Kinetic sculpting of the seven stripes of the *Drosophila even-skipped* gene

Augusto Berrocal[1,*], Nicholas Lammers[2,*], Hernan G. Garcia[1,2,3,4,$] and Michael B. Eisen[1,2,4,5,6,$]

1. Department of Molecular & Cell Biology, University of California at Berkeley, Berkeley, California
2. Biophysics Graduate Group, University of California at Berkeley, Berkeley, California
3. Department of Physics, University of California at Berkeley, Berkeley, California
4. Institute for Quantitative Biosciences-QB3, University of California at Berkeley, Berkeley, California
5. Department of Integrative Biology, University of California at Berkeley, Berkeley, California
6. Howard Hughes Medical Institute

* co-first authors
$ joint senior authors

Direct correspondence to: mbeisen@berkeley.edu

Abstract

We used the MS2 system to visualize the transcriptional dynamics of the *Drosophila melanogaster even-skipped* gene at single-cell and high spatiotemporal resolution spanning nuclear cycle 14. We observe extensive transcriptional bursting across the anteroposterior axis, initially in broad domains that are refined over the course of approximately 15 minutes to form the gene's iconic seven striped expression pattern. We developed a rigorous quantitative framework for analyzing and visualizing these data, that allows us to simultaneously infer the high-resolution temporal dynamics of the transcriptional state of roughly 3,000 nuclei, and characterize how the bursting process varies over time and space. We show that despite being created by the largely independent activity of five discrete enhancers, the seven *eve* stripes are sculpted by the same basic kinetic phenomena - a low burst frequency outside of stripes, and a high burst frequency inside the stripe that increases over time to sharpen the contrast between stripe and interstripe regions.

Introduction

The patterns of gene expression that choreograph animal development are formed dynamically by an interplay between processes - transcription, mRNA decay and degradation, diffusion, directional transport and the migration of cells and tissues - that vary in both space and time. However the spatial aspects of transcription have dominated the study of developmental gene expression, with the role of temporal processes in shaping patterns receiving comparably little attention (Bothma and Levine, 2013).

Gene expression patterns are dynamic on many levels. They form, change and disappear over time, often as cells, tissues and organs are forming and moving in the developing embryo (Lawrence, 1992). Furthermore the transcriptional process that creates these patterns is itself highly dynamic. The classical view of transcription as a switch or a tunable rheostat has been replaced in recent years by a picture in which mRNA synthesis occurs in bursts, with promoters switching stochastically between an ON state when polymerases are loaded and begin elongating, and an OFF state where few or no new transcripts are initiated (Figure 1A) (Blake et al., 2006; Chubb et al., 2006; Golding et al., 2005; Janicki et al., 2004; Zenklusen et al., 2008).

A slew of studies, from theoretical models (Ko, 1991; Peccoud and Ycart, 1995) to imaging-based analyses (Fukaya et al., 2016; Jones et al., 2014; Senecal et al., 2014; Xu et al., 2015), have shown that overall rates of mRNA synthesis can be adjusted by controlling the details of the bursting process. Transcriptional output can be altered by changing the frequency with which bursts occur, their duration, and the rate at which polymerases are loaded during a burst (specified by k_{on}, k_{off} and r, respectively; Figure 1A). Modulating any or all of these parameters over space and time could in principle produce arbitrarily complex output patterns. However, it remains unclear how diverse the kinetic strategies employed by different regulatory sequences actually are, and what, if anything, constrains how these different kinetic parameters are used by evolution to shape patterns of expression.

In this paper we set out to compare the ways that different enhancers that drive similar types of spatiotemporal patterns during animal development deploy transcriptional bursting to produce their outputs. We use as our model the *Drosophila melanogaster even-skipped* (*eve*) gene whose seven stripes ring the embryo in the cellularizing blastoderm (nuclear cycle 14; nc14) in the hour preceding gastrulation.

The *even-skipped* phenotype was originally described in a screen for mutants that alter the segmented pattern of *D. melanogaster* larvae (Nüsslein-Volhard and Wieschaus, 1980), and

named based on the absence of the 2nd, 4th, 6th and 8th abdominal segments. In the absence of *eve* activity, segmentation is abolished (Nüsslein-Volhard et al., 1985). Early imaging studies described *eve* as being expressed broadly in nc13 and early nc14 embryos before refining sequentially into four then seven stripes (Frasch and Levine, 1987; Macdonald et al., 1986). Further details of *eve* expression have been uncovered with improved labeling and imaging techniques, with the current view based on imaging of mRNA and protein levels in series of static images of fixed and stained embryos (Fowlkes et al., 2008; Surkova et al., 2008), which show an initial phase with broad domains in the anterior and posterior, followed by the formation of stripes from within these broad domains and, eventually, amplification of the stripe pattern.

The *eve* stripes are produced through the largely independent activity of five discrete enhancers (Figure 1B) that drive individual stripes (the stripe 1, stripe 2, and stripe 5 enhancers) or pairs of stripes (the stripe 3/7 and stripe 4/6 enhancers) (Goto et al., 1989; Harding et al., 1989; Small et al., 1991). These enhancers respond in different ways to canonical maternal factors Bicoid (Bcd) and Caudal (Cad), and gap genes Hunchback (Hb), Giant (Gt), Krüppel (Kr), Knirps (Kni) and Sloppy Paired 1 (Slp1), balancing activating and repressive inputs to generate novel output patterns. For example, the *eve* stripe 2 enhancer is activated in the anterior by Bcd and Hb, and repressed by Gt and Kr, ultimately expressing in a stripe of nuclei that fall between the domains occupied by these two repressors.

The dynamic nature of *eve* stripe formation was highlighted by recent work from the Levine and Gregor labs, who used the MS2 system (Bertrand et al., 1998; Forrest and Gavis, 2003; Garcia et al., 2013) to directly visualize and quantify transcription from an *eve* stripe 2 transgene (Bothma et al., 2014). They observed that the single stripe driven by this enhancer formed and dissipated rapidly, and, significantly, that it was generated by bursts of transcriptional activity in the nuclei that form the stripe.

The MS2 system, which exploits the interaction between the phage MS2 coat protein (MCP) and a short RNA stem loop to fluorescently label nascent transcripts as they are being synthesized, makes it possible to quantify transcription at the single cell level in living embryos. It has led to a series of experimental and theoretical studies of transcriptional bursting in *Drosophila* (Boettiger and Levine, 2013; Bothma et al., 2014; Fukaya et al., 2016; Holloway and Spirov, 2017; Little et al., 2013; Xu et al., 2015; Zoller et al., 2017) and other systems (Corrigan et al., 2016; Golding et al., 2005; Senecal et al., 2014; Zenklusen et al., 2008).

In this paper we use the MS2 system (Bertrand et al., 1998; Forrest and Gavis, 2003; Garcia et al., 2013) to directly visualize and quantify *eve* transcription as the seven stripes form,

revealing how a complex pattern of gene expression driven by multiple enhancers is built in space and time at the single cell level. Our objective in carrying out this work was twofold: first, to characterize the detailed dynamics of this classical and well-studied pattern, and, second to establish a rigorous systematic framework for analyzing such data, and conceptual paradigms for characterizing what we observe from this new type of experiment. Indeed, the advent of live imaging in the context of development calls for the establishment of a new language and new metrics for characterizing the formation of gene expression patterns in space and time.

We find that the formation of *eve* stripes across all five enhancers is dominated by the same regulatory strategy: the modulation of burst initiation rates (k_{on}) along the anterior-posterior (AP) axis with stripes occurring in regions of high k_{on} separated by regions of low k_{on}. The establishment of this pattern is gradual, and involves the coordinated increase in k_{on} within stripes and its reduction from an initially higher level in most interstripe regions. We further show that the pronounced anterior movement of stripes, particularly in the posterior, involves the dynamic regulation of k_{on} for individual nuclei that shift to and from a quiescent (low k_{on}) and active (high k_{on}) state as they enter and leave stripes. Thus, in this experiment and with our new set of analytical tools, we capture not only how single cell transcriptional activity encodes the formation of the stripes, but also how this activity is modulated in time in order to create and displace a complex pattern of gene activity across the embryo.

Results

Live imaging of *eve* expression

We used recombineering (Warming et al., 2005) to modify a bacterial artificial chromosome (BAC) (Venken et al., 2006) containing the *D. melanogaster eve* gene and all of its enhancers and regulatory elements (Venken et al., 2009), replacing the coding region with an array of MS2 stem loops followed by the *D. melanogaster yellow* (*y*) gene (Figure 1C). The 4,329 base pair *y* gene, which is longer than the endogenous *eve* transcript, is both phenotypically neutral and provides a means to increase the number of Pol II molecules loaded onto the gene in order to amplify signal. We inserted the engineered BAC into a targeted site on chromosome 3L using ΦC31 integrase-mediated recombination (Fish et al., 2007), and homozygosed the line, which showed no signs of adverse effects of the transgene.

We crossed males from this line with females carrying a transgene that expresses in embryos an RFP labeled histone to visualize nuclei, and an MCP-GFP fusion that binds the

MS2 RNA stem loops. The result is the direct visualization and quantification of nascent transcripts at the transgene locus as fluorescent puncta (Garcia et al., 2013). The temporal and spatial pattern of *eve* transgene transcription recapitulates the well-characterized dynamics of *eve* expression, most notably formation of the characteristic seven stripes in the late blastoderm (Figure 2; Video 1).

We used laser-scanning confocal microscopy to record, with high temporal resolution and high magnification, two color (MCP-GFP and histone RFP) movies of embryos from before nc14 through gastrulation at the end of nc14. We optimized our data collection strategy to sample multiple stripes (three to five) in each movie, to obtain high temporal resolution (one Z-stack every 16.8 seconds) and to have optimal signal to noise with minimal bleaching. In total, we collected 11 movies (Videos 2-12, available at http://eisenlab.org/eve7), with every stripe imaged at least five times (range 5-9; see Table 1).

We used a custom image processing pipeline (Bothma et al., 2014; Lammers et al.) to identify nuclei, track fluorescent puncta and extract fluorescence intensities in each nucleus over time. The resulting data (File S1) contains fluorescence traces from 2,961 nuclei at an interpolated time interval of 20s, representative examples of which are shown in Figure 3A.

We first sought to refine the previously characterized temporal dynamics of stripe formation (Surkova et al., 2008) using the increased temporal resolution (relative to earlier analyses of fixed embryos) of these data (Figure 3B). During nc14, we first observe *eve* transcription beginning approximately five minutes after the onset of anaphase. The initial transcription covers a broad swath of the embryo, from around 25% to 75% egg-length, with the highest activity in two domains roughly centered in the anterior and posterior halves of the embryo respectively. The greatest fluorescence signal during the first 25 minutes of nc14, when stripes are not yet fully formed, is in the most anterior region of *eve* transcription, in an area in which stripe 1 will eventually form.

Although the full seven stripe pattern is not fully formed until around 25 minutes, the three anterior-most stripes are already apparent as locally high areas of fluorescence intensity as early as 10 minutes. By 20 minutes stripes 1,2 and 3 have clearly formed, stripes 4 and 6 appear to split off from a large posterior domain, and stripe 7 forms *de novo* in the posterior. Stripe 5 appears last, emerging in an area of low transcriptional activity left behind following the splitting of stripes 4 and 6. The stripes persist for the 30 remaining minutes of nc14, gradually increasing in fluorescence intensity until they reach a peak at around 35 minutes into nc14.

The positions of stripes 1-3 along the AP axis are largely stable after they form, while stripes 4-6 show small anterior shifts. Stripe 7 makes a more dramatic movement towards the anterior, moving approximately 10% of egg-length, or around 45 um from its initial location, which is accompanied by a shift in the nuclei driving stripe expression (Figure 3C). This stripe movement, and the decoupling between stripes and nuclei, necessitated the development of a method, described below, to dynamically define the position of stripes throughout each movie.

A kinetic fingerprint of *eve* stripe formation

The fluorescence traces from individual nuclei show clear hallmarks of transcriptional bursting, with apparent stochastic fluctuations between states with low and high fluorescence output (Figure 3A). We characterize this process and probe transcriptional bursting in individual nuclei and stripes using the simple Markovian process presented in Figure 1A, in which a promoter switches stochastically between an OFF and an ON state with rates k_{on} and k_{off}, and loads polymerases in the ON state with a constant rate r.

In our implementation of the MS2 system, once a polymerase molecule transcribes the stem loops located at the 5' end of the gene, the MCP-GFP molecules bound to the stem loops produce a near constant fluorescent signal at the locus that persists until this polymerase completes its traversal of the gene. We estimated the polymerase transit time by examining the autocorrelation of the data across time (Lammers et al.) as approximately 140 seconds, consistent with a direct measurement of the rate of polymerase elongation of ~2,700 bp/min from (Fukaya et al., 2017) and the length of the construct (6,563bp).

We therefore model a promoter in the ON state as producing a single pulse of fluorescence whose duration is 140 seconds and height is proportional to the polymerase loading rate r (Figure 4A). When the temporal behavior of a single promoter is known, we can represent it as a vector of ON or OFF states (Figure 4B), with each ON period producing a pulse of fluorescence. The sum of these pulses is the expected observed fluorescence signal (Figure 4C).

These modeled traces have the features of real fluorescence traces: linear increases in intensity corresponding to periods when the promoter is ON; plateaus corresponding to period when transcriptional initiation is matched with previously initiated polymerases completing their transit of the gene; and linear signal decays corresponding to periods when the promoter is OFF but previously initiated polymerases are still transiting the gene (Bothma et al., 2014).

However, when given a fluorescence trace, it is not trivial to infer the promoter state sequence that generated it, owing to the time convolution between promoter state and fluorescence output. To solve this problem, we developed a memory-adjusted hidden Markov model (mHMM, described in (Lammers et al.)) that determines the values of k_{on}, k_{off}, and r, and fits the model to the data to infer the promoter state sequence that best describes the fluorescence data from each nucleus, thus providing a quantitative model of the bursting dynamics of *eve* across the nuclei we imaged. File S1 contains the inferred promoter state for each nucleus at every time point and the corresponding modeled fluorescence intensity, representative examples of which are shown in Figure 4D.

In the imaged embryos we imaged here, the MS2 BAC is heterozygous, contributed only by the father while the mother contributes the MCP-GFP. However DNA replication occurs within and average of 10 mins for loci in nc14 (McKnight and Miller, 1977), meaning that there are actually two sister chromatids with the MCP-GFP, and, because of sister chromatid cohesion we can not, in general, discriminate both alleles. As it is still unclear how the sister chromatids influence each other's transcription, we ran the mHMM twice, once assuming a single locus and once with two independent loci. The runs gave qualitatively similar results. For the analyses described below we used the two loci model, which has three states: OFF, one promoter ON, and two promoters ON. Except where noted we collapse to ON states to a single aggregate ON state.

The mHMM accomplishes two aims that are central to treating these data in a more rigorous and biologically meaningful manner. It allows us to describe the bursting behavior of any set of nuclei in quantitative terms, and, crucially, by providing a means to fit a sequence of ON and OFF states to the data from each nucleus, the mHMM also allows us to shift the focus in the analysis of individual traces from fluorescence, which only indirectly reflects the temporal behavior of a promoter, to the instantaneous promoter state.

Collectively these data represent a kinetic fingerprint of stripe formation: a record of every transcriptional burst that occurred during the formation of *eve* stripes in these embryos (Figure 5). To capture the full dynamics we also generated movies showing the transcription and bursting process (Video 13 showing fluorescence over time and space, Video 14 showing individual bursts highlight the kinetics of bursting during of stripe formation).

Dynamic determination of stripe positions

Before analyzing the data further we had to solve two practical problems. To compare the kinetic behavior of individual stripes, we had to determine which nuclei were in each stripe as a function of time, a process complicated by the movement of stripes relative to both the embryo and nuclei (Figure 3C). To analyze the data *in toto*, we also had to register the 11 movies relative to each other and to the embryo.

To solve the first problem, we developed a dynamic spatial binning algorithm to assign nuclei to stripes at each time point in nc14. For each movie, we first clustered nuclei based on position and intensity to partition the field of view into coarse stripe regions at single minute resolution through the period when all seven stripes are delineated (from 25 to 50 mins into nc14). Within each coarse spatio-temporal partition, we applied a weighted fitting routine to the observed distribution of fluorescence to obtain an estimate for the location of the stripe center. As the stripes are not perfectly perpendicular to the anteroposterior axis (the image x-axis), the fit position of the stripe center varies along the y-axis.

We defined the stripe as the region falling within 1.5% AP (approximately 7 um) of the stripe centers. Regions outside of the stripe centers were assigned to the nearest stripe and designated as anterior or posterior flanks depending upon their relative position. As stripe identities are not well defined prior to 25 minutes, we used the average position of the stripe centers from 25 to 30 minutes to categorize nuclei in these earlier time points. Stripe assignments for each nucleus over time are given in File S1.

For many of the figures discussed below, we used a registered AP position for each nucleus at every time point. Briefly, we calculated the mean x axis position for each stripe across movies. The registered AP position of each nucleus included a correction for slight deviations among movies in the position of each stripe along the AP axis and a correction for the variation of the stripe center along the y axis. Thus, our dynamical binning algorithm makes it possible to simultaneously track the behavior of individual nuclei while recognizing the fact that their participation in individual stripes is modulated both in space and time.

Eve stripes: transcriptional oases in bursting deserts

We used the output of the mHMM to examine the locations of transcriptional bursts along the AP axis and over time (Figure 6A). The most striking feature is the almost complete lack of observable transcriptional bursts in the regions between stripes from 25 minutes into nc14 (with the exception of the 5-6 interstripe which is discussed below).

These bursting deserts have several evident features. Their spatial boundaries are sharp, with the transition between the low bursting (quiescent) state and the frequently bursting (active) state occurring from one column of nuclei to the next, consistent with the classical descriptions of *eve* stripe patterns (Small et al., 1991). They also have sharp temporal boundaries: all of the bursting deserts, save that between stripes 6 and 7, form in regions where there was appreciable bursting early in nc14 that disappears at around 25 minutes into nc14 (Figure 6A,B)

To better understand how these transcriptional deserts form, we looked at the bursting history of the nuclei in these regions. The first thing we noticed was that, despite the initial period of activity, a large fraction of nuclei in interstripes do not burst at any point during nc14 (Figure 6C). With the exception of the 5-6 interstripe, these never-ON nuclei effectively form the ultimate boundaries between stripes, as essentially every nucleus within each stripe bursts at some point during nc14 (Figure 6C,D). The large number of never-ON nuclei is not simply due to a shorter period during which bursts occur, as again with the exception of the 5-6 interstripe, there is already a large difference in the number of nuclei that have burst in stripes compared to interstripes by 15 minutes into nc14 (Figure 6D), before interstripe bursting dissipates.

Despite the large number of never-ON nuclei in interstripes, a significant number of nuclei that ultimately end up in transcriptional deserts do burst in the first 25 minutes of nc14, but become quiescent after this time (Figure 6E, blue nuclei, and 6F). The number of such nuclei increasing along the AP axis (Figure 6D), from around 10% of nuclei in the 1-2 interstripe in the anterior to 70% in the 5-6 interstripe in the posterior. Indeed this variation in the number of interstripe nuclei that burst in the first 25 minutes is one of the most striking differences between the different stripes.

The 5-6 interstripe is particularly interestesting, as it has both the highest activity in the first 25 minutes and the highest rate of interstripe bursting after 25 minutes. Expression in the 5-6 interstripe is repressed by Kni acting on the stripe 4-6 enhancer (Clyde et al., 2003) and Gt acting on the stripe 5 enhancer (Fujioka et al., 1999). Given that the relatively common bursting is not observed in the stripe 4-5 interstripe, as would be expected if it were due to incomplete repression by Kni on the stripe 4-6 enhancers, we hypothesize that Gt is imperfectly repressing transcription from the stripe 5 enhancer.

Regardless of the particular molecular mechanisms at work at each enhancer and in each interstripe, there appears to be a unified regulatory strategy - large domains of negligible or reduced bursting - used to sculpt stripes.

Eve stripes are amplified by an increase in k_{on}

Given the reduction of k_{on} over time outside of stripes, we were next interested in examining whether bursting parameters varied within stripes over space and time. To do this, we used the assignment of nuclei to stripes discussed above, and directly estimated bursting parameters within each stripe and its anterior and posterior margins over time using the mHMM-inferred promoter states (Figure 7).

Within stripes, the average fluorescence (Figure 7A) and fraction of nuclei in the ON state (Figure 7B) peak at around 35 min into nc14, the result of a gradual increase in k_{on} in stripe nuclei over time (Figures 7D). This trend is also evident when looking at the spatial distribution of ON rates over time (Figure 6E).

There is comparably less modulation of the off rate within stripe centers over time (Figures 7E), although it is somewhat lower once stripes have formed compared to where it was in the first 25 mins. In contrast the we see a sharp increase in the off rate in the posterior flanks of each stripe (Figure 7E) that is most evident when looking at the average burst length (Figure 7F), which decreases about two-fold fairly consistently across stripes.

We were concerned that this apparent decrease in k_{off} was a byproduct of the way we were assigning nuclei to stripes, so we examined the spatial distribution of bursting parameters inferred directly on a nucleus-by-nucleus level (Figure 8A,B). Both the on rate and the off rate calculated on active (i.e. not quiescent) nuclei show striking patterns with high on rates in the stripe center decreasing along the margins, and an inverse pattern with off rates, in both cases maximizing ON time in the center of the stripes.

To obtain an orthogonal view of the spatial variation in bursting parameters, we reran the mHMM after binning nuclei by their stripe position, the results of which are shown in Figure 8C. This analysis is in general consistent with the results in Figures 7E, 7F, 8A, and 8B, with with a low-high-low pattern across each stripe in k_{on}, and a high-low-high or low-low-high pattern for k_{off}. In addition we see a low-high-low pattern for r from the mHMM, a pattern that is difficult to observe directly.

It is not surprising that there would be different bursting patterns across a stripe, as the more exterior nuclei experience higher levels of repressors, though the consistency of the pattern across stripes is striking. However, the tight coupling of variation in k_{on}, k_{off} (Figure 8-S1) and r was not expected, as these parameters are, at least in principle, unlinked. It has been previously suggested that an inverse coupling between k_{on} and k_{off} may be a general feature of transcriptional control (Zoller et al., 2017), however we suggest below that this may be a

consequence of a mismatch between the variables we use to parametrize the system and its actual kinetics. We also believe that this variation in the stripe margins is best thought of as a consequence of the manner in which stripes are formed - the repression of k_{on} in interstripe regions - rather than a driver of the pattern.

Regardless, there is a clear commonality in the ways that stripes are kinetically sculpted. Two complementary modulations of k_{on} - decreasing k_{on} in the region between stripes to create transcriptional deserts, while increasing k_{on} within each stripe - result in the enhancement of the transcriptional contrast between nuclei inside and outside of the seven *eve* stripes.

Stripe movement is associated with the dynamic switch of nuclei between active and quiescent states

Previous studies have described the anterior shift of the more posterior *eve* stripes during nc14 (Keränen et al., 2006). This shift has been proposed to arise from a combination of nuclear movement (nuclear flow) and movement in the pattern of regulators (expression flow), especially repressors, which are known to shift anteriorly during nc14 as well (Jaeger et al., 2004a, 2004b). We were interested in whether our ability to track nuclei and their bursting behavior over time might refine our understanding of this process.

We began by examining the nuclei in stripe 7, which, as discussed above, undergoes a dramatic movement of approximately 10 percent of embryo length in the final 25 minutes of nc14. Figure 9A shows the AP positions of the nuclei in one movie covering stripe 7 as a function of time, with the period in which they are part of the stripe highlighted.

Remarkably, although the stripe is clearly present throughout this period, no nuclei remain a part of the stripe for the entirety of this 25 minute period. As time progresses, nuclei at the posterior edge of stripe 7 shift from an active state, in which the promoter stochastically alternates between the ON and OFF state, to a quiescent state in which we observe no subsequent bursting. In contrast, nuclei just off the anterior edge of the stripe switch from a quiescent to an active state at roughly the same rate. This leads to a net overall anterior movement of the stripe, akin to treadmilling, at a velocity of approximately one percent of embryo length every three minutes.

The other stripes exhibit smaller and varied anterior shifts (Figure 9-S1), but in every case the shift is associated with a similar coupled gain of active nuclei along the anterior edge and loss along the posterior edge. This effect is most clearly seen in Figure 9B, which shows, for each time point where a nucleus initiates a new burst, the difference in activity (defined as the

fraction of the time the nucleus is in the ON state in the subsequent 10 min compared to the preceding 10 min). High (red) values represent nuclei that are increasing in activity while low (blue) values represent nuclei decreasing in activity. For all seven stripes there is a clear spatial pattern, with red nuclei along the anterior edge of the stripe and blue nuclei along the posterior edge, indicating a movement of stripes relative to nuclei, with new nuclei being added at the anterior edge while nuclei are dropped at the posterior edge of the stripe.

Thus, while we see some evidence for the modulation of k_{off} and r, the regulation of burst initiation - k_{on} - is the primary force in sculpting *eve* stripes. Reducing k_{on} is used to create the large areas of limited transcription that define the broad shape of the *eve* pattern; increasing k_{on} is used to boost mRNA production within stripes. And the spatially coordinated switching from low to high k_{on} states is used to reposition stripes along the AP axis. It is these three different manifestations of the regulation of the initiation of transcriptional bursts working together that create the complex and dynamic spatiotemporal gene expression pattern that is *even-skipped*.

Discussion

The most remarkable aspect of *eve* regulation is that what appears to be a regular, repeating pattern of nearly identical stripes is created by the largely independent activity of five separate enhancers responding to different combinations of activators and repressors (Arnosti et al., 1996; Fujioka et al., 1995, 1999; Small et al., 1991, 1992).

Our analysis of the time-dependent bursting behavior of nuclei across the embryo reveals that, despite this uniqueness of molecular inputs to each stripe, they employ the same basic regulatory strategy for creating a stripe of transcriptionally active nuclei. We show that all seven stripes are sculpted by flanking regions of high and increasing bursting activity with transcriptional deserts, consisting of a mix of nuclei that never burst during nc14 and nuclei whose bursting is suppressed after stripe formation. We also find that the partitioning of nuclei into stripes and interstripes is dynamic even after stripe formation, with many nuclei, especially in the posterior, switching between active and quiescent states, driving an anterior shift of stripe positions.

The primacy of k_{on} modulation

Although in principle complex patterns of transcription could be generated by regulating k_{on}, k_{off}, r, each of the key features of *eve* stripe regulation we observe here involves the modulation of k_{on}. The use of k_{on} as the primary parameter under dynamic control is consistent with previous

works in flies and other organisms where bursting parameters were inferred from single-molecule FISH experiments using theoretical models (Jones et al., 2014; Little et al., 2013; Senecal et al., 2014; So et al., 2011; Xu et al., 2015, 2016; Zenklusen et al., 2008; Zoller et al., 2017), or where MS2 traces were manually analyzed (Fukaya et al., 2016; Golding et al., 2005). However, these previous works were not capable of resolving the temporal regulation of bursting parameters. In contrast, our mHMM approach makes it possible to infer the promoter state at over development hence revealing k_{on}, k_{off} and r as as function of space and time.

Given the abundant evidence that *eve* stripe boundaries are set by repressors (Clyde et al., 2003), it is all but certain that the low k_{on} in transcriptional deserts is a product of the "short-range" repressive action of Gt, Kr, Kni, Hb and potentially other proteins. (These proteins are known as "short-range" repressors because they act locally, on the scale of at most a few hundred base pairs, to affect the ability of activators to bind nearby or to activate transcription when bound (Kulkarni and Arnosti, 2005; Li and Arnosti, 2011; Payankaulam and Arnosti, 2009) - a functional necessity for a gene where one *eve* enhancer can be active in the same nucleus that another nearby enhancer is being repressed.)

Our observation that, as stripes move towards the anterior, nuclei transition from a low k_{on} state to a bursting state along the anterior edge of the stripe, and from a bursting state to a low k_{on} state along the posterior edge, offers some mechanistic insights into repressor activity. Specifically, if we assume that these shifts are due to the anterior movement of the array of repressors (Jaeger et al., 2004a, 2004b) (see below), the rapid shifts to and from high and low k_{on} states at all stripes suggests both that the mechanism of action of each of the repressors is similar, and that repression is likely due to the direct effects of transcription factor binding making it possible for bursting to change rapidly as transcription factor nuclear concentrations are modulated.

The origins of the increase in k_{on} within stripes is less clear. The patterns and levels of both activators and repressors are changing rapidly during nc14 (Dubuis et al., 2013), creating multiple paths towards increasing k_{on}. Given the role of repressors in suppressing k_{on}, one possible factor in the increase in k_{on} within stripes is the sharpening of the patterns of repressors, which could reduce repressor levels within stripes resulting in a higher k_{on}.

Although k_{on} is the primary point of regulation, we also observe spatial regulation of r, which is significantly higher in stripe 1 than in the other six stripes, and k_{off}, which is consistently higher along the posterior border of several stripes. Thus, at least under some circumstances, other aspects of bursting than k_{on} are accessible to regulation, suggesting that there is either some

functional advantage attached to the reliance on k_{on} (such as less noise, lower energetic costs, or less susceptible to mutation), or that the regulation of k_{on} is more evolutionarily accessible.

Stripe movement is driven primarily by expression flow

Just as gene expression patterns are dynamic in time (Bothma et al., 2014) , they are dynamic in space, resulting in the movement of expression domains throughout the embryo during development (Jaeger et al., 2004a; Keränen et al., 2006). The anterior movement of *eve* stripes during nc14 has been previously described (Keränen et al., 2006), and proposed to arise from a combination of nuclear movement (nuclear flow) and movement in the pattern of regulators (expression flow), especially repressors, which are known to shift anteriorly during nc14 as well (Jaeger et al., 2004a, 2004b). While Keränen *et al.* (Keränen et al., 2006) concluded that the relative contributions of these two forces were roughly equal, our data suggest that, especially in the posterior, expression flow dominates the anterior shift of the *eve* stripes.

A typical nucleus in stripe 7 moves around one percent of embryo length in the final 25 min of nc14. The stripe, however, moves around six percent of embryo length during that time (see Figure 9A). Because we are tracking both the position and activity of individual nuclei, we can visualize expression flow in action. We see nuclei transition from low activity in the anterior interstripe to high activity in the stripe, from high activity in the stripe to low activity off the posterior flank of *eve* expression, and in some cases both.

This effect is most pronounced for the posterior stripes (Figure 9-S1) but is observed for the more anterior stripes as well, although the magnitude of the shift decreases for more anterior stripes. The difference in the amount of the effect we and Keränen *et al.* attribute to expression flow is likely an effect of differences in the data used. Because we are looking at instantaneous transcription rates while they looked at accumulated mRNA, there is a considerable temporal lag and integration of the transcriptional activity over the life time of *eve* mRNA in their data, which has the effect of underestimating the extent to which the stripes actually move.

We also note that the extent to which nuclear flow by itself would be expected to shift output patterns measured as a function of position in the embryo is unclear, as it would depend on the extent to which the repositioning of regulators drives movement of nuclei (which it is believed to do (Blankenship and Wieschaus, 2001)), and the corresponding effect that nuclear movement has on the positioning of regulators, which remains largely unknown.

One open question relates to the temporal relationship between changes in the position of the repressor array that drives stripe position and the transcriptional output of the stripes. Recent advances in the simultaneous monitoring of protein concentration and transcriptional output in living embryos should help answer this question in the near future (Bothma et al., 2018).

Appropriate parameterization of bursting

For the analyses presented here we employ a now fairly standard parameterization of bursting using intrinsic on, off and polymerase loading rates. That we see such a strong interdependence between k_{on}, k_{off} and fluorescence warrants a closer look at the origins of this correlation. Experiments with simulated data show that the mHMM can respond independently to variation in one but not the others of these parameters (Lammers et al.). As the mHMM explicitly embeds these rates in its model, we implemented an alternative method to fit pulses to the fluorescence data that directly fits bursts but does not rely on rates, and we see the same effect. Thus the correlation reflects a real property of the data and is not an artifact of the method we used to fit it.

Several recent analyses have looked at autocorrelation and various parameterizations of transcriptional bursting and noise, as well as high order cumulants of the fluorescence distribution to evaluate different models for transcriptional regulation against signatures of transcriptional kinetics (Desponds et al., 2016; Tostevin et al., 2012; Zoller et al., 2017). In particular (Zoller et al., 2017) conclude that there is an anti-correlation between k_{on} and k_{off}, and based on comparison of cumulants under different models, favor one with a constant correlation switching time $1/(k_{on} + k_{off})$ in which different output levels are achieved by trading off one of these parameters against the other.

One way to look at this more conceptually is to note that k_{on} and k_{off} are artificial parameters that describe a system in which the probability of a complex and mechanistically ill-defined biochemical event - the initiation of a transcriptional burst - is dependent on whether the promoter is actively loading polymerases or is not. Consider instead a process - say the physical interaction of an enhancer loaded with transcription machinery - and a promoter that is independent of the state of the promoter. If it occurs when the promoter is OFF, it turns it ON. If it occurs when the promoter is ON, it extends the duration of the current burst (e.g. by increasing the local supply of transcription components).

If these loading events occur randomly but with an adjustable frequency f_b, changing this f_b would adjust the apparent k_{on}, k_{off} and fluorescence in the correlated way we observe. For example, increasing f_b would increase the measured k_{on} as the rate at which new bursts are initiated while the promoter is in the OFF state is dependent directly on f_b, would decrease the measured k_{off} as the rate at which bursts are extended would increase, and would increase the total output of the promoter.

Whether this is the correct model and if so what its details are will require additional data and mechanistic insights. We raise it here because we believe our data suggest that the underlying model we and others use may not be parameterized in a appropriate manner, and that while this model is computationally useful, we wish to caution too strict a mechanistic interpretation of these parameters.

<u>A stripe is a stripe is a stripe is a stripe is a stripe is a stripe is a stripe</u>

The seven *eve* stripes, produced by the action of five distinct enhancers, form in different ways. Stripes 1, 2 and 3 appear early, with clear delineation by around 10 minutes into nc14. Some nuclei burst early in the region between the stripe 1, 2 and 3 anlage, but this activity is almost completely gone by 25 minutes. Within these three stripes there is a clear anterior to posterior gradient in the rate at which k_{on} increases (Figure 7D), perhaps reflecting their different positions along the Bcd morphogen gradient. Stripes 4, 5 and 6 become recognizable later than the more anterior stripes, although it is unclear why this is the case (the factors involved in the activation of these stripes are, remarkably, still unclear). Stripe 7 is unique in that it emerges late from an area in which there was little or no previous bursting. There are also subtle differences in bursting behavior between stripes, and within each stripe, both k_{on} and k_{off} differ between the middle and edges of each stripe.

Nonetheless, once stripes are formed after around 25 mins into nc14, when nuclei are in a stripe their bursting behavior is remarkably similar, no matter which stripe they happen to find themselves, no matter which enhancer is driving their expression, no matter which activator they are responding to and no matter which repressor is shaping their boundaries. The bursting kinetics of the different stripes are highly similar, and vary in a similar way across time (k_{on} increases) and across space (lower k_{on} at the margins; high k_{off} at edges, especially the posterior edge). The same dropping of nuclei at the posterior and adding them at the anterior drives their movement.

Before generating these data, we expected to find that the two pairs of stripes driven by common enhancers - 3 and 7, and 4 and 6 - might be more similar to each other than they are to the other five stripes. But we do not find this to be the case. Likewise, we expected to see different bursting behavior due to the different activators that regulate the different stripes, and different repressive mechanisms due to the the multiple gap proteins involved in repression. But we do not see such an effect either.

All in all, this commonality of processes used to sculpt *eve* stripes is the most striking finding of this paper, and is an emerging theme of studies of transcriptional kinetics. Zoller et al. recently reached a similar conclusion when comparing the dynamic transcription of different genes expressed in the cellular blastoderm (Zoller et al., 2017). Here we find it for different enhancers within the same gene. Whether this represents a constraint imposed by the features of the *eve* promoter or the *Drosophila* cellular blastoderm, or if it reflects convergence among different enhancers or promoters for a common optimal strategy, remains to be seen.

Characterizing dynamics patterns demands dynamics measurements

Although *eve* is arguably the gene expression pattern whose *spatial* regulation is best understood, by tracking its single-cell expression as stripes form during nc14, we have revealed several aspects of its *dynamics*, detailed above, that had not been previously appreciated. The most significant observations, the large number of nuclei that never turn on, the measurements of the rates at which the promoter turns on and off, and the shifts in the activity state of nuclei as the repressor arrays move anteriorly, could only have been made with this kind of real-time, single cell data. Further, our data uncovered a plethora of spatiotemporal features that lead to stripe formation whose regulatory rules and detailed underlying molecular mechanisms are yet to be uncovered.

That gene expression is a fundamentally dynamic process is not new information. However, the tools we have had at our disposal to study gene expression so far have tended to emphasize its static features, down to the language we use to describe the transcriptional output of a gene. In textbooks and the scientific literature, *eve* has a gene expression pattern consisting of seven stripes. But, as some earlier work emphasized, and we have directly visualized here, the transcriptional output of *eve*, rather than a single “pattern” is a rapidly changing as a function of time and space: it is dynamic at many time scales and across space and nuclear positions. Indeed, at no point does *eve* approach anything even remotely like a steady state.

We are at the dawn of a new period in the study of transcription, as new experimental techniques and advanced microscopy allow us to monitor transcriptional regulators, observe their behavior at the single-molecule level, and track the transcriptional output of a gene in living, developing animals. While the focus in this paper was on a single gene in a single species, we hope this and our accompanying work (Lammers et al.) have a broader impact by beginning to establish rigorous frameworks for quantifying, characterizing and visualizing the dynamics of transcription at the single-cell level during development that will be required in the era of live imaging of transcription in development.

Methods

Generation of MS2 tagged *eve* BAC

We used bacterial recombineering (Warming et al., 2005) to modify a bacterial artificial chromosome (BAC) (Venken et al., 2006) containing the *D. melanogaster eve* gene and all of its enhancers and regulatory elements (BAC CH322-103K22) (Venken et al., 2009), replacing the coding region with an array of 24 MS2 stem loops fused with *D. melanogaster yellow* gene [(Bothma et al., 2014)] (Figure 1B). We inserted our eve::MS2::yellow BAC-based construct in *D.melanogaster* genome at chromosome 3L through ΦC31 integrase-mediated recombination (see Generation of fly lines), and generated a viable homozygous fly line (w-; +; eve::MS2::yellow).

Specifics of recombineering

We modified a CHORI BAC CH322-103K22 derived from (Venken et al., 2009), which contained the entire *eve* locus and a GFP reporter instead of *eve* coding sequence (CH322-103K22-GFP). We replaced the GFP reporter with MS2::*yellow* (6665 bp) through a two step, scarless, *galK* cassette-mediated bacterial recombineering (Warming et al., 2005). Briefly, we transformed our starting CH322-103K22-GFP BAC in *E.coli* recombineering strain SW102. We then electroporated the strain with a *galK* cassette flanked by 50bp-long DNA homology arms homologous to the MS2::*yellow* (6665 bp) reporter. Upon electroporation, we selected transformants on M63 minimal media plates with galactose as a single carbon source. We achieved a correct replacement of GFP sequence by *galK* cassette in the BAC context (CH322-103K22-galK), validated by observing the digestion patterns produced by ApaLI restriction enzyme.

We purified CH322-103K22-galK BAC and transformed it into fresh *E.coli* SW102 cells. Next, we electroporated these cells with the purified MS2::*yellow* insert. We used M63 minimal media plates with 2-deoxy-galactose to select against bacteria with a functional *galK* gene, and used colony PCR to screen for colonies with a correct MS2::*yellow* insertion (CH322-103K22-MS2) replacing the *galK* cassette. We validated this insertion by observing ApaLI, XhoI, SmaI, and EcoRI restriction digestion patterns and through PCR and Sanger sequencing of the insertion junctions. We transformed our CH322-103K22-MS2 BAC in *E.coli* EPI300 cells to induce high copy number and purified it with a Qiagen plasmid Midiprep kit.

Generation of fly lines

We sent a sample of our purified CH322-103K22-MS2 BAC to Rainbow Transgenic Flies, Inc. for injection in *D.melanogaster* embryos bearing a ΦC31 AttP insertion site in chromosome 3L (Bloomington stock #24871; landing site VK00033; cytological location 65B2). We received the flies that resulted from that injection and used a balancer fly line (w- ; + ; +/TM3sb) to obtain a viable MS2 homozygous line (w- ; + ; MS2::*yellow*). We used line (yw; His::RFP; MCP::GFP) as the maternal source of Histone-RFP and MCP-GFP (Garcia et al., 2013).

Embryo Collection and Mounting

Embryo collection and mounting was done as specified in (Garcia and Gregor, 2018). In short, we set fly crosses between ~30 males (w-; +; eve::MS2::yellow) and ~80 females (yw; His::RFP; MCP::GFP) in a plastic cage capped with a grape juice agar plate. We collected embryos from cages two to ten days old by adding a fresh plate for 30 minutes and aging for 60 minutes to target embryos 90 min or younger. We embedded a gas-permeable Lumox Film (Sarstedt - Catalog # 94.6077.317) on a microscope slide hollowed on the center. Then, we coated the hydrophobic side of the Lumox film with heptane glue and let it dry. The film allows oxygenation of embryos during the 2-3h long imaging sessions while heptane immobilize them.

We soaked an agar plate with Halocarbon 27 oil, picked embryos with forceps, and laid them down on a 3 x 3 cm piece of paper tissue. We dechorionated embryos by adding 2 drops of bleach diluted in water (5.25%) on the paper tissue and incubating for 1.5 minute. We removed bleach with a clean tissue and rinsed with ~4 drops of distilled water. We then placed the tissue paper with dechorionated embryos in water, and picked buoyant embryos with a brush.

We lined ~30 apparently healthy embryos on the Lumox film slide and added 2-3 drops of Halocarbon 27 oil to avoid desiccation, and covered the embryos with a cover slip (Corning® Cover Glass, No.1, 18 x 18mm) for live imaging.

Imaging and Optimization of Data Collection

Movies of embryonic development were recorded on a Zeiss-800 confocal laser scanning microscope in two channels, (EGFP: 488 nm; TagRFP: 561 nm). We imaged embryos on a wide field of view, along their anterior-posterior axis, of 1024 x 256 pixels (202.8μm x 50.7μm), encompassing 3-5 stripes per movie. We optimized laser power, scanning parameters, master

gain, pinhole size and laser power to optimize signal to noise ratio without significant photobleaching and phototoxicity.

For imaging, the following microscope settings were used: 63X oil-objective, scan mode 'frame', pixel size of 0.2µm, 16 bits per pixel, bidirectional scanning at a speed of 7, line step of 1, laser scanner dwelling per pixel of 1.03µs, laser scanner averaging of 2, averaging method Mean, averaging mode Line, 488 nm laser power of 30µW (EGFP), 561 nm laser power of 7.5µW (TagRFP) (both powers were measured with the 10X air-objective), Master Gain in EGFP detector of 550V, Master Gain in TagRFP detector of 650V, Digital Offset in both detectors of 0, Digital Gain in both detectors of 1.0, and a pinhole size of 1 airy unit under the imaging conditions mentioned above (44µm, 0.7µm/section), laser filters EGFP:SP545 and TagRFP:LBF640. This resulted in an imaging time of 633 ms per frame and a full Z-stack of 21 frames in intervals of 0.5µm every 16.8s. We stopped imaging after 50 min into nuclear cycle 14, and took mid-sagittal and surface pictures of the whole embryo for localization of the recorded field of view along the embryo's AP axis.

Image processing

We used a Matlab computational pipeline based on (Garcia et al., 2013; Lammers et al.) to segment and extract numeric data from our raw movies. Briefly, this software segments and processes the images from the two channels (channel 1: MCP::GFP, channel 2: Histone::RFP) on which we collected our data. For segmentation of channel 1, we used Fiji-Weka Segmentation 3D software; this machine-learning-based method relies on the manual segmentation of a variety of MCP::GFP labeled transcriptional foci in a given 21 frame Z-stack from a single dataset (EVE_D11) to produce a model for the segmentation of all datasets recorded under the same imaging conditions. Next, we segmented and tracked the Histone::RFP labeled nuclei on channel 2. Subsequently, we assigned MCP::GFP labeled transcriptional foci to their corresponding Histone::RFP labeled nuclei. Since we collected whole embryo pictures of each of our datasets, we were able to match and locate the recorded fields of view to their right position along the body of their corresponding embryos. Finally, we extracted position and fluorescence values over time of all transcriptional foci to generate data structures ready to use in further analyses.

Stripe assignment

Stripes were called using a two step dynamic spatial binning approach. First, for each data set, 1-dimensional k-means clustering was employed to partition the field of view into coarse stripe regions for each ten minute period between 10 and 50 minutes into nc14. For simplicity, all observed fluorescence was projected onto the long imaging axis ("x"), which coincides (within 1-3 degrees) with the AP axis of the embryo. Spots were weighted according to their fluorescence, ensuring that the clustering routine returned regions that reflected the distribution of *transcriptional activity* along the AP axis, rather than the distribution of *transcribing loci*. Thus, stripe centers correspond to the activity-weighted "center of mass" for a given window of time and space. All results of the automated clustering routine were reviewed manually and corrected as needed.

The coarse stripe partitions returned by k-means clustering served as the basis for a polynomial fitting routine designed to track stripe centers in two dimensions over time. Stripe partitions--estimated at 10 minute intervals--were interpolated to obtain minute-by-minute estimates for domains of stripe activity spanning from 10 to 50 minutes into nc14. For each minute, a third degree polynomial was fit to the observed distribution of fluorescent spots within a given partition using the Matlab "polyfit" function to obtain an estimate for the location of the stripe center. By fitting stripe profiles along two dimensions (AP and DV axes), our approach captured cases in which stripes deviated from perfect DV alignment. Fitting a third degree polynomial was found to provide an optimal balance between structure and flexibility in tracking noisy, nonlinear stripe activity centers. As with the 1 dimensional clustering, spots were weighted by their fluorescent intensity. A gaussian smoothing kernel was also applied to raw fluorescence maps used for stripe fitting. This was found to improve the robustness of stripe fits to artifacts arising from the relative sparsity of the distribution of fluorescent spots within a given spatio-temporal region.

Stripe center regions were defined as the area falling within 1.5% AP (~5 um) of the estimated stripe center lines. Regions not within 1.5% AP of a stripe center were assigned to the nearest center and designated as anterior or posterior flanks depending upon their position relative to their "parent" stripe. Dynamic binning was not used for time points prior to 25 minutes, as stripe identities are not well defined. Instead, stable stripe regions were established using the average position of the stripe center from 25 to 30 minutes. Due to its late development, dynamic binning was not used for stripe 5 prior to 30 minutes into nc14.

<u>Memory-Adjusted Hidden Markov Model</u>

For this work we employed a statistical method that utilizes memory-adjusted Hidden Markov Models to infer bursting parameters from experimental fluorescence traces. The theory and implementation of this method are described in detail in (Lammers et al.). Briefly, parameters were inferred using a standard version of the Expectation Maximization (EM) Algorithm implemented using custom-written scripts in Matlab. Bootstrap sampling was used to estimate the standard error in our parameter estimates. Subsets of 8,000 data points were used to generate time-averaged parameter estimates. Sample sizes for windowed inference varied due to data set limitations. When possible, samples of 4,000 points were used. Inference was not conducted for spatio-temporal regions for which fewer than 1,000 time points were available.

Author contributions

AB designed and generated the labeled *eve* transgene, collected all of the imaging data and ran the initial image processing. NL and MBE performed all of the higher level analyses and generated figures. MBE wrote the paper. HGG and MBE conceived of the experiments. HGG and MBE provided funding and supervised every aspect of the project.

Data Availability

All of the raw and processed data described in this paper as well as links to code and computational notebooks are available at http://www.eisenlab.org/eve7.

Acknowledgements

This work was supported by an HHMI Investigator award to MBE. HGG was supported by the Burroughs Wellcome Fund Career Award at the Scientific Interface, the Sloan Research Foundation, the Human Frontiers Science Program, the Searle Scholars Program, the Shurl & Kay Curci Foundation, the Hellman Foundation, the NIH Director's New Innovator Award (DP2 OD024541-01), and an NSF CAREER Award (1652236). NL was supported by NIH Genomics and Computational Biology training grant 5T32HG000047-18. AB was supported by a University of California Institute for Mexico and the United States (UC MEXUS) - CONACyT Doctoral Fellowship.

Tables

Table 1. Summary of movies collected.

Embryo ID	Duration	Stripes	Link to Movie
EVE_D1	255 frames 71.2 min	1-4	http://www.eisenlab.org/eve7/movies/Video2.mp4
EVE_D2	254 frames 70.9 min	3-7	http://www.eisenlab.org/eve7/movies/Video3.mp4
EVE_D3	235 frames 65.6 min	3-6	http://www.eisenlab.org/eve7/movies/Video4.mp4
EVE_D4	246 frames 68.7 min	3-7	http://www.eisenlab.org/eve7/movies/Video5.mp4
EVE_D5	210 frames 58.6 min	4-7	http://www.eisenlab.org/eve7/movies/Video6.mp4
EVE_D6	196 frames 54.7 min	4-7	http://www.eisenlab.org/eve7/movies/Video7.mp4
EVE_D7	208 frames 58.1 min	3-7	http://www.eisenlab.org/eve7/movies/Video8.mp4
EVE_D8	232 frames 64.8 min	1-3	http://www.eisenlab.org/eve7/movies/Video9.mp4
EVE_D9	322 frames 89.9 min	1-4	http://www.eisenlab.org/eve7/movies/Video10.mp4
EVE_D10	267 frames 74.5 min	1-3	http://www.eisenlab.org/eve7/movies/Video11.mp4
EVE_D11	307 frames 85.7 min	1-4	http://www.eisenlab.org/eve7/movies/Video12.mp4

All videos frame per 16.76 sec

Videos

Video 1. Full expression pattern of *eve*-MS2 BAC

Maximum value projection of z-stacks of entire embryo carrying *eve*-MS2 BAC, MCP-GFP and histone-RFP imaged with a 40x objective.

Videos 2-12. Individual dataset movies

Maximum value projection of z-stacks of sections of embryos carrying *eve*-MS2 BAC, MCP-GFP and histone-RFP imaged with a 63x objective capturing 3-5 stripes as described in Table 1.

Video 13. Dynamics of fluorescence across all seven *eve* stripes

Measured fluorescence in every nucleus at every time point graphed at its AP (x-axis) and image y (y-axis) position with the size of each marker proportional to the fluorescence. Marker is colored green when the promoter is in the ON state and red when it is in the OFF state.

Video 14. Kinetic fingerprint of *eve* stripe formation

Nuclei are graphed at every time point graphed at its AP (x-axis) and image y (y-axis) position when the mHMM inferred that the promoter was in the ON state.

References

Arnosti, D.N., Barolo, S., Levine, M., and Small, S. (1996). The eve stripe 2 enhancer employs multiple modes of transcriptional synergy. Development *122*, 205–214.

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H., and Long, R.M. (1998). Localization of ASH1 mRNA particles in living yeast. Mol. Cell *2*, 437–445.

Blake, W.J., Balázsi, G., Kohanski, M.A., Isaacs, F.J., Murphy, K.F., Kuang, Y., Cantor, C.R., Walt, D.R., and Collins, J.J. (2006). Phenotypic consequences of promoter-mediated transcriptional noise. Mol. Cell *24*, 853–865.

Blankenship, J.T., and Wieschaus, E. (2001). Two new roles for the Drosophila AP patterning system in early morphogenesis. Development *128*, 5129–5138.

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

Bothma, J., and Levine, M. (2013). Development: Lights, Camera, Action — The Drosophila Embryo Goes Live! Curr. Biol. *23*, R965–R967.

Bothma, J.P., Garcia, H.G., Esposito, E., Schlissel, G., Gregor, T., and Levine, M. (2014). Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living Drosophila embryos. Proceedings of the National Academy of Sciences *111*, 10598–10603.

Bothma, J.P., Norstad, M.R., Alamos, S., and Garcia, H.G. (2018). LlamaTags: A Versatile Tool to Image Transcription Factor Dynamics in Live Embryos. Cell *0*.

Chubb, J.R., Trcek, T., Shenoy, S.M., and Singer, R.H. (2006). Transcriptional pulsing of a developmental gene. Curr. Biol. *16*, 1018–1025.

Clyde, D.E., Corado, M.S.G., Wu, X., Paré, A., Papatsenko, D., and Small, S. (2003). A self-organizing system of repressor gradients establishes segmental complexity in Drosophila. Nature *426*, 849–853.

Corrigan, A.M., Tunnacliffe, E., Cannon, D., and Chubb, J.R. (2016). A continuum model of transcriptional bursting. eLife Sciences *5*, e13051.

Desponds, J., Tran, H., Ferraro, T., Lucas, T., Romero, C.P., Guillou, A., Fradin, C., Coppey, M., Dostatni, N., and Walczak, A.M. (2016). Precision of Readout at the hunchback Gene: Analyzing Short Transcription Time Traces in Living Fly Embryos. PLoS Comput. Biol. *12*, e1005256.

Dubuis, J.O., Samanta, R., and Gregor, T. (2013). Accurate measurements of dynamics and reproducibility in small genetic networks. Mol. Syst. Biol. *9*, 639.

Fish, M.P., Groth, A.C., Calos, M.P., and Nusse, R. (2007). Creating transgenic Drosophila by microinjecting the site-specific phiC31 integrase mRNA and a transgene-containing donor plasmid. Nat. Protoc. *2*, 2325–2331.

Forrest, K.M., and Gavis, E.R. (2003). Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for nanos mRNA localization in Drosophila. Curr. Biol. *13*, 1159–1168.

Fowlkes, C.C., Hendriks, C.L.L., Keränen, S.V.E., Weber, G.H., Rübel, O., Huang, M.-Y., Chatoor, S., DePace, A.H., Simirenko, L., Henriquez, C., et al. (2008). A quantitative spatiotemporal atlas of gene expression in the Drosophila blastoderm. Cell *133*, 364–374.

Frasch, M., and Levine, M. (1987). Complementary patterns of even-skipped and fushi tarazu expression involve their differential regulation by a common set of segmentation genes in Drosophila. Genes Dev. *1*, 981–995.

Fujioka, M., Jaynes, J.B., and Goto, T. (1995). Early even-skipped stripes act as morphogenetic gradients at the single cell level to establish engrailed expression. Development *121*, 4371–4382.

Fujioka, M., Emi-Sarker, Y., Yusibova, G.L., Goto, T., and Jaynes, J.B. (1999). Analysis of an even-skipped rescue transgene reveals both composite and discrete neuronal and early blastoderm enhancers, and multi-stripe positioning by gap gene repressor gradients. Development *126*, 2527–2538.

Fukaya, T., Lim, B., and Levine, M. (2016). Enhancer Control of Transcriptional Bursting. Cell *166*, 358–368.

Fukaya, T., Lim, B., and Levine, M. (2017). Rapid Rates of Pol II Elongation in the Drosophila Embryo. Curr. Biol. *27*, 1387–1391.

Garcia, H.G., and Gregor, T. (2018). Live Imaging of mRNA Synthesis in Drosophila. Methods Mol. Biol. *1649*, 349–357.

Garcia, H.G., Tikhonov, M., Lin, A., and Gregor, T. (2013). Quantitative imaging of transcription in living Drosophila embryos links polymerase activity to patterning. Curr. Biol. *23*, 2140–2145.

Golding, I., Paulsson, J., Zawilski, S.M., and Cox, E.C. (2005). Real-Time Kinetics of Gene Activity in Individual Bacteria. Cell *123*, 1025–1036.

Goto, T., Macdonald, P., and Maniatis, T. (1989). Early and late periodic patterns of even skipped expression are controlled by distinct regulatory elements that respond to different spatial cues. Cell *57*, 413–422.

Harding, K., Hoey, T., Warrior, R., and Levine, M. (1989). Autoregulatory and gap gene response elements of the even-skipped promoter of Drosophila. EMBO J. *8*, 1205–1212.

Holloway, D.M., and Spirov, A.V. (2017). Transcriptional bursting in Drosophila development: Stochastic dynamics of eve stripe 2 expression. PLoS One *12*, e0176228.

Jaeger, J., Surkova, S., Blagov, M., Janssens, H., Kosman, D., Kozlov, K.N., Manu, Myasnikova, E., Vanario-Alonso, C.E., Samsonova, M., et al. (2004a). Dynamic control of positional information in the early Drosophila embryo. Nature *430*, 368–371.

Jaeger, J., Blagov, M., Kosman, D., Kozlov, K.N., Manu, Myasnikova, E., Surkova, S., Vanario-Alonso, C.E., Samsonova, M., Sharp, D.H., et al. (2004b). Dynamical analysis of

regulatory interactions in the gap gene system of Drosophila melanogaster. Genetics *167*, 1721–1737.

Janicki, S.M., Tsukamoto, T., Salghetti, S.E., Tansey, W.P., Sachidanandam, R., Prasanth, K.V., Ried, T., Shav-Tal, Y., Bertrand, E., Singer, R.H., et al. (2004). From silencing to gene expression: real-time analysis in single cells. Cell *116*, 683–698.

Jones, D.L., Brewster, R.C., and Phillips, R. (2014). Promoter architecture dictates cell-to-cell variability in gene expression. Science *346*, 1533–1536.

Keränen, S.V.E., Fowlkes, C.C., Luengo Hendriks, C.L., Sudar, D., Knowles, D.W., Malik, J., and Biggin, M.D. (2006). Three-dimensional morphology and gene expression in the Drosophila blastoderm at cellular resolution II: dynamics. Genome Biol. *7*, R124.

Ko, M.S.H. (1991). A stochastic model for gene induction. J. Theor. Biol. *153*, 181–194.

Kulkarni, M.M., and Arnosti, D.N. (2005). cis-regulatory logic of short-range transcriptional repression in Drosophila melanogaster. Mol. Cell. Biol. *25*, 3411–3420.

Lammers, N., Galstyan, V., Reimer, A., Meddin, S.A., Wiggins, C.H., and Garcia, H.G. Binary transcriptional control of pattern formation in development.

Lawrence, P.A. (1992). The Making of a Fly: The Genetics of Animal Design (Wiley-Blackwell).

Li, L.M., and Arnosti, D.N. (2011). Long- and short-range transcriptional repressors induce distinct chromatin states on repressed genes. Curr. Biol. *21*, 406–412.

Little, S.C., Tikhonov, M., and Gregor, T. (2013). Precise Developmental Gene Expression Arises from Globally Stochastic Transcriptional Activity. Cell *154*, 789–800.

Macdonald, P.M., Ingham, P., and Struhl, G. (1986). Isolation, structure, and expression of even-skipped: a second pair-rule gene of Drosophila containing a homeo box. Cell *47*, 721–734.

McKnight, S.L., and Miller, O.L., Jr (1977). Electron microscopic analysis of chromatin replication in the cellular blastoderm Drosophila melanogaster embryo. Cell *12*, 795–804.

Nüsslein-Volhard, C., and Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. Nature *287*, 795–801.

Nüsslein-Volhard, C., Kluding, H., and Jürgens, G. (1985). Genes affecting the segmental subdivision of the Drosophila embryo. Cold Spring Harb. Symp. Quant. Biol. *50*, 145–154.

Payankaulam, S., and Arnosti, D.N. (2009). Groucho corepressor functions as a cofactor for the Knirps short-range transcriptional repressor. Proc. Natl. Acad. Sci. U. S. A. *106*, 17314–17319.

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

Senecal, A., Munsky, B., Proux, F., Ly, N., Braye, F.E., Zimmer, C., Mueller, F., and Darzacq, X. (2014). Transcription factors modulate c-Fos transcriptional bursts. Cell Rep. *8*, 75–83.

Small, S., Kraut, R., Hoey, T., Warrior, R., and Levine, M. (1991). Transcriptional regulation of a pair-rule stripe in Drosophila. Genes Dev. *5*, 827–839.

Small, S., Blair, A., and Levine, M. (1992). Regulation of even-skipped stripe 2 in the Drosophila embryo. EMBO J. *11*, 4047–4057.

So, L.-H., Ghosh, A., Zong, C., Sepúlveda, L.A., Segev, R., and Golding, I. (2011). General properties of transcriptional time series in Escherichia coli. Nat. Genet. *43*, 554–560.

Surkova, S., Kosman, D., Kozlov, K., Manu, Myasnikova, E., Samsonova, A.A., Spirov, A., Vanario-Alonso, C.E., Samsonova, M., and Reinitz, J. (2008). Characterization of the Drosophila segment determination morphome. Dev. Biol. *313*, 844–862.

Tostevin, F., de Ronde, W., and ten Wolde, P.R. (2012). Reliability of frequency and amplitude decoding in gene regulation. Phys. Rev. Lett. *108*, 108104.

Venken, K.J.T., He, Y., Hoskins, R.A., and Bellen, H.J. (2006). P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. Science *314*, 1747–1751.

Venken, K.J.T., Carlson, J.W., Schulze, K.L., Pan, H., He, Y., Spokony, R., Wan, K.H., Koriabine, M., de Jong, P.J., White, K.P., et al. (2009). Versatile P[acman] BAC libraries for transgenesis studies in Drosophila melanogaster. Nat. Methods *6*, 431–434.

Warming, S., Costantino, N., Court, D.L., Jenkins, N.A., and Copeland, N.G. (2005). Simple and highly efficient BAC recombineering using galK selection. Nucleic Acids Res. *33*, e36.

Xu, H., Sepúlveda, L.A., Figard, L., Sokac, A.M., and Golding, I. (2015). Combining protein and mRNA quantification to decipher transcriptional regulation. Nat. Methods *12*, 739–742.

Xu, H., Skinner, S.O., Sokac, A.M., and Golding, I. (2016). Stochastic Kinetics of Nascent RNA. Phys. Rev. Lett. *117*.

Zenklusen, D., Larson, D.R., and Singer, R.H. (2008). Single-RNA counting reveals alternative modes of gene expression in yeast. Nat. Struct. Mol. Biol. *15*, 1263–1271.

Zoller, B., Little, S.C., and Gregor, T. (2017). Diverse spatial expression patterns emerge from common transcription bursting kinetics.

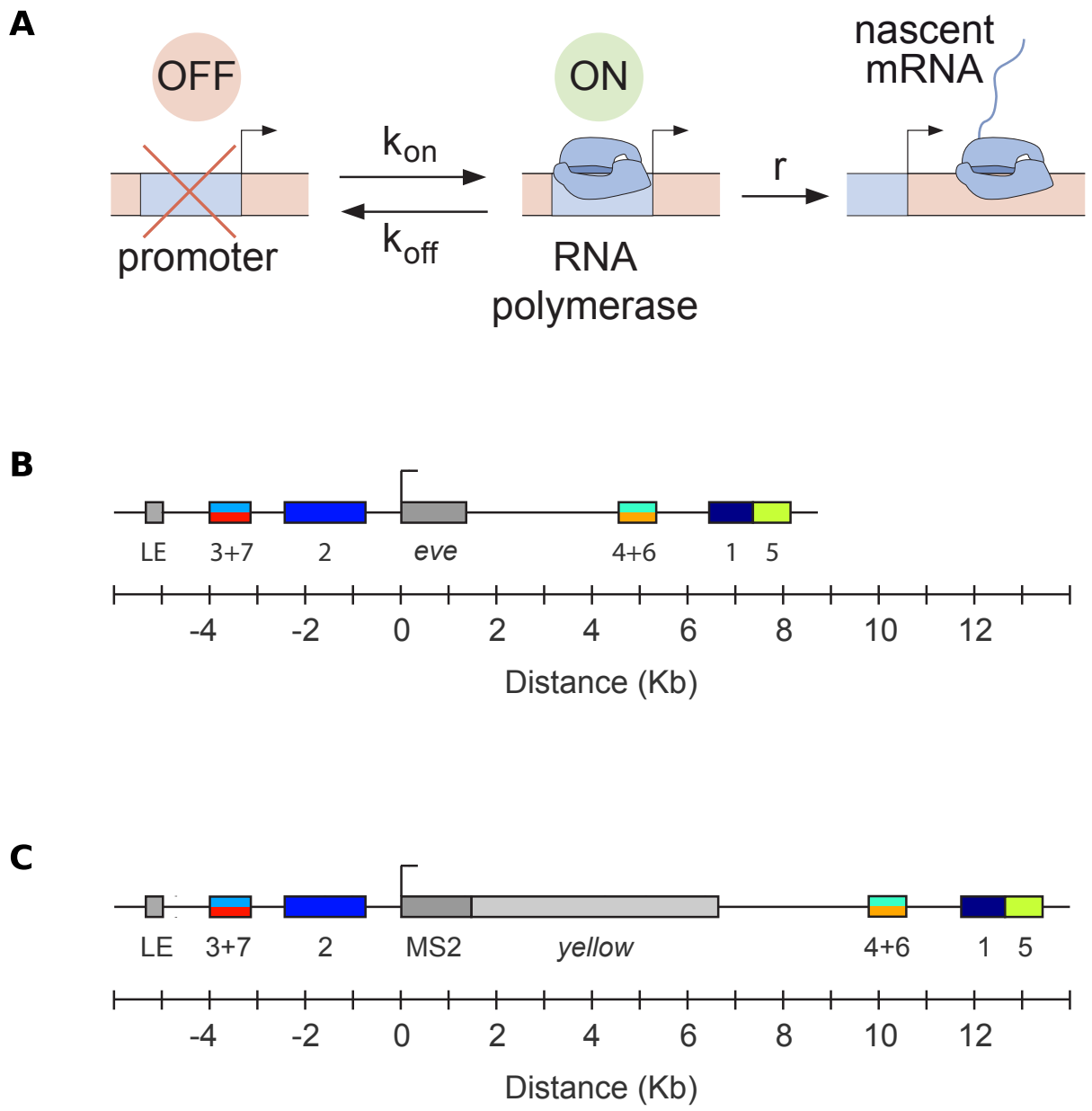

Figure 1. Visualizing live transcription from the seven stripes of D. melanogaster even-skipped.

A) Simple model of transcriptional bursting by promoter switching between ON and OFF states. B) Wild-type eve locus showing the five stripe enhancers (1,2,3+7,4+6,5) and the late enhancer (LE). C) Layout of the engineered eve BAC showing the location of the MS2 array and yellow gene.

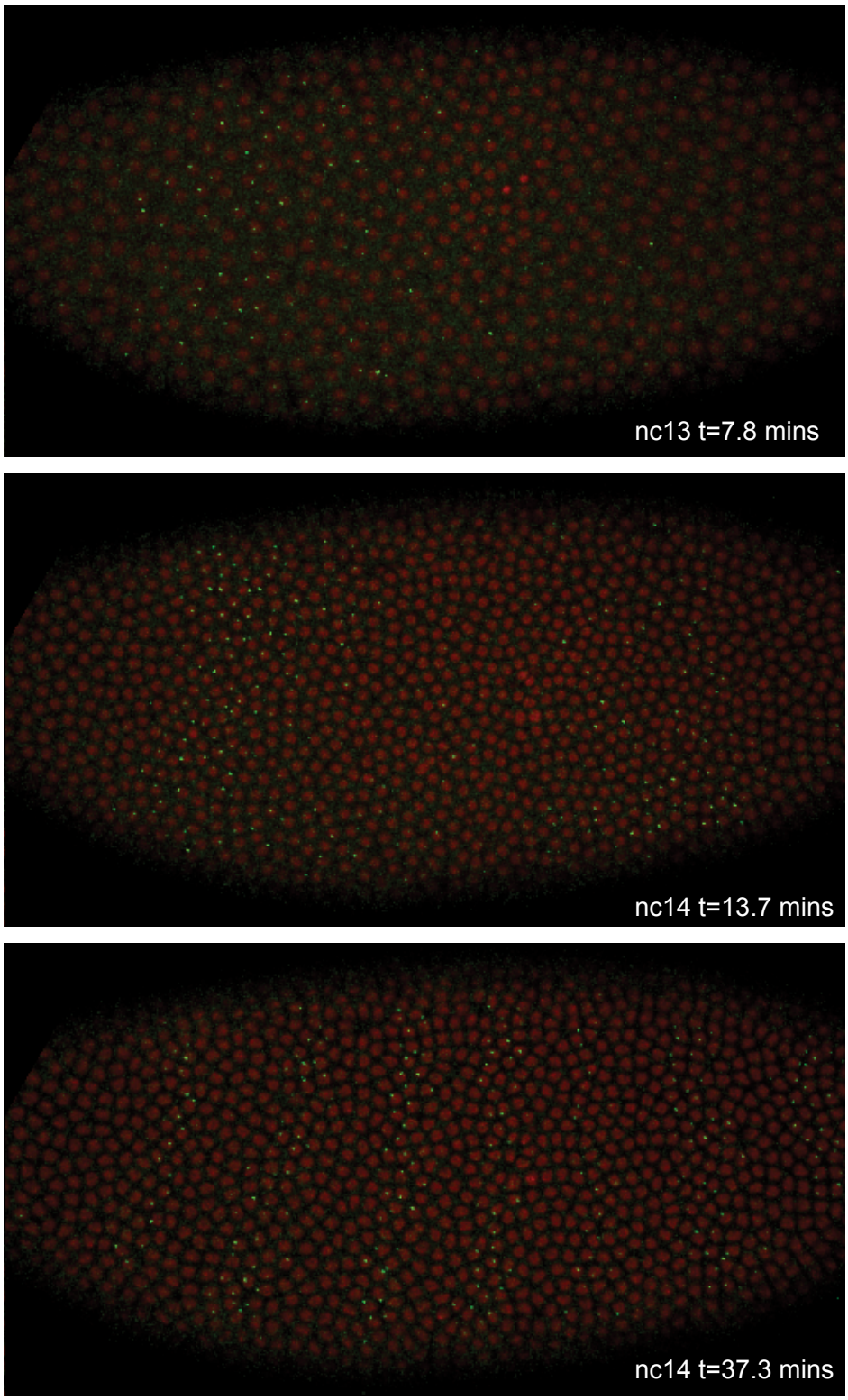

Figure 2. Live expression of even-skipped

Stills from maximum projection renderings of image stacks of an embryo spanning all seven stripes (link to full movie). This movie was collected at 40x resolution for illustration purposes only. Movies used for data analysis were collected at higher resolution as described in the text.

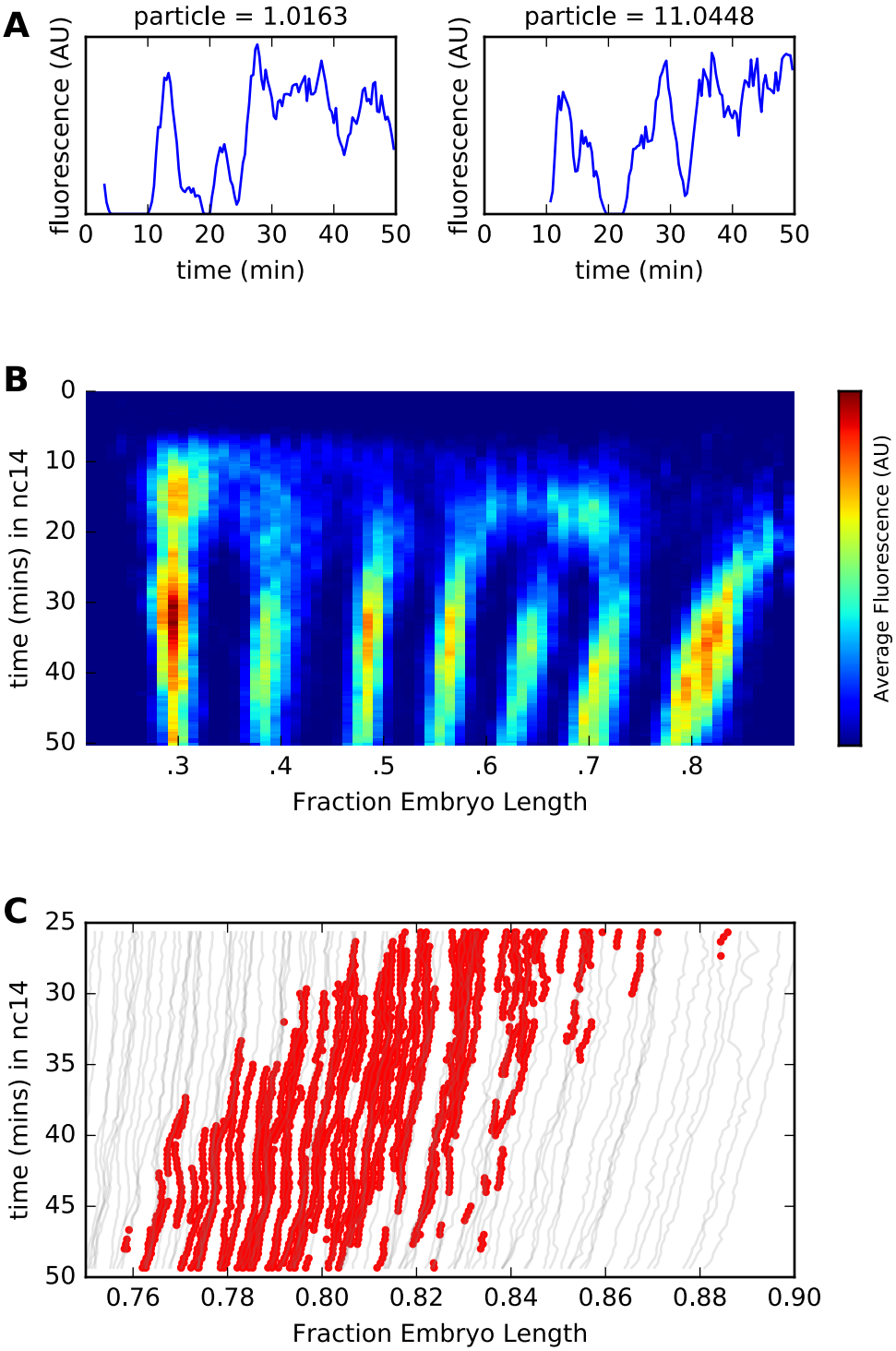

Figure 3. Spatiotemporal dynamics of *eve* expression

A) Fluorescence traces from three representative nuclei. B) Average fluorescence over space and time showing stripe formation, modulation and movement. The time resolution along the y-axis is 20s. The positions of nuclei along the x-axis were registered across movies based on the inferred position of stripe centers, and placed into bins of one percent embryo length, with the average fluorescence of all nuclei in each bin plotted. C) Traces of nuclei positions over time (gray lines) from stripe 7 region of movie EVE_D6 with timepoints where fluorescence is greater than 50,000 (the expected output for a promoter in the ON state for one 20s time interval) to illustrate stripe movement relative to nuclei.

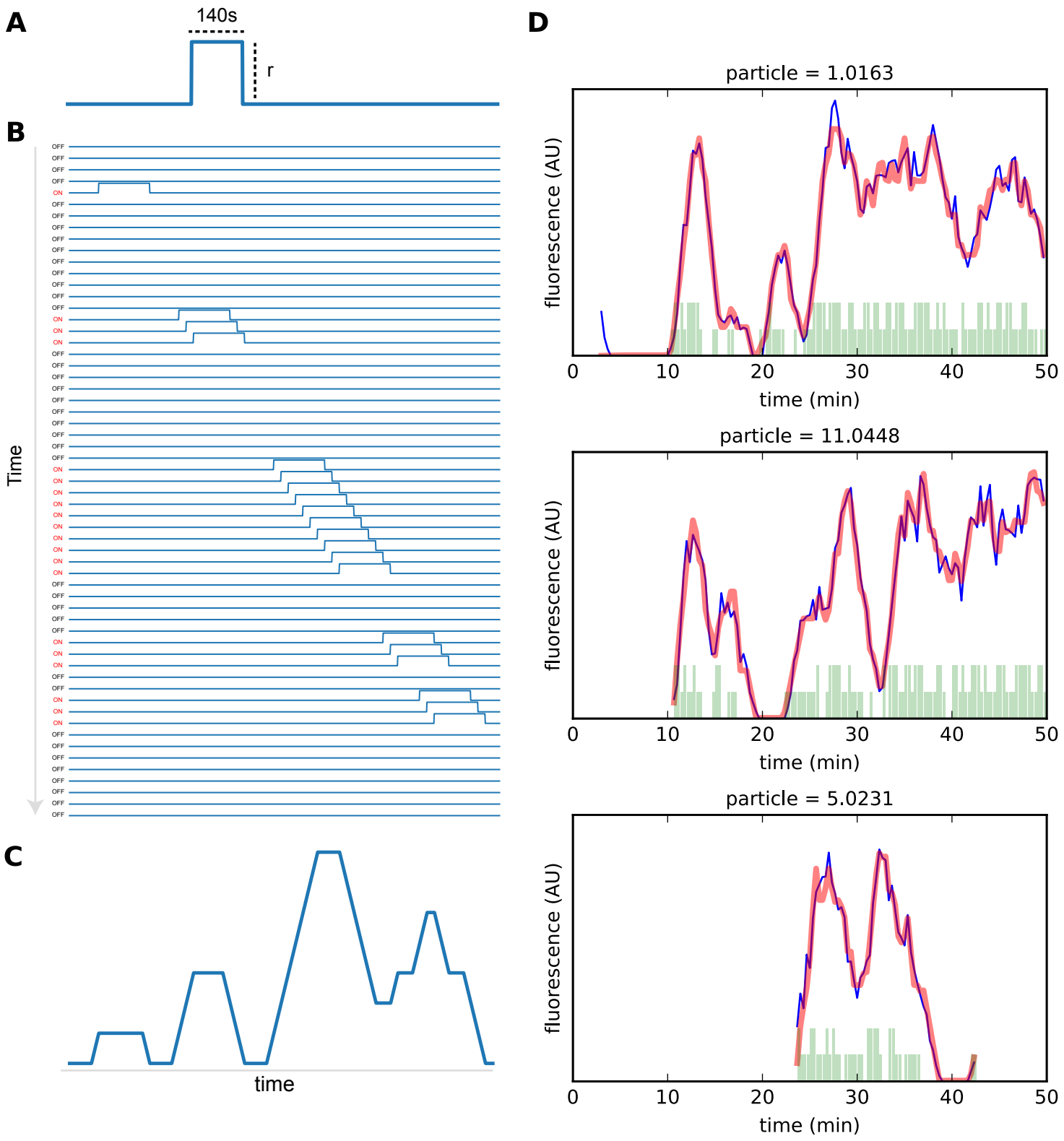

Figure 4. Modeling bursting in individual nuclei.

A) Fluorescence from a single time interval when promoter is in ON state. B) Representation of bursting model showing bursts of fluorescence produced when promoter alternates between ON to OFF states. C) Aggregate pattern from bursts shown in (B) to represent how complex patterns of fluorescence emerge from a simple underlying bursting process. D) Representative fluorescence traces from individual nuclei (blue lines), inferred bursting pattern (green bar means the promoter is ON; height represented number of engaged promoters) and fluorescence imputed by model (red line).

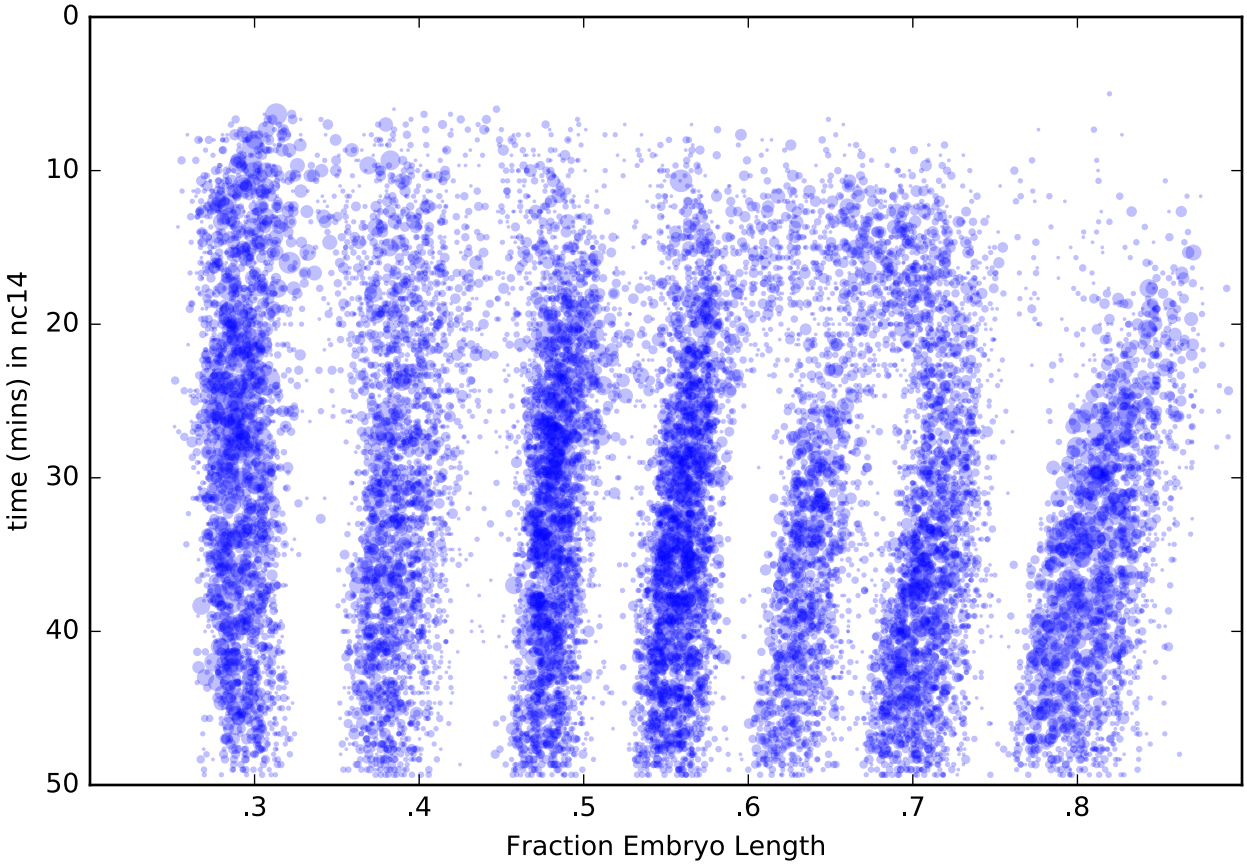

Figure 5. The kinetic fingerprint of *even-skipped* stripe formation

We inferred the location of every transcriptional burst in all 11 movies and plot here when (y-axis) and where along the anterio-posterior axis (plotted as fraction of embryo length) each burst occured. The size of the dot represents the duration of the burst. Collectively the data create a kinetic fingerprint of *eve* stripe formation.

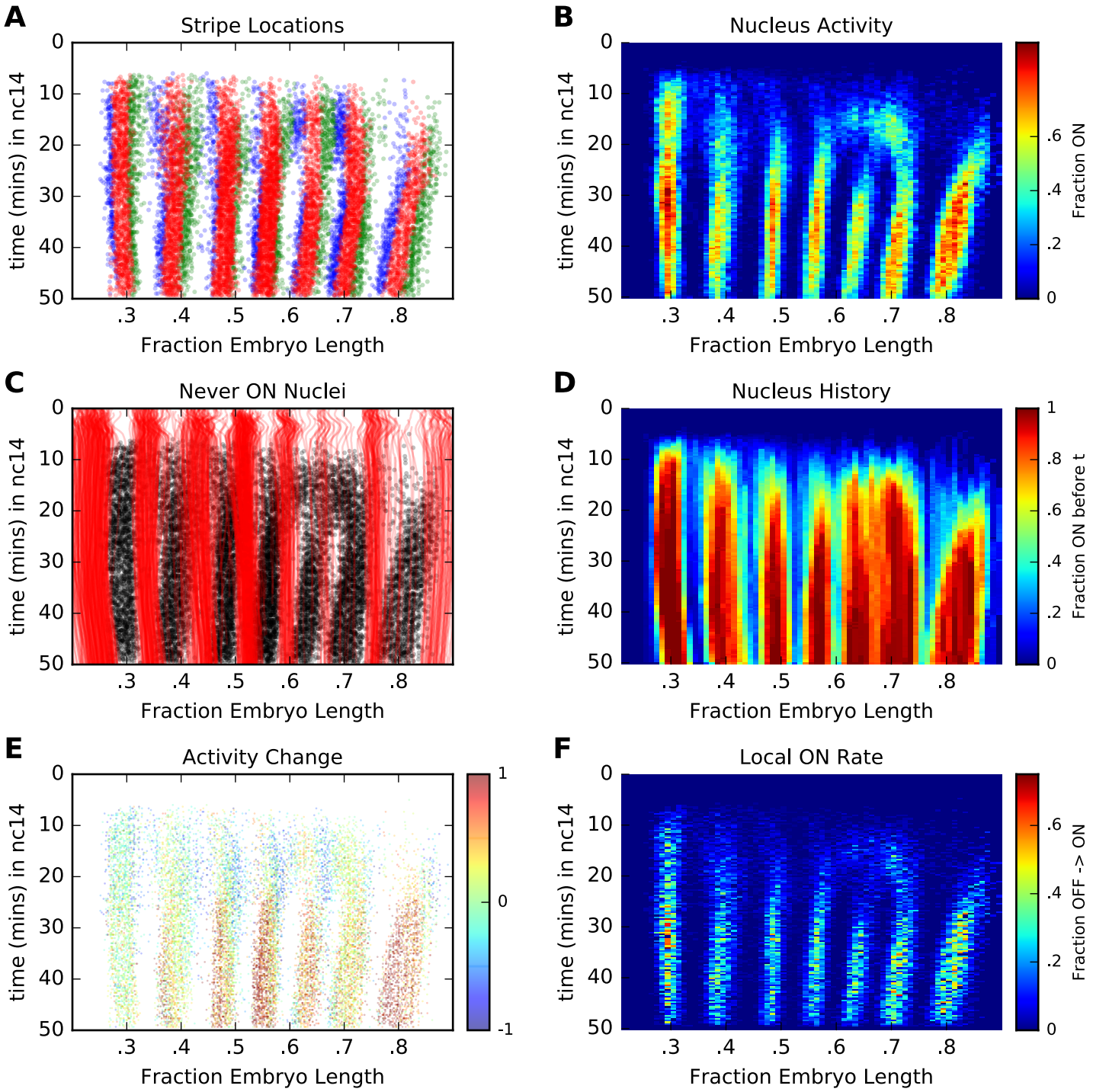

Figure 6. Stripes are formed from transcriptional deserts

A) Inferred positions of stripes (red) and stripe margins (ant. blue, post. green). Only locations of new transcriptional bursts are plotted. B) Mean activity (fraction of nuclei in an active state) over space and time. C) Locations of new bursts (black dots) in space and time along with spatiotemporal traces of nuclei that are in the OFF state throughout nc14 (red lines). D) Fraction of nuclei that have been on prior to given time (y-axis). E) Locations of new bursts colored by the change in activity of that nucleus from 0-25 minutes to 25-50 minutes into nc14. F) Local OFF to ON switching probability, computed by dividing the number of new bursts initiated in a given bin by the number of nuclei in OFF state.

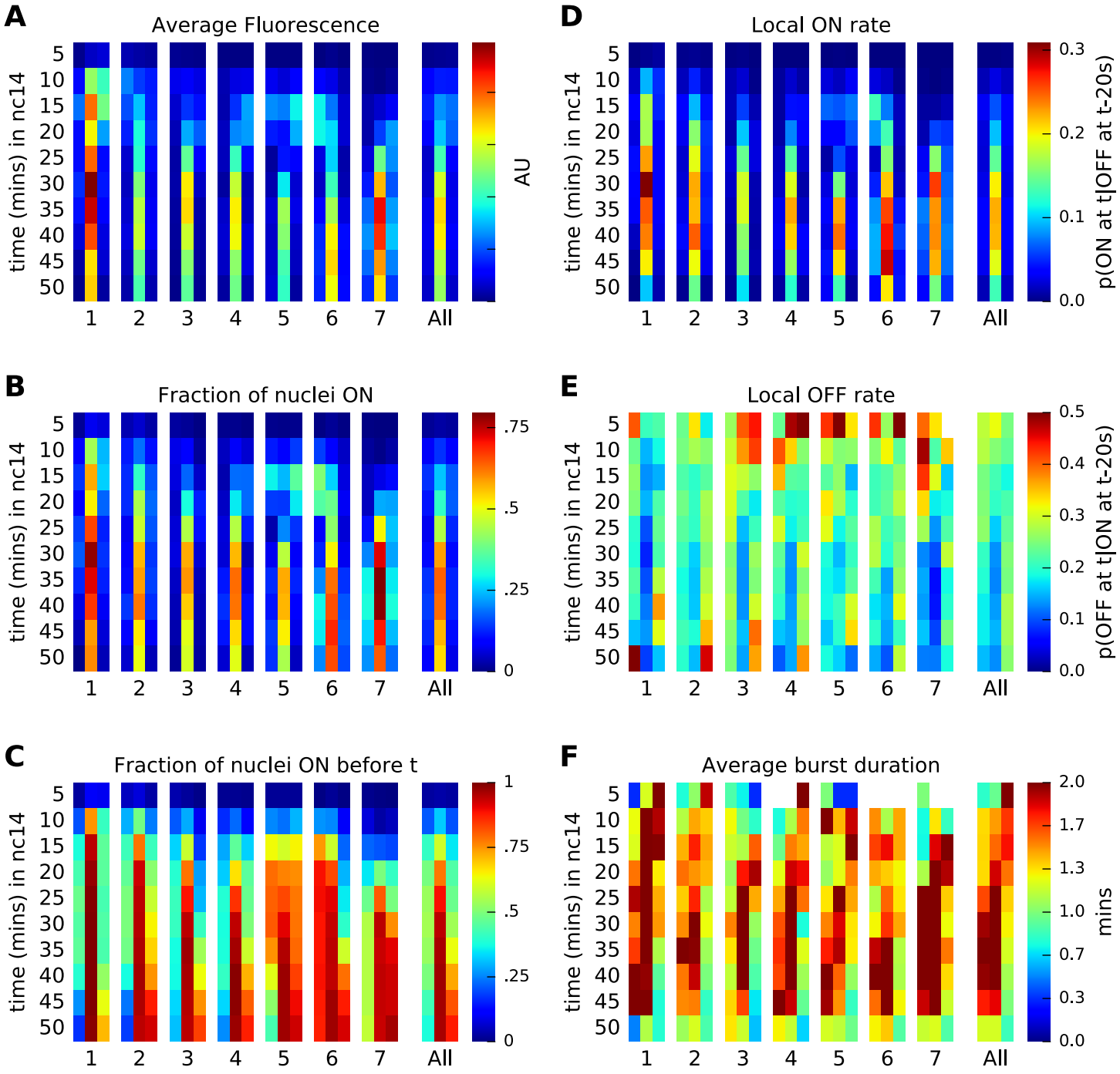

Figure 7. Variation in bursting parameters across time in stripes and stripe boundaries

Individual time points were binned based on stripe assignment and time (5 minute bins) and then various values were computed on these nuclei to highlight changes inside and outside of stripes over time. In each subpanel the left column is nuclei in the anterior flank of the stripe, the central column are stripe nuclei, and the right column are nuclei in the posterior flank. A) Average fluorescence. B) Activity (fraction of nuclei in ON state). C) Prior activity (fraction of nuclei that have been ON given time point or earlier). D) Local OFF to ON rate (fraction of nuclei in OFF state that switch to ON state). E) Local ON to OFF rate (fraction of nuclei in ON state that switch to OFF state). F) Average burst duration

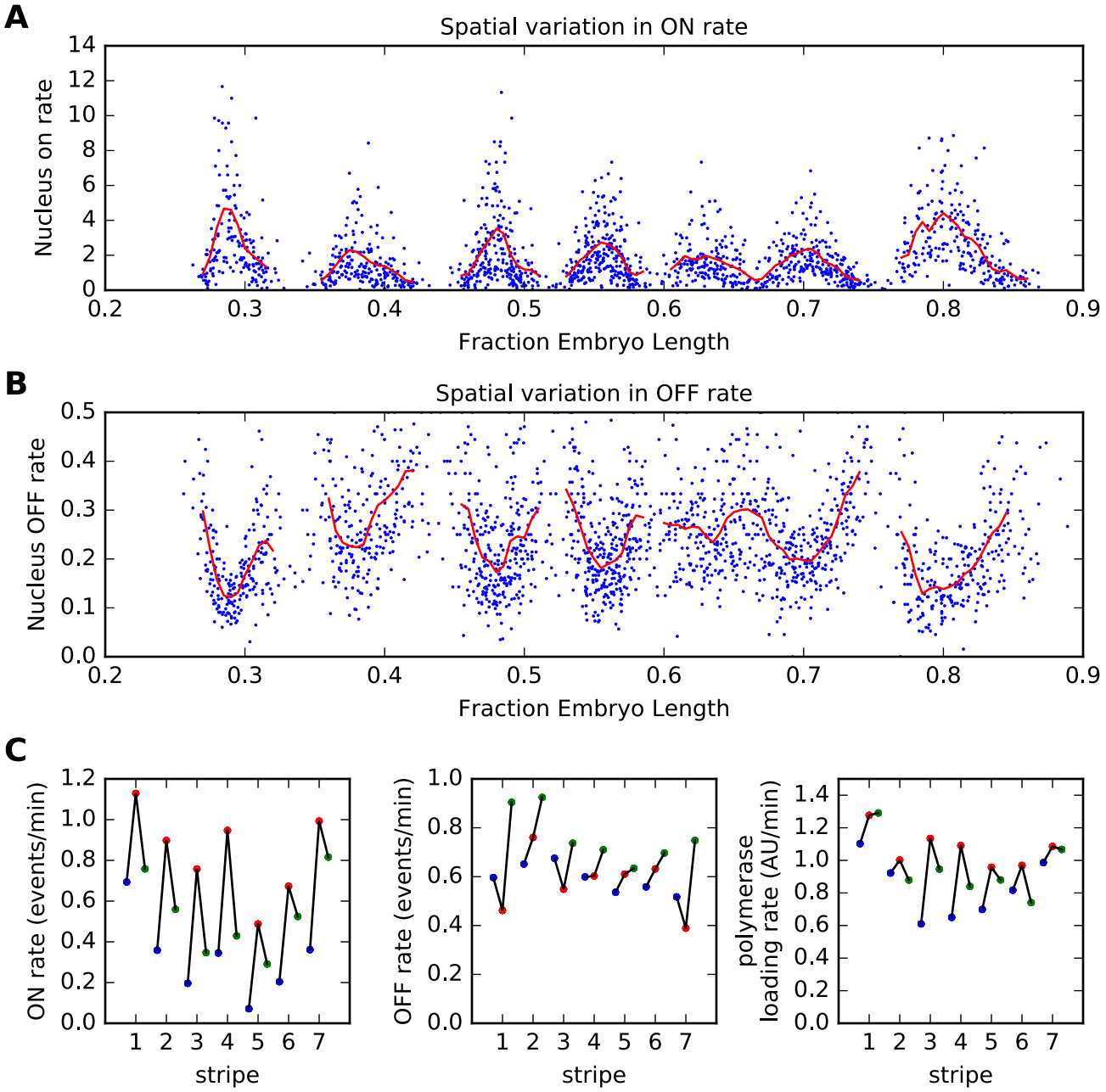

Figure 8. Variation in bursting parameters across stripe

A) Nucleus-specific ON rate for all nuclei (blue dots;). ON rate is number of OFF->ON transitions from mins 25-50 divided by number of time points in OFF state between first and last ON time point. Red lines are running averages in windows of .005 fraction of AP; breaks are points with fewer than 10 active nuclei. B) Nucleus-specific OFF rate for all nuclei (blue dots;). OFF rate is number of ON->OFF transitions from mins 25-50 divided by number of time points in ON state between first and last ON time point. Red lines are running averages in windows of .005 fraction of AP; breaks are points with fewer than 10 active nuclei. C) Stripe specific mHMM parameter inferences. Each triplet of points represents one stripe with the values for the posterior flank, stripe center, and anterior flank shown in blue, red and green.

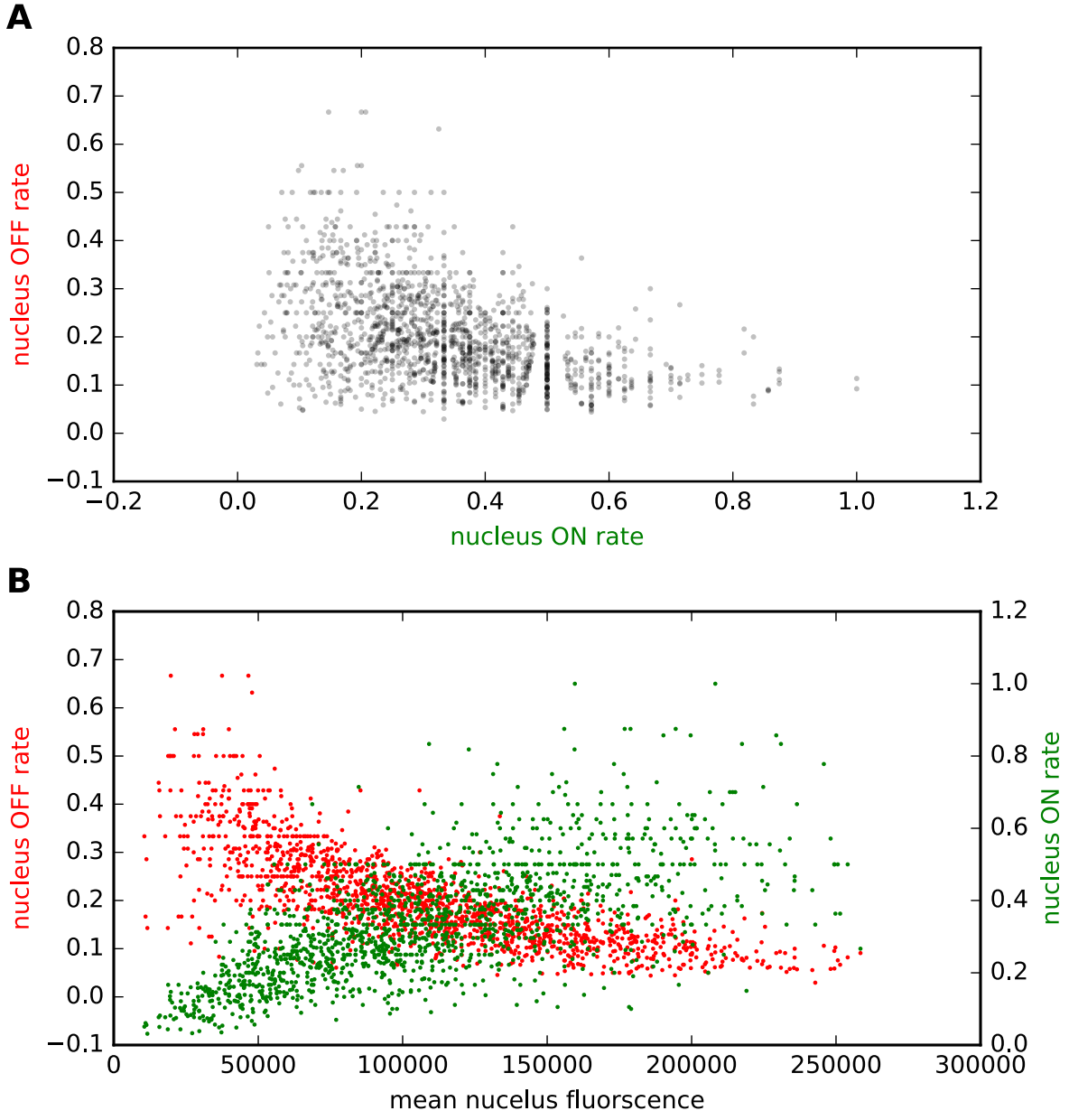

Figure 8-S1. Empirical on and off rates tightly corrlated with each other and fluorescence

We calculated empirical mean fluorescence, on rate (number of OFF to ON events divided by number of time points in OFF state) and off rate (number of ON to OFF events divided by number of time points in ON state) for all nuclei with at least one new burst between 25 and 50 mins in nc14. Only the period between active bursts was considered. For most nuclei this is the entire period from 25 mins through the end of their last burst. However for nuclei at stripe borders or in interstripe regions this is only a the non-quiescent portion of this time interval. We then compared these empircal ON and OFF rates to each other on a nucleus-by-nucleus level (A) or each individual to fluorescence (B), revealing high levels of correlation between these paramters.

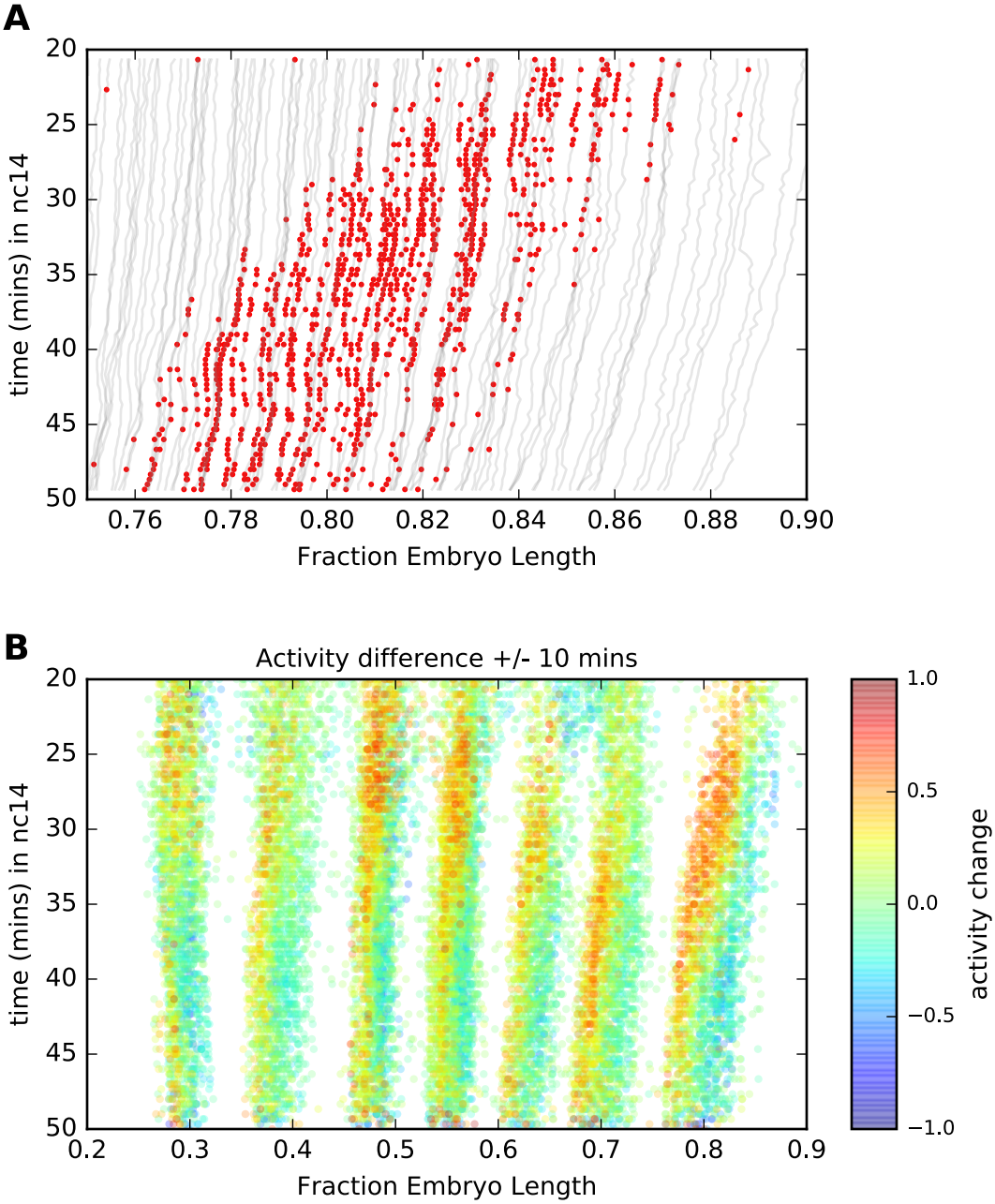

Figure 9. Stripe movement is driven by changes in nucleus state

A) Traces of nuclei positions over time (gray lines) from stripe 7 region of movie EVE_D6 with timepoints where new bursts initiated colored red to illustrate stripe movement relative to nuclei. B) Location of new bursts in space and time colored by the difference in their mean activity (fraction of timepoints in on state) in 10m after and 10m before the burst. Positive values represent a nucleus turning on or increasing activity, while blue values indicate a nucleus turning off or decreasing activity.

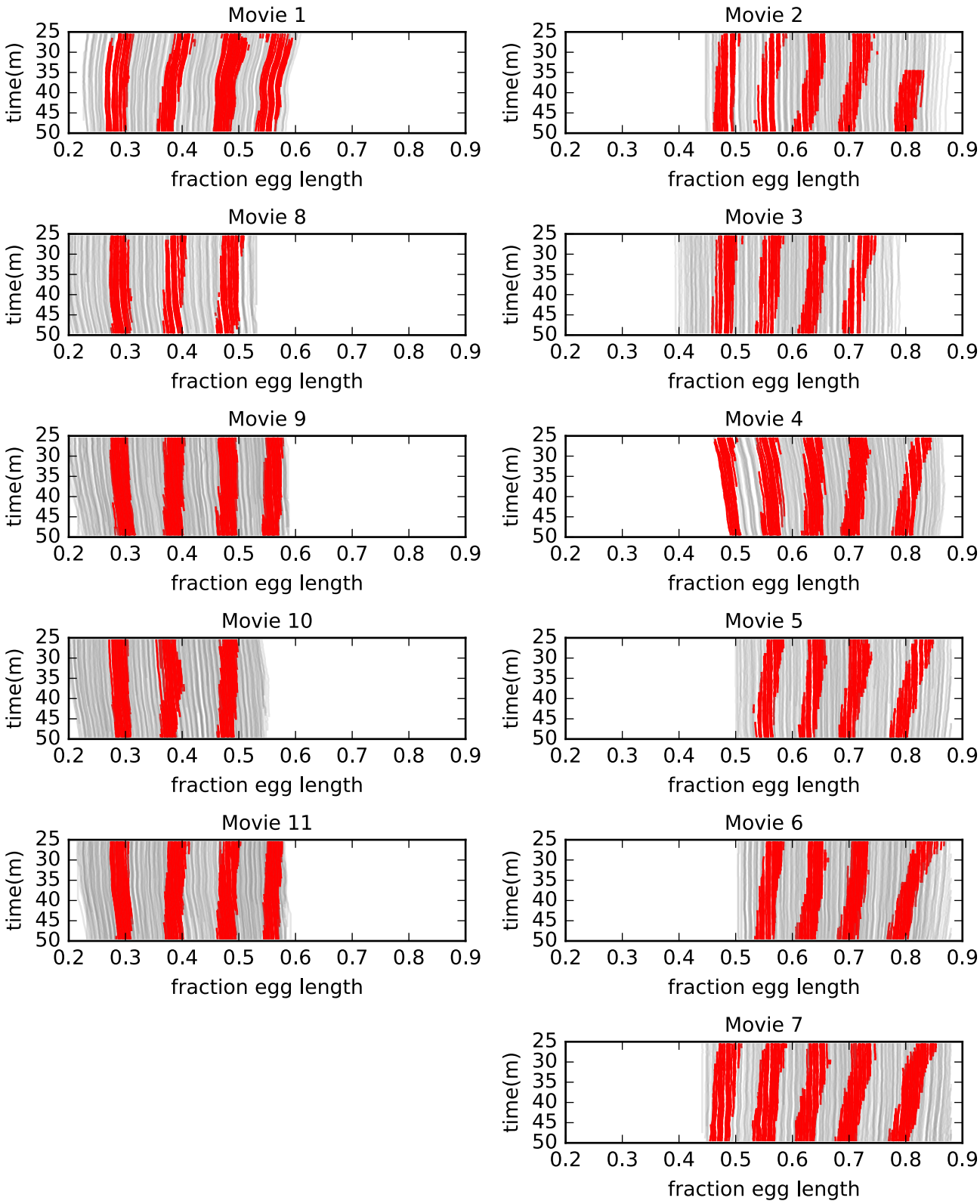

Figure 9-S1. Stripe movement

Traces of nuclei positions over time (gray lines) from all eleven movies with time points annotated as in stripes colored red. We plotted the nucleus-specific ON rate (number of OFF to ON events divided by the number of time points in OFF state) as a function of mean nucleus AP position for every nucleus across either all time points after 25 mins in nc14 (A) or timepoints between their first and last post 25 minute time point in the ON state. The red lines represent the smoothed mean.