

# Analysis of developmental gene expression using smFISH and *in silico* staging of *C. elegans* embryos

Laura Breimann<sup>1,2,3,#</sup>, Ella Bahry<sup>1,4</sup>, Marwan Zouinkhi<sup>5,§</sup>, Klim Kolyvanov<sup>1</sup>, Lena Annika Street<sup>2</sup>,  
Stephan Preibisch<sup>5,§,#</sup> and Sevinç Ercan<sup>2,#</sup>

<sup>1</sup>Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine, Berlin, Germany

<sup>2</sup>Department of Biology, Center for Genomics and Systems Biology, New York University, New York, NY, USA

<sup>3</sup>Department of Genetics, Harvard Medical School, Boston, MA, USA

<sup>4</sup>Helmholtz Imaging, Max-Delbrück-Center for Molecular Medicine (MDC), Berlin, Germany

<sup>5</sup>Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA, USA

§ Majority of the work was performed at MDC/BIMSB

# Corresponding authors

Keywords: dosage compensation, smFISH, *C. elegans*, X chromosome, sex determination, embryogenesis, condensin, DCC, transcription burst, transcription frequency, gene regulation, development

Contact: [se71@nyu.edu](mailto:se71@nyu.edu), [preibischs@janelia.hhmi.org](mailto:preibischs@janelia.hhmi.org), [Laura\\_Breimann@hms.harvard.edu](mailto:Laura_Breimann@hms.harvard.edu)

## Abstract

Regulation of transcription during embryogenesis is key to development and differentiation. To study transcript expression throughout *Caenorhabditis elegans* embryogenesis at single-molecule resolution, we developed a high-throughput single-molecule fluorescence in situ hybridization (smFISH) method that relies on computational methods to developmentally stage embryos and quantify individual mRNA molecules in single embryos. We applied our system to *sdc-2*, a zygotically transcribed gene essential for hermaphrodite development and dosage compensation. We found that *sdc-2* is rapidly activated during early embryogenesis by increasing both the number of mRNAs produced per transcription site and the frequency of sites engaged in transcription. Knockdown of *sdc-2* and *dpy-27*, a subunit of the dosage compensation complex (DCC), increased the number of active transcription sites for the X chromosomal gene *dpy-23* but not the autosomal gene *mdh-1*, suggesting that the DCC reduces the frequency of *dpy-23* transcription. The temporal resolution from *in silico* staging of embryos showed that the deletion of a single DCC recruitment element near the *dpy-23* gene causes higher *dpy-23* mRNA expression after the start of dosage compensation, which could not be resolved using mRNAseq from mixed-stage embryos. In summary, we have established a computational approach to quantify temporal regulation of transcription throughout *C. elegans* embryogenesis and demonstrated its potential to provide new insights into developmental gene regulation.

## Introduction

Temporal and spatial regulation of transcription plays a crucial role in defining cellular identities and maintaining proper function in multicellular organisms. Studies using single-molecule fluorescent in-situ hybridization (smFISH) have quantified mRNA molecules and revealed that transcription is a highly stochastic process that occurs in discrete bursts over time [1–4]. Simple models of transcription include two key parameters: burst size, which describes the number of RNAs transcribed from a single promoter during an active period, and burst frequency, which quantifies the number of transcription bursts occurring over a given timeframe [1,5,6].

In embryogenesis, zygotic transcription controls early processes, including sex determination, dosage compensation, and body patterning [7]. In *Drosophila* embryos, developmental genes were shown to be subject to burst frequency modulation [2,8–10]. Studies including different organisms hypothesize that burst size is a property of promoters but was also shown to be controlled by transcription factors and chromatin regulatory mechanisms [11–17].

Similar to other organisms, transcriptional bursting has been observed in *Caenorhabditis elegans* [18,19]. However, the mechanisms that control the parameters of embryonic transcription have not been addressed. In *C. elegans*, zygotic transcription starts around the 4-cell stage [20–23]. Sex is determined shortly thereafter, and in XX animals, leads to zygotic transcription of *sdc-2*, which is essential for hermaphrodite development and X chromosome dosage compensation [24].

SDC-2 binds to the X chromosomes around the ~40-cell stage and recruits the dosage compensation complex (DCC), which is composed of a specialized condensin I (hereafter DC) and several non-condensin proteins [25,26]. DCC represses X chromosome transcription ~2-fold in XX hermaphrodites to equalize overall output from that of XO males [27,28]. Embryos defective for dosage compensation die late in embryogenesis, suggesting that this is a critical developmental time for proper gene dosage [29].

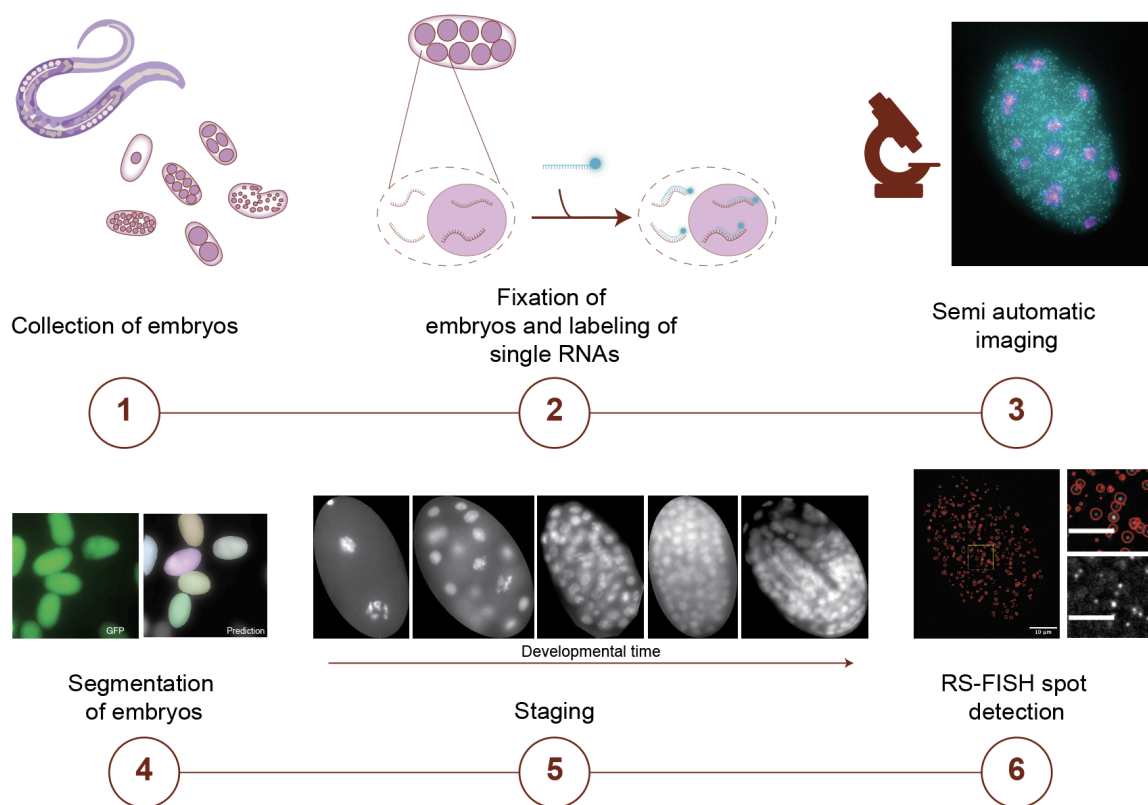
*C. elegans* is an excellent model for studying developmental gene regulation due to defined cell numbers and constant (stereotypic) lineage [30]. However, a major difficulty of performing smFISH at high throughput across embryogenesis is that there are no current methods to synchronize *C. elegans* embryos at specific stages. Accordingly, previous studies used smFISH in *C. elegans* to detect transcription qualitatively [31,32] with low throughput [19,33] or in early embryo stages [34,35].

To overcome the difficulty of obtaining smFISH measurements from distinct embryonic stages, we developed a computational approach to *in silico* stage embryos stained by smFISH probes and a nuclear marker (DAPI) acquired at random time points throughout embryogenesis. Our method suggests that the rapid zygotic activation of *sdc-2* is accomplished by increasing both the size and frequency of transcription bursts in early embryogenesis. We also quantified the effect of *sdc-2* and *dpy-27* knockdown on *dpy-23*, a dosage compensated X chromosomal gene, and used simultaneous staining for an autosomal gene *mdh-1* as an internal control. Our results show that the DCC represses *dpy-23* by reducing the frequency of active transcription sites per embryo. We compared the effect of a single DCC recruitment deletion using mRNA-seq from mixed-stage embryos versus our method and found that *in silico* staging coupled with high-throughput smFISH detects small temporal changes in gene expression during embryogenesis.

## Results

To obtain samples spanning all embryonic stages, we collected younger embryos by bleaching *C. elegans* adult hermaphrodites and older embryos by washing them off plates (Figure 1). We labeled single RNA molecules using smFISH probes targeting exonic or intronic regions of several genes on the X chromosome and autosomes (Supplemental Figure 1). We then semi-automatically imaged over 9000 smFISH-labeled embryos in 3D.

Figure 1



**Figure 1.** The experimental workflow. (1) Embryos are collected by bleaching adults and washing off older embryos. (2) Collected embryos are fixed, permeabilized, and labeled with smFISH probes. (3) The labeled embryos are imaged using a semi-automatic imaging setup. The analysis includes segmentation of single embryos (4), staging of embryos according to their developmental time (5), and detection of single RNAs using RS-FISH (6).

### Computational staging of *C. elegans* embryos and smFISH analysis

Imaging transcription using smFISH allows for additional visualization of nuclei by staining DNA with DAPI. Given that *C. elegans* follows a stereotypical developmental pattern, the number of nuclei directly correlates with its developmental stage [36,37]. While it is theoretically feasible to discern each embryonic stage ranging from a single nucleus to 558 nuclei at the conclusion of embryogenesis, such an undertaking would be exceedingly challenging for a high-throughput approach [38–40] or require staining of additional markers [41]. Therefore, we employ an autoencoder-based image classifier to learn representations from fixed, DAPI-stained *C. elegans* embryos to automatically stage them in an approach similar to [42]. We rely on a manually created training dataset for each developmental stage to achieve reliable predictions from the learned representations. Our training dataset was generated by categorizing embryos into different developmental stages.

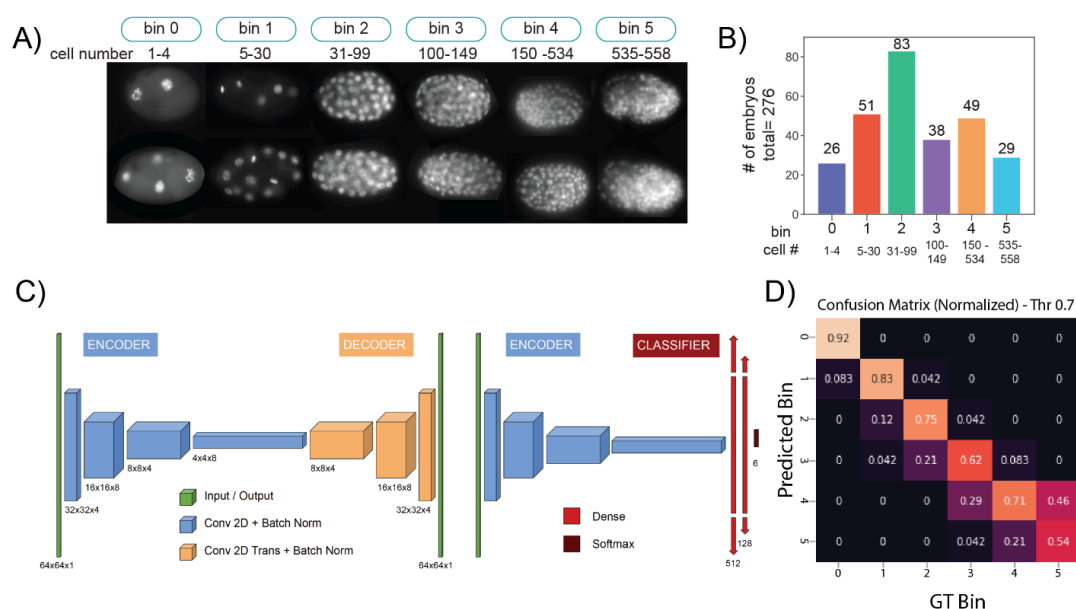
The categories reflected the following stages of gene regulation and development: Bin 0 contains 1-4 cell stage embryos, which is before zygotic genome activation (ZGA) for most genes, thus detecting maternal mRNA load [21,22]. Bin 1 includes embryos with 5 - 30 nuclei, which start zygotic transcription, undergo sex determination, and begin to express SDC-2 [24,43]. Bin 2 includes embryos with 31-99 nuclei when DCC starts localizing to the X chromosomes [24,44,45]. Bin 3 includes embryos from 100 - 149 nuclei when H4K20me1 on the X chromosomes increases compared to autosomes [46–48]. Bin 4 includes embryos with nuclei from 150 - 534, where the nuclei are dense, and counting is especially difficult. In bin 5 (535 - 558), cells hardly divide, but embryos undergo

several detectable morphological changes. We annotated at least 26 embryos per stage in the training set (for availability, see the methods section) (Figure 2B).

For automatic stage prediction, we adapted an autoencoder-based architecture (Figure 2C) [49,50]. An autoencoder learns representative features of an input dataset, called the latent space, from which it tries to reconstruct the input dataset. Thus, it can be used as a non-linear dimensionality reduction method as the latent space captures the essential features that describe the data. First, to train the autoencoder, the entire dataset was utilized in a self-supervised manner, with the autoencoder learning to reconstruct the input data as its output, while 250 embryo images were withheld for validation. Next, we discarded the decoder and attached two fully connected layers directly to the encoder's latent space output, forming an additional classifier arm designed to predict embryo stages. The classifier was initially trained by itself using the annotated training dataset, with the encoder weights frozen later, the entire network was trained end-to-end. Our approach provided sufficient prediction accuracy of the different embryonic stages (Figure 2D). Due to the dense packing of nuclei, early embryonic stages are more accurately predicted than older stages, which also holds true for the manual staging of embryos. While our approach has some uncertainty, especially for older embryonic stages, we decided that the ability to stage thousands of embryos automatically using only a small training set in a short time outweighs the benefits of the entirely (also uncertain) manual approach.

To quantify transcription in the *in silico* staged embryos, we implemented a pipeline for image analyses (Suppl. Fig. 2). Briefly, the raw smFISH signal was enhanced using a Median filter to reduce background noise. The individual transcripts were detected in 3D using the RS-FISH plugin [51]. The transcripts were filtered using a binary mask to exclude smFISH detections from neighboring embryos. Since imaging deeper into the tissue of embryos leads to loss of signal due to bleaching and scattering of light, we corrected the signal in Z by fitting a quadratic function to the detected spots. To compare transcripts from different embryos and conditions, we normalized the detections from each embryo by fitting a gamma function, where the maximum is set to 1.

Figure 2



**Figure 2 Prediction of embryonic stages using an autoencoder.**

**A)** Grouping all developmental stages into six bins with the included nucleus count labeled below each bin and additional example images of DAPI-stained embryos. **B)** Count of the annotated example images of each bin. The bin with the lowest count was used as a cut-off for all bins to have equal-sized training data. **C)** Illustration of the autoencoder architecture that was used for the classification of the stage bin. **D)** Confusion matrix for the predicted bins against a set of ground truth (GT) bins using a class-stratified leave-one-out validation. Values closer to 1 represent better predictions for the bin.

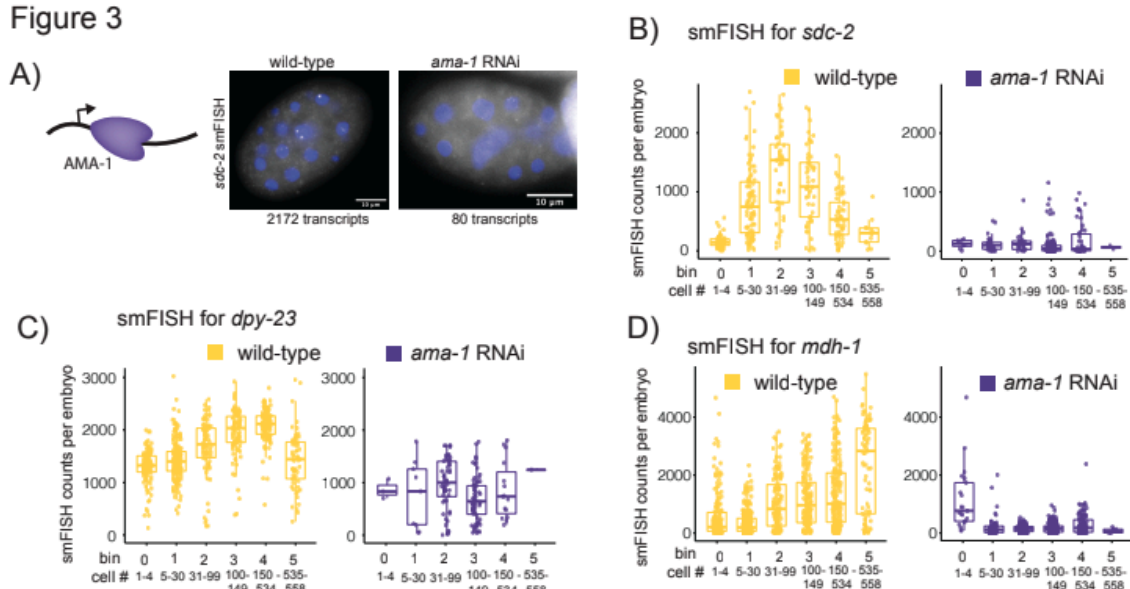
## RNA Polymerase II knockdown causes distinct effects on the mRNA levels of different genes

To validate our analysis of transcription in embryos, we depleted the large subunit of RNA Pol II (AMA-1) by feeding bacteria expressing double-stranded RNAs for interference (RNAi). Since *ama-1* RNAi leads to germline formation defects, we used a relatively short exposure to allow embryo development. This treatment results in partial knockdown of mRNAs (Figure 3A).

The *sdc-2* gene is zygotically transcribed downstream of the sex determination pathway [24,43]. Accordingly, smFISH analysis shows few transcripts in the 1-4 cell stage (Figure 3B, left panel). *sdc-2* is strongly activated at the 5-30 nuclei time-point, in agreement with its functions in early embryogenesis. *Ama-1* RNAi effectively prevents the activation of *sdc-2* transcription through embryogenesis (Figure 3B, right panel).

In contrast to the zygotically transcribed *sdc-2*, the maternally loaded *dpy-23* transcripts were reduced but were not eliminated by *ama-1* RNAi (Figure 3C). Interestingly, another maternally loaded mRNA, *mdh-1*, was present in 1-4 stage embryos but strongly reduced after the 4-cell stage, suggesting that the *mdh-1* mRNA has lower stability compared to *dpy-23* mRNA (Figure 3D). The distinct effects of *ama-1* knockdown on the three genes validate our methods and highlight the expression differences between *sdc-2*, *dpy-23*, and *mdh-1*.

Figure 3



**Figure 3 Detection of smFISH in wild-type and *ama-1* RNAi embryos**

A) Cartoon of the RNA Pol II with the largest *C. elegans* subunit AMA-1 and example images of wild-type or *ama-1* RNAi embryos, labeled with *sdc-2* smFISH. The number of detected smFISH spots for each example is written under the image. B) Plots for smFISH detection of X chromosomal mRNA *sdc-2*. Each spot represents the total mRNA count for one embryo, plotted against the binned nuclei number for wild type (yellow) or *ama-1* RNAi (dark purple). C) Same as in B but for the X chromosomal gene *dpy-23*. D) Plot as in B but for the autosomal mRNA *mdh-1*.

## Distinguishing transcription sites from mRNAs using smFISH data

smFISH probes against exons detect single mRNAs as distinct spots with similar intensities [3,52]. The images also capture a few spots with high intensity inside the nucleus, marking sites of transcription that contain multiple mRNAs [4]. RNAi leads to the degradation of mRNA but does not affect *de novo* transcription. Accordingly, while *ama-1* RNAi reduced the high-intensity sites in *sdc-2* smFISH due to Pol-II reduction, *sdc-2* RNAi reduced overall mRNAs, while the high-intensity spots remained prominent (Figure 4A). Thus, our analysis supports the conclusion that high-intensity smFISH spots are sites of active transcription (Figure 4A).

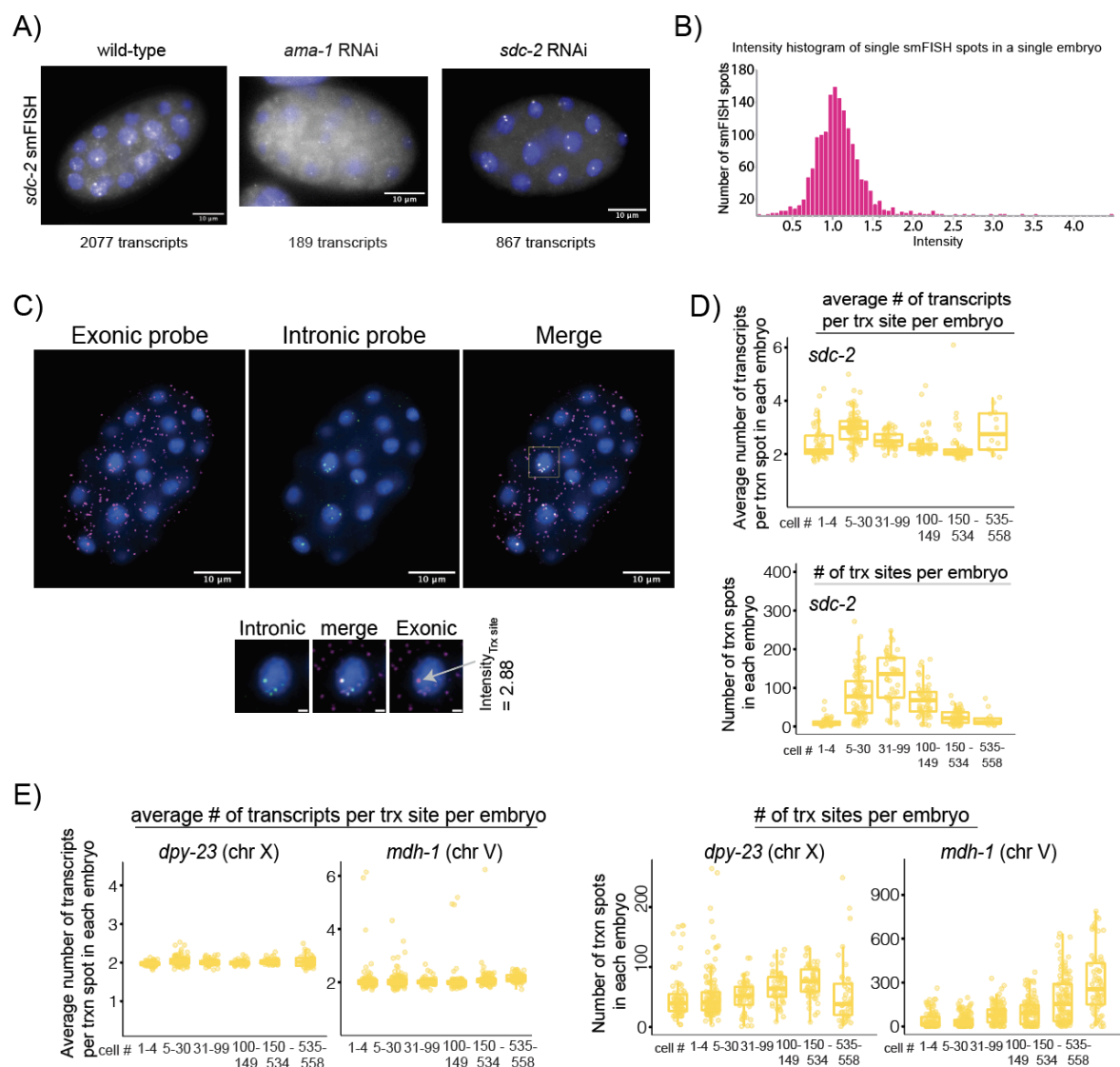
To distinguish transcription sites from individual transcripts, first we normalized the smFISH signal by setting the maximum of the intensity distribution of all detected mRNAs per embryo to 1 (Figure 4B, Suppl. Fig. 2B), which corresponds to a single mRNA as previously established [4]. We used a threshold of 1.7 to capture spots with more than 1.7 full-length RNAs and interpret them as sites of active transcription. This approach might underestimate the number of active transcription sites as genes that just started being transcribed would be excluded. To validate our approach, we used smFISH probes against introns, which mark the site of active transcription [4,52]. Visual inspection of several co-stained transcription sites confirmed this threshold (Figure 4C). Occasionally, more than two intronic spots were observed inside or outside a nucleus, possibly due to intron retention [53,54]. Since exonic probes allow the normalization of smFISH intensity between embryos, we used the exonic smFISH data to analyze transcription sites.

### **Zygotic activation of *sdc-2* involves regulation of both transcription burst size and frequency**

In the two-state random telegraph model of transcription kinetics, an increase in transcription can be accomplished by increasing the burst frequency (how often a promoter is turned on) or the burst size (the number of RNAs produced at a transcription site) [55,56]. These modeling approaches use experimental measurements in mammalian systems that are currently lacking for *C. elegans* [57,58]. Therefore, we reasoned that the number of transcription sites detected in each embryo (normalized to the number of nuclei) correlates with the frequency of transcription bursts, and the smFISH intensity detected at each transcription site correlates with the size of the transcription bursts.

Analysis of *sdc-2* in wild-type embryos showed that both the number of transcripts per transcription site and the number of transcription sites per embryo increase in the 5-30 cell stage when *sdc-2* is transcription activated for functions in sex determination and dosage compensation (Figure 4D). For *dpy-23* and *mdh-1*, the intensity of the smFISH signal at transcription sites was more similar, while the number of sites varied across embryogenesis (Figure 4E).

Figure 4



**Figure 4** Detecting transcription sites and parameters. **A)** Example images of embryos labeled with *sdc-2* smFISH probes. Embryos are either wild-type, *ama-1* RNAi, or *sdc-2* RNAi. The number of detected transcripts is stated under each image. **B)** Histogram of normalized smFISH detections (exon) for *dpy-23* for a single embryo. **C)** Example embryo showing the dual labeling of *dpy-23* exonic (magenta) and intronic (green) smFISH probes. The lower panel shows a detailed view of the yellow boxed nucleus with the scale bar indicating 1  $\mu$ m. The normalized intensity of the exonic transcription spot highlighted by the arrow is 2.88. The images were filtered using a Gaussian filter ( $\sigma = 5$ ). **D)** The upper plot shows the average number of *sdc-2* transcripts per transcription site per wild-type embryo. The lower plot shows the number of transcription sites per embryo. **E)** Left panel plots show the average number of *dpy-23* and *mdh-1* transcripts per transcription site in wild-type embryos—the plots on the right show the number of transcription sites per embryo.



## **Knockdown of the DCC subunits SDC-2 and DPY-27 leads to increased transcription of *dpy-23*, a dosage-compensated gene on the X chromosome**

The DCC subunit SDC-2 is the first to bind the X chromosomes and is required for the recruitment of the rest of the complex (Figure 5A) [24,59]. In hermaphrodite embryos, the knockdown of SDC-2 results in a ~1.8-fold increase of RNA Pol II binding at X chromosomal promoters as measured by GRO-seq [27]. We knocked down SDC-2 using RNAi and analyzed the transcription of a dosage-compensated gene, *dpy-23*. We used mock RNAi as a control because feeding is performed on different growth media and uses a different strain of bacteria compared to standard growth conditions used for the wild-type strain N2. We validated the level of *sdc-2* knockdown (Suppl. Fig. 3A) by quantifying the *sdc-2* smFISH signal (Suppl. Fig. 3B).

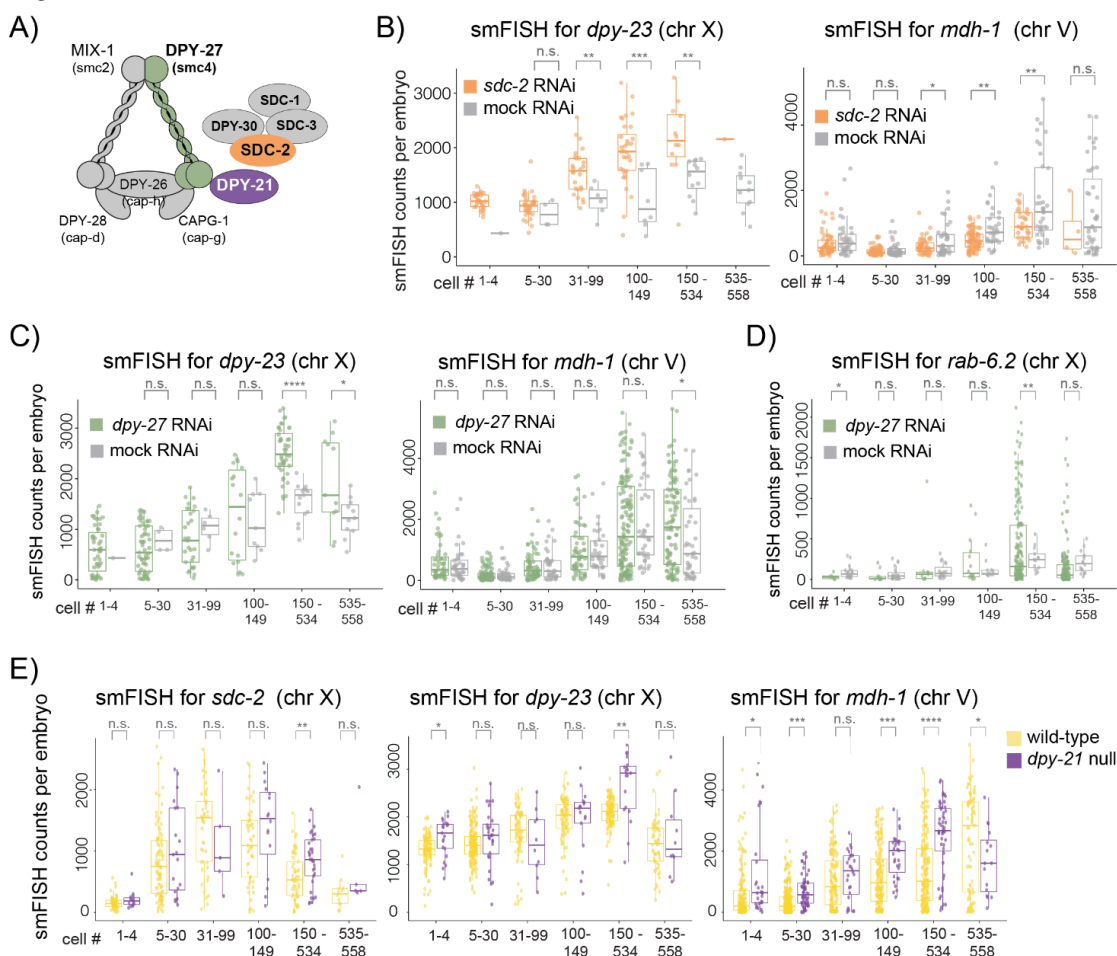
To capture the biological and technical variability between embryos, transcripts from an autosomal gene *mdh-1* were measured in the identical embryos in parallel to *dpy-23*. Since RNAi produces variable knockdown, having a large number of embryos was important to test for statistical significance. Where the sample numbers were low, we refrained from performing a statistical test. We found that *sdc-2* knockdown leads to upregulation of *dpy-23* but not *mdh-1* (Figure 5B). Notably, upregulation of *dpy-23* mRNA started from the 31-99 cell stage, immediately following DCC localization but before H4K20me1 enrichment on the X chromosomes. The apparent downregulation of *mdh-1* upon *sdc-2* RNAi may be due to the previously demonstrated repressive effect of *sdc-2* knockdown on autosomes [60].

To validate that the upregulation of *dpy-23* was due to reduced DCC binding, we performed RNAi to knock down the condensin core of the DCC, targeting the *smc-4* variant *dpy-27* (Suppl. Fig. 3C, D). *Dpy-27* RNAi also resulted in an increase in *dpy-23* transcripts (Figure 5C). To test if DCC-mediated repression can be observed for another X chromosomal gene whose transcription profile differs from that of *dpy-23*, we analyzed *rab-6.2*, a dosage-compensated gene with a lower transcription level and no maternal loading. There was higher variability in *rab-6.2* transcription in the *dpy-27* RNAi condition (Figure 5D). The reason for this variability is unclear.

## **Analysis of transcripts in a null mutant of DPY-21 H4K20me2 demethylase**

Condensin DC recruits the H4K20me2 demethylase DPY-21 to the X chromosomes, resulting in the enrichment of H4K20me1 starting at ~100-cell stage [46,48,61,62]. DPY-21 loss is not lethal thus, a homozygous null line is viable and shows a pronounced dumpy phenotype due to increased X chromosomal gene expression [46,59,62]. In the *dpy-21* null mutant embryos, the transcripts from two X chromosomal genes, *sdc-2* and *dpy-23*, increased in 150-534 nuclei (Figure 5E). Unlike other DCC subunits, DPY-21 is expressed in the germline and regulates H4K20me1 on the autosomes [62]. We observed an upregulation of *mdh-1* in the *dpy-21* null mutant (Figure 5E), which may reflect the repressive function of H4K20me1 on autosomal transcription.

Figure 5



**Figure 5 Measuring the effect of DCC knockdown on a dosage-compensated gene. A)** Cartoon of the DCC complex highlighting the subunits SDC-2, DPY-27, and DPY-21. **B)** Plots for smFISH detection of X chromosomal mRNA *dpy-23* and autosomal mRNA *mdh-1*. Each spot represents the total mRNA count for one embryo, plotted against the binned nuclei number for wild type (yellow) and *sdc-2* RNAi (orange). Data plotted for *mdh-1* includes measurements from additional embryos that were probed for *sdc-2*. **C)** Plots for the X chromosomal mRNA *dpy-23* or the autosomal RNA *mdh-1* in mock RNAi treatment (gray) or *dpy-27* KD (green). Data plotted for *mdh-1* includes embryos probed for *dpy-23* and *rab-6.2*. **D)** Plots for the X chromosomal gene *rab-6.2* in mock RNAi (gray) or *dpy-27* KD (green) condition. **E)** Plots for the X chromosomal genes *sdc-2* and *dpy-23* as well as the autosomal gene *mdh-1* in wild-type (yellow) or *dpy-21* KO (purple) conditions. Data plotted for *mdh-1* includes embryos probed for *sdc-2* and *dpy-23*. Significant P-values from Welch's unequal variances t-test are noted above the wild-type vs. knock-down condition for each time-point. P-values are denoted as n.s. ( $P > 0.05$ ), \* ( $P \leq 0.05$ ), \*\* ( $P \leq 0.01$ ), \*\*\* ( $P \leq 0.001$ ) and \*\*\*\* ( $P \leq 0.0001$ ). If one of the samples had less than two points, no P-value was calculated.

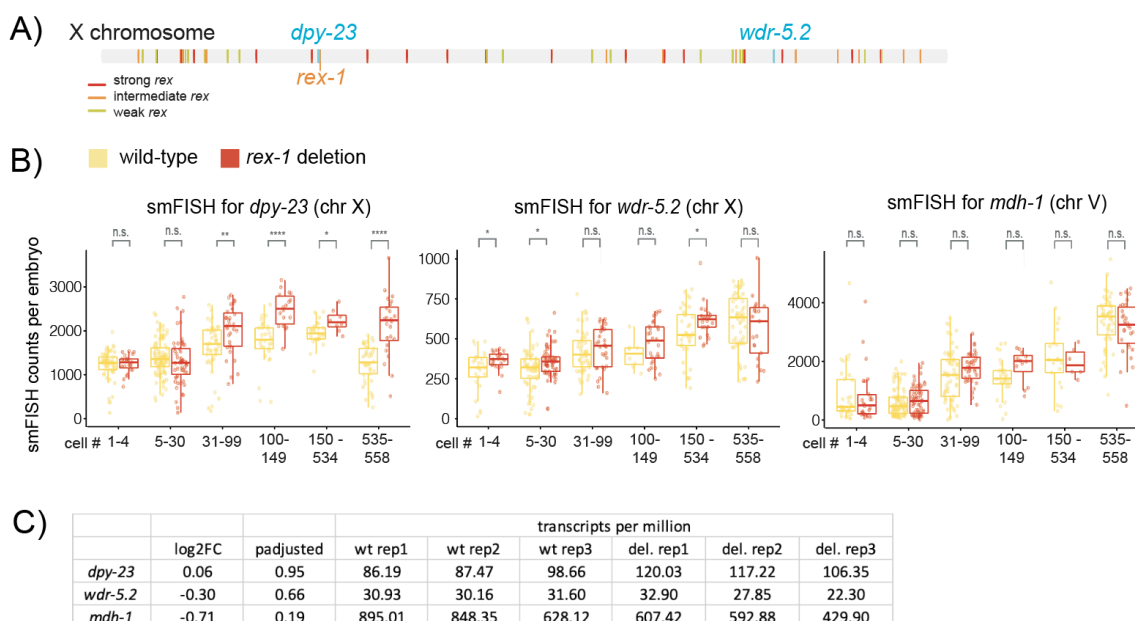
### smFISH combined with *in silico* staging detects the local and temporal effect of an individual DCC recruitment site deletion

In embryos, DCC binding across the length of the ~17.7 Mb X chromosome initiates at a set of ~60 recruitment elements on the X (*rex*) [26]. The DCC spreads out from these sites and exerts its repression chromosome-wide [28,63]. Individual *rex* sites collectively contribute to recruitment and repression, thus deletion of one or few *rex* sites causes subtle changes, as measured by mRNA-seq in mixed-stage embryos [64,65]. We wondered if our approach could reveal the contribution of an individual *rex* site in developmentally staged embryos.

We previously constructed a strain deleting *rex-1*, located ~6 kb downstream of *dpy-23* [64]. Using three different fluorophores (Suppl. Fig. 1B), we quantified two internal control genes simultaneously with *dpy-23*; *wdr-5.2*, which is far from *rex-1*, and the autosomal gene *mdh-1* (Figure 6A). Compared to the two internal controls, *dpy-23* transcripts increased significantly upon *rex-1* deletion (Figure 6B, Suppl. Fig. 5).

Next, we tested whether mRNA-seq could detect the same effect in mixed-stage embryos. The transcript per million (TPM) measure of *dpy-23* expression was slightly elevated in the deletion strain but did not reach statistical significance (Figure 6C). This is likely due to mixed-stage embryos containing a high proportion of early embryos that have not started dosage compensation. Thus, smFISH analysis of *in silico* staged embryos reveals the effect of a *rex* site that is missed by population assays.

**Figure 6**



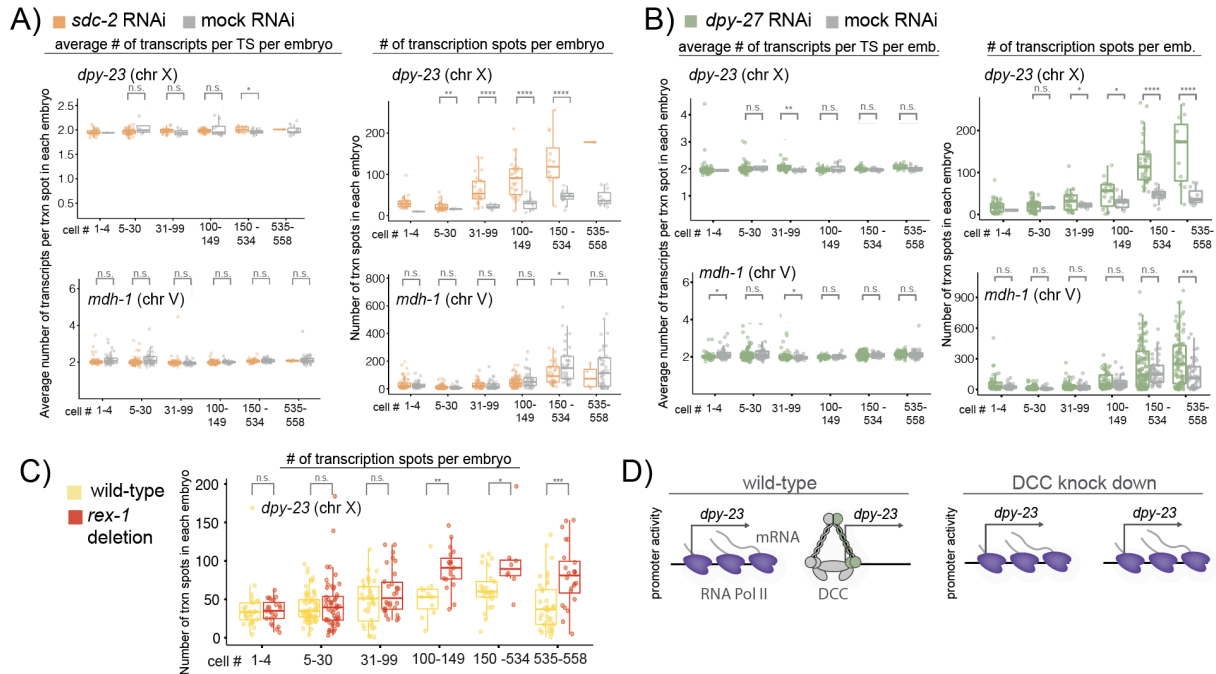
**Figure 6** Influence of a single *rex* site deletion on a neighboring gene. **a)** Cartoon of the X chromosome with approximate locations of the genes *dpy-23* and *wdr-5.2* (~10 Mbp away) and the location of *rex-1*. **b)** Plots for smFISH detection X chromosomal mRNAs (*dpy-23* and *wdr-5.2*) and one autosomal mRNA (*mdh-1*) in wild-type and *rex-1* deletion strains. The data are from the same embryos using 3-color imaging. Each spot represents the total mRNA count for one embryo, plotted against the binned nuclei number for wild-type (yellow) or *rex-1* deletion (red) strains. Significant P-values from Welch's unequal variances t-test are noted above the wild-type vs. deletion condition for each time point. P-values are denoted as n.s. ( $P > 0.05$ ), \* ( $P \leq 0.05$ ), \*\* ( $P \leq 0.01$ ), \*\*\* ( $P \leq 0.001$ ) and \*\*\*\* ( $P \leq 0.0001$ ). **c)** mRNA-seq data from mixed-stage embryos isolated from gravid adults was used for differential expression analysis. DESeq2 output of the log<sub>2</sub> fold change and adjusted p values are provided along with the expression values from the three biological replicates.

### SDC-2 and DPY-27 knockdown increases the number of *dpy-23* transcription sites per embryo

Next, we addressed how the DCC represses *dpy-23* expression. Upon *sdc-2* RNAi, the average *dpy-23* smFISH signal at the transcription sites did not change, but the number of transcription sites per embryo increased significantly (Figure 7A). The effect of *dpy-27* RNAi on *dpy-23* was similar to that of *sdc-2*, increasing the number of transcription sites per embryo rather than the number of transcripts at each site (Figure 7B). To ensure that our conclusion is not sensitive to threshold selection with which we define active transcription sites, we tested if an intensity threshold of 1.5 or 2.0 changes the result. Both thresholds recapitulated the main finding (Suppl. Fig. 4). In addition to the RNAi condition, in the *rex-1* deletion strain, the number of transcription sites increased significantly for the

neighboring gene *dpy-23* (Figure 7C) and not for the control genes (Suppl. Fig. 4). Thus, our results support the idea that the DCC represses *dpy-23* transcription by reducing its frequency of transcription.

Figure 7



**Figure 7 Analysis of transcription site parameters upon DCC knockdown. A)** Relative transcription parameter (average number of transcripts per transcription site per embryo and the number of transcription spots per embryo) changes for the X chromosomal gene *dpy-23* for *sdc-2* RNAi (orange) and mock RNAi (gray) conditions. The left plots show the average number of nascent transcripts per TSS, which is correlated with the burst size. The right plots show the number of actively transcribing promoters, which is correlated with the burst frequency. Data from *mdh-1* includes embryos co-probed with *dpy-23* and *rab-6.2*. **B)** Relative transcription parameter changes for *dpy-27* RNAi (green) and mock RNAi (gray) conditions. **C)** Relative number of transcription spots per embryo for the X chromosomal gene *dpy-23* for wild-type (yellow) and *rex-1* deletion (red) conditions. Significant P-values from Welch's unequal variances t-test are noted above the wild-type vs. knock-down condition for each time-point. P-values are denoted as n.s. ( $P > 0.05$ ), \* ( $P \leq 0.05$ ), \*\* ( $P \leq 0.01$ ), \*\*\* ( $P \leq 0.001$ ) and \*\*\*\* ( $P \leq 0.0001$ ). If one of the samples had less than two too few points, no P-value was calculated. **D)** Model for transcription regulation during dosage compensation.

## Discussion

smFISH analysis of transcription in developing organisms is powerful but presents experimental and computational challenges. In this work, we used widefield smFISH to quantify transcripts in each embryo and applied machine learning to stage a large number of embryos for analysis of transcription throughout embryogenesis in *C. elegans*.

### *Transcriptional activation of an early zygotic gene sdc-2*

Our analysis of *sdc-2* suggests that both the rate and frequency of transcription bursts increase significantly following zygotic genome activation (Figure 4). In *C. elegans*, the single fertilized cell undergoes several divisions to become a 30-cell-stage embryo in ~150 min [36,66,67]. Rapidly activated zygotic genes tend to be shorter [53]. Yet, *sdc-2* transcript (~10 kb) is longer than the median transcript in the genome (~2 kb) and encodes a large ~340 kD protein. Therefore, multiple mechanisms may activate *sdc-2* transcription to ensure robust production of SDC-2 protein for timely initiation of hermaphrodite development and dosage compensation.

### *Timing of DCC-mediated transcription repression*

Our current model is that the DCC starts repression soon after SDC-2-mediated recruitment of condensin DC to the X chromosomes and is augmented by enrichment of H4K20me1 in later embryogenesis [59,61]. Consistent with this model, we observed an earlier and stronger effect of *sdc-2* RNAi compared to *dpy-27* and *dpy-21* (Figure 5). The DCC-mediated repression starting after the 30-cell stage was also supported by the effect of *rex-1* deletion on the neighboring *dpy-23* gene (Figure 6). The deletion analysis also highlighted the power of *in silico* staging embryos to detect the temporal effect of a single *rex* site, which was missed by mRNA-seq analysis of mixed-stage embryos.

### *Mechanism of DCC-mediated transcription repression*

Our data suggest that the DCC reduces the number of *dpy-23* alleles that are actively transcribed in an embryo (Figure 7). Notably, *dpy-23* expression during development appears to be regulated by the frequency of transcription rather than the number of mRNAs being produced at each site (Figure 4E). Therefore, the DCC may repress *dpy-23* by modulating the mechanism that controls its frequency of transcription. It will be interesting to find another X chromosomal gene that is primarily activated by burst size to test if the DCC also has the ability to also modulate the size of transcription bursts.

How does the DCC reduce the frequency of *dpy-23* transcription? Previous research has shown that in addition to TFs [3,13–15,68–71] and chromatin [16,17,59,72–75], 3D organization of enhancers and promoters can modulate burst frequency [9–11,72,76–78]. The DCC has the potential to regulate one or more of these processes, as the core of the DCC is a structural maintenance of chromosomes (SMC) complex that controls the 3D DNA contacts on the X chromosomes [79,80] and is required for the compaction and regulation of histone modifications on chromatin [46,48,61,81–83].

The level and dynamics of condensin DC binding are important for transcription repression [59,81]. In a *dpy-21* null mutant where X chromosome transcription increases, the proportion of mobile DPY-27 proteins, as measured by fluorescence recovery after photobleaching (FRAP), decreases dramatically [59]. Therefore, we speculate that the transient association of condensin DC molecules with the *dpy-23* promoter may reduce RNA Pol II initiation at a fraction of the time, resulting in fewer alleles transcribing within a population of embryos (Figure 7D).

### *Fine-tuning the transcriptional output of an entire chromosome for dosage compensation*

Transcription is a stochastic process leading to differential mRNA and protein levels between cells and organisms with the same genetic background and sharing the same environment [84,85]. Even in a deterministic organism like *C. elegans* with a fixed lineage and cell numbers, transcription was described as stochastic in reaction to mutations [19] during development [18]. The presence of X chromosome dosage compensation mechanisms argues that the high level of transcriptional heterogeneity at the short time scale is averaged over time to produce the required dosage of gene transcripts.

In *D. melanogaster*, a complex of proteins and non-coding RNAs called the male-specific lethal (MSL) complex binds to and increases transcription from the X chromosome in males to accomplish dosage compensation [86,87]. A recent study used the MS2-MCP system to live-image transcription of four X chromosomal genes in female and male embryos and found that the amplitude of transcriptional bursts was increased in males [88]. How the different mechanisms of *D. melanogaster* and *C. elegans* dosage compensation (2-fold upregulation in the fly and 2-fold downregulation in the worm) modulate the burst and frequency parameters of transcription to fine-tune the average output of an entire chromosome is a fascinating question for future research.

## Methods

### Strains and Worm Growth

Unless noted, worms were grown and maintained using standard methods at 20-22°C on NGM plates (1L: 980 ml Milli-Q water, 3 g NaCl, 20 g agar, 2.5 g peptone, 1 ml 1 M CaCl<sub>2</sub>, 1 ml 5 mg/ml cholesterol in ethanol, 1 ml 1 M MgSO<sub>4</sub> and 25 ml 1 M KPO<sub>4</sub> buffer) containing OP50-1 strain of *E. coli* as food. Worm strains used in this study include N2 (wild-type, CGC), CB428 (*dpy-21* (e428) V, CGC), and ERC41 (*ers8*[delX:4394846-4396180]) [64].

### RNAi conditions

RNAi conditions used in this study include *pop-1* (W10C8.2), *sdc-2* (C35C5.1), *dpy-27* (R13G10.1), *ama-1* (F36A4.7), *rluc* (*Renilla luciferase*) and empty vector (L4440) all from the Ahringer library. Bacteria were sequenced before use to ensure the correct strain was used. RNAi plates (RNAi plates (1 L), 3 g NaCl, 17 g bacto agar, 3 g bacto peptone, 975 ml water, autoclave, cool and add: 1 ml cholesterol (5 mg/ml in EtOH), 1 ml 1M CaCl<sub>2</sub>, 1 ml 1M MgSO<sub>4</sub>, 25 ml 1 M K phosphate buffer, 1 ml IPTG (1 mM final concentration, made freshly), 1 ml Amp (50 mg/ml stock), 1 ml Tet (12 mg/ml stock)) were prepared freshly and usually not stored longer than one month in the dark. Single clones of bacteria were picked and cultured in a 5 ml overnight culture (LB + Amp Kan) at 37°C and shaking at 400 rpm. The next day, the small culture was added to typically 200 mL LB (Amp Kan) and continued shaking for 2-4 hours when the density of the culture reached OD 1. Then, 1 ml IPTG was added to the culture, and incubation continued for 2 hours. To reduce the bacteria volume by 1/10, bacteria were spun down at 1000 g and resuspended in M9 to have about 400 µl of bacteria per plate (very dense).

Worms were grown to the adult stage and bleached to retrieve embryos. Embryos were placed in a 15 ml tube with M9 at RT and were shaken overnight. The next day, starved L1s were placed on the freshly prepared RNAi plates at 20°C in the dark. For *ama-1* RNAi, L3/L4 were placed on RNAi bacteria to allow germline formation. After three days, adult worms were collected, and embryos were extracted, as described below.

### smFISH staining in *C. elegans* embryos

The smFISH labeling of *C. elegans* embryos was performed as reported in Bahry et al. [51] with some modifications.

#### *Probe design*

All smFISH probe sets (Custom Stellaris® RNA FISH Probe Set) were designed and manufactured by Biosearch Technologies using the Stellaris probe designer web tool (Biosearch Technologies) specifically for *C. elegans*. Probes were designed to target exons or introns and labeled Quasar 670, CAL Fluor Red 610, or Quasar 570. Probes were designed to have at least two-nucleotide space in between neighboring probes to avoid quenching. For exonic or intronic probe design, sequences listed at wormbase.org were used, and splice regions were masked with "n" nucleotides to avoid probes in these areas. If possible, the maximum amount of probes (48 per order) was selected for each exonic or intronic sequence. An example of the distribution of exonic and intronic probes for *dpy-23* mRNA can be found in Supplemental Figure 1a. Filter sets with narrow bandwidths were selected to separate the detection from the three wavelengths ( $\lambda$  570 nm, 610 nm, and 670 nm). We observed no bleedthrough between the three channels (Supplemental Figure 1b). Typically, far-red dyes performed better due to the autofluorescence of *C. elegans* embryos, especially in older stages. The number and color of probes are listed in Supplemental Table 1. The list of complete sequences can be found in the Supplements.

### *Preparation of worms*

Worms were synchronized by either bleaching or egg-laying and grown to the adult stage. At this step, having non-starved, healthy adult worms in sufficient amounts is crucial. Typically, at least 5 x 6 cm plates full of worms were used to collect embryos.

### *Collection of embryos*

For embryo collection, worms, and older embryos were washed off plates using M9 (5.8 g Na<sub>2</sub>HPO<sub>4</sub>, 3.0 g KH<sub>2</sub>PO<sub>4</sub>, 0.5 g NaCl, 1.0 g NH<sub>4</sub>Cl, Nuclease-free water to a final volume of 1000 ml) and collected in a 35 µm nylon filter. Worms were washed at least three times with H<sub>2</sub>O, then carefully transferred to a falcon tube using M9 and allowed to settle down. The supernatant was removed, and 5 ml of freshly prepared bleaching solution (2.5 ml 4N NaOH, 2.5 ml 5% NaClO, 5 ml Nuclease-free water) was added to the worms. The dissolving of the adult worms was closely observed, and after 3 minutes, the tubes were spun at 3000 g for 1 minute to collect the embryos. The supernatant was quickly removed, and the embryo pellet was vortexed. Then 10 ml of 1x PBS (with 0.05% of Triton X-100) was added, and the tube was centrifuged again for 3 min at 3000 g. This step was repeated twice until a clean embryo extract was left in the tube.

### *Fixing and permeabilization*

Embryos were resuspended in 1 ml fixation solution (4% paraformaldehyde (PFA) in 1xPBS (DEPC treated + autoclave + 0,05% of Triton X-100) and incubated at RT for 15 min while rotating. Next, the tube was submerged in liquid nitrogen for 1 minute to freeze crack the embryo eggshells. The tube was then transferred to a beaker with RT water to thaw, and once it was fully thawed, it was kept on ice for an additional 20 minutes. After this incubation, the tubes were spun down at 3000 g for 3 minutes, the supernatant was removed, and the embryos were washed twice with 1 ml of 1x PBS (with 0.05% Triton X-100). The embryos were resuspended in 70% EtOH and kept at 4 °C for at least 24 hours. Embryos can be kept at 4 °C for at least several weeks.

### *smFISH staining*

For smFISH staining of the previously fixed embryos, tubes were centrifuged at 3000 g for 3 minutes, and ethanol was carefully removed. The pellet can be loose at this step, so removing the supernatant should also be done in two stages. Embryos were then resuspended in 1 ml wash buffer (40 ml nuclease-free water, 5 ml deionized formamide, 5 ml 20x SSC) and vortexed. Tubes were centrifuged, as above, and the supernatant was removed. The embryos were then resuspended in 50 µl hybridization solution (50 µl H<sub>2</sub>O (RNase free), 37.5 µl ethylene carbonate (EC) 5 mg/ml, 25 µl formamide (at RT), 12.5 µl SDS (dissolved), 125 µl dextran sulfate 10%), and 1 µl of each probe set (12.5 µM stock solution) was added directly to the sample. Tubes were then vortexed lightly and incubated at 37°C in the dark overnight. The next day, 0.5 ml of the wash buffer was added, and the tubes vortexed and centrifuged to remove the supernatant. Next, 1 ml of wash buffer was added, and samples were incubated at 37°C for 30 minutes. After that, tubes were centrifuged again to remove the supernatant, and the embryo pellet was vortexed before adding 1 ml wash buffer. In this step, DAPI (5 ng/mL) was added to the wash buffer, and tubes were incubated at 37 °C for 30 minutes. After centrifugation, the wash buffer was removed, and samples were washed once with 2x SSC.

## Mounting

Most liquid was removed from the stained embryos to mount the stained embryo tubes. About 15  $\mu$ l of dense embryo solution (in 2x SSC) were used per microscopy slide and spread onto a coverslip (#1.5, 22 x 22 mm). The sample was left to dry for about 15 minutes, and then 15  $\mu$ l of ProLong<sup>TM</sup> Diamond Antifade Mountant (Thermo Fisher) was added to the sample. A glass slide (SuperFrost, VWR) was pressed onto the embryos, and mounting media was placed on the coverslip. Slides were left at RT in the dark for 24 hours to cure before sealing the sides with nail polish and then stored at 4°C. Images were then acquired within two weeks of the sample being prepared.

## Imaging smFISH embryos

Embryos were imaged on a Nikon Ti inverted fluorescence microscope with an EMCCD camera (ANDOR, DU iXON Ultra 888), Lumen 200 Fluorescence Illumination Systems (Prior Scientific), and a 100x plan apo oil objective (NA 1.4) using appropriate filter sets (Gold FISH (Chroma 49304), Red FISH (Chroma 49306), Cy5 Narrow (Chroma 49396), GFP, DAPI). Images were acquired with 90 z-stacks positions with 200 nm step-width using Nikon Elements software. Every field of view had 1024x1024 pixels (XY axes) and ~ 90 slices (Z), with 0.13  $\mu$ m lateral and 0.2  $\mu$ m axial resolution. Positions of embryos were marked in the software for semi-automatic imaging.

The 7447 images used in this study are provided at <https://data.janelia.org/3Mq4F>. The same site also provides 2505 smFISH images corresponding to additional conditions not used in this paper.

## Embryo instance segmentation

Single fields of view contained several embryos, which were segmented for downstream analysis. An instance segmentation was performed on the max projection of the GFP channel, which contained an autofluorescence signal marking the entire embryo (Supplemental Figure 2a). We used a 2D StarDist implementation, an approach for nuclei/cell detection with a star convex shape, to extract individual embryos [89]. A ground-truth dataset was created using a random forest classifier and a custom ellipsoid fit script implemented in Java, resulting in 1734 fully annotated images. This dataset was split into training (85% - 1168 images) and validation (15% - 206 images) sets. As preprocessing steps, each GFP max projection image was normalized and resized from 1024x1024 to 512x512. All ground-truth (GT) embryo instances were detected (true positives) with a Jaccard similarity score > 0.75 (Supplemental Figure 2a). Instances of false positives, where embryos that were not annotated as GT were detected, occurred in 7 (2.83%) predicted instances. Post-processing for each embryo included predicted label resizing back to the original size and creating a cropped individual embryo image (40 pixels padding) (Supplemental Figure 2a). Analysis scripts can be found here: [https://github.com/PreibischLab/nd2totif-maskembryos-stagebin-pipeline/blob/master/2\\_stardist\\_predict.py](https://github.com/PreibischLab/nd2totif-maskembryos-stagebin-pipeline/blob/master/2_stardist_predict.py)

## Staging of embryos

Classification of different embryos into selected age bins was done by training an autoencoder-based classifier for stage prediction (Figure 2). The developmental bins used for staging (1-4, 5-30, 31-99, 100-149, 150-534, 535-558) were a compromise, selected based on biological meaning and classification prediction capability. A validation set comprising 250 embryos was employed to validate the autoencoder training. The validation of the classifier results was done using a class-stratified leave-one-out approach. For preprocessing of all images (training set and full dataset), 3D DAPI channel images were first masked for each embryo, then the 21 central slices were extracted, and pixel intensities were normalized using only non-zero values. Then, slightly modified copies were generated using small shifts, flips, rotations, shear, and brightness changes



(augmentation). From those copies, 750 2D tiles of size 64x64 were extracted from random image regions for each embryo. All tiles of the embryo were used for training and inference.

First, an autoencoder was trained in an auto-associative, self-supervised manner to reconstruct its input, enabling it to learn a latent representation of the images. The autoencoder utilized mean squared error as the loss function and employed the Adam optimizer for training. In the subsequent step, the pre-trained encoder was integrated with a classifier designed to predict embryo stages, where the classification is achieved through a softmax output layer corresponding to the stage bins. This new network was initially trained with frozen encoder weights and then finally tuned by training the full network. For the classifier, cross-entropy was used as the loss function, and the Adam optimizer was used for optimization. A stage was predicted for each 64x64 tile of an embryo image, and the final stage for each embryo was determined by a majority vote of all tiles.

A training dataset was manually annotated using a custom-written ImageJ macro. We annotated ~100 embryos with this approach, but we faced a strong class imbalance due to the initial random sampling of embryos. Therefore, an active learning approach (human in the loop) was utilized. Embryos for additional manual annotation were chosen based on low certainty in the classifier's initial predictions, focusing on classes lacking sufficient examples. This method allowed us to move away from random sampling in subsequent iterations, strategically annotating stages with fewer examples to balance the distribution across classes (Figure 2B).

Scripts can be found here:

[https://github.com/PreibischLab/nd2totif-maskembryos-stagebin-pipeline/blob/master/4\\_stage\\_prediction.py](https://github.com/PreibischLab/nd2totif-maskembryos-stagebin-pipeline/blob/master/4_stage_prediction.py)

## Transcription detection

Detection of single RNA spots in 3D was performed using the Fiji plugin RS-FISH [51]. For this, images were preprocessed by subtracting a duplicated and median filtered (sigma = 19) image from the raw image to increase single spots and smooth background signals (Supplemental Figure 2B). Spot detection was performed using RS-FISH [51] with the following detection settings: -i0 0, -i1 65535, -a 0.650, -r 1, -s 1.09, -t 0.001, -sr 2, -ir 0.3, -e 0.5116, -it 33000, -bg 0. Spots were filtered using binary masks to exclude spots found from neighboring embryos using the Mask filtering option in RS-FISH with masks created using Stardist as described above. Computation was performed in parallel on the cluster and using AWS (Amazon Web Services). To correct for z-dependent signal loss, a quadratic function was fitted to the detected spots and used to correct the spot intensity throughout the embryo. A gamma function was fitted to the histogram of all found detections, and the maximum of the curve was set to 1 to normalize intensity detection between different embryos in order to allow quantification of transcript numbers (Figure 2B).

## Plotting and statistical analysis

Detected spots were plotted using custom-written scripts in R using ggplot2 and ggpubr. To test if mutant and wild-type conditions had significantly different RNA levels or intensities, a Welch's unequal variances t-test was used in R. Significant P-values were indicated in the plots as n.s. ( $P > 0.05$ ), \* ( $P \leq 0.05$ ), \*\* ( $P \leq 0.01$ ), \*\*\* ( $P \leq 0.001$ ) and \*\*\*\* ( $P \leq 0.0001$ ). If one of the samples had too few points, no P-value was calculated.

## mRNA sequencing and analysis

mRNA-seq and data analysis from the deletion strain ERC41 were performed using methods to generate and analyze published data from wild-type embryos [64]. mRNA-seq data is deposited to GEO under series number GSE265989 with relevant information provided in Supplemental File 1.

## Acknowledgments

Research in this manuscript and SE was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R35 GM130311. L. B. was supported by MDC-Berlin and the Joachim Herz Foundation (no. 850022), E. B. was supported by the HFSP grant RGP0021/2018-102 and MDC-Berlin, M. Z. was supported by EU H2020 Training grant 721890 ‘circRtrain’ and HHMI Janelia, S. P. was supported by HFSP grant RGP0021/2018-102, EU H2020 Training grant 721890 ‘circRtrain,’ MDC-Berlin and HHMI Janelia. We thank Sarah Albritton and Haoyu Wang for their contribution to the mRNA-seq data and the Gencore at the NYU Center for Genomics and Systems Biology for sequencing and raw data processing. We additionally thank Andrew Woehler and the BIMSB advanced imaging facility for their support and maintenance of the microscope. We also thank the HHMI Janelia Open Science Software Initiative for maintenance and help with RS-FISH.

## Competing interest

The authors declare no competing interest.

## Data availability

All raw smFISH images of *C. elegans* embryo, including the annotated data used for training, can be found here: <https://data.janelia.org/3Mq4F>. The RNA-seq data is deposited at Gene Expression Omnibus under series number GSE265989.

## Code availability

All custom code is either deposited or linked in this GitHub repository:  
[https://github.com/ercanlab/2024\\_Breimann\\_et\\_al](https://github.com/ercanlab/2024_Breimann_et_al)

## References

1. Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. Stochastic mRNA Synthesis in Mammalian Cells. *PLoS Biology*. 2006;4: e309. doi:10.1371/journal.pbio.0040309
2. Little SC, Tikhonov M, Gregor T. Precise Developmental Gene Expression Arises from Globally Stochastic Transcriptional Activity. *Cell*. 2013;154: 789–800. doi:10.1016/j.cell.2013.07.025
3. Senecal A, Munsky B, Proux F, Ly N, Braye FE, Zimmer C, et al. Transcription Factors Modulate c-Fos Transcriptional Bursts. *CellReports*. 2014;8: 75–83. doi:10.1016/j.celrep.2014.05.053
4. Halpern KB, Tanami S, Landen S, Chapal M, Szlak L, Hutzler A, et al. Bursty Gene Expression in the Intact Mammalian Liver. *Molecular Cell*. 2015; 1–11. doi:10.1016/j.molcel.2015.01.027
5. Singh A, Soltani M. Quantifying Intrinsic and Extrinsic Variability in Stochastic Gene Expression Models. *Plos One*. 2013;8: e84301. doi:10.1371/journal.pone.0084301
6. Munsky B, Neuert G, Oudenaarden A van. Using Gene Expression Noise to Understand Gene Regulation. *Science*. 2012;336: 183–187. doi:10.1126/science.1216379

7. Harrison MM, Marsh AJ, Rushlow CA. Setting the stage for development: the maternal-to-zygotic transition in *Drosophila*. *GENETICS*. 2023;225: iyad142. doi:10.1093/genetics/iyad142
8. Bakker R, Mani M, Carthew RW. The Wg and Dpp morphogens regulate gene expression by modulating the frequency of transcriptional bursts. *eLife*. 2020;9: R429-26. doi:10.7554/elife.56076
9. Hoppe C, Bowles JR, Minchington TG, Sutcliffe C, Upadhyai P, Rattray M, et al. Modulation of the Promoter Activation Rate Dictates the Transcriptional Response to Graded BMP Signaling Levels in the *Drosophila* Embryo. *Developmental Cell*. 2020; 1–23. doi:10.1016/j.devcel.2020.07.007
10. Fukaya T, Lim B, Levine M. Enhancer Control of Transcriptional Bursting. *Cell*. 2016;166: 358–368. doi:10.1016/j.cell.2016.05.025
11. Larsson AJM, Johnsson P, Hagemann-Jensen M, Hartmanis L, Faridani OR, Reinius B, et al. Genomic encoding of transcriptional burst kinetics. *Nature*. 2018; 1–17. doi:10.1038/s41586-018-0836-1
12. Hornung G, Bar-Ziv R, Rosin D, Tokuriki N, Tawfik DS, Oren M, et al. Noise–mean relationship in mutated promoters. *Genome Res*. 2012;22: 2409–2417. doi:10.1101/gr.139378.112
13. Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. Mammalian genes are transcribed with widely different bursting kinetics. *Science*. 2011;332: 472–474. doi:10.1126/science.1198817
14. Rullan M, Benzinger D, Schmidt GW, Miliás-Argeitis A, Khammash M. An Optogenetic Platform for Real-Time, Single-Cell Interrogation of Stochastic Transcriptional Regulation. *Mol Cell*. 2018;70: 745-756.e6. doi:10.1016/j.molcel.2018.04.012
15. Donovan BT, Huynh A, Ball DA, Patel HP, Poirier MG, Larson DR, et al. Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting. *Embo J*. 2019;38. doi:10.15252/embj.2018100809
16. Weinberger L, Voichek Y, Tirosh I, Hornung G, Amit I, Barkai N. Expression Noise and Acetylation Profiles Distinguish HDAC Functions. *Mol Cell*. 2012;47: 193–202. doi:10.1016/j.molcel.2012.05.008
17. Batenchuk C, St-Pierre S, Tepliakova L, Adiga S, Szuto A, Kabbani N, et al. Chromosomal Position Effects Are Linked to Sir2-Mediated Variation in Transcriptional Burst Size. *Biophys J*. 2011;100: L56–L58. doi:10.1016/j.bpj.2011.04.021
18. Lee C, Shin H, Kimble J. Dynamics of Notch-Dependent Transcriptional Bursting in Its Native Context. *Dev Cell*. 2019;50: 426-435.e4. doi:10.1016/j.devcel.2019.07.001
19. Raj A, Rifkin SA, Andersen E, Oudenaarden A van. Variability in gene expression underlies incomplete penetrance. *Nature*. 2010;463: 913–918. doi:10.1038/nature08781
20. Edgar LG, Wolf N, Wood WB. Early transcription in *Caenorhabditis elegans* embryos. *Development*. 1994;120: 443–451. doi:10.1242/dev.120.2.443
21. Baugh LR, Hill AA, Slonim DK, Brown EL, Hunter CP. Composition and dynamics of the *Caenorhabditis elegans* early embryonic transcriptome. *Development*. 2003;130: 889–900. doi:10.1242/dev.00302
22. Seydoux G, Fire A. Soma-germline asymmetry in the distributions of embryonic RNAs in *Caenorhabditis elegans*. *Dev Camb Engl*. 1994;120: 2823–34.
23. Seydoux G, Dunn MA. Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of *Caenorhabditis elegans* and *Drosophila melanogaster*. *Dev Camb Engl*. 1997;124: 2191–201.
24. Dawes HE, Berlin DS, Lapidus DM, Nusbaum C, Davis TL, Meyer BJ. Dosage compensation proteins targeted to X chromosomes by a determinant of hermaphrodite fate. *Science*. 1999;284: 1800–1804. doi:10.1126/science.284.5421.1800
25. Csankovszki G, Collette K, Spahl K, Carey J, Snyder M, Petty E, et al. Three Distinct Condensin Complexes Control *C. elegans* Chromosome Dynamics. *Curr Biol*. 2009;19: 9–19.

doi:10.1016/j.cub.2008.12.006

26. Albritton SE, Ercan S. *Caenorhabditis elegans* Dosage Compensation: Insights into Condensin-Mediated Gene Regulation. *Trends in Genetics*. 2018;34: 41–53. doi:10.1016/j.tig.2017.09.010
27. Kruesi WS, Core LJ, Waters CT, Lis JT, Meyer BJ. Condensin controls recruitment of RNA polymerase II to achieve nematode X-chromosome dosage compensation. *eLife*. 2013;2: e00808. doi:10.7554/eLife.00808
28. Kramer M, Rao P, Ercan S. Untangling the Contributions of Sex-Specific Gene Regulation and X-Chromosome Dosage to Sex-Biased Gene Expression in *Caenorhabditis elegans*. *Genetics*. 2016;204: 355–369. doi:10.1534/genetics.116.190298
29. Plenefisch JD, DeLong L, Meyer BJ. Genes that implement the hermaphrodite mode of dosage compensation in *Caenorhabditis elegans*. *Genetics*. 1989;121: 57–76.
30. Breimann L, Preusser F, Preibisch S. Light-microscopy methods in *C. elegans* research. *Current Opinion in Systems Biology*. 2019;13: 82–92. doi:10.1016/j.coisb.2018.11.004
31. Harterink M, Kim D hyun, Middelkoop TC, Doan TD, Oudenaarden A van, Korswagen HC. Neuroblast migration along the anteroposterior axis of *C. elegans* is controlled by opposing gradients of Wnts and a secreted Frizzled-related protein. *Development*. 2011;138: 2915–2924. doi:10.1242/dev.064733
32. Nishimura EO, Zhang JC, Werts AD, Goldstein B, Lieb JD. Asymmetric Transcript Discovery by RNA-seq in *C. elegans* Blastomeres Identifies *neg-1*, a Gene Important for Anterior Morphogenesis. *Plos Genet*. 2015;11: e1005117. doi:10.1371/journal.pgen.1005117
33. Seidel HS, Ailion M, Li J, Oudenaarden A van, Rockman MV, Kruglyak L. A Novel Sperm-Delivered Toxin Causes Late-Stage Embryo Lethality and Transmission Ratio Distortion in *C. elegans*. *Plos Biol*. 2011;9: e1001115. doi:10.1371/journal.pbio.1001115
34. Charles S, Aubry G, Chou H-T, Paaby AB, Lu H. High-Temporal-Resolution smFISH Method for Gene Expression Studies in *Caenorhabditis elegans* Embryos. *Anal Chem*. 2020. doi:10.1021/acs.analchem.0c02966
35. Parker DM, Winkenbach LP, Boyson S, Saxton MN, Daidone C, Al-Mazaydeh ZA, et al. mRNA localization is linked to translation regulation in the *Caenorhabditis elegans* germ lineage. *Development*. 2020;147: dev186817. doi:10.1242/dev.186817
36. Sulston JE, Schierenberg E, White JG, Thomson JN. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Developmental Biology*. 1983;100: 64–119. doi:10.1016/0012-1606(83)90201-4
37. Hench J, Henriksson J, Lüppert M, Bürglin TR. Spatio-temporal reference model of *Caenorhabditis elegans* embryogenesis with cell contact maps. *Developmental Biology*. 2009;333: 1–13. doi:10.1016/j.ydbio.2009.06.014
38. Bao Z, Murray JI, Boyle T, Ooi SL, Sandel MJ, Waterston RH. Automated cell lineage tracing in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences*. 2006;103: 2707–2712. doi:10.1073/pnas.0511111103
39. Wang Z, Ramsey BJ, Wang D, Wong K, Li H, Wang E, et al. An Observation-Driven Agent-Based Modeling and Analysis Framework for *C. elegans* Embryogenesis. Klymkowsky M, editor. *PLoS ONE*. 2016;11: e0166551-16. doi:10.1371/journal.pone.0166551
40. Santella A, Kolotuev I, Kizilyaprak C, Bao Z. Cross-modality synthesis of EM time series and live fluorescence imaging. *eLife*. 2022;11: e77918. doi:10.7554/eLife.77918
41. Wang S, Ochoa SD, Khaliullin RN, Gerson-Gurwitz A, Hendel JM, Zhao Z, et al. A high-content imaging approach to profile *C. elegans* embryonic development. *Development*. 2019;146: dev174029. doi:10.1242/dev.174029
42. Ulicna K, Kelkar M, Soelistyo CJ, Charras GT, Lowe AR. Learning dynamic image representations for self-supervised cell cycle annotation. *bioRxiv*. 2023; 2023.05.30.542796.

doi:10.1101/2023.05.30.542796

43. Miller LM, Plenefisch JD, Casson LP, Meyer BJ. *xol-1*: A gene that controls the male modes of both sex determination and X chromosome dosage compensation in *C. elegans*. *Cell*. 1988;55: 167–183. doi:10.1016/0092-8674(88)90019-0
44. Chuang P-T, Lieb JD, Meyer BJ. Sex-Specific Assembly of a Dosage Compensation Complex on the Nematode X Chromosome. *Science*. 1996;274: 1736–1739. doi:10.1126/science.274.5293.1736
45. Davis TL, Meyer BJ. SDC-3 coordinates the assembly of a dosage compensation complex on the nematode X chromosome. *Dev Camb Engl*. 1997;124: 1019–31.
46. Vielle A, Lang J, Dong Y, Ercan S, Kotwaliwale C, Rechtsteiner A, et al. H4K20me1 contributes to downregulation of X-linked genes for *C. elegans* dosage compensation. *PLoS genetics*. 2012;8: e1002933. doi:10.1371/journal.pgen.1002933
47. Liu T, Rechtsteiner A, Egelhofer TA, Vielle A, Latorre I, Cheung M-S, et al. Broad chromosomal domains of histone modification patterns in *C. elegans*. *Genome Research*. 2011;21: 227–236. doi:10.1101/gr.115519.110
48. Wells MB, Snyder MJ, Custer LM, Csankovszki G. *Caenorhabditis elegans* Dosage Compensation Regulates Histone H4 Chromatin State on X Chromosomes. *Mol Cell Biol*. 2012;32: 1710–1719. doi:10.1128/mcb.06546-11
49. Hinton GE, Salakhutdinov RR. Reducing the Dimensionality of Data with Neural Networks. *Science*. 2006;313: 504–507. doi:10.1126/science.1127647
50. Masci J, Meier U, Cireşan D, Schmidhuber J. Stacked Convolutional Auto-Encoders for Hierarchical Feature Extraction. 2011. pp. 52–59. doi:10.1007/978-3-642-21735-7\_7
51. Bahry E, Breimann L, Zouinkhi M, Epstein L, Kolyvanov K, Mamrak N, et al. RS-FISH: precise, interactive, fast, and scalable FISH spot detection. *Nat Methods*. 2022;19: 1563–1567. doi:10.1038/s41592-022-01669-y
52. Halpern KB, Itzkovitz S. Single molecule approaches for quantifying transcription and degradation rates in intact mammalian tissues. *Methods*. 2016;98: 134–142. doi:10.1016/j.ymeth.2015.11.015
53. Sivaramakrishnan P, Watkins C, Murray JI. Transcript accumulation rates in the early *Caenorhabditis elegans* embryo. *Sci Adv*. 2023;9: eadi1270. doi:10.1126/sciadv.adi1270
54. Coté A, Coté C, Bayatpour S, Drexler HL, Alexander KA, Chen F, et al. pre-mRNA spatial distributions suggest that splicing can occur post-transcriptionally. *Biorxiv*. 2021; 2020.04.06.028092. doi:10.1101/2020.04.06.028092
55. Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. *Nature*. 2000;403: 335–338. doi:10.1038/35002125
56. Molina N, Suter DM, Cannavo R, Zoller B, Gotic I, Naef F. Stimulus-induced modulation of transcriptional bursting in a single mammalian gene. *Proc Natl Acad Sci*. 2013;110: 20563–20568. doi:10.1073/pnas.1312310110
57. Darzacq X, Shav-Tal Y, Turrís V de, Brody Y, Shenoy SM, Phair RD, et al. In vivo dynamics of RNA polymerase II transcription. *Nat Struct Mol Biol*. 2007;14: 796–806. doi:10.1038/nsmb1280
58. Singh J, Padgett RA. Rates of in situ transcription and splicing in large human genes. *Nat Struct Mol Biol*. 2009;16: 1128–1133. doi:10.1038/nsmb.1666
59. Breimann L, Morao AK, Kim J, Jimenez DS, Maryn N, Bikkasani K, et al. The histone H4 lysine 20 demethylase DPY-21 regulates the dynamics of condensin DC binding. *J Cell Sci*. 2022;135. doi:10.1242/jcs.258818
60. Jans J, Gladden JM, Ralston EJ, Pickle CS, Michel AH, Pferdehirt RR, et al. A condensin-like dosage compensation complex acts at a distance to control expression throughout the genome. *Genes & Development*. 2009;23: 602–618. doi:10.1101/gad.1751109
61. Kramer M, Kranz A-L, Su A, Winterkorn LH, Albritton SE, Ercan S. Developmental Dynamics

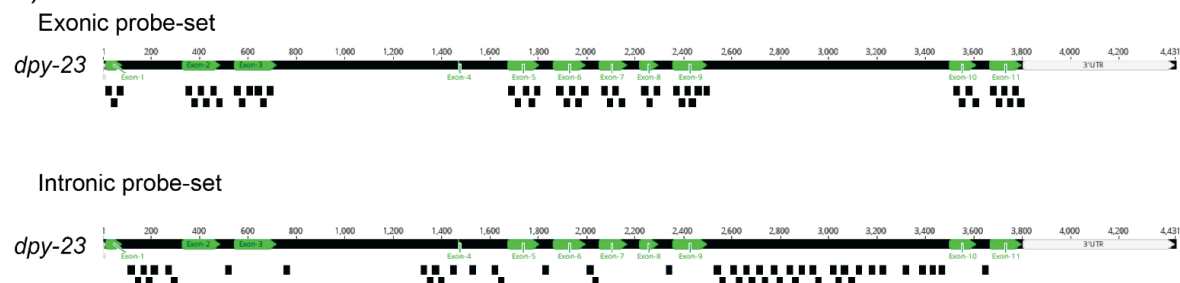
- of X-Chromosome Dosage Compensation by the DCC and H4K20me1 in *C. elegans*. Larschan E, editor. *PLoS genetics*. 2015;11: e1005698. doi:10.1371/journal.pgen.1005698
62. Brejc K, Bian Q, Uzawa S, Wheeler BS, Anderson EC, King DS, et al. Dynamic Control of X Chromosome Conformation and Repression by a Histone H4K20 Demethylase. *Cell*. 2017;171: 85-102.e23. doi:10.1016/j.cell.2017.07.041
63. Wheeler BS, Anderson E, Frøkjær-Jensen C, Bian Q, Jorgensen E, Meyer BJ. Chromosome-wide mechanisms to decouple gene expression from gene dose during sex-chromosome evolution. *Elife*. 2016;5: e17365. doi:10.7554/elife.17365
64. Albritton SE, Kranz A-L, Winterkorn LH, Street LA, Ercan S. Cooperation between a hierarchical set of recruitment sites targets the X chromosome for dosage compensation. *eLife*. 2017;6: e23645-29. doi:10.7554/elife.23645
65. Anderson EC, Frankino PA, Higuchi-Sanabria R, Yang Q, Bian Q, Podshivalova K, et al. X Chromosome Domain Architecture Regulates *Caenorhabditis elegans* Lifespan but Not Dosage Compensation. *Developmental Cell*. 2019;51: 192-207.e6. doi:10.1016/j.devcel.2019.08.004
66. Bao Z, Zhao Z, Boyle TJ, Murray JI, Waterston RH. Control of cell cycle timing during *C. elegans* embryogenesis. *Dev Biol*. 2008;318: 65–72. doi:10.1016/j.ydbio.2008.02.054
67. Hall DH, Herndon LA, Altun Z. WormAtlas Embryo Handbook - Introduction to *C. elegans* Embryo Anatomy. *WormAtlas*. 2017. doi:10.3908/wormatlas.4.1
68. Friedrich D, Friedel L, Finzel A, Herrmann A, Preibisch S, Loewer A. Stochastic transcription in the p53-mediated response to DNA damage is modulated by burst frequency. *Molecular Systems Biology*. 2019;15: e9068. doi:10.15252/msb.20199068
69. Lammers NC, Kim YJ, Zhao J, Garcia HG. A matter of time: Using dynamics and theory to uncover mechanisms of transcriptional bursting. *Curr Opin Cell Biol*. 2020;67: 147–157. doi:10.1016/j.ceb.2020.08.001
70. Larson DR, Fritsch C, Sun L, Meng X, Lawrence DS, Singer RH. Direct observation of frequency modulated transcription in single cells using light activation. *Elife*. 2013;2: e00750. doi:10.7554/elife.00750
71. Mir M, Reimer A, Haines JE, Li X-Y, Stadler M, Garcia H, et al. Dense Bicoid hubs accentuate binding along the morphogen gradient. *Gene Dev*. 2017;31: 1784–1794. doi:10.1101/gad.305078.117
72. Chen L-F, Lin YT, Gallegos DA, Hazlett MF, Gómez-Schiavon M, Yang MG, et al. Enhancer Histone Acetylation Modulates Transcriptional Bursting Dynamics of Neuronal Activity-Inducible Genes. *Cell Reports*. 2019;26: 1174-1188.e5. doi:10.1016/j.celrep.2019.01.032
73. Urban EA, Johnston RJ. Buffering and Amplifying Transcriptional Noise During Cell Fate Specification. *Frontiers Genetics*. 2018;9: 591. doi:10.3389/fgene.2018.00591
74. Nicolas D, Zoller B, Suter DM, Naef F. Modulation of transcriptional burst frequency by histone acetylation. *Proceedings of the National Academy of Sciences of the United States of America*. 2018;115: 7153–7158. doi:10.1073/pnas.1722330115
75. Wu S, Li K, Li Y, Zhao T, Li T, Yang Y-F, et al. Independent regulation of gene expression level and noise by histone modifications. *Plos Comput Biol*. 2017;13: e1005585. doi:10.1371/journal.pcbi.1005585
76. Bothma JP, Garcia HG, Esposito E, Schlissel G, Gregor T, Levine M. Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living *Drosophila* embryos. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;111: 10598–10603. doi:10.1073/pnas.1410022111
77. Walters MC, Fiering S, Eidemiller J, Magis W, Groudine M, Martin DI. Enhancers increase the probability but not the level of gene expression. *Proc National Acad Sci*. 1995;92: 7125–7129. doi:10.1073/pnas.92.15.7125
78. Bartman CR, Hsu SC, Hsiung CCS, Raj A, Blobel GA. Enhancer Regulation of Transcriptional

- Bursting Parameters Revealed by Forced Chromatin Looping. *Molecular Cell*. 2016; 1–12.  
doi:10.1016/j.molcel.2016.03.007
79. Crane E, Bian Q, McCord RP, Lajoie BR, Wheeler BS, Ralston EJ, et al. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature*. 2015;523: 240–244.  
doi:10.1038/nature14450
80. Kim J, Jimenez DS, Ragipani B, Zhang B, Street LA, Kramer M, et al. Condensin DC loads and spreads from recruitment sites to create loop-anchored TADs in *C. elegans*. *Elife*. 2021;11: e68745.  
doi:10.7554/elife.68745
81. Street LA, Morao AK, Winterkorn LH, Jiao C-Y, Albritton SE, Sadic M, et al. Binding of an X-specific Condensin Correlates with a Reduction in Active Histone Modifications at Gene Regulatory Elements. *Genetics*. 2019;212: genetics.302254.2019. doi:10.1534/genetics.119.302254
82. Lau AC, Nabeshima K, Csankovszki G. The *C. elegans* dosage compensation complex mediates interphase X chromosome compaction. *Epigenet Chromatin*. 2014;7: 31. doi:10.1186/1756-8935-7-31
83. Snyder MJ, Lau AC, Brouhard EA, Davis MB, Jiang J, Sifuentes MH, et al. Anchoring of Heterochromatin to the Nuclear Lamina Reinforces Dosage Compensation-Mediated Gene Repression. Larschan E, editor. *Plos Genet*. 2016;12: e1006341. doi:10.1371/journal.pgen.1006341
84. Raj A, Oudenaarden A van. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell*. 2008;135: 216–226. doi:10.1016/j.cell.2008.09.050
85. Tunnaciff E, Chubb JR. What Is a Transcriptional Burst? *Trends in Genetics*. 2020;36: 288–297.  
doi:10.1016/j.tig.2020.01.003
86. Samata M, Akhtar A. Dosage Compensation of the X Chromosome: A Complex Epigenetic Assignment Involving Chromatin Regulators and Long Noncoding RNAs. *Annu Rev Biochem*. 2018;87: 1–28. doi:10.1146/annurev-biochem-062917-011816
87. Kuroda MI, Hilfiker A, Lucchesi JC. Dosage Compensation in *Drosophila*—a Model for the Coordinate Regulation of Transcription. *Genetics*. 2016;204: 435–450.  
doi:10.1534/genetics.115.185108
88. Beadle LF, Zhou H, Rattray M, Ashe HL. Modulation of transcription burst amplitude underpins dosage compensation in the *Drosophila* embryo. *Cell Rep*. 2023;42: 112382.  
doi:10.1016/j.celrep.2023.112382
89. Schmidt U, Weigert M, Broaddus C, Myers G. Cell Detection with Star-convex Polygons. *Arxiv*. 2018. doi:10.1007/978-3-030-00934-2\_30

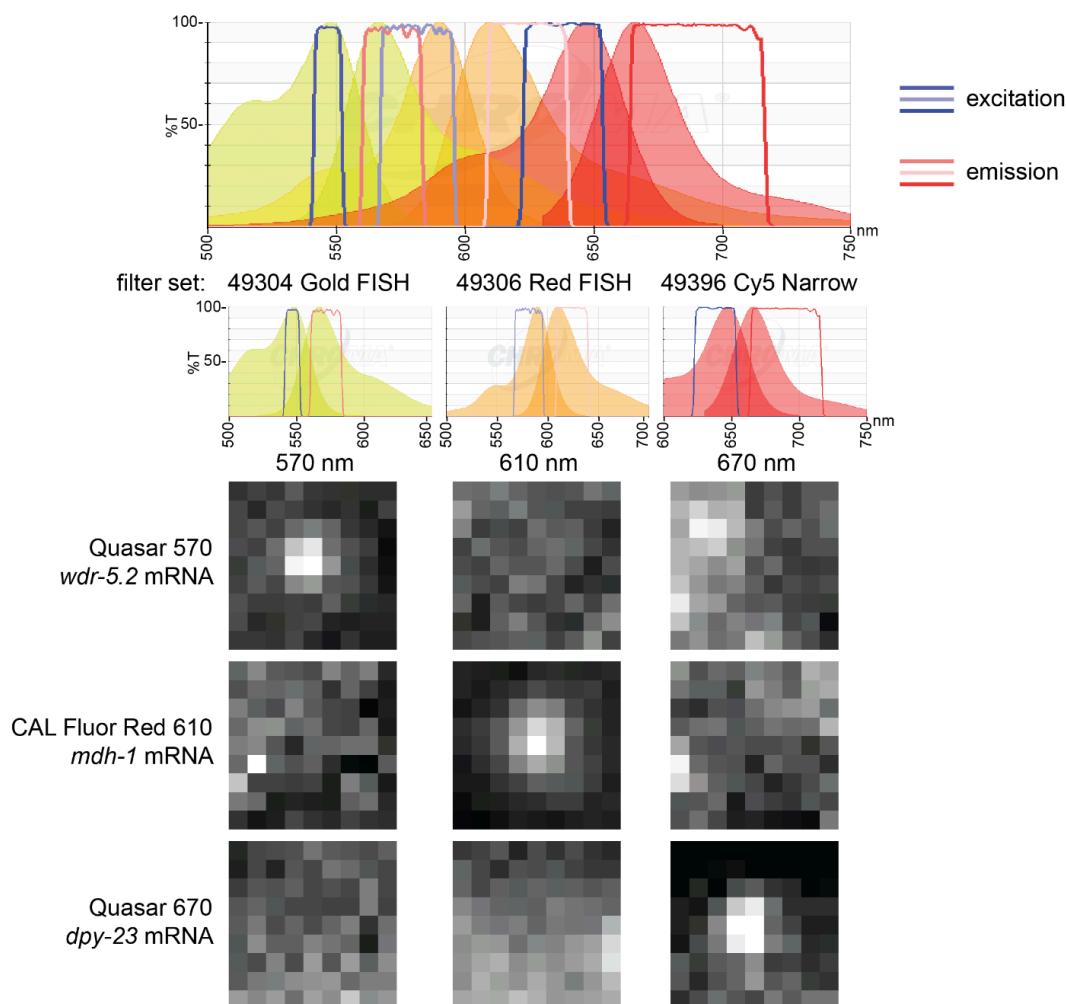
# Supplemental Material

## Supplemental Figure 1

A)



B)

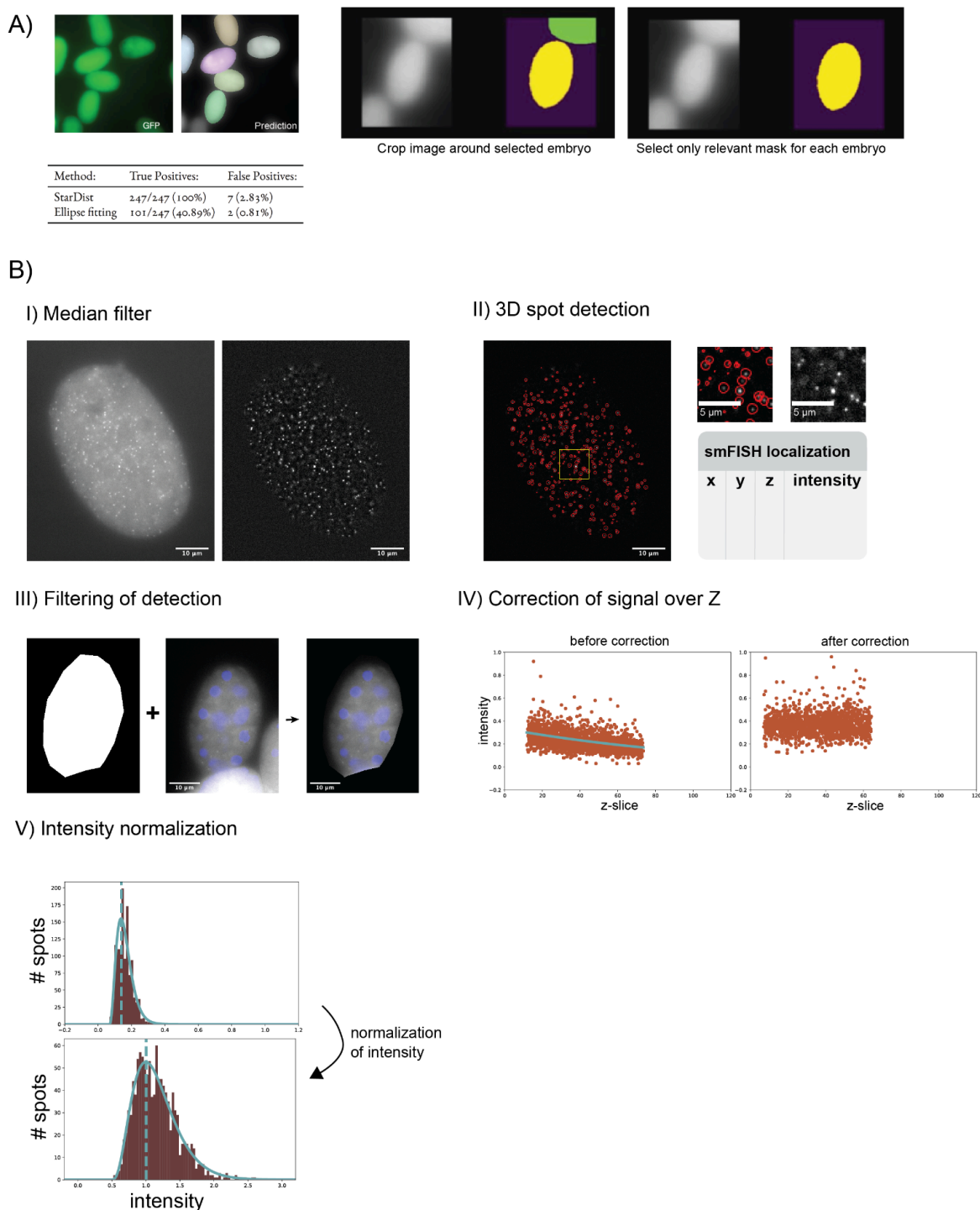


### Supplemental Figure 1 smFISH probe design

**A)** Example for exonic and intronic probes for DPY-23 mRNA. **B)** Curves of the fluorescent spectra of the three smFISH fluorophores (Quasar 570, CAL Fluor Red 610, and Quasar 670) and filter bandwidth for the three filters (Gold FISH, Red FISH, Cy5 Narrow) to specifically detect them. Spectra were created using the Chroma Spectra Viewer. Fluorescent images of smFISH spots for three different mRNAs with three different fluorophores. Images of the same spot in the different channels show the specificity of the signal and low bleed-through.

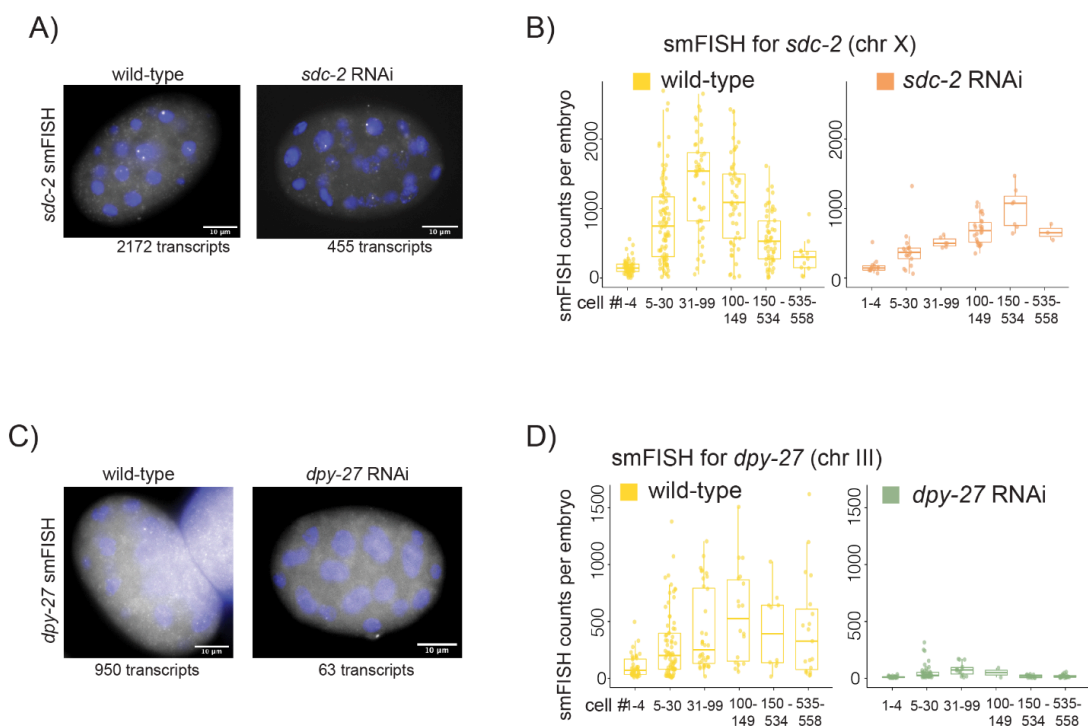


## Supplemental Figure 2



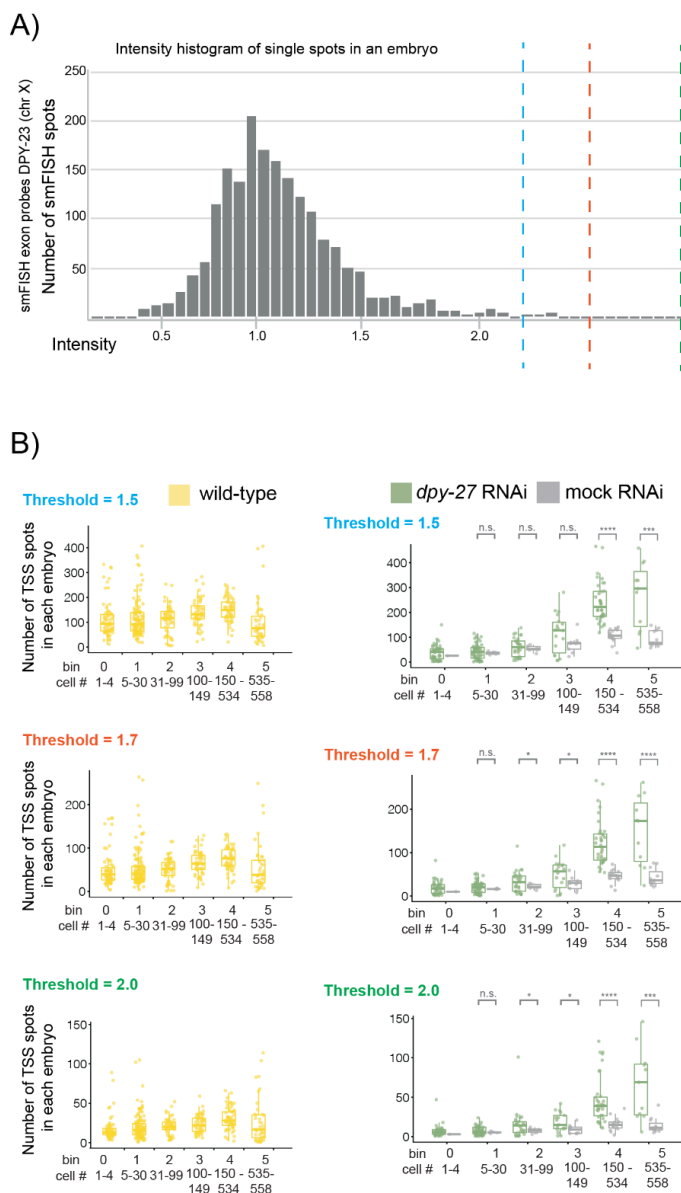
**Supplemental Figure 2 A)** Instance segmentation using StarDist. The neural network was trained by the ground truth data created in the ellipsoid fitting. The network predicted masks on the 2D max projections of GFP images. Masks for single embryos are curated, and the original image is cropped around each segmented embryo. Jaccard similarity score was computed per instance in the validation set. Accepted correctly predicted instances were defined by a Jaccard score  $> 0.75$ . Using StarDist, all (100%) instances in the GT images were found. However, in rare cases ( $< 3\%$ ), areas with low-quality embryos that are not desirable for analysis, as well as dirt, were detected. **B)** (I) The raw smFISH signal (left image) is filtered using a median filter ( $\sigma = 19$ ) to enhance the spots for detection (right image). (II) Single mRNAs are detected in 3D using RS-FISH, which records a table of localizations and intensity. (III) Masks (left image) are created using Stardist, which are used to filter localizations (right image). (IV) The image intensity drops as a function of distance to the coverslip ( $z$ ) (left plot). Through fitting a quadratic function, the intensities are corrected in  $z$  (right plot). (V) The intensity is normalized by fitting a gamma function, where the maximum is set to 1.

### Supplemental Figure 3



**Supplemental Figure 3** **A)** Example images of wild-type or *sdc-2* RNAi embryos labeled with *sdc-2* smFISH. The number of detected smFISH spots for each example is written under the image. **B)** Plots for smFISH detection of X chromosomal mRNA *sdc-2*. Each spot represents the total mRNA count for one embryo, plotted against the binned nuclei number for wild type (yellow) and *sdc-2* RNAi (orange). **C)** Example images of wild-type or *dpy-27* RNAi embryos labeled with *dpy-27* smFISH. The number of detected smFISH spots for each example is written under the image. **D)** Plots for the autosomal mRNA *dpy-27* in wild-type (yellow) or *dpy-27* RNAi (green) conditions.

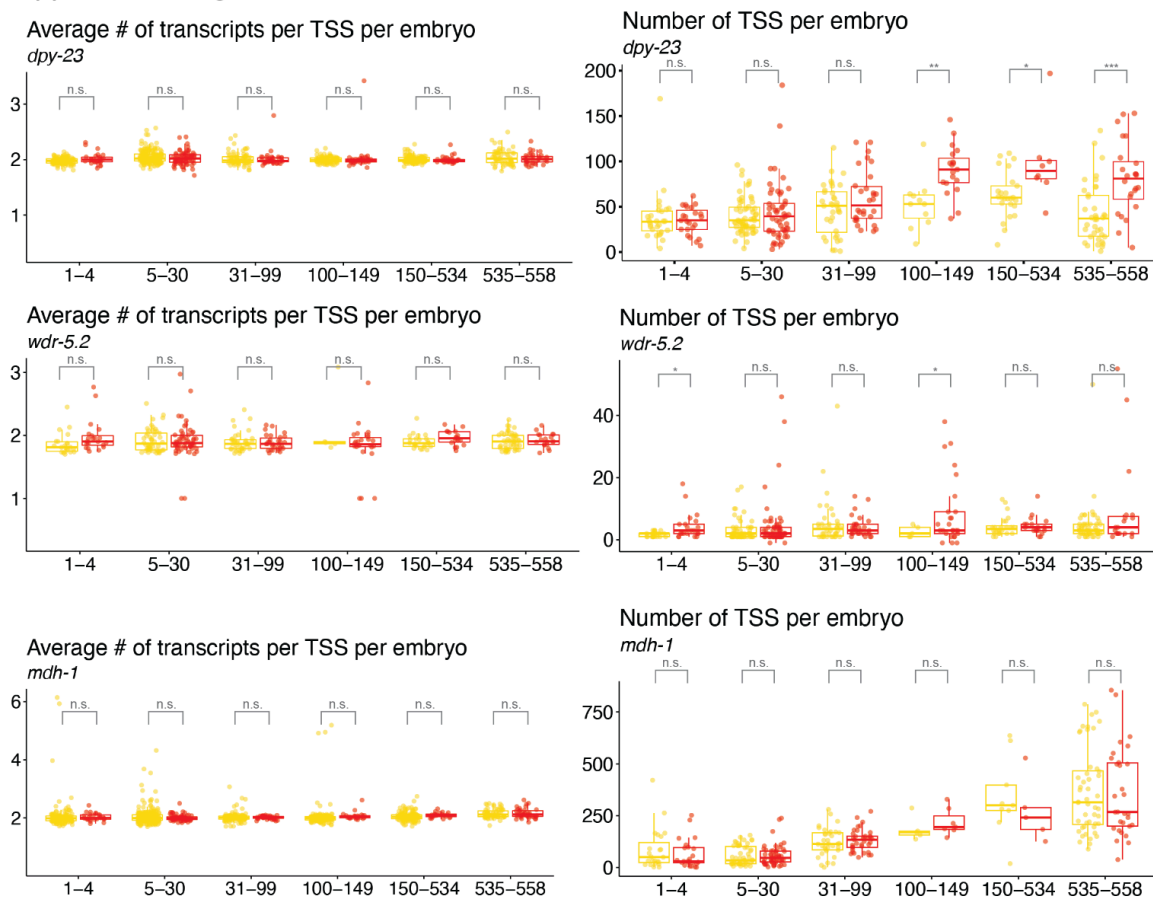
## Supplemental Figure 4



### Supplemental Figure 4 Analysis of threshold selection for transcription start site.

**A)** Histogram of the normalized exon intensities of all detected spots for one embryo. Different cut-offs for detection of high-intensity spots were selected at 1.5 (blue), 1.7 (red), and 2.0 (green). **B)** Relative change of actively transcribing promoters for X chromosomal genes for wild-type (yellow), *dpy-27* RNAi (green), and mock RNAi (gray) conditions. Three different thresholds were selected to classify transcription sites in the dataset. Significant P-values from Welch's unequal variances t-test are noted above the wild-type vs. knock-down condition for each time-point. P-values are denoted as n.s. ( $P > 0.05$ ), \* ( $P \leq 0.05$ ), \*\* ( $P \leq 0.01$ ), \*\*\* ( $P \leq 0.001$ ) and \*\*\*\* ( $P \leq 0.0001$ ). If one of the samples had less than two points, no P-value was calculated.

## Supplemental Figure 5



### Supplemental Figure 5 Analysis of transcription site parameters in *rex-1* deletion.

Relative transcription parameter (average number of transcripts per transcription site per embryo and the number of transcription spots per embryo) changes for the X chromosomal genes *dpy-23* and *wdr-5.2*, as well as autosomal gene *mdh-1* for wild-type (yellow) and *rex-1* deletion (red) conditions. Data are from embryos probed for the three genes simultaneously. Significant P-values from Welch's unequal variances t-test are noted above the wild-type vs. knock-down condition for each time-point. P-values are denoted as n.s. ( $P > 0.05$ ), \* ( $P \leq 0.05$ ), \*\* ( $P \leq 0.01$ ), \*\*\* ( $P \leq 0.001$ ) and \*\*\*\* ( $P \leq 0.0001$ ).

### Information on the imaging data

The complete raw imaging data can be found at <https://data.janelia.org/3Mq4F>. Images are split into subfolders of genotypes N2 (wild-type, CGC), CB428 (*dpy-21* (e428), SEA-12 (ERC41 (ers8[delX:4394846-4396180]))) or RNAi treatment (*sdc-2* (C35C5.1), *dpy-27* (R13G10.1), *ama-1* (F36A4.7), *ev* (empty-vector), *rluc* (Renilla luciferase)). Additionally, the embryo masks are provided in the file mask.zip. The manual annotation of the embryo stage and the image tiles used in the autoencoder's training can be found in the manual\_count.txt and tiles.zip, respectively.

Table 1 The number of probes and fluorophore of each probe-set

gene	Chr	transcript length (nt)	N2 FPKM		Vector RNAi FPKM	RNAi DPY-27 FPKM	DPY-21 KO FPKM		Mature RNA		Nascent RNA	
			Early embryo	Comma embryo	Mixed embryo	Mixed embryo	Early embryo	Mixed embryo	# of probes	Fl	# of probes	Fl
<i>dpy-23</i>	X	1980	68,64	76,93	61,98	79,23	66,81	89,03	48	Q670	48	Q570
<i>sdc-2</i>	X	9247	27,79	10,26	14,65	21,54	25,26	14,21	48	Q670	48	Q570
<i>wdr-5.2</i>	X	2040	23,06	34,85	29,67	39,67	13,66	30,47	48	Q570	-	-
<i>rab-6.2</i>	X	913	53,11	206,39	178,36	193,41	42,31	170,27	41	Q670	46	Q570
<i>mdh-1</i>	V	1149	123,26	588,64	634,14	599,16	121,53	601,70	41	C610	-	-
<i>ama-1</i>	IV	6020	271,49	92,02	92,88	99,41	255,66	104,29	48	C610	48	Q570
<i>dpy-27</i>	III	4677	58,82	77,45	69,78	26,95	80,74	56,82	48	C610	-	-

Table 2 DPY-23 exonic probes

acgaacaatccaccaatcat	caacttttccggcaacatga
cacttctcctttgtgattgt	ggcattccactcaaatagga
tcgtctc gatagattctcga	tcgttgatgccaaattgca
tgacgttgactcggaagggcg	cgactttccttcgatggtaa
cgaacttgctgtcgggcatg	ttgtttggatcatcacttcc
aagtacgagccatgttggtg	acagtcacatggccactg
ccacgcttcacatggaagaa	gcttcacgcattggtggaat
tgacgtgtcaccgcacaaat	tgctccgtctcaaattcgt
aaaacctggcagcgttgac	cgccgggtgggataaaagaa
tcggcgaagcgtttcaagaa	cagttggatccttagtgg
tccaaagtaagactgcatgg	ccaatgggatcacacggaat
agttgttctcacattctcc	cttcatcttgttacgagac
gagaatctcgtcgagcaact	gcttgaagtagacttgacg
tattctgggggtatccaaag	agcttttgagcaagaaggga
aaagtttcagcacaccagg	tggggttggaaatgcgaactt

atttgactgctcctcttt	gttgaacgccggatgtattt
aatttgccagtcacttgag	gctttcccttcatgcaaat
ttcggcggtacttaatacc	ttccacacaatggcattctc
aataacatccaggaagagct	aatttgctttccttcattc
gattcatgagcaagtgaca	agagaagatcgatttcgca
tccagatggagcaaacggaa	tccatttcttctcaacg
cacctcaagtagcgaactt	ctcaaagttcatgctgaccg
aatagttcagtttggtca	tagcatctggttcatacag
ccatttgatgacgtcatgat	ccgatctccaatgtaacga

Table 3 DPY-23 intronic probes

gaacatttccaaacatgcc	cttctgcccaattttcttag
cgcaccactttttctttaa	tgtacagctcaaagtgtgt
catttctttatttccctggt	ttctctgattgcagtttctt
aacagcgcggcaaacatttt	tagttaatgttctactgcc
gtgaacaaacagtcctggca	tatcccttttcggcaataat
tgtacaccatgcatttttc	cctttttagtaaatacceca
gatttctcttgtttgggtc	aaatttgttattcaagccc
agtggtgccctaatactttt	gttttcttttagttacaggt
ttgatactcctccttgaaca	tgtttcactgttacggacga
gcattttgttttgggcaac	acaattccaccaatttttga
gcgtgacgctttttcatat	attcaaaccttctctctgt
gtgtacaggtgggcagaata	ccacaaatcttgatctcgtg
ttgaacctgaaagcgtgagc	gaaattcaggggatcaagcc
gggttcgtgaataaagagcg	ctcttggttaactattgtgt
ccgtgcacacaaaacagct	cactttttgattgtgagctc
tggttatctagatttttct	ttaaggaagctgttggctcc
ctatgtattctcctgagaca	tgccaatctgcgacagattg
agtatgcacgtcgttattgt	acatttgcttagcagtcaga

tcccgaaactcagactaac	tgcagcacagtttgcata
caataaagtatcagttccc	cttcagtcaaggttggaaa
atgagagccatgattgtaa	atgctccaaaagcttagta
atgctcagtcagtgatgag	gcctgtgagaacttttcat
tctgccgcttcgtttaatc	ggtcaattattgagtgga
attcttgcccaaatggagc	
ctcccactttctcgtttc	

Table 4 RAB-6.2 exonic probes

taccaaagtccgacatttt	attaagctccttggccttc
agcttgaattcttcaacgg	ggcagaagtctcgaataca
actttgtccccgaggaaca	tgcttcacattgtagccggc
tgattagggagggtttccg	tccggcgatacagcggaaaa
aaggagtcgtacatgaagcg	ccttgatgattcctggtagg
tgggtggcctggtaggtgtg	tttgggtggctccactgggtc
ttactcaggaagtcgattcc	gattgggtccatcgtgacca
ttcggctcctcaagtacatg	cgggtgacaattgacgctga
tcccaaagttgaagacgaac	tagcaccagcacgatccctc
gaaacgttcctgtccagcgg	ttgctcaggaatacaagag
tgtaggatgggatgagtgag	atcacaacctttttgaca
acagcaactgtagaatcacg	aattatgctgtgagattgga
gttggatgtcataaacca	ggggcggaggatgctttaa
ggaagttgatgaaagagt	tcttcatgacaaaaccatta
tacggacgtcgtcaatccac	tttaaggggataattaata
attacatcactccacgctc	aaaacttactggcgatagcg
cttgttgcgaccaacatga	tttggggaatttaagata
gcctcttatcgaaagatcg	acacagaaaaaatcgataat
tctcctcgtcgggtgtgac	agaagccattcagagttcat
ccatgcataagttaataga	

caaagaaaattacacatgtt

ataataataatttattatg

Table 5 RAB-6.2 intronic probes

ctacggtagaatttcaatca	ttaaattgcattatccgccgt
ttagaaaacttgagaattca	taaaacttcagcctggagct
cgcaaaaactgatttcttac	tcatactgaaagttagccg
cggattttgaaatcatttct	aggttttaagctagccaaga
catcagaaaaacttctcca	ctccaacaatatctttaga
tgtgaagctttacagcatga	ttcagttaaagctatttct
aaatttctccagaggatcc	gatccaatttatgttaacgt
actatatacactactatagt	gaagctagttcaagtagtg
accaatatagttattatagt	taaacttgaggtaggtccag
tgccagtgactattgtagta	gcgtaggctctgattataaa
gctagaagcttttcgactta	atagagagcttgggcttgaa
tgtttttgagcttctcttc	tgtgtttaaggatgagctgg
tctaacttttctatcaatat	gcttatgtagtaggtcaac
aaataatggtttccttttgg	ttatgggagattcagcttat
tagttcaaaaatcaagattaa	atthttgctcaggaaaaatca
agagtttttcatgcccac	tttttagcctaatagtaacgt
tgaagctttttgctgttac	ctaattctgtttagaatgcg
ctgcaaaaaagttgggggta	cgcaaggtaagtttagcttg
tagcggaaactctcattcga	agatcagctcgaacaggcc
ccactacacaatttaattga	atagttcagctaagaccacg
tatgtaagctttatgtcagt	ggttaacaggtttcagaaaa
aaaaggtttttaaaacttac	tacacattgagtttattagt
aatcatggctttttgaagt	caaattaatagacctggagc

Table 6 WDR-5.2 exonic probes

tgctcttcatcttgagaagc	gtgttccagaaggggtgaaa
----------------------	----------------------



ggtttccgcaccaactattg	tgaatgatccagaggcaat
actggaagagatggctgact	acaccagatgcgaatagttt
gctataataggacgcattgg	aaatcgtgtttccattgcgt
taagaccatagctggggatg	gaactaacggatcctcgtg
ccaaattgtggaattcccat	tccgtcgcggttaaacaca
cgaagctgagctggcagac	taggatcccaggctagata
gcgtgaagaaagcataggc	cccaaattcgaactattcca
gtgcagagctataacttgg	acacaagtcccggtagtcca
actgggttcggcattggag	ctcttcgtcgatcaacgttt
agttttaggttagccaaag	tgacatgtgtgattggagga
gtgcgcatttgaatttcag	gtatatattttccgtttgg
tgatccctgagatcgacttt	gtattgttgaggtttgaggc
taccgtcgtctggagaaaa	ttggtagtcccacagtttg
atactacaatccgcagatcc	tactcttgagcactctcag
tacacaaaatccatcctcca	gtatttggaaatttcgtgcc
ccaagccgatgaccataag	taaccgaaaaattcgcggcc
gctccatgagaactcattga	gctgacaatccatttacctc
gcacgagacaatgagtttcg	acactttgtgatctccgag
ctttgactagtttgcacg	cttgctgaaggtccagat
cgaccactgctgacatcaaa	gccatctaattgtctgaagga
cccttcaaagtcttgacac	gacaatcggtcacatgaca
cagcagaacacgtagttgt	gcggcagaagcaattatgtt
ttatgactgggatgccaga	ttgatccgcataatctggctc

AMA-1 exonic probes: design ready catalog number: VSMF-6001-5

Table 7 AMA-1 intronic probes

gggttaaatacgggggtttc	tattctttggcaatttccgg
tcttgattttagcctaaa	cccactcaaagtacaattt
atfttcgaccgactttggag	ttgggggtcattttcgaac

ttgacaactttccactgga	atttgcgaagttttcctgg
attagctcattctcagcttt	ctcgaacagtgtacctttt
atgggcaaaatccggcaatt	gatgattttagtggcggttt
ttacgattttccagtctca	atggcaaatagctgtgcgat
cttatcggccaaaaagtgca	gacttttctgggtgcaaac
ctgttctaacgggatttgca	tttattccaacgcaactga
tcgggcttaaaaccactttt	acaggcttttgaacattcca
attccaattcttaagcgcat	gtctcaacagttggagtttt
gtcttgaatggttttccatg	gccgaatttccaccctaaaa
ttcctaaaccgcaattgtga	tttcgcacaatttcagccaa
ttttgagtcattttctggcg	gtacttttcgagttttcctt
tcacacaaaatcacgccagc	gcccaaatagacagatttca
ttttcctcaagttttggctg	ttttcgaacgttttctcgcg
aggtggtttccagggaaaaat	gatagattccccagaaatg
tccatcacaattttccagtt	gttgccaacttatcacattt
ttcctcgatcaaatttcgga	cttcggtgattttgctact
tccatcaaaaatccagtc	gcccctaatttctggagaat
tagcattttgagcgacactc	actgttacatttttcgcagt
ttttcaagtatccaggctta	agttttctccgaatattct
aacataggccaattttcgca	atttccgaacattttcaca
ctctccaatttttttagggt	gtccgtaactttgcttttg

Table 8 MDH-1 exonic probes

cggtgacccaaaacgcgaagt	agagtgattccccagatga
ggatgacgatcgagtaaccg	tgacatctgggaattgagtt
ttccgaatacagttccgtc	ttcttghtaacggtagcgtg
aacgagctccactggttggt	gtaagcgtcgggtctcagttc
agcattgtgggacgtcgaga	tgaaggaaagcgttctctcc
acaactcctcgagaatggt	tggacagtggcgaatggt

agcacaatcctggagctcaa	cttctcaatgatgactcctc
caacgctgaaaaggggtgga	acattgcagacgagagcttt
gacttttcatcagtgcacgc	tggatgtgatcacaagcggc
gtagtcgataccggggaagg	cttggttccgaaatgccagt
gcatggctccgaccaagaaa	cagccatggagacgaattgc
agcaagaagatctttgcgct	ccgtatgaaccatcggatgg
tccttgtgactgaagatct	gaaaatgagtccttgggga
cttggcatactcagcaagag	cctcgatggtaactgggaat
ctgggtttccaacaacgata	agacctgaacaatctcca
actggcagcaatgaatgca	cctggcaaaatcgtcgaag
agttcttggctgggatcttt	ttggtgtagcagcagatctt
tccaatctggatcggcaga	ctttagagcatcatctctct
ttgagcaagagcacggttat	ttagatgtggcatcatcgc
ttgtccggttttaagggcg	aagttaagggagcggttgg
acgttcttgacatttccgat	aaacgacaagacggcctact

Table 9 DPY-27 exonic probes

tgtcagcatatggctatca	aagccaattccgattctaca
ccaggatcattgtcaagttc	attccgctagtatcagtatc
cctcctgaattgatcaatgc	cgagttgcttctcaagact
ttcacgattattgtcgggg	gccattcagtgcaataac
tctttttgcgatgcatcttt	tctcttctgctttaactgc
tcatatcaattccagcacgt	acggtcacttctctatagtt
tcattagtgaatcgcctca	acttgacgtttcagatctgc
caagcattccttctcattc	attgctcaattgctgcttg
attgtgaacttttgcctcc	aaacattcggctccatttcc
caaagacttttaggtgctcc	ctgctcaattgagtccttat
ttgtttcagcataacggcac	taattgcagtttgcgtgca
ctttgcctcaagcatgatat	ctctccaactccaattgag

ccagctcagtgacaatacta	gtttgcaattgcgattgat
cttcgattccaatcgattc	tgacaaattgccaggtgga
agagattctcgcattctctg	ccagcttgcaaatcttgfta
attcgagttgcatgttttgc	gaagcgacacaatcatcgca
cgattgcagaagctcatctg	ccttcagcttatcgctaaac
tgtttgcaatgtccctftaa	tgaatctgccattctcaac
gctttgtcaaaacttcagctt	agtagtaccacaagaatgcga
gaagcccggaaataatccag	atcggtagattttccttctt
ctagtttgcaaacgtggta	catcgatttcatccatcacg
agaccaaagtccgagagttt	caggctgacattgttcaagt
ttcatcttctcttcaactt	tcggatccaccataatgtta
acgacaattggtggtggaga	ctgagcccattactcaaatc

Table 10 SDC-2 intronic probes

agaattgagatgttcagata	actatattttagaagctca
ttttgttctcgtcttgttca	gtctcaattcaaatgttcga
gtgtgtgtttcagttatfta	attgcagcttaaagctacat
gttgccatagaaaatgtaaa	attaaagagaaagaagccac
acggtttcagaatagaaaca	actgagcgtacttcatcttt
aaaaggtgttcataatttct	tttggattttaactaccat
tccaaaatttttagaaggt	tcactttccgactcgaaata
cttgctcgaaactcagtttt	tagcgctttagagtttttaa
ttataaaactttgagcaagc	ttctgattgttttagatga
ataaatatatttctgatgat	acgtacaaaactttgattca
attgcaatttgaactttta	ctgaaatggaacacactttt
tacagtttgacgtttgtggc	ggcaatttttgagcaatgga
ttcaaacagcaatctcac	ctgcaaaattattagttatt
tcacgcatttaaagaaaaca	acaaaattccttacctttct
gacgtatttgacttttagat	gtattaacaacgttttgcct

agaatacttattcagttgat	gtatcaaattattccatac
gatttcattactttttgca	ttttaaacgcaaatgctga
acccgaaataagecgtatcaa	tatattactacagcgactgc
cagattttcaaaaaaccca	attacgaaaaacttttaga
agtgatattgaagatatga	gctaccacaaaaatcaatgg
aataagcaagtcctcgtt	tctgactaaaattgaagttt
ttcagataatagaattatga	aaaactgggtacaaaactcac
actcattgtaatatgtgtt	gtaattatggtattgatgca
ggttaaagcgacaaaactca	attacgatataatttatga

Table 11 SDC-2 exonic probes

catctggagaatcggactcg	gagttcactagtacctctac
tcttctaattcctgttcgtt	ttctttacagctggctttac
ccagttttgatgttttgcg	ctcttccaaactcagtttct
ttggggaacaagtttgcca	gcattgcaacataaccgttc
tccgatgaacggttcatacg	atthtgaagctcttttagcc
tcaggaaacttatttcccca	gttttctcggacagcatatt
tggcttgcaaatgcgact	ggttggagcattgattgtt
ttttgcatccatgtttgact	gctccaccggtataaattag
aaatgtgtacacatctcccg	atthgtttttcgcacgtgg
gaaggtacaccctcaatttc	ccttcaaattcggagtcgac
ttttcacggfactggtagt	attcatctttctttcctct
gttcggaatctgtgaggaa	agcacgggtgataacatctt
atgcttgcaacttcaaaacc	tgtgaaagcatctcgatcgc
gctgattccaaacagagtcg	ccaaatccatgtagtcgttg
cttctcttcaaaacgcttc	aatagctctcatttcgatcc
ggcccaaacatatctttgag	gcttcttggtacttttcatt
gaagaccatttcaccatct	agtggtgcaagttttttggc
ccagttttctcacactaga	attctccatagcaacgtgac

ttctgacaacagcatacgt	cggfatactgatagcatctg
tacacaacctgattgagcgt	cgaatttgggcattctggat
cattcagcttgtgacttctg	agtgcataaattgtcgtct
tgctcacacacaaaacgtcc	ggaaccggaactatcgagtg
cacctcaaatgtcgtaca	aagctcagaaagttgccgac
agggatgcaatagtacacgc	tttcatatcgggcagagtg