Temporal Control of Transcription by Zelda in living *Drosophila* embryos

Jeremy Dufourt[1,5], Antonio Trullo[1,5], Jennifer Hunter[1], Carola Fernandez[1], Jorge Lazaro[1], Matthieu Dejean[1], Lucas Morales[1], Katharine N Schulz[2], Cyril Favard[3], Melissa M Harrison[2], 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/intro

Pioneer factors have the exquisite ability to engage their target sites at nucleosomal DNA, which leads to a local remodeling of chromatin and the establishment of a transcriptional competence. However, the direct impact of enhancer priming by pioneer factors on the temporal control of gene expression and on mitotic memory remains elusive. In *Drosophila* embryos, the maternally deposited activator Zelda (Zld) exhibits key pioneer factor properties and indeed regulates the awakening of the zygotic genome. The analysis of thousands of endogenous Zld bound regions in various genetic contexts, as well as the study of isolated synthetic enhancers with static approaches, led to the proposal that Zld could act as a quantitative developmental timer. Here we employ quantitative live imaging methods and mathematical modeling to directly test the effect of Zld on temporal coordination in gene activation and on mitotic memory. Using an automatic tracking software, we quantified the timing of activation in hundreds of nuclei and their progeny in *Drosophila* embryos. We demonstrate that increasing the number of Zld binding sites accelerates the kinetics of transcriptional activation regardless of their past transcriptional state. In spite of its known pioneering activities, we show that Zld is not a mitotic bookmarker and is neither necessary nor sufficient to foster mitotic memory. Fluorescent recovery after photo-bleaching and fluorescent correlation spectroscopy experiments reveal that, Zld is highly dynamic and exhibits transient binding to chromatin. We propose that Zld low binding rates could be compensated for by local accumulation of Zld in nuclear microenvironments *in vivo*, thus allowing rapid and coordinated gene activation.

1

Results and Discussion

Zelda fosters temporal transcriptional coordination

In the blastoderm embryo, two redundant enhancers control *snail* (*sna*) gene expression, a proximal (primary) and a distal (shadow enhancer)[1,2]. Both enhancers are bound by Zelda (Zld) at early nuclear cycles[3,4]. Expression driven by the long intact shadow enhancer (*sna-shadow*) leads to rapid transcriptional activation, precluding analysis of the impact of Zld on fine-tuning the timing of this activation (Figure S1, A-B). We therefore used a previously described truncated version of the *sna* shadow enhancer (*snaE*) that leads to a stochastic activation, compatible with mitotic memory tracking[5].

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 (Figure 1A). Upon transcription, MS2 loops are rapidly bound by the maternally provided MCP-GFP, which allows tracking of transcriptional activation in living embryos[6,7]. To decipher the role of Zld in *cis*, a varied number of canonical Zld-binding sites (CAGGTAG)[8,9] were added to the *snaE* in two locations, either close to the promoter (Zld 3') or at a distance (Zld 5') (Figure 1 A).

The pattern of expression of *snaE-extraZld* transgenes was dictated by the regulatory logic of the *sna* enhancer and was located in the presumptive mesoderm, similarly to the endogenous *sna* pattern (Figure 1 B-D'). Using fluorescent *in situ* hybridization, we could detect that the number of active nuclei within the *sna* pattern correlated with the number of Zld-binding sites present in the enhancer. However, this technique was performed on embryos fixed at various stages of nuclear cycle 14 (nc14), which lasts ≈55min. Thus, the impact of Zld on the probability to activate reporter expression could be due to small variations in developmental timing of these embryos. 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 (Figure 1 E-E') and tracks the time of transcriptional activation (Figure 1 F-F'). In the early fly embryo, transcriptional activators, like Dorsal, control gene expression in a graded manner[10]. It was therefore important to quantify temporal dynamics of gene activation in a spatially controlled pattern. For this purpose, we used the ventral furrow as a landmark to define dorso-ventral coordinates (Figure 1-G). Unless otherwise indicated, we studied temporal dynamics of gene activation in a region of 50μm around the ventral furrow to ensure non-limiting levels of the *sna* activators Dorsal and Twist[11] (Figure 1 G').

We initially examined the impact of Zld on the timing of gene activation. As identified with static approaches[8,12],

our dynamic data showed that increasing the number of Zld-binding sites resulted in precocious transcriptional activation (Figure 1 H-L) (Movie 1). Zld binding impacted not only the onset of gene activation but also the temporal coordination among a spatially defined pattern, referred to as synchrony[13] (Figure 1 M). For example, at precisely 10min into interphase 14, expression of the *snaE* transgene was not coordinated, whereas the number of active nuclei at this stage increased by adding Zld-binding sites (Figure 1 H-L). 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≈14min. Adding only a single CAGGTAG sequence, 5' to the enhancer accelerated this kinetic (t50≈10min). With two extra Zld-binding sites synchrony is further increased with t50≈8min. The activation kinetics of two extra canonical Zld sites mimicked the dynamics generated by the intact long shadow enhancer (Figure S1 B). The most dramatic synchrony shift was caused by the addition of a single CAGGTAG sequence close to the promoter (+1 Zld 3') (t50≈5.5min) (Figure 1M). This behavior was not changed when adding two additional Zld sites, suggesting that promoter proximal Zld binding causes a strong effect, that cannot be enhanced by the addition of more distant sites. 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 tracking the timing of activation along the dorso-ventral axis, we can detect the Dorsal nuclear spatial gradient[14] with temporal data (Figure S1 C-D). 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[12,15]. Adding extra Zld-binding sites boosted the spatio-temporal response to the dorso-ventral gradient (Figure S1C-D). These temporal data are consistent with the idea that Zld potentiates the binding of spatially restricted morphogens like Bicoid or Dorsal[4,12,16-18], and thus supports the previously described pioneering activities of Zld[19,20].

Altogether we have set up a live imaging-based approach to quantitatively measure the impact of Zld on transcriptional dynamics in a developing dividing embryo.

Accelerated transcription mediated by increased Zld-binding sites bypasses transcriptional memory

Our ability to track transcriptional activation in live embryos provided the unprecedented opportunity to determine activity through multiple cell 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

2

influences the timing of activation of the descendants in the following cycle[5]. In this previous study, memory was manually tracked, limiting the number of data points that could be processed. To address the challenge of following rapid movements during mitosis, a semi-automatic mitotic nuclei tracking software was developed (Figure 2A). 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 (green curves, Figure 2) from those coming from an inactive mother (red curves, Figure 2). Differences between the kinetics of activation of these two populations provide an estimation of transcriptional mitotic memory.

Given the association of Zld with regions of open chromatin[3] and its ability to shape the chromatin landscape in the early embryo[19,21], we reasoned that Zld could be 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, adding extra Zld-binding sites in *cis* augmented the speed of activation in both nuclei derived from active mothers and those from inactive mothers (Figure 2 C-D, Figure S1 E-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.

Modeling the impact of Zelda on memory

In an attempt to gain some insights into the role of Zld on mitotic memory, we developed a simple mathematical framework that described the delays of transcription (waiting times prior to the first detected initiation event) as a sequence of transitions (see Methods). The significant number of nuclei tracked in this study (>300 per genotype) provided the statistical power to examine their distribution. The distribution of waiting times for descendants of active nuclei and inactive nuclei were clearly distinct (Figure 2E). These distributions can be fitted by a mixture of three gamma distributions with scale parameters (number of transitions) 1, 2 and 3, suggesting that post-mitotic nuclei require at most three transitions to be come active. The first two moments of the waiting time distribution allow the calculation of two parameters; *a* and *b*, where *a* represents the number of transitions between discrete transcriptional states and *b* the duration of each transition (Table S1).

Fitting our data revealed that for *snaE*, the temporal behaviors of descendants of active and that of inactive

mothers differ mainly in parameter a, namely the number of transitions required to reach the transcription active state (Figure 2F). Corresponding parameter *b* estimates range between 2.4 to 3.8 min. This estimated memory time is sufficiently large to guarantee transmission of transcriptional state accross mitosis for both types of mothers (Table S1). The accelerated transcriptional activation observed with increased numbers of Zld-binding sites could similarly be due to either a decrease in the duration of the transitions (*b*) or the number of transitions (*a*). Our modeling suggests that the addition of extra Zld-binding sites accelerates the transitions between various states; hence *b* estimates are consistently smaller (Table S1). In summary, using a modeling approach, continuous waiting time data were converted into discrete metastable states that precede gene activation. This model suggests that transcriptional memory is supported by a sequence of relatively long-life metastable states preceding activation, which could be maintained by mitotic bookmarking.

Zelda is dispensable for transcriptional mitotic memory

To test whether Zld was necessary for transcriptional mitotic memory, activation kinetics for descendants from active and inactive nuclei were quantified by reducing maternal Zld expression using RNAi (Movie 2 and Movie 3). Validating this approach, germline expression of *zld*-RNAi[12,22] reduced the levels of *zld* mRNA to 2-4% of that found in control *white*-RNAi embryos (Figure S3 A).

Reduction of maternal Zld decreased the synchrony of activation of our transgenes with additional Zld-binding sites (Figure 3 A-F, Figure S3 B-D). Of note, temporal behaviors of the *white*-RNAi controls mimicked those observed in a wild-type genetic background (Figure1M), and thus provided an independent measure of synchrony. In the *zld*-RNAi, temporal coordination was reduced and all transgenes showed a t50≈12-15 min (Figure 3A-F, Figure S3 B-C). This confirms that the previously identified acceleration of transcriptional activation upon addition of extra Zld-binding sites is due to Zld activity on these synthetic enhancers.

We then quantified the level of memory in the *zld*-RNAi embryos with three *snaE<MS2* transgenes containing varying number of Zld-binding sites. To our surprise, in spite of the maternal reduction of Zld, a strong memory bias is still observed (Figure 3G-H, Figure S3D). We conclude although possessing several key properties of a pioneer-factor[19,20], Zld is not necessary to elicit heritability of active transcriptional states through mitoses. Of note, in Zld maternally depleted embryos, transcriptional activation is restricted to reduced region of about ≈25μm around the pseudo-furrow (embryos do not properly gastrulate, data not shown). These data confirm that outside of this domain, Zld potentiates Dorsal-dependent activation of *snaE* transgenes.

Upon maternal decrease of Zld, activation kinetics were lowered yet transcriptional mitotic memory was still present. This emphasizes the concept that synchrony is distinct from that of memory. Altogether our results show that Zld accelerates transcription and fosters synchrony but is not the basis for transcriptional mitotic memory.

Zelda is not sufficient to trigger transcriptional mitotic memory

To examine whether Zld binding to an enhancer was sufficient to trigger a memory bias, temporal activation from a second regulatory element (*DocE*) was examined (Figure S2). This region is located in a gene desert within the *Doc1/2* locus and is characterized by high Zld occupancy in early embryos (Figure S2 A-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* (Figure S2 C, C') and *Doc2* mRNA expression (data not shown). Upon *zld*-RNAi, expression from this enhancer was highly reduced thus confirming a role for Zld in driving expression from the *DocE* element (data not shown).

During nc13, our *DocE* transgene was stochastically activated, which is compatible with transcriptional mitotic memory tracking in three distinct dorsal ectoderm domains (Figure S2C). 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 (Figure S2 D-G). We therefore conclude that this *DocE* element does not trigger memory, despite being bound by Zld. Thus, experiencing transcription at a given cycle does not necessarily lead to a rapid *post*-mitotic reactivation in the following cycle.

Quantitative kinetic analysis reveals the dynamic properties of Zelda

Zld shares several key properties of a pioneer factor such as priming *cis*-regulatory elements prior to gene activation, which establishes competence for the adoption of many cell fates, and opening chromatin, which facilitates subsequent binding by classical transcription factor[19]. In addition to these properties, classical pioneer factors bind nucleosomes and are typically retained on the chromosomes during mitosis[23]. Based on these known characteristics of pioneer-like factors we had expected a role for Zld in retaining transcriptional memory through mitosis. Thus, our genetic data and modeling indicating that Zld was not the basis of memory was unexpected. To better understand this surprising result, we examined Zld dynamics in living embryos and particularly during mitosis (Figure 4). We took advantage of an endogenously tagged fluorescent version of Zld[24]. The *GFP-zld* flies are homozygous viable, thereby showing that the GFP-tagged version of Zld retains wild-type physiological properties.

First, GFP-Zld binding to DNA was examined through the mitotic cycle (Figure 4, Figure S4). Live imaging data showed that Zld is not obviously retained on the chromosomes during mitosis (Figure 4A-B), similar to what was shown in fixed embryos[25] (Figure S4A). While most GFP-Zld molecules are allocated to nuclei in interphase, they were evicted from nuclei and redistributed to the cytoplasm during mitosis (Figure 4B). Our live imaging approach revealed a highly dynamic behavior of Zld, whereby Zld comes back to the nucleus very rapidly at the end of mitosis (Figure 4A and Movie 4). Intra-nuclear FRAP experiments indeed confirm the significant mobility of Zld, with an apparent diffusion coefficient of ≈0.5µm²/s (Figure S4B). To obtain a precise measure of Zld diffusion properties including binding kinetics, we performed Fluorescent Correlation Spectroscopy (FCS) in living cycle14 embryos (Figure 4C). Auto-correlation curves for Zld were compared to a known mitotic bookmarking factor, Ash1, whose kinetics have been documented by FRAP and FCS approaches in early *Drosophila* embryos[26] and thus represent a point of reference (Figure S4C).

To estimate chromatin binding kinetics, FCS curves were fitted with a reaction-diffusion model that was specifically developed for this approach[27]. This model showed similar diffusion coefficient of ≈5µm²/s for both Zld and Ash1(Figure 4 D). Contrary to the mitotic bookmarking factor Ash1, the estimated pseudo first order association rate of Zld (k^*_{on}) is significantly lower than its dissociation rate (k_{off}) (Figure 4E, Figure S5D). We therefore conclude that Zld binding to chromatin is relatively transient, on the order of seconds (residence time of 0.35 seconds, Figure 4F). Moreover, Zld binds to chromatin with a significantly lower fraction than Ash1 (Figure 4G). These properties were unexpected given the strong impact of extra Zld-binding sites on accelerating the kinetics of gene activation (Figure 1). However, similarly transient bindings to chromatin have recently been reported for other key transcription factors[18,28] (e.g Bicoid). These transient bindings seem to be compensated for by increases in local concentrations, in particular nuclear microenvironments[18,28]. Live imaging with GFP-Zld, as well as immunohistochemistry on fixed embryos, revealed a heterogeneous distribution of Zld proteins in the nucleus (Figure 4H-J and Movie 5). Thus, similarly to Bicoid[18], multiple Zld proteins seem to aggregate in nuclear microenvironments *in vivo* (Figure 4). These findings are consistent with a phase separation model of transcriptional regulation[29], which could promote Zld cooperativity with other major transcriptional regulators to foster rapid and coordinated transcriptional initiation.

Using quantitative imaging approaches in living embryos combined with mathematical modeling, we have investigated a role for the pioneer-like factor Zld in regulating mitotic memory. Contrary to our expectations, we demonstrated that Zld is dispensable for

4

transcriptional mitotic memory. While memory may be potentiated by a reduction in the number of steps required for post-mitotic transcriptional activation, Zld accelerates activation by decreasing the time spent at each preceding step. Our data support a model whereby Zld binds transiently but frequently to chromatin, possibly 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 Zld to act as a quantitative timer for fine-tuning transcriptional activation during early *Drosophila* development. Transcriptional activation can be accelerated by other means, as exemplified by transcriptional mitotic memory. In this case, we propose that mitotic bookmarking factors prevent the decline to profound states during mitosis.

Materials and Methods

Drosophila stocks and genetics

The *yw* stock was used as a control. The germline driver *nos-Gal4:VP16* (Rorth, 1998) 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*[21], virgin females expressing RNAi, *MCP-GFP-His2Av-mRFP* were crossed with *MS2* containing transgenic constructs. The *GFP-zld* strain was obtained by CRISPR [24]. *GFP-ash1* [26] and *His2Av-mRFP* (BL23651).

Cloning and Transgenesis

The *snaE* enhancer transgene was described in Ferraro et al [5]. The 24XMS2 tag was inserted immediately upstream of the yellow reporter gene coding sequence. Extra Zld-binding sites (consensus sequences CAGGTAG) were added to the *wt snaE* enhancer by PCR using the primers indicated in Table S2. All transgenic flies inserted at the same locus using PhiC31 P[acman] strain, site specific insertion into the BL9750 line (VK33 site).

Live imaging

Embryos were dechorionated with tape and mounted between a hydrophobic membrane and a coverslip as described in[5]. 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, 16 bit/pixel image, with monodirectional scanning and 21z stacks 0.5µm apart. Under these conditions, the time

resolution is in the range of 22–25s per frame. 2 and 3 movies 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 Python™ made software, that will be described in another publication (*Trullo et al., in preparation*). Live imaging of *GFP-zld/+;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, a 40x oil objective and 512 x 512, 16bit/pixel image, with bi-directional scanning and 8 stacks 1 µm apart. Under these conditions, the time resolution is in the range of 5–7s per frame. Live imaging of *GFP-zld* embryos was acquired using a Zeiss LSM880 confocal microscope with the following settings: GFP protein was excited using a 488nm. A GaAsP-PMT array of an Airyscan detector was used to detect the GFP fluorescence, a 40x oil objective and 568 x 568, 8 bit/pixel image, with a 4x zoom and mono-directional scanning, with Super Resolution mode. Under these conditions and using a piezo Z, the time resolution is in the range of 3–4s per frame.

Manual tracking

Upon high reduction of maternal Zld, zygotic activation is perturbed and major developmental defects occur[9,25]. Abnormal nuclear shape (example movie 3 and control movie 2) 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 no activation seems to occurs 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 µg mL−1. Primary antibody dilutions for immunostaining were mouse anti-GFP (Roche IgG1κ clones 7.1 and 13.1) 1:200; rabbit anti-Zld antibody (gift from Dr.M.Harrison) 1:250; 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-rabbit Alexa555-conjugated (Life technologies, A31572); were used at a dilution 1:500. Fluorescent in situ hybridization was performed as described in[5]. 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[13]. Mounting was performed in Prolong Gold.

Fluorescence recovery after photobleaching
Fluorescence recovery after photobleaching (FRAP) in embryos was performed with the following settings: FRAP was performed on a Zeiss LSM780 using a 40X/1.3 Oil objective and a pinhole of 66 µm. Images (512_32 pixels), 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 130ms after 50 frames. In order 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 adjacent 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.:

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

and fitted using a slightly modified 19th order limited development of the Axelrod model for Gaussian profile illumination and diffusion[30,31], using the following equations:

$$I_{bl}(t) = \frac{1 - e^{-K}}{K}(1 - M) \quad (3)$$
$$+ M \sum_{n=1}^{19} \frac{(-K)^n}{n!}\left(1 + n + 2n\frac{t}{\tau}\right)^{-1}$$

$$M = \frac{I_{(t>30\tau)} - I_0}{1 - I_0} \quad (4)$$

with K, a constant proportional to bleaching deepness, M, the mobile fraction and τ, the half time of recovery. Diffusion coefficients of the different molecules were determined according to

$$D = \frac{\beta w^2}{4\tau} \quad (5)$$

With w the value of the radius at $1/e^2$ of the Gaussian beam (in our case, w=0.56µm) and β a discrete function of K tabulated in[32].

Fluorescence correlation spectroscopy
FCS experiments were performed on a Zeiss LSM780 microscope using a 40X/1.4 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 a Python™ shell. 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[33] using Df=414 µm²/s. Each time series was fitted with the reaction dominant model[27]:

$$G_D(\tau) = \frac{1}{2^{3/2}N}F_{eq}\left(1 + \frac{\tau}{\tau_{D_f}}\right)^{-1} \quad (6)$$
$$(1 + \frac{\tau}{\omega^2\tau_{D_f}})^{-1/2} + \frac{1}{2^{3/2}N}C_{eq}e^{-k_{off}\tau}$$

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 $\omega = w_z / w_{xy}$.
The fitting was performed using the Levenberg-Marquardt minimization algorithm. Since the photophysics of GFP (triplet state, dark states) and the existence of possible long time correlation (tau>1s) are not taken into account

6

in eq. (6), curves were fitted between ≈25µs and ≈1s. For each time series the value of k^*_{on}, k_{off} and D_f (reaction constants and diffusion coefficient) were estimated and averaged at the end over all the estimations.

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 split 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. 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.

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:

$$T_a = T_0 + T_r \qquad (7)$$

where T_0 is a deterministic incompressible lag, the same for all nuclei, and T_r is a random variable whose value fluctuates from one nucleus to another. Furthermore, T_r is defined such that it takes values close to zero with nonzero probability.

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, $1 \le i \le n-1$ 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 have 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. More complex models will be described, analyzed and identified elsewhere.

The distribution of the random time T_r depends on τ and on the number of transitions performed in order to reach

the ON state. If only one transition is utilized to reach the ON state, then T_r is exponentially distributed with the parameter τ^{-1}. Likewise, in the case where the ON state is reached after performing k transitions, then T_r is gamma distributed with shape parameter k and with scale parameter τ. In general, T_r is given by a mixture of gamma distributions with shape parameters 1,2,3,... and scale parameter τ, whose cumulative distribution function (cdf) reads:

$$
\begin{aligned}
F_r(X) = P[T_r \leqslant t] &= p_1\left(1 - \exp(-t/\tau)\right) \\
&+ p_2 \frac{1}{\Gamma(2)} \gamma(2, t/\tau) \\
&+ p_3 \frac{1}{\Gamma(3)} \gamma(3, t/\tau) + \cdots
\end{aligned}
\qquad (8)
$$

where Γ, γ are the complete and incomplete gamma functions, respectively. The mean and the variance of T_r are as such:

$$
\begin{aligned}
E[T_r] &= (p_1 + 2p_2 + 3p_3 + \ldots)\tau \\
Var[T_r] &= (p_1 + 2p_2 + 3p_3 + \ldots)\tau^2
\end{aligned}
\qquad (9)
$$

We define the following parameters of the mixed distribution that can be computed empirically from the mean and the variance, the first two moments of the distribution, as such:

$$b = Var[T_r]/E[T_r], \qquad a = (E[T_r])^2/Var[T_r] \qquad (10)$$

Using (9) we find that:

$$a = p\_1 + 2p\_2 + 3p\_3 + \cdots \qquad (11)$$

showing that the parameter a represents the average number of transitions.

Data analysis and parameter estimates for modeling

Data analysis and parameter estimates were performed using MATLAB and Optimization Toolbox Release 2013b, The MathWorks, Inc., Natick, Massachusetts, United States.

The time origin was first set at the end of mitosis. The deterministic waiting time T_0 was estimated as the time between the end of mitosis and the time when the first nucleus from a large population is activated. This estimate is accurate for a large number of nuclei, because the probability p_0 that T_r is close to zero was

supposed to be non-zero. The estimate is more accurate when the number of transitions is small, because p_0 is high for a gamma distribution with small shape parameter, as it is the case when the number of transitions is small. Then the origin of time was set at T_0 and T_r was determined for all nuclei. Parameters a, b were estimated with the formulas (10). We found that a is not higher than 3, which allowed us to restrict the analysis to only three unidirectional transitions.

The empirical cumulative distribution function of T_r was estimated using the Kaplan-Meier method. The mixed model (8) was fitted by minimization of an objective function O defined as the l^2 (sum of squares) distance between experimental data and model prediction. The optimization was performed using the MATLAB function lsqnonlin starting from 100, different, randomly chosen initial parameter values. The best optimum was kept as parameter estimate. The results of the fit are the three parameters p_1, p_2, p_3 and b (Table S1). Parameter a is also computed using (11). The values of a and b estimated from the moments and from the distribution fit can be different. The first method, based on the moments, is sensitive to rare events; therefore it is less reliable than the latter.

The result of the parameter estimation is given in the Table S1. The goodness of fit is illustrated in Figure 2 E.

References

1. Perry, M. W., Boettiger, A. N., Bothma, J. P. & Levine, M. Shadow Enhancers Foster Robustness of Drosophila Gastrulation. *Current Biology* **20,** 1562–1567 (2010).

2. 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).

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 Genetics* **7,** e1002266 (2011).

4. Nien, C.-Y. *et al.* Temporal Coordination of Gene Networks by Zelda in the Early Drosophila Embryo. *PLoS Genetics* **7,** e1002339–16 (2011).

5. Ferraro, T. *et al.* Transcriptional Memory in the Drosophila Embryo. *Current biology : CB* **26,** 212–218 (2016).

6. Garcia, H. G., Tikhonov, M., Lin, A. & Gregor, T. Quantitative Imaging of Transcription in Living Drosophila Embryos Links Polymerase Activity to Patterning. *Current Biology* **23,** 2140–2145 (2013).

7. Lucas, T. *et al.* Live Imaging of Bicoid-Dependent Transcription in Drosophila Embryos. *Current Biology* **23,** 2135–2139 (2013).

8. Bosch, J. R. T. The TAGteam DNA motif controls the timing of Drosophila pre-blastoderm transcription. *Development* **133,** 1967–1977 (2006).

9. 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).

10. Sandler, J. E. & Stathopoulos, A. Stepwise Progression of Embryonic Patterning. *Trends in Genetics* **32,** 432–443 (2016).

11. 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).

12. Foo, S. M. *et al.* Zelda potentiates morphogen activity by increasing chromatin accessibility. *Current biology : CB* **24,** 1341–1346 (2014).

13. Lagha, M. *et al.* Paused Pol II coordinates tissue morphogenesis in the Drosophila embryo. *Cell* **153,** 976–987 (2013).

14. 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).

15. Kanodia, J. S. *et al.* Pattern Formation by Graded and Uniform Signals in the Early Drosophila Embryo. *Biophysj* **102,** 427–433 (2012).

16. 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).

17. Hannon, C. E., Blythe, S. A. & Wieschaus, E. F. Concentration dependent chromatin states induced by the bicoid morphogen gradient. *eLife* **6,** 3165 (2017).

18. Mir, M. *et al.* Dense Bicoid hubs accentuate binding along the morphogen gradient. *Genes & Development* **31,** 1784–1794 (2017).

19. 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).

20. Iwafuchi-Doi, M. & Zaret, K. S. Cell fate control by pioneer transcription factors. *Development* **143,** 1833–1837 (2016).

21. Sun, Y. *et al.* Zelda overcomes the high intrinsic

22. Ni, J.-Q. *et al.* A genome-scale shRNA resource for transgenic RNAi in Drosophila. *Nature Methods* **8,** 405–407 (2011).

23. Caravaca, J. M. *et al.* Bookmarking by specific and nonspecific binding of FoxA1 pioneer factor to mitotic chromosomes. *Genes & Development* **27,** 251–260 (2013).

24. Hamm, E. D.,*et al..* A conserved maternal-specific repressive domain in Zelda revealed by Cas9-mediated mutagenesis in Drosophila melanogaster. 1–22 (2017). doi:10.1371/journal.pgen.1007120

25. 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. *Mol. Biol. Cell* **17,** 2356–2365 (2006).

26. 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).

27. Michelman-Ribeiro, A. *et al.* Direct Measurement of Association and Dissociation Rates of DNA Binding in Live Cells by Fluorescence Correlation Spectroscopy. 1–10 (2009).

nucleosome barrier at enhancers during Drosophilazygotic genome activation. *Genome Research* **25,** 1703–1714 (2015).

doi:10.1016/j.bpj.2009.04.027

28. Tsai, A. *et al.* Nuclear microenvironments modulate transcription from low-affinity enhancers. *eLife* **6,** e1006441 (2017).

29. 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).

30. 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).

31. 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).

32. Yguerabide, J., Schmidt, J. A. & Yguerabide, E. E. Lateral mobility in membranes as detected by fluorescence recovery after photobleaching. *Biophysj* **40,** 69–75 (1982).

33. Dertinger, T. *et al.* The optics and performance of dual-focus fluorescence correlation spectroscopy. *Opt Express* **16,** 14353–14368 (2008).

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 J.H 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 design and analysis. O.R performed mathematical modeling. M.H and K.S generated the *GFP-zld* line and commented the data. M.L wrote the manuscript with the help from J.D. All authors discussed, approved and reviewed the manuscript.

Competing interests
The authors declare no competing financial interests.

Materials & Correspondence
Correspondence and requests for materials should be addressed to M.L (email: Mounia.Lagha@igmm.cnrs.fr)

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 of extra Zelda (Zld)-binding sites (1 to 3) (black stars). The snaE contains a non-canonical Zld site (grey circle). Each enhancer controls a common minimal promoter (light 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 B'-D') and with a snail probe (green, B-D). Regardless of the number of Zld-binding sites at the sna enhancer, MS2 reporter expression is restricted to the snail endogenous domain. Dorsal views at nc14, anterior to the left, scale bars represent 30µm. (E-G') Workflow of single nucleus transcriptional activity automatic tracking. (E, E') Snapshot of a max-projected live imaging movie representing nuclei labeled with His2Av-mRFP transgene (E) and associated automatic segmentation (E'). (F) Max-projected image extracted from a typical movie showing the transcriptional foci (white dot). (F') False colored image representing the segmented transcriptionally inactive nuclei (blue) and active nuclei (red) and associated spots (yellow dot). (G) Max projected 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-L) Live imaging of transcriptional activity driven by the snaE upon increasing numbers of Zld-binding sites. Nuclei displaying transcriptional activity have been false-colored in yellow. Snapshots from live imaging movies are shown from nc11 to nc14.

(M) 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=499 nuclei), snaE+1Zld5' (4 movies, n=350 nuclei), snaE+2Zld (5 movies, n=365 nuclei), snaE+3Zld (4 movies, n=411 nuclei), snaE+1Zld3' (4 movies, n=349). Dashed line represent 50% of activation. 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. (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). 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). (E) Cumulative distribution functions. Data (cross) are compared to predictions (solid curves) for nuclei, resulting from active (green) and inactive (red) mothers. (F) Schematic representation of the discrete states required to reach transcription, when coming from an active or an inactive mother (see Methods).

Figure 3: Transcriptional memory persists in the absence of Zelda

(A-D') Snapshots of Drosophila embryos 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 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: 6 movies, n=302; snaE zld-RNAi: 4 movies n=86; snaE+3Zld white-RNAi 6 movies n=273; snaE+3Zld zld-RNAi 6 movies n=111). 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 gastrulation are considered.

Figure 4: Zelda kinetic properties

(A) Living GFP-zld/+;His2Av-mRFP/+ embryo imaged by confocal microscopy from interphase of nc13 to early interphase of nc14. Successive representative images are shown at the indicated timings (see also Movie 4). Scale bars represent 20µm.

(B) Average intensity profiles for histones (red), nucleoplasmic Zld (green) and cytoplasmic Zld (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) Average autocorrelation function (black dots) obtained from 94 FCS experiments performed on 12 nuclei from 2 GFP-zld nc14 embryos (red curves represent average ±SD) and the fitting range (blue dashed lines) (see methods).

(D-G) Kinetic parameters for Zld and Ash1, extracted after fitting FCS data with a reaction-dominant diffusion model. GFP-Ash1 autocorrelation curves are used as a control (see methods and figure S4 C-D). (D) Estimated diffusion coefficients (Df). (E) Estimated pseudo first-order association constant k*on and koff. Error bars represent SEM. (F) Estimated residence times (1/koff). (G) Estimated bound fractions (k*on/(k*on+koff). Error bars represent k*on and koff errors propagation .

(H) Time-lapse microscopy of a single GFP-zld nc14 stage nucleus revealing the existence of local accumulation of Zld (Movie 5). (I) Fluorescence tracking of the circled 'cluster' in panel (I) showing a stable (purple trace) and transient (red trace) behaviors. (J) Immunostaining using anti-GFP in GFP-zld embryo (left panel) and anti-Zld (right panel) in yw embryos at nc14 showing local concentration of Zld protein in the nucleus (yellow arrow). Scale bars represent 10µm.
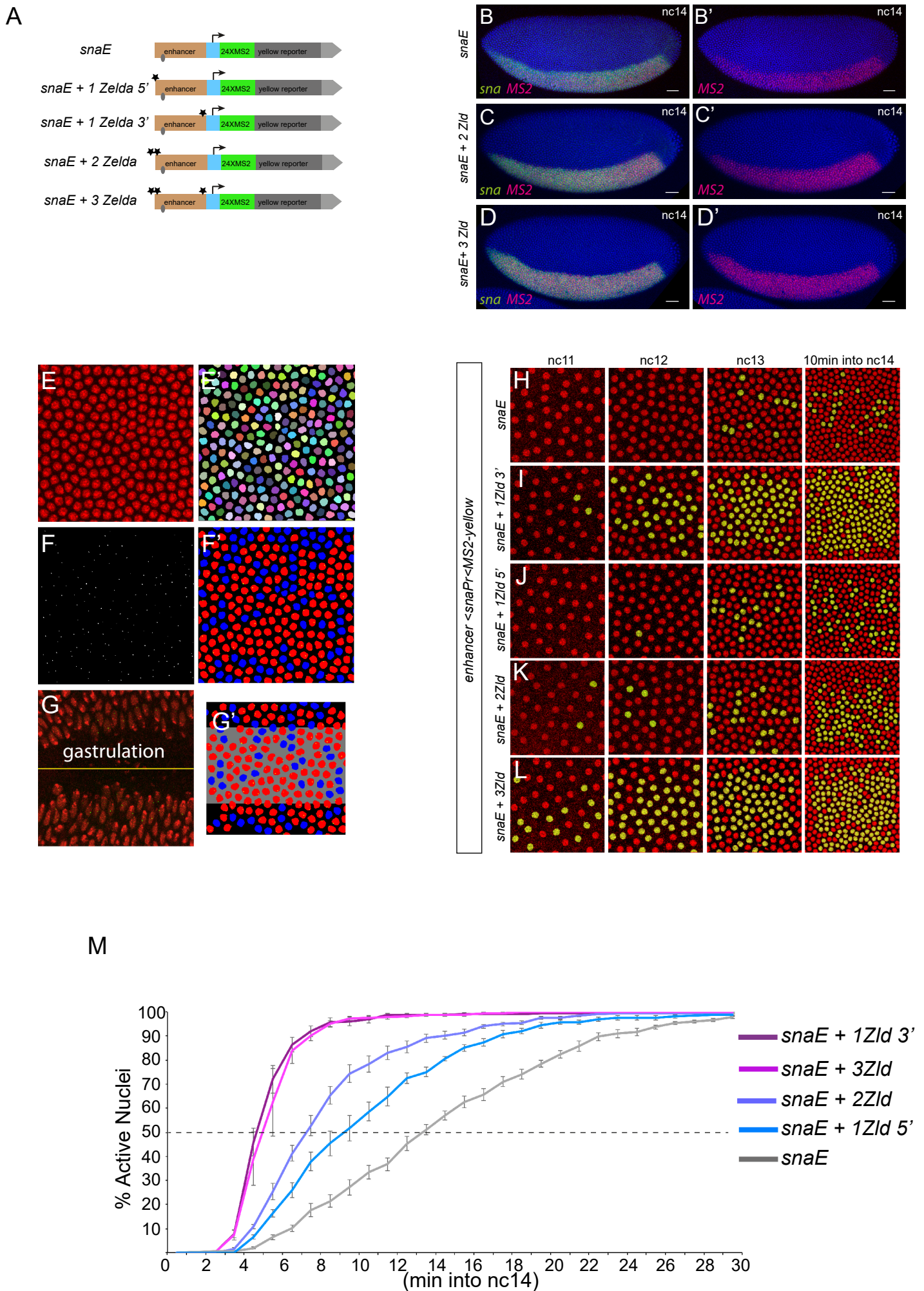
11

Figure1

Figure2

Figure3

Figure 4