Coordinated active repression operates via transcription factor cooperativity and multiple inactive promoter states in a developing organism

Virginia L. Pimmett [1], Maria Douaihy [1,2,4], Louise Maillard [1,4], Antonio Trullo [1], Pablo Garcia Idieder [1], Mélissa Costes [1], Jeremy Dufour [1,3], Hélène Lenden-Hasse [1], Ovidiu Radulescu [2] & Mounia Lagha [1]

Refining transcriptional levels via active repression in a euchromatic context represents a critical regulatory process. While the molecular players of active repression are well described, their dynamics remain obscure. Here, we used *snail* expression dynamics as a paradigm to uncover how repression, mediated by the Snail (Sna) repressor, can be imposed within a developing tissue. Combining live imaging and mathematical modeling, we show that Sna-mediated repression is cooperative and that cooperativity is primarily mediated by the distal enhancer. Repression shifts transcription bursting dynamics from a two-state ON/OFF regime to a three-state repressed regime with two temporally distinct OFF states. Mutating Sna binding sites suggests that repression introduces the long-lasting inactive state, which is stabilized by cooperativity. Our approach offers quantitative insights into the dynamics of repression and how transcription factor cooperativity coordinates cell fate decisions within a tissue.

Cell fate specification critically depends on differential gene expression. Cells are specified through the concomitant transcriptional activation of key lineage specifying factors and the repression of alternative fates. The coordination of gene activation is particularly important during the development of multicellular organisms where multipotent cells must choose distinct differentiation routes in an orchestrated manner. Thanks to functional genomics approaches, how the combinatorial action of transcription factor (TF) elicits the activation or repression of developmental promoters is relatively well understood. Downregulation of a gene is typically achieved by repressors, TFs that recruit co-repressors to reduce or silence gene

expression. Depending on their range of action, they 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. In contrast, short-range repressors function locally (50–150 bp) to inhibit the basal transcriptional machinery without interfering with more distant activators[1]. At the molecular level, several mechanisms have been proposed including direct competition between activators and repressors for a shared DNA binding site or 'quenching' of closely located activators and members of the basal transcription machinery[2]. A third and non-exclusive mechanism is the recruitment of histone deacetylases

[1]Institut de Génétique Moléculaire de Montpellier, University of Montpellier, CNRS UMR 5535, Montpellier, France. [2]Laboratory of Pathogens and Host Immunity, University of Montpellier, CNRS, INSERM, Montpellier, France. [3]Institut de Recherche en Infectiologie de Montpellier, CNRS UMR 9004, University of Montpellier, Montpellier, France. [4]These authors contributed equally: Maria Douaihy, Louise Maillard. ✉e-mail: ovidiu.radulescu@umontpellier.fr; mounia.lagha@igmm.cnrs.fr

(HDACs), condensing chromatin and restricting access to the promoter[1].

Two modes of transcriptional repression can be distinguished: the classical silencing in the context of heterochromatin or reduction in the context of a euchromatic environment, referred to as active repression. Contrary to Polycomb-mediated gene silencing, much less is known concerning active repression[3]. Yet, because of its fast establishment, reversibility, and capacity for partial reduction of expression, active repression stands as an optimal mode of gene expression control during periods of rapid decision-making. Failure of repression can lead to developmental defects and diseases such as cancer, as exemplified by the regulation of the Epithelial to Mesenchymal Transition (EMT). This fundamental cellular process is instructed by the conserved pro-EMT *snail* family, composed of Snail/Slug, Twist and ZEB1, acting as both activators and repressors[4]. Snail (Sna) is a zinc finger transcription factor primarily acting as a repressor, but reported to also act as an activator in some contexts[5]. Sna plays a critical role for correct completion of EMT, such as during *Drosophila* gastrulation or vertebrate neural tube formation[6]. Sna overexpression is sufficient to induce EMT[7] and tumorigenesis[8]. Thus, tuning Sna levels and the network induced by this repressor TF is critical.

While repressor identities are well known, their impact on transcriptional kinetics is much less described. Transcription is inherently dynamic and occurs in pulses known as transcription bursts[9]. Transcription bursts are due to the stochastic switching of the promoter between permissive active states (ON, from which RNA polymerase II (Pol II) can initiate) and inactive states (OFF) of multiple timescales[9]. The mean RNA production depends on the switching rates between these states and on the active state production rate[10]. Repressors can in principle modulate any or all state switching rates, resulting in fine-tuning of repression rather than an all-or-nothing process. However, the scheme of state switching and the specific parameter(s) modulated by repression remain unknown.

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, the EMT. Using novel theoretical approaches, we uncovered and quantified the timescales of the kinetic bottlenecks tuning transcription during repression. Based on experimental measurements, we propose a stochastic model of transcriptional activation and repression. In this model, repression adds a new long OFF state to the two-state unrepressed dynamic and modulates the stochastic switching rates cooperatively. Numerical simulations of this model suggest that the cooperativity between repressors contributes to coordination of repression within a tissue.

Results

Monitoring transcriptional repression in vivo
To decode the dynamics of transcriptional repression, we focused on a model gene, *snail* (*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. To examine the endogenous dynamics of *snail* expression in real time, we inserted a 24xMS2 array[11] into the 3′UTR of the endogenous *snail* gene using CRISPR-mediated recombination (*sna^MS2*; Fig. 1a). When both maternal and paternal alleles are tagged, the resulting *sna^MS2* flies are homozygous viable, and MS2 reporter expression matched the endogenous *sna* signals, as shown by single mRNA fluorescence in situ hybridization (smFISH) experiments (Supplementary Fig. 1A, B).

To image transcription, MS2 coat protein is fused to eGFP (MCP-eGFP) and provided maternally along with a fluorescently-tagged histone (His2A-mRFP). In combination with the paternally provided *sna^MS2* allele, transcription is visible as bright nuclear foci (Supplementary

Movie 1). Signal intensity was retrieved in 3D and tracked through nc13 and nc14, using mitosis as a temporal 'time-zero' reference for each nuclear cycle.

We characterized the dynamics of endogenous *sna* transcription at the single nucleus level in embryos heterozygous for *sna^MS2*. Transcription is detectable in living embryos as early as nc11 (Fig. 1b). While transcription was stable throughout nc11-13, the dynamics in nc14 evolved significantly over time. Reactivation after mitosis in nc14 is rapid and synchronous but declines after a short plateau (Fig. 1c, d, Supplementary Fig. 1C). 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 (Fig. 1e, Supplementary Fig. 1D).

We turned to identifying a putative repressor of *sna* in nc14. As Sna is a transcription factor known to act as a repressor in *Drosophila* and to form an autoregulatory loop in other species[12,13], we examined whether Sna formed an auto-repressive loop in the early embryo. To test Snail function while maintaining *snail* transcription as a readout, we created a protein-null allele (*sna^ΔATG*) (Fig. 1f, Supplementary Fig. 2A–D). Using smFISH, we observed that the loss of Sna resulted in a derepression of *sna* transcription in late nc14 (Fig. 1g). Thus, it appears that while *sna* transcriptional activity is stable during early embryogenesis, it undergoes rapid evolution in nc14 driven at least in part by Sna itself.

snail repression results in a non-stationary transcriptional regime
To access the kinetic parameters driving *sna* expression, we employed our previously developed deconvolution pipeline[14–16] to extract the sequence of Pol II initiation events from the MS2-MCP-GFP signal for each single nucleus. Critically, this process does not rely on the arbitrary assignment of bursts to the signal and is valid for signals that are both stationary (i.e. the underlying kinetic parameters are stable across time) and non-stationary (i.e. temporally variable kinetic parameters).

In brief, we consider the intensity trace of each spot to be 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[17] (Supplementary Fig. 3A, B); the resulting values are similar to those obtained from published Pol II speed measurements[18]. Using deconvolution, fluorescence traces can thus be converted to polymerase initiation events for single nuclei. Single nuclei can then be assessed as an ensemble of polymerase initiation events over time (Fig. 2a, b).

We extracted the waiting times between polymerase initiation events for each nucleus and quantified the mean waiting time between polymerase initiation events ($<\tau>$) in a sliding window (Fig. 2c). $<\tau>$ represents the inverse mean RNA production, and is thus directly related to the product of p_{ON}, or the probability to be in a productive state, and k_{ini}, or the initiation rate while in the productive state[19]. The temporal profile of $p_{ON} \cdot k_{ini} = 1/<\tau>$ is a readout of the stability of transcription. Using a sliding window, the temporal evolution of $p_{ON} \cdot k_{ini}$ (or RNA production rate) can be tracked across time. We examined the temporal evolution of the RNA production rate for *sna* in nc14 (Fig. 2d) and compared it to another target of Snail-mediated repression, *short gastrulation* (*sog*, Fig. 2e, Supplementary Movie 2)[20]. Unlike *sna*, which undergoes a partial repression, *sog* is fully silenced by the action of Sna[21] in the mesoderm (Fig. 2f, g). For both *sna* and *sog*, the mean waiting time between polymerase initiation events was non-stationary across nc14 (Fig. 2d, e).

As the population-level signal is non-stationary, we employed the Bayesian change point detection (BCPD) algorithm[22] to identify the timing of change between a transcriptionally unrepressed and repressed program at the single nucleus level (see "Methods").
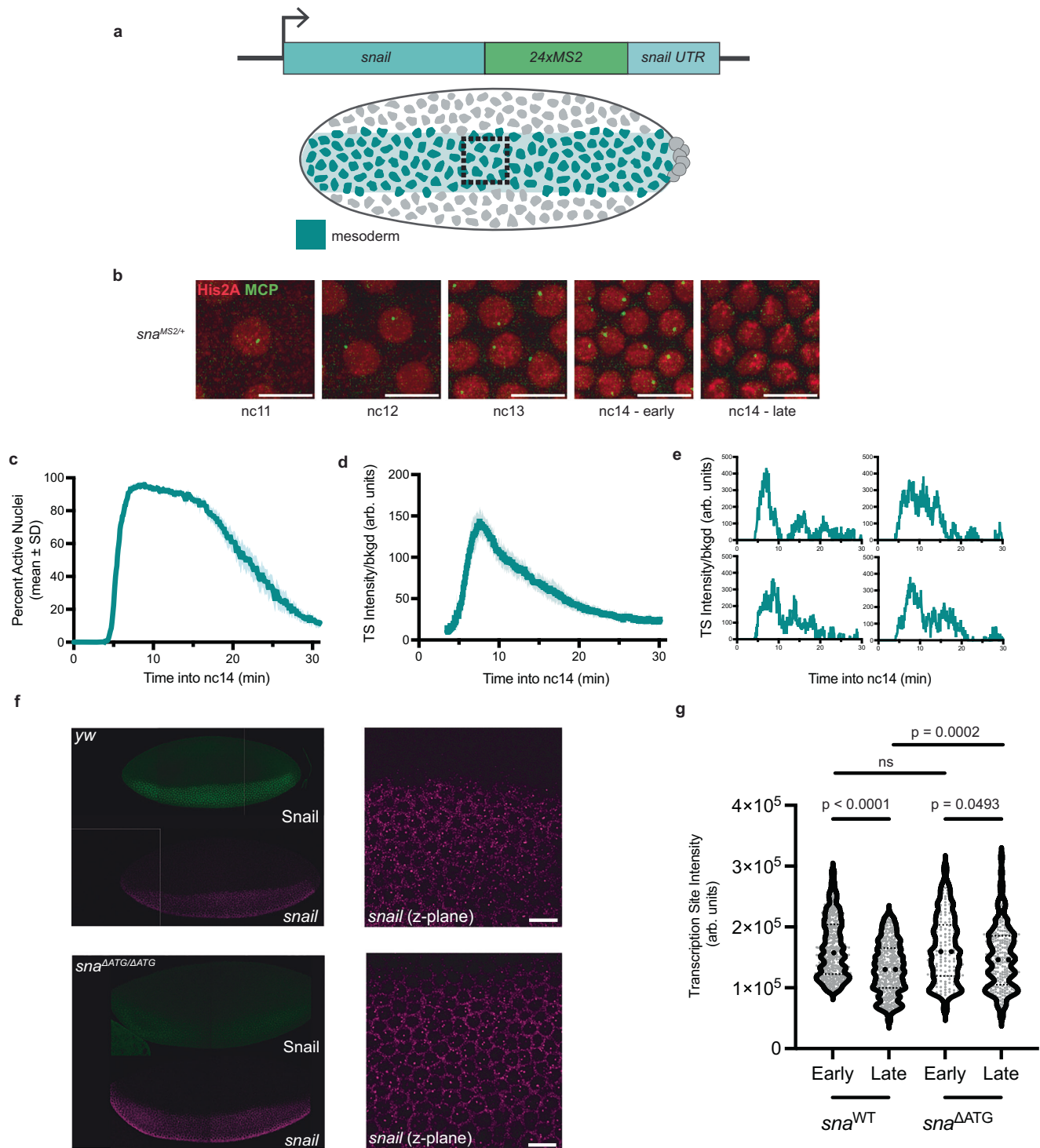
Fig. 1 | Monitoring transcriptional repression in vivo: the case of *sna* auto-regulation. a Schematic view of *sna^MS2^* allele (above) and expression domain in the embryonic mesoderm (below, teal). The box indicates the restricted imaging area. **b** Maximum intensity Z-projection of representative nuclei showing MS2/MCP-eGFP-bound puncta and nuclei (His2A-mRFP) in sequential nuclear cycles. Images were taken from a heterozygous embryo expressing *sna^MS2^* (Supplementary Movie 1). Scale bar represents 10 μm. **c** Instantaneous activation percentage (mean ± SEM) curves of ventral nuclei during the first 30 min of nc14. Time zero is from anaphase during nc13-nc14 mitosis. **d** Fluorescence intensity of actively transcribing nuclei (mean ± SEM) for *sna^MS2/+^* nuclei during the first 30 min of nc14.

Time zero is from anaphase during nc13-nc14 mitosis. **e** Sample single nucleus fluorescence traces in the first 30 min of nc14. Time zero is from anaphase during nc13-nc14 mitosis. **f** *sna^ΔATG/ΔATG^* embryos demonstrate absence of Snail protein (green) and increased nascent transcription activity by smFISH relative to control embryos. Scale bar represents 10 μm. **g** Quantification of endogenous *sna* and *sna^ΔATG^* transcription site intensity divided by background in early and late nuclear cycle 14 embryos via smFISH. Statistics: *sna^MS2/+^*: *N* = 6 embryos, *n* = 484 nuclei. smFISH: *N* = 3 *sna^ΔATG/ΔATG^* embryos for early and late smFISH time points as determined by membrane invagination. Significance is indicated using a Kruskal-Wallis test (two-sided) with Dunn's multiple comparisons.

a *snail*

b *sog*

c Δi = Waiting time between RNA Pol II initiation events

$1/\langle\tau\rangle = p_{ON} \cdot k_{ini}$

$\langle\tau\rangle$ = average Δi within the window Δt

d *snail*

e *sog*

f nc14 + 7 min   nc14 + 15 min   sna^MS2

g nc14 + 7 min   nc14 + 15 min   mesoderm   sog^MS2

h *snail*

Median repression: 13.6 min
Interquartile range: 6.6 min

i *sog*

Median repression: 14.7 min
Interquartile range: 4.8 min
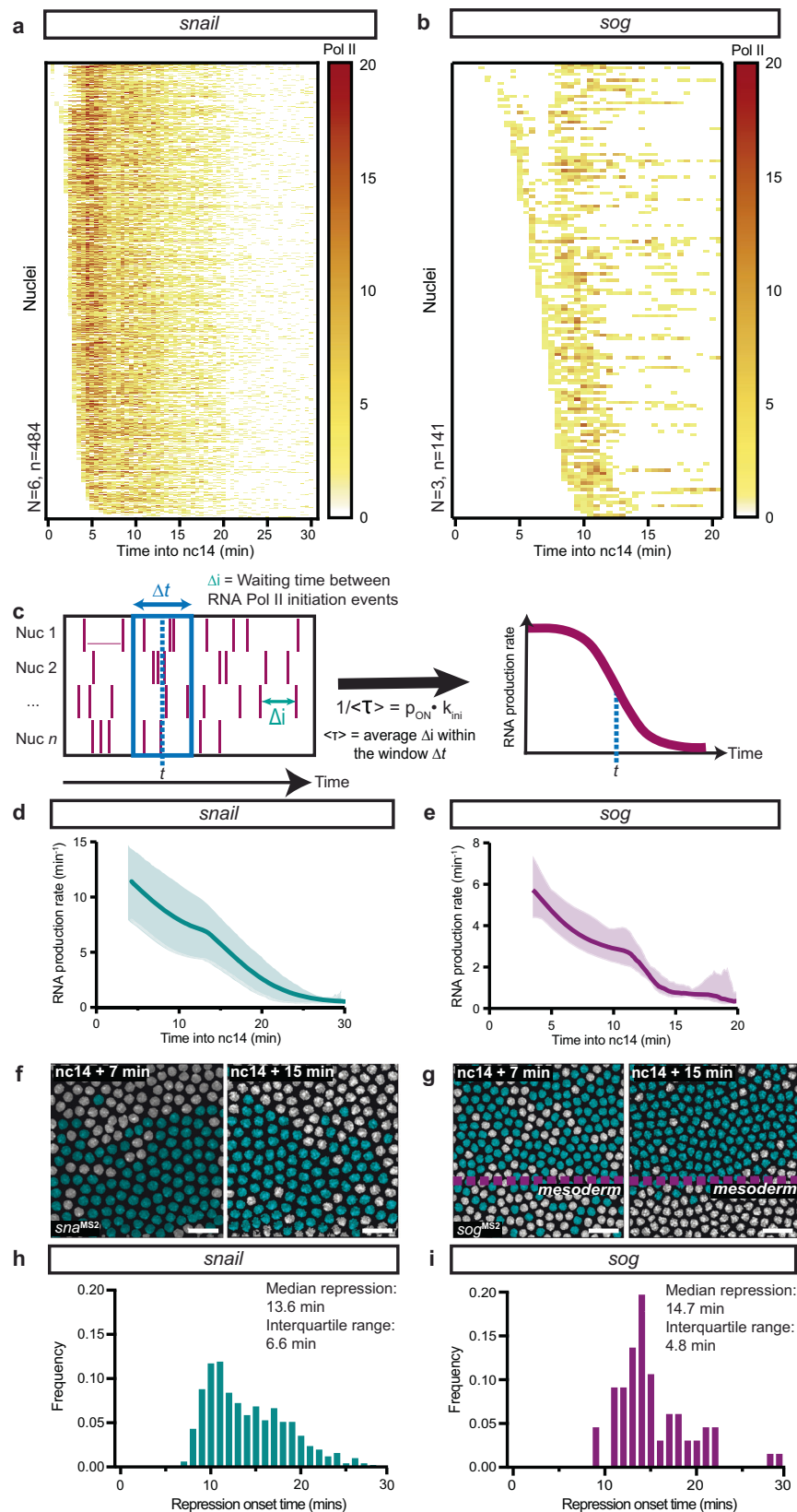
Importantly, this algorithm identified the point at which the repressive regime is stabilized (Supplementary Fig. 4A–D). Interestingly, the internuclear coordination of repression, represented by the breadth of repression onset times, was weaker for incomplete repression (*sna*, Fig. 2h) compared to complete repression (*sog*, Fig. 2i). We also examined whether repression was implemented similarly between both alleles of *sna* by analyzing *sna^MS2/sna^MS2* homozygous embryos. The median time where repression achieved stationarity (i.e. the underlying kinetic parameters stabilized) for the first- and second-activated allele was similar to both each other and a randomly selected pool of alleles (Supplementary Fig. 5A–F), indicating repression is implemented uniformly across alleles. We conclude that, although

Fig. 2 | Deconvolution of transcription in living embryos reveals gene-specific behavior at repression onset. Heatmap showing the number of polymerase initiation events in nc14 for $sna^{MS2/+}$ (**a**) and $sog^{MS2/+}$ (**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** Deconvolution of the transcriptional site intensities into RNA polymerase II initiation events over time. The average waiting time ($<\tau>$) between polymerase initiation events is calculated for all nuclei within a sliding time window (Δt). The inverse of $<\tau>$ 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. Stationarity is denoted by a slope ≈ 0. Kinetic parameter stability as a function of time for $sna^{MS2/+}$ (**d**) and $sog^{MS2/+}$ (**e**) transcription expressed as the product of the probability to be active (p_{ON}) and the RNA polymerase II initiation rate (k_{ini}). Error represents the upper and lower bounds of the 95% confidence interval from all movies. False-colored projections from live imaging of $sna^{MS2/+}$ (**f**) and $sog^{MS2/+}$ (**g**) embryos, with active nuclei indicated in teal and inactive in gray (Supplementary Movies 1, 2). scale bar is 20 μm. Distribution of switching times for initiation of stable repression in nuclear cycle 14 for sna^{MS2} (**h**) and sog^{MS2} (**i**) as determined using Bayesian Change Point Detection. The p value for the difference in distributions is 0.0012 (Kolmogorov-Smirnov test). Statistics: $sna^{MS2/+}$: $N = 6$ embryos, $n = 448$ nuclei. $sog^{MS2/+}$: $N = 3$ embryos, $n = 141$ nuclei.

repression is applied progressively, stationary repression is reached after a certain onset time, and variations in this time can be used to quantify the coordination of repression.

Snail-mediated active repression involves Sna cooperative action

To examine the interplay between endogenous Sna protein and its target genes *sna* and *sog*, we employed the LlamaTag system[23] and created a CRISPR *snail*^*Llama*^ (Fig. 3a, b). The LlamaTag system relies on a nanobody, targeting a fluorescent protein, fused to a protein of interest and the presence of a free fluorescent detector. Once Sna protein is translated in the cytoplasm, the Llama nanobody binds to free maternally-deposited GFP and is then imported in the nucleus. The subsequent increase in nuclear fluorescence signal provides a readout of Sna nuclear levels.

The *sna*^*Llama*^ CRISPR allele allowed us to quantify with high spatio-temporal resolution the levels of endogenous Sna protein (Fig. 3c, e, f) by tracking the nuclear GFP signal. Despite the early accumulation of *sna* transcripts (Fig. 1b), nuclear accumulation of Sna protein was detected weakly in nc13 (Fig. 3c, Supplementary Movie 3). By comparing signal within the mesoderm to that in a neighboring tissue where *sna* is not expressed (neurogenic ectoderm, Fig. 3d), we observed Sna nuclear protein levels in nc14 continuously increase in the presumptive mesoderm (Fig. 3e).

Sna has previously been demonstrated to bind known enhancers of both *sna* and *sog*[24] (Fig. 3f, g). Consistent with the idea that Sna acts as a transcriptional repressor, we observed an anticorrelation between Sna protein levels and RNA production rates of both *sna* and *sog* (Fig. 3h, j). We next wanted to quantitatively characterize their relationship. The input/output relationship is steep for both *sna* and *sog*, indicating thresholding effects. Such relationships are typically fitted with a Hill equation[25] that accesses a key parameter, the Hill coefficient *n* (representing the degree of cooperativity). The Hill coefficient for *sna* during nc14 was 6.07 (Fig. 3i). Similarly, the relationship between input Sna protein and output *sog* transcription also indicates high-degree cooperativity (Hill coefficient 7.6, Fig. 3k).

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

Sna cooperativity is differentially mediated through distinct cis-regulatory regions

We exploited previously characterized *sna* BAC reporter lines[26] to explore where Sna cooperative action operates. *Sna* is regulated by a pair of non-redundant enhancers, one proximal and one distal[27–29] (Fig. 4a). We performed quantitative imaging of the *sna* wild-type and enhancer deletion BACs in the mesoderm (Fig. 4b) and characterized the relationship between RNA production rate and the Snail endogenous protein using a Hill function. Compared to the control, loss of the distal (i.e. shadow[28], Fig. 4c, f) enhancer ($sna^{\Delta DIST}$, Fig. 4d) resulted in a complete loss of Sna cooperativity (Hill coefficient ≈1, Fig. 4g) while the loss of the proximal enhancer ($sna^{\Delta PROX}$, Fig. 4e) demonstrated increased Sna cooperativity (Hill coefficient ≈4) (Fig. 4h). Thus, it appears that Sna cooperative action is primarily decoded by the distal enhancer.

Sna repressor acts by introducing a second non-productive state and modulating the promoter's state switching rates

It remains an open question which aspect of the transcriptional kinetic cycle repression acts on. As the Hill equation is phenomenological in nature, we turned to transgenic reporter assays to establish a causal link between Snail binding and transcriptional repression, and to investigate the link between promoter dynamics and the imposition of repression.

The *sna* distal enhancer has a cluster of nine Sna binding sites (Fig. 5a, pink bars). We employed our previously established transgenic reporter platform[16] to investigate the effect of these Sna TF binding sites on transcription dynamics. We generated a series of transgenes where transcription is controlled by 4 types of enhancers (Fig. 5b–e): the wild type *sna* distal enhancer (Fig. 5b, *snail*^*Distal*^), a mutated distal enhancer with every other Sna binding site removed (i.e. 5/9) (Fig. 5c, *snail*^*DistalAlt*^), with 8/9 Sna binding sites removed (Fig. 5d, *snail*^*DistalMut*^), or a fragment of the distal enhancer with no Sna sites (Fig. 5e, *snail*^*DistalCore*^). Following deconvolution, we observed the *snail*^*Distal*^ transgene had globally lower numbers of polymerase initiation events compared to *snail*^*DistalAlt*^, *snail*^*DistalMut*^ and *snail*^*DistalCore*^ (Fig. 5f–i). Thus, the *sna* distal enhancer is derepressed in the absence of Sna binding sites.

To further investigate the underlying kinetics of transcription driven by each enhancer, we examined the distribution of waiting times between polymerase initiation events. Briefly, this distribution can be fitted with a multi-exponential function that provides insight into the number of promoter states as well as their duration (T_{State}) and probabilities (p_{State}), and the polymerase initiation rate (k_{ini})[14] (Supplementary Data 1).

The distribution of Pol II waiting times from the *snail*^*Distal*^ construct could not be fitted with a bi-exponential function, meaning a two-state 'random telegraph' model was insufficient to describe the promoter states (Supplementary Fig. 6A, B). A three-exponential fitting was sufficient, corresponding to a three-state model (Fig. 5j) with one productive (ON) state and two non-productive states on the order of minutes (OFF1) or seconds (OFF2) of approximately equal probability (Fig. 5k). The *snail*^*DistalAlt*^ transgene also required a three-state model fitting, but interestingly showed a substantial decrease in the length and probability of the OFF1 state, with a higher probability of the ON state to compensate (Fig. 5l, Supplementary Fig. 6C, D). Transcription from the *snail*^*DistaMut*^ enhancer, with 8/9 Sna binding sites mutated, was adequately fit by a two-state model (Fig. 5m, Supplementary Fig. 6E). Importantly, *snail*^*DistaMut*^ transcription dynamics exhibit a loss of the long OFF1 state, at the expense of an increase in both the duration and probability of the productive ON state as well as a small increase in the duration of the single non-productive OFF state. This bursting regime was recapitulated in the *snail*^*DistaCore*^ reporter, which was also well described by a two-state model with a longer and highly probable ON state and a single short non-productive state (Fig. 5n, Supplementary
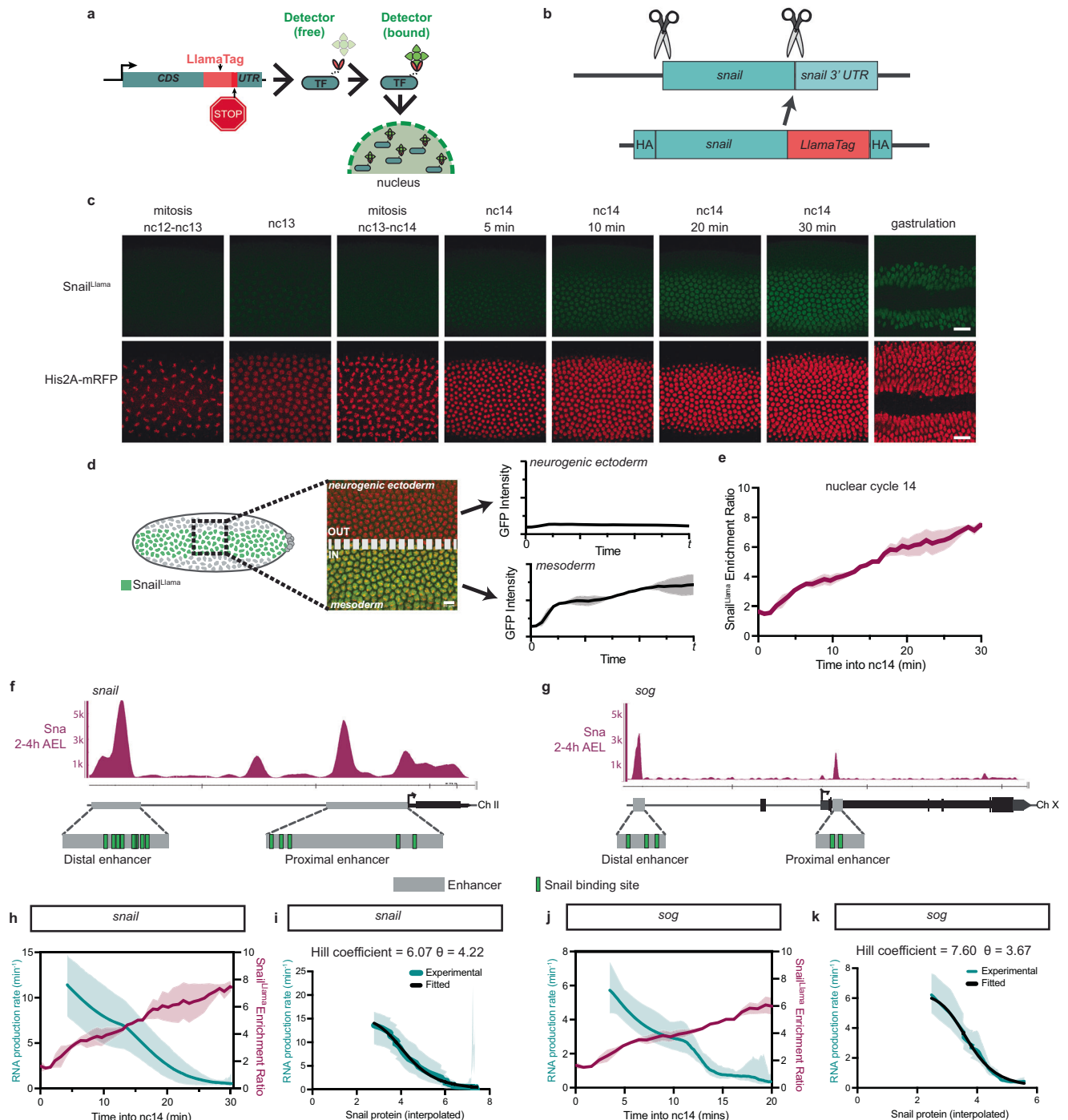
Fig. 3 | Capturing Snail protein dynamics and its correlation with transcription repression in living embryos. a Schematic demonstrating the principle of the LlamaTag system. **b** Schematic indicating CRISPR-mediated genome editing of the endogenous *sna* locus to introduce the LlamaTag nanobody. **c** Representative maximum intensity Z-projections of Sna^{Llama} nuclear signal (above) and nuclei (His2A-mRFP, below) from nuclear cycle 12–13 mitosis until gastrulation. See also Supplementary Movie 3. Scale bars represent 20 μm. **d** Schematic of Sna^{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. Scale bar represents 10 μm. **e** Nuclear Sna^{Llama} enrichment ratio in the first 30 min of nc14 calculated as the mean nuclear GFP signal in the mesoderm relative to the neurogenic ectoderm, expressed as mean ± SD. **f, g** Loci and embryonic enhancer sequences (gray) of *sna*

and *sog*, with Sna binding sites (green) indicated[24]. $p_{ON} \cdot k_{ini}$ (mean RNA production) expressed as a function of time for *sna*^{MS2} (**h**, teal) and *sog*^{MS2} (**j**, purple) and Snail protein enrichment ratio (red) in nc14. The mean RNA production expressed as mean ± SD. Hill fitting for *sna*^{MS2} (**i**) and *sog*^{MS2} (**k**) 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. Colored 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 nc14 (mean ± upper/lower bounds). Black line indicates Hill function fit. Statistics: *sna*$^{MS2/+}$: $N = 6$ embryos, $n = 448$ nuclei; Sna$^{Llama/+}$: $N = 3$ embryos; *sog*$^{MS2/+}$: $N = 3$ embryos, $n = 141$ nuclei.
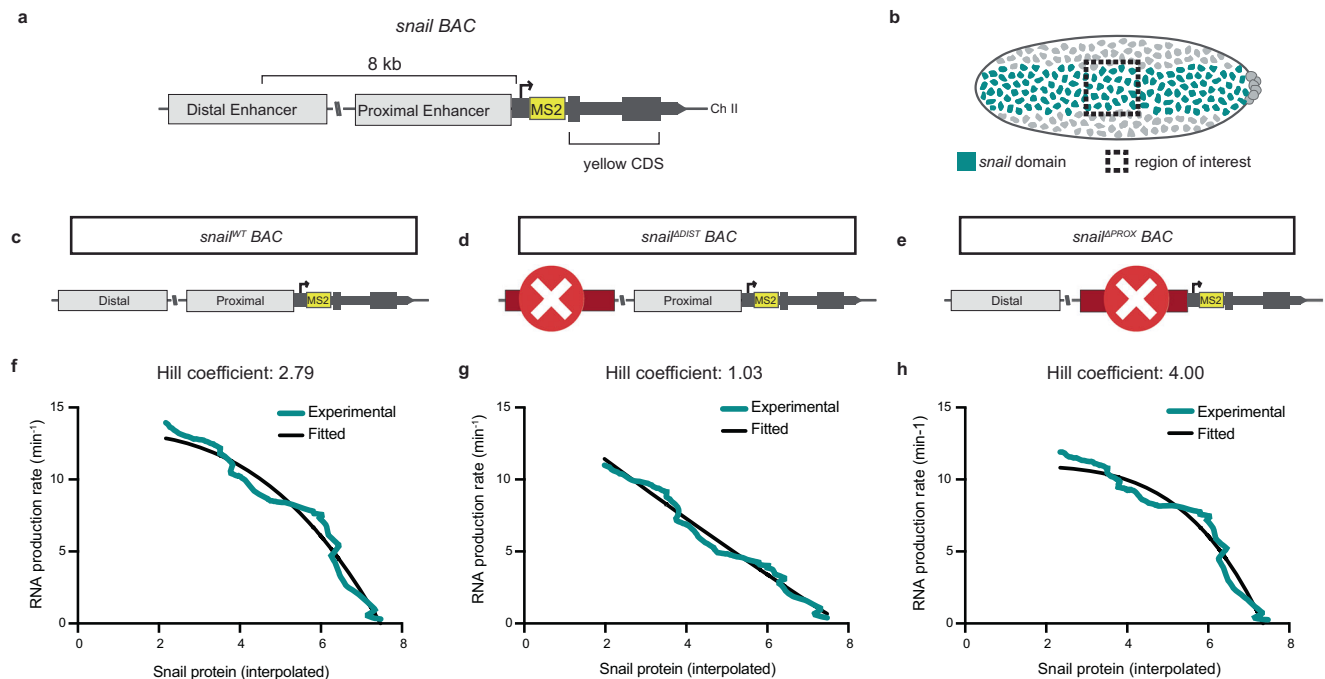
Fig. 4 | Sna cooperativity is differentially mediated by its proximal and distal enhancers. Schematic of *snail* BAC reporter construct (**a**) and expression domain (**b**, teal) with region of interest indicated (dashed box). Schematic of reporter constructs for *sna^WT^* (**c**), *sna^ΔDIST^* (**d**) and *sna^ΔPROX^* (**e**) from Bothma et al. (2015)[26]. Hill fitting for *sna^WT^* (**f**), *sna^ΔDIST^* (**g**) and *sna^ΔPROX^* (**h**) with the removed enhancer and Hill coefficient indicated. Colored 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 nc14. Black line indicates Hill function fit. Statistics: *snail^WT/+^* BAC $N = 4$ embryos, $n = 341$ nuclei; *snail^ΔDIST/+^* BAC $N = 5$ embryos, $n = 344$ nuclei; *snail^ΔPROX/+^* BAC $N = 3$ embryos, $n = 217$ embryos.
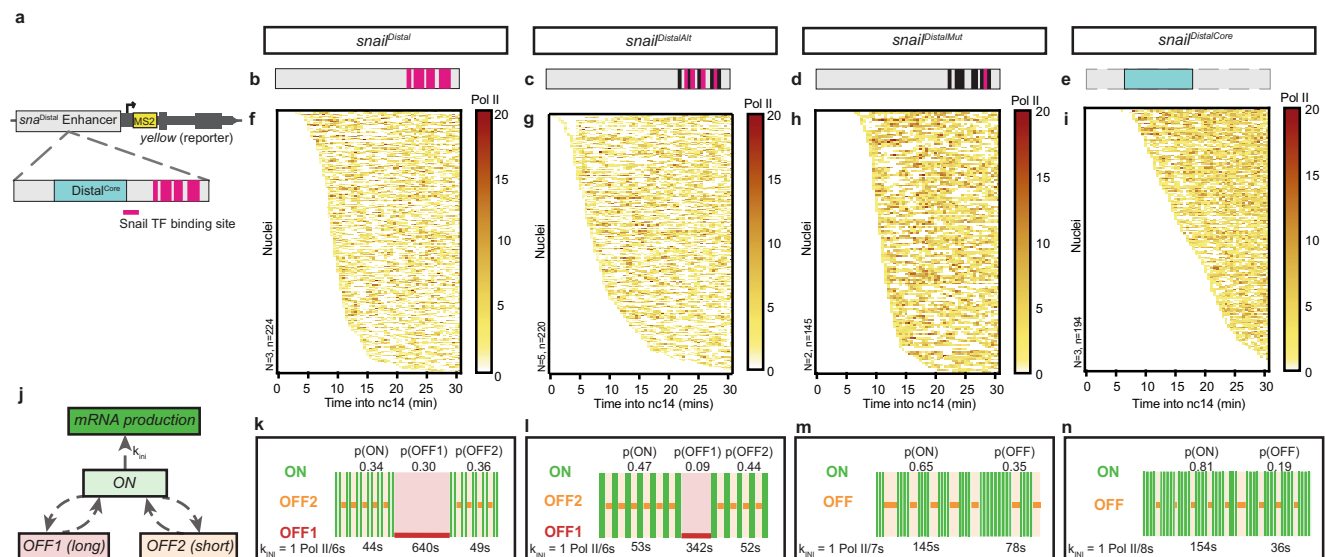

Fig. 5 | Snail repressor introduces a second non-productive state and modulates promoter state switching rates. a Schematic of the *sna* distal enhancer transgene with the 'core' enhancer module[49] (blue) and Snail binding sites (pink) indicated. Enhancer transgene schematics for the *snail^Distal^* enhancer (**b**), *snail^DistalAlt^* with mutated Sna binding sites (**c**), *snail^DistalMut^* with mutated Sna binding sites (**d**) and *snail^DistalCore^* with the 'core' enhancer module only (**e**). Heatmap showing the number of polymerase initiation events in nc14 for *snail^Distal^* (**f**), *snail^DistalAlt^* (**g**), *snail^DistalMut^* (**h**) and *snail^DistalCore^* (**i**) 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. **J** Topology of the three-state kinetic model with non-sequential OFF1 and OFF2 states. Representation of estimated bursting dynamics for *snail^Distal^* (**k**), *snail^DistalAlt^* (**l**), *snail^DistalMut^* (**m**) and *snail^DistalCore^* (**n**). Permissive ON state durations are depicted in green and inactive OFF states in red and orange, and probabilities of each state shown above. Statistics: *snail^Distal^* $N = 3$ embryos, $n = 224$ nuclei; *snail^DistalAlt^* $N = 5$ embryos, $N = 220$ embryos; *snail^DistalMut^* $N = 2$ embryos, $n = 145$ nuclei; *snail^DistalCore^* $N = 3$ embryos, $N = 194$ nuclei. See Supplementary Movies 4–7.

Fig. 6F). Interestingly, the initiation rate was consistent between all three reporter constructs, indicating that Sna binding does not affect Pol II firing from the ON state. Thus, the Sna repressor likely acts by introducing a new kinetic bottleneck, leading to a second non-productive state at the expense of the productive ON state and the pre-existing non-productive OFF state, and without affecting the polymerase initiation rate.

Multiple kinetic state topologies have been proposed for higher-order kinetic models[30] with multiple productive and non-productive states. Based on our previous work, and in agreement with the
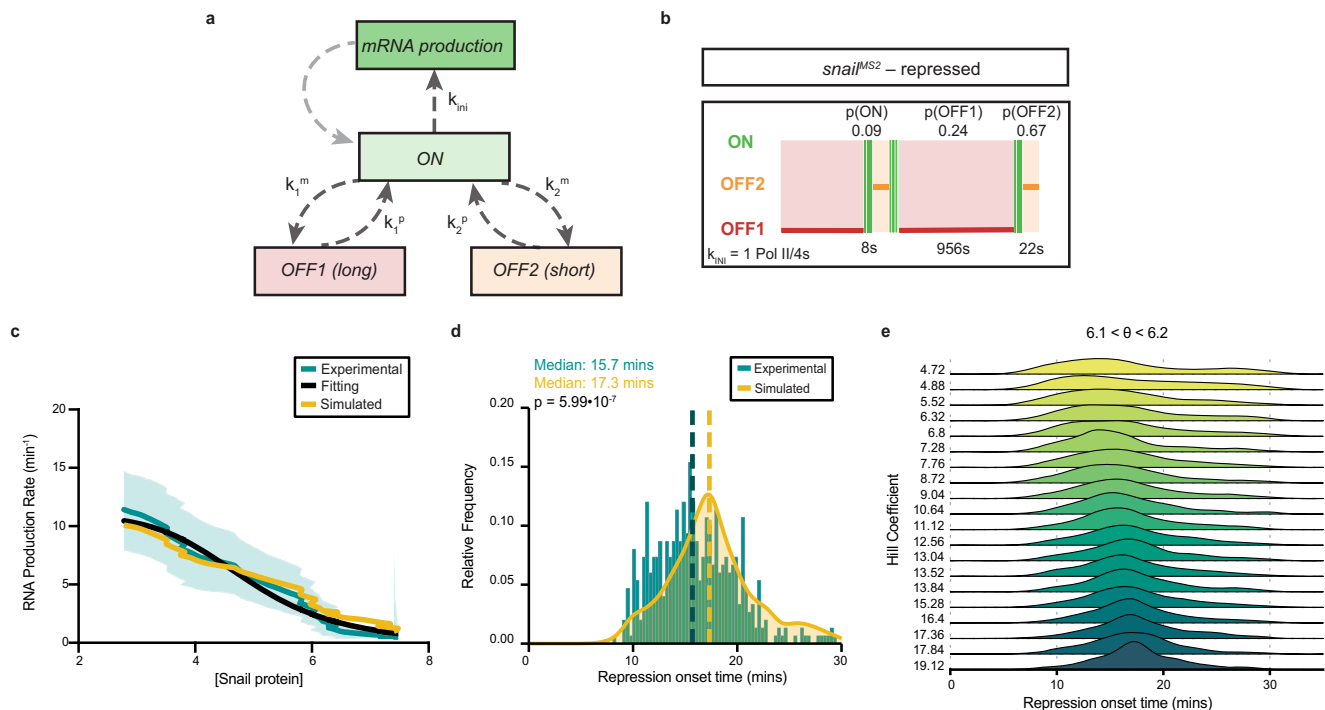
Fig. 6 | A minimal stochastic model reveals the impact of cooperativity on the coordination of transcriptional repression. a Scheme of the Markovian model. **b** Representation of estimated bursting dynamics for sna^{MS2} during stable repression in nc14 (Supplementary Movie 1). Permissive ON state durations are depicted in green and inactive OFF states in red and orange, and probabilities of each state shown above. **c** Best stochastic model fit (yellow) of the RNA production rate parallels the Hill function model fit (black) and the experimentally derived values (teal) over 30 min of minimal stochastic model fitted transcription. Error represents the upper and lower bounds of the 95% confidence interval from all movies. **d** 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 is indicated by dashed lines. As the model was not fitted to the experimental distribution, the good agreement suggests its validation. Significance determined by Brown-Mood median test (two-sided). **e** Distribution of simulated repression onset time for various Hill coefficients for θ constrained to 6.1 < θ < 6.2.

deconvolution findings, we favor a non-sequential three state model of promoter dynamics (Fig. 5j) where there is no direct transition between long OFF1 and short OFF2 state. In this model, the transitions from ON to OFF1 and from ON to OFF2 are independent. A sequential model[15,16,30] would require the promoter to systematically switch to OFF2 before transitioning to OFF1. We have shown that both the sequential and non-sequential models are compatible with the MS2 data[15,19]. Furthermore, they are also mechanistically indistinguishable when the two OFF states have distinct lifetimes (one much shorter than the other). In this case, the estimated lifetimes of OFF1 and OFF2 are very similar across the two models and under a wide range of conditions and genotypes (see Supplementary Text 1).

We also conclude that one of the two OFF states of the three-state model is suppressed when transcription is driven by the $snail^{DistaMut}$ and $snail^{DistaCore}$ enhancers. In the $snail^{DistalAlt}$ construction the long OFF1 state occurs rarely (p_{OFF1} = 0.09 compared to p_{OFF1} = 0.3 in $snail^{Distal}$), suggesting that repression introduces the long OFF1 state. This conclusion is further supported by the distinct orders of magnitude of the OFF state durations: T_{OFF1} is on the order of minutes, and T_{OFF2} on the order of seconds in the three-state constructions, whereas the single T_{OFF} in the two-state constructions is clearly closer to the T_{OFF2} values.

In summary, we find that the Sna repressor acts by introducing a second long non-productive state and modulating all state switching rates, but it does not affect the polymerase initiation rate in the permissive state.

The endogenous repressed state recapitulates transgenic reporter activity

Transgenic reporter experiments suggest that transcription bursting dynamics follow a three-state regime under repression, featuring an additional long OFF state that is not present in the absence of repression. Based on our findings with transgenic reporters, we hypothesized that autorepression of the *sna* locus would also require a three-state topology (Fig. 5j). We used the BCPD procedure previously outlined to isolate the stably repressed phase of endogenous sna^{MS2} transcription (Supplementary Fig. 7A, B) and fit it using various multi-exponential models. We found that a three-state model was required to capture the kinetics of endogenous sna^{MS2} during stable repression (Fig. 6a, b, Supplementary Fig. 7C, D). This contrasted the unrepressed nc13, where a bi-exponential fitting could recapitulate the data (Supplementary Fig. 8A–F).

Previous research has established a role for promoter-proximal polymerase pausing in the introduction of a novel promoter state[16]. We thus tested whether pausing was responsible for the introduction of a new rate-limiting step in transcription, by perturbing the levels of key pausing factors, Paf1[31], a subunit of the NELF complex[32] or Cyclin T[33], required for pause release. We employed two independent RNAi-mediated knockdowns of *paf1*, RNAi-mediated knockdown of *Nelf-A*, as well as overexpression of *Cyclin T*[33] in the early embryo. None of these perturbations altered the number of rate-limiting steps during stable *sna* repression (Supplementary Fig. 9A–O).

Collectively these results suggest that under repression, the endogenous promoter switches between three temporally distinct states: a competent ON state from which Pol II initiates at a given rate, and two inactive states, one short at seconds-scale and a longer minutes-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 attributed to a paused state.

A minimal stochastic model links cooperativity and coordination of repression

To gain quantitative insights into the relationship between Snail-mediated repression and transcriptional bursting, we developed a minimal stochastic model of transcriptional bursting. The model was designed to describe the active (nc13), the repression buildup (early nc14), and the stable repressed regimes (mid-nc14) of *sna* transcription. We therefore used a three-state promoter model comprising two transcriptionally inactive states, a long one (OFF1) and a short one (OFF2), and a permissive (ON) state (Fig. 6a, Supplementary Fig. 10A, B). We examined both the sequential and non-sequential three-state transcriptional models, and did not note a substantial difference in our findings (Supplementary Fig. 11, Supplementary Text 1).

To develop a minimal stochastic model, we first ensured that the state OFF1 is not accessible from the permissive ON during the active phase, corresponding to low concentrations of the putative repressor Sna. To achieve this, we considered that the parameter k_1^m (transition from ON to OFF1) decreases to zero when the concentration of Sna protein ([Sna]) approaches zero. Thus, when [Sna] is low, OFF1 can no longer be reached from ON, rendering the three-state model equivalent to a two-state model.

The parameters k_2^m and k_2^p (transition rate from ON to OFF2 and from OFF2 to ON) have smaller values in nc13 relative to nc14 (Supplementary Data 1) and were considered to increase with [Sna]. The parameter k_1^p (OFF1 to ON transition) is not available for nc13 because OFF1 is absent in the two-state model. However, it is available for nc14 and for the transcription regimes of *sna* enhancer reporter transgenes (Fig. 5). When these reporters are arranged in order of decreasing repression (Supplementary Fig. 10C–G), the value of k_1^p increases, which leads us to propose a decrease in k_1^p with increasing [Sna]. The limiting ([Sna] → 0) value of k_1^p for small concentration of [Sna] is a free parameter that is fitted from data. All the dependencies of state switching rates on [Sna] were modeled as Hill functions to account for cooperativity. Although there may be relationships between the values of these rates and the Hill coefficient (for instance, a large Hill coefficient may correlate with longer-lived states), to minimize the number of parameters and avoid overfitting, we considered a single coefficient for all the states and transitions. Finally, we considered that the polymerase initiation rate (k_{ini}) does not depend on [Sna].

After implementing the active nc13 and repressed nc14 phase parameters (see "Methods"), our model remained with only three free parameters: the limiting ([Sna] → 0) k_1^p value, the Hill coefficient n and the Sna repression threshold concentration θ. Gillespie simulations were used to generate synthetic nascent RNA data for this model and the three free parameters were fitted using the experimental RNA production rate data in nc14 (Fig. 6c).

Next, we simulated the distribution of the repression onset time within the mesodermal population utilizing the fitted parameters (Fig. 6d). The good agreement between the predicted and experimental repression onset time distributions validates the model.

The minimal stochastic model was then used to investigate the coordination of repression within a tissue. We simulated the model for many values of the Hill coefficient (corresponding to a spectrum of TF cooperativity) and Sna concentration threshold and computed the repression onset time distributions (Fig. 6e). The width of this distribution, quantitatively defined as the interquartile range, serves as a measure of repression coordination. In general, for a fixed concentration of repressor, the distribution of the repression onset time narrows as the Hill coefficient increases (Fig. 6e). Because a larger Hill coefficient indicates a higher degree of cooperativity, this suggests that Sna cooperativity contributes to the coordination of repression between nuclei.

The promoter model chosen to illustrate these phenomena is a non-sequential one. As discussed above (see also the **Supplementary Text** 1), given the separation between the two OFF states (with one

being much shorter), choosing a sequential model would not alter our conclusions.

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 *sna* 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 2 main features of active repression: (1) active repression introduces a new long OFF state to the two-state unrepressed dynamics; (2) repression modulates the stochastic switching rates cooperatively. Furthermore, we predict that the inter-nuclear coordination of repression is augmented by a high degree of repressor cooperativity.

Transcription kinetics during active repression

The analysis of the distribution of waiting times between polymerases 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. The molecular nature of these rate-limiting steps can only be an interpretation. Because of the comparison between varying number of Sna binding sites (Fig. 5), we propose that the long OFF1-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, apart from a few exceptions[34,35], activating TFs remain bound to DNA for up to tens of seconds[36] in both *Drosophila* and vertebrates. How can we reconcile typical TF residence time with the prolonged OFF1 state resulting from Sna binding? Given the dependency between OFF1 duration and the number of Sna binding sites and their arrangement, we propose Sna binds in a cooperative manner. Cooperative binding of Sna repressors to DNA may stabilize a longer OFF1 state. A similar scenario of a prolonged promoter state driven by TF cooperativity has recently been proposed in yeast[37]. Albeit reported for activation and not repression, the rationale is nonetheless similar: TF exchange, possibly via cooperative binding, can increase the duration of a rate-limiting step during transcription.

Our work shows that a simple two state bursting model is insufficient to accurately capture Sna repression dynamics. This contrasts with a recent study showing that Knirps-mediated repression operates via two states in *Drosophila* embryos[38]. However, an increasing number of examples have revealed that three state models can more accurately capture the multiple levels of transcriptional regulation, including in the context of the frequently observed long-lived repression periods in cells[39], envisaged as a 'deep off state' in the case of Polycomb-mediated silencing[40].

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 optimize inter-nuclear coordination of repression within a tissue.

It is well-demonstrated that TFs bind DNA cooperatively[41,42]. 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. Here, by directly measuring the relationship between nuclear protein concentration and transcriptional output in live embryos, we provide a quantitative estimation of cooperativity. By examining two transcriptional targets, *sna* 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. This is similar to that measured for the repressor Knirps in *Drosophila*[38].

Cooperativity can occur through protein-protein and protein-DNA interactions, which imposes a particular arrangement of TF binding sites, but can also be achieved with more flexible arrangements such as a local change in DNA structure[43] or mediated through competition with nucleosomes[44]. In principle, all these modes of cooperativity (DNA mediated, TF-TF interaction, or nucleosome-mediated) can lead to a high Hill coefficient. Although both *sna* enhancers contain Sna binding sites, they are clustered in the distal enhancer while their arrangement is more flexible in the proximal (Fig. 3). Interestingly, mutation analyses suggest that these so-called 'redundant' enhancers decode Sna repressor action very distinctly, supporting increased cooperative action at the distal enhancer compared to the proximal.

Future investigations, including promising single molecule technologies as single molecule footprinting assays[45] coupled to theoretical models[46], would be required to elucidate which TF co-occupy the same enhancer DNA molecule 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'[47,48], promise encouraging future insights.

In summary, by monitoring transcription and nuclear transcription factor levels in a developing embryo, we have uncovered kinetic bottlenecks governing repression. Our findings support a multiscale bursting model characterized by both short and long transcriptionally inactive periods. In this model, the initiation of repression results in prolonged non-productive periods governed by slow timescales. Looking ahead, we expect that the framework of analysis and results of this study will set a foundation for understanding repression dynamics in these more complex vertebrate models of development.

Methods

Fly Husbandry
All crosses were maintained at 25 °C. Transgenic and CRISPR lines were maintained as homozygous stocks unless otherwise noted (Supplementary Data 2). 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 *snail*^{Distal}-related transgenes, females heterozygous for *nos > MCP-eGFP-His2Av-mRFP* were crossed to males homozygous for the transgene of interest. For live imaging of RNAi and overexpression experiments, homozygous males carrying *mat-α>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, resulting in embryos heterozygous for *sna*^{MS2}. For imaging of *Sna*^{Llama}, *yw; P{w[+mC] = EGFP-STOP-bcd}* (hereafter named *bcd > GFP*[23]) was crossed to *His2A-RFPt/CyO*, followed by crossing of F1 virgin females to males bearing the *sna*^{llama} allele.

Generation of CRISPR knock-ins and transgenic fly lines
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 Data 3. 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. The dsRed cassette was retained in *sna*^{MS2} stocks. To create the *sna*^{Llama} allele, a CRISPR recombination matrix comprised of an 850 bp genomic homology arm followed by the *sna* coding sequence, a flexible linker and *Drosophila*-optimized GFP-targeting nanobody, the genomic *Drosophila* 3′ UTR, a

floxed 3xP3-dsRed selection cassette, and a 900 bp 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.

Generation of the *sna*^{ΔATG}/*CyO-Hb>lacZ* line was accomplished using a ssODN co-CRISPR approach detailed in Levi et al.[49] Briefly, gRNAs (Supplementary Data 3) were constructed targeting the *sna* and *w* coding sequences via PCR with Phusion polymerase (Invitrogen) and assembled into pCFD4 *sna*^{ΔATG}_*w*^{coffee} using NEBuilderHiFi DNA Assembly Kit (New England Biolabs). A ssODN repair matrix targeting the *sna* N-terminal coding sequence was designed to add an EcoRI site in parallel for the screening. It was co-injected into *y*¹,*M {vas-Cas9} ZH-2A* embryos along with a CRISPR repair matrix (pUC57-white [coffee] Addgene #84006) facilitating a conversion of the *w*⁺ allele into *w*^{coffee} and pCFD4 *sna*^{ΔATG}_*w*^{coffee}. F0 flies were single-crossed to females bearing *sp/CyO-Hb>lacZ* and resulting F1 screened for the *w*^{coffee} phenotype. F1 males were backcrossed to *sp/CyO-Hb>lacZ* balancer females and checked by genomic PCR and digestion (EcoRI) after several days to confirm the mutation. ssODNs were obtained from IDT Technologies.

The *sna*^{Distal}–*24xMS2-y* and *sna*^{DistalCore}–*24xMS2-y* minigenes have been previously described[50,51]. The *sna*^{DistalMut} and *sna*^{DistalAlt} enhancers were synthesized by Twist Biologicals. The *sna*^{DistalCore} sequence was removed from pBPhi *snaShadowCore > snaPr > 24xMS2-y* using restriction enzyme-mediated excision and the *sna*^{DistalMut} or *sna*^{DistalMut} enhancer was inserted using NEBuilder HiFi DNA Assembly Kit (New England Biolabs) upstream of the *sna* promoter sequence. Enhancer sequences are listed in Supplementary Data 4. Transgenic flies were generated by PhiC31-mediated recombination into the VK33 locus (BL 9750).

Injections were performed by the *Drosophila* Transgenesis Facility (Centro de Biología Molecular Severo Ochoa, Madrid) and FlyORF (Zurich). All stocks are homozygous with no observable viability defects, except for *sna*^{ΔATG}/*CyO-Hb>lacZ* which is heterozygous. All lines are listed in Supplementary Data 2.

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 an Airyscan module (Zeiss). Z-stacks comprised of 30 planes with a spacing of 0.5 µm were acquired at a time resolution of 4.64 s (*sna*^{MS2} and *sna* BAC reporters), 6.35 s (*sog*^{MS2} and derivatives), or 3.86 s (transgenic reporters) 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 and *sna* BAC reporter movies were performed with the following settings: GFP excitation by a 488-nm laser (8uW 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 640 × 640 pixels with bidirectional scanning. All *sog*^{MS2} (and derivative genotypes) 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 6.35 s and resolution was 800×800 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 640 × 640 pixels with bidirectional scanning. All *snail^Distal* transgene (and derivative genotypes) movies were performed with the following settings: GFP excitation by a 488-nm laser (10.5uW 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 3x zoom on the ventral region of the embryo centered (±25 μm) on the presumptive ventral midline. For all imaging conditions, Airyscan processing was performed using 3D Zen Black v3.2 (Zeiss).

Imaging of *Sna^Llama;His2A-RFP* was performed with an LSM880 (Zeiss). 18 Z-planes with a spacing of 1 μm were acquired with a time resolution of 42 s/z stack. Movies were performed with the following settings: GFP excitation by a 488 nm laser and RFP excitation by a 561 nm laser captured on a GaAsP-PMT array using a 40x Plan Apo oil lens (NA = 1.3) at 1x zoom with resolution of 512 × 512 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 region of analysis was maintained at 25 μm on either side of the presumptive ventral furrow. The intensity profile of the transcriptional sites imaged were extracted with a custom software developed in Python™[52–54] that has been previously published[16] (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 as two sister chromatids by the detection algorithm, potentially confounding distinguishing between sister chromatids and false detection events. A parameter was defined as the ratio between the convex hull surface determined by the two spots and their actual size. For sister chromatids this ratio is small (<4) since the two spots are close, while a false detection event 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 a user-defined threshold was established for determining the number of sequential inactive time points required for detection to be considered false.

For analysis of homozygous *sna^MS2/MS2* movies, a further post-processing tool was developed (https://github.com/ant-trullo/SistersSplitTool). To differentiate the alleles within the 'spot' signal, the position of each putative spot was identified relative to the center of mass of the nucleus across time. These values were organized into two clusters using the Gaussian mixture algorithm from which the tracking in 3D was reconstructed. False detection events generally appeared far from the spot positions and as such could be removed as outliers based on their spatial position. Finally, a manual inspection and correction tool was implemented via a graphical user interface to perform detailed corrections.

Single molecule FISH and smFISH-immunofluorescence

Embryos heterozygous for the allele of interest were fixed in 10% formaldehyde/heptane for 25 min with shaking followed by storage in methanol at −20 °C as previously described[16]. smFISH probes were designed and produced with primary labeling using Quasar 670 by LGC Biosearch Technologies Inc. Probe sequences are listed in Supplementary Data 5. smiFISH probes were designed following a modification of a previous methodology[55] and produced by IDT.

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). For smFISH-IF, the same protocol was performed with addition of the primary antibody (rabbit anti-snail 1:500)[56] in the probe mixture followed by secondary antibody (donkey anti-rabbit Alexa Fluor 488 1:500, Life Technologies) during PBT washes on the second day.

Fixed imaging

Fixed sample imaging was performed on an LSM 880 with an Airyscan module (Zeiss). Z-planes were acquired with 0.33 μm spacing to a typical depth of 25–30 μm from the apical surface of the embryo using laser scanning confocal in Airyscan super-resolution mode with a zoom of 3.0. DAPI excitation was performed with a 405 nm laser, secondary Alexa Fluor 488 excitation with a 488 nm laser (8μW), and Q670 with a 633 nm laser (11.5μW), 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™[52–54] that has been previously published[57]. Data was acquired in two channels, one for nuclei and the other for transcription, both in 3D (ZXY). 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, followed by manual inspection and selection. 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 frame. These Z-frames were then combined in 3D to have a preliminary 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 these points 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™[52–54] enabled by a graphical user interface (NucleiTracker3D, https://github.com/ant-trullo/NucleiTracker3D). Raw data consisting of a two channel TZXY series with Sna^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 labeled 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 neighboring nuclei, privileging combinations with the fewest sub-components. Segmented nuclei were tracked by sequentially projecting the *t-1* time point nuclear mask onto the frame of time point *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 Sna$^{\text{Llama}}$-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[58].

Deconvolution analysis of the data using BurstDeconv

We use BurstDeconv[14] to obtain, for each transcription site, the sequence of processive transcription initiation events. The method considers that the single site MS2 signal is a sum of identical single polymerase contributions translated in different positions corresponding to the initiation events. The signal is calibrated in terms of numbers of polymerases using smFISH[16], hence the amplitude of the signal polymerase contribution is equal to one. The dwell time, defined as the duration of the single polymerase contribution signal, is obtained using the autocorrelation function of the stationary single site MS2 signal in nc13[14,17]. The predicted mean RNA Pol II elongation rates are 35 bp/s in nc13 and 25 bp/s in nc14, and thus deconvolution was performed using a polymerase speed of 25 bp/s. Finally, the processive transcription initiation events positions are obtained by combining a genetic algorithm with a local optimisation procedure to minimise the least squares distance between the predicted and observed signal. One should note that this procedure, initially applied to stationary signals[14–16], can also be applied without modification to non-stationary signals.

Multi-exponential regression and multiple state transcription model reverse engineering using the survival function on stationary segments of the signal

The single-site MS2 signal can be approximated as stationary within segments corresponding to two regimes: the fully unrepressed regime during nuclear cycle 13 and the fully repressed regime during the terminal segment of the nuclear cycle 14 signal. On such segments we use the Kaplan-Meyer method to compute a survival function, representing the nonparametric cumulative distribution function of the waiting times separating successive transcription events. A multi-exponential fitting of the survival function using N exponentials ($N = 2,3$ for two and three states models, respectively) is followed by the reverse engineering of the transcription model using a symbolic method described previously[14,19]. The multi-exponential regression determines the number of states based on parsimony. We select the simplest model that provides a good fit to the survival function, as judged by three criteria: the least-squares objective function, the confidence interval of the Kaplan-Meier estimate, and the Kolmogorov-Smirnov test. The justification of three-state model topology is expanded upon in the **Supplementary Text** 1.

Determining the dwell time from autocorrelation

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

$$R \sim G(\Delta t + d) - 2G(\Delta t) + G(\Delta t - d) \tag{1}$$

where d is the dwell time, $G(x) = -xH(-x)$ and

$$H(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{2}$$

is the Heaviside function. A derivation of (1) has been previously published[17].

For the purposes of this research, the stationary signal from nuclear cycle 13 was used to fit the dwell time (Supplementary Fig. 8).

Estimating the time-dependent repression

The MS2 signal is non-stationary during the nuclear cycle 14. Typically, we notice a decrease of the signal amplitude, suggesting increasing repression. To characterize the time dependence of the repression, we use a moving window method to estimate the time dependence of the mean waiting time $<\tau>$ between successive processive transcription initiation events. As shown elsewhere[19], there is a general formula relating the mean $<\tau>$ and $p_{ON} \cdot k_{ini}$, where p_{ON} is the probability of the transcribing (ON) state and k_{ini} is the transcription initiation rate in this state. This formula is 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 $T \cdot p_{ON} \cdot k_{ini}$, because in the state ON, the promoter initiates with intensity k_{ini} and the total time spent in the ON state is $T \cdot p_{ON}$.

It follows that:

$$<\tau> = \frac{1}{p_{ON} \cdot k_{ini}} \tag{3}$$

For a stationary signal $p_{ON} \cdot k_{ini}$ and thus $<\tau>$ are functions of time. For increasing repression, $<\tau>$ is increasing (the initiation events are rarer). Therefore, the estimate of $<\tau>$ and implicitly that of $p_{ON} \cdot k_{ini}$ is a method to gauge repression. To estimate $<\tau>$, we define a narrow moving window centered on successive time frames and consider all the waiting times from all transcription sites contributing to signal observed in the moving window, gathering sites observed in several movies for enriched statistics. 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 $<\tau>$. We compute the uncertainty bounds of the mean for each movie independently using the central limit theorem with a 95% confidence interval, and then plot the minimum lower bound and the maximum upper bound over all the movies.

Bayesian change point detection for determining the onset of repression

The BCPD method[22] 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. The distribution of the run length $r(t)$, defined as the time since the most recent change point, is learned from the data using this method. At each time step, $r(t)$ increases by 1 if there is no change in the distribution, or it 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 single nucleus 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 Fig. 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. 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 autorepression by considering that the kinetic parameters

depend on the protein level according to decreasing and increasing Hill functions, respectively (Supplementary Fig. 4A). 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 Fig. 4A 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 Fig. 4B 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 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.

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[14]. 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[14].

Homozygous $sna^{MS2/MS2}$ data analysis

For $sna^{MS2/MS2}$ movies, deconvolution of the data was processed in two separate approaches: (1) the alleles were segregated based on the first (1st allele−paired) and second (2nd allele−paired) activated allele in a single nucleus, or (2) all alleles were pooled regardless of their 'mother' nucleus or activation order and randomly allocated to one of two pools ("1st allele−random" or "2nd allele−random"). For the paired allele pools, alleles were retained only if both alleles in a nucleus entered stable repression to avoid bias, which was >95% of nuclei analyzed. For the randomly assigned pools, all alleles that entered stable repression were retained.

The Hill equation model for the mean mRNA production rate

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

$$(p_{ON} \cdot k_{ini})([P]) = V_{min} + (V_{max} - V_{min}) \frac{\theta^n}{\theta^n + [P]^n} \tag{5}$$

where [P] is the repressor protein concentration, V_{max} and V_{min} are the maximal and minimal mRNA production rates, θ is the threshold repressor protein concentration at half-maximal repression, and n is the Hill coefficient. As suggested in its first utilisation by Hill[25], n does not necessarily take integer values. Values $n > 1$ indicate positive cooperativity, meaning that binding of one molecule of repressor to DNA facilitates the binding of additional molecules and/or enhances their repressive effect.

If the protein level and the product $p_{ON} \cdot k_{ini}$ (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 of the identity of the repressor or mechanism of repression.

Stochastic model of transcriptional repression

The Hill equation model for the mean mRNA production is useful for a first analysis of the data. It has the advantage of estimating a Hill coefficient which is a measure of cooperativity. This model does not propose mechanisms or dynamics for transcriptional repression. Therefore, we need a dynamical model of the transcriptional process under repression. Furthermore, the Hill equation cannot explain the stochastic fluctuations in mRNA production. Using the BCPD method, we extracted repression onset times from experimental data and observed that these times vary across different nuclei. A stochastic model is therefore needed to gain more insight into the distribution of these times.

We therefore propose a discrete state transcription model for the endogenous promoter, in which the transcription dynamics is described by the stochastic transition between states. To account for the effect of the repressor on transcription in the simplest way, we consider that the transition rates between discrete promoter states are modulated by the repressor in a Hill-dependent manner.

The objectives of this mathematical model are to (1) check whether our model can reproduce the distribution of repression onset times extracted using BCPD from the experimental data in nc14 and (2) predict the effect of cooperativity on coordinated repression (by scanning over different ranges of Hill coefficient and threshold values).

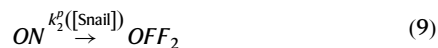To build this model, we consider the following experimental observations:

1. The BurstDeconv analysis of stationary segments of the single nucleus MS2 signal indicates a two-state model for nc13 and a three-state model for the last segment of nc14 during repression. The three-state model includes two OFF states: a long one, denoted OFF1, and a short one, denoted OFF2. This suggests a transition from a two-state system to a three-state system during nc14. We do not know a priori which of the states (OFF1 or OFF2) is newly introduced and which is a continuation of the OFF state from nc13/early nc14.

2. The transgene experiments show that when the number of Sna binding sites and therefore the repression is weakened, there is a transition from a three state model to a two state model. This is consistent with observation 1.

3. The BurstDeconv analysis of different genotypes suggests that the switching rate constants between discrete states in the three state model depend on repression as follows (see Supplementary Fig. 10A–G):
   a. $k_1^p = 1/T_{OFF1}$ decreases strongly with repression,
   b. k_1^m increases strongly with repression, with k_1^m being very small (its inverse corresponding to hours) when repression is very weak,
   c. k_2^m increases strongly with repression,
   d. $k_2^p = 1/T_{OFF2}$ has a mild increase with repression,

4. k_{ini} is very similar for nc13 and the last segment of nc14.

According to 3b, the long state OFF1 is very rare (practically not accessible) at weak repression, and therefore the long OFF1 likely corresponds to the new state induced by repression.

Here we hypothesized that weakening the repression by eliminating Sna binding sites in transgenes is equivalent to decreasing the concentration of Sna for endogenous promoters. We do not exclude the possibility that other TFs (activators or repressors) may also modulate the expression from the endogenous locus, but we consider that the main contribution to the observed variation of parameters is due to changes in the Sna concentrations.

Thus, the transcriptional process is described as a three state Markovian model and governed by the following set of chemical

reactions:

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

$$ON \xrightarrow{k_1^m([Snail])} OFF_1 \tag{7}$$

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

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

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

The Hill-like dependence of parameters $k_1^p([Snail])$, $k_2^p([Snail])$, $k_1^m([Snail])$ 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):

$$k_1^p([Snail]) = (k_1^p)_{rep} + \left((k_1^p)_{act} - (k_1^p)_{rep} \right) \frac{\theta^n}{\theta^n + [Snail]^n} \tag{11}$$

$$k_2^p([Snail]) = (k_2^p)_{act} + \left((k_2^p)_{rep} - (k_2^p)_{act} \right) \frac{[Snail]^n}{\theta^n + [Snail]^n} \tag{12}$$

$$k_2^m([Snail]) = (k_2^m)_{act} + \left((k_2^m)_{rep} - (k_2^m)_{act} \right) \frac{[Snail]^n}{\theta^n + [Snail]^n} \tag{13}$$

$$k_1^m([Snail]) = (k_1^m)_{rep} \frac{[Snail]^n}{\theta^n + [Snail]^n} \tag{14}$$

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.

The rate parameters $k_{ini}, (k_2^p)_{act}, (k_2^m)_{act}, (k_2^m)_{rep}, (k_2^p)_{rep}, (k_1^m)_{rep}$, $(k_1^p)_{rep}$ were chosen equal to the already estimated values, in the active nc13 and repressed nc14 regimes (Supplementary Data 1, 2 state and 3 state non-sequential model). All parameters except k_{ini} depend on [Sna].

The model remains with three free parameters n, θ, $(k_1^p)_{act}$ that were fitted to the $p_{ON}.k_{ini}$ vs [Sna] curves and to the distribution of repression onset times data. The result of the fit is given in Supplementary Data 7.

The Gillespie algorithm was used to numerically simulate the model. To account for post-mitotic lag in transcriptional reactivation, we assumed that transcription began after a lag time in nc14, which varied for each nucleus. The simulated model distribution of lag times was sampled from the experimental post-mitotic lag duration of the *sna* gene, which was obtained using the BurstDeconv deconvolution output. In these simulations, the protein signal was considered deterministic; in other words, different nuclei receive regulatory cues that depend on the time point only.

A synthetic MS2 signal was computed for each simulation using the dwell time parameter 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 $\{(k_1^p)_{act}, n, \theta\}$ 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 $\{(k_1^p)_{act}, n, \theta\}$ 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. Figure 6c shows the results of the best fit simulations. We then chose the best values of $\{(k_1^p)_{act}, n, \theta\}$ that fit the data according to the sum of squared error between $p_{ON}.k_{ini}$ of the data and the simulations. Figure 6d shows the distribution of the repression onset time computed for these parameters. The good agreement between predicted and experimental distributions validates the stochastic model.

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 Data 6. Analysis was performed using Microsoft Excel and Prism (Graphpad 9.1.1).

Snail binding site identification

Sna ChIP data was obtained from previously published data (GSE68983)[24]. To identify potential transcription factor binding sites, we employed the FIMO (Find Individual Motif Occurrences) tool[59] with Sna motifs obtained from the JASPAR database[60]. The significance threshold for motif matches was set at $p < 1e-3$. Sna binding sites were queried in the local neighborhood of the highest (>500) ChIP signal for Sna.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data supporting the findings of this study are available from the corresponding authors upon request.

Code availability

BurstDECONV source code is available through GitHub at https://github.com/oradules/BurstDECONV, and 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 post-processing tool developed for this study at https://github.com/ant-trullo/SpotsFiltersTool. NucleiTracker3D is available via GitHub at https://github.com/ant-trullo/NucleiTracker3D. The BCPD algorithm is available via GitHub at https://github.com/mariadouaihy/BCPD_inhomogeneous_transcriptional_signal. Raw imaging data is available upon request.

References

1. Courey, A. J. & Jia, S. Transcriptional repression: The long and the short of it. *Genes Dev.* **15**, 2786–2796 (2001).
2. 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).
3. Hanna-Rose, W. & Hansen, U. Active repression mechanisms of eukaryotic transcription repressors. *Trends Genet* **12**, 229–234 (1996).
4. Hemavathy, K., Ashraf, S. I. & Ip, Y. T. Snail/Slug family of repressors: Slowly going into the fast lane of development and cancer. *Gene* **257**, 1–12 (2000).
5. Rembold, M. et al. A conserved role for Snail as a potentiator of active transcription. *Genes Dev.* **28**, 167–181 (2014).

6. 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).

7. 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).

8. De Craene, B. et al. Epidermal Snail expression drives skin cancer initiation and progression through enhanced cytoprotection, epidermal stem/progenitor cell expansion and enhanced metastatic potential. *Cell Death Differ*. **21**, 310–320 (2014).

9. Meeussen, J. V. W. & Lenstra, T. L. Time will tell: Comparing timescales to gain insight into transcriptional bursting. *Trends Genet* **40**, 160–174 (2024).

10. Bintu, L. et al. Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev*. **15**, 116–124 (2005).

11. Bertrand, E. et al. Localization of ASH1 mRNA particles in living yeast. *Mol. Cell* **2**, 437–445 (1998).

12. Peiro, S. et al. Snail1 transcriptional repressor binds to its own promoter and controls its expression. *Nucleic Acids Res*. **34**, 2077–2084 (2006).

13. Boettiger, A. N. & Levine, M. Rapid transcription fosters coordinate snail expression in the *Drosophila* embryo. *Cell Rep*. **3**, 8–15 (2013).

14. 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*. **51**, e88 (2023).

15. Tantale, K. K. et al. Stochastic pausing at latent HIV-1 promoters generates transcriptional bursting. *Nat. Commun*. **12**, 4503 (2021).

16. 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).

17. Coulon, A. & Larson, D. R. Fluctuation Analysis. In *Methods in Enzymology* 572 159–191 (Elsevier, 2016).

18. Fukaya, T., Lim, B. & Levine, M. Rapid rates of Pol II elongation in the *Drosophila* embryo. *Curr. Biol*. **27**, 1387–1391 (2017).

19. Radulescu, O. et al. Identifying Markov chain models from time-to-event data: An algebraic approach. *Bull. Math. Biol*. **87**, 11 (2025).

20. 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).

21. 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).

22. Adams, R. P. & MacKay, D. J. C. Bayesian online changepoint detection. Stat.ML at http://arxiv.org/abs/0710.3742 (2007).

23. 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).

24. 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).

25. Hill, A. V. The possible effects of the aggregation of the molecules of hemoglobin on its dissociation curves. *J. Physiol*. **40**, iv–vii (1910).

26. Bothma, J. P. et al. Enhancer additivity and non-additivity are determined by enhancer strength in the Drosophila embryo. *eLife* **4**, e07956 (2015).

27. Dunipace, L., Ozdemir, A. & Stathopoulos, A. Complex interactions between cis-regulatory modules in native conformation are critical for Drosophila snail expression. *Development* **138**, 4075–4084 (2011).

28. Perry, M. W., Boettiger, A. N., Bothma, J. P. & Levine, M. Shadow enhancers foster robustness of *Drosophila* gastrulation. *Curr. Biol*. **20**, 1562–1567 (2010).

29. Dunipace, L., McGehee, J. M., Irizarry, J. & Stathopoulos, A. The proximal enhancer of the snail gene mediates negative auto-regulatory feedback in Drosophila melanogaster. *Genetics*. **230**, iyaf058 (2025).

30. Klindziuk, A. & Kolomeisky, A. B. Theoretical investigation of transcriptional bursting: A multistate approach. *J. Phys. Chem. B* **122**, 11969–11977 (2018).

31. Chen, F. X. et al. PAF1, a molecular regulator of promoter-proximal pausing by RNA polymerase II. *Cell* **162**, 1003–1015 (2015).

32. Wu, C. H. et al. NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in Drosophila. *Genes Dev*. **17**, 1402–1414 (2003).

33. 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).

34. Bellec, M. et al. The control of transcriptional memory by stable mitotic bookmarking. *Nat. Commun*. **13**, 1176 (2022).

35. Tang, X. et al. Kinetic principles underlying pioneer function of GAGA transcription factor in live cells. *Nat. Struct. Mol. Biol*. **29**, 665–676 (2022).

36. Lu, F. & Lionnet, T. Transcription factor dynamics. *Cold Spring Harb. Perspect. Biol*. **13**, a040949 (2021).

37. Pomp, W., Meeussen, J. V. W. & Lenstra, T. L. Transcription factor exchange enables prolonged transcriptional bursts. *Mol. Cell* **84**, 1036–1048.e9 (2024).

38. Zhao, J. et al. Optogenetic dissection of transcriptional repression in a multicellular organism. *Nat. Commun*. **15**, 9263 (2024).

39. Rodriguez, J. et al. Intrinsic dynamics of a human gene reveal the basis of expression heterogeneity. *Cell* **176**, 213–226.e18 (2019).

40. Szczurek, A. T., Dimitrova, E., Kelley, J. R., Blackledge, N. P. & Klose, R. J. The Polycomb system sustains promoters in a deep OFF state by limiting pre-initiation complex formation to counteract transcription. *Nat. Cell Biol*. **26**, 1700–1711 (2024).

41. Fountas, C. & Lenstra, T. L. Better together: How cooperativity influences transcriptional bursting. *Curr. Opin. Genet. Dev*. **89**, 102274 (2024).

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

43. Morgunova, E. & Taipale, J. Structural perspective of cooperative transcription factor binding. *Curr. Opin. Struct. Biol*. **71**, 171–179 (2017).

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

45. Sönmezer, C. et al. Molecular co-occupancy identifies transcription factor binding cooperativity in vivo. *Mol. Cell* **81**, 255–267.e6 (2021).

46. Estrada, J., Wong, F., DePace, A. & Gunawardena, J. Information integration and energy expenditure in gene regulation. *Cell* **166**, 234–244 (2016).

47. 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).

48. 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. *eLife* **12**, RP92083 (2024).

49. Levi, T., Sloutskin, A., Kalifa, R., Juven-Gershon, T. & Gerlitz, O. Efficient in vivo introduction of point mutations using ssODN and a Co-CRISPR approach. *Biol. Proced. Online* **22**, 14 (2020).

50. Ferraro, T. et al. Transcriptional memory in the *Drosophila* embryo. *Curr. Biol*. **26**, 212–218 (2016).

51. Dufourt, J. et al. Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs. *Nat. Commun*. **9**, 5194 (2018).

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

53. Harris et al. Array programming with NumPy. *Nature* **585**, 357–362 (2020).
54. Van Der Walt, S. et al. scikit-image: Image processing in Python. *PeerJ* **2**, e453 (2014).
55. Tsanov, N. et al. smiFISH and FISH-quant – a flexible single RNA detection approach with super-resolution capability. *Nucleic Acids Res.* **44**, e165 (2016).
56. Pimmett, V. L. et al. Optogenetic manipulation of nuclear Dorsal reveals temporal requirements and consequences for transcription. *Development* **152**, dev204706 (2025).
57. Dufourt, J. et al. Imaging translation dynamics in live embryos reveals spatial heterogeneities. *Science* **372**, 840–844 (2021).
58. 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.* **33**, 2839–2850.e7 (2023).
59. Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME suite. *Nucleic Acids Res* **43**, W39–W49 (2015).
60. 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).

Author contributions
Conceptualization: M.L., O.R., V.L.P., M.D. Investigation/Data acquisition: V.L.P., L.M., M.D., M.C., P.G.I. Image analysis: A.T., V.L.P., L.M. Mathematical modeling and code development: M.D., O.R. Stochastic model implementation: O.R. Design and supervision of sna^{Llama} plasmid construct cloning: J.D. Cloning of sna^{Llama} plasmid construct: H.L.-H. Funding acquisition: M.L., O.R. Project administration: M.L. Supervision: M.L., O.R. Writing original draft: V.L.P., M.L. Writing, review, and editing: O.R., M.D., L.M., A.T.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-62907-3.

Correspondence and requests for materials should be addressed to Ovidiu Radulescu or Mounia Lagha.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.