Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs

Jeremy Dufourt[1,5], Antonio Trullo[1,5], Jennifer Hunter[1], Carola Fernandez[1], Jorge Lazaro[1], Matthieu Dejean[1], Lucas Morales[1], Saida Nait-Amer[1], Katharine N Schulz[2], Melissa M Harrison[2], Cyril Favard[3], Ovidiu Radulescu[4] and Mounia Lagha[1,*].
[1]Institut de Génétique Moléculaire de Montpellier, University of Montpellier, CNRS, Montpellier, France
[2]Department of Biomolecular Chemistry, University of Wisconsin School of Medicine and Public Health Madison, WI USA
[3]Institut de Recherche en Infectiologie de Montpellier, CNRS, University of Montpellier, France
[4]DIMNP, UMR CNRS 5235, University of Montpellier, Montpellier, France
*correspondence: mounia.lagha@igmm.cnrs.fr
[5]Equal contribution

Abstract

Pioneer factors can engage nucleosomal DNA, which leads to local chromatin remodeling and to the establishment of transcriptional competence. However, the impact of enhancer priming by pioneer factors on the temporal control of gene expression and on mitotic memory remains unclear. Here we employ quantitative live imaging methods and mathematical modeling to test the effect of the pioneer-factor Zelda on transcriptional dynamics and memory in *Drosophila* embryos. We demonstrate that increasing the number of Zelda binding sites accelerates the kinetics of nuclei transcriptional activation regardless of their transcriptional past. Despite its known pioneering activities, we show that Zelda does not remain detectably associated with mitotic chromosomes and is neither necessary nor sufficient to foster memory. We further reveal that Zelda forms sub-nuclear dynamic hubs where Zelda binding events are transient.

We propose that Zelda facilitates transcriptional activation by accumulating in microenvironments where it could speed the duration of multiple pre-initiation steps.

Introduction

During the first stages of metazoan embryogenesis, the zygotic genome is largely quiescent, and development relies on maternally deposited mRNAs and proteins. Transcriptional activation of the zygotic genome occurs only hours after fertilization and requires specific transcription factors. In *Drosophila*, the transcriptional activator Zelda (Zld) plays an essential role in zygotic genome activation[1].

Zelda exhibits several pioneer factor properties such as (i) priming *cis*-regulatory elements prior to gene activation, which establishes competence for the adoption of cell fates, and (ii) opening chromatin, which facilitates subsequent binding by classical transcription factors. Indeed, hours before their activation, Zelda binds to thousands of genomic regions corresponding to enhancers of early-activated genes[2,3]. Zelda binding is associated with modifications of the associated chromatin landscape and overall chromatin accessibility[2-5].

Moreover, this ubiquitously distributed activator is able to potentiate the action of spatially restricted morphogens (*e.g* Bicoid and Dorsal)[6,7]. Consequently, target gene response is strengthened by Zelda binding, both spatially and temporally, for developmental enhancers[8] as well as for synthetic enhancers where input parameters are tightly controlled[9].

In addition to these properties, classical pioneer factors bind nucleosomes and are typically retained on the chromosomes during mitosis[10,11]. Mitotic retention of these transcription factors places them as ideal candidates for the transmission of chromatin states during cellular divisions, through a mitotic bookmarking mechanism.

However, the putative function of Zelda in mitotic memory has thus far not been examined. Moreover, all studies concerning Zelda's role as a quantitative timer have been performed on fixed embryos with limited temporal resolution. Thus, the effects of Zelda on the temporal dynamics of transcriptional activation and on transmission of active states through mitosis remain to be addressed.

Here we employ quantitative live imaging methods to directly measure the impact of Zelda on transcriptional dynamics *in vivo*. We show that increasing the number of Zelda binding sites at an enhancer fosters temporal coordination in *de novo* gene activation (synchrony), regardless of past transcriptional states. By monitoring transcription in Zelda-maternally depleted embryos, we reveal that Zelda is dispensable for mitotic memory. On the contrary, we found that Zelda allows bypassing the temporal disadvantage of arising from a transcriptionally inactive mother. Using a mathematical modeling framework, we propose that mitotic memory requires long lasting transitions between chromatin states, incompatible with the function of Zelda in accelerating these transitions.

By analyzing Zelda protein kinetics during many cell cycles, we report that while it is not retained on chromosomes during mitosis, this transcription factor exhibits a highly dynamic behavior. We discovered that intra-nuclear distribution of Zelda is not homogeneous, and that it accumulates in local hubs, where its residence time is surprisingly transient (in the second range). Taken together, our results provide fundamental insights into Zelda-mediated temporal control of transcription during zygotic genome activation *in vivo*.

Results and Discussion

Zelda fosters temporal transcriptional coordination

In the blastoderm embryo, two enhancers control *snail* (*sna*) gene expression, a proximal (primary) and a distal (shadow enhancer)[12,13]. Both enhancers are bound by Zelda (Zld) at early nuclear cycles (nc)[2,3]. When driven by its native enhancers, endogenous expression of *snail* is extremely rapid, precluding analysis of the impact of Zelda on fine-tuning the timing of its activation. We therefore used a previously described truncated version of the *sna* shadow enhancer (*snaE*) that leads to a stochastic activation, compatible with the tracking of transcriptional memory across mitosis[14].

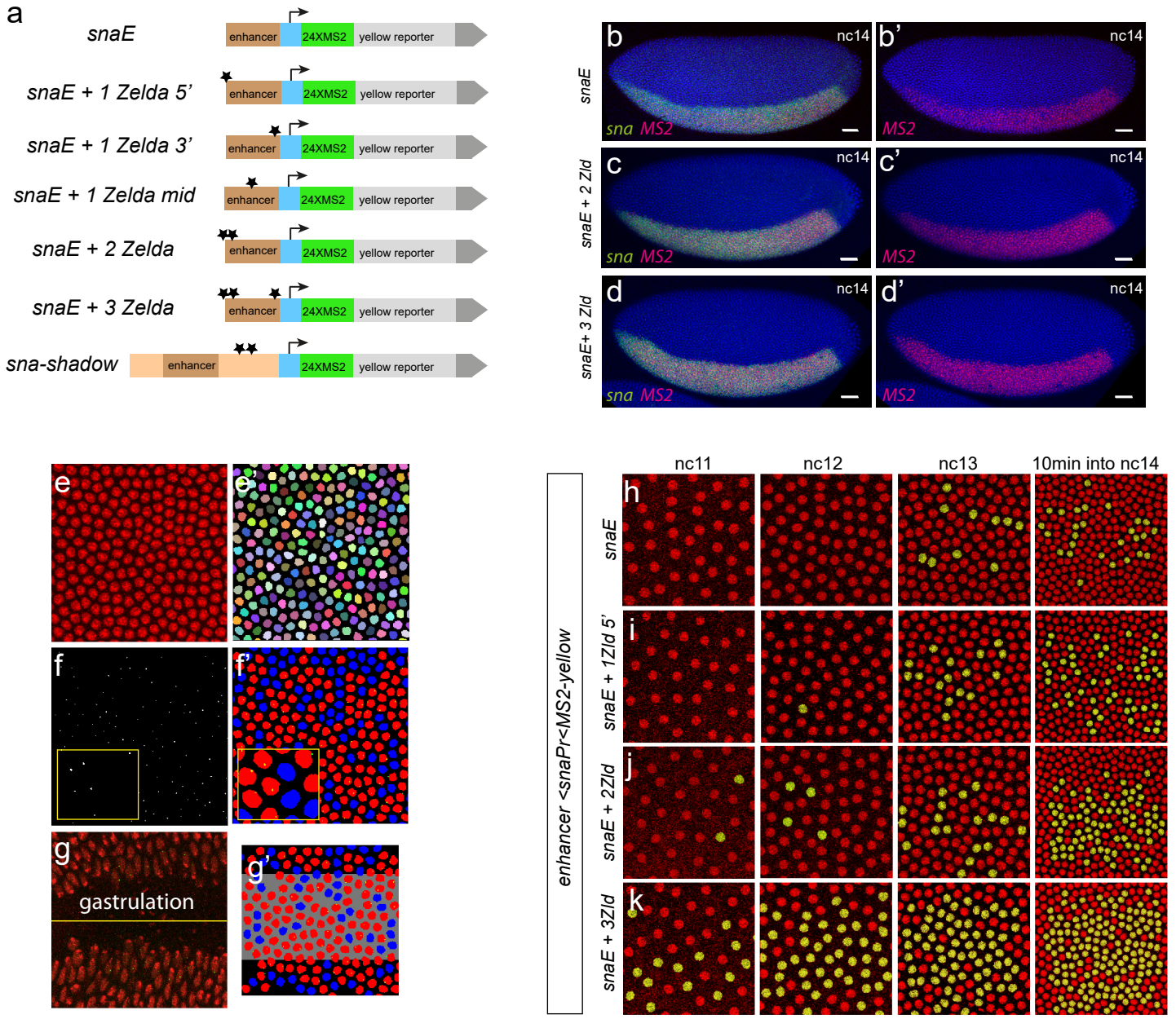
To follow transcriptional dynamics in a quantitative manner, we created a series of *enhancer<promoter<MS2-yellow* transgenes, with a unique minimal promoter (*sna* promoter) and a *MS2-yellow* reporter (Fig. 1a). Upon transcription, MS2 stem-loops are rapidly bound by the maternally provided MCP-GFP, which allows tracking of transcriptional activation in living embryos[15]. To decipher the role of Zelda in *cis*, a variable number of canonical Zld-binding sites (CAGGTAG)[16] were added to the *snaE* in three locations, either close to the promoter (Zld 3'), in the middle (Zld mid) or at a distance (Zld 5') (Fig. 1a).

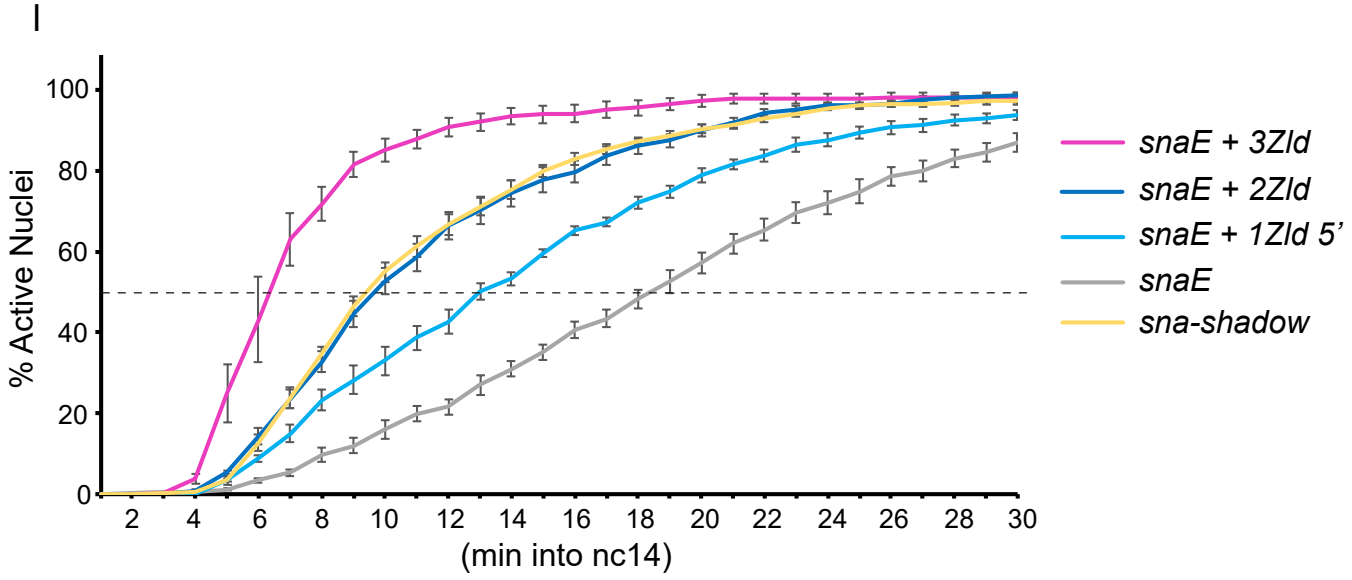
Using fluorescent *in situ* hybridization, we confirmed that the pattern of expression of *snaE-extraZld* transgenes was dictated by the regulatory logic of the *sna* enhancer, and expression was indeed evident in the presumptive mesoderm, similarly to the endogenous *sna* pattern (Fig. 1 b-d'). To precisely measure temporal dynamics, we took advantages of our MCP/MS2 system to track transcriptional activation in living embryos.

First, we implemented an automatic image analysis pipeline that segments nuclei in nc14 (Fig. 1 e, e') and tracks the time of transcriptional activation (Fig. 1 f, f'). In the early fly embryo, a subset of transcriptional activators, including Dorsal, control gene expression in a graded manner[17]. It was therefore important to quantify temporal dynamics of gene activation in a spatially controlled pattern. For this purpose, we used the ventral furrow (site of invagination during gastrulation process) as a landmark to define dorso-ventral coordinates (Fig. 1g, g' and Supplementary Fig. 1a). Unless otherwise indicated, we studied temporal dynamics of gene activation in a region of 50µm centered around the ventral furrow to ensure non-limiting levels of the *sna* activators Dorsal and Twist[18].

We initially examined the impact of Zelda on the timing of gene activation. As identified with static approaches[6,7,16], our dynamic data showed that increasing the number of Zld-binding sites resulted in precocious transcriptional activation (Fig. 1 h-l Supplementary Fig. 1b, c and Supplementary Movie 1-4). Zelda binding impacted not only the onset of gene

Figure1


a
snaE
snaE + 1 Zelda 5'
snaE + 1 Zelda 3'
snaE + 1 Zelda mid
snaE + 2 Zelda
snaE + 3 Zelda
sna-shadow
enhancer
24XMS2
yellow reporter
b
b'
c
c'
d
d'
nc14
sna MS2
MS2
snaE
snaE + 2 Zld
snaE+ 3 Zld
e
e'
f
f'
g
g'
gastrulation
nc11
nc12
nc13
10min into nc14
h
i
j
k
snaE
snaE + 1Zld 5'
snaE + 2Zld
snaE + 3Zld
enhancer <snaPr<MS2-yellow


l
% Active Nuclei
(min into nc14)
snaE + 3Zld
snaE + 2Zld
snaE + 1Zld 5'
snaE
sna-shadow

activation but also the temporal coordination among a spatially defined pattern (*i.e* the presumptive mesoderm, 50µm around the furrow), referred to as synchrony[19] (Fig. 1l and Supplementary Fig. 1c). For example, at precisely 10min into interphase 14, expression of the *snaE* transgene was not coordinated (~21% of active nuclei), whereas the number of active nuclei at this stage increased by adding Zld-binding sites (Fig. 1 h-k). The precise kinetics of gene activation, *i.e* synchrony curves, was quantified as a percentage of active nuclei within a defined spatial pattern for each transgene during the first 30min of nc14. *SnaE* generated a slow activation kinetic, where half of the pattern is activated at t50≈18min (Fig. 1l). Adding only a single CAGGTAG sequence, 5' to the enhancer accelerated this kinetic (t50≈13min). With two extra Zld-binding sites synchrony is further increased (t50≈9.5min). The activation kinetics of two extra canonical Zld sites mimicked the dynamics generated by the intact long shadow enhancer (t50≈9.5min) (Fig. 1l). The most dramatic synchrony shift was caused by the addition of a single CAGGTAG sequence close to the promoter (+1Zld 3') (t50≈6.5min) (Supplementary Fig. 1c). This behavior was not changed when two additional Zld sites were added, suggesting that promoter proximal Zelda binding causes a strong effect, that cannot be further enhanced by the addition of more distant sites (Fig. 1l and Supplementary Fig. 1c). By moving away the CAGGTAG site to the middle of the enhancer, synchrony was decreased (t50≈7.5min) (Supplementary Fig. 1c) suggesting a stronger role for Zelda when positioned closely to the promoter.

In this system, addition of a single CAGGTAG was sufficient to increase the activation kinetics. This is in sharp contrast to what has been reported for a synthetic enhancer, whereby activation only occurred when three or more Zld binding sites were added[9]. This difference is likely due to the ability of Zelda to function cooperatively with other activators of *sna*, which are not present in the synthetic enhancer previously examined[9].

While our initial analysis focused on the region surrounding the ventral furrow, live imaging data also provided insights into the temporal activation in the entirety of the *sna* domain. By temporally tracking activation along the dorso-ventral axis, we could observe the consequences of the Dorsal nuclear spatial gradient[20] (Supplementary Fig. 1d). Indeed, the first activated nuclei were located where activators are at their peaks levels, whereas the late activated nuclei were closer to the mesodermal border, as previously suggested by analysis of fixed embryos and modeling[7,21]. By comparing activation timing between 0-25µm and 25-50µm from the ventral furrow, we show that adding extra Zld-binding sites boosted the spatio-temporal response to the dorso-ventral gradient (Supplementary Fig. 1d). These temporal data are consistent with the idea that Zelda potentiates the binding of spatially restricted morphogens like Bicoid or Dorsal [6-8,22] and thus support the previously described pioneering activities of Zelda[5].

Zelda allows bypassing transcriptional memory

Our ability to track transcriptional activation in live embryos provided an unprecedented opportunity to determine activity through multiple nuclear cycles. Using the sensitized *snaE* transgene, we recently documented the existence of a transcriptional mitotic memory, whereby the transcriptional status of mother nuclei at nc13 influences the timing of activation of the descendants in the following cycle[14]. To address the challenge of following rapid movements during mitosis, we developed a semi-automatic mitotic nuclei tracking software (Fig. 2a) (Trullo et al., in preparation). By combining lineage information with transcriptional status, we quantified the timing of activation of hundreds of nuclei in nc14 in a spatially defined pattern. This allowed us to distinguish those transcriptionally active nuclei that arose from an active mother nucleus (green curves, Fig. 2b-d and Supplementary Fig. 1e, f) from those coming from an inactive mother nucleus (red curves, Fig. 2b-d and Supplementary Fig. 1e, f). Differences between the kinetics of activation of these two populations provide an estimation of transcriptional mitotic memory (Fig. 2b).

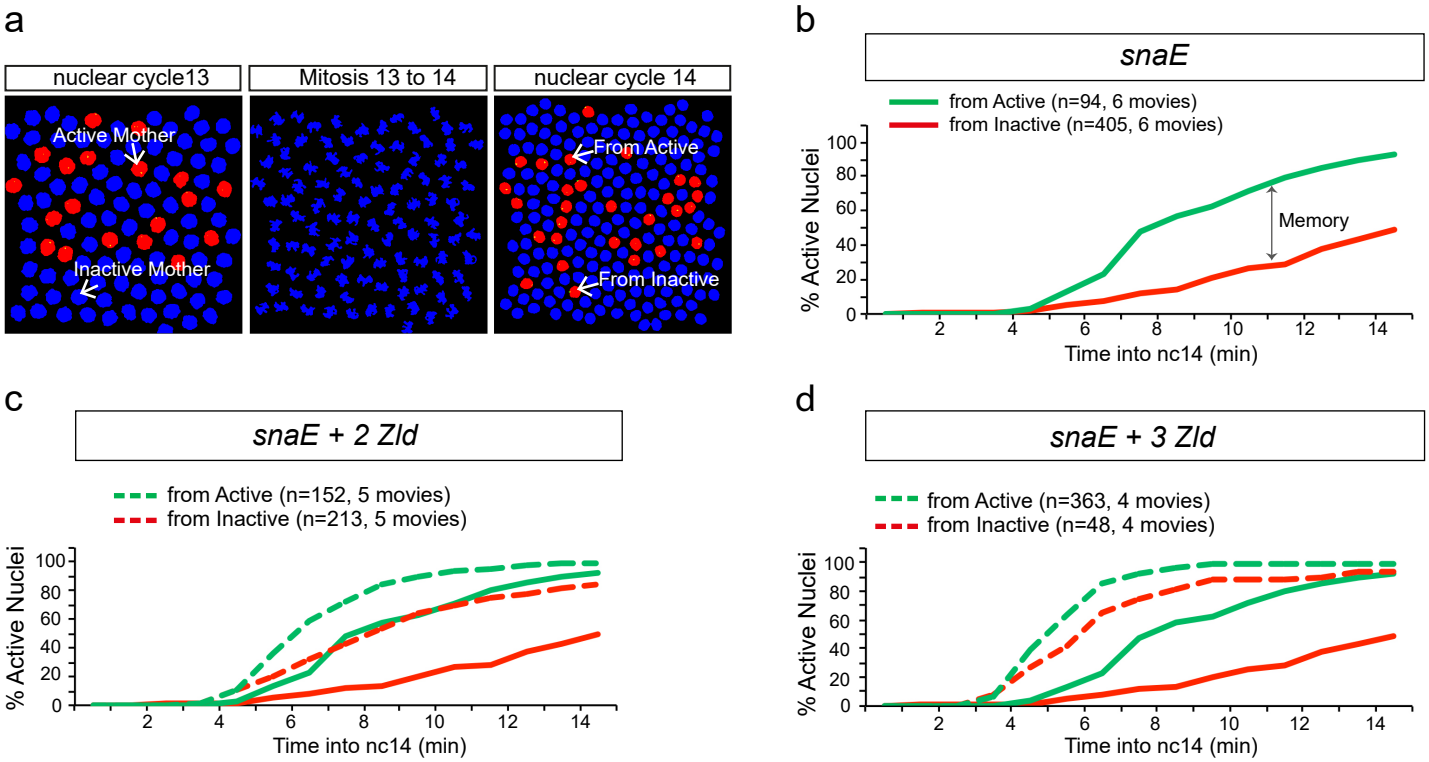
Based on the ability of Zelda to define hundreds of cis-regulatory regions through the establishment or maintenance of accessible chromatin[4,5], we reasoned that Zelda could be in part responsible for the establishment of transcriptional mitotic memory during early development. Thus, we leveraged our live imaging system to directly test the role of additional Zld-binding sites on transcriptional activation through mitosis. Contrary to our expectation, namely that extra Zelda transgenes would exhibit enhanced memory; adding extra Zld-binding sites in *cis* augmented the speed of activation in both nuclei derived from active mothers and those from inactive mothers (Fig. 2 c, d and Supplementary Fig. 1e, f). Increasing the number of Zld-binding sites reduced the differences in the timing of activation for descendants from active mothers and their neighboring nuclei, arising from inactive mothers. Our results suggest that extra Zld-binding sites accelerate transcriptional dynamics regardless of the past transcriptional status. Zelda seems thus capable of bypassing transcriptional memory by compensating for the negative effect of having an inactive mother nucleus.

Zelda is neither sufficient nor necessary to trigger transcriptional mitotic memory

We then sought to test whether Zelda was required for this memory. Activation kinetics for descendants from active and inactive nuclei were quantified after reducing maternal Zelda expression using RNAi. Validating this approach, germline expression of *zld*-RNAi[7,23] reduced the levels of *zld* mRNA to 2-4% of that found in control *white*-RNAi embryos (Supplementary Fig. 3a).

Reduction of maternal Zelda decreased the synchrony of activation of our transgenes with additional Zld-binding sites (Fig. 3a-f, Supplementary Fig. 3b-c and Supplementary Movie 5, 6). In *zld*-RNAi embryos, temporal coordination was reduced, and all transgenes showed a t50≈16-19 min (Fig. 3a-f, Supplementary Fig. 3b-c and Supplementary Movie 5, 6). This confirms that the previously identified acceleration of transcriptional activation upon addition of extra Zld-binding sites is due to Zelda activity on these enhancers. Further confirming the role of maternally contributed Zelda in potentiating Dorsal-dependent gene expression, transcriptional activation was

Figure2


a
nuclear cycle13
Mitosis 13 to 14
nuclear cycle 14
Active Mother
Inactive Mother
From Active
From Inactive
b
snaE
from Active (n=94, 6 movies)
from Inactive (n=405, 6 movies)
% Active Nuclei
Memory
Time into nc14 (min)
c
snaE + 2 Zld
from Active (n=152, 5 movies)
from Inactive (n=213, 5 movies)
% Active Nuclei
Time into nc14 (min)
d
snaE + 3 Zld
from Active (n=363, 4 movies)
from Inactive (n=48, 4 movies)
% Active Nuclei
Time into nc14 (min)

restricted to a reduced region of about ≈25µm around the 'pseudo-furrow' in Zelda maternally-depleted embryos (Supplementary Movie 6). Together these data demonstrate that expression of our *snaE* transgenes reflect the general properties of Zld-mediated gene expression.

We then quantified the level of memory in the *zld*-RNAi embryos with three *snaE<MS2* transgenes containing varying number of Zld-binding sites. In spite of the maternal reduction of Zelda, a strong memory bias was still observed (Fig. 3g, h, Supplementary Fig. 3d). We conclude that although possessing several key properties of a pioneer-factor, Zelda is not necessary to elicit heritability of active transcriptional states through mitosis.

Altogether our results show that Zelda accelerates transcription and fosters synchrony, despite playing no role in transcriptional memory maintenance.

To further examine whether Zelda binding to an enhancer was sufficient to trigger a memory bias, temporal activation from a second regulatory element (*DocE*) was examined (Supplementary Fig. 2). This region is located in a gene desert within the *Doc1/2* locus and is characterized by high Zelda occupancy in early embryos (Supplementary Fig. 2a, b)[3]. We created an *MS2* reporter transgene with this putative enhancer and found that it drives expression in the dorsal ectoderm in a pattern overlapping endogenous *Doc1* (Supplementary Fig. 2c, c') and *Doc2* mRNA expression (data not shown). Upon *zld*-RNAi, expression from this enhancer was strongly reduced, confirming a role for Zelda in driving expression from the *DocE* element (data not shown).

During nc13, our *DocE* transgene was stochastically activated in three distinct dorsal domains, where we could track transcriptional mitotic memory (Supplementary Fig. 2c). None of the three domains (anterior, central, posterior) revealed a bias in the timing of activation in nc14 for descendants of active mothers compared to those arising from inactive mothers (Supplementary Fig. 2d-g). We therefore conclude that this *DocE* element does not trigger memory, despite being bound by Zelda. Thus, in a model system that shows stochastic activation and that uses Zelda, we did not observe any mitotic memory, further confirming the lack of involvement of Zelda in memory. Moreover these results show that experiencing transcription at a given cycle does not necessarily lead to a rapid *post*-mitotic reactivation in the following cycle.

Modeling the impact of Zelda on synchrony and memory

To gain insights into the role of Zelda on transcriptional dynamics, we developed a simple mathematical framework that describes the waiting times prior to the first detected transcriptional initiation event in nc 14, as a sequence of discrete transitions (see Methods and Supplementary Methods). This model considers that the activation of transcription follows a stepwise progression through a series of nonproductive OFF events that precede activation (ON event) (Fig. 4a). Given the various well characterized steps required prior to productive transcriptional elongation (*e.g.* promoter opening, transcription factor binding, pre-initiation complex recruitment…), it is reasonable to consider a series of transitions that a promoter must 'travel' through prior to activation, with an allocated duration[11,24].

In this model, the average number of rate-limiting transitions is provided by the parameter *'a'* while the duration of the transitions by parameter *'b'* (expressed in seconds and considered the same for all transitions, a more complex model with heterogeneous 'b' is discussed in the Supplementary methods) (Fig. 4a). The model predicts that the distribution of waiting times prior to activation for a given gene in a set of nuclei can be described as a mixture of gamma distributions, in which the mixing occurs over the shape parameter. In this mixture model, the shape parameter is the number of transitions *'a'* and the constant scale parameter is '*b'* (Fig. 4a, Methods and Supplementary Methods). The model also predicts that the average number of steps (parameter '*a*') can be roughly computed from the first two moments of the distribution of waiting times (Eq.S3 Supplementary Methods). This rough but direct estimate shows that *'a'* is less than three for all genotypes. Thus, we have restricted our model to three limiting transitions and consequently to three nonproductive states (OFF1, OFF2 and OFF3) (Fig. 4a).

Taking advantage of the significant number of nuclei tracked in this study, we could fit this model to our data and estimate more accurately the parameters '*a'* and '*b'*, from populations of nuclei (descendants of active and descendants of inactive nuclei) (Supplementary Table1).

The distribution of waiting times for descendants of active nuclei and inactive nuclei were clearly distinct, indicating distinguishable scale and average shape parameters for these two populations (Fig. 4b). For *snaE*, the temporal behaviors of descendants of active and inactive mothers differed in both parameters '*a'* and '*b'*, with both being lower for descendants of active nuclei (Fig. 4c, d). Thus, our model suggests that descendants of active nuclei have fewer and shorter transitions prior to activation than descendants of inactive nuclei.

For transgenes with additional Zld-binding sites, the predictions for '*a*' are not obviously affected (Fig. 4c). This suggests that Zelda binding does not dramatically alter the number of transitions required for activation. By contrast, the estimates for '*b'* are systematically lower for descendants of active mothers and inactive mothers upon the addition of Zld-binding sites (Fig. 4d).

Thus, while the accelerated transcriptional activation observed with increased numbers of Zld-binding sites could be due to either a decrease in the duration of the transitions '*b*' or to the number of transitions '*a*' (Supplementary Fig. 4), our modeling suggests that Zelda accelerates transcriptional activation primarily by speeding the lifetime of the transitions. Consistent with these findings, upon maternal reduction of Zelda, the '*b'* parameter is augmented, when compared to estimates in *white RNAi* controls (Fig. 4d').

In contrast to nuclei expressing the *snaE* transgene, the relation between estimates of '*a*' for descendants of active mothers and those of inactive mothers is more complex for extra-Zld transgenes (Fig. 4c). In order to gain some insights, we have used numerical simulations of an extended version of our model that includes modified transition dynamics during mitosis (for example, backward transitions, see Supplementary

Figure3

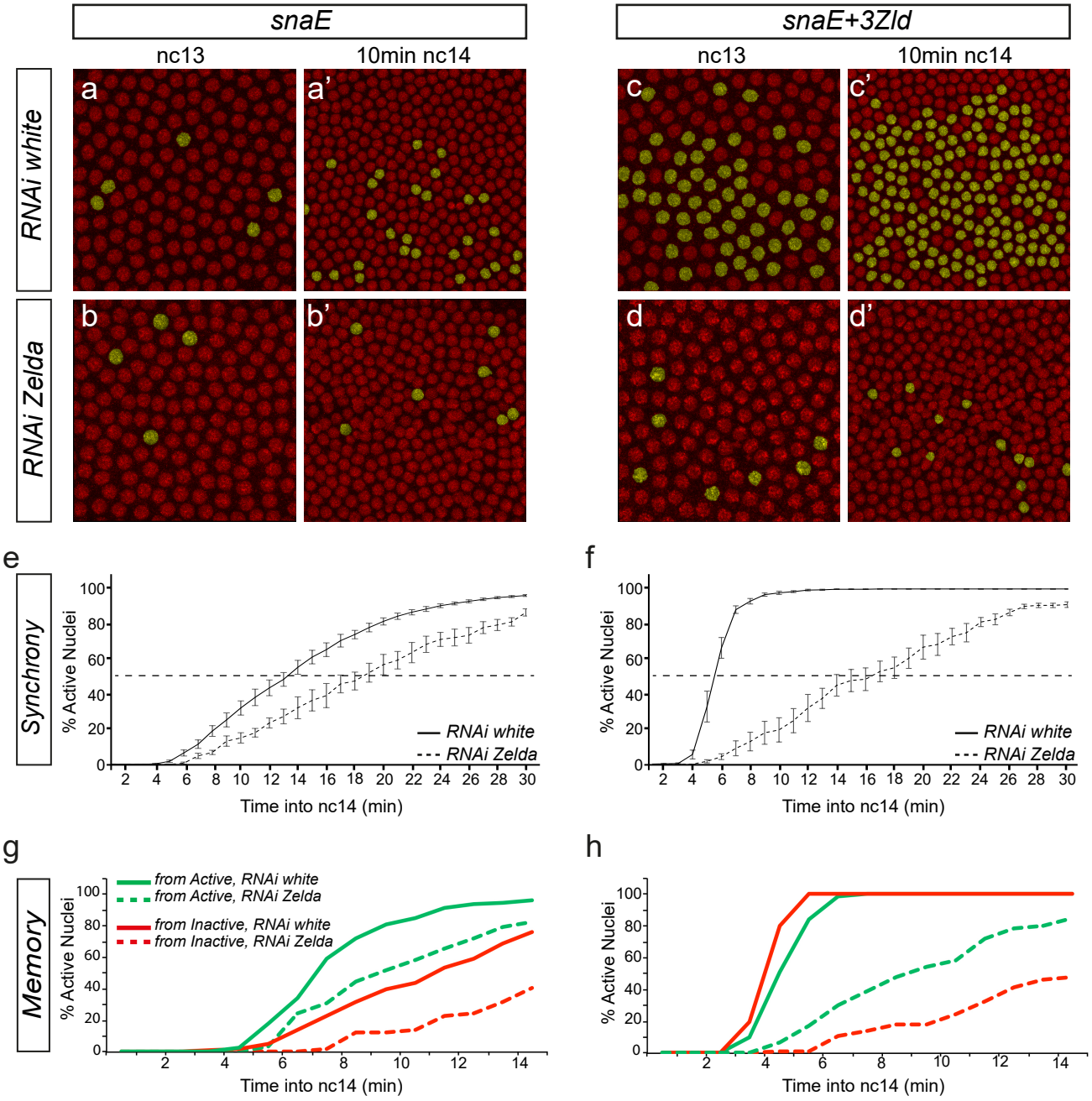
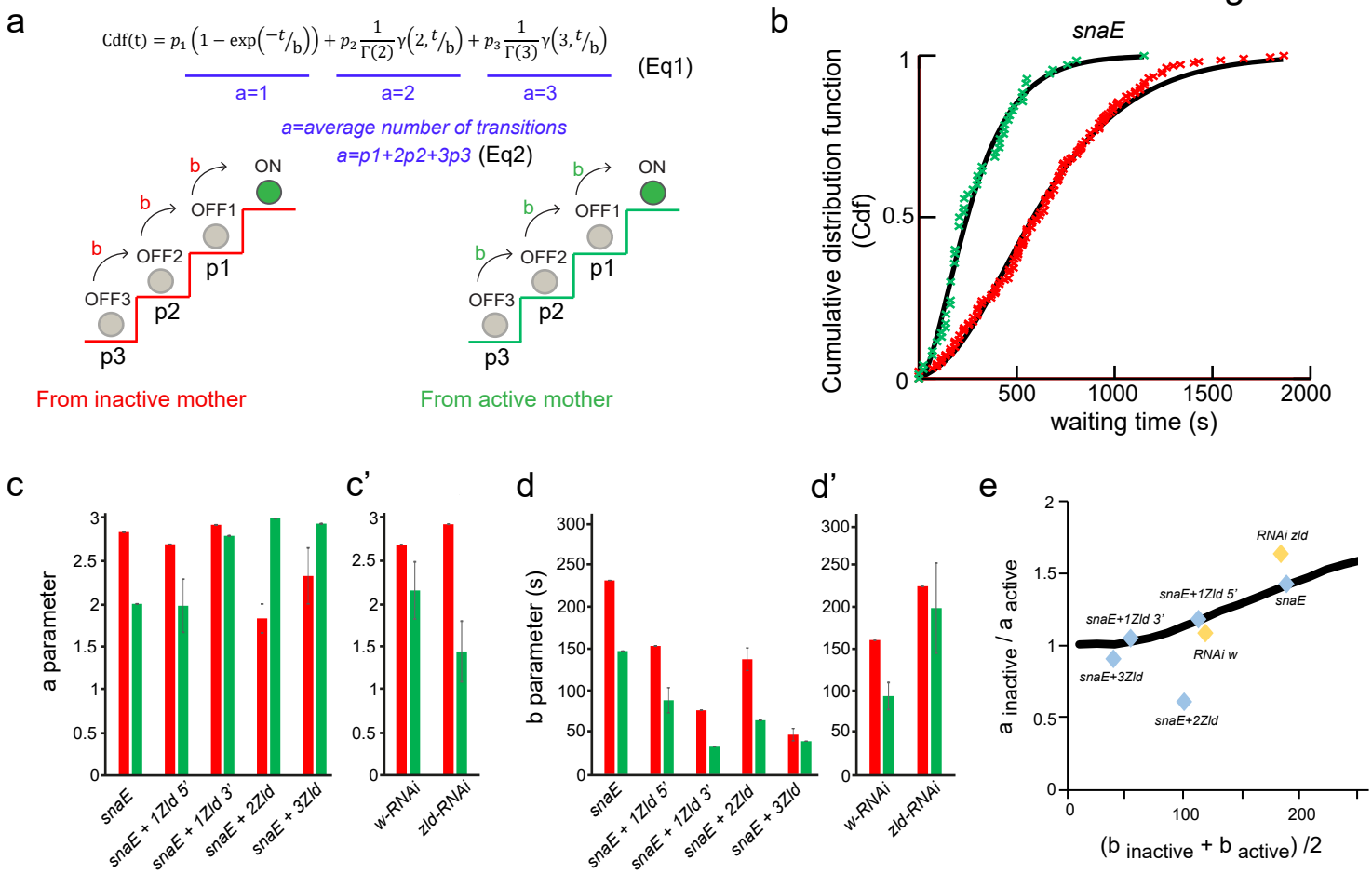

snaE
snaE+3Zld
nc13
10min nc14
nc13
10min nc14
RNAi white
RNAi Zelda
a
a'
b
b'
c
c'
d
d'
e
f
Synchrony
% Active Nuclei
Time into nc14 (min)
RNAi white
RNAi Zelda
g
h
Memory
from Active, RNAi white
from Active, RNAi Zelda
from Inactive, RNAi white
from Inactive, RNAi Zelda

Figure 4


a
$\mathrm{Cdf(t)} = p_1\left(1-\exp\left(-t/b\right)\right) + p_2\frac{1}{\Gamma(2)}\gamma\left(2, t/b\right) + p_3\frac{1}{\Gamma(3)}\gamma\left(3, t/b\right)$ (Eq1)
a=1
a=2
a=3
a=average number of transitions
a=p1+2p2+3p3 (Eq2)
ON
OFF1
OFF2
OFF3
b
p1
p2
p3
From inactive mother
From active mother
b
snaE
Cumulative distribution function (Cdf)
waiting time (s)
c
a parameter
snaE
snaE + 1Zld 5'
snaE + 1Zld 3'
snaE + 2Zld
snaE + 3Zld
c'
w-RNAi
zld-RNAi
d
b parameter (s)
d'
e
a inactive / a active
(b inactive + b active) /2
RNAi zld
snaE+1Zld 5'
snaE
snaE+1Zld 3'
RNAi w
snaE+3Zld
snaE+2Zld

methods). The simulations predict that the bias in *'a'* values (difference in the number of steps to reach active state, evaluated by $a_{inactive}/ a_{active}$ and referred to as memory bias) is correlated to '*b'* values (Supplementary Methods). Remarkably, we could obtain this correlation with our temporal data (Fig. 4e). Hence, enhancers with large '*b'* values (having slow transitions) tend to have a strong $a_{inactive}/ a_{active}$ ratio, whereas enhancers with small '*b'* values have a vanishing memory bias (Fig. 4e and Supplementary Methods). If we interpret transcriptional memory as the heritable reduction in the number of transitions required to reach the active state, this result suggests that stronger memory occurs when transitions are slower.

In summary, using a modeling approach, we converted our measurements of the timing of transcriptional activation in hundreds of single nuclei into discrete metastable states preceding gene activation. Although purely phenomenological, this model suggests that transcriptional memory is supported by a sequence of relatively long-lived metastable states preceding activation, which could be maintained by mitotic bookmarking. Zelda potentiates transcription by accelerating these states and is thus incompatible with mitotic memory, consistent with our observations in Zelda depleted embryos.

Quantitative kinetic analysis reveals the dynamic properties of Zelda

Based on the characteristics of pioneer factors, we had expected a role for Zelda in retaining transcriptional memory through mitosis. Thus, our genetic data and modeling indicate that Zelda was not the basis of memory. To better understand this result, we examined Zelda dynamics in living embryos and particularly during mitosis (Fig. 5a, b and Supplementary Movie 7). We took advantage of an endogenously tagged fluorescent version of Zelda[25]. The *GFP-zld* flies are homozygous viable, thereby showing that the GFP-tagged version of Zelda retains wild-type physiological properties.

GFP-Zld localization with chromatin was examined through the mitotic cycle (Fig. 5, Supplementary Fig. 5 and Supplementary Movie 7) and demonstrated that Zelda is not detectably retained on the chromosomes during mitosis (Fig. 5a, b and Supplementary Movie 7), similar to what was shown in fixed embryos[26]. This is in comparison to the transcription factor GAF (GAGA Associated Factor) (Supplementary Fig. 5a). While most GFP-Zld molecules are localized in nuclei in interphase, they were redistributed to the cytoplasm during mitosis (Fig. 5b). Our live imaging approach revealed a highly dynamic behavior of Zelda, whereby Zelda comes back to the nucleus rapidly at the end of mitosis prior to complete chromosome decompaction (Fig. 5a, b, Supplementary Fig. 5b and Supplementary Movie 7). When compared to Pol II-Ser5P by immuno-staining on embryos exhibiting a mitotic wave, GFP-Zld signal filled the nuclei prior to polymerase (Supplementary Fig. 5b). By comparing Zelda cell cycle dynamics to B-icoid, we found that these two DNA binding proteins are similarly evicted during mitosis and are quickly re-localized to chromatin at the end of mitosis (Supplementary Fig. 5c and Supplementary Movie 8). Nonetheless, quantitative comparisons revealed that Zelda entry into the nucleus at the end of mitosis was faster than that of Bicoid (Supplementary Fig. 5d).

To better characterize Zelda kinetics, we performed intra-nuclear Fluorescence Recovery After Photobleaching (FRAP) experiments and determined that Zelda recovery was remarkably fast in interphase nuclei (Fig. 5c). Because FRAP is not well suited for fast moving proteins, we performed Fluorescence Correlation Spectroscopy (FCS) in living cycle14 embryos (Fig. 5d, d') and estimated the kinetic properties of Zelda (Fig. 5e, f). FCS experiments for Zelda were compared to Ash1, whose kinetics have been documented by both FRAP and FCS in early *Drosophila* embryos[28] and thus serves as a well-defined point of reference (Supplementary Fig. 5e-f). As expected for chromatin-binding proteins, auto-correlation curves for Zelda and Ash1 show more than one unique characteristic time (Fig. 5d' and Supplementary Fig. 5e'). To estimate chromatin binding kinetics, FCS curves (86 curves for Ash1 and 109 curves for Zelda) were fitted with a reaction-diffusion model[29]. As previously revealed in Steffen et al.[28], we found that Ash1 is highly dynamic and estimate similar diffusion coefficients ($D_f \approx 4.5 \mu m^2.s^{-1}$ versus $D_{f(Steffen)} \approx 4.98 \mu m^2.s^{-1}$) (Fig. 5e), thus validating our FCS experiments. Similarly to other transcription factors (*e.g* Bicoid) we found that Zelda is a fast diffusing protein with apparent diffusion coefficient of $\approx 4.8 \mu m^2.s^{-1}$ (Fig. 5e).

We also extracted the dissociation rate (k_{off}) (see methods) and deduced the estimated residence time (RT) ($RT=1/k_{off}$) (Fig. 5f). Our FCS results reveal that Zelda binding to chromatin is highly transient, on the order of hundreds of milliseconds (RT= 0.24 seconds). However, we note that this FCS-estimated RT likely represents only a lower-bound limit for Zelda binding, whereas a FRAP approach allows for upper-bound estimations (see next paragraph).

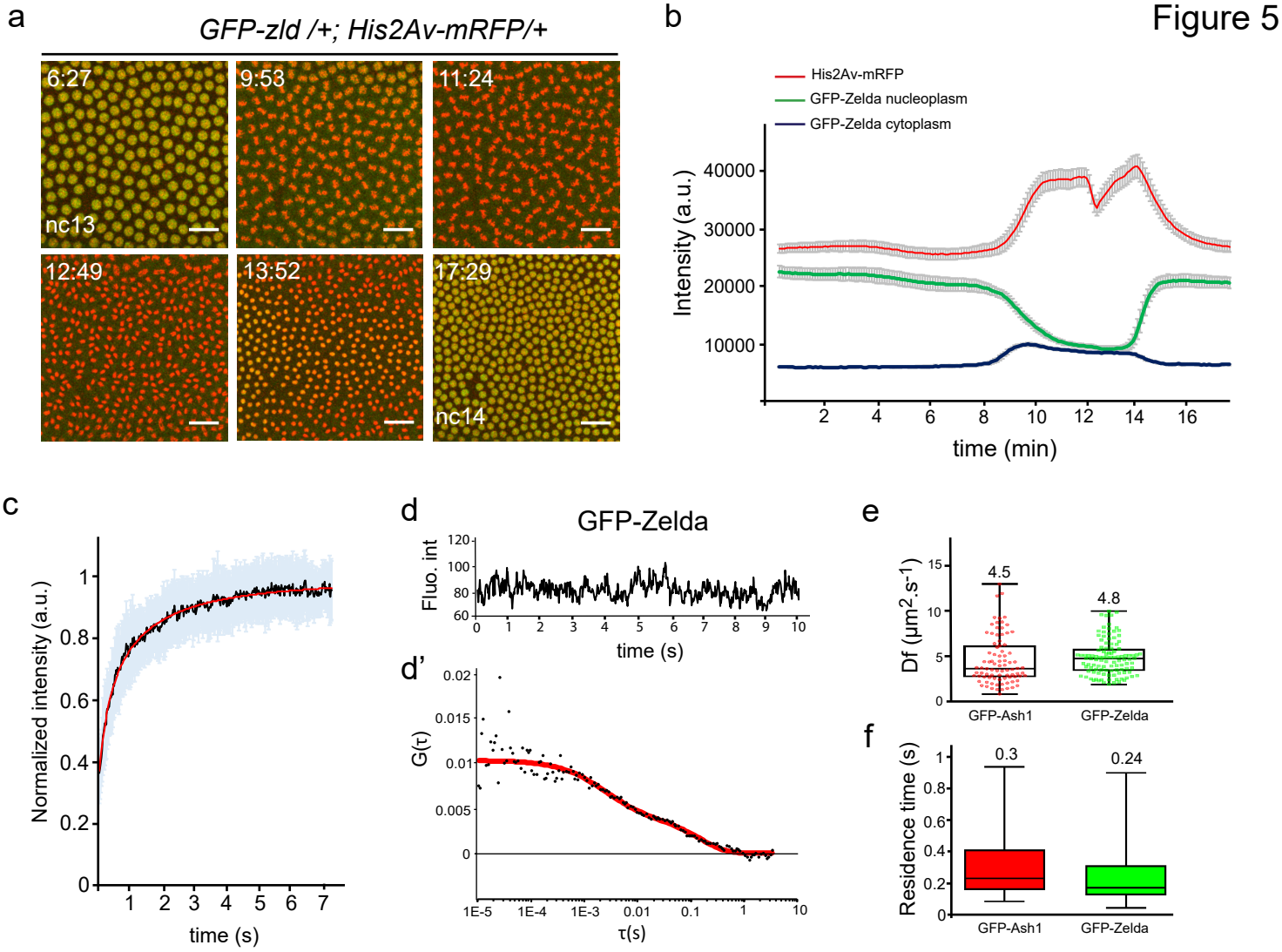
The low binding and fast diffusion rates for Zelda were unexpected given the large impact of a single extra Zld-binding site on accelerating the kinetics of gene activation (Fig. 1). However, similar transient chromatin bindings have been reported recently for other key transcription factors[8,30]. These transient bindings events seem to be compensated for by increased local concentrations, in particular nuclear microenvironments, referred to as subnuclear 'hubs'[8]. We thus explored Zelda spatial distribution within nuclei in living *Drosophila* embryos.

Zelda accumulates into subnuclear hubs

Live imaging with GFP-Zld, revealed a heterogeneous distribution of Zelda proteins in the nucleus (Fig. 6 a-c and Supplementary Movie 9-11). Thus, similarly to Bicoid[8], multiple Zelda proteins seem to cluster in local dynamic hubs that exhibit a diversity of sizes and lifetimes (Fig. 6a-b and Supplementary Movie 9- 11). For example, Zelda large hubs are more visible at early cycles and at the beginning of interphases (Fig. 6a and Supplementary Movie 9 and 11). These local inhomogeneities are unlikely GFP aggregate as we could observe them with a two-colored Zelda tagging approach (Fig. 6c). Hence Zelda hubs have been observed very recently with other methods[31].

These findings are consistent with a "phase separation" model of transcriptional regulation[32], which could promote Zelda cooperativity with other major

Figure 5


a
GFP-zld /+; His2Av-mRFP/+
6:27
9:53
11:24
nc13
12:49
13:52
17:29
nc14
b
His2Av-mRFP
GFP-Zelda nucleoplasm
GFP-Zelda cytoplasm
Intensity (a.u.)
time (min)
c
Normalized intensity (a.u.)
time (s)
d
GFP-Zelda
Fluo. int
time (s)
d'
G(τ)
τ(s)
e
Df (µm².s⁻¹)
4.5
4.8
GFP-Ash1
GFP-Zelda
f
Residence time (s)
0.3
0.24
GFP-Ash1
GFP-Zelda

transcriptional regulators, like Bicoid[8], which would then foster rapid and coordinated transcriptional initiation.

To decipher the link between Zelda hubs and transcription, we examined the spatial localization of nascent transcription with respect to Zelda hubs positions.

Immuno-FISH and live imaging experiments revealed that only a subset of our MS2 reporter nascent RNAs located within a Zelda dense region (Fig. 6d, Supplementary Fig. 6a, and Supplementary Movie 11). Furthermore, only limited co-localization was observed between dots of elongating Pol II (Ser2-P Pol II) and Zelda hubs (Supplementary Fig. 6b). Similar findings have been recently reported by following transcription from a *hunchback<MS2* transgene (*hb*)[31]. We therefore conclude that, at least with a synthetic transgene, long-lasting stable contacts between sites of transcription and Zelda dense regions are not detected. Given the important number of Zelda targets, rigorously connecting transcriptional activation to Zelda hubs would require broader analysis with adapted imaging methods. However, these observations are reminiscent of a dynamic kissing model recently described between an active gene and mediator clusters in ES cells[33].

Having demonstrated that Zelda was not homogeneously distributed, we wanted to determine if the dynamic properties of Zelda within hubs differed from its global dynamics.

One possibility was that high local concentrations of Zelda would correspond to long-lasting chromatin binding events that would result in increased residence time (long RT). Given the rapid movements of Zelda hubs (Supplementary Movie 9-12), capturing these long-lasting events by FCS was not feasible. Therefore, we performed FRAP experiments on long-lasting Zelda hubs (Fig. 6e and Supplementary Movie 12). We demonstrate that Zelda recovered as rapidly within long lasting hubs as it did in bulk nuclei (Fig. 6f and Fig. 5c). By fitting our FRAP data with a diffusion-reaction model, we estimate a residence time on the order of seconds in both cases (hubs and global FRAP) (Fig. 6g). Fitting the FRAP experiments with alternative models, such as one binding reaction only (Supplementary Fig. 6c-e) or two distinct binding reactions (slow and fast) (Supplementary Fig. 6f-h) similarly supported an estimated residence time for Zelda on the order of seconds and that this is independent of its local concentration (Supplementary Fig. 6i). Of note, the two reactions model estimates two residence times, where the fastest estimate is similar to the FCS-extracted residence time (Fig. 5f and (Supplementary Fig. 6i)). Previous studies on Bicoid have also identified two populations: one with a short characteristic time (RT on the order hundreds of milliseconds), described as non-specific binding events and a population with longer characteristic times (RT on the order of seconds) corresponding to specific binding events[8,34].

From the combined data, we conclude that Zelda chromatin binding is transient ranging from hundreds of milliseconds (likely unspecific binding) to seconds (likely specific binding) and this occurs regardless to its local concentration. How could Zelda exhibit a similarly transient chromatin binding at random positions and in regions where its concentration is high? An exciting possibility would be that dense regions could be enriched for Zelda target genes. Under this hypothesis, Zelda target genes could cluster together thereby increasing the probability to locally accumulate Zelda proteins. Alternatively, Zelda could spatially organize its targets in space through its pioneer factor properties and then maintain these clusters of targets possibly through its intrinsically disorders domains[35]. The latter scenario is consistent with the finding that Zelda participate to the *Drosophila* 3D chromatin organization during the awakening of the zygotic genome[36,37].

Using quantitative imaging approaches in living embryos combined with mathematical modeling, we have investigated a role for the pioneer-like factor Zelda in regulating mitotic memory. We demonstrated that Zelda is dispensable for transcriptional mitotic memory. While our modeling suggests that memory may be potentiated by a reduction in the number of steps required for post-mitotic transcriptional activation, we propose that Zelda accelerates this activation by decreasing the time spent at each preceding step. Our data support a model whereby Zelda binds transiently to chromatin in localized nuclear microenvironments to accelerate the various transitions required prior to transcriptional activation (*e.g* recruitment of transcription factors, recruitment of Pol II and general transcription factors). These dynamic properties allow the pioneer-like factor Zelda to act as a quantitative timer for fine-tuning transcriptional activation during early *Drosophila* development.

Acknowledgements

We are grateful to T.Fukaya for sharing the *MCP-eGFP-His2Av-mRFP* fly stock. We thank L.Ringrose for sharing *Ash1-GFP* fly stocks, C.Rushlow for *Zelda-RNAi* fly stocks and N.Dostatni for plasmids and eGFP-Bcd fly stocks. We thank G.Cavalli for providing the anti GAF antibody. We thank L.Ciandrini, E. Bertrand, R. Feil and T. Forne for helpful discussions and critical reading of the manuscript. We acknowledge the imaging facility MRI, member of the national infrastructure France-BioImaging supported by the French National Research Agency (ANR-10-INBS-04, «Investments for the future». A.T was a recipient of a FRM fellowship and is currently sponsored by the SyncDev ERC grant. Work in the lab of M.H. was supported by NIH R01GM111694. This work was supported by the ERC SyncDev starting grant to M.L and a HFSP-CDA grant to M.L.

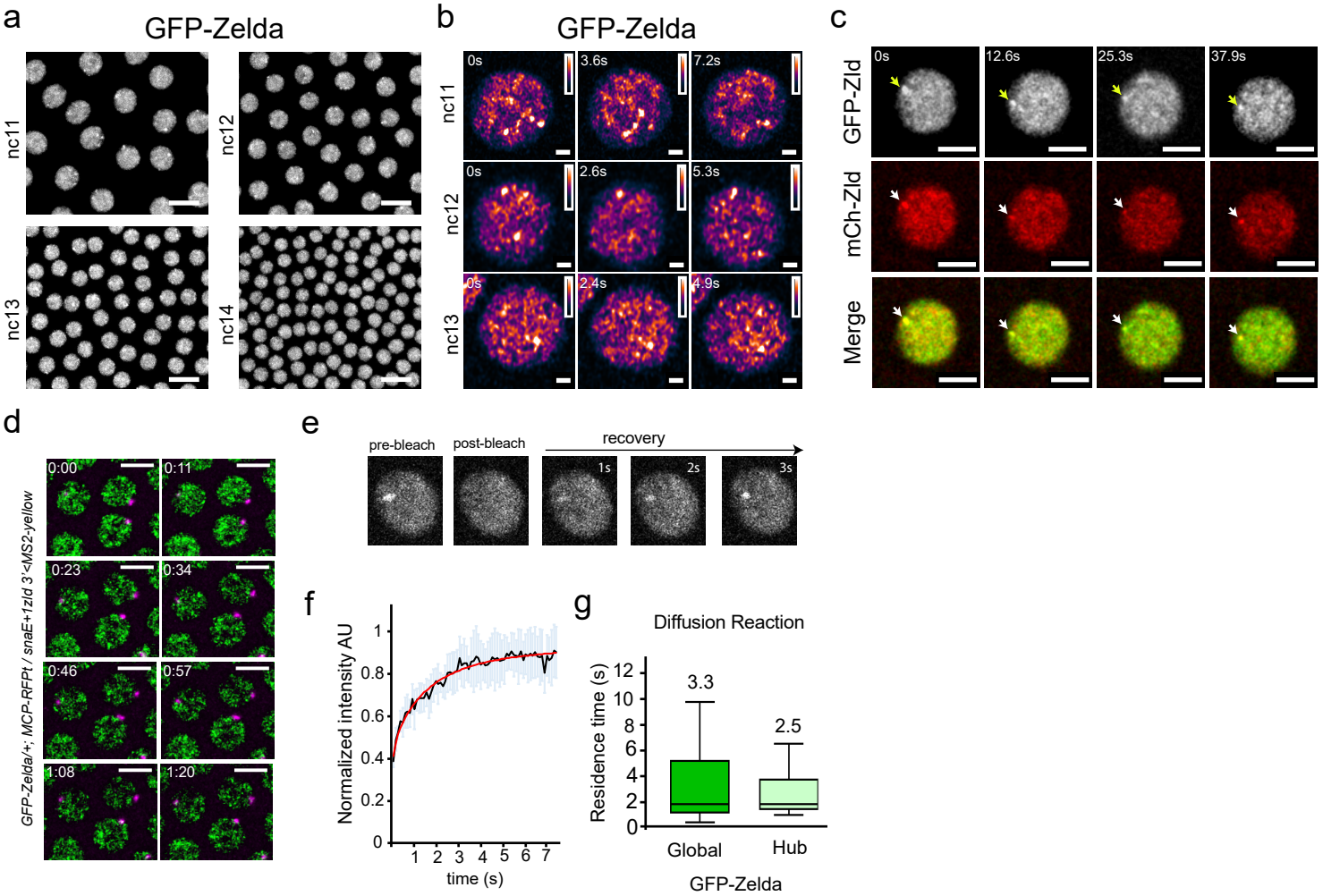
Author contributions

M.L conceived the project. M.L designed the experiment with the help from J.D. M.L, J.D, J.H, C.Fz, J.L, M.D, L.M, performed experiments. M.L, J.D and O.R analyzed the data and interpreted the results. A.T developed the software for image analysis. M.D and J.D performed fly handling. C.Fd advised J.D and A.T for FRAP and FCS experiments and performed the analysis. O.R performed mathematical modeling. M.H and K.S generated the *GFP-zld and mCherry-zld* lines and commented the data. M.L wrote the manuscript with help from J.D and O.R. All authors discussed, approved and reviewed the manuscript.

Competing interests

The authors declare no competing financial interests.

Figure 6


a
GFP-Zelda
nc11
nc12
nc13
nc14
b
GFP-Zelda
nc11
nc12
nc13
0s
3.6s
7.2s
0s
2.6s
5.3s
0s
2.4s
4.9s
c
GFP-Zld
mCh-Zld
Merge
0s
12.6s
25.3s
37.9s
d
GFP-Zelda/+; MCP-RFPt / snaE+1zld 3'<MS2-yellow
0:00
0:11
0:23
0:34
0:46
0:57
1:08
1:20
e
pre-bleach
post-bleach
recovery
1s
2s
3s
f
Normalized intensity AU
time (s)
g
Diffusion Reaction
Residence time (s)
3.3
2.5
Global
Hub
GFP-Zelda

Materials & Correspondence

Correspondence and requests for materials should be addressed to M.L (email: Mounia.Lagha@igmm.cnrs.fr). The data that support the findings of this study are available from the corresponding author upon request. Computer codes are available upon request.

Materials and Methods

Drosophila stocks and genetics

The *yw* stock was used as a control. The germline driver *nos-Gal4:VP16* (BL4937) was recombined with an *MCP-eGFP-His2Av-mRFP* fly line (gift from T.Fukaya). RNAi were expressed following crosses between this recombinant and *UASp-shRNA-w* (BL35573) or *UASp-shRNA-zld*[4], virgin females expressing RNAi, *MCP-GFP-His2Av-mRFP* were crossed with *MS2* containing transgenic constructs males. The *GFP-zld and mCherry-zld* strains were obtained by CRISPR[25]. *GFP-ash1*[28] and *His2Av-mRFP* (BL23651) and *eGFP-bcd* [27].

Cloning and Transgenesis

The *snaE* enhancer transgene was described in Ferraro et al[14]. The 24XMS2 tag was inserted immediately upstream of the yellow reporter gene coding sequence. We verified that our MS2 sequences do not contain any canonical binding site and we also compared synchrony of the *snaE* transgenes with our 24XMS2 to newly available MS2 loops, 24XMS2ΔZelda[38], kindly provided by Pr.Dostatni (Supplementary Fig. 7). Extra Zld-binding sites (consensus sequences CAGGTAG) were added to the *wt snaE* enhancer by PCR using the primers indicated in Table S3. Insertion of an extra Zld-binding site in the middle of the *snaE* enhancer was done by directed mutagenesis using primers indicated in Supplementary Table 3. All transgenic MS2 flies were inserted in the same landing site (BL9750 line) using PhiC31 targeted insertion[39].

Live imaging

Embryos were dechorionated with tape and mounted between a hydrophobic membrane and a coverslip as described in[14]. All movies (except when specified) were acquired using a Zeiss LSM780 confocal microscope with the following settings: GFP and RFP proteins were excited using a 488nm and a 561nm laser respectively. A GaAsP detector was used to detect the GFP fluorescence; a 40x oil objective and a 2.1 zoom on central ventral region of the embryo, 512 x 512 pixels, 16 bits/pixel, with monodirectional scanning and 21z stacks 0.5µm apart. Under these conditions, the time resolution is in the range of 22–25s per frame. Two and one movie respectively for *snaE + 3Zld* in *white*-RNAi and *zld*-RNAi background were acquired on a Zeiss LSM880, keeping the time resolution is in the range of 21–22s per frame. Image processing of LSM780 and LSM880 movies of the MS2-MCP-GFP signal were performed in a semi-automatic way using custom made software, developed in Python™, that will be described in another publication (*Trullo et al., in preparation*). Live imaging of *GFP-zld/+; His2Av-mRFP/+* and *GFP-bcd/+; His2Av-mRFP/+* embryos were performed using a Zeiss LSM780 confocal microscope with the following settings: GFP and RFP proteins were excited using a 488nm and a 561nm laser respectively. A GaAsP detector was used to detect the GFP fluorescence. 512 x 512 pixels, 16 bits/pixel, images were acquired using a 40x oil objective, with bi-directional scanning and 8 stacks 1µm apart. Under these conditions, the time resolution is in the range of 5–7s per frame. Live imaging of *GFP-zld* embryos (related to Fig. 6b) was acquired using a Zeiss LSM880 confocal microscope with the following settings: GFP protein was excited using a 488nm laser. A GaAsP-PMT array of an Airyscan detector was used to detect the GFP fluorescence. 568 x 568 pixels, 8 bits/pixel images were acquired using a 40x oil objective, 13-14z stacks 0.5µm apart, with a 4x zoom and bi-directional scanning, with fast Airyscan in optimal mode. Under these conditions and using a piezo Z, the time resolution is in the range of 2-4s per frame.

Live imaging of *GFP-zld/mCherry-zld* embryos (related to Fig. 6c) was acquired using a Zeiss LSM880 confocal microscope with the following settings: GFP and mCherry proteins were excited using a 488nm and a 561nm laser respectively. A GaAsP detector was used to detect the GFP fluorescence and a GaAsP-PMT array of an Airyscan detector was used to detect mCherry. 512 x 512 pixels, 16 bits/pixel images were acquired using a 40x oil objective, 5z stacks 1µm apart, with a 3.6x zoom and bidi-directional scanning. Under these conditions, the time resolution is in the range of 12-13s per frame. Live imaging of *GFP-zld, MCP-RFPt* (stock details available upon request) embryos containing *snaE+1Zld3'* (related to Supplementary Movie11) was acquired using a Zeiss LSM880 confocal microscope with the following settings: GFP and MCP-RFPt proteins were excited using a 488nm and a 561nm laser respectively. A GaAsP detector was used to detect the GFP fluorescence. 512 x 512 pixels, 16 bits/pixel, images were acquired using a 40x oil objective, 12z stacks 0.5µm apart, with a 4x zoom. Under these conditions, the time resolution is in the range of 11-12s per frame. Live imaging of *GFP-zld* embryos (related to Supplementary Movie10) was acquired using a Zeiss LSM880 using a 40X/1.3 Oil objective. Images (256_256 pixels), 12bits/pixel, 5z stacks 1µm apart, zoom 8x, were acquired every ≈550ms during 100 frames. GFP was excited with an Argon laser at 488nm and detected between 496–565nm.

Image analysis for transcriptional activation through multiple cell cycles

The image analysis software and associated graphical user interface developed for this study will be detailed in another methodological manuscript (Trullo et al, in preparation). Briefly, red (nuclei) and green (MS2 nascent RNA spots) channels were first maximum intensity projected. The analysis was divided into three parts: before mitosis (nc13), during mitosis and after mitosis (nc14). During interphases, nuclei were segmented (using mainly circularity arguments and water-shed algorithm) and tracked with a minimal distance criterion. Chromosomal splitting at the end of anaphase was automatically detected and used to define a nucleus-by-nucleus mitosis frame. During mitosis, nuclei were segmented by intensities and tracked by an overlap in consecutive frames criterion. By merging these three parts, we could give a label to each nc13 mother nucleus, their daughters in nc14 and an extra label to recognize one daughter from the other. MS2 spots were detected with a blob detection method with a user-defined threshold constant value. Detected spots were associated to the closest nuclei, inheriting their label. Finally, for each tracked nucleus, the timing of first transcriptional activation was recorded.

Analysis were done with the activation timing of the two daughters for synchrony and the first active daughter for memory.

Manual tracking

Upon high reduction of maternal Zelda, zygotic activation is perturbed, and major developmental defects occur. Abnormal nuclear shape (example Supplementary Movie 6 and control Supplementary Movie 5) precludes their automatic segmentation. Kinetics of transcriptional activation was manually analyzed for *zld*-RNAi embryos. We implemented a spot detection algorithm to export files with the detected MS2 spot. The thresholding of the MS2 spot is consistent with all the other automatic analysis. A spatial domain was determined, taking the 'pseudo-furrow' as a landmark and focusing on a 25µm region around it (note that activation seems not to occur outside of this area). Nuclei in this region are visually tracked frame-by-frame to detect when activation in nc13 occurs, when mitosis occurs and to recover the first activation frame in nc14. Approximately 2 to 3-fold more nuclei from inactive compared to active mother nuclei in nc13 were analyzed.

Immunostaining and RNA *in situ* hybridization

Embryos were dechorionated with bleach for 1-2 min and thoroughly rinsed with H2O. They were fixed in fixation buffer (500µl EGTA 0.5M, 500µl PBS 10X, 4ml Formaldehyde MeOH free 10%, 5ml Heptane) for 25 min on a shaker at 450 rpm; formaldehyde was replaced by 5 ml methanol and embryos were vortexed for 30s. Embryos that sank to the bottom of the tube were rinsed three times with methanol. For immunostaining, embryos were rinsed with methanol and washed three times with PBT (PBS 1X 0.1% triton). Embryos were incubated on a wheel at room temperature twice for 30 min in PBT, once for 20 min in PBT 1% BSA, and at 4 °C overnight in PBT 1% BSA with primary antibodies. Embryos were rinsed three times, washed twice for 30 min in PBT, then incubated in PBT 1% BSA for 30 min, and in PBT 1% BSA with secondary antibodies for 2 h at room temperature. Embryos were rinsed three times then washed three times in PBT for 10 min. DNA staining was performed using DAPI at 0.5 $\mu g.ml^{-1}$. Primary antibody dilutions for immunostaining were mouse anti-GFP (Roche IgG1κ clones 7.1 and 13.1) 1:200; rabbit anti-GFP (Life technologies A11122); mouse anti-Ser5-P Pol II (Covance H14, MMS-134R); rat anti-Ser2-P Pol II (3E10); rabbit anti-GAF (a gift from Dr.G.Cavalli) 1:250; Secondary antibodies (anti-rabbit Alexa 488-conjugated (Life technologies, A21206); anti-mouse Alexa 488-conjugated (Life technologies, A21202); anti-mouse IgM Alexa 555-conjugated (Life technologies, A21426); anti-rabbit Alexa 555-conjugated (Life technologies, A31572); were used at a dilution 1:500. Fluorescent *in situ* hybridization was performed as described in[14]. A dixogygenin-MS2 probe was obtained by *in vitro* transcription from a bluescript plasmid containing the 24-MS2 sequences, isolated with BamH1/BglII enzymes from the original Addgene MS2 plasmid (# 31865). *Snail* probe was described in[19]. Mounting was performed in Prolong Gold. Fluorescent *in situ* hybridization with *yellow* probes was performed as described in[40], probe sequences will be available upon request.

Fluorescence recovery after photobleaching

Fluorescence recovery after photobleaching (FRAP) in embryos at nc14 (Global) was performed with the following settings: a Zeiss LSM780 using a 40X/1.3 Oil objective and a pinhole of 66 µm. Images (512x32 pixels), 16bits/pixel, zoomed 6x, were acquired every ≈20ms during 400 frames. GFP was excited with an Argon laser at 488nm and detected between 507–596nm. Laser intensity was kept as low as possible to minimize unintentional photobleaching. A circular ROI (16x16 pixels) 0.07µm/pixels, was bleached using two laser pulses at maximal power during a total of ≈130ms after 50 frames. Hub FRAP was performed in embryos from nc10 to nc13 with the following settings: FRAP was performed on a Zeiss LSM880 using a 40X/1.3 Oil objective and a pinhole of 83 µm. Images (256x256 pixels), 12bits/pixel, zoom 8x, were acquired every ≈104ms during 100 frames. GFP was excited with an Argon laser at 488nm and detected between 496–565nm. Laser intensity was kept as low as possible to minimize unintentional photobleaching. A circular ROI (8x8 pixels) 0.1µm/pixels, was bleached using four laser pulses at maximal power during a total of ≈114ms after 3 frames. To discard any source of fluorescence intensity fluctuation other than molecular diffusion, the measured fluorescence recovery in the bleached ROI region (I_{bl}) was corrected by an unbleached ROI (I_{unbl}) of the same nucleus and another ROI outside of the nucleus (I_{out}) following the simple equation:

$$I_{bl_{corr}}(t) = \frac{I_{bl}(t) - I_{out}(t)}{I_{unbl}(t) - I_{out}(t)} \quad (1)$$

The obtained fluorescence recovery was then normalized to the mean value of fluorescence before the bleaching i.e.:

$$I_{bl_{norm}}(t) = \frac{I_{bl_{corr}}(t)}{\frac{1}{N}\sum_{n=1}^{50} I_{bl}(n)} \quad (2)$$

Fitting FRAP recovery curves

Three different analytical equations were used to fit the fluorescence recovery depending on the model of dynamics chosen.

For the pure reaction kinetics model, we started from the analytical expression of Spague et al.[41] and decline it for the case of two exchanging populations, leading to the two following equations:

one exchanging population

$$F(t) = 1 - C_{eq}e^{-k_{off}t} \quad (3)$$

two exchanging populations

$$F(t) = 1 - C_{eq_1}e^{-k_{off_1}t} - C_{eq_2}e^{-k_{off_2}t} \quad (4)$$

With $C_{eq} = k^*_{on} / (k_{off} + k^*_{on})$ as in their case.

Finally, for the reaction-diffusion model, we started from the analytical expression developed in the supplemental (eq.(S35)) of [29].

$$F(t) = F_{eq}F_D(t) + C_{eq}F_{exc}(t) \quad (5)$$

with C_{eq} defined as above and $F_{eq} = k_{off} / (k_{off} + k^*_{on})$. $F_D(t)$ is the fluorescence recovery due to diffusion and $F_{exc}(t)$ the fluorescence recovery due to exchange.
Since we used a Gaussian shape illumination profile, $F_D(t)$ is defined using a slightly modified version of the analytical equation of the 20th order limited development of the Axelrod model for Gaussian profile illumination and diffusion[42,43]:

$$F_D(t) = \frac{1 - e^{-K}}{K}(1 - M) + M\sum_{n=1}^{20}\frac{(-K)^n}{n!}\left(1 + n + 2n\frac{t}{\tau}\right)^{-1} \quad (6)$$

$$M = \frac{I_{(t>30\tau)} - I_0}{1 - I_0} \quad (7)$$

where K is a constant proportional to bleaching deepness, M is the mobile fraction and τ is the half time of recovery. To minimize the effect of mobile fraction on Ceq, M was kept between 0.9 and 1.1.

Diffusion coefficients of the different molecules were determined according to

$$D = \frac{\beta w^2}{4\tau} \quad (8)$$

with w the value of the radius at $1/e^2$ of the Gaussian beam (in our case, w=0.56µm or 0.4 µm) and β a discrete function of K tabulated in[44].

$F_{exc}(t)$ is defined as in [29], slightly modified with respect to the Gaussian illumination, leading to the following equation:

$$F_{exc}(t) = F_\infty - \left(\frac{1-e^{-K}}{K} - F_\infty\right)e^{-k_{off}t} \quad (9)$$

with K defined as previously.

Fluorescence correlation spectroscopy
FCS experiments were performed on a Zeiss LSM780 microscope using a 40X/1.2 water objective. GFP was excited using the 488nm line of an Argon laser with a pinhole of 1 airy unit. Intensity fluctuation measured for 10s were acquired and auto-correlation functions (ACFs) generated by Zen software were loaded in the PyCorrFit program[45]. Multiple measurements per nucleus, in multiple nuclei and embryos at 20°C were used to generate multiple ACF, used to extract kinetic parameters. The FCS measurement volume was calibrated with a Rhodamine6G solution[46] using D_f=414 µm²/s. Each time serie was fitted with the reaction dominant model[29]:

$$G_D(\tau) = \frac{1}{2^{3/2}N}F_{eq}\left(1 + \frac{\tau}{\tau_{D_f}}\right)^{-1}\left(1 + \frac{\tau}{\omega^2\tau_{D_f}}\right)^{-1/2} + \frac{1}{2^{3/2}N}C_{eq}e^{-k_{off}\tau} + G_\infty \quad (10)$$

where $F_{eq} = k_{off} / (k_{off} + k^*_{on})$, $C_{eq} = k^*_{on} / (k_{off} + k^*_{on})$, $\tau_{Df} = w^2_{xy} / 4\mu D_f$ and $\omega = w_z / w_{xy}$.

For each correlogram the value of k^*_{on}, k_{off} and D_f (reaction constants and diffusion coefficient) were estimated using the PyCorrFit program implemented with eq.(10).

Photophysics of the dye was neglected and we introduced a G∞ value to account for long time persistent correlation during the measurements. Based on the average number of Zelda molecules (N), connected to the G(0) value following the formula $G(0)=\frac{1}{2^{3/2}N}$ and on the illumination volume, we found that Zelda concentration is around $410^{+/-90}$ nM.

Mathematical model, general framework
We are interested in the time needed for post-mitotic transcription (re)activation. We model this time as the sum of two variables:

$$T_a = T_0 + T_r \quad (11)$$

where T_0 is a deterministic incompressible lag time, the same for all nuclei, and T_r is a random variable whose value fluctuates from one nucleus to another. The decomposition in Eq. (11) can be justified by the experimental observation that all the reactivation curves (Fig. 1l) start with a nonzero length interval during which no nuclei are activated. Furthermore, T_r is defined such that it takes values close to zero with nonzero probability. This property allows us to set the time origin to the instant when the first nuclei initiates transcription, in order to determine T_r. For large populations of nuclei, this stems to setting the time origin to T_0 (see Supplementary Methods).

The stochastic part of the (re)activation time is modeled using a finite state, continuous time, Markov chain.

The states of the process are $A_1, A_2, \ldots, A_{n-1}, A_n$. The states A_i, $1 \leq i \leq n-1$ are metastable and OFF, i.e. not transcribing. The state A_n is ON, i.e. transcribing. Each metastable state has a given lifetime, defined as the waiting time before leaving the state and going elsewhere. For the purposes of this paper, we considered that each of the states has the same lifetime denoted τ. Also, the topology of the transitions is considered linear: in order to go to A_{i+1} one has to visit A_i.

We have also tested a more complex model, with uneven lifetimes and therefore, more parameters. However, the complex model did not provide a sensibly better fit with data. Furthermore, the determination of individual parameters of the more complex model is uncertain: many different combinations of parameter values lead to very close fit. Following Occam's razor principle, we based our analysis on the simplest, homogeneous jump model. For consistency, detailed models are described in Supplementary Methods.

Figure 1: Zelda-binding sites provide quantitative temporal control of gene expression.

(**a**) Schematic representation of the truncated version of the *sna* shadow enhancer (*snaE*) and the variants carrying different numbers and locations of extra Zelda (Zld)-binding sites (black stars represent canonical Zld-binding sites: CAGGTAG). Each enhancer controls a common minimal promoter (blue box, *sna* promoter) that transcribes a *yellow* gene with 24 *MS2* repeats located at its 5' end. (**b-d'**) Transgenic embryos stained with a *MS2* probe (red, panels **b- d'**) and with a *snail* probe (green, **b-d**). Dorsal views at nc14, anterior to the left, scale bars represent 30µm. (**e-g'**) Workflow for automatic tracking of single nucleus transcriptional activity. (**e, e'**) Snapshot of a maximum intensity projected Z-stack from live imaging movie representing nuclei labeled with *His2Av-mRFP* transgene (**e**) and associated automatic segmentation, where each tracked nuclei is given a random color (**e'**). (**f**) Maximum intensity projected Z-stack image extracted from a typical movie showing the transcriptional foci (white dot). (**f'**) False-colored image, representing the segmented transcriptionally inactive nuclei (blue), active nuclei (red) and associated transcriptional foci (yellow dot). (**g**) Maximum intensity projected Z-stack image representing a typical gastrulation, used as a landmark to determine dorso-ventral (D/V) coordinates (yellow line). (**g'**) Representative image exhibiting the spatial domain (grey, here 25µm surrounding the ventral furrow) defined by precise D/V coordinates.

(**h-k**) Live imaging of transcriptional activity driven by the *snaE* upon increasing numbers of Zld-binding sites. Nuclei that displayed transcriptional activity were false-colored in yellow. Snapshots from maximum intensity projected Z-stack of live imaging movies are shown from nc11 to nc14. (**l**) Temporal coordination profiles during nc14. Synchrony curves were obtained by quantifying transcriptional activation during nc14 in a region of 50µm around the ventral furrow: *snaE* (6 movies, n=970 nuclei) (grey curve), *snaE+1Zld5'* (4 movies, n=700 nuclei) (light blue curve), *snaE+2Zld* (5 movies, n=723 nuclei) (dark blue curve), *snaE+3Zld* (4 movies, n=824 nuclei (purple curve)) and *snaE-shadow* (4 movies, n=612) (yellow curve). Dashed line represents 50% of activation. The time origin is the end of mitosis (start of nc14).

Error bars represent SEM.

Figure 2: Zelda accelerates gene activation regardless of transcriptional history

(**a**) Representative images from the Graphical User Interface implemented for automatic tracking of nuclei across multiple divisions. Chromosomal splitting at the end of anaphase is automatically detected for each nucleus. (**b-d**) Quantification of transcriptional memory in various *snaE<MS2* embryos. At each time point in nc14, the cumulative number of activated nuclei, expressed as a percentage of the total number of nuclei, is plotted after segregation into two populations: descendants of active mothers (green) and descendants of inactive mothers (red). The timing of activation of each nucleus has been normalized to its specific timing of mitosis. Temporal dynamics from *snaE* are depicted using solid curves, while those from transgenes with extra Zld-binding sites are represented as dashed curves (**c-d**).

Figure 3: Transcriptional memory persists in the absence of Zelda

(**a-d'**) Snapshots of maximum intensity projected Z-stack of *Drosophila* embryo movies expressing *snaE* (**a-b'**) or *snaE+3Zld* (**c-d'**) at nc13 (**a-d**) and nc14 (**a'-d'**), in a *white*-RNAi (**a, a', c, c'**) or in a *zld*-RNAi genetic background (**b, b', d, d'**). Nuclei exhibiting a spot of active transcription are false-colored in yellow.

(**e, f**) Temporal coordination of all activated nuclei in nc14 for *snaE* (**e**) and *snaE+3Zld* transgene (**f**), in a *white*-RNAi (solid curves) and in a *zld*-RNAi genetic background (dashed curves). Statistics: *snaE white*-RNAi: 9 movies, n=972; *snaE zld*-RNAi: 4 movies n=156; *snaE+3Zld white*-RNAi 6 movies n=541; *snaE+3Zld zld*-RNAi 5 movies n=160). Error bars represent SEM. (**g, h**) Kinetics of activation during the first 15min of nc14 driven by the *snaE* transgene or the *snaE+3Zld* transgene in *white*-RNAi (solid curves) or in *zld*-RNAi embryos (dashed curves). The kinetics of nuclei coming from transcriptionally active mothers in nc13 (green curves) is compared to

those arising from inactive mothers (red curves). In Figure 3, only nuclei located within a 25µm rectangle centered around the ventral furrow are considered.

Figure 4: Mathematical modeling of transcriptional memory
(**a**) Schematic representation of the discrete states (OFF1, OFF2 and OFF3) required to reach transcription (ON state), when coming from an inactive (red stairs) or an active (green stairs) mother nucleus. After mitosis, the probability for a nucleus to initiate the transcriptional activation process from a given state is provided by the parameter p, (p1=p(OFF1), p2=p(OFF2) and p3=p(OFF3). The distribution of the random time to transcription (t) is modeled as a mixture of gamma distributions, according to the displayed Eq. (1), where Cdf is the cumulative distribution function, Γ and γ are the complete and incomplete gamma functions, respectively. The average number of transitions *'a'* is provided by the sum of weighted probabilities (Eq2). The time of each jump from one state to another is provided by the parameter *'b'*. Parameters of this model are predicted separately for nuclei coming from active and inactive mothers. The parameter fitting is based on Cdf of the random time to transcription in nc14 (t) (see also Eq.S1 Supplementary Methods). (**b**) Cdf of the random time to transcription in nc14. The origin of time is the end of mitosis nc13 to nc14, (determined by our automatic software, see methods) added to the time required to detect the first activation (proper to each genotype). Data (cross) are compared to predictions (solid curves) for nuclei, resulting from active (green) and inactive (red) mothers. (**c, c'**) Mean '*a'* parameter; error intervals correspond to the variation among optimal and best suboptimal fits (see Supplementary Methods and Supplementary Table 1). (**d, d'**) Mean '*b'* parameter; error intervals correspond to variation among optimal and best suboptimal fits (see Supplementary Methods and Supplementary Table 1). (**e**) The ratio of parameter '*a'* in subpopulations from inactive and active mothers is positively correlated (Spearman correlation coefficient =0.82 p=0.034) to the average '*b'* parameter. Predictions of the mathematical model are represented as a black line.

Figure 5: Zelda global kinetic properties
(**a**) Living *GFP-zld/+;His2Av-mRFP/+* embryo imaged by confocal microscopy from interphase of nc13 to early interphase of nc14. Successive representative maximum intensity projected Z-stack images are shown at the indicated timings (in minutes) (see also Supplementary Movie 7). Scale bars represent 20µm.
(**b**) Average intensity profiles for histones (red), nucleoplasmic Zelda (green) and cytoplasmic Zelda (blue) measured from a nc13 embryo transitioning into nc14. An automatic tracking of fluorescence from a minimum of 114 nuclei and 10 cytoplasmic ROI generated these profiles. Error bars represent SD.
(**c**) FRAP mean curve (black) and the mean of all the fits (red curve) using reaction-diffusion models determined at the bleached spot for 25 nuclei from nc14 developing *GFP-zld* embryos. Error bars represent SD from different nuclei (light blue bars). (**d**) Example of a time trace obtained by FCS from a *GFP-zld* nc14 embryo showing no bleaching. (**d'**) Example of autocorrelation function (black dots) related to (**d**) (red curves represent fitting using reaction-diffusion model).
(**e-f**) Kinetic parameters for Zelda and Ash1 extracted after fitting FCS data with a reaction-diffusion model. (**e**) Box plot representing estimated diffusion coefficients (D_f). Numbers above each plot represent the mean. Centered lines represent the median. (f) Box plot representing estimated residence times ($1/k_{off}$). Numbers above each plot represent the mean. Centered lines represent the median.

Figure 6: Zelda hubs kinetic properties
(**a**) Living *GFP-zld* embryo imaged by confocal microscopy from interphase of nc11 to early interphase of nc14. Successive representative maximum intensity projected Z-stack images showing local inhomogeneities are shown at each cycle (see also Supplementary Movie 9). Scale bars represent 20µm.
(**b**) Nuclei zoom from living *GFP-zld* embryos imaged by confocal microscopy from interphase of nc11 to early interphase of nc13. Successive representative images showing local hubs are shown at each cycle. Scale bars represent 1µm.
(**c**) Nuclei zoom from living *GFP-zld/mCherry-Zld* embryos imaged by confocal microscopy at interphase of early nc11. Successive representative images showing Zelda subnuclear hubs are shown at each cycle (an example is indicted by an arrow). Scale bars represent 5µm.
(**d**) Living *GFP-Zld/+;MCP-RFPt/+* embryo containing the *snaE+1Zld3'* transgene (MS2 in magenta) imaged by confocal microscopy. Successive representative maximum intensity projected Z-stacks of slices containing the MS2 signal only (see also Supplementary Movie 11). Scale bars represent 5µm.
(**e**) Example of a FRAP experiment on a single hub. (**f**) FRAP mean curve (black) and the mean of all the fits (red curve) using diffusion reaction models determined at the bleached spot for 11 nuclei from nc10-13 developing *GFP-zld* embryos. Error bars represent SD from different nuclei (light blue bars).
(**g**) Box plot representing estimated residence times ($1/k_{off}$) extracted with diffusion reaction model of GFP-Zld in Global (dark green) compared to Hubs (light green). Numbers above each plot represent the mean RT. Centered lines represent the median.

References

1. Liang, H.-L. *et al.* The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. *Nature* **456,** 400–403 (2008).
2. Nien, C.-Y. *et al.* Temporal coordination of gene networks by Zelda in the early Drosophila embryo. *PLoS Genet* **7,** e1002339 (2011).
3. Harrison, M. M., Li, X.-Y., Kaplan, T., Botchan, M. R. & Eisen, M. B. Zelda Binding in the Early Drosophila melanogaster Embryo Marks Regions Subsequently Activated at the Maternal-to-Zygotic Transition. *PLoS Genet* **7,** e1002266 (2011).
4. Sun, Y. *et al.* Zelda overcomes the high intrinsic nucleosome barrier at enhancers during Drosophilazygotic genome activation. *Genome Research* **25,** 1703–1714 (2015).
5. Schulz, K. N. *et al.* Zelda is differentially required for chromatin accessibility, transcription factor binding, and gene expression in the early Drosophila embryo. *Genome Research* **25,** 1715–1726 (2015).
6. Xu, Z. *et al.* Impacts of the ubiquitous factor Zelda on Bicoid-dependent DNA binding and transcription in Drosophila. *Genes & Development* **28,** 608–621 (2014).
7. Foo, S. M. *et al.* Zelda potentiates morphogen activity by increasing chromatin accessibility. **24,** 1341–1346 (2014).
8. Mir, M. *et al.* Dense Bicoid hubs accentuate binding along the morphogen gradient. *Genes & Development* **31,** 1784–1794 (2017).
9. Crocker, J., Tsai, A. & Stern, D. L. A Fully Synthetic Transcriptional Platform for a Multicellular Eukaryote. *CellReports* **18,** 287–296 (2017).
10. Caravaca, J. M. *et al.* Bookmarking by specific and nonspecific binding of FoxA1 pioneer factor to mitotic chromosomes. *Genes & Development* **27,** 251–260 (2013).
11. Bellec, M., Radulescu, O. & Lagha, M. Remembering the past: Mitotic bookmarking in a developing embryo. *Current Opinion in Systems Biology* **11,** 41–49 (2018).
12. Perry, M. W., Boettiger, A. N., Bothma, J. P. & Levine, M. Shadow Enhancers Foster Robustness of Drosophila Gastrulation. *Current Biology* **20,** 1562–1567 (2010).
13. Dunipace, L., Ozdemir, A. & Stathopoulos, A. Complex interactions between cis-regulatory modules in native conformation are critical for Drosophila snail expression. *Development* **138,** 4566–4566 (2011).
14. Ferraro, T. *et al.* Transcriptional Memory in the Drosophila Embryo. *Current Biology* **26,** 212–218 (2016).
15. Pichon, X., Lagha, M., Mueller, F. & Bertrand, E. A Growing Toolbox to Image Gene Expression in Single Cells: Sensitive Approaches for Demanding Challenges. *Molecular Cell* **71,** 468–480 (2018).
16. Bosch, J. R. T. The TAGteam DNA motif controls the timing of Drosophila pre-blastoderm transcription. *Development* **133,** 1967–1977 (2006).
17. Sandler, J. E. & Stathopoulos, A. Stepwise Progression of Embryonic Patterning. *Trends Genet.* **32,** 432–443 (2016).
18. Ip, Y. T., Park, R. E., Kosman, D., Yazdanbakhsh, K. & Levine, M. dorsal-twist interactions establish snail expression in the presumptive mesoderm of the Drosophila embryo. *Genes & Development* **6,** 1518–1530 (1992).
19. Lagha, M. *et al.* Paused Pol II Coordinates Tissue Morphogenesis in the Drosophila Embryo. *Cell* **153,** 976–987 (2013).
20. Rushlow, C. A. & Shvartsman, S. Y. Temporal dynamics, spatial range, and transcriptional interpretation of the Dorsal morphogen gradient. *Current Opinion in Genetics & Development* **22,** 542–546 (2012).
21. Kanodia, J. S. *et al.* Pattern Formation by Graded and Uniform Signals in the Early Drosophila Embryo. *Biophysj* **102,** 427–433 (2012).
22. Hannon, C. E., Blythe, S. A. & Wieschaus, E. F. Concentration dependent chromatin states induced by the bicoid morphogen gradient. *eLife* **6,** 3165 (2017).
23. Ni, J.-Q. *et al.* A genome-scale shRNA resource for transgenic RNAi in Drosophila. *Nature Methods* **8,** 405–407 (2011).
24. Lionnet, T. & Singer, R. H. Transcription goes digital. *EMBO Rep.* **13,** 313–321 (2012).
25. Hamm, D. C. *et al.* A conserved maternal-specific repressive domain in Zelda revealed by Cas9-mediated mutagenesis in Drosophila melanogaster. *PLoS Genet* **13,** e1007120 (2017).
26. Staudt, N., Fellert, S., Chung, H.-R., Jäckle, H. & Vorbrüggen, G. Mutations of the Drosophila zinc finger-encoding gene vielfältig impair mitotic cell divisions and cause improper chromosome segregation. *Molecular Biology of the Cell* **17,** 2356–2365 (2006).
27. Gregor, T., Wieschaus, E. F., McGregor, A. P., Bialek, W. & Tank, D. W. Stability and nuclear dynamics of the bicoid morphogen gradient. *Cell* **130,** 141–152 (2007).
28. Steffen, P. A. *et al.* Quantitative in vivo analysis of chromatin binding of Polycomb and Trithorax group proteins reveals retention of ASH1 on mitotic chromatin. *Nucleic Acids Research* **41,** 5235–5250 (2013).
29. Michelman-Ribeiro, A. *et al.* Direct measurement of association and dissociation rates of DNA binding in live cells by fluorescence correlation spectroscopy. *Biophys. J.* **97,** 337–346 (2009).
30. Tsai, A. *et al.* Nuclear microenvironments modulate transcription from low-affinity enhancers. *eLife* **6,** e28975–18 (2017).
31. Mir, M. *et al.* Dynamic multifactor hubs interact transiently with sites of active transcription in Drosophila embryos. (2018). doi:10.1101/377812
32. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A Phase Separation Model for Transcriptional Control. *Cell* **169,** 13–23 (2017).
33. Cho, W.-K. *et al.* Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361,** 412–415 (2018).
34. Abu-Arish, A., Porcher, A., Czerwonka, A., Dostatni, N. & Fradin, C. High mobility of bicoid captured by fluorescence correlation spectroscopy: implication for the rapid establishment of its gradient. *Biophys. J.* **99,** L33–5 (2010).
35. Hamm, D. C., Bondra, E. R. & Harrison, M. M. Transcriptional Activation is a Conserved Feature of the Early Embryonic Factor Zelda that Requires a Cluster of Four Zinc Fingers for DNA Binding and a Low-Complexity Activation Domain. *Journal of Biological Chemistry* (2014). doi:10.1074/jbc.M114.602292
36. Hug, C. B., Grimaldi, A. G., Kruse, K. & Vaquerizas, J. M. Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell* **169,** 1–33 (2017).
37. Ogiyama, Y., Schuettengruber, B., Papadopoulos, G. L., Chang, J.-M. & Cavalli, G. Polycomb-Dependent Chromatin Looping Contributes to Gene Silencing during Drosophila Development. *Molecular Cell* **71,** 1–22 (2018).
38. Lucas, T. *et al.* 3 minutes to precisely measure morphogen concentration. (2018). doi:10.1101/305516
39. Venken, K. J. T., He, Y., Hoskins, R. A. & Bellen, H. J. P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. *Science* **314,** 1747–1751 (2006).
40. Tsanov, N. *et al.* smiFISH and FISH-quant – a flexible single RNA detection approach with super-resolution capability. *Nucleic Acids Research* **44,** e165–e165 (2016).
41. Sprague, B. L., Pego, R. L., Stavreva, D. A. & McNally, J. G. Analysis of Binding Reactions by Fluorescence Recovery after Photobleaching. *Biophys. J.* **86,** 3473–3495 (2004).
42. Escoffre, J. M., Hubert, M., Teissié, J., Rols, M. P. & Favard, C. Evidence for electro-induced membrane defects assessed by lateral mobility measurement of a GPi anchored protein. *Eur. Biophys. J.* **43,** 277–286 (2014).
43. Axelrod, D., Koppel, D. E., Schlessinger, J., Elson, E. & Webb, W. W. Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophysj* **16,** 1055–1069 (1976).
44. Yguerabide, J., Schmidt, J. A. & Yguerabide, E. E. Lateral mobility in membranes as detected by fluorescence recovery after photobleaching. *Biophysj* **40,** 69–75 (1982).
45. Müller, P., Schwille, P. & Weidemann, T. PyCorrFit-generic data evaluation for fluorescence correlation spectroscopy. *Bioinformatics* **30,** 2532–2533 (2014).
46. Dertinger, T. *et al.* The optics and performance of dual-focus fluorescence correlation spectroscopy. *Opt Express* **16,** 14353–14368 (2008).