Dissecting the dynamics of coordinated active transcriptional repression in a multicellular organism

Authors:

Virginia Pimmett[1], Maria Douaihy[1,2], Louise Maillard[1], Antonio Trullo[1], Jeremy Dufourt[1,3], Helene Lenden-Hasse[1], Ovidiu Radulescu[2#] and Mounia Lagha[1#]

[1] Institut de Génétique Moléculaire de Montpellier, Univ Montpellier, CNRS, Montpellier, France.

[2] Laboratory of Pathogens and Host Immunity, Univ Montpellier, CNRS, INSERM, Montpellier, France.

[3] Institut de Recherche en Infectiologie de Montpellier, Univ Montpellier, CNRS, Montpellier, France.

[#] Corresponding authors

Abstract: The ability to refine transcriptional levels via active repression in an euchromatic context represents a critical regulatory process during development. While the molecular players of active repression are well described, their dynamics remain largely obscure. By monitoring expression dynamics of the pro-EMT developmental gene *snail* in *Drosophila* embryos, we uncovered and quantified the timescale of kinetic bottlenecks tuning transcription during repression. Repression is associated with the transition of the promoter from two states to a three-state regime, comprising two temporally distinct inactive periods. Surprisingly, repression occurs without abrupt changes in Pol II initiation rates. By monitoring nuclear Sna protein levels, we show that Sna-mediated repression operates with high cooperativity, a feature dictating the degree of cell-cell coordination in the imposition of repression. Our approach offers quantitative insights into the dynamics of repression mediated by short-range repressors and how their cooperativity may coordinate cell fate decisions within a tissue.

Introduction

The selective adoption of specific cell fates critically depends on differential gene expression. Cells are specified through the concomitant transcriptional activation of key lineage specifying factors and the repression of alternative unwanted fates. This coordinated ON/OFF control of transcription is particularly important during the development of multicellular organisms where multipotent cells must choose distinct differentiation routes at specific timings and locations, in a coordinated manner. Thanks to functional genomics approaches, how the cis-regulatory code dictates the activation or silencing of developmental promoters is well-understood. Activating transcription factors typically bind to enhancers to elicit the recruitment of the preinitiation complex (PIC) and initiate transcription at the promoter. Conversely, downregulation of a gene is achieved via the action of repressors, transcription factors that recruit co-repressors to reduce or silence gene expression.

Depending on their range of action, repressors and their partners can be categorized into long or short-range repressors. Some (co)-repressors, such as Groucho/TLE, act over large distances and mediate long-range repression by silencing the entire locus. This mechanism generally involves the oligomerization of co-repressors that contribute to the spreading of repression and a subsequent altered chromatin structure. In contrast, short-range repressors function locally within 50-150 bp, to inhibit the basal transcriptional machinery without interfering with more distantly bound activators[1]. Short-range repression has been well-characterized genetically in the early *Drosophila* embryo during patterning. Transcriptional repressors including Giant, Krüppel, Knirps and Snail have been shown to act locally by recruiting the conserved co-repressor CtBP[2]. At the molecular level, several mechanisms have been proposed, including a direct competition between activators and repressors for a shared DNA binding site or the 'quenching' of closely located activators and members of the basal transcription machinery[3]. A third and non-exclusive mechanism is the recruitment of histone deacetylases (HDACs) by co-repressors, causing chromatin condensation and restricting access to the promoter. Many examples demonstrate CtBP-mediated repression requires interaction with HDAC1 in both vertebrates[4] and invertebrates[5] to control key cellular programs such as epithelial-to-mesenchymal transitions, EMT[6].

While the identity of repressors and co-repressors is well known, their impacts on the kinetics of transcription are much less described. Transcriptional bursting, characterized by alternating periods of activity and inactivity, is common to many promoters[7]. For such promoters, the mean RNA production is determined by the product of two parameters, the initiation rate in the active state and the probability of being in this state. Repressors can in principle gradually modulate these parameters, resulting in fine-tuning of repression rather than an all-or-nothing process. However, the kinetics of repression and which parameter(s) they modulate are still unknown.

Two modes of transcriptional repression can be distinguished: the classical silencing in the context of hostile chromatin (heterochromatin) or reduction in the context of a euchromatic environment, referred to as active repression. Polycomb-mediated silencing via chromatin compaction is very well-documented, particularly at the *Drosophila Hox* locus[8]. In contrast, much less is known concerning active repression[9]. Yet, because of its fast establishment, reversibility, and capacity for gradual, partial reduction of expression, active repression stands as an optimal mode of gene expression control during periods of rapid decision-making such as during cell specification.

Active repression can be total, leading to gene silencing, or partial (transcriptional attenuation), in order to reduce expression levels. In both cases, active repression represents a critical layer of gene regulation. Failure of repression can lead to developmental defects and disease such as cancer, as exemplified by the regulation of the process of EMT. This fundamental cellular process is instructed by the conserved pro-EMT Snail gene family, composed of Snail/Slug, Twist and ZEB1, encoding both activators and repressors. Snail (Sna) is a zinc finger transcription factor primarily acting as a repressor but reported to also act as an activator in some contexts[10]. Regardless of the tissue examined, Sna plays a critical role for correct completion of EMT, as for example during *Drosophila* gastrulation or vertebrate neural tube formation[11]. Sna overexpression is sufficient to induce EMT[12] and tumorigenesis (HCC, GC, BC, OC, melanoma)[13]. Sna levels correlate with the tumor grade and predict a poor prognosis in patients with metastatic cancers[14,15]. Thus, tuning Sna levels and the repression network induced by this repressor TF are functionally relevant.

In this study, we use *snail* expression dynamics as a paradigm to uncover how active repression, mediated by the Sna short-range repressor, can be imposed within a developing tissue. We take advantage of the power of quantitative live imaging to monitor endogenous *sna* transcription and protein dynamics in single nuclei prior to a major developmental decision, EMT. Using novel theoretical approaches, we uncovered and quantified the timescale of the kinetic bottlenecks tuning transcription during repression. Based on experimental measurements, we developed a minimal stochastic model of repression. Numerical simulations suggest that coordinated repression within a tissue relies on repressor TF cooperativity.

Results

Monitoring transcriptional repression in vivo

To decode the dynamics of transcriptional repression, we focused on a model gene, *sna*, which undergoes partial repression in the early blastoderm embryo. This gene encodes a key TF, instructing the mesodermal fate and subsequent EMT prior to gastrulation. Consistent with previous reports[16], we confirmed by smFISH that transcriptional activity reduces across nuclear cycle 14 (nc14) (**Figure 1G-H**). Critically, activity is not completely silenced but instead is only attenuated. This significant transcriptional attenuation but not silencing gives us the opportunity to explore the mechanisms and dynamics of repression in the context of active chromatin, in natural physiological conditions.

To examine the endogenous dynamics of *snail* expression in real time, we inserted a 24xMS2 stem-loop cassette[17] into the 3' UTR of the endogenous *snail* gene using CRISPR-mediated recombination (sna^{MS2}; **Figure 1A, Supplementary Movie 1**). The resulting flies are homozygous viable, indicating the MS2 insertion does not obviously perturb the gene function. Moreover, we confirmed that the MS2 reporter expression overlapped the endogenous *sna* signal using smFISH **(Supplementary Figure 1)**.

To image transcription, a protein binding the MS2 step loops fused to GFP (MCP-GFP) is maternally provided, along with a fluorescently-tagged histone (His2A-RFP). In combination with the paternally provided $snail^{MS2}$ allele, transcription is visible as bright nuclear foci within the nuclei. Signal intensity was retrieved in 3D and tracked through nc13 and nc14, using mitosis as a temporal 'time-zero' reference for each nuclear cycle.

Snail transcription has both stationary and non-stationary regimes

We began by characterizing the dynamics of endogenous *sna* transcription at the single nucleus level. Transcription is detectable in living embryos as early as nc11, although earlier activation cannot be ruled out (**Figure 1B**). In nc13, transcription synchronously reactivates shortly after mitosis and exhibits a high level of activity, with most nuclei remaining stably active throughout the cycle (**Figure 1C-E**). Like nc13, reactivation after mitosis in nc14 is rapid and synchronous (**Figure 1C, Supplementary Figure 1**). In contrast, activation is not maintained and instead declined after a short plateau (**Figure 1C,D, Supplementary Figure 1**). This change of regime and transcriptional attenuation is reflected in the individual nuclear traces as well, where TS intensity declined after peaking without completely vanishing (**Figure 1F, Supplementary Figure 1**). Thus, it appears that while *sna* transcriptional activity is stable and steady in nc13 (**Figure 1D**), it undergoes rapid evolution in nc14 and is therefore operating in a non-stationary regime.

To access the underlying kinetic parameters driving *sna* expression, we employed our previously developed deconvolution pipeline[18–20] that converts single nuclei MS2-MCP-GFP signal into a choreography of Pol II initiation events (**Figure 2A-B**). Critically, this process does not rely on the arbitrary assignment of bursts to the signal and is valid for both stationary and non-stationary signals.

In brief, we consider that the intensity trace of each spot is a convolution of multiple concurrently transcribing polymerases and model the contribution of a single polymerase, assuming full processivity, constant speed and a negligible retention time at the transcription site (see methods). To estimate the dwell time comprising Pol II elongation and transcript retention at the TS, we used signal autocorrelation[21] (**Supplementary Figure 2**) and published Pol II speed measurements[22].

We then extracted the waiting times between polymerase initiation events for each individual nucleus (**Supplementary Figure 3**). We quantified the mean waiting time between polymerase initiation events ($<\tau>$) in a sliding window across a given nuclear cycle (**Supplementary Figure 3**). The mean waiting time $<\tau>$ is directly related to the product of p_{ON}, or the probability to be in a productive state, and k_{ini}, or the initiation rate. Importantly, $<\tau>$ cannot provide p_{ON} or k_{ini} independently but can only describe the relationship between the two values. While this value appears stable in all window positions across nc13, it is clearly unstable in nc14 (**Figure 2C-D**). We therefore conclude that *snail* transcription kinetics exhibits two distinct temporal regimes, a time homogeneous behaviour in nc13, followed by a period in nc14 during which transcription rates are not homogeneous in time.

Active transcription employs a two-state promoter switching dynamic while repression induces a third state

In our previous works, we utilized the survival function of the waiting times between polymerase initiation events (defined as the complementary cumulative distribution function) to infer transcription kinetics from live imaging data[18–20]. Although the application of a survival function can be extended to non-stationary signals, our solutions for the inverse problem of determining kinetic models from data are valid under the assumption of stationarity. Similarly, alternative methods like Hidden Markov Models (HMMs) and autocorrelation models were also designed with this stationarity assumption[23–25]. In nc13, the signal is time homogeneous, and thus we could employ the survival function approach to extract kinetic parameters. The stationarity assumption does not hold for nc14, where the transcription kinetic parameters are time-dependent because repression builds in progressively.

While we cannot apply the survival function method to the entire signal, we sought to identify a time at which stable repression is established by applying a Bayesian change point detection (BCPD) algorithm[26] (see Methods and **Supplementary Figure 4**). Importantly, the window corresponding to stable repression shows a time homogeneous signal (**Figure 2D'**) and can be thus analysed using the survival function method to extract bursting kinetics.

To uncover *sna* transcription kinetics, and particularly its multiscale bursting, we compared the dynamics in two clearly distinct timings, nc13 (activation) and the effective repression in 14 that we extracted using BCPD.

In both cases, we performed parametric estimates of the empirical survival function using two- and three-exponential models. During activation, we found that a simple two-exponential fitting was able to reliably capture the data (**Figure 2E, Supplementary Table 4**). A two-exponential fitting corresponds to the simple 'telegraph' model of transcription, where a promoter switches between a non-permissive OFF state and a competent ON state, from which polymerase can be released into the gene body[27]. This result is consistent with the dynamics of the *sna* promoter when examined in a transgenic context with a unique minimal enhancer[20].

The parameters of the waiting time fitting provide the characteristic probabilities and durations of the ON and OFF states in nc13. The ON state is relatively highly probable at 0.61, with a duration of 142 seconds (**Figure 2G**). The OFF state has a probability of 0.39 and a duration of 91 seconds. We note that while the OFF state may seem incongruent with the stable activity visually apparent in **Supplementary Movie 1**, the non-productive state can be partially 'hidden' by the persistence of the MS2 signal from previously initiated polymerases. In the ON state, one polymerase is released into productive elongation approximately every 3 seconds. Collectively, our data demonstrate that *snail* can maintain a high level of transcriptional activity using a two-state kinetic regime.

In contrast to these dynamics, during stable repression, *snail* transcription undertakes a different kinetic regime. Indeed, the distribution of waiting times between polymerase initiation events clearly shows that a two-state fitting is not sufficient to capture the data but requires a three-state fitting **(Figure 2F,F', Supplementary Table 4**). This distribution also shows two distinct temporal scales of Pol II waiting times during repression, one in the order of minutes (OFF_1) and one in the order of seconds (OFF_2)(**Figure 2H**).

Various three-state topologies can be envisaged, such as a consecutive OFF state (stair) model or a non-consecutive model with alternative routes to distinct OFF states (for example, non-obligatory pausing), as shown previously by our lab and others[28,29]. We were unable to statistically differentiate between a topology where transcription passes systematically through sequential non-productive states and a topology where the two non-productive states are non-sequential and transit directly between them is not permitted (**Supplementary Figure 5**). While these topologies produce statistically equivalent MS2 signals, they provide slightly different values for the probabilities and durations of each inactive state (**Supplementary Table 4**). The probability and duration of the active state, as well as the initiation rate are however well-determined and consistent across all possible models.

In comparing the kinetic regimes of the active nc13 (**Figure 2, Supplementary Table 4**) and repression phase in nc14, no significant change in the initiation rate of RNA Polymerase II was observed (**Supplementary Table 4**). However, during repression, the probability of the ON state decreased from 0.39 to 0.09 and

the duration from 142s to 8s. Additionally, the duration of the long OFF (OFF1) increased from 91 to 956 s. Together, these indicate that repression is associated with an increase in the number and duration of the non-productive states and a reduction in the duration and probability of the productive state, but not with a change in polymerase initiation rate.

Moreover, the timescales of the two inactive states observed during repression (OFF1 and OFF2) are distinct from the lifetime of the unique OFF state obtained during activation in nc13. The duration of OFF2 is shorter than of OFF during activation, while OFF1 is longer than the inactive promoter state during activation. During *sna* activation, transcription operates via a simple 2 state regime, comprising a unique OFF state, which we and others have related to PIC disassembly. Indeed, mutations in the TATA box of core promoters, known to affect PIC stability, decrease the duration of the competent ON state, at the expense of the OFF state[19,20,30,31]. Because the PIC disassembled state must be present during both activation and repression, and because this inactive state cannot be shortened by the action of a repressor, we hypothesise that the short OFF state observed during repression (OFF2) cannot correspond to the PIC disassembled state. We therefore propose that during repression, OFF1 may reflect the PIC disassembled state while the shorter OFF2 may correspond to the binding of a repressor. With such an interpretation of promoter states, our data suggest that during *sna* repression, the lifetime of the PIC disassembled state significantly increases (by one order of magnitude). We note that the lifetime of the repressor-bound putative state, OFF2 (22 seconds) is compatible with the typical residence time of transcription factors[32].

Exploring the interplay between repression and polymerase pausing

We and others have observed that promoter proximal pausing can lead to the regulation of transcription through additional inactive promoter states[19,20,30]. Based on our previous work, we have proposed a non-systematic model of pausing, whereby only a subset of initiated Pol II would enter into a paused state (**Supplementary Figure 5**). Moreover, the *snail* promoter has been shown to progressively acquire pausing[33–35] at a timing concordant with that of stable *sna* repression. We therefore questioned whether *snail* repression was regulated at the level of promoter pausing.

Multiple factors control the entry of polymerase into a paused state and its release from this state. Among them, RNA Polymerase II-associated factor 1 (Paf1, also known as *antimeros* in *Drosophila*), plays a critical role in regulating pausing along with other functions during elongation[36]. Reducing Paf1 levels, through classical knockdown, acute degradation, or a small molecule inhibitor resulted in augmented pause release genome-wide[36–39]. Based on these studies, we reasoned that a depletion of Paf1 in *Drosophila* embryos would diminish pausing and favor pause release, thus possibly affecting transcriptional repression.

To investigate how repression dynamics would be impacted by promoter proximal pausing, we quantified *snail* transcription dynamics in embryos maternally depleted for Paf1 (**Figure 3, Supplementary Figure 6**). We employed two distinct RNAi strategies targeting a core component of the Paf1 complex (*Paf1*-i^{A} and *Paf1*-i^{B}; **Supplementary Movies 2-4**). Efficiency of knock-down was verified by qPCR (**Supplementary Figure 6**) and indicates the depletion of ~70% of the *Paf1* maternal pool. *Paf1*-i^{A} embryos showed similar post-mitotic reactivation compared to *white*-i controls (**Figure 3B**), but sustained transcription for longer and at a higher intensity (**Figure 3B,C**).

Both RNAi approaches caused a substantial reduction in the level of *snail* repression in nc14 (**Figure 3B,C**).

During the repression phase, we observe that transcription is regulated by three promoter states, comprising two inactive states, a short and longer-lived inactive state. We hypothesized that if one of these inactive states was linked to promoter proximal pausing, it should be affected when pausing is perturbed. Indeed, in a previous work we observed a switching from a three-state dynamic to a two-state upon perturbing pausing in *cis* or in *trans*[20]. Thus, we tested if perturbing pausing would switch *snail* repression dynamics to a simple two-state model.

Using BCPD, we could isolate a window of stable transcription in nc14 corresponding to stable repression of *snail*, in both *Paf1*-i^{A} conditions and in *white*-i control. By applying our fitting procedure to this specific temporal window, we observe that the *sna* promoter dynamics cannot be accurately captured by a simple ON/OFF two-state regime (**Figure 3D-F**) upon Paf1 depletion. Thus, when promoter pausing is perturbed, active repression still comprises three rate-limiting steps, including two inactive states.

To confirm these results, we employed an orthogonal genetic approach in which the release of paused polymerase into productive elongation was favored. Pause release requires the phosphorylation of Pol II CTD and the NELF complex by P-TEFb, a kinase-cyclin pair composed of Cyclin T (CycT) and Cdk9[40]. In an attempt to favor pause release, we overexpressed Cyclin T with the UAS/Gal4 system[41], leading to a 5.5-fold increase in its levels (**Supplementary Figure 7**). Next, we monitored *sna* transcription dynamics in this Cyclin T over-expression context (**Supplementary Movies 5-6**). While *sna* activation in early nc14 is not affected by overexpression of Cyclin T, *sna* repression appears to be slightly delayed (**Figure 3G,H**). Thus, two distinct genetic approaches favoring pause release (knockdown of Paf1 or overexpression of Cyclin T) affect, albeit modestly, the timing of *sna* repression. Moreover, upon Cyclin T overexpression, long inactive periods between polymerase are still observed (**Supplementary Figure 7**) and the distribution of Pol II waiting time cannot be fitted by a simple two state model (ON/OFF) (**Figure 3I-J**). Thus, affecting pause release by two distinct approaches does not affect the number of rate-limiting steps regulating *snail* during active repression. In all genetic contexts examined, *snail* transcription during the repressive phase appears to be regulated by three promoter states (ON, OFF1 and OFF2).

Collectively these results suggest that during repression, a promoter switches between three temporally distinct states: a competent ON state, from which Pol II initiates at a given rate, and two inactive states, a short second-scale and a longer minute-scale state. Because the three-state promoter topology persists in conditions where pause release is favored, we conclude that the extra rate-limiting step present during repression cannot be obviously attributed to a paused state. Instead, we hypothesized that the extra rate limiting state could correspond to the recruitment of a repressor and next investigated its identity.

Snail-mediated active repression involves Sna cooperative action

As Snail protein has been shown to directly repress neuroectodermal genes in the mesoderm[42,43], we hypothesized that the reduction in *sna* transcription could be caused by Snail protein itself. Snail is known to bind to snail enhancers[44] and genetic perturbations indeed point to Sna action via a negative feedback loop[16]. To examine the dynamics of Sna protein

endogenously, we employed the LlamaTag system and created a CRISPR *snail*^{Llama} allele (*sna*^{Llama}; **Figure 4A,B**).

The LlamaTag system[45] relies on the addition of a small nanobody tag to a protein of interest and the presence of a free fluorescent detector. Maternally-deposited free GFP in the embryo is captured by the mature nanobody, and the bound fluorescent complex provides a readout of protein localization and levels without the complexity of delayed fluorophore maturation.

The *sna*^{Llama} CRISPR allele allowed us to quantify with a high spatio-temporal resolution the levels of endogenous Sna nuclear protein across nc13 and nc14 (**Figure 4C,E-F**). Despite the high levels of *sna* transcripts accumulating as early as nc11 **(Figure 1B)**[46], stable nuclear accumulation of Sna protein was detected beginning in nc13 (**Figure 4C**). By comparing signal within the mesoderm to that in a neighboring tissue where *sna* is not expressed (neurogenic ectoderm, **Figure 4D,D'**), we could see that Sna nuclear protein levels in nc13 were only slightly above this background level (**Figure 4E**). In contrast, in nc14 the levels of nuclear Snail continuously increase in the presumptive mesoderm relative to the neurogenic ectoderm (**Figure 4F**). Consistent with the idea that Snail represses its own transcription by binding directly to its own enhancers[44], we observed an anticorrelation between *snail* endogenous transcription rates and Sna protein levels (**Figure 4G**).

Next, we sought to compare the dynamics of Sna-mediated repression between two targets, *sna* itself and another known target, the gene *short gastrulation* (*sog*, homolog of Chordin) (**Figure 5A,B**). We used a previously characterized *sog*^{MS2} CRISPR allele[47] and performed live imaging in the mesoderm in nc14. Like *sna*, *sog* is expressed in the presumptive mesoderm in early nc14 but then undergoes rapid repression by Sna[48] (**Figure 5D**). MS2-MCP fluorescent signal from *sog* promoter was analyzed in the same manner as the *sna* signal to retrieve overall transcription rates. We used the BCPD method to isolate windows of stable repression and plotted the time of stable repression onset in each individual (**Figure 5E,G**). We found that repression was much more coordinated for *sog* than for *sna*, as demonstrated by the breadth of repression onset times for each gene (**Figure 5E,G**).

Importantly, the relationship between mRNA production and protein levels can be fitted with small residuals by a non-linear function. The sigmoidal shape of the $p_{ON} \cdot k_{ini}$ graph (**Figure 5F, colored line)** suggests the use of a Hill function for the fit. The Hill coefficient for *sna* during nc14 was 6.07±0.30, (**Figure 5F**). Similarly, the relationship between input Sna protein and output *sog* transcription can be fitted with a Hill function with a high Hill coefficient (~7.6±0.55). However, in this case, the residuals from the Hill function are relatively high (**Figure 5H**), suggesting that the Hill function model may be too simplistic, and that Snail protein is not the sole repressor of *sog*, as has been noted by others[49].

Collectively, the relatively high Hill coefficients indicate that Snail protein may act cooperatively to elicit repression on its own *cis*-regulatory regions or on *sog* enhancers. We note however, that the Hill function model is purely phenomenological and does not account for the details of the underlying mechanism.

In order to gain quantitative insights into the relation between Snail-mediated repression and transcriptional bursting, we developed a minimal stochastic model. Based on our experimental results, we used a three-state promoter model, comprising two transcriptionally inactive states (OFF1) and (OFF2) and one permissive (ON) state (**Figure 6A**). As we notice (**Supplementary Table 4**) that

the parameter k_1^m (transition rate from ON to OFF1) remains constant between the nc13 active regime and the BCPD-isolated repressed regime in nc14, while the parameter k_1^p (transition rate from OFF1 to ON) corresponds to long timescales (91s and 956s in active and repressed regimes, respectively), we associate the OFF1 state to the PIC in our model while OFF2, that has a short lifetime in the repressed regime (22s) is associated to the repressor binding. As for the parameter k_1^p and the rest of the parameters that vary from active to repressed regimes (k_2^m and k_2^p), we hypothesized that they depend on the concentration of the repressor Sna. Furthermore, the repressor-dependent state (OFF2) is not accessible, therefore the corresponding rate leading to this state is vanishing ($k_2^m = 0$) and the three-state model functions as a two-state telegraph model in the absence of Sna (**Figure 6A**). Interestingly, our model predicts that, while the lifetime of the PIC disassembled state (OFF1) undergoes a change of one order of magnitude (parameter k_1^p), the lifetime of OFF2 changes very little when transitioning from the active to the repressed regime (parameter k_2^p) (**Figure 6B**, see Methods).

We used Hill functions to model the dependence between the transition rates and the Sna protein concentration (**Figure 6C, C'**). The parameters were chosen according to the switching parameters obtained from the nc13 active phase in the active regime of the model, and the BCPD-derived repressed phase of nc14 in the repressed regime of the model (**Figure 6C'**, see Methods). To approximate the model more closely to experimental data, we also incorporated the post-mitotic lag time obtained from the MS2 data. Gillespie simulations were used to generate synthetic nascent RNA data for this model. After parameter fitting, the model recapitulates the observed behaviour of mean repression in nc14 (**Figure 6D**) and the distribution of the repression onset time within the mesodermal population (**Figure 6E**). We should note that after imposing the active and repressed phase parameter values, our model remained with only three free parameters that were fitted using the experimental $p_{ON}.k_{ini}$ data only. The good agreement between the predicted and experimental repression onset time distributions validates the model.

Next, we investigated the model behavior for different sets of parameters of the Hill function. We considered the coordination of repression by describing the half-width, which describes where the distribution of the repression onset time reaches half of its maximum amplitude and determining the width of this region (**Figure 6F**), which was 9.3 minutes for *sna*. Interestingly, for Hill coefficients immediately smaller than the values obtained from experimental data, the distribution of the repression onset time becomes much broader (ie. the half-width is larger), implying a lack of inter-nuclear coordination of repression (**Figure 6G**). In general, for a fixed theta value, which represents here the threshold of half concentration of protein needed to switch into repression, the distribution of the repression onset time narrows as the Hill index increases (**Figure 6H**). Further, the repression onset time decreases as θ decreases for a fixed Hill coefficient. This suggests a monotonic relation between Sna cooperativity, protein level, and the synchrony of repression between nuclei.

Discussion

Understanding the mechanisms by which gene transcription is dynamically attenuated within a developing tissue is a fundamental question. Here we use live imaging and mathematical modeling to quantitatively address this question in the context of *Drosophila* early development. We focus on *snail* as a model gene to extract the kinetics of promoter switching during active repression, as well as the dynamics of the short-range

repressor protein it encodes. By monitoring nascent mRNA and protein levels from endogenous loci in live embryos, we unveil four main features of active repression: (1) active repression operates via three promoter states comprised of a competent ON state and two temporally distinct OFF states; (2) active repression does not modulate Polymerase initiation rates; (3) Snail protein acts cooperatively to regulate its target genes and (4) the inter-nuclear coordination of active repression is augmented by a high degree of repressor cooperativity.

Transcription kinetics during active repression

The analysis of distribution of waiting times between polymerase during the repression phase revealed the existence of two distinct OFF periods, one in the range of seconds and a prolonged one in the range of minutes. A promoter experiencing active repression appears to switch between three rate- limiting steps: a competent active ON state, and two OFF states. In a previous work with synthetic transgenes, we observed that pausing could create an extra rate-limiting step in promoter states[20]. In this study, we show another example of regulation by three promoter states, but here the extra rate-limiting promoter state does not appear to be obviously linked to paused polymerase. Instead, we believe that this extra state represents the action of a repressor, hypothesized here to be the short-range repressor Snail. Indeed, this supplementary kinetic bottleneck (short OFF state) is not present in earlier development where *snail* transcription is at maximum capacity. Our finding that repression is associated with long-lived inactive states are consistent with a recent study quantifying the transcriptional response to the repression mediated by Knirps[50].

The fact that affecting pausing does not obviously change the number of promoter states and only modestly delays repression was unexpected. Pausing is often seen as an entry point to rapidly impose a transcriptional response to stimuli[51,52]. The stimulus generally leads to a positive response as exemplified by enhanced transcription with heat shock, but it could in principle also be envisaged as a window of opportunity to impose rapid repression. Although mechanistically very different from pausing, it has been previously demonstrated that repressors take advantage of transient breaks in transcription at mitosis to impose repression[53]. Moreover, a functional relationship between pausing factors and Gro-mediated repression has been proposed in Drosophila cells[54,55]. Whether, similarly to Gro/TLE, Sna major co-repressor Ctbp can also affect pause release, remains to be investigated, but our results don't show that pausing significantly affects Sna-mediated repression in embryos. Thus, the relationship between active repression and pausing still remains an open question.

In simple terms, the mean mRNA production is determined by the product of two parameters, the initiation rate in the active state (k_{ini}) and the probability of being in this state (p_{ON}). In principle repression could affect both of these variables. However, our results demonstrate that repression occurs by primarily tuning p_{ON} and does not affect Pol II initiation rates. Numerous studies now show that Pol II initiation rates are relatively constant and are of the same order of magnitude across promoters, tissues, and species[7,56–58]. Our results are also consistent with existing quantifications for gap genes and their target in the *Drosophila* embryo that show a constant and stable Pol II initiation rate along the A/P axis[59,60]. We note however that a recent work examining dosage compensation in *Drosophila* suggests that the transcriptional hyperactivation observed in males can in part be mediated by a modest

increase in Pol II initiation rates[61]. However, this effect appears to be gene-specific and not a general phenomenon.

Our quantitative study provides knowledge on the number and timescale of the key rate-limiting steps regulating transcription, when a promoter is activated or repressed. The molecular nature of these rate-limiting steps can only be an interpretation. Because of the timescale and the comparison between activation and repression and thanks to our mathematical model and simulations, we propose that the shorter OFF2 state, apparent only in repression, corresponds to a repressor-bound state. While the residence time of specific repressors has yet to be detailed *in vivo*, it is well-demonstrated that activating transcription factors remain bound to DNA for up to tens of seconds[32] in both *Drosophila* and vertebrates, consistent with the duration of the OFF2 state. However, it is unclear what impact the highly cooperative nature of Sna repression has on its binding/unbinding dynamics. Sna binding sites cluster closely in the *sna* and *sog cis*-regulatory regions (**Supplementary Figure 8)**, which could foster cooperative interactions leading rapidly to high repressor presence on sequence. We also find that, contrary to binding (described by the parameter k_1^m), the unbinding of the repressor (described by the parameter k_1^p) does not depend on binding site occupancy (**Figure 6B**), suggesting that repressor unbinding consists of independent, affinity-controlled events. However, further mechanistic investigation remains to be performed to understand precisely how Sna binding site occupancy forms, blocks transcription, and breaks down.

Snail-mediated repression: a cooperative action

The cooperativity of repression manifests as a steep response of RNA production to protein concentration. Our mathematical model suggests that the high degree of cooperativity governing Sna-mediated repression might have been selected to ensure optimal inter-nuclear coordination of repression within a tissue.

It is well-demonstrated that transcription factors cooperate to bind to DNA[62]. Multiple examples point to the cooperative action of a combinatorial set of TFs to activate or silence cis-regulatory elements, but the underlying mechanisms remain unclear. Moreover, while recent advances in artificial intelligence allow the predictions of enhancer logic/grammar and even the *de novo* building of synthetic enhancers[63,64], these *in silico* approaches are not yet able to quantitatively measure or predict cooperativity. Here, by directly measuring the relationship between protein input and transcriptional output in live embryos, we provide a quantitative estimation of cooperativity. By examining two transcriptional targets, *snail* and *sog*, expressed in the same tissue at a similar developmental state, we obtained an estimation of Sna-mediated cooperativity, with Hill coefficients in the range of 6 and 7.6 for *sna* and *sog* respectively.

Cooperativity between transcription factors ranges from a Hill coefficient of 1.5-2 for most *E.coli* transcription factors[65] and the yeast GAL4 protein[66] to higher coefficient values in *Drosophila*. Recent work revealed a Hill coefficient of 3.3 for ftz[67], 6.4 for Hunchback[68,69], 7 for Bicoid[70], and 6.58 for another repressor Knirps[50] in the *Drosophila* blastoderm embryo. Although at the descriptive stage, these *in vivo* datasets represent a first important step towards a quantitative understanding of TF collective action at an enhancer.

Cooperativity can occur through physical protein-protein interactions, which imposes a certain arrangement of TF binding sites, but can also be achieved with more flexible arrangements in the absence of physical interactions[62]. Cooperativity can indeed rely on a local change in DNA structure[71] or can be mediated through competition with

nucleosomes[72]. In this latter model of nucleosome-mediated cooperativity, the cooperative binding of multiple TFs within a range of 150-200bp outcompete a nucleosome to evict it from an enhancer, something readily observed in the arrangement of Sna binding sites in the *sna* and *sog* enhancers (**Supplementary Figure 8**). In principle all of these modes of cooperativity (DNA mediated, TF-TF interaction or nucleosome-mediated) can lead to a high Hill coefficient. The use of multiple modes of cooperativity during active repression may play a role in a gene's capacity to ensure rapid, abrupt but coordinated changes in transcription levels.

Future investigations, including promising single molecule approaches such as single molecule footprinting assays[73] coupled to theoretical models[74] will be required to elucidate which TF co-occupy the same enhancer *in vivo*. Such TF co-occupancy mapping would be greatly enhanced by the quantification of TF binding kinetics. Recent advances in single molecule imaging, including the exciting possibility of imaging temporally-evolving repressor 'hubs'[75,76], promise encouraging future insights.

In summary, by monitoring transcription and nuclear transcription factor levels in a developing embryo undergoing cell fate decisions, we have uncovered kinetic bottlenecks governing repression. Combining experimental measurements with theoretical exploration allowed us to reveal and quantify cooperativity in Snail-mediated repression and identify the steps in the transcription cycle that repressors act upon. Recent advances in genome editing and imaging have allowed the deployment of quantitative imaging techniques in more complex organisms. Looking ahead, we expect that this framework of analysis and the results presented in this study will set a foundation for understanding repression dynamics in these more complex vertebrate models of development.

Data availability

BurstDECONV source code is available through Github at https://github.com/oradules/BurstDECONV, also on Zenodo at https://zenodo.org/record/7443044. SegmentTrack is available via github at https://github.com/ant-trullo/SegmentTrack_v4.0) and the special post-processing tool for this study is available at https://github.com/ant-trullo/SpotsFiltersTool. The BCPD algorithm is available via github at https://github.com/mariadouaihy/BCPD_inhomogeneous_transcriptional_signal. The stochastic model simulation code is available via github at https://github.com/mariadouaihy/stochastic_transcriptional_model.

Acknowledgements

We are grateful for Mattias Mannervik for sharing the *UAS-CyclinT* fly stock.

We thank members of the Lagha lab for their critical reading of the manuscript. We thank John Reinitz and Edouard Bertrand for insightful discussions. We are grateful to Amandine Palandri for help with fly handling. We thank Rachel Topno for assistance with autocorrelation-based dwell time estimates. We acknowledge the Montpellier Ressources Imagerie facility (France-BioImaging) and the Biocampus Drosophila facility of Montpellier.
This work was supported by the ERC SyncDev starting grant to M.L and the ANR HubDyn grant to M.L. M.D is supported by the CNRS and University of Chicago Joint PhD programme. L.M is supported by a fellowship from the LabMuse/University of Montpellier PhD program. M.L., O.R., and A.T. are sponsored by CNRS.

Contributions

Conceptualization: ML, OR, VP, MD

Investigation and data acquisition: VP, LM, MD

Design and supervision of *sna*^{Llama} plasmid construct cloning: JD

Cloning of *sna*^{Llama} plasmid construct: HL-H

Image analysis: AT, VP, LM

Mathematical modeling and code development: MD, OR

Funding acquisition: ML, OR

Project administration: ML

Supervision: ML, OR

Writing original draft: VP, ML

Writing, review, and editing: OR, MD, LM, AT

References

1. Courey, A. J. & Jia, S. Transcriptional repression: the long and the short of it. *Genes Dev.* **15**, 2786–2796 (2001).
2. Nibu, Y. *et al.* dCtBP mediates transcriptional repression by Knirps, Krüppel and Snail in the Drosophila embryo. *EMBO J.* **17**, 7009–7020 (1998).
3. Gray, S. & Levine, M. Short-range transcriptional repressors mediate both quenching and direct repression within complex loci in Drosophila. *Genes Dev.* **10**, 700–710 (1996).
4. Kim, T. W. *et al.* Ctbp2 Modulates NuRD-Mediated Deacetylation of H3K27 and Facilitates PRC2-Mediated H3K27me3 in Active Embryonic Stem Cell Genes During Exit from Pluripotency. *Stem Cells* **33**, 2442–2455 (2015).
5. Mannervik, M. & Levine, M. The Rpd3 histone deacetylase is required for segmentation of the Drosophila embryo. *Proc. Natl. Acad. Sci.* **96**, 6797–6801 (1999).
6. Jaiswal, A. & Singh, R. CtBP: A global regulator of balancing acts and homeostases. *Biochim. Biophys. Acta BBA - Rev. Cancer* **1878**, 188886 (2023).
7. Meeussen, J. V. W. & Lenstra, T. L. Time will tell: comparing timescales to gain insight into transcriptional bursting. *Trends Genet.* S0168952523002585 (2024) doi:10.1016/j.tig.2023.11.003.
8. Schuettengruber, B., Bourbon, H.-M., Di Croce, L. & Cavalli, G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* **171**, 34–57 (2017).
9. Hanna-Rose, W. & Hansen, U. Active repression mechanisms of eukaryotic transcription repressors. *Trends Genet.* **12**, 229–234 (1996).
10. Rembold, M. *et al.* A conserved role for Snail as a potentiator of active transcription. *Genes Dev.* **28**, 167–181 (2014).
11. Del Barrio, M. G. & Nieto, M. A. Overexpression of Snail family members highlights their ability to

promote chick neural crest formation. *Development* **129**, 1583–1593 (2002).

12. Ambrosini, A., Rayer, M., Monier, B. & Suzanne, M. Mechanical Function of the Nucleus in Force Generation during Epithelial Morphogenesis. *Dev. Cell* **50**, 197-211.e5 (2019).
13. De Craene, B. & Berx, G. Regulatory networks defining EMT during cancer initiation and progression. *Nat. Rev. Cancer* **13**, 97–110 (2013).
14. Kahlert, U. D., Joseph, J. V. & Kruyt, F. A. E. EMT- and MET-related processes in nonepithelial tumors: importance for disease progression, prognosis, and therapeutic opportunities. *Mol. Oncol.* **11**, 860–877 (2017).
15. Suzuki, T. *et al.* Role of epithelial-mesenchymal transition factor SNAI1 and its targets in ovarian cancer aggressiveness. *J. Cancer Metastasis Treat.* **9**, 25 (2023).
16. Boettiger, A. N. & Levine, M. Rapid Transcription Fosters Coordinate snail Expression in the Drosophila Embryo. *Cell Rep.* **3**, 8–15 (2013).
17. Bertrand, E. *et al.* Localization of ASH1 mRNA particles in living yeast. *Mol. Cell* **2**, 437–445 (1998).
18. Douaihy, M., Topno, R., Lagha, M., Bertrand, E. & Radulescu, O. BurstDECONV: a signal deconvolution method to uncover mechanisms of transcriptional bursting in live cells. *Nucleic Acids Res.*
19. Tantale, K. *et al.* A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat. Commun.* **7**, 12248 (2016).
20. Pimmett, V. L. *et al.* Quantitative imaging of transcription in living Drosophila embryos reveals the impact of core promoter motifs on promoter state dynamics. *Nat. Commun.* **12**, 4504 (2021).
21. Ferguson, M. L. & Larson, D. R. Measuring transcription dynamics in living cells using fluctuation analysis. *Methods Mol. Biol. Clifton NJ* **1042**, 47–60 (2013).
22. Fukaya, T., Lim, B. & Levine, M. Rapid Rates of Pol II Elongation in the Drosophila Embryo. *Curr. Biol.* **27**, 1387–1391 (2017).

23. Lammers, N. C. *et al.* Multimodal transcriptional control of pattern formation in embryonic development. *Proc. Natl. Acad. Sci.* **117**, 836–847 (2020).

24. Desponds, J. *et al.* Precision of Readout at the hunchback Gene: Analyzing Short Transcription Time Traces in Living Fly Embryos. *PLOS Comput. Biol.* **12**, e1005256 (2016).

25. Bowles, J. R., Hoppe, C., Ashe, H. L. & Rattray, M. Scalable inference of transcriptional kinetic parameters from MS2 time series data. *Bioinformatics* 1–10 (2021) doi:10.1093/bioinformatics/btab765.

26. Adams, R. P. & MacKay, D. J. C. Bayesian Online Changepoint Detection. Stat.ML at http://arxiv.org/abs/0710.3742 (2007).

27. Peccoud, J. & Ycart, B. Markovian Modeling of Gene-Product Synthesis. *Theor. Popul. Biol.* **48**, 222–234 (1995).

28. Radulescu, O. *et al.* Identifying Markov chain models from time-to-event data: an algebraic approach. Preprint at http://arxiv.org/abs/2311.03593 (2023).

29. Tunnacliffe, E. & Chubb, J. R. What Is a Transcriptional Burst? *Trends Genet.* **36**, 288–297 (2020).

30. Corrigan, A. M., Tunnacliffe, E., Cannon, D. & Chubb, J. R. A continuum model of transcriptional bursting. *eLife* **5**, 1–38 (2016).

31. Brouwer, I., Kerklingh, E., Van Leeuwen, F. & Lenstra, T. L. Dynamic epistasis analysis reveals how chromatin remodeling regulates transcriptional bursting. *Nat. Struct. Mol. Biol.* **30**, 692–702 (2023).

32. Lu, F. & Lionnet, T. Transcription Factor Dynamics. *Cold Spring Harb. Perspect. Biol.* **13**, a040949 (2021).

33. Chen, K. *et al.* A global change in RNA polymerase II pausing during the Drosophila midblastula transition. *eLife* **2**, e00861 (2013).

34. Lagha, M. *et al.* Paused Pol II coordinates tissue morphogenesis in the Drosophila embryo. *Cell* **153**, 976–87 (2013).

35. Saunders, A., Core, L. J., Sutcliffe, C., Lis, J. T. & Ashe, H. L. Extensive polymerase pausing during Drosophila axis patterning enables high-level and pliable transcription. *Genes Dev.* **27**, 1146–1158 (2013).

36. Chen, F. X. *et al.* PAF1, a Molecular Regulator of Promoter-Proximal Pausing by RNA Polymerase II. *Cell* **162**, 1003–1015 (2015).

37. Wang, Z. *et al.* Coordinated regulation of RNA polymerase II pausing and elongation progression by PAF1. *Sci. Adv.* **8**, eabm5504 (2022).

38. Soliman, S. H. A. *et al.* Enhancing HIV-1 latency reversal through regulating the elongating RNA Pol II pause-release by a small-molecule disruptor of PAF1C. *Sci. Adv.* **9**, eadf2468 (2023).

39. Adelman, K. *et al.* Drosophila Paf1 Modulates Chromatin Structure at Actively Transcribed Genes. *Mol. Cell. Biol.* **26**, 250–260 (2006).

40. Price, D. H. P-TEFb, a Cyclin-Dependent Kinase Controlling Elongation by RNA Polymerase II. *Mol. Cell. Biol.* **20**, 2629–2634 (2000).

41. Hunt, G. *et al.* Tissue-specific RNA Polymerase II promoter-proximal pause release and burst kinetics in a Drosophila embryonic patterning network. *Genome Biol.* **25**, 2 (2024).

42. Cowden, J. & Levine, M. The Snail repressor positions Notch signaling in the Drosophila embryo. *Development* **129**, 1785–93 (2002).

43. Hemavathy, K., Hu, X., Ashraf, S. I., Small, S. J. & Ip, Y. T. The repressor function of Snail is required for Drosophila gastrulation and is not replaceable by Escargot or Worniu. *Dev. Biol.* **269**, 411–420 (2004).

44. Koenecke, N., Johnston, J., He, Q., Meier, S. & Zeitlinger, J. Drosophila poised enhancers are generated during tissue patterning with the help of repression. *Genome Res.* **27**, 64–74 (2017).

45. Bothma, J. P., Norstad, M. R., Alamos, S. & Garcia, H. G. LlamaTags: A Versatile Tool to Image Transcription Factor Dynamics in Live Embryos. *Cell* **173**, 1810-1822.e16 (2018).

46. Sandler, J. E. & Stathopoulos, A. Quantitative single-embryo profile of drosophila genome activation and the dorsal–ventral patterning network. *Genetics* **202**, 1575–1584 (2016).

47. Whitney, P. H., Shrestha, B., Xiong, J., Zhang, T. & Rushlow, C. A. Shadow enhancers modulate distinct transcriptional parameters that differentially effect downstream patterning events. *Development* **149**, dev200940 (2022).

48. Bothma, J. P., Magliocco, J. & Levine, M. The Snail Repressor Inhibits Release, Not Elongation, of Paused Pol II in the Drosophila Embryo. *Curr. Biol.* **21**, 1571–1577 (2011).

49. Koromila, T. & Stathopoulos, A. Distinct Roles of Broadly Expressed Repressors Support Dynamic Enhancer Action and Change in Time. *Cell Rep.* **28**, 855-863.e5 (2019).

50. Zhao, J. *et al.* Optogenetic dissection of transcriptional repression in a multicellular organism. Preprint at https://doi.org/10.1101/2022.11.20.517211 (2022).

51. Abuhashem, A., Garg, V. & Hadjantonakis, A.-K. RNA polymerase II pausing in development: orchestrating transcription. *Open Biol.* **12**, 210220 (2022).

52. Core, L. J. & Adelman, K. Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev.* **33**, 960–982 (2019).

53. Esposito, E., Lim, B., Guessous, G., Falahati, H. & Levine, M. Mitosis-associated repression in development. *Genes Dev.* **30**, 1503–1508 (2016).

54. Kaul, A., Schuster, E. & Jennings, B. H. The Groucho Co-repressor Is Primarily Recruited to Local Target Sites in Active Chromatin to Attenuate Transcription. *PLoS Genet.* **10**, e1004595 (2014).

55. Martínez Quiles, M. L. & Jennings, B. H. The Groucho co-repressor inhibits progression through the early transcription elongation checkpoint to repress gene expression. Preprint at https://doi.org/10.1101/2022.09.17.508372 (2022).

56. Pelechano, V., Chávez, S. & Pérez-Ortín, J. E. A Complete Set of Nascent Transcription Rates for Yeast Genes. *PLoS ONE* **5**, e15442 (2010).

57. Hoppe, C. *et al.* Modulation of the Promoter Activation Rate Dictates the Transcriptional Response to Graded BMP Signaling Levels in the Drosophila Embryo. *Dev. Cell* **54**, 727-741.e7 (2020).

58. Petrenko, N. & Struhl, K. Comparison of transcriptional initiation by RNA polymerase II across eukaryotic species. *eLife* **10**, e67964 (2021).

59. Chen, P.-T., Zoller, B., Levo, M. & Gregor, T. Gene Activity as the Predictive Indicator for Transcriptional Bursting Dynamics. Preprint at http://arxiv.org/abs/2304.08770 (2023).

60. Berrocal, A., Lammers, N. C., Garcia, H. G. & Eisen, M. B. Unified bursting strategies in ectopic and endogenous even-skipped expression patterns. Preprint at https://doi.org/10.1101/2023.02.09.527927 (2023).

61. Forbes Beadle, L., Zhou, H., Rattray, M. & Ashe, H. L. Modulation of transcription burst amplitude underpins dosage compensation in the Drosophila embryo. *Cell Rep.* **42**, 112382 (2023).

62. Kim, S. & Wysocka, J. Deciphering the multi-scale, quantitative cis-regulatory code. *Mol. Cell* **83**, 373–392 (2023).

63. De Almeida, B. P. *et al.* Targeted design of synthetic enhancers for selected tissues in the Drosophila embryo. *Nature* (2023) doi:10.1038/s41586-023-06905-9.

64. Taskiran, I. I. *et al.* Cell-type-directed design of synthetic enhancers. *Nature* **626**, 212–220 (2024).

65. Hussein, R. & Lim, H. N. Direct comparison of small RNA and transcription factor signaling. *Nucleic Acids Res.* **40**, 7269–7279 (2012).

66. Luo, Y., North, J. A., Rose, S. D. & Poirier, M. G. Nucleosomes accelerate transcription factor dissociation. *Nucleic Acids Res.* **42**, 3017–3027 (2014).

67. Zhao, J., Perkins, M. L., Norstad, M. & Garcia, H. G. A bistable autoregulatory module in the developing embryo commits cells to binary expression fates. *Curr. Biol.* **33**, 2851-2864.e11 (2023).

68. Park, J. *et al.* Dissecting the sharp response of a canonical developmental enhancer reveals multiple sources of cooperativity. *eLife* **8**, e41266 (2019).

69. Gregor, T., Tank, D. W., Wieschaus, E. F. & Bialek, W. Probing the Limits to Positional Information. *Cell* **130**, 153–164 (2007).

70. Lucas, T. *et al.* 3 minutes to precisely measure morphogen concentration. *PLOS Genet.* **14**, e1007676 (2018).

71. Jolma, A. *et al.* DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature* **527**, 384–388 (2015).

72. Mirny, L. A. Nucleosome-mediated cooperativity between transcription factors. *Proc. Natl. Acad. Sci.* **107**, 22534–22539 (2010).

73. Sönmezer, C. *et al.* Molecular Co-occupancy Identifies Transcription Factor Binding Cooperativity In Vivo. *Mol. Cell* **81**, 255-267.e6 (2021).

74. Estrada, J., Wong, F., DePace, A. & Gunawardena, J. Information Integration and Energy Expenditure in Gene Regulation. *Cell* **166**, 234–244 (2016).

75. Mazzocca, M., Colombo, E., Callegari, A. & Mazza, D. Transcription factor binding kinetics and transcriptional bursting: What do we really know? *Curr. Opin. Struct. Biol.* **71**, 239–248 (2021).

76. deHaro-Arbona, F. J. *et al.* Dynamic modes of Notch transcription hubs conferring memory and stochastic activation revealed by live imaging the co-activator Mastermind. Preprint at https://doi.org/10.7554/eLife.92083.1 (2023).

77. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

78. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).

79. Van Der Walt, S. *et al.* scikit-image: image processing in Python. *PeerJ* **2**, e453 (2014).

80. Tsanov, N. *et al.* smiFISH and FISH-quant – a flexible single RNA detection approach with super-resolution capability. *Nucleic Acids Res.* **44**, e165–e165 (2016).

81. Dufourt, J. *et al.* Imaging translation dynamics in live embryos reveals spatial heterogeneities.

Science **372**, 840–844 (2021).

82. Birnie, A., Plat, A., Korkmaz, C. & Bothma, J. P. Precisely timed regulation of enhancer activity defines the binary expression pattern of Fushi tarazu in the Drosophila embryo. *Curr. Biol.* S0960982223004530 (2023) doi:10.1016/j.cub.2023.04.005.
83. Coulon, A. & Larson, D. R. Fluctuation Analysis. in *Methods in Enzymology* vol. 572 159–191 (Elsevier, 2016).
84. Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME Suite. *Nucleic Acids Res.* **43**, W39–W49 (2015).
85. Castro-Mondragon, J. A. *et al.* JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **50**, D165–D173 (2022).

Methods

Fly Husbandry

All crosses were maintained at 25°C. Transgenic and CRISPR lines were maintained as homozygous stocks. For live imaging of single MS2 allele crosses, homozygous males carrying the allele of interest were crossed with homozygous females bearing a *nos>MCP-eGFP-His2Av-mRFP* transgene. For live imaging of RNAi and overexpression experiments, homozygous males carrying *matα>gal4 ; nos>gal4, nos>MCP-eGFP-His2Av-mRFP* were crossed with homozygous females bearing the RNAi or overexpression transgene of interest. F1 virgin females were then crossed to males bearing the *sna*^{MS2} allele. For imaging of SnaIl^{Llama}, *yw; P{w[+mC] = EGFP-STOP-bcd}* (hereafter named *bcd*>GFP^{45}) was crossed to *His2A-RFPt/CyO*, followed by crossing of F1 virgin females to males bearing the *snail*^{Llama} allele. To image *Snail*^{Llama} / *snail-MS2* crosses, female *bcd>GFP*; *nos>MCP-RFPt* females were crossed to *snail-MS2* males, and the resulting F1 virgin females were crossed to males carrying the *snail*^{Llama} allele.

Generation of CRISPR knock-ins

Guide RNA sequences were selected using the CRISPR Optimal Target Finder site and cloned into pCFD3-dU6:3gRNA. The guide RNA sequences are listed in **Supplementary Table 1**. To create the *sna24xMS2* allele, a CRISPR recombination matrix comprised of a homology arm upstream of the 3' UTR, a 24xMS2 stem loop sequence (derived from Bertrand et al., 1998), a floxed 3xP3-dsRed selection cassette and a downstream homology arm. All genomic DNA was amplified using Phusion polymerase (Invitrogen), and the repair matrix was assembled in pBluescript-II SK(+). To create the *snail*^{Llama} allele, a CRISPR recombination matrix comprised of an 850bp genomic homology arm upstream of the stop codon of *sna* followed by a flexible linker and *Drosophila*-optimized GFP-targeting nanobody, the genomic Drosophila 3' UTR, a floxed 3xP3-dsRed selection cassette, and a 900bp genomic downstream homology arm. All genomic DNA was amplified using Phusion polymerase (Invitrogen), and the repair matrix was assembled in pBluescript-II SK(+). All matrices were sequenced prior to injection. Injections were performed by the Drosophila Transgenesis Facility (Centro de Biología Molecular Severo Ochoa, Madrid) and FlyORF (Zurich). All stocks are homozygous viable with no observable viability defects.

Live Imaging

Embryos were permitted to lay for 2 h prior to mounting for live imaging. Embryos were hand-dechorionated using tape and mounted on a hydrophobic membrane prior to oil immersion to prevent desiccation, followed by the addition of a coverslip. Live imaging of MS2 embryos was performed with an LSM 880 with Airyscan module (Zeiss). Z-stacks comprised of 30 planes with a spacing of 0.5 µm were acquired at a time resolution of 4.64s (*sna*^{MS2}) or 6.35s (*sog*^{MS2}) in fast Airyscan mode with laser power measured and maintained across embryos using a ThorLabs PM100 optical power meter (ThorLabs Inc.). All wild type background *sna*^{MS2} movies were performed with the following settings: GFP excitation by a 488-nm laser (6uW with 10x objective) and RFP excitation by a 561 nm were captured on a GaAsP-PMT array with an Airyscan detector using a 40x Plan Apo oil lens (NA = 1.3) and a 2.5x zoom on the ventral region of the embryo centered (±25 µm) on the presumptive ventral midline. Resolution was 640x640 pixels with bidirectional scanning. All *sog*^{MS2} movies were performed with the following settings: GFP excitation by a 488-nm (4.9uW with 10x objective) laser and RFP excitation by a 561 nm were captured on a GaAsP-PMT array with an Airyscan detector using a 40x Plan Apo oil lens (NA = 1.3) and a 2x zoom on the ventral/lateral region of the embryo including the ventral furrow. Time resolution was 6.35s and

resolution was 800x800 pixels with bidirectional scanning. All RNAi- and overexpression-related *sna*^{MS2} movies were performed with the following settings: GFP excitation by a 488-nm laser (7.7uW with 10x objective) and RFP excitation by a 561 nm were captured on a GaAsP-PMT array with an Airyscan detector using a 40x Plan Apo oil lens (NA = 1.3) and a 2.5x zoom on the ventral region of the embryo centered (±25 μm) on the presumptive ventral midline. Resolution was 640x640 pixels with bidirectional scanning. For all imaging conditions, Airyscan processing was performed using 3D Zen Black v3.2 (Zeiss).

Imaging of Snail^{Llama};His2A-RFP was performed with an LSM 880 (Zeiss). 18 Z-planes with a spacing of 1 µm were acquired with a time resolution of 42s/z stack. Movies were performed with the following settings: GFP excitation by a 488nm laser and RFP excitation by a 561nm laser captured on a GaAsP-PMT array using a 40x Plan Apo oil lens (NA = 1.3) at 1x zoom with resolution of 512x512 pixels. Laser power was measured and maintained across embryos using a ThorLabs PM100 optical power meter (ThorLabs Inc.).

Image Analysis for MS2-MCP movies

The intensity profile of the transcriptional sites imaged were extracted with a custom software developed in Python™ [77–79] that have been previously published[20] (SegmentTrackv4.0, https://github.com/ant-trullo/SegmentTrack_v4.0). However, for this study a post-processing tool was added (https://github.com/ant-trullo/SpotsFiltersTool). Transcription sites were infrequently resolved in the two sister chromatids by the detection algorithm, potentially confounding distinguishing between chromatids sisters and false detection. A parameter was defined as the ratio between the convex hull surface determined by the two spots and their actual size. For sister chromatids the ratio will be small (< 4) since the two spots are close, while a false detection will generally be far from the real spot and with a small volume, so the ratio will be large (>4). Some blinking activation could potentially be falsely discarded with an overly stringent criteria for sequential frames showing activity, so spots disappearing for a certain number of time frames were maintained following a user-defined thresholding process.

Single Molecule FISH

Embryos heterozygous for the allele of interest were fixed in 10% formaldehyde/heptane for 25 minutes with shaking followed by storage in methanol at -20°C as previously described [19]. smFISH probes were designed and produced with primary labelling using Quasar 670 by LGC Biosearch Technologies Inc. Probe sequences are listed in **Supplementary Table 2.** smiFISH probes were designed following a modification of the methodology in[80] and produced by IDT.

Embryos were prepared for smFISH as previously described[20]. Embryos were dehydrated with 2 × 5 min washes in 100% ethanol, followed by rehydration in PBT for 4 × 15 min and equilibration in 15% formamide/1 × SSC for 15 min. During equilibration, the probe mixture was prepared with a final concentration of 1 × SSC, 0.34 µg µL−1 E. coli tRNA (New England Biolabs), 15% formamide (Sigma), 5-µL probe, 0.2 µg µL−1 RNAse-free BSA, 2 mM vanadyl-ribonucleoside complex (New England Biolabs), and 10.6% dextran sulfate (Sigma) in RNAse-free water. The equilibration mixture was removed and replaced with probe mixture, and embryos were incubated overnight in the dark at 37°C with shaking. The following day, embryos were rinsed twice in equilibration mix and twice in PBT, followed by DAPI staining and three PBT washes before mounting in ProLong Gold mounting media (Life Technologies).

Fixed Imaging

Fixed sample imaging was performed on an LSM 880 with Airyscan module (Zeiss). Z-planes were acquired with 0.33 µm spacing to a typical depth of 25-30um from the apical surface of the embryo using laser

scanning confocal in Airyscan super-resolution mode with a zoom of 4.0. DAPI excitation was performed with a 405nm laser and Q670 with a 633nm laser, with detection on a GaAsP-PMT array coupled to an Airyscan detector. Airyscan processing was performed using 3D Zen Black v3.2 (Zeiss) prior to analysis. Embryos were staged based on membrane invagination.

Single Molecule FISH Analysis

To analyse smFiSH data we used a custom software developed in Python™ [77–79] that has been previously published[81]. Images were acquired in two channels, one for nuclei and the other for transcription, both in 3D (ZXY). The transcription channel was treated with a difference of Gaussian filter and the resulting image was thresholded; to find the optimal threshold value the algorithm performed a systematic study over a range of different threshold values that could be manually inspected and selected using a graphical user interface tool. The detected spots were composed of both transcriptional sites and single molecules that were further isolated into individual populations using a classifier and a visual tool for manual corrections when appropriate. The nuclei channel was pre-smoothed with a Gaussian filter and user-defined threshold individually for each z-slice to detect nuclei in 2D in each Z-frame. These Z-frames were than combined in 3D to have a first structure for nuclei. The following step was to find the smallest ellipsoid able to contain the detected 3D nucleus, which was then defined as the final nuclear volume. Finally, the intersection between the major axes of the ellipsoids was identified for each plane and used to simulate pseudo-cells with the Voronoi algorithm. Once the pseudo-cells were defined, the spatial position of transcription sites and single molecules was used to assign them to the appropriate 'cell'. As each transcription site and single molecule has an associated intensity, the equivalent number of mRNA molecules was then calculated for each transcription site.

Data Analysis of snail^{Llama}

Visualization and analysis of the time series data was performed using custom software developed in Python™[77–79] enabled by a graphical user interface (NucleiTracker3D). Raw data consisting of a two channel TZXY series with Snail^{Llama}-GFP intensity and His2A-mRFP as a reference nuclear marker. The reference nuclear channel was pre-smoothed using a Gaussian filter, and then thresholded with an Otsu algorithm. The resulting connected components were labelled in 3D treated with a 3D watershed algorithm to separate touching nuclei. Hyper-segmented nuclei were recognized using a classifier algorithm previously trained to identify 'sub-nuclear' fragments and combine them with neighbouring nuclei, privileging combinations with the fewest sub-components. Segmented nuclei were tracked by sequentially projecting the *t-1* timepoint nuclear mask onto the frame of timepoint *t* and tagging each nuclei *n* with the most coincident nuclear tag of the projected mask using a median filter, as nuclear motion between timepoints *t-1* and *t* was less than the nuclear radius in XYZ. This 3D-tracked nuclear mask was then projected on the Snail^{Llama}-GFP channel to retrieve the nuclear GFP fluorescence. For each time point the average nuclear out-of-pattern and in-pattern GFP signal intensity was retrieved. An enrichment ratio was then calculated by dividing the in-pattern by the out-of-pattern GFP signal intensity in a time-dependent manner[82].

Determining the dwell time from autocorrelation

The signal autocorrelation function is defined as $R(t,t') = Cov\big(x(t), x(t')\big)$, where $x(t)$ is the single site MS2 signal. For a stationary MS2 signal, this function depends only on $\tau = t - t'$ according to the relation:

$$R \sim H(\tau + d) - 2H(\tau) + H(\tau - d) \quad (1)$$

where d is the dwell time, $H(x) = -x\theta(-x)$ and

$$\theta(x) = \begin{cases} 1 & if\, x \geqslant 0 \\ 0 & if\, x < 0 \end{cases} \quad (2)$$

is the Heaviside function. A derivation of (1) has been previously published[83]. For the purposes of this research, the stationary signal from nuclear cycle 13 was used to fit the dwell time (**Supplementary Figure 2**).

Deconvolution analysis of the data using BurstDeconv

We use BurstDeconv[18] to obtain, for each transcription site, the sequence of processive transcription initiation events. We then analyse the distribution and the mean of the waiting time (t) separating successive transcription events. Because the process is time inhomogeneous, we perform the analysis in time windows of two types:

a) Constant repression windows obtained by the BCPD method.

Using BCPD we determine the onset of repression and perform the analysis on the last part of the movie (from the onset of repression). The distribution of the waiting time tau, in this region, provides the information needed for identification of three state and two state transcription bursting models according to methods published previously[18].

b) Narrow, moving window for $p_{ON}\, k_{ini}$ estimates.

We also define a narrow moving window centered on successive time frames and consider all the waiting times from all transcription sites active in the moving window. The width of the windows is 5-8 frames, i.e. 22.7-36.3 seconds for *sna*. This width is enough for including a sufficiently large number of waiting times for an accurate estimate of the mean $<\tau>$ but not enough for an accurate estimate of the distribution of τ needed for the application of BurstDeconv in each window. For this reason, we utilize the mean, rather than the distribution of waiting times. As discussed previously[27], there is a general formula relating the mean $<\tau>$ and $p_{ON}\, k_{ini}$, valid for all finite state Markov models, irrespective of their number of states. We reproduce here the reasoning leading to this formula. The mean number of transcription events on an interval $[0,T]$ is $T/<\tau>$. The same number is equal to $T\, p_{ON}\, k_{ini}$, because in the state ON, the promoter initiates with intensity k_{ini} and the total time spent in ON is $T\, p_{ON}$. It follows that:

$$< \tau > \; = \frac{1}{p_{ON} k_{\mathrm{ini}}} \quad (3)$$

In order to compute $<\tau>$ we combine the nuclei coming from different movies and for each narrow moving window we sum over all the waiting times between all initiation events that have contributed to the signal within the time window, even if they are not comprised in this window. We compute the uncertainty bounds of the mean for each movie independently by using the central limit theorem with

a 95% confidence interval, and then plot the min lower bound and the max upper bound over all the movies.

Bayesian Change Point Detection

The BCPD method introduced in[26] is used to determine the onset of repression by determining the probability of having a change point denoting a sudden change in the parameters that generate the data. This method learns from data the distribution of the run length $r(t)$ which is the time since the most recent change point. At each time step, $r(t)$ increases by 1 if there is no change in the distribution, or returns to zero when there is a change with a certain probability. It is based on a recursive message-passing algorithm for the joint distribution of observations and run lengths. The algorithm assumes that: i) the MS2 signal follows a normal distribution with unknown mean and variance, and ii) the run length advances without memory, according to a geometric distribution.

An illustration of the BCPD method is given in Supplementary Figure 4. For discretized times $t = 1,2,\ldots$ the run length $r(t)$ is defined by the following relation:

$$r(t+1) = \begin{cases} r(t)+1 & \text{if no change of parameters} \\ 0 & \text{if there is a change} \end{cases} \tag{4}$$

The BCPD method computes the conditional probability of the run length, given the observed values of the signal (**Supplementary Figure 4A**). The prior of this distribution is learned from the data. We have tested the method using synthetic data generated using the Gillespie algorithm and a two-state telegraph model. The model includes an RNA and a protein pool and considers self-repression by considering that the kinetic parameters depend on the protein level according to decreasing and increasing Hill functions, respectively (**Supplementary Figure 4B**). We have used the promoter transcription initiation events to compute a synthetic MS2 signal for each site. The traces generated by this model show change points corresponding to onset of the repression. These change points are correctly detected by the method as shown visually in **Supplementary Figure 4** where we confirm that the switching parameters are stable after the change point is found. Since this is a Bayesian approach, the change point is identified with a probability. A threshold was imposed such that the probability p must be > 0.8 for the change to be retained (**Supplementary Figure 4C** middle vs bottom panel). Three further criteria for detecting the change point were applied to change point detection on experimental data. First, due to the noisy nature of the data, a smoothing criterion was applied. The smoothing process involves convoluting the data with a box filter kernel of size 2 so that the rapid fluctuations in the experimental data are attenuated in the smoothed data. Second, a changepoint detection window was imposed such that the selected changepoint was the first change point to occur after 60% of the maximum in order to eliminate changepoints during the post-mitotic activation period. Third, a minimal filtering criterion was applied to individual traces. The filtering process includes the elimination of signals in the attenuated part that are either 0 everywhere or have less than 5% of the mean number of initiation events before the identified changepoint. Our implementation of the BCPD algorithm is available via github at:

https://github.com/mariadouaihy/BCPD_inhomogeneous_transcriptional_signal.

The part of the MS2 signal after the repressor onset checkpoint in nc14 and the full MS2 signal in nc13 were then used for inferring promoter models for the repressed and active phases, respectively, using BurstDeconv[18]. Model selection was performed using three fitting scores: the objective function, the

confidence interval of the empirical survival function and the Kolmogorov-Smirnov test comparing the empirical and predicted distribution of waiting time between successive transcription events[18].

qPCR analysis

To test changes in expression of pause-related genes, 0–2 h embryos were homogenized in Trizol (Invitrogen) and RNA was extracted as directed by the manufacturer. Reverse transcription was performed using Superscript IV (Invitrogen) with random hexamers. Measurements were performed in biological and technical triplicate. qPCR analysis was performed using LightCycler 480 SYBR Green I Master Mix (Roche) using primers listed in **Supplementary Table 5**. Analysis was performed using Microsoft Excel and Prism (Graphpad v9.1.1).

Deterministic model of repression based on a Hill function

The key to this model is the nonlinear relation between the repressor protein concentration and the mRNA production rate (the Hill function model):

$$p_{ON}\, k_{\text{ini}}\left([\text{protein}]\right) = V \frac{[\text{protein}]^n}{\theta^n + [\text{protein}]^n} \tag{5}$$

where [protein] is the protein concentration, V is the maximal mRNA production rate, ϑ is the threshold protein concentration at half-maximal repression, and n is the Hill index.

If the protein level and the product $p_{ON}k_{ini}$ (also described in text as the inverse of $<\tau>$, see Eq.(3)) are both known, Eq.(5) can be used to estimate the parameters V, θ, and n, with no assumption on the identity of the repressor or mechanism of repression.

Minimal stochastic model of repression

The deterministic model discussed in the previous section does not account for the observed single nuclei transcriptional bursting, nor for temporal variability. In order to model these phenomena, we developed a stochastic model of repression. The stochastic and the deterministic models are consistent with regard to their prediction of average properties of the population.

To be able to understand the impact of cooperativity on repression, we have developed a simple model of the transcriptional process under repression, in which the transition rates between discrete promoter states are modulated by the repressor in a Hill-dependent manner. The objective of this mathematical model is to 1) see if our model is able to re-obtain the same repression behaviour using a Hill function and cooperativity and 2) to predict the effect of cooperativity on coordinated repression synchrony (by scanning over different ranges of Hill coefficient and theta's).

In order to build this minimal model, we build on the results of the deconvolution analysis of unrepressed and fully repressed phases in nc13, and nc14 respectively. According to these results, the ON->OFF1 transition rate (k_1^m) is the same in the repressed and unrepressed phases (**Supplementary Table 4**). We therefore decided to keep this rate constant in the model. All the other rates are *a priori* repressor dependent.

The transcriptional process is described as a three state Markovian model and governed by the following set of chemical reactions:

$$OFF_1 \xrightarrow{k_1^p([\text{Snail}])} ON \tag{6}$$

$$ON \xrightarrow{k_1^m} OFF_1 \tag{7}$$

$$OFF_2 \xrightarrow{k_2^p([\text{Snail}])} ON \tag{8}$$

$$ON \xrightarrow{k_2^m([\text{Snail}])} OFF_2 \tag{9}$$

$$ON \xrightarrow{k_{\text{ini}}} ON + mRNA \tag{10}$$

where the first OFF state (*OFF1*) represents the PIC assembly/disassembly status. The second OFF state represents a repressor binding/unbinding. The state interpretation is based on the result of the parametric fit, see Results Section and below. The concentration of the protein snail ([Snail]) is obtained from our experimental data using the LlamaTag system.

The Hill-like dependence of parameters $k_1^p([Snail])$, $k_2^p([Snail])$ and $k_2^m([Snail])$ is given by the equations (we use increasing and decreasing Hill functions for the ON-> OFF, and OFF-> ON transitions, respectively) :

$$k_1^p([Snail]) = (k_1^p)_{rep} + ((k_1^p)_{act} - (k_1^p)_{rep})\frac{th^{nh}}{th^{nh} + [Snail]^{nh}} \tag{11}$$

$$k_2^p([Snail]) = (k_2^p)_{rep} + ((k_2^p)_{act} - (k_2^p)_{rep})\frac{th^{nh}}{th^{nh} + [Snail]^{nh}} \tag{12}$$

$$k_2^m([Snail]) = (k_2^m)_{act} + ((k_2^m)_{rep} - (k_2^m)_{act})\frac{[Snail]^{nh}}{th^{nh} + [Snail]^{nh}} \tag{13}$$

where $(k_1^p)_{act} > (k_1^p)_{rep}$, $(k_2^p)_{act} > (k_2^p)_{rep}$ are OFF-> ON rates, $(k_2^m)_{act} < (k_2^m)_{rep}$ are ON->OFF rates in active and repressed phases, respectively.

Since the transcription of *snail* goes from a simple two-state model in the active phase to a three-state model in the repressed phase, we set the switching parameter $(k_2^m)_{act}$ to zero, which means that this transition does not happen in the active phase.

The switching parameters k_{ini}, k_1^m, $(k_1^p)_{act}$, $(k_1^p)_{rep}$, $(k_2^p)_{rep}$, $(k_2^m)_{rep}$ were chosen equal to the already estimated values $(k_{ini})_{nc13\,\&\,rep,nc14}$, $(k_1^m)_{nc13\,\&\,rep,nc14}$, $(k_1^p)_{nc13}$, $(k_1^p)_{rep,nc14}$, $(k_2^p)_{rep,nc14}$, $(k_2^m)_{rep,nc14}$, in the active nc13 and repressed nc14 regimes (**Supplementary Table 4**, 2 state and 3 state non-sequential model).

The model remains with three free parameters *nh*, *th*, $(k_2^p)_{act}$that were fitted to the $p_{ON}{\cdot}k_{INI}$ and the distribution of repression onset times data. The result of the fit is given in the table below:

Symbol	Value	Justification
k_{ini}	0.351 s^{-1}	nc13 & nc14 deconvolution fit
k_1^m	0.007 s^{-1}	nc13 & nc14 deconvolution fit
$(k_1^p)_{act}$	0.011 s^{-1}	nc13 deconvolution fit
$(k_1^p)_{rep}$	0.001 s^{-1}	nc14 M2 BCPD deconvolution fit
$(k_2^m)_{act}$	0 s^{-1}	No OFF_2 in active state
$(k_2^m)_{rep}$	0.1125 s^{-1}	nc14 M2 BCPD deconvolution fit
$(k_2^p)_{rep}$	0.045 s^{-1}	nc14 M2 BCPD deconvolution fit
$(k_2^p)_{act}$	0.06 s^{-1}	free parameter, fit of the minimal stochastic model

The Gillespie algorithm was used to numerically simulate the model. To account for post-mitotic lag in transcriptional reactivation unique to each nucleus, the simulated model distribution of lag times was sampled from the experimental post-mitotic lag duration of the *snail* gene. The time points where the algorithm jumped to the elongation state *El* were used as the RNA Polymerase II initiation times. A synthetic MS2 signal was computed for each simulation using the parameters of the sna^{MS2} gene such that the simulated data can be compared with the experimental results in nc14.

In order to test the dependence of the repression onset time on the Hill function parameters, 154 value pairs of $\{(k_2^p)_{act}, nh, th\}$ were benchmarked with each simulation run for 484 sample traces and a duration of 30.93 min to match the time window studied for the expression of the experimental sna^{MS2} gene. Each simulation was then run for each set of $\{(k_2^p)_{act}, nh, th\}$ through the BCPD algorithm to obtain the repression onset time. No smoothing was applied to the signal used in the stochastic model since no noise was added. We then chose the best values of $\{(k_2^p)_{act}, nh, th\}$ that fit the data according to the sum of squared error between $p_{ON}\, k_{ini}$ of the data and the simulations. **Figure 6D-E** are the results of the best fit simulations with given fixed parameters as indicated. **Figure 6G** is the result of an additional 182 simulations with a fixed $(k_2^p)_{act}$0.06 s^{-1} from the optimal fit parameter and free $\{nh, th\}$. The repression onset time distribution was not fitted, therefore the good agreement between predicted and experimental distributions validates the stochastic model.

Snail binding site identification

Snail ChIP data was obtained from previously published data (GSE68983)[44]. To identify potential transcription factor binding sites, we employed the FIMO (Find Individual Motif Occurrences) tool[84] with Snail motifs obtained from the JASPAR database[85]. The significance threshold for motif matches was set at $p<1e-3$. Snail binding sites were queried in the local neighbourhood of the highest (>500) ChIP signal for Snail.

Pimmett_Figure 1

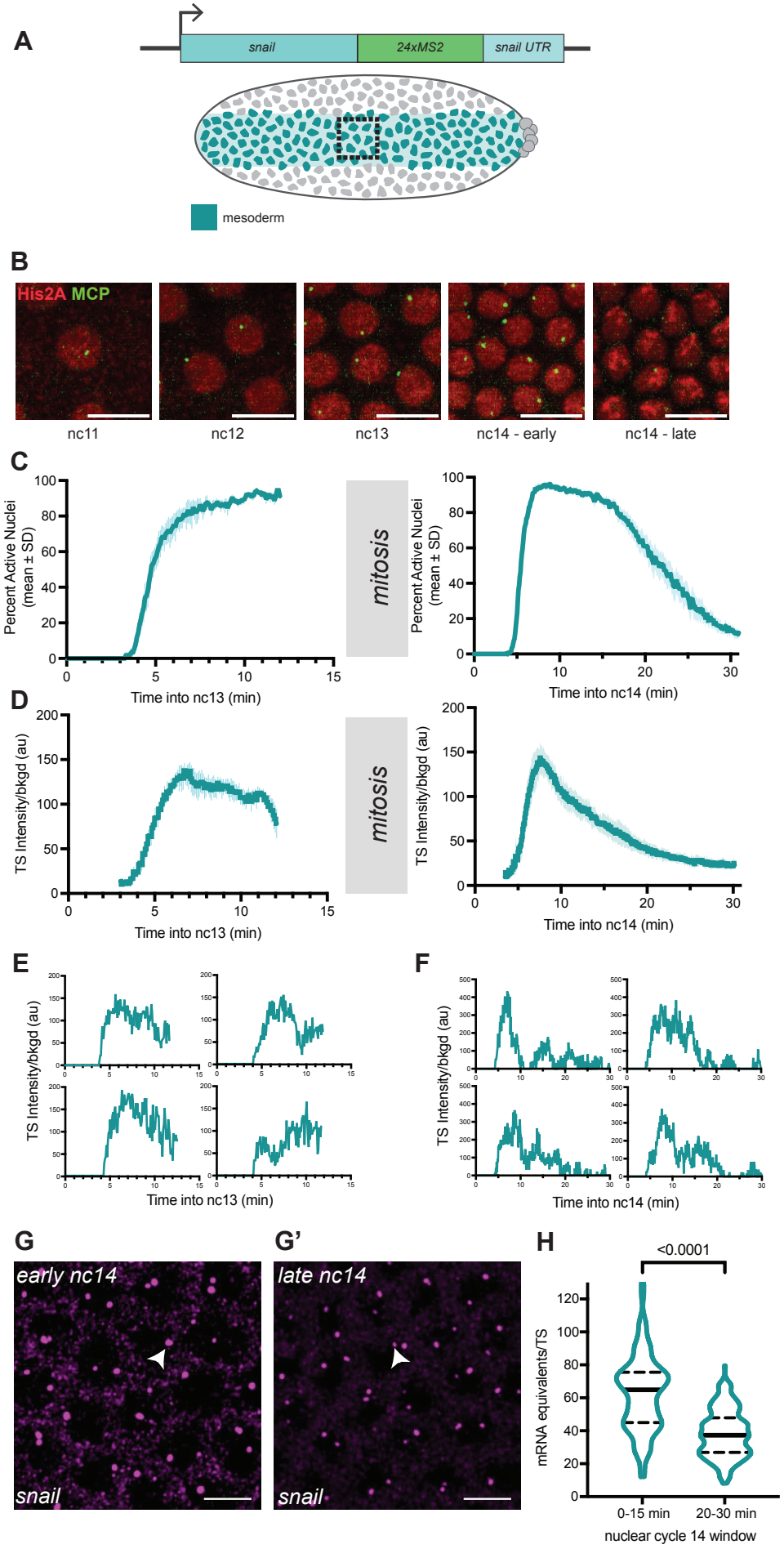

Figure 1: Live imaging of *snail* endogenous transcription demonstrates two regimes including a coordinated rapid transcriptional repression.

A) Schematic view of *snail*^{MS2} allele (above) and expression domain in the embryonic mesoderm (below, teal). The box indicates the restricted imaging area. B) Maximum intensity Z-projection of representative nuclei showing MS2/MCP-GFP-bound puncta and nuclei (His2A-RFP) in sequential nuclear cycles. Scale bar represents 10 µm. C) Instantaneous activation percentage (mean ± SEM) curves of ventral nuclei during nc13 and the first 30 min of nc14. Time zero is from anaphase during nc12-nc13 and nc13-nc14 mitosis respectively. D) Fluorescence intensity of actively transcribing nuclei (mean ± SEM) for *sna*^{MS2} nuclei during nc13 and the first 30 minutes of nc14. Time zero is from anaphase during nc12-nc13 and nc13-nc14 mitosis respectively. E-F) sample single nucleus fluorescence traces in nc13 (E) and the first 30 minutes of nc14 (F). Time zero is from anaphase during nc12-nc13 and nc13-nc14 mitosis respectively. G-G') Maximum intensity Z-projections of transcription site-containing volume in early (G) and late (G') *Drosophila* mesoderm showing endogenous *snail* expression with single molecule fluorescent *in situ* hybridization (smFISH) in nuclear cycle 14. Large puncta (arrowheads) represent nascent transcription. Scale bar represents 10 µm. H) Quantification of endogenous *snail* transcription site intensity in early and late nuclear cycle 14 embryos via smFISH.

Statistics: *sna*^{MS2} nuclear cycle 13: N=4 embryos, n=274 nuclei; *sna*^{MS2} nuclear cycle 14: N=6 embryos, n=484 nuclei. smFISH: N=3 embryos for early and late timepoints as determined by membrane invagination. Significance is indicated using a one-tailed Kruskal–Wallis test.

Pimmett_Figure 2

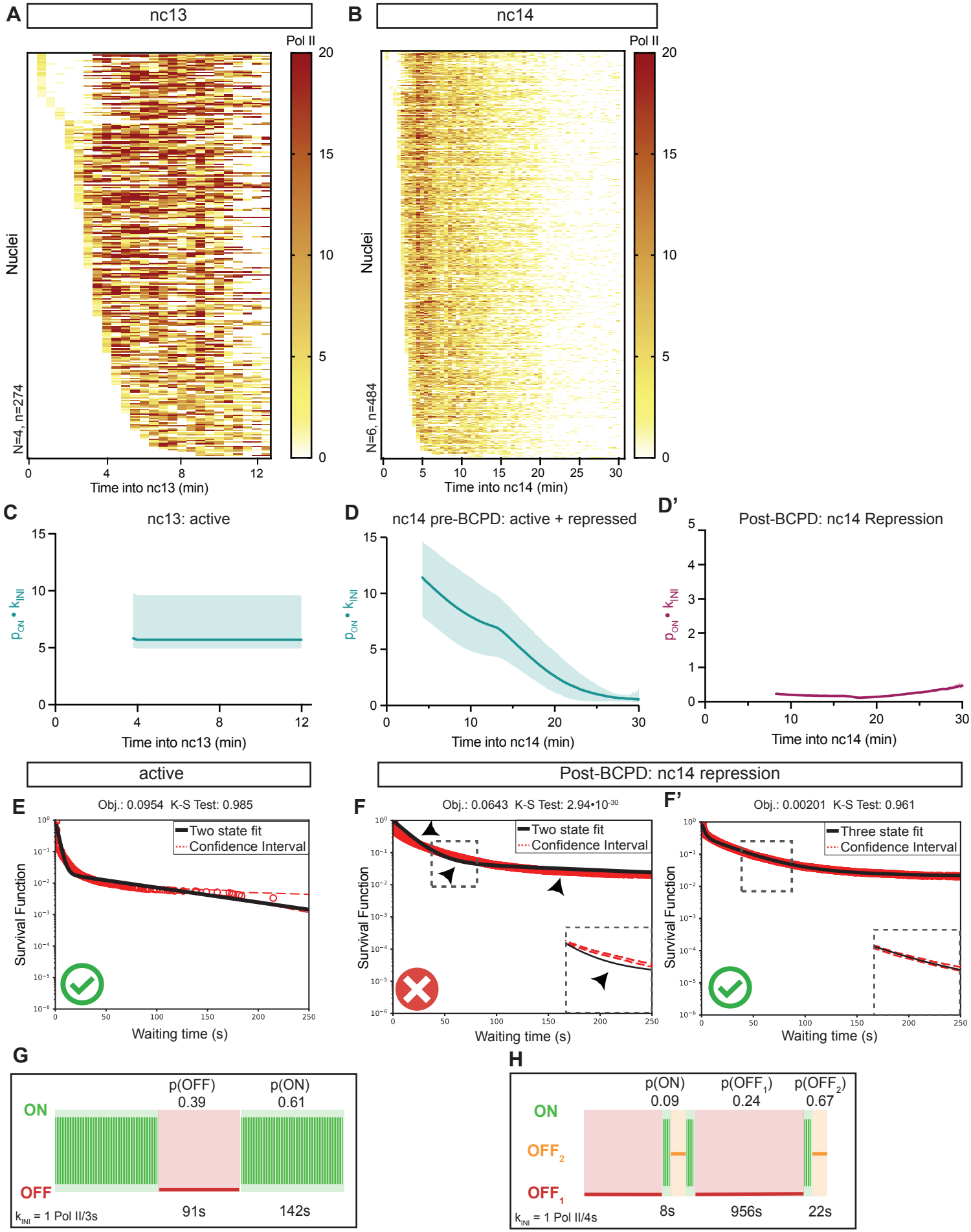

Figure 2: Active repression is associated with an extra promoter state but does not affect polymerase initiation rates.

A-B) Heatmap showing the number of polymerase initiation events for *snail* in nuclear cycle 13 (A) and nuclear cycle 14 (B) as a function of time. Each row represents one nucleus, and the number of Pol II initiation events per 30 s bin is indicated by the bin color. C-D) Kinetic parameter stability as a function of time for *snail* transcription expressed as the product of the probability to be active (p_{ON}) and the RNA polymerase II initiation rate (k_{INI}) in nuclear cycle 13 (C), the first 30 minutes of nuclear cycle 14 (D), and nc14 repression following Bayesian Change Point Detection (BCPD). Confidence intervals are provided as the minimum/maximum value across individual movies (C, D) or the minimum/maximum of the combined dataset (D'). E) Survival function of the distribution of waiting times between polymerase initiation events (red circles) for *snail* active phase (nuclear cycle 13) transcription with the two-exponential fitting of the population estimated using the Kaplan–Meyer method (black line). The dashed lines indicate 95% confidence interval estimated based on Greenwood's formula. The Kolmogorov-Smirnov (K-S) test verifies that the empirical waiting times come from the predicted distributions (small p-values mean that the null hypothesis is not valid). The objective function (Obj.) comes from the least squares error of the survival function (large values mean the error is high). A green check indicates accepted fitting. F-F') Survival function of the distribution of waiting times between polymerase initiation events (red circles) for the repressed transcriptional phase with the two-exponential (F, inset) and three-exponential (F', inset) fitting of the population estimated using the Kaplan–Meyer method (black line). The dashed lines indicate 95% confidence interval estimated based on Greenwood's formula. The Kolmogorov-Smirnov (K-S) test verifies that the empirical waiting times come from the predicted distributions (small p-values mean that the null hypothesis is not valid). The objective function (Obj.) comes from the least squares error of the survival function (large values mean the error is high). A red cross or green check indicates a rejected and accepted fitting respectively. G-H) Representation of estimated bursting dynamics for *snail* expression in the active phase (G) and repressed phase (H). Permissive ON state durations are depicted in green and inactive OFF states in red and orange, and probabilities of each state shown above. (See also **Supplementary Table 2**).

Statistics: *snail*^{MS2} active phase: N=4 embryos, n=188 nuclei; repressed phase: N=6 embryos, n=448 nuclei.

Pimmett_Figure 3

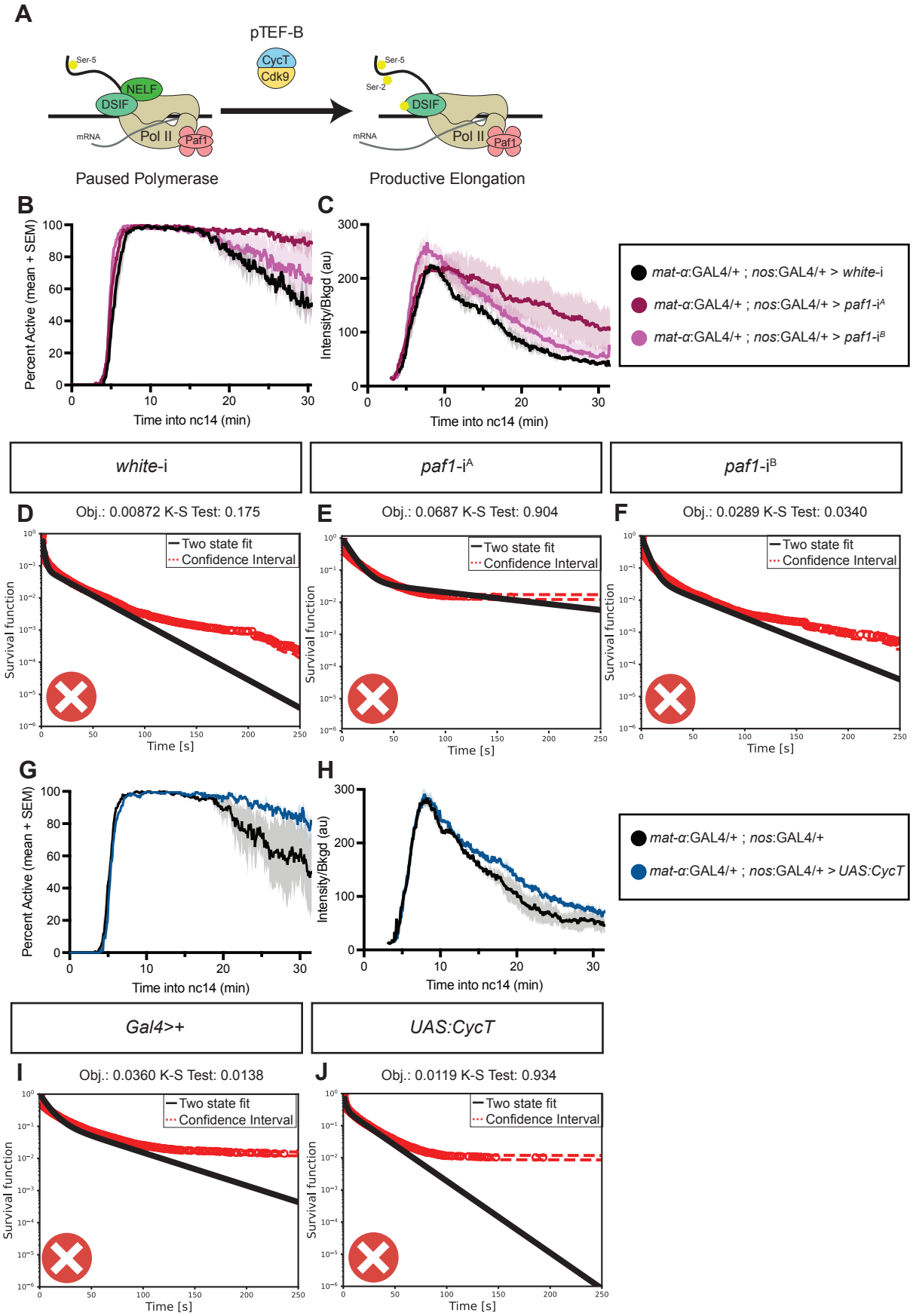

Figure 3: Promoter-proximal polymerase pausing has limited correlation with transcriptional repression.

A) Schematic of key components of promoter-proximal polymerase pausing imposition and release. B) Instantaneous activation percentage (mean ± SEM) curves of *snail* transcription in ventral nuclei during the first 30 min of nc14 for *white*-i control, *paf1*-i^A and *paf1*-i^B. Time zero is from anaphase during nc13-nc14 mitosis. C) Fluorescence intensity of *snail* transcription in actively transcribing nuclei (mean ± SEM) during the first 30 minutes of nc14 *white*-i control, *paf1*-$i^A$and *paf1*-i^B. Time zero is from anaphase during nc13-nc14. D-F) Survival function of the distribution of waiting times between polymerase initiation events (red circles) for the attenuated transcriptional phase with the two-exponential fitting of the population estimated using the Kaplan–Meyer method (black line) in the *white*-i control (D), *paf1*-i^A (E) and *paf1*-i^B(F). The dashed lines indicate 95% confidence interval estimated based on Greenwood's formula. The Kolmogorov-Smirnov (K-S) test verifies that the empirical waiting times come from the predicted distributions (small p-values mean that the null hypothesis is not valid). The objective function (Obj.) comes from the least squares error of the survival function (large values mean the error is high). A red cross indicates a rejected fitting. G) Instantaneous activation percentage (mean ± SEM) curves of *snail* transcription in ventral nuclei during the first 30 min of nc14 for GAL4>+ control and *UAS:CycT*. Time zero is from anaphase during nc13-nc14 mitosis. H) Fluorescence intensity of *snail* transcription in actively transcribing nuclei (mean ± SEM) during the first 30 minutes of nc14 for GAL4>+ control and *UAS:CycT*. Time zero is from anaphase during nc13-nc14. I-J) Survival function of the distribution of waiting times between polymerase initiation events (red circles) for the attenuated transcriptional phase with the two-exponential fitting of the population estimated using the Kaplan–Meyer method (black line) in the GAL4>+ control (I) and *UAS:CycT (J)*. The dashed lines indicate 95% confidence interval estimated based on Greenwood's formula. The Kolmogorov-Smirnov (K-S) test verifies that the empirical waiting times come from the predicted distributions (small p-values mean that the null hypothesis is not valid). The objective function (Obj.) comes from the least squares error of the survival function (large values mean the error is high). A red cross indicates a rejected fitting.

Statistics: *white*-i > *sna*^{MS2}: N=2 embryos, n=183; *paf1*-i^A > *sna*^{MS2}: N=2 embryos, n=126 nuclei; *paf1*-i^B > *sna*^{MS2}: N=2 embryos, n=177; Gal4>+ > *sna*^{MS2}: N= 2, n=191; *UAS:CycT* > *sna*^{MS2}: N=2, n=154. All numbers indicate repressed phase only.

Pimmett_Figure 4

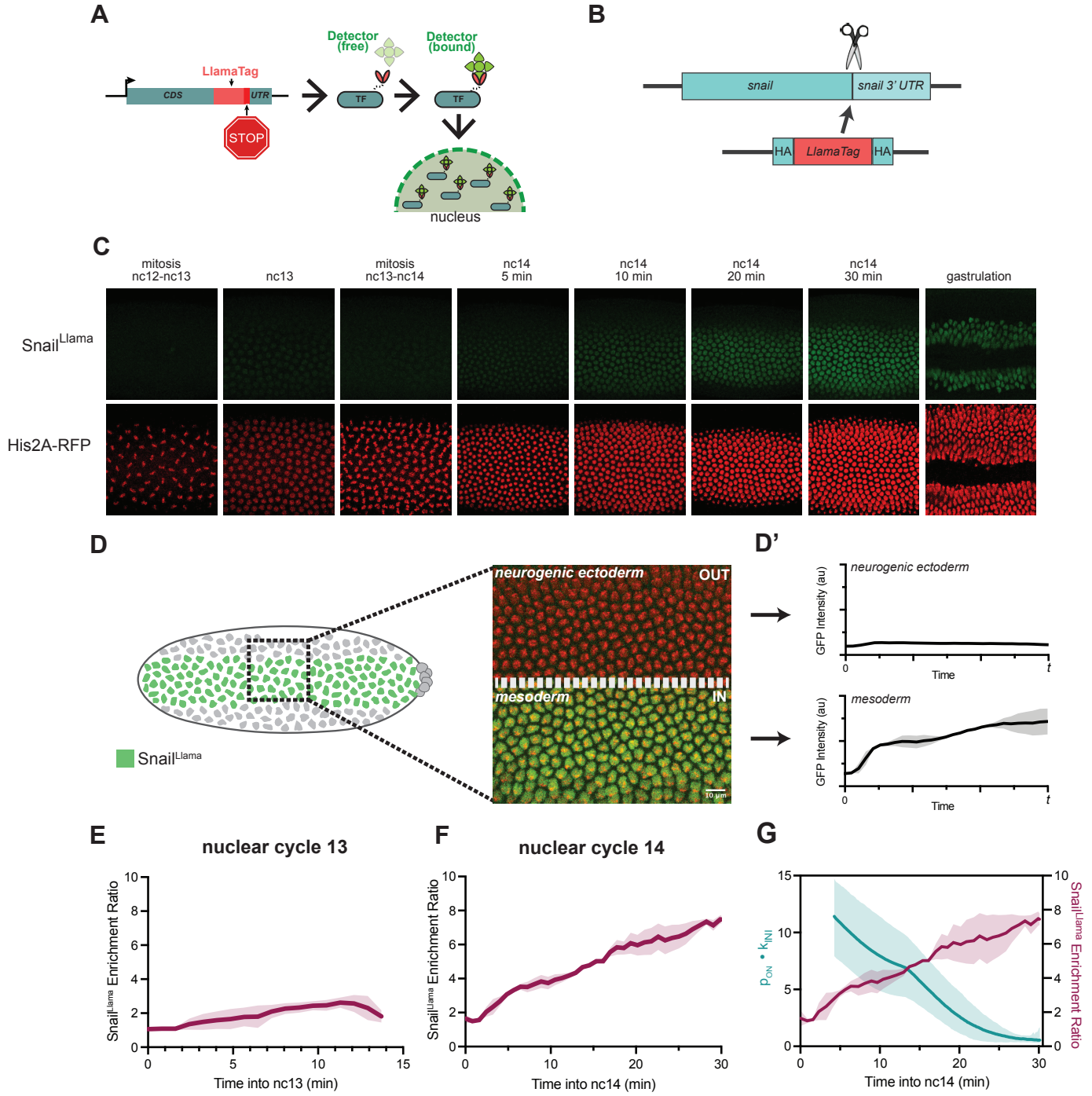

Figure 4: Capturing Snail protein dynamics and its correlation with transcription repression in living embryos.

A) Schematic demonstrating the principle of the LlamaTag system. B) Schematic indicating CRISPR-mediated genome editing of the endogenous *snail* locus to introduce the LlamaTag nanobody. C) Representative maximum intensity Z-projections of Snail^{Llama} nuclear signal (top) and nuclei labeled with His2A-RFP (bottom) from nuclear cycle 12-13 mitosis until gastrulation. D-D') Schematic of Snail^{Llama} analysis showing the imaging window on the embryo, with the mesoderm/neurogenic ectoderm boundary indicated. The nuclear GFP intensity for both the mesoderm and neurogenic ectoderm was quantified as a function of time (D'). E-F) Nuclear Snail^{Llama} enrichment ratio in nc13 (E) and the first 30 minutes of nc14 (F) calculated as the mean nuclear GFP signal in the mesoderm relative to the neurogenic ectoderm, expressed as mean ± SD. G) Transcription kinetic parameters expressed as a function of time (teal) and Snail protein enrichment ratio (red) for the first 30 minutes of nc14. Kinetic parameters expressed as mean ± upper/lower bounds and protein expressed as mean ± SD.

Statistics: *snail*^{MS2} repressed phase: N=6 embryos, n=448 nuclei; Snail^{Llama}: N=3 embryos.

Pimmett_Figure 5

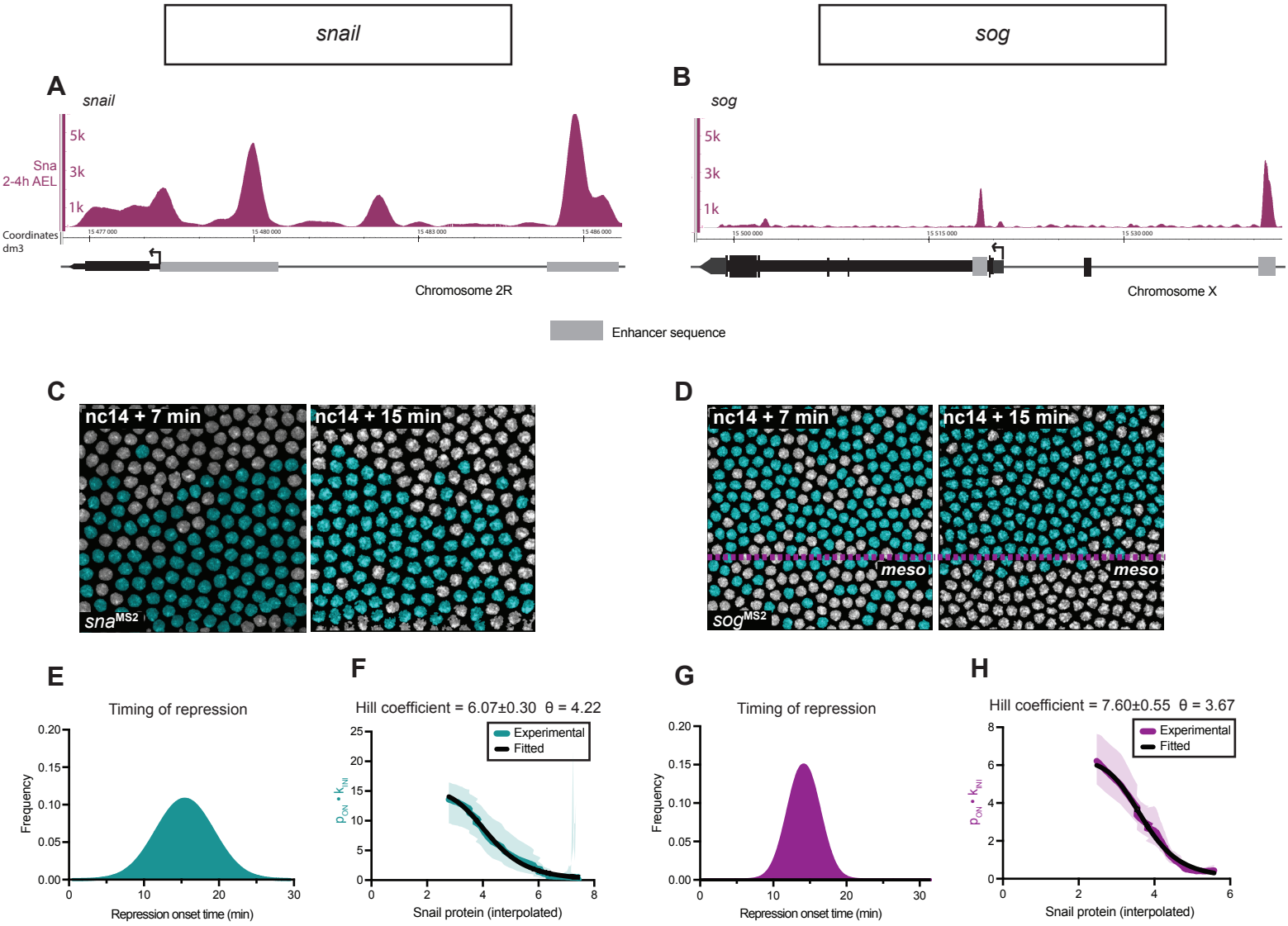

Figure 5: Quantifying the degree of coordination and cooperativity for Sna-mediated repression

A-B) Snail ChIP-seq enrichment profiles at 2-4h AEL for the endogenous *sna* (A) and *sog* (B) loci (above), with genomic features indicated (below). Sna binding sites in ChIP-enriched regions[44] identified with FIMO indicated with green bars. C-D) False-coloured projections from live imaging of *snail*^{MS2} (C) and *sog*^{MS2} (D) embryos, with active nuclei indicated in teal and inactive in grey (**Supplementary Movies 1,8**). E) Distribution of switching times for initiation of stable repression in nuclear cycle 14 for *sna*^{MS2} (E) as determined using Bayesian Change Point Detection. F) Hill fitting for *sna*^{MS2} with the Hill coefficient and θ indicated above. θ represents the repressor concentration reducing transcription intensity to half. The uncertainty interval of the Hill coefficient was computed as the standard deviation of the Hill coefficient coming from the fitting of different *sna*^{MS2} movies. Coloured line indicates experimental kinetic parameter stability as a function of time expressed as the product of the probability to be active (p_{ON}) and the RNA polymerase II initiation rate (k_{INI}) in the first 30 minutes of nuclear cycle 14 (mean ± upper/lower bounds). Black line indicates Hill function fit. G) Distribution of changepoint times for initiation of stable repression in nuclear cycle 14 for *sog*^{MS2} (G) as determined using Bayesian Change Point Detection. H) Hill fitting for *sog*^{MS2} with the Hill coefficient and θ indicated above. The uncertainty interval of the Hill coefficient was computed as the standard deviation of the Hill coefficient coming from the fitting of different *sog*^{MS2} movies. Coloured line indicates experimental kinetic parameter stability as a function of time expressed as the product of the probability to be active (p_{ON}) and the RNA polymerase II initiation rate (k_{INI}) in the first 17 minutes of nuclear cycle 14 (mean ± upper/lower bounds). Black line indicates Hill function fit.

Statistics: *sna*^{MS2} nuclear cycle 14: N=6 embryos, n=448 nuclei; *sog*^{MS2} nuclear cycle 14: N=3, n=141 nuclei. All numbers indicate repressed phase only.

Pimmett_Figure 6

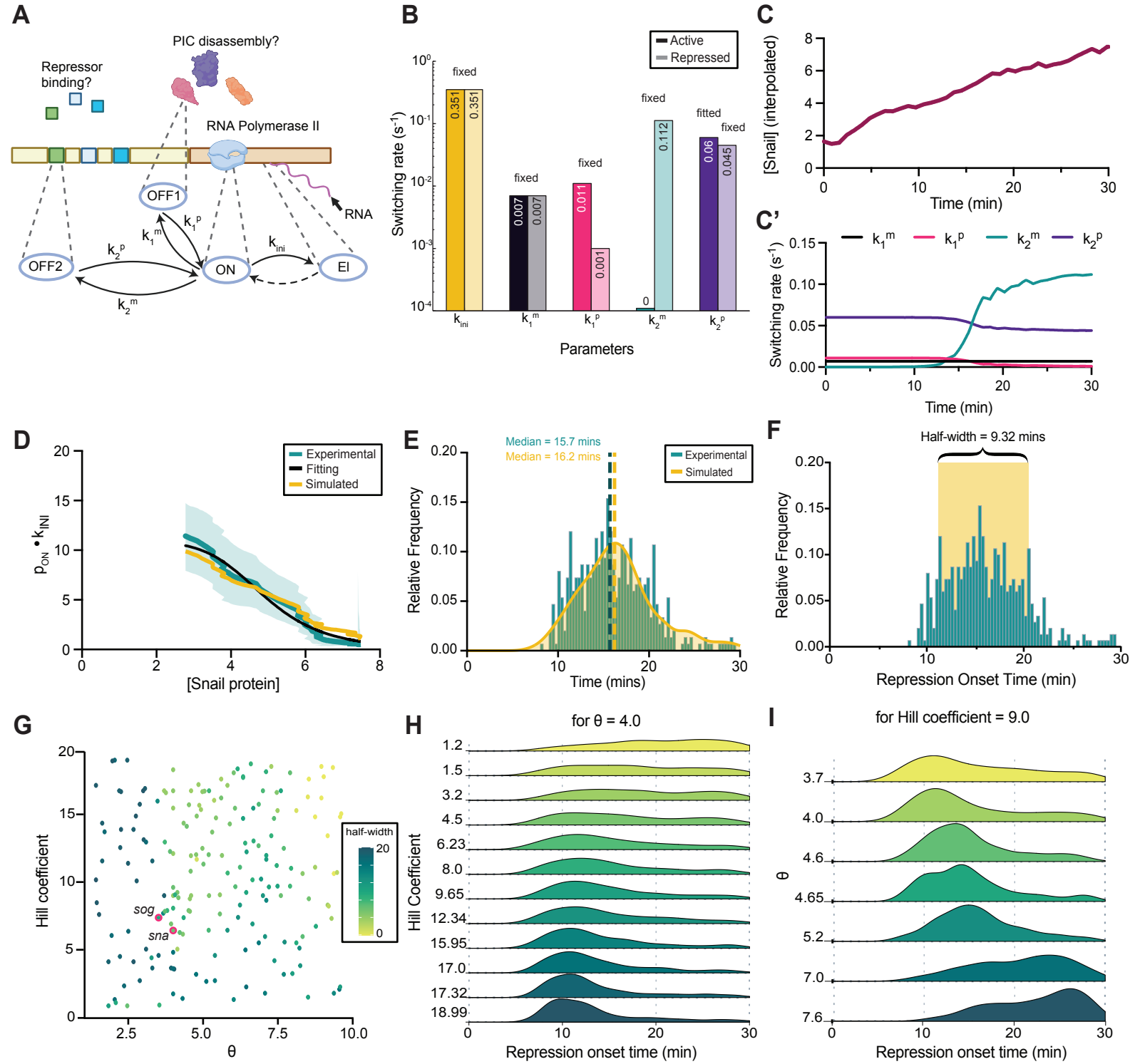

Figure 6: A minimal stochastic model reveals the impact of cooperativity on transcriptional repression.

A) Scheme of the Markovian model. The transitions to the first OFF state (OFF1) are interpreted as PIC assembly/disassembly, and the transitions to the second OFF state (OFF2) as repressor binding/unbinding. B) Bar plot showing the changes in kinetic parameters between activation and repression phases of the model. C-C') Parameters of the stochastic model showing evolution of protein level (C) and switching rates (C') as a function of time. D) Best fit of modelled $p_{ON} \cdot k_{ini}$ (yellow) resulting from simulations parallels the experimentally-derived values (teal) over 30 minutes of simulated transcription. E) Comparison of the BCPD-derived repression onset times for the experimentally-derived data (teal) and kernel density estimate for simulated data from C (yellow). Median repression onset time indicated by dashed lines. F) Half-height width for the distribution of the repression onset time for *sna*^{MS2} experimental data. G) Parameter scanning of the Hill coefficient and θ (repressor concentration reducing transcription intensity to half) for half-height width using the optimal fit value of k_p^2. Fill colour represents half-width of the repression onset time distribution. Modelled half-height width for experimental *sna* and *sog* θ and Hill coefficient are indicated (red). H) Distribution of simulated repression onset time for various Hill coefficients at a fixed θ of 4. Hill coefficient for simulation is indicated on the left. I) Distribution of simulated repression onset time for various θ values at a fixed Hill coefficient of 9.0.

Supplementary Figure 1: Profiling *sna*^{MS2} allele. A) *sna*^{MS2}/+ heterozygotes showing coincident labelling of *sna*^{MS2} allele and endogenous *snail* sequence at the transcription site (inset). Scale bar represents 10 μm. B) Sample membrane invagination (arrowheads) indicating partitioning of embryos into early (left) and late (right) nc14. C) Instantaneous activation profiles for individual *sna*^{MS2} movies during nc14. D) Average intensity of actively transcribing nuclei for individual *sna*^{MS2} movies (E1 to E6).

Supplementary Figure 2: Autocorrelation of the *sna* transcription signal in nuclear cycle 13. A) Sample trace demonstrating autocorrelation of a *snail* transcription trace in nc13. B) Dwell time of the *snail*^{MS2} signal in nuclear cycles 13 and nc14 compared to the theoretical prediction for a polymerase speed of 25bp/s and a retention time at the transcription site of 0s.

Supplementary Figure 3: Identifying steady state transcription using deconvolution. A) Deconvolution of each transcriptional site intensity into RNA polymerase II initiation events over time[18,19]. The average waiting time (Δi) between polymerase initiation events is calculated for all nuclei within a sliding time window (Δ*t*). The inverse of Δi is the product of the probability to be active, denoted p_{ON}, and the polymerase initiation rate (k_{INI}) for the given time window. The inverse value is plotted over time as a proxy for the stability of the underlying transcriptional kinetic regime.

Supplementary Figure 4: Bayesian Change Point Detection (BCPD) identifies active and repressed periods. A) Schematic depicting the run length, defined as the time from the last change of signal parameters, given the observed values of the signal. B) Synthetic data generated using the Gillespie algorithm with a two-state auto-repressive model. In this model, $k_{ON}$and k_{OFF} depend on the protein level according to decreasing and increasing Hill functions, respectively (top). A computed synthetic trace highlights a detected change point using a colour change (middle). Simulated data indicates a switch in global kinetic parameters (bottom). C) Synthetic traces generated using the Gillespie algorithm with their respective run length identified by the BCPD algorithm and classification into active and repressed states. D) Sample synthetic trace with no change in activation/repression is correctly detected by the BCPD algorithm. E) Cumulative distribution of BCPD algorithm-identified changepoints for synthetic data.

Supplementary Figure 5: Alternative *snail*^{MS2} three-state model topologies have identical fittings. A) Schematic of a non-sequential RNA Polymerase II passage topology. This model does not have a direct transit between the non-productive OFF1 and OFF2 states. B) Schematic of a sequential passage topology model where RNA Polymerase II passes directly between non-productive OFF1 and OFF2 states. C, D) Survival function of the distribution of waiting times between polymerase initiation events (red circles) for the repressed period of *sna*^{MS2} transcription in nc14 with the three-exponential fitting of the population according to the model in A (C) or model in B (D), estimated using the Kaplan–Meyer method (black line). The dashed lines indicate 95% confidence interval estimated based on Greenwood's formula. The Kolmogorov-Smirnov (K-S) test verifies that the empirical waiting times come from the predicted distributions (small p-values mean that the null hypothesis is not valid). The objective function (Obj.) comes from the least squares error of the survival function (large values mean the error is high). A green check indicates an accepted fitting. A green check indicates an accepted fitting. Statistics: *sna*^{MS2} nuclear cycle 14: N=6 embryos, n=448 nuclei.

Supplementary Figure 6: Survival functions for *white*-I, *paf1*-i^{A} and *paf1*-i^{B}. A) Schematic of *Drosophila* crosses used to generate data in Figure 3 and Supplementary Figure 6. B) RNAi knockdown efficiency expressed relative to *white*-i. C-E) Survival function of the distribution of waiting times between

polymerase initiation events (red circles) for *white*-i (C), *paf1-i*^{A} (D) and *paf1-i*^{B} (E) with the two-exponential fitting of the population estimated using the Kaplan–Meyer method (black line). The dashed lines indicate 95% confidence interval estimated based on Greenwood's formula. The Kolmogorov-Smirnov (K-S) test verifies that the empirical waiting times come from the predicted distributions (small p-values mean that the null hypothesis is not valid). The objective function (Obj.) comes from the least squares error of the survival function (large values mean the error is high). A green check indicates an accepted fitting. F-H) Survival function of the distribution of waiting times between polymerase initiation events (red circles) for *white*-i (F), *paf1-i*^{A} (G) and *paf1-i*^{B} (H) with the three-exponential fitting of the population estimated using the Kaplan–Meyer method (black line). The dashed lines indicate 95% confidence interval estimated based on Greenwood's formula. The Kolmogorov-Smirnov (K-S) test verifies that the empirical waiting times come from the predicted distributions (small p-values mean that the null hypothesis is not valid). The objective function (Obj.) comes from the least squares error of the survival function (large values mean the error is high). A green check indicates an accepted fitting. Statistics: *white*-i > *sna*^{MS2}: N=2 embryos, n=183; *paf1*-i^{A} > *sna*^{MS2}: N=2 embryos, n=126; *paf1*-i^{B} > *sna*^{MS2}: N=2 embryos, n=184.

Supplementary Figure 7: Survival functions for *Gal4>+* control and *UAS:CycT*. A) Schematic of *Drosophila* crosses used to generate data in Figure 3 and Supplementary Figure 7. B) Overexpression efficiency expressed relative to *Gal4>+*. C-D) Survival function of the distribution of waiting times between polymerase initiation events (red circles) for *Gal4>+* (C) and *UAS:CycT* (D) with the two-exponential fitting of the population estimated using the Kaplan–Meyer method (black line). The dashed lines indicate 95% confidence interval estimated based on Greenwood's formula. The Kolmogorov-Smirnov (K-S) test verifies that the empirical waiting times come from the predicted distributions (small p-values mean that the null hypothesis is not valid). The objective function (Obj.) comes from the least squares error of the survival function (large values mean the error is high). A red cross indicates a rejected fitting. E-F) Survival function of the distribution of waiting times between polymerase initiation events (red circles) for *Gal4>+* (E) and *UAS:CycT* (F) with the three-exponential fitting of the population estimated using the Kaplan–Meyer method (black line). The dashed lines indicate 95% confidence interval estimated based on Greenwood's formula. The Kolmogorov-Smirnov (K-S) test verifies that the empirical waiting times come from the predicted distributions (small p-values mean that the null hypothesis is not valid). The objective function (Obj.) comes from the least squares error of the survival function (large values mean the error is high). A green check indicates an accepted fitting. Statistics: *Gal4>+> sna*^{MS2}: N=2 embryos, n=191 nuclei; *UAS:CycT> sna*^{MS2}: N=2 embryos, n=154 nuclei.

Supplementary Figure 8: Sna binding profiles at selected loci. Illustrations depict the loci and regulatory sequences (in grey) of *sna* and *sog*, with Snail binding sites (highlighted in green) placed in regions where the ChIP signal of Snail[44] is elevated (>500). When Sna binding sites do not overlap, the number of base pairs separating each other is written above or below.

Supplementary Table 1: guide RNA sequences used to generate *snail*^{MS2} and Snail^{Llama} CRISPR alleles.

Supplementary Table 2: *Drosophila* lines used in this manuscript.

Supplementary Table 3: Single molecule fluorescent *in situ* hybridization probes for endogenous *snail* (related to Figure 1).

Supplementary Table 4: Kinetic parameters for promoters derived from deconvolution and multi-exponential regression fitting of live imaging data. Minimum and maximum values indicate the boundaries of the error interval. State durations are calculated from the provided switching rates (k_n) and time durations for each state are provided as 'T(state)'. State probability values are indicated as 'p(state)'. Bold indicates the most parsimonious appropriate fitting of the data. The table also provides the objective functions and one-sided Kolmogorov-Smirnov test results.

Supplementary Table 5: qPCR primers related to Supplemental Figures 6 and 7.

Supplementary Movie 1: Description: Live imaging of *snail*^{MS2} representative of nc13-14 beginning at mitosis. Nuclei are detected using His2Av-mRFP and MS2 using MCP-GFP.

Supplementary Movie 2: Description: Live imaging of *mat-α*:GAL4/+; *nos*:GAL4/+ > *white*-i > *snail*^{MS2} representative of NC14 beginning at mitosis. Nuclei are detected using His2Av-mRFP and MS2 using MCP-GFP.

Supplementary Movie 3: Description: Live imaging of *mat-α*:GAL4/+; *nos*:GAL4/+ > *paf1*-i^{A} > *snail*^{MS2} representative of NC14 beginning at mitosis. Nuclei are detected using His2Av-mRFP and MS2 using MCP-GFP.

Supplementary Movie 4: Description: Live imaging of *mat-α*:GAL4/+; *nos*:GAL4/+ > *paf1*-i^{B} > *snail*^{MS2} representative of NC14 beginning at mitosis. Nuclei are detected using His2Av-mRFP and MS2 using MCP-GFP.

Supplementary Movie 5: Description: Live imaging of *mat-α*:GAL4/+; *nos*:GAL4/+ > *+* > *snail*^{MS2} representative of NC14 beginning at mitosis. Nuclei are detected using His2Av-mRFP and MS2 using MCP-GFP.

Supplementary Movie 6: Description: Live imaging of *mat-α*:GAL4/CyO; *nos*:GAL4/+ > *UAS:CycT* > *snail*^{MS2} representative of NC14 beginning at mitosis. Nuclei are detected using His2Av-mRFP and MS2 using MCP-GFP.

Supplementary Movie 7: Description: Live imaging of *sna*^{Llama} representative of nc13 and nc14 beginning at mitosis. Nuclei are detected using His2Av-mRFP.

Supplementary Movie 8: Description: Live imaging of *sog*^{MS2} representative of NC14 beginning at mitosis. Nuclei are detected using His2Av-mRFP and MS2 using MCP-GFP.