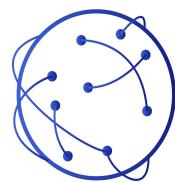
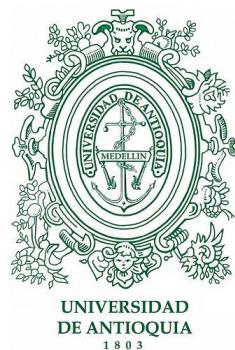


STUDY AND MODELING OF **DYNAMICAL PROPERTIES IN THE CONSERVATIVE GENE**
REGULATORY NETWORKS IN ANIMAL DEVELOPMENT

Lina Marcela Ruiz Galvis
Master student in Biology

Gloria Machado Rodriguez
Adviser

Boris Anghelo Rodriguez Rey
Co-adviser



FEnFiSDi

Fundamentos y Enseñanza de la
Física y los Sistemas Dinámicos

Acknowledgment

I cannot express enough thanks to my parents for their constant support. I offer my sincere appreciation for the learning opportunities provided by my teacher, Boris Anghelo Rodriguez Rey. My completion of this project could not have been accomplished without the support of my adviser Gloria Machado Rodriguez. Finally, thanks to my research group FenFisDi.

TABLE OF CONTENTS

FIGURES LIST	4
TABLES LIST	6
ABBREVIATION LIST	7
GLOSSARY	8
ABSTRACT	10
INTRODUCTION	11
Biological noise and gene expression as transcriptional bursting	11
Gene expression kinetics	13
Promoter activation and inactivation	14
Transcription and mRNA degradation	15
Translation and protein degradation	16
Gene regulatory networks and Biological noise in embryonic development	16
Methodological approaches to study the biological noise in GRN	19
Problem and justification	20
Objectives	20
General objective	20
Specific objectives	21
METHODOLOGY	21
Section 1: The regulatory systems	22
Section 2: The three-stage model of one gene with NGM and CME	22
Section 3: The CLE implementation for the three-stage model is a good approach for large systems	26
Section 4: Dynamical properties of the regulatory systems for individual cells	30
Section 5: Dynamical properties of the regulatory systems for coupled cells	32
RESULTS AND DISCUSSION	35
Section 1: The regulatory systems	35
Auto-Activated gene (Positive self-activated gene)	35
Activator-inhibitor (Turing System)	37
Section 2: The three-stage model of one gene with NGM and CME	37
Section 3: The CLE implementation for the three-stage model is a good approach for large systems	42
Section 4: Dynamical properties of the regulatory systems for individual cells	50
	2

Section 5: Dynamical properties of the regulatory systems for coupled cells	63
CONCLUSIONS	69
Section 2: The three-stage model of one gene with NGM and CME	69
Section 3: The CLE implementation for the three-stage model is a good approach for large systems	70
Section 4: Dynamical properties of the regulatory systems for individual cells	70
Section 5: Dynamical properties of the regulatory systems for coupled cells	71
SUPPLEMENTARY INFORMATION	72
REFERENCES	88

FIGURES LIST

Figure I1. Promoter states

Figure I2. Transcriptional and translational bursting

Figure I3. The three-stage model of gene expression

Figure I4. Some kinetic parameters of transcriptional bursting

Figure M1. General flux diagram of methodological sessions

Figure M2. The three-stage model of gene expression and its biochemical reactions

Figure M3. Next Generation Method (NGM)

Figure M4. The Chemical Master Equation (CME) for the three-stage model of gene expression

Figure M5. General flux diagram of methodological session 3

Figure M6. Chemical Langevin Equation for burst in mRNA and not in protein (CLE 1)

Figure M7. General flux diagram of methodological session 4

Figure M8. Evaluation of noise propagation from regulator to regulated gene

Figure M9. General flux diagram of methodological session 5

Figure M10. Chemical Langevin Equation for burst in mRNA and not in protein (CLE 1) with diffusion

Figure 1. The self-activated gene and the community effect

Figure 2. The Nodal community effect in the sea urchin embryo oral ectoderm

Figure 3. Activator - Inhibitor system

Figure 4. Temporal expression characteristics of one gene with noise simulated with NGM

Figure 5. Noise (FF) in mRNA expression of the self-activated gene for a range of some parameters values

Figure 6. Noise (FF) in protein expression of the self-activated gene for a range of k_{on} and k_{off} values

Figure 7. Noise (FF) in protein expression of the self-activated gene for a range of k_{off} and kd_{mRNA} values

Figure 8. Noise (FF) in protein expression of the self-activated gene for a range of kd_{mRNA} and kd_p values

Figure 9. Temporal expression characteristics of one gene with noise simulated with CLE1

Figure 10. Temporal expression characteristics of the self-activated gene with noise

Figure 11. Noise (FF) in the protein expression of the regulatory systems evaluated for a range of k_{off} and kon parameter values

Figure 12. Noise (CV^2) in protein expression of the regulatory systems evaluated for a range of k_{off} and kon parameter values

Figure 13. Mean of the protein expression of the regulatory systems evaluated for a range of k_{off} and kon parameter values

Figure 14. Regions that represent the four gene expression types

Figure 15. Dynamical characteristics of the four gene expression types at the mRNA level

Figure 16. Dynamical characteristics of the four gene expression types at the protein level

Figure 17. Noise (FF) in the protein expression of an Activation and Inhibition types of regulation with variations in the regulator

Figure 18. Noise (FF) in the protein expression of an Activation and Inhibition types of regulation with variations in the regulated gene

Figure 19. Temporal expression in regions 2 and 3-4 of a gene with and without regulator

Figure 20. Noise (FF) and mean of the protein expression of an inhibited gene for variations in its inhibitor

Figure 21. Temporal expression characteristics of both inhibitor and inhibited gene

Figure 22. Noise (FF) of the protein expression of a unicellular and multicellular approach

Figure 23. Mean, FF, and Entropy of the protein expression of a multicellular approach

Figure 24. Temporal expression characteristics for a multicellular approach in two regions of low and high noise

Figure 25. Estimates for a multicellular approach when it is changed the k_{on} and k_{off} parameters to represent the four regions of expression

Figure S1. Steady-state distribution of one gene with noise simulated with NGM

Figure S2. Temporal expression characteristics of one constitutive gene

Figure S3. Temporal expression characteristics of two genes according to parameter values reported in the literature

Figure S4. Noise (CV^2) in mRNA expression of the self-activated gene for a range of some parameters values

Figure S5. Noise (CV^2) in protein expression of the self-activated gene for a range of k_{on} and k_{off} values

Figure S6. Noise (CV^2) in protein expression of the self-activated gene for a range of k_{off} and kd_{mRNA} values

Figure S7. Noise (CV^2) in protein expression of the self-activated gene for a range of kd_{mRNA} and kd_p values

Figure S8. Noise (FF) in mRNA expression of the regulatory systems evaluated for a range of k_{off} and k_{on} parameter values

Figure S9. Noise (CV^2) in mRNA expression of the regulatory systems evaluated for a range of k_{off} and k_{on} parameter values

Figure S10. Mean of the protein expression of the regulatory systems evaluated for a range of k_{off} and k_{on} parameter values

TABLES LIST

Table I1. Some kinetic parameters of transcriptional bursting

Table M1. Three cases of the ratio between promoter rates (k_{on} / k_{off}) with three affinities or promoter velocities (ϵ)

Table M2. Propensity functions for the reactions of CLE1 system

Table M3. Parameters values range

Table M4. Some propensity functions for reactions of the regulatory systems

Table M5. Parameter values ranges

Table 1. Estimation of some kinetic parameters of transcriptional bursting

Table S1. Symbols of the parameters used in this study

Table S2. Activation and Inactivation promoter rate

Table S3. Transcription rate

Table S4. mRNA degradation rate

Table S5. mRNA transcriptional burst size

Table S6. Translation rate

Table S7. Protein degradation rate

Table S8. Extracellular Effective diffusion coefficients

Table S9. Default parameters values of the regulatory systems evaluated in this study

Table S10. Comparison of FF, CV2, and mean of molecules between an unregulated gene and a self-activated gene, when it is changed the parameters' values of the regulator

Table S11. Comparison of FF, CV2, and mean of molecules between an unregulated gene and a self-activated gene, when it is changed the parameters' values of regulated gene

ABBREVIATION LIST

- TF** : Transcription Factor
CME : Chemical Master Equation
CLE : Chemical Langevin Equation
NGM : Next Generation Method
FF : Fano Factor
GRN : Gene Regulatory Network
T-leap : tau leap method
JSD: Jensen–Shannon divergence
PDF: Probability Density Function
TGF β : Transforming Growth Factor beta
TCF: T-Cell Factors
SPK : Slow Promoter Kinetic
BMP: Bone Morphogenetic Protein

GLOSSARY

Biological noise: It refers to the heterogeneous expression of a gene among a group of cells. This is due to variations in the levels of expression in each cell.

Burst Expression: It is a type of gene expression in which the mRNA or the protein is produced in high levels for a while and then the production is ceased.

Constitutive Expression: It is the continuous production of a gene.

Transcription Factors: It includes a wide number of proteins that initiate and regulate the transcription of genes. They have DNA-binding domains that give them the ability to bind to specific DNA sequences called regulatory sequence (e.g. enhancer) or promoter sequences.

Promoter: Some transcription factors bind to a DNA sequence near the transcription start site. They help to form the transcription initiation complex for transcription initiation.

Regulatory sequence: Some transcription factors bind to regulatory sequences, such as enhancer sequences, and can either stimulate or repress transcription of the related gene. These regulatory sequences can be thousands of base pairs upstream or downstream from the gene being transcribed.

Gene Regulatory Network: It is the total map of regulatory interactions between genes. The genes composing these networks encode TFs, as well as signaling ligands and receptors for intercellular communication. Each gene also has sequences that control the expression of itself. Thus, the GRN components are elements of coding and non-coding DNA sequence.

Gene Regulatory Network Motif: It is a small set of genes that regulate each other in a particular way. This pattern of regulation appears in a GRN more frequently than it is expected by randomness.

Subcircuit: It is a set of functionally linked regulatory genes which together execute a developmentally defined job.

Three-state gene model: It summarizes the gene expression process in three main parts. Regulation to promoter level (i.e., promoter activation and inactivation), mRNA transcription and degradation, and protein translation and degradation.

ABSTRACT

Biological noise (or noise) refers to the heterogeneous expression of a gene among a group of cells. This is due to variations in the levels of expression in each cell. These variations are due to the kinetics of reactions that make part of the gene expression system. On the other hand, in development, organisms grow with the same spatial and temporal patterns, with few variations among individuals. For this reason, the noise should be filtered in the embryonic differentiation process. Here, we evaluated some gene regulatory systems as mechanisms for filtering biological noise in development. For this, each gene of the regulatory system was represented by a three-stage model for each cell. The regulation was represented by a Hill function. The systems were simulated with the Gillespie's Algorithm (NGM), Chemical Master Equation (CME), and Chemical Langevin Equation (CLE). The noise was measured with the Fano Factor, and it was estimated for each regulatory system in a range of kinetic parameters, and in a unicellular system without diffusion and in a multicellular system with diffusion. The regulatory systems evaluated do not reduce the biological noise. On the contrary, the noise remains at the same level as in the self-activation system or increases as in the Activator-Inhibitor system in comparison with an unregulated gene. But the diffusion of a paracrine signal throughout a colony of cells decreases the noise in the whole range of values of the parameters evaluated.

INTRODUCTION

Biological noise and gene expression as transcriptional bursting

Biological noise (or noise) refers to the heterogeneous expression of a gene among a group of cells. This is due to variations in the expression levels of the gene in each cell [1]. This variability in gene expression is even present in a genetically identical cell population (i.e., tissues) under the same and fixed environmental conditions [2–4]. Because this noise occurs at the single-cell level, it is commonly hidden in cell population studies where the expression is averaged [5]. However, single-cell and single-molecule approaches reveal the stochastic nature of gene expression [6–9].

There are two different sources of biological noise, *extrinsic* and *intrinsic noise* [10]. The extrinsic noise is due to cell-to-cell fluctuations of signaling cascades components [11], transcription factors (TFs) [12,13], and gene expression machinery (i.e., RNA polymerase) [14]. This is a consequence of environmental diversity, micro-fluctuations in the cellular environment, and asymmetric partitioning of molecular compounds in cell division [12,13]. On the other hand, the intrinsic noise is an inherent consequence of the stochastic nature of biochemical reactions involved in *gene expression* [15,16]. It means that chromatin modifications like promoter activation and inactivation, transcription, translation, and degradation reactions are sources of intrinsic noise [1,17].

Although extrinsic noise could be reduced by fixing the environmental conditions, there is always going to be intrinsic noise. Some studies have revealed that intrinsic factors have more weight in the biological noise than extrinsic ones [18,19]. For instance, intrinsic noise can play a significant role in determining the behavior of individual cells when the number of molecules is small [14,15]. But even when the number of molecules is not small, intrinsic noise plays a significant role when there is regulation of expression at the promoter level [20].

The regulation of expression at the promoter level means that genes do not have a constitutive expression [21]. Conversely, they have random transitions between an *active state* ('ON'), in which mRNA is transcribed, and an *inactive state* ('OFF'), in which mRNA is not transcribed (Fig. I1)[22]. The chromatin remodeling is an important process for the transition between these states [15]. For instance, the 'ON' state is associated with an open

active chromatin state, whereas the ‘OFF’ state is usually associated with a closed chromatin state in which the binding sites for TFs are not available [21].

Unlike the inactivation, promoter activation requires many sequential steps in the chromatin remodeling process [23,24]. These steps imply many different chromatin states and the transitions between them are probabilistic [15]. However, If one of these chromatin states is rate-limiting then the system may respond as if there is a single chromatin state transition [24]. And therefore, the active and inactive state could be summarized in the *two-states chromatin model* previously mentioned (Fig. I1)[15,24].

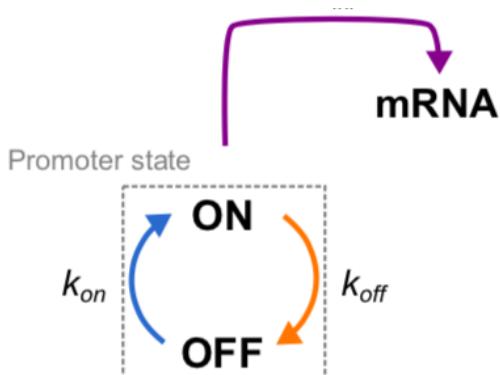


Figure I1. Promoter states. The promoter transits between the active (ON) and inactive (OFF) state with transition rates k_{on} and k_{off} , respectively. The mRNA is only produced in the active state (adapted from [24]).

As a consequence of the cyclical dynamics between the promoter activation and inactivation, transcription commonly occurs in a discontinuous manner over time [25]. It means, most genes are transcribed during short periods called *transcriptional bursts*, interspersed by silent intervals (Fig. I2)[26–28]. *Transcriptional bursting* is a widespread phenomenon that has been observed across many species [2,25]. And it is also possible to find *translational bursting* (Fig. I2)[29]. The frequency and size of these bursts affect the magnitude of temporal fluctuations in mRNA and protein content within the cell [11].

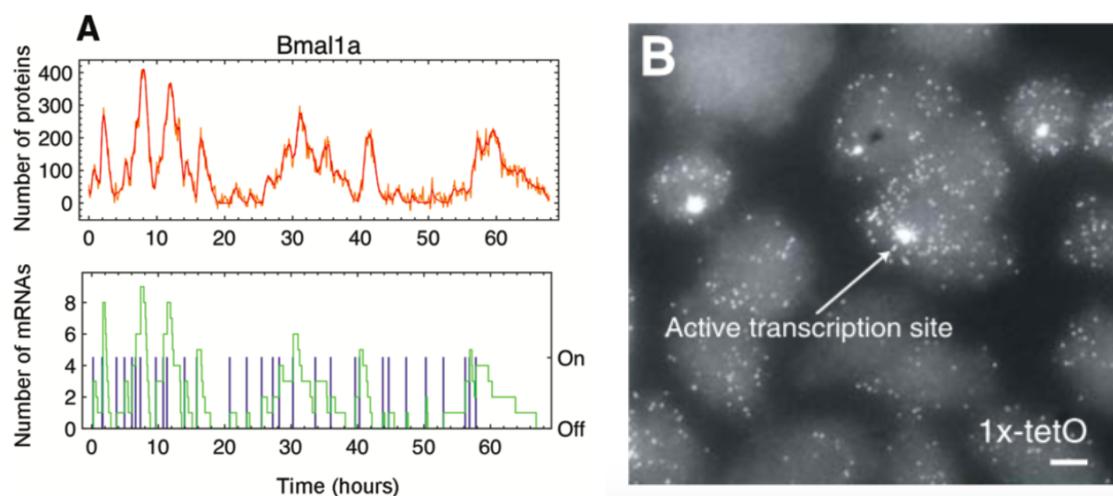


Figure I2. Transcriptional and translational bursting. A. Example of luminescence trace (orange), protein copy number (red), mRNA copy number (green), and promoter activity inferences (blue) for *Bmal1a* promoter (adapted from [25]), B. Cells from *E-YFP-M1-1x* cell line, containing the 1x-tetO promoter, where each mRNA is hybridized to FISH probe P1-TMR. The image was obtained by merging a three-dimensional stack of images (adapted from [2]). The occasional larger bright areas are recently activated transcription sites caused by a buildup of nascent mRNA that has not yet diffused away. These sites occur infrequently during brief periods when the gene is transcriptionally active. The rest of the time, the gene is in a transcriptionally inactive state, during which no mRNA molecules are synthesized and those synthesized earlier are degraded [2].

As it has been mentioned, gene expression is highly variable at the single-cell level and discontinuous transcriptional bursting is one of the primary sources of this variability [3,5,28]. This elevated level of noise is inconsistent with constitutive gene expression models [2]. However, there are genes associated with the maintenance of basic cellular function that present low variations [30]. For instance, the housekeeping genes MDN1, KAP104, and DOA1 in budding yeast are constitutively expressed and their expression levels are described by a Poisson distribution [21].

Gene expression kinetics

The kinetic and thermodynamic characteristics of each step in gene expression contribute to generate a characteristic kinetic profile (Fig. I3)[31]. This profile is affected by reactants concentration, the affinity between reactants, and macromolecular crowding inside the cell [32,33]. In this section it will be described some kinetic characteristics of gene expression when the expression is represented by a *three-stage gene model*. This model incorporates regulation to promoter level (i.e., promoter activation and inactivation), transcription and mRNA degradation, and translation and protein degradation (Fig. I3) [32].

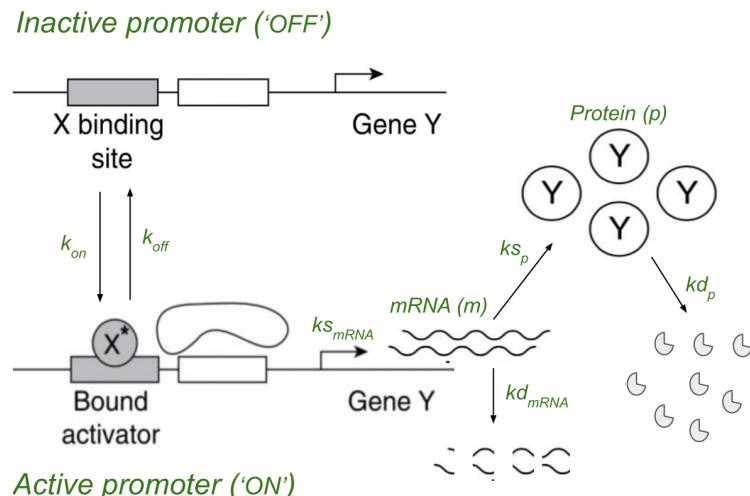


Figure I3. The three-stage model of gene expression. It is represented the inactive promoter and the binding-site, the active promoter once the transcription factor (TF) is bounded to the binding-site, the synthesis and degradation of mRNA, and the synthesis and degradation of proteins (adapted from [34,35]). The kinetic parameters of the model are k_{on} : promoter activation rate, k_{off} : promoter inactivation rate, ks_{mRNA} : transcription rate, kd_{mRNA} : mRNA degradation, ks_p : translation rate, kd_p : protein degradation.

1. Promoter activation and inactivation

The *two-states chromatin model* assumes that a gene can fluctuate randomly between ‘ON’ and ‘OFF’ promoter states, where mRNA can be transcribed only in the ‘ON’ state (Fig. I1) [24]. The fluctuations between these states are described by a *promoter activation rate* (k_{on}) at which the gene becomes active and a *promoter inactivation rate* (k_{off}) at which the gene becomes inactive [24]. k_{on} and k_{off} could also be interpreted as probabilities of activation and inactivation in a time interval, respectively [15]. The inverse of these transition rates describes the average waiting time of the gene in inactive ($T'off'$: *promoter off period*) and active ($T'on'$: *burst duration*) states, respectively (Fig. I4A) [24]. For some genes, the *promoter-off period* is on the scale of hours while the *burst duration* is in minutes (Fig. I4B).

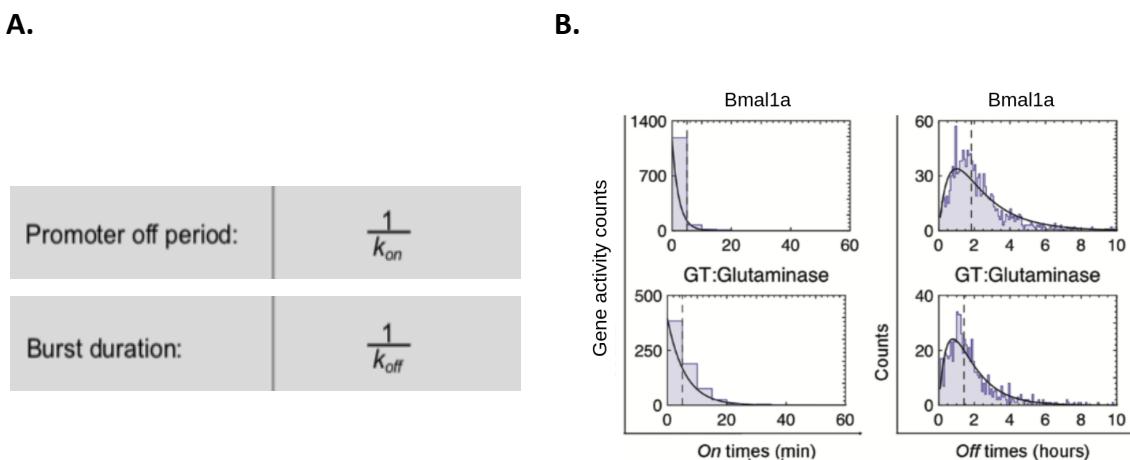


Figure I4. Some kinetic parameters of transcriptional bursting. A. From top to bottom: Promoter off period ($T'off'$) and Burst duration ($T'on'$) (adapted from [24]), B. Left: Distribution of $T'on'$ intervals of Bmal1a and Glutaminase promoters, black lines show exponential fits. Right: Distribution of $T'off'$ intervals of Bmal1a and Glutaminase promoters, black lines show best fits to “two-step” model. Black vertical dotted lines show medians of distributions (adapted from [25]).

k_{off} and $T'on'$ depend on the transcription itself [25,36], because the promoter remains active while the transcription is happening. The activation has approximately a similar duration each time [25,36]. Once the promoter is inactivated, it starts a refractory period in which the promoter is necessarily turned off for some time [25]. This period happens because turning on transcription requires several consecutive steps. These steps must be tightly coordinated to prepare the transcription machinery to be active [31].

In addition to the refractory period, k_{on} is also determined by TF-binding to regulatory sequences [36]. The TF-binding occurs following a trial-and-error sampling mechanism, in which a successful binding is randomly interspersed by many rounds of non-specific TF–chromatin collision events [29]. The TF-binding depends on TF concentration, the time a TF takes to find the binding site, the TF affinity to regulatory sequence, and the number of regulatory sequences [37–39]. The higher some of these factors are the less time the promoter remains off because k_{on} is higher (Fig. I4A)[25,31].

Finally, the noise happens, in part, because each cell in a sample pool is at a different time in the promoter activation process. Implying that, for each allele, some cells might be in the transcriptional ‘ON’ state, whereas other cells are in the ‘OFF’ state [28].

2. Transcription and mRNA degradation

Transcription kinetic is described by the transcription rate ks_{mRNA} , and this plays a role when the promoter is in the ‘ON’ state. ks_{mRNA} in mammalian cells vary over a very large range [31]. The maximal possible transcription rate is reached when elongation becomes limiting for initiation. That is, when the polymerase initiated at the promoter sterically hinders the next polymerase to take its place. Since the elongation rate in mammalian cells is about 63 nucleotides per second and since an RNA polymerase II may occupy a similar number of base pairs, the maximal transcription rate is about one transcript per second [31].

Commonly the transcriptional events occur over periods of minutes rather than seconds [27]. And in a particular cell type, protein-encoding genes can be expressed at widely different levels, producing from less than two mRNAs/cell to more than 10^5 mRNAs/cell. Additionally, a single mammalian cell contains 20–40 pg of total RNA but only 0.5–1.0 pg of mRNA (10^5 to 10^6 mRNA molecules) [5].

The transcriptional bursting efficiency is characterized by two parameters, *burst size* and *burst frequency* (Table I1). The burst size is the average number of synthesized mRNA molecules while a gene remains in the active state. And the burst frequency is the frequency at which a burst occurs per unit of time [24,30]. Another important parameter that determines the gene expression level is the Promoter Occupancy (Table I1). This is the average time a gene is active or the average number of cells in the active state in a time t.

Parameter	Equation	Ref.
Burst size (B.S)	$ks_{mRNA} / (k_{on} + k_{off})$	[40]
Burst frequency (B.F)	$k_{on} * k_{off} / (k_{on} + k_{off})$	[24]
Mean of mRNA molecules (M.M)	$B.F * ks_{mRNA} / kd_{mRNA}$	[21]
Promoter occupancy (P.O)	$k_{on} / (k_{on} + k_{off})$	[24]

Table I1. Some kinetic parameters of transcriptional bursting. Burst Size (B.S), Burst frequency (B.F.), Mean of mRNA molecules (M.M), Promoter Occupancy (P.O).

The number of mRNA molecules decays at a rate per unit of time known as the degradation rate (kd_{mRNA}) and the mRNA lifetimes are usually minutes [32]. From a kinetic perspective, the kd_{mRNA} is predicted to determine the net abundance of all mRNA molecules in the cell and the rates at which mRNA levels will react to input signals.

3. Translation and protein degradation

Protein synthesis occurs at a translation rate of ks_p . ks_p saturates at around 180 protein copies per mRNA per hour [41]. Inside the cell, protein numbers are often in the order of *hundreds* of molecules [2]. Proteins are, on average, five times more stable (median half-life of 46 h) than mRNAs (9 h) and spanned a bigger dynamic range [41]. Typically, proteins exist for at least several mRNA lifetimes, and protein fluctuations are determined by only time-averaged properties of mRNA fluctuations [32].

It has been observed a small, but statistically significant, negative correlation between total noise and the ratio between degradation rate of mRNA and protein ($\gamma = kd_{mRNA} / kd_p$) (a rank correlation of ≈ -0.2 with a P-value of 10^{-6}) [32]. For instance, in budding yeast, it has been reported that values of $\gamma > 1$ reduce protein fluctuations by allowing more averaging of the underlying mRNA fluctuations [32]. Proteins involved in transferring nucleotidyl groups have high median $\gamma > 5$, presumably because high stochasticity in these proteins can undermine many cellular processes [32].

Similarly, proteins that contribute to the structural integrity of protein complexes have a median $\gamma > 5$. It could be maybe because large fluctuations can prevent complete complex formation [32]. But surprisingly, TFs have a low median $\gamma > 1$. Although low γ does increase stochasticity, it can allow quick response times if the protein degradation rate is high. A high protein degradation rate may also keep numbers of TFs low to reduce deleterious nonspecific chromosomal binding [32].

Additionally, it has been pointed out that the degradation rate determines the occurrence of transcriptional/translational burst. For instance, some genes in mammalian cells have only obvious mRNA burst production because $k_{off} \gg kd_{mRNA} \sim kd_p$ [40], whereas some genes in yeast have only protein burst production because $kd_{mRNA} \gg k_{off} \sim kd_p$ [40,42]. Furthermore, some genes in bacteria have both mRNA and protein bursts because $k_{off} > kd_{mRNA} > kd_p$, and some do not have any obvious burst production because $k_{off} \sim kd_{mRNA} \sim kd_p$ [40,42,43].

Gene regulatory networks and Biological noise in embryonic development

The process of development and differentiation is regulated by site-specific DNA-binding proteins that direct transcriptional programs [15]. These proteins are known as TFs and they interact with regulatory sequences to govern the activity states of target promoters [15]. In turn, Gene Regulatory Networks (GRNs) are composed of this regulatory assemble. A GRN is the total map of regulatory interactions between genes [44]. The genes composing these networks not only encode TFs but also signaling ligands and receptors for intercellular communication [44]. Each gene also has sequences that control its expression. Thus, the GRN components are elements of coding and non-coding DNA sequences, and together they constitute the *regulatory genome* [44].

In each developmental control point, multiple inputs from regulatory genes and external signals are processed [45]. In this form, GRNs make part of short time regulation of development [46]. On the other hand, the architecture of developmental GRNs is the direct product of evolution [47]. And throughout the evolutionary time, developmental processes change when this architecture changes, resulting in a great or small change in the outcome of development [44]. This architecture is built of modular blocks called subcircuits [48] and GRN motifs [35]. Sometimes a modular block could be both a subcircuit or a GRN motif although they are not conceptually equal [48].

A subcircuit is a small set of functionally linked regulatory genes which together execute a job in the development [48]. It means, they are defined by their topologies, and the topology of a subcircuit directly indicates its function [48]. This implies that structurally similar subcircuits, but composed of different genes, are repeatedly encountered doing similar jobs [48]. Conversely, a GRN motif is a small set of genes that regulate each other in a particular way [35]. This pattern of regulation appears in a GRN more frequently than it is expected by randomness [35]. GRN motifs occur frequently and in many entirely unrelated developmental subcircuits. They contribute to the properties of subcircuits, for instance, they have implications for the kinetic with which subcircuits operate, and for other developmentally more important information processing functions [49].

Over the past two decades, experimental studies have established that many genes composing subcircuits show substantial and unavoidable stochastic fluctuations in their expression levels [50]. This noise can propagate throughout the GRN [29,40,51]. Even in tissues typically considered to consist of homogeneous populations of cells, intercellular variability in gene expression levels can be considerable. For example, the cells of a genetically identical population grown in the same environment have been shown to display substantial variability in the total number of mRNA molecules [24].

Additionally, it has been reported transcriptional burst expression in some developmental genes. For instance, it has been studied the transcriptional burst expression of genes *hind sight* (*hnt*) and *u-shaped* (*ush*) in *Drosophila*'s embryos. When the expression of these genes is modulated by a gradient of bone morphogenetic protein (BMP) signaling, which occurs in the patterning of ectodermal cell fates along the dorsal-ventral axis [36]. Also in *Drosophila*, it has been reported the transcriptional burst of gap gene *hunchback* in the anterior pole of embryos, when it is under the gradient of Bicoid protein [29]. In addition, *nanog* gene, which safeguards the embryonic stem cell (ESC) ground state, shows drastic transcriptional burst kinetics in mouse ESCs [29]. Finally, it is thought that most of the developmental regulators are transcribed in bursts [29].

It is important to highlight that although biological noise can be an obstacle for some types of cellular functions, it is also a useful feature for others [51]. For example, stochastic gene expression has been linked to phenotypic variability in the control of developmental

transitions in metazoans [25,29]. This is the case of the neural fate which is chosen in a stochastic way [24].

However, animal body structures and tissues exist in robust patterns with few variations among individuals [51]. These patterns ensure proper tissue form and function during early embryogenesis, development, homeostasis, and regeneration [52]. They arise seemingly out of the stochasticity of a few early cells in the embryo [52]. When precise control is needed, it may be necessary to reduce or filter the noise [40]. For this reason, it is expected that in the differentiation process and patterning formation the noise should be filtered in some way [51].

Several mechanisms of filtering noise in mammalian gene expression have been proposed. For instance, cellular compartmentalization was recently shown to work as a transcriptional noise filter through spatial partitioning of molecules in and out of the nucleus [29]. The nuclear retention of most transcripts is about twenty minutes, which is a similar time to scale transcription bursting or mRNA degradation. This can efficiently average out the stochasticity of mRNA concentration by a factor of three to four [29].

Schmiedel et al. [13] showed that microRNAs decreased protein expression noise for lowly expressed genes but increased noise for highly expressed genes. These authors estimated that hundreds of (lowly expressed) genes in mouse ESCs had reduced noise due to substantial microRNAs regulation [11]. Moreover, epigenetic modifications regulate noise through direct modulation of transcriptional bursting. Specifically, DNA methylation and histone deacetylation have been demonstrated to act as noise repressors, whereas histone acetylation does the opposite [61,62].

In addition, it has been reported the regulation of transcriptional bursting in development with consequences in reduction of noise [36]. For instance, in *Drosophila*'s embryos, BMP signaling modulates the k_{on} of their target genes and therefore their burst frequency [36]. As a result, there is an increase in k_{on} along the BMP gradient, but both the $k_{s_{mRNA}}$ and k_{off} remain unchanged [36]. This increase in k_{on} is associated with lower noise [36]. Finally, it has been proposed that feedback and feed-forward motifs are gene-specific solutions for noise regulation [10, 11].

The knowledge of how cells sustain their diverse dynamic behaviors, especially in the presence of biological noise, is still limited and there are a lot of open questions. For instance, how transcription factors influence the precision and robustness of gene regulation, how long and how frequently a gene is actively transcribed during differentiation [27], how transcriptional events are dynamically coordinated in cell populations [27], how cells efficiently and correctly process information in noisy environments [51], how the precise developmental events are controlled under the noisy condition [51], or how can these patterns be reproducible despite the biological noise in gene expression.

Methodological approaches to study the biological noise in GRN

The fluorescent *in situ* hybridization techniques allow to quantify transcriptomes and proteomes with single-molecule sensitivity [9,53–55]. Due to these techniques, it has been some advances for understanding noise in gene regulation in development. Some of these advances are the measures of some kinetic parameters and their effect in burst expression [56], the detection that important morphogens such as Bone Morphogenetic Protein (BMP) affect the burst kinetic of their target genes [20], the identification of noise sources [12], the evaluation of mechanism of noise filtering [57], and to identify the common presence of transcriptional bursts in gene activity in living embryos [37]. Therefore, *in situ* experiments allow to identify the parts of a system and their interaction with each other. Once you get the landscape, *in silico* experiments allow to explore the dynamical behavior of these systems.

For many organisms, *in situ* experiments have shown that gene expression can be described as a three-stage model [32]. For this model, the regulation of expression by a TF is commonly and well described by a Hill function [35]. The Hill function represents whether the gene is activated, inhibited, or both according to particular regulatory system [35]. Additionally, the Hill function should be incorporated in the model as a factor of promoter activation rate. In this way, it represents the effect of TF concentration in the burst frequency as was pointed out in [36].

It is possible to use several approaches to simulate a regulatory system with three-stage model for gene expression and Hill function for gene regulation. Here it is used three common approaches as the Chemical Master Equation (CME) [32,58], Next Generation Method (NGM) [59–61], and Chemical Langevin Equation (CLE) [62,63]. The first one allows to study the steady-state behavior of the system. The second one allows to explore the characteristics of temporal dynamics. The last one have a better performance than the second one. We used these stochastic simulation methods because they better represent the dynamic of small systems size (i.e., low number of molecules are around hundred) [64,65].

Previous studies commonly look at the temporal dynamics to get insights into the gene expression throughout time [10,50]. They combine this with the estimation of the steady-state distribution, because this characterized well the long term behavior of the systems [19,21,24,58]. From this distribution, it is possible to estimate dynamical properties such as noise and mean level of gene expression [2,18,19,21,65].

Noise is commonly measured with the Fano Factor (FF) and coefficient squared (CV^2). The FF is the ratio between the variance and the mean [2]. FF is key to quantify the deviation from Poisson statistics, which is the characteristic behavior of a constitutive gene [21,66]. For a Poisson behavior, the FF is one and it defines a “standard dispersion”, as a result, distributions with FF smaller/bigger than one are considered under/over-dispersed [66]. It

means that a FF different from one implies regulation to promoter level [21,66]. In such a way, FF measures the variability in the state system due to its regulation structures.

Conversely, CV^2 measures the variability due to system size or molecule number [20]. CV^2 is the ratio between the variance and the squared mean [20]. Noise measured with CV^2 tends to increase when the size of the system decreases [65]. This is because changes in the number of molecules are more significant when the number of molecules is small than when it is greater. This phenomenon is known as the finite-number effect [65].

Finally, it has been described that changes in some kinetic parameters affect the temporal dynamics of expression. Specifically, the promoter activation rate [54,65,67,68], promoter deactivation rate [56,65], mRNA degradation rate [22,56,65], and protein degradation rate [2,69]. Therefore, we will estimate the dynamical properties of some regulatory systems for a range of values of these parameters. Additionally, for a cell colony, the diffusion rate and colony size are parameters affecting the amount of variability in gene expression levels between cells [70].

Problem and justification

In this work, we studied the noise propagation in two regulatory systems commonly present in pattern formation and differentiation in animal embryonic development. The main questions are: How are the dynamical properties (i.e. noise and expression level) of these regulatory system in a cell throughout a range of parameters values (i.e. kinetic parameters)? Are there differences in the temporal dynamics of expression throughout this range of parameters values? How does the noise propagate from regulator to regulated gene? How are the dynamical properties (i.e. noise and expression level) of these regulatory system in a colony of cells throughout a range of parameters values (i.e. kinetic parameters, diffusion rate, and colony size)? Could we describe these regulatory systems as noise-filtering mechanisms?

None of the previous studies of noise in GRN have evaluated the noise for a range of parameters' values or have used real parameters values [51,71]. Nowadays, several experimental studies have determinated the values of some kinetic parameters affecting noise in gene expression [9,31,72–74]. This make possible to study the dynamical properties and noise propagation in GRN systems for a set of parameters. Additionally, some previous studies have studied self-regulated system but there is opposite conclusions about its noise controller capacities [51,71]. But none of the previous studies have studied the noise characteristics of Activator-Inhibitor system.

Objectives

1. General objective

To study the dynamical properties of some regulatory systems in embryonic animal development. Particularly, when biological noise is included in gene expression.

2. Specific objectives

1. To select, by biological interest and dynamical complexity, some regulatory systems in embryonic animal development.
2. To design dynamical equations that describe the expression of each gene in the selected regulatory systems in such a way that the biological noise is included in gene expression.
3. To study the dynamical properties of these systems according to parameters variations.

METHODOLOGY

The structure of methodology sections is designed to accomplish orderly each one of the specific objectives (Fig. M1). To accomplish the first objective, we described some criteria to select the regulatory systems in section one (Fig. M1A). To design the dynamical equations of the selected regulatory systems and accomplish the second objective, we followed the next two steps (Fig. M1B). In the first place, in section two, we used the Next Generation Method (NGM) and Chemical Master Equation (CME) to get insights into the three-stage gene model for gene expression.

In the second place, in section three, we evaluated the similitude in the results obtained by NGM and Chemical Langevin Equation (CLE) when they simulated the three-stage gene model of one self-activated gene (Fig. M1B). As it has been pointed out in previous work, the CLE has better performance than the NGM[75]. For this reason, we went on to this step to be sure that CLE is a good approach to simulate the selected regulatory systems. In this way, as we will see further for simulations in a colony of coupled cells, we can use CLE in cases that it is not possible to use NGM due to its low performance.

We calculated and plotted different measures to accomplish the third objective and to study the dynamical properties of the selected regulatory systems in different parameters values. For instance, in section four and five, we calculated dynamical properties such as the noise and the mean of molecules of mRNA and proteins in the steady-state (Fig. M1C). We also produced plots of the temporal dynamics of each system for some parameter values, the steady-state distribution of molecules, and the autocorrelation and velocities of the temporal dynamics.

We study these dynamical properties in decoupled cells as well as in coupled cells in sections four and five, respectively (Fig. M1C). In section four, these measurements represent the

dynamical properties of a regulatory system in one cell throughout time or in more than one decoupled cell at a given time. In section five, we simulated a population of isogenic cells in a circular colony similar to micropattern colonies. Each cell is represented by a regulatory system and they are coupled by the diffusion of a genetic product (i.e., a paracrine signal).

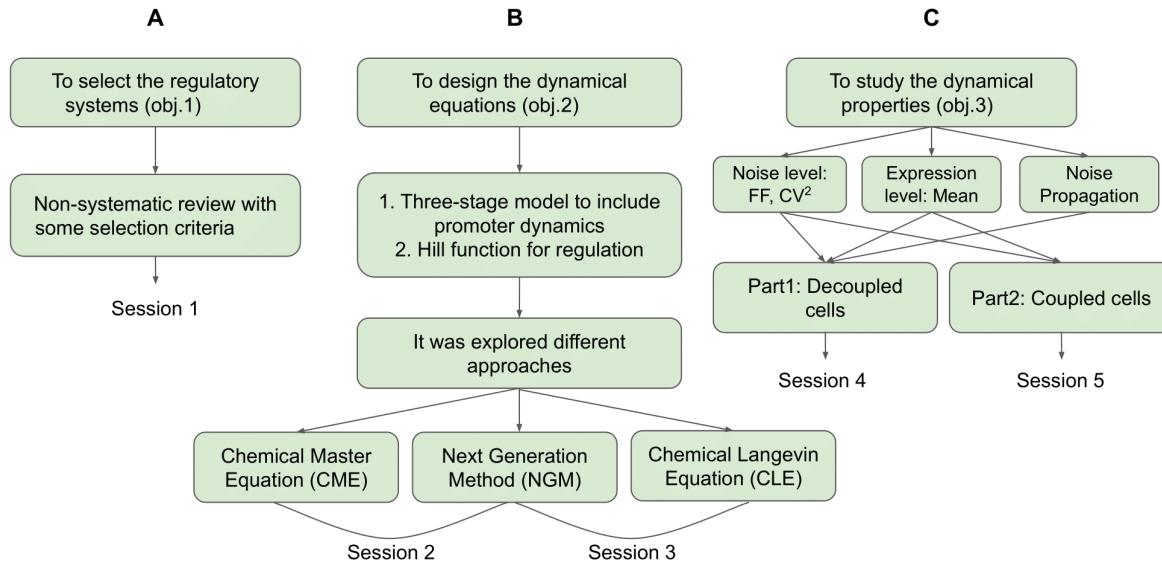


Figure M1. General flux diagram of methodological sessions. This diagram describes briefly the methodological process done for each objective, A. First objective (obj.1) was developed in session 1, B. Second objective (obj.2) was developed in session 2 and 3, C. Third objective was developed in session 4 and 5.

Finally, we used Wolfram Mathematica version 12.3.1.0 to write the computational code for each model and simulation. The results of each simulation were saved as CSV files and this software was also used to analyze them and to build the figures. All the scripts used are in https://github.com/LinaMRuizG/BiologicalNoise_DevelopmentalGeneRegulatoryNetworks.git as well as some CSV files of results.

Section 1: The regulatory systems

We made a non-systematic review of the literature to look for regulatory systems reported for embryonic development. To make the regulatory systems' selection we considered three aspects. First, they participate in embryonic pattern formation. Second, their genetic products act as paracrine signals that diffuse outside of cells. And third, some of their kinetic parameters are recorded in literature. Finally, we considered the amount of time it should be required for *in silico* experiments to select the number of systems to analyze.

Section 2: The three-stage model of one gene with NGM and CME

In this section, we simulated the three-stage model of one unregulated gene with NGM and used CME as an approximation to understand the steady-state behavior. This was the first step to build the dynamical system of equations and to explore the measurements to

characterize the dynamical properties of systems (i.e. kinetic parameters in Table I1, noise and expression level). This is also a short and first exploration of the effect of the kinetic parameters (i.e. promoter velocity and ratio between k_{on} and k_{off}) on this simplest system.

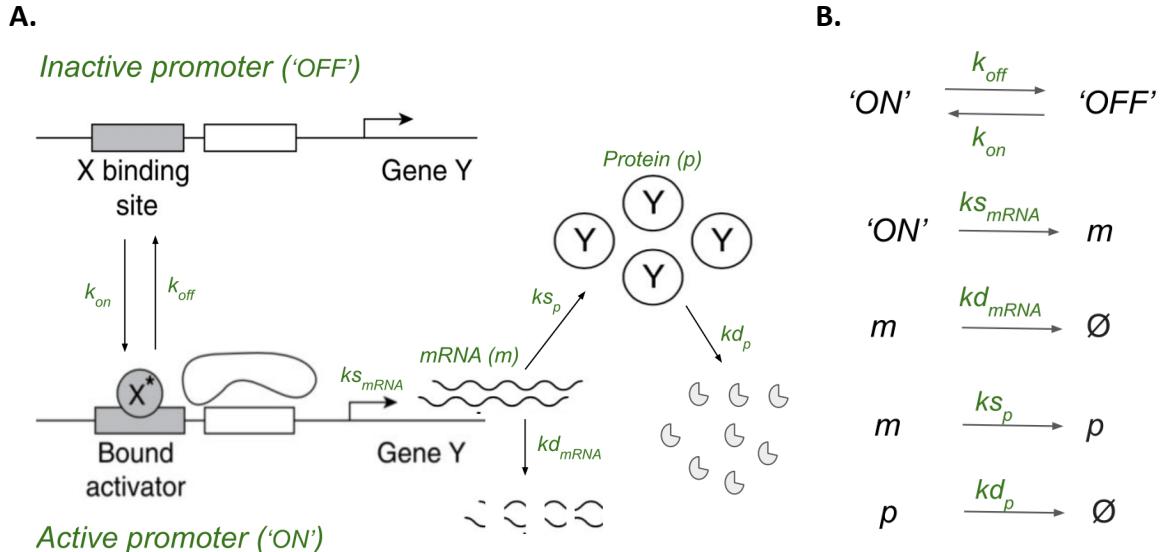


Figure M2. The three-stage model of gene expression and its biochemical reactions. A. Representation of the model (adapted from [34,35]), B. The biochemical reactions and their kinetic rate (above the arrows). The kinetic parameters of the model are k_{on} : promoter activation rate, k_{off} : promoter inactivation rate, ks_{mRNA} : transcription rate, kd_{mRNA} : mRNA degradation, ks_p : translation rate, kd_p : protein degradation.

We used the three-stage model of one gene to incorporate the dynamic at the promoter level in gene expression (Fig. M2). This model includes a couple of biochemical reactions at the promoter, mRNA, and protein level (Fig. M2A). The reactions are activation and inactivation of the promoter, synthesis and degradation of mRNAs, and synthesis and degradation of proteins. All these reactions are first-order chemical reactions (Fig. M2B).

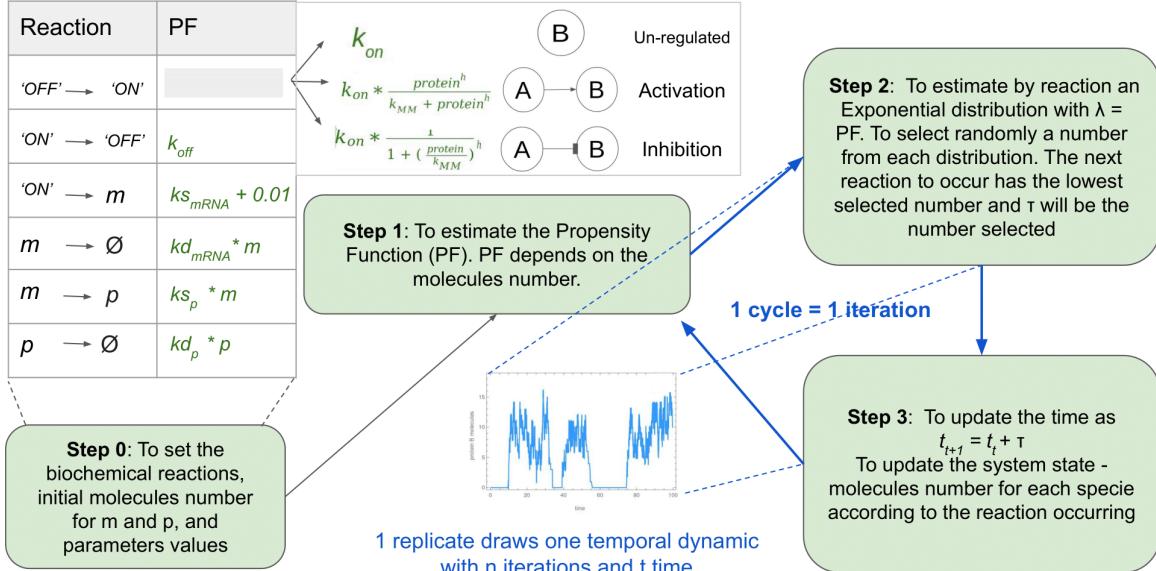


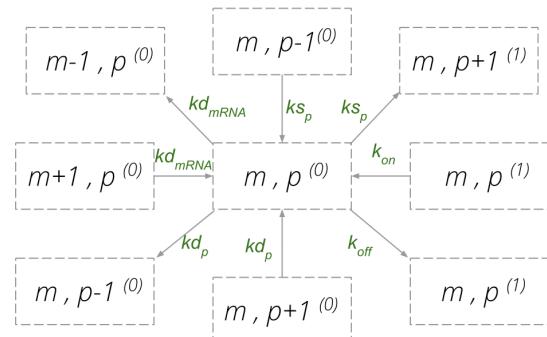
Figure M3. Next Generation Method (NGM). It summarizes the steps required in the NGM for the simulation of a regulatory systems using the three-stage model for gene expression and Hill function for gene regulation.

The Propensity Functions (PF) are estimated for each reaction after setting the biochemical reactions, the parameters' values, and the initial number of molecules for each specie (Fig. M3). Then in each iteration, it is selected one reaction and a time τ to occur (Fig. M3)[59]. The occurrence of a reaction adds or subtracts in one unit the number of molecules of mRNA or proteins, or changes the promoter state [59]. In this way, NGM tracks the change in time of m number of mRNA molecules, p number of protein molecules, and *active* or *inactive* promoter state [34][59]. On the other hand, the CME tracks the change in time of the *probabilities* to be in a state of m number of mRNA molecules, p number of protein molecules, and *active* or *inactive* promoter state (Fig. M4) [25,32,76,77]. In this section, we only explored the mRNA states although we simulated the whole system.

A.

$$\begin{aligned} \frac{d P_{m,p}^{(0)}}{d t} &= k_{off} P_{m,p}^{(1)} - k_{on} P_{m,p}^{(0)} \\ &+ kd_p [(p+1)P_{m,p+1}^{(0)} - n P_{m,p}^{(0)}] \\ &+ kd_{mRNA} [(m+1)P_{m+1,p}^{(0)} - m P_{m,p}^{(0)}] \\ &+ ks_p m [P_{m,p-1}^{(0)} - P_{m,p}^{(0)}] \end{aligned}$$

C.



B.

D.

$$\begin{aligned} \frac{d P_{m,p}^{(1)}}{d t} = & -k_{off} P_{m,p}^{(1)} + k_{on} P_{m,p}^{(0)} \\ & + kd_p [(p+1)P_{m,p+1}^{(1)} - pP_{m,p}^{(1)}] \\ & + ks_{mRNA} [P_{m-1,p}^{(1)} - P_{m,p}^{(1)}] \\ & + kd_{mRNA} [(m+1)P_{m+1,p}^{(1)} - mP_{m,p}^{(1)}] \\ & + ks_p m [P_{m,p-1}^{(1)} - P_{m,p}^{(1)}] \end{aligned}$$

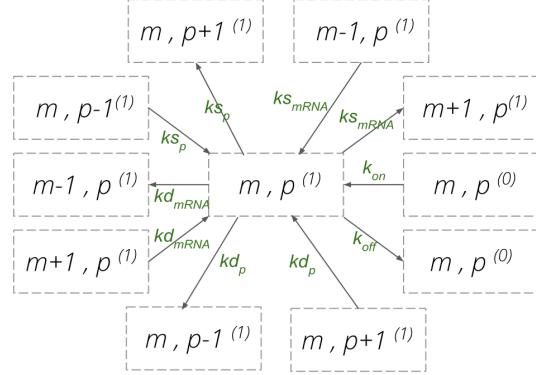


Figure M4. The Chemical Master Equation (CME) for the three-stage model of gene expression. A-B. The two coupled equations of the CME. They represent inactive (0) and active (1) promoter states, respectively. C-D. Representation of the transitions between states of mRNA (m) and protein (p) molecules for inactive (0) and active (1) promoter states, respectively. Each transition has a rate associated, k_{on} : promoter activation rate, k_{off} : promoter inactivation rate, ks_{mRNA} : transcription rate, kd_{mRNA} : mRNA degradation rate, ks_p : translation rate, kd_p : protein degradation rate (adapted from [32]).

We used these two approaches to simulate the system with different values of activation and inactivation promoter rate (k_{on} , k_{off}) and transcription rate (ks_{mRNA}). Specifically, we chose three cases to evaluate the effect on temporal dynamics of the ratio between promoter rates (k_{on}/k_{off}). The cases are: k_{on} one order of magnitude higher than k_{off} , both k_{on} and k_{off} with equal value, and k_{off} one order of magnitude higher than k_{on} (Table M1). For each case, we evaluated different affinities or promoter velocities (ϵ : 100, 1, 0.01 which are high, medium, and low, respectively) (Table M1). The ks_{mRNA} was increased in the first case from ten to one hundred.

	Cases	k_{on*} / k_{off*}	$\epsilon = 0.01$	$\epsilon = 1$	$\epsilon = 100$
k_{on} / k_{off}	Case 1	10 / 1	0.1 / 0.01	10 / 1	1000 / 100
	Case 2	1 / 1	0.01 / 0.01	1 / 1	100 / 100
	Case 3	1 / 10	0.01 / 0.1	1 / 10	100 / 1000

Table M1. Three cases of ratio between promoter rates (k_{on} / k_{off}) with three affinities or promoter velocities (ϵ). $k_{on/off} = k_{on*}/k_{off*} * \epsilon$.

For each set of parameters in Table M1, we simulated 100 replicates with 1000 iterations each. To explore the system, we plotted the temporal dynamic, we built the distribution of the differences between consecutive times of the temporal dynamic (hereafter referred to as *temporal differences distribution*, Δt). We also solved the CME and plotted the steady-state probabilities for each possible state of the system (hereafter referred to as *steady-state distribution*).

We measured the Factor Fano (FF), the variation coefficient squared (CV^2), and the mean of molecules number for both distributions. The measurements were made for each replicate and we averaged over them. For the steady-state distribution of molecules, FF and CV^2 are

measurements of biological noise. We also calculated some kinetic parameters shown in Table I1. We compared our results of a three-stage gene model with those for a constitutive gene. Finally, we simulated the three-stage model of one gene expression with biological noise for *plectin1* and *ush* genes with some parameters reported [25,36].

Section 3: The CLE implementation for the three-stage model is a good approach for large systems

The NGM is a scheme that simulates every reaction event and generates trajectories that follow the states probability distribution [77]. Since each reaction involves only a small change in molecular numbers, the process is time-consuming for large systems [40,77]. Therefore, in this section, we compared the CLE implementation with the NGM when both simulate a three-stage model (Fig. M5). The first one is known as a more efficient approach [40]. This was the second step to build the dynamical system to simulate the selected systems in a colony of cells.

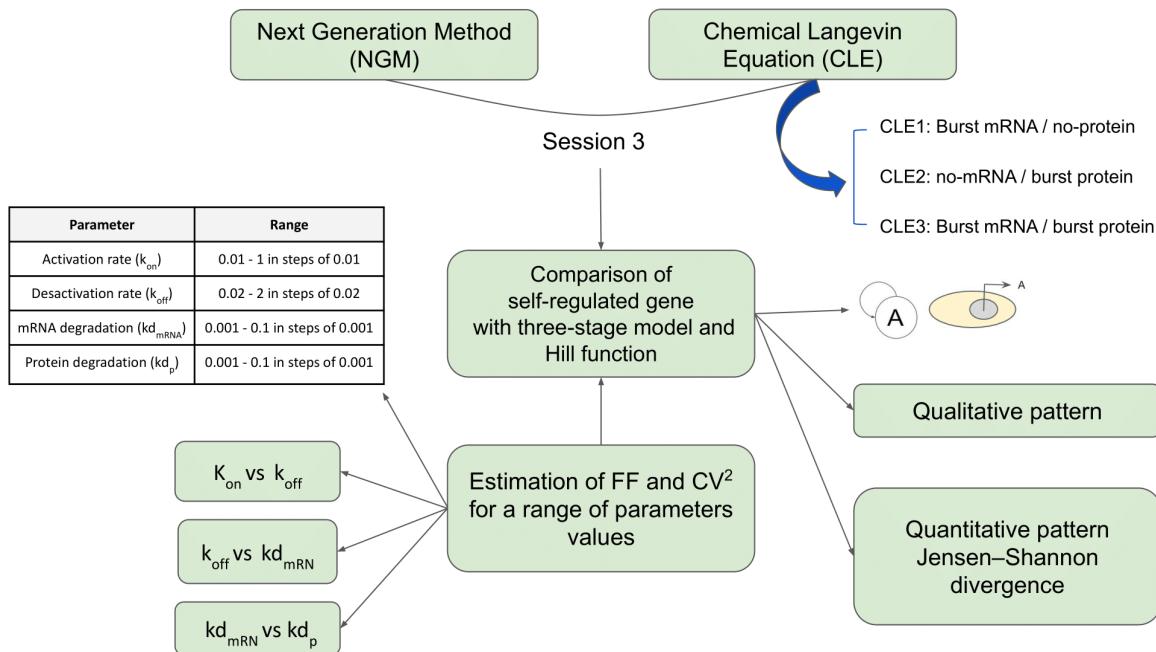


Figure M5. General flux diagram of methodological session 3. It summarizes the steps required for the comparison between NGM and CLE of a self-activated gene in a decoupled cell. The simulation of regulatory system uses the three-stage model for gene expression and Hill function for gene regulation.

We selected the self-activated regulatory system for the approaches comparison (Fig. M5). As it was developed in Yan *et al.* 2017 [40], we made simulations with the CLE implementation for burst production in mRNA and not in protein (CLE1) using the next equations when $k_{off} \gg kd_{mRNA} \sim kd_p$:

$$m(t + 1) = m(t) + (k_{on}\tau b_m \theta + \sqrt{k_{on}\tau b_m \theta(2b_m + 1)} N_1(0, 1)) - (kd_m \tau m(t) + \sqrt{kd_m \tau m(t)} N_2(0, 1)),$$

$$p(t + 1) = p(t) + (ks_p \tau m(t) + \sqrt{ks_p \tau m(t)} N_3(0, 1)) - (kd_p \tau p(t) + \sqrt{kd_p \tau p(t)} N_4(0, 1)).$$

Where m is mRNA molecules, p is protein molecules, τ is the step size, θ is the Hill function for regulation, k_{on} is the promoter activation rate, b_m is the mRNA burst size, $N_i(0,1)$ is the white noise, kd_m is the mRNA degradation rate, ks_p is the translation rate, and kd_p is the protein degradation rate. We also made simulations with CLE implementation for burst in protein but not in mRNA (CLE2) using the next equations when $kd_{mRNA} \gg k_{off} \sim kd_p$:

$$g(t + 1) = g(t) + k_{on}(1 - g(t))\theta\tau - k_{off}g(t)\tau,$$

$$p(t + 1) = p(t) + (g(t)ks_m\tau b_p + \sqrt{g(t)ks_m\tau b_p(2b_p + 1)} N_1(0,1)) - (kd_p\tau p(t) + \sqrt{kd_p\tau p(t)} N_2(0,1)).$$

Where g is the promoter state, k_{off} is the promoter inactivation rate, b_p is the protein burst size, and ks_m is the transcription rate. Finally, we made simulations with CLE implementation for burst both in mRNA and in protein (CLE3) using the next equation when $k_{off} \gg kd_{mRNA} \gg kd_p$:

$$p(t + 1) = p(t) + (k_{on}\tau b_m b_p \theta + \sqrt{k_{on}\tau b_m b_p \theta(2b_m b_p + 2b_p + 1)} N_1(0,1)) - (kd_p\tau p(t) + \sqrt{kd_p\tau p(t)} N_2(0,1))..$$

The Hill function for self-activation is $\frac{protein^h}{k_{aa} + protein^h}$ with k_{aa} equal to 100. In the CLE implementation, the *tau leap* (τ -leap) method is used to estimate the time step [40,77]. The τ -leap method accelerates the simulation with a long leaping-time step that allows several reaction events together [40]. The τ -leap is accurate as long as no propensity function changes its value significantly during any time step [77].

We followed the general description in Yan *et al.* 2017 [40] to calculate the τ (Table M2). For this, we used equations 1-3. The set of species $i \in S$ depends on the system of equations (Table M2). For instance, if the system is described by CLE1 there are two species: mRNA and protein molecules. The number of molecules of each species is represented by x_i . The relative tolerance ϵ was fitted to 0.03. The highest reaction order g_i for all species was 1. For each species i , the change in its state with j reaction (v_{ij}) was 1 for production and -1 for degradation.

$$\tau' = \text{Min}_{i \in S} \left\{ \frac{\max\{\epsilon x_i / g_i, 1\}}{|\mu_i|}, \frac{\max\{\epsilon x_i / g_i, 1\}^2}{\sigma_i^2} \right\} \quad (1)$$

$$\mu_i = \sum_j v_{ij} a_j \quad (2)$$

$$\sigma_i^2 = \sum_j v_{ij}^2 a_j \quad (3)$$

Table M2 shows the propensity function a_j for each one of the j reactions of each species for the CLE1 system. Some of these are critical reactions when their reactants become critical,

which means that reactants reach a low number of molecules and then could become zero within ten reactions. For instance, production is always non-critical but degradation is critical because a critical reactant could become zero when degradation occurs ten times. Reaction propensities involving critical reactants are summed up as

$$a_{0,cr} = \sum_{j \in J_{cr}} a_j$$

where J_{cr} means the j critical reaction of a critical reactant. In this case, an additional τ is estimated (τ'') as

$$\tau'' = \frac{1}{a_{0,cr}} \ln\left(\frac{1}{1-r_1}\right).$$

Where r_1 is a random number from a uniform distribution in the unit interval. Finally, τ was estimated as

$$\tau = \min\{\tau', \tau''\}.$$

Reaction (j)	Propensity Function (a_j)
mRNA production (non-critical)	$b_m k_{on} \theta$
mRNA degradation (critical)	$m(t) \gamma_m$
protein production (non-critical)	$k_{sp} m(t)$
protein degradation (critical)	$p(t) \gamma_p$

Table M2. Propensity functions for the reactions of the CLE1 system.

There is an additional change to equations 1-3 when some specie is produced in bursts. In this case, it should be taken by separate the production from the other reactions of this specie. This is well described in Yan *et al.* 2017 [40]. Finally, to avoid negative production we did not follow the description in [40], but we set a basal production of 0.01 each time the result was lower than zero.

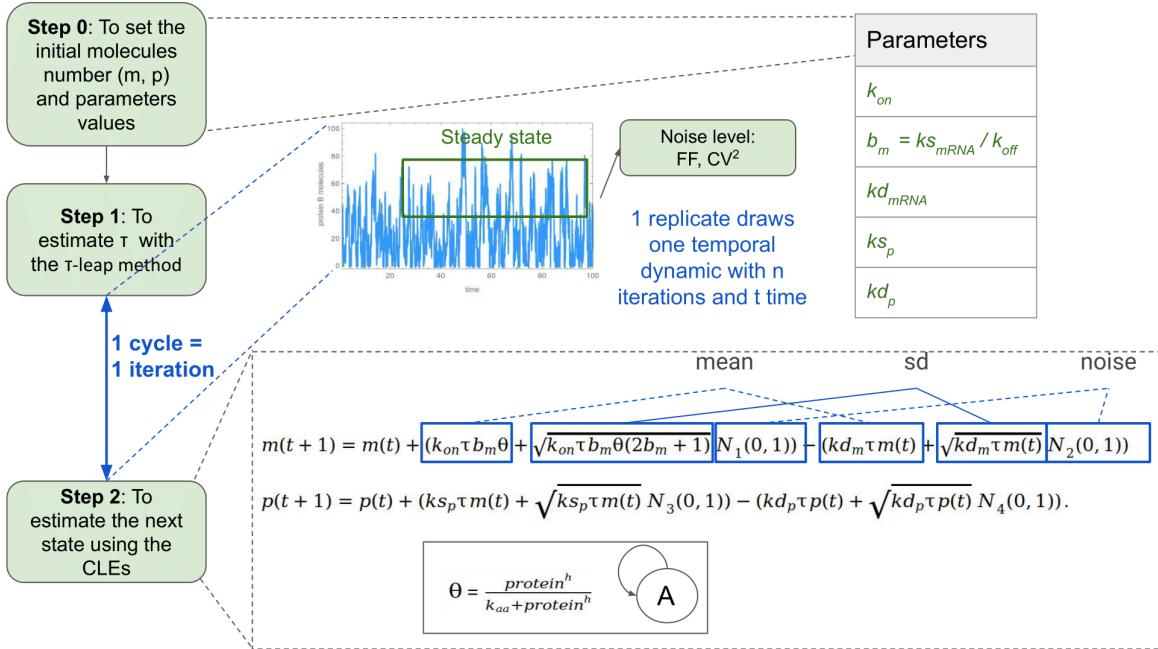


Figure M6. Chemical Langevin Equation for burst in mRNA and not in protein (CLE 1). It summarizes the steps required in the CLE method for the simulation of a self-activated gene in a decoupled cell. The simulation of regulatory system uses the three-stage model for gene expression and Hill function for gene regulation.*sd:standard deviation.

Figure M6 summarizes the CLE implementation for a self-regulated gene with burst in mRNA but not at protein. We simulated the temporal dynamics of this regulatory system for a range of parameters values described in Table M3 (Fig. M5). The parameter ranges were selected to include all the values found in the literature (Table S2-S6). For each set of parameters, we simulated 10 replicates with 1500 iterations each (Fig. M6).

We deleted the first thirty percent of data points from each temporal dynamic to be sure the FF and CV^2 measurements were estimated for the steady-state (Fig. M6). These estimates were averaged over the replicates. We plotted these estimates and compared qualitatively the pattern formed throughout the range of parameters (Fig. M5). We also compared them quantitatively by measuring the difference in their probability distributions with the Jensen–Shannon divergence (Fig. M5).

Parameter	Range
Activation rate (k_{on})	0.01 - 1 in steps of 0.01
Desactivation rate (k_{off})	0.02 - 2 in steps of 0.02
mRNA degradation rate (kd_{mRNA})	0.001 - 0.1 in steps of 0.001
Protein degradation rate (kd_p)	0.001 - 0.1 in steps of 0.001

Table M3. Parameter values ranges.

Additionally, we made a similar analysis to that in section two for one unregulated and one self-activated gene. We did this to get insights into the differences in the temporal dynamics of each system simulated with NGM and CLE1 implementation. For the unregulated gene, we simulated the same three cases of the previous section. For the self-activated gene, we used a set of parameters reported in the literature (Table S9). For each set of parameters, we simulated 10 replicates with 1500 iterations each. Finally, we plotted the temporal dynamic, the steady-state distribution, and the temporal differences distribution.

Section 4: Dynamical properties of the regulatory systems for individual cells

In this section, we accomplished the first approach of the third objective and studied the dynamical properties of the selected regulatory systems for decoupled cells (Fig. M7). This represents the dynamical properties of the system in one cell throughout the time or more than one decoupled cell at a given time. For this, we simulated the three-stage model of one unregulated gene, one self-activated gene, and two genes in an Activator-Inhibitor regulatory system (Fig. M7).

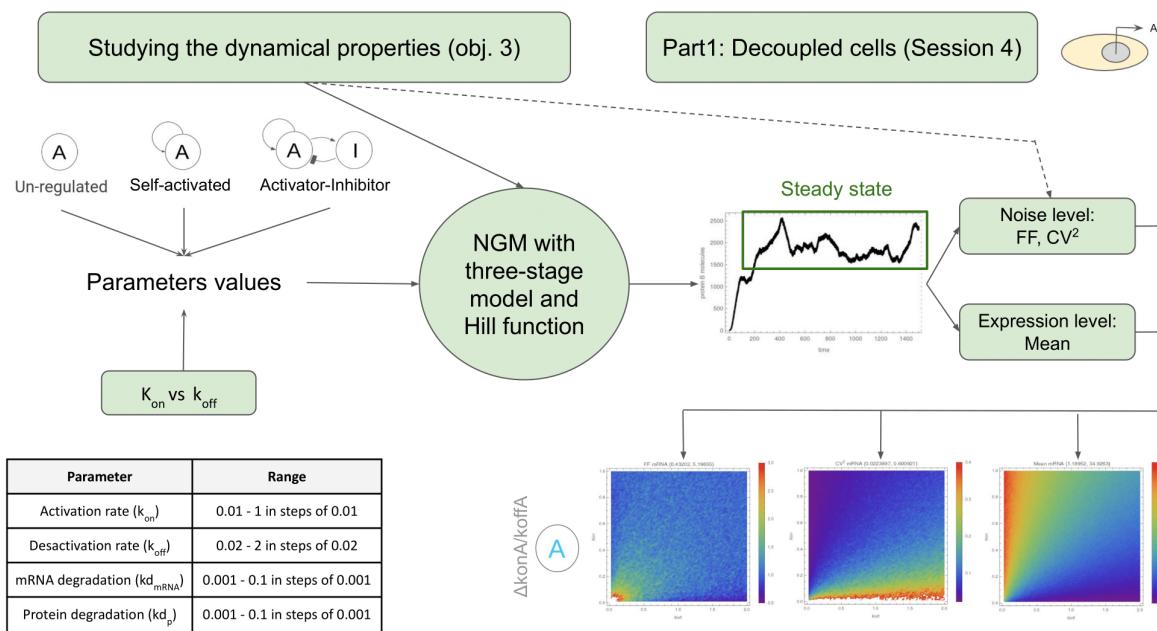


Figure M7. General flux diagram of methodological session 4. It summarizes the steps required for the evaluation of the dynamical properties (i.e. noise and expression level) in three regulatory systems in a decoupled cell. The simulation of regulatory systems uses the three-stage model for gene expression and Hill function for gene regulation.

The parameters evaluated and their range of values are in Figure M7. These parameters were selected based on previous studies [14,24,25]. The parameters values range was selected to include all the values found in experimental literature and previous models (Table S2-S8). However, there is a lack of knowledge about these kinetic parameters and many modelation studies assume their own values [74,78–82]. The default values for all parameters are in Table S9.

The regulatory systems were simulated with NGM and the set of reactions for each one of their genes are similar to that in Figure M2 and M1o. For an unregulated gene, k_{on} means the promoter activation frequency (Figure M2). But k_{on} means the maximal frequency of activation when regulation by a TF is included with a regulation function(θ), and it is modified by the TF (Table M4) (Fig. M3).

The regulation function includes the TF concentration (i.e., number of molecules by cell) and its affinity to promoter sequence (K_{ij} or K_{MM}). K_{MM} defines the TF expression level that is required to half-activate or repress the regulated gene (Fig. M3). The regulation function is included in the propensity function of promoter activation reaction (Table M4). We added a basal expression of 0.01 to propensity functions previously described in Figure M2 (Table M4) (Fig. M3).

regulatory system	Reaction	Propensity function
Auto-activated/ self-activated gene	Promoter activation	$k_{on} * \frac{protein^h}{K_{AA} + protein^h}$
	mRNA synthesis with basal production	$ks_{mRNA} + 0.01$
Activator - Inhibitor / Turing system Activator (A)	Promoter activation	$k_{onA} * \frac{proteinA^h}{K_{AA} + proteinA^h} * \frac{1}{1 + (\frac{proteinI}{K_{IA}})}$,
	mRNA synthesis with basal production	$ksA_{mRNA} + 0.01$
Activator - Inhibitor / Turing system Inhibitor (I)	Promoter activation	$k_{onI} * \frac{proteinA^h}{K_{AI} + proteinA^h}$
	mRNA synthesis with basal production	$ksI_{mRNA} + 0.01$

Table M4. Some propensity functions for reactions of the regulatory systems evaluated. The propensity function of promoter activation reaction includes the Hill regulation function (θ).

For each regulatory system and set of parameters, we simulated the system up to a final time of 1500 minutes (25 hours). It means that for some simulations the NGM is run by more than 400 000 iterations and for others only 5000. We calculated the FF, CV^2 , and mean of molecules in the steady-state (Fig. M7). We deleted the first thirty percent of data points from each temporal dynamic to be sure these measurements were estimated for the steady-state (Fig. M7 and Fig. M3). We averaged these estimates over ten replicates and we plotted their values for each set of parameters (Fig. M7 and Fig. M3).

After this, we selected from these plots four regions of interest because of their differences in FF, CV^2 , and mean of molecules in the steady-state. For each one of these regions and each gene of each regulatory system, we plotted the temporal dynamic of mRNA and protein, their autocorrelation and velocities, and their steady-state distributions.

We repeated the previously described analysis for the activation and inhibition types of regulation (Fig. M8). This was made to compare the differences in noise between regulated

and unregulated gene. We used NGM to simulate the regulatory systems for a set of parameters values (Fig. M8, Table M3). We measured the noise with FF and performed a t-test to estimate the differences in FF values for some parameters values (Fig. M8). We also repeated the previously described analysis with the same regulatory systems and activation/inhibition types of regulation for a constitutive gene. This means that we did not include promoter-level reactions to simulate a constitutive expression with NGM.

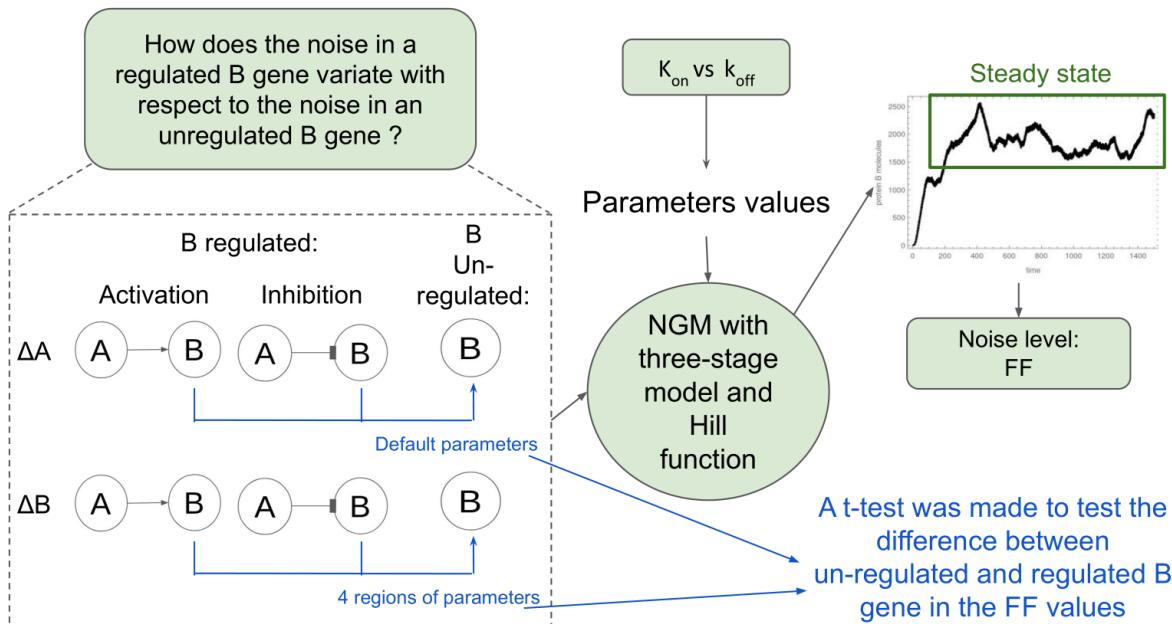


Figure M8. Evaluation of noise propagation from regulator to regulated gene. It summarizes the steps required for the estimation of the differences in noise between an unregulated and regulated B gene. The simulation of regulatory systems uses the three-stage model for gene expression and Hill function for gene regulation.

Section 5: Dynamical properties of the regulatory systems for coupled cells

In this section, we accomplished the second approach of the third objective. We studied the dynamical properties of the self-activated gene in different parameters values for coupled cells (Fig. M9). We simulated a population of isogenic cells in a circular colony similar to micropattern colonies [83] (Fig. M9). Each cell is represented by a self-activation system, the regulatory system in turn is represented by two CLEs, one for mRNA dynamics and another one for protein dynamics (Fig. M10).

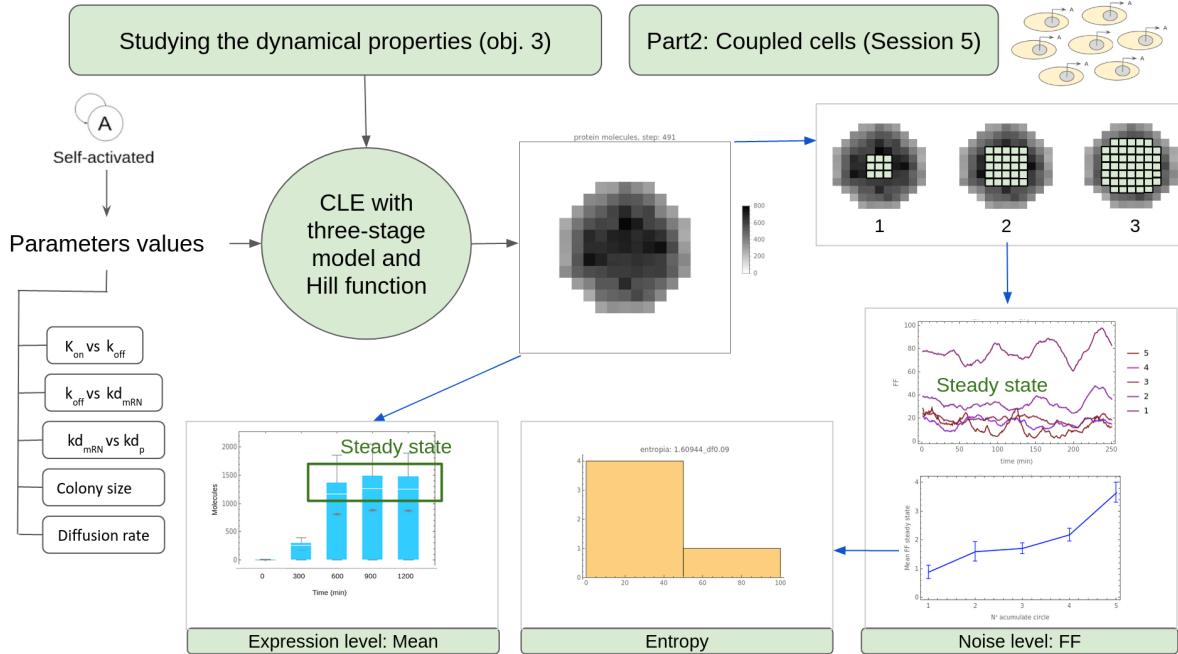


Figure M9. General flux diagram of methodological session 5. It summarizes the steps required for the evaluation of the dynamical properties (i.e. noise and expression level) in a self-activated gene in coupled cells. The simulation with CLE of regulatory system uses the three-stage model for gene expression and Hill function for gene regulation.

Cells are coupled by diffusion of a genetic product (i.e., a paracrine signal) (Fig. M10). The diffusion occurs at rate D toward n number of neighbors. The regulatory function θ is replaced by the first function described in Table M4 (Fig. M10). The parameters evaluated and their range of values are in Table M3 and M5. For this regulatory system and each set of parameters, we simulated 100 replicates up to 1500 minutes (25 hours) each (Fig. M10). Then, we measured the FF for the number of molecules in a set of cells in each time unit in the steady-state. The set of cells were delimited as the cells inside of consecutive circles (Fig. M9). The first circle is at the center of the colony and the last one surrounds all of it. We used the steady-state detection algorithm described in [84] to estimate the beginning of steady-state.

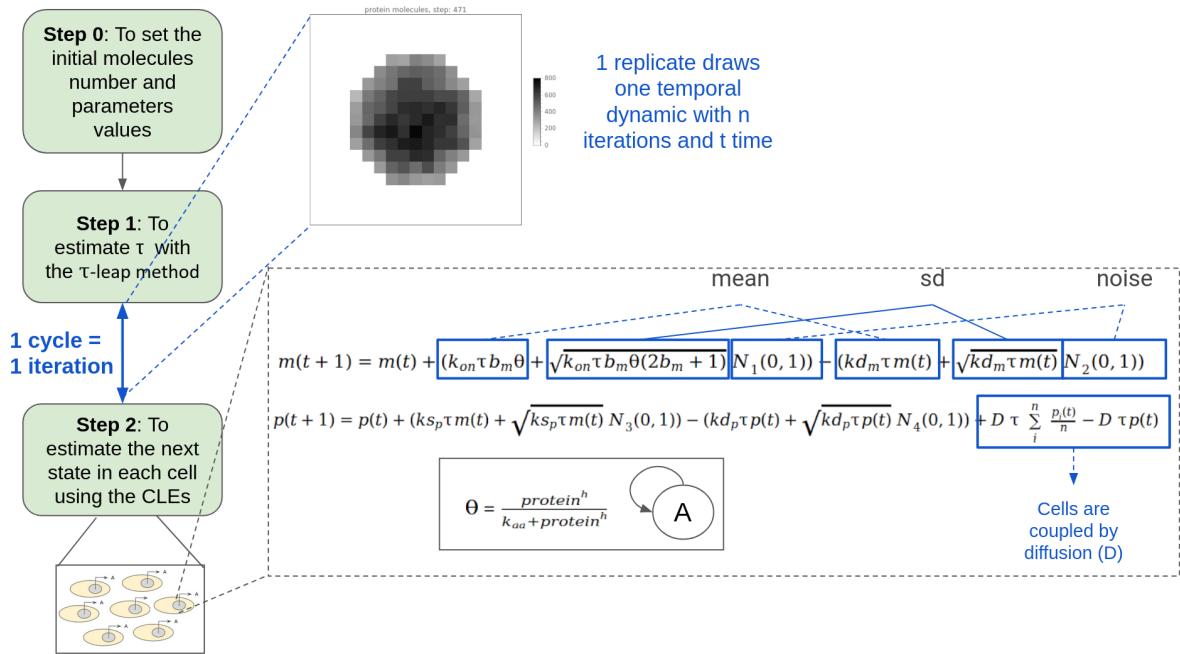


Figure M10. Chemical Langevin Equation for burst in mRNA and not in protein (CLE 1) with diffusion. It summarizes the steps required in the CLE method for the simulation of a self-activated gene in coupled cells. The simulation of regulatory system uses the three-stage model for gene expression and Hill function for gene regulation.*sd:standard deviation.

We plotted the temporal dynamics of FF by each circle (Fig. M9). For each circle, we estimated the mean of FFs in the steady-state and we plotted it with bars of its standard deviation (Fig. M9). Then, we made a histogram with these means of FF for all the consecutive circles in the colony and estimated the Shannon Entropy of these values as

$$H(x_i) = \sum_{i=1}^n P(x_i) \ln P(x_i),$$

where x_i is each one of the FF calculated for each one of the n circles. We also plotted these entropies for all the range of parameters values. For the last circle, we plotted the mean number of molecules in all the range of parameters values (Fig. M9).

Parameter	Range
Diffusion rate (D_a)	0.01 - 0.1 in steps of 0.01
Radius of the colony (colony size)	2 - 25 in steps of 5

Table M5. Parameter values range.

RESULTS AND DISCUSSION

Section 1: The regulatory systems

We selected two regulatory systems because of their importance and recurrent presence in differentiation and patterning formation in animal embryonic development. The nodes of these systems are generally genes encoding paracrine signals or morphogens. They had been previously studied and some of their kinetic parameters are well established for some genes (Table S1-9). We describe them below.

1. Auto-Activated gene (Positive self-activated gene)

In embryonic development, the major regulatory task is to specify spatial domains of gene expression. It means activating a unique set of regulatory genes in a given spatial domain of an embryo [46]. Sometimes the new regulatory state is locked down by deployment of positive feedback of a paracrine signal (Fig. 1A) [46]. For this, the gene encoding the signal responds to its signal transduction system. Thus, all cells in the domain both receive and emit the same signal. As a result, they perform a phenomenon known as “community effect” (Fig. 1B) [46,70]. In which, the cells are linked together by intercellular signaling and by the expression of the same set of downstream genes [46,70].

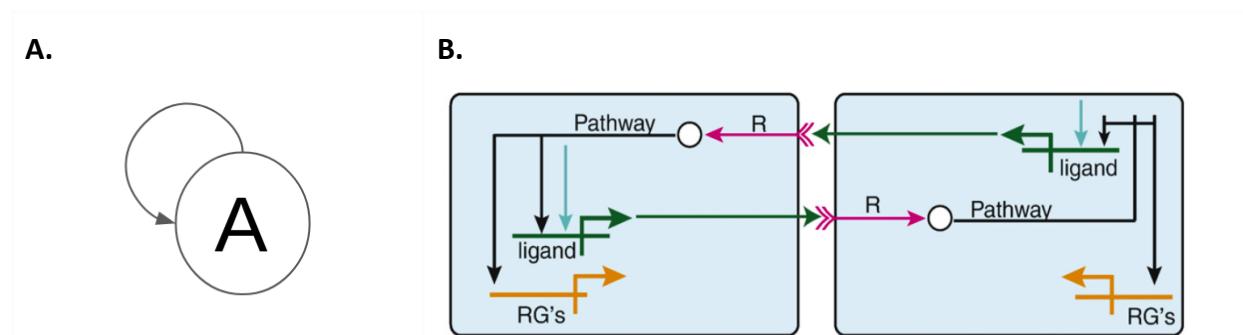


Figure 1. The self-activated gene and the community effect. A. self-activated gene, B. Two cells of the many constituting the field within which the community effect occurs. Each one expresses the signaling ligand gene (green) which activates a signal transduction pathway (black) after it binds to the receptor (double arrow tail). The result is feedback activation of the gene encoding the ligand, which also receives an independent initial input (blue arrow). The same downstream regulatory genes (RG's, orange) are consequently activated in both cells [69].

In the sea urchin (*S. purpuratus*) the community effect has been reported in the specification and delimitation of the oral ectoderm domain [70]. The *NODAL* gene is expressed in the hatching blastula stage (Fig. 2). This is necessary for the expression of downstream oral ectoderm regulatory genes, including *gsc* and *foxG* genes (Fig. 2). Many downstream regulatory genes fail to be expressed when the *NODAL* mRNA translation is blocked [70].

A.

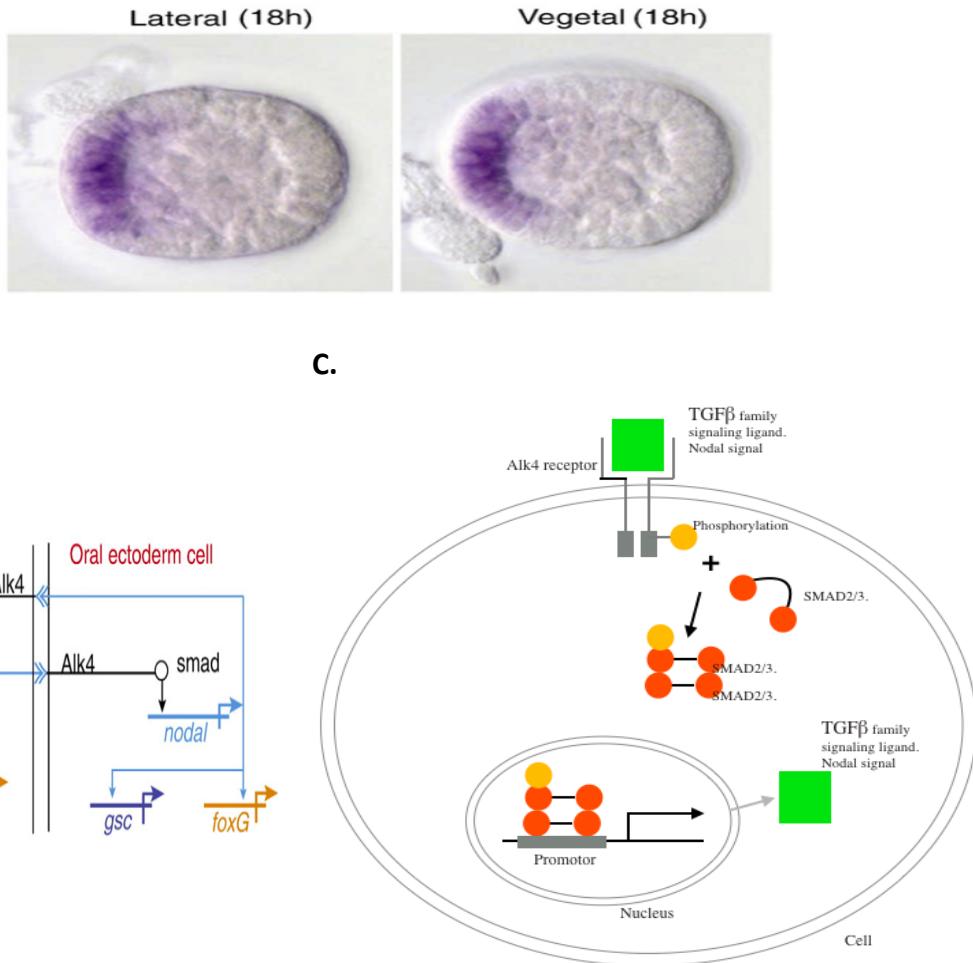


Figure 2. The Nodal community effect in the sea urchin embryo oral ectoderm. A. Lateral and vegetal view of a Nodal in the hatching blastula after 18 hours post fecundation ([85], [70]), B. NODAL self-activated gene and its downstream genes [70], C. The expression of NODAL gene requires the reception of Nodal signal and the nuclear translocation of smad2%.

In the sea urchin endomesoderm, the *wnt8* gene requires for its expression the β -catenin/complex that is generated by the reception of the Wnt signal [86]. In *Drosophila*, the gene encoding the Dpp TGF β ligand has been demonstrated to be self-activated in response to its signal. It happens in several domains of expression as in the midgut visceral mesoderm [87].

In *Xenopus* embryo, Bmp4 is expressed in ventrolateral regions. The immediate early response factor transducing this signal is Oaz. It provides an input into the *vent2b* gene, which provides direct input into the gene encoding Bmp4 ligand in the same domain [88]. eFGF gene is also expressed in the *Xenopus* embryo and its signal transduction system activates the *Xbra* gene. This gene in turn provides a positive input to the gene encoding the eFGF ligand [88].

2. Activator-inhibitor (Turing System)

Another important task in embryo development is self-organized fate patterning [83]. In 1952, Alan Turing put forward the reaction-diffusion model, in which two interacting and diffusing species of molecules can generate complex patterns [89]. Gierer and Meinhardt postulated that pattern formation in reaction-diffusion models requires a short-range activator. It enhances its production and the production of a long-range inhibitor (Fig. 3A)[90]. One example of this is the TGF β ligands Nodal and Lefty. They constitute an Activator-Inhibitor system in animals as different as sea urchin and mouse (Fig. 3B) [91].

During embryogenesis, Nodal signaling is active near its source and is inhibited by Lefty farther away (Fig. 3B). The lower mobility of Nodal allows its accumulation close to the site of secretion, whereas the high mobility of Lefty leads to rapid long-range dispersion. The Nodal/Lefty system fulfills two of the tenets of Activator-Inhibitor reaction-diffusion systems: (i) Nodal is a short-range to a mid-range activator that enhances their expression, and (ii) Lefty is a long-range inhibitor that is activated by NODAL (Fig. 3B)[91].

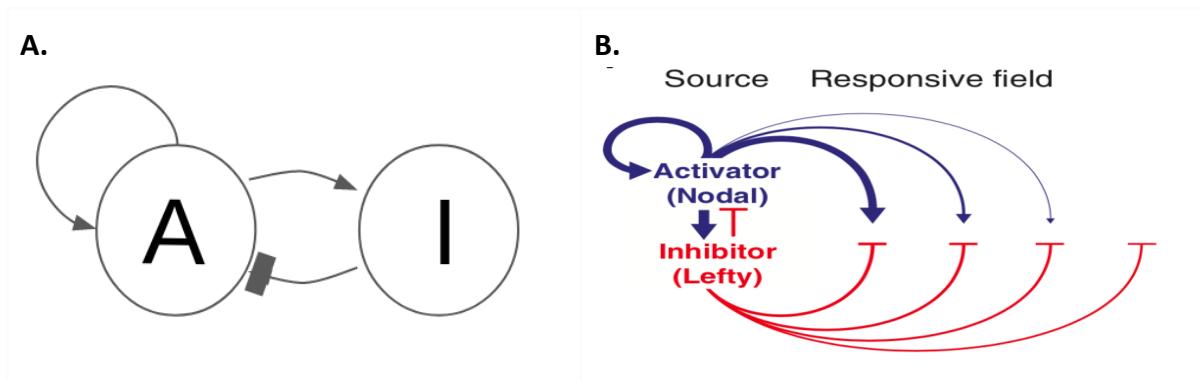


Figure 3. Activator-Inhibitor system. A. Activator (A) - Inhibitor (I), B. Model of the Nodal/Lefty Activator-Inhibitor reaction-diffusion system. In the source, Nodal signal (blue) activates their expression as well as the expression of Lefty (red), which inhibits Nodal production. Nodal signaling in the responsive field is inhibited by the long-range inhibitor Lefty. Both proteins diffuse throughout the responsive field (adapted from [91]).

The Nodal signaling pathway plays a fundamental role during the early development of all vertebrates [92]. For this reason, Nodal expression is tightly regulated in a highly dynamic fashion at discrete tissue sites during early development [92]. Nodal is required for initial specification of the anteroposterior axis in mice, mesoderm formation in mouse and *Xenopus*, left-right patterning in mice, and development of head and trunk mesoderm and endoderm tissues in zebrafish [92].

Section 2: The three-stage model of one gene with NGM and CME

The three-stage model of one unregulated gene with noise has a diversity of temporal dynamics for different parameters values (Fig. 4). The characteristics of these temporal dynamics are well represented by the *steady-state distribution* and the *temporal differences*

distribution (Fig. 4 A1-J1, A3-J3, respectively). For instance, the temporal dynamic is in discontinuous bursts with the highest noise for the lowest promotor velocity (or affinity) (Fig.4 A,E,H).

With low affinities, discontinuous bursts are not only present for the typical case with k_{off} bigger than k_{on} (Fig.4c). They are also produced when k_{off} is lower than k_{on} or when they are equal (Fig.4 A2 and E2) [76]. In nature, low affinities are commonly reported for a variety of genes (Table S2). One example is the promoter rates of *ush* gen with k_{on} of 0.25 and k_{off} of 0.04 [36]. They were measured in the patterning of ectodermal cells along the dorsal-ventral axis in *Drosophila* embryos [36].

Also with low affinities, temporal dynamic presents the highest values of FF and CV^2 calculated for the steady-state distribution (Fig.S1). This distribution is bimodal with a peak around zero and another one in a non-zero value. However, both FF and CV^2 decrease with the increase in k_{on} for the three cases with low affinities (Fig.S1 A,E,H). This is because the mean number of molecules increases with the rise of k_{on} (Table 1 A,E,H).

The steady-state distribution could be interpreted in two ways. First, different states in which a gene is expressed in one cell throughout time. Second, different states in which cells in a tissue are expressing a gene at a specific time. On the other hand, the distribution of the temporal difference also reflects the burst discontinuity in low affinities. It has FF, CV^2 , mean, and maximum value higher than those cases of higher affinities (Fig.4 A3,E3,H3).

The expression of a gene regulated at promoter level becomes similar to that of a constitutive gene when k_{on} higher than k_{off} and with medium to high affinities (i.e., case 1)(Fig.4 B1-C3, G1-3, J1-3; Fig.S1 B-C, G, J; Fig.S2B). The most important condition is to have k_{on} higher than k_{off} because it increases the promoter occupancy (Table 1). Even when cases 1 and 3 have the same burst size and frequency, they have different expression dynamics because of their differences in promoter occupancy (Table 1).

The similitude with a constitutive expression is characterized by two aspects. First one, the expression has a higher mean of mRNA molecules (Table 1). Second, the expression is more continuous (Fig. 4aC). This similitude is higher when ks_{mRNA} increases, which throws FF equals 1 and CV^2 below 1 (Fig.4aD; Fig.S1D; Fig.S2). In this way, it is not possible to distinguish between the expression of a constitutive gene and a promoter-regulated gene [20,21,25].

Case 2 illustrates well the general idea pointed out so far for all the cases. That is, the increase in the promoter affinity increases the burst frequency but decreases its size (Table 1). It means discontinuous bursting is associated with lower frequencies and bigger burst sizes. A more continuous expression is reached not only because the affinity is higher, but because k_{on} is higher than k_{off} .

Case 3 presents a discontinuous burst in the whole range of promoter affinities because k_{on} is lower than k_{off} (Fig.4c and Table 1). In comparison with the other cases, the promoter occupancy and mean of molecules are the lowest and CV^2 is the highest. But FF could be similar to the other cases (Table 1).

	$k_{on*} k_{off*}$	$\epsilon = 0.01$	$\epsilon = 1$	$\epsilon = 100$	$\epsilon = 1000$
Case 1		A	B	C	D*
$k_{on} k_{off}$		0.1 0.01	10 1	1000 100	1000 100
P.O.	10 1	0.909	0.909	0.909	0.909
B.S.		90.9	0.909	0.009	0.09
B.F.		0.009	0.909	90.9	90.9
M.M.		9.09	9.09	9.09	90.90
Case 2		E	F	G	
$k_{on} k_{off}$		0.01 0.01	1 1	100 100	
P.O.	1 1	0.5	0.5	0.5	
B.S.		500	5	0.05	
B.F.		0.005	0.5	50	
M.M.		5	5	5	
Case 3		H	I	J	
$k_{on} k_{off}$		0.01 0.1	1 10	100 1000	
P.O.	1 10	0.0909	0.0909	0.0909	
B.S.		90.9	0.9	0.01	
B.F.		0.009	0.909	90.9	
M.M		0.90	0.90	0.90	

Table 1. Estimation of some kinetic parameters of transcriptional bursting. For different values of k_{on} and k_{off} it is estimated the Promoter Occupancy (P.O), Burst Size (B.S), Burst Frequency (B.F.), and mean of mRNA molecules (M.M.). For all the cases ks_{mRNA} is 10, kd_{mRNA} is 1, ks_p is 1, and kd_p is 1, and k_{on}/k_{off} is equal to $k_{on*} * \epsilon / k_{off*} * \epsilon$. D*. Here the ks_{mRNA} is 100.

The *ush* and *plectin1* genes are examples of cases 1 and 3, respectively. The *plectin1* gene has higher noise than *ush* gene (Fig. S3). This is intriguing because of the function of the *plectin1* gene. This gene is implied in maintaining cell and tissue integrity, and it is a

scaffolding platform for the assembly, positioning, and regulation of signaling complexes (Table S2). This function is expected to be accomplished with low noise [32].

Two additional factors must be analyzed to understand the noise in gene expression. The first one is the effect of mRNA degradation. This is important because mRNA levels in the embryo are determined not only by the transcriptional rates but also by degradation rates [93]. The second one is to what extent the protein distributions follow the mRNA distributions. It has been found experimentally that when the protein decays faster than the mRNA, the two distributions are similar [14]. But when the protein lifetime is longer than the mRNA, the distributions could be different [14].

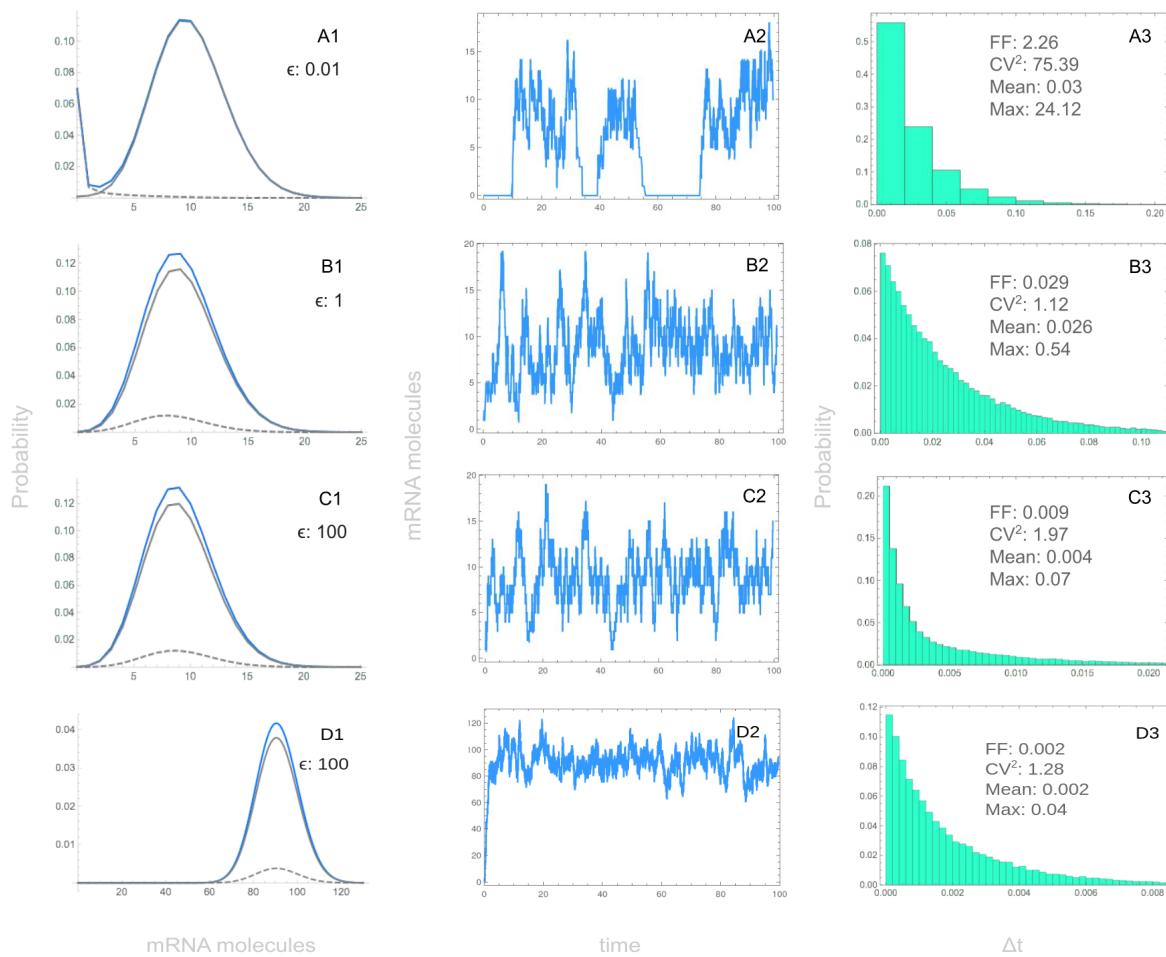


Figure 4a. Temporal expression characteristics of one gene with noise simulated with NGM. 1. Steady-state distribution of mRNA expression (blue line) estimated with the CME. It also shows the distribution for the active (solid grey line) and inactive (dashed grey line) promoter state. 2. Temporal expression of the system. 3. Temporal differences distribution and its FF, CV^2 , mean, and maximum value (Max). For case 1, the parameters evaluated are A1-C3. $k_{on^*} : 10$, $k_{off^*} : 1$, $k_{sRNA} : 10$, $k_{dRNA} : 1$, D1-D3. $k_{on^*} : 10$, $k_{off^*} : 1$, $k_{sRNA} : 100$, $k_{dRNA} : 1$. The promoter velocity value is indicated (ϵ : 0.01, 1, 100 which are low, medium, and high, respectively).

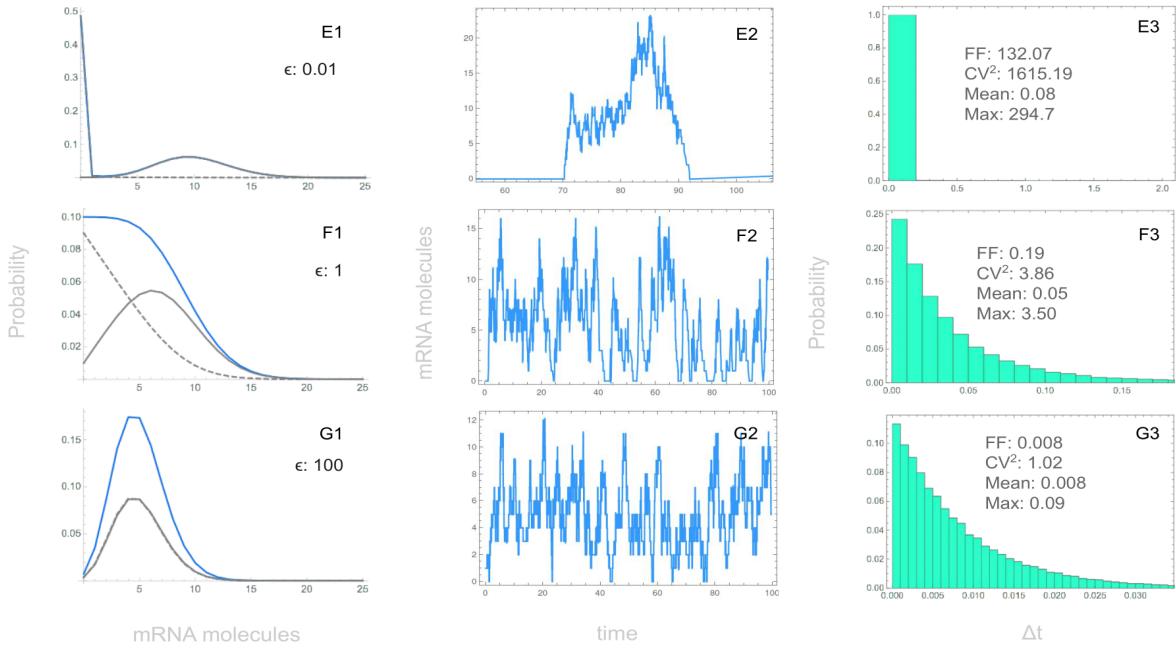


Figure 4b. Temporal expression characteristics of one gene with noise simulated with NGM. 1. Steady-state distribution of mRNA expression (blue line) estimated with the CME. It also shows the distribution for the active (solid grey line) and inactive (dashed grey line) promoter state. 2. Temporal expression of the system. 3. Temporal differences distribution and its FF, CV^2 , mean, and maximum value (Max). For case 2, the parameters evaluated are E1-G3. $k_{on^*} : 1$, $k_{off^*} : 1$, $ks_{mRNA} : 10$, $kd_{mRNA} : 1$. The promoter velocity value is indicated ($\epsilon: 0.01, 1, 100$ which are low, medium, and high, respectively).

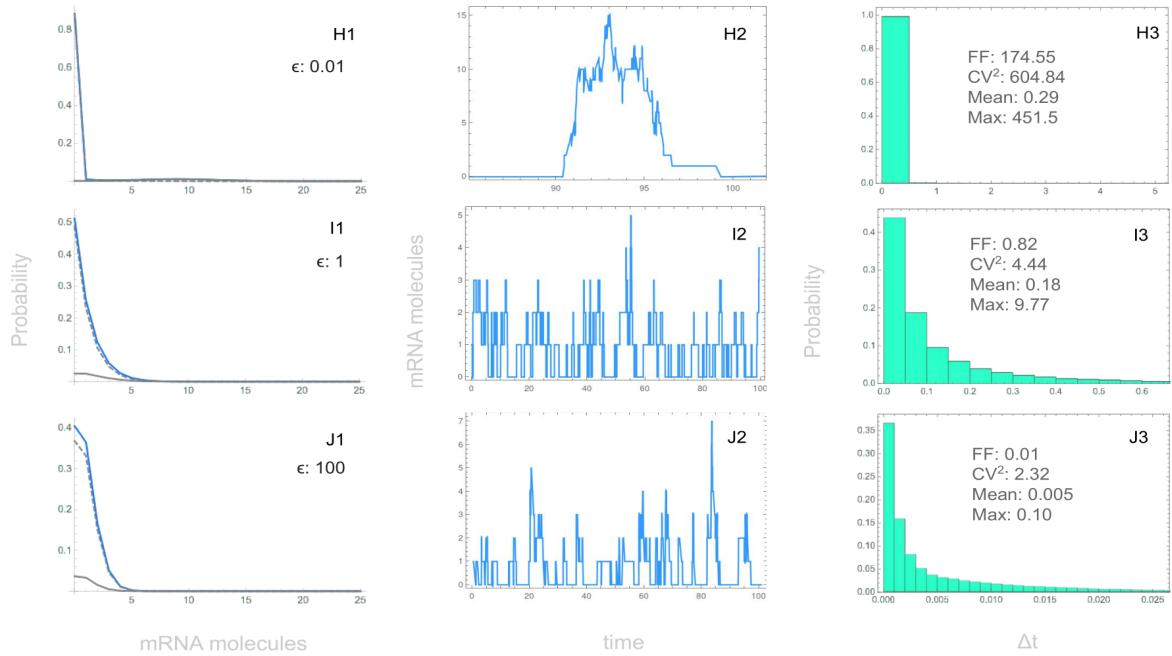


Figure 4c. Temporal expression characteristics of one gene with noise simulated with NGM. 1. Steady-state distribution of mRNA expression (blue line) estimated with the CME. It also shows the distribution for the active (solid grey line) and inactive (dashed grey line) promoter state. 2. Temporal expression of the system. 3. Temporal differences distribution and its FF, CV^2 , mean, and maximum value (Max). For case 3, the parameters evaluated are H1-J3. $k_{on^*} : 1$, $k_{off^*} : 10$, $ks_{mRNA} : 10$, $kd_{mRNA} : 1$. The promoter velocity value is indicated ($\epsilon: 0.01, 1, 100$ which are low, medium, and high, respectively).

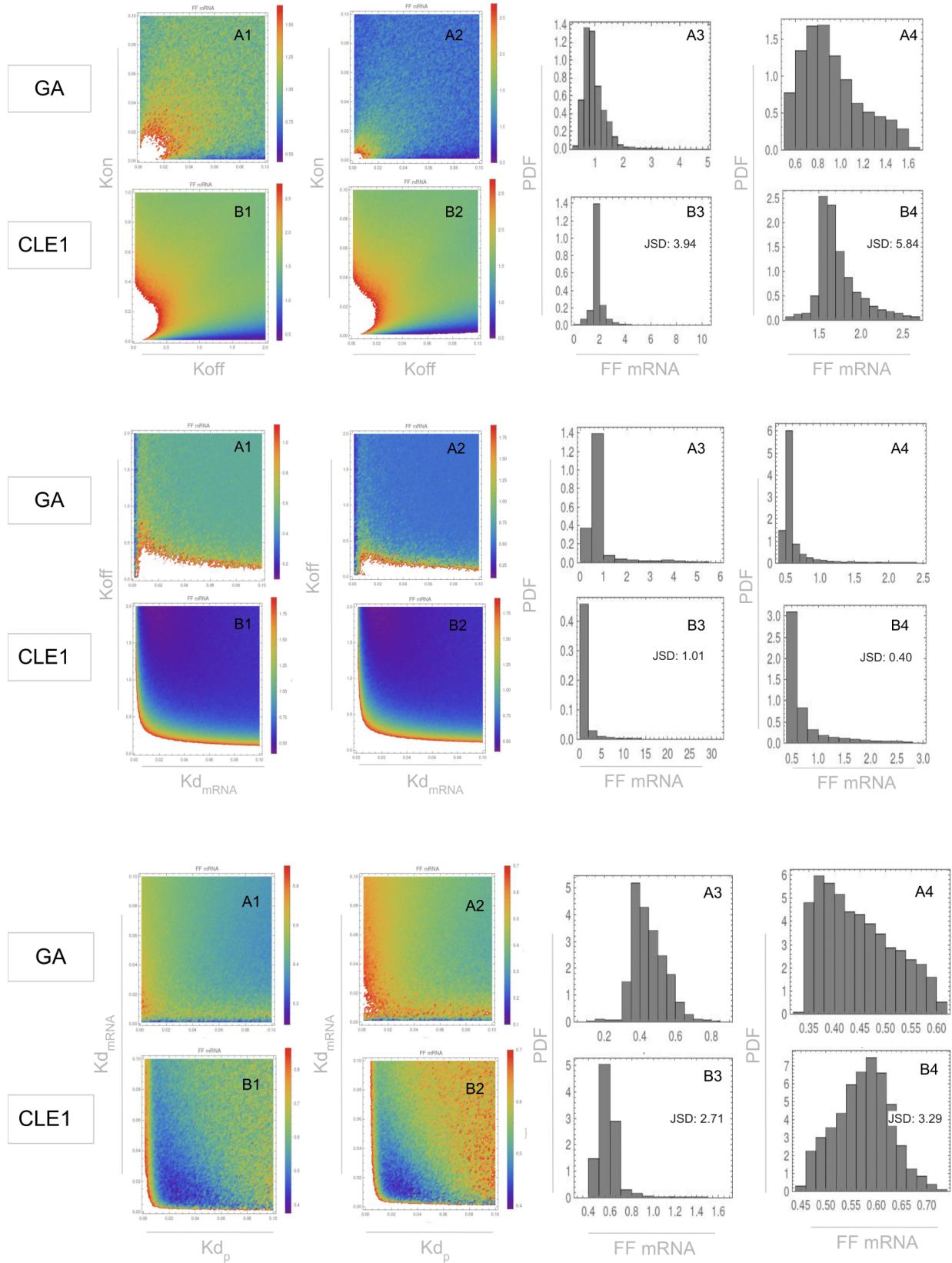
Section 3: The CLE implementation for the three-stage model is a good approach for large systems

Figures 5 and S4 show respectively the FF and CV^2 for mRNA with both NGM and CLE1 and or three different parameters. The similitude between NGM y CLE1 is higher in the distribution of CV^2 values (max-min JSD: 1.24-5.35) than in the distribution of FF values (max-min JSD: 0.4-5.84). But between NGM and CLE1, there is more similarity in the qualitative pattern of FF values than in the qualitative pattern of CV^2 values. Therefore, a high similitude in the qualitative pattern through the parameter values does not mean a high similitude in the JSD of the distribution of the values.

Figures 6 and S5 show respectively the FF and CV^2 for protein expression throughout the range of k_{on} and k_{off} parameters. They were simulated with NGM, CLE1, and CLE3. For the CLE2 case, it was not possible to get results because the simulations generated indeterminate values. For both FF and CV^2 , the highest similitude is between NGM and CLE1 in the qualitative pattern although it is low in the quantitative pattern (Fig.6B, Fig.S5B). CLE3 is different from NGM in the pattern of FF values more than in the pattern of CV^2 values (Fig.6C, Fig.S5C).

Figures 7 and S6 respectively show the FF and CV^2 for protein expression throughout the range of k_{off} and kd_{mRNA} values. They were simulated with NGM and CLEs. There is a high similarity between NGM and CLE3 in the distribution of the CV^2 values (7.15 JSD). But none of the CLEs is similar to NGM in the qualitative pattern of CV^2 values. The similarity between NGM and CLEs in the FF values distribution is lower than for the distribution of the CV^2 values. But there is a higher similarity between NGM and CLEs in the qualitative pattern of FF values. With CLE1 having the highest similarity.

Figures 8 and S7 show respectively the FF and CV^2 for protein expression throughout the range of kd_{mRNA} and kd_p values for NGM and CLEs. There is a high similarity between NGM and CLE3 in the distribution of the CV^2 values (6.99 JSD). But none of the CLEs is similar to NGM in the qualitative pattern of CV^2 values. The similarity between NGM and CLEs in the FF values distribution is lower than for the distribution of the CV^2 values. But there is a higher similarity between NGM and CLEs in the qualitative pattern of FF values. With CLE1 having the highest similarity.



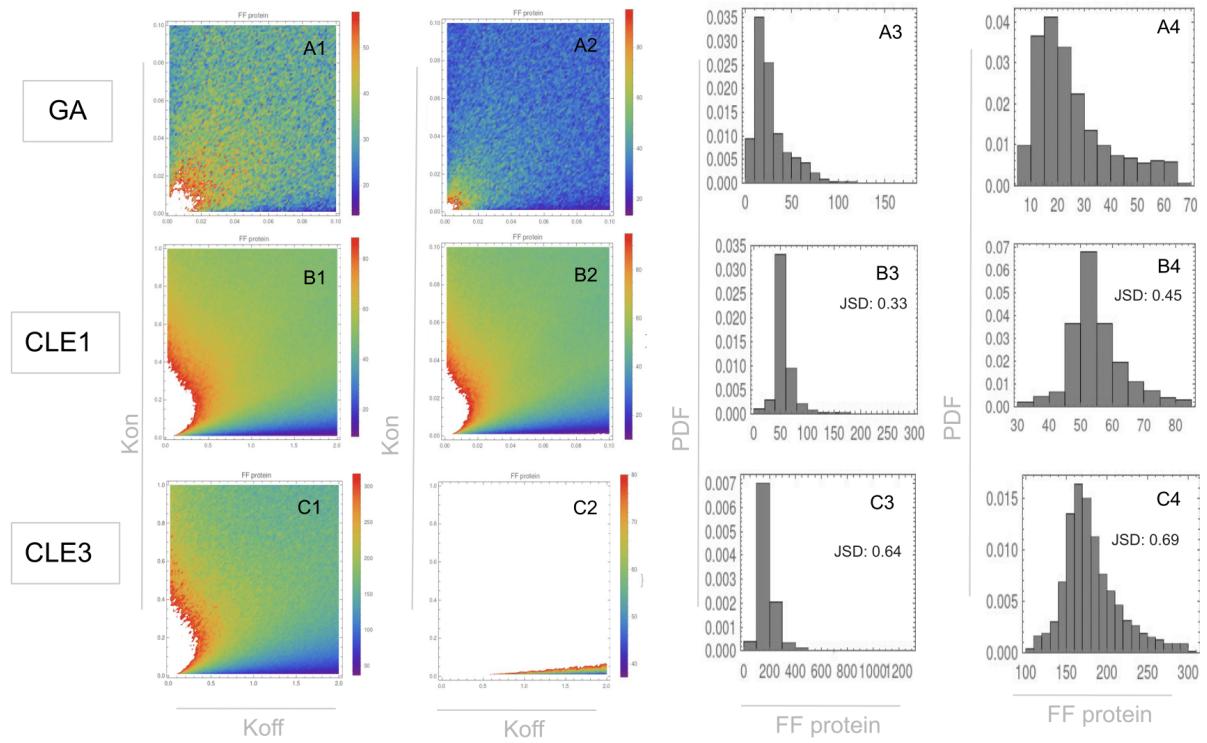


Figure 6. Noise (FF) in protein expression of the self-activated gene for a range of k_{on} and k_{off} values. A. Simulated with NGM, B. Simulated with the CLE for burst only in mRNA (CLE1), C. Simulated with the CLE for burst in both mRNA and protein (CLE3), 1. FF for different parameters values in their characteristic scale (white patches are values outside of the plot range), 2. FF for different parameters values in the same scale (white patches are values outside of the plot range), 3. Probability Density Function (PDF) for all the the values of FFs, 4. The same PDF as in 3 but for the interquartile range 0.05 to 0.95 of FFs values. JSD: Jensen–Shannon divergence between B and A PDFs.

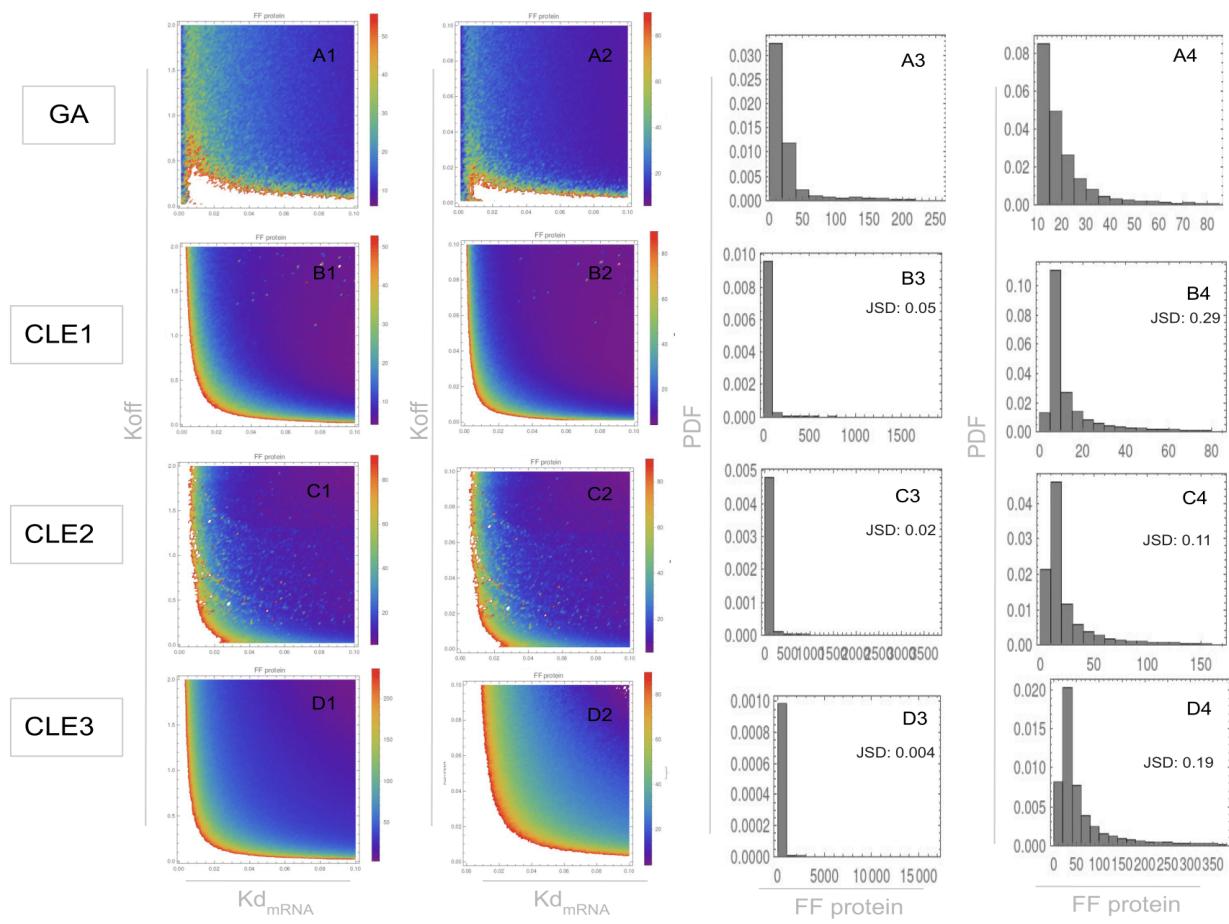


Figure 7. Noise (FF) in protein expression of self-activated gene for a range of k_{off} and kd_{mRNA} values. A. Simulated with NGM, B. Simulated with the CLE for burst only in mRNA (CLE1), C. Simulated with the CLE for burst only in protein (CLE2), D. Simulated with the CLE for burst in both mRNA and protein (CLE3), 1. FF for different parameters values in their characteristic scale (white patches are values outside of the plot range), 2. FF for different parameters values in the same scale (white patches are values outside of the plot range), 3. Probability Density Function (PDF) for all the values of FFs, 4. The same PDF as in 3 but for the interquartile range 0.05 to 0.95 of FFs values. JSD: Jensen–Shannon divergence between B and A PDFs.

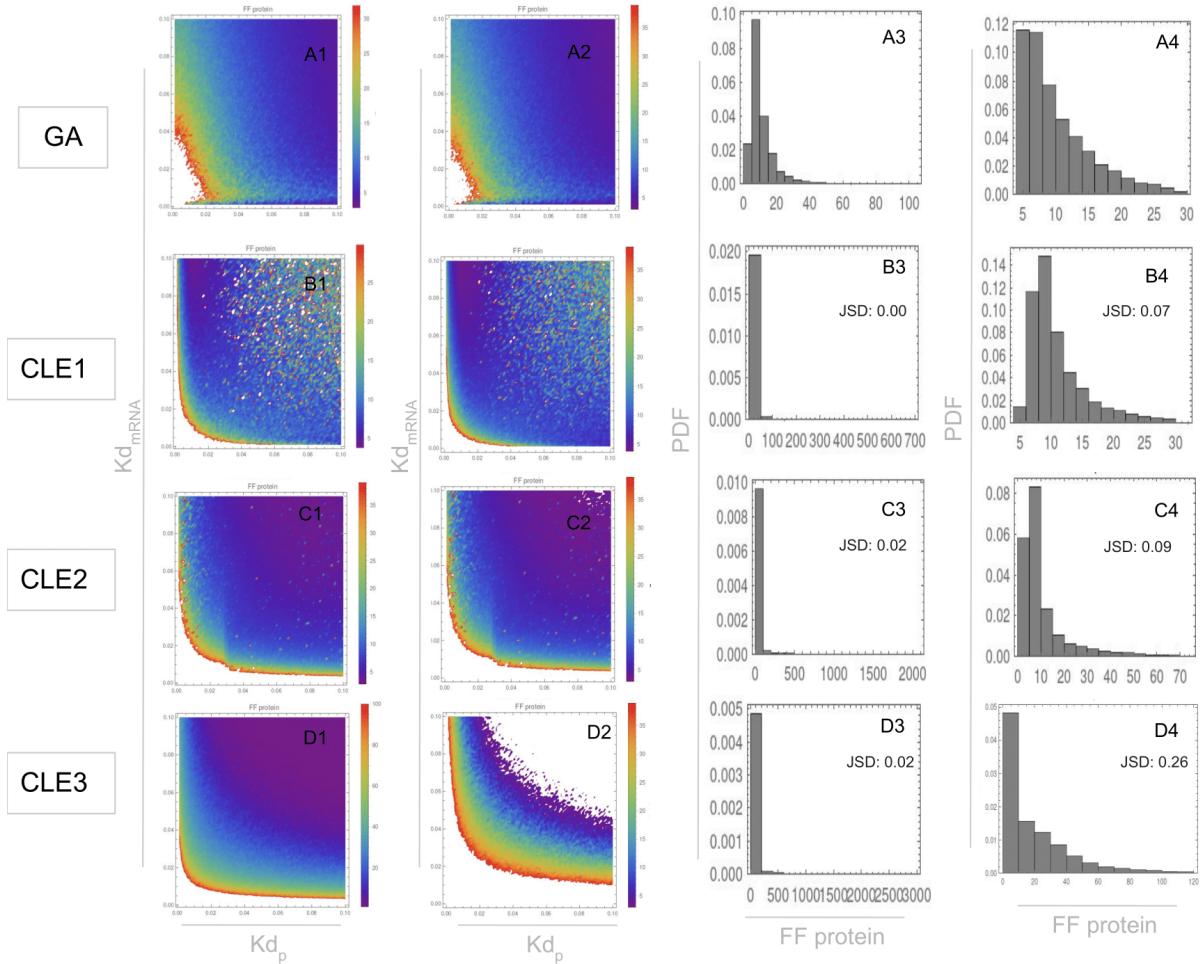


Figure 8. Noise (FF) in protein expression of self-activated gene for a range of kd_{mRNA} and kd_p values. A. Simulated with NGM, B. Simulated with the CLE for burst only in mRNA (CLE1), C. Simulated with the CLE for burst only in protein (CLE2), D. Simulated with the CLE for burst in both mRNA and protein (CLE3), 1. FF for different parameters values in their characteristic scale (white patches are values outside of the plot range), 2. FF for different parameters values in the same scale (white patches are values outside of the plot range), 3. Probability Density Function (PDF) for all the values of FFs, 4. The same PDF as in 3 but for the interquartile range 0.05 to 0.95 of FFs values. JSD: Jensen–Shannon divergence between B and A PDFs.

It is possible to point out several things from this comparison between NGM and CLEs. First, NGM and CLEs are more similar in the estimates of FF than in the estimates of CV^2 . This similarity occurs mainly in the qualitative pattern throughout the parameters' values. Second, CLE1 is the CLE with the highest similarity to NGM and it happens mainly for the pattern of FF values. This occurs because the default parameters values of the system (Table S9) fulfill the parameter requirements of the CLE1 implementation. Finally, we will estimate the noise by mean of FF and we will use the CLE1 to simulate isogenic cells in a circular colony.

Figure 9 shows the temporal dynamic of the three-stage model of one unregulated gene using the CLE1 implementation. Regarding steady-state distribution of mRNA, the FF increases for lower affinities in each case (Fig. 9A1- I1). This was expected from the same analysis with NGM in the previous section. The CV^2 estimates are always under one which is different from the previous analysis with NGM (Fig. 9A1-I1).

For the typical case (i.e., case 3 where $k_{on} \ll k_{off}$) and in all affinities, the mean values of mRNA molecules are higher when the CLE1 is used than the previously observed with NGM. But for the other cases, they are higher only for low affinities. It happens because CLE indicates the mean of the expression at each time point [77]. On the contrary, NGM indicates one particular expression state in each time point [77]. As a result, the burst frequency and size is higher with CLE than NGM (Fig. 4 and 9).

The temporal differences distribution is different between CLE and NGM (Fig. 4 A3-J3, Fig. 9 A3-I3). With CLE is not possible to identify discontinuous bursts as in NGM (Fig. 4 A2,E2,G2; Fig. 9 A2,E2,H2). This is because of the previously described differences between CLE and NGM. And additionally because with the τ -leap method, the time step does not mean the approximate time in which some reaction will occur [77]. But it means a temporal window in which many reactions can occur without changing considerably the number of the molecules of the species reacting [77]. For this reason, the temporal difference distribution will not be used in posterior sections.

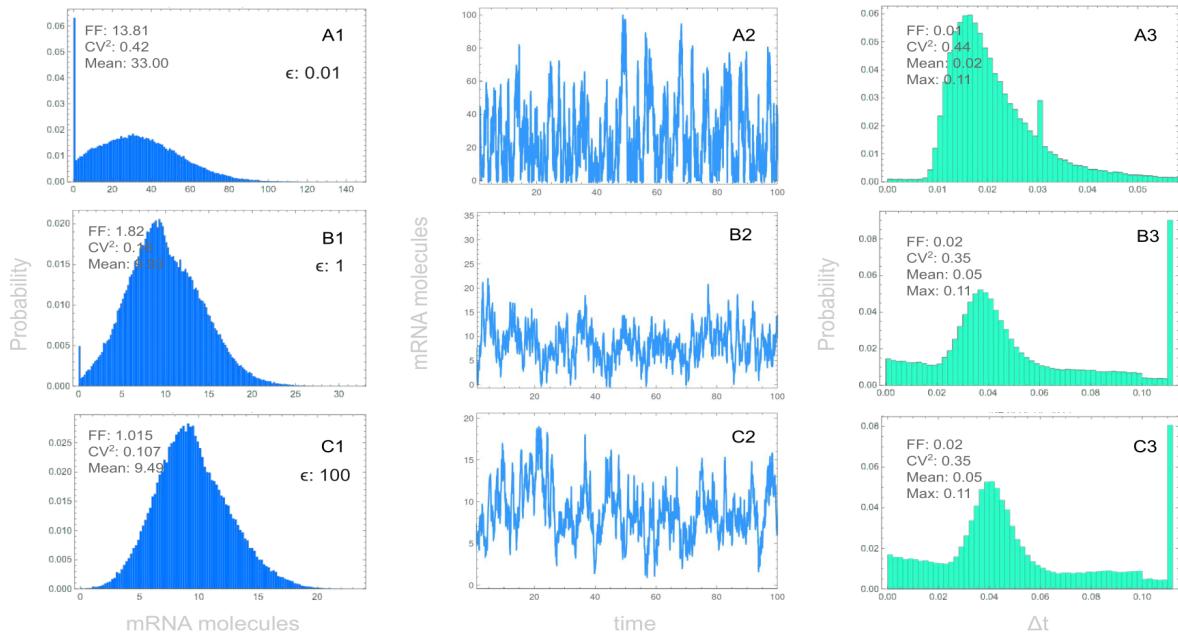


Figure 9a. Temporal expression characteristics of one gene with noise simulated with CLE1. 1. Steady-state distribution of mRNA expression with its FF, CV^2 , and mean, 2. Temporal expression of the system, 3. Temporal differences distribution and its FF, CV^2 , mean, and maximum value (Max). For case 1, the parameters evaluated are A1-C3. $k_{on*} : 10$, $k_{off*} : 1$, $k_{s_{mRNA}} : 10$, $k_{d_{mRNA}} : 1$. The promoter velocity value is indicated ($\epsilon: 0.01, 1, 100$ which are low, medium, and high, respectively).

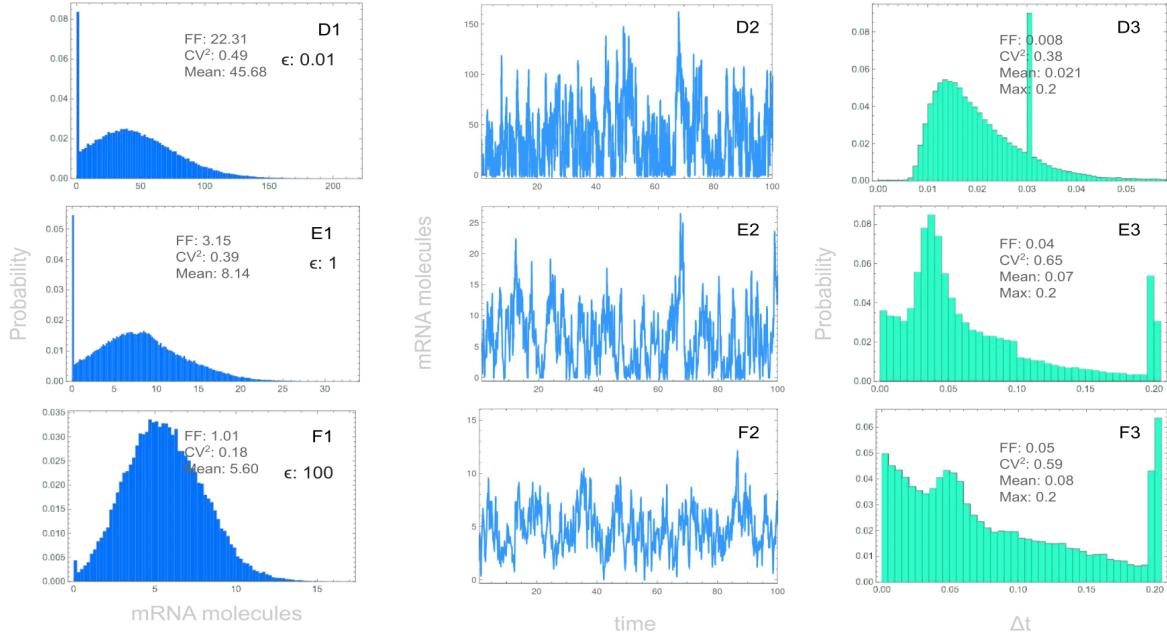


Figure 9b. Temporal expression characteristics of one gene with noise simulated with CLE1. 1. Steady-state distribution of mRNA expression with its FF, CV^2 , and mean, 2. Temporal expression of the system, 3. Temporal differences distribution and its FF, CV^2 , mean, and maximum value (Max). For case 2, the parameters evaluated are D1-F3. $k_{\text{on}*} : 1$, $k_{\text{off}*} : 1$, $k_{\text{s}_{\text{mRNA}}} : 10$, $k_{\text{d}_{\text{mRNA}}} : 1$. The promoter velocity value is indicated (ϵ : 0.01, 1, 100 which are low, medium, and high, respectively).

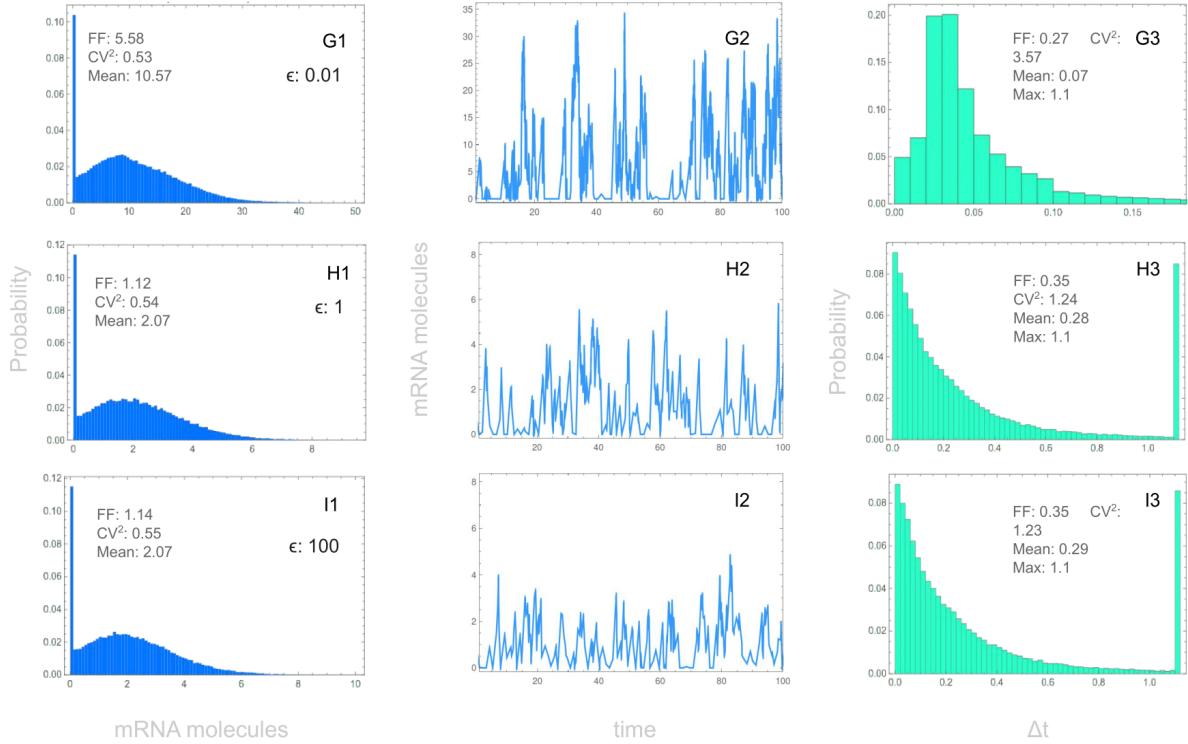


Figure 9c. Temporal expression characteristics of one gene with noise simulated with CLE1. 1. Steady-state distribution of mRNA expression with its FF, CV^2 , and mean, 2. Temporal expression of the system, 3. Temporal differences distribution and its FF, CV^2 , mean, and maximum value (Max). For case 3, the parameters evaluated are G1-I3. $k_{\text{on}*} : 1$, $k_{\text{off}*} : 10$, $k_{\text{s}_{\text{mRNA}}} : 10$, $k_{\text{d}_{\text{mRNA}}} : 1$. The promoter velocity value is indicated (ϵ : 0.01, 1, 100 which are low, medium, and high, respectively).

Figure 10 shows the temporal dynamics of one self-activated gene using CLE1 and NGM. Besides the differences between both methodological approaches, for both of them, the FF

is higher at protein level than at mRNA level (Fig. 10 A2,B2 and A4,B4). The first one has over-dispersion (i.e., $FF > 1$) and the second one has under-dispersion (i.e., $FF < 1$). This is unexpected because according to kinetic parameters this system should have burst at mRNA level but not at protein level (i.e., $k_{off} > (kd_{mRNA}, k_{on})$, and $kd_{mRNA} = kd_p$).

FF estimated with CLE and NGM has a similar qualitative pattern through the parameters evaluated here. This similarity is higher at the protein level than at the mRNA level. Conversely, the CV^2 estimated with CLE and NGM does not have a similar qualitative pattern. Thus, it is not a good estimator of noise when the CLE implementation is used. On the other hand, the differences between CLE and NGM depend on the parameters' values[12]. For this reason, we limited the next analysis to the same set and range of parameters evaluated here (Table M3). Finally, because the main advantage to use CLE is the increase of efficiency, we did not use it in the next section but in section 5 to simulate the colony of cells.

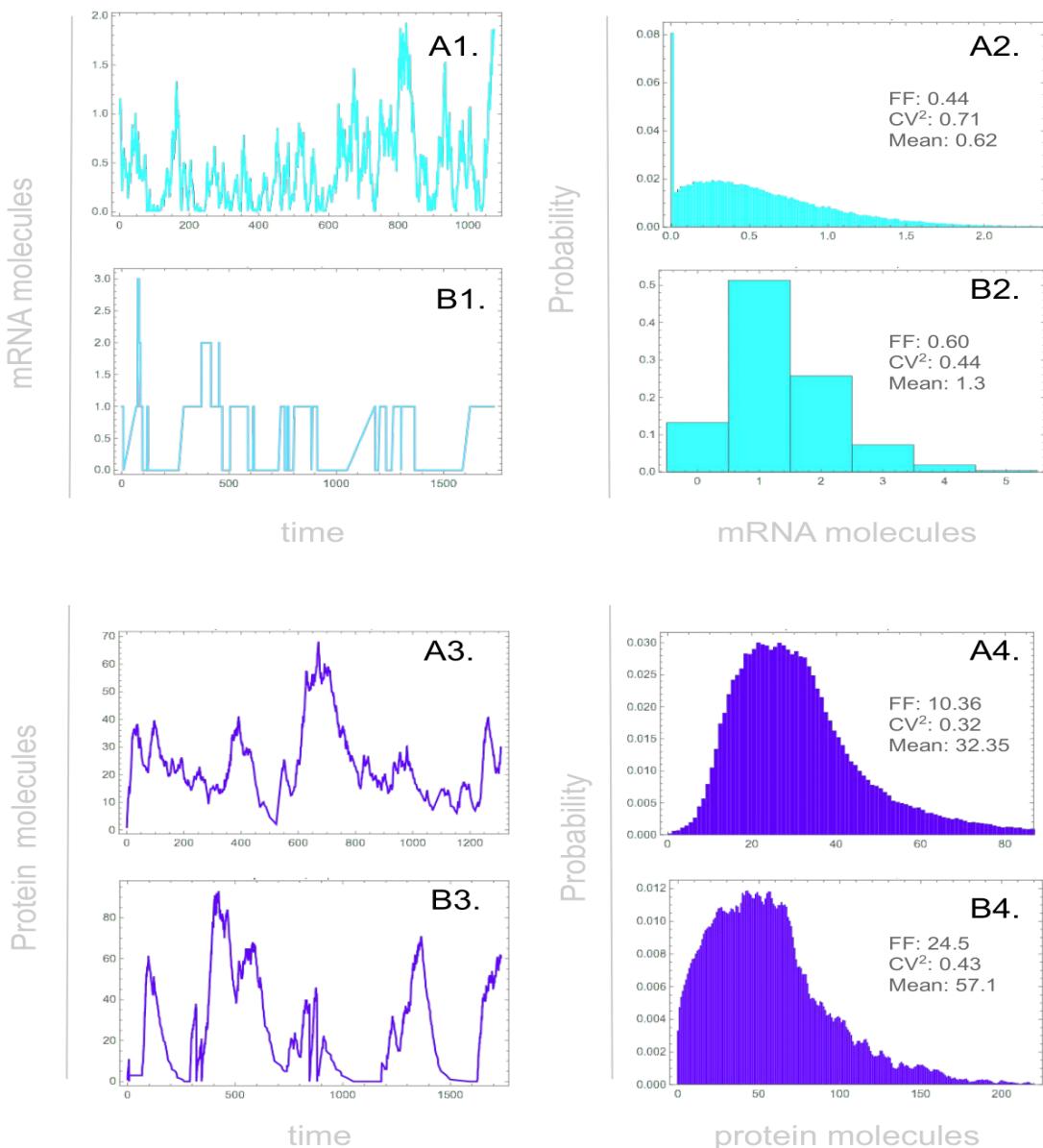


Figure 10. Temporal expression characteristics of the self-activated gene with noise. A. Simulated with CLE1, B. Simulated

with NGM, A1-B1. Temporal expression of mRNA, A2-B2. Steady-state distribution of mRNA expression, A3-B3. Temporal expression of protein, A4-B4. Steady-state distribution of protein expression. The parameters values are k_{on} : 0.01158, k_{off} : 2.082, $k_{s_{mRNA}}$: 0.696, $k_{d_{mRNA}}$: 0.0282, k_{s_p} : 1.386, k_{d_p} : 0.0282 as in [73]. We assumed a self-activation constant of 1 and Hill coefficient of 3.

Section 4: Dynamical properties of the regulatory systems for individual cells

Figures 11-13 and S8-S10 show FF, CV^2 , and mean of mRNA and protein molecules for the regulatory systems selected. The behavior of noise measured with FF at the mRNA level is qualitatively similar to that at the protein level (Fig. 11 and S8). For both of them, the unregulated gene has similar behavior to that of the self-activated gene (Fig. 11A-B and S8A-B). Therefore, the self-activation regulatory system does not reduce the noise. This goes in a contrary way to what has been postulated about self-activation as an amplifier of noise [29,94,95].

In both unregulated and self-activated genes, the noise (FF) is higher when k_{off} is bigger than k_{on} , k_{on} is bigger than k_{off} , and k_{off} is equal to k_{on} . But this happen for Slow Promoter Kinetic (SPK) as is shown in the lower left corner in Figure 11 A-C. From previous studies, it is well known that SPK owing to chromatin remodeling has an important role generating stochasticity in eukaryotic gene expression [24,58,65,72]. For instance, it has been reported that genes with characteristics associated with transcriptional bursting are those with a low value of k_{on} and a relatively high value of k_{off} [24]. Normally, with k_{on} one order of magnitude lower than k_{off} [25]. At the same time, for some genes, k_{on} is higher than k_{off} [36].

It is important to note that the noise measured with FF has a different behavior from the noise measured with CV^2 (Fig. 11-12 and S8-S9). The difference between FF and CV^2 is because CV^2 is more sensitive than FF to the number of molecules [20]. In fact, Figures 12-13 and S8-S10 show an inverse relationship between CV^2 and Mean values. But there is not a clear relationship between FF and Mean values. On the other hand, although CV^2 is consider as the most direct and unambiguous measure of noise, FF reveals trends that can not be observed with CV^2 when the molecules number is great [65]. CV^2 measured the noise and FF the noise strength [65].

This study extend the previous analysis about the relationship between mean number of molecules, noise (CV^2), and noise strength (FF). For instance, it is known that increasing gene expression through high k_{on} rates decreases noise and noise strength as shown in Figures 11-13 and S8-10 [96]. It happens because high k_{on} rates increases the mean of molecules and directs a more continuous expression as discussed below and in [36]. On the other hand, increasing gene expression through lengthening of burst duration (i.e., decreasing k_{off} rate) is associated with more noise strength as shown in Figure 11 [71,72]. But it happens only for some k_{on} values (Fig. 11 and S8). Finally, decreasing gene expression due to increase in k_{off} increases CV^2 for the biggest mean values and this is followed by a constant CV^2 value [65] (Fig. 12-13 and S9-S10). But again, it happens only for some k_{on} values (Fig. 12 and S9).

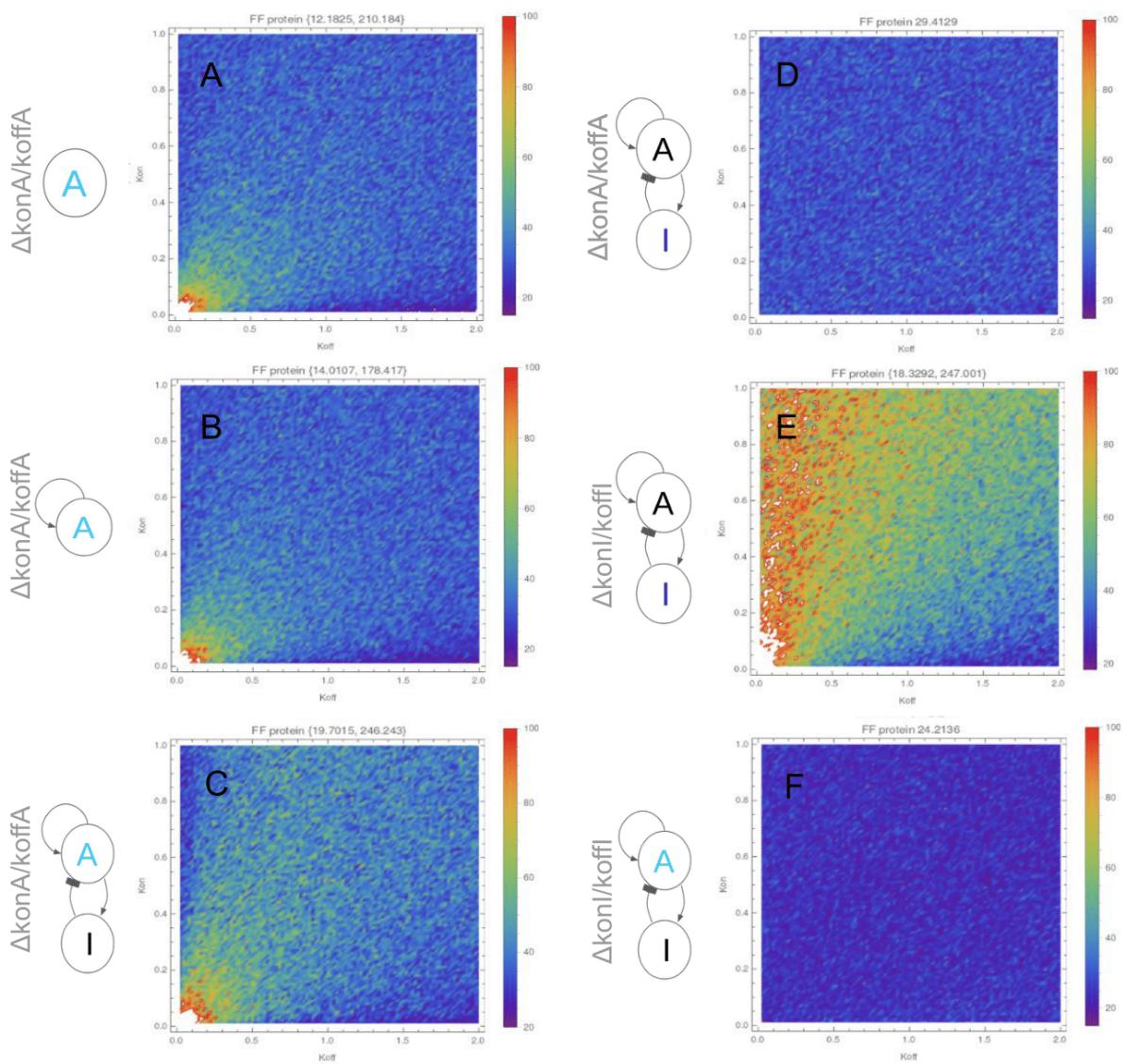


Figure 11. Noise (FF) in protein expression of the regulatory systems evaluated for a range of koff and kon parameter values. A. An unregulated gene without explicit regulation, B. self-activated gene, C. A gene of the Activator-Inhibitor regulatory system when it is changed the parameters of A, D. / gene of the Activator-Inhibitor regulatory system when it is changed the parameters of A, E. / gene of the Activator-Inhibitor GRN when it is changed the parameters of I, F. A gene of the Activator-Inhibitor GRN when it is changed the parameters of I. The values over the plots are the minimum and maximum or the mean (when there is only one value). *Simulated with NGM and the default parameters in Table S9.

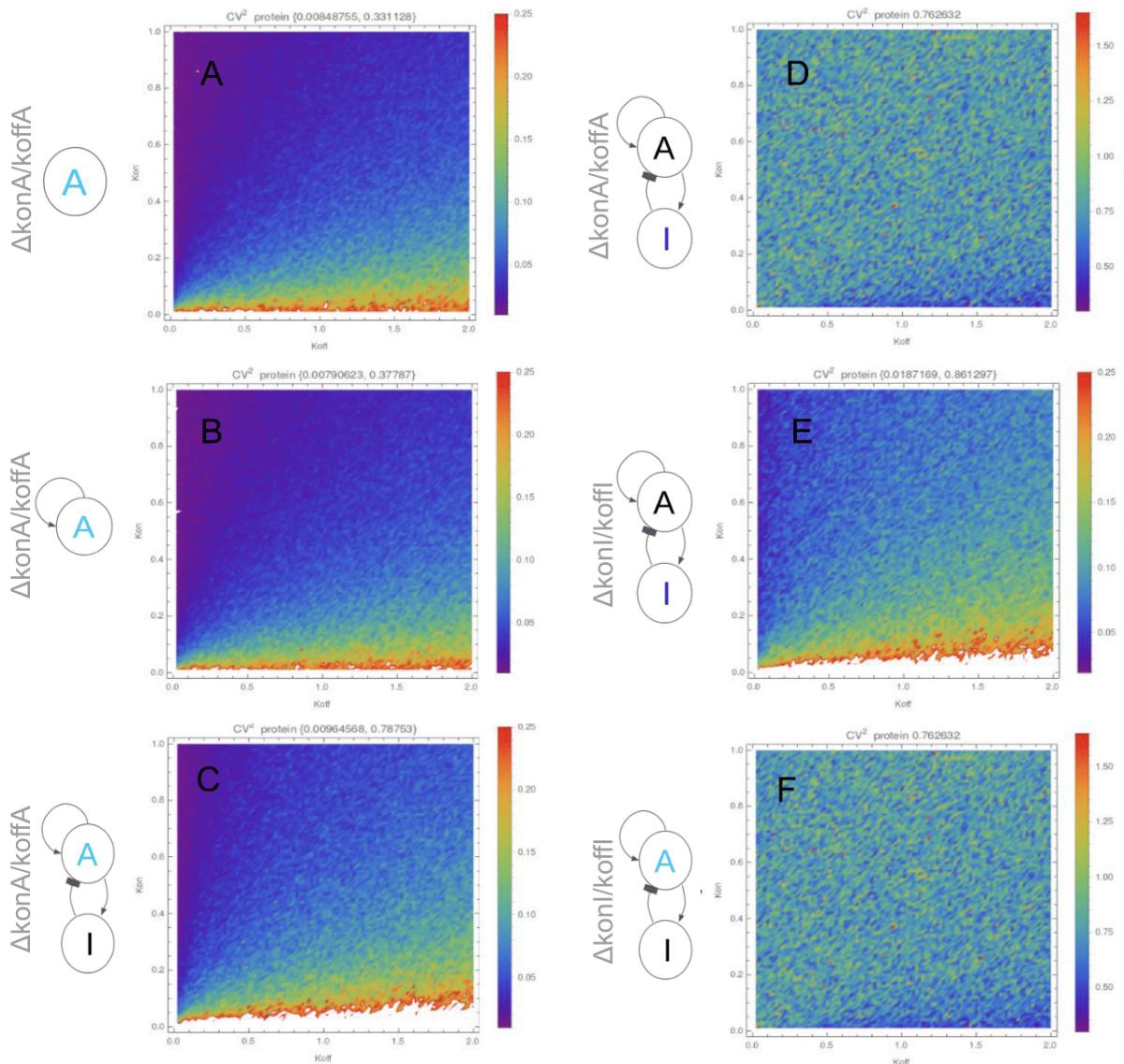


Figure 12. Noise (CV^2) in protein expression of the regulatory systems evaluated for a range of k_{off} and k_{on} parameter values. A. An unregulated gene without explicit regulation, B. self-activated gene, C. A gene of the Activator-Inhibitor regulatory system when it is changed the parameters of A, D. I gene of the Activator-Inhibitor regulatory system when it is changed the parameters of A, E. I gene of the Activator-Inhibitor GRN when it is changed the parameters of I , F. A gene of the Activator-Inhibitor GRN when it is changed the parameters of I . The values over the plots are the minimum and maximum or the mean (when there is only one value). *Simulated with NGM and the default parameters in Table S9.

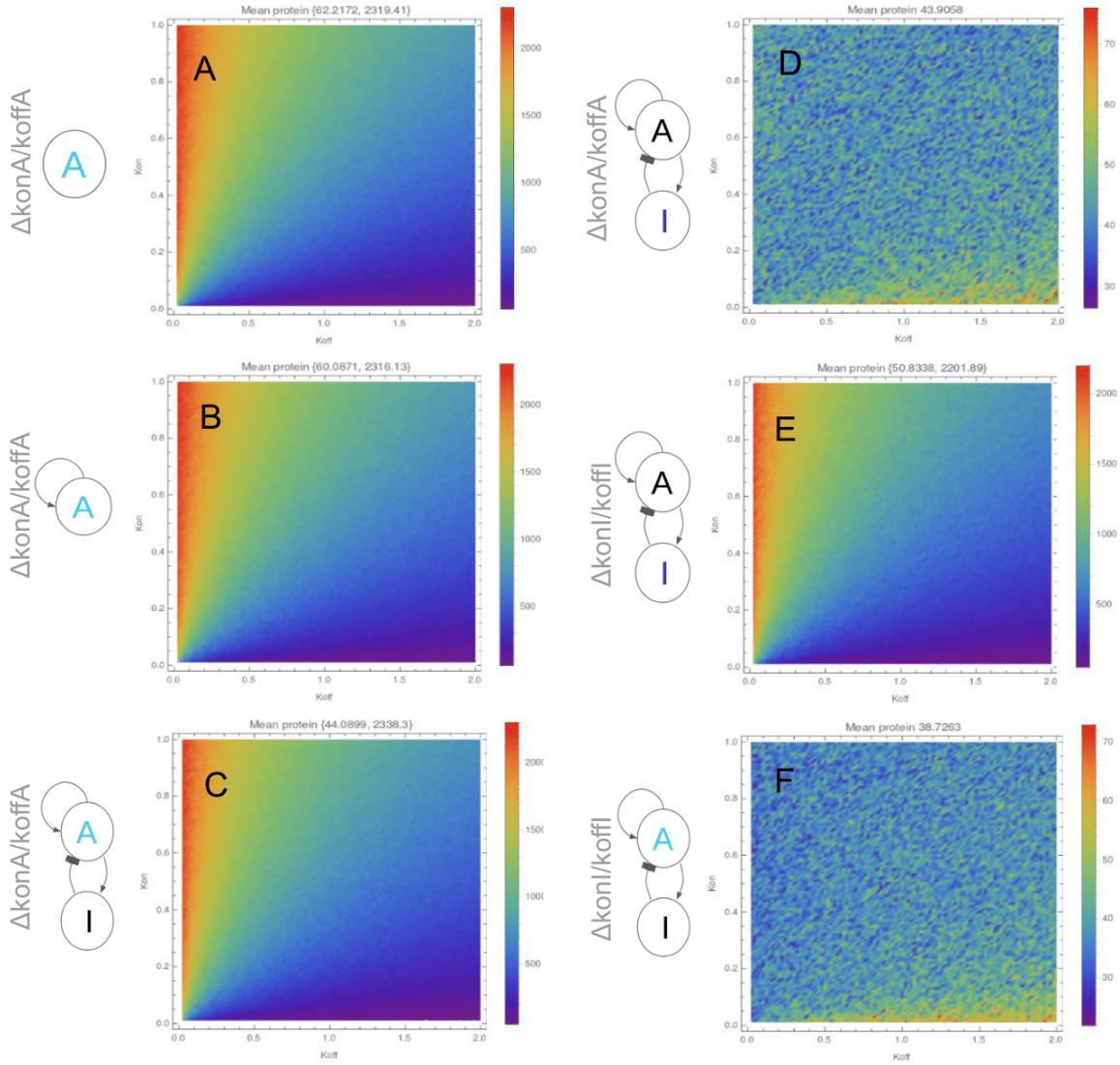


Figure 13. Mean of the protein expression of the regulatory systems evaluated for a range of koff and kon parameter values. A. An unregulated gene without explicit regulation, B. self-activated gene, C. A gene of the Activator-Inhibitor regulatory system when it is changed the parameters of A, D. / gene of the Activator-Inhibitor regulatory system when it is changed the parameters of A, E. / gene of the Activator-Inhibitor GRN when it is changed the parameters of I, F. A gene of the Activator-Inhibitor GRN when it is changed the parameters of I. The values over the plots are the minimum and maximum or the mean (when there is only one value). *Simulated with NGM and the default parameters in Table S9.

It is possible to identify qualitatively four regions of different gene expression types based on FF, CV^2 , and mean number of molecules for an unregulated gene (Fig.14). In figures 15 and 16 these regions are organized from left to right according to the ratio between k_{off} and k_{on} . Throughout the regions, the expression goes from a discontinuous burst to a more continuous both at the mRNA and protein level (Fig. 15-16 A). At the same time, the auto-correlation curve goes from high to low x-intercepts at both mRNA and protein levels (Fig. 15-16 B). The magnitude of the changes in expression (i.e., velocities) increases throughout the regions (Fig. 15-16 C).

The steady-state distribution is the plot that better describes the expression type (Fig. 15-16 D). When the expression is in discontinuous bursts this plot has a bias toward the right and

the probability to have low levels of expression is considerable (Fig. 15-16 D1-2). But the contrary, when the expression is more continuous the distribution is more gamma and there is no probability to have a low expression (Fig. 15-16 D3-4).

The FF can not be a measure to indicate if there is or no regulation at the promoter level. This is different from the previously pointed out in [21], but is the same that was indicated in section 2. First, because at the mRNA level when the ratio between k_{off} and k_{on} is low, the FF is closed to 1 and this is similar to the FF of a constitutive expression (Fig. 14 E3-5). Second, because at the protein level, the FF is higher than 1 for a constitutive expression (Fig. 15 E5).

Additionally, the FF can not be a measure to indicate if there is or no discontinuous burst expression. Because in both regions with expression in bursts (i.e., regions 1-2) the FF is very different (Fig. 15-16 E1-2). Indeed, an expression in burst could have a FF equal or lower than the FF for a continuous or constitutive expression (Fig. 15-16 E1, E4-5). This means that the burst expression of genes does not necessarily imply high noise (FF). As it has been previously mentioned, the burst expression implies high noise only for low promoter velocities [76].

As a summary, it is possible to identify four types of expression throughout the regions (Fig.14). In region 2, there are discontinuous bursts with low FF because they have low burst size and mean of molecules. In region 1, there are discontinuous bursts with high FF because they have high burst size and mean of molecules. In this region, there is a zone with the highest burst size and mean of molecules when k_{on} is bigger than k_{off} (Table 1). This support the fact that strong enhancers generate more bursts than weaker enhancers [97]. In region 4, there is a more continuous expression with high FFs. Finally, in region 3, there is an expression very similar that of a constitutive expression.

It is well known that developmental gene expression is noisy, and this is due to transcriptional bursting [98] and low copy numbers [64]. Developmental genes such as *ush* and *hnt* are expressed around 200 molecules [99]. Other genes such as Drosophila gap gene Kruppel (Kr) has CV^2 around 0.25 [57]. Therefore, because of molecules number and CV^2 estimates some developmental genes could be in regions 1 and 2. This is in accordance with their promoter activation and deactivation rates (Table S2).

.

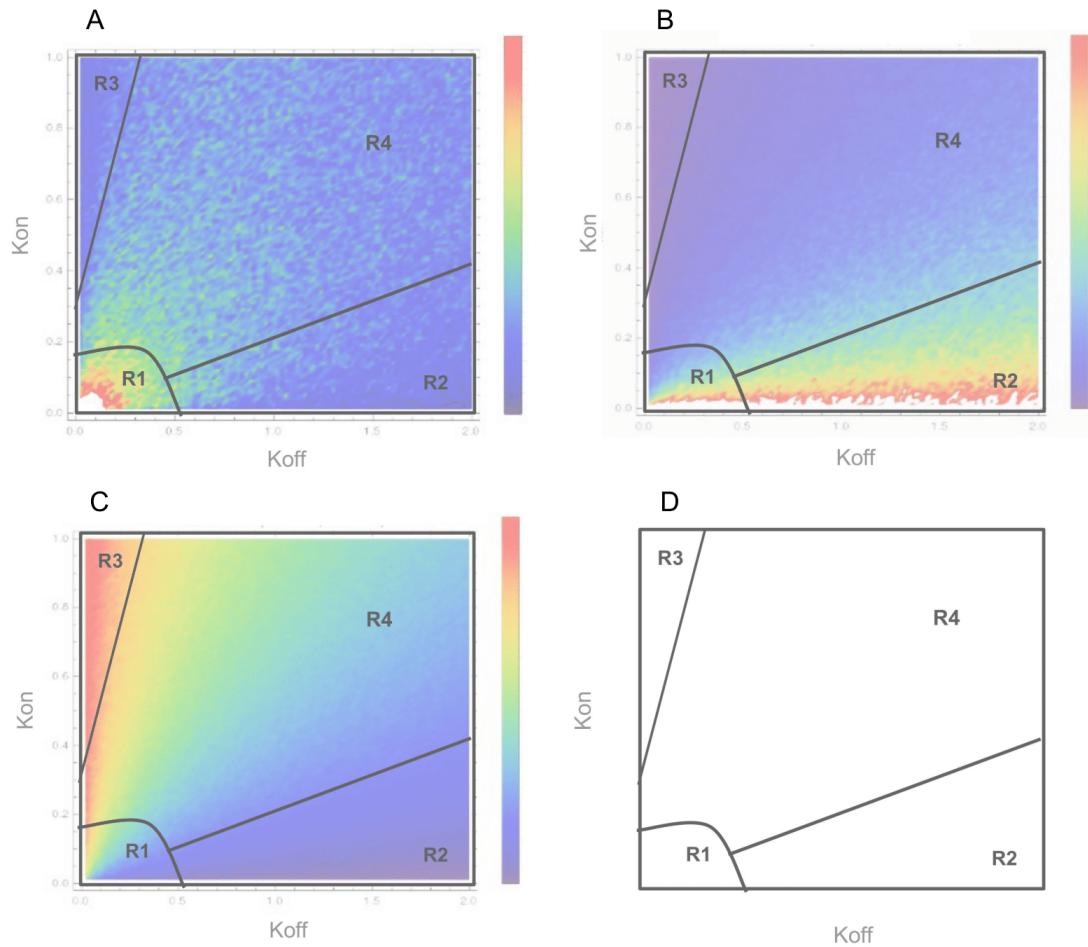


Figure 14. Regions that represent the four gene expression types. A. The four regions over the FF pattern of an unregulated gene, B. The four regions over the CV^2 pattern of an unregulated gene, C. The four regions over the mean pattern of an unregulated gene, D. The four regions. *Simulated with NGM and the default parameters in Table S9.

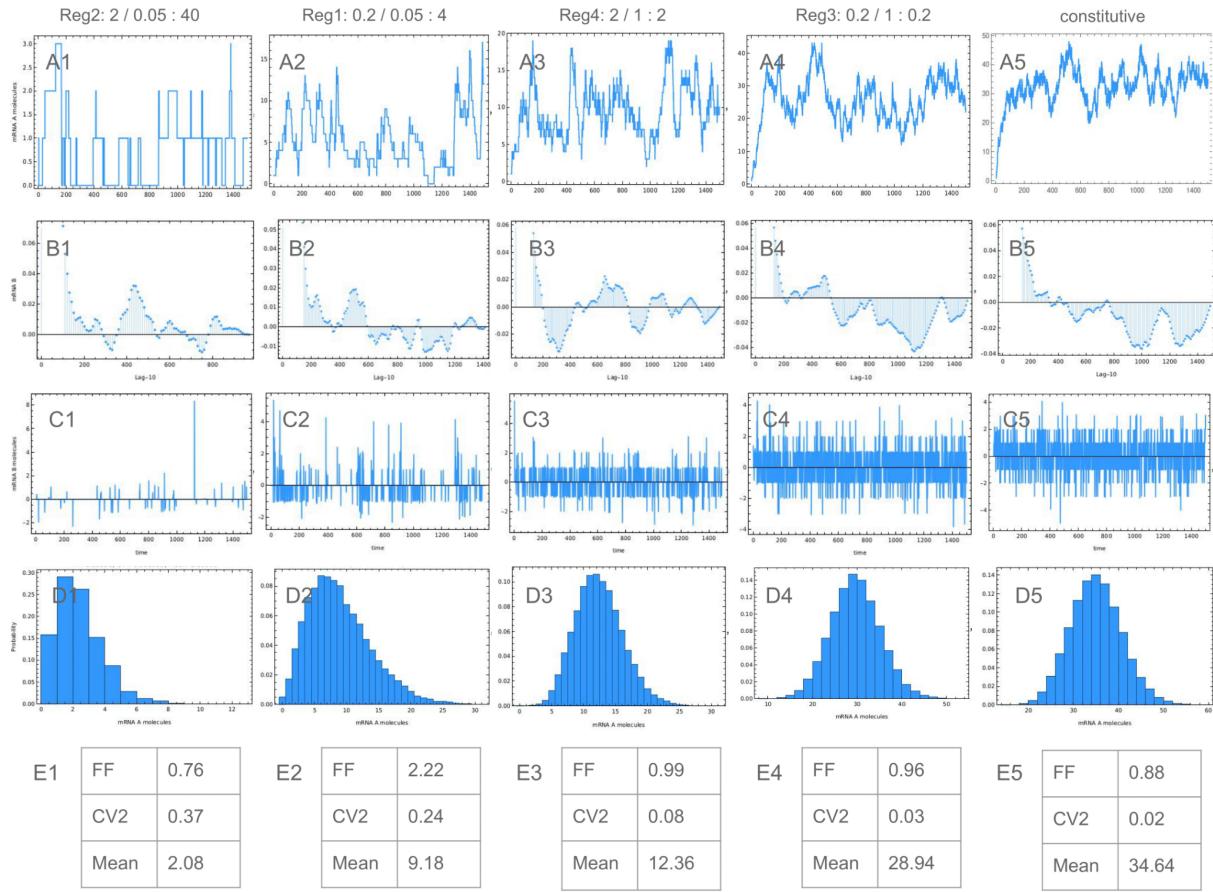


Figure 15. Dynamical characteristics of the four gene expression types at the mRNA level. A. Temporal expression, B. Autocorrelation of the temporal dynamic for a lag of 10, C. Velocities of the temporal dynamic, D. Steady-state distribution of the temporal dynamic, 5. All the plots for a constitutive gene. The regions are organized from left to right according to the ratio between k_{off} and k_{on} . *Simulated with NGM and the default parameters in Table S9.

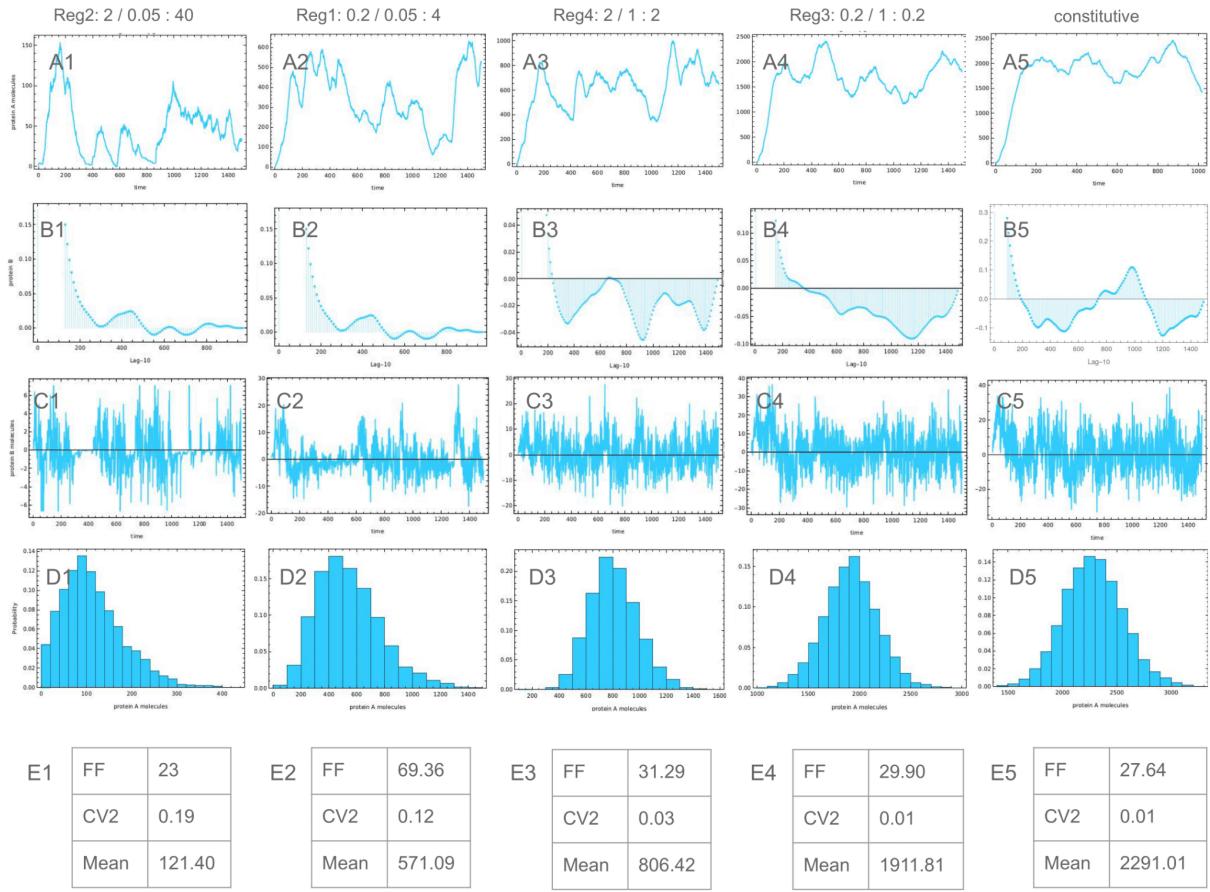


Figure 16. Dynamical characteristics of the four gene expression types at the protein level. A. Temporal expression, B. Autocorrelation of the temporal dynamic for a lag of 10, C. Velocities of the temporal dynamic, D. Steady-state distribution of the temporal dynamic, 5. All the plots for a constitutive gene. The regions are organized from left to right according to the ratio between k_{off} and k_{on} . *Simulated with NGM and the default parameters in Table S9.

Previous studies concluded that fluctuations in regulatory proteins (i.e., TFs and signaling molecules) propagate down a GRN [57]. This significantly alter the expression levels and noise of downstream target genes [57]. In this study, we also concluded this but we also propose new and important ideas about noise propagation as will be explained below.

For instance, the noise propagates through a regulatory connection from regulator to regulated gene independently of the amount of noise of regulator. Figure 17 A and C shows the noise pattern of both regulators when their promoter activation and deactivation rates are changed. Independently of noise pattern in regulators, the regulated genes increase significantly their noise compared with the noise in un-regulated gene (Table S10). But this increment is the same in the whole region of parameters values (Fig. 17 B and D).

The noise propagation through the regulatory connection depends on the regulatory system. The noise pattern of the self-activated system is equal to that of an unregulated gene. But the noise is bigger in an activated or inhibited gene than that of an unregulated gene (Fig. 17 and Table S10).

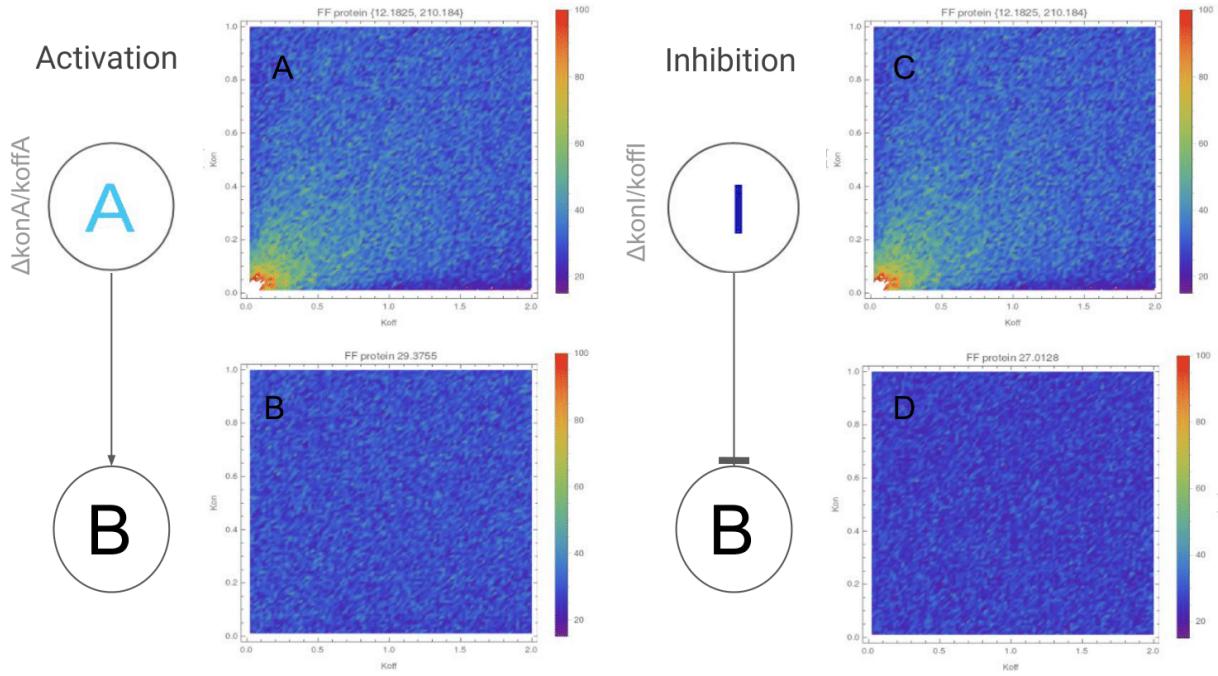


Figure 17. Noise (FF) in the protein expression of an Activation and Inhibition types of regulation with variations in the regulator. Here the parameters k_{off} and k_{on} of both regulators are changed. A. FF for Activator/Regulator, B. FF for the Activated/Regulated gene, C. FF for Inhibitor/Regulator, D. FF for the Inhibited/Regulated gene. The values over the plots are the minimum and maximum or the mean (when there is only one value). *Simulated with NGM and the default parameters in Table S9.

Noise is propagated through the regulatory connection depending on the expression type of the regulated gene. Figures 18 B and D show the noise pattern of both regulated genes when their promoter activation and deactivation rates are changed. In this case, the promoter rates of regulators are not changed. For both regulation types (i.e., activation or inhibition), the noise increases compared to the noise of an unregulated gene (Table S11). But this increment is not homogeneous throughout the parameters values or throughout regions of expression.

To explain this it will be used the activation case in figure 18 A-B. For this case, the noise is bigger in regions 3-4 than in regions 1-2 compared to unregulated gene (Table S11). This happens because the regulated gene has more continuous and higher expression in regions 3-4 than in regions 1-2 (Fig. 19A-B). For this reason in regions 3-4, the changes in regulator generate changes in regulated gene, and these changes are bigger than those expected for an unregulated gene in these regions (Fig. 19A and C). But the contrary in regions 1-2, the changes in regulator also generate changes in regulated gene, but these changes are similar to those expected for unregulated gene in these regions (Fig. 19B and D).

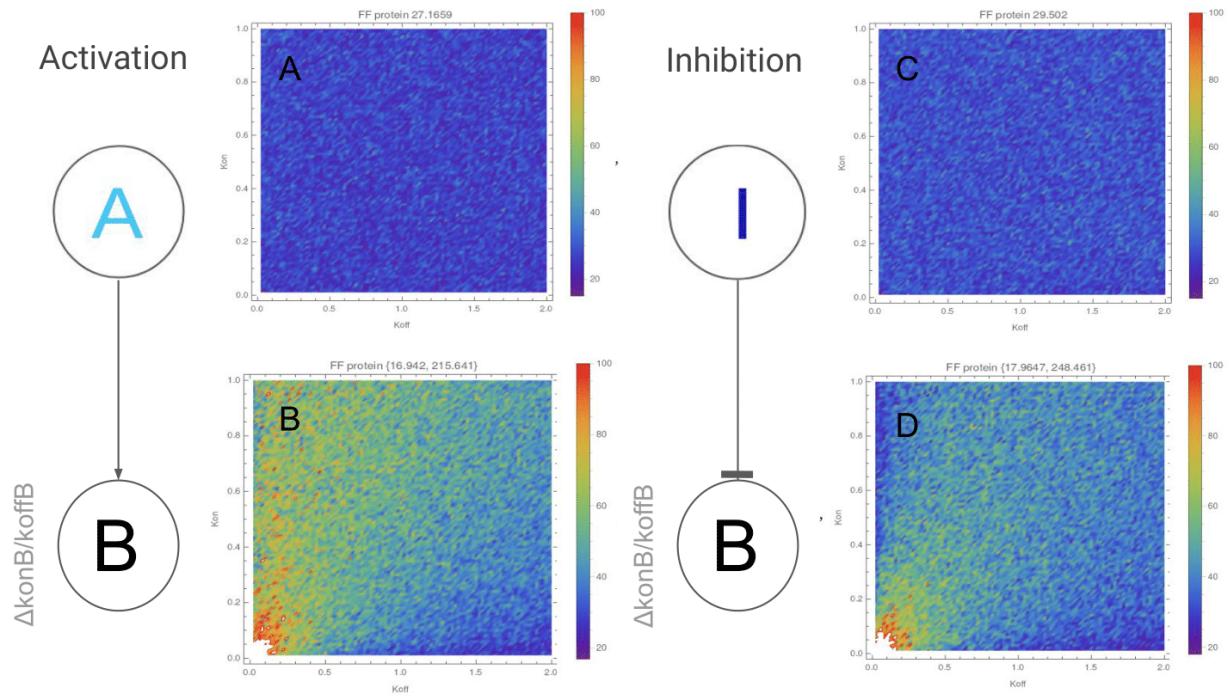


Figure 18. Noise (FF) in the protein expression of an Activation and Inhibition types of regulation with variations in the regulated gene. Here the parameters k_{off} and k_{on} of both regulated genes are changed. A. FF for Activator/Regulator, B. FF for the Activated/Regulated gene, C. FF for Inhibitor/Regulator, D. FF for the Inhibited/Regulated gene. The values over the plots are the minimum and maximum or the mean (when there is only one value). *Simulated with NGM and the default parameters in Table S9.

At the same time, noise is propagated through the regulatory connection depending on the expression type of regulator gene. As it is already shown in Figure 19 C-D, the regulator is expressed in discontinuous burst, and this generates burst expression in its regulated gene (Fig. 19C). But what happens when both regulator and regulated gene are expressed in a more continuous way. To explain this it will be used the inactivation case in Figure 20.

When both regulator and regulated gene are expressed in continuous way the inhibition occur without alter the expression type of inhibited gene and with both low FF and mean of molecules (Fig. 20 A-B). When regulator is expressed in big burst and regulated gene in a continuous way the inhibition occur altering the expression type of inhibited gene and with high FFs and low mean of molecules (Fig. 20 A-B). Finally, when regulator is expressed in small burst and regulated gene in a continuous way the inhibition did not occur well and there is not and alteration of the expression type of inhibited gene and there are low FFs and high mean of molecules (Fig. 20 A-B).

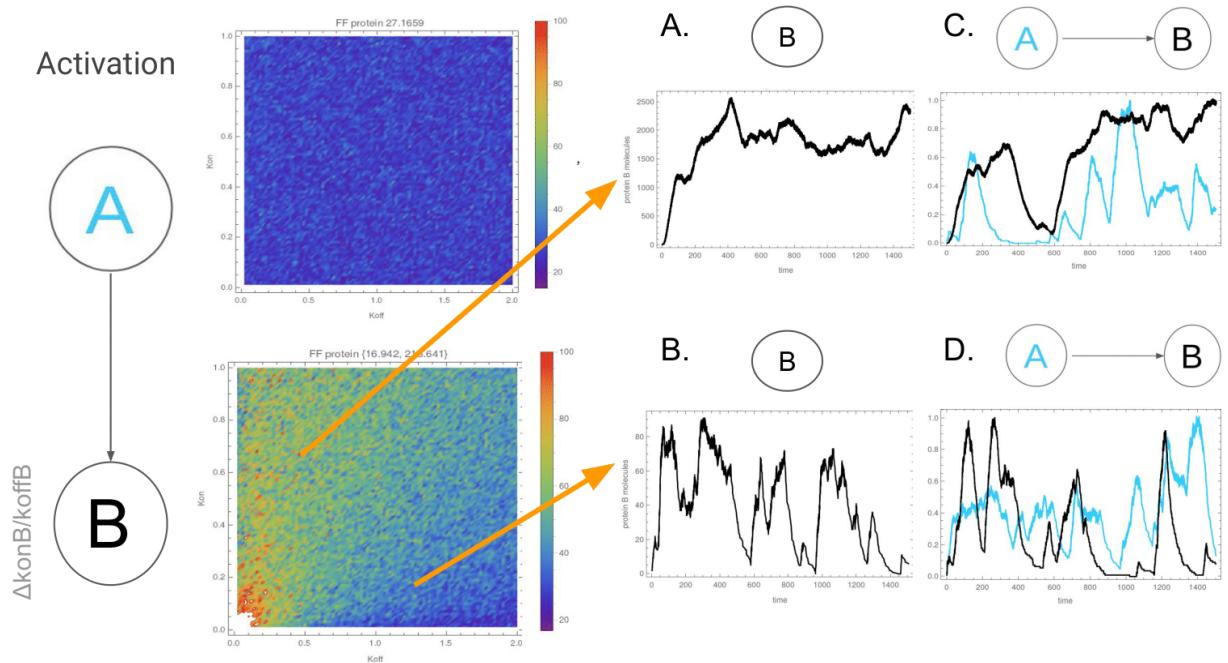


Figure 19. Temporal expression in regions 2 and 3-4 of a gene with and without regulator. A. Temporal expression of an unregulated gene (B) in regions 3 and 4, B. Temporal expression of an unregulated gene (B) in region 2, C. Temporal expression of activator (A) and activated (B) gene in regions 3 and 4, D. Temporal expression of activator (A) and activated (B) gene in region 2. The values over the plots are the minimum and maximum or the mean (when there is only one value). *Simulated with NGM and the default parameters in Table S9.

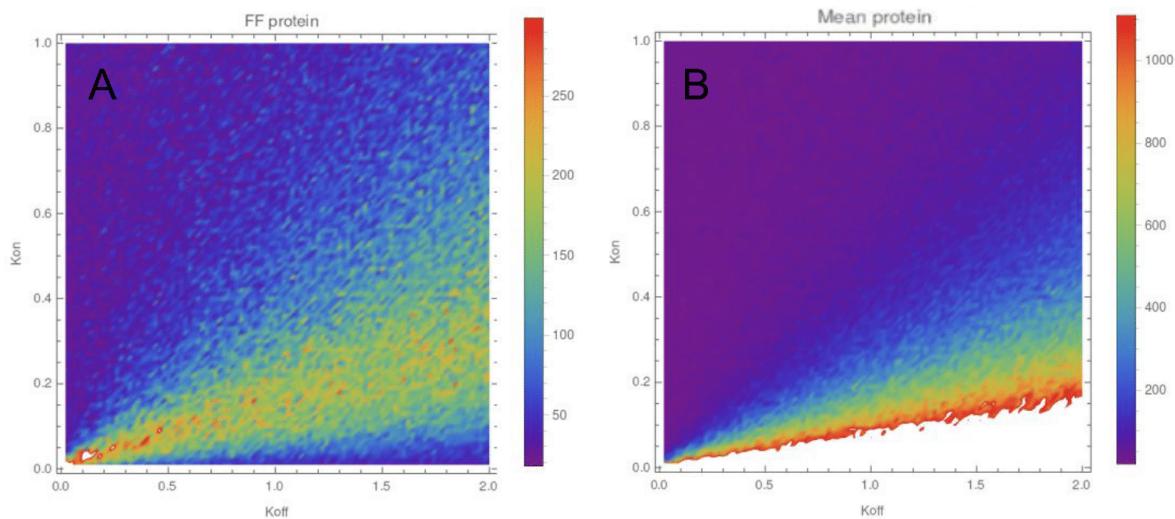


Figure 20. Noise (FF) and mean of the protein expression of an inhibited gene for variations in its inhibitor. Here it is changed the parameters k_{off} and k_{on} of the inhibitor but it is shown the noise of the inhibited gene when this has a k_{on} of 1 and k_{off} of 0.02. The other parameters are in Table S9 and it was simulated with NGM. A. Noise (FF) pattern, B. Mean of protein molecules.

The propagation of noise is associated with a reduction in the mean of molecules. As it was previously mentioned, if the regulator expresses in burst it makes the regulated gene expression in bursts as long as this last has a more continuous expression. This reduces for some period the levels of expression and consequently, it reduces the mean of molecules.

Figures 11 and S8 show that Activator-Inhibitor regulatory system does not reduce the noise in expression compared with an unregulated gene. On the contrary, the noise is higher than in an unregulated gene. This increase in noise in both genes is due to noise propagation from regulator to regulated gene (Fig. 11C-F, S8C-F). These fluctuations in expression of both genes could facilitate its function in development. For instance, it allows an heterogeneous expression of genes in time and space and this is implied in self-organized fate patterning [83]. Which means, the noise could be beneficial to drive phenotypic diversity as previously pointed out [71].

Previous studies mention that stochasticity and low intracellular copy number limit the cellular signal precision and then the gene regulation [1,100]. Here, we found that the effectiveness of regulation also depends on the continuity of regulator expression and it is independent of regulator amount of noise (FF). Figure 21D-E shows that when the inhibitor has a continuous expression it has a more effective inhibition of gene B than when it is expressed in bursts (Fig. 21C). The effectiveness of inhibition is defined as a higher reduction in the number of molecules of a regulated gene compared to its expression level in the unregulated case (Fig. 21 A-B). In Figure 21D-E both regulators have a different amount of noise but both reduce at similar level the gene B. This is because in Figure 24D the regulator has a low probability to have low expression levels. This means the regulator has a more continuous expression as in Figure 24E.

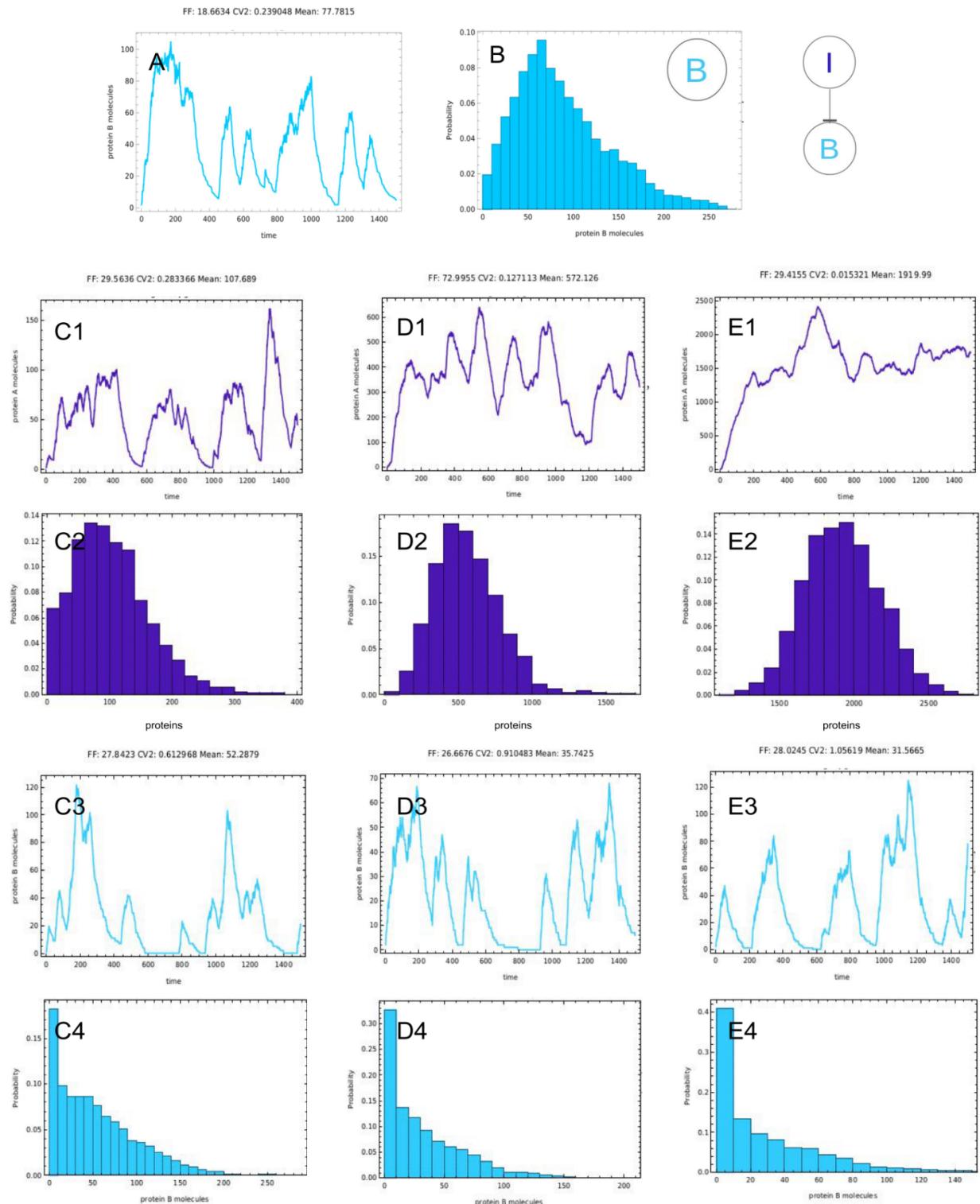


Figure 21. Temporal expression characteristics of both inhibitor and inhibited gene. I: Inhibitor gene (purple), B: Inhibited gene (blue), A-B: Temporal expression and steady-state distribution for the unregulated gene B, C. Temporal expression and steady-state distribution of molecules for both I and B genes, when I has a k_{on} of 0.05 and k_{off} of 2 and B has the default parameters, D. Temporal expression and steady-state distribution of molecules for both I and B genes, when I has a k_{on} of 0.05 and k_{off} of 0.2 and B has the default parameters, E. Temporal expression and steady-state distribution of molecules for both I and B genes when I has a k_{on} of 1 and k_{off} of 0.2 and B has the default parameters. *Simulated with NGM and the default parameters in Table S9.

1.

Transcriptional bursts are an important feature of gene activity in living embryos [28]. But how organism establish and maintain the precise levels of gene expression seen during development?. Some described mechanisms are redundancy in genetic circuits to achieve the precision required for proper development [57]. It has also been described that shadow enhancer filters the noise in TF due to input separation (i.e., each enhancer respond to a different TF). In this way, the target gene is less sensitive to levels of a single TF [57].

Other studies concluded that some GRN, such as interlinked Feed-forward loops, are effective in filtering noise [51]. But there are different studies pointing out different conclusions about self-activated gene. For instance in [101], feedback loop is a noise controller. On the contrary [71] points out that the noise is amplified by HIV-Tat positive feedback. In our study the self-activated gene did not have higher noise than the noise of an unregulated gene. These discrepancies could be due to the model assumptions. For example, in [71] model Tat protein affects the transcriptional rate but in our model, the feedback protein affects the promoter activation rate. The first one affects the burst size and then the noise but the second one only affects the burst frequency. Finally, according to this study, self-activation does not increase noise but activation by another gene does.

Section 5: Dynamical properties of the regulatory systems for coupled cells

The diffusion of a paracrine signal throughout a colony of cells decreases the noise (FF) in the whole range of parameters evaluated (Fig. 22). It also happens that the magnitude of this reduction depends on the parameters' values (Fig. 22 A3-C3). For instance, for the parameters k_{on} and k_{off} , the highest reduction happens in region 1 where the expression is in big bursts and high FFs (Fig. 22 A3). Also for these parameters, the lowest reduction happens in region 2 where the expression is in small burst and low FF (Fig. 22 A3). This could mean that there is a minimal noise that the diffusion could not reduce. In summary, for all the parameters evaluated it happens that the highest noise reduction occurs in the region of highest noise and the lowest reduction in the region of lowest noise (Fig. 22 A3-C3).

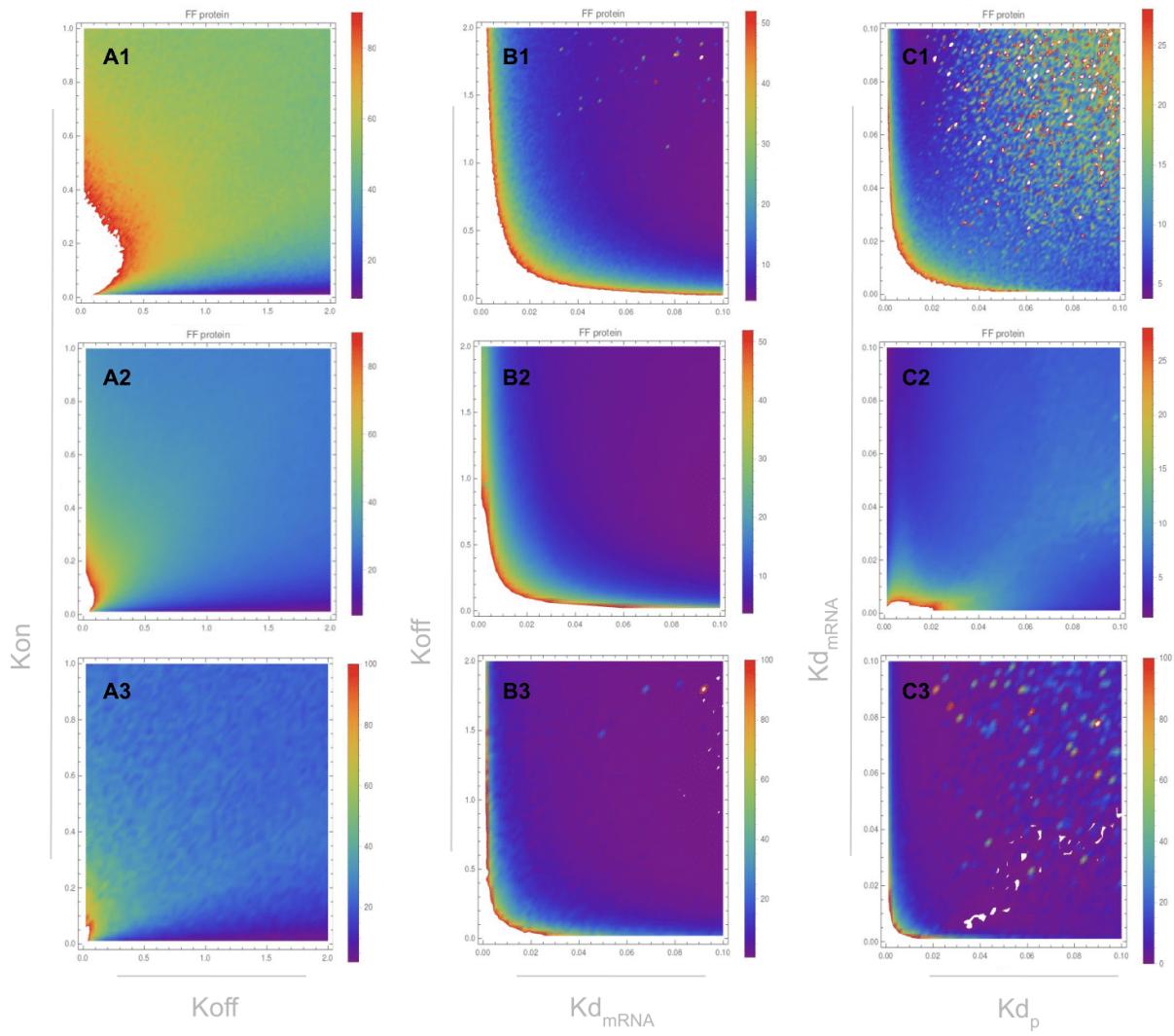


Figure 22. Noise (FF) of the protein expression of a unicellular and multicellular approach. 1. FF pattern for a unicellular approach, 2. FF pattern for a multicellular-colony approach, 3. Differences between the FF values between the two approaches, A. FF values when it is changed k_{on} and k_{off} , B. FF values when it is changed k_{off} and $k_{d_{mRNA}}$, C. FF values when it is changed $k_{d_{mRNA}}$ and k_{d_p} . *Simulated with CLE1 and the default parameters in Table S9.

The mean of molecules and noise (FF) pattern in a multicellular approach throughout the region of k_{on} and k_{off} parameters values is qualitatively similar to that of a unicellular approach (Fig. 22 A1-A2 and 23 A). On the other hand, in a multicellular approach and throughout the region of k_{off} and mRNA degradation rate ($k_{d_{mRNA}}$), the mean of molecules and noise (FF) present the highest values when both parameters are low. These estimates are also the highest when one of the parameters is low and the other one fluctuates between low and high values (Fig. 23 B1-B2). For instance, when $k_{d_{mRNA}}$ is low the translational efficiency (i.e., $k_s/k_{d_{mRNA}}$) increases. Then, the noise (FF) is higher because protein dynamics has greater fluctuations and a broader steady state distribution as mentioned in [65,75].

When both k_{off} and $k_{d_{mRNA}}$ are low, the gene expression is as in region 1 (i.e., big burst with high FFs). This occurs because the default parameter of k_{on} is 0.011 (Table S9). From the previous section, it is known that the increase in k_{off} decreases the burst size and noise (i.e.,

region 2). However, if $k_{d_{mRNA}}$ is very small it is possible to have big burst with big FF (Fig. 23 B1-B2).

The mean of molecules pattern in a multicellular approach throughout the region of $k_{d_{mRNA}}$ and k_{dp} present the highest values when both are low. These estimates are also the highest when one of the parameters is low and the other one fluctuates between low and high values (Fig. 23 C1). When both parameters are low, the mean of proteins is high because of the low degradation. When $k_{d_{mRNA}}$ is low, the mean of proteins is high, even if the protein degradation rate increases, because the increment in translational efficiency [65]. When k_{dp} is low, the mean of protein is high even if the mRNA degradation rate increases.

On the other hand, the FF pattern in a multicellular approach throughout the region of $k_{d_{mRNA}}$ and k_{dp} parameters values present the highest values when both are low (Fig. 23 C2). This happens because with the default parameters the gene expression is in region 2 (Tabla S9, Fig. 14). However, the expression becomes more similar to that in region 1 when the mRNA and protein degradation rates are lower than their default values.

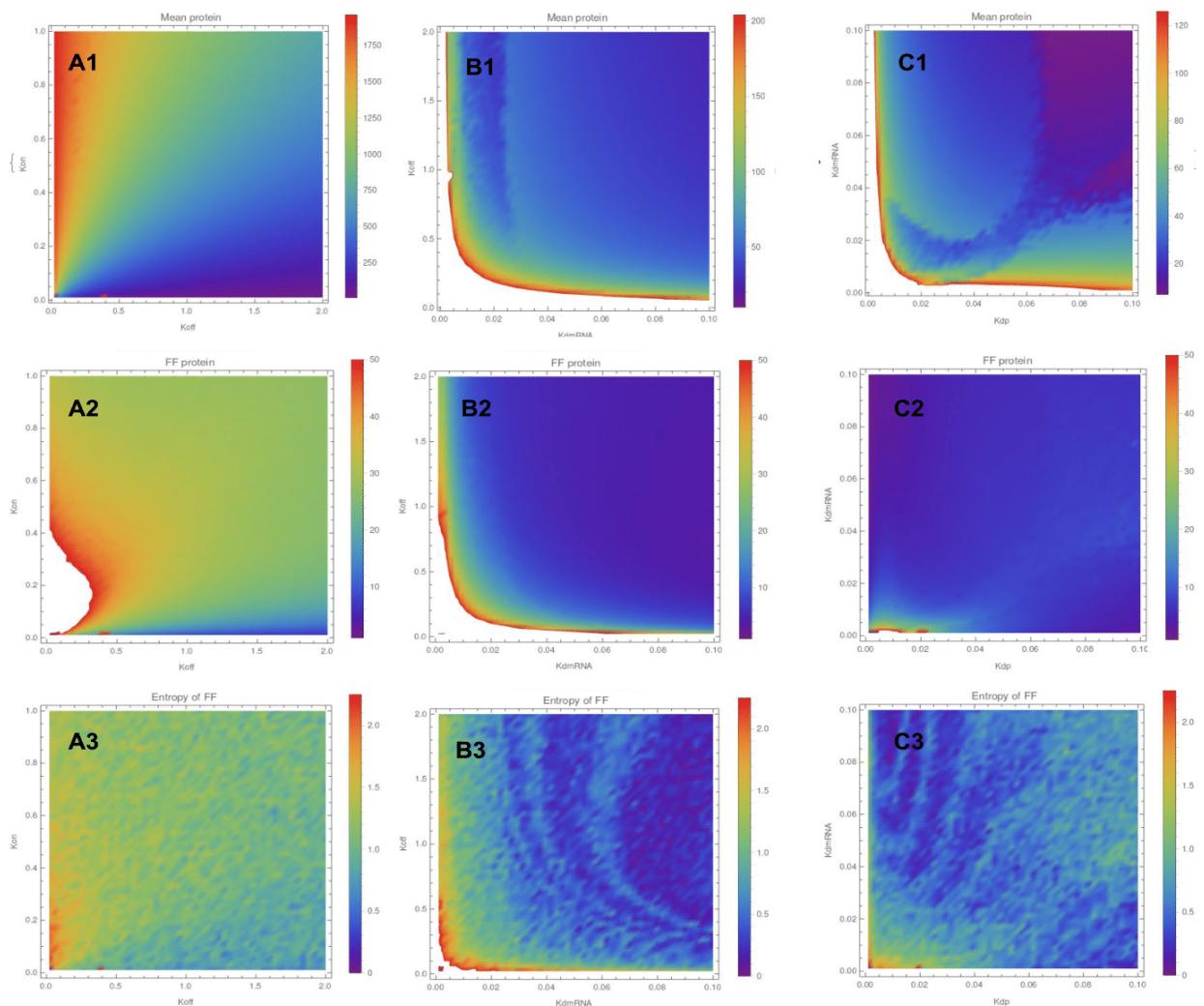


Figure 23. Mean, FF, and Entropy of the protein expression of a multicellular approach. 1. Mean of protein expression, 2. FF of protein expression, 3. Entropy of FFs values in consecutive, A. For parameters k_{on} and k_{off} , B. For parameters k_{off} and $k_{d_{mRNAf}}$, C. For parameters $k_{d_{mRNAf}}$ and k_{dp} . *Simulated with CLE1 and the default parameters in Table S9.

For all parameters evaluated, the entropy pattern of FFs by consecutive circles along the colony is very similar to the FF pattern (Fig. 23 A2-C2 and A3-C3). It means that the heterogeneity in FF values by consecutive circles along the colony is highest when the gene expression is in big bursts as in region 1 (Fig. 24 C-D, Gif. S1-S2). When the expression is in small bursts as in region 2, there is no difference between the FF calculated at the colony center and for the whole colony (Fig. 24 A-B, Gif. S3-4). It means that the FF values by consecutive circles along the colony are homogeneous or similar.

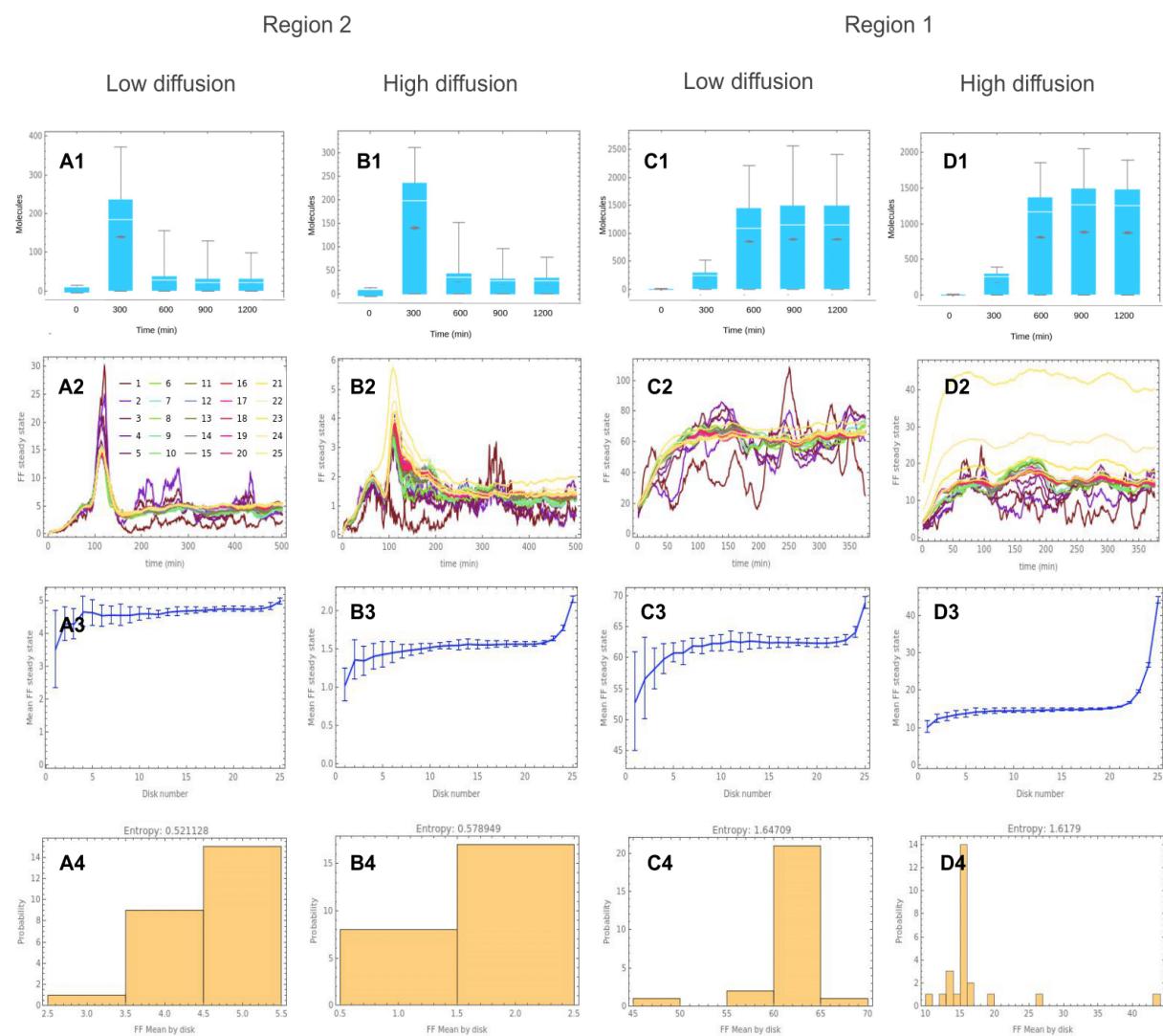


Figure 24. Temporal expression characteristics for a multicellular approach in two regions of low and high noise. Region 2: k_{on} is 0.01 and k_{off} is 2, Region 1: k_{on} is 0.1 and k_{off} is 0.06. 1. Temporal expression throughout time for all the cells in the colony. The boxplots for each time step have the mean (gray diamond), median (white line), and maximal values, 2. Temporal FF values for cells inside each consecutive circle (disk) in each time step. Disk 1(Brown) covers the cells at the most center and disk 25 (Yellow) covers the whole colony, 3. Mean of the temporal FF values for cells inside each consecutive circle (disk). Disk 1 covers the cells at the most central and disk 25 over the whole colony.

The increase in diffusion rate decreases the mean of molecules in the regions of expression (Fig. 25 A2-4). This does not happen for region 2 because it has a particular temporal dynamic that is not well described by the mean (Fig.24 A1-B1). Furthermore, the increase in diffusion rate decreases the noise (FF) in the four regions of expression (Fig. 25 B1-4). From the previous sections, we can see that diffusion does not affect the burst size or frequency (Table 1I). Therefore, the decrease in FF is because it increases the expression in those cells with low levels and decreases it in those with high levels.

As a consequence, there is a decrease in the mean (Fig. 25 A2-4), an increase in the median, and a decrease in the maximal values of expression (Fig. 24 A1-D1). However, the diffusion does not completely avoid the low expression levels and remains cells with very low expression (Fig.24 A1-D1). This could affect the effectiveness of regulation because, as was pointed out in section 4, this is lower when the regulator is expressed in burst with periods of low expression.

When the expression becomes more continuous and higher, it appears a zone after a minimum colony size where the increase in diffusion rate increases the FF values (Fig. 25 B2-4). This is because the cells in the center of the colony are more synchronized with each other than with the cells at the edge (Gif.S5-6). Therefore, the FF for the whole colony is higher than the FF for the colony center (Fig. 24 D2 and D3). For this reason, these zones are associated with the highest entropies (Fig. 25 C2-4). When the colony is larger this phenomenon starts to fade out and the expression in the entire colony becomes more homogeneous, therefore the entropy is lower (Fig. 25 C, Gif.S7-8).

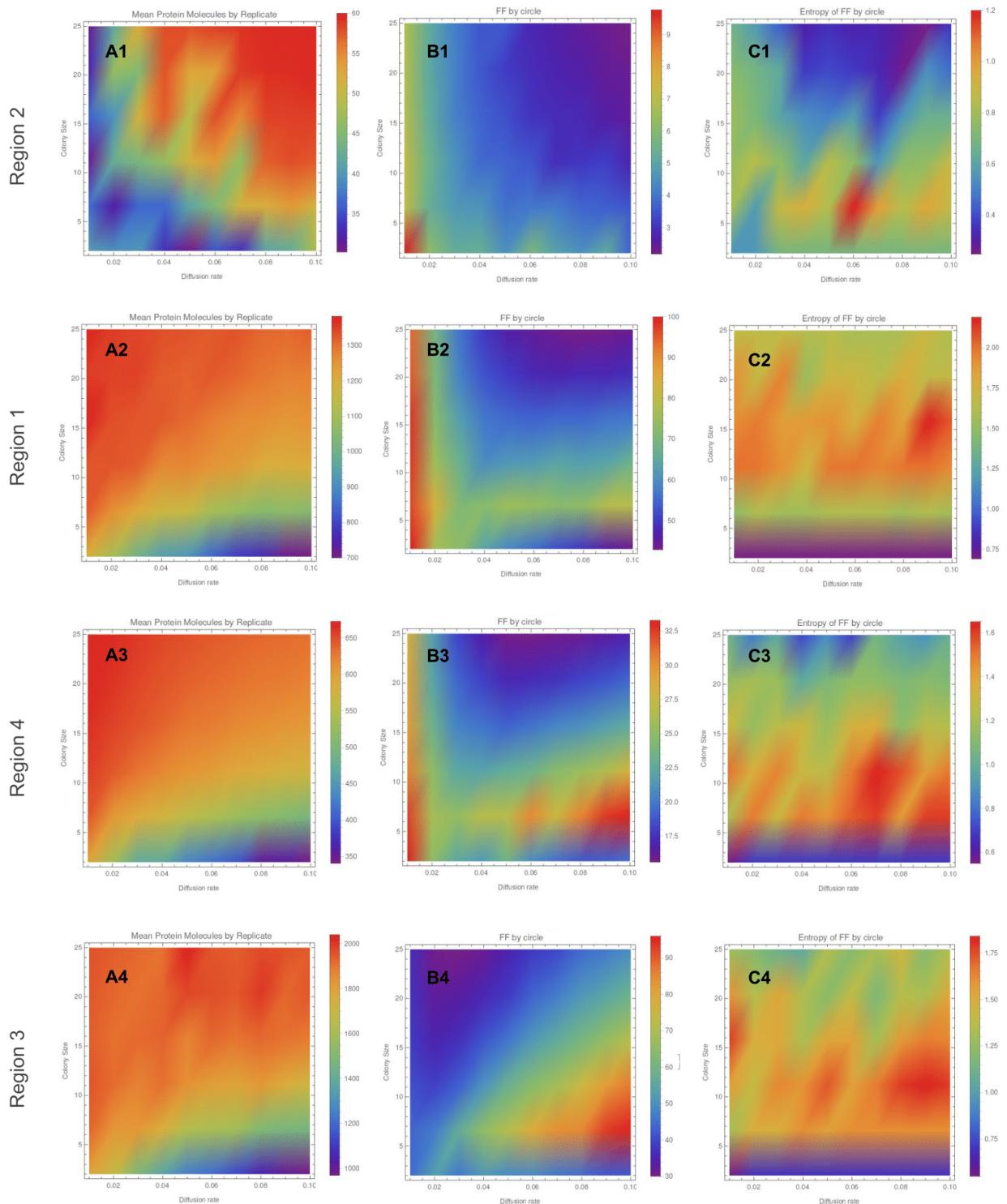


Figure 25. Estimates for a multicellular approach when it is changed the k_{on} and k_{off} parameters to represent the four regions of expression. 1. Region 2 with k_{on} of 0.01 and k_{off} of 2, 2. Region 1 with k_{on} of 0.1 and k_{off} of 0.06, 3. Region 4 with k_{on} of 0.8 and k_{off} of 1.8, 3. Region 5 with k_{on} of 1 and k_{off} of 0.02, A. Mean of protein expression, B. FF of protein expression, C. Entropy of FFs values in consecutive circles. *Simulated with CLE1 and the default parameters in Table S9.

This study apports new ideas about the relationship between the community effect, the colony size, and diffusion rate as will be described below. The community effect is the coordination of cells in a tissue to express the same set of genes that specifies a differentiation phenotype [70,73]. All cells in the community express the same set of target

genes because a cell signal regulates their transcription [102]. It is required a minimum number of cells for the effect to occur and below which the cells could not differentiate [73]. For instance, in the induction of tooth formation is necessary minimum 25% of inductive tooth mesenchyme cells in the tissue [102]. Additionally, the community effect depends on the strength of the interaction between cells that enforce a common fate, and this is associated with greater sensitivity to the external signal [103]. Because all the cell in the community express common signaling receptors, the most likely cause of the loss of common gene expression is the loss of cell signal secretion [102].

Some previous studies have modeled this phenomenon and they have included the promoter regulation in their models [70,73]. However, none of them have evaluated the effect of diffusion rate as evaluated here. For instance, in [73], they evaluated a cellular communication parameter which is not analogous to the diffusion rate evaluated here. Because the increase in their parameter increases the mean of molecules but our parameter decreases it. Finally, none of the previous studies have evaluated the community effect with different types of expression (Fig. 25).

Figure 25 shows three important things. First one, the increase in the number of cells increases the mean number of molecules in region 1, 3, and 4. Second one, the increase in diffusion generates an increment in the homogenization of gene expression level because a reduction in FF values. Third one, the increase in diffusion generates a decrement in the mean number of molecules . But there is a minimum diffusion rate and colony size above which both the increment in diffusion rate and colony size imply low FF and high mean number of molecules. In other words, due to the loss of signal molecules by diffusion it is required a minimum colony size to get a community effect with high molecules number and high homogeneity throughout the colony (i.e., low FF).

The number of cells and the interaction between them by diffusion not only affect the interpretation of endogenous signals but also the regulatory capacity of the signal. Because it affects the homogenization of signal expression level throughout the cell colony decreasing the FF. From section 4 we know that higher homogenization imply better regulation of target genes in the whole colony.

CONCLUSIONS

Section 2: The three-stage model of one gene with NGM and CME

1. The promoter affinity is a crucial kinetic parameter to get an mRNA expression in discontinuous bursts of high noise (FF). With low affinities, discontinuous bursts are not only present for the typical case when k_{off} is bigger than k_{on} but also can be produced even when k_{off} is lower than k_{on} or when they are equal.
2. The expression pattern of a gene regulated at the promoter level becomes more similar to the expression of a constitutive gene by mean of three mechanisms. Increasing both k_{on} and k_{off} , increasing k_{on} over k_{off} , and increasing $k_{s_{mRNA}}$. The similarity between them does not allow to distinguish between the expression of a constitutive gene and a promoter-regulated gene.
3. The temporal dynamic of gene expression depends on the kinetic parameters in the whole level of expression.
4. Because of the similarity of our results with those previously pointed out in the literature, we can conclude some things for the methodological approach. First, the three-stage model is a good approach to represent a gene expression with promoter regulation. Second, NGM and CME are good approximations to simulate the noise in gene expression. However, NGM is not a good approach for big systems because of its low performance.

Section 3: The CLE implementation for the three-stage model is a good approach for large systems

1. CLEs and NGM are more similar in the estimates of FF than in the estimates of CV^2 . This similarity occurs mainly in the qualitative pattern throughout the parameters' values. And this is only applicable at the protein level but not at the mRNA level. For this reason, CV^2 is not a good estimator of noise when the CLE implementation is used.
2. Among the CLEs, the CLE1 has the highest similarity with NGM because the default parameters values of the system are following the parameter requirements of the CLE1 implementation.
3. CLE is not useful to study the characteristics of the temporal dynamics of gene expression as they could be studied with NGM. It happens because CLE gives at every time point the gene expression average, while NGM gives at every time point a

particular realization. As a consequence, the frequency and burst size are different between CLE1 and NGM.

4. The difference between NGM and CLE implementation also depends on the parameters' values. For this reason, we limited the next analysis to the same set and range of parameters previously evaluated in section 3.
5. We will estimate the noise by mean of FF and we will use the CLE1 system of equations to simulate isogenic cells in a circular colony.

Section 4: Dynamical properties of the regulatory systems for individual cells

1. The regulatory systems evaluated do not reduce the biological noise. On the contrary, the noise remains at the same level as in the self-activation system or increases as in the Activator-Inhibitor regulatory system, compared to an unregulated gene.
2. It is possible to identify four types of expression: In region 2, there are discontinuous bursts with low FF because they have a low mean of molecules. In region 1, there are discontinuous bursts with high FF because they have a high mean of molecules. In region 4, there is a more continuous expression with high FFs. Finally, in region 3, the expression is very similar to that of a constitutive expression.
3. The noise propagates through the regulatory connection from the regulator to the regulated gene. This is independent of the amount of noise of the regulator and the regulation type (i.e., activation or inhibition). But it does not propagate in the self-activation regulatory system.
4. The noise is propagated through the regulatory connection depending on the expression type of the regulator gene. It means that when the regulator expression is in discontinuous bursts, this will drive a burst expression in its regulated gene as long as this has a more continuous expression. But the expression of the regulated gene will remain continuous if the regulator expression is more continuous.
5. The propagation of the noise is associated with a reduction in the mean of molecules. As it was previously mentioned, if the regulator expresses in burst, it causes bursts in the regulated gene expression. This reduces for some times the levels of expression and consequently, it reduces the mean of molecules.
6. Burst expression is associated with a lower expression than the continuous expression. But burst expression does not necessarily mean high noise, it is possible to have burst expression and a noise level lower than the noise for a constitutive expression.
7. The effectiveness of regulation depends on the continuity of the regulator expression and it is independent of the regulator noise (FF). Therefore, the regulation of the genes in development could happen correctly despite their noise. This means the discontinuous burst expression must be regulated. Additionally, the advantage of more continuous expression is not to reduce the noise but to improve the effectiveness of regulation.

Section 5: Dynamical properties of the regulatory systems for coupled cells

1. The diffusion of a paracrine signal throughout a colony of cells decreases the noise (FF) in the whole range of parameters' values evaluated. This reduction depends on the parameters' values. For instance, the highest noise reduction occurs in the region of highest noise and the lowest in the region of lowest noise. This last could mean that there is a minimal noise that diffusion could not reduce.
2. Noise (FF) and the mean of molecules follow the same pattern at the multicellular level and at the unicellular level. Therefore, the four regions of expression identified in section 1 apply for both unicellular and multicellular levels.
3. The increase in diffusion rate decreases the noise (FF) in the four regions of expression. This is because diffusion increases the expression in those cells with low levels and decreases it in those with high levels. As a consequence, there is a decrease in the mean, an increase in the median, and a decrease in the maximal values of expression. However, the diffusion does not completely avoid the low expression levels and some cells remain with very low expression. This could affect the effectiveness of regulation because, as was pointed out in section 4, the effectiveness is lower when the regulator is expressed in burst with periods of low expression.
4. The colony size has a threshold above which the changes in colony size do not affect the expression levels, noise (FF), and noise's entropy. Below this threshold, the production, although high, is lower than for the other colony size and it has low noise (FF) and noise's entropy. This could drive a good community effect and differentiation but further analysis is necessary to prove this.

SUPPLEMENTARY INFORMATION

Parameter	Symbol
Promoter active time	$T_{\text{'on'}}$
Promoter inactive time	$T_{\text{'off'}}$
Promoter activation rate	k_{on}
Promoter inactivation rate	k_{off}
Transcription rate	ks_{mRNA}
mRNA degradation rate	kd_{mRNA}
Translation rate	ks_p
Protein degradation rate	kd_p
Diffusion	D

Table S1. Symbols of the parameters used in this study.

kon value	Koff value	Gen/cell type	The function of the gene	Ref.
0.023- 0.033	0.26- 0.35	<i>Bmal1a</i> promoter in Mouse fibroblast single-cell	<i>It generates molecular circadian rhythms and it is the only clock gene without which the circadian clock fails to function in humans (wiki)</i>	[25]
0.018-0. 032	0.16- 0.28	<i>Glutaminase</i> promoter in Mouse fibroblast single-cell	<i>It generates glutamate from glutamine. Glutamate is the most abundantly used excitatory neurotransmitter in the CNS (Wikipedia)</i>	[25]
0.0225-0 .06	0.48- 0.58	<i>Plectin1</i> promoter of Mouse fibroblast single-cell	<i>It is implied in maintaining cell and tissue integrity, and it is a scaffolding platform for the assembly,</i>	[25]

			<i>positioning, and regulation of signaling complexes</i> (https://www.ncbi.nlm.nih.gov/gene/5339)	
0.023-0.045	0.22-0.36	<i>Serpine1</i> promoter of Mouse fibroblast single-cell	<i>It is an inhibitor of fibrinolysis and also functions as a component of innate antiviral immunity</i> (https://www.ncbi.nlm.nih.gov/gene/5054)	[25]
0.021-0.047	0.19-0.27	<i>Sh3kbp1</i> promoter of Mouse fibroblast single-cell	<i>It facilitates protein-protein interactions and has been implicated in numerous cellular processes including apoptosis, cytoskeletal rearrangement, cell adhesion, and in the regulation of clathrin-dependent endocytosis</i> (https://www.ncbi.nlm.nih.gov/gene/30011)	[25]
0.018-0.032	0.1-0.15	<i>Prl2C2</i> promoter of Mouse fibroblast single-cell		[25]
0.25 - 0.32	0.04 - 0.06	<i>Ush gene promoter in Drosophila embryos at the onset of nuclear cleavage cycle 14 (nc14) with high BMP level</i>	<i>A gradient of bone morphogenetic protein (BMP) signaling patterns ectodermal cell fates along the dorsal-ventral axis of vertebrate and invertebrate embryos. The BMP/pMad gradient activates different thresholds of gene activity, including the intermediate target u-shaped (ush)</i>	[36]
0.19-0.28	0.04 - 0.06	<i>Ush gene promoter in Drosophila embryos at the onset of nuclear cleavage cycle 14 (nc14) with medium BMP level</i>	---	[36]
0.07 - 1.7	0.04 - 0.06	<i>Ush gene promoter in Drosophila embryos at the onset of nuclear cleavage cycle 14 (nc14) with low BMP level</i>	---	[36]
0.46 - 0.5	1.9- 2.1	<i>hnt gene promoter in Drosophila embryos at the onset of nuclear cleavage cycle 14 (nc14) with high BMP level</i>	<i>A gradient of bone morphogenetic protein (BMP) signaling patterns ectodermal cell fates along the dorsal-ventral axis of vertebrate and invertebrate embryos. The BMP/pMad gradient activates different thresholds of gene activity, including the peak target gene hindsight (hnt)</i>	[36]
0.3 - 0.4	1.9- 2.1	<i>hnt gene promoter in Drosophila embryos at the onset of nuclear cleavage cycle 14 (nc14) with medium BMP level</i>	---	[36]
0,01158	2.082	For none in specific	<i>Their simulation parameters are all within biologically or biochemically relevant range, which usually spans orders of magnitude, such as the one for protein degradation rates. Where possible, they have referred to reaction rates that have been determined experimentally to choose parameter values for simulations</i>	[73]
0.045	0.005	<i>Human cyclin D1 (CCND1) in HEK-293 cells</i>	<i>Different cyclins exhibit distinct expression and degradation patterns which contribute to the temporal coordination of each mitotic event.</i>	[26]
	0.2 (0, 0.05)	For none in specific	---	[27][25]

1.0	1.0	NODAL	They indicate neither the dimensions nor the source of this date. It is an entirely theoretic study.	[74]
1.0	1.0	LEFTY	They indicate neither the dimensions nor the source of this date. It is an entirely theoretic study. They made a bifurcation diagram for k_{on} parameter between 0-1.4	[74]

Table S2. Activation and Inactivation promoter rate (k_{on} - 1/min, k_{off} - 1/min)

Value	Gen/cell type	The function of the gene	Ref.
0.1-0.28	<i>Plectin1</i> promoter of Mouse fibroblast single-cell	---	[25]
0.2 - 0.3	<i>Bmal1a</i> promoter in Mouse fibroblast single-cell	---	[25]
0.8 - 1.2	<i>Serpine1</i> promoter of Mouse fibroblast single-cell	---	[25]
1.8 - 2.0	<i>Sh3kbp1</i> promoter of Mouse fibroblast single-cell	---	[25]
2.0 - 3.0	<i>Glutaminase</i> promoter in Mouse fibroblast single-cell	---	[25]
3.4 - 4.3	<i>Prl2C2</i> promoter of Mouse fibroblast single-cell	---	[25]
0.31-0.78 kb/min	<i>Human cyclin D1 (CCND1)</i> in HEK-293 cells	---	[26]
9 - 15 Poll II / min (loading rate)	<i>Ush gene in Drosophila embryos at the onset of nuclear cleavage cycle 14 (nc14), it does not depend of BMP level</i>	---	[36]
37 - 45 Poll II / min (loading rate)	<i>hnt gene in Drosophila embryos at the onset of nuclear cleavage cycle 14 (nc14), it does not depend of BMP level</i>	---	[36]
0.696	Non-specific	It is for a gene making part of the community effect subcircuit	[73]

Table S3. Transcription rate (ks_{mRNA} - mRNAs/min)

Value	Gen/cell type	The function of the gene	Ref.
0,02082	Non-specific	It is for a gene making part of the community effect subcircuit	[73]
0.0083	NODAL	---	[70]
0.0041	LEFTY	---	[70]

Table S4. mRNA degradation rate (kd_{mRNA} - mRNAs/min)

Value	Gen/cell type	The function of the gene	Ref.
23 - 29	<i>Ush gene promoter in Drosophila embryos at the onset of nuclear cleavage cycle 14 (nc14) with high BMP level</i>	---	[36]
15 - 25	<i>Ush gene promoter in Drosophila embryos at the onset of nuclear cleavage cycle 14 (nc14) with medium BMP level</i>	---	[36]
5 - 16	<i>Ush gene promoter in Drosophila embryos at the onset of nuclear cleavage cycle 14 (nc14) with low BMP level</i>	---	[36]

Table S5. mRNA transcriptional burst size ($b_{mRNA} = ks_{mRNA} * T'_{on'} = ks_{mRNA} / Koff$)

Value	Gen/cell type	Function of the gene	Ref.
1.386	Non specific	It is for a gene making part of the community effect subcircuit	[73]
2 (proteins/min/mRNA)	Nodal	---	[70]
2 (proteins/min/mRNA)	Lefty	---	[70]

Table S6. Translation rate (ks_p - protein/min)

Value	Gen/cell type	The function of the gene	Ref.
0.02082	Non-specific	It is for a gene making part of the community effect subcircuit	[73]
0.1	Nodal	They indicate neither the dimensions nor the source of this date. It is an entirely theoretic study.	[74]
0.1	Lefty	They indicate neither the dimensions nor the source of this date. It is an entirely theoretic study.	[74]

Table S7. Protein degradation rate (kd_p - protein/min)

Value	Gen/cell type	The function of the gene	Ref.
$0.7 \pm 0.2 \mu m^2/s$ (Cyclops-GFP) $3.2 \pm 0.5 \mu m^2/s$ (Squint-GFP)	Nodal	---	[91]
$11.1 \pm 0.6 \mu m^2/s$ (Lefty1-GFP) $18.9 \pm 3.0 \mu m^2/s$ (Lefty2-GFP)	Lefty	---	[91]

Table S8. Extracellular Effective diffusion coefficients

Parameter	Symbol	Unity	Constitutive	Unregulated	Auto-activated	Activator-Inhibitor
-----------	--------	-------	--------------	-------------	----------------	---------------------

Promoter activation rate	k_{on}	1/min	---	0,01158	0,01158	0,01158
Promoter inactivation rate	k_{off}	1/min	---	2,08	2,08	2,08
Transcription rate	ks_{mRNA}	mRNA/min	0,696	0,696	0,696	0,696
mRNA burst size	b_m	mRNAs	---	0,333	0,333	0,333
mRNA degradation rate	kd_{mRNA}	mRNA/min	0,02082	0,02082	0,02082	0,02082
Translation rate	ks_p	protein/min	1,386	1,386	1,386	1,386
Protein burst size	b_p	protein/mRNA	---	66,57	66,57	66,57
Protein degradation rate	kd_p	protein/min	0,02082	0,02082	0,02082	0,02082
Auto-activation constant	k_{AA}	protein	---	---	100	100
Inhibition constant	k_{AI}	protein	---	---	---	100
Inhibition constant	k_{AI}	protein	---	---	---	100
Hill constant	h	---	---	---	3	3
Diffusion activator	Da	pix^2/mi n	---	---	0.01	----
Diffusion inhibitor	Di	pix^2/mi n	---	---	---	----

Table S9. Default parameters values of the regulatory systems evaluated in this study. The majority of these were from [73]. These parameters apply both for activator (A) and inhibitor (I). ° 1 pixel ~ 16 μm. mRNA and protein are molecules.

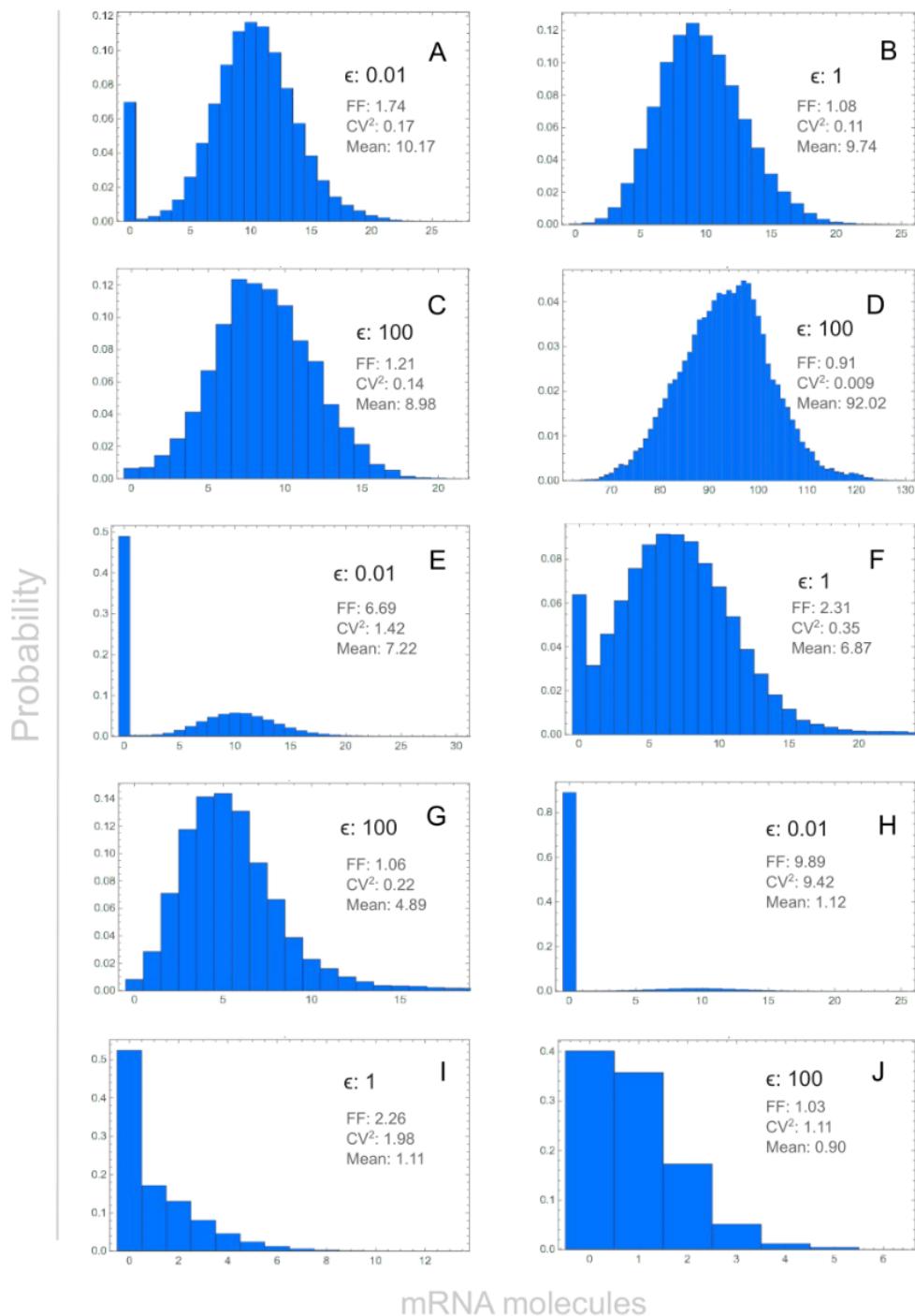
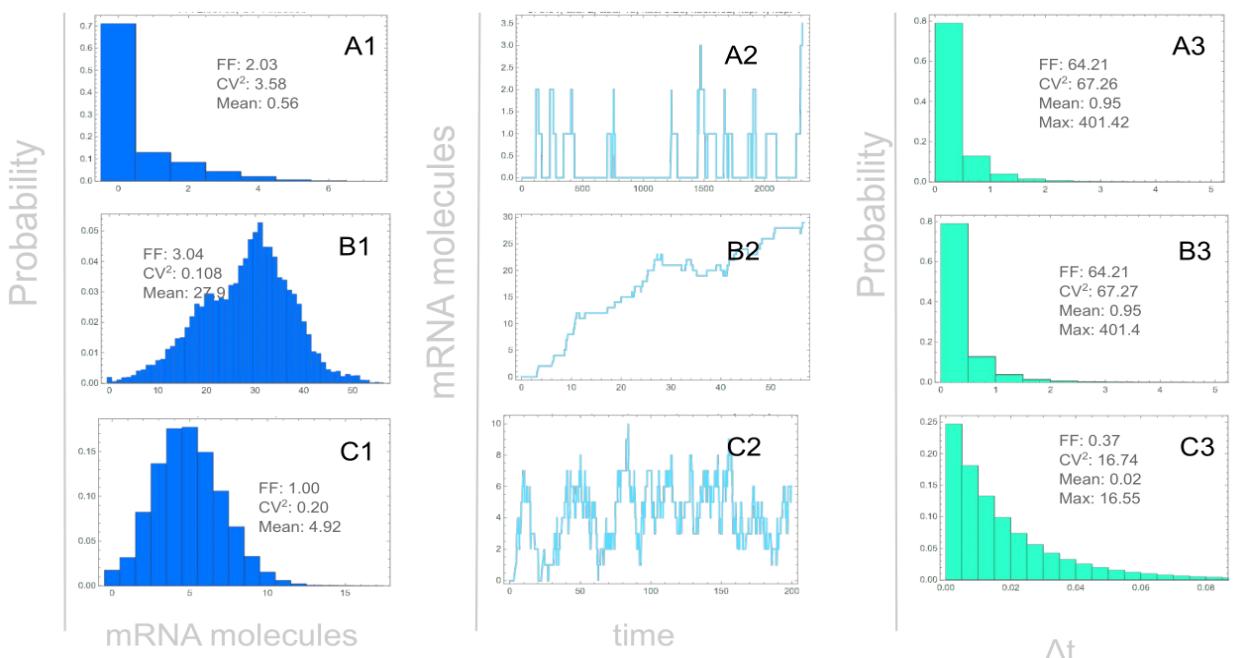
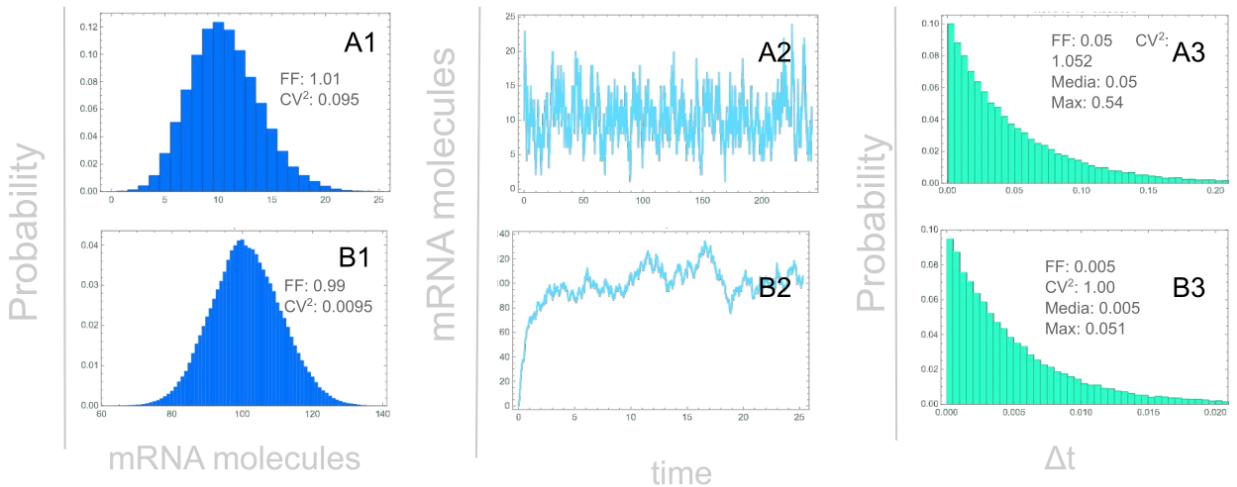


Figure S1. Steady-state distribution of one gene with noise simulated with NGM. A-J. They show the FF, CV^2 , and mean of mRNA molecules for each distribution. A-C. With k_{on} of 10, k_{off} of 1, ks_{mRNA} of 10, and kd_{mRNA} of 1, D. With k_{on} of 10, k_{off} of 1, ks_{mRNA} of 100, and kd_{mRNA} of 1, E-G. With k_{on} of 1, k_{off} of 1, ks_{mRNA} of 10, and kd_{mRNA} of 1, H-J. With k_{on} of 1, k_{off^*} of 10, ks_{mRNA} of 10, and kd_{mRNA} of 1. The promoter velocity value is indicated for each plot (ϵ of 100 (high), 1 (medium), and 0.1 (low)).



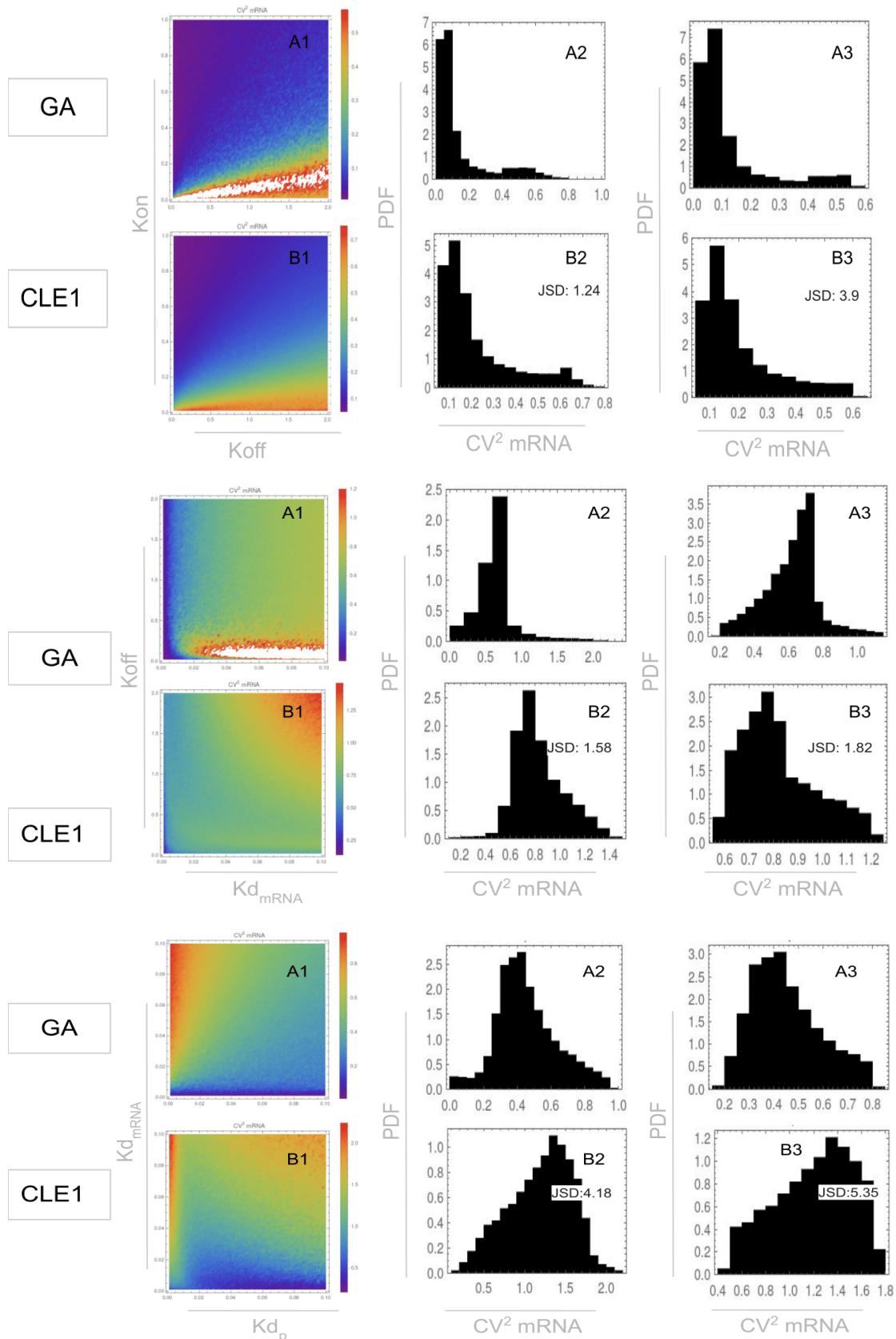


Figure S4. Noise (CV^2) in mRNA expression of self-activated gene for a range of some parameters values. TOP: k_{on} vs k_{off} , MIDDLE: k_{off} vs $k_{d,mRNA}$, BOTTOM: $k_{d,mRNA}$ vs $k_{d,p}$, A. Simulated with NGM, B. Simulated with the CLE for burst only in mRNA (CLE1), 1. CV^2 for different parameters values in the same scale (white patches are values outside of the plot range), 2. Probability Density Function (PDF) for all the values of CV^2 , 3. The same PDF as in 2 but for the interquartile range 0.05 to 0.95 of CV^2 values. JSD: Jensen–Shannon divergence between B and A PDFs.

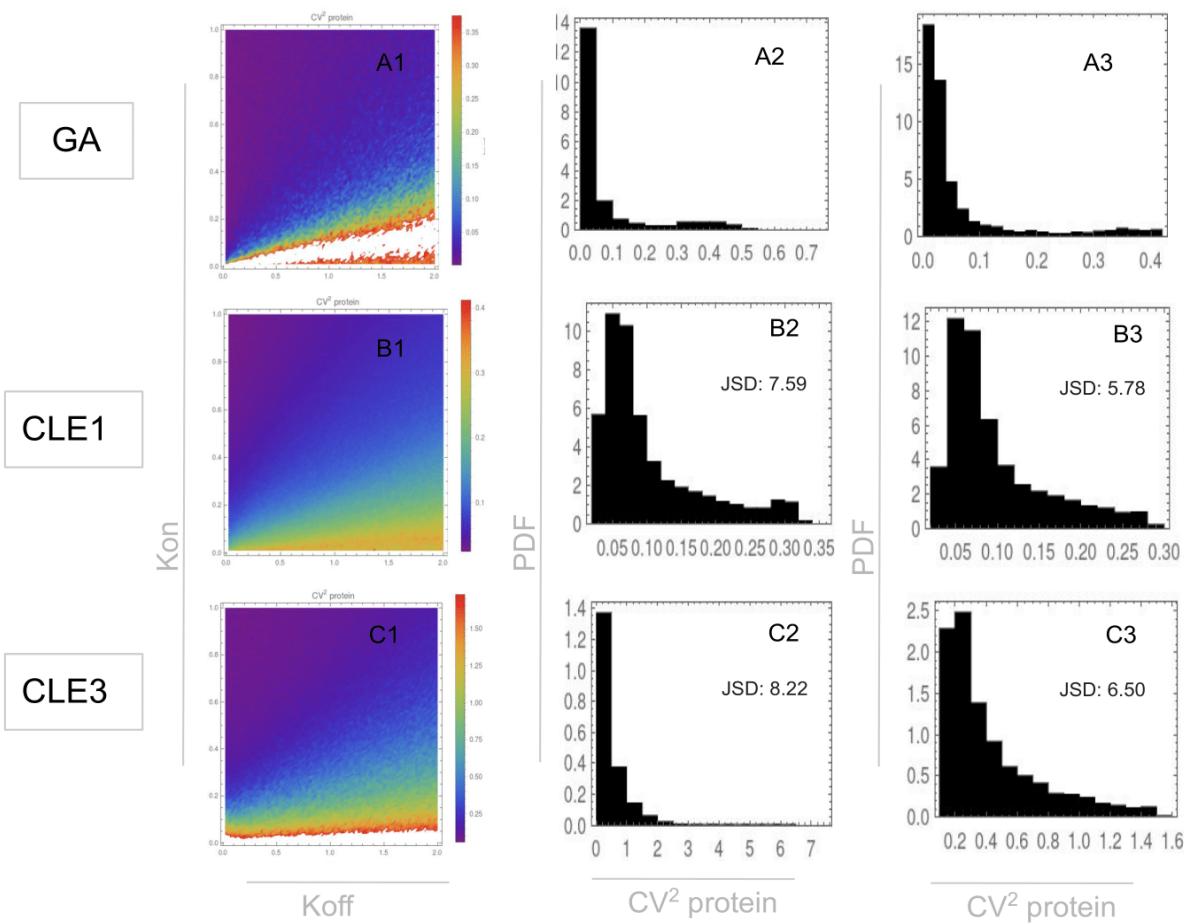


Figure S5. Noise (CV^2) in protein expression of the self-activated gene for a range of k_{on} and k_{off} values. A. Simulated with NGM, B. Simulated with the CLE for burst only in mRNA (CLE1), C. Simulated with the CLE for burst in both mRNA and protein (CLE3), 1. CV^2 for different parameters values in the same scale (white patches are values outside of the plot range), 2. Probability Density Function (PDF) for all the the values of CV^2 , 3. The same PDF as in 2 but for the interquartile range 0.05 to 0.95 of CV^2 values. JSD: Jensen–Shannon divergence between B and A PDFs.

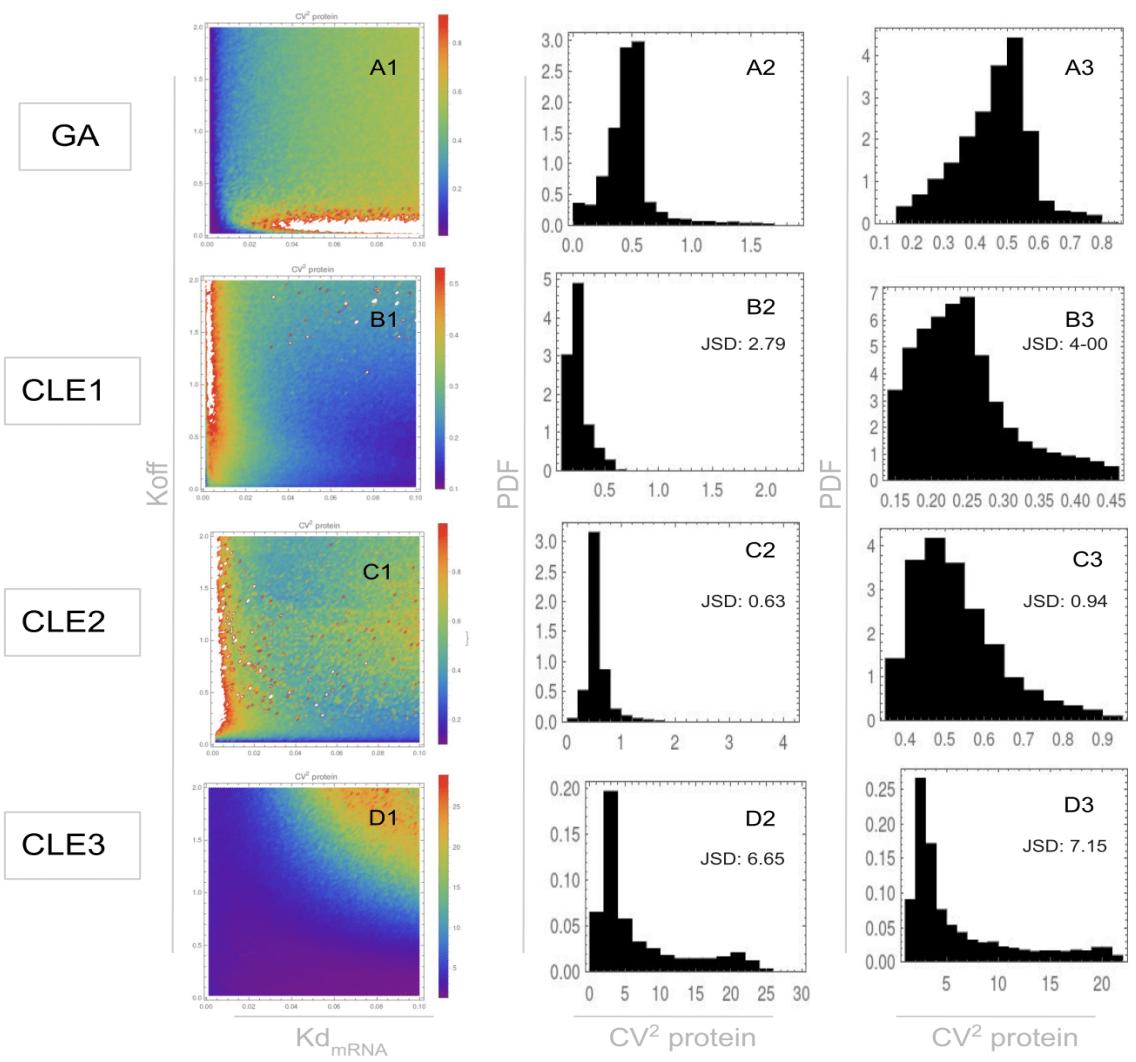


Figure S6. Noise (CV^2) in protein expression of self-activated gene for a range of k_{off} and kd_{mRNA} values. A. Simulated with NGM, B. Simulated with the CLE for burst only in mRNA (CLE1), C. Simulated with the CLE for burst only in protein (CLE2), D. Simulated with the CLE for burst in both mRNA and protein (CLE3), 1. CV^2 for different parameters values in the same scale (white patches are values outside of the plot range), 2. Probability Density Function (PDF) for all the values of CV^2 , 3. The same PDF as in 2 but for the interquartile range 0.05 to 0.95 of CV^2 values. JSD: Jensen–Shannon divergence between B and A PDFs.

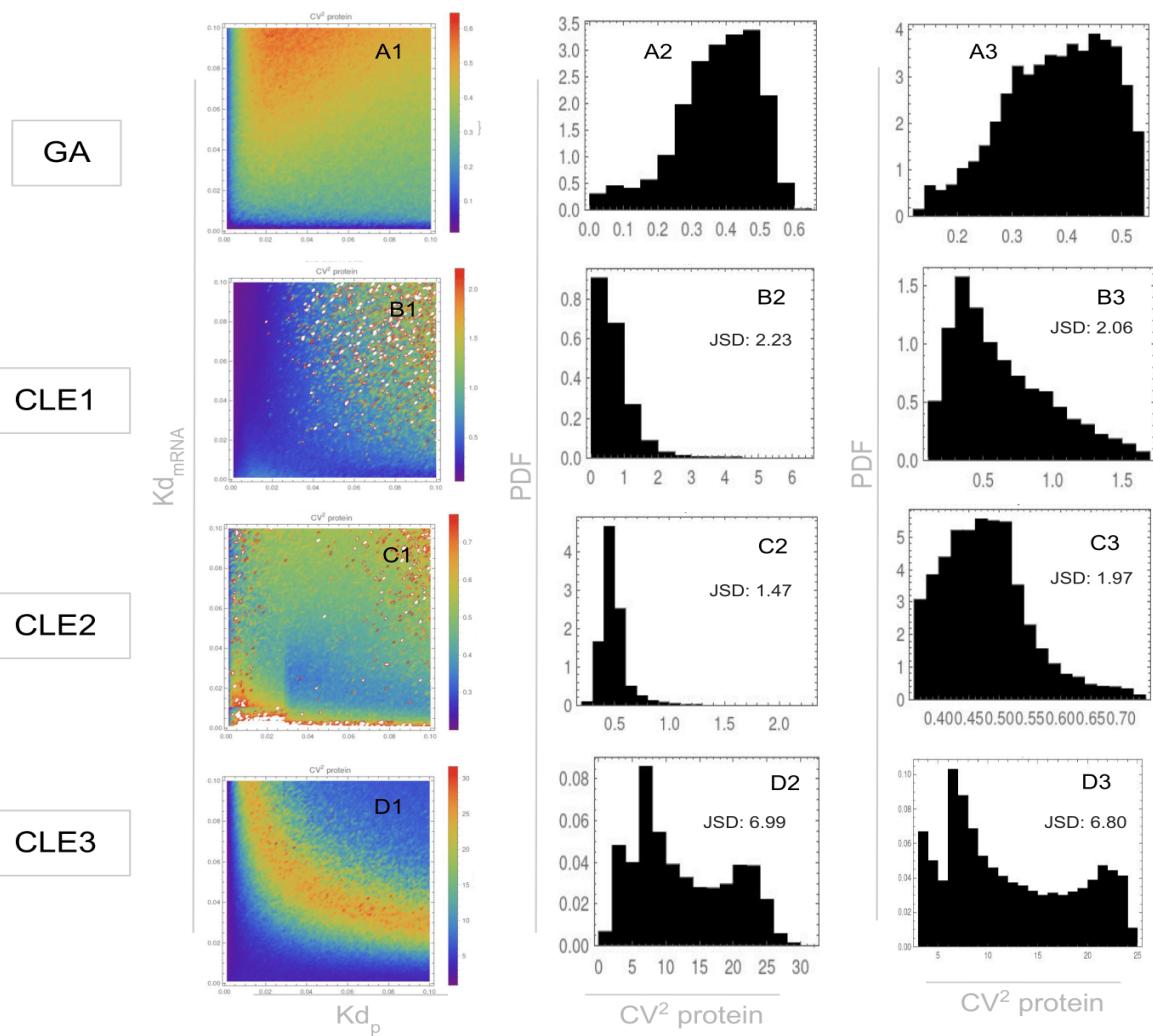


Figure S7. Noise (CV^2) in protein expression of self-activated gene for a range of kd_{mRNA} and kd_p values. A. Simulated with NGM, B. Simulated with the CLE for burst only in mRNA (CLE1), C. Simulated with the CLE for burst only in protein (CLE2), D. Simulated with the CLE for burst in both mRNA and protein (CLE3). 1. CV^2 for different parameters values in the same scale (white patches are values outside of the plot range), 2. Probability Density Function (PDF) for all the values of CV^2 , 3. The same PDF as in 2 but for the interquartile range 0.05 to 0.95 of FFs values. JSD: Jensen–Shannon divergence between B and A PDFs.

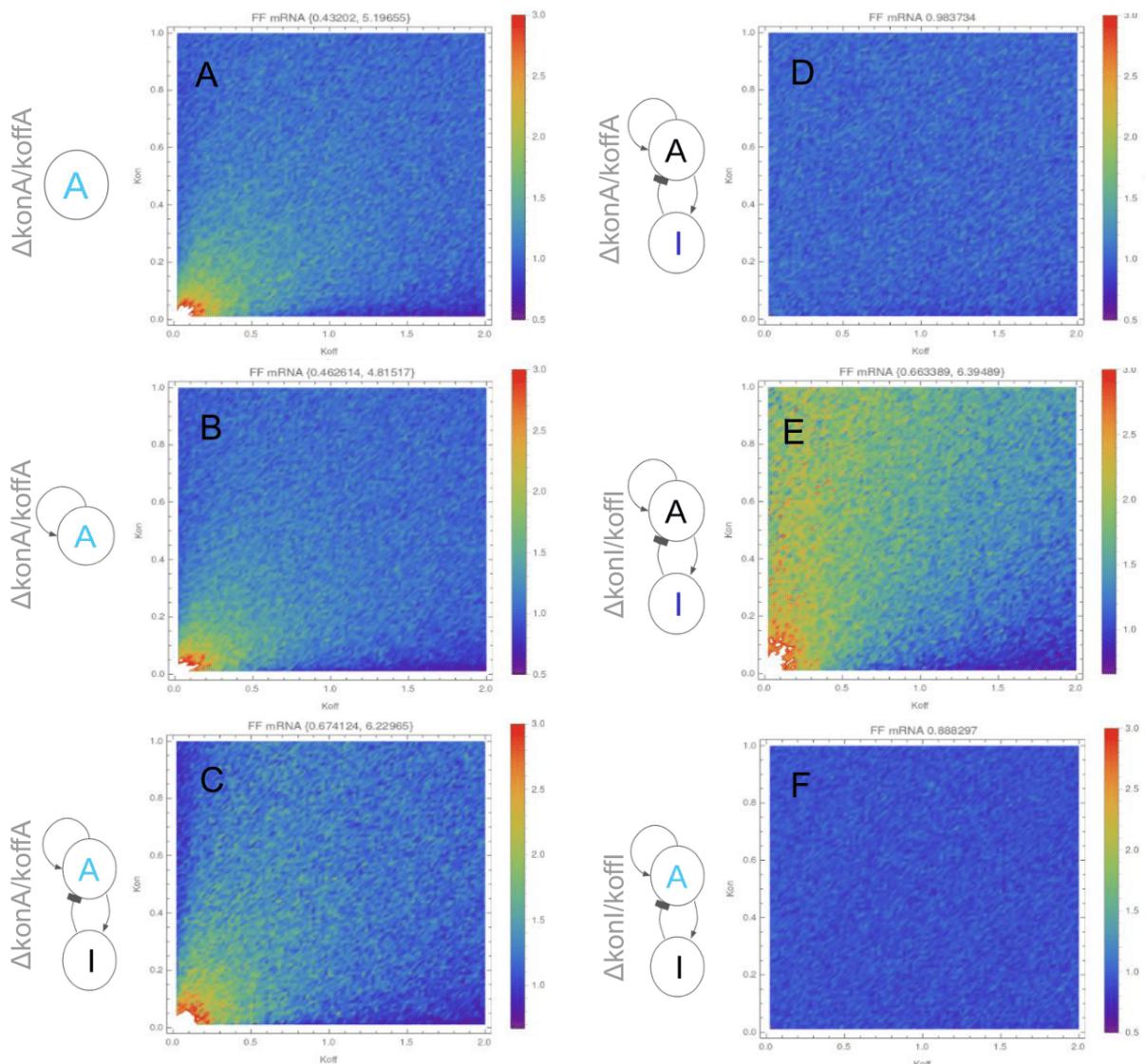


Figure S8. Noise (FF) in mRNA expression of the regulatory systems evaluated for a range of koff and kon parameter values. A. An unregulated gene without explicit regulation, B. self-activated gene, C. A gene of the Activator-Inhibitor regulatory system when it is changed the parameters of A, D. *I* gene of the Activator-Inhibitor regulatory system when it is changed the parameters of A, E. *I* gene of the Activator-Inhibitor GRN when it is changed the parameters of *I*, F. A gene of the Activator-Inhibitor GRN when it is changed the parameters of *A*. The values over the plots are the minimum and maximum or the mean (when there is only one value). *Simulated with NGM and the default parameters in Table S9.

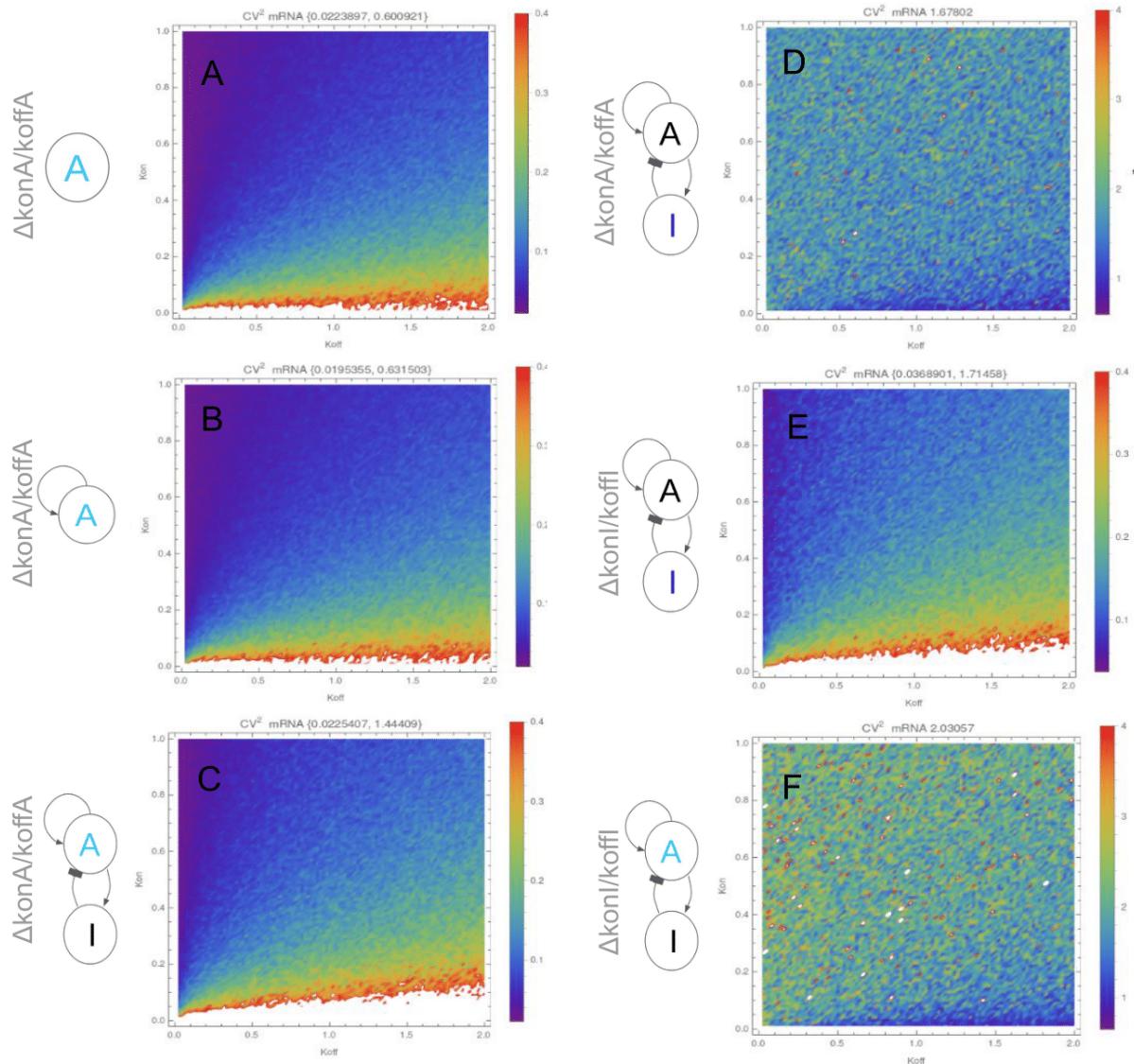


Figure S9. Noise (CV^2) in mRNA expression of the regulatory systems evaluated for a range of $koff$ and kon parameter values. A. An unregulated gene without explicit regulation, B. self-activated gene, C. A gene of the Activator-Inhibitor regulatory system when it is changed the parameters of A, D. I gene of the Activator-Inhibitor regulatory system when it is changed the parameters of A, E. I gene of the Activator-Inhibitor GRN when it is changed the parameters of I , F. A gene of the Activator-Inhibitor GRN when it is changed the parameters of I . The values over the plots are the minimum and maximum or the mean (when there is only one value). *Simulated with NGM and the default parameters in Table S9.

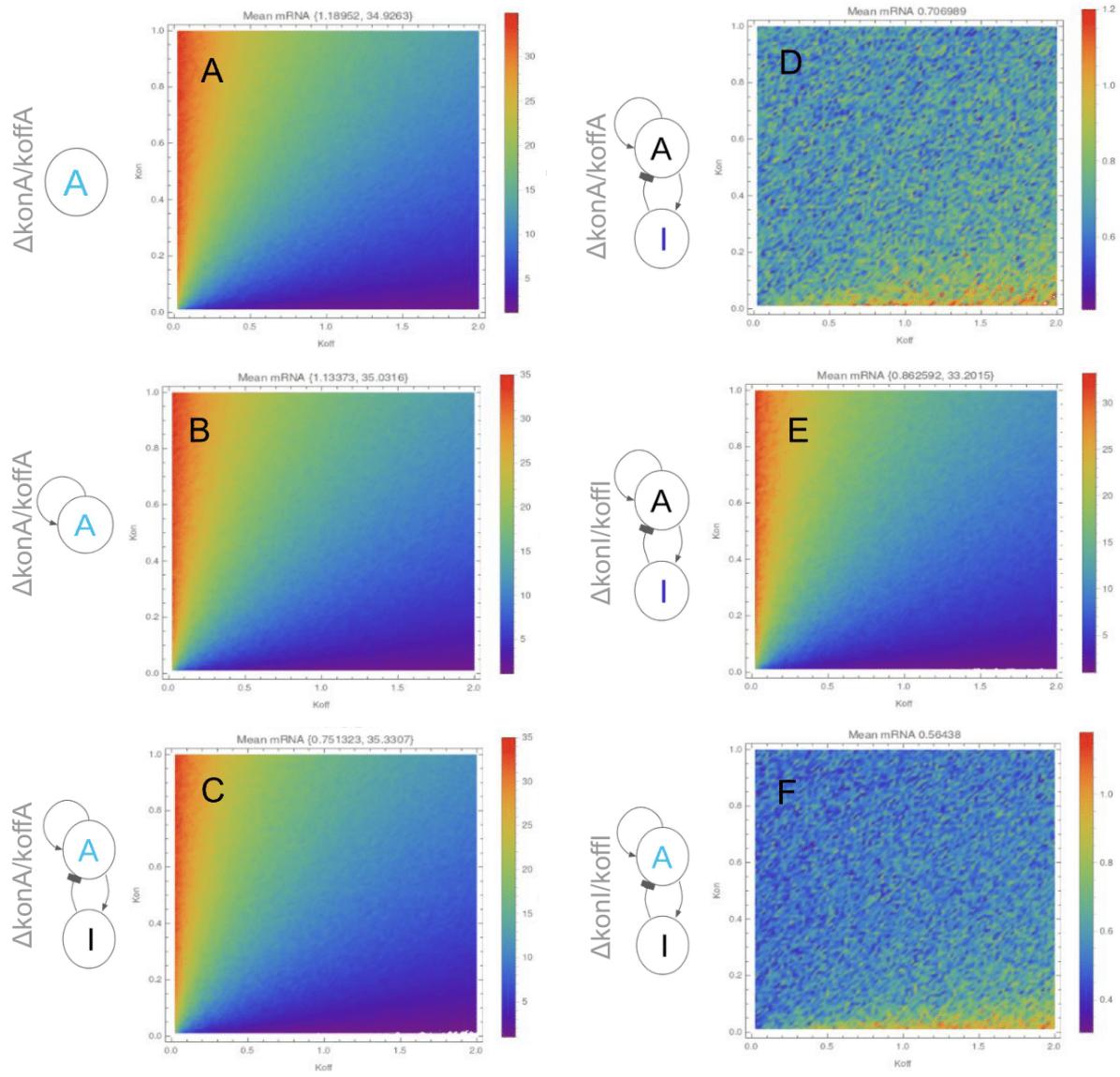


Figure S10. Mean of the protein expression of the regulatory systems evaluated for a range of k_{off} and k_{on} parameter values. A. An unregulated gene without explicit regulation, B. self-activated gene, C. A gene of the Activator-Inhibitor regulatory system when it is changed the parameters of A, D. I gene of the Activator-Inhibitor regulatory system when it is changed the parameters of A, E. I gene of the Activator-Inhibitor GRN when it is changed the parameters of I, F. A gene of the Activator-Inhibitor GRN when it is changed the parameters of I. The values over the plots are the minimum and maximum or the mean (when there is only one value). *Simulated with NGM and the default parameters in Table S9.

	One UnR	Region 2: 2 / 0.05		Region 1: 0.2 / 0.05	
		AB-Inh ΔA	AB-Activ ΔA	AB-Inh ΔA	AB-Activ ΔA
ff mRNA mean	0.58056	0.881078	0.881778	0.886391	0.973127
ff mRNA p-value		3.61043×10^{-65}	2.17028×10^{-66}	1.40931×10^{-78}	3.74532×10^{-89}
ff protein mean	18.6634	27.8423	26.8638	26.6676	28.8946
ff protein p-value		4.71534×10^{-36}	1.10156×10^{-33}	1.33103×10^{-33}	3.4124×10^{-41}
cv2 mRNA mean	0.413467	1.25589	1.17242	2.17629	1.67721
cv2 mRNA p-value		2.46473×10^{-135}	2.32398×10^{-136}	3.37747×10^{-160}	8.35052×10^{-155}
cv2 protein mean	0.239048	0.612968	0.57532	0.910483	0.704602
cv2 protein p-value		2.50675×10^{-105}	7.69888×10^{-97}	5.59265×10^{-142}	8.91567×10^{-123}
mean mRNA mean	1.40306	0.850672	0.892717	0.548221	0.721355
mean mRNA p-value		2.8863×10^{-95}	4.18086×10^{-86}	2.45385×10^{-153}	8.69044×10^{-126}
mean protein mean	77.7815	52.2879	54.1387	35.7425	46.6924
mean protein p-value		2.1333×10^{-60}	6.57189×10^{-53}	4.22959×10^{-129}	5.55217×10^{-85}

	One UnR	Region 4: 2 / 1		Region 3: 0.2 / 1	
		AB-Inh ΔA	AB-Activ ΔA	AB-Inh ΔA	AB-Activ ΔA
ff mRNA mean	0.58056	0.905395	0.964227	0.939738	0.982851
ff mRNA p-value		1.45894×10^{-80}	1.40397×10^{-89}	5.2504×10^{-90}	1.98119×10^{-92}
ff protein mean	18.6634	26.5037	29.1396	28.0245	29.1178
ff protein p-value		4.49922×10^{-30}	5.42299×10^{-43}	2.46276×10^{-39}	3.1628×10^{-45}
cv2 mRNA mean	0.413467	2.37962	1.56273	2.62344	1.72364
cv2 mRNA p-value		1.15441×10^{-159}	2.34347×10^{-158}	2.86362×10^{-162}	4.57944×10^{-158}
cv2 protein mean	0.239048	0.94881	0.684547	1.05619	0.741981
cv2 protein p-value		1.49618×10^{-141}	2.65196×10^{-129}	2.57385×10^{-152}	5.2967×10^{-131}
mean mRNA mean	1.40306	0.507813	0.708738	0.47698	0.67264
mean mRNA p-value		5.69097×10^{-158}	1.1135×10^{-137}	3.10908×10^{-160}	3.65506×10^{-139}
mean protein mean	77.7815	33.3741	46.2747	31.5665	44.0448
mean protein p-value		1.6961×10^{-138}	4.69738×10^{-93}	1.00983×10^{-145}	8.73366×10^{-102}

Table S10. Comparison of FF, CV2, and mean of molecules between an unregulated gene and regulated gene, when it is changed the parameters' values of the regulator. For each one of the four regions, it is shown the estimates (ff or cv2) and p-value of a T-test between both an Inhibited (AB-Inh ΔA) and Activated (AB-Activ ΔA) gene and an unregulated gene (One UnR), when it is changed the parameters' values of the regulator.

	Region 2: 2 / 0.05			Region 1: 0.2 / 0.05		
	One UnR	AB-Inh ΔB	AB-Activ ΔB	One UnR	AB-Inh ΔB	AB-Activ ΔB
ff mRNA mean	0.76479	0.924827	0.906001	2.19993	2.45642	2.42693
ff mRNA p-value		7.30615×10^{-20}	1.05821×10^{-14}		2.47049×10^{-7}	0.0000258117
ff protein mean	23.5258	28.3331	28.647	69.9036	78.938	79.5733
ff protein p-value		2.03576×10^{-10}	2.8436×10^{-8}		0.000201897	0.000215796
cv2 mRNA mean	0.365919	0.5939	0.570127	0.239866	0.306645	0.294591
cv2 mRNA p-value		1.11136×10^{-67}	1.38966×10^{-56}		1.60802×10^{-31}	1.06306×10^{-22}
cv2 protein mean	0.192034	0.298137	0.30039	0.121298	0.156925	0.155593
cv2 protein p-value		3.76441×10^{-39}	1.06403×10^{-32}		6.90153×10^{-17}	6.06589×10^{-15}
mean mRNA mean	2.11122	1.66806	1.70265	9.17688	8.15418	8.35582
mean mRNA p-value		8.32075×10^{-42}	6.44096×10^{-36}		2.69832×10^{-20}	1.78455×10^{-14}
mean protein mean	123.238	100.601	101.335	571.583	507.838	517.744
mean protein p-value		1.62666×10^{-28}	5.67605×10^{-27}		5.0432×10^{-20}	1.3955×10^{-15}

	Region 4: 2 / 1			Region 3: 0.2 / 1		
	One UnR	AB-Inh ΔB	AB-Activ ΔB	One UnR	AB-Inh ΔB	AB-Activ ΔB
ff mRNA mean	0.989938	1.18658	1.27562	0.956053	1.11032	1.51159
ff mRNA p-value		4.26029×10^{-15}	4.82942×10^{-33}		6.25476×10^{-10}	9.55942×10^{-51}
ff protein mean	30.7868	40.6901	46.7937	29.8583	37.4181	58.3129
ff protein p-value		3.28559×10^{-16}	3.35761×10^{-35}		4.76479×10^{-9}	1.06999×10^{-42}
cv2 mRNA mean	0.0800389	0.106948	0.112754	0.0329404	0.0400691	0.0561441
cv2 mRNA p-value		3.30234×10^{-32}	4.53385×10^{-50}		7.42942×10^{-15}	2.87576×10^{-56}
cv2 protein mean	0.0380981	0.0561885	0.0640566	0.0155553	0.0204915	0.0332442
cv2 protein p-value		9.27566×10^{-28}	4.516×10^{-47}		4.74258×10^{-12}	4.40289×10^{-47}
mean mRNA mean	12.3977	11.2426	11.4658	29.0345	27.9246	27.4816
mean mRNA p-value		1.20709×10^{-42}	9.34299×10^{-28}		7.05848×10^{-21}	2.11978×10^{-33}
mean protein mean	808.734	734.001	742.966	1915.72	1842.75	1806.53
mean protein p-value		3.49683×10^{-42}	1.81451×10^{-30}		1.70136×10^{-20}	2.85814×10^{-35}

Table S11. Comparison of FF, CV2, and mean of molecules between an unregulated gene and regulated gene, when it is changed the parameters' values of the regulated gene. For each one of the four regions, it is shown the estimates (ff or cv2) and p-value of a T-test between both an Inhibited (AB-Inh ΔB) or Activated (AB-Activ ΔB) gene and an unregulated gene (One UnR), when it is changed the parameters' values of the regulated gene.

Gif S1. Region1_size25_diffusion0.02. Expression levels of a self-activated gene with diffusion in a colony simulated with the CLE1.

Gif S2. Region1_size25_diffusion0.1. Expression levels of a self-activated gene with diffusion in a colony simulated with the CLE1.

Gif S3. Region2_size25_diffusion0.02. Expression levels of a self-activated gene with diffusion in a colony simulated with the CLE1.

Gif S4. Region2_size25_diffusion0.1. Expression levels of a self-activated gene with diffusion in a colony simulated with the CLE1.

Gif S5. Region3_size10_diffusion0.02. Expression levels of a self-activated gene with diffusion in a colony simulated with the CLE1.

Gif S6. Region3_size10_diffusion0.1. Expression levels of a self-activated gene with diffusion in a colony simulated with the CLE1.

Gif S7. Region3_size25_diffusion0.02. Expression levels of a self-activated gene with diffusion in a colony simulated with the CLE1.

Gif S8. Region3_size25_diffusion0.1. Expression levels of a self-activated gene with diffusion in a colony simulated with the CLE1.

Gif S9. Region4_size2_diffusion0.02. Expression levels of a self-activated gene with diffusion in a colony simulated with the CLE1.

[Gif S10. Region4_size2_diffusion0.1.](#) Expression levels of a self-activated gene with diffusion in a colony simulated with the CLE1.

REFERENCES

1. Elowitz MB. Stochastic Gene Expression in a Single Cell. *Science*. 2002;297: 1183–1186.
2. Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol*. 2006;4: e309.
3. Bengtsson M, Ståhlberg A, Rorsman P, Kubista M. Gene expression profiling in single cells from the pancreatic islets of Langerhans reveals lognormal distribution of mRNA levels. *Genome Res*. 2005;15: 1388–1392.
4. Rosenfeld N. Gene Regulation at the Single-Cell Level. *Science*. 2005. pp. 1962–1965. doi:10.1126/science.1106914
5. Zhong JF, Chen Y, Marcus JS, Scherer A, Quake SR, Taylor CR, et al. A microfluidic processor for gene expression profiling of single human embryonic stem cells. *Lab Chip*. 2008. pp. 68–74. doi:10.1039/b712116d
6. Jiang Y, Zhang NR, Li M. Modeling allele-specific gene expression by single-cell RNA sequencing. doi:10.1101/109629
7. Hensel Z, Feng H, Han B, Hatem C, Wang J, Xiao J. Stochastic expression dynamics of a transcription factor revealed by single-molecule noise analysis. *Nat Struct Mol Biol*. 2012;19: 797–802.
8. Hoppe C, Bowles JR, Minchington TG, Sutcliffe C, Upadhyai P, Rattray M, et al. Modulation of the Promoter Activation Rate Dictates the Transcriptional Response to Graded BMP Signaling Levels in the Drosophila Embryo. *Dev Cell*. 2020;54: 727–741.e7.
9. Yunger S, Rosenfeld L, Garini Y, Shav-Tal Y. Single-allele analysis of transcription kinetics in living mammalian cells. *Nature Methods*. 2010;7: 631–633.
10. Swain PS, Elowitz MB, Siggia ED. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*. 2002. pp. 12795–12800. doi:10.1073/pnas.162041399
11. Han R, Huang G, Wang Y, Xu Y, Hu Y, Jiang W, et al. Increased gene expression noise in human cancers is correlated with low p53 and immune activities as well as late stage cancer. *Oncotarget*. 2016. pp. 72011–72020. doi:10.18632/oncotarget.12457
12. Volfson D, Marciniak J, Blake WJ, Ostroff N, Tsimring LS, Hasty J. Origins of extrinsic variability in eukaryotic gene expression. *Nature*. 2006;439: 861–864.
13. Bowsher CG, Swain PS. Identifying sources of variation and the flow of information in biochemical networks. *Proc Natl Acad Sci U S A*. 2012;109: E1320–8.
14. Iyer-Biswas S, Hayot F, Jayaprakash C. Stochasticity of gene products from transcriptional pulsing. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2009;79: 031911.

15. Voss TC, Hager GL. Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nature Reviews Genetics*. 2014. pp. 69–81. doi:10.1038/nrg3623
16. Hensel Z, Feng H, Han B, Hatem C, Wang J, Xiao J. Stochastic expression dynamics of a transcription factor revealed by single-molecule noise analysis. *Nat Struct Mol Biol*. 2012;19: 797–802.
17. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A. Regulation of noise in the expression of a single gene. *Nat Genet*. 2002;31: 69–73.
18. Blake WJ, KAErn M, Cantor CR, Collins JJ. Noise in eukaryotic gene expression. *Nature*. 2003;422: 633–637.
19. Iyer-Biswas S, Hayot F, Jayaprakash C. Stochasticity of gene products from transcriptional pulsing. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2009;79: 031911.
20. Franz K, Singh A, Weinberger LS. Lentiviral vectors to study stochastic noise in gene expression. *Methods Enzymol*. 2011;497: 603–622.
21. Munsky B, Neuert G, van Oudenaarden A. Using gene expression noise to understand gene regulation. *Science*. 2012;336: 183–187.
22. Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, et al. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature*. 2006;441: 840–846.
23. Fuda NJ, Ardehali MB, Lis JT. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature*. 2009;461: 186–192.
24. Kim J, Marioni JC. Inferring the kinetics of stochastic gene expression from single-cell RNA-sequencing data. *Genome Biology*. 2013. p. R7. doi:10.1186/gb-2013-14-1-r7
25. Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. Mammalian genes are transcribed with widely different bursting kinetics. *Science*. 2011;332: 472–474.
26. Yunger S, Rosenfeld L, Garini Y, Shav-Tal Y. Single-allele analysis of transcription kinetics in living mammalian cells. *Nature Methods*. 2010;7: 631–633.
27. Chubb JR, Trcek T, Shenoy SM, Singer RH. Transcriptional pulsing of a developmental gene. *Curr Biol*. 2006;16: 1018–1025.
28. Jiang Y, Zhang NR, Li M. Modeling allele-specific gene expression by single-cell RNA sequencing. doi:10.1101/109629
29. Dong P, Liu Z. Shaping development by stochasticity and dynamics in gene regulation. *Open Biology*. 2017;7: 170030.
30. Voigt C. Synthetic Biology, Part A: Methods for Part/Device Characterization and Chassis Engineering. Academic Press; 2011.
31. Suter DM, Molina N, Naef F, Schibler U. Origins and consequences of transcriptional discontinuity. *Curr Opin Cell Biol*. 2011;23: 657–662.
32. Shahrezaei V, Swain PS. Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences*. 2008. pp. 17256–17261. doi:10.1073/pnas.0803850105
33. Tabaka M, Kalwarczyk T, Szymanski J, Hou S, Holyst R. The effect of macromolecular

- crowding on mobility of biomolecules, association kinetics, and gene expression in living cells. *Frontiers in Physics*. 2014. doi:10.3389/fphy.2014.00054
34. Gillespie DT. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics*. 2001. pp. 1716–1733. doi:10.1063/1.1378322
 35. Alon U. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. CRC Press; 2006.
 36. Hoppe C, Bowles JR, Minchington TG, Sutcliffe C, Upadhyai P, Rattray M, et al. Modulation of the Promoter Activation Rate Dictates the Transcriptional Response to Graded BMP Signaling Levels in the Drosophila Embryo. *Dev Cell*. 2020;54: 727–741.e7.
 37. Fukaya T, Lim B, Levine M. Enhancer Control of Transcriptional Bursting. *Cell*. 2016;166: 358–368.
 38. 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 U S A*. 2020;117: 836–847.
 39. Senecal A, Munsky B, Proux F, Ly N, Braye FE, Zimmer C, et al. Transcription factors modulate c-Fos transcriptional bursts. *Cell Rep*. 2014;8: 75–83.
 40. Yan C-CS, Chepyala SR, Yen C-M, Hsu C-P. Efficient and flexible implementation of Langevin simulation for gene burst production. *Sci Rep*. 2017;7: 16851.
 41. Schwahnäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global quantification of mammalian gene expression control. *Nature*. 2011;473: 337–342.
 42. Li G-W, Sunney Xie X. Central dogma at the single-molecule level in living cells. *Nature*. 2011. pp. 308–315. doi:10.1038/nature10315
 43. Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, et al. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science*. 2002;298: 799–804.
 44. Erwin DH, Davidson EH. The evolution of hierarchical gene regulatory networks. *Nature Reviews Genetics*. 2009. pp. 141–148. doi:10.1038/nrg2499
 45. Fernández P, Solé RV. Neutral fitness landscapes in signalling networks. *J R Soc Interface*. 2007;4: 41–47.
 46. Oliveri P, Davidson EH. Development. Built to run, not fail. *Science*. 2007;315: 1510–1511.
 47. Latchman DS. Gene Control. 2018. doi:10.4324/9780203833568
 48. Davidson EH. Emerging properties of animal gene regulatory networks. *Nature*. 2010;468: 911–920.
 49. The Regulatory Genome: Gene Regulatory Networks In Development And Evolution. [cited 2 Feb 2021]. Available: https://books.google.com/books/about/The_Regulatory_Genome.html?id=F2ibJj1LHGE
 50. Eldar A, Elowitz MB. Functional roles for noise in genetic circuits. *Nature*. 2010;467:

167–173.

51. Chepyala SR, Chen Y-C, Yan C-CS, Lu C-YD, Wu Y-C, Hsu C-P. Noise propagation with interlinked feed-forward pathways. *Scientific Reports*. 2016. doi:10.1038/srep23607
52. Pond KW, Doubrovinski K, Thorne CA. Wnt/β-catenin Signaling in Tissue Self-Organization. *Genes* . 2020;11: 939.
53. Foreman R, Wollman R. Mammalian gene expression variability is explained by underlying cell state. *Mol Syst Biol*. 2020;16: e9146.
54. Senecal A, Munsky B, Proux F, Ly N, Braye FE, Zimmer C, et al. Transcription factors modulate c-Fos transcriptional bursts. *Cell Rep*. 2014;8: 75–83.
55. Franz K, Singh A, Weinberger LS. Lentiviral vectors to study stochastic noise in gene expression. *Methods Enzymol*. 2011;497: 603–622.
56. Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. Mammalian genes are transcribed with widely different bursting kinetics. *Science*. 2011;332: 472–474.
57. Waymack R, Fletcher A, Enciso G, Wunderlich Z. Shadow enhancers can suppress input transcription factor noise through distinct regulatory logic. *eLife*. 2020;9. doi:10.7554/eLife.59351
58. Ali Al-Radhawi M, Del Vecchio D, Sontag ED. Multi-modality in gene regulatory networks with slow promoter kinetics. *PLoS Comput Biol*. 2019;15: e1006784.
59. Gibson MA, Bruck J. Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels. *The Journal of Physical Chemistry A*. 2000. pp. 1876–1889. doi:10.1021/jp993732q
60. Bernstein D. Simulating mesoscopic reaction-diffusion systems using the Gillespie algorithm. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2005;71: 041103.
61. Batmanov K, Kuttler C, Lhoussaine C, Saka Y. Self-organized Patterning by Diffusible Factors: Roles of a Community Effect. *Fund Inform*. 2012;118: 419–461.
62. Chepyala SR, Chen Y-C, Yan C-CS, Lu C-YD, Wu Y-C, Hsu C-P. Noise propagation with interlinked feed-forward pathways. *Scientific Reports*. 2016. doi:10.1038/srep23607
63. Yan C-CS, Chepyala SR, Yen C-M, Hsu C-P. Efficient and flexible implementation of Langevin simulation for gene burst production. *Sci Rep*. 2017;7: 16851.
64. Kohar V, Lu M. Role of noise and parametric variation in the dynamics of gene regulatory circuits. *NPJ Syst Biol Appl*. 2018;4: 40.
65. Kaern M, Elston TC, Blake WJ, Collins JJ. Stochasticity in gene expression: from theories to phenotypes. *Nat Rev Genet*. 2005;6: 451–464.
66. Foreman R, Wollman R. Mammalian gene expression variability is explained by underlying cell state. *Mol Syst Biol*. 2020;16: e9146.
67. Hoppe C, Bowles JR, Minchington TG, Sutcliffe C, Upadhyai P, Rattray M, et al. Modulation of the Promoter Activation Rate Dictates the Transcriptional Response to Graded BMP Signaling Levels in the Drosophila Embryo. *Dev Cell*. 2020;54:

727–741.e7.

68. Fukaya T, Lim B, Levine M. Enhancer Control of Transcriptional Bursting. *Cell*. 2016;166: 358–368.
69. Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. Mammalian genes are transcribed with widely different bursting kinetics. *Science*. 2011;332: 472–474.
70. Bolouri H, Davidson EH. The gene regulatory network basis of the “community effect,” and analysis of a sea urchin embryo example. *Dev Biol*. 2010;340: 170–178.
71. Wong VC, Bass VL, Bullock ME, Chavali AK, Lee REC, Mothes W, et al. NF- κ B-Chromatin Interactions Drive Diverse Phenotypes by Modulating Transcriptional Noise. *Cell Rep*. 2018;22: 585–599.
72. Hoppe C, Bowles JR, Minchington TG, Sutcliffe C, Upadhyai P, Rattray M, et al. Modulation of the Promoter Activation Rate Dictates the Transcriptional Response to Graded BMP Signaling Levels in the Drosophila Embryo. *Dev Cell*. 2020;54: 727–741.e7.
73. Saka Y, Lhoussaine C, Kuttler C, Ullner E, Thiel M. Theoretical basis of the community effect in development. *BMC Syst Biol*. 2011;5: 54.
74. Roussel CJ, Roussel MR. A mathematical model of the biochemical network underlying left-right asymmetry establishment in mammals. *Biosystems*. 2018;173: 281–297.
75. Yan C-CS, Chepyala SR, Yen C-M, Hsu C-P. Efficient and flexible implementation of Langevin simulation for gene burst production. *Sci Rep*. 2017;7: 16851.
76. Ali Al-Radhawi M, Del Vecchio D, Sontag ED. Multi-modality in gene regulatory networks with slow promoter kinetics. *PLoS Comput Biol*. 2019;15: e1006784.
77. Cao Y, Gillespie DT, Petzold LR. Efficient step size selection for the tau-leaping simulation method. *J Chem Phys*. 2006;124: 044109.
78. Klika V, Gaffney EA. History dependence and the continuum approximation breakdown: the impact of domain growth on Turing’s instability. *Proc Math Phys Eng Sci*. 2017;473: 20160744.
79. Seirin Lee S, Gaffney EA, Monk NAM. The influence of gene expression time delays on Gierer-Meinhardt pattern formation systems. *Bull Math Biol*. 2010;72: 2139–2160.
80. Kicheva A, Bollenbach T, Wartlick O, Jülicher F, Gonzalez-Gaitan M. Investigating the principles of morphogen gradient formation: from tissues to cells. *Curr Opin Genet Dev*. 2012;22: 527–532.
81. Middleton AM, King JR, Loose M. Wave pinning and spatial patterning in a mathematical model of Antivin/Lefty-Nodal signalling. *J Math Biol*. 2013;67: 1393–1424.
82. Dubrulle J, Jordan BM, Akhmetova L, Farrell JA, Kim S-H, Solnica-Krezel L, et al. Response to Nodal morphogen gradient is determined by the kinetics of target gene induction. *eLife*. 2015;4. doi:10.7554/eLife.05042
83. Chhabra S, Liu L, Goh R, Kong X, Warmflash A. Dissecting the dynamics of signaling events in the BMP, WNT, and NODAL cascade during self-organized fate patterning in human gastruloids. *PLoS Biol*. 2019;17: e3000498.

84. Kelly JD, Hedengren JD. A steady-state detection (SSD) algorithm to detect non-stationary drifts in processes. *J Process Control.* 2013;23: 326–331.
85. Peter IS, Davidson EH. Modularity and design principles in the sea urchin embryo gene regulatory network. *FEBS Lett.* 2009;583. doi:10.1016/j.febslet.2009.11.060
86. Minokawa T, Wikramanayake AH, Davidson EH. cis-Regulatory inputs of the wnt8 gene in the sea urchin endomesoderm network. *Dev Biol.* 2005;288: 545–558.
87. Staehling-Hampton K, Hoffmann FM. Ectopic decapentaplegic in the *Drosophila* midgut alters the expression of five homeotic genes, dpp, and wingless, causing specific morphological defects. *Dev Biol.* 1994;164: 502–512.
88. Koide T, Hayata T, Cho KWY. Xenopus as a model system to study transcriptional regulatory networks. *Proc Natl Acad Sci U S A.* 2005;102: 4943–4948.
89. Turing AM. The chemical basis of morphogenesis. 1953. *Bull Math Biol.* 1990;52: 153–97; discussion 119–52.
90. Gierer A, Meinhardt H. A theory of biological pattern formation. *Kybernetik.* 1972. pp. 30–39. doi:10.1007/bf00289234
91. Muller P, Rogers KW, Jordan BM, Lee JS, Robson D, Ramanathan S, et al. Differential Diffusivity of Nodal and Lefty Underlies a Reaction-Diffusion Patterning System. *Science.* 2012. pp. 721–724. doi:10.1126/science.1221920
92. Norris DP, Brennan J, Bikoff EK, Robertson EJ. The Foxh1-dependent autoregulatory enhancer controls the level of Nodal signals in the mouse embryo. *Development.* 2002;129: 3455–3468.
93. Alonso CR. A complex “mRNA degradation code” controls gene expression during animal development. *Trends in Genetics.* 2012. pp. 78–88. doi:10.1016/j.tig.2011.10.005
94. Palani S, Sarkar CA. Transient Noise Amplification and Gene Expression Synchronization in a Bistable Mammalian Cell-Fate Switch. *Cell Reports.* 2012. pp. 215–224. doi:10.1016/j.celrep.2012.01.007
95. Hornung G, Barkai N. Noise propagation and signaling sensitivity in biological networks: a role for positive feedback. *PLoS Comput Biol.* 2008;4: e8.
96. Franz K, Singh A, Weinberger LS. Lentiviral vectors to study stochastic noise in gene expression. *Methods Enzymol.* 2011;497: 603–622.
97. Fukaya T, Lim B, Levine M. Enhancer Control of Transcriptional Bursting. *Cell.* 2016;166: 358–368.
98. Fukaya T, Lim B, Levine M. Enhancer Control of Transcriptional Bursting. *Cell.* 2016;166: 358–368.
99. Hoppe C, Bowles JR, Minchington TG, Sutcliffe C, Upadhyai P, Rattray M, et al. Modulation of the Promoter Activation Rate Dictates the Transcriptional Response to Graded BMP Signaling Levels in the *Drosophila* Embryo. *Dev Cell.* 2020;54: 727–741.e7.
100. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A. Regulation of noise in the expression of a single gene. *Nat Genet.* 2002;31: 69–73.

101. Chepyala SR, Chen Y-C, Yan C-CS, Lu C-YD, Wu Y-C, Hsu C-P. Noise propagation with interlinked feed-forward pathways. *Scientific Reports*. 2016.
doi:10.1038/srep23607
102. Yang L, Angelova Volponi A, Pang Y, Sharpe PT. Mesenchymal Cell Community Effect in Whole Tooth Bioengineering. *Journal of Dental Research*. 2017. pp. 186–191.
doi:10.1177/0022034516682001
103. Nemashkalo A, Ruzo A, Heemskerk I, Warmflash A. Morphogen and community effects determine cell fates in response to BMP4 signaling in human embryonic stem cells. *Development*. 2017;144: 3042–3053.