1$) regulatory modes, and Fig. 1D illuminates 8 basic modes, under each of which regulators affect τ_P, τ_{S1}, τ_{S2}, τ_{S3}, p_1 (through J_1^+ or J_1^-), or p_2 (through J_2^+ or J_2^-) alone.

Compared with previous models where independent quantities characterized experimental observables, the current model focuses on a sequence of ordered events and is more concerned with the impact of biochemical reactions on the event durations and transitions. Thus, the quantities such as the transcription rate constant ($1/\tau_P$), the duration of the active phase [$\tau_A = \tau_P/(1 - p_1)$] and of the inactive phase ($\tau_I = \tau_S/p_1$) are all correlated (see below), which will have important implications as shown later.

Regulatory modes promoting transcriptional sensitivity and specificity

Gene transcription is mediated by a multitude of regulatory factors, such as TFs and cofactors. Without loss of generality, they are assumed to function by binding their cognate sites to affect biochemical reactions. For the general multibinding site scenario, the occupancy rate of regulators can be expressed as $\frac{[R]^{n_H}}{[R]^{n_H} + K_d^{n_H}}$, where $[R]$ denotes the regulator concentration, K_d is a constant, and Hill coefficient n_H could be a noninteger. As shown in Texts S4 and S5, if $[R']$ and K' are separately substituted for $[R]^{n_H}$ and $K_d^{n_H}$, either parts of original analytical expressions are identical to those in the case of $n_H = 1$, or relationships between specific quantities can be preserved. Consequently, the occupancy rate can simply be written as $\frac{[R]}{K_d + [R]}$, corresponding to the simplest single-binding site case, for the purpose of simplifying the exploration of gene regulation.

As mentioned above, bursting parameters can be modulated to various extents under 255 regulatory modes. In mode X, the mean transcription rate υ in steady state is approximated as

$$\upsilon \approx \upsilon_0 \frac{\beta_X^n + [R]^n}{\Omega_X^n + [R]^n}, \quad (1)$$

where Ω_X refers to $[R]$ at which υ is halfway between its maximum and minimum, β_X is associated with the basal transcription

without bound regulators, and n is a fitting number (see Text S6 for details). K_d/Ω_X reflects the sensitivity of transcription to changes in $[R]$ (Fig. 2A); high sensitivity allows for efficient regulation of transcription [46]. $F_\upsilon = \upsilon([R] = \infty)/\upsilon([R] = 0) = (\Omega_x/\beta_x)^n$ measures how adjustable the mean transcription rate is.

Regulators actually bind to both cognate and nontarget sites, and the latter may also induce mRNA production. To quantify the difference between the 2 scenarios, we introduced the specificity S defined as the ratio of the average expression resulting from specific binding of regulators to that resulting from nonspecific binding. S is a unimodal function of $[R]$ for nonzero β (Fig. 2B); its maximum S_m approximately equals F_υ, given that cognate binding has a much higher affinity than nontarget binding (see Text S6 for details). Collectively, K_d/Ω_X and S_m provide crucial information about transcriptional regulation.

Among 8 basic regulatory modes, $\Omega_X < K_d$ is realizable for X = MEE/EA/FI/CC/FA/IE, whereas $\Omega_X > K_d$ for X = MIA/FC (Fig. 1D and Table 3). In the former, the transcription is activated by reducing τ_P, τ_{S1}, τ_{S2}, τ_{S3}, p_2 (via J_2^-) and increasing p_1 (via J_1^+), which are achieved mainly by accelerating E_I^P and I_E, I_A and E_I, A_I and F_A, F_A and C_F, F_A and A_I, and I_E and E_I^P, respectively. In the latter, transcription is activated by increasing p_1 (via J_1^-) and reducing p_2 (via J_2^+), which can be acquired by slowing down I_A and E_I and A_F and F_C, respectively. Therefore, with $\Omega_X < K_d$, the transcriptional rate reaches its maximum before the regulator occupancy saturates, while the regulator binding promotes transcription by speeding up some state transitions, which involve the recruitment of TFs, enzymes, and Pol II and are part of E_I^P, I_E, A_I, F_A, or C_F. By contrast, with $\Omega_X > K_d$, the regulator binding contributes to transcription by slowing down some state transitions, which are engaged in E_I, I_A, A_F, or F_C. Notably, most reactions involved in E_I, I_A, A_F, and F_C are irrespective of regulator binding, like disintegration of the enhancer-promoter looping and scaffold complex. Together,

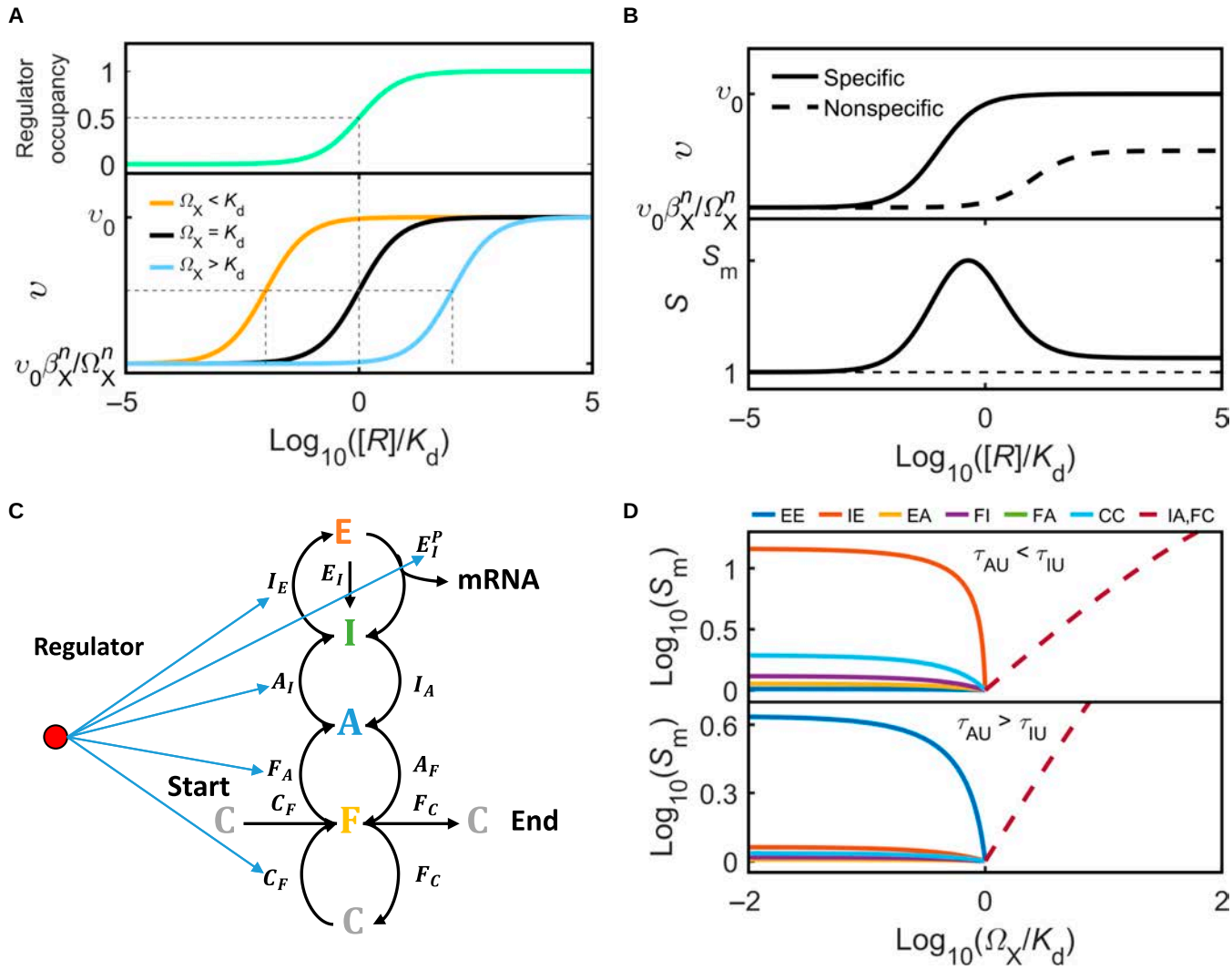


Fig. 2. Sensitivity to changes in regulator concentration ($[R]$) and resistance against crosstalk. (A) Schematic of the mean regulator occupancy rate (top) and steady-state transcription rate (bottom) versus the normalized concentration of regulators. K_d denotes the equilibrium dissociation constant for regulator binding, and K_d/Ω_X reflects the transcriptional sensitivity. (B) Schematic of the mean transcription rate (top) and specificity S (bottom) versus $[R]/K_d$. The equilibrium dissociation constant is K_d and $100K_d$, respectively, for cognate and nontarget binding. (C) Regulation of E_I^P, I_E, A_I, F_A, or C_F promotes the gene's sensitivity to the regulator when its binding facilitates state transitions. (D) S_m versus Ω_X/K_d under different regulatory modes. The default parameters are $\tau_{PU} = 1$ min, $\tau_{S1U} = 5$ min, $\tau_{S2U} = 10$ min, $\tau_{S3U} = 20$ min, $p_{2U} = 0.5$, and $p_{1U} = 0.5$, leading to $\tau_{AU} < \tau_{IU}$ (top); $\tau_{PU} = 10$ min, $\tau_{S1U} = 2$ min, $\tau_{S2U} = 5$ min, $\tau_{S3U} = 10$ min, $p_{2U} = 0.5$, and $p_{1U} = 0.9$, leading to $\tau_{AU} > \tau_{IU}$ (bottom). τ_{PB}, τ_{S1B}, τ_{S2B}, τ_{S3B}, p_{1B}, and p_{2B} are decided by Ω_X/K_d and regulatory modes (see Table 3).

Table 3. Gene regulatory function.

Mode	Changed parameter	Parameter change	υ_{max}	β	Ω
MEE	τ_P	$\tau_{PB} = \varepsilon_{EE}\tau_{PU}\ \varepsilon_{EE} < 1$	$\frac{1}{1-p_1}\frac{1}{\varepsilon_{EE}\tau_{AU}+\tau_I}$	$\varepsilon_{EE}K_d$	$\frac{\tau_{AU}+\tau_I}{\tau_{AU}+\frac{\tau_I}{\varepsilon_{EE}}}K_d < K_d$
MEA	τ_{S1}	$\tau_{S1B} = \varepsilon_{EA}\tau_{S1U}\ \varepsilon_{EA} < 1$	$\frac{1}{1-p_1}\frac{1}{\tau_A+\varepsilon_{EA}\frac{1}{p_1}(\tau_{S1U}+\tau_{S2})+\tau_{I2}}$	$\varepsilon_{EA}K_d$	$\frac{\varepsilon_{EA}\tau_{I1U}+\varepsilon_{EA}(\tau_A+\tau_{I2})}{\varepsilon_{EA}\tau_{I1U}+(\tau_A+\tau_{I2})}K_d < K_d$
MFI	τ_{S2}	$\tau_{S2B} = \varepsilon_{FI}\tau_{S2U}\ \varepsilon_{FI} < 1$	$\frac{1}{1-p_1}\frac{1}{\tau_A+\varepsilon_{FI}\frac{1}{p_1}(\tau_{S1}+\tau_{S2U})+\tau_{I2}}$	$\varepsilon_{FI}K_d$	$\frac{\varepsilon_{FI}\tau_{I1U}+\varepsilon_{FI}(\tau_A+\tau_{I2})}{\varepsilon_{FI}\tau_{I1U}+(\tau_A+\tau_{I2})}K_d < K_d$
MCC	τ_{S3}	$\tau_{S3B} = \varepsilon_{CC}\tau_{S3U}\ \varepsilon_{CC} < 1$	$\frac{1}{1-p_1}\frac{1}{\tau_A+\tau_{I1}+\varepsilon_{CC}\tau_{I2U}}$	$\varepsilon_{CC}K_d$	$\frac{\varepsilon_{CC}(\tau_A+\tau_{I1})+\varepsilon_{CC}\tau_{I2U}}{\tau_A+\tau_{I1}+\varepsilon_{CC}\tau_{I2U}}K_d < K_d$
MIE	p_1 via J_1^+	$p_{1B} = \varepsilon_{IE}p_{1U}\ \varepsilon_{IE} > 1$	$\frac{1}{1-p_{1U}}\frac{1}{\tau_{AU}+\frac{1-\varepsilon_{IE}p_{1U}}{\varepsilon_{IE}(1-p_{1U})}\tau_{IU}}$	$\frac{1-\varepsilon_{IE}p_{1U}}{\varepsilon_{IE}(1-p_{1U})}K_d$	$\frac{\tau_{AU}+\tau_{IU}}{\frac{\varepsilon_{IE}(1-p_{1U})}{1-\varepsilon_{IE}p_{1U}}\tau_{AU}+\tau_{IU}}K_d < K_d$
MIA	p_1 via J_1^-	$p_{1B} = \varepsilon_{IA}p_{1U}\ \varepsilon_{IA} > 1$	$\frac{1}{1-p_{1U}}\frac{1}{\tau_{AU}+\frac{1-\varepsilon_{IA}p_{1U}}{\varepsilon_{IA}(1-p_{1U})}\tau_{IU}}$	K_d	$\frac{\tau_{AU}+\tau_{IU}}{\tau_{AU}+\frac{1-\varepsilon_{IA}p_{1U}}{\varepsilon_{IA}(1-p_{1U})}\tau_{IU}}K_d > K_d$
MFC	p_2 via J_2^+	$p_{2B} = \varepsilon_{FC}p_{2U}\ \varepsilon_{FC} < 1$	$\frac{1}{1-p_1}\frac{1}{\tau_A+\tau_{I1}+\frac{\varepsilon_{FC}(1-p_{2U})}{1-\varepsilon_{FC}p_{2U}}\tau_{I2U}}$	K_d	$\frac{\tau_A+\tau_{I1}+\tau_{I2U}}{\tau_A+\tau_{I1}+\frac{\varepsilon_{FC}(1-p_{2U})}{1-\varepsilon_{FC}p_{2U}}\tau_{I2U}}K_d > K_d$
MFA	p_2 via J_2^-	$p_{2B} = \varepsilon_{FA}p_{2U}\ \varepsilon_{FA} < 1$	$\frac{1}{1-p_1}\frac{1}{\tau_A+\tau_{I1}+\frac{\varepsilon_{FA}(1-p_{2U})}{1-\varepsilon_{FA}p_{2U}}\tau_{I2U}}$	$\frac{\varepsilon_{FA}(1-p_{2U})}{1-\varepsilon_{FA}p_{2U}}K_d$	$\frac{\tau_A+\tau_{I1}+\tau_{I2U}}{(\tau_A+\tau_{I1})\frac{1-\varepsilon_{FA}p_{2U}}{\varepsilon_{FA}(1-p_{2U})}+\tau_{I2U}}K_d < K_d$

Note: $\upsilon = \upsilon_{max}(\beta+[R])/(\Omega+[R])$, $\tau_A = \tau_P/(1-p_1)$, $\tau_I = \tau_{I1}+\tau_{I2}$, $\tau_{I1} = (\tau_{S1}+\tau_{S2})/p_1$, $\tau_{I2} = p_2\tau_{S3}/[p_1(1-p_2)]$. "B" and "U" denote the cases with and without bound regulators, respectively. No "U" emerges if regulator binding has no influence. ε_X (X = EE, EA, FI, CC, IE, IA, FC, FA) is the ratio of the corresponding parameter with bound regulators to that without bound regulators.

the regulation of E_I^P, I_E, A_I, F_A, or C_F may be engaged more frequently to gain high transcriptional sensitivity (Fig. 2C).

The type of regulatory mode determines the trend in S_m changing with Ω_X. For $\Omega_X < K_d$, S_m drops with increasing Ω_X/K_d (Fig. 2D); S_m varies sharply around $\Omega_X/K_d = 1$ under MIE (MEE) for $\tau_{AU} < \tau_{IU}$ ($\tau_{AU} > \tau_{IU}$). The maximum of S_m is achieved via MIE for $\tau_{AU} < \tau_{IU}$ and via MEE otherwise, suggesting that regulating E_I^P or I_E is a desirable mode. For $\Omega_X > K_d$, S_m rises linearly with increasing Ω_X/K_d (on a log–log scale); obtaining a large S_m through MIA/FC requires that the binding sites should always be occupied, which is biologically implausible.

According to the event model, we directly infer that the regulation of E_I^P, I_E, A_I, F_C, or C_F can enhance the transcriptional sensitivity while promoting mRNA synthesis. The mode modulating E_I^P and/or I_E has a marked role in enlarging the transcriptional adjustability and reducing the crosstalk effects resulting from noncognate regulator binding.

Major regulatory modes underlying the observed transcriptional noise

Transcription is a complex stochastic process; for simplicity, we only explored intrinsic noise in transcription, neglecting variations in parameters between cells or across time (i.e., extrinsic noise). The relationship between the mean ($\langle m\rangle$) and variance (σ_m^2) of mRNA copy number or Fano factor ($F = \sigma_m^2/\langle m\rangle$) is a key representation of transcriptional regulation. $\langle m\rangle$, σ_m^2, and F in steady state take the following forms (see Text S7 for details):

$$\langle m\rangle = \frac{\tau_m}{\tau_P + \frac{1-p_1}{p_1}\tau_S}, \tag{2}$$

$$\sigma_m^2 \approx \langle m\rangle + \langle m\rangle^2 \frac{\left(\frac{\tau_I}{\tau_m}+\frac{\tau_A}{\tau_m}\right)\frac{\tau_S}{\tau_m} - \frac{\tau_A\tau_I}{\tau_m^2}}{\frac{\tau_I}{\tau_m}+\frac{\tau_A}{\tau_m}+\frac{\tau_A\tau_I}{\tau_m^2}}, \tag{3}$$

$$F \approx 1 + \frac{b}{c+d}\frac{d(c+d)(b-1)/b - cd}{c+d+cd}, \tag{4}$$

where τ_m is the lifetime of mRNA, τ_A (τ_I) is the mean duration of the active (inactive) phase, and τ_S is the mean interval between E_{S1} and E_{S2}: $\tau_A = \tau_P/(1-p_1)$, $\tau_I = \tau_S/p_1$, and $\tau_S = \tau_{S1}+\tau_{S2}+p_2\tau_{S3}/(1-p_2)$. The Fano factor is governed by the mean burst size b [the number of mRNAs per burst; $b = 1/(1-p_1)$], $c = \tau_A/\tau_m$, and $d = \tau_I/\tau_m$. A small burst size, a plateau-like bursting profile, or a short inactive phase can contribute to reducing F.

Figure 3A displays how F varies with $\langle m\rangle$ over a wide range of parameter values. For each curve, only one of the 6 parameters is altered, while the others are fixed at default values; we also performed simulations by directly setting τ_S rather than τ_{S1}, τ_{S2}, τ_{S3}, and p_2 and assumed τ_S to obey an exponential distribution. When y is varied, the corresponding F is denoted as F_y. As $F\tau_s$ differs slightly from F_x ($x = \tau_{S1}$, τ_{S2}, τ_{S3}, or p_2) (Fig. 3A, inset), only the curve for $F\tau_s$ is presented. With τ_P regulated alone, the mean height of burst profiles $\{h \approx \tau_m(1-e^{-c})(1-e^{-d})/[\tau_P(1-e^{-c-d})]$ for $\tau_P < \tau_m$ and $h \approx 1$ otherwise$\}$, rather

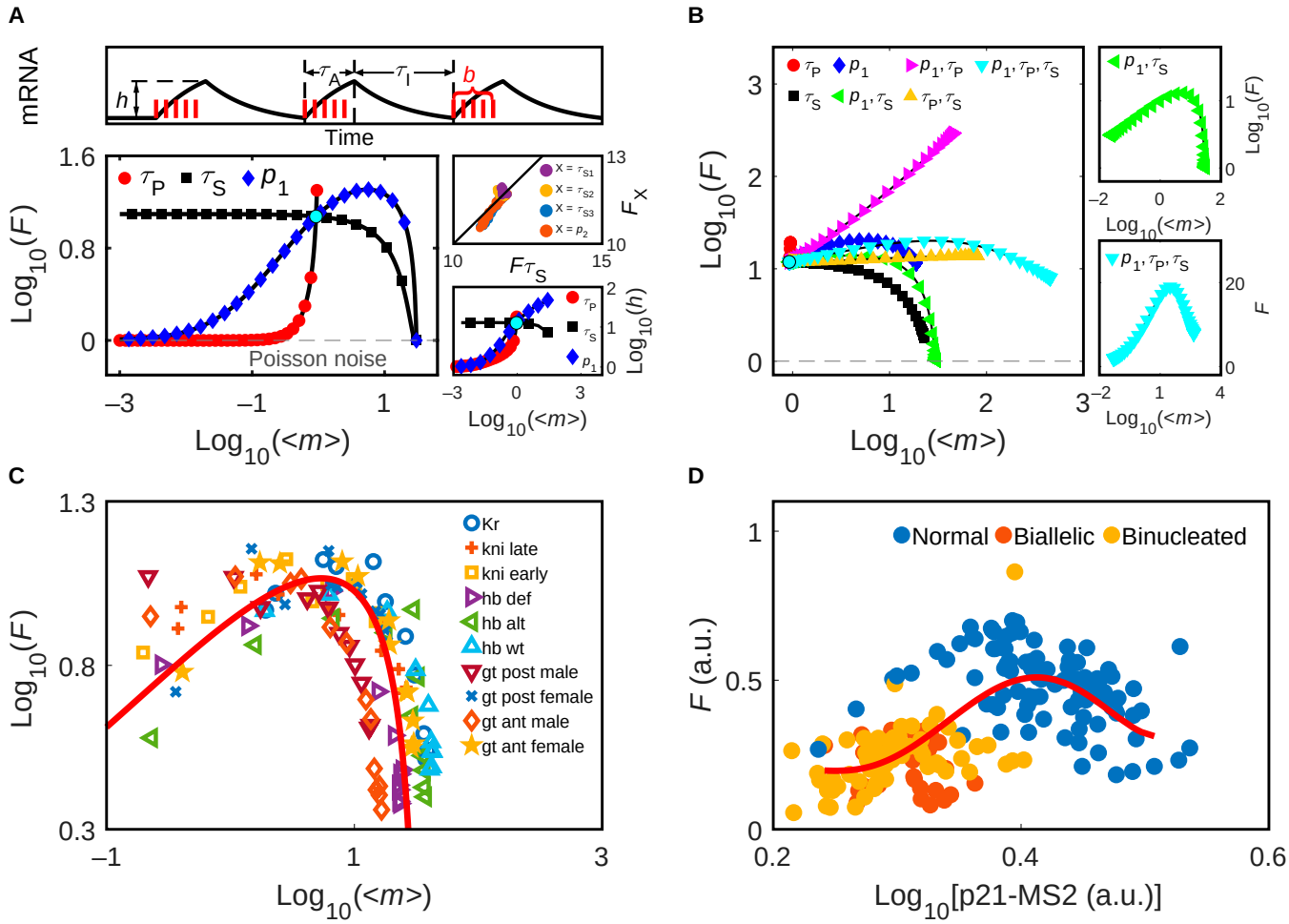


Fig. 3. Relationships between the mean mRNA number <*m*> and Fano factor *F* under various regulatory modes. Time is in units of minutes. (A and B) Solid lines are from analytic expressions (Eqs. G21 to G23 in Text S7), while symbols label numerical results. (A) *F* versus <*m*> for different parameter values. The dashed line represents $F = 1$ (Poisson noise). The default parameters without bound regulators are $p_{1U} = 0.95$ and $\tau_{PU} = 0.165$, and τ_{SU} is assumed to be exponentially distributed with an average of 100 min (cyan dot). On each curve, only one parameter (p_1, τ_P, or τ_S) varies, while the others are fixed at default values. The top panel depicts the sketch of transcriptional bursts (red bars represent Pol II entering productive elongation). The right bottom panel shows the corresponding burst height *h* versus *m*, while the right top one shows $F\tau$s versus F_X ($X = \tau_{S1}, \tau_{S2}, \tau_{S3}, p_2$) with $\tau_{S1U} = 10$, $\tau_{S2U} = 30$, $\tau_{S3U} = 60$, and $p_{2U} = 0.5$. (B) *F* versus <*m*> in different regulatory modes. Regulators affect the transcription by changing the following parameters, while the others remain fixed ($p_{1U} = 0.95$, $\tau_{PU} = 0.165$, $\tau_{SU} = 100$): τ_P (red circle; $\tau_{PB} = 0.01$), τ_S (black square; $\tau_{SB} = 1$), p_1 (blue diamond; $p_{1B} = 0.999$), τ_P and p_1 (pink rightward triangle; $\tau_{PB} = 0.01$ and $p_{1B} = 0.999$), p_1 and τ_S (green leftward triangle; $p_{1B} = 0.999$ and $\tau_{SB} = 1$), τ_P and τ_S (yellow upward triangle; $\tau_{PB} = 0.01$and $\tau_{SB} = 1$); and τ_P, p_1, and τ_S (cyan downward triangle; $\tau_{PB} = 0.01$, $p_{1B} = 0.999$, and $\tau_{SB} = 1$). Each data point represents an average over 10 thousand trajectories. Here, the regulator binding promotes transcription. Two right panels show the cases where both p_1 and τ_S are regulated with τ_P fixed ($p_{1U} = 0.7$, $\tau_{SU} = 500$, $p_{1B} = 0.999$, $\tau_{SB} = 1$, $\tau_P = 0.165$; top), or p_1, τ_S, and τ_P are regulated ($\tau_{PB} = 0.01$, $p_{1B} = 0.999$, $\tau_{SB} = 1$, $\tau_{PU} = 0.2$, $p_{1U} = 0.5$, $\tau_{SU} = 100$; bottom). (C and D) Experimental data on gap genes in early *Drosophila* embryos [47] (C) and on *CDKN1A* (measured by p21-MS2 signals) in MCF7 cells [49] (D). The red curve is a fit to all data with the event model. a.u., arbitrary units.

than burst size, varies prominently. At small τ_P, *F* rises sharply with increasing <*m*> due to little degradation of mRNA within a short active period. At large τ_P, the transcription rate is small such that a nascent mRNA has a high probability of being degraded before new mRNA production, and the active phase is much longer than the inactive phase, i.e., the transcription is nearly a Poisson process ($F = 1$). With τ_S modulated alone, *b* and τ_A show little change, whereas τ_I varies markedly, such that *F* drops toward one with increasing <*m*>. *F* varies slightly at small <*m*> but drops markedly at large <*m*> due to a continuous reduction in

$$\frac{dF}{d\langle m\rangle} \approx \frac{(1-p_1)^2 \frac{\tau_m^3}{\tau_P^3}\left[1+\frac{\tau_m}{\tau_P}\right]}{\left[m-(1-p_1)\frac{\tau_m^2}{\tau_P^2}-\frac{\tau_m}{\tau_P}\right]^2} - 1. \quad (5)$$

With p_1 altered alone, *b* rises with increasing <*m*>, and *h* rises toward saturation; thus, *F* first rises because of an increase in *b* and then drops because of saturation in *h* and plateau-like bursting. Strikingly, all numerical results agree perfectly with the analytical expressions.

The dependence of *F* on <*m*> is more complicated when 2 or more parameters vary concurrently. Notably, the curves for $F\tau_P$ and $F\tau_S$ constitute the boundaries, and the others lie in between and exhibit unique features (Fig. 3B and Fig. S7). Providing that the regulator binding promotes transcription, *F* rises markedly at small <*m*> together with a notable change in *h* or varies slightly when *h* changes little. If *F* falls at intermediate or large <*m*>, then at least τ_S or p_1 is regulated, and <*m*> tends toward τ_m/τ_P when *F* approaches 1; specifically, *F* drops rapidly when τ_P is sufficiently large or changes little. Together, the regulation of τ_P or p_1 is indispensable for a rise in

F, and its rapid rise is accompanied by sharp bursting; the regulation of τ_S or p_1 is essential to a drop in F, and its fast fall is associated with plateau-like bursting. Similarly, the numerical and analytic results match exactly.

Indeed, the fitting curves to experimental data resemble the typical curves above. Activated by the Bicoid (Bcd) protein, major gap genes in early *Drosophila* embryo are expressed at distinct levels along the anterior–posterior axis of its body due to the Bcd concentration gradient [47], sharing a similar trend in F versus $<m>$ (i.e., F first rises and then drops toward 1 with increasing $<m>$) (Fig. 3C; see also the right top panel in Fig. 3B). All the data points can be fitted by a single curve with the event model (see parameter inference in Text S4). For each gene, F is greater than 3.8 at small $<m>$ and falls rapidly after the maximum, indicating that τ_P is nearly consistent across body locations. With τ_P fixed, there should exist the upper bounds, $<m> < \tau_m/\tau_P$ and $F < 1 + \tau_m/\tau_P$ (see Text S7 for details). As these genes display similar limits in $<m>$ and F and have similar τ_m values [48], τ_P (and the frequency of transcription initiation) should also be comparable among them, which implies that variations in Bcd concentration have a minor impact on the major rate-limiting steps in Pol II recruitment and promoter-proximal pausing. The increase in Bcd concentration may primarily facilitate the activation of Pol II and its entry into productive elongation.

Another example involves the expression of *CDKN1A* activated by the tumor suppressor p53 in MCF7 cell lines [49]. F varies slightly at small $<m>$, rises to a peak, and then drops gradually (Fig. 3D); to fit the data with the event model, p_1, τ_S, and τ_P vary concurrently with small p_{1U} and τ_P (see also the right bottom panel in Fig. 3B). At small p_{1U}, the gene state tends to leave set E via E_I rather than E_I^P or leave set I via I_A rather than I_E, associated with slow reactions in recruiting, stabilizing, or releasing Pol II. At small τ_P, fast reactions are required before Pol II enters productive elongation. Thus, p53 may prefer to regulate *CDKN1A* expression via speeding up E_I^P through cellular signaling.

Further analysis of other experimental data reveals the diversity in transcriptional regulation. When the *ctgf* expression is mediated by transforming growth factor-β (TGF-β), the transcription rate rises with increasing the TGF-β concentration, and τ_A is nearly fixed [50], which requires that τ_P is inversely proportional to b. This further implies that the likelihood of Pol II entering productive elongation is prominently higher than its probability of escaping from the core promoter, i.e., the modulation lies in E_I^P rather than I_E. When glucagon-like peptide 1 (GLP-1)/Notch induces *sygl*-1 expression in *Caenorhabditis elegans*, the burst size is nearly irrelevant to the concentration of GLP-1/Notch ligand [51], i.e., τ_P varies markedly but b (p_1) is fixed, implying that the early elongation of Pol II until promoter-proximal pausing can be modulated. Obviously, different regulations of τ_P and p_1 are gene-specific, reflecting which steps are rate-limiting and which reactions are being adjusted in E_P.

Together, the event model enables the deduction of transcriptional modulatory mechanisms. The results above suggest that the changing trend of burst profiles is influenced by the adjustable range of Pol II-dependent reaction rates (e.g., recruitment and promoter-proximal elongation of Pol II). In the 3 limiting cases where the transcription occurs with fixed τ_A, b, or τ_P, the burst height, burst duration, and area under the curve are modulated, respectively.

Global constraints on transcriptional bursting

The results above unravel an important regulatory mode involving the concurrent modulation of p_1 and τ_S. Under this mode, the mean mRNA number ($<m> = \upsilon\tau_m$) and burst size b rise toward saturation with increasing $[R]$, while the burst frequency f first rises to a peak and then drops toward saturation, corresponding to $\tau_A < \tau_I$ and $\tau_A > \tau_I$, respectively (Fig. 4A). Ω_m, Ω_b, and Ω_f refer to the operating points where m, b, and f reach half of their maximum values, respectively. Because $\Omega_m < K_d$, $\Omega_b \approx K_d$ when E_I^P, I_E, A_I, F_A, or C_F is modulated, and $<m> = bf\tau_m$, $\Omega_f < \Omega_m < \Omega_b$ generally holds true, and thus f, $<m>$, and b reach their respective maximum values successively (see Text S4 and Fig. S8 for details). Furthermore, f is sensitive to changes in $<m>$ at low expression, while b is insensitive, and vice versa at high expression (Fig. 4B, top).

We also used $S_b = \frac{d\ln b}{d\ln m}$ and $S_f = \frac{d\ln f}{d\ln m}$ ($S_b + S_f = 1$; S_b and $S_f > 0$ for $\tau_A < \tau_I$), i.e., their sensitivity to changes in mRNA number, to dissect the contribution of size and frequency modulation to mRNA production. S_f and S_b drop and rise respectively with increasing $<m>$ (Fig. 4B, bottom), indicating that the modulation of burst frequency and size predominates in distinct regimes and a combination of both mechanisms emerges over intermediate ranges. A similar relationship between $f(b)$ and $<m>$ was observed across gap genes in early *Drosophila* embryos [47] (Fig. 4C), where their expression levels are graded along the anterior–posterior axis, responsible for body segmentation.

With high-throughput experimental data available, we calculated the mean mRNA numbers and Fano factors for 9196 genes in fibroblasts [52]; to exclude the influence of extrinsic noise, we preprocessed the data using the method in [53]. Each mRNA count is normalized by dividing it by the total count in each sample and scaled by the median count across all samples. Each data point in Fig. 4D denotes $<m>$ and F for one gene, and the red curve is a fitted trendline for all data. There exist lower bounds on Fano factor and burst size at each expression level; meanwhile, $\eta_m^2 = \sigma_m^2/<m>^2$ fluctuates around the mean for large $<m>$ (Fig. 4D, inset), while the total noise intensity in the original data from [52] remains nearly constant at high expression levels [54]. The lower bound on F is nearly irrelevant to $<m>$ at low expression but rises prominently with increasing $<m>$ at high expression. That is, higher expression is accompanied by increased burstiness. This phenomenon has been widely observed in mammalian cells [44,55–58], fly embryos [59–61], and even *Escherichia coli* [62,63], suggesting that transcriptional regulation may obey a universal principle across the genome, i.e., high-level expression is associated with the bursting with large size.

This global constraint results from the correlations among bursting parameters. Given that $\langle m\rangle = \frac{\tau_m}{\tau_P + \frac{1-p_1}{p_1}\tau_S}$, τ_P is expressed as a function of p_1, τ_S, and $<m>$, and, thus, $F = 1 + \frac{(1-p_1)\langle m\rangle(\langle m\rangle\tau_S - p_1\tau_m)\tau_S}{p_1^2\tau_m^2 + (1-p_1)\langle m\rangle\tau_S^2}$. For fixed $<m>$, the minimum of F, F_{min}, is determined by the ranges of τ_S and p_1; if $<m> > \tau_m/(2\tau_S)$, F_{min} equals $1 + \frac{(1-p_{10})\langle m\rangle(\langle m\rangle\tau_{S0} - p_{10}\tau_m)\tau_{S0}}{p_{10}^2\tau_m^2 + (1-p_{10})\langle m\rangle\tau_{S0}^2}$, where τ_{S0} is the minimum of τ_S and p_{10} is the maximum of p_1. F_{min} is greater than one and rises with increasing $<m>$ at high expression, as illustrated by the black curve in Fig. 4E, which is similar in trend to the black one in Fig. 4D (to justify that this global restriction is intrinsic to the transcriptional machinery, few constraints are

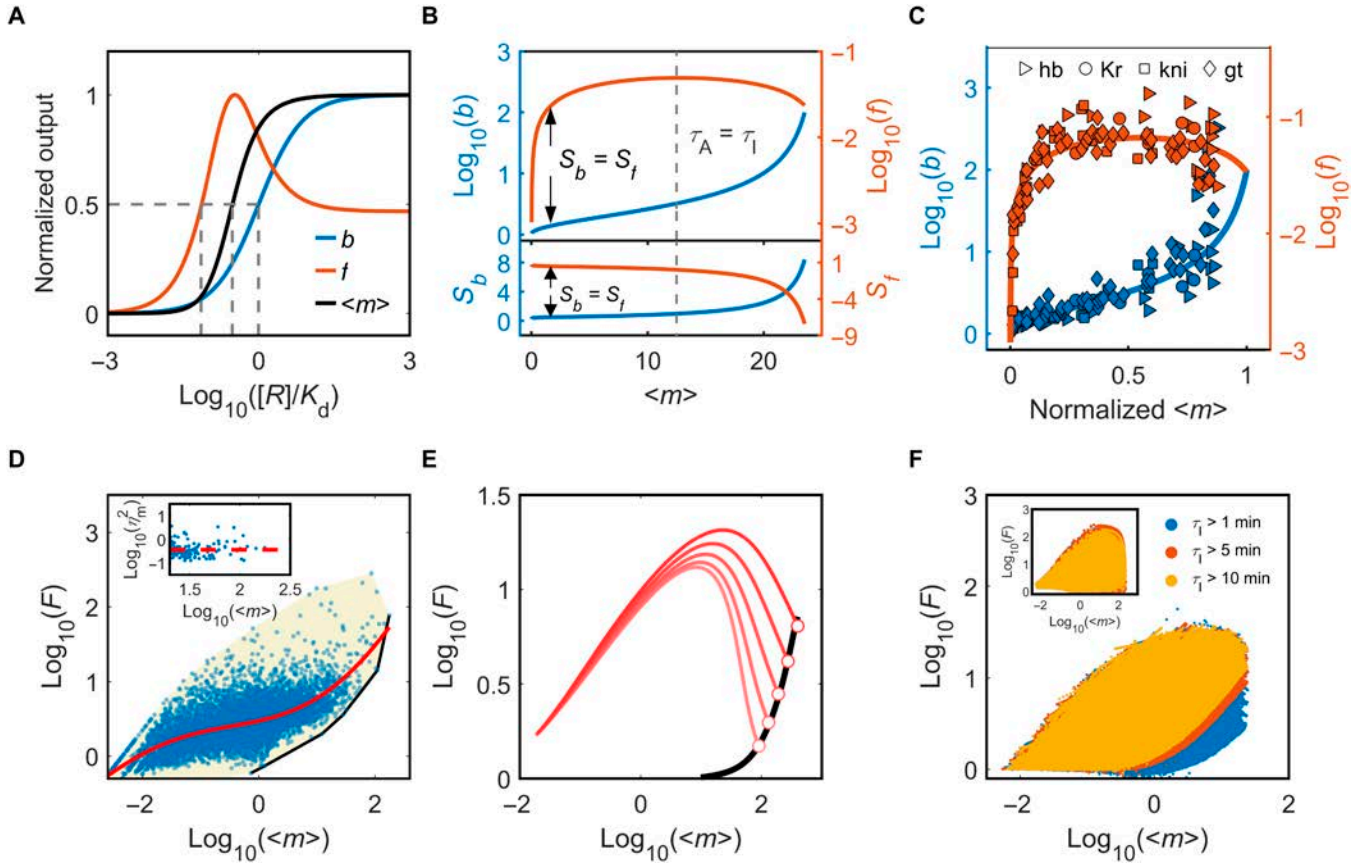


Fig. 4. Modulation of transcriptional bursting. Time is in units of minutes. (A) Mean mRNA count <*m*>, burst size *b*, and burst frequency *f*, normalized by their respective maximum values, versus the normalized regulator concentration [*R*]. The regulator binding affects p_1 and τ_S; $\tau_P = 0.2$, $p_{1B} = 0.995$, $p_{1U} = 0.5$, $\tau_{SU} = 250$, $\tau_{SB} = 5$ and $n = 1$. (B) Burst size and frequency versus the mean mRNA number (top); S_b and S_f versus <*m*> (bottom). The dashed line marks the case of $\tau_I = \tau_A$. The black arrows point to <*m*> with $S_f = S_b$. The parameters are the same as in (A). The curves in (A) and (B) are plotted on the basis of analytic expressions. (C) A fit to the experimental data (symbols) in [47] using the event model with $\tau_P = 0.139$, $p_{1U} = 0.401$, $p_{1B} = 0.995$, $\tau_{SU} = 498$, and $\tau_{SB} = 0.694$ (solid line). (D) Experimental data on the mean expression level and Fano factor in [52]. Each point is from an individual gene with the shadow showing the range. The red fitting curve is obtained by Gaussian regression. The black curve shows the lower bound on Fano factor at large <*m*>. The inset shows $\eta_m^2 = \sigma_m^2/\langle m\rangle^2$ versus <*m*>, with the red dashed line representing the mean. (E) The red lines show Fano factor versus <*m*>, and the black line shows the lower limit of Fano factor at large <*m*> with $\tau_S \geq 1$. The default parameters are the same as in (A). (F) Fano factor versus <*m*> with the event model or telegraph model (inset). Ten thousand sets of bursting parameters are used in each panel, where the duration of a burst varies between 1 and 1,000 min, the transcription rate is from 0.1 to 100/min and $\tau_A < \tau_I$. The refractory period is set to 1 (blue), 5 (red), or 10 min (yellow).

imposed on parameter values; otherwise, if $\tau_I > \tau_A$ were guaranteed, the 2 curves would match). A cluster of red curves is plotted in Fig. 4E, each beginning with the same parameters and ending with a point on the black line. Each curve represents the *F*-<*m*> curve for some gene, similar to that shown in Fig. 3B.

To recapitulate the lower bound on *F* at high expression levels in Fig. 4D, we systematically altered the values of p_1, τ_S, and τ_P and calculated the mean mRNA number and Fano factor in the case of $\tau_A < \tau_I$. The scatter plot of *F* versus *m* has similarities to that in Fig. 4D (Fig. 4F). The global constraint becomes more striking when there exists a longer refractory period in the inactive phase or a higher transcription rate, such as in the case of large τ_S and small τ_P, leading to sharp bursting peaks and enhancing the discernibility of burst profiles. By contrast, high-level expression is not necessarily associated with large Fano factor in the telegraph model and other phenomenological models (inset of Fig. 4F and Fig. S9), which arises from the independence of bursting parameters. Collectively, the global constraint reflects an inherent feature of transcriptional bursting.

Transient dynamics reflect the underlying regulatory mode

After showing the typical features of transcriptional bursting in steady state, we turned to explore the transcriptional response to a step rise in regulator concentration. A new steady state will be reached in diverse manners because of different regulatory modes and composition of the inactive phase. With the initial state completely silent (i.e., $p_1 \to 0$, $p_2 \to 1$, or $\tau_i \to \infty$ and no mRNA production), the mean probability $P_A(t)$ of gene activation and mean mRNA number are determined as follows (see Text S8 for details):

$$P_A(t) \approx f_{I_{\text{init}}}(t) * \delta(t) \sum_{i=0}^{\infty} \left[* f_A(t) * f_I(t) \right]^i * e^{-\frac{t}{\tau_A}}$$
$$\langle m \rangle(t) \approx \frac{1}{\tau_P} e^{-\frac{t}{\tau_m}} * f_{I_\text{init}}(t) * \delta(t) \sum_{i=0}^{\infty} \left[* f_A(t) * f_I(t) \right]^i * e^{-\frac{t}{\tau_A}} \quad (6)$$

where $f_A(t_A)$, $f_I(t_I)$, and $f_{I_\text{init}}(t_{I_\text{init}})$ separately denote the distribution functions of the duration of the active (inactive) phase

in the new steady state and the time taken to enter the active phase (Fig. 5A, top), with $f_A(t_A) = \frac{1}{\tau_A} e^{-\frac{t_A}{\tau_A}}$ and $f_I(t_I) \approx \Gamma(t_I, \alpha_I, \theta_I)$ (Gamma distribution), while * represents convolution. Depending on the initial condition, f_{I_init} can assume different forms (Fig. S10). $f_{first} = f_A * f_{I_init}$ represents the distribution function of the first burst duration t_{first}, which is the time from stimulus onset to completion of mRNA synthesis in the first active phase, and can be expressed as $\Gamma(t_{first}, \alpha_f, \theta_f)$. $f_T = f_A * f_I$ denotes the distribution function of burst duration in steady state (t_{ss}). The mean and coefficient of variability (CV) of t_{first} are τ_{first} and CV_{first}, respectively. The peak and steady-state values of P_A are separately P_{AP} and $P_{Ass} = \tau_A/T$ (T is the mean burst period) (Fig. 5A, middle); $<m>_p$ and $<m>_{SS} = b\tau_m/T$ denote the mean peak and steady-state values of mRNA count, respectively (Fig. 5A, bottom). CV_Z (Z = P, S1, S2, S3, S) refers to the CV for τ_Z in steady state.

Figure 5B shows the typical dynamics of $P_A(t)$: P_A either overshoots P_{Ass} or monotonically rises toward P_{Ass}. Without P_A overshooting, $<m>_p$ cannot exceed $<m>_{ss}$, and its dynamics also markedly depend on τ_m (Fig. 5C). To determine the condition for overshoot, we introduced $H_P = (P_{AP} - P_{Ass})/P_{Ass}$ and $H_m = (<m>_p - <m>_{ss})/<m>_{ss}$. P_A ($<m>$) must overshoot for $h_P T > 1$ ($h_m T > 1$), where $h_x = \frac{\alpha_x}{\delta_x + \tau_{first}} \frac{(\alpha_x - 1)^{\alpha_x - 1}}{\Gamma(\alpha_x)} e^{-(\alpha_x - 1)}$ (x = P or m) with $\alpha_P = \frac{1}{CV_{first}^2}$, $\alpha_m = \frac{(\tau_m + \tau_{first})^2}{\tau_m^2 + \tau_{first}^2 CV_{first}^2}$, $\delta_P = 0$, and $\delta_m = \tau_m$. H_P is governed by τ_{first}/T and CV_{first}; the area below the red curve of $h_p = 1/T$ corresponds to $H_p > 0$ (Fig. 5D). H_m also relies on τ_m/T, and the phase plane of $h_m = 1/T$ separates $H_m = 0$ from $H_m > 0$ (Fig. 5E). Strikingly, overshoot appears at small values of τ_{first}/T, CV_{first}, and τ_m/T.

On the other hand, the time taken to reach half of P_{Ass} ($<m>_{ss}$) is denoted as $t_{GA}(t_{re})$; t_{re} is affected by mRNA production and degradation. If the rising phase of $<m>(t)$ is governed by mRNA production over the time scale of τ_m due to strong correlation of bursts among individual simulations, t_{re} can be less than $\tau_m \ln 2$; otherwise, the dynamics of $<m>$ are controlled by mRNA degradation, leading to $t_{re} > \tau_m \ln 2$. Thus, decreasing τ_{first}/T and CV_{first} contributes to reducing t_{re}; changing T/τ_m has a dual effect, and exclusively increasing it leads to a first drop and then rise in t_{re} (Fig. 5F).

In general, τ_{first}/T and CV_{first} are modulated by the initial state distribution and constrained by the promoter architecture.

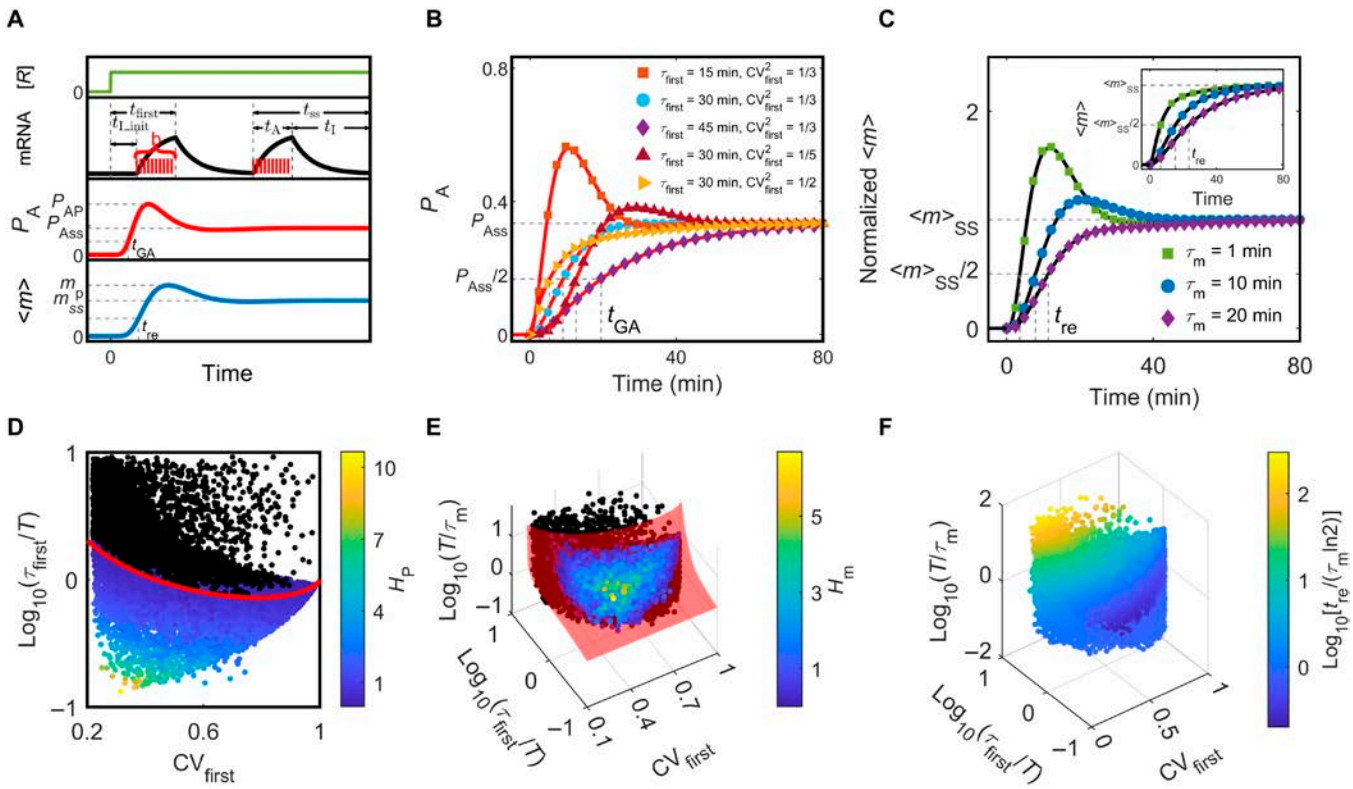


Fig. 5. Response patterns upon stimulation. Time is in units of minutes. (A) Schematic of time courses of [R], the mRNA number (black; red bars represent Pol II entering productive elongation), the probability of gene activation, and the mean mRNA count (from top to bottom). The panel in the second row also marks the durations of the active phase (t_A), inactive phase (t_I), first inactive phase (t_{I_init}), first burst (t_{first}), and a burst in steady state (t_{ss}). (B) Time courses of P_A with $\tau_A = 10$, $\alpha_I = 2$, and $\theta_I = 10$. The α and θ of f_{first} are 3 and 5 ($\tau_{first} = 15$, $CV_{first}^2 = 1/3$), 3 and 10 ($\tau_{first} = 30$, $CV_{first}^2 = 1/3$), 3 and 15 ($\tau_{first} = 45$, $CV_{first}^2 = 1/3$), 5 and 6 ($\tau_{first} = 30$, $CV_{first}^2 = 1/5$), and 2 and 15 ($\tau_{first} = 30$, $CV_{first}^2 = 1/2$), respectively. The data points are from simulation, while the solid curves are from analytical expression (Eq. H2 in Text S8). (C) Time courses of the mean mRNA number. f_A and f_I are the same as in (B); the α and θ of f_{first} are 3 and 5 ($\tau_{first} = 15$, $CV_{first}^2 = 1/3$) for $\tau_m = 1$, 10, or 20, and P_A overshoots. In the inset, the α and θ of f_{first} are 2 and 15 ($\tau_{first} = 30$, $CV_{first}^2 = 1/2$) for $\tau_m = 1$, 10, or 20, and P_A does not overshoot. $<m>(t)$ is normalized by $<m>_{ss}$. The data points are from simulation, while the solid curves are from Eq. H5. (D) Color-coded H_P controlled by τ_{first}/T and CV_{first}. The black points signify that P_A does not overshoot with $h_P T < 1$, while the red curve corresponds to $h_P T = 1$. Small values of τ_{first}/T and CV_{first} lead to large H_P. (E) Color-coded H_m controlled by τ_{first}/T, CV_{first}, and T/τ_m. The black points denote that m does not overshoot with $h_m T < 1$, while the red surface represents $h_m T = 1$. (F) Color-coded t_{re} controlled by τ_{first}/T, CV_{first}, and T/τ_m. In (D) to (F), the initial state is completely silent, and $\tau_A \in (0.1,10)$, $\tau_I \in (0.1,100)$, $CV_P = 1$, $CV_I \in (0.1,1)$ (i = S1, S2, S3), and $\tau_{I_init} = x\tau_I$ with $x \in (0.1, 10)$.

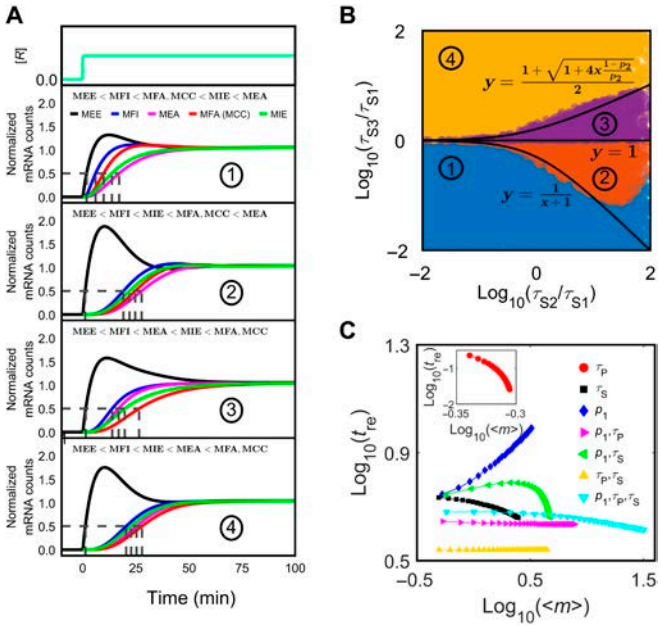


Fig. 6. Transcriptional responses in different regulatory modes. Time is in units of minutes. (A) Transcriptional responses under different regulatory modes and initial conditions. There emerge 4 scenarios in terms of t_{re}: MEE < MFI < MFA, MCC < MIE < MEA for $\tau_{S1} = 15$, $\tau_{S2} = 2$, and $\tau_{S3} = 4$ after the stimulation; MEE < MFI < MIE < MFA, MCC < MEA for $\tau_{S1} = 10$, $\tau_{S2} = 20$, and $\tau_{S3} = 6$; MEE < MFI < MEA < MIE < MFA, MCC for $\tau_{S1} = 3$, $\tau_{S2} = 10$, and $\tau_{S3} = 15$; and MEE < MFI < MIE < MEA < MFA, MCC for $\tau_{S1} = 5$, $\tau_{S2} = 20$, and $\tau_{S3} = 8$. The other parameters are fixed: $\tau_P = 1$, $\tau_m = 10$, $p_1 = 0.9$, and $p_2 = 0.5$. $<m>$ is normalized by its steady-state value. (B) Two-parameter phase diagram of $x = \tau_{S2}/\tau_{S1}$ and $y = \tau_{S3}/\tau_{S1}$ at $p_2 = 0.5$. The black curves constitute the boundaries. The blue region with $y < 1/(x+1)$, red region with $1/(x+1) < y < 1$, purple region with $1 < y < \left(1+\sqrt{1+4x(1-p_2)/p_2}\right)/2$, and yellow region with $y > \left(1+\sqrt{1+4x(1-p_2)/p_2}\right)/2$ correspond to the 4 scenarios in (A). (C) $<m>$ and t_{re} for different regulatory modes and bursting parameters. The color represents the bursting parameters regulated. Default parameters are $\tau_{PU} = 1$, $\tau_{SU} = 10$, $p_{1U} = 0.5$, $\tau_m = 5$, $CV_P^2 = 1$, and $CV_S^2 = 1$. Under a regulatory mode, the regulated parameters are decided by the regulator concentration, default parameters, and the corresponding parameters with bound regulators ($\tau_{PB} = 0.1$, $\tau_{SB} = 1$, and $p_{1B} = 0.95$).

When the gene state immediately enters the active phase upon stimulus onset due to regulation of Pol II-dependent reactions, τ_{first}/T and CV_{first} are separately close to τ_A/T and the CV of the active phase duration (C_A, which is often ~1), whereas τ_{first}/T and CV_{first} separately approach 1 and the CV of t_{ss} (i.e., CV_{ss}) when the state just leaves the active phase. If there are multiple rate-limiting steps in burst cycle before stimulus onset [11,44], the initial state is mostly distributed around those states, usually leading to $\tau_A/T < \tau_{first}/T < 1$ and small CV_{first}, whereas $\tau_{first}/T \sim 1$ and $CV_{first} \sim 1$ only if there is one rate-limiting step. The number of rate-limiting steps is governed by the promoter architecture [11,64,65] and is associated with the magnitude of p_2. For large p_2, the sum of all durations of E_{S3} within a burst cycle often accounts for the majority of the burst period, and its distribution is approximately exponential, implying that large p_2 is often associated with only one rate-limiting step (i.e., nucleosome clearance) (see Text S2). By contrast, formation of the scaffold complex and pause of Pol II are the major rate-limiting steps for small p_2. For instance, the nucleosome occupancy rate is rather low at target genes of white collar complex in *Neurospora crassa*, which is linked with small p_2 and relatively large $\tau_{S1} + \tau_{S2}$, leading to the observable overshoot [46]. Collectively, analyzing the transient response also provides insight into the modulatory mechanism and promoter architecture.

Notably, only a finite number of scenarios emerge from the comparison of transient transcriptional responses across different regulatory modes. When the gene is activated from the silent state to the same steady state (with identical f_A and f_I), for example, 4 scenarios appear upon examining the ordering of t_{re} values across 6 basic regulatory modes (Fig. 6A). These relationships primarily depend on the composition of the inactive phase, i.e., the values of p_2, τ_{S2}/τ_{S1}, and τ_{S3}/τ_{S1} (Fig. 6B). Nevertheless, the fastest response always occurs via MEE, as paused Pol II is poised to enter productive elongation upon an excitatory signal [66]; the second fastest response takes place via MFI, as the transcription machinery is ready for assembly on the promoter (Figs. S11 and S12). Under combinatory modes, a faster response can be induced to reach the same $<m>_{ss}$ when τ_P is regulated (Fig. 6C). Therefore, accelerating the reactions involving Pol II can elicit fast transcription from a completely silent state.

Moreover, it may be beneficial for an inducible gene to respond fast to stimuli and generate a broad range of outputs. The rapid response largely results from the overshoot caused by regulating τ_P via MEE, while regulating τ_P in the case of $\tau_A > \tau_I$ or regulating p_1 and τ_S in the case of $\tau_A < \tau_I$ elicits a broad range in $<m>$. Thus, integrating MEE with MX (X = IE, EA, FI, CC, or FA) allows for a fast response and highly tunable expression (Fig. 6C). In this sense, regulating E_I^P and I_E is conducive to inducible genes [67].

Discussion

The event model showcases several strengths when compared with previous phenomenological models. First, biochemical reactions involved in transcription are accurately mapped to the events in our model, and parameter values can be directly inferred from experimental data. Second, the event model retains the correlations among the bursting variables [transcription rate constant, burst size, and duration of the (in)active phase]. This naturally imposes constraints on burst profiles, guaranteeing bursty transcription in highly expressed genes [52,68,69] and enabling the modulation of burst frequency and burst size to separately predominate at low and high expression levels [47,57,70]. Third, 9 kinds of functional state transitions are distinctly separated in time scales from seconds to hours, and the durations and occurrence probabilities of events are independently modulated. These enable transcriptional bursting to manifest across a broad spectrum of temporal scales, exhibiting diverse dynamic patterns. With these features, the event model could have wide applications.

In principle, we can deduce the changes in (τ_P, τ_S, p_1) through the correlations among observables and further reveal the promoter architecture and molecular regulatory mechanisms. For example, E_P consists of sequential stages (initiation, promoter clearance, elongation, and promoter-proximal pausing). The tendency of Pol II to continually synthesize mRNAs and the duration of promoter-proximal pause determine p_1 and τ_P, respectively. For a group of gap genes in early *Drosophila* embryos [47], τ_P is nearly fixed, but p_1 is regulated, suggesting that regulators strongly affect the stability of the transcription complex and weakly affect the motion of Pol II along DNA templates, e.g., modulating the phosphorylation of Pol II

C-terminal domain. In the regulation of *ctgf* expression by TGF-β1 [50], the duration of the active phase varies only slightly, while the transcription rate constant rises with increasing the TGF-β1 concentration, suggesting that TGF-β1 affects the ratio of forward to backward rates of Pol II movement along DNA templates during the stalling and reverting of the EEC. In the regulation of *sygl*-1 expression in *C. elegans*, the burst size remains relatively constant, while τ_A experiences changes [51], suggesting that GLP-1/Notch influences the duration of a highly irreversible stage, e.g., early elongation of Pol II until promoter-proximal pausing. Together, although the major stages in the event model are similar across genes, the reversibility and rate limits of those stages govern the regulatory modes, reflected in the changes to the durations and ordering of the events.

Cells are exposed to various fluctuations, and some have to rapidly respond and adapt to new environments via transcription [67]. For inducible genes with rapid transcriptional activation [66], it is desirable to trigger highly tunable expression and allow transcription overshoot to lift the limitation of mRNA lifetime on response speed. This can be achieved when τ_P and/or p_1 (with $\tau_A > \tau_I$) or τ_S and/or p_1 (with $\tau_A < \tau_I$) are regulated. Meanwhile, transcriptional overshoot can be evoked when Pol II is poised for transcription, or there exist multiple rate-limiting steps, e.g., a refractory period and PIC recruitment during the inactive phase. The genes are indeed prone to overshoot and respond fast in the presence of a refractory period [46], helpful for the rapid and accurate transmission of information. On the contrary, eukaryotic genes with TATA box often burst with a large size and seldom overshoot [11], conducive to filtering environmental fluctuations, and the TATA box seems to be correlated with lack of a refractory period.

The event model engages various types of regulatory factors; nevertheless, only the concentration of one regulator was altered, while the concentrations of others were kept constant in simulations. When transcriptional modulation involves coordinating multiple regulators and the relationships between their concentrations are known, it is possible to elucidate the combinatorial control of bursting parameters. We could determine the dependence of bursting parameters on the concentration of one of the involved regulators and generate $<m>$-F curves under diverse conditions. Note that the conclusions drawn in Figs. 3A, 4D to F, and 5 rely solely on the values of τ_P, τ_{S1}, τ_{S2}, τ_{S3}, p_1, and p_2 and are not influenced by the single-regulator assumption. In Figs. 3C and D and 4A to C, the experimental data were collected when only one kind of regulator was monitored. If the concentrations of multiple regulators change simultaneously, the trends depicted in Fig. 3B may undergo changes, although the $<m>$-$F_{\tau P}$ and $<m>$-$F_{\tau S}$ curves still constitute the boundaries.

Transcriptional noise can be divided into intrinsic and extrinsic noise [71]. Here, the intrinsic noise is mainly manifested in the stochasticity in event orders, event durations, and mRNA degradation. Treating mRNA degradation as a one-step process is a simplification; however, it was reported that, in a broad range of parameter space, models assuming either multistep or one-step degradation of mRNA yield indistinguishable mRNA count distributions [16]. The current work did not explore extrinsic noise, which is influenced by various factors including cell volume, gene copies, DNA replication, and cell cycle progression [72,73]. For example, it has been shown that the presence of 2 uncorrelated alleles halves the noise intensity [56]. It was reported that extrinsic noise typically leads to a more dispersed mRNA count distribution, such as an increased probability of low numbers or a longer tail, and that extrinsic noise contributes substantially to the total noise at high expression levels, leading to a higher noise plateau [11]. It was also demonstrated that parameter variability among cells should be taken into account to fully interpret the experimental data [54]. Furthermore, when analyzing data obtained through single-cell measurements, technological noise is also unavoidable [53]. All these suggest that the event model has to be extended to incorporate more sources of noise, and it would be interesting to dissect the contributions of intrinsic and extrinsic noise.

The current work explored the generic case where the dwell times of regulators are much shorter than the time taken to switch between the inactive and active phases, regulator binding promotes transcription with monotonic gene-regulatory functions, and only one gene locus is involved [5,34,35,74]. More advanced models could be developed to probe the gene-specific transcriptional activity. If the concentration of regulators varies periodically (e.g., the pulsing of p53 and nuclear factor κB levels), how transcriptional bursting is modulated to reliably code signals could be the focus of further study.

In summary, the event model, built upon a small set of adjustable functional events, offers a clear and straightforward description of transcriptional progression. This approach facilitates the investigation of both shared and specific mechanisms for gene regulation. Modulation can occur at any stage of a burst cycle, but the underlying regulatory modes are limited in number because of intrinsic correlations among transcriptional events. Our results suggest that large Fano factor is required for transcriptional bursting at high expression levels and the transitions E_I^P *and* I_E, involved in the recruitment, pause and release, or elongation of Pol II, are pivotal regulatory targets for enhancing the transcriptional sensitivity, anti-crosstalk, adaptability, and responsiveness. Combining experimental data with the event model allows for deducing the molecular processes that predominantly govern transcriptional activity.

Methods

The development of the event model is an important part of this work, and the model is the basis of the whole study. Thus, we provided the details on building the model in the "Establishment of the event model" section. Texts S1 to S8 further present the derivation of formula and explanations. We made a detailed reference to the Supplementary Materials in the text.

Acknowledgments

We thank Y. Wang for comments on an early version of this paper and 2 anonymous reviewers for constructive comments and suggestions. The numerical calculations in this paper have been performed on the computing facilities at the High Performance Computing Center (HPCC) of Nanjing University.
Funding: This work was supported by the National Natural Science Foundation of China (11874209 and 12090052).
Author contributions: F.L., R.W., and W.W. conceived and designed the project. R.W. performed the study. F.L., R.W., and W.W. wrote the paper. All authors analyzed the data and read/edited the paper.
Competing interests: The authors declare that they have no competing interests.

Data Availability

All data needed to evaluate the conclusions in the paper are present in the paper. Custom codes will be available upon request to F.L.

Supplementary Materials

Text S1. Establishment of the event model.
Text S2. Structure of the event model.
Text S3. Simulation of the event model and phenomenological models.
Text S4. Regulation of transcriptional bursting.
Text S5. Comparison between the single-binding and multibinding sites cases.
Text S6. Gene regulatory function.
Text S7. Mean and variance of mRNA numbers.
Text S8. Transient response to stimulation.
Fig. S1. Events in transcriptional bursting.
Fig. S2. Duration distribution of the process that consists of 2 sequential processes.
Fig. S3. Duration distribution of a mixed process.
Fig. S4. Probability of repeating events E_P and E_{S3}.
Fig. S5. Relationships between the mean ($<m>$) and variance (σ_m^2) of mRNA numbers for different distributions of f_P and f_S.
Fig. S6. Relationships between the mean ($<m>$) and relative noise strength ($\eta_m^2 = \sigma_m^2/<m>^2$) of mRNA numbers for different f_P and f_S.
Fig. S7. Relationships between the mean ($<m>$) and variance (σ_m^2) of mRNA numbers under different regulatory modes.
Fig. S8. Transcriptional kinetics under different combination modes.
Fig. S9. Difference between the telegraph and event models under different regulatory modes.
Fig. S10. Response curves for different distribution functions of the time taken to activate the gene from the initial state.
Fig. S11. Regulatory range (v_{max}) and response speed for basic regulatory modes.
Fig. S12. Transcriptional sensitivity under different regulatory modes.

References

1. Neuert G, Munsky B, Tan RZ, Teytelman L, Khammash M, van Oudenaarden A. Systematic identification of signal-activated stochastic gene regulation. *Science*. 2013;339(6119):584–587.
2. Ezer D, Moignard V, Göttgens B, Adryan B. Determining physical mechanisms of gene expression regulation from single cell gene expression data. *PLOS Comput Biol*. 2016;12(8):e1005072.
3. Lis JT. A 50 year history of technologies that drove discovery in eukaryotic transcription regulation. *Nat Struct Mol Biol*. 2019;26(9):777–782.
4. Coulon A, Chow CC, Singer RH, Larson DR. Eukaryotic transcriptional dynamics: From single molecules to cell populations. *Nat Rev Genet*. 2013;14(8):572–584.
5. Lu F, Lionnet T. Transcription factor dynamics. *Cold Spring Harb Perspect Biol*. 2021;13(11):a040949.
6. Wissink EM, Vihervaara A, Tippens ND, Lis JT. Nascent RNA analyses: Tracking transcription and its regulation. *Nat Rev Genet*. 2019;20(12):705–723.
7. Chen X, Zhang D, Su N, Bao B, Xie X, Zuo F, Yang L, Wang H, Jiang L, Lin Q, et al. Visualizing RNA dynamics in live cells with bright and stable fluorescent RNAs. *Nat Biotechnol*. 2019;37(11):1287–1293.
8. Peccoud J, Ycart B. Markovian modeling of gene-product synthesis. *Theor Popul Biol*. 1995;48(2):222–234.
9. Pedraza JM, Paulsson J. Effects of molecular memory and bursting on fluctuations in gene expression. *Science*. 2008;319(5861):339–343.
10. Zhang J, Chen L, Zhou T. Analytical distribution and tunability of noise in a model of promoter progress. *Biophys J*. 2012;102(6):1247–1257.
11. Zoller B, Nicolas D, Molina N, Naef F. Structure of silent transcription intervals and noise characteristics of mammalian genes. *Mol Syst Biol*. 2015;11(7):823.
12. Corrigan AM, Tunnacliffe E, Cannon D, Chubb JR. A continuum model of transcriptional bursting. *eLife*. 2016;5:e13501.
13. Tantale K, Mueller F, Kozulic-Pirher A, Lesne A, Victor J-M, Robert M-C, Capozi S, Chouaib R, Bäcker V, Mateos-Langerak J, et al. A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat Commun*. 2016;7:12248.
14. Zhang J, Zhou T. Promoter-mediated transcriptional dynamics. *Biophys J*. 2014;106(2):479–488.
15. Szavits-Nossan J, Grima R. Steady-state distributions of nascent RNA for general initiation mechanisms. *Phys Rev Res*. 2023;5(1):013064.
16. Braichenko S, Holehouse J, Grima R. Distinguishing between models of mammalian gene expression: Telegraph-like models versus mechanistic models. *J R Soc Interface*. 2021;18(183):20210510.
17. Karmakar R. Control of noise in gene expression by transcriptional reinitiation. *J Stat Mech*. 2020;2020:063402.
18. Karmakar R, Das AK. Effect of transcription reinitiation in stochastic gene expression. *J Stat Mech*. 2021;2021:033502.
19. Cao Z, Filatova T, Oyarzún DA, Grima R. A stochastic model of gene expression with polymerase recruitment and pause release. *Biophys J*. 2020;119(5):1002–1014.
20. Muthukrishnan AB, Kandhavelu M, Lloyd-Price J, Kudasov F, Chowdhury S, Yli-Harja O, Ribeiro AS. Dynamics of transcription driven by the *tetA* promoter, one event at a time, in live *Escherichia coli* cells. *Nucleic Acids Res*. 2012;40(17):8472–8483.
21. Lloyd-Price J, Startceva S, Kandavalli V, Chandraseelan JG, Goncalves N, Oliveira SMD, Häkkinen A, Ribeiro AS. Dissecting the stochastic transcription initiation process in live *Escherichia coli*. *DNA Res*. 2016;23(3):203–214.
22. Weidemann DE, Holehouse J, Singh A, Grima R, Hauf S. The minimal intrinsic stochasticity of constitutively expressed eukaryotic genes is sub-Poissonian. *Sci Adv*. 2023;9(32):eadh5138.
23. Hansen AS, Zechner C. Promoters adopt distinct dynamic manifestations depending on transcription factor context. *Mol Syst Biol*. 2021;17(2):e9821.
24. Scholes C, DePace AH, Sánchez Á. Combinatorial gene regulation through kinetic control of the transcription cycle. *Cell Syst*. 2017;4(1):97–108.e9.
25. Wang Y, Liu F, Wang W. Dynamic mechanism for the transcription apparatus orchestrating reliable responses to activators. *Sci Rep*. 2012;2:422.
26. Wang Y, Ni T, Wang W, Liu F. Gene transcription in bursting: A unified model for realizing accuracy and stochasticity. *Biol Rev*. 2019;94(1):248–258.
27. Wang Y, Liu F, Wang W. Kinetics of transcription initiation directed by multiple *cis*-regulatory elements on the *glnAp2* promoter. *Nucleic Acids Res*. 2016;44(22):10530–10538.

28. Jia T, Kulkarni RV. Intrinsic noise in stochastic models of gene expression with molecular memory and bursting. *Phys Rev Lett*. 2011;106(5):058102.
29. Zhang Z, Deng Q, Wang Z, Chen Y, Zhou T. Exact results for queuing models of stochastic transcription with memory and crosstalk. *Phys Rev E*. 2021;103(6):062414.
30. Schwabe A, Rybakova KN, Bruggeman FJ. Transcription stochasticity of complex gene regulation models. *Biophys J*. 2012;103(6):1152–1161.
31. Nicolas D, Zoller B, Suter DM, Naef F. Modulation of transcriptional burst frequency by histone acetylation. *Proc Natl Acad Sci USA*. 2018;115(27):7153–7158.
32. Donovan BT, Huynh A, Ball DA, Patel HP, Poirier MG, Larson DR, Ferguson ML, Lenstra TL. Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting. *EMBO J*. 2019;38(12):e100809.
33. Zabidi MA, Stark A. Regulatory enhancer-core-promoter communication via transcription factors and cofactors. *Trends Genet*. 2016;32(12):801–814.
34. Voss TC, Hager GL. Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nat Rev Genet*. 2014;15(2):69–81.
35. Suter DM. Transcription factors and DNA play hide and seek. *Trends Cell Biol*. 2020;30(6):491–500.
36. Brown CR, Mao CH, Falkovskaia E, Jurica MS, Boeger H. Linking stochastic fluctuations in chromatin structure and gene expression. *PLOS Biol*. 2013;11(8):e1001621.
37. Muramoto T, Müller I, Thomas G, Melvin A, Chubb JR. Methylation of H3K4 is required for inheritance of active transcriptional states. *Curr Biol*. 2010;20(5):397–406.
38. Casamassimi A, Napoli C. Mediator complexes and eukaryotic transcription regulation: An overview. *Biochimie*. 2007;89(12):1439–1446.
39. Malik S, Roeder RG. Dynamic regulation of pol II transcription by the mammalian mediator complex. *Trends Biochem Sci*. 2005;30(5):256–263.
40. Jonkers I, Kwak H, Lis JT. Genome-wide dynamics of pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *eLife*. 2014;3:e02407.
41. Tome JM, Tippens ND, Lis JT. Single-molecule nascent RNA sequencing identifies regulatory domain architecture at promoters and enhancers. *Nat Genet*. 2018;50(11):1533–1541.
42. Shao W, Zeitlinger J. Paused RNA polymerase II inhibits new transcriptional initiation. *Nat Genet*. 2017;49(7):1045–1051.
43. Kumar N, Singh A, Kulkarni RV. Transcriptional bursting in gene expression: Analytical results for general stochastic models. *PLOS Comput Biol*. 2015;11(10):e1004292.
44. Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. Mammalian genes are transcribed with widely different bursting kinetics. *Science*. 2011;332(6028):472–474.
45. Hendy O, Campbell J Jr, Weissman JD, Larson DR, Singer DS. Differential context-specific impact of individual core promoter elements on transcriptional dynamics. *Mol Biol Cell*. 2017;28(23):3360–3370.
46. Li C, Cesbron F, Oehler M, Brunner M, Höfer T. Frequency modulation of transcriptional bursting enables sensitive and rapid gene regulation. *Cell Syst*. 2018;6(4):409–423.e11.
47. Zoller B, Little SC, Gregor T. Diverse spatial expression patterns emerge from unified kinetics of transcriptional bursting. *Cell*. 2018;175(3):835–847.
48. Garcia HG, Tikhonov M, Lin A, Gregor T. Quantitative imaging of transcription in living *Drosophila* embryos links polymerase activity to patterning. *Curr Biol*. 2013;23(21):2140–2145.
49. Hafner A, Reyes J, Stewart-Ornstein J, Tsabar M, Jambhekar A, Lahav G. Quantifying the central dogma in the p53 pathway in live single cells. *Cell Syst*. 2020;10(6):495–505.
50. Molina N, Suter DM, Cannavo R, Zoller B, Gotic I, Naef F. Stimulus-induced modulation of transcriptional bursting in a single mammalian gene. *Proc Natl Acad Sci USA*. 2013;110(51):20563–20568.
51. Lee CH, Shin H, Kimble J. Dynamics of Notch-dependent transcriptional bursting in its native context. *Dev Cell*. 2019;50(4):426–435.e4.
52. Larsson AJM, Johnsson P, Hagemann-Jensen M, Hartmanis L, Faridani OR, Reinius B, Segerstolpe Å, Rivera CM, Ren B, Sandberg R. Genomic encoding of transcriptional burst kinetics. *Nature*. 2019;565(7738):251–254.
53. Grün D, Kester L, van Oudenaarden A. Validation of noise models for single-cell transcriptomics. *Nat Methods*. 2014;11(6):637–640.
54. Grima R, Esmenjaud P-M. Systematic biases in transcriptional parameters inferred from single-cell snapshot data. *bioRxiv*. 2023;2023.06.19.545536.
55. Yunger S, Rosenfeld L, Garini Y, Shav-Tal Y. Single-allele analysis of transcription kinetics in living mammalian cells. *Nat Methods*. 2010;7(8):631–633.
56. Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. Stochastic mRNA synthesis in mammalian cells. *PLOS Biol*. 2006;4(10):e309.
57. Dar RD, Razooky BS, Singh A, Trimeloni TV, McCollum JM, Cox CD, Simpson ML, Weinberger LS. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proc Natl Acad Sci USA*. 2012;109(43):17454–17459.
58. Itzkovitz S, Lyubimova A, Blat IC, Maynard M, van Es J, Lees J, Jacks T, Clevers H, van Oudenaarden A. Single-molecule transcript counting of stem-cell markers in the mouse intestine. *Nat Cell Biol*. 2012;14(1):106–114.
59. Little SC, Tikhonov M, Gregor T. Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell*. 2013;154(4):789–800.
60. Pare A, Lemons D, Kosman D, Beaver W, Freund Y, McGinnis W. Visualization of individual *Scr* mRNAs during *Drosophila* embryogenesis yields evidence for transcriptional bursting. *Curr Biol*. 2009;19(23):2037–2042.
61. Boettiger AN, Levine M. Rapid transcription fosters coordinate snail expression in the *Drosophila* embryo. *Cell Rep*. 2013;3(1):8–15.
62. Engl C, Jovanovic G, Brackston RD, Kotta-Loizou I, Buck M. The route to transcription initiation determines the mode of transcriptional bursting in *E. coli*. *Nat Commun*. 2020;11:2422.
63. Zaslaver A, Bren A, Ronen M, Itzkovitz S, Kikoin I, Shavit S, Liebermeister W, Surette MG, Alon U. A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nat Methods*. 2006;3(8):623–628.
64. Valen E, Sandelin A. Genomic and chromatin signals underlying transcription start-site selection. *Trends Genet*. 2011;27(11):475–485.
65. Lenhard B, Sandelin A, Carninci P. Metazoan promoters: Emerging characteristics and insights into transcriptional regulation. *Nat Rev Genet*. 2012;13(4):233–245.

66. Muse GW, Gilchrist DA, Nechaev S, Shah R, Parker JS, Grissom SF, Zeitlinger J, Adelman K. RNA polymerase is poised for activation across the genome. *Nat Genet*. 2007;39(12):1507–1511.
67. Weake VM, Workman JL. Inducible gene expression: Diverse regulatory mechanisms. *Nat Rev Genet*. 2010;11(6):426–437.
68. Sanchez A, Golding I. Genetic determinants and cellular constraints in noisy gene expression. *Science*. 2013;342(6163):1188–1193.
69. Battich N, Stoeger T, Pelkmans L. Control of transcript variability in single mammalian cells. *Cell*. 2015;163(7):1596–1610.
70. Lammers NC, Galstyan V, Reimer A, Medin SA, Wiggins CH, Garcia HG. Multimodal transcriptional control of pattern formation in embryonic development. *Proc Natl Acad Sci USA*. 2020;117(2):836–847.
71. Swain PS, Elowitz MB, Siggia ED. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci USA*. 2002;99(20):12795–12800.
72. Cao Z, Grima R. Analytical distributions for detailed models of stochastic gene expression in eukaryotic cells. *Proc Natl Acad Sci USA*. 2020;117(9):4682–4692.
73. Fu X, Patel HP, Coppola S, Xu L, Cao Z, Lenstra TL, Grima R. Quantifying how post-transcriptional noise and gene copy number variation bias transcriptional parameter inference from mRNA distributions. *eLife*. 2022;11:e82493.
74. Zhang Z, English BP, Grimm JB, Kazane SA, Hu W, Tsai A, Inouye C, You C, Piehler J, Schultz PG, et al. Rapid dynamics of general transcription factor TFIIB binding during preinitiation complex assembly revealed by single-molecule analysis. *Genes Dev*. 2016;30(18):2106–2118.
75. Roeder RG. 50+ years of eukaryotic transcription: An expanding universe of factors and mechanisms. *Nat Struct Mol Biol*. 2019;26(9):783–791.
76. Yakovchuk P, Gilman B, Goodrich JA, Kugel JF. RNA polymerase II and TAFs undergo a slow isomerization after the polymerase is recruited to promoter-bound TFIID. *J Mol Biol*. 2010;397(1):57–68.
77. Warfield L, Ramachandran S, Baptista T, Devys D, Tora L, Hahn S. Transcription of nearly all yeast RNA polymerase II-transcribed genes is dependent on transcription factor TFIID. *Mol Cell*. 2017;68(1):118–129.e5.
78. Blair RH, Goodrich JA, Kugel JF. Single-molecule fluorescence resonance energy transfer shows uniformity in TATA binding protein-induced DNA bending and heterogeneity in bending kinetics. *Biochemistry*. 2012;51(38):7444–7455.
79. Bartman CR, Hsu SC, Hsiung CCS, Raj A, Blobel GA. Enhancer regulation of transcriptional bursting parameters revealed by forced chromatin looping. *Mol Cell*. 2016;62(2):237–247.
80. Chen H, Levo M, Barinov L, Fujioka M, Jaynes JB, Gregor T. Dynamic interplay between enhancer-promoter topology and gene activity. *Nat Genet*. 2018;50(9):1296–1303.
81. Saunders A, Core LJ, Lis JT. Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol*. 2006;7(8):557–567.